Demystifying Protocols:

A Comparison of Protocols Suitable for IP Telephony

By Tracy Venters
Sonus Networks
tracy.venters@sonusnet.com

INTRODUCTION

Often described as a "protocol soup," the dizzying array of standards addressing packet telephony networks is enough to leave even industry experts scurrying for the decoder ring. Casual conversations about IP telephony are sprinkled with frequent references to H.323, Megaco, SIP and MGCP. Dig into the subject a little deeper and one is likely to hear references to the makeup of these protocols, terms like: H.245, H.225, IPDC, SGMP and H.248. While it may provide an interesting debate for the engineering minded, trying to decipher which protocol is most suitable for a network application can be puzzling to a service provider.

This paper will attempt to sort through the soup, presenting the relative merits of four protocols, H.323, SIP, MGCP and Megaco in an unbiased fashion. While there are many protocols important to a converged voice and data network, these four are featured because they are currently the most popular for implementing voice over packet services. In case you are tempted to skip to the last page in search of the clear winner, you won't find it. Each protocol has its strengths and weaknesses with respect to issues such as ease of implementation, extensibility, suitability for various network applications, quality of service and security. The typical next-generation network may very well include all four.

PEER PROTOCOLS OR MASTER/SLAVE?

In order to properly compare these protocols, we must distinguish between the two models of distributing intelligence. H.323 and SIP operate between peer clients, while MGCP and Megaco operate between master and slave entities. This is an important distinction. For example it explains why it's unlikely to see a side-by-side comparison of SIP and MGCP. However, you might find yourself deciding between a SIP-based Integrated Access Device (IAD) or an MGCP-based IAD to deliver the same set of services to an analog phone. In this case, the difference lies in how much intelligence is built into the client device.

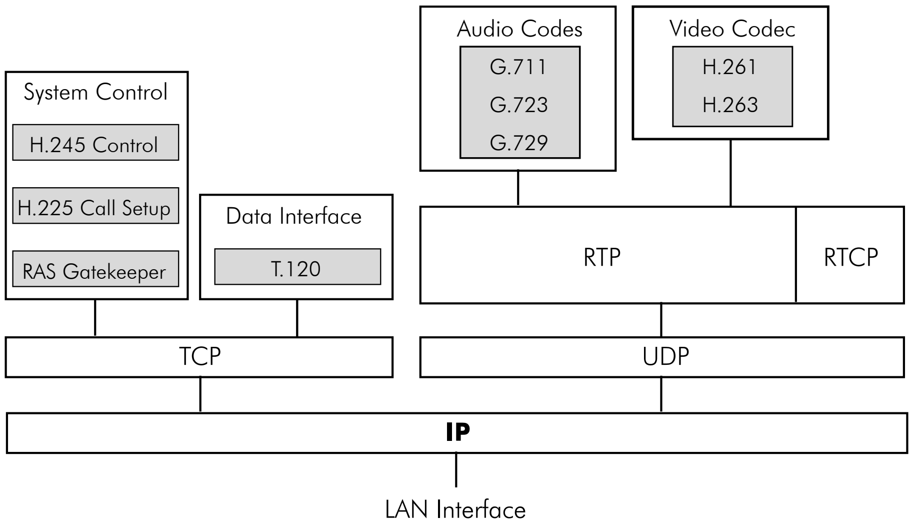
With slave devices such as MGCP- and Megaco-based media gateways, IADs and phones, the control model more closely aligns with traditional telephony equipment: the call agent (also referred to as the media gateway controller) must supply all instructions to the "dumb" end device, instructing it to wait for signals, collect digits, play tones, open ports and release connections. This is a simpler implementation, leading to lower-cost end devices.

SIP and H.323 are usually used in conjunction with signaling across networks, which requires peer-level control. However, SIP in particular is becoming popular for building smart terminals. SIP-based end points trade lower cost and ease of implementation in favor of a model that can deliver much richer services.

H.323 OVERVIEW

Developed by the ITU , H.323 actually encompasses several protocols, including: H.245 for media control, H.225 for connection establishment between endpoints, H.332 for large conferences, H.450 for supplementary services and RTP for transport (SIP also uses RTP for transport, which simplifies interworking between the two). H.323 was initially developed for multimedia conferencing over local area networks. Although it is the most widely deployed packet telephony standard today, H.323 is losing ground to SIP due to the complexity of the H.323 protocol suite. However, the standards continue to evolve to try to fit the needs of Internet telephony, including increasing efficiency and supporting more services.

H.323 Standard Series



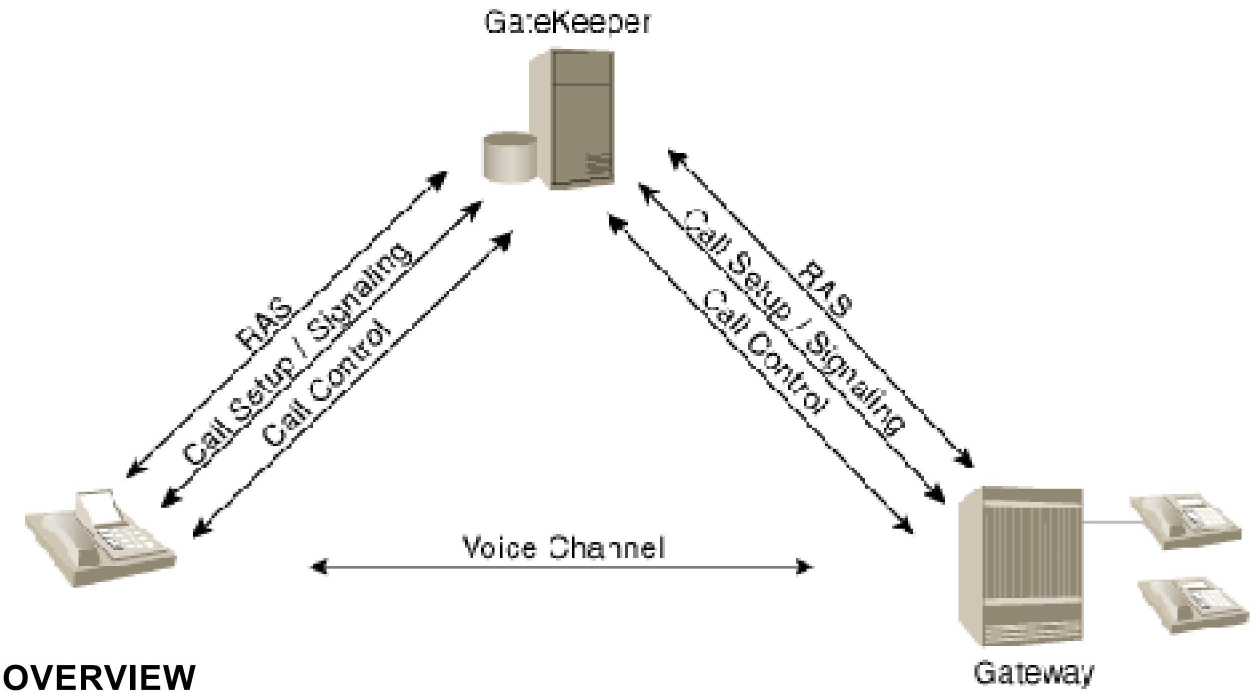
An H.323 system consists of four types of components:

- The H.323 **gatekeeper** is optional, providing admission control, address translation, bandwidth control and zoning.
- **Gateways** provide connection between packet and circuit-switched networks, providing call setup, control and media transcoding between the two networks.

---

[1] The International Engineering Task Force (IETF) is an international standards organization focusing on the architecture, operation and evolution of the Internet.

- **Terminals** must support H.225 (including RAS and Q.931 signaling) for call setup, H.245 for controlling the media, as well as speech compression and codec control. H.323 terminals can optionally support audio/video codecs and T.120 data conferencing logic.
- A **Multipoint Control Unit** (MCU) provides multimedia conferencing for three or more H.323 terminals. All parties interested in conferencing must go through the MCU to determine the codec in use. The MCU architecture is further divided into Multipoint Controller (MC) and Multipoint Processor (MP). The MC negotiates capabilities among all conference parties to determine an acceptable level of communication quality.



SIP OVERVIEW

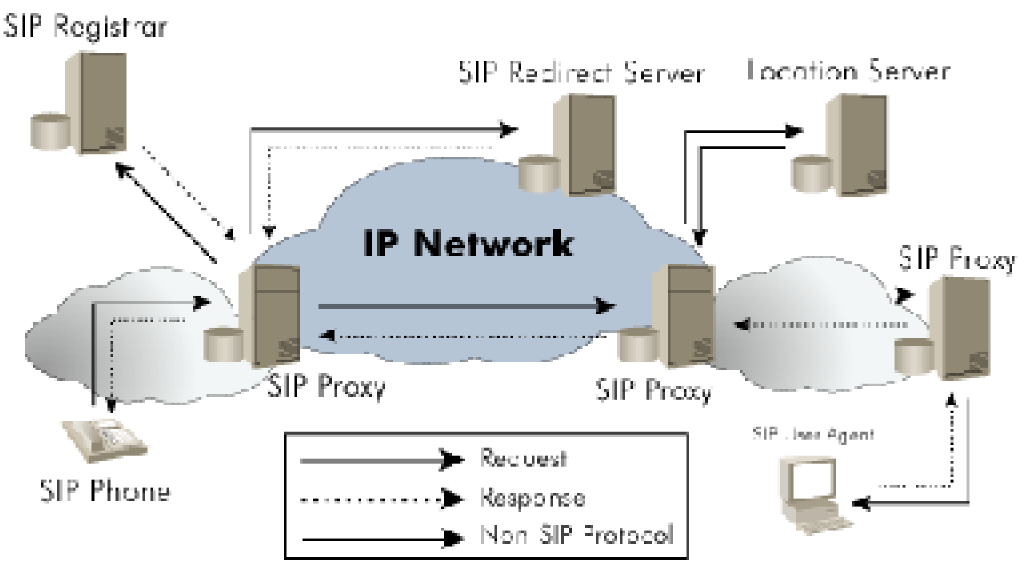
Developed by the IETF , Session Initiation Protocol (SIP) is a lightweight, text-based signaling protocol used for establishing sessions in an IP network. It reuses a lot of the constructs and concepts of Internet protocols such as HTTP and SMTP. Based on the principals learned from the Internet community, SIP is an application-independent protocol designed at the outset to be flexible and extensible. As the name implies, SIP deals generically with sessions. Those sessions can include voice, video or data. In fact, SIP does not specify the codec or transport protocol; the media it carries is described in a separate protocol called Session Description Protocol (SDP). SIP has been enthusiastically embraced for next-generation applications such as telephony over packet services, voice enabled e-commerce applications, presence management, instant messaging services and voice controlled web browsing.

---

[2] The International Engineering Task Force (IETF) is an international standards organization focusing on the architecture, operation and evolution of the Internet.

As shown in the diagram below, there are four main entities in SIP architecture. These can be implemented as stand-alone components or combined in a single platform, according to the scalability and architecture required.

- **User Agents** are SIP endpoints that initiate and respond to requests and communicate with other user agents to establish and release sessions. User agents can communicate with each other directly, however there is often one or more intermediate servers involved, either proxy or redirect servers.
- **Proxy servers** can be stateful or stateless, forwarding messages on to the user agents as well as performing services such as location services, authorization and accounting.
- **Redirect servers** are always stateless, they simply respond to a request with the location where the originating user can contact the desired party directly.
- User agents register their location with **Registrars**, allowing for a rich set of mobility features to be implemented using SIP.



SIP AND H.323 COMPARISON

Complexity

The several hundred pages of ITU H.323 specifications stand in stark comparison to the 128-page SIP specification, developed by the IETF. To setup an H.323 (Version 1) call using a gatekeeper, nearly 20 packets and 6 round trip delays are required. Although the number of messages may vary due to reliability requirements and feature interaction, a basic SIP call can be set up with just 4 packets and 1.5 round trip delays. In Version 2, H.323 specifies a FastStart method to reduce the number of delays and also allows the option of using UDP instead of TCP to reduce delay even further. SIP can be carried over any reliable or unreliable byte stream or datagram protocol; implementations using UDP and TCP are the most common.

Extensibility

Provisions for extending the H.323 protocol follow the typical mechanisms of traditional telephony protocols such as ISUP, meaning there are placeholders in the encoding where vendors can implement non-standard parameters. Of course, this limits the extension to only that which can be added in those particular fields. Also, there is not a mechanism for terminals to exchange information about which extensions they support, so the extensions only work between terminals from the same or cooperating vendors. SIP, which borrows heavily from Internet protocols such as HTTP, lends itself well to extensibility. In particular, SIP defines a formal mechanism for negotiating support of features. This mechanism allows endpoints to specify both the extensions they require as well as those that are desired but optional.

Quality of Service (QoS)

SIP was built from the ground up as a lightweight protocol, which facilitates minimal call-setup delay. Later versions of H.323 have introduced mechanisms to reduce the delay, although there is no guarantee that an end device will support these newer methods. Although its control messages usually run over TCP, H.323 specifies its own method for reliable transmission of control messages. SIP, which makes no assumptions of a reliable datagram service, specifies a simple retransmission scheme until acknowledgement is received. With both protocols, the media is exchanged over RTP so the available options for media stream QoS are similar. The difference is that H.323 builds in a rich set of capabilities for admission control, prioritization and bandwidth management. SIP relies on other protocols to take care of QoS.

Security

Both H.323 and SIP provide security mechanisms. Like its approach to QoS, H.323 specifies explicit security procedures in H.235. SIP borrows from established methods for security procedures, including HTTP for caller and called authentication, SSL for hop-by-hop encryption and PGP or S/MIME for end-to-end encryption and authentication. SIP is much more firewall friendly than H.323 because only one protocol and a single stage negotiation is involved.

Suitability for Telephony Applications

H.323 was initially developed for multimedia conferencing, so it is certainly suitable for such applications. H.323 has also been adopted and widely deployed internationally to take advantage of international calling arbitrage, providing cheaper rates for consumers and nice profits for service providers. There is no question that H.323 is suitable for carrying voice over IP networks. In addition to its rich conferencing abilities, H.323 supports a number of other services common to the business segment such as call forward and transfer, additional services continue to be added. H.323 does support a facility message that can be used to redirect a user to another location. For example, blind transfer uses the facility message and it could also be used for simple follow-me services. However, it does not qualify as providing true third-party call control to, for example, an application server.

Like H.323, SIP is also suitable for carrying voice and video over IP networks. Unlike H.323, however, SIP is not limited to TCP or UDP - it can also work well over other transport protocols such as ATM, frame relay or X.25. SIP supports at least as many services as H.323, including conferencing, transfer, call forward, call park/pick-up and hold. SIP is particularly strong for signaling across diverse networks; along with the help of (non-SIP) location services SIP proxy servers can be used to forward or redirect messages to the user location. In fact, SIP also has excellent capabilities in support of user mobility. SIP caller and callee preferences for communication can be different for each location. SIP also supports a variety of endpoints including wireline, wireless and soft phones, PDAs, and instant messaging applications. The most powerful feature of SIP is its origins; based on Internet protocols, SIP is extensible, relatively simple to understand and implement. It uses text encoding easily understood by looking at the messages, and is therefore adaptable by the worldwide army of Internet application developers.

Example SIP Message

```
INVITE sip:+1-978-5551212@111.122.133.144;user=phone SIP/2.0
   Via: SIP/2.0/UDP 155.166.177.188:5060
   From: 617-6661212 <sip:+1-617-666-
1212@155.166.177.188;user=phone>
   To: 978-5551212 <sip:+1-978-
5551212@111.122.133.144;user=phone>
   Call-ID: 12345678@155.166.177.188
   CSeq: 1 INVITE
   Content-Type: application/sdp
   Content-Length: 120
   v=0
   o=UAC 8761 9876 IN IP4 155.166.177.188
   s=Session SDP
   c=IN IP4 555.666.777.999
   t=0 0
   m=audio 49172 RTP/AVP 0
```

MGCP AND MEGACO

As mentioned above, MGCP and Megaco are master/slave protocols as opposed to SIP and H.323, which are peer protocols. But while it is easy to compare and contrast SIP and H.323, pitting MGCP and Megaco against each other in terms of suitability for various applications leaves fewer distinctions. The similarities can be traced to the roots of these protocols. The Megaco Working Group (WG) was originally formed under the IETF to provide an open standard for IP-based gateway device control. Several proposals were introduced, with an early leader being MGCP. The Megaco WG never adopted MGCP in its entirety, although key aspects of it were integrated in to what would become the Megaco Protocol completed in March 1999.

In parallel to the IETF efforts, the ITU-T Study Group 16 was working on a gateway control protocol initially known as H.GCP and later designated H.248. ITU SG-16 and the Megaco WG agreed to work toward a common standard and both groups approved Megaco/H.248 in June 2000.

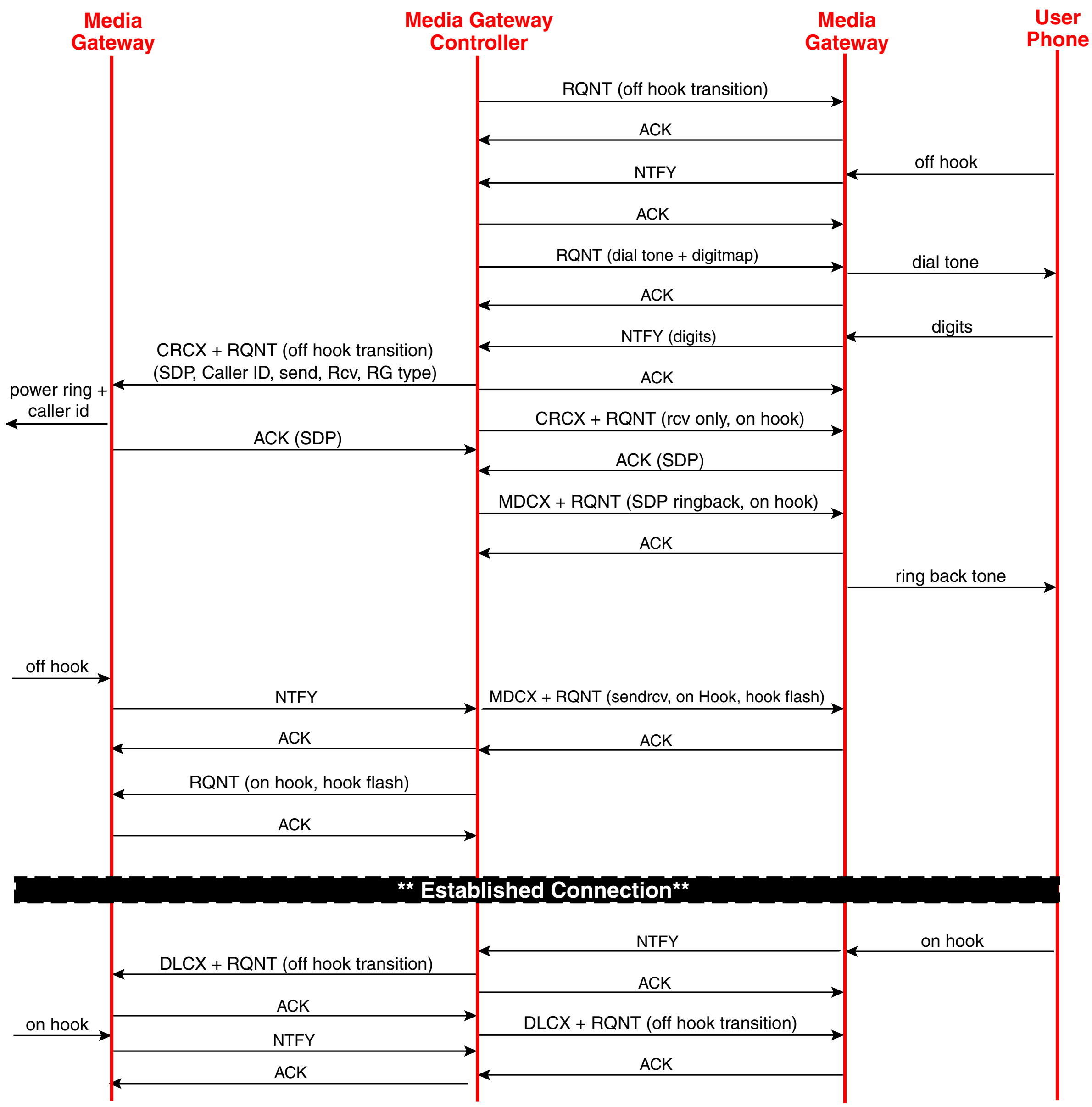
Although MGCP is technically not an official standard, it was released as a Request For Comments (RFC) document. After years of waiting for an official standard, many vendors began to implement gateways using MGCP. The CableLabs® standards body specifies standards for telephony over cable networks and developed the Network-based Call Signaling (NCS) protocol, which evolved out of MGCP. Many implementations are available today using NCS 1.0.

Before delving into the details of MGCP or Megaco, it is useful to define the set of functions expected of any media gateway control protocol. Both MGCP and Megaco provide mechanisms to support the following functions in a media gateway:

- Create, modify and delete connections involving bearer terminations in any combination. The bearer terminations may be TDM, analog, Ethernet, ATM or frame relay.
- Detect events and apply signals to media streams. When an event is detected, the media gateway controller specifies the media gateway's actions such as reporting the event or applying another signal.
- Collect digits according to the dial plan (digit map) specified by the media gateway controller. Digit maps increase efficiency; instead of reporting each digit as it is received, the media gateway may collect digits in a buffer according to the predetermined dial plan.
- Add or subtract media streams through out the session, allowing for services such as conferencing, call hold, call waiting and playing announcements.
- Report statistics collected for the call such as quality of service measurements.

Both MGCP and Megaco have the concept of endpoints and connections. The MGCP endpoint is called a termination in Megaco, and the MGCP connection is called a context in Megaco. A termination/endpoint is the origination or destination of a media stream. These media streams are associated through a connection/context. The procedure by which a connection is established differs slightly between the two; the Megaco model allows for more flexibility and finer control by the media gateway controller. In terms of resource reservation and control, media processing and stream management, Megaco has greater capabilities as well, which makes it a better protocol for applications such as multimedia conferencing. In both protocols, the control of similar types of endpoints is specified in a package. A package contains a list of signals, events, properties, and statistics that make sense to the endpoint it is intended for, for example an analog line or TDM trunk.

MGCP Message Flow



MGCP AND MEGACO COMPARISON

Complexity

In general MGCP and Megaco have constructs to accomplish the same types of tasks. However, since MGCP was a precursor to Megaco, Megaco has refined and extended many of the functions. Predictably, Megaco is a more complex protocol.

Extensibility

For both MGCP and Megaco, packages are key to the ability to extend the protocol. In both cases, the packages are not a part of the protocol specification, so the protocol does not need to be changed to extend or add new packages. One key benefit Megaco has over MGCP is the ability to negotiate the version between the media gateway and media gateway controller. With MGCP, version control and compatibility is a process outside the protocol itself.

QoS

Megaco is much more flexible when it comes to the underlying transport type. While MGCP defines only UDP as a transport layer for the signaling messages, Megaco allows TCP, UDP, SCTP and ATM. Megaco also has better resource allocation and stream management mechanisms.

Security

MGCP supports IPSEC to secure signaling packets in the underlying transport. Megaco also supports this, but adds another option of an authentication header if IPSEC is not available. For security of media streams, both protocols support encryption of audio messages for protection against eavesdropping and encryption and authentication of the source address to prevent source address spoofing.

Suitability for Telephony Applications

Because the functions of MGCP and Megaco concentrate only on the media stream, the same types of telephony applications can be provided by both, although the procedure for implementation may be much simpler in one protocol or the other. Conversely, a simple implementation is often offset by added flexibility. As previously mentioned, Megaco is better suited for multimedia conferencing and transport over diverse networks than MGCP.

SUMMARY

H.323, SIP, MGCP and Megaco are all integral protocols in the world of IP telephony. H.323 and SIP are used for peer-level communication, while MGCP and Megaco are used for communication between master and slave components. Similar services can be offered using various combinations of these protocols. For example, a multimedia conferencing application could be offered using the H.323 standards or using a combination; SIP between media gateway controllers, MGCP between the media gateway controller and media gateway.

H.323 is the most widely deployed of these four protocols and is well suited for voice and multimedia conferences. H.323 follows the path of traditional telephony protocols with voluminous specifications and limited extensibility. It implements its own mechanisms for dealing with issues such as quality of service and security. SIP, on the other hand, makes use of already established, popular Internet protocols such as HTTP and SMTP and relies completely on other protocols for encryption, quality of service and payload transport. This makes SIP a natural for integration between voice and data services - in addition to voice over IP, SIP supports a whole host of enhanced services such as voice enabled e-commerce applications, presence management, instant messaging services and voice-controlled web browsing.

MGCP and Megaco have many more similarities than differences. Both operate in a master/slave configuration where a media gateway controller instructs the media gateway to establish, control and release connections between one or more media streams. Megaco was the final derivative of several draft standards, including MGCP. Megaco refines and extends MGCP at the expense of adding complexity. Both MGCP and Megaco are suitable for many types of telephony applications.

CONCLUSION

For carriers and vendors, the decision as to which protocols to embrace must begin with the decision as to where to put the intelligence and control. Carriers may be most comfortable with the "dumb" terminal model, however SIP endpoints allow easier implementation for a rich class of services. The MGCP/Megaco model is particularly well suited for low-cost media gateways used for access such as IADs and IP telephones. In the end, despite all the debate, the consensus is that H.323, SIP, MGCP and Megaco will all be part of the next-generation network.