

그리드 환경에서 분산 포탈의 설계

허대영^o, 심규호^{*}, 황선태^{*}, 정갑주^{**}, 박형우^{***}

^o국민대학교 컴퓨터학부, ^{**}건국대학교 인터넷미디어공학부 ^{***}한국과학기술원
{phinoccio^o, simsaint^{*}, sthwang^{*}}@cs.kookmin.ac.kr {jeongk^{**}}@konkuk.ac.kr {hwpark^{***}}@kisti.re.kr

Design of Distributed Portal in GRID

Daeyoung Heo^o, Gyuho Sim^{*}, Suntae Hwang^{*}, Karpju Jeong^{**}, Hyoungwoo Park^{***}

^{*}Department of Computer Science, Kookmin University

^{**}College of Information and Communication, Konkuk University

^{***}Korea Institute of Science and Technology and Information

요 약

현재 응용 그리드와 그리드 인프라는 매우 밀접한 관계를 가지고 존재하여 있다. 이것은 개발하고 사용하는 데 있어 방해요소가 된다. 응용 그리드를 사용하기 위해 그리드 인프라에 대해 알아야 하고, 그리드 인프라를 구축하기 위해 응용 그리드를 알아야 하는 것이다. 포탈은 응용 그리드와 그리드 인프라 사이의 경계를 구분지어 주고, 양 진영을 중재해야 한다. 응용 그리드에서 이루어질 수 있는 서비스를 통합하여 일관적인 서비스를 포탈에서 제공하고자 한다. 하지만, 그리드 서비스는 데이터 전송이 필요한데, 웹 환경(WAN)에서는 대화식 작업의 응답 속도가 떨어져 작업 효율을 떨어뜨린다. 이러한 단점을 웹 환경에서 극복하고자 포탈의 일부 기능을 사용자의 지역 환경에서 서비스할 수 있는 방법을 제안한다.

1. 서 론

여러 곳에 분산되어 있는 컴퓨팅 자원을 이용하여 많은 양의 계산 량과 데이터를 필요로 하는 문제를 해결하는 그리드 응용프로그램은 웹 기반의 포탈 사이트를 통해 배포하는 방법이 효과적이다. 포탈 사용자는 웹이 가지는 영향력과 플랫폼에 독립적인 면, 그리고 분산된 자원관리가 용이하다는 점에서 손쉽게 접근할 수 있는 단일의 중앙 응용프로그램 서버에 무수히 많은 그리드 응용프로그램을 배포하여 서비스를 제공할 수 있는 메커니즘을 지원한다.[1]

본 논문은 사용자가 웹 환경을 통해 작업을 제어하고 감시하며, 그리드 응용프로그램과 함께 수행될 수 있는 통합된 환경의 포탈을 제안한다. 포탈에서 작업을 생성하고 수행할 때에는 대화식 작업은 WAN환경을 거쳐 사용자가 원하는 응답이 전달되기 때문에 작업 효율이 떨어진다. 사용자의 환경에서 서비스를 제공하여 작업 효율을 향상시켜 전체적인 성능을 높이고자한다.

2. 관련 연구

2.1. PSE(Problem Solving Environments)

PSE 설계를 위해 영향을 가지는 다음과 같은 목록을 언급한다.[2]

문제 지향성. PSE는 사용자가 컴퓨터에 대한 전문지식이 필요가 없게, 응용에 특화되도록 해야 한다.

통합성. PSE는 통합된 방법으로 문제와 솔루션의 이질성을 다루어 사용자가 사용할 수 있게 해야 한다.

협동성. 물리적으로 분산되어있는 참여자들과 함께 협업하는 데에 의사소통이 가능한 환경을 제공할 수 있는 솔루션을 제공해야 한다.

분산성. 분산된 자원을 공유할 때, 생기는 충돌을 해결하고 같은 자원을 사용하는 것처럼 해주어야 한다.

영속성. PSE는 오랜 시간동안 중단되지 않고 계속 되도록 내부적으로는 체크포인트를 하여, 중단되면 자동으로 다른 시스템에서 문제 해결 수행을 해야 한다.

유연성. 문제 해결 솔루션단위들을 조합하거나, 새로 만들어 낼 수 있어야 한다.

그래픽. 작업의 상태나 문제 해결 수행의 입력, 출력에 대한 값을 그래픽과 같은 것으로 보여줄 수 있어야 한다.

2.2. GENIUS Grid Portal

GENIUS(Grid Enabled web eNvironment for site Independent User job Submission)[3]는 사용자를 위한 포탈로, 클라이언트 소프트웨어를 필요로 하지 않는다. 사용자의 컴퓨터에서 단지 웹 브라우저를 통해 하나의 작업(job) 혹은 복합적인 작업을 수행할 수 있도록 지원한다.

작업은 JOL(작업 기술 언어)을 통해 자동화되고, 작업 수행을 위한 툴은 애플릿과 TightVNC를 이용하여 웹에 임베디드하여 제공하고 있다.

2.3. Grid Resource Broker Portal

Grid Resource Broker(GRB)[4] 포탈은 하나의 웹 게이트웨이로써, 그리드 서비스와 그리드 자원에 접근하기 위한 환경을 제공한다. 사용자 프로파일, 정보 서

비스, 파일전송, 작업 수행 및 추적 서비스를 제공한다. 입력 파일, 출력 파일의 연관성까지 모두 완전히 자동화된 작업을 만들어 실행시키고 파일 전송을 위한 여러 가지 방법을 지원한다.

3. 시스템 설계

현재 포탈에서는 자동화된 작업 수행 도중에 사용자와 대화가 필요하지 않게 되어있다. 이는 사용자의 실수로 인해 작업이 잘못 기술되었을 때, 이를 고칠 방법이 없다. 더욱이 실제 사용자는 완전히 자동화된 작업보다는 일부 작업들을 직접 실행하고 그 중간 결과를 추적하여 확인하는 방식으로 수행하기를 원한다. 본 논문에서는 사용자와 대화식 작업 수행을 지원할 수 있는 포탈을 설계하려 하는데, 이 때 다음 사항이 고려되어야 한다.

1. 작업의 효율성 - 대화식 작업에 대한 응답 속도
2. 일관적인 개념의 사용자 인터페이스

우선 포탈은 웹으로 정보를 보여주기 위해 최적화 되어 있어, 사용자와 대화식 작업 수행에 있어 그 응답 속도가 늦다. 사용자는 대화식 작업에 있어 빠른 작업을 요구한다. 빠른 응답이 이루어지지 않으면 사용자의 작업 효율을 떨어뜨리게 된다.

다음으로, 간단한 대화식 작업은 포탈에서 이루어질 수 있지만, 매우 복잡하게 구성된 대화식 작업은 포탈-웹에서 하기 어렵다. 이러한 작업은 그것을 위한 응용 프로그램을 필요로 한다. 이는 때에 따라 사용자가 둘 다 사용할 수도 있는데, 사용자에게 일관적인 개념의 인터페이스를 제공하여 두 가지 도구를 사용하기 위한 가지만 배워 가능하게 해야 한다.

3.1. 분산 포탈

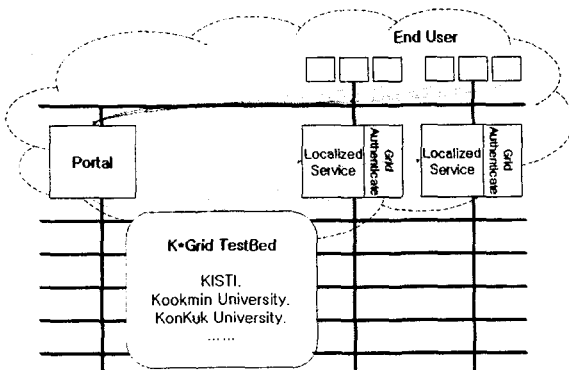


그림 1. 분산 포탈

포탈 서비스의 전체적인 성능 - 사용자에 대한 응답 속도 - 을 높이기 위해 다음 그림 2와 같이 설계한다.

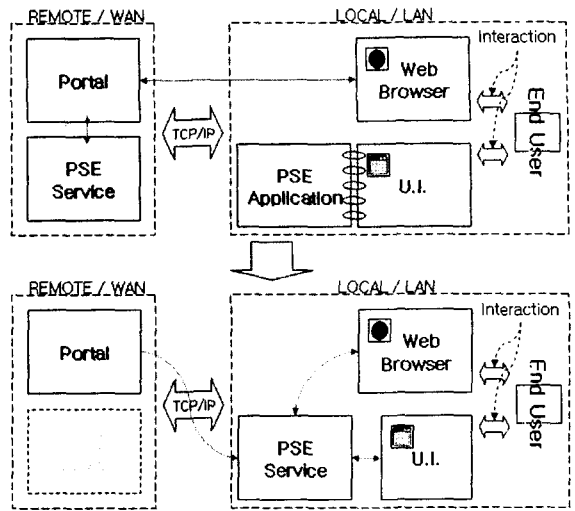


그림 2. 성능 향상을 위한 설계

포탈에서 대화식 작업을 수행하게 되면, 사용자는 WAN을 통해 오는 응답을 기다려야 한다. 이것을 LAN 환경으로 옮겨 그 응답 속도를 빠르게 하고, 서비스의 핵심 부분을 PSE의 UI부분과 웹 브라우저가 공유하도록 하여, 사용자의 선택에 의해 사용할 수 있게 한다. 이를 분산 포탈로 정의하고자 한다.

3.2. 사용자 인증

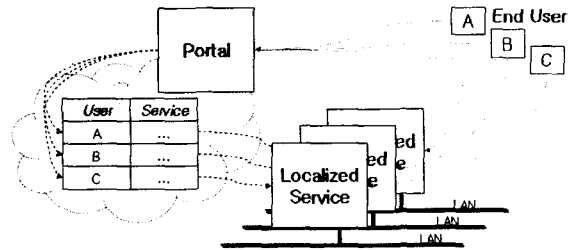


그림 3. 사용자 인증

포탈에서 일부 서비스를 LAN환경으로 이전하게 됨으로써 사용자가 어떤 LAN환경에서 작업하는 것에 대한 정보가 필요하다.

이 정보는 다음과 같은 절차를 필요로 한다. 사용자는 서비스의 성능을 높이기 위해 일부 서비스를 LAN환경에 설치하는데, 포탈은 이 서비스가 어디에 위치하였는지 IP 주소나 도메인 주소를 알고 있어야 한다.

서비스를 LAN환경에 설치한 기관 혹은 사람은 포탈에 이 정보를 등록해야 하는데 이를 기관 등록이라 한다. 기관 등록 후 사용자가 등록할 때에, 자신이 어느 기관의 소속인지 정보를 포탈에 제공하여 향상된 포탈의 서비스를 받을 수 있다.

3.3. 그리드 서비스

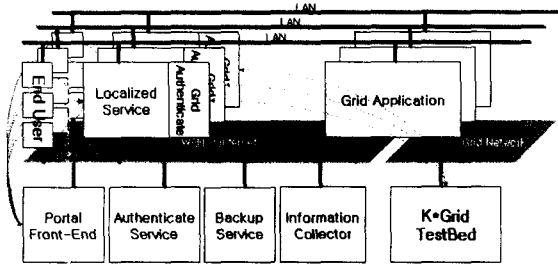


그림 4 포탈 시스템의 구조

기존의 포탈에서 하는 일과 마찬가지로 정보를 보여 주기 위한 역할이 필요하다. 이를 위한 서비스는 포탈에서 이루어진다. 그리고 인증을 따로 서비스로 운영하여 응용프로그램과 포탈을 하나의 인증으로 사용할 수 있도록 해야 한다.

사용자가 지역화된 서비스를 이용할 때, 서비스가 그리드 응용 프로그램 서비스가 일부 혹은 전부를 제공해야 한다. 이 때, 문제가 되는 것은 그리드 인증이다. 그리드 서비스는 사용자 혹은 기관 환경에서 이루어지는데, 이를 위해 각 기관별로 최소한 하나의 그리드 인증을 소유하여 공유하도록 해야 한다. 인증을 위해 필요한 여러 작업 중, 방화벽 문제가 가장 크다. 각 그리드 자원에 방화벽을 완화할 요구해야 하는데, 자원 관리자는 모든 사용자를 위해 완화할 수 없다. 따라서 기관별로 요청하여 그 것을 사용자들이 공유해야 한다.

3.4. 백업

지역화된 서비스는 사용자가 관리하는 환경에서 서비스가 된다. 사용자는 이러한 서비스를 관리 소홀이나, 실수로 인하여 서비스가 중단되도록 만들 우려가 있다. 이 지역 서비스는 실제 사용자가 작업을 수행한 정보를 보관하고 있어, 다음 두 가지를 고려해야 한다.

1. 저장장치의 장애
2. 지역 서비스의 불안정성

첫째로, 저장 장치에 장애가 발생하였을 때, 사용자의 작업(Job)중 태스크 단위와 같은 일부 정보가 그리드 자원에 있는 시뮬레이션 서버 - 자체적으로 체크포인트를 통해 안전하게 저장 -에 저장되어 있어 일부는 복원이 가능하다. 하지만 이 정보는 수행하기 위한 하위 정보로써 상위 정보의 작업을 복원할 수 없는데 이를 복원하기 위해 다음 두 가지방법을 제공하여 사용자가 선택하여 사용하도록 할 것이다.

- ▶ 다른 저장매체로 백업지원
- ▶ 안전한 백업 서버로 주기적인 백업지원

둘째, 지역서비스가 중단되었을 때, 사용자는 서비스를 사용할 수 없는 것 보다는 성능이 떨어지더라도 지속되는 것을 원한다. 포탈은 지역화된 서비스를 감시하

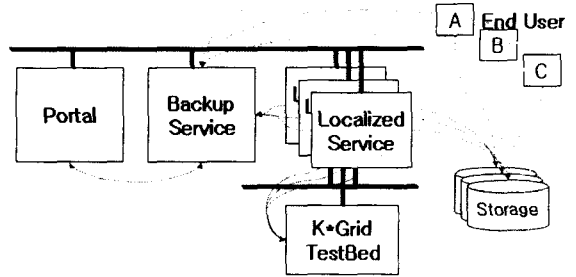


그림 5. 백업

여 중단되면 그 기관의 사용자를 모두 포탈에서 직접 서비스를 받게 하여야 한다. 백업 서비스를 통해 사용자가 지역 서비스에서 사용했던 정보를 포탈에 복원하여, 사용자 입장에서는 성능이 떨어진다는 것 외에는 똑같이 사용할 수 있게 해야 한다.

4. 결론

본 논문에서는 그리드 서비스를 웹 환경, 포탈에서 이용할 때, WAN환경을 거쳐 응답이 이루어지는 것을 LAN환경에서 응답이 이루어지도록 하여 전체적인 성능을 향상시키고자 포탈의 일부 기능을 사용자의 LAN환경에서 서비스하는 분산 포탈을 제안하였다. 본 논문에서 제안한 분산 포탈은 기존의 포탈과 비교해서 사용자의 입장에서는 달리 습득해야 할 변화가 없다.

그리고 분산 포탈이 되었을 때 필요한 사용자 인증 방법과, 사용자 환경에서 서비스되는 불안정성을 해소하는 방법을 제안하였다.

참고문헌

- [1] Michael Brim, Al Geist, Brian Luethke, Jens Schwidder, Stephen L. Scott, "M3C: Managing and Monitoring Multiple Cluster Computing and the GRID, p386, 2001
- [2] Gregor von Laszewski, Ian Foster, Jarek Gawor, Peter Lane, Nell Rehn, Mike Russell "Designing Grid-based Problem Solving Environments and Portals" Proceedings of the 34th Hawaii International Conference on System Sciences, 2001.
- [3] R. Barbra, "The GENIUS Grid Portal", Computing in High Energy and Nuclear Physics, 24-28 March 2003, La Jolla, California
- [4] Giovanni Aloisio, Massimo Cafaro, "Web-based access to the grid using the Grid Resource Broker Portal"