White Paper

MXM: SIP Support

Delivering SIP–H.323 Convergence and Co-Existence Using a Video PBX Architecture

VCON

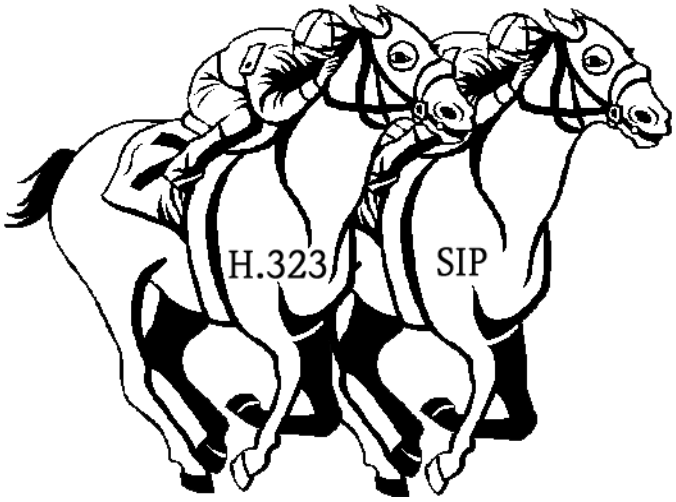
Table of Contents

Gordon Daugherty
Chief Marketing Officer
August 2004

Introduction

As new standards emerge and existing standards evolve, companies are often put into the very difficult position of predicting the future. However, this is like trying to shoot at a moving target. Sometimes one standard eclipses the other, while other times two or more standards converge into a broader, more encompassing standard. When in the early stages of standard validation, it is virtually impossible to predict the path and speed a market will take relative to standards. These statements certainly apply to the SIP and H.323 protocols and the evolution of the multimedia communications market.

If either the current H.323 or SIP is declared a "winning" protocol to the detriment of the other, customers' expectations will likely fall short of their potential realization. At the same time, customers' expectations are likely to be best met by a convergence and co-existence of these compatible means to a common goal of global IP communications. This paper briefly explains each of the protocols, describes a global perspective of what needs to be included in a complete, IP-based visual communications system and offers an alternative to choosing one protocol over the other while protecting all investments in either protocol over the lifetime of the solutions.



What are SIP and H.323?

At their core, SIP and H.323 are protocols defined and published as a result of vendor contributions to and ratification by respected international standards-setting bodies (IETF and ITU, respectively). Both standards are very modern protocols and show a great deal of promise. However, the scope and underlying concepts, as well as the early adopters of these standards differ significantly.

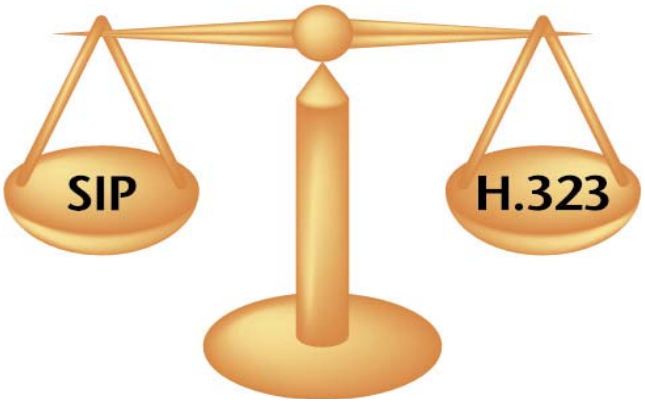
H.323

The body of work commonly known as "H.323" is an umbrella standard encompassing references to a range of sub-systems, protocols and algorithms for packet-based multimedia communications systems. Currently in its 4th full version release since approval of version 1, H.323 is continuously being enhanced through a submission and ratification process supervised by the International Telecommunications Union (ITU-T).

Originally, H.323 emerged by leveraging prior standards (H.320 and H.324) for video communications over circuit switched networks (i.e., ISDN and POTS, respectively). Approved in 1990, the H.320 standard ignited the videoconferencing industry by providing a means to interoperability and the bulk of the protocol remains in use today. In the first phases of development of H.323 it became clear that the standard had the potential to serve both voice and video communications systems and therefore, the video encoding components were never required as part of the most basic system. It also was important to the developers of H.323 technologies that the protocols include provisions for managing the network components in addition to end points, and for assuring network resilience in an otherwise non-deterministic environment such as IP.

The H.323 protocol defines an "entity" as the basic operational unit to and from which signals are sent and received, and media or data transmissions result (are made possible). There may be new entities defined but in the original and the present most expansive view, the entities include Terminals (otherwise known as "end points), Multipoint Controller (the "brains" behind a multipoint session), Multipoint Processor (the muscle behind a multipoint session), Gatekeepers (the overall mechanism by which policies are set and enforced in a H.323 network) and Gateways (the internetwork communications systems required for ensuring compatibility with solutions on any network other than the IP infrastructure). Gateways are further decomposed into the Media Gateway Controller and the Media Gateway.

In addition to entities with core/required functions and optional attributes, the H.323 protocol defines a large set of telephony-like capabilities over reliable transport for signaling. Inherent in every aspect of H.323 is the concept of "control" for both network operators and, to the extent the network operators permit, for users of systems that comply with the protocol. The architects of H.323 believed the control of entities and quality of service to be an essential element for providing (and having a revenue model associated with) network-based services.



SIP

Session Initiation Protocol (SIP) defines, as the names implies, how two IP nodes establish a session with or without media. In the original request for comment (RFC 2543), the authors state that SIP would not be "tied to any particular conference control protocol." As the interest in the lightweight protocol grew from many different perspectives, ancillary components were designed to augment the initial protocol. Since release and approval of the original protocol in 1999, the acronym "SIP" has been accepted as the unofficial name of the larger, more inclusive system.

One of the founding principles of SIP is that the protocol is to have as low of a resource demand as possible on any given device, thereby enabling communication with power-constrained and memory-constrained devices. This is accomplished, in part, by having a very simple, yet very flexible, message definition system. All messages are encoded in ASCII text format. The syntax for these messages must only conform with the Augmented Backus-Naur Form (ABNF), a syntactical notation designed by engineers for use in this protocol (not the same as any standardized ABNF). A byproduct of ASCII format is that anyone can read and potentially, edit message contents. SIP is extended by adding new header lines that may be used by different vendors to serve different purposes.

Although they began from different sets of objectives, and are evolving by different mechanisms, H.323 and SIP have a lot in common. Both are designed to treat underlying media "transparently." In other words, the audio, video and data compression algorithms used by both systems can continue to evolve independently of the control and communications system in which the media streams are sent. In both cases, video is an optional media. Voice can be format using any of the standards-based algorithms, such as G.711, G.721, G.723, PCM or proprietary options. Video streams can use any of the H.26x series protocols or proprietary and future standards.

The two standards also share a need for a diverse and complete ecosystem of vendors from which customers can choose and be guaranteed interoperability and stability in an infinite number of combinations found in enterprise and service provider networks.

A Complete Multimedia Communications System

A complete multimedia communications system must have the following components:

- Network bandwidth for transmission of signals and media,
- Network-based devices with which to provision services,
- Endpoints for user interface, network access, encoding and decoding,
- Network deployment and management tools,
- Security-assurance tools and
- A network administration platform for ongoing monitoring and management.

This paper assumes a basic understanding of the above components and services. However, it is appropriate to explore some of them further in the context of SIP-H.323 convergence and co-existence. As a general rule, the network transport itself and the network bandwidth is completely agnostic to any issues pertaining to SIP and H.323. Even issues such as network QoS, latency and jitter are present with both protocols and are addressed similarly.

However, moving up the OSI layers we can find opportunities for co-existence. Video-specific network components such as MCUs, gateways and PBXs can certainly play a role in the evolution of these protocols. They offer a hand-shake point for interoperability and/or protocol conversion. It is almost certain that during the upcoming migration or convergence period these components will play a vital role.

The videoconferencing endpoints present an interesting question. Will they continue to support one protocol or the other or will they also migrate to supporting both IP-based protocols much in the way that most videoconferencing systems today support both H.320 and H.323 (commonly referred to as “dual-mode”)? After all, much of the reason for dual-mode systems today is to ease the migration from ISDN (H.320) to IP (H.323). Perhaps this same thing will happen with IP-only endpoints supporting both SIP and H.323. But it is equally possible, and perhaps even more likely, that the endpoints will only support one IP video protocol and the migration or convergence will become possible via the video-specific network components mentioned above. This seems to better fit the IP networking model where intelligence and control tends to be more centralized.

At yet a higher layer we have tools for management and administration, which are already vital tools for the network administrator regardless of the protocol debate. Videoconferencing can be a demanding application to manage. This places extra burden on the administrator and puts extra emphasis on the tools he/she will use on a daily basis.

Why should you care about the evolution of these protocols?

Unfortunately, both SIP and H.323 have proponents and vendors that are building products and services complying with one or the other of these protocols. Carriers and application developers are currently dividing their limited resources to invest in both directions until such a time that one protocol dominates over the other or a new protocol emerges.

Unfortunately, most companies do not have the financial or human resources to support two parallel systems, forcing them to choose between one or the other. Without careful planning and continued analysis of the options on the market, a company could either opt to wait entirely on deploying IP based communications systems or pursue an avenue that includes risk of leading to a dead end.

Choosing a communications system should not become a religious debate about which protocol is better than the other and which will win. Successful communications systems must permit the technologists in forward

thinking organizations to deploy solutions without limiting their options to support new protocols in the future. If the hardware/software vendors and the network providers are mindful of the customer, they will do as much as possible to minimize risks and maximize investment protection without reducing functionality or control as a trade-off.

As mentioned above, effectively converging and supporting co-existence offers tremendous flexibility and, most importantly, investment protection. But it is most desirable that services are not compromised in the process. VCON clearly recognizes this need and is responding with an architecture that provides both application/service-level convergence along with SIP/H.323 protocol co-existence.

Convergence and Co-existence

Optimally, carriers and enterprise network managers should be able to deploy either SIP or H.323 in confidence and provide today's users with services available on today's platforms. Until the time if or when one of the two protocols dominates, or the two evolve sufficiently over time to completely satisfy the needs of all communities of interest, the best course of action is to support a converged environment wherein both SIP and H.323 can co-exist.

There are technical solutions that permit co-existence. The first, as suggested above in the context of internetworking facilities between circuit switched and packet switched networks, is to deploy H.323 to SIP gateways. These support interconnectivity with limited exposure. Proprietary solutions are available from several vendors and should have only incremental complexity relative to the complete communications system's operation. Over time, standards bodies will approve protocols for interconnection functions that will guarantee interoperability among different vendor solutions.

Convergence at the application or services level

While co-existence via gateway devices does achieve basic connectivity, it often leaves the user with only lowest common denominator services. In other words, services are exchanged for connectivity and interoperability. An alternative is to converge the two protocols at the application or services level.

In order to achieve application-level convergence, typically the applications and services must be logically separated from the device interface or access control points. With this, a common set of services can be delivered to multiple types of devices, each with a potentially different access protocol.

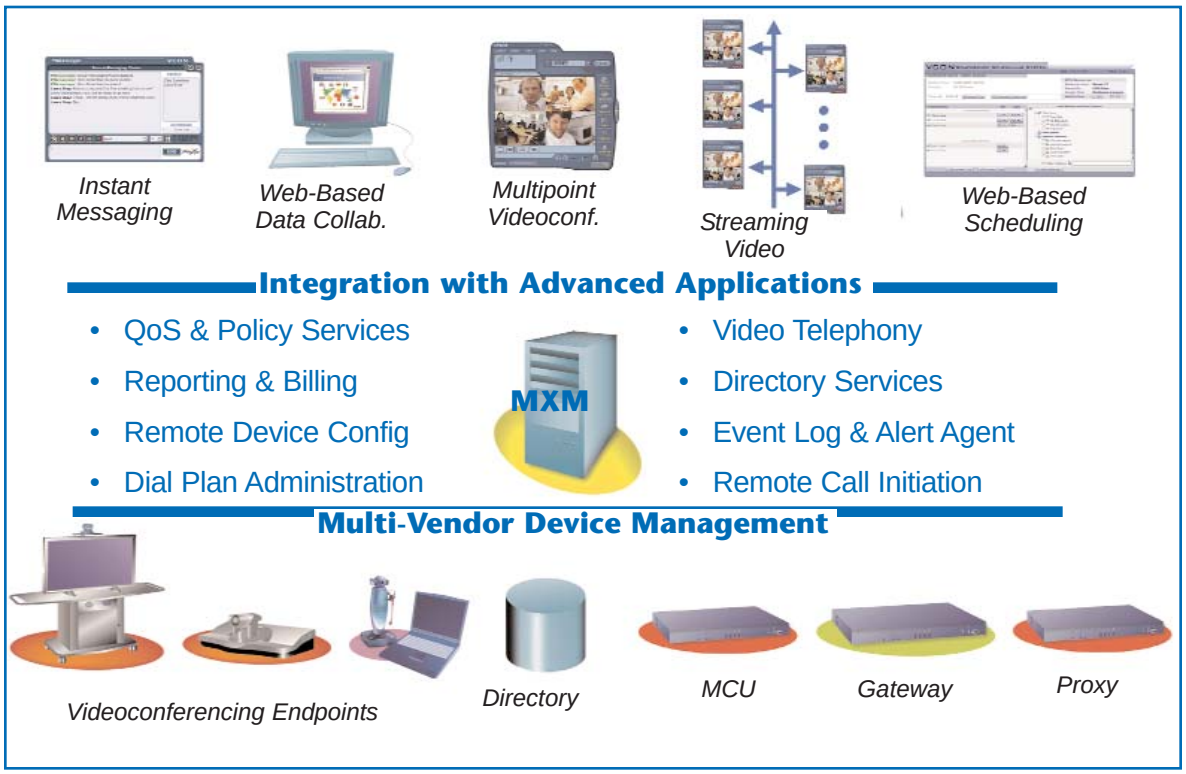
The VCON MXM Solution

In order to deliver a fully integrated solution for IP video, VCON has developed a client/server architecture. At the center of this solution is the VCON Media Xchange Manager (MXM). The MXM is a suite of integrated client/server applications and services that gives enterprises and service providers the opportunity to truly maximize the power of real-time, interactive visual communications across the enterprise. Using the MXM, video network administrators will be able to manage, administer, and monitor their videoconferencing networks from a remote console located anywhere on the network. And end users will have access to functionality and ease-of-use never before available. This new world of integrated IP video telephony will deliver the ease-of-use and advanced management capabilities that will finally allow for mass deployment.

The internal technical architecture of the VCON MXM is ideally suited for the convergence and co-existence of H.323 and SIP. This is derived from the fact that, at its core, the MXM is a video PBX. Most of the user-based services and applications are either embedded into this PBX core itself or built on top of it. The entry point to the MXM is via a proxy interface. With both a H.323 gatekeeper proxy and a SIP proxy as possible interfaces, a mixed network of H.323 and SIP devices can be deployed – all with similar services. H.323-specific devices supported by the MXM include desktop and conference room endpoints, MCUs and gateways. On the SIP side, concentration has been focused on the Microsoft XP Messenger client, which includes optional video services.

The following diagram depicts the logical separation of services, applications and interfaces within the MXM architecture.

The MXM is able to deliver a powerful set of services across a mixed network of H.323 and SIP devices. Key examples include the following:



Management and Administration

- Bandwidth limits (personal, zone and interzone)
- Policy services
- Call status monitor
- Reporting and billing
- Scheduling
- Remote call initiation
- Dial plan management

Video Telephony

- Call transfer
- Call forward
- Call pickup
- Ad hoc conferencing
- Hunting groups
- Simplified gateway dialing

Phased Rollout

VCON's introduction of convergence functionality is expected to take place over three phases:

	Phase 1	Phase 2	Phase 3
Conferences between like devices	Yes	Yes	Yes
Mixed Conferences		Yes	Yes
Common mgmt and admin features between H.323 and SIP	Yes	Yes	Yes
Video Telephony Features	H.323-all/SIP-some	Yes	Yes
Advanced Convergence Applications			Yes

Summary

Two standards for multimedia communications are competing for customer attention. Ultimately, macroeconomic forces on the marketplace—not the standards bodies or vendor community alone—will yield stable, proprietary or standards-based technology in which investments can be easily justified and yield a positive return on investment. Until the multimedia communications market reaches this state, customers can feel uneasy, afraid that if they choose one standard over the other they might be caught in the "wrong camp."

Many potentially powerful and efficient IP-based communications deployments are in limbo today, waiting for the outcome of what could be, in light of the similarities and differences between the SIP and H.323 standards, largely a symbolic victory. Approached with the objective of enabling standards convergence and coexistence, the technologies available today from VCON and its partners offer a low-risk opportunity to invest in a converged IP-based video network.

VCON is a recognized leader in the IP video market, regularly innovating new technologies and solutions that benefit both the end user and the network administrator. The phase that the industry is about to enter clearly presents continued challenges and risks along with the rich rewards and benefits. During such times it is the innovators the market will count on to bridge one era to the next. VCON's vision of the future and our focus on delivering true investment protection will provide tremendous benefits for years to come.

VCON

VCON Headquarters
Ph: +972-9-959-0059
Fx: +972-9-956-7244

VCON Americas
Ph: +1-512-583-7700
Fx: +1-512-583-7701

VCON Europe
Ph: +49-89-614-57-0
Fx: +49-89-614-57-399

VCON China
Ph: +86-10-65269791
Fx: +86-10-65269790

www.vcon.com