DOCUMENTO TÉCNICO

SIP — Introducción

Abril 2022

Índice

1 **Resumen** 3

2 **Introducción** 3

3 **¿Cómo funciona?** 3

3.1 Configuración peer-to-peer: la opción sencilla 3

3.2 Uso de un servidor SIP (PBX): más posibilidades 4

3.3 Uso de troncales SIP: asignación de un número de teléfono 5

4 **Dentro de una llamada SIP "normal"** 6

4.1 SDP: negociación del formato empleado 7

4.2 Llamadas en infraestructuras SIP complejas 7

5 **MFDT: envío de comandos en llamadas SIP** 8

6 **Entornos complejos y seguridad reforzada** 8

6.1 NAT traversal: encontrar el camino en redes complejas 8

6.2 Uso de cifrado con SIP 9

7 **Terminología de SIP** 10

1 Resumen

El protocolo de inicio de sesión (SIP) ofrece una interfaz adicional para la integración de sistemas de productos de seguridad. El protocolo SIP está ampliamente aceptado en el sector de las telecomunicaciones y aporta más flexibilidad para la interconexión y los usos cotidianos. Integradores de sistemas, desarrolladores y usuarios finales reclaman interfaces abiertas y estandarizadas, ya que les aportan un mayor valor al permitir el uso de los productos en diferentes sistemas. Los productos Axis compatibles con SIP están pensados tanto para usos de seguridad como para la comunicación.

La configuración de un sistema SIP puede resultar extremadamente sencilla. Sin embargo, en el caso de estructuras de red complejas o en situaciones que exijan un plus de seguridad o funciones adicionales de gestión de llamadas, es necesario utilizar servidores SIP y técnicas de NAT traversal, que requieren unos conocimientos técnicos más avanzados por parte del instalador o el técnico.

2 Introducción

El protocolo de inicio de sesión (Session Initiation Protocol, SIP) se utiliza para iniciar, mantener y finalizar sesiones multimedia entre distintas partes. Estas sesiones normalmente son de audio, pero a veces incluyen también vídeo. SIP es el protocolo estándar utilizado en aplicaciones de voz sobre IP (VoIP) y plataformas de comunicaciones unificadas.

Por ejemplo, los productos AXIS C3003-E Network Horn Speaker y AXIS I8016-LVE Network Video Intercom son compatibles con SIP, lo que abre la puerta a una nueva forma de conectar, integrar y controlar sus dispositivos de red Axis.

3 ¿Cómo funciona?

Para la comunicación a través de SIP, se necesitan al menos dos clientes SIP. Un cliente SIP puede ser un teléfono físico SIP, un softphone, un cliente móvil o un producto Axis compatible con Axis.

Cada cliente SIP tiene asignada su propia dirección SIP. Una dirección SIP es similar a una dirección de correo, pero con el prefijo “sip:”.

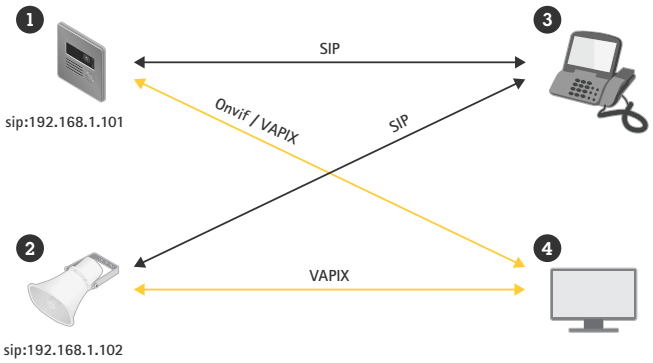
Por ejemplo, sip:bob@axis.com [sip:<usuario@]<proveedor>. Este identificador puede utilizarse en diferentes dispositivos y es similar a un número de teléfono vinculado a una tarjeta SIM, que puede usarse en varios dispositivos.

3.1 Configuración peer-to-peer: la opción sencilla

Un sistema SIP puede tener muchas formas. En su versión más sencilla, el sistema consta de dos agentes de usuario de SIP o más que se comunican directamente entre sí. Este modelo suele conocerse como configuración peer-to-peer, configuración de llamada directa o configuración local. En este caso, la dirección SIP suele tener el formato sip:<ip-local>, por ejemplo sip:192.168.0.90

Ejemplo: En una configuración sencilla, estos productos Axis (1, 2) pueden usar SIP para configurar la comunicación por audio y/o vídeo con otros dispositivos SIP (3) de la misma red sin necesidad de un servidor o PBX.

Al mismo tiempo, pueden estar conectados al sistema de gestión de vídeo (4) usando las API abiertas VAPIX u ONVIF Profile S como cualquier otro dispositivo Axis.



Para realizar una llamada peer-to-peer de un agente de usuario a otro de la red local, solo hace falta la dirección SIP que contiene la dirección IP de la unidad. Sin embargo, no todos los clientes SIP permiten las llamadas peer-to-peer.

3.2 Uso de un servidor SIP (PBX): más posibilidades

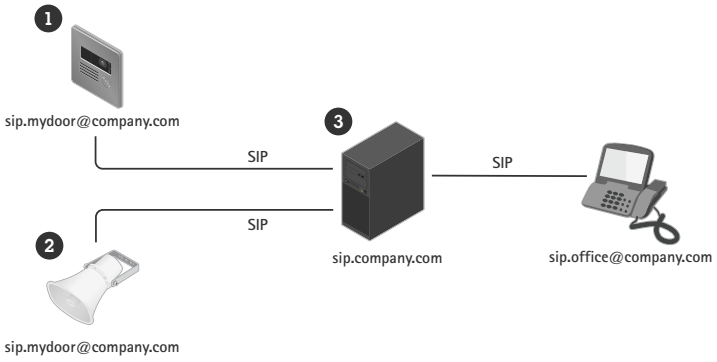
Una infraestructura VoIP basada en SIP es fácilmente escalable. Cuando llega el momento de crecer, podemos usar un servidor de registro, una centralita privada (PBX) o un servidor SIP como concentrador central. Los agentes de usuario de SIP se registran en el servidor de registro y luego contactan con otros agentes de usuario marcando una extensión en la centralita privada.

En este caso, una dirección SIP típica tendría este formato:<usuario>@<dominio>. Otra opción sería sip:<usuario>@<ip-servidor de registro>, por ejemplo sip:6007@ miservidorsip.net. Una centralita privada funciona como una central telefónica tradicional, que muestra el estado actual de los clientes y ofrece servicios como la transferencia de llamadas, buzón de voz y redireccionamientos, entre otros.

Un servidor SIP suele incluir un proxy, un servidor de registro y función de redireccionamiento. Los proxies dirigen las llamadas y aportan información lógica adicional a las llamadas entrantes. Los servidores de registro aceptan solicitudes de registro y actúan como servicio de localización del dominio que gestionan. Los servidores de redireccionamiento redirigen el cliente para contactar con una dirección SIP alternativa.

El servidor SIP puede configurarse como elemento local o puede estar deslocalizado. Puede estar alojado en una intranet o en un proveedor de servicios externo. Al realizar llamadas SIP entre diferentes sitios, normalmente las llamadas se enrutan a través de diferentes proxies SIP. Estos proxies consultan la ubicación de la dirección SIP de destino.

Ejemplo: Los productos Axis (1, 2) pueden conectarse a un servidor SIP (3) local o alojado en un proveedor de servicios externo. El servidor gestiona el establecimiento y la finalización de las llamadas entre dispositivos SIP en la red local o a través de Internet. Con esta configuración, la dirección SIP del dispositivo es independiente de su dirección IP y el servidor SIP permite el acceso al dispositivo siempre y cuando esté registrado en el servidor.

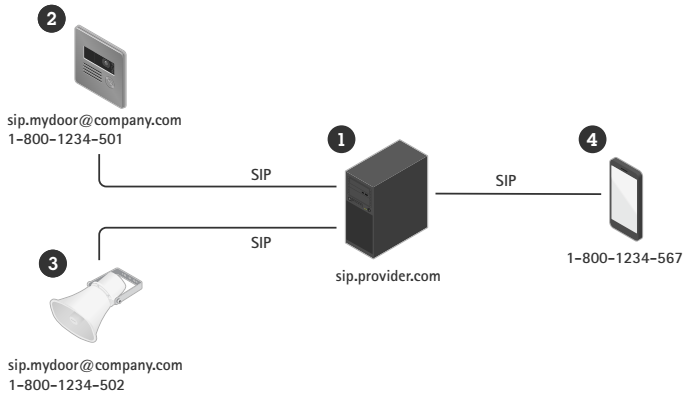


Para poder usar su dispositivo con un servidor SIP. debe crear una cuenta en el servidor con un ID de usuario y una contraseña determinados. Para registrar su dispositivo en el servidor, debe configurar una cuenta en el dispositivo introduciendo la dirección del servidor, el ID de usuario y la contraseña.

3.3 Uso de troncales SIP: asignación de un número de teléfono

Usando un troncal SIP, los agentes del usuario de SIP pueden incluso vincularse a una red de telefonía tradicional (PSTN). Eso abre la puerta a asignar un número de teléfono tradicional al agente de usuario de SIP.

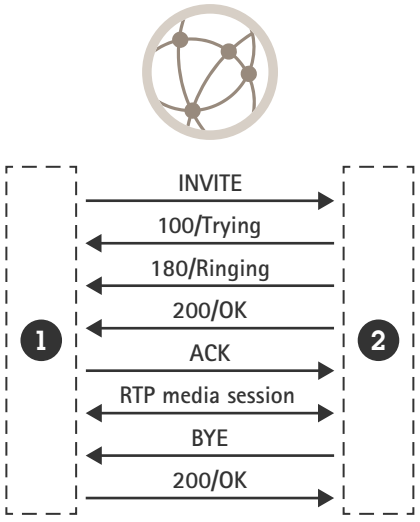
Ejemplo: Usando un troncal SIP (1) con un proveedor de servicio, puede asignar números de teléfono externos a sus dispositivos (2, 3). Esto permite realizar llamadas entre un altavoz de red o un videoportero y teléfonos tradicionales (4).



Al utilizarlo con un troncal SIP, el dispositivo se conecta al servidor del modo descrito arriba. El proveedor del servicio suele cobrar una tarifa adicional por los números externos.

4 Dentro de una llamada SIP “normal”

Para hacer una llamada SIP, se realizan varios pasos para intercambiar información entre los UA que inician y reciben la llamada.



Al iniciar una llamada, el UA iniciador (1) envía una solicitud o INVITACIÓN a la dirección SIP del UA destinatario (2). La INVITACIÓN contiene texto del Protocolo de descripción de sesión (SDP), en el que se describen los formatos multimedia disponibles y la información de contacto del iniciador de la llamada.

Tras recibir la INVITACIÓN, el destinatario confirma su recepción inmediatamente con una respuesta “100 TRYING”.

A continuación, el agente de usuario compara los formatos de contenidos descritos en el SDP con los suyos. Si se puede determinar un formato común, el agente de usuario indica al destinatario que hay una llamada entrante y envía una respuesta provisional, “180 RINGING”, al agente de usuario iniciador.

Cuando el destinatario responde a la llamada, se envía una respuesta “200 OK” al iniciador, para confirmar que se ha establecido la conexión. Esta respuesta contiene un SDP negociado en el que se indica al iniciador qué formato de contenido se debe usar y a dónde se deben enviar las transmisiones multimedia.

En ese momento, se configuran las transmisiones multimedia negociadas mediante el Protocolo de transporte en tiempo real (RTP), con parámetros basados en el SDP negociado, y los contenidos se transmiten directamente entre las dos partes. El iniciador envía una confirmación (ACK) a través de SIP para confirmar que ha configurado las transmisiones de contenido como se acordó. La sesión de SIP sigue activa, pero no participa en la transferencia de contenidos.

Cuando una de las partes decide finalizar la llamada, envía una nueva petición, en este caso BYE. Al recibir una petición BYE, la parte receptora responde con un “200 OK” y se detienen las transmisiones de contenidos RTP.

4.1 SDP: negociación del formato empleado

El protocolo de descripción de sesión (SDP) es un protocolo para describir los parámetros de inicialización de contenidos transmitidos. El texto SDP contiene información sobre qué formatos de contenidos (esto es, códecs) admiten los clientes y el orden preferido de selección de códec de los clientes.

Los códecs de audio más habituales para las llamadas SIP son PCMU, PCMA, G.722, G.726 y L16. Si tanto el iniciador como el destinatario admiten diferentes códecs solapados, normalmente se seleccionará el códec con la prioridad más alta en el lado del receptor. La selección de códecs afecta al ancho de banda, por lo que debe tenerse muy en cuenta para cumplir con los requisitos de compatibilidad con otros UA SIP y mantener unos requisitos de ancho de banda adecuados para cada caso. Por ejemplo, en una red local en la que todos los clientes admiten L16, el audio sin comprimir es una buena opción. Sin embargo, si va a accederse al UA SIP a través de Internet usando un teléfono móvil 3G, PCMU es una opción mejor.

4.2 Llamadas en infraestructuras SIP complejas

Si la configuración de la infraestructura SIP es compleja, el inicio es algo distinto, porque la sesión de SIP se configura paso a paso para cada salto. Sin embargo, una vez configurada la sesión SIP, el tráfico normalmente no se enruta, sino que viaja directamente entre las diferentes partes, como en el ejemplo anterior.

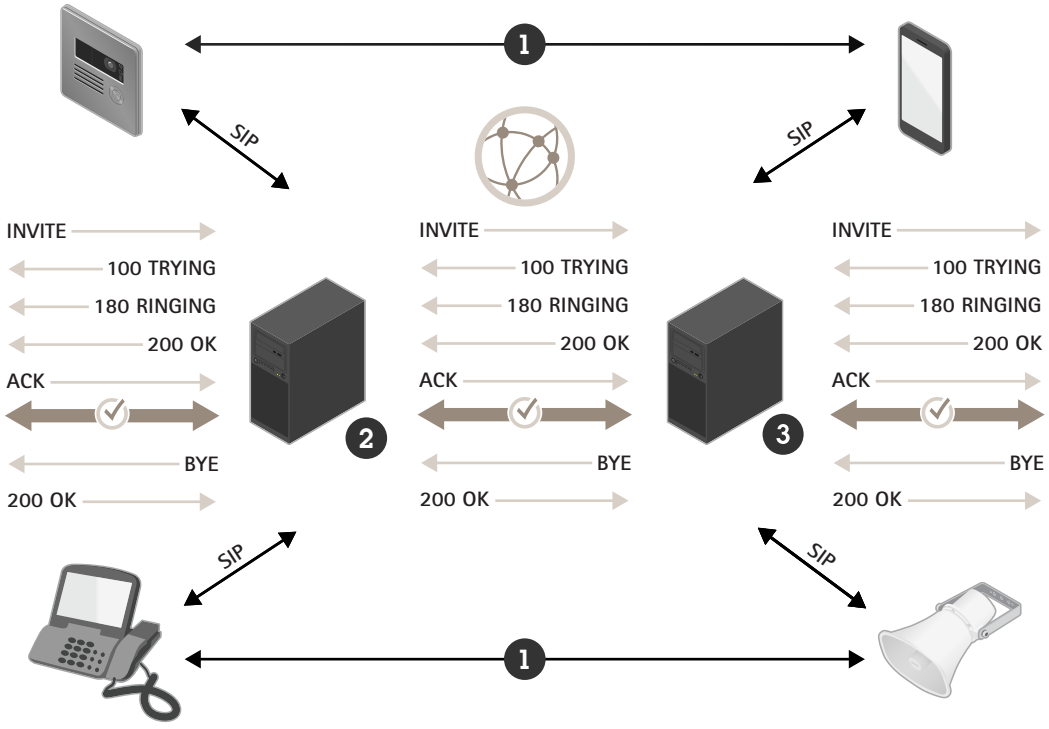


Figure 1. Configuración de llamadas en una infraestructura SIP compleja. Las transmisiones de contenidos RTP (1) viajan directamente entre las partes una vez que se ha configurado la sesión SIP entre el iniciador (2) y el destinatario (3).

5 MFDT: envío de comandos en llamadas SIP

La multifrecuencia de dos tonos (MFDT o DTMF) es un formato empleado para enviar información a través de una conexión telefónica. Las señales MFDT pueden enviarse en llamadas SIP y usarse para proporcionar instrucciones a un dispositivo SIP. Los caracteres que pueden utilizarse con la MFDT son los números de 0 a 9, las letras de A a D, * y #.

Por ejemplo, en una llamada a un portero automático con SIP, podría enviarse el carácter MFDT 5 desde el teclado del teléfono, configurado para que el receptor lo interprete como un comando para desbloquear la puerta.

Hay tres formas diferentes de enviar MFDT en una llamada SIP:

- El método tradicional dentro de banda, en el que la señal es en realidad un impulso de audio entrelazado con el flujo de audio. Sin embargo, es poco fiable y solo funciona con códecs sin comprimir.
- El método SIP INFO, en el que el carácter MFDT se envía en un mensaje SIP en el flujo de señalización. Es un método muy fiable y fuera de banda, pero su compatibilidad es limitada.
- El método RTP (RFC2833), en el que el carácter MFDT se codifica como paquete RTP y se envía fuera de banda. Es el procedimiento más aceptado y cuenta con una buena compatibilidad.

6 Entornos complejos y seguridad reforzada

Los entornos de red complejos, como las redes de empresa, pueden plantear dificultades cuando se utiliza SIP. Y lo mismo vale cuando quiere utilizar cifrado.

6.1 NAT traversal: encontrar el camino en redes complejas

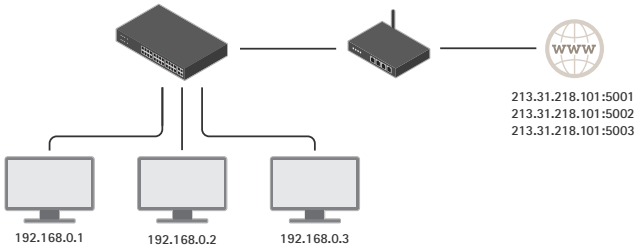
En un entorno de red más complejo, tal vez sea necesario utilizar la traducción de direcciones de red (Network Address Translation o NAT). La tecnología NAT es una representación pública de direcciones IP pertenecientes a una red local privada. Por tanto, todas las unidades de una subred privada comparten un prefijo de dirección IP común, por ejemplo 192.168.1.XXX. Esta es la dirección que utilizan al comunicarse entre sí. Cuando se comunican con otra red, esta dirección se convierte en la dirección pública del router, a las que se añade la asignación de puerto.

- 192.168.1.24 => 184.13.12.33:44221
- 192.168.1.121 => 184.13.12.33:24325, y así sucesivamente.

Como la tabla de traducciones se almacena en el router, en la mayoría de los casos un usuario externo no puede conocer la dirección de un dispositivo al que se ha aplicado NAT. Durante la comunicación por SIP, esto puede generar uno de los problemas siguientes:

- Imposibilidad de iniciar, actualizar o finalizar una sesión, esto es, no es posible llamar, retener una llamada o colgar.
- No hay flujos de contenido.
- Flujos de contenido unidireccionales.

NAT traversal: la NAT convierte la dirección de origen de cada paquete en una dirección IP pública con diferentes puertos de origen.



Para resolver estos problemas, SIP permite utilizar tres técnicas de NAT diferentes:

- STUN: un sistema para preguntar a un servidor de un lugar conocido cuál es la dirección pública de la unidad. El servidor STUN comunica la IP pública y la asignación de puertos utilizada para realizar la petición. El resultado se utiliza después en la señalización y la transferencia de contenidos, y la solución funciona en la mayoría de las situaciones.
- TURN: cuando se usa TURN, todo el tráfico pasa a través de un servidor conocido. Esta solución implica un coste extra, ya que la máquina en la que se aloja el servidor TURN debe tener suficiente potencia para enrutar todos los contenidos para cada cliente usando el servicio. Se trata de una solución más cara, pero puede ser útil en algunas situaciones en las que STUN no funciona.
- ICE: el protocolo ICE recopila todas las direcciones IP que puede encontrar en relación con un UA SIP e intenta calcular cuál debe usarse. Si se utiliza en combinación con STUN y TURN en el UA SIP iniciador y receptor aumenta las probabilidades de que las llamadas SIP se establezcan correctamente.

6.2 Uso de cifrado con SIP

El tráfico de señalización SIP normalmente se envía a través del protocolo UDP sin conexión. También puede enviarse a través de TCP, en cuyo caso también puede cifrarse utilizando la seguridad de la capa de transporte (TLS).

Para garantizar que se utiliza una conexión segura para una llamada, el protocolo SIP utiliza un esquema de direccionamiento conocido como SIP seguro (SIPS), que requiere configurar el modo de transporte en TLS. Al realizar una llamada, se añade a la dirección SIP marcada el prefijo “sips:” en lugar de “sip”. Por ejemplo, sips:bob@biloxi.ex.com en lugar de sip:bob@biloxi.ex.com. De este modo, cada salto debe estar protegido mediante TLS y requiere del extremo receptor el uso del mismo nivel de seguridad. Al llamar a una dirección con prefijo sip al utilizar solo TLS se garantiza que el primer salto está cifrado.

Para obtener el máximo nivel de seguridad, deben adoptarse las siguientes medidas:

- El modo de transporte debe ajustarse en TLS.
- El prefijo sips debe utilizarse en todo momento.
- Debe utilizarse SIP INFO para enviar tonos MFDT, ya que se envían por el canal cifrado.

Recuerde que no todos los clientes son compatibles con SIP seguro.

7 Terminología de SIP

3G	Tecnología de telecomunicaciones móviles de tercera generación
API	Interfaz de programación de aplicaciones
Códec	Codificador-descodificador
Teléfono físico	Aparato que se utiliza para realizar llamadas telefónicas, es decir, un teléfono en el sentido tradicional
Softphone	Programa de software que permite realizar llamadas telefónicas
ICE	Interactive Connectivity Establishment
IP	Internet Protocol (protocolo de Internet)
Cliente móvil	Programa de software de un dispositivo móvil que permite realizar llamadas telefónicas
NAT	Network Address Translation, Traducción de direcciones de red
PBX	Private Branch Exchange, Centralita telefónica privada
PSTN	Public Switched Telephone Network, es decir, una red de telefonía tradicional
RTP	Real-time Transport Protocol, protocolo de transporte en tiempo real
SDP	Session Description Protocol, protocolo de descripción de sesión
SIM	Subscriber Identity Module, módulo de identidad de suscriptor
SIP	Protocolo de inicio de sesión
Servidor SIP	Componente principal de una PBX IP, responsable de la configuración y la anulación de llamadas. También se conoce como proxy SIP o servidor de registro.
SIPS	SIP seguro
URI SIP (dirección SIP)	Uniform Resource Identifier, identificador de recursos uniforme. La dirección única del UA SIP.
STUN	Session Traversal Utilities for NAT
TCP	Transmission Control Protocol, protocolo de control de transmisión
TLS	Transport Layer Security, seguridad de la capa de transporte
TURN	Traversal Using Relays around NAT
UDP	User Datagram Protocol, protocolo de datagramas de usuarios
UA	User Agent, agente de usuario. Los dos extremos de una sesión de comunicación.
VoIP	Voz por IP

T10175991/ES/M2.2/2204

Acerca de Axis Communications

Axis contribuye a crear un mundo más inteligente y seguro a través de soluciones en red que mejoran la seguridad y suponen una nueva manera de hacer negocios. Como líder de la industria del vídeo en red, Axis pone a su disposición productos y servicios de videovigilancia y analítica, control de accesos y sistemas de audio e intercomunicación. Axis cuenta con más de 3800 empleados especializados en más de 50 países, y proporciona soluciones a sus clientes en colaboración con empresas asociadas de todo el mundo. Fundada en 1984, su sede central se encuentra en Lund, Suecia.

Para más información sobre Axis, visite nuestro sitio web *axis.com*.