

# 그리드 기반 정보검색 서비스를 위한 동적 가상 조직 관리 시스템

김 양 우<sup>†</sup> · 이 승 하<sup>\*\*</sup> · 김 혁 호<sup>\*\*</sup>

## 요 약

그리드 컴퓨팅 기반의 모든 응용 작업은 두 개의 요구사항을 필히 만족해야 하는데 그 중 하나는 가상조직의 동적 변화를 수용할 수 있어야 한다는 것이고, 또 다른 하나는 서로 다른 가상조직 사이에 서로 다른 보안 단계와 정책을 적용할 수 있어야 한다는 것이다. 본 논문에서는 우선 위의 요구사항들에 대한 두 개의 가상 시나리오를 통하여 그 중요성을 부각시키는 동시에 그리드 기반의 정보검색 시스템 시제품 개발을 통하여 그 요구사항들이 만족될 수 있음을 보여주었다. 가상조직의 동적 변화특성은 단지 그 조직에 속한 사용자들의 증가와 감소뿐만이 아니라, 전체 시스템을 구성하고 있는 서브시스템들 간에 컴퓨팅 파워에 대한 동적인 변화 요구에도 적용된다. 이는 어느 한 서브시스템으로부터의 컴퓨팅 파워 증가 요청을 시스템 내의 유휴 자원인 다른 서브시스템이 만족시켜 줄 수 있기 때문에 전체적인 시스템 상의 유연성이 유지된다는 장점을 갖는다.

본 논문에서는 글로버스 3.0에서 제공되는 보안기능과 가상조직을 이용하여 그리드 기반 정보검색 시스템을 구현하였고, 동적 가상조직을 위한 본 시스템의 확장성과 보안성이 어떻게 향상될 수 있는지를 보여 주었다. 이와 더불어 본 논문에서는 서로 다른 가상조직들을 관리하기 위한 가상조직 관리서비스를 구현하여 글로버스에 추가 서비스로 등록하였다.

키워드 : 그리드 컴퓨팅, 정보검색, 미들웨어, 검색엔진, 가상조직

## Dynamic Virtual Organization Management System for Grid Based Information Retrieval Service

Yang-Woo P. Kim<sup>†</sup> · Seung-Ha Lee<sup>\*\*</sup> · Hyuk-Ho Kim<sup>\*\*</sup>

## 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 paper, 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 paper 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.

Key Words : GRID Computing, Information Retrieval, Middleware, Search Engine, Virtual Organization

## 1. 서 론

그리드 정보검색[7]이란 그리드 기술에 기반을 둔 분산 정보검색 시스템을 말한다. 즉, 클러스터 구조로 전용의 컴퓨팅 자원들을 가지고 있는 기존의 전통적인 정보검색 시스

템들과는 달리, 그리드 정보검색 시스템은 한 가상조직 (Virtual Organization)내의 공유 가능한 유휴 자원들을 활용하여 전체 시스템을 동적으로 재구성 할 수 있다. 이와 더불어 글로버스[14]와 같은 그리드 미들웨어는 각 가상조직에 따라 정교한 다단계 보안 기능을 제공한다.

정보검색(Information Retrieval)은 인간이 필요로 하는 정보에 맞는 문서 또는 요약된 문서를 찾아주는 방법이라고 일반적으로 정의한다. 이러한 정보요구는 정보검색 시스템

<sup>†</sup> 종신회원 : 동국대학교 정보통신공학과 부교수 (주저자 및 서신저자)  
<sup>\*\*</sup> 준 회 원 : 동국대학교 정보통신공학과 박사과정  
 논문접수 : 2006년 8월 28일, 심사완료 : 2006년 11월 7일

이 가지고 있는 문서들을 대상으로 검색을 실행하는 질의어 형태로 표현된다. 현재 가장 잘 알려진 정보검색 시스템으로는 인터넷 검색엔진들[6][9]을 들 수 있으나 그 외, 예전부터 논문이나 도서를 검색해주는 온라인 도서관 정보검색 시스템이나 서지 정보검색 시스템들이 있고, 기업에서는 정보검색 시스템의 일종인 의사결정 시스템, 그리고 경영정보 시스템 등이 있다.

차세대 인터넷 인프라가 될 것으로 기대되는 그리드 컴퓨팅은 Global Grid Forum (<http://www.ggf.org>)에서 소개되었듯이 많은 과학계산용 응용에 성공적으로 접목되어 왔다. 그러나 본 논문에서는 과학계산 응용이 아닌 상용 응용이라 할 수 있는 정보검색 응용을 그리드 기술에 접목시켜 보았다. 상용 응용에 그리드 기술을 적용하기 위해서는 두 개의 중요한 요구사항들을 만족시켜야 하는데, 그 중 하나는 가상조직의 동적인 특성을 수용해야 한다는 것이고, 또 다른 하나는 서로 다른 가상 조직들 간에 서로 다른 보안 정책을 적용할 수 있어야 한다는 것이다.

가상조직(Virtual Organization, VO)이란 지역적으로 분산되어 있는 컴퓨팅 자원과 사용자들을 가상으로 네트워크를 통해 묶어 놓은 기술적 개념이다. 이러한 가상조직의 동적인 특성이란 사용자의 증가 또는 감소는 물론, 전체 시스템을 구성하고 있는 서로 다른 서브시스템들 간의 동적으로 변하는 컴퓨팅 파워에 대한 변경 요구를 포함한다[2][4]. 이는 전체적인 시스템 상의 유연성이라는 장점을 살려서 어느 한 서브시스템으로부터의 컴퓨팅 파워 증가 요청을 시스템 내의 유휴 자원인 다른 서브시스템이 동적으로 만족시켜 줄 수 있어야 한다는 것이다. 이러한 'On-Demand' 컴퓨팅 개념은 동적인 컴퓨팅 파워 변환을 위한 시스템 동적 재구성이 가능한 정보검색 시스템에 적용될 수 있다. 이는 동적인 확장성과 다단계 보안 적용이 가능한 정보검색 시스템을 위한 새롭고 유연한 컴퓨팅 패러다임을 제공할 수도 있음을 의미한다.

본 논문과 더불어 구현된 그리드 정보검색 시스템 시제품은 부분적으로 그 동안 저자들의 GGF(Global Grid Forum) GIR(Grid Information Retrieval) 워킹 그룹에서의 활동을 기반으로 한 것이다.

2. 동적인 가상조직을 위한 그리드 정보검색 시스템

본 논문에서는 동적 가상 조직 개념과 유연성 있는 보안 정책을 도입하여 새로운 정보검색 서비스 모델을 제시하였다. 우선, 두 개의 활용 케이스와 구현된 시스템을 통하여 이러한 시스템 모델의 유용성을 보여 줄 수 있도록 하였다.

2.1 활용 케이스 1: 가상 조직의 동적 재구성

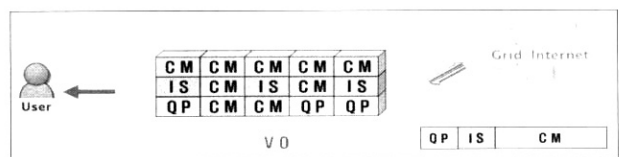
한 증권회사는 정보검색 시스템의 성능과 용량을 향상시키기 위하여 새로운 정보검색 시스템을 도입하기로 결정하였다. 그들은 유휴 컴퓨팅 자원들을 포함시키도록 가상조직을 동적으로 확장시키면 전체 컴퓨팅 시스템의 용량 증가와 시

스템 재구성이 동적으로 가능한 그리드 기술을 수용하는 새로운 정보검색 시스템을 도입하기로 하였다. 본 활용 케이스에서는 시스템 운영 상의 서로 다른 세 단계를 고려하였다.

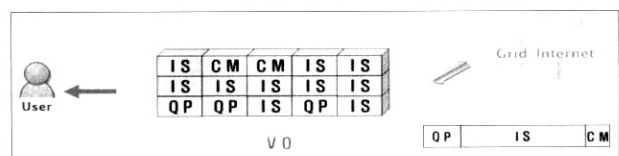
추가로 본 논문에서 사용하는 그리드 정보검색 시스템 모델은 컬렉션 관리자(CM)와 질의어 처리기(QP), 그리고 인덱스/검색기(IS) 등의 주요 세 요소로 구성되는 GGF의 GIR 요구사항 문서[7]를 기반으로 하였다. 이러한 세 부분으로 구성되는 GIR 모델에서 CM은 문서들에 대한 필요한 사용 권한, 변환, 임시 저장 등을 포함하는 문서 수집과 관리 기능을 주로 수행하고, QP는 사용자 질의어를 IS에게 전달하는 작업과 결과의 취합 및 포장, 그리고 연속적인 추가 정보검색이 필요한 장기 실행 질의어(long-running query)등의 관리 기능들을 수행한다.

첫째, 이 증권 회사의 정보검색 시스템은 매일 아침 증권 시장의 개장 전까지 증권시장의 상승과 하락에 영향을 줄 수 있는 모든 정보를 수집할 필요성이 있다. 이러한 정보 수집 작업은 증권 시장이 개장하기 전까지인 밤 사이에 이루어지며, 이 작업은 CM에 많은 컴퓨팅 파워를 요구한다. 그러나 같은 시간대에 QP와 IS는 할 작업이 없기 때문에 이들이 가지고 있는 컴퓨팅 파워는 대부분 남아 돌게 된다. (그림 1)에서 보듯이 전체적인 시스템의 이용률을 높이기 위해서는 이러한 유휴 컴퓨팅 자원들을 CM으로 재할당할 수 있어야 한다.

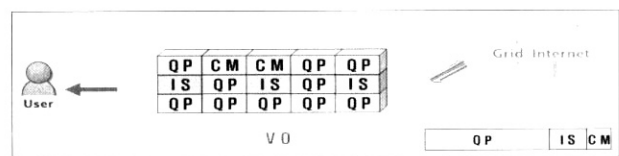
둘째, 이 회사의 정보검색 시스템은 그들의 고객이 투자를 하기 전에 증권 시장에 영향을 줄 수 있는 정치적, 사회적, 경제적, 그리고 국제 정세 등을 분석할 필요성이 있다. 이러한 유형의 서비스는 서로 다른 유형과 구조의 많은 문서들을 통합하기 위한 데이터 마이닝 기술을 필요로 한다. 그래서 (그림 2)에서 보듯이 지리적으로 분산되어 있는 데이터를 통합하고 실시간 분석 서비스를 제공하기 위해서는 IS에 높은 컴퓨팅 파워가 요구된다.



(그림 1) 한 가상조직 내에서 CM 확장이 필요한 모델의 한 예



(그림 2) 한 가상조직 내에서 IS 확장이 필요한 모델의 한 예

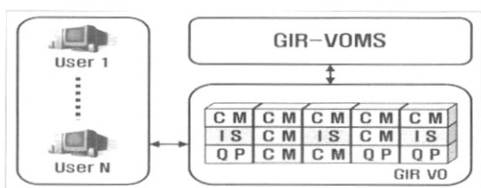


(그림 3) 한 가상조직 내에서 QP 확장이 필요한 모델의 한 예

셋째, 이 증권 투자 회사는 각 고객들에게 그들의 투자에 대한 분석 서비스를 고객들의 PDA를 통하여 제공하려 한다. 이 서비스는 고객들의 증권 투자에 대한 일주일 또는 한 달 간의 상승과 하락 폭에 대한 그래픽 데이터를 보여주는 것으로 (C x N)을 C는 1부터 전체 고객 수까지, 그리고 N은 각 고객이 투자한 종목의 수로 바꾸어 가며 합산한 개수만큼의 질의어 처리 능력을 요구한다. 이러한 많은 수의 복잡한 질의어를 처리하기 위해서는 한 가상조직 내에서 유희 한 컴퓨팅 파워를 찾아 더 많은 QP 컴퓨팅 파워를 제공해 줄 수 있는 동적 제어가 가능한 가상조직 관리 서비스가 있어야 한다.

2.1.1 동적 재구성을 위한 가상조직의 기능과 구조

본 연구에서 가상조직을 위한 그리드 정보검색 서비스는 그리드 정보검색 가상조직 관리 서비스 (GIR-VOMS, Grid Information Retrieval - Virtual Organization Management Service)에 의해서 동적으로 관리된다. GIR-VOMS는 정보검색 시스템 내 한 컴포넌트의 컴퓨팅 파워 증가요청을 감지하여 (그림 4)에서 보여 주듯이 동적 IR 시스템 재구성을 지원할 수 있다. 또한 새로운 컴퓨팅 자원들을 한 가상조직에 추가 함으로써 그리드 정보검색 시스템을 확장할 수도 있다.



(그림 4) 동적 재구성이 가능한 가상조직을 위한 그리드 정보검색 시스템 구성도

(1) GIR-VOMS 설계 고려사항

◦ 등록된 자원들의 관리와 감시

등록된 각 서비스들에 대한 관리와 감시 기능은 하나의 통합된 구조를 갖는 전통적인 정보검색 시스템에서는 요구되지 않지만, 각 컴포넌트들이 글로벌스에 등록된 독립적인 별개의 서비스(QP, IS, 그리고 CM)인 그리드 정보검색 시스템에서는 필요한 기능들이다. 새로이 추가된 자원을 특정 서비스에 할당하는 기능은 관리되어야 하며, 서비스 역할변경 요청 또한 항상 감시되어야 한다.

◦ IR 컴포넌트서비스 변환을 위한 시스템 재구성

특정 서비스를 수행하는 한 컴퓨팅 노드가 다른 서비스를 수행하도록 요구된다면 QP와 IS 그리고 CM간의 관계는 다시 재구성되어야 한다. 비록 각 서비스는 독립적으로 수행되는 것처럼 보일지라도 그들 사이에는 "생산자와 소비자"의 관계가 존재한다.

◦ 자원들의 예약과 스케줄링

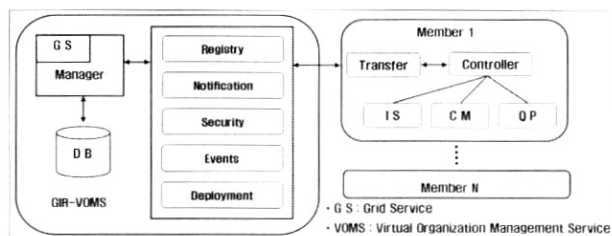
작업량이 많은 노드는 작업량이 많지 않은 다른 노드에게 서비스 변환을 요청할 수 있다. 그러나 그 요청은 요청을 받은 노드가 현재 수행중인 작업을 끝내기 전에는 허락될 수 없다. 그렇기 때문에 자원 예약과 스케줄링 기능이 필요하다.

◦ 사전 설치와 처리 결과의 전송

서비스 변환요청이 허가되기 전에 그 동안 처리된 결과는 데이터 저장 정책에 따라 로컬 저장소에 저장되거나 유사한 기능의 주변 서비스로 전송되어야 한다. 또한 한 서비스를 종료하고 새로운 서비스를 수행하기 전에 서비스의 사전 설치가 이루어져 있어야 한다.

(2) GIR-VOMS 시스템 구조

(그림 5)에서 보듯이 GIR-VOMS는 GIR 시스템을 모니터링하며 그 가상조직에서 실행되고 있는 서비스들에 대한 관리 제어 기능을 갖는다. GIR-VOMS는 자원들의 리스트를 보여 주는 "registry"와 서비스 변환을 요청하는 "notification", 인증과 권한 부여를 위한 "security", [11][12], 서비스 변환시 IS/CM/QP 사이의 관계를 재구성하는 "events", 그리고 새로운 멤버가 추가되었을 때 서비스들을 설치하는 "deployment" 기능들로 구성된다. 따라서 GIR-VOMS에 의해 특정 멤버 자원에 특정 서비스가 요청된다면 그 자원은 요청된 서비스 변환 작업을 수행하고 그 결과를 VOMS에게 통지한다.

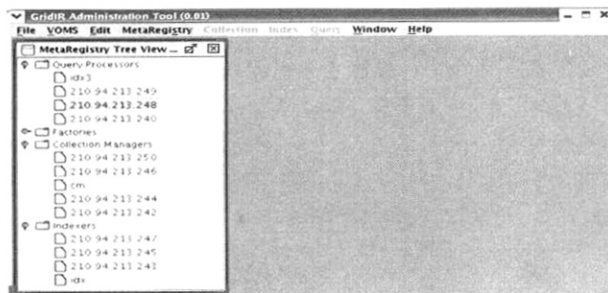


(그림 5) GIR-VOMS와 멤버들의 시스템 구조

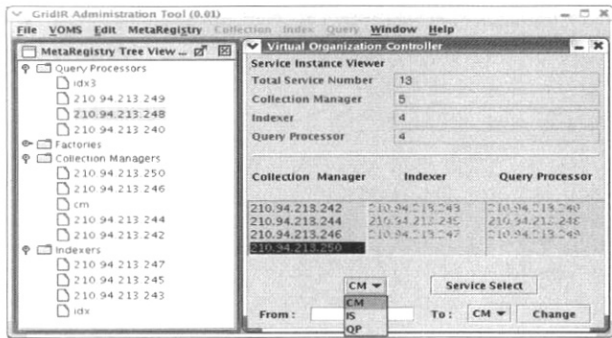
2.1.2 구현과 결과

아래의 (그림 6)은 필요에 따라 정보검색 컴포넌트 역할을 동적으로 변화시킬 수 있는 글로벌스 기반의 그리드 정보검색 시스템 위에서 동작하는 VOMS 인터페이스 툴을 보여주고 있다.

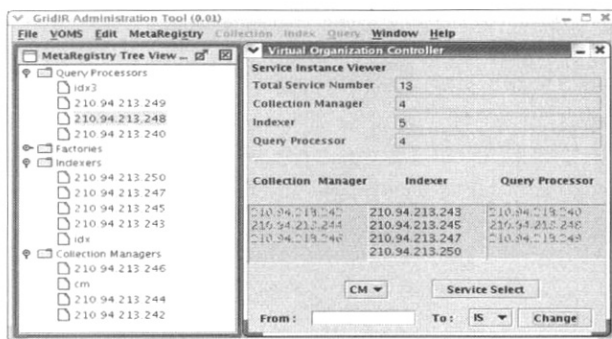
VOMS는 그리드 서비스로 등록된 후, 그리고 서비스 인스턴스가 생성된 후 수행될 수 있다 [4][5][8]. 실행되면 맨 처음 가상조직에 속한 현재 실행되고 있는 컴퓨팅 자원들의 모든 서비스 역할 정보를 수집한다. 한 예로, (그림 6)에 보여진 GIR 시스템은 4개의 QP들과 5개의 CM들, 그리고 4개의 IS로 등록된 서비스들을 가지고 있다. (그림 6)에서 VOMS를 클릭하고 'VOMS Controller'를 선택하면 등록된 서비스들을 변환할 수 있는 아래의 (그림 7)과 같은 창을 볼 수 있다.



(그림 6) 그리드 정보검색 관리 툴



(그림 7) CM을 IS로 서비스 변환시키는 가상조직 제어자



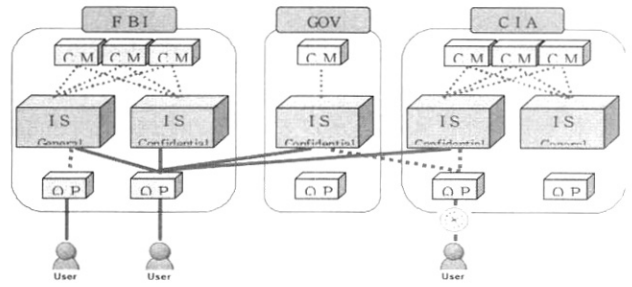
(그림 8) CM을 IS로 서비스 변환시킨 후

그림의 우측 편은 전체 서비스 개수와 CM, IS, 그리고 QP로 동작하는 각 서비스의 개수를 보여준다. 이와 더불어 각 서비스로 동작하는 자원의 정보를 나열한다. 한 예로 IP 값 210.94.213.250을 갖는 자원은 현재 CM으로 동작하고 있으나 만약 IS의 컴퓨팅 파워가 부족하다면 그 CM을 IS로 변환시킬 수 있다.

(그림 7)에서 보듯이, 먼저 리스트로부터 서비스 유형과 자원을 선택하고, 'Service Select' 버튼을 누르면 현재 CM으로 동작하고 있는 IP 210.94.213.250 자원은 그 서비스를 중단한다. 그 후 'From' 텍스트 박스를 선택하고 원하는 서비스를 선택한 후 'Change' 버튼을 누르면 잠시 후 CM으로 동작하던 자원은 IS로 동작하도록 변하며, 'Meta Registry' 리스트도 자동적으로 갱신된다. 이렇게 새로운 여분의 자원을 추가하는 대신에 가상조직의 동적 재구성을 통해 기존의 유휴 자원을 더 많은 컴퓨팅 파워를 요구하는 서비스로 변환시킬 수 있다. 이와 동시에 CM, IS, QP들 사이의 서비스 역할이 바뀌었기 때문에 그들 사이의 상호 관계도 동적으로 재구성된다.

## 2.2 활용 케이스 2: 보안 정책을 갖는 동적 가상 조직

CIA와 FBI에 테러방지를 위한 센터가 각각 있다고 가정하자. 이 경우 이들 사이 또는 다른 정부 조직간의 정보공유에는 심각한 보안 상의 이슈들이 관련 될 수 있다. 또한 이들의 활동 영역이 다르기 때문에 어느 수준의 보안이 제공되어야 하는지의 문제도 해결되어야 한다. 그리드 정보검색 시스템에서는 서로 다른 가상조직에 서로 다른 보안 단계를 적용함으로써 이러한 문제를 해결할 수 있다 [1][12][13].



(그림 9) 서로 다른 가상조직 보안정책을 갖는 그리드 정보검색 모델의 한 예

이 활용 케이스의 주된 고려 사항은 이들 조직들의 정보 시스템들이 통합되어 구성하게 되었을 때 각 기관은 자신의 보안 단계에 따라 허용된 데이터에만 접근 할 수 있어야 한다는 것이다. 이러한 보안 적용을 위해 우리의 단순한 모델에서는 높은 보안 단계를 갖는 IS들을 묶는 가상조직과 낮은 보안 단계를 갖는 IS들을 묶는 또 다른 가상조직을 각각 별도로 구성하도록 하였다. 이 경우 추가적인 행정기관들도 그들의 정보가 필요한 경우 동적으로 어느 한 가상조직에 추가될 수도 있다. 이러한 방식으로 어느 가상조직에 QP가 연결되느냐에 따라 서로 다른 단계의 보안과 경계를 갖는 정보검색이 가능해진다.

두 번째로, 정보보안 위반의 많은 경우가 내부인에 의한 것으로 알려져 있다. 이러한 유형의 정보 유출 문제는 기술적인 요소보다는 인간적인 요소에 기인하는 것으로 완전히 막는다는 것은 불가능하나, 각 가상조직에 '신용 위임장'(proxy credential)을 이용하면 감소시킬 수 있다. 즉, QP, IS, CM, 또는 VO와 같은 특정 자원에 유효 기간을 설정하여 데이터베이스에 접근 할 수 있는 사람들을 시간적으로 제한할 수 있다.

세 번째로, 특정 데이터를 관리하는 IS들로 구성된 동적 가상조직을 이용하여 보호된 데이터에 대한 접근 활동을 미리 정의된 형태로 보고하는 에이전트 기능을 구현할 수도 있다.

### 2.2.1 서로 다른 보안 정책을 갖는 그리드 정보검색

#### 시스템의 특징과 구조

본 논문에서 소개하는 그리드 정보검색의 보안 시스템은 '정책 관리 서비스'와 '인증 서비스'를 이용하여 구현하였다. 즉, 한 가상조직에 포함되는 사용자들과 자원들은 글로벌스 [14]의 그리드 보안 구조(Grid Security Infrastructure, GSI)에서 제공하는 인증 서비스를 이용하여 일단 상호 인증되며, 우리가 개발한 정책 관리 서비스가 인증된 사용자 목록을 대상으로 보안 정책을 관리하고 결정한다.

#### (1) 설계 고려사항

##### ◦ 클라이언트 사용자들의 인증

이미 위의 사용 케이스에서 언급하였듯이 각 문서에 서로 다른 보안 정책이 적용되기 때문에 모든 사용자는 특정 인증원으로부터 인증을 받아야 한다. 이와 더불어 인증서 사용된 구별이름(distinguished name, DN)은 그리드 정보검색

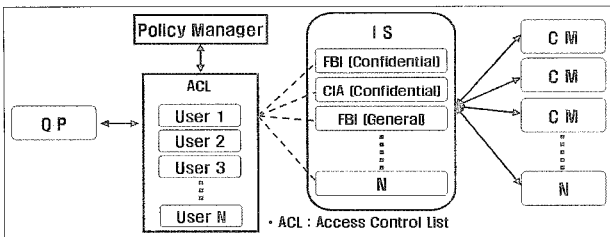
시스템에서 사용자들을 구별하기 위한 유일한 이름이어야 한다. 마찬가지로 제한된 시간 동안의 접근 제어 기능을 위한 '신용 위임장'(proxy credential)을 사용하기 위해서는 각 사용자에게 새로운 프록시(예, 매 12시간 마다)가 발급되어야 한다.

• 인증과 정책 서비스 관리

안전한 정보 검색 서비스를 위해서는 각 서비스마다 적절한 보안 단계가 적용되어야 하고, 각 인스턴스 마다 보안 정책이 적용되어야 한다. 이 정책을 기준으로 각 사용자는 오직 이에 해당되는 IS들을 대상으로만 정보검색 기능들을 실행시킬 수 있다.

(2) 시스템 구조

(그림 10)에서 보듯이 서로 다른 보안 정책을 갖는 IS들은 정책 관리자 (Policy Manager, PM)에 의해서 관리된다. 각 IS에 적용된 정책을 기반으로 정책 관리자는 허용된 사용자들의 목록인 접근 제어 리스트(Access Control List, ACL)를 만들고 유지한다. 만약 한 사용자가 특정 QP를 통해서 정보검색 서비스를 수행하고 싶다면 일단 해당되는 IS들의 ACL이 먼저 체크되어 동적 가상조직이 만들어지고 그 가상조직 내에서 검색 서비스가 이루어 진다.



(그림 10) 제안된 시스템의 구조

(3) 보안 정책의 생성

각 IS마다 서로 다른 보안 정책을 적용하기 위해서는 각 IS를 위한 정책관리자를 사용할 수 있어야 한다. 또한 정책관리자를 사용하기 위해서는 정책관리자는 물론 인증 서비스가 함께 글로벌스 컨테이너에서 먼저 실행이 되어야 하고 이렇게 해야 IS서비스가 정책관리자를 불러 사용할 수 있게 된다. 이 서비스들은 아래와 같은 단계를 거쳐서 생성될 수 있다.

• 프록시의 생성

서비스들과 사용자들 사이의 통신을 위한 일반적인 프록시는 12시간의 유효기간을 가지며 아래의 명령어를 이용하여 생성된다 [10-12].

```
% grid-proxy-init
```

• 보안 서비스의 생성

아래의 명령어를 이용하면 인증 서비스와 정책관리자를 글로벌스 컨테이너에서 생성시킬 수 있다.

```
% 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
```

• 보안 서비스 정책 참조

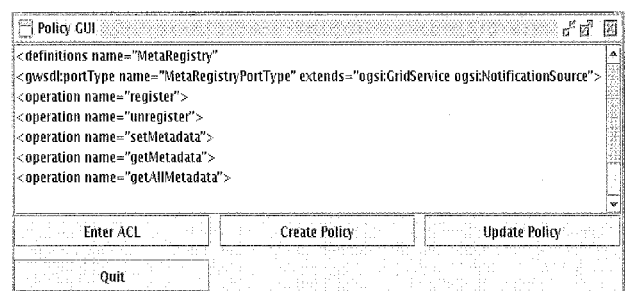
보안 정책의 생성과 심화된 단계의 인증을 위한 XACML 정책의 생성을 위해서 GUI 기반의 사용자 환경을 추가하였다. 'PolicyGUI'는 접근 제어를 위한 파일을 만들고 정책과 서비스에 관한 정보를 같이 이용하여 'PolicyManagerService'를 초기화시킨다. 'PolicyGUI'는 다음과 같이 실행시킬 수 있다[1][8][13].

```
% java org.cnidr.ogsa.appSec.policyManager.Policy
GUIhttp://localhost:8080/ogsa/services/gidrir/MetaRegistry/mr
http://localhost:8080/ogsa/services/appSec/PolicyManager/pm
$GRIDIR_LOCATION/src/schema/metaRegistry/metaRegis-
try.gwsdl
```

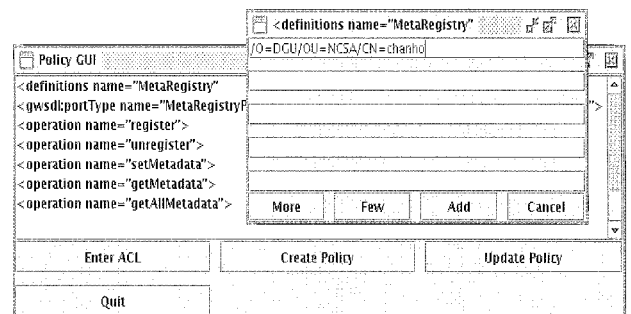
위의 명령어들을 실행시키면 (그림 11)과 같은 GUI 창이 뜬다.

다음으로 서비스 목록에서 정책 서비스를 선택하고 'Enter ACL'을 누르면 (그림 12)에서 보여주듯이 접근이 허용된 사용자의 서비스 유형과 포트, 그리고 구별이름(DN)을 입력시킬 수 있는 텍스트 창이 뜬다.

이 프로세스는 주어진 서비스에 대해 이미 인증을 받은 사용자들만을 대상으로 허용된다. ACL에 입력시킨 후 'Add' 버튼을 누르고 정책 생성을 위해 'Create Policy' 버튼



(그림11) 정책관리자 툴



(그림 12) 접근 제어 리스트에 항목 추가

을 누른다. 이렇게 하면 정책 파일이 생성되고, 이 새롭게 생성된 파일이 주어진 서비스를 처리하기 위한 정책이란 것을 '정책관리서비스'에게 알려준다. 현재의 정책을 새로운 정책으로 변경시키는 작업 또한 'Update Policy' 버튼을 이용하여 처리할 수 있고, 이 경우 이 작업은 보안 기능을 사용할 모든 IS들에게 적용되어야 한다.

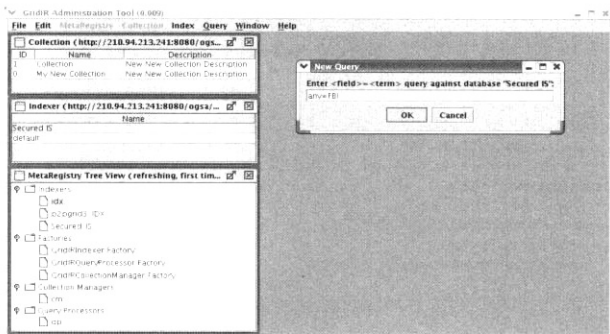
◦ 그리드 정보검색 서비스 인스턴스들의 생성

마지막으로, CM, IS, QP와 같은 그리드 정보검색 구성요소들의 인스턴스들을 생성시키고, 이러한 인스턴스들을 동적으로 제어하는 'MetaRegistry' 또한 글로벌스 컨테이너에 생성시켜야 한다[5][8][14].

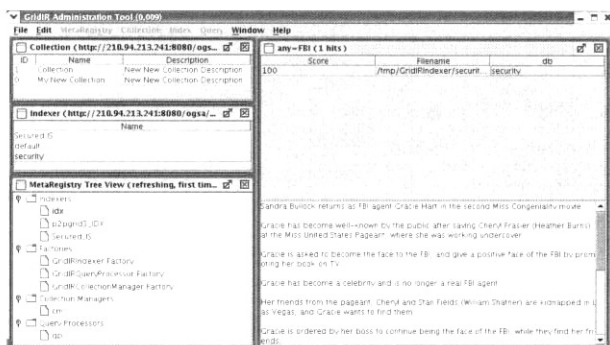
2.2.2 시스템 구현과 결과

일단 두 개의 서로 다른 보안 정책을 갖는 IS들을 구성하였다. 이들 중 하나는 보안 정책이 적용된 FBI에 대한 색인된 데이터를 가지고 있고, 다른 하나는 아무런 보안정책이 적용 안 된 FBI에 대한 영화 정보를 색인 시킨 데이터를 가지고 있다. 따라서 보안 정책을 만족하는 허가된 사용자라면 정보검색 기능은 두 개의 색인들을 대상으로 처리될 것이며, 보안이 적용된 다른 데이터에 대한 접근이 불허된 사용자라면 단순히 영화 정보를 갖고 있는 인덱스를 대상으로 정보검색을 수행할 것이다.

(그림 13)에서 보듯이 보안이 적용된 IS 하나와 보안이 적용 안된 IS 하나씩, 전부 두 개의 IS들을 정보검색 시스템에서 사용하였고, 이들은 모두 한 개의 CM과 한 개의 QP



(그림 13) 새로운 질의어의 생성



(그림 14) 보안 정책이 적용 안 된 일반 사용자의 검색 결과



(그림 15) 보안 정책이 적용된 허가된 사용자의 검색 결과

를 사용하도록 구성하였다. 다음으로 'any=FBI'와 같은 질의어를 한 번은 허가된 사용자를 위해서 다른 한 번은 일반 사용자를 위해서 실행시켰다.

(그림 14)에서 보듯이 보안 정책이 없는 일반 사용자의 검색결과는 영화에 대한 자료를 가지고 있는 IS에서 검색된 한 개이나, 보안 정책이 설정된 사용자의 검색 결과는 그림 15에서 보듯이 영화에 대한 것 하나와 FBI 소개를 다루고 있는 것 등 두 개가 검색된다. 두 경우 모두 "FBI"라는 검색어의 빈도수에 따라 정확도가 결정된다.

3. 관련 연구

본 논문에서 우리는 보다 나은 전체 시스템 활용도와 작업분배를 위해서 정보검색 서비스들을 유휴 컴퓨팅 노드에 동적으로 할당할 수 있다는 것을 보여주었다. 이렇게 원격으로 서비스를 할당하는 기능은 그리드 서비스의 구성과 설치에 관한 기능과 유사하다고 할 수 있다. 이런 분야에 대한 연구는 기존에 몇 개가 존재하지만[15-18, 20, 21], 그 중에 대표적인 것으로는 DistAnt[16][20] 프로젝트와 HAND[21] 프로젝트를 들 수 있다. DistAnt(Distributed Ant)는 사용자 기반의 자동화된 어플리케이션 배치 시스템이다. DistAnt는 공통의 어플리케이션 배치 게이트웨이, 배치 기술서, 파일 시스템 구조와 자원 기술서를 제공하기 위해 그리드 자원 속성들을 가상화시킴으로써 이기종의 그리드 환경에서 어플리케이션을 배치할 수 있다. 그리고 HAND (Highly Available Dynamic Deployment Infrastructure for Globus Toolkit 4)는 두 가지 접근 방법을 통해 동적으로 서비스를 배치한다. 첫 번째 접근 방법은 서비스 레벨 배치 (HAND-S)로써, 동작중인 여러 개의 서비스들을 비활성화한 후 새로운 서비스를 설치하고 컨테이너 전체의 재 시작 없이 이들 서비스들만 다시 활성화시킨다. 두 번째 접근 방법은 컨테이너 레벨 배치 (HAND-C)로써, 새로운 서비스를 설치할 때 전체 컨테이너 (재 시작과 재설정)를 다시 활성화 시키는 접근 방법이다. 일반적으로 많은 그리드 서비스 컨테이너들은 컨테이너를 재구성시키지 않고 서비스를 생성시키는 실시간 설치(hot deployment) 기능을 제약적으로만 지원하고 있다. 또한 그 중의 일부[16][18]는 특정 응용만을 위한 것으로 일반적인 기

능들이 없고 따라서 글로벌스와 같은 상이한 미들웨어를 갖는 이기종 시스템에 이식할 수가 없다.

일반적으로 그들의 그리드 서비스 설치 툴의 개발 목적은 주로 새로운 서비스의 원격 설치를 쉽게 하자는 것이고, 반면에 본 논문에서 취하고 있는 그리드 정보검색 컴포넌트 서비스들의 사전 설치와 작동 방식은 동적 시스템 재구성 과 균형 있는 작업 분배를 위한 것이다.

그리드 검색 엔진과 관련된 또 하나의 진행되고 있는 연구[19]가 있으나 이것은 자원정보에 대한 일관된 관점을 제공하기 위해 사용자들에게 유용한 자원을 찾아주거나, 그 자원들의 용량을 알려주거나, 또는 사용상의 조건을 알려주는 등사용자들이 상위 단계에서 그리드 자원정보 검색을 위한 포탈처럼 사용하기 위한 것이다. 이 연구는 글로벌스의 자원 모니터링 및 발견 서비스와 같이 다양한 그리드 미들웨어의 서브시스템들이 제공하는 보다 정형화된 자원정보를 검색하는 반면에, 본 논문에서의 그리드 정보검색 시스템은 검색 영역을 한 가상조직 내로 한정시키는 보다 일반적이고 동적 시스템 재구성이 가능한 분산 검색 엔진으로 사용된다.

#### 4. 결론 및 향후 연구 과제

본 논문에서는 상용 그리드 컴퓨팅의 두 가지 주된 요구 사항으로 알려진 동적 가상조직과 보안 기능을 추가한 새로운 그리드 정보검색 시스템 환경을 구현하였다. 우선 두 개의 활용 케이스를 만들어 이러한 시스템이 어떻게 유용할 수 있는지를 보여 주었고, 글로벌스 3.0 상에 구현된 그리드 정보검색 시스템 시제품에 이 기능들이 구현하여 추가하였다. 본 그리드 정보검색 시스템은 여러 추가적인 장점을 가지나 그 중에서 두 개의 주요 장점을 아래에 논의하였다.

첫 번째 장점은 기업과 정부 조직들이 그들 내부의 정보 검색 시스템을 구축 시 비용을 절감할 수 있다는 것이다. 시스템 확장 시 고가의 서버나 클러스터를 새로이 추가하는 대신에 그리드 분산 컴퓨팅 기술을 이용하여 기존에 존재하는 여분의 시스템들을 동적으로 가상조직에 포함시킴으로써 높은 컴퓨팅 성능을 기대할 수 있다. 또한 특정 서비스에 대한 동적 과부하가 시스템을 사용 불능 상태로 만들 수 있는 기존의 정보검색 시스템과는 달리 그리드 정보검색 시스템은 활용도가 낮은 컴퓨팅 자원을 이용한 동적 부하 분배 기술을 사용하거나 기존에 존재하던 자원들을 그 가상조직에 추가하여 시스템을 동적으로 재구성할 수 있다.

두 번째 장점은 서로 다른 보안 정책을 갖는 정보검색 시스템들을 한 가상조직으로 통합 시 발생하는 보안 문제들을 해결함으로써 보안이 필요한 문서나 정보들을 보다 효율적으로 관리할 수 있다는 점이다. 이것은 섬세하지 못한 인터넷 기반의 제어 방식과는 달리 서로 다른 기업간에 또는 정부기관 간에 다단계 보안 정책을 적용하여 서로간에 보안된 정보를 공유할 수 있음을 의미한다.

현재 작동하는 시제품은 글로벌스3.0 위에 구현되었으나 새로운 버전인 4.0으로의 이동은 현재 작업 중으로 조만 간

에 완료될 것이고, 이러한 작업은 GGF(Global Grid Forum)의 그리드 정보검색 워킹 그룹 활동의 일환으로 이루어지고 있다. 본 워킹 그룹은 대륙간 이기종 환경에서 대용량 분산 정보검색 시스템 테스트 베드를 구축하고 연동시켜 그 결과를 HPC Storage Challenge [22]에 발표할 계획으로 진행 중에 있다.

#### 참고 문헌

- [1] Ian Foster, Carl Kesselman, Steven Tuecke, "Security Architecture for Computational Grids." 5th ACM Conference on Computer and Communications Security Conference, 1998.
- [2] Dr. Bernhard, R. Katzy, "Design and Implementation of Virtual Organizations,," Proceeding of 31st International Conf. on System Science, vol.4, pp.142 - 151, IEEE, 1998.
- [3] Ian foster, Carl Kesselman, Steven Tuecke, "The Anatomy of the Grid: Enable Scalable Virtual Organizations," International Journal of Supercomputer Applications, 2001.
- [4] Carl Kesselman, Jeffrey M Nick, Steven Tuecke, Ian Foster, "Grid Service for Distributed System Integrations." IEEE Trans. on Computer, vol. 35, Issue 6, pp37 - 46, 2001.
- [5] Ian Foster, Carl Kesselman, Jeffrey M. Nick, Steven Tuecke, "The Physiology of the Grid: An Open Grid Service Architecture for Distributed System Integration." GGF Open Grid Service Infrastructure Working Group, 2002.
- [6] A. Arasu, J. Cho, H. Garcia-Molina, A. Paepcke, and S. Raghavan, "Searching the Web," ACM Transactions on Internet Technology, 1(1), pp2 - 43, 2001.
- [7] Garniel, Kevin; Newby, Gregory B. & Nassar, Nassib "Grid Information Retrieval Requirements (GFD.27)." Lamont, Illinois: Global Grid Forum, 2003.
- [8] Radu Prodan, Thomas Fahringer, "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, 2003.
- [9] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," In Proceedings of 7th World Wide Web Conference, 1998.
- [10] S. Cannon, S. Chan, D. Olson, C. Tull, V. Welch, L. Pearlman, "Using CAS to Manage Role-Based VO Sub-Groups." Proc. Computing in High Energy Physics 03 (CHEP '03), 2003.
- [11] L. Pearlman, C. Kesselman, V. Welch, I. Foster, S. Tuecke, "The Community Authorization Service: Status and Future." Proc. Computing in High Energy Physics 03(CHEP '03), 2003.
- [12] I. Foster, N. T. Karonis, C. Kesselman, S. Tuecke,

“Managing Security in High-Performance Distributed Computing.” Cluster Computing, 1(1):95-107, 1998.

[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 Globus Toolkit, The Globus alliance. (see <http://www.globus.org>)

[15] J McCabe, TJ Harmer, P Donachy, RH Perrott, “Grid Service Configuration and Lifecycle Management,” Proceedings of the First International Conference on e-Science and Grid Computing (e-Science’05), 2005.

[16] Wojtek Goscinski and David Abramson’ “Application Deployment over Heterogeneous Grids using Distributed Ant,” Proceedings of the First International Conference on e-Science and Grid Computing (e-Science’05), 2005.

[17] Fenglian Xu, M. Hakki Eres, David J. Baker, and Simon J. Cox, “Tools and Support for Deploying Applications on the Grid,” Proceedings of the 2004 IEEE International Conference on Services Computing (SCC’04), 2004.

[18] Roberto Santinelli, Flavia Donno’ “Installing and Configuring Application Software on the LHC Computing Grid,” Proceedings of the First International Conference on e-Science and Grid Computing(e-Science’05), 2005.

[19] Marios Dikaiakos , Yannis Ioannidis and Rizos Sakellariou, “Search Engines for the Grid: A Research Agenda,” Grid Computing: First European Across Grids Conference, Spain, February 13-14, 2004. Lecture Notes in Computer Science, Volume 2970, pp.49-58, 2004.

[20] Goscinski. W, Abramson. D, “Distributed Ant: a system to support application deployment in the grid,” Proceedings of Fifth IEEE/ACM International Workshop on Grid Computing, pp.436-443, Nov. 2004.

[21] Li Qi, Hai Jin, Ian Foster, Jarek Gawor, “HAND : Highly Available Dynamic Deployment Infrastructure for Globus Toolkit 4,” Submitted for Publication, 2006.

[22] HPC Storage Challenge, (see [http://sc06.supercomp.org/conference/hpc\\_storage\\_challenge.php](http://sc06.supercomp.org/conference/hpc_storage_challenge.php))

### 김 양 우

e-mail : ywkim@dongguk.edu

1984년 연세대학교 전자공학과(공학사)

1986년 Syracuse Univ.

컴퓨터공학전공(공학석사)

1992년 Syracuse Univ.

컴퓨터공학전공(공학박사)

1992년~1996년 한국전자통신연구원

선임연구원



1996년~현재 동국대학교 정보통신공학과 부교수

관심분야: 분산 그리드 컴퓨팅 시스템, 컴퓨터구조

### 이 승 하

e-mail : lesh915@dongguk.edu

2001년 동국대학교 정보통신공학과(학사)

2003년 동국대학교 정보통신공학과(석사)

2003년~현재 동국대학교 정보통신공학과 박사과정

관심분야: P2P, GRID, 검색엔진, 원격실행



### 김 혁 호

e-mail : hulegea@dongguk.edu

2003년 대전대학교 정보통신공학과(학사)

2005년 동국대학교 정보통신공학과(석사)

2005년~현재 동국대학교 정보통신공학과 박사과정

관심분야: GRID, 검색엔진, P2P, Ubiquitous

