# Digital Identity as a Basis for Internet Security Infrastructure

Ing. Radovan Semančík
*Business Global Systems, a.s.*
*semancik@bgs.sk*

## Abstract

*The proliferation of web services and distributed, loose coupled systems brings the need for a complex security infrastructure. Digital identity technologies are introduced as a potential good foundation for next generation security infrastructure. This paper provides overview of extended traditional security systems such as unified user management. The analysis of PKI mechanisms and their applications is provided and the limitations and problems are shortly described. Digital identity systems are described as a potential solution to these problems.*

## 1. Introduction

Traditional security systems are reaching the upper limit of their abilities. The ad-hoc user management techniques are becoming infeasible in today's distributed computing environments. The typical enterprise information system consists of many applications that feature incompatible user management mechanisms. The situation gets even worse when considering inter-enterprise communication, commonly used in the e-commerce scenarios. Similar situation can be observed in the public Internet environment, where a typical user posses several user accounts in different systems with different authentication credentials. It could be anticipated that the situation will get even more complex in the near future, when the number of systems on the Internet will grow and the business-to-business communication will be more frequent and sensitive.

## 2. Unified User Management

Unification of user management in different applications is one of the major concerns in today's enterprise information systems. Common user management systems use a central repository of user information based on directory services. The *de facto* standard in this area is LDAP [1] directory access protocol. Applications that support LDAP protocol could directly access the repository, but it is not sure that the directory structure and schemas used by the repository and required by the application are compatible. To address this problem several systems were proposed, ranging from metadirectories to user provisioning systems. These systems are based on replication and synchronization of various data sources. As also noted by the work in other areas [2], this approach is not suitable for dynamic data structures and has a scalability limits.

Another concern is user authentication and credentials synchronization. The directory services were not designed as an authentication servers and even though they are commonly used as such, it is not a systematic approach. Directories are not suitable to store dynamic information such as user session credentials nor the LDAP protocol was designed as an authentication protocol. It is clear that standalone authentication and session management service is needed to supplement directories in the user management system.

## 3. Public Key Infrastructure

Public Key Infrastructure (PKI) is an asymetric cryptography-based mechanism, that provides users with digital certificates asserting authenticity of their public keys. The public key certificates (PKC) are commonly expressed in X.509 [3] format, which became industry standard. Digital certificates are issued by certificate authorities, that act as a trusted third parties in PKI infrastructure. Every user in PKI posses an asymetric key pair and one or more certificates for his public key. User then authenticates to the server by presenting a public key certificate and a proof of possession of corresponding private key.

The PKI mechanisms could be found in many important Internet security systems, but especially noteworthy are the TLS (SSL) protocol, S/MIME mail security system and IPsec IP security extension. Several countries also included in their legislation an electronic signature system based on PKI mechanisms. It is clear, that X.509-based PKI is an important part of many security systems and infrastructures.

Public key certificates used in PKI have generally long lifetimes and therefore certificate attributes has a static character. Practically, user information in public key certificates is limited to the certificate holder identifier and the burden of obtaining user's attributes is left to the relying application. Original X.509 design assumed global X.500 [4] directory service and the X.509 subject and issuer identifiers are in X.500 distinguished name format. It was assumed that the name would have global meaning and that there will be a naming authority that would assign unique names to subjects. This approach was not followed in practice and many existing PKI

applications regard subject and issuer identifiers to be just a set of unstructured attributes. The content of these attributes is a major privacy concern. Information included in the subject identifier is accessible to any site, to which user presents her or his certificate and no access control is possible. Therefore, this information should be kept minimal from the privacy point of view. Since there is no direct and standard method to obtain more information about particular user, PKI implementers tend to include as much information as possible in the public key certificate itself.

To address some of these issues, X.509 recommendation proposed the use of attribute certificates. These relatively short-term certificates bind subject identifier with a value of an attribute. Such attribute certificates can form a privilege management infrastructure (PMI), that can work together with public key certificates in the PKI. The PMI is undoubtedly an enhancement to the basic PKI, but it still features several major problems. First of all, the burden of attribute certificate processing is on the end user's system. The user's application must implement the appropriate privacy policy for the use of attribute certificates and it must be implemented consistently in each application. Another problem is related to the dynamic attributes. The attribute certificates are not suitable for dynamic, fast changing user attributes like location or user presence. More that that, user's identity could be associated with specific attributes, that are in fact services. One such example could be a calendar service, that keeps and manages user's personal schedule. The PMI was not designed and is not suitable for conveying such dynamic and complex attributes.

The PKI mechanisms place considerable amount of processing on the user side. User agent presents public key certificate, proof of private key possession, attribute certificates, enforces policy constraints and maintains security of the entire process. Although this scenario may be appropriate in some circumstances, it could be inconvenient and even dangerous for typical Internet user. Considering the low security of common information systems [5], that is especially apparent in personal computers and workstations, security processing on end user workstation may not be secure enough. Even the use of cryptographic smart cards does not sufficiently improve the situation. The key material could be stored safely on the smart card, but once the legitimate user is authenticated to the smart card, the attacker has the same access to the smart card functions as the legitimate user has.

Similar analysis of the PKI limitations and drawbacks can be found in the literature [6][7]. The PKI mechanisms are undoubtedly important part of Internet security infrastructure, but PKI is by no means a complete system. The public key certificates can be efficiently used for a network node authentication or a strong user authentication, and could provide a service layer on which a more general digital identity systems are built.

## 4. Digital Identity

Increasing requirements on the user mobility, device independence and system integration does have considerable impact on the architecture of security infrastructure. Inter-organization interactions require user profiles to be always available and accessible in a standard and secure manner. User profile could consist of many types of attributes, ranging from primitive data types to complete dynamic services.

Considering the enterprise environment, it is desirable to store user profiles on a central server system. This system should enforce attribute access policy and therefore it could be beneficial to combine it with user authentication and authorization services. Such a system is called i*dentity server* and it provides a basic building block of digital identity infrastructure. The identity server authenticates users by any authentication method and sets up user sessions. This session can be used to transfer the authentication status to other systems, allowing effective single sign-on. The authentication status is commonly conveyed in the form of *SAML assertion*. The Security Assertion Markup Language (SAML) [8] is an OASIS specification of language, that is used to express security statements and transport them to other systems. The SAML security assertions are in fact short-living digital certificates expressed in XML, that assert subject's authentication status, attribute possession or authorization decision. The SAML assertions are used in several digital identity systems, for example Shibboleth [9] and Liberty [10].

Considering the simple enterprise scenario (Figure 1), user authenticates to the identity provider (identity server) by any available authentication protocol and the session is established between the user and the identity provider. When user accesses the content service, authentication status is transferred to the content provider (content service) as a part of the request. The transfer is accomplished by SAML authentication assertion, that

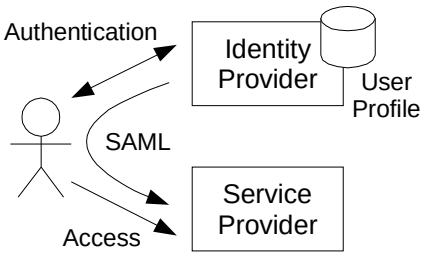


**Figure 1. Simple SAML scenario**

is issued by the identity provider. Service provider then establishes an authenticated session with the user and no further communication with identity server is required. Communication with the identity server could be desirable in a case of a user logout, changed security level or attribute exchange.

One of the most interesting digital identity system is the Liberty Alliance Project [10]. The first phase of Liberty Alliance specifications [11] is focused on federated simplified sign-on mechanism based on SAML. It supports standard unmodified web clients as well as "Liberty-enabled" clients. For more detailed analysis of Liberty Alliance phase 1 specifications see [12]. The drafts of the Liberty phase 2 specifications [13] add several features to the previous specifications of which the discovery service and the identity profile service are the most interesting additions. The profile service would allow to store user profiles centrally on the identity server (as can be seen on Figure 1) and to distribute profile data in a controlled manner to authorized content providers.

The full capabilities of digital identity systems cannot be seen in the simple enterprise example. Even the traditional techniques (e.g. directory service with kerberos authentication) could be feasible in the enterprise environment, where all involved systems are under single administration control. Although, the requirements are quite different in the global Internet environment.

The primary difference between enterprise environment and the Internet is the need to communicate across organizational boundaries. Figure 2 illustrates such a case, where user in an enterprise environment needs to use an application (service) deployed in different organization (service provider). In many cases, the service provider does not need to know the real identity of the user. The knowledge of specific attribute (e.g. membership in a group) could be sufficient to make an authorization decision. For that reason, different pseudonym in assigned to the user for each participating

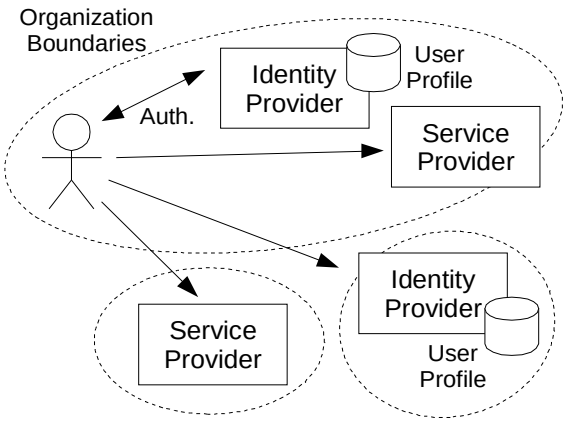


**Figure 2. Digital identity on the Internet**

content provider. Content provider has a handle, that can be used to obtain user attributes, but that does not leak any user information. As the user handle is different for each content provider, they cannot collude to trace and correlate user actions. Note that identity provider can gather some general information on user actions, and therefore must be trusted not to misuse that information. Since the identity provider stores and manages user profile, great amount of trust is assumed.

Different identity providers may have different trust levels, and store different parts of user profile. Figure 2 illustrates a case, when a user profile is split between two identity providers. For example, general profile could be stored by user's employer identity provider, but the user's medical records should be stored at a different identity provider system with higher trust level. This approach lowers the risk of the central profile repository compromise, that may have catastrophic consequences. Note that the situation depicted on the Figure 2 improves the overall situation only slightly, because the user authenticates only to the employer's identity provider and accesses the other identity provider directly. The employer identity provider can impersonate the user and misuse his or her data even if they are stored at a different system. To overcome this problem, identity providers may require additional authentication to access the more sensitive parts of the user profile.

The disadvantage of the digital identity technologies is the need for an ultimate trust in the identity providers. Identity provider has an access to all user attributes and may observe some of the user's actions. This situation can be mitigated by splitting user profile to several identity providers, although the problem is not yet fully resolved. Also the non-repudiation features of digital identity systems are very limited and the implementers are left with the traditional auditing mechanisms. Digital identity systems are also pure on-line systems and all interaction should be done in real-time.

The digital identity technologies are undoubtedly a great improvement over traditional security systems, especially in the area of distributed web applications. In our opinion, the most important effect of these technologies will be seen in a world of web services. The computer-to-computer interactions based on the web service mechanisms may play an important role in the area of e-business applications, but the key point that is still missing is the security infrastructure. As show in the section 3, the PKI techniques alone are not a sufficient solution and a mechanism is needed to provide more complex and featureful infrastructure. We hope, that digital identity could be a base for such infrastructure.

## 5. Conclusion

Traditional security mechanisms are reaching their limits in scalability, flexibility, manageability and cannot be easily transformed into a more complex infrastructures. Attempts to build a security infrastructure using public key cryptography such as PKI may provide a good foundation, but is not a complete solutions by itself. The digital identity technologies are introduced as a potential base of such infrastructure. These technologies provide many features applicable

to the contemporary Internet environment as well as promising future perspective. Digital identity systems used in practice today provide only the minimal features, especially cross-domain simplified sign-on functionality and minimalistic user profile access. Future work is definitely needed in both research and development fields to utilize the full potential of these technologies.


[1] Yeong, Y., Howes, T., Kille, S.: *Lightweighted Directory Access Protocol*, RFC 1777, 1995

[2] Czajkowski, K., Fitzgerald, S., Foster, I., Kesselman, C.: *Grid Information Services for Distributed Resource Sharing*, Proceedings of the Tenth IEEE International Symposium on High-Performance Distributed Computing (HPDC-10), IEEE Press, August 2001.

[3] *Information technology - Open Systems Interconection - The Directory: Public-key and attribute certificate frameworks*, ITU-T Recommendation X.509, 2001

[4] *Information technology - Open Systems Interconection - The Directory*, ITU-T Recommendation X.500, 2001

[5] Brunnstein, K.: *About Inherent Insecurity of Contemporary ICT Systems, and about Future Safety and Security Requirements*, Proceedings of the Information Security Summit, Prague, May 2002.

[6] Davis, D.: *Compliance Defects in Public-Key Cryptography*, Proceedings of the 6th USENIX UNIX Security Symposium, July 1996.

[7] Ellison, C., Scheier, B.: *Ten risks of PKI*, Computer Security Journal, 2000, http://www.counterpane.com/pki-risks.html

[8] Hallam-Baker, P., Maler, E.: Assertions and Protocol for the OASIS Security Assertion Markup Language, OASIS Standard, 2002

[9] Erdos, M., Cantor, S.: *Shibboleth Architecture DRAFT v05*, http://shibboleth.internet2.edu/docs/draft-internet2-shibboleth-arch-v05.pdf

[10] Liberty Alliance Project, http://projectliberty.org/

[11] Hodges, J.: *Liberty Architecture Overview*, Liberty Alliance Project Specification, 2002

[12] Semančík, R.: *Internet Applications Security*, Written part of Ph.D. exam, 2002

[13] Liberty Alliance Project, Phase 2 draft specifications, http://projectliberty.org/specs/main.html