

디지털인프라중심의 연구망구조 디자인

노민기

한국과학기술정보연구원

mknoh@kisti.re.kr

Design of Research Network for Digital Infrastructure

Minki Noh

Korea Institute Science Technology and Information

요약

디지털전환에서는 생산품을 소비자에게 전달하는 유통 방식부터 소비자의 제품 선택 방식, 그리고 소비자의 제품선택의 유형과 필요성을 미리 파악하여 제품을 소비자에게 권하는 마케팅까지 혁신적으로 바뀌며, 이를 통해 생산-선택-유통-전달-분석에 대한 일련의 과정과 연계과정에 대한 큰 변화를 말한다. 과학기술현장 디지털전환은 데이터를 연구에 소비하는 연구자와 데이터를 생산하는 관측/실험현장와의 연계와 데이터흐름을 파악하여 효율적 연결과 활용현장으로의 전승, 연구자에게 필요한 데이터를 분류하고, 연구에 필요한 시기에 맞춰 제공하는 적시성이 중심이 연구환경의 변화라고 할수 있다.

I. 서론

과학기술분야의 디지털전환은 제4세대 데이터중심형과학의 발전과 함께 연구자들이 과학기술 데이터를 위치, 용량 등 물리적인 제한없이 활용하고 최종 연구산출물을 획득을 위해 수행되는 일련의 과정을 효율적으로 연계하는 연구환경의 변화이다. 이때 변화된 연구환경은 실험과 관측에 의해서 데이터 생산하는 인프라, 생산된 데이터를 관리하고 가공할 수 있는 저장공간과 계산, 분석시스템을 포함하는 컴퓨팅 인프라, 그리고 이를 연결하는 네트워크 인프라가 연계되는 디지털리서치인프라(Digital Research Infrastructure)와 생산, 수집, 분석, 관리되는 데이터를 대상으로 인공지능(AI), 디지털분석 기능 등을 통해 사용자들이 인프라와 데이터로의 접근이 용이하고, 연구목적에 맞게 활용성과 적시성을 제공하는 디지털전환플랫폼(Digital Transformation Platform)로 구성될 수 있다.

과학기술디지털전환(Digital Transformation on S&T)은 디지털리서치 인프라와 디지털전환플랫폼으로 구성되어 데이터위치-데이터분석/처리-데이터획득까지의 일련의 과정이 효율적으로 연계되어 연구자들이 해당연구에 최적의 데이터를 필요한 시간에 획득하여 최종 연구목표를 달성할 수 있는 연구환경으로의 변화라고 할수 있다.

본 논문에서는 국가과학기술 디지털리서치인프라를 구성하는 국가연구망(National Research Network)을 중심으로 과학기술데이터 생산지와 연구현장을 연결하는 환경적 구성을 연구하고, 일반적인 노드 또는 시스템간의 전송을 위한 네트워킹 기능이나 기술보다 디지털인프라를 구성하는 요소간 연계 역할에 집중하여 디지털인프라 구성요소의 역할수행에 필요한 네트워크 구성에 대한 디자인을 통해 데이터 생산지에서 데이터 소비자까지 데이터흐름에 대한 효율성을 제한한다.

II. 본론

관측장비와 실험장치에서 생산되는 데이터를 저장하는 스토리지와 데이터를 분석하는 컴퓨팅자원, 발생하는 대용량데이터를 연구자들이 쉽게 접근하기

위해 구성하는 통신장비, 통신장비간 연결, 그리고 통신을 통해 다른 도메인에서 연구자들이 필요한 결과를 처리하고, 이를 최종적으로 연구자들의 도메인까지의 전승, 대용량 과학기술 데이터를 생산지에서 최종 활용지역까지 전송하기 위한 네트워크 구성은 매우 다양하고, 복잡하다.

단일한 도메인에서 수행되는 연구보다 다수의 도메인이 연계되어 수행하는 커뮤니티형 협력연구가 발전하고 있는 현재의 과학기술현장에서 과학기술데이터의 생산과 분석, 가공과 저장, 연구현장까지의 서비스까지 이를 구성하는 시스템과 기능, 필요한 기술은 다양하고, 각 시스템과 도메인을 연결하고 데이터를 전송하는 네트워크의 구성, 기능, 기술은 더 복잡한 상태이다. <그림 1>

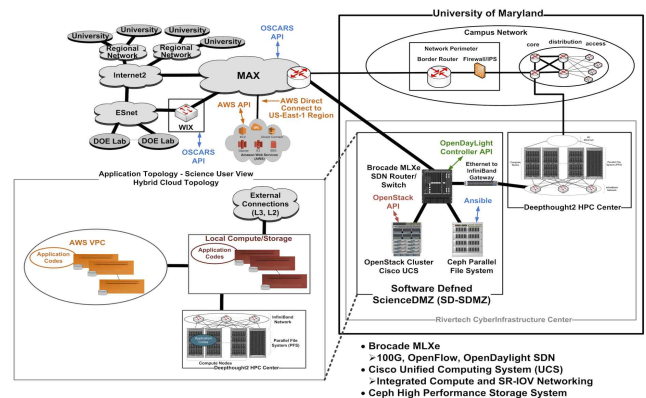


그림1. 대용량데이터처리 네트워크 구성 (출처: fasterdata.es.net)

1. 디지털전환중심 데이터전달구성

디지털전환에서 주요한 역할구성 생산-유통-전달-활용/분석을 기준으로 네트워크의 구성을 디자인하는 것은 각 구성요소의 요소기능 강화보다는 전체 데이터흐름을 매우 효율적으로 연결할수 있으며, 기능과 요소기술의 역할도 빠르게 적용가능하다.

첫 번째는 연구자들이 과학기술데이터를 자신이 수행하는 연구에 활용하기

데이터를 관측과 실험을 통해 직접 생산하는 역할이다. 이를 데이터생산(DS: Data Source)으로 역할로 정의할 수 있다.

두 번째는 데이터를 분석하거나 고성능 컴퓨팅자원으로 옮겨 데이터를 가공, 분석, 분류 등 가공을 하는 역할이며, 이는 데이터 생산지역 자체에서 혹은 대용량의 컴퓨팅자원을 보유한 지역에서, 또는 분산된 자원을 포함한다. 데이터 처리(DP: Data Processing)은 기능적으로 트랜잭션, 분산, 실시간, 일괄, 다중처리를 포함하며, 구조화된 데이터, 구조화되지 않은 데이터처리를 포함한다.

세 번째는 가공된 데이터를 연구에 직접 활용하는 역할을 수행한다. 대부분 연구자들이 위치한 지역에서 수행되며, 연구지역의 특성상 대용량의 데이터 송수신보다는 데이터 생산지와 가공지의 접근과 제어, 그리고 가공된 결과를 공동연구를 수행하는 타 연구자(지역)와 협력을 하는 역할을 수행하는 데이터의 활용(DU: Data Uptake)이다.

이처럼 연구환경에서 디지털전환 중심의 데이터활용과 역할측면에서의 구성은 데이터 생산-가공-활용으로 크게 분류할 수 있다. 이러한 역할측면에서 구성된 데이터의 특성은 아래와 같이 분류될 수 있다.

구성요소	데이터흐름(역할)	표기
데이터소스	관측/실험데이터 생성	DS
데이터처리	데이터 저장/제공	DP
데이터활용	활용/공유/	DU

표 1. 네트워크 역할중심 구성요소

위와 같이 데이터의 소스와 처리, 활용을 중심으로 과학기술데이터를 재구성하면 아래의 그림과 같다.

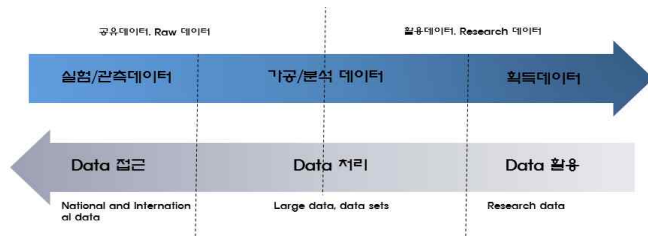


그림2. 디지털전환인프라 데이터 처리 및 관리

2. 디지털전환중심네트워크인프라

과학데이터의 대용량 관측과 실험에서 발생하는 데이터를 대용량데이터로 분류하고, 데이터의 처리와 저장, 관리하는 데이터의 규모는 연구데이터 인프라로 구성하며, 이러한 데이터는 과학기술데이터의 서비스로 활용된다.

역할중심의 요소와 데이터의 분류에 따라 구성된 구성요소와 데이터의 전송과 전달은 네트워킹을 통해 연결되고, 연계되어야 한다. 이때 크게 생산-가공-활용 3구역간의 디지털전환 연구환경 구성요소간 네트워킹으로 분류하고, 단위 시스템 또는 연구지역간 연결을 포함한다고 가정한다. 구성요소간 역할의 분류는 각 시스템이나 장비, IT자원등 지역적이나 기능적인 분류가 아니기 때문에 전송과 전달기능보다 더 많은 기능이 주어져야 한다. 이는 디지털 전환네트워크의 데이터흐름(Data Flow)로 정의하여 구성한다. 데이터흐름을 중심으로 각 구성요소의 역할을 중심으로 각 구성요소간 네트워킹의 역할을 정의하면 다음과 같다.

1) 접근(Access)

데이터생산지역(DS)은 실험과 관측 시설 혹은 해당 장비를 보유하고 있는 연

구기관을 의미한다. 이때 연구자들이 DS에 접근하는 관측과 실험시설의 운영과 생산데이터의 모니터링, 그리고 데이터 저장 및 가공지역으로의 데이터송신을 위한 접근이다. 이때 DS와 데이터가공 및 처리(DP)와의 연결에서 기본적으로 데이터 전송량의 예측과 안정적(실험, 관측시간동안) 전송이다.

DS-DP의 데이터흐름은 제어를 위한 접속과 모니터링, 실험 및 관측데이터의 대용량데이터 전송으로 이루어지며, 이는 대역폭중심으로 데이터흐름을 정의할 때 단방향 대용량데이터전송으로 이루어진다.

○ 기존망과의 구조적 차별성

- DS에 대한 접속, DS-DP간 대용량 전송 용량과 성능 보장
- 광네트워킹 기술을 중심으로 최소한의 네트워크 대역폭할당 기능

○ 성능저하 없는 경량화된 네트워크 보안 기술 적용

- 네트워크 트래픽 접근제어(ACL), 위협 트래픽 모니터링 등 성능저하가 발생되지 않은 보안기술 적용

2) 교환(Exchange)

DS에서 생산된 데이터는 일반적으로 DP로 전달된다. 이때 DP가 되는 대상은 다음과 같다.

① 대용량 데이터처리 역할을 수행할 수 있는 자원이 구성된 지역으로 다양한 대용량데이터의 1차 목적지가 된다.

이때 DS의 수에 따라 DP에 전달되는 데이터량은 증가한다.

② 연구현장에 생산된 데이터를 연구목적에 맞도록 분석, 가공할 자원이 있는 경우 DS-DP의 전송역할과 같다. DP-DS의 데이터량은 해당연구에 필요한 데이터로 고정된다.

○ 광역네트워크(WAN) 연결을 위한 경로설정

- 광패스(Light-Path/Layer1), 가상랜(Vlan 및 Trunk/Layer2), Virtual Circuit 및 SDN (Software-Defined Networking) 기술 적용
- 전송거리에 대한 응답지연시간에 따라 시스템 버퍼, BDP(Bandwidth Delay Product), MTU(Maximum Transmit Unit) 적용

3) 분류(Grouping)

데이터의 흐름을 중심으로 데이터의 생산과 활용지역뿐 아니라, 연구분야별로 데이터의 흐름을 나눌 수 있다. DS-DP의 데이터흐름이 데이터처리의 역할에 따른 분류라고 한다면, 연구분야 및 커뮤니티형태의 협업연구의 경우 연구분야에 따라 해당 데이터를 송수신할 때 해당 데이터의 분류를 통해 더욱 효율적인 데이터전달 역할을 수행할 수 있다.

데이터생산지역에서 데이터의 분류는 연결된 DS에서 전송되는 데이터의 합만큼 분류하여 구성되며 이때 생산지역과 가공지역에서는 생산된 데이터를 1차전송이라고 할 수 있다. 1차전송의 역할과 특징은 저장된 데이터와 전송된 데이터가 매우 대용량이며, 데이터 생산시간에 밀접한 관계를 가지고 있다. 처리된 데이터를 필요한 DU에 전달할 경우 DU에 포함된 연구자(기관)수 만큼을 분류해서 동시에 전송하는 방법도 효율적인 전송방법중 하나일 것 이다.

○ 전달되는 DU의 구성이나 데이터공유형태에 따라 달라지며, 이 경우 I2S 위칭 기술등 실시간으로 경로 및 대역폭을 조정할수 있는 네트워킹 기술이 필요하다.

구성요소	흐름	네트워킹 역할	네트워킹 기술
DS	→ DP	접근	자원할당
	→ DP'	교환	경로(Path)구성
DP	→ DP'	교환	경로(Path)구성
	→ DU	분류	전송성능

DU	→ DU'	분류	도메인구성
----	-------	----	-------

표 2. 네트워크 구성요소간 데이터 흐름

III. 결론

본 논문에서는 과학기술디지털전환 (Digital Transformation of Science & Technology)으로 구성된 연구환경 구축을 목적으로 각 구성요소간 데이터 흐름을 정의하고 이를 연계하는 네트워킹인프라와 기술을 통해 효율적인 연구환경 구축을 제시하였다. 과학데이터의 관측, 실험시설부터 발생하는 데이터를 효율적으로 데이터처리지역과 활용지역으로 전달하고, 연계된 자원을 통해 저장-분석/계산과 활용의 일련의 과정을 효율적으로 구성하는 것을 디지털전환 인프라의 역할로 디자인하였다. 디지털전환 중심의 네트워킹 디자인은 데이터 전송을 위한 시스템 또는 지역간 기능 제시에서 벗어나 전체적인 역할의 데이터흐름을 통해 접근성과 전송성능의 향상, 네트워크 자원의 효율적 구성을 디자인하여 우수한 연구환경 조성이 가능해진다.

이러한 디지털전환인프라로 업그레이드된 연구환경에서 데이터를 연구에 활용하는 과학자들은 자신의 연구에 집중하여 맞춤형 연구환경에서 우수한 연구결과가 창출이 가능하다. 또한, 디지털전환에 요구되는 네트워크 기술과 기능을 데이터의 생산-처리-활용에 역할에 맞도록 세부적으로 구성하고 디자인하는 것이 향후 추가적으로 진행되어야 할 연구이다.

참 고 문 헌

- [1] Developing a Digital research infrastructure strategy for Canada. 2014. <https://www.canaried.ca/>
- [2] KREONET 첨단연구사례집(2005-2020)
- [3] globus.org, last modified Sept. 18, 2019, accessed Sept. 18, 2019. <https://www.globus.org/>
- [4] perfsonar.net, last modified Sept. 12, 2019, accessed Sept. 18, 2019, <https://www.perfsonar.net/>
- [5] Shamir, A. "On the security of DES," Advances in Cryptology, Proc.Crypto '85, pp. 280-285, Aug. 1985.
- [6] NIST, "Announcing the Advanced Encryption Standard(AES),"FIPS PUB ZZZ, 2001, (<http://www.nist.gov/aes>).