XILINX®

WP138 (v1.0) March 21, 2001

Voice-Data Convergence—Voice Over IP

Author: Amit Dhir

Introduction

This paper gives an overview of voice-data convergence technologies and how Xilinx high volume programmable devices can be used to overcome some of the significant challenges facing the designers of these systems. The Xilinx products targeted at these high-volume applications include XC9500XL™ and CoolRunner™ CPLDs and Spartan™-II FPGAs.

This appendix starts with an overview of voice-data convergence technologies and the benefits they bring to the users. We will then describe the product architectures that are used to implement VoIP gateways and IP phones. The final topic will be to show how Spartan-II devices can be used in these applications.

Making Calls Today Using the Public Switched Telephone Network (PSTN)

While all of us are familiar with the steps involved in making a telephone call, before we discuss convergence technologies it is important that we also understand the significance of the underlying infrastructure that is used to make a call today.

Current voice telephony is based on a circuit-switched infrastructure. This means that when a call is placed, the PSTN dedicates 64 Kbps connection, with a reserved end-to-end bandwidth. This bandwidth is allocated for you for the duration of the call on a fixed path or channel.

This mode of operation is referred to as connection oriented, and there is a set of mechanisms that are required to establish these connections. While we are not aware of it, picking up the phone and making a call involves a significant amount of activity within the telephone system. These activities map to user actions in the following way:

- **User picks up the phone:** The local central office (CO) detects the off hook condition, verifies that sufficient resources are available within the subscriber switch and then gives the user dial tone.
- **User dials number:** After dialing is complete the subscriber switch creates a connection request that is forwarded through the packet network that is used for managing connections. This process is referred to as signaling, and the uses the Common Channel Signaling System No. 7 (SS7) protocol. Resources are allocated as the connection request propagates through the network.
- **User waits for party to answer:** If end-to-end resources are available the user will hear a ring indication. If there is congestion in the network the user will hear a fast busy signal, an “all circuits are busy” announcement or some other indicator.

From the handset to the CO the voice is transmitted as analog signals over a two-wire connection. The CO performs 2-wire to 4-wire conversion and may require line echo cancellation for certain calls, especially long distance. At the CO, voice is digitized at 8 KHz rate using 8-bit companded format. The digitized voice reaches the destination CO via bridges, switches and the backbone. Voice is then converted to analog signals and sent to the receiving handset. The process is the same from the remote handset. The PSTN assigns a fixed path or channel for voice transfer and uses SS7 with DTMF tones to setup and tear down the calls.

A voice call generally does not utilize the full channel bandwidth. While PSTN supports full duplex transfer, phone calls involve one person talking and the other listening and vice versa. Also, there are many periods of silence where the network transmits no information. Hence, there is a pending issue about how we could better utilize the network bandwidth—and with certain features, voice calls can be carried out with less bandwidth.

Voice Data Convergence

Convergence when used in the context of networking refers to the ability to transfer data and voice and or video traffic on a single network. VoIP (Voice over Internet Protocol) is the transmission of voice traffic in packets. It is also known as Internet telephony, IP telephony, packet-voice and packetized voice.

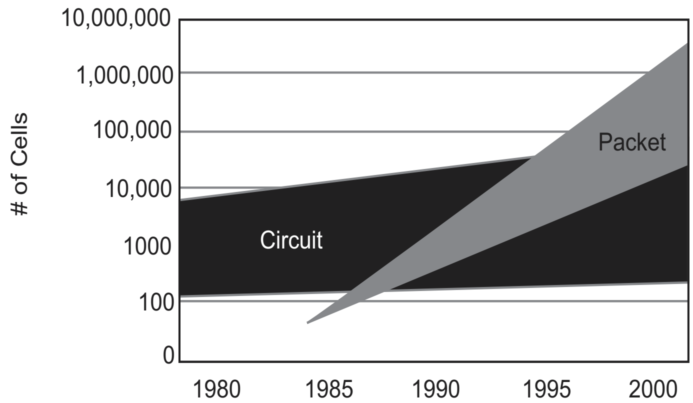

Figure 1: **The Rise in the Number of Calls Made Through the Packet-Switched Network vs. Circuit-Switched Network**

In VoIP networks, the packetization of voice happens in real-time. VoIP also decreases the bandwidth utilized significantly, since multiple packets can be transmitted simultaneously. The SS7 and TCP/IP networks are used together to set up and tear down the calls. Address Resolution Protocol (ARP) is also used in this process.

Process of creating IP packets:

- **Step 1:** An analog voice signal is converted to a linear pulse code modulation (PCM) digital stream (16 bits every 125 µs)
- **Step 2:** The line echo is removed from the PCM stream. It is further analyzed for silence suppression and tone detection.
- **Step 3**: The resulting PCM samples are converted to voice frames and a vocoder compresses the frames. G.729a creates a 10 ms long frame with 10 bytes of speech. It compresses the 128 Kbps linear PCM stream to 8 Kps.
- **Step 4:** The voice frames are integrated into voice packets. First a RTP packet with a 12-byte header is created. Then an 8-byte UDP packet with the source and destination address is added. Finally, a 20-byte IP header containing source and destination gateway IP addresses is added.
- **Step 5:** The packet is sent through the Internet where routers and switches examine the destination address, and route and deliver the packet appropriately to the destination. IP routing may require jumping from network to network and may pass through several nodes.
- **Step 6:** When the destination receives the packet, the packet goes through the reverse process for playback.

The IP packets are numbered as they are created and sent to the destination address. The receiving end must reassemble the packets in their correct order (when they arrive out of order) to create voice. The IP addresses and telephone numbers must be mapped properly.

Description	PSTN	Internet
Designed for	Voice Only	Packetized data, voice & video
Bandwidth Assignment	64Kbps (dedicated)	Full-line bandwidth over a period of time
Delivery	Guaranteed	Not Guaranteed
Delay	5-40ms (distance dependent)	Not predictable (usually more than PSTN)
Cost for the Service	Per minute for charges for long distance, monthly lat rate fr local access	Monthly flat rate for access
Voice quality	Toll quality	Depends on customer equipment
Connection Type	Telephone, PBX, switches with frame relay and ATM backbone	Modem, ISDN, T1/E1, Gateway, Switches, Routers, bridges, Backbone
Quality of Service	Real-time delivery	Not real-time delivery
Network Management	Homogeneous and interoperable at newtork and user level	Variuos styles with interoperability established at network layer only
Network Characteristics (Hardware)	Switching systems for assigned bandwidth	Routers & bridges for layer 3 and 2 switching
Network Characteristics (Software)	Homogeneous	Variuos interoperable software systems
Access Points	Telephones, PBX, PABX, switches, ISDN, high-speed trunks	Modem, ISDN, T1/E1 Gateway,high-speed DSL and cable modems

Figure 2: **PSTN (Circuit-Switched) vs. IP (Packet-Switched) Networks**

Motivation

Today we are seeing the integration of voice, video, and data in multiple applications. Using a unified packet network helps reduce bandwidth consumption by 8:1 in favor of packet-based networks.

The motivation for the deployment of convergence technologies is reduction in voice communication costs and enhanced services. The level of cost reduction can be significant—a study commissioned by Cisco Systems indicated that a business replacing an aging 100 user legacy voice network would see a 169 percent ROI over three years. These cost reductions come in several forms.

By eliminating the voice network infrastructure, the costs of maintaining both voice and data networks are eliminated. Convergence technologies also provide an efficient means of multiplexing voice and data traffic over private, wide area data networks. Individuals and businesses without private WANs can realize similar benefits through long-distance suppliers that offer lower long distance rates through the use of VoIP technology.

VoIP also provides enhanced features which come in the form of more flexible call routing, which is needed for unified messaging and call center applications. Another feature, which has been talked about for years, but is only just starting to become practical, is networked multimedia applications.

The advent of Internet telephony has come from consumers and business travelers looking for cheap international and inter-state calls. Web users are demanding free voice and video communications. Voice is also the logical step from ubiquitous Internet mail and instant messengers.

In addition, V-commerce (video and voice commerce) initiatives are coming to market. The ability to use voice and video as part of the Web experience helps sell and support consumers. It helps websites improve site stickiness in portals, communities, directories and audio ads—and making communication portals. Corporations can reduce cost and improve services through using VoIP services for distance learning, customer support and remote sales presentations.

With the growing demand for information appliances in today's homes there is a growing need for low-cost, integrated voice-data-voice access from consumer devices attached to the Internet.

However, there are several hurdles to overcome before there can be complete voice data convergence. The challenge involves creating a single network infrastructure that can efficiently handle the requirements of two classes of traffic that have fundamentally different characteristics. Voice and video (multimedia) streams require a more or less constant amount of bandwidth and are sensitive to delay variations in the network. The data traffic on the other hand is transmitted in bursts and relatively insensitive to network delay. With the connectionless nature of data networks, traffic competes for bandwidth on a real-time basis.

The Effect on Long Distance Carriers

The question now becomes one of, how does this technology effect the current long distance telephony market? Long distance carriers cannot afford to let this, large, profitable component of their business disappear. While there has been a considerable amount of pricing pressure on long distance services lately, these price reductions have been a result of cost reductions that the carriers have achieved and have passed through to consumers. Consequently, the carriers have been able to maintain the profitability of their long distance services.

Unfortunately Pandora's Box has been opened and consumers have heard of the concept of free long distance over the Internet. This combined with increasing competition in this space will keep the carriers motivated to continue to improve the efficiencies of their networks. This trend works in the favor of aggressive deployment of VoIP deployment by the carriers.

The other trend that is closely related to this are attempts by the carriers to retain revenue by bundling value-added services with long distance. While this is in many cases a packaging exercise, bundling Internet access for example, the enhanced services enabled by converged network technologies provides carriers with opportunities for truly differentiated product offerings.

VoIP Market Data

The VoIP market is just starting to take off with just $61 million of gateway equipment sold in 1998. While predicting growth in embryonic markets is always risky, research by Cahners In-Stat Group projects that sales will grow at a 280 percent average annual rate, reaching $3.8 billion in 2003. The VoIP market is expected to grow from 7.7 billion minutes in 2000, to 500 billion minutes by 2005, according to Probe Research. Probe Research also forecasts that the market for VoIP gateway equipment will increase from $1.2 billion in 2000 to $10 billion by 2005.

Challenges in Designing IP-Based Voice Systems

There is a rapid shift to standards-based architectures. While corporate telephony (PBX) has been based on proprietary designs, IP-telephony products are all based on IP—an open standards-based technology. Designers have to adhere to standards, placing a tougher load on product-validation and testing. Also, IP standards are constantly evolving and enhancements are being added.

The voice quality for VoIP products must match the quality of circuit-switched voice systems. The factors effecting voice quality are line noise, echo, voice coder used, and network delay.

The additional features and functions provided by a packet-switched network needs to be similar to a circuit-switched network that has been in use for several years. Features such as call waiting, toll-free numbers, credit-card billing, caller ID, and three-way calling will have to be supported by the network infrastructure.

Quality of Service

An important feature of converged networks is that they support Quality of Service (QoS) features for the transfer voice and video streams. QoS refers to a network's ability to deliver a guaranteed level of service to a user. The service level typically includes parameters such as minimum bandwidth, maximum delay, and jitter (delay variation).

An essential concept with QoS mechanisms is that they must be negotiated up front before the data transfer begins, a process referred to as signaling. The purpose of this negotiation process is to give the network equipment that is responsible for providing QoS an opportunity to determine if the required network resources are available and in most cases reserve the required resources before granting a QoS guarantee to the client. The overall process consists of the following steps:

- Client requests resource
- Network determines if request can be fulfilled
- Network signals yes or no to request
- If yes it reserves resources to meet need
- Client begins transferring data

Another contentious issue in the quest for converged networks is what is the appropriate layer of the protocol stack to merge the traffic.

The old way to combine the traffic was at Layer I using separate TDM circuits for voice and data traffic. The problem with this approach is that it is cumbersome to configure and made inefficient use of bandwidth since there was no statistical multiplexing between separate circuits.

Up until recently, the vision for voice data convergence was through the use of asynchronous transfer mode (ATM) at Layer II. The convergence of voice and data was the reason that ATM was developed over a decade ago. ATM has built in quality of service features that were defined just for this application. However, ATM has a fixed cell length and that leads to added overhead. Also, one must manage ATM and IP networks.

The most recent trend is to merge voice and data traffic at Layer III over IP networks. This approach takes advantage of new IP QoS features such as the Resource Reservation Protocol (RSVP) and Differentiated Services (DiffServ) technology. These technologies can also take advantage of layer two QoS features if available.

The Internet Engineering Task Force (IETF) has developed several technologies that are being deployed to add quality of service features to IP networks.

- **Resource Reservation Protocol (RSVP)** as defined in RFC 2205 is used by a host to request specific qualities of service from the network for particular application data streams or flows. RSVP is also used by routers to communicate QoS requests to all nodes along the path of the flow and to establish and maintain state. RSVP requests usually result in resources being reserved in each node along the data path.
- **Resource Allocation Protocol (RAP)** is a protocol defined within the IETF. RAP will be used by RSVP capable routers to communicate with policy servers within the network. Policy servers are used to determine who will be granted network resources and which requests will have priority in cases where there are insufficient network resources to satisfy all requests.
- **Common Open Policy Service (COPS)** is the base protocol used for communicating policy information within the RAP framework. COPS is defined in RFC 2748.
- **Differentiated Services (DiffServ)** as defined in RFCs 2474, 2475, 2597, and 2598 uses the Type of Service (TOS) field within the IP header to prioritize traffic. DiffServ defines a common understanding about the use and interpretation of this field.
- **Real-time Protocol (RTP)** is used for the transport of real-time data, including audio and video. Using the User Datagram Protocol (UDP) for transport, it is used in both media-on-demand and Internet telephony applications. RTP consists of a data and a control part; the latter is called Real Time Control Protocol (RTCP). The data part of RTP is a thin protocol providing timing reconstruction, loss detection, security and content identification.

- **Real Time Streaming Protocol (RTSP)** as defined in RFC 2326, and is a control extension to RTP. It adds VCR like functions such as rewind, fast forward and pause to streaming media.

VoIP Market Segments

There are currently five segments for VoIP products.

- **Enterprise VoIP gateways:** These devices are CPE deployed between a PBX and a WAN access device (typically a router) to provide call set-up, call routing and to convert voice into IP packets.
- **Service provider VoIP gateways:** These devices are used to aggregate incoming VoIP traffic and route the traffic accordingly; the role is somewhat analogous to a traditional Class 5 CO switch.
- **VoIP routers:** These products are standard routers with voice cards inserted into a chassis. The voice cards perform the packetization and compression functions, and the router then directs the packets to their ultimate destination.
- **VoIP remote access servers:** These products are updated versions of the systems that are used to provide dial-up access for ISP subscribers. In a VoIP the are used as dial-out VoIP gateways.
- **VoIP end stations:** These devices are telephone handsets that include a VoIP gateway and a LAN interface.

While each of these market segments has different product configurations, they all have the VoIP gateway technology common. The VoIP components are:

- Hardware components:
  - Digital signal processor
  - Controller
  - Codec
  - AFE (analog front end)
  - WAN interface
- Software components:
  - Voice processing elements: Speech coders, echo cancellers, voice activity detection (VAD), comfort noise generator (CNG), telephony
  - Protocol stack related elements: H.323 and TCP/IP protocol stack
  - Voice and LAN/WAN interface management
  - RTOS

VoIP Gateway Technologies

Figure 3 illustrates the functional architecture of a VoIP gateway and its three major functional blocks.

- **Voice processing:** This includes all of the functions required to encode voice data samples and packetize them for transmission.
- **Telephony signaling gateway:** The Telephony Signaling Gateway (TSG) subsystem performs the functions for establishing, maintaining, and terminating a call.
- **Network interface protocols**: The TCP/IP protocol suite provides the transmission structure for VoIP. Sitting under TCP/IP can be any number of OSI Layer II protocols including ATM, frame relay, or Ethernet.

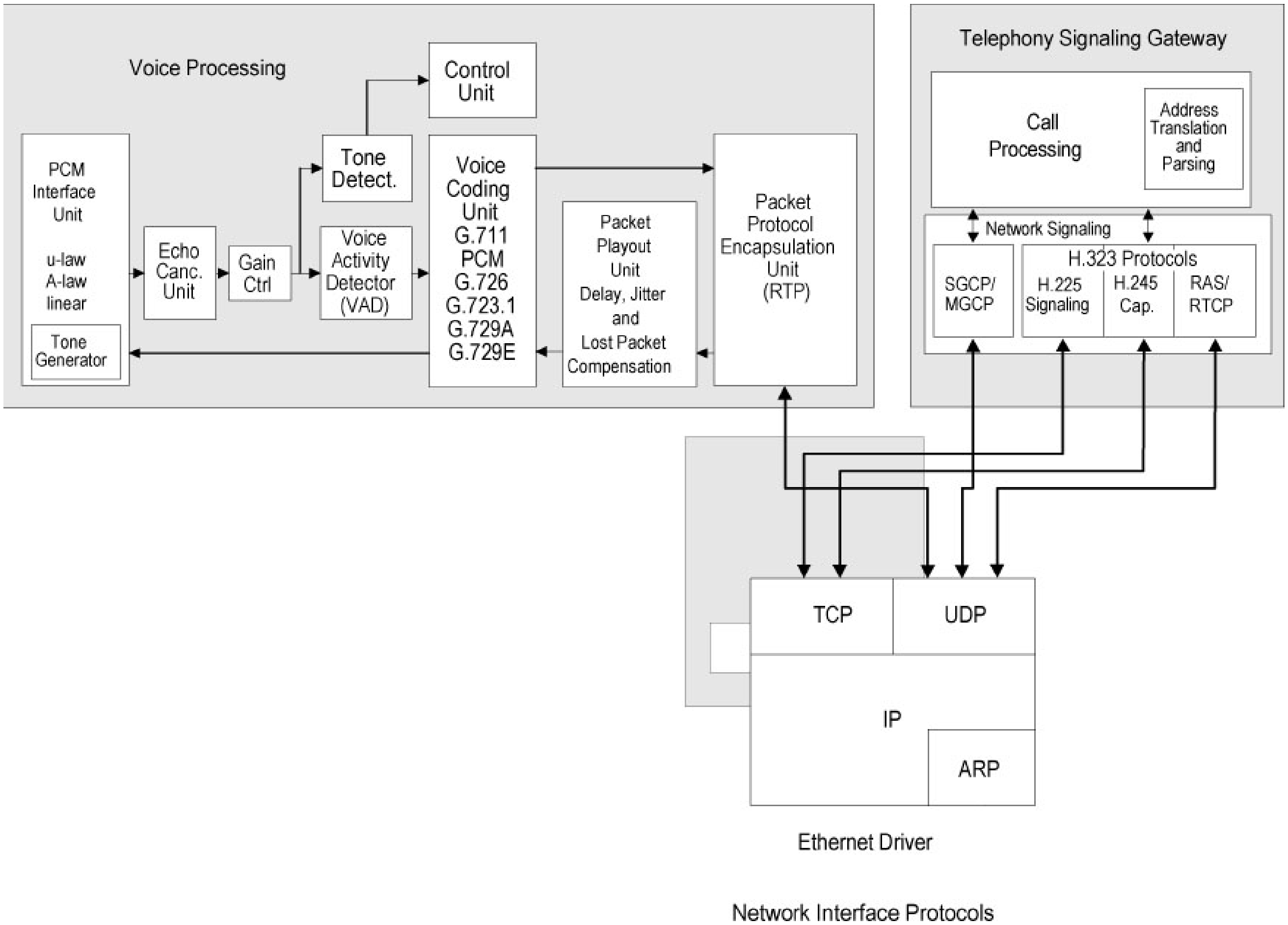

Figure 3: **VoIP Gateway Architecture**

VoIP Voice Processing

Voice processing functions include the following:

- **PCM interface:** Conditions PCM data and includes functions such as companding and resampling. This block also includes the Tone Generator, which generates DTMF tones and call progress tones.
- **Echo cancellation unit:** Performs echo cancellation on sampled, full-duplex voice port signals in accordance with the ITU G.165 or G.168 standard.
- **Voice activity detector:** Suppresses packet transmission when voice signals are not present. If no activity is detected for a period of time, the voice encoder output will not be transported across the network. Idle noise levels are also measured and reported to the destination so that “comfort noise” can be inserted into the call.
- **Tone detector:** Detects the reception of DTMF tones and discriminates between voice and facsimile signals.
- **Voice coding unit:** Compresses the voice data streams for transmission. There are several different codecs used to compress voice streams in VoIP applications. Each has been targeted at a different point in the tradeoff between data rate (compression ratio), its processing requirements, the processing delay, and audio quality. Table 1 compares the various ITU codecs with respect to these parameters. The mean opinion score (MOS) parameter is a measure of audio quality that is rated by a subjective testing process, where a score of five is excellent and a score of one is bad.

Table 1: **Voice Coding Standards**

Standard	Description	Data Rate (Kbps	Delay (ms)	MOS
G.711	Pules Code Modulation (PCM)	64	0.125	4.8
G.721, G.723, G.726	Adaptive Differential PCM (ADPCM)	16, 24, 32, 40	0.125	4.2
G.728	Low-Delay Codebook Excited Linear Predictors (LD-CELP)	16	2.5	4.2
G.729	Conjugate-Structure Algebraic CELP (CS-ACELP)	8	10	4.2
G.723.1	Another CS-CELP codec: Multipulse-Maximum Likelihood Quantizer (MP-MLQ)	5.3, 6.3	30	3.5, 3.98

- **Voice play-out:** Buffers the packets that are received and forwards them to the voice codec for play-out. This module provides an adaptive jitter buffer and a measurement mechanism that allows buffer sizes to be adapted to the performance of the network.
- **Packet voice protocol:** Encapsulates the compressed voice and fax data for transmission over the data network. Each packet includes a sequence number that allows the received packets to be delivered in the correct order. This also allows silence intervals to be reproduced properly and lost packets to be detected.

Delay and Echo Issues in VoIP

End-to-end delay is a significant issue when voice data is transported across IP networks. Delay can cause two types of problems echo, and speaker (talker) overlap. Echo becomes a problem if the delay exceeds 50 ms. Delay induced echo can be overcome by the use of echo cancellation technology. Speaker overlap, a problem of one talker stepping on the other talker's speech, becomes a problem if delay exceeds 250 ms and cannot be corrected.

Delay is introduced from several parts of the signal chain and has three major components algorithmic (accumulation) delay, processing delay, and network delay. Algorithmic delay comes primarily from the voice coding algorithms ranges from 125 µs to 30 ms. Processing delays are primarily due to packetization and can be up to 30 ms. Network delay itself has three primary components and includes transmission time due to signal propagation (for example 15 ms for New York to Los Angeles), delays through routers (on the order of 10 µs each), and the jitter buffering performed at the receiving end (70 to 100 ms).

Echoes are caused by signal reflections from far-end analog hybrids that are used to convert from a four wire to two-wire interface (signal reflections of the speaker's voice from the far-end telephone equipment back to the speaker's ear). These echoes can become objectionable if the loop (round-trip) delay is greater than 50 ms. Since loop delays will be greater than this value for virtually all VoIP connections, all VoIP gateways include an echo cancellation function. Echo cancellation performance is defined in ITU G.165 specification. An echo canceller is made up of the following major functional blocks:

- **Correlator**: Used to determine the loop delay
- **FIR filter:** Used to filter the echo components from the received signal
- **Speech detector:** Used to suppress updates of coefficients when there is no speech

Jitter is a variable inter-packet timing caused by the network a packet transverses. Jitter can be removed by collecting packets and holding them long enough. This allows the slowest packet to arrive in time to be played in the correct sequence—but this causes delays. For VoIP to be successful, a system must minimize delay and remove jitter.

Packet loss is another QoS issue for packetized voice. IP networks cannot provide a guarantee that packets will be delivered at all or in an order. Packet losses greater than ten are generally not tolerable, as packets will be dropped under peak loads and during periods of congestion. Normal TCP-based retransmission schemes are not suitable due to the time sensitivity of voice transmissions. Some approaches for compensating packet loss is by the interpolation of speech by replaying the last packet and by sending of redundant information.

VoIP Telephony Signaling

Telephony signaling functions include the following: (see Figure 4)

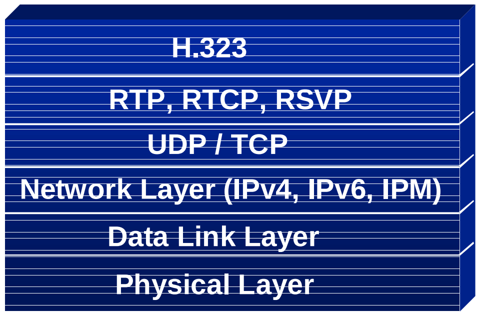

Figure 4: **VoIP Protocol Structure**

- **Call processing:** Performs the state machine processing for call establishment, call maintenance and call tear down. This also includes Address Translation and Parsing, which determines when a complete number has been dialed and makes this dialed number available for address translation.
- **Network signaling:** Performs signaling functions for establishment, maintenance and termination of calls over the IP network. There are two standards two widely used standards: H.323 and SGCP/MGCP.
  - **H.323 Protocols** H.323 is an ITU standard (Figure 5) that describes how multimedia communications occur between user terminals, network equipment, and assorted services on local and Wide Area IP networks. The following H.323 standards are used in VoIP gateways:
    - **H.225:** Call Signaling Protocols. Performs signaling for establishment and termination of call connections based on Q.931.
    - **H.245:** Control Protocol. Provides capability negotiation between the two end-points such as voice compression algorithm to use, conferencing requests, etc.
    - **RAS:** Registration, Admission, and Status Protocol. Used to convey the registration, admissions, bandwidth change, and status messages between IP Telephone devices and servers called Gatekeepers, which provide address translation and access control to devices.
    - **RTCP:** Real-time Transport Control Protocol. Provides statistics information for monitoring the quality of service of the voice call.
  - **SGCP/MGCP Protocols:** Simple Gateway Control Protocol (SGCP) is a standard that describes a master/slave protocol for establishing VoIP calls. The slave side or client resides in the gateway (IP telephone) and the master side resides in an entity referred to as a Call Agent. SGCP has been adopted by the Cable Modem industry as part of the DOCSIS standard. SGCP is evolving to the Multimedia Gateway Control Protocol (MGCP).

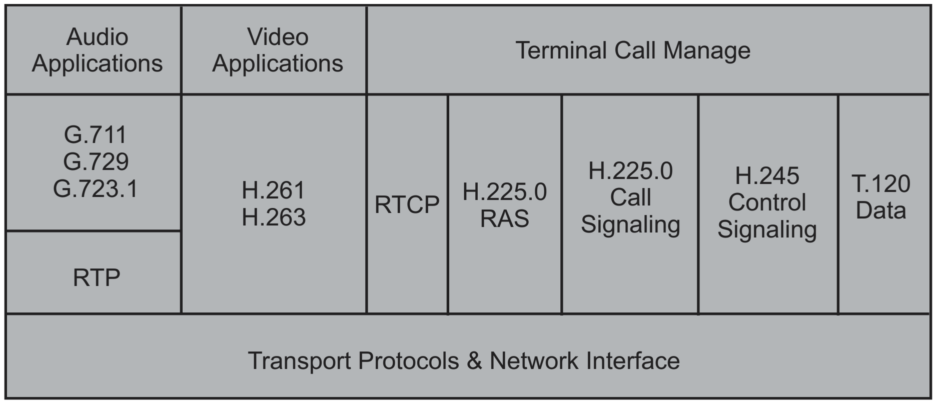

Figure 5: **H.323 Protocol Stack**

VoIP Variations

Fax over IP (FoIP)

For the most part, FoIP uses the same technologies that VoIP uses. The key difference is that unlike voice conversations, fax transmissions can be distributed in a non real-time fashion through the Internet to fax gateways where they are transferred in real-time to legacy fax equipment. FoIP service vendors provide servers for buffering fax traffic, distribution service including multi-way distribution, and FoIP gateways located strategically around the country.

Voice over DSL (VoDSL)

VoDSL is technology for transporting VoIP traffic across a DSL connection to a specialized VoIP gateway maintained by a Competitive Local Exchange Carrier (CLEC). Figure 6 shows how the Voice Service Gateway interconnects VoIP traffic to a Class 5, voice switch, and the PSTN. Subscribers can then either use IP Phones or specialized VoDSL gateways and standard phones to access both local and long distance telephone service. (Figure 6).

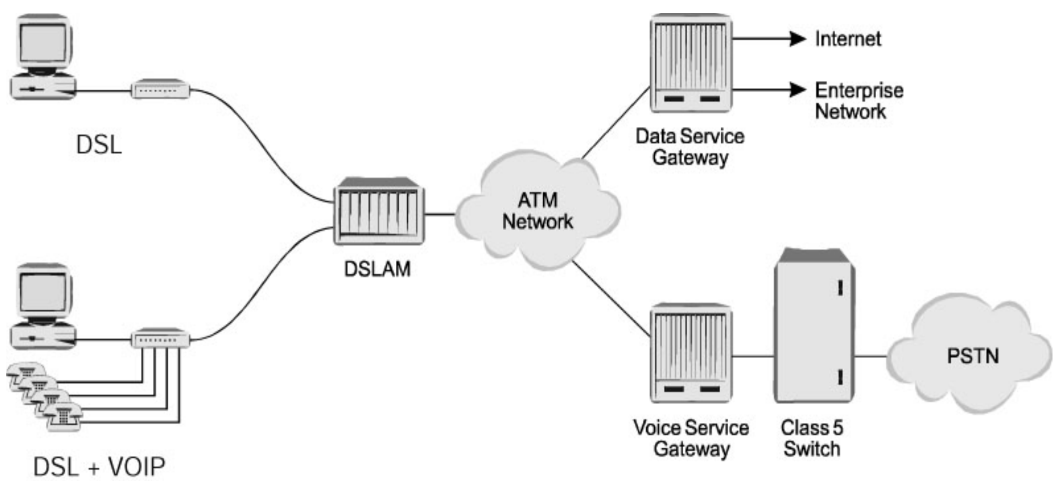

Figure 6: **VoDSL Network Topology**

VoDSL has two basic components:

1. **Voice gateway:** Voice packets are de-packetized and converted to a standards-based format (GR-303, TR-08, V5.X) for delivery to a Class-5 voice switch.
2. **Integrated Access Device (IAD)/CPE (Customer Premise Equipment):** This serves as a DSL modem and interfaces between a DSL network service and customer's voice and

data equipment. The packetization of voice traffic takes place on this unit utilizing standards-based technology. The CPE prioritizes the voice packets over data calls to ensure toll-quality voice delivery and then sends the packet over the DSL line.

There are several standards groups involved with VoDSL such as ANSI, ETSI, DAVC (Digital Audio Video Council), ATM Forum, ITU, and ADSL Forum.

Voice over Cable (VoCable)

The ability to deliver Internet and voice services over the cable infrastructure. Cable companies are struggling to upgrade buried cables from half-duplex to full-duplex. Half-duplex is the capability to transmit broadband in the downstream (from the headend cable companies to the subscriber) direction only. It makes it cumbersome for premium TV services such as pay-per-view that require upstream communication. Internet service is extremely inconvenient as e-mail messages and HTTP requests have to be sent via phone lines.

VoCable Issues:

- **Direct connect:** Cable networks were originally designed to broadcast only one signal to many recipients.
- **Security:** Prioritization of packet traffic with QoS and security features for voice communication are not completely addressed in even the second generation DOCSIS standard. Despite supporting data-encryption and authentication algorithms, cable-based IP telephony is susceptible to illegal wire tapping and inadvertent chat conditions.
- **Power consumption:** Telephones draw power from POTS lines, providing a life-line service even in case of power outages. VoIP phones connected to the cable, must have at least four hours of battery back-up.
- **Billing:** While most cable TV customers receive the same bill every month, telephone billing is very complex. Cable billing has special billing for pay-per-view requests only. Telephone bills include monthly service fees, long distance, and international charges, charges based on time and day, premium services, and pay-per-use services.

Implementing High-Capacity VoIP Gateways

Currently VoIP gateways support capacities of tens to hundreds of lines, but system vendors are increasing the densities to the hundreds to thousands range in anticipation of VoIP moving from the trial to adoption phase.

Creating these high-capacity systems is challenging due to the processing horsepower required to handle the channels. The current approach to this problem is to use arrays of high performance DSPs. Still for enterprise and carrier class systems a large number of DSPs are needed. For example, a single 100 MHz TMS320C5421 can support up to 12 channels, and more than eighty of them would be required to support a thousand lines.

At the system level, these high capacity gateways are implemented as CompactPCI systems with a H.110 telephony extensions for TDM data. The various functions are partitioned in to the following cards:

- **Voice processing card:** Consisting of DSPs and support logic
- **Line interface card:** Containing DS1, DS3, or ATM port interfaces and a management processor, which typically runs the SS7 signaling software
- **Signaling processor card:** Usually a RISC based single board computer

These boards interact through both the CompactPCI bus and the H.110 CT bus as shown in Figure 7. The CompactPCI bus is used primarily for management and status functions by the signaling CPU card. Fast Ethernet ports on the voice processing cards and the signaling CPU cards are used to bring packetized voice traffic and control traffic, respectively, to each. The H.110 CT bus is used to transfer PCM voice streams to the line interface cards so that they can be transferred to the PSTN.

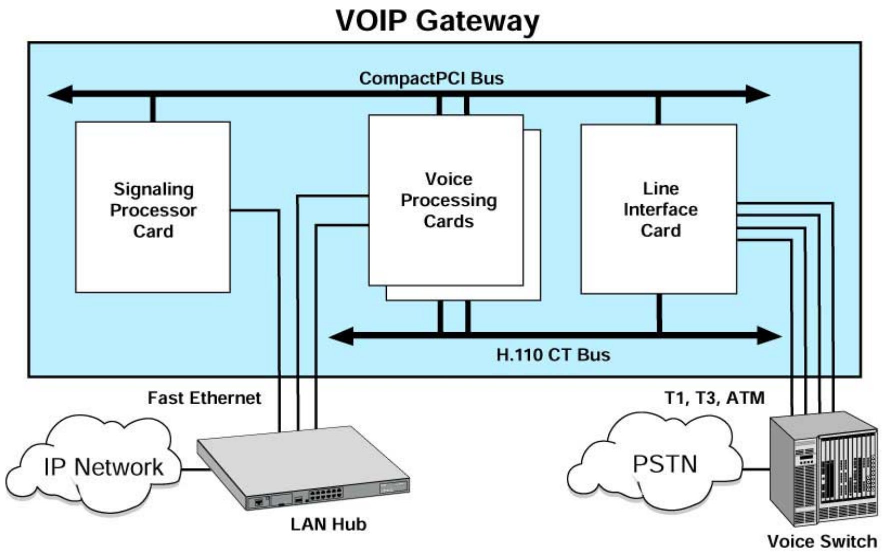

Figure 7: **High-Capacity VoIP Gateway Architecture**

- A typical voice-processing card includes:
- Eight 250 MHz DSPs each of which has 16 MB SDRAM
- Two 80 MHz processors for control, signaling and data processing functions
- H.110 compliant bus interface
- Two 10/100 Base-TX Ethernet interfaces

The manufacturer claims that this board is capable of supporting voice-processing functions for more than 192 channels. Boards such as this depend on complex datapath configurations to achieve their promised throughput. This requires a significant amount of complex glue logic including four PCI bridges and data path FIFOs.

IP Phones

IP phones are telephones that connect to a local area network rather the traditional phone jack. They are essentially a telephone with a built-in VoIP gateway and LAN interface circuitry. Processing functions in these systems are usually split between a DSP which handles voice processing functions, and a RISC processor which handles signaling, system management, and network protocol processing.

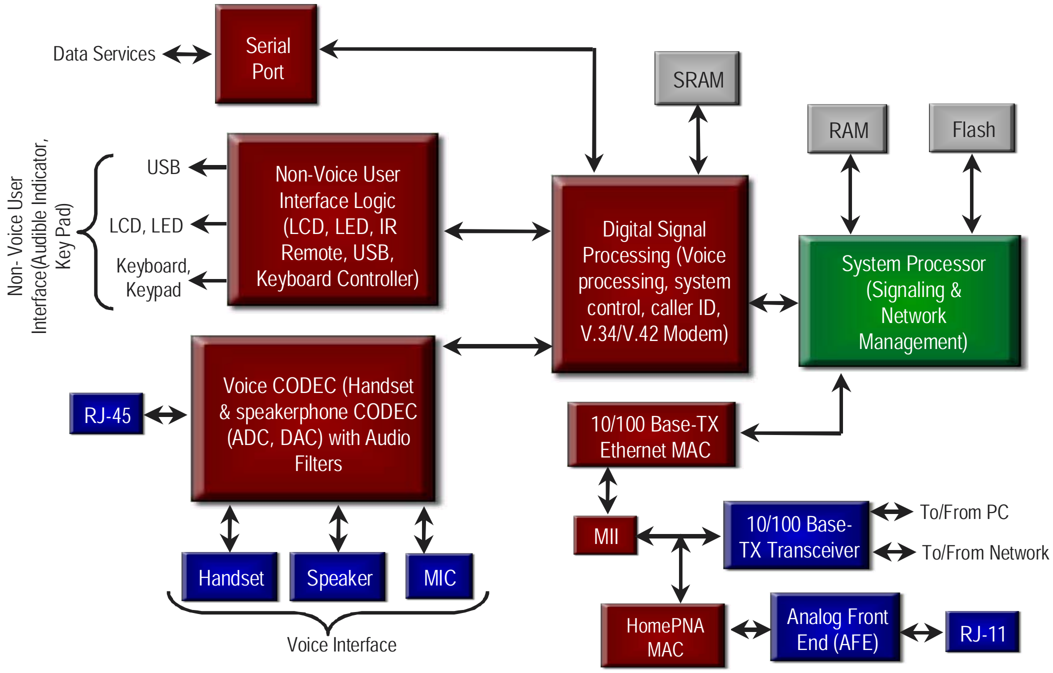

Figure 8: **IP Phone Block Diagram**

Figure 8 illustrates the architecture of a typical IP phone. Aside from the processing and network interface logic, the following functional blocks are also included:

- **Voice codec:** This block also includes the analog to digital and digital to analog conversion functions.
- **User interface logic:** This includes interfaces for the keypad, a status display, and an audio indicator used for ringing.
- **Optional data port:** This is typically a serial port and would be used for value added functions such a PDA synchronizing.

Xilinx Value Proposition

IP Phones

In spite of the high level of integration that is provided in the IP phone chipsets from ASSP manufacturers, there are numerous opportunities for value added differentiation using Spartan-II devices. This is because these chipsets represent a lowest common denominator solution, with few interfaces. Examples of enhanced features include:

- **PDA interfaces:** While both chipsets provide RS-232 serial ports for PDA interconnect only the Lucent offering provides a USB interface, and neither provides USB 2.0 support.
- **Non-Ethernet LAN interfaces:** While Fast Ethernet is the natural choice in a corporate environment, the use of IP phones in a small or home office environment will require support for one or more of the new home networking or wireless LAN standards such as HomePNA, IEEE 802.11b, HiperLAN2, and HomeRF.
- **Enhanced user interfaces:** While each of these devices provides support for a small LCD display, vendors wishing to provide a product with an enhanced user interface will have to provide an external interface solution.
- **Advanced features:** A natural extension to a phone that is network connected is to provide web browsing and e-mail support. The limited display and I/O peripheral support of these chipsets mandates the use external support logic to address this class of product.

High-Capacity Gateways

The Spartan-II FPGA devices are also useful in high-capacity gateway applications. Example of these applications include:

- **System level glue:** Implementing specialized PCI host bridges, DSP to processor interface logic, CT interfaces, data path switching, and FIFO functions.
- **Echo cancellation:** Xilinx has several application notes that demonstrate how much more effective FPGAs are then DSPs at implementing functions such as high-performance FIR filters and correlators.
- **Voice coding:** There are several IP vendors that provide voice coding cores for Xilinx devices. The ISS ADPCM core for example can process eight full duplex data streams and supports the G.721, G.723, G.726, G.726a, G.727, and G.727a ITU standards.

QoS Infrastructure

There are excellent opportunities for Spartan-II devices in QoS infrastructure applications. The key opportunity is in implementing the complex glue logic functions needed to interface network processors to switch fabrics or other ASSPs such as port interfaces. Spartan-II devices can be also be used as application specific coprocessors for network processors. In these applications they would be used to accelerate complex frame processing algorithms such as: traffic classification, traffic scheduling and shaping, complex policies, and queue management.

Conclusion

Internet telephony has grown up, and is now part of the mainstream communication scene. While a complicated technology, it provides cost and bandwidth savings to the consumer and the enterprise. The Xilinx Spartan-II devices present a low risk, cost effective way for system designers to develop VoIP gateways and solutions that can be taken directly to production.

Revision History

The following table shows the revision history for this document.

Date	Version	Revision
03/21/01	1.0	Initial Xilinx release.