AVAYA

Ethernet Routing Switch
Engineering

The Large Campus
Technical Solution Guide

Abstract

This Technical Solution Guide defines the recommended designs for a Large Converged Campus infrastructure. The document provides an overview of the best design practices to implement a network capable of supporting converged applications and services.

The audience for this Technical Solution Guide is intended to be Avaya Sales teams, Partner Sales teams and end-user customers. All of these groups can benefit from understanding the common design practices and recommended components for a converged campus network design.

For any comments, edits, corrections, or general feedback, please contact Dan DeBacker (ddebacke@Avaya.com).

Table of Contents

Figures

Tables

Document Updates

October 2010

Updated and rebranded document to reflect release 7.0 on the ERS 8800/8600.

Conventions

This section describes the text, image, and command conventions used in this document.

Symbols:

Tip – Highlights a configuration or technical tip.

Note – Highlights important information to the reader.

Warning – Highlights important information about an action that may result in equipment damage, configuration or data loss.

Text:

Bold text indicates emphasis.

Italic text in a Courier New font indicates text the user must enter or select in a menu item, button or command:

> ERS5520-48T# *show running-config*

Output examples from Avaya devices are displayed in a Lucinda Console font:

> ERS5520-48T# *show running-config*
>
> ! Embedded ASCII Configuration Generator Script
> ! Model = Ethernet Routing Switch 5520-24T-PWR
> ! Software version = v5.0.0.011
> enable
> configure terminal

1. Converged Campus Design Solutions

The Converged Campus architecture is built using the fundamental strategic values of the Avaya Data Solutions organization. By adhering to these core values, Avaya provides a solid infrastructure on which the enterprise can build upon. With this solid infrastructure, the enterprise can solve their business challenges by enabling services easily and without worry. Avaya offers a unique value proposition in its ability to provide this infrastructure while still offering best-in-class total cost of ownership.

Figure 1.1: Avaya Data Solutions Strategic Values

The Converged Campus solutions have been broken down into Small, Medium, and Large to address specific requirements of the Enterprise. A major objective of the Small / Medium / Large Campus solutions is to provide a blueprint and starting point for the customer network design. By providing solutions that have been architected, validated, and documented, the building block for the network is now in place and ready for the specific customization required by each individual network. This customization comes in the form of specific VLANs required, protocols being used, number of edges to connect, and application requirements for the infrastructure.

This solution guide provides optimal network designs and general best practices when implementing and administering the network. The end result is a network that can sustain both normal data traffic as well as any converged applications deployed in the enterprise.

	Note that all design recommendations and best practices within this guide should be reviewed against the available features on the Ethernet switching platforms being deployed and should also be reviewed against the release notes for the versions of software being used. As feature enhancements are introduced and bugs are fixed, it is imperative to understand the capabilities and limitations of the switches and software being implemented. This ensures that the design being deployed utilizes the features and functions of the switches to their maximum effectiveness.

1.1   Avaya Converged Enterprise Architecture

The architecture shown in Figure 1.2 includes all areas of the Avaya Converged Enterprise solution. This guide focuses specifically on the Converged Campus architecture for edge switching and core switching. There are many permutations of possible designs when deploying infrastructure from Avaya, but this guide highlights the major design concepts that need to be addressed.

The ultimate goal of these designs is to provide a highly reliable infrastructure with sub-second seamless failover preventing any interruption of traffic on the network. The value in this is two-fold. First, in the event of a failure, no loss of connectivity or traffic will be experienced by the end user. Secondly, and probably just as important, is the ability to provide near hitless software upgrades for the core of the network.
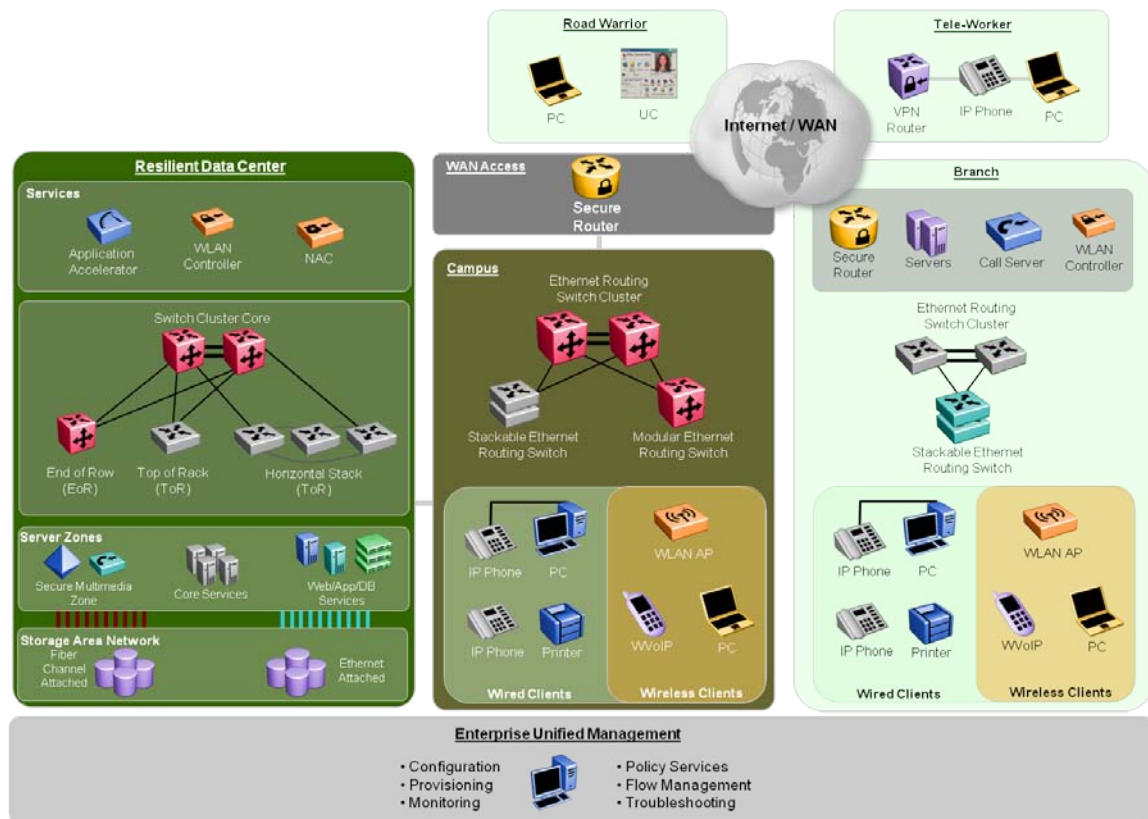
Figure 1.2: Converged Enterprise Architecture

1.2 Chassis versus Stackable

Several factors come into play when choosing the edge switching solution. Consider the following criteria when selecting the edge product while keeping in mind that the stacking technology continues to evolve and is getting closer and closer to simulating a modular chassis solution in many respects.

Switch reliability is a key concern. In the past, modular switches were thought to be more reliable with redundant power supplies, redundant fan trays, and redundant switch fabrics and CPUs. However, the evolution of the stackable switch has reduced the disparity between the two platforms by employing a resilient stacking architecture, supporting internal or external redundant power supplies, and providing features such as auto unit replacement and new unit quick configuration. Both solutions can provide an equally highly reliable edge solution today.

Scalability of the edge switch includes the ability to add ports easily, increase bandwidth out of the closet, and add protocol and features within the closet. A chassis solution typically adds ports by adding new input/output (I/O) modules in the chassis, while stackable switches add ports by adding switches to the existing stack. Both solutions limit the total number of ports supported in a single stack/chassis. The stackable switches provide more flexibility when adding bandwidth out of the closet. A stack can be broken up into two or more stacks, thus increasing bandwidth out of the closet very easily.

As stackable switches are added the closet, each one must be powered individually, which uses several outlets in the closet. In contrast, only two to four outlets are usually required for a chassis.

The same protocols and features are for the most part available on both platforms; however, scalability of those protocols is normally greater in a chassis solution. It is easier to redeploy stackable switches as a stack or stand-alone unit, whereas the modular chassis requires additional hardware to support the I/O modules.

Serviceability and manageability differences between the two solutions are minimal. With both solutions, you can add ports easily, perform software upgrades, retain multiple configurations, and manage the stack or chassis as a single entity.

Rack space can also be a consideration when selecting the edge-switching platform. Typically, a stackable solution takes up less total rack space than a chassis solution in both height and depth. However, stackable switches require rear access for power connections and stacking connections, whereas a chassis solution requires only front access.

The final consideration between the two solutions is price. Usually, a chassis solution is slightly more expensive than a stackable solution due to the additional Switch Fabric/CPU (SF/CPU), chassis, and power supplies needed. In summary, both solutions offer great reliability and scalability. Customers must decide which provides the optimal solution for their organization.

1.3 Layer 2 versus Layer 3 at the Edge

The process of choosing between Layer 2 and Layer 3 at the edge can take many different twists. When considering the differences between the two, it is imperative to keep in mind the end goal of 99.999 percent network availability. There are several ways to design a Converged Campus network. The goal is to design a network that provides high reliability, fast convergence, and yields the lowest possible total cost of ownership (TCO). The TCO is derived by adding the initial cost of equipment/installation (CAPEX – Capital Expenditures) and the ongoing administration and support of the network (OPEX – Operating Expenditures). Over the long run, the OPEX is often higher than the CAPEX, so the goal is to help reduce OPEX by making the network easy to administer and troubleshoot.

The two major areas to consider when deciding between Layer 2 and Layer 3 at the edge are (1) IP routing and (2) intelligence, which can be thought of as operating at Layers 3 to 7. Intelligence can further be defined as the ability to provide traffic management (QoS and content-aware switching) and security, which includes end user authentication and policy enforcement. The goal is to centralize the routing and distribute the intelligence to provide a high-performing and secure network along with easy and simplified management. However, one must also consider the number of users being aggregated. The ability to distribute ARP tables across the network may prove a more efficient design. There are no absolute numbers to tell you whether your network should centralize all routing or distribute the routing, but guidelines are provided in the design recommendations below.

A Layer 2 edge solution, when combined with a strong distributed intelligence features, is easier to implement, administer, and troubleshoot. In addition, sub-second failover and no penalties on performance make Layer 2 the clearly superior choice in the Converged Campus design.

Avaya, however, recognizes that a Layer 2 solution is not always possible or may not fit every network design. The Avaya edge switch portfolio includes products that support a Layer 3 edge into a Layer 3 core/distribution. There are no performance penalties for implementing Layer 3 at the edge. The switches provide outstanding performance whether implemented as Layer 2 or Layer 3. The main difference is seen in the complexity of laying out the Layer 3 design and the ongoing administration and troubleshooting of such a network.

In summary, Avaya provides the flexibility for both approaches. Some customers choose a Layer 3 edge design solution for various reasons – no VLAN propagation, same configuration replicated, smaller broadcast domains, security/access control lists (ACL), for example – and they have the necessary routing expertise to support such a network. Other customers prefer a centralized routing and filtering/ACL approach, which may reduce the overall complexity of the network administration by not distributing Layer 3 throughout the network.

2. Large Campus Design

The Large Campus design is intended to support a network with more than 2500 network devices. The lower limit of 2500 is not a hard number, but rather a general figure upon which to base designs. Please take note that these numbers are network attached devices such as PCs, IP phones, printers, access points, etc. and **_not_** users. Attempting to base a network design on users is becoming increasingly difficult as more devices are being converged onto the infrastructure; therefore, recommendations are based on network-attached devices.

The **_Large Campus Solution_** includes the following key components along with design and best practice recommendations for:

> ➢ Ethernet Switching Infrastructure

> ➢ Network Access Control

> ➢ Network Management

ERS 5000 Stackable
- **High End Stackable**
- **10/100/1000/10Gig**
- **PoE**

ERS 4500 Stackable
- **Mid-Range Stackable**
- **100FX, 10/100, 10/100/1000/10Gig**
- **PoE**

ERS 2500 Stackable
- **Entry Level Stackable**
- **10/100**
- **PoE on Half the ports**

ERS 8300 Modular
- **High Density Modular**
- **10/100/1000/10Gig**
- **PoE**

ERS 8800/8600 Switch Cluster Core
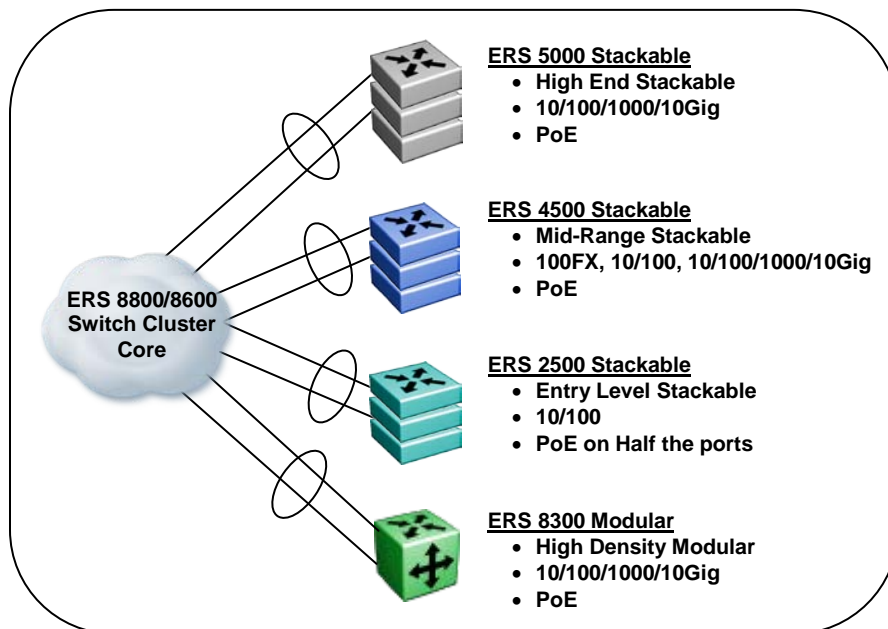
Figure 2.1: Large Campus Ethernet Infrastructure

The following business requirements are taken into consideration when selecting the products used in this specific solution:

> ➢ Effective and Efficient Edge Switching

> ➢ Scalability

> ➢ Cost-effective Without Compromising Performance

> ➢ High Availability

> ▪ Edge with Resilient Stacking

> ▪ Switch Clustering Core

➢ Simple to Build and Run

➢ Energy Efficiency

To meet these business requirements, the Large Campus Design will use the following platforms:

➢ Ethernet Routing Switch 8800/8600 at the core

➢ Ethernet Routing Switch 2500 / 4500 / 5000 / 8300 at the edge

➢ Network Access Control with Identity Engines

➢ Network Management
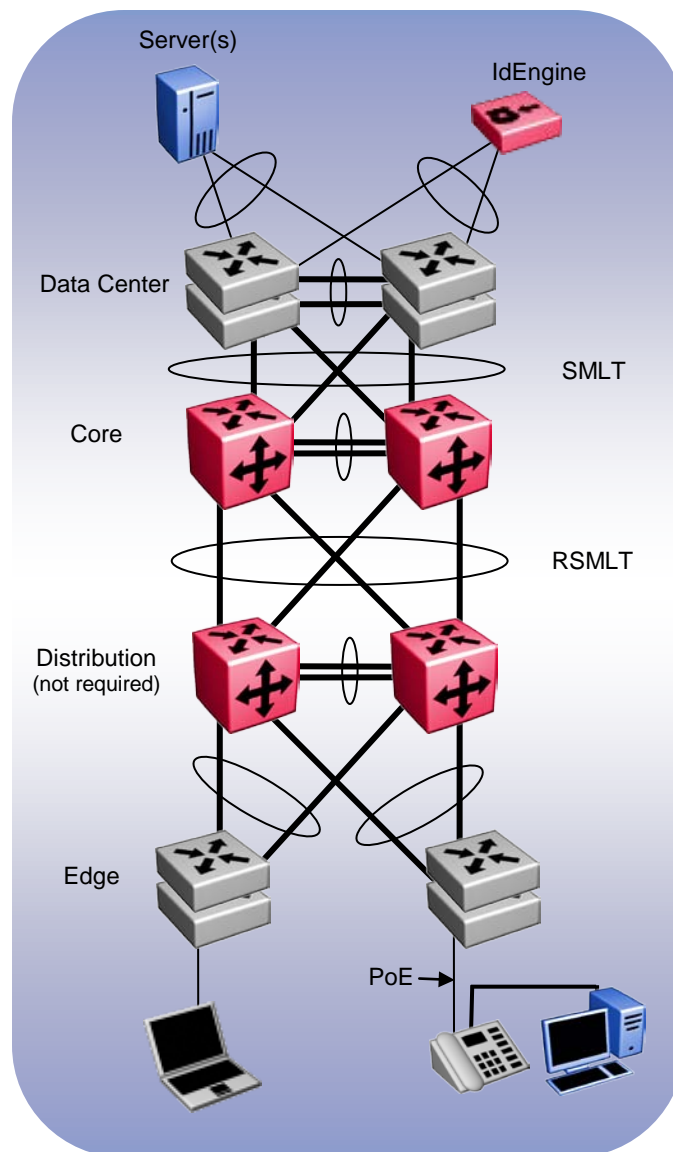
Figure 2.2: Large Campus Design Topology Example

2.1 Core Switching

The ERS 8800/8600 platform serves as the Core switching platform for the Large Campus solution. The features and functionality highlighted here represent the basic requirements for the Large Campus. Please refer to the product documentation for a detailed explanation of these and all the other features of the ERS 8800/8600 platform.

- ➢ Core Switching Hardware
- ➢ Software Licenses
- ➢ Switch Clustering
- ➢ VLACP
- ➢ SLPP
- ➢ Filter Untagged Frames
- ➢ Spanning Tree

- ➢ VLANs
- ➢ DHCP Relay
- ➢ Quality of Service
- ➢ Layer 3
- ➢ VRRP with Backup Master
- ➢ Server Connectivity
- ➢ Multicast

2.1.1 Platform Redundancy

The core-switching layer is the most critical component of the Converged Campus design and therefore it is imperative to ensure the most redundancy and resiliency possible. Scalability from the low-end to the high-end is imperative to provide the most cost-effective solution possible without compromising the design goals stated earlier.

Providing redundancy in the hardware is the basic building block to creating the highly resilient core and therefore the platforms deployed must provide redundancy in hardware components. These redundancy features will vary between chassis-based and stackable/standalone products. This does not imply that one platform is more or less redundant than any of the other platforms; it simply states that the way redundancy is achieved will vary based on the hardware being used.

The ability to hot-swap any and all hardware components is an absolute necessity and should be one of the highest criteria when evaluating hardware platforms.

For chassis systems, key hardware redundancy components include; power supplies, fan trays, switch fabric/CPUs, and I/O modules. For stackable/standalone systems, the key hardware redundancy components include; external/internal redundant power supplies, resilient stacking architecture and the ability to hot-swap any switch without interrupting traffic flow from/to other switches in the stack.

2.1.2 Core Switching Hardware

The ERS 8800/8600 product portfolio offers a chassis-based solution with many different I/O module options to fit the need of the Large Campus Core. Being able to accommodate high-density Gigabit connections (copper or fiber), along with 10 Gigabit offer the flexibility needed for most core switching solutions.

ERS 8800/8600

- ➢ 3 slot, 6 slot, 10 slot, or 10 slot CO chassis
- ➢ Dual switching fabrics
- ➢ N+1 power supply redundancy (AC and DC options available)
- ➢ All components fully hot swappable
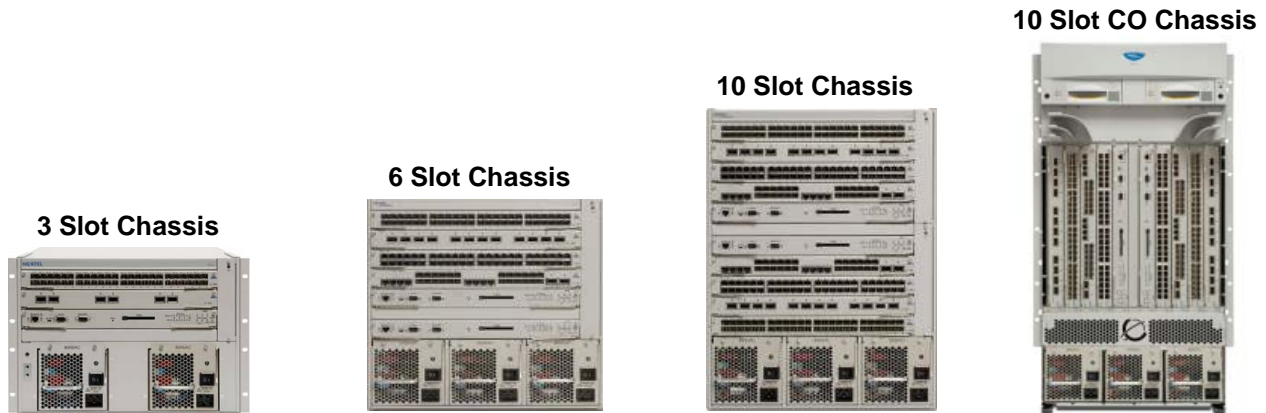
Figure 2.3: ERS 8000 Chassis Options

Module	Ports	Type
8895SF	0	256Gbps Switch Fabric with 10/100/1000 OOB mgmt port
8692SF w/ Mezz	0	256Gbps Switch Fabric with 10/100 OOB mgmt port
8648GTRS	48	48 port 10/100/1000BaseT
8648GBRS	48	48 port Gigabit SFP baseboard
8634XGRS	34	8 port 10/100/1000BaseT, 24 SFP, 2 XFP (LAN phy)
8612XLRS	12	12 port 10Gigabit XFP baseboard

8612XLRS   **8648GBRS**   **8648GTRS**   **8634XGRS**   **8895 SF/CPU**

Table 2.1: ERS 8800 Modules

The ERS 8800/8600 I/O modules support a variety of pluggables:

  ➢ 100Mbps – 100FX

  ➢ Gigabit – SX, LX, XD, ZX, BX, CWDM, and copper

  ➢ 10 Gigabit – SR, LR, ER, ZR

2.1.3 ERS 8800/8600 High Availability Mode

From a software perspective, the ERS 8800/8600 supports High Availability (HA) mode when both Switch Fabric/CPU (SSF/CPU) modules are installed. With HA enabled, both CPUs are active. The CPUs exchange topology data, so in the event of an SSF/CPU failure, the functioning SSF/CPU can continue passing traffic with sub-second recovery. In the event of a failure of the master CPU, the backup CPU takes over system control with sub-second convergence and minimal or no interruption to user applications/traffic.

Data Synchronized	Release 4.0	Release 3.7/4.1	Release 5.0	Release 5.1	Release 7.0
L1/Port Configuration	Yes	Yes	Yes	Yes	Yes
Syslog	Yes	Yes	Yes	Yes	Yes
RMON	No	3.7.1	Yes	Yes	Yes
L2/VLAN Parameters	Yes	Yes	Yes	Yes	Yes
SMLT	Yes	Yes	Yes	Yes	Yes
Spanning Tree	Yes	Yes	Yes	Yes	Yes
802.3ad/802.1X	N/A	Yes	Yes	Yes	Yes
ARP Entries	Yes	Yes	Yes	Yes	Yes
Static/Default Routes	Yes	Yes	Yes	Yes	Yes
VRRP	No	Yes	Yes	Yes	Yes
RIP	No	Yes	Yes	Yes	Yes
OSPF	No	Yes	Yes	Yes	Yes
BGP	No	No	Yes *	Yes *	Yes *
VRF	No	No	Yes	Yes	Yes
BFD	No	No	No	Yes*	Yes*
MPLS, LDP, RSVP-TE	No	No	Yes *	Yes *	Yes *
Filters	No	Yes	Yes	Yes	Yes
L2 Multicast (IGMP)	Yes	Yes	Yes	Yes	Yes
L3 Multicast (MSDP)	No	No	No	Yes*	Yes*
L3 Multicast (PIM-SM, SSM)	No	No	Yes *	Yes *	Yes *
* Warm standby support – features can be configured with HA enabled, however, when a failover occurs, protocols will be re-started.					

Feature support in HA mode is dependent on the software version as show in Table 2.15. For more details, please refer to the ERS 8800/8600 Administration Guide (NN46205-605).

Table 2.2: ERS 8800/8600 High Availability Feature Support

2.1.4 Advanced and Premier Software License

The ERS 8600 with Release 5.0 and ERS 8800/8600 with Release 7.0 have adopted a software licensing format. There are three licenses available; the Base license (required with every chassis), the advanced license, and the premier license. The premier license encompasses all features within the premier category as well as the advanced license category. There is no upgrade license available to go from advanced to premier.

The Advanced Software License is required to enable the following features:

- ➢ Packet Capture (PCAP)
- ➢ IPv6 Routing
- ➢ BGP (more than 10 peers)
- ➢ MSDP
- ➢ BFD

The Premier Software License is required to enable the following features:

- ➢ All Advanced License features
- ➢ VRF-Lite
- ➢ MP-BGP
- ➢ IP VPN MPLS RFC 2547
- ➢ IP VPN-Lite
- ➢ Multicast Virtualization for IGMP, PIM-SM/SSM

The ERS 8800/8600 has a built-in trial period for these features. For the first 60 days after upgrading to a major release, the features are available. After that time period expires, a valid license will be required for these features.

For details on software licensing, see the *Converged Data Networks Licensing Guide* (NN48500-540).

2.1.5 Switch Clustering using Split MultiLink Trunking (SMLT)

Switch Clustering using Split MultiLink Trunking (SMLT) provides industry-leading technology for the resiliency of the Converged Campus design. Providing redundant links that are forwarding traffic with no Spanning Tree allows the ultimate design in a converged environment. Sub-second failover and the simplicity of a network without Spanning Tree reduces TCO and ensures converged applications will function flawlessly. A vital feature of Switch Clustering is its ability to work with any end device (3rd party switch, servers, etc.) that supports a form of link aggregation.

Switch Clustering also provides the ability to perform virtual hitless upgrades of the core switches (cluster). With all connections to the cluster dually attached, a single core switch can be taken out of service without interrupting end user traffic. This switch then can be upgraded and brought back into service. By performing the same function on the other switch, after the upgraded switch is back online, the entire cluster can be upgraded without taking a service outage and with minimal interruption to traffic flows on the network.

2.1.6 Switch Clustering Terminology

There are different design options to be considered with the deployment of Switch Clustering:

  ➢ Single Link Trunking (SLT)

  SLT is a port-based option allowing large-scale deployments of SLT from a single Switch Cluster. Every port (saving at least two for the IST) can be used for SLT groups terminating into the cluster, with each SLT group consisting of a maximum of two uplinks (one per core Ethernet Routing Switch). For most typical deployments, the ability to have two connections per edge switch/stack is more than sufficient bandwidth, and allows a single cluster to handle many environments. The flexibility of the Avaya edge switch solutions allows for uplinks ranging from 10 Mbps to 10 Gbps (uplinks within the same SLT group must be of the same media type and link speed).

  ➢ Split MultiLink Trunking (SMLT)

  The MLT-based SMLT option allows for increased scaling of the number of links within a single SMLT group. The number of links supported in an SMLT group is the same number of MLT links supported on the Ethernet Routing Switch platform being used for the Switch Cluster. The SMLT links can be spread across the Switch Cluster – usually in an even dispersion, but this is not an absolute requirement. One MLT group must be used to create the IST between the two switches used to form the Switch Cluster.

Both SLT and SMLT can be configured on the same Switch Cluster.
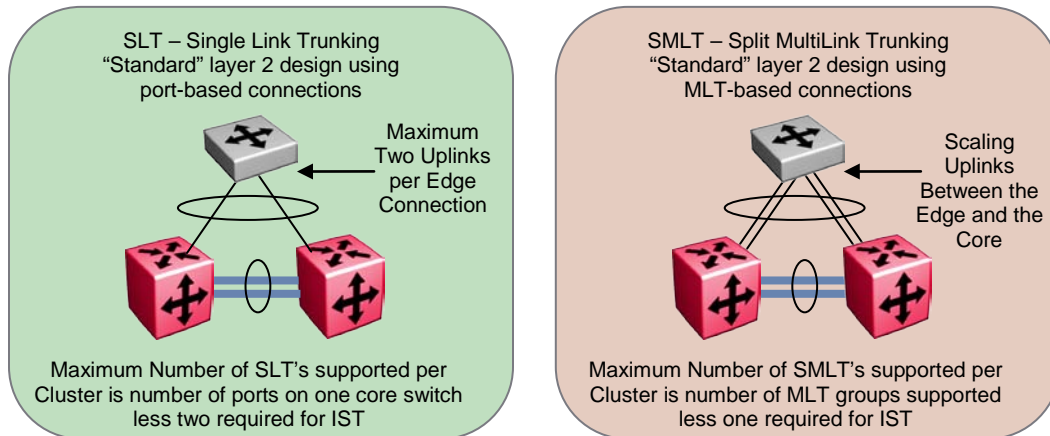
Figure 2.4: SLT and SMLT Terminology

Switch Model	Links per MLT Group	MLT Groups per Switch or Stack	MLT-based SMLT Groups			Port-based SLT Groups		
			Copper	Fiber (1GbE)	Fiber (10GbE)	Copper	Fiber (1GbE)	Fiber (10GbE)
ERS 8600 Legacy Modules	8	32	31	31	31	382	238	22
ERS 8800/8600 R and RS Modules	8	128	127	127	127	382	238	22
ERS 8300	4	31	30	30	30	382	398	67
ERS 5000	8	32	31	31	31	398	190	62

Note: Advanced Software License required on ERS 8300 and ERS 5000 for Switch Clustering (SMLT/SLT)

Table 2.3: MLT/SMLT/SLT Scaling Capabilities

2.1.7 Switch Clustering Topologies

There are three supported topologies with Switch Clustering. The use of each of these topologies will depend on the overall design of the network.

➢ Triangle – Single Switch Cluster at the core with the edge directly connected via SLT or SMLT

➢ 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. Full Mesh topologies can be scaled with additional pairs of Switch Clusters.

The following sections highlight the supported topologies that can be used with the ERS 8800/8600 as the Switch Cluster core.

2.1.7.1   Triangle Switch Cluster

The triangle Switch Cluster is comprised of a single ERS 8800/8600 switch for each IST peer. This configuration can terminate Edge closet connections using MLT-based SMLT or port-based SLT.
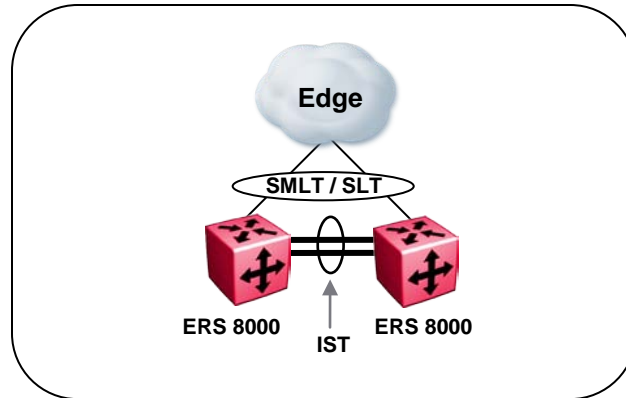
Figure 2.5: Triangle Switch Cluster

2.1.7.2   Square / Full Mesh Switch Cluster

The square or full mesh Switch Cluster extends the scalability of the ERS 8800/8600 by connecting Switch Cluster Cores using SMLT or RSMLT. This configuration can terminate Edge closet connections using MLT-based SMLT or port-based SLT. All rules from triangle configurations still apply
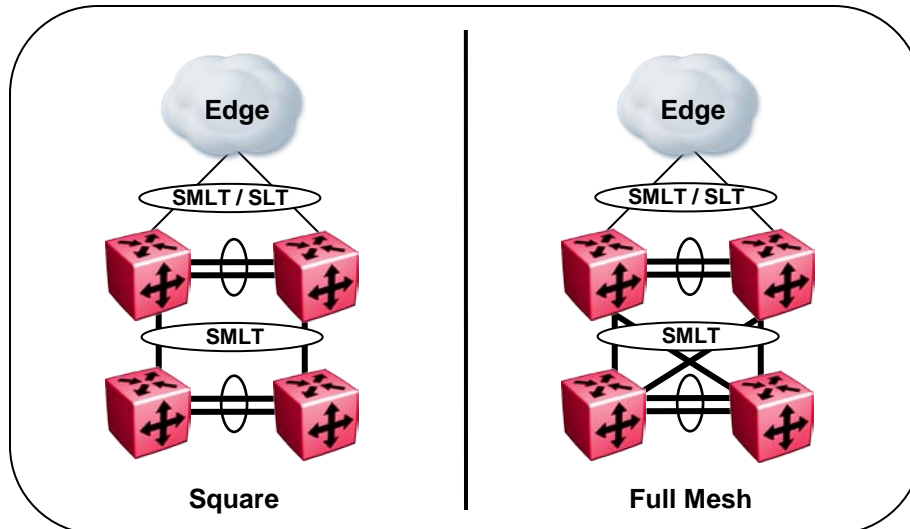
Figure 2.6: Square / Full Mesh Switch Cluster Topologies

The square or full mesh Switch Cluster provides the basic infrastructure to overlay various logical designs. Whether running all Layer 2, all Layer 3, or a combination of Layer 2 on one Switch Cluster and Layer 3 on the other Switch Cluster, this design accommodates those needs.

Interoperability between Switch Clusters is also available. For a detailed description of the various topologies supported by Avaya, please refer to Switch Clustering Supported Topologies and Interoperability with ERS 8800/8600/5000/8300/1600 (NN48500-555).

2.1.8 Switch Clustering Reference Architecture

In order to easily identify different aspects of the Switch Cluster design, the following reference architecture will be used throughout the discussion on best practice recommendations. The following diagram depicts a six-switch core for completeness, showing the triangle, square and full mesh topologies. Please note that this is not a requirement for implementing Switch Clustering.

> ➤ Access SLT/SMLT's are connections from the core out the edge closets and are normally in a standard triangle configuration.

> ➤ Core SMLT/RSMLT connections exist between Switch Clusters and can be formed using either the square or full mesh topologies. SMLT connections are used for the core so that bandwidth can easily be increased by adding another connection to the MLT group that forms the SMLT.

The major difference between the Access and the Core will be in the Loop Prevention mechanisms recommended for each. The core is obviously more critical to the overall network and is also a much more controlled environment; therefore, the best practice implementation will differ between the access and the core. A more detailed discussion on these techniques will follow in the coming sections.
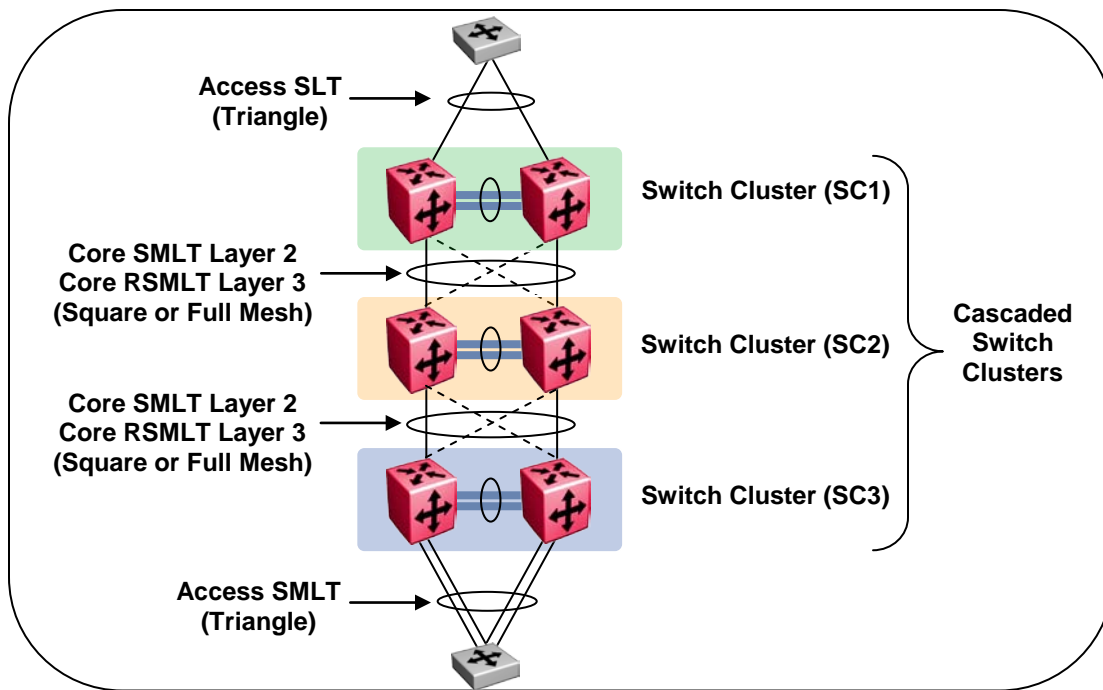
Figure 2.7: Switch Clustering Reference Architecture

2.1.9 Two Tier versus Three Tier Architecture

When designing a Converged Campus solution, there are two major topologies that can be implemented. The two-tier architecture, in which all edge switches terminate into the core of the network, and the three tier architecture, in which the edge switches terminate into a distribution layer network. The distribution layer network then terminates into the core. Three-tier architecture is usually required when the existing cable plant cannot support a two-tier deployment because of fiber distances or physical layout of the fiber.

From a switching/routing perspective, there are two options to be considered: either Layer 2 at the edge with Layer 3 in the core/distribution, or Layer 3 at the edge with Layer 3 in the core/distribution. Avaya provides Ethernet switching platforms that can provide either design alternative. There is no right answer for all possible designs; however, the Avaya design philosophy is always to keep the architecture simple without compromising resiliency and scalability. This translates into easier management and an overall lower total cost of ownership (TCO) by centralizing routing in the core and distributing intelligence across the network.

Avaya recommends deploying two-tier architecture whenever possible. This simplifies the network, reduces the amount of equipment required, and does not compromise scalability and resiliency. The two-tier architecture supports either Layer 2 or Layer 3 at the edge. Avaya recommends keeping Layer 2 VLANs at the edge and routing between VLANs at the core.

If three-tier architecture is deployed, Avaya recommends using Layer 3 between the distribution and core layers, utilizing RSMLT when possible for these connections. The same rules apply to the connections between the access and distribution layers (for less than 3000 users, use Layer 2; for more than 3000 users, use Layer 3).

With any of these options, it is critical to deploy an end-to-end QoS strategy to ensure that mission-critical applications are able to provide the required quality of experience for the users. A detailed discussion of QoS is covered in a later section of this document.

2.1.10 Two Tier Design – Core to Edge

With the basic two-tier design, the edge switches connect directly into the core. In the Converged Campus, the core is a Resilient Switch Cluster consisting of a minimum of two Ethernet Routing Switches with sufficient port density to accommodate dual homing of all edge switches.

An Inter-Switch Trunk (IST) ties the pair of Ethernet Routing Switches together to form the Resilient Switch Cluster. The IST is a critical component of the Switch Cluster and therefore must be highly resilient. The architecture of Switch Clustering and the traffic flow through the cluster is such that there is not a high volume of traffic across the IST, so resiliency of the connection is more important than total bandwidth.

The architecture is very flexible and can accommodate most design scenarios. The standard recommendation is to have Layer 2 at the Access and Layer 3 at the core. This architecture does not preclude the ability to extend Layer 3 to the Access if that is desired.
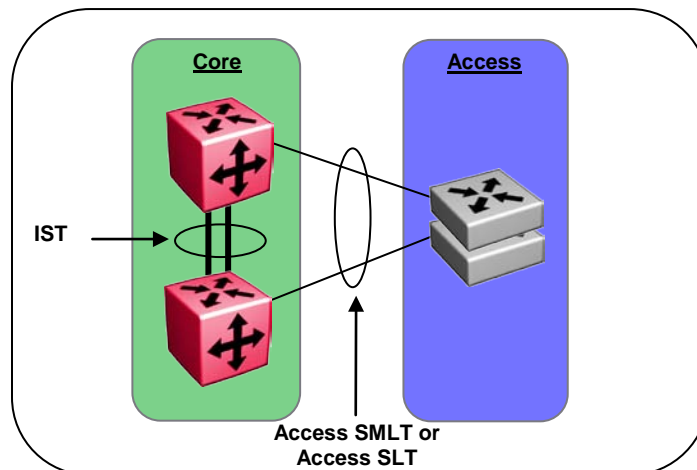
Figure 2.8: Switch Clustering – Two Tier Architecture

2.1.11 Three Tier Design – Core to Distribution to Edge

With a three-tier design, a distribution layer is inserted between the edge and the core. The need for a distribution layer can be attributed to existing physical infrastructure such as fiber plant layout or the requirement to connect multiple buildings on a single campus together – in that case, the building cores would be the distribution layer and connect back to a centralized core.

In situations where three tiers are necessary, there are options on the deployment of Switch Clustering for resiliency between the layers. The edge to distribution considerations are the same as described in the above section. Between the distribution and core layers, there are different options available based on the architecture deployed:

>   Layer 2 between Distribution and Core – SMLT

In the attempt to centralize routing functionality and distribute the intelligence throughout the network, it is easy to keep a simple Layer 2 architecture between these two layers of the network. In this design, the distribution to core connectivity mimics that of the edge to the core described in the above section. The main difference lies in the ability to fully mesh the distribution to the core. A fully meshed solution provides the highest level of resiliency possible with still maintaining sub-second failover and recovery. A square or full mesh is mandatory to maintain full resiliency and bandwidth between distribution and core.

➢ Layer 3 between Distribution and Core – Routed SMLT

If routing is desired between the distribution and core layers, deploy routed SMLT (RSMLT) to maintain sub-second failover and recovery while running a standard IGP routing protocol such as RIP or OSPF. RSMLT builds on the SMLT technology by providing an active-active router concept to SMLT networks with routing enabled on the core VLANs. In the case of a routing switch failure, RSMLT takes care of packet forwarding at Layer 2 while the routing protocol converges at the Layer 3 level. This allows the non-stop forwarding of traffic in the event of any failure with no disruption to the user. Another huge advantage of RSMLT is the ability to extend Layer 2 subnets – something that is not possible if strictly using Layer 3 routing between the core and distribution.
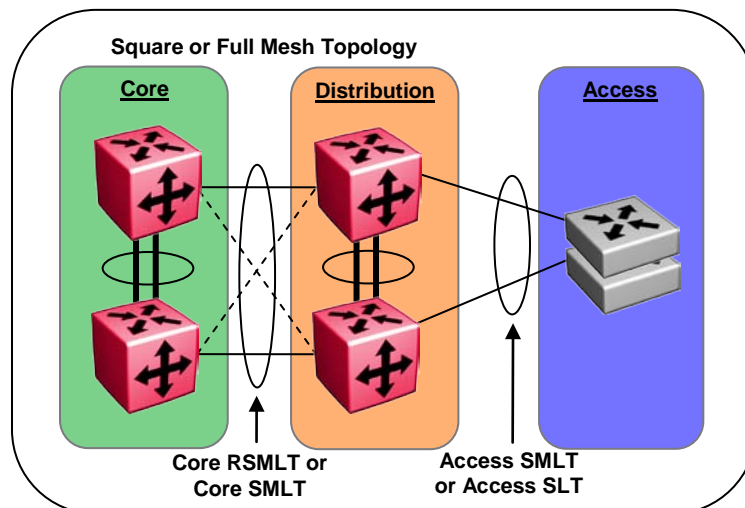
Figure 2.9: Switch Clustering – Three Tier Architecture

2.1.12 Switch Cluster Core Configuration Guidelines

When configuring Switch Clustering on the ERS 8800/8600 review the following:

- ➢ IST is two port MLT (minimum) and up to eight port MLT (maximum)

  - ▪ The number of links required for the IST is based on the amount of bandwidth required during a failure scenario. The amount of traffic that traverses the IST during normal operations is minimized with the Switch Clustering architecture and all connected devices being dual-homed.

  - ▪ Ports within the MLT must be same speed

  - ▪ Mixed media MLTs (copper and fiber) are supported for the IST

  - ▪ It is strongly recommend to distribute MLT connections between modules in the chassis for added resiliency (DMLT – distributed MLT)

  - ▪ Ports assigned to an MLT (IST) are indexed by a number starting at zero (0). The lowest port position (module 1 port 1) for an MLT link is assigned an index of zero. The next MLT link in the second lowest position gets an index of one (1) and so on. This index is used by the MLT algorithm to assign a flow over a particular MLT link. Therefore, Avaya recommends that you mate the lowest port position of one MLT link in one switch with the lowest port position of the peer switch. Follow this rule for all successive MLT links. This will help to ensure that the MLT algorithm always resolves a flow over the same link between the two switches.

  - ▪ It is recommended that the IST links terminate on non-blocking ports in each chassis whenever possible. This will ensure that all critical management traffic will be received properly by each switch regardless of the utilization of the IST links. At the same time, having the IST comprised of oversubscribed ports is supported and is sometimes the only option available. With multiple ports comprising the IST and the fact that multiple failures would likely be required to overrun the IST, these designs will work in most cases, however, it is important to understand potential limitations of such designs.

  - ▪ Do not use the IST IP addresses as next hop addresses for any static routes

- ➢ If only two uplinks are required per edge closet, use port-based Single Link Trunking (SLT). The number of SLT's configurable equals the number of ports in the switch less the number of ports which are used for the IST.

  - ▪ Begin SLT ID's at 129

- ➢ If more than two uplinks are required per edge closet, use MLT-based SMLT. The number of SMLT's configurable is 128 in R-mode and 32 in mixed mode. One MLT-based SMLT will be used for the IST.

  - ▪ Use MLT 1 for the IST.

  - ▪ SMLT ID's 2 to 128 and should correspond with MLT ID – although this is not required it is highly recommended to simplify the design and troubleshooting in the future

- ➢ Both SLT and SMLT connections can be configured on the same Switch Cluster simultaneously.

> It is possible to overlap the ID numbers when using MLT-based SMLT and port-based SLT, Avaya recommends avoiding this and follow the recommendations in Table 2.4.

Switch Model	Software Version	MLT-based SMLT ID's	Port-based SLT ID's
ERS 8600 Legacy Modules	3.x and higher	1-32	33-512
ERS 8800/8600 R and RS Modules	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
ERS 1600	2.1 and higher	1-7	8-512

Table 2.4: SMLT ID Recommended Values

> Create a separate IST VLAN and use a private address space for the IST VLAN IP addresses with a small subnet mask (i.e. 30-bit mask). This VLAN is only required for IP communications between IST peers.

> - To simplify network management, OSPF with passive interfaces can be enabled on the IST VLAN

> Verify that all VLANs participating in SLT/SMLT are configured on both IST peer switches and are tagged on both ends of the IST.

> All SLT/SMLT uplinks shall be 802.1Q tagged, as this will easily facilitate adding additional VLANs to the edge without impacting traffic. This will also ensure that any switches added to the network in default configuration will not cause loops, as untagged frames will be discarded at the Switch Cluster core – see details that follow in the sections on Filter Untagged Frames.

> STP will automatically be disabled on all ports participating in SLT/SMLT on the ERS 8800/8600, this includes both the IST and SLT/SMLT ports – make sure to also disable STP on the edge switch uplinks, as this will not be done automatically.

> If multicast routing (PIM-SM) is enabled on the Switch Cluster (ERS 8800/8600 and ERS 8300), enable PIM-SM on the IST VLAN to insure fast recovery of multicast traffic.

> When configuring a Core SMLT square or full mesh (SMLT between two pairs of Switch Clusters), use the same SMLT ID on both sides of the square/mesh for operational simplification.

> The ERS 8800/8600 supports SMLT while using a single Switch Fabric/CPU module in the chassis. If this configuration is required, ensure that the hardware I/O modules (specific to E-modules) are at the correct hardware revision to support this feature and enable the single CPU SMLT feature in the software. This ensures SMLT ports on the I/O modules are disabled if the Switch Fabric/CPU module fails or is removed from the chassis. There are no hardware restrictions when using R or RS modules.

2.1.13 VLANs

VLANs provide an easy mechanism for traffic separation, a way to minimize the size of broadcast domains, and can help isolate different protocols from each other. In most cases, the VLAN is considered equal to a broadcast domain; for example, a specific IPv4 subnet is assigned to a single VLAN. From an administrative point of view, VLAN to subnet mapping makes each very easy to identify quickly. The number of VLANs and the type of VLANs deployed can vary greatly from design to design. Consider these points when creating the VLAN strategy:

> The need for isolation of different protocols on the network

> VLAN types to be used – port, protocol, MAC, subnet

> Traffic separation for voice and data

> Size of the broadcast domain/number of users

> VLANs by geographic area – per closet, per floor, per building

> Network services required per VLAN – DHCP, UDP forwarding, etc.

All Voice and Data VLANs will be configured on both core ERS 8800/8600 switches. Any other services that will be attached directly to the Switch Cluster core will also need VLANs as well as the IST.

> By default, VLAN 1 is created and all ports are members – this VLAN cannot be deleted. Do not use this VLAN for any production traffic, but only as a repository for unused ports.

> 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) for the ERS 8800/8600. This reduces the amount of re-ARPs that need to occur.

> Any VLAN with an IP Address can be used to manage the switch; however, to simplify management of the switch, create a management VLAN.
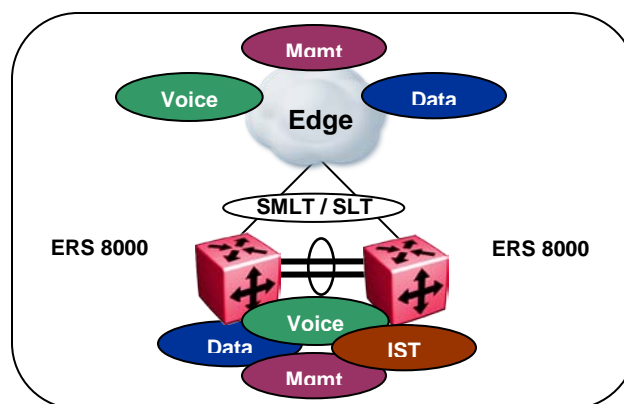
Figure 2.10: ERS 8000 VLANs

Order of precedence for identifying packets ingressing the switch

| Hardware Platform | Highest Priority ────────────▶ Lowest Priority | | | | |
	802.1Q Tagged Packet	Subnet Based VLAN	Protocol Based VLAN	MAC Based VLAN	Port Based VLAN
ERS 8800/8600	✓	✓	✓	✓	✓
ERS 8300	✓	N/A	✓	N/A	✓
ERS 5000	✓	N/A	✓	N/A	✓
ERS 1600	✓	N/A	✓	N/A	✓

Table 2.5: VLAN Support

2.1.14 Discard Untagged Frames

The recommendation to always enable 802.1Q VLAN tagging on uplink ports is critical to using the discard untagged frames feature. This provides protection against a factory defaulted or incorrectly configured device being connected to the ERS 8800/8600. The core will automatically drop all packets that are not 802.1Q tagged, adding another level of protection against potential loops.

> The Discard Untagged Frames feature should be enabled on the IST/SLT/SMLT ports
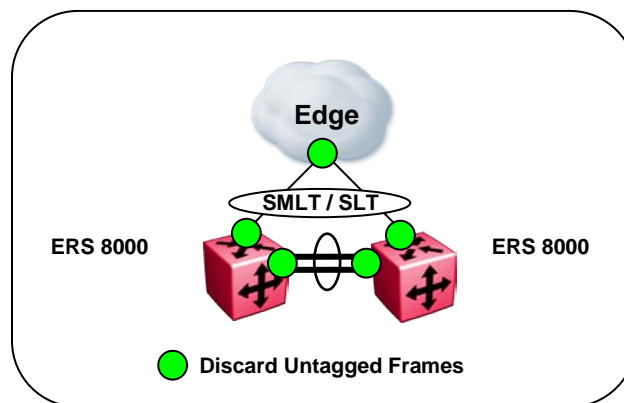
Figure 2.11: ERS 8000 Discard Untagged Frames

2.1.15 Spanning Tree

Spanning Tree is automatically disabled on all IST / SLT / SMLT ports on the ERS 8800/8600 Switch Cluster Core in order for Switch Clustering to function properly. Avaya recommends leaving all the other ports as spanning tree enabled to insure protection from a hub or other network device causing a loop in the core.
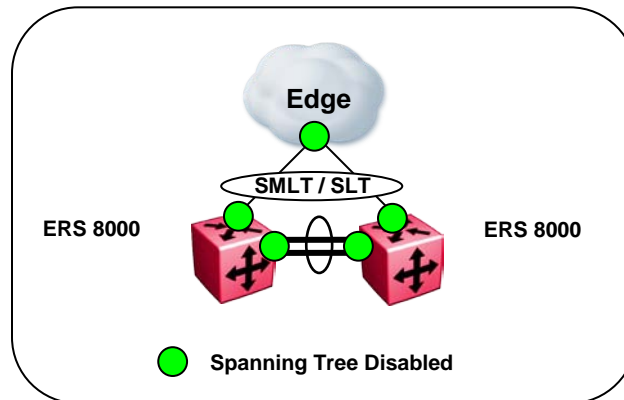
Figure 2.12: ERS 8000 Spanning Tree

2.1.16 Control Plane Rate Limit (cp-limit)

Control plane rate limit (cp-limit) controls the amount of multicast control traffic, broadcast control traffic, and exception frames that can be sent to the CPU from a physical port (i.e. OSPF hello, RIP update, etc.). It protects the CPU from being flooded by traffic from a single, unstable port. This differs from normal port rate limiting which limits non-control multicast traffic and non-control broadcast traffic on the physical port that would not be sent to the CPU (i.e. IP subnet broadcast, etc.). The cp-limit feature is configured on a per-port basis within the chassis.

The CP-Limit default settings are:

  ➢ Default state is enabled on all ports

  ➢ When creating the IST, cp-limit is disabled automatically on the IST ports

  ➢ Default multicast packets-per-second value is 10,000

  ➢ Default broadcast packets-per-second value is 10,000

If the actual rate of packets-per-second sent from a port exceeds the defined rate, then the port is administratively shut down to protect the CPU from continued bombardment. An SNMP trap and a log file entry are generated indicating the physical port that has been shut down as well as the packet rate causing the shut down. To re-activate the port, one must first administratively disable the port and then re-enable the port.

Having cp-limit disable IST ports in this way could impair network traffic flow, as this is a critical port for SMLT configurations. Avaya recommends that an IST MLT contain at least 2 physical ports, although this is not a requirement. Avaya also recommends that CP-Limit be disabled on all physical ports that are members of an IST MLT – this is the default configuration. Disabling CP-Limit on IST MLT ports forces another, less- critical port to be disabled if the defined CP-Limits are exceeded. In doing so, you preserve network stability should a protection condition (CP-Limit) arise. Although it is likely that one of the SMLT MLT ports (risers) would be disabled in such a condition, traffic would continue to flow uninterrupted through the remaining SMLT ports.
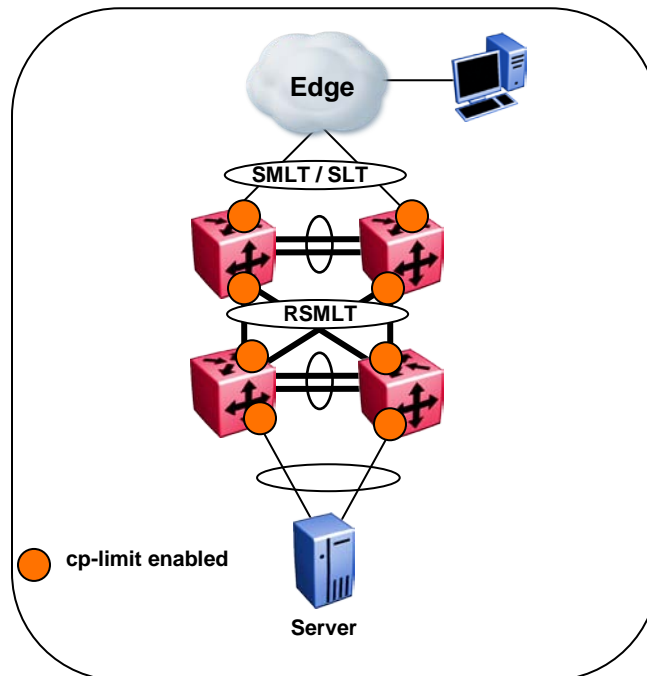
Figure 2.13: cp-limit Recommendations

- ➢ Leave enabled on all SLT/SMLT ports in the distribution/core

- ➢ Disable for all ports in the IST – ports participating in the IST should never be shut down under any circumstances

- ➢ In multi-tiered core environments, Avaya recommends that edge closet switches have cp-limit values less than the values used on the core links. This way, if an offending device does transmit malicious traffic, the edge switches will get triggered because of lower values, thus preventing the important core links from shutting down. This will also aid in isolating problems.

- ➢ Please note that cp-limit will only work on ports which belong to VLANs that have Layer 3 enabled.

For edge and server connected ports, if the connected device is determined to produce traffic to the levels for which cp-limit is configured, the connected port will be disabled when it starts transmitting.

	Recommended cp-limit Values	
	Broadcast	Multicast
Aggressive		
Workstation	1000	1000
Server	2500	2500
Non-IST Interconnection	7500	7500
Moderate		
Workstation	2500	2500
Server	5000	5000
Non-IST Interconnection	9000	9000
Relaxed		
Workstation	4000	4000
Server	7000	7000
Non-IST Interconnection	10000	10000

Table 2.6: cp-limit Recommended Values

CAUTION - Altering cp-limit values from their defaults during normal network operation can cause the links to become disabled. Avaya strongly recommends that to obtain a baseline of the network traffic across the links, choose the right value, and apply.

2.1.17 Extended CP-Limit (Ext-CP-Limit)

The ERS 8800/8600 supports the ext-cp-limit feature which goes one step further than cp-limit by adding the ability to read buffer congestion at the CPU as well as port level congestion on the I/O modules. This feature will protect the CPU from any traffic hitting the CPU by shutting down the ports which are responsible for sending traffic to CPU at a rate greater than desired.

To make use of ext-cp-limit, configuration must take place at both the chassis and port level. The network administrator must predetermine the number of ports that should be monitored when congestion occurs. Ext-cp-limit can be enabled on all ports in the chassis, but when congestion is detected, ext-cp-limit will monitor the most highly utilized ports in the chassis. The number of highly utilized ports monitored is configured in the MaxPorts parameter as described below.

When configuring ext-cp-limit at the chassis level, the following parameters are available:

➢ MinCongTime (Minimum Congestion Time) sets the minimum time, in milliseconds, the CPU frame buffers can be oversubscribed for before triggering the congestion algorithm.

➢ MaxPorts (Maximum Ports) sets the total number of ports that need to be analyzed from the may-go-down port list.

➢ PortCongTime (Port Congestion Time) sets the maximum time, in seconds, a port's bandwidth utilization can exceed the threshold. When this timer is exceeded, the port is disabled – this parameter is only used by SoftDown.

> ➤ TrapLevel Sets the manner in which a SNMP trap is sent if a port becomes disabled.
>> ▪ None - no traps are sent (default value)
>> ▪ Normal - sends a single trap if ports are disabled.
>> ▪ Verbose - sends a trap for each port that becomes disabled.

When configuring ext-cp-limit at the port level, the following parameters are available:

> ➤ HardDown disables the port immediately once the CPU frame buffers are congested for a certain period of time.

> ➤ SoftDown monitors the CPU frame buffer congestion and the port congestion time for a specified time interval – the ports are only disabled if the traffic does not subside after the time has been exceeded. The network administrator can configure the maximum number of SoftDown ports to be monitored.

> ➤ CplimitUtilRate defines the percentage of link bandwidth utilization to set as the threshold for the PortCongTime – this parameter is only used by SoftDown.

The following figures detail the flow logic of the HardDown and SoftDown operation of ext-cp-limit.

System Congestion Detected

↓

MinCongTimer Starts Counting

↓

MinCongTimer Expires (default 3 sec)

Send a system Congestion Alarm

Shutdown Ports In HardDown
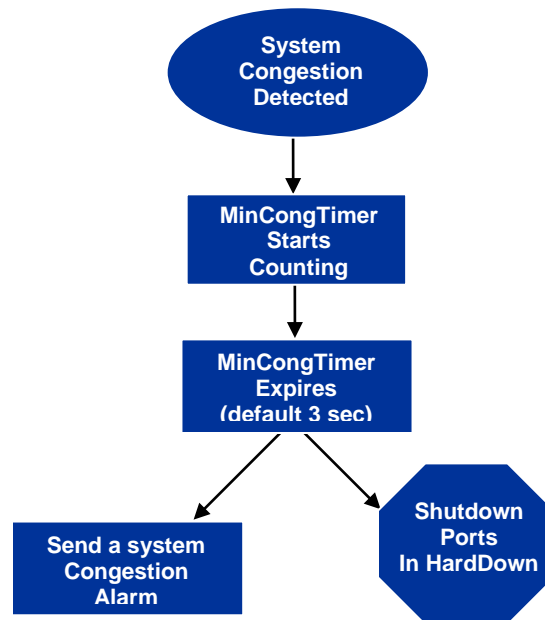
Figure 2.14: Ext-cp-limit HardDown Operation

System Congestion Detected

↓

MinCongTimer Starts Counting Down

↓

MinCongTimer Expires (default 3

↓

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

Continues until congestion subsides

Shutdown the ports which exceed their threshold + sends alarm

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

Figure 2.15: Ext-cp-limit SoftDown Operation

> Do not use the HardDown option of ext-cp-limit. VLACP and SLPP features are better and more complete and should be used in lieu of the HardDown option.

> Please note that ext-cp-limit will only work on ports which belong to VLANs that have Layer 3 enabled.

> Enable SoftDown option on all ports except IST on the ERS 8800/8600 Release 4.1 or later 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.

2.1.18 VLACP

Avaya has developed Virtual LACP (VLACP) that provides a true end-to-end failure detection mechanism between directly connected switches or connectivity across intermediary networks. This feature now adds a greater level of resiliency and flexibility to the Converged Campus design when used in conjunction with MLT, DMLT, and SMLT.

Please note that LACP and VLACP are two totally independent features and one does not require the other for implementation. LACP provides standards-based link aggregation capabilities and point-to-point failure detection, while VLACP provides end-to-end failure detection only.
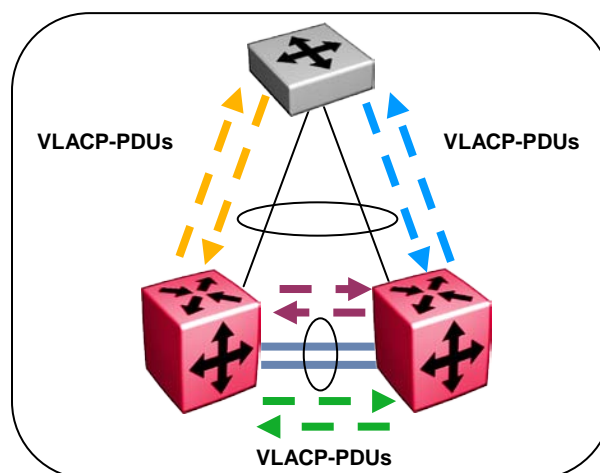
Figure 2.16: Virtual Link Aggregation Control Protocol (VLACP)

The following highlights the features associated with VLACP:

➢ Designed to operate end-to-end, regardless of whether the switches are directly connected or have an intermediary connection between them.

➢ VLACP is strictly a heart-beat end-to-end detection mechanism with no link aggregation capabilities.

➢ Global and port-based configuration parameters

➢ Very light load on CPU processing

➢ On each port that has VLACP enabled, VLACP PDUs are sent periodically. If VLACP PDUs are not received on a particular link, that link is administratively disabled after a configurable timeout period.

➢ Can run independently as a port-to-port protocol or on top of MLT, DMLT, and SMLT.

VLACP is a critical feature when deploying resilient networks with Switch Clustering. VLACP can detect end-to-end failures as described above and can also disable links to a switch that might still have link connectivity but cannot process traffic due to unexpected switch lockups. VLACP can also prevent loops in the network if uplinks are plugged into the wrong ports; VLACP is enabled only on uplinks, therefore, when the uplink is plugged into the wrong port without VLACP enabled, the link at the other end will be administratively disabled and no traffic will use that link.

When VLACP is used on directly connected switches, it is recommended to use a reserved multicast MAC address for the VLACP PDU. This reserved MAC address will not be flooded/forwarded by a switch that receives it. In the case where a factory-defaulted switch is connected, the VLACP packets would not be flooded back down to the core (giving a false positive that the links should be brought up), thus ensuring integrity of the network.

Enabling VLACP on the IST is critical for Switch Cluster designs. This will protect the integrity of the core in the case where one of the switches was inadvertently reset to factory default. Without VLACP on the IST, data packets are sent to the defaulted switch, causing loops and various other types of forwarding issues. By enabling VLACP on the IST, these packets would not be sent across, as the ports would be administratively down. Also, by having VLACP on all the SMLT/SLT ports, the edge switches would not send or receive packets from the defaulted switch, thus somewhat alleviating a much larger problem.

In the Large Campus design, VLACP will be used between the ERS 8800/8600 Switch Cluster and all uplinks to the ERS 2500/4500/5000/8300 switches at the edge as well as on the IST.

➢ Globally configure VLACP to use the reserved multicast MAC of 01-80-C2-00-00-0F

➢ For the IST links, use the long timeout

▪ Slow periodic timer of 10000 msecs * timeout scale of 3

▪ Make sure these values match on both ends of the links

➢ For the SLT/SMLT links, use the short timeout

▪ Fast periodic timer of 500 msecs * timeout scale of 5

▪ Make sure these values match on both ends of the links
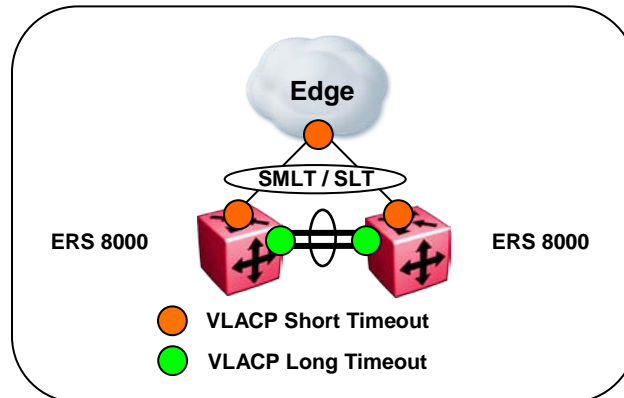
Figure 2.17: ERS 8000 VLACP

2.1.19 Simple Loop Prevention Protocol (SLPP)

Simple Loop Prevention Protocol (SLPP) provides active protection against Layer 2 network loops on a per-VLAN basis. SLPP uses a lightweight hello packet mechanism to detect network loops. SLPP packets are sent using Layer 2 multicast and a switch will only look at its own SLPP packets or at its peer SLPP packets. It will ignore SLPP packets from other parts of the network. Sending hello packets on a per VLAN basis allows SLPP to detect VLAN based network loops for un-tagged as well as tagged IEEE 802.1Q VLAN link configurations. Once a loop is detected, the port is shutdown. The SLPP functionality is configured using the following criteria:

➢ SLPP TX Process – the network administrator decides on which VLANs a switch should send SLPP hello packets. The packets are then replicated out all ports which are members of the SLPP-enabled VLAN. It is recommended to enable SLPP on all VLANs.

➢ SLPP RX Process – the network administrator decides on which ports the switch should act when receiving an SLPP packet that is sent by the same switch or by its SMLT peer. You should enable this process only on Access SMLT/SLT ports and never on IST ports or Core SMLT/SLT ports in the case of a square/full mesh core design.

➢ SLPP Action – the action operationally disables the ports receiving the SLPP packet. The administrator can also tune the network failure behavior by choosing how many SLPP packets need to be received before a switch starts taking an action. These values need to be staggered to avoid edge switch isolation – see the recommendations at the end of this section.
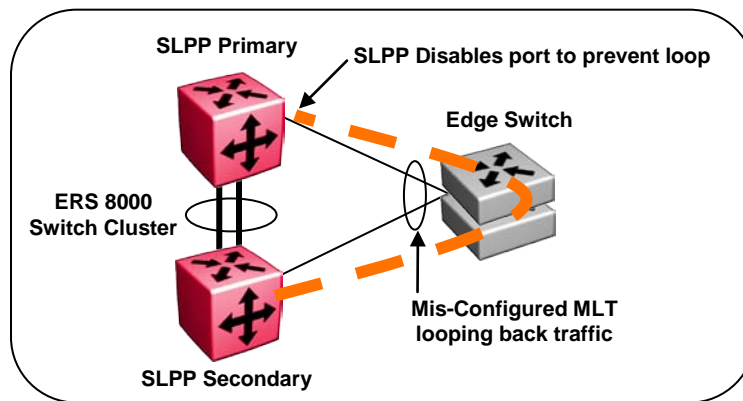
Figure 2.18: Simple Loop Prevention Protocol (SLPP)

Loops can be introduced into the network in many ways. One way is through the loss of an MLT configuration caused by user error or malfunctioning equipment. This scenario may not always introduce a broadcast storm, but because all MAC addresses are learned through the looping ports, does significantly impact Layer 2 MAC learning. Spanning Tree would not in all cases be able to detect such a configuration issue, whereas SLPP reacts and disables the malfunctioning links, limiting network impact to a minimum.

The desire is to prevent a loop from causing network problems while also attempting to not totally isolate the edge where the loop was detected. Total edge closet isolation is the last resort in order to protect the rest of the network from the loop. With this in mind, the concept of an SLPP Primary switch and SLPP Secondary switch has been adopted. These are strictly design terms and are not configuration parameters. The Rx thresholds are staggered between the primary and secondary switch, therefore the primary switch will disable an uplink immediately upon a loop occurring. If this resolves the loop issue, the edge closet still has connectivity back through the SLPP secondary switch. If the loop is not resolved, the SLPP secondary switch will disable the uplink and isolate the closet to protect the rest of the network from the loop.

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.

To accommodate different design requirements, the following table details the different options for SLPP RxThreshold values for edge connections:

	Recommended SLPP RxThreshold Values	
	Primary	**Secondary**
Aggressive	5	50
Moderate	2 times the number of VLANs Minumum SLPP value = 5 Maximum SLPP value = 100	10 times the number of VLANs Minumum SLPP value = 50 Maximum SLPP value = 500
Relaxed	3 times the number of VLANs Minumum SLPP value = 5 Maximum SLPP value = 100	15 times the number of VLANs Minumum SLPP value = 50 Maximum SLPP value = 500

Table 2.7: SLPP Recommended Values – Access Edge

In the Large Campus design, SLPP will be configured on the ERS 8800/8600 Switch Cluster(s) as described below:

➢ Designate one ERS 8800/8600 core switches as primary and the other as secondary. This is strictly from a design perspective. There are no actual configuration parameters to create the primary or secondary.

➢ Use default values for Ethertype and transmit interval

➢ Enable SLPP globally and on all VLANs on the Switch Cluster **except** for IST VLAN

➢ SLPP Rx should never be configured on IST ports

➢ Enable SLPP Rx (threshold of 5) on all non-SMLT/SLT ports in the ERS 8800/8600 except the IST ports as added protection against possible loops in the core

➢ Do not enable auto recovery – once the port is disabled by SLPP, it will need to be re-enabled manually after the loop has been fixed

➢ SLPP Rx-threshold is NOT reset upon any activity, but is a cumulative count. This can cause a situation in which multiple different loop events, can 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 counters.
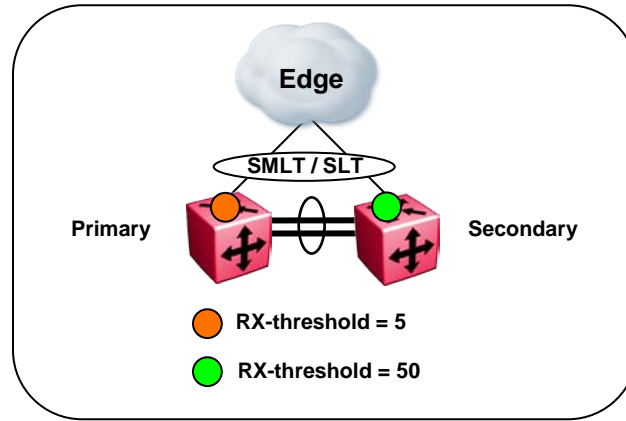
Figure 2.19: ERS 8000 SLPP in Triangle Topology

For Square and Full Mesh configurations that use a bridged core (Layer 2 VLANs extend from the edge through all switches in the core), it is recommended to enable SLPP on the primary switches as shown below. Enabling SLPP on half of the core will still prevent any possible loops and will not allow the possibility of the entire core being shut down by a loop at the edge of the network.

	Recommended SLPP RxThreshold Values		
	🟠	🟢	🟣
Aggressive	5	50	300
Moderate	2 times the number of VLANs (5-100)	10 times the number of VLANs (50-500)	400
Relaxed	3 times the number of VLANs (5-100)	15 times the number of VLANs (50-500)	500

Table 2.8: SLPP Recommended Values – Bridged Core
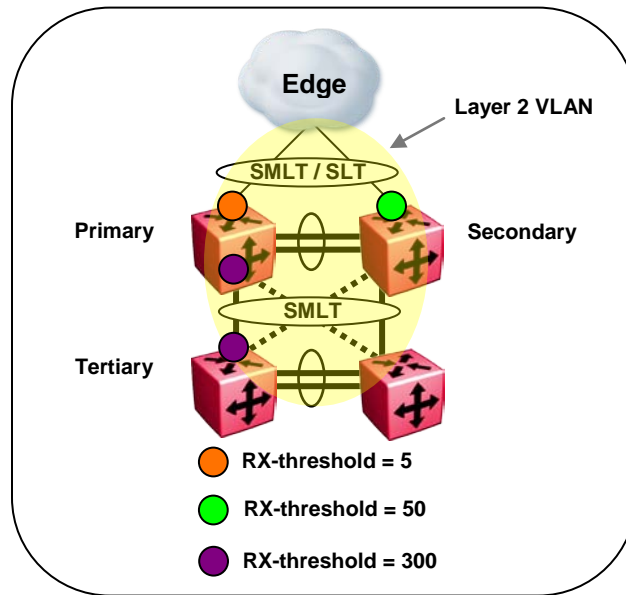
Figure 2.20: SLPP in Square/Full Mesh Bridged Core

For Square and Full Mesh configurations that use a routed core, it is recommended to use or create a separate core VLAN and enable SLPP on that VLAN and the square or full mesh links between the Switch Clusters. Loops created in the core will be caught and loops at the edge will not affect core ports. If using RSMLT between the Switch Clusters, then enable SLPP on the RSMLT VLAN.

Since SLPP will only be enabled on one or two VLANs in the core, changing the RX-threshold values will not likely be necessary.
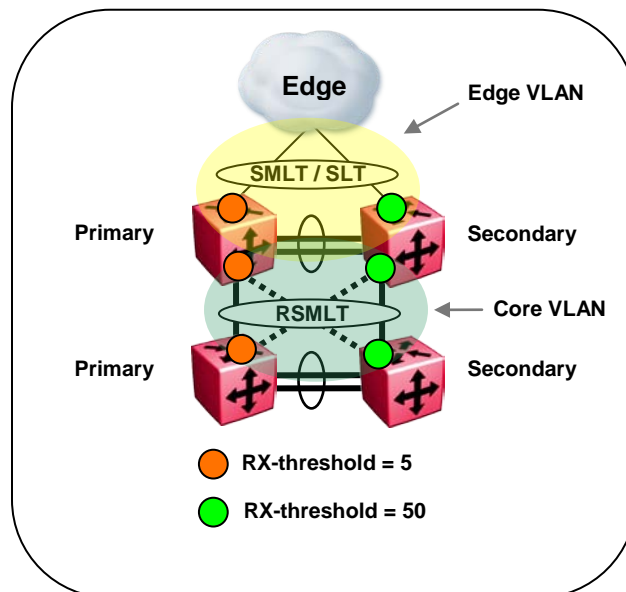
Figure 2.21: SLPP in Square/Full Mesh Routed Core

2.1.20 Quality of Service

Differentiated Services (DiffServ) is the industry and Avaya standard for the implementation of QoS. DiffServ provides QoS on a per hop behavior in the Ethernet switches by marking the header of individual packets with a DiffServ Code Point (DSCP). This DSCP then provides an indication to the Ethernet switch as to the priority of each packet and into which queue the packet should be placed. Note that the network infrastructure must support QoS end to end. Without a full end-to-end deployment, QoS cannot provide the necessary actions to ensure priority through every hop of the network. By default, the edge switches remark all QoS bits to zero and do not honor any markings.
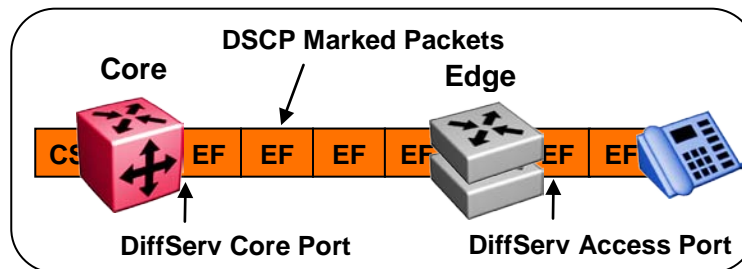
Figure 2.22: Quality of Service

Whether QoS needs to be deployed in the network depends on the applications and the business-critical nature of those applications. In order to deploy an effective QoS strategy within the Enterprise, it is imperative to understand the types of applications and the traffic patterns. For most installations, QoS is required for an IP Telephony deployment or any other application that is time/delay sensitive (e.g., video conferencing). QoS can also be employed for mission-critical applications such as Enterprise Resource Planning (ERP) tools. It is not necessary to provide QoS for every application on the network, only those that require special treatment. Note that even if there is sufficient bandwidth available in the network, QoS is still required for real-time applications to ensure minimum latency. It is also imperative to understand the potential impact of network component failures (uplinks, switches, modules) where an end-to-end QoS deployment will ensure the high priority traffic (business applications) will be forwarded and best effort (or lower priority traffic) discarded in the event of congestion.

There are various strategies for deploying QoS throughout the infrastructure, and Avaya provides several tools to streamline the implementation. The Ethernet switches have the ability to either mark or honor the DSCP within each Ethernet packet. Many end devices now have the ability to set their own DSCP and thus set their own priority across the network. For example, the Avaya IP phones set their DSCP to Expedited Forwarding for all voice traffic, which maps into the Premium Service Class queue.

Care should be taken when simply honoring the markings from end stations. Current Windows operating systems can also mark DSCP, and therefore savvy users could prioritize their traffic on the network that honors the DSCP marking. It is a better practice for the edge switch to re-mark the DSCP to one that is controlled by the network administrator. This can be accomplished by using VLAN prioritization, which marks all packets in a specific VLAN with the same priority (prioritizing the voice VLAN), or by using the filtering capabilities of the Ethernet switches to mark packets individually based on filtering criteria established by the network administrator.

To assist in qualifying these types of applications and the associated QoS levels, Avaya has created a QoS matrix and has standardized Avaya Service Classes across all platforms. This matrix is intended as a guideline for the implementation of QoS:

Avaya Service Class	Target Applications and Services	Tolerance to:		
		Loss	Delay	Jitter
Critical	Super user Telnet, Critical heartbeats between routers/switches	Very Low	Very Low	N/A
Network	ICMP, OSPF, BGP, RIP, ISIS, COPS, RSVP DNS, DHCP, BootP, high priority OAM	Low	Low	N/A
Premium	VoIP, T.38 Fax over IP, Lawful Intercept, CES Real-time VPN service (CIR > 0, EIR = 0)	Very Low	Very Low	Very Low
Platinum	Video Conferencing, Interactive Gaming Real-time VPN service (CIR > 0, EIR > 0)	Low	Low	Low
Gold	Streaming audio, video on demand Broadcast TV, video surveillance	Low-Med	Med-High	High
Silver	Credit card transactions, wire transfers Instant Messaging Low Loss/Delay Data VPN service (CIR > 0, EIR > 0)	Low	Low-Med	N/A
Bronze	E-mail Non-time-critical OAM&P	Low	Med-High	N/A
Standard	Best effort applications Best effort VPN (CIR >= 0, EIR > 0)	Med	High	N/A

Table 2.9: Quality of Service Matrix

The following highlights the IP header/DSCP and the DSCP/ToS/IP precedence mapping to the Avaya Service Classes. There are 64 possible different DSCP markings that can be utilized for QoS in the network, along with four different per hop behaviors:

➢ Expedited Forwarding – voice services

➢ Assured Forwarding – real-time and non-real-time applications

➢ Class Selector – used to support legacy routers

➢ Default Forwarding – best effort

IP Header

6 Bytes	6 Bytes	4 Bytes	2 Bytes	64-1500 Bytes
Dest MAC	Source MAC	802.1q Tag	Protocol Type	Data

Differentiated Services (DS) Field

Version 4 bits	Length 4 bits	TOS 8 bits	Total Length 16 bits	More IP Header

```
0   1   2   3   4   5   6   7
        DSCP            |  CU

1   0   1   1   1   0     CU
```

Code Point Space	USE
XXXXX0	Defined Code Points
XXXX11	Experimental or Local use
XXXX01	Future Defined Code Points
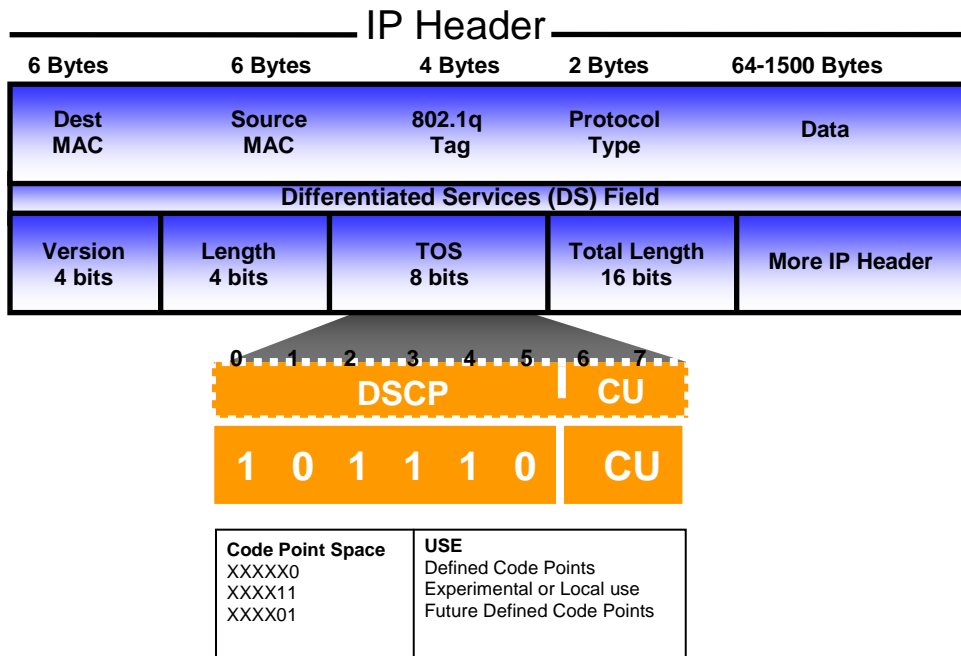
DSCP Marking

Differentiated Services Code Point – Six bits of the DS field are used to select the per hop behavior that packet experiences at each node. There are 64 possible code points.

Drop Precedence	Class 1	Class 2	Class 3	Class 4
Low	001010	010010	011010	100010
Medium	001100	010100	011100	100100
High	001110	010110	011110	100110

Figure 2.23: IP Header – DSCP Definition

Table 2.10 depicts the Avaya mapping of DSCP, Type of Service (ToS), and IP precedence to the Avaya Service Classes, along with their mapping into the DSCP per hop behavior.

DSCP	TOS	IP Precedence	Binary	NNSC	PHB
0x0	0x0	0	000000 **00**	Standard	CS0
0x0	0x0	-	000000 **00**		DE
0x8	0x20	1	001000 **00**	Bronze	CS1
0xA	0x28	-	001010 **00**		AF11
0x10	0x40	2	010000 **00**	Silver	CS2
0x12	0x48	-	010010 **00**		AF21
0x18	0x60	3	011000 **00**	Gold	CS3
0x1A	0x68	-	011010 **00**		AF31
0x20	0x80	4	100000 **00**	Platinum	CS4
0x22	0x88	-	100010 **00**		AF41
0x28	0xA0	5	101000 **00**	Premium	CS5
0x2E	0xB8	-	101110 **00**		EF
0x30	0xC0	6	110000 **00**	Network	CS6
0x38	0xE0	7	111000 **00**	Critical	CS7
DSCP and TOS are in HEX IP Precedence in decimal NNSC: Avaya Service Class				PHB: Per Hop Behavior	

Table 2.10: Default Avaya DSCP / ToS / IP Mapping

DSCP markings happen at the edge of the network, either by the end device themselves or by the edge switch. At the core of the network, the ERS 8800's must honor those DSCP markings. All IST/SLT/SMLT ports should be configured as DiffServ Trusted ports, which honor the DSCP marking in the packets.

Separate QoS filters can be created for any locally attached devices that require special treatment on the network. As shown in the figure below, the ports used to connect to the Server would be configured with QoS filters to mark specific DSCP values based on the application and/or traffic type.
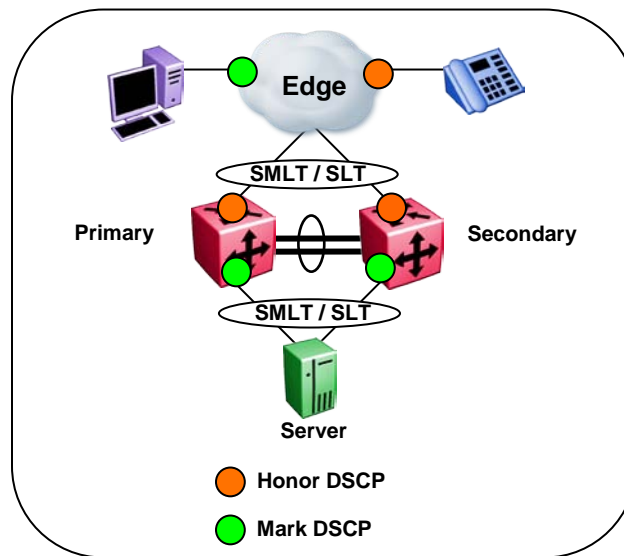
Figure 2.24: ERS 8000 Core QoS

2.1.21 DHCP Relay

DHCP is used to provide IP addresses to end stations in the network. In the design scenario where multiple VLANs are used, DHCP relay will be required to get the DHCP information from the end user VLAN to the DHCP server (which would normally sit on a core server VLAN).

> ➢ Enable per VLAN with the physical VLAN IP address as the DHCP relay agent

> ➢ Multiple DHCP server addresses are supported (up to 10) per relay agent

> ➢ Do not use the VRRP virtual IP address as the DHCP Relay agent, always use the physical IP address of the VLAN

2.1.22 VRRP with Backup Master

VRRP provides redundancy for end user's default gateway and should be utilized for each VLAN configured that host end stations. Along with VRRP, Backup Master should be enabled on the Switch Cluster to provide active-active routing and forwarding of traffic.
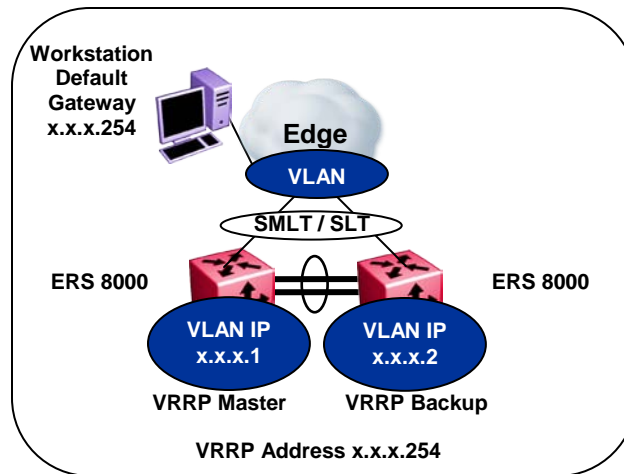
Figure 2.25: ERS 8000 VRRP

- ➤ Enable VRRP and Backup Master on each VLAN
- ➤ Configure VRRP priority higher than 100 (i.e. 200) to set VRRP Master
- ➤ Stagger VRRP Masters between ERS 8000's in the core
- ➤ Leave VRRP priority at default (100) for VRRP Backup
- ➤ Do not configure the virtual address as a physical interface that is used on any of the routing switches – use a third address, for example:
    - ▪ Physical IP address of VLAN on Switch 1 = x.x.x.1
    - ▪ Physical IP address of VLAN on Switch 2 = x.x.x.2
    - ▪ Virtual IP address of VLAN a = x.x.x.254

2.1.23 RSMLT Layer 2 Edge

RSMLT L2 Edge offers an alternative to VRRP for end user default gateway redundancy. VRRP and RSMLT L2 Edge can be used on the same Switch Cluster on different VLANs, but do not use both VRRP and RSMLT L2 Edge on the same VLAN simultaneously.

The RSMLT implementation does not use a Virtual IP address but instead uses physical IP addresses on each ERS 8800/8600 for redundancy. RSMLT L2 Edge stores the RSMLT peer MAC/IP address-pair in its local configuration file and restores the configuration if the peer does not restore after a simultaneous reboot of both RSMLT peer switches. It is imperative to save the configuration file on each ERS 8800/8600 when RSMLT L2 Edge if first implemented to ensure that the peer MAC/IP address-pair is saved in the configuration file.

Each ERS 8800/8600 is able to forward on behalf of itself as well as its peer in the Switch Cluster. This makes very efficient use of bandwidth and resources and also ensures seamless failover and recovery in the event of a failure.

Avaya recommends using RSMLT L2 Edge in place of VRRP as it provides several advantages, including:

> RSMLT is only limited by the number of IP interfaces on the ERS 8800/8600

  ▪ VRRP is limited to 250 instances

> RSMLT requires significantly less control traffic

> RSMLT is much less intensive on CPU resources
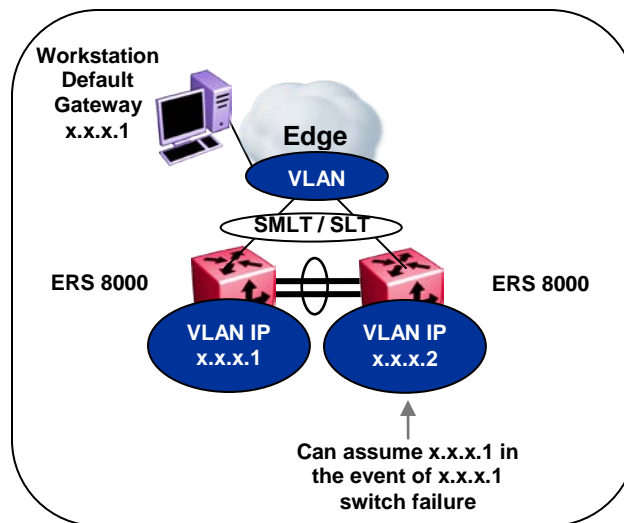
Figure 2.26: ERS 8000 RSMLT L2 Edge

When implementing RSMLT L2 Edge, make sure to:

> Enable RSMLT on each VLAN

> Enable RSMLT-Edge-Support on each VLAN

> Configure the Hold-up timer to 9999 (infinity) – this timer defines how long the RSMLT switch maintains forwarding for its peer

2.1.24 Layer 3 Routing

The Large Campus Design will most likely require a unicast routing protocol, the ERS 8800/8600 presently supports OSPF, RIP, and BGP. OSPF is the Avaya recommended unicast routing protocol for the Large Campus design. The efficiencies of OSPF along with its scalability and overall market adoption make it the best alternative for routing.

Please consult the Release Notes for the ERS 8800/8600 for route scalability, number of IP interfaces, number of OSPF areas, etc. supported per switch/stack.

When designing an OSPF network, review the following criteria:

> Timers must be consistent across the entire network – **use default values**

> Configure core switches as designated routers

  ▪ Higher rtrpriority is designated router

  ▪ Configure rtrpriority in increments of 10 for future flexibility

> Configure **OSPF router ID** the same as the management address for operational simplicity

> Must enable **ASBR** to utilize route redistribution

> Use **MD5 authentication** on any untrusted OSPF links

> Use **OSPF area summarization** to reduce routing table sizes

> Use **Stub or NSSA** areas as much as possible to reduce CPU overhead

> Use **OSPF passive interfaces** to ensure routing updates are not sent to the edge switches (assuming no routers exist there)

> Use **OSPF active interfaces** only on intended route paths. Typically, you should configure wiring closet subnets as OSPF passive interfaces unless they form a legitimate routing path for other routes.

> Limit the number of **OSPF areas per switch to as few as possible** to avoid excessive shortest path calculations. Be aware that the switch has to execute the Djikstra algorithm for each area separately.

> Ensure that the **OSPF dead interval** is at least four times the OSPF hello interval

When designing a RIP network, review the following criteria:

> For user VLANs where RIP is enabled, disable default supply and listen to reduce the amount of broadcast traffic on those VLANs

> Disable RIP supply and learn on all access ports

> Use triggered updates to help reduce routing convergence times

2.1.24.1 Equal Cost Multipath (ECMP) for Layer 3 link load balancing

ECMP enables the ability to load balance Layer 3 links and provide redundancy for routing in situations where there are at least two equal paths from the source to the destination network. ECMP uses a very similar algorithm as Multilink Trunking to distribute traffic among the links.

The number of paths is configurable. The ERS 8800/8600 supports up to eight paths in R mode (using R or RS modules) and up to four paths with all other legacy modules. The ERS 8300 and ERS 5000 support up to four equal cost paths. Note that ECMP is not supported on the ERS 5510 switches. ECMP supports and complements OSFP, RIP, and Static routes.

2.1.24.2 Routed Split Multilink Trunking (RSMLT)

Avaya's RSMLT permits rapid failover for core topologies by providing an active-active router concept to core Split MultiLink Trunking (SMLT) networks. RSMLT scenarios include SMLT triangles, squares, and SMLT full-mesh topologies, with routing enabled on the core VLANs. In the event of core router failures, RSMLT manages packet forwarding, thus eliminating dropped packets during the routing protocol convergence. RSMLT supports the following:

- IP Unicast Static Routes
- OSPF
- RIP v1/v2
- BGP

When RSMLT is enabled on a VLAN (on both aggregation devices) the cluster switches simply inform each other (over IST messaging) of their physical IP/MAC on that VLAN; thereafter the two Cluster switches take mutual ownership of each other's IP address on that VLAN; which means each cluster switch will:

- Reply to ARP request for both it's IP and it's Peer's IP on that VLAN

- Reply to pings to it's IP and it's Peer's IP on that VLAN

- Route IP traffic which is being directed to the physical MAC of it's IP or the physical MAC of it's peer's IP on that VLAN
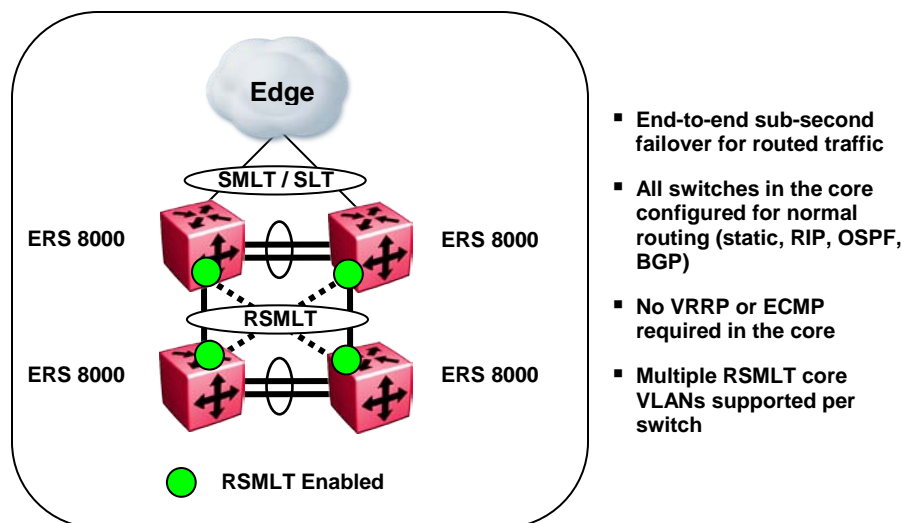
Figure 2.27: ERS 8000 Routed Split Multilink Trunking (RSMLT)

RSMLT runs under the existing routing protocols, so no tuning of the IGP is necessary and there is no direct interaction between RSMLT and the IGP.

- ➢ Based on SMLT, so all SMLT rules apply
- ➢ Configured on a per VLAN basis
- ➢ VLAN must be routable and part of the SMLT links and IST link

> ➢ If possible, ensure that destination networks are directly accessible from each of the switches participating in RSMLT. This guarantees sub-second failover

> ➢ The hold-up timer defines how long the RSMLT-peer switch keeps routing for its peer after a peer switch failure. Leave the hold-up timer at default of 180 seconds.

> ➢ The hold-down timer defines how long the switch waits before activating Layer 3 forwarding for it's peers' MAC address. Leave the hold-up timer at default of 60 seconds.

> ➢ In an RSMLT-based VLAN, the maximum number of supported SMLT/SLT's is 32. The RSMLT-based VLAN can have as many non-SMLT/SLT ports as needed.

2.1.24.3 RSMLT Dual Core VLANs

By using Avaya technologies in new innovative ways, the availability and resilency of the network can be increased substanstially. The following uses RSMLT in a configuration with dual core VLANs to minimize traffic interruption in the event of losing the OSPF designated router (DR).

When the DR fails, the backup designated router (BDR) is immediately elected as the new DR on the broadcast segment. As soon as the new DR is elected (via the Hello protocol) all routers must immediately perform a shortest path first (SPF) run and issue new link state advertisements (LSAs) for the segment. Since the DR is the only router generating the Network LSA for the segment, a new Network LSA needs to be generated by the new DR (the old one no longer being valid) and every router on that segment also needs to refresh their own router LSA.

The SPF run is done as soon as the router has detected a new DR. If the router is quick, it will perform the SPF run before receiving the new LSAs for the segment; as such all OSPF routes across this particular segment are not possible after the first SPF run. Whereas if the router is slow, by the time it detects the new DR and does it's SPF run, it already has received most (if not all) of the new LSAs from its neighbors and it's routing table will be still be able to route most (if not all) routes across the segment.

In either case, if the DR fails, two SPF runs are necessary to restore routes across the segment. But the OSPF holddown timer will not allow two consecutive SPF runs to occur within the value of the timer. That timer defaults to 10secs on the ERS 8800 and can be reduced to a minimum of 3 seconds, however reducing this timer is not recommended. Every time the DR is lost, traffic interruption of up to 10 seconds can occur.

The solution for this scenario is to

> ➢ Create a second OSPF Core VLAN, forcing different nodes to become 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 – the two VLANs back each other up from a routing perspective
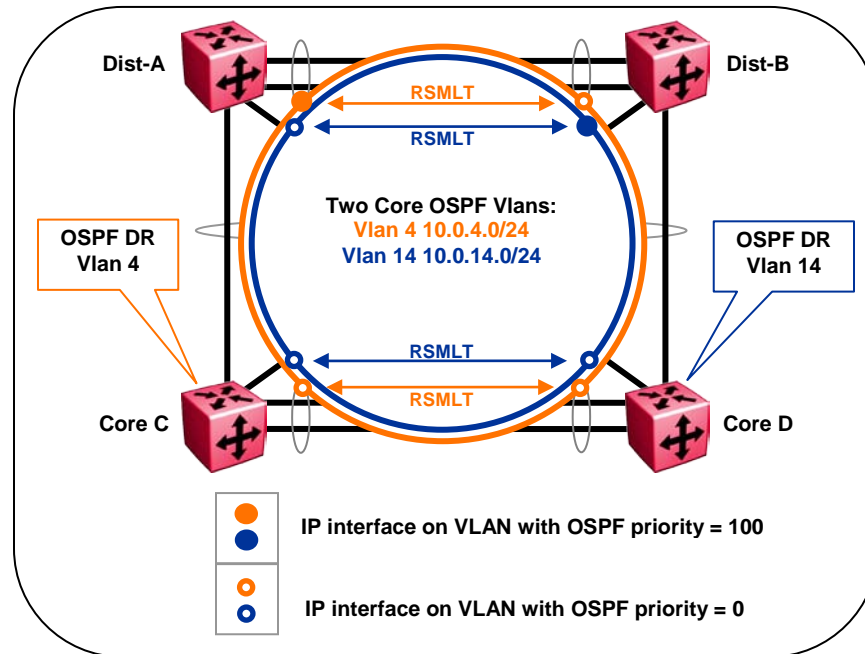
Figure 2.28: RSMLT with Dual Core VLANs

2.1.25 Multicast

Multicast routing is used to distribute multicast traffic within the network. Sources of multicast traffic must be routed on different IP subnets to reach the multicast subscribers. The Avaya Ethernet Routing Switches are uniquely designed to support multicast traffic flows in a very efficient manner. Ingressing packets are classified in hardware and only replicated where needed.

Avaya recommends using PIM as the multicast routing protocol of choice. PIM only sends multicast to areas of the network that have specifically requested the multicast stream. This is a much more efficient use of the available bandwidth in the network. PIM is usually the protocol of choice when there are a sparse number of users requesting the multicast stream. PIM-SSM is usually the protocol of choice for applications such as TV distribution or applications that require transmission acknowledgement. PIM uses the underlying routing table of the unicast routing protocol for its route table – for large scale networks, it is best to use OSPF.

DVMRP is a flood-and-prune technology in which the multicast streams are sent throughout the network and then pruned back to only those areas needing the multicast. This technology works very well where there are large groups of users requesting the same multicast stream. DVMRP is presently only supported in the ERS 8800 and there are no plans to add this functionality to any other ERS products. DVMRP is supported in a Switch Cluster triangle topology (ERS 8800/8600 Switch Cluster core with edge switches directly connected). There is no immediate or future support for DVMRP over SMLT in a square or full mesh topology.

For details on Multicast Routing, please refer to the *Resilient Multicast Routing Using Split Multilink Trunking for the ERS 8600 Technical Configuration Guide* (NN48500-544).

The Large Campus Solution will likely have multicast application requirements. The ERS 8800/8600 supports PIM-SM with Switch Clustering and IGMP over SMLT/SLT. The network design accommodates the multicast needs by using PIM-SM for multicast routing in the core and IGMP Snooping and Proxy at the edge. When implementing Multicast routing in the Large Campus, follow the recommendations outlined below:

> Use PIM-SM as the Layer 3 Multicast routing protocol

> Enable IGMP Snooping and Proxy on the edge switches
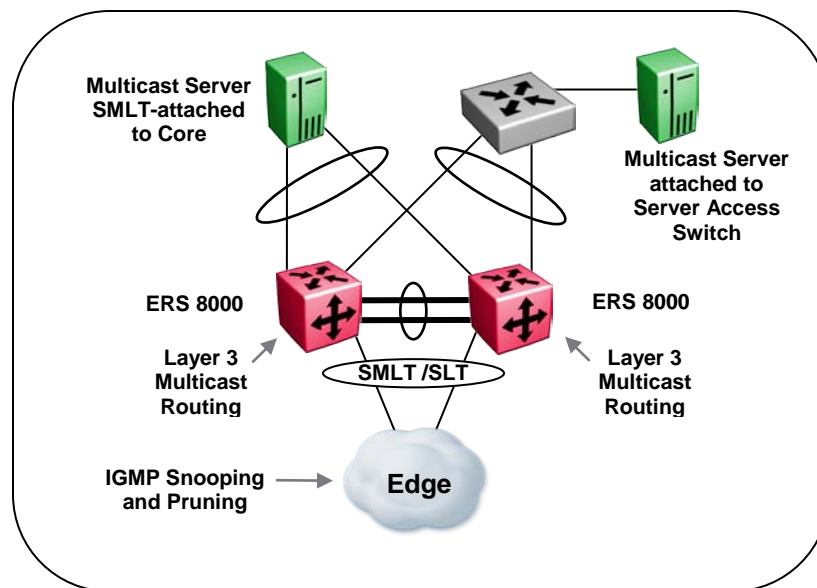
Figure 2.29: ERS 8000 Multicast Routing

2.1.26 Server Connectivity

The preferred method for connecting servers in the Large Campus Solution is to use a subtended Server Access Switch/Stack (ERS 2500/4500/5000). For configuration details, see the Data Center Server Access Solution Guide (NN48500-577). Servers can also be directly attached to the core if required. However, keeping the core functionality separated from server connectivity is preferred when possible to aide in fault isolation, maintenance, and troubleshooting.

Server virtualization is also fully supported with the Large Campus solution. Avaya and VMware have fully certified a solution as described in the Resilient Data Center Server Edge Solutions for VMware ESX Server Technical Configuration Guide (NN48500-542).

General design considerations:

> Disable Spanning Tree on the ports connected to the servers (MLT/SMLT/SLT)

> Server NIC Teaming configurations supported include:

   - Broadcom NICs using FEC/GEC Generic Trunking

   - Intel NICs using Static Link Aggregation

   - HP NICs using SLB with automatic Transmit Load Balancing

Quantity of Servers supported when connecting to a subtended Server Access Switch/Stack:

> ➢ MLT-attached servers will be number of MLT groups supported on the Server Access Switch/Stack less one needed for the SMLT/SLT to the Switch Cluster Core.
>   - ▪ ERS 5000 supports 32 MLT groups with a maximum of 8 ports per group
>   - ▪ ERS 4500 supports 32 MLT groups with a maximum of 8 ports per group
>   - ▪ ERS 2500 supports 6 MLT groups with a maximum of 4 ports per group

Quantity of Servers supported when connecting directly to Switch Cluster Core:

> ➢ SMLT-attached servers equals the number of MLT groups supported on the core less the number of MLT groups already in use by Edge switch connections via SMLT or MLT.
>
> ➢ SLT-attached servers will be number of ports available on the Switch Cluster Core
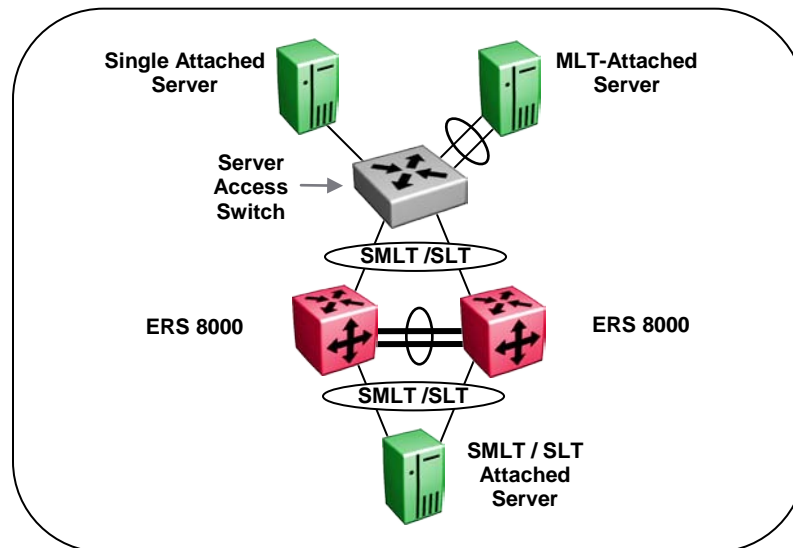
Figure 2.30: ERS 8000 Server Connectivity

2.2 Edge Switching

The Edge Switching provides the end user connectivity for the Large Campus Solution. Avaya provides flexibility in which product(s) can be used at the edge; ERS 2500, ERS 4500, ERS 5000.

The Edge switching portfolios support a multitude of features, all of which are not presented in this solution. The features and functionality highlighted here represent the basic requirements for the Large Campus Solution. Please refer to the product documentation for a detailed explanation of these and all the other features of the above mentioned platforms.

This section will detail the switching options for the edge and also provide the best practice recommendations. Platform differentiations and feature differences will be highlighted as applicable in each section for Edge Switching.

- ➢ Edge Switching Products
- ➢ Stacking of Edge Switches
- ➢ Virtual LANs (VLANs)
- ➢ Link Aggregation
- ➢ Spanning Tree
- ➢ BPDU Filtering

- ➢ VLANs
- ➢ DHCP Relay
- ➢ Quality of Service
- ➢ Layer 3
- ➢ VRRP with Backup Master

2.2.1 Edge Switching Products

The ERS 2500 product portfolio offers the following four different switch models.

ERS 2526T

- ➢ 24 ports 10/100
- ➢ 2 combo ports (SFP or 10/100/1000)

ERS 2526T-PWR

- ➢ 24 ports 10/100 (12 with PoE)
- ➢ 2 combo ports (SFP or 10/100/1000)

ERS 2550T

- ➢ 48 ports 10/100
- ➢ 2 combo ports (SFP or 10/100/1000)

ERS 2550T-PWR

- ➢ 48 ports 10/100 (24 with PoE)
- ➢ 2 combo ports (SFP or 10/100/1000)

The ERS 4500 product portfolio offers the following eleven different switch models.

ERS 4526T

> 24 ports 10/100

> 2 combo ports (SFP or 10/100/1000)

ERS 4526T-PWR

> 24 ports 10/100 with PoE

> 2 combo ports (SFP or 10/100/1000)

ERS 4550T

> 48 ports 10/100

> 2 combo ports (SFP or 10/100/1000)

ERS 4550T-PWR

> 48 ports 10/100 with PoE

> 2 combo ports (SFP or 10/100/1000)

ERS 4524GT

> 24 ports 10/100/1000

> 4 shared SFP ports

ERS 4524GT-PWR

> 24 ports 10/100/1000 with PoE

> 4 shared SFP ports

ERS 4548GT

> 48 ports 10/100/1000

> 4 shared SFP ports

ERS 4548GT-PWR

> 48 ports 10/100/1000 with PoE

> 4 shared SFP ports

ERS 4526GTX

- ➤ 24 ports 10/100/1000
- ➤ 2 XFP ports (10 Gig LAN Phy)

ERS 4526GTX-PWR

- ➤ 24 ports 10/100/1000 with PoE
- ➤ 2 XFP ports (10 Gig LAN Phy)

ERS 4526FX

- ➤ 24 ports 100FX
- ➤ 2 combo ports (SFP or 10/100/1000)

The ERS 5000 product portfolio offers the following ten different switch models.

ERS 5650TD

- ➤ 48 ports 10/100/1000
- ➤ 2 XFP ports (10Gig)

ERS 5650TD-PWR

- ➤ 48 ports 10/100/1000 with PoE
- ➤ 2 XFP ports (10Gig)

ERS 5698TFD

- ➤ 96 ports 10/100/1000
- ➤ 6 Shared SFP ports
- ➤ 2 XFP ports (10Gig)

ERS 5698TFD-PWR

- ➤ 96 ports 10/100/1000 with PoE
- ➤ 6 Shared SFP ports
- ➤ 2 XFP ports (10Gig)

ERS 5632FD

- ➤ 24 SFP ports
- ➤ 8 XFP ports (10Gig)

ERS 5530-24TFD

- ➤ 12 10/100/1000 ports
- ➤ 12 Shared (SFP or 10/100/1000)
- ➤ 2 XFP ports (10Gig)

ERS 5520-48T-PWR

- ➤ 48 10/100/1000 ports with PoE
- ➤ 4 Shared SFP ports

ERS 5520-24T-PWR

- ➤ 24 10/100/1000 ports with PoE
- ➤ 4 Shared SFP ports

ERS 5510-48T

- ➤ 48 10/100/1000 ports
- ➤ 2 Shared SFP ports

ERS 5510-24T

- ➤ 24 10/100/1000 ports
- ➤ 2 Share SPF ports

ERS 8300

➢ 6 slot or 10 slot chassis

➢ Dual switching fabrics with integrated uplink ports for maximum density

➢ N+1 power supply redundancy (AC and DC options available)

➢ All components fully hot swappable

10 Slot Chassis

6 Slot Chassis

Module	Ports	Type
8393SF	8	**288Gbps Switch Fabric with 8 1GbE SFP ports**
8394SF	2	**288Gbps Switch Fabric with 2 10GbE XFP ports**
8348TX	48	**48 port 10/100BaseT**
8348TX-PWR	48	**48 port 10/100BaseT with 802.3af PoE**
8324GTX	24	**24 port 10/100/1000BaseT**
8348GTX	48	**48 port 10/100/1000BaseT**
8348GTX-PWR	48	**48 port 10/100/1000BaseT with 802.3af PoE**
8348GB	48	**48 port 1GbE SFP**
8308XL	8	**8 port 10GbE XFP**

8308XL   **8394SF**   **8393SF**   **8348GTX-PWR**

2.2.2 Stacking of Edge Switches

The ERS 2500, ERS 4500, and ERS 5000 switches support a common resilient stacking architecture. The stack is created by using the stacking cables and stacking ports on the ERS switches. Switches are cabled together in the manner shown below so that every switch has two stacking connections for utmost resiliency. The shortest path algorithm used for stacking allows for the most efficient use of bandwidth across the stack. The maximum number of switch can be allowed to stack is 8 and can be of any model (mix and match) within the same product family.

A failure in any unit of the stack will not adversely affect the operation of the remaining units in the stack. Replacement of the failed switch is easy with the Auto-Unit Replacement feature. This allows for a new switch to be put into the stack and will automatically get the right software image and configuration without user intervention – the replacement switch must be the exact model of the failed switch. In addition, the SW image has to be right for the AUR to work properly. Please refer to Product documentation for more info on AUR.

- ➢ Enable Stack Monitor on all Edge stacks which will alert on any stack size changes
- ➢ ERS 2500
  - ▪ 4Gbps stacking bandwidth per switch, 32Gbps maximum per stack
  - ▪ Requires a stack license per unit
- ➢ ERS 4500
  - ▪ 40Gbps stacking bandwidth per switch, 320Gbps maximum per stack
  - ▪ Enable Forced Stack Mode when using stacks of two switches
- ➢ ERS 5000
  - ▪ ERS 5500 – 80Gbps stacking bandwidth per switch, 640Gbps maximum per stack
  - ▪ ERS 5600 – 144Gbps stacking bandwidth per switch, 1.1Tbps maximum per stack
  - ▪ Hybrid stacks of 5500 and 5600 switches are supported
  - ▪ Enable Forced Stack Mode when using stacks of two switches
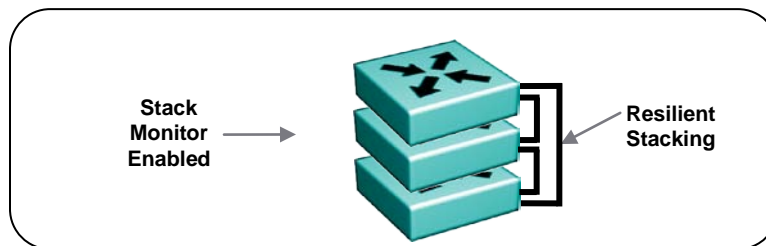
Figure 2.31: Edge Stacking

2.2.3 Power over Ethernet

The use of Power over Ethernet (PoE) has become increasingly popular over the past few years and is standardized by the IEEE as 802.3af. Many end devices, such as wireless access points, security cameras, and security card readers now support power over Ethernet; however, the largest market driver for PoE today is IP Telephony. The ability to centrally power IP phones from the Ethernet switch has made PoE ubiquitous in IP Telephony environments.

There are two components to PoE:

➢ Power Sourcing Equipment (PSE) which is normally the Ethernet switch providing the PoE

➢ Powered Device (PD) which is the end device using the power
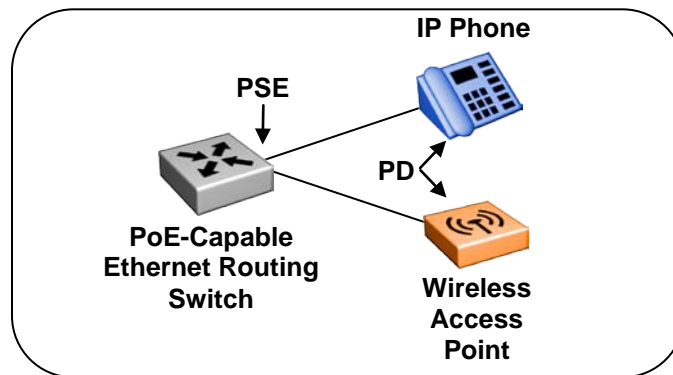
Figure 2.32: Power over Ethernet

The PSE will not send power out of the Ethernet switch port until it is able to verify the end station is PoE capable. The 802.3af standard specifies a resistive discovery mechanism for the PSE to perform this function. This resistive method sends out two low voltage electrical pulses and once the PSE has successfully discovered the PD, it will begin providing power. Legacy (pre-standard) PoE devices use a capacitive detection mechanism to detect the PD. All Avaya PoE switches support both resistive and capacitive detection mechanisms as many pre-standard PD's supported only capacitive discovery.

Power can be provided over the used data pairs (1, 2, 3, and 6) or the unused pairs (4, 5, 7, and 8) in a UTP copper cable. The 802.3af standard mandates that PD's must be able to accept power from either option. The Avaya Ethernet switches (2500, 4500, 5000, and 8300) provide power over the used pairs.

When designing a Converged Campus network, it is imperative to understand the PoE requirements in order to provide adequate power to the end devices. The 802.3af standard specifies a maximum of 15.4 watts per device depending on power classification, with many devices using only 4 to 8 watts.

The 802.3af standard defines the different classes of detection, as defined in Table 2.11. The Avaya Ethernet switches do not use power classification to provide PoE to the PD's. The Ethernet switches use a pool of power per switch or per module. As devices come online and begin to use power, the overall pool of power is decremented. Power management and power priority are used to control the amount of power and what ports have priority to that power.

Class	Usage	Maximum power level with 100 m cable of Cat 5 at	
		Output of PSE	Input of PD
0	Default	15.4 Watts	0.44–12.95 Watts
1	Optional	4.0 Watts	0.44–3.84 Watts
2	Optional	7.0 Watts	3.84–6.49 Watts
3	Optional	15.4 Watts	6.49–12.95 Watts

Table 2.11: PoE Classes of Power Input/output

➢ Four classes of DC detection (Classes 0, 1, 2 and 3) per IEEE 802.3af

➢ Four classes of AC detection (Classes 0, 1, 2 and 3) per IEEE 802.3af

➢ One class of Capacitive detection (Class 2) for pre-IEEE Avaya IP phones

The difference in output power of the PSE and input power of the PD is to account for loss in the Ethernet cable.

When designing the network, take into consideration the potential number of end devices requiring PoE and the amount of power each of these devices requires. This will help decide the number of PoE-capable ports required in the wiring closet.

The network administrator can also use features within the Ethernet switches to control the PoE across the switch. Power Management will enable/disable or limit the PoE on a per port/module basis. Power Priority allows certain ports to have priority over other ports when requiring PoE. In certain configurations, there may not be enough PoE available. The power priority feature will allow designated ports to get power over lower priority ports. This will ensure the critical PoE devices will always have power.

Another critical aspect of PoE is redundant power in the wiring closet. When relying on the Ethernet switch to provide power to the end devices, it becomes even more important to have redundant power available in the closet in case of an electrical circuit failure. The addition of redundant power can also be used by the Ethernet switches to add overall power capability and increase the total amount of power available for end devices using PoE.

In all cases, Avaya strongly recommends having separate electrical circuits available in the closet for the various AC feeds into the Ethernet switching equipment as well as the Redundant Power Supply Units (if applicable).

Avaya created a PoE Calculator tool which assists during the design phase. By simply selecting the number of Avaya IP phones and wireless access points, or by entering in information for third party PoE devices, the calculator will provide information regarding the power required, number of switches, and if necessary, the amount of redundant power required. This tool eliminates the guess work around the design of the PoE infrastructure. The PoE Calculator can be found on the support site of Avaya.com or by searching for document number NN48500-520.

A new PoE standard is now emerging. 802.3at, or also commonly referred to as PoE+ will increase the amount of power supplied by the PSE. Avaya will offer products supporting 802.3at in the near future. Please note that this will require new Ethernet switching hardware to support 802.3at and the increased power. More information on this emerging standard can be found at http://www.ieee802.org/3/at/

The design of a network capable of PoE will have several variations depending on the Ethernet switching equipment that is chosen. The following highlights the various options available with Avaya PoE Ethernet switches:

Ethernet Routing Switch 8300

This chassis system provides both 10/100 and 10/100/1000 48 port I/O modules capable of PoE. When utilizing PoE, make sure to engineer the power requirements of the chassis properly. The amount of PoE per module is configurable up to 800 watts per module, along with the ability to specify port priority for PoE. The total PoE power required will dictate the type of input power for the chassis. The ERS 8300 provides different power options as indicated in Table 2.12.

ERS 8300 Six Slot   **ERS 8300 Ten Slot**

Power Supply	Power Supply Rating	# of Power Supplies	Redundancy	PoE Available
8301AC	110-120 VAC 20 Amp 1140 watts	1	No	400 watts
		2	Yes 1+1	400 watts
		3	Yes 2+1	800 watts
	200-240 VAC 20 Amp 1770 watts	1	No	800 watts
		2	Yes 1+1	800 watts
		3	Yes 2+1	1600 watts
8302AC	100-120 VAC 15 Amp 850 watts	1	No	200 watts
		2	Yes 1+1	200 watts
		3	Yes 2+1	400 watts
	200-240 VAC 15 Amp 1400 watts	1	No	400 watts
		2	Yes 1+1	400 watts
		3	Yes 2+1	800 watts

Table 2.12: ERS 8300 Power over Ethernet Options

Ethernet Routing Switch 5600

The PoE capable ERS 5600 series stackable switches are available in a 48-port and a 96-port version. The ERS 5600 offers built-in, hot swappable redundant power supply options in both AC and DC varieties. It is also capable of providing full 15.4watts per port on every port in the switch along with full N+1 redundant power simultaneously. The available configurations for power options are specified in Table 2.13.

ERS 5650TD-PWR

ERS 5698TFD-PWR

Switch Model	PoE with one power supply	PoE with two power supplies	PoE with three power supplies
ERS 5650TD-PWR (600W)	370 watts total 7.7 watts/port	740 watts total 15.4 watts/port	N/A
ERS 5650TD-PWR (1000W)	740 watts total 15.4 watts/port	740 watts total * 15.4 watts/port	N/A
ERS 5698TFD-PWR (1000W)	740 watts total 7.7 watts/port	1480 watts total 15.4 watts/port	1480 watts total * 15.4 watts/port
* Full 15.4 watts on every port with N+1 power redundancy			

Table 2.13: ERS 5600 Power over Ethernet Options

Ethernet Routing Switch 5500

The PoE capable ERS 5520 stackable switch is available in both a 24-port and a 48-port version. The ERS 5520 provides up to 320 watts per switch on standard 110 VAC power. To provide more power and/or redundant power, use the ERS Redundant Power Supply 15 (RPS 15) to augment the ERS 5520. The RPS 15 can support up to three ERS 5520 switches. The available configurations for power options are specified in Table 2.14.

ERS 5520-48T-PWR

ERS 5520-24T-PWR

Switch Model	PoE on Standard AC	RPS 15 Power Sharing	RPS 15 RPSU
ERS 5520-24T-PWR	320 watts total 13.3 watts/port	740 watts total 15.4 watts/port	320 watts total 13.3 watts/port
ERS 5520-48T-PWR	320 watts total 6.7 watts/port	740 watts total 15.4 watts/port	320 watts total 6.7 watts/port

Table 2.14: ERS 5500 Power over Ethernet Options

Ethernet Routing Switch 4500

The PoE capable ERS 4500 stackable switches are available in 10/100 and 10/100/1000 48-port versions. The ERS 4500 provides up to 370 watts per switch on standard 110 VAC power. To provide more power and/or redundant power, use the ERS Redundant Power Supply 15 (RPS 15) to augment the ERS 4500. The RPS 15 can support up to three ERS 4500 switches. The available configurations for power options are specified in Table 2.15.

ERS 4526T-PWR   **ERS 4550T-PWR**

ERS 4524GT-PWR   **ERS 4548GT-PWR**

ERS 4526GTX-PWR

Switch Model	PoE on Standard AC	RPS 15 Power Sharing	RPS 15 RPSU
ERS 4526T-PWR	370 watts total 15.4 watts/port	740 watts total 15.4 watts/port	370 watts total 15.4 watts/port
ERS 4550T-PWR	370 watts total 7.7 watts/port	740 watts total 15.4 watts/port	370 watts total 7.7 watts/port
ERS 4524GT-PWR	360 watts total 15.0 watts/port	740 watts total 15.4 watts/port	360 watts total 15.0 watts/port
ERS 4548GT-PWR	320 watts total 6.7 watts/port	740 watts total 15.4 watts/port	320 watts total 6.7 watts/port
ERS 4526GTX-PWR	360 watts total 15.0 watts/port	740 watts total 15.4 watts/port	360 watts total 15.0 watts/port

Table 2.15: ERS 4500 Power over Ethernet Options

Ethernet Routing Switch 2500

The PoE capable ERS 2500 switches are available in both a 24-port and a 48-port version. With both of these ERS 2500 switches, PoE is provided on half the ports (ports 1-12 of the 24 port switch and ports 1-24 on the 48 port switch). The ERS 2500 provides up to 165 watts per switch on standard 110 VAC power. The ERS 2500 does not support a redundant power option. The available configurations for power options are specified in Table 2.16.

ERS 2526T-PWR   **ERS 2550T-PWR**

Switch Model	PoE on Standard AC	RPS 15 Power Sharing	RPS 15 RPSU
ERS 2526T-PWR	165 watts	N/A	N/A
ERS 2550T-PWR	165 watts	N/A	N/A

Table 2.16: ERS 2500 Power over Ethernet Options

Redundant Power Supply 15 (RPS 15)

The RPS 15 provides redundant power to the Avaya stackable Ethernet switches (both PoE and non-Poe). The RPS 15 is comprised of the following components:

- ➢ RPS 15 Chassis (supports up to three 600 watt power supplies)
- ➢ 600 Watt Power Supply
- ➢ DC-DC Converter (only required for some switches – see table below)
- ➢ DC cable to connect power supply to Ethernet switch

The RPS 15 supports two different DC cable types. The first (AA0005018) is used with all Ethernet switches that have a built-in DC-DC converter and can provide a single power connection to one Ethernet switch. The second type of cable, which comes in two models (AA0005020 – 25' and AA0005021 – 10') is used with all Ethernet switches that require the addition of the DC-DC converter module. This second cable type can provide a single power connection for up to four Ethernet switches.

The RPS 15 can be added to an Ethernet switch or stack of Ethernet switches while the switches are powered up and running. There is no need to power off the switch to connect the RPS 15 cable.
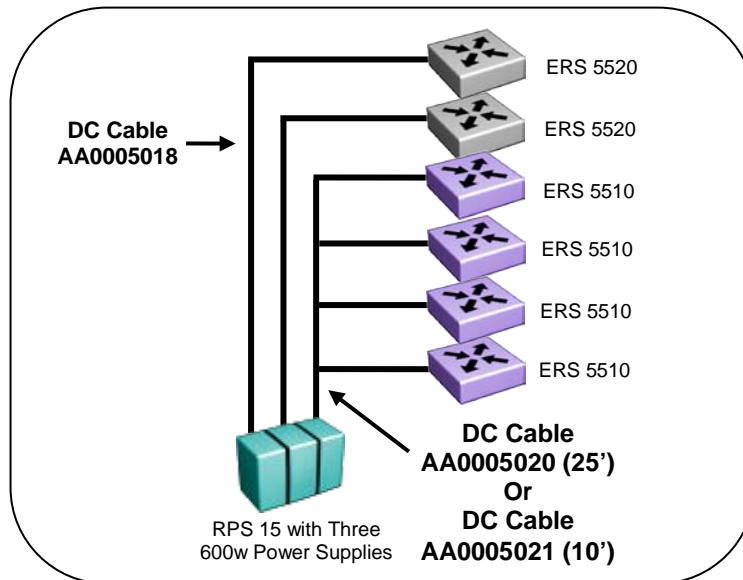
Figure 2.33: Redundant Power Supply 15 (RPS15)

Table 2.17 provides information on the required components when using the RPS 15 with the various Ethernet switching options.

Switch Model	PoE Capable Switch	RPS 15 Chassis	RPS 15 600w Power Supply	DC-DC Converter	DC Cable for Built-In Converter	10' or 25' DC Cable
ERS 5510	No	1	1 per 4 switches	Required	N/A	Required
ERS 5520	Yes	1	1	Built-In	Required	N/A
ERS 5530	No	1	1	Built-In	Required	N/A
ERS 4526FX	No	1	1 per 4 switches	Required	N/A	Required
ERS 4526T	No	1	1 per 4 switches	Required	N/A	Required
ERS 4526T-PWR	Yes	1	1	Built-In	Required	N/A
ERS 4550T	No	1	1 per 4 switches	Required	N/A	Required
ERS 4550T-PWR	Yes	1	1	Built-In	Required	N/A
ERS 4524GT	No	1	1 per 4 switches	Required	N/A	Required
ERS 4524GT-PWR	Yes	1	1	Built-In	Required	N/A
ERS 4548GT	No	1	1 per 4 switches	Required	N/A	Required
ERS 4548GT-PWR	Yes	1	1	Built-In	Required	N/A
ERS 4526GTX	No	1	1 per 4 switches	Required	N/A	Required
ERS 4526GTX-PWR	Yes	1	1	Built-In	Required	N/A
ES 470-24T-PWR	Yes	1	1	Built-In	Required	N/A
ES 470-48T-PWR	Yes	1	1	Built-In	Required	N/A
ES 470-24T	No	1	1 per 4 switches	Required	N/A	Required
ES 470-48T	No	1	1 per 4 switches	Required	N/A	Required

Table 2.17: RPS 15 Configuration Options

The following table highlights the PoE requirements of several Avaya end devices.

PoE Device	Average PSE
1100 Series Phones	
IP Phone 1110 PEC NTYS02xxE6	2.80 watts
IP Phone 1120E PEC NTYS03xA -- NTYS03xCE6	7.00 watts
IP Phone 1120E PEC NTYS03xDE6 (see Note 7)	4.60 watts
IP Phone 1120E PEC NTYS03xEE6 (see Note 8)	4.20 watts
IP Phone 1140E PEC NTYS05xA -- NTYS05xCE6	7.30 watts
IP Phone 1140E PEC NTYS05xCE6 Rel 50 & higher (see Note 7)	4.80 watts
IP Phone 1140E PEC NTYS05xEE6 (see Note 8)	4.30 watts
IP Phone 1150E PEC NTYS06xxE6	7.00 watts
IP Phone 1165E PEC NTYS07xxE6 (see Note 9)	3.80 watts
1200 Series Phones	
1210 phone w/ integrated 3 port 10/100 switch	3.40 watts
1220 phone w/ integrated 3 port 10/100 switch	3.40 watts
1230 phone w/ integrated 3 port 10/100 switch	3.40 watts
1600 Series Phones	
Avaya IP Phone 1603 Class 2 w/ Adapter	4.32 watts
Avaya IP Phone 1608 Class 2	4.66 watts
Avaya IP Phone 1616 Class 3 (see Note 1)	6.22 watts
Avaya IP Phone 1616 w/ BM32 Class 3	6.82 watts
Avaya IP Phone 1616 Class 2 (see Note 1)	3.17 watts
Avaya IP Phone 1616 w/ BM32 Class 2	4.37 watts
Avaya IP Phone 1692 Class 3	5.20 watts
Gig Adapter for 1600 Series Phones	3.16 watts
4600/5600 Series Phones	
Avaya IP Phone 4601, 4602, 5601, 5602 Class 2	3.50 watts
Avaya IP Phone 4602SW, 5602SW Class 2	4.10 watts
Avaya IP Phone 4606, 4612, 4624 Class 0	5.00 watts
Avaya IP Phone 4610SW, 5610SW Class 2	4.00 watts
Avaya IP Phone 4620 Class 3	7.70 watts
Avaya IP Phone 4620SW Class 3 (see Note 2)	5.90 watts
Avaya IP Phone 4620SW, 5620SW Class 2	4.60 watts
Avaya IP Phone 4621SW, 4622SW Class 2	4.90 watts
Avaya IP Phone 4625SW Class 3	7.80 watts
Avaya IP Phone 4630SW Class 3	11.80 watts
Gig Adapter for 1600 Series Phones	3.16 watts

PoE Device	Average PSE
9600 Series Phones	
Avaya IP Phone 9610 Class 2	3.72 watts
Avaya IP Phone 9620 Class 2	4.90 watts
Avaya IP Phone 9620L Class 1	2.20 watts
Avaya IP Phone 9620C Class 2	4.60 watts
Avaya IP Phone 9630 Class 2	5.20 watts
Avaya IP Phone 9630G Class 2	4.83 watts
Avaya IP Phone 9640 Class 2	4.52 watts
Avaya IP Phone 9640G Class 2	4.83 watts
Avaya IP Phone 9650 Class 2	4.60 watts
Avaya IP Phone 9650C Class 2	4.52 watts
Avaya IP Phone 9670G Class 2	6.24 watts
Phase 0 Phones	
IP Phone 2004	3.30 watts
IP Phone 2004 w/ external 3 port 10/100 switch	11.00 watts
Phase I Phones	
IP Phone 2004 w/ integrated 3 port 10/100 switch	3.30 watts
IP Phone 2002 w/ integrated 3 port 10/100 switch	3.30 watts
Phase II Phones	
IP Phone 2001 w/ integrated 3 port 10/100 switch	3.20 watts
IP Phone 2002 w/ integrated 3 port 10/100 switch	3.20 watts
IP Phone 2004 w/ integrated 3 port 10/100 switch	3.40 watts
IP Phone 2007 w/ integrated 3 port 10/100 switch	7.00 watts
2000 Series Audio Conference Phone IP Audio Conference Phone 2033	11.6 watts
LG-Avaya Phones	
LG-Avaya IP Phone 6804	2.30 watts
LG-Avaya IP Phone 6812	2.30 watts
LG-Avaya IP Phone 6830	2.30 watts
LG-Avaya IP Phone 8540	4.50 watts
Wireless LAN	
WLAN 8120 AP	9.30 watts
WLAN 2220 AP	8.50 watts
WLAN 2230 AP	10.00 watts
WLAN 2330 – 2330A - 2332 AP	8.00 watts

Table 2.18: PoE Consumption for Avaya IP Phones and Access Points

2.2.4 Physical Layer Considerations/Fiber Fault Detection

Avaya provides several options for uplink connectivity over fiber – this does not necessarily preclude the use of copper for uplink connections; however, due to the distance limitations of copper (100 m), fiber is normally the media of choice. Both ends of a link generally must use the same transceiver type (that is, SX to SX), but they do not necessarily have to be from the same manufacturer or vendor – Avaya GBICs, SFPs, XFPs interoperate with most all third-party vendors' products of the same transceiver type.

Avaya also supports the interoperability of CWDM with both XD and ZX GBICs. More specifically, one end of the link can use a ZX GBIC and the other end of the link a CWDM SFP, or one end can use an XD GBIC and the other end of the link a CWDM SFP. The following rules apply:

  ➢ XD GBIC with 40 Km CWDM SFP
  ➢ ZX GBIC with 70 Km CWDM SFP

Avaya offers SFPs and XFPs with a Digital Diagnostic Interface (DDI). These SFPs can provide significant diagnostic information when enabled by the Ethernet Switches via Digital Diagnostic Monitoring (DDM). The ERS 8800/8600 with Release 5.1 and higher supports DDM and allows the switch to monitor SFP and XFP laser operating characteristics. ERS 8600/8800 support for Digital Diagnostic Interfaces (DDI—an interface that supports DDM) involves data collection and alarm and warning monitoring. Static data collection includes SFP vendor information, DDI support information, and DDI alarm and warning threshold values. Dynamic data collection includes temperature, supply voltage, laser bias current, transmit power, and receive power. DDM works at any time during active laser operation without affecting data traffic. The switch only checks warning and alarm status bits during initialization and during requests for dynamic data. If an alarm or warning is asserted or cleared, the switch logs a message and generates a trap. The switch maps DDM warning and alarm messages into Warning and Fatal message categories for system logging purposes. If you activate the ddm-alarm-portdown option, DDI shuts down the corresponding port if a high or low alarm occurs on the port. This DDM functionality will be enabled on the other Ethernet Routing Switches in future software releases.

Review the fiber requirements of the network and select the appropriate GBIC/SFP/XFP based on those fiber specifications.

Fiber	62.5μ MMF		50μ MMF			9μ SMF	Wavelength
MHz*Km	160	200	400	500	2000	-	
10GBASE-SR	26m	33m	66m	82m	300m	-	850nm
10GBASE-LR/LW	-	-	-	-	-	10km	1310nm
10GBASE-ER/EW	-	-	-	-	-	40km	1550nm
10GBASE-ZR/ZW	-	-	-	-	-	80km	1550nm

Table 2.19: XFP Specifications

Transceiver	Speed	Fiber Type	Wavelength	Minimum Range	Maximum Suggested Range
1000SX	1 Gigabit	MMF – 62.5μ	850 nm	2-275 m	1.0 km
1000SX	1 Gigabit	MMF – 50μ	850 nm	2-550 m	1.0 km
1000LX *	1 Gigabit	MMF – 62.5μ	1310 nm	2-550 m	8.5 km
1000LX *	1 Gigabit	MMF – 50μ	1310 nm	2-550 m	8.5 km
1000LX	1 Gigabit	SMF – 9μ	1310 nm	2m-10 km	32.0 km
1000XD	1 Gigabit	SMF – 9μ	1550 nm	See Notes	50.0 km
1000ZX	1 Gigabit	SMF – 9μ	1550 nm	See Notes	70.0 km
CWDM	1 Gigabit	SMF – 9μ	1470-1610 nm	See Notes	See Note

* LX over MMF may require mode conditioning patch cables (single mode/multimode hybrid)

Notes:

1000XD GBIC – if range is less than 25 km, use a 5 dB in-line attenuator

1000ZX GBIC – if range is less than 25 km, use a 10 dB in-line attenuator, and if range is less than 50 km, use a 5 dB in-line attenuator

CWDM ranges vary with the SFP or GBIC that is used. Two versions of SFP are available (40 km and 70 km) and one version of GBIC (120 km). Please refer to product documentation for optical specifications and possible maximum ranges for CWDM.

Table 2.20: GBIC / SFP Specifications

It is imperative to preserve the integrity of the uplinks from the edge closet to the core/distribution layer. In the case of a single fiber fault – either the transmit or the receive – the link must be automatically disabled at both ends. If this does not happen, data could be passed to a port that is not operating properly, and that data would be lost. This issue can cause severe performance degradation and eventually render the network inoperable. To protect the network, it is important to properly enable some form of single fiber fault detection on all uplink ports. VLACP can also be enabled as another single fiber fault detection mechanism.

All uplink ports should be configured with a method to detect a single fiber fault and disable the affected ports. There are two options for enabling this feature, depending on the switching platform being utilized at the edge:

➢ Switches supporting Autonegotiation on the uplink ports should have that feature enabled on both ends of the uplink.

▪ On the Gigabit ports, Remote Fault Identification (RFI) is enabled by using Autonegotiation. RFI removes the link from a port in the event of a single fiber fault on the link connected to that port.

- On 100 Mbps ports, Far End Fault Identification (FEFI) is enabled by using Autonegotiation. FEFI removes the link from a port in the event of a single fiber fault on the link connected to that port.
  - ➢ 10-Gigabit does not support autonegotiation; however, RFI is enabled as part of the physical layer implementation automatically.

In certain scenarios where the link may span across a providers LAN extension service, detecting a link failure in the LAN extension core will not work using Autonegotiation and RFI. This will not work end to end between a pair of Avaya switches because RFI only operates between direct connected switches. Hence, if there is a failure in the LAN extension core, the link on both Avaya switches will still be running. To solve this problem, enable VLACP on the Avaya switches. The VLACP protocol is forwarded between the Avaya switches. If the switch does not receive any VLACP updates, a link will be declared out-of-service and traffic will be forwarded through another working link. A detailed discussion of VLACP is covered in an upcoming section.

2.2.5 Autonegotiation

The autonegotiation standard for Ethernet allows end stations to connect at their most optimal speed and duplex – anything from 10 Mbps half-duplex up to 1 Gbps full duplex. This feature allows different end stations with different connectivity capabilities to connect to a single network without the intervention of the network administrator.

Autonegotiation must be enabled on fiber ports to support Remote Fault Identification (RFI) for Gigabit and Far End Fault Identification (FEFI) for 100FX connections. These features shut down a port in the case of a single fiber fault at the remote end of the link. A more detailed description is covered in the section on fiber fault detection.

10-Gigabit Ethernet does not have the concept of autonegotiation. By default, 10-Gigabit links are full duplex. The RFI component of autonegotiation is built into the 10-Gigabit physical interface and therefore is automatically enabled.

It is critical to verify that both ends of the link are capable of supporting autonegotiation. If one end is not able to support it, then Autonegotiation must be disabled on both ends of the link. Having autonegotiation enabled on one end and disabled on the other end is a common configuration error that can cause severe performance degradation of that connection. Excessive FCS (Frame Check Sequence) errors on a port are a common indicator of a speed/duplex mismatch between the two devices.

In certain situations, it is useful to autonegotiate to a specific speed and duplex value by controlling which capabilities are being advertised from the switch. Avaya introduced the Custom Auto-Negotiation Advertisement (CANA) feature to accommodate this need. This feature allows the administrator to control which advertisements are made by the switch. For example, if the switch only advertises a 100 Mbps full duplex capability on a specific link, then the link is only activated if the neighboring device is also capable of autonegotiating a 100 Mbps full duplex capability. This prevents mismatched speed/duplex modes if customers disable autonegotiation on the neighboring device.

- ➢ Enable autonegotiation on all switch-to-switch ports to ensure single fiber fault detection.
- ➢ Enable autonegotiation on all end station ports, ensuring that those stations are capable of supporting autonegotiation.

- ➢ Where required, use Customized Auto-Negotiation Advertisements (CANA) to control end station connectivity speed and duplex.

- ➢ Disable autonegotiation on problematic devices. Because the autonegotiation standard is rather broad, there are some devices that will not connect properly when autonegotiation is enabled on both ends.

- ➢ When using autonegotiation, always have the most recent Network Interface Card (NIC) driver from the manufacturer.

2.2.6 Link Aggregation

To increase both resiliency and bandwidth from the edge-wiring closet, Avaya recommends implementing link aggregation of the uplinks. Avaya supports multiple options for link aggregation on the Ethernet switches, MultiLink Trunking (MLT), Switch Clustering using Split MultiLink Trunking (SMLT), and 802.3ad.

Multiple physical links compose logical trunk

Logical MLT or 802.3ad Trunk
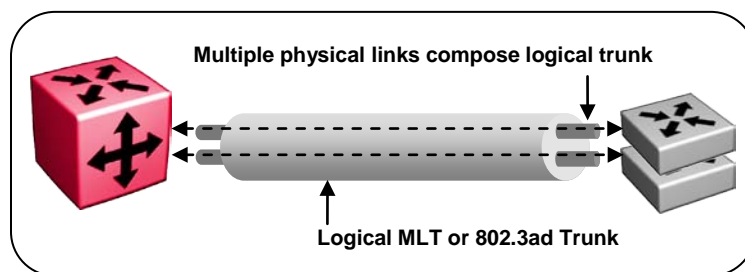
Figure 2.34: Link Aggregation

MultiLink Trunking (MLT) provides the ability to group multiple physical links into a single logical link (see Table 2.12 for the specific number of links and groups supported per switch or stack). MLT automatically increases bandwidth from the wiring closet by utilizing all the physical links in a logical group. A failure of any physical links results in automatic sub-second failover of the traffic to the remaining links within the MLT group. After the failed link has been repaired, recovery of that link back into the MLT group is also accomplished in sub-second time.

Distributed MultiLink Trunking (DMLT) adds the ability to terminate the physical links of the trunk group on different switches within a stack or on different modules within a chassis. This feature significantly increases the resiliency of the uplinks out of the wiring closet. A failure of the switch or module on which one of these physical links terminates will not cut off communication from the wiring closet to the core/distribution layer.

The IEEE standard for link aggregation is 802.3ad. This standard uses a Link Aggregation Control Protocol (LACP) to aggregate multiple physical links into a single logical link aggregation group (LAG) from the Ethernet switch. Adherence to the IEEE standard will help to ensure interoperability of link aggregation between different vendor equipment (switches, NICs, etc.).

Although IEEE 802.3ad-based link aggregation and MLT provide similar services, MLT is statically defined, whereas IEEE 802.3ad-based link aggregation is dynamic and provides added functionality through the Link Aggregation Control Protocol (LACP). LACP dynamically detects whether links can be aggregated into a link aggregation group and does so when links become available. IEEE 802.3ad was designed for point-to-point link aggregation only. The Avaya Ethernet switches support the formation of 802.3ad link aggregation groups across different

switches in a stack or different I/O modules in a chassis to provide additional resiliency in the event on a switch failure in a stack or an I/O module failure in a chassis.

The added functionality of end-to-end checking is a resiliency feature that should be used within a Converged Campus design. Because MLT is statically defined, there is no mechanism for checking between directly connected switches. In the rare event that a switch may become inoperable but the link status remains up, data could be inadvertently sent to that switch, causing it to be lost. A checking mechanism automatically removes that port from the aggregation group, ensuring data is not sent to an inoperable switch, and therefore ensuring no data loss.

There are limitations to LACP that must be considered:

> LACP scaling can be somewhat limited, please see Table 2.12 below for Ethernet Switch support of LACP and scaling numbers

> LACP was designed to operate between two directly connected switches and will not work in an end-to-end fashion if there are any intermediary devices (Optical ring, Service Provider Network, etc.).

> Failover times of LACP are higher than that of MLT – default value is a 30 second poll and 3 polls must be missed to disable a port. Thus, it will take LACP 90 seconds to failover by default. Fast timers are available which reduce the poll time to 1 second intervals, thus reducing failover time to 3 seconds.

For the purposes of the Large Campus solution, a MultiLink Trunk Group (MLT) will be created on the Edge switch/stack connecting to the ERS 8800/8600 Switch Cluster core. Each closet containing edge switches/stacks will have an MLT created to the core of the network.
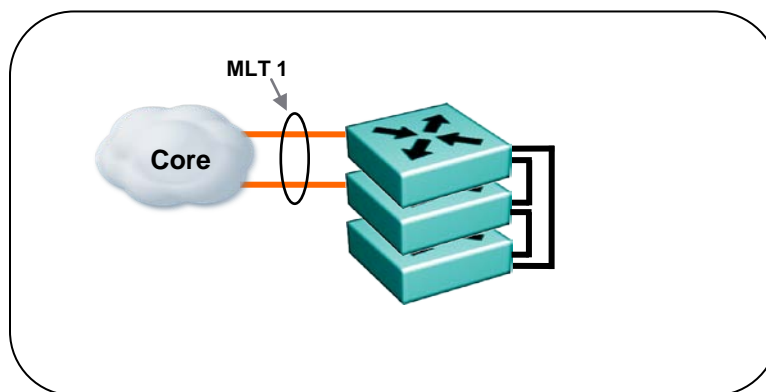
Figure 2.35: Edge Switch Link Aggregation

> Create an MLT group with id 1 and name it MLT-to-ERS8000. The MLT id on the edge stack can be the same across the network. This id is only locally significant to that stack and is not propagated outside the local stack.

> Spanning Tree Protocol must be disabled for the MLT when using it with SMLT/SLT in the core of the network.

> The uplink ports will be assigned as members of the MLT group. The uplink ports are 802.1Q tagged and members of the Data, Voice, and Management VLANs, along with any other VLANs used at the Edge.

> When in a stack configuration at the edge, use DMLT (Distributed MLT). This adds another layer of resiliency to the design by spreading the MLT across different units in the stack.

> Autonegotiation should be enabled (on by default) on the uplink ports to enable Remote Fault Indication (prevention of single fiber faults)

> All links in an MLT group must be of the same speed and duplex – Avaya does not support the mixing of different speed links within the same MLT group.

> For specific configuration rules about link aggregation, see the product documentation.

> Avaya recommends using MLT whenever possible, however, if using LACP review the following:

   ▪ LACP supports a maximum of eight active links, all other links (nine and above) are put into standby

   ▪ Active/Standby is defined by ActorPortPriority (higher actor priority = lower port priority)

   ▪ If actor priority is the same, lower MAC = higher priority

   ▪ Use the same ID number for the ActorAdmin key and MLT ID

The following table depicts the support and scaling of MLT and 802.3ad on Avaya Ethernet switches.

Switch Model	Links per Group	Groups per Switch or Stack	802.3ad Support	LACP-MLT Scaling	LACP-SMLT Scaling
ERS 8600 Legacy Modules	8	32	Rls.3.7	32	32
ERS 8800/8800 R and RS Modules	8	128	Rls 3.7	128	128
ERS 8300	4	31*	Rls 4.1	31	31
ERS 5000	8	32	Rls 4.1	32	32
ERS 4500	4	8	Rls 5.0	32	N/A
ERS 2500	4	6	Rls 4.0	6	N/A
* Up to seven Fast Ethernet groups and/or 31 Gigabit groups					

Table 2.21: LACP / VLACP Support and Scaling

2.2.7 VLANs

The ERS 5000 series supports up to 1024 VLANs while the ERS 4500 and ERS 2500 series support up to 256 VLANs. The following details the basic VLAN configuration parameters to consider and configure:

  ➤ Assign general use VLANs by geographic location if possible (by closet or floor). This practice limits the need to bridge VLANs throughout the network, thus reducing administrative overhead. It also aids in troubleshooting any network or application issues that may arise. An exception to this practice occurs when creating a VLAN for a specific protocol or application that may need to be bridged throughout the network.

  ➤ VLAN config control is strict – Globally enabled by default. In this mode, you must first remove port members from the default VLAN 1 prior to adding these port members to a new VLAN.

  ➤ Data VLAN and Voice VLAN to separate traffic. These VLANs can be modified to suit the specific environment needs and additional VLANs may be required for specific network services or user connectivity.

  ➤ Port Membership – All switch ports will be removed from the default VLAN (1) and added as members to the Data VLAN and Voice VLAN – the PVID for the ports must be the Data VLAN. See Section 2.2.11 Quality of Service for a discussion about dynamically creating the voice VLAN using Auto Detect Auto Config (ADAC) on the ERS stackables.

  ➤ Always enable 802.1Q VLAN tagging on uplink ports and have all VLANs (except VLAN 1) added to them. Even if only one VLAN is used at the edge, this enables the addition of more VLANs in the future without disrupting existing traffic. Enabling 802.1Q VLAN tagging adds a Q tag to every packet on the uplink in order to maintain the VLAN separation across the link.

  ➤ Avoid using the default VLAN whenever possible. This helps to minimize the possibility of accidentally creating Layer 2 loops in the network. Because the same default VLAN exists on every switch, it is very easy to incorrectly connect all these VLANs together and create unexpected traffic flows in the network. Please note that you cannot delete the default VLAN, therefore, it is recommended to remove valid port members from the default VLAN.

  ➤ Management – Create a separate management VLAN.
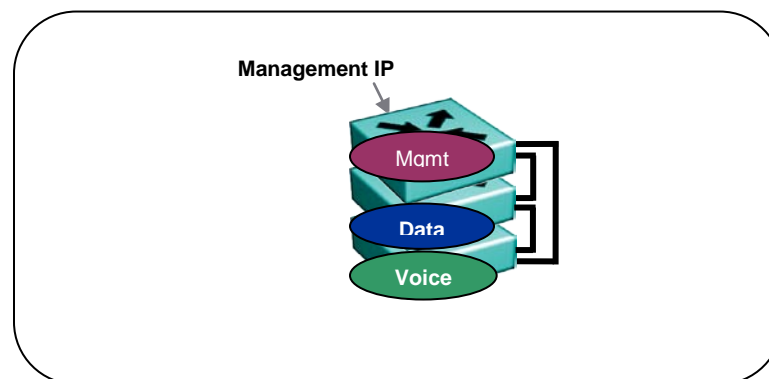
Figure 2.36: Edge Switch VLANs

2.2.8 Filter Untagged Frames

To provide protection against a factory defaulted or mis-configured device being connected to the uplink ports on the Edge switch/stack, which can result in a loop, the Filter Untagged Frames feature should be enabled on the MLT ports.
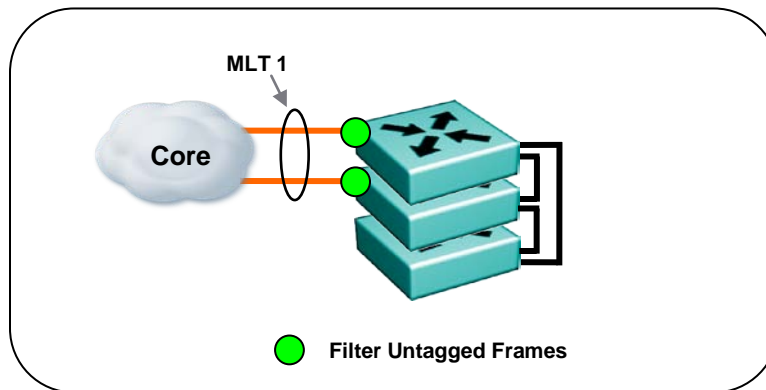
Figure 2.37: Edge Switch Filter Untagged Frames

2.2.9 Spanning Tree Protocol

The IEEE 802.1D Spanning Tree protocol is used to prevent loops in the network. Usually, these loops occur when the design includes redundant connections from the edge to the core, or when multiple wiring closets are inadvertently interconnected. A loop in the network causes severe congestion and eventually renders the network inoperable.

Although there are newer versions of Spanning Tree protocols, such as 802.1w Rapid Spanning Tree (RSTP) and 802.1s Multiple Spanning Tree Groups (MSTP), these protocols are still based on the legacy 802.1D STP fundamental architecture and therefore have limitations. Although RSTP offers faster failover than normal 802.1D Spanning Tree, it still has the same problem as that of 802.1D: all redundant or looped paths are blocked.

MSTP does allow load balancing of VLANs over redundant paths; however, this requires configuration of every switch to assign cost or weight to all available paths for each VLAN, which can lead to administrative difficulties when there are a large number of switches and/or subnets in the network.

Other issues to be considered in Spanning Tree environments include the need to set a root bridge for the network. The root should be configured on one of the core switches in the network. When using VRRP for default gateway redundancy in conjunction with Spanning Tree, the VRRP Master should be configured on the same switch as the Spanning Tree root (for optimal link performance). Utilizing the Backup Master feature on the Avaya switches will also aide in the optimal link usage with VRRP – both VRRP and VRRP Backup Master are discussed in more detail in the Layer 3 section of this document. If PIM-SM is configured on the network, the Designated Router (DR) needs to also be configured on the switch that is the Spanning Tree root.

Avaya does recommend using the Spanning Tree protocol on all end station connections in order to safeguard the network from hubs or other devices that could be inserted into the network at the

end station. A modification to the normal learning of spanning tree is available in all Avaya edge switches. This feature is known as Fast Start or Fast Learning, and is the recommended setting for all end station ports.

The BPDU filtering feature also adds a level of protection against inadvertent loops in the network. This feature was originally developed to prevent an unwanted root selection process when a new device was added to a Spanning Tree network and/or to prevent unknown devices from influencing an existing spanning tree topology. A more detailed discussion on BPDU filtering follows in the next section.
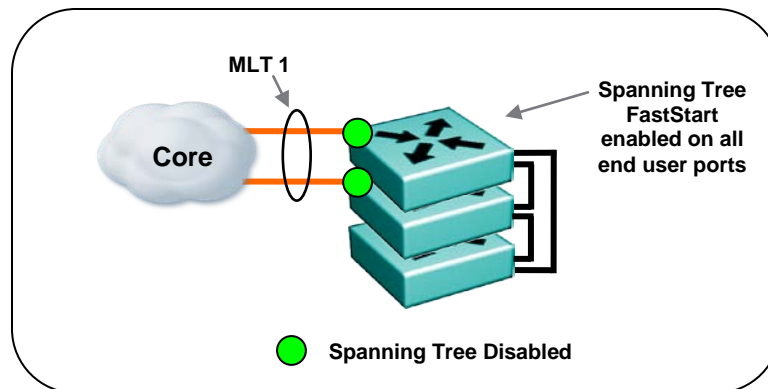
Figure 2.38: Edge Switch Spanning Tree

> ➢ Enable Spanning Tree Fast Start/Fast Learning on all end station ports.

> ➢ Enable BPDU filtering on all end station ports.

> ➢ Never enable Fast Start/Learning on any uplink ports; this could cause unexpected behaviors on the entire network.

> ➢ When using Spanning Tree, pay attention to the root bridge. Ensure the root bridge is one of the core switches by configuring the Spanning Tree priority.

> ➢ When using SMLT to connect the edge to the distribution/core, always disable Spanning Tree on the uplink ports/MLT of the edge switch.

Avaya recommends using Split MultiLink Trunking (SMLT) to interconnect closets to the core of the network, thus eliminating the need for the Spanning Tree protocol on uplinks. When using SMLT between the edge switch and the core or distribution switch, two or more redundant paths to two separate core/distribution switches are utilized in an active-active fashion without the need for Spanning Tree to prevent loops. Traffic is distributed over all available paths using either MLT, 802.3ad, or any other form of link aggregation. If one or more of the paths fail, including link and/or switch failures, SMLT provides sub-second failover to the remaining path(s).

2.2.10 BPDU Filtering

BPDU filtering is normally used as a protection mechanism within a Spanning Tree network. This feature blocks an unwanted root selection process when a device is added to the network and also blocks BPDU packets from ingressing the port. This feature is also a valuable protection mechanism in a Switch Cluster network as it will block loops created by connecting Edge closets directly together, therefore it is recommended to enable BPDU Filtering on all end user ports.

➢ Set Timeout to 0 – requires manual intervention to re-enable the port
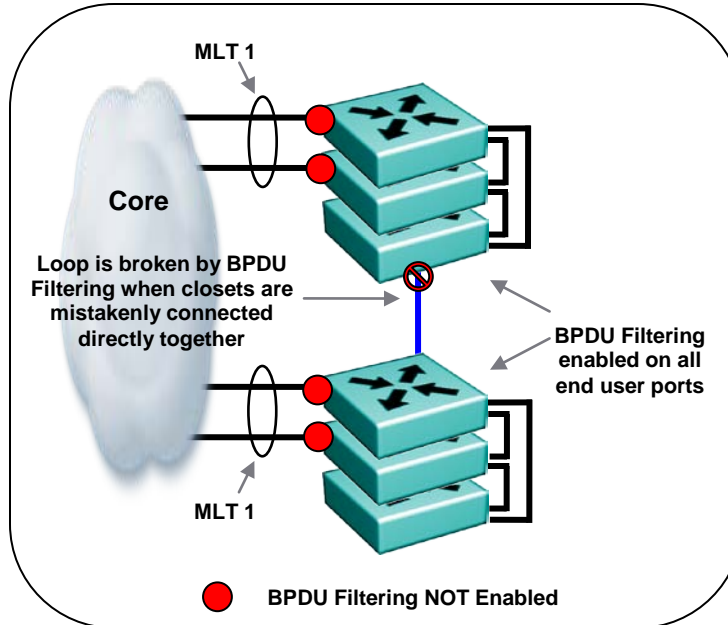
Figure 2.39: Edge Switch BPDU Filtering

2.2.11 VLACP

This Avaya feature provides an end-to-end failure detection mechanism which will help to prevent potential problems caused by mis-configurations, a switch being defaulted, or links connected incorrectly. In the Large Campus design, VLACP will be used between the ERS 8800/8600 Switch Cluster and all uplink ports to the ERS 2500/4500/5000 switches at the edge.

➢ Globally configure VLACP to use the reserved multicast MAC of 01-80-C2-00-00-0F

➢ For the SLT/SMLT links, use the short timeout

- Fast periodic timer of 500 msecs * timeout scale of 5

- Make sure these values match on both ends of the links
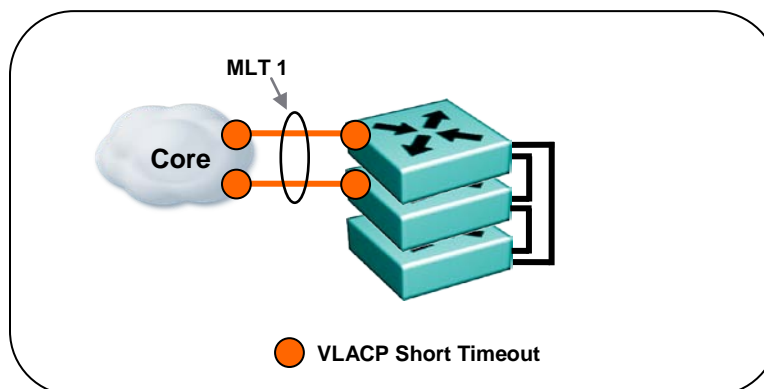
Figure 2.40: Edge Closet VLACP

2.2.12 Rate Limiting

Port level rate limiting limits packets with broadcast and/or multicast addresses to control the amount of traffic on a port. Rate limiting is used to protect the network from non-CPU bound traffic. This functionality is configured on a per port basis. For each port, the network administrator can configure a rate limit for broadcast traffic and a rate limit for multicast traffic. These rates are the maximum allowed amount of that traffic type on that specific port. When traffic exceeds the configured threshold, it is dropped. The design recommendations below detail the way broadcast/multicast traffic is calculated (%, pps, kbps) and the values that should be used. Be aware that these are rule-of-thumb values for the "typical" enterprise network.

It is extremely important to understand the network and application environment before configuring the rate limiting feature. In certain environments, there will naturally be a higher rate of a traffic type due to the applications being used. For example, in a network that utilizes multimedia communications such as streaming video and video on demand, there will likely be a higher rate of multicast traffic and rate limiting this traffic may adversely affect the applications being used.

If rate limiting is implemented, it should be done at the edge of the network closest to the user as possible. This will have the greatest effect on overall traffic as limiting will occur before hitting the core of the network.

The following details the implementation of rate limiting on each of the ERS platforms and the recommended values to be used.

ERS 2500 / 4500 / 5000

- ➢ 1 – 10% of port speed
- ➢ Recommendation → 10%

ERS 8300

- ➢ 1-100% of port speed
- ➢ Recommendation → 10%

ERS 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

2.2.13 Quality of Service

To provide appropriate QoS treatment for Voice traffic, DiffServ will be enabled on the Edge Switches/Stacks. The QoS capabilities of the Edge switching platforms are detailed below. Although Voice traffic is normally, the driver for a QoS implementation, any real time delay intolerant or business critical traffic can take advantage of QoS over the network.

From a generic QoS perspective, the Edge switches support the following:

The ERS 2500 presently can be configured to honor DSCP markings within the packets entering the switch and placing them in the appropriate egress queue.

> Configure the ERS 2500 to honor DSCP markings on all ports to ensure proper QoS

The ERS 4500 and ERS 5000 series have advanced QoS features and can be configured to mark and/or honor DSCP markings within the packets entering the switch and place them in the appropriate egress queue.

> Configure the ERS 4500 or ERS 5000 to prioritize any traffic type that requires QoS throughout the network. This can be done by creating filters to identify the traffic (tcp port, udp port, src/dst IP, etc.) and then apply the appropriate DSCP marking to the packets.

When deploying IP Telephony, Avaya provides different options for the configuration of the Ethernet switches to ease the deployment and later moves, adds, and changes that occur within the normal Enterprise network. Each of these options will be described briefly here and references to Solution Guides and Configuration Guides are provided for more in-depth coverage.

> The ADAC (Auto Detect Auto Config) feature will automatically discover the IP phone once plugged into the Ethernet edge switch. Once discovered, the Ethernet edge switch will automatically provision the appropriate Voice VLAN and QoS to the port as well as on the uplink to the core.

ADAC will use 802.1AB LLDP or the phone's MAC address as the discovery mechanism. This fact allows ADAC to be IP phone agnostic and will work with any 3rd party device.

For a detailed discussion of ADAC and VoIP deployment, please refer to the *Avaya IP Phone Set InterWorking with Avaya ES and ERS Switches TCG* NN48500-517.
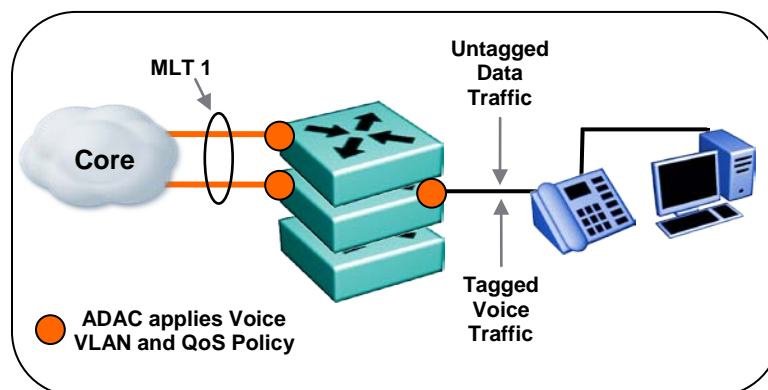
Figure 2.41: ADAC in Tagged Frames Mode

> The Avaya Automatic QoS feature will automatically discover the IP phone once plugged into the Ethernet edge switch. Once discovered, the Ethernet edge switch will automatically honor the QoS marking (DSCP) of the Avaya IP phone.

> A big advantage of Avaya Automatic QoS is its ability to recognize the 2050 Softphone and provide QoS treatment at the edge for these devices.

> Please note that if Voice VLANs are required they will need to be manually configured on the Ethernet edge switches, both for the end stations and the uplinks.

> For a detailed discussion of ADAC and VoIP deployment, please refer to the *Avaya Automatic QoS Technical Configuration Guide* NN48500-576.
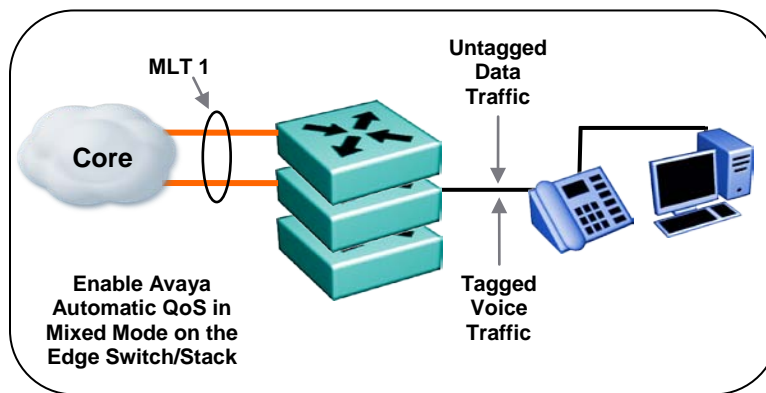
Figure 2.42: Avaya Automatic QoS

2.2.14 Security

The Edge switches support security features to help protect the network from denial of service (DoS) attacks. DHCP Snooping, Dynamic ARP Inspection, and IP Source Guard should be enabled on all edge switches to help ensure the protection and integrity of the Campus network. See Section 2.5 for a more detailed discussion about security features.

> DHCP Snooping prevents DHCP spoofing by creating a DHCP binding table to help ensure that no rogue DHCP servers can be inserted into the network

> Dynamic ARP Inspection examines ARP packets to prevent man-in-the-middle attacks

> IP Source Guard filters clients with invalid IP addresses

> Untrusted ports are normally the end user access ports on the switch

> Trusted ports are normally the uplinks from the edge switch to the core

> Dynamic ARP Inspection and IP Source Guard use the binding table created by DHCP Snooping, therefore, in order to use these features, DHCP Snooping must be enabled.

> IP Source Guard should not be enabled on uplink ports from the Edge to the Core, only enable on Edge access ports (untrusted ports) where DHCP Snooping and Dynamic ARP Inspection are enabled.
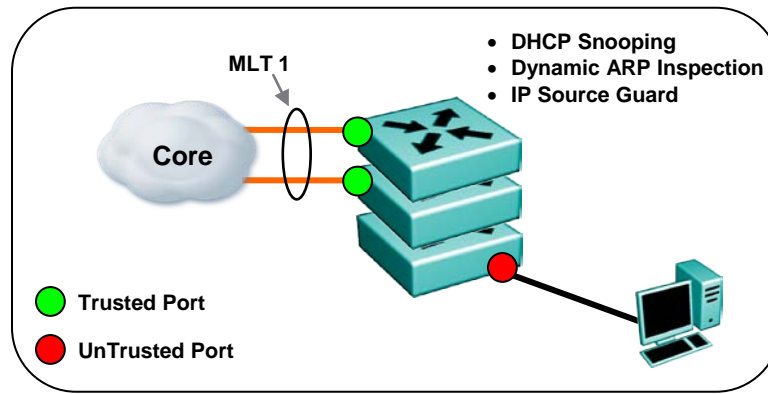
Figure 2.43: Edge Switch Security

2.2.15 Multicast

The multicast protocol distributes traffic to all subscribers of a multicast group. The Edge Switch/Stack is functioning at Layer 2 in this design and therefore IGMP features must be properly configured for the most efficient use of the bandwidth available.

> IGMP Snooping

The Avaya Ethernet switches can sense IGMP (Internet Group Multicast Protocol) host membership requests for each specific multicast group. Only host ports requesting a multicast stream receive that stream; the switch automatically prunes the other ports and does not send multicast traffic to hosts that did not request it, thus making efficient use of the available bandwidth to each of the hosts on the switch.

> IGMP Proxy

The Avaya Ethernet switch provides a single proxy report upstream for all members within the same multicast group on the same switch/stack. By consolidating all the IGMP host membership requests into a single request, the switch does not flood the network needlessly with multiple copies of the same request. IGMP Snooping must be enabled for this feature to work.

> IGMP Static Router Port

This feature allows unknown multicast traffic to be forwarded only to the statically defined multicast router port. The traffic will not be flooded to all ports and will not be sent to dynamically learned multicast router ports. Static router ports serve two purposes: to get (1) multicast traffic and (2) IGMP reports to multicast routers that may not be discoverable through passive detection; for example, when there are two queriers and one is elected and the other becomes silent (per the IGMP standard).

If IGMP Snooping/Proxy is not enabled, multicast traffic is flooded to all ports on the switch that are in the same VLAN.
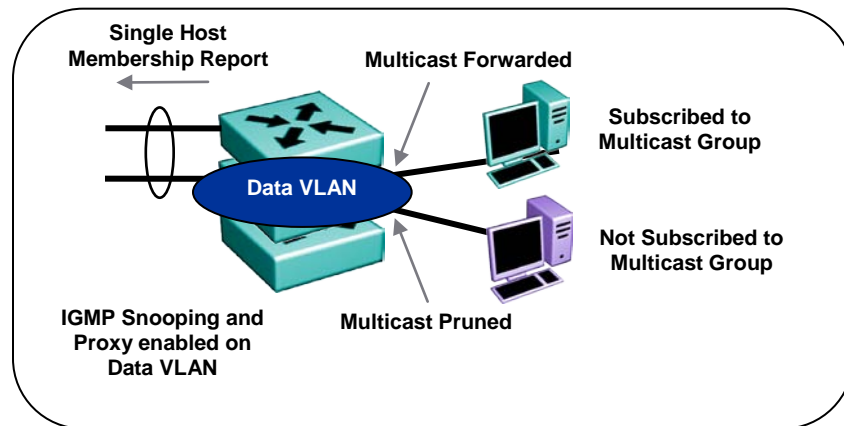
Figure 2.44: Edge Switch Multicast

➢ Enable IGMP Snooping on VLANs that have Multicast applications running on them – this prunes multicast traffic from end user ports that are not subscribed to the multicast group

➢ Enable IGMP Proxy on the VLANs that have Multicast applications running on them – this will consolidate IGMP reports into a single report to the upstream network and is thus a much more efficient use of the bandwidth

➢ IGMP Proxy must have IGMP Snooping enabled before it takes effect

➢ IGMP Snooping and Proxy require an IGMP Querier to be present on the network. This function is performed by the Layer 3 Router with a Multicast routing protocol enabled or with the layer 2 multicast querier function available on the ERS 8800/8600 and ERS 8300

➢ If there are no Multicast applications being used on the network, leave IGMP snooping and IGMP proxy disabled (default value) on all the VLANs.

➢ Flooding unknown multicast traffic is a behavior that is configurable (on or off) on some models of the Ethernet edge switches. The "vlan igmp unknown-mcast-no-flood" command provides this functionality. Turning off flooding ensures that static router ports become the destination of unknown multicast traffic.

2.3 Network Access Control

Avaya's Identity Engines is the framework for role-based network access control. Within this framework there are several options to best accommodate the needs of the Enterprise customer, from simple MAC authentication to full 802.1X authentication and posture assessment (end station compliancy to corporate security policies) of the end user's workstation. With all the methods available, the end result is to ensure users are allowed on the network and permitted access to resources based on identity and credentials.

This section will describe the backend infrastructure required (Identity Engines) along with the options available for end user authentication.

2.3.1.1 Identity Engines

The Avaya Identity Engines portfolio integrates with any current network infrastructures to provide the central policy decision needed to enforce role-based Network Access Control (NAC). This is accomplished by combining the best elements of a next-generation RADIUS/AAA server, the deep directory integration found in application identity offerings, and one of the industry's most advanced standards-based policy engines. All this is done out-of-band for maximum scalability and cost effectiveness.

The centralized policy engine sits in the data center to provide centralized authentication and authorization for wired, wireless, and VPN network devices. It is closely aligned with Avaya and third-party Ethernet switching, WLAN and VPN products as it provides centralized integrated security services for these network devices.

Coupled with the centralized policy engine is a suite of complementary products that enable 802.1X rollouts for wired and wireless networks, while unifying those policies with existing VPN rules to achieve audit and compliance goals. These products offer a holistic network identity management solution which involves all aspects of managing how users access networks. Benefits include admission control, temporary user provisioning, policy decisions and directory integration.

Identity Engines Ignition Server

A state-of-the-art network identity management solution with a powerful policy engine to centralize, streamline and secure access across the network. The Identity Engines Ignition Server offers a new level of accuracy, with identity- and policy-based control over who accesses the network, where, when, how, and with what type of device. Easy to deploy and use, it is a powerful, scalable foundation for network access control, guest access, secure wireless, compliance, and more

Identity Engines Ignition Posture

Identity Engines Ignition Posture provides endpoint health and posture checking that works in the real world. Most posture checking products today are inflexible, add-on layers that are expensive to support and frustrating for network users. In contrast, this product provides policy flexibility and integration with the Identity Engines Ignition Server to ensure that it is easier to support and less frustrating for users.

Identity Engines Ignition Guest Manager

Because guests and visitors often have legitimate reasons to access networks, Identity Engines Ignition Guest Manager makes it easy and safe for organizations to let front-desk staff create guest user accounts. Simple delegation rules ensure front-desk personnel can give guests access to only specified network resources, and each guest account expires automatically after a designated period.

Identity Engines Ignition Analytics

Identity Engines Ignition Analytics is a powerful reporting application that allows organizations to perform in-depth analysis of network activity including ingress and usage. With over 25 preconfigured audit, compliance and usage reports, organizations can easily produce multiple custom reports to fulfill its specific reporting requirements.
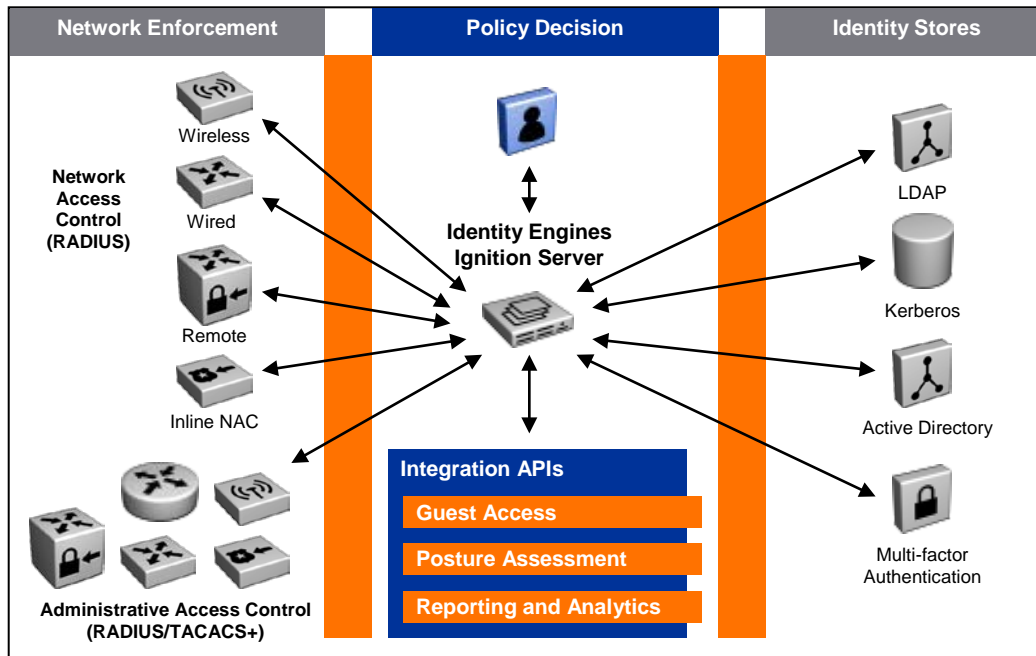
Figure 2.45: Identity Engines Portfolio Architecture

2.3.1.2 MAC Based Authentication

The Media Access Control (MAC) address-based security feature is based on Avaya BaySecure local area network (LAN) Access for Ethernet, a real-time security system that safeguards Ethernet networks from unauthorized surveillance and intrusion. You can use the MAC address-based security feature to configure network access control, based on the source MAC addresses of authorized stations.

➢ Local Authentication

The MAC Address security feature allows the administrator to specify a list of MAC addresses that are authorized to access the switch. You can also specify the ports that each MAC address is allowed to access. The options for allowed port access include: NONE, ALL, and single or multiple ports that are specified in a list.

➢ Auto-learning or Static Authentication

Auto-learning allows a switch to automatically add a MAC address to the MAC security table without user intervention. You can also limit the number of MAC addresses allowed per port.

Static authentication allows the manual creation of static MAC entries on a per port basis.

➢ Centralized Authentication

This feature functions the same way as local authentication; however, the list of allowed MAC addresses is stored in a RADIUS server. This is a much more manageable approach to MAC security. Dynamic VLAN assignment is supported for MAC authenticated clients.
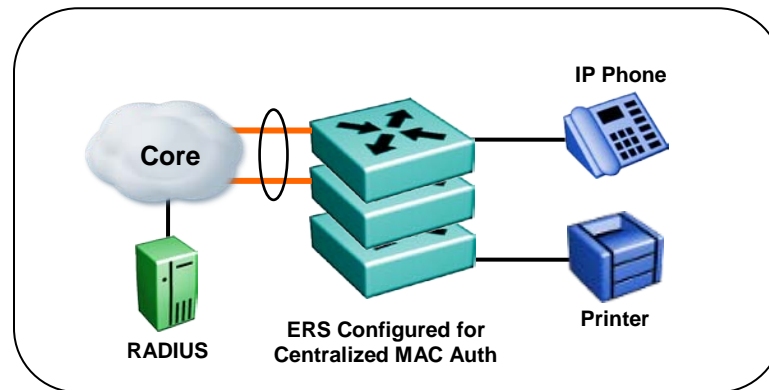
Figure 2.46: MAC Based Authentication

2.3.1.3   802.1X Extensible Authentication over LAN (EAPoL)

The Ethernet Routing Switches use an encapsulation mechanism to provide security, referred to as the Extensible Authentication Protocol over LAN (EAPOL). This concept uses the Extensible Authentication Protocol (EAP) as described in the IEEE 802.1X standard allowing the set up of network access control on internal LANs. EAP allows the exchange of authentication information between an end station or server connected to the switch and an authentication server, such as a RADIUS server. The EAPOL-based security feature operates in conjunction with a RADIUS-based server to extend the benefits of remote authentication to internal LAN clients.

A variation of 802.1X allows for the configuration of a guest VLAN, which allows non-authenticated users on the network, but also allows the administrator to control what the guest VLAN users can access. Other variations of 802.1X allow for multiple authenticated users on the same physical Ethernet switch port, or the ability to support both 802.1X and MAC authenticated users on the same port. This flexibility in using 802.1X makes it much easier to deploy scenarios with IP Telephony where the PC is connected to the IP phone and they share a single Ethernet switch port in the closet.

> Guest VLAN

Allows non-EAP users connected on EAP-enabled ports access to a guest network. This feature allows network access to users through the guest VLAN. This VLAN is port-based, configured on a per switch/stack basis, and is enabled per port. Once a user completes EAP authentication, the user is moved from the Guest VLAN to the authenticated VLAN.

> Single Host Single Authentication (SHSA)

For an EAP-enabled port with SHSA, at any time only one client (single MAC address) is authenticated on a port, which is assigned to only one port-based VLAN. Only a particular client who completes EAP negotiations on the port is allowed access to that port for traffic.

▪ Dynamic VLAN assignment for the client is supported by RADIUS attributes passed back to the Ethernet switch

▪ Guest VLAN is supported in conjunction with SHSA
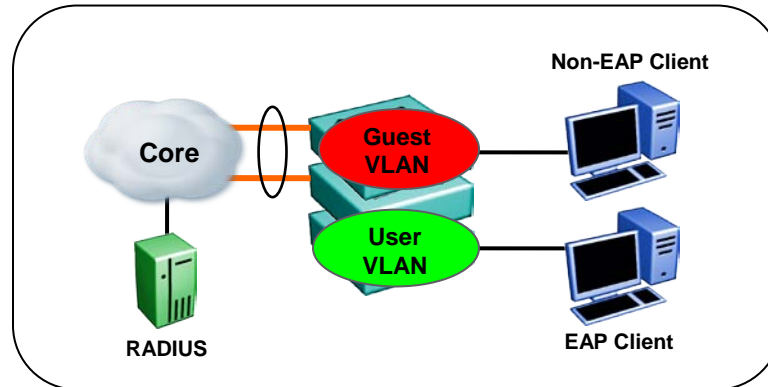
Figure 2.47: 802.1X SHSA

➢ Multiple Host Multiple Authentication (MHMA)

For an EAP-enabled port with MHMA, a finite number of clients or devices with unique MAC addresses are allowed access to a port. Each client must complete EAP authentication to enable the port to allow traffic from the corresponding MAC address. Only traffic from the authorized hosts is allowed on that port.

- Dynamic VLAN assignment for the client is supported by RADIUS attributes passed back to the Ethernet switch – note that the dynamic VLAN is supported for the first authenticated client only, subsequent client authentications will automatically be put into that initial dynamically assigned VLAN.

- Guest VLAN is supported in conjunction with MHMA – after first successful EAP authentication, the Guest VLAN is no longer accessible on that port.
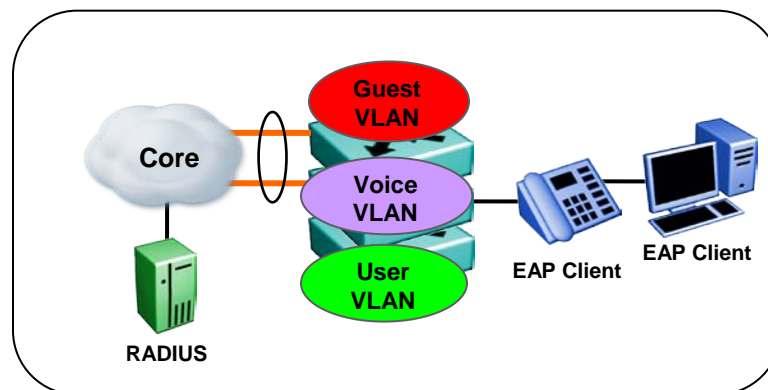
Figure 2.48: 802.1X MHMA

➢ Non-EAP MAC Authentication

Allows non-EAP users connected on EAP-enabled ports access to the network. This feature uses MAC authentication for the non-EAP user. This authentication can be local MAC authentication on the Ethernet switch or centralized MAC authentication via RADIUS – depending on Ethernet switch feature support. Please note that user-based policies can be used with MAC Authentication. The non-EAP hosts are permitted on the port even if there are no EAP authenticated hosts on the port.

- Guest VLAN and Non-EAP are mutually exclusive (ERS 4500 with Release 5.3 software removes this limitation)

➢ Non-EAP IP Phone Authentication

The "non-eap-phone-enable" feature allows Avaya IP phones (heritage Nortel versions only) on EAP-enabled ports without the need for the phone MAC to be pre-configured in the MAC list (local or centralized). The switch will look for the phone signature enclosed in the DHCP Discover packet. If the proper signature is found, the switch will register the MAC address of the IP Phone as an authenticated MAC address and will let the phone traffic pass on the port.
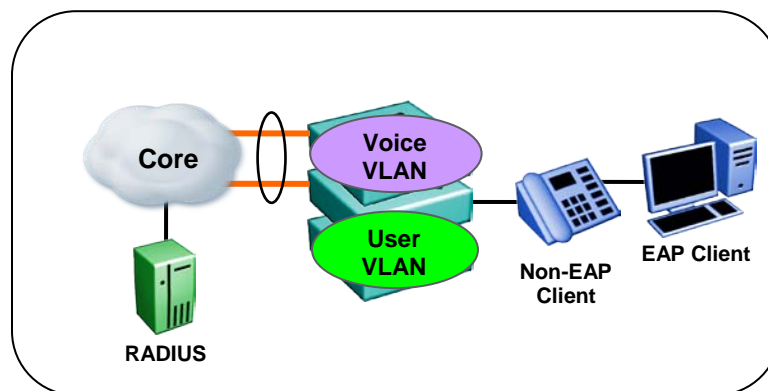
Figure 2.49: 802.1X Non-EAP Phone Authentication

➢ Multiple Host Single Authentication (MHSA)

Multiple Host Single Authentication (MHSA) is a more restrictive implementation of support for non-EAPOL hosts on EAPOL-enabled ports. For an EAPOL-enabled port configured for MHSA, one EAPOL user must successfully authenticate before a finite number of non-EAPOL users or devices with unique MAC addresses can access the port without authentication. The MHSA feature is intended primarily to accommodate printers and other passive devices sharing a hub with EAPOL clients.

AUTHENTICATION FEATURES	ERS 2500 V 4.4	ERS 4500 V 5.5	ERS 5000 V 6.2
Single Host Single Authentication (SHSA)	Y	Y	Y
Multi Host Single Authentication (MHSA)	Y	Y	Y
Multi Host Multi Authentication (MHMA)	Y	Y	Y
Guest VLAN (GV)	Y	Y	Y
RADIUS Assigned VLANs for EAP MACs	Y	Y	Y
Port-mirroring on 802.1x (EAP) ports	N	Y	Y
EAP/NEAP Separation	N	N	Y
User Based Policies (UBP)	N	N	Y
User Based Policies v2 (UBPv2)	N	N	Y
NEAP User Based Policies	N	N	Y
NonEap Support on EAP enabled ports (NEAP)	Y	Y	Y
IP-Phone Authentication based on DHCP Signature	Y	Y	Y
RADIUS Assigned VLANs for NonEap MACs	Y	Y	Y
NEAP & Guest VLAN on same port (VOIP VLANs)	Y	Y	Y
EAP/NEAP with Fail-Open VLAN (FOV)	Y	Y	Y
EAP/NEAP Last Assigned VLAN	Y	Y	Y
EAP/NEAP MultiVLAN Capabilities	N	Y	Y
EAP Accounting	Y	Y	Y
RFC 3576 RADIUS Request Server	Y	Y	Y
RADIUS Interim Accounting Updates	Y	Y	N
EAP/NEAP with VLAN names	Y	Y	Y
NEAP Re-authentication	Y	Y	N
RADIUS EAP NEAP different servers	Y	Y	N
NEAP RADIUS Accounting	Y	Y	N
RADIUS Management Accounting	N	Y	N
Radius Request use Management IP	Y	Y	Y
RADIUS Timeout Mode	N	Y	N
RADIUS Management Accounting Tacacs Support	N	Y	N

Table 2.22: Supported Authentication Features

2.4 Troubleshooting and Monitoring

Understanding what is happening during the normal course of operations and knowing what to look for during abnormal times can help to maintain connectivity or restore operations quickly. This section highlights a few critical and often used troubleshooting tools. For details on all the options available, refer to the Troubleshooting documentation for each ERS product.

2.4.1 Packet Capture (PCAP)

The ERS 8800 supports a Packet Capture Tool (PCAP) tool that captures ingress and egress packets on selected I/O ports. With this feature, you can capture, save, and download one or more traffic flows through the Ethernet Routing Switch 8800/8600. The captured packets can then be analyzed offline for troubleshooting purposes. This feature is based on the mirroring capabilities of the I/O ports. To use PCAP, you must have the Advanced Software License. All captured packets are stored in the Secondary CPU, used as the PCAP engine. The Master CPU maintains its protocol handling and is not affected by any capture activity.

2.4.2 Port Mirroring

The Ethernet Routing Switches offer a port mirroring feature that helps you monitor and analyze network traffic. Port mirroring supports both ingress (incoming traffic) and egress (outgoing traffic) port mirroring. When you enable port mirroring, ingress or egress packets are forwarded normally from the mirrored (source) port, and a copy of the packets is sent to the mirroring (destination) port. Port mirroring capabilities and scalability vary between the ERS platforms.

As a general rule, you can mirror different speed ports and different physical media

- ➢ 10/100 to Gig – no issues
- ➢ Gig to 10/100 – may not see all packets if exceeding 100 Mbps on the Gig link
- ➢ Copper to Fiber – no issues
- ➢ Fiber to Copper – no issues

Port mirroring capabilities of the ERS stackable portfolio:

Ethernet Switch	Port Mirroring	Mirrored Ports (Source)					Mirroring Ports (Dest)
		Total Number	Ingress	Egress	Both	MAC-based	
ERS 2500	1-to-1	1	Yes	Yes	Yes	No	1
ERS 4500	Many to 1	Many	Yes	Yes	Yes	Yes	1
ERS 5500	Many to 1	Many	Yes	Yes	Yes	Yes	1
ERS 5600	Many to Many	Many	Yes	Yes	Yes	Yes	1

Many = all ports on a switch or stack

Many to Many applies to ERS 5600 hardware only and is implemented as four instances of Many to 1 mirroring. This functionality works only on ERS 5600 switches or pure stacks and will not work in a hybrid stack of ERS 5600s and ERS 5500s.

Table 2.23: Stackables Port Mirroring Capabilities

Port mirroring capabilities of the ERS modular portfolio:

Ethernet Switch	Port Mirroring	Mirrored Ports (Source)					Mirroring Ports (Dest)
		Total Number	Ingress	Egress	Both	MAC-based	
ERS 8300	Many to 1	Many	Yes	Yes	Yes	No	1
ERS 8600 Legacy Modules	Many to 1	Many	Yes	Yes	Yes	Yes	64
ERS 8800/8600 R Modules (4.0/4.1)	Many to 1	1 per LANE	Yes	Yes	Yes	Yes	Many
ERS 8800/8600 R Modules (5.0 and later)	1 to 1	1	1	1	Yes	Yes	1
	1 to Many	1	1	1	Yes	Yes	Many
	1 to MLT	1	1	Not Supported	Yes	Yes	MLT
	Many to 1	See Ingress/ Egress	20 ports per LANE	1 port per LANE	Yes	Yes	1
	Many to Many	See Ingress/ Egress	20 ports per LANE	1 port per LANE	Yes	Yes	Many
	Many to MLT	See Ingress/ Egress	20 ports per LANE	Not Supported	Yes	Yes	MLT
ERS 8800/8600 RS Modules (5.0 and later)	1 to 1	1	1	1	Yes	Yes	1
	1 to Many	1	1	1	Yes	Yes	Many
	1 to MLT	1	1	1	Yes	Yes	MLT
	Many to 1	See Ingress/ Egress	20 ports per LANE	20 ports per LANE	Yes	Yes	1
	Many to Many	See Ingress/ Egress	20 ports per LANE	20 ports per LANE	Yes	Yes	Many
	Many to MLT	See Ingress/ Egress	20 ports per LANE	1	Yes	Yes	MLT

ERS 8800/8600

Legacy modules, all ports in the same octapid must be mirrored to same destination

1 to Many = mirroring of one port to all ports in a destination VLAN

1 to MLT = mirroring of one port to all ports in a destination MLT

Many to Many = mirroring of many ports to all ports in a destination VLAN

Many to MLT = mirroring of many ports to all ports in a destination MLT

Table 2.24: Modular Port Mirroring Capabilities

Table 2.25 and Table 2.26 show the assignments of ports within an ERS 8800/8600 switch to the Octapids (for E/M modules) and Lanes (for R and RS modules). Note that ports belonging to the same Octapid group must be mirrored to the same destination port.

ERS 8800 Modules	Ports/Octapid	Port Assignments/Octapid (8 Octapids/Module)
8608 (GBE, GTE, SXE) 8608 (GBM, GTM)	1	1, 2, 3, 4, 5, 6, 7, 8
8616 (SXE, GTE)	2	1-2, 3-4, 5-6, 7-8, 9-10, 11-12, 13-14, 15-16
8624FXE	8	1-8, 9-16, 17-24
8632 (TXE, TXM)	8 per copper 1 per GBIC	1-8, 9-16, 17-24, 25-32, 33 (GBIC), 34 (GBIC)
8648 (TXE, TXM)	8	1-8, 9-16, 17-24, 25-32, 33-40, 41-48
8672 (ATME, ATMM)	4 with OC3 2 with DS3 1 with OC12	1-4, 5-8 with OC3 1-2, 3-4 with DS3 1,2 with OC12
8683POSM	2 with OC3 1 with OC12	1-2, 3-4, 5-6 with OC3 1, 2, 3 with OC12
8681XLR / 8681XLW	1 port uses all 8 Octapids	1

Table 2.25: ERS 8800 I/O Module Port to Octapid Mapping

ERS 8800 Modules	Ports/Lane	Port Assignments/Lane (3 Lanes/Module)
8630GBR	10	1-10, 11-20, 21-30
8648GTR	24	1-24, 25-48 (only uses 2 Lanes)
8683XLR / 8683XZR	1	1, 2, 3
8648GBRS	16	1-16, 17-32, 33-48
8648GTRS	24	1-24, 25-48 (only uses 2 Lanes)
8634XGRS	16, 16, 2	1-16, 17-32, 33-34
8612XLRS	4	1-4, 5-8, 9-12

Table 2.26: ERS 8800 I/O Module Port to Lane Mapping

2.4.3 Remote Logging

All of the ERS platforms support a remote logging feature. This provides an enhanced level of logging by replicating system messages on a syslog server. System log messages from several switches can be collected at a central location, alleviating the network manager from querying each switch individually to interrogate the log files. It also ensures that information is not lost when a switch becomes inoperable. The level of logging and the details provided differ between ERS platforms – refer to the System Monitoring or Troubleshooting documentation for each of the products to obtain more details.

2.4.4 Stackables Tools

The ERS stackables have numerous built-in tools that offer information on the health and well-being of the stack:

Stack Health Check

The stack health check feature provides a view into the overall operation of the stack; with information on the cascade connections, if the stack is resilient (return cable connected) or non-resilient, number of units in the stack and the model of each unit along with its unit number. This feature shows you a quick snapshot of the stack configuration and operation.

Stack Monitor

You use the Stack Monitor feature to analyze the health of a stack by monitoring the number of active units in the stack. With the Stack Monitor feature, when a stack is broken, the stack and any disconnected units from the stack send SNMP traps. If the stack or the disconnected units are still connected to the network, they generate log events and send trap messages to notify the administrator of the event. After the problem is detected, the stack and disconnected units continue to generate log events and send traps at a user-configurable interval until the situation is remedied (or the feature is disabled).

Stack Loopback Test

The stack loopback test feature allows the customer to quickly test the switch stack ports and the stack cables on the ERS units. This feature helps you while experiencing stack problems to determine whether the root cause is a bad stack cable or a damaged stack port and prevents potentially good switches being returned for service. You can achieve this by using two types of loopback tests – internal and external.

Stack Port Counters

The stack port counters show statistics of the traffic traversing the stacking connectors, including the size of packets, FCS errors, filtered frames, etc.

Environmental Information

This feature displays environmental information about the operation of the switch or units within a stack. The show environmental command does not require any configuration, and it reports the power supply status, fan status and switch system temperature.

CPU and Memory Utilization

The CPU utilization feature provides data for CPU and memory utilization. You can view CPU utilization information for the past 10 seconds, 1 minute, 1 hour, 24 hr, or since system boot up. The switch displays CPU utilization as a percentage. You can use CPU utilization information to see how the CPU is used during a specific time interval. The memory utilization

provides you information on what percentage of the dynamic memory is currently used by the system. The switch displays memory utilization in terms of megabytes available since system boot up. This feature does not require a configuration. It is a display-only feature.

2.4.5 Key Health Indicator Collection

The ERS 8800/8600 with release 7.0 support Key Health Indicators (KHI) that allow for the collection of statistics and information about the health of the system for troubleshooting purposes related to system failure. The Key Health Indicator (KHI) feature identifies a small number of key health indicators that allow quick assessment of the overall operational state of the ERS. These indicators do not provide complete coverage of all possible failure scenarios. Rather, KHI is a diagnostic tool for the health of the switch. Further debugging is required to correctly understand the system state and actions required to remedy the situation.

KHI provides global health information for the switch, including:

> Chassis health indication

> CPU performance health indication

> Port state change indication

> Forwarding health indication

> IP interface configuration and operation information

> Protocol information

> Management information: Log, TCP, UDP and Users

The switch stores the information locally and displays the information as requested by the user using show commands.

KHI supports multiple KHI types that track specific switch areas or subsystems. Each KHI type keeps track of the last ten events for the specific subsystem (for example, protocol going down or loss of connection) in a rolling history. KHI creates a reference point using a time stamp, and then tracks events from that point forward. Clear commands are provided to reestablish fresh timelines.

Generally, the KHI information allows you to track the source of a problem to a particular subsystem. Once this determination is made, you can use specific statistics for that subsystem (for example, OSPF-specific statistics and show commands) to further locate the source of the issue.

To configure KHI, you can enable or disable the feature globally. In addition, you can enable or disable some of the KHI types individually.

This additional control is provided for KHI types that have a greater impact on loaded systems.

The main configuration actions for KHI are:

> Enabling or disabling KHI (at global or feature-level)

> Displaying statistics

> Clearing statistics/history to establish a new timeline

Currently, EDM does not support KHI configuration.

2.5 Security Features

The security of the Converged Campus is of paramount importance to the overall design. Avaya offers a variety of security mechanisms, both within the Ethernet switching platforms and externally, that work in conjunction to provide the highest level of security possible. This section focuses on features built into the Ethernet Routing Switches that guard against attacks on the network.

 ➢ Broadcast and Multicast Rate Limiting

 To protect the switch and other stations from a high number of broadcasts and multicasts, the switch has the ability to limit the broadcast/multicast rate. This feature can be configured on a per-port basis. By default, this feature is disabled, and should only be enabled in accordance with the recommendations from the previous section.

 ➢ Directed broadcast suppression

 The Ethernet Routing Switches provides the ability to enable or disable forwarding for directed broadcast traffic on an IP-interface basis. A directed broadcast is a frame sent to the subnet broadcast address on a remote IP subnet. When you disable or suppress directed broadcasts on an interface, all frames that are sent to the subnet broadcast address for a local router interface are dropped. Directed broadcast suppression protects hosts from possible Denial of Service (DoS) attacks.

 ➢ Prioritization of control traffic

 The Ethernet Routing Switches use a very sophisticated prioritization scheme for scheduling received control packets (BPDUs, OSPF Hellos, etc.) on physical ports. This scheme involves two levels with both hardware and software queues and guarantees proper handling of these control packets regardless of the load on the switch. In turn, this guarantees the stability of the network. More specifically, it guarantees that the applications that heavily use broadcasts are handled with a lower priority. Note that you cannot use the CLI to view, configure, or modify these queues. Setting the queues and determining the type of packets entering each queue is Avaya confidential.

 ➢ ARP limitation

 The ARP request threshold limits the ability of the ERS 8800/8600 to source ARP requests for workstation IP addresses it has not learned within its ARP table. The default setting for this function is 500 ARP requests per second. For networks experiencing excessive amounts of subnet scanning caused by a virus, Avaya recommends changing the ARP request threshold to a value between 100 and 50. This will help protect the CPU from causing excessive ARP requests, help protect the network, and lessen the spread of the virus to other PCs.

 ➢ Multicast Learning Limitation

 This feature protects the CPU from multicast data packet bursts generated by malicious applications such as viruses. Specifically, it protects against those viruses that cause the CPU to reach 100 percent utilization or that prevent the CPU from processing any protocol packets or management requests. If more than a certain number of multicast streams enter the CPU through a port during a sampling interval, the port is shut down until the user or administrator takes appropriate action.

➢ High Secure Mode

To protect the ERS 8800/8600 against IP packets with an illegal source address of 255.255.255.255 from being routed (per RFC 1812 Section 4.2.2.11 and RFC 971 Section 3.2), the ERS 8800 supports a configurable flag, called *high secure*. By default, this flag is disabled. Note that when you enable this flag, the desired behavior (not routing source packets with an IP address of 255.255.255.255) is applied.

➢ Access Policies

Access policies let you control management access by setting policies for services to prevent or allow access to the switch. You can specify which hosts or networks can access the switch through FTP, http, rlogin/rsh, SSH, Telnet, and TFTP. You can set the access level (ro|rw|rwa).

➢ Configurable Software Daemons

The ERS 8300 and ERS 8800/8600 provide the ability to enable or disable various access methods. On the ERS 8800/8600, you enable or disable ftp, tftp, telnet, rlogin, SSH, or SNMP. The ERS 8300 allows you to enable or disable ftp, tftp, telnet, rlogin, or SSH. The ERS 8800/8600 also has a High Secure Mode in which all daemons are disabled; i.e., telnet, ftp, tftp, rlogin, and SNMP are disabled.

For the ERS 5000, ERS 4500, ERS 2500, HTTP, telnet, and SNMP can be enabled or disabled through standard configuration.

➢ Route Policies

The ERS 8800/8600, ERS 8300, and ERS 5000 support IP RIP/OSPF accept/announce policies. This provides extra security by either blocking specific subnets or selecting to announce specific subnets.

➢ Port Lock Feature

This feature lets you lock a port or prevent other users from changing port parameters or modifying port action. Locked ports cannot be modified in any way until the port is unlocked.

➢ Out-of-band Management

Each Switch Fabric on the ERS 8300 and ERS 8800/8600 provides an Out-of-band Management port. Traffic on this port is completely separated from the user traffic and provides a high secure network for management.

➢ Access Security

The following features are supported for access to ERS switch:

- RADIUS authentication
- TACACS+
- SSH
- SSL
- Password security and CLI logging

➢ Stopping IP Spoofed Packets

Spoofed IP packets are stopped by configuring the ERS 8800/8600 to ensure that IP packets are forwarded only if they contain the correct source IP address of your network. A spoofed packet is one that comes from the Internet into your network with a source address equal to one of the subnet addresses used on your network. Its source address belongs to one of the address blocks or subnets used on your network. With anti-spoofing protection, you have a filter rule/configuration assigned to the external interface, which examines the source address of all outside packets crossing that interface. If that address belongs to internal network or firewall itself, the packet is dropped. The correct source IP addresses consist of the IP network addresses that have been assigned to the site/domain. It is particularly important that you do this throughout your network, especially at the external connections to the existing Internet/upstream provider. By denying all invalid source IP addresses, you minimize the chances that your network will be the source of a spoofed DoS attack.

This will not stop DoS attacks coming from your network with valid source addresses, however. In order to prevent this, you need to know which IP network blocks are in use. You then create a generic filter that:

- Permits your sites' valid source addresses
- Denies all other source addresses

➢ Reverse Path Checking

The reverse path checking feature (available only on ERS 8800/8600 R and RS modules), when enabled, prevents packet forwarding for incoming IP packets that have incorrect or forged (spoofed) IP addresses. Reverse path checking guarantees that traffic received on one interface was sent by a station from this interface, which prevents address spoofing. With this mode enabled, the Ethernet Routing Switch 8600 performs a reverse path check to determine if the packet IP address is verifiable. If the address is not verifiable, the packet is discarded.

You configure reverse path checking on a per-IP interface basis. When reverse path checking is enabled, the Ethernet Routing Switch 8800/8600 checks all routing packets which come through that interface. It ensures that the source address and source interface appear in the routing table, and that it matches the interface on which the packet was received.

You can use one of two modes for reverse path checking:

- Exist-only mode: In this mode, reverse path checking checks whether the incoming packet's source IP address exists in the routing table. If the source IP entry is found, the packet is forwarded as usual; otherwise, the packet is discarded.

- Strict mode: In this mode, reverse path checking checks whether the incoming packet's source IP address exists in the routing table. If the source IP entry is not found, reverse path checking further checks if the source IP interface matches the packet's incoming interface. If they match, the packet is forwarded as usual, otherwise, the packet is discarded.

➢ DHCP Snooping

Dynamic Host Configuration Protocol (DHCP) snooping provides security to the network by preventing DHCP spoofing. DHCP spoofing refers to an attacker's ability to respond to DHCP requests with false IP information. DHCP snooping acts like a firewall between untrusted hosts and the DHCP servers, so that DHCP spoofing cannot occur. This feature classifies ports into trusted (where the DHCP server exists – which could simply be the uplink to the core) and untrusted (end user ports). DHCP requests are only forwarded to and thru the trusted ports. Any DHCP replies from untrusted ports are automatically dropped. A binding table is also created with the source MAC, IP address, VLAN, port, and lease time.

➢ Dynamic ARP Inspection

Dynamic Address Resolution Protocol (Dynamic ARP) inspection is a security feature that validates ARP packets in the network. Without dynamic ARP inspection, a malicious user can attack hosts, switches, and routers connected to the Layer 2 network by poisoning the ARP caches of systems connected to the subnet and by intercepting traffic intended for other hosts on the subnet. Dynamic ARP inspection prevents this type of *man-in-the-middle* attack. It intercepts, logs, and discards ARP packets with invalid IP-to-MAC address bindings. The address binding table is dynamically built from information gathered in the DHCP request and reply when DHCP snooping is enabled. The MAC address from the DHCP request is paired with the IP address from the DHCP reply to create an entry in the DHCP binding table.

When Dynamic ARP inspection is enabled, IP traffic on *untrusted* ports is filtered based on the source MAC and IP addresses seen on the switch port. The switch forwards IP traffic when the source MAC and IP address matches an entry in the address binding table. Otherwise, the IP traffic is dropped. For dynamic ARP inspection to function, DHCP snooping must be globally enabled.

➢ IP Source Guard

IP Source Guard provides security to the network by filtering clients with invalid IP addresses. IP Source Guard is a Layer 2 (L2), port-to-port basis feature that works closely with information in the Dynamic Host Control Protocol (DHCP) snooping binding table. When you enable IP Source Guard on an untrusted port with DHCP snooping enabled, an IP filter entry is created or deleted for that port automatically, based on IP information stored in the corresponding DHCP snooping binding table entry. When a connecting client receives a valid IP address from the DHCP server, a filter is installed on the port to allow traffic only from the assigned IP address. A maximum of 10 IP addresses are allowed on each IP Source Guard-enabled port. When this number is reached, no more filters are set up and traffic is dropped.
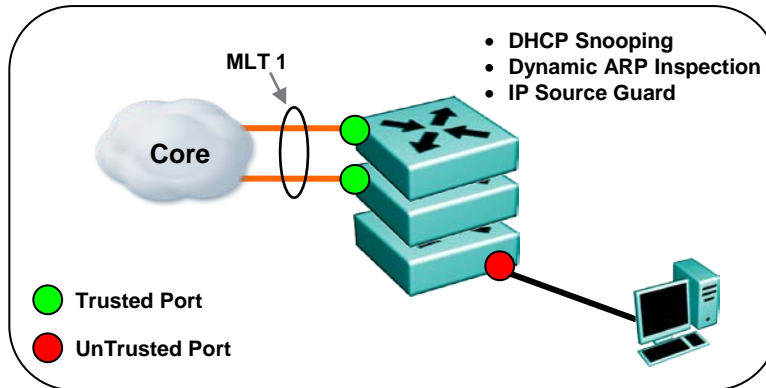
Figure 2.50: Edge Switch Security

➢ Untrusted ports are normally the end user access ports on the switch

➢ Trusted ports are normally the uplinks from the edge switch to the core

➢ Dynamic ARP Inspection and IP Source Guard use the binding table created by DHCP Snooping, therefore, in order to use these features, DHCP Snooping must be enabled.

➢ IP Source Guard should not be enabled on uplink ports from the Edge to the Core, only enable on Edge access ports (untrusted ports) where DHCP Snooping and Dynamic ARP Inspection are enabled.

2.6 Network Management

For the Large Campus, a full suite of management tools are typically requested. There is usually a large number of devices that need to be managed. The key is to simplify the management process while still supporting a large number of devices and potentially complex network designs.

2.6.1 Access Security

Ensuring secure access to the infrastructure is extremely important for maintaining the integrity of the Large Campus network. There are a few basic guidelines as part of the Avaya best practices as highlighted here.

- ➢ SSH (secure shell) provides encrypted communication to the Ethernet Switches, this is preferable over Telnet which is clear text

- ➢ A web interface is available on the ERS products – if this interface is used it is strongly recommended to implement SSL and/or change the port number to which the switch will be managed via the web

- ➢ Create access lists to ensure only certain clients or subnets are able to manage the switches

- ➢ Implement RADIUS for authorization and accounting

    - ▪ RADIUS fallback ensures access to switches in the event the RADIUS server is not reachable
    - ▪ Change all local passwords from factory default

- ➢ TACACS+ is also supported for authorization and accounting

- ➢ Configure remote syslog capabiliites to ensure all critical log file information is kept and recorded off the switch(es).

- ➢ Configure SNMPv3 for secure SNMP communications to all Ethernet Switches

    - ▪ Make sure community strings are changed from default

2.6.2 Network Management

Avaya provides a comprehensive set of solutions and tools to enable a system-wide life cycle management. The objective is to offer management solutions that are efficient, survivable, consistent, and drive simplicity. This is accomplished by addressing all areas of FCAPS.

FCAPS is the ISO Telecommunications Management Network model and framework for network management. FCAPS is an acronym for *Fault, Configuration, Accounting, Performance* and *Security* which are the management categories into which the ISO model defines network management tasks.

The goal of fault management is to recognize, isolate, correct and log faults that occur in the network. It uses trend analysis to predict errors so that the network is always available. This is established by monitoring different aspects of the network for abnormal behavior.

Configuration Management covers configuration changes to the network, including adding new devices, removing old ones and modifying the configuration of others. It ensures that changes are conducted consistently in an ordered manner no matter how many changes are required.

Accounting Management gathers usage statistics, and by using the statistics, the users can be billed and usage quota enforced. For non-billed networks, "administration" replaces "accounting". The goals of administration are to administer the set of authorized users by establishing passwords, permissions, and to administer network operations.

Performance Management addresses the throughput, percentage utilization, error rates and response times of the network. By collecting and analyzing performance data, the network health is monitored. Trends indicate capacity or reliability issues before they become service affecting. Performance thresholds can trigger alarms to proactively correct potential bandwidth issues.

Security Management is the process of controlling access to assets in the network. Data security can be achieved mainly with authentication i.e. verifying the identity of the person accessing the network, authorization (that is determining what the user can than do) and finally encryption (that is protecting traffic from unauthorized interception when it passes across the network).
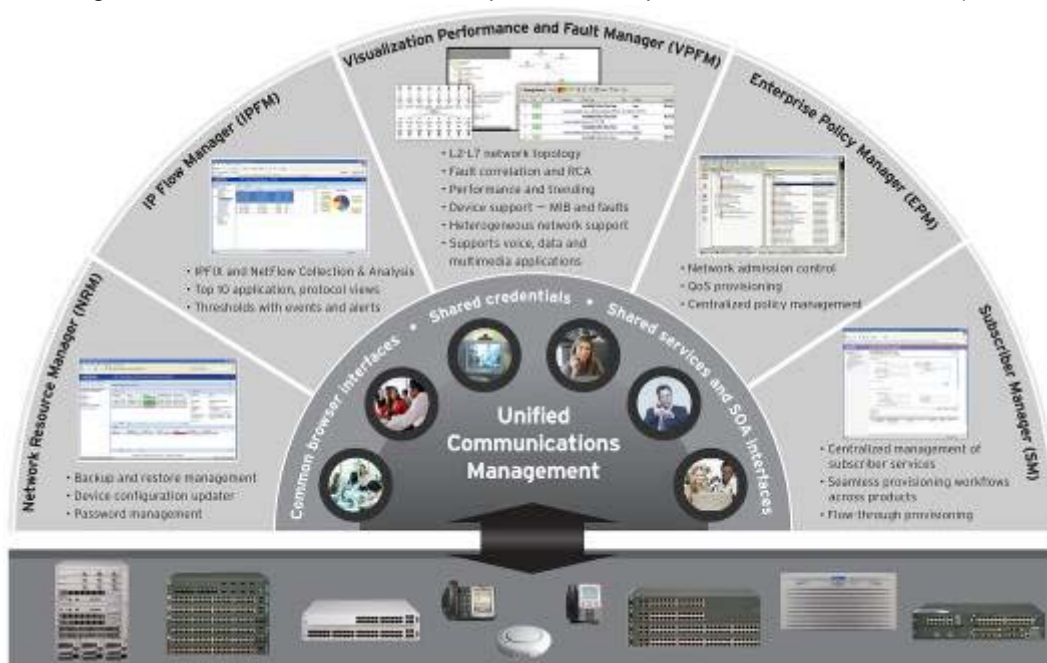
Figure 2.51: Unified Communications Management

2.6.3 Unified Communications Management (UCM)

In essence, Avaya Unified Communications Management is the 'glue' that binds together distributed network management components to establish a true unified communications environment. UCM, which follows the FCAPS framework, provides comprehensive and simplified network management across voice, data and multimedia applications through integrating management features and capabilities including visualization, configuration, fault and performance management, subscriber services and security.

UCM is a centralized and integrated set of management tools and applications. The key and distinguishing element of UCM is the in-built Common Services that are an integral part of each of the different network management applications that form UCM. Common Services allow network management applications to integrate with each other so that common components (for example, user data, database information and certificate management) can be shared, without requiring that the same definitions and configurations be repeated for each application. The result is a fully

integrated, single point of contact, providing a unified view of the network, while streamlining workflows and the management environment, and reducing installation and configuration time, system operations and maintenance.

Each product offering in the Avaya UCM solution has two components; first, the application which provides the application-specific functions by purpose. These components include:

- **Visualization Performance and Fault Manager** (data & voice)
    - Multi-vendor network discovery, root cause analysis, network topology maps
- **Configuration and Orchestration Manager** (data)
    - Multi-element configuration for Ethernet Routing Switches
- **Enterprise Policy Manager** (data)
    - Network access control policies, bandwidth management, QoS
- **IP Flow Manager** (data)
    - IPFIX collection, analysis and reporting
- **Subscriber Manager** (voice)
    - Centralized management of telephony subscriber services

Second is Common Services (UCM-CS). All applications in the UCM solution have the same UCM-CS which offer a common set of tools, interfaces, service subscriptions, user authentication, and user definition / management. Integration between components occurs at the UCM-CS layer.

The UCM key features include:

- Application co-residency – lower CAPEX/OPEX
- Single unified management domain – decreased complexity
- Integrated workflow – reduced errors
- Centralized authentication and navigation – improved user experience
- Shared data – reduced errors, secure
- Centralized management via browser – simple to use
- Single Sign-On – simple to use, secure
- Simplified system admin configuration – simple to use
- Flexible XML Architecture – investment protection

2.6.4 Visualization Performance & Fault Management

Visualization Performance & Fault Manager (VPFM) – part of the Unified Communications Management (UCM) solution – provides a heterogeneous network discovery and visualization tool that uses standards-based and proprietary protocols to provide an end-to-end view of the network. The functionality available with VPFM includes:

> Network Device and End-node Discovery

- Proprietary-based (Avaya SONMP, Cisco CDP)
- Standards-based (IP, SNMP, IEEE 802.1AB, ad etc.)
- Application Discovery
- Servers, PC, etc.

> L2 and L3 Network Topology Visualization

- Device relationships and their interconnectivity
- Switch views
- Campus Views
- VoIP Service (Application) Views

> Fault Management

- Event Correlation and Analysis
- Event Handling and Scripting
- Trap Retention and Exporting
- Custom Propagation Rules

> Device Performance Monitoring

- LAG Performance Monitoring
- Threshold support
- Trending, Graphing, and data Exporting
- Custom Device Classification (scope) Management
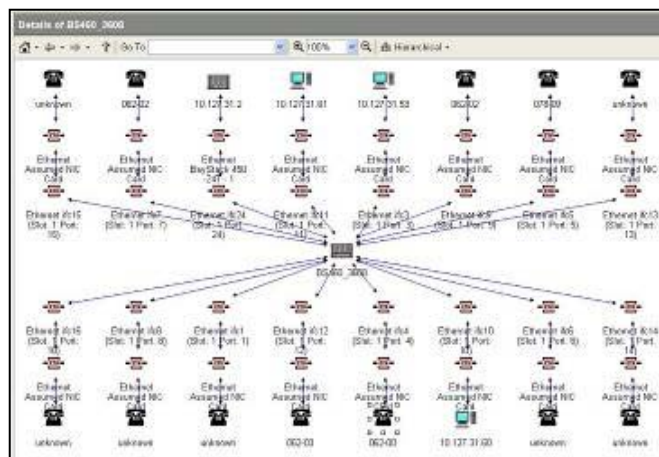
> Enhanced Diagnostics

- L2 and L3 Diagnostic management

Figure 2.52: VPFM Topology View

2.6.5 Configuration & Orchestration Manager

Configuration & Orchestration Manager (COM) delivers centralized configuration, provisioning and troubleshooting for enterprise-wide devices and technologies. It is a real-time Web-based application, enabling anytime, anywhere access for multiple authorized users. With role-based access privileges and audit trails, a high level of security is ensured.

COM simplifies network-wide configuration by offering a topology-driven GUI with workflow-based visibility. Complex configurations such as VLAN, SMLT/MLT, VRF and multicast are easily managed using templates and wizards.

In addition, COM can provide network discovery, password management and serve as a platform for downloading, launching and hosting off-box device management. With centralized off-box element management, COM also streamlines the managing of inventory, updates as well as device and user access controls.
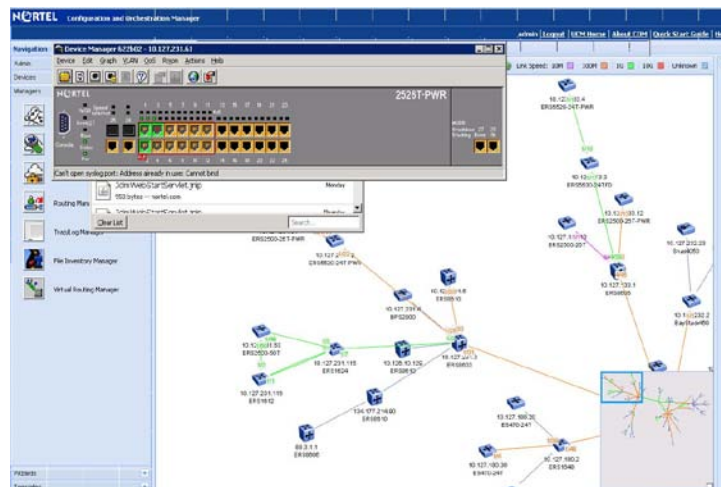
Figure 2.53: COM Topology View

COM also provides a launch point for Enterprise Device Manager (EDM) and off-box EDM functionality. EDM is the on-box management platform for all Ethernet switching products. EDM replaces Java Device Manager (JDM) and delivers a more streamlined approach to single element management. EDM uses a web-based interface to simplify configuration of the network element.

EDM comes as part of the software image for the Ethernet Routing Switches at no additional cost. The look and feel of EDM is very similar to that of legacy JDM, so transitioning to the new management platform should be nearly seamless. Because EDM is an on-box configuration tool, there are certain features that cannot be done on-box (syslog, trap receiver, audit logs, mib browser, etc.) and therefore requires an off-box application – COM provides this off-box application through EDM plug-ins. With COM and EDM, the network administrator has a full suite of tools required to quickly and easily perform configuration and orchestration of the network infrastructure.

2.6.6 Enterprise Policy Manager

Enterprise Policy Manager (EPM) is a network-level application that allows administrators to manage network bandwidth, prioritize traffic streams, and set network access policies. EPM enables critical applications to receive the proper QoS by deploying common policies from one central management application. This ensures consistency across the network along with simplifying the deployment of an enterprise-wide QoS policy.

EPM can be used to define Policy Enforcement Points in the network. An operator will associate multiple ports throughout the network across many switches to a Role within EPM. A user can then quickly define a Policy (consisting of a Traffic Condition, Action, and Schedule) into a role which will once applied immediately apply that policy to all of the switches/ports in that Role.
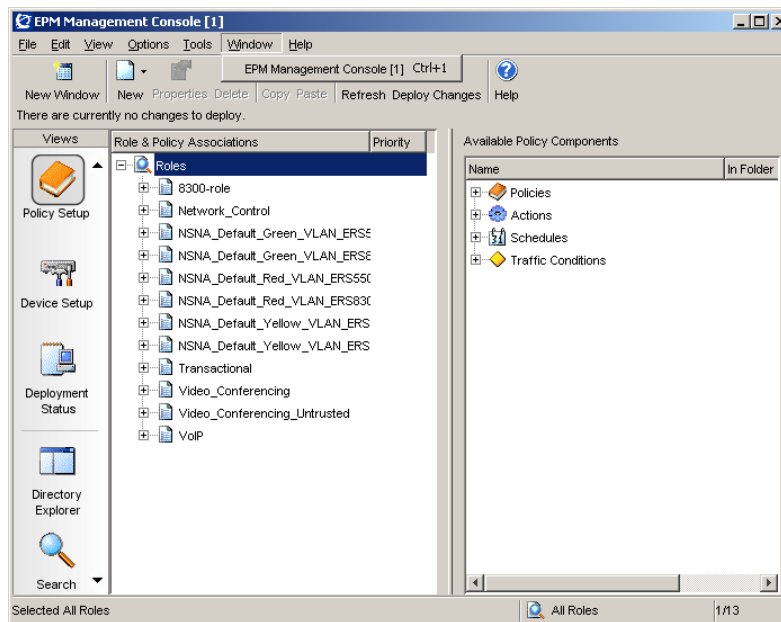
Figure 2.54: Enterprise Policy Manager

2.6.7 IP Flow Manager

IP Flow Manager (IPFM) is the external IPFIX collector used to analyze the various services and applications on the network. IPFIX is a very powerful tool for collecting application flows across the network. This enables the network administrator to gain insight in regard to usage monitoring of the network, user activity, network abuse, top users and top protocols, and can help answer the questions about true end-to-end application performance.

An IP flow is defined as a set of packets sent over a period of time that have some common properties. These properties include:

> Source IP address

> Destination IP address

> Protocol type

> Source protocol port

> Destination protocol port

> Ingress VLAN ID

> Ingress port and observation point (VLAN or port)

IPFM turns IPFIX data into useful and easily understood information. This valuable information now increases the visibility into whom and which applications are consuming network resources and bandwidth. It also is used as a capacity planning tool, allowing the network administrator to proactively make changes to avert any potential bandwidth issues that may arise with new applications being rolled out continually.
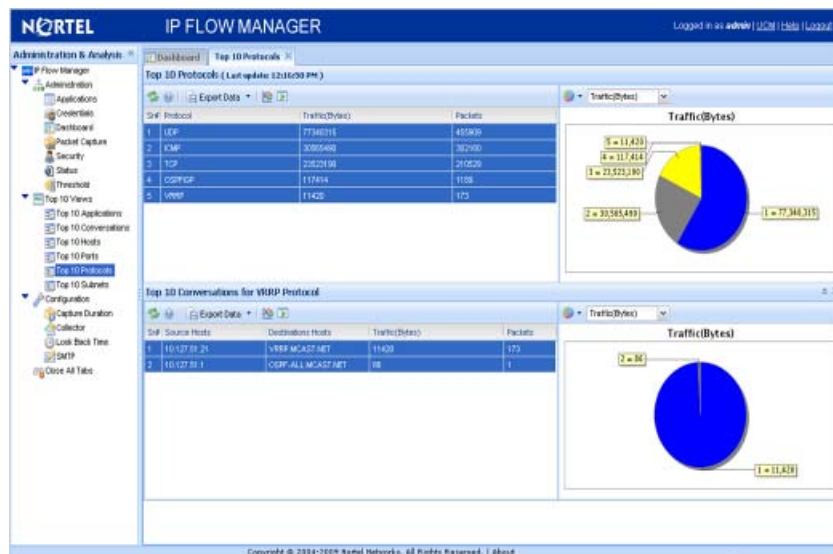
Figure 2.55: IP Flow Manager

2.6.8 Proactive Voice Quality Management

Business challenge – Unified communications can give your business a competitive advantage by empowering the end user with productivity-improving applications, but deploying and maintaining Voice over IP (VoIP) can be challenging. The problem facing many customers is summarized in the benchmark survey Benchmarking VoIP Performance Management by the Aberdeen Group (March 2008). "Since transitioning from old-fashioned dedicated analog telephony to VoIP is a major initiative at many organizations, managing performance of VoIP services is becoming increasingly important. Managing the performance of voice services delivered over IP networks is possible through execution of a variety of strategies; but the challenge that organizations face is in identifying the right strategy and developing capabilities needed for successful execution of that strategy." How do you know that your network is ready to handle the real-time bandwidth required? How can you monitor the end-user experience versus the network performance? How will you troubleshoot problems in real-time as they occur and reduce the number of support requests?

The Solution – PVQM (Proactive Voice Quality Management). Co-developed by Avaya and NetIQ, Proactive Voice Quality Management ensures that you have the right tools in place to support the right processes to stay on top of your VoIP service quality. Count on Avaya and NetIQ to provide your VoIP Service Level Management needs.

What is Proactive Voice Quality Management? Proactive Voice Quality Management offers a life-cycle approach that provides the necessary management tools to support each phase of an IP Telephony project: assessment, pre-deployment, ongoing monitoring and reporting. PVQM then provides a set of interrelated technologies that enable an effective Service Level Management solution for IP Telephony. PVQM focuses on the end-user Quality of Experience (QoE) using standards-based technologies ensuring you can stay on top of your service quality needs.
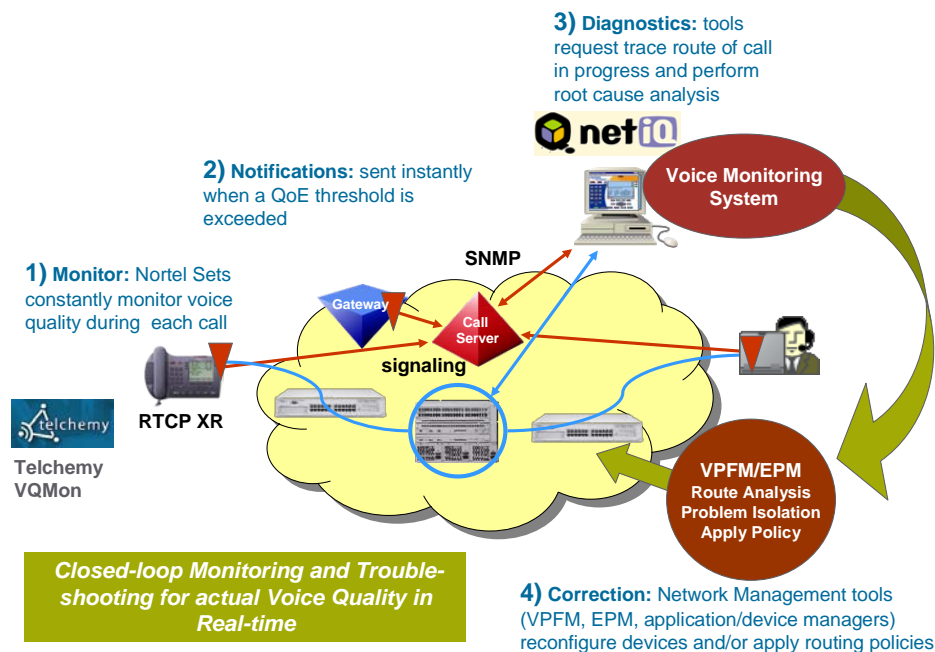

Figure 2.56: PVQM

Network assessment and pre-deployment – Is your network ready for VoIP? According to a Gartner Group study, 85 percent of locations will require upgrades to their data networks to properly run voice. Before deploying VoIP, you need to have the proper equipment and identify if you need network upgrades before deployment. Fully supporting Avaya Network Health Checks, NetIQ's Vivinet Assessor determines quickly and easily how well VoIP will work on the network prior to deployment. Before you invest in costly training and pilot deployments, Vivinet Assessor predicts the overall call quality that you can expect from the network and generates polished, customizable reports detailing the network's VoIP readiness. Also in this phase, Enterprise Policy Manager provides centralized management of your network QoS through an intuitive graphical user interface. Deploy QoS in a consistent manner across your enterprise network using Enterprise Policy Manager.

Monitoring and reporting - Managing VoIP availability and quality needs – Once VoIP is deployed, you must monitor the VoIP environment to maximize uptime, reliability and quality. NetIQ's AppManager Suite is a robust platform and suite of modules that provide comprehensive monitoring, management and reporting for voice solutions. NetIQ AppManager for Avaya extends the suite to optimize the health and availability of Avaya's VoIP platforms. Call quality is monitored from an end-user perspective and in real time with embedded software from Telchemy. Telchemy's VQmon is integrated into Avaya's IP phones to provide the metrics needed to support management and QoS reporting protocols and facilitate problem diagnosis. NetIQ AppManager hen combines platform, system and call quality metrics for Avaya Call Servers with information about the availability and health of network devices and overall network performance for VoIP. Avaya's Visualization Performance Faults Manager can also be deployed providing integrated infrastructure management for your data network, data services and IP Telephony. With Avaya Communication Server 1000 (CS 1000), VoIP quality is monitored from an end-user perspective providing a high degree of correlation between how network performance affects actual service quality. End-user QoE is monitored and reported in real-time using industry-standard technology (ITU G.107 and IETF 3611) providing the first open solution on the market. This allows operators to focus on real problems in the network. Furthermore, critical insight is provided to highlight the underlying conditions that led to service quality degradation. This allows operators to map VoIP service quality back to the underlying network infrastructure once and for all!

PVQM Summary – Successful unified communications implementation with VoIP requires integrated management solutions that allow you to take control of your entire voice network and server infrastructure. Understanding how data traffic will affect voice applications — before deployment — and then continually monitoring and diagnosing the status of IP Telephony devices will help maximize success. For all stages of deployment, NetIQ's products provide the most comprehensive solution available on the market. NetIQ delivers assessment, monitoring and diagnostic products to help you ensure a successful implementation and accelerate your overall return on VoIP investment. In addition, Enterprise Policy Manager and Visualization Performance Faults Manager complement your investment in Avaya data solutions, providing point-and-click management of your converged network.

3. Configuration Example

The following example will be used throughout the remainder of this document as the reference topology for configuration of all the best practice recommended parameters.

Configuration Design Details:

VLAN Name	VLAN ID	IP Subnet
IST - Core	3999	1.1.1.4/30
Core-SMLT-1	4	10.0.4.0/24
Core-SMLT -2	14	10.0.14.0/24
IST - Distribution	3999	1.1.1.8/30
Mgmt	2000	10.20.0.0/24
Data1	2101	10.21.1.0/24
Voice1	2102	10.21.2.0/24
Data2	2201	10.22.1.0/24
Voice2	2202	10.22.2.0/24

Core SMLT VLAN
➢ Each Square topology will have two SMLT VLANs configured with RSMLT and OSPF enabled to provide better receovery in the event of losing the OSPF DR

Management VLAN
➢ Configured as Layer 2 across the network

Layer 3 Edge Routing
In order to provide examples of both RSMLT Layer 2 Edge and VRRP, the following configurations will be used:
➢ Data VLANs are using RSMLT Layer 2 Edge
➢ Voice VLANs are using VRRP / Backup Master

Figure 3.1 shows the entire test topology that was used in the validation of the Large Campus design. The configuration examples that follow are from the top shaded portion of the diagram. The bottom portion's configuration would be identical to the top with the only difference being VLAN id's and IP addresses.
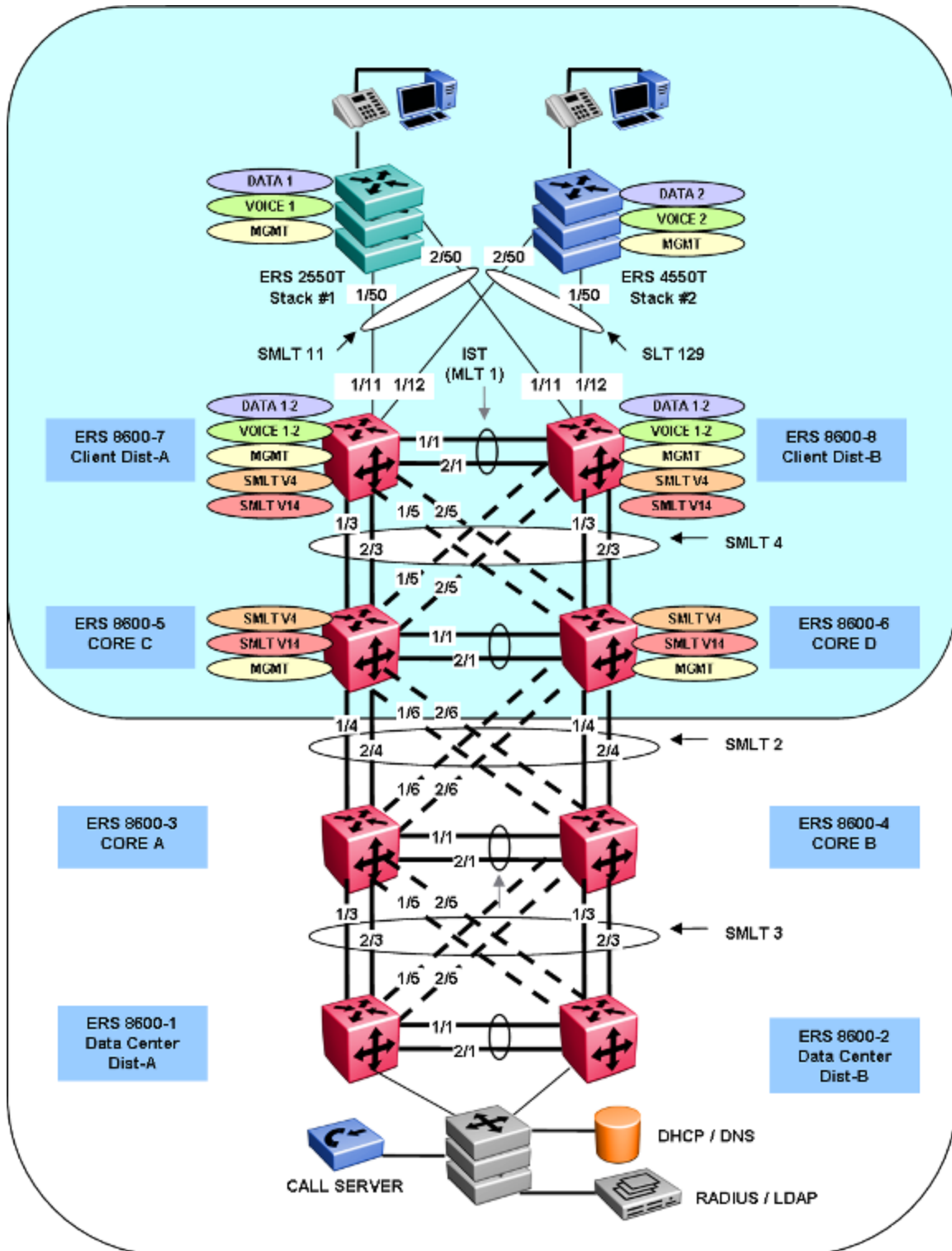
Figure 3.1: Configuration Topology

3.1 Software Versions & Upgrade Policy

The Large Campus Solution was validated with specific software versions as indicated below. These software releases provided stable interoperability of products at the time of release of Large Campus Solution. However, post release, individual products will undergo minor updates or major release upgrades. This section provides guidance on how to benefit from periodic software changes while still maintaining the overall integrity of the solution.

In general, upgrading of products to the next major release moves the solution outside the boundaries of validation undertaken for the particular solution. Major product upgrades will be validated in future updated versions of the Large Campus Solution Guide. If upgrades to the next major revisions are required, consult with Avaya and proceed with caution. Review the Release Notes carefully to ensure changes in the newly released software will not adversely affect the overall network.

In contrast to the above, you can update or patch products within the Large Campus Solution to the next minor software release with the worry of moving the solution outside the validation boundaries. This allows you to take advantage of bug fixes as they become available.

Software versions used for the Large Campus validation:

- ➢ ERS 8800/8600 Release 7.0
- ➢ ERS 8600 Release 5.1.0
- ➢ ERS 8300 Release 4.1.2.0
- ➢ ERS 5000 Release 6.0.3
- ➢ ERS 4500 Release 5.2.2
- ➢ ERS 2500 Release 4.2.2

Minor software releases for the above products would be as follows, with "x" being incremented as required for each maintenance release. When downloading software from the Avaya support site, each version of software is clearly identified by type (Major or Maintenance).

- ➢ ERS 8800/8600 Release 7.0
- ➢ ERS 8600 Release 5.1.x
- ➢ ERS 8300 Release 4.1.2.x
- ➢ ERS 5000 Release 6.0.x
- ➢ ERS 4500 Release 5.2.x
- ➢ ERS 2500 Release 4.2.x

Major software releases for the above products would be as follows. If upgrading to these releases, please consult Avaya and the Release Notes as these have not been validated with the LargeCampus Solution at this time.

- ➢ ERS 8300 Release 4.2
- ➢ ERS 5000 Release 6.1
- ➢ ERS 4500 Release 5.3
- ➢ ERS 2500 Release 4.3

3.2 Switch Cluster Core Configuration

This section covers the major configuration requirements of the Switch Cluster Core. Please note that other features are available and can be implemented, for the purpose of this document, the basics are covered and the rest is left to the product documentation.

3.2.1 Create VLANs in the Core

Create SMLT VLANs and IST VLAN on ERS8600-5 & ERS8600-6 Core Switches

```
Core-C:5# config vlan 4 create byport 1 name "CoreSMLT-1"
Core-C:5# config vlan 4 fdb-entry aging-time 21601
Core-C:5# config vlan 14 create byport 1 name "CoreSMLT-2"
Core-C:5# config vlan 14 fdb-entry aging-time 21601
Core-C:5# config vlan 2000 create byport 1 name "Mgmt"
Core-C:5# config vlan 3999 create byport 1 name "istVlan"

-------------------------------------------------------------------------------

Core-D:5# config vlan 4 create byport 1 name "CoreSMLT-1"
Core-D:5# config vlan 4 fdb-entry aging-time 21601
Core-D:5# config vlan 14 create byport 1 name "CoreSMLT-2"
Core-D:5# config vlan 14 fdb-entry aging-time 21601
Core-D:5# config vlan 2000 create byport 1 name "Mgmt"
Core-D:5# config vlan 3999 create byport 1 name "istVlan"
```

Add IP addresses to the SMLT and IST VLANs

```
Core-C:5# config vlan 4 ip create 10.0.4.1/255.255.255.0
Core-C:5# config vlan 14 ip create 10.0.14.1/255.255.255.0
Core-C:5# config vlan 2000 ip create 10.20.0.7/255.255.255.0
Core-C:5# config vlan 3999 ip create 1.1.1.5/255.255.255.252

-------------------------------------------------------------------------------

Core-D:5# config vlan 4 ip create 10.0.4.2/255.255.255.0
Core-D:5# config vlan 14 ip create 10.0.14.2/255.255.255.0
Core-D:5# config vlan 2000 ip create 10.20.0.8/255.255.255.0
Core-D:5# config vlan 3999 ip create 1.1.1.6/255.255.255.252
```

3.2.2 Create the Switch Cluster Core

Create MLT 1 for IST

```
Core-C:5# config mlt 1 create
Core-C:5# config mlt 1 name "istToCoreD"
Core-C:5# config mlt 1 add ports 1/1,2/1
Core-C:5# config mlt 1 perform-tagging enable
Core-C:5# config vlan 3999 add-mlt 1

-------------------------------------------------------------------------------

Core-D:5# config mlt 1 create
Core-D:5# config mlt 1 name "istToCoreC"
Core-D:5# config mlt 1 add ports 1/1,2/1
Core-D:5# config mlt 1 perform-tagging enable
Core-D:5# config vlan 3999 add-mlt 1
```

Create IST

```
Core-C:5# config mlt 1 ist create ip 1.1.1.6 vlan-id 3999
Core-C:5# config mlt 1 ist enable

-------------------------------------------------------------------------------

Core-D:5# config mlt 1 ist create ip 1.1.1.5 vlan-id 3999
Core-D:5# config mlt 1 ist enable
```

Create SMLT to ERS 8800 Distribution Switches

```
Core-C:5# config mlt 4 create
Core-C:5# config mlt 4 add ports 1/3,1/5,2/3,2/5
Core-C:5# config mlt 4 name "smltToClientDist-AB"
Core-C:5# config mlt 4 perform-tagging enable
Core-C:5# config mlt 4 smlt create smlt-id 4

-------------------------------------------------------------------------------

Core-D:5# config mlt 4 create
Core-D:5# config mlt 4 add ports 1/3,1/5,2/3,2/5
Core-D:5# config mlt 4 name "smltToClientDist-AB"
Core-D:5# config mlt 4 perform-tagging enable
Core-D:5# config mlt 4 smlt create smlt-id 4
```

Add VLANs to IST and SMLT

```
Core-C:5# config vlan 4 add-mlt 1
Core-C:5# config vlan 4 add-mlt 4
Core-C:5# config vlan 14 add-mlt 1
Core-C:5# config vlan 14 add-mlt 4
Core-C:5# config vlan 2000 add-mlt 1
Core-C:5# config vlan 2000 add-mlt 4

--------------------------------------------------------------------------------
Core-D:5# config vlan 4 add-mlt 1
Core-D:5# config vlan 4 add-mlt 4
Core-D:5# config vlan 14 add-mlt 1
Core-D:5# config vlan 14 add-mlt 4
Core-D:5# config vlan 2000 add-mlt 1
Core-D:5# config vlan 2000 add-mlt 4
```

3.2.3 Enable CP-Limit and Ext-CP-Limit in the Core

CP-Limit are enabled by default and automatically being disabled on ports where IST is enabled

Enable Ext-CP-Limit globally

```
Core-C:5# config sys ext-cp-limit extcplimit enable

--------------------------------------------------------------------------------
Core-D:5# config sys ext-cp-limit extcplimit enable
```

Enable Ext-CP-Limit on SMLT

```
Core-C:5# config ethernet 1/3,1/5,2/3,2/5 ext-cp-limit SoftDown
threshold-util-rate 40

--------------------------------------------------------------------------------
Core-D:5# config ethernet 1/3,1/5,2/3,2/5 ext-cp-limit SoftDown
threshold-util-rate 40
```

3.2.4 Enable VLACP and SLPP in the Core

Enable VLACP globally

```
Core-C:5# config vlacp enable

-------------------------------------------------------------------------------

Core-D:5# config vlacp enable
```

Enable VLACP on IST

```
Core-C:5# config ethernet 1/1,2/1 vlacp macaddress 01:80:c2:00:00:0f
Core-C:5# config ethernet 1/1,2/1 vlacp slow-periodic-time 10000
Core-C:5# config ethernet 1/1,2/1 vlacp timeout long
Core-C:5# config ethernet 1/1,2/1 vlacp timeout-scale 3
Core-C:5# config ethernet 1/1,2/1 vlacp enable

-------------------------------------------------------------------------------

Core-D:5# config ethernet 1/1,2/1 vlacp macaddress 01:80:c2:00:00:0f
Core-D:5# config ethernet 1/1,2/1 vlacp slow-periodic-time 10000
Core-D:5# config ethernet 1/1,2/1 vlacp timeout long
Core-D:5# config ethernet 1/1,2/1 vlacp timeout-scale 3
Core-D:5# config ethernet 1/1,2/1 vlacp enable
```

Enable VLACP on SMLT

```
Core-C:5# config ethernet 1/3,1/5,2/3,2/5 vlacp macaddress
01:80:c2:00:00:0f
Core-C:5# config ethernet 1/3,1/5,2/3,2/5 vlacp fast-periodic-time 500
Core-C:5# config ethernet 1/3,1/5,2/3,2/5 vlacp timeout short
Core-C:5# config ethernet 1/3,1/5,2/3,2/5 vlacp timeout-scale 5
Core-C:5# config ethernet 1/3,1/5,2/3,2/5 vlacp enable

-------------------------------------------------------------------------------

Core-D:5# config ethernet 1/3,1/5,2/3,2/5 vlacp macaddress
01:80:c2:00:00:0f
Core-D:5# config ethernet 1/3,1/5,2/3,2/5 vlacp fast-periodic-time 500
Core-D:5# config ethernet 1/3,1/5,2/3,2/5 vlacp timeout short
Core-D:5# config ethernet 1/3,1/5,2/3,2/5 vlacp timeout-scale 5
Core-D:5# config ethernet 1/3,1/5,2/3,2/5 vlacp enable
```

Enable SLPP for SMLT VLANs on SMLT ports

```
Core-C:5# config slpp add 4
Core-C:5# config slpp add 2000
Core-C:5# config ethernet 1/3,1/5,2/3,2/5 slpp packet-rx-threshold 5
Core-C:5# config ethernet 1/3,1/5,2/3,2/5 slpp packet-rx enable

-------------------------------------------------------------------------------
```

```
Core-D:5# config slpp add 4
Core-D:5# config slpp add 2000
Core-D:5# config ethernet 1/3,1/5,2/3,2/5 slpp packet-rx-threshold 50
Core-D:5# config ethernet 1/3,1/5,2/3,2/5 slpp packet-rx enable
```

☠ SLPP should only be enabled on the SMLT ports and not on the IST port members.

3.2.5 Enable Discard Untagged Frames

Enable Discard Untagged Frames on Core Switches

```
Core-C:5# config ethernet 1/1,1/3,1/5,2/1,2/3,2/5 untagged-frames-
discard enable

--------------------------------------------------------------------------------

Core-D:5# config ethernet 1/1,1/3,1/5,2/1,2/3,2/5 untagged-
frames-discard enable
```

3.2.6 Quality of Service

Enable QoS Trusted Interfaces on the Switch Cluster

```
Core-C:5# config ethernet 1/3,1/5,2/3,2/5 enable-diffserv true

--------------------------------------------------------------------------------

Core-D:5# config ethernet 1/3,1/5,2/3,2/5 enable-diffserv true
```

3.2.7 Layer 3 Configuration in the Core

3.2.7.1 RSMLT

For RSMLT configuration, the following parameters will be used:

VLAN Name	VLAN ID	IP Address	Hold Down Timer	Hold Up Timer
Core SMLT-1	4	10.0.4.1 10.0.4.2	60	180
Core SMLT-2	14	10.0.14.1 10.0.14.2	60	180

Enable RSMLT

```
Core-C:5# config vlan 4 ip rsmlt enable
Core-C:5# config vlan 14 ip rsmlt enable

--------------------------------------------------------------------------------

Core-D:5# config vlan 4 ip rsmlt enable
Core-D:5# config vlan 14 ip rsmlt enable
```

3.2.7.2   OSPF

Enable OSPF Globally and on each VLAN

```
Core-C:5# config ip circuitless-ip-int 1 create
10.0.0.5/255.255.255.255
Core-C:5# config ip circuitless-ip-int 1 ospf enable
Core-C:5# config ip ospf router-id 10.0.0.5
Core-C:5# config ip ospf enable
Core-C:5# config vlan 4 ip ospf enable
Core-C:5# config vlan 4 ip ospf priority 100
Core-C:5# config vlan 14 ip ospf enable
Core-C:5# config vlan 14 ip ospf priority 0

--------------------------------------------------------------------------------

Core-D:5# config ip circuitless-ip-int 1 create
10.0.0.6/255.255.255.255
Core-D:5# config ip circuitless-ip-int 1 ospf enable
Core-D:5# config ip ospf router-id 10.0.0.6
Core-D:5# config ip ospf enable
Core-D:5# config vlan 4 ip ospf enable
Core-C:5# config vlan 4 ip ospf priority 0
Core-D:5# config vlan 14 ip ospf enable
Core-D:5# config vlan 14 ip ospf priority 100
```

3.2.7.3 PIM-SM

Assuming candidate RPs and BSRs are configured on Core-C and Core-D switches.

Enable PIM-SM on IST and SMLT VLANs

```
Core-C:5# sys mcast-smlt square-smlt enable
Core-C:5# config ip pim enable
Core-C:5# config ip pim fast-joinprune enable
Core-C:5# config vlan 3999 ip pim enable
Core-C:5# config vlan 4 ip pim enable
Core-C:5# config vlan 14 ip pim enable

Core-C:5# config ip circuitless-ip-int 1 pim enable
Core-C:5# config ip pim candrp add grp 224.0.0.0 mask 255.0.0.0 rp
10.0.0.5
Core-C:5# config ip pim interface 10.0.0.7 cbsrpreference 100 enable

------------------------------------------------------------------

Core-D:5# sys mcast-smlt square-smlt enable
Core-D:5# config ip pim enable
Core-D:5# config ip pim fast-joinprune enable
Core-D:5# config vlan 3999 ip pim enable
Core-D:5# config vlan 4 ip pim enable
Core-D:5# config vlan 14 ip pim enable
Core-D:5# config vlan 2101 ip pim enable
Core-D:5# config vlan 2201 ip pim enable

Core-D:5# config ip circuitless-ip-int 1 pim enable
Core-D:5# config ip pim candrp add grp 224.0.0.0 mask 255.0.0.0 rp
10.0.0.6
Core-D:5# config ip pim interface 10.0.0.7 cbsrpreference 50 enable
```

3.3 Distribution Configuration

This section will cover the major configuration requirements of the Switch Cluster Distribution.

3.3.1 Create VLANs in the Distribution

Create SMLT VLANs, Edge VLANs and IST VLAN on ERS8600-1 & ERS8600-2 Distribution Switches

```
Dist-A:5# config vlan 4 create byport 1 name "Core-SMLT-1"
Dist-A:5# config vlan 4 fdb-entry aging-time 21601
Dist-A:5# config vlan 14 create byport 1 name "Core-SMLT-2"
Dist-A:5# config vlan 14 fdb-entry aging-time 21601
Dist-A:5# config vlan 2000 create byport 1 name "Mgmt"
Dist-A:5# config vlan 2101 create byport 1 name "Data-1"
Dist-A:5# config vlan 2102 create byport 1 name "Voice-1"
Dist-A:5# config vlan 2201 create byport 1 name "Data-2"
Dist-A:5# config vlan 2202 create byport 1 name "Voice-2"
Dist-A:5# config vlan 3999 create byport 1 name "istVlan"

-------------------------------------------------------------------------------

Dist-B:5# config vlan 4 create byport 1 name "Core-SMLT-1"
Dist-B:5# config vlan 4 fdb-entry aging-time 21601
Dist-B:5# config vlan 14 create byport 1 name "Core-SMLT-2"
Dist-B:5# config vlan 4 fdb-entry aging-time 21601
Dist-B:5# config vlan 2000 create byport 1 name "Mgmt"
Dist-B:5# config vlan 2101 create byport 1 name "Data-1"
Dist-B:5# config vlan 2102 create byport 1 name "Voice-1"
Dist-B:5# config vlan 2201 create byport 1 name "Data-2"
Dist-B:5# config vlan 2202 create byport 1 name "Voice-2"
Dist-B:5# config vlan 3999 create byport 1 name "istVlan"
```

Add IP addresses to the Edge and Management VLANs

```
Dist-A:5# config vlan 4 ip create 10.0.4.3/255.255.255.0
Dist-A:5# config vlan 14 ip create 10.0.14.3/255.255.255.0
Dist-A:5# config vlan 2000 ip create 10.20.0.7/255.255.255.0
Dist-A:5# config vlan 2101 ip create 10.21.1.1/255.255.255.0
Dist-A:5# config vlan 2102 ip create 10.21.2.1/255.255.255.0
Dist-A:5# config vlan 2201 ip create 10.22.1.1/255.255.255.0
Dist-A:5# config vlan 2202 ip create 10.22.2.1/255.255.255.0
Dist-A:5# config vlan 3999 ip create 1.1.1.9/255.255.255.252

-------------------------------------------------------------------------------

Dist-B:5# config vlan 4 ip create 10.0.4.4/255.255.255.0
Dist-B:5# config vlan 14 ip create 10.0.14.4/255.255.255.0
Dist-B:5# config vlan 2000 ip create 10.20.0.8/255.255.255.0
Dist-B:5# config vlan 2101 ip create 10.21.1.2/255.255.255.0
Dist-B:5# config vlan 2102 ip create 10.21.2.2/255.255.255.0
```

```
Dist-B:5# config vlan 2201 ip create 10.22.1.2/255.255.255.0
Dist-B:5# config vlan 2202 ip create 10.22.2.2/255.255.255.0
Dist-B:5# config vlan 3999 ip create 1.1.1.10/255.255.255.252
```

3.3.2 Create the Distribution Switch Cluster

Create MLT 1 for IST

```
Dist-A:5# config mlt 1 create
Dist-A:5# config mlt 1 name "istToClientDistB"
Dist-A:5# config mlt 1 add ports 1/1,2/1
Dist-A:5# config mlt 1 perform-tagging enable
Dist-A:5# config vlan 3999 add-mlt 1
-------------------------------------------------------------------------------
Dist-B:5# config mlt 1 create
Dist-B:5# config mlt 1 name "istToClientDistA"
Dist-B:5# config mlt 1 add ports 1/1,2/1
Dist-B:5# config mlt 1 perform-tagging enable
Dist-B:5# config vlan 3999 add-mlt 1
```

Create IST

```
Dist-A:5# config mlt 1 ist create ip 1.1.1.10 vlan-id 3999
Dist-A:5# config mlt 1 ist enable

-------------------------------------------------------------------------------

Dist-B:5# config mlt 1 ist create ip 1.1.1.9 vlan-id 3999
Dist-B:5# config mlt 1 ist enable
```

Create SMLT to ERS 8800 Distribution switches

```
Dist-A:5# config mlt 4 create
Dist-A:5# config mlt 4 add ports 1/3,1/5,2/3,2/5
Dist-A:5# config mlt 4 name "smltToCore-AB"
Dist-A:5# config mlt 4 perform-tagging enable
Dist-A:5# config mlt 4 smlt create smlt-id 4
Dist-A:5# config mlt 11 create
Dist-A:5# config mlt 11 add ports 1/11
Dist-A:5# config mlt 11 name "smltToEdge1"
Dist-A:5# config mlt 11 perform-tagging enable
Dist-A:5# config mlt 11 smlt create smlt-id 11

-------------------------------------------------------------------------------

Dist-B:5# config mlt 4 create
Dist-B:5# config mlt 4 add ports 1/3,1/5,2/3,2/5
Dist-B:5# config mlt 4 name "smltToCore-AB"
```

```
Dist-B:5# config mlt 4 perform-tagging enable
Dist-B:5# config mlt 4 smlt create smlt-id 4
Dist-B:5# config mlt 11 create
Dist-B:5# config mlt 11 add ports 1/11
Dist-B:5# config mlt 11 name "smltToEdge1"
Dist-B:5# config mlt 11 perform-tagging enable
Dist-B:5# config mlt 11 smlt create smlt-id 11
```

Create SLT to ERS 8800 Distribution switches

```
Dist-A:5# config ethernet 1/12 name sltToEdge1
Dist-A:5# config ethernet 1/12 perform-tagging enable
Dist-A:5# config ethernet 1/12 smlt 129 create

--------------------------------------------------------------------------------

Dist-B:5# config ethernet 1/12 name sltToEdge1
Dist-B:5# config ethernet 1/12 perform-tagging enable
Dist-B:5# config ethernet 1/12 smlt 129 create
```

Add VLANs to IST, SMLT and SLT

```
Dist-A:5# config vlan 4 add-mlt 1
Dist-A:5# config vlan 4 add-mlt 4
Dist-A:5# config vlan 14 add-mlt 1
Dist-A:5# config vlan 14 add-mlt 4
Dist-A:5# config vlan 2000 add-mlt 1
Dist-A:5# config vlan 2000 add-mlt 4
Dist-A:5# config vlan 2000 add-mlt 11
Dist-A:5# config vlan 2101 add-mlt 1
Dist-A:5# config vlan 2101 add-mlt 11
Dist-A:5# config vlan 2102 add-mlt 1
Dist-A:5# config vlan 2102 add-mlt 11
Dist-A:5# config vlan 2201 add-mlt 1
Dist-A:5# config vlan 2201 add port 1/12
Dist-A:5# config vlan 2202 add-mlt 1
Dist-A:5# config vlan 2202 add port 1/12

--------------------------------------------------------------------------------
Dist-B:5# config vlan 4 add-mlt 1
Dist-B:5# config vlan 4 add-mlt 4
Dist-B:5# config vlan 14 add-mlt 1
Dist-B:5# config vlan 14 add-mlt 4
Dist-B:5# config vlan 2000 add-mlt 1
Dist-B:5# config vlan 2000 add-mlt 4
Dist-B:5# config vlan 2000 add-mlt 11
Dist-B:5# config vlan 2101 add-mlt 1
Dist-B:5# config vlan 2101 add-mlt 11
```

```
Dist-B:5# config vlan 2102 add-mlt 1
Dist-B:5# config vlan 2102 add-mlt 11
Dist-B:5# config vlan 2201 add-mlt 1
Dist-B:5# config vlan 2201 add port 1/12
Dist-B:5# config vlan 2202 add-mlt 1
Dist-B:5# config vlan 2202 add port 1/12
```

3.3.3 Enable CP-Limit and Ext-CP-Limit in the Distribution

CP-Limit are enabled by default and automatically being disabled on ports where IST is enabled

---

Enable Ext-CP-Limit globally

```
Dist-A:5# config sys ext-cp-limit extcplimit enable

--------------------------------------------------------------------------------

Dist-B:5# config sys ext-cp-limit extcplimit enable
```

Enable Ext-CP-Limit on SMLT to Core and Edge Switches

```
Dist-A:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 ext-cp-limit
SoftDown threshold-util-rate 40

--------------------------------------------------------------------------------

Dist-B:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 ext-cp-limit
SoftDown threshold-util-rate 40
```

3.3.4 Enable VLACP and SLPP in the Distribution

Enable VLACP globally

```
Dist-A:5# config vlacp enable

--------------------------------------------------------------------------------

Dist-B:5# config vlacp enable
```

Enable VLACP on IST

```
Dist-A:5# config ethernet 1/1,2/1 vlacp macaddress 01:80:c2:00:00:0f
```

```
Dist-A:5# config ethernet 1/1,2/1 vlacp slow-periodic-time 10000
Dist-A:5# config ethernet 1/1,2/1 vlacp timeout long
Dist-A:5# config ethernet 1/1,2/1 vlacp timeout-scale 3
Dist-A:5# config ethernet 1/1,2/1 vlacp enable

---------------------------------------------------------------------------

Dist-B:5# config ethernet 1/1,2/1 vlacp macaddress 01:80:c2:00:00:0f
Dist-B:5# config ethernet 1/1,2/1 vlacp slow-periodic-time 10000
Dist-B:5# config ethernet 1/1,2/1 vlacp timeout long
Dist-B:5# config ethernet 1/1,2/1 vlacp timeout-scale 3
Dist-B:5# config ethernet 1/1,2/1 vlacp enable
```

Enable VLACP on SMLT to Core and Edge Switches

```
Dist-A:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 vlacp macaddress
01:80:c2:00:00:0f
Dist-A:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 vlacp fast-
periodic-time 500
Dist-A:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 vlacp timeout short
Dist-A:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 vlacp timeout-scale
5
Dist-A:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 vlacp enable

--------------------------------------------------------------------------

Dist-B:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 vlacp macaddress
01:80:c2:00:00:0f
Dist-B:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 vlacp fast-
periodic-time 500
Dist-B:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 vlacp timeout short
Dist-B:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 vlacp timeout-scale
5
Dist-B:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 vlacp enable
```

Enable SLPP for SMLT VLANs on SMLT ports

```
Dist-A:5# config slpp add 4
Dist-A:5# config slpp add 2000
Dist-A:5# config slpp add 2101
Dist-A:5# config slpp add 2102
Dist-A:5# config slpp add 2201
Dist-A:5# config slpp add 2202
Dist-A:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 slpp packet-rx-
threshold 5
Dist-A:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 slpp packet-rx
enable

---------------------------------------------------------------------------
```

```
Dist-B:5# config slpp add 4
Dist-B:5# config slpp add 2000
Dist-B:5# config slpp add 2101
Dist-B:5# config slpp add 2102
Dist-B:5# config slpp add 2201
Dist-B:5# config slpp add 2202
Dist-B:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 slpp packet-rx-
threshold 50
Dist-B:5# config ethernet 1/3,1/5,1/11,1/12,2/3,2/5 slpp packet-rx
enable
```

☠ SLPP should only be enabled on the SMLT ports and not on the IST port members.

3.3.5 Enable Discard Untagged Frames

Enable Discard Untagged Frames on Distribution Switches
| ```
Dist-A:5# config ethernet 1/1,1/3,1/5,1/11,1/12,2/1,2/3,2/5 untagged-
frames-discard enable

-------------------------------------------------------------------------------

Dist-B:5# config ethernet 1/1,1/3,1/5,1/11,1/12,2/1,2/3,2/5
untagged-frames-discard enable
``` |

3.3.6 Quality of Service

Create trusted interfaces (honoring QoS marking) for the IST, and SMLT/SLT uplink ports on the ERS 8800 Switch Cluster.

Enable QoS Trusted Interfaces on the Switch Cluster
| ```
Dist-A:5# config ethernet 1/1,1/3,1/5,1/11,1/12,2/1,2/3,2/5 enable-
diffserv true

-------------------------------------------------------------------------------

Dist-B:5# config ethernet 1/1,1/3,1/5,1/11,1/12,2/1,2/3,2/5 enable-
diffserv true
``` |

3.3.7   Layer 3 Configuration in the Distribution

Enable RSMLT and OSPF. Please note, that these are not required features for the Large Campus solution, but are normally required in most instances. These are all independent features that can be used with or without each other.

3.3.7.1   RSMLT

VLAN Name	VLAN ID	IP Address	Hold Down Timer	Hold Up Timer
Core SMLT-1	4	10.0.4.3 10.0.4.4	60	180
Core SMLT-2	14	10.0.14.3 10.0.14.4	60	180
Data-1	2101	10.21.1.1 10.21.1.2	60	infinity
Data-2	2201	10.22.1.1 10.22.1.2	60	Infinity

Enable RSMLT

```
Dist-A:5# config vlan 4 ip rsmlt enable
Dist-A:5# config vlan 14 ip rsmlt enable
Dist-A:5# config vlan 2101 ip rsmlt enable
Dist-A:5# config vlan 2101 ip rsmlt holdup-timer 9999
Dist-A:5# config vlan 2201 ip rsmlt enable
Dist-A:5# config vlan 2201 ip rsmlt holdup-timer 9999
Dist-A:5# config ip rsmlt rsmlt-edge-support enable

---------------------------------------------------------------------------
Dist-B:5# config vlan 4 ip rsmlt enable
Dist-B:5# config vlan 14 ip rsmlt enable
Dist-B:5# config vlan 2101 ip rsmlt enable
Dist-B:5# config vlan 2101 ip rsmlt holdup-timer 9999
Dist-B:5# config vlan 2201 ip rsmlt enable
Dist-B:5# config vlan 2201 ip rsmlt holdup-timer 9999
Dist-B:5# config ip rsmlt rsmlt-edge-support enable
```

3.3.7.2 VRRP

For the VRRP configuration, the following parameters will be used:

VLAN Name	VLAN ID	VRID	IP Address	Dist-A Priority	Dist-B Priority	Adv Interval	HoldDown Timer
Voice1	2102	21	10.21.2.254	200	Default	10	60
Voice2	2202	22	10.22.2.254	Default	200	10	60

Enable VRRP and Backup Master

```
Dist-A:5# config vlan 2102 ip vrrp 21 address 10.21.2.254
Dist-A:5# config vlan 2102 ip vrrp 21 backup-master enable
Dist-A:5# config vlan 2102 ip vrrp 21 priority 200
Dist-A:5# config vlan 2102 ip vrrp 21 adver-int 10
Dist-A:5# config vlan 2102 ip vrrp 21 holddown-timer 60
Dist-A:5# config vlan 2102 ip vrrp 21 enable

Dist-A:5# config vlan 2202 ip vrrp 22 address 10.21.2.254
Dist-A:5# config vlan 2202 ip vrrp 22 backup-master enable
Dist-A:5# config vlan 2102 ip vrrp 22 adver-int 10
Dist-A:5# config vlan 2102 ip vrrp 22 holddown-timer 60
Dist-A:5# config vlan 2202 ip vrrp 22 enable


----------------------------------------------------------------------


Dist-B:5# config vlan 2102 ip vrrp 21 address 10.21.2.254
Dist-B:5# config vlan 2102 ip vrrp 21 backup-master enable
Dist-B:5# config vlan 2102 ip vrrp 21 adver-int 10
Dist-B:5# config vlan 2102 ip vrrp 21 holddown-timer 60
Dist-B:5# config vlan 2102 ip vrrp 21 enable

Dist-B:5# config vlan 2202 ip vrrp 22 address 10.21.2.254
Dist-B:5# config vlan 2202 ip vrrp 22 backup-master enable
Dist-B:5# config vlan 2202 ip vrrp 22 priority 200
Dist-B:5# config vlan 2102 ip vrrp 22 adver-int 10
Dist-B:5# config vlan 2102 ip vrrp 22 holddown-timer 60
Dist-B:5# config vlan 2202 ip vrrp 22 enable
```

3.3.7.3   DHCP Relay

To enable DHCP and DHCP relay, enter the following assuming the DHCP server IP address is 10.20.0.10.

Enable DHCP Reply Agent

```
Dist-A:5# config vlan 2101 ip dhcp enable
Dist-A:5# config vlan 2102 ip dhcp enable
Dist-A:5# config vlan 2201 ip dhcp enable
Dist-A:5# config vlan 2202 ip dhcp enable

Dist-A:5# config ip dhcp-relay create-fwd-path agent 10.21.1.1 server
10.20.0.10 mode bootp_dhcp state enable
Dist-A:5# config ip dhcp-relay create-fwd-path agent 10.21.2.254 server
10.20.0.10 mode bootp_dhcp state enable
Dist-A:5# config ip dhcp-relay create-fwd-path agent 10.22.1.1 server
10.20.0.10 mode bootp_dhcp state enable
Dist-A:5# config ip dhcp-relay create-fwd-path agent 10.22.2.254 server
10.20.0.10 mode bootp_dhcp state enable

--------------------------------------------------------------------

Dist-B:5# config vlan 2101 ip dhcp enable
Dist-B:5# config vlan 2102 ip dhcp enable
Dist-B:5# config vlan 2201 ip dhcp enable
Dist-B:5# config vlan 2202 ip dhcp enable

Dist-B:5# config ip dhcp-relay create-fwd-path agent 10.21.1.2 server
10.20.0.10 mode bootp_dhcp state enable
Dist-B:5# config ip dhcp-relay create-fwd-path agent 10.21.2.254 server
10.20.0.10 mode bootp_dhcp state enable
Dist-B:5# config ip dhcp-relay create-fwd-path agent 10.22.1.2 server
10.20.0.10 mode bootp_dhcp state enable
Dist-B:5# config ip dhcp-relay create-fwd-path agent 10.22.2.254 server
10.20.0.10 mode bootp_dhcp state enable
```

3.3.7.4 OSPF

Each Edge VLAN will be configured with OSPF passive interface on the Switch Cluster core.

The following shows the OSPF configuration for VLAN 2101. These steps would be repeated for other edge VLAN (ex. 2201, 2202)

Enable OSPF Globally and on each SMLT-CORE VLANs and Edge VLANs

```
Dist-A:5# config ip circuitless-ip-int 1 create
10.0.0.7/255.255.255.255
Dist-A:5# config ip circuitless-ip-int 1 ospf enable
Dist-A:5# config ip ospf router-id 10.0.0.7
Dist-A:5# config ip ospf enable
Dist-A:5# config vlan 4 ip ospf priority 0
Dist-A:5# config vlan 4 ip ospf enable
Dist-A:5# config vlan 14 ip ospf priority 0
Dist-A:5# config vlan 14 ip ospf enable

Dist-A:5# config vlan 2101 ip ospf interface-type passive
Dist-A:5# config vlan 2101 ip ospf enable
Dist-A:5# config vlan 2102 ip ospf interface-type passive
Dist-A:5# config vlan 2102 ip ospf enable

---------------------------------------------------------------------

Dist-B:5# config ip circuitless-ip-int 1 create
10.0.0.8/255.255.255.255
Dist-B:5# config ip circuitless-ip-int 1 ospf enable
Dist-B:5# config ip ospf router-id 10.0.0.8
Dist-B:5# config ip ospf enable
Dist-B:5# config vlan 4 ip ospf priority 0
Dist-B:5# config vlan 4 ip ospf enable
Dist-B:5# config vlan 14 ip ospf priority 0
Dist-B:5# config vlan 14 ip ospf enable

Dist-B:5# config vlan 2101 ip ospf interface-type passive
Dist-B:5# config vlan 2101 ip ospf enable
Dist-B:5# config vlan 2102 ip ospf interface-type passive
Dist-B:5# config vlan 2102 ip ospf enable
```

3.3.7.5 PIM-SM

Enable PIM-SM on IST and SMLT VLANs

```
Dist-A:5# sys mcast-smlt square-smlt enable
Dist-A:5# config ip pim enable
Dist-A:5# config ip pim fast-joinprune enable
Dist-A:5# config vlan 3999 ip pim enable
Dist-A:5# config vlan 4 ip pim enable
Dist-A:5# config vlan 14 ip pim enable
Dist-A:5# config vlan 2101 ip pim enable
Dist-A:5# config vlan 2201 ip pim enable

Dist-A:5# config ip circuitless-ip-int 1 pim enable

----------------------------------------------------------------------

Dist-B:5# sys mcast-smlt square-smlt enable
Dist-B:5# config ip pim enable
Dist-B:5# config ip pim fast-joinprune enable
Dist-B:5# config vlan 3999 ip pim enable
Dist-B:5# config vlan 4 ip pim enable
Dist-B:5# config vlan 14 ip pim enable
Dist-B:5# config vlan 2101 ip pim enable
Dist-B:5# config vlan 2201 ip pim enable

Dist-B:5# config ip circuitless-ip-int 1 pim enable
```

3.4 Edge Configuration

This section will cover the major configuration requirements of the Edge. Please note that other features are available and can be implemented, for the purpose of this document, the basics are covered and the rest is left to the product documentation.

The configuration of one Edge stack is covered in this section. All configurations can be applied to the other Edge switches/stacks. Where differences in configuration occur, a note is provided to highlight these.

3.4.1 Create VLANs at the Edge

The following example creates the Data VLANs on each Edge Stack along with the Management VLAN. The Voice VLAN is not created in this step when ADAC is used. If Avaya Automatic QoS is used, then the Voice VLAN would need to be created and applied to the proper ports during this step of the configuration.

Create Edge VLANs

```
4500-2(config)# vlan create 2101 name data1 type port

4500-2(config)# vlan create 2000 name mgmt type port
```

```
4500-2(config)# vlan ports 1/50,2/50 tagging tagAll filter-untagged-frame
enable
4500-2(config)# vlan members remove 1 1/1-3/50
4500-2(config)# vlan members add 2101 1/1-3/50
4500-2(config)# vlan members add 2000 1/50,2/50
```

> 👍 The ERS 2500 and ERS 4500 switches use the VLAN configuration mode of *strict*
> (default setting). In this mode, you must first remove port members from the default
> VLAN 1 prior to adding these port members to a new VLAN. The VLAN configuration
> mode is set by using the command *vlan configcontrol*
> *<automatic|autopvid|flexible|strict>*

3.4.2 Create the MLT on the Edge

Create MLT 1

```
4500-2(config)#mlt 1 name Dist-AB member 1/50,2/50 learning disable
4500-2(config)#mlt 1 enable
```

> 👍 When configuring the MLT group, Spanning Tree must be disabled on the MLT. The
> *learning disable* on the MLT command disables Spanning Tree for the MLT.

3.4.3 Create Management IP Address for Edge

In this configuration, the Edge stacks are Layer 2. Only the Management VLAN needs and IP
address. Please also note that this IP address is used to manage the entire stack of switches and
also the address used as the source address for Syslog entries, SNMP, and other switch/stack
management functions.

Add IP address and assign the Management VLAN

```
4500-2(config)#interface vlan 2000
4500-2(config-if)#ip address 10.20.0.12 255.255.255.0
4500-2(config-if)#exit
4500-2(config)#vlan mgmt 2000
```

3.4.4 VLACP at the Edge

Configure the VLACP MAC and enable VLACP globally

```
4500-2(config)#vlacp macaddress 180.c200.f
4500-2(config)#vlacp enable
```

👍 It is recommended to use the reserved multicast MAC address of 01:80:c2:00:00:0f for the VLACP MAC address. Enter the hex value *180.c200.f*.

Enable VLACP on MLT

```
4500-2(config)#interface FastEthernet all

4500-2(config-if)#vlacp port 1/50,2/50 timeout short

4500-2(config-if)#vlacp port 1/50,2/50 fast-periodic-time 500

4500-2(config-if)#vlacp port 1/50,2/50 timeout-scale 5

4500-2(config-if)#vlacp port 1/50,2/50 enable

4500-2(config-if)#exit
```

👍 Both ends of the link must be configured for VLACP with the same parameters. The VLACP configuration for the Switch Cluster Core was shown earlier in this section.

3.4.5 Enable STP Fast Start / BPDU Filtering at the Edge

Spanning Tree Fast Start and BPDU Filtering are recommended for all Edge access ports. These features should never be enabled on the uplinks from the Edge to the Switch Cluster Core.

Enable STP Fast Start and BPDU Filtering on all Access ports

```
4500-2(config)#interface fastEthernet 1/1-1/49,2/1-2/49,3/1-3/50

4500-2(config-if)#spanning-tree learning fast

4500-2(config-if)#spanning-tree bpdu-filtering timeout 0

4500-2(config-if)#spanning-tree bpdu-filtering enable

4500-2(config-if)#exit
```

3.4.6 Enable Rate Limiting

The Rate Limiting configuration on the Edge switches is different between the ERS 2500 and the ERS 4500 / 5000. For that reason, both configuration examples are provided below.

The rate limit parameter on the ERS 4500 / ERS 5000 is expressed as percentage of port speed whereas for the ERS2500 it is expressed in packets per second (pps). The Rate Limiting configuration on the ERS 2500 is per switch, not per port.

The values used in this example are recommendations but may vary depending on the network environment and applications being used.

Enable Rate Limiting to 10% of total traffic for both broadcast and multicast traffic

```
4500-2(config)#interface fastEthernet all

4500-2(config-if)#rate-limit port 1/1-1/49,2/1-2/49,3/1-3/50 both 10
```

```
4500-2(config-if)#exit
```

Enable Rate Limiting to 262143 pps for both broadcast and multicast traffic

```
2500-1(config)#interface fastEthernet all

2500-1(config-if)#rate-limit both 262143

2500-1(config-if)#exit
```

3.4.7 Quality of Service

Two options for configuring QoS in regard to IP Telephony are shown below. The first option uses the ADAC feature, while the second option shows the Avaya Automatic QoS configuration. Use one or the other, but not both simultaneously.

Enable ADAC

```
4500-2(config)#adac enable op-mode tagged-frames voice-vlan 2202 uplink-port
1/50

4500-2(config)#adac port 1/1-1/48,2/1-2/48,3/1-3/48 enable
```

ⓘ The Data VLAN must be created and provisioned on the port BEFORE enabling ADAC. If the Data VLAN is added to the port after enabling ADAC, once ADAC is disabled, the PVID of the port will be reset to 1 (default VLAN).

Enable Avaya Automatic QoS

```
4500-2(config)#qos agent nt-mode mixed
```

ⓘ The Voice VLAN must be created and provisioned manually when using this feature. Please refer to section 3.2.1 for the configuration steps for this.

The CS1000 or BCM must also be configured to support Avaya Automatic QoS (also known as NT on NT) in order for the phones to use the correct DSCP values.

3.4.8 Enable Security Features

DHCP Snooping, Dynamic ARP Inspection, and IP Source Guard should be enabled on the Edge switch/stack. DHCP Snooping must be enabled to use Dynamic ARP Inspection and IP Source Guard. Certain restrictions apply to IP Source Guard configuration as noted below. In certain cases, IP Source Guard may not be used when it is in conflict with other, more critical features required at the Edge.

Enable DHCP Snooping

```
4500-2(config)#ip dhcp-snooping enable

4500-2(config)#ip dhcp-snooping vlan 2201
```

```
4500-2(config)#ip dhcp-snooping vlan 2202

4500-2(config)#interface fastEthernet 1/50,2/50

4500-2(config-if)# ip dhcp-snooping trusted

4500-2(config-if)#exit
```

Enable Dynamic ARP Inspection

```
4500-2(config)#ip arp-inspection vlan 2201

4500-2(config)#ip arp-inspection vlan 2202

4500-2(config)#interface fastEthernet 1/50,2/50

4500-2(config-if)# ip arp-inspection trusted

4500-2(config-if)#exit
```

IP Source Guard should not be enabled on the uplink ports from the Edge stack, but only on the Edge access ports. Also note that IP Source Guard cannot be enabled if Baysecure (MAC security) or EAPOL is enabled on the Edge access ports. IP Source Guard can only be enabled on ports configured with DHCP Snooping and Dynamic ARP Inspection.

Enable IP Source Guard

```
4500-2(config)#interface fastEthernet all

4500-2(config-if)# ip verify source interface fastEthernet 1/1-1/49,2/1-
2/49,3/1-3/50

4500-2(config-if)#exit
```

3.4.9 Enable Multicast Features

If there is a requirement to run multicast on the network, it is recommended to enable IGMP Snooping and Proxy at the edge. Presently, the ERS 5000 does not support IGMP over SMLT/SLT, so traffic in the core will be flooded; however, traffic at the edge will be pruned as expected. If no multicast traffic is expected on the network, leave both Snooping and Proxy disabled (default configuration).

In this example, IGMP is enabled on the Data VLAN but not on the Voice or Management VLAN.

Enable Multicast

```
4500-2(config)#vlan igmp 2201 snooping enable

4500-2(config)#vlan igmp 2201 proxy enable
```

3.4.10 Enable EAPOL Features

3.4.10.1   802.1X SHSA (Single Host Single Authentication)

Assuming VLAN 2200 is being configured as a guest VLAN

Enable EAPOL – 802.1X SHSA

```
4500-2(config)#interface fastEthernet 1/1-1/49,2/1-2/49,3/1-3/50
4500-2(config)#eapol guest-vlan port 1/1-1/49,2/1-2/49,3/1-3/50 vid 2200
4500-2(config)#eapol guest-vlan port 1/1-1/49,2/1-2/49,3/1-3/50 enable

4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 use-
radius-assigned-vlan
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 enable
4500-2(config)#eapol status auto
4500-2(config)#exit

4500-2(config)#eapol guest-vlan vid 2200
4500-2(config)#eapol guest-vlan enable
4500-2(config)#eapol multihost use-radius-assigned-vlan


4500-2(config)#eapol enable
```

3.4.10.2   802.1X MHMA (Multiple Host Multiple Authentication

Assuming VLAN 2200 is being configured as a guest VLAN

Enable EAPOL – 802.1X MHMA

```
4500-2(config)#interface fastEthernet 1/1-1/49,2/1-2/49,3/1-3/50
4500-2(config)#eapol guest-vlan port 1/1-1/49,2/1-2/49,3/1-3/50 vid 2200
4500-2(config)#eapol guest-vlan port 1/1-1/49,2/1-2/49,3/1-3/50 enable
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 eap-
packet-mode unicast
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 use-
radius-assigned-vlan
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 eap-mac-
max 2
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 enable
4500-2(config)#eapol status auto
4500-2(config)#exit

4500-2(config)#eapol guest-vlan vid 2200
4500-2(config)#eapol guest-vlan enable
4500-2(config)#eapol multihost eap-packet-mode unicast
4500-2(config)#eapol multihost use-radius-assigned-vlan
```
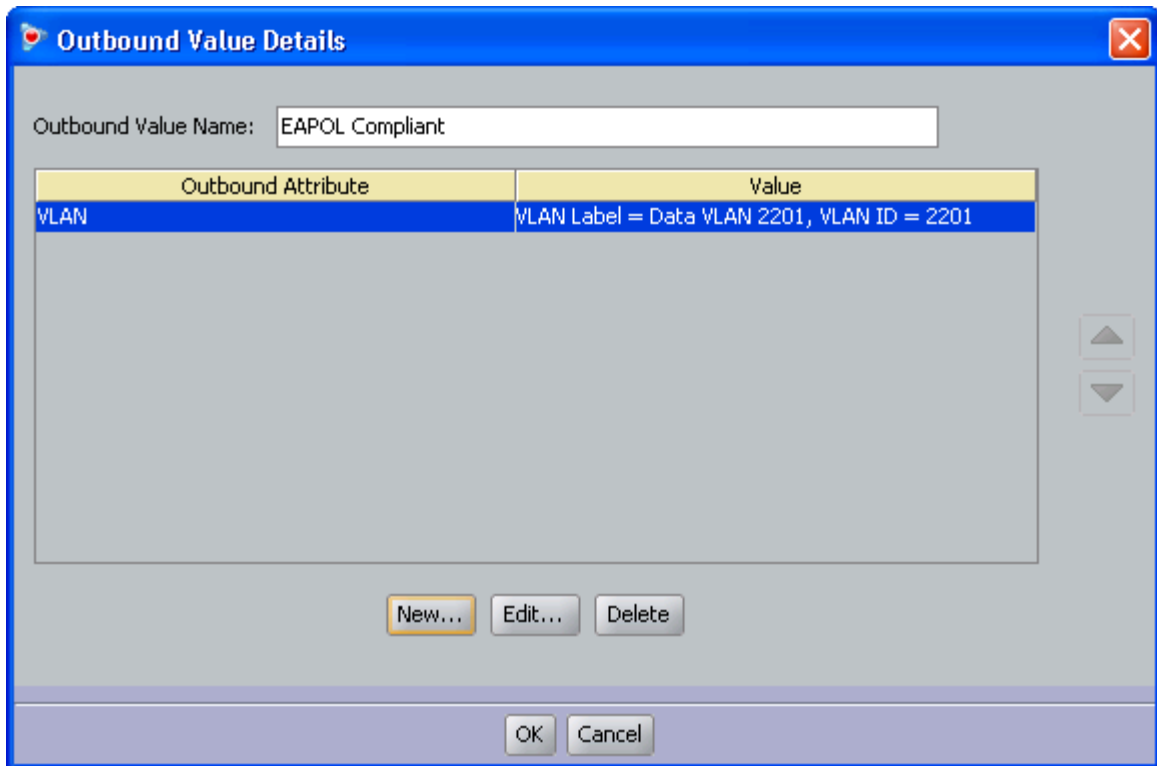

```
4500-2(config)#eapol enable
```

3.4.10.3 8021.X Non-EAP MAC Authentication

Enable EAPOL – Non-EAP MAC Authentication

```
4500-2(config)#interface fastEthernet 1/1-1/49,2/1-2/49,3/1-3/50
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 allow-non-
eap-enable
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 radius-
non-eap-enable
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 non-eap-
use-radius-assigned-vlan
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 enable
4500-2(config)#eapol status auto
4500-2(config)#exit

4500-2(config)#eapol multihost allow-non-eap-enable
4500-2(config)#eapol multihost radius-non-eap-enable
4500-2(config)#eapol multihost non-eap-use-radius-assigned-vlan
4500-2(config)#no eapol multihost non-eap-pwd-fmt
4500-2(config)#eapol multihost non-eap-pwd-fmt mac-addr

4500-2(config)#eapol enable
```

3.4.10.4 802.1X Non-EAP IP Phone Authentication

Enable EAPOL – Non-EAP IP Phone Authentication

```
4500-2(config)#interface fastEthernet 1/1-1/49,2/1-2/49,3/1-3/50
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 non-eap-
phone-enable
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/49,3/1-3/50 enable
4500-2(config)#eapol status auto
4500-2(config)#exit

4500-2(config)#eapol multihost non-eap-phone-enable

4500-2(config)#eapol enable
```

3.4.10.5 802.1X MHMA - ADAC

Note: MLT is not supported on ADAC uplink port.

Enable EAPOL – 802.1X MHMA - ADAC

```
4500-2(config)#adac voice-vlan 2202
4500-2(config)#adac op-mode untagged-frames-advanced
4500-2(config)#adac uplink-port 1/50

4500-2(config)#adac enable

4500-2(config)#interface fastEthernet 1/1-1/49,2/1-2/50,3/1-3/50
4500-2(config)#adac port 1/1-1/49,2/1-2/50,3/1-3/50 enable
4500-2(config)#eapol guest-vlan port 1/1-1/49,2/1-2/50,3/1-3/50 vid 2200
4500-2(config)#eapol guest-vlan port 1/1-1/49,2/1-2/50,3/1-3/50 enable
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/50,3/1-3/50 eap-
packet-mode unicast
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/50,3/1-3/50 use-
radius-assigned-vlan
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/50,3/1-3/50 eap-mac-
max 2
4500-2(config)#eapol multihost port 1/1-1/49,2/1-2/50,3/1-3/50 enable
4500-2(config)#eapol status auto
4500-2(config)#exit

4500-2(config)#eapol guest-vlan vid 2200
4500-2(config)#eapol guest-vlan enable
4500-2(config)#eapol multihost eap-packet-mode unicast
4500-2(config)#eapol multihost use-radius-assigned-vlan
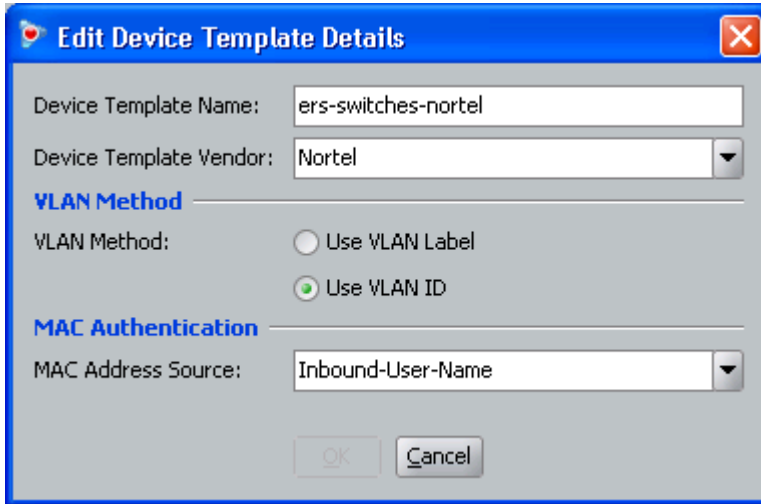

4500-2(config)#eapol enable
```

```
Phone setting:
DHCP? (0-No, 1-Yes): 1
DHCP: 0-Full, 1-Partial: 1
S1 IP: 10.100.2.11
S1 PORT: 4100
S1 ACTION: 1
S1 RETRY COUNT: 1-3
S2 IP: 10.100.3.11
S2 PORT: 4100
S2 ACTION: 1
S2 RETRY COUNT: 1-3
Voice VLAN? 0-No, 1-Yes: 1
VLAN Cfg? 0-Auto, 1-Man: 1
Voice VLAN ID: 2304
VLANFILTER? 0-No, 1-Yes: 0
PC Port? 1-ON, 0-OFF: 1
```

3.5 IdEngine Configuration

3.5.1 Create Provisioning Values

3.5.1.1 Dynamic VLAN Assignment

3.5.1.2 Device Template Attribute for MAC Authentication

Modify Device Template attribute using **Inbound-User-Name** if using user name for MAC authentication.

3.5.2 Create Directory Service

3.5.2.1 Create Active Directory

3.5.2.2 Internal User

This profile is used for username authentication.

3.5.2.3 Internal Device

This profile is used for MAC authentication.

3.5.3 Create Access Policy

3.5.3.1 RADIUS Authentication

Authetication Policy : Select desired authentication protocols to be included

Identity Routing : Enable Directory set and choose available configured Directory Sets (ex: Large Campus, Large-Campus and Internal Store, etc)

3.5.3.2 MAC Authentication

3.5.4 Create Authenticator

3.5.4.1 Radius Access and MAC Authentication

Authenticator IP address below is switch 4500-2 in-band management IP address.

Note: To authenticate devices that are authenticated through internal store services (ie. Internal users or devices) – sub authenticator has to be configured.

Sub Authenticator Details

Name:	John Doe - Internal User
Authenticator Attribute:	Inbound-User-Name
Attribute Value:	jdoe
Sub Authenticator Type:	Any
Access Policy:	EAPOL 4500-2

☐ Enable MAC Auth

○ Do Not Use Password
◉ Use RADIUS Shared Secret As Password
○ Use This Password _____ [Show]

Access Policy: default-radius-device

[OK] [Cancel]

Authenticator Details

Name:	4500-2
Enabled:	✓
IP Address:	10.20.0.12
Bundle:	No
Container:	default
Authenticator Type:	any
Vendor:	Nortel
Device Template:	ers-switches-nortel

RADIUS Settings
RADIUS Access Enabled: ✓
Access Policy: EAPOL 4500-2

TACACS+ Settings
TACACS+ Access Enabled: ✗

MAC Authentication
MAC Auth Access Enabled: ✓
Access Policy: Building A Room X Phone

[Edit...] [Delete]

Sub Authenticators Associated With 4500-2

Sub Authenticators are only applicable for the RADIUS protocol

Sub Authenticator	Attribute	Value	Access Policy
John Doe - Internal User	Inbound-User-Name	jdoe	EAPOL 4500-2

3.6 Key Health Indicators Configuration

To configure Key Health Indicators (KHI), you enable and disable the feature globally or only some of the KHI types. Controlling KHI types has a greater impact on loaded systems.

The main configuration actions for KHI are:

- ➢ Enabling or disabling KHI (at global or feature-level)
- ➢ Displaying statistics
- ➢ Clearing statistics/history to establish a new timeline

Note: EDM does not support KHI configuration.

3.6.1 Displaying KHI information

Display all KHI information
`Core-B:5# config sys set khi info`

3.6.2 Configuring global KHI

The Ethernet Routing Switch 8600 provides a global boot delay parameter for KHI. If the system begins collecting statistics at boot-up, the transitions that the system initially experiences do not provide an appropriate baseline of normal. To provide a valid baseline, you can configure the **boot-delay** parameter to specify how long the system takes to stabilize before KHI begins collecting statistics.

Enable KHI globally
`Core-B:5# config sys set khi khi-enable true` `Core-B:5# config sys set khi boot-delay 5`

3.6.3 Configure Management KHI

Management KHI tracks TCP connections, CLI users, and KHI log status.

Configure management
`Core-B:5# config sys set khi mgmt-khi-enable true`

3.6.4   Configure Chassis KHI

Chassis KHI displays the chassis key health indicators, such as temperature, fans, power supply, slots and CPU state.

Configure chassis
Core-B:5# **config sys set khi chassis-khi-enable true**

3.6.5   Configure Performance KHI

Performance KHI displays the performance key health indicators, such as utilization status for CPU and switch fabric.

Configure performance
Core-B:5# **config sys set khi performance-khi-enable true**

3.6.6   Configuring Protocol KHI

Protocol KHI tracks the health of the following protocols:

- ➢ OSPF
- ➢ BGP
- ➢ IST/SMLT
- ➢ PIM
- ➢ IGMP
- ➢ VLACP
- ➢ RTM and FDB table statistics

Protocol KHI also provides statistics and historical data for protocol and neighbor state transitions. It also allows for the establishment of reference timestamps and reference data to track protocol health in the network. It supports VRFs.

Every protocol has a large number of parameters that can be tracked, but only the key parameters are tracked by the KHI. Protocol information is collected and displayed on-demand, creating minimal overhead. The information is not stored in any separate database (except reference data), so that memory utilization is also minimal.

To ensure the validity of the KHI information, ensure that it is in sync with the output from the protocol show commands, and verify that the timestamps are relevant.

Configure protocol

```
Core-B:5# config sys set khi
```

3.6.7 Configuring Forwarding KHI

Forwarding KHI provides a history of the last 10 Forwarding KHI events and tracks the following on each chassis slot:

- ➢ Asic Resets
- ➢ RSP State Error Events
- ➢ RSP Stats Error Events
- ➢ F2X (F2I, F2E) Error Events

Configure forwarding
| ```
Core-B:5# config sys set khi forwarding-khi-enable true
``` |

3.6.8 Configure IP Interface KHI

IP Interface KHI provides the total configured and total operational IP interface count. It also provides a history of the last 10 IP Interface Up/Down events.

Configure IP interface
| ```
Core-B:5# config sys set khi
``` |

3.6.9 Configure Port KHI

Port KHI tracks the following information:

- ➢ Overall system statistics (unicast, multicast and broadcast Rx, Tx packets) for the preceding 2 minutes
- ➢ Port Up/Down Events
- ➢ SMLT Port Up/Down Events
- ➢ IST Port Up/Down Events
- ➢ Port Errors

Configure port
| ```
Core-B:5# config sys set khi port-khi-enable true
``` |

4. Appendix

4.1 Link Aggregation Algorithms

All the variations of MultiLink Trunking (MLT, DMLT) and 802.3ad use a hashing algorithm that distributes traffic across the physical links of the link aggregation group. Traffic is distributed on a per session basis, so packets never arrive at the destination out of order. The hashing algorithm does not necessarily have to match on both ends of the link, and therefore, MLT/DMLT is interoperable with most third-party vendors (both switches and server NICs) as is 802.3ad. As a general rule, the algorithms are based on source and destination MAC or source and destination IP address. The specific traffic distribution algorithms are as follows:

ERS 8800/8600

> For any bridged packet (except IP), distribution is based on source and destination MAC:

  ▪ MOD (DestMAC[5:0] XOR SrcMAC[5:0], number of active links)

> For bridged and routed IP or routed IPX, the distribution is based on source and destination network address:

  ▪ MOD (DestIP(X)[5:0] XOR SrcIP(X)[5:0], number of active links)

> The MLT Hashing algorithm has been modified to be based upon Layer4 fields, for IPv4 UDP/TCP traffic for R-modules only, versus Source and Destination IP addresses. This hash change provides a better distribution of traffic for many environments, with no impact on R-module performance. MLT hashing for any module type is determined at the ingress port, although the effect is seen at the egress MLT. That is, the traffic ingressing a port determines the destination is an MLT and sends the traffic to the correctly hashed MLT port, regardless of MLT port type or even a mixed port type. Therefore in a mixed chassis, which algorithm that will be used will based upon ingress port for the traffic, not the MLT configuration.

  For IPv4 TCP/UDP traffic:

  ▪ 64-bit key = (SrcPort (16 bits), DstPORT (16 bits), DstIP (LSB 16 bits), SrcIP (LSB 16 bits))

  For non TCP/UDP IPv4 traffic:

  ▪ 64-bit key = (DstIP (32 BITS), SrcIP (32 BITS))

> Multicast flow distribution over MLT is based on source-subnet and group addresses. To determine the port for a particular Source, Group (S,G) pair, the number of active ports of the MLT is used to MOD the number generated by the XOR of each byte of the masked group address with the masked source address. This feature was introduced in release 3.5. The feature is not enabled by default and must be enabled in order for IP multicast streams to be distributed.

ERS 8300

> For Non Ipv4, distribution is a function of the packet MAC Source Address (SA) and the MAC Destination Address (DA).

- Mac Hash [6:0] = XOR(SA[6:0], DA[6:0])

- Hash Index [2:0] = The Sum of MAC Hash bits (MAC_Hash[6] + MAC_Hash[5] + MAC_Hash[4] + MAC_Hash[3] + MAC_Hash[2] + MAC_Hash[1] + MAC_Hash[0]).

- The target port is chosen based on the Hash Index and number of active ports in the trunk.

➤ For IPv4 packets, distribution is a function of the IPv4 source IP address and the destination IP address.

- L3 Hash[6:0] = XOR(SIP[22:16],SIP[6:0],DIP[22:16],DIP[6:0])

- Hash Index [2:0] = The Sum of L3 Hash bits (L3_Hash[6] + L3_Hash[5] + L3_Hash[4] + L3_Hash[3] + L3_Hash[2] + L3Hash[1] + L3_Hash[0])

- The target port is chosen based on the Hash Index and number of active ports in the trunk.

➤ For any Layer 2 multicast/unknown unicast/broadcast:

- V[3:0] = (Unk_V[11:8]) XOR(Unk_V[7:4])XOR(Unk_V[3:0])

- Unk_Trunk_Dist_Index[3:0] = (Src_Port[3:0])XOR(Src_Dev[3:0])XOR ({Src_Dev[6:4]})XOR({Src_Port[5:4],2'd0})XOR(Unk_V[3:0])

➤ For any registered multicast (source port, source device, VID, VIDX):

- Reg V[11:0] = VID[11:0] XOR (VIDX[0:11] - (This is VID[11:0]XOR with VIDX[11:0] flipped)

- V[3:0] = (Reg_V[11:8])XOR (Reg_V[7:4])XOR(Reg_V[3:0])

- Reg_Trunk_Dist_Index[3:0] = (Src_Port[3:0])XOR(Src_Dev[3:0])XOR ({Src_Dev[6:4]})XOR({Src_Port[5:4],2'd0})XOR(Unk_V[3:0])

ERS 5500

➤ Two options are available to choose from for load balancing algorithms; basic MAC-based or advanced IP-based load balancing. This option is configured using the CLI.

MAC-based algorithm (Basic)

- (Least 3 bits SRC MAC XOR Least 3 bits DST MAC) MOD Active Trunk Links

IP-based algorithm (Advanced)

- (Least 3 bits SRC IP XOR Least 3 bits DST IP) MOD Active Trunk Links

- The link selection algorithm is applied to unicast traffic when both source and destination MAC or IP addresses have been learned by the system. For broadcast, unknown unicast, and multicast traffic, only a single link is selected (currently the lowest link) for packet transmission.

ERS 4500 & ERS 5600

➤ Two options are available to choose from for load balancing algorithms; basic MAC-based or advanced IP-based load balancing. This option is configured using the CLI.

➤ The numbers in the brackets of the formula represent the bits being used in the calculation. For example, DA[42:40] is the Destination MAC address bits 42 and 40.

MAC-based algorithm (Basic)

- Index =
  DA[42:40]^DA[34:32]^DA[26:24]^DA[18:16]^DA[10:8]^DA[2:0]^SA[42:40]^SA[34:32]^SA[26:24]^SA[18:16]^SA[10:8]^SA[2:0]^VLAN[10:8]^VLAN[2:0]^Ethertype[10:8]^Ethertype[2:0]^SRC_MODID[2:0]^SRC_PORT_TGID[2:0]

IP-based algorithm (Advanced)

- Index =
  SIP[122:120]^SIP[114:112]^SIP[106:104]^SIP[98:96]^SIP[90:88]^SIP[82:80]^SIP[74:72]^SIP[66:64]^SIP[58:56]^SIP[50:48]^SIP[42:40]^SIP[34:32]^SIP[26:24]^SIP[18:16]^SIP[10:8]^SIP[2:0]^TCP_UDP_SPORT[10:8]^TCP_UDP_SPORT[2:0]^DIP[122:120]^DIP[114:112]^DIP[106:104]^DIP[98:96]^DIP[90:88]^DIP[82:80]^DIP[74:72]^DIP[66:64]^DIP[58:56]^DIP[50:48]^DIP[42:40]^DIP[34:32]^DIP[26:24]^DIP[18:16]^DIP[10:8]^DIP[2:0]^TCP_UDP_DPORT[10:8]^TCP_UDP_DPORT[2:0];

➢ Variable Definition

- ^ = XOR operator

- DA = Destination MAC Address

- SA = Source MAC Address

- VLAN = VLAN tag

- Ethertype = Ethernet Type Field

- SRC_MODID is the ASIC system number identifier. The ModuleId is zero based. Each ASIC inside the switch is using one ModuleId, starting from zero to the number of ASICs used by the switch. While a switch is part of a stack the ModuleId's for the ASICs are 4*(StackUnitNumber -1) + Local ASIC Number -1

- SRC_PORT_TGID = Ingress port number or the trunk number (zero based). The port number is the ASIC port number, not the front panel port number.

- SIP = Source IP Address

- DIP = Destination IP Address

- TCP_UDP_DPORT = TCP or UDP Destination Port

➢ Index is the MLT/LAG link member index starting from zero. For broadcast, and unknown unicast, traffic is forwarded thru the DLF (Destination Lookup Failure) link based on the active trunk configuration. It is the lowest member of an active trunk group. If the advanced load balancing mode is selected for non-IP packets, load balancing falls back to MAC-Based.

➢ The port number is the ASIC port number, not the front panel port number. For the unshared ports on the ERS 4500 24/26 ports models the physical port number is the logical port number minus one. For the unshared ports on the ERS 4500 48/50 ports models, the physical port number for the first 24 ports is the logical port number minus one, for the front panel ports 25-48 the physical port number is the front panel port number (logical number) minus 25.

> ➢ In a mixed stack configuration of ERS 5500 and ERS 5600, the algorithm used will be based on the type of switch the packet ingresses.

ERS 2500

> ➢ Two options are available to choose from for load balancing algorithms; basic MAC-based or advanced IP-based load balancing. This option is configured using the CLI.

> MAC-based algorithm (Basic)

> ▪ (Least 3 bits SRC MAC XOR Least 3 bits DST MAC) MOD Active Trunk Links

> IP-based algorithm (Advanced)

> ▪ (Least 3 bits SRC IP XOR Least 3 bits DST IP) MOD Active Trunk Links

> ➢ For broadcast, and unknown unicast, traffic is forward thru the DLF (Destination Lookup Failure) link based on the active trunk configuration. In general, it is the lowest member of an active trunk group.

> ➢ For non-IP packets, load balancing falls back to MAC-Based.