THE PREVALENCE OF SAML WITHIN THE EUROPEAN UNION
An Empirical Study

Bernd Zwattendorfer, Thomas Zefferer and Arne Tauber
E-Government Innovation Center (EGIZ), Inffeldgasse 16a, 8010 Graz, Austria

Keywords: Security Assertion Markup Language, SAML, Authentication, eID, STORK.

Abstract: Various European countries have set up national eID infrastructures that allow citizens to securely authenticate at e-Government or e-Banking services. In a converging European society, interoperability between national eID solutions becomes an important issue. The EU large scale pilot STORK tackles this issue and implements an interoperability layer that connects national infrastructures. The secure, reliable, and efficient exchange of identity information is thus a key feature of the STORK interoperability layer. Several protocols exist that are basically able to implement this feature. In private sector applications, SAML is frequently used for the exchange of identity and authentication data. To verify whether this protocol has also proven itself in the public e-Government domain, a survey on existing national eID solutions based on SAML has been carried out. The survey was based on a comprehensive questionnaire that was sent out to 14 Member States of the European Union. The collected results revealed that SAML is prevalently used in most national eID solutions and hence perfectly suitable to build the basis of the STORK interoperability layer.

1 INTRODUCTION

The Internet plays an important role in many aspects of our daily life. During the past years, an increasing number of security sensitive services have been mapped to the digital world and are now provided online. These services have in common that they usually require a reliable user identification and authentication process with a certain level of assurance. Reliable user identification and authentication over the Internet is no trivial task. For several reasons, common username and password based authentication schemes have to be regarded as weak (Kessler, 1997). Hence, security sensitive services usually rely on stronger two-factor authentication schemes that incorporate some kind of cryptographic hardware token (e.g. smart card, mobile phone) often supported by public key infrastructures (PKI).

Most European countries use strong authentication schemes and appropriate PKI solutions for their national eID infrastructures. National eID infrastructures provide citizens a unique electronic ID that allows for secure and reliable authentication in e-Government or e-Business processes. A comprehensive overview of existing national eID solutions is provided by (MODINIS, 2006), (IDABC, 2009) and (Siddhartha, 2008).

Identity management becomes even more complex in cross-border scenarios, in which parties from two or more different countries are involved. During an authentication process, involved parties need to exchange identity and authentication data. Depending on the authentication process, this data can be considerably complex. It is thus reasonable to have a common understanding and to appropriately structure the data to be exchanged according to a well-defined standard.

The need for a standard that facilitates the structured exchange of identity and authentication data has been recognized early. Therefore, the Security Assertion Markup Language (SAML) (Lockhart and Campbell, 2008) has been introduced in 2002. It is an XML based framework, which is heavily used by the industry and aims to facilitate the exchange of identity information (Naedele, 2003). Although SAML is quite popular, also alternative standards and frameworks exist. Especially WS-Federation has to be mentioned in this context.

In a converging European economy and society, country specific eID solutions hinder the development of eID based cross-border services. This aspect is emphasized in the Digital Agenda for Europe (European Commission, 2010) and has also been confirmed by a recent report of the (OECD, 2011) highlighting the

Zwattendorfer B., Zefferer T. and Tauber A..
THE PREVALENCE OF SAML WITHIN THE EUROPEAN UNION - An Empirical Study.
DOI: 10.5220/0003938005710576
In *Proceedings of the 8th International Conference on Web Information Systems and Technologies* (WEBIST-2012), pages 571-576
ISBN: 978-989-8565-08-2
Copyright © 2012 SCITEPRESS (Science and Technology Publications, Lda.)

need for governments and other stakeholders to create solutions for mutual recognition of national digital identity management approaches. To overcome this issue, the EU LSP STORK (Leitold and Zwattendorfer, 2010) has been started in 2008 as a 3-year project. STORK aims to establish an interoperability layer for national eID solutions in order to facilitate eID based cross-border services. The capabilities of the developed STORK modules have been demonstrated by several pilot applications (Tauber et al., 2011) (Knall et al., 2011).

The (cross-border) exchange of authentication data is one of the key features of the STORK interoperability framework. The selection of an appropriate standard or framework was therefore one of the key decisions during the conceptual design. SAML was known to be a popular industry standard. However, it was unclear if SAML was also able to meet the requirements of the public sector, particularly in a cross-border context. To get an overview of gained experiences with SAML in public sector applications, an assessment and study has been made covering all European countries participating in the STORK project. The basic findings of this study, the followed methodology, and consequences of the obtained results are discussed in this paper.

2 SAML

In the current identity and authentication landscape the Security Assertion Markup Language (SAML) plays a major role. For being able to follow the empirical research and study that has been carried out in this paper, a brief introduction to SAML and its architecture is given.

2.1 Basics

SAML constitutes an XML-based standard that has been developed by OASIS[1] and especially designed for the secure exchange of authentication and authorization data of a given *subject*. A subject in SAML terminology defines the main actor for whom identity and authentication data needs to be exchanged. Usually this term concerns a natural person but it can also be a web service or a system in general (Lockhart and Campbell, 2008).

According to the general SAML architecture, authentication or authorization data is typically exchanged between one *identity provider* and one or more *service providers*. The identity provider is usually responsible for the subject's authentication and the issuance of so-called SAML assertions for authentication requesting service providers. A SAML assertion is an XML-based security token, which assures that a certain subject has been successfully authenticated using specific means at a certain point in time and owning specific attributes if authorization is required. Service providers that receive such assertions verify it and grant or deny access to the resources that have been requested by the subject. In SAML terminology, identity providers are also called *asserting party* or *SAML authority*, service providers can also be named *relying party*.

Summarizing, the most important features of SAML are:

- Single Sign-on (SSO).
- Identity Federation.
- Web Services and other Industry Standards.

2.2 Architecture

SAML highly profits from its modular architecture. Due to this modularity, various components can be put together and appropriate solutions for different use cases can be modeled. In the nested architectural model statements (as part of assertions) specify the most detailed and profiles the highest abstract level. The following sub-sections give a brief introduction into the individual SAML components (Lockhart and Campbell, 2008).

2.2.1 Assertions

So-called SAML assertions (Lockhart and Campbell, 2008) constitute the core component of SAML. SAML assertions contain specific information about a subject, e.g. related special attributes or information indicating that the subject has been successfully authenticated. In typical scenarios, assertions are issued by an identity provider and consumed by a service provider, which uses the included information for access control decisions for the subject.

Basically, three different types of SAML assertions can be distinguished although the wrapping XML-fragment is common to all of them. A differentiation on the assertion is made based on the statements included. The SAML specification distinguishes between the following three statements: Authentication Statement, Attribute Statement, Authorization Decision Statement.

Authentication statements are usually created by an identity provider if a subject has been authenticated successfully. The statement contains informa-

[1]Organization for the Advancement of Structured Information Standards, http:// www.oasis-open.org.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

tion at what point in time and by which means the subject has authenticated and specifies the validity period of the assertion. Attribute statements wrap specific attributes belonging to the authenticated subject. Additionally, authorization statements can give information whether the subject is permitted to gain access to a certain resource or not. Although authorization statements have especially been designed for access control, in practice mainly attribute statements are used for authorization. There is also no strict regulation to use only one statement per assertion. Hence, different statements can be mixed. However, there is a limitation to one authentication statement only.

2.2.2 Protocols

SAML Protocols (Lockhart and Campbell, 2008) define the next layer in this modular architecture. They specify which assertion is transmitted between two providers or entities and also define how this transmission takes place. SAML assertions can be either pulled from or pushed by an identity provider. If pulled, the service provider requests an assertion from the identity provider. If using the push method, the identity provider sends unsolicited assertions to the service provider without any further request.

2.2.3 Bindings

SAML Bindings (Lockhart and Campbell, 2008) depict the transport protocol used for carrying the SAML protocol messages. These protocols remain untouched by the SAML specification and are just used for transportation. Typical examples for such transport protocols are HTTP or SOAP web services.

2.2.4 Profiles

SAML Profiles (Lockhart and Campbell, 2008) combine all inner parts of the modelling architecture to model certain use cases. The most popular use case or profile respectively depicts the so-called *Web Single Sign-On Profile*, which enables users SSO by using web browsers.

3 METHODOLOGY

The main objective and hypothesis, respectively, of this paper work has been the verification to what extend SAML is used in national eID systems. To approve this, an empirical study based on questionnaires, which had been distributed to specific European Union Member States was chosen as adequate mean.

Before starting any empirical analysis, an agreement on the underlying sample must be obtained. To achieve the best results, actually no sample should be taken but in contrast the whole basic population should be explored. However, in most cases such explorations are not very economic and are very time intensive due to a huge population basis. Therefore, usually a representative sample out of the basic population is extracted which saves costs and time but certainly is not as accurate as a census and hence gives probabilistic results only.

3.1 Sample Selection

Basically, sample extraction does not define a trivial task and must be chosen carefully. According to (Kaya and Himme, 2009), sample extraction must be carried out by applying the following steps:

1. Determination of the basic population.
2. Determination of the basic selection.
3. Definition of the sample.
4. Definition of the selection process.
5. Carrying out the selection.

Mapping these steps to our analysis, the general basic population would be defined by all countries over the world (Step 1). However, since probably not all countries have national eID solutions deployed, we limit the basic population to countries that have already national eID infrastructures in place or are planning to do so in the next years (Step 2). Because for the whole world this would also not be a feasible and trivial approach, the basic selection is narrowed down to European countries as a sample set (Step 3). Out of this sample set, the actual sample has to be chosen and analyzed. In theory, a sample should be selected completely random. However, in practice this usually can never be achieved and one's sights must be lowered. Hence, the underlying sample of an empirical study can be selected on various aspects. For that, several approaches exist, e.g. based on random or non-random selection. For our analysis, an approach based on convenience sampling[2] was chosen (Step 4). This approach has been chosen due to practical aspects. Since this work was carried out within the STORK project, all countries participating in this LSP were chosen for building a representative sample[3](Step 5). In fact, this final sample consisted of 14 countries, involving Austria, Belgium, Estonia, France, Germany, Italy, Luxembourg, Netherlands,

[2]Convenience sampling defines a non-random selection method where elements of the sample are readily available or nearby. (Black, 2010).

Portugal, Slovenia, Spain, Sweden, United Kingdom, and Iceland (as the only non EU country).

3.2 Structure of the Questionnaire

After selecting the appropriate sample and before being able to do any analysis, information and data must be gathered. This section briefly describes the structure of the questionnaire that had been sent out to the STORK partners for information gathering. Actually, the questionnaire can be divided into two parts. The first part containing administrative questions requesting contact details in case any questions arose as well as motivating the questionnaire, the second part asking specific questions to SAML implementations and specifications. Besides contact details, the administrative part concluded with a short introduction into SAML. This introduction overviewed the architecture and the most important parts of SAML giving a concise definition on the various SAML terms used in the SAML specification to gain a common understanding amongst all partners.

The actual questionnaire is built upon 17 questions, involving both closed (offering the recipient predefined questions) and open (possibility to enter free text) questions. Basically, both general questions as well as more technically detailed questions regarding SAML were asked. For instance, general questions concerned the national eID system and its use of SAML, or practical details or issues during SAML adoption. In addition, detailed questions specifically focused on the SAML specification, e.g. asking which SAML versions are deployed or - regarding to the SAML architecture - which protocols, bindings, or profiles are in use. The order of the detailed questions were aligned according to the natural nesting of the SAML architectural components, hence a top-down approach was followed. According to this approach, questions regarding to the use cases (SAML profiles) were asked first. Referring to section 2, SAML profiles constitute the outermost part of the architectural model. Following the intuitive structure of this model, questions concerning bindings, protocols and assertions as the innermost part were stated next. In the following, details on the assertion structure or whether parts are signed or encrypted were questioned. At the end, open questions rounded up the questionnaire in order to give the participants the possibility to provide hyperlinks for further information beyond the scope of this questionnaire.

[3]Please note that this work has already been carried out during the first project phase of STORK when no Member State enlargement had been defined. An enlargement of additional five Member States was negotiated in the second year of the project phase only, hence those five countries are not part of our analysis.

4 RESULTS

This section reports on the results extracted and synthesized from the returned questionnaires. The complete analysis and evaluation of the questionnaires is based on common descriptive statistical methods.

All of the sent out questionnaires were returned, thus we could rely on a response rate of 100% for our analysis[4]. Although a sample of 14 countries does not represent all European Union Member States, at least a very good approximation can be given for the complete EU. However, the derived findings could deliver a complete picture of the use of SAML in the STORK Member States and built a fundamental basis for the standards chosen in the STORK specification.

In general, 11 out of 14 countries are either using SAML (Austria, Belgium, France, Portugal, Spain, UK, and Iceland) or are planning to do so (Italy, the Netherlands, Slovenia, and Sweden) There, SAML is used in broad context on domestic or regional level in the national eID systems. This finding corresponds to a quota of about 80% of SAML using countries and can be seen as an evidence for the prevalence of SAML in the European Union. The only countries of the interviewed sample that do not rely on SAML are Estonia, Germany and Luxembourg.

The first country deploying SAML in its national eID infrastructure has been Austria which has started the adoption of SAML (version 1.0) already in 2002. Belgium and France were next having started in 2003 and 2004 respectively (version 1.0 and 1.1). A couple of countries have been deploying SAML in 2008, mostly based on the SAML version 2.0. Most countries rely on national guidelines for their SAML implementation. However, there exists also a number of countries which have deployed SAML on regional level only (e.g. Catalonia in Spain) and hence rely on sector/application specific policies.

Besides general questions on the use and the adoption of SAML, more detailed questions affecting the SAML specifications were asked to the participating countries. The aim of those questions was to get insight which out of the box SAML profiles, bindings, or protocols are in place in the Member States. The dominant version of SAML (planned or already deployed) is the current version 2.0 used by app. 61% of the asked countries. 23% use SAML version 1.1 and 15% still version 1.0.

[4]Not all questionnaires were completely filled out though as some countries do not rely on SAML.

Concerning the different types of used SAML profiles we limit our evaluation to the profiles offered by SAML version 2.0, as only two profiles are standardized in the versions 1.0 and 1.1. SAML 2.0 provides 13 standardized profiles in total where only seven are used over all Member States. The most popular profile constitutes the *Web Browser SSO Profile* which is implemented by six countries. The country, which bases its eID implementation on the most profiles, is Belgium with eight deployed planned profiles (all SAML versions). All other countries use or plan to use between one and four standardized profiles. The average profiles used per country are 2.66.

In comparison, the average number of SAML bindings used per country is 2.4. The number of bindings per country is nearly equally distributed, reaching from one to four bindings per country, having again Belgium leading together with Slovenia. The most common binding is the *HTTP Post Binding* with seven mentions, followed by the *HTTP Redirect Binding* with five, the *SAML SOAP Binding* with four, and the *HTTP Artifact Binding* with three. All those bindings refer to the SAML 2.0 specifications. As only few countries still use earlier SAML versions we skip again a detailed analysis on them.

Additionally, all countries rely on SAML protocols for the transport of SAML assertions. Again, Belgium uses the most predefined SAML protocols, namely four. All other countries rely on between one and three out of the box protocols. The most popular protocol is the *Assertion Query and Request Protocol* in version 2.0 succeeded by the *Authentication Request Protocol*, the *Artifact Resolution Protocol* and the *Single Logout Protocol*. Taking profiles, bindings, and protocols together, all countries just rely on predefined SAML components and did not implement any country specific solution for SAML adoption.

Digging a little bit deeper into the modular SAML architecture, after analyzing profiles, bindings, and protocols we evaluated the use of SAML assertions and its statements. Nearly all responding countries (87.5%) have an authentication statement included in their assertion (except Austria). Thereby, for authentication different authentication methods are invoked in all countries, reaching from simple username/password mechanisms to more secure and high sophisticated smart card and PKI based solutions. In contrast to that, all participating countries include at least one attribute statement in their assertion. Belgium, Italy and the UK even include more than one. In our questionnaire, we also asked which kind of and how many attributes are wrapped in an attribute statement. Most countries use SAML assertions only for identification of natural persons (in Austria, France, and Spain legal persons can also be authenticated in the national eID infrastructure), hence the most common attributes are a unique/sectoral identifier, first name, last name, and date of birth. Regarding the maximum number of attributes (mandatory or optional) within an attribute assertion, UK's assertion can take up to nine attributes, Autria's seven, and Italy's six. The average number of attributes in one assertion is 4.75.

Concerning section 2, besides authentication and attribute statements also authorization statements can be incorporated in a SAML assertion. However, this feature is rarely used amongst the participating countries. Only Belgium is planning to regulate access control using this SAML possibility.

All Member States take care about the security of the identification and authentication data transmitted. In fact, all countries sign their SAML assertion using XML-DSig as signature syntax and processing algorithm. However, in contrast no country actually encrypts the assertion. Also no encryption algorithm is used for encrypting single attributes. To additionally improve security, SAML assertions have only a certain period of validity. However, this validity period greatly varies between the Member States, reaching from 5min (Belgium, France, and Spain) to a couple of hours (Iceland and UK) or even several days (Italy).

5 CONCLUSIONS

The aim of the work carried out in this paper was to prove if SAML is also a dominant standard for exchanging identification and authentication data in national eID concepts across the European Union. To verify this, a questionnaire containing general questions to the national eID infrastructure and additionally more detailed questions regarding the structure of SAML components (profiles, bindings, protocols, and assertions) was sent out to all 14 partners of the STORK project in its early phase. All partners representing a Member State replied to this questionnaire, hence the evaluation of those questionnaires is based on a response rate of 100%. Based on these results and findings, SAML can be seen as an important standard in the field of eID across Europe. The prevalence of SAML amongst the interviewed Member States led also STORK to set up its interoperability framework for cross-border identification and authentication on SAML 2.0. According to the findings resulting from the described empirical study, the *Web Browser SSO Profile* and the *HTTP Post Binding* were chosen as basic SAML components as they are used most frequently. Although the Assertion Query and Request

Protocol was stated most in the returned questionnaires, the protocol of choice in STORK was the *Authentication Request Protocol*. The reason was because the *Authentication Request Protocol* especially focuses on the transfer of secure authentication data. The SAML assertions transmitted within STORK contain one authentication statement and one attribute statement, giving the possibility of including a large number of various personal attributes. Since STORK focuses on authentication of natural persons only, support for legal persons is not given yet[5]. As sensitive data is transmitted within STORK, the SAML messages are digitally signed but not encrypted according to the current status of the participating Member States. However, to further improve security the use of the so-called SAML *Holder-of-Key Binding* (Lockhart and Hardjono, 2010) has been specified.

Although SAML plays an important role in European eID related applications and authentication processes, still some gaps could be identified in this work. First, although OASIS specified a high number of profiles, bindings, and protocols, only a small number of those SAML components are really used in production. Hence, a couple of ”exotic” specifications only exist on paper and not in real environments. Second, SAML only specifies protocols for either attribute transfer or authentication. There does not exist a SAML protocol handling both processes in one request/response interaction yet, hence currently for supporting such a use case the implementation of two protocols is required. In STORK, this gap was overcome by using the SAML extension mechanisms in the protocols and introducing new XML elements in the *Authentication Request Protocol* (Alcalde-Morano et al., 2011). By doing that, the exchange of authentication data AND attribute data for a specific subject within one request/response message interaction becomes possible. This gap and its according solution are currently reported to OASIS for standardization discussions (Reible, 2011).

Summarizing, the results of our study have shown that SAML defines a key component in the European Union eID landscape when data exchange is required. Nevertheless, although the SAML specification has been amended and improved over a couple of years, still lacking capabilities can be found and identified.

[5]One main objective of the STORK successor project STORK-2 will be the cross-border recognition of legal identities (European Commission, 2011).

REFERENCES

Alcalde-Morano, J., Hernndez-Ardieta, J., Johnston, A., Martinez, D., Zwattendorfer, B., and Stern, M. (2011). *STORK D5.8.3b Interface Specification*. STORK Consortium.

Black, K. (2010). *Business Statistics for Contemporary Decision Making*. Wiley, 6th edition.

European Commission (2010). *A Digital Agenda for Europe, COM (2010) 245*. European Commission (EC).

European Commission (2011). *ICT Policy Support Programme (PSP) Work Programme 2011*. European Commission (EC).

IDABC (2009). *eID Interoperability for PEGS: Update of Country Profiles*. IDABC.

Kaya, M. and Himme, A. (2009). Möglichkeiten der Stichprobenbildung. In *Methodik der empirischen Forschung*. Gabler.

Kessler, G. C. (1997). Passwords - Strengths and Weaknesses. In *Internet and Networking Security*. Auerbach.

Knall, T., Tauber, A., Zefferer, T., Zwattendorfer, B., Axfjord, A., and Bjarnason, H. (2011). Secure and Privacy-preserving Cross-border Authentication: the STORK Pilot 'SaferChat'. In *Proceedings of the Conference on Electronic Government and the Information Systems Perspective*. Springer.

Leitold, H. and Zwattendorfer, B. (2010). STORK: Architecture, Implementation and Pilots. In *ISSE 2010 Securing Electronic Business Processes: Highlights of the Information Security Solutions Europe 2010 Conference*, pages 131–142.

Lockhart, H. and Campbell, B. (2008). *Security Assertion Markup Language (SAML) V2.0 Technical Overview*. OASIS Committee Draft 02.

Lockhart, H. and Hardjono, D. (2010). *SAML V2.0 Holder-of-Key Web Browser SSO Profile Version 1.0*. OASIS Committee Specification 02.

MODINIS (2006). *The Status of Identity Management in European eGovernment initiatives*. MODINIS.

Naedele, N. (2003). Standards for XML and Web Services Security. *IEEE Computer*, 36(4):96–98.

OECD (2011). *Digital Identity Management for Natural Persons: Enabling Innovation and Trust in the Internet Economy - Guidance for Government Policy Makers*. OECD Publishing.

Reible, V. (2011). STORK Results: What's in it for Industry. Technical report, STORK.

Siddhartha, A. (2008). National e-ID card schemes: A European overview. *Information Security Technical Report*, 13(2):46–53.

Tauber, A., Zwattendorfer, B., and Zefferer, T. (2011). STORK: Pilot 4 Towards Cross-border Electronic Delivery. In *Electronic Government and Electronic Participation - Joint Proceedings of Ongoing Research and Projects of IFIP EGOV and ePart 2011*. Springer.