IGI PUBLISHING **ITJ 4915**

701 E. Chocolate Avenue, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.igi-global.com

This paper appears in the publication, International Journal of Grid and High Performance Computing, Volume 1, Issue 2
edited by **Emmanuel Udoh and Frank Zhigang Wang © 2009, IGI Global**

Using Policy-Based Management for Privacy-Enhancing Data Access and Usage Control in Grid Environments

Wolfgang Hommel, Leibniz Supercomputing Centre, Germany

ABSTRACT

Preventing the misuse of personally identifiable information and preserving user privacy are key issues in the management of IT services, especially when organizational borders are crossed. However, in Grid environments only few of the recent advantages in research areas such as privacy enhancing technologies and federated identity management have been adopted so far. In this article, we first present an analysis of the differences between Grid environments and previous models of inter-organizational collaboration. Based on requirements derived thereof, we demonstrate how existing policy-based privacy management architectures can be extended to provide Grid-specific functionality and integrated into existing infrastructures. Special emphasis is put on privacy policies which can be configured by users themselves, and distinguishing between the initial data access and the later data usage control phases. We also discuss the application of this approach to a XACML-based privacy management system. [Article copies are available for purchase from InfoSci-on-Demand.com]

Keywords: *Data Usage Control; Grid Governance; Policy-Based Access Control; User-centric Privacy Management*

MOTIVATION

As an essential part of their terms of use and privacy statements, organizations define which information about a customer and its users they demand in order to provide a certain service, and for which purposes the collected data will be used. Often, this

information is required for accounting and billing purposes as well as service personalization. Generally, it thus includes personally identifiable information (PII), i.e. data that can be used to uniquely identify a single person.

To prevent misuse of this sensitive data, such as selling addresses to marketing agencies, legislative regulations exist which restrict how the collected information may be used. Although privacy laws differ between countries and dedicated regulations exist for industrial sectors such as finance and healthcare, one common principle is that data must only be used for purposes which the user has been informed about and agreed to.

As intra-organizational solutions so-called privacy management systems have successfully been deployed over the past few years. Whenever a user's data is about to be accessed, rule sets are evaluated to determine whether the current access attempt is in accordance with the privacy policy the user has agreed to. Basically, such systems can be viewed as an extension of traditional access management systems in order to enforce the purpose limitation principle.

In inter-organizational service usage scenarios, such as Grid computing, privacy protection becomes a much more complicated issue, because multiple organizations – typically also located in different countries – are involved and service providers need to retrieve user data from the user's home organization in an automated manner.

Instead of a single organization's privacy policy, various demands must now be fulfilled. For example, there may be Grid-wide privacy policies, such as those specified by a virtual organization (VO), which must often be combined with service provider or user home organization specific policies, as well as policies eventually specified by the users themselves.

Privacy management thus becomes a two-tiered process: First, users must decide which of their data may be submitted to a service provider at all, and second they must be able to monitor and control how their data is being used.

In the research areas of privacy enhancing technologies (PET) and federated identity management (FIM) various solutions to these issues have been suggested, with several of them already being used in production environments; a short overview will be given in the next section.

However, these solutions are not suitable for certain characteristics of Grid environments, such as the concept of VOs, and cover only the user's PII, thus neglecting sensitive data submitted along with Grid jobs, such as medical records used as input data for those programs. In this article, we will discuss these differences of Grid environments and point out the relevant shortcomings of existing approaches and Grid-specific requirements.

Furthermore, we demonstrate how existing policy-based privacy management can be adapted to provide the additional functionality required in Grids. A proof of concept based on the policy language XACML is presented and the integration of the discussed privacy management components into existing infrastructures is outlined. This article is concluded by an outlook to our future research.