



An Introduction to VoIP

An L1 Associates Technology Whitepaper

For more information on VoIP please visit www.L1Associates.com

Please direct any comments or questions about this whitepaper to enquiries@L1Associates.com

Background

Today's economic climate has made it more crucial than ever for businesses, whether operator or enterprise, to ensure that investment delivers value or can provide a competitive advantage.

Communication trends are shifting towards data, with data packets surging across the Internet, corporate intranets, and extranets, and the foundation for communications is indisputably becoming IP internetworking.

The emergence of voice video & data 'convergence' has caused business decision makers to rethink their current network strategies. In trying to reduce their operating costs, increase productivity and improve Return On Investment, businesses are cautious about investment timing and are faced with tough technology choices.

What is VoIP

Around 1995 there was a revolutionary advancement in the way traditional voice traffic was carried when voice, video and data were transmitted over a single 'converged' network. The terms Voice Over IP (VoIP) and IP Telephony (IPT) were introduced to describe this method of transport of voice signals as well as fax, DTMF tones, signalling traffic, and network management traffic.

At the heart of the VoIP model is network convergence, whereby voice & multimedia traffic are converted to 'packets' allowing them to be integrated with data traffic on a common shared IP transport architecture. This convergence reduces costly and complex network layers and with careful design can produce savings in bandwidth, equipment expenditure and equipment administration.

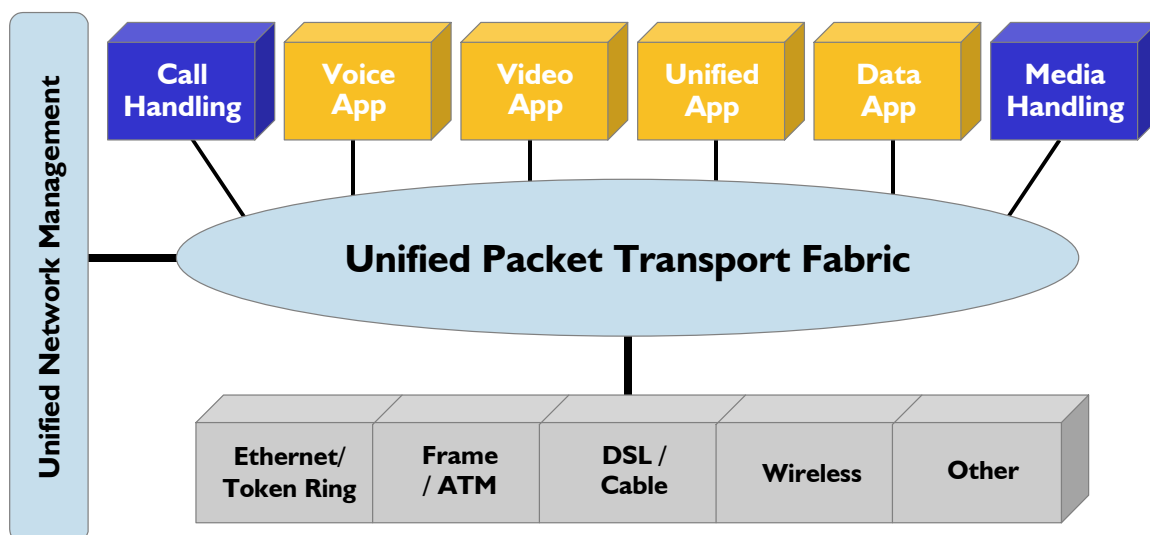


Figure 1: Converged Network Model

VoIP also allows network and application functions to be dispersed throughout network with bearer independent access to services. Complex call processing and applications can be centralised and low complexity media handling functions can be flexibly distributed throughout the network.

Benefits

In the current market climate IT & Telecoms investments are exposed to a greater degree of scrutiny than ever before and the business case justification must clearly identify the total life cost as well as the business benefits & business risks.

Many key drivers exist that speed up the introduction of VoIP solutions which when implemented can realise many benefits such as::

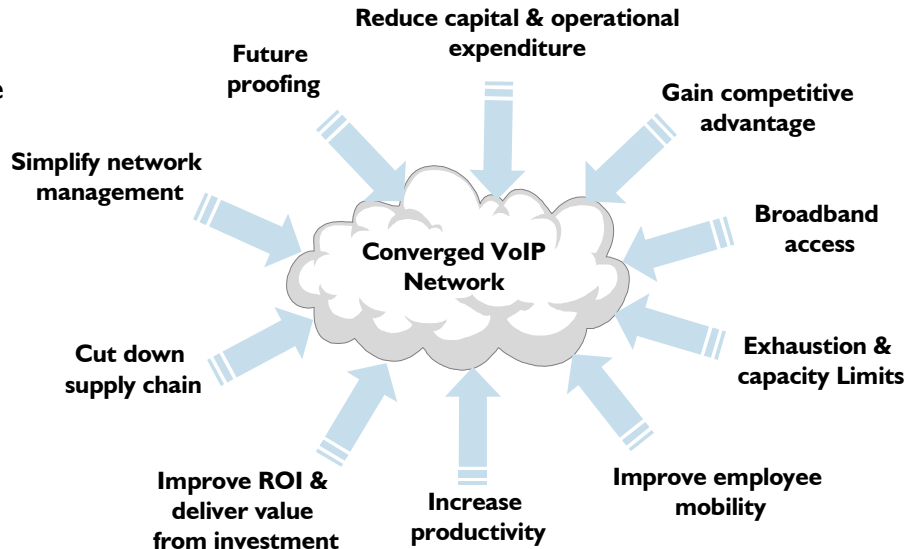


Figure 2: VoIP Drivers & Benefits

- Carrier & leased circuit charges can be reduced using national & international toll bypass, by running voice & video traffic over the LAN / WAN / Virtual Private Networks (VPN).
- Enhanced business communication brought about by IP PBXs, multimedia collaborative applications, unified messaging & videoconferencing can improve employee productivity.
- IT department productivity can be improved with centralised network administration, and the simplified moves, adds & changes that IP provides.
- Common building wiring for both voice and data traffic can provide for rapid office expansion and the introduction of new office locations.
- Remote offices can take advantage of centralised IP PBX architectures with distributed IP phones providing consistent user experience for all regardless of size or location.
- The 'IP anywhere' concept improves mobility for road-warriors and home-workers who gain access to the office services and benefit from enhanced unified messaging services.
- Capital expenditure is future proofed by investing in highly robust, scaleable, and resilient converged infrastructure.

Services & Applications

VoIP by its nature is primarily concerned with the provision of voice and telephony services on IP networks. Traditional telephony services on VoIP networks today include Long Distance Toll bypass, Voice Conferencing, Call Centres, and PBX networking.

In addition to traditional voice services VoIP provides an environment that allows voice to be integrated with other media types both at the transport layer and at the service layer.

New integrated applications are in their infancy and already new service opportunities such as powerful unified messaging, IP video conferencing & other IP video applications, IP Phone & PC Softphone features, multimedia contact centres, collaboration applications, and mobility enabled presence services, can all help to greatly enhance an organisation's communication capabilities.

The extent and pace at which more converged services will emerge is not yet totally clear. What is clear however is the quantum leap convergence has brought to voice service creation and data integration compared to the past history of ISDN and Intelligent Networking.

Architecture

VoIP equipment can be introduced gradually for either capacity replacement or expansion. Intersite voice traffic can be migrated off of leased lines or from PSTN routes onto the existing data network. IP PBXs can be deployed to control IP phones and softphones on remote sites. Remote workers with IP access are enabled for voice calling. The diagram below shows a simplified architecture evolution for the introduction of VoIP.

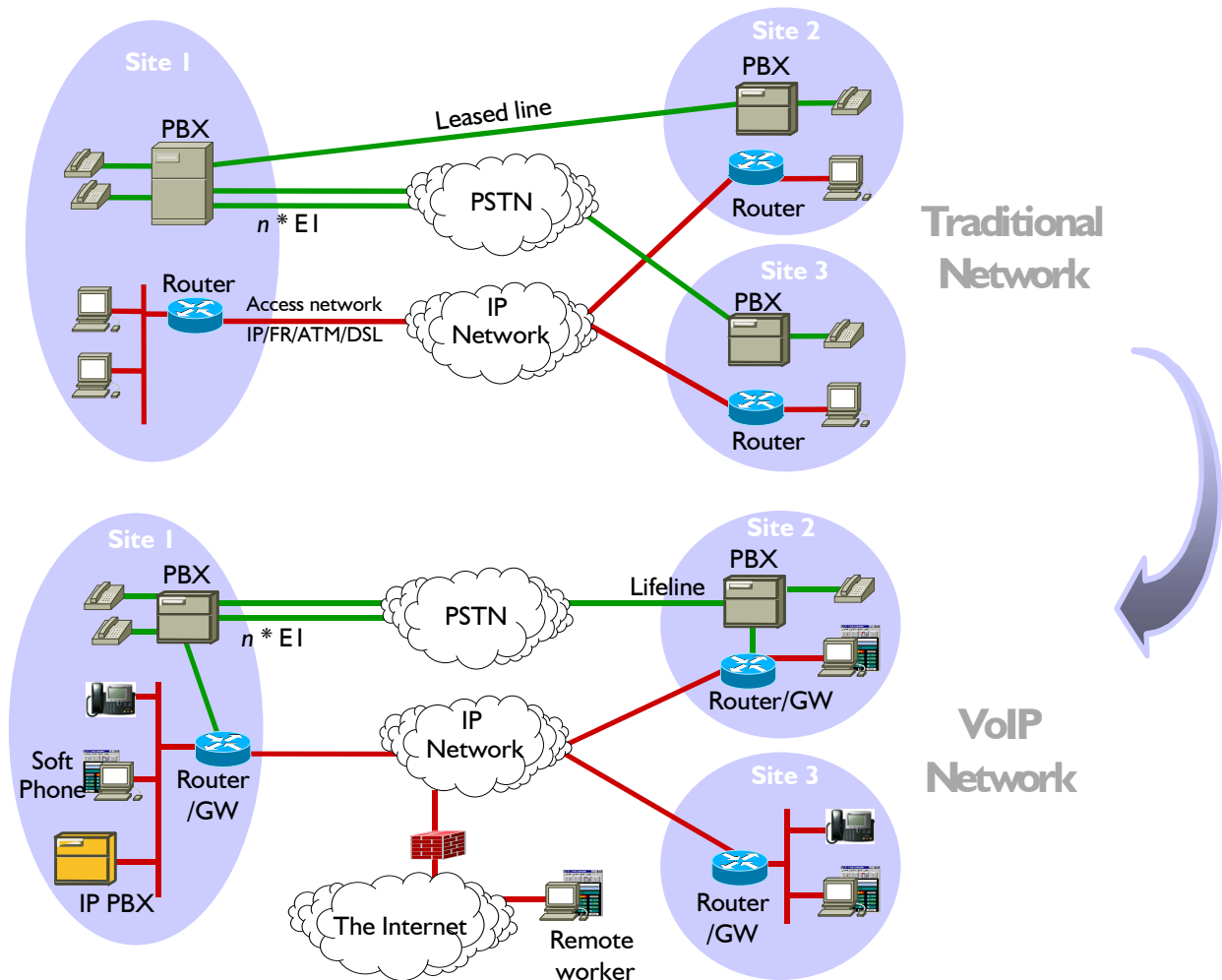


Figure 3: Architecture Evolution

Key Network Components

A number of new specialised equipment types have emerged to support the transport and control of voice communication on IP networks.

Call Agent (CA): Examples of CAs include Gatekeeper (GK), Media Gateway Controller (MGC), SIP Server, or Softswitch (SS). CAs typically perform

- Translation between ITU E.164 telephone numbers and IP network addresses
- Receiving & generating signalling messages for call routing to the appropriate destination
- Admission control and subscriber screening to validate access to the network or service
- Control access to network resources such as conference bridges
- Provide bandwidth management for MGs and end devices
- Call Record generation for performance and accounting

Media Gateway (MG): a device under the control of a Call Agent that converts the media (e.g. voice, video, fax) between circuit switched networks and IP networks.

- Encode and decode digitised media signals using a codec (e.g. G.711, G.729, etc.)
- Performs packetisation and depacketisation to and from IP
- Controls any echo in the signal
- Plays announcements & tones
- The egress MG buffers packets to account for jitter (i.e. delay variation)
- In some cases terminates the voice signalling channel (e.g. PRI, QSIG, SS7) although this can also be performed by a dedicated Signalling GW or a Softswitch Call Agent.

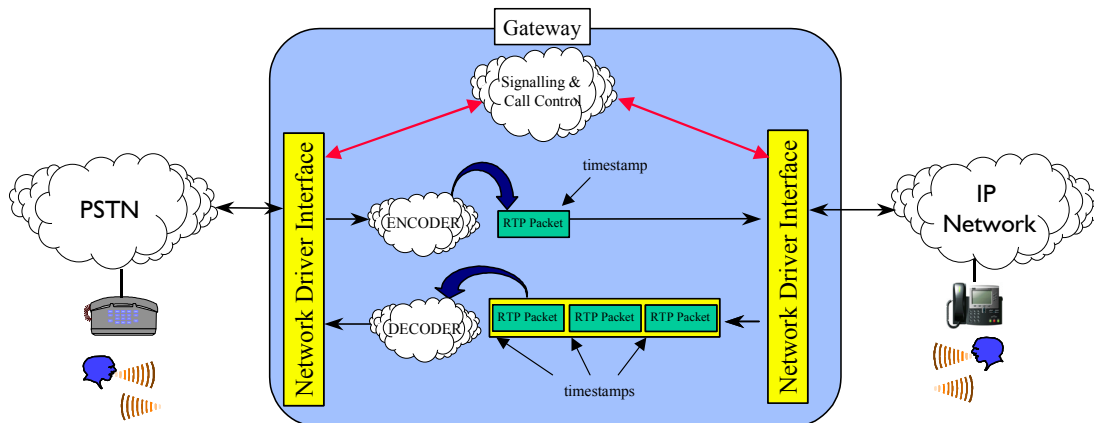


Figure 4: MG Operation

IP PBX: Either traditional PBXs with a VoIP gateway function or software based new generation server platforms. These can provide CA, MG or media resource (e.g. announcements) functions.

IP Phones: Similar to a POTS (Plain Old Telephony System) phones or business sets but contain MG function and have an Ethernet connection.

PC Softphone: such as Microsoft Messenger and Netmeeting that provide a software voice calling function usually from a PC.

Other Devices: include media servers such as audio servers or IVRs (Interactive Voice Response), MCU (Multipoint Control Unit) for conferencing, Presence servers, databases for authentication (e.g. RADIUS, DIAMETER) and directory servers (e.g. LDAP). Other traditional data equipment such as DNS, DHCP, Firewalls etc. may also be used.

Media Coding & Transport

The MG converts analogue or TDM (Time Division Multiplex) digital voice traffic to / from a format suitable for transport across IP networks. A sample of the digital signal is taken, typically 5 to 30msec in size, and encoded using a codec. The codec can compress in order to maximise available bandwidth, however compression can reduce the quality of the original signal.

Audio Codecs: common examples include G.711 (PCM 56/64kbit/s i.e. no compression), G.723.1 (5.3 and 6.4kbit/s), G.726 (16, 24, 32 & 40kbit/s), G.728 (16kbit/s), G.729 (8/13kbit/s), GSM AMR (4.75 to 12.2kbit/s), MPEG 4 Audio (MP3 and AAC).

Video Codecs: common examples include H.261 (for 64kbps and above), H.263 (for 64kbps and below), and MPEG 4.

The encoded information is then encapsulated within an IP packet and these packets are then transported across the network to their destination.

RTP/RTCP (Real Time Protocol / Real Time Control Protocol) is used as the IP transport protocol as specified in IETF RFC 1889 & 1890, and is typically run on UDP. RTP provides time reconstruction (using time stamps), loss detection (using sequence numbers), delivery monitoring, and content identification for the application to implement reliability and flow/congestion control. Flow and congestion control information of RTP is provided by RTCP sender & receiver reports.

Upon receipt at the destination the packets are ordered into the correct sequence. The MG uses a buffer to temporarily store all received packets and to reorder late or out of sequence packets.

If the voice is to be converted back to switched circuit voice then the voice information is extracted from the packet and a decoder is applied. Where a packet is lost and does not arrive then either a 'blank' packet is inserted causing distortion to the original voice signal when decoded, or a concealment technique can be applied that 'estimates' the content of packet which can minimise this distortion.

Modem, fax and DTMF (Dual Tone Multi Frequency) tones can be subject to distortion, delay, and transport failure when sent across packet networks. When no compression and low loss networks are used then tones can be transported inband. Otherwise the MG monitors TDM trunk & IP interface for these tones and upon detection treats the call appropriately (e.g. disables codecs & echo control, fixes jitter buffers etc.).

Another initiative for fax support is the ITU T.38 recommendation which defines an IP network protocol used by T.38 fax devices which demodulate / modulate and translate T.30 fax signals into Internet fax packets (IFP).

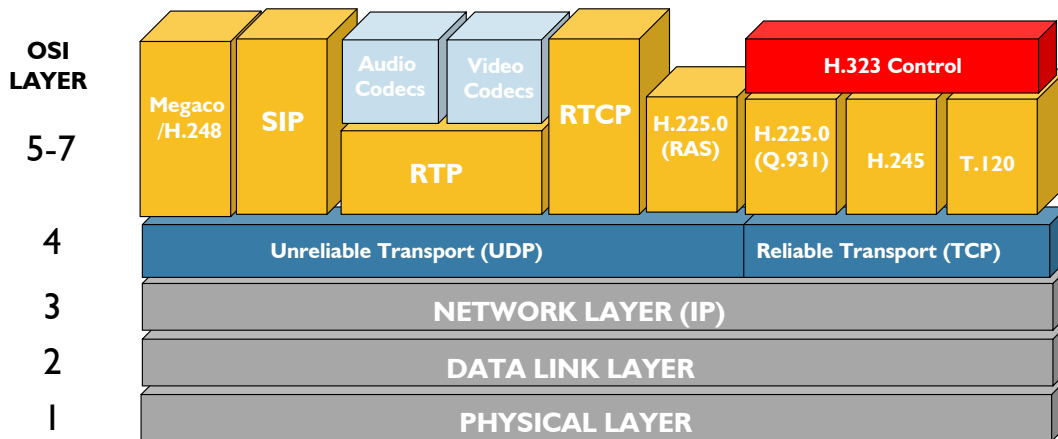
Industry Standards

Around 1996 industry standards bodies started to ratify early protocol standards that specified how these various entities should communicate and interact with each other. Today there are many protocols that exist within the IPT arena not just for voice but for also for video telephony. The following are some of the key protocols and standards bodies.

H.323 Developed by Study Group 16 of the ITU-T, H.323 is an umbrella standard encompassing many subcomponent standards and annexes for transmitting multimedia (voice, video and data) across packet based networks. The two main subcomponent protocols are H.225 (call control), & H.245 (bearer control and capabilities exchange). H.225 consists of two main parts Q.931 (basic call control as used in ISDN networks) and RAS (Registration, Admission & Status). H.323 also references other standards such as Real Time Protocol (RTP), G.xxx audio codecs, H.26x video codecs and T.120 real time data conferencing protocol.

SIP Developed by IETF, SIP is a mechanism to initiate, terminate & modify sessions in an IP network. It uses a client / server architecture and the protocol is request-response based. It enables personal mobility by tracking down users and delivering calls to an endpoint. It is a lightweight, text-based protocol and reuses much of the construct of other internet protocols such as HTTP and SMTP. SIP does not know about the underlying details of a session and relies on IETF protocol Session Description Protocol (SDP) to describe the session. It also interworks with other IETF protocols such as Megacop, RTP, RTSP, RSVP and SAP.

MGCP & Megacop/H.248 MGCP (Media Gateway Control Protocol) & Megacop/H.248 are relatively low level, master/slave protocols used between Call Agents and MGs. MGCP is not officially a standard but does exist as an informational RFC (RFC 3435). Megacop (IETF RFC 3015) & H.248 (ITU Rec.) are the same protocol developed by an IETF and ITU collaboration. It is derived from, and draws heavily from MGCP but adds several new enhancements.

Protocol Stack**Figure 5: OSI Protocol Stack****Comparison of Key Standards**

H.323 is the most mature of the protocols described and gained the earliest momentum in the IP Telephony market. It is a well specified protocol originally designed for the requirements of multimedia communication (voice, video and data conferencing) over IP.

H.323 takes advantage of traditional circuit switched voice protocols, for example the call establishment protocol H.225 is based on ISDN Q.931 and the H.450 series of standards define the support of ISDN like supplementary services.

SIP based solutions have been around for many years but recent enhancements and an increasing adoption by the marketplace have seen a more widespread take up and it is forecast that SIP shall ultimately replace H.323 as the peer protocol of choice for most IP Telephony applications.

SIP was designed originally to set up and terminate calls (i.e. a session with media streams) between two parties. Its model has many close similarities with the Internet in that the protocol reuses much of HTTP & SMTP, and the addressing scheme is URL based.

Media gateway control protocols such as MGCP and Megacop/H.248 evolved to satisfy deficiencies in the gateway models allowing MGs to be more efficiently used in larger scale networks and their current development activity is much less than that of H.323 and SIP.

Whilst H.323 and SIP contrast substantially as peer protocols, there are fewer differences between MGCP and Megacop/H.248. These similarities stem from the fact that they perform the same architecture function of MG control, and that Megacop/H.248 has its roots in MGCP.

Design Challenges

IP networks were not originally designed to support services such as voice and as such certain design challenges have arisen.

Voice Quality

The quality of voice must in many cases be equivalent or better to that achieved on the PSTN. Voice quality is affected by delay & delay variation, packet loss and choice of QoS techniques as well as by the codec choice and the presence of echo.

Voice quality is best measured subjectively using a Mean Opinion Score method although objective measurements can also be carried out.

IP Quality of Service

Standard IP networks provide 'best effort' data delivery services by default. Time critical VoIP services requires a significant change to the way IP networks are designed. The IP network design for the model shall have the following goals:

- An end to end focus - an understanding that different domains (LAN, Access & Core) have different characteristics and solutions.
- A network that efficiently utilises the bandwidth available - bandwidth is not free and will always have a cost associated with it.
- A fully resilient network - if toll quality voice services are to be delivered, the network will require full redundancy and availability.
- Low end-to-end delay, low delay variation and low packet loss - ensuring these characteristics are met will ensure voice quality.

Delay: End to end or 'Total Transmission Delay' is the sum of the compression, decompression, processing, buffering, queuing, transmission and the network delays. When this total delay exceeds a set amount (e.g. 150ms), the speakers experience problems on interactive dialogue.

Delay Variation: or 'Delay Jitter' is the variability in arrival time of a packet and when a packet does not arrive in time it may have to be discarded. It cannot be re transmitted, as it would delay proceedings too much. A jitter buffer in the egress MG exists to provide buffering of packets allowing time for late packets to arrive.

Packet Loss: In some cases packets may not reach their destination in IP networks, although the IP network can be well engineered it is difficult to totally eliminate this problem particularly in networks running multiple services (voice, video & data). Lost packets can also degrade the quality of the voice since parts of the original signal are lost. Lost packets can sometimes be 'concealed' in the MG by transmitting 'estimated' packets in their place. High packet loss (e.g. 1- 5%) can also adversely affect fax services.

Traditional data services are normally unaffected by characteristics such as delay, delay variation and packet loss and they are not controlled by default. QoS mechanisms have to be employed to assure voice (& real time video services where implemented) will be delivered as a priority regardless of the current network conditions. There are two main methods to address these issues:

- Over provisioning - over provisioning the bandwidth in the network to guarantee delivery.
- QoS mechanisms - the deployment of quality of service mechanism(s) within the IP network to guarantee delivery.

Purely over provisioning networks does not adequately address the voice quality issue on their own, for example TCP by nature is a bandwidth 'hungry' protocol and will attempt to utilise all available bandwidth at a given moment. Over provisioning also requires very accurate network information and the ability to dynamically assign bandwidth instantly.

Multiple QoS mechanisms are available today to address the demands of today's applications. The choice depends on factors such as the application requirements, network element functionality and domain. For instance, certain protocols are LAN specific, some focus on different characterises and some guarantee delivery (at the detriment to other possible business critical traffic). QoS protocols can be categorised mainly as:

- Resource reservation - network resources are apportioned according to an application's QoS request, and subject to bandwidth management policy. Some examples of resource reservation include RSVP and MPLS.
- Prioritisation - network traffic is classified and apportioned network resources according to bandwidth management policy criteria. To enable QoS, network elements give preferential treatment to classifications identified as having more demanding requirements. Some examples of prioritisation include Diffserve, LLQ (low latency queuing), CBWFQ (class-based weighted-fair queuing), 802.1q.

Critical to the success of the network is the correct choice and correct configuration of a QoS mechanism. Some QoS techniques are designed for less demanding applications than high quality real time voice so it is essential that each technique is examined in relation to the technical requirements of the network and for each application. In additional voice services require guaranteed delivery of call control and signalling packets as top priority (usually in order that network traffic can be removed gracefully in times of overload or operational difficulty).

Choice of codec

Codecs can degrade the original signal by compressing voice and can also introduce delay. It is recommended that G.711 PCM be used in order to simply fax & DTMF tone support and provide PSTN equivalent toll quality voice.

Packetising voice introduces packet overheads and increases the size of the PCM 64kbit/s voice signal. Although high capacity broadband core and access networks can help, these overheads can still be cost prohibitive. If this is the case compression based codecs such as G.729 or G.723.1 and methods for transporting fax (e.g. T.38) and DTMF must be employed.

Header compression techniques can sometimes be used to keep the bandwidth down although these are not recommended where flexibility in packet routing is not available and where higher resilience under failure conditions is required.

Voice Activity Detection

Voice activation detection (VAD) also known as 'silence suppression' takes advantage of the fact that most conversations include about 50% silence (e.g. one party is not speaking at any one time). VAD is a software application within the MG that monitors signals for voice activity so that when silence is detected for a specified amount of time, packet transmission is halted avoiding the transmission of "silent packets" over the network. Although VAD conserves bandwidth it does affect the user experience and perception of call quality.

Echo cancellation

When a two-wire telephone cable connects to a four-wire interface, a special electric circuit called a hybrid is used to do the conversion. But in doing the conversion a small percentage of telephone energy is reflected back to the caller creating an echo. If the Total Transmission Delay is high then the caller hears the echo and this has to be removed by the egress MG.

Firewalls & NAT

Common functions in any IP network are Firewalls and NAT. The configuration of devices containing these functions can be problematic for VoIP since VoIP protocols can 'hide' transport addresses within messages. Additionally the use of encryption, VPNs and secure tunnelling can prevent multimedia sessions from traversing 'simple' FWs or NATs. Solutions to this problem are very specific to the individual application and include use of devices such as VoIP protocol or VPN aware Firewalls, Application Servers, Media Proxies, and Application Server Agents.

Addressing / Directories

Telephone numbers & IP addresses have to be managed in a way that it is transparent to the user. Some devices (e.g. PCs) may not have telephone numbers and directories have to include mappings between E.164 numbers, IP addresses, email addresses, URLs, & SIP addresses.

Network & Service Migration

Migration must address service characteristics from a user's point of view as they are likely to experience a change to normal service. Not all traditional telephony features map to VoIP networks and some may have a different user look and feel to that of traditional networks.

Tight control of operational, customer and network management is crucial. A migration plan must account for geographical & environmental issues, user issues, traffic routing and dial plans, particularly as there may be some parallel running and requirement to 'fallback' at short notice.

Conclusions

VoIP has established itself within the voice & data industries. Standards have matured to the point where robust, scalable and reliable products can be readily integrated within existing networks.

VoIP brings together the best of the internet and voice worlds creating an environment that facilitates high value, cutting edge, mobility enabled, converged communications services.

There is still however a significant amount of activity within standards organisations to improve and develop existing protocols particularly in the SIP arena. This activity is considered a refreshing change to the stagnation that existed with the development of voice services prior to VoIP.

Only with a comprehensive knowledge of convergent technology and real implementation experience can a company generate accurate business cases and provide the highest probability of success for implementation.

Careful design and integration can reduce the risk and delay with deploying IPT and VoIP solutions particularly as every user of the technology will have their own requirements with respect to services, legacy infrastructure, and operational support arrangements.

Almost all businesses now agree that the benefits brought about by convergence are too good to ignore realising that it is not a question of 'if' but 'when' to invest in converged technology.

Some businesses are delaying IT and telecoms investment and are preparing IP infrastructures in anticipation for a gradual migration process.

Other more shrewd businesses recognise the importance of competitive advantage and that the benefits of deploying converged solutions today can open up exciting new opportunities, and deliver true value from their investments.

What is clear is that we are entering a new era in the evolution of voice communications.
