

From Transcoding to Media Conditioning

IP Media Stream Processing for Mobile, NGN, and IMS Architectures

Overview

The groundwork has been laid for seamless connectivity between fixed and wireless telecommunications networks, and now the focus is on optimizing the transmission media streams across these converged networks. This whitepaper discusses the delivery of media conditioning in IP communication networks with respect to service provider requirements and the impact on network architecture.

CONTENTS

Executive Summary *pg. 2*

Transcoding Reduces Complexity *pg. 2*

Other Media Conditioning *pg. 3*

Architectural Considerations *pg. 5*

Radisys Solutions for Media Conditioning *pg. 10*

Executive Summary

The groundwork has been laid for seamless connectivity between fixed and wireless telecommunications networks, and now the focus is on optimizing the transmission media streams across these converged networks. The goal is to deliver an improved user experience, such as better Voice over IP (VoIP) call quality, while consuming less network bandwidth.

Playing a crucial role, codecs based on encoding/decoding technologies define the algorithms used to convert data streams. For example, many VoIP solutions use the G.711 codec to encode analog voice to digital voice for 64 Kbps circuit-switched networks or Real Time Protocol (RTP) packet streams for IP networks, and vice-versa. Other codecs, like G.729, perform a similar conversion, but the algorithms achieve a comparable digital voice representation using less bandwidth.

In fact, the number of codecs found in networks today is actually increasing to meet the needs of new service models delivered by both wireline and wireless networks. For instance, in certain pockets of the evolving network where end-to-end IP broadband connectivity is available, the industry is seeing increased adoption of High-Definition (HD) audio codecs, like G.722, which deliver a higher fidelity audio experience. Another example is the growing number of consumers watching video on mobile phones, which is driving the need for video optimization functions in the network that reduce the video content received from a content provider and accommodate the limited bandwidth requirements and small screens of a handheld device.

However, the downside to codec proliferation is they are increasing the complexity of traditional network equipment that must process them. One solution is transcoding, the process of converting from one encoding format to another, which allows operators to standardize on a smaller subset of codecs. This way, fewer codecs must be supported in the core network, thus minimizing equipment capital expenditures (CapEX) and ongoing operating expenditures (OpEX). Additionally, the same equipment performing transcoding on IP media streams in real time can perform other media conditioning functions, including video transrating and voice quality enhancements.

This whitepaper discusses the delivery of media conditioning in IP communication networks with respect to service provider requirements and the impact on network architecture.

Transcoding Reduces Complexity

The explosion in the number and variety of audio and video codecs used in modern telecommunication networks is making it difficult for carriers. Although some codec standards help carriers by requiring less bandwidth, the need to support numerous codecs increases network management complexity and requires greater media processing investment. Exacerbating the situation, the rising popularity of VoIP technology and the high demand for IP

communications over core networks is leading to the creation of new codec standards.

For instance, the G.711 codec (circa 1972) is still widely used to convert voice signals to a digital stream, but other codecs have emerged to fulfill different needs, particularly low bandwidth requirements. Along with G.711, carriers must support additional speech codecs, such as G.726, G.729, AMR, EVRC and iLBC, due to the lack of standardization, as well as high-definition (HD) audio codecs like G.722 or AMR-WB.

However, carriers can benefit from codec standardization in their core networks by means of transcoding. Imagine a carrier providing a hosted Interactive Voice Response (IVR) application for many users, where audio transmissions are all encoded using different codecs (e.g., G.711, G.729a and AMR). Rather than transmitting media streams based on three different codecs through the IP WAN, a transcoder function, which transcodes the G.729a and AMR formatted calls, can be deployed into a common G.711 format for transmission across the IP WAN, as shown in Figure 1.

Transcoding offers a number of advantages. First, the majority of network elements don't have to support the full breadth of codecs, thus lowering the computing requirements and, consequently, equipment cost. Second, transcoding allows devices, like the IVR platform and the mobile phone in the example, to communicate with each other even when they support different codecs.

Other Media Conditioning

In addition to audio transcoding, carrier networks often require other media stream processing capabilities, referred to as media conditioning in this paper. These requirements might include:

- Voice Quality Enhancement (VQE)
- Video Transcoding and Transrating

Generally, media conditioning functions require powerful platforms and architectures that deliver both fast digital signal processing and real-time performance. These functions are often best performed on hardware platforms leveraging digital signal processor (DSP) technologies. Sufficient real-time performance is needed to minimize the delay introduced by applying a media conditioning function against the media stream. As well, carrier class communications systems require fully redundant services and fast service restoration times.

Voice Quality Enhancement

Voice quality is an important requirement for many network operators that offer telephony services based on VoIP technology. Despite providing tremendous economic benefits, VoIP also presents new voice

Complexity and Quality

Codecs, such as G.729a, have a lower bit rate than G.711,¹ but this bandwidth savings may come at a cost. This is illustrated in Table 1, where the bit rate, complexity and voice quality are compared. Although G.711 consumes eight times more bandwidth than G.729a, it requires 1/37th the computing processing power, as shown in the relative complexity column. In addition, G.711 delivers higher voice quality, as measured by the mean opinion score (MOS), the leading subjective measurement of voice quality. The G.711 listener rating is "very satisfied," while G.729a is "satisfied," using the Perceptual Speech Quality Measure (PSQM).

Codec	Bit-rate	Relative Complexity*	MOS Score
G.711 (u-law)	64 Kbps	1 (best)	4.41 (best)
G.729a	8 Kbps (best)	37	4.10
AMR	12.2 Kbps	63	4.29
G.722	64 Kbps	29	

Table 1. Codec Comparison

* Normalized to G.7111

¹Source: <http://en.wikipedia.org/wiki/G.729>

Source: *Codec and Media Specification by PacketCable™ 2.0 (Document PKT-SP-CODEC MEDIA-I09-100527, May 27, 2010 at <http://www.cablelabs.com/specifications/PKT-SP-CODEC MEDIA-I09-100527.pdf>)*

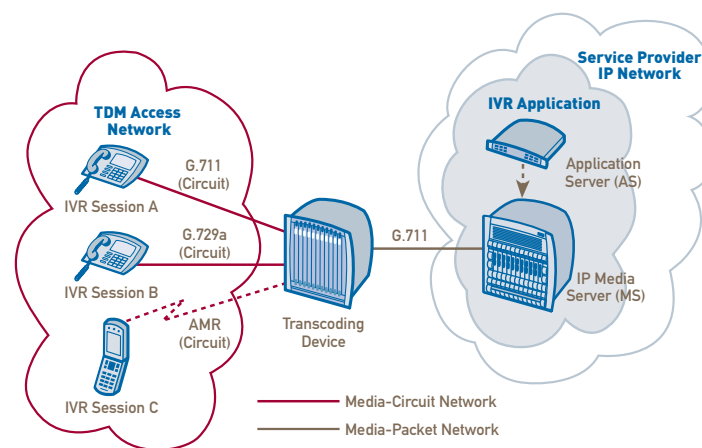


Figure 1. Transcoding for a Hosted IVR Application

quality challenges such as dropped packets, variable packet delay, and packet reordering and corruption. In-network quality enhancement solutions, generally referred to as Voice Quality Enhancement (VQE), are available to effectively address these issues, thereby increasing end-user satisfaction.

VQE solutions typically incorporate features designed to overcome background noise, packet loss, acoustic echo and variable network delays, the four most common sources of audio quality problems in a VoIP network. A comprehensive VQE solution also measures VoIP quality metrics, which are used in ongoing voice quality measurement associated with service level agreements. Many VQE features require sophisticated real-time digital signal processing, particularly noise reduction and acoustic cancellation, which together lead to improved voice quality often measured as a mean opinion score (Figure 2).

Video Transcoding and Transrating

A prior example illustrated some of the codec and transcoding issues associated with the delivery of audio calls and services. In the video world, the diversity of various video stream and container formats introduces additional dimensions of disparity, including video picture size, frame rate, sampling speed and bandwidth. New handsets, ongoing video compression research, and substantially better end-point processing capacity is allowing engineers to introduce new video codec standards with improved “video quality per bit.” All this innovation is accelerating consumer demand for video across mobile networks, as described in the sidebar (Figure 3).

A process similar to transcoding, transrating re-encodes the video streams to a lower bit-rate without changing video content, which ultimately saves a significant amount of bandwidth. For example, transrating could be used to convert high-resolution video content to a lower resolution (and lower bandwidth) video stream suitable for a handheld device with a small screen or limited availability of bandwidth, particularly in wireless networks. Therefore, video transrating and re-scaling is an important requirement for operators deploying a 3 screen strategy for television, PC desktop and mobile devices.

Bearer Channel Adaptation

In addition to the audio and video media stream processing, media conditioning might include broader bearer channel adaptation features such as normalizing between IPv4 and IPv6 addressing

What is MOS?

The leading subjective measurement of voice quality is the mean opinion score (MOS), based on a large number of people listening to audio and giving their opinion of the call quality. MOS scores, ranging from “very satisfied” to “not recommended,” are mapped to “R” factors, which can be generated electronically and account for network impairments and delays. Acknowledged for its high degree of correlation to subjective MOS testing, the Perceptual Evaluation of Speech Quality (PESQ) standard is often used to automate call monitoring in test equipment.



Figure 2. MOS Diagram

Mobile Traffic Explosion

According to Cisco, mobile data traffic will double every year through 2014, increasing 39 times between 2009 and 2014. Cisco also forecasts that almost 66 percent of the world’s mobile data traffic will be video by 2014, as shown in Figure 3.

Transrating is a crucial technology for curbing the amount of network bandwidth needed to support this demand, which is accelerating at a phenomenal rate. Transrating enables service providers to send video streams from a single video source to a variety of devices, from small-screen devices with limited display capability, to PCs or larger screens.

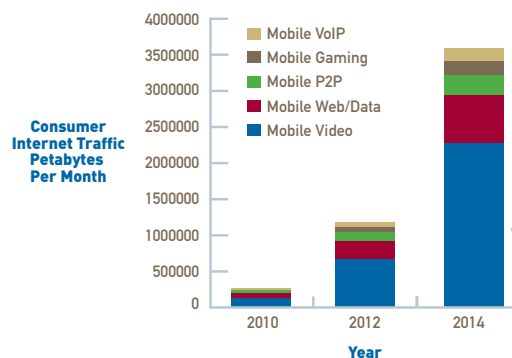


Figure 3. Mobile Traffic Forecast²

²Source: Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2009-2014 (Feb. 2010) at http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html

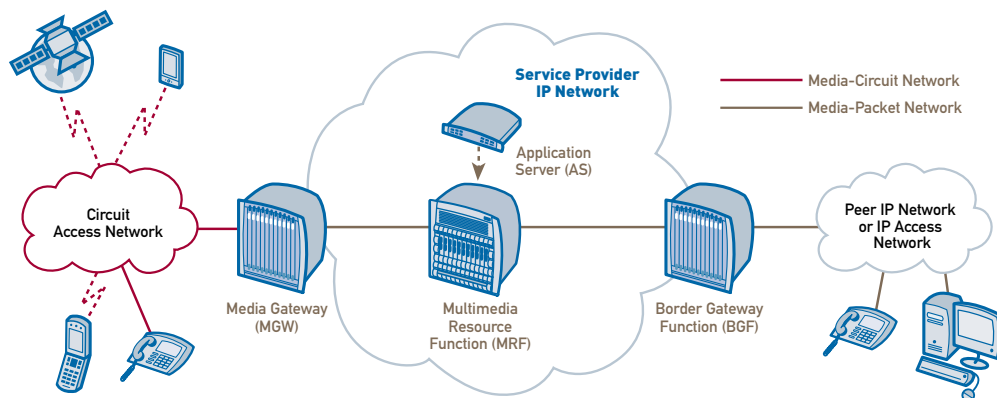


Figure 4. Network Locations for Media Conditioning

schemes, adapting media streams for security, and implementing access and bandwidth controls. All these requirements are also applicable to the scope of media conditioning in a modern next-generation IP network.

Architectural Considerations

Media conditioning requirements are broad based, arising from a wide range of factors such as codec diversity, network bandwidth consumption and media quality enhancement. For instance, the need for audio transcoding stems from the disparate codec support in terminals (e.g., IP phones often use HD audio codecs) and various networks. The challenge is to implement audio transcoding and other media conditioning functions at the lowest cost, while optimizing quality with maximized flexibility and expandability. Every transcoding operation requires processing against a media stream, which introduces a slight delay that could negatively impact user experience. Hence, it is important for network designers to address transcoding and media stream processing requirements based on informed architectural decision making.

Network Locations for Media Conditioning

Three alternative network locations have emerged as the key contenders for performing media conditioning in a packet communications network. These network locations and functional definitions are consistent

with specifications put forward by leading standards organizations such as the 3rd Generation Partnership Project (3GPP) specifications for IP Multimedia Subsystem (IMS), Telecoms & Internet converged Services & Protocols for Advanced Networks (TISPAN) and the MultiService Forum (MSF). The three network locations include the:

- **Multimedia Resource Function (MRF)**, also known as an IP media server, provides media related functions such as media manipulation (e.g., voice stream mixing) and playing of tones and announcements, under the control of the application servers in a network.
- **Media Gateway (MGW)** provides media conversion functions between Time Division Multiplex (TDM) circuits and Real Time Protocol (RTP) media packet streams in the IP core network.
- **Border Gateway Function (BGF)** provides security with IP network connectivity at the border between autonomous IP networks.

Figure 4 shows the location of these network elements in a typical next-generation network. The MRF is normally associated with service processing in the service provider core, while the MGW and BGF elements are associated with the access edge. The following section describes the particular merits and implementation characteristics related to the network location (i.e., service provider core or access edge) where media conditioning is performed.

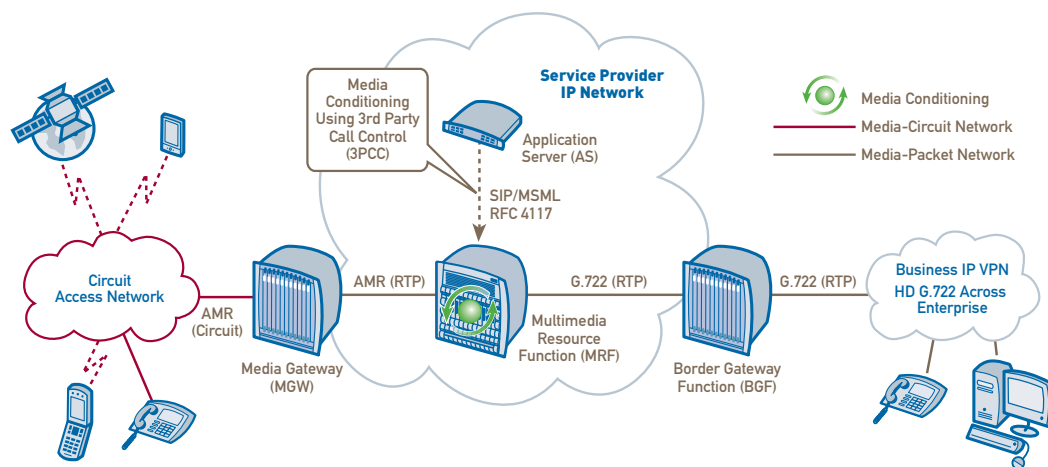


Figure 5. Media Conditioning in the Core using 3rd Party Call Control (3PCC)

Media Conditioning in the Service Provider Core

Located in the service provider core, the MRF already performs automatic transcoding today depending on the service processing requirements. Hence, the MRF is also a desirable network location for many media conditioning functions in the network.

The MRF provides media processing services for IP communication applications, including media play/record, phone digit collection and audio/video mixing (see sidebar). However, in this primary role, the media server may also need to perform automatic transcoding or voice quality enhancement. For example, an IP media server supporting heterogeneous conference call mixing uses transcoding to normalize the call legs to a common codec before it mixes the audio and then re-encodes each conference leg to the end-points codec requirements.

When media conditioning is controlled by an application server, it supports a topology called 3rd Party Call Control (3PCC), as shown in Figure 5. The topology makes a distinction between signaling and application logic processing as separate and distinct functions from real-time media processing performed by the MRF against Real-Time Protocol (RTP) audio and video packet streams. The application server “controls” the MRF, typically using Session Initiation Protocol (SIP-based) commands. Media conditioning

IP Media Servers

Carriers deliver numerous telecommunication services and capabilities such as unified communications, multimedia conferencing, voicemail and ringback tones. While all these services have differing application logic and signaling requirements, the underlying media processing requirements share many common characteristics, such as playing an audio or video clip, collecting digits from a phone, bridging multiple signals into a conference mix. IP media servers provide a common, shared media processing resource for supporting a wide range of audio, video, fax and speech applications in VoIP and IMS networks.

In addition to supporting service-level media processing, media servers also perform various media conditioning functions, such as automatic transcoding or voice quality enhancement. Because of the central location in an IP services core, and their sophisticated IP media stream processing capabilities and capacities, IP media servers are a fitting network element for integrating media conditioning.

controlled by 3rd party application logic facilitates granular control to apply specific media conditioning treatments within the IP media server against each RTP media stream.

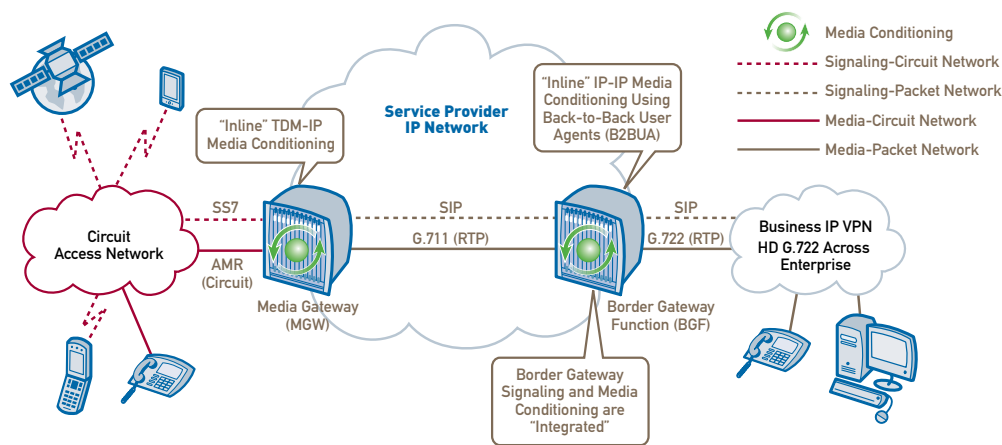


Figure 6. *Inline Media Conditioning at Edge*

Transcoding control can be achieved using a full-featured IP media server interface, such as SIP with Media Server Markup Language (MSML–RFC 5707), used for rich IVR, conferencing or voice quality enhancement (VQE) service creation and feature control. Alternatively, a SIP-based interface, such as RFC 4117 could also be used, which can support transcoding for simpler two leg or two party RTP connections.

There are several benefits of performing media conditioning tasks, such as transcoding, on the MRF in the service provider core, including:

- Media processing is consolidated into a core network function, which enables functional synergies, centralized resource pooling, simplified management, and economies of scale.
- The close proximity and integration of service-level media processing and stream-level media conditioning facilitates new service deployment and expands capabilities.
- Edge devices, like media gateways and border gateway functions, don't have to support multiple codecs, which saves cost on network edge equipment with a high number of network instances.
- Voice quality is often better than performing transcoding on multiple edge elements because there are fewer transcoding operations in the end-to-end call path.

It is worth noting that many basic point-to-point calls typically do not require media processing applied by a centralized MRF. Furthermore, if all calls that require media conditioning are routed into the service provider core, backhaul bandwidth requirements may be unnecessarily large and possibly introduce network bottlenecks. While centralized media conditioning offers benefits for calls already requiring MRF processing, most network architects should also consider incorporating media conditioning in access edge elements as well.

Media Conditioning at the Access Edge

Situated on the edge of IP networks, media gateways and border gateway functions, pictured in Figure 6, are also logical options for transcoding or media conditioning. Their location is ideal for achieving codec “normalization” (e.g., converting all access codec varieties into G.711) since transcoding is performed before media streams enter the core network.

Media gateways (MGW) are deployed at the border of circuit-based access networks, like the public switched telephone network (PSTN) or 2G cellular networks, and the IP core. The primary function of a media gateway is to convert media circuits on one side into RTP over IP media packet streams on the other side. In performing this function, the media gateway may also perform transcoding or other basic media conditioning enhancements, like acoustic echo cancellation. However, media gateways have limitations that impact their ability to effectively perform transcoding for the entire network.

- Primarily supporting G.711, media gateways often lack the computing capacity to take on more complex codecs, a critical requirement for transcoders.
- More importantly, media gateways are not a viable network location with the growing number of IP-to-IP media conditioning operations required in a modern IP network.
- Media gateways are inherently distributed, making the provisioning of media conditioning in each media gateway element hard to predict, resulting in over-provisioning and waste of capital.

The **Border Gateway Function (BGF)** is located at the border of two IP networks in order to provide security with IP network connectivity. For example, a BGF can be deployed between service provider's IP network and an enterprise IP VPN.

Media conditioning at the access edge is often achieved using an "inline" approach. Media gateways provide inline TDM-IP media conditioning, whereas a BGF performs inline IP-to-IP media conditioning typically using a Back-to-Back User Agent (B2BUA) topology, as shown in Figure 6. Unlike the 3PCC topology, a B2BUA approach does not require a 3rd party controlling element. Instead, media conditioning running on a network edge element is typically configured once, and then the BGF processes both the signaling and media streams passing through.

Figure 6 also shows an increasingly common situation where a large enterprise might standardize their IP-based phonesets or softphones using a high-definition G.722 audio codec. While internal calls would achieve great audio quality, the enterprise still needs to make external calls to customers and suppliers using a more common G.711 baseband codec. Hence, the BGF peering point between the service provider's network and the Business IP VPN is an ideal location for performing G.722-to-G.711 transcoding.

A Border Gateway Function that processes both signaling and media streams is sometimes called an integrated Border Gateway Function, or more commonly, a **Session Border Controller (SBCs)**.

Supporting SIP and RTP

The two widely used protocols for VoIP signaling control and media plane processing are the Session Initiation Protocol (SIP) and the Real-Time Transport Protocol (RTP), respectively. Both protocols can be supported using two separate network elements (decomposed model) or by one network element (integrated model).

The decomposed model is illustrated in Figure 5, where the IP media server handles RTP flows, and the application server manages the SIP flows. The integrated model example in Figure 6 shows media gateways or session border controllers handling both RTP and SIP flows.

The decomposed model allows service providers to easily scale SIP and RTP traffic processing independently and avoid burdening an edge element, like an SBC, with two workloads. Alternatively, the integrated model does not require an extra network element

(i.e., signaling control), which can simplify network deployment, presuming the edge element has sufficient computing capacity.

There is industry consensus that SBC resources are best applied to border gateway functions, as confirmed by IMS architecture, the same IMS standards do not specify transcoding and media conditioning as an SBC function. Although an SBC occupies an optimal network location for media conditioning, including media conditioning in an integrated SBC product presents some design and economic drawbacks:

- SBC hardware platforms are primarily designed for signaling security applications and network connectivity, and they are typically not well-suited for media processing.
- SBCs often lack the computing resources to cost-effectively scale transcoding and media conditioning functionality, a burden that can lower their capacity and increase price per port.

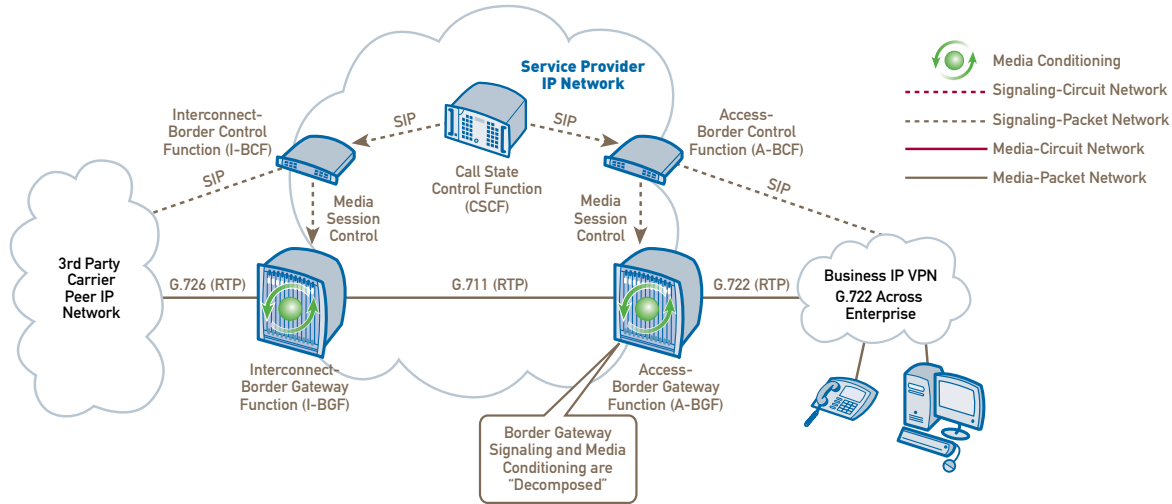


Figure 7. Media Conditioning at the Edge using 3rd Party Call Control (3PCC)

Alternatively, the Border Gateway Function could be decomposed, which separates signaling gateway and signaling security functions, performed in a Border Control Function (BCF), from media encryption and media conditioning in the decomposed BGF. Figure 7 shows how this implementation starts to resemble the 3PCC approach applied to the access edge, for both Access Border Gateway Functions (A-BGC) to enterprise IP VPNs, and Interconnect Border Gateway Functions (I-BGF) with 3rd party carrier peer IP networks.

A key benefit of a decomposed border gateway architecture is the independent scaling of border gateway and border signaling elements. This allows the network architect increased flexibility in designing a cost-efficient access edge architecture, while providing the signaling network more granular control of media conditioning.

In addition, by decoupling the border signaling requirements from the BGF, the remaining media plane processing requirements in the BGF increasingly align with the product specifications of an MRF. Therefore, MRF or IP media server products, which are already optimized for fast, efficient IP packet processing, will see an increasing role for applying media conditioning in access edge applications as well.

Media Conditioning Placement

Criteria	Preferred Transcoding Location	Comments
Reduce edge element complexity	Core	Edge elements are simpler and more cost efficient, lowering CapEx and OpEx
Increase multi-service media processing capabilities	Core	Centralization of media conditioning resources with multi-service media processing creates application synergies and economies of scale
Decrease transcoding delay	Core	Fewer transcoding steps minimizes delay and improves voice quality
Reduce network backhaul bandwidth	Edge	Many calls never need media processing services of an MRF, therefore forcing all calls into the core for media conditioning needlessly increases backhaul bandwidth
Enable codec normalization	Edge	Transcoding is performed before streams enter the core, providing efficient codec reduction
Eliminate application servers for media conditioning call setup	Edge	Controlling elements are not required for inline edge transcoding applications
Increase edge element scalability	Edge	3PCC is optional for edge-based deployments

Table 2. Comparison of Edge and Core Deployment Scenarios

In summary, Table 2 examines various criteria when deploying media conditioning at the service provider network edge or in the core network.

Radisys Solutions for Media Conditioning

Radisys is a global leader in carrier-class IP media processing with its existing IP media server products. Radisys Conveda Media Servers already support extensive IP media processing features for a broad range of IP service applications including multimedia conferencing, ringback tones, unified communications, IP contact centers and Interactive Voice and Video Response (IVVR applications). In addition, Radisys Conveda media servers support a comprehensive media conditioning solution with both automatic transcoding and voice quality enhancements as integrated components of existing IP media processing features. The carrier-class Radisys CMS-9000 (Figure 8) delivers industry-leading performance and capacities—up to 22,800 RTP media streams in a 13 rack unit equipment chassis with extensive fault-tolerance capabilities and features.

Radisys Conveda Media Servers are designed and well-suited for 3rd Party Call Control (3PCC) media conditioning requirements. While the Radisys Conveda CMS-9000 Media Server is ideally suited for large capacity service provider core media conditioning, the more economical Conveda CMS-3000 Media Server or Conveda Software Media Server might be good choices for access edge media conditioning using 3PCC requiring smaller capacities.

Media conditioning can be controlled using a feature-rich SIP with Media Server Markup Language (MSML—RFC 5707) interface. Alternatively, Radisys Conveda media server also support RFC 4117 for two-party call transcoding applications.

While performing multi-service media processing and media conditioning in the same core network element offers many architectural benefits, some carriers might instead be looking for an economical, dedicated, transcoding-only solution. For these requirements, Radisys now offers a cost-effective Transcoding Processing Card (TPC-I), which could also be hosted in a Radisys Conveda CMS-9000 media server system.



Figure 8. *Radisys Conveda CMS-9000 Media Server*



Figure 9. *Radisys Promentum® ATCA application-ready platforms for BGF, SBC, and video optimization applications*

The TPC-I card is ideal for service provider core network deployments, as well as decomposed access edge deployments where RFC 4117 3PCC is available.

For telecom equipment manufacturers building Integrated Border Gateway Function (BGF), Session Border Controller (SBC), and network elements for video optimization, Radisys offers our application-ready Radisys Promentum® family of AdvancedTCA® products (Figure 9). Radisys offers a complete portfolio of ATCA chassis systems, packet processing cards, CPU blades, and switching blades, along with pre-integrated middleware solutions, to enable telecom equipment manufacturers to rapidly develop integrated BGF, SBC, or video optimization networking products.

Exceptional Media Processing

For over thirty years, transcoding has been a critical technology for bridging TDM- and IP-based networks. Moving forward, it is playing a vital role in delivering new services, while optimizing network bandwidth in modern IP telecommunication networks. At the same time, the scope of media stream processing is reaching well beyond transcoding, and today it includes other media conditioning functions such as voice quality enhancement, video transrating and other bearer channel adaptation features.

Helping carriers deliver cost-effective and higher quality services, Radisys offers leading-edge media processing hardware and platforms for media conditioning applications. With over twenty years of experience in the telecom industry, Radisys is at the forefront of IP media processing and network architecture design.

The Radisys logo, consisting of the word "radisys" in a lowercase, sans-serif font, with a registered trademark symbol (®) to the upper right.

Corporate Headquarters

5435 NE Dawson Creek Drive
Hillsboro, OR 97124 USA
503-615-1100 | Fax 503-615-1121
Toll-Free: 800-950-0044
www.radisys.com | info@radisys.com



©2011 Radisys Corporation.
Radisys, Trillium, Continuous Computing and Conveda
are registered trademarks of Radisys Corporation.
*All other trademarks are the properties of their respective owners.
January 2011