제 19 회 한국정보처리학회 춘계학술발표대회 논문집 제 10 권 제 1 호 (2003. 5)

그리드 환경을 위한 Grid Configuration System 에 관한 연구

홍원택, 이종숙, 박형우
한국과학기술정보연구원 슈퍼컴퓨팅센터
e-mail : wthong@kisti.re.kr

A Research on Grid Configuration System for Grid Environment

Won Taek Hong, Jong Sook Lee, Hyoung Woo Park
Supercomputing Center, Korea Institute of Science and Technology Information

요 약

그리드 기술은 많은 요소 기술들을 포함하는 특징을 갖고 있기 때문에 복잡한 그리드 환경을 보다 편리하게 구축하고 이용하려는 노력이 계속 진행되고 있다. 본 논문에서는 그리드 미들웨어를 포함한 관련 소프트웨어의 통합 캐기징 기술에 대한 연구 동향을 살펴보고, 국내 그리드 환경을 위한 확장성 있는 Grid configuration system 을 설계해 본다. 제안하는 Grid Configuration System 은 컴포넌트들의 설정 파라미터를 관리하는 유용한 기능들을 제공하여 복잡한 그리드 환경을 일관성있게 관리할 수 있게 해 줌으로써 관리자에게 그리드 환경 관리의 용이함을 제공해 줄 것이다.

1. 서론

과학 연구, 공학 설계, 교육 등의 공통된 목적을 추구하기 위해 분산된 커뮤니티들간의 협업이 이루어져야 한다. 이러한 연구 자원을 제공하기 위해서는 자원 공유라는 근본적인 개념을 충족시켜야 한다. 그리드 기술은 물리적 또는 지리적으로 분산되어 있는 IT 전문가, 첨단 연구장비, 차세대 응용, 그리고 첨단 네트워크를 마치 " 하나의 로컬 시스템처럼(single virtual system)" 활용할 수 있게 하는 고성능 인프라를 구축하는 기술로 요약할 수 있다. 이러한 관점에서 그리드는 현재의 인터넷 기반의 IT 인프라를 한 단계 전진시킬 수 있는 기술로 많은 관심의 대상이 되고 있다.[1] 그리드 환경을 구축하기 위해서는 그리드 미들웨어를 포함하여 추가적으로 필요한 응용 프로그램들의 설치 및 운영이 필요하다. 그리드 환경을 구축하고자 하는 관리자들에게 이러한 일들이 처음에는 어려움으로 느껴질 수 있다. 따라서 그리드 기술 보급과 함께 그리드 환경 구축에 필요한 많은 과정을 보다 편리하게 제공해 주고자 하는 노력이 계속되고 있다. 본 논문에서는 그리드 사용자에게 환경 구축을 용이하게 해주는 그리드 관련 소프트웨어의 통합 패키지에 관한 연구 동향을 살펴보고, NMI-R2 에서 소개된 Gridconfig Tools 을 분석한다. 또한 그리드 관련 소프트웨어의 설정 파일을 일관성 있게 관리할 수 있는 확장된 개념의 Grid Configuration System 을 설계하고 분석해 본다.

2. 통합 패키지 관련 연구

그리드는 많은 요소 기술들을 포함하는 특징을 갖고 있는 기술이므로 그리드 미들웨어를 포함한 응용 프로그램들을 통합적으로 제공하고자 하는 노력이 계속되고 있다. VDT(Virtual Data Toolkit), Grid Starter Kit, NMI 가 그 예가 될 수 있다.

2.1 VDT(Virtual Data Toolkit)[2]

VDT 는 Griphyn project 에 관련된 연구 그룹과 실험의 요구 사항을 지원하는 일련의 소프트웨어 집합체이다. 현재 VDT 는 크게 서버, 클라이언트, SDK (Software Development Kit)의 3 부분으로 나뉘어진다. 서버는 Globus gatekeeper, Condor 와 같은 서버 상에서 요구하는 소프트웨어를 포함하고, 클라이언트는 원격

지 그리드 사이트에서 작업을 수행하기 위해 사용하는 소프트웨어를 포함한다. SDK 는 새로운 소프트웨어를 개발하기 위한 라이브러리들을 포함한다.

2.2 Grid Starter Kit[3]

UK Grid Support Centre 는 UK 그리드 테스트 베드에서의 그리드 미들웨어와 분산 자원 관리의 적용, 운영 및 유지 등을 지원한다. 이러한 지원을 위해 UK 내에서의 계산 그리드 환경을 구성하기 위해 필요한 관련된 문서, 소프트웨어 등을 제공한다. *Grid Starter Kit* 은 UK e-Science Grid product 이고, 주요 그리드 소프트웨어를 제공한다. 주요 소프트웨어는 보안, 작업 관리, 지역 자원 관리를 위한 Globus, Condor, SRB 들이다.

2.3 NMI(NSF Middleware Initiative)[4]

NSF(National Science Foundation)는 응용 과학자 및 연구원들이 Internet 을 사용하여 효과적으로 장비, 자료를 공유하고 동료와 협업이 가능하도록 지원하는 NMI(NSF Middleware Initiative)의 창설을 지원해 주기로 했다. NMI 배포판은 소프트웨어 패키지, 오브젝트 클래스들, 연구 및 엔터프라이즈 컴퓨팅을 지원하는 실제 사례들 등으로 구성된다. NMI 의 이러한 시도는 공개된 소스와 공개된 표준을 사용하여 소프트웨어 패키지를 만들려는 목적에 부합하는 것이다. 통합적으로 배포판에 포함된 소프트웨어와 툴들은 21 세기 협업을 위한 모델이다. GriPhyN, the Network for Earthquake Engineering Simulation(NEES), multi-site TeraGrid supercomputing system 과 같은 대규모의 NSF 과학 및 공학 프로젝트는 머지않아 NMI 를 통해 안정된 소프트웨어 지원을 받게 될 것이다. NMI 에서 사용하는 Gridconfig Tools 은 그리드 관련 소프트웨어 컴포넌트들의 설정을 관리하는 설정 툴의 집합이다. Gridconfig Tools 의 개요는 그림 1 와 같다. DB tables 은 MySQL database 로 구현되고, database API 와 컴포넌트에 관련된 스크립트들은 python 으로 구현된다. MySQL 과 Python 간의 연동을 위한 MySQLdb python 모듈이 추가적으로 필요하다. 설정 파일의 파리미터는 DB table 의 질의를 통한 정보와 GUI 를 통한 사용자 입력을 통해 반영된다. 현재 Gridconfig Tools 을 통해 생성된 설정 파일들은 특정 디렉토리에 위치하게 되어, 관리자가 수동으로 이러한 설정 파일들을 컴포넌트 별로 반영해 주어야 하는 단점이 있다.

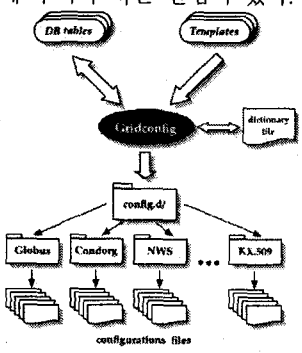

그림 1 Gridconfig Tools 의 개념도

3. Grid Configuration System

3.1 연구동기

현재 KISTI 에서는 그리드 미들웨어 개발, 국가 그리드 기반 구축 등 그리드 관련 과제가 수행되고 있다. 국가 그리드(N*Grid)는 IT + BT, IT + NT, IT + ET, IT + ST, 그리고 IT + 전통산업을 위한 고성능 IT 중심 연구 개발 지원 인프라이다. 그리드 미들웨어는 이러한 인프라를 효과적으로 이용하여 그리드 환경에서 효율적인 원격 고성능 컴퓨팅을 지원한다. 즉, 과학기술분야 전문가들이 분산되어 있는 컴퓨팅 자원, 데이터 자원, 인적 자원 등을 통합하여 연구를 수행할 수 있게 하는 국가 그리드(N*Grid) 환경을 제공하고자 한다. 현재 국내 그리드 테스트베드는 KISTI 를 중심으로 7 개 기관이 참여하여 구축되어지고 있다. 글로버스 기반의 멀티 클러스터 형태 테스트 베드로 제공중이며 사용자 환경은 현재 명령어 기반으로 이루어지고 있고, 포탈 서비스가 테스트되어 서비스될 예정이다. Grid Configuration System 을 설계하기 위하여 먼저 그리드 환경 구축 로드맵을 서술하고자 한다.

그림 2 의 그리드 환경 구축 로드 맵은 그리드 환경을 구축하기 위해 진행되는 과정의 일부이다. 우선 기존의 그리드 미들웨어와 응용 프로그램을 운영 및 시험하고, 영역을 확장하여 클러스터 시스템을 위한 그리드 환경을 구축한다. 이어서 클러스터 시스템을 관리하는 관리자를 위한 관리 툴을 개발함으로써 보다 효율적인 관리 체제를 제공하고자한다. 궁극적으로는 그리드 미들웨어 과제를 통해 개발된 미들웨어 툴킷으로 국가 그리드 기반을 구축하고자하는 것이 본 목적이다. 그러한 단계 중에 시스템 관리자에게 관리의 용이함을 제공해 주기 위해 *Grid Configuration System* 은 필요할 것으로 생각된다. 그리드 관련 소프트웨어 관리 툴에 대한 전체 시스템 설계하고, 시나리오를 생각해 봄으로써 안정적인 그리드 환경을 제공할 수 있을 것으로 기대된다.

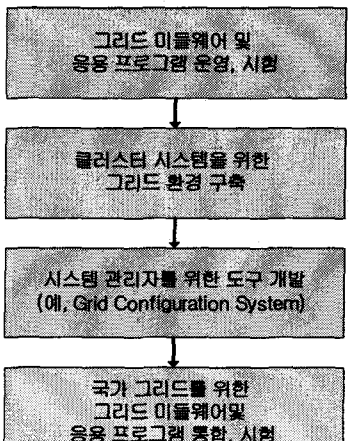

그림 2 그리드 환경 구축 로드맵

3.2 설계 사상

그리드 시스템에서 주요 관리 이슈 중의 하나는 많은 컴포넌트들에 대한 설정 파라미터들을 관리하는 것이다. 설정 파일 자체는 일반적으로 ASCII 형태로 쓰여지지만, 구조는 잘 정의되지 않는다. 또한 파라미터들이 여러 위치에 중복되어 저장된다. 관리자는 중요한 그리드 소프트웨어 파라미터 값의 목록을 작성하거나 일관성 여부를 검증하는 것이 쉽지 않다.

Grid Configuration System 은 설정에 대해 실용적인 방법론을 제공한다. 첫째, 설정 파일의 본래 포맷은 변경되어서는 안된다. 둘째, 관리자는 이러한 포맷 파일을 직접 편집해서는 안된다. 대신 제어되는 데이터베이스로부터 설정 파일을 생성하기 위해 구조(structure)가 요구된다. 데이터베이스가 질의되어 그리드 설정 파라미터가 무엇인지 결정된다. 설정 파일 템플릿과 데이터베이스를 사용하여 Grid Configuration System 은 필요한 설정 파일들을 생성한다. 이 파일들을 변경하기 위해, 단지 DB 의 파라미터를 변경하고, 모든 설정 파일을 재생성하면 된다. 일관성은 이러한 식으로 보장된다.[5]

3.3 제안하는 전체시스템 구조

확장성 있는 Grid Configuration System 을 지원하기 위해 새로운 그리드 관련 소프트웨어 추가 시 전체 시스템에 큰 영향을 미치지 않고 반영될 수 있어야 하고, 변경된 관리 파라미터들을 포함하는 설정 파일들이 자동으로 적용될 수 있어야 한다. 생성된 설정 파일들은 관리자가 수동으로 반영할 필요없이 설정파일의 특성에 따라 분류되어 자동적으로 반영될 수 있는 기능이 필요하다. 또한, Grid Configuration System 의 확장성을 추구하기 위해 관리자가 추가한 새로운 소프트웨어 컴포넌트의 템플릿을 생성할 수 있는 기능이 요구된다.

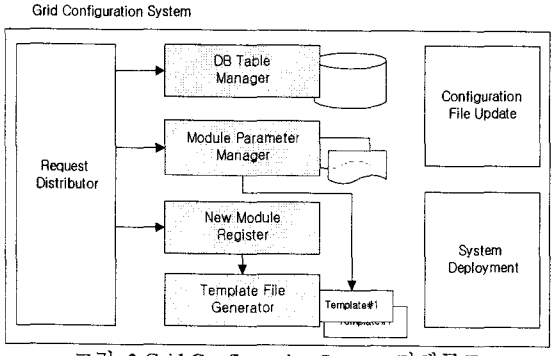

그림 3 Grid Configuration System 전체구조

- Request Distributor : 관리자의 요구사항을 받아 해당 모듈로 분기시켜주는 역할을 한다. DB table Manager, Module Parameter Manager, New Module Register 의 모듈로 분기된다.
- DB Table Manager : 유지하고 있는 DB table 을 관리하기 위한 모듈로서, 각 소프트웨어 컴포넌트들에 대해 Subcomponet, Service, Keyval 의 범주로 나뉘어 정보를 관리한다. 일반적인 DB query 에 대한 기능을 제공하고 DB entry 의 추가, 삭제 기능을 제공한다.
- Module Parameter Manager: 그리드 관련 소프트웨어에 대한 파라미터 설정을 하는 모듈로서 컴포넌트에 따라 설정파일을 생성한다. 이 모듈은 설정 파일을 생성하기 위해 이미 생성된 템플릿 파일을 참고한다.
- New Module Register: 그리드 환경에서 필요한 관련 소프트웨어를 추가할 경우 전체 시스템에 반영될 수 있도록 새로운 모듈을 등록하는 역할을 한다.
- Template File generator: 새로운 모듈이 추가될 경우 설정 파일에 대한 Template File 이 필요하다. 사용자로부터 주요 설정 파라미터 항목을 입력받아 템플릿 파일을 생성한다.

3.4 주요 기능

그림 3 에서 볼 수 있듯이 관리자는 Request Distributor 를 통해 요구사항을 의뢰한 후 Grid Configuration System 내부에서 DB table manager, Module parameter manger, New module register 으로 분기되어 원하는 기능을 수행할 수 있다. 주요 기능은 다음과 같다.

3.4.1 컴포넌트 파라미터 설정

Grid Configuration System 은 각각의 소프트웨어 컴포넌트마다 파라미터를 설정할 수 있게 된다. 관리자로부터 입력 받은 파라미터와 DB table 에 유지하고 있는 Key value 를 이용하여 Grid Configuration System 은 필요한 설정 파일들을 생성해 낸다. Grid Configuration System 은 소프트웨어 컴포넌트들의 설정 파일들을 일관성 있게 관리할 수 있다. 이러한 설정 파일들은 성격에 따라 분류될 수 있다. 예를들어, Globus 의 경우 Security 설정 파일, Startup script 파일, Job manager 설정 파일, 일반적인 Shell script 파일 등으로 분류될 수 있다. 새로운 그리드 관련 소프트웨어 모듈이 추가될 때, 관리 대상이 되는 설정 파일 역시 추가된다.

3.4.2 DB tables 을 통한 정보 추출

일반적인 그리드 관련 소프트웨어 패키지를 기준으로 할 때, DB tables 은 표 1 과 같이 Component, Subcomponent, Service, Keyval 4 개의 table 들로 설계될 수 있다. 이러한 DB table 들을 통해 각 컴포넌트의 서비스에서 사용되는 포트정보, 사용자 정보, 설정 파일의 정보를 일관성 있게 유지하여 추출할 수 있다. 추가적인 변경을 필요로 하는 정보는 DB query 를 통해 유지되는 정보를 변경함으로써 반영할 수 있어서 최종적으로 설정 파일의 생성시 일관성을 유지할 수 있게된다.

표 1 DB tables 의 구성

DB table	구성요소
Component	globus, condorg, nws, gsissh, gpt, kx.509, gridconfig
Subcomponet	gram, mds, gridftp, nws, condorg, gsissh, gpt, kx509, gridconfig
Service	globus-gatekeeper, slapd, gridftpd, condor_master, nws_nameserver, nws_forecast, nws_memory, nws_sensor, slapd-nws, sshd, gpt, kca, kx509, gridconfig
Keyval	port, user, configuration file list

3.4.3 변경된 설정의 적용

관리자의 요구에 따라 변경된 설정 파일들은 컴포넌트 별로 반영하여 그리드 환경에 적용할 수 있어야 한다. 관리자가 일일이 이러한 변경된 부분을 부분적으로 반영하는 하는 것 보다는 설정파일의 특성에 따라 분류되어 자동적으로 반영될 수 있는 기능이 제공될 것이다. 또한, Grid Configuration System 의 확장성을 추구하기 위해 관리자가 추가한 새로운 소프트웨어 컴포넌트의 템플릿을 생성할 수 있는 기능을 제공할 것이다.

4. 결론 및 향후계획

그리드 환경을 구축할 수 있는 통합 패키지을 제공함으로써 분산된 커뮤니티들간의 협업이 보다 수월해지고, 그리드 컴퓨팅을 보다 쉽게 수행할 수 있을 것이다. NMI, Virtual Data Toolkit, Grid Starter Kit 등이 그러한 예이다. 또한 그리드 환경을 지원하는 그리드 관련 소프트웨어 컴포넌트들의 설정 파라미터들을 효과적으로 관리할 수 있는 시스템이 요구된다. 본 논문에서는 확장성 있는 Grid Configuration System 의 구조를 설계하고 기능들을 분석해 보았다. 향후 Grid Configuration System 에 대한 연구를 통해 그리드 환경에서 관리자에게 일관성 있는 관리 기능을 제공해 줄 수 있을 것으로 기대된다.

참고문헌

[1] 2002 년 국가그리드기반구축 최종보고서
[2] Virtual Data Toolkit, http://www.lsc-group.phys.uwm.edu/vdt/
[3] Grid Starter Kit, http://esc.dl.ac.uk/StarterKit/
[4] NSF Middleware Initiative, http://www.nsf-middleware.org
[5] NMI Software Documentation, http://www.nsf-middleware.org/documentation/NMI-R2/0/
[6] I. Foster and C. Kesselman, "The Grid: Blueprint for a New Computing Infrastructure", Morgan Kaufmann, 1998
[7] The Globus Project, http://www.globus.org
[8] Alan Blatecky, " NMI Overview" , http://www.grids-center.org/NET2002.ppt
[9] Tom Garritano, " GRIDS Center Overview " , http://www.grids-center.org/Net2002 Talk 4-18-02.ppt
[10] Ian Foster, " GRIDS Center, Enabling Collaborative Science " , http://www-fp.grids-center.org/GRIDS Overview November 20011.ppt