December 2003

Middleware for Secured Video-Conferencing

Tarun Abhichandani
Claremont Graduate University

Bengisu Tulu
Claremont Graduate University

Samir Chatterjee
Claremont Graduate University

Jill Gemmil
University of Alabama at Birmingham

MIDDLEWARE FOR SECURED VIDEO-CONFERENCING

Tarun Abhichandani
Network Convergence Laboratory
Claremont Graduate University
tarun.abhichandani@cgu.edu

Bengisu Tulu
Network Convergence Laboratory
Claremont Graduate University
bengisu.tulu@cgu.edu

Samir Chatterjee
Network Convergence Laboratory
Claremont Graduate University
samir.chatterjee@cgu.edu

Jill Gemmill
University of Alabama at Birmingham
jgemmill@uab.edu

Abstract

Video-conferencing over IP networks is rapidly becoming a popular application. Currently, there are two standards for signaling that are used in such applications. H.323 is the signaling standard from ITU-T (used by most commercial video-conferencing system) and SIP, which is an IETF approved standard for voice and video communications. In this paper, we present federated security mechanisms as developed within a large project (Vide.Net) on Internet2. We discuss an actual SIP client architecture. Issues and techniques for authentication and authorization in SIP and H.323 based systems are explained. Moreover, we provide insights towards building a federated authentication/authorization model for secured video-conferencing. This federated model utilizes emerging SAML technology that promotes single sign-on authentication and is a novel approach for inter-realm authentication. Call flows depicting behavior of secured video-conferencing are enumerated.

Keywords: Authentication, authorization policies, federated administration, middleware, H.323, SIP, video-conferencing.

Introduction

Real-time applications that send and receive media (audio, video, instant messaging) are rapidly converging on the Internet. Among them, video-conferencing is a popular application that lets diverse group of people located at distributed sites to communicate with each other using video and audio. For video-conferencing, we need signaling protocols as well as media handling capabilities. Session Initiation Protocol (SIP) (Rosenberg, et al., 2002) and H.323 (http://www.itu.int, 2000) have been used for Voice over IP (VoIP) with SIP gaining popularity as a flexible session oriented protocol approved by the Internet Engineering Task Force (IETF). However, in the video-conferencing space, we could not find many academic or commercial applications that use SIP[1]. Most commercial video systems use H.320 protocol over ISDN lines or H.323 over ethernet. Only recently have we started to see the migration of these products to IP-based networks. Not only there is a need to develop and deploy SIP-based video-conferencing applications but also there are several requirements within the higher education community that must be met. These requirements include security, enterprise-level authentication, and having proper authorization policies in place to facilitate inter-campus video communications. Privacy and confidentiality of users is also needed.

The paper starts with explaining the design of a SIP-based video-conferencing application. This prototype implementation is being carried out as a collaborative project at Vide.Net (http://www.vide.net) over Internet2. Further, the paper substantiates requirements for a federated authentication system. After evaluating why a federated authentication is necessary it examines Shibboleth architecture, based on Security Assertion Markup Language (SAML) assertions, which provides us with capabilities of implementing federated security structure. In conclusion, the paper illustrates SIP-based and H.323-based call

[1] MSN Messenger is a SIP client from Microsoft.

flows using Shibboleth as a security mechanism. This paper is meant for researchers and practitioners who are interested in understanding developments in middleware on various applications.

Security Considerations for CGUsipClientv1.1

CGUsipClientv1.1 was developed by Network Convergence Lab (NCL) at Claremont Graduate University (CGU) using Dynamicsoft SIP stack (Dynamicsoft, 2001) and Java Media Framework (JMF) 2.1.1 Sun libraries (http://java.sun.com/products/java-media/jmf/, 2003). Dynamicsoft provides a comprehensive SIP stack including all the authentication mechanisms included in the latest RFC (Rosenberg, et al., 2002). JMF 2.1.1 libraries provides support for capturing and storing media data, controlling the type of processing that is performed during playback, and performing custom processing on media data streams (http://java.sun.com/products/java-media/jmf/, 2003). CGUsipClientv1.1 architecture has two main Java packages – *cgusip.client* and *cgusip.utils*. The *utils* package handles the existing instances of SIP connections and calls. The client package has three main components: *gui*, *sip*, and *media*. Structural details of SIP stack, media framework and client are explained in (Tulu, et al., 2003).

There is a growing need for binding authentication systems with applications such as video-conferencing for secured collaboration. The SIP authentication procedure is derived from HTTP Digest authentication (Franks, et al., 1999). It is a challenge-based mechanism; when a server receives a request, it may challenge the initiator of the request to provide assurance of its identity. The challenge contains a nonce value that is a string uniquely generated and used for one challenge only. Both the requestor and the server share a secret password, and the requestor uses this password, together with the nonce value, to compute a response value. The requestor sends the request again with the computed response value, which is used by the server to authenticate the request. The Digest mechanism uses a function to compute the response. Several algorithms can be indicated in the challenge but the default one is MD5 (Peterson, et al., 2003; Stefano, et al., 2002). CGUsipClientv1.1 provides authentication using native mode SIP authentication that uses the Digest mechanism with MD5 hashing (Tulu, et al., 2003).

University campuses and colleges use a variety of local authentication systems. LDAP and UNIX systems are few examples. A number of different passwords and PIN based systems are also in use. A complicated authentication system uses tokens that could be based on one-time passwords or clock-based passwords. Challenge-response mechanism is also used for authentication purposes. Indirect authentication as in RADIUS is also popular (Smith, 2001). These systems have an agent that accept tokens from a user and passes them to a server for authentication. Kerberos is another type of secured authentication system that uses encryption and tickets (Walla, 2000). A more sophisticated system is the Public Key Infrastructure (PKI) that uses public and private keys and digital certificates (Mel and Baker, 2001). More advanced authentication system may use biometric scanning like fingerprints or retina scans (Walla, 2000).

Applications that provide access to protected resources like video conferencing require users to be authenticated. If users belong to a local domain then they would have been assigned these credentials. Depending on these attributes the domain would decide on allowing access to protected resources. If a user does not belong to a local domain, there are related management issues that the local domain has to consider in addition to authorization policies for users that are external. Further, the end-user has to remember several usernames and passwords. Web services, as an example, allow users to access various applications and software components via standard web protocols. If more than one organization is involved in that environment and access to the site is restricted to some set of people then authentication warrants substantial management. Should every organization create a username and password for all the users? Is this scalable? Is it possible to make one organization trust the other and allow a user to reach restricted resources of a domain if they are authenticated in some other domain? Can one organization retrieve information about another organization's user to make an authorization decision? These questions show the need for a federated administration plan.

Federated Identity Management

The concept of federation was first developed within the Shibboleth project in Internet2 (Erdos and Cantor, 2001). Shibboleth, an open source project of Internet2/MACE, is meant to develop policy structures and authentication architectures to support a multi-boundary sharing of web resources in higher education. Shibboleth presents a framework that a multi-boundary security administration should provide for. Elements of the framework are:

- Cooperative authentication and authorization decisions between various domains leading to federated administration.

- Access control decisions based on attributes stored for the user.
- Standards based authentication and authorization transactions.
- Scaleable trust and privacy sets leading to community decision-making.
- Extensive pool of attribute storage for the user.

Shibboleth utilizes recent research implementations to achieve these elements.

- Implementation of Single Sign-On (SSO) solution.
- Using standards-based assertion language, SAML.
- Defining services in addition to SAML that produce or consume assertions.

Single Sign-On (SSO)

In a SSO solution, an initial action of user authentication will allow users to access all resources that they have permission to access without any need for a subsequent authentication. Authentication information with a specific structure will be given to the user as a proof of authentication to be provided to any other service that asks for authentication. SSO provides coordination between the authentication system and other services in an enterprise. This approach will reduce the possibility of sign-on failures caused by user errors, improve security by reducing the need for a user to handle and remember multiple sets of authentication information, reduce the system administration time, and improve system security by enhancing the ability of system administrators to maintain the integrity of user accounts. SSO requires applications to use a common security mechanism and make use of the user credentials for all the session access control requirements. One of the application using SSO features is Pubcookie, developed by University of Washington. Further documentation on Pubcookie can be accessed at (http://www.pubcookie.org, 2003).

Security Assertion Markup Language (SAML)

As explained, different campuses and enterprises are likely going to implement different types of authentication systems. These systems differ in their capability as well as complexity. However, if we assume that a client authenticates against some kind of a system, then this event or fact should be recorded and an "assertion" should be created which basically states that a particular end user has authenticated at a specific time using a specific system. Such assertions can be very valuable in inter-realm authentication as is often the case with video-conferencing. Further, these assertions should be capable of providing details stored in attributes when asked for by any domain, origin or destination. An emerging technology called SAML has been developed to the assertions (Daniel, 2002). Further information on SAML can be found at (http://www.opensaml.org/, 2002).

Services in Shibboleth

Shibboleth exchanges user attribute information between administrative domains using SAML (Daniel, 2002). SAML is designed for bilateral exchange of assertions between two domains. Shibboleth, using SAML, builds a community in which authentication and authorization is enforced among various domains (http://www.simc-inc.org/archive0002/February02/Speakers/hill/tsld023.htm, 2002). Shibboleth provides following services (http://shibboleth.internet2.edu, 2003):

- Shibboleth Attribute Requestor (SHAR).
- Shibboleth Indexical Reference Establisher (SHIRE).
- Where Are You From (WAYF) service.

Further discussions on Shibboleth can be found at (Erdos and Cantor, 2001).

SAML/SIP Role-based Authorization

The usage scenario envisioned in this paper consists of a user in one domain establishing a session with a user in another domain through the use of a role-based authorization mechanism. The requirements of this mechanism are that it functions seamlessly across domains; is as flexible and granular as possible to facilitate multiple access levels; focuses on the role of the user rather than identity; does not create additional security vulnerabilities; and does not burden system administrators

with additional responsibilities. Furthermore, care must be taken to minimize changes to and ensure interoperability with existing protocols. An additional goal will be to ensure that such authorization can occur in an end-to-end manner between user agents.

To meet these requirements, entities will assert roles (defined by their attributes) between domains. These assertions will be carried by the SIP messages. The format used for the assertions will be based on the SAML. Since SAML was designed so that it can be carried within other protocols, the marriage of SAML and SIP is a logical and useful next step. New conceptual entities need to be defined within the SIP architecture, such as role-based Policy Decision Points (PDP) and Policy Enforcement Points (PEP). In most real-time communication sessions, the target user is an individual PDP by default, as the ultimate decision to accept or decline a call lies with the user (http://shibboleth.internet2.edu, 2003). The amount of information transferred across domains about the user should also be in accordance with the privacy policies of the local domain and the user. Thus, the information transferred to the remote domain about the user should be just the minimum required for authorization decisions to be made. This approach extends the ability of SIP networks to provision privacy in a way not presently possible (Chatterjee S. et al., 2003).

The SIP proxy server can provide the role of verifying user authentication and authorization for locally defined policies; perhaps a pre-existing SSO credential could be used for transparent login. Upon attempting to access a remote resource such as a target SIP proxy server, the target resource would then need to communicate with the calling user's home attribute authority, which would contain the attributes and release policies associated with that user (Chatterjee S. et al., 2003). This specific manner in which the transfer of assertions/attributes would take place has to be defined. Mappings from SAML request-response message exchanges into standard messaging or communication protocols are called SAML protocol bindings (or simply bindings) (Daniel, 2002). Hence in order to use SAML, SIP bindings for SAML need to be defined. In SAML parlance, a set of rules describing how to embed and extract SAML assertions into a framework or protocol are called profiles of SAML. A SIP profile of SAML describes how the originating PEP can embed SAML assertions/artifacts in SIP messages, communicated to the destination PDP and subsequently processed by the destination PDP and the PEP. Thus, as has been suggested earlier, some changes to SIP may be required to support this model.

Federated authentication and authorization in video-conferencing systems

We now describe the call flow details for SIP-based video-conferencing systems using federated identity management as shown in Fig. 4.

1. The user "logs on" to the network using the organization's centralized authentication service. A web browser might be used to do so, and could leverage HTTP-based Shibboleth installations.
2. The User Agent (UA) is invoked, credentials are deposited and the UA registers with a SIP registrar.
3. The browser connects to the web resource. This would typically be a global white pages directory of users who could be dialed using SIP Uniform Resource Identifiers (URIs). The resource may be protected by Shibboleth.
4. The Presence Server returns the SIP URI of the person to be contacted. Between steps 3 and 4, Shibboleth is invoked. The target web server SHIRE is presented an anonymized "handle" for the requesting user and the address of the user's Attribute Authority. The SHAR obtains the necessary attributes associated with the handle, and those attributes are passed to a Resource Manager for a Grant/Deny decision to provide the target SIP URI. (Details of Shibboleth can be found in (Erdos and Cantor, 2001))
5. The UA sends the INVITE. The INVITE may contain a credential or the origin proxy may intercept and require a re-INVITE with proxy provided credentials.
6. The Origin Proxy passes the INVITE to destination proxy.
7. The Target Proxy queries the Relying Authority asking for authorization. Note that this interaction is outside SIP domain. There are two ways that this can be done. One can use SAML bindings over Hypertext Transfer Protocol (HTTP) or Simple Object Access Protocol (SOAP). Another alternate way would be to define SAML bindings specifically for SIP. (Such a proposal has recently been made in (Peterson, et al., 2003)).
8. Relying Authority contacts the Issuing Authority and gets back an Attribute Assertion.
9. The Relying Authority conveys a decision to the Target Proxy, based on the released attributes and local policies. In this case the Relying Authority acts as the PDP while Target Proxy is the PEP.
10. Based on decision, the INVITE is passed to the SIP UA.
11. Session begins.

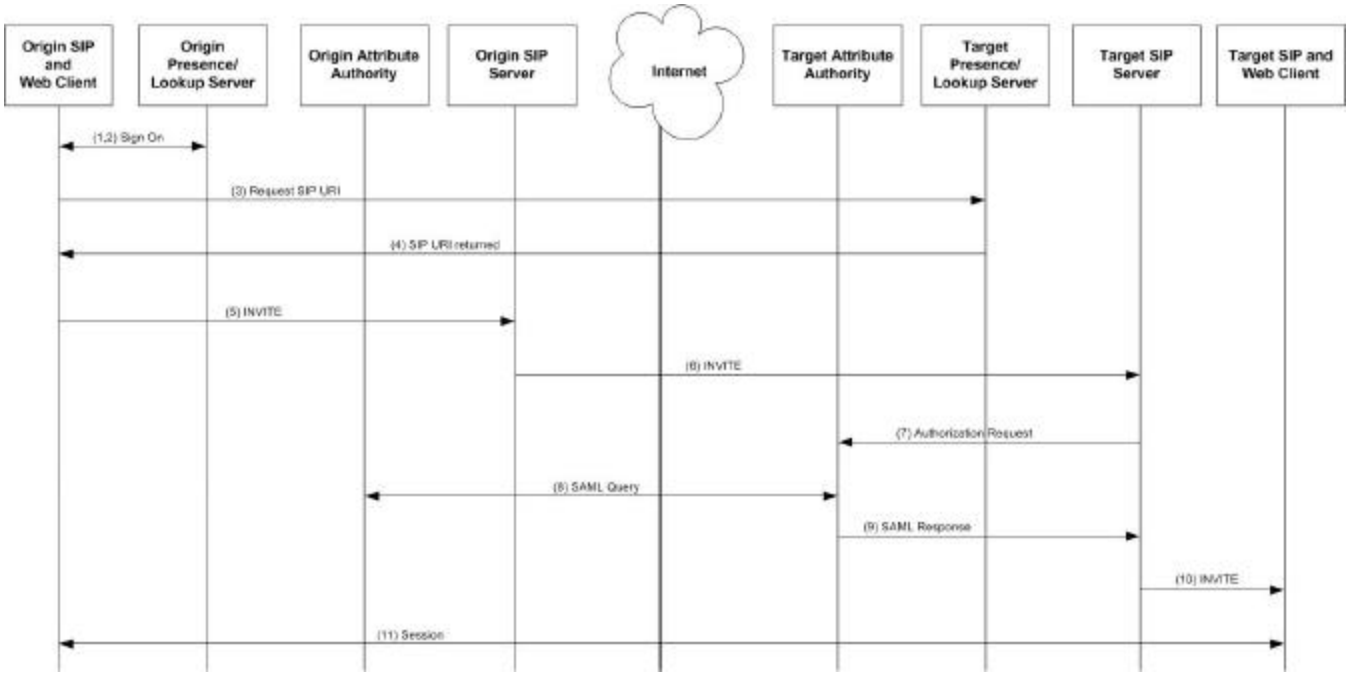

Figure 1. Call flow signaling diagram in SIP showing authentication, authorization using SAML (adopted from (Chatterjee S. et al., 2003)).

A similar call flow for H.323 is shown in Figure 5.

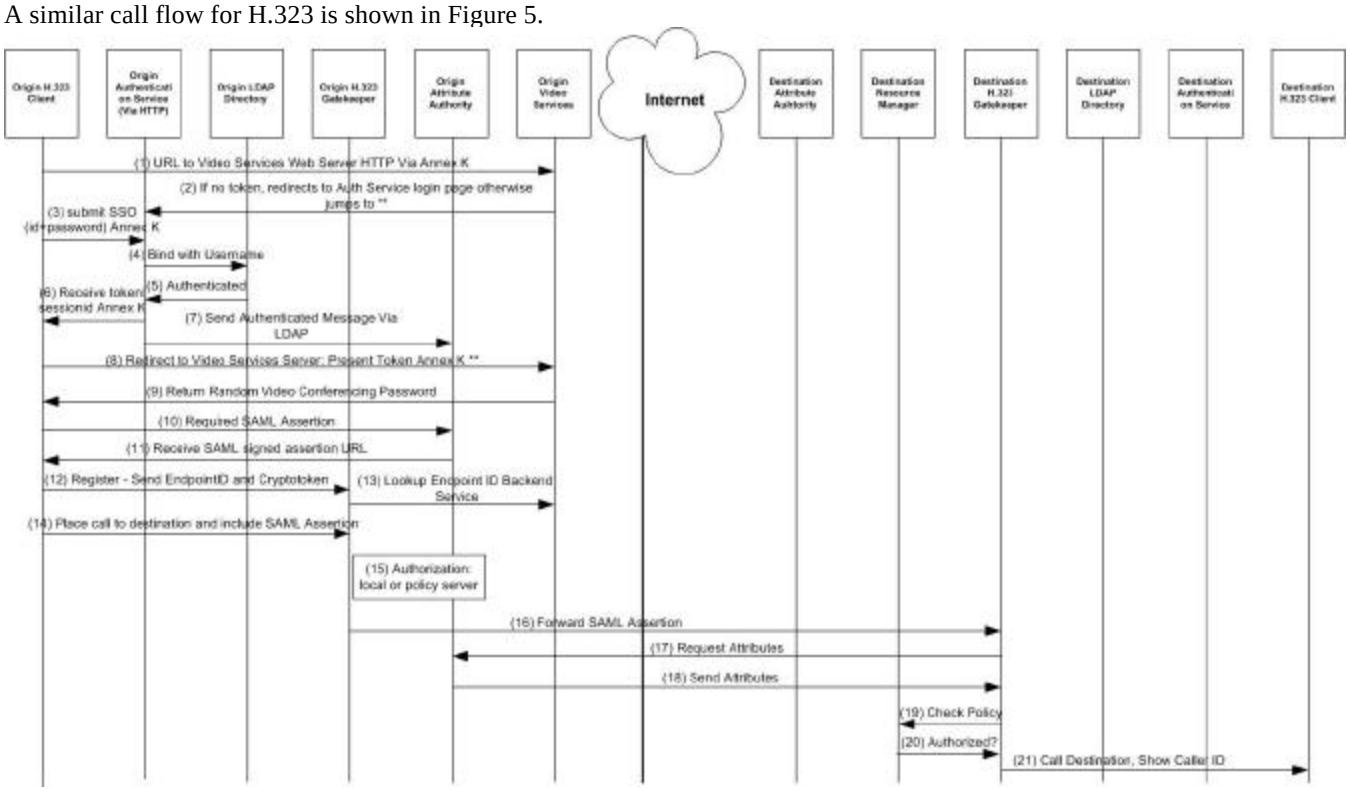

Figure 2. Call flow and details of implementation for H.323-based video-conferencing systems.

Conclusion

Our collaborative research group (Gemmill, et al., 2003) has been jointly developing several of the above mentioned technologies. Some outcomes of our research include:

- A SIP-based video-conferencing desktop client (See Figure 6).
- Directory-enabled dialing services for each campus domain.
- Implementing authentication in commercial H.323 video-conferencing system (this is being developed by a leading commercial vendor, RADVISION).
- Ongoing implementation of federated security in CGUsipClientv1.1.

As part of Internet2 Middleware program, our group is primarily responsible for developing the necessary middleware that will provide security services for large-scale video-conferencing systems. Since both SIP and H.323 protocols have become dominant signaling standards for voice and video over the Internet, we are building actual clients that will have the above discussed authentication and authorization mechanisms. The use of SAML assertions to create a federated identity management system is expected to scale very well. This article has described the high-level challenges we are facing for implementing a federated security middleware and the of how we are implementing such systems. Future articles will report on the case study and performance aspects of these systems as they get widely adopted by higher educational institutions for campus-wide video-conferencing.

Figure 3. Snapshot of CGUsipClientv1.1 while on a call.

References

Chatterjee S., Sicker D., Gemmill J. "Federated Identity Management and Role-based Authorization for SIP-based Video-conferencing Systems (Working Paper)", 2003,

Daniel, B. "Plan on SAML for Identity Management," *Network World*, August 2002,

Dynamicsoft *Dynamicsoft Proxy Server 5.2 Administrator's guide*, 2001.

Erdos, M., and Cantor, S. "Shibboleth Architecture," Draft v04, Internet2, November 26 2001.
Franks, J., Hallam-Baker, P., Hostetler, J., Lawrence, S., Leach, P., Luotonen, A., and Stewart, L. "HTTP Authentication: Basic and Digest Access Authentication," RFC 2617, Internet Engineering Task Force, June 1999.
Gemmill, J., Chatterjee, S., Johnson, T., and Verharen, E. "ViDe.Net Middleware for Scalable Video Services for Research and Higher Education," *Proceedings of the ACM Southeastern Conference*, 2003,
http://java.sun.com/products/java-media/jmf/ "Java Media Framework API," (2003:May 30), 2003,
http://shibboleth.internet2.edu "Shibboleth," (2003:May 29), 2003,
http://www.itu.int "Packet based multimedia communications systems," Recommendation H.323, International Telecommunications Union, November 2000.
http://www.opensaml.org/ "SAML 1.0 Specification Set," November 5 2002.
http://www.pubcookie.org "Pubcookie," (2003:May 29), 2003,
http://www.simc-inc.org/archive0002/February02/Speakers/hill/tsld023.htm "Shibboleth: SAML-based federated administration,"), 2002,
Mel, H.X., and Baker, D. *Cryptography Decrypted*, Addison-Wesley Publication Co, Boston, USA, 2001.
Peterson, J., Polk, J., and Sicker, D. "Role-based Authorization Requirements for the Session Initiation Protocol," SIPPING Working Group Internet Draft Internet Engineering Task Force, February 24 2003.
Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, N., and Schooler, E. "SIP: Session Inititation Protocol," RFC 3261, Internet Engineering Task Force, June 2002.
Smith, R.E. *Authentication: From Passwords to Public Keys*, Addison-Wesley Publication Co, Boston, USA, 2001.
Stefano, S., Luca, V., and Donald, P. "SIP Security Issues: The SIP Authentication Procedure and Its Processing Load," *IEEE Network* (16:6), 2002,
Tulu, B., Abhichandani, T., Chatterjee, S., and Li, H. "Design and Development of a SIP-Based Video Conferencing Application," *Proceedings of the 6th IEEE International Conference on High Speed Networks and Multimedia Communications HSNMC'03*, Estoril, Portugal, 2003,
Walla, M. "Kerberos Explained," *Windows 2000 Advantage*, May 2000,