

TIGRIS 상에서 작업 관리 서비스의 설계 및 구축

A Design and Implementation of Job Management Service on TIGRIS

명훈주, 이필우
한국과학기술정보연구원

Hun Joo Myung, Pilwoo Lee
Korea Institute of Science and Technology Information

요약

그리드란 지역적으로 분산되어 있거나 혹은 다양한 조직(organization)에 소속된 여러 자원들을 네트워크로 연동하여 사용자에게 하나의 자원처럼 이용할 수 있는 인프라스트럭처이다. 전통적으로 그리드 상에서 작업 관리 서비스는 작업을 제출하고 모니터링 하는 중요한 역할을 하고 있으며, 자원 정보 서비스와 데이터 전송 및 관리 서비스와 더불어 중요한 요소 중의 하나이다. 본 논문에서는 TIGRIS 상에서 작업 관리 서비스의 설계 및 구현 방법을 제시한다.

Abstract

Grid is infrastructure that connect distributed resources or ones in various organizations and enable user to use them as a virtual enormous resources. Traditionally, Job Management Service in Grid plays a role of job submission window and job monitoring and a important part of Grid Services with Resource Information Service and Data Transfer & Management Service. In this paper, we describe a design and implementation of Job Management Service on TIGRIS.

I. 서론

그리드란 지역적으로 분산되어 있거나 혹은 다양한 조직들(organization)에 소속된 여러 자원들을 네트워크로 연동하여 사용자들이 하나의 자원처럼 이용할 수 있게 해 주는 인프라스트럭처(infrastructure)이다[3]. 이 인프라스트럭처를 구성하는 주 요소 중의 하나가 그리드 미들웨어인데, 일반적으로 그리드 미들웨어가 제공하는 기능은 크게 보안, 자원 할당 및 작업 실행, 자원 정보 검색 및 모니터링, 데이터 전송 및 관리 등으로 나누어진다. 이 가운데 자원 할당 및 작업 실행 서비스는 사용자에게 작업을 제출할 수 있는 창구 역할을 할 뿐만 아니라 작업을 수행할 수 있는 자원을 할당하고 자원에서 수행되는 작업을 모니터링하는 중요한 역할을 하고 있다.

KISTI 슈퍼컴퓨팅센터에서는 2008년도에 도입되는 슈퍼컴퓨터 4호기를 비롯하여 기관 내 슈퍼컴퓨터들을 그리드 인프라로 편입시키고, 기존의 슈퍼컴퓨터 사용자에게 그리드 사용자 환경을 제공하기 위한 TIGRIS (Tera-scale Infrastructure for K* Grid Service) 프로젝트가 진행 중에 있다. TIGRIS는 KISTI 슈퍼컴퓨터들 뿐만이 아니라, 추후 서울대, 부산대 등의 국내의 여러 슈퍼컴퓨터들을 포함시켜 테라 플롭스급의 성능을 갖게 되는 계산 그리드를 구축할 계획이다.

본 논문에서는 TIGRIS 상에서 제공되는 작업 관리 서비스들의 기능을 정의하고 아키텍처를 설계하며 구현 이슈들에 대

해서 기술한다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 TIGRIS 상에서 작업 할당 및 작업 제출에 관련된 요구사항을 정리하고 그와 관련된 기존 연구들을 살펴보고, 3장에서는 작업 관리 서비스의 기능 정의 및 설계를 제시하며, 4장에서는 결론을 맺는다.

II. 기존 연구

1. TIGRIS 상에서의 요구사항

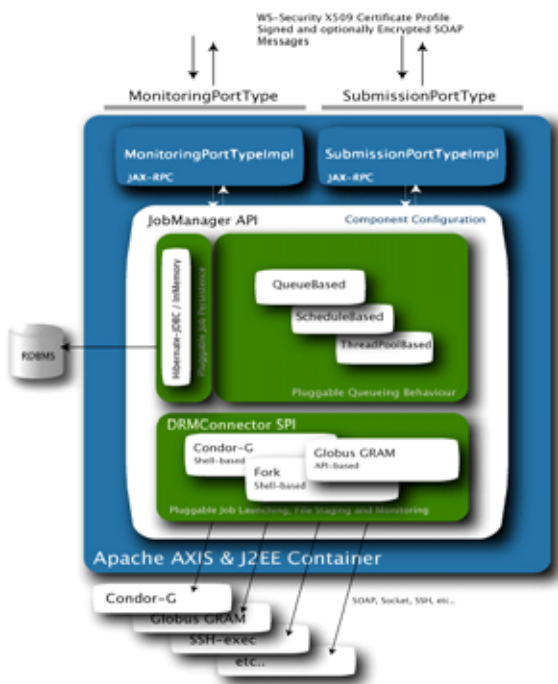
TIGRIS 프로젝트의 요구사항 정의 단계에서 작업 할당 및 작업 제출 기능과 관련하여 정리된 요구사항들은 다음과 같다.

- TIGRIS에서 수행되는 작업의 종류는 크게 3가지이다- 단순 작업, Grid-MPI 작업, HTC 작업(high throughput computing) 작업. 이와 같은 여러 종류의 작업을 수행 시킬 단일 창구가 필요하다.
- 사용자가 편리한 작업 수행을 할 수 있도록, 작업을 수행 시킬 적합한 자원을 찾아내어 수행시키는 글로벌 스케줄러(Global scheduler)가 필요하다.
- TIGRIS 프로젝트에서 새로 개발할 파일 관리 서비스와 연동하여 파일 스테이징(file staging) 등 다양한 기능을 제공한다.

2. 관련 연구

2.1 Gridsam

Gridsam[5]은 영국 e-Science 프로젝트에서 진행 중인 오픈 소스 그리드 미들웨어 OMIII에 포함되어 있는 서비스 중의 하나이다. Gridsam은 웹 서비스 (Web Service)를 인터페이스로 하여 작업 제출 및 모니터링을 제공하는 서비스이며, JSDL[10] 언어 및 WS-I+를 지원한다. 또한 모듈 단위의 설계(Modular design)를 통하여 GT2.4,[4] Condor[3], Grid Engine 6[8] 등 여러 스케줄러를 위한 플러그인과 연동이 가능하다.



▶▶ 그림 1. Gridsam의 구조

Gridsam에서는 사용자 수준에서 단계별로 적절한 작업 상태 정보를 알 수 있으며 데이터베이스를 이용하여 작업 정보를 저장하여 작업 복구 기능을 제공한다. 그러나 현재 GT4와 연동 가능한 모듈이 존재하지 않아 GT4와 연동이 불가능하다.

2. Condor-G

Condor-G[7]는 Condor 프로젝트와 글로버스 프로젝트에서 나온 기술들을 결합한 것으로서, 강력한 태스크 브로커(Task broker)기능을 제공하며 실제 여러 그리드 커뮤니티에서 널리 쓰이는 소프트웨어이다. Condor-G는 분산되어 있는 사이트에서 수행되는 수천 개의 작업을 관리할 수 있으며, 작업 모니터링, 로깅(logging), 결함 내성, 크리덴셜 관리 등을 제공한다. 또한 매치메이킹(matchmaking) 방식의 메타 스케

줄링을 기능을 제공하며, DAGMan를 이용하여 워크플로우 기능도 제공한다. 그러나 Condor-G와 WS-GRAM 연동할 경우에 Condor-G의 로깅 정보는 단지 Condor-G에 해당되는 정보만 기록될 뿐이지, 하위 레이어인 WS-GRAM에서 발생하는 유용한 로깅 정보는 Condor-G에 전달되지 않아, 문제 발생 시에 사용자가 참고할 만한 정보를 찾는 데 어려움이 있다.

III. 작업 관리 서비스 기능 정의 및 설계

1. 기능 정의

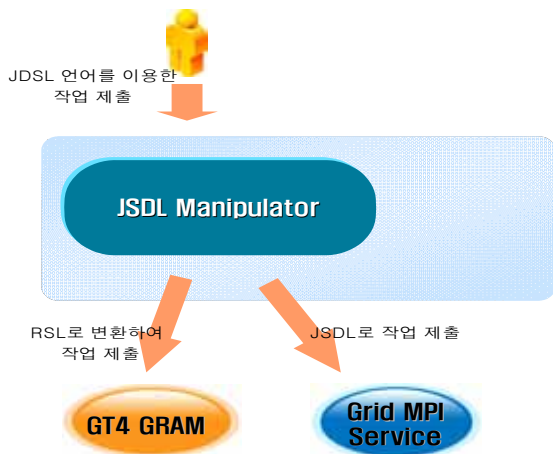
관련 연구에서 살펴본 바와 같이 TIGRIS 상의 요구사항들을 기존의 그리드 관련 소프트웨어들이 충족시키지 못하는 단계로 작업관리서비스를 개발하기로 결정하고 다음과 같이 기능을 정의하였다.

- Grid MPI 작업, HTC 작업, 단순 작업 등을 제출할 수 있는 단일 창구 기능을 제공하고, 작업 기술 언어를 GGF 표준안인 JSDL (Job Submission Description Language)로 사용한다.
- 파일관리 서비스와 연동하여 작업 수행 시 파일 스테이징 기능을 제공한다.
- 사용자의 크리덴셜(credential)로부터 생성된 프락시(proxy) 유효기간에 구애받지 않고 장시간 걸리는 작업 수행을 지원한다.
- 글로벌 스케줄러와 연동하여 사용자의 요구사항을 받아 수행할 자원을 선택하여 자동적으로 작업을 수행하는 기능을 제공한다.

2. 설계상의 이슈들

2.1 작업 창구 역할 기능

위에서 언급한 대로 작업 관리 서비스는 단순 작업, HTC 작업, Grid MPI 작업에 대해 창구역할을 한다. Grid MPI는 TIGRIS 프로젝트에서 자체 개발하고 있는 서비스로, KISTI에서 세계 최초로 개발한 내성 결함 MPI 라이브러리를 비롯한 여러 MPI 라이브러리들을 GT4 기반에서 수행 가능하게 한다. 이 서비스는 JSDL를 작업 기술 언어로 채택하고 있으며, HTC 작업, 단순 작업은 WS-GRAM에 의뢰해 수행되게 되므로, 이와 관련한 메카니즘은 다음 그림과 같다.



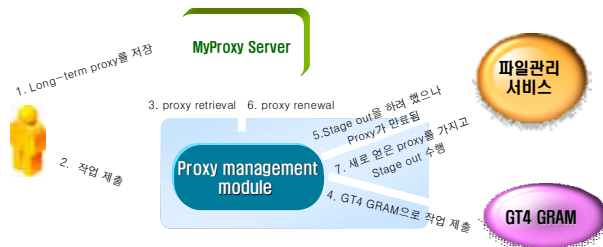
▶▶ 그림 2. 작업관리서비스의 작업 제출 창구 기능

2.2 파일 스테이징 기능을 제공

작업 관리에서 파일 스테이징 기능은 WS-GRAM과 유사하나 rft (reliable file transfer)서비스를 이용하지 않고 TIGRIS 프로젝트에서 개발하고 있는 파일 관리 서비스를 이용해 이루어진다.

2.3 장시간 작업 수행

일반적으로 글로벌스로 구축된 그리드 환경에서 생성된 프락시의 유효기간의 기본 값은 12시간이다. 이 기본 값은 수개월 이상이 걸리는 계산과학자들의 응용을 수행하기에는 턱없이 부족한 값이다. 이에, TIGRIS에서는 MyProxy[9]의 재생 (renewal) 기능을 이용하여 최대 1년까지 (크리덴셜의 유효기간이 1년일 경우) 작업을 수행시키고 결과 데이터를 원하는 위치에 전송할 수 있다.

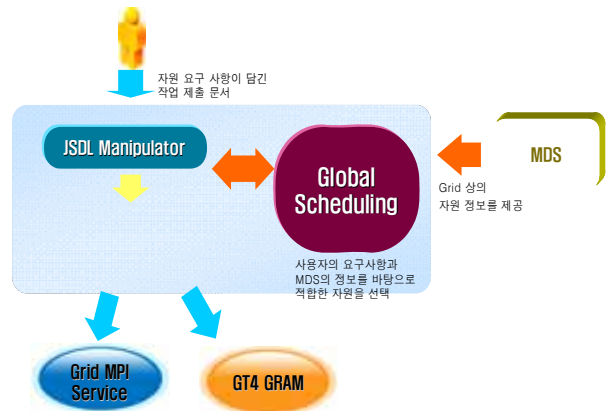


▶▶ 그림 3. 작업관리서비스의 장시간 작업 수행 기능

2.4 글로벌 스케줄러와의 연동

TIGRIS에서는 그리드 사용자들이 작업 제출하는데 있어 보다 편리하게 할 수 있도록 글로벌 스케줄러를 제공하고 있다. 작업 관리 서비스는 사용자로부터 자원 요구 사항이 담긴 JDSL 문서를 받으면, 이 문서를 다시 글로벌 스케줄러에게 넘겨서 이에 적합한 자원을 추천받고, 이를 토대로 WS-GRAM

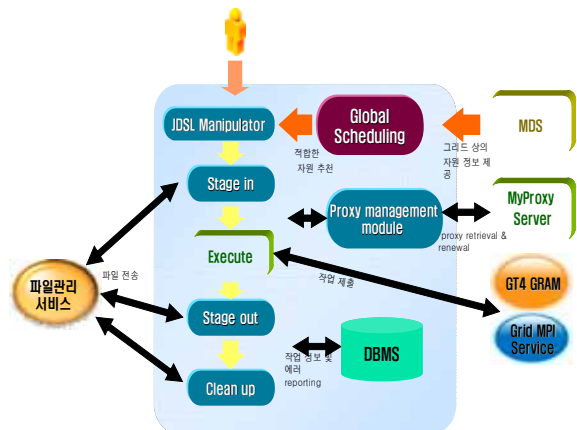
혹은 Grid MPI 서비스에서 작업을 수행하게 한다.



▶▶ 그림 4. 작업관리서비스와 글로벌 스케줄러와의 연동

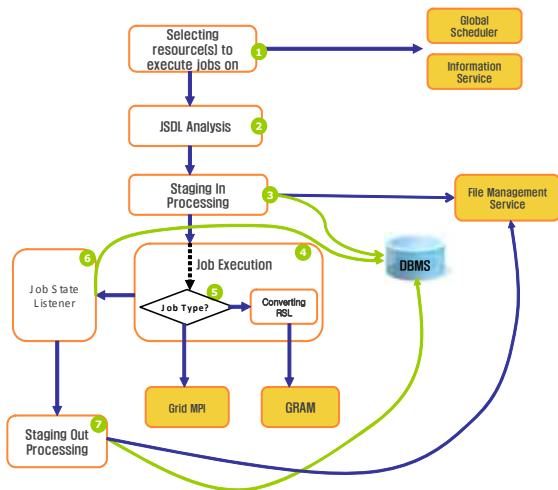
3. 작업관리서비스 아키텍처

위에서 언급한대로 작업 관리 서비스는 그리드의 거의 모든 서비스들과 연동하여 동작한다.



▶▶ 그림 5. 작업관리서비스의 구조

작업 수행 전후로 이루어지는 파일 스테이징 인/아웃은 파일 관리 서비스와 연동하여 이루어지고, 글로벌 스케줄러와 연동하여 사용자에게 자원 추천부터 수행까지 자동적으로 이루어지는 기능을 제공한다. 또한 위에서 언급한대로 MyProxy의 재생 기능을 이용하여 장시간 수행 작업을 지원하며, DBMS를 이용하여 작업관리서비스가 동작하는 단계 혹은 컴포넌트에서 발생하는 정보들을 저장하고, 필요에 따라 사용자에게 정보를 제공한다.



▶▶ 그림 6. 작업관리서비스의 동작흐름도

[10] <http://www.ogf.org/documents/GFD.56.pdf>, JSDL specification v1.0

IV. 결 론

앞에서 살펴본 바와 같이 우리는 성공적인 TIGRIS 구축 및 서비스를 위해 요구사항을 수집하고, 그에 따라 작업관리 서비스의 기능을 정의하고 아키텍처를 설계했으며 이에 따라 구현 중에 있다. 작업 관리 서비스를 비롯한 TIGRIS의 기본 서비스들을 2007년 말에 구현 완료할 예정이며, 통합 테스트를 거쳐 그리드 인프라에 올라갈 예정이다. TIGRIS 프로젝트는 단순 연구에서 끝나는 선도 연구 프로젝트가 아니라, 실제 슈퍼 컴퓨터들을 한데 묶어 그리드를 서비스하게 되는 프로젝트들이다. 이들 서비스들이 제공하는 기능도 중요하지만, 얼마나 안정적으로 서비스할 수 있느냐가 더 큰 이슈가 될 것이다. 이것은 서비스 개발 완료 후 충분한 통합 테스트 시간을 가지면서 안정적인 그리드 서비스 개발에 있어서 유의하여야 할 점들을 정리하게 될 것이다.

■ 참고 문헌 ■

- [1] S. Tuecke, K. Czajkowski, I. Foster, J. Frey, S. Graham, C. Kesselman, T. Maguire, T. Sandholm, P. Vanderbilt, and D. Snelling, "Open Grid Services Infrastructure (OGSI) Version 1.0.", Global Grid Forum Draft Recommendation, 6/27/2003.
- [2] I. Foster, and C. Kesselman, "The Grid: Blueprint for a New Computing Infrastructure", Morgan-Kaufman, 1999.
- [3] <http://www.cs.wisc.edu/condor/>
- [4] <http://www.globus.org>
- [5] <http://gridsam.sourceforge.net/2.0.1/index.html>
- [6] <http://www.omii.ac.uk/>
- [7] <http://www.cs.wisc.edu/condor/condorg/>
- [8] <http://gridengine.sunsource.net/>
- [9] <http://grid.ncsa.uiuc.edu/myproxy/>