

VOICEGENIE

VoiceGenie 7
RTSP Users' Guide

Sept 14th, 2005



VoiceGenie Technologies Inc.
1120 Finch Ave. W. • Toronto, Ontario • M3J 3H7 • Canada
T. +1.416.736.4151 • F. +1.416.736.1551 • support@voicegenie.com
www.voicegenie.com

VoiceGenie Contacts

VoiceGenie Technologies Inc.
1120 Finch Avenue West
Toronto, Ontario
Canada
M3J 3H7

T. +1.416.736.4151
F. +1.416.736.1551
support@voicegenie.com

<http://www.voicegenie.com/index.html>

Proprietary / Copyright Information

This material contains proprietary and/or copyright information of VoiceGenie Technologies Inc. and may not be copied, used, or disclosed without the permission of VoiceGenie Technologies Inc.

© COPYRIGHT 2005 VoiceGenie Technologies Inc.

Notice

Every effort was made to ensure that the information in this document was complete and accurate at the time of printing. However, this information is subject to change without notice.

Please note that VoiceGenie makes no warranties with respect to future releases. The information provided herein is for informational purposes only. VoiceGenie reserves the right to change product release schedules and/or features allocated to a product release at any time.

Trademarks

All trademarks are the property of their respective owners. Where those designations appear in this document, and VoiceGenie is aware of a trademark claim, the designations have been printed in initial caps or all caps.



Table of Contents

Proprietary / Copyright Information	3
Notice	3
Trademarks	3
Table of Contents	4
Revision History	5
1 Introduction	6
1.1 Functional Overview	6
1.2 VoiceGenie Components	6
1.2.1 <i>Media Platform (MP)</i>	6
1.2.2 <i>Speech Resource Management (SRM)</i>	6
2 Architecture	7
3 Installation	8
4 System Configuration Overview	8
5 Media Platform Provisioning	9
6 Application Development	10
6.1.1 <i>Sharable RTSP URI</i>	10
6.1.2 <i>Multiple RTSP URIs</i>	11

Revision History

Version	Date	Change Summary	Author/Editor
1.0	March 29, 2005	Initial release	Johnson Tse
1.1	April 19 th , 2005	Revised Version for VoiceGenie 7 Release	Andrew Ho
1.2	Sept 14 th , 2005	2 nd Revision	Rob Marchard

1 Introduction

The VoiceGenie Media Platform is capable of delivering streamed audio content to a caller as part of a VoiceXML application. The Media Platform makes use of the Real-Time Streaming Protocol (RTSP) connector to interface to third party RTSP format content servers. This content can be live or pre-recorded. The purpose of this document is to provide information on how to utilize the RTSP connector, and the feature itself.

1.1 Functional Overview

The Real Time Streaming Protocol (RTSP, described in <http://www.ietf.org/rfc/rfc2326.txt>) is a popular protocol for streaming real-time media from a content server to listeners. There are two primary use cases in which the VoiceGenie platform might be used to deliver streaming content:

- **Listening to a live audio broadcast** --- Examples of these are sporting events, presidential speeches, radio broadcasts, etc.
- **Listening to archived audio** --- Examples of these are on-line sampling of music, or listening to radio programs from previous weeks.

The RTSP connector allows a VoiceGenie installation to deliver these classes of applications to a caller. The RTSP connector is implemented as part of the VoiceGenie Speech Resource Management (SRM) infrastructure, and can therefore take advantage of the availability and resilience features provided by that infrastructure. It is also a highly optimized implementation – for example, one of the features of the VoiceGenie RTSP implementation is the sharing of the RTSP session data when one or more callers listen to the same streaming source on a VoiceGenie platform.

1.2 VoiceGenie Components

Before explaining the VoiceGenie architecture and describing examples of configurations, the following subsections describe the functionality of each VoiceGenie component. For a detailed VoiceGenie 7 system overview, please refer to the [VoiceGenie 7 Media Platform Users Guide](#).

1.2.1 Media Platform (MP)

The VoiceGenie Media Platform contains a VoiceXML interpreter that allows incoming/outgoing calls to interact with callers using VoiceXML dialogs. The Media Platform connects to Voice over IP (VoIP) networks or, using PSTN media processing boards, to TDM telephony networks.

1.2.2 Speech Resource Management (SRM)

The Speech Resource Management server is a component that keeps track of all ASR and Text-to-Speech (TTS) engines and distributes requests from the Media Platform to these engines.

The VoiceGenie SRM Server is the component that interfaces with a third-party RTSP server or proxy. It is responsible for accepting the audio data from the RTSP server and forwarding the stream to the Media Platform.

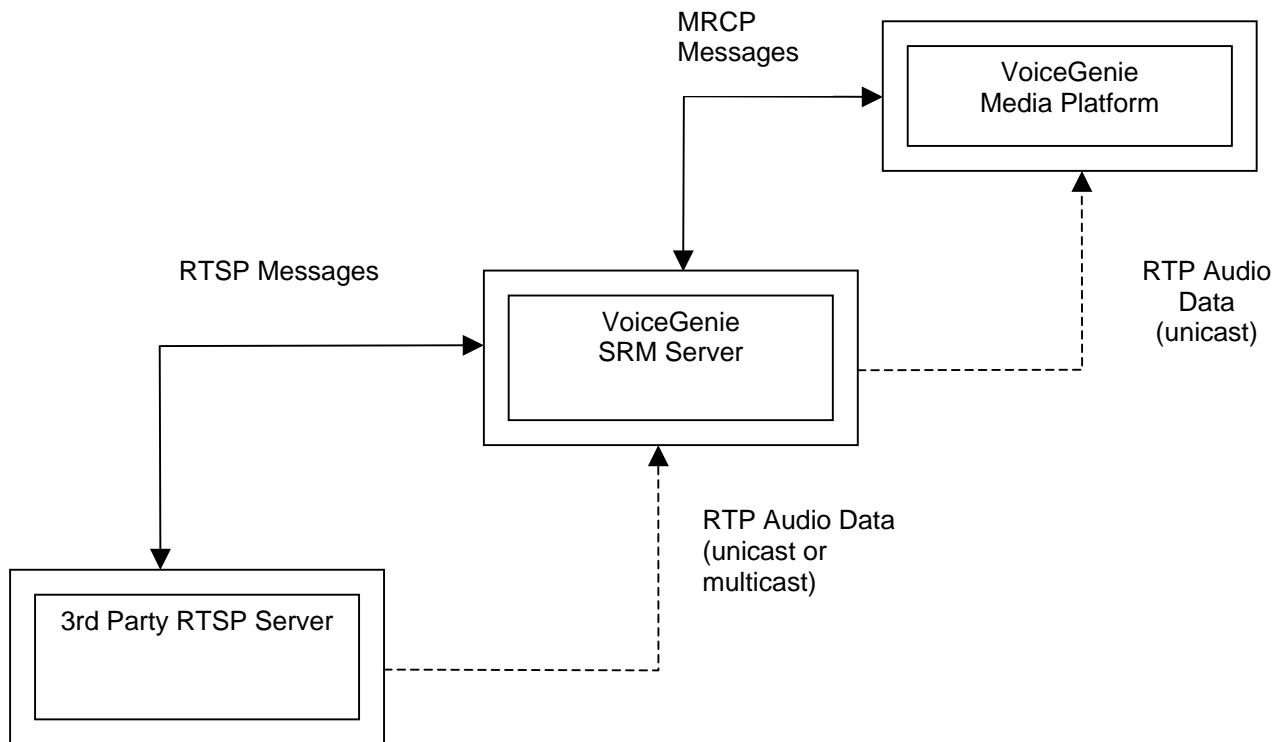
The following RTSP Methods are supported by the RTSP Engine module:

- RTSP DESCRIBE
- RTSP SETUP
- RTSP TEARDOWN
- RTSP PLAY

2 Architecture

The SRM Server is capable of receiving audio either in *unicast* or *multicast* mode. This is controlled by setting the `vrmserver.rtspts.usemulticast` property in the SRM Server configuration.

The SRM Server +-always forwards the audio data in unicast mode to the VoiceGenie Media Platform.



If the VoiceGenie Media Platform receives more than one call requesting the same RTSP resource that is shareable (can specified on the application), then the SRM Server will establish only one RTSP session to the third-party RTSP server. When the audio data is received by the SRM Server, it will forward the audio data for the different calls to the Media Platform. In this optimization, only one copy of the audio data is transferred between the third-party RTSP server and the SRM Server.

3 Installation

To use the RTSP server, the VoiceGenie RTSP package must be installed. Please see the [VoiceGenie 7 Platform Installation Guide](#) for details.

As a pre-requisite, the deployment must have the following installed:

- VoiceGenie 7 Software or later
- VoiceGenie SRM Server
- VoiceGenie RTSP package
- Any third-party RTSP software

While the RTSP package and the SRM Server must be installed on the same server, the other components may be installed on separate servers to have a multi-tiered architecture.

4 System Configuration Overview

The VoiceGenie RTSP component is implemented in a shared object that is loaded by the SRM Server. The SRM Server exposes the RTSP engine as a TTS resource.

Once the package is deployed successfully, the system will be configured automatically. From the Cluster Management Platform (CMP) System Management Console (SMC), browser to the Speech Resource Manager -> SRM Server submenu. The SRM Server configuration be pre-populated, similar to what is shown below:

<input checked="" type="checkbox"/>	vrmserver.rtspts.enginename	rtspsynthesizer
<input checked="" type="checkbox"/>	vrmserver.rtspts.virtual_dir	rtspsynthesizer
<input checked="" type="checkbox"/>	vrmserver.rtspts.numsession	200
<input checked="" type="checkbox"/>	vrmserver.rtspts.audiotimeout	7000
<input checked="" type="checkbox"/>	vrmserver.rtspts.usemulticast	false
<input checked="" type="checkbox"/>	vrmserver.rtspts.multicastaddr	225.219.201.15
<input checked="" type="checkbox"/>	vrmserver.rtspts.sharetcpsocket	true
<input checked="" type="checkbox"/>	vrmserver.rtspts.multicastttl	0

The following table describes each of the parameters – these can be changed using the VoiceGenie SMC interface:

Parameter Name	Default Value	Description
vrmserver.rtsptts.enginename	rtspsynthesizer	The engine name to be used in the CMP health monitor
vrmserver.rtsptts.virtual_dir	rtspsynthesizer	The MRCP URI to be exposed by the SRM server for incoming TTS requests
vrmserver.rtsptts.numsession	User input	Maximum number of concurrent calls accepted by the SRM server
vrmserver.rtsptts.audiotimeout	7000	Timeout value (in milliseconds) before MRCP SPEAK-COMplete is thrown, if no data is received within the specified timeout period.
vrmserver.rtsptts.usemulticast	false	Specify whether to support Multicasting audio for receiving data from the RTSP Streaming server to the VoiceGenie RTSP Server. Please note that the audio data is always sent in unicast mode between the VoiceGenie RTSP server and the VoiceGenie Media Platform
vrmserver.rtsptts.multicastaddr	225.219.201.15	Multicast address to be used (usemulticastaddr must be set to true for this parameter to take effect)
Vrmservice.rtsptts.multicastttl	0	Time To Live (TTL) value for multicast packets. Set to 0 if the servers are located in the same subnet.
vrmservice.rtsptts.sharetcpsocket	true	Reuse the same TCP socket for different TTS sessions

5 Media Platform Provisioning

After the RTSP package is installed, from the SMC menu, the RTSP portion of the MediaPlatform -> SpeechResources menu will contain the following entries:

Speech Resource Name:	<input type="text" value="RTSPTTS"/>	ID: 262 (42)
Hostname/IP:	<input type="text" value="10.0.0.70"/>	Port: <input type="text" value="22000"/>
Resource URI:	<input type="text" value="rtsp://10.0.0.70/rtspsynthesizer"/>	
Resource Type:	<input type="text" value="TTS"/>	
Parameter Name:	<input type="text"/>	Value: <input type="text"/> <input type="button" value="Add"/>
<input type="button" value="Update"/> <input type="button" value="Delete"/> <input type="button" value="Target as Primary"/> <input type="button" value="Target as Backup"/>		

At this point, the user should click on the "Target as Primary" button and select the server that has the Media Platform installed. Click on the "Update" button to confirm your choice.

As the RTSP Server cannot accept SSML SPEAK requests, the VoiceXML interpreter needs to include RTSPTTS in the non_ssml_engines list. This configuration is performed by the VoiceGenie installation package.

6 Application Development

The application developer needs to specify the TTSENGINE property with the value RTSPTTS on the application. If you have specified a different Speech Resource Name in the SRM configuration, use this name in place of RTSPTTS.

TTSENGINE property can be specified at any of the following locations, in the order of precedence:

- At the prompt level, where the property is specified as an attribute in the <prompt> tag in the application;
- At the page level, where the <property name="ttsengine" value="RTSPTTS"/> is specified inside the <vxml> tag
- At the system level, where the <property name="ttsengine" value="RTSPTTS"/> is specified in the defaults.vxml file

To expose an RTSP resource in the VoiceXML application, the application developer makes use of the engine as a normal TTS resource. However, instead of putting in the text for a user prompt, the RTSP URI for the audio source is specified instead. The following is an example of a VoiceXML application using RTSP:

```
<vxml version="2.0" xmlns="http://www.w3.org/2001/vxml">
<property name="ttsengine" value="RTSPTTS"/>
  <form>
    <block>
      <![CDATA[
        rtsp://10.0.0.106:8554/sounds/paper
      ]]>
    </block>
  </form>
</vxml>
```

In the above example, a third-party RTSP server is installed on the server 10.0.0.106. It exposes the audio resource using the URI `rtsp://10.0.0.106:8554/sounds/paper`.

6.1.1 Sharable RTSP URI

When used in live streaming, it is a useful optimization to share RTSP audio among callers. This feature is enabled by adding a "?shareable=true" suffix to the URI in the application. If this sharing is enabled, a new caller trying to listen to the RTSP stream will start listening to the RTSP stream at the time when the attempt is made, instead of starting from the beginning.

The following is an example of how to utilize this feature:

```
<vxml version="2.0" xmlns="http://www.w3.org/2001/vxml">
<property name="ttsengine" value="RTSPTTS"/>
  <form>
    <block>
      <![CDATA[
        rtsp://10.0.0.91:2004/sample_100kbit.mov?shareable=true
      ]]>
    </block>
  </form>
</vxml>
```

```
</form>  
</vxml>
```

6.1.2 Multiple RTSP URIs

Here is a sample VoiceXML application showing the use of multiple RTSP streams on the same page:

```
<?xml version="1.0"?>  
<vxml version="2.0" xmlns="http://www.w3.org/2001/vxml">  
<property name="ttsengine" value="RTSPTTS"/>  
<form>  
<block>  
<![CDATA[  
rtsp://10.0.0.171:8554/sting?shareable=true  
]]>  
</block>  
<audio src="empty.vox"/>  
<block>  
<![CDATA[  
rtsp://10.0.0.171:8554/welcome?shareable=false  
]]>  
</block>  
<audio src="empty.vox"/>  
<block>  
<![CDATA[  
rtsp://10.0.0.171:8554/longtts  
]]>  
</block>  
</form>  
</vxml>
```

Note the use of 'empty.vox' – this is an empty audio file that is used to break up the different URIs into separate TTS requests when the VoiceXML interpreter processes the application. This restriction will be removed in a future release.

When combined with other features of VoiceXML, complex applications incorporating RTSP can be built. Please refer to the online documentation for VoiceXML on the [VoiceGenie Developer Workshop](#) for a complete list of features supported by the VoiceGenie 7 software.