

Enterprise Grid 환경에서의 Integrated Development Environment Grid Portal 구축 제안

홍필두⁰, 이용우
서울시립대학교 전자전기컴퓨터공학부
iamhpd@metalab.uos.ac.kr⁰, ywlee@uos.ac.kr

Construction of the Grid Portal for the Integrated Development Environment in Enterprise Grid.

Pil-Du Hong⁰, Yong-Woo Lee
Faculty of Electrical & Computer Engineering, The University Of Seoul

요 약

기업의 computing process에서는 기업의 주요업무 및 전략을 분석 정의하여 이것을 computing resource에 실행한다. 이를 위하여, 일반적으로, 여러 개발자들이 공유개발환경을 통하여 프로세스를 개발하고 이를 운영한다. 현재 Enterprise Computing Environment에서는 Grid Computing이 필연적 과제이다. Grid Computing에서, Grid Portal은 통합사용자인터페이스를 제공하여 주는 중요한 요소이다. 그러나, 현재의 대부분의 Grid Portal들은 Job Submit 처리를 중점적으로 구현하는데 주력하고 있기 때문에, 기업 computing process의 중요한 요소인 통합적 개발환경에 대한 지원이 부족하다. 본 연구에서는, 이런 부족한 부분을 개선하여, Enterprise Grid 환경을 위한 통합개발환경의 Grid Portal의 모델을 제안한다.

1. 서 론

그리드 컴퓨팅(Grid Computing)이란, 지리적으로 분산된 고성능 컴퓨터, 대용량 저장장치 및 데이터 베이스, 첨단 실험장비등의 자원들을 고속 네트워크에 연결하여 상호공유, 이용할 수 있도록 하는 기반 구조이다.

Grid Computing은 다양한 형태의 자원들을 공유하여 융통성 있는 공유관계를 설정하고, 공유 자원에 대한 효과적 제어 등을 할 수 있는 가상의 조직(Virtual Organization)이라는 개념으로 볼 수 있으며[1], 위치와 물리적 자원이 분산되어 있더라도 사용자에게는 하나의 View로 인식하여, 보다 편리하게 사용할 수 있게 하자는 관련 연구가 진행되고 있다.

이러한 Grid Computing은 사용목적에 따라 Computation Grid, Data Grid, Access Grid로 나누어진다. Computation Grid는 이 기종 간의 자원을 공유하여, 분산처리를 효율적으로 지원하기 위한 환경을 제공하며, Data Grid는 지리적으로 분산되어 있는 자료를 하나의 Data처럼 사용할 수 있는 환경을 제공한다. 마지막으로 Access Grid는 가상의 협업환경을 제공함으로써, 시간적 공간적 제약을 뛰어넘어, 협업할 수 있는 환경을 제공한다.

이와 같은 Grid Computing에 대하여 수많은 Grid Consortiums, Open Forums과 Grid Conferences가 진행되고 있으며, 계정 자원관리 등을 다루는 Globus와 같은 Grid의 core service의 연구, Volunteer Computing분야로서 Grid Systems의 관련된 연구, MPI등 Grid Programming Environments의 연구들이 다양한 Grid 연구를 통하여 이루어 지고 있다.

지역적 물리적으로 분산 병행 처리되는 Grid는 최종 사용자 또는 운영자에게 Grid Portal을 통하여 Grid Computing resource를 효과적으로 제어 사용할 수 있기 위한 인터페이스를 제공하고 있으며, 이를 구현하기 위한 다양한 Development kit 및 다양한 Test bed가 세계적으로 운영되고 있다.

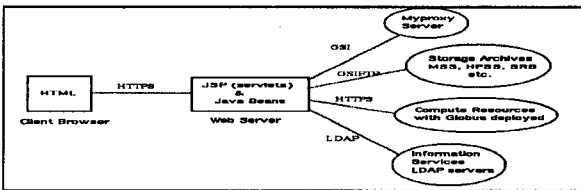
한편 기업의 computing process는 기업의 주요업무 및 전략을 분석 정의하여 이것을 computing resource에 실행하기 위하여 여러 개발자들이 공유개발환경을 통하여 프로세스를 개발하고 이를 운영하는 업무가 일반적이다. 즉 프로세스의 개발에서 운영까지 어떻게 효율적으로 관리 될 수 있는가 하는 문제는 Enterprise Computing에서의 중요한 관점이다. 또한 기업내부의 본사와 지사간, 또한 기업과 기업의 협업환경을 구현하기 위하여 Grid Computing의 도입은 필연적인 과제이기도 하다.

Grid Computing이 Enterprise Computing에 도입 운영되기 위하여는 하나의 인터페이스를 통하여 전체 시스템이 통합 관리될 수 있는 Enterprise개발 환경의 Grid Portal이 구현되어야 한다. 그러므로 본 논문에서는 현재 Grid Portal이 Job Submit resource monitoring만 지원되는데 반하여, 제안하는 Grid Portal은 전체적인 Process의 개발과 운영의 관점에서 Integrated development environment를 지원하는 기능의 구축을 제시하고, 이런 제안을 Seoul IDEG (Integrated Development Environment Grid) Portal로 구현해 봄으로써 유용성을 제시하였다.

2. 기존 Grid Portal의 문제점

Grid Computing에서 분산된 자원을 통합적으로 운용하기 위하여 관리자 및 사용자에게 하나의 One View를 제공하는 실체가 Grid Portal이며, 호주 멜버른대의 G-Monitor라는 시스템 자원 모니터링 중심의 Grid Portal[2], 미국 인디애나 주립대의 과학기술연구중심의 Science Portal[3], 영국 런던대의 XML 기술을 이용하여 Web Service를 바탕으로 data grid를 제어 및 모니터링 하는 Grid Portal[4]의 다양한 예를 들 수 있다. Grid Portal을 구축하기 위하여 다양한 Development Kits이 발표되어 있으며 GPDK(Grid Portal Development Kit)[5]의 사례가 있다. 그밖에 많은 Project차원의 Grid Portal에 대한 연구가 진행되고 있으며 GridSphere Project[6]는 Globus Toolkit 3.X Project와 Web Service 연계 등의 과제를 해결하고자 진행되고 있다. Grid Portal을 구축하기 위하여 반드시 Development Kit을 이용하여야 하는 것은 아니며, Self made로 구현한 사례 등도 많이 있으며, 단 Kit을 사용시 Grid자원의 추가시 Plug & Play기능, 구축시간의 단축 등의 이점을 가질 수 있다.

일반적인 Grid Portal은 분산된 Grid 자원에 대하여 Security, User profiles, Job Submission, File Transfer, Information Service등을 지원하고 있으며 [그림1]은 Grid portal이 어떻게 동작하고 있는지 동작 Mechanism을 보여주고 있다.



[그림1] Grid Portal 동작 Mechanism

이와 같이 관련 연구와 프로젝트에서는 Grid Resource의 통합된 인터페이스를 통하여 Job을 Submission하는 관점에서의 운용적인 측면에 대한 연구가 활발한 편이다. 하지만 실제 프로세스가 설계, 개발, 배포 및 운영 이라는 모든 과정을 지원하는 부분은 아직 미흡하다.

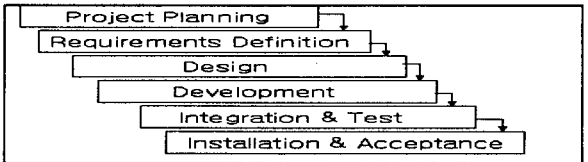
반면 본 논문에서는 기업의 Computing관점에서, 업무분석, 설계, 프로세스구축, Test, 배포, 운용 등 SDLC (Software Development Life Cycle)전반에 걸쳐 Integrated Development Environment를 지원하는 Grid Portal의 구축을 제안하며 기존 Grid Portal과 구분하여 IDEG Portal이라 정의하겠다.

3. Enterprise Grid 환경에서 유용한 Grid Portal의 구축 설계

기업의 있어서 Grid Computing은 필연적 과제이다. 기업의 사례에서 해당 기업의 업종 등 기업목적에 따라 구체적인 사안들은 다를 수 있다. 어떤 은행의 사례로 하나의 예를 들어보자. 은행의 주요 업무는 일반적인 중앙집중형태의 Client/Server Processing이 이루어지고 있다. 하지만 은행에서 경제예측 부서 등에서 경제 지표 및 환율 예측, 이자 예측 등 몇몇 업무는 은행이 보유한 주 전산기(mainframe)하나의 Computing Resource이상을 요구하

는 고성능, 고자원의 Computing Power이 지원된다면 좀더 쉽게 업무를 처리할 수가 있다. 이에 대하여 회사내의 글로벌 네트워크 상 컴퓨팅 자원을 하나로 묶어 Computing Power를 사용한다면 보다 쉽게 문제해결을 할 수 있다. 이와 같이 기업에서도 Grid computing needs는 많이 있으며 이와 같은 사례는 금융계의 국한된 문제가 아니라 일반적 기업의 사례일 수도 있다. 또한 기업과 기업간 Biz to Biz에서도 많은 Grid computing needs를 가지고 있다.

기업의 Enterprise Computing 분야에서는 과학기술 분야와는 달리 기업의 전략에 의하여 수많은 업무가 발생하며 이 업무를 지원하기 위한 Computing지원이 따르게 된다. 이는 기업만의 전략적 핵심업무에 대하여 이를 Computer processing하기 위한 요건의 정의, 설계, 개발, Test및 실무부서에서의 배포 및 교육 등 다양한 수반 업무가 발생하며, 단순히 Process를 수행하는 것뿐 아니라 Process를 설계 구현하는 전반적인 개발과정을 지원해주기 위한 통합적 개발환경이 구축되어 있다. 즉 어떻게 Grid환경의 Computing을 Enterprise환경에서 운영하는가의 문제뿐 아니라, Grid 환경에서 프로세스를 분석하고, 설계하며 구현 및 테스트를 하여야 하는 문제도 고려하여야 한다.



[그림2] Integrated development environment Grid Portal의 SDLC단계별 기능 설계

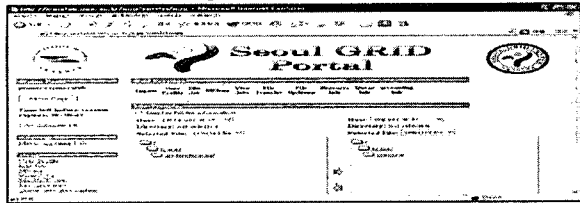
Grid portal은 Grid Computing의 환경이 복잡하다는 문제에서부터, 단순하고 편리한 interface를 제공하여야 한다는 설계의 목표로 삼는다. 이를 위하여 첫째 모든 개발자 및 운영자가 Web이나 PC interface에 익숙하므로 웹이나 일반적 PC Client Program으로 작성되어야 하고, 둘째 Grid를 다룰 수 있는 middle-tier의 infra (예를 들면 Globus)를 지원하여야 하며, 셋째 Grid Portal사용자는 Grid의 자세한 부분이나 자원 기능 등의 자세한 사항에 대하여 반드시 익숙할 필요는 없으며, 이 기존 System이 Grid 자원으로 추가 삭제 되더라도 Plug & Play 방식으로 Grid Portal로 인식되어야 한다. IDEG Portal에서도 위에 제시한 4가지 기본원칙을 가지고 Grid Portal을 설계하였다.

IDEG Portal은 Integrated Development Environment를 위한 다음 기능들을 지원하도록 목표하였다[그림2]. 1) Project Planning을 위하여 원활한 communication을 지원하기 위하여 Grid Portal내의 Access Grid를 연계하였다. 2) Requirements Definition을 위하여 요건자료 요건 자료를 Grid Portal내부에 library화 하였다. 3) Design을 위하여 Grid Portal에 직접 UML등의 설계 프로세스를 지원하는 것보다, Well-defined 되어 있는 MS-Visio등을 ActiveX를 이용하여 상호 연계하였다. 4) Development를 위하여 Web상에서 code-Editor및 mpi-program Helper를 구현하였고, 또한 collaboration developing을

위하여 source-code의 통제 locking을 지원하는 소스대 출승인절차와 반납처리 절차를 구현하였다. 5)Integration & Test를 위하여 통합된 Compiler 수행을 포함하였으며, 디버그 툴을 지원하여 원활한 test를 수행할 수 있도록 하였다. 6) Installation & Acceptance를 위하여 Grid Portal의 통제하에 실행모듈의 형상관리를 하며, GridFTP등을 통하여 실제 운영환경으로의 배포를 수행할 수 있도록 하였다. 배포된 프로세스는 기존 Grid Portal의 Resource Monitoring등을 통하여 운영상태 전반을 관리할 수 있도록 하였다. 이에 대하여 test bed로서 현재 운영중인 서울 그리드 센터의 슈퍼컴퓨터 256노드를 활용하여 Seoul IDEG(Integrated Development Environment Grid) Portal을 구현하였다.

4. Seoul IDEG-Portal의 구현

서울 그리드 센터의 슈퍼컴퓨터는 Pentium IV 2GHz PC 256대로 구성된 Beowulf Style PC Cluster이다. 운영체제로는 RedHat Linux 7.3을 채택하였고 73GB의 스카시 하드디스크를 가진 4대의 master 서버(NFS Server)를 통해 252대의 Diskless 계산노드(Cluster Node)가 Network Booting을 하도록 구성되어 있다.



[그림3] 서울 그리드 센터의 Grid Portal

이를 바탕으로 Seoul GRID Portal이 Monitoring Server와 Web Server에서 운영되고 있다[7]. [그림3] 이를 기본 프레임으로 Seoul IDEG Portal을 구현하고, Integrated development environment를 지원하기 위하여 병렬처리 프로그램 언어인 MPI를 Portal에서 직접 Editing 할 수 있는 Editor, 도움말을 얻을 수 있는 Program Helper, 병렬결과파일 및 디버그가 가능한 compiler 및 dbx모듈 등을 구현하였다. 또한 Portal에서 ActiveX기술을 응용하여 서울 그리드 센터내의 Access Grid Client를 호출하여 연동하였다. 또한 다수 개발자의 소스코드 공유를 위한 제어환경을 연동하여 구현하였다.

5. 기존 Grid Portal과 Seoul IDEG Portal과의 성능 비교

Seoul IDEG Portal은 기업의 Computing 환경에서 개발 및 운영환경에 적합하도록 설계된 Grid Portal이다. 이와 현재 구현되어 있는 일반 Grid Portal과는 많은 기능적인 차이가 있다. 또한 그 기능적 차이는 Enterprise Computing을 수행하는 데 있어서 매우 중요한 기능으로 기존 Grid Portal에서는 간과한 부분이며 보완되어야 할 기능이다. [표1]은 기존 Grid Portal과 본 논문에서 구현 제시한 새로운 Grid Portal의 기능적인 성능을 비교한 도표이다. 이에 결과와 같이 Seoul IDEG-Portal이 기능적으로 우수함을 알 수 있다.

6. 결론

Function	Seoul IDEG-Portal	GPOK Grid Portal	General IDE Tool Kit	Venue AGTK
Account Manager	Yes	Yes	No	No
User Profile	Yes	Yes	No	No
Resource Broker	Yes	Yes	No	No
MPIRun	Yes	Yes	No	No
View job	Yes	Yes	No	No
File Transfer	Yes	Yes	No	No
File Up/Down	Yes	Yes	No	No
Resource Info	Yes	Yes	No	No
Queue Info	Yes	Yes	No	No
Privilege Manager	Yes	No	No	No
Warning Message	Yes	No	No	No
All Node Status	Yes	No	No	No
Source Editor	Yes	No	Yes	No
Help Browser	Yes	No	Yes	No
Compiler	Yes	No	Yes	No
Debug Manager	Yes	No	Yes	No
Venue AG Relay	Yes	No	No	Yes
Quick Bridge AG Relay	Yes	No	No	Yes
VNC Relay	Yes	No	No	Yes

[표1] 기존 Grid Portal과 Seoul IDEG-Portal과 성능비교

본 논문에서 Enterprise Grid 환경에서의 Integrated Development Environment Grid Portal 구축을 제안하였고 이를 Seoul IDEG-Portal을 통하여 가능성을 제시하였다. 이는 효율적인 Grid Portal을 제안하였지만, 실제 환경에서의 구축에서는 보다 많은 변수와 문제점을 가질 수도 있다. 이와 같은 사항들은 향후 과제로 남긴다.

참고문헌

[1] I. Forster, C. Kesselman, S. Tuecke. (2001) ,The Anatomy of Grid
 [2] Martin Placek and Rajkumar Buyya, "G-Monitor: A Web Portal for Monitoring and Steering Application Execution on Global Grids" ,2004
 [3] Dennis Gannon, Middleware Technology to Support Science Portals: a Gateway to the Grid,2003
 [4] Irene Pompili1, Claudio Zunino1, Andrea Sanna1, and Giacomo Piccinelli2, A Web Services based system for data grid,2004
 [5]http://doesciencegrid.org/projects/GPOK/
 [6]http://www.gridisphere.org/
 [7] Yong Woo Lee, Seoul Grid Portal: A Grid Resource Management System for Seoul Grid Testbed, Lecture Notes in Computer Science ,Volume 3251 / 2004