WLCG/EGEE 티어-2의 설계 및 구축

Design and Construction of WLCG/EGEE Tier-2

곽재혁, 이필우
한국과학기술정보연구원

Jae-Hyuck Kwak, Pillwoo Lee
Korea Institute of Science and Technology Information

요약

WLCG/EGEE 프로젝트는 응용 연구자들에게 지역적으로 분산된 컴퓨팅 그리드 인프라스트럭처를 제공하는 것을 목적으로 한다. 현재, WLCG/EGEE 인프라스트럭처는 40여개국 200여개 사이트의 2만여 개의 CPU 자원과 약 3PB의 디스크 자원을 제공하는 세계 최대의 그리드 인프라스트럭처로서 고에너지 물리학 분야의 초강입자 가속기 실험을 포함하여 9개 응용 연구 분야에서 활용되고 있다. 본 논문에서는 EGEE-II 프로젝트와의 국제 협력의 결과로서 추진된 WLCG/EGEE 티어-2의 설계 및 구축에 대해서 설명한다.

Abstract

WLCG/EGEE project has a main goal of providing application scientists with access to a geographically distributed computing Grid infrastructure. Currently, WLCG/EGEE infrastructure is the world's largest Grid infrastructure, providing over 20,000 CPUs and about 3 PB disk at 200 international sites and used by over 9 application domains including LHC experiments from high-energy physics. This paper addresses the design and construction of WLCG/EGEE tier-2 propelled by international collaboration with EGEE-II project.

I. 서 론

EGEE(Enabling Grids for E-sciencE)[1] 프로젝트는 유럽연합(EU)내의 학계 및 산업계의 다양한 과학 기술 및 비즈니스 분야에서 활용될 수 있는 프로덕션 레벨의 e-Infrastructure를 구축하는 것을 목표로 2004년에 시작되었다. EGEE 프로젝트는 현재 2단계 프로젝트인 EGEE-II가 2006년부터 진행되고 있고, 전세계 32개 국가의 90개 이상의 파트너가 참여하고 US 및 아시아 지역의 국가 그리드 프로젝트와도 협력하고 있다.

본 논문에서는 한국과학기술정보연구원(이하 KISTI)과 EGEE-II와의 국제 협력의 결과로서 WLCG/EGEE 티어-2를 설계하고 구축하였다. WLCG/EGEE 인프라스트럭처는 현재 2만여 개 이상의 CPU 자원과 약 3PB의 디스크 자원을 제공하는 세계 최대의 그리드 인프라스트럭처이며, 2008년 완공 예정인 LHC(Large Hadron Collider) 실험 데이터의 시뮬레이션, 프로세싱, 분석을 위한 컴퓨팅 인프라스트럭처로 활용될 예정이다. WLCG/EGEE 인프라스트럭처는 멀티 티어 구조로 되어 있고, 각 티어마다 책임이 분리되어 있는데, 티어-0는 가속기 데이터 생산과 분산, 티어-1은 분산 데이터 저장 및 처리, 티어-2는 데이터 시뮬레이션 및 분석을 담당하게 된다.

본 논문은 다음과 같이 구성되어 있다. II장은 EGEE-II 프로젝트에 대해서 설명한다. III장은 WLCG/EGEE 티어-2를 설계 및 구축하기 위한 요구 사항을 분석하고, 응용 사례로서 ALICE(A Large Ion Collider Experiment)[2] 실험 환경을 구축하였다. 마지막으로 IV장은 결론과 향후 계획을 제시한다.

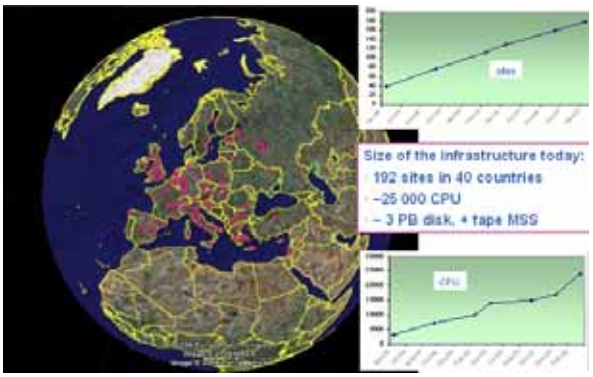
II. EGEE 프로젝트 소개

EGEE 프로젝트는 응용 연구자들에게 일일 24시간 지역적으로 분산된 컴퓨팅 그리드 인프라스트럭처를 제공하기 위한 목적으로 EU 집행위원회의 제6차 프레임워크 프로그램에 의해서 시작되었다. EGEE 프로젝트는 [표 1]에서 보는 바와 같이 크게 네트워킹, 서비스, 공동연구개발 활동으로 나누어 진행되고 있는데, 여기서 SA1이 WLCG/EGEE 그리드 서비스를 운영하는 역할을 하며, JRA1이 WLCG/EGEE 미들웨어인 gLite[3]의 연구 및 개발을 담당하게 된다.

[표 1] EGEE 프로젝트 활동

활동		내용
Networking	NA1	Project Management
	NA2	Dissemination, Outreach and Communication
	NA3	User Training and Induction
	NA4	Application Identification and Support
	NA5	Policy and International Cooperation
Service	SA1	European Grid Operations, Support and Management
	SA2	Networking Support
	SA3	Middleware Integration, Testing and Certification
Joint Reserach	JRA1	Middleware Re-Engineering
	JRA2	Quality Assurance

SA1은 [그림 1]에서 보는 것처럼 현재 2만여 개 이상의 CPU 자원과 약 3PB의 디스크 자원을 제공하는 프로덕션 그리드 환경을 구축하고 운영하며, 24 X 7 지원 서비스를 제공한다. EGEE 사용자 지원은 티어-0인 CERN의 OCC(Operations Coordination Centre)의 조율하에 각 지역의 티어-1인 ROCs(Resouce Operation Centres)에 의해서 제공되며, ROCs는 각 지역의 티어-2인 RCs(Resource Centres)를 통합 및 관리하는 역할을 한다. 아시아 지역의 경우는 대만의 ASGC가 티어-1으로서 APROC(Asia Pacific ROC)[4]를 운영되고 있으며, KISTI는 티어-2로서 WLCG/EGEE 그리드 서비스 환경을 구축하였다. III장에서는 KISTI에 구축된 WLCG/EGEE 티어-2 설계 및 구축 현황에 대해서 알아본다.



▶▶ 그림 1. WLCG/EGEE 인프라스트럭쳐 현황

III. WLCG/EGEE 티어-2 설계 및 구축

1. 요구사항 분석

WLCG/EGEE 티어-2를 구축하기 위한 하드웨어 요구 사항은 [표 2]에서 보는 바와 같이 티어-2가 지원하는 응용 연구 분야에 따라 달라질 수 있으므로 먼저, 응용의 작업 특성을 파악할 필요가 있다.

[표 2] LHC 최소 자원 요구 사항[14]

	ALICE	ATLAS	CMS	LHCb
SE GB/cpu	30	20	50	?
WN Disk GB/job	2.5	2	1	5
WN memory MB/job	600	300~1GB(for pileup at selected sites)	500	500
Longest job(@~2GHz)	8h	24h	72h~1w(for Oscar)	24h
SW installation space(GB)	0.5(in shared area)	15	0.7(prod)~20(analysis) in shared area	0.5

WLCG/EGEE 티어-2를 구축 및 운영하기 위해서는 WLCG/EGEE 미들웨어인 gLite를 설치해야 한다. 현재, gLite는 Scientific Linux(이하 SL)[5] 3.0 환경에서만 설치되므로 SL이 설치된 노드가 준비되어야 하며, 또한, NTP(Network Time Protocol)등을 통한 시간 동기화가 이루어져야 한다.

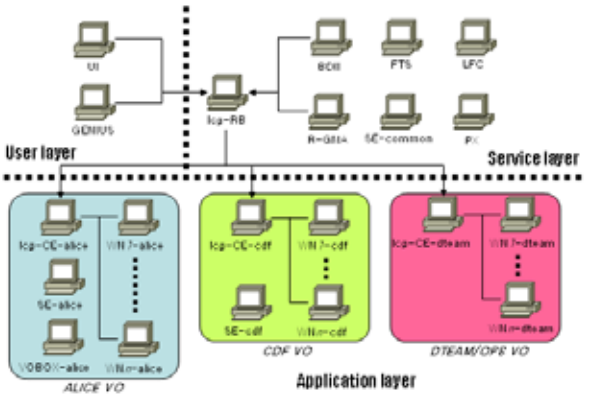
gLite는 기능별 컴포넌트로 분리되어 있어서 티어 구성과 응용 요구 사항에 따라 선택적으로 설치될 수 있으며, 또한, 응용 계산 모델에 따라 gLite와 연동하여 실행되는 응용 미들웨어가 설치되어야 할 수도 있다. [표 3]은 gLite 컴포넌트별 기능과 설치 여부를 정리하였다.

[표 3] gLite 컴포넌트별 기능 및 설치 여부

컴포넌트 타입	기능	티어-2 필수?	KISTI 설치?
UI	사용자 인터페이스	O	O
RB	자원 브로커	O	O
BDII	자원 정보 시스템	O	O
RGMA	자원 정보 시스템	O	O
CE	클러스터 프론트 노드	O	O
WN	클러스터 계산 노드	O	O
SE	스토리지 노드	O	O
FTS	파일 전송 서비스	X	X
LFC	논리 파일 카탈로그	X	O
PX	프록시 서버	X	O
VOBOX	VO-BOX 노드	응용 의존적	O (ALICE)

2. 설계 및 구축

본 논문에서는 WLCG/EGEE 티어-2를 크게 사용자 계층, 서비스 계층, 응용 계층으로 나누었다[그림 2].



▶▶ 그림 2. WLCG/EGEE 티어-2 아키텍쳐

사용자 계층은 응용 연구자가 WLCG/EGEE 인프라스트럭쳐에 접속하기 위한 사용자 인터페이스를 제공하는 계층으로 응용 연구자는 프록시 인증서를 생성하고 작업을 자원 브로커

에 제출하게 된다. 서비스 계층은 응용 연구자가 제출한 작업이 수행되기 위한 인프라 서비스 인터페이스를 제공하는 계층으로 자원 브로커는 자원 정보 시스템을 통해 응용 연구자가 제출한 작업 요구사항에 맞는 자원을 선택하고 작업을 제출하게 된다. 또한, 실험 데이터의 저장 및 관리를 위해 스토리지 노드, 논리 파일 카탈로그 등의 데이터 컴포넌트를 제공한다. 응용 계층은 작업이 실행되는 계층으로 각 응용별로 계산 클러스터, 스토리지, 응용 미들웨어 환경을 포함한다.

WLCG/EGEE 티어-2를 사용자 계층, 서비스 계층, 응용 계층으로 나누고 각 응용별로 자원을 독립적으로 구성한 것은 각 계층별로 업무를 분리하고 운영을 자율화하기 위함이다. 사용자 계층은 사용자 지원 업무에, 서비스 계층은 인프라 서비스 운영 업무에, 응용 계층은 응용 자원 운영 업무에 편입될 수 있으며, 업무의 중복을 피하고 업무별로 권한과 책임이 분리될 수 있다.

본 논문에서는 WLCG/EGEE 티어-2를 구축하기 위해서 다음과 같은 단계를 수행하였다.

(1) gLite 노드 준비

WLCG/EGEE 티어-2를 구축하기 위해서는 먼저 gLite 노드를 위한 하드웨어 자원을 준비해야 하며, KISTI에 설치된 gLite 노드의 하드웨어 사양은 [표 4]와 같다.

[표 4] gLite 노드의 하드웨어 사양

CPU	Intel Pentium4 2.0GHz
Memory	2GB/node
Disk	80GB/node, 500GB external
Network	1Gbps
OS	Scientific Linux 3.0.4
# Compute Nodes	30

(2) WLCG/EGEE 티어-2 등록

WLCG/EGEE 티어-2가 되기 위해서는 먼저 APROC에 새로운 RC로서 등록해야 하며, 이때 사이트 정보, 관리자 정보, 지원 응용 분야, 자원 정보 등이 제공된다. KISTI는 KR-KISTI-GCRT-01이라는 이름으로 가입되었다. APROC에 새로운 RC로 등록되면 GOCDB[6]에도 자동적으로 등록되게 되는데, 이를 통해 WLCG/EGEE 참여 사이트 정보를 전체적으로 확인할 수 있다.

(3) gLite 컴포넌트 설치

gLite는 YAIM[7]을 통해서 설치 및 관리된다. YAIM은 /opt/glite/yaim에 설치되며, etc/site-info.def를 수정해야 한다. KISTI에 설치된 gLite 컴포넌트는 [표 3]과 같다.

(4) WLCG/EGEE 티어-2 인증

새로운 RC 구축이 끝나면, APROC는 EGEE 사이트 인증 과정을 시작하며, 설치된 gLite 컴포넌트에 대한 테스트를 약 1~2주 동안 수행하게 된다. WLCG/EGEE 사이트 인증 과정을 통과하면, 새로운 RC는 인증(Certified)되고, WLCG/EGEE 자원 정보 시스템(top-BDII)[8]에 업데이트되어 WLCG/EGEE 인프라스트럭쳐에 포함되게 된다.[그림 3]

▶▶ 그림 3. WLCG/EGEE 사이트 인증

(5) WLCG/EGEE 티어-2 운영

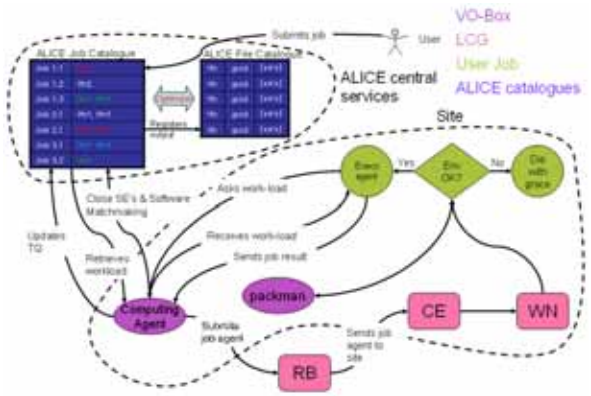
WLCGEGEE 사이트는 GStat[9]과 SAM[10]에 의해서 주기적으로 모니터링된다. 특히, WLCG/EGEE 사이트가 SAM 테스트를 통과하지 못하면, 자동적으로 WLCG/EGEE 자원 정보 시스템에서 일시적으로 제거되어 무의미한 작업 제출을 사전에 방지하게 된다.

3. 응용 사례: ALICE 컴퓨팅 환경 구축

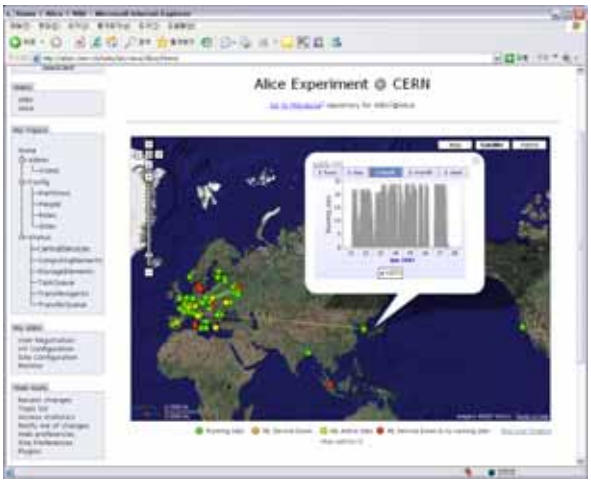
본 논문에서는 WLCG/EGEE 티어-2의 응용 사례로서 ALICE 컴퓨팅 환경을 구축하였다. ALICE 컴퓨팅 환경은 AliEn[11]이라는 응용 미들웨어 기반으로 수행된다. 따라서, gLite와 AliEn은 연동될 필요가 있으며, gLite의 VOBOX가 이를 담당한다. AliEn은 VOBOX와 WN에 의해서 공유되어야 하며, 일반적으로 SE의 /opt/exp_soft/alice/alien에 설치되고 NFS로 마운트하여 공유하게 된다.

ALICE 컴퓨팅 작업은 [그림 4]에서처럼 VOBOX에 설치된 AliEn 에이전트에 의해서 ALICE의 TQ(Task Queue)로부터

패치되며, RB를 통해 자원에 제출된다. ALICE 작업 스크립트는 실행 환경 검사를 통해 ALICE 작업이 실행되기 위한 환경을 동적으로 설치하며, 결과 파일은 AliEn 에이전트에 직접 전달되어 ALICE의 TQ를 업데이트한다.



▶▶ 그림 4. ALICE 컴퓨팅 작업 제출 흐름



▶▶ 그림 5. ALICE 컴퓨팅 모니터링 환경

V. 결론 및 향후 계획

WLCG/EGEE 인프라스트럭쳐는 세계 최대의 인프라스트럭쳐로서 특히, 내년에 LHC가 본격적으로 가동되면 매년 수십 TB~수백 TB의 실험데이터가 생성되어 본격적인 그리드 운영에 들어가게 될 것으로 보인다. 사실, 각국에서 앞다투어 그리드 인프라스트럭쳐 구축을 추진하고 있지만, 베스트 프락티스를 찾기가 어려웠던 것이 사실이다. LHC가 가동되면 WLCG/EGEE 인프라스트럭쳐는 입자물리학 분야에서 전세계적인 데이터 그리드 환경이 구축된 중요한 사례로서 기록될 것이고, 향후에 그리드 기반 응용 연구를 활성화하는데 도움을 줄 수 있을 것으로 기대된다. 이런 점에서 WLCG/EGEE 티어-2를 구축하여 WLCG/EGEE 인프라스트럭쳐에 참여하는 것은 중요한 의미를 지닌다.

현재 KISTI는 CERN과 ALICE 티어-2 센터를 구축하기 위한 논의를 진행되고 있다. 아시아 지역에서는 티어-1인 ASGC를 포함하여 ALICE를 지원하는 티어가 많지 않으므로 ALICE와 연구 협력은 WLCG/EGEE와의 협력에서 보다 실질적인 효과를 기대할 수 있을 것으로 생각된다.

▌참 고 문 헌▐

[1] EGEE-II, http://www.eu-egee.org
[2] ALICE, http://aliceinfo.cern.ch
[3] gLite, http://www.cern.ch/glite
[4] Asia Pacific ROC, http://www.twgrid.org/aproc
[5] Scientific Linux, https://www.scientificlinux.org
[6] GOCDB, http://goc.gridops.org
[7] YAIM, https://twiki.cern.ch/twiki/bin/view/LCG/YaimGuide301
[8] http://grid-deployment.web.cern.ch/grid-deployment/gis/lcg2-bdii/dteam/lcg2-all-sites.conf
[9] GStat, http://goc.grid.sinica.edu.tw/gstat
[10] SAM,, https://lcg-sam.cern.ch:8443/sam/sam.py
[11] AliEn, http://alien.cern.ch
[12] Saiz, P. and others, AliEn - ALICE environment on the GRID, Nucl. Instrum. Meth., A502 (2003) 437-440.
[13] Bagnasco, S. and others, AliEn - EDG Interoperability in ALICE, CHEP 03, La Jolla, Ca, USA, March 2003.
[14] Ian Bird, LCG Minimum Resource Requirements, http://ibird.home.cern.ch/ibird/LCGMinResources.doc
[15] gLite User Guide, https://edms.cern.ch/document/722398/