



QoS in a SOHO Virtual Private Network for IP Telephony

This chapter provides information about implementing QoS in an AVVID-enabled Small Office Home Office (SOHO) Virtual Private Network (VPN) for IP telephony. It includes the following:

- [Overview](#)
- [QoS Toolset](#)
- [Solutions](#)
- [Summary](#)



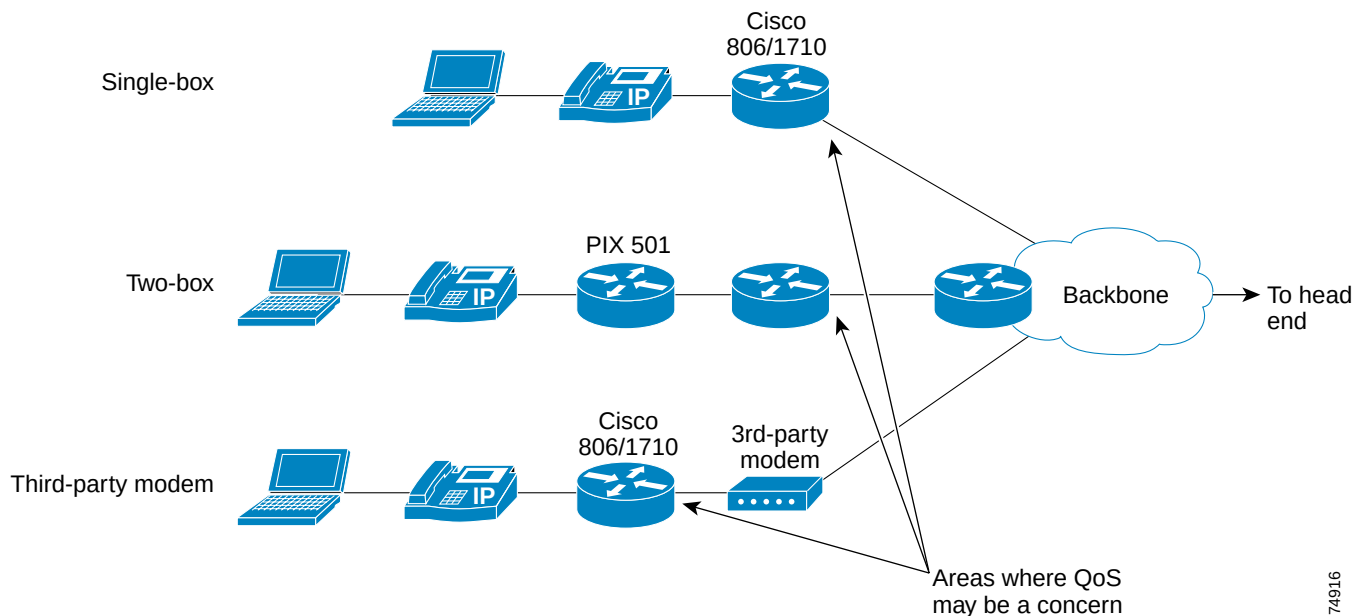
Note

This chapter contains references to other documents. These references are included as tips in the text. The URL for each referenced document is located in [Appendix A, “Reference Information.”](#) In some cases, an internal document is referenced. For copies of internal documents, please see your Cisco Systems representative.

Overview

As IP telephony becomes accepted in the enterprise, the ability to extend its functionality to SOHO environments is becoming more desirable. This chapter provides in-depth design and configuration guidance for implementation of QoS in a SOHO environment. It focuses on the CPE, or enterprise, implementation of QoS features and functionality. This document does not cover end-to-end QoS design and implementation as it does not include SP design and configuration guidance.

There are many points in SOHO networks where QoS mechanisms are required to manage loss, delay, and delay variation. [Figure 5-1](#) illustrates areas where QoS mechanisms are required to control the impact of loss, delay, and delay variation on voice performance.

Figure 5-1 Areas of a SOHO Network Where QoS is Needed

74916

QoS Toolset

The challenges of packet loss, delay, and delay variation can be address through the application of various QoS tools.

This section provides information about the use of QoS tools in a SOHO environment. For general information about the QoS Toolset, see the [“What is the Quality of Service Toolset?”](#) section on page 1-12.

Classification

Within a SOHO AVVID network, you should classify voice bearer and signaling traffic.

For general recommendations for classification, see [“Classification Recommendations”](#) section on page 1-15.

Classification of Voice Bearer traffic

In a SOHO AVVID network, you must classify packets that contain voice traffic so that they can be placed into the appropriate queues. The recommended DSCP PHB label of Expedite Forward (EF) for should be used for VoIP traffic. To remain backwardly compatible with IP Precedence, use an IP Precedence value of 5 and a CoS marking of 5.

These markings can be used as selection criteria for entry into the priority queue, where it exists, or the queue with the highest service weight and lowest drop probability in a WRR/WRED scheduling scheme.

Classification of Voice Signaling Traffic

In a SOHO AVVID network, signaling traffic should be assigned a DSCP PHB label of AF31, an IP Precedence value of 3, and a CoS marking of 3.

Scheduling

The IOS feature of LLQ/CBWFQ is used in a SOHO environment to schedule the voice bearer and voice control traffic preferentially over all other types of traffic if there is congestion on the link.

For general recommendations for classification, see [“Scheduling Recommendations” section on page 1-22](#).

Provisioning

In a SOHO/VPN environment there are many aspects to consider when provisioning the CPE device, such as TX ring sizing, link fragmentation and interleave, traffic shaping, and bandwidth calculation.

TX Ring Sizing

On ATM interfaces, which are the transport medium for many DSL environments, the default depth of the TX ring is very large. However, a large TX ring and LLQ/CBWFQ do not work well together.

When congestion occurs and LLQ/CBWFQ is engaged (to give voice packets preferential treatment), the TX ring must be emptied before the first expedited voice packet can get serialized by the physical interface and be transmitted. If the TX ring is very large, it can take some time to empty before the voice packets are serviced. This results in very short periods of poor voice quality followed by long period of excellent voice quality. The solution is to reduce the TX ring size.

Link Fragmentation and Interleave

In DSL environments that use ATM as the transport medium, MLP with fragmentation and interleave is used as a Layer 2 LFI mechanism. Depending on the underlying (Layer 2) transport, other LFI mechanisms may be required.

Today's cable deployments use link speeds in excess of 768 Kbps. At link speeds greater than 768 Kbps, the variation in delay introduced by random-sized packets is not of significant concern and LFI is not required.

Fragment Sizing for MLPPP over ATM

When ATM is the underlying transport technology in a DSL network, there are 53 byte cells for transport with 48 bytes of usable payload. Therefore, the fragment size selected must be divisible by 48 with a remainder of zero so that partial cells are **not** transmitted. MLPPP with fragmentation and interleave uses a maximum delay parameter and calculates the fragment size based on the bandwidth of the interface. Some manipulation of the interface bandwidth statement and max delay statement is required to arrive at a fragment size that cleanly maps into ATM cells.

**Tip**

Currently, IOS does not automatically optimize the fragment sizes (which are indirectly defined by the MLP fragment-delay parameter). Arriving at such optimal sizes requires a fair bit of math, which is well documented in the *Multilink PPP over Frame Relay and ATM* white paper (internal).

[Table 5-1](#) provides a list of often used PVC speeds and the corresponding fragment delay and interface bandwidths required to achieve the desired fragment size and delay.

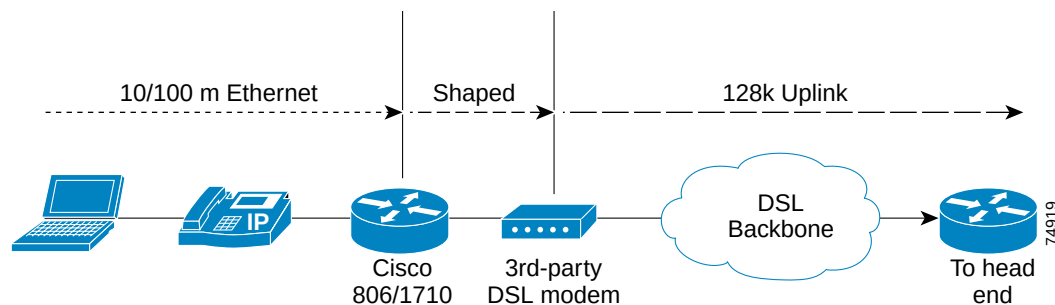
Table 5-1 Common ATM Virtual Circuit Speeds and Maximum Delay and Bandwidths

PVC Speed (in Kbps)	Fragmentation Size (in cells)	PPP Multilink Fragment Delay (in msec)	Bandwidth (in Kbps)	Real Delay (in msec)
56	2	12	57	13.7
64	2	10	68	12.0
128	4	11	132	12.0
192	6	11	202	12.0
256	7	10	260	10.5
320	9	10	337	10.8
384	11	10	414	11.0
448	12	10	452	10.3
512	14	10	529	10.5
576	16	10	606	10.7
640	17	10	644	10.2
704	19	10	721	10.4
768	21	10	798	10.5

Traffic Shaping

Traffic shaping is required in SOHO environments to insure that the SP guaranteed rate is not exceeded. Traffic in excess of the guarantee could result in drops or queuing that would defeat the QoS measures (LLQ/CBWFQ, and LFI) that have been applied at the edge of the network. To avoid congestion in the virtual path through the SP network, traffic in the direction of the SP should be rate-limited to the negotiated SP guaranteed rate.

Figure 5-2 Traffic Shaping



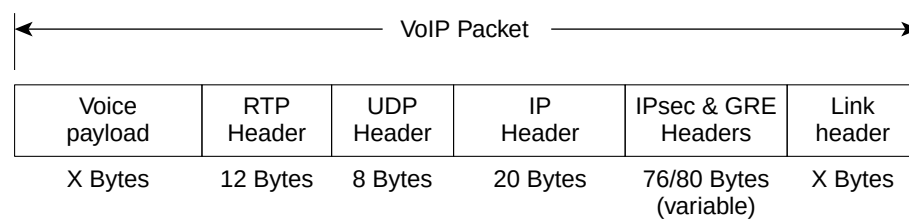
- In DSL over ATM (PPPoA) environments, ATM traffic shaping per Virtual Circuit (VC) is inherently available. No additional steps beyond defining the VC type and speed are required.
- In DSL over Ethernet (PPPoE) environments, traffic shaping on the egress Ethernet interface of the first QoS capable device is required.
- In Cable environments, the Data over Cable Service Interface Specification (DOCSIS) specification version 1.1 allows us to operate within the guaranteed rates to which the SP has agreed and is policing.

Bandwidth Calculation

QoS is not a substitute for sufficient bandwidth. It is important to understand the exact bandwidth requirements of the priority traffic so that links can be provisioned accordingly. For VoIP traffic in a SOHO environment, you must consider the CODEC sample size, and the IP, UDP, RTP, and IPsec overhead when provisioning the link (as shown in Table 5-1).

Because of the overhead associated with the encapsulating protocols involved in an IPsec solution, the bandwidth requirements are considerably more than one might think. For example, a G.711, or 64 Kbps, call requires 128 Kbps of bandwidth over an ATM (PPPoA) connection.

Figure 5-3 VoIP Bandwidth Utilization Including IPsec



VoIP with IPsec MLPPP over ATM			
CODEC	Plus IP UDP RTP and IPsec	Plus PPP	Plus ATM cells 53b cells 48b payload equals
G.711 at 50 pps	112 kbps	114.40 kbps	127.20 kbps
G.729A at 50 pps	54.4 kbps	56.8 kbps	63.6 kbps

74920

Solutions

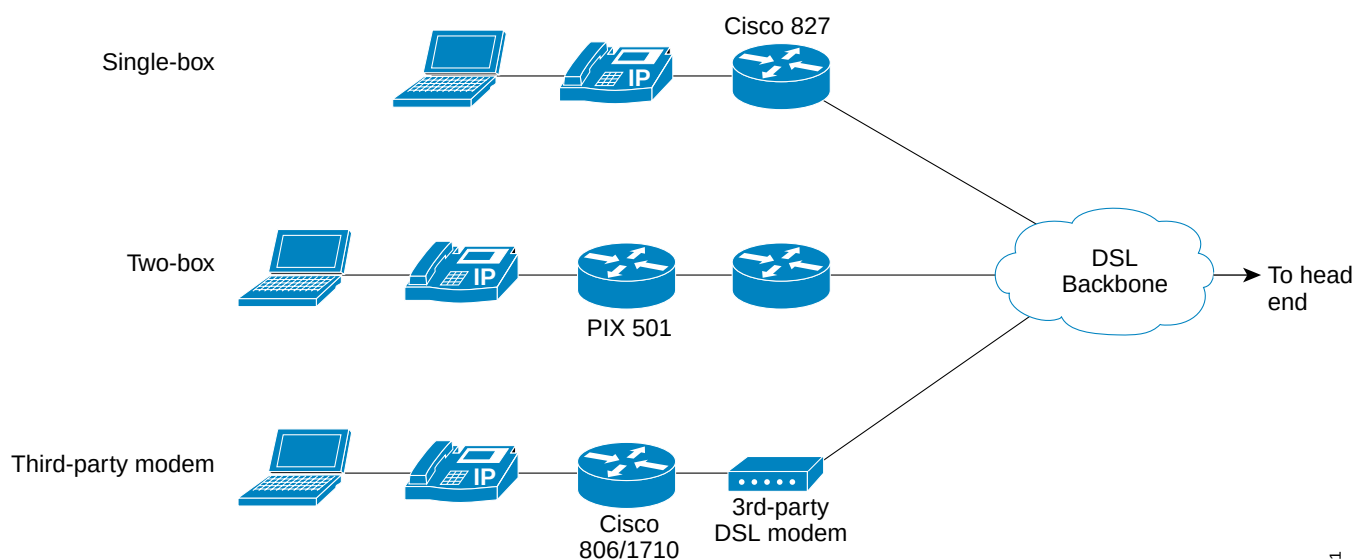
Now that you understand the problems that you need to solve (classification, scheduling, and provisioning) and the tools that you can use to resolve those problems (LLQ/CBWFQ for classification and scheduling, MLPPP with fragmentations and interleave for LFI), as well as an understanding of the bandwidth requirements of VoIP in an IPsec environment, let's see how you apply these tools in the following environments:

- DSL
- Cable
- Other

Application of QoS to DSL in a SOHO Environment

In SOHO VPN environments that use DSL as the transport, there are three alternatives for deployment (as illustrated in Figure 5-4). They are a one-box solution, a two-box solution, and a solution that uses a third-party modem.

Figure 5-4 Typical DSL Deployments



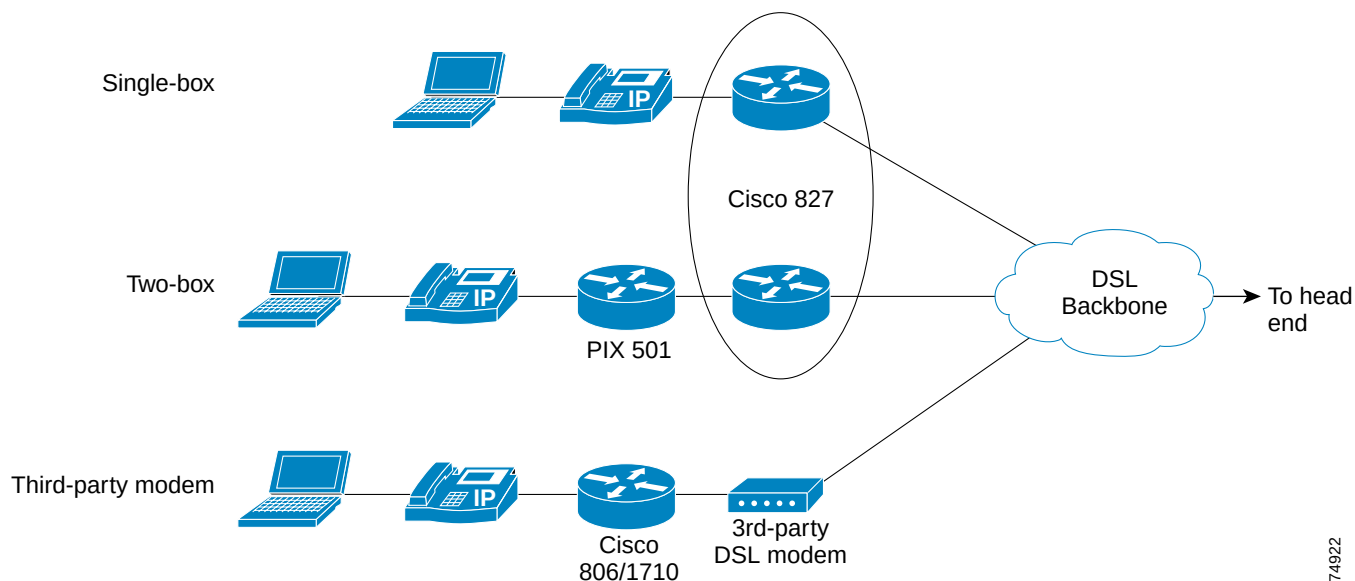
74921

There are two commonly used methods of delivering DSL to the SOHO, DSL over ATM (PPPoA) and DSL over Ethernet (PPPoE). PPPoA implementations are the only DSL environments where you have the complete set of features required to address all of the QoS requirements of the solution. In a PPPoE environment (as of the writing of this paper), the full set of features required to address QoS from the CPE, or last mile, perspective does not exist.

One and Two Box DSL Solution

The one-box and two-box solutions require the same set of features from a QoS perspective. The only difference is that in a one box solution, VPN and QoS are in the same box and a two box solution, the VPN functionality is moved out of the DSL termination device. In both of these solutions (where QoS is being provided by the DSL termination devices), all three aspects of QoS must be provided: classification, scheduling (including LFI and traffic shaping), and provisioning.

Figure 5-5 One- and Two-Box DSL Deployments



For the one- and two-box DSL solutions (illustrated in [Figure 5-5](#)):

- The IP phone handles classification of the VoIP bearer and control traffic. The IP phone marks its traffic at Layer 2 with a CoS of 5 for bearer traffic and 3 for control traffic. It marks its traffic at Layer 3 with a DHCP PHB label of EF for bearer traffic and AF31 for control traffic.
- A Cisco 827, or equivalent router, handles scheduling.
- LLQ/CBWFQ is used to give the VoIP bearer and control traffic the guarantees for loss, delay, and delay variation that they require.

[Example 5-1](#) shows the IOS configuration used to address the QoS requirements for VoIP in a SOHO environment.

Example 5-1 DSL PPPoA QoS Configuration

```
class-map match-all VOICE
  match ip dscp EF
class-map match-all VOICE-CONTROL
  match ip dscp AF31
!
policy-map TELEWORK
  class VOICE
    priority 64
  class VOICE-CONTROL
    bandwidth 8
  class class-default
    fair-queue
```

```

!
interface ATM0
  no ip address
  pvc 1/100
    vbr-rt 128 128
    tx-ring-limit 3
    encapsulation aal5mux ppp dialer
    dialer pool-member 1
interface Dialer0
  bandwidth 132
  ip address negotiated
  ip nat outside
  encapsulation ppp
  no ip mroute-cache
  load-interval 30
  dialer pool 1
  dialer-group 1
  service-policy output TELEWORK
  no cdp enable
  ppp authentication chap callin
  ppp chap hostname 827a
  ppp chap password 7 104D000A0618
  ppp multilink
  ppp multilink fragment-delay 11
  ppp multilink interleave

```

In this example:

- The **class-map**, **policy-map**, and **service-policy** commands are used for scheduling.
- The TX ring must be tuned so that voice traffic (given preferential treatment by the LLQ/CBWFQ configuration) does not wait behind a large amount of other traffic before it is serialized onto the physical interface. The **tx ring limit 3** command is used to tune the TX ring to a reasonable depth.
- The majority of DSL deployments deliver an upstream link that is less than 768k. To solve this problem, use MLPPP with fragmentation and interleave. The **ppp multilink** commands are used to enable MLPPP. You must set the fragment size so that it easily maps into the 48 bytes available for data in an ATM cell. This is accomplished by manipulating the interface bandwidth statement and the MLPPP fragment delay statement to arrive at a fragment size that fits into ATM cells without causing half full cells. The **interface bandwidth** and **ppp multilink fragment-delay** commands are used to achieve the desired fragment size in this example.



Tip

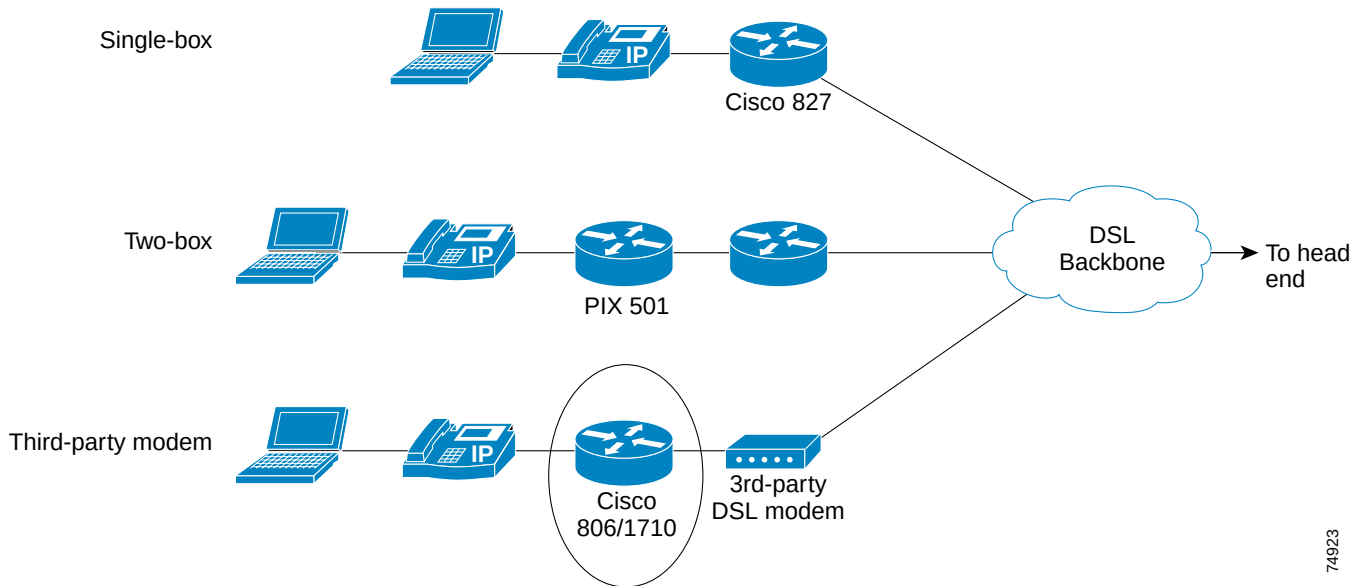
Table 5-1 lists the bandwidth and fragment delay combinations required for many common ATM VC speeds. A detailed analysis of this requirement and how to arrive at the specified bandwidth and delay values can be found in the *Multilink PPP Over Frame Relay & ATM* white paper (internal).

- Traffic shaping is handled by the PVC definition. The **vbr-rt 128 128** command is used to define the ATM virtual circuit and to shape the traffic to the specified rate.
- This example is provisioned to support a single G.729 call (64 Kbps of bandwidth including the encapsulating technologies overhead). The **priority 64** command in the service policy definition section is used to provision the LLQ (PQ) for this purpose.

Third-Party Modem Solution

Figure 5-6 illustrates where QoS is required when a third-party DSL modem is used as the DSL termination device. In this environment QoS cannot be guaranteed.

Figure 5-6 DSL with Third-party Modem



For the third-party modem solution (illustrated in Figure 5-6):

- The IP phone handles classification of the VoIP bearer and control traffic. The IP phone marks its traffic at Layer 2 with a CoS of 5 for bearer traffic and 3 for control traffic. It marks its traffic at Layer 3 with a DHCP PHB label of EF for bearer traffic and AF31 for control traffic.
- A Cisco 806 or 1710, or equivalent router, handles scheduling.
- LLQ/CBWFQ is used to recognize and preferentially schedule VoIP bearer and control traffic in the direction of the DSL modem.
- Class-based policing (shaping) is used to rate limit the traffic transmitted towards the DSL modem to the rate that you are guaranteed from the SP on the DSL link.

There are many factors that are out of our control when you do not directly terminate the DSL connection. Example 5-2 illustrates the best that can be done in this environment. The biggest issue left to be resolved in this configuration is the lack of LFI. Packets can be fragmented in the direction of the DSL modem. However, IOS today does not have the ability to interleave on this segment of the network so you can not insure that a VoIP packet will not wait for a maximum MTU packet to serialize out the physical link on the DSL modem. This would introduce a large delay and delay variation into the VoIP traffic and compromise the voice quality (due to packet loss as a result of jitter buffer under and over runs).

Example 5-2 DSL QoS Configuration with a Third-party Modem

```
class-map match-any CBS-256kbps
  match any
class-map match-all VOICE
  match ip dscp EF
class-map match-all VOICE-CONTROL
  match ip dscp AF31
!
policy-map WAN-EDGE
  class VOICE
    priority percent 33
  class VOICE-CONTROL
    bandwidth percent 2
class class-default
  fair-queue
  random-detect dscp-based
policy-map CBS-256kbps
  class CBS-256kbps
    bandwidth 256
    shape peak 256000
    service-policy WAN-EDGE
!
interface Ethernet1
  ip address dhcp
  service-policy out CBS-256kbps
```

In this example:

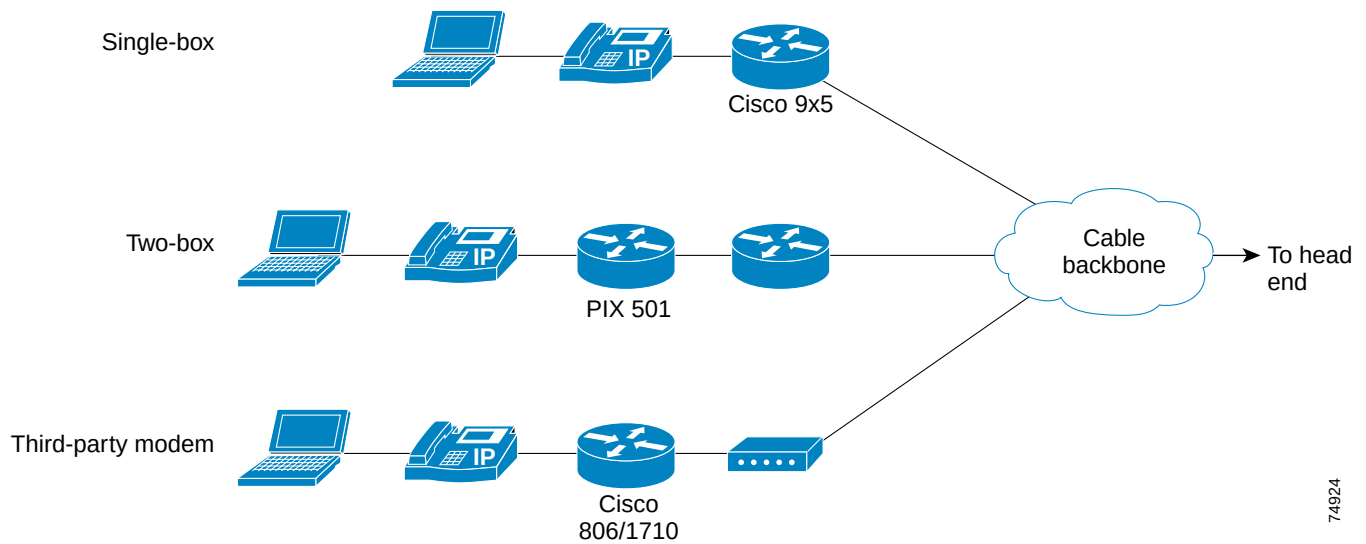
- The concept of hierarchical service policies and classed-based policing is introduced. The **class-map match-any** and **policy-map** commands are used to provide class-based policing. The symbolic name of “CBS-256kbps” represents class-based shaping (CBS) and the link speed.

**Note**

This configuration addresses most of the QoS requirements of a solution where you are not in direct control of the DSL termination. However, LFI is not addressed, so large data packets could introduce variability in delay and affect voice quality.

Application of QoS to Cable in a SOHO Environment

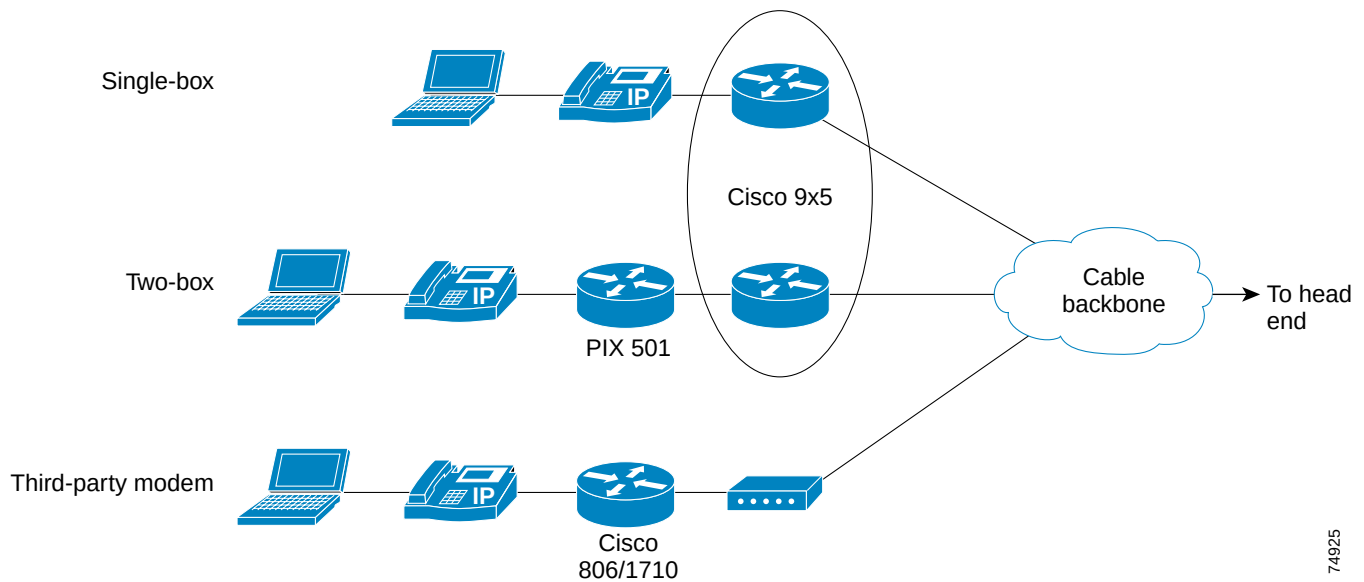
In SOHO VPN environments that use Cable as the transport, there are three alternatives for deployment (as illustrated in [Figure 5-7](#)). They are a one-box solution, a two-box solution, and a solution that uses a third-party modem.

Figure 5-7 Typical Cable Deployment

74924

One and Two Box Cable Solution

The one-box and two-box solutions require the same set of features from a QoS perspective. The only difference is that in a one box solution, VPN and QoS are in the same box and a two box solution, the VPN functionality is moved out of the cable termination device.

Figure 5-8 One- and Two-Box Cable Deployments

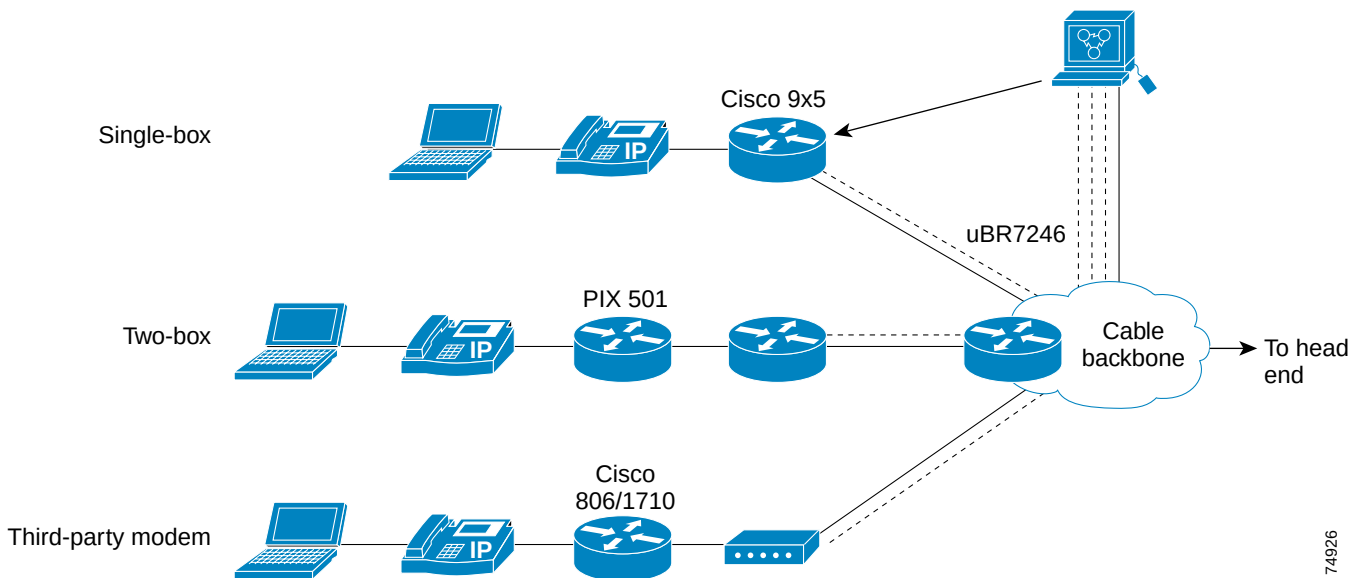
74925

For the one- and two-box cable solutions (illustrated in Figure 5-8):

- The IP phone handles classification of the VoIP bearer and control traffic. The IP phone marks its traffic at Layer 2 with a CoS of 5 for bearer traffic and 3 for control traffic. It marks its traffic at Layer 3 with a DHCP PHB label of EF for bearer traffic and AF31 for control traffic.
- QoS is provided via DOCSIS version 1.1. DOCSIS 1.1 provides Layer 2 QoS services that enable a cable solution to provide bandwidth guarantees for VoIP traffic.

It is important to note that QoS in cable environments is not accomplished through the Modular QoS CLI configuration. Instead, a DOCSIS policy is downloaded to the cable router when it boots. Then the router enforces the policy as it transmits traffic. Figure 5-9 illustrates this download process.

Figure 5-9 DOCSIS 1.1 Configuration Downloaded to a Cable Router



Example 5-3 highlights specific options in the cable modem DOCSIS 1.1 configuration that are required to provide the bandwidth and scheduling guarantees for VoIP traffic.

Example 5-3 Cable DOCSIS 1.1 Extracts for QoS

```

18 (Maximum Number of CPE)      = 10
22 (Upstream Packet Classification Block)
  T01 (IP ToS Range & Mask)    = 160.160.224      Match IP Prec 5
  T02 (IP Protocol)            = 256
  T02 (Source MAC Addr)        = 0-8-21-3a-29-7e    Match MAC
23 (Downstream Packet Classification Block)
  T01 (IP ToS Range & Mask)    = 160.160.224      Match IP Prec 5
  T05 (Destination Address)     = 10.91.20.4        Match IP address
  T06 (Destination Mask)       = 255.255.255.255
24 (Upstream Service Flow Block)
  S19 (Unsolicited Grant Size)  = 300              Bandwidth
  S20 (Nominal Grant Interval)  = 20000            Timeslot

```

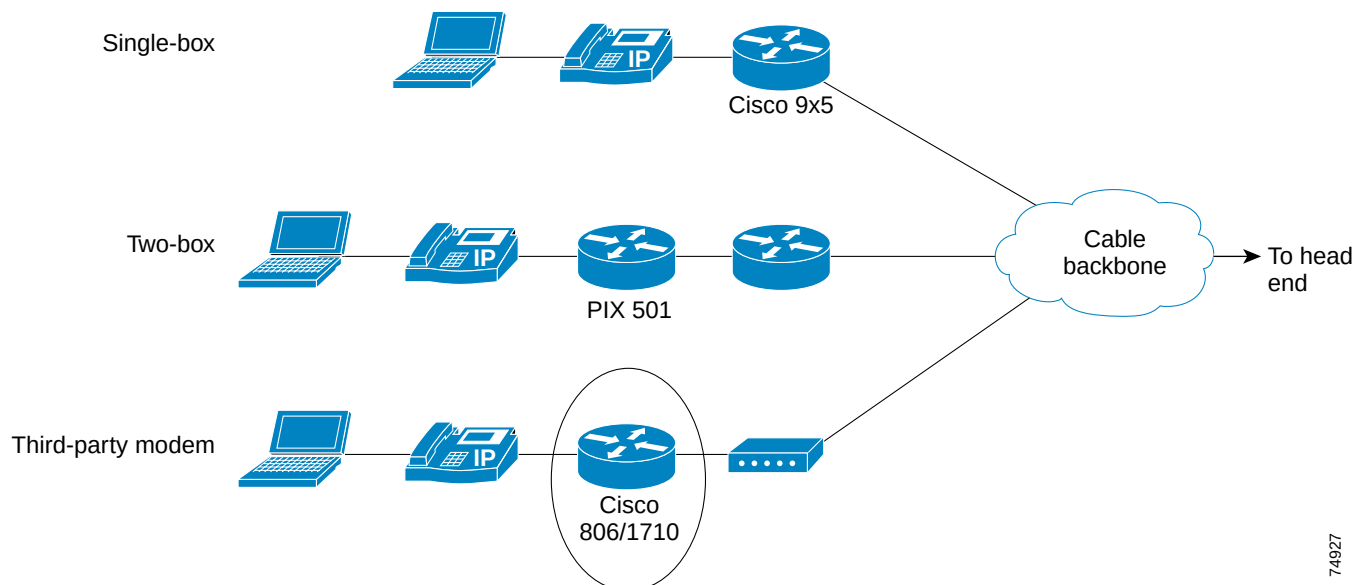
In this example:

- Options 22 and 23 specify traffic source and destination.
- Option 24 specifies the bandwidth guarantee being granted to the VoIP traffic.

Third-Party Modem Solution

In environments where the cable termination device is not QoS capable, the same restrictions and limitations exist as in a cable environment where a third-party cable modem is used.

Figure 5-10 Cable QoS with a Third-party Modem



For the third-party modem solution (illustrated in [Figure 5-10](#)):

- The IP phone handles classification of the VoIP bearer and control traffic. The IP phone marks its traffic at Layer 2 with a CoS of 5 for bearer traffic and 3 for control traffic. It marks its traffic at Layer 3 with a DHCP PHB label of EF for bearer traffic and AF31 for control traffic.
- A Cisco 806 or 1710, or equivalent router, provides preferential scheduling for the VoIP bearer and control traffic that has been classified or tagged by the IP phone.
- LLQ/CBWFQ is used to give the preferential treatment.
- The router also shapes its traffic to the guaranteed bandwidth that the service provider is delivering.

There are many factors that are out of our control when you do not directly terminate the cable connection. [Example 5-4](#) illustrates the best that can be done in this environment.

Example 5-4 Cable QoS Configuration with a Third-party Modem

```
class-map match-any CBS-256kbps
  match any
class-map match-all VOICE
  match ip dscp EF
class-map match-all VOICE-CONTROL
  match ip dscp AF31
!
policy-map WAN-EDGE
  class VOICE
    priority percent 33
  class VOICE-CONTROL
    bandwidth percent 2
  class class-default
    fair-queue
    random-detect dscp-based
```

```

policy-map CBS-256kbps
  class CBS-256kbps
    bandwidth 256
    shape peak 256000
    service-policy WAN-EDGE
!
interface Ethernet1
  ip address dhcp
  service-policy out CBS-256kbps

```

This configuration cannot provide LFI, which means that if the link speeds available in the cable plant are less than or equal to 768 Kbps, the delay through the SP network could introduce packet loss due to jitter buffer over and under runs. In most cable environments, LFI is not needed. However, there is no method in this environment of informing the QoS-enabled router of changes in the load present on the cable segment. For QoS to be effective in this case you must shape to a worst-case rate to protect VoIP bearer and control traffic from congestion in the cable segment. The QoS provided in this environment may not be sufficient to guarantee voice quality as the QoS-enabled router does not have vision into the state of the cable network.

Application of QoS over Other Technologies

There are SOHO solutions other than DSL and Cable that are viable solutions for a QoS-enabled SOHO environment that supports IP telephony. [Figure 5-11](#) illustrates alternate environments for SOHO link delivery.

Figure 5-11 Typical Other Deployments



Last Mile Wireless, IDSL, etc.

In environments where the device terminating the SP link is another technology, such as last mile wireless or DSL with ISDN as the Layer 2 transport (IDSL), you can provide scheduling services via LLQ/CBWFQ for the traffic that has been tagged with the appropriate DSCP PHB labels by the IP phone. You can also shape the traffic to the speeds guaranteed to by the SP. You cannot, however, provide LFI for these types of connections. This means that on links where either the link serialization rate or virtual circuit speeds are equal to or less than 768 Kbps, you cannot provide sufficient QoS services to guarantee voice quality.

[Example 5-5](#) illustrates the best that can be done in this environment.

Example 5-5 Configuration for Last Mile Wireless, IDSL, and Others

```

class-map match-any CBS-256kbps
  match any
class-map match-all VOICE
  match ip dscp EF
class-map match-all VOICE-CONTROL
  match ip dscp AF31
!

```

```

policy-map WAN-EDGE
  class VOICE
    priority percent 33
  class VOICE-CONTROL
    bandwidth percent 2
  class class-default
    fair-queue
    random-detect dscp-based
policy-map CBS-256kbps
  class CBS-256kbps
    bandwidth 256
    shape peak 256000
  service-policy WAN-EDGE
!
interface Ethernet1
  ip address dhcp
  service-policy out CBS-256kbps

```

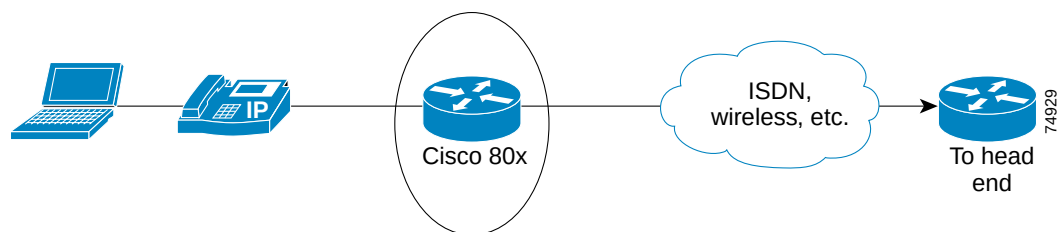
In this example:

- The **class-map VOICE**, **class-map VOICE-CONTROL**, and **service-policy WAN-EDGE** commands are responsible for providing LLQ/CBWFQ for preferential treatment of the VoIP bearer and control traffic.
- The **policy-map CBS-256kbps** and **service-policy out CBS-256kbps** commands are responsible for shaping the traffic towards the SP link, protecting the VoIP traffic from unexpected loss and queuing delays due to exceeding the SP guaranteed bandwidth.

Voice over ISDN

VoIP over ISDN is an acceptable solution for a SOHO solution. [Figure 5-12](#) illustrates how a SOHO connected via ISDN would look from a device perspective.

Figure 5-12 VoIP over ISDN



This is not a solution where IPsec is encrypted through the Internet. It is a private network extension using ISDN. Prior to IOS 12.2(7a), this was not a viable solution because the LLQ/CBWFQ Modular QoS CLI service policies were not available for ISDN connections. With IOS 12.2(7a) and beyond, you now have access to the full suite of tools required to guarantee voice quality over an ISDN connection.

[Example 5-6](#) illustrates the IOS configuration required to provide QoS in this environment.

Example 5-6 VoISDN QoS Configuration Example

```

class-map match-all VOICE
  match ip dscp EF
class-map match-all VOICE-CONTROL
  match ip dscp AF31
!

```

```

policy-map VOICE-AND-DATA
  class VOICE
    priority percent 33
  class VOICE-CONTROL
    bandwidth percent 10
!
interface BRI0/0
  encapsulation ppp
  dialer pool-member 1
  ppp authentication chap
interface Dialer1
  encapsulation ppp
  dialer pool 1
  dialer remote-name routerB-dialer1
  dialer-group 1
  dialer string 12345678
  service-policy output VOICE-AND-DATA
  ppp authentication chap
  ppp chap hostname routerA-dialer1
  ppp chap password cisco
  ppp multilink
  ppp multilink fragment-delay 10
  ppp multilink interleave
  ppp multilink links minimum 2

```

In this example:

- The **class-map**, **policy-map**, and **service-policy** commands enable LLQ/CBWFQ to provide preferential treatment for the VoIP bearer and control traffic.
- The **ppp multilink** commands enable MLPPP/LFI to use link fragmentation and interleave to protect VoIP traffic from variable delay due to serialization of random packet sizes.



Tip

A detailed study of VoIP over ISDN can be found in the *VoIP over ISDN* white paper (internal).

Summary

There are three major categories of SOHO environments: DSL, Cable, and others. These environments have varying levels of QoS support.

- DSL over ATM environments support the full suite of QoS features. In a DSL over ATM (PPPoA) environment, LLQ/CBWFQ is used to give VoIP traffic preferential treatment and MLPPP with fragmentation and interleave is employed to break data traffic into small fragments and protect VoIP traffic from delay due to serialization of large packets on the physical interface. DSL over ATM also inherently supports traffic shaping, as the rates that the ATM virtual circuit can support must be defined when the virtual circuit is configured. This insures that the DSL terminating device will not transmit traffic at a rate in excess of the SP guaranteed rate.
- Cable solutions that are DOCSIS 1.1 enabled can also provide the QoS services required to support VoIP applications. The QoS mechanisms required are inherent to DOCSIS 1.1.
- Other solutions (such as DSL over Ethernet (PPPoE), DSL over ISDN (IDSL), or implementations where a third-party device terminates the SP link) support an incomplete set of QoS services and are not ideal candidates for environments where guaranteed VoIP performance is required. However, many QoS problems can be overcome in these environments and VoIP for internal enterprise communications can be supported.