



# Nextport Voice Tuning White Paper

---

## Introduction

The Nextport Voice Tuning White Paper describes the new Voice Tuning Configuration Capability (also known as Voicecap) that has been added to the Cisco AS5350, AS5400, AS5400HPX, and AS5850 voice gateways. This white paper discusses the reasons for adding the new Voice Tuning Configuration Capability and explains its usage.

This white paper presents the following information:

- [Voice Tuning Configuration Capability - page 2](#)
- [VoIP Gateway Architecture - page 2](#)
  - [SPE Functionality - page 3](#)
- [Typical Problems with Switched Telephone Networks - page 4](#)
  - [Power Variations - page 4](#)
  - [Echo - page 4](#)
  - [Variations in a Switched Telephone Network - page 5](#)
- [Echo Cancellation - page 6](#)
  - [Echo Return Loss - page 6](#)
  - [Echo Canceller Adaptive Filter - page 6](#)
- [Trade-offs - page 7](#)
- [Primary Uses for Voicecap - page 8](#)
  - [Setting Minimum ERL - page 8](#)
  - [Adding Attenuation - page 8](#)
  - [Dynamic Attenuation - page 9](#)
  - [Comfort Noise Generator Enable/Disable - page 9](#)
- [Voicecap Upgrades and Availability - page 10](#)
- [Glossary - page 11](#)

## Voice Tuning Configuration Capability

The large number of Voice over IP (VoIP) networks, and the widely diversified types of installation environments for them, has given rise to a need for much more flexible and configurable VoIP gateways. The Nextport Voice Tuning Configuration Capability was created to provide greater flexibility and configurability to VoIP gateway service providers.

The Nextport Voice Tuning Configuration Capability was designed for VoIP gateway systems that use Nextport Digital Signal Processors (DSPs) as the service-processing elements (SPEs) to process voice signals.

The Nextport Voice Tuning Configuration Capability is very similar to a modem configuration string and should be familiar to anyone who has worked with modem configuration strings. A modem configuration string is often called a Modemcap; therefore, the Nextport Voice Tuning Configuration Capability is called a Voicecap.

The Voicecap allows a user to have low level access to the SPEs to facilitate actions like fine tune the voice signal power levels and adjusting echo canceller performance. One of the primary uses of the Voicecap is to reduce as much of the echo produced by a voice signal as possible.

The Voicecap is configured using the Cisco IOS™ command line interface (CLI). Refer to *NextPort Voice Tuning and Background Noise Statistics Feature Module* document for details on the Voicecap configuration procedures.

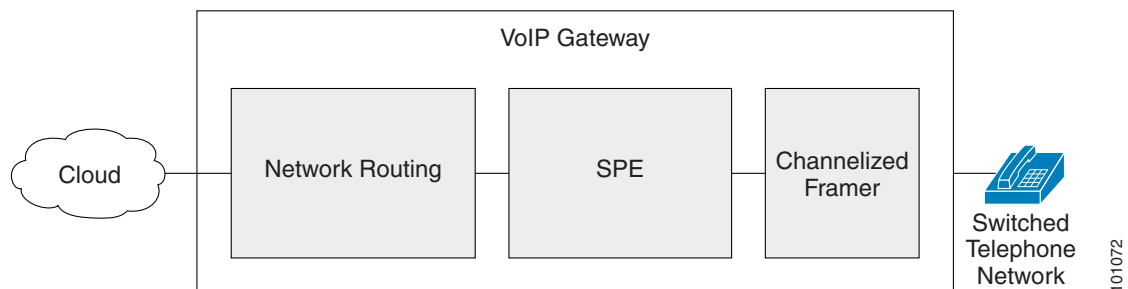
## VoIP Gateway Architecture

In a digital telephone system, the VoIP gateway receives voice signals from a switched telephone network, converts the analog signals to digital signals, and sends them to a network cloud or core. It also receives voices signals from the network core and sends them to the switched telephone network.

For our purposes, the VoIP gateway has three blocks that perform different functions; the Channelized Framer block, the SPE block, and the Network Routing block.

Figure 1 shows a VoIP gateway connected to a switched telephone network and to the network cloud or core. It also shows where these three blocks fit into the VoIP gateway.

**Figure 1** VoIP Gateway



In a Nextport system, the interface between the switched telephone network and the VoIP gateway is usually a channelized TDM framer card. The framer card receives a time division multiplexed (TDM) signal from the switched telephone network and separates the aggregate TDM signal into multiple, separate, single voice stream channels.

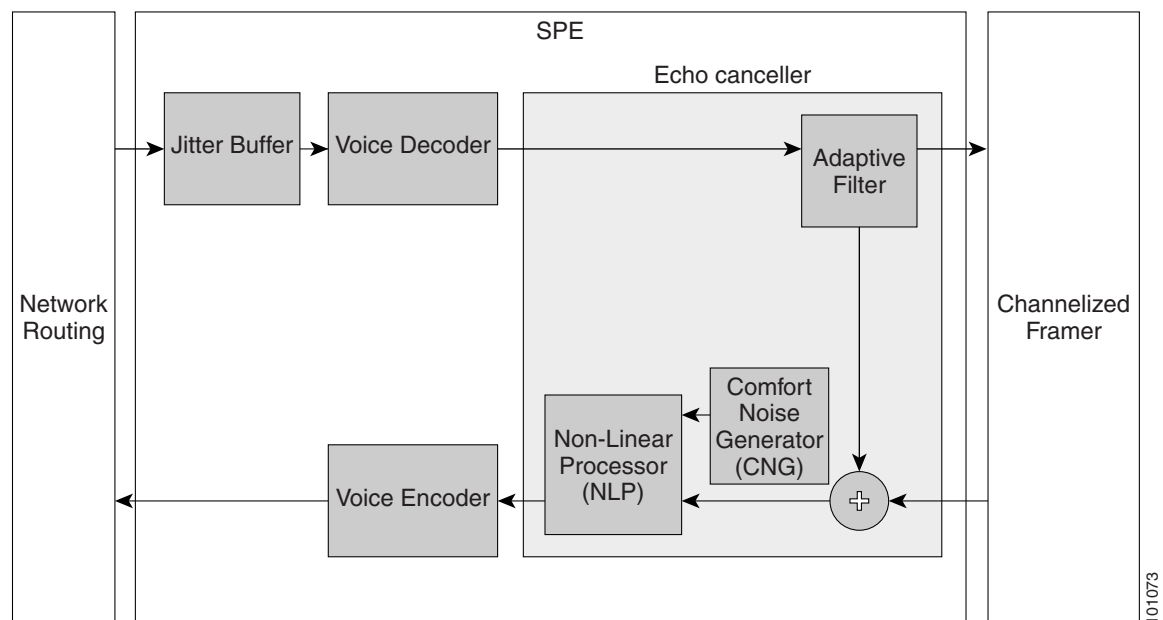
Each voice stream is then passed to the SPE. The SPE converts the voice stream from a TDM format to a packet format. The SPE sends the voice stream packets to the network routing block, which sends the voice stream to the network cloud.

The major portion of this new configurability is concerned with ensuring that the SPE works in harmony with the public switched telephone network (PSTN) to which it is connected.

## SPE Functionality

For the purposes of this White paper, the SPE has three primary functional components; the jitter buffer, the voice encoder/decoder, and the echo canceller. [Figure 2](#) shows the SPE functional components.

**Figure 2** SPE



The jitter buffer removes the timing discrepancies that are introduced by the packet network.

The voice encoder/decoder pair handles compression/decompression to the voice signal and performs other functions including the following functions:

- Voice Activity Detection (VAD)—detects silence periods that should not be transmitted
- Comfort Noise Generation (CNG)—replicates noise indicated in silence packets
- Concealment—covers for lost packets

The echo canceller removes echo that occurs in the telephone network system. One of the primary purposes of the Voicecap is to make optimization of the echo canceller possible.



### Note

The echo canceller also does CNG, but this should not be confused with the CNG done by the voice encoder/decoder pair.

Echo canceller CNG is discussed in the [Comfort Noise Generator Enable/Disable](#) section on [page 9](#).

# Typical Problems with Switched Telephone Networks

The average power level of voice signals as they travel to and from a VoIP gateway can vary greatly depending on the switched telephone network that sends them. It is these wide variations in the quality of switched telephone networks that make it necessary to have more flexible voice tuning capabilities.

## Power Variations

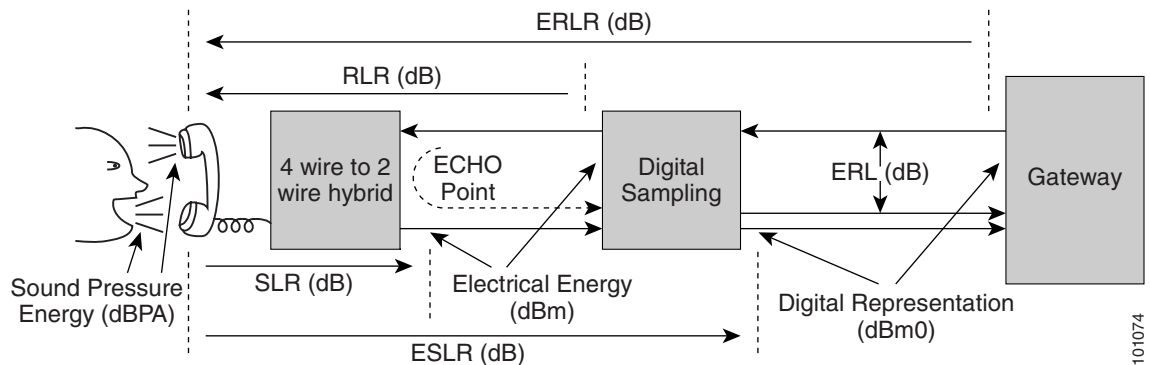
When a voice signal reaches the VoIP gateway, it has been changed from an analog signal to a digital signal, and the power level has been changed into a digital representation of the power level of the original analog signal. However, there may be gains or losses to the signal as it travels from the sender to the VoIP gateway. This change in the power level between the talker and the VoIP gateway is called the Equipment Send Loudness Rating (ESLR).

For example, if the average power level of the voice signal when it leaves the telephone mouthpiece is  $-4.7\text{dBPA}$ , and the average power level of the voice signal when it reaches the VoIP gateway is  $-14.7\text{dBm0}$  after telephone network gains and losses, then the ESLR is  $-10\text{dB}$ .

As the voice signal travels from the VoIP gateway to the listener, there are also gains and losses, and these gains and losses can vary greatly depending on the telephone network. The change in the power level between the VoIP gateway and the listener is called the Equipment Receive Loudness Rating (ERLR).

Figure 3 illustrates the path between a talker or listener and a VoIP gateway.

**Figure 3** Path between Talker or Listener and VoIP Gateway



## Echo

Echo occurs when power from the voice signal that is sent to the listener from the VoIP gateway is reflected back to the talker. Echo can be caused by impedance mismatching in the analog portion of the telephone network. The amount of echo that is reflected depends on the power level of the signal that is sent and the degree of impedance mismatching that exists in the telephone network.

The ratio of the power level of the signal that is sent to the power level of the echo that is reflected is called the Echo Return Loss (ERL). The ERL also varies greatly depending on the telephone network that is connected to the VoIP gateway.

## Variations in a Switched Telephone Network

There are a large number of variations in a telephone network that can be addressed by optimizing the VoIP gateway using a Voicecap. However, there are trade-offs as well as improvements that will occur when attempting to optimize a VoIP gateway using a Voicecap.

To understand these trade-offs, we need to understand what the characteristics of an ideal switched telephone network would be when it interfaces with a VoIP gateway.

The better the conditions of the switched telephone network are, the better the overall signal quality will be in the entire system. As a general rule, a VoIP network cannot produce better voice quality signals than what is sent to it by the switched telephone network. The more organized and balanced that the switched telephone network is, the better the VoIP network will be, and the easier it will be to configure the gateway.

The ideal switched telephone network for interfacing with a VoIP gateway would have the following characteristics:

- The average power level of the voice signal on the switched side of the network would be equal to the average power level of the voice signal on the packet side of the network.
- The Equipment Send Loudness Rating (ESLR) on the switched side of the network would be equal to the ESLR on the packet side of the network.
- The power levels of noise signals on the line would be significantly lower than the power levels of voice signals on the line so that the VoIP gateway could distinguish noise signals from voice signals.
- The Echo Return Loss (ERL) for the telephone network would be as high as possible.
- The echo return delay would be as short as possible.

### ESLR Balancing

The closer the ESLRs are on both sides of the network, the easier it is to distinguish echo from valid voice signals. Where ESLRs are different, the Voicecap allows you to make adjustments to improve performance.

### Noise Recognition

The delineation between noise and voice signals can be done using a power level threshold setting. All incoming signals with power levels above this threshold are treated as voice signals and all incoming signals with power levels below this threshold are treated as noise. A standard default threshold already exist for this; however, the Voicecap allows you to adjust this threshold for any given network.

# Echo Cancellation

One of the primary and most effective uses of the Voicecap feature is adjusting the parameters of the echo canceller. Echo cancellation is performed by the echo canceller block in the SPE (see [Figure 2](#)). The Voicecap configuration allows you to fine tune the echo canceller to work at an optimum level.

## Echo Return Loss

As stated earlier, Echo Return Loss (ERL) is the ratio of the power level of the transmitted voice signal to the power level of the echo signal that is generated by the VoIP gateway. ERL varies greatly depending on the switched telephone network that is connected to the VoIP gateway.

In an absolutely perfect case, there would be no echo at all, which would yield an ERL of infinity, but there is always going to be echo whenever you have an analog trunk line connected to a digital network.

The Voicecap allows the user to set the minimum ERL threshold for the echo canceller. The performance of the VoIP gateway will degrade sharply when the minimum ERL threshold is not met.

**Note**

---

When using the Voicecap to configure echo cancellation, it is not the average ERL that is important, but the worst-case ERL.

---

If the worst-case ERL is smaller than the set minimum ERL threshold, lowering the threshold will allow the VoIP gateway to perform better at a wider range of voice signal power levels.

## Echo Canceller Adaptive Filter

The echo canceller in the SPE has an adaptive filter (see [Figure 2](#) or [Figure 5](#)) that learns to remove echo by modeling the echo that it detects. This model has a finite duration referred to as the echo canceller coverage time. The adaptive filter cannot learn the nature of any echo energy that does not return during this coverage time. The shorter that the coverage time setting is, the more accurate the model will be.

Echo becomes more noticeable as the echo delay increases. At long delays, even a low amount of echo is noticeable. The echo return delay must be as short as possible, and the echo canceller coverage time must be set to the worst-case value of the echo return delay at which it is expected to perform well.

# Trade-offs

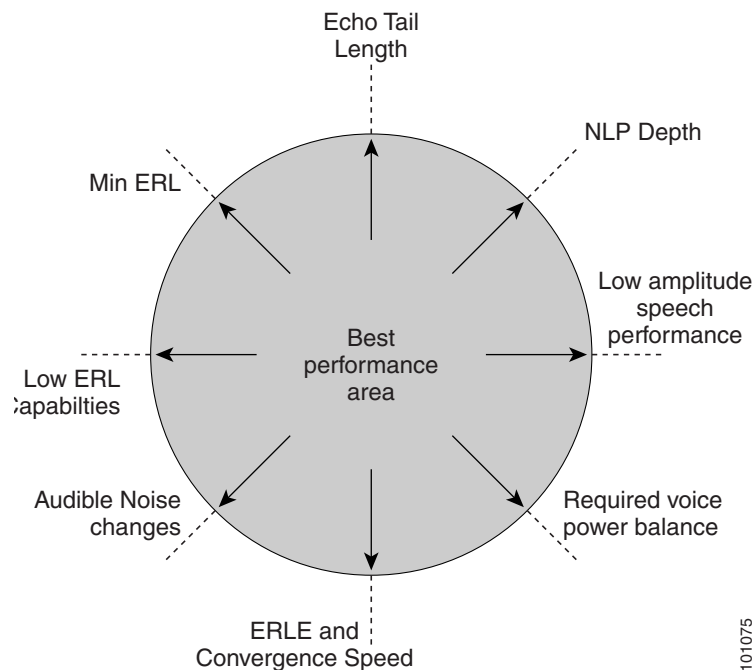
There are many different parameters that can be set using Voicecap configuration to help optimize a telephone network. Determining the best values for the Voicecap settings can be very complex.

As a general rule, the initial installation of a new gateway should be done without applying any Voicecap. The default Voicecap settings are optimized for the most common situations.

Changing the Voicecap settings is almost always a trade-off situation. [Figure 4](#) illustrates some of these trade-offs. Adjusting the Voicecap settings only moves the area over which the system performs optimally. The area does not increase.

If unique situations arise where Voicecap configuration seems necessary, please contact the Cisco Technical Assistance Center (TAC).

**Figure 4** Trade-offs using Voicecap Configuration



# Primary Uses for Voicecap

The following sections discuss the primary uses for the Nextport Voice Tuning Configuration Capability or Voicecap. The Voicecap has uses beyond this, but these uses are rare. It is best to contact Cisco Technical Assistance Center (TAC) to address those features.

## Setting Minimum ERL

The minimum (or worst-case) ERL setting was discussed previously in the [Echo Return Loss](#) section.

The Voicecap allows the user to set the minimum ERL threshold to a specific value in decibels (dB). Performance of the VoIP gateway will degrade sharply when ERLs lower than this threshold are experienced.

However, setting the minimum ERL threshold is a trade-off. The lower the minimum ERL setting is, the more critical the balance between the transmitted signals and the received signals becomes. If the minimum ERL is set too low, and the received signal is lower in power than the transmitted signal, the received signal may become distorted. This is referred to as clipping or squelching.

## Adding Attenuation

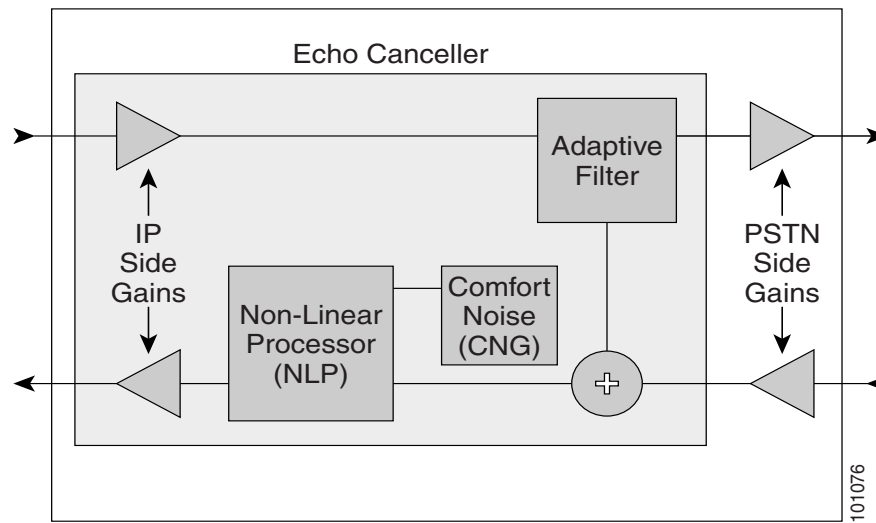
Voicecap allows the user to add gain or attenuation on either side of the of the VoIP gateway—the IP side or the PSTN side (see [Figure 5](#)).

Adding attenuation can be used to achieve the following goals:

- ensuring that both the transmitted and the received signals are at the same power levels at the echo canceller block
- ensuring that the network ERL conforms to the minimum ERL threshold setting at the echo canceller block

Setting the initial attenuation level can also be used in conjunction with the dynamic attenuation feature discussed in the next section.

Figure 5 Echo Cancellor



## Dynamic Attenuation

When the user adds attenuation on the PSTN side of the telephone network to ensure that the minimum ERL setting is sufficient, the dynamic attenuation feature can automatically remove this added attenuation during calls where the minimum ERL setting is sufficiently met.

The dynamic attenuation feature is useful for networks that experience significantly varying ERLs from call to call. This can occur if several different switched telephone network providers are using the same VoIP gateway and their loss plans either do not exist or do not coincide.

The dynamic attenuation feature is a workaround to be used only when circumstances warrant it, because it can add an unneeded variable to a stable network. By default the dynamic attenuation feature is disabled.

## Comfort Noise Generator Enable/Disable

When the SPE is transmitting a voice signal, but is not simultaneously receiving a voice signal, it still receives the echo of the transmitting voice signal as well as background noise. During these periods, the echo canceller attempts to remove as much of the echo as possible, using a device called the non-linear processor (NLP). The NLP adds attenuation to the receive path whenever there is a transmitting voice signal, but no receiving voice signal. The problem with this is that the added attenuation affects not only the echo, but the background noise as well.

This sudden drop in the noise level is audible and can sound like the phone has been disconnected. To lessen this effect, the comfort noise generator adds noise when the NLP is engaged. The need or desirability of comfort noise varies with installation, configuration, and user demographics. Therefore, the echo canceller's comfort noise generator can be enabled or disabled using the Voicecap.

## Voicecap Upgrades and Availability

The Voicecap makes software upgrades easy. The SPE has a separate software image than the network routing components that use an IOS software image. Since the SPE software image is separate from the IOS image, it can be upgraded without changing the IOS software.

The Voicecap has been designed, so that the parameters pass directly to the SPE. When the SPE software is upgraded, any new Voicecap settings can be used without having to upgrade the IOS software.

An early release of the Nextport Voicecap is available in IOS Releases 12.3(2)T and 12.3(1)M. In these early releases, all Voicecap settings are configured by directly setting locations in the SPE.

This initial release of Voicecap does not perform conversions or range checking. Therefore, special care must be taken. Future IOS releases will perform conversions and range checking.

# Glossary

attenuation	Decrease of signal power in a transmission; a positive attenuation is a negative gain.
CLI	Command Line Interface
CNG	Comfort Noise Generator
dB	Decibels
DSP	Digital Signal Processor
ERLR	Equipment Receive Loudness Rating
gain	Increase of signal power in a transmission; a positive gain is a negative attenuation.
ESLR	Equipment Send Loudness Rating
IOS	Cisco Internetwork Operating System, Cisco IOS™
IP	Internet Protocol
NLP	Non-Linear Processor
PSTN	Public Switched Telephone Network
SPE	Service-processing Element
TAC	Cisco Technical Assistance Center
TDM	Time Division Multiplexing
VAD	Voice Activity Detection
Voicecap	Nextport Voice Tuning Configuration Capability
VoIP	Voice Over IP