ETSI GS INS 010 V1.1.1 (2014-03)

Identity and access management for Networks and Services; Requirements of a global distributed discovery mechanism of identifiers, providers and capabilities

Disclaimer

This document has been produced and approved by the Identity and access management for Networks and Services ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference
DGS/INS-0010

Keywords
access, control, ID, management, network, service

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://ipr.etsi.org).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Group Specification (GS) has been produced by ETSI Industry Specification Group (ISG) Identity and access management for Networks and Services (INS).

Introduction

The analysis presented in GS INS 006 [i.1] concludes that there is a need for the development of a global discovery mechanism of identifiers, providers and capabilities. This work item investigates the requirements of such a mechanism and examines if any existing systems or mechanism meet these requirements and can - fully or partially - support the design of its architecture.

The present document is based on the principle that the global discovery mechanism provides only discovery of information that is somehow related to an identity, and is not by any means involved in any other kind of identity management procedures like information exchange, trust between service providers etc. It should be noted that it is not among the goals of the present document to enforce this principle towards the development of such a mechanism.

1 Scope

The scope of the present document is the identification of the requirements to develop a global distributed discovery of identifiers, providers and capabilities.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the reference document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are necessary for the application of the present document.

Not applicable.

2.2 Informative references

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ETSI GS INS 006: "Identity and access management for Networks and Services; Study to Identify the need for a Global, Distributed Discovery Mechanism".

NOTE: See http://www.etsi.org/deliver/etsi_gs/INS/001_099/006/01.01.01_60/gs_ins006v010101p.pdf.

[i.2] SAML Specifications.

NOTE: See http://saml.xml.org/saml-specifications.

[i.3] International DOI® Foundation, Digital Object Identifier.

NOTE: See http://www.doi.org/.

[i.4] Liberty Alliance Project, Project liberty.

NOTE: See http://www.projectliberty.org.

[i.5] Internet2 Middleware Initiative, Shibboleth®.

NOTE: See http://shibboleth.internet2.edu.

[i.6] Eduserv, OpenAthens.

NOTE: See http://www.openathens.net.

[i.7] InCommon Discovery Service.

NOTE: See https://spaces.internet2.edu/display/InCFederation/Discovery+Service.

[i.8] Handle System®.

NOTE: See http://www.handle.net.

[i.9] IF-MAP.

NOTE: See http://www.if-map.com.

[i.10] Jon Crowcroft, Steven Hand, Richard Mortier, Timothy Roscoe, Andrew Warfield, "Plutarch: An Argument for Network Pluralism", ACM SIGCOMM, 2003.

[i.11] Ion Stoica, Robert Morris, David Liben-Nowell, David R. Karger, M. Frans Kaashoek, Frank Dabek, Hari Balakrishnan, Chord: "A Scalable Peer-to-peer Lookup Service for Internet Applications", ACM SIGCOMM, 2001.

[i.12] Petar Maymounkov, David Mazières, Kademlia: "A Peer-to-Peer Information System Based on the XOR Metric", IPTPS, 2002.

[i.13] Sylvia Ratnasamy, Paul Francis, Mark Handley, Richard Karp, Scott Shenker: "A Scalable Content-Addressable Network", ACM SIGCOMM, 2001.

3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

CAN	Content-Addressable Network
DDNS	Dynamic DNS
DDoS	Distributed Denial-of-Service
DHT	Distributed Hash Table
DNS	Domain Name System
DNSSEC	DNS Security Extension
DOI®	Digital Object Identifier
DS	Discovery Service
IdM	Identity Management
ID-WSF	Identity Web Services Framework
IF-MAP	Interface for Metadata Access Points
IP	Internet Protocol
ISG	Industry Specification Group
SAML	Security Assertion Markup Language
STORK	Secure IdenTity AcrOss BoRders LinKed
VID	Virtual Identity
XRDS	EXtensible Resource Descriptor Sequence
XRI	EXtensible Resource Identifier

4 Requirements of a Global Distributed Discovery mechanism of identifiers, providers and capabilities

This clause presents the requirements of a Discovery System (DS) capable of supporting a Global Distributed Discovery mechanism of identifiers, providers and capabilities as identified by the ETSI INS ISG group.

1) Independent

1) It is recommended that the DS remains unaffected by the peculiarities of the various IdM operations between an identity producer and an identity consumer.

2) It is recommended that the DS exists as an independent entity which only provides discovery services and not as part of a specific identity management system or infrastructure.

3) The DS' components ought to be independent of each other, so the behavior of a component does not affect the others.

4) It is recommended that the DS remains unaffected by any commercial or other interests that may try to affect its functionality for business or any other non-technical reasons.

5) It is recommended that the ownership of this mechanism is assigned to a global non-profit organization. This organization ought to only supervise the smooth operation of the DS. It is also recommended that the owner organization does not interfere with the DS' functionality nor has access or rights on any of the information registered in it.

2) Distributed

1) It is recommended multiple entities host and run the components required to provide the discovery service. A global entity (non-profit organization) may be present to play a supervisory role.

2) It is required to ensure that there is no single point of failure in the architecture of the DS.

3) It is necessary that the architecture of the DS is scalable and not affected in any way (e.g. functionality, privacy, trust, security, etc.) by its size or the amount of stored information.

4) It is recommended that the DS supports dynamic join and leave of DS components (storage hosts), managed by different domains, holding a portion of the global space and managing the operations related to it (for performance reasons).

3) Global

1) The services of the DS are not to be confined within a specific context. It is necessary to ensure that the DS's services are accessible from everywhere and everyone (domains, federations, countries, networks, etc).

2) It is required for the DS to avoid any conflict with local laws and regulations. Moreover it ought to be able to adjust to diverse regulations that apply across various locations.

3) It is recommended that external entities (users, providers etc) have the ability to freely choose the place (network location within the DS) where they wish to store/register their private data. Also these entities ought to have the ability to easily migrate their private data to other places.

4) The DS ought to provide the means to associate identity related data irrespectively of the location that these data may reside.

5) It is required that the DS architecture ensures that the discovery process is able locate all the information that is relevant to a request. (This does not mean that all this information will be included in the response).

6) The DS ought to behave in a consistent way and always be in position to provide the best available response for a given request.

7) It is recommended that the DS has the ability to provide information about a user (or any other entity) regardless his intervention or accessibility. It is required that this action reflects user's policies and preferences (e.g. the user has previously authorized the DS to act on his behalf).

8) The response time of a request is not critical but it should be reasonable.

4) Privacy Enabled

1) It is required that the DS protects the privacy of the stored information against any kind of internal or external security issues e.g. misconfiguration, attacks from internal or external malicious parties, etc.

2) It is required that the DS always provides the necessary means for an external entity (users, providers, etc) to control its registered information e.g. insert, manage and withdraw information from the infrastructure, apply policies, allow or restrict access to certain information or parties etc.

3) Conflicting policies and rules ought to be acknowledged by the DS and presented to the appropriate entity (owner user, provider etc).

4) It is recommended that the DS supports various options when handling an incoming discovery request e.g. act independently, request the owner's consent, etc.

5) It is recommended that discovery responses provide the minimum of information required to handle a request (minimal disclosure).

6) It is required that discovery operations (requests and responses) are not traceable by unauthorized third parties, so the identity of the entities which perform the operations cannot be guessed.

7) It is recommended that external entities (users, providers etc) have the ability to freely choose the place (network location within the DS) where they wish to store/register their private data. They also ought to have the ability to easily migrate their private data to other places *(This requirement is also relevant with the section "Global").* It is necessary that the DS keeps track of all the requests and actions performed on specific information and has the ability to present this activity to the owner of the information.

5) Secure

1) An authentication mechanism is required to control the access of authorized entities e.g. users.

2) It is recommended that the information stored in the DS is encrypted, easily invalidated and recoverable with minimum or no cost or harm.

3) A security framework is required to protect the DS from any kind of external attacks, e.g. DDoS.

4) It is required that the DS provides protection against data traffic monitoring that could lead to any kind of malicious or unauthorized actions e.g. construction of user behavior models.

5) It is recommended that the DS is able to support the establishment of secure connections/channels for any kind of communication with internal or external entities (users, providers etc). Such communication may involve actions like discovery requests, data registration, user account management etc.

6) Trusted

1) It is recommended that the DS is able to prove its trustworthiness to all external entities that interact with it.

2) It is recommended that the DS has the necessary means to monitor its components behaviour and take actions to prevent or stop inappropriate actions.

3) It is recommended that the DS does not interfere with trust relations between external entities that use its functionality.

4) An access control framework is required to provide the means to validate external parties which wish to register information in the DS. It is recommended that the DS is able to accept and evaluate any kind of input from any available trust frameworks or other sources in order to validate external entities that want to register information in its infrastructure.

5) It is recommended that the DS does **not** validate external entities which issue discovery requests to the DS. Exceptions may apply for external parties that have been reported to act maliciously.

6) It is recommended that the DS is able to notify the owners of the information requested in a discovery request and ask permission to include their data in the response.

7) Open and Extensible

1) It is recommended that the DS has the ability to support, adopt and provide additional functionality in the form of external frameworks.

2) It is recommended that external entities (users, providers etc) have the ability to create and manage multiple accounts in the DS.

8) Interoperable

1) It is required that the DS' services are not restricted to only specific technologies, protocols and formats etc. It is necessary that the DS has the ability to handle any kind of identity information, and all types of existing formats.

2) It is recommended that the DS does not introduce new types of identifiers and formats that need to be enforced in existing networks and services procedures. The creation of new formats that do not affect the operation of existing networks (e.g. for DS' internal purposes) is acceptable.

3) It is recommended that globally accepted semantics are adopted to describe hosted data, requests and responses in a unified manner (e.g. SAML [i.2], DOI® [i.3], etc).

4) It is recommended that among the DS's capabilities is the ability to transform external entities' actions and choices to globally accepted semantics - information.

5) It is required that the DS defines and follows a clear process for the insertion of new globally accepted elements (e.g. description of a new identity attribute).

9) User friendly

1) It is recommended that the use of the DS should be intuitive to the user.

2) It is recommended that the DS search mechanism supports finding elements from their semantic description, using various and different ontologies (and vocabularies) to regulate them.

5 Existing discovery systems & mechanisms vs the Requirements of a global DS

This clause will present the most relevant discovery systems and mechanisms and evaluate whether they can support the development a global discovery mechanism of identifiers, providers and capabilities.

5.1 Federated Identity Management Frameworks

The notion "federation" can be interpreted in many ways. Any kind of collaboration between two or more parties can be characterized as a federation. In the present document the word federation describes a group of domains, which agree on a specific set of rules and share/exchange information within a closed Circle of Trust. Based on this interpretation we will analyse the federated IdM systems and their ability to support the development of a Global Distributed Discovery mechanism of identifiers, providers and capabilities.

The majority of the federated IdM systems design their own proprietary Discovery Systems (DS) based on the procedures and services of the federation. Such DS are the ID-WSF Discovery Service proposed in Liberty Alliance [i.4] project, the "Where Are You From (WAYF)" service from the early versions of Shibboleth [i.5] and Athens [i.6], the InCommon Discovery Service [i.7] from Shibboleth etc. (More information about these systems are available on [i.1]). Despite the fact that these DS have large diversities, there are foundational architectural similarities and principles based on which all of them are designed. This clause will examine these DSs as one category and evaluate their capability to support a Global Distributed Discovery mechanism of identifiers, providers and capabilities.

Advantages

1) Use of well established protocols

Federated IdM systems and their DSs generally use widely accepted protocols and formats to support communication between its participants. This offers easy adaptation of new components, features and entities, etc.

2) User friendly

The majority of the procedures needed to operate a DS in a federation are not presented to the end users and are handled by the providers that form the federation. Such procedures are for example the association of identities and identity related data issued by different providers, registration of information in the DS, policy enforcement etc. Hiding these complex operations from the users, allow federated IdM to construct more user-friendly interfaces, which require the minimum of users' intervention to fully operate.

3) High levels of Security, Privacy and Trust management **in small scale systems**

Federated IdM systems and their DSs offer very good levels of security. Minimum user intervention not only offers better/higher adaptation of the system but also keeps the system secure from malicious actions or misconfiguration of inexperienced users. All participants (users, providers, etc.) are known to the federation thus the chance of a malicious internal party causing harm is minimized. Multiple security frameworks can be applied and adopted by everyone to ensure that the information stored and exchanged in the federation is safe.

Regarding privacy management, all the data stored in the federations are governed by strict and clear policies and rules, approved by all participants. Actions performed on private data like association, disclosure, sharing etc are based on commonly agreed predefined rules.

Considering trust management, all participating providers are evaluated and meet a minimum set of requirements in order to be accepted in a federation. Multiple levels of trust can be defined to describe the relations between various entities and in all cases the metrics used to describe trust are approved and understood by everyone. Only the providers that participate in a federation are eligible to register or exchange information data with its DS. Thus the DS can be considered to be a trustworthy component.

Disadvantages

1) Not scalable

The main disadvantage of the federation approach towards the creation of a Global Distributed Discovery mechanism of identifiers, providers and capabilities is scalability. Federations are usually the result of business agreements that serve very specific "non-technical" purposes. Accordingly, their IdM solutions are always tightly coupled with very specific requirements and features of the services they provide. The federation approach can provide adequate solutions only in small scale where the data and entities that can be trusted and organized are few. In large or global scale, this approach is unable to form the required Circle of Trust and handle and organize the large amount of diverse formats, processes, protocols, etc.

Considering the DSs designed by federated systems, the majority of them are not operational outside federations' borders. Due to trust issues (private information sharing), any efforts to interconnect different federations and DSs are only possible with the formation of a new bigger federation.

2) Low levels of Security, Privacy and Trust management **in large or global scale systems**

The creation of a global Discovery System implemented through a global federation would require from all participants to agree on a minimum set of trust and privacy requirements and also share/store private information in the DS. In global scale, this requires the design and implementation of frameworks (security, privacy and trust) that could encompass diverse requirements from the all the various entities (providers, services, local laws, etc.) that participate in this federation.

Even if such a federation could be formed, there will still be cases where specific entities (providers, countries etc) will refuge to join and cooperate with others or cases where some entities will not be allowed to participate in the DS (who will have the responsibility to overview such a federation?), thus their information will never be available for discovery.

3) Association of identity information

One of the majors issues that IdM systems face today is the association of identity data which are somehow related (e.g. belong to the same user) and reside across multiple network places. The federation approach does not have the ability to collect and associate every piece of information that might belong to a single entity (e.g. user). The only information that can be organized in a federation is the data that hold in the providers that participate in it.

4) Ownership - One organization (or group of organizations) cannot host a global federation

The creation of a global DS based on the federation approach, would require a globally trusted organization (or group of organizations) to organize and supervise all activities. Such an action will set the supervisory organization in an abnormal strong position and would give it the ability to have access to all information, and control all the procedures that involve identity management.

What the Global DS can adopt from federated IdM systems:

A global DS cannot be based on a centralized architecture which collects, hosts and organizes all identity related information. Instead it will most likely be build around a distributed infrastructure where multiple participants contribute to the system but at the same time have the responsibility of hosting their own data and choose how to organize and share them. To create this infrastructure, the participants may need to agree and collaborate under a specific set of rules, in other words create some kind of federation which is not in the form described above. The DS can adopt some of the existing federations' techniques and procedures that would allow it to build this infrastructure in such a way that it meets all the requirements presented in clause 4.

5.2 User-Centric Identity Management Frameworks

Identity management systems that belong to the user-centric category give to the end users full control over their private data. These systems usually develop mechanisms to support direct communication and data submission/exchange between a user and a service provider. Identifiers proposed by these systems usually follow the format "identifier@location" and directly point to a location where the service provider may find additional information about the submitted identifier. Examples of such identifiers include the eXtensible Resource Identifiers (XRI), handles, Virtual Identities (VID), etc.

The discovery process on such systems e.g. XRDS usually relies on infrastructures that can discover metadata about a given resource and also forward a request to the appropriate location. This clause will examine if this type of systems and their DSs can support or provide partial solutions for creation of a Global Distributed Discovery mechanism of identifiers, providers and capabilities.

Advantages

1) Use of well established protocols

User centric IdM systems and their DSs generally use widely accepted protocols and formats. Since information are directly submitted by the end user to many different service providers, for interoperability issues, all involving parties follow widely accepted standards (e.g. SAML) to be able to exchange data.

2) Association of identity information

The ability of a global DS to provide adequate discovery services depends on the extend (level, quality) of the association between identity data that are somehow related/relevant and reside in various places. The way networks were built until today, resulted in the fact that, in global scale, large amounts of identity related data that belong to the same entity are stored across multiple places. Thus it is impossible for someone else (service providers, federations etc), other than the owner of these data, to locate, associate and manage them. Only the user centric approach provides the means to perform this action.

Disadvantages

1) Low levels of Security, Privacy and Trust management

User centric systems can offer very good mechanisms for security privacy and trust management, however the majority of the procedures are performed by the end users who in most cases are inexperienced and unfamiliar with technical details. Actions like misconfiguration, submission of wrong data etc, may lead to serious private information exposure and can also attract many malicious parties to perform e.g. phishing attacks. A Global DS where end users gather and organize their information by themselves is vulnerable to such threats.

The same vulnerabilities also affect privacy. In theory, user centric systems provide high levels of privacy since all the information are managed by the owners. In practice though, the absence of highly experienced entities in the process of privacy management usually leads to information exposure.

Finally trust management is not examined in user centric systems, since it is left to the user to decide whom to trust and give his personal information. Users usually rely on external trust information to make these decisions (e.g. reputation systems).

2) Not user friendly

The complexity of the IdM mechanisms and procedures is a major issue for the adaptation of user centric systems. Procedures like association of different identities registration of information in a DS, policy enforcement etc. are usually supported through complicated interfaces and processes which users find complex and tend to avoid or ignore.

What the Global DS can adopt from user-centric systems:

User centric systems provide to the owners of the identity information the means to independently organize their private data. The DS should examine and if necessary adopt user-centric techniques and practices that will give the end-users a central point of control for easy management over their private information. The creation of such a central point of control may also facilitate the collection of all the scattered unassociated identity data that today reside in multiple network places and belong to the same entity (e.g. user).

5.3 Discovery Frameworks

5.3.1 DNS, DDNS, DNSSEC

These systems were designed to provide and support a very specific service: the translation of domain names to IP addresses. Exploiting them to support discovery of identity related information will introduce high risks in term of security, privacy and functionality, since both systems have diverse requirements and operations.

What the Global DS can adopt from DNS, DDNS and DNSSEC frameworks:

The DNS system is a working example of a global discovery system in which multiple organizations can contribute to its infrastructure. However in its current form, it cannot host or contribute to the development of a Global DS for identities. The reason is that the DNS discovery process relies on caching information to various servers. This functionality is not acceptable in a Global DS for identities for privacy reasons.

The DNSSEC could be adopted to ensure the integrity of the participants that form the global DS infrastructure and enhance the overall system's security.

5.3.2 HANDLE

The Handle System® [i.8] is a large scale system which is able to locate a specific object irrespectively of its network location however it presumes that the requester already knows a persistent identifier about this object. HANDLE system has the ability to create a digital object that contains a group of other digital objects. This means that the framework can support association of multiple identifiers to a unique object (e.g. associating multiple identities to a user). The system in its current form is not able to organize user profiles or associate multiple identities, thus in order to support this functionality there are additional components that need to be developed.

What the Global DS can adopt from the HANDLE system:

HANDLE is a fully operational global discovery system with an open architecture that can be enhanced with additional functionality. It is globally deployed and it is governed by a non profit organization which only supervises its smooth operation. Each participant is able to organize and store its own data. With these features, HANDLE system is a good candidate to form the basis of the Global DS.

5.3.3 IF-MAP

IF-MAP [i.9] is a system which collects real time information about an object and notifies any updates about it, to all entities that have requested (through a subscribe process) to monitor its status. A discovery mechanism is not present in this system and all requesters are required to have previous knowledge of the corresponding IF-MAP server in order to monitor a specific object. The system has the ability to create updated profiles of objects and aggregate various data related to them. Current implementations provide solutions for closed contexts with strictly defined borders. The system in its current form is unable to operate outside such a context (e.g. for privacy reasons) and support a global scale discovery service.

What the Global DS can adopt from the IF-MAP:

In the Global DS, an entity (user, service, etc) would likely sustain a profile composed by multiple identities, attributes and other information. The IF-MAP framework may contribute its solutions to the Global DS for the aggregation and management of these data.

5.3.4 Plutarch

In the Plutarch [i.10] system there are no global names, and the name resolution mechanism involves the discovery of the appropriate context in which the target host exists. The discovery process is held through epidemic-style gossip advertisements and queries across contexts, and the result may be more than one candidate contexts. Even though this system is compatible with existing networks and technologies and can support communication among heterogeneous contexts, its main purpose is not identity management. Aspects like privacy (e.g. its discovery mechanism advertises queries to multiple places) or association of multiple elements (e.g. identities) are not examined.

What the Global DS can adopt from Plutarch:

Depending on its final architecture, there might cases where the Global DS would have to search for information stored in unknown locations. The epidemic-style discovery process of Plutarch may provide an adequate solution for such cases.

5.3.5 Distributed Hash Tables (Chord, Kademlia, CAN, etc.)

The Distributed Hash Table (DHT) systems, such as Chord [i.11], Kademlia [i.12], and CAN [i.13], are large structures intended to manage a large and totally distributed space of mappings from identifiers to objects, like in a typical hash table. Each of them is based on an overlay network and provides a stable mechanism to allow new components (nodes) to dynamically join and leave the structure, thus adding or removing storage space in a dynamic way. The overlay network mechanisms also ensure the scalability of the global system. The discovery operations are performed through the search of identifiers (keys) by crossing the DHT structure to reach the node that has the requested object.

What the Global DS can adopt from DHTs:

The Global DS can benefit from DHT functionality by incorporating its distribution mechanisms to achieve scalability, using its overlay network mechanisms for the dynamic building of the global system, and the ease distribution of the identifier namespaces in separate domains.

6 Conclusion

The present document is an effort to identify the requirements of a global distributed discovery mechanism of identifiers, providers and capabilities. Furthermore it examines which of the existing systems or mechanisms have the ability to support these requirements and thus contribute to the creation of this mechanism.

However it cannot be assumed, that the requirements presented in this study are the only ones or even the correct ones. Identity discovery and generally identity management are two highly complicated research areas influenced by multiple parameters like technological, geographical, ethical, legal, etc. Only by defining an architecture and developing a global discovery mechanism, it would be possible to establish in practice its final requirements.

Authors & contributors

The following people have contributed to the present document:

Rapporteur:
Konstantinos Lampropoulos, University of Patras, Greece (klamprop@ece.upatras.gr)

Other contributors:
Pedro Martinez-Julia, University of Murcia, Spain (pedromj@um.es).

Antonio F. Skarmeta, University of Murcia, Spain (skarmeta@um.es).

History

Document history		
V1.1.1	March 2014	Publication