

WHITEPAPER

SIP – eine Einführung

September 2024

Zusammenfassung

Das Session Initiation Protocol (SIP) stellt eine zusätzliche Schnittstelle für die Systemintegration für Sicherheitsprodukte dar. SIP ist ein häufig eingesetzter Standard in der Telekommunikationsbranche, der flexiblere Zwischenverbindungen und mehr Einsatzmöglichkeiten im Alltag ermöglicht. Offene, standardisierte Schnittstellen werden von Systemintegratoren, Entwicklern und Endnutzern gefordert. Sie erhöhen den Wert von Produkten, da diese in einer Vielzahl von Systemen eingesetzt werden können. SIP-fähige Axis Produkte sind für Sicherheits- ebenso wie für Kommunikationslösungen vorgesehen.

Die Einrichtung eines SIP-Systems kann sehr einfach sein. Bei komplizierten Netzwerk-Topologien, oder wenn Sicherheitsanforderungen zu erfüllen und eine zusätzliche Rufbehandlung erforderlich sind, bedarf es jedoch SIP-Server- und NAT-Traversal-Techniken, für die der Installateur oder Techniker ein umfassendes technisches Verständnis haben muss.

Inhalt

1	Einführung	4
2	So funktioniert es:	4
2.1	Peer-to-Peer-Setup – die einfache Art	4
2.2	Verwenden eines SIP-Servers (Nebenstellenanlage, PBX) – Erweiterung der Möglichkeiten	5
2.3	Verwendung einer SIP-Verbindung – Zuweisung einer Rufnummer	5
3	Vereinheitlichte Kommunikation (UC)	6
4	Was bei einem normalen SIP-Anruf geschieht	7
4.1	SDP – Verhandlung des zu verwendenden Formats	7
4.2	Anrufe in einer komplexen SIP-Infrastruktur	8
5	DTMF – Senden von Befehlen in SIP-Anrufen	8
6	Komplexe Umgebungen und höhere Sicherheit	9
6.1	NAT Traversal – Navigation in komplexen Netzwerken	9
6.2	Verschlüsselung in Verbindung mit SIP	10
7	SIP-Terminologie	10

1 Einführung

Das Session Initiation Protocol (SIP) dient zur Initiierung, Aufrechterhaltung und Beendigung von Multimediasitzungen zwischen verschiedenen Parteien. Diese Sitzungen bestehen üblicherweise aus Audio, enthalten aber manchmal auch Video. SIP ist das Standardprotokoll für VoIP-Anwendungen (VoIP=Voice over IP) und auf vereinheitlichte Kommunikationsplattformen (UC) (siehe Abschnitt 3).

SIP bietet eine Möglichkeit, Ihre Netzwerkprodukte von Axis zu verbinden, zu integrieren und anzuschließen. Sie wird von allen Axis Netzwerk-Lautsprechern, allen Axis Netzwerk-Sprechanlagen und ausgewählten Axis Systemgeräten und Axis Kameras unterstützt.

2 So funktioniert es:

Für eine Kommunikation über SIP werden mindestens zwei SIP-Clients benötigt. Ein SIP-Client kann ein SIP-Hardphone, ein Softphone, ein Mobile Client oder ein SIP-fähiges Axis Produkt sein.

Jedem SIP-Client wird eine eigene SIP-Adresse zugewiesen. Eine SIP-Adresse ähnelt einer E-Mail-Adresse, hat aber das Präfix „sip:“.

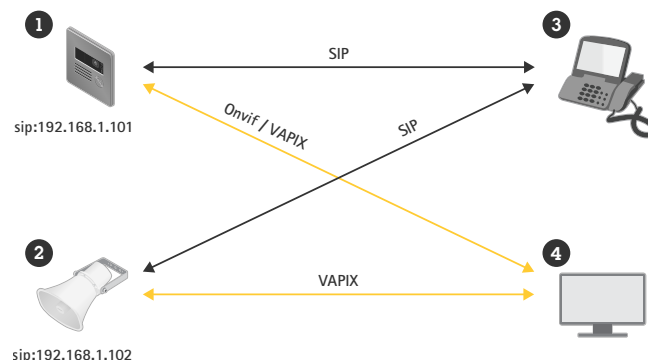
Zum Beispiel sip:bob@axis.com [sip:<benutzer@><anbieter>]. Diese Kennung kann bei unterschiedlichen Geräten verwendet werden, ähnlich wie bei einer Telefonnummer, die zu einer SIM-Karte gehört.

2.1 Peer-to-Peer-Setup – die einfache Art

SIP-Systeme können verschiedene Formen haben. In seiner einfachsten Form besteht das System aus zwei oder mehr SIP-Nutzeragenten (User Agents, UA), die direkt miteinander kommunizieren. Dies wird als Peer-to-Peer-Setup, Direktruf-Setup oder lokales Setup bezeichnet. Eine typische SIP-Adresse hat die Form sip:<lokal-ip>, zum Beispiel SIP:192.168.0.90.

Beispiel: In einem einfachen Setup können diese Axis-Produkte (1, 2) SIP für die Einrichtung von Audio- bzw. Videokommunikation mit anderen SIP-Geräten (3) im gleichen Netzwerk verwenden, ohne dafür einen Server oder eine Nebenstellenanlage zu benötigen.

Zugleich können sie wie jedes andere Axis-Gerät mithilfe offener API, mit VAPIX oder dem ONVIF Profil S an das Video Management System (4) angeschlossen werden.



Für einen Peer-to-Peer-Anruf von einer UA zu einer anderen in einem lokalen Netz ist nur die SIP-Adresse erforderlich, die die IP-Adresse der UA enthält.

2.2 Verwenden eines SIP-Servers (Nebenstellenanlage, PBX) – Erweiterung der Möglichkeiten

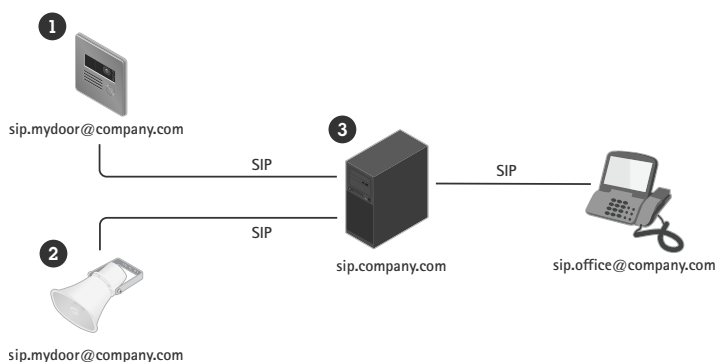
Eine SIP-basierte VoIP-Infrastruktur ist sehr gut skalierbar. Die nächste Größenordnung ist die Verwendung eines SIP-Servers oder einer Nebenstellenanlage (Private Branch Exchange, PBX) als zentraler Knotenpunkt. Die SIP-UAs registrieren sich beim Registrar des Servers und können dann die anderen UAs einfach anrufen, indem sie die Nebenstellenummer der TK-Anlage wählen.

Eine typische SIP-Adresse in diesem Fall hätte die Form sip:<benutzer>@<domäne>. Eine Alternative wäre sip:<benutzer>@<registrar-ip>, zum Beispiel sip:6007@meinsipserver.net. Eine Nebenstellenanlage funktioniert wie eine herkömmliche Telefonzentrale: Sie zeigt den aktuellen Status der Clients an, erlaubt Rufumleitungen, Sprachnachrichten, Weiterleitungen und vieles mehr.

Ein SIP-Server enthält meist einen Proxy, einen Registrar und eine Umleitungsfunktion. Proxies leiten Anrufe weiter und stellen zusätzliche Logik für eingehende Anrufe bereit. Registrare nehmen Registrierungsanfragen an und fungieren als Standortdienst für die von ihm verwaltete Domäne. Umleitungsserver leiten den Client an eine alternative SIP-Adresse um.

Der SIP-Server kann als lokale Einheit eingerichtet werden oder sich außerhalb des Standorts befinden. Er kann vor Ort oder in der Cloud gehostet werden. Bei standortübergreifenden SIP-Anrufen werden die Anrufe normalerweise zunächst über eine Reihe von SIP-Proxies geleitet. Diese Proxies fragen den Standort der SIP-Adresse ab, die erreicht werden soll.

Beispiel: Axis Produkte (1, 2) können lokal oder extern mit einem SIP-Server (3) verbunden werden. Der Server übernimmt den Aufbau und Abschluss der Anrufe zwischen SIP-Geräten innerhalb des lokalen Netzwerks oder über das Internet. In diesem Setup ist die SIP-Adresse des Geräts von seiner IP-Adresse unabhängig. Der SIP-Server ermöglicht den Zugang zu dem Gerät, solange es im Server registriert ist.



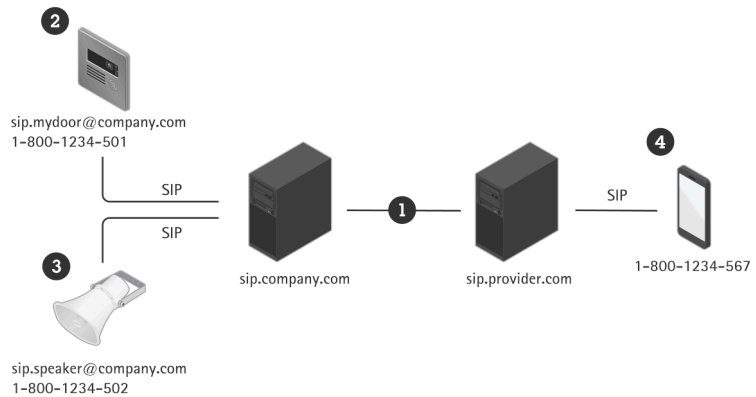
Um Ihr Gerät mit einem SIP-Server zu nutzen, müssen Sie beim Server einen Account mit einer spezifizierten Benutzer-ID und einem Passwort einrichten. Zur Anmeldung Ihres Geräts beim Server müssen Sie einen Account am Gerät einrichten und die Server-Adresse, die Benutzer-ID und das Passwort eingeben.

2.3 Verwendung einer SIP-Verbindung – Zuweisung einer Rufnummer

Mit einer SIP-Verbindung können SIP-UAs in das reguläre Telefonnetz (PSTN) geschaltet werden. So können Sie dem SIP-UA auch eine herkömmliche Telefonnummer zuweisen.

Cloudbasiertes SIP-Trunking ist ein moderner Ansatz, der das Internet für die Bereitstellung von Anrufen und anderen Kommunikationsdiensten nutzt. Diese Methode macht physische Telefonleitungen überflüssig und erleichtert so die Integration in Cloud-Lösungen und VoIP-Systeme.

Beispiel: Mithilfe der SIP-Verbindung (1) eines Serviceanbieters können Sie Ihren Geräten (2, 3) externe Telefonnummern zuweisen. Auf diese Weise können Sie Anrufe zwischen einem Netzwerk-Lautsprecher oder einer IP-Türsprechanlage und regulären Telefonen (4) tätigen.



Bei Verwendung zusammen mit einer SIP-Verbindung stellt das Gerät wie oben beschrieben eine Verbindung zum Server her.

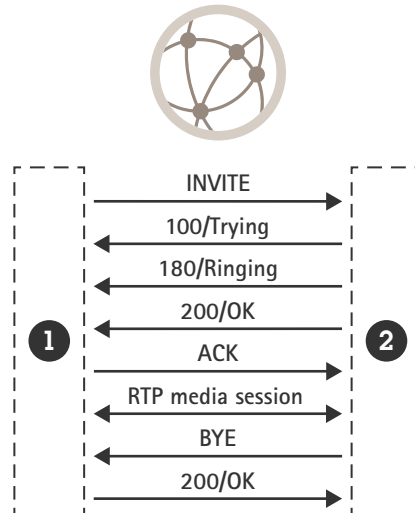
3 Vereinheitlichte Kommunikation (UC)

Vereinheitlichte Kommunikation (UC) bezeichnet die Integration verschiedener Kommunikationstools und -technologien in ein einziges, zusammenhängendes System. SIP spielt bei UC eine wichtige Rolle, da es die nahtlose Interaktion zwischen verschiedenen Kommunikationskanälen, wie Sprache, Video, Instant Messaging und Präsenzinformationen, aktiviert. Durch die Nutzung von SIP können Unternehmen eine vereinheitlichte Kommunikationsumgebung schaffen, die die Zusammenarbeit verbessert, die Produktivität steigert und eine konsistente Benutzer, Anwender über mehrere Geräte und Plattformen hinweg ermöglicht.

UC kann lokal in On-Prem-Lösungen oder über Cloud-Lösungen (UC as a Service - UCaaS) bereitgestellt werden. Beispiele von Dienstleistern für Cloud-Lösungen sind Cisco Webex, Microsoft Teams und Zoom.

4 Was bei einem normalen SIP-Anruf geschieht

Um bei einem SIP-Anruf Daten zwischen den Benutzeragenten auszutauschen, die den Anruf tätigen und empfangen, ist eine Abfolge von Schritten erforderlich.



Beim Tätigen des Anrufs sendet der Initiator-UA (1) eine Anfrage oder ein INVITE an die SIP-Adresse des Empfänger-UA (2). Das INVITE enthält einen SDP-Hauptteil (Session Description Protocol), in dem die verfügbaren Medienformate sowie die Kontaktdaten des Anrufers beschrieben werden.

Beim Erhalt des INVITE wird dieses vom Empfänger umgehend mit einer 100 TRYING-Antwort bestätigt.

Der empfangende UA vergleicht die angebotenen, im SDP beschriebenen Medienformate mit seinen eigenen. Wird ein gemeinsames Format gefunden, benachrichtigt der UA den Empfänger über den eingehenden Anruf und sendet eine vorläufige Antwort an den anrufenden Benutzeragenten – 180 RINGING.

Nimmt der Empfänger den Anruf an, wird eine 200 OK-Antwort an den Anrufer gesendet, um zu bestätigen, dass eine Verbindung aufgebaut wurde. Diese Antwort enthält ein verhandeltes SDP mit den zu verwendenden Medienformaten sowie dem Ziel der Medienstreams.

Die ausgehandelten Medienstreams werden nun mit RTP (Real-time Transport Protocol) sowie anhand der auf dem ausgehandelten SDP beruhenden Parameter eingerichtet, sodass die Medien direkt zwischen den beiden Parteien übertragen werden. Der Anrufer sendet über SIP eine Bestätigung (ACK), um darauf hinzuweisen, dass die Medienstreams wie vereinbart eingerichtet wurden. Die SIP-Sitzung ist zwar weiterhin aktiv, jedoch nicht mehr an der Medienübertragung beteiligt.

Beschließt eine der Parteien, den Anruf zu beenden, sendet sie eine neue Anfrage – BYE. Der Empfänger bestätigt den Erhalt des BYE mit einem 200 OK und die RTP-Medienstreams werden beendet.

4.1 SDP – Verhandlung des zu verwendenden Formats

Das Session Description Protocol (SDP) ist ein Format zur Beschreibung der Initialisierungsparameter für Streaming-Medien. Das SDP-Gehäuse enthält Informationen darüber, welche Medienformate (Codecs) von den Clients in welcher Reihenfolge bevorzugt ausgewählt werden.

Typische Audio-Codexs für SIP-Anrufe sind PCMU, PCMA, G.722, G.726 und L16. Falls Initiator und Empfänger mehrere überlappende Codexs unterstützen, wird normalerweise der Codex mit der höchsten Priorität auf Empfängerseite ausgewählt. Die Wahl der Codexs wirkt sich am Ende auf die Bandbreite aus, also will dies gut überlegt werden, um die Kompatibilitätsanforderungen anderer SIP-UAs zu erfüllen und einen passenden Bandbreitenbedarf für den jeweiligen Anwendungsfall zu erhalten. So funktioniert etwa in einem lokalen Netz, in dem alle Clients L16 unterstützen, unkomprimiertes Audio recht gut. Soll jedoch über das Internet über ein Mobiltelefon auf den SIP-UA zugegriffen werden, ist PCMU die bessere Wahl.

4.2 Anrufe in einer komplexen SIP-Infrastruktur

Bei einem komplexeren Setup der SIP-Infrastruktur ergibt sich ein etwas abweichender Aufbau, da die SIP-Sitzung für jeden Hop schrittweise eingerichtet wird. Sobald die SIP-Sitzung aber einmal eingerichtet ist, werden die Signale normalerweise nicht geroutet, sondern laufen direkt zwischen den verschiedenen Parteien, wie im obigen Beispiel.

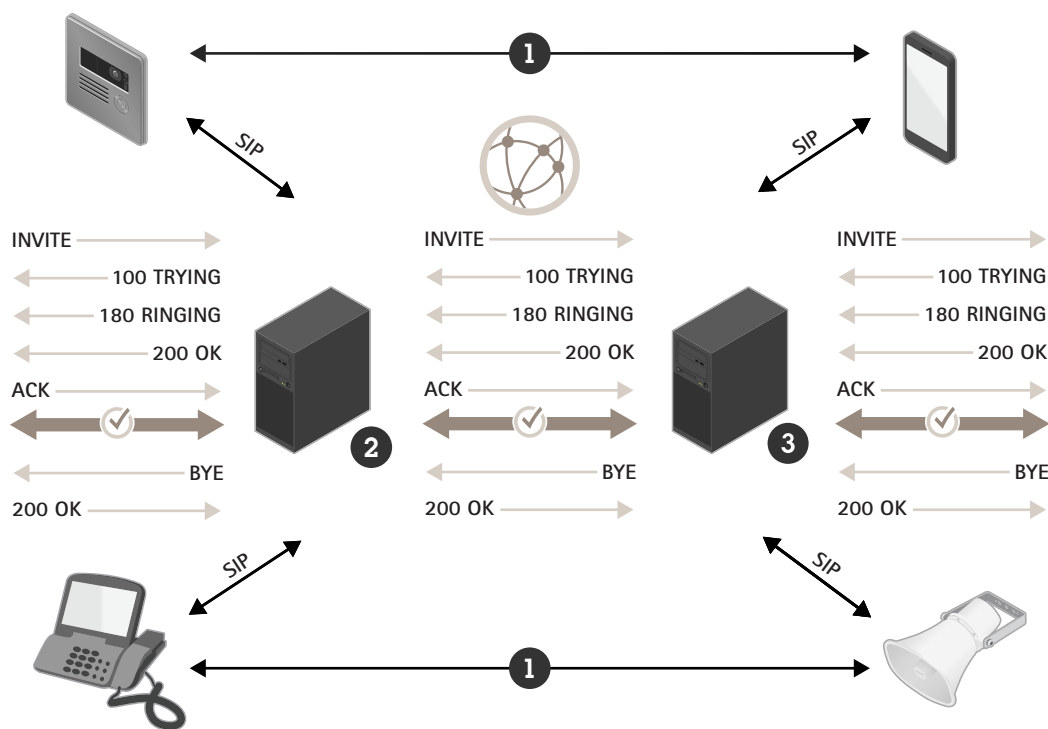


Figure 1. Anruf-Setup in einer komplexen SIP-Infrastruktur. Media RTP-Streams (1) verlaufen direkt zwischen den Parteien, nachdem die SIP-Sitzung zwischen Initiator (2) und Empfänger (3) eingerichtet wurde.

5 DTMF – Senden von Befehlen in SIP-Anrufen

Dual-Tone Multiple-Frequency (DTMF) ist ein Format zum Senden von Informationen über eine Telefonverbindung. DTMF-Signale können in SIP-Anrufen gesendet werden, um SIP-Geräten Anweisungen zu erteilen. Der DTMF-Zeichensatz besteht aus den Ziffern 0-9, den Buchstaben A-D, * und #.

So könnte z. B. während eines Anrufs an eine SIP-fähige IP-Türsprechanlage das DTMF-Zeichen „5“ über die Telefontastatur senden, was so konfiguriert werden kann, dass sie der Empfänger als Entriegelungsbefehl der Tür interpretiert.

Es gibt drei Möglichkeiten zum Senden von DTMF in einem SIP-Anruf:

- Das traditionelle Band-interne Verfahren, bei dem das Signal eigentlich ein in den Audiostream verschachtelter Audioimpuls ist. Dies ist jedoch unerwünscht und funktioniert nur bei nicht komprimierten Codecs.
- Das Verfahren SIP INFO, bei dem das DTMF-Zeichen in einer SIP-Nachricht im Signalstream gesendet wird. Dieses Verfahren ist sehr zuverlässig und bandextern, wird aber nur begrenzt unterstützt.
- Das RTP-Verfahren (RFC2833), bei dem das DTMF-Zeichen als RTP-Paket verschlüsselt und bandextern gesendet wird. Dies ist der De-facto-Standard, der umfassend unterstützt wird.

6 Komplexe Umgebungen und höhere Sicherheit

In komplexen Netzwerkkumgebungen wie Unternehmensnetzwerken können bei Verwendung von SIP Schwierigkeiten auftauchen. Ebenso, falls Sie Verschlüsselung einsetzen möchten.

6.1 NAT Traversal – Navigation in komplexen Netzwerken

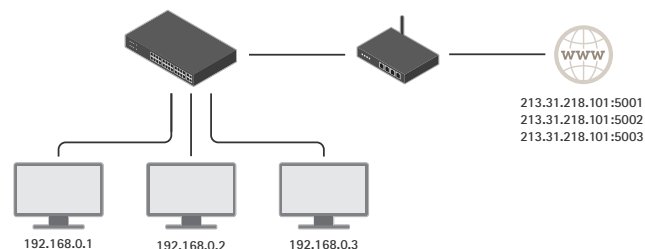
Eine komplexere Netzwerkkumgebung erfordert möglicherweise Network Address Translation (NAT). NAT ist eine Möglichkeit zur öffentlichen Darstellung von IP-Adressen in einem privaten lokalen Netz. Dabei teilen sich alle Einheiten in einem privaten Subnetz ein gemeinsames IP-Adresspräfix wie etwa 192.168.1.XXX. Unter dieser Adresse kommunizieren sie miteinander. Bei der Kommunikation mit einem anderen Netzwerk wird diese Adresse in die öffentliche Adresse des Routers übersetzt und mit einer Portzuweisung versehen:

- 192.168.1.24 => 184.13.12.33:44221
- 192.168.1.121 => 184.13.12.33:24325, und so weiter.

Weil die Übersetzungstabelle im Router gespeichert ist, kann in der Regel kein externer Benutzer die Adresse eines NAT-Einzelgeräts feststellen. Das kann bei der Kommunikation über SIP zu folgenden Problemen führen:

- Unfähigkeit, eine Sitzung zu eröffnen, zu aktualisieren oder zu beenden, d.h. keine Möglichkeit zum Anrufen, Halten oder Beenden.
- Kein(e) Medienstream(s).
- Eindirektionale Medienstreams.

NAT Traversal: NAT ändert die Quelladresse jedes Pakets in eine öffentliche IP-Adresse mit unterschiedlichen Quellports.



Zur Lösung dieser Probleme unterstützt SIP drei NAT-Techniken:

- STUN – Hierbei wird ein Server an einem bekannten Ort nach der öffentlichen Adresse der Einheit gefragt. Der STUN-Server meldet die öffentliche IP und Portzuweisung für die Anfrage zurück. Das Ergebnis wird daraufhin für die Signal- und Medienübertragung verwendet, was in den meisten Situationen funktioniert.
- TURN – Bei TURN verläuft der gesamte Datenaustausch über einen bekannten Server. Das bedeutet einen zusätzlichen technischen Aufwand, weil der Rechner, der den TURN-Server hostet, leistungsfähig genug sein muss, um die gesamten Medien jedes Clients, der den Service nutzt, weiterleiten zu können. Dies ist eine teurere Lösung, die aber in manchen Situationen funktionieren kann, die STUN nicht bewältigt.
- ICE – Das ICE-Protokoll sammelt alle IP-Adressen mit Verbindung zu einem SIP-UA, die es finden kann, und versucht zu berechnen, welche er verwenden soll. In Verbindung mit STUN und TURN am anrufenden sowie am angerufenen SIP-UA erhöht dies die Chancen für erfolgreiche SIP-Anrufe.

6.2 Verschlüsselung in Verbindung mit SIP

SIP-Signale werden normalerweise über das verbindungslose UDP-Protokoll ausgetauscht. Sie können aber auch über TCP gesendet werden, wobei sie außerdem gemäß TLS (Transport Layer Security) verschlüsselt werden können.

Um sicherzustellen, dass ein Anruf über eine sichere Verbindung getätigt wird, verwendet das SIP-Protokoll das sogenannte Adressierungsschema Secure SIP (SIPS), wofür die Beförderungsart auf TLS gesetzt werden muss. Bei einem Anruf erhält die gewählte SIP-Adresse das Präfix „sips:“ anstelle von „sip“; zum Beispiel: sips:bob@biloxi.ex.com anstatt sip:bob@biloxi.ex.com. Dadurch wird gefordert, dass jeder Hop mit TLS gesichert sein muss und dass die Empfängerseite dieselbe Sicherheitsstufe anwendet. Der Anruf einer Adresse mit Präfix „sip“ über TLS stellt lediglich sicher, dass der erste Hop verschlüsselt ist.

Für größte Sicherheit sollten folgende Maßnahmen ergriffen werden:

- Die Beförderungsart sollte auf TLS gesetzt sein.
- Es sollte grundsätzlich das Präfix „sips“ verwendet werden.
- Zum Senden von DTMF-Tönen sollte SIP INFO verwendet werden, da dieses über den verschlüsselten Kanal gesendet wird.

Bitte beachten Sie, dass nicht alle Clients Secure SIP unterstützen.

7 SIP-Terminologie

API	Anwendungsprogrammierschnittstelle (engl. Application Programming Interface)
Codec	Coder-Decoder
Hardphone	Hardware für Telefonanrufe, d. h. ein Telefon
ICE	Interactive Connectivity Establishment
IP	Internet Protocol
Mobile Client	Softwareprogramm auf einem Mobilgerät, das Telefonanrufe tätigt
NAT	Network Address Translation
Nebenstellenanlage	Private Branch Exchange

PSTN	Public Switched Telephone Network, das normale Telefonnetz
RTP	Real-Time Transport Protocol
SDP	Session Description Protocol
SIM	Subscriber Identity Module
SIP	Session Initiation Protocol
SIP-Server	Hauptkomponente einer IP-Nebenstellenanlage, die Anrufaufbau und -beendigung abwickelt. Wird auch als SIP-Proxy oder Registrar bezeichnet.
SIPS	Secure SIP
SIP URI (SIP-Adresse)	Uniform Resource Identifier. Die eindeutige Adresse des SIP-UA.
Softphone	Softwareprogramm, das Telefonanrufe tätigt
STUN	Session Traversal Utilities für NAT
TCP	Transmission Control Protocol
TLS	Transport Layer Security
TURN	Traversal Using Relays around NAT
UA	User Agent. Beide Enden einer Kommunikationssitzung.
UC	Vereinheitlichte Kommunikation
UDP	User Datagram Protocol
VoIP	Voice over IP

Über Axis Communications

Axis ermöglicht eine intelligente und sichere Welt durch Lösungen zur Verbesserung der Sicherheit und Geschäftsperformance. Als Unternehmen für Netzwerktechnologie und Branchenführer bietet Axis Lösungen in den Bereichen Videosicherheit, Zutrittskontrolle sowie Intercoms und Audiosysteme. Sie werden verstärkt durch intelligente Analyseanwendungen und unterstützt durch gute Schulungen.

Axis beschäftigt rund 4.000 engagierte Mitarbeiter in über 50 Ländern und arbeitet weltweit mit Technologie- und Systemintegrationspartnern zusammen, um den Kunden Lösungen anbieten zu können. Axis wurde 1984 gegründet und der Hauptsitz befindet sich in Lund, Schweden