

HP TeleShare: Integrating Telephone Capabilities on a Computer Workstation

Using off-the-shelf parts and a special interface ASIC, an I/O card was developed that provides voice, fax, and data transfer via a telephone line for the HP 9000 Model 712 workstation.

by S. Paul Tucker

Integration of the telephone and the computer workstation is a natural step in the evolution of the electronic office. It allows the user to perform telephone transactions without having to change from the keyboard and mouse environment to the telephone and handset environment and vice-versa. This capability provides obvious benefits to a wide customer audience, especially those dealing with customer service and support. The HP TeleShare option card for the HP 9000 Model 712 workstation represents HP's first integrated telephony product. Coupled with multimedia technologies such as audio, video, and HP SharedX,¹ HP TeleShare provides the user with a powerful arsenal of communications tools. This article will focus mainly on the hardware aspects of the HP TeleShare product.

Features provided by HP TeleShare include:

- Two-line support, with each line configurable for voice, fax, or data
- Workstation audio support and mixing (stereo headset with built-in microphone included)
- Dual-tone multifrequency (DTMF) tone generation and detection
- Telephone line status and control
- Call progress support
- Caller-ID support
- V.32bis modem (14,400 bits/s) with V.42bis and MNP5 (Microcom Networking Protocol) compression and error correction
- Fax Group 3 Class II up to 14,400 bits/s.

Background

HP TeleShare began as an experimental interface card for the HP 9000 Series 300 workstations. It had simple voice-only telephone capabilities, including single-source audio record and playback, and it was perceived as useful (and entertaining) to those engineers who were fortunate enough to have the opportunity to use it. At some point, further investigation was needed and the HP TeleShare project team was formed. It was determined that fax and data modem capabilities were needed with close coupling to the workstation's audio capability. Dual telephone lines were included so the user could talk on one line and at the same time use the other line for faxing or data. The standard analog phone line interface was chosen over digital (i.e., ISDN) because of

the relatively insignificant number of digitally equipped PBX systems.

The first incarnation of the current product was an external RS-232-driven box with stereo inputs for computer, line-in, and microphone audio, and stereo outputs for computer, line-out, and headphones. It provided dual-line operation and employed two DSP (digital signal processor) subsystems for maximum flexibility and performance in voice and data modes. Audio mixing was provided by dedicated analog hardware, and any combination of audio inputs could be sent to any output complete with treble and bass control. The audio capabilities were so good that HP TeleShare engineers always had their CD players plugged into the box and their headphones on. This forced development of an automatic audio mute feature when an incoming call was detected. Since modem functionality was a primary goal, the command interface for the box used a partial modem AT command set, along with some proprietary extensions for new functionality like setting audio gains, setting audio mix values, telling a DSP to reboot, and so on. These commands were delivered over the RS-232 interface and received the typical OK and error responses.

While the external box was well received in the lab and by customers, it was postponed indefinitely in favor of a lower-cost internal version with a proprietary interface available on a single workstation, the Model 712. The same DSP subsystem used in the external box was carried over to the Model 712 option card and some effort was exerted to leverage as much as possible of the external box's software interface and feature set into the new design.

Architecture

HP TeleShare is made up of two independent DSP subsystems that communicate with the workstation host through an interface chip called *XBAR* (see Fig. 1). Each DSP is coupled to a hybrid chip called a *data access arrangement*, which provides direct connection to a standard two-wire analog telephone line. HP TeleShare is tightly coupled with the workstation audio system to provide the highest degree of audio flexibility. For instance, line-in audio (perhaps from a CD player) could be sent to telephone line 0 while the party on that line is on hold. Simultaneously, the workstation

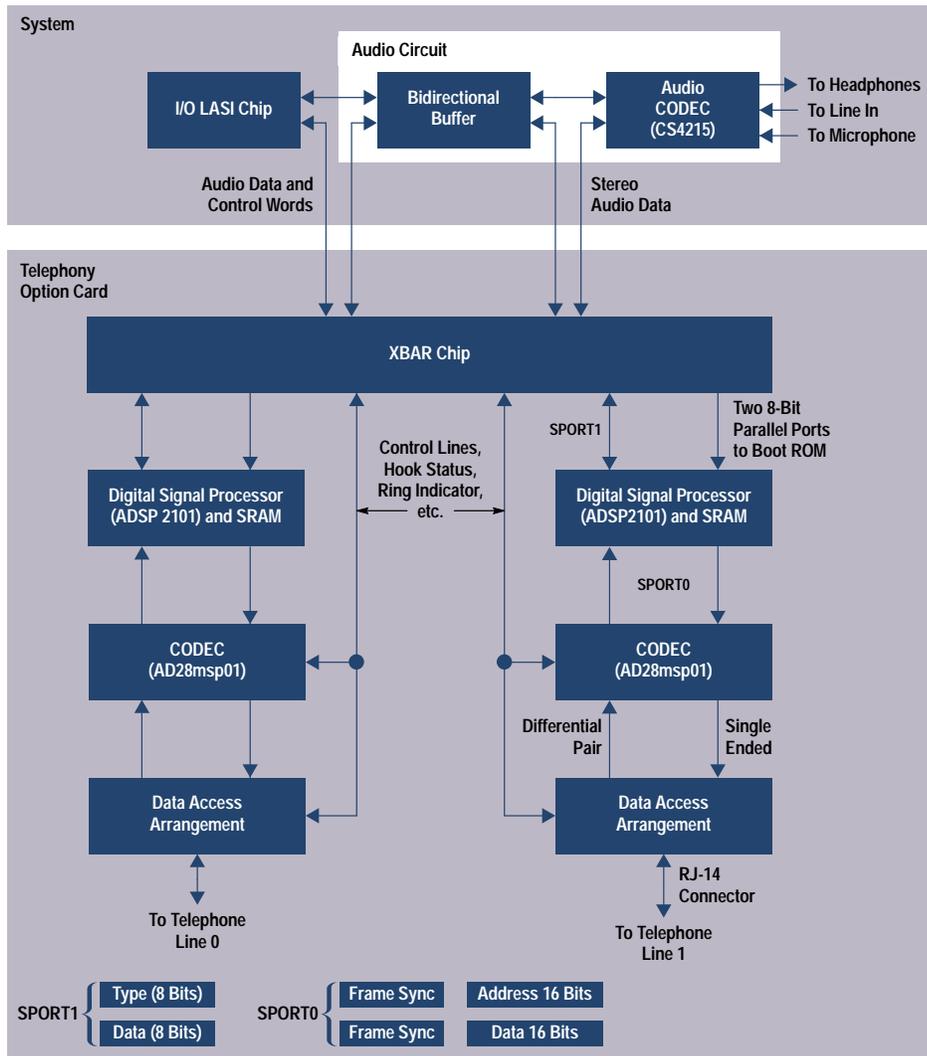


Fig. 1. Block diagram of HP TeleShare.

user could be conversing with or faxing a message to the party on telephone line 1. During this time comments from the party on line 0 could be recorded to disk for later playback.

XBAR. The XBAR ASIC (application-specific integrated circuit) is a custom VLSI part packaged in a low-cost 80-pin QFP (quad flat pack). It was designed by the HP TeleShare team specifically for use with the Model 712 workstation and performs all of the interface functions required by the HP TeleShare card. The XBAR chip communicates with the system I/O chip (LASI) and the audio CODEC (coder/decoder) through a pair of proprietary serial interfaces. If HP TeleShare is not present in the system, audio data and CODEC control words pass through the bidirectional buffer between the system I/O chip and the audio CODEC. When the HP TeleShare card is installed, XBAR is effectively placed between the system I/O chip and the CODEC, forcing all audio to be routed through XBAR in either direction. The serial interface between XBAR and the audio CODEC carries 16-bit stereo audio data.

The serial interface between XBAR and the I/O chip multiplexes 16-bit system audio data (to and from disk) and control words for XBAR. In addition, this interface is used for modem data, voice-mode AT commands and responses, and DSP application code downloaded from the host system. On

the DSP side, XBAR has two 13.824-MHz serial ports, each designed specifically for interfacing with the DSPs. These ports are used for passing audio samples, modem data, application programs, and commands and responses to and from the DSPs.

XBAR's configuration can be changed by writing to the control registers in the XBAR-to-LASI I/O serial interface address space. In voice mode, XBAR is configured to pass each audio data sample from the LASI I/O chip and CODEC (coder/decoder) to a DSP, whereupon the DSP will return responses to each of those sources. XBAR can also send audio data samples from DSP to DSP for conferencing between lines when both lines are in voice mode. In data and fax modes, XBAR sends appropriately formatted data to the DSP and receives similar data in return.

Although XBAR supports stereo audio at up to a 48-kHz sample rate, DSP bandwidth limitations require all audio data to and from the telephone lines to be left-channel only, sampled at 8 kHz. This is not a serious limitation, since telephone-quality audio only requires a sample rate around 7.2 kHz for full reproduction and is inherently a single-channel signal.

In addition to the DSP serial ports, XBAR also has a pair of byte-wide parallel ports that connect to the DSPs' boot ROM

ports. This allows DSP boot code (not to be confused with DSP application code) to be downloaded from the host system. This provides additional flexibility and eliminates the cost and board-space limitations associated with external ROMs.

XBAR has several asynchronous control signals that are connected to downstream hardware, including reset lines for the DSP subsystems and hook control and telephone status (such as the ring indicator) signals to and from the data access arrangement chips.

The biggest challenge in XBAR's design was purely logistical because we had a lot of different data types (e.g., stereo data and telephone command and status data) to handle and very little time to implement them in XBAR. There are no less than 52 separate data types that XBAR must recognize and generate for the two DSP serial interfaces alone, with a slightly smaller number required for the system I/O interface. To provide these types, each transaction between XBAR and a DSP consists of 16 bits, with the upper eight bits providing type information and the lower eight bits providing the associated data. To prevent data overruns, XBAR requires a data acknowledge (Ack) word back from the appropriate DSP for every transaction.

Audio data samples in HP TeleShare are 16 bits long. Since XBAR sends eight bits of data at a time, audio samples must be broken into two pieces: an upper half, or most-significant byte (MSB) and a lower half, or least-significant byte (LSB). Using this model, it requires two transfers from XBAR and two Acks back to send one sample of audio data to a DSP. Sending one set of stereo system and line-in audio samples to a DSP requires eight output transfers (four transfers for the system sample and four transfers for the line-in sample), with an Ack back after each transfer. The DSP will then send mixed audio samples back for the system and the CODEC, requiring an additional eight transfers, for a total of 24 transfers per sample. This has to happen at an 8-kHz sample rate (once every 125 microseconds). Fortunately, XBAR can handle these transactions, but order must be maintained exactly or audio quality will suffer. Other data types, such as AT commands and responses, are given lower priority during audio frames and are queued until audio transfer is finished.

Digital Signal Processor. The DSP used by TeleShare is an Analog Devices ADSP2101. This is a programmable single-chip microcomputer optimized for digital signal processing, and operates at 16.67 MHz. The 2101 operates on 16-bit data and uses a 24-bit instruction word. It has 1024 words of data RAM and 2048 words of program RAM on the chip. The part has two data address generators and a program sequencer, which allows program and data accesses to occur simultaneously in a single cycle. Dual data operand fetches can also occur in a single cycle since program memory can also be used to store data. The part can address up to 16K words of data and 16K words of program memory, both of which are supplied on the HP TeleShare board in the form of six external SRAMs.

The DSP has two independent serial ports, SPORT0 and SPORT1, which support multiple data formats and frame rates and are fully programmable. In the HP TeleShare design, SPORT1 on each DSP is dedicated to communication with

XBAR, while SPORT0 is dedicated to communication with an AD28msp01 telephone line CODEC. Each full transfer to or from one of the SPORT lines triggers an associated interrupt in the 2101, allowing programs to act on the incoming data as it arrives.

Analog Devices supplies a complete set of software development tools for the 2100 microprocessor family, including a C compiler with a DSP function library.

CODEC. The Analog Devices AD28msp01 provides HP TeleShare with a multiple-sample-rate CODEC specifically designed for use in modem designs. This device supports sample rates of 7.2 kHz, 8.0 kHz, and 9.6 kHz, and has an 8/7 mode* for sampling at 8.23 kHz, 9.14 kHz and 10.97 kHz. For voice mode operation, 7.2 kHz and 8.0 kHz are all that is required, but the sophisticated algorithms used by modern modem standards often require the other rates. The CODEC uses 16-bit sigma-delta conversion technology and includes resampling and interpolation filtering along with transmit and receive phase adjustments.

Each CODEC has one serial port which is connected directly to SPORT0 on the associated DSP. This port operates in free-running mode once it is properly initialized and continually sends 16-bit data samples from the telephone line to the DSP. All transfers to the DSP consist of a serial output frame sync followed by a 16-bit address word, then a second frame sync followed by a 16-bit data word. These address and data pairs are transmitted at the selected sample rate and trigger SPORT0 receive interrupts in the DSP. The DSP transfers data to the CODEC using the same mechanism as just described (in the other direction, of course). The address portion of each transfer coming from the DSP identifies the data as a control word (for programming the part) or as a data word to be sent through the on-chip digital-to-analog converter (DAC) and transmitted to the data access arrangement chip. Data from the CODEC to the DSP is identified as either a control word or as a data word from the on-chip analog-to-digital converter (ADC).

The AD28msp01 CODEC is attached to the telephone line through the data access arrangement chip. Transmit data outputs are differential for noise reduction, while the receive data input is single-ended.

Data Access Arrangement. HP TeleShare uses the TDK 73M9002 data access arrangement chip as its telephone line interface. This part provides all the necessary line monitoring, filtering, isolation, protection, and signal conversion functions for connection of high-performance analog modem designs to the PSTN (public switched telephone network) in the United States, Canada, and Japan. The 73M9002 incorporates, on a two-to-four-wire hybrid, ring detection circuitry, off-hook relay, and on-hook line monitoring for caller-ID support (see "Caller-ID" on page 72).

The 73M9002 comes with FCC (Federal Communications Commission) part 68 DOC CS-03 and JATE (Japan Approvals Institute for Telecommunications Equipment) protection circuitry built in, and is compliant with UL 1459 2nd Edition

* The 8/7 mode is a capability required by some modem applications. It simply adds some sampling bandwidth. For example, in 8/7 mode the normal 8-kHz sample rate becomes 9.14 kHz ($8 \times 8/7$).

Caller-ID

Caller-ID information is sent between the first and second power ringing signals. The data is sent a minimum of 500 milliseconds after the first ring and ends at least 200 milliseconds before the second ring begins. This leaves 2.9 to 3.7 seconds of time for data transmission. The data is sent at 1200 baud using frequency shift keying (FSK) modulation. All data is 8-bit ASCII.

Two standard formats exist for Caller-ID information: single message format and multiple message format. In general, both formats can be described using Fig. 1.

The message type is 0x4 (hexadecimal 4) for single message format. The message length is variable and indicates the number of message words in the message body. The final word is a checksum word, used for error checking. Single message format provides the receiver with date, time, and calling number data.

The message type is 0x80 (hexadecimal 80) for multiple message format. The message length is variable as before, but provides the receiver with date, time, calling number, and calling name data if available. In the absence of calling name data, a P indicating private or an O indicating out of area or unavailable will be sent.

Caller-ID detection requires on-hook line monitoring, which the HP TeleShare data access arrangement chip fully supports. HP TeleShare can detect and display both message formats.

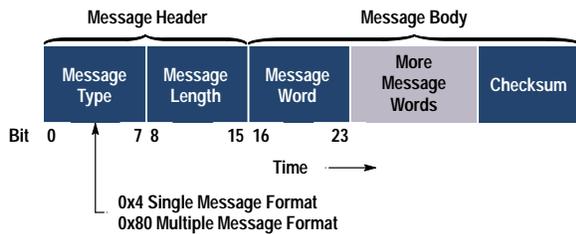


Fig. 1. Caller-ID message format.

with the addition of an external slow-blow fuse. The off-hook relay is controlled by a TTL-level input from XBAR. This relay determines when the phone is off the hook and can be pulsed for use as a pulse dialer. Another TTL-level input from XBAR is used to enable on-hook line monitoring.

The ring detection circuitry is capable of detecting ringing signals that comply with Ringing Type B from the FCC Part 68 regulations. The detected ring signal appears at a pair of differential outputs which are also connected to XBAR.

The 73M9002 provides telephone connectivity to the DSP subsystem through the CODEC's analog receive and transmit lines and is attached to the telephone line through a standard RJ-14 connector.

Operating Modes

HP TeleShare is capable of operating in three modes: voice, fax, and data. The modes are selected through a graphical user interface by the workstation user. The mode application software is downloaded through XBAR to a DSP as needed and runs continually until a reset of that DSP is performed. The voice mode code was developed solely by HP and can be run on one or both lines simultaneously. The fax and data modem code was developed with a third party and because of licensing restrictions, only one line can be configured as a fax or data modem at a time. Combinations of voice and fax or data are fully supported.

Voice Mode Operation. When configured in the voice mode, HP TeleShare essentially operates like an enhanced telephone. Digital mixing of microphone, line-in, telephone, and recorded audio (from system disk) is supported for both playback and recording. This capability allows numerous interesting audio configurations including placing a line on hold with music, recording conversations, playing back recorded audio over the phone, and so on. While in voice mode, HP TeleShare provides the user with caller-ID information if it is available. In addition, DTMF (dual-tone multi-frequency) tone and pulse dialing are supported, along with DTMF tone detection for unattended phone functions like answering machines or voicemail (see "Call Progress, DTMF Tones, and Tone Detection" on page 73).

Dialing and hook manipulation actions are performed through the GUI (graphical user interface), but at the lowest level these actions are sent to the DSP as standard AT commands like ATDT (attention dial tone) and ATH n (n = 0 is on-hook or hang up, and n = 1 is take telephone off hook). Special functions like audio mixing are also controlled with low-level AT-type commands, but are manipulated using sliders in the GUI.

The voice mode application firmware is driven primarily by DSP SPORT interrupts. Every incoming 16-bit SPORT0 word from XBAR triggers an interrupt, which in turn causes the SPORT1 interrupt service routine to execute. Likewise, every 16-bit SPORT0 word from the CODEC causes the SPORT0 interrupt service routine to execute. The SPORT1 interrupt service routine is responsible for audio I/O with XBAR and queuing AT commands as they arrive. Commands arrive asynchronously, that is, they can arrive at any time, while audio arrives in 8-piece bundles every 125 microseconds (one frame) as described earlier. Normally, every piece of data received by SPORT1 causes an interrupt, but the firmware disables these interrupts for the rest of a frame once it recognizes the first piece of audio data. Otherwise, at least eight context switches would occur every frame, which would render the system useless. Once the SPORT1 interrupt service routine has received all of the audio samples, it is responsible for transmitting the new audio back to XBAR for routing to the workstation (i.e., headphones and/or disk).

The SPORT0 interrupt service routine is responsible for receiving and transmitting telephone-line audio and mixing all audio data, including DTMF tones. Before mixing can occur in the DSP, all of the LSBs must be appended to the MSBs. Remember that each 16-bit sample transferred between XBAR and the DSP is divided so that the most-significant byte contains the data type and the least-significant byte contains the data. Thus, all the data from XBAR is put back into 16-bit linear format before transfer to the CODEC.

The audio input and output amplitude matrices, built by the user via the GUI, are used to determine what the final mix will sound like. The DSP firmware processes each output in sequence by adding together any inputs that are on to create a total value for each output. Any gain adjustments are made at this time as well. When this is completed for all outputs, the resulting 16-bit values are broken into MSBs and LSBs, if required.

Audio data that is meant for XBAR is transmitted during the next XBAR audio frame. Audio data meant for output to the

Call Progress, DTMF Tones, and Tone Detection

HP TeleShare's voice mode firmware has the ability to detect a number of tones used commonly in telephone communications, including DTMF tones and call progress tones like busy, ringback (the ringing sound you hear when you call someone), and dial tone.

DTMF Tones

Dual-tone multifrequency (DTMF) tones are made up of two separate tones, as the name suggests, and can be accurately generated using easily understood principles. The DTMF standard specifies two sets of distinct tones, called row frequencies and column frequencies (see Fig. 1). The row frequencies correspond to the horizontal rows on a standard telephone touchpad. The column frequencies correspond to the vertical columns on the touchpad, plus an additional column to the right of the last touchpad column.

This makes eight separate frequencies, which combine for a total of sixteen DTMF tones (see Fig. 2).

Generation of a DTMF tone is accomplished by creating a sinusoid for each of the two frequencies, row and column, and then adding the results. In a digital implementation, the sinusoids are computed and added on a sample-by-sample basis. HP TeleShare uses a five-coefficient Taylor series approximation for the sinusoid generation. The sinusoid samples are updated and added at 8 kHz, or every 125 microseconds, and the sum of the sinusoid samples is used as the current DTMF sample.

Tone Detection

Tone detection is accomplished through the use of a 512-point fast Fourier transform (FFT), which is implemented in the ADSP2101 C-language run-time library. The FFT, when given a set of samples of an input signal over some time interval, returns the frequency spectrum of the signal during the interval. This can be done in almost real time with a DSP, making it very useful for detecting incoming tones. The following important rules and relationships should be noted concerning sample rate, input points, output points, time, frequency, and the FFT in general:

- The FFT requires complex (real and imaginary) data for input (two arrays).
- The imaginary input array may be filled with zeros if unused.
- The output data is complex (two arrays).
- The frequency spectrum returned covers half of the sampling frequency.
- Only the first half of output data is used, and the other half is a mirror image.
- The output frequency resolution is equal to (sampling rate)/(number of input points).

Using an 8-kHz sampling rate and 512 points causes the FFT to return a spectrum from 0 to 4 kHz, with 512 complex output points. The second 256 output points can be ignored since they are the mirror image of the first 256. The output will have a resolution of 15.625 Hz per point, using the formula above. These output points will be referred to as bins since they include spectral data on either side of each point.

HP TeleShare calculates magnitude-squared values for each bin by squaring the real and imaginary values at each point and adding them. The magnitude-squared

		Column Frequencies			
		1209 Hz	1336 Hz	1477 Hz	1633 Hz
Row Frequencies	697 Hz	1	2	3	A
	770 Hz	4	5	6	B
	852 Hz	7	8	9	C
	941 Hz	*	0	#	D

Fig. 1. Dual-tone-multifrequency digits and the frequencies associated with them.

		Tone	Frequency (Hz)	512-Point FFT Index
Call Progress Tones	}	Dial Tone	350 + 440	22.4 + 28.16
		Busy	480 + 620	30.72 + 39.68
		Ringback	440 + 480	28.16 + 30.72
		DTMF 1	697 + 1209	44.61 + 77.38
		DTMF 2	697 + 1336	44.61 + 85.5
		DTMF 3	697 + 1477	44.61 + 94.53
		DTMF 4	770 + 1209	49.28 + 77.38
16 DTMF Tones	}	DTMF 5	770 + 1336	49.28 + 85.5
		DTMF 6	770 + 1477	49.28 + 94.53
		DTMF 7	852 + 1209	54.53 + 77.38
		DTMF 8	852 + 1336	54.53 + 85.5
		DTMF 9	852 + 1477	54.53 + 94.53
		DTMF *	941 + 1209	60.22 + 77.38
		DTMF 0	941 + 1336	60.22 + 85.5
		DTMF #	941 + 1477	60.22 + 94.53
		DTMF A	697 + 1633	44.61 + 104.51
		DTMF B	770 + 1633	49.28 + 104.51
		DTMF C	852 + 1633	54.53 + 104.51
		DTMF D	941 + 1633	60.22 + 104.51

Fig. 2. Call progress and DTMF tones recognized by HP TeleShare.

values correspond roughly to the power of the signal in each bin. Once the powers are known for each bin in the spectrum, they can be analyzed to see if any DTMF or call progress tones are present.

As an example, suppose the telephone has been taken off-hook in preparation for dialing and HP TeleShare is configured to check for dial tone. 512 samples of the input signal would be stored in the real input array, while the imaginary array is filled with zeros. Next, the FFT function is called, returning the real and imaginary arrays. The magnitude-squared values of the first 256 bins are computed using the two output arrays. The two frequencies that make up a dial tone are 350 and 440 Hz (see Fig. 2), so FFT indexes (or bin numbers) must be computed for these frequencies:

$$350/15.625 = 22.4 \quad 440/15.625 = 28.16$$

An effective method of checking for the existence of a particular frequency is to compare the power present at that frequency with the total power of the spectrum. This is done quite easily with magnitude-squared values since they represent power in each bin already. Total power is simply the sum of all the magnitude-squared values for the first 256 FFT return values. Divide this into the power of the frequency being checked for, and the result is the percentage of total power for that frequency. For example, when checking for 350 Hz, compute the sum of the power values for bins 22 and 23 since the real index (22.4) falls between them, and then divide by the total power. The result is the percentage of the total power present around 350 Hz. The same can be done for 440 Hz, using bins 28 and 29.

Once the percentage of total power is calculated, a comparison can be made to see if the power in each frequency meets *match criteria*. The HP TeleShare firmware typically uses 35% of total power as a match condition. In other words, if the power present at the desired frequencies is 35% or more of the total power, dial tone has been detected. Otherwise, no dial tone is found.

The number of bins used in the comparison and the match criteria can be fine-tuned for a particular application. The match criteria can include other tests and can be relaxed or tightened as needed. The number of bins used can be influenced by the total number of points in the FFT and by a preprocessing tool that does windowing. Windowing is used to create a finite-length sequence from a continuous sequence. It is basically a digital filter that truncates an infinite-length input sequence while preserving its frequency characteristics. Since we are grabbing finite pieces (sequences) of data, we need to window the data.

telephone line is immediately sent to the CODEC without being split in half. Since both interrupt service routines run at 8 kHz, there is no need to worry about sample rate changes. DTMF audio data is only available for mixing when a tone is being generated. A new DTMF sample is generated during every SPORT0 interrupt and is based on the sample rate (always 8 kHz) and the time elapsed since the tone began.

All of these interrupts and audio manipulations require almost all of a DSP's processing bandwidth and can effect some areas of system performance. Because of DSP bandwidth limitations, DTMF detection can have a slight, but noticeable effect on the audio quality heard by the user. However, in unattended modes like answering machines or voicemail (where DTMF detection could be used for such things as navigation), this should not be a concern. The default configuration has DTMF detection disabled, since the typical user will never use it, and the current GUI does not support it.

Fax and Data Modem Operation. The fax and data modem functionality was codeveloped by HP and Digicom Systems Incorporated and uses their SoftModem technology.

The fax mode allows transfers up to 14,400 bits/s and covers Group 3 Class II and all fallbacks. Data mode supports transfers up to 14,400 bits/s (V.32bis) and can reach peak rates of 57,600 bits/s with compression.

Conclusion

HP TeleShare effectively combines telephone communications capability with a low-cost computer workstation. Context switches between the display and the telephone are minimized by integrating the telephone into the computer system and providing an easy-to-use graphical user interface. Voice, fax, and high-speed data modes are supported using flexible digital signal processing technology.

References

I. D. Garfinkel, B. Welti, and T. Yip, "HP SharedX: A Tool for Real-Time Collaboration," *Hewlett-Packard Journal*, Vol. 45, no. 2, April 1994, pp. 23-36.