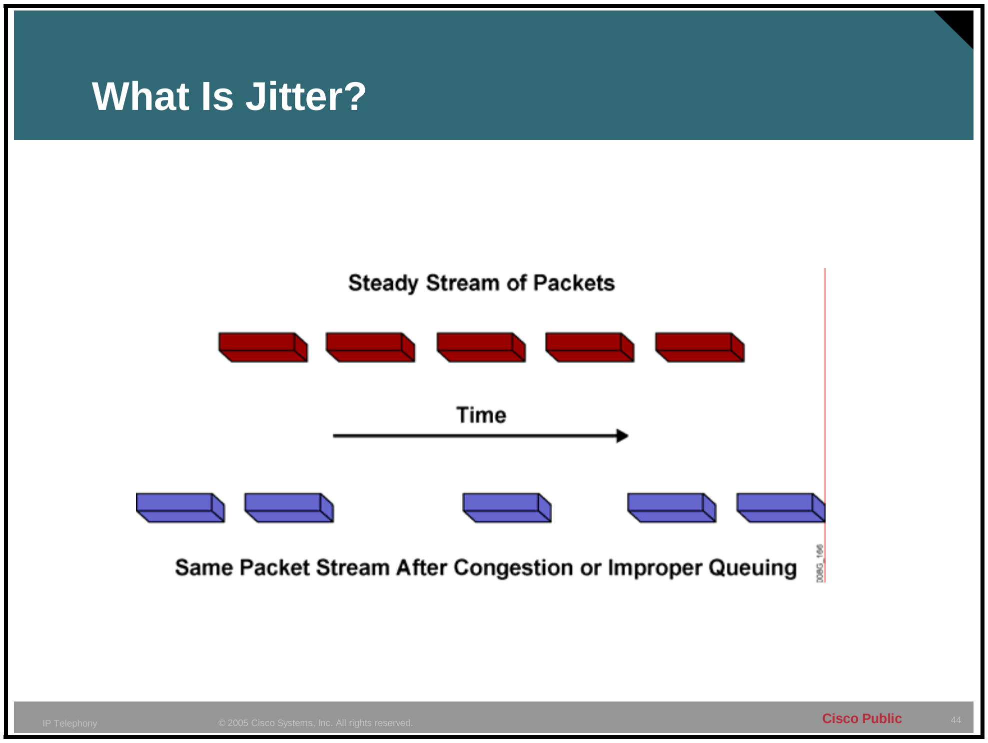
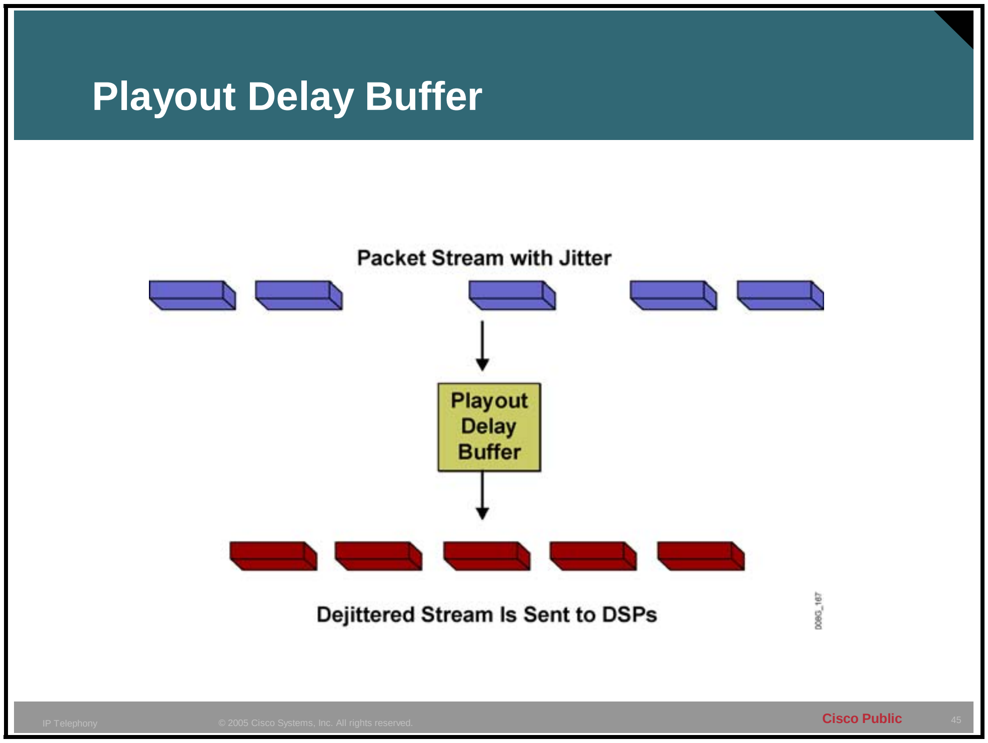
Jitter

Understanding Jitter

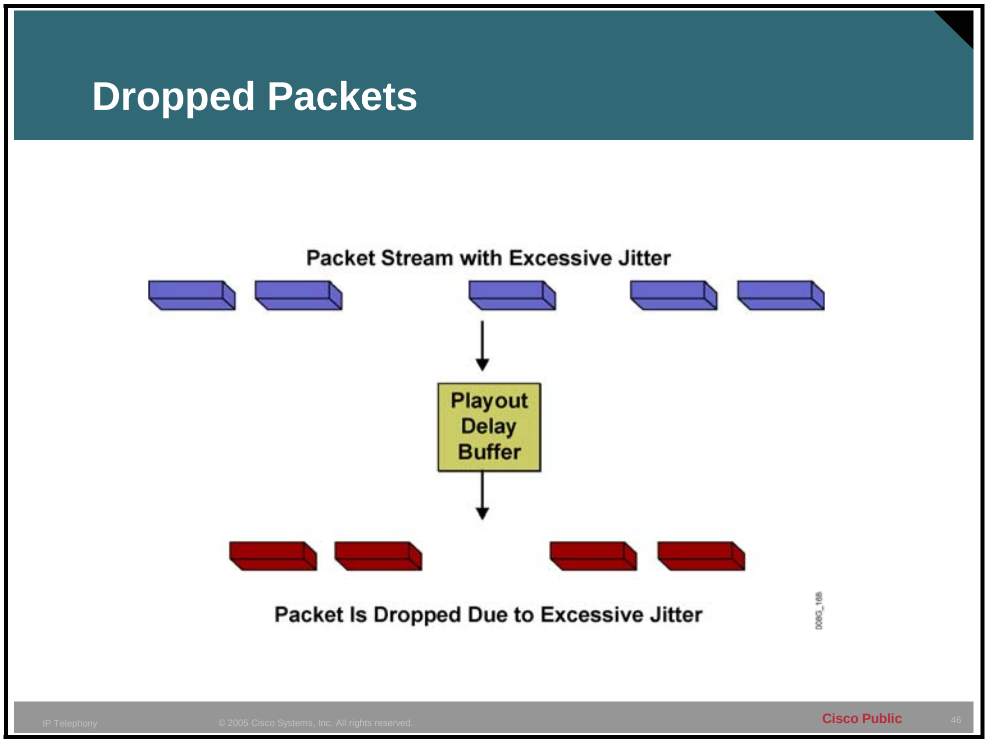
Jitter is an undesirable effect caused by the inherent tendencies of TCP/IP networks and components. This topic describes the cause and effect of jitter.



Jitter is defined as a variation in the delay of received packets. The sending side transmits packets in a continuous stream and spaces them evenly apart. Because of network congestion, improper queuing, or configuration errors, the delay between packets can vary instead of remaining constant, as shown in the figure. This variation causes problems for audio playback at the receiving end. Playback may experience gaps while waiting for the arrival of variable delayed packets.



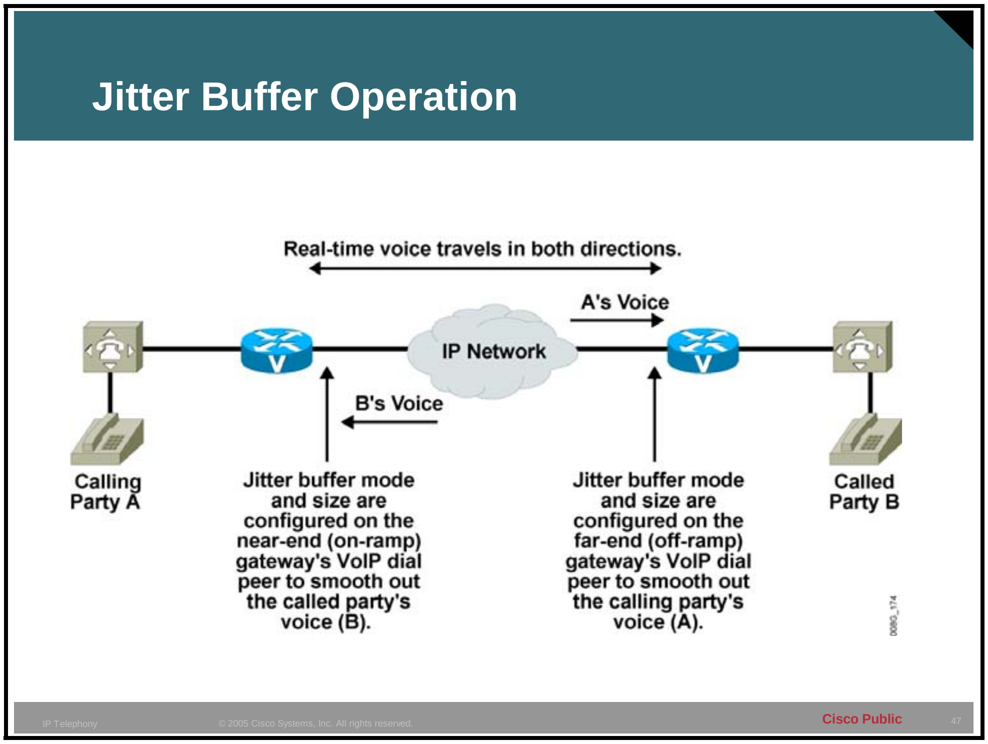
When a router receives an audio stream for VoIP, it must compensate for any jitter that it detects. The playout delay buffer mechanism handles this function. Playout delay is the amount of time that elapses between the time a voice packet is received at the jitter buffer on the DSP and the time a voice packet is played out to the codec. The playout delay buffer must buffer these packets and then play them out in a steady stream to the DSPs. The DSPs then convert the packets back into an analog audio stream. The playout delay buffer is also referred to as the dejitter buffer.



If the magnitude of jitter is so great that packets are received out of range of the playout delay buffer, the out-of-range packets are discarded and dropouts appear in the audio. For losses as small as one packet, the DSP interpolates what it calculates the audio should be, making the problem inaudible through the Cisco IOS Packet Loss Concealment (PLC) service.

Overcoming Jitter

Cisco voice networks compensate for jitter by setting up a buffer, called the "jitter buffer," on the gateway router at the receiving end of the voice transmission. This topic explains how to overcome jitter.



The jitter buffer receives voice packets from the IP network at irregular intervals. Occasionally, the voice packets are out of sequence. The jitter buffer holds the packets briefly, reorders them if necessary, and then plays them out at evenly spaced intervals to the decoder in the DSP on the gateway. Algorithms in the DSP determine the size and behavior of the jitter buffer based on user configuration and current network jitter conditions. The DSP uses this information to maximize the number of correctly delivered packets and minimize the amount of delay.

The size of the jitter buffer and the amount of delay is configurable by the user with the **playout-delay** command. Proper configuration is critical. If voice packets are held for too short a time, variations in delay may cause the buffer to underrun (become empty) and cause gaps in speech. However, packets that arrive at a full buffer are dropped, also causing gaps in speech.

To improve voice quality, the speech gaps are hidden by several different techniques that synthesize packets to replace those that were lost or not received in time. Depending on the contiguous duration of the gaps, the missing voice frames are replaced by prediction from the past frames (usually the last frame), followed by silence if the condition persists (for more than 30 to 50 ms, for example). The **show call active voice** command output gives buffer overflow and concealment statistics, which are a good indication of the network effect on audio quality.

Example: Overcoming Jitter

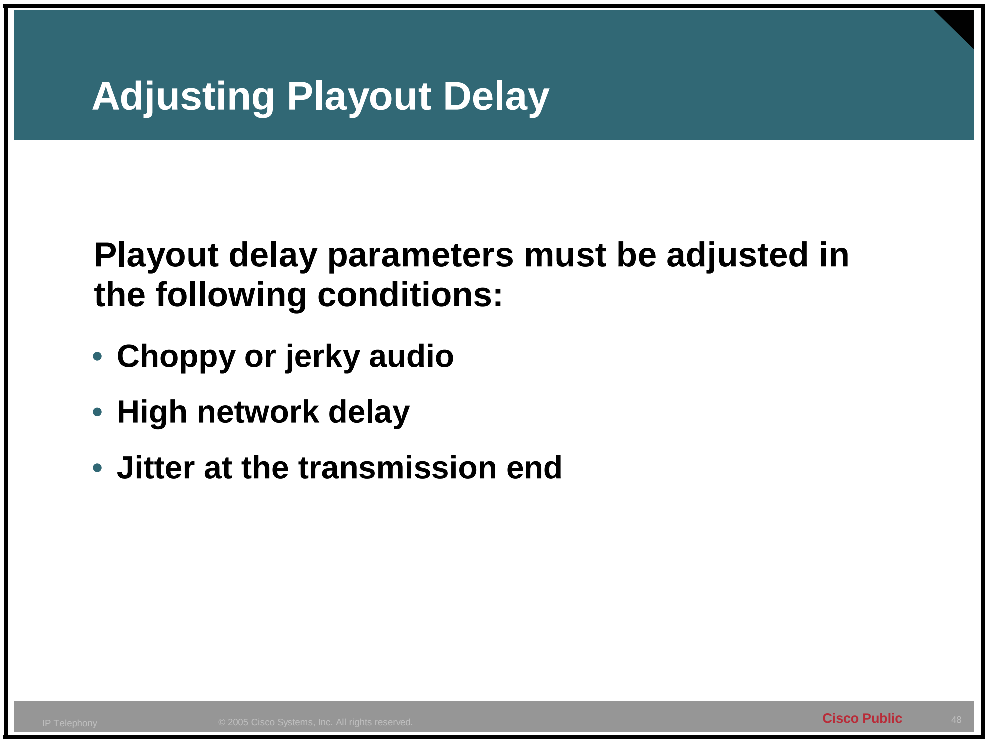
In an example that demonstrates how packets can be lost, a jitter buffer is configured with a maximum playout delay of 40 ms. On the network, packets are delayed from their source; perhaps a media server stops sending packets for 60 ms, or there is severe network congestion.

The jitter buffer empties while waiting for input from the network. Input does not arrive until *after* the maximum playout delay time is reached and there is a noticeable break in voice transmission. Now, the media server sends packets *to* the jitter buffer at a faster rate than the packets *leave* the jitter buffer; this makes the jitter buffer fill up. The jitter buffer discards subsequent packets, resulting in a choppy voice signal.

Even though the size of the jitter buffer is configurable, it is important to note that if the buffer size is too large, the overall delay on the connection may rise to unacceptable levels. You must weigh the benefit of improving jitter conditions against the disadvantage of increasing total end-to-end delay, which can also cause voice quality problems.

Adjusting Playout Delay Parameters

This topic lists the symptoms that lead to adjusting playout delay parameters.

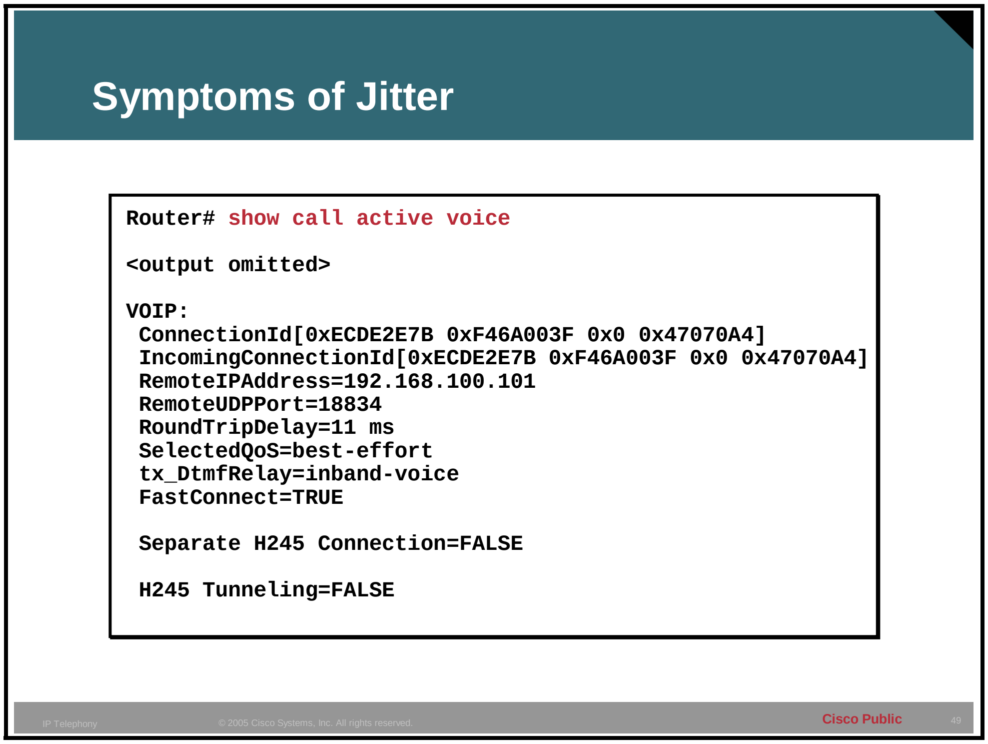


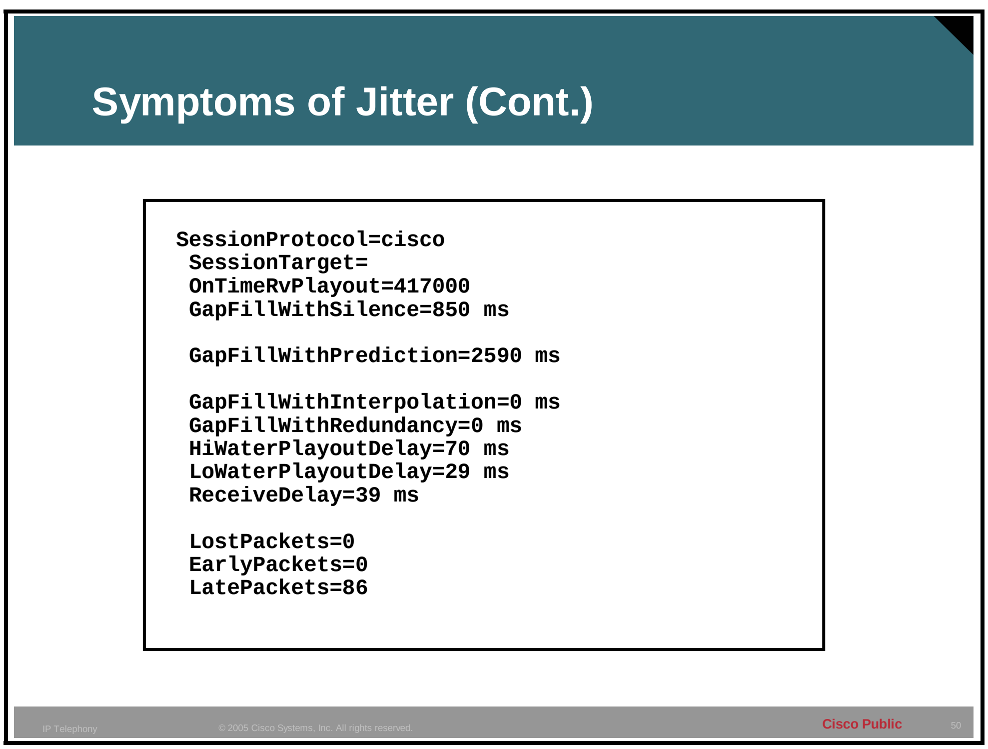
The conditions that require you to adjust playout delay parameters are as follows:

- **Choppy or jerky audio:** Gaps in speech patterns that produce choppy or jerky audio suggest that you should increase the minimum playout delay, increase the maximum playout delay, or both, if you are using adaptive mode. For fixed mode, you must increase the nominal value.
- **High network delay:** High overall network delay suggests that you should reduce the maximum playout delay in adaptive mode, or reduce the nominal delay in fixed mode. You must watch for loss of voice quality. The maximum delay value sets an upper limit on adaptive playout delay, which in many cases is the major contributor to end-to-end delay. In many applications, it may be preferable to have the system or the user terminate the call, rather than allow an arbitrarily large delay. The data received with jitter outside this limit will show up in the late packet count in the **show call active voice** playout statistics.
- **Jitter at the transmission end:** A noisy but well-understood network or interworking with an application that has lots of jitter at the transmission end, from a source such as a unified messaging server or interactive voice response (IVR) application, suggests selection of fixed mode.

Symptoms of Jitter on a Network

This topic provides examples of output for the **show call active voice** command, which can be used to determine the size of jitter problems.





The figure shows sample output for the **show call active voice** command. Several fields in the **show call active voice** command output that can help you to determine the actual size of the jitter problems are as follows:

- **ReceiveDelay:** The playout delay for jitter compensation plus the average expected delay after the frame is available for playout to the decoder. The current low-water mark and high-water mark statistics for the receive delay are available in the output.
- **GapFillWith:** These fields refer to the amount of *concealment*—or packet synthesizing—that took place in this call, to replace the voice packets that were lost or not received in time.
- **LostPackets:** The actual number of packets that were lost; that is, the packets *not* received at the egress gateway. This is detected using the **sequence number** field in the RTP packets.
- **EarlyPackets:** The actual number of packets that arrived *earlier* than the current minimum delay packet. They cause the dejitter algorithm to readjust the minimum delay packet used in jitter estimation.
- **LatePackets:** The actual number of packets that arrived later than the current playout delay setting. The information in these packets is discarded.

Average Jitter Statistics

```
# show call active voice

<output omitted>
.
.
.
VOIP:
 ConnectionId[0xECDE2E7B 0xF46A003F 0x0 0x47070A4]
 IncomingConnectionId[0xECDE2E7B 0xF46A003F 0x0 0x47070A4]
 RemoteIPAddress=192.168.100.101
 RemoteUDPPort=18834
 RoundTripDelay=26 ms
 SelectedQoS=best-effort
 tx_DtmfRelay=inband-voice
 FastConnect=TRUE

 Separate H245 Connection=FALSE

 H245 Tunneling=FALSE
```

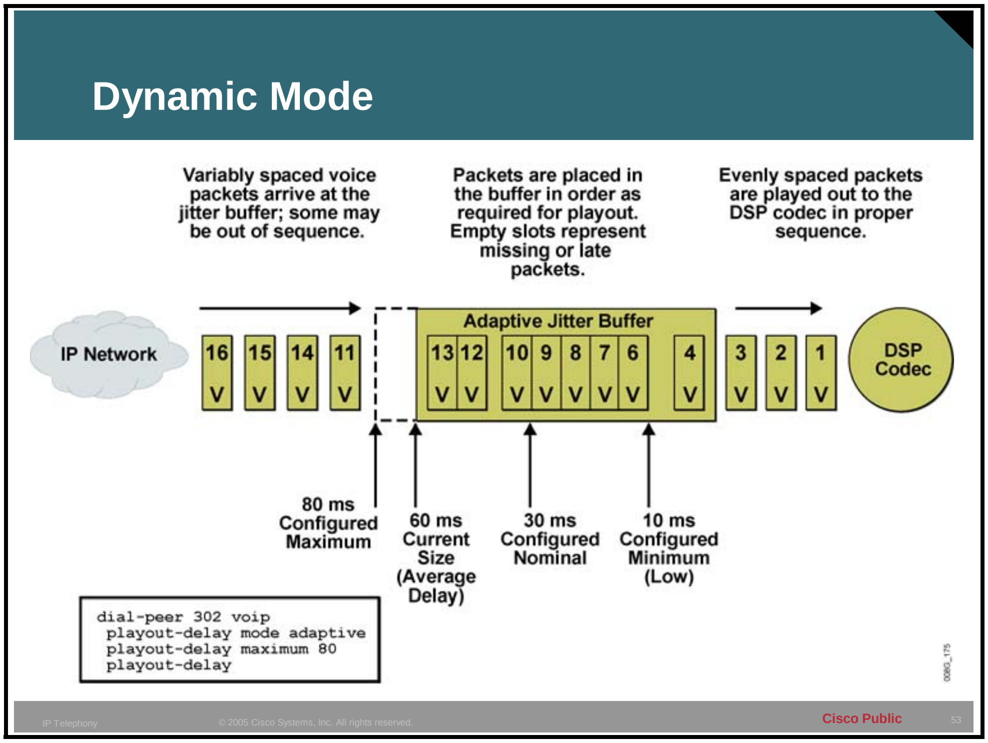
Average Jitter Statistics (Cont.)

```
SessionProtocol=cisco
SessionTarget=
OnTimeRvPlayout=482350
GapFillWithSilence=1040 ms     <------------  Increased
GapFillWithPrediction=3160 ms  <------------  Increased
GapFillWithInterpolation=0 ms
GapFillWithRedundancy=0 ms
HiWaterPlayoutDelay=70 ms
LowWaterPlayoutDelay=29 ms
ReceiveDelay=43 ms             <------------  Increased
LostPackets=0
EarlyPackets=0
LatePackets=105                <------------  Increased
```


The sample output in this figure displays average jitter statistics when poor voice quality was perceived on the network. The GapFillWithSilence line indicates that too many consecutive packets were lost or late, and the DSP could not predict and fill in the gaps. The GapFillWithPrediction line indicates that packets were late or lost, and the DSP filled in the lost audio with prediction. The ReceiveDelay line indicates the average one-way delay the packets are experiencing as per the time stamps in the RTP header. The LatePackets line indicates the number of packets that were too late to be processed by the DSP. As these fields increase, the playout delay buffer should be increased in size.

Dynamic Jitter Buffer

This topic describes the dynamic jitter buffer mode.



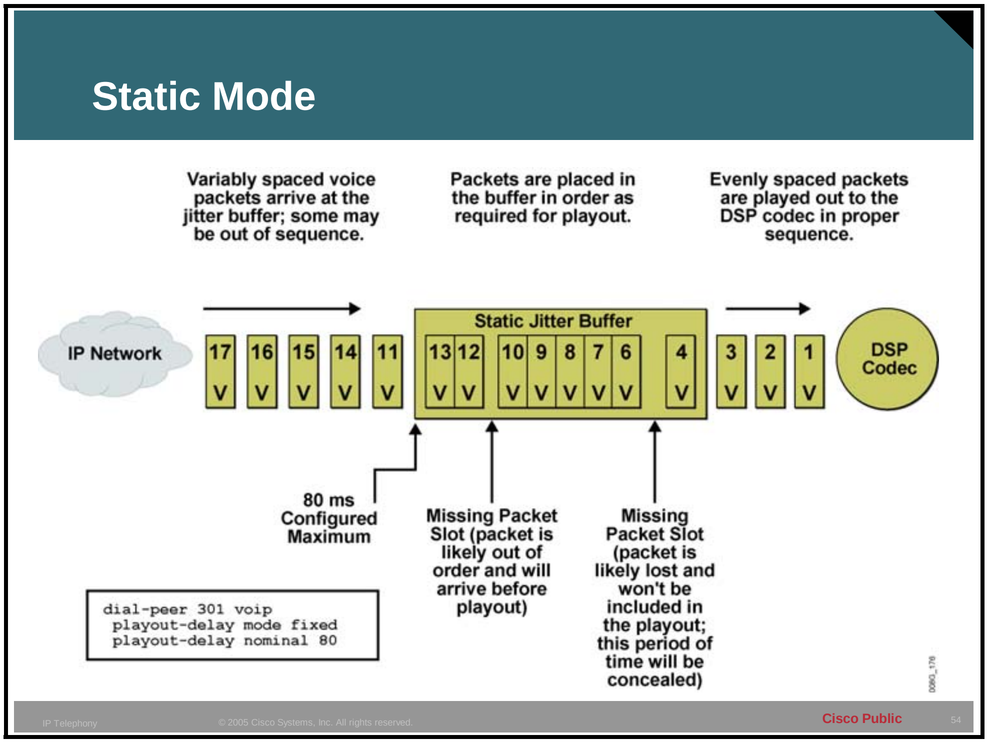
The **playout-delay** command allows you to select a jitter buffer mode (**static** or **dynamic**) and specify certain values that are used by the DSP algorithms to adjust the size of the jitter buffer. During a voice call, the algorithms read time stamps in the RTP headers of sample packets to determine the amount of delay that the jitter buffer will apply to an average packet; that is, as if there is no jitter at all in the network. This is called the average delay.

Example: Dynamic Jitter Buffer

When you configure the **playout-delay mode adaptive** option, the DSP algorithms in the codec take samples throughout the voice call and adjust the value of the average delay as network jitter conditions change. The size of the jitter buffer and the amount of delay applied are adjusted upward or downward, as needed. This adjustment ensures the smooth transmission of voice frames to the codec within the minimum and maximum limits that you configure. The algorithms are designed to *slowly* reduce the amount of delay and *quickly* increase the amount of delay during adjustment. As a result, voice quality is achieved at the risk of longer delay times.

Static Jitter Buffer

This topic describes the static mode buffer.



When you configure the **playout-delay mode fixed** option, you can specify the nominal delay value, which is the amount of playout delay applied at the beginning of a call by the jitter buffer. This is also the maximum size of the jitter buffer throughout the call.

Example: Static Jitter Buffer

Configuring static or "fixed" playout delay limits the size of the buffer. The figure shows an example of adjusting the nominal size of the buffer to 80 ms. Care should be taken when using the jitter buffer in static mode. Variations in arrival times of voice packets may be put at risk when network conditions change. In addition, if the static jitter buffer is configured to be too large, overall one-way delay is increased, exacerbating delay and echo problems.

