Audio transcoding, particularly for voice, is a critical but often overlooked function in telecom applications such as session border controllers, media gateways and servers or media resource appliances. The key technology behind conference call applications such as Zoom, Microsoft Teams and countless other voice-over-IP applications (VoIP), voice processing is now essential in delivering quality sound efficiently.

This white paper outlines the trends driving the need for network voice processing and offers an alternative to the conventional host media processing (HMP) solution.

Using a PCI Express media acceleration card and embedded voice firmware offers dramatically improved performance while taking up less space, consuming less power and providing a lower cost per channel.
As telecom networks complete their transition to an all-IP environment, service providers and network operators are finding a need for IP media servers and advanced flow management devices, to meet next generation demands. Traditionally, telephony systems have been based on host media processing (HMP), using a general-purpose computer to process a call’s media stream. However, the rise of interest in VoIP in recent years, has driven demand for integrated components such as session border controllers, quality of service analytic engines and intelligent flow optimizers. These are typically developed and deployed on 1U or 2U standard rack mounted servers (RMS) for simplicity.

While the role of IP media gateways and media servers is clear, as developers and users of border flow management devices consider where to go next, one obvious step is to integrate advanced media stream processing into the platform. A key concern however, is scalability. The pressure is on to find cost and power-efficient ways to scale audio processing to reflect the needs of both wired and wireless networks.

But, what are the options?

**Adding Voice Transcoding to a Session Border Controller**

A good example of a flow management application is the session border controller (SBC), an often-quoted example of a class of equipment known as network security gateways. SBCs are characteristic of bump-in-the-wire devices that form a bridge between trusted and untrusted networks or enterprises. Carriers rely on them to govern their interconnection. Their job is to regulate incoming IP traffic, block undesirable or unauthorized flows, and allow trusted network traffic through, via the most efficient path.

In its early form, the SBC had just one main application – to provide security for VoIP calls. However, as usage spread and codecs and protocols proliferated, additional functions were required.

As a gateway point, many SBC users are seeking additional media format translation in addition to stream management. Even simple requirements like dual tone multi-frequency (DTMF) monitoring require that the media streams are decoded and analyzed. The ability to have voice transcoding within the box helps simplify the communications flow for an operator, hence provides a competitive advantage for the equipment vendor.

Unfortunately, voice stream processing in real time at high channel counts is a strenuous task, and adding this function can significantly reduce the processing power available to the main service, leading to a reduction in overall capacity.

**Possible Solutions**

Adding media processing functionality to an application can be done in a number of ways:

- An additional system or device linked to the original appliance
- An internal software solution, adding functionality to existing software
- An internal media processing accelerator offering hardware-accelerated transcoding

In the Figure 1 example, using an external media gateway is perhaps the simplest solution to envisage.

The border gateway terminates principal traffic streams, and redirects media to the external gateway for transcode via external ports. Media can come back into the border gateway for egress filtering.

The disadvantage is that this is costly, amplifying both capital and running costs for the operator. It uses extra power and rack space, which can present provisioning issues. It also takes up valuable physical network interfaces.
off the border gateway, and still requires application development that controls and configures media stream handling on a stream-by-stream basis as additions, moves and changes of voice service occur.

The other two solutions allow for an embedded function. An internal solution, for instance using commercially available HMP software, automatically makes use of internal processing resources. In the case of voice transcoding, this may be a great solution for a moderate number of simultaneous channels, however it does not scale effectively.

At upwards of 1200 simultaneous channels of G.729 encoding, the software solution approaches 50% utilization of a typical server, starving the original application of processing resource. Effectively this means that additional servers would be required to offer higher densities of voice transcoding, and the cost of the commercial software that is usually charged on a per-channel basis soon mounts up.

This could hardly be viewed as adding functionality. The end result is that much more processing resource would be required. Although it is possible to add more servers to address this issue, accepting a reduction in capacity even for an improvement in functionality is often difficult to manage from a product line perspective. It results in a downgrade of capacity within the same product offering so cannot really be regarded as enhancing functionality.

Matters get even worse when considering field upgrades since a customer must accept that a given installation would no longer be able to carry the same traffic.

The solution? To use a plug-in media processing accelerator, as evidenced by extensive take-up throughout the telecom industry. Using a PCI Express add-in card with DSPs and software-definable transcode capability offers a smart way to offload the voice processing function from the main processor and greatly increase the total transcoding availability within the server.

Not only does this keep the function internal to the network element, it also avoids the loss of central processing resource that would otherwise be required to run a fully software solution. Ideally this would be able to take account of new voice compression schemes as they emerge.

Introducing the PCIE-8130

The PCIE-8130 from SMART Embedded Computing is a PCI Express media processing board that offers high performance voice transcoding based on digital signal processing (DSP) technology.

Each board features an array of low power DSP devices running a network-proven, optimized software framework that provides transcodes, signalling and call management functionality. Application developers interact with the board via a simple object-oriented application programmers interface (API).

The transcoding performance scales linearly according to the number of DSPs that are fitted: a half-length PCIe card variant offers options for 4 or 6 DSPs; while the full-length PCIe card variant features up 12 DSPs.

But even with 4 DSPs and consuming less than 60W of power, the PCIE-8130 delivers a voice transcoding performance comparable to a typical server consuming 300W or more.

SMART EC’s PCIE-8130 is designed for NEBS carrier grade and data center environments, depending on the server enclosure, and so offers a common solution for both enterprise and telecom environments.
A Better Solution

High density voice processing is hugely sought-after in 4G and 5G networks and underpins the massive increase in remote working and online collaboration tools. The demand for products for applications such as session border controllers, media gateways and servers or media resource functions, content optimization, and interactive voice response systems has grown substantially.

The challenge is to achieve greater flexibility and speed while simultaneously increasing productivity, optimizing costs and achieving high product quality: an effective solution to efficient audio transcoding.

A key factor in delivering voice and other audio services, SMART’s PCIE-8130 offers an elegant way to offload the voice processing function from the server host while retaining high-quality audio.

When considering a more agile, resilient and centrally managed approach to network applications, using the PCIE-8130 rather than additional servers has many benefits:

- It takes up less space
- It consumes much less power
- It can easily be retro-fitted to existing deployed systems as a true feature addition

The PCIE-8130 costs less than a comparable server + commercial host media processing combination for the same performance. It offers a lower total cost of ownership and a much simpler upgrade and deployment experience.

Additional Resources

Solution Brief

This SMART Embedded Computing Solution Brief describes an example scenario comparing a real-life application taking alternative approaches to adding audio transcode capability.

Technical Specification

The PCIE-8130 datasheet describes the card's hardware and software capabilities, including codec functionality.
About SMART Embedded Computing
SMART Embedded Computing (SMART EC) is part of the SMART Global Holdings, Inc family of companies. We are a global leader in the design and manufacture of highly reliable embedded computing solutions for a broad range of defense, industrial IoT (IIoT), edge computing, and communications customers. Building on the acquired heritage of industry leaders such as Motorola Computer Group and Force Computers, SMART EC is a recognized leading provider of advanced computing solutions including application-ready platforms, single board computers, enclosures, blades, enabling software and professional services.

For more than 40 years, customers have trusted us to help them accelerate time-to-market, reduce risk and shift development efforts to the deployment of new, value-add features and services that build market share.

Our engineering and technical expertise is backed by world-class manufacturing, global sales offices and advanced worldwide logistics capabilities that can significantly reduce time-to-market and help customers gain a clear competitive edge.