실험망을 위한 파일 전송 기술 동향 분석

이한나, 이동우, 엄영익
성균관대학교 임베디드 소프트웨어학과
e-mail:{hannah1210, lightof, yieom}@skku.edu

Analysis of Technical Trend for Scientific Networks

Han-Na Lee, Dongwoo Lee, Young Ik Eom
College of Information and Communication Engineering, Sungkyunkwan University

요 약

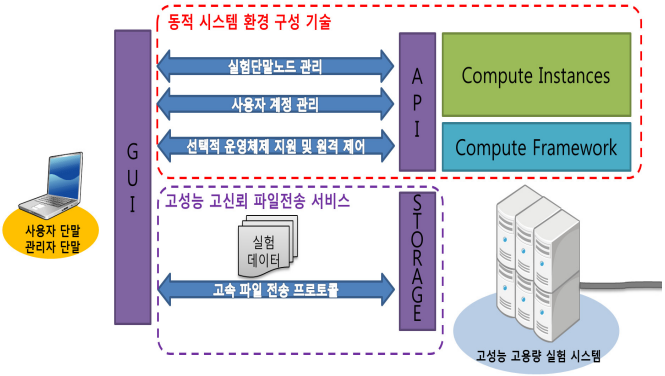
과학 기술 연구 개발을 위한 국제 협업이 활발하게 일어나면서 국내에도 글로벌 과학기술협업연구망인 GLORIAD를 구축하게 되었다. 첨단 과학 기술 연구에 쓰이는 데이터가 교환되는 만큼 신속하며 정확한 데이터 교환이 요구된다. 따라서 신속하며 높은 신뢰도를 가지고 대용량의 데이터를 전송할 수 있는 파일 전송 기술이 필요하다. 본 논문에서는 GLORIAD 실험망 구축에 적용 가능한 파일 전송 기술의 동향을 조사하고, 분석하는 과정을 통해서 최적의 파일 전송 기술을 찾아낸다.

1. 서론

인터넷의 급속한 기술 발전과 함께 고성능 네트워크를 기반 시설로 하여 대용량의 데이터 분석이 요구되는 과학 분야가 크게 늘고 있다. 고에너지 물리학, 기상학 등의 첨단 과학 기술 연구에서는 실시간 대용량 데이터 전송 처리를 통한 사이버 인프라가 요구된다. 분산된 지역 간의 원활한 원격 협업 연구를 위해서도 연구망의 구축은 국가 과학 기술 발전에 중요한 요소가 되었다. 국제협력을 통한 과학기술 연구개발을 위한 최초의 글로벌 과학기술협업연구망인 GLORIAD를 국내에도 구축하게 되었다. (그림 1)은 국내 GLORIAD 실험망에서 사용될 컴퓨팅 시스템이다. 이 시스템은 고성능/고신뢰 파일 전송 서비스를 통해 사용자와 스토리지에 저장되어 있는 실험 데이터를 교환하게 된다. 실험 데이터 교환에 있어서 중요하게 다루어야 할 부분은 데이터의 정확성과 전송의 신속성이다. 실험 데이터가 손상된다면 실험 결과에 영향을 미칠 수 있기 때문이다. 따라서 신뢰성을 보장하고, 신속하며, 대용량의 데이터를 처리할 수 있는 파일 전송 서비스가 요구된다.

본 논문에서는 대용량 파일 전송 서비스들에 대해 조사하고, 분석 과정을 통해 최적의 파일 전송 기술이 무엇인지 알아보도록 한다. 본 논문은 다음과 같이 구성되어 있다. 2절에서는 파일 전송 서비스의 동향을 알아보고, 3절에서는 이 파일 전송 서비스들을 분석한다. 그리고 마지막으로 4절에서는 본 논문에 대한 결론을 내며 마친다.

†"이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단-차세대정보컴퓨팅기술개발사업의 지원을 받아 수행된 연구임(No. 2012-0006423)."



(그림 1) 실험망 컴퓨팅 시스템

2. 기술 동향

본 절에서는 파일 전송 기술인 Rapidant, Aspera[1], Globus Toolkit[3]에 대해 알아본다.

2.1 Rapidant

Rapidant는 국내 회사에서 만든 대용량 데이터 고속 전송 솔루션으로, 특정 운영체제(OS)에 상관없이 간단하게 소프트웨어 설치만으로 대용량 데이터를 고속 전송이 가능하게 설계 되었다. 대용량 파일 고속 전송을 위해 병렬 TCP와 RUDP(Reliable UDP) 방식을 활용하였다. UDP에 데이터 신뢰 전송 및 혼잡 제어가 보완된 RUDP와 다수의 TCP 채널을 사용하여 전송률을 높인 병렬 TCP를 사용함으로써 TCP 전송의 문제점들을 해결하였다. 또한, 네트워크 공정성 알고리즘을 적용하여 네트워크의 부하도 최소화되었다. 이로 인해 기존 파일 전송 대비 10배 이상

빠른 처리 속도를 보여주고 있다. 전송 속도 향상으로 글로벌 통합 서비스 체계를 구축하고 전송 안정화로 향상된 고객 서비스를 제공하고 있다.

2.2 Aspera

Aspera는 초고속 대용량 파일 전송 솔루션이다. Aspera는 대용량의 데이터를 거리와 네트워크 상태에 관계없이 빠른 속도로 전송하는 서비스를 제공하는 차세대 초고속 파일전송 소프트웨어다. 기존 파일 전송 속도에 비해 최대 1000배의 속도를 제공하며, 기존 네트워크 인프라를 변경하지 않아도 된다. 신뢰 전송에 쓰이는 FTP나 HTTP 같은 프로토콜에서 일어나던 TCP의 병목 현상을 해결하여 네트워크 상태에 상관없이 빠른 속도의 전송을 제공한다.

2.3 Globus Toolkit

Globus Toolkit는 그리드 컴퓨팅과 웹 서비스를 궁극적으로 하나로 통합하자는 취지에서 시작됐다. Globus Toolkit는 그리드와 그리드 응용 툴을 지원하는 서비스들과 소프트웨어 라이브러리들의 집합이다. 보안, 정보 자원 관리, 데이터 관리, 통신, 장애 감지, 그리고 이식성을 제공하는 소프트웨어를 포함하고 있다. 유용한 그리드 응용 프로그램들과 프로그래밍 툴들을 개발하기 위해 함께 또는 독립적으로 사용될 수 있는 콤포넌트들의 집합으로 패키지화 되어있다. Globus Toolkit는 GridFTP를 이용하여 대용량의 파일을 전송한다. GridFTP[2]는 고성능의 안전하고 견고한 프로토콜이다. 병렬 데이터 전송을 사용하여 전체 대역폭을 향상시켜 전송률을 높였다. 큰 파일과 많은 양의 작은 파일들에 대해 수동 및 자동 TCP 버퍼 크기 협상을 지원한다. 따라서 TCP 버퍼/윈도우 크기를 최적으로 설정한다면 데이터 전송 성능을 크게 향상시킬 수 있다. 그 밖에도 GridFTP는 데이터 전송의 third-party를 제어하고, 부분 파일 전송, 스트라이핑 데이터 전송을 한다는 특징을 가지고 있다.

3. 분석

<표 1>에는 2절에서 조사한 파일 전송 서비스들의 솔루션 별로 전송 방법, 혼잡 제어 방법 등을 비교하였다.

파일이나 데이터 전송에서 주로 사용되는 TCP는 신뢰성 있는 전송과 혼잡 제어를 제공한다. 하지만 TCP 프로토콜은 RTT(Round Trip Time)와 패킷 손실이 큰 원격 지간의 통신에 있어서 상대적으로 좋지 않은 성능을 보인다. 따라서 대용량의 데이터를 원거리로 전송해야하는 경우에 TCP는 적합하지 않기 때문에 비교 분석된 3개의 파일 전송 기술 모두 TCP의 문제점을 보완해줄 수 있는 다른 전송 프로토콜[4]을 사용하고 있다.

Rapidant의 경우는 여러 개의 TCP 채널을 이용하여 전송률을 높인 병렬 TCP와 데이터 전송의 신뢰성을 제공하는 RUDP를 함께 사용한다. 하지만 혼잡이 일어날 때 패킷이 손실되며 병렬 TCP로 혼잡 제어한다. Aspera의 경우는 UDP를 통해 데이터를 전송하며 혼잡 상황에서 큐잉 지연이 일어난다. Globus Toolkit의 경우에는 병렬 TCP를 사용하는 GridFTP로 데이터 전송이 이루어지며, 병렬 TCP로 혼잡 제어 한다. 세 가지 방법 모두 네트워크 공정성을 고려하도록 되어있다.

<표 1> 솔루션 별 전송 프로토콜 및 혼잡 제어 비교

솔루션 명	전송 프로토콜	혼잡 제어	비고
Rapidant	RUDP, 병렬 TCP	패킷 손실, 병렬 TCP	네트워크 공정성 고려, 플랫폼 독립적 설계
Aspera	UDP	큐잉 지연	네트워크 공정성 고려
Globus Toolkit	GridFTP (병렬 TCP)	병렬 TCP	네트워크 공정성 알고리즘

4. 결론

국내에 구축되는 GLORIAD 실험망에서 대용량의 데이터를 실시간으로 정확하게 처리하기 위한 파일 전송 기술 동향에 대해 조사하였다. 실험망은 정확한 실험 결과를 만들어야하기 때문에 손실이 일어나지 않는 신뢰성을 보장받는 데이터 전송을 전제로 한다. 본 논문에서는 Rapidant, Aspera, Globus Toolkit 3개의 기술을 조사하였다. 3개의 기술 중에 Globus Toolkit가 사용하는 GridFTP는 가장 안전하고 견고한 전송을 제공한다. 따라서 신뢰성 있고 안정적인 실험망 구축을 위해서는 Globus Toolkit이 가장 적합하다. 하지만 GridFTP는 병렬 TCP를 사용하기 때문에 병렬 TCP가 갖는 전송 성능상의 단점을 가지게 된다. 프로토콜 중에 UDP를 기반으로 신뢰성을 보장하는 대용량 데이터 전송 프로토콜인 UDT를 쓰면 이런 단점을 보완할 수 있을 것이다.

참고문헌

[1] Aspera, http://asperasoft.com/company/why-aspera/

[2] 김은성, 박형우, 이상산, “고성능 데이터 전송을 위한 GridFTP 성능 향상에 관한 연구“ 한국정보과학회, 2003.

[3] William Allcock, John Bresnahan1, Rajkumar Kettimuthu, Michael Link, Catalin Dumitrescu, Ioan Raicu, Ian Foster “The Globus Striped GridFTP Framework and Server” SC'05, ACM Press, 2005.

[4] 박종선, 김승해, 황건준, 조기환 ”UDT 기반의 대용량 데이터 전송도구 설계 및 구현” 전자공학회지, 2012.