

EGEE

Security and Privacy of Accounting Information with DGAS

SECURE FLOW OF ACCOUNTING INFORMATION AND ACCESS AUTHORIZATION MECHANISM

Document identifier:	UNOFFICIAL
Date:	October 6, 2006
Activity:	JRA1: Middleware Engineering and Integration
Lead Partner:	INFN
Document status:	DRAFT
Document link:	

Abstract: This document gives a brief description of the Distributed Grid Accounting System (DGAS), with a particular focus on security and privacy-related issues. It describes the flow of accounting information and the measures taken to ensure its integrity and confidentiality, as well as the authorization mechanism applied to queries to Home Location Register (HLR) for information retrieval.

Delivery Slip

	Name	Partner	Date	Signature
From	A.Guarise, G. Patania, R.M.Piro, A.Werbrouck	INFN Torino		
Reviewed by				
Approved by				

Document Change Log

Issue	Date	Comment	Author
Initial version	October, 2006	-	R.M.Piro, A.Guarise, G. Patania, A.Werbrouck

Document Change Record

Issue	Item	Reason for Change
--------------	-------------	--------------------------

Copyright ©Members of the EGEE Collaboration. 2006. See <http://eu-egEE.org/partners> for details on the copyright holders.

EGEE (“Enabling Grids for E-science in Europe”) is a project funded by the European Union. For more information on the project, its partners and contributors please see <http://www.eu-egEE.org>.

You are permitted to copy and distribute verbatim copies of this document containing this copyright notice, but modifying this document is not allowed. You are permitted to copy this document in whole or in part into other documents if you attach the following reference to the copied elements: “Copyright ©2005. Members of the EGEE Collaboration. <http://www.eu-egEE.org>”

The information contained in this document represents the views of EGEE as of the date they are published. EGEE does not guarantee that any information contained herein is error-free, or up to date.

EGEE MAKES NO WARRANTIES, EXPRESS, IMPLIED, OR STATUTORY, BY PUBLISHING THIS DOCUMENT.

CONTENTS

1	INTRODUCTION	4
2	FLOW OF ACCOUNTING INFORMATION IN DGAS	4
3	INFORMATION RETRIEVAL WITH DGAS	5
3.1	USER HLRS AND RESOURCE HLRS	6
3.2	USER QUERIES AND PRIVACY	6
4	CONCLUSIONS	6

1 INTRODUCTION

The Distributed Grid Accounting System (DGAS), previously called DataGrid Accounting System [1, 2], is an accounting toolkit conceived and designed to be completely Grid-oriented. It is based on fully distributed client/server infrastructure without having a central repository of accounting information, relying upon a network of independent accounting servers used to keep the accounting/transaction records.¹

Since user-level accounting information is strictly confidential the protection of privacy is an important issue when designing an accounting system. According to EU directives and national laws, for example, the publication of data that could be used to track down the activities of a single individual is usually not permitted. Personal data (basically any information relating to an identified or identifiable natural person) has thus to be handled with proper measures to ensure confidentiality and allow access only to authorized persons.

Nonetheless, any user should have the right to access his/her own accounting information. VO managers as well may need to access accounting information regarding their users to guarantee a proper management of the VOs resources.

In this document we describe the flow of accounting information and the measures taken to ensure its integrity and confidentiality, as well as the authorization mechanism applied to queries to Home Location Register (HLR) for information retrieval.

Note, that in most cases, an informed user consent for the collection, storage and processing of user-level accounting information may still be required. This has to be addressed by proper agreements between users, their VOs and the sites that provide the Grid services. Such agreements are, however, out of scope of this document, that assumes that appropriate agreements are in place and users have given their consent. In this document we instead address the question how DGAS can help in ensuring that these agreements can be fulfilled while preserving the users' confidentiality and privacy.

The remainder of this document is organized as follows. Section 2 describes the flow of accounting information in DGAS and discusses the applied security mechanisms. Section 3 concentrates on information retrieval and the authorization mechanisms implemented for privacy and confidentiality. The document is concluded with some final remarks in Section 4.

2 FLOW OF ACCOUNTING INFORMATION IN DGAS

The flow of accounting information in DGAS is depicted in Fig. 1. The accounting information is collected (step 1) by dedicated metering sensors on the Computing Element (CE). These sensors parse LRMS² accounting log files to gather usage information for the single jobs. These log files, however, contain for most LRMS types only local information (that is, for example, local job ID, local user ID, ...). Grid-related job information (User DN, User FQAN, Grid Job ID, Grid CE ID) is taken from an additional log file written by the CE's Grid middleware.³

The collated usage records (URs) are then sent (step 2) from the CE to the DGAS Home Location Register (HLR) server that manages the resource's (CE's) account. This "Resource HLR" can be a remote accounting server, but the general deployment foresees to have one Resource HLR server at each site (in this case it is also called a "Site HLR").

The transmission of a UR from the CE to the Site HLR is done via the Globus Security Infrastructure (GSI) [3] using the CE's host certificate. This means that the entire UR is encrypted during transmission

¹DGAS also provides a network of independent servers for resource pricing in order to enable the deployment of a Grid resource market, but resource pricing is optional in DGAS and of no importance for the purpose of this document.

²Local Resource Management System, e.g. PBS and LSF.

³This describes the new approach that we proposed with a dedicated log file written by BLAH for gLite and the gatekeeper for LCG. The Condor developers have also agreed on providing such an additional log file.

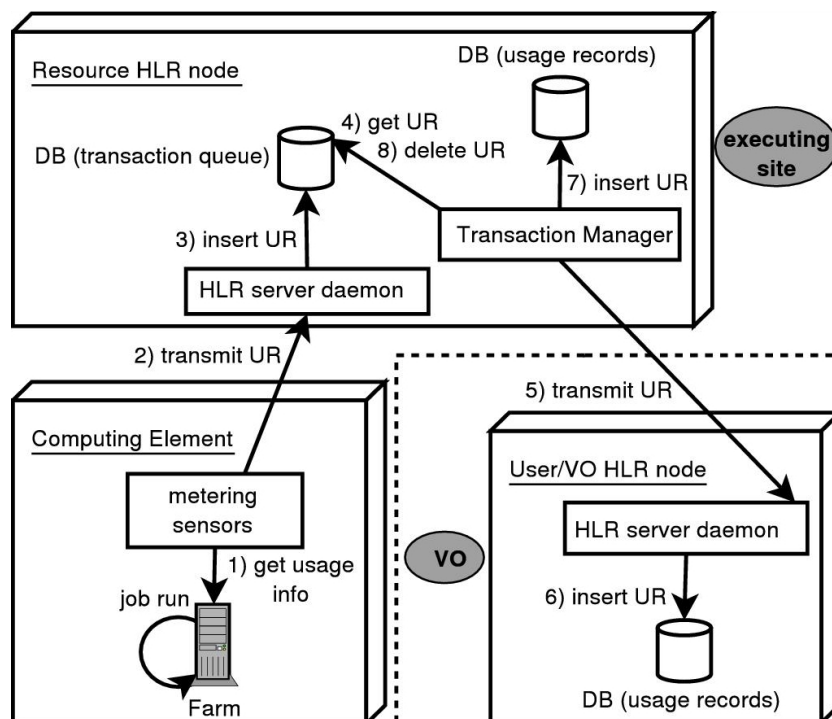


Figure 1: Flow of accounting information in DGAS.

and cannot be read without the Site HLR's host certificate, even if intercepted by an unauthorized third party. Additionally, this provides a means for assuring that only authorized CEs can submit job usage records to a given Site HLR, since GSI connections can be recognized by the initiating certificate (the CE's host certificate in this case) and be selectively refused. Allowing usage records to be transmitted by authorized entities only, is important to ensure their authenticity.

On the Site HLR URs are first inserted into a transaction queue (step 3) and then asynchronously processed by a Transaction Manager daemon (steps 4, 5, 7 and 8). The exact technical details of this procedure is out of scope of this document. The only step that is important in the context of security is step number 5 in which a UR is forwarded from the Resource HLR to the HLR server that manages the user's account ("User HLR" or "VO HLR" if all users of a VO are managed by the same HLR). This communication is done via GSI as well using the Resource HLRs certificate, hence URs are encrypted in their entirety and User HLR servers can receive URs only from authorized Resource/Site HLRs for inserting them in their own database (step 6).

Note that is the site that can configure whether URs should be forwarded from the Site HLR to the User HLRs or not (since some sites stated the wish not to let accounting information leave their administrative domain, although we think that Grid users should have the right to easily retrieve accounting information regarding the jobs they submitted). It is even possible to forward accounting information from the Site HLR to the User HLRs only for specific VOs. This is important in case only some of the VOs sign an agreement on the processing of the accounting information of their users.

3 INFORMATION RETRIEVAL WITH DGAS

The Home Location Register (HLR) service does not only store accounting information, but it also processes remote queries for information retrieval. Usage information can be obtained from the single HLR servers for single jobs as well as in aggregate form (per user, per resource, per VO).

3.1 USER HLRs AND RESOURCE HLRs

DGAS associates accounting information to previously registered user and resource accounts and foresees two logical types of HLR servers: User HLRs and Resource HLRs, as already described in Section 2. A Resource HLR stores information from a resource owner's or site manager's point of view and is the DGAS server that resource owners, or site managers, can query for information concerning their resources.⁴

The reason of this division is straightforward. In order to guarantee a reasonable scalability there will be many HLR servers on the Grid, and different CEs will be registered with different Resource HLR servers. Hence it is advisable that all accounting information concerning a given user be forwarded from the different Site HLRs to the HLR that manages the user's account ("User" HLR) in order to be able to compute accounting statistics for the single Grid users although they submit jobs to many different Grid resources.

With a distributed accounting system with duplicated usage records each Grid participant (user or resource owner) needs to query only a single HLR server in order to have an exhaustive accounting view, nonetheless preserving a reasonable scalability that cannot be achieved through a single centralized accounting repository.

3.2 USER QUERIES AND PRIVACY

Each User and Resource/Site HLR server can be queried by command line clients to retrieve accounting information. For querying an HLR server the user has to authenticate with a valid user certificate proxy. All communication for this purpose is encrypted using the user certificate proxy.

Access to private information is granted only to authorized users. That is, users can generally access only information regarding their own jobs. Users registered as HLR administrators with an HLR server can access all information available from that server. This is useful above all for VO administrators (and site managers) that might need to access information about the resources consumed by their fellow users (or the resources provided by their CEs).⁵

In case users of several VOs are managed by the same User HLR server (like, for example, several VOs can be handled by the same VOMS server) it is also possible to register their managers as VO administrators (instead of HLR admin with full access rights). In this case each manager can get only accounting information regarding its own VO, not that of other VOs managed by the User HLR server. The same mechanism can in theory be used on Site HLRs, that is a site administrator might allow a VO manager of a particular VO to query the Site HLR only for information about that VO. Before granting such access rights, however, an agreement between sites, users and VO managers will be necessary.

This authorization mechanism, that should ensure the confidentiality of the accounting information even if it has been forwarded to the User HLRs, is currently being improved, such that authorization decisions can also be based on VOMS attributes that is VOMS groups and roles (extracted from the User FQAN that comes with a VOMS certificate proxy). The VOMS-based authorization mechanism has already been implemented but will need further testing before being released.

4 CONCLUSIONS

DGAS was designed for low-level usage accounting (tracking single jobs and precisely mapping the usage information to Grid resource IDs (CE IDs), Grid user IDs (user DNs) and Grid job IDs (not only

⁴Each HLR server, however, can manage both user and resource accounts if required.

⁵Note that giving the VO administrators access to private user information might require a specific VO agreement on the treatment of accounting information, but, as explained in the introduction, this is out of scope of this document.

local user IDs and LRMS job IDs).

The strengths of DGAS lie in its distributed and flexible architecture that significantly improves its scalability, but also its security layer and its strict authorization mechanism that provide the possibility for users (and VO admins) to obtain detailed usage information for their jobs, nonetheless respecting their privacy. The security layer and authorization mechanism have been briefly described in this document.

More detailed information on DGAS can be found at: <http://www.to.infn.it/grid/accounting/>

REFERENCES

- [1] R.M. Piro, A. Guarise, and A. Werbrouck. "An Economy-based Accounting Infrastructure for the DataGrid". *Proceedings of the 4th International Workshop on Grid Computing (GRID2003)*, held in conjunction with the SC2003 Conference, Phoenix, Arizona, November 17, 2003.
- [2] DGAS Home. <http://www.to.infn.it/grid/accounting/>
- [3] I. Foster et al. "A Security Architecture for Computational Grids". In *Proc. of the 5th ACM Conference on Computers and Security*, pp. 83-91, ACM Press, 1998.