So You Want to...
Add Voice Transcoding to a Session Border Controller
(or other Network Gateway) Application?

Network security gateways, of which session border controllers (SBC) are an oft-quoted example, are characteristic of “bump in the wire” devices that form a bridge between trusted and untrusted networks or enterprises. Their job is to analyze and characterize incoming IP traffic, block undesirable or unauthorized flows, and let through approved traffic. In communications networks, a lot of this traffic is media streams. As this is a gateway point, many SBC users are also interested in providing additional media format translation in addition to the stream management. Even simple requirements like DTMF tone monitoring require that the media streams are decoded and analysed.

POSSIBLE SOLUTIONS
Media streams tagged for analysis or for transcode are sent to specific media processing tasks. This could be:
- An external IP:IP media gateway device
- An internal software solution, perhaps running within a local virtual machine
- An internal media processing accelerator offering hardware-accelerated voice transcoding

Using an external media gateway is perhaps the simplest to envisage. The border gateway terminates principal traffic streams, and redirects media to the external gateway for transcoding via external ports. Media can come back into the border gateway for egress filtering. The disadvantage is that this is costly, uses rack space and extra power, 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. The other two solutions allow for this function to be taken inside the box of the border gateway.

An internal software solution, for instance using commercially available “Host Media Processing” software from suppliers like Dialogic, necessarily makes use of internal processing resources. While this would be a great solution for a moderate number of simultaneous channels, it does not scale effectively. At upwards of 1200 simultaneous channels of G.729 encoding, the software solution requires 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.

For higher channel densities, the best solution is adding a hardware-assisted accelerator. This keeps the function internal to the border gateway AND avoids the loss of central processing resource that would otherwise be required to run a software solution.