Advanced Grid Authorisation using Semantic Technologies – AGAST

Norman Gray[1 2], Richard O Sinnott[2], Thomas Doherty[2] and Jeff Lusted[1]

[1] Department of Physics and Astronomy, University of Leicester, UK,
[2] Department of Physics and Astronomy, University of Glasgow, UK
norman@astro.gla.ac.uk

Abstract. Collaborative research requires flexible and fine-grained access control, beyond the common all-or-nothing access based purely on authentication. Existing systems can be hard to use, and do not lend themselves naturally to federation. We present an access-control architecture which builds on RDF's natural strength as an integration framework, which uses RDF scavenged from X.509 certificates, and policies expressed as ontologies and SPARQL queries, to provide flexible and distributed access control. We describe initial implementations.

Collaborative research often demands finer-grained security that goes beyond the authentication-only paradigm of many e-Infrastructure systems. There are solutions for finer-grained access control, but in our experience these are often fragile, inflexible and difficult to establish, maintain and federate (for further discussion of these problems, see [1], which focuses on just one use-case). In this paper we present results of the JISC-funded AGAST project (www.nesc.ac.uk/hub/projects/agast), illustrating an approach to the specification and enforcement of security policies that is based on semantic technologies and the natural integratability of RDF data, and which addresses many of the limitations of existing security solutions. We illustrate the approach by describing sample client implementations in two scientific disciplines.

There is a good deal of contemporary interest in authentication ('AuthN') and identity technology, rather less so in authorisation ('AuthZ'). This is at least partly because this is hard, and that in turn is because the networked identities (on which access-control decisions must be made) are and will remain heterogeneous, time-varying, and distributed. The AGAST architecture, based as it is on the natural 'webbiness' of RDF and the associated standards of the Semantic Web, is a much more natural fit to this distributed, and opportunistically available identity and access-control information (ACI).

Existing technology in this area relies on distributed Attribute Authority (AA) architectures such as Shibboleth (shibboleth.internet2.edu), and centralised ones such as VOMS [2]. The typical scenario is as follows: the AAs communicate attributes to a Policy Enforcement Point (PEP, using the terminology of X.812 [3]) using SAML or X.509 Attribute Certificates (AC); once there, the role of the Policy Decision Point (PDP) is taken by a PERMIS system (www.permis.org), which implements a role-based access control system to determine whether the available attributes indicate that the user has a suitable role in a hierarchy represented by the PERMIS policy. Crucially, the role hierarchy is static, and the available attributes are fixed and globally pre-agreed; the mapping of attributes to PERMIS roles is also fixed, and fundamentally based on string comparison; there is no way to cope with attributes deliberately or accidentally having different meanings at different AAs (for example a 'lecturer' at one institution may be allowed to borrow an unlimited number of books, but the same role elsewhere may borrow only a limited number). That is, integration is hard, and though federation can be made to work, it is not natural to do so.

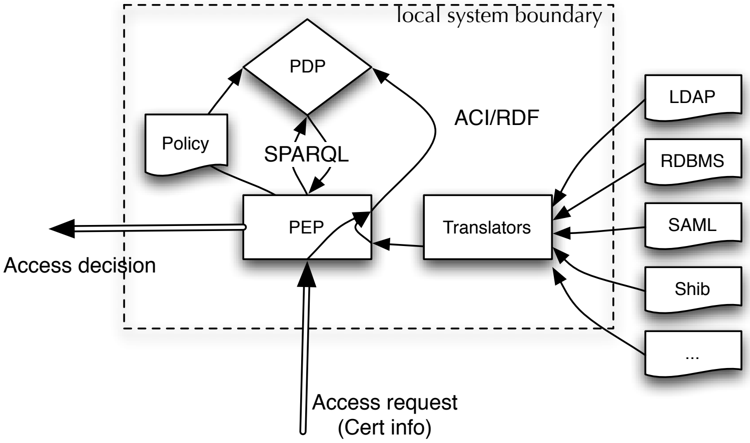

Fig. 1. The Semantic PDP architecture

We can use the natural integration features of RDF – its 'webbiness' – to support much more flexible access control policy specification and subsequent enforcement. For example: we may decide that a particular confidential resource can be seen only by members of the class `ns1:Medic`, which consists of the union of the classes `ns1:Doctor` and `ns1:Nurse` in some local or standard ontology. We then wish to integrate the class `ns2:Krankenschwester`, which is an element of an 'ontology' which may be an LDAP schema or AC attribute set, or be otherwise derivable from a X.509 certificate; the 'translators' in Fig. 1 translate this information into RDF. If we then state *as an element of local policy* that `ns2:Krankenschwester` is a subclass of `ns1:Nurse`, then immediately anyone whose (remote) credentials indicate they are of type `ns2:Krankenschwester` is (locally deemed to be) provably a `ns1:Nurse` and thus also a `ns1:Medic`; in this way distinct (PERMIS or other) role hierarchies can be aligned and combined straightforwardly and robustly. This is a simple example: the point is that expressing the access problem in RDF makes the relevant mappings almost trivially expressible. Crucially, all of the components of this chain of reasoning could potentially have come from different sources (interoperability), locally translated into RDF (the 'RDF-native' ACI of FOAF+SSL would be a natural fit here). In this scenario, the policy information is expressed in the set of available classes (and the mappings implied by the translators), the logical relationships between the classes, the grant of access to members of one of the classes, and the query which asks whether a user is a `ns1:Medic`. Reference [4] uses a similar service-based PDP, but without the semantic policies.

This architecture has been implemented in the AGAST project through a RESTful web-service acting as a PDP: *Qadi*. A PEP client of the service may define a policy as an ontology, as illustrated above. When a user attempts to access the service, the PEP obtains or extracts information about that user as RDF, and uploads that to a newly-created 'decider' (this may simply require posting the X.509 user- or attribute-certificate used for authentication). It then asks the decider 'is the user X in the class of entities permitted access?', to which the Qadi service can reply with a simple yes or no, which the PEP can act upon by permitting or denying the user access. The policy is expressed in the combination of the uploaded ontology and the SPARQL query used to interrogate it.

We have implemented case-studies using this architecture, in the context of clinical trials (VOTES, `www.nesc.ac.uk/hub/projects/votes`), and remote processing in astronomy. We will discuss both use-cases in detail, confirming the basic feasibility of the approach, and confirming that the server software is indeed substantially easier to set up and configure, and the client software easier to integrate, than existing well-developed approaches.

In summary: users have and will retain multiple identities and numerous pre-existing permissions to resources, and semantic web approaches can integrate these resources, and allow us to reason with them, *in a natural way*.

References

1. Richard O Sinnott, Thomas Doherty, Norman Gray, and Jeff Lusted. Semantic security: Specification and enforcement of semantic policies for security-driven collaborations. In *7th HealthGrid Conference, Berlin, Germany*, Studies in Health Technology and Informatics. IOS Press, 2009.
2. R. Alfieri, R. Cecchini, V. Ciaschini, L. dell'Agnello, Á. Frohner, A. Gianoli, K. Lõrentey, and F. Spataro. VOMS, an authorization system for virtual organizations. In *Grid Computing*, volume 2970 of *Lecture Notes in Computer Science*, pages 33–40. Springer, 2004.
3. Information technology – open systems interconnection — security frameworks for open systems: Access control framework (ITU-T Rec X.812 = ISO/IEC-10181-3:1996). International Standard ISO-10181-3/X.812, 1996.
4. Marcin Adamski, Michal Kulczewski, Krzysztof Kurowski, Jarek Nabrzyski, and Alastair Hume. Security and performance enhancements to OGSA-DAI for grid data virtualization. *Concurrency and Computation: Practice and Experience*, 19:1–11, 2007.