

Patterns for VoIP Signaling Protocol Architectures

Juan C. Pelaez¹, Eduardo B. Fernandez¹, and Christian Wieser²

jpelaez@fau.edu, ed@cse.fau.edu, chwieser@ee.oulu.fi

¹ Department of Computer Science and Engineering
Florida Atlantic University
Boca Raton, FL 33433, USA

² Department of Electrical and Information Engineering
University of Oulu
Oulu, Finland

Abstract

Voice over IP (VoIP) is an alternative to traditional circuit-switched telephony that allows human voice and video to travel over existing packet data networks along with traditional data packets. H.323 and SIP are the two standard protocols used today for signaling and call control in VoIP, and they are essential for providing total access and for supporting IP-based services. In order to provide complete end-to-end connectivity, interworking between SIP and H.323 is necessary since both of these protocols have been extensively deployed. We present here a pattern for the H.323 protocol architecture followed by a Hybrid VoIP signaling protocol pattern that combines H.323 and SIP architectures. These patterns can be used to guide the design of VoIP systems and products as well as to simulate these systems.

1. Introduction

Voice over IP (VoIP) is an alternative to traditional circuit-switched telephony that allows human voice and video to travel over existing packet data networks along with traditional data packets. VoIP technology uses IP-based networks to establish and manage communication sessions between terminal devices.

Two types of protocols are used in VoIP: signaling protocols and media transport protocols. Signaling allows call information to be carried across network boundaries providing session setup, control and teardown. VoIP signaling protocols can be generally divided into two main groups, client-server and user-to-user. In the latter group, SIP and H.323 are the two most popular. Media exchange (Client-server type) protocols are out of the scope of this paper.

H.323 defines a family of protocols specified by the ITU research group [ITU06]. The standard provides a foundation for signaling in order to exchange voice, video and data communications in an IP-based network. H.323 supports Secure Real-Time protocol (SRTP) for media confidentiality, and Multimedia Internet Keying (MIKEY) for key exchange. It is important to emphasize that the signaling is only protected up to the gateway.

SIP is a more recent standard for multimedia conferencing over IP. The standard was defined by the Internet Engineering Task Force (IETF) [Ros02] and is conceptually simpler than H.323. SIP is used for creating, modifying and terminating sessions between endpoints. SIP supports SRTP for securing media traffic and TLS and S/MIME for signaling protection. Although most VoIP implementations today use the H.323 protocol for IP services, SIP is gaining more acceptance in the network telephony market due partly to its flexibility and lower implementation costs. It is possible to use each protocol alone or both protocols within the same network in order to provide universal connectivity.

Protocol standards imply an abstract architecture that can be used to guide the implementation of systems or products. We can describe such abstract architectures by using patterns. The abstraction power of patterns is useful to understand complex standards, to compare standards, and to analyze if a given product complies with the standard [Fer06]. An architectural pattern is also useful for simulation. We introduce here two patterns for call control and signaling in VoIP: the H.323 Signaling Protocol Architecture and the Hybrid SIP/H.323 VoIP Signaling Protocol Architecture. The latter addresses the interoperability and coexistence of H.323 and SIP in VoIP networks. Figure 1 shows a pattern diagram which relates these architectural patterns to other existing patterns in VoIP. The patterns described in this paper are indicated with a double contour. Since the Hybrid pattern subsumes the SIP pattern we do not discuss this latter separately. The patterns for Secure VoIP Calls and Signed Authenticated Calls can be found in [Fer07b]. Patterns for cryptography can be found in [Bra00].

We start with the H.323 protocol pattern in Section 2. We introduce the Hybrid pattern in Section 3. We end this paper with some conclusions.

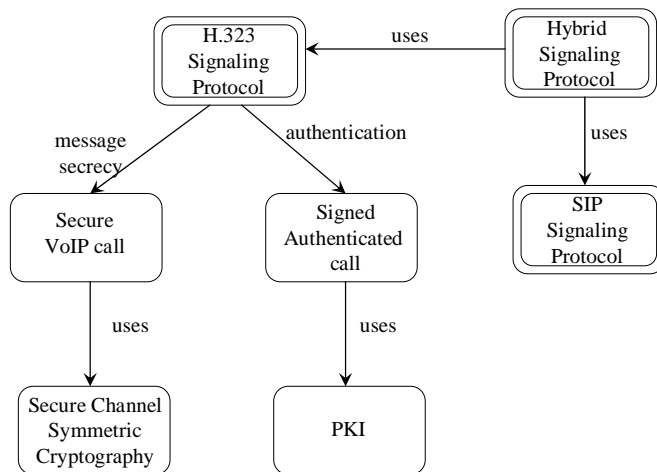


Figure 1 Relationships between VoIP architectural and security patterns

2. H.323 Signaling Protocol Architecture

This pattern describes an abstract generic architecture to support the H.323 VoIP signaling protocol. This protocol is used to set up and terminate voice calls, and to support the transport of voice, video and data packets over IP-based networks.

Context

In VoIP networks, voice and signaling are multiplexed and travel as normal data inside LANs, WANs or the Internet. Signaling protocols are required in packet networks for transport and control. The VoIP infrastructure is designed to initially support simultaneous users and is capable of scaling. This pattern assumes the availability of multiple Internet Service Provider (ISP) partners to provide edge termination for diverse users.

Problem

The H.323 protocol is rather complex and requires a combination of components to perform its functions. How can we structure the components and procedures required for delivering multimedia communication services (i.e. voice, video and data) across packet-based networks (e.g. Internet, intranets and Local Area Networks (LANs)? The solution to this problem is affected by the following **forces**:

- We need to define an abstract architecture that can be used to guide the design of products and systems.
- There is a need to maintain a stable and reliable transmission throughout VoIP conversations.
- Incompatible VoIP products are the result of the absence of industry standards within this technology. Standards need precise and clear expressions but the standards documents are textual and long descriptions that are hard to follow [ITU06].
- Interoperability with other multimedia service networks and terminals is vital in VoIP. Terminal devices in disparate networks (e.g. softphone communicating with an analog phone) communicate frequently.
- In order to transport real-time data over VoIP, call signaling is needed to set up the connections between terminal devices.

Solution

Define an abstract architecture using terminals (to make/receive calls), gateways (to connect different networks), gatekeepers (for control and setup) and multi-point control units (MCU) (for conferencing).

H.323-based networks have the ability to manage available resources for call routing via H.323 gatekeepers. Gatekeepers are used for address resolution, terminal devices admission control (based on bandwidth availability, concurrent call limitations, or registration privileges), bandwidth management, and zone management (the routing of calls originating or terminating in the gatekeeper zone, including multiple path reroute). Gateways coordinate calls by communicating with gatekeepers using the Registration, Admission, and Status (RAS) protocol [Cis02]. Gatekeepers are the central part of an H.323 network.

Structure

Figure 2 shows the UML class diagram of the H.323 architecture. The components inside the dotted lines indicate the specific units of the standard while the external units are the network components that participate in the whole system. The **Layer 2 Switch** provides connectivity between H.323 components and the rest of the system. The **Gateway** takes a voice call from a circuit-switched Public Switched Telephone Network (**PSTN**) and places it on the IP network. The PSTN uses **PBX** switches and **Analog Phones**. The **Internet** (IP network) contains **Routers** that connect to each other and **Firewalls** to filter traffic to the **Terminal Devices** (i.e. where users interact with the system). The gateway also queries the **Gatekeeper** via the Internet with caller/callee numbers and the gatekeeper translates them into routing numbers based upon service logic. Gatekeepers act like central managers providing call setup and routing the calls throughout the network to other voice devices. The **MCU** (Multipoint Control Unit) is used for conferencing. **Softphones** are applications installed in Terminal Devices (e.g. PCs or wireless devices) used to send/receive calls.

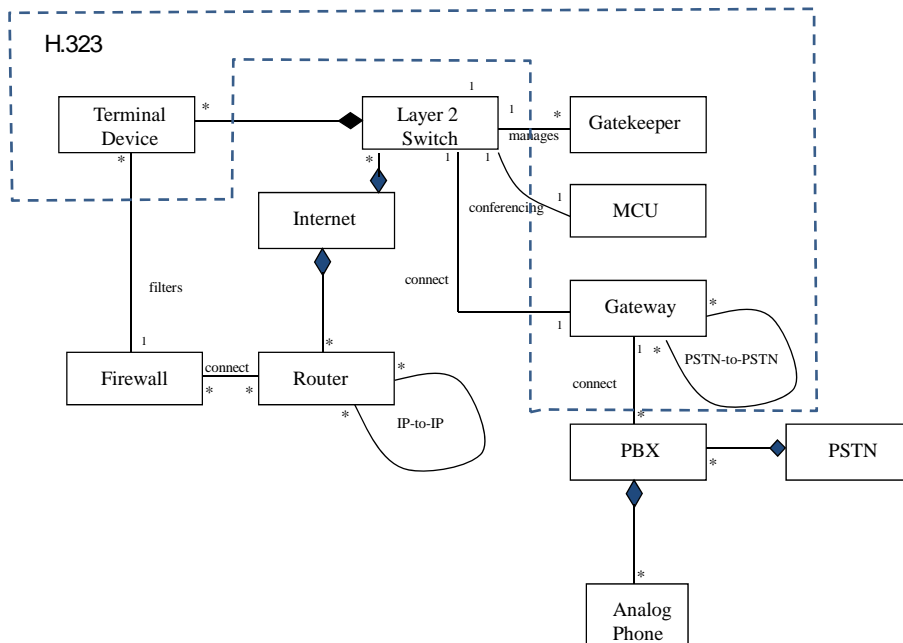


Figure 2 Class diagram for an H.323 architecture

Dynamics

The sequence diagram in **Figure 3** shows the necessary steps for call connection between terminal devices in H.323. When Terminal devices (i.e. Caller and Callee) are communicating with each other in disparate networks, more than one gatekeeper may be necessary, as shown in the figure.

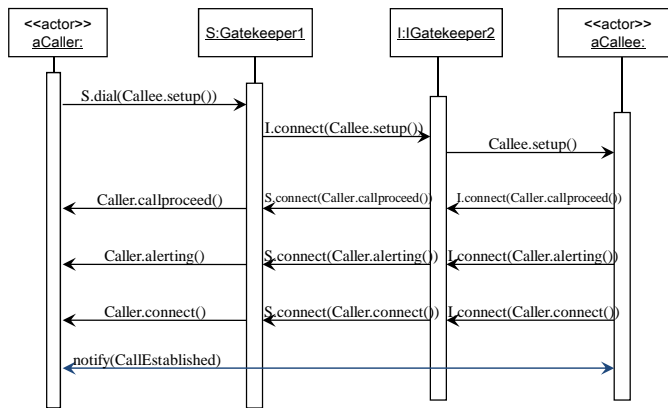


Figure 3 Sequence diagram for call connection in H.323

Known uses

Cisco [Cis00], and others.

Implementation

Most existing VoIP implementations use the H.323 protocol, and the same protocol is used for IP video (although SIP is more popular for new systems). H.323 references many other the ITU Telecommunication Standardization Sector (ITU-T) [ITU06] protocols such as:

- H.225.0 protocol is used to describe call signaling, the media (audio and video), the stream packetization, media stream synchronization and control message formats.
- H.245 control protocol for multimedia communication, describes the messages and procedures used for opening and closing logical channels for audio, video and data, capability exchange, control and indications.
- H.450 describes the Supplementary Services
- H.235 describes security in H.323
- H.239 describes dual stream use in videoconferencing, usually one for live video, the other for presentation
- H.460.17-19 describes firewall traversal in H.323
- H.261 H.263 H.264 describes video encoding

Consequences

This pattern has the following advantages:

- It is possible to establish a phone conversation between different domains and to use all type of telephony devices throughout IP networks.

- It contains components that can maintain a stable and reliable transmission between two or more VoIP users.
- The reference architecture lets vendors produce compatible products.
- The components in the H.323 architecture enable network interoperability (e.g. packet switch-to-circuit switch).

This pattern has the following disadvantages:

- H.323 is a rather complex protocol and even its abstract architecture is rather complex.
- H.323 defines several associated protocols and many services require interactions between those sub-protocols, which increases complexity and decreases scalability [Dal99].
- Security features are more easily implemented in SIP when compared to H.323 because of the SIP client-server operation mode.
- H.323 protocols frequently use ASN.1 encoding roles. The H.323 ASN.1 parser showed vulnerable to implementation level attacks [OUS04].

Related patterns

This pattern is related to the Hybrid VoIP Signaling Protocol Pattern described next and to the security patterns indicated in Figure 1. The model of Figure 2 is based on an early model in [Pel04]. An E/R model for H.323 is given in [OUS04].

3. Hybrid VoIP Signaling Protocol Architecture

The Hybrid Protocol Architecture pattern describes an abstract architecture to combine the H.323 and SIP architectures using a shared infrastructure of interworking functions between both protocols. This architecture allows coexistence and transparent translation of signaling between both architectures.

Context

In VoIP networks, voice and signaling are multiplexed and travel as normal data inside LANs, WANs or the Internet. Signaling protocols are required in packet networks for transport and control. The VoIP infrastructure is designed to initially support simultaneous users and is capable of scaling. This pattern assumes the availability of multiple Internet Service Provider (ISP) partners to provide edge termination for diverse users.

Problem

Some environments have a large variety of users and require the use of multiple signaling protocols. How do we design an architecture that provides support for both SIP and H.323 calls, is capable of scaling, and can be made secure?

The solution to this problem is affected by the following **forces**:

- VoIP must be able to accommodate multiple existing and potential signaling protocols.

- Provision must be made for the necessary security mechanisms in order to protect the VoIP network and its users. As VoIP operates on a converged network, voice and video packets are subject to the same threats than those associated with data networks [Pel04]. In contrast to PSTN, signaling in VoIP is sent through the public Internet. This leads to easy access to the voice packets (i.e. call interception) by attackers and its consequent security problems [Wie06].
- Because the interworking function combines both SIP and H.323 functionality, security considerations for both of these protocols apply [Sch04].
- Because all data elements in SIP or H.323 have to terminate at the interworking function, the resulting security cannot be expected to be end-to-end. Thus, the interworking function terminates not only the signaling protocols but also the security in each domain [Sch04].
- VoIP networks must be based on industry standards so as to provide functionality between disparate networks and product compatibility.
- Interoperability and coexistence between SIP and H.323 is essential in order to support new deployments that might use SIP as a substitute VoIP signaling protocol.
- In deployments where both protocols are used, it is important that there are no performance limitations related to the call mix between SIP and H.323 calls, and that there is no significant deviation in calls-per-second measurements compared to a homogeneous SIP or H.323 network [Cis02].

Solution

The key component of this architecture is defined as the interworking function (IWF) which provides this SIP-H.323 translation. The main functionality of this interworking function includes user registration, address translation, establishment of call connect, and service provision. This functionality can be implemented as part of a VoIP network server such as an H.323 Gatekeeper, a SIP Proxy, or a Softswitch, which might include a gatekeeper and SIP Proxy. Or, the functionality can be implemented via an external SIP-H.323 signaling gateway [Rad01].

Structure

The class diagram in **Figure 4** shows a hybrid configuration where a SIP architecture (in red and withing dotted lines) is combined with an H.323 network (in green). The gateway in the H.323 side takes the voice call from the PSTN network and routes it to the called party. In a similar way, the SIP protocol uses redirect or proxy servers for call routing. The proxy server provides security mechanisms for terminal devices such as access control, authentication and authorization. A hybrid signaling setting may involve two types of Endpoints: H.323 Terminal devices and SIP User Agents. **User Agents** (UAs), are combinations of User Agent Clients (UAC) and User Agent Servers (UAS). The UA is the phone in the SIP side and the **Register server** receives registrations and requests updates of the **Location server**, which keeps track of the UAs. A UAC is responsible for initiating a call by sending a URL-addressed INVITE to the intended recipient. A UAS receives requests and sends back responses. The **Proxy server** is connected to a **VoIP gateway** (to make possible a call from a regular telephone to an IP phone) and to other proxy servers. The registrar and location server may be integrated in the proxy server. Once the call has been established, the RTP media packets flow between endpoints.

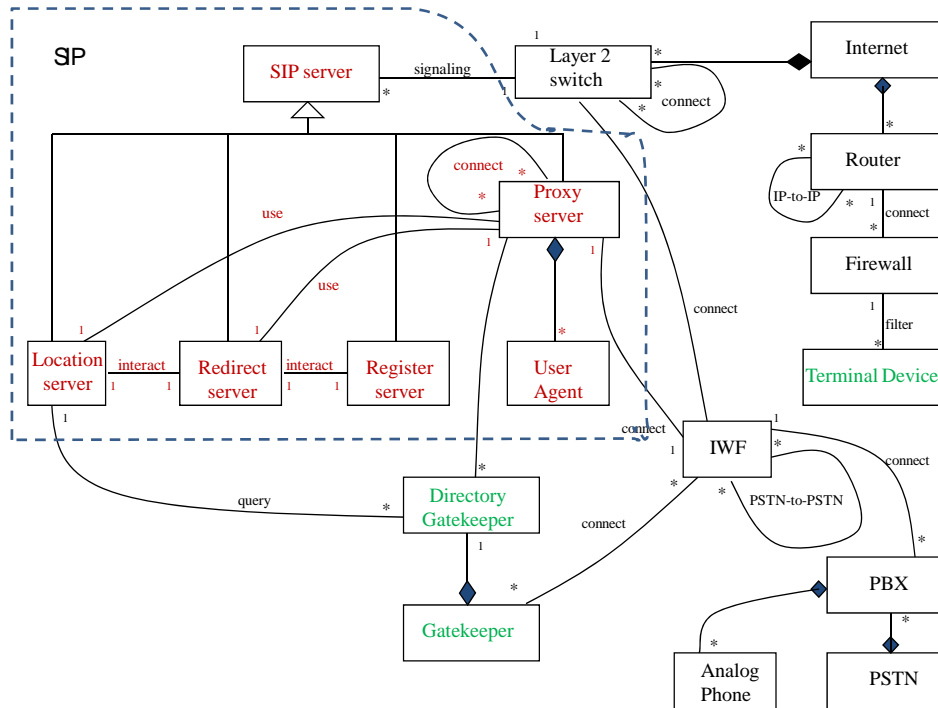


Figure 4. Hybrid VoIP Signaling Protocol Architecture

Dynamics

The sequence diagram in Figure 5 shows the necessary steps for calling from a SIP phone to an H.323 terminal device using call signaling between a SIP Proxy Server and a H.323 gatekeeper. As mentioned earlier, these servers provide registration and address resolution services. For all phases of the voice call, the interworking function component provides signaling translation.

Implementation

The implementation of VoIP services (e.g. Call waiting) must be uniform, consistent, and must effectively work with other signaling protocols. Existing H.323-based systems must update their existing signaling gateways in order to support additional SIP-based services.

The capability for H.323 gatekeepers and SIP proxies to interwork in VoIP sharing routing capabilities is crucial. While the SIP Proxy Server could supply routing information to SIP gateways, this scheme allows a packet voice carrier to not only use its existing routing structure on its H.323 gatekeepers, but to also take advantage of the H.323 Resource Availability Indicator (RAI) functionality for a more efficient network. The SIP Proxy Server actually acts like another gatekeeper to the H.323 network [Cis02]. Likewise, the SIP proxy server provides registration, address resolution, and a session initiation services. It is important to note that this type of message exchange between SIP and H.323-based components is only for signaling purposes. VoIP uses another standard the Real-Time Protocol (RTP) for transport of media packets between terminal devices.

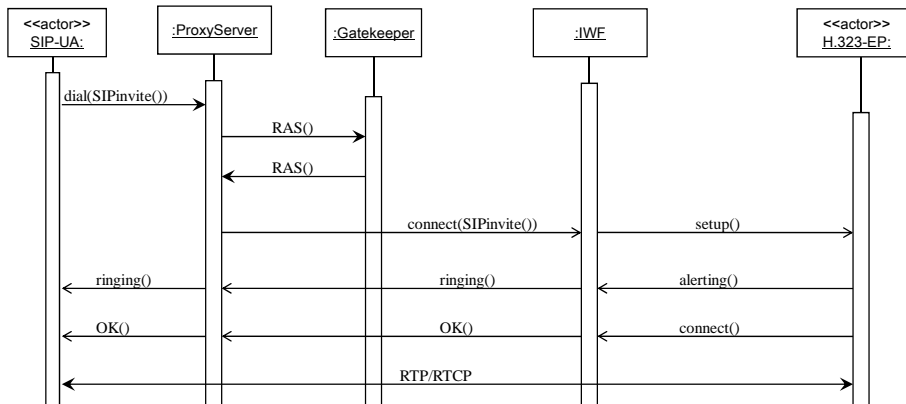


Figure 5 Sequence diagram for call connection in Hybrid configurations

Additionally to what is shown in Figure 5, the interworking function may provide and update the H.323 gatekeeper with the addresses of SIP UAs. Similarly, the interworking function can provide information about H.323 endpoints to a SIP registrar. This allows the SIP proxy using this SIP registrar to direct calls to the H.323 endpoints via the interworking function [Sch04] and vice versa. Provisions for communication between the SIP Proxy Server and H.323 gatekeepers, along with other features that allow SIP and H.323 to coexist on a common network, allow users the ability to build out hybrid networks that include both SIP and H.323 traffic.

Interworking can be achieved via multi-protocol endpoints (such as IP Phones that support both SIP and H.323) or via network bridging entities (such as Softswitches or Signaling Gateways) [Rad01]. When signaling messages are sent from an H.323 gatekeeper, the interworking function translates them into the corresponding SIP messages and routes them to the equivalent SIP component. By providing translation of signaling messages from SIP to H.323 and vice versa, the hybrid connects calls between VoIP devices using disparate signaling protocols. In this setting, address resolution and registration functions for both H.323 and SIP protocols, are supported by the interworking component.

An interworking function contains functions [Sch04] from the following list, inter alia:

- Mapping of the call setup and teardown sequences;
- Registering H.323 and SIP endpoints with SIP registrars and H.323 gatekeepers;
- Resolving H.323 and SIP addresses;
- Maintaining the H.323 and SIP state machines;
- Negotiating terminal capabilities;
- Opening and closing media channels;
- Mapping media-coding algorithms for H.323 and SIP networks;
- Reserving and releasing call-related resources;
- Processing of mid-call signaling messages;

- Handling of services and features.

In order to transfer of voice traffic over a packet switched network, some vendor offer a hybrid or IP enabled design using an existing telephone switch (TDM type), while others are based on a pure IP or IP centric architecture and only trunk into the local telephone switch. There are two basic architecture types: IP-enabled and IP-centric [Gha02]. The two are differentiated by:

- Where the PSTN to IP conversion takes place
- Where supporting call center functions, such as queuing, queue slots, prompting, music-on-hold, and announcements, take place.

IP Centric

This type of VoIP architecture is designed around an IP based core-switching system. These solutions have distributed IP devices that function together to perform the functions of a PSTN.

The TDM switching infrastructure is replaced with the VoIP infrastructure, and call control. In a multi-site deployment, a centralized call control server controls VoIP streams, whether they arrive at the main facility or the remote sites. In this deployment, the call control server at headquarters controls all inbound communications independent of location [Tip04].

IP enabled

IP enabled architectures allow traditional Time Division Multiplexing (TDM) switches to deliver voice over an IP network. This approach is an extension of the traditional TDM environment. In this approach the TDM switch is IP- enabled through the addition of IP trunk and/or line cards. This solution is usually considered by companies in order to extend the investment life of installed TDM switches [Tip04].

The implementation of the hybrid pattern is feasible using the IP-centric solution. However, it is far more difficult in an IP-enabled infrastructure, which requires several PSTN-to-IP gateways. In either case, it requires compatibility between the applications at endpoints. A key deciding factor in determining when and how to move to VoIP is the migration strategy. IP-enabled approaches will appeal to the more conservative and heavily invested call center environment, while IP-centric methods will be the choice of smaller, newer, and more aggressive centers who can bear some reliability risk to move faster to a more advanced platform [Gha02].

Known Uses

Vendor solutions that fall into the IP-enabled category include Aspect, Avaya, Nortel, and other switch providers. Cisco uses IP-centric solutions.

Consequences

The advantages of this pattern include:

- The hybrid pattern provides call connection between VoIP devices using disparate signaling protocols.
- The hybrid architectural pattern provides protocol flexibility for users to incorporate SIP networks on established converged infrastructures, while keeping H.323 functionality within their networks and interoperability with traditional PSTN networks.

- By using this pattern H.323-based carriers are able to incorporate new VoIP services to the existing infrastructure by allowing the SIP Server to register and acquire the address resolution with the H.323 gatekeeper using the RAS protocol.
- It is possible for a SIP Proxy Server to obtain updated routing information from VoIP gateways deployed in the hybrid network by enabling the server to communicate with an H.323 gatekeeper using the RAS protocol.
- This solution provides origination and termination of VoIP calls and more flexible billing and usage options for several other service providers.
- This solution support additional SIP-based services for H.323-based users that want to use their existing signaling gateway infrastructure.
- This optimized routing structure provides shorter post-dial delay and more efficient usage of gateway resources.

Possible disadvantages include:

- The combined architecture is complex.
- The deployment of hybrid protocol networks may be affected by QoS and bandwidth issues.

Related Patterns

The Hybrid VoIP Signaling Protocol pattern includes as a subpattern the H.323 pattern which was previously introduced. The model of Figure 4 is based on an early model in [Pel04]. The pattern is also related to the concept of Attack patterns [Fer07a] and to the following security patterns [Fer07b]:

- The Secure VoIP Call pattern.
- The Authenticated Call and other similar authentication patterns used to establish trust relationships between the VPN endpoints.
- The VoIP Tunneling pattern.
- The Network Segmentation pattern.

4. Conclusions

We have presented two patterns that describe the architectures implied by the two main VoIP protocols. The H.323 Signaling Protocol Architecture pattern offers a way to complement and support the transport of data, voice and video packets in VoIP systems. The Hybrid Signaling Protocol pattern allows architectural and protocol flexibility by supporting both H.323 and SIP. These patterns complement our work in VoIP security patterns [Fer07b] and provide a model of the environment where specific VoIP security patterns can be implemented, thus adding security to the structure. Patterns describing generic architectures can guide systems development, be used to evaluate existing designs, be a basis for simulation, and be a pedagogical tool. As indicated earlier, they can also be used to understand and compare standards.

Future work will include extending these architectures to describe security aspects of the Tactical Internet that includes wireless aspects and the development of attack patterns for that

environment. Another possibility is the development of simpler patterns that can be used as components in complex architectures, e.g. an IWF pattern (a suggestion of the shepherd).

Acknowledgements

We thank our shepherd Alan O’Callaghan for valuable comments and suggestions that significantly improved this paper.

References

- [Bra00] A. Braga, C. Rubira, and R. Dahab, “Tropyc: A pattern language for cryptographic object-oriented software”, Chapter 16 in *Pattern Languages of Program Design 4* (N. Harrison, B. Foote, and H. Rohnert, Eds.). Also in *Procs. of PLoP’98*, http://jerry.cs.uiuc.edu/~plop/plop98/final_submissions/
- [Cis02] Cisco Systems. “H.323 and SIP Integration”, March 2002. <http://www.cisco.com>
- [Dal99] I. Dalgic, H. Fang. “Comparison of H.323 and SIP for IP Telephony Signaling.” http://www.cs.columbia.edu/~hgs/papers/others/1999/Dalg9909_Comparison.pdf
- [Fer06] E.B.Fernandez and N. Delessy, ""Using patterns to understand and compare web services security products and standards", *Proceedings of the IEEE Int. Conference on Web Applications and Services (ICIW'06)*, Guadeloupe , February 2006.
- [Fer07a] E. B. Fernandez, J. C. Pelaez, and M. M. Larrondo-Petrie. “Attack patterns, a new forensic and design tool.” *Procs. of the Third Annual International Conference on Digital Forensics* (IFIP WG 11.9). January 28 – 31, 2007. Orlando FL. USA
- [Fer07b] E.B.Fernandez, J.C. Pelaez, and M.M. Larrondo-Petrie, “Security patterns for voice over IP networks”, *Procs. of the 2nd IEEE Int. Multiconference on Computing in the Global Information Technology (ICCGI 2007)*, March 4-9, Guadeloupe, French Caribbean.
- [Gha02] A. Gharakhanian. “Which VoIP Architecture Makes Sense For Your Contact Center?.” August 2002. www.vanguard.net
- [ITU06] International Telecommunication Union. “Packet-based multimedia communication systems.” ITU-T recommendation H.323. June, 2006.
- [OUS04] Oulu University Secure Programming Group, University of Oulu, Finland, PROTOS Test-Suite: c07-h2250v4, October 2004, <http://www.ee.oulu.fi/research/ouspg/protos/testing/c07/h2250v4/index.html>

[Pel04] J. C. Pelaez. *Security in VoIP networks*. Master's thesis. Department of Computer Science and Engineering, Florida Atlantic University, August 2004.

[Rad01] Radvision. "An Overview of H.323 - SIP Interworking." October 2001.
www.radvision.com

[Ros02] J. Rosenberg. "SIP: Session Initiation Protocol." Network Working Group. Request for Comments: 3261. June, 2002.

[Sch04] H. Schulzrinne. "Session Initiation Protocol (SIP)-H.323 Interworking Requirements draft-agrawal-sip-h323-interworking-reqs-07." Network Working Group Internet-draft. October 2004.

[Tip04] TippingPoint Technologies, Inc. "Intrusion Prevention: The Future of VoIP Security." June 2004. <http://www.tippingpoint.com>

[Wie06] C. Wieser, J. Roning, and A. Takanen, "Security analysis and experiments for Voice over IP RTP media streams", *Procs. of the 8th Intl. Symp. on System and Information Security (SSI'2006)*, Sao Jose dos Campos, Brazil, November 08-10, 2006.