COMDIAL®
The Rising Star

BUSINESS SOLUTIONS

Comdial Voice Over IP Voice Quality White Paper

Comdial Corporation, Inc.
1180 Seminole Trail
Charlottesville, VA 22901-2829

BUSINESS SOLUTIONS

Introduction

Voice over IP (VoIP) is new technology that allows the passing of speech information over IP data networks. In contrast to common traditional telecommunications, which moves voice traffic over the special Public Switched Telephone Network (PSTN), VoIP technology means that the voice shares the same network with the data in a true voice-data convergence. The Comdial IP-FX system uses VoIP networks to interconnect one another. VoIP has all the well publicized advantages inherent in having a unified data network infrastructure to manage all communications needs. Care must be taken, however to ensure that the voice quality is preserved when a data network is used to pass voice information. This paper explains the basics of Voice over IP technology and discusses some of the issues involved in maintaining good voice quality.

Traditional Telephony

In Comdial Telecommunications Systems and in most modern telecommunications equipment voice information is transported from phone to phone in a digital format. The speech information is first captured as analog signals, then transformed from analog signals to "coded" bytes of data by a process called Pulse Code Modulation (PCM). This PCM data flows in a continuous "stream" of data bits, 64,000 bits a second.

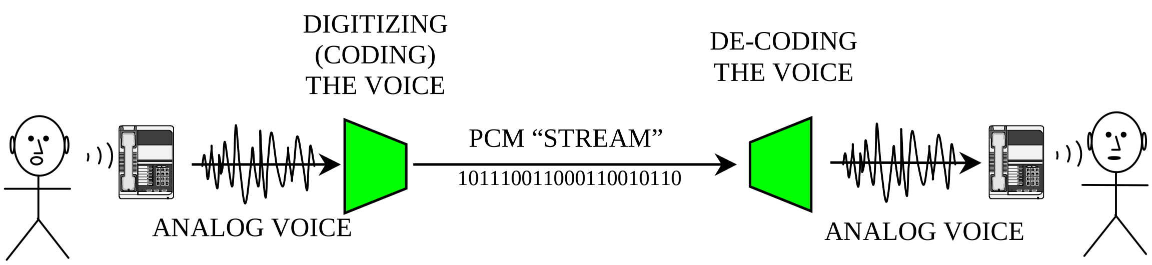

Figure 1. Voice to PCM Data Stream and Back

Usual conversations occur in both directions simultaneously so each end has to have a coder and a decoder. These two are referred to as the *CODEC*. The CODEC implements standards-based mathematical algorithms to code the voice information as data. There are a variety of standardized CODEC's. Traditional telephony in the U.S. uses the "G.711" CODEC in its "µ-law" form. The µ-law form of G.711 is common in North America. The alternate form, A-law is common in other parts of the world. G.711 provides very good speech quality and is often referred to as "*toll quality*."

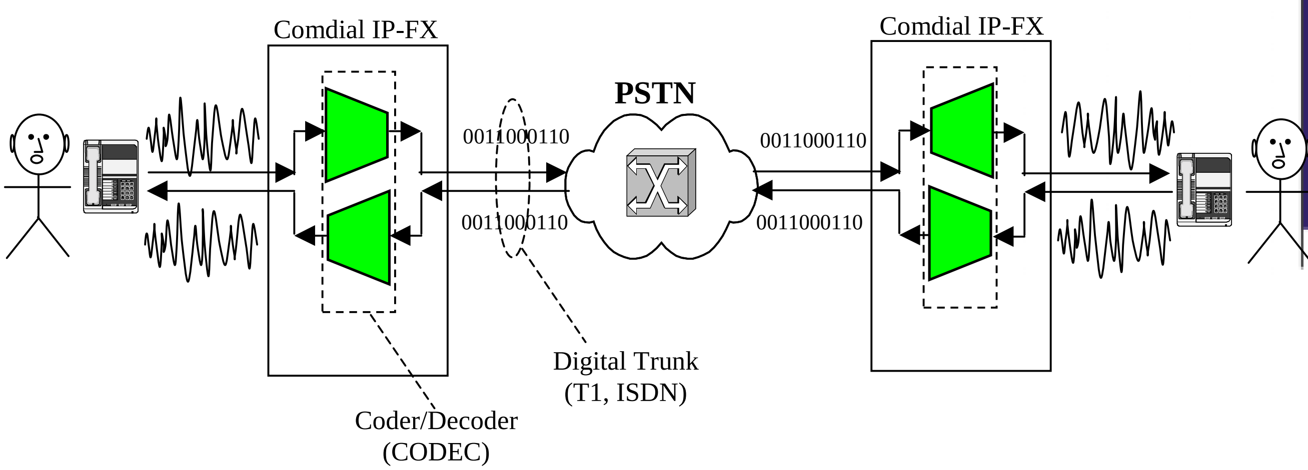

Figure 2. End-to-End Bi-Directional PCM Connection

The infrastructure of the Public Switched Telephone Network (PSTN) has many analog lines but from Central Office (CO) to CO, and in many cases from CO to businesses, voice is transported as streams of PCM data bits travelling over copper and/or fiber trunks (T1's, T3's, OC3's, etc). Internally, the CO's route the PCM data streams to the correct destination. This is referred to as *Circuit Switched* telephony. When a call is made, the CO *switches in* the correct pathway for the PCM data to travel to the correct destination. This same data pathway is used for the duration of the call. The size of the data "pipe" is exactly the size required by the transmit and receive PCM data streams.

VoIP

Voice over IP (VoIP) uses a different mechanism to transfer speech information. VoIP moves voice information over IP data networks (LAN's, WAN's, etc.). Unlike the PCM data streams in *Circuit Switched* telephony, VoIP data travels over networks in *packets*. In VoIP, the digitized voice is bundled into IP packets and sent out onto the network for delivery. Routers, switches and other network equipment direct the packets to their destination IP address. This is *Packet Switched* telephony. The voice packets must share the data network with any other data traffic already on the network (e.g. email, file transfers, Web/Internet accesses).

The Comdial IP-FX system uses microprocessors and high performance Digital Signal Processors (DSP's) to convert the system's internal PCM-encoded speech into IP packets for transmission across an IP data network. The DSP's provide additional CODEC capabilities and signal processing often required in VoIP transmission.

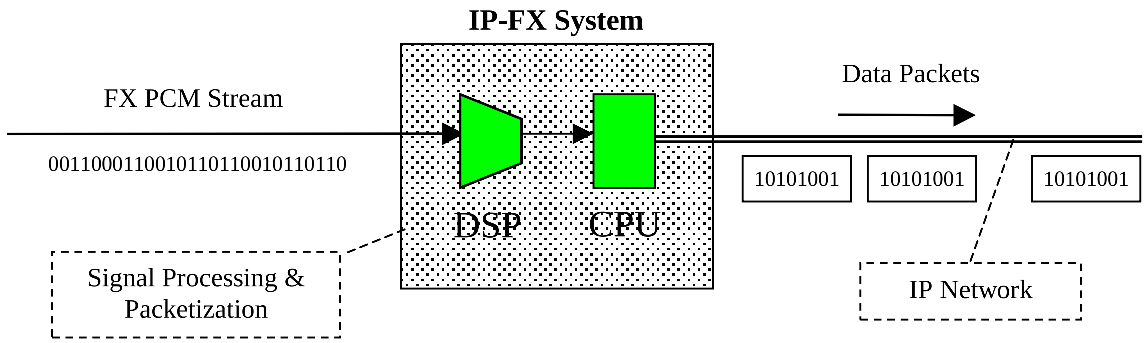

Figure 3. PCM to Packets

The IP-FX systems can be networked together using a data network as the connection method in a true convergence of voice and data.

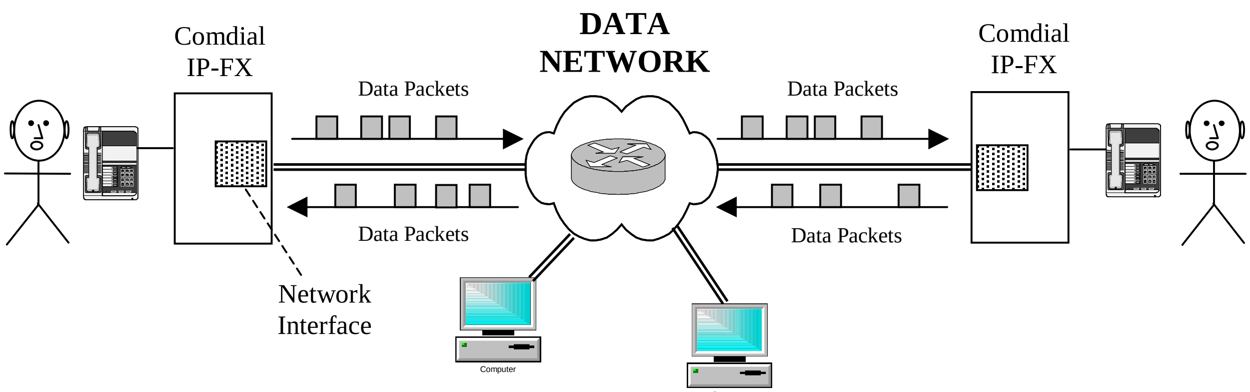

Figure 4. Networking IP-FX Systems Using a Data Network

Voice Quality With VoIP

Voice quality in VoIP depends on several interrelated things. First, the voice must be coded and decoded with a CODEC that has good intelligibility characteristics. Secondly, the voice data must travel from the source to the destination without delay, loss or corruption. In traditional *Circuit Switched* telephony, a continuous data "pipe" is provided through the PSTN to guarantee the flow of the PCM voice data. *Packet Switched* telephony on the other hand must overcome a variety of obstacles and impairments that data networks can provide to the regular and timely delivery of voice data packets to the far end. The transport of the voice packets is constrained by the amount of *bandwidth* available in the network connection, the *delay* that the packet experiences and any packet *loss* or *corruption* that occurs. In general, the measure of the ability of a data network to transport voice data packets quickly and consistently is referred to as the network's *Quality of Service* (QoS). It is very important to have a network with high QoS to ensure good voice quality. The next several sections will address the factors which affect QoS.

Bandwidth and CODEC's

The bandwidth (BW) is the measure of the number of bits per second that can flow through a network link at a given time. As various devices (e.g. PC's) use the data network, the bandwidth is consumed. Transfer of large files can cause momentary loss of bandwidth. *The available bandwidth* is the bandwidth that is left over after the various network devices' network usage has been accounted for. For a VoIP connection, lack of available bandwidth means that the connection will not support the added packet flow of the voice packets. Typically, the section with the smallest physical bandwidth is also the section with the smallest *available* bandwidth. This section will limit the VoIP data packet capacity.

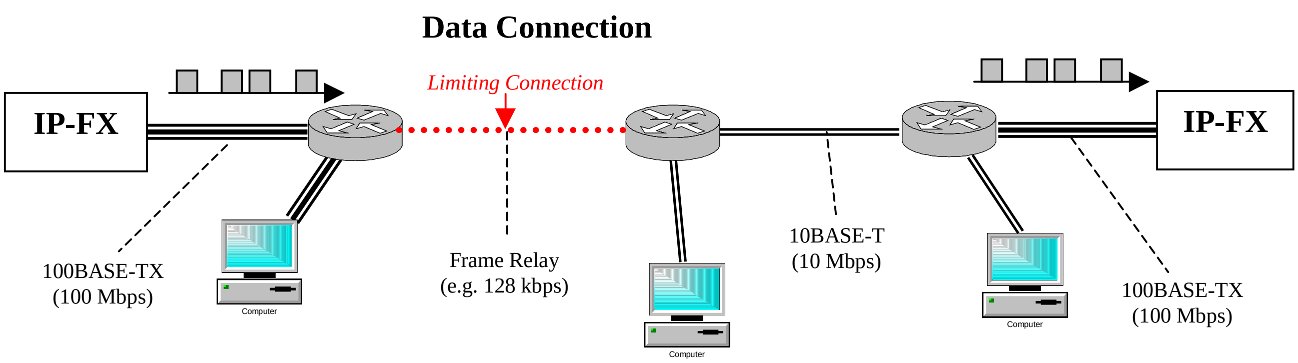

Figure 5. Different Bandwidth Sections in a Network

A data network supporting VoIP must be designed to have adequate bandwidth across the *entire VoIP connection path*, end-to-end. A 100 Base-TX network (at 100 Mbits/s) with ten times the bandwidth of a 10 Base-T network is only as good as a 10 Base-T network if one section of the network is the lower bandwidth technology. Network topology (or layout) and equipment type (e.g., switches vs. routers) also will dictate the available bandwidth of the end-to-end connection.

It is *critical* to have adequate network bandwidth to support the maximum VoIP data traffic. The bandwidth required is the product of the number of voice channels supported (simultaneous calls) and the bandwidth usage of each voice channel. The bandwidth usage per channel is determined primarily by the CODEC used and its associated overhead.

To provide several options, Comdial offers several different CODEC's (see Table 1). Different CODEC's use different amounts of compression to reduce the amount of voice data that must be sent, thereby reducing the amount of network bandwidth required to pass the voice packets. Table 1 lists the number of kilobits per second (kbps) required by the different CODEC's for a full-duplex Voice over IP conversation. (As a reference, a typical 33.6 modem passes 33.6 kbps of data in each direction.)

As the table indicates, CODEC's such as G.723 and G.729 significantly reduce the data bandwidth required. They accomplish this by using advanced algorithms that model voice in a highly compressed way. There is in general a tradeoff between using a high compression CODEC (with its low BW usage) and voice quality. The high compression CODEC's typically have slightly reduced voice quality, and introduce additional delay due to the added computational effort. The highest bandwidth is required by the minimal compression G.711 CODEC, which is the standard *toll quality* CODEC. Conversely, some of the high-compression low-bandwidth CODEC's are often considered to be comparable to cell phones in voice quality.

CODEC	Nominal Data Rate (one way only) (kbps)	Bi-directional (Total) IP Bandwidth Usage (kbps)		
		Small Jitter Tolerance	Medium Jitter Tolerance	Large Jitter Tolerance
G.711 (A-Law or u-Law) (*Toll Quality*)	64	221	175	159
G.726 (version 1)	32	157	111	95
G.726 (version 2)	16	125	79	63
G.729AB	8	63	40	32
G.723.1A (version 1)	6.3	44	44	29
G.723.1A (version 2)	5.3	44	44	29

Table 1. CODEC's Available on the IP-FX

The table indicates for example that one channel of G.711 VoIP traffic will use a maximum of 221 kbps of bandwidth. One 128 kbps frame relay channel would not support this CODEC. A 10Base-T Ethernet link could support 10 of these channels and utilize less than 25% of the available bandwidth. (In the real world the full theoretical bandwidth cannot be utilized due to congestion issues.)

It is important to note that these figures are for a VoIP call that is IP end-to-end. If a portion of the connection path is over Frame Relay or ATM, additional overhead may be added to "encapsulate" the IP packets into a different format. In this case the bandwidth requirement may increase by as much as 50%. This will depend upon the method used to pass the IP packets.

Voice Activity Detection

The IP-FX implements a signal processing algorithm that can reduce the bandwidth requirements shown in Table 1 by 20 to 50%. The feature is called Voice Activity Detection (VAD). With VAD, the speech processing circuit monitors the speech data for the presence of silence. In typical conversations, one of the parties is usually silent. When silence is detected, the DSP sends special "silence" packets to the far-end, thereby saving bandwidth. High-compression CODEC's G.723 and G.729 have VAD-like functionality built in to the compression algorithms and so do not gain additional benefit from the IP-FX's VAD feature.

Packet Delay

The second factor which significantly affects VoIP quality is the *delay* that the VoIP data packets experience as they traverse the network. Regular packet delays longer than 100 milliseconds begin to interfere with normal conversation. Longer delays can cause echoes, which can make normal conversation difficult. The total end-to-end packet delay is the result of the many incremental delays along the connection path. Some small delay occurs during the digital signal processing of the speech signals by the IP-FX. As the VoIP packets travel across the network, a number of sometimes large delays can occur.

The CODEC's specified for use with VoIP for the processing of the speech impose small processing delays. The high compression CODEC's (G.723.1A and G.729AB) require greater processing power and so introduce incrementally more delay than the low-compression CODEC's. The packet delay experienced during signal processing is typically on the order of tens of milliseconds.

VoIP packet delay in the IP network can be quite large and can vary significantly from one network to another. Each router traversed in a packet's travel across a network adds an additional delay. The router's or switch's design and the configuration of its "queueing" (buffering) size will affect the delay each router adds to the voice packets. On the reception end, packet (or "playout") buffers can add more delay. These network delays can only be mitigated though careful network architecture, equipment selection and configuration. Minimizing the number of router "hops" along the path reduces delay. Small router queue sizes and high bandwidth connections also reduce delay. (Unfortunately, small router queue sizes may result in lost packets.)

Echo

Echo can result when the cumulative delay from transmitter to receiver and back is too long. Using VoIP as a means to pass signals adds some delay, as indicated above, though not enough, typically, to cause the user to experience echo. If the VoIP call is a call over a network with very bad delay (e.g. the Internet) or if the call goes out over a long-distance PSTN line that has excessive delay (e.g. using a satellite connection) then echo can occur. To suppress the echo, the Comdial IP-FX implements a standards-based G.168 echo canceller to help reduce the echo signal returning from the far-end.

Jitter

In addition to the average delay that all the VoIP packets experience, *individual* packets can experience a little extra delay (or *jitter*) relative to the other packets in the data stream. This is due to instantaneous network usage and congestion or to traversing different routes through the network. Jitter causes some packets to arrive at different times at the destination relative to one another, sometimes causing them to arrive out of sequence. Obviously this will cause the CODEC to have difficulty recreating the speech on the receiving end, and voice quality will be impaired.

Comdial provides a way to reduce the jitter effect by using a *jitter buffer* to buffer incoming packets. The IP-FX then reassembles the packets in the correct order. The IP-FX allows the configuration of this jitter buffer to a small, medium or large size. This buffer makes it possible to reorganize late packets at the cost of an incremental delay.

Lost Packets

Another factor that will affect Voice over IP quality is lost or damaged packets. In the process of traversing the network occasionally a packet can get lost. This can occur if a router queue along the network path is overloaded. In this case the router will typically discard the packet. It is also possible for a data packet to become damaged or corrupted during its travel through the network. In both cases, these packets are unusable. The packets in VoIP use a protocol called UDP. This protocol assumes that the packets arrive correctly. The packets are not retransmitted if their delivery fails or they are received in a corrupted state.

In the Comdial IP-FX the voice CODEC on the receiving end will extrapolate new packets to fill in the gaps caused by the missing packets. If the packet damage or loss is severe, voice quality will begin to degrade. The impact on voice quality will depend upon the CODEC used.

Excessive packet loss may also be mitigated by using a large *jitter tolerance* setting (see Table 1). This increases the size of each individual voice packet, resulting in less overhead in the data packets and a slightly improved network utilization.

In any case, lost packets indicate poor network quality or conditions and signals the need for network re-engineering efforts.

General QoS

In general, the Quality of Service of a data network determines whether VoIP can function with adequate voice quality. The public Internet's problems with low bandwidth and especially long and uncontrollable delay have made VoIP over the Internet difficult. For years PC-base IP phones have attempted to use the Internet to make calls as a way of *toll-bypass*. These calls have suffered from poor voice quality. In general, there is no guaranteed Quality of Service over the Internet, and QoS is key to maintaining good voice quality. It is *very important* with VoIP to have a well designed data network to ensure high voice quality. Correct network architecture, equipment selection and equipment configuration is important to guarantee a high QoS for Voice over IP. In general, this is only possible on private data networks.

The networking industry is continually developing various standards to help guarantee the Quality of Service that time-critical data such as VoIP requires. Some of these include:

- Differentiated Services ("DiffServ") which instructs the network routers to route based on priority bits in the packet header (the same bits as the IP Precedence of IPv4).
- Integrated Services (and RSVP) to set up end-to-end virtual pathways that have reserved bandwidth similar to Circuit-Switched telephony.

- Multi-Protocol Labeling Switching (MPLS) which uses labels inserted into the packets to route traffic in an efficient way.
- Subnet Bandwidth Management (IEEE 802.1p) which uses priority bits in the packet header to force the routers to give packets priority through routing at “Layer 2”.

It will be important to utilize network equipment that complies with these QoS standards to guarantee the necessary QoS for high voice quality. It is important to investigate the QoS mechanisms that the router and switch vendors are offering with their equipment and to use them. In some cases, QoS can be ensured by appropriately configuring specific ports of the network router or switch that are connected to the VoIP equipment.

In general, when the VoIP packets travel from one kind of network to another, care must be used to be sure that the QoS mechanisms are maintained end-to-end. This is true, for example, when two “edge” IP networks, each with an FX are inter-connected by an ATM or Frame Relay connection. The equipment that converts (or encapsulates) the IP packets to the Frame Relay or ATM medium must preserve the QoS information embedded in the packet headers.

In any Voice over IP deployment, it is important to work closely with the network personnel and equipment to configure the available QoS mechanisms and to instigate performance monitoring for delay, congestion, lost packets, etc. as a means of monitoring the QoS of the actual network.

Conclusion

Traditional telephony passes speech from end-to-end as continuous bit streams. By converting speech to data packets, the Comdial IP-FX enables speech to pass over data networks, providing the cost advantages of one unified infrastructure for data and voice. There are various voice processing CODEC’s and other signal processing options to help configure the IP-FX to match the available data network. Fundamentally, the VoIP quality will be dictated by the Quality of Service available on the data network used. It is very important with VoIP to have a properly designed and configured data network to ensure high voice quality.

Cary Lancaster
Senior Principal Engineer
Comdial Corporation, Inc.
clancaster@comdial.com

COMDIAL®
The Rising Star

BUSINESS SOLUTIONS

Glossary

ATM - Asynchronous Transfer Mode. A cell-based fixed packet length protocol for transporting data cells or packets of fixed length over a network. ATM has inherent QoS mechanisms in its design.

Bandwidth - The measure of the number of bits per second that can flow through a network link at a given time.

Circuit Switched – Modern digital telephony moves voice information as constant "streams" of PCM data bits. When a call is made, the CO connects together (or switches) the PCM streams from each phone.

CO - Central Office. The Telephone Company site where standard telephone calls are directed to their correct destination.

CODEC - Coder/Decoder. An algorithm or device, typically based on an Industry standard that both codes and compresses voice data.

DSP - Digital Signal Processor. An integrated circuit device similar to a microprocessor but optimized for mathematical processing such as the algorithms used with VoIP.

G.711 - The ITU-T standard CODEC that is used in most ,modern telecommunications equipment. This CODEC is considered the reference for speech quality by which other CODEC's are judged. It is known as "toll quality".

IP - Internet Protocol. An IETF protocol used in packet networks to pass data packets.

Jitter - The variable delay that individual packets can experience when traversing a network.

PCM - Pulse Code Modulation. A standardized method of coding analog speech into binary (digital) code words.

PSTN - The Public Switched Telephone Network. The entire public telephone system, which enables directing (or switching) calls from one user to another.

QoS - Quality of Service. A measure of a networks quality with respect to its ability to move data packets from their source to destination without delay, jitter, corruption, etc.

VAD - Voice Activity Detect. A processing algorithm that monitors the speech for the presence of silence and sends silence packets out-of-band (as a unique code) to the far-end, rather than sending empty data packets.

VoIP - Voice over IP. The method of voice communication that uses IP networks as a medium to move the voice information represented as data packets.

µ-Law - One form of the ITU-T's G.711 CODEC that is typically used in the United States. A-law is another version that is used in most of the rest of the world.