VoIP Encryption Briefing Paper – Version 1

Introduction

VoIP, like all other services that use the internet, requires protection from unauthorised use and intrusion.

Please also see the Comms Council UK best practice papers on:

Security considerations for provisioning end-user equipment and
Protecting IP PBXs

This paper discusses the pros and cons of using encryption to protect:

1. The voice media i.e. the actual conversation between the parties making the call.
2. The signalling i.e. the information used to set-up, manage and end calls.

The use and need for encryption in VoIP based telephony using the public Internet is intuitively obvious and for some users it is required to protect the sensitive content of the voice communications (e.g. PCI compliance).

This document aims to provide an Comms Council UK member with a balanced view of the pros and cons of encryption and an understanding of the approach and challenges involved with providing encrypted services over real time protocols. The document will also provide an overview of the encryption technologies available and how they might operate in an ITSP environment.

About Comms Council UK

Comms Council UK is a membership-led organisation that both represents and supports telecommunications companies that provide services to both business and residential customers in the UK. We keep Britain talking in its various guises by providing or reselling voice services over data networks (VoIP) as well as other "over the top" applications including instant messaging and video.

The membership is a mixture of network operators, service providers, resellers, suppliers and consultants involved in a sector that is diversifying rapidly. Comms Council UK (formerly known as ITSPA) represents its members at a policy level, builds coalitions to collaborate on industry initiatives and provides a platform to help members prepare for change, learn about new trends and develop new business relationships.

Contact

For more information, please contact:

Comms Council UK
team@commscouncil.uk
020 3397 3312

DISCLAIMER: The information contained in this guidance is for your information only and is not intended to be relied on. It does not constitute legal professional advice, nor is it a substitute for you obtaining your own legal professional advice relevant to your circumstances. Comms Council UK accepts no liability whatsoever for any errors, omissions or statements contained in this guidance or for any loss which may arise from your use of this guidance. Comms Council UK owns absolutely and exclusively all copyright, moral rights and any other proprietary rights for their full terms throughout the world in respect of the information contained in this guidance.

The Risk

Adding encryption to any VoIP service adds cost and complexity and therefore needs to be balanced against its actual effect in reducing the risk of a successful attack. Different ITSPs have differing markets to serve and operate different networks that will therefore have differing perceptions of risk – a consumer service provider may have fewer concerns over voice interception than a provider targeting the financial services sector for example. To date, the occurrence of successful external intrusion of either media or signalling is an unknown but the fact that it is technically feasible, creates cause for concern. The high publicity created by recent thefts of un-encrypted static data creates a raising of consciousness for all forms of data to be encrypted.

The interception of voice calls is probably what the lay person feels is meant by encrypting his telephone calls. But it is important not to oversell the undoubted benefits of encryption, as to do so can lead to a false sense of security by lay users. It would be good practice for VoIP companies marketing "secure" and "encrypted" services to provide a layman's description, detailing under what circumstances and to what extent their calls are actually protected.

As discussed below, in all but a very few kinds of specialised applications, a VoIP call will not be validated and encrypted end-to-end – successful encryption can only be partially achieved in everyday, public services. This is because the majority of calls will cross networks and terminate on hardware not in the control of the originating provider. Additionally, in almost all circumstances the providers themselves will have access to the conversations.

Even so, it's clear that the pressure to provide encryption of real-time media and signalling for end users will only increase and it is equally clear that actual end-user benefits can be achieved by providing it.

Benefits of Using Encryption

Apart from the obvious benefits that encryption provides to protect the user's privacy from having their conversations listened into, there are several other key benefits to using encryption for VoIP services.

Encrypting SIP traffic prevents SIP ALG from identifying, inspecting and interfering with the SIP traffic. This means that with the majority of routers, encrypted SIP traffic will pass through and the SIP ALG will not be used even if it is active. For an ITSP this reduces the impact of faulty SIP ALG routers, reduces support costs and improves the overall experience of VoIP for end users. (Please note that if SIP ALG eventually improves, some features, to assist VoIP may be unavailable).

One of the key challenges for an ITSP is reducing toll fraud that occurs as a result of the hacking of SIP account credentials. For hosted providers in particular, encryption provides a mechanism to reduce account hijacking among other measures taken by ITSPs. Mutual TLS authentication can be used for the majority of VoIP phones that come pre-installed with a factory installed device certificate. The TLS connection will only be established if the certificate on the device is validated and matched to the correct account stored in the provider's database. Even if you know the MAC address and SIP account details you would not be able to use them on anything other than the designated device. Mutual TLS is the answer to the question, "how does an ITSP provide a VoIP service with zero fraud from a SIP account hijacking?"

Many consider the use of encryption with VoIP to be limited since the PSTN is not encrypted. Indeed, PSTN calls can be intercepted with crocodile clips on a telephone line but the majority of the PSTN is a closed network governed by individual companies controlling and limiting access that makes it difficult to intercept and record calls without major assistance or extreme skill. For end users, the portion of the network between themselves and their provider is much more of a wild west in terms of governance, access controls and opportunity for interception. The risk of call interception to a VoIP user is reduced by using encryption even for calls to PSTN.

The overall benefits of using encryption for VoIP services are clear and very much focused on protecting end users.

Real World Selection of Encryption Technologies

For the ITSP, the challenge with implementing encryption begins with the options available based on the capabilities of the phones and devices they wish to support or supply as part of their services. An ITSP that provides multiple services covering hosted and SIP trunking with customer based PBXs will have a wider set of equipment that needs to be taken into account. The ITSP also has to consider their own SBC or service infrastructure capabilities.

Generally speaking, all physical devices fall into the following 3 categories:

- Equipment with no encryption capabilities
- Equipment with limited capabilities or lagging behind industry requirements
- Equipment that provide up to date VoIP encryption

The overwhelming majority of the SIP phones supplied by ITSPs that provide up to date encryption are based on SIP over TLS to encrypt the signalling and Secure RTP to encrypt the media.

For software based VoIP there is a similar process for selection except that there may be a wider set of encryption options available to them. For in-house developed software the libraries and code they use or choose to develop will set the limitations of choice they have. For pure software services such as Skype
or WhatsApp, the capabilities of their engineering team means that they can implement any encryption mechanism they choose.

In the real world, an ITSP will be severely constrained in their choices or ability to provide encrypted services if they wish to offer a wide range of equipment for customers or allow users to bring their own SIP compliant equipment to their network. An ITSP offering a single handset manufacturer in a hosted model has the simplest of paths to offering encrypted services. There is an increased effort and expense to support more equipment and a decision will need to be made based on the tradeoff of supporting more devices versus the cost of doing so.

Good and Bad Encryption

It is not illegal to supply telephony services with voice encryption enabled. End user rights to privacy need to be considered and unencrypted telephone calls offer little protection from interception. The reality for ITSPs operating with SIP and the type of encryption they can make use of is server key based and lawful intercepts can still take place.

The fight over the rights to use certain types of encryption might one day take place between governments and the people. However, in practical terms, end to end encryption that protects conversations from interception is not trivial for an ITSP to implement or integrate with their existing services. For example call recording is not possible and it's unlikely that an ITSP offering Hosted VoIP or Sip Trunking would offer End to End encryption services. End to end encryption with ZRTP would not work for PSTN calls and would really only work for App based VoIP services.

Appendix: Background Material on Encryption Mechanisms

Technical options for encryption

Encrypting the Signalling

Signalling is handled by the Session Initiation Protocol (SIP). The primary function of SIP is handling the registration of devices (so that their location is known and they may receive calls) and for setting up and terminating calls. SIP also handles other functions including Instant Messaging and Presence. SIP also handles device authentication. Good practice dictates that whenever a device registers, makes a call or sends an IM, then the request should be authenticated. SIP uses an authentication mechanism known as HTTP digest which enables a device to authenticate with a server without ever sending the device's password over the network. This provides a reasonable level of security. SIP is defined in a number of standards documents; the primary definition is in RFC 3261, a lengthy and detailed document. RFC 3261 defines 3 network transports for SIP. These are:

- UDP (User Datagram Protocol). This is a connectionless transport protocol where each SIP message is exchanged on a *best efforts* basis. It is up to the higher protocol levels to check that a request has been received and to manage re-transmission. UDP imposes a limit on the size of a SIP message which can make delivering some services such as Instant Messaging difficult as the default delivery mechanism requires the entire IM to be delivered in a single SIP message. When SIP messages are sent over UDP they are sent in clear text and the content of those messages may be monitored. Despite these limitations, the majority of SIP deployments use UDP.

- TCP (Transmission Control Protocol). TCP is a connection oriented transport protocol. This means that the protocol itself ensures that all transmitted data is received in the correct order, relieving the higher protocols of this task. TCP removes the limitation on SIP message size, but it still transmits message in clear text.

- TLS (Transport Layer Security). TLS is an encrypted, connection oriented transport protocol. SIP messages sent over a TLS connection are protected against unauthorised monitoring. TLS, in common with TCP removes the size limitation on SIP messages which apply when UDP is used. TLS is widely used, all secure web sites, including banking web sites use TLS to protect the data sent between the browser and the web server. Many VoIP system support TLS as a SIP transport.

There are three modes of operation for a TLS connection:

1. Encryption Only: provides no certificate verification. Traffic will still be encrypted between the devices but without using certificates to verify each other's identity.

2. Server Only: single ended authentication mode where only the server side of the TLS connection is validated using certificates. This is common for web browsers and SIP clients that need to verify that they are connecting to the proper server before logging in with their username and password.

3. Mutual: involves a mutual authentication where both sides verify each other's certificate prior to sending any encrypted traffic.

There are many benefits to using TLS to encrypt SIP signalling traffic over and above the obvious benefit of protecting SIP messages from unauthorised monitoring. These include reducing the risk of fraud as most SIPVicious attacks do not use TLS (yet).

Encrypting the Media

The media component of VoIP calls is transported using the Realtime Transport Protocol (RTP). RTP runs over UDP. UDP is a good choice for media transmission as timely and efficient delivery of each packet is more important than ensuring that all packets are received and acknowledged. RTP transmits media (both audio and video) in clear text.

A related protocol, Secure RTP (SRTP), provides an encrypted media transport. SRTP defines a mechanism for encrypting and authenticating every RTP packet. This protects the media stream from unauthorised monitoring and also from replay attacks and RTP injection attacks. SRTP uses a symmetric encryption algorithm where the same key is used to encrypt and decrypt the media stream. The SRTP standard defines a mechanism to use AES, a widely used and trusted encryption algorithm. SRTP uses a different encryption key for each media stream. In a voice call this means that two encryption keys are used (one for transmitted audio, one for received). If a video stream is added then four keys will be used. SRTP is designed specifically for encrypting media, it mirrors the operation of UDP by using a best effort delivery. SRTP tolerates dropped packets and will not attempt to re-transmit those packets. This makes SRTP a much better choice for encrypting media streams than other options such as VPNs which are designed primarily for data. Most VPNs add too much overhead, they can double the size of each media packet. VPNs are designed to provide a reliable connection ensuring that all packets are delivered in the right order. This is exactly what is needed for data, but the overhead can reduce voice or video quality.

SRTP is widely implemented and is supported by many IP phones and soft-phones. The SRTP standard defines 3 key lengths when AES is used, 128 bits, 192 bits and 256 bits. The keys are ephemeral, generated at the start of the call and discarded at the end. Most SRTP implementations used 128 bit keys. In 2012, it was estimated that a brute force attack on a 128 bit AES key using currently available technology would take 1.02 x 10^{18} years (1 billion billion years), the universe is only 13.75 billion years old. While technology has improved in the last 5 years, a brute force attack remains completely infeasible. This means that 128 bit keys offer a very high level of encryption, particularly as keys are discarded at the end of the call.

SRTP does not define how the keys needed for a call are set up, this is left for a separate key exchange protocol.

SDES

Security Descriptors for Media Streams (SDES) is the most widely implemented key exchange protocol. Despite its name, SDES is not related to the old and discredited Data Encryption Standard (DES) encryption algorithm. SDES defines a mechanism for exchanging SRTP keys via the SIP signalling steam. Specifically SDES operates by adding additional a=crypto attributes to the Session Description Protocol (SDP) offer/answer used to set up the media streams in a voice or video call. To protect the key exchange, SDES must be used only when the signalling stream is encrypted using TLS.

ZRTP

ZRTP is an alternative key exchange protocol developed by Phil Zimmermann, creator of PGP encryption. ZRTP differs from SDES in that the key exchange operates over the media stream rather than the signalling stream. ZRTP was designed to provide peer-to-peer key exchange and security. This

means that encryption keys are agreed between the end-points of a call and that the process cannot be viewed by an intermediate device such as a SIP call router or a PBX. Keys agreed using SRTP are utilised by SRTP to encrypt a call. In practice ZRTP is difficult to implement in service provider networks as most calls are handled by intermediate devices which perform important roles, such as providing voice mail. ZRTP also complicates implementing call recording services which in many networks are required to meet regulatory and compliance needs.

MIKEY and DTLS

MIKEY (Multimedia Internet Keying) defines another mechanism for establishing encryption keys for use with SRTP. DTLS (Datagram TLS) is a version of TLS designed to work over UDP. While both have some technical advantages over other media encryption methods, they are less widely implemented.