

# 그리드3: 데이터 그리드

권기환<sup>0,1</sup> 김복주<sup>1</sup> 한대희<sup>1</sup> 조기현<sup>1</sup> 오영도<sup>1</sup> 손동철<sup>1</sup> 이지수<sup>2</sup>

1. 경북대학교 고에너지물리연구소(CHEP) 2. 한국과학기술정보연구원 슈퍼컴퓨팅센터  
{kihwan<sup>0</sup>, hanbi, cho, son}@knu.ac.kr, {bockjoo, ydoh}@chep12.knu.ac.kr, jysoo@hpcnet.ne.kr

## KNU Experiences with Grid3

Kihwan Kwon<sup>0,1</sup> Bockjoo Kim<sup>1</sup> Daehee Han<sup>1</sup> Kihyeon Cho<sup>1</sup> Youngdo Oh<sup>1</sup> Dongchul Son<sup>1</sup> Jysoo Lee<sup>2</sup>

1. Center for High Energy Physics, Kyungpook National University, Daegu 702-701, Korea  
2. KISTI, Supercomputing Center, P.O. Box 122, Daejeon 305-806, Korea

### 요 약

그리드3는 데이터 그리드로서 고에너지물리를 포함하는, 많은 양의 데이터를 다루는 7개 분야에 컴퓨팅 및 스토리지 리소스를 제공할 목적으로 미국에서 시작되었다. 경북대학교 고에너지물리연구소는 이 프로젝트에 참여하여 그리드3 데이터 그리드를 경북대학교에 구축하고 다른 미국의 26개 사이트와 컴퓨팅 자원을 공유하며 실제 고에너지물리 실험의 데이터 분석에 일조하고 있다. 한국은 27개 Grid3사이트 중 유일하게 북미대륙에 위치하지 않는 사이트로서 네트워크와 데이터 전송의 문제점을 테스트해 볼 수 있는 사이트로서의 의의가 크다. 이 논문에서는 경북대의 그리드3 경험을 바탕으로 그리드3의 구축과정을 비롯한 한 그리드3를 소개하며 한국 데이터 그리드 측면에서 그리드3의 발전방향을 모색해본다.

## 1. 서 론

고에너지물리연구는 가속기를 이용하여 우주 생성 및 진화에 관한 학문을 연구하는 기초과학분야이다. 고에너지물리실험은 검출기 설계, 제작, 신호처리 및 자료 수집, 분석에 이르기까지의 일련의 작업을 가속기가 있는 외국에서 국제 공동연구로 수행한다. 고에너지 물리연구에서는 입자가속기에서 배출되는 막대한 양의 데이터를 저장, 처리하기 위한 계산능력과 대용량의 저장장치를 확보하는데 어려움을 겪어왔다. 또한, 실제 실험에 참여, 연구, 분석하는 인력 또한 전세계에 흩어져 있어 실험의 효율성을 떨어뜨리고 있다. 이러한 어려움을 해결하고자 차세대 인터넷이라 불리는 그리드에 기반을 둔 데이터 그리드를 활용하고자 노력하고 있다.

그리드3(Grid3)[1]는 미국의 iVDGL, GriPhyN, PPDG, USCMS, USATLAS등과 같은 기관들의 공동프로젝트이다. 이 프로젝트의 목표는 실제로 가용한 그리드를 개발하고 통합하여 LHC (Large Hadron Collider)[2]기관에 실제 설치하여 응용하는 것이다. 또한, 이 그리드를 LHC 기관이 아닌 외국에 까지 확대한다는 것이다. 경북대학교 고에너지물리연구소는 iVDGL의 일원으로써 이 프로젝트에 참여하여 실제로 경북대학교내에 그리드3 데이터 그리드를 구축하고 미국의 여러 사이트들과 자원을 공유하며 고에너지물리 실험을 수행하고 있다. 이 논문에서는 경북대학교가 그리드3 데이터 그리드구축에 참여한 경험을 바탕으로 그리드3를 소개하고, 한국 고에너지물리 데이터 그리드 측면에서의 그리드3의 발전방향을 모색해본다.

## 2. 본 론

### 2.1 그리드3 개요

그리드3는 현재 미국 26개 사이트와 한국 1개 사이트(경북대 고에너지물리연구소)로 구성되어 있다. 경북대는 3대의 머신을 그리드3에 할당하여 미국 사이트들과 연동하고 있다. 현재 그리드3는 총체적으로 2000 CPU 이상을 제공하고 있다. 최대 작업 처리량으로 500~900여 개의 작업이 동시에 수행되며, 작업 완성률은 75%에 달한다. 현재 고에너지물리를 비롯한 천문학, 천체물리, 생화학 등 7개 분야의 데이터 분석과 시뮬레이션을 제공한다. 현재 100여명 이상이 등록되어 사용하고 있다.

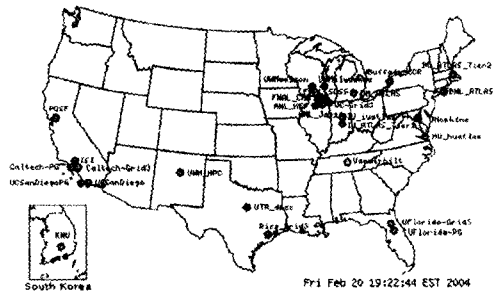


그림 1. 그리드3 사이트: 경북대(KNU)와 미국의 26개 사이트로 구성되어 있다[1].

### 2.2 Grid3의 설치 과정

그리드3의 설치과정은 그리 복잡하지 않다. 그리드3는 최소한의 공통적인 환경을 요구하는데, 바로 그리드3가 설치될 공간과 어플리케이션 및 데이터가 저장되는 장소 그리고 임시 디렉토리 와 같은 것이다. 이것은 어

편 응용프로그램이 똑바로 설치되고 문제없이 실행될 수 있도록 하기 위해 서로 협력해야 하는 부분이다. 그러므로 그리드3를 설치하기 전에 이 네 개의 특별한 디렉토리를 지정해 주어야 하는데, 이 디렉토리들은 헤드노드(작업을 배치시스템에 제출하는 머신)와 워커노드(실제 작업이 수행되는 노드)에서 접근이 가능해야 한다. 그리드3는 어느 특정한 어플리케이션을 위한 그리드가 아니며 사용자는 자신이 사용하고자 하는 어플리케이션을 어플리케이션 디렉토리에 설치함으로써 그리드3 사이트에서 작업을 수행할 수 있다. 기본적으로 그리드3는 PACMAN[3]이라는 유틸리티를 이용하여 간단하게 설치 가능하다.

### 2.3 그리드3 사이트 모니터링

Grid3 사이트의 모니터링은 모나리사(MonaLisa)[4]와 갱글리아(Ganglia)[5]를 통하여 이루어진다. 모나리사는 분산되어 있는 Grid3사이트들을 모니터링 할 수 있는 툴이며 기본적으로 Java JINI 기술을 이용하여 개발되었다. 그림 2는 모나리사의 스크린샷이다.

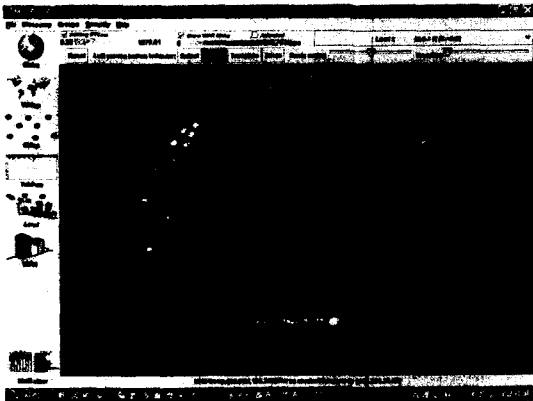


그림 2. 모나리사(MonaLisa)를 이용한 그리드3 사이트 모니터링. 경북대(KNU)는 그림 하단에 표시되고 있다.

각 사이트를 클릭함으로써 각 사이트의 정보를 더욱 자세하게 알 수 있다. 즉 헤드노드의 작업량은 얼마인지, 각 가상기관이 이 사이트에서 얼마만큼의 작업을 수행하고 있는지 알 수 있다. 그리고 각 사이트들 간의 네트워크 정보도 실시간으로 나타난다. 예를 들면, 각 사이트를 연결하는 선은 RTT (Round Trip Time)을 실시간으로 표시한다. 또한, 모나리사는 갱글리아와 연동되어 각 노드에 대해서 자세한 정보를 수집해서 보여준다. 현재 인디아 나대학의 iGOC(Grid Operation Center)에서 갱글리아를 통하여 그리드3 사이트들을 모니터링하고 있다(그림 3 참조).

### 2.4 인증시스템

그리드3에서 이용하는 인증방식은 공개키(PKI) 방식이

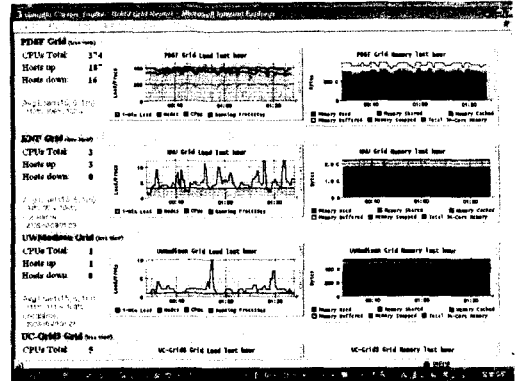


그림 3. 갱글리아(Ganglia)를 통한 그리드3 사이트 모니터링. 경북대(KNU Grid)는 가운데에 나타나고 있다.

다. 그리드 사용자들은 미국 DOE (Department of Energy) 인증기관에서 서명한 인증서를 가져야만 그리드3에 작업을 수행할 수 있다. 그리드3 헤드노드도 마찬가지로 DOE에서 서명한 호스트 인증서를 가져야만 헤드노드로서 기능할 수 있다. 글로벌스에서 사용자들의 인증서를 관리하는 것이 그리드맵파일(grid-mapfile)인데 이것의 문제는 그리드맵파일의 유지보수가 쉽지 않다는 것이다. 새로운 사용자가 생길 때마다 관리자는 그 패키지를 재 설치함으로써 그 새 파일을 다운로드 받아야 했다. 당연히 시스템 관리자 와 그리드맵 관리자는 많은 일거리가 생기게 된다. 이 문제를 해결하기 위해서 그리드3는 VOMS (Virtual Organization Membership Service)[6]를 운영한다. VOMS는 가상기관들이 사용자들을 그룹으로 정돈할 수 있게 해주며 이 정보를 각 사이트에 안전하게 전달되도록 해준다. 그리드3 태스크포스트팀은 각 가상기관마다 따로 VOMS를 운영하며 각 가상기관이 그들 자신의 유저를 등록할 수 있게 하였다. VOMS 소프트웨어의 핵심은 토크 서비스 데이터 베이스이며, 클라이언트는 edg-mkgridmap파일이다. 이것은 데이터를 인증해서 그리드맵파일에 쓸 수 있는 권한을 가지며 규칙적으로 실행된다.

### 2.5 그리드3의 인포메이션 서비스

MDS는 글로벌스에서 제공하는 기본적인 모니터링 툴이다. 그리드3 사용자는 이 MDS를 통하여 필요한 리소스를 검색할 수 있다. 또한 그리드3은 각 사이트의 기본적인 정보를 MDS를 통하여 공표한다. 즉, 어플리케이션을 설치할 수 있는 디렉토리는 어떤 사이트에서 어디에 위치하는지 알려면, MDS 관련 명령을 통해서 찾을 수 있다. 이 외에 이 사이트가 어느 가상기관(VO)에 속하는지 등을 기록하여 공표한다.

### 2.6 그리드3의 성과 및 문제점

총체적으로 볼 때, 그리드의 성과는 미국 26개 사이트

와 경북대 사이트가 참여하여 2300~2700여개의 CPU를 제공하고 500~900여개의 작업이 동시에 수행될 수 있다는 것이다. 경북대 사이트의 경우 다음과 같은 성과가 있었다.

가. 경북대 사이트는 유일하게 북미대륙에 위치하지 않은 그리드3 사이트로서, 진정한 의미의 그리드를 보여 주었다.

나. 현재 고에너지물리 데이터그리드를 구축하고 있는 경북대의 경우 관련 인프라를 재 점검해볼 수 있는 계기가 되었다. 예를 들면, 네트워크의 역질 (reverse mapping) 셋업이 잘못되어 미국의 사용자들이 경북대 머신에 인증할 수 없는 경우가 생겨 작업을 수행하지 못하는 경우가 발생하였다. 그리고 네트워크가 아직은 모든 어플리케이션을 감당할 만큼 빠르지 않아서 몇몇 어플리케이션은 경북대 사이트에서 실행할 수 없었다.

다. 피닉스에서 개최된 슈퍼컴퓨팅2003에서 시연함으로써 경북대 그리드3사이트를 잘 알릴 수 있었다.

그리드3의 운영자들이 밝힌 그리드3의 문제점은 다음과 같이 요약될 수 있다. 이러한 문제들은 크고 작은 그리드를 구성하려는 누구나가 부딪힐 수 있는 문제들이다.

가. 응용프로그램/사이트 디버깅: 모든 응용 프로그램을 어느 정도 안정성있게 돌릴 수 있었으나, 이렇게하기 위해서 각 사이트에서 각 응용프로그램의 문제점들을 해결해야 했다. 이것은 사이트가 굉장히 많을 때는 가능하지 않은 방법이다.

나. 사이트 및 서비스 환경설정: 환경설정 문제 그리고 같은 성격의 문제들을 해결하는데 굉장히 많은 시간이 투입되었다. 핵심 프로그램이 업데이트될 때마다 다시 안정된 시스템으로 돌아갈 때까지 많은 노력과 시간이 걸렸다 (예를들면, VOMS, MDS, MonaLisa).

다. 디스크 공간 할당과 클린업: Grid3는 디스크 관리 서비스를 제공하지 않았다. 특히 SC2003 동안에 많은 작업이 수행되었고, Share 혹은 Temporary 디스크가 아무런 예고없이 가득차버리는 경우가 발생했다. 한시적인 데이터 스테이징 파일들은 어플리케이션에 따라 공간사용량이 굉장히 다르다. 이런 디스크 공간 관리를 시스템 관리자가 직접 수행해야만 했다.

라. 헤드 노드, 즉 Gatekeeper 노드의 부하(Load): 사이트 헤드노드에 굉장히 많은 CPU 부하가 걸리는 경우가 많았다. 이 사이트에 굉장히 많은 작업이 submit 되고 있어서 부하가 많이 걸리는 경우도 있었고, 헤드노드에서 돌아가는 그리드 모니터링 서비스 때문에 그런 경우도 있었다.

마. 리눅스 버전: 리눅스 7.3 보다 낮은 버전과 혼용하여 사용하는 경우 문제가 있었다. 이것은 리눅스 레드햇 버전 문제가 아니라 Glibc 버전, 패키징, shared library 버전, 환경설정 파일의 위치, 리눅스 커널 버전, gcc run time 라이브러리, NIS 사용등과 관련이 있을 것이다.

바. 사이트 공유 시스템 디렉토리 퍼미션: Grid3 사이트에서 어플리케이션 프로그램을 수행하기 위해서 공유 파일 시스템과 구체적인 디렉토리에 write 퍼미션이 필요하다. 이 퍼미션을 안전하고도 바르게 설정하는 것은 실수하기 쉬운 부분이다. 어떤 사이트들은 필요하지도

않는데 세상 모든 사람들에게 쓸 수 있는 권한을 부여하였다.

## 2.7 한국 측면에서의 그리드3의 발전 방향

경북대 사이트의 그리드3에 대한 발전 방향은 다음과 같이 요약될 수 있을 것이다.

가. 네트워크 향상: 현재의 네트워크는 데이터 그리드를 수행하기에 아직 충분치 못하며 업그레이드가 필요한 실정이다. 그리고 네트워크 환경을 올바르게 설정하고 네트워크 퍼포먼스 측정으로 주변의 네트워크 상황을 잘 이해할 필요가 있다.

나. 스토리지 서비스와 데이터관리: 그리드3는 현재 스토리지에 대한 인프라가 미약한 편이다. 데이터 그리드에서 대용량의 데이터 저장공간과 데이터관리는 필수적이며, 그리드와 저장 장치를 잘 접목하는 것은 큰 이슈이다. 현재 경북대의 대용량 테이프 저장자원(43TB)을 SRB(Storage Resource Broker)[7]를 이용해서 통합관리할 예정이다. 이것을 그리드3에 접목할 수 있다면, 그리드3 스토리지 인프라와 데이터 관리에 향상을 가져올 수 있을 것이다.

다. 고에너지물리실험에 참여하는 한국내 다른 대학들과도 그리드3를 설치하여 상호 연동하고 자원을 공유하며, 모나리사나 갱글리아로 자원을 모니터링 해야 할 것이다.

## 3. 결 론

고에너지물리에서는 많은 양의 데이터를 처리하고 세계 각 지에 분산되어 연구하는 연구자들이 쉽게 공동으로 연구할 수 있도록 데이터 그리드를 구성하기 위하여 노력한다. 그리드3는 미국에서 개발된 데이터 그리드로서 경북대 고에너지물리연구소는 ivdgl의 일원으로서 그리드3를 경북대에 구축하여 미국 26개 사이트와 컴퓨팅 자원을 공유하고 있으며, 그리드3상에서 고에너지물리실험을 수행하고 있다.

이 그리드3 프로젝트에 참여함으로써 우리는 현재 우리가 처한 그리드 자원을 다시 한 번 점검할 수 있는 계기가 되었다. 문제점으로 파악된 것은 네트워크 상태가 아직은 좋지 못하다는 것이다. 어떤 어플리케이션의 경우 느린 네트워크 속도 때문에 실행되지 못하는 경우도 있었다. 또한, 기본적으로 DNS의 역질의(reverse query)에 문제가 있어서, 미국의 사용자들이 우리의 사이트에 인증되지 못하는 경우도 발생하였다. 이런 문제들의 원인을 파악하고 해결함으로써 한국 고에너지물리 데이터 그리드 구현에 한층 더 가까워질 수 있을 것이다.

## 4. 참고문헌

- [1] <http://www.ivdgl.org/grid2003/>
- [2] <http://http://lh-c-new-homepage.web.cern.ch/lh-c-new-homepage/>
- [3] <http://physics.bu.edu/~youssef/pacman/>
- [4] <http://monalisa.cacr.caltech.edu/>
- [5] <http://ganglia.sourceforge.net/>
- [6] <http://hep-project-grid-scg.web.cern.ch/hep-project-grid-scg/voms.html>
- [7] <http://www.npaci.edu/dice/srb/>