DataGrid

EARTH OBSERVATION APPLICATION SECURITY USECASES

TECHNICAL NOTE

Document identifier:	**DataGrid-09-TN**
Date:	**17 June 2003**
Work package:	**WP9 Earth Observation Applications**
Partner(s):	**IPSL, ESA, KNMI,. ENEA**
Lead Partner:	**ESA**
Document status	**DRAFT**

Abstract:

This technical note presents and discusses Earth Observation Application security usecases and requirements for VO and group membership, authentication and authorization and access control to shared datasets in a grid middleware environment. It is intended as input for discussion and eventual design of the EDG security architecture and implementation

Delivery Slip

	Name	Partner	Date	Signature
From	J. Linford	ESA		
Verified by				
Approved by				

Document Log

Issue	Date	Comment	Author
0.0.1	17/06/2003	First draft	ESA
0.0.2	20/06/2003	Contributions from IPSL and KNMI	ESA

Document Change Record

Issue	Item	Reason for Change
0.0.2	Various	Comments & feedback from C. Leroy, W. Som de Cerff

Files

Software Products	User files
Word	DataGrid-wp9-TN_security_usecases.doc

CONTENT

1. INTRODUCTION

1.1. OBJECTIVES

This technical note aims to describe EO security usecases and requirements for grid as input to the development of the EDG security architecture design and implementation. In particular it aims to describe usecases and requirements related to :

- VO and group membership
- authentication, authorization and access control to datasets shared between one or more groups in one or more VOs

This will be achieved by analysing the security aspects and issues to be addressed for the EO GOME processing and validation usecase, in order to describe and illustrate the features and capabilities required for future versions of the EDG grid middleware security design and implementation.

1.2. APPLICATION AREA

The principal area of investigation is the advance RTD (Research and Technological Development) of computing grids and testbeds for the deployment of resource- and data-intensive scientific applications.

Both the task and the resulting deliverable revolve around the Earth Observation community and application area. However, in order to be useful to the widest possible range of applications, the core of the chosen approach is intentionally application-independent. Therefore, the contents of the document are neither exclusive nor specific to Earth Observation: they may be equally applicable in both the HEP and Biomedical application areas, among others.

1.3. REFERENCE DOCUMENTS

Reference documents

R1 DataGrid Wp9 GOME Use Case UseCase
http://edms.cern.ch/document/346050
(KNMI IPSL ESA), WP9.4 Use Case Annex to Requirements specification: EO application requirements for Grid, May 2002, annex to EU deliverable D9.1

R2 Earth Observation application requirements specification*
http://edms.cern.ch/document/332411
DataGrid WP9 Deliverable Document D9.1, December 2001, EU deliverable

1.4. DOCUMENT AMENDMENT PROCEDURE

This document is under the responsibility of KNMI, ESA/ESRIN and IPSL. Amendments, comments and suggestions should be sent to the person in charge of the document.

1.5. TERMINOLOGY

Definitions

Glossary

X.509 A standard public key security cryptography

The DataGrid Project Glossary is at : http://eu-datagrid.web.cern.ch/eu-datagrid/Glossary.htm

2. EXECUTIVE SUMMARY

This technical note presents and discusses Earth Observation Application security usecases and requirements for VO and group membership, authentication and authorization and access control to shared datasets in a grid middleware environment.

The document describes EO security usecases and requirements for grid as input to the development of the EDG security architecture design and implementation. In particular it aims to describe usecases and requirements related to :

- VO and group membership
- authentication, authorization and access control to datasets shared between one or more groups in one or more VOs

Relevant security aspects and issues for the EO GOME processing and validation usecase are analysed in order to describe and illustrate the features and capabilities required for future versions of the EDG grid middleware security design and implementation.

3. USECASE SCENARIO - CONTROLLED ACCESS TO SHARED DATA AND RESOURCES

The usecase involves three scientific investigation groups which need controlled access to shared datasets stored on the grid Storage Elements and registered using the Replica Location Service.

3.1. USECASE DESCRIPTION

Two groups (KNMI, ESA) process Level1 data to produce Level2 data and the other group (IPSL) validates the Level2 products. Table 1 illustrates the three datasets involved and the required access control.

Table 1. Datasets and access control among three scientific groups within the EO VO

Dataset	IPSL		KNMI		ESA	
Level1			r		r	
Level2	r				r	w
Level2b	r		r	w		
Validation	r	w	r		r	

- KNMI and ESA need read-only access to Level1 data and read-only access to Validation data
- KNMI needs read-write access to Level2b data
- ESA needs read-write access to Level2 data
- IPSL needs read-only access to Level2 data and read-write access to Validation data

However, it must still be possible (i.e. for any of the users) to create replicas of read-only data

3.2. REQUIREMENTS

3.2.1. The middleware should provide access control at the group level[1]

3.2.2. When a user obtains grid authorization he/she must also be able to choose the VO and one or more groups within the VO

3.2.3. A job submitted to the grid via the Workload management System, or a request to the Replica Management System must carry information about the user's access permissions to allow the middleware to determine which group(s) the user belongs to

3.2.4. The middleware should control the creation, modification and deletion of both the physical files on the Storage Elements and the associated information in

[1] For example, in the usecase there are three groups defined: NNO, OPERA and VAL (example of section 3.4.2.1)

the Replica Catalogues in a secure, coordinated and consistent way, according to the user's access permissions

3.2.5. Whenever a file is created the access permissions associated with it must also be defined and must remain associated with the file

3.2.6. All subsequent accesses to the file, as well as any Replica Management operations performed on it, must be checked for consistency with the user's access permissions

3.2.7. Jobs running on the grid must carry with them the default access permissions to be applied to all file and replica management operations

3.3. CURRENT SITUATION

3.3.1. VO Management

The information for each VO is kept in an LDAP database. Each VO has an assigned administrator, who is authorized to add or remove users in the database. When a user requests a certificate he/she chooses the VO which the certificate will be valid for.

3.3.2. Access control

The EDG release 1.4 does not allow the definition of groups within VOs and therefore cannot make any distinction between the three scientific groups within the EO VO.

Neither the user credential nor proxy certificate provide any VO or group information - only the user's identity and organization are provided.

Security for jobs running on the CE is implemented using the standard unix permissions which allows three levels of access control: `user` (owner of the file), `group` (members) and `other` (everyone else).

A single group is created for the VO on each CE and a number of pooled user accounts are also created which are shared among the VO users. Therefore all EO VO members get mapped to the same group, identified as '`eo`', and the user account mapping varies, e.g. '`eo001`', '`eo002`', ... '`eo099`', being assigned on a first-come-first-served basis.

The unix user-level access control cannot be used for providing any sort of coordinated access control among the VO members. Group-level access control (i.e. for multiple groups within a VO) is not possible since there is only a single group allowed within each VO. Owner-level access control is also not possible due to the arbitrary mapping of VO users to the local site pooled accounts.

In general, for EDG jobs that need to write maintainable datasets (i.e. re-writeable by the creator), the default access permissions must be set to allow read-write group access, since it cannot be guaranteed the owner will map to the same local account on different CEs. For instance, a job which created a file on a SE which is later re-submitted again (i.e. to redo the processing) is sent by the RB to a different CE; if the job is mapped to a different user account it may be unable to re-write the same file (because the file was first created by a different user).

3.3.3. Modification of access permissions

A job running on the grid is automatically assigned group and owner ids by default and can override these using the unix commands `chown`, `chmod`, `chgrp`, `umask`, etc. (depending on the user privileges). However, these commands are of limited usefulness, since both group-level and owner-level access control are not usable.

3.4. PROPOSED SCENARIO

Please note it is not intended to suggest an implementation, the solutions in the following scenario are used to illustrate the required features.

3.4.1. VO Management

The VO directory/database should be modified to allow the configuration of groups within the VO. In particular, it must be possible to create/modify/delete groups and to assign users to one or more groups.

When new users join the VO they should be allowed choose one or more groups within the VO and to specify whether *read-only* or *read-write* access is required for each group.

To allow the VO manager to authorize users' requests for group membership, a representative associated with each group should be provided (e.g. email address, telephone) when the group is first created in the VO database. The VO manager can then contact the group representative to verify whether the requested permissions should be granted.

3.4.2. Access control

Prior to running one or more grid jobs, after the user has generated the initial proxy certificate, he/she must be able to obtain access permissions for one or more VO/groups. The access control middleware should contact the relevant VO servers to authorize the requested permissions. The access permissions obtained should be attached in some way to the proxy certificate and thus remain available for middleware services that need it. For end-user interfaces (e.g. portals) it must be possible to automate this behind a single grid login procedure.

3.4.2.1. Example: controlled shared access to EO GOME data and validation products

The access permissions set out in section 3.1 suggest the creation of three groups of users within the VO : NNO (Level2 data production), OPERA (Level2b data production) and VAL (Level2 products validation). Users from KNMI, IPSL and ESA register with the VO and choose which of the groups they needs to join and the type of access (read-only or read-write). Table 2 shows the groups and access rights assigned to the users in order to provide the access control required in section 3.1.

- ESA users have read-write membership to the NNO group and read-only membership to the VAL group
- KNMI users have read-write membership to the OPERA group and read-only membership to the VAL group
- IPSL users have read-write membership to the VAL group and read-only membership to the NNO and OPERA groups

Table 2. VO group membership and access rights for EO GOME usecase

Groups	VAL		OPERA		NNO	
Access type	R	RW	R	RW	R	RW
Users						
ESA	x					x
KNMI	x			x		
IPSL		x	x		x	

3.4.2.2. Example: shared access between different VOs and groups

This example illustrates users from two different application areas, Earth Observation and Earth Science needing to control access to shared data across two distinct VOs.

The two VOs are *Earth Observation (EO)* and *Earth Science (ES)*, which each have associated groups. ES applications produce meteorological data products, while EO applications produce oceanographic data. An environmental monitoring application (oil slicks), needs to use both datasets.

The environmental application (ENV) needs to run grid jobs which merge *ES/meteo* and *EO/ocean* data and creates new *EO/oilslick* data. The researcher has joined the *ocean* and *olislick* groups in the *EO* VO, and the *meteo* group in the *ES* VO. When joining the VOs the researcher requests read-only access to *ocean* and *meteo* data and read-write access to *oilslick* data. The VOs, groups and access rights scheme required for this example are shown in Table 3.

Table 3. VO group membership and access rights for data merging example

VOs	ES		EO			
Groups	METEO		OCEAN		OILSLICK	
Access type	R	RW	R	RW	R	RW
User application						
ENV	x		x			x

During the grid authorization/login procedure the researcher specifies the VO/groups to use for the session. The middleware contacts the associated VO servers to authorize and obtain the access control information, which it attaches it to the proxy certificate (as an ACL?). The proxy certificate then accompanies the job, providing the relevant access control information to other middleware components that require it (e.g. Storage Element, Replica Manager). Whenever the job access data stored on a grid SE the SE access middleware first checks the job's access capabilities.

Whenever the OILSLICK job writes new data to the SE (and carry out RM operations) the SE and RM middleware can determine the group to use (i.e. only OILSLICK group can write). In other cases the job (capabilities) may have write-access to more than one group. In such cases the job needs to expressly set the access permissions to be used before accessing the SE/RM.

3.4.3. Modification of access permissions

It must be possible for a job running on the grid to modify the current access permissions, as long as the capabilities has sufficient privileges to allow it.