# Voice and Fax over Internet Protocol (V/FoIP)

## Definition

A voice-over–Internet protocol (VoIP) application meets the challenges of combining legacy voice networks and packet networks by allowing both voice and signaling information to be transported over the packet network. A fax-over–Internet protocol (FoIP) application enables the interworking of standard fax machines with packet networks. It accomplishes this by extracting the fax image from an analog signal and carrying it as digital data over the packet network.

## Overview

Organizations around the world seek to reduce rising communications costs. The consolidation of separate voice, fax, and data resources offers an opportunity for significant savings. Accordingly, the challenge of integrating voice, fax, and data is becoming a rising priority for many network managers. Organizations are pursuing solutions that will enable them to take advantage of excess capacity on broadband networks for voice, fax, and data transmission, as well as utilize the Internet and company Intranets as alternatives to costlier mediums.

This tutorial discusses the principles related to implementing real-time voice- and fax-over-packet networks. An overview of the embedded software architecture is presented, and a system is described for sending voice, fax image, data, and signaling information over the packet network. Benefits to designers and manufacturers of this embedded approach are lower cost of goods sold, faster time to market, and lower development costs. Customers can gain a considerable advantage in time to market in building their communication systems.

This tutorial references a general class of packet networks, as the modular software objects allow networks such as asynchronous transfer mode (ATM), frame relay, and Internet/intranet (Internet protocol [IP]) to transport voice and fax. Currently, the Frame Relay Forum and the International Telecommunication Union (ITU) have defined protocols for transmission of fax over a packet network. However, the principles described are equally applicable to ATM networks.

# Topics

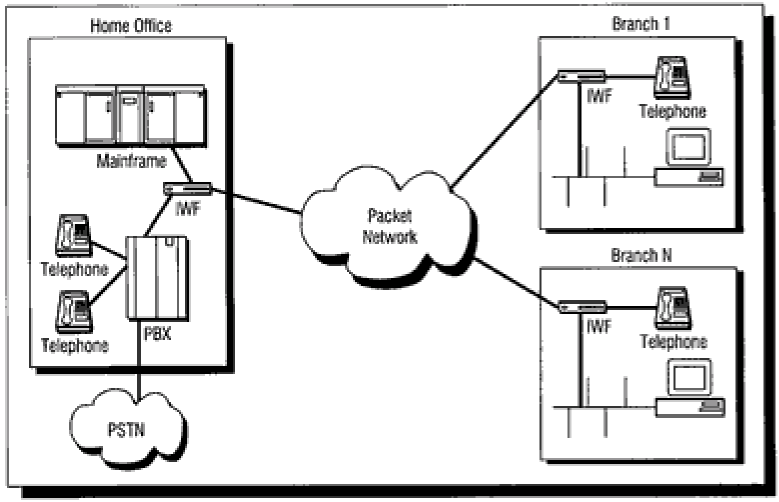


# 1. VoIP Applications

A wide variety of applications are enabled by the transmission of VoIP networks. This tutorial will explore three examples of these applications.

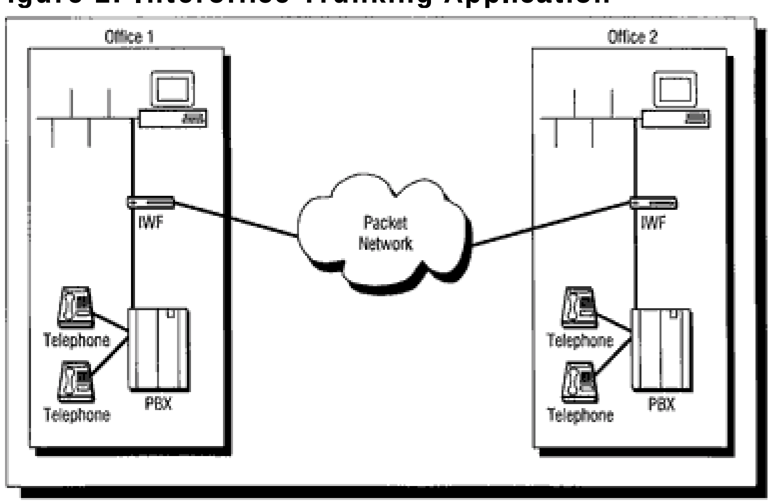
The first application, shown in *Figure 1*, is a network configuration of an organization with many branch offices (e.g., a bank) that wants to reduce costs and combine traffic to provide voice and data access to the main office. This is accomplished by using a packet network to provide standard data transmission while at the same time enhancing it to carry voice traffic along with the data. Typically, this network configuration will benefit if the voice traffic is compressed as a result of the low bandwidth available for this access application. Voice over packet provides the interworking function (IWF), which is the physical implementation of the hardware and software that allows the transmission of combined voice and data over the packet network. The interfaces the IWF must support in this case are analog interfaces, which directly connect to telephones or key systems. The IWF must emulate the functions of both a private branch exchange (PBX) for the telephony terminals at the branches, as well as the functions of the telephony terminals for the PBX at the home office. The IWF accomplishes this by implementing signaling software that performs these functions.

**Figure 1. Branch Office Application**



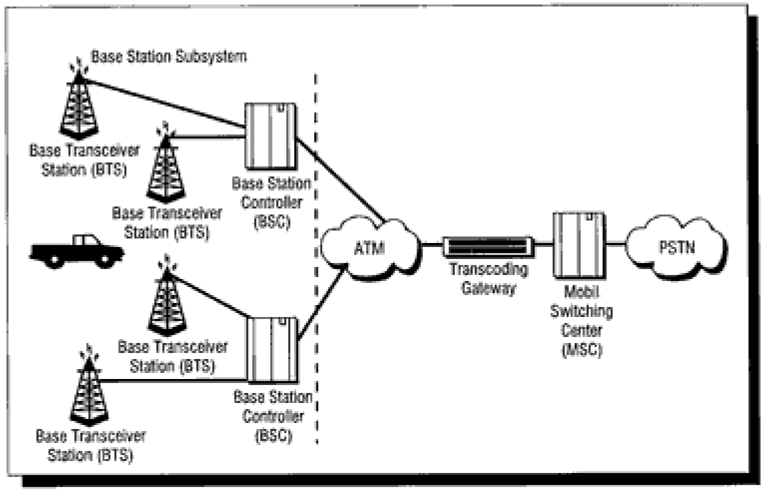

A second VoIP application, shown in *Figure 2*, is a trunking application. In this scenario, an organization wishes to send voice traffic between two locations over the packet network and replace the tie trunks used to connect the PBXs at the locations. This application usually requires the IWF to support a higher-capacity digital channel than the branch application, such as a T1/E1 interface of 1.544 or 2.048 Mbps. The IWF emulates the signaling functions of a PBX, resulting in significant savings to companies' communications costs.

**Figure 2. Interoffice Trunking Application**




A third application of VoIP software is interworking with cellular networks, as shown in *Figure 3*. The voice data in a digital cellular network is already compressed and packetized for transmission over the air by the cellular phone. Packet networks can then transmit the compressed cellular voice packet, saving a tremendous amount of bandwidth. The IWF provides the transcoding function required to convert the cellular voice data to the format required by the public switched telephone network (PSTN).

**Figure 3. Cellular Network Interworking Application**




# 2. VoIP QoS Issues

The advantages of reduced cost and bandwidth savings of carrying voice-over-packet networks are associated with some quality-of-service (QoS) issues unique to packet networks.

## Delay

Delay causes two problems: echo and talker overlap. Echo is caused by the signal reflections of the speaker's voice from the far-end telephone equipment back into the speaker's ear. Echo becomes a significant problem when the round-trip delay becomes greater than 50 milliseconds. As echo is perceived as a significant quality problem, voice-over-packet systems must address the need for echo control and implement some means of echo cancellation.

Talker overlap (or the problem of one talker stepping on the other talker's speech) becomes significant if the one-way delay becomes greater than 250 milliseconds. The end-to-end delay budget is therefore the major constraint and driving requirement for reducing delay through a packet network.

The following are sources of delay in an end-to-end, voice-over-packet call:

## Accumulation Delay (Sometimes Called Algorithmic Delay)

This delay is caused by the need to collect a frame of voice samples to be processed by the voice coder. It is related to the type of voice coder used and varies from a single sample time (.125 microseconds) to many milliseconds. A representative list of standard voice coders and their frame times follows:

- G.726 adaptive differential pulse-code modulation (ADPCM) (16, 24, 32, 40 kbps)—0.125 microseconds
- G.728 LD–code excited linear prediction (CELP)(16 kbps)—2.5 milliseconds
- G.729 CS–ACELP (8 kbps)—10 milliseconds
- G.723.1 Multirate Coder (5.3, 6.3 kbps)—30 milliseconds

## Processing Delay

This delay is caused by the actual process of encoding and collecting the encoded samples into a packet for transmission over the packet network. The encoding delay is a function of both the processor execution time and the type of algorithm used. Often, multiple voice-coder frames will be collected in a single packet to reduce the packet network overhead. For example, three frames of G.729 code words, equaling 30 milliseconds of speech, may be collected and packed into a single packet.

## Network Delay

This delay is caused by the physical medium and protocols used to transmit the voice data and by the buffers used to remove packet jitter on the receive side. Network delay is a function of the capacity of the links in the network and the processing that occurs as the packets transit the network. The jitter buffers add delay, which is used to remove the packet-delay variation to which each packet is subjected as it transits the packet network. This delay can be a significant part of the overall delay, as packet-delay variations can be as high as 70 to 100 milliseconds in some frame-relay and IP networks.

# Jitter

The delay problem is compounded by the need to remove jitter, a variable interpacket timing caused by the network a packet traverses. Removing jitter requires collecting packets and holding them long enough to allow the slowest packets to arrive in time to be played in the correct sequence. This causes additional delay.

The two conflicting goals of minimizing delay and removing jitter have engendered various schemes to adapt the jitter buffer size to match the time-varying requirements of network jitter removal. This adaptation has the explicit goal of minimizing the size and delay of the jitter buffer, while at the same time preventing buffer underflow caused by jitter.

Two approaches to adapting the jitter buffer size are detailed below. The approach selected will depend on the type of network the packets are traversing.

The first approach is to measure the variation of packet level in the jitter buffer over a period of time and incrementally adapt the buffer size to match the calculated jitter. This approach works best with networks that provide a consistent jitter performance over time, such as ATM networks.

The second approach is to count the number of packets that arrive late and create a ratio of these packets to the number of packets that are successfully processed. This ratio is then used to adjust the jitter buffer to target a predetermined, allowable late-packet ratio. This approach works best with the networks with highly variable packet-interarrival intervals—such as IP networks.

In addition to the techniques described, the network must be configured and managed to provide minimal delay and jitter, enabling a consistent QoS.

## Lost-Packet Compensation

Lost packets can be an even more severe problem, depending on the type of packet network that is being used. Because IP networks do not guarantee service, they will usually exhibit a much higher incidence of lost voice packets than ATM networks. In current IP networks, all voice frames are treated like data. Under peak loads and congestion, voice frames will be dropped equally with data frames. The data frames, however, are not time sensitive, and dropped packets can be appropriately corrected through the process of retransmission. Lost voice packets, however, cannot be dealt with in this manner.

Some schemes used by voice-over-packet software to address the problem of lost frames are as follows:

- interpolate for lost speech packets by replaying the last packet received during the interval when the lost packet was supposed to be played out; this scheme is a simple method that fills the time between noncontiguous speech frames; it works well when the incidence of lost frames is infrequent; it does not work well if there are a number of lost packets in a row or a burst of lost packets

- send redundant information at the expense of bandwidth utilization; this basic approach replicates and sends the nth packet of voice information along with the (n+1)th packet; this method has the advantage of being able to correct for the lost packet exactly; however, this approach uses more bandwidth and also creates greater delay

- use a hybrid approach with a much lower bandwidth voice coder to provide redundant information carried along in the (n+1)th packet; this reduces the problem of the extra bandwidth required but fails to solve the problem of delay

### Echo Compensation

Echo in a telephone network is caused by signal reflections generated by the hybrid circuit that converts between a four-wire circuit (a separate transmit and receive pair) and a two-wire circuit (a single transmit and receive pair). These reflections of the speaker's voice are heard in the speaker's ear. Echo is present even in a conventional circuit-switched telephone network. However, it is acceptable because the round-trip delays through the network are smaller than 50 milliseconds and the echo is masked by the normal side tone every telephone generates.

Echo becomes a problem in voice-over-packet networks because the round-trip delay through the network is almost always greater than 50 milliseconds. Thus, echo-cancellation techniques are always used. ITU standard G.165 defines performance requirements that are currently required for echo cancellers. The ITU is defining much more stringent performance requirements in the G.IEC specification.

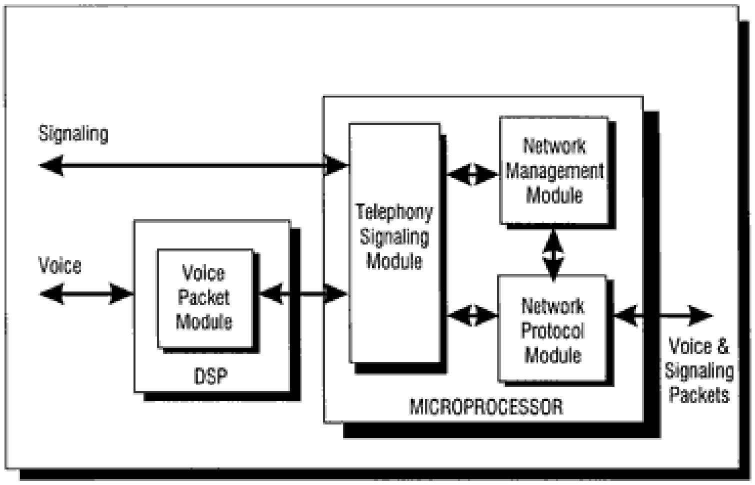
Echo is generated toward the packet network from the telephone network. The echo canceller compares the voice data received from the packet network with voice data being transmitted to the packet network. The echo from the telephone network hybrid is removed by a digital filter on the transmit path into the packet network.

## 3. VoIP–Embedded Software Architecture

Two major types of information must be handled to interface telephony equipment to a packet network: voice and signaling information.

As shown in *Figure 4*, VoIP software interfaces to both streams of information from the telephony network and converts them to a single stream of packets transmitted to the packet network. The software functions are divided into four general areas.

**Figure 4. VoIP Software Architecture**




## Voice Packet Software Module

This software, also known as the voice-processing module, typically runs on a digital-signal processor (DSP), prepares voice samples for transmission over the packet network. Its components perform echo cancellation, voice compression, voice-activity detection, jitter removal, clock synchronization, and voice packetization.

## Telephony-Signaling Gateway Software Module

This software interacts with the telephony equipment, translating signaling into state changes used by the packet protocol module to set up connections. These state changes are on-hook, off-hook, trunk seizure, etc. This software supports ear, mouth, earth, and magneto (E&M) Type I, II, III, IV, and V; loop or ground start foreign exchange station (FXS); foreign exchange office (FXO); and integrated services digital network (ISDN) basic rate interface (BRI) and primary rate interface (PRI).

## Packet Protocol Module

This module processes signaling information and converts it from the telephony-signaling protocols to the specific packet-signaling protocol used to set up connections over the packet network (e.g., Q.933 and voice-over-frame relay signaling). It also adds protocol headers to both voice and signaling packets before transmission into the packet network.

## Network-Management Module

This module provides the voice-management interface to configure and maintain the other modules of the voice-over-packet system. All management information is defined in American National Standards Institute (ANSI).1 and complies with signaling network-management protocol (SNMP) V1 syntax. A proprietary voice packet management information base (MIB) is supported until standards evolve in the forums.

The software is partitioned to provide a well-defined interface to the DSP software usable for multiple voice packet protocols and applications. The DSP processes voice data and passes voice packets to the microprocessor with generic voice headers.

The microprocessor is responsible for moving voice packets and adapting the generic voice headers to the specific voice packet protocol that is called for by the application, such as real-time protocol (RTP), voice over frame relay (VoFR), and voice telephony over ATM (VToA). The microprocessor also processes signaling information and converts it from supported telephony-signaling protocols to the packet network signaling protocol [e.g. H.323 IP, frame relay, or ATM signaling].

This partitioning provides a clean interface between the generic voice-processing functions, such as compression, echo cancellation, and voice-activity detection, and the application-specific signaling and voice protocol processing.

# 4. Voice Packet Module

This section describes the functions performed by the software in the voice packet module, also known as the voice-processing module, which is primarily responsible for processing the voice data. This function is usually performed in a DSP. The voice-processing module consists of the following software:

- **PCM interface**—This receives pulse code modulation (PCM) samples from the digital interface and forwards them to appropriate DSP software modules for processing, forwards processed PCM samples received from various DSP software modules to the digital interface, and performs continuous phase resampling of output samples to the digital interface to avoid sample slips.

- **tone generator**—This generates dual-tone multifrequency (DTMF) tones and call progress tones under command of the host (e.g., telephone, fax, modem, PBX, or telephone switch) and is configurable for support of U.S. and international tones.

- **echo canceller**—This performs G.165–compliant echo cancellation on sampled, full-duplex voice port signals. It has a programmable range of tail lengths.

- **voice activation detector/idle noise measurement**—This monitors the received signal for voice activity. When no activity is detected for the configured period of time, the software informs the packet voice protocol. This prevents the encoder output from being transported across the network when there is silence, resulting in additional bandwidth savings. This software also measures the idle noise characteristics of the telephony interface. It reports this information to the packet voice protocol to relay this information to the remote end for noise generation when no voice is present.

- **tone detector**—This detects the reception of DTMF tones and performs voice/fax discrimination. Detected tones are reported to the host so that the appropriate speech or fax functions are activated.

- **voice codec software**—This compresses the voice data for transmission over the packet data. It is capable of numerous compression ratios through the modular architecture. A compression ratio of 8:1 is achievable with the G.729 voice codec (thus, the normal 64–kbps PCM signal is transmitted using only 8 kbps).

- **fax software**—This performs a fax-relay function by demodulating PCM data, extracting the relevant information, and packing the fax-line scan data into frames for transmission over the packet network. Significant bandwidth savings can be achieved by this process.

- **voice playout unit**—This buffers voice packets received from the packet network and sends them to the voice codec for playout.

The following features are supported:

- a first in, first out (FIFO) buffer that stores voice code words before playout removes timing jitter from the incoming packet sequence

- a continuous-phase resampler that removes timing-frequency offset without causing packet slips or loss of data for voice- or voiceband-modem signals

- a timing jitter measurement that allows adaptive control of FIFO delay

The voice-packetization protocols use a sequence-number field in the transmit packet stream to maintain temporal integrity of voice during playout. Using this approach, the transmitter inserts the contents of a free-running, modulo-16 packet counter into each transmitted packet, allowing the receiver to detect lost packets and to reproduce silence intervals during playout properly.

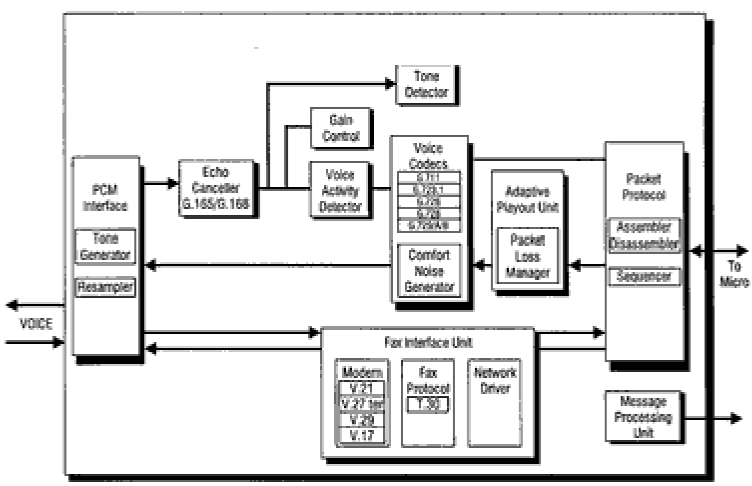
- **packet voice protocol**—This encapsulates compressed voice and fax data for end-to-end transmission over a backbone network between two ports.

- **control interface software**—This coordinates the exchange of monitor and control information between the DSP and host via a mailbox mechanism. Information exchanged includes software downline load, configuration data, and status reporting.

- **real-time portability environment**—This provides the operating environment for the software residing on the DSP. It provides synchronization functions, task management, memory management, and timer management.

*Figure 5* diagrams the architecture of the DSP software. The DSP software processes PCM samples from the telephony interface and converts them to a digital format suitable for transmission through a packet network.

**Figure 5. Voice Packet Module**




# 5. Signaling, Protocol and Management Modules

The VoIP software performs telephony signaling to detect the presence of a new call and to collect address (dial digit) information, which is used by the system to route a call to a destination port. It supports a wide variety of telephony-signaling protocols and can be adaptable to many environments. The software and configuration data for the voice card can be downloaded from a network-management system to allow customization, easy installation, and remote upgrades.

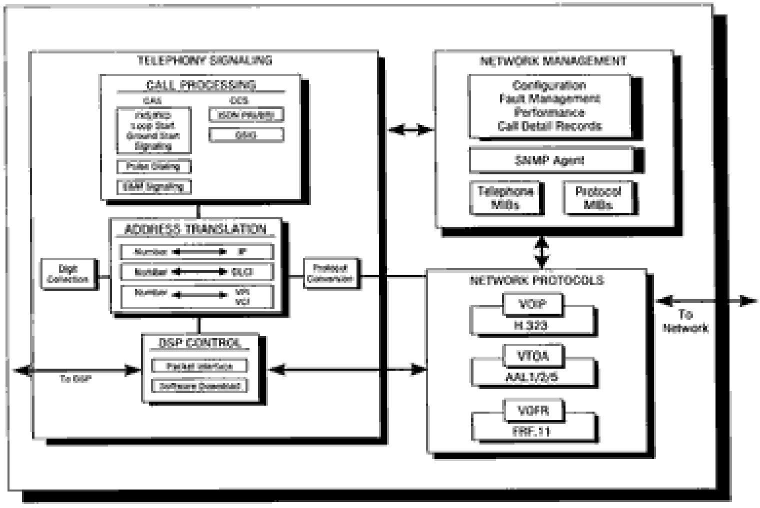
The software interacts with the DSP for tone detection and generation, as well as mode of operation control based on the line supervision, and interacts with the telephony interface for signaling functions. The software receives configuration data from the network-management agent and utilizes operating-system services.

## Telephony-Signaling Gateway Module

*Figure 6* diagrams the architecture of the signaling software, which consists of the following components:

- **telephony interface unit software**—This periodically monitors the signaling interfaces of the module and provides basic debouncing and rotary digit collection for the interface.
- **signaling protocol unit**—This contains the state machines implementing the various telephony-signaling protocols, such as E&M.
- **network control unit**—This maps telephony-signaling information into a format compatible with the packet voice session establishment signaling protocol.
- **address translation unit**—This maps the E.164 dial address to an address that can be used by the packet network (e.g., an IP address or a data-link connection identifier (DLCI) for a frame-relay network).
- **DSP interface driver**—This relays control information between the host microprocessor and DSPs.
- **DSP downline loader**—This is responsible for downline load of the DSPs at start-up, configuration update, or mode changes (e.g., switching from voice mode to fax mode when fax tones are detected).

**Figure 6. Signaling Modules**




## Network-Protocol Module

- **IP signaling stack**—This involves H.323 call control and transport software, including H.225, H.245, RTP/real-time conferencing protocol (RTCP) transport protocol, transmission control protocol (TCP), IP, and user datagram protocol (UDP).

- **ATM signaling protocol stack**—ATM Forum VT0A voice-encapsulation protocol. ATM Forum–compliant, user-network interface (UNI) signaling protocol stack for establishing, maintaining, and clearing point-to-point and point-to-multipoint switched virtual circuits (SVCs).

- **frame-relay protocol stack**—This includes Frame Relay Forum VoFR voice-encapsulation protocol, permanent virtual circuit (PVC) and SVC support, local management interface (LMI), congestion management, traffic monitoring, and committed information range (CIR) enforcement.

## Network-Management Module

The network-management software consists of three major services addressed in the MIB:

- physical interface to the telephone endpoint

- voice channel service for the following:

    - processing signaling on a voice channel

- converting between PCM samples and compressed voice packets

- call-control service for parsing call-control information and establishing calls between telephony endpoints

- The VoIP software is configured and maintained through the use of a proprietary voice service MIB.

# 6. VoIP Summary

A VoIP software architecture has been described for the interworking of legacy telephony systems and packet networks. Some of the key features enabling this application to function successfully are as follows:

- an approach that minimizes the effects of delay on voice quality

- an adaptive playout to minimize the effect of jitter

- features that address lost-packet compensation, clock synchronization, and echo cancellation

- a flexible DSP system architecture that manages multiple channels per single DSP

Carrying VoIP networks provides the most bandwidth-efficient method of integrating these divergent technologies. While the challenges to this integration are substantial, the potential savings make the investment in a quality implementation compelling.

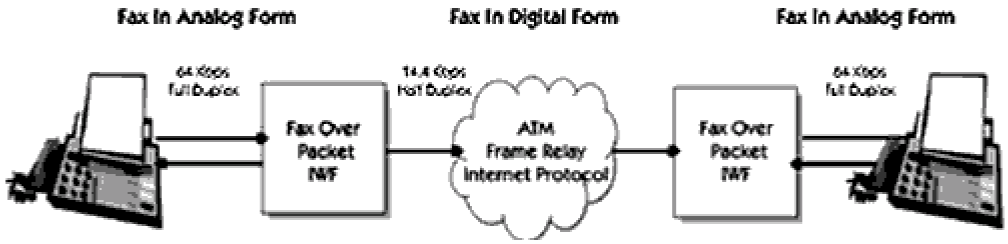
# 7. FoIP Applications

Traditionally, there have been two approaches for sending FoIP networks: real-time methods and store-and-forward methods. The primary difference in service between these two approaches lies in the delivery and method of receipt confirmation. The Frame Relay Forum has defined a real-time protocol for the transmission of fax-over–frame relay networks. Likewise, the ITU and Internet Engineering Task Force (IETF) are working together to continue to evolve both the real-time FoIP network standard (T.38) as well as the store-and-forward FoIP network standard (T.37). Both T.37 and T.38 were approved by the ITU in June, 1998. Furthermore, T.38 is the fax transmission protocol selected for H.323.

There are tremendous opportunities for cost savings by transmitting fax calls over packet networks. Fax data in its original form is digital. However, it is modulated and converted to analog for transmission over the PSTN. This analog form uses 64 kbps of bandwidth in both directions.

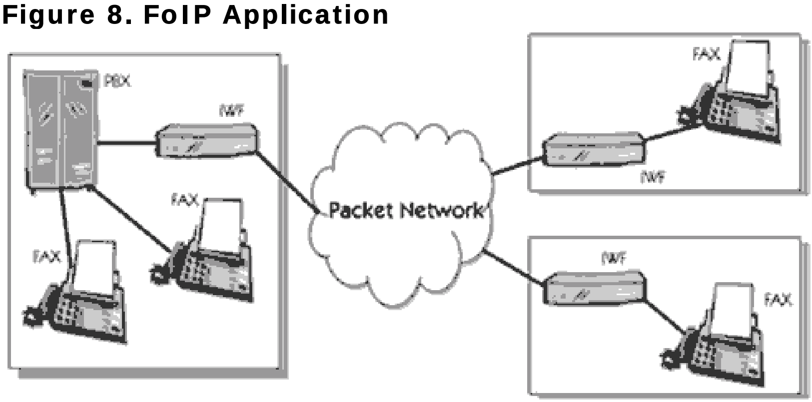
The FoIP IWF reverses this analog conversion, transmitting digital data over the packet network and then reconverting the digital data to analog for the receiving fax machine. This conversion process reduces the overall bandwidth required to send the fax, because the digital form is much more efficient, and the fax transmission is half-duplex (i.e., only one direction is used at any time). The peak rate for a fax transmission is 14.4 kbps in one direction. A representation of this process is shown in *Figure 7*.

**Figure 7. FoIP Conversion Process**




An application for fax over packet, shown in *Figure 8*, is a network configuration of a company with numerous branch offices that wants to use the packet network, instead of the long-distance network, to provide fax access to the main office. The IWF is the physical implementation of the hardware and software that enables the transmission of fax over the packet network. The IWF must support analog interfaces that directly interface to fax machines at the branches and to a PBX at the central site. The IWF must emulate the functions of a PBX for the fax machines.

**Figure 8. FoIP Application**




# 8. PSTN Fax-Call Procedure

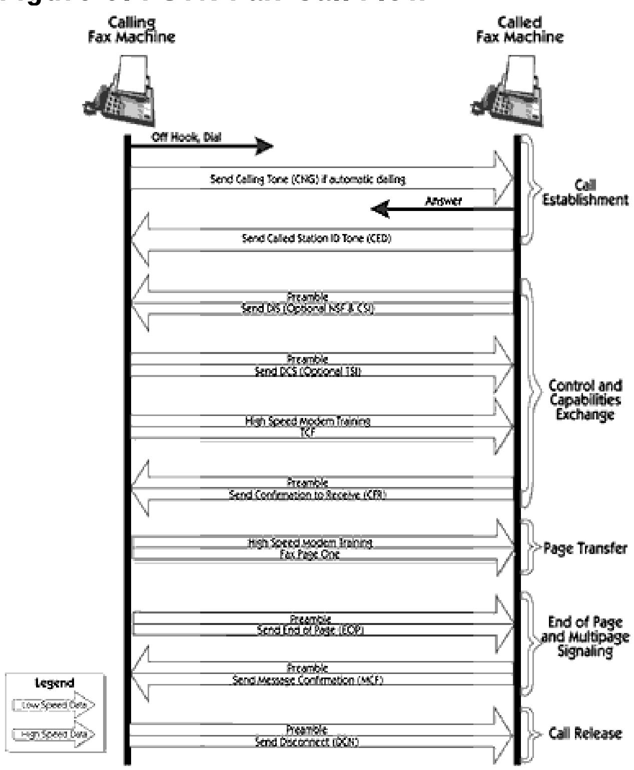
This module will describe the stages of a standard fax call over the PSTN so that the processing required for a reliable fax transmission over a packet network can be explored. Fax machines in common use today implement the ITU recommendations T.30 and T.4 protocols. The T.30 protocol describes the formatting of non-page data, such as messages that are used for capabilities negotiation. The T.4 protocol describes formatting of page image data.

T.30 and T.4 have evolved substantially over time and are now quite complex, given that they attempt to describe the behavior of an evolving set of fax machines. The timing related to the message interaction and phases of the call is critical and is one of the major causes of problems in the transmission of FoIP networks.

The PSTN fax call is divided into five phases, as shown in *Figure 9*. This example assumes that the call is accomplished without errors. The procedure becomes somewhat more complicated if errors occur or if there is a need for modem retraining. The five phases are as follows:

- call establishment
- control and capabilities exchange
- page transfer
- end of page and multipage signaling
- call release

**Figure 9. PSTN Fax-Call Flow**



## Call Establishment

The fax call is established either through a manual process, according to which someone dials a call and puts the machine into fax mode, or by automatic procedures, according to which no human interaction is required. In both cases, the answering fax machine returns an answer tone, called a CallEd station IDentification (CED), which is the high-pitched tone that you would hear when you call a fax machine. If the call is automatically dialed, the calling station will also indicate the fax call with a calling tone (CNG), which is a short, periodic tone that begins immediately after the number is dialed. These tones are generated to allow a human participant to realize that a machine is present on the other end call. They are also, although not reliably, used to recognize the presence of a fax call.

## Control and Capabilities Exchange

The control and capabilities exchange phase of the fax call is used to identify the capabilities of the fax machine at the other end of the call. It also negotiates the acceptable conditions for the call. The exchange of control messages throughout the fax call are sent using the low-speed (300 bps) modulation mode. Every control message is preceded by a one-second preamble, which allows the communication channel to be conditioned for reliable transmission.

The called fax machine begins the procedure by sending a digital identification signal (DIS) message, which contains the capabilities of the fax machine. An example of a capability that could be identified in this message is the support of the V.17 (14,000 bps) data signaling rate. At the same time, the called subscriber information (CSI) and nonstandard facilities (NSF) messages are optionally sent.

NSF are capabilities that a particular fax manufacturer has built into a fax machine to distinguish its product from others. These facilities are not required to be supported for interoperability.

Once the calling fax machine receives the DIS message, it determines the conditions for the call by examining its own capabilities table. The calling machine responds with the digital command signal (DCS), which defines the conditions of the call.

At this stage, high-speed modem training begins. The high-speed modem will be used in the next phase of the fax call to transfer page data. The calling fax machine sends a training check field (TCF) through the modulation system to verify the training and ensure that the channel is suitable for transmission at the accepted data rate. The called fax machine responds with a confirmation to receive (CFR), which indicates that all capabilities and the modulation speed have been confirmed and the fax page may be sent.

## Page Transfer

The high-speed modem is used to transmit the page data that has been scanned in and compressed. It uses the ITU T.4 protocol standard to format the page data for transmission over the channel.

## End-of-Page and Multipage Signaling

After the page has been successfully transmitted, the calling fax machine sends an end-of-procedures (EOP) message if the fax call is complete and all of the pages have been transmitted. If only one page has been sent and there are additional ones to follow, it sends a multipage signal (MPS). The called machine would respond with message confirmation (MCF) to indicate the message has been successfully received and that it is ready to receive more pages.

## Call Release

The release phase is the final phase of the call, in which the calling machine sends a disconnect message (DCN). While the DCN message is a positive indication that the fax call is over, it is not a reliable indication, as the fax machine can disconnect prematurely without ever sending the DCN message.

# 9. FoIP QoS

The advantages of reduced cost and bandwidth savings of carrying FoIP networks are associated with some QoS issues that are unique to packet networks and can affect the reliability of the fax transmission.

## Timing

A major issue in the implementation of FoIP networks is the problem of inaccurate timing of messages caused by delay through the network. The delay of fax packets through a packet network causes the precise timing that is required for many portions of the fax protocol to be skewed and can result in the loss of the call. The FoIP protocol in the IWF must compensate for the loss of a fixed timing of messages over the packet network so that the T.30 protocol operates without error.

There are two sources of delay in an end-to-end, FoIP call: network delay and processing delay.

- **network delay**—This is caused by the physical medium and protocols that are used to transmit the fax data and by buffers used to remove packet jitter on the receiving end. This delay is a function of the capacity of the links in the network and the processing that occurs as the packets transit the network. The jitter buffers add delay when they remove the packet-delay variation of each packet as it transits the packet network. This delay can be a significant part of the overall delay, as packet-delay variations can be as high as 70 to 100 milliseconds in some frame-relay networks and even higher in IP networks.

- **processing delay**—This is caused by the process of demodulating and collecting the digital fax information into a packet for transmission over the packet network. The encoding delay is a function of both the processor execution time and the amount of data collected before sending a packet to the network. Low-speed data, for instance, is usually sent out with a single byte per packet, as the time to collect a byte of information at 300 bps is 30 milliseconds.

## Jitter

Delay issues are compounded by the need to remove jitter, a variable interpacket timing caused by the network that a packet traverses. An approach to removing the jitter is to collect packets and hold them long enough so that the slowest packets to arrive are still in time to be played in the correct sequence. This approach, however, causes additional delay. In most FoIP protocols, a time stamp is incorporated in the packet to ensure that the information is played out at the proper instant.

## Lost-Packet Compensation

Lost packets can be an even more severe problem, depending on the type of packet network that is being used. In a VoIP application, the loss of packets can be addressed by replaying last packets and other methods of interpolation. A FoIP application, however, has more severe constraints on the loss of data, as the fax protocol can fail if information is lost. This problem varies, depending on the type of fax machine used and whether error-correction mode is enabled.

Two schemes that are used by FoIP software to address the problems of lost frames are as follows:

- repeating information in subsequent frames so that the error can be corrected by the receiver's playout mechanism
- using an error-correcting protocol such as TCP to transport the fax data at the expense of added delay

# 10. FoIP Software Architecture

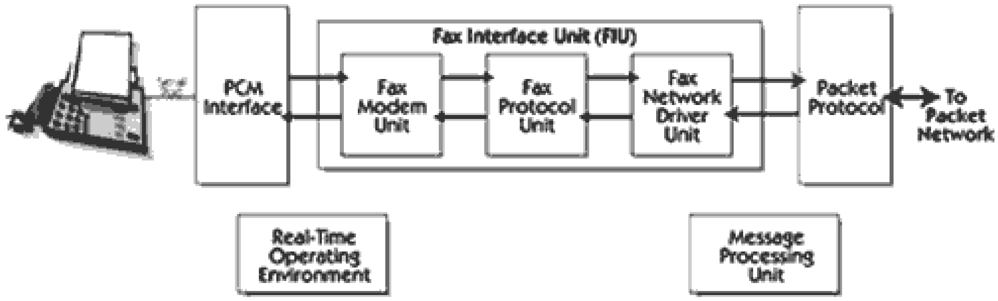
The facsimile interface unit (FIU) is the software module that resides within a FoIP IWF. It demodulates voiceband signals from an analog interface and converts them to a digital format suitable for transport over a packet network. It also remodulates data received from the packet network and transmits it to the analog interface. In doing so, the FIU performs protocol conversion between Group-3 facsimile protocols and the digital facsimile protocol employed over the packet network.

The FIU, shown in *Figure 10*, consists of the following three units:

**fax-modem unit (FM)**—This processes PCM samples based on the current modulation mode and supports the following functions:

- V.21 Channel 2 (300 bps) binary signaling modulation and demodulation
- high-level data link control (HDLC) framing (0 bit insertion/removal, cyclic redundancy check (CRC) generation/checking)
- V.27 ter (2400/4800 bps) high-speed data modulation and demodulation
- V.29 (7200/9600 bps) high-speed data modulation and demodulation
- V.17 (7200/9600/12000/14400 bps) high-speed data modulation and demodulation
- V.33 (12000/14400 bps) high-speed modulation and demodulation
- CED detection and generation
- CNG detection and generation
- V.21 Channel-2 detection

**Figure 10. FoIP Module**




**fax protocol unit (FP)**—This compensates for the effects of timing and lost packets caused by the packet network. The FP prevents the local fax machine from timing out while waiting for a response from the other end by generating HDLC flags. If, after a time out, the response from the remote fax machine is not received, it also sends a command repeat (CRP) frame to resend the frame. This unit monitors the facsimile transaction timing, the direction of current transmission, and the proper modem configuration. It performs the following functions:

- protocol processing (group-3 facsimile)
- examination/alteration of binary signaling messages to ensure compatibility of the facsimile transfer with the constraints of the transmission channel
- network channel interface data formatting
- line state transitions

**fax network driver unit (FND)**—This assembles and disassembles fax packets to be transmitted over the network and is the interface unit between the FP and network modules. The control information packets consist of header and time stamp information. In the direction of the PCM to the packet network, the FND collects the specified number of bytes and transmits the packet to the network. In the receive direction, the FND provides data with the proper timing (as generated on the transmit side and reproduced through the received time stamp information) to the rest of the FIU. The FND formats the network packets for transmission to the network based on the specific network protocol. The FND delays the data to remove timing jitter from the packet arrival times and performs the following functions:

- formatting of control information
- formatting of fax data
- properly timed playout of data

- elastic (slip) buffering
- lost-packet compensation

# 11. FoIP Summary

A fax-over-packet software architecture has been described for the interworking of fax machines and packet networks. Some of the key features enabling these applications to function successfully are as follows:

- an approach that addresses the effect of delay through the network
- a process that minimizes the effect of jitter
- features that address lost-packet compensation

Though the QoS issues associated with carrying FoIP networks are significant, the future of this approach will be driven by the substantial cost savings and exciting applications made possible with FoIP software technology.

# Self-Test

1. The consolidation of separate voice and data networks offers an opportunity for _____________.

    a. utilization of extra broadband bandwidth for voice and data transmission

    b. reduced delay over a telephone call

    c. reduced computer and telephone hardware requirements

2. Voice-over-packet technology may be used to transfer information over both broadband and wireless networks.

    a. true

    b. false

3. A QoS issue unique to packet networks is ________________.

    a. interworking

    b. compression

    c. jitter

4. Signal reflections generated by the circuit that converts between a four-wire and a two-wire circuit can result in ____________________.

   a. jitter

   b. echo

   c. delay

5. Developers seeking to incorporate voice-over-packet technology face which of the following challenges?

   a. still-evolving technical standards

   b. network phenomena such as delay, jitter, echo, and lost packets

   c. integration of incompatible technologies

   d. all of the above

6. The fax-over-packet IWF reduces overall bandwidth because the fax transmission is ________________.

   a. half-duplex

   b. real time

   c. low-speed data

   d. only b and c

   e. all of the above

7. A major cause of problems in the transmission of fax-over-packet networks is _____________.

   a. properly dividing the four phases of the call

   b. incompatible fax machines

   c. using T.30 and T.4 protocols

   d. the timing related to the message interaction

   e. all of the above

8. Fax-over-packet software addresses the problems of lost frames by _______________.

   a. using an error-correcting protocol, such as TCP, to transport the fax data at the expense of added delay

   b. replaying last packets and other methods of interpolation

   c. repeating information in subsequent frames so that the error can be corrected by the receiver's playout mechanism.

   d. only a and c

   e. all of the above

9. Processing delay is caused by the process of modulating and correcting the digital fax information into a packet for transmission over the packet network.

   a. true

   b. false

10. The fax network driver delays the data and performs which of the following?

   a. lost-packet compensation

   b. formatting of fax data

   c. removal of timing jitter from fax arrival times

   d. only a and b

   e. all of the above

# Correct Answers

1. The consolidation of separate voice and data networks offers an opportunity for _____________.

   **a. utilization of extra broadband bandwidth for voice and data transmission**

   b. reduced delay over a telephone call

   c. reduced computer and telephone hardware requirements

   See Topic 1.

2. Voice-over-packet technology may be used to transfer information over both broadband and wireless networks.

   **a. true**

   b. false

   See Topic 1.

3. A QoS issue unique to packet networks is ________________.

   a. interworking

   b. compression

   **c. jitter**

   See Topic 2.

4. Signal reflections generated by the circuit that converts between a four-wire and a two-wire circuit can result in _____________________.

   a. jitter

   **b. echo**

   c. delay

   See Topic 2.

5. Developers seeking to incorporate voice-over-packet technology face which of the following challenges?

   a. still-evolving technical standards

   b. network phenomena such as delay, jitter, echo, and lost packets

   c. integration of incompatible technologies

   **d. all of the above**

   See Topic 2.

6. The fax-over-packet IWF reduces overall bandwidth because the fax transmission is ________________.

   **a. half-duplex**

b. real time

c. low-speed data

d. only b and c

e. all of the above

See Topic 7.

7. A major cause of problems in the transmission of fax-over-packet networks is _____________.

a. properly dividing the four phases of the call

b. incompatible fax machines

c. using T.30 and T.4 protocols

**d. the timing related to the message interaction**

e. all of the above

See Topic 9.

8. Fax-over-packet software addresses the problems of lost frames by _______________.

a. using an error-correcting protocol, such as TCP, to transport the fax data at the expense of added delay

b. replaying last packets and other methods of interpolation

c. repeating information in subsequent frames so that the error can be corrected by the receiver's playout mechanism.

**d. only a and c**

e. all of the above

See Topic 9.

9. Processing delay is caused by the process of modulating and correcting the digital fax information into a packet for transmission over the packet network.

a. true

**b. false**

See Topic 9.

10. The fax network driver delays the data and performs which of the following?

a. lost-packet compensation

b. formatting of fax data

c. removal of timing jitter from fax arrival times

d. only a and b

**e. all of the above**

See Topic 10.

# Glossary

**ADPCM**
adaptive pulse-code modulation

**ANSI**
American National Standards Institute

**ATM**
asynchronous transfer mode

**BRI**
basic rate interface

**CED**
CallED station Identification

**CELP**
code excited linear prediction

**CFR**
confirmation to receive

**CIR**
committed information range

**CNG**
calling tone

**CRC**
cyclic redundancy check

**CRP**
command repeat

**CSI**
called subscriber information

**DCN**
disconnect message

**DCS**
digital command signal

**DIS**
digital identification signal

**DLCI**
data-link connection identifier

**DSP**
digital signal processor

**DTMF**
dual-tone multifrequency

**E&M**
ear & mouth, earth, and magneto

**EOP**
end of procedures

**FIFO**
first in, first out

**FIU**
facsimile interface unit

**FM**
fax-modem unit

**FND**
fax network driver unit

**FoIP**
fax over Internet protocol

**FP**
fax-protocol unit

**FXO**
foreign exchange office

**FXS**
foreign exchange station

**HDLC**
high-level data link control

**IETF**
Internet Engineering Task Force

**IP**
Internet protocol

**ISDN**
integrated services digital network

**ITU**
International Telecommunications Union

**IWF**
interworking function

**LMI**
local management interface

**MCF**
message confirmation

**MIB**
management information base

**MPS**
multipage signal

**NSF**
nonstandard facilities

**PBX**
private branch exchange

**PCM**
pulse code modulation

**POTS**
plain old telephone service

**PRI**
primary rate interface

**PSTN**
public switched telephone network

**PVC**
permanent virtual circuit

**QoS**
quality of service

**RTCP**
real-time conferencing protocol

**RTP**
real-time protocol

**SVC**
switched virtual circuit

**TCF**
training check field

**TCP**
transmission control protocol

**UDP**
user datagram protocol

**UNI**
user network interface

**V/FoIP**
voice and fax over Internet protocol

**VoFR**
voice over frame relay

**VoIP**
voice over Internet protocol

**VToA**
voice telephony over ATM