

Features and Services

Revised: December 21, 2009

This chapter briefly describes all the services and features that are part of the Services Ready Large Branch Network design and that meet the business criteria outlined in “[Large Branch Design Considerations](#)” section on page 4. The building blocks of the Cisco Enterprise Branch Architecture framework are described as they apply to the Services Ready Large Branch Network.

Contents

- [Branch Network Components](#), page 1
- [WAN Services](#), page 9
- [LAN Deployment Model](#), page 22
- [Network Fundamentals](#), page 31
- [Security Services](#), page 53
- [Management Services](#), page 68
- [Voice Services](#), page 73
- [Optimization Services](#), page 92

Branch Network Components

Cisco offers a broad and versatile portfolio of routers, switches, and IP Phones. There are three product lines of routers and four product lines of switches for the branch office. Each product line offers different performance and features, enabling enterprise IT departments to meet a wide range of functional requirements. [Figure 1](#) provides an overview of the various Cisco Integrated Services Routers Generation 2 (Cisco ISRs G2) that are commonly deployed in the branch office.

Figure 1 Branch Office Integrated Services Router Portfolio

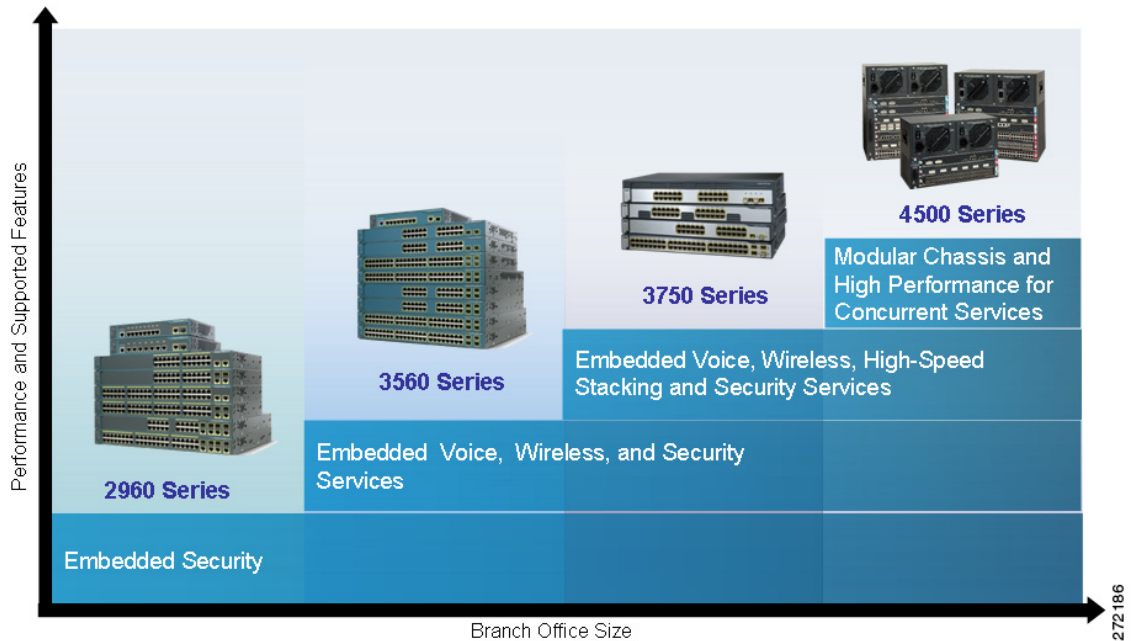


To learn more about each router product line, see the *Cisco Router Guide*:

http://www.cisco.com/en/US/prod/collateral/routers/ps5855/prod_brochure0900aecd8019dc1f.pdf

Figure 2 provides a high-level overview of the various Catalyst switches that are commonly deployed in the branch office.

Figure 2 Branch Office Catalyst Switch Portfolio

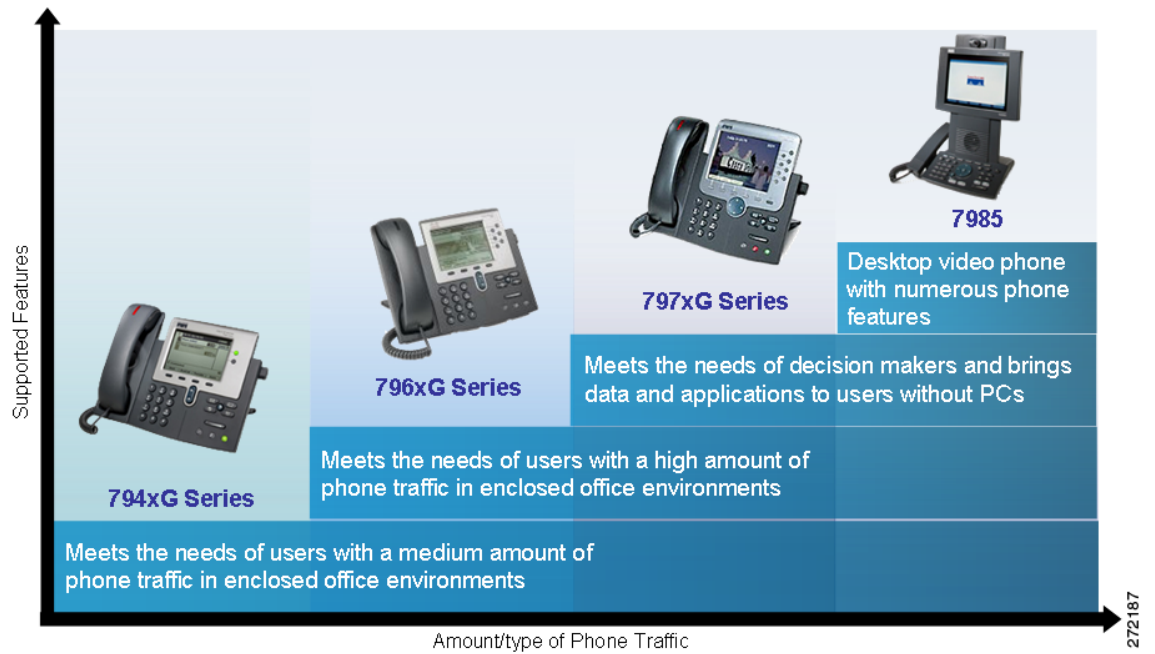


To learn more about each switch product line, see the *Cisco Catalyst Switch Guide*:

http://www.cisco.com/en/US/prod/switches/ps5718/ps708/networking_solutions_products_genericcontent0900aecd805f0955.pdf

There are four desktop IP Phone product lines that are suited for the branch office. Each phone offers different functions and capabilities, as shown in **Figure 3**.

Figure 3 Branch Office Cisco Unified IP Phone 7900 Series Portfolio



To learn more about each IP Phone, visit:

http://www.cisco.com/en/US/products/sw/voicesw/products_category_buyers_guide.html#number_1

Selecting Network Components

Selecting the appropriate routing and switching platforms for a branch office involves numerous considerations. The most important considerations are:

- Branch office size: The platform must support required port densities for the expected number of end-user devices.
- Features and services: The platform must support required networking services, interfaces, and modules.
- Performance: The platform, including features and services, must handle wire speeds required by branch applications.
- Scalability: The platform must have extra slots for DRAM, flash, interface and module expansion.
- Resiliency: The platform must support high availability and fault tolerance.

In accordance with the business criteria outlined in the “[Large Branch Design Considerations](#)” section on page 4, Cisco 3925 and Cisco 3945 Integrated Services Routers (ISRs) were selected for the Services Ready Large Branch Network.

The Cisco 3945 ISR, shown in [Figure 4](#), is ideal for medium-sized and large business and enterprise branch offices. It offers embedded video, WAN optimization, network awareness, voice, wireless, switching, and security features. Built for performance, it delivers multiple concurrent services at a wire speed of up to a T3/E3 rate. High availability is supported through online network module insertion and removal, redundant internal power supply, field-replaceable components, and Cisco IOS software features for redundant system design.

Figure 4 Cisco 3945 Integrated Services Router



To learn more about the Cisco 3945 ISR, visit:

<http://www.cisco.com/en/US/products/ps10541/index.html>

The Cisco 3925 ISR, shown in [Figure 5](#), offers similar functionality to that of the Cisco 3945 ISR. It differs from the Cisco 3945 ISR in the following ways:

- Performance: Half the wire speed with concurrent services (up to one-half the T3/E3 rate)
- Resiliency: Absence of redundant internal power supply
- Scalability: Two fewer network module slots

Figure 5 Cisco 3925 Integrated Services Router



To learn more about the Cisco 3925 ISR, visit:

<http://www.cisco.com/en/US/products/ps10542/index.html>

Catalyst 3560 and Catalyst 3750 series switches were selected for the Services Ready Large Branch Network. Several different models are available in each product family. The selection of a specific model depends on the desired number of ports, support for PoE, and Gigabit Ethernet connectivity, and will vary from enterprise to enterprise.

The Catalyst 3560 series switch, shown in [Figure 6](#) is an ideal access layer switch for branch-office environments, combining both 10/100/1000 and PoE configurations and enabling the deployment of new applications such as IP telephony, wireless access, and video surveillance. It offers Fast Ethernet and Gigabit Ethernet connectivity and concurrent QoS, ACL, port security, link aggregation, and VLAN functionality at forwarding rates of up to 32 Gb/s. For scalability, the Catalyst WS-C3560G-48PS-S and Catalyst WS-C3560G-48TS-S models provide up to forty-eight 10/100/1000 ports and four small form-factor pluggable (SFP) ports. The Catalyst WS-C3560G-48PS-S also adds the PoE option. The main selection criterion for the Catalyst 3560 switch is support for the PoE option; however, the Catalyst WS-C3560G-48TS-S model was tested to provide an option for connecting devices that do not require PoE.

Figure 6 Catalyst WS-C3560G-48PS-S/Catalyst WC-C3560G-48TS-S Switch



To learn more about the Catalyst 3560 switch series, visit:

<http://www.cisco.com/en/US/products/hw/switches/ps5528/index.html>

The Catalyst 3750 series switch, shown in [Figure 7](#), is an ideal distribution layer switch for large branch-office environments, combining 10/100/1000 configurations and Cisco StackWise technology. Cisco StackWise technology unites Catalyst 3750 switches into a single logical unit through special stack interconnect cables, achieving high throughput and availability. In addition to Cisco StackWise technology, the switch supports a redundant power supply for added resiliency. The Catalyst WS-C3750G-24TS-S offers Fast Ethernet and Gigabit Ethernet connectivity and concurrent QoS, ACL, port security, link aggregation, and VLAN functionality at forwarding rates of up to 32 Gb/s. It provides up to twenty-four 10/100/1000 ports and four SFPs. The Catalyst WS-C3750G-24PS-S offers the same functionality, and also adds PoE.

Figure 7 Catalyst WS-C3750G-24TS-S/Catalyst WS-C3750G-24PS-S Switch



To learn more about the Catalyst 3750 switch series, visit:

<http://www.cisco.com/en/US/products/hw/switches/ps5528/index.html>

Cisco offers a variety of IP Phones. Selection of the appropriate phone depends on its intended usage. The most important selection criteria for Cisco Unified 7900 Series office worker IP Phones are:

- Display: The applications used on the phone determine the need for backlight, color, and touch screen.
- Line count: The expected usage determines the required number of phone lines or telephony features.
- Physical features: The amount and type of phone traffic and the applications determine the required number of buttons, the functionality of the navigation wheel, and the need to support key expansion modules.
- Video: Video conferencing requires video capabilities.

When considering an IP Phone, in general, there are numerous other features to evaluate (e.g., QoS, codec). However, all office worker Cisco 7900 Series Unified IP Phones implement the same core features required of an enterprise class IP Phone. Therefore, the above criteria are the primary considerations when selecting from the various options. To learn more about the features of the Cisco Unified IP Phones, see the *Cisco Unified IP Phone Features A - Z*:

http://www.cisco.com/en/US/docs/voice_ip_comm/cuipph/all_models/phone_a_to_z/english/user/guide/az_user.html

Business criteria outlined in the “[Large Branch Design Considerations](#)” section on page 4 specify five different use cases for IP Phones in a branch office: moderate call volume user, heavy call volume user, decision maker, video conferencing user, and conference room. For each of the first three use cases two different phones were selected.

The Cisco Unified IP Phone 7942G and Cisco Unified IP Phone 7945G, shown in [Figure 8](#), were chosen for the moderate call-volume use case. Both phones support:

- High-fidelity audio
- High-resolution display for advanced XML applications and double-byte characters/Unicode
- IEEE 803.af PoE (Class 2) or local power supply
- Access to two phone lines (or combination of line access and telephony features)
- Integrated Ethernet switch and 10/100BASE-T Ethernet connection through an RJ-45 interface for LAN connectivity
- Standards-compliant Session Initiation Protocol (SIP) support.

In addition, the Cisco Unified IP Phone 7945G phone offers Gigabit Ethernet VoIP telephony technology and a large backlit color display.

Figure 8 Cisco Unified IP Phones 7942G and 7945G



The Cisco Unified IP Phone 7962G GE and Cisco Unified IP Phone 7965G, shown in [Figure 9](#), were selected for the high call-volume use case. Both phones support the same features and differences as the Cisco Unified IP Phone 7942G and Cisco Unified IP Phone 7945G phones, and both phones support four additional phone lines.

Figure 9 Cisco Unified IP Phones 7962G and 7965G



272193

The Cisco Unified IP Phone 7971G GE and Cisco Unified IP Phone 7975G, shown in [Figure 10](#), were selected for the decision-maker use case. Both phones support the following features:

- High-fidelity audio
- Gigabit Ethernet VoIP telephony technology
- Backlit high-resolution, color touch screen for easy access to communications information
- XML applications
- Integrated Ethernet switch and 10/100/1000BASE-T Ethernet connection via an RJ-45 interface for LAN connectivity
- IEEE 802.3af Power (Class 3) over Ethernet (PoE) or a local power supply
- Standards-compliant SIP phone support

In addition, the Cisco Unified IP Phone 7975G features a high-resolution screen, high-fidelity wideband audio, and Internet Low Bit Rate Codec (iLBC) support for use in lossy networks.

Figure 10 Cisco Unified IP Phones 7971G-GE and 7975G



272194

[Table 1](#) provides a high-level feature comparison of the six IP Phone models.

Table 1 Comparison of Cisco Unified IP Phone Models for Large Branch Offices

Use Case	Moderate Call Volume		Heavy Call Volume		Decision Maker	
	7942G	7945G	7962G	7965G	7971G-GE	7975G
Cisco Unified IP Phone	7942G	7945G	7962G	7965G	7971G-GE	7975G
Display	Grayscale	Color	Grayscale	Color	12-bit Color	16-bit Color
Touch screen	No	No	No	No	Yes	Yes
Wideband speaker	Yes	Yes	Yes	Yes	No	Yes
Wideband handset	Yes	Yes	Yes	Yes	Accessory	Yes
Wideband headset	Supported	Supported	Supported	Supported	Supported	Supported
iLBC	Yes	Yes	Yes	Yes	No	Yes
Navigation cluster	2-way	4-way + Select	2-way	4-way + Select	4-way	4-way + Select
Gigabit Ethernet	No	Yes	No	Yes	Yes	Yes
Line keys	2	2	6 (+KEM)	6 (+KEM)	6 (+KEM)	8 (+KEM)
KEM support ¹	No	No	Yes	Yes	Yes	Yes

1. KEM: Key Expansion Module.

The Cisco Unified IP Phone 7985G, shown in [Figure 11](#), was selected for the video-conferencing use case. The phone supports personal desktop video for instant, face-to-face communications, incorporates all the components required for video calls (camera, LCD screen, speaker, keypad, and handset), provides integrated Ethernet switch and 10/100BASE-T Ethernet connection through an RJ-45 interface for LAN connectivity, and has dedicated buttons that control the video features: Self View, Picture in Picture, Video Mute, Display, and Brightness.

Figure 11 Cisco Unified IP Phone 7985G

The Cisco Unified IP Conference Station 7936, shown in [Figure 12](#), was selected for the conference room scenario. The conference station offers a regular telephone keypad plus three soft keys, menu navigation keys, and a backlit, pixel-based LCD display.

Figure 12 Cisco Unified IP Conference Station 7936



WAN Services

A number of WAN technologies are available to meet the diverse business requirements of an enterprise. This guide does not address considerations and issues pertaining to enterprise WAN design. However, certain aspects of WAN deployment, such as basic connectivity and routing, affect configuration of the branch office router and influence the use of specific features and services in the branch network. To ensure its relevance and applicability, the Services Ready Large Branch Network was validated with the most commonly deployed enterprise WANs. For detailed guidance on WAN design and implementation see the Cisco WAN design documents at:

http://www.cisco.com/en/US/netsol/ns817/networking_solutions_program_home.html.

Today enterprises have five common WAN connectivity options for the branch office. Each option, as shown in [Figure 13](#), has its own set of benefits and trade-offs.

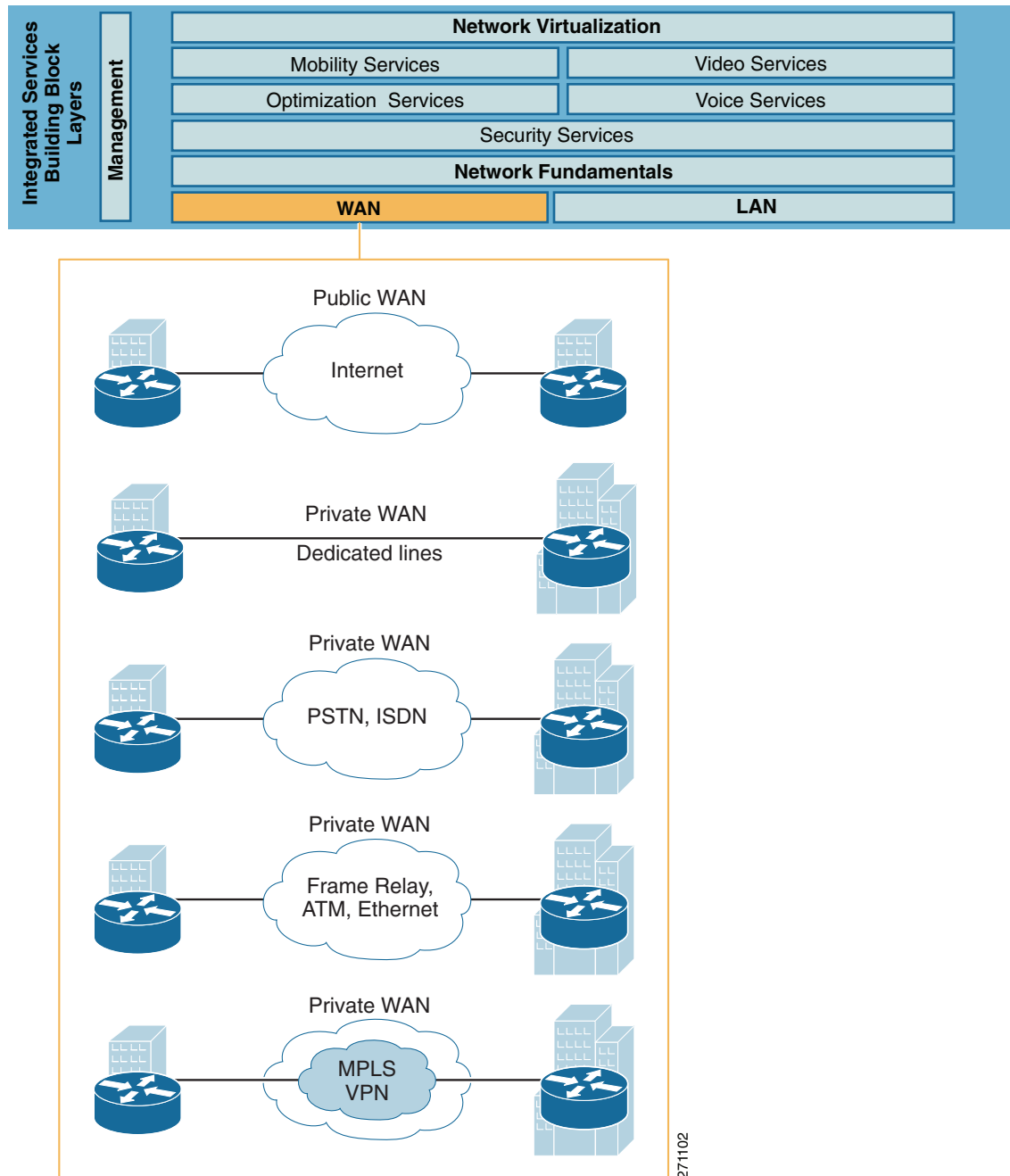
Private WAN

- **Dedicated leased lines:** Permanent point-to-point links connecting two fixed points across a provider network. In general, the links are based on Layer 1 (SONET/SDH, T1/E1, T3/E3, xDSL) technology. Today, because of the availability of cheaper alternatives, only branches that have special business requirements, that are geographically near a central site, or that are limited by availability of local connection options, favor dedicated lines.
- **Circuit-switched transmission service:** Dynamically created point-to-point links over telephone wires. The links are typically based on analog dialup or ISDN technology. Today, because of bandwidth limitations and lengthy call setup, they are mainly used for voice services or as a primary link backup.
- **Packet-switched transition service:** Virtual point-to-point or point-to-multipoint links that are established over a provider-administered Layer 2 network. The provider network is based on Frame Relay, ATM, or Ethernet technology. Although this is the most widely used connectivity option for branch offices, Frame Relay and ATM as services are declining in popularity because of MPLS based alternatives. Using Ethernet implemented over SONET or using Ethernet switches is gaining popularity in the form of carrier Ethernet services (L2VPN) such as Ethernet Private Line (EPL), Ethernet Virtual Private Line (EVPL), or Ethernet-LAN (E-LAN).
- **Label-switched transmission service:** Virtual any-to-any links running on top of a packet or circuit-switched network. The provider network is based on MPLS technology, which is emerging as the foundation of next-generation WANs that can deliver a wide range of advanced services such as Layer 3 VPN (L3VPN), or as transport mechanisms for carrier Ethernet services (L2VPN) mentioned above.

Public WAN

- Internet broadband link: Shared any-to-any links over the Internet. This has become an attractive connectivity option in recent years for smaller branch offices as VPN technologies has matured and as broadband connectivity has become more widely available. For large branch offices, this connectivity option is mainly used as a primary link backup. In general, broadband links are based on dialup, cable, and terrestrial or satellite wireless technologies.

Figure 13 WAN Service Options



Selecting WAN Service

A WAN includes transmission service available from a service provider and an access link to the service provider network. Selecting the appropriate provider network service and the access link involves many considerations. For a branch office, the most important considerations are:

- Purpose: The WAN service must provide seamless access to any site in the enterprise.
- Geographic scope: The WAN service must provide access to both regional and global sites.
- Traffic profile: For the Services Ready Large Branch Network, the WAN service must support up to 45 Mb/s of data, voice, and video traffic.
- Quality guarantee: The WAN service must provide a mechanism to ensure quality of service (QoS).
- Security: The WAN service must provide a mechanism to ensure traffic privacy.
- Existing infrastructure: The WAN service must be consistent with or must leverage existing WAN deployment.
- Availability: Selection of the WAN service must take into account local availability.
- Cost: The WAN service cost must be evaluated based on how well it meets the above considerations.

Table 2 lists advantages and disadvantages of the most commonly used WAN transmission services for a branch office.

Table 2 Common WAN Transmission Service Options for a Large Branch Office

Service Type	Advantage	Disadvantage	Appropriate for Branches
Leased Line	<ul style="list-style-type: none"> • Secure and private • Uncontended bandwidth • Reliable and predictable • Supports any protocol 	<ul style="list-style-type: none"> • Expensive • Point-to-point • Fixed bandwidth 	<ul style="list-style-type: none"> • Geographically close to campus or data center • With critical applications that require guaranteed bandwidth
Frame Relay (FR) Service	<ul style="list-style-type: none"> • Cost effective • Adjustable bandwidth • Extensive coverage • Secure and private • Reliable and resilient • Flexible and scalable • IP and non-IP protocols 	<ul style="list-style-type: none"> • Variable bandwidth, latency, and jitter • Point-to-point • Inefficient QoS 	<ul style="list-style-type: none"> • With legacy FR WAN deployment • With hub-and-spoke WAN topology • With non-IP applications

Table 2 Common WAN Transmission Service Options for a Large Branch Office (continued)

Service Type	Advantage	Disadvantage	Appropriate for Branches
Layer 3 Virtual Private Network Service (MPLS L3VPN)	<ul style="list-style-type: none"> • Same benefits as Frame Relay except for support of non-IP protocols • Any-to-any connectivity • QoS provisioning • Traffic engineering • Support wide variety of IP applications 	<ul style="list-style-type: none"> • Potentially costly migration • Proprietary to service provider network • Limited global availability • Supports only IP 	<ul style="list-style-type: none"> • Most medium and large branch offices
Layer 2 Virtual Private Wire Service (VPWS)	<ul style="list-style-type: none"> • Same benefits as Frame Relay • Transparent LAN integration • Low administrative overhead 	<ul style="list-style-type: none"> • Potentially costly migration • Limited availability • Limited scalability • Point-to-point 	<ul style="list-style-type: none"> • With enterprise control over WAN routing • With non-IP applications

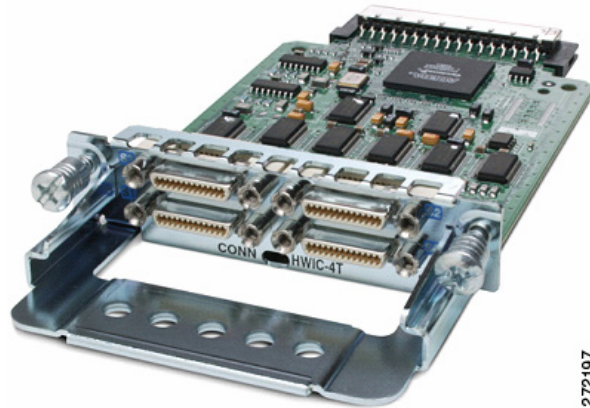
In addition to these general considerations, a WAN service must meet the business criteria outlined in the “[Large Branch Design Considerations](#)” section on [page 4](#). To ensure its relevance and applicability, the Services Ready Large Branch Network was validated with all the WAN service options listed in [Table 2](#). Specific design considerations related to each WAN service type are described in the following sections:

- [Leased-line Deployment, page 15](#)
- [Frame Relay Service Deployment, page 17](#)
- [L3VPN Service Deployment, page 18](#)
- [VPWS Services, page 20](#)

To access the WAN service, a branch office needs a local loop to the nearest location where the provider makes the service available. Typically, this is a dedicated leased line to the edge of the provider’s network. To support 100 to 240 active users, the following connection types and bandwidth options are appropriate:

- Multiple T1 carrier lines connected to an HWIC-4T interface, shown in [Figure 14](#)

Figure 14 4-Port Serial High-Speed WAN Interface Card (HWIC-4T) with 4 T1 High-Speed Serial Ports



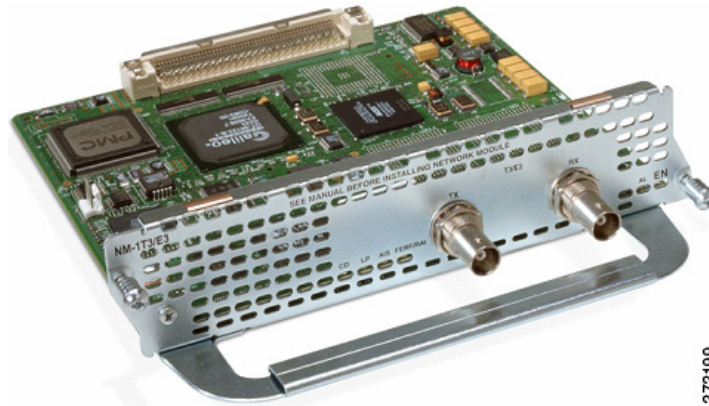
272197

To learn more about the HWIC-4T interface card, visit:

http://www.cisco.com/en/US/prod/collateral/modules/ps5949/ps6182/product_data_sheet0900aecd80274416.html

- Gigabit Ethernet line connected to an onboard small form pluggable (SFP) port.
- T3/E3 carrier line connected to a NM-1T3/E3 interface shown in [Figure 15](#).

Figure 15 1-Port Network Module (NM-1T3/E3) with One T3 High-Speed Serial Port and One E3 High-Speed Serial Port



272199

To learn more about the NM-1T3/E3 interface card, visit:

http://www.cisco.com/en/US/prod/collateral/modules/ps2797/ps4909/product_data_sheet09186a008010fba2.html

The specific selection of a WAN access link depends on the number of end user devices, the branch traffic profile, the applications used in the branch, and the available budget. The Services Ready Large Branch Network was validated with the two interface cards and the onboard SFP port described previously.

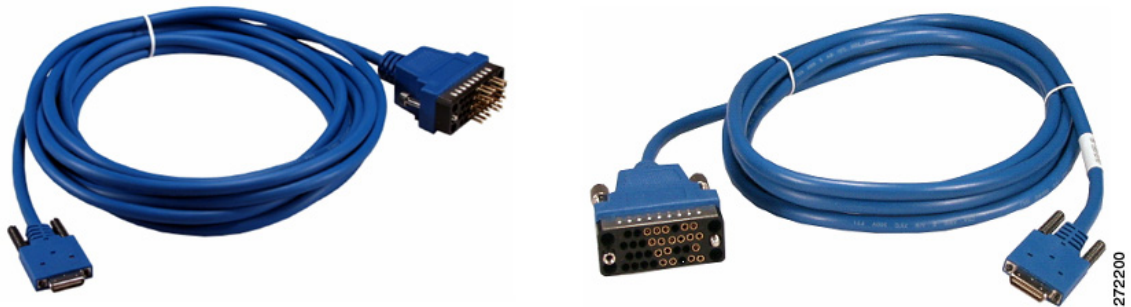
Physical layer standards define the mechanical connection and electrical signaling to connect the branch router to the service provider network, which are typically done through a channel service unit (CSU)/data service unit (DSU) device that provides termination for digital signals, clocking, and

synchronization, and that converts T-carrier line frames into frames that the LAN can interpret and vice versa. The branch router typically uses serial communication to connect to the CSU/DSU. The specific serial standard and socket type depend on the CSU/DSU equipment supplied by the service provider.

The Services Ready Large Branch Network was validated with the following serial communication specifications:

- V.35 shown in [Figure 16](#). This serial specification is typically used to connect a Cisco router to a T1/E1 and fractional T1/E1 through a CSU/DSU. A single V.35 connector can achieve up to 2.048 Mb/s speed.

Figure 16 Male (CAB-SS-V35MC) and Female (CAB-SS-V35FC) V.35 Connectors



To learn more about Cisco High-Speed Serial Interface options, visit:

http://www.cisco.com/en/US/prod/collateral/modules/ps5949/ps6182/product_data_sheet0900aecd80274416.html

The Gigabit Ethernet interface has an SFP slot on which to interchange different copper or optical SFP modules. Selection of a specific SFP depends on the distance to the nearest service provider point of presence and the type of 1000BASE-X Ethernet available. The Services Ready Large Branch Network was validated with the following SFP module:

- GLC-LH-SM, shown in [Figure 17](#). This is a connector for long-wavelength/long-haul (1000BASE-LX/LH) single-mode fiber and contains a Class 1 1300-nm laser that can reach up to 6.2 miles.

Figure 17 GLC-LH-SM Small Form-Factor Pluggable Module



To learn more about Cisco SFP modules, visit:

http://www.cisco.com/en/US/docs/routers/7200/install_and_upgrade/gbic_sfp_modules_install/5067g.html

Each T3/E3 port on the NM-T3/E3 module consists of a pair of 75-ohm BNC coaxial connectors (Type RG-59), one for transmit data and one for receive data. The module provides an integrated DSU that allows T3/E3 lines to be directly terminated on a Cisco router, eliminating the need for external DSU equipment.

[Table 3](#) summarizes the WAN access line types, bandwidth, physical connection for the link, and ISR interface or module that provides access to the provider network.

Table 3 **WAN Access Link Summary**

WAN Access Line Type	Bandwidth	Physical Connection	Cisco ISR Interface or Module
T3/E3 line	43/34 Mb/s	BNC terminated coaxial cable	NM-1T3/E3
4 T1/E1 lines	6 Mb/s	V.35 cable	HWIC-4T
Gigabit Ethernet line	Shaped to 12 Mb/s	LX/LH single mode fiber	Onboard Gigabit Ethernet SFP

Each deployment scenario was also validated with a backup link to the WAN. The details are described in the [“High Availability, Rapid Recovery, and Disaster Recovery”](#) section on page 32.

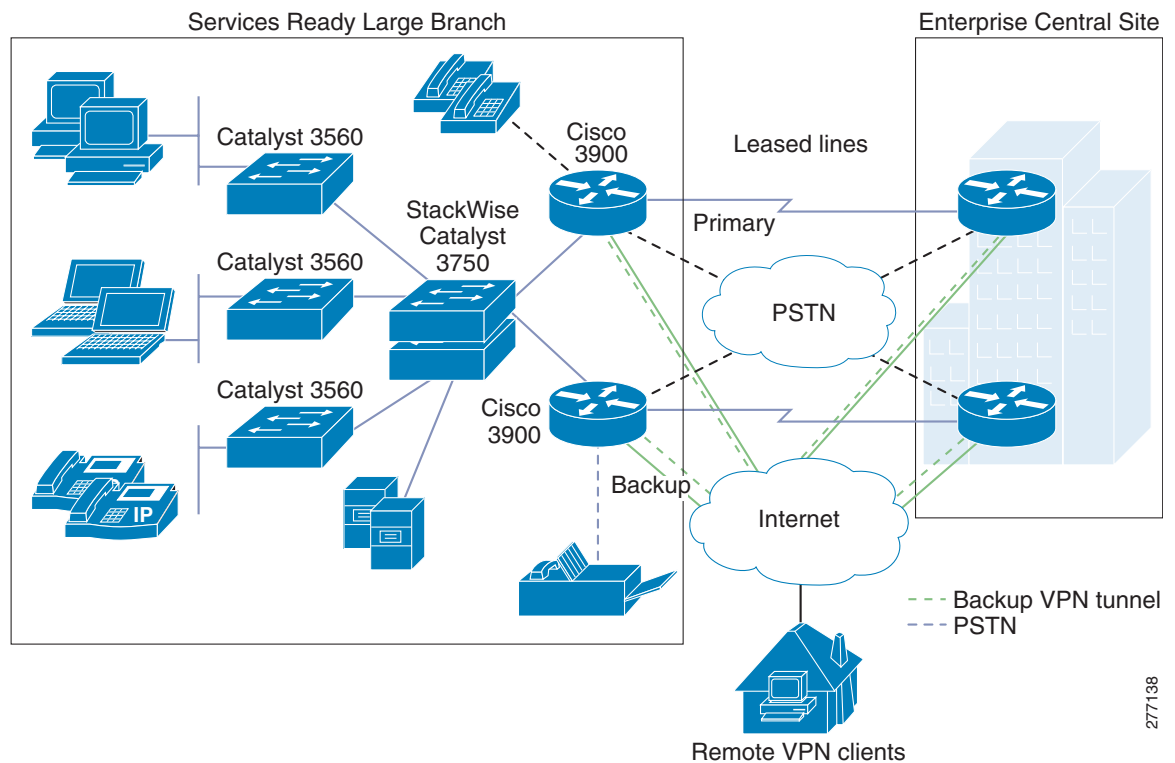
The routing and addressing aspects of each WAN deployment are described in the [IP Addressing and IP Routing](#), page 41.

Leased-line Deployment

When a branch office requires a permanent dedicated connection, a point-to-point leased line is used to provide a preestablished digital circuit from the branch through the service provider network to the central site. The service provider reserves the circuits for exclusive use by the enterprise. For a branch office, leased lines are typically available in fractional, full, or multiple T1/E1 or T3/E3 capacities. They are generally priced based on bandwidth and distance between the two connected endpoints. The cost of a leased-line WAN can become significant when it is used to connect a branch to many sites over increasing distance. Therefore, leased-line WANs are typically used to connect the branch to a central site, only when it is over a geographically short distance; when branch applications have critical bandwidth, latency, and/or jitter requirements; or when no acceptable alternatives are available in the geographic area. However, leased lines are used extensively to connect branches to a local point of presence (POP) that serves as an entry point into a service provider network offering other types of WAN transmission services.

[Figure 18](#) shows the Services Ready Large Branch Network leased-line deployment scenario.

Figure 18 Services Ready Large Branch Network Leased Line Deployment



All traffic must be encapsulated by a data link layer protocol while it is crossing the WAN. The protocol defines how data is encapsulated into frames and the mechanism for transferring the frames between the branch and a central site. Selection of the data link layer protocol depends on the WAN technology and the communicating equipment in use. For leased-line WAN links, the following are the most prevalent data link protocols:

- Point-to-Point Protocol (PPP): The most popular encapsulation protocol for transporting IP traffic over point-to-point links. PPP provides asynchronous and synchronous encapsulation, network protocol multiplexing, link configuration, link quality testing, error detection, and option negotiation for capabilities such as network layer addresses or data-compression algorithms.
- Multilink Point-to-Point Protocol (MLPPP): A method for splitting, recombining, and sequencing datagrams across multiple PPP links. It combines multiple physical links into one logical link to increase available bandwidth. To learn more about PPP and MLPPP, visit: http://www.cisco.com/en/US/tech/tk713/tk507/tsd_technology_support_protocol_home.html
- Gigabit Ethernet (GigE): Various standards capable of carrying standard Ethernet frames at a rate of 1 Gb/s. GigE employs the same Carrier Sense Multiple Access with Collision Detection (CSMA/CD) protocol, same frame format, and same frame size as its lower speed predecessors. One of the optical fiber standards (1000BASE-X) is typically used for local loop connectivity.

The Services Ready Large Branch Network was validated with the following combination of leased lines and encapsulation protocols:

- T3/E3 line with PPP
- Four T1 lines with MLPP
- 1000BASE-LX/LH GigE

Frame Relay Service Deployment

The traditional alternative to permanent leased lines has been virtual circuits provisioned over a service provider-administered Frame Relay network. A branch office is connected to the network by attaching a point-to-point link from the branch router (DTE) to the provider's nearest Frame Relay switch (DCE). When connections are in place for both the branch and a central site, a virtual circuit is set up to allow communication between the two locations. The virtual circuit is typically configured to stay active all the time. A virtual circuit is identified by Data Link Connection Identifier (DLCI), which ensures bidirectional communication from one DTE device to another and which guarantees data privacy. A number of virtual circuits can be multiplexed into a single physical line for transmission across the network. Therefore, it is relatively easy to connect one branch office to multiple destinations.

Frame Relay is an any-to-any service over a network shared by many subscribers. The sharing allows service providers to offer lower monthly rates in comparison to dedicated leased lines. The data rate is also more flexible. Instead of one fixed rate, bursts are allowed if the network has available capacity. The downside to a shared network is a potential drop in service when traffic increases. To provide acceptable performance, service providers usually offer a minimum committed rate that is guaranteed to a subscriber. Frame Relay can provide speeds from 56 kb/s to 43 Mb/s, depending on the capability of the service provider's network.

While Frame Relay is considered legacy today, it is used extensively to implement enterprise WANs. Its primary advantages are cost and deployment flexibility. In comparison to leased lines, bandwidth is cheaper because it is shared, and only a short local loop is required to connect the branch to the nearest Frame Relay switch. Adding virtual circuits or increasing bandwidth is simple and fast.

The leased-line connection to the Frame Relay network typically uses one of the following Frame Relay encapsulation mechanisms:

- Frame Relay (FR) protocol: Specifies how data moves between the DTE and DCE over a single line.

To learn more about FR, visit:

http://www.cisco.com/en/US/tech/tk713/tk237/technologies_tech_note09186a008014f8a7.shtml

- Multilink Frame Relay (MLFR): Enables multiple lines to be aggregated into a single bundle of bandwidth.

To learn more about MLFR, visit:

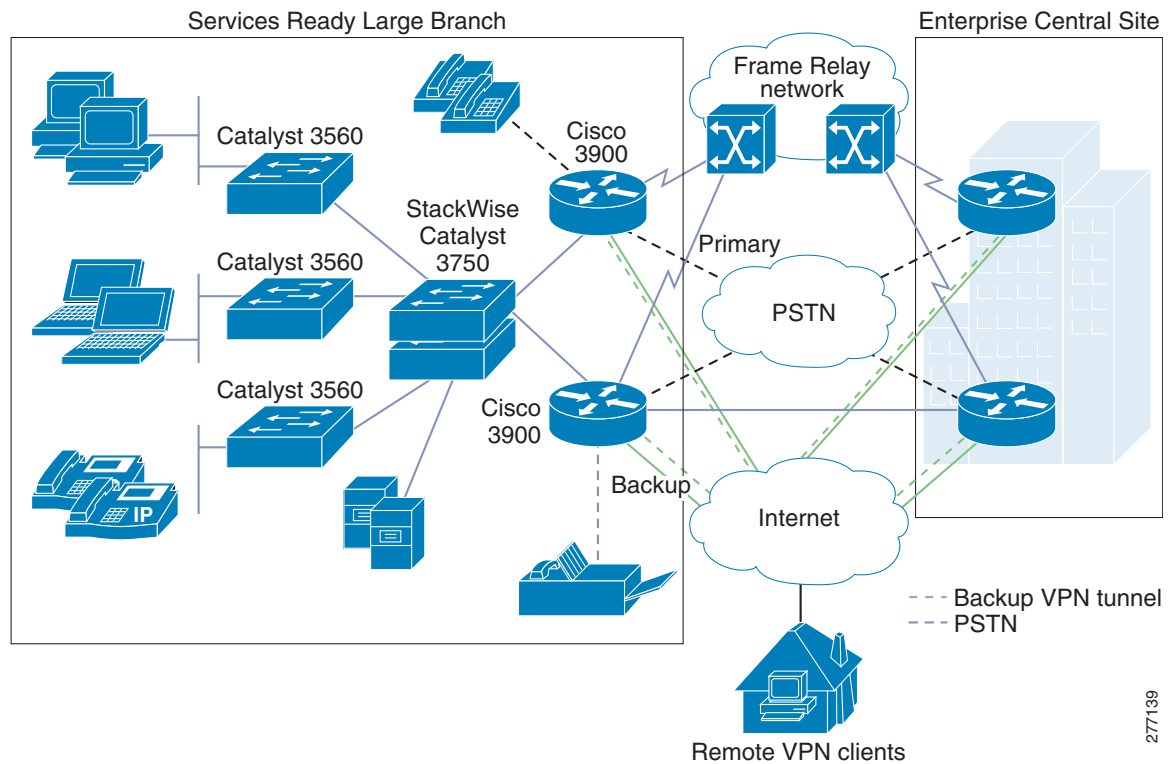
http://www.cisco.com/en/US/docs/ios/12_0s/feature/guide/17s_mfr.html

The Services Ready Large Branch Network was validated with the following combination of Frame Relay encapsulation protocols:

- T3/E3 line with FR protocol
- Four T1 lines with MLFR protocol

Figure 19 shows the Frame Relay private WAN deployment scenario.

Figure 19 Services Ready Large Branch Network Frame Relay Service Deployment



277139

L3VPN Service Deployment

Many enterprises are turning to MPLS-based WAN services because they offer cost-effective, scalable, and flexible alternatives to the traditional Frame Relay (or ATM) based private WANs. MPLS is a label-based protocol that operates between the data link layer (Layer 2) and the network layer (Layer 3). A label is imposed on a packet at the edge of the MPLS network and is removed at the other end. Label forwarding is performed by a lookup on the incoming label, which is then swapped for the outgoing label and forwarded to the next hop. Routing decisions and reachability information are based on IP addresses. Therefore, Layer 3 is also the foundation for any services offered by MPLS-based networks. Virtual Private Network (VPN) technology combined with MPLS provides traffic security and privacy. There are two general types of VPNs: enterprise-managed and service provider-managed. Layer 3 MPLS VPN (L3VPN) is a service provider-managed VPN service.

In an L3VPN WAN deployment, the provider's MPLS network routes the enterprise IP traffic. A provider edge (PE) router directly connects to the customer edge (CE) router in the branch office. The PE router communicates with the CE router via the routing protocol selected by the enterprise (RIP, OSPF, BGP, and so on). Thus, the PE router learns all of the enterprise routes and forwards the packets based on that information. The PE router also exchanges reachability information with other PE routers in the MPLS network by running Multiprotocol Border Interior Gateway Protocol (M-iBGP) in the MPLS network core.

L3VPN services offer several unique advantages over traditional private WANs:

- They offer scalable any-to-any connectivity. A CE router peers with a PE router that maintains the full mesh topology. Unlike Frame Relay (or ATM), there is no complex virtual circuit topology to manage. Adding a new site to the mesh involves no other connections beyond the one connection to the PE router.
- Two branches can have overlapping address space if they are members of different VPNs.
- MPLS is IP aware and has a single control plane that matches the physical topology of the network. This allows better mapping of traffic into available resources or rapid redistribution of traffic in response to changes in the topology.
- Service providers are leveraging IP QoS to offer a full range of service guarantees for critical traffic.

The main limitation of MPLS stems from its dependence on IP. Only IP-based traffic is supported, and all other protocols must use a tunneling mechanism.

To learn more about Layer 3 MPLS VPN, visit:

http://www.cisco.com/en/US/solutions/ns340/ns414/ns465/net_design_guidance0900aec80375d78.pdf

http://www.cisco.com/en/US/docs/net_mgmt/vpn_solutions_center/1.1/user/guide/VPN_UG1.html

The leased-line connection to the PE device typically uses one of the following data link layer encapsulation mechanisms:

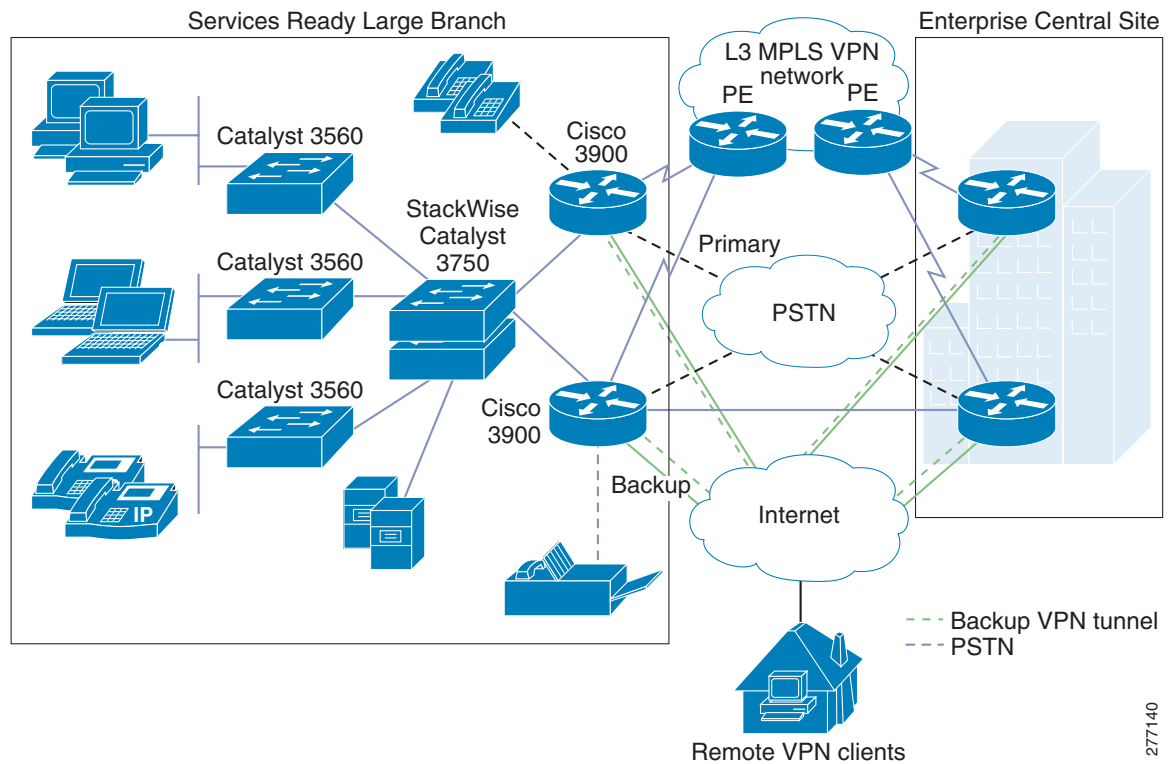
- PPP: Described in the “Leased-line Deployment” section on page 15.
- MLPPP: Described in the “Leased-line Deployment” section on page 15.

The Services Ready Large Branch Network was validated with the following combination of access links to a PE device:

- T3/E3 line with PPP
- Four T1 lines with MLPPP

Figure 20 shows the L3VPN private WAN deployment scenario.

Figure 20 Services Ready Large Branch Network L3VPN Deployment



277140

VPWS Services

For enterprises that want to retain control over Layer 2 connectivity, service providers offer Layer 2 VPNs. The following sections describe the most typically offered services.

MPLS Switched WAN Services

- Layer 3 VPNs: Described in the “[L3VPN Service Deployment](#)” section on page 18.
- Layer 2 VPNs: Emulation of Layer 2 connectivity over MPLS network
 - Virtual Private LAN Service (VPLS): The branch office Ethernet LAN is extended to the provider edge (PE) device. The provider network then emulates the function of a LAN switch to connect all customer LANs into a single bridged LAN. VPLS is a point-to-multipoint service.
 - Virtual Private Wire Service (VPWS, also called *PWE3 pseudowire*): The service provider network emulates point-to-point connections from the branch over the underlying MPLS tunnel. In general, the network emulates existing Frame Relay, ATM, Ethernet, HDLC, or PPP links. The enterprise keeps the same Layer 2 connections to the service provider, but instead of the data being carried natively over a Frame Relay or ATM service, the data is encapsulated and routed over the provider’s MPLS backbone.

Ethernet Switched WAN Services

- Permanent Point-to-Point Ethernet Line: Dedicated Ethernet circuit. The permanent point-to-point Ethernet switched WAN series are described in the “Leased-line Deployment” section on page 15.
- Virtual Ethernet Connections: Connectivity over a service provider’s shared Ethernet network.
 - E-Line: Point-to-point Ethernet services (single link configuration)
 - Ethernet Private Line (EPL): Dedicated point-to-point virtual line. The connection from the branch goes to a dedicated User Network Interface (UNI) device. Multiple EPLs require multiple UNIs. EPL is an alternative to dedicated leased lines.
 - Ethernet Virtual Private Line (EVPL): Multipoint-to-point virtual lines. A single UNI multiplexes multiple virtual connections. EVPL is an alternative to Frame Relay or ATM PVCs.
 - E-Tree: Point-to-multipoint Ethernet services (hub-and-spoke configuration)
 - Ethernet Private Tree (EP-Tree): Single point-to-multipoint virtual lines.
 - Ethernet Virtual Private Tree (EVP-Tree): Multipoint-to-multipoint virtual lines.
 - E-LAN: Multipoint-to-multipoint Ethernet service (full-mesh configuration)
 - Ethernet Private LAN (EP-LAN): Single multipoint-to-multipoint virtual lines.
 - Ethernet Virtual Private LAN (EVP-LAN): Multiple multipoint-to-multipoint virtual lines.

Selecting the most appropriate Ethernet-switched WAN service from this list involves several considerations. One of the first decision points is between L3VPN or L2VPN service. Table 4 provides a high-level comparison of the two options. Ultimately, the decision depends on the amount of control that the enterprise wants to retain over its WAN deployment.

Table 4 High-Level Comparison Between L2VPNs and L3VPNs

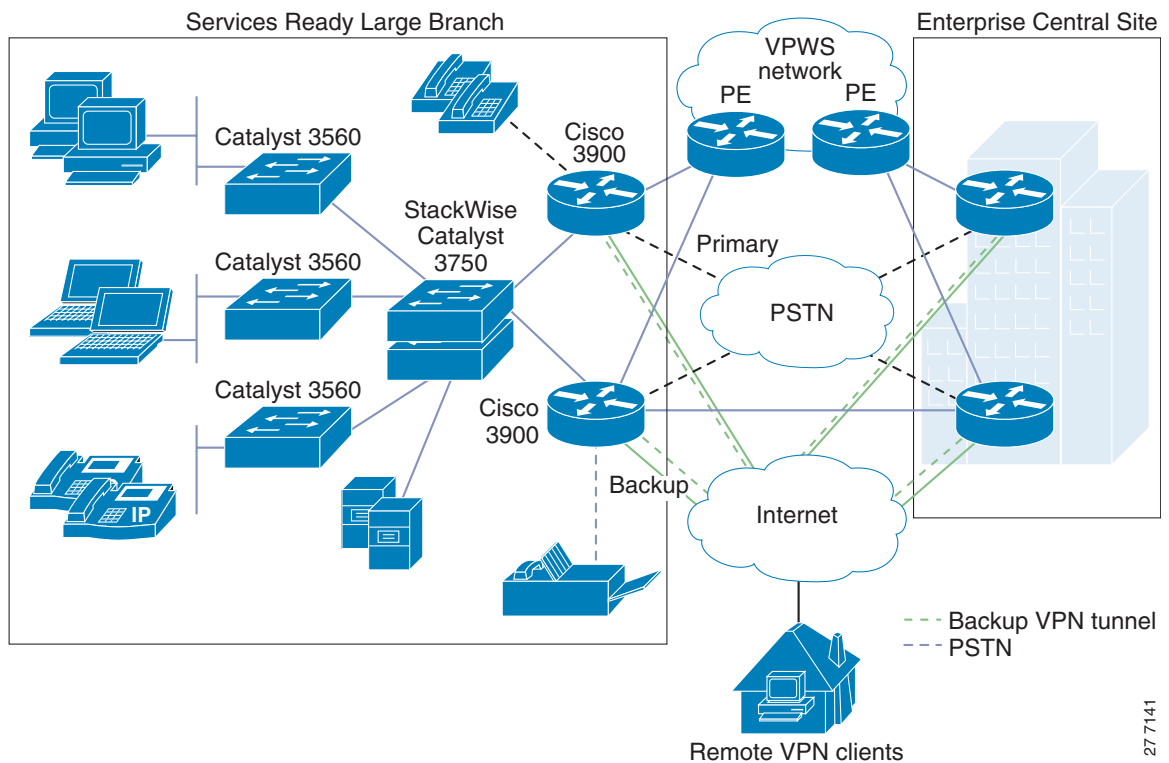
L2VPN	L3VPN
Provider forwards frames, based on Layer 2 information	Provider forwards packets, based on Layer 3 information
Provider involved in routing	Provider not involved in routing
Supports only Ethernet as access technology	Supports any access technology
Enterprise controls Layer 3 policies (routing, QoS)	Provider controls Layer 3 policies (routing, QoS)
Supports any Layer 3 protocol	Supports only IP
Limited scalability	Scalable

The Services Ready Large Branch Network was validated with Virtual Private Wire Services (VPWS). In this deployment, the service provider network acts as a Layer 2 switch. It maps incoming traffic to pseudowires based on Layer 2 headers. Figure 21 shows a VPWS deployment scenario.

To learn more about Layer 2 MPLS VPNs, visit:

http://www.cisco.com/en/US/technologies/tk436/tk891/technologies_white_paper0900aecd80162178_ns585_Networking_Solutions_White_Paper.html

Figure 21 The Services Ready Large Branch Network VPWS Deployment



VPWS services allow the enterprise to keep its existing WAN infrastructure and to transparently connect to the service provider's Ethernet network, providing a transparent migration path to VPLS services. The leased-line connections to the PE device continue to use the typical Layer 2 encapsulation mechanism:

- PPP: Described in the [“Leased-line Deployment”](#) section on page 15.
- MLPPP: Described in the [“Leased-line Deployment”](#) section on page 15.
- Gigabit Ethernet (GigE): Described in the [“Leased-line Deployment”](#) section on page 15.

The Services Ready Large Branch Network was validated with the following combination of access links to a PE device:

- T3/E3 line with PPP
- Four T1 lines with MLPPP
- 1000BASE-LX/LH GigE

LAN Deployment Model

LAN services provide connectivity for converged data, voice, and video communication. Consequently, a properly designed LAN is a fundamental requirement for performing day-to-day business functions at the branch office. Of the various ways to architect a LAN, a hierarchical design is best suited to meet the business criteria outlined in the [“Large Branch Design Considerations”](#) section on page 4.

A typical hierarchical design is broken into three logical layers:

- **Access layer:** Interfaces with end devices, such as PCs, IP Phones, printers, and servers. The access layer provides access to the rest of the network, and it controls which devices are allowed to communicate on the network.
- **Distribution layer:** Aggregates the data that is received from the access layer switches, provides for data separation and forwards traffic to the core layer for routing to its final destination. It controls the flow of traffic, delineates broadcast domains, and provides resiliency.
- **Edge layer:** Aggregates the data that is received from the distribution layer switches and serves as an entry and exit point between the LAN and WAN. This is typically the branch router.

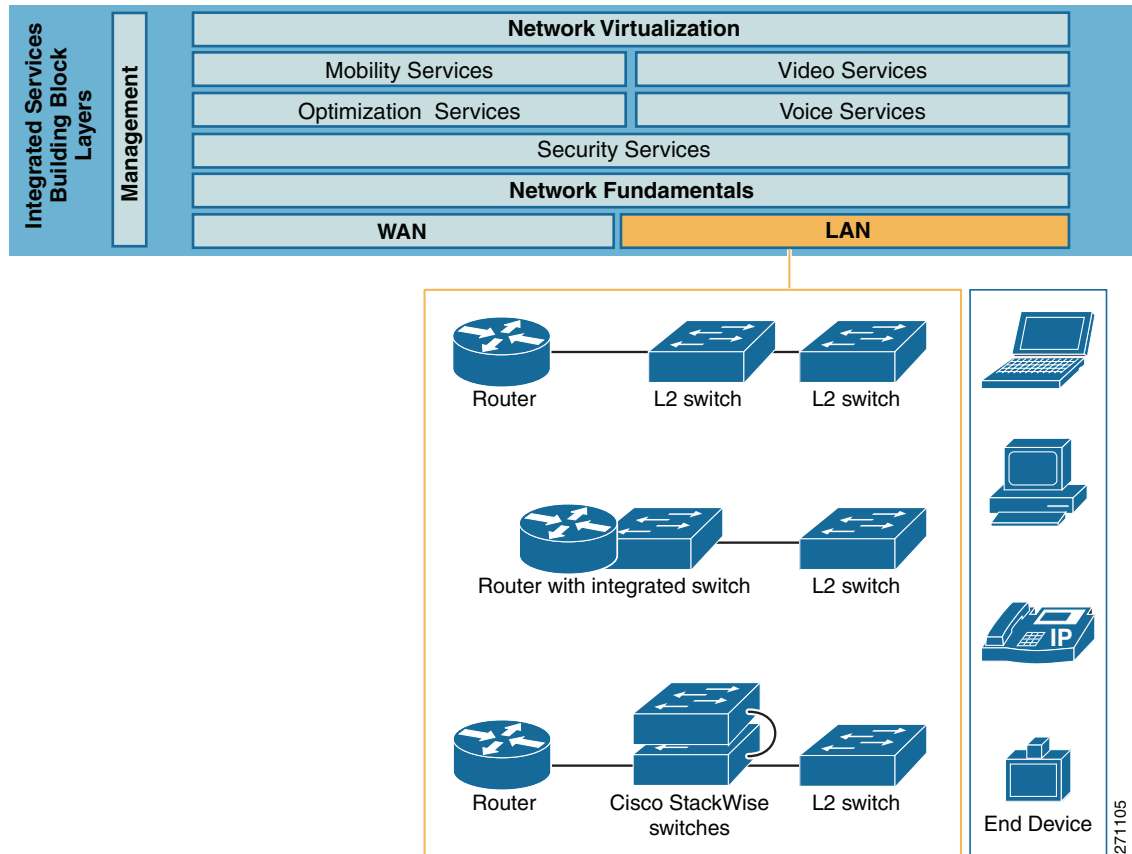
This design has the following benefits:

- **Scalability:** The modularity of the design provides room for easily adding devices as the network grows.
- **Resiliency:** Connecting the access layer switches to multiple distribution switches ensures path redundancy.
- **Performance:** Hierarchical layering enables fewer higher performing switches to aggregate traffic from many lower performing switches. The need for fewer higher performing switches results in both cost savings and optimal use of network devices.
- **Security:** Different security policies can be implemented at various levels of the hierarchy
- **Manageability:** All switches in one layer perform the same function, making it easy to propagate changes.

Hierarchical LAN design is only a logical layout of network devices. A large branch office has three prominent physical implementation options, shown in [Figure 22](#), that map into the logical hierarchical design:

- Access router that is connected to physically separate distribution and access switches
- Access router with integrated distribution switches and physically separate access switches
- Access router that is connected to physically separate distribution switches stacked in a Cisco StackWise topology and connected to physically separate access switches

Figure 22 LAN Connectivity Options for Large Branch Office



Although it is feasible to use both integrated and unstacked switch configurations for a branch office of 100 to 240 users, these deployments do not meet the requirements highlighted in the “[Large Branch Design Considerations](#)” section on page 4. Specifically, they lack the desired resiliency and scalability. Therefore, only the Cisco StackWise switch configurations were considered for the Services Ready Large Branch Network.



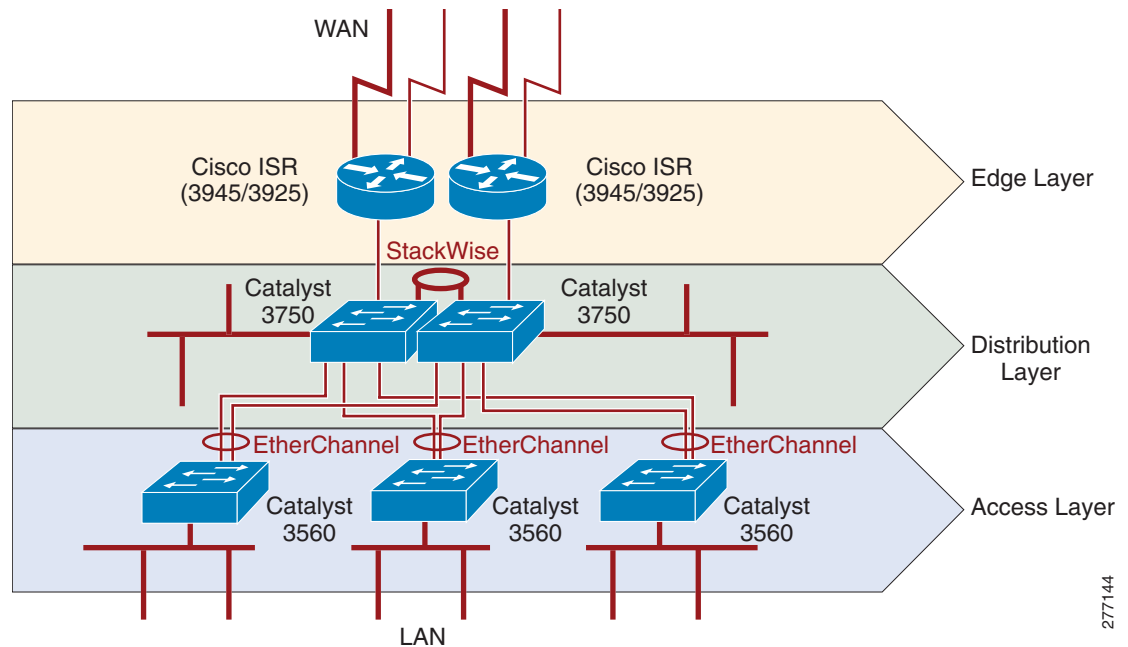
Note

Stacked switches must all be running the same Cisco IOS software release.

For a more in-depth discussion of various branch LAN deployment options and features, see the following:

- [LAN Baseline Architecture Branch Office Network Reference Design Guide](#)
- [LAN Baseline Architecture Overview--Branch Office Network](#)

The “[Selecting Network Components](#)” section on page 3 briefly describes the Catalyst 3750 and Catalyst 3560 switches that were selected for the Services Ready Large Branch Network LAN. The Catalyst 3560 switch was chosen for the access layer, and the Catalyst 3750 switch for the distribution layer. [Figure 23](#) shows a high-level physical topology diagram for the LAN. The actual number of switches in the distribution and access layer, and the cabling arrangements depend on the number of end devices. The Services Ready Large Branch Network used 1.25 end devices per user, assuming that most PCs are connected to the switch through an IP Phone. [Figure 23](#) shows one possible physical configuration for a 240-user branch office.

Figure 23 Hierarchical LAN Design

Switches must support many features to facilitate interoffice connectivity. Features of the Catalyst 3560 and 3750 switches that were leveraged by the Services Ready Large Branch Network are described in the following sections:

- [Virtual LANs, page 26](#)
- [VLAN Trunks and VLAN Trunking Protocol, page 28](#)
- [Power-over-Ethernet, page 30](#)
- [Spanning Tree Protocol, page 30](#)
- [Cisco StackWise Interconnects, page 30](#)
- [EtherChannel Link Aggregation, page 31](#)

In addition, the following features of the Catalyst switches are described in other parts of this guide:

- Layer 2 security in the [“Threat Protection, Detection, and Mitigation” section on page 65](#)
- Layer 2 Quality of Service (QoS) in the [“Quality of Service” section on page 46](#)
- Authentication services in the [“Access Control” section on page 56](#)

Each layer of the hierarchical design serves specific functions or provides important services. Access layer switches facilitate the connection of end node devices to the network. Most of these devices are equipped with a single network interface card (NIC) and therefore form only one connection to the network. If a device has multiple NICs, it can be wired to two or more access layer switches for increased resiliency. For the Services Ready Large Branch Network, the access layer provides the following functions:

- Voice, data, black hole, and management VLANs: Provide traffic separation and broadcast domains for voice, data, and management traffic.
- Uplink connections with VLAN Trunking Protocol (VTP) trunks to the distribution layer switches: Extend VLANs to distribution switches and across the entire network.
- VTP client: Accepts VLAN configuration propagated by the distribution switches.

- Layer 2 security: Controls the number and identity of devices that can connect to the network.
- QoS: Guarantees network resources for voice traffic and enforces proper usage of QoS by end devices.
- Authentication services: Authenticates the connecting device with RADIUS server.
- Power over Ethernet: Provides power to the connected IP Phones.
- Spanning Tree Protocol (STP): Eliminates any accidentally introduced loops from the network.
- High availability and link aggregation via EtherChannel to distribution layer switches: Provide bandwidth bundling and alternate paths to distribution layer switches in case of failure.

Distribution layer switches control traffic flow from access switches and are aggregation points for LAN management. When certain types of client/server applications are deployed in the branch, such as print services, building access control, or video surveillance, many of these applications run on servers that can be equipped with multiple NICs and connected to the resilient distribution layer switches. The switches provide necessary bandwidth and high availability for these essential applications. Moreover, branch offices may provide remote workers with access to web, application, or database servers located in the branch. It is a security “best practice” to isolate these servers into a demilitarized zone (DMZ) VLAN. The distribution layer is the most appropriate place in the network to configure the DMZ VLAN. Finally, if the branch office requires wireless access points, the distribution layer can access ports for these devices. For the Services Ready Large Branch Network, the distribution layer provides following functions:

- Voice, data, black hole, management, and DMZ VLANs: Switches VLAN traffic between access switches, and provides DMZ VLAN for servers accessible by home office workers.
- Uplink connections with VTP trunks to the two routers: Extends VLANs to edge routers.
- VTP server: Propagates VLAN information across the LAN.
- STP: Eliminates any accidentally introduced loops from the network.
- Cisco StackWise configuration between distribution layer switches: Enables distribution switches to act as a single logic switch.


Note

Only limited interVLAN routing is required between virtual LANs of the Services Ready Large Branch Network. Therefore, no Layer 3 functionality is enabled on the distribution layer switches. All interVLAN routing is performed by the branch router in the edge layer.

The edge layer provides:

- Connectivity, security, and management services described throughout this guide
- High availability via HSRP or GLBP described in the [“High Availability, Rapid Recovery, and Disaster Recovery”](#) section on page 32

Virtual LANs

A VLAN defines a group of logically connected devices that act as an independent LAN while sharing the same physical infrastructure with other VLANs. Each VLAN is a logically separate IP subnet. A switch can carry multiple VLANs, and a VLAN can be extended across multiple Layer 2 and Layer 3 devices. VLANs offer several benefits:

- Security: Traffic in a VLAN is separated from all other traffic by Layer 2 tags.

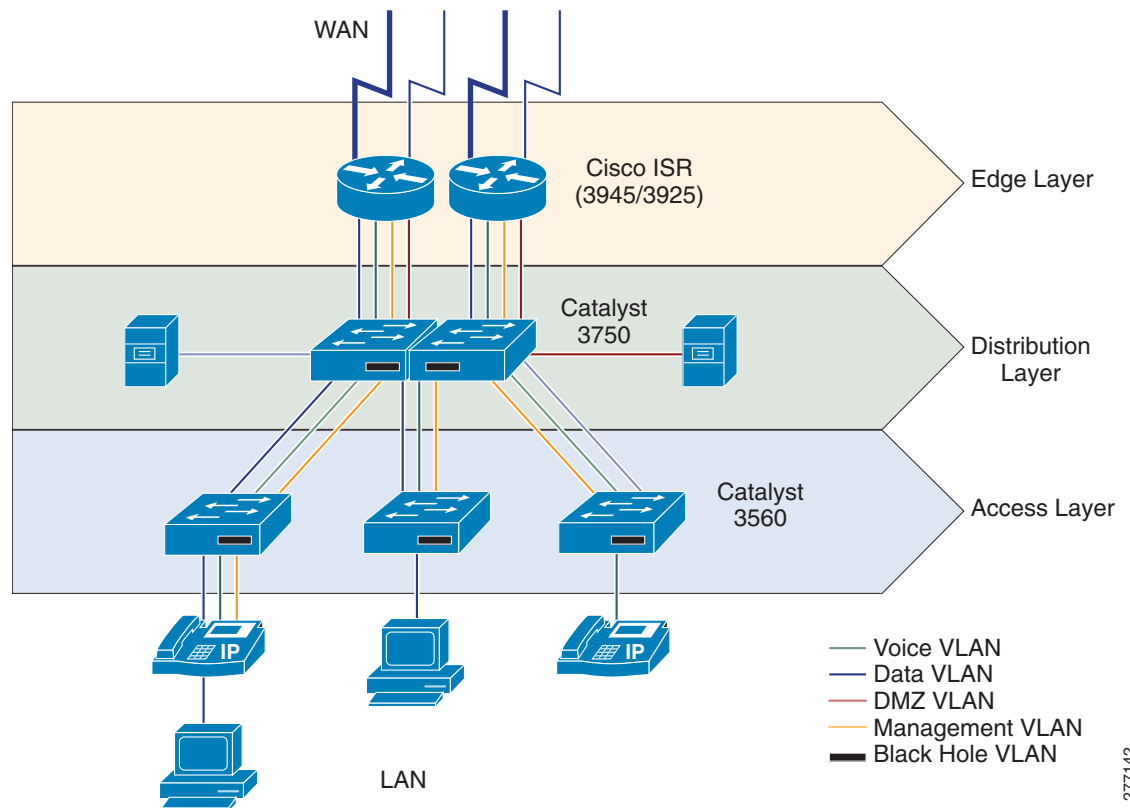
- Performance: VLANs reduce unnecessary traffic and use bandwidth more efficiently by delimiting broadcast domains.
- Management: VLANs are managed globally, and configuration is propagated across the network.

Several VLANs were defined for the Services Ready Large Branch Network:

- Data VLAN: Carries traffic generated by laptops, PCs, and servers.
- Voice VLAN: Carries traffic generated by IP Phones, and singles out voice traffic for QoS.
- DMZ VLAN: Special VLAN for web, application, and database servers accessible by home office users.
- Management VLAN: Carries traffic for managing networking devices.
- Black Hole VLAN: All unused ports are assigned to this VLAN. This is a security best practice.

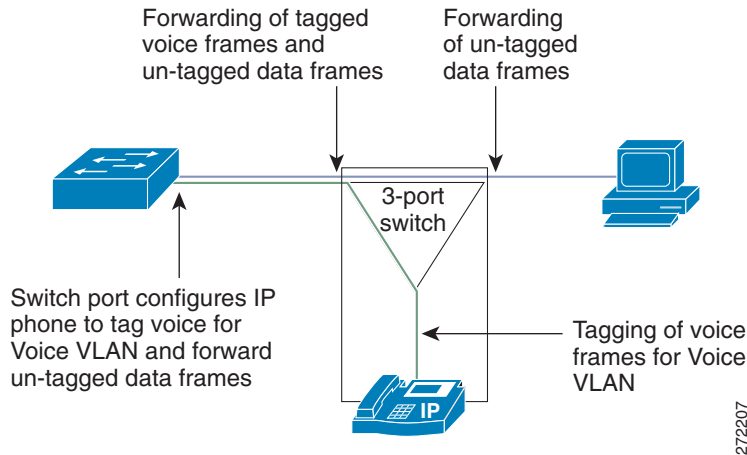
Figure 24 shows the VLAN configuration for the Services Ready Large Branch Network.

Figure 24 VLAN Design



Cisco IP Phones contain integrated three-port switches, as shown in Figure 25. An access layer switch instructs the phone to tag voice traffic for voice VLAN and to forward data frames for tagging at the switch port. This allows the switch port to carry both voice and data traffic and to maintain the VLAN separation. The link between the switch port and the IP Phone acts as a trunk for carrying both voice and data traffic.

Figure 25 *Integrated Switch in Cisco Unified IP Phone 7900 Series*



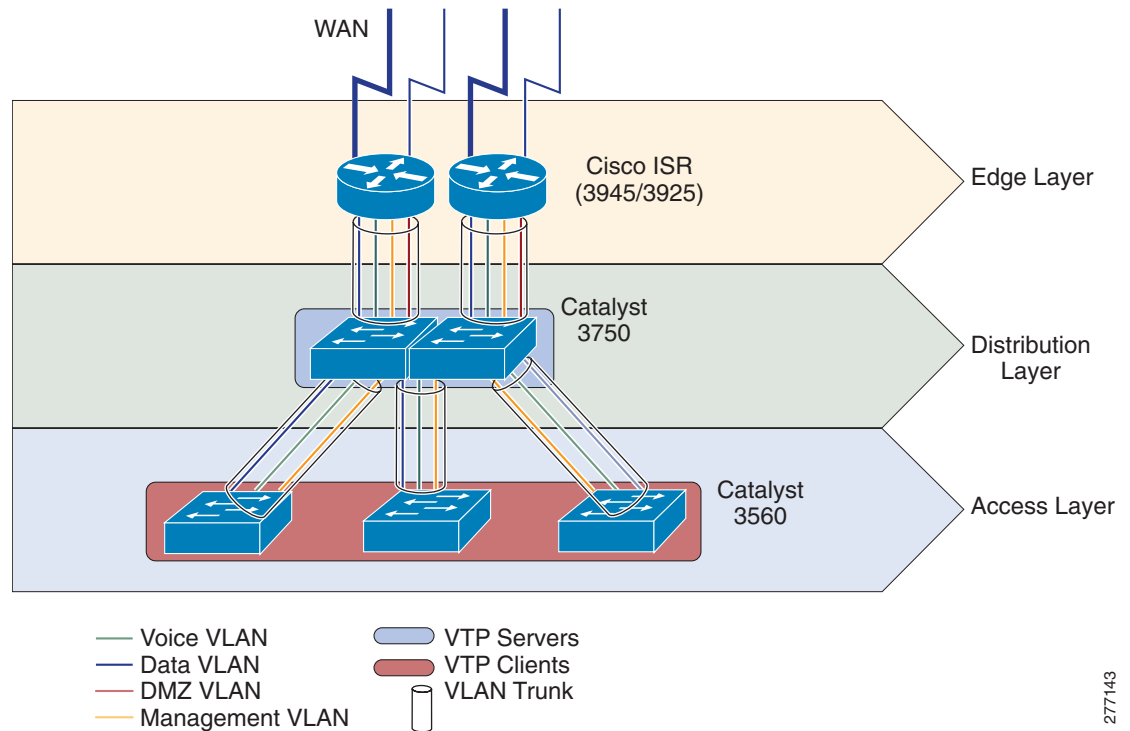
The DMZ VLAN and the black hole VLAN are described in the [“Security Services”](#) section on page 53. The Management VLAN is described in the [“Management Services”](#) section on page 68. In addition to the VLANs that were defined for the Services Ready Large Branch Network, other VLANs could be required. If the branch office has wireless access points, they should be connected to the distribution layer switches and the traffic generated through these devices should be assigned to the wireless VLAN. Moreover, some networks could continue to use older equipment that does not support 802.1Q frame tagging. Isolate these devices in their own native VLAN that supports both untagged and tagged traffic.

VLAN Trunks and VLAN Trunking Protocol

VLAN trunks are point-to-point links between two Ethernet interfaces that carry traffic for multiple VLANs. They are used to extend VLANs across the entire network. VLAN Trunking Protocol (VTP) propagates VLAN information from one switch (server) to other switches in the network (clients). VTP maintains VLAN configuration consistency by managing the addition, deletion, and changes to VLANs across multiple switches.

Figure 26 shows VLAN trunks that are defined for the Services Ready Large Branch LAN.

Figure 26 VLAN Trunks and VTP Configuration



A switch can be configured as a VTP server, as a VTP client, or in transparent mode. A VTP server distributes and synchronizes VLAN information to VTP-enabled switches. VTP clients act on that information. VTP transparent switches are unaffected, but they pass VTP advertisements to other switches. The VTP domain delimits the portion of the LAN managed by a single VTP server.

The Services Ready Large Branch Network consists of a single VTP domain. Distribution layer switches were configured as VTP servers, and access layer switches were configured as VTP clients shown in Figure 26.

VTP version 2 was used in validating the Services Ready Large Branch Network.



Note

Always check the revision number of a new switch before bringing adding it to the network, regardless of whether the switch is going to operate in VTP client mode or operate in VTP server mode. To reset the revision number, do one of the following:

- Reboot the switch
or
- Temporarily change the domain name of the new switch and then change it back to its valid domain name.

In using VTP, it is possible to run into a “VTP bomb,” which can happen when a VTP server with a higher revision number of the VTP database is inserted into the network. The higher VTP database number will cause VLAN information to be deleted from all switches. Therefore, it is important to make sure that the revision number of any new switch introduced into the network is lower than that of the VTP server.

Power-over-Ethernet

Power-over-Ethernet (PoE) provides power to devices that are attached to the switches such as IP Phones or wireless access points. All access layer switches in the Services Ready Large Branch Network are provided with the PoE option. In the distribution layer, only one switch is provided with PoE. It is assumed that distribution layer switches will primarily be used to connect to other network devices or servers that do not require PoE. However, because there maybe wireless access points connected to the distribution layer one of the switches provides PoE. Although all access layer switches should provide PoE, one non-PoE Catalyst 3560 was inserted into the Services Ready Large Branch Network for validation completion.

Spanning Tree Protocol

Spanning Tree Protocols (STPs) are used to detect and prevent traffic loops or duplicate frames in a network with redundant paths. The Services Ready Large Branch Network, by design, does not have loops. However, to prevent accidental loops that frequently occur in the wiring closet or when users connect desktop switches to the network, Rapid VLAN Spanning Tree (RVST) protocol was enabled on all the switches in the network. One of the distribution layer switches served as the root bridge for the protocol.

To learn more about STP, visit:

http://www.cisco.com/univercd/cc/td/doc/product/rtrmgmt/sw_ntman/cwsi2/cwsiug2/vlan2/tpapp.htm

Cisco StackWise Interconnects

Cisco 3750 Catalyst switches can be interconnected with Cisco StackWise technology to provide high-bandwidth throughput and fault tolerance. The switches are united into a single logical unit by special interconnect cables that create a bidirectional closed-loop path. The stack behaves as a single switching unit that is managed by a master switch chosen from one of the member switches. The master switch automatically creates and updates all the switching tables. A working stack can accept new members or delete old ones without service interruption. [Figure 27](#) shows the StackWise interconnect. [Figure 28](#) shows Catalyst 3750 switches arranged in the stacked configuration.

The Services Ready Large Branch Network uses Cisco StackWise technology to interconnect distribution layer switches.

Figure 27 *Cisco StackWise Interconnect*



Figure 28 Catalyst 3750 in Cisco StackWise Configuration



To learn more about Cisco StackWise technology, visit:

http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps5023/prod_white_paper09186a00801b096a.html

EtherChannel Link Aggregation

EtherChannel allows multiple physical Ethernet links to combine into one logical channel. The logical channel provides load sharing of traffic among the links and redundancy in the event that one or more links in the channel fail.

EtherChannel configuration is used extensively in the Services Ready Large Branch Network to aggregate bandwidth and to provide fault tolerance. A 2-port EtherChannel is created for each access layer switch by connecting the switch to each of the two distribution layer switches. This cross-stack EtherChannel ensures that, when a distribution layer switch fails, there is an alternate path to the rest of the network. Moreover, cross-stack EtherChannel increases the uplink bandwidth because distribution switches interconnected with StackWise technology act as a single switch.

To learn more about Cisco EtherChannel technology, visit:

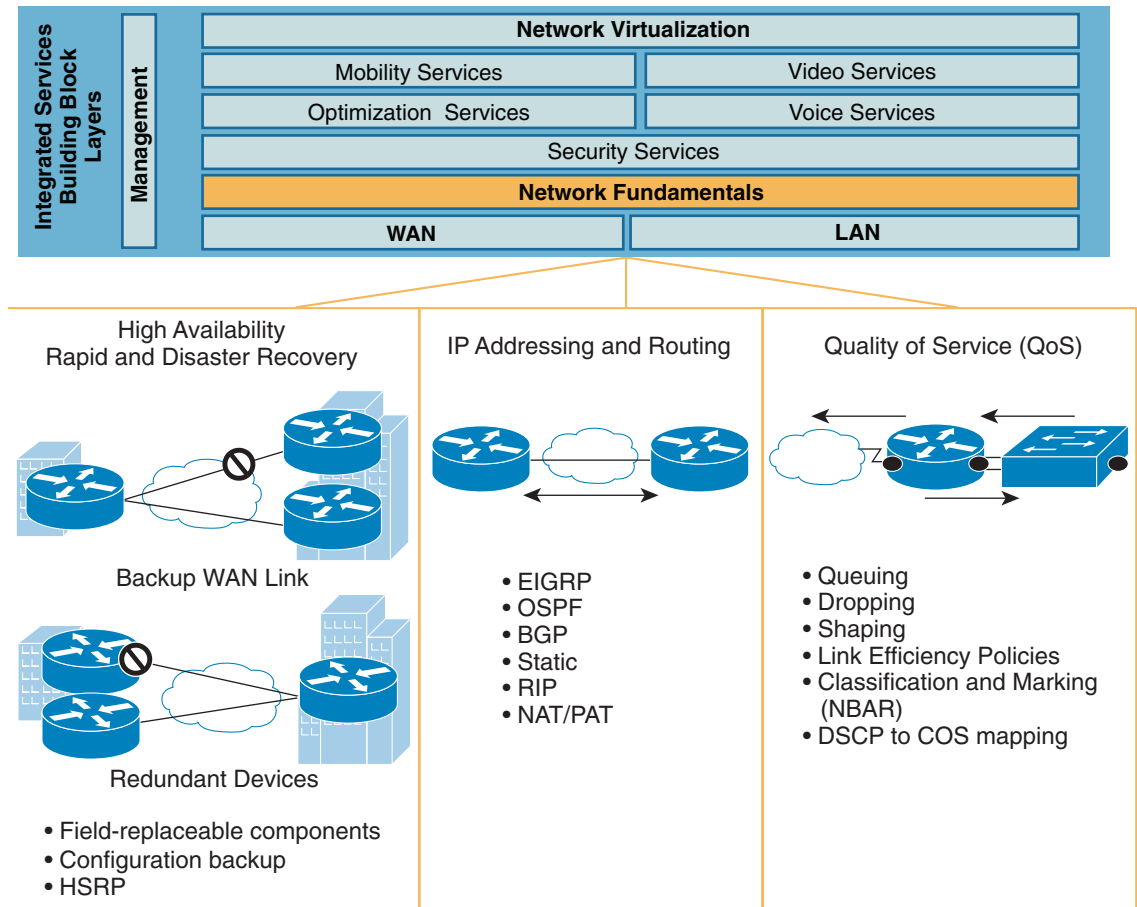
http://www.cisco.com/en/US/tech/tk389/tk213/technologies_white_paper09186a0080092944.shtml

Network Fundamentals

Network fundamentals are the basic services required for network connectivity. These services are described in the following sections and shown in [Figure 29](#):

- [High Availability, Rapid Recovery, and Disaster Recovery, page 32](#)
- [IP Addressing and IP Routing, page 41](#)
- [Quality of Service, page 46](#)

Figure 29 Basic Connectivity Services



271433

High Availability, Rapid Recovery, and Disaster Recovery

Network uptime and recovery time are critical for many types of enterprise branches. Designing the branch network for high availability ensures that network services continue to function if a single device or link failure occurs. The Services Ready Large Branch Network achieves high availability through full device and link redundancy. At every layer of the network, there is an alternate path and a backup device for failover.

Rapid recovery is the ability of a network service to quickly recover from downtime. The Services Ready Large Branch Network achieves rapid recovery by using modular, field-replaceable components with online insertion and removal (OIR) capabilities and using hierarchical network design that enables alternate devices or links to be added into the network with minimal disruption.

Disaster recovery is the process of restoring network services to full function after a failure-induced downtime. The Services Ready Large Branch Network enables disaster recovery by storing redundant copies of all device configurations on external storage devices. In addition, a Cisco SmartNet contract is recommended to provide around-the-clock, global access to the Cisco Technical Assistance Center (TAC), and 2-hour or next-business-day hardware replacement.

For the purposes of this guide, there are subtle differences between high availability, rapid recovery, and disaster recovery. High availability enables service continuity in case of a single device or link failure. The system automatically switches to an alternate device or a link. Rapid recovery enables restoration

of full service, minimal downtime in the event of a nonredundant device failure, or recovery from a multidevice failure. Disaster recovery enables restoration of a service after a failure of the network as a whole (for example, in case of a fire).

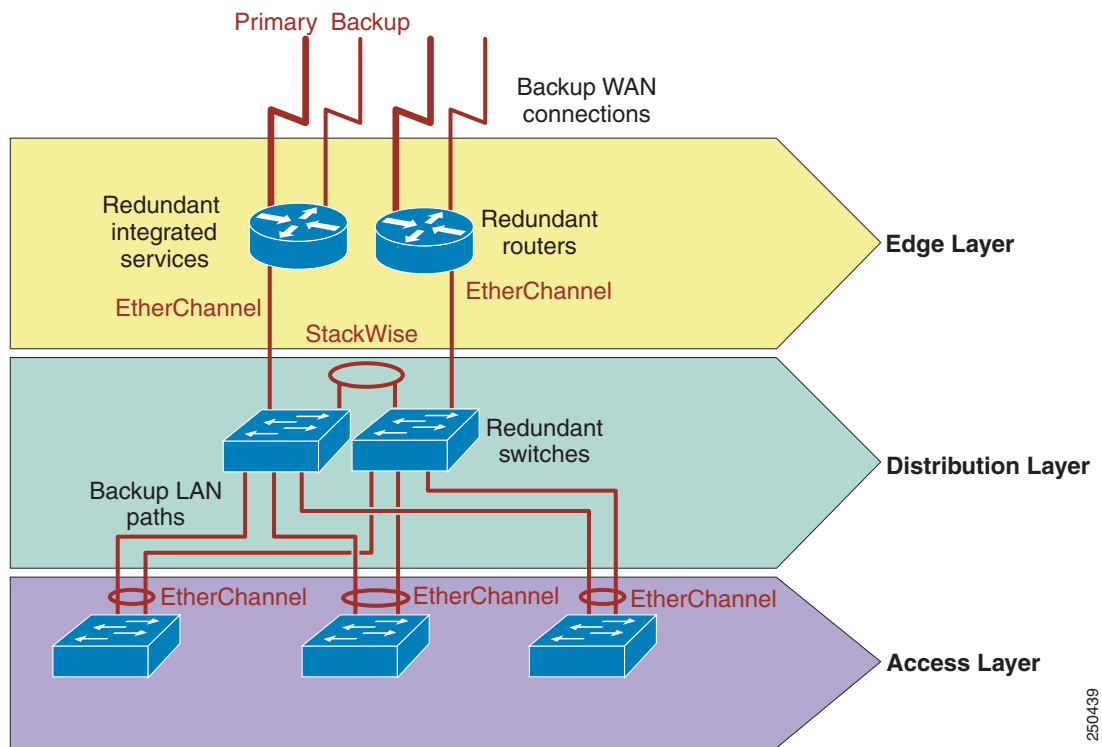
The benefits of a network design that provides high availability, rapid recovery and disaster recovery include the following:

- Availability: Network services are available to users when needed and as expected.
- Minimal time to repair: There are minimal disruptions when outages or failures occur.
- Transparent maintenance: Planned maintenance may be performed with minimal downtime.

The various mechanisms and features used in the different layers of the hierarchical network design to achieve high availability and rapid recovery are shown in Figure 30 and described in the following sections:

- Backup WAN Link, page 35
- Redundant Services, page 38
- Redundant Edge Router, page 40
- Backup LAN Links, page 41
- Redundant Distribution Switch, page 41

Figure 30 High Availability and Rapid Recovery Components



250439

The general purpose of high availability is to ensure operational continuity over a given period of time. High availability is expressed as a percentage of uptime in a given year (for example, 99.99 percent). Many external factors influence this availability measure (for example, cooling, backup power, administrator’s skills, available spares, and so on), and therefore it varies from one network implementation to another. Excluding these external factors, a system should be able to guarantee a specific degree of availability. The Services Ready Large Branch Network had a design goal of less than

one hour of downtime per year. However, the current version of the network has not been validated to determine whether the design goal has been met. Validation will be implemented and documented in the future updates to this guide.

Highly available networks are designed to tolerate small errors in favor of continued, uninterrupted system operation. A small error in the context of a Services Ready Large Branch Network implies outages that do not exceed a *significant* outage threshold. At or below the significant threshold level, users generally ignore errors in the system. [Table 5](#) lists the target significant outage thresholds. To achieve the target threshold levels, the system would require automatic stateful switchover to redundant hardware and software components for most of its services (especially voice and real-time video). At present, several documented limitations prevent stateful switchover for some of the network components in the Services Ready Large Branch Network. See the [“Redundant Edge Router” section on page 40](#) for a list of services that do not support stateful switchover. Therefore, in the context of the Services Ready Large Branch Network, high availability includes the concept of *acceptable* outage threshold. This is a threshold level at or below which users generally tolerate errors in the system and quickly reestablish their communication sessions. If outage events are infrequent (no more than one per week) and the acceptable threshold level is met during the outage, the branch network can be considered highly available.

Table 5 Target Significant Outage Thresholds

Application	Significant Outage Threshold	Acceptable Outage Threshold
IP Transport	10 seconds	Reconnection within 60 seconds
Voice	2.5 seconds	Redial within 60 seconds
Real-time Video	100 milliseconds	Redial within 60 seconds
Video on Demand	10 seconds	60 seconds

All switch and router configuration files should be stored on an external storage device to enable disaster recovery. The Services Ready Large Branch Network used two different methods of storing copies of configuration files:

- Backup to centrally located TFTP server
- Password protected USB flash drive

For more information about backup and restore of configuration files to/from TFTP server, visit:

http://www.cisco.com/en/US/products/sw/iosswrel/ps1835/products_tech_note09186a008020260d.shtml

The TFTP backup and recovery method provides fast and convenient access to the configuration files if they are needed for disaster recovery. However, because a centrally located server may not be accessible in all circumstances, locally stored USB flash token is also provided in the Services Ready Large Branch Network. Aladdin Knowledge Systems USB eToken, shown in [Figure 31](#), was selected for this purpose. It requires authentication to access the configuration files encrypted and stored on the device. The eToken itself should be stored in a secure, fire- and temperature-resistant container at the branch office.

Figure 31 Aladdin Knowledge Systems USB eToken and Cisco ISR



To learn more about the Aladdin eToken, visit:

http://www.cisco.com/en/US/prod/collateral/modules/ps6247/product_data_sheet0900aecd80232473.html

Backup WAN Link

Any of the WAN connectivity options that are described in “WAN Services” section on page 9 can be used as a backup link mechanism. In practice, however, PSTN and Internet based connections are primarily used for this purpose. The main considerations when selecting the backup link are:

- Service provider: The backup link should go through a different service provider network than the primary link. There should be no or minimal sharing of back-end infrastructure by the providers.
- Service availability: Selection of backup link service must take into account local availability.
- Availability and recovery requirements: The properties and type of service expected for the backup connection.
- Cost: The backup link cost must be evaluated based on how well it meets the availability requirements.

Table 6 lists advantages and disadvantages of the most commonly used backup connections for a branch office.

Table 6 Common WAN Backup Link Options for a Large Branch Office

Service Type	Advantages	Disadvantages	Appropriate for Branches
ISDN (PRI or BRI)	<ul style="list-style-type: none"> Concurrent data and voice transmission Symmetric and dedicated bandwidth Works over telephone wires 	<ul style="list-style-type: none"> Call setup Limited bandwidth 	<ul style="list-style-type: none"> Telephone wires are the only connection option, and the office is too far from POP for xDSL link. Voice is the primary traffic (use PRI). Diversify service provider for backup.
xDSL	<ul style="list-style-type: none"> Concurrent data and voice transmission Dedicated bandwidth Works over telephone wires Relatively high bandwidth 	<ul style="list-style-type: none"> Quality dependent on wiring and distance to POP Asymmetric bandwidth 	<ul style="list-style-type: none"> Appropriate for most branch offices.
Cable	<ul style="list-style-type: none"> High bandwidth 	<ul style="list-style-type: none"> Asymmetric bandwidth Shared bandwidth Less secure 	<ul style="list-style-type: none"> Require high bandwidth.
3G	<ul style="list-style-type: none"> Easy installation Small antenna No cabling 	<ul style="list-style-type: none"> Limited bandwidth Limited availability Unreliable link 	<ul style="list-style-type: none"> Locations without wiring. Diversify service providers for backup.
Satellite	<ul style="list-style-type: none"> Global coverage 	<ul style="list-style-type: none"> Link delay Unreliable link Large antenna 	<ul style="list-style-type: none"> Remote locations. Diversify service provider for backup.

In addition to these general considerations, a backup link must meet the business criteria outlined in the “Large Branch Design Considerations” section on page 4. At present, the Services Ready Large Branch Network has been validated only with SHDSL as a backup WAN link. In future updates to this guide, some of the other options listed in Table 6 will be validated and documented.

All WAN deployments described in the “WAN Services” section on page 9 provide a backup link to the central site. The traffic is encrypted and directed over the Internet as shown in Figure 40. The backup link connects the branch to the nearest location where the provider makes access to the Internet service available. The link can be set to standby mode and used only for backup when the primary WAN link fails, or it can stay active and provide access to the Internet using a split tunneling mechanism. Both of these options were validated in the design.

For the Services Ready Large Branch Network, the following connection option was selected for backup:

- A single broadband G.SHDSL link connected to the Cisco HWIC-4SHDSL interface is shown in Figure 32

Figure 32 4-Port Symmetric High-Speed DSL (SHDSL) WAN Interface Card (HWIC-4SHDSL)



To learn more about the Cisco HWIC-4SHDSL interface card, visit:

http://www.cisco.com/en/US/prod/collateral/modules/ps5949/ps7175/product_data_sheet0900aecd80581fa0.html

Physical connectivity for the xDSL line consists of one or multiple telephone wires terminated at a DSL access multiplexer (DSLAM) in the provider’s nearest point of presence (POP). The Cisco HWIC-4SHDSL comes with a cable that directly connects its single RJ-45 port to two telephone lines terminated at one of the supported DSLAMs. Table 7 identifies the WAN backup link, bandwidth, physical connection for the link, and Cisco ISR interface that provides access to the Internet provider’s network.

Table 7 WAN Backup Line Option

WAN Backup Line Type	Bandwidth	Physical Connection	ISR Interface or Module
SHDSL with IMA	4.6 Mb/s	Two twisted-pair telephone wires	HWIC-4SHDSL

- xDSL Connection

Digital subscriber line (DSL) technology is a popular option for connecting home office workers and small branch offices to the enterprise network. In a large branch office, it is used mainly as a backup link. DSL creates an always-on connection that uses existing telephone wires to transport high-bandwidth data and to provide IP-based services. A DSL modem converts digital signals to and from analog signals. At the telephone company POP, a DSLAM is used to redigitize the signal and forward it to the Internet service provider. There are various DSL standards, all under the general name xDSL, for various x. The Services Ready Large Branch Network office used single-pair high-speed DSL (G.SHDSL).

The universal choice of Layer-2 encapsulation protocol for use on xDSL lines is asynchronous transfer mode (ATM). ATM adaptation layer (AAL) is a mechanism for segmenting upper-layer information into ATM cells at the transmitter and reassembling them at the receiver. AAL5 provides support for segmenting and reassembling routed/switched protocols over ATM permanent virtual circuits (PVCs) using Logical Link Control Layer (LLC)/Subnet Access Protocol (SNAP) or virtual channel multiplexing (VCMUX). LLC/SNAP adds an extra header that allows multiplexing of multiple protocols over the same PVC circuit. VCMUX allows multiple virtual circuits (VCs) on the xDSL link and maps each protocol to a different VC. For simplicity, AAL5+SNAP encapsulation was chosen for the Services Ready Large Branch Network.

Inverse multiplexing over ATM (IMA) allows bundling of several xDSL lines to form a single logical link of higher combined bandwidth. Two telephone lines were bundled together in the Services Ready Large Branch Network to create a bandwidth of 4.6 Mb/s.

To learn more about IMA, visit:

http://www.cisco.com/en/US/tech/tk39/tk356/technologies_q_and_a_item09186a0080111162.shtml

In summary, the Services Ready Large Branch Network used the following xDSL configuration:

- G.SHDSL with 2-line IMA and AAL5+SNAP encapsulation

Redundant Services

The two branch routers have identical integrated services. This redundancy ensures high availability of individual components. The degree of failure recovery that can be provided by a redundant component depends on its ability to take over the load of the failed component (switchover). Switchover may be a manual operation (for example, CLI-invoked) or an automatic software- or hardware-initiated operation. Stateful switchover allows services to maintain a state between the active component and the standby component. This facilitates the speed and transparency of the switchover event. To achieve the significant outage threshold defined in [Table 5 on page 34](#), a service must support stateful switchover.

In general, Cisco IOS-based software-only services (for example, NAT and IPS) do not provide fault detection mechanism at the component level. Therefore, failure of software-only components in this context means that the Cisco IOS software failed as a whole. However, the goal of the Services Ready Large Branch Network validation effort is to minimize software failures of individual components. For the rest of this section, *services* refers to those services that have either dedicated hardware or are software components with built in fault-tolerance mechanism. All other software services are labeled as *Cisco IOS-based*.

The type of load sharing between redundant components also impacts the level of high availability of those components. There are two general cases:

- Redundant components do not share load (Active-Standby configuration).
- Some type of load sharing is enabled between the two components (Active-Active configuration).

The non-load sharing configuration (Active-Standby) is easier to configure and manage because it involves switchover only when there is a failure. In a load-sharing configuration, the traffic must be segmented to go to the appropriate component. That is, it is switched over to the active component in case of failure, and then switched back to load sharing after the component becomes operational. [Table 8](#) lists the switchover mechanisms available for the various hardware and fault-tolerant software components.

Table 8 Support of Various Dedicated Hardware and Fault-Tolerant Software Services for Failure-Forced Switchover

Service	Active-Standby		Active-Active	
	Switchover	Mechanism	Switchover	Mechanism
Cisco IOS-based	Varies by service	HSRP	Varies by service	GLBP
Cisco WAAS ¹	Stateless	Cisco WCCP ²	Stateless	Cisco WCCP
Cisco Unity Express	None (RR only)	N/A	None	N/A
VPN Module	Stateful	HSRP	Stateless	GLBP
Digital Voice Card	Stateless	HSRP	None	N/A
Analog Voice Card	Stateless	HSRP	None	N/A

1. WAAS = Wide Area Application Services.

2. WCCP = Web Cache Coordination Protocol.

Cisco IOS-based services are described in the “[Redundant Edge Router](#)” section on page 40. Cisco WAAS provides the Active-Standby stateless switchover mechanism through HSRP.

Cisco Unified Communications Manager Express (Cisco Unified CME) provides a built-in mechanism for high availability. One Cisco Unified CME is designated as active, and another is as backup. Switchover happens when the active designated Cisco Unified CME fails. To learn more about the Cisco Unified CME high-availability mechanism, visit:

http://www.cisco.com/en/US/docs/voice_ip_comm/cucme/admin/configuration/guide/cmestm.html#wpmkr1030108

Cisco Unified Survivable Remote Site Telephony (Cisco Unified SRST) is a backup client for the centralized Cisco Unified Communications Manager (Cisco Unified CM) server. If the WAN link fails, Cisco Unified SRST takes over to provide local telephony services. To ensure that Cisco Unified SRST continues to provide service in the event that both the WAN link and the primary router fail, Cisco Unified SRST is configured with the IP address of the HSRP virtual gateway. The virtual IP address ensures that local IP Phones properly register to Cisco Unified Communications Manager by using HSRP virtual IP as the default gateway. The voice gateway does not provide any fault-tolerance mechanism. However, in the presence of either Cisco Unified SRST or Cisco Unified CME, the voice gateway on the router with the active Cisco Unified SRST (or Cisco Unified CME) will be used.

Cisco Unity Express (CUE) does not provide any switchover mechanism. However, the module can be taken out from the failed router, inserted into the standby router using its inline insertion feature, and reconfigured to make previously recorded voice mail available. This option is not possible when using CUE Advanced Integration Module 2 (AIM2), because it is internal to the router and the router must be powered down to install it.

The VPN service module relies on HSRP to track the state of the interfaces or protocols on those interfaces, and to switch the router state to backup if one of the interfaces fails. The VPN service module offers stateful failover. However, because Zone-based Policy Firewall (ZPF) does not provide stateful switchover, it is not possible to leverage the stateful switchover of the VPN service module. HSRP is described in the “[Redundant Edge Router](#)” section on page 40.

The digital voice interface card provides stateless switchover through HSRP tracking. However, this switchover capability was not validated in the current version of this guide. Validation will be performed in a future update.

The analog voice interface card does not provide any switchover capabilities. However, if the card fails on the active router, a manual switchover (for example, shutting off the active router) can force the backup router with the working card to the active state. This enables rapid recovery of the lost service.

Redundant Edge Router

The Services Ready Large Branch Network features two routers to ensure continuous network availability. There are two options for configuring the routers: Active-Active and Active-Standby. In Active-Active configuration, both routers share the load. In Active-Standby configuration, one of the routers is in a standby mode until the active router fails, at which point the standby becomes active. Active-Standby router configuration provides network resiliency, but it also decreases network efficiency while the standby router sits idle. However, an Active-Standby network design is simpler to set up, manage, and troubleshoot than an Active-Active design. In the current version of this guide, only the Active-Standby configuration has been validated. With future updates, the Active-Active with GLBP configuration will be tested, and its configuration provided. The following Active-Standby options have been validated with the Services Ready Large Branch Network:

- Active-Standby configuration with Hot Standby Router Protocol (HSRP) and rapid recovery of voice mail

There are no Cisco IOS-based services in the Services Ready Large Branch Network that support stateful switchover; therefore, the goal of the design is to provide acceptable outage targets, as defined in [Table 5](#). In general, any active user sessions utilizing a stateful component that does not provide a stateful switchover will be terminated. However, users should be able to reestablish terminated sessions within 60 seconds. Stateless Cisco IOS-based services include:

- NAT/PAT
- Zone-based Policy Firewall (ZPF)
- Voice Gateway
- DHCP
- AAA

With HSRP, one of the two Cisco 3900 ISRs in the Services Ready Large Branch Network is designated as active, while the other is designated as standby. There is a path to the standby router in the event that the active router fails. Using HSRP, the two routers are connected to the same Ethernet segment. The routers work together to present the appearance of a single virtual router on the LAN. The routers share the same virtual IP address, and if the active router fails, the hosts on the LAN are able to continue forwarding packets to the standby router. The process of transferring the routing responsibilities from one device to another is transparent to the user.

One of the two routers was configured as primary for data traffic and standby for voice traffic, and the other router was configured as primary for voice traffic and backup for data traffic. This configuration maximizes the utilization of router resources and available WAN bandwidth by providing load sharing based on traffic type.

Several different events can trigger HSRP switchover. In the Services Ready Large Branch Network these events are router failure and primary WAN interface failure. There are two potential failure cases for the primary WAN interface: either the WAN link fails, or the WAN interface card fails. When an interface is tracked by HSRP, it is assigned a priority. When the WAN link goes down, the router that is currently active will switch over to its backup WAN link. The active router's overall priority is decremented by the interface priority value, but so is the standby router's, since they are most likely connected to the same primary WAN link service provider and since both interfaces are tracked by HSRP. Therefore, switchover does not happen, and the active router uses the backup WAN link. If, however, the interface card on the active router fails (or one primary link fails but

another continues to function), then HSRP forces a switchover to the backup router because its primary WAN interface is still up. In this situation, the active router's overall priority is decremented by the interface priority value, but the standby router's overall priority stays the same. Therefore, HSRP forces a switchover to the backup router.

To learn more about HSRP visit:

http://www.cisco.com/en/US/tech/tk648/tk362/technologies_tech_note09186a0080094a91.shtml

Backup LAN Links

The EtherChannel technology described in the “[EtherChannel Link Aggregation](#)” section on page 31 provides redundant paths between access switches and distribution layer switches. Moreover, if servers are connected to the distribution layer switches as described in the LAN design section, the servers could be equipped with two NICs to provide redundant paths to the distribution layer switches. At the access layers, IP Phones provide two Ethernet ports that could be used as redundant paths to access layer switches. In addition, wireless access points can be added to the network to provide redundant paths to PCs that are equipped with wireless access cards (or wired connections if a second NIC is provided for the PCs). Redundancy for IP Phones and PCs connected to the access layer is not a business requirement for a typical branch office and therefore not considered in the Services Ready Large Branch Network.

Redundant Distribution Switch

The Services Ready Large Branch Network provides two distribution switches in the Cisco StackWise configuration to ensure continuous LAN connectivity. The Cisco StackWise configuration is described in the “[Cisco StackWise Interconnects](#)” section on page 30. One of two switches is chosen to be the master switch to manage all switching tables. EtherChannel aggregates the two links to each of the switches into a single virtual link. If one distribution switch fails, the remaining link provides a path to the active switch, and takes over the failed switch's load.

IP Addressing and IP Routing

Cisco offers a broad portfolio of IP routing and addressing technologies. Only some of these technologies are relevant to branch offices. To meet the design criteria in the “[Large Branch Design Considerations](#)” section on page 4, the Services Ready Large Branch Network was deployed with the following IP routing and addressing services enabled in the Cisco IOS software on the routers:

- [Routing Protocols, page 43](#)
- [Multicast, page 45](#)
- [DHCP, page 45](#)
- [NAT and PAT, page 46](#)

When assigning IP addresses to the various devices in the branch office, it is important to follow the IP addressing scheme and conventions set for the entire enterprise network. Today, enterprises use classless IP addressing, classless IP routing protocols, and route summarization. The Services Ready Large Branch Network uses a private addressing scheme allocated from the 10.0.0.0/21 address pool that has 2046 available hosts. The design assumes that a single user will need two IP addresses: one for the PC and another for the IP Phone. Two backup subnets are provided to create non-overlapping DHCP pools in case the primary DHCP server fails. The other addresses are used for server and network devices, or are left unallocated.

The address pool is divided among VLANs as follows:

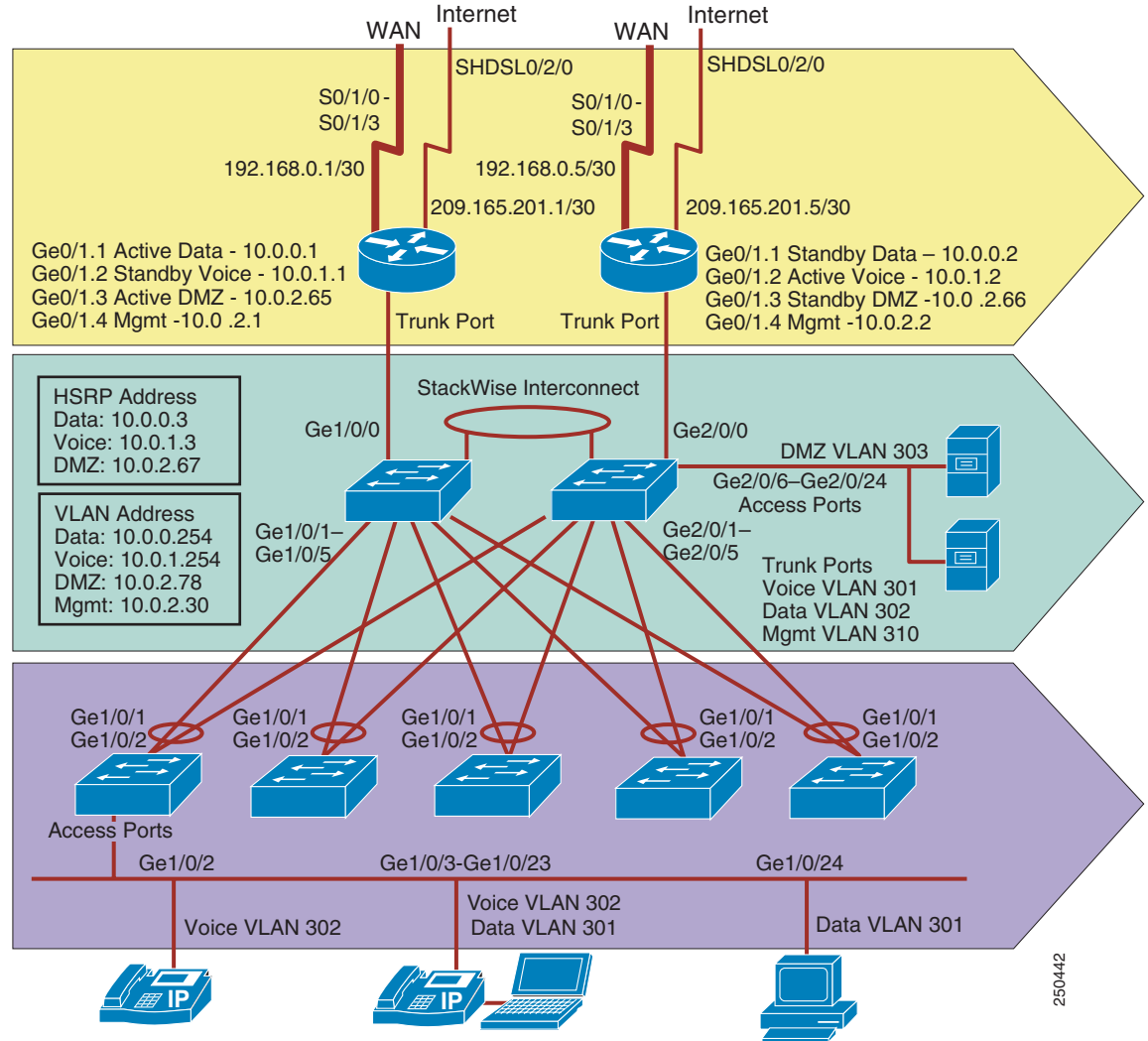
- Voice VLAN: 254 addresses
- Data VLAN: 254 addresses
- DMZ VLAN: 14 addresses
- Backup Voice VLAN: 254 addresses
- Backup Data VLAN: 254 addresses
- Management VLAN: 30 addresses
- Black hole VLAN: 30 addresses

[Table 9](#) shows the address assignment, and [Figure 33](#) shows the corresponding topology. The addressing scheme is only an example. Each enterprise should follow its own addressing scheme.

Table 9 *Sample Address Assignment Scheme for the Services Ready Large Branch Network*

Component	Network
Data VLAN	10.0.0.0/24
Voice VLAN	10.0.1.0/24
Management VLAN	10.0.2.0/27
Black Hole VLAN	10.0.2.32/27
DMZ VLAN	10.0.2.64/28
Backup Data VLAN	10.0.3.0/24
Backup Voice VLAN	10.0.4.0/24

Figure 33 Sample Address Assignment for the Services Ready Large Branch Network



250442

Routing Protocols

Several routing protocols are relevant to the branch office. Although there are design differences among these routing protocols, all have a common goal of stability, availability, fast convergence, and high performance. However, no one protocol is best suited for all situations, and trade-offs must be considered when deciding on the appropriate one. The following are the most common routing protocols:

- **Static routing:** Manually defined routes as next hops to various destinations. Static routes are generally used in very small networks or when the routing is managed by the service provider. In a branch, a static route is typically used to forward traffic to the Internet service provider network.

For more information about static routes, visit:

http://www.cisco.com/en/US/tech/tk365/technologies_tech_note09186a00800ef7b2.shtml

- **Routing Information Protocol version 2 (RIPv2):** Distance vector protocol now considered a legacy. It should be used only in small legacy networks that have little need to grow.

For more information about RIP, visit:

http://www.cisco.com/en/US/tech/tk365/tk554/tsd_technology_support_sub-protocol_home.html

- Enhanced Interior Gateway Routing Protocol (EIGRP): Enhanced distance vector protocol proprietary to Cisco. Unlike traditional distance vector protocols, EIGRP does not age out routing entries or uses periodic updates. The Distributed Update Algorithm (DUAL) algorithm is used to determine the best path to a destination network. The EIGRP protocol maintains a topology table that includes both the best path and any loop-free backup paths. When a route becomes unavailable, the DUAL algorithm finds the best backup path to the destination. The protocol uses bandwidth and delay to select the preferred path, and can optionally include link reliability and jitter. EIGRP works best in small to medium-sized networks that have a flat design and use only Cisco routers.

For more information about EIGRP, visit:

http://www.cisco.com/en/US/tech/tk365/tk207/tsd_technology_support_sub-protocol_home.html

- Open Shortest Path First (OSPF): Link state protocol standardized by IETF. OSPF floods link state information to its neighbors and builds a complete view of the network topology. The Shortest Path First (SPF) algorithm is used to determine the best path to a destination. The protocol uses bandwidth to determine the best path, or can be optionally forced to use a manually defined cost for a path. OSPF works best in networks that are large, have a hierarchical design, have a mixture of Cisco and non-Cisco routers, are expected to grow to a large scale, or require fast convergence time.

For more information about OSPF, visit:

http://www.cisco.com/en/US/tech/tk365/tk480/tsd_technology_support_sub-protocol_home.html

- External Border Gateway Protocol (eBGP): An exterior gateway path vector protocol. The eBGP protocol is used to exchange routing information between different autonomous systems. In general, eBGP is not used in branch routers unless there are special considerations, such as connecting to two service providers and actively using both links or when routing information needs to be exchanged with the service provider when there are downstream routers, especially for MPLS-based WANs.

For more information about eBGP, visit:

http://www.cisco.com/en/US/tech/tk365/tk80/tsd_technology_support_sub-protocol_home.html

Choosing the appropriate routing protocol in most cases depends on the routing protocol currently used in the enterprise network. Therefore, to ensure its relevance and applicability, the Services Ready Large Branch Network was validated with all of the routing protocols listed.

In all WAN deployments, with the exception of Layer 3 Virtual Private Network (L3VPN), the enterprise manages routing. RIPv2, EIGRP, or OSPF is used to route traffic on the primary link. Both the primary and backup links have a default static route to either the PE or the ISP router. With a standby mode backup interface configuration, the backup default route is automatically inserted into the routing table only after the backup interface becomes active. With an active mode backup interface configuration, the primary default route was assigned a lower cost than the backup default route. The primary default route became active and started directing Internet traffic to the central site only after the backup link failed, and its default route was removed from the routing table.

The eBGP protocol was added to route the backup WAN link traffic. Generally, either eBGP or an IGP can be used to advertise the customer's public routes to the provider's ISP router. In the case of eBGP, the branch routers directly connected networks are distributed into the global instance of the local eBGP process. Then, an eBGP session between the branch router and the provider's ISP router is used to advertise the networks to the provider's ISP router as standard BGP updates. In the case of IGPs, the protocols advertise the directly connected networks to the provider's ISP router. In general, the service provider disables routing updates from the ISP router to prevent Internet routing tables from propagating into the branch router. If this disabling is not provided, the branch router can filter out routing updates in order to minimize the size of its routing table.

VPN access by the Services Ready Large Branch Network is accomplished by the following:

- Split Tunneling

The Services Ready Large Branch Network provides direct access to the Internet through split tunneling. To access the Internet, NAT and PAT are used to map the branch network private addresses to public addresses. See the “NAT and PAT” section on page 46. Split tunneling is accomplished by running a separate routing process for the Internet-bound traffic. There are four options for split tunneling in the Services Ready Large Branch Network, depending on the type of VPN used for the primary link and whether the backup interface is in active or standby mode. The “Routing Protocol Implementation” section on page 30 provides detailed configurations. The following are the four different options:

- Active/Standby Primary/Backup WAN links with DMVPN
- Active/Standby Primary/Backup WAN links with GETVPN
- Active/Active Primary/Backup WAN links with DMVPN
- Active/Active Primary/Backup WAN links with GETVPN

- Remote User Access

In the Services Ready Large Branch Network, remote office workers have direct access to the DMZ VLAN over SSL VPN. The users connect to the SSL VPN gateway that is running in the branch office.

Multicast

IP multicast was enabled in the Services Ready Large Branch Network for applications that take advantage of multicast technologies, such as video conferencing, corporate communications, distance learning, and distribution of software. Cisco Protocol Independent Multicast (PIM) was used to forward multicast traffic. The protocol leverages the router's unicast routing table populated by IGP protocols to maintain a multicast routing table that is used strictly for multicast traffic. PIM does not send routing updates, and it relies on IGP protocols to keep routing information up-to-date.

There are several modes of operation for PIM. In dense mode, the router floods multicast traffic to all interfaces except the one through which the multicast packet arrived. In sparse mode, multicast receivers request multicast traffic to be forwarded to their network segment. This information is propagated between the PIM-enabled network nodes. Sparse-dense mode allows an interface to be configured in both modes in order for different multicast groups to leverage either propagation mechanism.

To learn more about multicast, visit:

http://www.cisco.com/en/US/tech/tk828/tech_brief09186a00800a4415.html

DHCP

Dynamic Host Control Protocol (DHCP) was enabled in the Services Ready Large Branch Network to automatically assign and manage end device IP addresses from specified address pools within the router. There are two DHCP servers on each of the routers. At present, it is not possible to set the servers in an Active/Standby configuration. However, the DHCP protocol allows for a level of control that enables one of the DHCP servers to become the primary address assigner for the network. This can be accomplished by customizing the DHCP lease reservation mechanism.

When a DHCP-enabled end device is connected to the network, the end device first sends out a DHCP discovery request. Then, any available DHCP server offers a lease for an IP address to the end device. However, before the IP address can be assigned, the DHCP server must first check that no other device is currently using this same address. To perform this check, the DHCP server pings the address and waits

for the response. By increasing the amount of time one of the servers has to wait for the ping response and increasing the number of pings it sends, it is possible to ensure that one of the servers always leads the other with the IP lease offer from its pool of addresses. When the end device receives a lease offer, it then returns a formal request for the offered IP address to the originating DHCP server. The server confirms that the IP address has been exclusively allocated to the end device. If the primary DHCP server fails, the secondary server provides the IP address from its address pool, although the secondary DHCP server takes slightly longer than the primary DHCP server.

Any servers running in the branch should use static addressing. Only PCs and IP Phones should rely on DHCP for address assignment. There is a special consideration for IP Phones. They must be registered with Cisco Unified Communications Manager (Cisco Unified CM). If the active router fails, a lease renewal would force the phones to reregister with the Cisco Unified CM or Cisco Unified Survivable Remote Site Telephony (Cisco Unified SRST) agent, which would make the phones unavailable for the period of reregistration. To avoid this reregistration process, the Services Ready Large Branch Network splits the available IP addresses into two non-overlapping address pools, with each assigned to one of the DHCP processes on the active and standby routers. The DHCP lease is set to one month. If the active router experiences a catastrophic failure, but manages to be repaired and come back up within two weeks, no phone reregistration will occur.

To learn more about Cisco IOS DHCP server, visit:

http://cco.cisco.com/en/US/docs/ios/12_0t/12_0t1/feature/guide/Easyip2.html

NAT and PAT

To access the Internet directly from the branch office, Network Address Translation (NAT) or Port Address Translation (PAT) is needed to map the private addresses of the branch network to valid public addresses. When a packet comes to the router, NAT rewrites the source address in the IP header. The router tracks this translation. When return traffic comes back, the destination address will be rewritten to its original value. PAT adds the ability to rewrite port numbers, thereby increasing the number of times that a single public address can be used for translation. NAT and PAT were enabled to allow multiple hosts from the private branch network to access the Internet by using a single shared public IP address and various port numbers.

To learn more about NAT and PAT (also referred to as *NAT Overloading*), visit:

http://www.cisco.com/en/US/tech/tk648/tk361/technologies_tech_note09186a0080094831.shtml

Quality of Service

- [Classification and Marking, page 51](#)
- [Policing and Markdown, page 52](#)
- [Scheduling, page 52](#)
- [Shaping, page 53](#)
- [Scavenger Class QoS, page 53](#)
- [Security Services, page 53](#)

An enterprise branch must support a variety of user applications, and some applications are more sensitive than others to packet delay, loss, and jitter that exceed tolerable levels when multiple users share limited network resources. Business-critical applications tend to be sensitive to delays and packet loss, real-time applications have strict delay and jitter requirements, and other types of applications may

impose additional requirements. QoS is a set of tools and techniques for managing network resources in order to provide different priorities to different applications or to guarantee them a certain level of performance.

For more information about QoS and the various tools available in Cisco IOS software see the *Enterprise QoS Solution Reference Network Design Guide* at:

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

QoS policies vary from one enterprise to another, as each policy reflects particular business and organizational objectives. To meet the business criteria outlined in the “[Large Branch Design Considerations](#)” section on page 4, the Services Ready Large Branch Network adopted a hierarchical QoS model that is configured to support eight classes of traffic flows. The eight-class model specifically includes voice, interactive video, call signaling, internetwork control, transactional data, bulk data, best effort, and scavenger classes, as shown in [Table 10](#). The designated classification conforms to the Cisco QoS Baseline and RFC 3246.

Table 10 QoS Eight-Class Model

Application	Layer 3 Classification			Layer 2 CoS/MPLS EXP
	IPP	PHB ¹	DSCP	
Internetwork control	6	CS6	48	6
Voice	5	EF	46	5
Interactive video	4	AF41, AF42	34, 36	4
Call signaling/ Critical Data	3	AF31, CS3	24, 25, 26	3
Transactional data	2	AF21, AF22	18, 20	2
Bulk data	1	AF11, AF12	10, 12	1
Scavenger	1	CS1	8	1
Best effort	0	0	0	0

1. PHB = per hop behavior.

Each class of traffic carries a specific service level requirement. For the eight classes selected, the requirements are as follows:

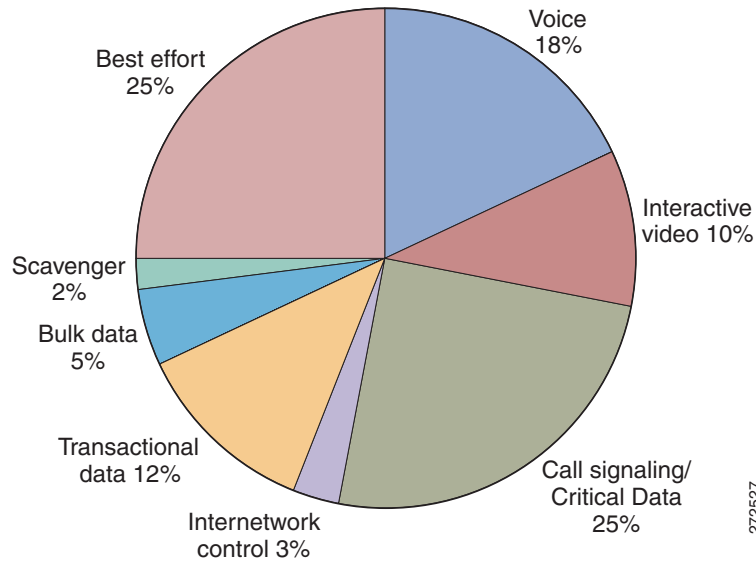
- Voice
 - Loss should be no more than 1 percent.
 - One-way latency (mouth-to-ear) should be no more than 150 ms.
 - Average one-way jitter should be targeted under 30 ms.
- Interactive Video
 - Loss should be no more than 1 percent.
 - One-way latency should be no more than 150 ms.
 - Jitter should be no more than 30 ms.
 - Overprovision interactive video queues by 20 percent to accommodate bursts.

- Call Signaling/Critical Data
 - Voice control traffic requires 150 bps (plus Layer 2 overhead) per phone of guaranteed bandwidth. A higher rate may be required, depending on the call signaling protocol(s) in use.
 - Mission-critical data traffic must have an adequate bandwidth guarantee for the interactive foreground operations that it supports.
- Internetwork Control
 - IGPs are usually adequately protected with the Cisco IOS internal PAK_Priority mechanism; we recommend that EGPs such as BGP have an explicit class for IP routing with a minimal bandwidth guarantee.
- Transactional Data
 - Transaction data traffic should have an adequate bandwidth guarantee for the interactive, foreground operations it supports.
- Bulk Data
 - Bulk data traffic should have a moderate bandwidth guarantee, and should be constrained from dominating a link.
- Best Effort
 - Adequate bandwidth should be assigned to the best-effort class as a whole, because the majority of applications will default to this class; reserve at least 25 percent for best-effort traffic.
- Scavenger
 - Scavenger traffic should be assigned the lowest configurable queuing service; for instance, in Cisco IOS this would mean assigning a Class-Based Weighted Fair Queuing (CBWFQ) of 1 percent to the scavenger class.

Figure 34 shows allocation of bandwidth to the eight QoS classes. The Eight-Class QoS Model allocates bandwidth to the general traffic categories as follows:

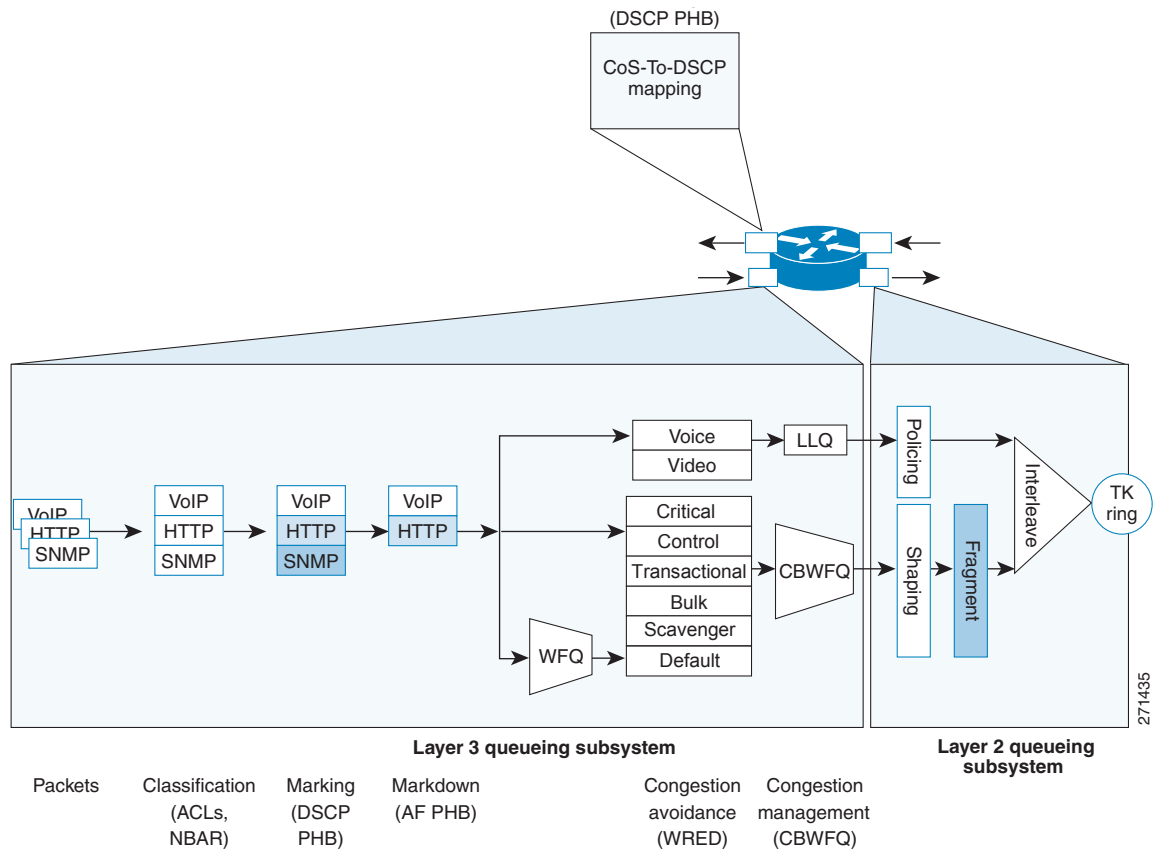
- Real-time traffic (voice and interactive video): 28 percent
- Scavenger and bulk traffic: 7 percent
- Best effort traffic: 25 percent
- Critical data traffic: 25 percent
- Transactional Data and Internetwork traffic: 15%

Figure 34 Bandwidth Allocation for Eight-Class QoS Model



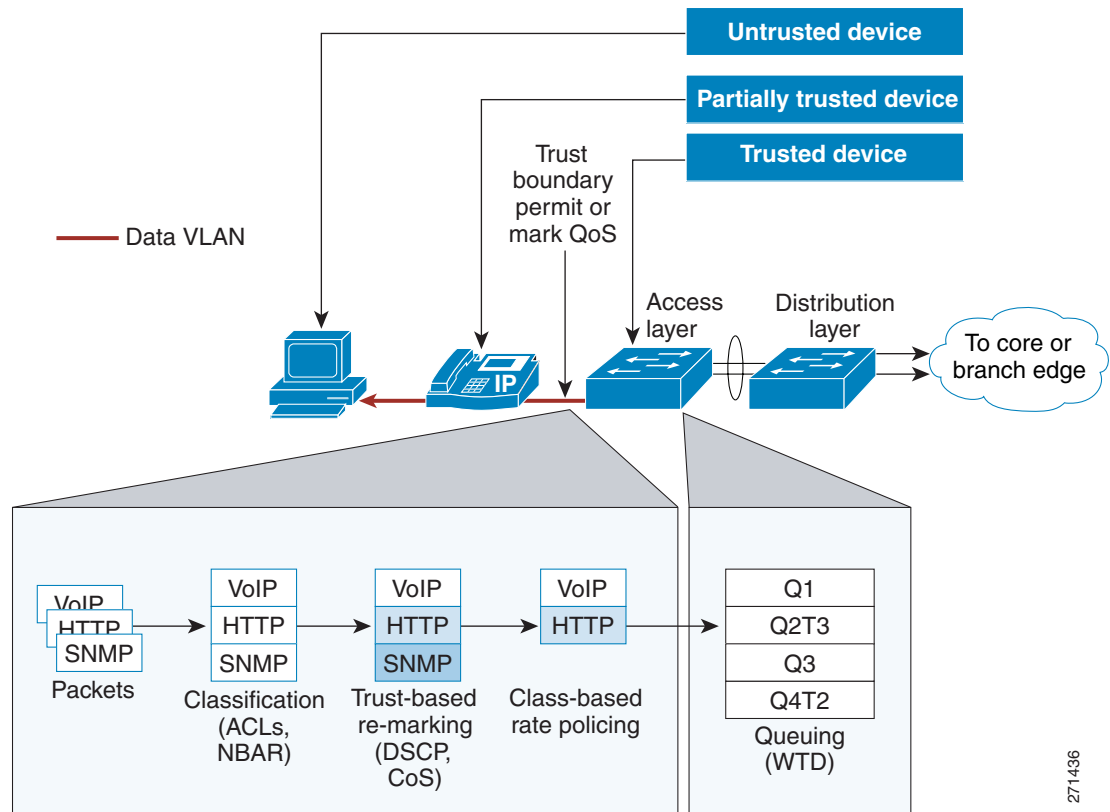
There are various ways to enable QoS in an enterprise branch network. The Eight-Class QoS policy is implemented in two logically different places in the network. A part of the policy is implemented at the access and distribution layers, and another part is implemented at the WAN edge layer. [Figure 35](#) and [Figure 36](#) shows summaries of QoS features that are part of the Services Ready Large Branch Network and their different implementation points. This design conforms to the Differentiated Services (DiffServ) architecture, as defined in RFC 2475.

Figure 35 WAN Router



271435

Figure 36 LAN Switch



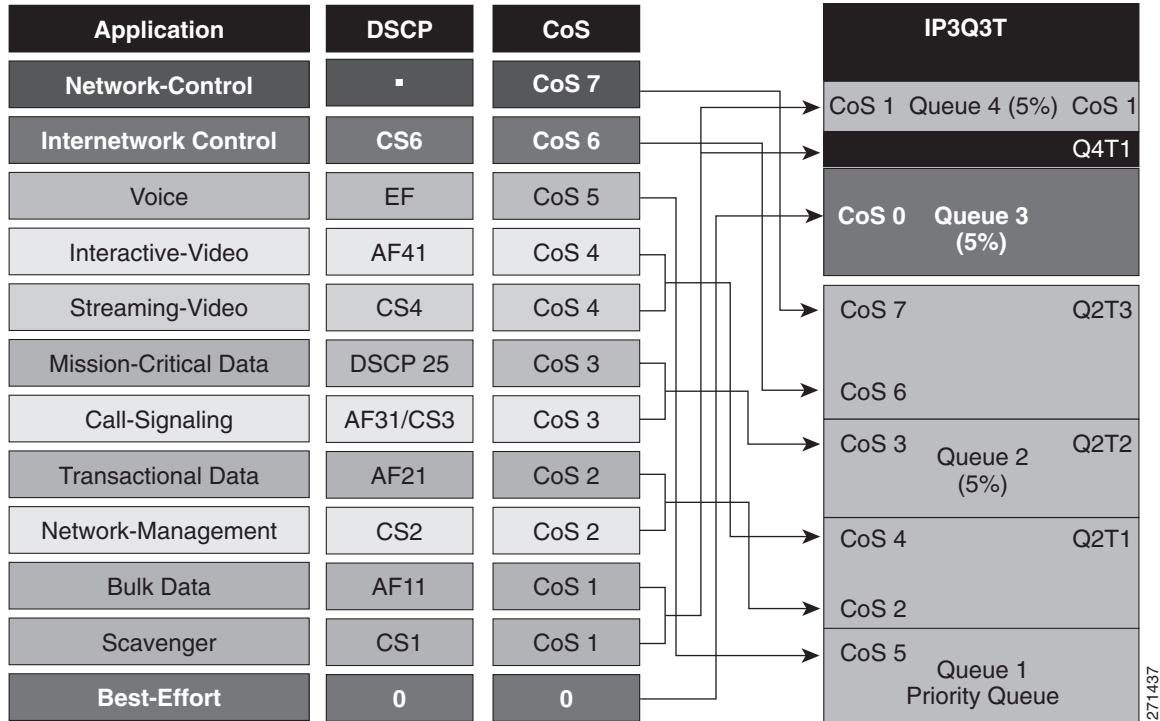
271436

Regardless of the implementation point, the design incorporated a set of common QoS design principles. These principles are described in the following sections.

Classification and Marking

Classification identifies packets belonging to a certain traffic class, based on one or more TCP/IP header fields as defined in the Access Control List (ACL), or in application signatures via Network Based Application Recognition (NBAR). Marking tags the classified traffic by modifying either the 802.1Q/p class of service (CoS) Ethernet header field for incoming traffic or the DSCP per-hop behavior (PHB) header bits for outgoing traffic. Applications are classified and marked as close to their sources as technically and administratively feasible. Access layer switches remark all the packets coming from PC endpoints, servers, and so on, with appropriate CoS/DSCP values. Voice and signaling packets coming out of Cisco IP Phones are trusted, but all the packets coming from PCs attached to the IP Phones are re-marked. Figure 37 shows assignment of different traffic flows to corresponding DSCP PHB and 802.1Q/p CoS classes. In addition, the assignment of each class to the corresponding Catalyst 3560 queue is shown.

Figure 37 Traffic Flow to QoS Class Mapping



271437

Policing and Markdown

Policing determines whether packets are conforming to administratively defined traffic rates, and marks, re-marks, or drops nonconforming traffic flows. Excess traffic is marked down according to the Assured Forwarding PHB Group (RFC 2597) rules. Traffic flows are policed and marked down as close to their sources as possible. Traffic leaving access layer switches was rate limited. Policing is enabled on the outgoing WAN interface.

Scheduling

Scheduling determines how a frame or packet exits a device. The Weighted Random Early Detection (WRED) algorithm provides for congestion avoidance on network interfaces by providing buffer management and allowing TCP traffic to throttle back before buffers are exhausted. This helps avoid tail drops and global synchronization issues, thereby maximizing network utilization and TCP-based application performance.

Queuing techniques such as weighted fair queuing (WFQ), CBWFQ, and low latency queuing (LLQ) are necessary to ensure that critical applications are forwarded even during network congestion. Real-time applications such as voice or video that need to be forwarded with the least latency and jitter use LLQ. Non-delay-sensitive traffic can use CBWFQ. Best-effort data has several queues using WFQ.

Queuing comes into effect automatically only when the amount of traffic exceeds the available bandwidth.

Shaping

Shaping delays excess traffic that is above an administratively defined rate. It uses a buffer to hold packets when the data rate is higher than expected. Shaping was performed on the WAN interface.

Scavenger Class QoS

QoS can also provide network security by using scavenger class QoS. The scavenger class QoS strategy identifies known worms and attacks. Traffic patterns from that end user that are considered “unusual” or as “normal traffic but at an unusually high rate” are marked as scavenger class (CS1) in the DSCP field and allowed to pass through the switch. Through the use of the scavenger class, QoS can be used as a security mechanism to limit the arrival rate of any traffic that is destined for the firewall or Cisco IOS IPS configurations. The Services Ready Large Branch Network also uses scavenger class QoS for excess traffic on the data VLAN.

Security Services

Security services help to protect the branch network from unauthorized, malicious, or inadvertent use of network resources. The challenge in designing the network is to find a balance between the need to keep networks open to support critical business requirements and the need to protect business-sensitive information. The Services Ready Large Branch Network strikes this balance by using technology and best practices that provide protection against the most common security threats.

Cisco offers a large number of products, features, and recommendations for securing a network. This design blueprint focuses on security guidelines and security features for services that are integrated into the branch office router and branch office switch. For comprehensive coverage of the subject, see the *Enterprise Branch Security Design Guide* at:

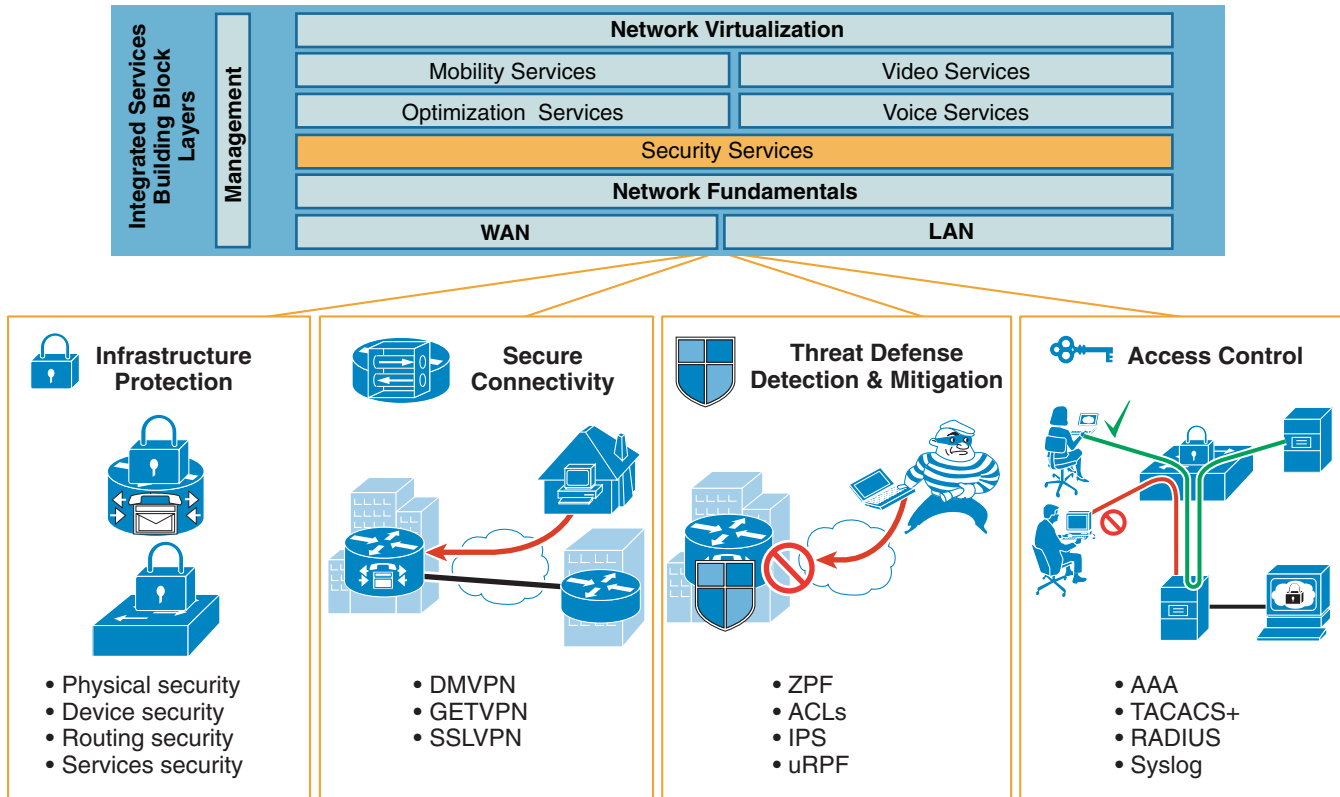
http://www.cisco.com/en/US/docs/solutions/Enterprise/Branch/E_B_SDC1.html

Providing effective security starts with establishment of a security policy for the branch network. A security policy provides a set of rules by which people who have access to the network resources must abide. RFC 2196 Site Security Handbook provides a good starting point for development of a branch office security policy. In addition, SANS Institute (www.sans.org) provides guidelines for developing comprehensive security policies for enterprises of various sizes.

Security services for a large branch office network are described in the following sections and shown in [Figure 38](#):

- [Infrastructure Protection, page 54](#)
- [Access Control, page 56](#)
- [Secure Connectivity, page 58](#)
- [Threat Protection, Detection, and Mitigation, page 65](#)

Figure 38 Security Services Building Blocks



In addition to following the guidelines and implementing security features recommended in this guide, it is important to emphasize that providing security for the branch network is an ongoing activity. Security threats evolve, and vulnerabilities are uncovered almost daily. Therefore, it is critical for the branch network to undergo continuous monitoring, periodic security assessment, and policy review.

While technology can create high enough barriers to prevent security breaches, the most costly security violations tend to be caused by either low-tech methods or unauthorized employees. Therefore, it is also critical to provide physical security and to ensure that security procedures are enforced at every level in the enterprise.

Infrastructure Protection

Infrastructure protection provides proactive measures to protect the branch routers and switches from direct attacks and indirect misuse. Infrastructure protection assists in maintaining network service continuity and availability. To protect network devices, the following methods are used in the Services Ready Large Branch Network:

- Physical security: Place routers and switches in a locked, temperature- and humidity-controlled room or cabinet accessible only by authorized administrators.
- Device security: Harden network devices.
 - Securing unused ports: Any ports not in use are disabled, autonegotiated trunking is turned off, and the ports are placed into the black hole VLAN.
 - Enabling Secure Shell (SSH): SSH is enabled and Telnet is disabled to prevent snooping and unauthorized access by unwanted parties. SSH is configured with five login retries.

- Enabling secure web access: HTTPS access should be used for management applications.
- Enabling VTY, console, and AUX timeouts, and ACLs: Set all VTY, console, and AUX ports with timeouts to automatically drop any idle sessions after 300 seconds. ACLs are applied to restrict access to the network devices and permit only specific protocols for administrative and monitoring purposes.
- Providing banner message: It is a security best practice to provide a banner to inform unauthorized users that access to the device is restricted.
- Routing protocol security:
 - Configure protocol authentication: MD5 algorithm is used to authenticate routing protocol packets. In addition, RIPv2 has all interfaces, except for the primary, set to passive mode.
- Network services security:
 - Turning off unnecessary services: Turning off unnecessary services means disabling any known potentially hazardous interface features and any global services not specifically required in the network. [Table 11](#) lists services available on the branch router that should be disabled if not used.

Table 11 Router Services That Should Be Disabled If Unused

Feature	Description	Default	Action
Cisco Discovery Protocol (CDP)	Layer 2 device discovery protocol	Enabled	Disable
TCP small servers	TCP network services	Disabled	
UDP small servers	UDP network services	Disabled	
Finger	User lookup service	Disabled	
Identification service	Device identification service	Disabled	
BOOTP	Legacy service for obtaining IP addresses	Enabled	Disable
Autoloading	Autoloading of configuration from TFTP	Disabled	
Classless routing	Forwarding packets with no specific route to the best supernet route	Enabled	Disable unless required
HTTP server	Used for web-based configuration	Enabled	Disable and use HTTPS
HTTPS server	Used for web-based configuration	Enabled	Disable if not used
FTP server	Used to copy configuration files	Disabled	
DNS server	Name resolution	Enabled	Disable or enable explicit server if needed
PAD	Packet assembler/disassembler	Disabled	

Table 11 Router Services That Should Be Disabled If Unused (continued)

Feature	Description	Default	Action
IP source routing	Packet-specified routing	Enabled	Disable on all interfaces
Proxy ARP	Proxy for Layer 2 address resolution	Enabled	Disable on all interfaces
IP redirects	ICMP ¹ redirect message	Enabled	Disable on WAN interfaces
ICMP unreachable	Incorrect IP address notification	Enabled	Disable on WAN interfaces
Directed broadcast	Packet specified broadcast	Enabled	Disable on all interfaces
ICMP mask reply messages	Replies to subnet mask queries	Disabled	Disable on WAN interfaces
MOP	Maintenance Operation Protocol for loading Cisco IOS images	Disabled	

1. ICMP = Internet Control Message Protocol.

To simplify the steps for providing network device protection, the Services Ready Large Branch Network used the AutoSecure feature of Cisco IOS software. It is a single interactive command that disables all nonessential system processes and services as previously described. In addition, it enables several services that improve security, including:

- Tuning of scheduler interval and allocation
- TCP syn wait time
- TCP keepalive messages
- ICMP unreachable messages
- Enables Cisco Express Forwarding (CEF)
- Provides antispoofing
- Blocks all IANA-reserved address blocks
- Blocks all private address blocks

To learn more about AutoSecure, visit:

http://www.cisco.com/warp/public/cc/pd/iosw/prodlit/cas11_ds.htm

Access Control

Access control is a mechanism for verifying user identity, restricting access to network resources, and auditing usage. Three independent security processes—authentication, authorization, and accounting—are used for this purpose. The processes perform the following functions:

- Provide a method for identifying users, verifying their identity, and granting/denying access to the network resources through mechanisms such as login and password or challenge and response.

- Provide a method for controlling access to network resources by authenticated users through mechanisms such as user groups, various access levels, privileges, or explicit user/group resource assignment (and vice versa).
- Provide a method for auditing the network to ensure compliance with security policies or to monitor attempts of unauthorized use.

Cisco offers several mechanisms to perform the authentication, authorization, and accounting processes independently as well as an integrated architectural framework that consistently enforces security policies across the entire network. The Services Ready Large Branch Network used a mixture of independent mechanisms and an integrated framework to reinforce and expand access control coverage. Authentication Authorization Accounting (AAA) service is used as the integrated framework to perform the eponymous identity and access control processes.

When AAA is activated, the network device on which it is running verifies security information and reports user activity to the RADIUS or TACACS+ security server on the network. The Services Ready Large Branch Network was validated with both RADIUS and TACACS+. The two servers provide the following functions:

- **RADIUS:** Distributed client/server system implemented through AAA that secures networks against unauthorized access. RADIUS clients run on routers and switches and send authentication requests to a central RADIUS server that contains all user authentication and network service access information.

To learn more about RADIUS, visit:

http://www.cisco.com/en/US/docs/ios/12_2/security/configuration/guide/scfrad.html

- **TACACS+:** Security application implemented through AAA that provides centralized validation of users attempting to gain access to a router or network access server. TACACS+ services are maintained in a database on a TACACS+ daemon running, typically, on a UNIX or Windows NT workstation. TACACS+ provides for separate and modular authentication, authorization, and accounting facilities.

To learn more about TACACS+, visit:

http://www.cisco.com/en/US/docs/ios/12_2/security/configuration/guide/scftplus.html

Authentication

Authentication identifies the user through a login and verifies the user's identity through a password (or challenge/response in case of a software process). Authentication is the first gate that must be crossed to gain access to the system. If the login is found, the user is identified. If the password matches, then the user's identity is verified. If the login is not found or the password does not match, then the user is denied access. The following measures were taken to provide authentication in the Services Ready Large Branch Network:

- **Password management:** Password management ensures that only approved users can access a device or services within the network. Strong passwords that are at least 8 characters, combining letters, numbers, and symbols and avoiding dictionary words, numbers, or dates are recommended. Passwords should be changed frequently. The Services Ready Large Branch Network uses Type 5 encryption for storing administrative passwords in the configuration file as well as the Cisco IOS password encryption feature. In addition, all devices mandate a minimum of an 8-character password length.
- **VTY, console, and AUX passwords:** All access mechanisms on all devices are guarded by administrative passwords.

- AAA authentication: A list of authentication methods that are applied to the various interfaces is created. The method list defines the types of authentication to be performed and the sequence in which they will be performed. All authentication methods, except for local, line password, and enable authentication, are defined through AAA.

Authorization

In the simplest terms, authorization defines the network resources accessible to an authenticated user. There are two orthogonal methods for implementing authorization. Either the user is associated with all resources accessible to that user, or a resource is associated with all users that have access to that resource. A user can have different privilege levels for a resource (for example, list, read, write, execute). To simplify management and speed up the authorization process, users are assigned to groups (for example, administrator). Group membership defines which resources can be accessed by the user. Temporal authorization provides a mechanism to grant count- or time-based access to specified resources. The following measures were taken to provide authorization in the Services Ready Large Branch Network:

- AAA authorization: Assembles a set of attributes that describe what the user is authorized to perform. These attributes are compared to the information contained in a database for a given user, and the result is returned to AAA to determine the user's actual capabilities and restrictions. The database is located on a server at the central site. As with authentication, a named list of authorization methods is created and is applied to various interfaces.

Accounting

As the name implies, accounting tracks access by users to various resources. Accounting is used to audit the network to ensure full compliance with security policies or to identify security breaches. The following measures were taken to provide accounting in the Services Ready Large Branch Network:

- Enabling logging: Access control of Simple Network Management Protocol (SNMP) and syslog on the router and switches is configured to ensure that there is a tracking mechanism when any unusual activity occurs. For more information about logging see the “[Management Services](#)” section on [page 68](#).
- AAA accounting: Provides a method for collecting and sending security server information used for auditing, and reporting, such as user identities, start and stop times, executed commands, and packet and byte counts. As with authentication and authorization, a named list of accounting methods is created and applied to various interfaces.

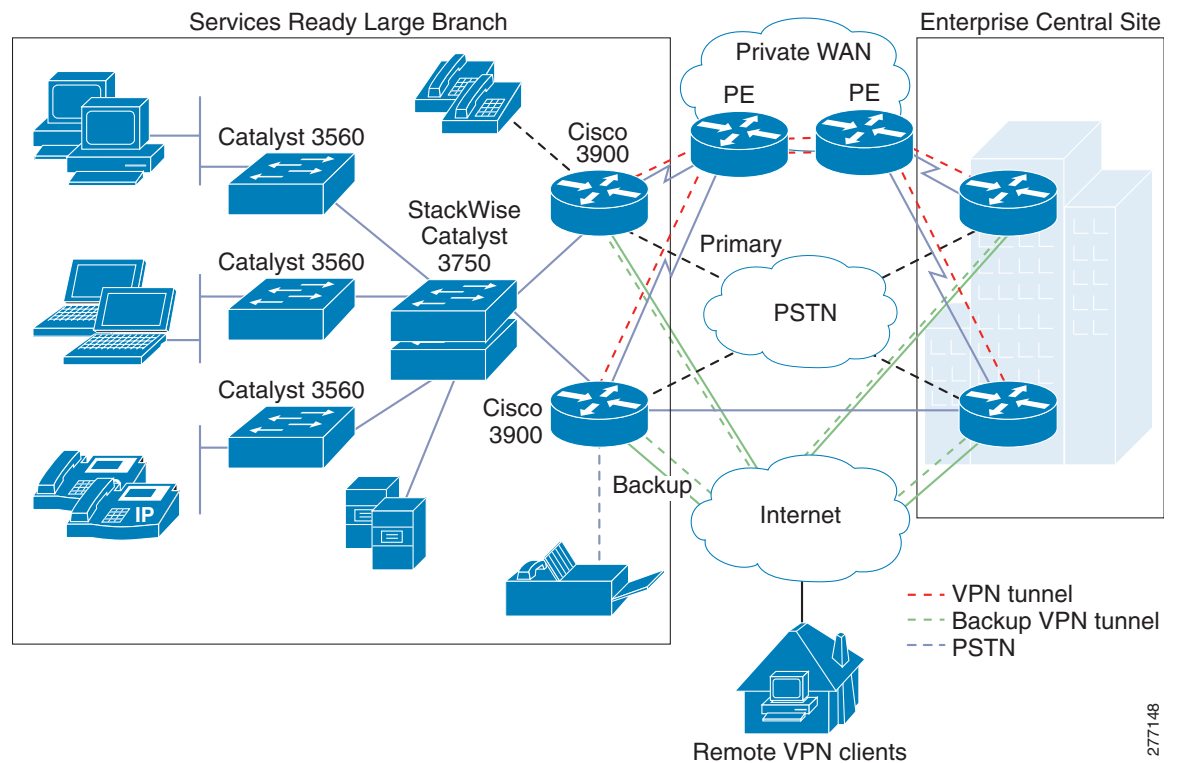
For more information about AAA, visit:

http://www.cisco.com/en/US/docs/ios/12_2/security/configuration/guide/scfaaa.html

Secure Connectivity

Secure connectivity protects against information theft or alteration of end-user data on public shared transport mediums. A Virtual Private Network (VPN) provides the means for securely and privately transmitting data over such a medium. There are two types of VPNs: provider-provisioned and enterprise-provisioned. The Frame Relay, Layer 3 VPN (L3VPN), and Layer 2 VPN (L2VPN) services described in the “[WAN Services](#)” section on [page 9](#) are examples of provider-provisioned VPNs. This section focuses on WAN-based VPN technologies in the context of a branch office, as shown in [Figure 39](#).

Figure 39 The Services Ready Large Branch Network Private WAN Deployment



IP-based WAN VPNs routed over the Internet have in recent years become an attractive alternative to traditional Layer 2 WAN deployments. IP VPNs offer low cost, secure, flexible, and scalable site-to-site connectivity. There are a number of WAN VPN options, and selecting the appropriate one involves many considerations. For a branch office the most important of these considerations are:

- WAN topology: Support for full-mesh or partial-mesh WAN designs.
- Scalability: Number of branch offices in the network and plans for future expansion.
- Availability: Local availability of WAN services that can support VPN deployments.
- Multicast: Requirement to support multicast traffic.
- Security: Type of encryption, key exchange, and authentication required, if any.
- Multiprotocol: Support for non-IP protocols.
- Quality of Service: End-to-end QoS requirements.
- Dynamic routing: Required support for dynamic routing protocols.
- High availability: Degree of resiliency required of a VPN.

To provide traffic separation on a public network, VPN uses a tunneling mechanism such as generic routing encapsulation (GRE), IPsec, Point-to-Point Tunneling Protocol (PPTP), or Layer 2 Tunneling Protocol version 3 (L2TPv3). Direct IPsec and GRE are the most typically deployed tunneling protocols for branch office VPNs. A tunneling protocol combined (or supported natively) with authentication and encryption mechanism, forms the basis of enterprise-provisioned VPNs. Table 12 provides an overview of the most commonly used IP-based WAN VPNs in a branch office. SSL-based VPNs are typically used for traffic that traverses the Internet. In the Services Ready Large Branch Network, SSL VPN is used to connect home users to the branch network.

Table 12 *Typical VPNs Provisioned in a Large Branch Office*

VPN Type	Advantages	Disadvantages	Appropriate for Branch
IPSec with direct encapsulation	<ul style="list-style-type: none"> • Multivendor interoperability 	<ul style="list-style-type: none"> • Limited support for mesh topology • No dynamic routing • No multicast • IP only • No QoS 	When interoperability with non-Cisco products is required
IPsec with VTI ¹ encapsulation	<ul style="list-style-type: none"> • QoS • Multicast • Dynamic routing • Lower overhead than GRE • Ease of use 	<ul style="list-style-type: none"> • Limited interoperability • IP only 	Small number of sites.
IPSec with GRE encapsulation	<ul style="list-style-type: none"> • Non-IP protocols • Multicast • QoS • Dynamic routing 	<ul style="list-style-type: none"> • Limited support for mesh topology • Overlay routing 	When non-IP protocols are required.
Easy VPN	<ul style="list-style-type: none"> • Simple configuration 	<ul style="list-style-type: none"> • No mesh topology • No dynamic routing • No multicast • IP-only 	Ease of management and simplicity of configuration are a priority.
DMVPN ²	<ul style="list-style-type: none"> • Multicast • Simpler configuration than IPsec+GRE • Small scale on-demand meshing • Easier to scale 	<ul style="list-style-type: none"> • Limited support for meshed topology • IP-only • Overlay routing • No spoke-to-spoke QoS 	<ul style="list-style-type: none"> • Internet-based primary WAN links. • Backup WAN link.

Table 12 *Typical VPNs Provisioned in a Large Branch Office (continued)*

VPN Type	Advantages	Disadvantages	Appropriate for Branch
GETVPN	<ul style="list-style-type: none"> • Tunnel-less VPN • Full-mesh connectivity • Routing • Efficient multicast • Advanced QoS • Scalable 	<ul style="list-style-type: none"> • Public WAN deployments • IP only 	<ul style="list-style-type: none"> • Appropriate for most branch offices. • MPLS/IP WANs. • Traditional Layer 2 WANs that need added security.
SSLVPN	<ul style="list-style-type: none"> • Clientless solution • Ease of use 	<ul style="list-style-type: none"> • Limited support for application-level protocols • Lower performance than IPsec alternatives 	<ul style="list-style-type: none"> • Remote users connecting to the branch.

1. VTI = Virtual Tunnel Interface.

2. DMVPN = Dynamic Multipoint Virtual Private Network.

In addition to these general considerations, a VPN solution must meet the business criteria outlined in the “[Large Branch Design Considerations](#)” section on page 4. Those requirements specify support for multicast and dynamic routing protocols. Because IPsec with direct encapsulation, IPsec with VTI, and Easy VPN do not support multicast and dynamic routing, they were excluded from branch office considerations. Moreover, IPsec with GRE encapsulation is a less general case of Dynamic Multipoint Virtual Private Network (DMVPN). Therefore, the only VPN solutions evaluated for the Services Ready Large Branch Network are DMVPN, Group Encrypted Transport Virtual Private Network (GETVPN) and SSL VPN.

GETVPN is appropriate for the primary WAN link, and DMVPN is appropriate for the Internet backup link for all WAN deployment scenarios described in the “[WAN Services](#)” section on page 9. However, existing hub-and-spoke WAN designs may already have DMVPN deployed. Therefore, DMVPN was validated on the primary link for leased line, Frame Relay, and VPWS WAN services. It should be noted that leased-line, Frame Relay, and Virtual Private Wire Service (VPWS) offer a degree of data privacy by providing traffic isolation. However, it is common to add a VPN to improve overall security and to enable enterprises to meet regulatory requirements such as Health Insurance Portability and Accountability Act (HIPAA), Sarbanes-Oxley Act, and Payment Card Industry (PCI) security standards. In summary, the following VPN deployment scenarios were tested for the Services Ready Large Branch Network:

- GETVPN on the primary link, DMVPN on the backup link, and SSL VPN for remote user access
- DMVPN on the primary link, DMVPN on the backup link, and SSL VPN for remote user access

Each VPN technology is described in more detail later in this section.

The foundation of a secure VPN is based on three independent security measures: data confidentiality, data integrity, and authentication. Each VPN solution listed in [Table 12](#) uses a different combination of technologies to provide these security measures. The following technologies are used in the Services Ready Large Branch Network:

- **Data Confidentiality:** Protects data from unauthorized interception. There are two general mechanisms for providing confidentiality:
 - **Encryption:** Reorders bits of the original message, making it incomprehensible to people not authorized to view the information. There are numerous encryption algorithms of various strengths. The following were used in the Services Ready Large Branch Network:
 - Triple Data Encryption Standard (3DES): Symmetric encryption mechanism that uses three different keys to encrypt a message. 3DES was used with both DMVPN and GETVPN.
 - Advanced Encryption Standard (AES)-256: Symmetric encryption mechanism that uses 256-bit key for encryption. AES-256 was used with both DMVPN and GETVPN.
 - **Tunneling:** Encapsulates original packet in a new packet and sends the composite packet over the network. The following mechanisms are used to provide tunneling:
 - Generic Routing Encapsulation (GRE): Encapsulates an original IP packet in a new IP packet whose source and destination become the two virtual endpoints of the GRE tunnel. The traffic in a GRE tunnel is not encrypted. However, GRE offers several advantages such as ability to carry both IP and non-IP traffic and the ability to support multicast. Therefore, GRE is typically placed inside an IPsec tunnel for greater security. This is the mechanism used by DMVPN.
 - IP Security (IPsec): IPsec is a framework for various security features. There are two main protocols within IPsec: tunnel mode protocol (also known as Authentication Header [AH]), and transport mode protocol (also known as Encapsulating Security Payload [ESP]). HA provides unencrypted tunneling and therefore was not used in the Services Ready Large Branch. ESP tunneling provides both encryption and authentication. In addition, ESP encrypts the original IP header. Standalone ESP is the mechanism used by GETVPN.
- **Data Integrity:** Guarantees that no tampering or alteration of the data occurs while it travels between the source and destination. The following algorithms are used for both DMVPN and GETVPN:
 - **Message Digest 5 (MD5):** A 128-bit hash algorithm. A hashing key is produced on the original message, appended to the end, and then encrypted. The recipient recomputes the hash to detect any alterations.
 - **Secure Hash Algorithm 1 (SHA-1):** A 160-bit hash algorithm. SHA-1 works on the same principle as MD5.
- **Authentication:** Verifies the identity of both endpoints that are communicating. VPN can use a variety of methods to perform authentication, such as login and password, smart cards, or biometrics. Most typically, digital certificates are used. The services-ready method used the following VPN authentication method:
 - **Preshared Key (PSK):** A secret key that is shared between the endpoints using a secure channel. A PSK is entered into each peer manually, and is used to authenticate the peer. In the Services Ready Large Branch Network, the secure channel for key exchange is provided by the following mechanism:
 - Diffie-Hellman Group 2 (DH2): 3DES and MD5 encryption and hashing algorithm with 1024-bit key

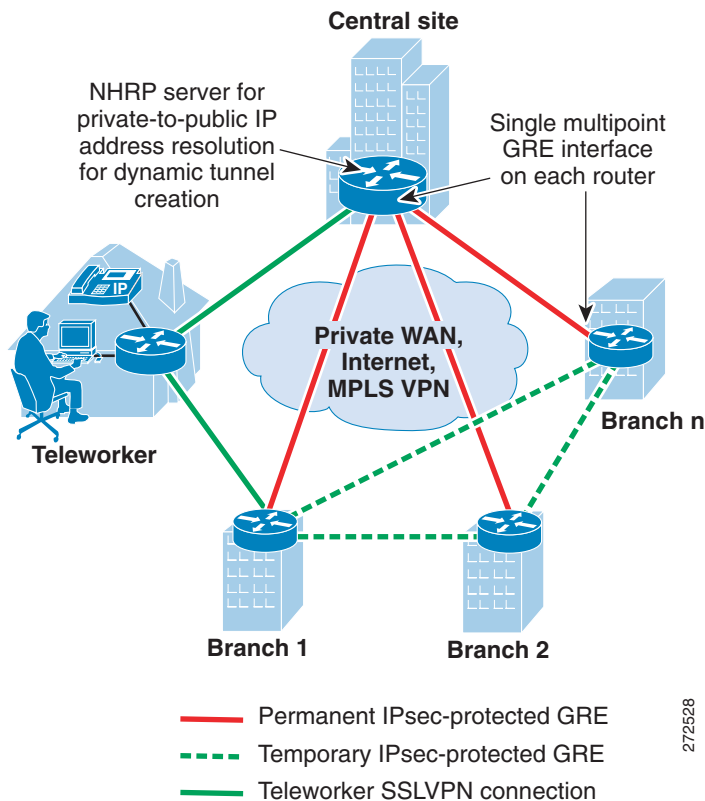
A secure communication channel between two endpoints is also referred to as a *security association* (SA). It is a security best practice to provide a lifetime limit for the SA. Typically, the lifetime is short enough to prevent attackers from gathering enough data to break the encryption ciphers. The lifetime

data volume thus depends on effective bandwidth and the encryption algorithm. It is also important to frequently change encryption keys when using the preshared key infrastructure. For the Services Ready Large Branch Network, both lifetimes are provided in Table 13.

In addition to security measures, VPNs differ in the way they manage keys, provide point-to-point or multipoint communication, and allow for dynamic creation of VPN tunnels. The three VPNs used in the Services Ready Large Branch Network offer the following functions:

- DMVPN is IPsec- and GRE-based VPN. It enables dynamic spoke-to-spoke tunnel creation in a traditional hub-and-spoke WAN design. DMVPN leverages multipoint GRE (mGRE) to establish multiple tunnel endpoints and to create an overlay non-broadcast multi-access (NBMA) network. While a traditional hub and spoke GRE configuration would require a separate tunnel between endpoints, mGRE allows multiple endpoints to have a single tunnel interface in the same subnet. Next Hop Resolution Protocol (NHRP) is used to provide tunnel-to-physical address lookup, facilitating dynamic configuration of GRE tunnels between endpoints. NHRP operates in a client/server configuration. NHRP Server typically runs on the hub, and each spoke router (NHRP Client) registers its tunnel-to-physical address mapping with the server. When a spoke wants to communicate on the NBMA mGRE subnet, it first sends a request to the NHRP Server to map a tunnel endpoint to a physical address. When the physical address is known, a GRE tunnel is established, and a regular routing process determines the path to the endpoint. Figure 40 shows DMVPN hub-and-spoke and spoke-to-spoke architecture.

Figure 40 DMVPN mGRE architecture

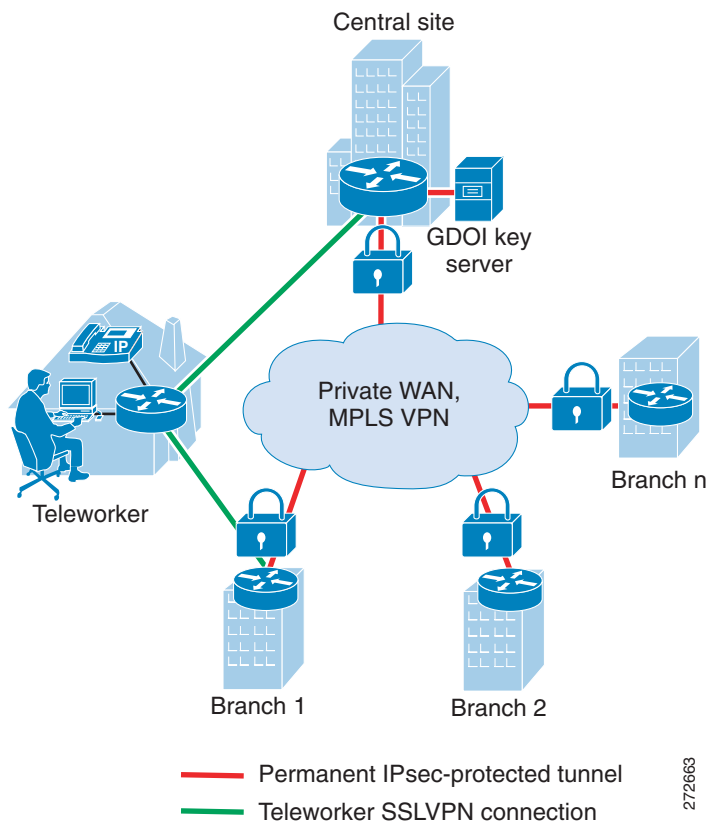


To learn more about DMVPN visit:

http://www.cisco.com/en/US/docs/ios/12_2t/12_2t13/feature/guide/ftgreips.html

- Group Encrypted Transport VPN (GETVPN) combines IPsec and Group Domain of Interpretation (GDOI) key server to encrypt traffic on a private WAN. Traditional VPN gateways directly authenticate each other and set up IPsec sessions that are private to the pair. This approach does not scale well when the network provides any-to-any connectivity or has large number of VPN gateways. GDOI server facilitates management and distribution of digital certificates or pre-shared cryptography keys. It authenticates group members and distributes keys and policies. GETVPN is a tunnelless VPN and therefore should be used in private WANs such as MPLS or traditional Layer 2 WANs. GETVPN can be used in conjunction with DMVPN or IPsec/GRE to simplify key management for a public WAN VPN. GETVPN uses IPsec ESP to provide confidentiality, integrity, and replay protection for packets flowing between VPN gateways. Figure 41 shows any-to-any GETVPN connectivity.

Figure 41 Any-to-Any GETVPN connectivity



To learn more about GETVPN, visit:

http://www.cisco.com/en/US/docs/ios/12_4t/12_4t11/htgetvpn.html

- Secure Socket Layer Virtual Private Network (SSL VPN): Leverages Secure Socket Layer (SSL) and its successor Transport Layer Security (TLS) to provide remote-access VPN capability, using the SSL/TLS function that is already built into a modern web browser. SSL VPN allows users from any Internet-enabled location to launch a web browser to establish remote-access VPN connections. Encryption is a component of the SSL/TLS framework; AAA is used to authenticate the remote users.

To learn more about SSL VPN, visit:

<http://www.cisco.com/en/US/docs/security/asa/asa72/configuration/guide/webvpn.html>

Table 13 summarizes all the security mechanisms used for GETVPN and DMVPN in the Services Ready Large Branch Network.

Table 13 Security Mechanisms for DMVPN and GETVPN

Mechanism	DMVPN	GETVPN
Peer authentication	Preshared key	Preshared key
Encryption	3DES, AES-256	3DES, AES-256
Integrity algorithm	SHA-1, MD5	SHA-1, MD5
Key exchange	DH2	DH2
Tunneling	GRE inside IPsec ESP	IPsec ESP
SA lifetime ¹	86400 seconds	86400 seconds
	28800 seconds	28800 seconds
	3600 seconds	3600 seconds
Rekey lifetime	300 seconds	300 seconds

1. The SA lifetime value depends on the aggregate amount of data that passes through VPN gateways. This will vary from enterprise to enterprise. To determine appropriate SA value follow instructions provided at:
http://www.cisco.com/en/US/prod/collateral/iOSSwrel/ps6537/ps6586/ps6635/ps7180/white_paper_c11-471053.html

Encryption is a CPU-intensive process. The Services Ready Large Branch Network uses the VPN and SSL advanced integration module to support the required 100-240 users in the branch. The Cisco VPN and SSL service module provides up to 40 percent better performance for IPsec VPN over the router built-in IPsec encryption, and up to twice the performance for SSL VPN encryption. The AIM2 supports both SSL encryption and VPN IPsec encryption with either Data Encryption Standard (DES) or Advanced Encryption Standard (AES) in its hardware.

Threat Protection, Detection, and Mitigation

Threat protection, detection, and mitigation are security mechanisms for protecting the branch network from security policy violations and from malicious attacks on the network infrastructure. In the context of this document, threats are security breaches in which the primary goal is information theft or tampering. Reconnaissance and unauthorized access fall into this category. Attacks are intentional or unintentional activity to disrupt the operation of the network. Denial of service and malicious code fall into this category. Prevention proactively blocks both threats and attacks. Detection identifies threats and attacks that are currently in progress. Mitigation stops current threats and attacks, and prevents recurrence. Attackers can be either individuals external to the enterprise or someone within the organization. Internal attackers are much more difficult to spot and block because they have more information and more options for launching an attack. In addition, both types of attackers can use low-tech methods, such as social engineering, to gain unauthorized access. It is therefore critical to have a solid security policy for the branch office and to educate all users to follow the established security measures. Security policy was described in the “Security Services” section on page 53.

Services Ready Large Branch Network uses the following security mechanisms to prevent external attacks:

- **Zone-based Policy Firewall (ZPF):** Prevents external threats and attacks. Firewalls provide stateful security and application inspection for each protocol entering or leaving a branch network. A stateful inspection firewall uses a combination of access control with application inspection to ensure that only approved responses get through the firewall. ZPF assigns the router interfaces to

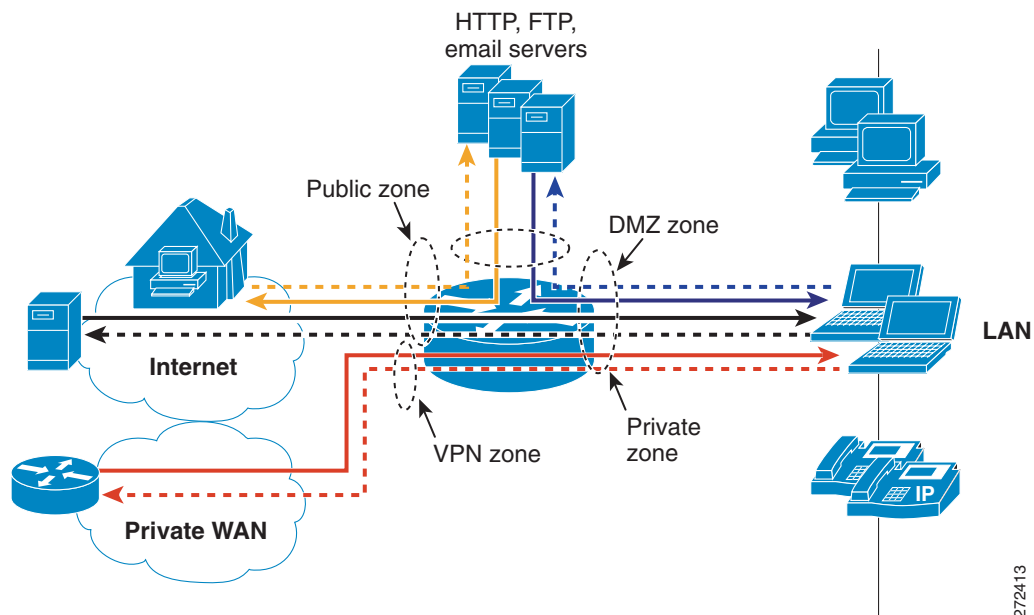
various zones and applies inspection policies to traffic flowing between the various zones. Inter-zone policies offer considerable flexibility and granularity, enabling different inspection policies for different host groups connected to the same router interface. An interface can be easily added or removed from a zone. Four security zones were defined for the Services Ready Large Branch Network: demilitarized zone (DMZ), Public zone, VPN zone, and Private zone as shown in Figure 42. The following traffic is inspected and permitted to pass:

- From Private zone to Private zone, all traffic passes without any inspection.
- From Private zone to Public zone HTTP, FTP, DNS, HTTPS, SSH, and ICMP traffic is inspected and allowed, but the rest of the traffic is blocked.
- From Public zone to Private zone, no traffic is allowed.
- From Public zone to DMZ zone, only HTTP, HTTPS, and DNS are allowed.
- From Private zone to VPN zone, all traffic passes with inspection.
- From VPN zone to Private zone, all traffic passes with inspection.

To learn more about Zone-based Policy Firewall, visit:

http://www.cisco.com/en/US/products/sw/secursw/ps1018/products_tech_note09186a00808bc994.shtml

Figure 42 Security Zones



- Unicast Reverse Path Forwarding (uRPF): Leverages routing tables to validate source addresses that are expected to be seen on an interface. Packets are forwarded only if they match the router's best path to the source. This ensures that packets coming into an interface are from valid hosts that have a corresponding entry in the routing table. Packets with source addresses that cannot be reached via the input interface are dropped.

To learn more about uRPF, visit:

<http://www.cisco.com/web/about/security/intelligence/unicast-rpf.html>

The following security mechanisms are used to prevent internal threats and to control access to network resources in the Services Ready Large Branch Network:

272413

- Standard and extended access control lists (ACLs): Control whether a router permits or denies packets to pass, based on criteria in the packet header. Standard ACLs filter packets based on source IP address only. Extended ACLs filter packets on source and destination IP addresses, port numbers, and protocol type. ACLs are used extensively within the Services Ready Large Branch Network to permit or deny access between the different firewall zones.
- Layer 2 security: Prevents various attacks or access violations that could be launched through the branch switches
 - 802.1x: Client-server-based access control and authentication protocol that restricts unauthorized devices from connecting to a LAN through publicly accessible ports. The authentication is provided by a RADIUS server.
 - Port Security: Switch port limits the number of MAC addresses that are able to connect to a switch, and ensures that only approved MAC addresses are able to access the switch. It prevents MAC address flooding and ensures that only approved users can log on to the network.
 - DHCP Snooping: Switch port forwards DHCP requests only from trusted access ports and drops all other types of DHCP traffic. DHCP snooping eliminates rogue devices from behaving as the DHCP server.
 - Dynamic Address Resolution Protocol (ARP) Inspection (DAI): Maintains a binding table containing IP and MAC address associations dynamically populated using DHCP snooping. This feature ensures the integrity of user and default gateway information so that traffic cannot be captured. This feature mitigates ARP spoofing and ARP poisoning attacks.
 - IP Source Guard: When a client receives a valid IP address from the DHCP server, or when a static IP source binding is configured by the user, a per-port and VLAN access control list (PVACL) is installed on the port. This process restricts the client IP traffic to the source IP addresses configured in the binding; any IP traffic with a source IP address except that in the IP source binding is filtered out. This filtering limits a host's ability to attack the network by claiming a neighbor host's IP address.
 - Bridge Protocol Data Unit (BPDU) Guard: Prevents loops if another switch is attached to a PortFast port. When BPDU Guard is enabled on an interface, the interface is shut down if a BPDU is received on the interface. To assume the root bridge function, a device would be attached to the port and would run STP with a lower bridge priority than that of the current root bridge. If another device assumes the root bridge function in this way, it renders the network suboptimal. This is a simple form of a denial-of-service (DoS) attack on the network.

To detect and mitigate various external and internal attacks, the Services Ready Large Branch Network uses the following mechanisms:

- Cisco Intrusion Prevention System (IPS): Monitors packets and sessions as they flow through the branch, and scans each packet to match any of the IPS signatures. When IPS detects suspicious activity, it can shunt the offending packets before network security can be compromised. When an IPS signature is matched, one or more of the following actions are taken:
 - An alarm is sent to a syslog server or a centralized management interface.
 - The packet is dropped.
 - The connection is reset.

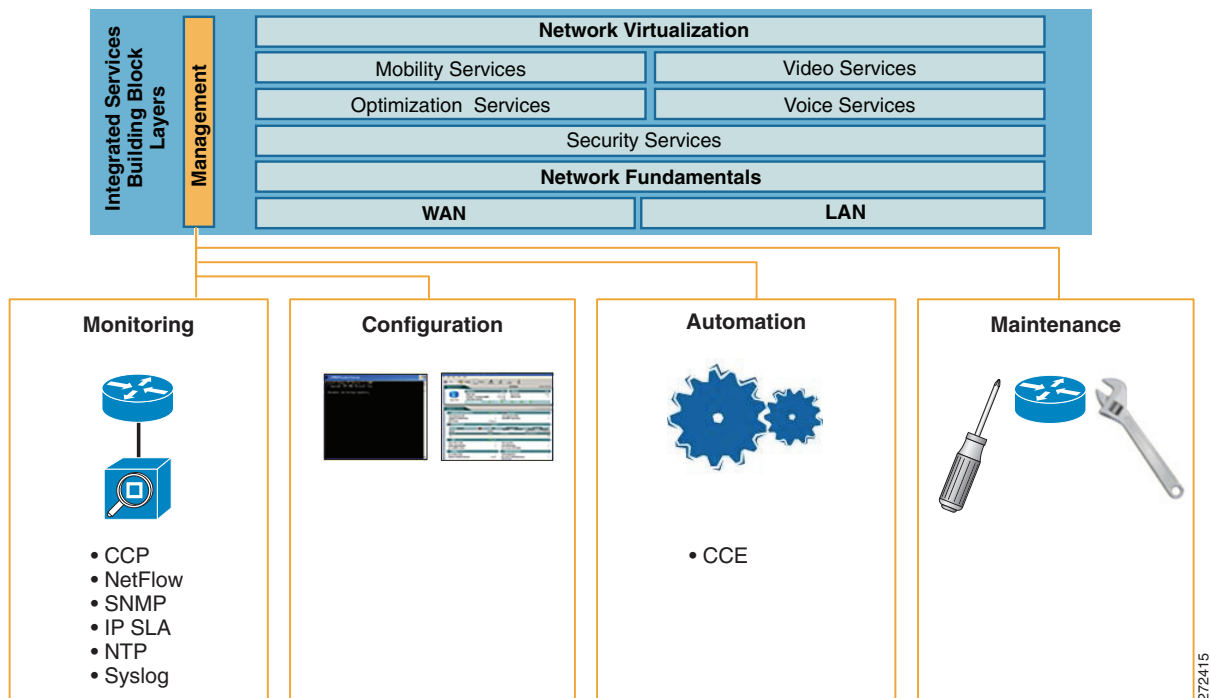
The Services Ready Large Branch Network is configured to take different actions depending on which attack signature is matched. An advanced signature set was used to identify various attacks. IPS is configured on all outside and inside interfaces. Traffic, regardless of whether it is a WAN link to the public or an internal LAN link, is inspected. See the “[System Testing](#)” chapter or page for the various attacks that were validated for the Services Ready Large Branch Network.

- Network Based Application Recognition (NBAR): Recognizes certain type of attacks and drops packets involved in a denial-of-service attacks such as SQL Slammer, and worms such as CODE RED and NIMDA.

Management Services

Management services include activities related to configuration, monitoring, automation, and maintenance of a branch office network, as shown in [Figure 43](#).

Figure 43 Management Services for a Branch Network



Cisco offers numerous tools for performing network management in the branch office. At this time, only a subset of those tools has been validated for the Services Ready Large Branch Network. The primary focus was on monitoring the branch router. Future updates to this guide will address configuration management, automation, and maintenance for all the branch network devices.

Monitoring services for the Services Ready Large Branch Network are described in the following sections:

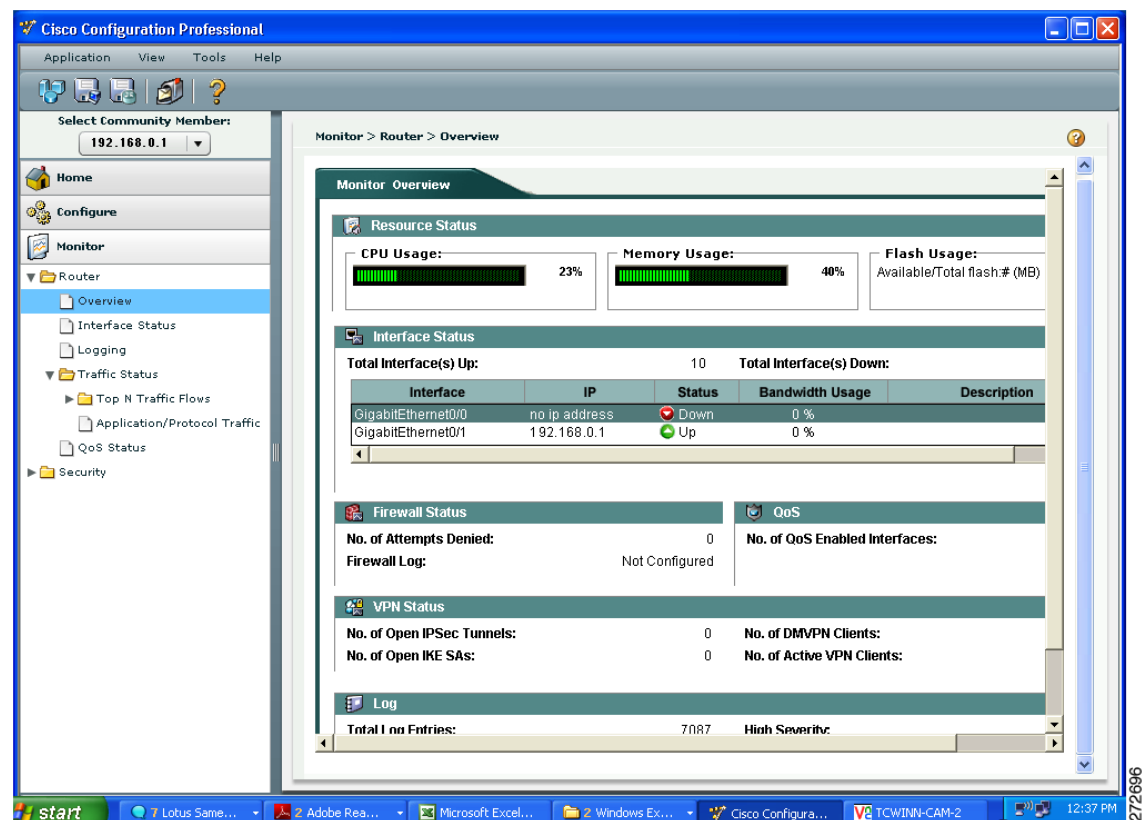
- [Cisco Configuration Professional, page 69](#)
- [Simple Network Management Protocol, page 70](#)
- [Syslog, page 70](#)
- [NetFlow, page 70](#)
- [Network Based Application Recognition, page 71](#)
- [IP Service Level Agreement, page 71](#)
- [Network Time Protocol, page 71](#)
- [Cisco Configuration Engine, page 71](#)

Configuration management in the Services Ready Large Branch Network was done primarily through the command line. However, several services have a web-based graphical interface that was used to configure those services. Configuration of all networking devices is extensively documented in the “System Implementation” chapter.

Cisco Configuration Professional

Cisco Configuration Professional, shown in Figure 44, is a web-based device management tool embedded within the Cisco IOS software. Cisco Configuration Professional simplifies router, security, Unified Communications, wireless WAN, and basic LAN configuration through intelligent wizards. It enables faster configuration and monitoring of the branch router without requiring knowledge of the Cisco IOS command-line interface (CLI). In the Services Ready Large Branch Network, Cisco Configuration Professional was used for monitoring only.

Figure 44 Cisco Configuration Professional



In monitor mode, Cisco Configuration Professional provides an overview of router status and performance metrics such as the Cisco IOS release number, interface status (up or down), and CPU and memory usage. The monitor mode also allows users to view the number of network access attempts that were denied by Cisco IOS Firewall, and provides easy access to the firewall log. Additionally, VPN status, such as the number of active IPsec tunnels, can be monitored.

For more information about Cisco Configuration Professional, visit:

http://www.cisco.com/en/US/prod/collateral/routers/ps9422/data_sheet_c78_462210.html

Simple Network Management Protocol

Simple Network Management Protocol (SNMP) provides a standardized framework and a common language for the monitoring and management of devices in a network. In the Services Ready Large Branch Network, SNMP version 3 traps were enabled to log various events on the routers and switches.

To learn more about configuring SNMP visit:

http://www.cisco.com/en/US/docs/ios/12_2/configfun/configuration/guide/fcf014.html

Syslog

Syslog is a protocol for sending logging messages on a network. Various devices log status, events, alerts, and errors, using syslog components that forward the log messages to a syslog service. A syslog service simply accepts messages and stores them in files or prints them to a console. Syslog was used extensively in the Services Ready Large Branch Network for security accounting and for monitoring the status of various devices.

To learn more about Cisco IOS software syslog messages, visit:

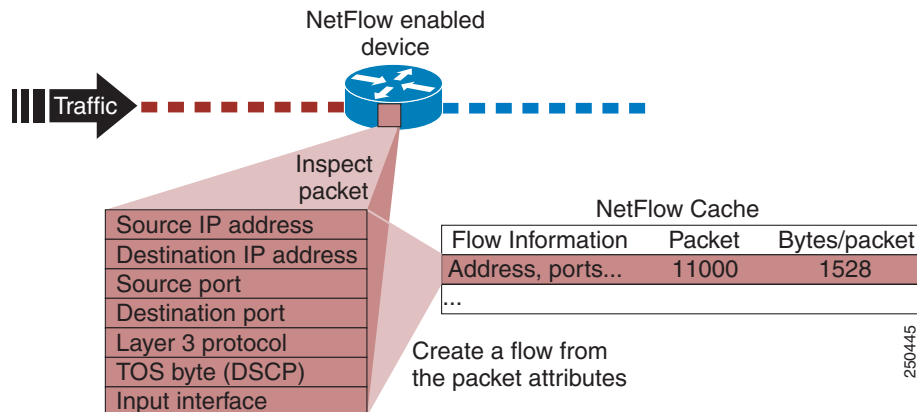
http://www.cisco.com/en/US/docs/ios/12_3/sem1/system/messages/123semv1.html

http://www.cisco.com/en/US/docs/ios/12_3/sem2/system/messages/123semv2.html

NetFlow

NetFlow version 9 technology is used to monitor and measure specific traffic flows and to provide an aggregate view of all network activity. With NetFlow, network administrators can view detailed time and application-specific usage of the network. This information is essential for network planning, security analysis, application optimization and delivery, and traffic engineering. A typical NetFlow record includes source and destination IP addresses, TCP/UDP port numbers, type of service (ToS), packet and byte counts, time stamps, input and output interfaces as shown in Figure 45, TCP flags and routing information. NetFlow data is exported from the router to a centrally located NetFlow collection server for analysis. This typically consumes 1 to 5 percent of bandwidth. The Services Ready Large Branch Network used Netflow version 9.

Figure 45 Data Captured by NetFlow



For more information about NetFlow and third-party NetFlow data analysis tools, visit:

http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6555/ps6601/prod_white_paper0900aec80406232.html

Network Based Application Recognition

Network Based Application Recognition (NBAR) is a Cisco IOS classification engine that can recognize a wide variety of applications, including web-based applications and client-server applications that dynamically assign TCP or User Datagram Protocol (UDP) ports. After the application is recognized, the network can invoke specific services for the application. In the Services Ready Large Branch Network, NBAR was used to support QoS features described in “[Quality of Service](#)” section on page 46. NBAR identifies and stops command worms, such as SQL Slammer, NIMDA, and Arctic, from propagating through the network.

To learn more about NBR, visit:

http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6558/ps6612/ps6653/prod_qas09186a00800a3ded_ps6616_Products_Q_and_A_Item.html

IP Service Level Agreement

The IP service level agreement (IP SLA) feature of Cisco IOS software is used to verify service guarantees, to increase network reliability by validating network performance, and to proactively identify network issues. In the Services Ready Large Branch Network, IP SLAs were used to measure:

- End-to-end response time (delay) between the branch router and the central location router
- Packet delay variability (jitter) for traffic flowing between the branch and the central location

Both IP SLA metrics are critical to ensure high-quality voice services. To learn more about IP SLAs visit:

http://www.cisco.com/en/US/technologies/tk648/tk362/tk920/technologies_white_paper0900aec8017f8c9_ps6602_Products_White_Paper.html

Network Time Protocol

Network Time Protocol (NTP) is used to synchronize clocks among network devices. This synchronization allows events to be correlated when system logs are created and when other time-specific events occur. All devices in the Services Ready Large Branch Network used NTP to synchronize their clocks. The NTP server was hosted at the central site.

To learn more about NTP, visit:

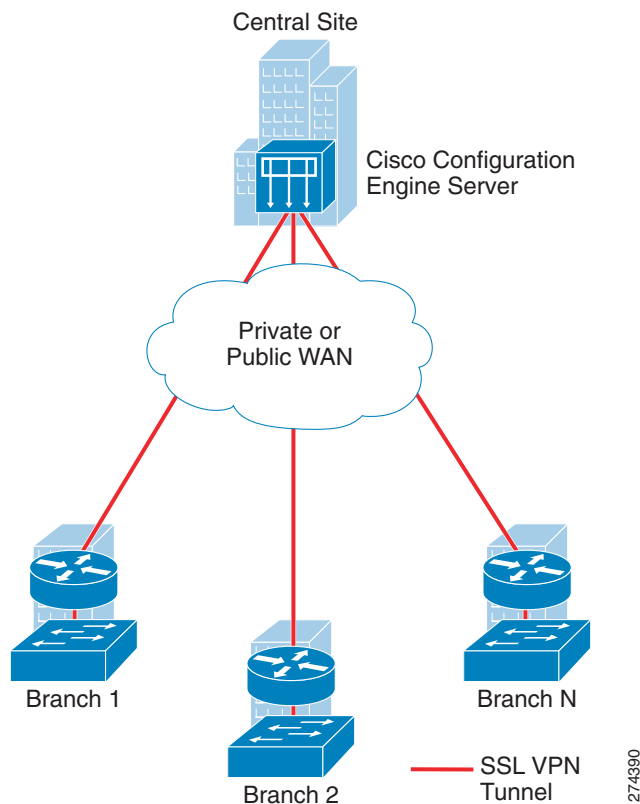
http://www.cisco.com/en/US/tech/tk869/tk769/technologies_white_paper09186a0080117070.shtml

Cisco Configuration Engine

The Cisco Configuration Engine (CCE) automates installation and provisioning of Cisco devices during their initial deployment and in subsequent reconfigurations. It securely distributes software images and device configuration files to one or multiple devices on a local LAN or over the WAN. In the Services Ready Large Branch Network, a centrally hosted CCE server was used to distribute Cisco IOS images and device configuration files to the branch routers and switches. During the initial deployment, the primary benefit of the CCE is consistent Cisco IOS image and configuration distribution across multiple

branch networks. Once the network becomes operational, the CCE provides a simple, secure, and fast way to reconfigure all branch devices without the assistance of an on-site technician. Moreover, the ability to configure multiple devices from a single toolkit is less error-prone than individual configuration of each device. Figure 46 shows the deployment of CCE in the Services Ready Large Branch Network.

Figure 46 Deployment of CCE in the Services Ready Large Branch Network.



Each device that is to be provisioned with the CCE is assigned a unique Cisco Network Services (CNS) identifier and pre-loaded with a bootstrap configuration. Prior to powering up of the device, the CNS ID must be registered with the centrally hosted CCE server. After the device is powered up, it contacts the CCE server and requests to be provisioned. The CCE server uploads and activates the appropriate Cisco IOS image and configures the device for operation. This provisioning can be further simplified by configuring a centrally hosted DHCP server to provide the bootstrap configuration through the DHCP option 150.

- This document is accompanied by several CCE toolkits that can be used to deploy the Services Ready Large Branch Network. Because this deployment blueprint covers several hardware components and networking services that are functional alternatives of one another, the following six sample CCE toolkits, covering different combinations of technologies, are provided: Gigabit Ethernet WAN interface, active primary and standby backup WAN links, OSPF routing, DMVPN over primary and backup WAN links, and Cisco CME with SCCP configured IP phones and H.323 trunking to central site.
- A four T1 WAN interface bundle with MLPPP encapsulation, active primary and standby backup WAN links, EIGRP routing, GETVPN over primary and DMVPN over backup WAN links, and Cisco CME with SIP configured IP phones and SIP trunking to central site.

- A four T1 WAN interface bundle with MLFR encapsulation, simultaneously active primary and backup WAN links, EIGRP routing, DMVPN over primary and backup WAN links, and Cisco SRST with SCCP configured IP phones and H.323 trunking to central site.
- T3 WAN interface with Frame Relay encapsulation, simultaneously active primary and backup WAN links, OSPF routing, GETVPN over primary and DMVPN over backup WAN links, and Cisco SRST with SIP configured IP phones and SIP trunking to central site.
- Two 24-port distribution switches configured in Cisco StackWise topology with 5 two-link EtherChannel trunks and DMZ VLAN on the access ports.
- A 48-port access switch with a two-link EtherChannel trunk and Data and Voice VLANs on the access ports

The sample CCE toolkits are intended to provide:

- Full and validated router and switch configurations for the Services Ready Large Branch Network
- Alternative configurations of the various technologies of the Services Ready Large Branch Network
- Starting points for customization of the Services Ready Large Branch Network configuration

The Services Ready Large Branch Network used CCE version 3.0 to deploy the branch router and switch Cisco IOS images and configurations. To learn more about CCE, visit:

http://www.cisco.com/en/US/prod/collateral/netmgtsw/ps6504/ps4617/data_sheet_c78-502925.html

Voice Services

The availability of higher bandwidth and more reliable QoS guarantees enable enterprises to combine voice and data on the same converged IP network. IP-based voice services offer new, business-relevant functionality and are more cost effective than traditional telephone services.

Today, branch offices have two fundamental options for converged telephony:

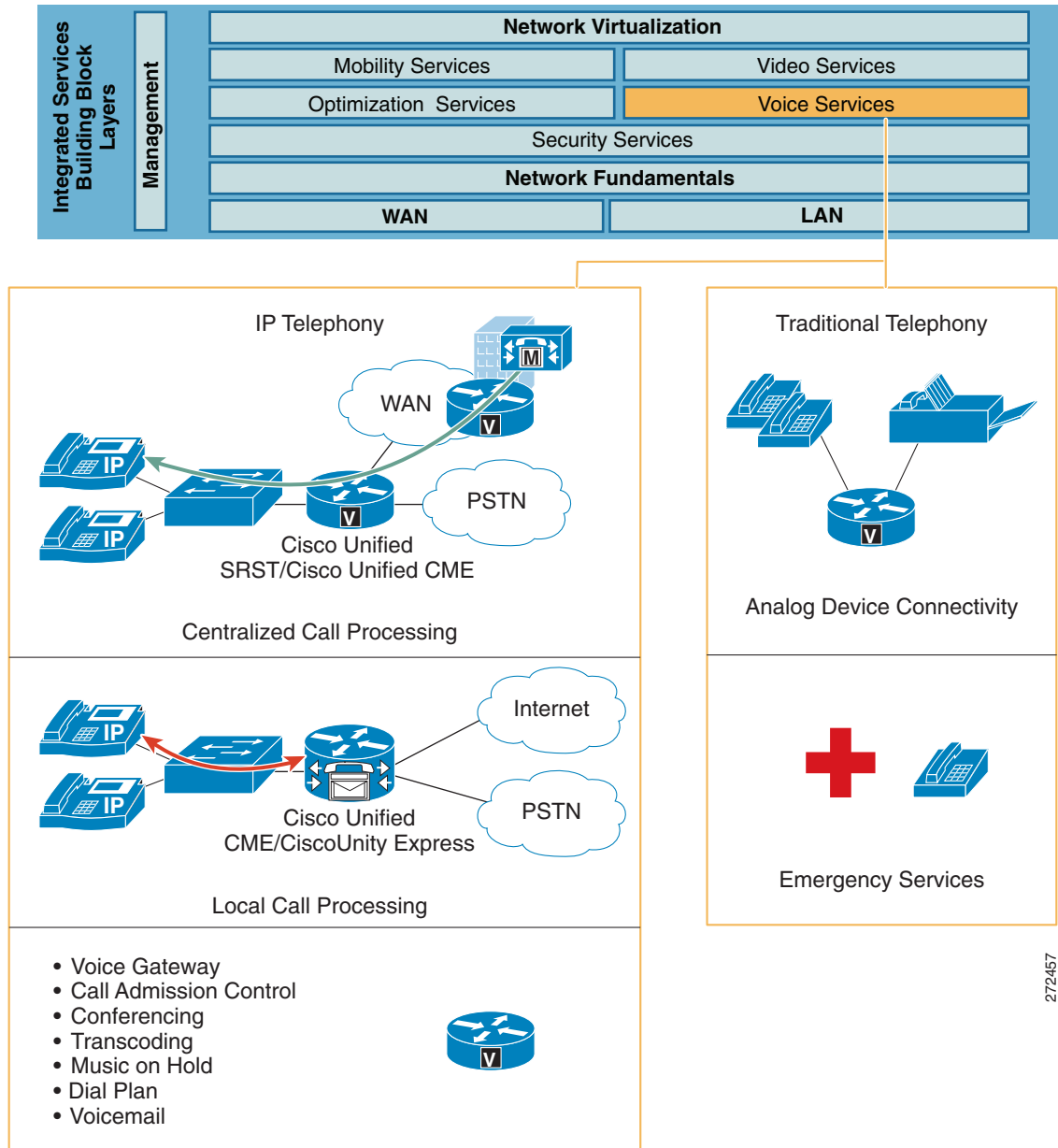
- Voice over IP (VoIP): Traditional telephony devices such as analog phones, faxes, PBXs, and public switched telephone network (PSTN) attached to an IP network. A voice-enabled router digitizes and packetizes the voice and signaling traffic from the traditional devices and transports the traffic over the IP network.
- IP Telephony: IP-based telephony devices connected to an IP network that natively digitize and packetize voice and signaling traffic. A voice-enabled router transports the traffic over the IP network.

IP telephony was the primary focus of the Services Ready Large Branch Network. However, a small number of analog phones and fax machines were connected to the network and used for VoIP as well as traditional PSTN connectivity.

Voice services for a large branch office network are described in the following sections and shown in [Figure 47](#):

- [Voice Quality Considerations, page 74](#)
- [WAN Capacity Considerations, page 76](#)
- [IP Telephony, page 79](#)
- [Traditional Telephony, page 91](#)

Figure 47 Voice Services



272457

Voice Quality Considerations

The following fundamental packet propagation criteria must be satisfied in order to provide high-quality voice service:

- **Delay:** Delay is defined as the finite amount of time necessary for a packet to reach the receiving endpoint after being transmitted from the sending endpoint. For voice, this delay is defined as the amount of time it takes for sound to leave the mouth of the speaker and be heard in the ear of the listener. The ITU G.114 and Cisco recommend a maximum one-way, mouth-to-ear delay of 150 ms for high-quality voice.

- Delay Variability (jitter): Jitter is the difference in the end-to-end delay between packets. Cisco recommends a maximum jitter of less than 30 ms for high-quality voice.
- Packet loss: Packet loss is a relative ratio of packets successfully sent and received to the total number of packets transmitted. The amount of packet loss that can be tolerated is user-dependent; however, on average, packet loss should be kept to less than 1 percent to ensure high-quality voice service.

Table 14 summarizes packet propagation criteria that must be met to support high-quality voice.

Table 14 Not-to-Exceed Packet Propagation Criteria for High-Quality Voice Service

Propagation Factor	Not-to-exceed Value
Delay (Latency)	150 ms
Delay variability (Jitter)	30 ms
Packet Loss (Packet Drops)	1 percent

For more information about controlling voice quality, visit:

http://www.cisco.com/en/US/netsol/ns341/ns396/ns172/ns103/networking_solutions_white_paper09186a00801b1c5a.shtml

Another factor affecting voice quality is the codec used to digitize the voice signal. Cisco voice devices typically use the following two codecs:

- G.711: Provides encoding that does not perform any compression and requires 64 kb/s of bandwidth (not including overhead) for a single voice call. The mean opinion score (MOS), a metric used to measure voice quality, for G.711 is 4.1.
- G.729a: Provides encoding with compression and requires 8 kb/s of bandwidth (not including overhead) for a single voice call. Compression reduces the amount of required bandwidth, but affects the quality of the transmitted voice signal. However, the MOS score for G.729a is 3.9, which is a barely perceptible difference in comparison to G.711, and therefore the codec provides an acceptable tradeoff for the significant reduction in consumed bandwidth.

The selection of the appropriate codec depends on the desired level of voice quality, the amount of available bandwidth, and the number of concurrent voice calls that must be supported. In the Services Ready Large Branch Network, the G.729a codec is used for voice calls that will traverse the WAN links because it will provide bandwidth savings on these lower-speed links. The G.711 codec is used for LAN calls. To compensate for the quality factors described previously, it is critical that QoS be enabled in the branch network. The “Quality of Service” section on page 46 provides detailed information on QoS implemented in the Services Ready Large Branch Network. All voice traffic was given 18 percent of the available bandwidth and was assigned for low latency queuing (LLQ). Call signaling was combined with other mission-critical data and was assigned 20 percent of the available bandwidth.

Traffic shaping is required for multiple-access, nonbroadcast media such as Frame Relay, where the physical access speed varies between two endpoints and several branch sites are typically aggregated to a single router interface at the central site. Shaping at the branch router alleviates potential congestion when the central site oversubscribes bandwidth or when the branch WAN link allows bursting beyond the Frame Relay committed information rate (CIR). The Services Ready Large Branch Network used traffic shaping to limit the traffic sent out on the WAN interfaces to a rate lower than the line rate. The specific settings for traffic shaping vary from implementation to implementation and depend on the central site router provisioning and the Frame Relay configuration. IP SLAs described in the “Management Services” section on page 68 ensured that the desired delay and jitter were maintained on the WAN link.

WAN Capacity Considerations

Three types of calls must be considered when provisioning the branch office for voice: PSTN (traditional), LAN (private exchange), and WAN (toll-bypass) calls. PSTN calls are needed for external communication, LAN calls are for intraoffice communication, and WAN calls enable communication with the rest of the enterprise. Knowing the number of PSTN calls and WAN calls helps to determine the number of voice lines and WAN bandwidth needed for the branch office. Traditionally, basic oversubscription ratios or Erlang traffic models have been used to determine the number of voice lines required for PSTN and WAN calling. Basic oversubscription ratios are typically based on call records collected from other existing offices of similar size and function, and applied to the new office. They equate the number of users to the number of PSTN and WAN calls required for calling. The business criteria outlined in the [Large Branch Design Considerations, page 4](#) specified the following oversubscription ratios:

- 5:1 user-to-active call ratio
- 4:1 WAN-to-LAN call ratio
- 4:1 WAN-to-PSTN call ratio

[Table 15](#) lists the requirements of the number of active calls for three sample office sizes.

Table 15 **Active Calls for Typical 120-, 180-, and 240-User Branch Offices, Using Oversubscription Ratios**

Active Calls	120-User Branch	180-User Branch	240-User Branch
WAN	16	24	32
PSTN	4	6	8
LAN	4	6	8
Total calls	24	36	48

Alternatively, an Erlang traffic model can provide a more accurate method for determining the number of external voice lines (PSTN and WAN) required for a branch office. There are several variants of the Erlang model, depending on the intended telephone use in the branch office. The following example uses the Extended Erlang B to determine the number of voice lines required for the Services Ready Large Branch Network.

The Extended Erlang B traffic model takes into account the additional traffic load caused by blocked callers that immediately try to call again if their calls are blocked. The four variables involved are recall factor, busy hour traffic (BHT), blocking, and lines:

- Recall factor: Percentage of calls that immediately retry if their calls are blocked.
- Busy hour traffic (BHT): Number of hours (in Erlangs) of call traffic during the busiest hour of operation of a telephone system.
- Blocking: Failure rate of calls because of an insufficient number of available lines. For example, 0.03 means three calls blocked per 100 calls attempted.
- Lines: Total number of external lines needed.



Note

An *Erlang* is a unit of measurement of voice traffic. Strictly speaking, an Erlang represents the continuous use of one voice path or line. In practice, it is used to describe the total traffic volume in one hour.

If an average user calls for 12 minutes during the busy hour, external calls account for 10 minutes of those calls (or 10 min/60 min/hr = 0.17 Erlang), half of blocked calls immediately retry, blocked calls are no more than 3 percent of total calls, there is a 4:1 WAN-to-LAN call ratio, and there is a 4:1 WAN-to-PSTN call ratio, the Extended Erlang B calculator at <http://www.erlang.com/calculator/exeb/> suggests the total number of external lines for 120-, 180-, and 240-user branch office as shown in Table 16.

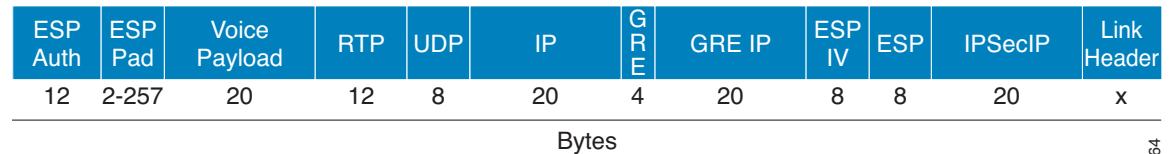
Table 16 Active Calls for Typical 120-, 180-, and 240-User Branch Offices, Using Extended Erlang B Traffic Model

Active Calls	120 User Branch	180 User Branch	240 User Branch
Busy Hour Traffic (Erlangs)	20	30	40
WAN	21	29	36
PSTN	7	9	12
LAN	7	9	12
Total calls	35	47	60

The critical assumption in the Extended Erlang B model is the amount of BHT per user (0.17 Erlang in the preceding example), which varies between enterprises, and even between branch offices within an enterprise. Therefore, Table 16 is provided only as an example. The Services Ready Large Branch Network used active call counts derived from the oversubscription ratios shown in Table 15.

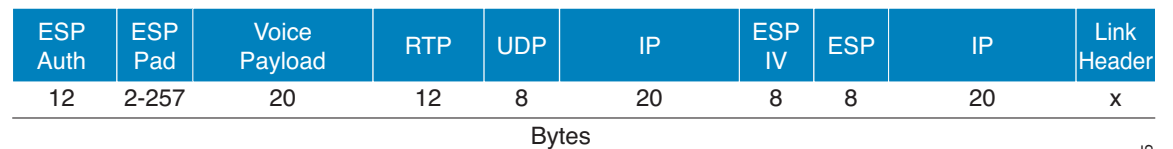
Real-time Transport Protocol (RTP) is the primary protocol for transporting real-time traffic such as voice or interactive video. The minimum amount of bandwidth required to place a given number of calls over the WAN can be derived from the number of RTP streams. The size of each RTP stream depends on the WAN type, the associated encapsulations (Frame Relay, PPP, MLPP, Ethernet, IPsec, GRE), and the voice sampling rate. Figure 48 shows packet size for a G.729a RTP packet with DMVPN encapsulation. Figure 49 shows the packet size for G.729a RTP packet with GETVPN encapsulations.

Figure 48 RTP Packet for G.729a Codec with DMVPN Encapsulation



272664

Figure 49 RTP Packet for G.729a Codes with GETVPN Encapsulation



272665

An RTP packet contains 40 bytes of RTP and UDP header information. Because most information in these headers is identical (for example, the same source/destination IP address/UDP port numbers and the same RTP payload type), compressed RTP (cRTP) can be used to eliminate redundant header information in each frame. Using cRTP reduces the 40-byte header to only 2 or 4 bytes, allowing more

calls to be placed over the same link speed. [Table 17](#) shows sample bandwidth requirements for RTP and cRTP streams with the various Services Ready Large Branch Network WAN encapsulations. The Cisco Voice Codec Bandwidth Calculator that was used to calculate the necessary bandwidth requirements is available at:

<http://tools.cisco.com/Support/VBC/do/CodecCalc1.do>,

Although cRTP reduces the amount of required bandwidth, it is a CPU intensive process that may impact the overall performance of the router. Therefore, cRTP is appropriate only when voice traffic represents more than 33 percent of the load on the link, when the link uses a low bit-rate codec (such as G.729), and when no other real-time application (such as video conferencing) is using the same link.

Table 17 Bandwidth Requirement for a Single Call with Various WAN Encapsulation Methods

	Frame Relay, PPP, MLPP		Ethernet	
	RTP (kbps)	cRTP (kb/s)	RTP (kb/s)	cRTP (kb/s)
DMVPN	56	40	60	N/A
GETVPN	46	30	50	N/A

The Services Ready Large Branch Network used cRTP to minimize bandwidth consumption only on the 4 T1 connections; other WAN connection types used RTP. The QoS model allocates 18 percent of bandwidth to voice traffic. [Table 18](#) shows the amount of bandwidth required for voice communication and the total bandwidth that is required to support branch offices of 120, 180, and 240 users with various WAN encapsulation methods. The total number of active voice calls is derived from the oversubscription ratios shown in [Table 15](#). In general, each call has two streams for audio traffic; one stream from caller to callee, and another stream in the reverse direction.

Table 18 Bandwidth Requirements for Voice Traffic and Total Bandwidth for a Services Ready Large Branch Network with 120, 180, and 240 User Counts

	Frame Relay, PPP, MLPPP				Ethernet	
	RTP Voice (Mbps)	RTP Total (Mb/s)	cRTP Voice (Mp/s)	cRTP Total (Mb/s)	RTP Voice (Mb/s)	RTP Total (Mb/s)
120-User Services Ready Large Network (16 simultaneous WAN calls)						
DMVPN	0.9	3.8	0.6	2.7	0.9	4.1
GETVPN	0.7	3.1	0.5	2.0	0.8	3.4
180-User Services Ready Large Network (24 simultaneous WAN calls)						
DMVPN	1.3	5.7	0.9	4.1	1.4	6.1
GETVPN	1.1	4.7	0.7	3.1	1.2	5.1
240-User Services Ready Large Network (32 simultaneous WAN calls)						
DMVPN	1.8	7.6	1.3	5.4	1.9	8.2
GETVPN	1.4	6.3	0.9	4.1	1.6	6.8

[Table 18](#) shows that the following user counts are appropriate for the various WAN connection options of the Services Ready Large Branch Network:

- T3/E3 line: Up to 240 users with RTP

- 4 T1 lines: Up to 240 users with cRTP and 200 users with RTP
- 1000BASE-LX/LH GigE shaped to 12 Mb/s: Up to 240 users with RTP

Besides considering provisioning of bandwidth for voice bearer traffic, you should consider bandwidth requirements for call control traffic. For centralized call control described below, the following calculations can be used to determine the amount of required bandwidth in a VPN network:

- SCCP Phone Traffic with VPN:

Bandwidth (bps) = 415 * (number of IP Phones and gateways in the branch)

- SIP Phone Traffic with VPN:

Bandwidth (bps) = 619 * (number of IP Phones and gateways in the branch)

A 240-user Services Ready Large Branch Network requires less than 100 kb/s for SCCP phone traffic, and 150 kb/s for SIP phone traffic, which is well below the 5 percent maximum assumed in the preceding calculations and well below the 20 percent maximum allocated through the QoS mechanism.

For the local call control described below the following calculation can be used to determine the amount of required bandwidth in a VPN network:

Bandwidth (b/s) = 116 * (number of telephone lines)

A 240-user Services Ready Large Branch Network requires less than 32 kb/s for H.323 or SIP control traffic, which is also well below the 5 percent maximum assumed in the above calculations and well below the 20 percent maximum allocated through the QoS mechanism.

In most cases, an Internet-based backup link (for example, xDSL) does not provide enough bandwidth and link quality to support voice traffic. Therefore, the Services Ready Large Branch Network uses PSTN as the backup link for voice traffic.

To learn more about voice communication in a VPN network see the *Voice and Video Enabled IPsec VPN (V3PN) Design Guide* at:

http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/V3PN_SRND/V3PN_SRND.html

IP Telephony

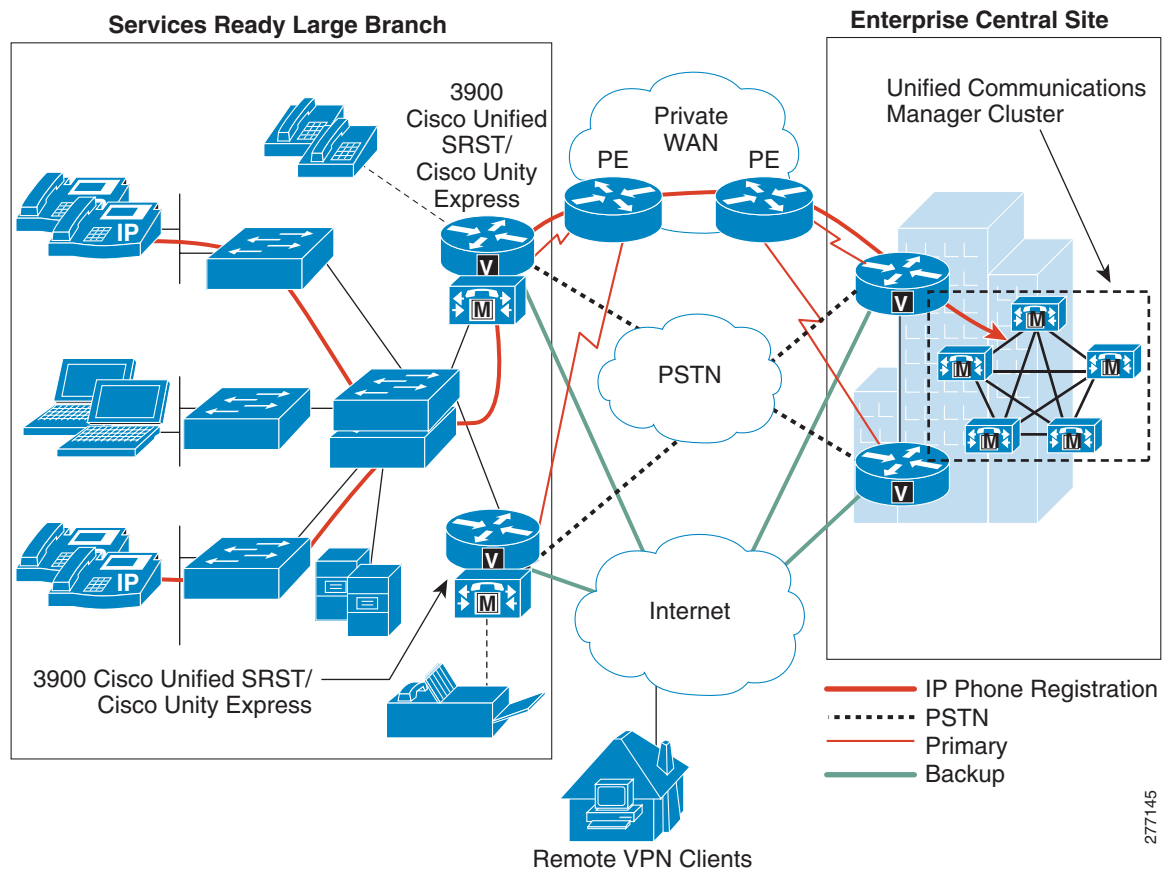
- [Centralized Call Control, page 80](#)
- [Local Call Control, page 81](#)
- [Selecting a Call Control Model, page 82](#)
- [IP Phones, page 83](#)
- [Voice Gateway, page 84](#)
- [Call Admission Control, page 87](#)
- [Conferencing and Transcoding, page 89](#)
- [Music on Hold, page 90](#)
- [Dial Plan, page 90](#)
- [Voice Mail and Auto Attendant Services, page 91](#)

The call control agent is a component of IP telephony that is responsible for overall coordination of all audiovisual communication. The agent has three typical deployment models: single site, multisite centralized, and multisite distributed call control (local). The Services Ready Large Branch Network assumes the presence of an enterprise central site; therefore, only the multisite centralized and distributed call control models were evaluated.

Centralized Call Control

The centralized call control model consists of a centrally located Cisco Unified Communications Manager (Cisco Unified CM) cluster that provides services for many branch offices and uses the WAN to transport voice traffic between the sites. The WAN also carries call signaling traffic between the central site and the branches. The Centralized Call Processing Model shown in [Figure 50](#) depicts the centralized call control deployment with a Cisco Unified CM cluster as the call control agent at the central site and with a WAN connection to the Services Ready Large Branch Network. The branch relies on the centralized Cisco Unified CM cluster to handle its call control. Applications such as voice mail and music on hold (MOH) are provided in the branch to reduce the amount of traffic traversing the WAN.

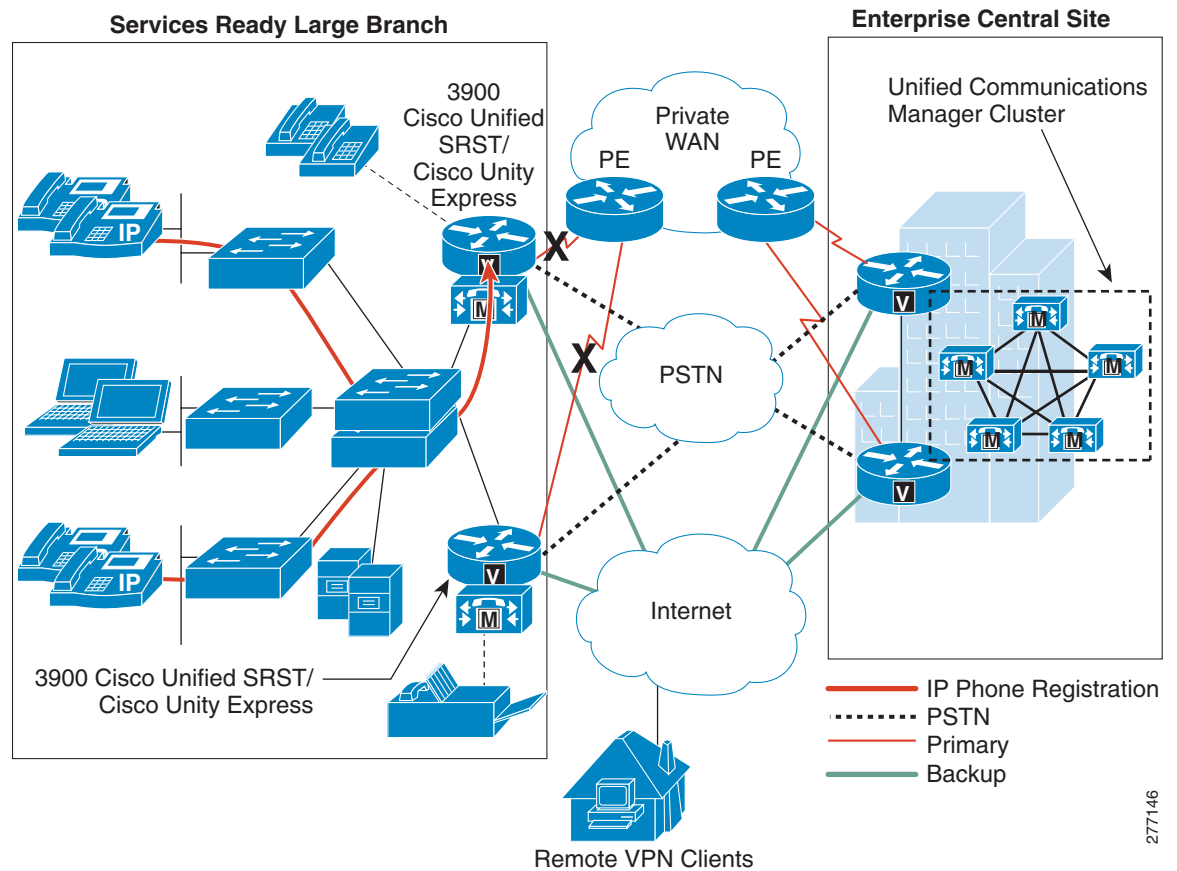
Figure 50 Centralized Call Control Model



Under normal operations shown on the left in [Figure 50](#), the branch office connects to the central site via a WAN, which carries data traffic, voice traffic, and call signaling. IP Phones at the branch exchange call signaling information with the Cisco Unified CM cluster at the central site. The voice gateway in the router forwards both types of traffic (call signaling and voice) transparently and has no “knowledge” of the IP Phones in the branch.

If the WAN link to the branch office fails, or if some other event causes loss of connectivity to the Cisco Unified CM cluster, the branch IP Phones reregister with the branch router that is running Cisco Unified Survivable Remote Site Telephony (Cisco Unified SRST) agent, as shown in Figure 51. The Cisco Unified SRST queries the IP Phones for their configuration and uses this information to build its own configuration automatically. The branch IP Phones can then make and receive calls either internally or through the PSTN. The phone displays the message “Unified CM fallback mode,” and some advanced Cisco Unified CM features are unavailable and are dimmed on the phone display. When WAN connectivity to the central site is reestablished, the branch IP Phones automatically reregister with the Cisco Unified CM cluster and resume normal operation. The branch Cisco Unified SRST router deletes its information about the IP Phones and reverts to its standard gateway configuration.

Figure 51 Cisco Unified SRST Mode for Centralized Call Control Model



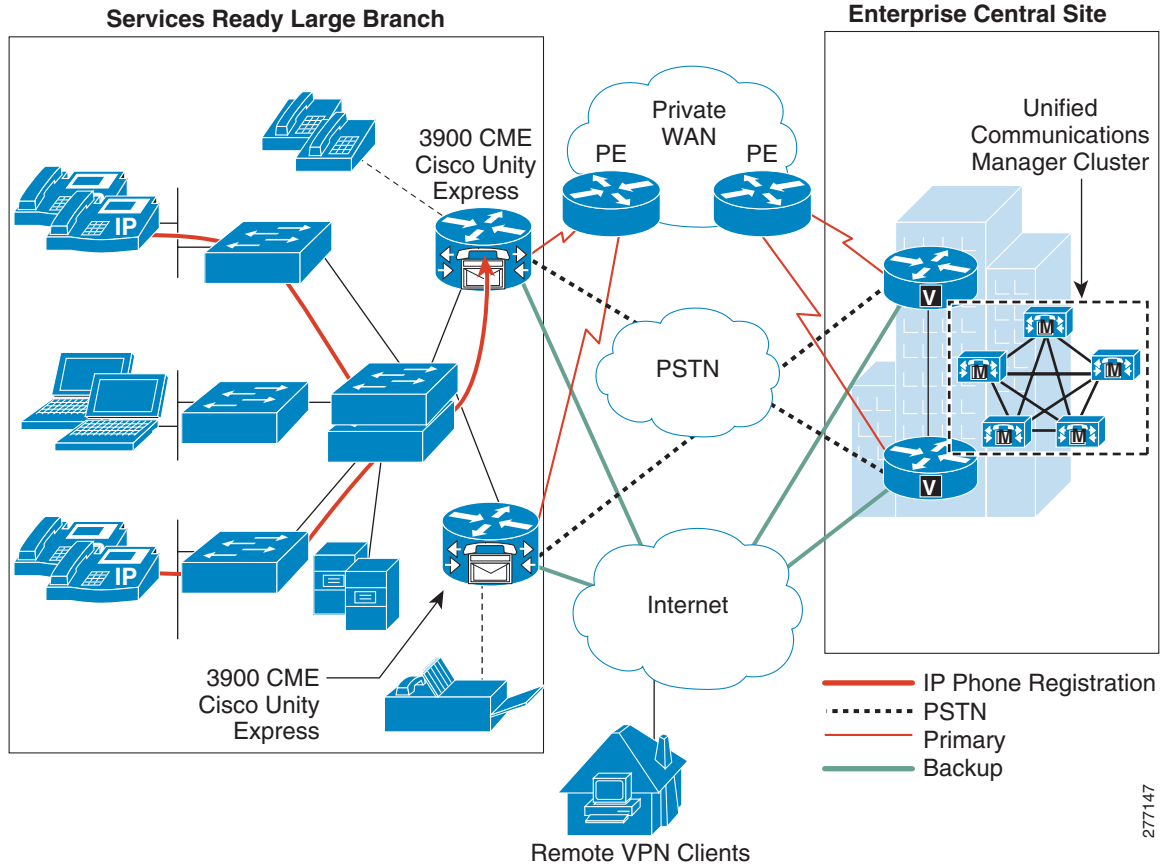
To learn more about Cisco Unified CM, visit:

http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/6x/uc6_1.html

Local Call Control

In the local call control model, each branch has its own Cisco Unified Communications Manager Express (Cisco Unified CME) connected to a WAN that carries voice traffic between the enterprise branches and central site. The PSTN serves as a backup connection between the sites if the WAN connection fails or has no more bandwidth available for additional calls. All call functionality is provided locally through Cisco Unified CME, and all IP Phones are registered locally, as shown in Figure 52. Applications such as voice mail and music on hold are provided in the branch router.

Figure 52 Distributed Call Control Model



277147

The local call control model eliminates dependency on out-of-the-branch control elements that would otherwise have to be accessed over the WAN. Thus, a WAN link failure has no effects on functionality provided by the IP telephony network; the network changes only the path over which the external WAN calls are routed.

To learn more about Cisco Unified CME, visit:

http://www.cisco.com/en/US/docs/voice_ip_comm/cucme/srnd/design/guide/cmesrnd.html

Selecting a Call Control Model

Although the local call control model has better availability properties than the centralized model, this advantage comes at an expense of additional functionality and management. Selecting the appropriate model involves numerous considerations. Table 19 describes the general trade-offs between the two models.

Table 19 Trade-offs Between Centralized and Local Call Control Models

Factor	Centralized Model	Local Model
WAN link characteristics	Needs more bandwidth and is more sensitive to link delay	Needs less bandwidth and is less sensitive to link delay
High availability	Impacted by WAN link failure	No WAN dependencies

Table 19 Trade-offs Between Centralized and Local Call Control Models (continued)

Factor	Centralized Model	Local Model
Feature set	More features	Fewer features
Scalability	Scales better	Scales poorly
Management	Centralized	Per-branch office

When deciding between the two deployment models, you must consider the overall enterprise voice deployment and any existing voice systems already in use. The Services Ready Large Branch Network was validated with both centralized call control using Cisco Unified CM with Cisco Unified SRST and with local call control using Cisco Unified CME.

IP Phones

Cisco IP Phones described in the “[Selecting Network Components](#)” section on page 3 can operate in either Skinny Call Control Protocol (SCCP) or Session Initiation Protocol (SIP) mode. The main trade-off between SCCP and SIP is in the functionality supported and third-party interoperability. SCCP is a Cisco proprietary protocol with a large number of Cisco voice products supporting its various features. SIP, on the other hand, is based on an open standard and has been adapted by a larger number of VoIP vendors. The Services Ready Large Branch Network has been tested with both SIP and SCCP phones, with both the centralized call control model and the local call control model.

In addition to the IP Phones described previously, the Services Ready Large Branch Network also uses Cisco IP Communicator, a software-based application that runs on a PC. The Cisco IP Communicator, shown in [Figure 53](#), only uses SCCP for call signaling.

Figure 53 Cisco IP Communicator

To learn more about the Cisco IP Communicator product, visit:

http://www.cisco.com/en/US/prod/collateral/voicesw/ps6788/phones/ps5475/product_data_sheet0900aec8064efe0.html

Voice Gateway

Both VoIP and IP telephony networks require a gateway to convert voice and signaling information between the traditional PSTN system and an IP-based system. The gateway must interpret PSTN analog or digital signaling to provide connectivity. A Cisco IOS voice gateway provides a full range of signaling options. Analog signaling and Basic Rate Interface (BRI)-based digital signaling provide PSTN connectivity for branch offices with a relatively small number of users. Only T1 or E1 digital voice trunks support the required number of users in a large branch office. [Table 20](#) displays the various Cisco IOS digital signaling options that pertain to the Services Ready Large Branch Network.

Table 20 Cisco IOS Software Support for High Density Digital Signaling Protocols

Signaling	Description	Typical Use
T1 CAS	Channel Associated Signaling	Used widely in North America to connect to the PSTN or PBXs. Several variations of this signaling exist, including T1 FXS, T1 FXO, and T1 E&M. T1 FXS and T1 FXO support loop start and ground start signaling. T1 E&M signaling supports delay dial, wink, and immediate dial.
T1 FGD	Feature Group D	The T1 CAS variations generally cannot convey caller ID, but T1 FGD can convey caller ID. It is used to connect to the PSTN where caller ID is required and where PRI is not an option. T1 FGD is an asymmetric protocol.
T1/E1 PRI	ISDN Primary Rate Interface	An ISDN connection to the PSTN carrying 23 (T1) or 30 (E1) simultaneous voice calls, introducing the terms <i>23B+D</i> and <i>30B+D</i> . T1/E1 PRI uses the Q.931 ISDN specification. Calls are controlled via a dedicated signaling channel (D channel).
T1 PRI NFAS	Nonfacility Associated Signaling	A variation of PRI available only on T1 that uses a single D channel to control multiple spans of T1s with only B channels (voice calls).
E1 R2	Regional System 2 (R2) CAS protocol	Used in South America and Asia for PSTN connectivity. There are numerous country-specific variations of the R2 protocol.

The Services Ready Large Branch Network used a T1 PRI trunk to connect the branch network to the PSTN. The T1 connection is provided by the local telephone company and runs to the nearest central office (CO) in the area. In future updates to this guide, some of the other options listed in [Table 20](#) will be validated and documented. The following interface card was used to connect to the PSTN and to provide the T1 PRI trunk:

- 1 Port T1/E1 Multiflex Trunk Voice/WAN interface card (VWIC2-1MFT-T1/E), shown in [Figure 54](#).

Figure 54 1 Port T1/E1 Multiflex Trunk Voice/WAN Interface Card



The VWIC2-1MFT-T1/E1 interface card provides up to 23 individual PSTN channels. Both oversubscription ratio and Extended Erlang B calculations, provided in [Table 15](#) and [Table 16](#) respectively, show that a typical large branch office does not require more than 12 PSTN lines. Therefore, only a fractional T1 connection is necessary to meet the requirements of the Services Ready Large Branch Network outlined in the “[Large Branch Design Considerations](#)” section on [page 4](#). Consequently, only two one-half T1 lines with 12 channels each were connected to the two branch routers. Although only one of the routers actively routes voice calls that originated on an IP Phone, each router provides connectivity for analog phones and faxes. These analog devices function independent of the Active/Standby high-availability mechanism and therefore both T1 lines are active.

To learn more about the Multiflex Trunk Voice WAN Interface Card, visit:

http://www.cisco.com/en/US/prod/collateral/routers/ps5855/product_data_sheet0900aecd8028d2db.html

Digital signal processor (DSP) technology provides voice compression, echo cancellation, tone generation, and voice packetization functions for servicing voice interfaces and converting voice signals for transport over IP networks. A digital PSTN voice port must have access to DSP resources in order to digitize and packetize the analog signal coming from the PSTN line. In the Services Ready Large Branch Network, the DSP resources were provided by the router because the VWIC2-1MFT-T1/E1 card has no DSPs. The number of required DSP modules depends on the amount and type of voice traffic in the branch. [Table 21](#) shows the number of required packet voice DSP modules (PVDMs) for branch offices with various active call requirements. The number of PSTN calls corresponds to the number of estimated active calls calculated from oversubscription ratios, as shown in [Table 15](#). Transcoding, conferencing, and analog phone/fax connectivity are described in later sections. The DSP calculator is available at:

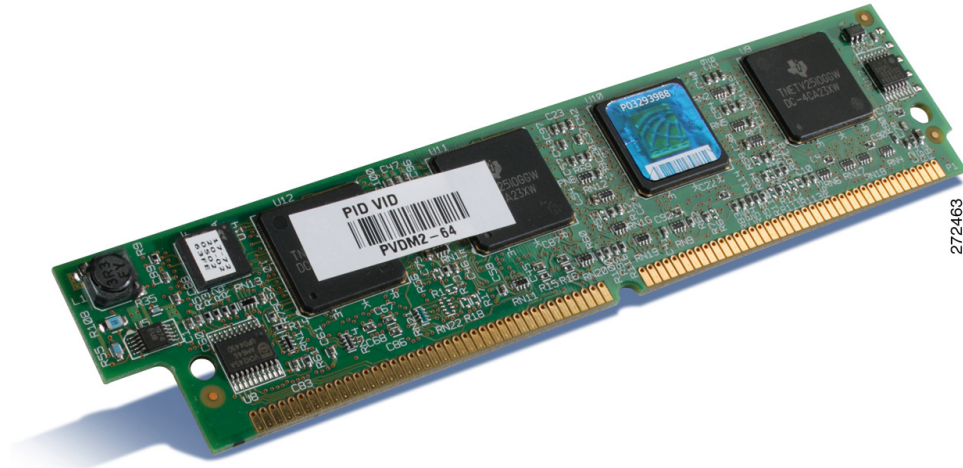
http://www.cisco.com/cgi-bin/Support/DSP/cisco_dsp_calc.pl

Table 21 DSP Requirements for Various Voice Configurations, Assuming G.711 Encoding and Fax Pass-through

Configuration	4 PSTN calls, 4 analog devices, 3 transcoding sessions, 1 conference calls	6 PSTN calls, 4 analog devices, 4 transcoding sessions, 2 simultaneous conference calls	8 PSTN calls, 4 analog devices, 5 transcoding sessions, 3 simultaneous conference calls
Number of onboard DSPs	3	3	5
PVDMs	PVDM3-32 + PVDM3-8	PVDM3-48	PVDM3-64 + PVDM3-8

The Services Ready Large Branch Network used 4 DSPs with the PVDM3-64 module. The PVDM3-64 module, shown in [Figure 55](#), supports up to 64 G.711 channels or 32 channels for medium-complexity codecs such as G.729a or Fax Relay.

Figure 55 64-Channel Packet Fax/Voice DSP Module



Besides physical connectivity and signal conversion, you must consider other PSTN services when configuring the voice gateway. The T1 PRI signaling mechanism that was selected for the Services Ready Large Branch Network supports the following PSTN services:

- Direct Inward Dial (DID): Enables callers to dial directly to an extension without the assistance of an operator or automated call attendant. In the Services Ready Large Branch Network, several users, the main office, and the PSTN voice mail access have DID numbers. Calls to other users in the office terminate on an automated attendant.
- Traditional fax services continue to be a widely used mechanism for document delivery. Physical integration of fax into the Services Ready Large Branch Network is described in the [“Analog Device Connectivity” section on page 91](#). In addition to physical connectivity of fax machines, the voice gateway must support a mechanism for interoperability of analog fax with IP telephony networks.

In its original form, fax data is digital and is contained in High-Level Data Link Control (HDLC) frames. However, to transmit across a traditional PSTN, these digital HDLC frames are modulated onto an analog carrier. While this analog carrier is necessary for effective faxing in PSTN environments, it is not ideal for the type of digital transport used by IP packet networks. Therefore, specific transport methods have been devised for successful transport of fax transmissions over an IP infrastructure.

The two main methods for transporting fax over IP are pass-through and relay. Pass-through is the simplest method. It works by sampling and digitizing the analog fax signal just as a voice codec does for human speech. While there are a number of codecs available, pass-through always uses the G.711 codec for carrying fax information because it offers the least distortion of the analog fax signals. Fax pass-through works only with the call control protocols of H.323 and SIP. Because fax pass-through utilizes the call control protocol for its switchover, this is the only pass-through solution that can work with third-party devices.

Relay is the other main method for transporting fax over IP. Relay strips off the analog carrier from the fax signal, in a process known as *demodulation*, to expose the fax HDLC data frames. The pertinent information in these HDLC frames is then removed and efficiently packaged in a fax relay protocol to be transported to the gateway on the other side. After it is received on the other side, the fax information is pulled from the relay protocol, reconstructed into fax HDLC frames, and modulated on to an analog carrier for transmission to a fax machine.

Cisco supports two versions of Fax Relay, T.38 and Cisco Fax Relay. An ITU standard, T.38 allows Cisco gateways to interoperate with third-party devices that also support the T.38 specification. In most scenarios, T.38 Fax Relay uses the call control protocol to switch from voice mode to T.38 fax

relay mode. Fax Relay mode, and more specifically T.38, is the preferred method for transporting fax traffic. The Services Ready Large Branch Network used both T.38-based fax relay and fax pass-through.

Two VoIP-enabled endpoints must use a common protocol stack to perform speech coding, call setup, signaling, data transport, and other functions related to voice communication. To ensure its relevance and applicability, The Services Ready Large Branch Network was validated with the following VoIP protocol stacks:

- **H.323:** Defines a suite of protocols, algorithms, and hardware specifications for audiovisual communication over IP-based network. The suite provides a complete protocol stack and defines precisely what is allowed and what is forbidden. H.323 includes speech coding algorithms such as G.711; RTP-based data transport; RTCP for controlling data channels; H.225 protocol for registration, admission, and status control; Q.931 call signaling protocol; and H.245 call control protocol.
- **Session Initiation Protocol (SIP):** Defines a protocol for setting up audiovisual connections over an IP network. Unlike H.323, which provides a complete protocol stack, SIP is a single, extensible module that has been designed to interwork with existing network-based applications. It is a text-based protocol modeled on HTTP.
- **Skinnny Client Control Protocol (SCCP):** Lightweight protocol used to set up calls between Cisco IP Phones and a voice gateway proxy (for example, Cisco Unified CME). The proxy communicates with the H.323 gateway, using H.225 and H.245 signaling, and the IP Phone using the SCCP protocol. The IP Phone requires less processing overhead because most of the H.323 processing resides in the proxy.

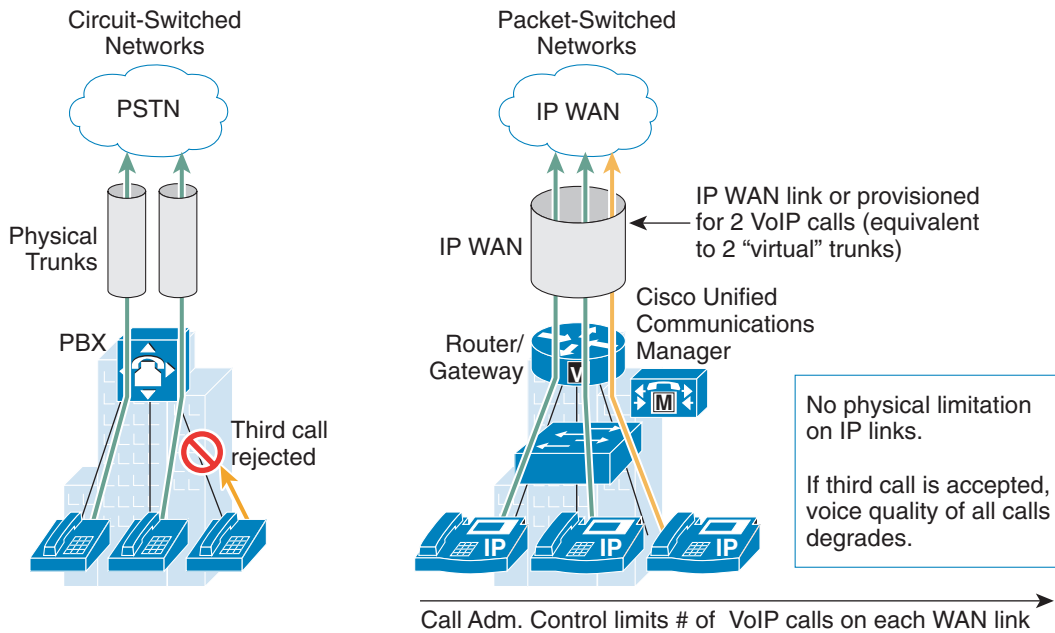
The choice between H.323 and SIP depends on the enterprise and is often based on feature requirements as well as interoperability with existing systems (for example, PBX, voicemail). In the Services Ready Large Branch Network, the following four combinations of call control agent, IP Phone protocol, and gateway-to-gateway protocol were validated:

- Cisco Unified CME with SCCP endpoints and H.323 trunk
- Cisco Unified CME with SIP endpoints and SIP trunk
- Cisco Unified SRST with SCCP endpoints and H.323 trunk
- Cisco Unified SRST with SIP endpoints and SIP trunk

Call Admission Control

Call Admission Control (CAC) maintains high voice quality over an IP WAN by limiting the number of calls that are admitted. Traditional telephony circuits, in which physical channels limit the number of calls allowed to connect to the PSTN, do not have this requirement. When VoIP calls traverse an IP WAN, calls are packet streams and there are no physical limitations that control the number of calls admitted to the WAN link. An IP WAN link can easily be oversubscribed, and the voice quality of all connected calls can be degraded, as shown in [Figure 56](#).

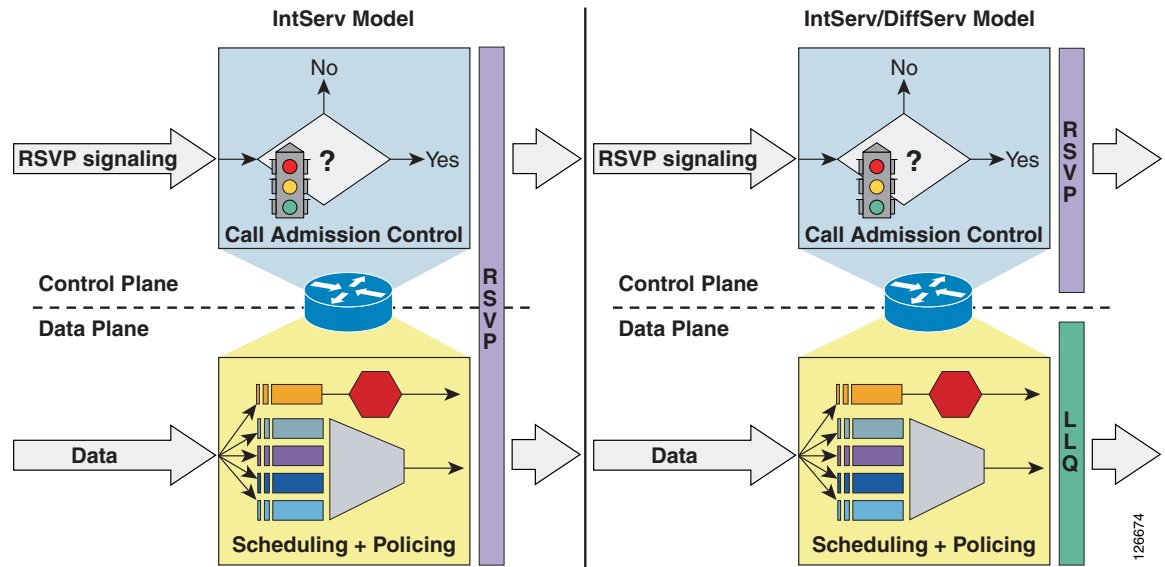
Figure 56 Traditional Versus VoIP Call Admission Control



Resource Reservation Protocol (RSVP) is a mechanism for dynamically setting up end-to-end QoS across a heterogeneous network. A resource reservation is created by exchanging signaling messages between the source and destination devices that are processed by intermediate routers along the path. The signaling messages "reserve" bandwidth at the intermediate routers for each unidirectional data flow. RSVP can propagate RSVP requests across routers that do not support the protocol. There are two operational models for RSVP, as described below and shown in [Figure 57](#).

- **IntServ:** Controls resource reservation at both control and data planes. In the control plane, RSVP admits or denies the reservation request. In the data plane, it classifies the data packets, polices them based on the traffic description contained in the RSVP messages, and queues them in the appropriate queue.
- **IntServ/DiffServ:** Controls resource reservation at the control plane only. This means that the CAC function is separate from the scheduling and policing functions, which can be performed by the low latency queuing (LLQ) algorithm according to predefined class maps, policy maps, and service policies. With the IntServ/DiffServ model, it is therefore possible to add RSVP CAC to a network that is already using a differentiated services approach to QoS. RSVP admits or rejects calls, based on a preconfigured bandwidth amount, but the actual scheduling is based on the preexisting LLQ criteria such as the DSCP value of each packet.

Figure 57 *RSVP Operational Models: IntServ and IntServ/DiffServ*



The Services Ready Large Branch Network used the IntServ/DiffServ RSVP mechanism to control the number of calls placed on the network, but relied on the established QoS policy explained in the [“Quality of Service” section on page 46](#) to control actual packet scheduling. This model is appropriate for the Services Ready Large Branch Network because all LLQ-destined traffic is controlled by RSVP.

At present, RSVP is supported only in the centralized call control model with Cisco Unified SRST. To simulate the function of RSVP for the local call control model with Cisco Unified CME, a simple maximum call limit was placed on the WAN voice gateway.

Conferencing and Transcoding

Conferencing joins multiple participants into a single call. The number of media streams connected to a conference corresponds to the number of participants. A conference bridge mixes the streams together and creates a unique output stream for each connected participant. The output stream for a given participant is the composite of the streams from all connected participants minus their own input stream. The conference bridge is controlled by Cisco Unified CM or Cisco Unified CME. A conference bridge is allocated from the onboard DSPs. The Services Ready Large Branch Network was designed to support up to three simultaneous conferencing sessions. Cisco Unified CME provides conferencing locally through the branch router, while the centralized call control model leverages the conferencing functionality of the Cisco Unified CM in the central site.

Transcoding converts an input stream from one codec into an output stream that uses a different codec. It may also connect two streams that utilize the same codec but with a different sampling rate. Transcoding is typically used to convert between a G.711 voice stream and the low bit-rate compressed voice stream G.729a. The Services Ready Large Branch Network used transcoding to support endpoints that are configured for G.711 only. This condition exists when G.729a is used in the system but there are devices that do not support this codec, or there is a device with G.729a support that may be configured to not use G.729a. The Services Ready Large Branch Network was designed to support up to five simultaneous transcoding sessions.

The G.711 codec was used for LAN calls to maximize call quality and the G.729a coded was used for calls that traverse a WAN to maximize bandwidth efficiency. The G.729a codec is supported on all Cisco Unified IP Phone models and therefore G.711 to G.729a transcoding is not required.

Music on Hold

Music on hold (MOH) provides music to callers when their call is placed on hold, transferred, parked, or added to an ad-hoc conference. The integrated MOH feature allows both internal and external users to place users on hold with music streamed from a streaming source. There are two types of MOH transport mechanism: unicast and multicast. The Services Ready Large Branch Network used unicast to transport MOH data in the local call control mode (Cisco Unified CME). In the case of centralized call processing, multicast is used to transport MOH data. Multicast MOH consists of streams that are sent from the MOH source to a multicast group IP address, to which endpoints requesting an MOH audio stream can join. A multicast MOH stream is a point-to-multipoint, one-way audio RTP stream between the MOH source and the multicast group IP address. Multicast MOH conserves system resources and bandwidth because it enables multiple users to use the same audio source stream.

In the case of SCCP phones, the multicast was enabled on the branch router. In the case of SIP phones, multicast was configured at the central Cisco Unified CM, and the branch router simply forwarded the traffic as it would any other multicast application.

In the Services Ready Large Branch Network, the MOH source was an audio file stored on the branch router, except for the centralized deployment option with SIP phones.

Dial Plan

The dial plan is one of the key elements of an IP telephony system, and is an integral part of all call control agents. Generally, the dial plan is responsible for instructing the call control agent on how to route calls. Specifically, the dial plan in the Services Ready Large Branch Network performs the following functions:

- Endpoint addressing: Reachability of internal destinations is provided by assigning directory numbers (DNs) to all endpoints (such as IP Phones, fax machines, and analog phones) and applications (such as voice mail systems, auto attendants, and conferencing systems).
- Path selection: A secondary path can be used when the primary path is not available. The secondary path is made by rerouting over the PSTN during an IP WAN failure.



Note Cisco Unified CME does not support path selection.

- Digit manipulation: In some cases, it is necessary to manipulate the dialed string before routing the call; for example, when rerouting over the PSTN, a call originally dialed using the access code, or when expanding an abbreviated code (such as 0 for the operator) to an extension.

Additional functions are possible and will be validated in the future update to this guide:

- Calling privileges: Different groups of devices can be assigned to different classes of service by granting or denying access to certain destinations. For example, lobby phones might be allowed to reach only internal and local PSTN destinations, but executive phones could have unrestricted PSTN access.
- Call coverage: Special groups of devices can be created to handle incoming calls for a certain service according to different rules (top-down, circular hunt, longest idle, or broadcast).

The automated alternate routing (AAR) feature enables Cisco Unified CM to establish an alternate path for the voice data when the preferred path between two endpoints within the same cluster runs out of available bandwidth, as determined by the locations mechanism for call admission control. If a phone in one branch calls a phone in another branch, and the available bandwidth for the WAN link between the branches is insufficient, then AAR reroutes the call through the PSTN.

Voice Mail and Auto Attendant Services

All voice mail in the Services Ready Large Branch Network is stored locally in the branch for both centralized and distributed call control models. The Cisco Unity Express network module shown in [Figure 58](#) was used for voice mail services. Cisco Unity Express provides cost-effective voice and integrated messaging and automated attendant for enterprise branch offices with up to 240 users. The Cisco 3945 ISR used the Cisco Unity Express network module, shown in [Figure 58](#), while the Cisco 3925 ISR used the Advanced Integration Module 2 (AIM2) form factor, as shown in [Figure 59](#).

Figure 58 Cisco Unity Express Network Module



Figure 59 Cisco Unity Express Advanced Integration Module 2



Traditional Telephony

In the Services Ready Large Branch Network, traditional telephony is used to provide traditional fax services, emergency response, and call backup options as described in the following sections.

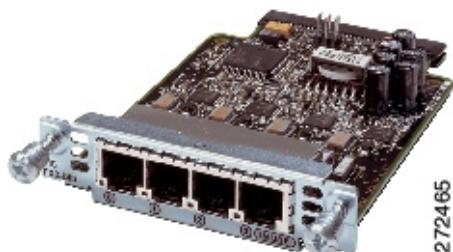
Analog Device Connectivity

There are various reasons to continue using some forms of traditional telephony in a branch office. For example, fax services continue to be widely used, and analog phones connected directly to a voice gateway can provide a backup of last resort. The Services Ready Large Branch Network used the following voice interface card for connecting traditional voice devices into the network:

- 4-port FXS/DID Voice Interface Card (VIC-4FXS/DID)

There are four FXS ports on the VIC-4FXS/DID card shown in [Figure 60](#). The ports were used for connecting a mixture of analog phones and faxes.

Figure 60 *Four Port FXS/DID Voice Interface Card*



Emergency Services

Emergency services are of great importance in a proper deployment of a voice system. The Services Ready Large Branch Network was validated with the 911 emergency network as deployed in Canada and the United States. The design and implementation described are adaptable to other locales. Please consult with your local telephony network provider for appropriate implementation of emergency call functionality.

In general, a local exchange carrier has a dedicated network for the 911 service. In the Services Ready Large Branch Network, the T1 PRI trunk connected the branch to the 911 service managed by Public Safety Answering Point (PSAP). In this configuration, the voice gateway provides emergency response location (ERL) and emergency location identification (ELIN) number.



Note

Advanced Emergency Services with ERL and ELIN information are currently supported only with Cisco Unified SRST. Cisco Unified CME implements 911 services by forwarding the call to a PSTN without any additional information.

To learn more about Emergency Services see:

http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/6x/e911.html

Optimization Services

This section covers the following optimization services topics:

- [Selecting a Cisco WAE Module, page 95](#)
- [Cisco WAAS General Design Considerations, page 96](#)
- [Cisco WAAS High-Availability and Rapid Recovery Considerations, page 97](#)
- [Cisco WAAS Security Considerations, page 98](#)
- [Cisco WAAS Management Considerations, page 98](#)

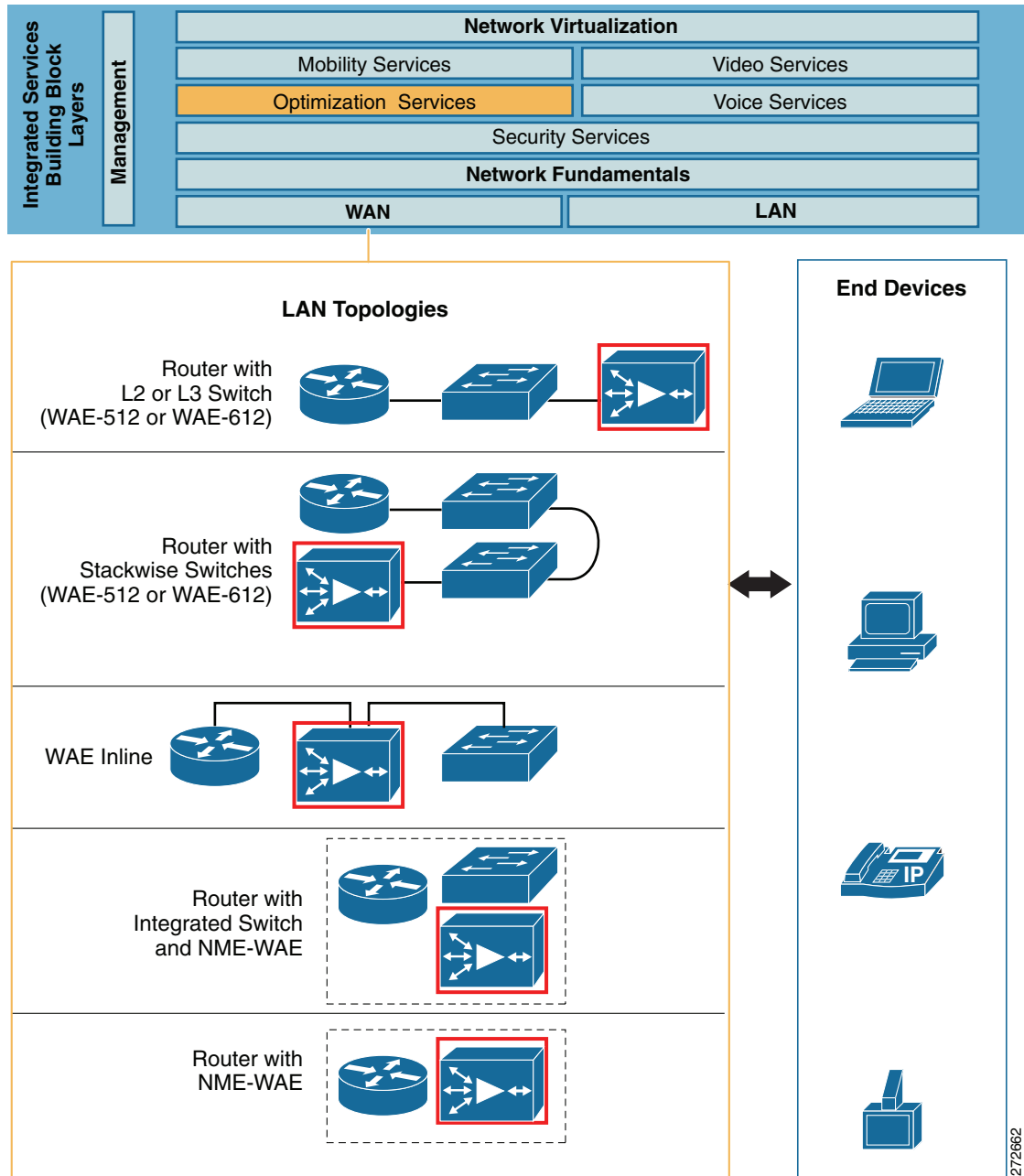
Limited bandwidth, high latency, packet drops, and frequent congestion limit the performance of client/server applications that communicate over the WAN. Because of these unfavorable WAN link characteristics, enterprises have traditionally opted to deploy distributed applications and enterprise middleware in the branch offices to ensure adequate performance. However, running applications in the branch offices increases management costs, complicates disaster recovery, and hampers regulatory compliance. Today, more and more enterprises are turning to WAN and application optimization technologies to deploy client/server applications over the WAN with a LAN-like performance.

The Cisco Wide Area Application Services (Cisco WAAS) technologies and products give enterprise branches LAN-like access to centrally hosted applications, servers, storage, and multimedia services. Cisco WAAS provides the enterprise branch with application delivery, acceleration, WAN optimization, and local service solutions to optimize performance of any TCP-based application in a WAN or metropolitan area network (MAN) environment. The Cisco WAAS software solution runs on the Cisco Wide Area Application Engine (Cisco WAE) family of hardware platforms. In general, there are three types of optimization services:

- WAN optimization: Provides mechanisms for improving performance of TCP based applications. Three techniques are widely used to optimize TCP applications:
 - Compression: Cisco WAAS uses a persistent version of the Lempel-Ziv (LZ) lossless compression algorithm.
 - TCP Optimization: Cisco WAAS uses an optimized version of TCP known as Transport Flow Optimization (TFO).
 - Caching: Cisco WAAS uses a data redundancy elimination (DRE) technique to cache duplicated data patterns.
- Application acceleration: Provides mechanisms for reducing “chattiness” of enterprise applications or protocols
 - Application based: Cisco WAAS provides adapters for optimizing specific applications (video, e-mail)
 - Protocol based: Cisco WAAS provides adapters for optimizing specific application layer protocols (SSL)
- Wide Area File Services: Provide faster access to files stored on a network file system.
 - Object caching: Cisco WAAS caches or pre-positions entire files and keeps the local copy synchronized to ensure accuracy.
 - Request prediction: Cisco WAAS inspects Common Internet File System (CIFS) and Network File System (NFS) messages and prefetches data by predicting follow-on requests.

Cisco WAAS offers different deployment options at the branch, as shown in [Figure 61](#).

Figure 61 Cisco WAAS Deployment Options



The Services Ready Large Branch Network focuses on services that are integrated into the branch router. Designs featuring standalone Cisco WAE appliances are not considered in this guide. However, designs with standalone Cisco WAE appliances are equally viable and are described in the *Enterprise Branch Wide Area Application Services Design Guide* at:

<http://www.cisco.com/en/US/docs/solutions/Enterprise/Branch/WAASBr11.html>.

Moreover, all switches in the Services Ready Large Branch Network are external to the router; therefore, only the integrated Cisco WAE without the integrated switch, shown in [Figure 61](#), was evaluated.

Selecting a Cisco WAE Module

Cisco WAAS is a symmetric solution that requires one Cisco WAE device in the branch and another at the central site. This guide focuses on Cisco WAE deployment in the branch. Selection, design, and configuration of the central site Cisco WAE are not considered in this guide. For more information on Cisco WAAS central site deployment, see the *Cisco Enterprise Data Center Wide Area Application Services (Cisco WAAS) Design Guide* at:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/WAASDC11.html

Table 22 lists and compares three Cisco WAE network module models that can be deployed in the Cisco 3900 ISRs.

Table 22 Cisco WAE Feature Comparison

Cisco Device	Max Optimized TCP Connections	Max CIFS Sessions	Drive Capacity (GB)	RAM (GB)	Max Recommended WAN Links	Max Optimized Throughput (Mb/s)	Optimization
NME-WAE-302	250	N/A	80	0.5	4	90	LZS, TFO, DRE
NME-WAE-502	500	500	120	1	4	150	LZS, TFO, DRE, applications, and protocols
NME-WAE-522	750	750	160	2	8	200	LZS, TFO, DRE, applications, and protocols

Only the Cisco NME-WAE-522 network module supports enough concurrent TCP connections and WAN bandwidth to meet the needs of a 100- to 240-user branch office. Therefore, NME-WAE-522, shown in Figure 62, was the only module validated for the Services Ready Large Branch Network. Two types of software licenses are available for the Cisco NME-WAE-522 network module:

- Transport license: Provides the WAN optimization features including Data Redundancy Elimination (DRE), Lempel-Ziv (LZ) compression, and Transport Flow Optimizations (TFO).
- Enterprise license: Includes the transport license functions plus application-specific accelerations such as Common Internet File System (CIFS) services and print services, disk encryption, and TCP Flow Agent for NetQoS integration.

The Enterprise license was used in the Services Ready Large Branch Network.

Figure 62 Cisco NME-WAE-522 Wide-Area Application Engine



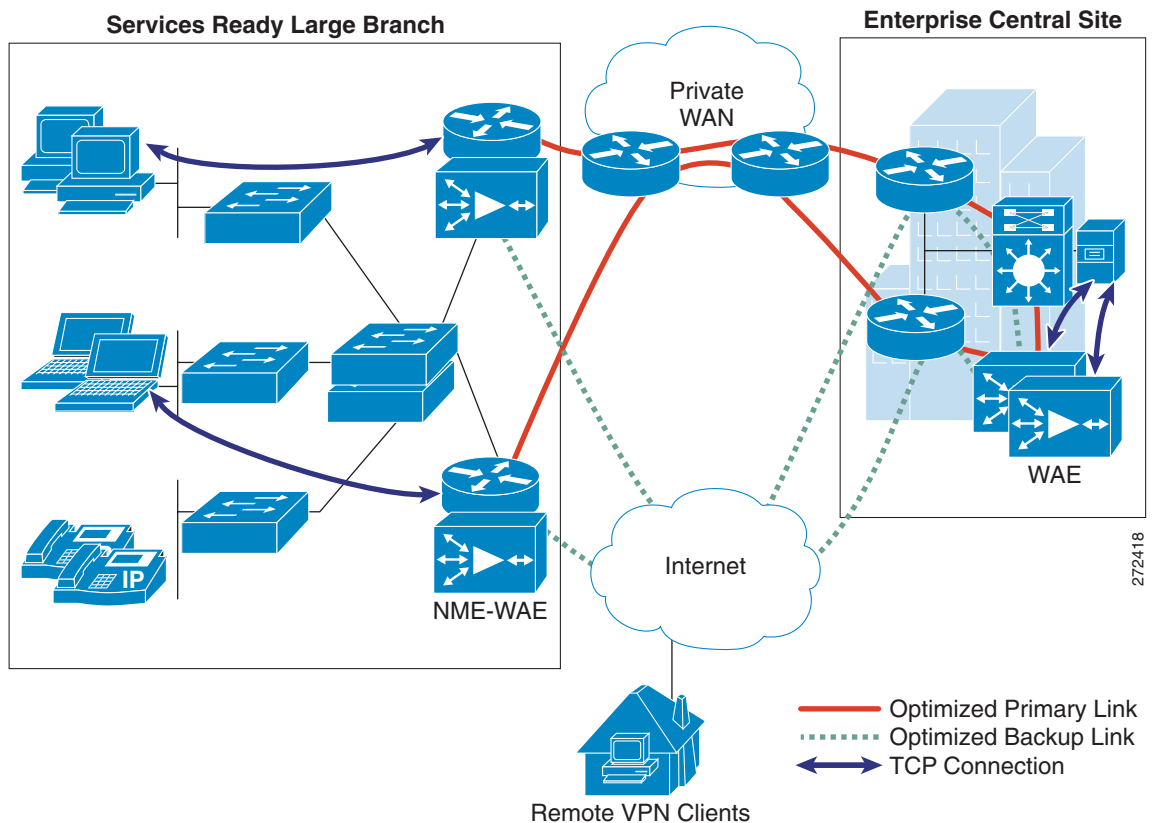
To learn more about the Cisco WAE network modules, visit:

http://www.cisco.com/en/US/prod/collateral/contnetw/ps5680/ps6870/product_data_sheet0900aecd8058218c.html

Cisco WAAS General Design Considerations

Figure 63 shows the Cisco WAAS optimization path for the Services Ready Large Branch Network. Traffic is optimized on both the primary link and the backup link in the event that the backup link becomes active.

Figure 63 Cisco WAAS Optimization Path



Cisco WAAS provides two options for controlling redirection and processing of application traffic:

- Dynamic mode: Uses Cisco Web Cache Communication Protocol (Cisco WCCPv2)
- Static mode: Uses policy-based routing (PBR)

Cisco WCCP is the preferred mechanism for interception and redirection because it is easier to configure, supports high availability, and provides clustering capabilities. To meet the design criteria in the [“Large Branch Design Considerations” section on page 4](#), the Services Ready Large Branch Network used only Cisco WCCPv2. Cisco WCCP is a Cisco IOS software feature that enables routing platforms to transparently redirect traffic. It provides several configuration options for traffic handling and distribution to the two Cisco WAE engines in the Services Ready Large Branch Network. The configuration involves:

- Traffic forwarding mechanism: Cisco WCCP traffic is forwarded to the Cisco WAE module using either GRE encapsulation or Layer 2 (L2) redirection. L2 redirection is more appropriate when Cisco WAAS is deployed as an external appliance connected to a switch; therefore, only GRE encapsulation was used in the Services Ready Large Branch Network.
- Traffic redirection: Cisco WCCP uses service groups to redirect traffic for further processing to the appropriate Cisco WAE module. These service groups are determined by the web cache and configured for identification by Cisco WCCP. The Cisco WAAS TCP promiscuous mode uses Cisco WCCP service groups 61 and 62 for traffic redirection. Service group 61 is in the path of packet flow for one direction, and service group 62 is in the path of packet flow for the opposite direction.
- Intelligent traffic filtering: Application traffic policies (ATPs) are defined to optimize only specific types of application traffic. All other traffic is pass-through. The following traffic was optimized in the Services Ready Large Branch Network:
 - HTTP
 - FTP
 - CIFS

The NME-WAE-522 network module has an internal interface through the router backplane and an external Fast/Gigabit Ethernet interface on the faceplate of the module. The internal interface is the recommended mode for deployment unless special considerations require the external interface. The Services Ready Large Branch Network used only the internal interface.

None of the WAN deployment scenarios described in the [“WAN Services” section on page 9](#) required modification of the default Cisco WAAS TFO transmit/receive buffers.

Cisco WAAS High-Availability and Rapid Recovery Considerations

- Cisco WAAS offers several mechanisms to guarantee rapid error recovery: Cisco WAAS DRE cache is persistent and loosely synchronized, enabling quick recovery in the event of a reboot or software restart.
- Cisco WAAS Device Manager offers the ability to back up individual devices for fast restore onto a standby/replacement device.

Cisco WAAS Security Considerations

Zone-based Policy Firewall was configured to support Cisco WAAS traffic. The Cisco WAE network module was placed into the Private zone of the firewall. Traffic from and to Cisco WAAS was encrypted and decrypted before moving to and from the VPN zone.

Cisco WAAS Management Considerations

Cisco WAAS Central Manager was used to configure the two Cisco WAE network modules. See the [“System Implementation”](#) chapter on [page 1](#).