IP Protocols: The Protocols that Matter For Voice, Video and Data over IP Networks H.323, SIP, MEGACO and MGCP

I. Introduction

Industry standards play a key role in driving the broadbased deployment of IP-based "converged" networks for real-time multimedia (voice, video and data) communications. There are four key industry standards for IP communications that provide enabling technology for real-time communication over packet-based IP networks. H.323, SIP, MEGACO and MGCP. Global IP communication networks of the future will consist of user devices and networking infrastructure products that are built around all of these protocols. Perhaps the most important aspect of these standards is that they were designed to facilitate "interoperability." The success of IP communication depends on providing connectivity between anyone, using any device, calling from anywhere. By building "standards-compliant" products, equipment from different vendors and networks can interoperate seamlessly.

Another important contribution of standards is that they provide a "pyramid of knowledge." It doesn't make any sense for every developer to have to reinvent the underlying enabling technology that is embodied in a standard. Nor does it make sense for every developer to have to devote people, time and money resources to become an expert in evolving standards. Instead, by licensing the underlying IP communication protocols from a third party, equipment vendors, application developers and service providers can focus their core competencies on developing innovative new products and services that provide compelling value-added functionality to an IP network.

II. What is H.323?

H.323 was the first IP communications protocol to be introduced to support real-time multimedia communication over IP networks. Today, H.323 is the most widely-deployed standard. H.323 is an "umbrella" specification developed by a consortium of computing, telephony and computer networking experts in the internationally recognized ITU (International Telecommunications Union). The standard provides a framework for developing H.323-compliant products and services. H.323 is referred to as an "umbrella" specification because it encompasses many "protocols"(other recommendations). The specification is broad in scope and covers standalone devices, embedded personal computer technology, point-to-point and multipoint conferencing. H.323 is linked to a number of communications standards for multimedia conferencing over a range of networks. The ultimate goal behind the standard is to make end-to-end interoperability a reality.

The H.32X series of protocols also includes H.320 for ISDN communications and H.324 for PSTN communications. Other recommendations under the H.323 umbrella include H.225.0 for packet and synchronization, H.245 for call control, H.235 for security, H.450 for supplementary services, H.261 and H.263 for video codecs, G.711, G.722, G.728, G.729, and G.723 for audio codecs, and the T.120 series of multimedia communications protocols.

The core-enabling building block underpinning all H.323-compliant multimedia systems is the protocol stack. The H.323 stack is the most basic, low-level software product that is the result of translating the specification into actual program code. All H.323-compliant products contain an embedded H.323 protocol stack. The specification defines four different H.323 entities as the functional units of a complete H.323 network:

Gatekeepers – Network management tools that control who gets access to which services and when. Gatekeepers also monitor service usage, network bandwidth and other network administration tasks. They perform the critical control, administration and management functions needed to maintain the integrity of enterprise local and wide area networks..

Multipoint Control Units (MCUs) - A multipoint control unit is designed to support simultaneous conferences between three or more locations. In H.323, the multipoint-session dynamics are very flexible. The standard allows for a variety of ad hoc conferencing scenarios, in addition to the traditional method of scheduled resource usage.

Gateways - Connect IP and circuit-switched networks; voice-only and/or multimedia. Gateways are essential network components for routing voice/video/data between H.323 based conferencing systems on IP networks and H.320-compliant systems over ISDN networks.

Terminals – End user devices that provide real-time, two-way communication; voice-only and/or multimedia

Developers use enabling software to build MCUs, gateways and terminals. Since the standard only provides a development framework and does not define specific implementations, developers can use core H.323 software to develop H.323-compliant products with unique value-added features.

III. Brief History of H.323

Telecommunications design specifications for circuit-switch networks evolved gradually over nearly a century, and, for a substantial portion of this time, it was with the support (and dictates) of government. With this government support, telecommunications products became extremely reliable and end-point equipment interoperability was at 100%.

The telecommunications industry and the computer industry had little interaction and even less influence over what each other was doing. In contrast to the high level of reliability/interoperability within the telecommunications industry, the computer industry became known for releasing products too early. Customers tolerated a low level of reliability and interoperability, accepting defacto standards whenever available (e.g., lowest cost of ownership). It was not until the industry-wide adoption of the ITU's H.320 standard for multimedia communications over ISDN that the computer industry had any involvement in developing specifications published by international telecommunications standards bodies.

Collaboration between telecommunications and computer industry leaders rose dramatically during the development of H.323. The result is the rapid growth of the specification that now draws upon the experience and innovation of both industries. Global adoption of the ITU-T H.323 assures developers, manufacturers and their customers' interoperability and highly functional products and services more quickly than would otherwise be possible. And the fact that H.323 promises these products and services on non-guaranteed Quality of Service packet-based networks is even more significant.

IV. What is SIP?

The Session Initiation Protocol, or SIP, is a new IETF signaling protocol for establishing real-time calls and conferences over Internet Protocol networks that is growing quickly in popularity. Each session may include different types of data such as audio and video although currently most of the SIP extensions address audio communication. As a traditional text-based Internet protocol, it resembles the hypertext transfer protocol (HTTP) and simple mail transfer protocol (SMTP). SIP uses Session Description Protocol (SDP) for media description.

SIP is independent of the packet layer. The protocol is an open standard and is scalable. It has been designed to be a general-purpose protocol. However, extensions to SIP are needed to make the protocol truly functional in terms of interoperability. Among SIP basic features, the protocol also enables personal mobility by providing the capability to reach a called party at a single, location-independent address.

V. What is the Relationship Between SIP and H.323?

Both SIP and H.323 define mechanisms for call routing, call signaling, capabilities exchange, media control, and supplementary services. SIP is a new protocol that promises scalability, flexibility and ease of implementation when building complex systems. H.323 is an established protocol that has been widely used because of its manageability, reliability and interoperability with PSTN. There is a general consensus among standards organizations, companies and technology experts that standardized procedures need to be specified to allow seamless interworking between the two protocols. Bodies such as TIPHON (ETSI), aHIT (IMTC) and IETF are working to address this topic.

RADVISION is leading the ITU-T SG16 initiative to define SIP guidelines, and is working closely with other ITU members to achieve true end-to-end connectivity between the two protocols on the same network infrastructure providing global services to end-users. To achieve an interoperable system, the following items need to be addressed:

?? Topology Definition: The architecture needs to be defined in terms of the respective entities. Call scenarios and interfaces must then be developed based on the defined topology.

?? Supported Services: Advanced services such as conference calls and supplementary services should be incorporated.

- ?? Supported "Data Capabilities": H.245 channels to Session Description Protocol (SDP) mapping are essential to respond to issue of supporting audio and multimedia support.
- ?? Protocol Versions: The protocol versions of both SIP and H.323 need to be addressed in order to define the scope of the work.

VI. SIP Architecture

SIP's basic architecture is client/server in nature. The main entities in SIP are the User Agent, the SIP Proxy Server, the SIP Redirect Server and the Registrar.

The User Agents, or SIP endpoints, function as clients (UACs) when initiating requests and as servers (UASs) when responding to requests. User Agents communicate with other User Agents directly or via an intermediate server. The User Agent also stores and manages call states.

SIP intermediate servers have the capability to behave as proxy or redirect servers. SIP Proxy Servers forward requests from the User Agent to the next SIP server, User Agent within the network and also retain information for billing/accounting purposes. SIP Redirect Servers respond to client requests and inform them of the requested server's address. Numerous hops can take place until reaching the final destination. SIP's tremendous flexibility allows the servers to contact external location servers to determine user or routing policies, and therefore, does not bind the user into only one scheme to locate users. In addition, to maintain scalability, the SIP servers can either maintain state information or forward requests in a stateless fashion.

VII. What is MGCP?

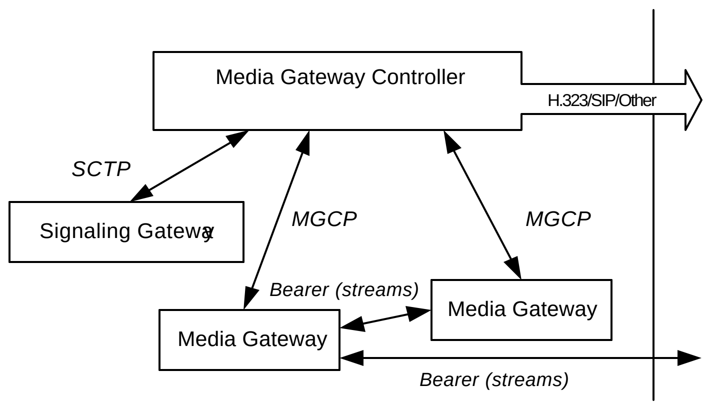
The Media Gateway Control Protocol, or MGCP, was designed to address the requirements of production IP telephony networks that are built using "decomposed" VoIP gateways. MGCP-based VoIP solutions separate call control (signaling) intelligence and media handling. MGCP functions as an internal protocol between the separate components of a decomposed MGCP-compliant VoIP gateway. More specifically, MGCP is a protocol used by external call control elements called Media Gateway Controllers (MGCs) for controlling Media Gateways (MGs). Decomposed MGCP-compliant VoIP gateways appear to the outside as a single VoIP gateway. Example VoIP gateways include:

Trunking gateways that interface the public telephone network and VoIP network
Residential gateways that provide traditional analog (RJ11) interfaces to VoIP networks
Access gateways that provide traditional analog (RJ11) or digital PBX interfaces to VoIP networks

The evolution of the MGCP specification was largely influenced by "political" conflicts between proponents of alternative proposals for decomposed gateway architectures. In particular, MGCP owes its origin to the confluence of the SGCP (Simple Gateway Control Protocol) and IPDC (Internet Protocol Device Control) protocols.

VIII. Relationship Between MGCP, SIP and H.323

MGCP is a complementary protocol to both SIP and H.323. It was designed specifically as an internal protocol between MGCs and MGs for decomposed gateway architectures. In the MGCP model, an MGC handles call processing by interfacing with the IP network via communications with an IP signaling device such as an H.323 gatekeeper or SIP Server and with the circuit-switch network via an optional signaling gateway. Using an H.323 analogy, the MGC implements the "signaling" layers of H.323 and presents itself as an "H.323 Gatekeeper" or as one or more "H.323 Endpoints". Within the MGCP approach, MGs focus on the audio signal translation function, performing conversion between the audio signals carried on telephone circuits and data packets carried over the Internet or other packet networks.



MGCP Model

IX. What is MEGACO/H.248?

MEGACO/H.248 is the official industry standard protocol for interfacing between external call agents called Media Gateway Controllers (MGCs) and Media Gateways (MGs). The standard is the result of a unique collaborative effort between the IETF and ITU standards organizations. Derived from MGCP (which, it turn, was derived from the combination of SGCP and IPDC), MEGACO draws heavily from MGCP plus introduces several enhancements. Even though MGCP was deployed first, MEGACO/H.248 is expected to win wide industry acceptance as the official standard for decomposed gateway architectures sanctioned by both the IETF and ITU. MGCP is currently being maintained under the auspices of the PacketCable™ and the Softswitch Consortium™.

MEGACO offers these key enhancements as compared to MGCP:

- ?? Supports multimedia and multipoint conferencing enhanced services
- ?? Improved syntax for more efficient semantic message processing
- ?? TCP and UDP transport options
- ?? Allows either Text or Binary encoding
- ?? Formalized extension process for enhanced functionality
- ?? Expanded definition of PACKAGES

X. MEGACO/H.248 Architecture & Entities

MEGACO has the same architecture as MGCP its commands are similar to MGCP commands. However, a main difference between the two implementations is that with MEGACO, commands apply to Terminations relative to a Context, rather than to individual Connections, as is the case with MGCP. Connections are achieved by placing two or more Terminations into a common Context. It is the concept of a Context that facilitates support of multimedia and conferencing calls. The Context can be viewed as a mixing bridge that supports multiple media streams for enhanced multimedia services.

MEGACO Packages include more detail than MGCP Packages. MEGACO packages define additional Properties and Statistics along with Event and Signal information that may occur on Terminations. With MEGACO, the primary mechanism for extension is by means of Packages. To accommodate expanded functionality, MEGACO specifies rules for defining new packages.

Below is an "at-a-glance" comparison between MGCP and MEGACO.

MGCP	MEGACO
- Endpoints, Connections	- Terminations, Context
- 1 transaction = 1 command	- 1 transaction = N actions 1 action = N commands
- 1 type of response	- 2 types of response (new Transaction Pending)
- Commands : Endpoint Configuration Notification Request Notifications Create Connection Modify Connection Delete Connection Audit Endpoint Audit Connection Restart in progress	- Commands : Add Modify Subtract Move Audit Value Audit Capabilities Notify Service Change
- Grammar : Text encoded (BNF)	- Grammar : Text encoded (ABNF) + Binary form (ASN1)

XI. The Future

In the telecommunications industry, traditional telephony services have been relatively unchanged since their introduction. In an IP environment, it is apparent that communication services will continue to evolve and expand and offer us new services and applications that will enable any device, anytime, anywhere communication.

The key to enabling seamless communication and connectivity in a converged network is not just the protocols themselves, but the interworking solutions between the protocols. It is clear, that the global network of the future will be a hybrid of communication platforms with multiple vendors and an infinite number of devices with various capabilities and features. All of the protocols mentioned herein will be part of that new environment as well as others that will be introduced in the interim. Interoperability and interworking are critical to making the use of next-generation networks as reliable and as robust as traditional telephony is today.

That is why it is more important than ever to choose an experienced partner with proven solutions like RADVISION for products and technology for voice, video and data over IP. RADVISION is recognized globally as the experts in real-time voice and video over IP (V^2oIP?). The company has a long-standing reputation as a champion of interoperability and a driving force behind industry standards for IP-centric communication. With a broad suite of enabling technology toolkits and network infrastructure products, RADVISION has the solution for migration from traditional telephony networks to converged networks.

###

APPENDIX

The Standards Bodies and Industry Consortia

There are a number of organizations all working toward "complete interoperability" in the multimedia communications arena and each has their own charter.

The ITU-T. The main international body responsible for telecommunications standards is the ITU-T (formerly the CCITT). The ITU-T is chartered by the United Nations and all UN countries are represented. Processes are clearly defined and standards have to go through specific revisions, reviews and approval stages before being ratified ("determined" is the ITU term).

The ITU-T is divided into Study Groups (SG's), which get a four-year mission, couched as "questions", to solve a specific problem. The ITU-T SG 16 was assigned several questions relating to multimedia conferencing and packet based telephony. All the SG 16 standards carry the "H." prefix (e.g. H.320, H323, H.324).

The IETF. The IETF is a voluntary organization open to all. The IETF's main charter is to define the protocols that are to be supported on the public Internet. Internet standards are termed "Request for Comments" (RFC's) and are numbered. RFC's are officially short term and need to be renewed and given a new RFC number on a regular basis to stay alive.

The IETF functions mostly through Working Groups (WG) that address specific issues. A group of people interested in a specific item convenes, communicating mostly via email. A person either volunteers or is chosen as the editor and starts creating the draft for comments. If there is a consensus, the WG can offer it as an RFC and get an RFC number. Session Initiation Protocol, or SIP, is one of the latest IETF standards that is being used for multimedia communications.

The IMTC is the major consortium and represents vendors, carriers and end users. The IMTC has about 150 members. Its claim to fame was the adoption of the H.323 and G.723.1 standards for VoIP by the Voice over IP Forum (which is part of the IMTC). The IMTC's most significant contribution is in interoperability testing. The InterOP and SuperOp events are open to members who come with their equipment and test it with other vendors.

ETSI/TIPHON. A European body affiliated to the EC, which regulates telephony standards in Europe. EuroISDN is an example of an ETSI standard. ETSI members used to be national carriers (PTT's), Telecommunications Ministries and large European vendors.

ETSI created a new body called TIPHON, which is focused on IP Telephony. TIPHON's mission was to provide inputs from PTT's and other Carriers for the standardization bodies. Practically speaking, TIPHON is open to any interested party (including non-European vendors). TIPHON is organized into six working groups covering issues like general architecture, testing, and billing. To date TIPHON has not written any standard but has provided proposals and submissions as liaisons to the ITU-T.