

Science DMZ 환경에서의 컨테이너 기반 협업연구 플랫폼 구성 및 성능 측정

홍원택*, 이상권***, 김기현*, 문정훈*

*한국과학기술정보연구원 과학기술연구망센터

**과학기술연합대학원대학교 과학기술정보과학학과

{wthong, sglee, kkh1258, jhmoon}@kisti.re.kr

Deployment and performance measurement of containerized collaborative research platform in Science DMZ environment

Wontaek Hong*, Sangkwon Lee*, Kihyeon Kim*, Jeonghoon Moon*

*Advanced Kreonet Center, KISTI

**Univ. of Science and Technology

요약

고성능 네트워크를 기반으로 하는 과학응용 협업연구 환경에서 대용량 데이터의 고속 전송을 위해 활용되고 있는 전용의 데이터 전송 노드는 이용률을 감안하여 연산을 위한 부수적인 목적을 위해 활용이 가능하다. 본 논문에서는 이러한 분산된 데이터 전송 노드들을 활용한 컨테이너 기반 협업연구 플랫폼의 구성 사례를 소개하고 해당 플랫폼 상에서 분산 학습을 실험하고 소요시간 측면에서 성능을 측정하여 제시한다.

1. 서론

기하급수적으로 증가하는 연구 데이터의 공유 및 전송이 요구되는 과학응용 협업연구 분야에서 고성능 IT 연구 환경의 구축을 위해 많은 노력들이 행해지고 있다. 이러한 노력들은 컴퓨팅, 스토리지, 네트워크 자원의 질적 양적 측면에서의 향상과 클라우드 서비스에서의 베어메탈, 가상머신, 컨테이너 기술 등을 활용한 서비스 제공 측면에서의 변화를 모두 수용하기 위한 방향으로 발전해 가고 있다.

이와 관련하여 본 논문에서는 과거의 제약적인 과학 협업응용 연구 환경을 살펴보고, 최근 과학응용 분야에서 활발히 적용되고 있는 ESnet의 Science DMZ 개념[1]을 따라 국내 주요 출연연을 대상으로 구축된 연구 데이터 고속 전송 환경 하에서 데이터 전송 노드(DTN)의 활용도를 향상시킬 수 있는 컨테이너 기반 협업응용 연구 플랫폼을 소개하고 해당 플랫폼 상에서 분산 학습 실험을 수행하여 성능을 측정한다.

2. 본론

최근까지도 캠퍼스 또는 개별 연구소의 과학응용 분야 연구자들은 연구에 필요한 IT 자원들을 각각의

실험실에 직접 구축하는 경우가 많고 이러한 개별적인 연구 환경의 구축은 IT 자원의 구축 및 관리에 수반되는 시간, 공간, 경제적인 문제를 발생시킴으로써 본연의 연구 업무에 집중하기 원하는 많은 연구자들을 좌절하게 만들었다. 설령, 요구되는 고성능 서버, 저장 장치들이 준비되었다 하더라도, 원거리에 존재하는 협업 파트너와 연구를 진행하기 위해서는 실험실 네트워크에서 백본 네트워크로 연결되는 구간에서 발생하는 Last 1 mile problem에 의해 대용량의 연구 데이터의 전송 및 공유가 어려웠다.

이러한 문제를 해결하고자 [2]에서는 고에너지물리 분야의 특정 응용을 대상으로 고성능 컴퓨팅 및 전송 서버 개념을 도입하여 캠퍼스 또는 실험실에서 발생하는 Last 1 mile problem을 해결하고자 하였다. 이러한 접근의 주요 핵심은 고속 전송 프로토콜이 탑재된 고성능 서버, 저장 장치들을 고성능 백본 네트워크에 일종의 프락시 개념으로 구축하여 제공함으로써 캠퍼스 또는 개별 실험실에서의 병목 현상을 임시적으로 해결하는 것이었다. 그러나 이러한 시도는 대용량 파일의 전송 성능 향상 측면에 초점을 두었으므로, 서로 다른 과학응용 분야의 다수의 사용자가 동시에 원하는 만큼, 원하는 시점에 이용할 수 있도록 하기에는 제약사항이 있었다. 특히, 자원 스케줄

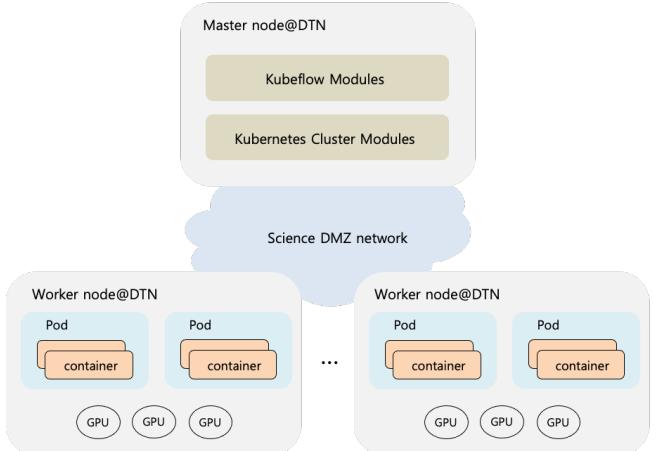
링 및 모니터링, 사용자의 인증 및 인가 등의 기능적 요소가 결여되어 실제 서비스로서의 한계가 있었다.

비슷한 맥락에서 과학기술협업연구 플랫폼[3]은 고성능 백본 망에서 가상머신 기반의 클라우드 컴퓨팅 개념을 활용하여 컴퓨팅, 저장 장치 등을 포함한 연구 자원의 제공을 통해 [2]에서의 서비스 관점에서 결여된 부분을 해결함으로써 과학응용 연구자들에게 대용량 데이터의 고속 전송을 포함한 연구 환경의 편의를 제공한다. 특히, 다수의 사용자를 지원하기 위한 사용자 인증 및 인가, 자원 풀링, 스케줄링 등의 필요한 개념들은 오픈 소스 기반의 클라우드 관리 플랫폼의 핵심 기능들과 매칭되어 제공된다.

현재까지도 가상머신 기반의 클라우드 컴퓨팅 환경은 과학응용의 특성에 따라 활발히 이용되고 있고, 최근에는 가상머신 대비 경량화된 컨테이너 기반의 인공지능, 머신러닝, 딥러닝 등의 환경을 필요로 하는 과학응용 분야의 요구가 많이 발생해 오고 있다. 한편, 연구망 측면에서는 ESnet에서 제안된 Science DMZ 개념[1]을 바탕으로 연구 자원들이 밀집된 실험실 및 연구소들의 네트워크 상황이 더욱 고도화되고 있고, 그 중심에는 전용의 데이터 전송 노드(DTN)가 배치되어 대용량 데이터의 고속 전송을 수행한다. 하지만, [4]에서도 지적되었듯이 사이트의 특성에 따라 일부의 DTN들은 시스템의 평균 이용률이 낮아 전송에 참여하지 않는 일정 시간동안 전송이 아닌 다른 목적으로 활용할 수 있고, 그러한 시도들 중에는 GPU 자원들을 제공하는 컨테이너 기반 클라우드 플랫폼 구축에 활용이 가능할 것이다.

현재 국가과학기술연구망에서는 Science DMZ 네트워크를 기반으로 분산되어 위치한 DTN 노드들을 대상으로 대덕연구단지 내의 출연연들이 참여하는 컨테이너 기반 협업응용 연구지원용 클라우드 플랫폼을 개발, 구축 및 운영해 오고 있고, 핵심적인 기능으로는 대용량 파일 고속 전송 툴이 탑재된 DTN 중심의 고속 전송 기능과 GPU 자원의 활용이 가능한 인공지능 과학응용연구 환경을 제공한다[5]. 해당 플랫폼은 그림 1에서와 같이 Science DMZ 네트워크 환경 하에 배치된 DTN 노드들을 대상으로 컨테이너 기반의 GPU 자원 제공을 위해 Kubernetes 주요 컴포넌트들이 마스터와 워커 노드들에 설치 및 설정된다. 이를 바탕으로 Jupyter 노트북, TensorFlow 환경을 제공하는 계산 Pod 생성, 파일시스템 볼륨 생성, 참여 노드 관리 및 대시보드 등의 세부 기능들을 제공한다. 추가적으로 본 논문에서 초점을 두고 있는 참여

기관별 “자원 연합”이라는 관점에서 자원 풀링화된 노드들을 활용하기 위해 Kubeflow 모듈들이 설치 및 적용된다.

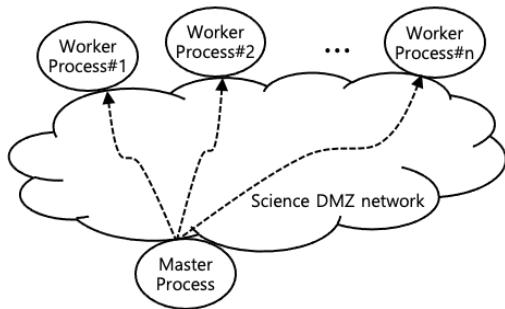


(그림 1) DTN을 활용한 협업연구 플랫폼

본 논문에서는 그림 1의 플랫폼 환경에서 TensorFlow 클러스터를 구성하고, 해당 클러스터에서 분산 처리 수행의 성능을 측정한다. 일반적으로 분산 처리를 위한 TensorFlow 클러스터 환경은 동일한 로컬 네트워크에 존재하는 워커 노드들을 대상으로 구동되지만, 그림 1에서는 참여하는 노드들이 Science DMZ 네트워크 상에 지역적으로 분산된 DTN 노드들로 구성되는 차이점이 존재한다. 이러한 접근은 friction-free Science DMZ 네트워크를 기반으로 DTN 노드들이 참여하므로 가능하고, 결과적으로 참여하는 노드들의 유연한 “자원 연합”을 유도 할 수 있다.

그림 2에서처럼 기본적으로 Kubernetes 기반의 컨테이너 자원들이 분산되어 존재할 때, TensorFlow 클러스터 상에서 분산 처리를 수행하기 위해서는 연산을 분산 배치하는 마스터 프로세스와 실질적인 연산을 수행하는 워커 프로세스들이 다른 서버에서 실행된다. 그리고, 이를 유연하게 구동하기 위해 Kubeflow를 도입하여 풀링화된 분산된 자원들을 원하는 시점에 이용한다. Kubeflow는 Kubernetes 기반의 머신러닝 워크플로우의 구성을 돋고 배포하기 위한 일종의 머신러닝 툴킷이다. 본 실험에서는 이러한 환경에서 Kubeflow의 “TFJob” 오브젝트를 생성함으로써 Science DMZ 네트워크 상에 분산된 노드들을 연계하여 TensorFlow job을 구동하는 실험을 수행한다. 추가적으로 “affinity” 속성을 이용하여 참여하는 노드들 중 이용 가능한 2개의 노드들을 지정

한다.



(그림 2) 분산 TensorFlow 클러스터 구동

본 실험에서 이용하는 EMNIST 데이터 세트는 “NIST Special Database 19”에서 파생된 손으로 쓴 문자, 숫자의 집합으로써 기존 MNIST 데이터 세트와 직접 일치하는 28x28 픽셀 이미지 및 구조로 변환되어 있다[6]. 이 전체 세트는 6개의 다른 데이터 세트로 나뉘고 본 논문에서는 47개의 클래스와 131,600 개의 샘플을 포함하는 “Balanced” 데이터 세트를 이용한다. 또한 NIST에서 제공하는 기본 포맷인 Binary 포맷과 Matlab 포맷 대신 학습을 수월하기 위해 Kaggle(<https://www.kaggle.com/>)에서 제공하는 csv 포맷의 Train/Test 데이터 세트를 이용한다. 실험에서 이용되는 학습 모델은 하나의 은닉 계층과 소프트맥스 계층을 갖는 CNN 모델로서 공개된 버전의 분산 MNIST 학습/검증 코드를 그대로 가져와 EMNIST “Balanced” 데이터 세트를 적용할 수 있도록 수정한 코드를 이용하여 처리 시간을 비교한다.

위의 실험 환경에서 지정된 2개의 노드들은 각각 TITAN Xp, GTX 1080Ti를 장착한 DTN 노드들이다. Science DMZ 네트워크로 연계된 노드들의 클러스터링 효과를 확인하기 위해 단일 노드로 동작하는 경우와 멀티 노드로 구성되는 경우로 나누어 2000회의 학습을 각각 2회 반복하여 수행한 후 소요된 평균 시간을 측정하였다. 측정 결과 단일 노드의 구성에서는 약 2993초가 소요되었고, 멀티 노드의 구성에서는 약 2179초가 소요됨을 확인할 수 있었다. 이는 비록 로컬 네트워크 환경은 아니지만 지리적으로 분산되어 학습에 참여하는 워커노드 수의 증가에 따른 성능 향상에서 기인한다.

3. 결론

본 논문에서는 Science DMZ의 핵심 구성 요소인

DTN을 활용하여 구축된 컨테이너 기반 협업응용 연구 플랫폼 상에서 지리적으로 분산된 워커노드를 이용하여 EMNIST “Balanced” 데이터 세트를 대상으로 분산 학습 소요시간 측면에서 성능을 측정하였다. 이러한 시도는 DTN 자체의 데이터 고속 전송 기능을 제공함과 동시에 DTN의 활용도를 높이고, 해당 플랫폼에 참여하는 노드들의 “자원 연합”을 유도하기 위해 활용될 수 있을 것이다.

ACKNOWLEDGMENT

본 연구는 2021년도 한국과학기술정보연구원(KISTI) 주요사업 과제로 수행한 것입니다.

참고문헌

- [1] E. Dart et al., "The Science DMZ: A Network Design Pattern for Data-Intensive Science", SC'13 Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, Denver, USA, Nov. 2013.
- [2] W. Seok et al., "A Study on End-to-End Performance Enhancement for Remote Large Data Transfer", Journal of KICS, vol. 32, no.6, pp. 367-374, 2007.
- [3] Y. Kwon and W. Hong, "Design and Prototyping of Scientific Collaboration Platform over KREONET", KIPS Tr. Comp. and Comm. Sys., vol. 4, no. 9, pp. 297-306, 2015.
- [4] Z. Liu et al., "Cross-Geography Scientific Data Transferring Trends and Behavior", Proceedings of the 27th International Symposium on High-Performance Parallel and Distributed Computing, New York, USA, Jun. 2018.
- [5] K. Kim et al., "Connecting Method Research of Distributed Computing for AI Research based on ScienceDMZ", Journal of KICS, In press, 2021.
- [6] G. Cohen et al., "EMNIST: an extension of MNIST to handwritten letters", 2017, Retrieved from <http://arxiv.org/abs/1702.05373>