ScienceDMZ 기반 빅데이터 고속도로 체계 및 AI 컴퓨팅 체계 연계 구축 방안 연구

문정훈, 김기현, 석우진
한국과학기술정보연구원
jhmoon@kisti.re.kr, kkh1258@kisti.re.kr, wjseok@gmail.com

A Study on the big data super highway and AI computing environment integrate based on ScienceDMZ

Moon Jeong Hoon, Kim Ki Hyeon, Seok Woo Jin
Korea Institute of Science and Technology Information

요 약

본 논문은 데이터 집약형 과학의 발달로 대용량 데이터의 생산이 급증하고 있으며, 또한 IT 기술의 발달로 인하여 실험, 관측, 계산 장비들의 고도화에 따른 연구결과의 대용량와 복잡화 되어 가는 추세에 있다. 이러한 동향은 4차 산업 시대에 전세계적인 현상이며, 선진국을 중심으로 여러 나라들이 컨소시엄을 구성한 국제 프로젝트들의 등장으로 엑사바이트 규모 이상의 데이터가 생산되고 있다. 이러한 거대 규모의 데이터는 한 두 곳의 데이터 센터에서 처리하기 어려운 수준이며, 이것은 전송, 저장, 공유의 문제를 의미하기도 한다. 따라서 이러한 과학 빅데이터들은 전세계 주요 데이터 센터들에 분산되어 저장되며, 실제 연구자들은 연구에 투자하는 시간보다 이렇게 분산되어 있는 데이터들을 모으고, 분류하는데 더 많은 시간을 사용하고 있다. 따라서 본 논문에서는 데이터 집약형 과학분야의 빅데이터를 위한 빅데이터 고속도로 체계인 ScienceDMZ의 구축과 이러한 빅데이터를 기반으로 AI 연구를 위한 AI 컴퓨팅 체계의 연계 구축 방안을 제안한다.

Ⅰ. 서 론

본 논문에서는 데이터 집약형 과학분야의 비약적인 발전과 IT 기술의 발전으로 인하여 과거에 비해 수백배 이상의 과학 데이터들이 생성되고 있다. 이런 대용량 데이터의 생성은 실험, 관측 장비로부터 저장을 위한 데이터 센터로의 전송과 데이터 센터에서 연구자에게 전송을 효과적으로 해야 하지만 기존 네트워크의 전송 성능의 한계로 인하여 연구자에게 적절한 전송 환경이 제공되지 못하고 있다. 특히 미국 과학재단에서는 향후 과학분야에서 중점적으로 지원해야하는 데이터 집약형 과학 분야를 7개로 분류하고 있는데 입자물리, 천문관측 및 천체물리, 바이오분야의 게놈 데이터, 의료데이터, 지진, 기상 분야를 포함하는 지구과학 데이터, 인공지능분야, 가상현실 분야이다. 이들 분야 중 입자물리는 스위스 CERN에서 발생하는 데이터를 전세계 데이터 센터로의 전송과 분석을 통해 새로운 입자를 발견하고 있으며, 천문분야의 LSST 프로젝트[1]는 남미 칠레의 관측 망원경으로부터 관측된 자료를 전세계 연구자들과 공유하는 프로젝트로서 하루 15TB~30TB가 발생하고 있다. 이러한 거대 데이터는 한 두 곳의 데이터 센터에서 처리할 수 있는 수준이 아니기 때문에 전세계의 빅데이터 센터로 분산 저장되고 있다. 또한 분산 저장된 데이터들을 연구자의 입장에서 하나의 데이터 셋으로 구성하여 원하는 연구를 할려면 고속의 전송과 분석 작업이 효과적으로 연계되어 진행되어야 한다. 이러한 과학분야의 빅데이터를 효과적으로 처리하기 위해서는 전송 성능의 향상과 이와 연계된 AI 컴퓨팅 작업이 함께 이루어지는 것이 바람직한데 기존 TCP/IP 프로토콜에서 TCP는 아직까지 데이터 전송에서 가장 적합한 프로토콜이다[2]. 그러나 망에서 여러 가지 요인들로 인하여 전송 성능의 한계가 있다. 그중에서 특히 불특정한 이유로 발생하는 패킷로스는 그림 1에서와 같이 고대역망에 대해서 최저 성능까지 떨어졌다가 다시 회복되는 과정을 거쳐 기존 성능으로 복구가 되는데 이러한 현상은 망 자체가 리부팅 되는 수준으로 성능저하가 일어난다.[3]

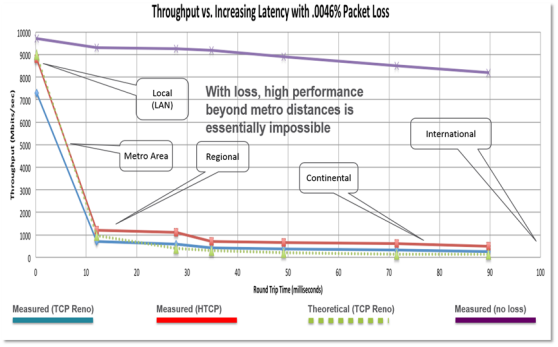

그림. 1. 패킷로스 발생에 따른 네트워크 전송 성능 [3]

그러나 대부분의 과학분야 응용프로그램들은 TCP 기반으로 발전해 왔기 때문에 새로운 기술 체계로의 전환은 어렵다. 따라서 이러한 문제를 개선하기 위해 기존 네트워크 환경의 변화 없이 빅데이터 고속도로 체계를 구축하는 기술인 ScienceDMZ 기술의 적용을 통해 이러한 전송 문제를 해결한다. 또한 인공지능의 경우 GPU 기반의 컴퓨팅 환경이 필수적이지만 한 연구자 또는 대학에서 보유하고 있는 GPU 자원의 한계로 인하여 고성능의 계산은 어려운 실정이다. 이러한 연구 환경에 대해 ScienceDMZ기반의 빅데이터 고속도로를 통한 고성능 전송 환경을 분산 환경에 적용하여 분산되어 있는 전산 자원들을 오케스트레인션 기법을 통하여 연동함으로써 분산되어 있는 GPU자원들을 단일 자원화하여 사용하게 된다.

본 논문에서는 이러한 빅데이터 고속도로 체계와 AI 컴퓨팅 체계를 연계하여 데이터 집약형 과학분야의 빅데이터의 전송과 이를 기반으로 AI 컴퓨팅 환경의 개선을 위해 출연연을 중심으로 ScienceDMZ의 구축과 GPU기반의 연계 자원들의 연계 구축을 하였다.

Ⅱ. 본론

본 논문에서는 출연연 중심의 빅데이터 고속도로 체계 구축을 위해 ScienceDMZ 기술을 적용하여 구축 하였다. ScienceDMZ는 E2E 환경의 연구자원의 고성능 전송 수준으로 연동하는 기술로서 분산된 대용량 데이터의 획기적인 전송 성능과 공유를 위한 기술 체계이다. [4]

ScienceDMZ의 기본 구성 요건은 다음의 3가지를 만족시키는 환경이 되어야 한다. 1) 연구 전용 트래픽과 일반 업무용 트래픽의 분리, 2) 데이터 전송을 위한 전송 전용 노드의 구축 (DTN-Data Transfer Node), 3) 성능 보장을 위한 모니터링 시스템. 앞서 언급되었듯이 패킷로스는 고대역망의 성능에 치명적인 영향을 끼치는 요소인데 주로 엘리펀트 트래픽으로 구성된 연구전용 데이터보다 마이스 데이터들로 구성이 되는 업무용 트래픽에서 많이 발생한다. 이메일, 웹서핑, 프린터 등이 이에 해당된다. 이 두가지 상이한 트래픽을 분리함으로써 패킷로스 프리 네트워크를 구성한다. 또한 패킷로스의 발생은 고성능 스위치 장비 보다는 방화벽과 보안 장비에서 주로 발생하기 때문에 망분리를 하면서 방화벽을 우회하는 방법을 통해 패킷로스의 발생을 최소화 한다. 방화벽을 우회한 사이트들 간에 신뢰구역(Trust-Zone)을 형성하여 스위치 장비에서 ACL을 통해 접근제어를 하는 동시에 독립망으로 구성한다. 이것은 기존 네트워크 환경의 변화 없이 고성능 전송 성능을 달성하는 기술로서 현재 국가과학기술연구망의 다양한 사용자들의 요구를 통해 출연연 중심으로 대학과 연구소를 중심으로 구축이 진행 중에 있다. 다음의 그림2는 ScienceDMZ의 기본 구성도이다.[4]

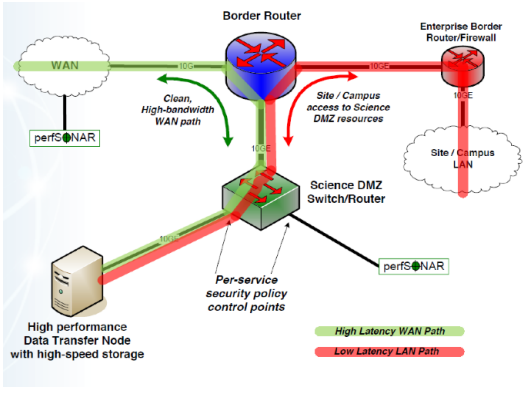

그림. 2. ScienceDMZ 기본 구성도

데이터 전송 전용 노드인 DTN은 일종의 전송 가속기로서 OS의 고속화 튜닝과 SSD/NVMe, 고성능 NIC, GPU를 장착하며, ScienceDMZ 구축 사이트의 WAN 환경에 가장 가깝게 구축한다. DTN을 통해 그림2에서와 같이 보더 라우터에서 ScienceDMZ 트래픽과 일반 트래픽을 분리하여 연구전용 데이터들을 고속으로 전송하는 역할을 하게 된다. 앞서 언급한대로 ScienceDMZ 구성요건을 토대로 구축된 ScienceDMZ 사이트에서 DTN을 통한 고속 전송 결과와 일반 전송 결과는 다음의 표 1과 같다.

ScienceDMZ 환경에서 DTN의 역할은 전송 가속기로서의 역할과 분산 컴퓨팅 노드의 역할을 하게된다. 특히 AI를 위한 분산 컴퓨팅 환경을 위해 각 DTN 노드에 GPU를 장착하고 쿠버네티스 기반의 오케스트레이션을 통해 자원의 컴퓨팅 자원의 통합과 리소스 관리를 한다.

src	dst	speed	status
KISTI	KRICT	742.96Mbps/1GE(74.2%)	S-DMZ
KISTI	KRIBB	785.12Mbps/1GE(78.5%)	S-DMZ
KISTI	KIOM	744Mbps/1GE(74.4%)	S-DMZ
KISTI	KIER	70Mbps/100ME(70%)	S-DMZ
KISTI	KARI	2Gbps/10GE(20%)	normal
KISTI	KAIST	5Gbps/40GE(12.5%)	normal

표. 1. ScienceDMZ & 일반망에서의 전송 성능 결과

쿠버네티스 기반의 분산된 DTN들을 활용한 AI 컴퓨팅을 위한 연계 방안은 다음의 그림 3과 같으며, 쿠버네티스 마스터를 구축하고, DTN을 워커노드로 m 마스터에 연동을 한다. 쿠버네티스 환경을 포털로 구축하여 워커노드의 CPU, GPU, Memory, Storage 자원을 관리하고, Jupyter Notebook을 통해 GPU 자원을 할당받아 파이선, R 등을 통해 AI 연산을 수행한다. Storage 기능은 DTN 자체의 SATA, SSD, NVMe를 활용할수도 있으며, Rook을 통해 Ceph 스토리지를 관리하여 대용량 데이터를 저장한다.

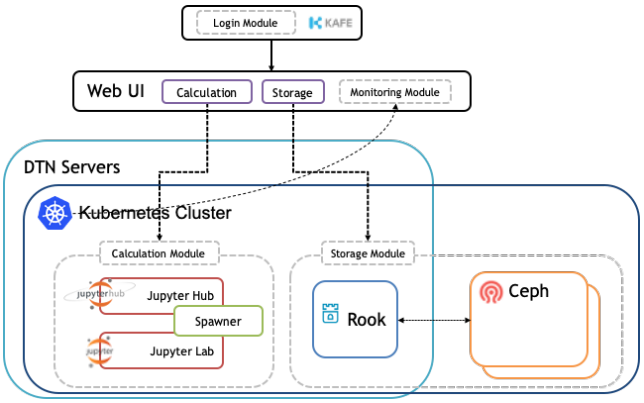

그림. 3. ScienceDMZ 환경에서 Kubernetes기반 AI 플랫폼 아키텍쳐

ScienceDMZ 기술의 적용을 통한 빅데이터 고속도로 체계와 AI 컴퓨팅 환경의 연계는 현재 KISTI외 9개 출연연을 중심으로 파일럿 환경으로 구축되어, 국내외 다양한 빅데이터 센터로부터 연구데이터의 전송을 하고 있으며, 각 분야별 다양한 인공지능 연구자들의 부족한 IT자원을 쿠버네티스 기반으로 연계된 GPU-enabled DTN을 통해 활용하고 있다. 아래의 그림4는 빅데이터 고속도로 체계와 AI 컴퓨팅 체계의 연계 파일럿 구성도이다.

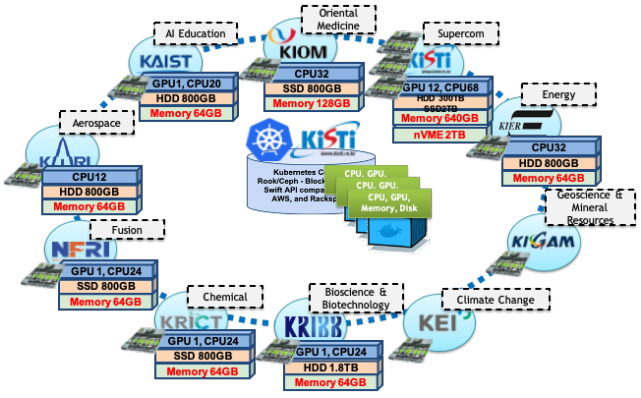

그림 4 출연연 중심 빅데이터 고속도로 체계와 AI 컴퓨팅 체계의 연계 구성도

ACKNOWLEDGMENT

국가과학기술연구회 “출연연 중심 AI 융합을 위한 빅데이터 Super Highway 융합클러스터” 사업

참 고 문 헌

[1] 지구 남반구 관측 천문 국제 프로젝트, URL, https://www.lsst.org/

[2] J. Postel. “Transmission Control Protocol. Request for Comments(Standard) 793”, Internet Engineering Task Force, September, 1981

[3] E. Dart, L. Rotman, B. Tierney, “The ScienceDMZ: A Network Design Pattern for Data-Intensive Science”, Scientific Programming Volume 22, Issue 2, Pages 173-185

[4] 미국국가연구망 ScienceDMZ, URL, ScienceDMZhttps://fasterdata.es.net/science-dmz/DTN/