Implementing scalable and consistent authorisation across multi-SP environments

Publication Date 2019-03-11
Authors: AARC Consortium Partners;AppInt members;Marcus Hardt (ed.)

Document Code: AARC-I047
DOI:

Grant Agreement No.: 730941
Work Package: JRA1

The research leading to these results has received funding from the European Community's Horizon2020 Programme under Grant Agreement No. 730941 (AARC2).

Abstract

The purpose of this document is to provide information to infrastructures for efficiently implementing access restrictions that are required by the individual communities and e-Infrastructures. The suggestions are given within the setting of the AARC BPA. In this scenario, user communities make use of an SP-IdP-Proxy (including User Attribute services) in order to manage access to resources (end services). The suggestions given address two different topics. One is about providing an interoperable schema to use for expressing authorisation information. This is an extension of the recommendations provided in AARC-G002 - Expressing group membership and role information and AARC-G027 - Specification for expressing resource capabilities. The other topic concerns the organisational architecture for conveying authorisation information. All information within this latter area are derived from the more detailed Deliverable DJRA1.2 on authorisation models.

Table of Contents

1. Introduction

Research Infrastructures and generic e-Infrastructures, referred to in the rest of this document as Infrastructures, provide services to increasing numbers of research communities.

The purpose of this document is to guide infrastructures in the efficient implementation of the access restrictions that are required by the individual communities and e-Infrastructures. The guidance presented in this document addresses two different topics. One is about providing an interoperable schema to use for expressing authorisation information. This is an extension of the recommendations provided in [AARC-G002] (Expressing group membership and role information) and [AARC-G027] (Specification for expressing resource capabilities). The other topic concerns the organisational architecture for conveying authorisation information. The information within this latter area is derived from the more detailed Deliverable [AARC2-DJRA1.2] on authorisation models.
The guidance is given within the setting of the AARC BPA [BPA]. In this scenario, user communities make use of an SP-IdP-Proxy (including Attribute management, possibly via Attribute Authorities). The users are given access to resources (end services) via infrastructure SP-IdP-Proxies.
Hence, we will assume the existence of a community SP-IdP-Proxy, which must comply with [SNCTFI]. Furthermore, authorisation information (in whichever way it will be expressed or conveyed) needs to be defined in a clear and interoperable way, such that it can be correctly interpreted across different infrastructures. Some of this information may come from the user's home organisation: Examples include the affiliation within the Home Organisation and the specific entitlement values required for accessing services such as GÉANT's Trusted Certificate Service [TCS]. However, communities cannot rely solely on information from Home IdPs, because this information alone may not be sufficient, because it is relatively static (can be updated only by the home organisation), or because the home organisation is not authoritative for the required information. Thus, in practice, authorisation models will primarily make use of community-managed authorisation attributes.

1.1. Conventions

The key words “MUST”, “MUST NOT”, “REQUIRED”, “SHALL”, “SHALL NOT”, “SHOULD”, “SHOULD NOT”, “RECOMMENDED”, “MAY”, and “OPTIONAL” in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Scope

All guidance given in this document is technology agnostic, but the context in which it is defined, will require translations between different technological realms. Guidance on translations of attributes and their values between SAML and OIDC can be found in the OIDCre Whitepaper [OIDCRE]. Guidance on usage with X.509 can be found in [AARC-G010].
The recommendations provided in this document relate to both the schema and the organisational flow for conveying authorization informatio. Note that different services exchanging information using attributes described in this document need to agree on their precise meaning. Hence, in addition to the general syntax described in this document, they will need a profile for the accepted attributes and their values.
The term “Attribute” is quite general and may be applied differently depending on the context. This will be the topic of Chapter 3. In Chapter 4 we will introduce the authorisation models which describe the flow of authorisation information, from managing attributes, to authorisation decision and enforcement.

3. Different types of authorisation information

Authorisation information can be classified into two types:

1. User-attributes (Such attributes are often aggregated from different sources) such as:
   - Affiliation within the Home Organisation and/or the Community [AARC-G025]
   - Assurance, i.e. how well a user is known (e.g. [RAF], [AARC-G021])
   - Group and role information (these primarily come from the Community) [AARC-G002].
2. Capabilities such as:
   - Capabilities describing what a user is entitled to do at a specific resource [AARG-G027].

The difference between entitlements[1] used for describing user-attributes as opposed to those that are used for describing capabilities will be clarified in Section 3.1.
The different authorisation models described in Chapter 4 also rely on this distinction of authorisation information.

[1] Technically speaking, groups, roles and capabilities can all be expressed using the same attributes or claims (for example using the eduPersonEntitlement SAML attribute).

3.1. Expression of user-attributes

Expression of authorisation information for user-attribute-based information is described in the REFEDS Assurance Framework [RAF], [AARC-G021] and [AARC-G002]. For example [AARC-G002] states that group membership information should be expressed as:[2]

```
<NAMESPACE>:group:<GROUP>[:<SUBGROUP>]...[:role=<ROLE>]#<GROUP-AUTHORITY>
```

The following example describes membership of a top-level group, "parent-group":

```
urn:example:example-ri.org:group:parent-group#auth-x.example-ri.org
```

The example below expresses a membership with a specific role, i.e. manager, in a group named child-group which is a subgroup of parent-group:

```
urn:example:example-ri.org:group:parent-group:child-group:role=manager\ #auth-x.example-ri.org
```

3.2. Expression of capabilities

Expression of capabilities follows the AARC Specification for expressing resource capabilities [AARG-G027]:

```
<NAMESPACE>:res:<RESOURCE>[:<CHILD-RESOURCE>]…[:act:<ACTION>[,<ACTION>]…]\ #<AUTHORITY>
```

For example the right to perform the actions create and delete on the storage resource at a **vm_dashboard** could be issued by the example-ri.org like this:

```
urn:example:example-ri.org:res:vm_dashboard:storage:act:create,delete#\  auth-x.example-ri.org
```

4. Authorisation models

Authorisation models describe the organisational flow of authorisation information. Any other information needed by the service to fulfil actions such as personalisation, accounting, traceability, is out of the scope of this document. The organisational flow of authorisation information follows this lifecycle:

- Definition of authorisation information at one or more Attribute Authorities (AA)
- Aggregation of authorisation information
- Use of authorisation information for making an authorisation decision
- Enforcement of the authorisation decision

The information provided in this chapter is based on the analysis of the authorisation models from nine different scientific communities or infrastructures. Details are available in DJRA1.2 ("Scalable, integrated authorisation models for SPs"[3]) [AARC2-DJRA1.2].

[2] Note that the **role** component is scoped to the rightmost (child)group.

[3] Please note that this document does not make use of the "P*P" terminology adopted in the deliverable document

From the models identified in DJRA1.2, we have identified three main approaches for managing authorisation, using an SP-IdP-Proxy.

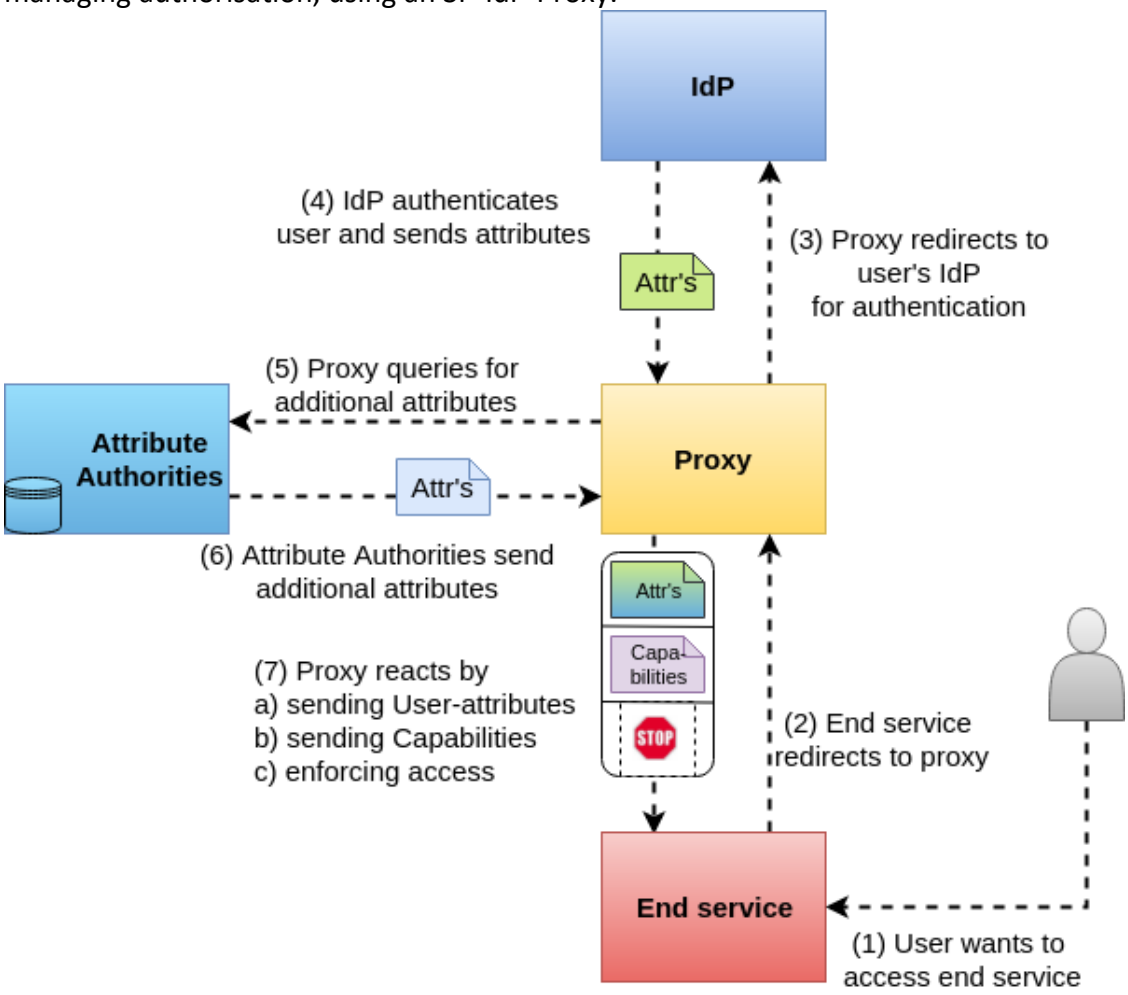

Figure 1: Flow for a user who wants to access an end service in a BPA-compliant infrastructure. The different authorisation schemes are indicated in step 7, where the proxy either (a) sends a combination of group or role information as authorisation information to the end service, which can then make a decision based on that information; or (b) takes the authorisation decision and re-expresses it as, for example, capabilities; or (c) takes the authorisation decision and enforces it by denying access to the end service. Note that option (c) may be followed by option (a) or (b) in case the proxy grants the user access to the end service.

We propose the following three authorisation models:

1. Centralised Policy Information Point: the proxy aggregates user-attributes and makes them available to the end-services (Step 7a in Image 1)
2. Centralised Policy Management and Decision Making: the proxy conveys the authorisation decision to the end-services (Step 7b in Image 1)
3. Centralised Policy Management and Decision Making and Enforcement: the proxy enforces the decision directly at the proxy (Step 7c in Image 1)

It should be mentioned that user-attributes and/or capabilities in the models above can be communicated to end services following either a “push” or “pull” approach [AARC-G006]. For example, in the SAML 2.0 Web SSO flow, attributes are pushed to end-services, whereas in

the OpenID Connect authorisation code flow, the end-services can either query the UserInfo endpoint or rely on the pushed id_token. These three models are described in more detail in the following sections.

4.1. Centralised Policy Information Point

In this model, the proxy aggregates the information and makes them available to the end services so they can make the authorisation decision. This allows the service to perform fine-grained access control, because all information necessary for an informed decision is available. However, scalability may become an issue for large deployments. For example, it may become non-trivial to consistently update authorisation with large numbers of services distributed across the globe, as the authorisation policy needs to be replicated to every service. Additionally, services may see user-specific authorisation data, such as group membership, that might be intended for other services. This may be problematic with regard to the "data minimisation principle". Furthermore, this puts the onus on the services to correctly interpret and act on the obtained authorisation information.

4.2. Centralised Policy Management and Decision Making

In this model the proxy makes the authorisation decision and encodes this decision into resource-specific authorisation information, typically in the form of capabilities. This allows the decision at the proxy to be based on additional information which the proxy might prefer not to send to the services. This is generally simpler for the end services to implement, since the complexity of interpretation of the authorisation information is handled by the proxy. In contrast to 4.1, this puts the onus to correctly interpret and act on the authorisation information obtained on the proxy. Note that in this model:

1. the proxy is creating and/or translating authorisation statements
2. the proxy may need to make a mix of capabilities and user attributes available for the service to be able to properly enforce the authorisation decision.

4.3. Centralised Policy Management, Decision Making and Enforcement

In this model, just like in the previous scenario, the proxy makes the authorisation decision, but additionally, it also enforces that decision. This allows the integration of services that might not be capable of doing any authorisation, with only little modification. However, it requires the proxy to understand the authorisation policy of the end services. Often this type of authorisation enforcement is only used for certain parts (e.g. a global black- or whitelist) while using the other models for the rest of the authorisation. For example, in case the proxy grants the user access to the end service, this model may be followed by the models described in sections 4.1 and 4.2.

5. Summary of Recommendations

This chapter provides a summary of the recommendations in this document.

- The release of attributes to services SHOULD follow the "data minimisation" principle (GDPR Article 5 (1.e) [GDPR, GDPR-INFO]. This might influence the choice between attribute- and capability-based access-control.
- To support traceability, as required by [SIRTFI, SNCTFI], implementations SHOULD at least do one, but preferably both, of the following:
  - maintain and send a pseudonymous unique identifier for the user from proxies down to the services.
  - maintain and send a unique ID that identifies the job or associated session.
- Group and subgroup membership and roles (where applicable):
  - SHOULD be expressed using [AARC-G002]
  - Subgroups (where applicable) SHOULD be used for expressing finer grained access permissions.
  - Roles (where applicable) SHOULD be used to specify additional rights inside the corresponding (sub)group. → [AARC2-DJRA1.3]
- Resource Specific Capabilities (where applicable) SHOULD be expressed according to [AARC-G027].
- Assurance information (where applicable) SHOULD be expressed:
  - Using the REFEDS Assurance Framework [RAF] and [AARC-G021] which extends [RAF] with additional assurance profiles recommended to be used between infrastructures.
  - In conjunction with specifications focusing on authentication, such as the REFEDS Single Factor Authentication (SFA) [REFEDS-SFA] and the REFEDS Multi-Factor Authentication (MFA) [REFEDS-MFA] profile.
- Affiliation information (where applicable) SHOULD be expressed according to [AARC-G025].

Considerations on the different models:

1. Authorisation implementations SHOULD support the Centralised Policy Information Point model for end services that require full control over the authorisation process. Authorisation implementations MUST be aware that in this model it is easy to send more data than required to end services. Filtering the information released to specific services MAY be a solution.
2. Authorisation implementations SHOULD support the Centralised Policy Management and Decision Making model for simplifying the authorisation process for the end services. Authorisation implementations MUST be aware that the onus for correctly interpreting and acting upon authorisation information is on the proxy.
3. Authorisation implementations SHOULD only use the Centralised Policy Management, Decision Making and Enforcement model for a partial authorisation decision (e.g. central suspension), but combine it with one of the two models above.
4. Depending on the requirements of the Service Providers reached through the proxy, it is possible to use a hybrid approach, combining any of the three models above, in a single authorisation flow. In all these flows the proxy can supplement the attributes from the authenticating IdP with information from AAs. The three different approaches address whether and how this information is passed on to the end services.

References

[AARC-G002] AARC Recommendation on Expressing group membership and role Information; https://aarc-project.eu/guidelines/aarc-g002

[AARC-G006] Best Practices for managing authorization; https://aarc-project.eu/guidelines/aarc-g006

[AARC-G010] Best practices and recommendations for attribute translation from Federated authentication to X.509 credentials; https://aarc-project.eu/guidelines/aarc-g010

[AARC-G021] Exchange of specific assurance information between Infrastructures; https://aarc-project.eu/guidelines/aarc-g021

[AARC-G025] AARC Recommendation on Exchange of affiliation information between Infrastructures; https://aarc-project.eu/guidelines/aarc-g025

[AARC-G027] AARC Specification for expressing resource capabilities; https://aarc-project.eu/guidelines/aarc-g027

[AARC2-DJRA1.2] AARC2 Deliverable on Scalable, integrated authorisation models for SPs; https://aarc-project.eu/wp-content/uploads/2018/07/AARC2-DJRA1.2_V4-FINAL.pdf

[AARC2-DJRA1.3] AARC2 Deliverable on VO Platforms for Research Collaboration; https://aarc-project.eu/wp-content/uploads/2018/10/AARC2-DJRA1.3-v2.pdf

[BPA] AARC Blueprint Architectures; https://aarc-project.eu/architecture

[GDPR] General Data Protection Regulation on eur-lex; https://data.europa.eu/eli/reg/2016/679/2016-05-04

[GDPR-INFO] Informational website for the General Data Protection Regulation; https://gdpr-info.eu

[OIDCRE] REFEDS White Paper for implementation of mappings between SAML 2.0 and OpenID Connect in Research and Education OIDCre Whitepaper; https://wiki.refeds.org/display/GROUPS/OIDCre

[RAF] REFEDS Assurance Framework; https://refeds.org/assurance

[REFEDS-SFA] REFEDS Single Factor Authentication Profile; https://refeds.org/profile/sfa

[REFEDS-MFA] REFEDS Multiple Factor Authentication Profile; https://refeds.org/profile/mfa

[RFC2119] Key words for use in RFCs to indicate Requirement levels; https://tools.ietf.org/html/rfc2119

[SNCTFI] Scalable Negotiator for a Community Trust Framework in Federated Infrastructures; https://www.igtf.net/snctfi

[SIRTFI] Security Incident Response Trust Framework for Federated Identity; https://refeds.org/sirtfi

[TCS] Trusted Certificate Service; https://www.geant.org/Services/Trust_identity_and_security/Pages/TCS.aspx