Small Enterprise Design Profile (SEDP)—Network Foundation Design

This chapter describes the Small Enterprise Design Profile network design, which is a well designed and validated network architecture that is flexible, and cost effective to support a wide range of network foundational services. Key features of this network design include the following:

- High availability
- Single fabric, multi services
- Differentiated services
- Layer 2 and Layer 3 access

This chapter provides design guidance to build a highly resilient, manageable, and cost-effective small enterprise network that provides a solid foundation for seamless integration and operation of applications and network services. The network has been specifically designed to meet the challenges of the small enterprise environment.

Building Unified Small Enterprise Network Infrastructure

Cisco has years of experience developing high performance, highly available, multi service networks. The key to developing a robust design is applying a proven methodology. The following design principles were applied to develop the Small Enterprise Network Design architecture:

- Hierarchy
  - Clarifies the role of each device in each tier
  - Simpler to deploy, operate, and manage the network
  - Reduces fault domains at every tier
- Modularity
  - Enables growing the network on demand basis
- Resiliency
  - Meet users expectation of network always being available.
- Flexibility
  - Allows intelligent traffic load-sharing by using all network resources

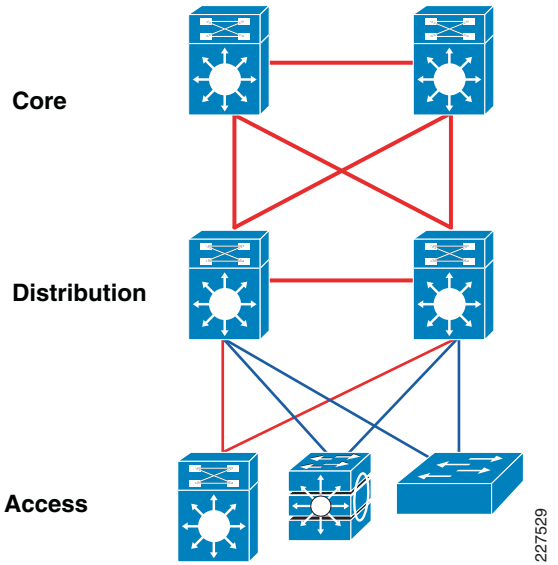
The Unified Campus network is designed to be highly available, and cost effective, while delivering capabilities necessary to enable advanced services, such as IP telephony, video, security, wireless LANs. The network design includes the following key features;

- Hierarchical design with collapsed Core
- Quality-of-service (QoS) to ensure real-time data (telephony, video) are given higher priority
- Application of resilient design principles
- Multi cast
- Routed access
- Redundancy

Hierarchical Network Design

The three-tier hierarchical model (see Figure 1) is the approach typically employed to achieve a high performance, highly available, scalable network design. This design employs the four key design principles of hierarchy, modularity, resiliency and flexibility.

Figure 1 Three-Tier Hierarchical Model



Each layer in the three-tier hierarchical model has a unique role to perform:

- Access Layer—The primary function of an access-layer is to provide network access to the end user. This layer often performs OSI Layer-2 bridge function that interconnects logical Layer-2 broadcast domains and provides isolation to groups of users, applications, and other endpoints. The access-layer interconnects to the distribution layer.
- Distribution Layer—Multi-purpose system that interfaces between access layer and core layer. Some of the key function for a distribution layer include the following:
  - Aggregate and terminate Layer-2 broadcast domains
  - Provide intelligent switching, routing, and network access policy function to access the rest of the network.
  - Redundant distribution layer switches provides high availability to the end-user and equal-cost paths to the core. It can provide differentiated services to various class-of-service applications at the edge of network.

- Core Layer—The core-layer provides high-speed, scalable, reliable and low-latency connectivity. The core layer aggregates several distribution switches that may be in different buildings. Backbone core routers are a central hub-point that provides transit function to access the internal and external network.

Table 1 lists the key functions of each layer.

Table 1 Key Functions of Hierarchical Network Layer Devices

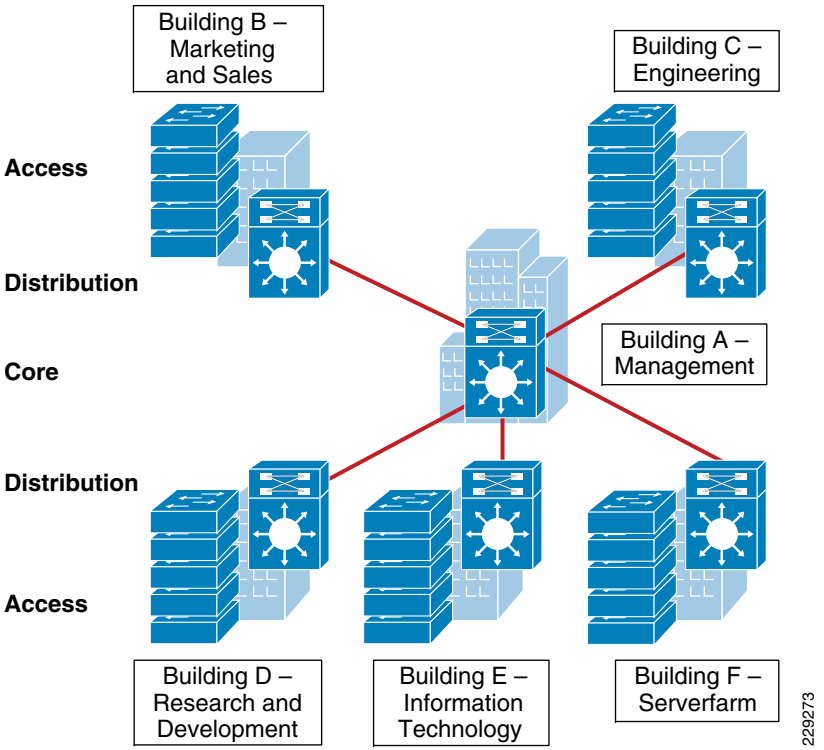
Key Function	Access	Distribution	Core
Network Transit	Rest of the network.	Internal and External network	
Intelligent Services	PoE, IEEE 802.1AD, Mobility, AutoQoS, Auto-SmartPort Macro(ASP)	Route optimization Network and System Virtualization Layer-2 Interconnect	
Forwarding Decision	Layer 2/Layer 3		Layer 3
Security Services	CISF, 802.1x, NAC, ACL etc.	CISF, ACL, Route Filter, CoPP etc.	ACL, Route Filter, CoPP etc.
QoS Services	Classification, Marking, Policer and Queueing	Classification, Marking, and Queueing	

To learn more about typical network designs, refer to the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/campover.html

Figure 2 illustrates a sample network diagram for a multi-building small enterprise network design.

Figure 2 Multi Building Small Enterprise Network Design



Collapsed Core Network Design

The three-tier hierarchical design maximizes performance, network availability, and the ability to scale the network design. Most small enterprise campus' do not grow significantly larger over time, and most small enterprise campus are small enough to be well served by a two-tier hierarchical design, where the core and distribution layers are collapsed into one layer. The primary motivation for the collapsed core design is reducing network cost, while maintaining most of the benefits of the three-tier hierarchical model.

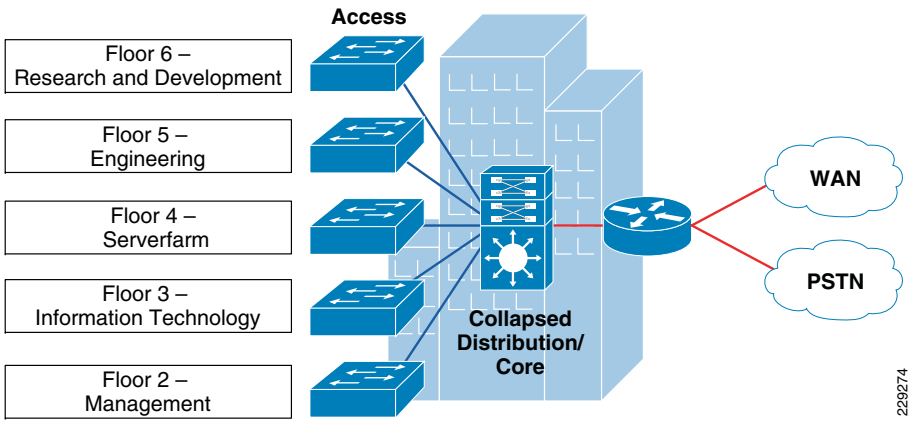
Deploying a collapsed core network results in the distribution layer and core layer functions being implemented in a single device. The collapsed core/distribution device must provide the following:

- High speed physical and logical paths connecting to the network
- Layer-2 aggregation and demarcation point
- Define routing and network access policies
- Intelligent network services—QoS, Network virtualization, etc.

Note If the main site or a remote site campus has multiple buildings, and is expected to grow over time, then implementing the three-tier hierarchical model is a better choice.

Figure 3 illustrates a sample network diagram for a single main site building.

Figure 3 Main Site—Collapsed Core Network Design

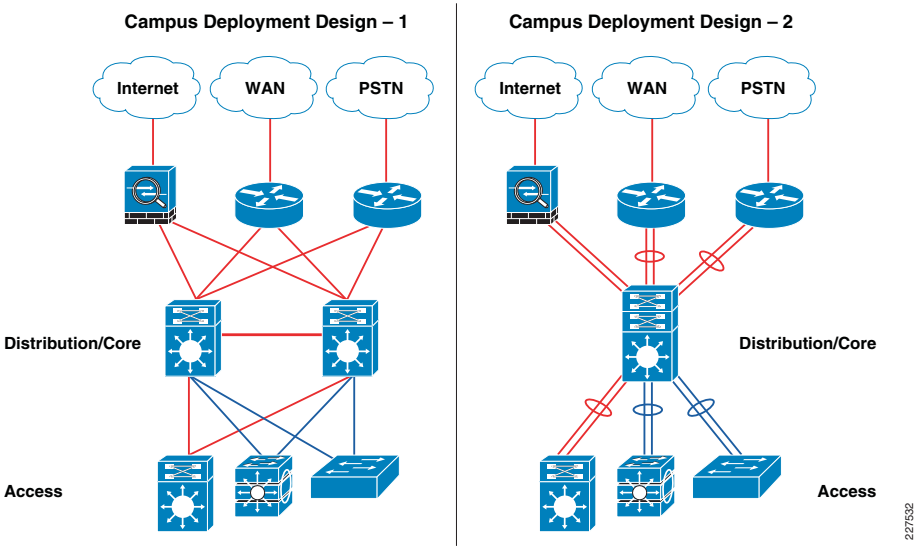


Main Site Network Design

If the main site has multiple buildings and it is expected to grow significantly over time, then implementing the three-tier hierarchical model is a good choice. For small size main sites that are unlikely to grow significantly, the collapsed core model is more cost effective. The Small Enterprise Design Profile uses the collapsed core network design in the main site.

The collapsed core network (see Figure 4) may be deployed with redundant core/distribution router, or consolidated core/distribution router.

Figure 4 Small Enterprise Design —Collapsed Core Network Models



The redundant design is more complex, because all of the core/distribution functions must be implemented on two routers in a complimentary fashion. To learn more about the redundant designs, refer to High Availability Campus Recovery Analysis Design Guide at the following URL:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/HA_recovery_DG/campusRecovery.html

The main site is designed with a consolidated core/distribution router to maximize performance, while keeping costs affordable (design 2). The consolidated collapsed core model has the following benefits:

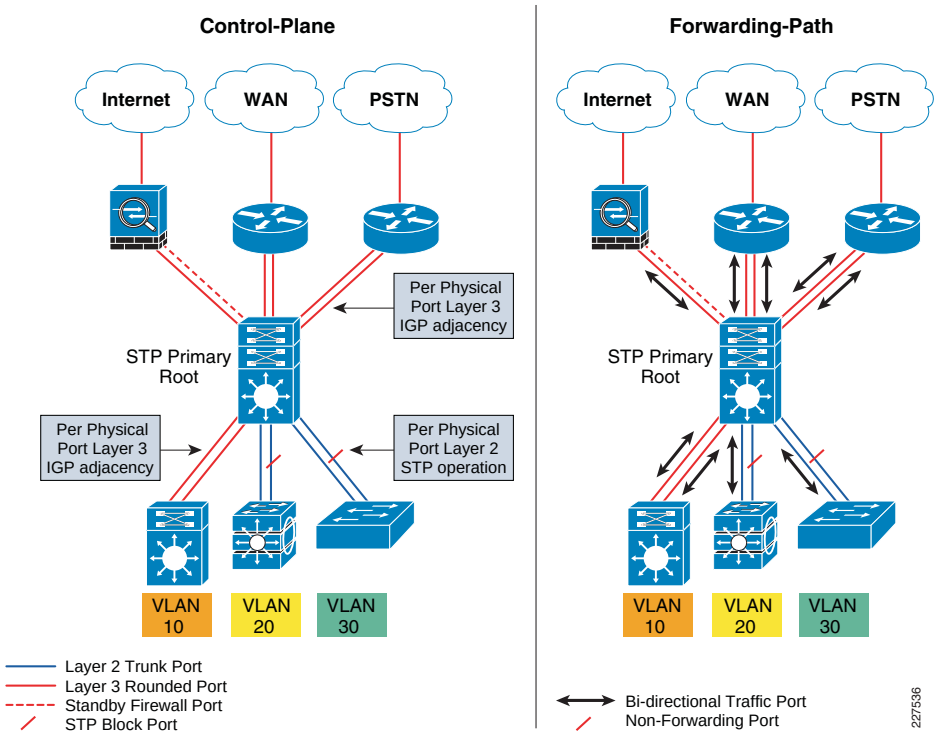
- Simplifies network protocols (eases network operations)
- Enables symmetric forwarding paths
- Delivers deterministic network recovery performance

With this design, the default behavior of Layer-2 and Layer-3 network control protocols is to create a redundant view between two systems. The core router builds a ECMP routing topology which results in symmetric forwarding paths beyond the main site.

Default Layer-2 configuration eliminates the need for FHRP, automatically eliminating the asymmetric forwarding behavior which causes unicast flooding in the network. This simplifies the network operation, since there is no need to configure or tune FHRP protocols.

The disadvantage of this Layer-2 network design is that the network is under-utilized. This is due to the way Layer-2 protocols are designed to build loop-free network topologies. When two Layer-2 bridges are directly connected, the STP protocol will block low-priority STP physical port in the forwarding table. Figure 5 illustrates the control-plane, and the forwarding-plane for this design.

Figure 5 Design Model 2 – Developing Control and Forwarding Paths



This design suffers from two challenges:

- Multiple-routing adjacencies between two Layer-3 systems. This configuration doubles the control-plane load between each of the Layer-3 devices. It also uses more system resources like CPU and memory to store redundant dynamic-routing information with different Layer-3 next-hop addresses connected to same router.
- As depicted in Figure 5, STP protocol blocks one of the physical ports in the Layer-2 network. Since this design employs point-to-point links between the collapsed core and peer devices, the solution is to tune the network to enable a single control plane, to improve forwarding efficiency and resource utilization. The recommendation is to aggregate all physical ports into a single logical channel-group. This logical aggregated Ethernet bundle interface is known as EtherChannel.

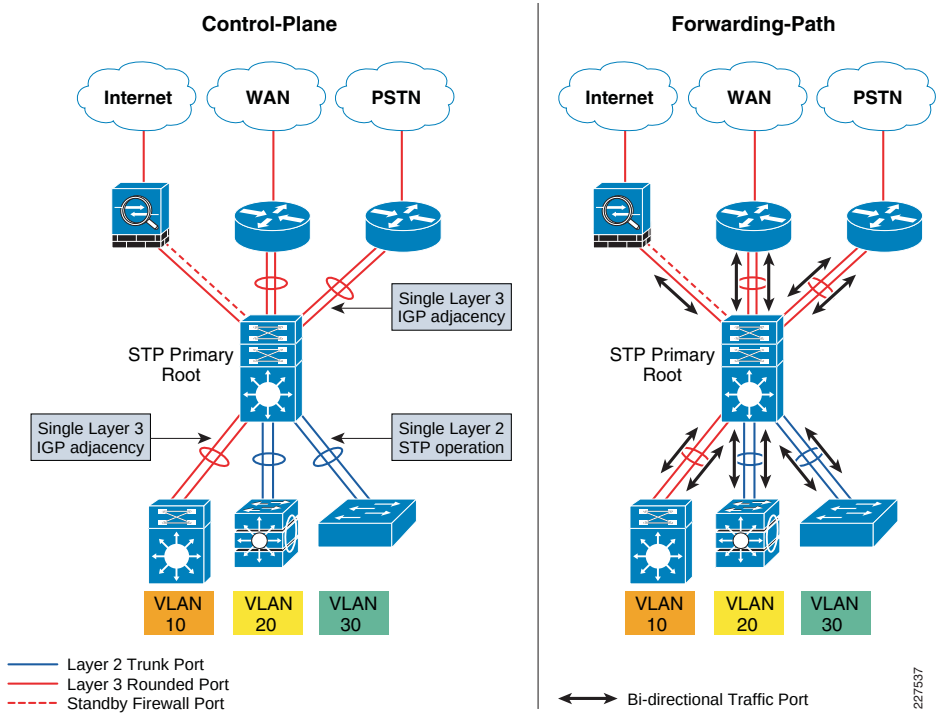
EtherChannel Fundamentals

EtherChannel provides inverse-multiplexing of multiple ports into a single logical port to a single neighbor. This technique increases bandwidth, link efficiency, and resiliency. EtherChannel technology operates on the MAC layer. Upper layer protocols require a single instance to operate over the logical interface. EtherChannel provides efficient network operation and graceful recovery to higher layer protocols during bundle port failure and restoration.

The control-plane depicted in Figure 5 builds redundant Layer- 2 or Layer-3 network information over each physical links. Each device builds common network prefix entries with a different next-hop path pointing to same next hop device. Implementing EtherChannel results in a network topology with a single destination entry for single next-hops, via the egress logical EtherChannel port. EtherChannel reduces storing redundant network entries in the database and forwarding tables, which automatically improves network convergence times and system resources utilization.

EtherChannel helps improve the overall network stability and availability. Failure of individual physical link will cause network topology recomputation, restoration, and may be rerouted. Such process requires CPU interruption that could impact the overall application performance. EtherChannel significantly simplifies the network response to a individual link failure. If an individual link in EtherChannel fails, the interface will not trigger any network topology changes. All underlying hardware changes remain transparent to higher-layer protocols, thus minimizing impact to network and application performance, and improving network convergence. Figure 6 illustrates how enabling EtherChannel in Layer-2 and Layer-3 network simplifies control-plane and forwarding-plane.

Figure 6 Design Model 2 – Optimized Control and Forwarding Paths with EtherChannel



Resilient Distributed System

The Small Enterprise Design Profile uses the Cisco Catalyst 4500 with next-generation Supervisor-6E in the consolidated core/distribution layer. It is chosen for its price performance, and the high availability features within the device. The Cisco Catalyst 4500 switch supports redundant supervisor engines and provides Stateful Switchover (SSO) and Non-Stop Forwarding (NSF) capabilities. SSO ensures the Layer-2 and Layer-3 protocol state-machines and network forwarding entries on the standby supervisor engine are maintained, and can quickly assume control-plane responsibilities and

gracefully restore the control-plane in the event of a primary supervisor failure. While the control-plane is gracefully recovering, the NSF function continues to switch traffic in hardware.

The Cisco Catalyst 6500 platform is an enterprise-class system providing integrated network services for large scale and high-speed networks. For large, multi building sites, or in situations where future scalability is important, the Catalyst 6500 is a better choice for core/distribution layer switch. The design principles remain the same when deploying a Catalyst 6500.

Main Site Access-Layer Edge Services

The access layer is the first tier or edge of the network. It is the layer where end-devices (PCs, printers, cameras, etc.) attach to the small enterprise network. It is also the layer where devices that extend the network out one more level are attached; IP phones and wireless access points (APs) are examples of devices that extend the connectivity out from the access switch. The wide variety of devices that can connect and the various services and dynamic configuration mechanisms required, make the access layer the most feature-rich layer of the small enterprise network. Figure 7 illustrates a main site network deployment with various types of trusted and untrusted endpoints.

Figure 7 Access-Layer Trust Boundary and Network Control Services

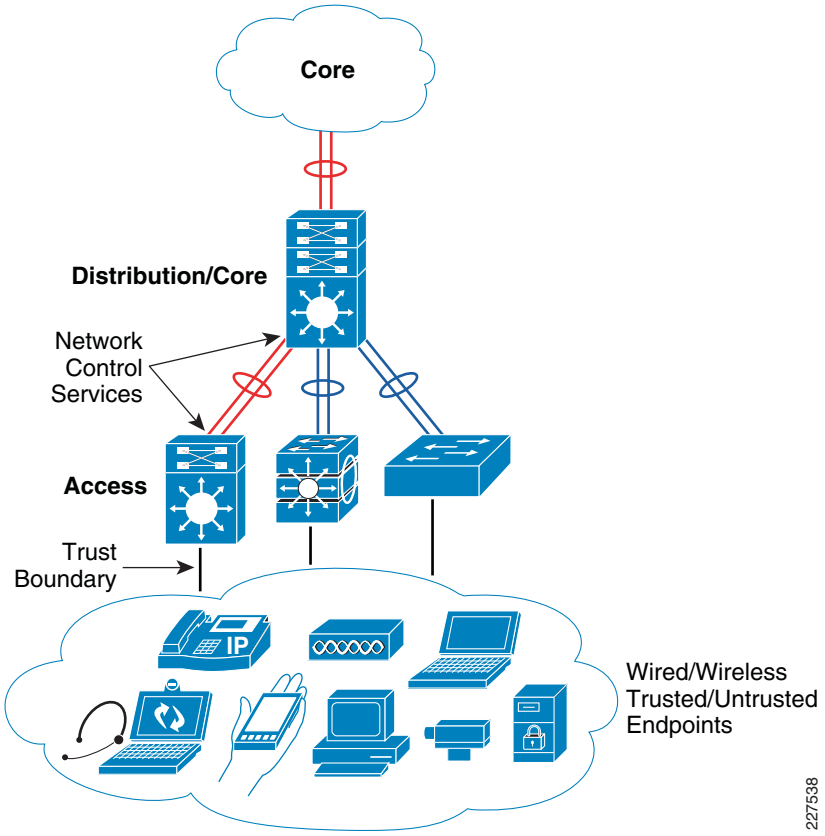


Table 2 examples of the types of services and capabilities that need to be defined and supported in the access layer of the network.

Table 2 Access-layer Services and Capabilities

Service Requirements	Service Features
Discovery and Configuration Services	802.1AF, CDP, LLDP, LLDP-MED
Integrated Security Services	IBNS (802.1X), CISF – Port-Security, DHCP Snooping, DAI and IPSG
Network Identity and Access	802.1X, MAB, Web-Auth
Application Recognition Services	QoS marking, policing, queueing, deep packet inspection NBAR
Intelligent Network Control Services	PVST+, Rapid PVST+, EIGRP, OSPF, DTP, PAgP/LACP, UDLD, FlexLink, Portfast, UplinkFast, BackboneFast, LoopGuard, BPDUGuard, Port Security, RootGuard
Energy Efficient Services	Power over Ethernet, EnergyWise, Energy efficient systems
Management Services	Auto-SmartPort Macro, Cisco Network Assistant

The access layer provides the intelligent demarcation between the network infrastructure and the computing devices that use the infrastructure. It provides network edge security, QoS, and policy trust boundary. It is the first point of negotiation between the network infrastructure and the end devices seeking access to the network.

A flexible network design, and the demand for mobility are two requirements which drive the access layer design. A flexible network design allows any legitimate device to be connected anywhere in the network (eg IP Phone, printer, video surveillance camera, digital signage, etc). Network users expect to be able to move around their devices (laptops, PDAs, printers, etc) and gain network access wherever necessary.

In order to allow devices to be moved within the network and ensure they associate with the correct network policies and services; the following access services are integrated into the small enterprise architecture:

- Ability to physically attach to the network and be associated with or negotiate the correct Layer-1 and Layer-2 network services—PoE, link speed and duplex, subnet (VLAN or SSID)
- Ability to provide device identification and, where needed, perform network access authentication
- Ability for the network to apply the desired QoS policies for the specific user, device or traffic flow (such as RTP streams)
- Ability for the network to apply the desired security policies for the specific user or device
- Ability for the network and device to determine and then register the location of the attaching device

- Ability for the device to negotiate and register the correct end station parameters (such as DHCP), as well as register for any other necessary network services (such as register for Unified Communications presence and call agent services)

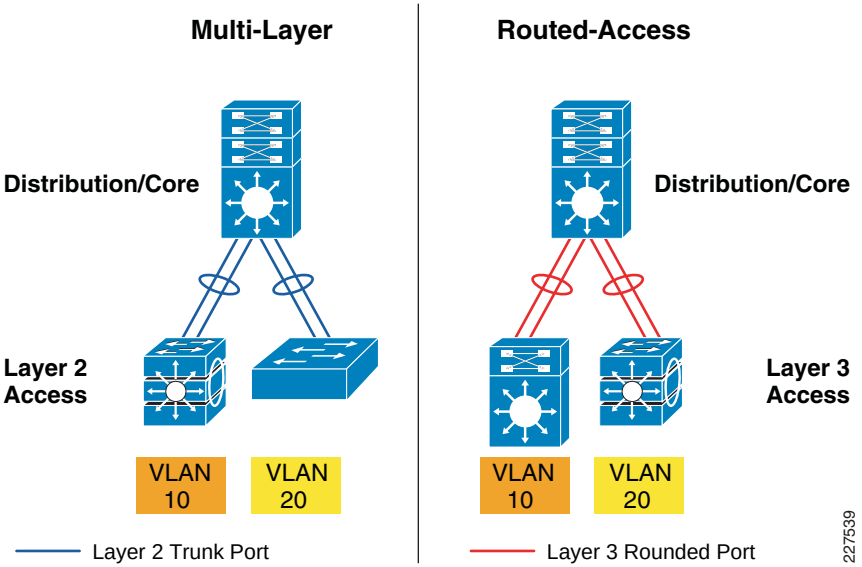
The basic steps for deploying edge access switch features are as follows:

1. Configure the baseline switching foundation
2. Protect the network infrastructure
3. Protect the end devices and their application data flows
4. Apply the necessary network policies (QoS) to provide for the required service levels.
5. Create the final template macro to allow for simplified configuration

Access-Layer Network Control Services

Properly designing the distribution block ensures the stability of the overall architecture. In the collapsed core model, the access-distribution block includes the access and distribution layers. Each of these layers has specific service and feature requirements. The network control plane choice (i.e., routing or spanning tree protocols) are central to determining how the distribution block fits within the overall architecture. The Small Enterprise Design Profile includes two designs for configuring the access-distribution block: multi-layer and routed-access. See Figure 8.

Figure 8 Access-Distribution Deployment Model



While both of these designs use the same basic physical topology and cabling plant, there are several key differences:

- Where the Layer-2 and Layer-3 boundaries exist
- How the network redundancy is implemented
- How load-balancing works

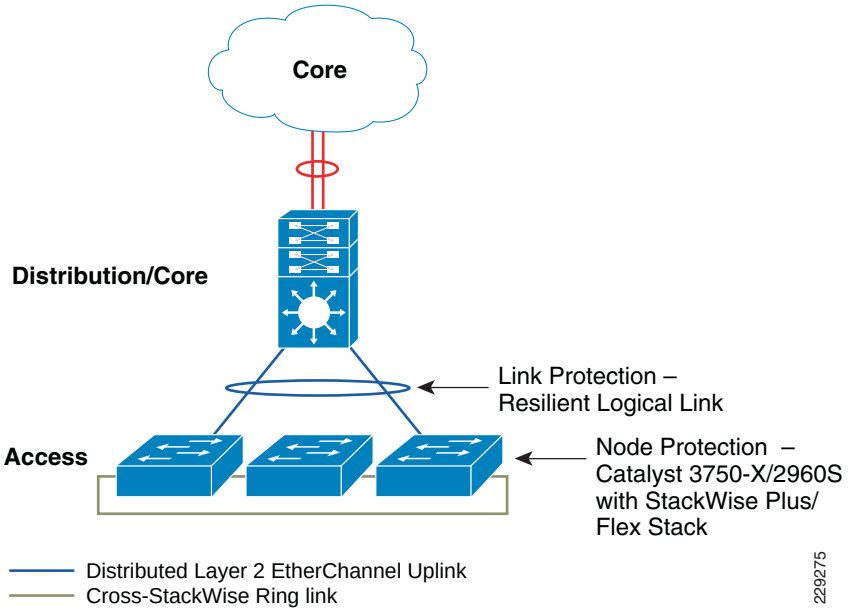
A complete configuration description of each access-distribution block model is provided in "Logical Multi-Layer Network" section on page -16 and the "Deploying Routed-Access Network" section on page -18 of this document.

Resilient Access-Layer Network and System

The access-layer provides endpoint connectivity to the rest of the network. Typical access switches like the Cisco Catalyst 2900 Series and Cisco Catalyst 3500 Series switches becomes single-point-of-failure (SPOF), if the hardware fails or if there is a software upgrade. Disrupting communication to mission critical endpoints (e.g., physical security camera) may be unacceptable.

The Small Enterprise Design Profile is designed with 2 to 4 uplink ports for each access switch, providing link-failure protection. For mission critical endpoints, this design employs the Cisco StackWise Plus and FlexStack solution in the access. It is designed to physically stack and interconnect multiple Layer-2 or Layer-3 switches using special cables. Stacking multiple switches into a logical ring creates a single unified and resilient access-layer network topology (see Figure 9). The next-generation Cisco Catalyst 2960-S FlexStack can be deployed in Layer-2 network domain and the Cisco Catalyst 3750-X StackWise Plus is deployed for routed access implementations.

Figure 9 Resilient, Scalable and Efficient Access-Layer Network Design



Main Site Data Center Network Design

The serverfarm is a central location which houses servers and storage. These resources must be available to users throughout the small enterprise network. The serverfarm may be collocated at the small enterprise network, or in a nearby site. Typically, small enterprise network are unable to afford high-speed redundant WAN links between the serverfarm and the remote sites. This makes the design vulnerable to service outage at the remote sites, in the event of WAN link failure. The Small Enterprise Design Profile recommends

which services should be placed in the centralized serverfarm, and which services should be distributed (i.e., hosted at each remote site). The key criteria to consider when making this decision include:

- Scalability—The compute capacity and storage capacity of the centralized serverfarm must be sufficient to handle peak loads.
- Network Load—The overall network design, from serverfarm to remote site must have enough capacity to carry the anticipated traffic (data and control traffic) to ensure good application performance.
- Redundancy—A WAN link failure will result in service outage between serverfarm and the remote site. This can impact network data services and can expose security issue.
- Synchronization—Some applications which are hosted locally will require content synchronization with a centralized server.This can often be scheduled for off hours, to avoid adding traffic to WAN links during normal working hours.

Table 3 provides a sample list of centralized and distributed servers.

Table 3 Sample List of Centralized and Distribution Servers

Data Center Model	Server Function	Deployment Location
Centralized	Database server(i.e., Oracle, Sybase, etc)	Main Site Data Center
	Cisco Unified Call Manager	
	Cisco Presence Server	
	Cisco Digital Media Manager (DMM)	
	E-mail Messaging Server	
Distributed	Hosted services – Web, FTP, DHCP, DNS, NTP	Main and Remote Site Data Center
	Access-Control – Cisco Access Control Server	
	Cisco Video Surveillance Operation Management	
	Media Storage Server	

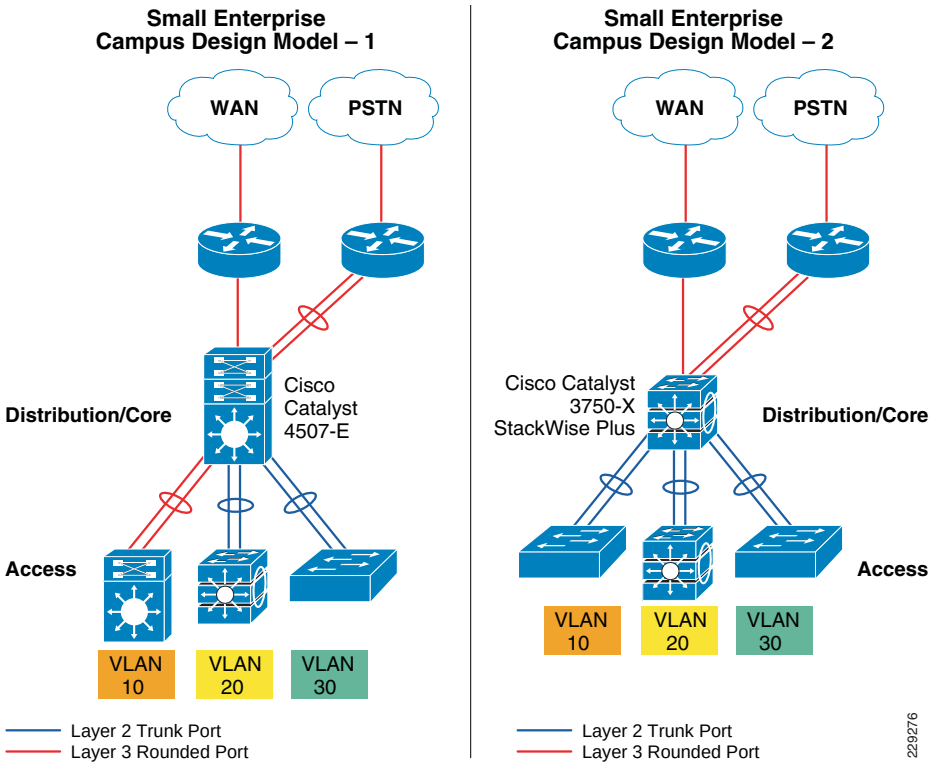
Remote Site Network Design

The Small Enterprise Design Profile includes two remote site designs. One for larger remote site, and another for medium to smaller remote sites. The typical remote site is a single building with a limited population which makes the collapsed core network design a suitable choice.

Remote Site Collapsed Core Network Design

The key criteria to consider when designing a remote site network are the network size, bandwidth capacity and high-availability requirements. The Small Enterprise Design Profile includes two models: one for smaller remote site and another for larger remote site. Both designs offer high capacity, performance, and availability. Figure 10 illustrates the two remote site network design models.

Figure 10 Remote Site—Collapsed Core Network Models



Design Model-1 is for a larger remote site. The network design is the same as the main site network design, with the same performance capabilities, scalability options, and high availability features.

Design Model-2 is for a medium to small remote site. The primary difference is the use of the Cisco Catalyst 3750-X Stack Wise Plus switch in the collapsed core/distribution layer. The 3750-X Stack Wise Plus deploys up to nine 3750-X switches in a ring topology as a single virtual switch. Each chassis replicates the control functions, and provides packet forwarding. If the master switch fails, another switch will quickly assume the control plane 'master' function. This results in a cost effective, high performance, scalable solution, with built in resiliency.

- Performance—Provides wire-rate network connection to access switches
- Scalable—May deploy up to 9 switches in a stack to aggregate a reasonable number of access switches
- High Availability—Stack provides a virtual switch, with distributed control plane, delivering subsecond system failure recovery times

The Cisco 3750-X StackWise Plus delivers high performance routing and switching capability and robust IOS feature support. The control-plane and forwarding paths functions for the Cisco 3750-X StackWise Plus in the collapsed core network design remain the same. However, the switching architecture of the Cisco 3750-X StackWise differs significantly from the high-end distributed and modular switch platforms like Cisco Catalyst 4500-E and 6500-E Series switches.

For more information about the Cisco 3750-X StackWise-Plus architecture, refer to the following URL:

http://www.cisco.com/en/US/partner/prod/collateral/switches/ps5718/ps5023/prod_white_paper09186a00801b096a_ps7077_Products_White_Paper.html

Remote Site Access-Layer Design

The access-layer network designate the remote site is the same as at the main site. The same devices are available, and the same design choices may be deployed to achieve a high performance, secure and resilient access layer. To simplify the overall system design, and network operations, it is recommended to use consistent design and platform selections in the access-layer role, at the main and remote sites. This will allow a common configuration template and simplify operations and troubleshooting procedures.

Deploying Network Foundation Services

The two-tier hierarchical design delivers a reliable and resilient, scalable, and manageable foundation network design. This subsection provides design and deployment guidelines for the small enterprise campus core layer, and access-distribution block.

The access-distribution block, as described in the "Main Site Data Center Network Design" section on page -6, uses a combination of Layer-2 and Layer-3 switching to provide a balance of policy and access controls, availability, and flexibility in subnet allocation and VLAN usage. Deployment guidelines are provided to implement multi-layer, and routed access designs in the access-distribution block.

Implementing EtherChannel in Network

Etherchannel is used throughout the network design, and the implementation guidelines are the same for multi-layer, and routed-access models, and in the WAN edge design. As recommended in the "EtherChannel Fundamentals" section on page -4, there should be single logical point-to-point EtherChannel deployed between collapsed core and access-layer. The EtherChannel configuration on each end of the link in the access-distribution block must be consistent to prevent a link bundling problem. EtherChannels use link bundling protocols to dynamically bundle physical interfaces into a logical interface.

The following are the benefits of building EtherChannel in dynamic mode:

- Ensure link aggregation parameters consistency and compatibility between switches.
- Ensure compliance with aggregation requirements.
- Dynamically react to runtime changes and failures on local and remote Etherchannel systems
- Detect and remove unidirectional links and multi-drop connections from the Etherchannel bundle.

EtherChannels can be deployed in dynamic or static modes. Both EtherChannel modes can coexist in a single system; however, the protocols (PagP, LACP) can not interoperate with each other.

- Cisco proprietary link aggregation—Cisco's implementation of Port Aggregation group Protocol (PAgP) is supported on all the Cisco Catalyst platforms. The PAgP protocol is not supported when the Cisco Catalyst 2960-S or 3750-X Series switches are deployed in FlexStack or StackWise mode. The PAgP protocol can operate in the different channel-group modes shown in Table 4 to initialize link bundling process.

Table 4 Cisco's Proprietary PAgP Channel-Group Mode

	Distribution	Access-switch and WAN Edge	EtherChannel State
channel-group mode	auto	auto	Non-Operational
	desirable (recommended)	desirable (recommended)	Operational State
	desirable	auto	
	auto	desirable	

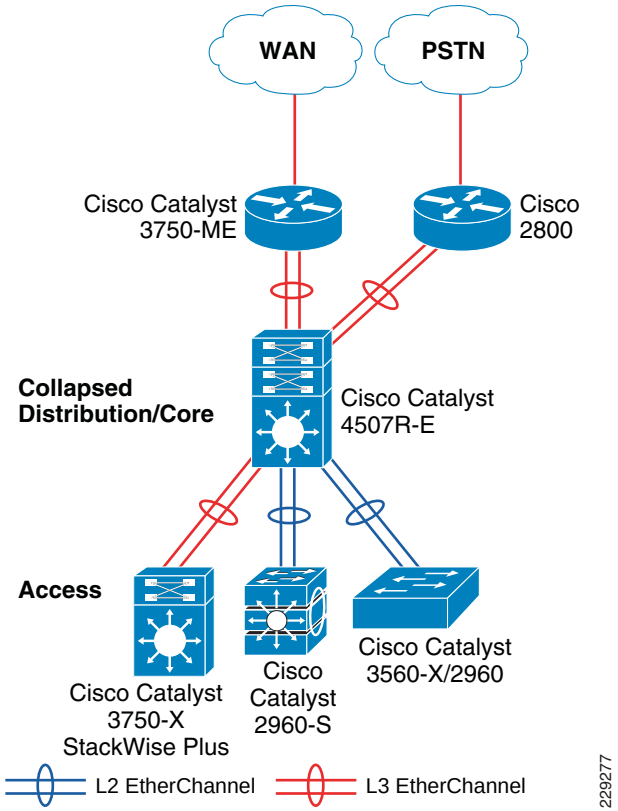
- IEEE 802.3ad link aggregation—Link Aggregation Control Protocol (LACP) is based on IEEE 802.3ad specification to operate in vendor-independent network environment. LACP link bundling protocol is developed with same goal as Cisco's PAgP. Cisco Catalyst switches in FlexStack and StackWise Plus mode must use LACP to dynamically bundle. LACP can operate in the following different channel-group modes to initialize the link bundling process. See Table 5.

Table 5 IEEE 802.3ad LACP channel-group Mode

	Distribution	Access-switch and WAN Edge	EtherChannel State
channel-group mode	passive	passive	Non-Operational
	active (recommended)	active (recommended)	Operational State
	active	passive	
	passive	active	

- Static Mode—Each system statically bundles selected physical ports into a logical port-channel. In static mode, Etherchannel consistency check is not performed between two switches, which may lead to network protocol instability or network outage due to mis-configuration. This mode is not recommended and should only be considered when EtherChannel is required but side of the link does not support PAgP or LACP link aggregation protocol.

Figure 11 Implementing EtherChannel in Main Site Network



The following sample configuration shows how to build Layer 2 and Layer 3 EtherChannel configuration and bundling physical ports into appropriate logical EtherChannel-group:

```
cr24-4507-1#config t
Enter configuration commands, one per line.  End with CNTL/Z.
cr24-4507-1(config)#interface Port-channel1
cr24-4507-1(config-if)# description Connected to cr24-3750ME-1
cr24-4507-1(config-if)#
cr24-4507-1(config-if)#interface Port-channel11
cr24-4507-1(config-if)# description Connected to cr24-2960-1
cr24-4507-1(config-if)# switchport
cr24-4507-1(config-if)#
cr24-4507-1(config-if)#interface Port-channel16
cr24-4507-1(config-if)# description Connected to cr25-3750s-1
cr24-4507-1(config-if)# switchport
cr24-4507-1(config-if)#
cr24-4507-1(config-if)#interface range Gig 3/3 , Gig 4/3
cr24-4507-1(config-if-range)# description Connected to cr24-3750ME-1
cr24-4507-1(config-if-range)# channel-protocol pagp
cr24-4507-1(config-if-range)# channel-group 1 mode desirable
cr24-4507-1(config-if-range)#
cr24-4507-1(config-if-range)#interface range Gig 1/1 , Gig 2/1
cr24-4507-1(config-if-range)# description Connected to cr24-2960-1
cr24-4507-1(config-if-range)# channel-protocol pagp
cr24-4507-1(config-if-range)# channel-group 11 mode desirable
cr24-4507-1(config-if-range)#
cr24-4507-1(config-if-range)#interface range Gig 1/6 , Gig 2/6
cr24-4507-1(config-if-range)#description Connected to cr26-3750s-1
cr24-4507-1(config-if-range)# channel-protocol lacp
cr24-4507-1(config-if-range)# channel-group 16 mode active
```

Enabling EtherChannel on each switch endpoint will automatically form a logical connection and can be verified using following CLI command:

```
cr24-4507-1#show etherchannel summary | inc Po
Group  Port-channel  Protocol    Ports
1      Po1(RU)         PAgP      Gi3/3(P)    Gi4/3(P)
11     Po11(SU)        PAgP      Gi1/1(P)    Gi2/1(P)
16     Po16(SU)        LACP      Gi1/6(P)    Gi2/6(P)
```

EtherChannel Load Balancing

EtherChannel load-sharing is based on a polymorphic algorithm. On per protocol basis, load sharing is done based on source XOR destination address or port from Layer 2 to 4 header and ports. For higher granularity and optimal utilization of each member-link port, an EtherChannel can intelligently load-share egress traffic using different algorithms. EtherChannel load balancing method support varies on Cisco Catalyst platforms. Table 6 summarizes the currently supported EtherChannel load-balancing methods.

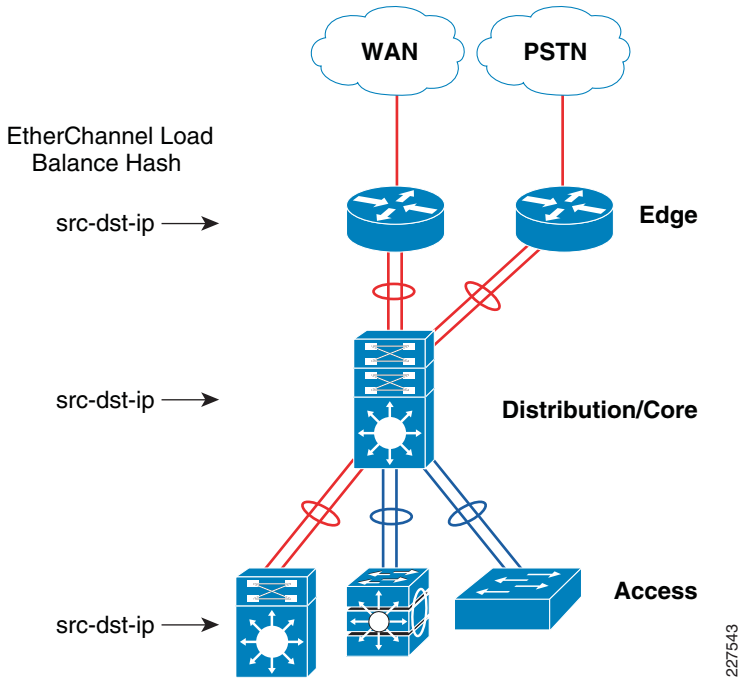
Table 6 EtherChannel Load Balancing Support Matrix

Packet Type	Classification Layer	Load Balancing Mechanic	Supported Cisco Catalyst Platform
Non-IP	Layer 2	src-dst-mac	29xx, 35xx, 3750, 4500
IP		src-mac	
		dst-mac	
		src-dst-mac	
IP	Layer 3	src-ip	
		dst-ip	
		src-dst-ip	
IP	Layer 4	src-port	4500
		dst-port	
		src-dst-port	

EtherChannel load-balancing mechanisms function on a per-system basis. By default, EtherChannel will use the hash computation algorithm. The network administrator can globally configure the load balancing mechanism. In Cisco Catalyst platforms, EtherChannel load balancing is performed in hardware and it cannot perform per-packet-based load balancing among different member links within EtherChannel. Bandwidth utilization of each member-link may not be equal in default load balancing mode. The Ether Channel load balancing method should be changed to source and destination IP address-based throughout the main and remote site network for the following reasons:

- One cannot optimize load balancing using hash tuning in a general network deployment model. This is due to variations in application deployment and usage patterns.
- EtherChannel does not take into account the bandwidth of each flow. Instead, it relies on the statistical probability that the load is equally distributed across the links of the port-channel group, given a large number of flows of relatively equal bandwidths. However, this may not always be true. Tuning the load-balancing to source-and-destination IP address allows for statistically-superior load-distribution. When loads are balanced in this manner, packets belonging to a single flow will retain their packet order. See Figure 12.

Figure 12 EtherChannel Load-Balance Method



The following output provides sample configuration guideline for changing the default port-channel load-balance setting to source-destination-ip based. Aside from Layer-2 or Layer-3 EtherChannel mode, similar configuration must be applied on each system in the access-distribution block and WAN edge.

```
cr24-4507-1(config)#port-channel load-balance src-dst-ip
```

```
cr24-4507-1#show etherchannel load-balance
EtherChannel Load-Balancing Configuration:
        src-dst-ip
EtherChannel Load-Balancing Addresses Used Per-Protocol:
Non-IP: Source XOR Destination MAC address
  IPv4: Source XOR Destination IP address
  IPv6: Source XOR Destination IP address
```

The following additional EtherChannel design and configuration must be taken into consideration for an optimal EtherChannel design:

- Enable single EtherChannel between access-layer and distribution system. Enabling more than a single Ether Channel in a collapsed core network design imposes the same limitations as discussed in non-EtherChannel scenario in Figure 5.
- For optimal load sharing and hashing computation, it is recommended to bundle the number of physical ports in powers of 2 (i.e., 2, 4, and 8).
- EtherChannel is a logical interface in Cisco Catalyst platform. EtherChannel scalability in collapsed core and distribution must be taken into account. The Cisco Catalyst 4500 can support up to 64 EtherChannels, whereas the Cisco Catalyst 3750 StackWise can support up to 48 EtherChannels per-system.

Deploying Core Network Layer

This section provides implementation and best practice guidelines for deploying the core-layer in both the main and remote enterprise sites. Proper design of the core network layer ensures reachability, transparency and availability. This section focuses on building a unicast routing topology.

Routing Protocol

Enabling routing in the small enterprise network is a simple task. However, the network physical layout must be carefully planned and designed to ensure flexible, stable and efficient routing. Developing a hierarchical network addressing scheme enables a stable, efficient and scalable design.

- Hierarchical network addressing—Structured IP network addressing in small enterprise LAN/WAN network is a must to make network scalable, stable.
- Routing protocol—Cisco IOS supports wide range of Interior Gateway Protocol (IGP). It is recommended to deploy a single choice of routing protocol across the network infrastructure. This solution guide does not recommended any particular IGP to deploy in the small enterprise network architecture as it significantly varies based on different network infrastructure. However it will provide some key points to be considered when selecting unicast routing protocol.
- Hierarchical routing domain—Routing protocols must be designed in a hierarchical model that allows network to scale and operate with greater stability. Building routing boundaries and summarizing the network addresses minimizes topology size and synchronization procedure, which improves the overall network resource utilization and reconvergence.

Routing Protocol Selection Criteria

- Efficient address allocation—Hierarchical addressing enables efficient use of address space, since groups are contiguous.
- Improves routing efficiency—Using contiguous ip addresses enables efficient route summarization. Route summarization simplifies the routing database, and computations during topology changes. This reduces the network bandwidth used by the routing protocol, and improves routing protocol performance by reducing network convergence time.
- Improves system performance—Hierarchical, contiguous ip addressing reduces router memory usage by eliminating dis-contiguous and non-summarized route entries. It saves on CPU cycles needed to compute the routing database during topology changes. This contributes to a more stable routing network, and simplifies the task of network operations and management.

Cisco IOS supports many Interior Gateway Protocols (IGP), including EIGRP and OSPF, either of which are suitable for large network deployments. While OSPF is capable of greater scale, it is also more complex, and hence more difficult to configure, operate and manage. The Small Enterprise Design Profile is designed and validated using EIGRP, since it is a stable, high performance, efficient protocol, which is simple to implement and manage. The same design principles apply whether using EIGRP or OSPF.

Table 7 lists some of the EIGRP and OSPF side-by-side feature comparison information.

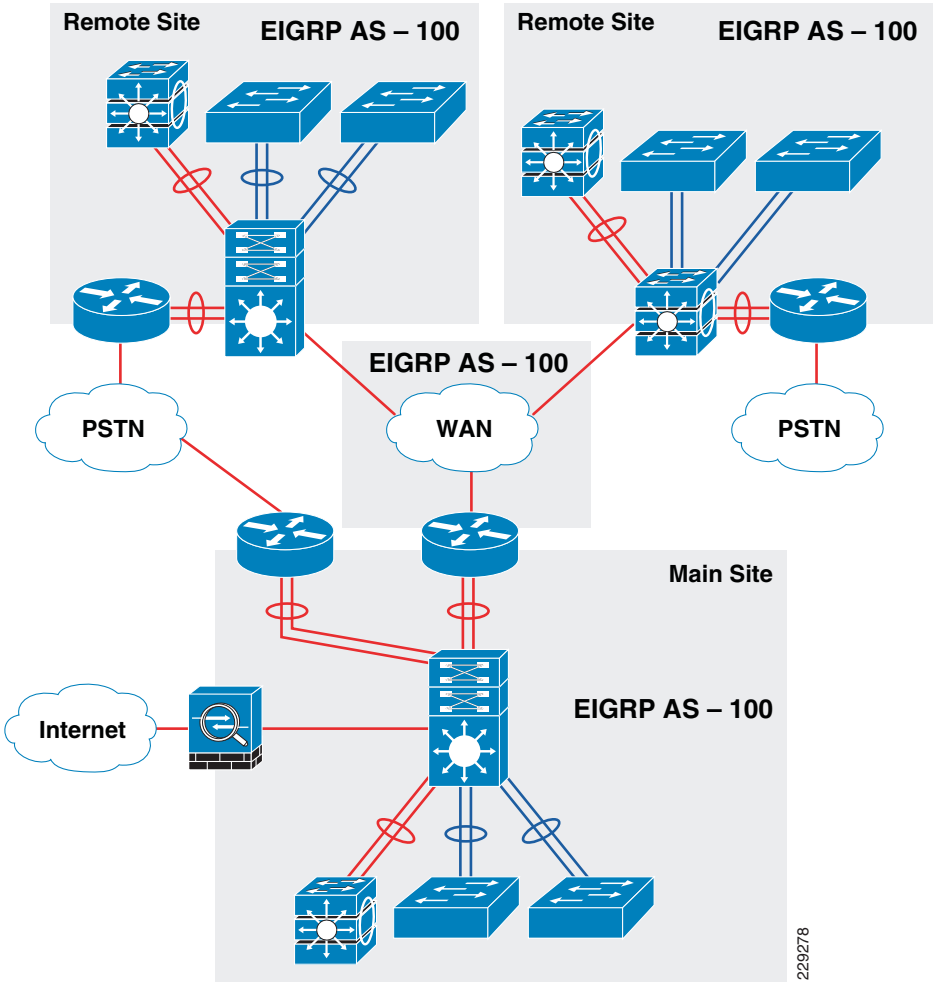
Table 7 EIGRP and OSPF feature comparison chart

Feature	EIGRP	OSPF
Classless Routing	Both routing protocols support classless routing that allows to partition networks into VLSM.	
Loop Prevention	Built-in mechanic to prevent routing loop in network.	
Robust metric	Aggregated link Bandwidth + Delay	Aggregated link Bandwidth
Efficient routing	Partial update	
Multi-access routing adjacency	Full-mesh	Hub-n-spoke
Hierarchical Routing	No. All routers considered in backbone. Non-backbone or routers in non-transit path can be deployed in Stub role.	OSPF area is divided in multiple routing domains. Backbone area maintains complete summarized network topology; non-backbone area can be transit or non-transit OSPF routers.
Network convergence	Both routing protocol offers rapid network recovery during link failure.	
Graceful-Restart Support	Yes	Yes. Cisco and IETF based
Route Summarization	Flexibility to manual summarized on any routing node.	Can only be performed on ABR or ASBR
Load-Balancing	Support equal and un-equal cost load balancing	Equal-cost path only.
Standard	Cisco proprietary	IETF standard

Designing End-to-End EIGRP Routing Domain

EIGRP is a balanced hybrid routing protocol that builds neighbor adjacency and a flat routing topology on a per-autonomous-system (AS) basis. The LAN/WAN infrastructure of Small Enterprise Design Profile should be deployed in a single EIGRP AS to prevent route redistribution, loops, and other problems that may occur due to misconfiguration. See Figure 13.

Figure 13 End-to-End EIGRP Routing Design in Small Enterprise Architecture



Implementing EIGRP Routing

The main site is the central hub in the network. Each remote site is connected to the main site over the WAN infrastructure. The main site network includes the Internet gateway and provides access to the central data-center. Since main and remote sites use the collapsed core design, the routing configuration of the core routers is the same.

The following is a sample configuration to enable EIGRP routing process at the edge of the main site collapsed core network. EIGRP is enabled in the remote sites network with the same configuration:

```
cr24-4507-1(config)#interface Loopback0
cr24-4507-1(config-if)# ip address 10.125.100.1 255.255.255.255

cr24-4507-1(config-if)#interface Port-channel1
cr24-4507-1(config-if)# description Connected to cr24-3750ME-1
cr24-4507-1(config-if)#no switchport
cr24-4507-1(config-if)# ip address 10.125.32.4 255.255.255.254

cr24-4507-1(config-if)#interface Port-channel2
cr24-4507-1(config-if)# description Connected to cr24-2851-1
cr24-4507-1(config-if)#no switchport
cr24-4507-1(config-if)# ip address 10.125.32.6 255.255.255.254
cr24-4507-1(config)#interface Vlan200
cr24-4507-1(config-if)# description Connected to cr24_ASA_Inside_Port
cr24-4507-1(config-if)# ip address 10.125.33.9 255.255.255.0

cr24-4507-1(config)#router eigrp 100
cr24-4507-1(config-router)# no auto-summary
cr24-4507-1(config-router)# eigrp router-id 10.125.100.1
cr24-4507-1(config-router)# network 10.125.0.0 0.0.255.255

cr24-4507-1#show ip eigrp neighbor port-channel 13
EIGRP-IPv4:(100) neighbors for process 100
H   Address                 Interface       Hold Uptime   SRTT   RTO  Q
Seq
                                            (sec)         (ms)       Cnt
Num
1   10.125.33.10Vl200111d00h       1   200  0  171
0   10.125.32.7Po2161d02h          1   200  0  304
2   10.125.32.5Po1141d02h          2   200  0  25038
```

EIGRP Adjacency Protection

Implementing summarization in the EIGRP routing process automatically enables EIGRP routing process on each interface that is summarized. By default, the router transmits and accept EIGRP hello messages from remote device to form an adjacency on all EIGRP enabled interfaces. This behavior needs to be modified to ensure a secure, efficient and stable routing design:

- System efficiency—There is no need to send EIGRP hellos on an interface where there is no trusted EIGRP neighbor. In a large network, sending EIGRP hello messages periodically to such interfaces consumes unnecessary CPU resource. EIGRP route processing should only be enabled on interfaces where trusted network devices are connected. All other interfaces can be suppressed in passive mode. The following configuration shows how to automatically disable EIGRP processing on all the Layer-3 interfaces and only enable on the trusted interface. This design principle must be applied on each EIGRP router, including distribution and core routers:

```
cr24-4507-1(config)#router eigrp 100
cr24-4507-1(config-router)# network 10.125.0.0 0.0.255.255
cr24-4507-1(config-router)# passive-interface default
cr24-4507-1(config-router)# no passive-interface Port-channel1
```

```
cr24-4507-1(config-router)# no passive-interface Port-channel2
cr24-4507-1(config-router)# no passive-interface Vlan200

cr24-3560r-1#show ip eigrp interface
EIGRP-IPv4:(100) interfaces for process 100

                    Xmit           Queue    Mean    Pacing Time
Multicast Pending
Interface    Peers      Un/Reliable  SRTT   Un/Reliable   Flow Timer
Routes
Vl2001       0/01   0/1     50           0
Po1 1        0/02   0/1     50           0
Po2 1        0/04   0/1     50           0

cr24-4507-1#show ip protocols | inc Passive|Vlan
  Passive Interface(s):
    Vlan1
    Vlan101
    Vlan102

    Vlan103
    Vlan104
```

- Network Security—Sending unnecessary EIGRP Hello messages opens a security vulnerability in two ways. An attacker can detect EIGRP operation and send flood of EIGRP hello messages to destabilize the network. Or an attacker could establish a "fake" EIGRP adjacency and advertise a best metric default-route into the network to black hole and compromise all critical traffic. Each EIGRP system should implement MD5 authentication, and each EIGRP neighbor should validate MD5 authentication is enabled on adjacent systems. This provides a secure method of transmitting and receiving routing information between devices in the network. Following is a sample configuration to enable EIGRP neighbor authentication using MD5:
  - Distribution

```
cr24-4507-1(config)#key chain eigrp-key
cr24-4507-1(config-keychain)# key 1
cr24-4507-1(config-keychain-key)#   key-string <password>

cr24-4507-1(config)#interface Port-channel1
cr24-4507-1(config-if)# description Connected to cr24-3750ME-1
cr24-4507-1(config-if)# ip authentication mode eigrp 100 md5
cr24-4507-1(config-if)# ip authentication key-chain eigrp 100
eigrp-key
```

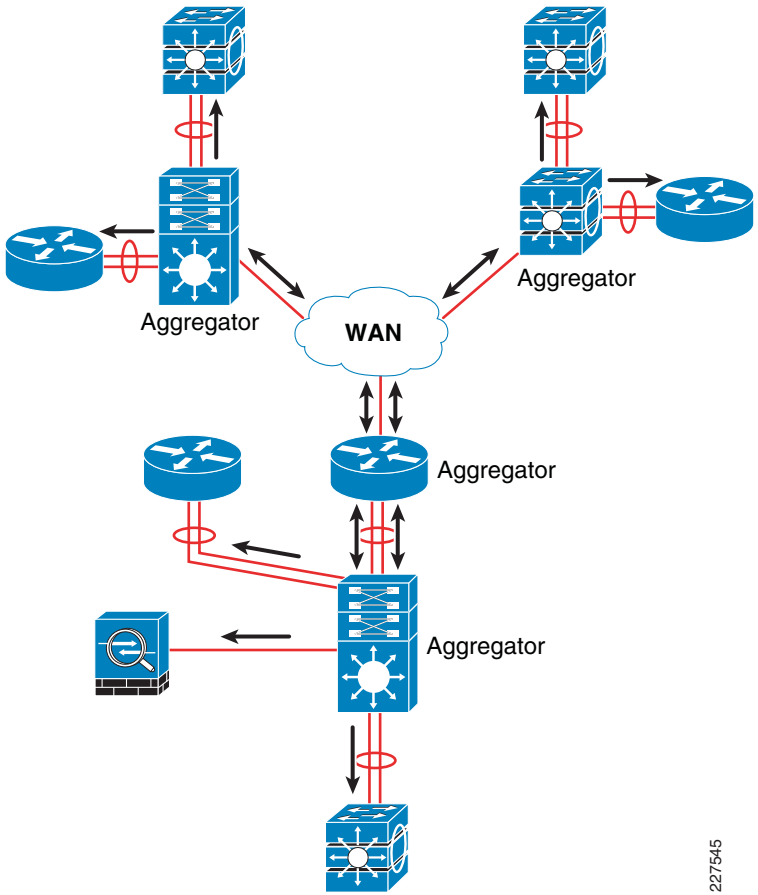
  - WAN Aggregation

```
cr24-3750ME-1(config)#key chain eigrp-key
cr24-3750ME -1(config-keychain)# key 1
cr24-3750ME -1(config-keychain-key)#   key-string <password>

cr24-3750ME -1(config)#interface Port-channel1
cr24-3750ME -1(config-if)# description Connected to cr24-4507-1
cr24-3750ME -1(config-if)# ip authentication mode eigrp 100 md5
cr24-3750ME -1(config-if)# ip authentication key-chain eigrp 100
eigrp-key
```

- System Stability—As mentioned in Table 8, EIGRP allows network administrator to summarize multiple individual and contiguous networks into a single summarized network before advertising to neighbors. Route summarization improves performance, stability, and convergence times, and it makes the network easier to manage operate and troubleshoot.

EIGRP provides the flexibility to summarize at any point in the network. Proper design requires determining which routers will serve as Aggregators, and advertise summarized network information to peers. Routers which connect multiple access devices, or connect to the WAN edge should be made Aggregators. Figure 14 provides an example small enterprise network with route aggregator devices identified with the direction of route summarization illustrated.

Figure 14 Route Aggregator and Summary Route Advertisement Direction



The following sample configuration shows EIGRP route summarization. In this example, the entire access-layer network is summarized into a single classless network and advertised to the WAN edge, the ASA firewall and the PSTN gateway:

- Distribution

```
cr24-4507-1(config)#interface Port-channel1
cr24-4507-1(config-if)# description Connected to cr24-3750ME-1
cr24-4507-1(config-if)# ip summary-address eigrp 100 10.125.0.0
255.255.0.0

cr24-4507-1(config-if)#interface Port-channel2
cr24-4507-1(config-if)# description Connected to cr24-2851-1
cr24-4507-1(config-if)# ip summary-address eigrp 100 10.125.0.0
255.255.0.0

cr24-4507-1(config-if)#interface Vlan200
cr24-4507-1(config-if)# description Connected to cr24_ASA_Inside_Port
cr24-4507-1(config-if)# ip summary-address eigrp 100 10.125.0.0
255.255.0.0

cr24-4507-1#show ip protocols | inc Address|10.125.0.0
Address Family Protocol EIGRP-IPv4:(100)
  Address Summarization:
    10.125.0.0/16 for Port-channel1, Vlan200, Port-channel2
```

- WAN Aggregation

  Verifying main site EIGRP summarized route status at WAN aggregation layer as follows:

```
cr24-3750ME-1#show ip route 10.125.0.0 255.255.0.0
Routing entry for 10.125.0.0/16
  Known via "eigrp 100", distance 90, metric 1792, type internal
  Redistributing via eigrp 100
  Last update from 10.125.32.4 on Port-channel1, 1d04h ago
  Routing Descriptor Blocks:
  * 10.125.32.4, from 10.125.32.4, 1d04h ago, via Port-channel1
      Route metric is 1792, traffic share count is 1
      Total delay is 20 microseconds, minimum bandwidth is 2000000
Kbit
      Reliability 255/255, minimum MTU 1500 bytes
      Loading 1/255, Hops 1
```

Tuning EIGRP Protocol Timers

EIGRP uses Hello messages to form adjacencies and determine if neighbors are alive. EIGRP adjacency is declared down if it fails to receive Hello messages within the Hold down timer interval. All the prefixes discovered from a dead neighbor are removed from the routing table. By default, EIGRP transmits a Hello message every 5 seconds to notify neighbors that it is still alive. The EIGRP hold-down timer gets reset each time the router receives a EIGRP Hello message. Default EIGRP adjacency hold-down timer is 15 seconds.

Lowering EIGRP hello and hold-down timer intervals improves network convergence times (i.e. time to detect and respond to an outage). For small enterprise network design it is recommended to use the default EIGRP Hello and Hold timer values for the following reasons:

- EtherChannel benefits—EIGRP operates over the Layer-3 EtherChannel. In the event of a single member-link failure condition, layer 2 will respond more quickly than the routing protocol, and switchover traffic from the impacted link to an alternate member link. EIGRP routing is not impacted by individual link member and no change in the routing table is required. Thus reducing the EIGRP timers will not result in quicker convergence, and may adversely impact system stability.
- High availability—The Cisco Catalyst 4507R-E and 3750-X Stack Wise Plus layer 3 switches support graceful-restart protocol extensions which enables a redundant module or member switch to gracefully assume the active role while maintaining adjacency with neighbors, during a active supervisor failure condition. The backup

supervisor requires sufficient time to detect a failure and initiate graceful recovery with neighbors. Implementing aggressive timers may abruptly terminate adjacency and cause network outage before a stateful switch over is accomplished. Thus, default EIGRP Hello and Hold timers are recommended on Cisco Catalyst 4507R-E and 3750-X Stackwise Plus Series Layer-3 platforms.

Deploying Multi-Layer Network

Multilayer design is one of the two access-distribution block designs included in the Small Enterprise Design Profile. This section provides implementation and best practices guidelines the multi-layer design. The deployment and configuration guidelines for the multi-layer access-distribution block are the same for both main and remote site networks.

Spanning-Tree in Multilayer Network

Spanning Tree (STP) is a Layer-2 protocol that prevents logical loops in switched networks with redundant links. The Small Enterprise Design Profile uses Etherchannel (point-to-point logical Layer-2 bundle) connection between access-layer and distribution switch which inherently simplifies the STP topology and operation. In this design, the STP operation is done on a logical port, therefore, it will be assigned automatically in forwarding state.

Over the years, the STP protocols have evolved into the following versions:

- Per-VLAN Spanning Tree Plus (PVST+)—Provides a separate 802.1D STP for each active VLAN in the network.
- IEEE 802.1w – Rapid PVST+—Provides an instance of RSTP (802.1w) per VLAN. It is easy to implement, proven in large scale networks that support up to 3000 logical ports and greatly improves network restoration time.
- IEEE 802.1s – MST—Provides up to 16 instances of RSTP (802.1w) and combines many VLANs with the same physical and logical topology into a common RSTP instance.

Following is the example configuration to enable STP protocol in multi-layer network:

Distribution

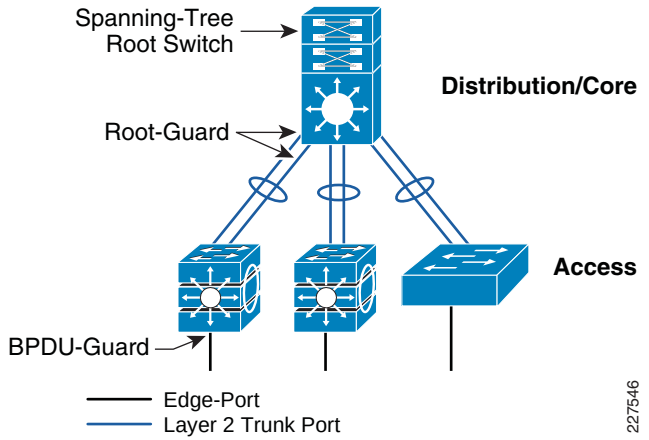
```
cr24-4507-1(config)#spanning-tree mode rapid-pvst

cr24-4507-1#show spanning-tree summary | inc mode
Switch is in rapid-pvst mode
```

Access-Layer Switch

```
cr24-2960-1(config)#spanning-tree mode rapid-pvst
```

Default STP parameters optimize the network for packet forwarding. Best practice design includes hardening STP parameters in the access and distribution switch to protect against STP misconfiguration, or malicious user by deploying spanning-tree toolkit in the access-distribution block. See Figure 15.

Figure 15 Hardening Spanning-Tree Toolkit in Multi-Layer Network



The following is the configuration deploys spanning-tree toolkit in the access-distribution block:

Distribution

```
cr24-4507-1(config)#spanning-tree vlan 1-4094 root primary
cr24-4507-1(config)#interface range Gig 1/1 - 2 , Gig 2/1 - 2
cr24-4507-1(config)#spanning-tree guard root
```

Access

```
cr26-2960S-1(config)#interface GigabitEthernet1/0/1
cr26-2960S-1(config-if)#description CONNECTED TO UNTRUSTED-PC
cr26-2960S-1(config-if)#spanning-tree bpduguard enable
```

Other STP Toolkit Consideration

When the access-distribution block multi-layer design is deployed using the recommended best practices, it automatically minimizes the need for deploying the following additional spanning-tree toolkit technologies:

- UplinkFast—Improves the network convergence time by providing direct access to the root switch link failure. UplinkFast is not necessary in this design, because there is no alternate STP path and RSTP protocol natively includes rapid recovery mechanism.
- BackBone Fast—Provides rapid convergence from indirect Layer-2 link failures in a redundant distribution switch configuration. This is feature is not necessary for the same reason as stated for UplinkFast.
- LoopGuard—Protects Layer-2 networks from loops that occur due to any malfunction that prevents normal BPDU forwarding. A STP loop is created when a blocking port in a redundant topology erroneously transitions to the forwarding state. This usually happens because one of the ports in a physically redundant topology

(not necessarily the blocking port) stopped receiving BPDUs. Because there is single point-to-point STP forwarding port in this design, enabling Loopguard does not provide any additional benefit. UDLD protocol must be implemented to prevent STP loop that may occur in the network due to network malfunction, mis-wiring, etc.

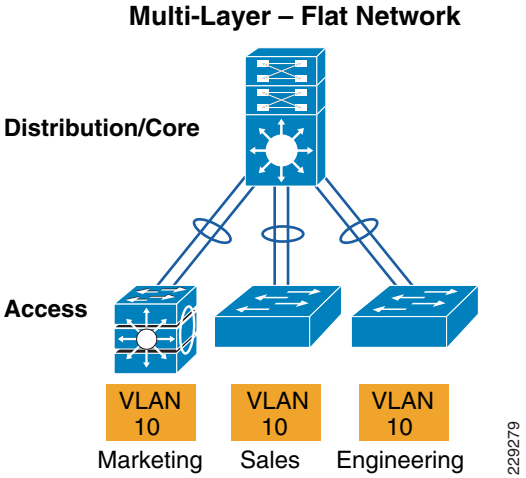
Logical Multi-Layer Network

VLAN assignment can have a significant impact on network performance and stability. There are three basic ways to assign VLANs within the access-distribution block.

Flat Logical Network Design

Spanning a single VLAN across multiple access-layer switches is much simpler with a single collapsed core-distribution device versus a design with redundant distribution devices. The flat multi-layer design has a single VLAN across multiple access devices, as shown in See Figure 16.

Figure 16 Multi-Layer Flat Network Design



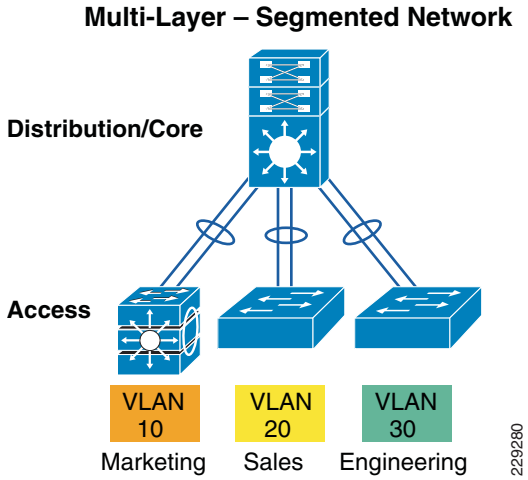
A flat multi-layer network deployment introduces the following challenges:

- Scalability—Spanning the same VLAN in different access-layer switches will create a large Layer-2 broadcast domain that dynamically discovers and populates MAC address entries for endpoints that may not need to communicate. In a large network, this may become a scalability issue (i.e. memory required to hold large CAM table).
- Performance—In a large network, spanning a large number of broadcast domains will impact the performance of all network devices in the access-distribution block, because the switch will have to process many more broadcast packets such as ARP.
- Security—The flat muli0layer design widens the fault domain which increases possible attacks to a larger number of users.The number of users is not necessarily due to the number switches spanned and applications during DoS or viruses attack.

Segmented Logical Network Design

Best practice design includes identifying meaningful groups within the user community, and assigning a unique VLAN to each group. These groups may be departments, user groups, or any other logical grouping of users. Enabling a unique VLAN for each group will segment the network and build a logical network structure. All network communication between groups will pass through the routing and forwarding policies defined at the distribution layer. See Figure 17.

Figure 17 Multi-Layer Segmented Network Design

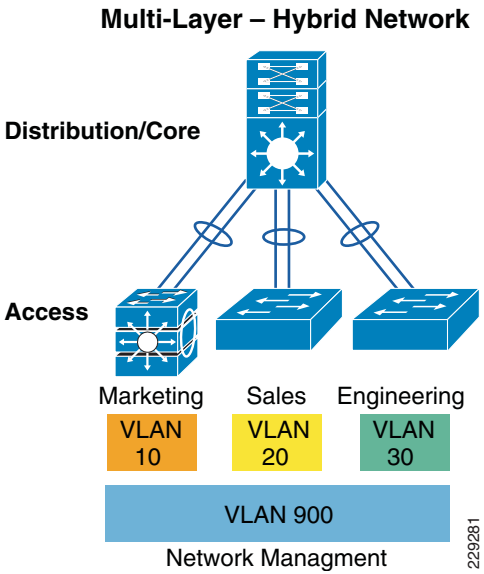


A segmented VLAN design is the solution to the challenges described in the flat network design. VLAN segmentation improves the scalability, performance, and security of the network.

Hybrid Logical Network Design

The segmented logical network design improves scalability, performance and security, and addresses the challenges of a flat network design. In real world deployments, there is usually a need for some users or applications to communicate with all users (eg system administrator). The hybrid network design is the segmented design, with the addition of a exceptional VLAN which spans the entire access-distribution block. See Figure 18.

Figure 18 Multi-Layer Hybrid Network Design



Cisco recommends the segmented VLAN network design and optionally hybrid network for centralized users or applications that requires distributed function across the access-layer network.

Following are the sample VLAN configuration steps in the access and the distribution layer switches.

Distribution

VLAN Trunking Protocol (VTP) is a Cisco proprietary Layer -messaging protocol that manages the addition, deletion, and renaming of VLANs on a network-wide basis. Cisco's VTP simplifies administration in a switched network. VTP can be configured in three modes: server, client, and transparent. Set the VTP domain name and change the mode to the transparent mode as follows:

```
cr24-4507-1(config)#vtp domain campus
cr24-4507-1(config)#vtp mode transparent

cr24-4507-1(config)#vlan 10
cr24-4507-1(config-vlan)#name cr24-3750-Mktg-Dept
cr24-4507-1(config-vlan)#vlan 20
cr24-4507-1(config-vlan)#name cr24-3560-Sales-Dept
cr24-4507-1(config-vlan)#vlan 30
cr24-4507-1(config-vlan)#name cr24-2960-Engg-Dept
```

Access

Set VTP domain name and change the mode to the transparent mode as follows:

```
cr24-3750-1(config)#vtp domain campus
cr24-3750-1(config)#vtp mode transparent

cr24-3750-1(config)#vlan 10
cr24-3750-1(config-vlan)#name cr24-3750-Sales-Dept
```

Implementing Layer 2 Trunk

In a typical network design, a single access switch will have more than one VLAN, for example a Data VLAN and a Voice VLAN. The network connection between Distribution and Access device is a trunk. VLAN's tag their traffic to maintain separation between VLANS across the trunk. By default on Cisco Catalyst switches, the native VLAN on each layer 2 trunk port is VLAN 1, and cannot be disabled or removed from VLAN database. The native VLAN remains active on all access switches layer 2 ports.

There are two choices for encapsulating the tagged VLAN traffic on the trunk: IEEE 802.1Q or Cisco ISL. It is recommended to implement trunk encapsulation in static mode instead of negotiating mode, to improve the rapid link bring-up performance. Not all Cisco Catalyst platforms support ISL encapsulation; therefore IEEE 802.1Q is recommended, and validated in the access and distribution switches.

Enabling the Layer-2 trunk on a port-channel, automatically enables communication for all of the active VLANs between the access and distribution. This means an access-switch which has implemented, for example, VLANs 10 to 15, will receive flood traffic destined for VLANs 20 to 25, which are implemented on another access switch. RPVST +, using logical ports, operates on a per-VLAN basis to load balance traffic. In a large network, it is important to limit traffic on Layer-2 trunk ports to only the assigned VLANS, to ensure efficient and secure network performance. Allowing only assigned VLANs on a trunk port automatically filters rest.

The default native VLAN must be properly configured to avoid several security risks—Attack, worm and virus or data theft. Any malicious traffic originated in VLAN 1 will span across the access-layer network. With a VLAN-hopping attack it is possible to attack a system which does not reside in VLAN 1. Best practice to mitigate this security risk is to implement a unused and unique VLAN ID as a native VLAN on the Layer-2 trunk between the access and distribution switch. For example, configure VLAN 802 in the access-switch and in the distribution switch. Then change the default native VLAN setting in both the switches. Thereafter, VLAN 802 must not be used anywhere for any purpose in the same access-distribution block.

Following is the configuration example to implement Layer-2 trunk, filter VLAN list and configure the native-VLAN to prevent attacks on port channel interface. When the following configurations are applied on port-channel interface (i.e., Port-Channel 11), they are automatically inherited on each bundled member-link (i.e., Gig1/1 and Gig2/1):

Distribution

```
cr24-4507-1(config)#vlan 802
cr24-4507-1(config-vlan)#name Admin-Hopping-VLAN

cr24-4507-1(config)#interface Port-channel 11
cr24-4507-1(config-if)# description Connected to cr24-3750-1
cr24-4507-1(config-if)# switchport
cr24-4507-1(config-if)# switchport mode trunk
```

```
cr24-4507-1(config-if)# switchport trunk allowed vlan 101-110,900
cr24-4507-1(config-if)# switchport trunk native vlan 802

cr24-4507-1#show interface port-channel 11 trunk

Port        Mode             Encapsulation  Status        Native vlan
Po11        on               802.1q         trunking      802

Port        Vlans allowed on trunk
Po11        101-110,900

Port        Vlans allowed and active in management domain
Po11        101-110,900

Port        Vlans in spanning tree forwarding state and not pruned
Po11        101-110,900
```

Access-switch

```
cr24-3750-1(config)#vlan 802
cr24-3750-1(config-vlan)#name Admin-Hopping-VLAN

cr24-3750-1(config)#interface Port-channel 1
cr24-3750-1(config-if)# description Connected to cr24-4507-1
cr24-3750-1(config-if)# switchport
cr24-3750-1(config-if)# switchport mode trunk
cr24-3750-1(config-if)# switchport trunk allowed vlan 101-110,900
cr24-3750-1(config-if)# switchport trunk native vlan 802
```

Unidirectional Link Detection

UDLD is a Layer 2 protocol that works with the Layer 1 features to determine the physical status of a link. At Layer 1, auto-negotiation takes care of physical signaling and fault detection. UDLD performs tasks that auto-negotiation cannot perform, such as detecting the identity of neighbors and shutting down misconnected ports. When both auto-negotiation and UDLD are enabled, Layer 1 and Layer 2 detection works together to prevent physical and logical unidirectional connections and the malfunctioning of other protocols.

Copper media ports use Ethernet link pulse as a link monitoring tool and are not susceptible to unidirectional link problems. Because one-way communication is possible in fiber-optic environments, mismatched transmit/receive pairs can cause a link up/up condition even though bidirectional upper-layer protocol communication has not been established. When such physical connection errors occur, it can cause loops or traffic black holes. UDLD functions transparently on Layer-2 or Layer-3 physical ports. UDLD operates in one of two modes:

- Normal mode—If bidirectional UDLD protocol state information times out; it is assumed there is no-fault in the network, and no further action is taken. The port state for UDLD is marked as undetermined. The port behaves according to its STP state.
- Aggressive mode—If bidirectional UDLD protocol state information times out, UDLD will attempt to reestablish the state of the port, if it detects the link on the port is operational. Failure to reestablish communication with UDLD neighbor will force the port into the err-disable state. That must be manually recovered by user or the switch can be configured for auto recovery within specified interval of time.

Following is the configuration example to implement UDLD protocol:

Distribution

```
cr24-4507-1(config)#interface range Gig 1/2 , Gig 2/2
cr24-4507-1(config-int)#udld port

cr24-4507-1#show udld neighbor
PortDevice Name   Device ID     Port IDNeighbor State
------     ----------------    ------------ -------- --------------
Gi1/2    FOC1318Y06V    1            Gi1/0/49Bidirectional
Gi2/2    FOC1318Y06J     1            Gi3/0/49Bidirectional
```

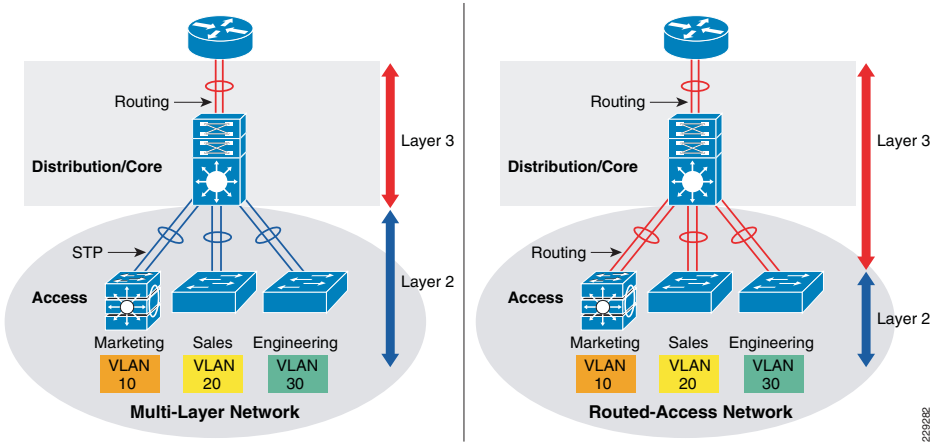
Access

```
cr26-2975-1(config)#interface Gig 1/0/49 , Gig 3/0/49
cr26-2975-1(config-if)#description Connected to cr24-4507-1
cr26-2975-1(config-if)#udld port

cr26-2975-1#show udld neighbor
PortDevice Name   Device ID     Port IDNeighbor State
------     ----------------    ------------ -------- --------------
Gi1/0/49 FOX1216G8LT     1            Gi1/2      Bidirectional
Gi3/0/49 FOX1216G8LT     1            Gi2/2      Bidirectional
```

Deploying Routed-Access Network

This section provides implementation and best practices guidelines to deploy routed-access in the access-distribution block. The routed access design moves the boundary between Layer 2 and Layer 3 from the distribution layer to the access layer as seen in Figure 19.

Figure 19 Control Function in Multi-Layer and Routed-Access Network Design

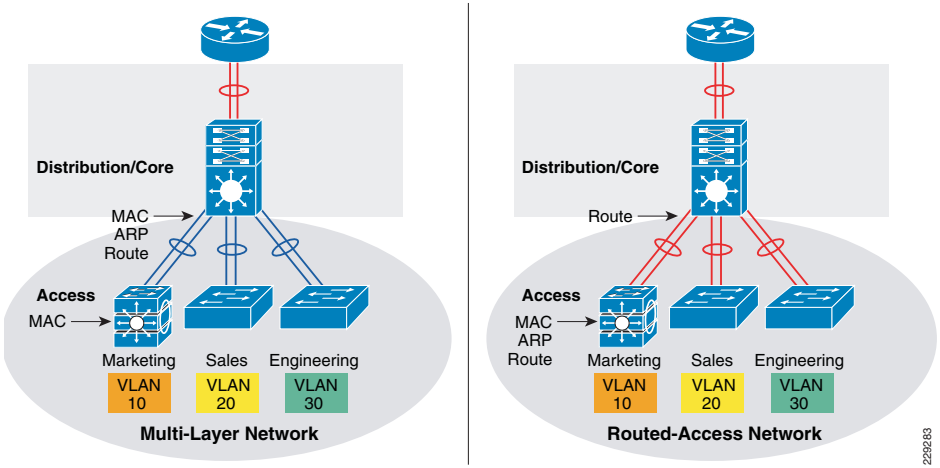


Routing in the access-layer simplifies configuration, optimizes distribution performance, and improves end-to-end troubleshooting tools. Implementing routing in the access-layer replaces Layer-2 trunk configuration with single point-to-point Layer-3 interface in distribution layer. Placing Layer-3 function one tier down on access-switches, changes the multilayer network topology and forwarding path. Implementing Layer-3 function in the access-switch does not require a physical or logical link reconfiguration; the same EtherChannel in access-distribution block can be used.

At the network edge, Layer-3 access-switches provides an IP gateway and become the Layer-2 demarcation point to locally connected endpoints that could be logically segmented into multiple VLANs. Following are the benefits of implementing routed-access in the access-distribution block:

- Eliminates the need to implement STP and the STP toolkit in the distribution layer. As a best practice, STP toolkit must be hardened at the access-layer.
- Shrinks the Layer-2 fault domain, which minimizes the number of endpoints affected by a DoS/DDoS attack.
- Improves Layer-3 uplink bandwidth efficiency by suppressing Layer-2 broadcasts at the access edge port.
- Improves performance by reducing resource utilization in collapsed core-distribution layer. In a large multilayer network, the aggregation layer may consume more CPU cycles due to the large number of MAC and ARP discovery and processing and storing required for each end-station. Routed-access reduces the load of this Layer-2 processing and storage in the distribution layer, by moving the load to layer-3 access-switches. Figure 20 illustrates where Layer-2 and Layer-3 forwarding entry processing and storage takes place when access-distribution block is implemented as multi-layer versus routed-access network.

Figure 20 Forwarding Entry Development in Multi-tier Network



While the routed access design is appropriate for many small enterprise networks it is not suitable for all environments. Routed access does not allow a VLAN to span multiple access switches. Refer to following URL for detailed design guidance for the routed access distribution block design:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/routed-ex.html

Implementing EIGRP Routing in Access-Distribution Block

The Small Enterprise Design Profile uses EIGRP routing protocol, and all the devices in the LAN and WAN sub-networks are deployed in a single AS. This subsection focuses on implementing EIGRP in the access-distribution block. All the deployment and configuration guidelines in this section are the same for deploying in the main or remote site network.

Following is the example configuration to enable basic EIGRP routing in the distribution layer and in the access layer:

Distribution

```
cr24-4507-1(config)#interface Port-channel13
cr24-4507-1(config-if)# description Connected to cr24-3560r-1
cr24-4507-1(config-if)#no switchport
cr24-4507-1(config-if)# ip address 10.125.32.0 255.255.255.254

cr24-4507-1(config)#router eigrp 100
cr24-4507-1(config-router)# no auto-summary
cr24-4507-1(config-router)# eigrp router-id 10.125.100.1
cr24-4507-1(config-router)# network 10.125.0.0 0.0.255.255

cr24-4507-1#show ip eigrp neighbor port-channel 13
EIGRP-IPv4:(100) neighbors for process 100
```

```
H   Address                 Interface       Hold Uptime    SRTT    RTO    Q
Seq
                                             (sec)          (ms)
Cnt Num
3   10.125.32.1             Po13              14 00:02:14     2
200   0   385
```

Access

```
cr24-3560r-1(config)#interface Loopback0
cr24-3560r-1(config-if)# ip address 10.125.100.4 255.255.255.255
cr24-3560r-1(config-if)#
cr24-3560r-1(config-if)#interface Port-channel1
cr24-3560r-1(config-if)# description Connected to cr24-4507-1
cr24-3560r-1(config-if)# no switchport
cr24-3560r-1(config-if)# ip address 10.125.32.1 255.255.255.254


cr24-3560r-1(config)#ip routing

cr24-3560r-1(config)#router eigrp 100
cr24-3560r-1(config-router)# no auto-summary
cr24-3560r-1(config-router)# eigrp router-id 10.125.100.4
cr24-3560r-1(config-router)# network 10.125.0.0 0.0.255.255

cr24-3560r-1#show ip eigrp neighbor port-channel 1
EIGRP-IPv4:(100) neighbors for process 100
H   Address                 Interface       Hold   Uptime    SRTT    RTO  Q
Seq
                                          (sec)            (ms)
Cnt Num
0   10.125.32.0             Po1               13     00:10:00     1
200   0   176
```

Building EIGRP Network Boundary

EIGRP creates and maintains a single flat routing network topology between EIGRP peers. Building a single routing domain enables complete network visibility and reach ability between all of the elements within the network.(access, distribution, core, serverfarm, WAN, etc)

In a tiered design, the access layer always has a single physical or logical forwarding path to the distribution layer. The access switch will build a forwarding topology pointing to same distribution switch as a single Layer-3 next-hop. Since the distribution switch provides a gateway function to the access switch, the routing design can be optimized with the following two techniques to improve performance and network convergence in the access-distribution block:

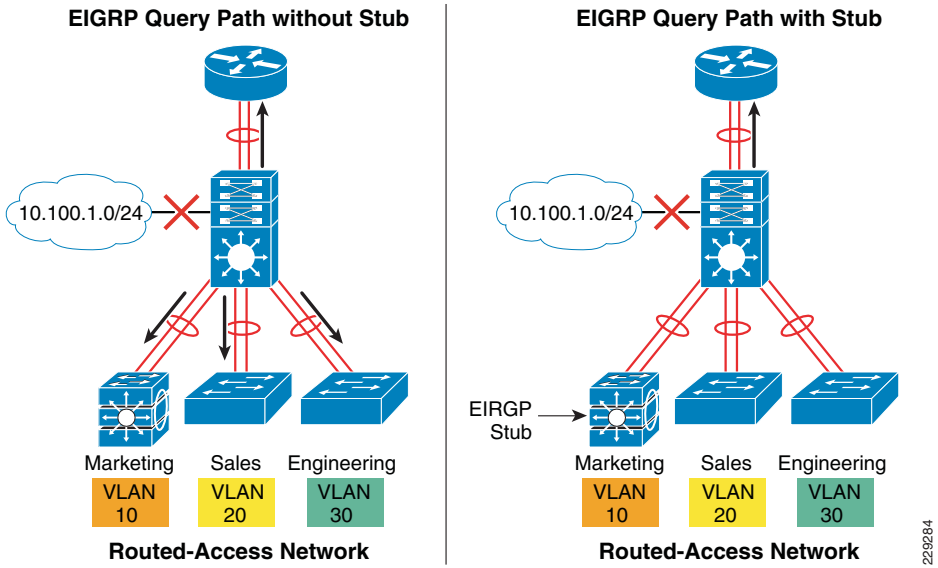
- Deploy Layer 3 access-switch in EIGRP stub mode
- Summarize network view to Layer-3 access-switch for intelligent routing function

Deploy Layer 3 Access-Switch in EIGRP Stub Mode

The Layer-3 access switch can be deployed to announce itself as a stub router that acts as a non-transit router and does not connect any other Layer-3 stub or non-stub routers. Announcing itself as a non-transit stub Layer-3 router is one way to notify the distribution router that it should not include the Layer-3 access switch in the EIGRP topology recomputation process. This optimized recomputation process will prevent unnecessary EIGRP network queries, which reduces network traffic, and simplifies the route computation.

As illustrated in Figure 21, implementing EIGRP stub function in the access switches, greatly reduces the number of EIGRP network queries.

Figure 21 EIGRP Query Path with and without Stub Implementation



EIGRP stub router in Layer-3 access-switch can announce routes to a distribution-layer router with great flexibility.

EIGRP stub router can be deployed to announce routes dynamically discovered or statically configured. Best practice design is to deploy EIGRP stub router to announce locally learned routes to aggregation layer.

Following is the example configuration to enable EIGRP stub routing in the Layer-3 access-switch, no configuration changes are required in distribution system:

Access

```
cr24-3560r-1(config)#router eigrp 100
cr24-3560r-1(config-router)#eigrp stub connected

cr24-3560r-1#show eigrp protocols detailed

Address Family Protocol EIGRP-IPv4:(100)
  EIGRP metric weight K1=1, K2=0, K3=1, K4=0, K5=0
```

```
  EIGRP maximum hopcount 100
  EIGRP maximum metric variance 1
  EIGRP NSF-aware route hold timer is 240
  EIGRP stub, connected
  Topologies : 0(base)
```

Distribution

```
cr24-4507-1#show ip eigrp neighbors detail port-channel 13
EIGRP-IPv4:(100) neighbors for process 100
H   Address                 Interface       Hold Uptime   SRTT   RTO  Q
Seq
                                            (sec)         (ms)       Cnt
Num
1   10.125.32.1             Po13              13 00:19:19   16   200  0
410
   Version 12.2/3.0, Retrans: 0, Retries: 0, Prefixes: 11
   Topology-ids from peer - 0

   Stub Peer Advertising ( CONNECTED ) Routes
   Suppressing queries
```

Summarizing Stub Routed-Access Network

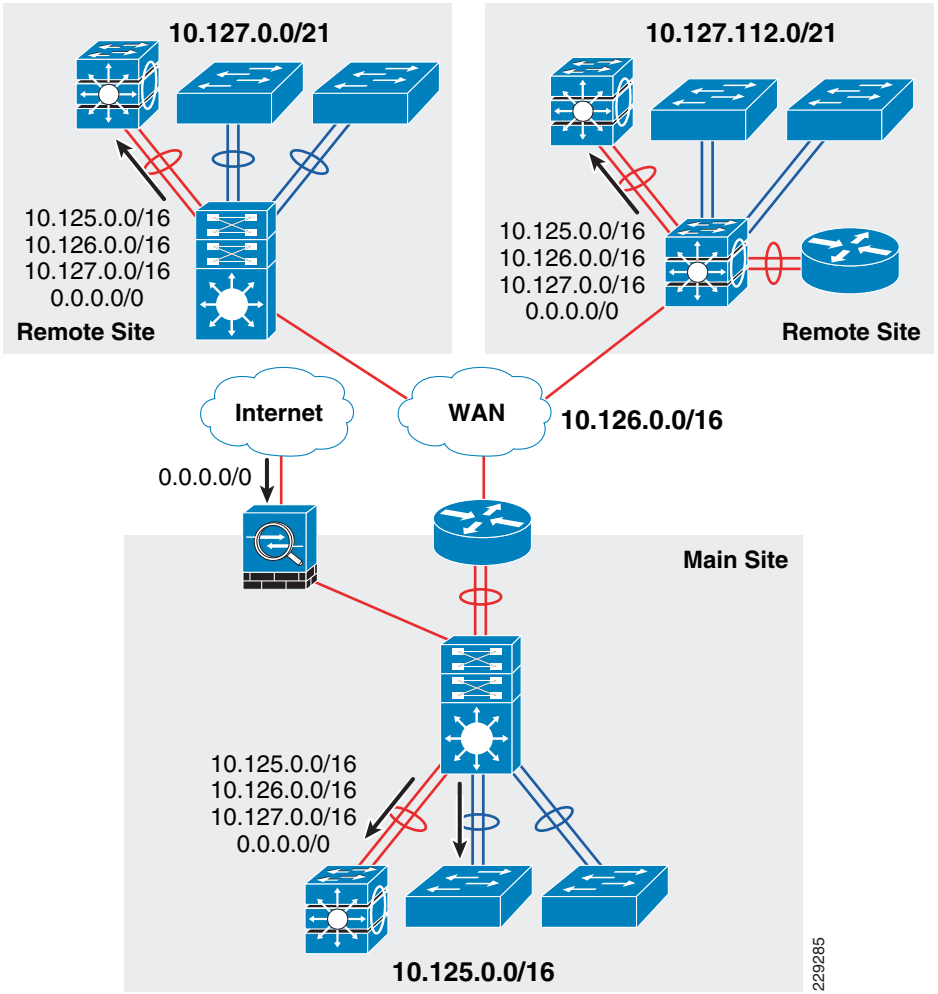
Enabling the EIGRP stub function on the access switch does not change the distribution router behavior of forwarding the full EIGRP topology table. The Distribution router must be configured to advertise summarized routes that do not compromise end-to-end reach ability, and help access switches maintain minimal routing information. In a network with a well designed IP addressing scheme, the aggregation system can advertise summarized routes in a classless address configuration, that reduce individual network advertisements, improve network scalability and network convergence. The distribution router must have full network topology information to ensure efficient reachability paths. Therefore, it is recommended to summarize at the distribution router, and not summarize at the access-layer.

Route summarization must be implemented on the distribution layer of main and each remote site network. This includes devices such as the WAN aggregation in the main site. The distribution router must advertise the following summarized network information to Layer 3 access-switch:

- Local Network—Distribution router can be implemented in hybrid access-distribution configuration that interconnects several multi-layer or routed-access enabled access-layer switches. Independent of route origination source (connected or dynamic route) and network size within the access-distribution block, the distribution router in main and remote site network must advertise a single, concise and summarized Layer 3 network to each Layer 3 access-switch and to core devices.
- Remote Network—Summarized network will be propagated dynamically across the network. Single summarization of all remote networks may be advertised to local Layer 3 access-switches, since it improves bandwidth efficiency. During a network outage, Layer 3 access-switch may drop traffic at the network edge instead of transmitting it to the distribution router to black hole traffic.
- WAN Network—Announcing a single summarized WAN network provides flexibility to troubleshoot and verify network availability.
- Default Network—When Layer 3 access-switch receives un-known destination traffic from the edge that does not match any of the above mentioned summarized networks, then it is sent to the distribution router to make a forwarding decision. The distribution router performs a forwarding table lookup and may forward to appropriate path or black hole the traffic. In a typical small enterprise network environment, a default route is announced by an Internet edge system, to forward all internet traffic. Distribution router must propagate this default route to the Layer 3 access-switch.

Figure 22 illustrates a summarized EIGRP network advertisement, by route aggregation system, that provides end-to-end internal and external network reachability.

Figure 22 End-to-End Routed-Access Network



Following is configuration example to deploy summarized and filtered Layer-3 network information to Layer-3 access-switch.

Distribution

```
interface Port-channel13
 description Connected to cr24-3560r-1
 dampening
 ip address 10.125.32.0 255.255.255.254
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp-key
 ip summary-address eigrp 100 10.125.0.0 255.255.0.0 5
 load-interval 30
 carrier-delay msec 0
!
!configure ACL and route-map to allow summarized route advertisement to
Layer 3 access-
switch
!
access-list 1 permit 0.0.0.0
access-list 1 permit 10.126.0.0
access-list 1 permit 10.127.0.0
access-list 1 permit 10.125.0.0

!
route-map EIGRP_STUB_ROUTES permit 10
 match ip address 1
!
router eigrp 100
 distribute-list route-map EIGRP_STUB_ROUTES out Port-channel13


cr24-4507-1#show ip protocols | inc Outgoing|filtered
  Outgoing update filter list for all interfaces is not set
    Port-channel13 filtered by
```

Access

```
cr24-3560r-1#show ip route eigrp
     10.0.0.0/8 is variably subnetted, 15 subnets, 4 masks
D       10.126.0.0/16 [90/3328] via 10.125.32.0, 01:37:21, Port-channel1
D       10.127.0.0/16 [90/3584] via 10.125.32.0, 01:37:21, Port-channel1
D       10.125.0.0/16 [90/1792] via 10.125.32.0, 01:34:29, Port-channel1
D*EX 0.0.0.0/0 [170/515072] via 10.125.32.0, 00:03:15, Port-channel1
cr24-3560r-1#
```

EIGRP Adjacency Protection

EIGRP adjacency protection guidelines discussed earlier for the core network, apply equally to routed access in the access-distribution block. The two challenges, system efficiency, and network security also apply equally to the routed access design, and the same solution is applied:

- System efficiency—EIGRP hello transmission must be blocked on an interface where there are no trusted EIGRP neighbors, to reduce CPU utilization and prevent network attacks. EIGRP routing process should only be enabled on interfaces where trusted enterprise devices are connected. All other interfaces can be suppressed in passive mode.

Following is the example configuration on Layer-3 access-switch that advertises networks enabled on SVI interfaces; however, keeps them in passive mode and explicitly allows EIGRP function on uplink port-channel to distribution router. Same configuration principle must be applied on each EIGRP router including distribution and core routers:

```
cr24-3560r-1(config)#router eigrp 100
cr24-3560r-1(config-router)# network 10.125.0.0 0.0.255.255
cr24-3560r-1(config-router)# passive-interface default
cr24-3560r-1(config-router)# no passive-interface Port-channel1
cr24-3560r-1#show ip eigrp interface
EIGRP-IPv4:(100) interfaces for process 100

                        Xmit           Queue   Mean   Pacing Time
Multicast Pending
Interface        Peers      Un/Reliable  SRTT   Un/Reliable   Flow
Timer  Routes
Po1                  1        0/0         1       0/1           50
0

cr24-3560r-1#show ip protocols | inc Passive|Vlan
  Passive Interface(s):
    Vlan1
    Vlan11
    Vlan12
    Vlan13
    Vlan14
```

- Network Security—EIGRP adjacency between distribution and Layer-3 access-switch must be secured. Following is the example configuration to enable EIGRP neighbor authentication using MD5:

Distribution

```
cr24-4507-1(config)#key chain eigrp-key
cr24-4507-1(config-keychain)# key 1
cr24-4507-1(config-keychain-key)#   key-string <password>

cr24-4507-1(config)#interface Port-channel13
cr24-4507-1(config-if)# description Connected to cr24-3560r-1
cr24-4507-1(config-if)# ip authentication mode eigrp 100 md5
cr24-4507-1(config-if)# ip authentication key-chain eigrp 100 eigrp-key
```

Access

```
cr24-3560r-1(config)#key chain eigrp-key
cr24-3560r-1(config-keychain)# key 1
cr24-3560r-1(config-keychain-key)#   key-string <password>
cr24-3560r-1(config)#interface Port-channel1
cr24-3560r-1(config-if)# description Connected to cr24-4507-1
cr24-3560r-1(config-if)# ip authentication mode eigrp 100 md5
cr24-3560r-1(config-if)# ip authentication key-chain eigrp 100 eigrp-key
```

Tuning EIGRP Protocol Timers

EIGRP protocol functions the same in routed-access as it does in the core network. It is highly recommended to retain default EIGRP hello and hold timers on distribution and Layer 3 access-switch and rely on EtherChannel and SSO-based recovery mechanisms, that offers sub-second network convergence, during individual link or supervisor failure scenarios.

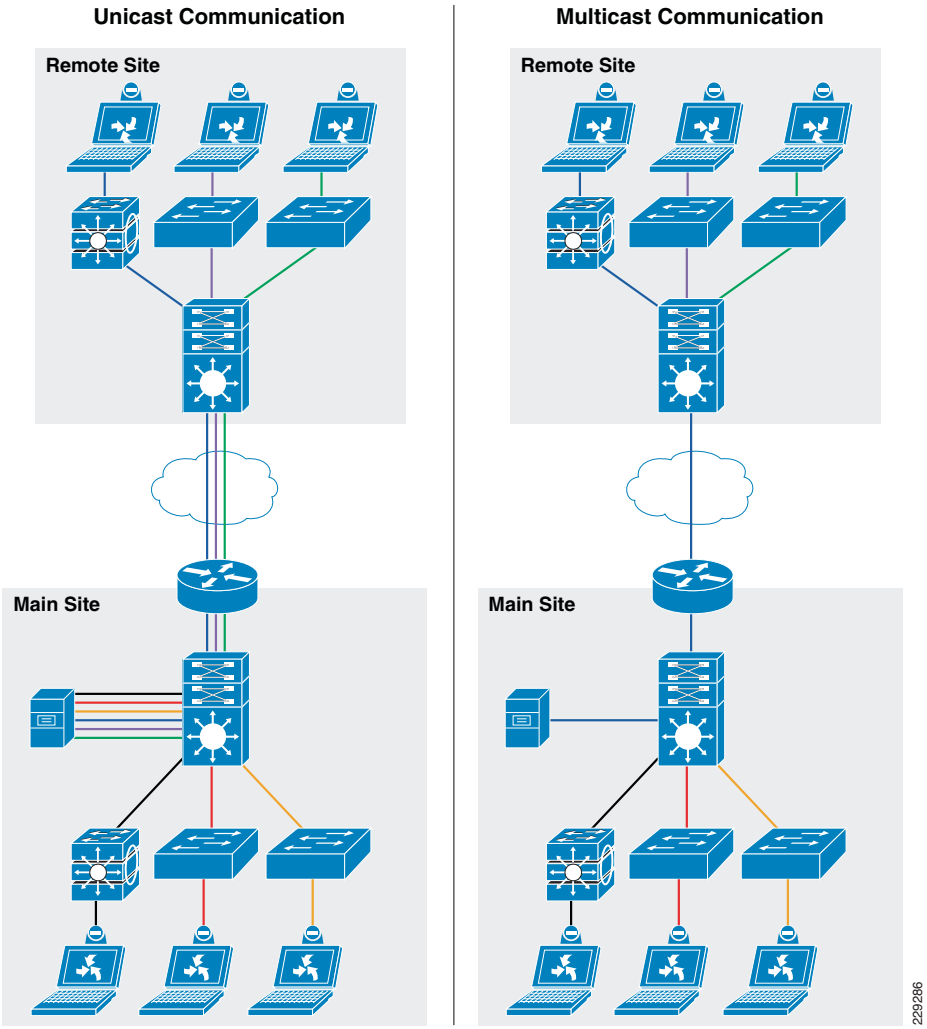
Deploying Multicast in Network

Communications in a IP network can be:

- Unicast—One source sends a message to one destination
- Broadcast—One source sends a message to all destinations
- Multicast—One source sends a message to a subset of destinations

IP multicast allows a source to transmit a message as a group transmission to a subset of hosts on the network. Many collaboration applications, such as video conferencing, distance learning, software distribution, utilize multicast techniques. IP multicast improves network bandwidth utilization, by reducing un necessary duplicate traffic. Multicast improves efficiency by reducing data processing on the source server, and sending a single flow into the network. Multicast packets are replicated in the network where paths diverge, by Protocol Independent Multicast (PIM)-enabled routers, and other supporting multicast protocols. See Figure 23.

Figure 23 Unicast versus Multicast Communication in Network



Multicast IP Addressing

The Internet Assigned Numbers Authority (IANA) controls the assignment of IP multicast addresses. A range of class D address space is assigned for IP multicast applications. All multicast group addresses fall in the range of 224.0.0.0 through 239.255.255.255. In IP multicast packets, the destination IP address is in the multicast group range, while the source IP address is always in the unicast address range. The multicast IP address space is further divided into several pools for well-known multicast network protocols, and inter-domain multicast communications as shown in Table 8.

Table 8 Multicast Address Range Assignments

Application	Address Range
Reserved – Link Local Network Protocols	224.0.0.0/24
Globally Scope – Group communication between organization and Internet	224.0.1.0 – 238.255.255.255
Source Specific Multicast (SSM) – PIM extension for one-to-many unidirectional multicast communication	232.0.0.0/8
GLOP – Inter-domain Multicast group assignment with reserved global Autonomous System (AS)	233.0.0.0/8
Limited Scope – Administratively scope address that remains constrained within local organization or AS. Commonly deployed in enterprise and other organization.	239.0.0.0/8

For the Schools SRA network design, the multicast IP addresses must be selected from the Limited Scope pool (239.0.0.0/8).

Multicast Routing Design

Each device between a multicast source and receiver must enable dynamic multicast. The technique for creating a multicast forwarding table is different than unicast routing and switching techniques. Multicast requires Multicast Routing Protocol (MRP) and Dynamic Group Membership (DGM) to enable communication.

Multicast Routing Protocol

IP multicast delivers source traffic to multiple receivers using the least amount of network resources, without placing additional burden on the source or the receivers. Multicast packet replication in the network is performed by Cisco routers and switches enabled with Protocol Independent Multicast (PIM) and other multicast routing protocols.

The network must build a packet distribution tree that specifies a unique forwarding path between the source subnet and each multicast group members subnet. A primary goal for the tree is to ensure that only one copy of each packet is forwarded on each branch of the tree. The two basic types of multicast distribution trees are source trees and shared trees:

- Source trees—The simplest form of a multicast distribution tree is a source tree, with the source at the root and the receivers at the branches. Because this tree uses the shortest path through the network, it is also referred to as a shortest path tree (SPT).
- Shared trees—A shared tree uses a single common root placed at a chosen point in the network. This shared root is called a Rendezvous Point (RP).

PIM protocol has two modes which support both types of multicast distribution trees:

- Dense Mode—This mode assumes that most routers in the network will distribute multicast traffic to each multicast group. PIM-DM builds distribution trees by initially flooding the entire network and then pruning back the small number of paths without receivers.

- Sparse Mode—This mode assumes that relatively few routers in the network will be involved in each multicast group. The hosts belonging to the group are usually widely dispersed, as would be the case for most multicast over the WAN. PIM-SM begins with an empty distribution tree and adds branches only as the result of explicit IGMP requests to join.

It is recommended to deploy multicast in PIM-SM in the small enterprise network design. All the recommended platforms in this design support PIM-SM mode on physical or logical (SVI and EtherChannel) interfaces.

Dynamic Group Membership

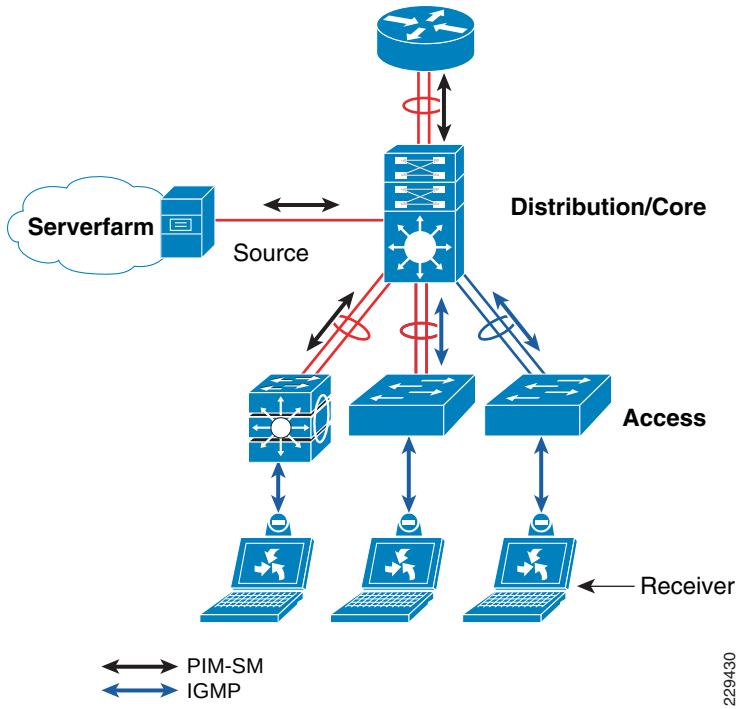
Multicast receiver registration and deletion is done via Internet Group Management Protocol (IGMP) signaling. IGMP operates between a multicast receiver in the access-layer and a collapsed core router at the distribution layer in the main or the remote site.

In a multi-layer design, the layer 3 boundary is at the distribution switch. Multi-layer access-switches do not run PIM, and therefor flood the traffic on all ports. This multi-layer access-switch limitation is solved by using IGMP snooping feature, which is enabled by default. Best practice is to not disable IGMP snooping feature.

In a routed-access network design, the Layer-3 boundary is at the access-layer and IGMP communication is between receiver and access-switch. Along with unicast routing protocol, PIM-SM must be enabled on the Layer 3 access-switch to communicate with RP in the network.

Figure 24 demonstrates multicast source and receiver registration procedure and how shared-tree is dynamically developed for multicast data delivery.

Figure 24 Multicast Source and Receiver Registration Procedure



Deploying PIM-SM

PIM-SM distributes information about active sources by forwarding data packets on the shared tree. Because PIM-SM uses shared trees initially, it requires the use of a RP. It is recommended to deploy the RP close to the multicast source (collapsed core-distribution router in the main site is a good choice). Multicast sources centrally deployed in main site will register themselves with the RP and then data is forwarded down the shared tree to the receivers that could be located anywhere in the network.

PIM-SM Rendezvous Point

PIM-SM distributes information about active sources by forwarding data packets on the shared tree. Because PIM-SM uses shared trees initially, it requires the use of a RP. It is recommended to deploy the RP close to the multicast source (collapsed core-distribution router in the district office is a good choice). Multicast sources centrally deployed in district office will register themselves with the RP and then data is forwarded down the shared tree to the receivers that could be located anywhere in the network.

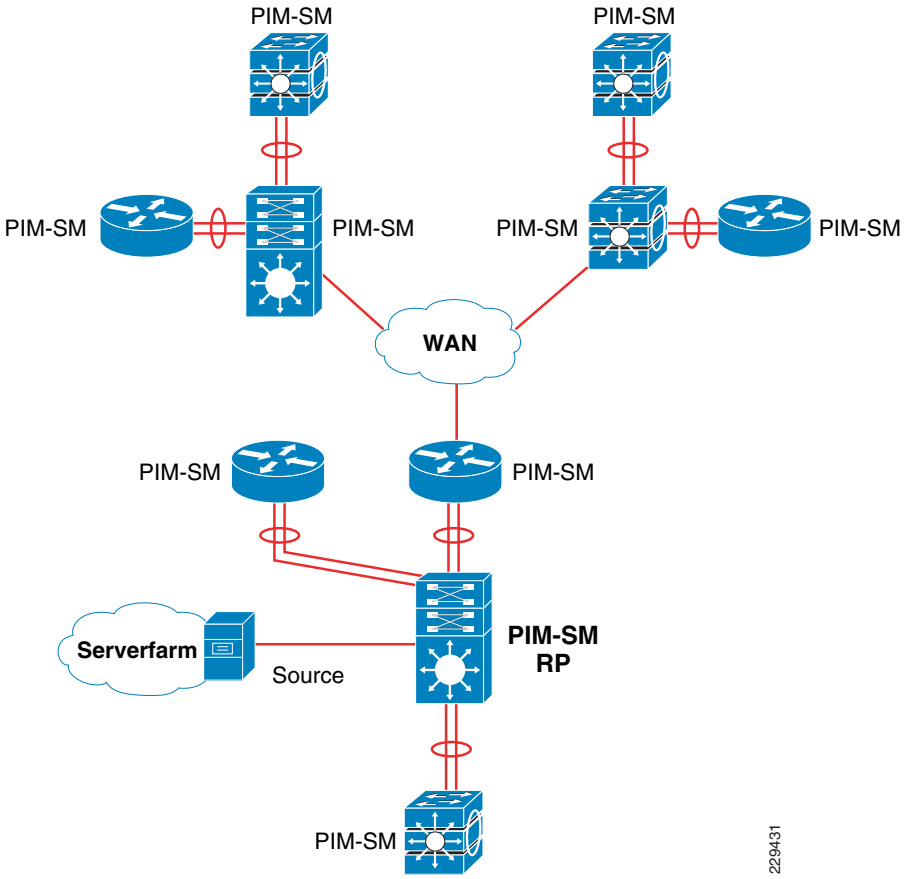
PIM-SM supports RP deployment in the following three different modes in the network:

- Static—As the name implies, RP must be statically identified and configured on each PIM router in the network. RP load-balancing and redundancy can be achieved using Anycast RP.

- Auto-RP—Dynamic method to discover and announce RP in the network. Auto-RP implementation is beneficial when there are multiple RPs and groups that often change in the network. To prevent network reconfiguration during change, RP mapping agent router must be designated in the network to receive RP group announcements and arbitrate conflicts. This capability is part of PIM version 1 specification.
- BootStrap Router (BSR)—Performs same task as Auto-RP but different mechanism. This capability is pare of PIM version 2 specification. Auto-RP and BSR cannot coexist or interoperate in the same network.

In a small to mid-size multicast network, static RP configuration is best overall, due primarily to the amount of administrative overhead that Auto-RP or BSR introduce. Static RP implementation offers same RP redundancy and load sharing and a simple ACL can be applied to deploy RP without compromising multicast network security. See Figure 25.

Figure 25 PIM-SM Network Design in Network Infrastructure



Following is an example configuration to deploy PIM-SM RP in the small enterprise network. Similar static PIM-SM configuration must be enabled on each Layer-3 PIM router or an access-switch in the remote sites:

Distribution - RP

```
cr24-4507-1(config)#interface Loopback1
cr24-4507-1(config-if)# description RP
cr24-4507-1(config-if)# ip address 10.125.100.100 255.255.255.255

cr24-4507-1(config)#ip multicast-routing
cr24-4507-1(config)#ip pim rp-address 10.125.100.100
```

Layer 3 Access

```
cr24-3560r-1(config)#ip multicast-routing distributed
cr24-3560r-1(config)#ip pim rp-address 10.125.100.100
```

Remote Site Core

```
cr36-3750s-1(config)#ip multicast-routing distributed
cr36-3750s-1(config)#ip pim rp-address 10.125.100.100
```

Upon successful PIM-SM RP implementation throughout the enterprise network, PIM-SM must be enabled on Layer-3 edge and core network-facing ports. The following sample configuration provides a simple PIM-SM implementation guideline to be implemented on every intermediate Layer-3 systems between receiver and source:

Distribution - RP

```
! Main Site - Access Network
cr24-4507-1(config)#interface range Vlan101 - 140
cr24-4507-1(config-if-range)# ip pim sparse-mode

! Main Site - Data Center Network
cr24-4507-1(config)#interface range Vlan141 - 150
cr24-4507-1(config-if-range)# ip pim sparse-mode

! Layer 3 Core and Routed-Access Port-Channel
cr24-4507-1(config)#interface range Port-channel 1, Port-channel 13,
Port-channel 15
cr24-4507-1(config-if-range)# ip pim sparse-mode


cr24-4507-1#show ip pim interface
Address          Interface Ver/   Nbr    Query  DR      DR
                                         Mode   Count   Intvl  Prior
10.125.32.4      Port-channel1v2/S   1      30     1     10.125.32.4
<omitted>
10.125.1.1       Vlan101         v2/S 0      30     1      10.125.1.1
```

```
cr24-4507-1#show ip mroute sparse
(*, 239.192.51.8), 02:33:37/00:03:12, RP 10.125.100.100, flags: SJC
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    Vlan111, Forward/Sparse, 02:04:33/00:02:44, H
    Vlan101, Forward/Sparse, 02:04:59/00:02:58, H
    Port-channel15, Forward/Sparse, 02:04:59/00:02:47, H
    Vlan131, Forward/Sparse, 02:04:59/00:02:32, H
    Port-channel13, Forward/Sparse, 02:04:59/00:03:12, H
    Vlan121, Forward/Sparse, 02:04:59/00:02:14, H
    Vlan146, Forward/Sparse, 02:21:26/00:02:01, H
```

Layer 3 Access

```
! Main Site - Layer 3 Access Network
cr24-3560r-1(config)#interface range Vlan11 - 20
cr24-3560r-1(config-if-range)# ip pim sparse-mode

! Routed-Access Port-Channel
cr24-4507-1(config)#interface Port-channel 1
cr24-4507-1(config-if)# ip pim sparse-mode

cr24-3560r-1#show ip pim interface
Address          Interface Ver/   Nbr    Query  DR     DR
                                  Mode   Count  Intvl  Prior
10.125.32.1Port-channel1v2/S   1      30     1      10.125.32.0
10.125.11.1      Vlan11v2/S   0      30     1      10.125.11.1
```

Implementing IGMP

By default the Layer-2 access-switch will dynamically detect IGMP hosts and multicast-capable Layer-3 routers in the network. The IGMP snooping and multicast router detection functions on a per VLAN basis, and is globally enabled by default for all the VLANs. The IGMP configuration can be validated using the show command on the Layer-2 access-switch:

```
cr24-2960-1#show ip igmp snooping
Global IGMP Snooping configuration:
-----------------------------------------------
IGMP snooping               : Enabled
IGMPv3 snooping (minimal)    : Enabled
Report suppression           : Enabled
TCN solicit query            : Disabled
TCN flood query count        : 2
Robustness variable          : 2
Last member query count      : 2
Last member query interval   : 1000

cr24-2960-1#show ip igmp snooping mrouter
Vlan    ports
------    ------
 101    Po1(dynamic)
 102    Po1(dynamic)

cr24-2960-1#show ip igmp snooping group
Vlan      GroupType        Version     Port List
------------------------------------------------------------------------
101       239.192.51.1igmp        v2     Fa0/1, Po1
101       239.192.51.2igmp        v2     Fa0/2, Po1
```

Multicast routing function changes when the access-switch is deployed in routed-access mode. PIM operation is performed at the access layer, therefore multicast router detection process is eliminated. The following output from a Layer-3 switch verifies that the local multicast ports are in router mode, and provide a snooped Layer-2 uplink port-channel which is connected to the collapsed core router, for multicast routing:

```
cr24-3560r-1#show ip igmp snooping mrouter
Vlan    ports
------    ------
  11    Router
  12    Router

cr24-3560r-1#show ip igmp membership | inc Channel|Vl
 Channel/Group                  Reporter        Uptime       Exp.Flags
Interface
 *,239.192.51.8                 10.125.11.2000:17:52  02:45 2A     Vl11
 *,239.192.51.9                 10.125.11.13100:17:52  02:43 2A     Vl12
```

Multicast Security—Preventing Rogue Source

This section provides basic multicast security configuration guidelines to prevent an unauthorized host in the network from acting like a rogue source in the network and sending multicast traffic.

In a PIM-SM network, an unwanted traffic source can be controlled with the pim accept-register command. When the source traffic hits the first-hop router, the first-hop router (DR) creates (S,G) state and sends a PIM Source Register message to the RP. If the source is not listed in the accept-register filter list (configured on the RP), then the RP rejects the Register and sends back an immediate Register-Stop message to the DR. The drawback with this method of source-filtering is that the pim accept-register command on the RP, PIM-SM (S,G) state is still created on the source's first-hop router. This can result in traffic reaching receivers local to the source and located between the source and the RP. Furthermore, the pim accept-register command works on the control plane of the RP, which could be used to overload the RP with “fake” register messages, and possibly cause a DoS condition.

Best practice is to apply the pim accept-register command on the RP in addition to other edge-filtering methods, such as simple data plane ACLs on all DRs and on all ingress points into the network. While ingress ACLs on the DR are sufficient in a perfectly configured and operated network, best practice includes configuring the pim accept-register command on the RP in the main site as a secondary security mechanism in case of misconfiguration on the edge routers.

Following is the sample configuration with a simple ACL which has been applied to the RP to filter only on the source address. It is also possible to filter the source and the group with the use of an extended ACL on the RP:

Distribution-RP

```
cr24-4507-1(config)#ip access-list extended PERMIT-SOURCES
cr24-4507-1(config-ext-nacl)# permit ip 10.125.31.80 0.0.0.15 239.192.0.0
0.0.255.255

cr24-4507-1(config)#ip pim accept-register list PERMIT-SOURCES
```

Multicast Security—Preventing Rogue RP

Any router can be misconfigured or maliciously advertise itself as a multicast RP in the network. with the valid multicast group address. With a static RP configuration, each PIM-enabled router in the network can be configured to use the static RP for the multicast source and ignore any Auto-RP or BSR multicast router announcement.

Following is the sample configuration that must be applied to each PIM-enabled router in the main and remote sites, to accept PIM announcements only from the static RP and ignore dynamic multicast group announcement from any other RP:

Distribution-RP

```
cr24-4507-1(config)#ip access-list standard Allowed_MCAST_Groups
cr24-4507-1(config-std-nacl)# permit 224.0.1.39
cr24-4507-1(config-std-nacl)# permit 224.0.1.40
cr24-4507-1(config-std-nacl)# permit 239.192.0.0 0.0.255.255

cr24-4507-1(config)#ip pim rp-address 10.125.100.100 Allowed_MCAST_Groups
override

cr24-4507-1#show ip pim rp mapping
PIM Group-to-RP Mappings
Acl: Allowed_MCAST_Groups, Static-Override
    RP: 10.125.100.100 (?)
```

Deploying QoS in Network

IP networks forward traffic on a best-effort basis by default. The routing protocol forwards packets over the best path, but offers no guarantee of delivery. This model works well for TCP-based data applications that adapt gracefully to variations in latency, jitter, and loss. The Small Enterprise Design Profile is a multi-service network design which supports voice and video as well as data traffic on a single network. Real time applications (such as voice, video) require packets delivered with in specified loss, delay and jitter parameters. Quality-of-Service (QoS) is a collection of features which allows the network to dedicate network resources for higher priority real time applications, while reserving sufficient network resources to service lower priority traffic. QoS accomplishes this by providing differentiated services, depending on the traffic type. For a detailed discussion of QoS, refer to the Enterprise QoS SRND at the following URL:

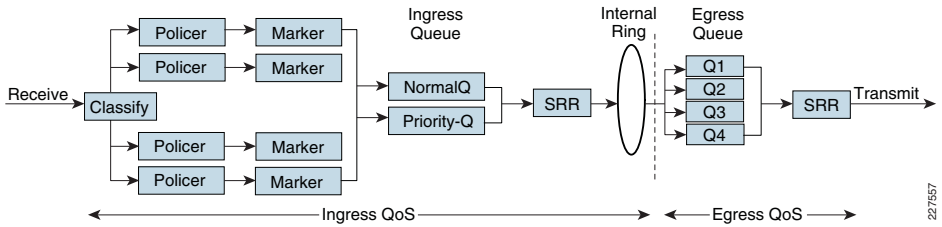
http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND/QoS-SRND-Book.html

While design principles are common, QoS implementation varies between fixed-configuration switches and the modular switching platforms like the Cisco Catalyst 4500-E/6500-E. This section discusses the internal switching architecture and the differentiated QoS structure on a per-hop-basis.

QoS in Catalyst Fixed Configuration Switches

The QoS implementation in Cisco Catalyst 2960, 2960S, 3560-X and 3750-X Series switches is similar. There is no difference in ingress or egress packet classification, marking, queuing and scheduling implementation among these Catalyst platforms. The Cisco Catalyst switches allow users to create a policy-map by classifying incoming traffic (Layer 2 to Layer 4). Catalyst switches allow attaching the policy-map to an individual physical port or to logical interfaces (SVI or port-channel). This creates a common QoS policy which may be used in multiple networks. To prevent switch fabric and egress physical port congestion, the ingress QoS policing structure can strictly filter excessive traffic at the network edge. All ingress traffic from edge ports passes through the switch fabric and congestion may occur at the egress ports. Congestion in access-layer switch can be prevented by tuning queuing scheduler and Weighted Tail Drop (WTD) drop parameters. See Figure 26.

Figure 26 Fixed Configuration Catalyst QoS Architecture



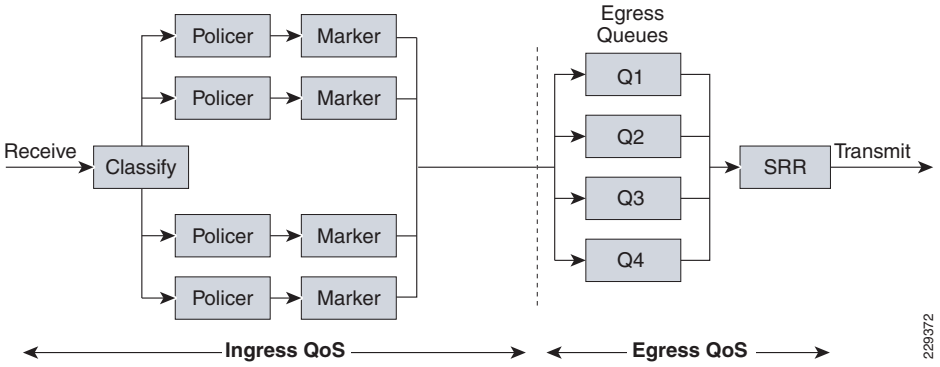
The main difference between these platforms is the switching capacity which ranges from 1G to 10G. The switching architecture and some of the internal QoS structure differs between these switches also. Following are some important differences to consider when selecting the access switch:

- The Cisco Catalyst 2960 does not support multilayer switching and does not support per-VLAN or per-port/per-VLAN policies.
- The Cisco Catalyst 2960 can police to a minimum rate of 1 Mbps; all other switches including next-generation Cisco Catalyst 2960-S Series within this product family can police to a minimum rate of 8 kbps.
- Only the Cisco Catalyst 3560-X and 3750-X support IPv6 QoS.
- Only the Cisco Catalyst 3560-X and 3750-X support policing on 10-Gigabit Ethernet interfaces.

- Only the Cisco Catalyst 3560-X and 3750-X support SRR shaping weights on 10-Gigabit Ethernet interfaces.

The next-generation Cisco Catalyst 2960-S Series platform introduces modified QoS architecture. To reduce the latency and improve application performance, the new Cisco 2960-S platform does not support ingress queueing and buffer function in harware. All other ingress and egress queuing, buffer and bandwidth sharing function remain consistent as the Cisco Catalyst 2960 platform. Each physical ports, including StackPort, have 2 MB buffer capacity to prevent traffic drop during congestion. This buffer allocation is static and cannot be modified by the user. However, when the Cisco Catalyst 2960-S is deployed in FlexStack configuration mode, there is a flexibility to assign different buffer size on egress queue of StackPort. Figure 27 illustrates QoS architecture on Catalyst 2960-S Series platform.

Figure 27 QoS Implementation in Catalyst 2960-S Switches

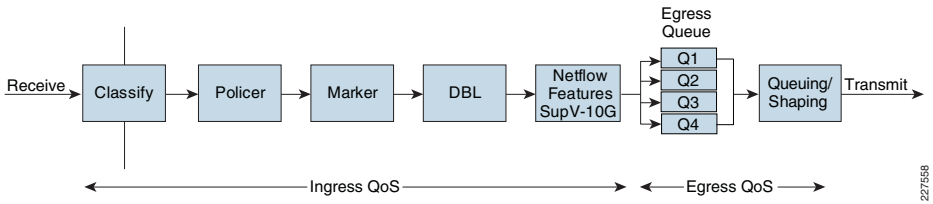


QoS in Cisco Modular Switches

Cisco Catalyst 4507R-E and 6500-E are high density, resilient switches for large scale networks. The Small Enterprise Design Profile uses the Cisco Catalyst 4507R-E in the main and larger remote site designs, hence all the QoS recommendations in this section will be based on 4500-E architecture. Cisco Catalyst 4500-E Series platform are widely deployed with classic and next-generation supervisors.

The classification function in the classic supervisor module is based on incoming DSCP or CoS setting in the pack, which was assigned by the access-layer switch. Catalyst 4500-E with classic supervisor performs ingress and egress QoS function based on internal mapping table that performs DSCP, ToS, or CoS interworking. Classic supervisor relies on trust model configuration; redirection of ingress traffic to an appropriate queue is based on the trust model defined on the edge port. See Figure 28.

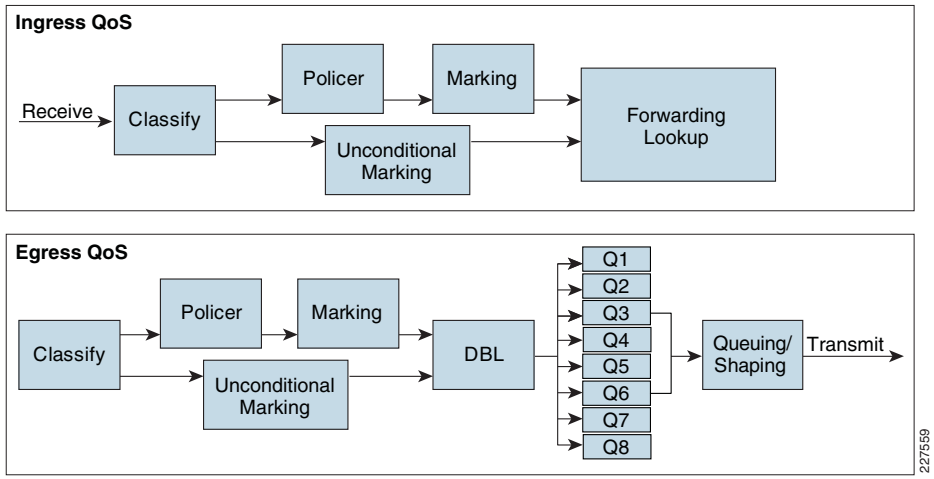
Figure 28 Catalyst 4500-E– Classic Supervisor QoS Architecture



The Cisco Catalyst 4500-E with next generation Sup-6E (see Figure 29) is designed to offer better differentiated and preferential QoS services for various class-of-service traffic. New QoS capabilities in the Sup-6E enable administrators to take advantage of hardware-based intelligent classification and take action to optimize application performance and network availability. The QoS implementation in Sup-6E supports Modular QoS CLI (MQC) as implemented in IOS-based routers that overall enhances QoS capabilities and eases implementation and operations. Following are some of the key QoS features which differentiate the Sup-6E versus classic supervisors:

- Trust and Table-Map—MQC based QoS implementation offers a number of implementation and operational benefits over classic supervisors that rely on Trust model and internal Table-map as a tool to classify and mark ingress traffic.
- Internal DSCP—The queue placement in Sup-6E is simplified by leveraging the MQC capabilities to explicitly map DSCP or CoS traffic in hard-coded egress Queue structure,. For example, DSCP 46 can be classified with ACL and can be matched in PQ class-map of an MQC in Sup-6E.
- Sequential vs Parallel Classification—With MQC-based QoS classification, the Sup6-E provides sequential classification rather than parallel. Sequential classification method allows the network administrator to classify traffic at egress based on the ingress markings.

Figure 29 Catalyst 4500-E—Supervisor 6-E QoS Architecture

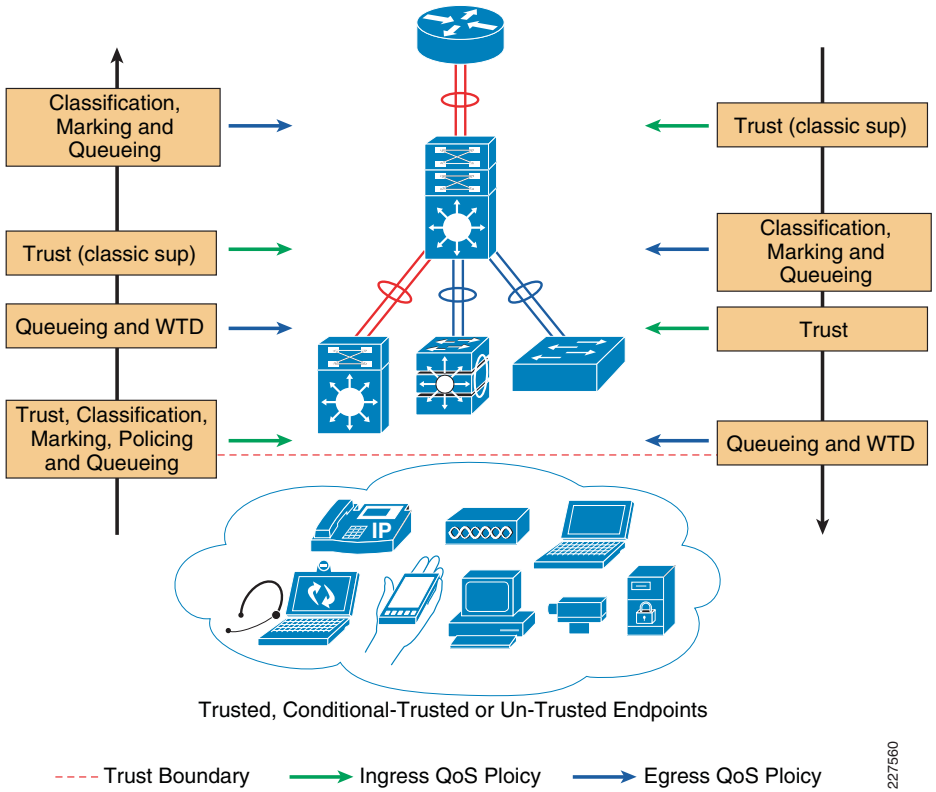


QoS Framework

QoS needs to be designed and implemented considering the entire network. This includes defining trust points, and determining which policies to enforce at each device within the network. Developing the trust model, guides policy implementations for each device.

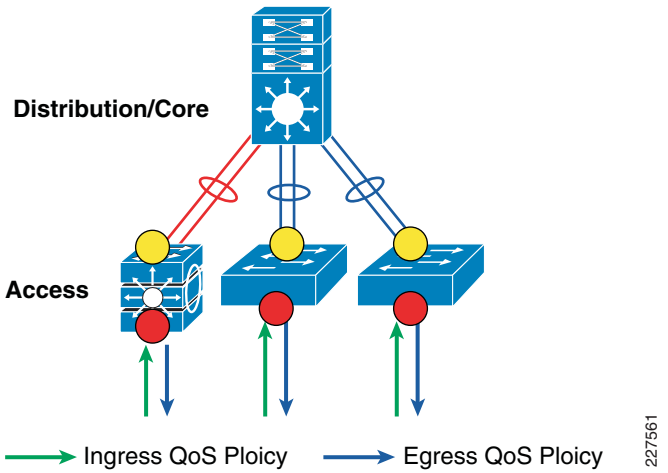
Figure 30 depicts QoS trust model that guides QoS policy implementation in the main and remote site networks.

Figure 30 Small Enterprise Network QoS Framework



The devices (routers, switches) within the internal network are managed by the system administrator, and hence are classified as trusted devices. Access-layer switches communicate with devices that are beyond the network boundary and within the internal network domain. QoS trust boundary at the access-layer communicates with various devices that could be deployed in different trust models (Trusted, Conditional-Trusted, or Un-Trusted). This section discusses the QoS policies for the traffic that traverses access-switch QoS trust boundary. The QoS function is unidirectional; it provides flexibility to set different QoS polices for traffic entering the network versus traffic that is exiting the network. See Figure 31.

Figure 31 Small Enterprise Network Edge QoS Boundary



QoS Trust Boundary

The access-switch provides the entry point to the network for end devices. The access-switch must decide whether to accept the QoS markings from each endpoint, or whether to change them. This is determined by the QoS policies, and the trust model with which the endpoint is deployed.

End devices are classified into one of three different trust models; each with its own unique security and QoS policies to access the network:

- Untrusted—An unmanaged device that does not pass through the network security policies. For example, employee-owned PC or network printer. Packets with 802.1p or DSCP marking set by untrusted endpoints are reset to default by the access-layer switch at the edge. Otherwise, it is possible for an unsecured user to take away network bandwidth that may impact network availability and security for other users.
- Trusted—Devices that passes through network access security policies and are managed by network administrator. For example, secure PC or IP endpoints (i.e., servers, cameras, DMP, wireless access points, VoIP/video conferencing gateways, etc). Even when these devices are network administrator maintained and secured, QoS policies must still be enforced to classify traffic and assign it to the appropriate queue to provide bandwidth assurance and proper treatment during network congestion.
- Conditionally-Trusted—A single physical connection with one trusted endpoint and a indirect untrusted endpoint must be deployed as conditionally-trusted model. The trusted endpoints are still managed by the network administrator, but it is possible that the untrusted user behind the endpoint may or may not be secure. For example, Cisco Unified IP Phone + PC. These deployment scenarios require hybrid QoS policy that intelligently distinguishes and applies different QoS policy to the trusted and untrusted endpoints that are connected to the same port.

Deploying QoS

The ingress QoS policy at the access-switches needs to be established, since this is the trust boundary, where traffic enters the network. The following ingress QoS techniques are applied to provide appropriate service treatment and prevent network congestion:

- Trust—After classifying the endpoint the trust settings must be explicitly set by a network administrator. By default, Catalyst switches set each port in untrusted mode when QoS is enabled.
- Classification—IETF standard has defined a set of application classes and provides recommended DSCP settings. This classification determines the priority the traffic will receive in the network. Using the IETF standard, simplifies the classification process and improves application and network performance.
- Policing—To prevent network congestion, the access-layer switch limits the amount of inbound traffic up to its maximum setting. Additional policing can be applied for known applications, to ensure the bandwidth of an egress queue is not completely consumed by one application.
- Marking—Based on trust model, classification, and policer settings the QoS marking is set at the edge before approved traffic enters through the access-layer switching fabric. Marking traffic with the appropriate DSCP value is important to ensure traffic is mapped to the appropriate internal queue, and treated with the appropriate priority.
- Queueing—To provide differentiated services internally in the Catalyst switching fabric, all approved traffic is queued into priority or non-priority ingress queue. Ingress queueing architecture assures real-time applications, like VoIP traffic, are given appropriate priority (eg transmitted before data traffic).

Implementing QoS Trust Mode

By default, QoS is disabled on all Catalyst switches and must be explicitly enabled in global configuration mode. The QoS configuration is the same for a multilayer or routed-access deployment. The following sample QoS configuration must be enabled on all the access-layer switches deployed in main and remote sites.

```
cr24-2960-1(config)#mls qos
cr24-2960-1#show mls qos
QoS is enabled
QoS ip packet dscp rewrite is enabled
```

Upon enabling QoS in the Catalyst switches, all physical ports are assigned untrusted mode. The network administrator must explicitly enable the trust settings on the physical port where trusted or conditionally trusted endpoints are connected. The Catalyst switches can trust the ingress packets based on 802.1P (CoS-based), ToS (ip-prec-based) or DSCP (DSCP-based) values. Best practice is to deploy DSCP-based trust mode on all the trusted and conditionally-trusted endpoints. This offers a higher level of classification and marking granularity than other methods. The following sample DSCP-based trust configuration must be enabled on the access-switch ports connecting to trusted or conditionally-trusted endpoints.

Access (Multilayer or Routed-Access)

Trusted Port

```
cr24-2960-1(config)#interface FastEthernet0/5
cr24-2960-1(config-if)# description CONNECTED TO IPVS 2500 - CAMERA
cr24-2960-1(config-if)# mls qos trust dscp

cr24-2960-1#show mls qos interface f0/5
FastEthernet0/5
trust state: trust dscp
trust mode: trust dscp
trust enabled flag: ena
COS override: dis
default COS: 0
DSCP Mutation Map: Default DSCP Mutation Map
Trust device: none
qos mode: port-based
```

Conditionally-Trusted Port

```
cr24-2960-1(config)#interface FastEthernet0/3
cr24-2960-1(config-if)# description CONNECTED TO PHONE
cr24-2960-1(config-if)# mls qos trust device cisco-phone
cr24-2960-1(config-if)# mls qos trust dscp

cr24-2960-1#show mls qos interface f0/3
FastEthernet0/3
trust state: trust dscp
trust mode: trust dscp
trust enabled flag: ena
COS override: dis
default COS: 0
DSCP Mutation Map: Default DSCP Mutation Map
Trust device: cisco-phone
qos mode: port-based
```

UnTrusted Port

As described earlier, the default trust mode is untrusted when globally enabling QoS function. Without explicit trust configuration on Fas0/1 port, the following show command verifies current trust state and mode:

```
cr24-2960-1#show mls qos interface f0/1
FastEthernet0/1
trust state: not trusted
trust mode: not trusted
```

```
trust enabled flag: ena
COS override: dis
default COS: 0
DSCP Mutation Map: Default DSCP Mutation Map
Trust device: none
qos mode: port-based
```

Implementing QoS Classification

When creating QoS classification policies, the network administrator needs to consider what applications are present at the access edge (in the ingress direction) and whether these applications are sourced from trusted or untrusted endpoints. If PC endpoints are secured and centrally administered, then endpoint PCs may be considered trusted endpoints. In most deployments, this is not the case, thus PCs are considered untrusted endpoints for the remainder of this document.

Not every application class, as defined in the Cisco-modified RFC 4594-based model, is present in the ingress direction at the access edge; therefore, it is not necessary to provision the following application classes at the access-layer:

- Network Control—It is assumed that access-layer switch will not transmit or receive network control traffic from endpoints; hence this class is not implemented.
- Broadcast Video—Broadcast video and multimedia streaming server are centrally deployed at the main and multicast traffic is originated from trusted serverfarm servers and is unidirectional to remote site endpoints (and should not be sourced from remote site endpoints).
- Operation, Administration and Management—Primarily generated by network devices (routers, switches) and collected by management stations which are typically deployed in the trusted serverfarm network, or a network control center.

All applications present at the access edge need to be assigned a classification, as shown in Figure 34. Voice traffic is primarily sourced from Cisco IP telephony devices residing in the voice VLAN (VVLAN). These are trusted devices, or conditionally trusted, if users also attach PC's, etc to the same port. Voice communication may also be sourced from PC's with soft-phone applications, like Cisco Unified Personal Communicator (CUPC). Since such applications share the same UDP port range as multimedia conferencing traffic (UDP/RTP ports 16384-32767) this soft-phone VoIP traffic is indistinguishable, and should be classified with multimedia conferencing streams. See Figure 32.

Figure 32 QoS Classes

Application	PHB	Application Examples	Present at Campus Access-Edge (Ingress)?	Trust Boundary
Network Control	CS6	EIGRP, OSPF, HSRP, IKE		
VoIP	EF	Cisco IP Phone	Yes	Trusted
Broadcast Video		Cisco IPVS, Enterprise TV		
Realtime Interactive	CS4	Cisco TelePresence	Yes	Trusted
Multimedia Conferencing	AF4	Cisco CUPC, WebEx	Yes	Untrusted
Multimedia Streaming	AF3	Cisco DMS, IP/TV		
Signaling	CS3	SCCP, SIP, H.323	Yes	Trusted
Transactional Data	AF2	ERP Apps, CRM Apps	Yes	Untrusted
OAM	CS2	SNMP, SSH, Syslog		
Bulk Data	AF1	Email, FTP, Backup	Yes	Untrusted
Best Effort	DF	Default Class	Yes	Untrusted
Scavenger	CS1	YouTube, Gaming, P2P	Yes	Untrusted

227562

MQC offers scalability and flexibility in configuring QoS to classify all 8 application classes by using match statements or an extended access-list to match the exact value or range of Layer-4 known ports that each application uses to communicate on the network. The following sample configuration creates an extended access-list for each application and then applies it under class-map configuration mode.

```
cr24-3560r-1(config)#ip access-list extended MULTIMEDIA-CONFERENCING
cr24-3560r-1(config-ext-nacl)# remark RTP
cr24-3560r-1(config-ext-nacl)# permit udp any any range 16384 32767
cr24-3560r-1(config-ext-nacl)#!
cr24-3560r-1(config-ext-nacl)#ip access-list extended SIGNALING
cr24-3560r-1(config-ext-nacl)# remark SCCP
cr24-3560r-1(config-ext-nacl)# permit tcp any any range 2000 2002
cr24-3560r-1(config-ext-nacl)# remark SIP
cr24-3560r-1(config-ext-nacl)# permit tcp any any range 5060 5061
cr24-3560r-1(config-ext-nacl)# permit udp any any range 5060 5061
cr24-3560r-1(config-ext-nacl)#
cr24-3560r-1(config-ext-nacl)#ip access-list extended TRANSACTIONAL-DATA
cr24-3560r-1(config-ext-nacl)# remark HTTPS
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 443
cr24-3560r-1(config-ext-nacl)# remark ORACLE-SQL*NET
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 1521

cr24-3560r-1(config-ext-nacl)# permit udp any any eq 1521
cr24-3560r-1(config-ext-nacl)# remark ORACLE
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 1526
```

```
cr24-3560r-1(config-ext-nacl)# permit udp any any eq 1526
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 1575
cr24-3560r-1(config-ext-nacl)# permit udp any any eq 1575
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 1630
cr24-3560r-1(config-ext-nacl)#
cr24-3560r-1(config-ext-nacl)#ip access-list extended BULK-DATA
cr24-3560r-1(config-ext-nacl)# remark FTP
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq ftp
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq ftp-data
cr24-3560r-1(config-ext-nacl)# remark SSH/SFTP
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 22
cr24-3560r-1(config-ext-nacl)# remark SMTP/SECURE SMTP
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq smtp
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 465
cr24-3560r-1(config-ext-nacl)# remark IMAP/SECURE IMAP
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 143
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 993
cr24-3560r-1(config-ext-nacl)# remark POP3/SECURE POP3
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq pop3
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 995
cr24-3560r-1(config-ext-nacl)# remark CONNECTED PC BACKUP
cr24-3560r-1(config-ext-nacl)# permit tcp any eq 1914 any
cr24-3560r-1(config-ext-nacl)#
cr24-3560r-1(config-ext-nacl)#ip access-list extended DEFAULT
cr24-3560r-1(config-ext-nacl)# remark EXPLICIT CLASS-DEFAULT
cr24-3560r-1(config-ext-nacl)# permit ip any any
cr24-3560r-1(config-ext-nacl)#
cr24-3560r-1(config-ext-nacl)#ip access-list extended SCAVENGER
cr24-3560r-1(config-ext-nacl)# remark KAZAA
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 1214
cr24-3560r-1(config-ext-nacl)# permit udp any any eq 1214
cr24-3560r-1(config-ext-nacl)# remark MICROSOFT DIRECT X GAMING
cr24-3560r-1(config-ext-nacl)# permit tcp any any range 2300 2400
cr24-3560r-1(config-ext-nacl)# permit udp any any range 2300 2400
cr24-3560r-1(config-ext-nacl)# remark APPLE ITUNES MUSIC SHARING
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 3689
cr24-3560r-1(config-ext-nacl)# permit udp any any eq 3689
cr24-3560r-1(config-ext-nacl)# remark BITTORRENT
cr24-3560r-1(config-ext-nacl)# permit tcp any any range 6881 6999
cr24-3560r-1(config-ext-nacl)# remark YAHOO GAMES
cr24-3560r-1(config-ext-nacl)# permit tcp any any eq 11999
cr24-3560r-1(config-ext-nacl)# remark MSN GAMING ZONE
cr24-3560r-1(config-ext-nacl)# permit tcp any any range 28800 29100
cr24-3560r-1(config-ext-nacl)#
```

Creating class-map for each application services and applying match statement:

```
cr24-3560r-1(config)#class-map match-all VVLAN-SIGNALING
cr24-3560r-1(config-cmap)# match ip dscp cs3
cr24-3560r-1(config-cmap)#
cr24-3560r-1(config-cmap)#class-map match-all VVLAN-VOIP
cr24-3560r-1(config-cmap)# match ip dscp ef
cr24-3560r-1(config-cmap)#
cr24-3560r-1(config-cmap)#class-map match-all MULTIMEDIA-CONFERENCING
cr24-3560r-1(config-cmap)# match access-group name
MULTIMEDIA-CONFERENCING
cr24-3560r-1(config-cmap)#
cr24-3560r-1(config-cmap)#class-map match-all SIGNALING
cr24-3560r-1(config-cmap)# match access-group name SIGNALING
cr24-3560r-1(config-cmap)#
cr24-3560r-1(config-cmap)#class-map match-all TRANSACTIONAL-DATA
cr24-3560r-1(config-cmap)# match access-group name TRANSACTIONAL-DATA

cr24-3560r-1(config-cmap)#
cr24-3560r-1(config-cmap)#class-map match-all BULK-DATA
cr24-3560r-1(config-cmap)# match access-group name BULK-DATA
cr24-3560r-1(config-cmap)#
cr24-3560r-1(config-cmap)#class-map match-all DEFAULT
cr24-3560r-1(config-cmap)# match access-group name DEFAULT
cr24-3560r-1(config-cmap)#
cr24-3560r-1(config-cmap)#class-map match-all SCAVENGER
cr24-3560r-1(config-cmap)# match access-group name SCAVENGER
```

Implementing Ingress Policer

It is important to limit how much bandwidth each class may use at the ingress to the access-layer for two primary reasons:

- Bandwidth Bottleneck—To prevent network congestion, each physical port at trust boundary must be rate-limited. The rate-limit value may differ based on several factors—end-to-end network bandwidth capacity, end-station, and application performance capacities, etc.
- Bandwidth Security—Well-known applications like Cisco IP telephony, use a fixed amount of bandwidth per device, based on codec. It is important to police high-priority application traffic which is assigned to the high-priority queue, otherwise it could consume too much overall network bandwidth and impact other application performance.

In addition to policing, the rate-limit function also provides the ability to take different actions on the excess incoming traffic which exceeds the established limits. The exceed-action for each class must be carefully designed based on the nature of application to provide best effort service based on network bandwidth availability. Table 9 provides best practice policing guidelines for different classes to be implemented for trusted and conditional-trusted endpoints at the network edge.

Table 9 Best Practice Policing Guidelines

Application	Policing Rate	Conform-Action	Exceed-Action
VoIP Signaling	<32 kbps	Pass	Drop
VoIP Bearer	<128 kbps	Pass	Drop
Multimedia Conferencing	<5Mbps[1]	Pass	Drop
Signaling	<32 kbps	Pass	Drop
Transactional Data	<10 Mbps[1]	Pass	Remark to CS1
Bulk Data	<10 Mbps[1]	Pass	Remark to CS1
Best Effort	<10 Mbps[1]	Pass	Remark to CS1
Scavenger	<10 Mbps[1]	Pass	Drop

1. Rate varies based on several factors as defined earlier. This table depicts sample rate-limiting values.

As described in the "QoS in Catalyst Fixed Configuration Switches" section on page -28, the policer capabilities differ in Cisco Catalyst switching platforms. When deploying policer policies on the access-layer switches the following platform limitations must be taken into consideration:

- The Catalyst 2960 can only police to a minimum rate of 1 Mbps; all other platforms, including next-generation Cisco Catalyst 2960-S Series, within this switch-product family can police to a minimum rate of 8 kbps.
- Only the Cisco Catalyst 3560-X and 3750-X support policing on 10 Gigabit Ethernet interfaces.

The following sample configuration shows how to deploy policing for multiple classes on trusted and conditionally-trusted ingress ports in access-layer switches.

Trusted or Conditionally-Trusted Port

```
cr24-3560r-1(config)#policy-map Phone+PC-Policy
cr24-3560r-1(config-pmap)# class VVLAN-VOIP
cr24-3560r-1(config-pmap-c)# police 128000 8000 exceed-action drop
cr24-3560r-1(config-pmap-c)# class VVLAN-SIGNALING
cr24-3560r-1(config-pmap-c)# police 32000 8000 exceed-action drop
cr24-3560r-1(config-pmap-c)# class MULTIMEDIA-CONFERENCING
cr24-3560r-1(config-pmap-c)# police 5000000 8000 exceed-action drop
cr24-3560r-1(config-pmap-c)# class SIGNALING
cr24-3560r-1(config-pmap-c)# police 32000 8000 exceed-action drop
cr24-3560r-1(config-pmap-c)# class TRANSACTIONAL-DATA
cr24-3560r-1(config-pmap-c)# police 10000000 8000 exceed-action
policed-dscp-transmit
cr24-3560r-1(config-pmap-c)# class BULK-DATA
cr24-3560r-1(config-pmap-c)# police 10000000 8000 exceed-action
policed-dscp-transmit
cr24-3560r-1(config-pmap-c)# class SCAVENGER
cr24-3560r-1(config-pmap-c)# police 10000000 8000 exceed-action drop
cr24-3560r-1(config-pmap-c)# class DEFAULT
cr24-3560r-1(config-pmap-c)# police 10000000 8000 exceed-action
policed-dscp-transmit
```

All ingress traffic (default class) from untrusted endpoint be must be policed without explicit classification that requires differentiated services. The following sample configuration shows how to deploy policing on untrusted ingress ports in access-layer switches:

UnTrusted Port

```
cr24-3560r-1(config)#policy-map UnTrusted-PC-Policy
cr24-3560r-1(config-pmap)# class class-default
cr24-3560r-1(config-pmap-c)#  police 10000000 8000 exceed-action drop
```

Implementing Ingress Marking

Accurate DSCP marking of ingress traffic at the access-layer switch is critical to ensure proper QoS service treatment as traffic traverses through the network. All classified and policed traffic must be explicitly marked using the policy-map configuration based on an 8-class QoS model as shown in Figure 32.

Best practice is to use a explicit marking command (set dscp) even for trusted application classes (like VVLAN-VOIP and VVLAN-SIGNALING), rather than a trust policy-map action. A trust statement in a policy map requires multiple hardware entries, while the use of an explicit (seemingly redundant) marking command, improves the hardware efficiency.

The following sample configuration shows how to implement explicit marking for multiple classes on trusted and conditionally-trusted ingress ports in access-layer switches:

Trusted or Conditionally-Trusted Port

```
cr24-3560r-1(config)#policy-map Phone+PC-Policy
cr24-3560r-1(config-pmap)# class VVLAN-VOIP
cr24-3560r-1(config-pmap-c)# set dscp ef
cr24-3560r-1(config-pmap-c)# class VVLAN-SIGNALING
cr24-3560r-1(config-pmap-c)# set dscp cs3
cr24-3560r-1(config-pmap-c)# class MULTIMEDIA-CONFERENCING
cr24-3560r-1(config-pmap-c)# set dscp af41

cr24-3560r-1(config-pmap-c)# class SIGNALING
cr24-3560r-1(config-pmap-c)# set dscp cs3
cr24-3560r-1(config-pmap-c)# class TRANSACTIONAL-DATA
cr24-3560r-1(config-pmap-c)# set dscp af21
cr24-3560r-1(config-pmap-c)# class BULK-DATA
cr24-3560r-1(config-pmap-c)# set dscp af11
cr24-3560r-1(config-pmap-c)# class SCAVENGER
cr24-3560r-1(config-pmap-c)# set dscp cs1
cr24-3560r-1(config-pmap-c)# class DEFAULT
```

```
cr24-3560r-1(config-pmap-c)# set dscp default
```

All ingress traffic (default class) from an untrusted endpoint must be marked without a explicit classification. The following sample configuration shows how to implement explicit DSCP marking:

Untrusted Port

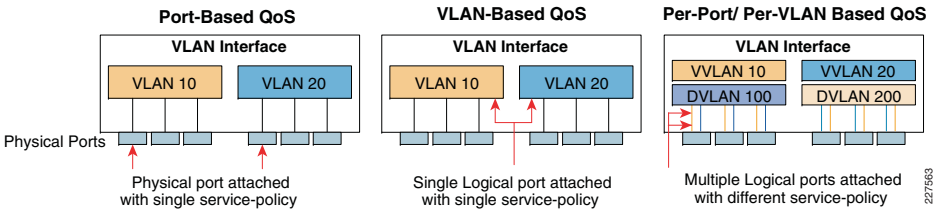
```
cr24-3560r-1(config)#policy-map UnTrusted-PC-Policy
cr24-3560r-1(config-pmap)# class class-default
cr24-3560r-1(config-pmap-c)#  set dscp default
```

Applying Ingress Policies

After creating a complete policy-map with all the QoS policies defined, the service-policy must be applied on the edge interface of the access-layer to enforce the QoS configuration. Cisco Catalyst switches offer three simplified methods to apply service-policies. Depending on the deployment model, any of these methods may be used:

- Port-based QoS—Applying service-policy on a per physical port basis will force traffic to pass-through the QoS policies before entering the network. Port-based QoS functions on a per-physical port basis even if the port is associated with a logical VLAN.
- VLAN-based QoS—Applying service-policy on per VLAN basis requires the policy-map to be attached to a logical Layer-3 SVI interface. Every physical port associated with the VLAN will require an extra configuration to enforce the QoS policies defined on a logical interface.
- Per-Port/Per-VLAN-based QoS—Not supported on all the Catalyst platforms and the configuration commands are platform-specific. Per-port/per-VLAN-based QoS creates a nested hierarchical policy-map that operates on a trunk interface. A different policy-map can be applied on each logical SVI interface that is associated to a single physical port.

Figure 33 Depicts All Three QoS Implementation Method



The following sample configuration shows how to deploy port-based QoS on the access-layer switches:

```
cr24-3560r-1(config)#interface fastethernet0/4
cr24-3560r-1(config-if)# description CONNECTED TO PHONE+PC

    cr24-3560r-1(config-if)# service-policy input Phone+PC-Policy
```

```
cr24-3560r-1#show policy-map interface f0/4 | inc Service|Class
Service-policy input: Phone+PC-Policy
Class-map: VVLAN-VOIP (match-all)
Class-map: VVLAN-SIGNALING (match-all)
Class-map: MULTIMEDIA-CONFERENCING (match-all)
Class-map: SIGNALING (match-all)
Class-map: TRANSACTIONAL-DATA (match-all)
Class-map: BULK-DATA (match-all)
Class-map: SCAVENGER (match-all)
 Class-map: DEFAULT (match-all)
    Class-map: class-default (match-any)
```

Applying Ingress Queueing

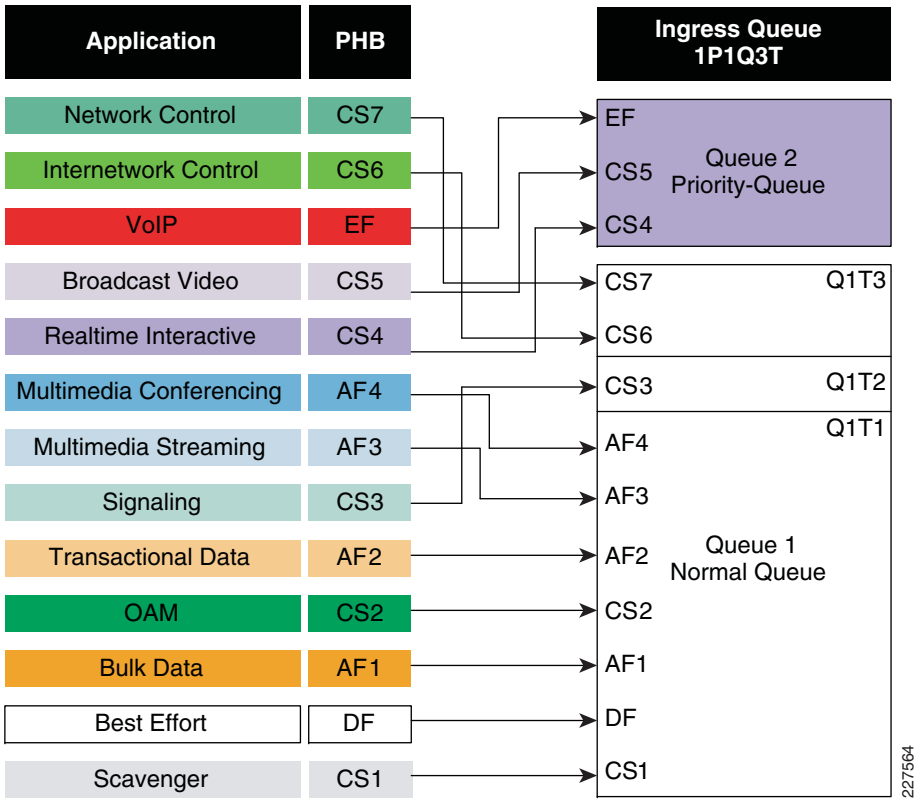
Fixed configuration Cisco Catalyst switches (2960 and 3xxx) not only offer differentiated services on the network ports but also internally on the switching fabric. Note, Cisco Catalyst 2960-S Series platform do not support ingress queueing and buffer allocation. After enabling QoS and attaching inbound policies on the physical ports, all the packets that meet the specified policy are forwarded to the switching fabric for egress switching. The aggregate bandwidth from all edge ports may exceed switching fabric bandwidth and cause internal congestion.

Cisco Catalyst 2960 and 3xxx platforms support two internal ingress queues: normal queue and priority queue. The ingress queue inspects the DSCP value on each incoming frame and assigns it to either the normal or priority queue. High priority traffic, like DSCP EF marked packets, are placed in the priority queue and switched before processing the normal queue.

The Catalyst 3750-X family of switches supports the weighted tail drop (WTD) congestion avoidance mechanism. WTD is implemented on queues to manage the queue length. WTD drops packets from the queue, based on dscp value, and the associated threshold. If the threshold is exceeded for a given internal DSCP value, the switch drops the packet. Each queue has three threshold values. The internal DSCP determines which of the three threshold values is applied to the frame. Two of the three thresholds are configurable (explicit) and one is not (implicit). This last threshold corresponds to the tail of the queue (100 percent limit).

Figure 34 depicts how different class-of-service applications are mapped to the Ingress Queue structure (1P1Q3T) and how each queue is assigned a different WTD threshold.

Figure 34 Catalyst 2960 and 3xxx Ingress Queueing Model



The DSCP marked packets in the policy-map must be assigned to the appropriate queue and each queue must be configured with the recommended WTD threshold as defined in Figure 34. The following ingress queue configuration must be enabled in global configuration mode on every access-layer switch.

```
cr25-3750-1(config)#mls qos srr-queue input priority-queue 2 bandwidth 30
 ! Q2 is enabled as a strict-priority ingress queue with 30% BW

cr25-3750-1(config)#mls qos srr-queue input bandwidth 70 30
 ! Q1 is assigned 70% BW via SRR shared weights
 ! Q1 SRR shared weight is ignored (as it has been configured as a PQ)

cr25-3750-1(config)#mls qos srr-queue input threshold 1 80 90
 ! Q1 thresholds are configured at 80% (Q1T1) and 90% (Q1T2)
 ! Q1T3 is implicitly set at 100% (the tail of the queue)
 ! Q2 thresholds are all set (by default) to 100% (the tail of Q2)

! This section configures ingress DSCP-to-Queue Mappings
cr25-3750-1(config)# mls qos srr-queue input dscp-map queue 1 threshold 1
0 8 10 12 14
 ! DSCP DF, CS1 and AF1 are mapped to ingress Q1T1
cr25-3750-1(config)# mls qos srr-queue input dscp-map queue 1 threshold 1
16 18 20 22
 ! DSCP CS2 and AF2 are mapped to ingress Q1T1
cr25-3750-1(config)# mls qos srr-queue input dscp-map queue 1 threshold 1 26 28 30
34 36 38
 ! DSCP AF3 and AF4 are mapped to ingress Q1T1
cr25-3750-1(config)#mls qos srr-queue input dscp-map queue 1 threshold 2
24
 ! DSCP CS3 is mapped to ingress Q1T2
cr25-3750-1(config)#mls qos srr-queue input dscp-map queue 1 threshold 3
48 56
 ! DSCP CS6 and CS7 are mapped to ingress Q1T3 (the tail of Q1)
cr25-3750-1(config)#mls qos srr-queue input dscp-map queue 2 threshold 3
32 40 46
 ! DSCP CS4, CS5 and EF are mapped to ingress Q2T3 (the tail of the PQ)


cr25-3750-1#show mls qos input-queue
Queue:       12
-----------------------------------
buffers   :9010
bandwidth :7030
priority  :030
threshold1:80100
threshold2:90100

cr25-3750-1#show mls qos maps dscp-input-q
   Dscp-inputq-threshold map:
     d1 :d2    0        1         2         3          4         5
6          7         8         9

------------------------------------------------------------------------
---------------
      0 :    01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01
      1 :    01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01
      2 :    01-01 01-01 01-01 01-01 01-02 01-01 01-01 01-01 01-01 01-01
      3 :    01-01 01-01 02-03 01-01 01-01 01-01 01-01 01-01 01-01 01-01
      4 :    02-03 02-01 02-01 02-01 02-01 02-01 02-03 02-01 01-03 01-01
      5 :    01-01 01-01 01-01 01-01 01-01 01-01 01-03 01-01 01-01 01-01
      6 :    01-01 01-01 01-01 01-01
```

Deploying Egress QoS

The QoS implementation for egress traffic toward the network edge on access-layer switches is much simpler than the ingress traffic QoS. The egress QoS implementation provides optimal queueing policies for each class and sets the drop thresholds to prevent network congestion and application performance impact. Cisco Catalyst switches support 4 hardware queues that are assigned the following policies:

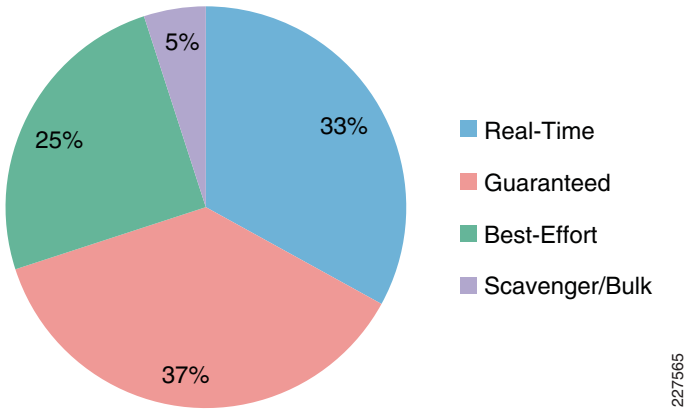
- Real-time queue (to support a RFC 3246 EF PHB service)
- Guaranteed bandwidth queue (to support RFC 2597 AF PHB services)
- Default queue (to support a RFC 2474 DF service)
- Bandwidth constrained queue (to support a RFC 3662 scavenger service)

As a best practice each physical or logical link must diversify bandwidth assignment to map with hardware queues:

- Real-time queue should not exceed 33 percent of the link's bandwidth.
- Default queue should be at least 25 percent of the link's bandwidth.
- Bulk/scavenger queue should not exceed 5 percent of the link's bandwidth.

Figure 35 shows best practice egress queue bandwidth allocation for each class.

Figure 35 Engress QoS

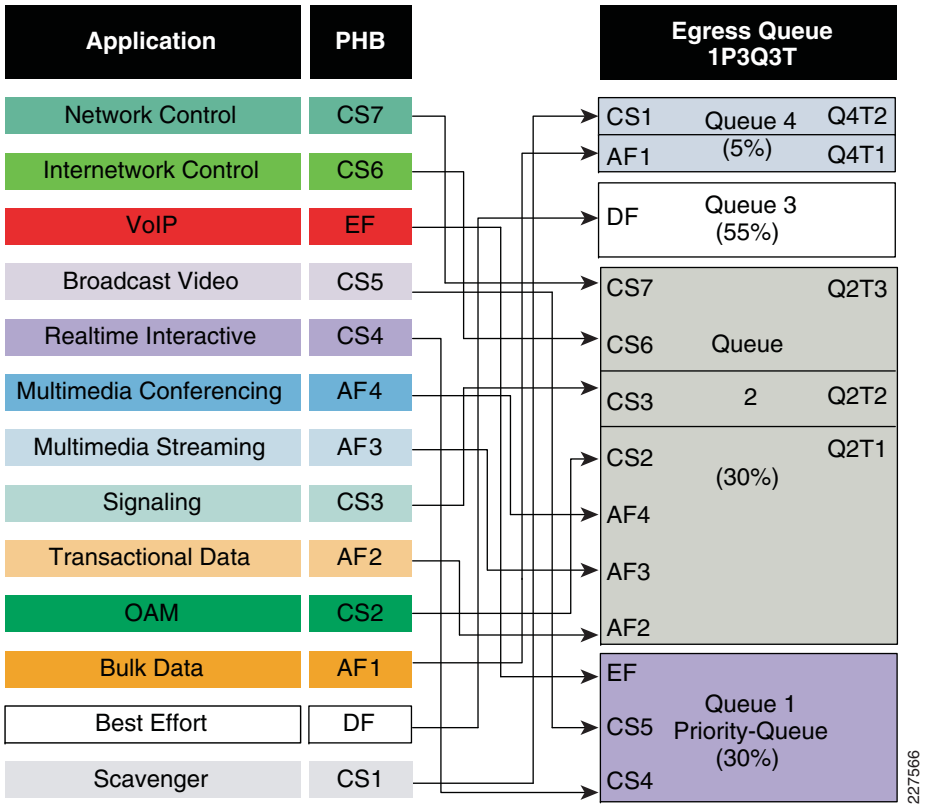


Given these minimum queuing requirements and bandwidth allocation recommendations, the following application classes can be mapped to the respective queues:

- Realtime Queue—Voice, broadcast video, and realtime interactive may be mapped to the realtime queue (per RFC 4594).
- Guaranteed Queue—Network/internetwork control, signaling, network management, multimedia conferencing, multimedia streaming, and transactional data can be mapped to the guaranteed bandwidth queue. Congestion avoidance mechanisms (i.e., selective dropping tools), such as WRED, can be enabled on this class. If configurable drop thresholds are supported on the platform, these may be enabled to provide intra-queue QoS to these application classes, in the respective order they are listed (such that control plane protocols receive the highest level of QoS within a given queue).
- Scavenger/Bulk Queue—Bulk data and scavenger traffic can be mapped to the bandwidth-constrained queue and congestion avoidance mechanisms can be enabled on this class. If configurable drop thresholds are supported on the platform, these may be enabled to provide inter-queue QoS to drop scavenger traffic ahead of bulk data.
- Default Queue—Best effort traffic can be mapped to the default queue; congestion avoidance mechanisms can be enabled on this class.

The egress queueing is designed to map traffic, based on DSCP value, to four egress queues. as shown above. The egress QoS model for a platform that supports DSCP-to-queue mapping with a 1P3Q8T queuing structure is depicted in Figure 36.

Figure 36 Access-Layer 1P3Q3T Egress Queue Model



DSCP marked packets are assigned to the appropriate queue and each queue is configured with appropriate WTD threshold as defined in Figure 36. Egress queueing is the same on network edge port as well as on uplink connected to internal network, and it is independent of trust mode. The following egress queue configuration in global configuration mode, must be enabled on every access-layer switch in the network.

```
cr25-3750-1(config)#mls qos queue-set output 1 buffers 15 30 35 20
! Queue buffers are allocated
cr25-3750-1(config)#mls qos queue-set output 1 threshold 1 100 100 100
100
```

```
! All Q1 (PQ) Thresholds are set to 100%
cr25-3750-1(config)#mls qos queue-set output 1 threshold 2 80 90 100 400
! Q2T1 is set to 80%; Q2T2 is set to 90%;
 ! Q2 Reserve Threshold is set to 100%;
 ! Q2 Maximum (Overflow) Threshold is set to 400%
cr25-3750-1(config)#mls qos queue-set output 1 threshold 3 100 100 100
400
! Q3T1 is set to 100%, as all packets are marked the same weight in Q3
 ! Q3 Reserve Threshold is set to 100%;
 ! Q3 Maximum (Overflow) Threshold is set to 400%
cr25-3750-1(config)#mls qos queue-set output 1 threshold 4 60 100 100 400
! Q4T1 is set to 60%; Q4T2 is set to 100%
 ! Q4 Reserve Threshold is set to 100%;
 ! Q4 Maximum (Overflow) Threshold is set to 400%


 ! This section configures egress DSCP-to-Queue mappings
cr25-3750-1(config)# mls qos srr-queue output dscp-map queue 1 threshold
3 32 40 46
 ! DSCP CS4, CS5 and EF are mapped to egress Q1T3 (tail of the PQ)
cr25-3750-1(config)# mls qos srr-queue output dscp-map queue 2 threshold
1 16 18 20 22
 ! DSCP CS2 and AF2 are mapped to egress Q2T1

cr25-3750-1(config)# mls qos srr-queue output dscp-map queue 2 threshold 1 26 28 30
34 36 38
 ! DSCP AF3 and AF4 are mapped to egress Q2T1
cr25-3750-1(config)#mls qos srr-queue output dscp-map queue 2 threshold 2
24
 ! DSCP CS3 is mapped to egress Q2T2
cr25-3750-1(config)#mls qos srr-queue output dscp-map queue 2 threshold 3
48 56
 ! DSCP CS6 and CS7 are mapped to egress Q2T3


cr25-3750-1(config)#mls qos srr-queue output dscp-map queue 3 threshold 3
0
 ! DSCP DF is mapped to egress Q3T3 (tail of the best effort queue)
cr25-3750-1(config)#mls qos srr-queue output dscp-map queue 4 threshold 1
8
 ! DSCP CS1 is mapped to egress Q4T1
cr25-3750-1(config)# mls qos srr-queue output dscp-map queue 4 threshold
2 10 12 14
 ! DSCP AF1 is mapped to Q4T2 (tail of the less-than-best-effort queue)

 ! This section configures interface egress queuing parameters
cr25-3750-1(config)#interface range GigabitEthernet1/0/1-48
cr25-3750-1(config-if-range)# queue-set 1
 ! The interface(s) is assigned to queue-set 1
cr25-3750-1(config-if-range)# srr-queue bandwidth share 1 30 35 5
 ! The SRR sharing weights are set to allocate 30% BW to Q2
 ! 35% BW to Q3 and 5% BW to Q4
 ! Q1 SRR sharing weight is ignored, as it will be configured as a PQ
cr25-3750-1(config-if-range)# priority-queue out
 ! Q1 is enabled as a strict priority queue

cr25-3750-1#show mls qos interface GigabitEthernet1/0/27 queueing
GigabitEthernet1/0/27
Egress Priority Queue : enabled
Shaped queue weights (absolute) :  25 0 0 0
Shared queue weights  :  1 30 35 5
The port bandwidth limit : 100  (Operational Bandwidth:100.0)
    The port is mapped to qset : 1
```

Table 10 and Table 11 summarize the ingress and egress QoS policies at the access-layer for several types of validated endpoints.

Table 10 Summarized Network Edge Ingress QoS Deployment Guidelines

Endpoint	Trust Model	DSCP Trust	Classification	Marking	Policing	Ingress Queueing
Unmanaged devices, printers, etc.	UnTrusted	Don't Trust. Default.	None	None	Yes	Yes
Managed secured devices, Servers etc	Trusted	Trust	8 Class Model	Yes	Yes	Yes
Phone	Trusted	Trust	Yes	Yes	Yes	Yes
Phone + Mobile PC	Conditionally -Trusted	Trust	Yes	Yes	Yes	Yes
IP Video surveillance Camera	Trusted	Trust	No	No	No	Yes
Digital Media Player	Trusted	Trust	No	No	No	Yes
Core facing Uplinks	Trusted	Trust	No	No	No	Yes

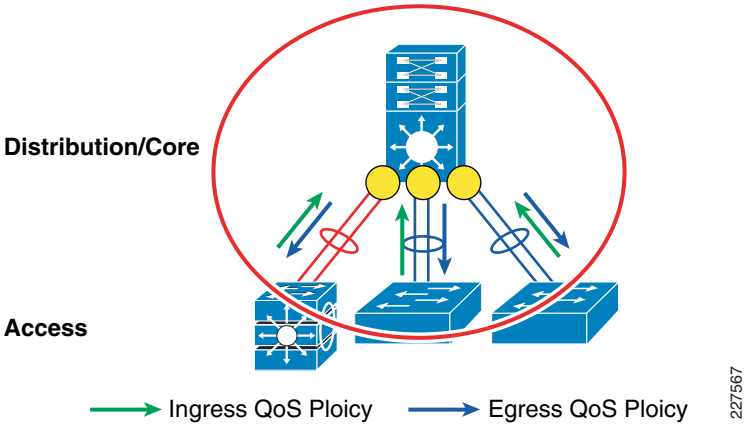
Table 11 Summarized Network Edge Egress QoS Deployment Guidelines

Endpoint	Trust Model	Classification / Marking / Policing	Egress Queueing	Bandwidth Share
Unmanaged devices, printers etc	UnTrusted	None	Yes	Yes
Managed secured devices, Servers etc	Trusted	None	Yes	Yes
Phone	Trusted	None	Yes	Yes
Phone + Mobile PC	Conditionally-Trusted	None	Yes	Yes
IP Video surveillance Camera	Trusted	None	Yes	Yes
Digital Media Player	Trusted	None	Yes	Yes
Core facing Uplinks	Trusted	None	Yes	Yes

Deploying Network Core QoS

All connections between internal network devices that are deployed within the network domain boundary are classified as trusted devices and follow the same QoS best practices recommended in the previous section. Ingress and egress core QoS policies are simpler than those applied at the network edge, See Figure 37.

Figure 37 Network Core QoS Boundary



The core network devices are considered trusted and rely on the access-switch to properly mark DSCP values. The core network is deployed to ensure consistent differentiated QoS service across the network. This ensures there is no service quality degradation for high-priority traffic, such as IP telephony or video.

The QoS implementation at the main and remote large site differ from the remote small site, due to different platforms used as the collapsed core router (Catalyst 4500-E vs Catalyst 3750-X StackWise Plus).

Deploying Main or Remote Large Site Ingress QoS

The collapsed core at main site is deployed with Cisco Catalyst 450R-E with Supervisor-6E, whereas the remote large site collapsed core is deployed with Cisco Catalyst 4507R-E with either Supervisor-6E, Supervisor-6LE or Supervisor-V. The next-generation Sup-6E and Sup-6LE module has a redesigned QoS implementation which matches Cisco IOS routers. No ingress QoS configuration is required, since QoS is enabled by default, and all ports are considered trusted.

The Cisco Catalyst 4507R-E with Supervisor-V requires ingress QoS configuration similar to trusted endpoints in the access-layer. Following is a sample configuration which enables QoS in the Catalyst 4507R-E with Supervisor-V:

```
cr35-4507-1(config)#qos
! Enables QoS function in the switch

cr35-4507-1#show qos
QoS is enabled globally
IP header DSCP rewrite is enabled
```

After QoS is globally enabled, all interfaces are in the untrusted mode by default. QoS trust settings must be set on each Layer 2 or Layer 3 port that is physically connected to another device within the network trust boundary. When Cisco Catalyst 4500 is deployed in EtherChannel mode, the QoS trust settings must be applied to every physical

member-link and logical port-channel interface. Best practice is to enable trust DSCP settings on each physical and logical interface that connects to another internal trusted device (e.g., access-layer switches in wiring closet or data-center, a router, wireless LAN controller (WLC)).

```
cr35-4507-1(config)#interface range Po11 , Gi1/2 , Gi2/2
cr35-4507-1(config-if-range)#description Connected to cr35-2960-1
cr35-4507-1(config-if-range)#qos trust dscp

cr35-4507-1#show qos interface Port-channel 11
QoS is enabled globally
Port QoS is enabled
Administrative Port Trust State: 'dscp'
Operational Port Trust State: 'dscp'
Trust device: none
Default DSCP: 0 Default CoS: 0
```

Additional ingress QoS techniques (such as classification, marking, and policing) are not required at the collapsed core layer since these functions are already performed by the access-layer switches. The architecture of Catalyst 4500-E with classic or next-generation Supervisor do not need ingress queueing since all of the forwarding decisions are made centrally on the supervisor. There are no additional QoS configurations required at the collapsed core-layer system.

Deploying Remote Small Site Ingress QoS

The remote small site is deployed using Cisco Catalyst 3750-X StackWise Plus as the collapsed core switch. The QoS implementation remains the same whether deployed as 3750-X StackWise or as a standalone switch. By default, QoS is disabled on the 3750-X switch. Following is a sample configuration to enable QoS in global configuration mode:

```
cr36-3750s-1(config)#mls qos
! Enables QoS function in the switch

cr36-3750s-1#show mls qos

QoS is enabled
QoS ip packet dscp rewrite is enabled
```

After QoS is globally enabled, all interfaces are in the untrusted mode by default. QoS trust settings must be set on each Layer 2 or Layer 3 port that is physically connected to another device within the network trust boundary. When Cisco Catalyst 3750-E StackWise Plus is deployed in EtherChannel mode, the QoS trust settings must be applied to every physical member-link. Best practice is to enable trust DSCP settings on each physical and logical interface that connects to another internal trusted device (e.g., access-layer switches in wiring closet or data-center, a router, wireless LAN controller (WLC)).

```
cr36-3750s-1(config)#int range gi1/0/49 , gi3/0/49
cr36-3750s-1(config-if-range)# description Connected to cr36-2960-1
cr36-3750s-1(config-if-range)#mls qos trust dscp

cr36-3750s-1#show mls qos interface Gi1/0/49
GigabitEthernet1/0/49
trust state: trust dscp
trust mode: trust dscp
trust enabled flag: ena
COS override: dis
default COS: 0
DSCP Mutation Map: Default DSCP Mutation Map
Trust device: none
qos mode: port-based
```

Additional ingress QoS techniques (such as classification, marking, and policing) are not required at the collapsed core layer since these functions are already performed by the access-layer switches. The ingress queueing and DSCP-Ingress-Queue function in 3750-X StackWise Plus must be enabled to allow differentiation between normal versus high-priority traffic. The ingress queuing configuration is consistent with the implementation at the access-edge. Following is a sample configuration for the ingress queues of the Catalyst 3750-X StackWise collapsed core switch:

```
cr36-3750-1(config)#mls qos srr-queue input priority-queue 2 bandwidth 30
 ! Q2 is enabled as a strict-priority ingress queue with 30% BW

cr36-3750-1(config)#mls qos srr-queue input bandwidth 70 30
 ! Q1 is assigned 70% BW via SRR shared weights
 ! Q1 SRR shared weight is ignored (as it has been configured as a PQ)

cr36-3750-1(config)#mls qos srr-queue input threshold 1 80 90
 ! Q1 thresholds are configured at 80% (Q1T1) and 90% (Q1T2)
 ! Q1T3 is implicitly set at 100% (the tail of the queue)
 ! Q2 thresholds are all set (by default) to 100% (the tail of Q2)

! This section configures ingress DSCP-to-Queue Mappings
cr36-3750-1(config)# mls qos srr-queue input dscp-map queue 1 threshold 1
0 8 10 12 14
 ! DSCP DF, CS1 and AF1 are mapped to ingress Q1T1
cr36-3750-1(config)# mls qos srr-queue input dscp-map queue 1 threshold 1
16 18 20 22
 ! DSCP CS2 and AF2 are mapped to ingress Q1T1
cr36-3750-1(config)# mls qos srr-queue input dscp-map queue 1 threshold 1 26 28 30
34 36 38
 ! DSCP AF3 and AF4 are mapped to ingress Q1T1
cr36-3750-1(config)#mls qos srr-queue input dscp-map queue 1 threshold 2
24
 ! DSCP CS3 is mapped to ingress Q1T2
cr36-3750-1(config)# mls qos srr-queue input dscp-map queue 1 threshold 3
48 56
```

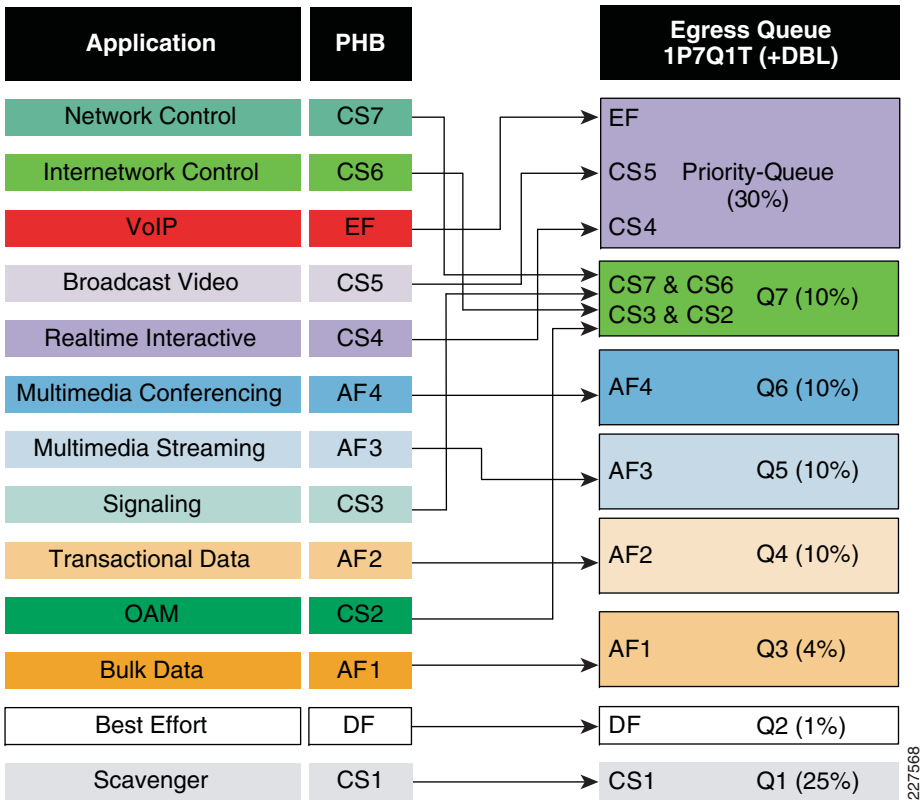
```
 ! DSCP CS6 and CS7 are mapped to ingress Q1T3 (the tail of Q1)
cr36-3750-1(config)# mls qos srr-queue input dscp-map queue 2 threshold 3
32 40 46
 ! DSCP CS4, CS5 and EF are mapped to ingress Q2T3 (the tail of the PQ)

cr36-3750s-1#show mls qos input-queue
Queue     :       1       2

-------------------------------------
buffers   :      90      10
bandwidth :      70      30
priority  :       0      30
threshold1:      80     100
threshold2:      90     100

cr36-3750s-1#show mls qos maps dscp-input-q
   Dscp-inputq-threshold map:
     d1 :d2    0        1         2        3          4         5
6          7         8         9

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

      0 :    01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01
      1 :    01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01 01-01
      2 :    01-01 01-01 01-01 01-01 01-02 01-01 01-01 01-01 01-01 01-01
      3 :    01-01 01-01 02-03 01-01 01-01 01-01 01-01 01-01 01-01 01-01
      4 :    02-03 02-01 02-01 02-01 02-01 02-01 02-03 02-01 01-03 01-01
      5 :    01-01 01-01 01-01 01-01 01-01 01-01 01-03 01-01 01-01 01-01
      6 :    01-01 01-01 01-01 01-01
```

Deploying Main Site Egress QoS

The main site is deployed with Cisco Catalyst 4507R-E with Supervisor-6E as the collapsed core router. Egress QoS from the collapsed core router provides optimized queueing and drop thresholds to drop excess low-priority traffic and protect high-priority traffic.

The Supervisor-6E supports up to 8 traffic classes for QoS mapping. It also supports a platform-specific congestion avoidance algorithm to provide Active Queue Management (AQM) with Dynamic Buffer Limiting (DBL). DBL tracks the queue length for each traffic flow in the switch. When the queue length of a flow exceeds its limit, DBL drops packets or sets the Explicit Congestion Notification (ECN) bit in the TCP packet header. With 8 egress (1P7Q1T) queues and DBL capability in the Sup-6E, the bandwidth distribution for each class changes, as shown in Figure 38.

Figure 38 Small Enterprise Main Site Network



Implementing QoS policies on Sup-6E-based Catalyst 4500 platform follows IOS (MQC)-model. The egress QoS implementation bundles the queueing and policing functions on EtherChannel based networks. To provide low-latency for high priority traffic, all lower priority traffic must wait until the priority-queue is empty. Best practice includes implementing a policer along with the priority-queue to provide more fair treatment for all traffic.

The following sample configuration shows how to create an 8-class egress queueing model and protect from high-priority traffic consuming more bandwidth than global policies allow. The egress QoS service-policy must be applied to all the physical EtherChannel member-links connected to different service-blocks (i.e., WAN edge, serverfarm, access-layer switches, etc).

```
! Creating class-map for each classes using match dscp statement as
marked by edge systems
cr24-4507-1(config)#class-map match-all PRIORITY-QUEUE
cr24-4507-1(config-cmap)# match  dscp ef
cr24-4507-1(config-cmap)# match  dscp cs5
cr24-4507-1(config-cmap)# match  dscp cs4
cr24-4507-1(config-cmap)#class-map match-all CONTROL-MGMT-QUEUE
cr24-4507-1(config-cmap)# match  dscp cs7
```

```
cr24-4507-1(config-cmap)# match  dscp cs6
cr24-4507-1(config-cmap)# match  dscp cs3
cr24-4507-1(config-cmap)# match  dscp cs2
cr24-4507-1(config-cmap)#class-map match-all
MULTIMEDIA-CONFERENCING-QUEUE
cr24-4507-1(config-cmap)# match  dscp af41  af42  af43
cr24-4507-1(config-cmap)#class-map match-all MULTIMEDIA-STREAMING-QUEUE
cr24-4507-1(config-cmap)# match  dscp af31  af32  af33
cr24-4507-1(config-cmap)#class-map match-all TRANSACTIONAL-DATA-QUEUE
cr24-4507-1(config-cmap)# match  dscp af21  af22  af23
cr24-4507-1(config-cmap)#class-map match-all BULK-DATA-QUEUE
cr24-4507-1(config-cmap)# match  dscp af11  af12  af13
cr24-4507-1(config-cmap)#class-map match-all SCAVENGER-QUEUE

cr24-4507-1(config-cmap)# match  dscp cs1


! Creating policy-map and configure queueing for class-of-service
cr24-4507-1(config)#policy-map EGRESS-POLICY
cr24-4507-1(config-pmap)# class PRIORITY-QUEUE
cr24-4507-1(config-pmap-c)# priority
cr24-4507-1(config-pmap-c)# class CONTROL-MGMT-QUEUE
cr24-4507-1(config-pmap-c)# bandwidth remaining percent 10
cr24-4507-1(config-pmap-c)# class MULTIMEDIA-CONFERENCING-QUEUE
cr24-4507-1(config-pmap-c)# bandwidth remaining percent 10
cr24-4507-1(config-pmap-c)# class MULTIMEDIA-STREAMING-QUEUE
cr24-4507-1(config-pmap-c)# bandwidth remaining percent 10
cr24-4507-1(config-pmap-c)# class TRANSACTIONAL-DATA-QUEUE
cr24-4507-1(config-pmap-c)# bandwidth remaining percent 10
cr24-4507-1(config-pmap-c)# dbl
cr24-4507-1(config-pmap-c)# class BULK-DATA-QUEUE
cr24-4507-1(config-pmap-c)# bandwidth remaining percent 4
cr24-4507-1(config-pmap-c)# dbl
cr24-4507-1(config-pmap-c)# class SCAVENGER-QUEUE
cr24-4507-1(config-pmap-c)# bandwidth remaining percent 1
cr24-4507-1(config-pmap-c)# class class-default
cr24-4507-1(config-pmap-c)# bandwidth remaining percent 25
cr24-4507-1(config-pmap-c)# dbl

! Attaching egress service-policy on all physical member-link ports
cr24-4507-1(config)#int range Gi1/1 - 6 , Gi2/1 - 6
cr24-4507-1(config-if-range)# service-policy output EGRESS-POLICY
```

EtherChannel is an aggregated logical bundle interface that does not perform queueing and relies on individual member-links to queue egress traffic. The policer to rate-limit priority class traffic must be implemented on EtherChannel and not on individual member-links since it governs the aggregate egress traffic limits. The following additional policy-map must be created to classify priority-queue class traffic and rate-limit the traffic to 30 percent of egress link capacity:

```
cr24-4507-1(config)#class-map match-any PRIORITY-QUEUE
cr24-4507-1(config-cmap)# match  dscp ef
cr24-4507-1(config-cmap)# match  dscp cs5
cr24-4507-1(config-cmap)# match  dscp cs4

cr24-4507-1(config)#policy-map PQ-POLICER
cr24-4507-1(config-pmap)# class PRIORITY-QUEUE
cr24-4507-1(config-pmap-c)# police cir 300 m conform-action transmit
exceed-action drop

cr24-4507-1(config)#interface range Port-Channel 1 , Port-channel 11 - 17
cr24-4507-1(config-if-range)#service-policy output PQ-POLICER
```

Deploying Remote Large Site Egress QoS

The remote large site is deployed with Cisco Catalyst 450R-E and either Supervisor-6E, Supervisor-6LE or Supervisor-V as the collapsed core router. If the remote large site network is deployed with Sup-6E or Sup-6LE, then the configuration is the same as described in the previous section.

The QoS deployment and implementation guidelines differ when the Cisco Catalyst 4500-E is deployed with the classic Supervisor-V module. The SupV supervisor can have up to four egress queues like the Cisco Catalyst 2960, 2960-S and 35xx/37xx-X Series switches. Before forwarding egress traffic, each packet must be internally classified and placed in the appropriate egress-queue. Placing traffic into different class-of-service queues, will offer traffic prioritization and guaranteed bandwidth to the network. The following sample configuration shows how to implement egress QoS on the Catalyst 4500-E with Supervisor-V:

```
cr35-4507-1(config)#qos dbl
 ! DBL is globally enabled
cr35-4507-1(config)#no qos dbl dscp-based 32
cr35-4507-1(config)#no qos dbl dscp-based 40
cr35-4507-1(config)#no qos dbl dscp-based 46
 ! DBL is explicitly disabled on DSCP CS4, CS5 and EF
 ! as these DSCP values are assigned to the PQ
 ! and as such should never experience congestion avoidance drops
cr35-4507-1(config)#qos dbl exceed-action ecn
 ! DBL will mark IP ECN bits in the event of congestion

! This section configures the DBL policy-map
cr35-4507-1(config)#policy-map DBL
cr35-4507-1(config-pmap)# class class-default
cr35-4507-1(config-pmap-c)# dbl
 ! DBL is enabled on all flows
```

```
 ! (with the exception of DSCP CS4, CS5 and EF)
 ! This section configures the DSCP-to-Queue mappings


cr35-4507-1(config)#qos map dscp 8 10 12 14 to tx-queue 1
 ! DSCP CS1 and AF1 are mapped to Q1 (the less than best effort queue)
cr35-4507-1(config)#qos map dscp 0 to tx-queue 2
 ! DSCP DF is mapped to Q2 (the best effort/default queue)
cr35-4507-1(config)#qos map dscp 32 40 46 to tx-queue 3
 ! DSCP CS4, CS5 and EF are mapped to Q3 (the PQ)
cr35-4507-1(config)#qos map dscp 16 18 20 22 to tx-queue 4
 ! DSCP CS2 and AF2 are mapped to Q4 (guaranteed BW queue)
cr35-4507-1(config)#qos map dscp 24 26 28 30 to tx-queue 4
 ! DSCP CS3 and AF3 are mapped to Q4 (guaranteed BW queue)
cr35-4507-1(config)#qos map dscp 34 36 38 to tx-queue 4
 ! DSCP AF4 is mapped to Q4 (guaranteed BW queue)
cr35-4507-1(config)#qos map dscp 48 56 to tx-queue 4
 ! DSCP CS6 and CS7 are mapped to Q4 (guaranteed BW queue)

 ! This section configures all the EtherChannel member-link for egress
queuing
cr35-4507-1(config)#interface range Gig1/1 - 6 , Gig2/1 - 6
cr35-4507-1(config-if-range)# tx-queue 1
cr35-4507-1(config-if-tx-queue)# bandwidth percent 5
 ! Q1 (less than best effort queue) is assigned 5% BW
cr35-4507-1(config-if-tx-queue)# tx-queue 2
cr35-4507-1(config-if-tx-queue)# bandwidth percent 35
 ! Q2 (default/best effort queue) is assigned 35% BW
cr35-4507-1(config-if-tx-queue)# tx-queue 3
cr35-4507-1(config-if-tx-queue)# priority high
cr35-4507-1(config-if-tx-queue)# bandwidth percent 30
 ! Q3 is enabled as a PQ and assigned 30% BW
cr35-4507-1(config-if-tx-queue)# tx-queue 4
cr35-4507-1(config-if-tx-queue)# bandwidth percent 30
 ! Q4 (guaranteed BW queue) is assigned 30% BW
cr35-4507-1(config-if-range)# service-policy output DBL
 ! DBL policy-map is attached to the interface(s)

cr35-4507-1#show qos dbl
QOS is enabled globally

DBL is enabled globally on DSCP values:
     0-31,33-39,41-45,47-63
DBL flow includes vlan
DBL flow includes layer4-ports
DBL uses ecn to indicate congestion
DBL exceed-action probability: 15%
DBL max credits: 15
DBL aggressive credit limit: 10
DBL aggressive buffer limit: 2 packets


cr35-4507-1#show qos maps dscp tx-queue
DSCP-TxQueue Mapping Table (dscp = d1d2)
d1 : d2  0   1   2    3   4   5   6    7   8   9
------------------------------------------------
 0 :    02 01 01 01 01 01 01 01 01 01
 1 :    01 01 01 01 01 01 04 02 04 02
 2 :    04 02 04 02 04 02 04 02 04 02
 3 :    04 02 03 03 04 03 04 03 04 03
 4 :    03 03 03 03 03 03 03 03 04 04
 5 :    04 04 04 04 04 04 04 04 04 04
 6 :    04 04 04 04

cr35-4507-1#show qos interface Gig1/2
QoS is enabled globally
Port QoS is enabled
Administrative Port Trust State: 'dscp'
Operational Port Trust State: 'dscp'
Trust device: none
Default DSCP: 0 Default CoS: 0
Appliance trust: none
Tx-Queue    Bandwidth    ShapeRate    Priority    QueueSize
              (bps)        (bps)                  (packets)
  1         50000000     disabled    N/A         2080
  2         350000000     disabled    N/A2080
  3         300000000disabled    high2080
  4         300000000disabledN/A2080
```

Deploying Remote Small Site Egress QoS

Collapsed Core— Catalyst 3750-X StackWise Plus

The remote small site is deployed with Cisco Catalyst 3750-X StackWise Plus as the collapsed core router. The Catalyst 3750-X can have up to four egress queues. Before forwarding egress traffic, each packet is placed in the appropriate egress-queue as shown in Figure 36. The Catalyst 3750-E switch supports Shaped Round Robin (SRR) packet schedule service which can be deployed in two different modes:

- Shaped—To provide guaranteed bandwidth, the shaped egress queue reserves some of the bandwidth of the port for each queue. Traffic load exceeding the shape parameter gets dropped. The queue cannot take advantage of excess bandwidth capacity when other queues are not using their bandwidth allocations.

- Shared—Shared mode also provides guaranteed bandwidth for each queue; however, it allows the flexibility of using excess bandwidth when there is any available.

The following sample configuration shows how to implement egress QoS on the Catalyst 3750-X:

```
cr36-3750s-1(config)#mls qos queue-set output 1 buffers 15 30 35 20
! Queue buffers are allocated
cr36-3750-1(config)#mls qos queue-set output 1 threshold 1 100 100 100
100
! All Q1 (PQ) Thresholds are set to 100%
cr36-3750s-1(config)#mls qos queue-set output 1 threshold 2 80 90 100 400
! Q2T1 is set to 80%; Q2T2 is set to 90%;
 ! Q2 Reserve Threshold is set to 100%;
 ! Q2 Maximum (Overflow) Threshold is set to 400%
cr36-3750s-1(config)#mls qos queue-set output 1 threshold 3 100 100 100
400
! Q3T1 is set to 100%, as all packets are marked the same weight in Q3
 ! Q3 Reserve Threshold is set to 100%;
 ! Q3 Maximum (Overflow) Threshold is set to 400%
cr36-3750s-1(config)#mls qos queue-set output 1 threshold 4 60 100 100
400
! Q4T1 is set to 60%; Q4T2 is set to 100%
 ! Q4 Reserve Threshold is set to 100%;
 ! Q4 Maximum (Overflow) Threshold is set to 400%

 ! This section configures egress DSCP-to-Queue mappings
cr36-3750s-1(config)# mls qos srr-queue output dscp-map queue 1 threshold
3 32 40 46
 ! DSCP CS4, CS5 and EF are mapped to egress Q1T3 (tail of the PQ)
cr36-3750s-1(config)# mls qos srr-queue output dscp-map queue 2 threshold
1 16 18 20 22
 ! DSCP CS2 and AF2 are mapped to egress Q2T1
cr36-3750s-1(config)#mls qos srr-queue output dscp-map queue 2 threshold 1 26 28 30
34 36 38
 ! DSCP AF3 and AF4 are mapped to egress Q2T1
cr36-3750s-1(config)#mls qos srr-queue output dscp-map queue 2 threshold
2 24
 ! DSCP CS3 is mapped to egress Q2T2
cr36-3750s-1(config)#mls qos srr-queue output dscp-map queue 2 threshold
3 48 56
 ! DSCP CS6 and CS7 are mapped to egress Q2T3
cr36-3750s-1(config)#mls qos srr-queue output dscp-map queue 3 threshold
3 0
 ! DSCP DF is mapped to egress Q3T3 (tail of the best effort queue)
cr36-3750s-1(config)#mls qos srr-queue output dscp-map queue 4 threshold
1 8
 ! DSCP CS1 is mapped to egress Q4T1
cr36-3750s-1(config)# mls qos srr-queue output dscp-map queue 4 threshold
2 10 12 14
 ! DSCP AF1 is mapped to Q4T2 (tail of the less-than-best-effort queue)

 ! This section configures interface egress queuing parameters
cr36-3750s-1(config)#interface range GigabitEthernet1/0/1-48
cr36-3750s-1(config-if-range)# queue-set 1
 ! The interface(s) is assigned to queue-set 1
cr36-3750s-1(config-if-range)# srr-queue bandwidth share 1 30 35 5
 ! The SRR sharing weights are set to allocate 30% BW to Q2
 ! 35% BW to Q3 and 5% BW to Q4
 ! Q1 SRR sharing weight is ignored, as it will be configured as a PQ
cr36-3750s-1(config-if-range)# priority-queue out
 ! Q1 is enabled as a strict priority queue

cr36-3750s-1#show mls qos interface GigabitEthernet1/0/49 queueing
GigabitEthernet1/0/49
Egress Priority Queue : enabled
Shaped queue weights (absolute) :  25 0 0 0
Shared queue weights  :  1 30 35 5
The port bandwidth limit : 100  (Operational Bandwidth:100.0)
    The port is mapped to qset : 1
```

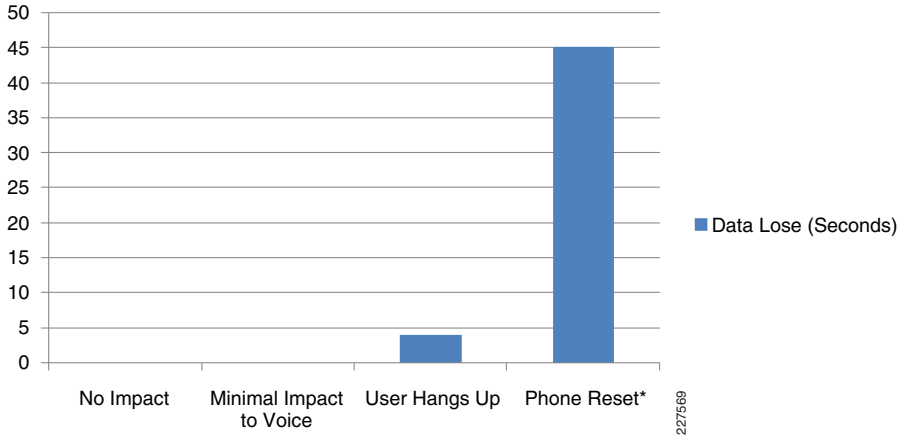
Building a Resilient Network

The Small Enterprise Design Profile is a high performance, resilient and scalable network design. A network outage may be caused by the system, human error, or natural disaster. The small enterprise network is designed to minimize the impact of a failure regardless of the cause. Network outages may be either planned or unplanned.

- Planned Outage—Planned network outage occurs when a portion of the network is taken out of service as part of a scheduled event (e.g., a software upgrade).
- Unplanned Outage—Any unscheduled network outage is considered an unplanned outage. Such outages may be caused by internal faults in the network, or devices due to hardware or software malfunctions.

The network is designed to recover from most un planned outages in less than a second (milliseconds). In many situations, the user will not even notice the outage occurred. If the outage lasts longer (several seconds), then the user will notice the lack of application responsiveness. The network is designed to minimize the overall impact of a unplanned network outage, and gracefully adjust and recover from many outage conditions. Figure 39 shows an example of a real-time VoIP application and user impact depending on duration of outage event.

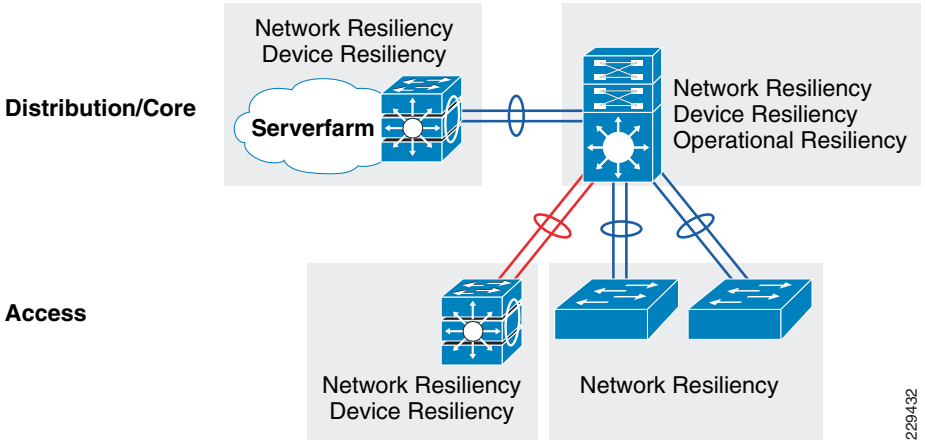
Figure 39 VoIP User Impact for Minor and Major Network Outage



Several techniques are used to make the network design more resilient. Deploying redundant devices and redundant connections between devices, enables the network to recover from fault conditions. Identifying critical versus non critical applications, and network resources optimizes cost performance, by focusing on the most important elements of the network design. The resiliency of a system design is often categorized as follows:

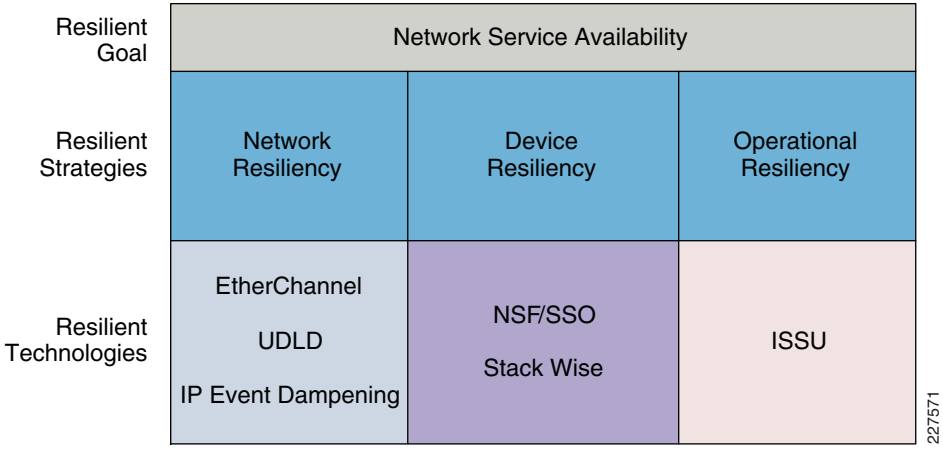
- Network Resiliency—Provides redundancy during physical link outages (e.g., fiber cut, bad transceivers, incorrect cabling, etc).
- Device Resiliency—Protects network during device outage triggered by hardware or software (e.g. software crash, non-responsive supervisor, etc).
- Operational Resiliency—Capabilities which provide network availability even during planned network outage conditions (e.g., ISSU features which enable software upgrades while device is operating).

Figure 40 Resiliency Deployment Strategy



The high availability framework is based upon the three resiliency categories described in the previous section. Figure 41 shows which technologies are implemented to achieve each category of resiliency.

Figure 41 High-Availability Categories and Technologies



Redundant Hardware Components

Redundant hardware implementations vary between fixed configuration and modular Cisco Catalyst switches. Selective deployment of redundant hardware is an important element of the Small Enterprise Design Profile which delivers device resiliency.

Redundant hardware component for device resiliency varies between fixed configuration and modular Cisco Catalyst switches. To protect against common network faults or resets, all critical main and remote site campus network devices must be deployed with similar device resiliency configuration. This subsection provides a basic redundant hardware deployment guideline at the access-layer and collapsed core switching platforms in the campus network.

Redundant Power System

Redundant power supplies protect the device from power outage or power supply failure. Protecting the power is not only important for the network device, but also the endpoints that rely on power delivery over the Ethernet network. Redundant power supplies are deployed differently depending on the switch type:

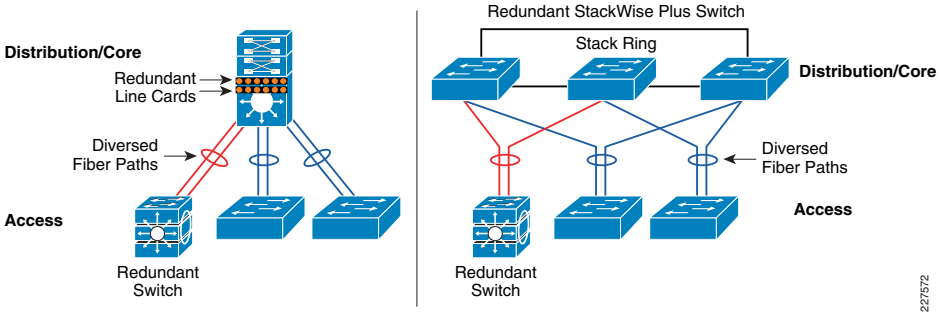
- Modular Switch—Dual power supplies can be deployed in the modular switching platforms like the Cisco Catalyst 4500-E. By default, the Cisco Catalyst 4500 power supply operates in 1+1 redundant mode (both power supplies are active).
- Fixed Configuration Switch—Fixed configuration switches are deployed with internal power supplies and they may also use Cisco RPS 2300 external power supply. A single Cisco RPS 2300 power supply has modular power supplies and fans to deliver power to multiple switches. Deploying internal and external power-supplies provides a redundant power solution for fixed configuration switches.

Redundant Network Connectivity

Redundant network connections protect the system from failure due to cable or transceiver faults. Redundant network connections attached to a single fixed configuration switch or network module in the Cisco Catalyst 4500 switch do not protect against internal device hardware or software fault.

Best practice design is to deploy redundant network modules within the Catalyst 4500 switch and the Cisco 3750-X StackWise Plus solution in the small remote site collapsed core network. Deploying the 3750-X StackWise Plus in critical access-layer switches in the serverfarm network and in the main site is also best practice. Connecting redundant paths to different hardware elements provides both network and device resiliency.

Figure 42 Redundant Network Connectivity



Redundant Control-Plane

The processing software operation is different in standalone or StackWise fixed configuration switches, and on a supervisor module of a modular switch. Network communication and forwarding operations can be disrupted when the processing unit fails, causing a network outage. Network recovery techniques vary based on the different platforms.

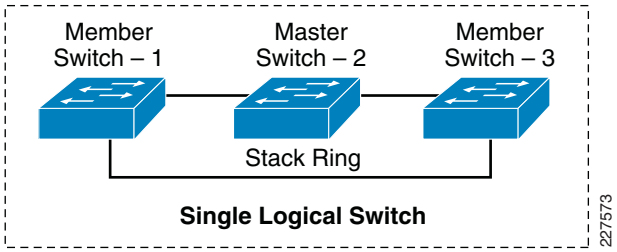
The standalone and non-stackable fixed configuration switches like the Cisco Catalyst 2960 or 3560-E feature power redundancy and network resiliency support; however they do not protect against a processing unit failure. During a processing unit failure event, all endpoints attached to the switch are impacted and network recovery time is undeterministic.

Device resiliency in Cisco StackWise and modular switching platforms provides 1+1 redundancy with enterprise-class high availability and deterministic network recovery time.

Cisco StackWise Plus

Cisco Catalyst 3750-X switches can be deployed in StackWise mode using a special stack cable. Up to nine switches can be integrated into a single stack that delivers distributed forwarding architecture and unified single control and management plane. Device level redundancy in StackWise mode is achieved via stacking multiple switches using the Cisco StackWise technology. One switch from the stack is selected automatically to serve as the master, which manages the centralized control-plane process. Cisco StackWise solution provides 1:N redundancy. In the event of a active master-switch outage, a new master is selected automatically. See Figure 43.

Figure 43 Cisco Stack Wise Switching Architecture



Since Cisco StackWise enables up to 9 switches to appear as one logical switch, it has centralized management and control functions. Most Layer 2 and Layer 3 functions are centrally performed, however Layer-2 topology development is distributed (i.e., each switch performs the function independently). Table 12 lists network protocol functions and identifies which are centralized and which are distributed.

Table 12 Cisco StackWise Centralized and Distributed Control-Plane

Protocols		Function
Layer 2 Protocols	MAC Table	Distributed
	Spanning-Tree Protocol	Distributed
	CDP	Centralized
	VLAN Database	Centralized
	EtherChannel - LACP	Centralized
Layer 3 Protocols	Layer 3 Management	Centralized
	Layer 3 Routing	Centralized

Cisco StackWise Plus solution offers network and device resiliency with distributed forwarding. In the event of a master switch outage, Non-Stop Forwarding (NSF) enables packet forwarding to continue based on current state information, while a new master switch is selected. New master switch selection is accomplished in the range of 700 to 1000 milliseconds; the amount of time to reestablish the control-plane and develop distributed forwarding will vary depending on the size and complexity of the network.

Following is a best practice to reduce Layer-3 disruption in the event of a master switch outage: Determine the master switch with the higher switch priority, and isolate the uplink Layer-3 EtherChannel bundle path by using physical ports from member switches (i.e. don't use the master switches ports for Etherchannel uplinks). With NSF capabilities enabled, this design decreases network downtime during a master-switch outage.

An understanding of SSO and StackWise components and failover events associated with NSF provides significant insight in designing a network that enables supervisor redundancy. The following subsection uses the above concepts and principles to identify the design parameters, and applies them to develop a best-practice hierarchical network with the highest availability.

Cisco FlexStack

The next-generation Catalyst 2960-S Series Layer 2 access-switch introduces high-speed, low-latency stacking capability based on "pay as you grow" model. Following the Catalyst 3750-X StackWise Plus success, the Catalyst 2960-S model offers high availability, increased port-density with unified single control-plane and management to reduce the cost for small enterprise network. However, the architecture of FlexStack on Catalyst 2960-S Series platform differs from StackWise Plus. The Cisco FlexStack is comprised with hardware module and software capabilities. The FlexStack module must be installed on each Catalyst 2960-S switches that are intended to be deployed in stack-group. Cisco FlexStack module is hot-swappable module providing flexibility to deploy FlexStack without impacting business network operation.

Cisco FlexStack allows up to four Catalyst 2960-S Series switches into a single stacking group; it is recommended to deploy each switch member with dual FlexStack cable to provide increased 20G bidirectional stack bandwidth capacity and FlexStack redundancy. The FlexStack protocol dynamically detects switch member and allows it to join the stack group if all stacking criteria is met. The unique data forwarding architecture and FlexStack QoS is on per-hop basis, the unknown unicast, broadcast, and multicast traffic will be flooded between stack group switch members. The FlexStack protocol detects and breaks the loop between the FlexStack group switches. Once the destination switch member is determined, Catalyst 2960-S use shortest egress stack port path to forward traffic. Any packet traverses across FlexStack is encapsulated with 32 bytes of FlexStack header carrying unique information to provide centralized control-plane and distributed forwarding design.

Cisco Modular Switch

The Cisco Catalyst 4500-E modular switch supports redundant supervisors, and Stateful Switch Over (SSO). When deployed along with NSF, the 4500-E provides a enterprise-class highly available system with network and device resiliency.

SSO is a Cisco IOS service used to synchronize critical forwarding and protocol state information between redundant supervisors configured in a single chassis. With SSO enabled, one supervisor in the system assumes the role of active and the other supervisor becomes the hot-standby. Each is ready to backup the other, thus providing 1:1 hot redundancy to protect from a control-plane outage. Since both supervisors are active, the system benefits by using the physical ports from both supervisors during normal operation. SSO synchronizes system services such as DHCP snooping, Switched Port Analyzer (SPAN), security access control lists (ACLs), and QoS policies so ensure the switch provides the same level of protection and service after a supervisor failover event.

NSF enables packets to continue to be forwarded using existing routing table information, during switchover. NSF also provides graceful restart to the routing protocol such that during the failover, the routing protocol remains aware of the change and does not react by resetting its adjacency. If the routing protocol were to react to the failure event, and alter routing path information, the effectiveness of stateful switch over would be diminished.

Operational Resiliency Strategy

Designing the network to recover from unplanned outages is important. It is also important to consider how to minimize the disruption caused by planned outages. These planned outages can be due to standard operational processes, configuration changes, software and hardware upgrades, etc.

The same redundant components which mitigate the impact of unplanned outages can also be used to minimize the disruption caused by planned outages. The ability to upgrade individual devices without taking them out of service is enabled by having internal component redundancy (such as with power supplies, and supervisors) complemented with the system software capabilities. Two primary mechanisms exist to upgrade software in a live network:

- Full-image In-Service Software Upgrade (ISSU) on the Cisco Catalyst 4500-E leverages dual supervisors to allow for a full, in-place Cisco IOS upgrade. This leverages the NSF/SSO capabilities of the switch and provides for less than 200 msec of traffic loss during a full Cisco IOS upgrade.
- Network and device level redundancy, along with the necessary software control mechanisms, guarantee controlled and fast recovery of all data flows following a fault condition, and provide the ability to manage the fault tolerant infrastructure during planned outage events.

Validating operational resiliency is beyond the scope of this design guide, refer to CCO documentation for deployment guidelines.

Deploying High Availability in Network

Many of the design features of the Small Enterprise Design Profile which were described in "Deploying Network Foundation Services" section on page -8, contribute to the network high availability capabilities. This section focuses on how to implement additional features which complete the high availability design in small enterprise network design.

Network Resiliency

Etherchannel and UDLD are two design features which are included in the network foundation services, which contribute to network resiliency.

Implementing IP Event Dampening

Poor signaling or a loose connection may cause continuous port-flap (port alternates between active state and inactive state). A single interface flapping can impact the stability and availability of the network. Route summarization is one technique which mitigates the impact of a flapping port. Summarization isolates the fault domain with a new metric announcement by the aggregator and thereby hides the local networks fault within the domain.

A best practice to mitigate local network domain instability due to port-flap, is implementing IP Event Dampening on all layer 3 interfaces. Each time the Layer-3 interface flaps the IP dampening tracks and records the flap event. Upon multiple flaps, a logical penalty is assigned to the port and suppresses link status notification to IP routing until the port becomes stable. IP event dampening is a local function and does not have a signaling mechanism to communicate with a remote system. It can be implemented on each individual physical or logical Layer-3 interface: physical ports, SVI or port-channels. Following is a example configuration to implement IP Event Damenping:

Distribution/Core

```
cr24-4507-1(config)#int range Port1 , Gig5/6 , Gig6/6 , Vlan 101 - 110
cr24-4507-1(config-if-range)#dampening

    cr24-4507-1#show interface dampening | be Port
```

```
    Port-channel1 Connected to cr24-3750ME-1
    Flaps Penalty    Supp ReuseTm   HalfL  ReuseV   SuppV  MaxSTm    MaxP
Restart
      0       0   FALSE       0            5       1000       2000
20            16000       0
```

The following output illustrates how the IP event dampening keeps track of port flaps and makes a decision to notify IP routing process based on interface suppression status:

```
cr24-4507-1#debug dampening interface
cr24-4507-1#show logging | inc EvD|IF-EvD


12:32:03.274: EvD(GigabitEthernet5/6): charge penalty 1000, new accum.
penalty 1000, flap count 2
12:32:03.274: EvD(GigabitEthernet5/6): accum. penalty 1000, not
suppressed
12:32:03.274: IF-EvD(GigabitEthernet5/6): update IP Routing state to
DOWN, interface is not suppressed


cr24-4507-1#show interface dampening | be 5/6
Flaps Penalty    Supp ReuseTm   HalfL  ReuseV   SuppV  MaxSTm    MaxP
Restart
2       0   FALSE       0           5      1000       2000      20
16000       0
```

In a multilayer access-distribution design, the Layer-2 and Layer-3 demarcation is at the collapsed core-distribution device. IP event dampening is enabled on per-logical VLAN (SVI) interface basis on the collapsed core device. IP event dampening becomes more effective when each access-layer switch is deployed with a unique set of Layer-2 VLANs.

Assigning unique VLANs on each access-layer switch also helps IP event dampening to isolate the problem and prevent network faults triggered in a multilayer network. The following output illustrates how IP event dampening keeps track of individual logical VLAN networks associated to same Layer-2 physical trunk ports. When a Layer-2 trunk port flaps, the state of SVI also flaps, and forces dampening to track and penalize unstable interfaces:

```
12:58:41.332: EvD(Vlan101): charge penalty 1000, new accum. penalty 2627,
flap count 3
12:58:41.332: EvD(Vlan101): accum. penalty 2627, now suppressed with a
reuse intervals of 7
12:58:41.332: IF-EvD(Vlan101): update IP Routing state to DOWN, interface
is suppressed


    cr24-4507-1#show interface dampening
    Vlan101 Connected to cr24_2960_Dept_1_VLAN
    Flaps Penalty    Supp ReuseTm   HalfL  ReuseV   SuppV  MaxSTm    MaxP
Restart
      3       71   FALSE       0           5      1000       2000      20
16000       0
```

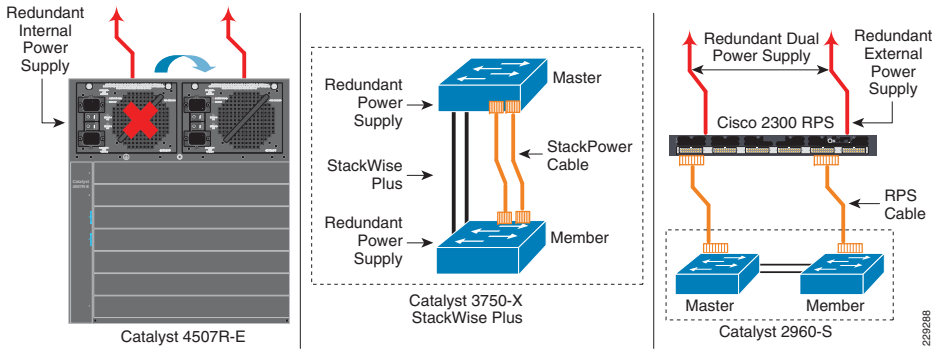
Device Resiliency

As described earlier, redundant hardware is an important technique for achieving device resiliency. The small enterprise network design applies hardware redundancy considering the cost/performance tradeoffs.

Implementing Redundant Power Supply

Redundant power supplies can prevent a system outage due to power outage, power supply or fan hardware failure. All Cisco Catalyst switching platforms supports robust 1+1 redundant power capabilities that can be deployed with internal or external power source management. See Figure 44.

Figure 44 Cisco Catalyst Internal and External Power Redundancy Option



Catalyst 4500-E—Redundant Internal Power Supply

The Cisco Catalyst 4500-E provides power to internal hardware components and external devices like IP phones. All the power is provided by the internal power supply. Dual-power supplies in the Catalyst 4500-E can operate in one of two different modes:

- Redundant Mode—By default, Catalyst 4500-E power supply operates in redundant mode offering 1+1 redundant option. The system determines power capacity and number of power supplies required based on the power required for all internal and external power components. Both power supplies must have sufficient power to support all the installed modules and operate in 1+1 redundant mode.

```
cr24-4507-1(config)#power redundancy-mode redundant


cr24-4507-1#show power supplies
Power supplies needed by system    :1
Power supplies currently available :2
```

- Combined Mode—If the system power requirement exceeds the capacity of a single power supply, then both power supplies can be combined to increase the capacity. In this mode, the power system does not provide 1+1 power redundancy. The following global configuration will enable power supplies to operate in combined mode:

```
cr24-4507-1(config)#power redundancy-mode combined

cr24-4507-1#show power supplies
Power supplies needed by system:2
Power supplies currently available:2
```

Catalyst 3750-X— Cisco StackPower Redundancy

The next-generation Catalyst 3750-X Series platform introduces innovative Cisco StackPower technology to provide power redundancy solution for fixed configuration switches. Cisco StackPower unifies the individual power supplies installed in the switches and creates a pool of power, directing that power where it is needed. Up to four switches can be configured in a StackPower stack with the special Cisco proprietary StackPower cable. The StackPower cable is different than the StackWise data cables and is available on all Cisco Catalyst 3750-X models.

During individual power supply fault from the stack can regain power from global power pool to provide seamless operation in the network. With the modular power supply design in Catalyst 3750-X series platform the defective power supply can be swapped without disrupting network operation. The Cisco StackPower can be deployed in following two modes:

- Sharing Mode—All input power is available to be used for power loads. The total aggregated available power in all switches in the power stack (up to four) is treated as a single large power supply. All switches in stack can share power with available power to all powered devices connected to PoE ports. In this mode, the total available power is used for power budgeting decisions and no power is reserved to accommodate power-supply failures. If a power supply fails, powered devices and switches could be shut down. Default mode.
- Redundant Mode—The power from the largest power supply in the system is subtracted from the power budget, which reduces the total available power, but provides backup power in case of a power-supply failure. Although there is less available power in the pool for switches and powered devices to draw from, the possibility of having to shut down switches or powered devices in case of a power failure or extreme power load is reduced. It is recommended to budget the required power and deploy each Catalyst 3750-X switch in stack with dual power supply to meet the need. Enabling redundant mode will offer power redundancy as a backup during one of the power supply unit failure event.

Since Cisco StackWise Plus can group up to nine 3750-X Series switches in the stack-ring, the Cisco StackPower must be deployed with two power stack group to accommodate up to four switches. Following sample configuration demonstrate deploying Cisco StackPower redundancy mode and grouping the stack-member into power stack group, to make new power configuration effective, it is important that network administrator must plan downtime as all the switches in the stack ring must be reloaded:

```
cr36-3750X-xSB(config)# stack-power stack PowerStack
cr36-3750X-xSB(config-stackpower)#mode redundant

cr36-3750X-xSB(config)#stack-power switch 1
cr36-3750X-xSB(config-switch-stackpower)#stack-id PowerStack
%The change may not take effect until the entire data stack is reloaded

cr36-3750X-xSB(config)#stack-power switch 2
cr36-3750X-xSB(config-switch-stackpower)#stack-id PowerStack
%The change may not take effect until the entire data stack is reloaded
```

Catalyst 2960 (External Power Redundancy)

The Cisco Redundant Power Supply (RPS) 2300 can support up to six RPS ports to provide seamless power backup to critical access-layer switches in the campus network. Upto two devices can be protected by Cisco RPS 2300 against power or power supply failure. Additional power resiliency on Cisco RPS 2300 can be added by deploying dual power supply to backup to two devices simultaneously. Note that external power redundancy requires special RPS cable and specific 2960 models currently do not support external power redundancy. Deploying external power redundancy on Cisco Catalyst 2960 and 2960-S with Cisco RPS 2300 is performed automatically and do not require any extra configuration. Cisco Catalyst 3560-X and 3750-X switches can be used if RPS 2300 configuration is required.

Implementing Redundant Control Plane System

The collapsed core device in the main and remote sites (Catalyst 4500-E or 3750-X StackWise Plus) is deployed with redundant supervisor, or StackWise Plus to enable graceful recovery from switch hardware outage. Any access-switch which is deemed critical may be deployed as StackWise Plus and FlexStack to improve device resiliency. The implementation for each switch is different, and is discussed separately in the sections which follow.

Resilient Cisco FlexStack and StackWise Plus

Starting in Cisco IOS Release 12.2(53)SE1, Cisco Catalyst 2960-S supports FlexStack and it can be used when higher port-density, increased uplink bandwidth capacity, and the resiliency at Layer-2 access is required. Starting in Cisco IOS Release 12.2(53)SE2, Cisco Catalyst 3750-X supports StackWise Plus, and is used when a resilient Layer 2 or Layer 3 access-switch is required. Cisco 3750-X StackWise Plus is deployed for the collapsed core in the small remote site network.

Cisco Catalyst 3750-X and 2960-S switches are provisioning dynamically in the stack group by the StackWise or FlexStack protocols. Cisco IOS automatically adjusts the interface addressing and its associated configuration based on the number of provisioned switches in the stack.

```
cr26-2960s-1#show run | inc provision

switch 1 provision ws-c2960s-48ts-s
switch 2 provision ws-c2960s-48ts-s
```

Master Switch Election

The centralized control-plane and management plane is managed by the master switch in the stack. By default, the master switch selection within the ring is performed dynamically by negotiating several parameters and capabilities between each switch within the stack. Each StackWise-capable switch is by default configured with priority 1.

```
cr26-C2960S-1# show switch
Switch/Stack Mac Address : 0022.bdc4.1d80
                                           H/W   Current
Switch#  Role   Mac Address     Priority Version  State
-----------------------------------------------------------
*1       Master 0022.bdc4.1d80     1       1       Ready
 2       Member 0026.0ac1.3e00     1       1       Ready
```

As described in previous section, the Cisco StackWise Plus architecture is not SSO-capable. This means that all the centralized Layer-3 functions must be reestablished with the neighbor switch during a master-switch outage. To minimize the control-plane impact and improve network convergence, the Layer-3 uplinks should be diverse, originating from member switches instead of the master switch. The default switch priority must be increased manually after identifying the master switch and switch number. The new switch priority becomes effective after switch reset. It is recommended to modify default switch priority on the Catalyst 2960-S FlexStack group as Catalyst 3750-X StackWise.

```
cr26-3750r-1(config)#switch 1 priority 15
Changing the Switch Priority of Switch Number 1 to 15
cr26-3750r-1(config)#switch 2 priority 14
Changing the Switch Priority of Switch Number 2 to 15

cr26-3750r-1#show switch
Switch/Stack Mac Address : 0023.eb7b.e580
                                           H/W   Current
Switch# RoleMac AddressPriorityVersionState
------------------------------------------------------------
  1        Member0023.eb7b.e580150       Ready
* 2        Master0026.5284.ec80140       Ready
```

StackWise Layer-3 MAC Management

To provide a single unified logical network view in the network, the MAC addresses of Layer-3 interfaces on the StackWise (physical, logical, SVIs, port channel) are derived from the Ethernet MAC address pool of the master switch in the stack. All the Layer-3 communication from the StackWise switch to the endpoints (like IP phone, PC, servers and core network system) is based on the MAC address pool of the master switch.

```
cr26-3750r-1#show switch
Switch/Stack Mac Address : 0026.5284.ec80
                                           H/W   Current
Switch# Role   Mac AddressPriority     Version       State
-----------------------------------------------------------------------------
----------------
    1Member0023.eb7b.e580     1              0          Ready

* 2 Master  0026.5284.ec80      1              0         Ready

cr26-3750r-1#show version
.  .  .
Base ethernet MAC Address       : 00:26:52:84:EC:80
.  .  .
```

After a master-switch outage, the new master switch in the stack assigns new MAC addresses to all Layer-3 interfaces, from the local MAC address pool. Once the new MAC address is assigned, it will force the switch to generate a gratuitous ARP in the network to make sure no other system is using the same MAC address. The default timer to retain the MAC address from the failed master switch is four minutes. While the new MAC address is not assigned on Layer-3 interface and not being propagated and updated in the network, the traffic will blackhole in the network.

```
cr26-3750r-1#reload slot 2
Proceed with reload? [confirm]
Switch 2 reloading...

cr26-3750r-1#show switch
Switch/Stack Mac Address : 0023.eb7b.e580
                                           H/W   Current
Switch# Role   Mac AddressPriority     Version       State
-----------------------------------------------------------------------------
----------------
* 1 Master  0023.eb7b.e580     1              1          Ready
   2Member000.0000.0000      0              1         Removed
```

To prevent this network instability, the old MAC address assignments on Layer-3 interfaces can be retained even after the master switch fails. The new active master switch can continue to use the MAC addresses assigned by the old master switch, which prevents ARP and routing outages in the network. The default stack-mac timer settings must be changed in Cisco Catalyst 2960-S FlexStack and 3750-E StackWise Plus switch mode using the global configuration CLI mode as shown below:

```
cr26-3750r-1(config)#stack-mac persistent timer 0

cr26-3750r-1#show switch
Switch/Stack Mac Address : 0023.eb7b.e580
Mac persistency wait time: Indefinite

                                           H/W   Current
Switch# Role   Mac AddressPriority     Version       State
```

```
------------------------------------------------------------------------
---------------
    1Master0023.eb7b.e580     1               0          Ready
* 2 Master  0026.5284.ec80      1               0          Ready
```

Non-Stop Forwarding (NSF)

The Cisco Catalyst 3750-X switch in StackWise Plus mode is not SSO-capable. When the master switch fails, the new master switch is required to reform the Layer-3 adjacencies with the neighbors in the network. The forwarding architecture in StackWise switch is designed to provide non-stop forwarding during the master switch outage using NSF technology. Each 3750-X switch in the stack maintains distributed Layer-3 FIB from the old master switch and continues to forward upstream traffic, until they are updated by the new master in the stack ring.

To enable NSF capability, explicit configuration must be enabled under the routing process. NSF-aware feature is enabled by default on all Layer-3 Ethernet switches to function in helper mode to perform graceful recovery during NSF-capable Cisco 3750-X master switch outage. NSF-capable system can also operate in NSF aware role:

NSF Capable Layer-3 Switch

```
cr36-3750s-1(config)#router eigrp 100
cr36-3750s-1(config-router)#nsf

cr36-3750s-1#show ip protocols | inc NSF
*** IP Routing is NSF aware ***
  EIGRP NSF-aware route hold timer is 240
  EIGRP NSF enabled
     NSF signal timer is 20s
     NSF converge timer is 120s
```

NSF-Aware Layer-3 Switch

```
cr24-3560r-1#show ip protocols | inc NSF
*** IP Routing is NSF aware ***
  EIGRP NSF-aware route hold timer is 240
```

NSF Timers

As depicted in the above show commands, the default NSF-aware system hold timer is 240 seconds. Lowering the timer value may abruptly terminate graceful recovery, causing network instability. Best practice is to use the default NSF hold timer, unless it is observed that NSF recovery takes longer than 240 seconds.

600 seconds after a graceful-recovery starts on a NSF-aware system, NSF clears the route stale marking and resumes using the synchronized routing database.

```
! NSF Aware received graceful-restart message from new master switch
11:56:15.365: %DUAL-5-NBRCHANGE: EIGRP-IPv4:(100) 100: Neighbor
10.125.32.3 (Port-channel15) is resync: peer graceful-restart
11:56:15.365: EIGRP: NSF: AS100, NSF or GR initiated by 10.125.32.3 at
00:00:00, flags 0x4

! NSF route hold timer expires and searches and removes all stale route
entries
received graceful-restart message from new master switch
12:00:15.392: EIGRP: NSF: AS100. route hold timer expiry
12:00:15.392: DUAL: Search for outdated routes from 10.125.32.3
```

Resilient Cisco Catalyst 4500-E

A modular switching platform like the Cisco Catalyst 4507R-E is fully NSF/SSO-capable, providing 1+1 control plane redundancy. In the Catalyst 4507R-E, all the intelligent Layer-2 and Layer-3 functions are performed centrally on the supervisor module. Deploying redundant supervisor in SSO mode in same system will allow the primary supervisor to fully synchronize the adjacencies, forwarding, configuration, counters, and more information on redundant hot-standby supervisor.

The Cisco Catalyst 4507R-E ports are independent of the supervisor state. Because of this hardware design, during a supervisor switchover, the ports connected to the failed supervisor do not go down. Because paths and ports are not down, hardware keeps forwarding the packet to a valid next-hop while supervisor switchover is occurring.

The configuration and implementation guidelines for implementing NSF/SSO on the Cisco Catalyst 4507R-E are the same for main and remote site network designs.

Increasing Supervisor Uplink Port Availability

There are restrictions on which supervisor uplink ports can be actively configured. Multiple ports can be simultaneously active on the supervisor. However Cisco IOS Release 12.2(25)SG or later is required for concurrent use of both 10G and 1G. The Small Enterprise Design Profile uses the 1G interface to connect to the Cisco 3750-MetroE WAN aggregation switch. To use 10G port in 1G mode with redundancy, the following configuration must be applied on collapsed core Catalyst 4507R-E switch:

```
cr24-4507-1(config)#hw-module uplink mode shared-backplane
cr24-4507-1(config)#hw-module module 3 port-group 1 select
gigabitethernet
cr24-4507-1(config)#hw-module module 4 port-group 1 select
gigabitethernet

cr24-4507-1#show hw-module uplink
 Active uplink mode configuration is Shared-backplane
```

Stateful Switchover (SSO)

SSO redundancy mode in the Cisco Catalyst 4507R-E supervisor is turned on by default starting with Cisco IOS Release 12.2(20)EWA. To provide 1+1 redundancy, all the technical specifications between active and standby supervisor must be identical. Also note that the Cisco Catalyst 4507R-E and 4510R-E are the only models that support supervisor redundancy. SSO is supported on all supervisors running IOS except Sup II-Plus-TS. The

NSF-awareness feature is supported by all the supervisors supporting EIGRP, OSPF, IS-IS, and BGP routing protocols, while the NSF-capable feature is supported only on supervisors IV, V, and V-10G. NSF/SSO on Catalyst 4500-E requires a minimum boot ROM version and must be the same on both supervisors. For additional details on hardware requirements, refer to the Release Notes at the following URL:

http://www.cisco.com/en/US/docs/switches/lan/catalyst4500/12.2/53SG/configuration/NSFwSSO.html#wp1135767

```
cr24-4507-1(config)#redundancy
cr24-4507-1(config-red)# mode sso
cr24-4507-1(config-red)# main-cpu
cr24-4507-1(config-r-mc)# auto-sync standard

cr24-4507-1#show module | inc Chassis | 6-E | SSO
Chassis Type : WS-C4507R-E


 3     6  Sup 6-E 10GE (X2), 1000BaseX (SFP)     WS-X45-SUP6-E
JAE1132SXQ3
 4     6  Sup 6-E 10GE (X2), 1000BaseX (SFP)     WS-X45-SUP6-E
JAE1132SXRQ

 3   Active Supervisor   SSOActive
 4   Standby Supervisor  SSOStandby hot
```

The active supervisor dynamically detects the secondary supervisor installed in the same chassis and initiates several SSO dependency configuration checks. If the SSO dependency check fails, then the standby supervisor falls back into RPR mode. For example, IOS release mismatch between two supervisors may not allow SSO to synchronize.

If the SSO dependency configuration checks successfully pass, then SSO communication between both supervisors goes through several synchronization states before it transitions to hot-standby state as illustrated in the following output:

```
cr24-4507-1#show redundancy states
my state = 13 -ACTIVE
peer state = 8  -STANDBY HOT
. . .

Redundancy Mode (Operational) = Stateful Switchover
Redundancy Mode (Configured)  = Stateful Switchover
Redundancy State              = Stateful Switchover
    Maintenance Mode = Disabled
    Manual Swact = enabled
    Communications = Up
    . . .
```

All the state-machines and dynamic information of SSO-capable protocols are automatically synchronized to the standby supervisor module without any additional operational requirement. The hot-standby supervisor takes over the ownership of control-plane process when the active supervisor outage or removal from the chassis is detected.

Non-Stop Forwarding (NSF)

All the state-machines and dynamic information of SSO-capable protocols are automatically synchronized to the standby supervisor module. The hot-standby supervisor takes over the ownership of control-plane process if the active supervisor suffers an outage or is removed from the chassis.

NSF-Capable Layer 3 Switch

```
cr24-4507-1(config)#router eigrp 100
cr24-4507-1 (config-router)#nsf

cr24-4507-1#show ip protocols | inc NSF
*** IP Routing is NSF aware ***
  EIGRP NSF-aware route hold timer is 240
  EIGRP NSF enabled
     NSF signal timer is 20s
     NSF converge timer is 120s
```

NSF-Aware Layer 3 Switch

```
cr24-3560r-1#show ip protocols | inc NSF
*** IP Routing is NSF aware ***
  EIGRP NSF-aware route hold timer is 240
```

Summary

Designing the LAN network aspects for the small enterprise network design establishes the foundation for all other aspects within the design profile including WAN, Security, and Mobility.

This chapter reviews LAN design models recommended by Cisco, as well as where to apply these models within the various locations of a small enterprise network. Each of the layers is discussed and design guidance is provided on where to place and how to deploy these layers. Finally, key network foundation services such as routing, switching, QoS, multicast, and high availability best practices are given for the entire small enterprise design.