Application Guide

Q-SYS Softphone SIP

A primer on SIP telephony and the Q-SYS softphone (Designer 5.2)

Application Engineering Team

Rev. A, November 2016

1. Introduction to SIP

The Q-SYS softphone is a great way to integrate a Q-SYS Core processor to a voice over IP (VoIP) telephone system for conferencing and telephone paging.

The terminology around VoIP can seem quite mysterious, so it can often be hard to know if the Q-SYS Softphone is compatible with an end user's VoIP system. In this application guide we will try to clarify and simplify this topic.

Defining the Q-SYS Softphone

The Q-SYS Softphone is just what the name implies … a phone that is implemented in software. It connects to a VoIP network and can optionally register with a server that is responsible for that system's call negotiation. Once registered (if required), the softphone can make and receive calls just like any other VoIP phone.

In most VoIP systems, the server and phones come as a package. The "language" spoken by the phone system, i.e., the protocol, can vary widely from one system to another. The Q-SYS softphone, however, does not make use of these many proprietary languages. It uses the Session Initiation Protocol (SIP), a VoIP interoperability standard for controlling and establishing calls. Once the call is established it uses the Real Time Protocol (RTP) to transmit and receive voice data. Even most proprietary VoIP systems understand SIP or can with a system upgrade.

Basic SIP telephony has two major device types:

- User Agent Client (UAC). This device generates requests and directs them to servers.
- User Agent Server (UAS). This device receives and processes requests and sends responses to other devices on the network. A UAS device can serve in these roles:
  - Proxy server. A proxy server receives SIP commands and forwards them to the appropriate devices on the network. Thus, it is responsible for handling calls on the network.
  - Registrar. A UAC may not always stay in the same location, so the registrar detects their locations on the network. Those locations are stored to a third server role, the location server.
  - Location Server—The location server keeps a record of all UAC locations, as detected by the registrar.
  - Redirect Server—If the intended recipient of a SIP message has moved or is otherwise unreachable, the redirect server will tell the originating UAC that it must try another route.

Note that these server roles do not have to be carried out all by different physical devices. In fact, the same device can host all of these functions and serve as a UAC as well.

Most VoIP phones and softphones must act as both UAS and UAC so they can provide all the functionality required of them. A device that combines these duties is referred to as a user agent (UA) or VoIP endpoint.

To make calls outside a local VoIP network, it is often necessary to route them to the Internet or to other local VoIP networks. SIP trunks are often created to accomplish this. One can think of SIP trunks as virtual wires bridging the expanse between two domains. SIP is used to control the sessions between them.

The Q-SYS softphone is a simple SIP endpoint or UA device. It is used to make and receive calls through interaction with the various UAS roles. The Q-SYS softphone does *not* serve as a SIP trunk.

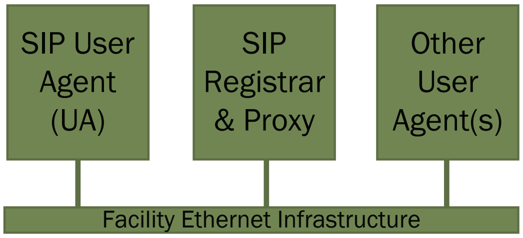


Figure 1. Simple VoIP network with only user agents and proxy

Basics of SIP Communication

The primary functions of SIP are the setup, teardown, and modification of calls. SIP is based on the HTTP model and makes use of simple requests and replies. It can use either UDP or TCP transport, but all SIP devices on a network must be configured to use the same transport the proxy server uses. The Q-SYS softphone in Q-SYS Designer 5.2.39 and higher support both TCP and UDP SIP communication; earlier versions use only UDP.

SIP typically uses port 5060 for unsecure sessions and 5061 for secure sessions using Transport Layer Security (TLS). In some VoIP systems this can be changed if necessary.

SIP Authentication

The vast majority of VoIP systems require some form of authentication from a device on the network before it can make or receive calls. See Figure 2. A SIP UA (i.e., an endpoint) sends a **Register** message with no username or password to the UAS that is serving as registrar. If the system requires authentication, the registrar will reply with a challenge or **Unauthorized** message. The challenge message will include a nonce, which is an MD5 hash value the endpoint will use to encrypt its credentials when it replies. If the username and password are correct, the registrar will respond with a **200 OK** message and the endpoint's location will be updated and stored in the location server. Only then will the endpoint be able to send or receive calls.

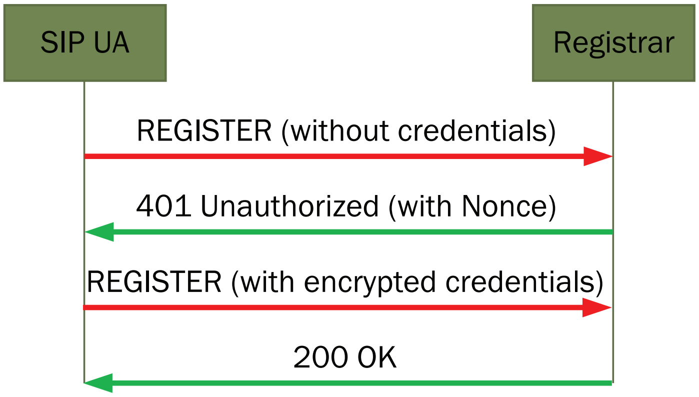

SIP UA
Registrar
REGISTER (without credentials)
401 Unauthorized (with Nonce)
REGISTER (with encrypted credentials)
200 OK

Figure 2. Message flow in simple digest authentication.

SIP Call Flow

Call Setup

With the endpoint registered, calls can then be attempted to or from it. Each call begins with an **INVITE** message to or from the proxy server. The INVITE message includes information about where the call is to be directed and often also includes information about the streaming audio formats supported by the caller.

The different audio encoding and decoding formats are called codecs. The range of codecs used by a particular VoIP system will often be based on available network bandwidth and licensing cost. In general, transmitting higher quality audio streams requires more network bandwidth.

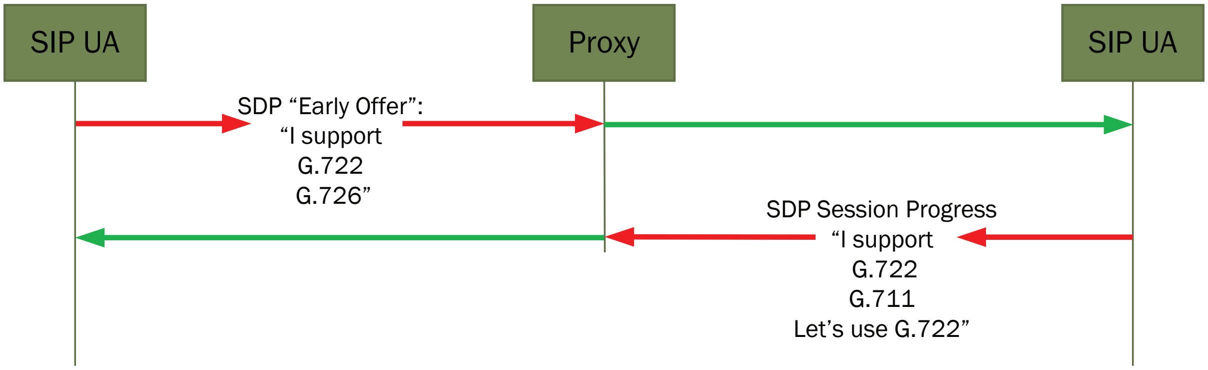


Figure 3. Simplified codec negotiation.

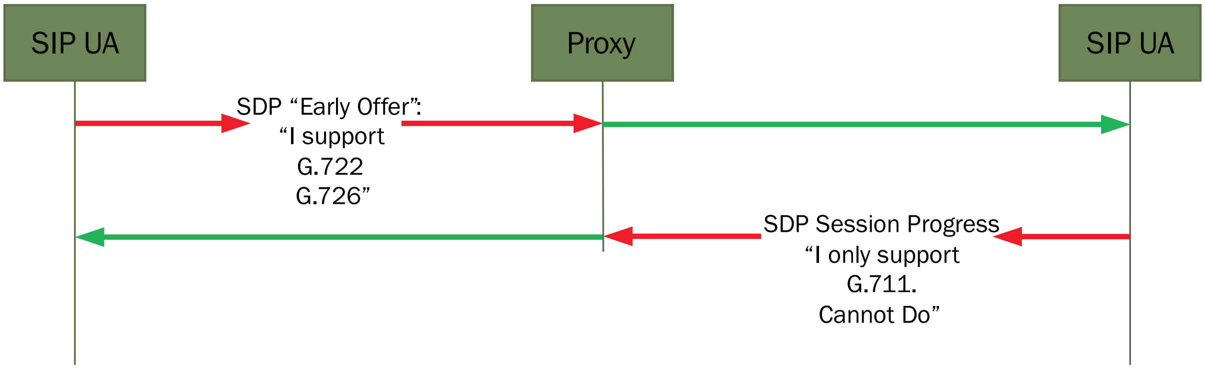
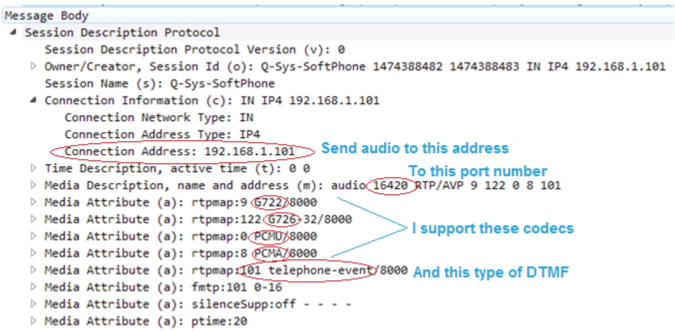


Figure 4. Simplified unsuccessful codec negotiation; there is no codec in common

The session description protocol (SDP) advertises what codecs a device supports and typically lists them in preferred order from highest to lowest quality (Figure 4). Including the SDP information in the **INVITE** message is called "early offer." The Q-SYS softphone requires "early offer" to successfully receive a phone call.

The SDP information details the supported DTMF type and also includes contact and port information that tells the far end how to route RTP audio, once a codec is chosen.



```
Message Body
◢ Session Description Protocol
    Session Description Protocol Version (v): 0
  ▷ Owner/Creator, Session Id (o): Q-Sys-SoftPhone 1474388482 1474388483 IN IP4 192.168.1.101
    Session Name (s): Q-Sys-SoftPhone
  ◢ Connection Information (c): IN IP4 192.168.1.101
      Connection Network Type: IN
      Connection Address Type: IP4
      Connection Address: 192.168.1.101
  ▷ Time Description, active time (t): 0 0
  ▷ Media Description, name and address (m): audio 16420 RTP/AVP 9 122 0 8 101
  ▷ Media Attribute (a): rtpmap:9 G722/8000
  ▷ Media Attribute (a): rtpmap:122 G726-32/8000
  ▷ Media Attribute (a): rtpmap:0 PCMU/8000
  ▷ Media Attribute (a): rtpmap:8 PCMA/8000
  ▷ Media Attribute (a): rtpmap:101 telephone-event/8000
  ▷ Media Attribute (a): fmtp:101 0-16
  ▷ Media Attribute (a): silenceSupp:off - - - -
  ▷ Media Attribute (a): ptime:20
```

Send audio to this address

To this port number

I support these codecs

And this type of DTMF

Figure 5. Content of INVITE SIP message with Session Description

When the **INVITE** message is received, the proxy server will have the location server determine how to route the call to the appropriate party. If it is an internal call, it might be to another phone or softphone on the same network. It may be to a peer on the other side of a SIP trunk (on another VoIP network). It may be to a phone connected to the public switched telephone network (PTSN), in which case the VoIP call is established to a media gateway that will provide the appropriate signaling to the analog telephony network.

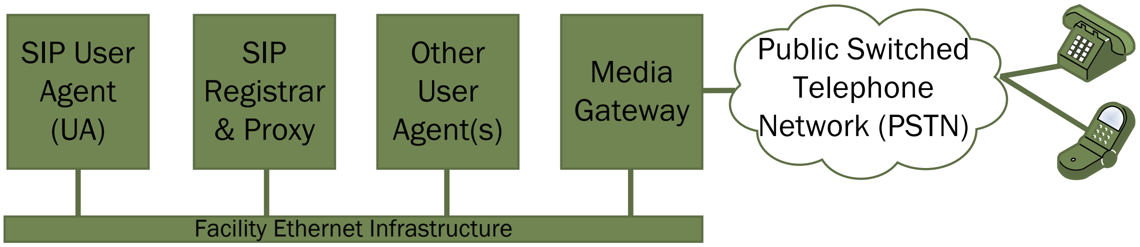

SIP User Agent (UA)
SIP Registrar & Proxy
Other User Agent(s)
Media Gateway
Public Switched Telephone Network (PSTN)
Facility Ethernet Infrastructure

Figure 6. Simple VoIP Network with Media Gateway to PSTN

Once the proxy knows where to direct the **INVITE** message, it will forward the message to the appropriate device in the appropriate manner. The softphone or phone being called will respond to the proxy and then the proxy will in turn respond to the caller. When the details of the upcoming call are established, the proxy's response to the caller will be a **RINGING** message.

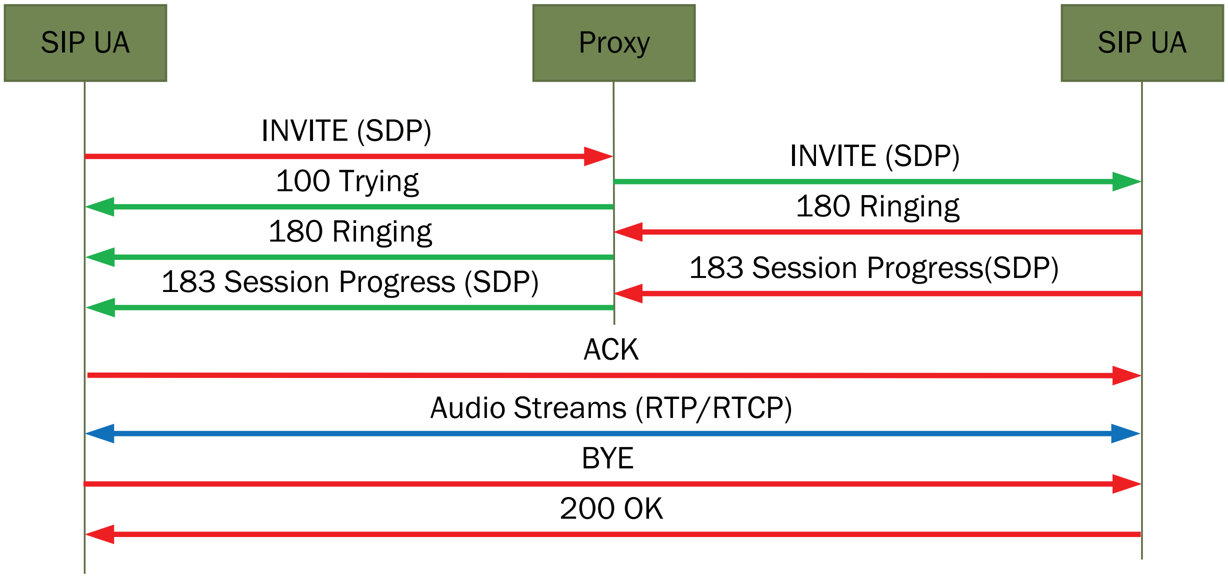


Figure 7. SIP, RTP/RTCP message flow for simple SIP call

Once the call is answered at the far end, the session initiation protocol has done its job and the peers now set up the call, with the two parties now directly exchanging the audio streams necessary for communication. The proxy server may establish a session timer, a SIP-based method of periodically checking the status of the call. If not, the SIP is finished until it's time to tear down the call.

The Call Itself

To send and receive streaming audio, the call uses another set of protocols: Real Time Protocol (RTP) and Real Time Control Protocol (RTCP). They are used to send and supervise the audio streams to each device participating in the call. As noted before, each device lists the audio codecs it supports in a preferred order, and the call will use the highest priority format common to both devices. If they have no codecs in common, the call cannot take place.

DTMF Considerations

Sometimes the caller must negotiate voice prompts or conference call codes using the keypad. In VoIP telephony the calling device typically does not play these tones and send them via RTP to the far end. Instead, control commands that identify the keys pressed are sent to the far end. This is called "out of band" DTMF. The Q-SYS softphone supports two common formats for out of band DTMF: RFC2833 and SIP INFO. RFC2833 uses **Event** messages in the RTP stream to convey DTMF signals, while SIP INFO sends INFO messages in the SIP flow. A VoIP device sending actual audio tones in the RTP stream is called "in-band" DTMF (to be supported in a future Q-SYS softphone release). If the call goes to the PTSN through a media gateway, the gateway is responsible for converting the tones into analog audio.

Tearing Down the Call

When the call is finished, of course, someone hangs up. The device that does so sends a **BYE** message to signal the end of the call. Each peer responds appropriately and the RTP streams stop.

Secure SIP (TLS)

Some VoIP systems require Transport Layer Security (TLS). This form of secure SIP takes extra care to confirm endpoints are authorized and in turn encrypts the SIP and RTP traffic on the VoIP network. This makes it virtually impossible to intercept and decode with network tools used by would-be hackers.

Q-SYS softphone support for TLS is planned in a future release of Q-SYS Designer.

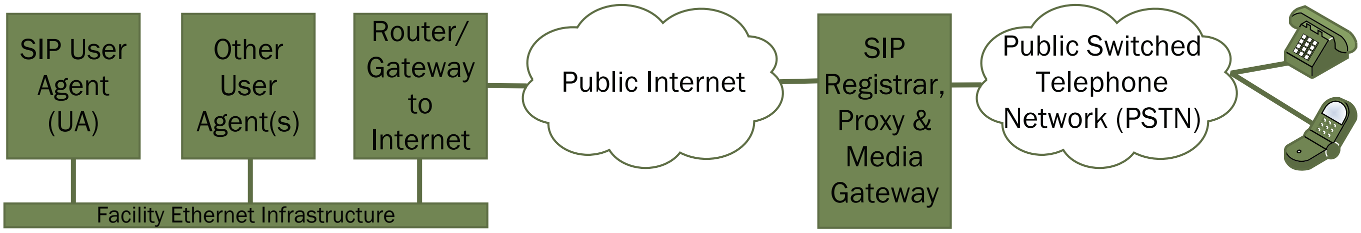

SIP User Agent (UA)
Other User Agent(s)
Router/ Gateway to Internet
Facility Ethernet Infrastructure
Public Internet
SIP Registrar, Proxy & Media Gateway
Public Switched Telephone Network (PSTN)

Figure 8. Simple diagram of a hosted SIP proxy solution

Hosted SIP Solutions

The current trend is having the SIP registrar and proxy lie on the Internet side of the typical router or firewall that sits between a local area network (LAN) and the Internet itself. This is often called a "hosted" or "cloud-based" SIP solution. While the basic SIP exchange for registration and call flow remain the same, the introduction of the router or firewall can complicate matters.

Firewalls

A firewall's job is to protect a network from outside attacks and unauthorized access to internal resources. It typically works by allowing traffic to flow only on UDP/TCP ports that are known to be used for necessary services and blocking all others. Although the port used for SIP is static (typically 5060), the ports used for the audio portion of the call (RTP/RTCP) are dynamic. This requires that the firewall opens a wide range of ports, which can compromise the security of the network.

Network Address Translation (NAT)

Routers typically employ Network Address Translation (NAT) when forwarding network packets from the LAN to outside networks. When doing so, the router replaces the IP address of a computer on the LAN with its own public IP address so hosts on the Internet know where to reply. The router is responsible for recording what packets were forwarded and then making sure any replies make it back to the originator on the LAN. If the NAT router receives an unsolicited packet, however, it may not know what local device should receive it. For example, a call into the LAN from a cloud-based proxy may not reach the intended recipient. Even if the call is negotiated properly from the SIP perspective, the router still may not know how to forward the resulting RTP/RTCP traffic, resulting in a call with no audio.

In addition to NAT, some routers employ port address translation (PAT), meaning the UDP or TCP port number may change as a message moves from the public to the private side of the router. This, of course, can compound the problem. The Q-SYS softphone offers STUN to help get around these issues.

STUN

Session Transversal Utilities for NAT (STUN) is a simple mechanism to help a peer on a private network discover its public IP address. In this case the Q-SYS softphone would periodically contact a STUN server on the Internet (public side of the NAT router) on a known UDP port. The STUN server sends a response to the softphone that includes the IP address presented to the STUN server (the softphone's public address) as well as the originating port number (the public port number in the case of PAT). The Q-SYS softphone can then write the public IP address into the SDP contact information. The originating port number is used for the RTP and RTCP, which has already been "punched" through the NAT router or firewall.

VLAN Tagging Concerns

Many VoIP or SIP endpoints support attachment to the general data network for some functions (such as DHCP addressing, passing data to a secondary Ethernet port, etc.) while they apply a "Tag" to the SIP and RTP traffic. This allows the tagged traffic to be prioritized and handled in different ways by the IT infrastructure. Note that the Q-SYS softphone does not offer this option to tag its transmissions. If a tag is required, the switch port to which the Q-SYS softphone is connected must be configured to add the tag upon ingress. Egress traffic should present to the Q-SYS softphone untagged.

Summary

This document is intended to explain the basic operating principles of SIP telephony and ultimately tie this understanding to the Q-SYS softphone. These concepts also provide a great tool for determining if the Q-SYS softphone is compatible with a given VoIP system:

- ✔ The Q-SYS softphone is a SIP endpoint, meaning that it does not support other VoIP protocols such as Skinny (SCCP) or H.323.
- ✔ The Q-SYS softphone does not serve as or connect directly to a SIP trunk.
- ✔ The Q-SYS softphone supports both TCP and UDP SIP communication; the SIP port number is configurable.
- ✔ The Q-SYS softphone requires Session Description Protocol (SDP) information in the INVITE packet from the proxy. This is known as “Early Offer”
- ✔ The Q-SYS softphone supports both RFC2833 and SIP INFO out-of-band DTMF signaling.
- ✔ The Q-SYS softphone supports Session Transversal Utilities for NAT (STUN) to accommodate connection to SIP proxies on the public Internet. It may not always be required, however.
- ✔ The Q-SYS softphone does not apply a VLAN tag to SIP/RTP/RTCP traffic.
- ✔ The Q-SYS softphone does not support secure SIP (TLS/SRTP).

VoIP Compatibility Checklist

The Q-SYS softphone is compatible with a number of VoIP providers, including Cisco, Avaya, and a number of hosted solutions (such as VoIPo, etc.). To get an idea if the Q-SYS softphone will properly register with the VoIP system in a given facility, the following questions must be asked of the VoIP administrator:

- ✔ Does this system support SIP services and allow for the registration of SIP endpoints?
- ✔ Can the system be configured to provide “Early Offer” on incoming calls to the SIP endpoint?
- ✔ Does the system support RFC2833 or SIP INFO methods for out-of-band DTMF?
- ✔ Please confirm this system DOES NOT require TLS/SRTP.
- ✔ Is this a private or hosted SIP solution?

If all the above are true, use the form on the following pages to gather the information needed to properly register and make calls with the Q-SYS softphone.

Q-SYS Softphone SIP Integration Worksheet

Please fill out this form with the requested information. This will allow us to preconfigure the Q-SYS core and greatly simplify the process of integrating a Q-SYS softphone.

IP Address

Will this SIP endpoint be given a static IP address or will it be DHCP?

If static, please provide:

IP Address: ____

Net mask: ____

Either DHCP or static:

DNS Server, if needed: ____

Proxy information

Address or hostname of SIP Proxy: ___________________________________ (if host, specify DNS above)

SIP Transport (select one): □ TCP □ UDP

SIP port to be used (5060 default): ________________________________

SIP Domain: ____

Subscriber information

Subscriber Number: __________________________________

Digest Username: ___________________________________

Digest Password: ___________________________________

Required Voice Codec Support (check all that apply)

□ G.722

□ G.726 40K

□ G.726 32K

□ G.726 24K

□ G.726 16K

□ G.711 ulaw

□ G.711 alaw

□ GSM

□ G.726 40k(AAL2)

□ G.726 32k(AAL2)

□ G.726 24k(AAL2)

□ G.726 16k(AAL2)

□ BroadVoice16 (BV16)

□ BroadVoice32 (BV32)

□ Speex

□ Raw PCM Signed Linear (16-bit)

DTMF support (check all that apply)

□ RFC2833 (RTP Event)

□ SIP Info

□ In-band DTMF only

QSC

qsc.com

AET Application Engineering Team

(800) 854-4079 or (714) 957-7100
Outside the U.S. +1 (714) 754-6175
Fax: +1 (714) 754-6174

QSC, LLC
1675 MacArthur Boulevard
Costa Mesa, CA 92626 USA