

# 대용량 과학데이터 전송을 위한 플랫폼 구현

이민선\* · 유관종\*\*

## Implementation of a Platform for the Big Scientific Data Transfers

Min-Sun Lee\* · Kwan-Jong Yoo\*\*

### 요약

지난 수십 년간, 컴퓨터 기술의 발전은 응용연구의 데이터집약형 연구라는 새로운 패러다임으로 변화를 가져왔다. 지리적으로 떨어진 연구자들에게 대용량 실험 데이터를 저장, 공유, 분석하는 환경이 요구되고, 우리나라는 연구 경쟁력 강화를 위해 국가과학기술연구망(KREONET)과 대용량데이터센터 등 국가 차원의 인프라를 서비스하고 있다. 특히, KREONET는 2016년부터 세계 주요 연구지까지 최대 100Gbps의 속도로 연결하는 네트워크를 제공하고 있다. 본 논문에서는 국제간 협업을 지원하는 초고속 장거리 전송의 연구 환경으로서, 고성능 전송 노드와 계산 자원을 고속의 네트워크에 연동한 대용량 데이터전송용 전송플랫폼 시연 결과를 소개하고자 한다.

### ABSTRACT

Over the last several decades, the advances in computer engineering technology have led a new paradigm of data-intensive research in the field of scientific applications. A collaborative work environment for storing, sharing and analyzing data is required for researchers of geographical distance. The Korean government supports the Research & Education network(KREONET) and the Global Science experimental Data hub Center(GSDC) to strengthen the nation's competitiveness. The KREONET has upgraded its backbone to 100Gbps to accommodate demand to transfer data fast among the global major experimental sites. This paper introduces the test result between high performance nodes reserved for big data transfer.

### 키워드

R&E network, science DMZ, SDN, 클라우드, Collaborative Science  
연구망, 사이언스 DMZ, 소프트웨어 정의 네트워크, 클라우드, 협업 연구

## 1. 서론

1965년, (주)인텔의 공동설립자인 고든 무어는 컴퓨터의 중심부품인 반도체 집적회로의 성능이 18개월마다 2배로 증가할 것이라 예측했고[1], “무어의 법칙”이라 불리는 이 이론은 수십 년간 컴퓨터 기술 개발의 역사와 속도를 말하는 상징적인 표현처럼 사용되어왔다. 컴퓨팅 속도는 이제 초당 페타 플롭스( $10^{15}$ 년

의 부동 소수점 연산을 하는 속도)의 단위를 다루고 있다. 이러한 고성능 컴퓨팅 및 네트워크를 통한 데이터전송기술의 발전은 첨단 과학 응용분야에도 큰 영향을 주었다. 천문연구자들은 대규모 천체 관측 데이터를 수집하여 우주의 초미세 구조 연구를 수행하며, 기상 기후 연구자들은 기상 관측을 통해 얻은 측정 데이터로 기상 예측이나 예보에 활용한다[2, 3]. 월드 와이드웹(www)의 탄생이 입자물리연구소의 고에너지

\* 충남대학교 컴퓨터공학과(mleeh@cnu.ac.kr)

\*\* 교신저자 : 충남대학교 컴퓨터공학과

• 접수일 : 2018. 07. 13

• 수정완료일 : 2018. 07. 29

• 게재확정일 : 2018. 08. 15

• Received : Jul. 13, 2018, Revised : Jul. 29, 2018, Accepted : Aug. 15, 2018

• Corresponding Author : Kwan-Jong Yoo

Dept. Computer Science & Engineering, Chungnam National University,

Email : kjyoo@cnu.ac.kr

지물리 데이터 교환에서 시작되었음은 이미 잘 알려진 사실이다. 입자가속기나 전파망원경을 통해 수집한 대용량의 실험데이터가 연구자의 실험실까지 전달되는 것은 생각보다는 쉽지 않다. 실험에 필요한 연구 장비는 연구자로부터 지리적으로 멀리 떨어져 있고, 더구나 최근에는 실험의 규모와 정확도를 높이기 위해 여러 나라가 공동연합체를 구성해 운영하는 대형 국제 프로젝트로 진행되기 때문에 실험데이터를 전세계 수많은 개별연구자에게 배포하는 과정도 중요한 연구과정이 되었다. 이들 연구 분야의 데이터는 수십~수 백 테라바이트의 규모이고, 세계최대의 가속기인 거대강입자가속기(Large Hadron Collider: LHC<sup>1)</sup>)에서는 연간 40 ~ 50 페타바이트의 데이터를 만드는 것으로 알려져 있다.

LHC의 컴퓨팅 모델은 계층(Tier)구조를 가지며, 실험장비가 위치한 스위스 CERN이 Tier 0이며, 주요 국가 연구소 및 슈퍼컴퓨팅 센터가 운영하는 13개 Tier 1 센터가 있다. 또한 160개의 Tier 2 센터는 전세계의 대학 및 대규모 실험실에서 운영한다. 약 300개에 달하는 Tier 3 클러스터가 있으며, 모든 Tier 그룹은 연구/교육용으로 서로 연결되어 있다.

최근 몇 년간 세계 연구/교육망은 10Gbps의 백본에서 100Gbps로 증속되는 추세에 있으며, 2016년 우리나라의 KREONET<sup>2)</sup> 백본도 100Gbps로 증속되었다. 본 논문은 고속네트워크 기반의 전송 플랫폼을 구현하여 시연한 결과를 소개하고자 한다. 먼저, 2장은 대용량 과학 데이터 전송 관련 기술을 소개하고, 3장에서는 데이터 전송을 위한 플랫폼 구현과 테스트 결과를 소개한다. 마지막장에서는 연구에서 얻은 결론과 향후 연구방향을 제시한다.

## II. 관련기술 소개

### 2.1 소프트웨어 정의 네트워크(SDN)

소프트웨어정의네트워킹(SDN)기술은 SDN을 초기에 시작한 Nicira가 VM웨어에 인수되면서부터 알려지기 시작했다. 기존의 데이터센터가 컴퓨팅서버와 저

장장치를 가상화하였지만, 실제 네트워킹에 필요한 수작업으로 인해 전체 효율성이 떨어지는 문제를 안고 있었다. 결국 각각의 컴포넌트 뿐 아니라 데이터센터 전체를 가상화하려는 노력이라 할 수 있다. 네트워크의 경계가 없어지면 장비가 어디에 있던지 어떤 종류이던지 연결이 가능해진다. 이제 SDN은 네트워크를 공급자위주에서 사용자중심으로 또 하드웨어에서 소프트웨어로 패러다임을 바꾸고 있다.

네트워크 장비는 다른 어떤 자원보다 더 하드웨어 종속성이강해 장비 업체마다 다른 주문 형 반도체(ASIC)를 사용하는 등 폐쇄적인 설계로 유명해서 복수업체의 제품을 사용할 때 통합관리를 힘들게 하였다. 이에 네트워크의 하드웨어종속성을 없애고자 Open-flow가 출현하였고 결국 이는 네트워크 장비업체가 제공하는 비싼 소프트웨어 없이도 범용의 프로그래밍언어로 네트워크 관리 기능을 구현한 애플리케이션 개발이 가능하게 하였다[4].

SDN은 네트워크 장비의 역할에서 데이터전송 처리만 남기고 그 외 모든 Configuration, 액세스제어목록(ACL), 배포 등의 기능을 컨트롤러에게 맡겨 네트워크 인프라를 쉽고 빠르게 운영하고자한다. 유럽의 글로벌 통신업자들이 연합하여 시작된 네트워크 기능 가상화(NFV) 기술은 통신업자들이 사용하고 있는 네트워크 장비의 여러 기능을 분리시켜 소프트웨어적으로 제어 및 관리할 수 있는 가상화시키는 기술이며 특히 데이터 센터 내에 위치한 대용량 서버, 스토리지, 그리고 스위치를 분리하여 접근 가능한 네트워크 구조를 만든다. 이를 통해 네트워크 장비 비용 절감은 물론 새로운 서비스 시장 투입에 필요한 시간 단축(time-to-market), 유연하고 혁신적인 서비스가 가능하다.

대용량 데이터 전송이 필요한 응용분야에서는 망의 동적 할당을 통해 특정 전송에 대해 보장된 대역폭을 보장 받을 수 있었다. 이제 SDN기술이 적용된 경로 구축 소프트웨어를 통해 다중 도메인 환경에서 종단간(end-to-end) 동적 경로를 생성하여 높은 전송률로 대용량의 데이터를 전송한다[5].

1) Large Hadron Collider-CERN, <https://home.cern/topics/large-hadron-collider>

2) KREONET, <http://kreonet.net/>

### 2.2 Science DMZ

미국 에너지부(Department of Energy)의 연구망인 ESnet<sup>3)</sup>에서 개발된 Science DMZ는 장비나 환경설정 혹은 보안정책이 첨단 응용과학분야에 최적화되도록 설계된 네트워크의 일부로 캠퍼스(혹은 실험실)의 로컬 네트워크 주변에 구축된다[6].

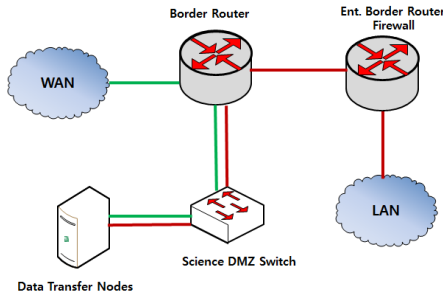


그림 1. Science DMZ 구성도  
Fig. 1 Science DMZ diagram

Science DMZ의 장점은 다음과 같다. 첫째, 과학용 네트워크는 로컬네트워크 내 다른 범용 네트워크와 방화벽을 공유하지 않아도 된다. 과학데이터 전송은 웹서핑이나 이메일 보내기 등 일반 사무용 작업과는 다른 요구사항이 있기 때문에, Science DMZ 구조는 과학데이터 전송트래픽을 범용의 트래픽과 분리하여 그림 1과 같이 데이터 전송 노드를 통해 전달한다. 이는 결과적으로 과학 연구자와 일반사용자 모두를 만족시키는 성능의 향상을 준다. 둘째, Science DMZ 장비와 구성은 패킷 손실 없이 고속의 전송을 가능하게 한다. 셋째, 보안이나 접근 정책 등을 과학적 사용자 목적에 맞게 조정이 가능하다. 마지막으로 SDN, 100G 기가비트 서비스 등, 최신의 네트워크 기술의 적용이 쉽다[7, 8].

### 2.3 Flash I/O Network Appliance(FIONA)<sup>4)</sup>

고대역폭(10~40Gbps)의 연구망을 사용하는 연구자들에게 경제적인 가격대에 대용량 데이터 연구에 최적화된 단일(혹은 클러스터링 된) 노드로 개발된 PC가 FIONA이다. 40Gbps 네트워크 인프라 디버깅용으

로 제작되었지만 현재 많은 연구자들이 데이터전송노드로 활용하고 있다. FIONA는 수십에서 수백 테라바이트의 데이터를 저장하고 전송하는데 최적화 되어있으며, 다음 표 1과 같은 사양과 예산으로 초기에 제작되어 배포되었다. 약 \$7,000정도의 예산으로 최적의 데이터전송노드 FIONA를 구축할 수 있는 것으로 보고되었다. 그림 2는 현재 대전 한국과학기술정보연구원안에 설치된 FIONA이고, 상세한 사양은 표 2에서 설명하고 있으며, 표 3은 데이터전송노드의 튜닝내용을 보여준다.

표 1. DTN 사양 및 예산  
Table 1. DTN Specification and budget

Specification
Flash disk(2~16TB)
Hard disk (5 - 120 TB), 6-20 cores
Network Interface card: two 40G/10G
Optional GPUs
Budget
about \$2,000 : perfSONAR 40G mem-to-mem testbox
up to \$7,000 : plus flash disk for DTN
\$15,000 or more : plus 100TB for local server, 2-3U for network switch rooms

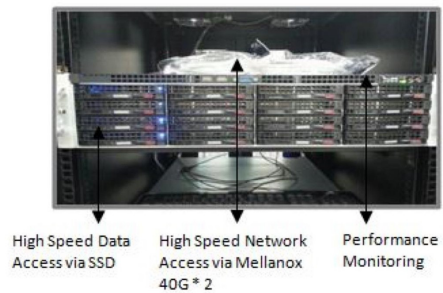


그림 2. KREONET-DTN 하드웨어  
Fig. 2 Data transfer node H/W on KREONET

### 2.4 Fast Data Transfer(FDT)

FDT는 광역네트워크(WAN)에서 디스크 속도로 읽고 쓸 수 있는 효율적인 데이터 전송용 응용 프로그램으로, JAVA로 작성되어 있고, 대부분의 플랫폼에서 실행가능하다. FDT의 특징은 독립적인 쓰레드

3) ESnet, <https://www.es.net>

4) FIONA at PRP project, <http://prp.ucsd.edu/>

를 사용하여 각각의 물리 장치에서 읽고 쓴다. TCP 소켓을 통해 데이터세트를 지속적으로 스트림하며, 필요시 TCP 스트림에서 병렬로 데이터를 전송한다. 결과적으로 FDT를 사용하면 전체 파일 세트를 대량으로 스트리밍 할 수 있으므로 수천 개의 파일 세트를 최고의 속도로 주고받을 수 있다[9-11].

표 2. 하드웨어 구성 장비  
Table 2. Hardware overview

Intel Xeon™ E5-2643v3 (3.4GHz, 20MB, 60Core )
128GB DDR4 ECC REG PC4017000R (8 x DDR4 Slots, Max 1TB)
Supernmicro MBD-X10SRL-F Server Motherboard
Mellanox ConnectX-3 Pro EN Single-port 40/56 Gigabit Ethernet - Driver : mix4_en - version: 3.4-1.0.0(25 Sept 2016) - Firmware-version: 2.36.5150
SSD Samsung 850 Evo 1TB - Cache : 1GB - Sequential Read : 540MB/s - Sequential Write : 520MB/s
CentOS 7.2 : 3.10.0-327.36.3.el.x86_64

표 3. DTN 튜닝  
Table 3. DTN tuning

Networking tuning (NIC, OS) BIOS I/O Scheduler File System(EXT4 Tuning) RAID controller (SSD+SATA) Virtual memory Subsystem SSD Tuning Benchmarking DTN Performance Test
--

### III. 구현 및 테스트

#### 3.1 데이터전송 플랫폼 구현

연구자 중심의 고성능 ICT 연구 환경을 제공하기 위해 고성능 전송 및 보안에 관련된 문제를 먼저 고려하여야한다. 많은 응용분야에서 TCP기반의 전송을

선호하고, TCP는 패킷손실이 전체 네트워크 성능에 큰 영향을 주기 때문에, 범용의 데이터 트래픽과 연구자의 과학데이터가 분리된 구조의 Science DMZ 기술을 적용하였다. 고성능의 데이터 전송을 위해서는 데이터 노드의 역할이 중요하기 때문에, 저장 및 접근제어에 관한 기능을 함께 포함하여 고려하였다. 실제 네트워크 장비(라우터, 스위치 등)에서 기본적으로 제공하는 에러(error) 모니터링은 패킷카운터를 제공하지 않기 때문에 perfSONAR 프로그램을 활용하여[8], 실시간 성능 검증을 시도하였다.



그림 3. 성능모니터링 매트릭스  
Fig. 3 Performance monitoring

저가의 서버를 통해 구성이 가능한 FIONA(표2 참조)를 구축하였고, PCIe에 40GE NIC을 통해 100GE에 연결하였다. 최적화된 성능을 얻기 위해 표3과 같은 튜닝작업도 수행하였다. 그림 3은 perfSONAR 프로그램이 제공하는 툴(Maddash)로 대전의 데이터 전송 노드와 연동된 해외 노드들 간의 성능을 모니터링하는 매트릭스이다.

#### 3.2 FDT를 사용한 국제구간 시연

지난 2016년 11월, 미국 유타주의 솔트레이크에서는 슈퍼컴퓨팅 컨퍼런스(SC'16)가 개최되었다. 이 기간 중, 한국 대전에 위치한 한국과학기술정보연구원과 솔트레이크 행사장의 캘리포니아공과대학교(Caltech) 부스를 초고속망으로 연결하였다. KREONET에서는 그림 2에서 소개된 FIONA기반의 데이터 전송 노드를 사용하여 “고성능 데이터 전송 시스템” 시연에 참여하였고, 평균 9.157Gb/s와 최고 9.488Gb/s의 전송속도를 보였다(그림 5 참조). 대전과 솔트레이크 시는

10,000마일 떨어진 거리로, 대전에서 미국 시카고 Starlight에 연동되어 솔트레이크 행사장 네트워크 SCinet까지 직접 10기가(Gigabit Ethernet)로 연결되었다.

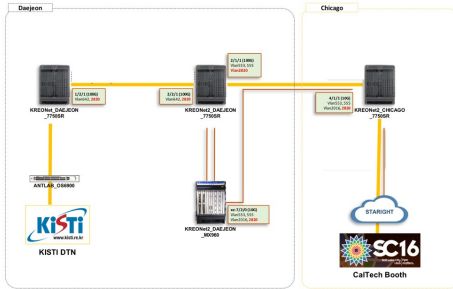


그림 4. 대전-솔트레이크간 네트워크  
Fig. 4 Network - Daejeon & Salt Lake City

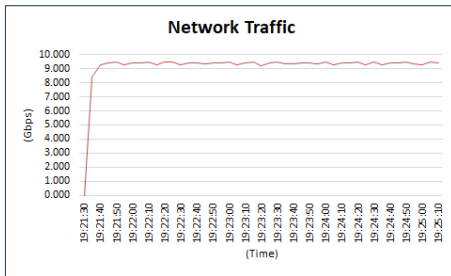


그림 5. 10기가 네트워크 데이터 전송 결과  
Fig. 5 Data Transfer over 10Gbps

### 3.3 iperf를 사용한 국내구간 네트워크 성능측정

SC'16 시연이후, FTP 프로토콜 계층에서의 성능향상 방식[11]을 적용하여, KREONET의 거점지(PoP)인 대전과 서울에 그림 6과 같은 10G기반의 고성능데이터 전송 플랫폼을 구현하였다. 전송 데이터는 데이터 전송노드(DTN-A, 대전)를 경유하여 DTN-A에 저장된 후 DTN-B(서울)로 전송된다. 예지네트워크가 전체 전송 성능에 최소한의 영향을 미치도록 설계된 구조로, 두 DTN은 WAN구간에 가장 가깝게 위치하도록 하였다.

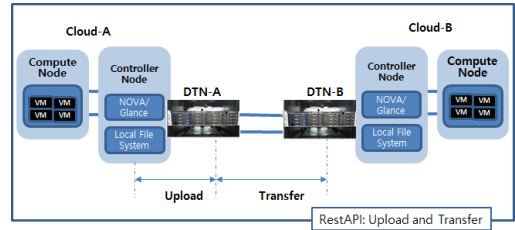


그림 6. 고성능데이터 전송  
Fig. 6 High performance data transfer

대전과 서울사이에 구현된 플랫폼의 네트워크 전송 속도를 확인하기 위하여, iperf 테스트를 진행하였고, 메모리-메모리간 성능은 평균 9.88Gbps (10GE기반)의 높은 속도를 보였다. 대전 사이트의 DTN은 그림 3에서 보여주듯이, 현재 미국 PRP프로젝트의 주요 사이트와 연동되어 있으며, 한국과 미국 간 약 7-8Gbps(10GE기반) 전송 성능을 보여준다.

## IV. 결론 및 향후 과제

대용량 데이터를 다루는 응용연구자들에게 데이터를 저장하고 원격지의 실험지에서 가져오고 또 동료 연구자에게 전달하는 일련의 작업은 무척이나 번거로운 일이지만 반드시 필요한 연구과정이다. 이런 연구자들에게 전송거리가 긴 국제 WAN구간에서 효율적으로 데이터를 전송하는 방안은 가장 중요한 이슈라 할 수 있다.

본 논문에서는 빠르게 대용량화하는 과학실험 데이터의 전송을 위한 플랫폼을 구현하여, 국제 및 국내네트워크상의 성능을 측정하였다. 대용량 과학데이터의 전송 성능을 보장하기 위해 Science DMZ 기술을 적용하였고, 네트워크의 성능측정을 완료하였다. 이제 100GE 환경에서 대륙 간 초고속 전송 및 분산된 실제(real) 실험 데이터의 효과적인 전송기술 적용이 필요하다. 본 논문에서 구현된 플랫폼은 대용량과학데이터의 국제간 전송에 더 효과적일 것으로 기대한다. 왕복시간(RTT)이 큰 전송의 경우, 에지네트워크의 혼잡성이 전체 성능에 큰 영향을 작용하기 때문에 향후, 특정 응용연구분야의 성능 향상을 정량적으로 검증하는 연구를 추가적으로 진행하고자 한다.

References

[1] G. Moore, "Cramming more components onto integrated circuits," *Electronics Mag.*, vol. 38, no. 8, 1965, pp. 114-117.

[2] M. Song, Y. Kang, and H. Kim, "The study on the design and optimization of storage for the recording of high speed astronomical data," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 12, no. 1, Feb. 2017, pp. 75-84.

[3] H. Yoon, "Development of Contents on the Marine meteorology service by meteorology and climate big data," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 11, no. 2, Feb. 2016, pp. 125-138.

[4] W. Stallings, "Software-defined networks and Openflow," *The internet protocol Journal*, vol. 16, no. 1, Mar. 2013, pp. 2-14.

[5] H. Newman, A. Barczyk, A. Mughal, S. Rozsa, R. Voicu, I. Legrand, S. Lo, D. Kcira, R. Sobie, I. Gable, C. Brown, Y. Savard, T. Tam, M. Hay, S. Mckee, R. Hocket, B. Meekhof, and S. Timoteo, "Efficient LHC Distribution across 100Gbps networks," In *SC Companion: High performance computing*, Salt Lake City, USA, Nov. 2012, pp. 1594-1599.

[6] E. Dart, L. Rotman, B. Tierney, M. Hester, and J. Zurawski, "The Science DMZ: A Network Design Pattern for Data-Intensive Science," In *Proc. of the Int. Conf. on High Performance Computing, Networking, Storage and Analysis*, Denver, USA, Nov. 2013, pp. 1-10.

[7] J. Park, S. Kim, and M. Noh, "Performance enhancement method through Science DMZ data transfer node tuning parameters," *Trans. on Computer and Communication Systems*, vol. 7, no. 2, 2018, pp. 33-40.

[8] J. Moon and M. Lee, "A Study to Data-intensive Science data transfer over Science DMZ," In *Proc. of Korea Computer Congress*, Jeju, Korea, 2016, pp. 1359-1361.

[9] Z. Maxa, B. Ahmed, D. Kcira, I. Legrand, A.

Mughal, M. Thomas, and R. Voicu, "Powering physics data transfers with FDT," *J. of Physics : Conference series*, vol. 331, no. 5, 2011, pp. 1-6.

[10] M. Lee, "A study on the throughput guarantee with TCP traffic control," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 11, no. 3, Mar. 2016, pp. 303-308.

[11] W. Seok, W. Hong, J. Kwak, and J. Moon, "DTN-aware Store-and-Forward Data Transfer for Exabyte scale science big-data," *J. of Korean Institute of Communications and Information Sciences*, vol. 42, no. 10, 2017, pp. 1991-1998.

저자 소개

**이 민 선 (Min-Sun Lee)**



1989년 숙명여자대학교 물리학과 졸업(이학사)

1997년 미국 네브라스카주립대학교 대학원 전기공학과 졸업(공학석사)

2012년 ~ 2016년 : 한국과학기술정보연구원 연구원  
2016년 ~ 현재 충남대학교 SW중심대학사업단 산학 협력중점교수

※ 관심분야 : 가상화 기술, 클라우드 컴퓨팅, 네트워크 통신 기술 등

**유 관 종 (Kwan-Jong Yoo)**



1976년 서울대학교 전산학과 졸업 (이학사)

1978년 서울대학교 전산학과 졸업 (이학석사)

1985년 서울대학교 전산학과 졸업 (이학박사)

1987년 ~ 1989년 : 충남대학교 전자계산소 소장

2001년 ~ 2006년 : 충남대학교 BK사업단 사업단장

2011년 ~ 2012년 : 충남대학교 교무처장

1979년 ~ 현재 충남대학교 컴퓨터공학과 교수

※ 관심분야 : 멀티미디어 어플리케이션, 클러스터 시스템, 병렬처리 기술 등