



Secure Voicemail Transcription: Speech to Text with Cisco Unity Connection Voice Messages

This solution overview explains how the Cisco[®] SpeechView (speech-to-text) feature of the Cisco Unity[®] Connection unified messaging solution handles security for voice message transcriptions. You will learn about:

- Security during message transport
- Security measures in effect with the third-party external transcription service

You do not need any prior knowledge of SpeechView or Cisco Unity Connection to understand the security discussion in this overview.

Cisco SpeechView Introduction

SpeechView converts voice messages to text and delivers the text version of the voice message to your email inbox, allowing you to read your voice messages and take immediate action. The application is a feature of Cisco Unity Connection, so the original audio version of each voice message remains available to you anywhere, anytime. SpeechView transcribes and sends voice messages within minutes of being left in your Cisco Unity Connection voice mailbox—you do not need to learn any commands or take special action to receive text versions of your voice messages.

You can learn more about SpeechView at <http://www.cisco.com/go/speechview>.

Challenge

Accurate voicemail transcription services generally require human intervention to transcribe complex voice messages, because even the most powerful computers cannot accurately decipher the intricacies in the way we speak to each other in our everyday language. Therefore, many organizations seeking voicemail transcription are faced with a choice: use an automated service or use human assistance to improve accuracy. Often, transcription services involve sending voice messages to a third-party service, causing security concerns within most organizations regarding the measures in place that secure the message during transport to and from the service and protection of the message content during transcription.

Business Benefits

Cisco SpeechView solves these challenges and removes the trade-offs that organizations must make with typical voicemail transcription services. Cisco has partnered with a third-party external transcription service, SpinVox Ltd. (a subsidiary of Nuance Communications, Inc.), to provide accurate, secure transcriptions of voice messages left in Cisco Unity Connection voice mailboxes, which are then delivered to your email inbox. Easy to use and secure, SpeechView improves responsiveness.

SpeechView benefits include:

- You can learn who called and what they said at a glance.
- You do not need to dial in to retrieve messages, or take notes on the message content.
- You have nothing new to learn—your experience is the same as for regular email messages.
- Messages delivered in both audio and text format allow you to decide the best way to manage them.
- You can prioritize and sort both voice and email messages from a single email inbox.

SpeechView security features include:

- All data that is transmitted is encrypted.
- Security measures that comply with ISO certification and data protection and privacy protocols are applied at the physical, network, and application layers.
- User data is kept anonymous.

The next sections provide details about how Cisco and SpinVox partner to provide security throughout the entire message transcription process.

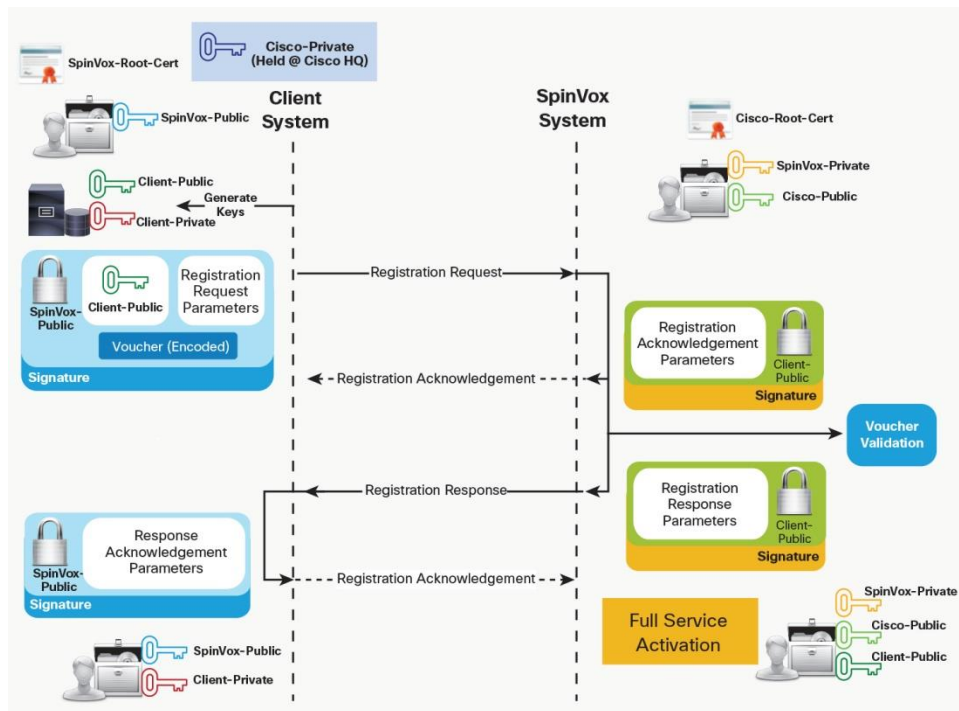
Initial Service Registration

When SpeechView is initially configured, the Cisco Unity Connection server registers with SpinVox. Figure 1 illustrates the following process:

1. The Cisco Unity Connection server generates the Client-Private and Client-Public keys.
 - a. Cisco and root certifications are 2048-bit RSA.
 - b. SpinVox and client keys are 1024-bit RSA.
 - c. Client keys may be refreshed periodically by reregistering through the Cisco Unity Connection administration GUI.
2. The Client-Public key, registration request, and voucher are packaged in a signed Secure/Multipurpose Internet Mail Extensions (S/MIME) message and delivered to SpinVox. Refer to the “Outbound Request” section of this document for an example of a registration request.
3. SpinVox acknowledges the message and validates the voucher.
4. When SpinVox has validated the voucher, a registration response is sent back to the Cisco Unity Connection server. Refer to the “Inbound Response” section of this document for an example of a registration response.
5. When the Cisco Unity Connection server receives the response, the SpeechView feature is active and ready to begin transmission of messages to SpinVox.

For specifications of the S/MIME standard, please visit the IETF RFC 3851 at <http://tools.ietf.org/html/rfc3851>.

Figure 1. Cisco SpeechView Service Registration



Outbound Request

The following is an example of the XML file sent to SpinVox in the initial registration request from Cisco Unity Connection:

```

<?xml version="1.0" ?>
<request>
<interface-version>10</interface-version>
<id>1234567890abcdefghijklmnopqrstuvwyz</id>
<registration>
<registration-request>
<registration-acknowledgement>>true</registration-acknowledgement>
<enterprise-name>Acme</enterprise-name>
<name>Joe Bloggs</name>
<contact-phone-number>15551234567</contact-phone-number>
<contact-email>joe@acme.com</contact-email>
<voucher-code>1234567890-abcdef-0987654321</voucher-code>
<language>en-US</language>
<language>es-ES</language>
<codec>G711-A</codec>
<date-timestamp>Wed, 30 Jan 2008 16:34:33 +0000</date-timestamp>
<reply-address>example@acme.com</reply-address>
</registration-request>
</registration>
</request>

```

Inbound Response

The following is an example of the XML file received by Cisco Unity Connection from SpinVox indicating a successful registration:

```
<?xml version="1.0" ?>
<request>
<interface-version>10</interface-version>
<id>1234567890abcdefghijklmnopqrstuvwxyz</id>
<registration>
<registration-response>
<response-acknowledgement>>true</response-acknowledgement>
<enterprise-name>Acme</enterprise-name>
<registration-request-id>zyxwvu9876</registration-request-id>
<enterprise-identification>1725Acme</enterprise-identification>
<status>Accepted</status>
<text>Activation successful</text>
<language>en-US</language>
<injection-address>1725Acme@integration.partnerpartner.com</injectionaddress>
<codec>G711-A</codec>
</registration-response>
</registration>
</request>
```

Message Flow

After successfully registering with SpinVox, the Cisco Unity Connection users configured for the SpeechView feature will begin to receive voice message transcriptions. All message flow to and from SpinVox uses S/MIME. The process follows:

1. A Cisco Unity Connection user receives a voice message.
2. Cisco Unity Connection packages the message, encrypts it using the SpinVox-Public key, and signs it with the Client-Private key. The following is an example of an encrypted message sent by Cisco Unity Connection to SpinVox:

Encrypted

```
Date: Wed, 30 Jan 2008 16:34:33 +0000
Message-Id: <1234567890abcdefg@cisco.com>
From: example@cisco.com
To: enterprise-id@cisco-unity.integration.partner.com
Subject: New message
Content-Type: multipart/mixed; boundary="----
=_NextPart_000_0001_01C7F52D.834C7D30"

Content-Transfer-Encoding: 7bit

-----=_NextPart_000_0001_01C7F52D.834C7D30 Content-Type: application/pkcs7-mime;
smime-type=enveloped-data;

TWAaqamd2oYigKrIcSPZCyBruZJil0tAK8mNJKdNgBNHarS0mvhpca0CtKlTyW2wAEaS3kax5Aro
KABIKiRJKiBIF9aEAoAkSZIKiQKAJEMSJIKCgCRJkiSJAoAkSZIKCYJ0YU0oAEiSJEmSKABIKiRJ
```

<snip>
KABIKiRJkiBIF9aEAoAkSZIKiQKAJEmSJIKCgCRJkiSJAoAkSZIKCYJ0YU0oAEiSJEmSKABIKiRJ
TWAaqamd2oYigKrIcSPZCyBruZJil0tAK8mNJKdNgBNHarS0mvhpca0CtKlTyW2wAEaS3kax5Aro

-----=_NextPart_000_0001_01C7F52D.834C7D30
Content-Type: application/pkcs7-mime; smime-type=signed-data;

KABIKiRJkiBIF9aEAoAkSZIKiQKAJEmSJIKCgCRJkiSJAoAkSZIKCYJ0YU0oAEiSJEmSKABIKiRJ
TWAaqamd2oYigKrIcSPZCyBruZJil0tAK8mNJKdNgBNHarS0mvhpca0CtKlTyW2wAEaS3kax5Aro
<snip>
TWAaqamd2oYigKrIcSPZCyBruZJil0tAK8mNJKdNgBNHarS0mvhpca0CtKlTyW2wAEaS3kax5Aro
KABIKiRJkiBIF9aEAoAkSZIKiQKAJEmSJIKCgCRJkiSJAoAkSZIKCYJ0YU0oAEiSJEmSKABIKiRJ

The following is the preceding message decrypted:

Decrypted

Date: Wed, 30 Jan 2008 16:34:33 +0000
Message-Id: <1234567890abcdefg@cisco.com>
From: example@cisco.com
To: enterprise-id@cisco-unity.integration.partner.com
Subject: New message
Content-Type: multipart/mixed; boundary="-----
_NextPart_000_0001_01C7F52D.834C7D30"

Content-Transfer-Encoding: 7bit

This is a multi-part message in MIME format.

-----=_NextPart_000_0001_01C7F52D.834C7D30
Content-Type: audio/wav
Content-Transfer-Encoding: base64
Content-Duration: 18
Content-Disposition: inline; filename="example.wav"
UklGRi9QAABXQVZFZm10IBQAAAAxAAEAQB8AAAFkGAABBAAAAAGBAAWZHY3QEAAAaWIkBAGRhdGH7
TWAaqamd2oYigKrIcSPZCyBruZJil0tAK8mNJKdNgBNHarS0mvhpca0CtKlTyW2wAEaS3kax5Aro
<snip>
KABIKiRJkiBIF9aEAoAkSZIKiQKAJEmSJIKCgCRJkiSJAoAkSZIKCYJ0YU0oAEiSJEmSKABIKiRJ

-----=_NextPart_000_0001_01C7F52D.834C7D30
Content-Type: text/xml
Content-Disposition: inline; filename="message.xml"

```
<?xml version="1.0" ?>  
<request>  
  <interface-version>10</interface-version>  
  <id>1234567890abcdefghijklmnopqrstuvwxy</id>  
  <conversion>  
    <conversion-request>  
      <conversion-acknowledgement>>false</conversion-acknowledgement>  
      <enterprise-identification>site1234</enterprise-identification>
```

```
<message-class>Voicemail</message-class>
<audio-max-length>180</audio-max-length>
<audio-offset>100</audio-offset>
<confidence>
  <threshold>95</threshold>
  <low-confidence-action>0</low-confidence-action>
</confidence>
<user-language>en-US</user-language>
<message-language>en-US</message-language>
<alt-language-support>true</alt-language-support>
<priority>1</priority>
<text-max-length>500</text-max-length>
<result-case>proper</result-case>
<return-audio>>false</return-audio>
<source-device>cell</source-device>
<user-information>
  <calling-party>K8mNJKdNgBNHarS0mvhpca0Ct</calling-party>
  <called-party>AkSZIkCYJ0YU0oAEAkSZIkCYJ</called-party>
</user-information>
</conversion-request>
</conversion>
</request>
```

-----=_NextPart_000_0001_01C7F52D.834C7D30--

3. The message is transcribed and returned to the Cisco Unity Connection server encrypted with the Client-Public key and signed with the SpinVox-Private key.

The following is an example of the encrypted response followed by the same message decrypted:

Encrypted

```
Date: Wed, 30 Jan 2008 16:34:33 +0000
Message-Id: <1234567890abcdefg@cisco.com>
From: cisco-us-cisco@integration.partner.com
To: example@example.com
Subject: New message
Content-Type: multipart/mixed; boundary="-----
=_NextPart_000_0001_01C7F52D.834C7D30"
```

Content-Transfer-Encoding: 7bit

-----=_NextPart_000_0001_01C7F52D.834C7D30 Content-Type: application/pkcs7-mime; smime-type=enveloped-data;

```
TwAAqamd2oYigKrIcSPZCyBruZJil0tAK8mNJKdNgBNHarS0mvhpca0CtKlTyW2wAEaS3kax5Aro
KABIKiRJKiBIF9aEAoAkSZIKiQKAJEmSJIKCgCRJkiSJAoAkSZIkCYJ0YU0oAEiSJEmSKABIKiRJ
<snip>
KABIKiRJKiBIF9aEAoAkSZIKiQKAJEmSJIKCgCRJkiSJAoAkSZIkCYJ0YU0oAEiSJEmSKABIKiRJ
TwAAqamd2oYigKrIcSPZCyBruZJil0tAK8mNJKdNgBNHarS0mvhpca0CtKlTyW2wAEaS3kax5Aro
```

-----=_NextPart_000_0001_01C7F52D.834C7D30 Content-Type: application/pkcs7-mime;
smime-type=signed-data;

KABIKiRJkiBIF9aEAoAKSZIKiQKAJEmSJIKCgCRJkiSJAoAKSZIKCYJ0YU0oAEiSJEmSKABIKiRJ
TwAAqamd2oYigKrIcSPZCyBruZJil0tAK8mNJKdNgBNHarS0mvhpca0CtKlTyW2wAEaS3kax5Aro
<snip>
TwAAqamd2oYigKrIcSPZCyBruZJil0tAK8mNJKdNgBNHarS0mvhpca0CtKlTyW2wAEaS3kax5Aro
KABIKiRJkiBIF9aEAoAKSZIKiQKAJEmSJIKCgCRJkiSJAoAKSZIKCYJ0YU0oAEiSJEmSKABIKiRJ

Decrypted

Date: Wed, 30 Jan 2008 16:34:33 +0000
Message-Id: <1234567890abcdefg@cisco.com>
From: cisco-us-cisco@integration.partner.com
To: example@example.com
Subject: New message Content-Type: multipart/mixed; boundary="-----
_NextPart_000_0001_01C7F52D.834C7D30"
Content-Transfer-Encoding: 7bit
This is a multi-part message in MIME format.

-----=_NextPart_000_0001_01C7F52D.834C7D30
Content-Type: text/xml
Content-Disposition: inline; filename="message.xml"
<?xml version="1.0" ?>
<request>
<interface-response>1.0.1</interface-response>
<id>1234567890abcdefghijklmnopqrstuvwxy</id>
<conversion>
 <conversion-response>
 <response-acknowledgement>>false</response-acknowledgement>
 <enterprise-identification>site1234</enterprise-identification> <text>"This
is the converted text" </text>
 <count>
 <word>5</word>
 <character>26</character>
 </count>
 <user-information>
 <calling-party>K8mNJKdNgBNHarS0mvhpca0Ct</calling-party>
 <called-party>AKSZIKCYJ0YU0oAEAKSZIKCYJ</called-party>
 </user-information>
 <scrid>20081203153307-xxxxxxx-12345-1234</scrid>
 <status-code>1</status-code>
 <status-description>Converted</status-description>
 </conversion-response>
</conversion>
</request>

-----=_NextPart_000_0001_01C7F52D.834C7D30--

4. When the message is returned to the Cisco Unity Connection server, it is deleted from SpinVox's databases within 48 hours.

Message Processing

A message arriving at SpinVox is flagged for automated transcription, and the following steps occur in the transcription:

1. Audio is processed and transcription is performed by the machine.
2. The transcription is written to the SpinVox database.
3. The message is returned to the Cisco Unity Connection server.
4. Data is deleted from the SpinVox database within 48 hours. The data is kept for 48 hours to allow time for troubleshooting, if necessary.

Agent Security information

Numerous security policies are in place to govern the transcription process at SpinVox:

- Audio stays on the central processing system.
- Facilities undergo a rigorous selection process and are subject to regular security audits.
- Nuance and SpinVox personnel:
 - Are screened, tested, and vetted
 - Undergo security, privacy, and compliance training
 - Sign nondisclosure and confidentiality agreements with SpinVox
- PC hardware at quality-control facilities:
 - Is hardened as per generally accepted industry best practices
 - Has only essential programs and services enabled
 - Has working files flushed after use
 - Has cut, copy, and paste functions disabled
 - Has controlled Internet access

The security measures in place for the Cisco SpeechView solution should satisfy any organization's concerns about sending voice messages outside of the company firewall in order to use a third-party external transcription service. Cisco has worked closely with SpinVox to help ensure that security is a priority throughout every step of the transcription process. As a result, SpeechView offers our Cisco Unity Connection customers an accurate, secure, and easy-to-use speech-to-text solution.

For more information about Cisco SpeechView, please visit <http://www.cisco.com/go/speechview>.

Cisco Capital

Financing to Help You Achieve Your Objectives

Cisco Capital can help you acquire the technology you need to achieve your objectives and stay competitive. We can help you reduce CapEx. Accelerate your growth. Optimize your investment dollars and ROI. Cisco Capital financing gives you flexibility in acquiring hardware, software, services, and complementary third-party equipment. And there's just one predictable payment. Cisco Capital is available in more than 100 countries. [Learn more](#).




Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

 Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: www.cisco.com/go/trademarks. Third party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)