

그리드 기반 가상대형장비실험을 위한 통합데이터베이스 시스템의 설계 및 구현

김근태⁰ 김한준

서울시립대학교 전자전기컴퓨터공학부

{kobi⁰, khj}@uos.ac.kr

Design and Implementation of Integrated Database System for Grid-based Virtual Instrument Experiments

Keuntae Kim⁰, Han-joon Kim

Department of Electrical and Computer Engineering, The University of Seoul

요 약

본 논문에서는 그리드 기술과 웹 서비스 기술을 이용하여 원격지에 위치한 고가의 대형 실험장비와 데이터 공유를 위한 통합데이터베이스 시스템을 제안한다. 제안된 통합데이터베이스 시스템은 오픈 분산 컴퓨팅 환경을 통해서 지리적으로 분산되어있는 전체 실험시설들의 실험장비와 데이터베이스를 총괄 관리한다. 또한 인증과 권한 기능을 도입하여, 인증 받은 실험자에게만 권한 별로 차등하여 실험장비 이용과 데이터베이스의 접근을 제공하며, 공유된 실험결과데이터에 대해서 효율적인 검색과 데이터의 다운로드, 실험결과 그래프를 제공한다. 국내 6개의 건축대형장비를 연결하는 통합시스템의 프로토타입의 개발을 통해 가상대형장비실험의 효율성을 확인하였다.

1. 서 론

유비쿼터스 컴퓨팅 시대의 도래와 함께 독립적으로 존재했던 IT자원이 필요할 때 서로 빌려 사용할 수 있는 '활용'모델로 변화하고 있다. 현재 생물학, 물리학, 화학 연구, 건축, 토목, 공학 응용 연구 등의 분야에서 문제점으로 지적되고 있는 고가의 '대형 실험 장비 확보 문제와 많은 시간이 소요되는 실험 수행, 그리고 접근성의 어려움을 해결하고자 그리드 기술과 웹 서비스를 도입, 가상화를 통하여 원격지의 실험 장비를 자신의 실험장치처럼 사용하고자 한다.

그리드는 슈퍼컴퓨터와 같은 고성능 자원을 가상 조직으로 묶어 자유롭게 사용함으로써 엄청난 양의 계산 자원을 필요로 하는 과학분야를 지원하는 목적으로 나타났다[1]. 그리드의 중요 특징으로는 초고속 네트워크에 기반 한 차세대 인터넷 서비스이면서, 분산 컴퓨팅을 위한 개방형 표준을 지향하면서 텍스트 중심의 공유에서 벗어나 글로벌 리소스의 공유를 가능케 하여 웹 서비스와 결합되어 어플리케이션의 공유까지 가능한 IT 기반 인프라인 점을 들 수 있다[1][2]. 또한 그리드는 자원 공유를 가능하게 하고 자원 접근을 쉽게 하며, 중복투자를 방지하여 동 분야 연구의 효율성을 높이고 타 분야와의 연계를 활성화시킬 수 있다[3].

가상대형장비실험은 그리드 환경내의 여러 자원들을 웹에서 접속 할

수 있도록 하는 것이다. 대형연구시설과 컴퓨팅시설, 디지털 도서관이 초고속 정보통신 네트워크에 의해서 연결되며, 그리드 시스템으로 통합되어 전체 연구시설이 마치 하나의 연구시설처럼 작동하는 연구시설이다. 이를 통해서 협업이 가능해지며, 실험적 연구, 이론 개발과 검증, 데이터 저장, 모델-기반 시뮬레이션, 고성능 계산 및 교육이 통합되어, 지금까지 성취할 수 없었던 건설, 토목 기술 연구 역량을 보유할 수 있으며, 통합된 연구 인프라를 이용하여 지금까지 수행할 수 없었던 혁신적인 연구를 수행할 수 있다[6].

본 연구에서는 그리드 컴퓨팅 기술을 활용하여 지리적으로 분산되어 있는 대형실험장비를 활용하기 위함에 있다. 그래서 실험메타정보, 실험결과데이터 및 실험과정 전반을 지원하는 데이터 스키마와 프로토타입을 제안하게 되었다.

2. 가상대형장비실험을 위한 통합데이터베이스 시스템

통합데이터베이스 시스템은 원격지의 실험실을 초고속 통신망으로 연결하여 대화형 메시지 전달, 전자공책, 원격 실험장비 모니터링 기능을 이용하여 원격지에서 수행하는 실험과정에 직접 참여하게 하며 실험 자료를 데이터베이스로 구축하여 정보의 공유를 실현하는데 목적이 있다.

2.1 요구사항

대형장비를 이용한 가상실험을 실현하기 위한 통합데이터베이스 시

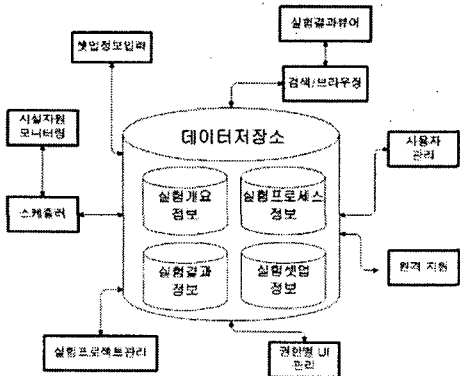
본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 육성 지원 사업(HTA-2005-C1090-0502-0016)의 연구결과로 수행되었음.

시스템은 아래와 같은 기능을 요구한다.

- 1) 스케줄링 및 자원 모니터링: 실험장비는 공동 사용 기기 이므로, 실험장비에 대한 사용 일과 시간을 다른 사람이 볼 수 있도록 한다. 자원이 어떻게 사용되고 있는지 자원의 상태를 모니터링 할 수 있어야 한다.
- 2) 인증 및 권한: 단일화된 사용자 인증 서비스로 사용자 인증을 통한 보안 유지, 접근 제어를 한다. 권한 레벨에 따라 장비의 제어 나 데이터의 열람의 범위가 제한되어야 한다.
- 3) 원격 제어 및 통신: 실험장비를 원격지에서 작동하기 위하여, 실험자가 통합데이터베이스 시스템에 접속하여 원격지의 실험장비를 셋팅, 제어가 가능해야 한다.
- 4) 실험데이터 공유: 실험 및 가상실험을 통하여 생성된 데이터를 데이터베이스에 저장하며, 공유된 실험데이터만 다른 연구자들이 열람 할 수 있게 한다. 실험자가 공유여부를 결정할 수 있어야 한다.
- 5) 실험데이터 검색: 메타데이터를 통해 원하는 데이터를 효율적으로 검색할 수 있어야 하며, 모든 연구자가 손쉽게 실험데이터를 검색할 수 있어야 한다.
- 6) 사용자 권한에 따른 UI 및 데이터 접근의 차별화: 인증 및 권한 서비스를 이용하여 사용자 권한에 따른 UI의 차별화를 통해 데이터와 실험장비에 대한 악의적인 접근을 차단해야 한다.

2.2 시스템 설계

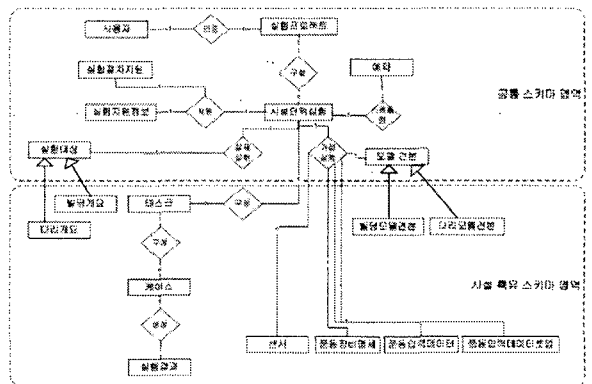
본 연구에서 통합데이터베이스 시스템은 위에서 제시한 요구사항을 반영하여 그림 1와 같은 아키텍처와 그림 2와 같은 데이터스키마를 설계 하였다.



(그림 1) 통합데이터베이스 시스템의 소프트웨어 아키텍처

통합데이터베이스 시스템의 아키텍처는 데이터 저장소와 9 개의 모듈로 구성된다. 모듈은 실험프로젝트관리, 스케줄러, 시설자원모니터링, 셋업정보입력, 사용자 관리, 권한별 UI 관리, 검색/브라우징, 실험결과뷰어, 원격지원 등이 있으며, 모듈은 아래와 같은 역할을 하게 된다.

‘실험프로젝트관리’ 모듈은 특정의 목적을 달성하기 위해 하나 이상의 실험시설을 활용하기 위해서 ‘실험프로젝트’를 관리하는 모듈이다. ‘스케줄러’, ‘시설자원 모니터링’ 모듈은 특정 시설에 대한 실험을 예약하기 위하여 예약기능, 날짜 중복 예약 방지, 실험예약 정보 조회, 실험요청 승인의 절차를 지원하는 기능을 가진다. 실험 시설에 대한 정보에 접근하여 실험요청자가 요청한 해당 시설을 요청한 날짜에 사용 가능한지 여부를 판단한다. ‘셋업정보입력’ 모듈은 실험준비를 위해 실험시설에 따라 다양한 데이터를 입력해줘야 한다. ‘사용자관리’ 모듈과 ‘권한별 UI 관리’ 모듈은 실험시설을 사용/관리하는 사용자들의 정보를 관리하는 것이다. 사용자를 확인 하는 인증과 레벨에 따른 권한으로 장비의 제어나 데이터의 열람의 범위를 제한하며, 사용자 등급별 메뉴 차별화로 악의적인 접근을 차단할 수 있다. ‘검색/브라우징’, ‘실험결과뷰어’ 모듈은 실험데이터 검색 방법으로 키워드 검색과 브라우징 검색인 트리 메뉴 구조를 이용하여 실험데이터 검색이 가능하며, 결과데이터를 업/다운로드를 제공하며, 결과데이터로 그림 3과 같이 결과데이터 차팅기능을 제공한다. ‘원격지원’ 모듈은 원격지의 실험자와 실험시설, 실험자와 시설의 보조실험자의 통신을 확립하는 모듈이다.



(그림 2) 데이터저장소 및 원격실험 지원을 위한 데이터스키마 (ER 다이어그램)

데이터 스키마는 크게 두 가지로 나눌 수 있는데, 공통 스키마 영역과 시설 특유 스키마 영역이다(그림2 참조). 공통 스키마 영역은 모든 실험시설에서도 동일하게 쓰이는 스키마를 나타내고 있으며, 시설 특

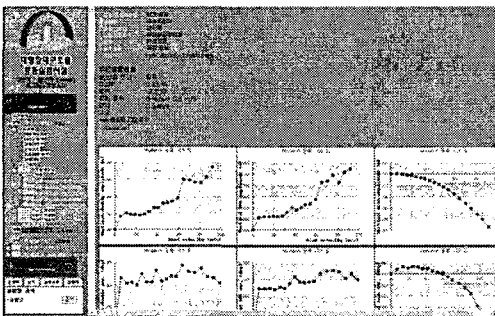
유 스키마 영역은 '설정정보'와 '실험결과' 등을 포함하는데, 각 실험 시설 별로 요구사항에 따라 별도의 데이터스키마를 설계해야 한다.

2.3 구현 및 활용

가상대형장비실험의 통합데이터베이스 시스템은 전북대학교에 위치한 "대형장대구조를 공동실험센터"(이하 '공동실험센터')를 프로토타입으로 개발하였다.

구현환경은 리눅스에서 GT4(globus toolkit 4)를 이용하여 OGSA환경을 구축하였고, DBMS로는 MySQL를 사용하였다. 또한 Apache와 Php5로 웹 인터페이스 기반의 통합데이터베이스 시스템을 구현하였다.

'공동실험센터'를 가정하여 시나리오를 묘사해보면, 원격지의 실험자가 통합데이터베이스 시스템에 접속을 한다. 프로젝트를 정의한 후 해당 실험센터로 접속한다. 실험센터에서는 스케줄링 기능을 통하여 사용되지 않는 날을 선택하여 예약하게 된다. 예약 시에는 실험의 기본적인 개요정보가 입력되고, 실험종류를 선택한다. 센터장의 승인에 의하여 실험시설 사용 여부가 결정된다. 센터장의 승인 후 원격지의 실험자가 실험장비의 설정정보를 입력한다. 선택한 실험종류에 따라서 "Task"와 "Case"가 다르게 생성된다. 실험이 진행 중 발생하는 실험 결과데이터는 실험센터의 데이터베이스에 저장된다. 실험 종료 후 실험결과 데이터는 데이터 저장소로 저장하며, 실험자가 결과 데이터의 공개를 원할 경우에만 공유한다. 공유한 데이터에 대해서 모든 연구자들이 결과 데이터를 Case별로 열람할 수 있으며, 각 Case별 실험결과에 대하여 차팅기능을 제공한다(그림3 참조).



(그림 3) 실험결과 디스플레이

3. 평가 및 분석

통합데이터베이스 시스템은 사용자별 맞춤형 시스템으로 설계되었다. 전체 실험시설의 총괄 관리가 가능하며, 웹 인터페이스로 접근성이 높다. 트리 형태의 메뉴 구조로 사용자가 사용이 용이하며, 스케줄링과 자원모니터링 기능을 통해서 실험센터의 실험장비의 사용 여부를 확인 할 수 있으며, 미리 예약을 통해 선정을 할 수 있다. 예약을 통해 승인을 받은 실험자만이 실험장비를 다룰 수 있다. 실험자는 결

과 데이터의 공유를 결정하여, 원치 않는 데이터 유출을 막을 수 있다. 또한 공유된 데이터의 검색에 있어서 키워드 검색뿐만 아니라, 트리 구조의 메뉴로 브라우징 검색이 가능하며, 결과데이터의 그래프를 실시간으로 볼 수 있으며, 다중비교가 가능하다.

사용자의 권한과 인증 서비스를 통해 허락되지 않은 사용자에 대한 접근을 막으며, 권한에 따른 결과데이터의 열람과 실험장비의 제어의 통제를 가능하게 하였다.

실험시설에 따라 실험장비와 데이터의 차이가 있기 때문에, 통합데이터베이스 시스템은 실험시설에 따라 다른 스키마를 사용한다.

4. 결론 및 향후 연구

본 연구는 그리드 서비스와 웹 서비스를 이용하여 원격지에 있는 실험시설의 장비를 가상화하여 내 실험시설처럼 쓰는데 목적이 있었다. 1차적으로 공동실험센터를 프로토타입으로 통합데이터베이스 시스템을 만들었다. 통합데이터베이스 시스템을 통하여 원격지의 실험자와 실험시설간의 협업이 가능하게 되었다.

추후 다른 실험센터들에도 적용시키며 향후 전체 실험시설장비를 총괄할 수 있는 통합데이터베이스 시스템으로 원격지의 실험자가 다양한 실험시설을 사용 가능하도록 개발 중이다.

5. 참고 문헌

- [1] I. Foster, C. Kesselman, J. Nick, and S. Tuecke. "The physiology of the grid: An open grid services architecture for distributed systems integration, open grid service infrastructure" wg, global grid forum, June 2002.
- [2] 함재근, 명춘주, 김형진, 이종숙, "웹 서비스를 통한 그리드의 진화", 2005
- [3] 한국정보산업연합회, 이상호, "그리드 컴퓨팅과 비즈니스 적용전략", 한국정보산업연합회, 2005.
- [4] I. Foster, C. Kesselman, J. Nick, S. Tueckem, "Grid Services for Distributed System Integration" IEEE Computer, 35(6), 2002.
- [5] I. Foster, D. Gannon, H. Kishimoto - Editors, "The Open Grid Services Architecture", GGF OGSA Working Group, 2003.
- [6] 김등근, 이필우, 황일선, "그리드 컴퓨팅", 한국정보과학회 정보과학회지 제 20권 2호, 2002, pp5-7
- [7] 강경우, 박형우, "그리드 연구개발 동향", 한국정보과학회 정보과학회지 제 20권 2호, 2002, pp.22-33
- [8] K. Czajkowski, S. Fitzgerald, I. Foster and C. Kesselman, "Grid Information Services for Distributed Resource Sharing, 2001