Published in Applied Computing, Summer 2000

A Revolution in Voice Networks – VoIP

By Jon Kenton – Telecom Market Development Manager, Motorola Computer Group
& Philippe Chevallier – System Architect, Motorola Computer Group

Introduction

For many years the networks that have been in use to transmit our telephone conversations have been based on the same basic technology, switched circuits. These networks have evolved many times including one major evolution from analogue to digital circuits. We are currently however at the onset of a major Revolution that is sweeping across the telecom industry and bringing in many new players who hope to take advantage of a movement to transmit voice calls over the internet. We are all now involved one way or the other in the "packet revolution". It is called many things, Voice over Packet, Voice on the Net, Internet Telephony or Voice over IP (Internet Protocol) or **VoIP**.

This revolution, brought about by the explosive growth of the Internet, is driving a major shift in the spending of the infrastructure dollar towards high-speed packet technologies. It has also made it apparent that the days of shoehorning data onto the voice network are ending and that it is time to start adapting voice traffic to the Internet.

History

Since its inception the telecommunications network has been based upon circuit switching and for good reason. Until the 1970s the only traffic on the network was voice telephony and the transmission of voice telephony has essentially been based on creating a physical path (or circuit) between two handsets by closing a series of relays or switches. From the days when an operator manually patched you through their mechanical switchboard, to the present when your phone call is automatically switched through enormously expensive high-speed digital switches, the phone system infrastructure was designed for a single function, circuit switched voice. In the 1960s however, with the advent of commercial computers, these same phone switches began to be used for data communications albeit awkwardly. The only solution was the analog modem. Over the years the analog modem overlaid data onto the phone network at ever-increasing speeds but has always been limited by the basic issue of bandwidth. This is primarily because the phone network has been limited to audio frequencies from 300Hz to 3400Hz and is also very "noisy". Although people can make out a voice conversation through certain levels of static a modem has a more difficult job.

In the early 1980s the telephone networks evolved into a purely digital system. From Central Office (CO) to CO the analog voice signal was encoded as a digital stream and then switched digitally until the stream reached its destination CO switch where it was converted back to analog. The connection from the CO to the home (the local loop) was still basically an analogue two-wire connection. This worked reasonably well until increasing loads were put upon the switching infrastructure by dialup data from a new and rapidly growing phenomenon, the Internet.

The Internet

The Internets growth and popularity surprised everybody especially the capacity planners for the regional network operators. Nobody truly anticipated the explosive growth in bandwidth demand and the overall effect on the circuit switched infrastructure. While the Internet itself is based on packet switched architecture, private individuals and businesses would access it almost exclusively via modems and the established circuit switched local loop. The usage model developed by the telco capacity planners was based on the typical length of a voice call, which was in the order of minutes. With US local calls being free, and ISPs offering inexpensive “all you can eat” Internet service, these dialup connections were often “offhook” for hours. Some people even left them on 24 hours a day. This severely impacted the capacity for regular voice calls tying up expensive ports at the switching offices.

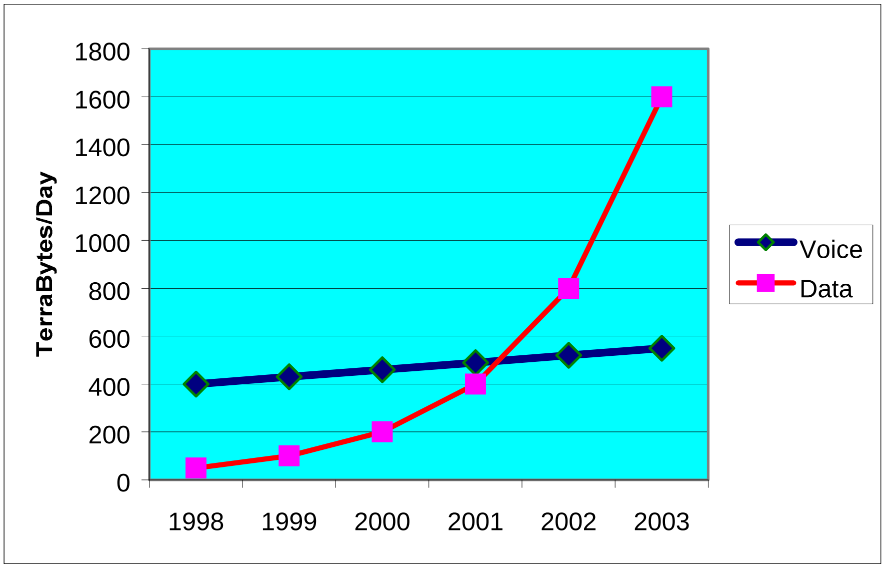


Fig 1: Voice growth vs Data growth

The new network solution

As more and more people hooked up to “the net” the issue was exacerbated and it continues to escalate. In the U.S it is estimated that total data bandwidth will have exceed total voice bandwidth by the year 2002 (figure 1.). The solution to these problems has been the evolution of a new digital packet switched network that can route the data traffic separately and keep it away from the circuit switched voice network.

The local access problems now have many solutions also, with the advent of Cable modems, satellite connections and DSL (digital subscriber line) services. These technologies all have the benefits of higher speeds often in the 1-5Mbps range as well as direct access straight into the Internet without ever having to touch the conventional telephone network.

Voice on the Internet

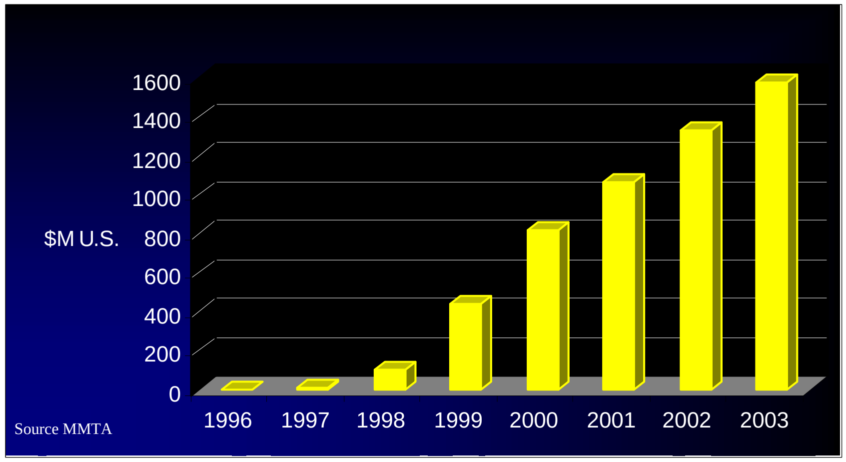
So now we have two distinct networks one for voice and one for data. As well as the fundamental circuit/packet differences the billing structure is very different indeed. We are used to paying more the further we want to call especially for international connections. On the Internet there is no distance component, once we pay for our connection into the net it costs the same to access a site in the next building as it does to get to one clear across the globe.

This caused many to think that with data having been carried on the voice networks for so many years, why not put the voice on the data network. This has lead to many opportunities where new companies have setup operations as voice providers when they own only a packet switched data infrastructure and build interconnections into the conventional phone networks. Service revenues from VoIP have been small by comparison to existing services but are forecast to exceed the $5billion mark by 2003.

A whole new architecture has been built and continues to evolve to support the Internet telephony revolution. As you would expect there are many hardware and software components, the next section provides a breakdown of the various architectural components and how they all fit together. The growth in sales of these key components is another indication of how rapidly this new infrastructure is taking hold and will begin to start taking service revenues from the existing networks. Figure 2 shows the growth of VoIP gateways which are responsible for the interconnect of the old and new networks and according to the MMTA have a CAGR approaching 40%.

Fig 2:
VoIP Gateways

Growth 1996 – 2003

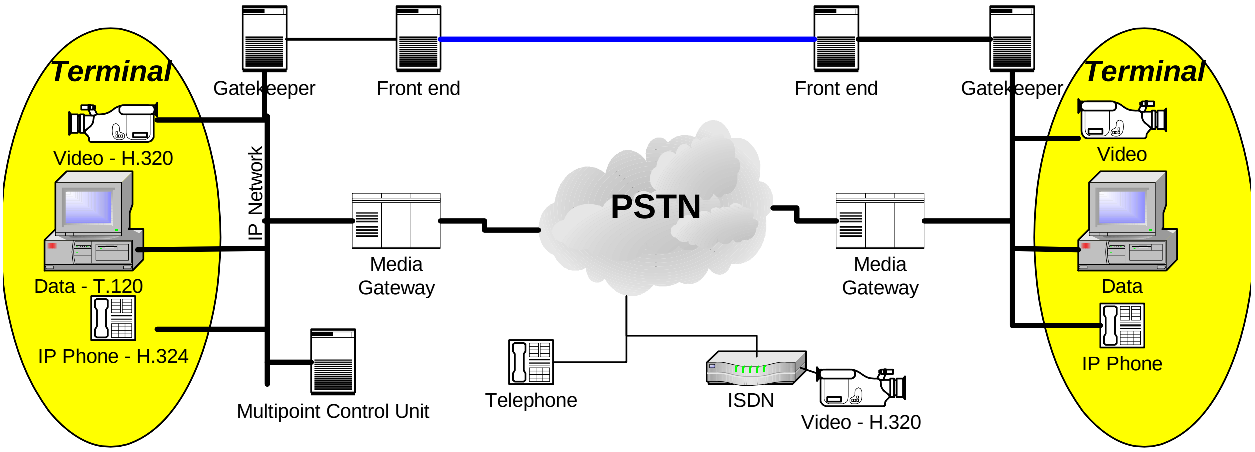


The fundamental building block components for VoIP

The term Voice over IP refers to a delivery mechanism that allows any IP based network to carry voice with higher or lower quality depending upon how the network is being engineered. On a dedicated Local Area Network (LAN) for instance, the voice quality can be higher than on the regular Public Switched Telephone Network (PSTN). However, if a network should become saturated, compression mechanisms become very important to maintain QoS (Quality of Service) levels. To communicate over an IP network, signaling protocols similar to the ones used by the PSTN such as SS7 or Channel 7 have been created. To setup, tear down and communicate over an IP network, VoIP uses four different types of elements:

1. Terminals
2. Gateways
3. Gatekeepers
4. Multipoint Control Units.

All these components have a different role within the network, although some of them are optional such as the gatekeeper, and all can be either built into a single system or spread over multiple systems at different physical and geographic locations.

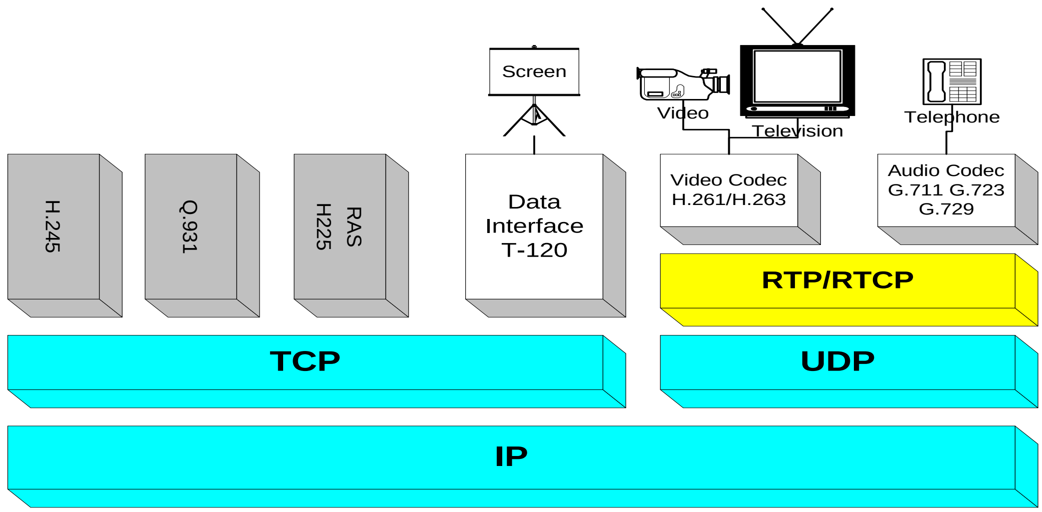


H.323 Network model

Terminals

A Voice over IP Terminal or client is the communications element that helps connect to real live calls. There are a number of client types found today and they must all be capable of supporting voice communications at a minimum but can optionally provide support for video or even data communications. The primary terminal type in use today is a software package such as Microsoft's NetMeeting, running on a PC. This presents the user with an interface through which they can make calls across the Internet. In the background it is responsible for the setup and teardown of the calls as well as the encoding and decoding of the transmitted/received voice such that the PCs microphone and speakers can be used instead of a regular handset. If you wish to use the PSTN and a regular telephone to make use of the Internet for voice calls a “virtual terminal” is used. This is a software module that would be part of a VoIP gateway deployed by the service provider and this provides the interfaces & protocol endpoint for the communication. This would then be converted by further functions inside the gateway into a regular phone connection.

So the terminal is the end user service that provides real time, two-way voice, video or data communications with other VoIP terminals. It communicates with the VoIP gateways using H.245 for call control, Q.931 for call setup and RAS for registration and administration with its local Gatekeeper. It interfaces with various elements such as a regular Plain Old Telephone Set (POTS) or microphone/speaker for the audio side, or to a camera/monitor for video transmission.



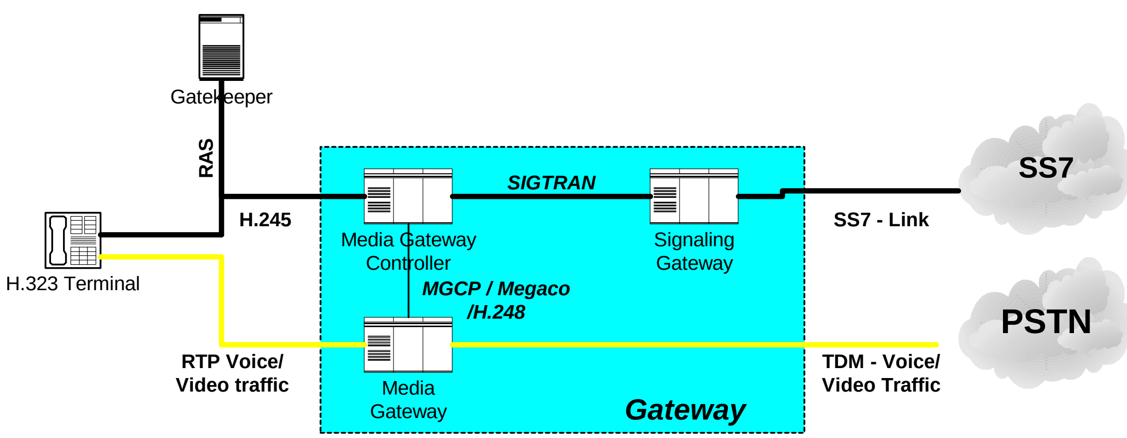
H.323 Terminal Software Model

Gateways

The whole concept of VoIP could not be totally viable if Internet telephony users could not talk to regular telephones. VoIP Gateways, provide the interconnection between the old school of traditional telephony and the digital world of Internet telephony. This way users of both technologies are able to communicate with one another. The primary function of the Gateway device is to provide translation services for the previously mentioned 'virtual' terminal as well as different transmission formats, communications procedures, and audio codecs.

The Gateway is a two-way interface between the telephone network and the IP-based network, therefore optional when there is no need to interconnect with a regular public switched telephone network, such as in an enterprise LAN only architecture. It assumes responsibility for setting up and taking down voice channels between H.323 and the PSTN network such as T1, B-ISDN, SS7, etc.

The Gateway model can be a single box that provides this interface but the Gateway model can be decomposed into three separate components and running on three different platforms:



H.323 Gateways

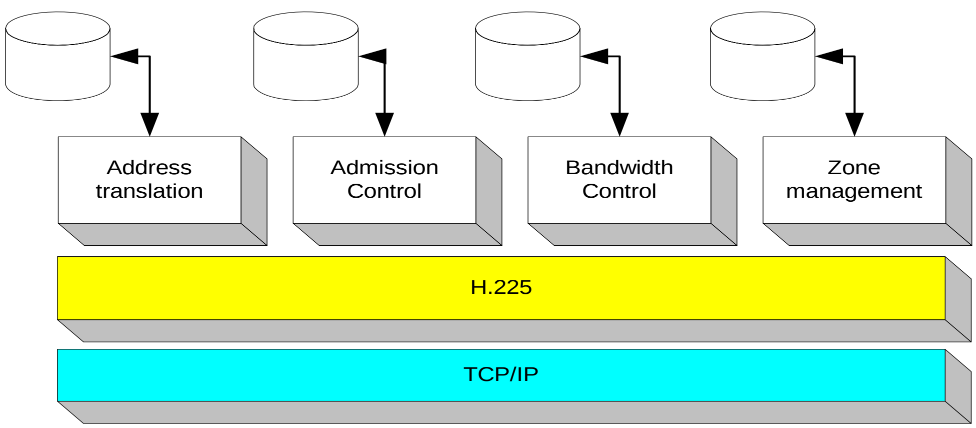
1. **Media Gateway**: provide the voice traffic translation between an IP based G.723.1 at 6.3kbps to G.711at 64kbps. From one side it is connected to a Local Area Network such as Ethernet 10/100BT,

and on the other side it assumes the connection to the telephone network as a T1 trunk or ISDN line for video communication with H.320 compliant video equipment. This platform is required to remain active all the time to prevent any discontinuity of service between two end points. A High availability platform is required with a minimum downtime and allows maintenance operation while the system is operational. This node controls jitter, delay, echo cancellation, or any other component that constitute quality of service (QoS).

2. **Media Gateway Controller**: provide the overall control of the gateway. It communicates with the Gatekeeper for database information regarding mapping between IP address and phone network.
3. **Signaling gateway**: responsible for the interface between SS7 signaling network and the VoIP signaling such as H.323.

Gatekeepers

Call control is key to any network managing voice communications; Gatekeepers provide these functions within the IP network. Many of these functions are provided by complex database management systems and include billing, address translation, routing and bandwidth management.
Database software constitutes the primary element of this platform. However, due to the kind of information this platform is required to provide, a fault tolerant or high availability platform is required with a reliable database management environment. Gatekeepers are linked together using a “border-element” or Super-Gatekeeper using the **H.225 Annex G** protocol definition.



Gatekeeper S/W modules

Gatekeeper mandatory functions

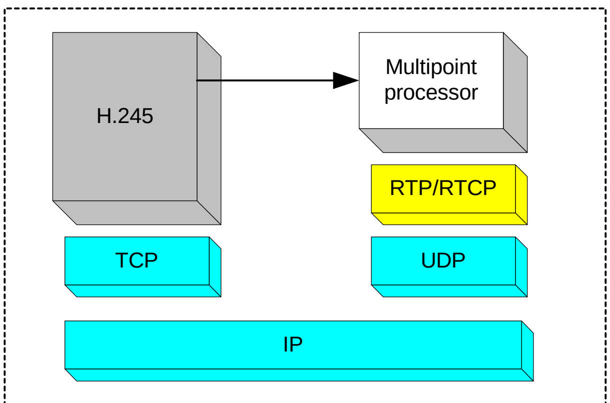
1. Address translation
   This function allows any endpoint to retrieve a transport address from an alias address or vice versa. On an IP network, this feature prevents H.323 endpoints (terminal, gateway, etc..) from keeping the alias to IP address translation locally, therefore preventing connection errors or unknown IP address.
2. Admission Control
   The Gatekeeper provides connection admission control based upon network bandwidth availability, authorization, or any other criteria that is relevant for a specific implementation.
3. Bandwidth Control
   The gatekeeper manages the bandwidth over the network and allows the best communication quality among endpoints.
4. Zone Management
   Defines which H.323 endpoints are currently managed by this Gatekeeper. Each endpoint is responsible for requesting a registration to this module, therefore benefiting from all the features provided by the Gatekeeper.

Overall, the Gatekeeper is a simple heavy-duty computer with high availability characteristics.

Multipoint Control Unit

One of the Internet's inherent advantages is the ability to build cooperative and collaborative environments. This ability is used heavily in the enterprise. The MCU is the component of the VoIP architecture that allows users to further collaborate by taking part in either telephone or videoconferences.

The MCU acts as an endpoint in the network that provides the capability for three or more H.323 terminals to participate in a conference. It consists of two parts: a Multipoint Controller and a Multipoint Processor (optional). The gatekeeper can explicitly invoke the MCU when more two or more endpoints are participating to the same conference call.



MCU

The MP provides the mixing and switching of all the audio/video and data among all the H.323 endpoints under the H.245 control. It has the same audio/video Vocoder functions as included into all other H.323 terminals or gateways, therefore allowing all automatic call attendant functions. For instance, it is possible to send a text to speech message to all attendants, or provide a DTMF function for specific features that the MCU would provide. Since the MCU has no geographic location requirement and since it uses the same element as other pieces of the H.323 network, it can be located on a local Gateway or Gatekeeper.

Standards and Protocols

There are two active groups working on the standards for VoIP. They are the ITU-T (International Telecommunication Union) Study Group 16, and the IETF (Internet Engineering Task Force). The IETF has a number of groups working on different aspects of IP telephony, including the MMUSIC (Multiparty Multimedia Session Control) working group. Two major standards are being recommended and deployed to satisfy the same needs H.323 (ITU) & SIP (IETF). SIP encoding is based on textual formats such as HTML or SMTP, etc and while simpler is not currently as widely employed as the H.323 standard. H.323 is based more on a traditional protocol stack with complex encoding and layering e.g. ASN1 within the traditional 7 layer OSI stack. The OSI approach had been limited by its cumbersome size and the compute power needed for such complex processing. With today's technology this is no longer a limitation that could hamper success of H.323 based solutions.

- **ITU-T - H.323**
  The ITU recommendation was ratified in its first version in 1996 and approved in its version 2 in February 1998. This recommendation describes the multimedia communication mechanism over a

network with no quality of service. Due to its latency and, non-deterministic aspect, an IP or IPX network over Ethernet, Fast Ethernet or token ring is considered to constitute such a network. H.323 recommendation uses the following protocols and recommendations:

- **RTP**: (Real-time Transport Protocol) provides end-to-end delivery services of real-time audio and video. It provides payload-type identification, sequence numbering, times-tamping, and delivery monitoring.
- **RTCP**: (Real-time Transport Control Protocol) is the counterpart of RTP that provides control services. It provides feedback on the quality of payload. Other functions include carrying transport level identifier for synchronization of audio and video data.
- **H.225.0 RAS**: Registration, Administration and Status. This layer is used between the IP Gateway and the IP Gatekeeper to exchange call setup and call tear down information, IP address mapping, etc…
- **H.225.0 Call Signaling**: used to establish a voice channel between two H.323 endpoints. The call-signaling channel is opened between two H.323 endpoints or between and endpoint and the gatekeeper.
- **H.245.0 Control Signaling**: used to exchange end-to-end control messages governing the operation of the H.323 endpoint. These control messages carry information related to the following:
  - Capabilities exchange,
  - Opening and closing of logical channels used to carry media streams,
  - Flow control messages,
  - General commands and indication
- **H.323 Annex D**: Real time fax over H.323 systems.

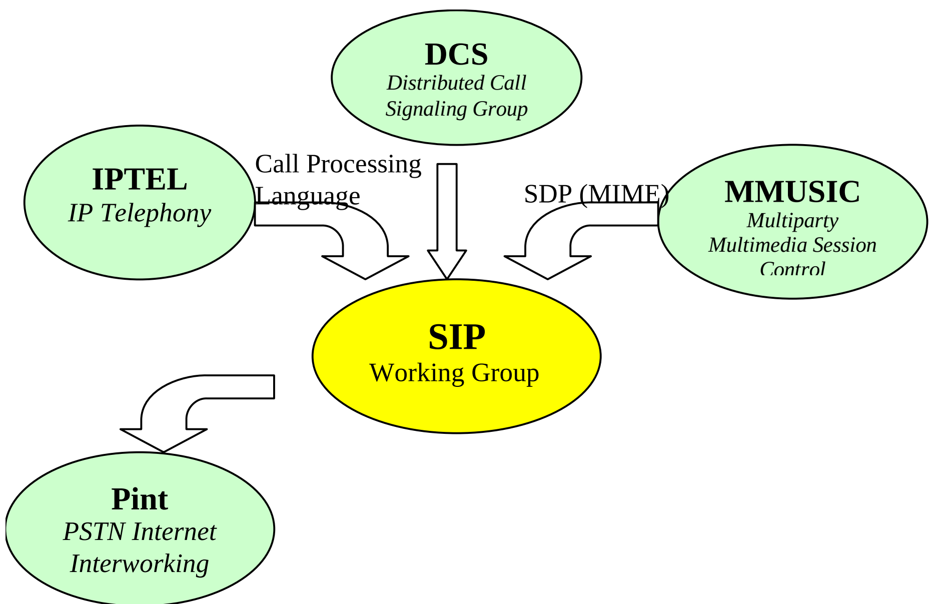
  This recommendation allows the transmission of a fax over a packet based communication system supporting H.323 standard. It describes the procedure for opening of a channel and the sending of T.38 (fax format) packets as would normally be done between two physical or logical fax endpoints.
- **H.450.x:** Supplementary services for H.323. These recommendations specify signaling procedures for optional features that H.323 endpoints (terminal, media gateway) may support. These services are fully deployed and used by today's telephone infrastructure.
  - H.450.2: Call transfer
  - H.450.3: Call diversion
  - H.450.4: Call hold
  - H.450.5: Call park and call pick up
  - H.450.6: Call waiting
  - H.450.7: Message waiting indication

- **IETF and SIP**

The IETF working group MMUSIC has developed several standards within Internet conferencing and telephony. The ITU would seem to have had a head start but IETF has been working on developing standards specifically with telephony in mind from the outset. The ITU approach considered the end to end protocols first and then looked to add the telephony specifics. The primary ratified and currently most utilized IETF standard is known as, SIP (Session Initialization Protocol) and is published under RFC 2543.

Last year, the IETF spun off SIP development into an independent working group to concentrate specifically on the work started within MMUSIC. The SIP group still maintains a close relationship with MMUSIC along with other related working groups such as Iptel (IP Telephony), Pint (PSTN Internet Interworking) and DCS (Distributed Call Signaling).

SIP takes a very different approach to VoIP than the H.323 standard. It is more akin to a series of recommendations than hard protocol boundaries. Even though its approach is simpler, in contrast to the ITU's H.323, additional services such as call transfer, call waiting, etc. are already included within the initial recommendation.



IETF Working Group relationship

When studying the differences between these two protocol sets lots of pros and cons can be identified and only the future will tell which one will become the ultimate basis for the world's VoIP environments. The SIP versus H.323 issue looks not dissimilar to the old X.400 versus SMTP battle, and currently real deployments seem to be in favor of H.323.

Other protocols are being developed by the IETF to communicate between the various pieces of the Media Gateway.

- SIGTRAN
  Part of the IETF, the SIGTRAN working group's primary charter is to address the transport of packet based PSTN signaling over IP Network. Within the VoIP architecture, it addresses more specifically the communication protocol between the Media gateway and the Signaling gateway.

- MeGaCo & MGCP Media Gateway Controller Protocol

Since the MG (Media Gateway) and the MGC (Media Gateway Controller) can be two distinct network elements, these Protocols are used to setup and tear down voice channel based on the signaling protocol (SS7, H.245, Q.931). However, this time the IETF and ITU-T (Study Group 16,) have agreed to publish a combined effort between the two working groups. This will allow an unified media gateway control protocol that is currently to be specified from the ground up. It will integrate all services needed to support VoIP media gateways for audio, video and data communications as well as packet networks such as ATM. Most of the large vendors such as Cisco, Lucent, Nortel, etc. have already committed to support such a protocol. This recommendation will be published by the ITU under the H.248 standard and is currently known as H.GCP.

Quality of Service

Among the most important issues that the IP telephony world has needed to resolve is that of Quality of Service (QoS). We have come to rely on the existing Public Switched Telephone Network and expect similar standards of service from any new networks. Due to the variable availability of bandwidth as well as the networks possible loss of data, VoIP has to address issues such as audio quality, signaling and call reliability in order to match the PSTN. Following is a description of the different elements involved in this challenge.

- **Voice Compression:** To minimize the amount of data transmitted over the packet network, the voice payload is compressed on its way into the network and decompressed when it reaches its destination. To perform this operation, special DSP (Digital Signal Processing) chipsets such as those developed by Motorola and Texas Instruments, are being used. This voice compression decompression functionality is also known as a “vocoder”. The issue related to the use of DSPs is due to the trade off between the compression required to limit the amount of data to be transmitted versus the audio quality. A high compression can result in a high degradation of the voice quality. Also, during the compression/decompression operation, an additional delay other than that caused by the network, is generated. To compare Vocoders, each algorithm is associated with its MOS (Mean Opinion Score). MOS is a number from 0 to 5. The higher the MOS is, the better the voice quality. For instance G.711 compression takes about 0.75ms and has a MOS of 4.4. A G.723.1 has a MOS of 3.5 and a delay of 30ms, however G.711 needs 64kbps of bandwidth as opposed to 5.3kbps with G.723.1.

The following table captures three kinds of compression algorithms currently available on IP telephony networks.

Compression Algorithm	Data Rate	MOS Score	Delay (ms)
G.711	64kbps	4.4	0.75
G.726	32kbps	4.2	1
G.723	5.3kbps	3.5	30

- **Latency Delay:** During a conversation between two phones or two endpoints, each physical element constitutes a delay to the propagation of the payload. Such elements being the Gateway, network, switches, DSPs, etc… This delay needs to be minimized so as not to become significantly annoying and obviously noticeable to the user.
- **Jitter:** Due to the IP based network, packets are not transmitted in a deterministic manner. If packets are not being delivered at regular time intervals, a conversation could be bursty and not very pleasant. To minimize this deficiency, a buffer is inserted at the receiving end that ensures delivery of the payload at a constant data rate. This is called the Jitter. Jitter requires fine-tuning and can be difficult to administer for each type of network and Jitter buffers only add to the latency issues.

- **Echo cancellation:** VoIP terminals can be subject to echoes caused by sound loops created between a speaker and microphones located on that same unit. It is necessary to cancel this echo first to minimize the amount of overhead this would generate to the network as well as the discomfort to the conversation itself. Echo-canceller algorithms such as G.168 are inserted within the IP network to eliminate such "noise".
- **Throughput**: Each voice channel requires a minimum network bandwidth. If the network does not provide such bandwidth, the call can be terminated.
- **Packet Loss**: During network congestion, some packets could be discarded this can cause significant voice degradation and eventually loss of service.
- **Signaling quality:** To establish a voice communication, it is important to provide fast and reliable call setup and call tear down of a voice channel so a caller does not experience undue delays in making calls.
- **Availability:** A voice network is expected to be operational whenever anybody needs to use it. We now take this as read when using the PSTN. VoIP networks need to meet this same expectation.

To help quality of service within a network, working groups such as Diffserv (IETF) are developing mechanism to improve the delays a network can generate. Methodologies are implemented to help routers to speed up their decision algorithm, and minimize the overall time delay of such network. The IETF has recommended a new protocol called MPLS (Multi-Protocol Label Switching) to help in this area**.**

What the future may hold

The IP telephony environment is required to provide the same type of services currently being offered by any Public Switched Telephone Network. Such services include audio/video conferencing, 911, toll free (800, 888), directory services, etc. Such services are as of today limited by the current VoIP implementation. New standards and recommendation are being developed as listed bellow:

➢ **IETF – SIP**
The SIP working group is currently working on new capabilities based on the call control specification, multiparty services and MIB support, etc…

➢ **ITU - H.323**
The ITU is also working on specifying new services to meet the current IP telephony trend as follows:

- **H.225.0 Annex G: Inter Domain Communication.** Defines the protocol used to interconnect two "Super-Gatekeepers" and share the telephony information between two IP networks.
- **H.323 Annex C:** H.323 on ATM
- **H.323 Annex D1:** Generic Functional Protocol
- **H.323 Annex D2:** Call Transfer Supplementary service
- **H.323 Annex D3:** Call Diversion Supplementary Service
- **H.323 Annex E** "Multiplexed Call Signaling Transport" defines multiplexed transport layer. First goal being UDP based H.323 with optional reliability methods such as
- **H.323 Annex F** "Simple Endpoint Type"
- **H.341 "Multimedia MIB"**

Regarding audio Vocoders, new DSP algorithms are being introduced to allow high quality audio traffic. These type of algorithms are meant to be used for radio broadcast, teleconferencing, or any other type of audio transfer that requires high quality audio on a low bandwidth network. One of them is known as G.722.1 and has been developed by PictureTel. It consists of a wideband audio coding at 50 to 7000Hz audio bandwidth as oppose to a traditional narrowband 300-3400Hz, and can be transferred over the IP network 16, 24 or 32 kbps.

Hardware reliability and availability

Linked to QoS of the software and protocol environment is also the establishment of a reliable and highly available hardware platform. In the world of conventional telephony we have come to expect “infinite dialtone” where the phone always works no matter what. Voice over IP must be able to meet the same standards of availability, however many of these new technologies are being built upon open standard systems. The advantage is time to market, as the long development times of the older proprietary telephony switch are untenable in this new environment. As such many developers are looking to technologies that can provide them the stable platform required on which to base their “carrier grade” VoIP solutions. CompactPCI is proving itself as the open standard technology of choice in this area.

The technological benefits of CompactPCI are now well recognized. Growing out of the mainstream PCI community the foundation is a solid one, based on tried and tested chipsets. The physical limitations of PCI slot cards made them impossible to safely be included within designs requiring high levels of availability and reliability. Coupling the PCI electrical specification with the IEEE 1101 mechanical standards gives CompactPCI the robustness required for carrier grade infrastructure deployment. One of the ultimate “killer benefits” provided by CompactPCI is its intrinsic ability to provide the hardware fundamentals required for high availability architectures. This is due in no small part to the Hot Swap specification which defines 3 models, Basic, Full and High Availability. The high availability model, while inheriting the attributes of the others (physical hotswap and software control) adds the ability for the application to control the system at a slot-by-slot level. This creates an environment where the software can have ultimate control of the hardware, literally “turning off” a slot that may have a failing component. This type of isolation is vital for the fault management functions required of a high availability solution aiming for the 5NINES goal, which translates to only 5 minutes of planned, or unplanned downtime each year.

At this time, Motorola Computer Group’s CPX8000 range of CompactPCI platforms is possibly the only CompactPCI solution proving capable of delivering these levels of availability. The CPX8000 family of CompactPCI® systems, when combined with appropriate software, is designed for critical telecom infrastructure applications that must meet 5NINES (99.999 percent) availability. CPX8000 systems have built-in redundancy for active system components—including system-slot central processing unit (CPU) boards—enabling active modules to be exchanged for repair or upgrade while the system continues to operate. Designed as a carrier-grade platform for operation in network equipment building standards (NEBS) and European Telecommunications Standards Institute (ETSI) environments, the CPX8000 family is particularly well suited for switching applications and deployment within unattended sites.

Conclusion

The market for VoIP solutions continues to increase building upon the huge growth and expansion of the Internet overall. This is underlined by manufacturers both established and startup who are rapidly releasing products to meet the needs of the growing community of Internet voice service providers. The standards will undoubtedly continue to evolve, adding features to match the existing infrastructure as well as new ones that capitalize on the packet data foundation of the new networks. As dependence on this new environment increases so will the demand for higher quality of service and highly available platforms. While there have been many attempts in the past to introduce new revolutionary technologies many have not be able to make the grade, VoIP has already established a significant foothold. The VoIP revolution continues unabated, Voice on the Net is here to stay – watch this space.