IP Centric Conferencing

A RADVISION White Paper

Contents

Introduction

During the latter part of the 80's businesses began conducting videoconferencing using room systems over proprietary networks. Later on, the H.320 standard supported by ISDN was adopted as the videoconferencing network of choice. Both the proprietary and the ISDN systems required substantial investment and high management and operation costs.

The dramatic growth of the Internet and IP networks, together with the universal adoption of standards for multimedia conferencing over packet networks, have caused businesses to switch to voice and videoconferencing over IP networks. The reduced cost of videoconferencing equipment, affordable desktop stations, cheaper bandwidth, ever-improving video quality and the availability of enhanced video-related services are driving businesses to adopt **IP Centric Conferencing** as a cost-effective means to inter- and intra-organizational communications.

Implementing IP-centric voice and videoconferencing solutions is less complex than implementing ISDN-based solutions. Nevertheless, there are still careful choices that need to be made to ensure that the necessary infrastructure is in place to support IP Centric Conferencing or in selecting services from a Conference Service Application Provider. This paper aims to clarify some of the issues and technologies associated with IP Centric Conferencing.

IP Centric Conference Benefits

In order to keep pace with today's dynamic, global marketplace, businesses, organizations, and institutions of all types and sizes can benefit from IP Centric Conferencing Global IP-centric visual collaboration tools removes the barriers of geographic separation and brings people together. Improved communications facilitates project management, speeds development and manufacturing cycles, improves sales contact, and enhances service responsiveness. IP Centric Conferencing is an excellent mechanism for managing change, increasing business competitiveness, reducing cost, and enhancing revenue opportunities.

Industries especially suited for IP Centric Conferencing applications include large, multi-location, global Fortune 1000 companies, government agencies, distance education both in business and school markets, and telemedicine.

What is IP Centric Conferencing?

IP Centric Conferencing is real-time, multimedia voice, video and data conferencing over IP networks. This section introduces you to some basic IP Centric Conferencing terms and concepts and provides a brief look at the history of this industry.

Note: H.323 is an ITU standard and is currently the most prevalent standard used for IP conferencing. There are new and emerging standards and these are discussed in a separate section below. In this white paper, unless specifically stated to the contrary, H.323 terms and concepts have been used.

Terms and Concepts

The main types of conferencing on IP are:

- **IP point-to-point conferencing** -- real-time communication between two endpoints over an IP connection.
- **IP multipoint conferencing** -- three or more endpoints communicate in real-time over an IP connection.
- **Voice conference** -- interactive voice-only communication between two or more endpoints
- **Videoconference** -- two-way real-time audio and visual communication between two or more points.
- **Data collaboration** -- conference participants view shared files/applications.
- **Data sharing** -- multiple conference participants can actually work in an application running on one of the participant's machines.

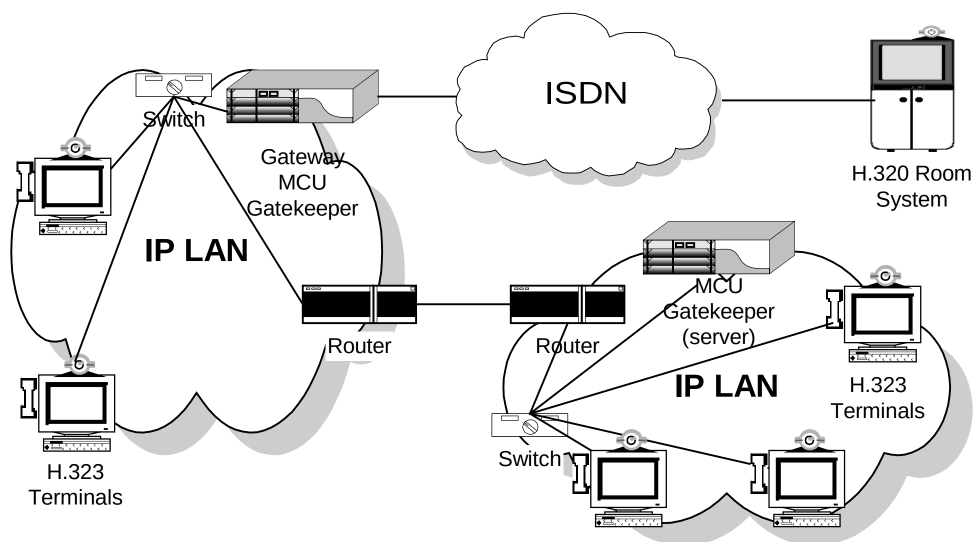

Figure 1 – An IP Centric Conference

The main networking devices that are used during a voice or video conference are: (i) *terminals* that are voice- and, optionally, video-enabled, (ii) *multipoint control units* (MCU) that control multiple users and handle media processing, (iii) *gateways* that convert media between legacy circuit-switched networks and IP networks, and (iv) *gatekeepers* that process calls by tracking and managing call progress as well as handling conversion between addressing schemes.

History

Up until a few years ago, conferencing involved separate networks. A data network was needed for transferring data files; PSTN was used for talking by telephone; ISDN lines were needed for videoconferencing using room systems.

The ability to use packet-based IP networks for real-time communication is fueling interest in multimedia communication that supports converged voice, video and data communications. An increasing number of carriers and service providers are introducing new value-added converged services. With promises of ever-increasing and affordable bandwidth, IP is much more versatile and lends itself to cost effective voice and videoconferencing with the option of visual data collaboration.

Technology Choices

In order for participants to communicate over an IP network, all networking devices and terminal equipment need to be interoperable or to “interwork”. Standards help to achieve interoperability. Global IP networks will most certainly be built around multiple IP communications standards (such as H.323, SIP, MGCP and MEGACO.) Inter-standard or multi-protocol bridges or gateways help to achieve interworking.

This section discusses the IP standards related to multimedia conferencing and then details the key devices needed to establish and sustain a videoconference.

Standards

The ITU and IETF are the two dominant standards bodies that establish and control the standards for multimedia communications. The main standards are:

- **ITU-T H.323** -- a complete suite of protocols that facilitate multimedia communication over IP packets. Incorporated within this standard are video codecs such as H.261and H.263, audio codecs, such as G.711, G.723 and others.
- **ITU-T T.120** -- contains a series of communication and application protocols and services that provide support for real-time, multipoint data communications. These multipoint facilities are important building blocks for a whole new range of collaborative applications, including desktop data conferencing, multi-user applications, and multi-player gaming.
- **IETF MGCP** and **MEGACO** -- address the requirements of production IP telephony networks that use decomposed gateways. The ITU-T and IETF have recently begun to collaborate in the specification of a converged protocol called MEGACO/H.248.
- **IETF SIP** -- a signaling protocol that focuses on session initiation, modification and termination.
- **IETF RTP/RTCP** -- the predominant standard transport of any real-time stream of data such as audio or video.

H.323 was the first standard developed to enable real-time communication over packet-based IP networks and was instrumental in driving the transition from ISDN to IP. Today, H.323 is the most widely deployed standard for IP and is supported by a global coalition of companies working in a joint effort to provide interoperability. SIP is a new signaling protocol that is growing in popularity.

Whereas H.323 was designed initially for interactive multimedia communication, SIP was designed initially for voice over IP (VoIP) applications. MGCP and MEGACO/H.248 are complementary protocols to both SIP and H.323. MGCP and MEGACO/H.248 are “internal protocols” designed specifically for interfacing between “intelligent” Media Gateway Controllers (MGC) and “dumb” Media Gateways.

A key point for future growth of the IP Centric Conferencing is the interworking of all these protocols allowing for seamless, end-to-end connectivity between all types of endpoint and network devices.

Network Devices

Since H.323 is currently the most widely deployed standard for IP Centric Conferencing, this section discusses network solution components in terms of H.323.

Terminals

An H.323 terminal is an endpoint that enables real-time voice or video communications with another H.323 terminal, gateway, or MCU on the network. The terminal must support basic H.323 protocols such as H.245, Q.931, RAS (for communication with a gatekeeper) and RTP/RTCP. For a full multimedia conference the terminal must have audio and video capabilities, such as support of the H.323 compliant audio and video codecs, as well as appropriate peripheral equipment such as microphones, speakers, and a video camera.

Multipoint Controller Units / Multipoint Controllers / Multipoint Processors

MCU/MC/MPs are the enablers of a multipoint conference. H.323 differentiates between a multipoint controller (MC) and a multipoint processor (MP). The MC controls conference setup including capabilities negotiations with all terminals and the opening and closing of channels for the audio, video and data streams. The MP centrally processes audio, video and/or data. The MP mixes, switches and processes the streams controlled by the MC. One MC is mandatory while one or more MPs are optional. An MP can be combined with an MC in a single network device called an MCU, or one or more MPs can be distributed over the network.

Gateways

Gateways essentially allow intercommunication between IP networks and legacy networks. They provide transcoding facilities by receiving, for example, an H.320 stream from an ISDN line, converting it to an H.323 stream and sending it to the IP network. Gateways can also perform call setup and clearing on both sides of an IP to switched-circuit connection. As many video conferencing systems are still ISDN-bound, the gateway is likely to continue to be an essential device in any IP Centric Conferencing network.

Gatekeepers

Gatekeepers “manage” the network. Besides handling address translation (translating complex IP addresses to people-friendly aliases), gatekeepers may offer an array of services such as call routing, call transfer and forwarding, line hunting, LDAP and DNS support, CDR generation (for billing) and so on. One or more gatekeepers may reside anywhere on the network, fully integrated into another networking device (such as a gateway) or operating as a standalone software application on a desktop computer.

Firewalls and Proxies

There is a risk that key information may be exposed over an H.323 network since terminals that signal each other directly must have direct access to each other’s IP address. Firewalls and proxies are used alone or together to prevent this exposure. To be effective on H.323, a network firewall should support H.323 signaling. Alternatively, proxies can limit the address information that is exposed. Typically, firewalls and proxies are used together to prevent security breaches.

Deployment Issues

There is a range of network and conferencing issues that should be considered when deploying an IP Centric Conferencing network. These considerations should help to determine whether an organization would update its existing network, create its own new IP Centric network or seek the services of a Conference Application Service Provider (CASP).

Bandwidth

A critical factor for successful IP Centric Conferencing is ensuring that there is sufficient bandwidth to maintain satisfactory levels of voice, video and data transmission. Both traditional data network bandwidth considerations and usage issues need to be taken into account when determining the required amount of bandwidth. For example:

- **Optimization** -- the use of features such as silence suppression can result in a bandwidth saving of at least 50%.
- **Overhead** -- such as Packet Overhead, where each packet that an IP network transmits has a header that requires approximately 40 bytes.
- **Usage Modeling** -- calculating the Erlang factor (the ratio of the number of conference terminals in the network to the number of conference terminals actually in use at any given time) and Busy Hour Calls (at certain times of the day more people will be in a conference than at other times of the day).

Quality of Service (QoS)

QoS relates to the guaranteed quality of the media being delivered. With traditional circuit-switch telephone networks we expect to hear what someone says immediately and without distortion. On a packet network, the guaranteed level of performance depends on a set of transmission parameters such as delay, jitter and bandwidth that is assigned to selected traffic on the network. QoS is a key factor for successful implementation of multimedia conferences over any packet-based network.

Seemingly negligible environmental and human factors often determine the success or failure of a videoconference. These factors include type of terminal, acoustic echo cancellation, lighting, camera quality, background noise, silence suppression, relative position of the camera, screen and participant, and setup time.

As bandwidth increases, the QoS should improve. In addition, three major initiatives are attempting to solve QoS issues (RSVP and DiffServ from the IETF, and 802.1p from IEEE). However, the work is not complete and the necessary bandwidth is not yet universally available, so it will be some time before QoS is fully resolved within and between enterprises.

Scalability and Expandability

IP networks are fundamentally different from ISDN networks. IP networks have a distributed and flexible architecture that spans LAN, WAN and/or the Internet. The IP infrastructure is location- and service-provider independent. The inherent scalability of IP allows bandwidth to be increased, equipment to be added and services to be improved without making any fundamental changes to the underlying infrastructure.

One of the main advantages of IP Centric Conferencing is that the communication system goes where the IP network goes. This allows companies to scale from internal conferencing over the LAN to multi-location, global conferencing over a WAN. The software driven design of the equipment and the network results in a fully scalable, location-independent, cost-effective communication system.

By definition, videoconferencing incorporates audio conferencing. Most organizations have already adopted voice conferencing as part of their corporate culture. As videoconferencing becomes more mainstream, because the IP infrastructure allows for a seamless transition from voice to video, more and more organizations will be able to expand from voice to enhanced videoconferencing services.

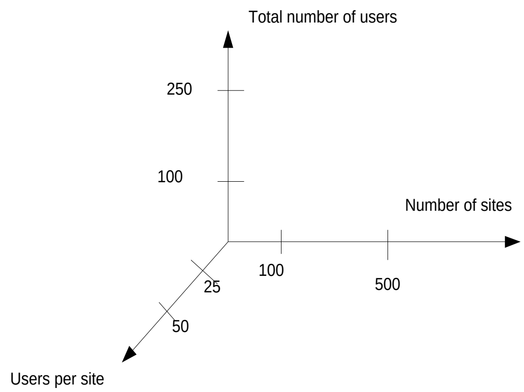

Figure 2 - Project the growth over time, in each dimension, and see whether the network and the products you use on the network will be able to scale sufficiently to support future growth

Management

The IP Centric network has it own set of unique practical management and administration issues.

Conference management facilities are required to configure the services and properties of any given conference. These can include allocated bandwidth, load balancing, cascading, supported codecs, transcoding priorities, continuous presence, video switching and far-end camera control. To compensate for the non-intuitive underlying IP address system, devices such as H.323 gatekeepers provide address translation and directory server services. Rationalized and easy-to-use dial plans further reduce the complexity of inter-zone dialing on the IP network.

A conference can be reserved, scheduled or initiated ad hoc. Once in progress, a conference chairperson monitors the conference, and may disconnect participants, invite others, focus and lock the camera on a single participant and so on.

As voice and video conferencing becomes more tightly integrated with the web and web-based applications, web-based management tools provide a comprehensive single system view of all users, sites, equipment, features and services.

Security

Firewalls and/or proxies are used to protect the network from unwanted access attempts. Firewall techniques identify and filter traffic based on source or destination address, protocol type, or IP port numbers.

In H.323, security is achieved through the H.235 protocol. H.235 provides privacy and authentication. H.235 version 2 makes it almost impossible for someone to have unauthorized access to the IP network, to listen to a conversation, or to identify the destination number of the call.

Connectivity with Legacy Systems

In order to communicate from an IP network to a legacy system two essential types of devices are needed—gateways and gatekeepers. Gateways bridge between circuit-switched networks and packet networks. Legacy networks and most telephones use digits (numbers). IP networks use complex IP addresses. Gatekeepers provide the address translation facilities and route calls coming in from the legacy network via a gateway, to the destination endpoint on the IP network.

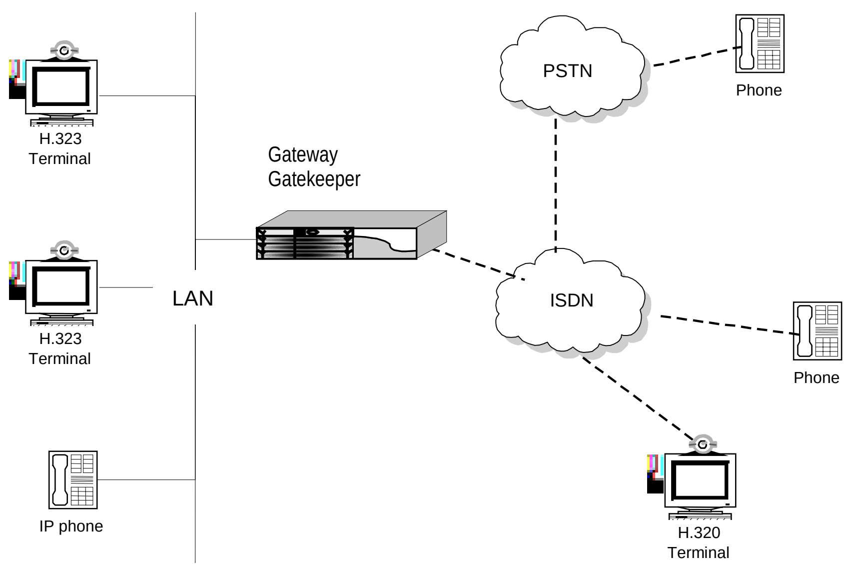

Figure 3 – IP and Legacy Systems Connectivity

Interoperability and Interworking

Network managers, systems integrators and developers, regard H.323 interoperability as the ability to incorporate any new H.323 compliant product and “drop” it into an existing H.323-ready network with minimum planning and no risk of network disruption. For developers, H.323 interoperability means that all H.323 compliant products conform to H.323 specifications and that all messages and parameters can be exchanged and correctly interpreted by H.323 endpoints to produce a predictable result.

With the emergence of new standards such as SIP, a new challenge emerges over the IP Centric network. Devices that support different standards need to communicate in point-to-point or multipoint conferences. Interworking devices such as bridges and gateways will facilitate the seamless communication between dissimilar protocols.

Media Streams

Multipoint conferencing on IP supports a number of media transmission models such as:

- Centralized multipoint conferences in which all participating terminals communicate in a point-to-point fashion with an MCU. The terminals transmit their control, audio, video, and/or data streams to the MCU.
- Decentralized multipoint conferences in which the participating terminals multicast their audio and video to all other participating terminals without using an MCU. The terminals are responsible for combining the received audio streams; and selecting one or more of the received video streams for display.
- Cascaded conferences provide an economical architecture for supporting many conference participants. In this architecture, one MCU with active participants connects to another MCU with active participants to create one large conference.

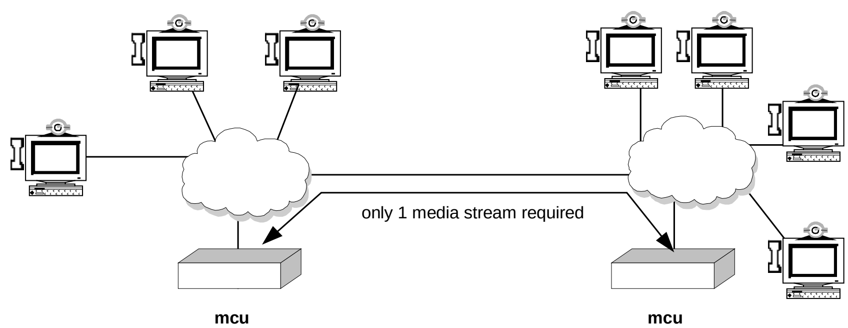

Figure 4 - A Cascaded Conference

Signaling and Media

In H.323 centralized multipoint conferences all endpoints exchange call signaling with the MCU. The H.245 control channel is opened between the endpoints and the MC (within the MCU). The audio, video and data channels are opened between the endpoints and the MP. The call signaling channel (Q.931) for call setup and teardown uses one of the well-known ports, while H.245 signaling, audio, video and data channels are dynamically negotiated between endpoints. In a Decentralized Multiple Conference the audio and video channels are multicast to all endpoints.

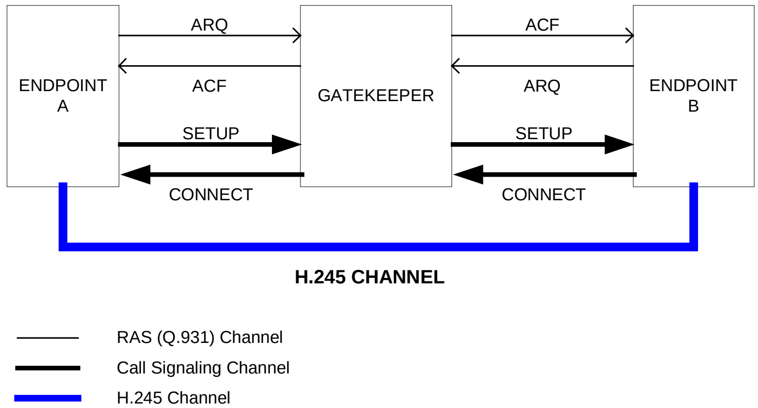

Figure 5 – H.245 Routed Call Signaling

Control signals and data are sent using reliable transport (TCP) to ensure that signals are not lost and arrive in a proper order. Because audio and video streams are time-sensitive, they are sent using unreliable transport (UDP). Even though delayed audio and video packets are dropped, this generally has a small impact on the quality of the audio and video.

Media Transcoding

Audio and video codecs are used to compress and decompress the audio and video data. Different protocols and devices support different codecs. In order for participants to speak to each other, or to see each other, the data streams usually have to be transcoded (translated) from one format to another. During call setup, capabilities are exchanged between the endpoints to establish what codecs each side supports. Gateways and some MCUs have built-in capabilities for prioritizing and performing audio and video transcoding.

Another consideration is that, until quite recently, most MCUs forced each conference participant to the lowest common denominator for call capabilities. For instance, if one participating endpoint could only send/receive QCIF calls at 128K bandwidth, all other participants in the same conference would be forced to send/receive the same. This limitation is changing as increased transcoding capabilities are being introduced into some MCUs.

Video Presentation

During a multipoint videoconference, the MCU receives video streams simultaneously from the participants. In order to make sense of the conference, the MCU needs to send out a video stream that will be received as a single picture on the participating terminals. A number of options are available. The option chosen is usually application-dependent. For example, in a distance learning conference the video image would be locked on the teacher. In a business conference, the video image switches to the most prominent (loudest) speaker. Continuous Presence is another mode of video presentation where multiple sites are mixed into different areas of the screen and viewed as a single image.

Data Collaboration

Data collaboration and sharing provides an extra dimension complementing voice and videoconferences. Conference participants can share information with fellow participants during the conference by viewing other participants' graphic presentations and slide shows, by exchanging diagrams, pictures and texts using white boards by transferring files rapidly and by participating in text chats.

In order to participate in a multipoint conference with data collaboration the terminals and the MCU must be compliant with T.120.

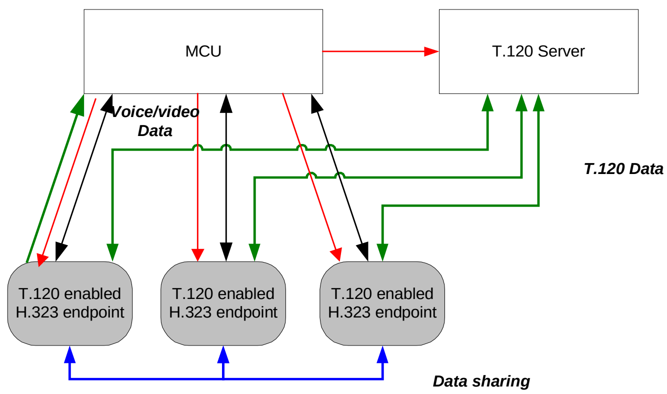

Figure 6 - Data Collaboration during a Multipoint Conference

Scheduling Issues

A conference can be ad hoc (unreserved) or scheduled. When scheduling a conference many aspects need to be coordinated–multiple sites, time zones, equipment availability, people and bandwidth. Given the complexity of this task, it is desirable to use an automatic scheduling service for IP Centric Conferencing. There are an emerging number of easy-to-use, web-based voice and videoconference solutions that offer scheduling and enhanced services such as reservations, billing and resource control across multiple sites and multiple time zones.

Ad hoc or unreserved conferencing provides a more spontaneous model. In this case, a conference identity is established and participants are invited to join the conference by simply dialing a given conference number.

A Real IP Conferencing Solution

The RADVISION viaIP platform provides a single architecture for multiple, integrated IP conferencing applications. The platform is central to the IP network infrastructure and allows any combination of one or more MCU and/or gateway boards. An NT board holds the service applications such as the Enhanced Communication Server (ECS) gatekeeper, Data Collaboration Server (DCS) and scheduling applications.

The viaIP mcu can operate in a centralized, distributed (one or more MPs), or cascading topology (two or more joined mcu's). One or more simultaneous conferences can be held on the same mcu. A configurable conference profile (service) determines the characteristics of a particular conference. These characteristics include whether the conference is voice or video, the quality of the video, the video bitrate, the video and picture formats, the maximum number of participants allowed, silence suppression support, transcoding priorities, continuous presence and support of data collaboration. Conferences can be ad hoc or scheduled. Participants dial in or are invited to ad hoc conference. Scheduling applications arrange scheduled conferences.

The viaIP mcu supports T.120 data collaboration together with the DCS, a data collaboration server.

A gateway is necessary to support point-to-point and multipoint conferences that span a PSTN network and an IP network. The most important function of the viaIP gateway is to perform all necessary protocol conversions. The viaIP gateway provides incoming call routing services such as IVR support, DID, TCS4 and default extension. Included in the many features of the viaIP gateway are far end camera control, audio transcoding, transparent support for the H.261 and H.263 video codecs, echo cancellation and RAI/RAC load balancing.

The ECS performs the essential tasks of a gatekeeper, such as address translation, and also provides a range of PBX-like and enhanced services. Services include forwarding, transfer, conference hunting, routing modes, bandwidth usage and so on. Enhanced directory services include support of LDAP and DNS. By generating comprehensive call detail records the ECS provides the information needed for 3rd party billing applications. For security, the ECS can support a proxy, such as the Cisco Proxy, and in future versions the ECS will provide authentication and integrity verification under the H.235 version 2 standard.

All viaIP products are administered through a web-based element manager. Using a web browser, from anywhere in the IP network, the network manager has full remote control and can configure, manage and monitor all elements of the viaIP.

Summary

The ability to use packet-based networks for real-time communications is fueling interest in converged voice, video and data communications. With the promise of ever-increasing and affordable bandwidth, IP lends itself to cost-effective voice and videoconferencing.

As the decision is made to implement IP-centric voice and videoconferencing solutions, careful choices must be made to ensure that the necessary infrastructure is in place to support it. Network devices and terminal equipment must interoperate and network and conferencing issues like bandwidth, quality of service, security, data collaboration and legacy system connectivity must be successfully addressed.

RADVISION's *via***IP** multifunctional platform provides a cost-effective solution for the IP communication requirements of both service providers and large enterprises. *via***IP** products integrate multimedia gateway, multipoint conferencing, data collaboration, and advanced gatekeeper intelligence into a single platform. *via***IP** is built around RADVISION's award winning, industry standard H.323 technology and provides the scalability and proven interoperability needed for delivering enhanced V^2oIP services for converged voice, video and data networks. RADVISION's market-leading IP-centric products were specifically designed to leverage the inherent advantages of IP network architectures and deliver the superior price/performance compared to legacy, ISDN-based solutions.

Glossary

Abbreviation	Description
CDR	Call Detail Report
DCS	Data Collaboration Server
DID	Direct Inward Dialing
DiffServ	Differentiated Services
DNS	Domain Name System
ECS	Enhanced Communication Server
G.711, G.723	ITU standards for audio codecs
H.235	ITU standard for security
H.261,H.263	ITU standards for video codecs
H.320	The ITU standard for videoconferencing over digital networks such as ISDN
H.323	The ITU standard for videoconferencing over packet switched networks
IETF	The Internet Engineering Task Force
ITU	International Telecommunications Union
IVR	Interactive Voice Response
LDAP	Lightweight Directory Access Protocol
MCU	Multipoint Controller Unit
RAS	Registration Admission Status protocol
RTP/RTCP	Realtime Transport Protocol/ Realtime Transport Control Protocol
SIP	Session Initiation Protocol
TCP	Transmission Control Protocol
TCS4	A Direct Inward Dialing routing method for H.320 calls.
UDP	User Datagram Protocol