radisys®

IP Media Servers for Next-Generation Contact Centers

Overview

Next-generation contact centers deliver services across multiple modes of communication: voice, fax, e-mail, instant messaging, web, images, and video. Combining all these communication channels over a ubiquitous, cost-efficient IP network provides rich new features, while simultaneously lowering telecommunication costs.

The enabler is Voice over Internet Protocol, a technology that converts voice into IP packets that are transmitted, alongside traditional data packets, on the IP network. This white paper explains how moving from circuit-based voice networks to VoIP technology significantly improves customer service and productivity by exploiting open computing systems, standards-based interfaces, and decomposed functional architectures.

A fundamental element in the infrastructure of next-generation contact centers is the IP media server, which processes and integrates real-time audio, video, and fax media streams for hundreds of call agents. This white paper also describes the deployment of IP media servers in next-generation contact centers, their capabilities, and their benefits to contact-center operators, hosted service providers, system integrators, and independent software vendors (ISVs).

CONTENTS

Greater Contact Center Efficiency

Contact centers are implementing cost-effective services and features with unified IP networks that simultaneously carry voice calls using VoIP on the same IP network, rather than traditional solutions where voice calls are carried on separate circuit-switched networks.

This integration of voice and data can increase the efficiency of many existing capabilities, while introducing new features in contact center functions in the areas of agent interaction, inbound call processing, and outbound call processing, as summarized on the following page.

Benefits of VoIP

Many contact centers are switching to VoIP and open computing architectures in order to reduce cost and increase flexibility. The savings are significant because voice calls routed over the Internet using public IP networks typically cost much less than paying service providers to use their circuit switching networks. And developers have more opportunities to create highly customized services that combine different media types such as voice, video, and data. This flexibility stems from using a common infrastructure to process voice traffic and data/computing applications, which gives developers access to more media at one time.

Many of today's VoIP solutions are based on open system components interconnected using standards-based interfaces, and this opens the door for more technology vendors offering "best-in-class" system components, compared to a handful of suppliers offering closed, proprietary platforms. As a result, contact centers can take advantage of an expanded ecosystem with competitively-priced products to reduce capital expenditures and ongoing operating costs.

Efficiency Opportunities

Improved Interaction between Customer and Call Agent	
Agent Interaction	Using VoIP and IP media processing, contact center applications can assign and control a "personalized mixer" for each call agent in the contact center. The application can then connect customers waiting in a queue to the call agent's mixer, add listen-only connections for supervisors, establish "whisper" connections from the supervisor to the call agent only (i.e. so customer does not hear "whisper" conversations), as well as links to conversation recorders. This flexibility permits contact center applications to customize the call agent interaction in terms of media, modes and participants.
The Right Agent or Supervisor on the Call	
Agent Interaction	Seamlessly transfer a customer to specialized agents or the supervisor faster, which shortens call duration and increases customer satisfaction.
Richer Automatic Speech Recognition (ASR)	
Inbound	Deploy speech recognition to 1) improve call routing in automatic call distribution (ACD) applications, 2) capture commands and selections in interactive voice response (IVR) applications, and 3) facilitate speech-to-text conversion.
Accelerated IVR Development Using Text-to-Speech (TTS)	
Inbound	Avoid the effort and logistics of pre-recording audio prompts, menu choices and announcements by using Text-to-Speech (TTS), which converts a text string in an IVR script to audible speech.
Wide Call Agent Distribution	
Inbound	Efficiently install contact centers in remote locations, extend services to new customers and benefit from lower contact center costs.
Embedded FAX	
Inbound Outbound	Automatically detect and save an incoming fax to a storage device. Automated outbound fax capabilities are much more productive then the old-fashioned approach of printing documents and cover page, going to the fax machine and inserting, sending and confirming a successful transmission.
Auto-dialer Application Support	
Outbound	Improve call progress analysis and follow-up action after determining whether the called number is a busy line, a fax machine, an answering machine or a human.
Quick Connect	
Outbound	Reduce abandon rates by connecting the call with the agent prior to connecting with the customer, thus eliminating the lengthy pauses associated with dialer-generated calls that many customers associate with incoming sales.
Video Interaction with Call Agent	
Agent Interaction	Integrate 'How to' videos to explain complex concepts to online customers, which saves agent time and provides a detailed response to customer inquiries.
Agent Interaction	Conduct video contact calls (requires compatible video call software at both customer and call agent terminals).
Agent Interaction	Seamlessly switch between voice and video communication modes, such that an audio-only call can be upgraded to a video call in mid-session.

Traditional Contact Center Solutions

Starting with Alexander Graham Bell, voice calls were carried over time-division multiplexed (TDM) circuit-switched networks. This technology has had a long run and still makes up a significant portion of today's contact center telecommunications infrastructure. Figure 1 highlights the historical differences between voice and data networks, which create challenges for developers who are integrating these environments for contact center applications.

The top half of Figure 1 illustrates a traditional voice network. In this example, a phone or a fax machine instructs the Public Switched Telephone Network (PSTN) to place a call by dialing digits. These digits tell the PSTN how to route the call to the call center through various switching elements. The call center telecom network also uses circuit-switched technology, which is deployed in smaller scale equipment such as private branch exchanges (PBXs) and key systems for smaller locations. Larger multi-site contact centers may have dedicated voice trunks between locations, enabling the PBX equipment to implement call transfers and load balancing between locations.

The underlying technology used for point-to-point calls, whether across the PSTN or call center internal voice networks, is a dedicated end-to-end circuit. Although PSTN circuits actually work quite well, they use bandwidth very inefficiently. A voice call requires only a 64 kilobit per second (Kbps) channel; however, it's usually transmitted over a circuit capable of carrying many megabits of data, which represents a tremendous amount of underutilization. This explains why traditional circuit-switched infrastructure, in today's era of ever increasing broadband communications technology, is becoming increasingly inefficient and expensive to operate.

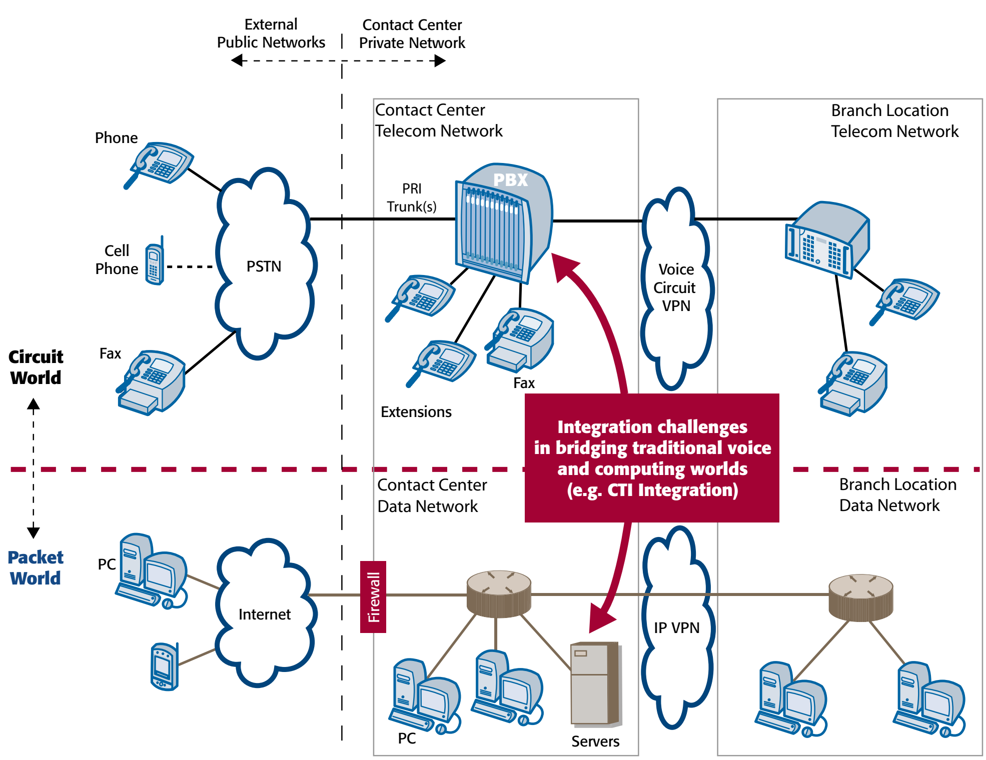

Figure 1. *Two Worlds: Circuit and Packet*

Early Efforts in CTI Integration

One of the first computer telephony integration (CTI) innovations for inbound contact centers was the "pop-up screen." The phone number digits of the incoming call were routed from the circuit world (top of Figure 1) to the packet world (bottom of Figure 1). This linkage supported database queries for customers' account information, which was displayed on the call agent's screen when the call arrived. Early CTI efforts provided immediate business value and improved customer service, but the underlying differences in circuit and packet technologies, protocols and interfaces made these basic capabilities challenging and expensive to implement.

Getting to VoIP

VoIP technologies allow real-time voice and video communications to be handled in the exact same technology and networking environment as data, immediately reducing, and possibly eliminating, the classic barriers to CTI. By supporting voice, data and real-time video on a single broadband IP connection, VoIP technology is enabling cost-efficient deployment models like "work-at-home" agent pools and offshore contact centers.

In fact, it could be argued that VoIP technology was the catalyst driving the ongoing trend towards offshore contact centers. VoIP long-distance backhauls are substantially lower cost than circuit-based long distance trunks and thus, making remote offshore contact centers financially viable.

Next-Generation IP Contact Center Solution

Possibly a more important and broader benefit of VoIP systems are their underlying open architectures with functionally separate components that are interconnected using standards-based interfaces.

Case Study: Contact Center Systems Integrator

With the assistance of a systems integrator, a large government tax agency collapsed nine contact centers to a single centralized IP-based contact center (IPCC). Today, the agency services inbound calls and when at full capacity, the facility will accommodate 2,000 call agents. In addition to consolidating the contact centers, the agency sought to automate call screening, route customers to the appropriate agent faster and maintain records of calls as mandated by the government.

Key Features: The systems integrator designed a front-end IVR application to capture customer information and route calls to the proper department with Music-on-Hold until an agent is free. The responding call agent is set up in a personalized call agent audio mixer, and new customers are added or removed from the conference. When a customer is added to the conference, recording commences and lasts until the call ends, and then the customer is removed from the conference. Conferences are recorded on a per transaction basis and stored on centralized hard disk drives.

Implementation: After evaluation of various competing products, the systems integrator selected the Radisys Convedia CMS-9000 Media Server because of its advanced personalized audio mixing capabilities. Contact center application servers control the media server using Session Initiation Protocol (SIP) and Media Sessions Markup Language (MSML), since this combination handles both inbound IVR dialog scripting as well as personalized call agent audio mixing. This VoIP system provides the contact center a cost reduction over TDM-based telephony and new value-added features supported by IP-based infrastructure.

Future: The government agency is planning to implement an IVR-enabled voice mail feature to facilitate callbacks to clients who don't want to wait for an available agent. They will also introduce outbound calling by leveraging Convedia Call Progress Detection capabilities for contacting the customers. Down the road, the agency will develop "Manager to Call Agent" coaching and mentoring, to augment traditional training, using the expanded Convedia Media Server's personalized mixing capabilities.

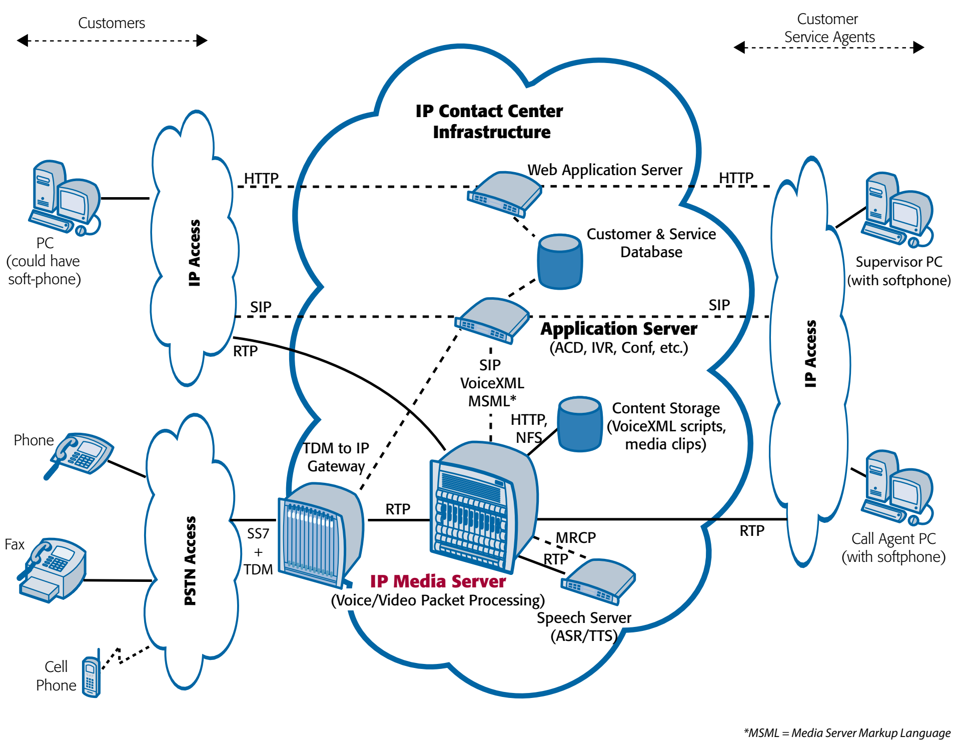

Figure 2. *Next Generation IP Contact Center*

IP-based systems typically separate functions like application and signal processing from media processing, which results in open architectures built upon scalable 'best-of-breed' components that are interconnected using standards-based interfaces. For example, the next generation IP contact center illustrated in Figure 2 employs gateways, web servers, application servers and IP media servers that work together using the open standards. These standards include Session Initiation Protocol (SIP) for signaling; VoiceXML for IVR dialog scripting, Media Sessions Markup Language (MSML) for even richer IVR and conferencing feature control; Media Resource Control Protocol (MRCP) for speech server control, and Real-Time Transport Protocol (RTP) for the audio and/or video media packets.

In this architecture, the application server is where the ACD, IVR, and agent conferencing applications run (including end-to-end signaling control and call flow logic). The application server then controls the IP media server to perform the required real-time media processing on the RTP audio and video media streams. The separation of call control and media processing also makes it easier to scale the components independently, while allowing multiple contact center applications to share the same IP media server, which further improves utilization and efficiency. This architecture also benefits and accelerates service introduction, as new service capabilities are often isolated changes or additions to the application server.

The Use of IP Media Servers in Contact Center Applications

The IP media server is the workhorse for supporting real-time processing of voice and video media streams in an IP contact center solution. IP contact center applications, including automatic call distribution (ACD) applications, interactive voice response (IVR) dialog processing, recording and call agent audio mixing, all share and control the diverse audio and video RTP media processing capabilities of the IP media server. The applications typically use open standards-based control protocols such as SIP and extensible markup language (XML) scripting; VoiceXML for IVR dialogs; and MSML for feature-rich conferencing control. The IP media server interprets the command requests and then terminates or originates the RTP audio and video media streams to complete the media processing request on behalf of the application server.

IP media servers deliver a common set of real-time RTP media processing capabilities across different contact center missions, including inbound call processing, call agent interaction and outbound call. Some of the key features and capabilities are described in this section:

Inbound Call Processing

Automatic Call Distribution (ACD) applications generally capture information from the caller to determine the most appropriate agent pools for call processing. The IP media server, under the control of the ACD application server, can play an audio prompt such as "Please press "1" for English and "2" for Spanish." After the caller presses a phone button, the media server collects the DTMF digits pressed by the caller, reports the collected digits to the application server, which can then decide how to route the call either to the English-speaking or Spanish-speaking agent pools.

Case Study: Hosted Contact Center Service Provider

Supporting hundreds of customers in Europe, a hosted contact center service provider is migrating to a next-generation platform with VoIP capabilities to offer enhanced customer service capabilities. The service provider expects the unified IP-based infrastructure to create a competitive advantage by lowering cost and offering new features to customers.

Key Features: The new hosted contact center platform will now offer "click to out-dial conference" capabilities to rapidly initiate conference calls from within the web-based contact center application environment. Using VoIP audio conferencing delivers communication cost savings, while the web-based interface is designed for ease-of-use for rapid adoption of the new out-dial conferencing conferencing. The service provider will also enhance its existing web-based contact center software to allow the moderator to see who is on the call, along with conference call controls (mute, accept, drop).

Implementation: The service provider is developing its own conferencing application based on an open source application server platform. Rapid integration between the application server and media server was achieved using open standards-based control protocols. Conferencing application logic, running on the open source application servers will control the Radisys media servers to terminate and mix the audio participant call legs in a unified IP-based call center architecture.

For IVR applications, generalized dialogs are written as XML scripts and processed by the IP media server. These scripts can define complex IVR dialog trees that play a series of prompts or menu choices and then collect and report the digits selected. Prompts may also insert dynamic "variables" at run time, like the current time: "The time is <hour>:<minute>."

IVR dialogs can also employ Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) capabilities supported by the IP media server. Rather then typing in a digit on the keypad for DTMF detection, IP media servers, in conjunction with ASR technology on an adjunct speech server, allow the caller to say "1" or "yes" to go through the IVR dialog. Instead of recording a library of audio prompts, TTS permits the IVR dialog designer to code a "text string" that is interpreted by the speech server to create an audible RTP media stream. TTS eliminates the need to make time-consuming recordings and provides a means to change dialogs quickly.

If the inbound call is actually a video call, the IP media server can support many of the same capabilities, but in a video environment. For example, the server may play a video announcement, video menu prompts or simply provide instructions to the caller in a video format.

And if an inbound call is a fax machine, the IP media server can receive and store the fax document.

Call Agent Voice and Video Interaction

After the call is pre-processed using the ACD and IVR technologies previously described, the call is placed in an agent queue. The IP media server can play announcements, advertising or music-on-hold while the caller is waiting to be connected.

Once an actual connection is made between the caller and the call agent, the IP media server provides a wide range of personalized audio mixing capabilities. Audio mixing supports a flexible contact interaction environment, as shown in Figure 3.

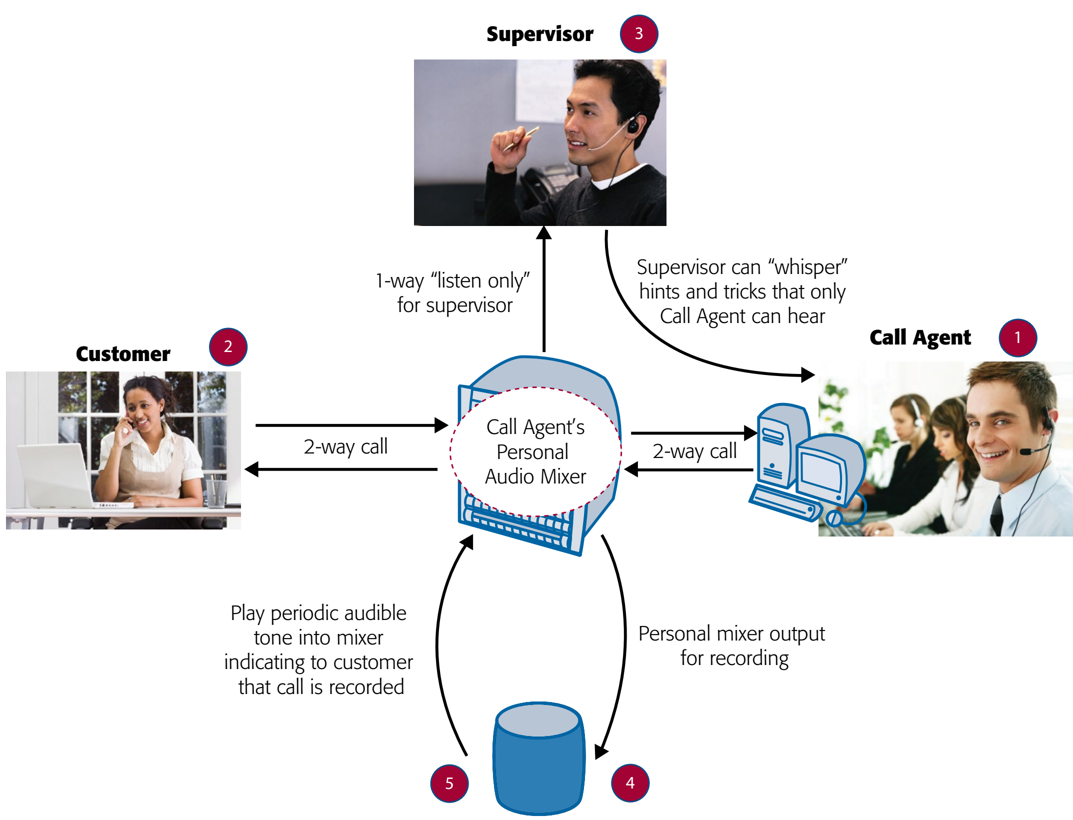

Figure 3. *Call Agent Personalized Mixing*

1. Call agents are associated with their own personalized audio mixer for the duration of a call or possibly for their entire work shift.

2. New callers (customers) are then connected into the mixer to begin their conversation with the call agent.

3. The supervisor is added to the same mixer with a "listen only" connection. The supervisor can overhear the call and "whisper" hints, tips or tricks to the call agent without the customer hearing this side of the conversation.

4. A recording function is added to the mixer, so the entire conversation is recorded for archival purposes. Maintaining a recording is often a legal requirement for stock trading and government applications.

5. A periodic audible tone (for example, a subtle "beep" played into the mix every 15 seconds) can also be inserted to alert the customer that the conversation is being recorded.

Similar personalized mixing capabilities may be performed using video calls and video conferencing. Although video technology is available for contact center solutions today, the wide-spread use of real-time video communications in contact center applications hinges on a pervasive adoption of standards-based video capabilities by end-user terminal devices.

Outbound Call Processing

Facilitating the design of IP-based outbound call center and out-dialer applications, IP media servers support Call Progress Tone Detection and Call Progress Analysis. This allows the IP media server port to "listen" to the call progress tones when an outbound call is initiated and perform the following actions:

1. If the called number is busy, the call is dropped and the number is redialed at a future time.

2. If the call reaches a fax machine, a pre-recorded fax document is sent.

3. If the call connects to an answering machine, a pre-recorded announcement is played. TTS technology can be used to dynamically generate announcements using variables: "This is a message from your physician, <Dr. Smith>, on <date>, <time> requesting that you…"

4. Ultimately, the objective is to connect with a live human. In this case, the IP media server will immediately notify the application server to insert the call into a call agent queue. While the caller is waiting, the IP media server can play music-on-hold, announcements, or even an audio/video advertisement.

Media Processing Delivered by IP Media Server	*VoiceXML Script Processing*	*MSML Script Processing*	*Play Audio*	*Record Audio*	*Personalized Audio Mixing*	*Collect/Report DTMF Digits*	*ASR (Convert Speech to Text)*	*TTS (Convert Text to Speech)*	*Call Progress Analysis*	*Detect Fax*	*Store Incoming Fax to Storage*	*Send Outgoing Fax from Storage*	*Play Video*	*Record Video*	*Video Switching/Mixing*	*Transcoding (If Required)*
Inbound Call Processing																
Interactive Voice Response																
Play Announcement/Prompt																
using pre-recorded audio/video clip	✓	✓	✓													
by generating speech from text string in XML script	✓	✓						✓								
Collect User Response																
by detecting digitis	✓	✓				✓										
by detecting speech commands	✓	✓					✓									
collect voicemail message	✓	✓		✓												
Inbound Fax Processing																
If Incoming Fax -> Receive and store fax message		✓								✓	✓					
Call Agent Voice and Video Interaction																
Personalized Audio Mixing		✓	✓	✓	✓											✓
Personalized Video Mixing		✓	✓	✓	✓								✓	✓	✓	✓
Outbound Call Processing																
Call Progress Tone Detection and Processing																
If Busy -> Try again later			✓						✓							
If Answering Machine -> Leave a message			✓	✓					✓							
If Human -> Connect call to agent			✓						✓							
Outbound Fax Processing																
If Fax Machine -> Send a fax			✓						✓	✓		✓				

Figure 4. *IP Media Processing Reuse Across Contact Center Applications*

Media Processing Reuse—The Key to Good IP Media Server Design

A good IP media server is designed in such a way that individual RTP media processing capabilities can be grouped together and reused to fulfill the large variety of media processing requests from the contact center application servers. Figure 4 summarizes the various media processing capabilities of the IP media server, and how they are invoked and reused across inbound call processing, call agent interaction and outbound call processing scenarios.

Benefits in the Contact Center

IP media servers are a fundamental component in many VoIP installations, and their importance is growing in IP-based contact center infrastructures. In addition to enabling a host of features and capabilities, IP media servers also deliver some macro benefits to hosted service providers and vendors. These benefits come from open standards, interoperable components and a large ecosystem that drives lower cost and greater vendor choice.

Radisys Convedia Media Servers

Radisys Convedia media server solutions offer powerful media processing capabilities to support IP Contact Center applications. Convedia media servers in a next generation VoIP contact center solution overcome the complexity of circuit-switched voice networks and CTI technology by supporting all media types over a common IP infrastructure.

- *Convedia CMS-9000 Media Server*
  The carrier-class CMS-9000 delivers industry-leading performance and capacities up to 24,000 RTP media streams in a 12 rack unit equipment chassis with extensive fault-tolerance capabilities and features. This solution is ideally suited for high capacity IP contact centers supporting hundreds of call agents.
- *Convedia CMS-3000 Media Server*
  This entry-level, hardware-based, one rack unit network appliance is based on the same digital signal processor (DSP) architecture as the CMS-9000. It's the ideal product for smaller contact centers or development environments.
- *Convedia Software Media Server*
  The new Convedia Software Media Server for Linux*-based computing platforms offers the same media processing capabilities and SIP control interface behavior as our DSP-based hardware products. The Convedia Software Media Server can be scaled for different capacities and performance requirements. It runs on RedHat* Linux platforms ranging from commercial-off-the-shelf Linux appliances to blade servers, and up to AdvancedTCA platforms. It's designed for CPU-intensive IP contact center applications requiring VoiceXML-based IVR dialogs and record/playback capabilities.

**Other names and brands may be claimed as the property of others.*

Macro Benefits Derived from IP Media Servers

Benefits for Hosted Service Providers

- Leverages proven technology
- Accelerates service innovation
- Reduces cost

Benefits for Solution Vendors

- Decreases time to market
- Increases ability to meet future requirements
- Eases integration by using open-standard protocols

Radisys Convedia CMS-9000 Media Server

Radisys Convedia CMS-3000 Media Server

Radisys Convedia Software Media Server

A Path to Greater Contact Center Efficiency

VoIP and open systems are fuelling a new round of innovation in contact center features and capabilities. Contact centers migrating from traditional circuit-switched communications to VoIP technologies are creating opportunities to significantly reduce communications costs and increase efficiency. This transition is also changing the way system integrators, hosted contact center service providers and independent software vendors approach their businesses.

One of the key components in the modern next-generation contact center is the IP media server, which delivers real-time audio and video packet processing for a complete range of contact center application requirements, including ACD, IVR, feature-rich agent conferencing modes, embedded fax and video. Radisys offers a complete product portfolio of IP media processing products and technology, from entry-level software-based media servers to the largest carrier-class hardware media servers, all delivering operational and economic benefits for contact center operators and their communication solution vendors.

Case Study: Contact Center Software Vendor

An IP contact center software vendor is developing a next-generation IP contact center platform using a single IP-based network infrastructure. The key objective was to rapidly introduce a new feature-rich solution offering with attractive economics through the integration of third-party best-in-breed technologies. The vendor wanted a flexible and scalable platform that can be expanded to meet future customer requirements.

Key Features: To accelerate time-to-market for the new solution offering, this vendor decided to integrate Radisys media servers into their solution, rather then developing their own real-time IP media processing subsystem. This next-generation contact center application uses Session Initiation Protocol (SIP) and Media Sessions Markup Language (MSML) for a standards-based control interface to the Convedia media server, providing a feature-rich inbound and outbound call center environment for VoIP call processing. Through the use of a common IP-based infrastructure and open standards, this vendor's solution also easily integrates with third-party technologies such as Automatic Speech Recognition (ASR), Text-to-Speech (TTS) and third party applications, and includes recording systems, customer relationship management (CRM) and enterprise resource planning (ERP).

Implementation: The contact center solution integrates the Radisys Convedia CMS-3000 product for all real-time VoIP audio packet processing using a SIP/MSML control interface.

radisys.

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