

메타데이터 카탈로그 GUI 툴 설계

허태상^o, 안선일, 김한기, 이상도, 이세훈, 유상수, 황순욱
한국과학기술정보연구원

tshuh@kisti.re.kr, siahn@kisti.re.kr, hgkim@kisti.re.kr, doyaa2@gmail.com, sehooi@kisti.re.kr,
sangsuryu@kisti.re.kr, hwang@kisti.re.kr

A Design of GUI Tool for Metadata Catalogue

Taesang Huh^o, Sunil Ahn, HanGi Kim, Sangdo Lee, Sehoon Lee, Soonwook Hwang
Korea Institute of Science and Technology Information

요 약

그리드 컴퓨팅 환경에서 세계의 많은 과학기술연구자들은 gLite 미들웨어 기반의 어플리케이션을 이용하고 있다. 미들웨어에서 데이터를 관리, 분석하기 위해서는 메타데이터 카탈로그 서비스를 이용하고 있으며, 유닉스, 그리드, 데이터베이스 환경에 익숙하지 않은 과학기술연구자들은 이 서비스를 이용하는데 어려움을 겪고 있다. 본 논문은 과학기술연구자로 하여금 gLite의 공식 컴포넌트인 메타데이터 카탈로그 서비스 이용 시 접근 제약을 최소화하고 사용자 친화적인 그래픽 사용자 인터페이스를 제공하기 위한 메타데이터 카탈로그 GUI 툴 설계에 대해 논의하고자 한다.

1. 서 론

메타데이터 카탈로그 GUI(Graphic User Interface)는 EGEE(Enabling Grids for E-science) gLite 미들웨어 공식 메타데이터 서비스인 AMGA(Arda Metadata Grid Application) 서비스를 그 대상으로 한다. AMGA는 ARDA 팀에 의해서 그리드 상에서 고에너지 실험 분석을 위한 LCG 프로젝트를 통해 설계 및 개발되었다. 메타데이터 카탈로그는 조직, 계층 구조, 사용자, 그룹 및 ACL(Access Control List) 관리, SQL과 유사한 쿼리 언어, 복제기능으로 구성된다. gLite 미들웨어는 GSI(Grid Security Infrastructure)를 따르고 있으며, X.509 인증서/프록시를 통해 액세스할 수 있다.[1-4]

현재 AMGA 서비스를 이용하기 위해서 그리드 사용자들은 그리드 사용자 계정을 가지고 전문 머신을 이용하여야 하기 때문에 방화벽, 보안과 같은 환경적 설정을 항상 고려해야 한다. 또한 그리드 사용자는 전문적인 그리드 질의, 유닉스 명령어, gLite명령어 등을 통해서만 데이터의 접근이 가능하기 때문에 그리드 환경에 익숙하지 않은 과학기술연구자는 메타데이터 카탈로그를 이용시 많은 고충을 겪는다. 이에 e-Science 어플리케이션의 몇 개는 메타데이터의 관리 및 편집을 위해 별도로 GUI 기능 일부를 개발하여 사용하기도 하였다.[5]

본 논문에서는 그리드 UI 사용 없이 AMGA 메타데이터를 이용하려는 사용자가 네트워킹이 가능한 자신의 PC에 설치된 메타데이터 카탈로그 GUI 툴과 볼스 프록시(Voms Proxy)만으로 별도의 유닉스 명령어, 구문(Syntax)에 대한 전문성 없이도 자신의 메타데이터 관리, 편집 및 개발을 쉽게 할 수 있도록 GUI 툴을 설계하였다. 또한 범용적인 사용을 위해 사용자 OS환경을 고려

하여 사용자가 많이 쓰는 GUI 툴의 사용 패턴을 적용한 사용자 친화적이고 독립적 서비스 체제로 설계하였다. 본 논문에서 제시하는 GUI 툴의 스키마 브라우저(Schema Browser)와 완성형 CLI(Command Line Interface)는 과학기술연구자에게 AMGA 컬렉션(Collection), 스키마(Schema), 엔트리(Entry)에 대한 권한 관리 및 편집 기능을 제공할 수 있다.

2. 관련연구

2.1. AMGA

AMGA는 gLite 미들웨어 gLite 3.1의 일부 컴포넌트로서 그리드 메타데이터에 대한 액세스 인터페이스를 제공한다. 시작은 고에너지 물리학 연구 영역에 적합한 메타데이터 카탈로그로 설계되었고, 지금은 메타데이터 카탈로그를 필요로 하는 모든 연구영역으로 확대되어, 다양한 사용자 요구를 수용하여 한국의 KISTI에서 개발을 진행 중에 있다.[6] 세계 각국에서 Biomedical 연구와 스위스 Cern의 Atlas, LHCb 실험, 독일 DKRZ의 기상연구, EGEE의 Health e-Child, 프랑스 IN2P3의 신약후보물질 탐색 등 여러 응용 프로그램들은 그리드 작업에 대한 메타데이터 저장소로써 AMGA를 활용하고 있다.[7-10] 특히 Biomedical 응용 프로그램은 데이터의 비밀성과 무결성이 중요하므로 데이터의 권한에 따른 사용자 보안관리가 매우 중요하다. AMGA는 액세스하는 사용자 레벨에 따라 세분화된 보안 기능도 제공할 수 있다. gLite 배포용 AMGA 2.0에서는 그림 1과 같이 MD 서버를 중심으로 두 가지 형태의 프론트 엔드(front-end) 버전이 있다. 첫째는 웹서비스 프론트 엔드로써 이는 gSOAP를 사용하고 있고 둘째는 TCP 스트리밍 베이스의 프론트 엔드가 있다.

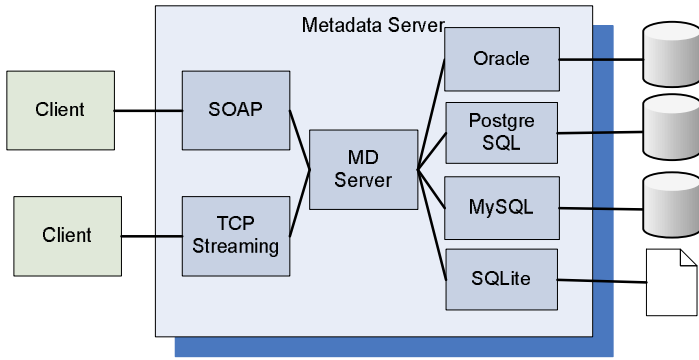


그림 1. AMGA Implementation[19]

백엔드(Backend)는 저장소 역할을 수행하는 영역으로 PostgreSQL, MySQL, Oracle, SQLite와 같은 RDB로 구성되어 사용될 수 있다.[6][11]

AMGA에서 제공하는 보안은 유닉스 스타일의 승인과 컬렉션, 엔트리 제어에 사용되는 ACL이 있다. 또한 보안 접속의 SSL 방식을 따르며, 클라이언트 인증은 사용자 이름/패스워드, X.509 인증서, 그리드 프록시 인증서 방식을 따른다. 그리드에는 아래 그림과 같은 가상조직관리(Virtual Organization Management)시스템이 있어 가상 조직마다 액세스 컨트롤을 관리할 수 있다.[12]

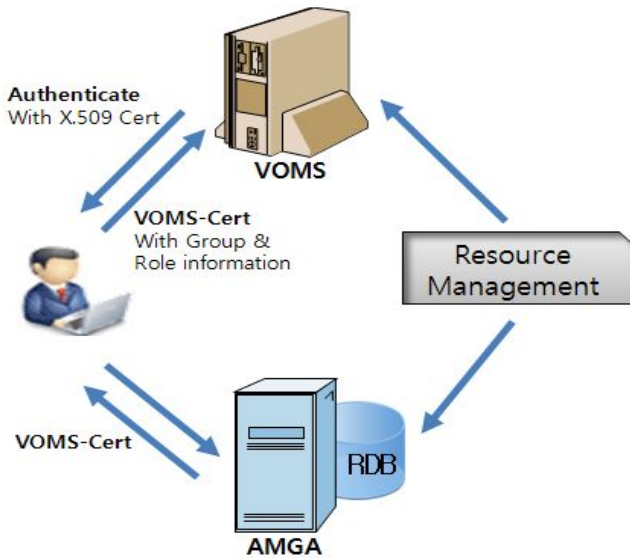


그림 2. VOMS(Virtual Organization Management System)의 액세스 컨트롤

또한 AMGA 서비스는 데이터와 데이터에 대한 메타데이터를 관리하는 강력한 인터페이스를 아래와 같이 제공한다.

- 엔트리의 의미 체계를 설명하기 위해 속성의 집합을 정의할 수 있는 스키마를 제공
- 많은 응용프로그램 영역을 지원하기 위한 동적인 스키마 적용이 가능한 유연성을 제공
- 컬렉션과 서브 컬렉션을 정의할 수 있도록 계층적 메

타데이터 구조를 제공

- 상당 수의 파일과 스키마를 관리할 수 있는 확장성을 제공
- 콘텐츠의 비밀성과 데이터 무결성을 보장하는 할 수 있는 보안성을 제공(액세스 컨트롤은 스키마와 엔트리 베이스로 지원됨, 그리드 증명서, VO 서비스, SSL의 보안 접속)

2.2. 요구사항 분석

메타데이터 카탈로그 GUI 툴 개발을 위한 요구사항은 많은 과학기술연구자들이 겪었던 UI의 불편함에서 출발한다. Window OS 어플리케이션 사용자들은 CLI에 이용하는 데 많은 어려움이 따른다. 일반사용자가 AMGA 서비스를 이용하기 위해 필요한 CLI를 최대한 GUI로 전환함을 개발의 최종 목표로 한다. AMGA가 보유한 기능을 중심으로 아래와 같이 4가지로 요구사항을 재구성했다.

○ 사용자/그룹 관리

메타데이터 카탈로그의 사용자/그룹 관리 기능은 사용자, 그룹의 리스팅, 생성, 위임 관리를 할 수 있도록 하며, 사용자와 그룹간의 관계를 관리할 수 있어야 함.

○ 액세스 컨트롤 지원

메타데이터 카탈로그 웹서비스의 AMGA API 접근 권한 제어를 위해 폼스 프록시를 제공하는 사용자만 로그인 가능해야 한다. 또한 메타데이터 카탈로그의 컬렉션 관리의 승인하는 로직과 컬렉션 관리에 대한 ACL(그룹 리스팅, 그룹추가/삭제)의 관리할 수 있어야 함.

○ 컬렉션 관리

사용자의 컬렉션 관리를 편리하게 할 수 있도록 트리 형태의 브라우징이 가능한 컬렉션 및 서브 컬렉션 관리 기능과 컬렉션의 생성, 삭제, 위임 기능을 개발한다. 또한 컬렉션은 ACL 관리에 따라서 제어할 수 있어야 하며, 엔트리 가져오기에 대한 대규모 메서드 처리가 가능해야 함.

○ 메타데이터 관리

메타데이터 관리 기능은 기본적으로 엔트리의 리스팅 및 검색이 가능해야 하며, 엔트리 생성 시 엔트리 속성값을 가지는 엔트리를 삽입할 수 있도록 개발한다. 또한 엔트리의 수정, 위임과 스키마 관리가 비전문적 그리드 사용자가 쉽게 할 수 있어야 함.

3. 메타데이터 카탈로그 GUI 툴 설계

AMGA에서 메타데이터는 데이터를 위한 데이터로써 실제 실험 데이터는 파일형태로 이루어지고 본 툴은 메타데이터 카탈로그를 사용자 편의를 중심으로 설계했다. 사용자는 자신의 PC환경(MAC/Window/Linux)에 구애받지 않는 클라이언트용으로 설계되었고, 그 아키텍처는 아래 그림과 같다. 메타데이터 카탈로그 GUI 툴은 그리드 API를 이용하여, AMGA의 프론트 엔드 중 SOAP과 연동하여 데이터에 접근한다.

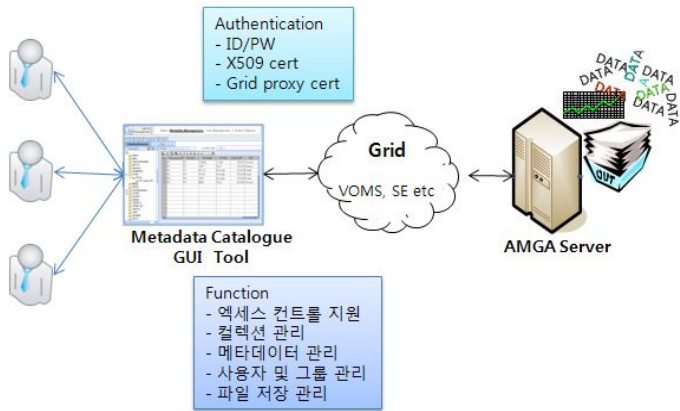


그림 3. 메타데이터 카탈로그 관리 개념도

서버/클라이언트 아키텍처에서 주요 클라이언트 컴포넌트로 인증서 관리(Authentication Manager), 응용프로그램 관리(Application Manager), 모니터링(Monitoring), 권한관리(Privilege Manager), 파일관리(File Manager), 응용프로그램 업데이트 관리(Application Update Manager) 모듈로 구성된다.[그림 4]

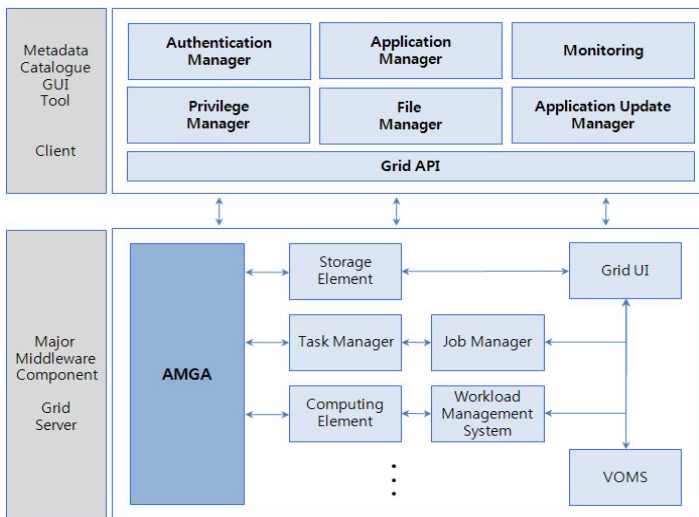


그림 4. Server/Client 아키텍처 다이어그램

인증서 관리는 사용자의 머신사용에 대한 username/password 관리와 X.509 인증서의 등록, 그리드 프록시 인증서 등록을 위한 모듈이고 응용프로그램 관리는 GUI의 관리할 수 있는 컴포넌트로 데이터 편집, 각종 쿼리 지원, 개인화 서비스가 가능한 모듈이다. 모니터링은 사용자의 메타데이터의 변경, 스키마의 변경을 모니터링하고, 권한관리는 메타데이터 카탈로그 컬렉션의 ACL(그룹리스팅, 그룹추가/삭제)을 관리할 수 있는 모듈이다. 파일관리는 파일의 등록, 삭제를 관리하고, 응용프로그램 업데이트 관리는 AMGA의 버전에 변경을 모니터링하여 새로운 클라이언트 툴을 설치할 수 있도록 지원하는 모듈이다. 메타데이터 카탈로그 GUI 툴의 주요 핵심 모듈은 다양한 메타데이터 조작을 가능하게 한다. 그림 5는 메타데이터의 주요 기능별로 권한위임을 표현한 그림이다. 컬렉션 관리에서 컬렉션을 추가, 읽기, 변경, 삭제하고

ACL과 엔트리를 관리할 수 있다. ACL 관리는 소유자(Owner)의 관리, 모드(Mode)의 관리 및 그룹의 관리(추가, 읽기, 변경, 삭제)가 가능하다. 엔트리 관리에서는 엔트리의 추가, 읽기, 변경, 삭제, 검색이 가능하고 스키마 관리에 대해서 애트리뷰트(Attribute)의 추가, 읽기, 변경, 삭제가 가능하다.

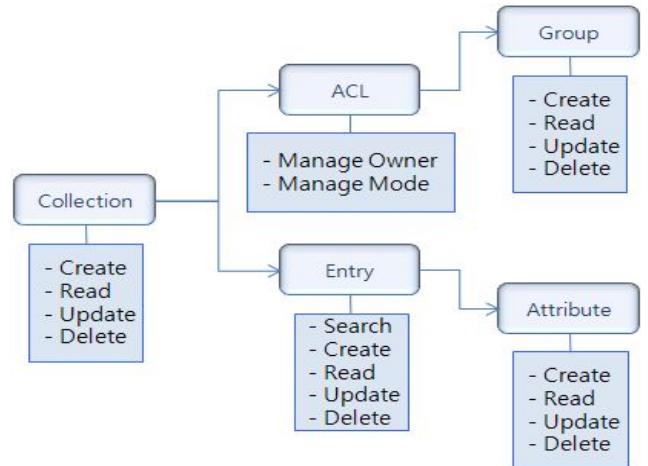


그림 5. 메타데이터 주요 기능(권한위임) 다이어그램

그림 6는 메타데이터 관리가 가능한 스키마 브라우저 화면으로 컬렉션은 디렉토리별로 메타데이터를 관리하도록 되어 있어, 사용자는 윈도우 탐색기와 일반 상용 데이터 개발/관리 툴과 유사한 인터페이스를 제공받을 수 있을 뿐만 아니라, SQL 쿼리 이용이 편한 사용자는 완성형 CLI 기능도 병행해서 제공받을 수 있게 설계하였다.

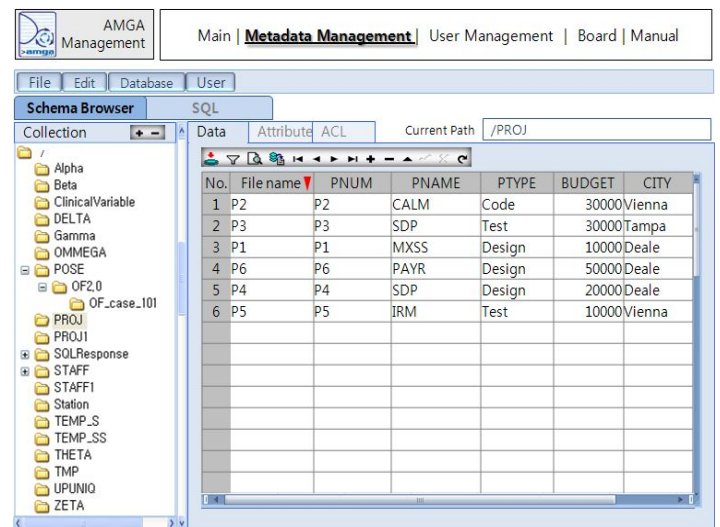


그림 6. 메타데이터 카탈로그 GUI 툴 주요 화면 설계 [예시]

4. 결론

기존 사용자들은 자신의 메타데이터와 데이터 파일을 디렉토리별로 접근하기 위해 유닉스 환경의 명령어와 그리드 SQL을 사용하여 왔다. 이는 시스템 사용이 미숙한 일반 사용자로 분류될 수 있는 사용자에게는 다소 낮설

고 사용방법을 숙지하는데 어려움이 따랐다. 본 연구에서는 메타데이터 카탈로그를 활용하는 그리드 과학기술 연구자에게 보다 편리한 사용자 인터페이스를 제공하기 위한 프로그램 설계에 대해 논의하였다.

본 연구에서는 메타데이터 카탈로그 GUI 툴의 본 개발에 앞서 프로그램 레이아웃에 해당하는 설계에 대해서 요구사항의 분석과 기초 설계를 중심으로 설명하였다. GUI 툴은 gLite AMGA 메타데이터 카탈로그에 접근하는 AMGA API를 토대로 일반사용자그룹과 관리자의 권한에 적합한 기능으로 설계하였다.

메타데이터 카탈로그 GUI 툴은 AMGA 데이터 관리를 사용자 친화적 인터페이스를 제공함으로써, 몇 가지 기대효과를 갖는다. 첫째, 시스템 명령어, 그리드 명령어, 쿼리 등에 비전문적인 과학기술연구자에게 데이터 활용을 극대화하여 연구 생산력을 높일 수 있도록 지원해준다. 둘째, 사용빈도가 높은 데이터 입/출력 및 검색을 별도의 SQL 명령어 없이 빠르고 쉽게 할 수 있도록 지원해 준다. 셋째, AMGA 메타데이터 카탈로그 서비스 이용을 활성화 할 수 있다.

메타데이터 카탈로그 GUI 툴 개발이 완료되고 안정화가 이루어지면, gLite AMGA 클라이언트로 공개 배포할 계획이며, 사용자의 요구와 AMGA 버전 업데이트에 따라 향후 버전관리가 지속적으로 필요할 것이다.

참 고 문 헌

- [1] Enabling Grids for E-science: <http://public.eu-egee.org/>
- [2] Gagliardi F, Jones B, Grey F, Bégin M-E, Heikkurinen M. "Building an infrastructure for scientific grid computing: Status and goals of the EGEE project", Philosophical Transactions: Mathematical, Physical and Engineering Sciences, vol.363, pp.1729 - 1742, 2005.
- [3] The gLite middleware: <http://glite.web.cern.ch/glite/>
- [4] Laure E, Fisher SM, Frohner A, Grandi C, Kunszt P, Krenek A, Mulmo O, Pacini F, Prezl F, White J, Barroso M, Buncic P, Hemmer F, Di Meglio A, Edlund A. "Programming the grid with gLite", Computational Methods in Science and Technology, vol.12(1), pp.33 - 45, 2006.
- [5] 황순욱, 남덕운, 김법균, "e-Science 국내외 기술 동향", 정보처리학회, 정보처리학회지, vol.15 no.2, pp.43-57, 2008.
- [6] AMGA. ARDA Metadata Grid Application. Available at: <http://amga.web.cern.ch/amga/>
- [7] Montagnat J, Jouvenot D, Pera C, Frohner A', Kunszt P, Koblitz B, Santos N, Loomis C. "Bridging clinical information systems and grid middleware: A medical data manager", Proceedings of the HealthGrid, Valencia, Spain, 2006, June.
- [8] T. Doherty, "Development of gLiteWeb Service Based Security Components for the ATLAS Metadata Interface", EGEE User Forum, CERN, Switzerland, 2006.
- [9] J. P. Baud, S. Lemaitre, "The LHC File Catalogue (LFC)", HEPPIX, Karlsruhe, Germany, 2005.
- [10] Ahn. Sunil, Kim. Namgyu, Lee. Seehoon, "Performance analysis and optimization of AMGA for the large-scale virtual

screening", Software, practice and experience, vol.12, pp.1055-1072, 2009.

- [11] Box D, Ehnebuske D, Kakivaya G, Layman A, Mendelsohn N, Nielsen HF, Thatte S, Winer D. Simple object access protocol (SOAP) 1.1. World Wide Web Consortium, Note NOTE-SOAP-20000508, 2000, May.
- [12] Ganga. Gaudi/Athena and Grid Alliance. Available at: <http://ganga.web.cern.ch/ganga>
- [13] Alfieri R, Cecchini R, Ciaschini V, dellAgnello L, Frohner A', Gianoli A, Lo~rentey K, Spataro F. "VOMS, an Authorization System for Virtual Organizations", Lecture Notes in Computer Science, vol. 2970. Berlin, 2004.