그리드 어카운팅 시스템 설계 및 구현

장경익[0] 이관옥 김법균 황호전 안동언 정성종 *장행진
전북대학교 컴퓨터공학과 *KISTI
{mainclass[0], royaloak, kyun, hjhwang}@duan.chonbuk.ac.kr, {duan, sjchung}@moak.chonbuk.ac.kr,
*hjjang@hpcnet.ne.kr

Design and Implementation for GRID Accounting System

Kyung-Ik Jang[0], Kwan-Ok Lee, Beob-Kyun Kim, Ho-Jeon Hwang
Sung-Jong Chung, *Haeng-Jin Jang
Dept. of Computer Engineering, Chonbuk National University
*KISTI

요 약

인터넷이 발달됨에 따라 분산된 환경에서 각각 컴퓨팅 자원들을 공유하기 위한 기술들이 도입되고 있다. 고속 cpu, 대용량 storage, 실시간 가시화 장비등을 통합하여 하나의 시스템처럼 사용하는 GRID 기술은 차세대 인터넷의 필요충분 조건이다. 본 논문에서는 국가적 GRID 환경을 구축하는데 필요한 기술중 유료화 정책과 관련있는 그리드 어카운팅 시스템을 설계 및 구현에 관하여 기술한다.

1. 서 론

컴퓨터 자원들의 지속적인 초고속화, 대용량화되는 추세에 힘입어 기초과학분야 및 응용 연구들이 가능하게 되었다. 그러나 지금도 생명공학, 유체역학, 기상기후예측 등의 분야에서는 단일자원으로 해결하기 어려운 계산 및 저장자원을 필요로 하고 있다. 따라서 지리적으로 분산되어 있는 컴퓨팅 및 스토리지 자원등을 통합하여 하나의 자원처럼 사용하기 위한 방법에 관한 연구는 필연적이다. GRID는 차세대 인터넷이 추구하는 고품질, 실시간 가시화, 대용량 정보처리 및 협업연구 등이 가능하기 때문에 선진국의 경우 현재 핵심 어플리케이션을 중심으로 고속의 슈퍼컴퓨팅환경과 접목시켜 국가 차세대 인터넷 인프라를 구축하고 있는 추세이다. 국내에서도 현재 그리드 응용 연구가 활발히 진행되고 있다.

GRID 아키텍처 중 상용 서비스를 실시하기 위해서 반드시 필요한 부분이 바로 어카운팅 시스템이다. 어카운팅 시스템은 각 사이트 자원의 상태정보 및 각 자원에 대한 접근 및 사용권한, 해당 사이트에 대한 인증, 사용된 자원에 대한 비용 산출 서비스가 필수적이다.

본 논문에서는 Globus Toolkit[2]을 기반으로 GRID 어카운팅 시스템중에서 remote site에 job을 submission 하는 방법과 job을 수행후 accounting 정보를 수집하는 방법에 관한 구조를 설계 및 구현하였다.

2. 관련연구

2.1 Poland VUS

폴란드의 슈퍼컴퓨팅 센터에서는 Globus 를 기반으로 구현하지 않고 독자적으로 구현한 VUS 시스템은 중간에 Server Machine을 두고 User Machine과 Execution Machine사이에 Virtual User ID를 발급하고 공유함으로서 사용자를 쉽게 구분한다. 이 모델의 경우, 모든 어카운팅 정보를 VUS가 유지 관리하므로 VUS에 부하가 집중된다는 단점이 있어, 이로 인해 파생되는 여러 문제점이 있을 수 있다.

2.2 NASA 의 Template Account

Free Market System을 제안한 T.Hacker가 설계한 모델로 사용자 계정간의 바인딩에 대한 부분을 모델링한 것이다[5]. 자원을 제공하는 사이트에서 template account을 발급하여 실제 사용자의 계정과 일시적으로 바인딩시키는 형식을 취하고 있다.

그러나, 이 모델의 경우 template account가 현재 사용되고 있는 운영체제에서는 지원하지 못하다는 점에서 현실성이 없다. 이 모델이 성공하기 위해서는 운영체제 수준에서 template account를 필요할 때 발급하고 사용되지 않을 때는 삭제하는 기능을 지원해야 한다. 또한 ID에 관한 모델만 제시 했을뿐 비용을 계산하기 위한 어떠한 메커니즘이나 방안도 제시하지 않았다.

3. 어카운팅 시스템 설계

3.1 어카운팅 시스템 구성

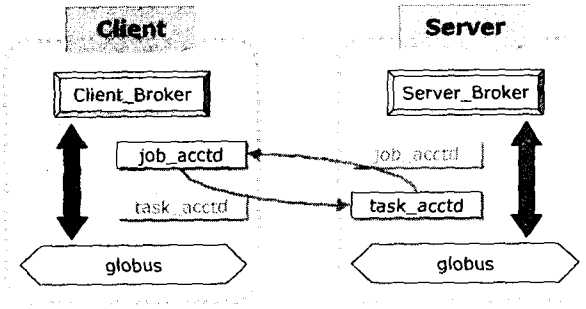
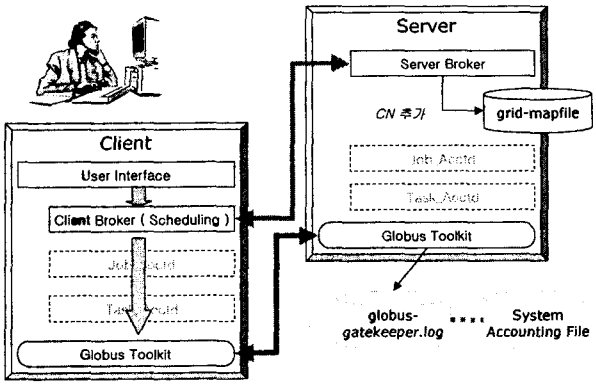


그림 1은 GRID 어카운팅 시스템의 전체적인 구조를 나타내며, 어카운팅 시스템은 job을 submission 하는 Client 와 실제로 수행이 되는 Server로 나누어 진다. 이 시스템은 Client와 Server 양쪽 모두 Globus를 기반으로 하며 상위에 데몬형태로 존재하는 job_acctd, task_acctd가 있으며 Client에는 Client_Broker가 job submission 대행하고 Server에는 Server_Broker가 job 수행을 대행한다.

3.2 Job Submission

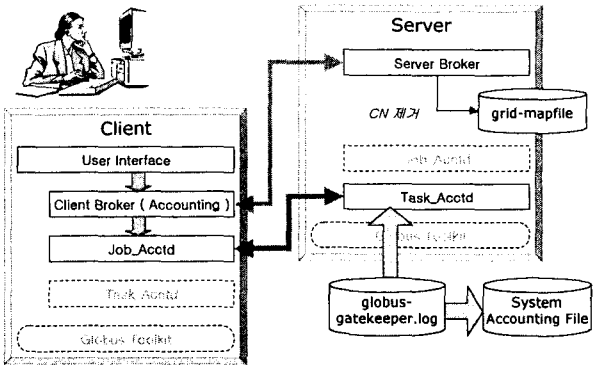
job을 submission 하기 위해서 Client_Broker는 리소르를 제공하는 각 사이트의 Server_Broker와 협상을 한다. 각 리소스 제공 사이트들은 사이트 정책에 따라서 job의 수행 여부를 판단하고 Client_Broker에게 응답을 한다. 일단 Job을 수행하기로 결정한 Server는 수행하기 전에 자신의 grid-mapfile[2]에 사용자의 CN을 등록해 줘야 한다. Client_Broker는 Server_Broker에게 사용자 CN에 관한 정보를 전달하고 Server_Broker는 CN을 받아서 grid-mapfile에 추가한다. 또한 Job이 수행된 후 Client_Broker는 사용자 CN 삭제 요청을 해서 Server 측의 grid-mapfile에서 제거할 수 있다. 그림 3은 이 과정을 나타낸다.



[그림 2] Job Submission

3.3 어카운팅 정보 획득 방법

job을 Submit 할 때 Client_Broker는 항상 로그를 남긴다. 현재 Globus 자체에는 Submit한 정보를 남기지 않기 때문에 Client_Broker에서 로그를 남기도록 설계하였다. job이 수행된 후 Client측의 Job_Acctd는 로그파일을 참고하여 Server측의 Task_Acctd에게 해당 사용자의 어카운팅 정보를 요청한다. 이때 Client측이 Server측에 넘겨주는 정보는 사용자 CN, Host, job submission Time 등이다. 요청을 받은 Server 측의 Task_Acctd는 Client로부터 받은 정보를 참고하여 globus-gatekeeper.log로 부터 시스템에서 Job이 수행될때 사용된 UID/GID와 시간을 획득한다. 그 UID/GID 및 시간을 참고하여 시스템 자체에서 Job을 수행할 때 남기는 로그파일 (예 : pacct file on AIX machine)에서 실제적인 어카운팅 정보를 획득 한다. 그림 4는 이 과정을 보여준다.



[그림 3] Gathering Accounting Infomation

3.4 각 구성요소 요약

3.4.1 Client_Broker

실제적인 사용자의 입력을 받아 job의 scheduling과 submission을 담당하며 사용자의 CN을 execution machine(Server)에 등록/삭제 요청 등을 한다.

3.4.2 Server_Broker

Client의 scheduling이 진행될 때, Client_Broker와 협상을 하며, Client측의 scheduling이 결정되면 Client Broker의 요청을 받아 grid-mapfile에 사용자 CN을 등록 하거나 삭제한다.

3.4.3 Job_Acctd

job이 수행된 후 사용되며 execution machine측의 Task_Acctd에게 어카운팅 정보를 요청하고, 그로부터 받은 어카운팅 정보를 통합 및 사용자에게 보고한다.

3.4.4 Task_Acctd

Client측의 Job_Acctd로부터 사용자 정보를 받아서 그에 해당하는 어카운팅 정보를 분석 및 수집하여 다시 Job_Acctd로 전송한다.

4. 시스템 구현

시스템을 구현하고 테스트 한 기종은 IBM RS/6000 SP 9076-550 머신을 사용했으며 운영체제는 IBM AIX 4.23이다. 이식성이 강한 Java를 사용했고 1.1.6버전으로 구현하였다[1]. Server측과 Client측 모두 로그파일을 남기고 사용하기 위한 데이터 베이스를 구축하기 위해 MySql 3.23.52 소스파일을 컴파일 해서 사용했다. 특히 Globus를 이용한 서비스 프로그래밍을 쉽게 할 수 있도록 개발된 API인 Cog Kit[9]을 사용하였다. Cog Kit은 Java에서 Globus의 거의 모든 함수를 사용할 수 있도록 되어 있고 그에 따른 GUI도 제공한다. 그래서 좀 더 발전적인 GRID 서비스 프로그래밍을 할 수 있도록

도와준다. 마지막으로 Globus는 버전 1.4를 사용했다.

구현하는데 있어서 Java로 리모트 프로세스를 호출할 수 있도록 하기 위해 RMI 기술을 적용했다. 로그파일의 경우 IBM AIX 머신은 job이 수행된 후 표 1에 나열된 파일들에 로그가 남는다.

[표 1] 어카운팅 로그에 관한 파일

/var/adm/wtmp	connection 어카운팅 정보
/var/adm/pacct	process 어카운팅 정보
/var/adm/acct/nite/tacct	disk 어카운팅 정보
/var/adm/qacct	queue 어카운팅 정보

[표 2] Task_Acct가 생성하는 정보

자료형	이름	내 용
uit_t	ta_uid	user id
char	ta_name[8]	login name
float	ta_cpu[2]	cpu time(mins)
float	ta_kcore[2]	kcore-mins
float	ta_io[2]	chars xferred (KB)
float	ta_rw[2]	blocks read/writen
float	ta_con[2]	connect time(mins)
float	ta_du	disk usage
long	ta_qsys	queueing sys charges(pgs)
float	ta_fee	fee for special services
long	ta_pc	count of processes
unsigned short	ta_sc	count of login sessions
	ta_dc	count of disk samples

표 1에서 열거된 파일들은 Task_Acctd가 Job_Acctd의 어카운팅 정보 요청에 의해 어카운팅 정보를 분석하기 위해서 사용되는 로그파일들이다. Task_Acctd는 이들을 이용해 어카운팅 정보를 완성한 다음 정보를 요청했던 사이트로 보내준다. 이 어카운팅 정보의 구조는 표 2와 같다. 이 구조는 Client의 Job_Acctd가 유지하는 정보와 같다. 또한 Client측은 여러 사이트에서 보내져 오는 이러한 정보를 통합하여 전체 어카운팅 정보를 생성하여 유지한다.

5. 결론 및 향후 연구과제

본 논문에서는 어카운팅 시스템을 실제 설계 및 구현하는데 있어서 모든 프로그램을 java를 사용하여 구현하였다. 그렇기 때문에 가상머신이 존재하는 모든 OS에서 사용할 수 있으며 그만큼 이식성이 높다. 또한 어카운팅 시스템에서 각 로컬의 리소스 사용정책에 따라 협상하는 방법을 사용하기 때문에 각 사이트의 특징을 최대한 반영할 수 있었다.

현재 설계되고 구현된 어카운팅 시스템에서 예상되는 문제점은 서로 다른 Client로부터 2명이상의 사용자가 동시에 job을 submit하여 수행했을 경우 시스템 어카운팅 정보에서 각 레코드를 구분할 수 있는 방법이 모호하다. 이를 해결하기 위해 Server_Broker가 grid-mapfile에 사용자 CN을 등록할 때 2명이상의 사용자가 동일한 로컬 UID에 매핑되지 않도록 각각의 사용자를 구분하여 등록하는 방법이 있어야 한다.

향후 어카운팅 시스템에서 보완하고 연구해야할 과제는 이기종 간의 어카운팅 메커니즘과 Client 와 Server가 통신을 하는 동안 발생할 수 있는 보안문제등을 해결해야 할 것이다.

6. 참고 문헌

[1] I. Foster, C. Kesselman(eds), Q.677, "The Grid : Blueprint for a New Computing Infrastructure" Mogan Kaufmann Publishers, 1998.

[2] I. Foster, et al, "Globus: A Metacomputing Infrastructure Toolkit" Intl. J. Supercomputer Applications, 11(2), 1997. pp115-128

[3] K. Czajkowski, I. Foster, S. Fitzgerald, "Grid Information Service", 2001.

[4] M.Lawenda, N. Meyer, Q.31, "VUS specification", Poznan SuperComputing and Networking Center March. 2001.

[5] Thomas J. Haker, Brian D. Athey, Q.16,"Account Allocations on the Grid", Center for Parallel Computing University of Michigan. 2000.

[6] Auditing and Accounting on AIX SG24-6020-00 Redbook,published October-24-2000, http://www.ibm.com/redbooks

[7] http://www.girdforum.org

[8] http://www.globus.org

[9] http://www.cogkits.org

[10] http://www.gridforumkorea.org