Session Initiated Protocol (SIP): A Five-Function Protocol

SIP is an application-layer control protocol that can establish, modify, and terminate multimedia sessions (conferences) such as Internet telephony calls. SIP can also invite participants to already existing sessions, such as multicast conferences. Media can be added to (and removed from) an existing session. SIP transparently supports name mapping and redirection services, which supports personal mobility-users can maintain a single externally visible identifier regardless of their network location.

White Paper
by Paul Stalvig

Introduction

> " SIP is an application-layer control protocol that can establish, modify, and terminate multimedia sessions (conferences) such as Internet telephony calls. SIP can also invite participants to already existing sessions, such as multicast conferences. Media can be added to (and removed from) an existing session. SIP transparently supports name mapping and redirection services, which supports personal mobility-users can maintain a single externally visible identifier regardless of their network location.

—Rosenberg, J., et al. (2002, June). RFC 3261 - SIP: Session Initiation Protocol. Internet RFC/STD/FYI/BCP Archives: http://www.faqs.org/rfcs/rfc3261.html

"SIP, SIP, SIP" that's all we hear lately. Why? Simply put, the Session Initiated Protocol (SIP) is the means to an end, bringing about the convergence of communications. SIP is the conduit that will bring together the enterprise and telecommunications to provide IP-based communications. Enterprises have worked for many years to provide triple-play — voice, video, and data — services. The enterprise has grown from data services and is working towards adding voice and video services over the network. Meanwhile, the telecommunications industry has grown from providing voice, then video, and now data services. As telecommunications migrates toward being able to provide comprehensive IP services, a standard method of providing sessions is needed.

Many protocols came before SIP, such as H.323. While H.323 provided an all-in-one solution, it was not flexible enough to provide for much beyond voice services. SIP also brings Web 2.0 concepts into fruition, but in a completely different manner.

SIP is a primitive in that it does not provide services; it provides a conduit for services. For example, consider SIP a regular telephone line. The telephone line doesn't provide a service; it provides the ability to connect services. For example, we can talk – choose a language, use a modem for data, or even pump the modem full of multimedia services. As long as the end-points (users) agree on what is being sent and how to view, listen, and so on, SIP can provide the session.

SIP is a component of a complete multimedia architecture, and relies on other Internet Engineering Task Force (IETF) protocols. The SIP typically will use the Real-Time Session Protocol (RTSP) to provide transportation and quality of service (QoS) feedback for streaming media. Other standardized protocols control access to the public-switched telephone network, as well as describing multimedia sessions. While SIP uses these protocols, it is not tied to them if a better solution arises. This is one of the great things about SIP over its predecessor protocols (H.323); it does not have to be redefined to move onto something better.

The key to SIP is that it provides only five functions: user location, user availability, user capabilities, session setup, and session management. That is all SIP does. That being said, SIP is flexible and open enough to allow developers to build their own "hooks" into SIP. This flexibility has given SIP an advantage over other "telecommunications protocols," and is why many enterprises are eager to develop, implement, and use SIP.

SIP in a Nutshell

SIP is responsible for user location, availability, capabilities, and session setup and management. SIP does not determine what services are being sent back and forth, nor does it affect how information is passed. Whether it uses radio waves, wired-networks, satellites, etc., SIP only needs to be able to communicate. SIP is a text-based protocol, but also carries a lot of non-text information. Also, SIP uses request/response transactions. The "session," however, is not a maintained connection as many enterprises have come to think of sessions.

SIP works between the session and application layers of the OSI model and is not confined to any one IP version. This means that it can work in and between IPv4 and IPv6 models. With the desire to keep SIP as flexible as possible, most of SIP's message and header field syntax is derived from the HTTP/1.1 specification, but is not tied to the HTTP/1.1 protocol.

SIP Functionality

Exactly what does SIP do? SIP provides the set-up/establishing, tying together, and tear-down/ terminating of multimedia communications. SIP does this by providing five different functions:

User Location

SIP determines user locations by a registration process. When a soft-phone is activated on a laptop, it sends out a registration to the SIP server announcing availability to the communications network. Voice over-IP (VoIP) phones, cellular phones, or even complete teleconferencing systems can be registered as well. Depending on the registration point chosen, there may be several different locations registered simultaneously.

User Availability

User availability is simply a method of determining whether or not a user would be willing to answer a request to communicate. If you "call" and no one answers, SIP determines that a user is not available. A user can have several locations registered, but might only accept incoming communications on one device. If that is not answered, it transfers to another device, or transfers the call to another application, such as voicemail.

User Capabilities

With all the various different methods and standards of multimedia communications, something is needed to check for compatibility between the communications and the users' capabilities. For example, if a user has an IP phone on their desk, a white-board conference via that device would not work. This function also determines which encryption/decryption methods a user can support.

Session Setup

SIP establishes the session parameter for both ends of the communications — more specifically, where one person calls and the other answers. SIP provides the means to setup and/or establish communications.

Session Management

This function provides the greatest amount of user awe. Provided a device is capable, a user could transfer from one device to another — such as from an IP-based phone to a laptop — without causing a noticeable impact. A user's overall capabilities would change — such as being able to start new applications such as white-board sharing — perhaps affecting the voice quality temporarily as SIP re-evaluates and modifies the communications streams to return the voice quality. With SIP session management, a user can also change a session by making it a conference call, changing a telephone call to a video conference, or opening an in-house developed application. And finally, SIP terminates the communications.

Although SIP has five functions, it is currently easier to think of SIP as the setup, management, and tear down of IP-based communications. The user location and capabilities functions could easily become absorbed into the session setup function. SIP maintains the distinct five functions to remain open and provide for an unknown future.

Summary

The intent of this paper is to offer the reader a basic understanding of the Session Initiated Protocol. SIP provides five functions, uses a layered approach that gives it flexibility, and has an openness that enables new applications to utilize it. SIP's open nature guarantees its large presence in IP-based multimedia communications architectures.