A Web based Voice over IP Application

Audience

This document covers the architecture, design and working of Voice over Internet Protocol based communication systems followed by the Quality of Service section. The qos section focuses on methods to improve voice compression, security and jitter. At the end, the future of VoIP based telephone systems and its market is discussed in detail.

The reader is expected to have a basic knowledge of computer networks and to be able to understand the architecture and design of communication systems. An elementary background of OSI model and packet compression is desired to have a better understanding of quality issues.

This document can be used by students, researchers, network architects, and technical authors.

Table Of Contents

Table of Tables

Table Of Figures

1

1. Introduction

Voice communication using the voice over Internet Protocol (VOIP) is based on Session Initiation Protocol (SIP) has been on the surge over the recent years as it eliminates the dependency on dedicated infrastructure for traditional telephony communication systems. The advent of session initiation protocol in communication service has led to the development of various open source technologies implementing SIP, which allows Real Time Communication (RTC) and multimedia transfer over the IP networks. Voice over internet protocols relate to transmitted packets over voice traffic. This technology draws the big gap from the past circuit-switch transmission across networks that are available to the public. The main motive of the communication industry with regards to VoIP is to operate based on two primary principles. The two principals include data traffic and unification (Chakraborty et al., 2019). Another primary reason for IP telephony is the international existence of protocols in network devices and for the users. Additionally, an IP network goes after resource shared techniques, while PSTN is a facility meant to lower costs for every network user.

There are numerous approaches to offering quality service in internet protocol networks and their components. The best method of providing an efficient IP network is via provision as the compared complicated quality of service protocols. In a case in a network setting, a 30% provocation to the link comprising a traffic situation, in such instances, packets must establish a corrective measure as fast as possible.

The designer should consider the number of channels determining the capacity of routers delivering small voice packets forehand and the links of network occupancies. The network could behave similarly to a conventional network with lower occupancy levels. Generally, VoIP enables people to convert and compress voice signals and transport such voice information through a protocol-enabled network similar to Ethernet and wireless LAN and its components. On the other hand, VoIP applies internet protocols to manage voice packets over an IP network for efficiency and convenience (Choudhury et al., 2020). Today voice over internet protocol is the most widely used technology in communication. It sends packets through a packet switch-based network, making it easy to communicate through phone calls.

Interactive communication such as voice calling, call conferencing, have become very popular and have become a need for all kinds of industries. In a business office, which requires collaborative work, VoIP can replace traditional telephony for peer-to-peer communication. The wireless cellular communication technologies have to abide by regulated protocols defined by the authorities (3GPP) and require licensed bandwidth. The use of VOIP for communication has advantages over traditional Public Switched telephone network (PSTN), as it can be economically and easily employable in home or offices, provided the availability of Internet, thereby mitigating the need for a dedicated and standardized infrastructure.

In this research work, we have proposed an end-to-end VOIP architecture that leverages SIP based technologies and open source tools. Further, we discuss our proposed solution developed by integrating these tools to develop a scalable and reliable system, which is capable of providing high quality VOIP services. A client application is developed based on WebRtc to offer user interaction with the system. The proposed application can be integrated to a web or mobile app, after which the existing application will be capable of providing communication services. In this work, we have achieved end-to-end encrypted transportation of SIP messages over TLS to provide a secure communication mechanism. Furthermore, the application can be containerized and multiple instances of it can be deployed to scale up the business. It can be further extrapolated to provide state of the art VoIP features such as voice bots.

2

2. Technology and Protocols

2.1 Voice Signaling Protocols

Signaling protocols play a critical role when making a phone call using mobile phones connected to an internet source. It empowers components of the network to enable communication with one another. A phone call may be prescribed as a multimedia connection between those participating in multimedia participants for telephony regarding interest protocols. On the other hand, a link defines signaling conjoined with a call. The primary duties of signaling protocols take four main dimensions; establishment of a session, which is accepting, redirecting or accepting a call from senders (Chen et al., 2018). On the other hand, user location defines the first option of finding a place for callee by the caller. Call participant management enables endpoints to leave and join an existing session when permitted.

2.2 Session Initiation Protocol (SIP)

SIP defines communication protocols applicable in signaling and yielding to communication in a multimedia session like instant messaging, online gaming, and other services to the general public. Since their messages have body and headers, they are considered similar to web protocol HTTP. SIP applies a 5060 port to serve as its default protocol for UDP or TCP (Chen et al., 2018). All these components are interpreted as an authorized protocol for video over internet protocol and telephony.

2.3 WebRTC

The WebRTC development is driven by the idea to provide a simple way to into Real Time communication media, directly in the browser's page. This allows audio communication to work inside a web application eliminating the need to install plugins and download native apps. The inclusion of WebRTC enables high quality RTP media flowing between peer-to-peer communication and call conferencing on browsers such as Chrome, Mozilla etc. It also provides a secured media path to establish a session setup or message signaling. The signaling mechanism can negotiate with connection parameters and media features, described in SDP. This mechanism is based on JavaScript Session Establishment Protocol (JSEP) with capabilities of establishing a connection and handshaking. This protocol supports secured and encrypted signaling transport over DTLS/UDP or TLS/TCP stack. The real time media transport uses UDP datagram stream in encrypted form. To fulfill this need, WebRTC can communicate through Stream Control Transport Protocol (SCTP), Datagram Transport Layer Security and Secure Real Time Transport (SRTP). SRTP provides secured audio channels by interacting with the key management protocol DTLS that gives way to media and application data encryption by the exchange of encrypted keys.

The WebRTC implements media stream, RTC peer connection and RTC data channel APIs. The media stream API allows browser based applications to access the microphone. RTC peer connection negotiates with the server, facilitates session establishment and bandwidth management. The RTC data channel API provides end-to-end bi-directional secured data channels and implements a successful establishment of sessions within any type of topologies and network scenarios like VPNs. The signaling and media transport behind NAT, firewall traversals is deployed on public/private networks and tested.

3

3. Design and Architecture

3.1 Elements of a VoIP Network

SIP defines the server network. It is also right to note that many SIP endpoints can communicate, minus the involvement of SIP infrastructure. This suggestion is not easy to achieve in actual general practice. The following are the elements of a network infrastructure.

i. **Proxy server**

Despite being a mediator entity, proxy reacts as UAS servers, and also the UAC clients are important in raising requests in place of many clients (Chen et al., 2018). It does the routing with the motive of transferring the assigned task to another different entity closer to the other user-targeted.

ii. **User Agent**

This is a network component that serves the motive of receiving and generating messages. The user agent transmits SIP messages on this platform while the client acts as the server agent. Aspects of the SIP network save the information since its capable of identifying issues of compatibility. The other components of a network include a redirect server and gatekeeper.

3.2 VoIP Architecture

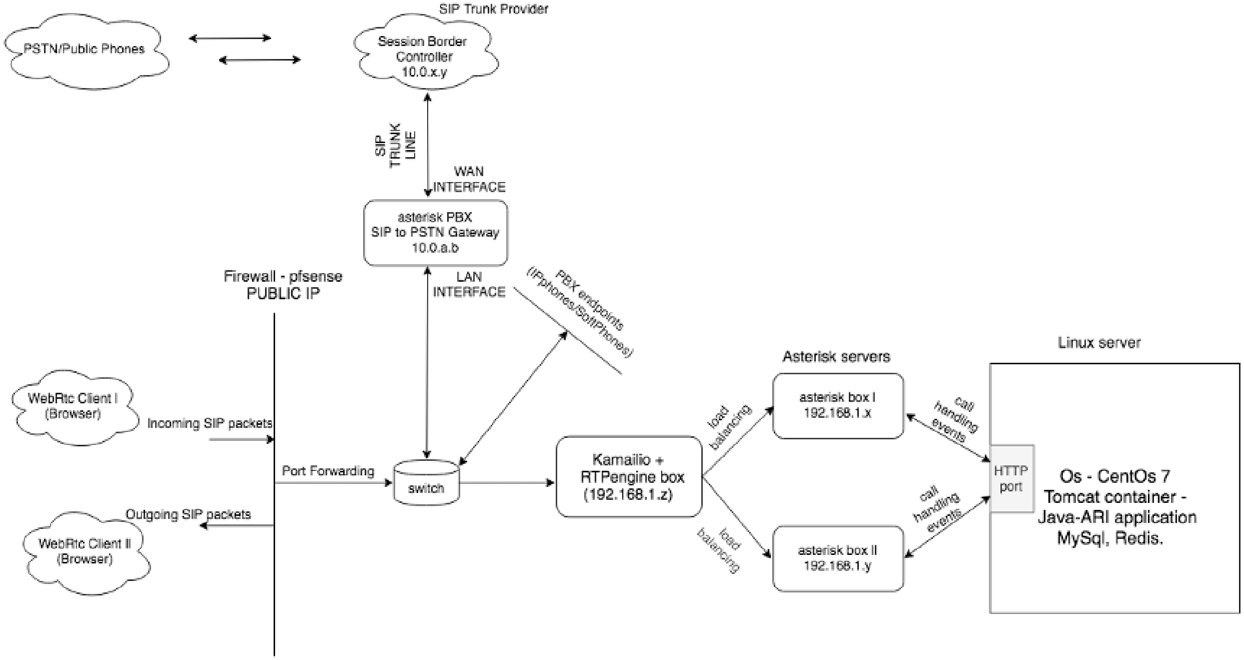

SIP Trunk Provider
PSTN/Public Phones
Session Border Controller 10.0.x.y
SIP TRUNK LINE
WAN INTERFACE
asterisk PBX SIP to PSTN Gateway 10.0.a.b
Firewall - pfsense PUBLIC IP
LAN INTERFACE
PBX endpoints (IPphones/SoftPhones)
Linux server
Asterisk servers
WebRtc Client I (Browser)
Incoming SIP packets
load balancing
asterisk box I 192.168.1.x
call handling events
Port Forwarding
switch
Kamailio + RTPengine box (192.168.1.z)
HTTP port
Os - CentOs 7 Tomcat container - Java-ARI application MySql, Redis.
call handling events
Outgoing SIP packets
WebRtc Client II (Browser)
load balancing
asterisk box II 192.168.1.y

Figure 1: Architecture of a VoIP System

Voice over internet protocol comprises equipment of end-users, call processor, server, gateway, and network components. The end-user becomes very important to access VoIP systems hence straightforward communication with the end-users. The end-user equipment helps in initiating and maintaining signals required to initiate a call of networks of internet protocols. It is also helpful in converting voice to data packets, network surveillance capability, network communications, and monitoring the communication process over internet-connected devices (Taruk et al., 2018). Cables and workstations installed with softphones help facilitate the process of communication. The IP-enabled phone is an example of end-user equipment. Multimedia computers and workstations connected with internet sources are examples of other end-user equipment.

The common component of networks includes cables, switches, firewalls, wireless, and other IP-enabled phones connected to a network source. Numerous available institutions and IP networks are enabling the installation of systems of VoIP. Due to the influence of IP data networks being higher than data networks, a network needs to be strong enough to support VoIP networks. Example component gateway is defined as an optional component system of VoIP helpful in call routing, handling of call connections, managing multipoint control units, and managing terminals (Taruk et al., 2018). Gateway works as an interface between VoIP networks and public switched telephone networks. The gateway comprises a signaling gateway which is helpful in the management of signal traffic media signals is managed by the signal gateway. The multipoint control unit establishes conference calls between two or more people for groups' video, voice calls, or even group data conferencing. It provides a combination of audio, data conference, and video conference depending on the network strengths of the end-users taking part in such kinds of communication. On the other hand, call processors define software running on the operating systems of the end-users helpful in the setup and operation of such call setups. It also helps in user authorization, bandwidth control, and signal monitoring.

The above architecture describes a working prototype of bidirectional communication wherein the SIP signaling and establishment of peer-to-peer SIP sessions is handled by Kamailio and the RTP sessions are maintained by RTPengine. Media channel creation, bridge creation, is provided by Asterisk for the flow of RTP packets. The browser based client side RTC application is created by WebRtc, which interacts with Kamailio securely over Transport Layer Secured (TLS), and capable of media negotiation, NAT traversal and encoding/decoding of SIP messages.

The incoming SIP packets from WebRtc client are transported over TCP or UDP through a firewall, exposed to the outside world to filter out the SIP packets. These SIP packets consist of a header containing SIP URI and other parameters. It may contain a SIP body describing RTP stream address and media negotiation schemes, which is defined by Session Description Protocol (SDP). The filtered packets are forwarded to Kamailio using Port Forwarding technique at firewall layer, where it performs initial checks including verification of domain, authorized certificates and authentication of clients credentials. On successful completion of the above processes, Kamailio registers the sip clients on Asterisk server. The Asterisk Server stores the client information in its cache and monitors their status by sending keep-alive. The keep-alive mechanism plays a crucial role in maintaining NAT entry at the firewall layer for the clients sitting behind NAT. Asterisk shares this information to our application by sending live events over HTTP, provided by its built-in application called Stasis. This establishes a permanent connection to our application through HTTP port - which is capable of sending and receiving both synchronous and asynchronous events from Asterisk. The REST library provides capabilities to connect to Asterisk via HTTP client and enables our application to send and receive various call handling events based on REST API models. This application also utilizes Hibernate - an Object Relational Model (ORM) to access and query MySQL databases for call detail records. Furthermore Redis, an in memory caching mechanism is used to store ongoing calls data such as call-id, call originator (sender) and receivers channel ids, call status etc. to handle events and delegate events to WebRTC clients. It stores objects (key-value pairs) in cache memory; this property of Redis is leveraged to access concurrent call data providing a faster way to respond to the events generated on concurrent calls. To prevent over flooding of cache, after the SIP session is destroyed or the call is ended the corresponding object is removed from cache. The application is mainly responsible for SIP channel creation from Asterisk to the other endpoint, followed by bridge creation and finally, addition of both party's channels to the same bridge, enabling the flow of media between the two endpoints. It is also capable of maintaining users status, handling various call events such as call hang-up, call recording, call transfer etc. The SIP signals generated by the application traverse back to Kamailio through asterisk, where they are further processed and routed to the appropriate WebRtc clients. This process enables the application to act as a SIP UAS (User Agent Server), which is capable of sending and receiving signals to multiple concurrent SIP sessions. Furthermore, this application is capable of controlling the SIP sessions and handling events generated by clients based on use cases. The architecture can be further scaled up according to the business needs by load balancing Asterisk servers and containerizing our application to deploy multiple instances.

The other part of the architecture leverages SIP-trunk facility, which enables the PBX endpoints to send and receive voice calls via Internet Telephony Service Providers (ITSP). This replaces the traditional telephone lines, PRI (primary rate interface) or DAHDI (Digium Asterisk Hardware Device Interface) that carries the voice calls over the PSTN and Public Land Mobile Network (PLMN) supporting cellular technologies like GSM/2G, 3G/4G/LTE. On the other hand, SIP trunks are virtual phone lines that utilize a packet switch network, where voice calls are broken down into digital packets and routed across networks to its final destination. Each SIP trunk provides multiple numbers of SIP channels; each channel is capable of handling a single SIP session at a time. These channels can be considered as call lines utilizing exactly one channel to establish a call. Since a SIP trunk can hold an unlimited number of channels, depending on the number of concurrent calls business tends to make, channels can be further increased. In this project, Asterisk is configured as a PBX system that implements SIP trunk services and features. The driving factors behind implementing a PBX system to offer businesses the ability to choose local Direct Inward Dialing (DID) numbers, toll free numbers for their business headquarters and other offices. This also allows customers to have customizable inbound and outbound caller id functionality to specify the users or business. In addition to this, it supports advanced voicemail features and Interactive Voice Response (IVR) facilities. This implementation can be easily scaled to match the needs, without any additional hardware or physical lines.

3.3 Call Flow

The following are the stages in the audio connection between gateways, gatekeepers, and terminals. The first stage is the registration of gatekeeper and terminal registration. Routing of calls between the gatekeeper and the terminal. Capacity exchange and initial communication are other components of gatekeeper protocols. Establishing audio communication and audio transmission are other stages in audio connection. The connection between clients follows the following steps to establish the two endpoints with gatekeepers in a communication platform. The A terminal begins registration admission through gatekeeper message exchange initiation. A terminal gets the opportunity to connect with B through information provided by the gatekeeper (Taruk et al., 2018). Terminal B obtains and sends-up messages from terminal A. under the request from the gatekeeper, the call preceding, and a message is sent by terminal B. Connect messages and alerts are sent by terminal B. Establishment of real-time protocol paths is also another novel idea in making communication across numerous communicating parties.

Below figure shows the INVITE message to originate a call between two user agents or clients. The header parameters are as follows

1. Call-Id: It is a unique identifier for each call and used to keep track of all the subsequent messages.
2. From/To: FROM field specifies the caller Id and TO field specifies the callee Id.
3. Contact: It contains a SIP URI which is used to define the contact information of the originator for any subsequent requests. For example, user@192.168.1.11
4. Via: This field keeps track of all the nodes/proxies a SIP message has crossed to reach the final destination.

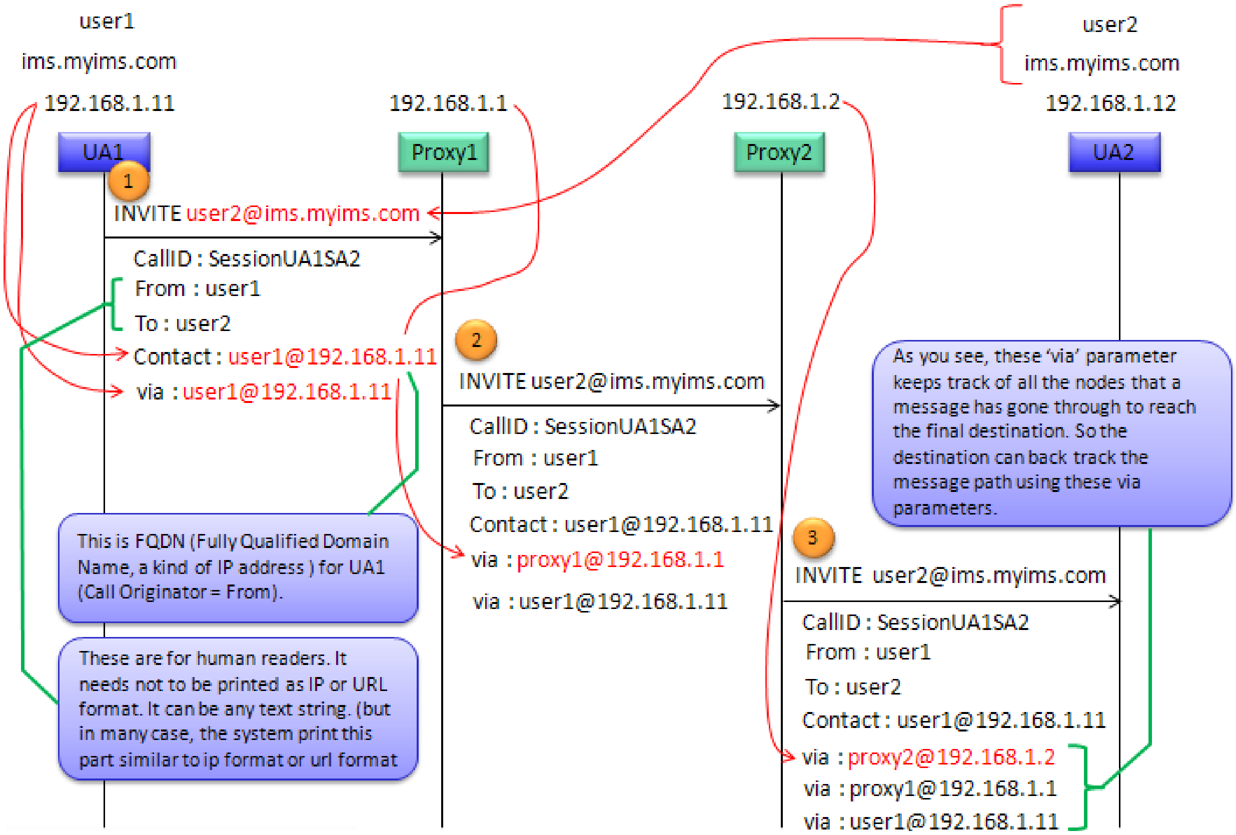

user1
ims.myims.com
192.168.1.11
192.168.1.1
192.168.1.2
user2
ims.myims.com
192.168.1.12
UA1
Proxy1
Proxy2
UA2
1
INVITE user2@ims.myims.com
CallID : SessionUA1SA2
From : user1
To : user2
Contact : user1@192.168.1.11
via : user1@192.168.1.11
2
INVITE user2@ims.myims.com
CallID : SessionUA1SA2
From : user1
To : user2
Contact : user1@192.168.1.11
via : proxy1@192.168.1.1
via : user1@192.168.1.11
3
INVITE user2@ims.myims.com
CallID : SessionUA1SA2
From : user1
To : user2
Contact : user1@192.168.1.11
via : proxy2@192.168.1.2
via : proxy1@192.168.1.1
via : user1@192.168.1.11
This is FQDN (Fully Qualified Domain Name, a kind of IP address) for UA1 (Call Originator = From).
These are for human readers. It needs not to be printed as IP or URL format. It can be any text string. (but in many case, the system print this part similar to ip format or url format
As you see, these 'via' parameter keeps track of all the nodes that a message has gone through to reach the final destination. So the destination can back track the message path using these via parameters.

Figure 2: Sip Message Flow (source: ***researchgate.net*****)**

3.4 Available VoIP Services in the Market

Based on the infrastructure employed by an owner of a network, we can establish different types of VoIP. The first type of VoIP is computer to computer. This service offers internet telephony through software like instant message and Skype. In this case, the caller and receiver should use their computer to achieve these motives. For these calls to be made a reality there are a few requirements like computer VoIP services, sound card, internet connection, and softphone software (Chen et al., 2018). When using VoIP services, the user cannot call either a mobile phone or landline. The receiver of such calls should have internet access to make these kinds of calls.

When using VoIP services, the user cannot call either a mobile phone or landline. The receiver of such calls should have internet access to make these kinds of calls. The other type of VoIP service is computer to phone and vice versa. These are hardware-based services and software-based services. Softphone hardware is applicable in routing the call to the protocol of the internet and also hands off to conventional telephone networks and all their components. One must subscribe and make payments for them to get the opportunity to use these services. Examples of these platforms include MSN, Google Talk, and Skype, which allow users to call landlines from their personal computers and communicate effectively (Taruk et al., 2018). VoIP subscription services, analog terminal adapters, and modems convert the call signal to digital and transform it into an analog signal of such networks. Phone to computer fails to allow emergency call users. They should replace such phones with computers connected to the internet to enable efficient communication between the two parties.

Phone to phone is another type of VoIP system. It is a hardware-based service permitting the receiver and the caller to call one another through the internet and social networks. Many telecommunication companies use the method to maintain long-distance calls between two communicating parties (Taruk et al., 2018). VoIP converts audio sound to data packets and passes all these pieces of information over an internet platform and a network.

3.5 Open Source VoIP Tools

Kamailio is an open source SIP server integrated that features most of the functionalities of VOIP services. Kamailio is a C based framework for UNIX/LINUX platforms to build real time communication platforms solving several common use cases such as authentication, authorization, registration, location services, accounting, NAT traversal. It is capable of interacting with other VOIP tools such as WebRTC, RTPengine, Asterisk and FreeSwitch etc., which play a crucial role in developing a robust architecture for our application.

Furthermore, it supports asynchronous communication over TCP, UDP and SCTP via TLS for voice and video communication. It offers load-balancing mechanism for PBX systems and media servers like Asterisk or FreeSwitch, handles SIP transaction failover mechanism, least cost routing and authentication against MySQL, Postgres, Oracle, LDAP. In addition to this, module extensions can be written in Lua, javaScript, Perl or Python.Kamailio has a modular architecture that provides powerful functionalities in real world deployments. The plethora of modules makes it easier for developers to add new functionalities in their VOIP services. The idea of implementing Kamailio in this project is to focus on performance, security and scalability. Since Kamailio has built in support to manage the number of worker processes per socket – separately for TCP and UDP, this provides scalability to our application and traffic management is made easier by segregating UDP and TCP workers. Kamailio also interacts with RTPengine to enable media streams to be proxies via an RTP proxy. RTPengine is responsible for directing RTP streams to correct endpoints, negotiation of media capabilities described in SDP, transcoding and setting other connection related parameters present in SDP such as IP of media server, codec related parameters etc. These proxies listen on a separate UDP port and manage a user-defined range of RTP ports that is selected randomly as the call session is established.

Therefore, a separate Linux box is dedicated to Kamailio and RTPengine to increase the robustness of the architecture and enhance the ability to handle multiple calls as per the need. The various other features of this reliable technology can be leveraged according to the project requirements in future.

Asterisk acts as a multimedia server with the ability to provide functionalities such as voicemail, services, call transfer, auto-attendants, transcoding etc. It utilizes and supports multiple databases to store users' information. Generally, MySQL server v5.6 is used as a database management system to maintain the information between multiple instances of Asterisk.

Below figure shows the top open source PBX tools. Specifically, Asterisk and Freeswitch are the leading the race in developing modern VoIP based communication systems.

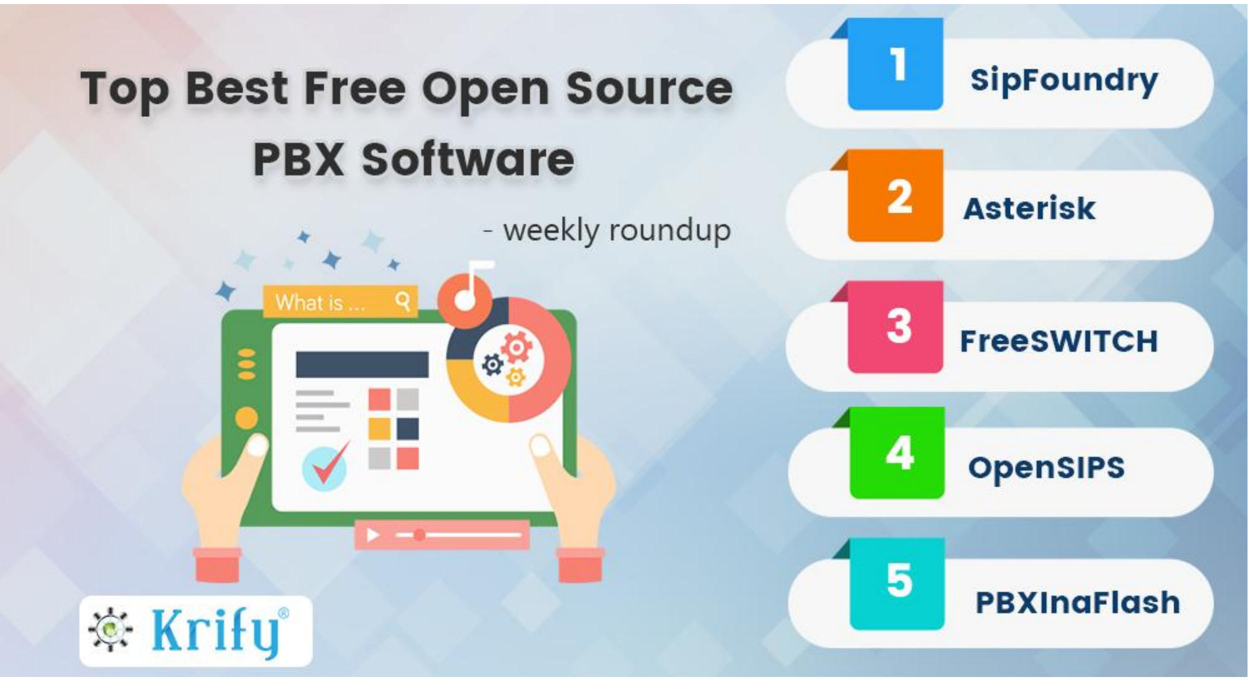


Figure 3: Open Source Tools for VoIP (source: *krify.co***)**

4

4. Session Initiation Protocol

4.1 SIP Messages

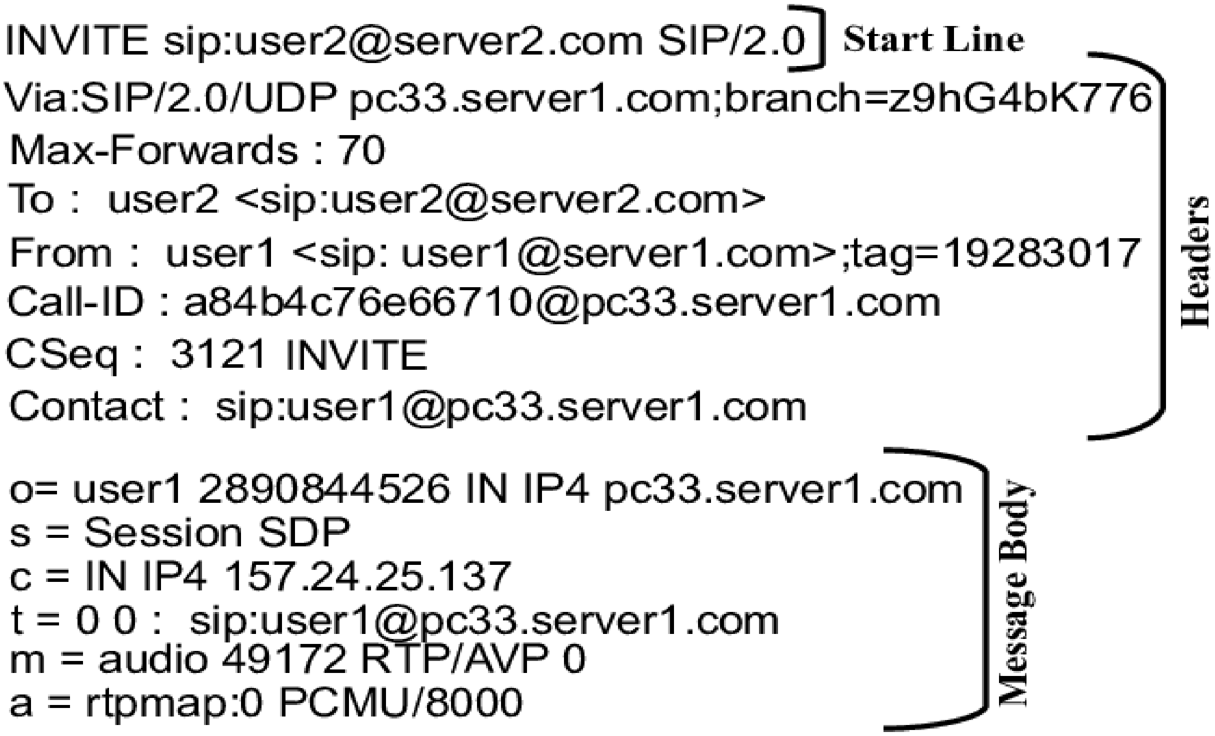
The internet engineering task force established SIP protocol. It terminates or initiates sessions of VoIP with two or more participants in a particular system. This peer-to-peer application protocol is initiating, creating, terminating, and modifying interactive media communication between users of a particular system. Since SIP is very flexible, it can be used for data packets, audio and audio communication, and data transmission. Hypertext transfer protocols and session initiation protocols are similar, and they are made of the clients' servers. The client requests the server where the request will be processed, and a response is immediately sent to the sender of the request.SIP is applied with applications like apple chart, MSN messenger, and instant messaging. Application of user description protocols helps in negotiations for codec identification, hence the support of user mobility through redirect server ad proxy used in the users' current location.

SIP comprises a network server and user agent. User-agent defines the endpoint acting on behalf of the user like server and clients. The client is also defined as the user agent and assists in initiating SIP requests, while the servers, who are also regarded as agents, receive the request, make meaning out of such requests and respond appropriately in the place of the user of such devices. The public network serves proxy servers, redirect servers, and registration services. The user's current location can be used in uploading current locations (Taruk et al., 2018).

The redirect server receives a request and determines the next step as the registration server uploads the current user location. The proxy server also has the task of returning the next hop server to clients instead of forwarding such requests without following the set due process.

There are six known steps for establishing communication in a protocol of SIP. The process commences with the user's registration, location, and initiation. Media determination becomes the second step, while the determination of willingness to accept or reject calls forms the third step of the call flow. Call setup, modification, and handling form the fourth and five stages of SIP call flow. Termination of the call becomes the last stage in the SIP call flow.

SIP enables numerous messages between a SIP server and the client. Some of the listed messages include INVITE, which initiates a call through a SIP user invitation into a session call. ACK is another listed message which is essential for a reliable and reasonable exchange of messages of invitations (Kaul & Jain, 2019). On the other hand, CANCEL is a listed message that enables users to cancel pending requests sent by clients to a server. BYE is another listed message which indicates the termination of a call. REGISTER is another listed message offering information concerning user location to SIP servers (Kaul & Jain, 2019). OPTION defines another listed message that helps obtain information about all the capabilities of a call. INFO forms parts of the last listed messages which indicate process out of bond information of great significance.



```
INVITE sip:user2@server2.com SIP/2.0
Via:SIP/2.0/UDP pc33.server1.com;branch=z9hG4bK776
Max-Forwards : 70
To :  user2 <sip:user2@server2.com>
From :  user1 <sip: user1@server1.com>;tag=19283017
Call-ID : a84b4c76e66710@pc33.server1.com
CSeq :  3121 INVITE
Contact :  sip:user1@pc33.server1.com

o= user1 2890844526 IN IP4 pc33.server1.com
s = Session SDP
c = IN IP4 157.24.25.137
t = 0 0 :  sip:user1@pc33.server1.com
m = audio 49172 RTP/AVP 0
a = rtpmap:0 PCMU/8000
```

Figure 4: SIP INVITE Message (source: *researchgate.net***)**

4.2 Comparison Between SIP and H323

H.323 and SIP offer similar services by exhibiting varied strengths and relying on their uses. Both H.323 and SIP perform duties like call set up, call waiting, transfer, and even holding off calls. They also perform the duties of Call Park, exchange of capabilities, and call identification. H.323 is majorly used in multimedia conferencing in support of video conferencing, data collaboration, white boarding, and support voicing (Shahadi et al., 2021). H323 is considered very complex and has a higher overhead, making it less efficient when attempting to use voice over internet protocols.

H.323	**SIP**
Complex protocols	Easy and simple to implement
Contextual representation use	Less modular
Very scalable	Less scalable
Less backward compatibility	Do not require a backward compatibility
IETE backed	A larger share of the market
Easy to detect loops	Difficult to detect loops
Has numerous elements	It contains only 37 elements

Table 1: Comparison of H.323 and SIP Protocols (source: Kyaw Myat, World Academy of Science, Feb 2009)

5

5. Challenges in Implementing VoIP

5.1 Point of Failures

There are numerous challenges that impact the implementation of voice over internet protocols in daily operations. Since the internet was not established to transfer real-time data, network equipment drops most of its packets resulting in qos that cause issues when transferring such information. Switching networks is less time consuming as compared to routing internet protocols. The use of reservation protocols is a possible way of handling these challenges.

The other issue identified in the voice over internet protocols is multiprotocol label switching which uses routing points to speed up packets. The Internet is not managed by a single centralized operator who has the possibility of coordinating and controlling the flow of calls; however, the internet is made up of varied services and networks; hence very difficult to achieve real-time data traffic (Kaul & Jain, 2019). This is another problem that has really challenged the use of the internet protocol. To achieve effective and efficient real-time communication, it is important to access internet bandwidth. The network comprises links and individual nodes generating a large amount of traffic when in use. If nodes send more packets higher than the network itself, congestion occurs, hence network latency and jitters. A network should be able to support the real-time transmission of video and voice for better service delivery and accurate data connection, and increased interaction on these networks.

When discussing the challenges of voice over network protocols, then we cannot afford to ignore power failures and failures in the back system of a firm. The past telephones operated on internal battery storage of about 48 volts and therefore is able to operate even if power goes off. They need to make use of an uninterruptible power supply as its backup systems. The system further needs a proper back up assessment done to ensure the amount of power needed for the operation of VoIP is always in place.

The implementation of VoIP is also challenged by the use of softphone firmwares. This software is prone to attacks by viruses. Spyware and malware are other vulnerability factors on these software platforms. Computers and other devices of users can be attacked by these threats in case they visit affected websites even if they fail to open such affected links. The past PSTN telephone has the ability to provide emergency calls since it is connected to emergency service providers and physical locations (D'Arienzo & Musto, 2021). These locations can be very helpful in tracking the locations of the users in case of anything. People who use VoIP networks can be available at all points across the globe hence making emergency calls very challenging with the ease of tracing the identity and location of the caller (D'Arienzo & Musto, 2021). There is even a lack of standard VoIP environment even if the VoIP providers are capable of offering emergency call services.

The current wireless devices currently are accompanied with algorithms for equivalent privacy security. WEP has been identified to be very weak and often very easy to attack and crack with even the free software that is available in online web pages. There is a great improvement with the introduction of VoIP with the current introduction of wifi alliances, also known as wifi-protected access (Nguyen et al., 2019). Gateway and VoIP compression algorithm is capable of being encrypted to offer improved security to handle the problems of network layer or network encryption and latency issues.

5.2 Algorithms for Compressing Voice Packets

Codec or the decompression and compressions are applicable in converting analog video and voice data to digital data that could stream along with these analog signals. It helps save network bandwidth by providing network compression. There are numerous different available video codec. They include

G.711. This codec was approved in 1965 and provided the easier and the simplest ways of converting an analog signal to digital form (Kaul & Jain, 2019). It uses pulse code modulation with less than 1% agreed on loss factor. This algorithm is the worst for bandwidth but the best inequality since it has 54kbit/s. The other algorithm is G.722, which got approvals in 1988 with a higher quality digital coding at 64bits with 7 kHz for audio spectrum (Kaul & Jain, 2019). It is mostly used for professional purposes like teleconferencing and application in IP mobile phones and personal digital computers.

The other algorithm is G. 722.1, designed by Picturetel as a wideband coder. T operates at 32 bits/s or the 24 bits/s. it has a workload of 20ms; hence it can encode frames of 20ms. This algorithm supports window messenger since it is G722.1. G.723.1 is the other algorithm that got its approval in 1995 for H.323 communications (D'Arienzo & Musto, 2021). It applies to video cell phones. This algorithm is difficult to use in modern transmission since it is not designed for musical use. It is recommended for use in 3G wireless multimedia and narrow-band video under the telecommunication and telephony union recommendations.

G. 726 is another algorithm that obtained its approval in 1990 and is mostly applied in adaptive differential modulation of codes. The encoding techniques of G.711 are done in words of 2,3 and 4 bits hence the 24,32 and 64 Kbits/s. G. 726 is another algorithm that is known to use lower delay. It is mostly used in modern and fax transmission (D'Arienzo & Musto, 2021). It is also used in H.323 conferencing of videos.

6

6. Quality of Service

6.1 VoIP QoS

Quality of service defines the ability to measure and control data transmission. This motive would assist in the elimination of errors and provision of predictable outcomes to chosen network traffics over the underlying innovative results. Video and voice packets do not tolerate packet loss, unlike the past packets of data with minimal delays in the delivery of packets. Therefore there is no need to handle issues to deal with quality of service (Nguyen et al., 2019). Reliable arrival of packages is important in conveying traffic over internet protocol networks. Therefore the purpose of achieving no jitters, latency, and improved characteristics one must therefore consider the quality of service since their impacts on the voice over internet protocols.

Delay defines the amount of time spent transmitting data packets from one source to another destination intended for the data to be sent. Communication delays should be controlled, and sound never goes beyond 15ms to ensure high quality of communication ('Voice over Internet Protocol Security,' 2005). Three major factors are the common causes of delays that are in the algorithm of equipment used in communication. Codec algorithms are another factor that leads to delays in communications. The three leading causes of delays include algorithm processing which defines the amount of time spent encoding a coded voice message ('Voice over Internet Protocol Security,' 2005).

The other reason for the delay is overhead, needed by a codec to critically examine the entire parts of a frame. The last reason for the delay is the frame delay which defines the time needed for sending systems to send one frame piece of information; it is also essential that the higher the system delay o is the higher level of compression.

Packet loss is another area of concern in this study. It results from hybrid circuits in which a shift from 4 wires to 2 wires takes place. This occurrence occurs in the event of a packet drop in a network, resulting in network loss and other critical pieces of information. The loss of a pocket significantly impacts the quality of service on a system of over-voice internet protocols. The level of Los that can be accepted on VoIP should be less than 1%, and anything beyond this value is not accepted in a system ('Voice over Internet Protocol Security,' 2005). Congestion and buffer size in a network is a significant cause of packet drops. All the congestions that can be seen in a network should be carefully analyzed and corrective actions done to ensure efficiency and effectiveness in communication.

Jitter is another issue that should be critically analyzed when studying voice over internet protocols. Jitters define the variation in arrival of inter packets, hence introducing a variable transmission to help address the issues of over the internet delays. Since the VoIP datagram protocols cannot guarantee the delivery of parcels in time, hence inconsistency in the arrival of packages ('Voice over Internet Protocol Security,' 2005). Jitter buffers can help do away with jitters. Jitter buffers also increase the number of delays in a system. A network should be able to provide the following for it to be considered very reliable and efficient in support of VoIP traffic and its reliable nature. A network should be able to provide the following for VoIP traffic reliability; a latency for packet forwarding, which is very tolerable for the conversation of VoIP (Nguyen et al., 2019). A VoIP session jitters that can accommodate packet forwarding. It should also guarantee bandwidth and VoIP session capacity for network congestion. To maintain higher quality, a network should provide low jitters and latency. To ensure high-quality VoIP services, then we should be able to control all these parameters.

6.2 Security in VoIP

IP applicable in transporting traffic over switched data packets is prone to attack by different hackers and vulnerabilities. A hacker may use the available hacking software to identify, modify playback and even steal voice traffic traveling through internet protocols. The hackers can trap the call, establish callers' identities and their locations and even establish the kind of information such speakers are communicating (Nguyen et al., 2019). Eavesdropping is another security breach of internet protocols. In this situation, a hacker can listen to the parties in a communication. Despite the end-user being deployed on networks that are internally protected, attacks are a common occurrence. The ending networks that are not covered are prone to attack by hackers who could compromise information security.

Software like the soft phone on the endpoint is prone to attacks in its operating systems and all the components that exist from such systems. The operating system needs anti-virus software and other virus detection software to ensure the security of calls through a network system. Software of PSTN is more prone to viral attacks. Therefore the use of IPSec can be helpful in the prevention of interferences with calls. It also protects the network by the use of active firewalls that may not be corrupted by the commonly available software commonly used by hackers.

Figure 5, states the security measures to keep a VoIP system protected from hackers on the internet specially UDP flood attack. Below are a few best practices to achieve a secured VoIP network.

1. **Securing Credentials:** It is critically important to have a strong password for VoIP phones and updating them periodically. Also, Two-factor authentication is a good practice to ensure the safety
2. **Preventing Ghost calls:** Hackers around the world scan for the open SIP ports and try to infiltrate into the systems causing UDP flood. To prevent this, a proxy server Kamailio can be used to block the incoming calls from untrusted IP sources.
3. **Data encryption:** Ensure end-to-end encryption of all calls processing data sent among servers

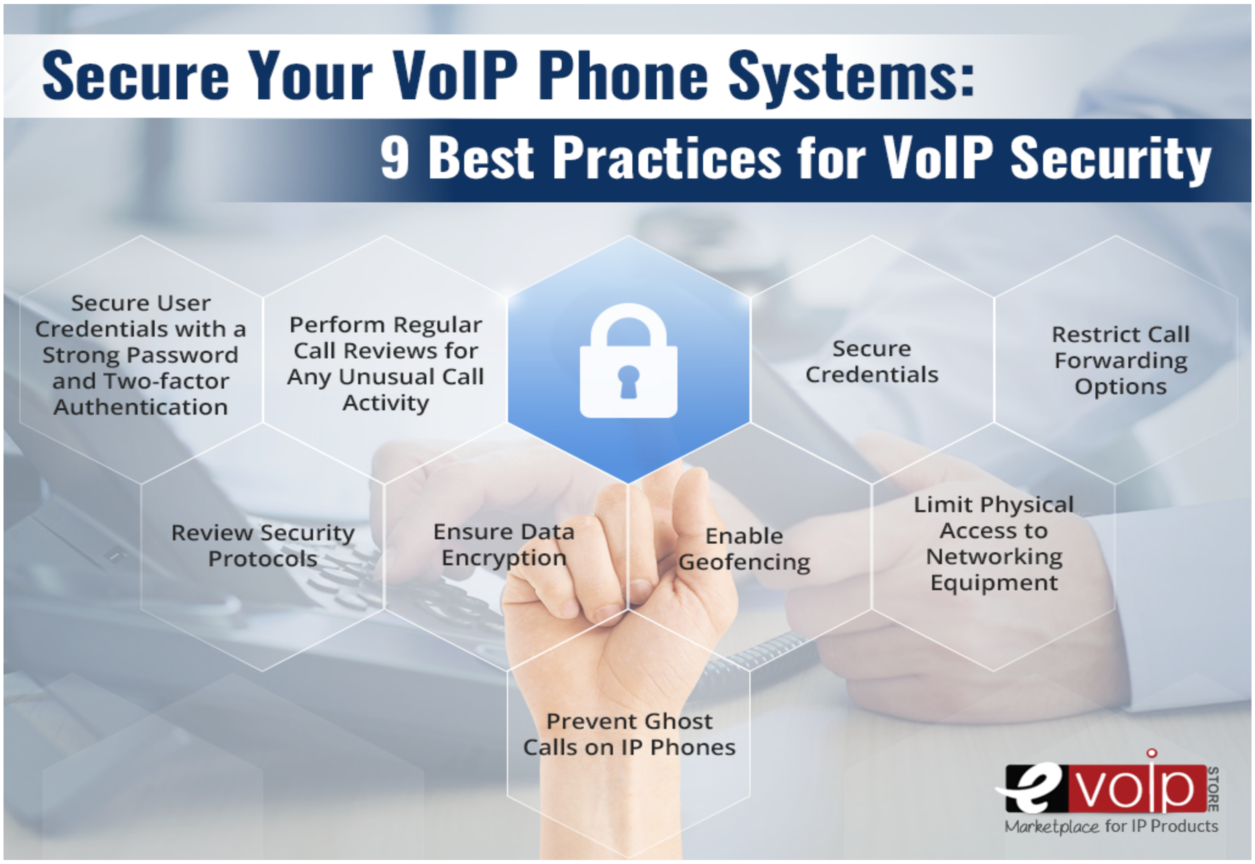

Secure Your VoIP Phone Systems:
9 Best Practices for VoIP Security
Secure User Credentials with a Strong Password and Two-factor Authentication
Perform Regular Call Reviews for Any Unusual Call Activity
Secure Credentials
Restrict Call Forwarding Options
Review Security Protocols
Ensure Data Encryption
Enable Geofencing
Limit Physical Access to Networking Equipment
Prevent Ghost Calls on IP Phones
evoip STORE
Marketplace for IP Products

Figure 5: VoIP Security Best Practices (source: *evoipstore.com***)**

4. **Geo-fencing:** Hackers may hack-in and use a VoIP phone to make international calls which could rack up the charges. Hence, it is a good idea to enable geo-fencing and allow only a few specific phones to make international calls.

In conclusion, these firewalls protect calls from any unauthorized users. It also guarantees call and endpoint protection from other unsecured networks prone to attack by hackers. It also ensures the proper billing by other payment service providers and protects such information from unauthorized users. It also protects caller behavior or additional caller information from unauthorized access on such sites. To ensure abreast with the changing technologies internationally, a more reliable and less expensive communication means is considered.

7

7. Future of VoIP Telephony

VoIP is very beneficial to companies' individuals and mostly educational centers. Some of the advantages of VoIP include saving on cost. These systems are often deployed on already existing networks, reducing the cost of purchasing network equipment and connecting such devices. These pieces of equipment are often very cheap and can be obtained from many vendors compared to the older equipment of PSTN (D'Arienzo & Musto, 2021). It also reduces the cost of making phone calls over longer distances. Application software like live window messenger and Skype permits personal computers to call other personal computers freely or considerably reduced costs. It also enables the making of calls between mobile phones and landlines at considerably reduced costs.

It also has the advantage of having advanced multimedia applications. Networks on VoIP enable the application of multimedia like whiteboarding and video conferencing as well as the transfer of files and hence the establishment of a secured network connection. VoIP network further integrates data and voice. All these inclusion helps in saving operational manpower and effective communication among a varied communication site (D'Arienzo & Musto, 2021). It also promotes the effective application of communication links between different sites or network platforms.

The other advantage is the global widespread use of internet protocols. In the event of the availability of numerous computers and mobile phones linked to the internet throughout the world, then the process of communication would be very fast and very effective. The presence of a gateway allows calls to apply VoIP and video calls from various people.

It also has the advantage of reducing the cost of ownership. VoIP comprises voice communication traffic and data integration into one network, reducing the cost of network maintenance and reducing infrastructural ownerships (D'Arienzo & Musto, 2021). This system brings together voice communication and data into one single network, reducing the cost of maintaining and acquiring infrastructure. It also integrates various network elements like the call server, client-server and application server.

VoIP is efficient in its role as a network resource. It efficiently uses network bandwidth by doing away with a silent conversation, increasing inefficient data throughput, and reducing repetition in human speeches. It also offers greater operational flexibility. Internet protocol network is made up of numerous varied layers of components that are separate and can be brought together to form a single system that operates independently (Kaul & Jain, 2019). This motive enables services, systems, and applications to be changed, hence extensible, customized, and flexible systems which are better for service delivery.

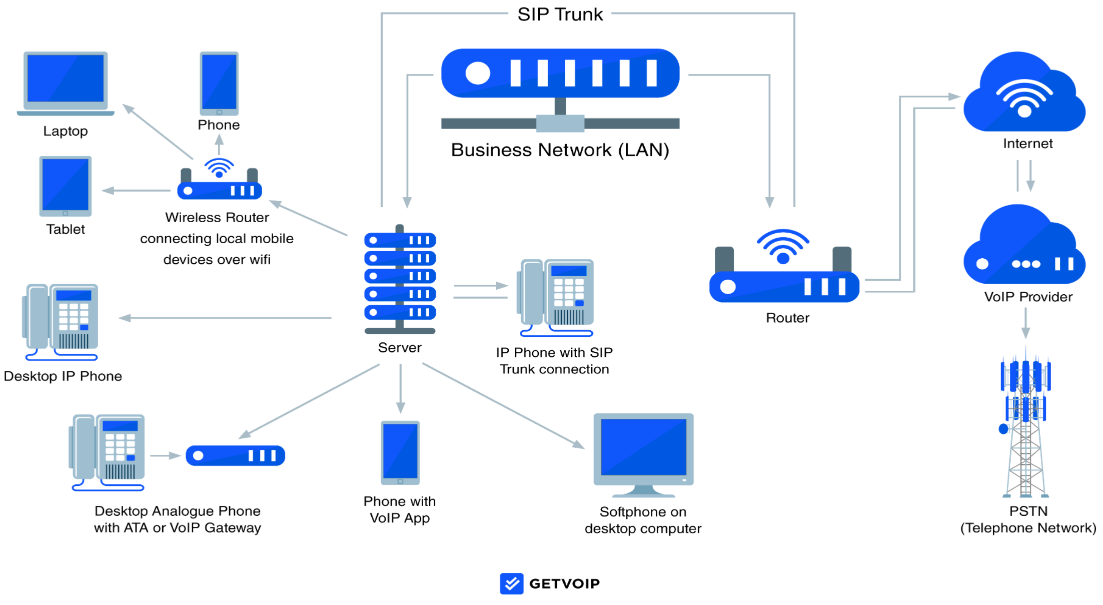


Figure 6: VoIP based communication ecosystem (source: *getvoip.com***)**

The above figure demonstrates a complete VoIP based ecosystem that can be used at offices, homes etc. Rather it can be implemented everywhere, where there is a reliable internet connection as it allows users to communicate via the internet. This can be easily scaled as per the needs to make local and international calls at affordable rates. Most importantly there are numerous features which could be added for example call queuing, as compared to traditional phones with limited features.

The proposed solution leverages the existing networks and different other standard protocols such as RTP and RTCP to develop a low cost VOIP solution. The existing proprietary solutions increase the implementation cost and cost of scaling up the service is huge in these cases. On the other hand, I have focused on the inclusion and coupling of multiple services to provide a lucrative and reliable VOIP service. For future work, a cluster-based solution could be implemented to further increase the availability of the solution and avoid redundancy. Moreover, incorporating technologies such as LDAP to enhance the security mechanism that could provide a robust and reliable communication system.

Acronyms

3GPP	Third Generation Partnership Project
PLMN	Public Land Mobile Network
PSTN	Public Switched Telephone Network
QOS	Quality of Service
RTC	Real Time Communication
RTP	Real Time Protocol
SIP	Session Initiation Protocol
UAS	User Agent Server
VoIP	Voice Over Internet Protocol

References

Pavel Segeč, Peter Palúch, Jozef Papán, Milan Kubina. “The Integration of WebRTC and SIP: a way of enhancing real-time, interactive multimedia communication”. *ICETA 2014, 12th IEEE International Conference. December 2014*

Ahmadreza Montazerolghaem, *Student Member, IEEE*, Mohammad Hossein Yaghmaee Moghaddam, *Senior Member, IEEE*, and Alberto Leon-Garcia, *Fellow, IEEE. “*Open-SIP : Towards Software Defined SIP Networking”. *IEEE March 2018*

Sarwar Khan and Nouman Sadiq. “Design and Configuration of VoIP based PBX using Asterisk Server and OPNET Platform”. *5th International Electrical Engineering Congress, March 2017.*

“Voice over Internet Protocol (Voip) Security.” 2005.

Chakraborty, T., Misra, I. S., & Prasad, R. (2019). Overview of VoIP technology. In *VoIP Technology: Applications and Challenges* (pp. 1-24). Springer, Cham.

Chen, C. M., Xiang, B., Wang, K. H., Yeh, K. H., & Wu, T. Y. (2018). A robust mutual authentication with a key agreement scheme for session initiation protocol. *Applied Sciences*, *8*(10), 1789.

D'Arienzo, M., & Musto, G. (2021). A comparative analysis of protocols for VoIP services. *International Journal of Communication Networks and Distributed Systems*, *26*(2), 159-175.

Choudhury, P., Kumar, K. R., Nandi, S., & Athithan, G. (2020). An empirical approach toward characterization of encrypted and unencrypted VoIP traffic. *Multimedia Tools and Applications*, *79*(1), 603-631.

Kaul, S., & Jain, A. (2019). Opus and session initiation protocol security in voice over IP (VOIP). *European Journal of Engineering and Technology Research*, *4*(12), 27-37.

Nguyen, H. T., Battle, A., Jain, N., Kesavan, S. P., Bhatia, H., Gamblin, T., ... & Bremer, P. T. (2019). Visualizing hierarchical performance profiles of parallel codes using call flow. *IEEE transactions on visualization and computer graphics*, *27*(4), 2455-2468.

Shahadi, H. I., Kod, M. S., Qasem, B., & Farhan, H. R. (2021, August). Real-Time Scheme for Covert Communication Based VoIP. In *Journal of Physics: Conference Series* (Vol. 1997, No. 1, p. 012020). IOP Publishing.

Taruk, M., Budiman, E., Rustam, M. R., Azis, H., & Setyadi, H. J. (2018, November). Quality of Service Voice over Internet Protocol in Mobile Instant Messaging. In *2018 2nd East Indonesia Conference on Computer and Information Technology (EIConCIT)* (pp. 285-288). IEEE.