Building Residential VoIP Gateways: A Tutorial
Part One: A Systems-Level Overview

By T Y Chan and Debbie Greenstreet,
VoIP Group,
Texas Instruments Incorporated

While voice-over-IP (VoIP) products have been deployed in the market for over seven years, recent announcements by service providers such as Vonage, AT&T, Sprint and others have created a flurry of activity by consumer equipment manufacturers racing to roll out residential VoIP gateway products. These low-cost devices are usually standalone boxes that provide VoIP functionality for POTS (plain old telephone system) via a broadband modem (usually cable or DSL). They serve as a bridge between the TMD/analog POTS world, and the IP-centric, packet-based world of the Internet.

As with most consumer products, their designers are usually faced with meeting aggressive product cost targets along with tight development schedules. The product feature shopping list often includes features not only specific to the basic VoIP gateway functionality but to other ancillary functions as well. These include data bridging and routings, such as found in common residential router products, emerging voice and signaling security features such as voice encryption and IPSec, and quality of service (QoS) features necessary to troubleshoot and maintain residential VoIP services.

This article is the first in a series intended to assist engineers by providing detailed design considerations for all major portions of VoIP residential gateway products. This part serves as a functional overview of the major components often required in today's VoIP residential gateways. It goes into detail about the key elements necessary for a quality voice over IP call, and highlights some of the design considerations of the telephony circuitry (which will be explored in additional detail in a subsequent article).

Overviews of the security, data routing and QoS monitoring elements are provided as well, to serve as an introduction to more detailed design analyses that will follow in subsequent articles.

Mimicking POTS

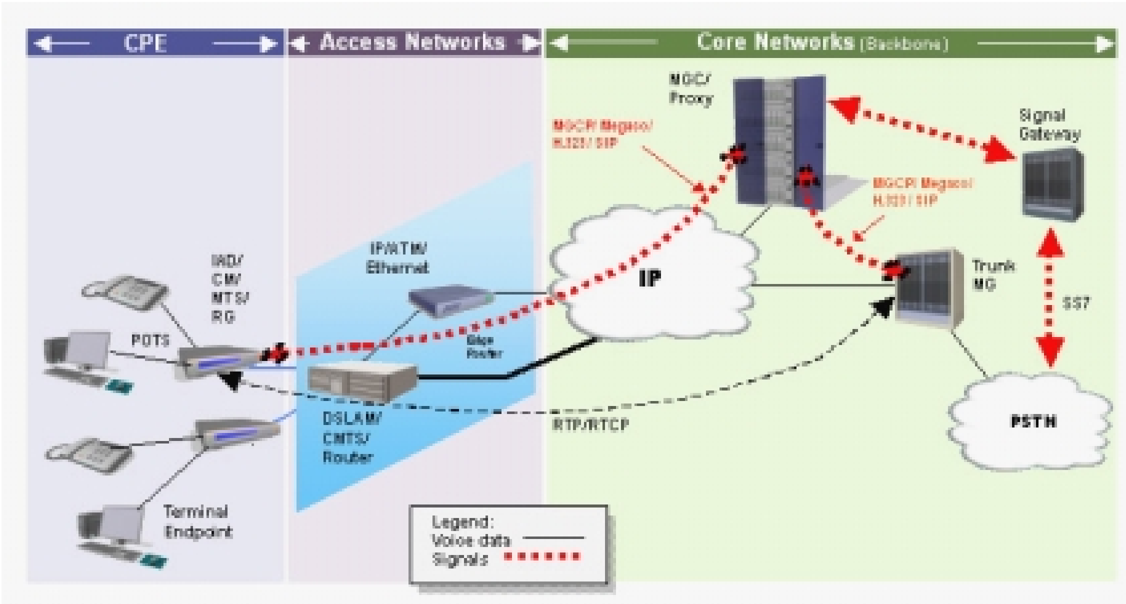

Fig. 1: Network's-Eye View Of A VoIP Call

Since the majority of today's installed base of handsets are still POTS-type units, most Voice over IP calls originating from customers premise equipment (CPE), come from a POTS phone connected to a voice gateway. The unit at the remote location could be another VoIP CPE device or simply a POTS phone connected to the PSTN. The flow of a sample VoIP call through the network is depicted in Fig. 1.

Of course, for any VoIP system to be successful, it must first provide the equivalent experience that end users have grown accustomed to with current POTS systems. To replicate the user's traditional interface with the PSTN, tone generation and detection functions are necessary. Dialed digits must be accurately collected and replayed at the receiving end to successfully execute a call.

Tone generation/detection is not only necessary at the beginning of the call. Tone-driven features like voicemail and calling card functions, tone generation/detection must be handled mid-call as well. Therefore the VoIP processing must also support the ability to successfully transmit dual-tone multiple-frequency (DTMF) tones in-band; however, vocoders with compression may distort these tones. To avoid these potential distortions, designers need to turn to advanced techniques such as IETF request for comment RFC2833 when passing tones in conjunction with the use of low bit rate vocoders. A tone generation function is necessary to provide depressed tone playback and call progress tones. The ability to detect tones and properly switch processing for fax and data modem signals is also a requirement, as all types of telephony currently available on the PSTN must be supported.

Dealing with Echo

Combating problems with echo are an essential element to VoIP adoption in the traditional telco world. As VoIP moves to replace the PSTN system, it must adopt robust echo-cancellation techniques to meet the demands of packet networks.

Echo is present in conventional POTS networks, and the PSTN employs echo cancellers throughout the system. Line echo is caused when a connection involves conversion between a four-line to a two-line telephony hybrid. Echo is generated toward the packet network from the telephone network.

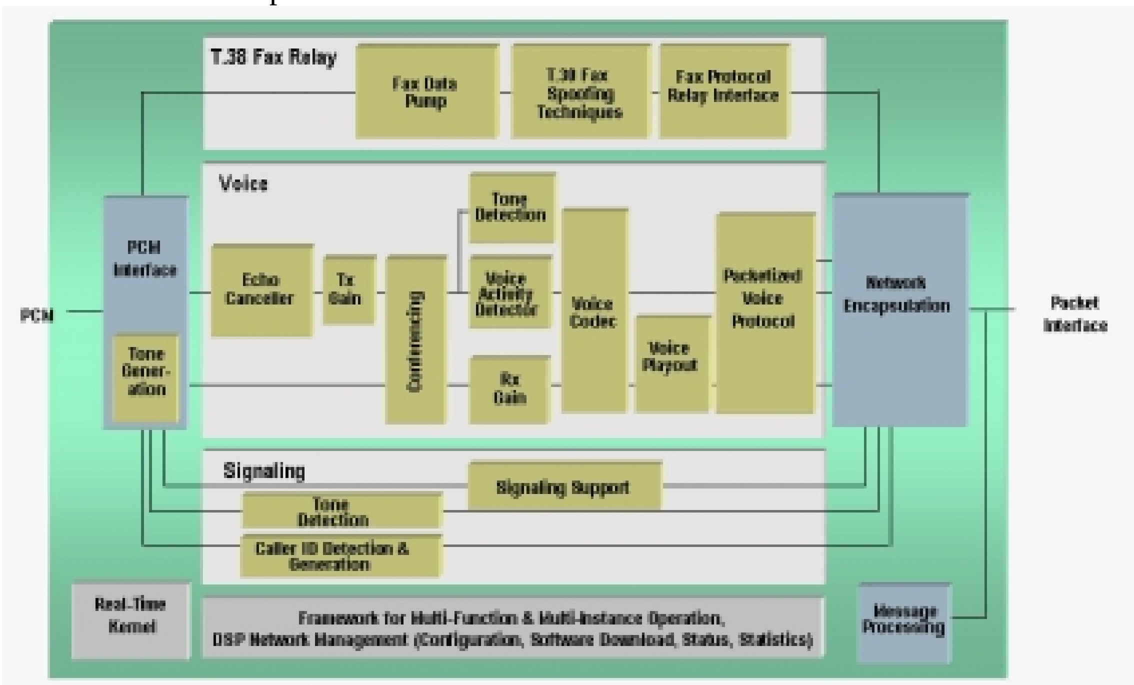

Fig. 2: Block Diagram Of A TDM-IP Gateway

PSTN specifications dictate that echo cancellation functionality is necessary when the delay exceeds 50 ms. However, because the IP network portion of the VoIP solution almost always adds more than 50 ms of round trip delay, line echo cancellation is essential when VoIP solutions interface to the PSTN. To do this, the echo canceller, as shown in Fig. 2, compares the voice data received from the packet network with voice data being transmitted to the packet network. The echo from the telephone network is removed by a digital filter on the transmit path into the packet network.

The echo-cancellation tail length, that is the length of the echo required to be cancelled by the processor, varies among different VoIP applications. The tail-length requirement is determined by the distance between the gateway equipment (residential gateway) and the four-to-two line hybrid. Typically this ranges from an 8-ms tail size for residential/SOHO applications to 32 ms, compared to up to 128-ms tail sizes for carrier applications.

Since most phone calls established via residential VoIP gateways will, at some point, terminate to PSTN equipment, line echo cancellation is required. For residential gateways, a typical length of 8- to 16-ms echo-cancellation tail-length capability is usually sufficient. As a minimum quality benchmark, the echo-cancellation functionality should be compliant to ITU G.165 and G.168 standards.

Voice Encoding

Voice encoding is necessary to convert the analog signal to voice packets. This often includes compression to reduce the 64-kbit/s stream produced by G.711-PCM-encoding stream (used by most traditional PSTN trunk lines) to a lower bit rate for more efficient transport across both the network and the subscriber's "last mile" link.

Typical vocoders used in VoIP systems today include G.729ab and G.723.1. The G.729ab vocoder offers data rates as low as 8kbit/s and the G.723.1 at 5.3 and 6.3 kbit/s. The tradeoff between these low bit rate vocoders and G.711 is reduced bandwidth utilization vs. slightly higher voice quality. The G.729a is an optimized implementation of the very common G.729 voice compression algorithm. It is important to note that G.729 is the base algorithm and that G.729 is interoperable with G.729a. G.729ab, the appendix B portion of this algorithm, incorporates the voice activity detection function in the vocoder itself.

Detecting Voice

Voice activity detection (VAD) and related silence suppression, whether incorporated in the codec or as an external software function, should also be supported as a configurable (enable/disable) feature in VoIP designs. The VAD monitors the received signal for voice activity. When no activity is detected for a specific period of time the software prevents unnecessary packetization and subsequent transmission of silence. This function also measures the idle noise characteristics of the telephony interface and noise measurements are subsequently relayed to the receiving gateway. Comfort noise generation (CNG), the playout of low-level background noise to the receiver, is recommended for user confidence in the call connection. If the call appears too quiet, users may anticipate that the call has been disconnected.

Residential gateways must also support the use of fax relay techniques. Fax relay offers bandwidth reduction and a more robust, reliable means of connecting fax over IP calls, and is a very popular feature for SOHO and SMB equipment. Fax relay functionality involves demodulation of the facsimile scan data, encapsulation into IP packets, and subsequent demodulation of the fax IP packets at the receiving gateway. This requires support of the T.30 fax protocol implemented between the fax machine and the VoIP gateway, as well as T.38 fax IP packet encapsulation for IP transmission.

POTS interfaces

Residential VoIP gateways interface to traditional telephony equipment through FXS signaling to a pulse code modulation (PCM) interface. This interface receives PCM samples from an analog codec interface and forwards them to the appropriate functions, such as those described above. Conversely, the interface forwards processed PCM samples received from the DSP to the digital interface. The PCM interface performs continuous re-sampling of output samples to avoid sample slips.

Playout

When voice and fax samples have been processed, they must be packetized. VoIP systems typically employ real-time packets (RTPs). On the receive side a voice playout unit is necessary to buffer received voice packets and forward them to the vocoder for playout to the user. This playout unit also serves as a jitter buffer/manager to queue several packets, avoiding packet under-run or over-run.

Implementing The Features

The features described above are typically implemented in software, usually on a DSP. They can, however, be implemented in a RISC processor. This is advisable only when the additional processing requirements for the typical RISC functions are minimal. For RISC-only VoIP architectures, the available CPU cycles (MIPS) must be carefully managed between voice processing and network/telephony signal processing. The functions described so far represent the telephony signal processing tasks required for supporting VoIP media streams. The VoIP gateway must also support signaling protocols, for both the telephony side and the packet side of the gateway.

Packet And Telephony Network Signaling

Translation of the telephony signals to packets is only a part of the VoIP gateway solution. A gateway must also support telephony control signals, such as on-hook and off-hook functions as well as network control signals or protocols such as Session Initiated Protocol (SIP) -- both of which place very different demands on the host processor and its software. There is a comprehensive set of processing tasks that are typically executed on a RISC processor that translates the telephony signals/protocols to the packet protocols and vice versa.

Analog Phone And PSTN interfaces

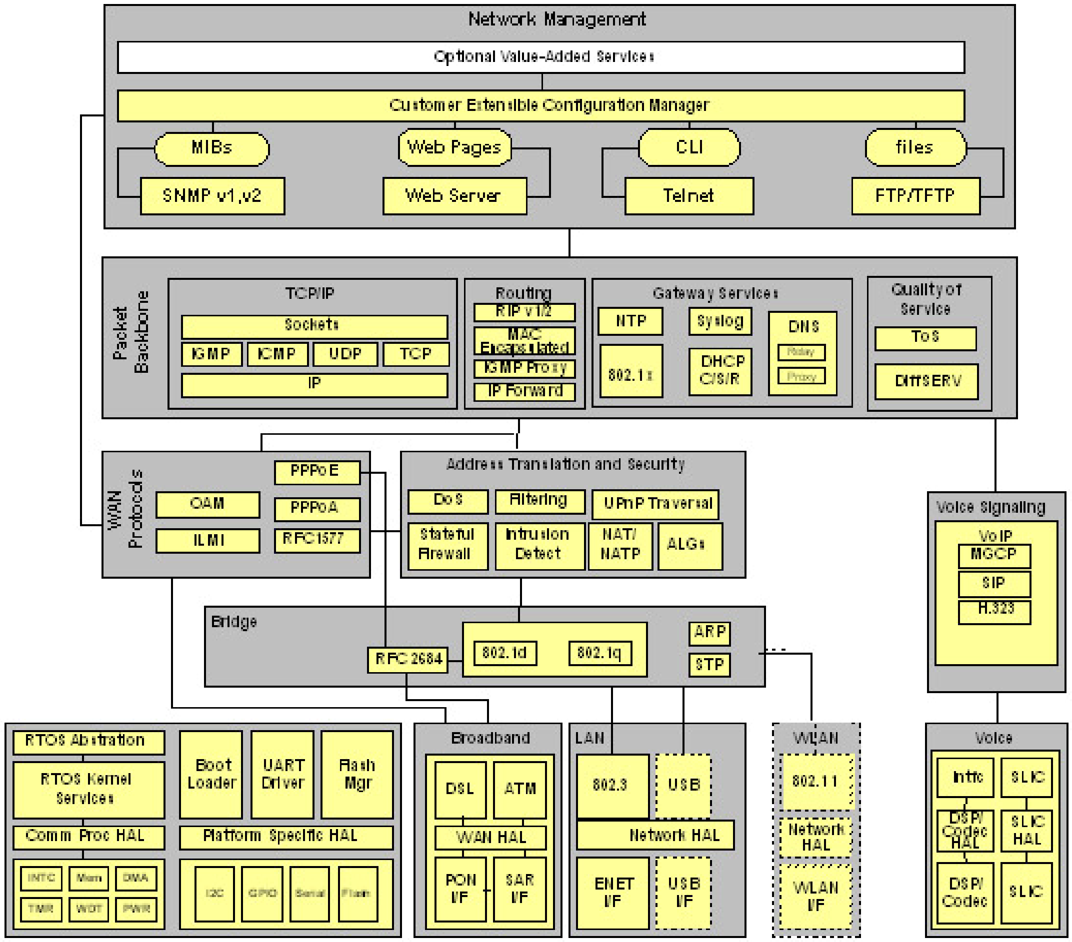

Fig. 3: Software Functions In A Typical VoIP Gateway

In most applications, a software-based state machine-like service that serves as a call controller handles these functions (See Fig. 3). The call control executes the necessary functions through a stage of each call in the gateway. Another critical element of the signaling process is the network protocol stack itself. SIP is a very popular protocol for the residential gateway market; however there are some deployments of Media Gateway Control Protocol (MGCP) and H.323.

Supplementary Services And Device Provisioning

VoIP gateways used in residential applications require the support of functions typically available in phone services today. This includes features such as call waiting, call forwarding, visual message waiting indicator and call transfer. Software is required to interpret these commands from the network and execute the function through the gateway to the telephone.

As a remote device in the service provider network, the residential gateway must be able to be configured either on premises or remotely, but not require separate monitors or other equipment. This configuration requires software in the gateway to accept and process the provisioning and it is desirable that the user interface be simple and easy to use. This provisioning software is not insignificant. There is also a preference in the market that residential gateway devices have the capacity for dual image (program load) storage, such that a program update can be downloaded without deleting the current image. This has impact on the overall software program design, as well as FLASH and SDRAM requirements.

Fig. 3, again, shows a comprehensive view of the functional blocks required in a complete VoIP residential gateway system. In addition, device drivers, Ethernet interfaces, real time operating systems (RTOS), and IP stacks should not be overlooked.

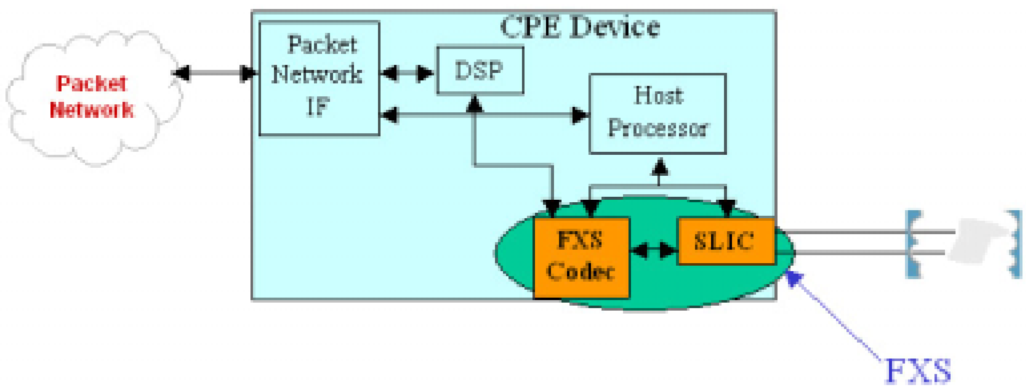

Fig. 4: A Foreign Exchange Station (FXS) Supporting A POTS Phone

To connect the old reliable analog phone to the voice gateway, a Foreign Exchange Station (FXS) is needed (see Fig. 4). In some applications, for outside calling using the PSTN, an FXO connection is needed (see later Fig. 5).

The FXS consist of two parts, a codec and a SLIC. A codec is comprised of an ADC and a DAC. The ADC is used convert the analog signal from the analog phone into digital signal for transmission onto the VoIP network. The DAC is to convert digital signals to analog levels to drive the analog phone. The sampling rate for ADC and DAC is usually in the 8-kHz range in order to achieve an audio bandwidth of 4 kHz. The Subscriber Line IC (SLIC) device emulates PSTN networks' voltage levels. It needs to detect on-hook, off-hook and generate ringing voltages which can range to 120 V. Its main function is to combine the analog signal with the PSTN voltages.

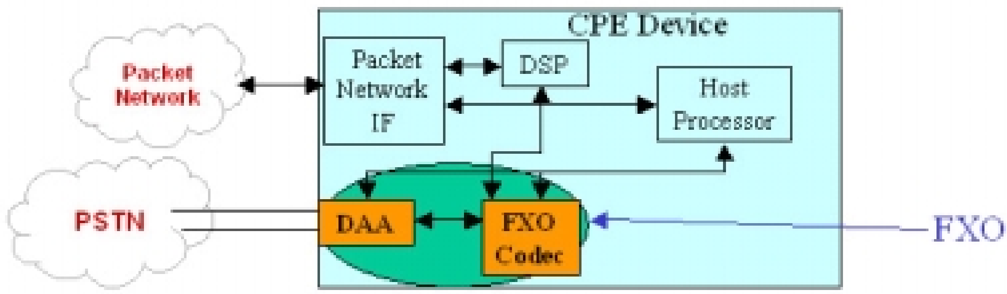

Fig. 5: A Foreign Exchange Office (FXO) In A Residential Gateway

For the case for when a voice gateway CPE device needs to connect to a local phone company, this requires a Foreign Exchange Office (FXO) interface (fig. 5). The FXO consists of a codec and the data access arrangement (DAA). The codec has the same functionality as in the FXS, while the DAA emulates a (POTS) phone. Its main function is to remove the high voltage dc bias, and it only passes the analog ac signal through when coming from the PSTN system, by applying a loop closure towards the PSTN.

Uses for FXO ports are:

- Lifeline for power failure: used when there is no power to the voice gateway, which prevents calls from connecting to the packet network. In this case, the analog phone (through FXS) connects directly to the FXO port through a relay
- Call redirection: used in the case when the subscriber dials a number that is unreachable through the packet network. In order to complete the call, the voice gateway will redirect the call through the FXO port. For a user-friendly gateway the CPE device dials the digit to the FXO port, which prevents customers from having to redial the numbers
- Remote VoIP calling: when a customer is not at home, they can still make a VoIP call by calling their home number through the PSTN network. The voice gateway picks up the call through the FXO port, and forwards it to the VoIP network.

Additional details on FXS and FXO design circuitry will be discussed in upcoming articles.

Voice Gateway Data Functions

In CPE residential applications, the voice gateway is normally connected on the LAN side of a broadband modem. If the household has more than one PC, the voice gateway can be a standalone device terminating IP connections by connecting to a router or hub. If the home has a single PC, then introducing a voice gateway will involve creating a home network and purchasing a router or hub. To ease the adoption of packet voice services, the most appropriate configuration for a voice gateway is to include a data routing function for connecting another PC. This way, the PC connects to the LAN side and the modem connects to the WAN side of the broadband modem. In this type of configuration, the voice gateway should include data routing functionality.

In deciding on the functionality and performance of data functions, it is useful to understand the application and configuration of the broadband connection. With the exception of VDSL, most residential broadband modems have capacities well below 50 Mbit/s. Therefore, in designing a voice gateway, it is necessary to understand the end-user application in order to determine the appropriate price/performance goals. Some of the data functions that should be included are:

- Routing
- NAT, NAPT, dynamic and static
- Firewall
- DHCP client / server
- PPPoE
- TFTP

Including these popular and useful functions makes the transition to VoIP easier for the consumer to connect with their broadband service. Further, incorporating a voice gateway into a router or hub lowers the cost of ownership by not requiring customer to purchase a separate box. When the household purchases another PC, a switch can be purchased to network the PCs and voice gateway together.

VoIP Security Elements

Secure voice communications is receiving a great deal of attention by service providers deploying residential VoIP services. Secure VoIP implementations can leverage many security elements already established for data communications. One of the key functions of the current Internet security infrastructure is monitoring the integrity of the data transmitted. This element covers both the assurance that the message between two entities has not been tampered with, as well as the authentication of the recipient. A similar element is the support for non-repudiation, which is the rejection of a digitally-signed message (by secure keys). The confidentiality level of Internet security ensures that the recipient and the transmitter of the message are the only ones that may view the contents of such a message. The authorization function of the security element suite assures a network user access to a particular network service only upon verifying identity.

Depending upon the level of security concern by end users or service providers, various levels of security features may be required. One common feature is encryption of the voice payload itself. Another level of security might require encryption of the signaling messages that establish the phone call.

Pulling it All Together

While there is pressure to put together a residential gateway solution at the lowest possible cost, it imperative that the components selected achieve optimal quality and performance. The voice processing, network and telephony signaling, POTS interface and Ethernet interface are the minimum functions required to develop these gateways. It is also essential to understand the types of supplementary services and the extent of provisioning functions required in order to ensure that the product is complete. Regional considerations and programmability requirements will dictate the type, and ultimately the cost of the POTS interface. In addition, for residential gateways requiring advanced features such as data-routing functionality or voice encryption or authentication, require additional processing power. If including these features, a designer must take care to ensure that the proper amount of processing power and a sufficiently-flexible architecture is available to support such requirements. Subsequent articles in this series will go into further detail on the issues and tradeoffs of these features.

About The Authors

Debbie Greenstreet dgreenstreet@ti.com is the product management director for TI's voice-over-packet group. She is responsible for product definition and direction of voice-over-cable and SME voice gateway products. Debbie has more than 18 years of experience in the networking and telecommunications field, in hardware and software design, as well as program and product management at companies such as Hyundai Network Systems and Raytheon. She earned a BEE at the University of Virginia.

T Y Chan tychan@ti.com is a senior member, technical staff with TI's voice-over-packet group, serving as the system engineering manager for VoIP customer premise equipment products. Since joining TI in 1989, T Y has held various positions in the company, including engineering manager for DVD solutions and system manager for PC processors. He earned both a BSEE and MSEE at Bradley University.