Author: Yangwoo Kim
February 25, 2006
Updated May 8, 2006

Grid Information Retrieval System for Dynamically Reconfigurable Virtual Organization

Status of This Memo

This memo provides experimental information to the Grid community regarding implementation issues for Grid Information Retrieval (GIR) system with dynamically reconfigurable virtual organizations (VOs). It does not define any standards or technical recommendations. Distribution is unlimited.

Copyright Notice

Abstract

Under foundational precepts of Grid computing, two important requirements that all Grid application systems should satisfy are to accommodate the dynamic nature of Virtual Organizations (VOs), and to enforce different levels of security among different VOs. For the research described in this document, we developed two different use-case scenarios addressing the two requirements, and then showed how the requirements can be met by implementing a Grid information retrieval (GIR) system prototype. The dynamic nature of VO applies not only to increasing and decreasing number of users, but also to the dynamically changing requirement of computing power among the different subcomponents that consist in overall system configuration. This implies that a request to increase computing power by a certain subcomponent can be satisfied by other idling subcomponents taking advantage of overall system flexibility.

This document describes how we implemented a Grid IR system using VO and security mechanisms provided by Globus toolkit 3.0, and shows how GIR system scalability and security can be improved for dynamic VOs. In order to manage different VOs, we implemented VO management service (VOMS), and registered it to Globus as an additional service.

Contents

1. Introduction

Grid information retrieval (GIR, see GAMIEL)[6][7] describes an information retrieval system based on Grid technology[3][5]. Unlike more conventional contemporary IR systems with dedicated computing resources in cluster architecture, a Grid IR system can be dynamically reconfigured using idling shared computing resources within a VO. Moreover, Grid middleware such as Globus [13][15] provides sophisticated multilevel security mechanisms for VOs.

Information retrieval (IR) may be generally defined as a means of matching documents (or document extracts, summaries or similar content) to human information needs. These information needs are typically expressed as queries, which are matched against document representations held by an IR system. While Web-based search engines are perhaps the best known IR systems today, there are many other types. Historically, bibliographical IR systems helped information seekers to find journal articles or similar documents of interest. Online library catalogs are IR systems that seek to match journals, books or other content types to human information needs. Within organizations, several varieties of IR systems or systems with IR components may be used, including decision support systems, management information systems, and more traditional bibliographical retrieval systems.

Grid computing has been introduced as a next generation Internet infrastructure and there have been many active, successful attempts to integrate science applications to Grid computing, as highlighted in the Global Grid Forum (http://www.ggf.org). Herein this document, we will describe our effort to apply Grid computing technology to Information Retrieval (IR), which is often thought of as a commercial application area, rather than a scientific area. In order to apply Grid computing to a commercial application area, we have to meet two main requirements. One of them is to accommodate the dynamic nature of Virtual Organizations (VOs), and the other is to enforce different levels of security among different VOs [1][2][11][13].

VO is a technical concept for virtual binding of geographically distributed computing resources and users. The dynamic nature of a VO applies not only to increasing and decreasing numbers of users, but also to the dynamically changing requirement of computing power among the different subcomponents that make up the overall system configuration [2][4]. This implies that a request to increase computing power by a certain subcomponent can be satisfied by other idling subcomponents taking advantage of overall system flexibility. This ‘On-Demand’ computing concept can be applied to information retrieval system where dynamic system reconfiguration is possible for dynamic computing power change. It may provide a new flexible computing paradigm for IR where dynamic scalability and subdivided security enforcement are possible [1][8].

2. Grid Information Retrieval System for Dynamic Virtual Organizations

In this document, we introduce a new information retrieval service model [6][7] by incorporating the dynamic VO concept and flexible security mechanisms. We discuss two use-case scenarios to show the usefulness of such a system, and then to demonstrate system implementation.

2.1 Dynamic Reconfiguration of Virtual Organization

2.1.1 Use Case One: A Scenario for Reconstruction of Various Dynamic Virtual Organizations

“The stock” security corporation decided to adopt a new Grid information retrieval system to improve IR system performance and capacity. They expected that the new IR system would employ Grid technology such that overall computing capacity and configuration could be dynamically scaled up by expanding the VO to include idling computing components. Three phases of this use case are envisioned, during different phases of the system’s operation.

First, before the opening of stock market every morning, the corporation's IR system needs to gather all information that may affect stock market rising or falling. We consider a GIR system modeled after the GIR Requirements document [7], in which the three main components are a collection manager (CM), query processor (QP) and indexer/search (IS). In the three-part GIR model, the CM is responsible primarily for ingesting documents, including any needed authorization, document transformation, staging or temporary storage. The query processor deals with sending user queries to the IS components, merging & presenting results, and maintaining long-running queries for ongoing information needs [6][7]. In GIR, the information gathering operation is usually carried out over night until the opening of the stock market, and it requires very high computing power for the CM (collection manager) component. However, at the same time, the QP (query processor) and IS (indexer/searcher) components were idling most of the time and thus have redundant computing power. The redundant computing nodes can be reassigned to act as CMs for better overall system utilization as shown in Fig. 1.

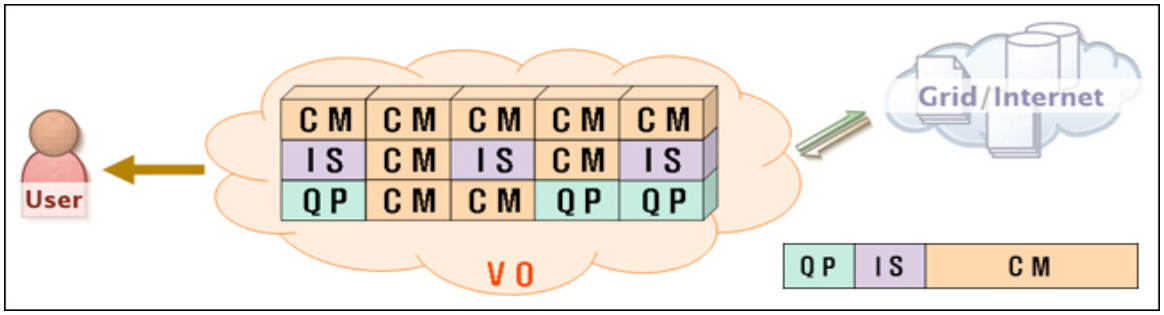

Figure 1 : An Example Model for Collection Manager Expansion within a Virtual Organization

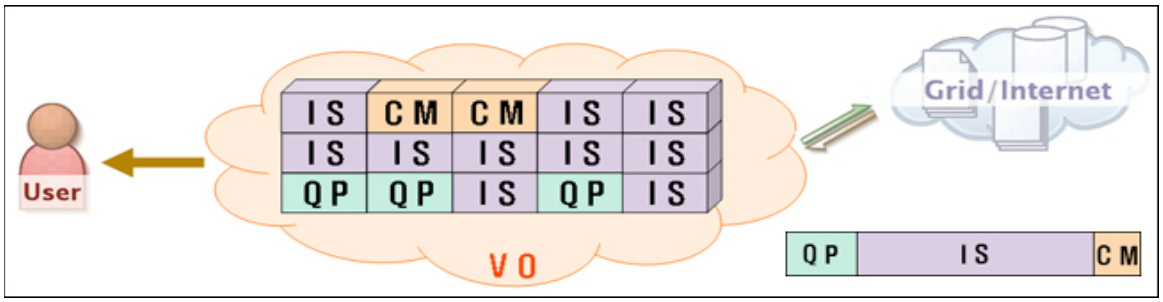

Figure 2 : An Example Model for Index/Search Expansion within a Virtual Organization

Second, the security corporation's IR system needs to analyze political, economical, social, and other international situations that may affect a stock market before their customers make investments. This kind of service needs data mining techniques for integrating a lot of documents with different types and structures. Therefore, there is higher computing power requirement for IS to integrate the geographically distributed databases when providing real time analysis services to customers as shown in Fig. 2.

Third, the security investment corporation wants to provide an analysis service to their customers for their investments using PDAs (personnel digital assistants). This service shows the recent week's or month's graphical data for rising/falling of stocks they invested. This kind of service requires more than "summation of (***C*** * ***N***) where ***C*** is from 1 to total number of customers, and ***N*** is a number of different stocks each customer is holding" number of query processing capacity. To be able to process the large number of complex queries, there must be a dynamically

controllable VO manager service that can provide more QP computing power using redundant computing resources within the VO.

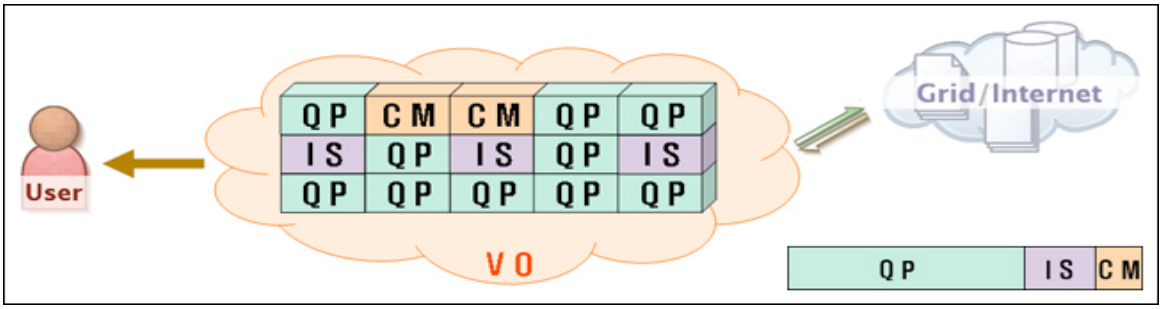

Figure 3 : An Example Model for Query Processor Expansion within a Virtual Organization

2.1.2 Features and Architecture of a Virtual Organization for Dynamic Reconfiguration

In our work, the Grid information retrieval service for a VO is dynamically managed by GIR-VOMS (Grid Information Retrieval – Virtual Organization Management Service). GIR-VOMS can provide dynamic IR system reconfiguration as shown in Fig. 4 when it detects a request for more computing power by certain IR component. It may also expand a GIR system by adding new computing resources to a VO.

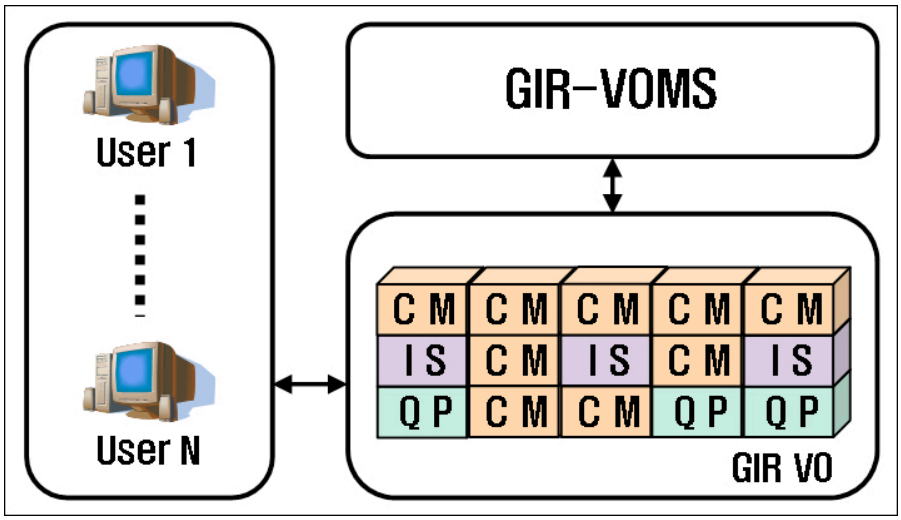

Figure 4 : GIR System Architecture for Dynamically Reconfigurable Virtual Organization

(1) GIR-VOMS Design Considerations

- Management and monitoring of registered resources
  The management and monitoring for each registered service is not required for current conventional IR systems because they have single integrated architecture. However, management and monitoring are required in a GIR system because the IR system components (QP, IS, and CM) are independent, separate services registered to Globus. Allocation of new joined resource to a specific service should be managed, and service role change request should also be monitored.

- System reconfiguration for IR component service change

Relationships among QP, IS, and CM components should be reconfigured when a computing node running a specific service is requested to run a different service. Even though it seems that each service is running independently of each other, there is a "Producer-Consumer" relationship among them.

- Scheduling and reservation of resources
  A specific overloaded node may initiate a service change request to another node where workload is not high. But the request is not granted until the requested node finishes processing of current service. Therefore, there is a need for resource reservation and scheduling functions.

- Prior installation and transfer of processing results
  Intermediate results should be saved on local storage or transferred to a similar nearby service depending on data policy, before a service change request is granted by a component node. Also, installation of the new service is required before a new service is executed, with the current one being ended.

(2) GIR-VOMS System Architecture

As shown in Fig. 5, GIR-VOMS is monitoring a GIR system and has management control over services running in a VO. GIR-VOMS consists of a "registry" for listing resources, "notification" for requesting service change, "security" for authentication and authorization [9][11][12], "events" for reconfiguration of relationship among IS/CM/QP at service change, and "deployment" for installation of services when a new member is added. Hence, when a specific service is requested by GIR-VOMS to a specific member resource, the resource carries out the service change operation as requested, then notifies its executed action back to the VOMS.

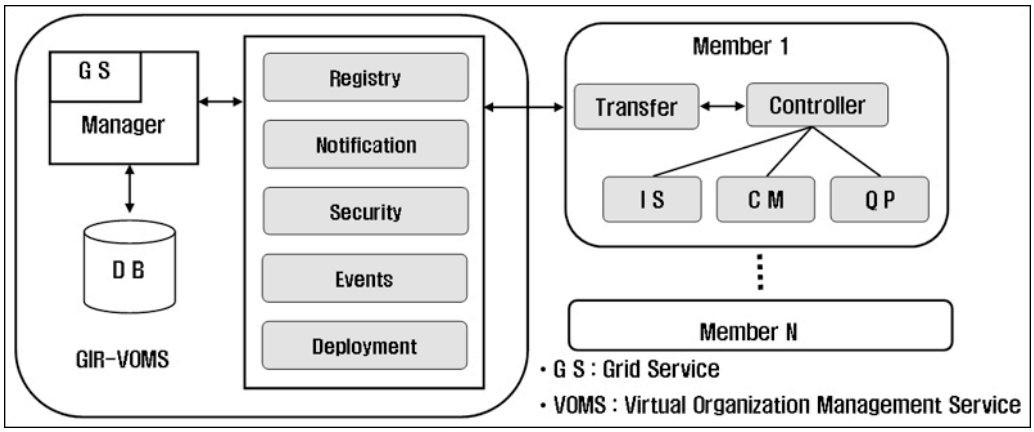

Figure 5 : System Architecture of GIR-VOMS and Members

2.1.3 Implementation and Results

Fig. 6 below shows our VOMS interface tool built on top of a Globus based Grid IR system that can dynamically change IR component role as needed.

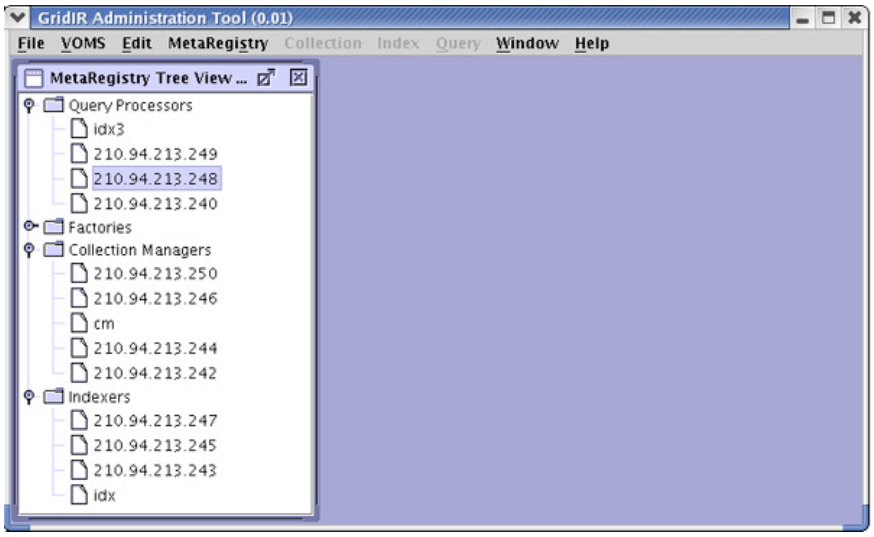

Figure 6 : Grid Information Retrieval Administrator Tool

VOMS can be executed after registering as a Gird service and after creating a service instance [4][5][8]. Then it gathers all the service role information for computing resources currently running and belonging to a VO. For example, the GIR system shown in Fig. 6 has four QPs, five CMs, and four ISs registered as services. In Fig. 6 above, if we click VOMS and then choose 'VOMS Controller' we will be able to see the screen below as in Fig. 7, in which registered services may be changed. The right hand side of the window shows the total number of services, as well as number of each service for CM, IS, and QP. It also shows a resource information list for each service. For example the resource with IP 210.94.213.50 is currently working as a CM. Next, consider what happens when we experience a shortage of computing power for the IS activities, so convert the CM into an IS.

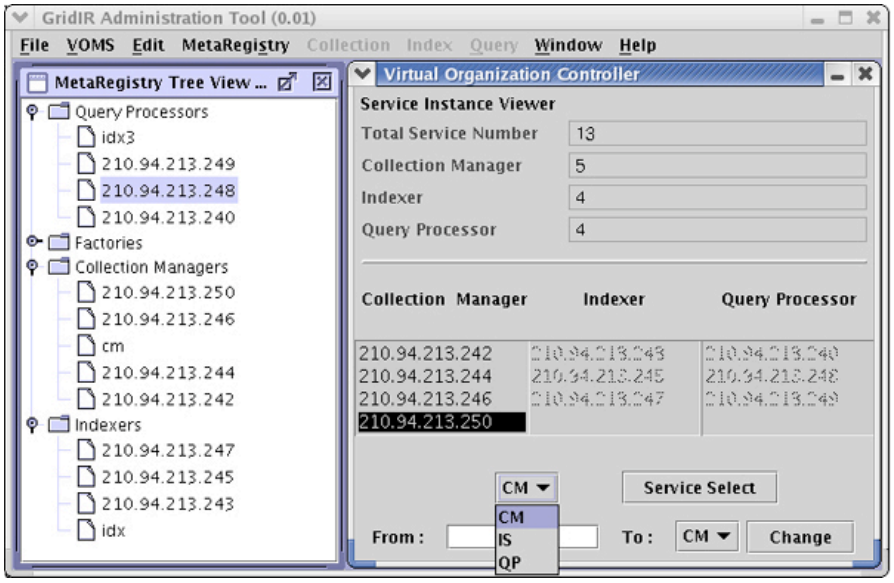

Figure 7 : Virtual Organization Controller showing Service Change from CM to IS

As shown in Fig. 7, we first select service type and a resource from the list, then click 'Service Select' button. Then the resource with IP 210.94.213.250, which is currently working as a CM, stops its service, and the text box called 'From' is selected.

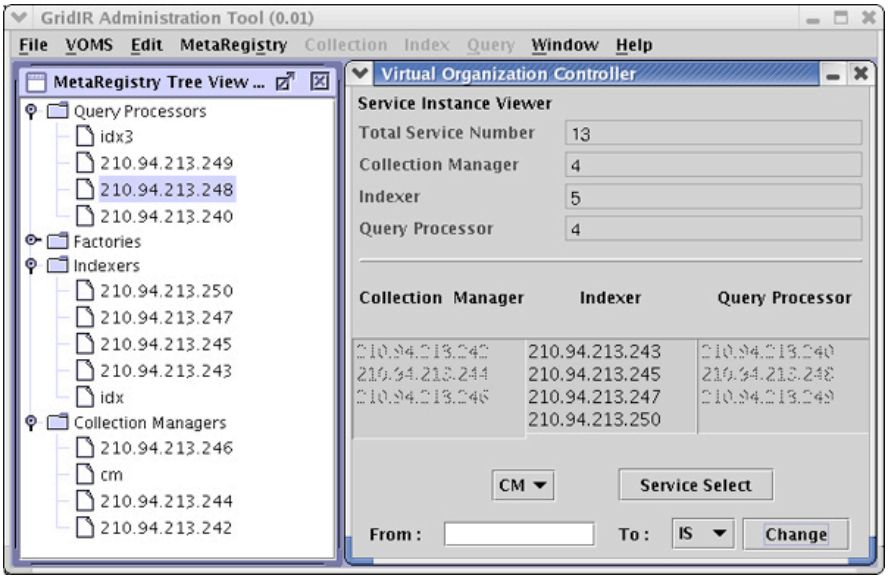

Figure 8 : After the Service Change from CM to IS

Now, we select a desired service, then click 'Change' button. A few moments later, the resource with IP 210.94.213.250 changes from CM to IS, and the 'Meta Registry' list is also updated automatically as in shown Fig. 8. In this way, dynamic reconfiguration of VO is possible converting to a service that requires additional computing power, rather than adding extra resources. At the same time, interrelationships among the CM, IS, QP components are dynamically reconfigured as their service roles are changed.

2.2 Dynamic Virtual Organization with Security Policy

2.2.1 Use Case Two: Grid IR Scenario for Different VO Security Policies

Consider the separate special task centers for terrorist attack prevention at the CIA and FBI. Information sharing between them, as well as with other government offices, involves serious security issues. Additionally, there is a problem to be solved regarding what level of security should be provided between those two centers since their operation areas are different. In Grid IR, this kind of problem can be solved by organizing different VOs with different security levels [1][9][13].

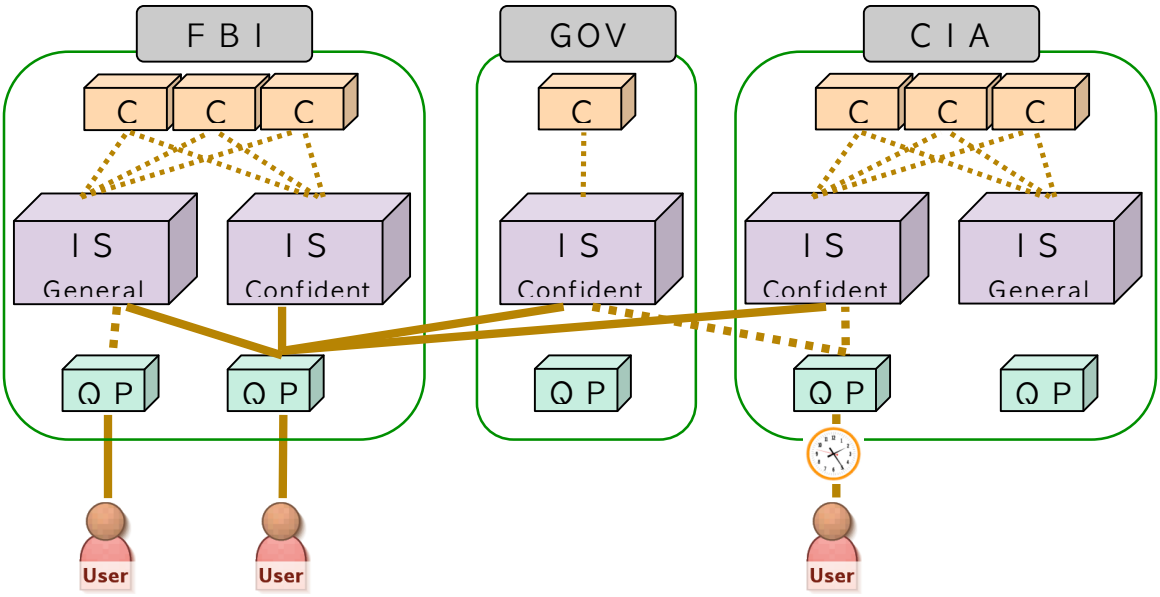

Figure 9 : An Example Model for Grid IR with different VO Security Policies

A primary consideration in this use case is that each organization should be allowed to access only the data permitted by its security level when the organizations are integrated together to form a larger VO. In our simplified model, in order to enforce this, different VOs would be formed, one with higher security level ISs and the other with lower security level ISs. Additional administrative offices may be dynamically added to a VO when their information is needed. In this way, information retrieval with different levels of security and different boundaries are possible depending on which VO a QP is connected to.

Second, a large percentage of information security violations are caused by internal staff members. This kind of information leaking problem is caused by human factors rather than technical factors, so it is almost impossible to prevent the problem entirely. But it can be reduced by using a 'proxy credential' for each VO. In other word, persons with authorized access to a database can be changed as defined by the term of validity to certain resources like QPs, ISs, CMs or VOs.

Third, an agent function for reporting well refined data can be implemented by dynamically configuring VO consisting of ISs for certain specific data.

2.2.2 Features and Architecture of Grid IR with Different VO Security Policies

The security system for Grid IR has been implemented using a 'policy management service' and an 'authentication service'. In other words, users and resources included in a VO are mutually authenticated using the authentication service provided by Grid Security Infrastructure (GSI) in Globus toolkit [9][13][15]. Then, the policy management service that we implemented determines and manages security policy against the authenticated user list.

(1) Design Considerations

- Authentication for client users
  As mentioned in the above scenario, since different security policies are applied to each document, every client has to be authenticated from a proper certification authority (CA). Moreover, the 'distinguished name' (DN) used at the time of authentication is a unique name used in Grid IR system for distinguishing clients. Likewise, a new proxy has to be

issued for each user (e.g., for every 12 hours) in order to use a 'proxy credential' for limited time access operation.

- Authentication and policy service management
  For a secure IR service operation, a proper security level should be applied for each IS service, and the policy has to be applied for each instance. Based on the policies, each client can only perform information retrieval operations against corresponding ISs.

(2) System Architecture

As can be seen in Fig. 10, ISs with different security policies are managed by a 'Policy Manager' (PM). Based on policies applied to each IS, the policy manager makes and keeps an Access Control List (ACL), which is a list of permitted users. If a user client wants to perform an IR service through a certain QP, the ACLs of ISs are checked first to form a VO dynamically, and then the requested IR operation is performed within the VO.

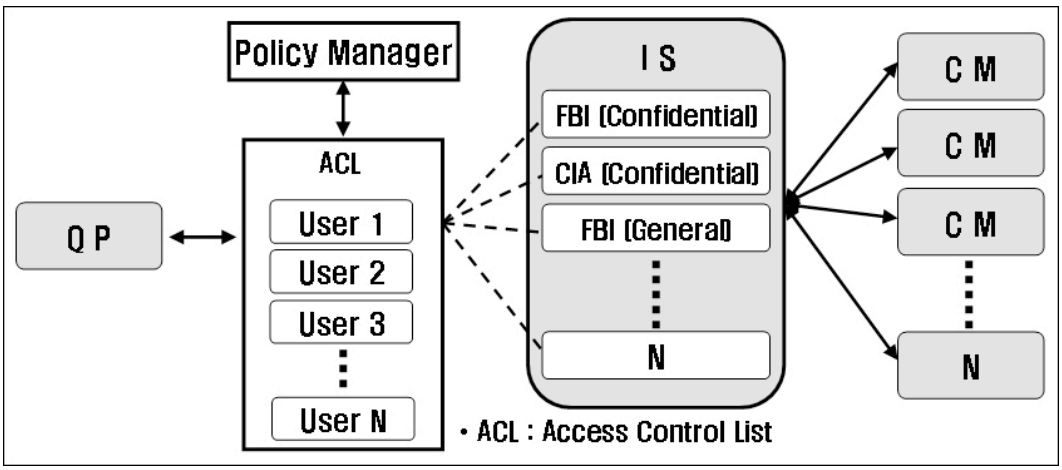

Figure 10 : Architecture of the Proposed System

(3) Creation of Security Policy

We should be able to use a Policy Manager (PM) for each IS in order to apply different security policies to each IS. Moreover, in order to be able to use PM, the PM as well as authentication service should be executed within a Globus container so that IS service can call and use the PM. This service can be generated through steps such as these.

- Generation of Proxy

  General proxy used for communication between services and clients has 12 hours of life time and it is generated using the instruction below [9][11][12].

  % grid-proxy-init

- Generation of Security Service

  Authentication service and PM can be generated in the Globus container using the instructions below.

  % java org.globus.ogsa.client.CreateService –gsiSecConv sig –auto none
  http://localhost:8080/ogsa/services/appSec/PolicyManager pm

% java org.globus.ogsa.client.CreateService –gsiSecConv sig –auto none
http://localhost:8080/ogsa/services/appSec/AuthorizationService authz

- Conferring Security Service Policy

A GUI environment is added for generation of security policy as well as XACML policy for deeper level of authentication. 'PolicyGUI' makes a file for access control and initializes 'PolicyManagerService' using information regarding policy and services together. 'PolicyGUI' can be executed as follows.

% java org.cnidr.ogsa.appSec.policyManager.PolicyGUI
http://localhost:8080/ogsa/services/gridir/MetaRegistry/mr
http://localhost:8080/ogsa/services/appSec/PolicyManager/pm
$GRIDIR_LOCATION/src/schema/metaRegistry/metaRegistry.gwsdl

Executing the above instructions, a GUI window like one shown in Fig. 11 will pop up.

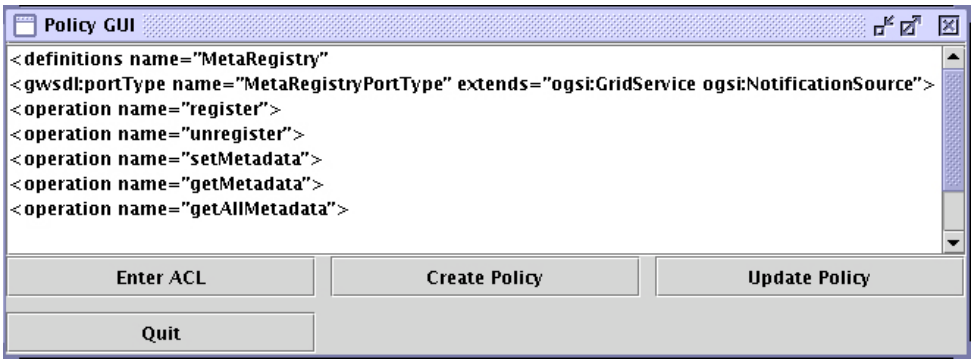

Figure 11 : Policy Administrator Tool

Next, select a policy service from service list, and press 'Enter ACL'. Then a text box will pop up for input of service type, port, and DN (Distributed Name) who are permitted for access as shown below in Fig. 12.

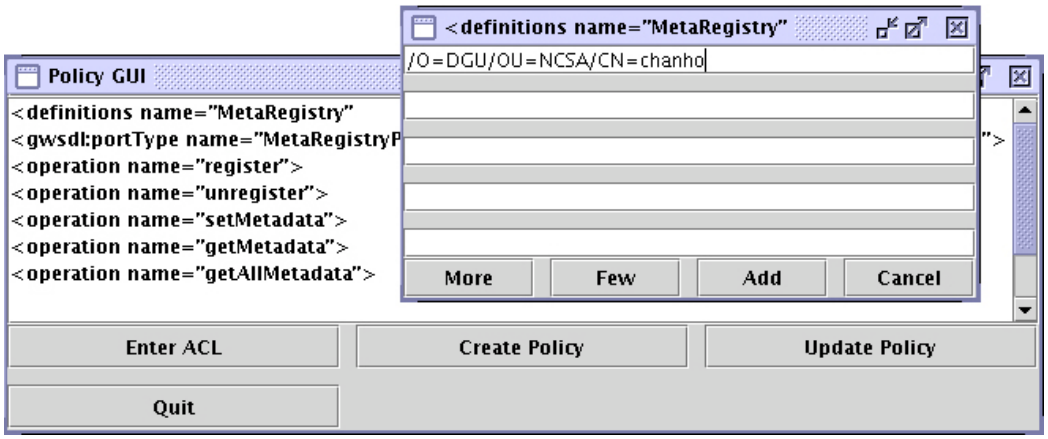

Figure 72 : Access Control List Addition

This process will only be allowed for users who already have certification for the given service. After inputting the ACL press 'Add' button, then press 'Create Policy' button for generation of policy. This will result in the creation of policy file, and it will notify the

'Policy Manager Service' that the new generated file is a policy for handling of a given service. An update of a current policy to a new one can also be done by pressing 'Update Policy' button. In addition, this process has to be applied to all ISs that security mechanism is to be used.

- Generation of Instances for Grid Information Retrieval Services

  Finally, instances for GIR components such as QP, IS, CM are generated, and a 'MetaRegistry' that dynamically control those instances is also generated at Globus container.

2.2.3 Implementation and Results

There are two ISs with different security policies. One of them contains indexed data about the FBI with security policy applied; the other contains indexed data for a movie about the FBI without any security policy applied. Therefore, an IR operation will be processed against both indexes for an authorized user who meets the security policy, and against only the index that contains movie information in the case of users who are not authorized for the other data.

As shown in Fig. 13, we run two ISs in a GIR system, one with security policy applied and the other being a default IS. Both ISs are using one CM and one QP, Then we executed a query such as 'any=FBI' twice, one for secured user and the other for ordinary user.

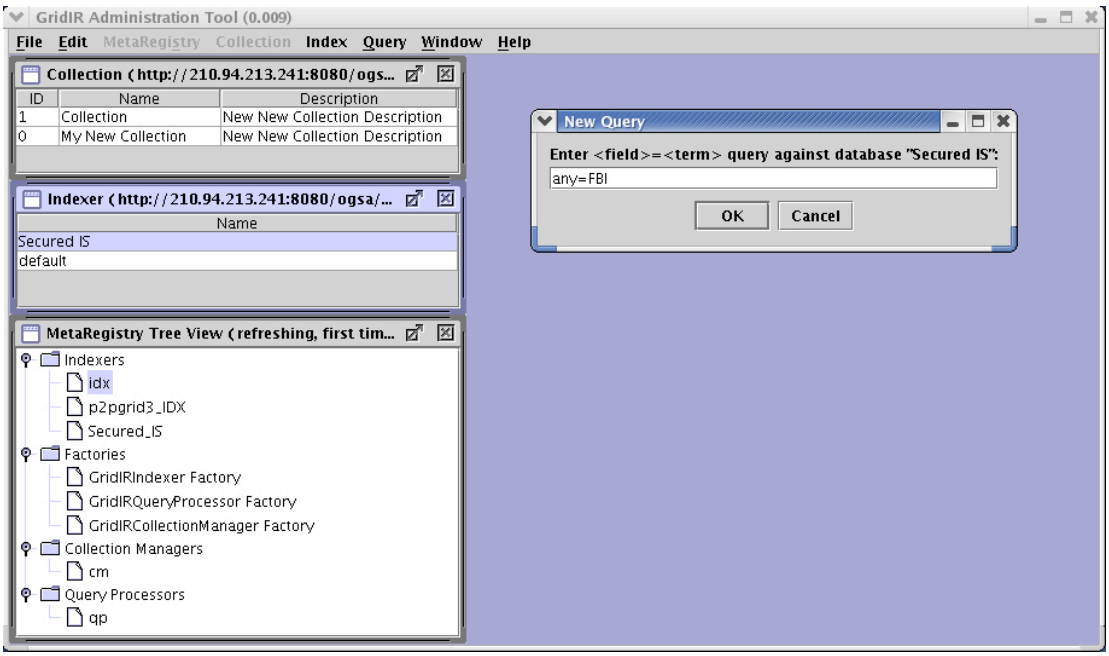

Figure 13 : Generation of New Query

For the ordinary user without a security policy, there will be only one result coming from the IS containing movie data as in Fig. 14. But in case of secured user, there will be two results, one about the movie, and the other containing an FBI introduction document as in Fig. 15. In both cases, the relevance score is determined depending on the number of occurrence of the query term "FBI."

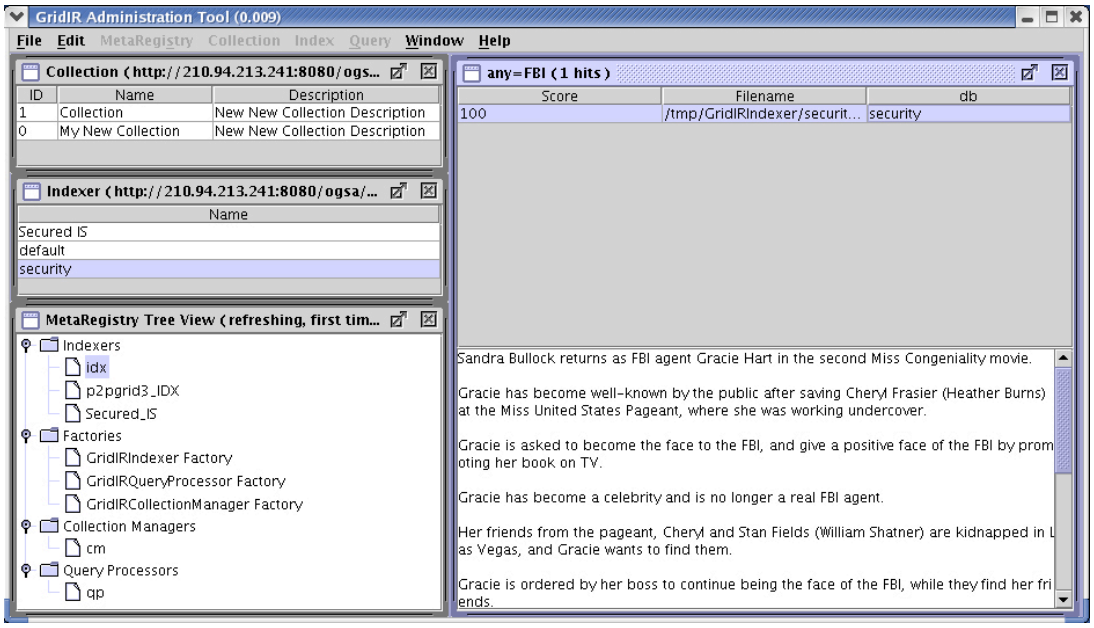

Figure 14 : Retrieval Result for an Ordinary User without a Security Policy

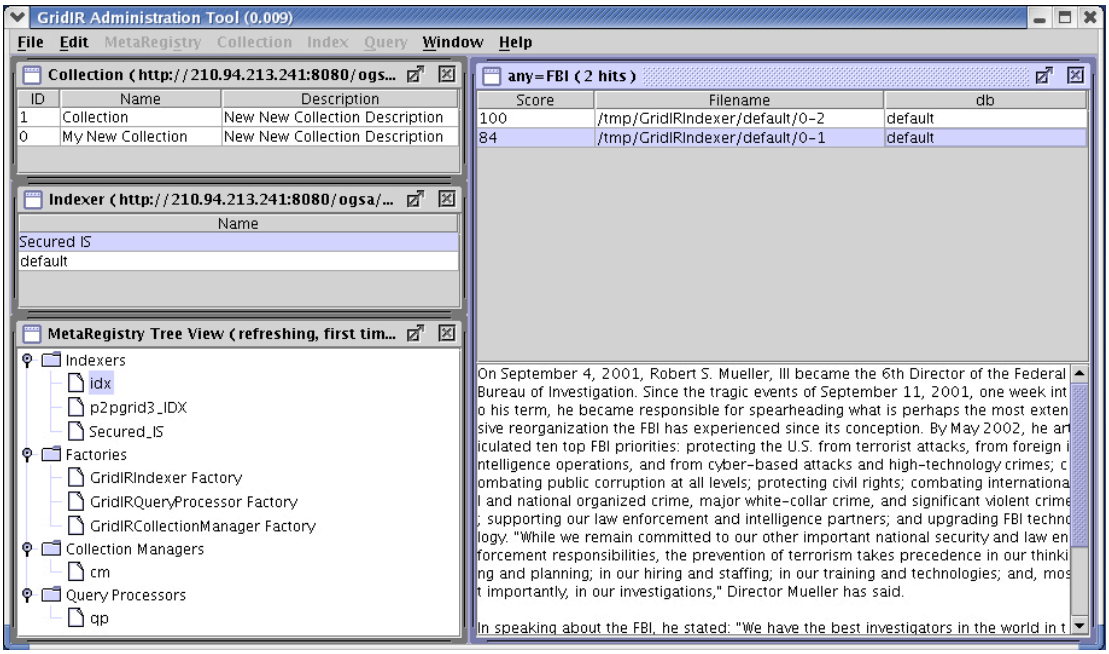

Figure 15 : Retrieval Result for a Secured User with a Security Policy

3. Conclusions

In this document, we implemented a new Grid IR system environment by adding dynamic VO and security mechanisms, which are known to be two major characteristics of Grid computing. First, we made two use-case scenarios showing how that kind of GIR system can be useful. Then we

implemented and added the two mechanisms to a GIR system prototype built on Globus3.0. This GIR system has several additional advantages, but there are two major advantages discussed below.

The first major advantage is that the overall cost of ownership for internal IR system will decrease for corporations or government organizations. Instead of buying additional expensive servers or clusters as capacity need grows, existing redundant systems can be utilized and dynamically added to a VO for more power using Grid distributed computing technology. Moreover, unlike other conventional IR systems where dynamic overload for a specific service may cause system outages, a GIR system can employ dynamic load balancing by using low utilized computing resources or by adding additional existing resources to a VO.

The second major advantage is that by solving security problems that arise when IR systems with different security policies are merged into a VO, security documents or information can be more efficiently managed. This implies that corporations or government organizations can share their secured information with other cooperating companies or other government organizations applying multilevel security policies, instead relying on their Intranet access controls or other coarse-grained methods.

Contributor Information

Yangwoo Kim
Dongguk University
Korea
ywkim@dongguk.edu

Pilwoo Lee
KISTI, Korea Institute of Science and Technology Information
Korea
pwlee@kisti.re.kr

Gregory B. Newby
Arctic Region Supercomputing Center, University of Alaska Fairbanks
United States
newby@arsc.edu

Intellectual Property Statement

Full Copyright Notice

References

1. Ian Foster, Carl Kesselman, Steven Tuecke. 1998. "*Security Architecture for Computational Grids."* 5th ACM Conference on Computer and Communications Security Conference
2. Dr. Bernhard, R. Katzy. 1998. "*Design and Implementation of Virtual Organizations*.", Proceeding of 31th International Conf. on System Science, vol. 4, pp. 142 – 151, IEEE.
3. Ian foster, Carl Kesselman, Steven Tuecke. 2001. "*The Anatomy of the Grid: Enable Scalable Virtual Organizations*." International Supercomputer Applications.
4. Carl Kesselman, Jeffrey M Nick, Steven Tuecke, Ian Foster. 2001. *"Grid Service for Distributed System Integrations." IEEE Trans. on Computer, vol. 35, Issue 6, pp 37 – 46*
5. Ian Foster, Carl Kesselman, Jeffrey M. Nick, Steven Tuecke. 2002. "*The Physiology of the Grid: An Open Grid Service Architecture for Distributed System Integration." GGF* Open Grid Service Infrastructure Working Group.
6. Nassar, Newby, Gamiel, Dovey, Morris. 2003. "*Grid Information Retrieval Architecture."* GGF Grid Information Retrieval Working Group.
7. Gamiel, Kevin; Newby, Gregory B. & Nassar, Nassib. 2003. "*Grid Information Retrieval Requirements* (GFD.27)." Lamont, Illinois: Global Grid Forum.
8. Radu Prodan, Thomas Fahringer. 2003. "*From Web Service to OGSA: Experiences in Implementing an OGSA-based Grid Application*." Proceeding of the Fourth International Workshop on Grid Computing (GRID'03) IEEE.
9. Von Welch, Ian Foster, Carl Kesselman, Olle Mulmo, Laura Pearlman, Steven Tueke, Jarek Gawor, Sam Meder, Frank Siebenlist. 2004. "*X.509 Proxy Certificates for Dynamic Delegation*." 3rd Annual PKI R&D Workshop.
10. S. Cannon, S. Chan, D. Olson, C. Tull, V. Welch, L. Pearlman. Mar. 2003. "*Using CAS to Manage Role-Based VO Sub-Groups*." Proc. Computing in High Energy Physics 03 (CHEP '03).
11. L. Pearlman, C. Kesselman, V. Welch, I. Foster, S. Tuecke. Mar. 2003. "*The Community Authorization Service: Status and Future*." Proc. Computing in High Energy Physics 03 (CHEP '03)
12. I. Foster, N. T. Karonis, C. Kesselman, S. Tuecke. 1998. "*Managing Security in High-Performance Distributed Computing*." Cluster Computing, 1(1):95-107.
13. The Grid Security Team, Globus Toolkit Version 4 Grid Security Infrastructure: A Standards Perspective, Sep. 2005. (see http://www.globus.org/toolkit/docs/4.0)
14. The Java CoG Kit 4.1.3 Release. Argonne National Laboratory. (see http://www.cogkit.org/user)
15. The Globus Toolkit, The Globus alliance. (see http://www.globus.org)
16. The Globus Toolkit 3 Programmer's tutorial. The Globus alliance. (see http://www.casa-sotomayor.net/gt3-tutorial)