

그리드 고성능 네트워크 구축과 해외 동향

글_ 노민기 | 초고속연구망사업실 연구원 | mknoh@kisti.re.kr
김승혜 | 초고속연구망사업실 연구원 | shkim@kisti.re.kr

1. 서 론

그리드 환경의 자원은 네트워크에 분산되어 연결되어 있으며 분산환경에서의 연구는 고성능 네트워크를 통하여 효율적으로 사용 할수 있도록 구성된다. 여기에서 네트워크 환경 즉 그리드 네트워크는 그리드 응용연구의 연산과 데이터전송의 효율과 성능유지를 위한 기본적인 통신환경이며 기존의 일반적인 연결망주의 네트워크(Best-Effort Network)보다 더 고속의 처리(High Throughput)와 안정성이 요구되는 환경을 제공하여야 한다.

즉 그리드네트워크의 구성은 광대역폭의 네트워크 인프라와 네트워크기술의 제공(Grid optical network Technology)을 통하여 네트워크 기반을 구성하고 그리드 연구 트래픽에 대한 효율적인 전달(Grid traffic engineering & TE) 및 그리드 어플리케이션에서 요구되어지는 프로토콜과 그 프로토콜에 대한 운영(High-performance transport Protocols and Service) 기능과 네트워크의 자원관리와 할당(Provide management like Access control, Scheduling)의 네트워크 전달기능을 수행하며 이러한 자층(Layer)위에 그리드 응용 어플리케이션과 미들웨어의 운영의 효과적인 지원(Driving Grid Application & Middleware)을 제공할수 있는 네트워크 기반과 기술을 제공하여 그리드응용 어플리케이션에 중요한 요소인 End-to-end 사용자간의 네트워크의 성능향상과 그리드응용 트래픽의 특성에 맞는 효율적인 네트워크를 제공하는 것을 목적으로 한다.

2. 그리드 응용 특성별 네트워크 환경

1) 컴퓨팅 그리드(Computational Grid) 네트워크 환경

컴퓨팅 그리드에 속하는 시스템들은 많은 자원들을 연결하여 대용량 연산을 해결할 수 있도록 하는 장치들을 의미한다. 이들 중에서도 분산 슈퍼 컴퓨팅 관련 그리드는 작업의 전체 수행 시간을 줄이기 위하여 많은 컴퓨터들을 동시에 이용하는 것을 말한다.

컴퓨팅 그리드의 네트워크의 특성은 시스템의 연구활용에 있어 서비스의 응답시간과 가용성(availability)과 함께 데이터의 일관성유지를 위해 네트워크의 지연시간이 최소화되어야 한다. 또한 네트워크의 장애에 대해 반복적인 조직과 데이터의 안정성이 유지되어야 한다. 또한 주어전 시간 안에 가능한 많은 작업을 처리 하기 위해 대용량의 파일송수신을 제공하여야 한다.

2) 데이터 그리드(Data Grid)

대용량 데이터의 처리분야인 데이터그리드는 대규모 시뮬레이션, 관측장비의 저장 및 로그데이터, 수치데이터 등의 자료를 발생하며 이를 네트워크를 활용하여 협업연구의 지원, 원거리 데이터의 저장 등에 활용한다. 따라서 분산된 여러 곳에서 생성되고 한 곳에서 처리하는 형태 혹은 분산된 대용량의 데이터를 여러 곳에서 처리하는 형태를 지원하기 위해서는 고용량의 대역폭과 서비스 또한 데이터를 효율적으로 전달하는 프로토콜의 연구 및 서비스가 네트워크 인프라에서 제공되어야 한다.

3) 액세스 그리드(Access Grid)

엑세스 그리드는 본질상 분산처리를 필요로 하는 응용을 위한 그리드이다. 엑세스 그리드 역시 사용분야에 따라 나눌 수 있다. On Demand 그리드는 사용자 작업이 사용자 요구에 따라 자원들을 동적으로 연결하는 것이 가능한 그리드이다. 사용자는 작업 처리 중에 결과를 좀더 정밀하게 보기 위해 더 많은 자원을 요구할 수 있다. 협업환경 그리드는 사용자들이 원격지에서 협력하여 일을 처리할 수 있게 하는 작업 공간을 제공한다. 멀티미디어 그리드는 실시간 멀티미디어 응용을 위한 인프라를 제공한다. 멀티미디어 그리드는 다양한 스트림이 동시에 교환되기 때문에 네트워크의 성능이 반드시 보장되어야 하며 이를 위해서는 지연시간(Delay time)의 최소화, 지연시간 변동(Jitter)의 최소화, 엑세스 그리드의 표준프로토콜인 멀티캐스트(Multicast)망의 구성 및 서비스가 제공되어야 한다.

3

해와 그리드 네트워크 현황

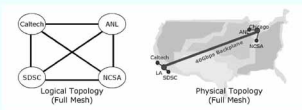
해외에서 활발히 진행되고 있는 그리드 응용 사례 연구들에서 사용되고 있는 초고속 네트워크 현황을 통해 네트워크를 좀더 효과적으로 이용하고 각 응용 결과들의 최대 효과를 보기 위해 어떠한 네트워크 통해 어떠한 방식으로 운용하고 있는가를 살펴본다.

1) TeraGrid

2001년 미국과학재단(NSF)은 테라바이트급 슈퍼 그리드 컴퓨팅 구축을 위해 슈퍼컴퓨팅 시스템인 DTF(Distributed Terascale Facility) 구축 계획을 추진하였으며 이는 DTF의 컴퓨팅 그리드로부터 자원을 임의로 사용할 수 있도록 하는 것을 목적으로 하고 있다. TeraGrid에는 일리노이즈 소재의 NCSA(National Center for Supercomputing Applications)와 샌디에고의 SESO(San Diego Supercomputer Center), 시카고의 ANL(Argonne National Laboratory), 파사디나의 Caltech(California Institute of Technology), FSO(Pittsburgh Supercomputing Center)가 참여하고 있다.

각 참여 기관에는 시각화, 시뮬레이션, 데이터 관리, 공동 연구 지원 등을 포함하여 서로 다른 4개의 지역에 초당 1.2TF(Teraflops)를 수행할 수 있는 시스템이 나누어져 있다. TeraGrid 사이트는 용량과 기능면에서 광대역성을 제공하고 있다. ANL 사이트에서는 상대적으로 작은 용량(1.25TF)의 클러스터로 시각화 서비스를 제공하며, Caltech는 0.4TF 클러스터를 이용하여 데이터 분석 능력을 제공 할 것이다. NCSA와 SESO에서의 대규모 클러스터 시스템처럼 ANL과 Caltech 클러스터들은 각각의 모든 노드들은 1Gb/s(Gigabit Ethernet) 인터페이스를 가지며 Gigabit 스트림의 여러 채널을 이용하여 다른 TeraGrid 노드들과 데이터를 주고받을 수 있다. 유사하게 FSO의 LAN은 백본에 연결하기 위해 여러 개의 Gigabit Ethernet 채널을 사용한다.

TeraGrid 백본은 코스텔벨레스와 시카고 사이의 광통신 기반구조와 ANL과 NCSA에서 시카고로 연결짓는 광통신 기반구조, 그리고 Caltech와 SDSC에서 로스앤젤레스로 연결짓는 광통신 기반구조로 구성되었다. 불필요한 전자장치와 계층을 피함으로써 지연을 최소로 구성하고 여러 개의 10Gb/s 광채널을 채용하여 대역폭을 최대로 하는 Full Mesh 내부연결로 구성되었다. 향후 시카고 대학과 Northwestern 대학, UI-Chicago의 협력기관에게까지 네트워크를 확장할 계획이 있다. 클러스터의 각각의 노드들은 Gb/s 의 LAN 인터페이스를 가지며 Gigabit Ethernet Lan 스위치에 접속되어 있다. 여기에서 보여지고 있는 각각의 Lan 스위치는 클러스터로부터 16Gb Ethernet 연결이 들어오고 하나의 10Gb Ethernet을 통해 중앙의 클러스터 집중 스위치로 빠져나가며 3개의 채널을 통해(Border 라우터) 백본으로 연결된다. 두 계로 분리되어있는 이유는 지역 클러스터 환경변화와 경전, 실험 등으로부터 백본의 기능에 영향을 미치지 않게 하기 위해서이며 다른 하나는 LAN과 WAN에서의 네트워크 트래픽을 스위칭, 라우팅을 위한 장비가 매우 다르기 때문이다. Border 라우터와 허브 라우터를 위해 사용되어지는 것 같은 멀티프라이즈 IP 라우터들은 필요량의 버퍼링과 긴 지연, 고 대역폭의 WAN을 위한 요구와 관련되어 있다. 반면에 LAN 스위치들은 짧은 지연을 위해 최적화 되었다.



TeraGrid 네트워크 구성

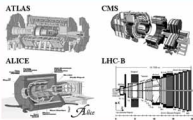
2) LHC(Large Hadron Collider)

스위스의 CERN에 건설 중인 거대 강입자 충돌형 가속기(LHC)를 이용하는 입자물리학 실험으로 물질의 궁극적인 구조와 법칙을 규명하기 위한 차세대 실험물리학 과제로 2005년부터 실험을 수행할 예정이다. 이 실험체계에는 세계의 연구소, 대학에 걸쳐 약 6,000여명의 연구자가 참여할 것으로 예상되고 있으며 필요한 Computing 요구량 해도 연간 데이터 생성량이 5~8Peta바이트에 이르고 분석을 위한 디스크 요구량은 약 10

Peta바이트, 요구 계산량은 최근의 가장 빠른 PC의 프로세서 200,000의 컴퓨팅 능력을 요구하고 있다. LHC에는 4개의 주요 검출기가 존재하며 각각의 검출 프로세스를 위하여 컴퓨팅 자원을 필요로 하고 있다.

참여기관으로는 CERN, European Union, Intel, Entersys, KPNQwest와 10여개국 이상이 참여하고 있다.

모든 데이터들은 처음에 CERN에서 생성되어지며 연구에 참여하는 실험장비의 필요 대역폭은 2006년에 CERN-Tier1 간에 필요로 하는 수치이다. 이러한 필요 대역폭의 산출은 Tier1-Tier1, Tier2-Tier1, Tier2-Tier0의 트래픽에 대해서도 화장을 하여야 할 것이다. 대역폭은 TierN에서부터 TierN+1까지의 Bulk 데이터



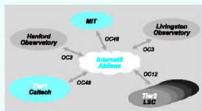
LHC 프로젝트의 4개의 검출기

전문분야가 아니라 연구원 각자의 개인분석, 데이터가 생성되고 분석이 수행되는 장소에서 CERN으로의 트랙, 재가공된 데이터의 배포, 협력 연구를 위해서도 필요 되어진다. LHC 프로젝트에 제공되는 네트워크는 네트워크 트래픽의 포화를 방지하기 위해서는 채용된 네트워크 대역폭이 요구되는 네트워크 처리량의 최소 2배를 포함하여 네트워크는 50%이상의 점유가 되지 않아야 하므로 하나의 실험 장비의 필요 대역폭은 대개 1.5~3Gbps에 추정하고 있다. 연동네트워크의 대역폭은 파이버 페어(Fibre pair)당 800Gbps의 용량을 가지는 테라케이블과 현재의 DWDM 기술을 이용하는 4개의 파이버 페어의 설치가 2004~5년까지 계획되어 있으며 이 파이버를 경유하여 End to End 기기비트 이더넷이 지원되며, 이 기반을 중심으로 네트워크 엔지니어링(Traffic Engineering), 프로토콜 스택 최적화, 네트워크를 통한 사용자 모니터링(Traffic monitoring, Measurement), 그리드 미들웨어의 사용 등 고성능 대역폭의 효율적인 사용을 위해 필요 되어지고 있다.

3) LIGO(Laser Interferometer Gravitational wave Observatory)

미국 NSF가 지원하고 MIT와 CIT가 공동으로 추진하는 중력파의 실제 검출을 위한 대규모 순수 물리학 분야의 양자광학의 기초이론을 검증하는 대규모 실험 프로젝트이다. 중력파 관측 장치를 이용하여 쌍을 이루는 중성자 별들은 포함하여 다양한 근원지로부터 발생한 파동 검출하고, 지구상의 두 곳의 위치에 4km의 측정거리를 갖는 초대형 레이저 간섭계를 구축하여 10e 11m의 절대정밀도측정의 구현 목표를 한다. 또한 일반 상대성 이론에 대한 실험적 검증을 위해 이용된다. LIGO 실험에는 초대형 레이저 간섭계는 Hanford 관측소와 Livingston 관측소 두 곳이 있다. 이러한 관측소에서 실험, 측정된 데이터를 가지고 실험을 수행하고 있는 기관은 UWM, UTB, CIT 이다.

저의 모든 장비들은 대부분 수십 GB의 SCSI 디스크를 가지고 있으며 1TB 이상의 저장장치를 가지고 LAN에 1,000Mbps로 연결되어 있으며 위의 LIGO 실험을 수행하기 위해 End to End 네트워크 대역폭 요구량은 최소 9Mbps가 요구된다. 초대형 레이저 간섭계간은 미국의 internet2의 백본연구망인 Abilene을 사이에 두고 각 155Mbps로 연결되어 있으며, UWM은 MREN을 백본망으로 하여 155Mbps로 연결되어 있다. CIT는 Caltech과 ATM 백본망을 두고 2.5Gbps로 연결되어 있다.



UGO 네트워크 구성

4) GriPhyN(Grid Physics Network)

GriPhyN 프로젝트는 IT 연구자들과 ATLAS, CMS, LIGO, SLAS의 물리학자들과의 컴퓨터 과학의 협업 연구이다. 이 프로젝트는 전세계에 흩어져 있는 다양한 공동 연구 조직의 수천명의 과학자들이 data intensive 계산 요구에 부합하기 위해 Petascale Virtual Data Grids를 생성하는 것에 초점이 맞춰져 있다. GriPhyN 프로젝트는 Virtual Data Toolkit(VDT)를 이용하여 Petascale Virtual Data Grids(PVDG)를 구성하기 위해 요구되는 기본적인 IT 발전을 수행하는 것에 초점이 맞추어져 있다. GriPhyN 프로젝트의 네트워크 성능은 최소대역폭을 약 5.3Gb/s의 고용량과 매우 낮은 지연시간의 제공을 기본으로 하고 있다. 또한 DiffServ와 같은 QoS 메커니즘과 구성의 효과 측면에서 이용 가능한 모든 대역폭을 획득하기 위해 bulk flow를 허용하는 QoS를 중심으로 제공하고 있으며 QoS가 설정된 라우터 및 스위치에 대한 실시간 모니터링 기능을 수행하고 있다.

5) Hyunhae/Genkai Project

Hyunhae/Genkai Project는 국내의 연구 기관들과 일본의 대학 및 연구기관이 KJCN네트워크를 구성하여 만든 AG(Access Grid) 프로젝트이며 가상 박물관, 디지털 TV/Video 전송기술, 광범위한 컴퓨팅 시스템, 원격 교육 및 협업 지원 시스템(Access Grid System), 고에너지 물리 실험 데이터 전송을 주요연구내용으로 하고 있다. 현재 Hyunhae/Genkai Project는 일본의 ISIT와 Kyushu 대학을 중심으로 진행되고 있으며 우리나라와는 한국과학기술원(KIST), 부산대학교, 전남대학교, 울산대학교, 경북대학교 등이 참여하고 있다. Hyunhae/Genkai Project를 위해 구성된 KJCN(Korea-Japan Cable Network)는 12개의 fiber pair로 구성되어 있으며, 50Gbps에서 시작하여 향후 2.88Terabit까지 확장할 계획이다. 현재 우리나라와 일본간의 데이터 트래픽 용량은 500Mbps 정도로 나타나고 있다.

4 KISTI 그리드 기가 네트워크 구축 현황

국내·외 첨단 응용 연구 기반 제공과 국가 그리드 기반 구축 (K*GRID) 프로젝트의 성공적 수행의 네트워크 환경을 지원하기 위하여 그리드 NOC(Network Operation Center)에서는 초고속연구망(KREONET)의 네트워크 자원을 중심으로 그리드 응용 연구 기관 및 그리드 자원 제공 기관 간에 광 대역폭의 기가급 네트워크를 제공함으로써 국내·외 연구망의 네트워크 자원을 활용하고 그리드응용 및 슈퍼 컴퓨팅, 클러스터 자원 공유를 통한 그리드 기반에서의 응용 연구 등의 연구 과제를 수행 할 수 있는 고성능 네트워크 인프라를 제공을 수행하고 있다.

1) 네트워크 연동 및 지원 형태

각 연구 기관 및 연구 장비에 대해 네트워크 자원의 직접 제공과 연구에 필요한 일정한 네트워크 자원할당을 기본구상으로 대상 기관에 제공되는 네트워크는 초고속연구망(KREONET) 12개 지역 센터에 연동을 기본으로 하여 물리적인 연결을 하여 지역망 센터중심의 자원할당을 통해 네트워크 장비를 중심으로 라우팅(Routing) 및 스위칭(Switching)을 기반으로 네트워크에 논리적으로 연동한다.



Diagram of international research network links

또한 컴퓨팅 자원공유, 액세스 그리드를 통한 협업연구, 첨단장비 공동 활용과 같은 그리드 네트워크 환경 구축 및 자원을 활용한 그리드 응용연구의 실현 및 구현에 네트워크를 제공하는 방식은 DWDM(Dense Wavelength Division Multiplexing) 네트워크의 신규 연동 및 MSP(Multi Service Provisioning Platform)를 활용하여 지원기관에 대한 람다(Lambda) 할당 및 스위치 포트에 의한 대역폭 할당을 통하여 그리드 응용 연구에 기가급 네트워크를 지원한다.

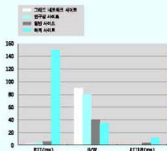
2) 그리드 기가 네트워크 구축 내용

- 고성능 클러스터 컴퓨팅 그리드 자원(WDM 신규 지원 및 자원할당)
: KISTI, 부산대, 서울시립대, 서울대, 포항공대

- 슈퍼컴퓨터 컴퓨팅 그리드 자원 : 6개 슈퍼컴퓨터
: KISTI, 포항공대, 전북대, 동명정보대, 명지대, 금호그룹
- Access Grid(협업그리드) 구축 고도화(국내 최고의 협업 연구 인프라 구축 및 성능 향상)
: 광주과학기술원, KISTI
- 첨단 과학 기술 실험 장비 그리드 구축 기관
: 한국기초과학연구원(초고전압투과 전자현미경: UHV TEM)
- 국가적 고 에너지 물리 실험 데이터의 전송 및 결과 분석연구에 활용될 데이터 그리드 테스트베드 구축
: 경북대학교

3) 그리드 네트워크 연동 효과

사용자간의 End to End 기기네트워크의 재공에 따라 기존의 연동보다 연동속도의 고속화와 라우팅의 경로가 단순함에 따라 지연시간(Delay)의 감소와 단순화된 라우팅 경로를 제공한다. 또한 각 주요연동기관의 기기네트워크를 연동은 캠퍼스 라우팅(Campus Routing) 즉 기관 전체에 초고속연구망(KREBONET)을 사용할 수 있는 연동하였으므로 그리드유저뿐 아니라 슈퍼컴사용자들도 같은 속도 및 대역폭의 회선을 사용할 수 있도록 구성하였으므로 다양한 연구 및 사용자에 대한 지원 가능하게 되었다. 그외에 국내 연구자들이 국제협력 연구 활동을 보다 편리하고 효율적으로 제공하기 위하여 STAR TAP/미망(155Mbps)을 중심으로 북미 지역의 연구망 접속을 제공하여 초고속선도망(KOREN)과의 1Gbps 고대역폭의 연동을 통한 Renater/유럽, APAN JP/일본, SingAREN /싱가포르 국제망을 활용 가능한 네트워크를 구축하여 제공 한다. 국제 협업을 수행할 수 있도록 국제 연구망과의 직접적인 연동을 초고속연구망의 국제 게이트웨이와 APN Teetbed 국제망간 연동을 통하여 제공하고 있다.



기기 네트워크 성능 측정결과

4 결론

전 세계적으로 그리드 환경에서의 연구와 응용의 개발에 대한 지원과 이에 필요한 고성능 네트워크 기반 및 기술에 대한 연구가 활발히 진행중이 있으며 기술적인 많은 발전을 보이고 있으며 특히 세계 선진국을 중심으로 각 국가의 연구망 자원을 활용하여 네트워크 기반을 제공하고 있다. 이에 발맞추어 국내에서도 초고속연구망(KREBONET)/초고속선도망(KOREN)/APN Teetbed를 통해 국가적인 그리드 네트워크 기반을 제공하고 있으며 고성능 기기 네트워크 환경을 구축 활용하고 있다. 현재 세계적인 수준으로 구축되어 있는 국내 기기네트워크의 활용은 본고에서 살펴본 해외 그리드 연구 프로젝트별로 구축되어 있는 네트워크의 형태와 특성에 따른 구축 및 제공 기술을 참고하여 국내 실정에 맞게 적용함으로써 구축된 기기네트워크를 활용한 효율적인 그리드응용에 대한 지원을 기대할수 있다. 또한 이에 요구되는 네트워킹 기술의 적용과 발전은 성공적인 국가적인 그리드 인프라 구축뿐만 아니라 향후 다양한 고성능 어플리케이션의 출현과 이에 요구되는 차세대 네트워킹기술에 대한 선도적인 역할도 성공적으로 수행 할수 있으리라 본다.