Chapter 77
Semantic-Based Access Control for Data Resources in Open Grid Services Architecture – Data Access and Integration (OGSA-DAI)

Vineela Muppavarapu
Wright State University, USA

Soon M. Chung
Wright State University, USA

ABSTRACT

This paper proposes a semantic-based access control system for the data resources in the Open Grid Services Architecture - Data Access and Integration (OGSA-DAI). OGSA-DAI is a widely used middleware for integrating data resources in Grids. However, the identity-based access control in OGSA-DAI causes substantial overhead for the resource providers in virtual organizations (VOs), because the access control information of individual users has to be maintained by each resource provider. To solve these problems, the authors propose a semantic-based access control system using Shibboleth and ontology. Shibboleth, an attribute authorization service, is used to manage the user attributes, and the Web Ontology Language (OWL) is used to represent the ontology of the data resources and users. By using ontology, VOs can resolve the differences in their terminologies and specify access control policies based on concepts and user roles, instead of individual resources and user identities. As a result, the administration overhead of the resource providers is reduced considerably. In addition, the eXtensible Access Control Markup Language (XACML) is used to specify the access control policies uniformly across multiple VOs. The authors also developed an XACML policy administration tool that allows the administrators to create, update, and manage XACML policies. The performance analysis shows that our proposed system adds only a small overhead to the existing security mechanism of OGSA-DAI.

DOI: 10.4018/978-1-5225-5191-1.ch077

1. INTRODUCTION

Grid is an integration infrastructure for sharing and coordinated use of diverse resources in dynamic, distributed virtual organizations (VOs) (Foster et al., 2001). A Data Grid is an architecture for the access, exchange, and sharing of data in the Grid environment. It facilitates the coordination and sharing of a large number of geographically distributed heterogeneous data sets across different domains in a controlled and secure manner (Foster & Grossman, 2003). Distributed data resources can be diverse in their formats, schema, quality, access mechanisms, ownership, access policies, and capabilities. As more and more organizations are participating in the Data Grids and sharing their resources, the complexity and heterogeneity of the systems is increasing constantly, and there is a clear need for standardized mechanisms to manage access control for the shared resources (Clemente et al., 2006).

The Database Access and Integration Services-Working Group (DAIS-WG) of the Global Grid Forum (renamed to Open Grid Forum (OGF)) is developing the standards for Grid interface to data resources (Malaika et al., 2003). The Open Grid Services Architecture - Data Access and Integration (OGSA-DAI) (Atkinson et al., 2005) provides the first implementation for these emerging standards, and supports query/transaction processing.

Currently, OGSA-DAI supports role-based access control (RBAC) (Ferraiolo & Kuhn, 1992; Sandhu et al., 1996) via a role-map file that maps individual Grid users to database roles. In this case, each resource provider has to maintain a role-map file to authorize access to its resources. This method of access control is not suitable for VOs, because both users and resources are dynamic in VOs. Managing up-to-date access control information of the users by each and every resource provider is a difficult task. Multiple entries in multiple role-map files may need to be updated if new users are allowed to access multiple data resources or if the access privileges of current users change. This puts an unnecessary burden on the resource providers when both the users and resource providers belong to multiple VOs. Furthermore, there are unnecessary overheads on the server side whenever users make invalid requests. This is because users are mapped to local roles and connected to the resource without first verifying their requests against their access privileges.

Several improvements to the current OGSA-DAI's access control method have been proposed (Pereira et al., 2006, 2007; Muppavarapu et al., 2010). In Pereira et al. (2006, 2007), we described how the Community Authorization Service (CAS) (Pearlman et al., 2002) can be used to provide RBAC in the OGSA-DAI framework. In Muppavarapu et al. (2010), we described how Shibboleth (Welch et al., 2005) and eXtensible Access Control Markup Language (XACML) (OASIS XACML, 2005) can be used for specifying the RBAC policies for distributed administration. Shibboleth is an attribute authorization service and is designed to provide user attributes to the resources for access control, and it mainly targets the Internet-based resources (Welch et al., 2005). XACML is a standard of the Organization for the Advancement of Structured Information Standards (OASIS) for describing the access control policies uniformly across different security domains (OASIS XACML, 2005). We used the Core and Hierarchical RBAC profile of XACML (OASIS, 2005) to specify the access control policies for multiple VOs.

However, all these approaches have two major problems. The first problem is that different organizations may use different terminologies in describing their resources and user roles. For example, different agencies, such as FBI and CIA, might use different terms to describe the same role: FBI may use the term *agent*, whereas CIA may use the term *field-agent*. These agencies will fail to match the relevant information when a specific term used in the request is different from those used in the resources and access control policies, despite having the same meaning (Davies et al., 2006). The second problem is