Using SAML to Protect the Session Initiation Protocol (SIP)

Hannes Tschofenig and Rainer Falk, Siemens AG
Jon Peterson and Jeff Hodges, NeuStar, Inc
Douglas Sicker, University of Colorado at Boulder
James Polk, Cisco Systems

Abstract

The Security Assertion Markup Language (SAML) standard supports the expression of security assertions such as authentication, role membership, or permissions. SAML assertions may be used to realize single-sign-on between Web servers located in different domains. After a short introduction to SAML, this article describes the application of SAML to protect Session Initiation Protocol (SIP) signaling.

The Session Initiation Protocol (SIP) is an application-layer signaling protocol defined by the Internet Engineering Task Force (IETF). It is used in particular for setting up voice-over-IP (VoIP) calls, but it may also be applied to such uses as establishing instant messaging sessions. The protocol allows for locating of other SIP-aware entities, and establishing, maintaining, and terminating communication sessions. The base SIP specification [1] reuses other protocols to provide security protection, including Transport Layer Security (TLS) and username/password DIGEST-based authentication. This allows unilateral or mutual authentication between a SIP client and a SIP proxy, thus preventing unauthorized use of services offered by a SIP proxy. These security mechanisms are used to protect SIP signaling traffic, while protection for data traffic is provided by different mechanisms, such as MIKEY for authentication and key exchange and SRTP for media traffic.

In many environments, it is desirable to reuse already deployed authentication credentials and user/policy databases in architecting the security for new SIP-based service deployments. This has been done specifically for the SIP-based 3GPP IP-based multimedia subsystem (IMS), which reuses the cellular 3GPP security infrastructure, which comprises on the client side a user authentication module (e.g., ISIM, the subscriber authentication module for IMS), and the authentication and user profile servers referred to as the home location register (HLR) and home subscriber server (HSS), respectively.

This article describes a more generic solution for SIP authentication and authorization that is based on asserted traits. This approach provides a richer framework for authorization, and allows for greater privacy of users. Technically speaking, the core part consists of the Security Assertion Markup Language (SAML), an XML-based format for expressing attributes, and a query-response protocol to request or retrieve assertions (or references to assertions called artifacts). The initial driver for SAML was the Web-based single-sign-on usage scenario [2].

This article is structured as follows: we describe SAML and its use for Web-based single-sign-on. After describing general SIP authentication, we describe how SAML can be included in SIP signaling and bound close to a SIP context.

SAML-Based Web Single-Sign-On

This section briefly describes the Security Assertion Markup Language (SAML), an XML-based framework for representing and exchanging security information, and how it is used to realize Web-based single sign-on. [2].

A SAML assertion encodes security information about an entity. An assertion may contain multiple assertion statements. There are three kinds of SAML assertion statements:

- Authentication statements describe a subject authentication event (e.g., when, by whom, via which authentication mechanism)
- Attribute statements provide details of the subject (e.g., the department in which the subject works)
- Authorization decision statements indicate whether the subject has permission to access a particular resource

Thus, one can employ SAML to encode statements such as, "Alice has these profile attributes and her domain's certificate is available over there, and I'm making this statement, and here's who I am."

One can then cause such an assertion to be conveyed to an interested party who can use this information to make decisions about such things as access to certain resources. For example, one might make use of the contents of a SAML statement to determine whether a user should be provided access to a restricted Web resource. This is done in a particular "context of use." Such a context of use could be, for example, deciding whether to accept and act upon a SIP-based invitation to initiate a communication session.

The specification of how SAML is employed in a particular context of use is known as a "SAML profile." The specification of how SAML assertions and/or protocol messages are concretely conveyed in, or over, another protocol is known as a "SAML binding." Typically, a SAML profile specifies the SAML binding(s) that are to be used in its context. Specifica-

0890-8044/06/$20.00 © 2006 IEEE

tion of both (or either) SAML profiles and SAML bindings are by definition built on the foundation provided by the SAML specifications, namely, the SAML Assertions and Protocols specification, also known as "SAML Core" specification.

Web single sign-on was the domain of the initial SAML profiles. We briefly describe a typical scenario below.

The three main entities involved in the Web single sign-on are:

- A user running a Web browser
- An asserting party (often termed an identity provider)
- A relying party (generally, a Web site)

Typically, the subject's browser is interacting with the relying party, whereupon the relying party wishes to obtain attestation from an identity provider as to the user's identity. The relying party sends a SAML Authentication Request message to the identity provider via an HTTP Redirect Response routed through the user's browser. The identity provider authenticates the user if it has not already done so, and sends a SAML Authentication Response message containing a SAML assertion back to the relying party. The subject of this SAML assertion typically denotes the identity of the browser's user. The relying party evaluates the information conveyed in the SAML assertion against local policy and decides whether or not to provide the user access. An assertion obviously needs integrity protection. It is either transported over a secure connection, using for example SSL/TLS, or it may be digitally signed, or both.

The overall message flow of this SAML profile is depicted in Fig. 1.

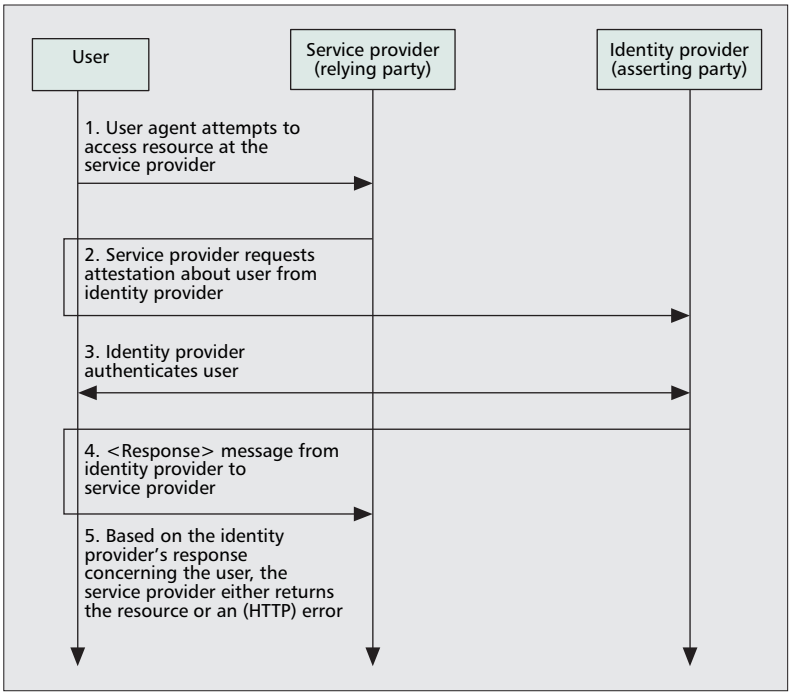


Figure 1. *SAML in an HTTP environment.*

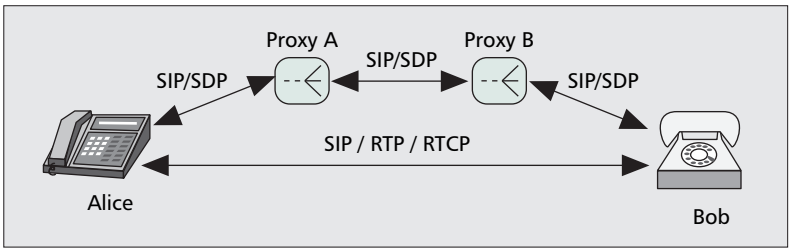


Figure 2. *SIP trapezoid.*

Local to each Web site, the session semantics can be realized with well-known technologies for maintaining local session state, for example using "cookies," or via "URL rewriting." The security of the SAML-based Web single-sign-on profiles is based on following model. A trust relationship exists between the asserting and relying parties. The assertions can contain information addressing them to a specific relying party, thus misuse can be detected. For example, the original relying party cannot simply reuse an assertion by presenting it to another relying party (assuming that the relying party is properly evaluating the assertion contents and rejecting those not explicitly addressed to it).

As we demonstrate in this article, applying SAML to SIP necessitates deriving a new SAML profile rather than simply reusing the existing Web single sign-on profiles, since the communication model of the two protocols are fundamentally different.

Authentication for SIP

Developing a security solution for SIP requires that we consider the communication model of SIP; specifically differentiating between hop-by-hop and end-to-end issues (and combinations such as end-to-middle, middle-to-middle), and security for signaling and for media traffic.

In the classical SIP trapezoid, as shown in Fig. 2, SIP signaling communication between two endpoints involves SIP proxies, while the media traffic is exchanged directly.

The communication between a SIP User Agent ("UA" in Fig. 2) and an outbound proxy ("Proxy A") can be secured in a variety of ways depending upon the deployment environment (e.g., SOHO, enterprise networks or public VoIP providers, and 3GPP IMS). Security between proxies is likely to be accomplished using TLS or IPsec.

Authentication of Alice to Bob (in an end-to-end fashion) and the establishment of media traffic is typically more complicated if the communication is between two generic endpoints. This is due to the lack of a global public key infrastructure, including the limited usage of client certificates. Additionally, existing deployments often use shared secrets, for example, via Kerberos or other deployment dependent authentication protocols (e.g., UMTS Authentication and Key Agreement protocol in the 3GPP environment [3]).

Since end-to-end authentication is not a realistic option today, it seems that a migration path using an interim solution is required. The core idea of this interim solution is accomplished with the introduction of a SIP Authentication Service logical role, which is typically played by a SIP outbound proxy. SIP user agents send requests through an Authentication Service, which:

- Authenticates the user according to a set of practices
- Creates and cryptographically signs an authentication token for the user
- Shares that identity with others

This approach is illustrated in Fig. 3, where Alice sends a SIP INVITE message through the outbound Proxy A. Alice is authenticated and authorized by Proxy A, which also acts as a

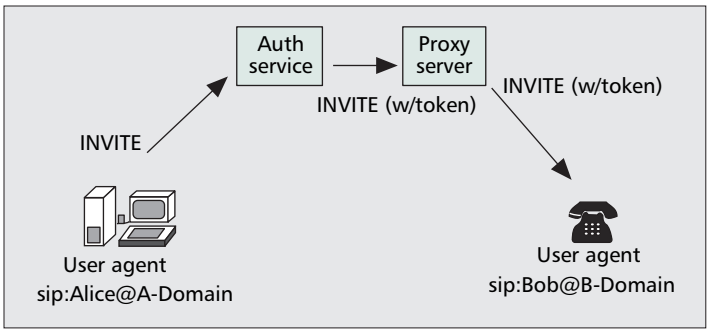


Figure 3. *Usage of authenticated identities in SIP.*

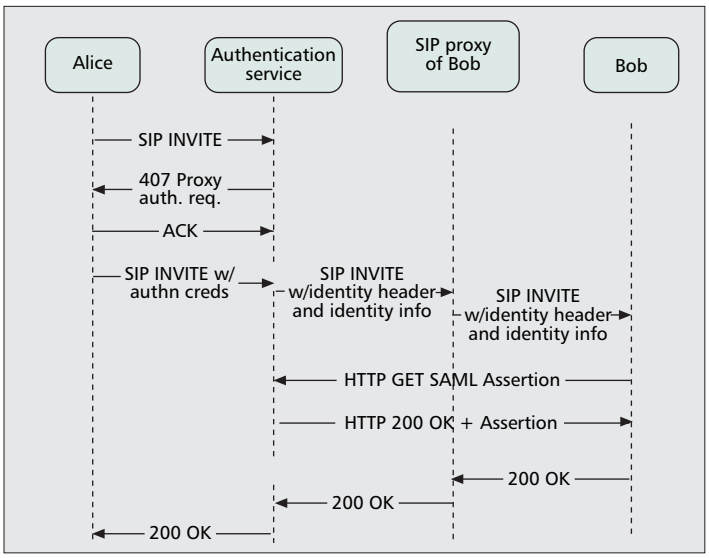


Figure 4. *SIP SAML usage.*

SIP Authentication Service. Thus it will assert the user's identity, and that she was authenticated, by adding a digitally signed token to the SIP message. The digital signature is computed over a number of additional fields of the SIP message in order to protect their integrity and that of the overall message. Then, the SIP message, including the asserted identity (denoted as "token" in Fig. 3), is sent to proxy B. Proxy B (and possibly Bob's user agent) inspects the content of the token. The processing steps require verification of the digital signature. Note that the main advantage for Proxy B and Bob is that they only require knowing, or having to discover, the certificates of the Authentication Services they interact with, rather than certificates of each individual end user. This obviously aids in scaling.

Since the Authentication Service asserts the user's identity, this approach seems to play against SIP's user privacy features. In SIP, privacy is the withholding of identity information from recipients of a SIP message. Private requests can still lead to a SIP dialog, but should not allow the originator of a message to be contacted by the recipient outside of the dialog. Hence, these functions also have to be integrated with the goal of asserting only that a particular user was authenticated, but without stating that authenticated identity.

Services like network authentication and privacy need to be inserted into the path of the SIP request. While the client can complete some privacy functions, others are the responsibility of the network. From a deployment perspective it seems that it is best if these services are colocated with a local outbound proxy for an administrative domain; although privacy may require services provided outside the local administrative domain (e.g., onion routing). SIP Authentication Services should use the same credentials as are provided when a user registers and it is suggested that the Authentication Service and registrar might be colocated. Forcing requests to pass through an Authentication Service has the effect of creating a de facto notion of SIP single sign-on. The development of the full details of the mechanisms described in this section is the subject of ongoing work within the IETF SIP working group [4]. Further information can be found on the group's Web site.

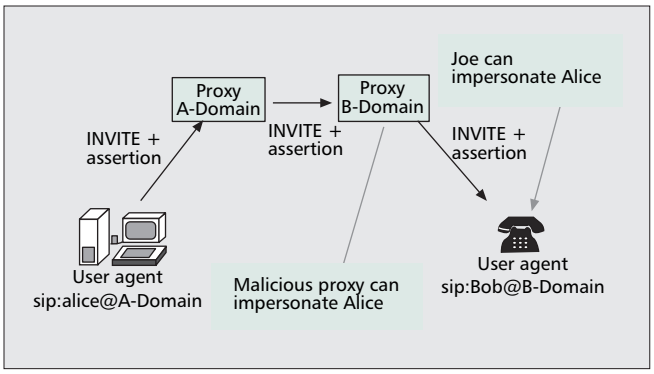
Introducing SAML to SIP

In the previous section we described how basic identity information could be asserted. However, additional identity information can also be asserted according to the same style. The requirements and scenarios for asserting additional identity information, referred to as roles or traits, are described in [5]. Fortunately, we can borrow from the existing identity assertion work of SAML. This work defines the conveyance of references to SAML assertions in SIP messages in order to build the foundation for trait-based authorization.

Figure 4 illustrates the SIP SAML profile via an example of Alice wanting to call Bob. Alice authenticates with the Authentication Service, which then forwards Alice's SIP INVITE message on to Bob's inbound SIP proxy. This SIP message includes Alice's identity information as blessed by her proxy, along with a reference to a SAML assertion, which asserts various traits of Alice and points to Alice's domain certificate. If the assertion and domain certificate pass verification by Bob's inbound proxy, then the call setup continues.

- In steps 1 and 2, an authentication and authorization process is executed between Alice's SIP UA and the Authentication Service, also known as the Asserting Party. User authorization by the Asserting Party is important in order to be able to create the SAML assertion and the respective attributes. The SIP UA must ensure that the Asserting Party is genuine.
- In step 3, the Asserting Party verifies the identity information in the SIP INVITE message received from Alice's UA and generates a SAML assertion attesting to various of Alice's attributes per local configuration. The assertion is held for later retrieval by the Relying Party, and the SIP INVITE message is modified per the processing rules of [6, 7]. This includes placing an HTTP URL reference to the aforementioned SAML assertion in the "identity-info" header field of the SIP INVITE message.
- In step 4, Bob receives the SIP INVITE message. It extracts the URL from the identity-info field and dereferences it using an HTTP GET.
- In step 5, the SAML assertion is returned in the HTTP response message. Bob receives it and verifies it, and the domain certifies the assertion references, according to the processing rules of [6, 7].
- In step 6, if the verification of step 5 succeeds, a SIP response with a success status code ("200 OK") is returned to Alice's UA, and call setup proceeds.

A SIP message, such as an INVITE, may traverse zero or more intermediaries on its journey from the initiator to the ultimate recipient (e.g., Bob's UA in the above examples). Any of these entities, including Bob and his UA, may not be entirely trustworthy. Since a SIP INVITE message as described above contains an HTTP URL with which one may easily retrieve the associated SAML assertion, any of the entities that handle the SIP message will be able to retrieve the assertion and associated domain certificate. Additionally,



■ Figure 5. *Vulnerability without reference integrity.*

since the HTTP-based infrastructure also commonly involves proxies, one of those entities could intercept a returned assertion. The attacker could then conceivably attempt to impersonate the subject (e.g., Alice) to some SIP-based target entity (Fig. 5).

Such an attack is implausible for several reasons. The primary reason is that a message constructed by an impostor using a stolen assertion, which conveys the public key certificate of a legitimate domain, will not verify because the impostor will not have the corresponding private key with which to generate the signed SIP Identity header value.

Also, due to the assertion content stipulated in [6], termed a "SAML assertion profile," the assertion will not be useful to arbitrary parties. This is because the assertion:

- Is digitally signed, thus causing any alterations to break its integrity, making them detectable
- Does not contain an authentication statement
- Identifies the targeted relying party
- Identifies the assertion issuer
- Explicitly stipulates its validity period
- Contains or refers to the originating user's domain's public key certificate

For property (1), the assertion is to be signed by the same key used to sign the SIP Identity header, which is stipulated in [6]. This binds the assertion to the subject identity (i.e., the caller's) being asserted by the caller domain's outbound SIP proxy/AS. Property (2) means that no parties faithfully implementing [5, 6] should be relying on SAML assertions (as specified in [5]) as sufficient in and of themselves to allow access to resources. Due to properties (3) and (4), an entity receiving such an assertion is able to ascertain whether the assertion was targeted to them, as well as who originated it, and thus make an informed decision whether to proceed with SIP session establishment. Property (5) is simply the well-known and oft-used technique of having security tokens explicitly reflect the time period within which they may be relied upon. In addition to all the above, with property (6) the assertion refers to or actually contains the user's domain's public key certificate. Note that this linkage is protected by the signature on the assertion, and the reference to the assertion in the SIP message's identity-info field (as well as several other SIP header fields) is protected via signature. Thus there is a verifiable chain from the SIP message to the user's domain's public key certificate. If any of the links in the chain do not verify, then a relying party should not continue with SIP session establishment.

Summary

This article has described a method for using the Security Assertion Markup Language (SAML) in collaboration with SIP. The flexibility of SAML assertions allows for the encoding not only of identity information about the user, but also generic authentication and authorization attributes in order to accommodate richer authorization mechanisms and enable trait-based authorization. The authorization decision might therefore be based on traits in addition to the identity. For individuals interested in following the development of SAML as a method for protecting SIP, please refer to the IETF working group cited in [4].

References

[1] J. Rosenberg *et al.*, "SIP: Session Initiation Protocol," RFC 3261, June 2002.
[2] See http://www.oasis-open.org/committees/download.php/6837/sstc-saml-tech-overview-1.1-cd.pdf
[3] 3GPP TS33.102 "3G Security — Security Architecture," Release 7, Dec. 2005.
[4] IETF SIP Working Group, http://www.ietf.org/html.charters/sip-charter.html
[5] J. Peterson *et al.*, "Trait-based Authorization Requirements for the Session Initiation Protocol (SIP)," Feb. 2005 (work in progress). draft-ietf-sipping-trait-authz-01
[6] H. Tschofenig *et al.*, "Using SAML for SIP," June 2006 (work in progress), draft-ietf-sip-saml-00.txt
[7] J. Peterson and C. Jennings, "Enhancements for Authenticated Identity Management in the Session Initiation Protocol (SIP)," Oct. 2005 (work in progress), draft-ietf-sip-identity-06.txt

Biographies

HANNES TSCHOFENIG received a university diploma in computer science from the University of Klagenfurt, Austria in 2001. He then joined the Siemens research labs, where he currently works as a senior research scientist. His primary research interests lie in network security, with a focus on mobile communications. He is Chair of the IETF ECRIT and co-chair of the IETF DIME working group. He is author/co-author of a number of RFCs.

RAINER FALK received his doctoral degree in 2000 from the Munich University of Technology. He works as a senior research scientist at Siemens Corporate Technology in the area of mobile communications systems security. His research interests include security for future mobile communication systems, mobile Internet, reconfiguration, cognitive radio, ad hoc/mesh networks, as well as for WLAN and WiMax.

JON PETERSON is senior technical industry liaison at NeuStar, Inc. He previously worked at Level(3) Communications on SIP architecture and softswitch technology. He has done extensive work on real-time communications in various standards bodies, especially the IETF. He has authored or co-authored several RFCs; he is a co-author of RFC3261, which defines SIP.

DOUGLAS C. SICKER [SM] (douglas.sicker@colorado.edu) holds B.S., M.S., and Ph.D. degrees from the University of Pittsburgh. He is an assistant professor in the Computer Science Department with a joint appointment in the Interdisciplinary Telecommunications Program at the University of Colorado at Boulder. Prior to this he worked in the Office of Engineering and Technology at the Federal Communications Commission. His research interests include cognitive radio networks, network security, and public policy.

JAMES POLK is a senior consulting engineer at Cisco Systems. He is an active contributor and writer of specifications in the IETF, TIA, NENA, and IEEE Standards organizations, emphasizing priority and emergency architectures, Geo-Location, QoS, IP-E911/112, international emergency preparedness, and government/military networking. He is the co-author of "Preferential Emergency Communications" and the author of *Session Initiation Protocol Fundamentals*.

JEFF HODGES is a senior architect of Digital Identity at NeuStar, Inc. He is working in the areas of identity solutions, distributed infrastructure, and security. He is an editor of several Liberty Alliance specifications, as well as a contributor to the OASIS SAML specifications. His earlier protocol design work was with LDAPv3 security specifications.