CHAPTER 6

IPSec VPN QoS Design

IPSec VPNs are the most widely deployed VPNs and are found in three main contexts:

- Site-to-site IPSec VPNs
- Teleworker IPSec VPNs
- Remote-access client (mobility) IPSec VPNs

QoS considerations for site-to-site and teleworker IPSec VPNs are examined in this design chapter (as QoS is rarely—if ever—deployed in remote-access client IPSec VPN scenarios). These considerations include the following:

- IPSec modes of operation
- Bandwidth and delay increases because of encryption
- IPSec and cRTP incompatibility
- IP ToS byte preservation through IPSec encryption
- QoS and Anti-Replay interaction implications

Following a discussion of these considerations, design recommendations for site-to-site and teleworker (DSL and cable) solutions are presented in detail.

Whereas MPLS technologies provide VPN services, such as network segregation and privacy, by maintaining independent virtual router forwarding tables, IPSec achieves such VPN services through encryption.

As defined in RFCs 2401 through 2412, IPSec protocols provide mechanisms to enable remote sites to cost effectively connect to enterprise intranets or extranets using the Internet (or a service provider's shared IP networks). Because of IPSec protocol encryption, such VPNs can provide the same management and security policies as private networks.

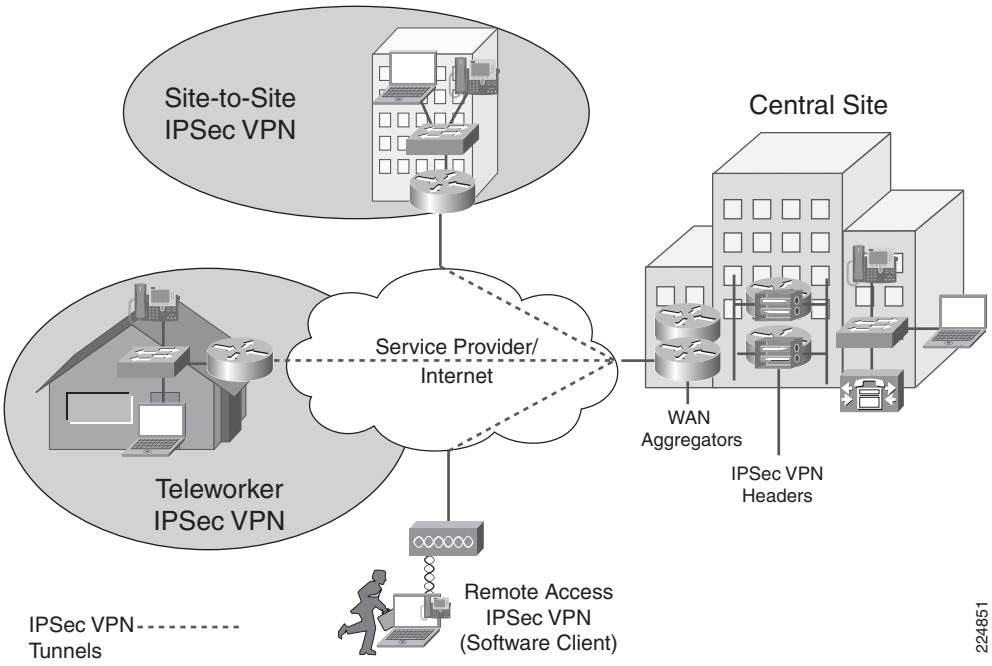
IPSec VPN services are built by overlaying a point-to-point mesh over the Internet using Layer 3–encrypted tunnels. This architecture requires security appliances, such as (hardware or software) firewalls or routers that support IPSec tunnels, to be installed at both ends of each tunnel. Encryption/decryption is performed at these tunnel endpoints, and the protected traffic is carried across the shared network.

Three main design contexts for IPSec VPNs exist, as shown in Figure 6-1:

- **Site-to-site VPNs**—Tunnels are maintained by hardware devices to connect multiple users at remote branches to (one or more) central sites.
- **Teleworker VPNs**—Tunnels are maintained by hardware devices to connect (typically) a single user at his or her residence to a central site.

- **Remote-access clients**—Tunnels are established by software to connect mobile users at airports, hotels, or similar places to a central site using WLAN hotspots, LAN ports, or modems.

Figure 6-1 IPSec VPN Design Contexts



Enabling converged services, such as voice and video, on an IPSec VPN has been dubbed V3PN. V3PN is essentially the overlaying of QoS technologies over IPSec VPNs to provide the required service levels to voice and video applications. As such, V3PN solutions relate to only two of the three IPSec VPN design contexts: site-to-site VPNs and telecommuter VPNs. (Little, if any, QoS is available in remote-access client networks.)

This chapter discusses QoS design considerations and recommendations for both site-to-site and telecommuter V3PN solutions.

Note It is beyond the scope of this chapter to detail IPSec encryption operation and configuration; a working knowledge of IPSec is assumed.

Site-to-Site V3PN QoS Considerations

Attractive pricing is usually the driver behind deploying site-to-site IPSec VPNs as an alternative to private WAN technologies. Many of the same considerations required by private WANs need to be taken into account for IPSec VPN scenarios because they usually are deployed over the same Layer 2 WAN access media.

IPSec VPNs also share some similar concerns with MPLS VPNs. For instance, the enterprise's end-to-end delay and jitter budgets depend significantly on the service provider's SLAs. Therefore, enterprises deploying V3PN solutions are recommended to utilize Cisco Powered Network IP Multiservice service providers, as discussed in Chapter 5, "MPLS VPN QoS Design."

However, IPSec VPNs present many unique considerations for QoS design, including the following (each is discussed in detail throughout the rest of the chapter):

- IPSec VPN modes of operation
- Packet overhead increases because of encryption
- cRTP and IPSec incompatibility
- Prefragmentation
- Bandwidth provisioning
- Logical topologies
- Delay budget increases because of encryption
- ToS byte preservation
- QoS Pre-Classify feature
- Pre-encryption queuing
- Anti-Replay implications
- Control plane provisioning

IPSec VPN Modes of Operation

Three principal modes of IPSec VPN operation exist:

- IPSec Tunnel Mode (No IP GRE Tunnel)
- IPSec Transport Mode with an Encrypted IP GRE Tunnel
- IPSec Tunnel Mode with an Encrypted IP GRE Tunnel

The advantages, disadvantages, features, and limitations of these options are discussed next.

IPSec Tunnel Mode (No IP GRE Tunnel)

This option does not utilize an IP GRE tunnel. With this option, only IPSec unicast traffic can be transported. (IP multicast traffic cannot be transported between IPSec peers without configuring an IP GRE tunnel.)

This configuration might be sufficient to support application requirements; its advantage lies in lower CPU overhead (primarily at the headend IPSec VPN router) compared with alternative IPSec design options.

IPSec security associations (SAs) are created for each access list line matched. An access list must be specified in the crypto map to designate packets that are to be encrypted. Such an access list typically entails several lines to define the application(s) to be encrypted by the five ACL tuples: source/destination IP address, protocol, and source/destination port numbers. When not encrypting a GRE tunnel, it is possible to create a separate SA for each application or access-list line match or to create an SA that carries all traffic that matches an ACL range (which is recommended). Each SA has its own Encryption Security Protocol (ESP) or Authentication Header (AH) sequence number.

Anti-Replay drops can be eliminated or minimized by constructing access lists that create a separate security association for each class of traffic being influenced by per-hop QoS policies. (Anti-Replay is an IPSec standard feature that discards packets that fall outside a receiver's 64-byte sliding window because such packets are considered suspect or potentially compromised—it is discussed in greater detail later in this chapter.)

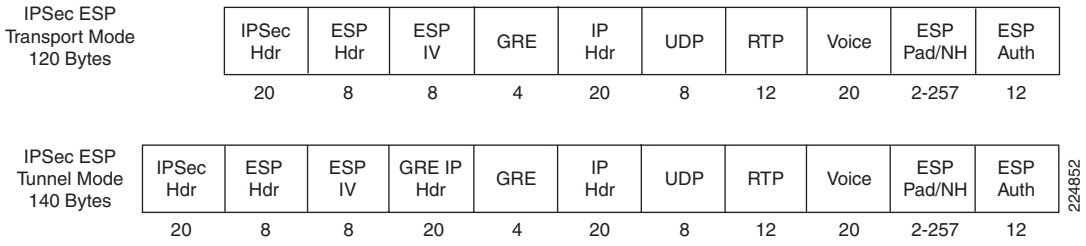
The Cisco IOS feature of prefragmentation for IPSec VPNs (also discussed later in this chapter) is supported in IPSec tunnel mode (no IP GRE tunnel) as of Cisco IOS Release 12.2(12)T.

IPSec Transport Mode with an Encrypted IP GRE Tunnel

IPSec transport mode (encrypting an IP GRE tunnel) is a commonly deployed option because it provides all the advantages of using IP GRE, such as IP Multicast protocol support (and, thus, also the support of routing protocols that utilize IP Multicast) and multiprotocol support. Furthermore, this option saves 20 bytes per packet over IPSec tunnel mode (encrypting an IP GRE tunnel) because an additional IP header is not required. Figure 6-2 illustrates IPSec transport mode versus tunnel mode when encryption is performed in an IP GRE tunnel.

The IPSec peer IP addresses and the IP GRE peer address must match for transport mode to be negotiated; if they do not match, tunnel mode is negotiated.

Figure 6-2 IPSec Transport Mode Versus Tunnel Mode for a G.729 VoIP Packet



The Cisco IOS prefragmentation feature for IPSec VPNs (discussed later in the chapter) is *not* supported for transport mode because the decrypting router cannot determine whether the fragmentation was done before or after encryption (for example, by a downstream router between the encrypting and decrypting routers).

Although IPSec transport mode saves a small to moderate amount of link bandwidth, it does not provide any reduction in packets per second switched by the router. Therefore, because the number of packets per second primarily affects CPU performance, no significant CPU performance gain is realized by using IPSec transport mode.

IPSec tunnel mode is the default configuration option. To configure transport mode, it must be specified under the IPSec transform set, as shown in Example 6-1.

Example 6-1 Enabling IPSec Transport Mode

```
!
crypto ipsec transform-set ENTERPRISE esp-3des esp-sha-hmac
 mode transport             ! Enables IPSec Transport mode
!
```

IPSec Tunnel Mode with an Encrypted IP GRE Tunnel

IPSec tunnel mode (encrypting an IP GRE tunnel) is the primarily recommended IPSec VPN design option. Although it incurs the greatest header overhead of the three options, it is capable of supporting IP Multicast (with the capability to run a dynamic routing protocol within the IP GRE tunnel for failover to an alternative path), and it supports prefragmentation for IPSec VPNs.

When configured with a routing protocol running within an IP GRE tunnel, the routing protocol's Hello packets maintain the security associations between the branch and both (assuming a redundant configuration) headend routers. There is no need to create a security association with a backup headend peer if the primary peer fails.

Note The design principles in this chapter were proven by scalability testing in the Cisco Enterprise Solutions Engineering labs. These large-scale testing methods were designed to test worst-case scenarios. From a design standpoint, these entailed enabling the following:
— Strong Triple-Digital Encryption Standard (3DES) encryption for both Internet Key Exchange (IKE) and IPSec
— IP GRE with IPSec tunnel mode
— Diffie-Hellman Group 2 (1024 bit) for IKE
— Secure Hash Algorithm (SHA) 160-bit RFC 2104 Keyed-Hashing for Message Authentication (HMAC) with RFC 1321 Message Digest 5 (MD5)
— (MD5)-HMAC (both hash algorithms truncated to 12 bytes in the ESP packet trailer)
— Preshared keys
If an enterprise chooses to implement less stringent security parameters, to use IPSec transport mode instead of tunnel mode, or to not implement IP GRE tunnels, the designs continue to be applicable from functional and scalability standpoints.

Packet Overhead Increases

The addition of tunnel headers and encryption overhead increases the packet sizes of all encrypted applications: voice, video, and data. This needs to be taken into account when provisioning LLQ or CBWFQ bandwidth to a given class.

For example, consider voice. The two most widely deployed codecs for voice are G.711 and G.729. Each codec typically is deployed at 50 pps (generating 20-ms packetization intervals).

The Layer 3 data rate for a G.711 call (at 50 pps) is 80 kbps. IP Generic Routing Encapsulation (GRE) tunnel overhead adds 24 bytes per packet. The IPSec Encapsulating Security Payload (ESP) adds another 56 bytes. The combined additional overhead increases the rate from 80 kbps (clear voice) to 112 kbps (IPSec ESP tunnel-mode encrypted voice).

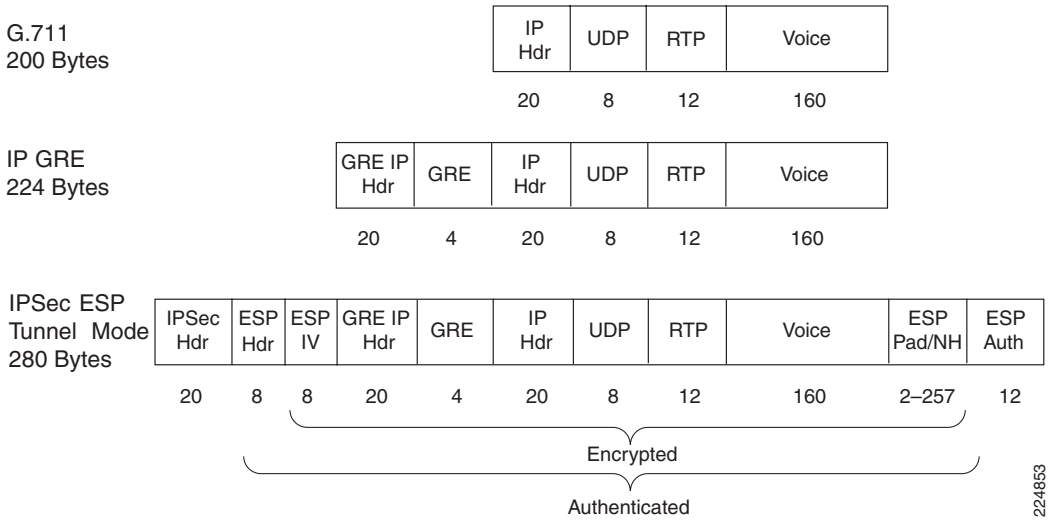
The calculation is as follows:

	200 bytes per packet (G.711 voice)
	24 bytes per packet (IP GRE overhead)
+	56 bytes per packet (IPSec ESP overhead)
	280 bytes per packet
×	8 bits per byte
	2240 bits per packet
×	50 packets per second
	112,000 bits per second

The additional overhead represents a 40 percent increase in the bandwidth required for an encrypted G.711 call.

The 280-byte packet's header, data, and trailer fields for an IPSec tunnel-mode ESP encrypted G.711 call are shown in Figure 6-3.

Figure 6-3 Anatomy of an IPSec-Encrypted G.711 Packet



The Layer 3 data rate for a G.729 call (at 50 pps) is 24 kbps. IP GRE tunnel overhead adds 24 bytes per packet. IPSec ESP adds another 52 bytes. The combined additional overhead increases the rate from 24 kbps (clear voice) to just less than 56 kbps (IPSec ESP tunnel-mode encrypted voice).

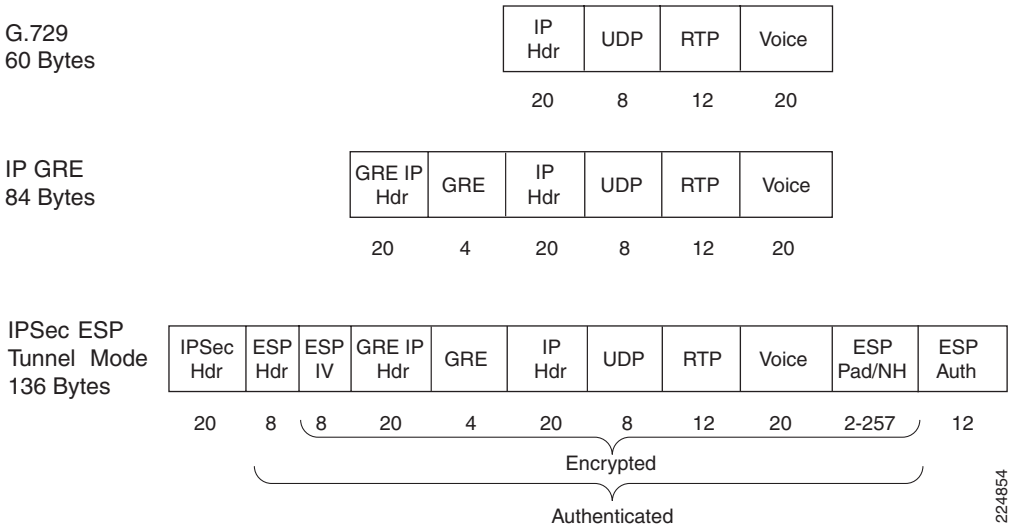
The calculation is as follows:

60 bytes per packet (G.729 voice)
24 bytes per packet (IP GRE overhead)
+ 52 bytes per packet (IPSec ESP overhead)
136 bytes per packet
× 8 bits per byte
1088 bits per packet
× 50 packets per second
54,400 bits per second

The additional overhead represents a 227 percent increase in the bandwidth required for an encrypted G.729 call.

The 136-byte packet's header, data, and trailer fields for an IPSec tunnel-mode ESP encrypted G.729 call are shown in Figure 6-4.

Figure 6-4 Anatomy of an IPSec-Encrypted G.729 Packet



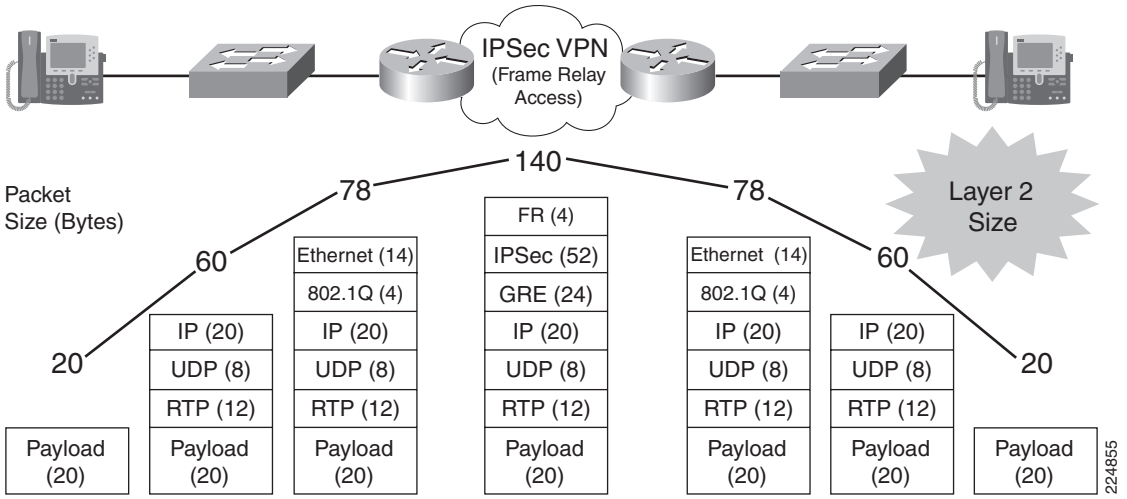
It is important to note that these bandwidth allocations are Layer 3 bandwidth requirements and *do not include* Layer 2 overhead (which is media dependent). Therefore, Layer 2 overhead needs to be added on top of the Layer 3 requirements in provisioning LLQ and CBWFQ bandwidth. This is illustrated in Figure 6-5, where Ethernet overhead (Ethernet plus 802.1Q trunking) and Frame Relay overhead are added and removed from the packet in transit.

Key Layer 2 overhead values are reiterated in Table 6-1.

Table 6-1 Layer 2 Encapsulation Overhead

Layer 2 Encapsulation	Overhead
Ethernet	14 bytes (+ 4 for 802.1Q)
Frame Relay	4 bytes (+ 4 for FRF.12)
MLP	10 bytes (+ 3 for MLP LFI)
ATM	5 bytes per 53-byte cell + cell padding (variable)

Figure 6-5 Packet Size Changes of a G.729 IPSec-Encrypted Packet



Therefore, the calculation, inclusive of Layer 2 overhead, is as follows. This example assumes that a G.729 call will be encrypted over a slow speed (≤768-kbps Frame Relay link), which requires FRF.12 fragmentation and interleaving.

60 bytes per packet (G.729 voice)
24 bytes per packet (IP GRE overhead)
52 bytes per packet (IPSec ESP overhead)
4 bytes per packet (FR overhead)
+ **4 bytes per packet (FRF.12 overhead)**
44 bytes per packet
× 8 bits per byte
1152 bits per packet
× 50 packets per second
57,600 bits per second (rounded up to 58 kbps)

In summary, it is important always to include Layer 2 overhead in accurate bandwidth provisioning for IPSec-encrypted applications.

cRTP and IPSec Incompatibility

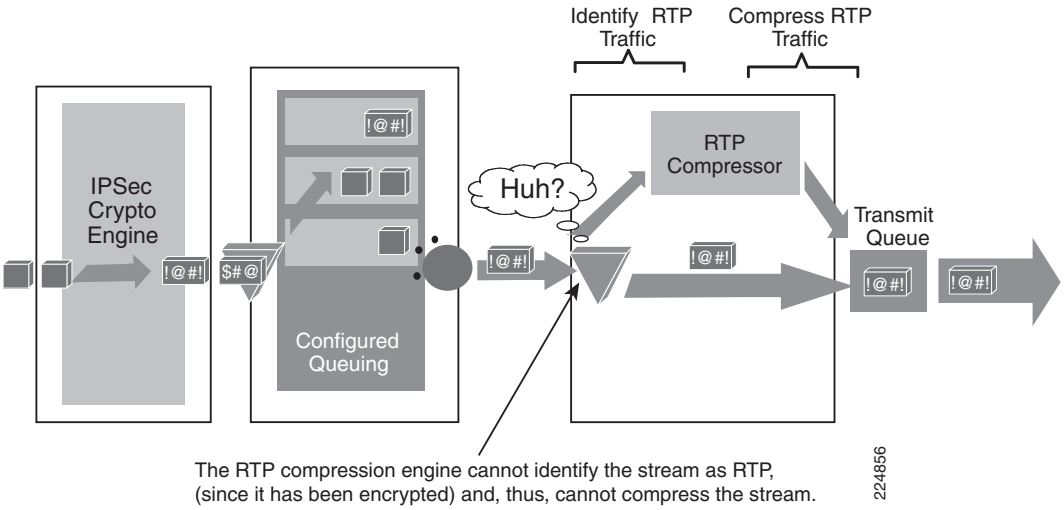
The significant increases in bandwidth required by IPSec encryption lead many administrators to consider the use of IP RTP header compression (cRTP) to offset these increases.

However, one of the caveats of encryption is that key portions of the original IP packet that could be referenced for QoS (and other) purposes are no longer readable. Such is the case with cRTP.

cRTP and IPSec are inherently incompatible standards. The original IP/UDP/RTP header already is encrypted by IPSec by the time the RTP compressor is called upon to perform the compression. Therefore, because cRTP cannot associate the encrypted IP/UDP/RTP packet with a known media stream, compression cannot occur and cRTP bandwidth savings cannot be realized. The encrypted IP/UDP/RTP packet simply bypasses the compression process and continues (uncompressed) to the transmit queue.

This is illustrated in Figure 6-6.

Figure 6-6 IPSec and cRTP Incompatibility



It is important to recognize that cRTP functions on a hop-by-hop basis, whereas IPSec can span multiple intermediate (Layer 3) hops between IPSec endpoints. This distinction further exacerbates incompatibility between the features.

Although developments are under way to address these incompatibilities, at the time of this writing, cRTP cannot be utilized to achieve bandwidth savings in an IPSec VPN environment.

Prefragmentation

A problem arises when a packet is nearly the size of the maximum transmission unit (MTU) of the outbound link of the encrypting router and then is encapsulated with IPSec headers. The resulting packet is likely to exceed the MTU of the outbound link. This causes packet fragmentation after encryption, which makes the decrypting router reassemble in the process path.

Cisco IOS Release 12.2(13)T introduced a new feature: prefragmentation for IPSec VPNs. Prefragmentation increases the decrypting router's performance by enabling it to operate in the high-performance CEF path instead of the process path.

This feature enables an encrypting router to predetermine the encapsulated packet size from information available in transform sets, which are configured as part of the IPSec security association. If it is predetermined that the packet will exceed the MTU of the output interface, the packet is fragmented before encryption. This function avoids process-level reassembly before decryption and helps improve decryption performance and overall IPSec traffic throughput.

Prefragmentation for IPSec VPNs is enabled globally by default for Cisco VPN routers running Cisco IOS Release 12.2(13)T or higher.

Bandwidth Provisioning

Chapter 1, "Quality of Service Design Overview,"presented the 33 Percent LLQ Rule, along with the design rationale behind the recommendation. Furthermore, the rule was expressed as a conservative design recommendation that might not be valid under all constraints. Provisioning for VoIP over IPSec on slow links sometimes poses constraints that might preclude applying the 33 Percent LLQ Rule.

As shown in Table 6-2, the percentage of LLQ required on 64-, 128-, and 256-kbps links for a single encrypted G.729 call exceeds the recommended 33 percent LLQ limit. Enterprises considering such deployments must recognize the impact on data traffic traversing such links when encrypted voice calls were made—specifically, data applications would slow down significantly. If that is considered an acceptable trade-off, not much can be said or done. Otherwise, it is recommended to increase bandwidth to the point that encrypted VoIP calls can be made and still fall within the 33 percent bandwidth recommendation for priority queuing.

When planning the bandwidth required for a branch office, consider the number of concurrent calls traversing the IPSec VPN that the branch is expected to make during peak call periods. This varies based on the job function of the employees located at a branch. For example, an office of software engineers would be expected to make fewer calls than an office of telemarketers. A typical active call ratio may be one active call for every six people (1:6), but this could range from 1:4 or 1:10, depending on the job function of the employees. Given the 512-kbps link from Table 6-2 as an example, with a target of 3 G.729 calls, this link theoretically could support a branch office of between 12 and 30 people. As with all other topologies, call admission control must be administered properly to correspond to the QoS policies deployed within the network infrastructure.

Table 6-2 G.729 Calls by Link Speeds (FRF.12 Is Enabled on Link Speeds ≤768 kbps Only)

Line Rate (kbps)	Maximum Number of G.729 Calls	LLQ Bandwidth (kbps)	LLQ Bandwidth (Percentage)
64 (FRF.12)	1 (58 kbps)	58	91%
128 (FRF.12)	1 (58 kbps)	58	46%
256 (FRF.12)	2 (58 kbps)	116	46%
512 (FRF.12)	3 (58 kbps)	174	34%
768 (FRF.12)	4 (58 kbps)	232	31%
1024	6 (56 kbps)	336	33%
1536	9 (56 kbps)	504	33%
2048	12 (56 kbps)	672	33%

Note Although VoIP has been discussed as a primary example of bandwidth provisioning considerations when deploying QoS over IPSec VPNs, it is important to recognize that VoIP is not the only application that might require such considerations; this exercise needs to be performed for *any* application that is being encrypted.

Unlike VoIP, however, other applications—such as video and data—have varying packet rates and sizes. Therefore, crisp provisioning formulas might not apply. Moderate traffic analysis and best guesses, along with trial-and-error tuning, are usually the only options for factoring bandwidth provisioning increases for non-VoIP applications.

Logical Topologies

Similar to private WANs, but unlike fully meshed MPLS VPNs, IPSec VPNs typically are deployed in a (logical) hub-and-spoke topology.

The QoS implications of hub-and-spoke topologies include that access rates to remote sites need to be constrained and shaped at the hub, to avoid delays and drops within the provider's cloud. Shaping at the IPSec VPN hub is done in a manner similar to that of private WAN NBMA media, such as Frame Relay or ATM. Refer to the Headend VPN Edge QoS Options for Site-to-Site V3PNs section, later in this chapter, and also to Chapter 3, "WAN Aggregator QoS Design."

IPSec VPNs are not limited to hub-and-spoke designs. They also can be deployed in partial-mesh or even fully meshed topologies. In such cases, shaping is recommended on any links where speed mismatches occur (similar to private WAN scenarios).

Another alternative is to deploy IPSec VPNs via Dynamic Multipoint Virtual Private Networks (DMVPN), which establish site-to-site IPSec tunnels as needed and tear them down when they no longer are required. As with the previously discussed logical topologies, shaping is required on DMVPN NBMA links with speed mismatches. Specifically, shapers are required to be created dynamically and applied to *logical* DMVPN tunnels to offset any speed mismatches attributed to *physical* NMBA links. However, as of the time of this writing, no shaping or queuing solution exists to guarantee QoS SLAs over DMVPN topologies (although Cisco IOS solutions currently are being evaluated and are in development).

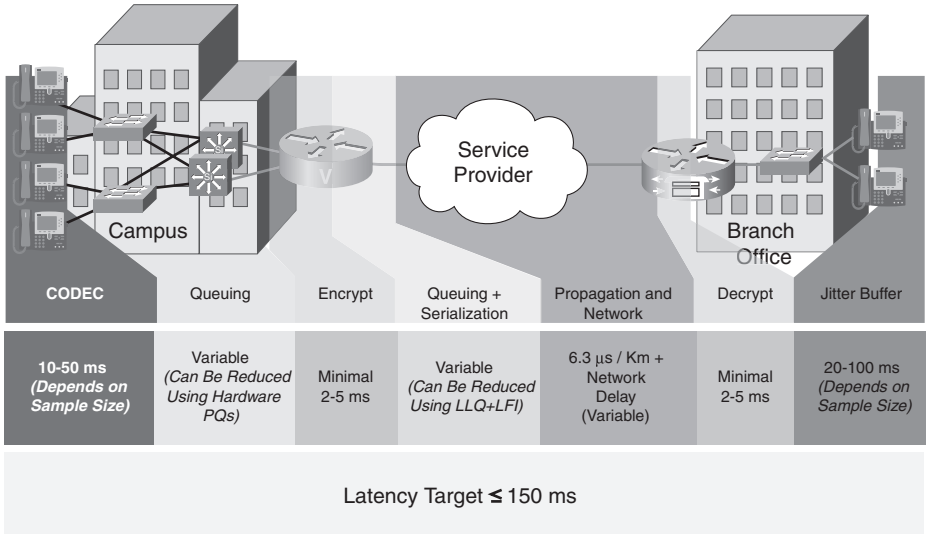
Delay Budget Increases

As previously discussed, the delay budget for a typical IP Telephony implementation includes fixed and variable components. The ITU G.114 specification's target value for one-way delay is 150 ms. In an IPSec VPN deployment, however, two additional delay components must be factored into the overall delay budget:

- Encryption delay at the origination point of the IPSec VPN tunnel
- Decryption delay at the termination point of the IPSec VPN tunnel

Performance and scalability testing results suggest that, in most cases, the additional delay caused by encryption and decryption is approximately 4 to 10 ms (combined). These incremental delays are shown in Figure 6-7.

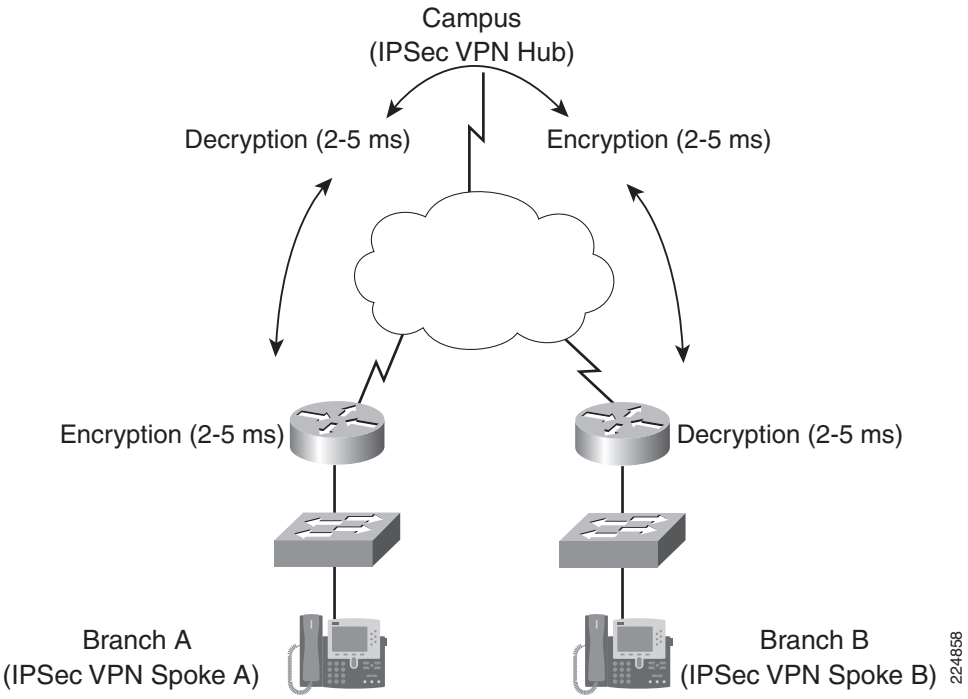
Figure 6-7 IPSec Encryption/Decryption Incremental Delays



A conservative planning estimate would be 10 ms for encryption delay and 10 ms for decryption delay.

This delay might not seem significant for a campus-to-branch call (hub-to-spoke), but the delay might be more relevant in branch-to-branch (spoke-to-spoke) scenarios because encryption and decryption might occur twice (depending on the logical topology of the VPN). This is illustrated in Figure 6-8.

Figure 6-8 IPSec VPN Spoke-to-Spoke Encryption/Decryption Delays



Note Not only do encryption delays need to be factored into spoke-to-spoke IPSec VPN scenarios, but queuing and serialization delays for both legs of the tunnel do as well (as they also would in private WAN spoke-to-spoke scenarios).

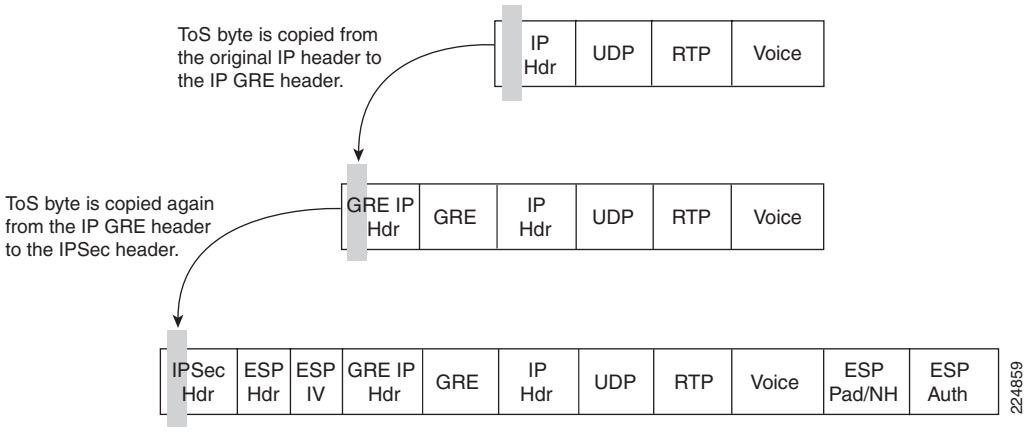
ToS Byte Preservation

For the majority of QoS designs discussed thus far, classification is performed based on DSCP markings in the ToS byte of the IP packet. However, when an IP packet is encrypted through IPSec, the original ToS byte values also are encrypted and, thus, unusable by QoS mechanisms that process the packet (post encryption).

To overcome this predicament, the IPSec protocol standards inherently have provisioned the capability to preserve the ToS byte information of the original IP header by copying it to the IP headers added by the tunneling and encryption process.

As shown in Figure 6-9, the original IP ToS byte values are copied initially to the IP header added by the GRE encapsulation. Then these values are copied again to the IP header added by IPSec encryption.

Figure 6-9 IP ToS Byte Preservation



This process compensates for the fact that the original IP header (including the ToS byte) is actually unreadable (because of encryption) and allows the packet to be processed by (post encryption) QoS mechanisms in the same manner as any other packet.

Additionally, this process underscores the importance of ensuring that the encrypted traffic is marked properly (at Layer 3) before encryption.

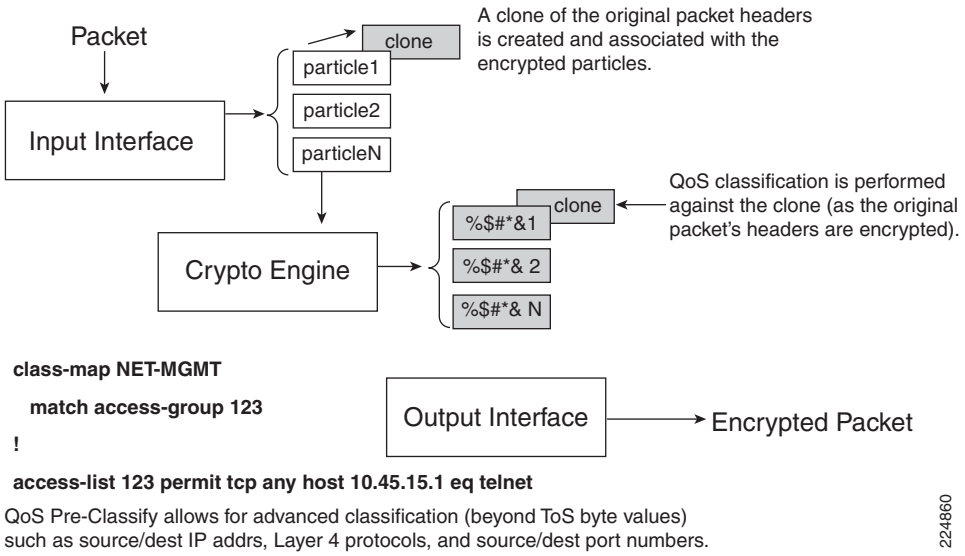
QoS Pre-Classify

The QoS Pre-Classify feature often is confused with ToS byte preservation. QoS Pre-Classify is a Cisco IOS feature that allows for packets to be classified on header parameters other than ToS byte values after encryption.

Because all original packet header fields are encrypted, including source or destination IP addresses, Layer 4 protocol, and source or destination port addresses, post-encryption QoS mechanisms cannot perform classification against criteria specified within any of these fields.

A solution to this constraint is to create a clone of the original packet's headers before encryption. The crypto engine encrypts the original packet, and then the clone is associated with the newly encrypted packet and sent to the output interface. At the output interface, any QoS decisions based on header criteria, except for ToS byte values—which have been preserved—can be performed by matching on any or all of the five access-list tuple values of the clone. In this manner, advanced classification can be administered even on encrypted packets. The process is illustrated in Figure 6-10.

Figure 6-10 QoS Pre-Classify Feature Operation



A key point to remember regarding QoS Pre-Classify is that it is applicable only at the encrypting router's output interface. The fields preserved by QoS Pre-Classify are not available to any routers downstream; the clone never leaves the router performing the encryption, thus ensuring the integrity and security of the IPSec VPN tunnel.

QoS Pre-Classify is supported in all Cisco IOS switching paths and is recommended to be enabled on some platforms even when only the ToS byte values are being used for classification. Testing has shown that when hardware-based encryption cards are combined with QoS, the Cisco IOS Software implementation of the QoS Pre-Classify feature slightly enhances performance, even when matching only on ToS byte values. Furthermore, enabling QoS Pre-Classify by default eliminates the possibility that its configuration will be overlooked if the QoS policy later is changed to include matching on IP addresses, ports, or protocols.

Design recommendations for the QoS Pre-Classify feature can be summarized as follows:

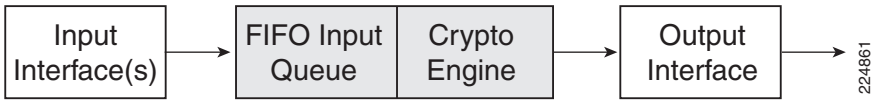
- Enable QoS Pre-Classify on all branch IPSec VPN routers that support the feature.
- Enable QoS Pre-Classify on headend IPSec VPN routers only when both the VPN termination and QoS policies reside on the same device.

Pre-Encryption Queuing

The hardware crypto engine within a Cisco VPN router's chassis can be viewed as an internal interface that processes packets for encryption or decryption.

Before Cisco IOS Release 12.2(13)T, packets to be encrypted were handed off to the crypto engine in a *first-in-first-out* (FIFO) basis. No distinction was made between voice packets and data packets. The FIFO queuing for crypto engines is illustrated in Figure 6-11.

Figure 6-11 FIFO Crypto Engine QoS



Consider a Cisco 2651XM router deployed at a branch site configured with a full-duplex Fast Ethernet interface, a Serial E1 interface (also full duplex), and an AIM-BP encryption accelerator. The Fast Ethernet interface connects to the branch's LAN, and the serial interface connects to the Internet. These factors could limit throughput (causing a bottleneck) in this scenario:

- The clock rate of the slowest interface (in bits per second—in this case, 2 Mbps transmitted or 2 Mbps received over the E1 interface)
- The packet-forwarding rate of the router's main CPU (in packets per second)
- The crypto engine encryption/decryption rate (in packets per second)

The performance characteristics of these items further are influenced by the traffic mix, including the rates and sizes of the IP packets being switched through the router and the configured Cisco IOS switching path (process-switched, fast-switched, or CEF-switched).

Note In most hardware platforms, the packets-per-second capabilities of the router are more important for planning purposes than bits per second switched through the router. For example, if the average packet size of packets switched through the router increases from 128 bytes to 256 bytes, the packet-per-second capabilities of the main CPU are not necessarily cut in half.

The control plane requirements also factor into the CPU's utilization. These requirements are determined by the routing protocol(s) in use, the number of routes in the routing table, the overall network stability, and any redistribution requirements. Management requirements such as NTP and SNMP also add to the CPU tax. Additionally, Cisco IOS HA, QoS, multicast, and security features all consume CPU resources and must be taken into account.

Another factor in the equation is the ratio of packets switched through (and originated by) the router in relation to the number of packets selected by the crypto map's access list for encryption or decryption. If an IP GRE tunnel is being encrypted, this tends to be a large percentage of encrypted packets to total packets; if an IP GRE tunnel is not being encrypted, the ratio could be quite small.

Hardware crypto engines can become congested when their packet-processing capabilities are less than those of the router's main CPU and interface clock speeds. In such a case, the crypto engine becomes a bottleneck, or a congestion point. The crypto engine might be oversubscribed on either a momentary or (worse case) sustained basis. Such internal precrypto congestion could affect the quality of real-time applications, such as VoIP.

Cisco internal testing and evaluation has shown it to be extremely difficult for conditions to arise that cause hardware crypto engine congestion. In nearly all cases, the Cisco VPN router platform's main CPU is exhausted before reaching the limit of the crypto engine's packet-processing capabilities.

Nevertheless, Cisco provides a solution to this potential case of congestion in the rare event that a hardware crypto engine is overwhelmed so that VoIP quality will be preserved. This feature, low-latency queuing (LLQ) for IPSec encryption engines, was introduced in Cisco IOS Release 12.2(13)T.

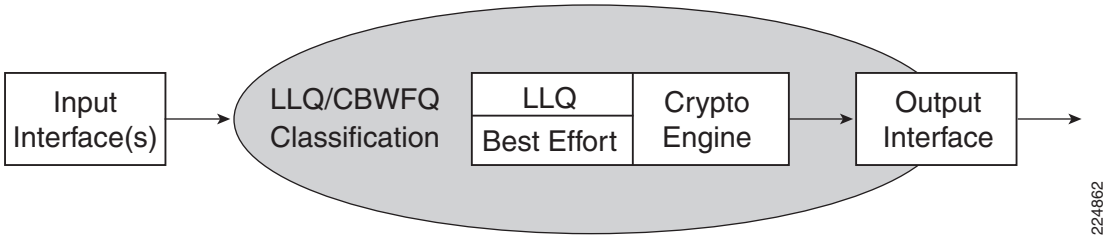
The LLQ for Crypto Engine feature provides a dual-input queuing strategy for packets being sent to the crypto engine:

- A priority or low-latency queue
- A best-effort queue

This feature is targeted at alleviating any effects of momentary or sustained oversubscription of the hardware crypto engines that could result in priority traffic (such as voice and video) experiencing quality issues. This feature is illustrated in Figure 6-12.

Note Because software-based crypto adds unacceptable latency and jitter, there are no plans to incorporate this feature for software crypto. Hardware accelerators for IPSec encryption are highly recommended.

Figure 6-12 LLQ (Dual-FIFO) Crypto Engine QoS



The classification component to segregate traffic between the priority (LLQ) queue and the best-effort queue is based on the MQC service policy on the output interface(s).

No additional configuration is required to enable LLQ for crypto engines; it is enabled internally by the presence of a service policy with an LLQ **priority** command that is applied to an output interface of an IPSec VPN router.

Traffic specified in the service policy to be assigned to the interface's priority queue (LLQ) automatically is sent to the crypto engine's LLQ. Traffic included in any CBWFQ bandwidth classes (including the default class) automatically is assigned to the crypto engine's best-effort queue.

It is possible to configure different service policies, each with different traffic assigned to an LLQ, on different interfaces. For example, perhaps voice is assigned to the LLQ of Serial1/0 and video is assigned to the LLQ of an ATM PVC. Assuming that both voice and video are to be encrypted, the question arises, which type of traffic (voice or video) will be assigned to the crypto engine's LLQ?

Because the crypto engine acts like a single interface inside the VPN router, encrypting and decrypting all outbound and inbound traffic streams for each interface on which crypto is applied, in the case of multiple service policies (on different interfaces) the crypto engine maps *all* interface priority queues (LLQ) to its LLQ and all other queues to its best-effort queue. Therefore, *both* voice and video would be assigned to the crypto engine's LLQ.

In short, the LLQ for Crypto Engine feature ensures that if packets are dropped by momentary or sustained congestion of the crypto engine, the dropped packets will be of appropriately lower priority (not VoIP packets).

Although the feature is enabled by the presence of a service policy with an LLQ **priority** statement, as with interface queuing itself, crypto-engine queuing does not actually engage prioritization through the dual-FIFO queuing strategy until the crypto engine itself experiences congestion.

The LLQ for Crypto Engine feature in Cisco IOS Software is not a prerequisite for deploying QoS for IPSec VPN implementations in a high-quality manner. As indicated previously, internal Cisco evaluations have found it extremely difficult to produce network traffic conditions that resulted in VoIP quality suffering because of congestion of the hardware crypto engine.

In general, the LLQ for Crypto Engine feature offers the most benefit under one of the following conditions:

- When implementing Cisco IOS VPN router platforms that have a relatively high amount of main CPU resources relative to crypto engine resources (these vary depending on the factors outlined earlier in this discussion).
- When the network experiences a periodic or sustained burst of large packets (for example, for video applications).

To summarize, high-quality IPSec VPN deployments are possible today without the LLQ for Crypto Engine feature in Cisco IOS software. The addition of this feature in Cisco IOS software further ensures that high-priority applications, such as voice and video, can operate in a high-quality manner even under harsh network conditions.

Anti-Replay Implications

IPSec offers inherent message-integrity mechanisms to provide a means to identify whether an individual packet is being replayed by an interceptor or hacker. This concept is called connectionless integrity. IPSec also provides for a partial sequence integrity, preventing the arrival of duplicate packets. These concepts are outlined in RFC 2401, "Security Architecture for the Internet Protocol."

When ESP authentication (**esp-sha-hmac**) is configured in an IPSec transform set, for each security association, the receiving IPSec peer verifies that packets are received only once. Because two IPSec peers can send millions of packets, a 64-packet sliding window is implemented to bound the amount of memory required to tally the receipt of a peer's packets. Packets can arrive out of order, but they must be received within the scope of the window to be accepted. If they arrive too late (outside the window), they are dropped.

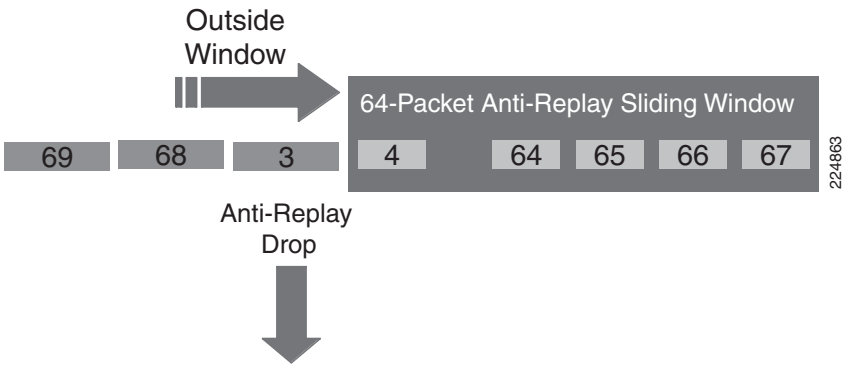
The operation of the Anti-Replay window protocol is as follows:

1. The sender assigns a unique sequence number (per security association) to encrypted packets.
2. The receiver maintains a 64-packet sliding window, the right edge of which includes the highest sequence number received. In addition, a Boolean variable is maintained to indicate whether each packet in the current window was received.
3. The receiver evaluates the received packet's sequence number:
   - If a received packet's sequence number falls within the window and was not received previously, the packet is accepted and marked as received.
   - If the received packet's sequence number falls within the window and previously was received, the packet is dropped and the replay error counter is incremented.
   - If the received packet's sequence number is greater than the highest sequence in the window, the packet is accepted and marked as received, and the sliding window is moved "to the right."
   - If the received packet's sequence number is less than the lowest sequence in the window, the packet is dropped and the replay error counter is incremented.

In a converged IPSec VPN implementation with QoS enabled, lower-priority packets are delayed so that higher-priority packets receive preferential treatment. This has the unfortunate side effect of reordering the packets to be out of sequence from an IPSec Anti-Replay sequence number perspective. Therefore, there is a concern that through the normal QoS prioritization process, the receiver might drop packets as Anti-Replay errors, when, in fact, they are legitimately sent or received packets.

Figure 6-13 provides a visualization of the process. In this example, voice packets 4 through 67 have been received, and data packet 3 was delayed and transmitted following voice packet 68. When the Anti-Replay logic is called to process packet 3, it is dropped because it is outside the left edge of the sliding window. Packets can be received out of order, but they must fall within the window to be accepted.

Figure 6-13 Anti-Replay Operation



Anti-Replay drops can be eliminated in a pure IPSec tunnel design (no encrypted IP GRE tunnel) by creating separate security associations for voice and data; voice and data packets must match a separate line in the access list referenced by the crypto map. This is implemented easily if the IP phones are addressed by network addresses (such as private RFC 1918 addresses) separate from the workstations.

However, if IPSec tunnel mode (with an encrypted IP GRE tunnel) is used for a converged network of voice and data, Anti-Replay drops impact data packets instead of voice packets because the QoS policies prioritize voice over data.

Consider the effect of packet loss on a TCP-based application: TCP is connection oriented and incorporates a flow-control mechanism within the protocol. The TCP application cannot see why a packet was dropped. A packet dropped by a service policy on a congested output interface is no different to the application than a packet lost by an Anti-Replay drop. From a *network perspective*, however, it would be more efficient to drop the packet *before* sending it over the WAN link (where bandwidth is the most expensive, only to have it dropped by the Anti-Replay mechanism on the receiving IPSec VPN router), but the location or nature of the packet loss is immaterial to the TCP driver.

Anti-Replay drops of data traffic flows are usually in the order of 1 percent to 1.5 percent on IPSec VPN links that experience sustained congestion and have queuing engaged (without any additional policy tuning).

Output drops on the output WAN interface, however, tend to be few, if any, and certainly are far fewer than those dropped by Anti-Replay. This is because Anti-Replay triggers packet drops more aggressively than the output service policy, which is a function of the size of the output queues and the number of defined classes.

By default, each CBWFQ class receives a queue with a length of 64 packets. This can be verified with the **show policy interface** verification command. Meanwhile, the receiving IPSec peer has a *single* 64-packet Anti-Replay window (per IPSec Security Association) with which to process all packets from all LLQ and CBWFQ bandwidth classes.

So, it stands to reason that the Anti-Replay mechanism on the receiving VPN router will be more aggressive at dropping packets delayed by QoS mechanisms preferential to VoIP than the service policy at the sending router. This is because of the size mismatch of the queue depth on the sender's output interface (multiple queues of 64 packets each) compared to the width of the receiver's Anti-Replay window (a single sliding window of 64 packets per SA). As more bandwidth classes are defined in the policy map, this mismatch increases. As mentioned, this is an inefficient use of expensive WAN/VPN bandwidth because packets are transmitted only to be dropped before decryption.

The default value of 64 packets per CBWFQ queue is designed to absorb bursts of data traffic and delay rather than drop those packets. This is optimal behavior in a non-IPSec–enabled network.

When IPSec authentication is configured (**esp-sha-hmac**) in the network, the scenario can be improved by reducing the queue limit (max threshold) of the bandwidth classes of the sender's output service policy so that it becomes more aggressive at dropping packets than buffering or delaying them. Extensive lab testing has shown that such queue-limit tuning can reduce the number of Anti-Replay drops from 1 percent to less than a tenth of percent (< 0.1 percent). This is because decreasing the service policy queue limits causes the sender's output service policy to become more aggressive in dropping instead of significantly delaying packets (which occurs with large queue depths). This, in turn, decreases the number of Anti-Replay drops.

As a rule of thumb, the queue limits should be reduced in descending order of application priority. For example, the queue limit for Scavenger should be set lower than the queue limit for a Transactional Data class, as is shown later in Example 6-3 through Example 6-6.

Note In many networks, the default queue limits and IPSec Anti-Replay performance are considered acceptable. A modification of queue-limit values entails side effects on the QoS service policy and related CPU performance. When queue limits are tuned, a careful eye should be kept on CPU levels.

Note As of IOS 12.3(14)T another potential remedy to QoS/IPSec Anti-Replay issues became available with the ability to expand or disable the IPSec Anti-Replay window. However it should be kept in mind that the IETF IPSec standards define the Anti-Replay window-sizes to be 64 packets and as such, this would not be a standards-compliant solution. On the other hand, the presented solution of tuning of the QoS queue-limits is fully standards-compliant.

The documentation for IOS feature to allow the expanding or disabling of IPSec Anti-Replay is at: http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123newft/123t/123t_14/gt_iarwe.htm

Control Plane Provisioning

As discussed in Chapter 3, "WAN Aggregator QoS Design," Cisco IOS Software has an internal mechanism for protecting control traffic, such as routing updates, called PAK_priority.

PAK_priority marks routing protocols to DSCP CS6, but it currently does not protect IPSec control protocols, such as ISAKMP (UDP port 500). Therefore, it is recommended to provision an explicit CBWFQ bandwidth class for control plane traffic, including ISAKMP, as shown in Example 6-2.

Example 6-2 Protecting IPSec Control Plan Traffic

```
!
class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6
  match access-group name IKE                ! References ISAKMP ACL
!
!
policy-map V3PN
 ...
  class INTERNETWORK-CONTROL
   bandwidth percent 5                       ! Control Plane provisioning
...
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp      ! ISAKMP ACL
```

```
!
```

Site-to-Site V3PN QoS Designs

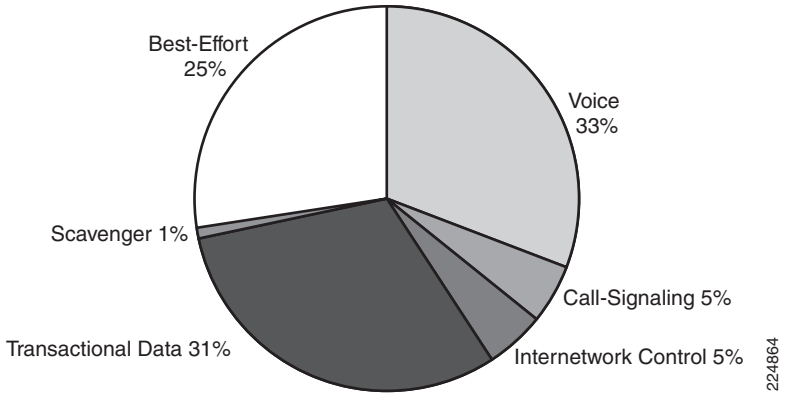
As with WAN and MPLS VPN QoS models, site-to-site V3PN QoS models can range from a basic number of classes (in this case, the minimum number of recommended classes is 6) to a complex QoS Baseline model (11 classes). Each enterprise must determine present needs and comfort level of QoS complexity, along with future needs, to more easily migrate to progressively more complex QoS models, as required.

Six-Class Site-to-Site V3PN Model

The six-class V3PN model technically should be referred to as V2PN because it includes provisioning for only voice (not for video) over an IPSec VPN. Voice is protected explicitly with LLQ; call signaling and control plane traffic also explicitly are protected through CBWFQ classes. This model includes a preferential data class (Transactional Data) and a deferential class (Scavenger, which is squelched to the minimum configurable amount: 1 percent). If the queue limits are tuned to minimize Anti-Replay drops, the queue limit for Transactional Data should be set higher than the queue limit for class default.

An example six-class V3PN model, which is suitable for link speeds up to and including T1/E1 speeds, is illustrated in Figure 6-14 and detailed in Example 6-3.

Figure 6-14 Six-Class Site-to-Site V3PN Model



Example 6-3 Six-Class Site-to-Site V3PN Model Configuration Example

```
!
class-map match-all VOICE
  match ip dscp ef                              ! VoIP
 class-map match-any CALL-SIGNALING
  match ip dscp cs3                             ! New Call-Signaling
  match ip dscp af31                            ! Old Call-Signaling
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                             ! IP Routing
  match access-group name IKE                   ! References ISAKMP ACL
 class-map match-all TRANSACTIONAL-DATA
  match ip dscp af21 af22                       ! Transactional-Data
 class-map match-all SCAVENGER
  match ip dscp cs1                             ! Scavenger
```

```
!
!
 policy-map SIX-CLASS-V3PN-EDGE
  class VOICE
    priority percent 33                          ! VoIP gets 33% LLQ
  class CALL-SIGNALING
    bandwidth percent 5                          ! Call-Signaling provisioning
  class INTERNETWORK-CONTROL
    bandwidth percent 5                          ! Control Plane provisioning
  class TRANSACTIONAL-DATA
    bandwidth percent 31                         ! Transactional-Data provisioning
    queue-limit 20                               ! Optional: Anti-Replay tuning
  class SCAVENGER
    bandwidth percent 1                          ! Scavenger class is throttled
    queue-limit 1                                ! Optional: Anti-Replay tuning
  class class-default
    bandwidth percent 25                         ! Best Effort needs BW guarantee
    queue-limit 16                               ! Optional: Anti-Replay Tuning
!
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp          ! ISAKMP ACL
!
```

Note Currently, only distributed platforms support the **queue-limit** command in conjunction with WRED commands. However, this command combination will be available on nondistributed platforms with the release of Consistent QoS Behavior, as discussed in Chapter 3, "WAN Aggregator QoS Design."

Verification commands:

- **show policy**
- **show policy interface**

Eight-Class Site-to-Site V3PN Model

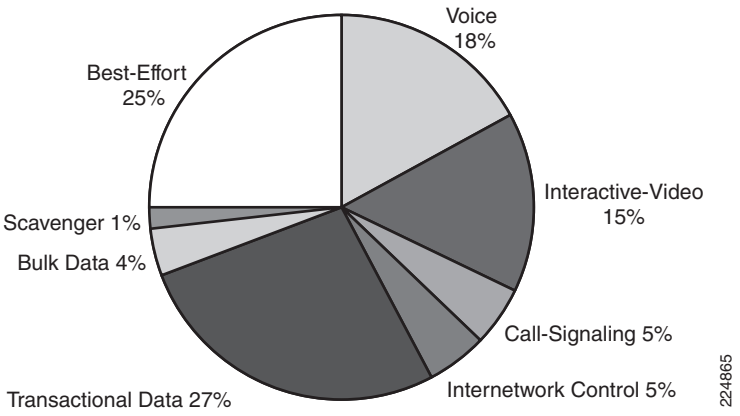
With the addition of a class for Interactive-Video, this model more accurately lives up to its V3PN name. Interactive-Video (as with VoIP) must be provisioned adequately to include IPSec encryption overhead, but (unlike VoIP) there are no clean formulas for calculating the required incremental bandwidth. This is because video packet sizes and packet rates vary significantly and are largely a function of the degree of motion within the video images being transmitted.

On an unencrypted topology, the guideline for provisioning for Interactive-Video is to overprovision the LLQ by 20 percent. Although this conservative guideline holds true in most videoconferencing scenarios, which usually consist of relatively minor motion (meetings, conferences, talking heads, and so on), in some cases this rule proves inadequate over an IPSec VPN. In such cases, the LLQ might need to be provisioned to the stream's rate plus 25 percent or more. The need to increase the bandwidth for Interactive-Video's LLQ will be apparent if drops (in excess of 1 percent) appear under this class when using the verification command **show policy interface**.

The other class added to this model is a separate class for Bulk Data applications (which will be constrained if congestion occurs, to prevent long sessions of TCP-based flows from dominating the link). Notice that the queue limit of the Bulk Data class has been reduced to below the queue limit of the Best-Effort class. This is because of the relative ranking of application priority: During periods of congestion, the Bulk Data class (large TCP-based file operations that operate mainly in the background) is prevented from dominating bandwidth away from the Best-Effort class (as a whole).

This policy is suitable for link speeds of 3 Mbps or higher. However, in an IPSec environment, it is good to remember that load-sharing GRE tunnels over multiple physical interfaces exacerbate Anti-Replay drops. Whenever possible, a single physical interface should be used to achieve these higher speeds. Another consideration to bear in mind is that higher-end platforms, such as the 2691 and 3700- or 7200-series routers, are required to perform crypto at higher speeds. The Eight-Class Site-to-Site V3PN model is illustrated in Figure 6-15 and detailed in Example 6-4.

Figure 6-15 Eight-Class Site-to-Site V3PN Model



Example 6-4 Eight-Class Site-to-Site V3PN Model Configuration Example

```
!
 class-map match-all VOICE
  match ip dscp ef                          ! VoIP
 class-map match-all INTERACTIVE-VIDEO
  match ip dscp af41 af42                   ! Interactive-Video
 class-map match-any CALL-SIGNALING
  match ip dscp cs3                         ! Old Call-Signaling
  match ip dscp af31                        ! New Call-Signaling
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                         ! IP Routing
  match access-group name IKE               ! References ISAKMP ACL
 class-map match-all TRANSACTIONAL-DATA
  match ip dscp af21 af22                   ! Transactional-Data
 class-map match-all BULK-DATA
  match ip dscp af11 af12                   ! Bulk Data
 class-map match-all SCAVENGER
  match ip dscp cs1                         ! Scavenger
!
 policy-map EIGHT-CLASS-V3PN-EDGE
  class VOICE
    priority percent 18               ! VoIP gets 18% LLQ
  class INTERACTIVE-VIDEO
    priority percent 15               ! IP/VC gets 15% LLQ
class CALL-SIGNALING
    bandwidth percent 5               ! Call-Signaling provisioning
 class INTERNETWORK-CONTROL
    bandwidth percent 5               ! Control Plane provisioning
  class TRANSACTIONAL-DATA
    bandwidth percent 27              ! Transactional-Data provisioning
    queue-limit 18                    ! Optional: Anti-Replay tuning
  class BULK-DATA
    bandwidth percent 4               ! Bulk-Data provisioning
    queue-limit 3                     ! Optional: Anti-Replay tuning
```

```
  class SCAVENGER
    bandwidth percent 1               ! Scavenger class is throttled
    queue-limit 1                     ! Optional: Anti-Replay tuning
  class class-default
    bandwidth percent 25              ! Best Effort needs BW guarantee
    queue-limit 16                    ! Optional: Anti-Replay Tuning
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp    ! ISAKMP ACL
!
```

Verification commands:

- **show policy**
- **show policy interface**

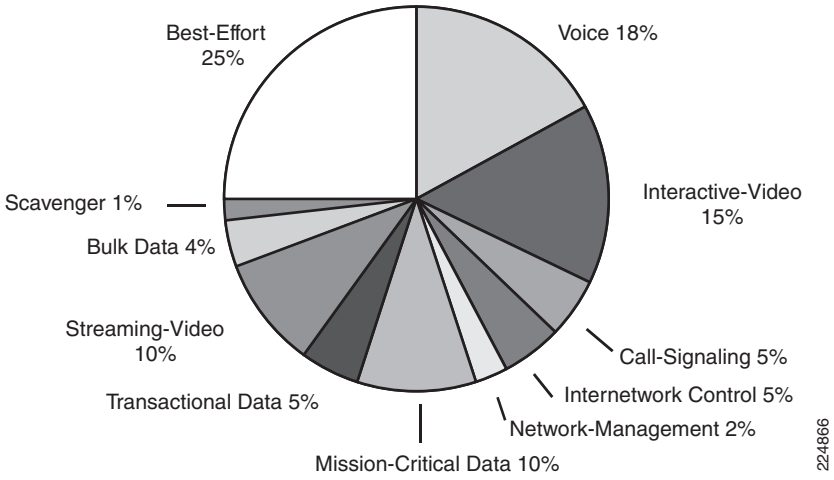
QoS Baseline (11-Class) Site-to-Site V3PN Model

Building on the previous model, three new classes are added: Network-Management, Mission-Critical Data, and Streaming-Video.

This model also is suitable only for high-speed (3 Mbps and above) links, with the same caveats regarding multiple physical links aggravating Anti-Replay drops and the requirement of using newer platforms to perform crypto at these speeds. The queue limits have been tuned to reflect relative application priority.

The QoS Baseline V3PN model, suitable for 3-Mbps link speeds and higher, is illustrated in Figure 6-16 and detailed in Example 6-5.

Figure 6-16 QoS Baseline (Eleven-Class) Site-to-Site V3PN Model



Example 6-5 QoS Baseline (Eleven-Class) Site-to-Site V3PN Model Configuration Example

```
!
 class-map match-all VOICE
  match ip dscp ef                          ! VoIP
 class-map match-all INTERACTIVE-VIDEO
  match ip dscp af41 af42                   ! Interactive-Video
 class-map match-any CALL-SIGNALING
```

```
  match ip dscp cs3                          ! Old Call-Signaling
  match ip dscp af31                         ! New Call-Signaling
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                          ! IP Routing
  match access-group name IKE                ! References ISAKMP ACL
 class-map match-all NET-MGMT
  match ip dscp cs2                          ! Network Management
 class-map match-all MISSION-CRITICAL-DATA
  match ip dscp 25                           ! Interim MC Data
 class-map match-all TRANSACTIONAL-DATA
  match ip dscp af21 af22                    ! Transactional Data
 class-map match-all STREAMING-VIDEO
  match ip dscp cs4                          ! Streaming Video
 class-map match-all BULK-DATA
  match ip dscp af11 af12                    ! Bulk Data
 class-map match-all SCAVENGER
  match ip dscp cs1                          ! Scavenger
!
!
policy-map QOSBASELINE-V3PN-EDGE
  class VOICE
    priority percent 18              ! VoIP gets 18% LLQ
  class INTERACTIVE-VIDEO
    priority percent 15              ! IP/VC gets 15% LLQ
  class CALL-SIGNALING
    bandwidth percent 5              ! Call-Signaling provisioning
  class INTERNETWORK-CONTROL
    bandwidth percent 5              ! Control Plane provisioning
  class NET-MGMT
    bandwidth percent 2              ! Network Management provisioning
  class MISSION-CRITICAL-DATA
    bandwidth percent 10             ! Mission-Critical Data provisioning
    queue-limit 6                    ! Optional: Anti-Replay tuning
  class TRANSACTIONAL-DATA
    bandwidth percent 5              ! Transactional-Data provisioning
    queue-limit 3                    ! Optional: Anti-Replay tuning
  class STREAMING-VIDEO
    bandwidth percent 10             ! Streaming-Video provisioning
    queue-limit 6                    ! Optional: Anti-Replay tuning
  class BULK-DATA
    bandwidth percent 4              ! Bulk-Data provisioning
    queue-limit 3                    ! Optional: Anti-Replay tuning
  class SCAVENGER
    bandwidth percent 1              ! Scavenger throttling
    queue-limit 1                    ! Optional: Anti-Replay tuning
class class-default
    bandwidth percent 25             ! Best Effort needs BW guarantee
    queue-limit 16                   ! Optional: Anti-Replay tuning
!
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp     ! ISAKMP ACL
!
```

Verification commands:

- **show policy**
- **show policy interface**

At remote sites, these policies can be applied to the physical interfaces connecting them to the service provider (provided that the SLAs are for line rates—otherwise, shapers must be used). For example, if the service provider guarantees a full T1 rate to a remote site and the access medium is Frame Relay, the

service policy can be applied directly to the main Frame Relay interface. If the service provider guarantees only ≤768 kbps across the same link, Frame Relay traffic shaping (either legacy or class-based FRTS) combined with FRF.12 must be used at the remote site.

For the central site(s) WAN aggregators, however, unique considerations exist. These are discussed in the next section.

Headend VPN Edge QoS Options for Site-to-Site V3PNs

IPSec V3PNs can be configured in various ways at the central sites. Some enterprises simply overlay V3PNs on top of their existing private WANs; others subscribe to service providers that offer classes of service within their clouds. Many enterprises deploy VPN headends behind WAN aggregation routers to distribute CPU loads, while some perform encryption and QoS on the same box. Each of these options presents considerations on how V3PN policies optimally are applied on WAN aggregation routers.

For enterprises that have overlaid IPSec VPNs on top of their private WAN topologies, the V3PN policies should be applied to the leased lines or Frame Relay/ATM PVCs, as described in Chapter 3, "WAN Aggregator QoS Design."

For enterprises that are subscribing to service providers that offer PE-to-CE QoS classes (including enterprises that are deploying IPSec VPNs *over* MPLS VPNs), V3PN policies need to be applied on the CE-to-PE links (complete with any re-marking that the service provider requires to map into these service provider classes), as described in Chapter 6, "IPSec VPN QoS Design."

For enterprises that are subscribing to service providers that do not offer explicit QoS (beyond an SLA) within their cloud and are using VPN headends behind WAN aggregation routers, the V3PN service policies would be applied to the WAN aggregator's (CE-to-PE) physical links. This prioritizes packets (by applications) and relies on the service provider's SLA to ensure that the delivery largely reflects the priority of the packets as they are handed off to the service provider. Such a configuration would require only a single QoS policy for a WAN aggregator (albeit, on a high-speed interface), but at the same time, it would involve an increased dependence on the service provider's SLA to deliver the desired QoS service levels.

When the VPN headend routers have adequate CPU cycles to perform QoS, another option exists: hierarchical MQC policies that shape and queue (within the shaped rate) and are applied on a per-tunnel basis. A sample of per-tunnel hierarchal QoS policies is shown in Example 6-6. As with previous shaping design recommendations, the shaper is configured to shape to 95 percent of the remote site's line rate.

Note It is critical to keep an eye on CPU levels when IPSec VPN encryption *and* per-tunnel QoS policies are applied on the same router. CPU levels, in general, should not exceed 75 percent during normal operating conditions. Configuring hierarchical shaping and queuing policies on a per-tunnel (per-SA) basis to a large number of sites could be very CPU intensive, especially when such sites experience periods of sustained congestion.

Example 6-6 Per-Tunnel Hierarchical Shaping and Queuing MQC Policy for VPN Headends/WAN Aggregators

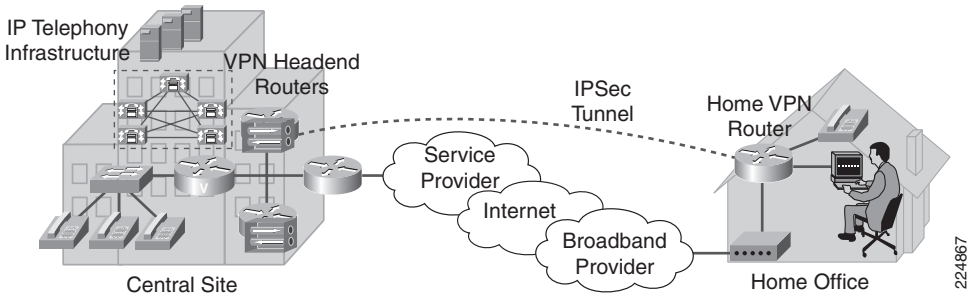
```
!
policy-map SHAPING-T1-TUNNEL
  class class-default
   shape average 1460000 14600 0       ! Shaped to 95% of T1 line-rate
   service-policy V3PN-EDGE            ! Nested queuing policy
!
```

```
!
interface Tunnel120
 description VPN Pipe to V3PN Site#120 (T1 Link)
 bandwidth 1536
 ip address 10.10.120.1 255.255.255.252
 ip mtu 1420
 service-policy output SHAPING-T1-TUNNEL     ! Policy applied to tunnel int
 qos pre-classify                            ! Performance recommandation
 tunnel source 192.168.1.1
 tunnel destination 192.168.2.2
 crypto map VPN
!
```

Teleworker V3PN QoS Considerations

Organizations constantly are striving to reduce costs and improve productivity and employee retention. Teleworker solutions address these organizational objectives by giving employees the ability to work from home with compatible quality, functionality, performance, convenience, and security, as compared to working in an office location. Teleworker solutions that provide such functionality have been branded "enterprise class" by Cisco Marketing and are illustrated in Figure 6-17.

Figure 6-17 Enterprise Teleworker Design



Telecommuter solutions include these main benefits:

- **Increased productivity**—On average, employees spend 60 percent of their time or less at their desks, yet this is where the bulk of investment is made in providing access to corporate applications. Providing similar services at an employee's residence, for a relatively minor investment, significantly can increase productivity gains.
- **Business resilience**—Employees can be displaced from their normal workplace by natural events (such as winter storms, hurricanes, or earthquakes), health alerts (such as SARS), man-made events (such as travel restrictions or traffic conditions), or simply family-related events, such as sick children or home repairs. These disruptions significantly can impact an organization's processes. Providing employees with central site–equivalent access to applications and services in geographically dispersed locations (such as home offices) creates a built-in back-up plan to keep business processes functioning in unforeseen circumstances.
- **Cost savings**—A traditional remote worker setup involves toll charges for dial-up and additional phone lines. Integrating services into a single, broadband-based connection can eliminate these charges while delivering superior overall connectivity performance.
- **Security**—Demands for access to enterprise applications outside the campus are stretching the limits of security policies. Teleworking over IPSec VPNs offers inherent security provided by encryption of all traffic, including data, voice, and video. Also critical is integrating firewall and

intrusion-detection capabilities, along with finding ways to easily accommodate both corporate and personal users who share a single broadband connection (the "spouse-and-children" concern, which will be discussed shortly).

- **Employee recruitment and retention**—In the past, enterprises recruited employees in the locations where corporate offices were located. Today enterprises need the flexibility of hiring skilled employees wherever the skills exist and need to integrate remote workers into geographically dispersed teams with access to equivalent corporate applications.

Although QoS designs for IPSec V3PN teleworker scenarios share many of the same concerns of site-to-site V3PN scenarios, additional specific considerations relating to teleworker deployment models and broadband access technologies need to be taken into account.

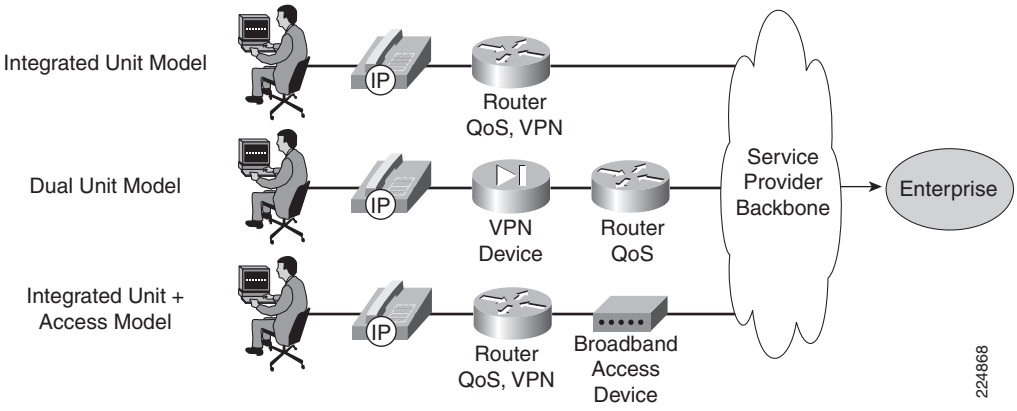
Teleworker Deployment Models

Enterprise teleworker deployment models must provide the following services:

- **Basic services**—These include NAT, DHCP, IP routing, and multiple Ethernet connections for home office devices and broadband connection (attachment to the WAN circuit cable, DSL, ISDN, or wireless network).
- **QoS services**—These include classification, prioritization, and, in some cases, shaping of traffic.
- **VPN/security services**—These include encryption of traffic to the main site and firewall functionality for the home office.

Given these requirements, there are three main deployment models for Enterprise teleworker V3PN solutions, as illustrated in Figure 6-18. These include the Integrated Unit Model, the Dual-Unit Model, and the Integrated Unit + Access Model.

Figure 6-18 Enterprise Teleworker Deployment Models



Integrated Unit Model

In the Integrated Unit Model, a single device (such as a Cisco 837 or 1700-series router) provides basic services, QoS services, and VPN/security services. Furthermore, the device connects directly to the broadband media.

Note The Cisco routers used in these scenarios require an IP/FW/PLUS 3DES Cisco IOS Software feature set. This feature set would include support for LLQ/CBWFQ with class-based hierarchical traffic shaping, and also support for Easy VPN (EZVPN), PKI, IDS, AES, URL filtering, and EIGRP.

Advantages include the following:

- Single-device deployment and management
- Adaptability for service provider fully managed services (transport, QoS, IP Telephony application)
- Potential cost savings

Disadvantages include the following:

- Availability of a single device at an appropriate cost with the features and performance required
- No single unit for some broadband access circuit types

The Integrated Model is a preferred model for service providers offering a fully managed V3PN teleworker service.

From a QoS design perspective, this model is highly similar to a site-to-site V3PN, except that it interfaces with a broadband media instead of a traditional WAN access media.

Dual-Unit Model

In this model, one device (such as a Cisco 831, 837, or 1700-series router) provides basic services and QoS, and a second device (a Cisco router or a PIX 501 or even a VPN 3002) performs security/VPN services.

Advantages include the following:

- **Granularity of managed services**—Service providers can manage broadband access while enterprises manage VPN/security and private network addressing because these are two different units.
- **Media independence**—Because the VPN/security device is separate from the router connecting to the broadband circuit, the same VPN device can be used for cable, DSL, wireless, and ISDN by changing the router model or module in the router. This is especially valuable if one enterprise must support teleworkers with different broadband circuit types (such as DSL, cable, and ISDN).

Disadvantages include the following:

- Two units packaged for deployment
- Ongoing management of two devices
- The cost for two devices

From a QoS design perspective, this model is no different from the previous (Integrated Unit) model.

Integrated Unit + Access Model

In this third model, a single router (such as a Cisco 831 or 1700-series router) provides basic services, QoS services, and VPN/security services. However, the router does not connect directly to the broadband access media; it connects (through Ethernet) to a broadband access device (such as a DSL or cable modem).

Advantages include the following:

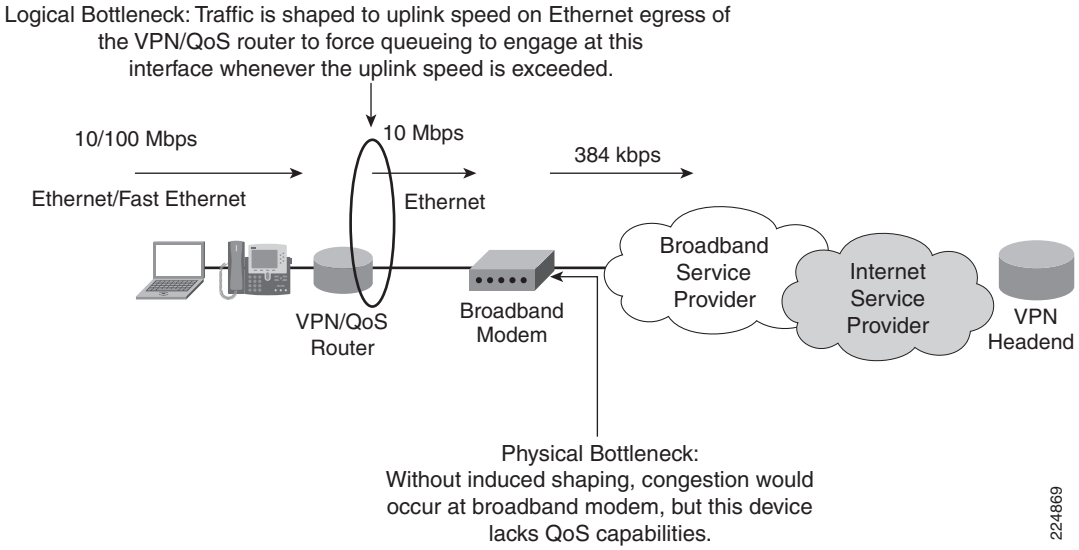
- Cost savings realized by using existing broadband-access devices

- Simplified provisioning because one size fits all, regardless of broadband access media
- Solution support even when no router interface for a specific broadband circuit type is available

Disadvantages include the following:

- Increased troubleshooting complexity because most broadband-access devices (modem) are not intelligent and, therefore, cannot be queried, managed, or controlled.
- Additional QoS complexity because, in this model, the router does not control the broadband circuit and, thus, must perform hierarchical shaping, as shown in Figure 6-19.

Figure 6-19 Hierarchical QoS Requirements for Integrated Unit + Access Teleworker Deployment Model



Because of the media-specific encapsulation overhead requirements (discussed in the following sections), it is recommended to shape to 95 percent of broadband link for cable and 70 percent of the uplink rate for DSL.

A hierarchical shaping policy that forces queuing to engage for a 384-kbps cable broadband connection is shown in Example 6-7.

Example 6-7 Hierarchical Shaping and Queuing MQC Policy for a 384-kbps Cable Connection

```
!
policy-map SHAPE-384-CABLE
 class class-default
   shape average 364800 3640          ! Shapes to 95% of 384 kbps cable link
   service-policy V3PN-TELEWORKER     ! Nested V3PN Teleworker queuing policy
!
...
!
interface Ethernet0
 service-policy output SHAPE-384-CABLE     ! Shaper applied to LAN interface
!
```

The Integrated Unit + Access Model is a preferred model for enterprise V3PN teleworker deployments because it completely abstracts any service provider– or access media–specific requirements (a one-size-fits-all solution).

Broadband-Access Technologies

In North America, there are four main broadband-access technology options for telecommuter scenarios: DSL, cable, ISDN, and wireless. DSL and cable are by far the dominant broadband technologies.

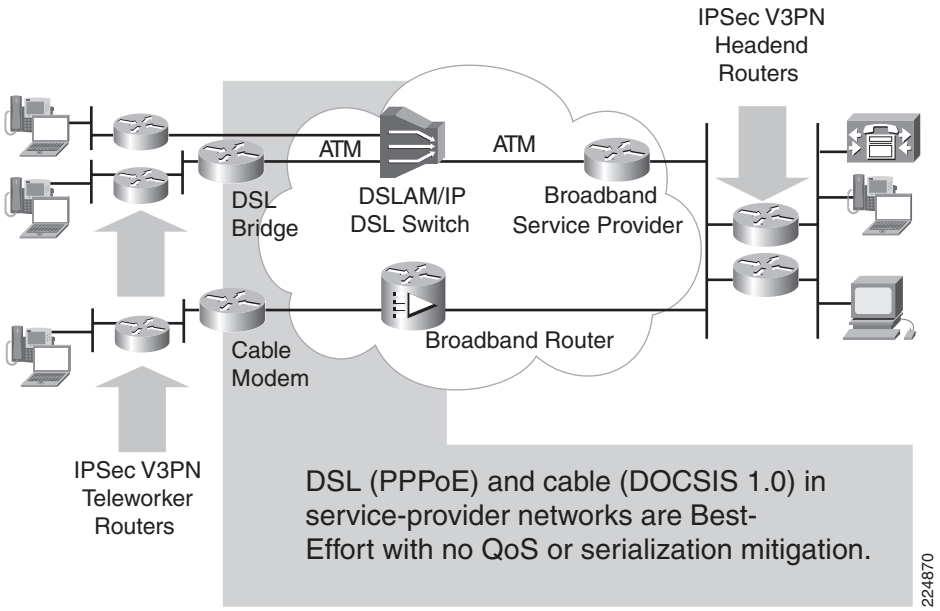
Because of per-minute costs, ISDN is used considerably less; however, ISDN flat rate is becoming available and will make ISDN a good option for areas where DSL or cable is not available. Last-mile wireless is a new option that is being piloted in certain areas to determine viability.

The minimum recommended broadband data rate for most deployments is 160 kbps (uplink)/860 kbps (downlink). Data rates below this speed require more troubleshooting by support staff and are less likely to provide acceptable voice quality. The recommended data rate for V3PN teleworker deployments is 256 kbps (uplink)/1.4 Mbps (downlink) or higher rates. Although V3PN can be deployed at rates less than 160 kbps/860 kbps, generally the voice quality at that service level is in the cell phone quality range, and support costs are higher.

Because QoS design for ISDN was discussed in Chapter 3, "WAN Aggregator QoS Design," and because wireless as a last-mile broadband technology has yet to gain wide deployment, this section focuses only on DSL and cable broadband technologies.

DSL and cable topologies are illustrated in Figure 6-20. Cable is a shared medium between the teleworker's cable modem and the broadband provider's cable headend router. DSL is a dedicated circuit between the teleworker's DSL modem (bridge) and the DSL Access Multiplexer (DSLAM). Both cable and DSL offerings utilize shared uplinks between these aggregation devices and the service provider's core network. QoS typically is not provisioned within the broadband service provider's cloud in either medium. This lack of QoS within the broadband cloud underscores the requirement of adequate QoS provisioning at the endpoints of the VPN tunnels.

Figure 6-20 DSL and Cable Topologies



Digital Subscriber Line

DSL service features a dedicated access circuit and offers a service similar to Frame Relay or Asynchronous Transfer Mode (ATM), in which a single permanent virtual circuit (PVC) is provisioned from the home office to the service provider aggregation point. DSL has a variety of speeds and encoding schemes.

Most service providers today offer residential Asynchronous Digital Subscriber Line (ADSL). ADSL provides for asymmetric speeds (with the downstream rate greater than the upstream rate). Asymmetrical links are a better choice and benefit for the telecommuter because the greater the downlink bandwidth is, the less the need is for QoS. (Slower-speed uplinks slow TCP sender transmission rates to more manageable levels.) Because QoS is not generally available by broadband service providers, such uplink/downlink speed mismatches are desirable.

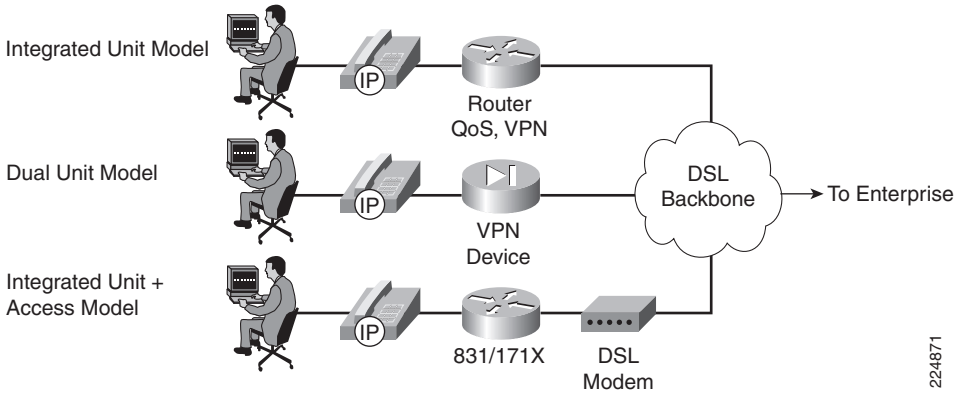
Residential access speeds are generally 128 to 384 kbps upstream and 608 kbps to 1.5 Mbps downstream. For ADSL circuits with upstream speeds less than 256 kbps, G.729 VoIP codecs are recommended.

ADSL also utilizes a single best-effort PVC using RFC 2516–based PPP over Ethernet (PPPoE) encapsulation. In DSL networks, delay and jitter are very low but are not guaranteed. Because PPPoE is used, no LFI is available in service provider DSL networks. In the future, QoS at Layer 2 might be available across service provider networks using familiar ATM variable bit-rate (VBR) definitions.

Single-pair high bit-rate DSL (G.SHDSL) is the new high-speed standard for business DSL. Most residences will continue to be served by ADSL, while small business and branch offices likely will use G.SHDSL. G.SHDSL offers varying rates controlled by the service provider; the upstream and downstream rates are the same speed (symmetric). G.SHDSL is seen as an eventual replacement for T1s in the United States and will become increasingly available from more service providers.

The telecommuter deployment options for DSL circuits include all three teleworker deployment models: Integrated Unit, Dual-Unit, and Integrated Unit + Access, as shown in Figure 6-21.

Figure 6-21 Teleworker Deployment Models for DSL



Cable

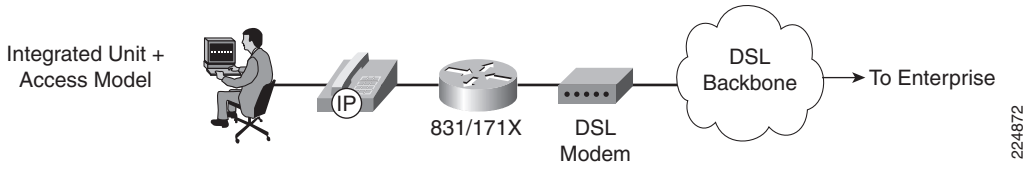
Cable offers a shared service with symmetric speeds varying from 100 kbps to 4 Mbps. In the past, delay and jitter varied too greatly over cable, making it unsuitable for VoIP.

The common installed base of cable services today is made up of Data-over-Cable Service Interface Specifications (DOCSIS) 1.0. No LFI is available with DOCSIS 1.0, although DOCSIS 1.1 defines fragmentation and interleaving mechanisms.

DOCSIS 1.1 also provides the capabilities to shape traffic at Layer 2 before transmission over the cable network. Although the circuit and frequencies physically are shared, access to the medium can be controlled by the headend so that a device can be guaranteed specified bandwidth.

At the time of this writing, the only recommended deployment model for cable is the Integrated Unit + Access Model, as shown in Figure 6-22.

Figure 6-22 Teleworker Deployment Model for Cable



Bandwidth Provisioning

A few key differences with respect to bandwidth provisioning need to be taken into account for broadband teleworker scenarios, as compared to site-to-site VPNs.

The first is that usually only a single call needs to be provisioned for (unless two teleworkers share a single residential broadband connection, which is rarely the case). If bandwidth is low, G.729 codecs are recommended. The second key bandwidth-provisioning consideration is the inclusion of the overhead required by the broadband access technologies, whether DSL or cable.

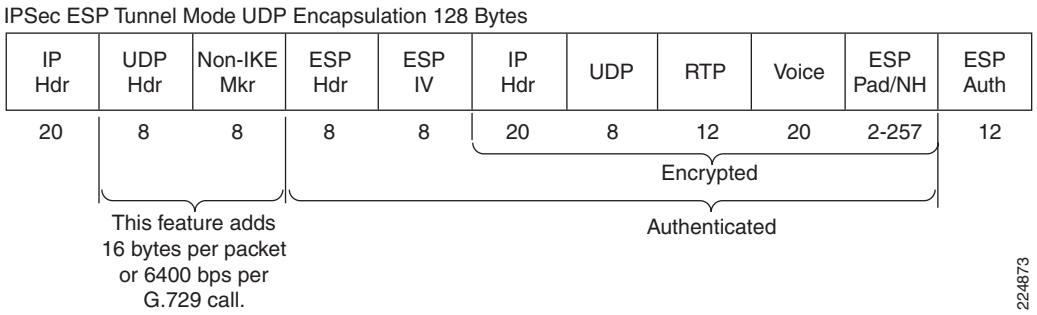
Note Sometimes multicast support is not required in a teleworker environment. For example, routing protocols (which depend on multicast support) might not be required by teleworkers (because default gateways are usually sufficient in this context). In such cases, IPSec tunnel mode (no encrypted IP GRE tunnel) can be used to achieve additional bandwidth savings of 24 bytes per packet.

For the remainder of this chapter, IPSec tunnel mode (no encrypted IP GRE tunnel) is the assumed mode of operation.

NAT Transparency Feature Overhead

Beginning in Cisco IOS Release 12.2(13)T, NAT transparency was introduced and is enabled by default, provided that both peers support the feature. It is negotiated in the IKE exchange between the peers. This feature addresses the environment in which IPSec packets must pass through a NAT/pNAT device, and it adds 16 bytes to each voice and data packet. The overhead that this feature adds is shown in Figure 6-23.

Figure 6-23 NAT Transparency Feature (Layer 3) Overhead



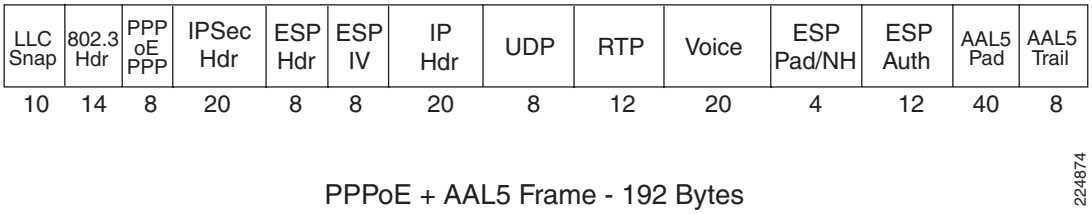
No additional overhead is caused by NAT transparency on a G.729 call over DSL (PPPoE/AAL5) because there is enough AAL5 cell padding to absorb the 16 additional bytes per packet (discussed in more detail in the following section). In cable implementations, these additional 16 bytes increase the number of bits on the wire and, thus, the overall bandwidth consumption. At the headend, NAT transparency increases the bandwidth consumption of VoIP traffic by 1 Mbps for approximately every 82 concurrent G.729 calls; on the teleworker router, NAT transparency increases the bandwidth required by 6.4 kbps per G.729 call.

Unless there is a need to implement this feature, the recommendation is to disable it when bandwidth conservation is a requirement. This feature can be disabled with the **no crypto ipsec nat-transparency udp-encapsulation** global configuration command.

DSL (AAL5 + PPPoE) Overhead

For DSL broadband connections, PPPoE is the most commonly implemented deployment. Given the 112-byte, Layer 3 size of an IPSec (only) encrypted G.729 VoIP call, 40 bytes of PPP, PPPoE, Ethernet, and ATM AAL5 headers and trailers are added. Additionally, the pre-ATM Software Segmentation and Reassembly (SAR) engine is required to pad the resulting 152-byte packet to the nearest multiple of 48 (each fixed-length ATM cell can transport only a 48-byte payload). Figure 6-24 shows the resulting 192-byte pre-ATM packet.

Figure 6-24 IPSec-Encrypted G.729 Packet Size Through AAL5 + PPPoE Encapsulation



The 192 bytes of cell payload are incorporated through ATM SAR into the (48-byte) payloads of four 53-byte ATM cells (192 / 48 = 4 cells).

Therefore, the bandwidth required for an IPSec (only) encrypted G.729 VoIP call over DSL is calculated as follows:

53 bytes per cell
× 4 cells per IPSec-encrypted VoIP packet
212 bytes per (192-byte) AAL5/PPPoE IPSec VoIP packet
× 50 packets per second
10,600 bytes per second

× 8 bits per byte
84,800 bits per second

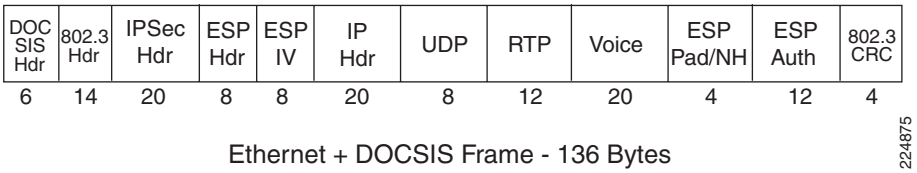
Thus, an encrypted G.729 call requires 85 kbps of bandwidth, while an encrypted G.711 requires seven ATM cells or 148,400 bps on the wire.

This results in the LLQ configuration values of 85 kbps (**priority 85**) for a G.729 VoIP call and 150 kbps for a G.711 (**priority 150**) VoIP call, respectively, for DSL.

Cable Overhead

For cable deployments, the Layer 2 overhead is less than that for DSL. The IPSec packet is encapsulated in an Ethernet header (and trailer) that includes a 6-byte DOCSIS header, as shown in Figure 6-25.

Figure 6-25 IPSec-Encrypted G.729 Packet Size Through Ethernet and DOCSIS Encapsulation



If baseline privacy is enabled (baseline privacy encrypts the payload between a cable modem and the headend), the extended header is used, adding another 5 bytes.

The packet size of a G.729 call (with a zero-length extended header) is 136 bytes or 54,400 bps (at 50 pps); a G.711 call is 280 bytes or 112,000 bps (at 50 pps).

To simplify configuration and deployment, the values of 64 kbps (for G.729) and 128 kbps (for either G.729 or G.711) can be used for priority queue definition for cable.

Asymmetric Links and Unidirectional QoS

With both DSL and cable, the uplink connection can be enabled with QoS, either in the form of a service policy on the DSL (ATM PVC) interface or through a hierarchical MQC service policy that shapes the uplink and prioritizes packets within the shaped rate on the Ethernet interface. This half of the link is under the enterprise's control and easily can be configured.

The downlink connection is under the control of the broadband service provider, and any QoS policy must be configured by the service provider; however, most service providers do not offer QoS-enabled broadband services. This is usually because DSL providers often have implemented non-Cisco equipment (DSLAM or other ATM concentration devices) that typically have few or no QoS features available. Cable providers, on the other hand, might have an option to enable QoS as DOCSIS 1.1 becomes more widely deployed.

Fortunately, most service offerings are asymmetrical. For example, consider a circuit with a 256-kbps uplink and 1.5-Mbps downlink. The downlink rarely is congested to the point of degrading voice quality.

Testing in the Cisco Enterprise Solutions Engineering labs has shown that when congestion is experienced on the uplink, the resulting delay of (TCP) data packet acknowledgments automatically decreases the arrival rate of downlink data traffic so that downlink congestion does not occur. To summarize the results of these tests: Asymmetrical links are preferred (over symmetrical links) as long as QoS is enabled on the uplink, provided that the lower of the two speeds (the uplink) is adequate for transporting both voice and data.

Some service providers for business-class services offer symmetrical links in an effort to compete with Frame Relay providers; 384 kbps/384 kbps and 768 kbps/768 kbps are examples. With no QoS enabled on the service provider edge, this offering is not optimal for the enterprise. An asymmetrical link such as 384 kbps/1.5 kbps is a better choice for the V3PN networks.

Broadband Serialization Mitigation Through TCP Maximum Segment Size Tuning

The majority of broadband deployments are DSL with PPPoE and cable with DOCSIS 1.0. As previously noted, neither of these technologies includes any mechanisms to fragment data packets and interleave voice packets at Layer 2 to minimize the impact of serialization and blocking delay on voice packets (which is a recommended requirement on link speeds ≤768 kbps).

The DOCSIS 1.1 specification for cable includes fragmentation and interleaving support, and DSL providers can implement MLP over ATM (which includes MLP LFI support). However, these do not represent the majority of the currently deployed base of broadband networks.

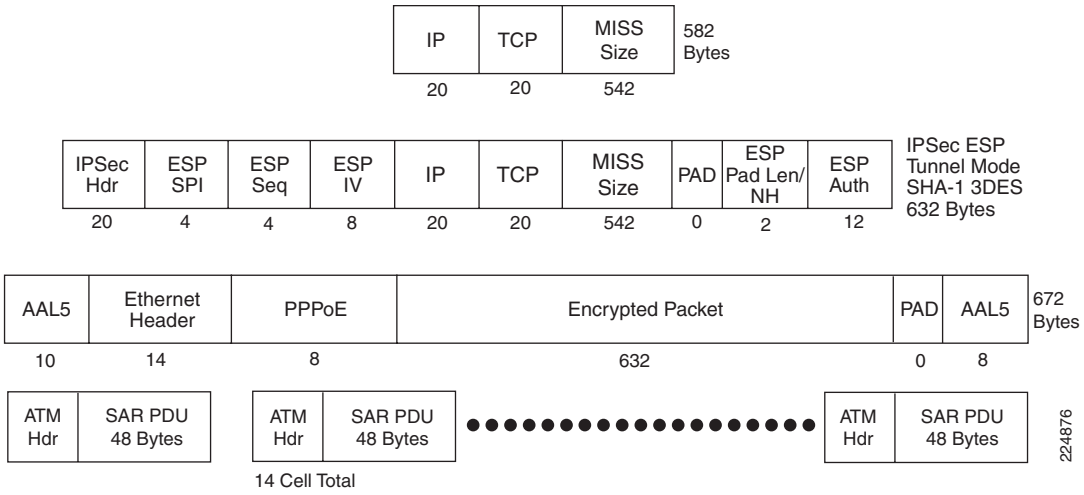
An alternative way to mitigate serialization delay on broadband circuits is provided by adjusting the TCP Maximum Segment Size (MSS). The TCP MSS value influences the resulting size of TCP packets and can be adjusted with the **ip tcp adjust-mss** interface command. Because the majority of large data packets on a network are TCP (normal UDP application packets, such as DNS and NTP, average less than 300 bytes and do not create a significant serialization delay issue), this command effectively can reduce serialization delay in most teleworker scenarios when no Layer 2 fragmentation and interleaving mechanism is available.

Note It is not recommended to run UDP-based video applications on broadband links ≤768 kbps that do not support Layer 2 LFI in a V3PN teleworker scenario. This is because such applications regularly generate large UDP packets that, obviously, are not subject to TCP MSS and thus can cause significant and unmitigatable serialization delays to VoIP packets.

The recommended TCP MSS value of 542 was calculated to eliminate the IPSec crypto algorithm (3DES) padding and the ATM AAL5 padding in DSL implementations (cable implementations have 3DES padding but no AAL5). Therefore, a TCP MSS value of 542 bytes is valid for cable but is optimized for DSL. Figure 6-26 illustrates a TCP packet with an MSS size of 542.

Note Both the ESP and AAL5 pad lengths are 0. A TCP packet with a TCP MSS value of 542 fits exactly into 14 ATM cells, with no wasted bytes because of cell padding.

Figure 6-26 Optimized TCP MSS Value for DSL (542 Bytes)



The evident negative aspect of using this technique to minimize the impact of serialization delay is the decreased efficiency of large data transfers, coupled with an increase in the number of packets per second that the router must switch. The higher packets-per-second rate is not as much of an issue at the remote router as at the headend, where hundreds or thousands of remote connections are concentrated. However, adjusting the TCP MSS value provides a means for the network manager to deploy voice over broadband connections, which do not support any LFI mechanism at rates less than 768 kbps.

Split Tunneling

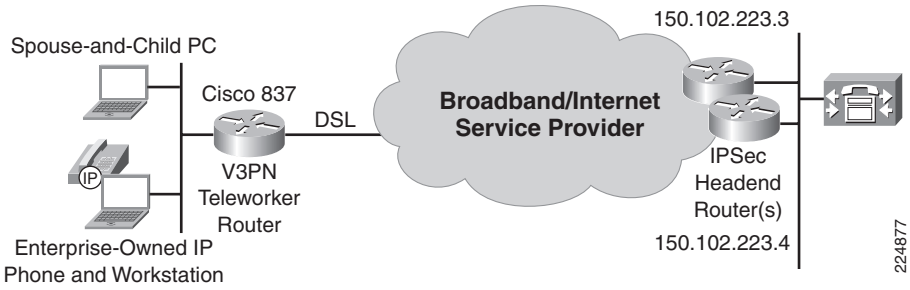
The teleworker is usually not the only user of the residential broadband connection. The worker's spouse and children also might utilize this link from other devices, such as additional PCs or laptops or even gaming units. In such cases, only "work-related" traffic would require encryption and could be given a preferential treatment, through QoS, over general Internet traffic.

In this situation, prioritization could be made based on the destination address of the traffic. One method to accomplish this would be to create a separate bandwidth class (for example, named CORPORATE) and provision this class with a dedicated CBWFQ queue.

Given the sample topology illustrated in Figure 6-27, a suitable QoS configuration is broken down as follows:

- **VOICE**—A single call's VoIP packets assigned to an LLQ, with the bandwidth allocation based on codec and broadband media (DSL or cable).
- **CALL-SIGNALING**—Call-Signaling traffic in a CBWFQ queue, allocated 5 percent.
- **INTERNETWORK**—Internetwork-Control traffic, such as routing protocol updates and IKE keepalives, in a CBWFQ queue, allocated 5 percent.
- **CORPORATE**—All other traffic that is destined to the enterprise through the IPSec tunnel, in a CBWFQ queue, allocated 25 percent.
- **INTERNET/Class-Default**—Traffic to the Internet (outside the IPSec tunnel) defaults to this fair-queued class.

Figure 6-27 Split Tunnel Example Topology



The following configuration fragment provides a means to implement this policy. It assumes that the headend IPSec crypto peers reside on network 150.102.223.0/29. In this example environment, two peers are configured at 150.102.223.3 and 150.102.223.4. Packets within the IPSec tunnel—those matching the CORP-INTRANET access control list (identifying RFC 1918 private addresses that are assumed in this example to represent the intranets behind the headends)—will be encapsulated in an ESP (IPSec-IP protocol 50) IP header. A class-map CORPORATE is configured that references the CORPORATE extended access list, pointing to the VPN headend subnet.

A policy map named V3PN-SPLIT-TUNNEL is created that includes classes for VOICE, INTERNETWORK-CONTROL, and CALL-SIGNALING, along with a bandwidth class for CORPORATE.

In this example, the VOICE class is provisioned for a G.711 codec over a (384-kbps) DSL uplink, allocating 150 kbps (or 40 percent, whichever syntax is preferred) for VoIP. Example 6-8 shows the relevant configuration fragment.

Example 6-8 V3PN-SPLIT-TUNNEL Policy Example

```
!
crypto map SPLIT-TUNNEL-CRYPTO-MAP 1 ipsec-isakmp
 set peer 150.102.223.3
 set peer 150.102.223.4
 set transform-set TS
 match address CORP-INTRANET           ! References CORP-INTRANET ACL
 qos pre-classify                      ! Enables QoS Pre-Classify
!
!
class-map match-all VOICE
  match ip dscp ef                     ! VoIP
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                    ! IP Routing
  match access-group name IKE          ! References ISAKMP ACL
 class-map match-any CALL-SIGNALING
  match ip dscp cs3                    ! New Call-Signaling
  match ip dscp af31                   ! Old Call-Signaling
 class-map match-all CORPORATE
  match access-group name CORPORATE    ! References CORPORATE ACL
!
policy-map V3PN-SPLIT-TUNNEL
  class VOICE
    priority 150                       ! Encrypted G.711 over DSL (PPPoE/AAL5)
  class INTERNETWORK-CONTROL
    bandwidth percent 5                ! Control Plane provisioning
  class CALL-SIGNALING
    bandwidth percent 5                ! Call-Signaling provisioning
  class CORPORATE
    bandwidth percent 25               ! "Work-related" traffic provisioning
    queue-limit 15                     ! Optional: Anti-Replay Tuning
```

```
  class class-default
    fair-queue
    queue-limit 15                    ! Optional: Anti-Replay Tuning
!
...
!
ip access-list extended CORP-INTRANET       ! CORP-INTRANET ACL (RFC 1918)
 permit ip any 10.0.0.0 0.255.255.255
 permit ip any 172.16.0.0 0.15.255.255
 permit ip any 192.168.0.0 0.0.255.255
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp      ! ISAKMP ACL
!
ip access-list extended CORPORATE           ! CORPORATE ACL (VPN Head-ends)
 permit esp any 150.102.223.0 0.0.0.7
!
```

Teleworker V3PN QoS Designs

To review, three deployment models exist for teleworker V3PN scenarios:

- Integrated Unit Model
- Dual-Unit Model
- Integrated Unit + Access Model

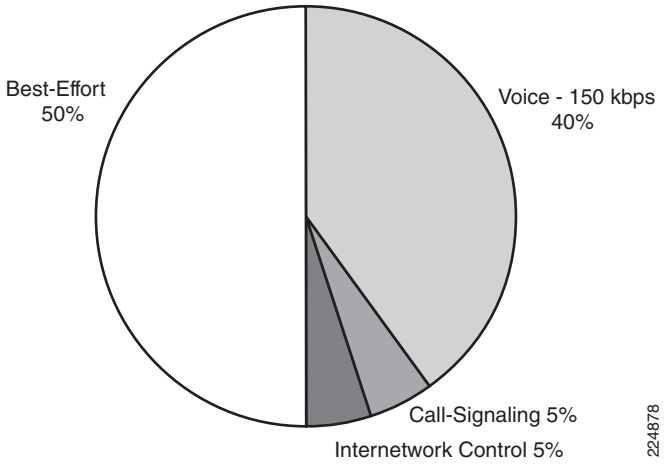
Furthermore, there are two main broadband deployment media: DSL and cable. DSL supports all three teleworker deployment models, but cable supports (at the time of writing) only the Integrated Unit + Access Model.

Integrated Unit/Dual-Unit Models—DSL Design

The key point to remember when provisioning QoS for V3PN over DSL (PPPoE/ATM AAL5) is the significant bandwidth overhead required by these protocols (85 kbps is needed per G.729 call, and 150 kbps is needed per G.711 call).

Some additional points to keep in mind are that since the service policy is being applied to a low-speed ATM PVC, the Tx-ring should be tuned to 3 (as discussed in Chapter 3, "WAN Aggregator QoS Design") and also that because no LFI mechanism exists for DSL, TCP-MSS tuning can be done on the dialer interface to mitigate serialization delay. An Integrated Unit/Dual-Unit V3PN teleworker example for a 384-kbps DSL circuit is illustrated in Figure 6-28 and detailed in Example 6-9.

Figure 6-28 Integrated Unit/Dual-Unit V3PN Teleworker Model for 384-kbps DSL Uplink



Example 6-9 Integrated Unit/Dual-Unit V3PN Teleworker QoS Design Example for a 384-kbps (PPPoE/ATM) DSL Uplink

```
!
 class-map match-all VOICE
  match ip dscp ef                      ! VoIP
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                     ! IP Routing
  match access-group name IKE           ! References ISAKMP ACL
 class-map match-any CALL-SIGNALING
  match ip dscp cs3                     ! New Call-Signaling
  match ip dscp af31                    ! Old Call-Signaling
!
!
 policy-map V3PN-TELEWORKER
  class VOICE
    priority 150                       ! Encrypted G.711 over DSL (PPPoE/AAL5)
  class INTERNETWORK-CONTROL
    bandwidth percent 5                ! Control Plane provisioning
  class CALL-SIGNALING
    bandwidth percent 5                ! Call-Signaling provisioning
  class class-default
    fair-queue
    queue-limit 30                     ! Optional: Anti-Replay Tuning
!
...
!
interface ATM0
 no ip address
 no atm ilmi-keepalive
 dsl operating-mode auto
 dsl power-cutback 0
!
interface ATM0.35 point-to-point
 description Outside PPPoE/ATM DSL Link
 bandwidth 384
 pvc dsl 0/35
  vbr-nrt 384 384
  tx-ring-limit 3                              ! Tx-Ring tuned to 3
  pppoe max-sessions 5
  service-policy output V3PN-TELEWORKER        ! MQC policy applied to PVC
  pppoe-client dial-pool-number 1
```

```
!
...
!
interface Dialer1
 description Dialer for PPPoE
 ip address negotiated
 ip mtu 1492
 encapsulation ppp
 ip tcp adjust-mss 542                        ! TCP MSS value tuned for slow-link
!

...
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp       ! ISAKMP ACL
!
```

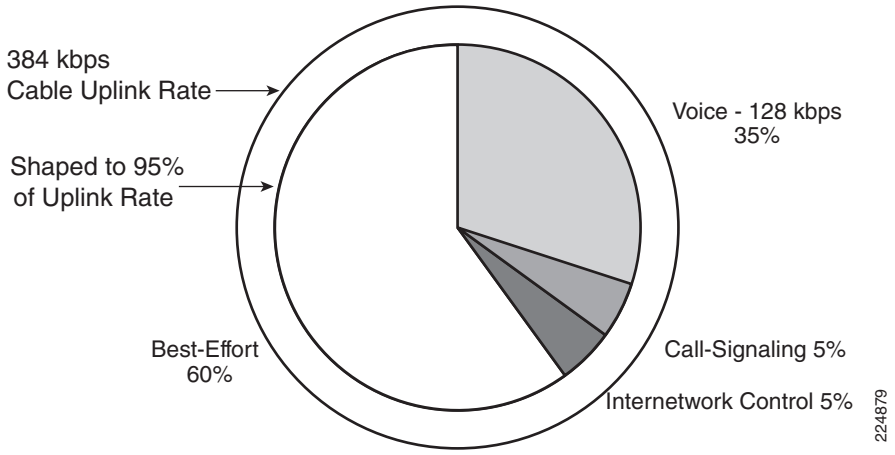
Verification commands:

- **show policy**
- **show policy interface**
- **show atm pvc**

Integrated Unit + Access Model—DSL/Cable Designs

Hierarchical MQC policies are required for Integrated Unit + Access Models to shape and queue (within the shaped rate) on the outbound Ethernet interface. For cable, the shaped rate should be 95 percent of the broadband link's speed (as detailed in Chapter 3, "WAN Aggregator QoS Design" in the "Frame Relay" section ["Committed Information Rate" subsection]). For DSL, the shaped rate should be 70 percent of the uplink's speed (to account for the increased bandwidth overhead required by DSL). Furthermore, on slow-speed (≤768 kbps) links, TCP-MSS should be tuned on both the inbound and outbound Ethernet interfaces. An Integrated Unit + Access Model for a 384-kbps cable teleworker uplink is shown in Figure 6-29 and detailed in Example 6-10.

Figure 6-29 Integrated Unit + Access V3PN Teleworker Model for 384-kbps Cable Uplink



Example 6-10 Integrated Unit + Access Model—Cable Design Example

```
class-map match-all VOICE
```

```
  match ip dscp ef                     ! VoIP
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                    ! IP Routing
  match access-group name IKE          ! References ISAKMP ACL
 class-map match-any CALL-SIGNALING
  match ip dscp cs3                    ! Old Call-Signaling
  match ip dscp af31                   ! New Call-Signaling
!
!
 policy-map V3PN-TELEWORKER
  class VOICE
    priority 128                      ! Encrypted G.711 over Cable
  class INTERNETWORK-CONTROL
    bandwidth percent 5               ! Control Plane provisioning
  class CALL-SIGNALING
    bandwidth percent 5               ! Call-Signaling provisioning
  class class-default
    fair-queue
    queue-limit 30                    ! Optional: Anti-Replay Tuning
!
!
policy-map SHAPE-384-CABLE
 class class-default
   shape average 364800 3640          ! Shapes to 95% of 384 kbps cable link
   service-policy V3PN-TELEWORKER     ! Nested V3PN Teleworker queuing policy
!
...
!
interface Ethernet0
 description Inside Ethernet Interface
 ip tcp adjust-mss 542                          ! TCP MSS value tuned for slow-link
!
interface Ethernet1
 description Outside Ethernet Interface
 ip address dhcp
 ip tcp adjust-mss 542                          ! TCP MSS value tuned for slow-link
 service-policy output SHAPE-384-CABLE          ! Shaper applied to LAN interface
!
```

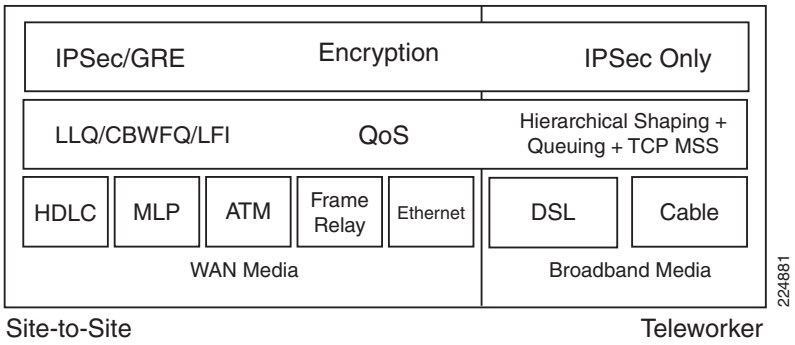
Verification commands:

- **show policy**
- **show policy interface**

Summary

IPSec VPNs, the most commonly deployed VPN solutions today, are found in three main contexts: site-to-site VPNs, teleworker VPNs, and remote-access VPNs. The overlaying of QoS technologies on top of IPSec VPNs is dubbed V3PN, for voice- and video-enabled Virtual Private Networks. This chapter presented considerations and design recommendations for V3PN deployments in site-to-site and teleworker contexts. A summary of the design recommendations for encryption and QoS for site-to-site and teleworker IPSec V3PNs is illustrated in Figure 6-30.

Figure 6-30 IPSec V3PN Site-to-Site and Teleworker Design Summary Comparison



Some site-to-site considerations that were discussed include the bandwidth implications of various IPSec modes of operations and the incompatibility of cRTP and IPSec. The interrelation of IPSec VPN logical hub-and-spoke topologies and the effect these have on spoke-to-spoke delay budgets also were examined. Subsequently, the ToS byte preservation mechanism was overviewed along with the QoS Pre-Classify Cisco IOS Software feature, which allows for the classification of packets already encrypted (on the same router) through ACLs. QoS and Anti-Replay implications then were discussed, illustrating how QoS policies that reorder packets potentially can exacerbate Anti-Replay drops (an IPSec message-integrity mechanism). Techniques for minimizing such undesired QoS/Anti-Reply interaction effects, such as reducing the queue length of data queues, were presented. Next, the need for control plane provisioning was highlighted, along with basic designs for doing so.

Several site-to-site QoS models were detailed, ranging from a six-class V3PN QoS model to a complex 11-class V3PN QoS Baseline model. WAN aggregator considerations specific to IPSec VPN deployments were examined next, including QoS provisioning for IPSec over private WANs, per-tunnel hierarchical shaping and queuing, and recommendations for decoupled VPN headend/WAN aggregation deployment models, where encryption and QoS are performed on different routers.

Attention then shifted to teleworker scenarios and the three main teleworker deployment models: the Integrated Unit Model, the Dual-Unit Model, and the Integrated Unit + Access Model. The two main broadband media types, DSL and cable, were broken down to ascertain the bandwidth-provisioning implications of each media. Neither DSL nor (DOCSIS 1.0) cable includes any mechanism for serialization delay mitigation, so TCP maximum segment-size tuning was considered as an alternative mechanism to achieve this. Split-tunneling designs, to address spouse-and-child requirements, were introduced.

Teleworker V3PN designs then were detailed for Integrated Unit and Dual-Unit models over DSL, in addition to an Integrated Unit + Access Model solution for cable.

References

Standards

- RFC 1321, “The MD5 Message-Digest Algorithm”
  http://www.ietf.org/rfc/rfc1321.txt
- RFC 1918, “Address Allocation for Private Internets”
  http://www.ietf.org/rfc/rfc1918.txt
- RFC 2104, “HMAC: Keyed-Hashing for Message Authentication”

http://www.ietf.org/rfc/rfc2104.txt

- RFC 2401, "Security Architecture for the Internet Protocol"
  http://www.ietf.org/rfc/rfc2401.txt
- RFC 2402, "IP Authentication Header"
  http://www.ietf.org/rfc/rfc2402.txt
- RFC 2403, "The Use of HMAC-MD5-96 Within ESP and AH"
  http://www.ietf.org/rfc/rfc2403.txt
- RFC 2404, "The Use of HMAC-SHA-1-96 within ESP and AH"
  http://www.ietf.org/rfc/rfc2404.txt
- RFC 2405, "The ESP DES-CBC Cipher Algorithm with Explicit IV"
  http://www.ietf.org/rfc/rfc2405.txt
- RFC 2406, "IP Encapsulating Security Payload (ESP)"
  http://www.ietf.org/rfc/rfc2406.txt
- RFC 2407, "The Internet IP Security Domain of Interpretation for ISAKMP"
  http://www.ietf.org/rfc/rfc2407.txt
- RFC 2408, "Internet Security Association and Key Management Protocol (ISAKMP)"
  http://www.ietf.org/rfc/rfc2408.txt
- RFC 2409, "The Internet Key Exchange (IKE)"
  http://www.ietf.org/rfc/rfc2409.txt
- RFC 2410, "The NULL Encryption Algorithm and Its Use with IPSec"
  http://www.ietf.org/rfc/rfc2410.txt
- RFC 2411, "IP Security Document Roadmap"
  http://www.ietf.org/rfc/rfc2411.txt
- RFC 2412, "The OAKLEY Key Determination Protocol"
  http://www.ietf.org/rfc/rfc2412.txt

Books

- Kaeo, Merike. *Designing Network Security.* Indianapolis: Cisco Press, 2003.
- Malik, Saadat. *Network Security Principles and Practices*. Indianapolis: Cisco Press, 2002.
- Mason, Andrew. *Cisco Secure Virtual Private Networks.* Indianapolis: Cisco Press, 2001.

Cisco IOS Documentation

- IP Security and Encryption overview (Cisco IOS Release 12.2)
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fsecur_c/fipsenc/scfencov.htm
- Configuring IPSec network security (Cisco IOS Release 12.2)

http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fsecur_c/fipsenc/scfipsec.htm

- Configuring Internet Key Exchange Security Protocol (Cisco IOS Release 12.2)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fsecur_c/fipsenc/scfike.htm

- Prefragmentation for IPSec VPNs (Cisco IOS Release 12.2[13]T)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftprefrg.htm

- Quality of service for Virtual Private Networks (Cisco IOS Release 12.2[2]T)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t2/ftqosvpn.htm

- Low-latency queuing (LLQ) for IPSec encryption engines (Cisco IOS Release 12.2[13]T)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/llqfm.htm

- IPSec NAT transparency (Cisco IOS Release 12.2[13]T)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftipsnat.htm

- IPSec and quality of service feature (Cisco IOS Release 12.3[8]T)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123newft/123t/123t_8/gtqosips.htm

- Configuring broadband access (Cisco IOS Release 12.2)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fwan_c/wcfppp.htm

- PPP over Ethernet Client (Cisco IOS Release 12.2[2]T)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t2/ftpppoec.htm