

## Introducción a VOIP

VOIP is an acronym which stands for "Voice Over IP". Most of us are familiar with the "Public Switched Telephone System" (PSTN), which allows us to contact people around the globe by dialling a sequence of numbers. VOIP offers an alternative, which works by routing digitised voice signals over IP networks, such as Company Intranets, or in some cases the public Internet.

On the face of it, the PSTN hasn't really changed much in more than 100 years. There have been many technology changes and improvements, such as tone dialling and Caller ID, but as far as the user is concerned, it's still a matter of dialling (more recently, pressing) a sequence of numbers, and getting connected to the person who's number was dialled. However, what happens behind the scenes to make this happen has changed considerably in recent years.

VOIP isn't a particularly new technology; there are papers and patents about the subject dating back several decades, and there was some early VOIP software available as early as 1991. The basic principle is pretty simple; it is essentially the same technology that is used to stream music across the internet. Voice sounds are picked up by a microphone and digitised by the sound card. The digitised audio is then compressed using an audio codec. This works by removing redundant and unneeded data, while maintaining the legibility of the audio, to make the stream compact enough to be sent in real time over the network. The term codec is short for "enCODer/DECoder". The sounds are encoded at the sending end, sent over the network and then decoded at the receiving end, where they are played back over speakers or a headset.



The only requirements are a network connection between the two computers of an adequate speed, and matching codecs at each end.

Regular "off the shelf" PCs equipped with microphones, sound cards, headsets and a broadband connection fit the bill perfectly.

It is necessary, of course, that the two talking parties agree to use the same codec before making a call, so that the compression results in audio streams that can be decompressed properly by the system at the far end. Codec's are always in a state of flux, as anyone with a digital music player will know - mp3, wma, ogg, mp4, and aac are all file extensions that you may have seen on compressed music files from online music stores, and they are all different. Some music players will play all of them, some only a few, and some will play only a single specific type. Thankfully, there is some common ground in the telephony world that means that VOIP systems can usually negotiate with each other to find a codec that both sides can understand. Commonly used telephony codec's include G.711, G.729 and G.726, though there are many others, including proprietary systems. These codec's differ in two main ways. First, the amount of CPU power required to perform the compression and decompression, which has an impact on the type of hardware needed in the system (PC, PBX or telephone) and second, the size of the compressed audio stream or file, and therefore the amount of network bandwidth needed to transport the data between the two parties. This has an impact on the network infrastructure. It's a trade off - spend more CPU cycles on compressing the signal so that it is smaller, and you can spend less on network bandwidth.

Conversely, if you can afford huge amounts of network bandwidth, you can get away with slower, cheaper terminal equipment.

To be useful, a VOIP system needs some method for establishing and managing a connection, for example, calling the other computer, finding out if they accept the call, and closing the connection when a user hangs up. Because VOIP allows two way communication, and even conference calls, this part is more complex than simple audio streaming. Call management - session initiation, call setup and tear down, is one area in which VOIP systems fundamentally differ, and two VOIP users must be using the same system (or compatible ones) in order to be able to call each other.

Because most domestic Internet users don't have a permanent Internet address, domestic VOIP systems don't generally work by calling another computer direct - it is akin to having a telephone number that changes from time to time. Instead, each user of the service registers with an intermediate server, which maintains a record of their IP address all the time they are connected. A small application can be installed on each user's PC, which manages this data in conjunction with the server.

Another reason for using an intermediate server is that it eases the problem of getting VOIP to work through firewalls. Many firewalls block any data from the Internet that is not sent in response to a specific request. This makes it impossible to call another computer direct; because the called computer did not request any data from the caller, the call request would be blocked. By establishing a connection with a server, the VOIP software opens a channel of communication through which other computers can call it. Communication may continue using the server, or information may be passed via the server that allows the two computers to open a direct connection between them and continue using this communications channel.



### **Standards**

There are several 'standards' for communicating with Voice over IP. These can be split into 'open standards' that are available for anyone to use, and proprietary systems. H.323 and SIP fall into the former category, while Skype uses its own proprietary system.

H.323 is a standard for teleconferencing that was developed by the [International Telecommunications Union \(ITU\)](#). It supports full multimedia audio, video and data transmission between groups of two or more participants, and it is designed to support large networks. H.323 is network-independent: it can be used over networks using transport protocols other than TCP/IP. H.323 is still a very important protocol, but it has fallen out of use for consumer VOIP products due to the fact that it is difficult to make it work through firewalls that are designed to protect computers

running many different applications. It is a system best suited to large organizations that possess the technical skills to overcome these problems.

SIP (Session Initiation Protocol) is an [Internet Engineering Task Force \(IETF\)](#) standard signalling protocol for teleconferencing, telephony, presence and event notification and instant messaging. It provides mechanisms for setting up, managing, and tearing down connections, but not for transporting the audio or video data. It is probably now the most widely used protocol for managing Internet telephony. Like all IETF protocols, SIP is defined in a number of RFCs (Request For Comments), principally RFC 3261.

A SIP-based VOIP implementation may send the encoded voice data over the network in a number of ways. Most implementations use Real-time Transport Protocol (RTP), which is defined in RFC 3550. Both SIP and RTP are implemented on UDP, which, as a connectionless protocol, can cause problems with certain types of routers and firewalls. Usable SIP phones therefore also need to use STUN (for Simple Traversal of UDP over NAT), a protocol defined in RFC 3489 that allows a client behind a NAT router to find out its external IP address and the type of NAT device. Thanks to STUN, setting up SIP-based VOIP hardware or software behind a home or small office firewall should be a simple affair, but in practise it can still be troublesome.