국가와 국민을 위한
데이터 생태계 중심기관 KISTI

# Researches on implementing workflows for applied research based on big-data based(HEP) API networks through the Research Platform

KISTI/KREONET
Jeonghoon Moon
23rd Jan 2024

KiSTi 한국과학기술정보연구원
Korea Institute of Science and Technology Information
www.kisti.re.kr

# Background KREONET

[1]     Introduction of KREONET

## Map of KREONET 2023

Korea Research Environment Open NETwork
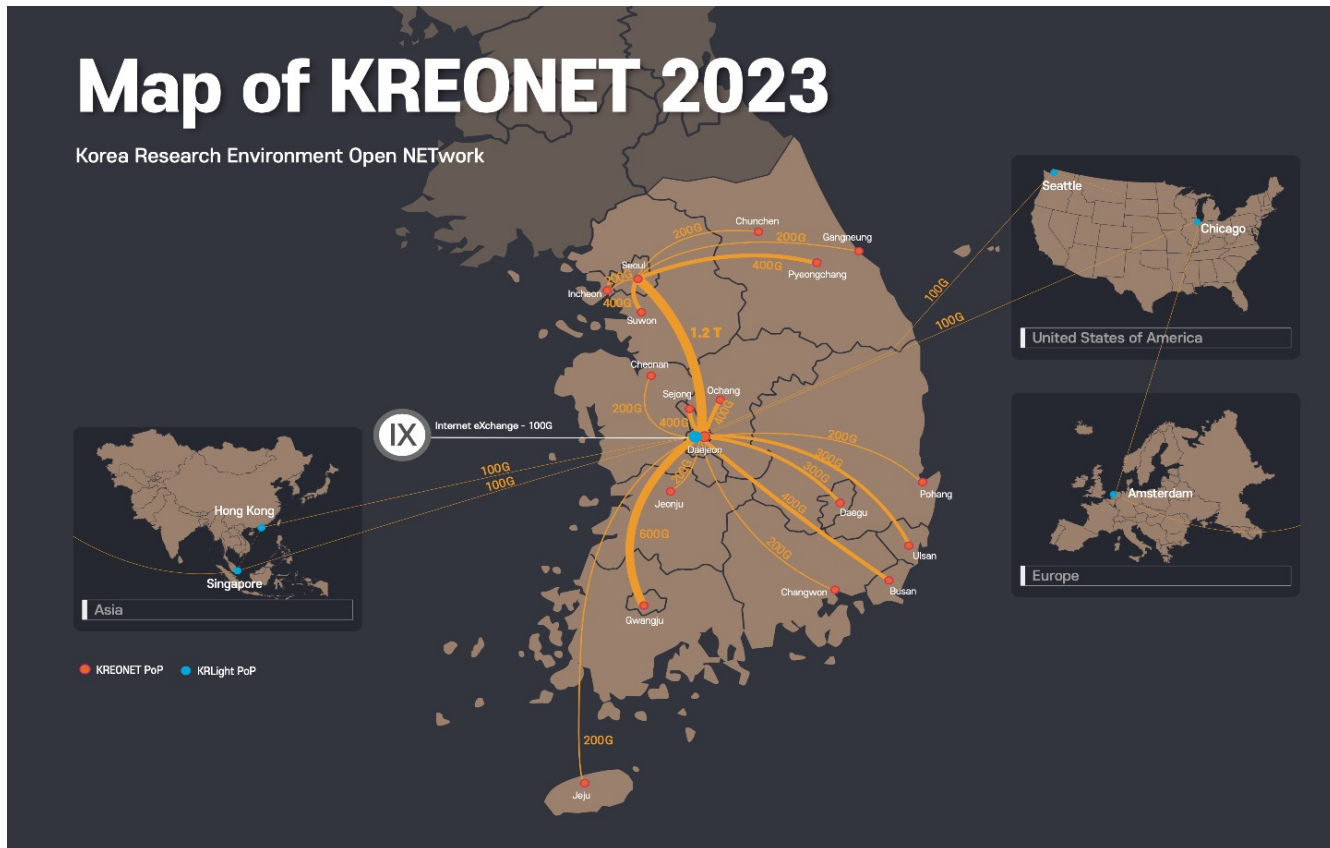
- **Max 1.2Tbps Backbone and 17 local gigapops for domestic**
  - Among Seoul, Daejeon, Kwangjoo, Busan, Daegu: 200Gbps
- **100Gbps 5 Gigapops for international**
  - Hongkong, Seattle, Chicargo, Amsterdam, Singapore
- KREONET-NOC, 24*365 Operating service
- Domestic IX (KT, SKB, Sejong Telecom) and International GIX (2G) connection service, COGENT(10G, Seattle)
- Connection for Commercial and public Cloud service
- **203 member of Institutes (over 0.5 million member)**

- **GLORIAD, Since 2005**

# KREONET 17 local GigaPops
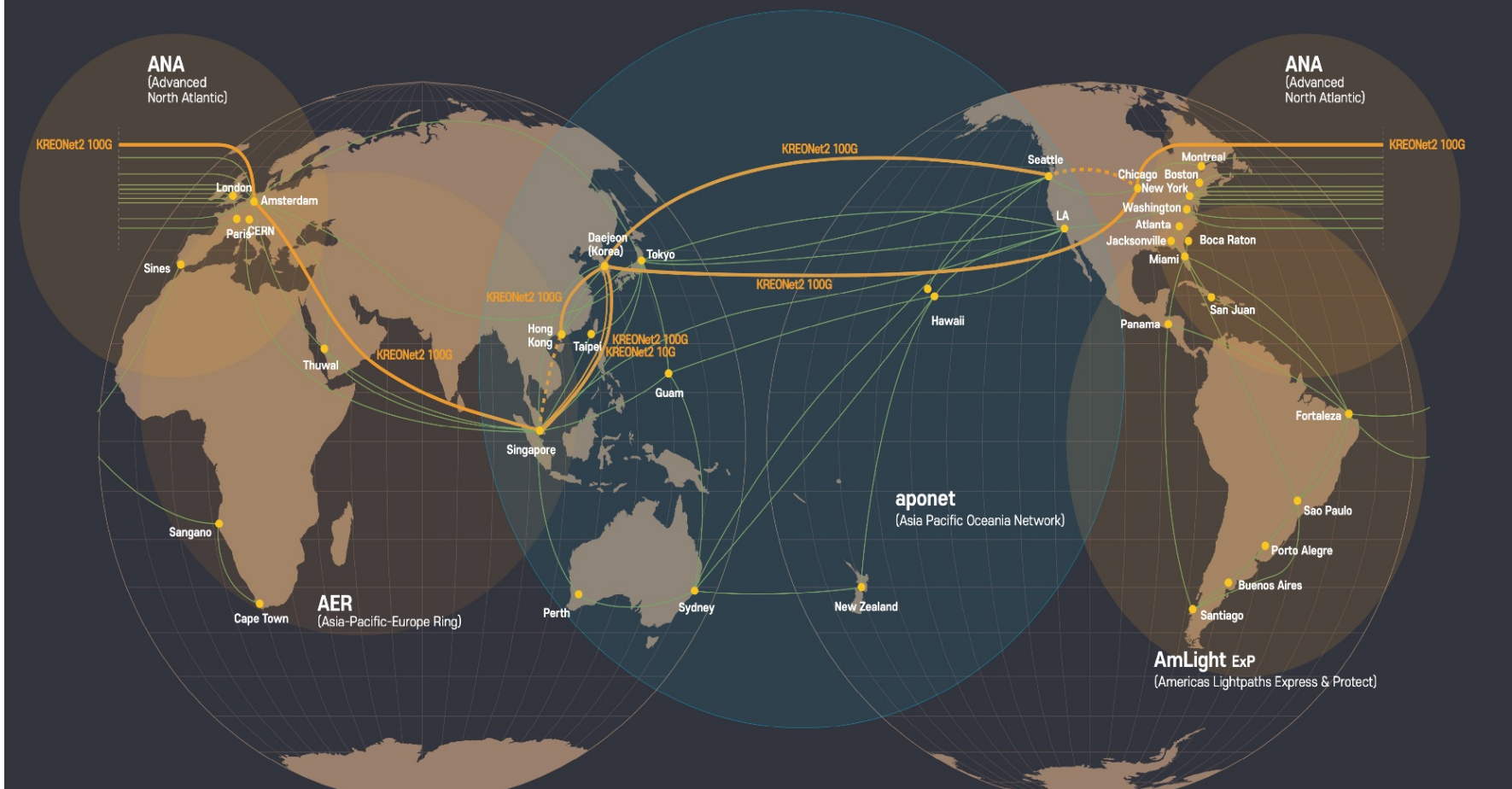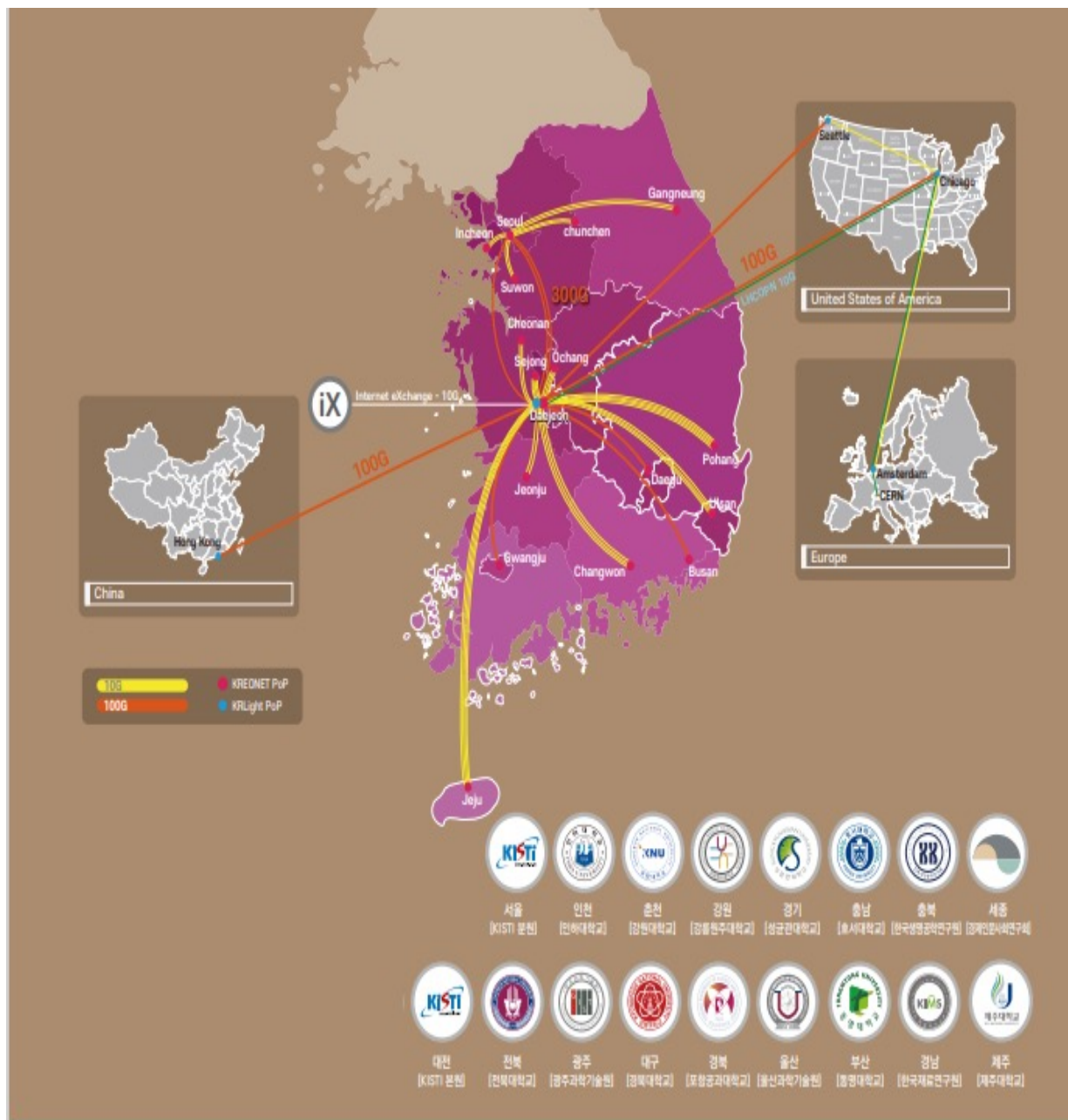


| No. | Local Center | Institute |
|-----|--------------|-----------|
| 1 | Daegu | KNU |
| 2 | Kyungki | SKKU(Suwon Campus) |
| 3 | Busan | TMU(동명대학교) |
| 4 | Sejong | NRC(경제인문사회연구회) |
| 5 | Incheon | Inha Univ. |
| 6 | Kyungbook | Postech |
| 7 | Kyungnam | KIMS |
| 8 | Kwangjuu | GIST |
| 9 | Kangwon | GWNU |
| 10 | Chungnam | Hoseo Univ. |
| 11 | Chuncheon | Kangwon Univ. |
| 12 | Ulsan | UNIST |
| 13 | Jeonbuk | JBNU |
| 14 | Jeju | Jeju Univ. |
| 15 | Chungbook | KRIBB(Ochang Branch) |
| 16 | Daejeon | KISTI Main |
| 17 | Seoul | KISTI Branch |

# KREONET Applications based on big data

| AREA | Institutes | Contents | Bandwidth |
|---|---|---|---|
| High Energy Physics | Seoul National University and 14 Institutes (서울대학교 격자게이지 이론연구단 외 14개 기관) | - 입자가속기 중심의 실험데이터를 국내연구진의 글로벌 전송/활용 | 1G-10G |
| Astronomy Space | Korea Astronomy and Space Science (천문연구원, 국토지리정보연구원(전파망원경센터)) | - 울산, 세종, 제주, 서울에 구축된 전파망원경 관측 데이터의 상관센터 중심의 데이터 공유 | 100G-40G |
| Astronomy Space | Korea Astronomy and Space Science 천문연구원(SDO) | - NASA 태양위성 관측 이미지를 글로벌 전송 및 국내 우주기상센터간 전송 | 10G |
| Astronomy Space | Korea Astronomy and Space Science 천문연구원(SKA, LSST) | - 칠레, 남아프리카공화국 에 위치한 광시야 망원경 관측 데이터의 공유 | 10G |
| Genome/Bio | Seoul National University Hospital and 8 Institutes 서울대병원 CMI센터외 8개 기관 | - ICGC 유전체 데이터 중심 RDC센터 구축 | 10G |
| Meteorology | The National Weather Service and 5 Institutes (기상청외 5개 기관) | - 기상청 슈퍼컴퓨터 및 기상데이터 공유 및 연구 활용 | 1G-40G |
| KSTAR | National Fusion Research Institutes NFRI외 3개 기관 | - KSTAR를 활용한 데이터 저장 및 연구활용 | 100G |
| Basic Science Facility (기초연구 관측장비) | Institute of Basic Science (기초연구원) | - 대형현미경으로 관찰된 고해상도이미지 및 데이터 전송 | 10G |
| Education | Pusan University and 5 Institute (부산대학교외 5개기관) | - 데이터, 계산자원, 저장자원의 연계를 통한 대규모 인력교육(원격) | 1G-10G |
| Building Construction | KOCED-CI(1,2차 실험센터) | - 지진실험 등 대규모 건설/건축 실험장비를 활용한 연구 및 교육 | 1G |
| Particle Accelerator | Pohang Accelerator (포항가속기센터(4세대)) | - 포항가속기센터에서 연간 실험/발생되는 데이터의 공유 | 10G |

K*GENOME

Genome RDC
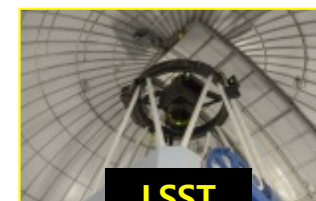
LSST

eKVN

HEP (CMS, ALICE)

KBSI TEM
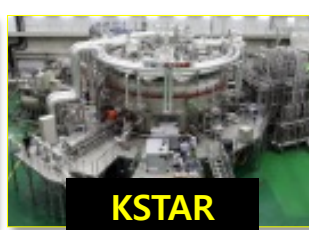
KSTAR

GSDC

Supercomputer(5th)

SDO

# Korea Research Platform

- Since 2015, Global partner of PRP project
- Since 2018, Leading of APAN APRP WG
- Since 2021, Expanding to 25 Korea National Research Institute

**National Scale**

**Establish a high-reliability & high-speed transfer system without boundaries between participants**

☞ For 25 Korea National Research Institutes & 4 University of Advanced Science & Technology
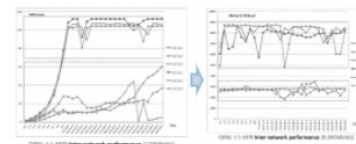


**Key elements**

- Borderless single domain transfer performance level
  ☞ Esnet ScienceDMZ level
- A Unified authentication system without boundaries

**Example of Nuclear Fusion Research**

Improve transfer performance about 75 times after ScienceDMZ deploy
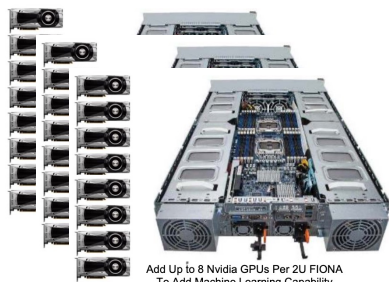☞ To expand the national scale

- Korea Research Platform expanding to 25 National Research Institutes
- HPC(Supercomputer, Cloud, Storage) over HPN and global federation

**HPC over HPN: A High bandwidth distributed HPC**

**Korea Research Platform (1G ~ 100Gb/s)**

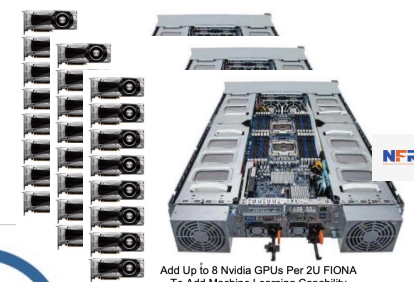**Faster workflow with big data**

**GPU-DTNs**

Add Up to 8 Nvidia GPUs Per 2U FIONA
To Add Machine Learning Capability

DATA center

Nurion

KISTI Supercomputer

**GPU-DTNs**

Add Up to 8 Nvidia GPUs Per 2U FIONA
To Add Machine Learning Capability

DATA center

- **Korea Research Platform**
  - Several prototype KRPs
    - GPU intensive, CPU intensive, Storage intensive type

- **Applications**
  - Analysis of Urban flooding by LSTM
  - Analysis of Solar Visual data by High intensive GPU(Astronomy)
  - Several distributed computing task based on AI & GPU



KRP Login



Web UI



Platform Architecture

# Big-data based API networks through the Research Platform
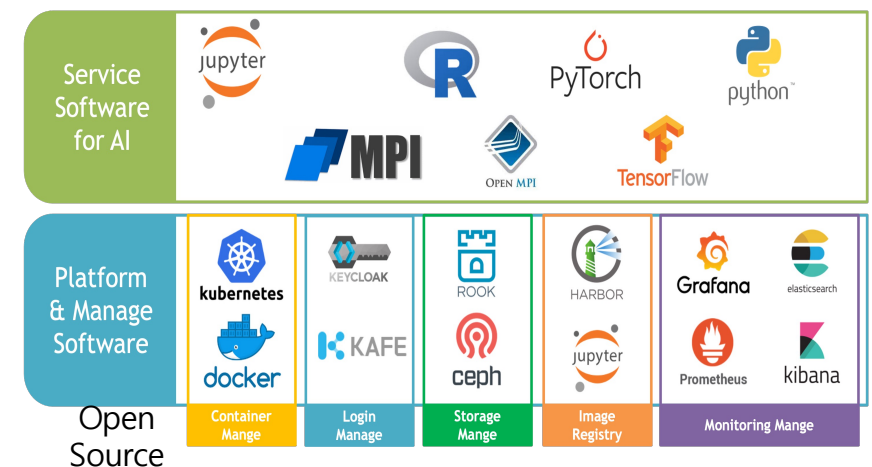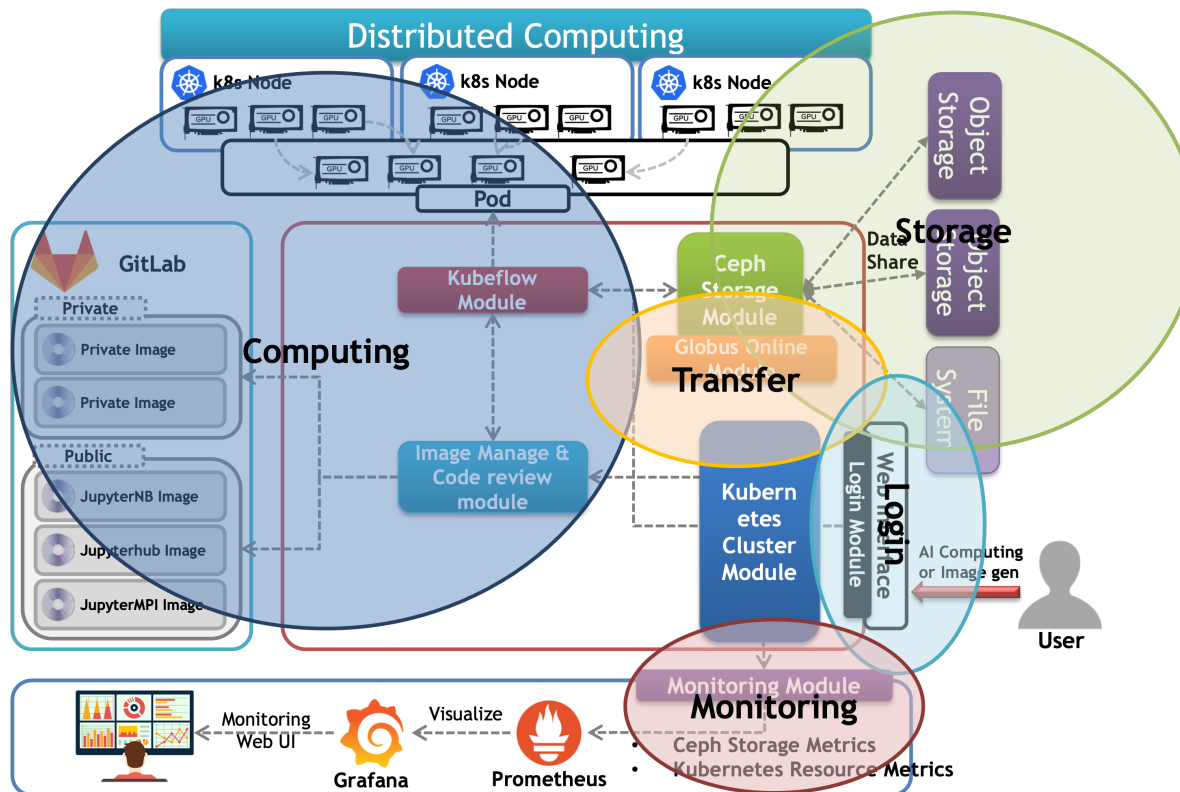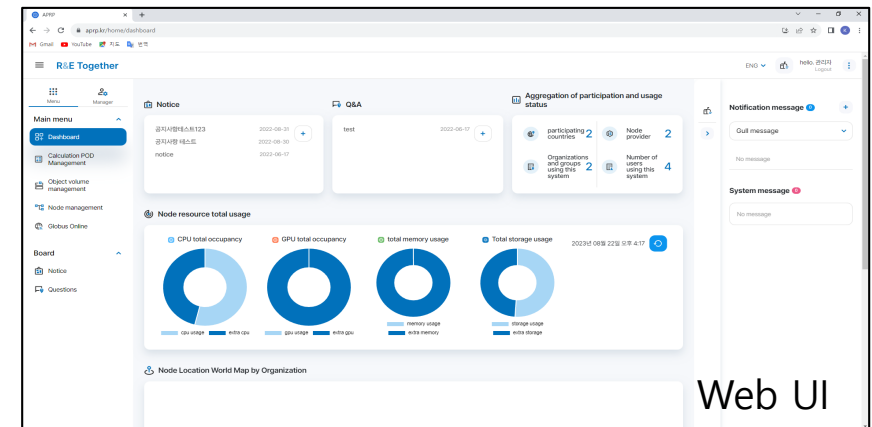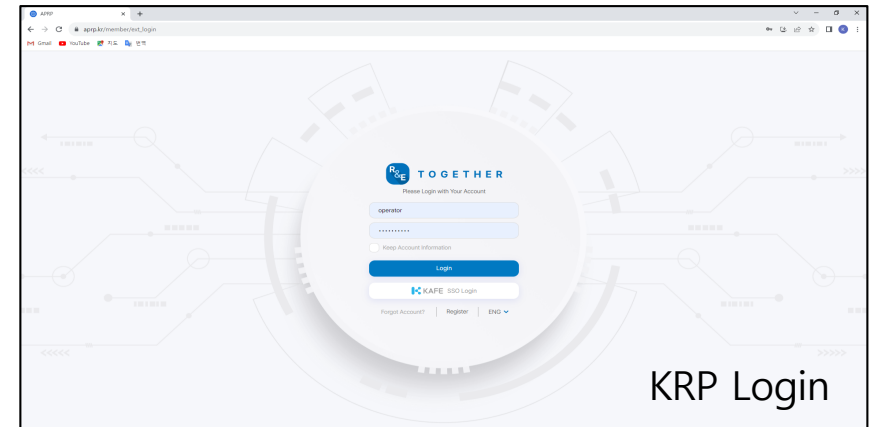
- Workflow based on research platform
  - Separate User & Manager
  - Combination of ML/DM & Ensemble relation

# Primitive API networks component by Kubernetes
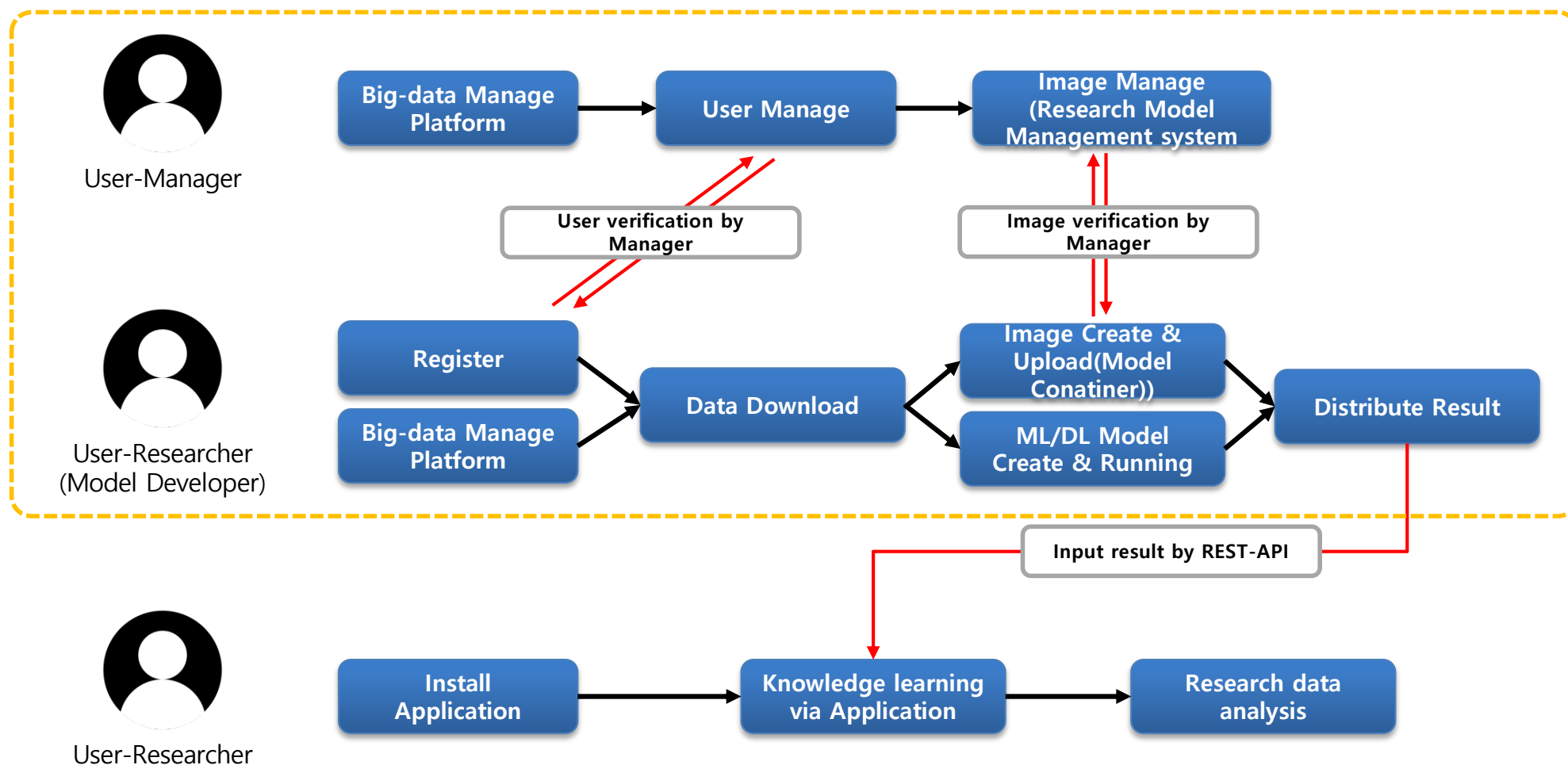
- Kubernetes and Artificial Intelligence for API networks through the Research Platform(Functions)
- Network communication function between each environment model composed of containers
- **Primitive API function and can be used to implement various services.**

| | |
|---|---|
| **POST** /uploads Uploads ⌄ | ① Input file upload of Model |
| **GET** /list List ⌄ | ② Output uploaded file list |
| **POST** /action Action ⌄ | ③ Computation of Model using Input file |
| **POST** /action_json Action ⌄ | ④ Using JSON data and running model |
| **GET** /action/{action_id} Action ⌄ | ⑤ Individual running Status check |
| **GET** /downloads/{action_id} Downloads ⌄ | ⑥ Individual running Output download |
| **POST** /ref_action Ref Action ⌄ | ⑦ Individual running Output download |

* For more detail of API function: See the appendix

# Ensemble of models A and B

- ## Worker1 - Input the results from Model A into Worker2 - Model B to execute calculations



worker 1

worke1 Address

Model A Job ID

worker 2

Model B Job ID

Result of Model B

▶ Execute calculations using the results of Model A as input to Model B

▶ Check the calculation results of Model B

# Computational automation through the API Networks

| _id | action_id | model_code | input_json | input_file | output_file | status | starting_date | finished_date | worker |
|---|---|---|---|---|---|---|---|---|---|
| 655223d3c80ce395c6e91fcc | b2305b97-a8c9-403a-81a1-44acb6ac33 **Model A** | finedust-001 | ⊖{<br>  "temperature": 0.902,<br>  "rainfall": 0.665,<br>  "wind_speed": 0.838,<br>  "wind_direction": 0.878<br>} | | output_b2305b97-a8c9-403a-81a1-44acb6ac33ef | finish | 2023-11-01 05:00:00 | 2023-11-01 05:00:01 | 10.10.10.10:8080 **Model A** |
| 655223d7c80ce395c6e91fcd | 5695974b-9fcb-4355-afee-2b9f966c471 **Model B** | disease-001 | | output_b2305b97-a8c9-403a-81a1-44acb6ac33ef | output_b2305b97-a8c9-403a-81a1-44acb6ac33ef-output | finish | 2023-11-01 05:30:00 | 2023-11-01 05:30:02 | 10.10.10.20:8080 **Model B** |
| 655223d8c80ce395c6e91fce | dde98ab1-c518-4823-9994-411d981ebd50 | finedust-001 | ⊖{<br>  "temperature": 0.782,<br>  "rainfall": 0.525,<br>  "wind_speed": 0.747,<br>  "wind_direction": 0.864<br>} | | output_dde98ab1-c518-4823-9994-411d981ebd50 | finish | 2023-11-02 05:00:00 | 2023-11-02 05:00:01 | 10.10.10.10:8080 |
| 655223e6c80ce395c6e91fd0 | cc14ff36-a43b-4c2a-90c3-3abb47370677 | disease-001 | | output_dde98ab1-c518-4823-9994-411d981ebd50 | output_dde98ab1-c518-4823-9994-411d981ebd50-output | finish | 2023-11-02 05:30:00 | 2023-11-02 05:30:02 | 10.10.10.20:8080 |
| 65522434c80ce395c6e91fd1 | 87017e34-bab5-4c70-8ffa-23c873203fd7 | finedust-001 | ⊖{<br>  "temperature": 0.877,<br>  "rainfall": 0.392,<br>  "wind_speed": 0.715,<br>  "wind_direction": 0.867<br>} | | output_87017e34-bab5-4c70-8ffa-23c873203fd7 | finish | 2023-11-03 05:00:00 | 2023-11-03 05:00:01 | 10.10.10.10:8080 |

1  {"temperature": 0.931, "rainfall": 0.558, "wind_speed": 0.75, "wind_direction": 0.936, "finedust": 3.265} ← Result

| _id | action_id | model_code | input_json | input_file | output_file | status | starting_date | finished_date | worker |
|---|---|---|---|---|---|---|---|---|---|
| 65522437c80ce395c6e91fd2 | 57d87993-10c5-4928-94b1-d94f926c292f | disease-001 | | output_87017e34-bab5-4c70-8ffa-23c873203fd7 | output_87017e34-bab5-4c70-8ffa-23c873203fd7-output | finish | 2023-11-03 05:30:00 | 2023-11-03 05:30:02 | 10.10.10.20:8080 |

1  {"disease": 2.282} ← Result

## ■ computational automation: Model A

```python
7    IP = "192.168.6.52"
8    PORT = "8080"
9    HOST = f"http://{IP}:{PORT}"
10   INPUT_JSON = {
11       "temperature": round(uniform(0.75, 0.95), 3),
12       "rainfall": round(uniform(0.35, 0.75), 3),
13       "wind_speed": round(uniform(0.65, 0.85), 3),
14       "wind_direction": round(uniform(0.85, 0.95), 3)
15   }
16
17   def main():
18       url = HOST + "/action_json"
19
20       response = post(url=url, json=INPUT_JSON)
21       if response.status_code == 200:
22           return True
23       else:
24           return False
```

← Server Address of Model A

← Setting of input parameter(Model A)

← Transfer data to API of Model A

Every morning 05:00 - Perform prediction model calculations

16

## computational automation: Model A

```
43    def main():
44        mongo = MONGODB_CONTROL()
45        json_data = mongo.get_spectific_model("finedust-001")
46        if json_data:
47            ref_action_id = json_data[0]['action_id']
48            ref_ip = json_data[0]['worker']
49        else:
50            return False
51
52
53        url = HOST + "/ref_action"
54        json_data = {
55            "ref_ip": ref_ip,
56            "ref_action_id": ref_action_id
57        }
58
59        response = post(url=url, json=json_data)
60        if response.status_code == 200:
61            return True
62        else:
63            return False
```

Check the latest calculation results of Model A

Transfer data to API of Model B

Every morning 05:00 - Perform exponential model calculations using the calculation results of the prediction model

# Related Project

[1]    Smart Agriculture
[2]    Smart decision-making system
[3]    Smart Hospital development project

## 1. Korea Rural Development Administration project (Grant March 2022)

- Title: Development of integrated linkage system for agricultural big data and utilization model

- Participants: Seoul National Univ., KISTI

- Total Budget: About 450,000$(3years)

- KISTI: 150,000$(For 3years)

- Research Contents (KISTI/KREONET part)
    - Agriculture big data transfer
    - Running Crop model on the Research Platform (Using CPU & GPU)
    - Establishment Research Platform for end-user(Farmer) & developer(Agriculture Researcher)

CALS

서울대학교 농업생명과학대학
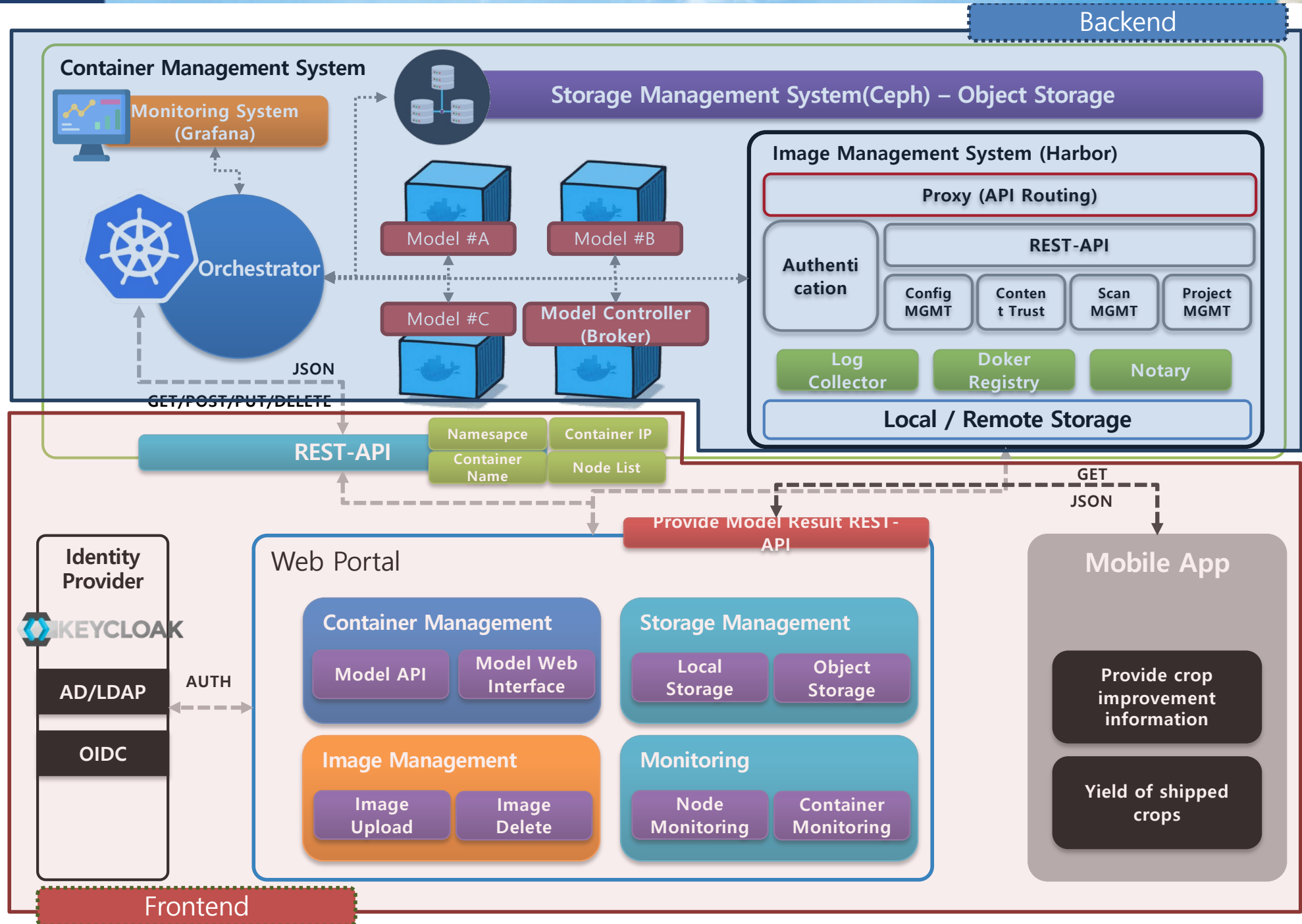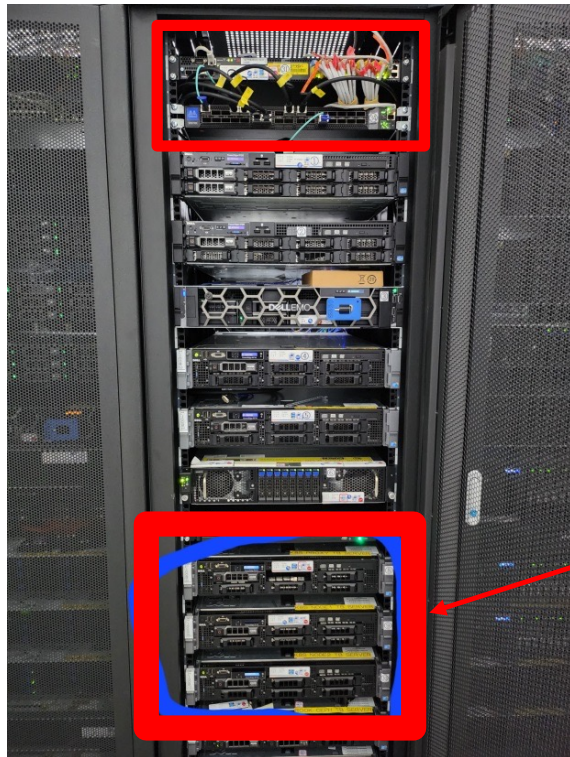College of Agriculture and Life Sciences

KiSTi
www.kisti.re.kr

Korea Institute of
Science and Technology Information

농촌진흥청
Rural Development Administration

- 3 DELL R710 server, Network switch
  (CPU 12core, Memory 72GB, 1G NIC)
- 1 Gygabyte storage server
  (CPU 104core, Memory 256GB, 1G NIC, DISK 6.4TB)
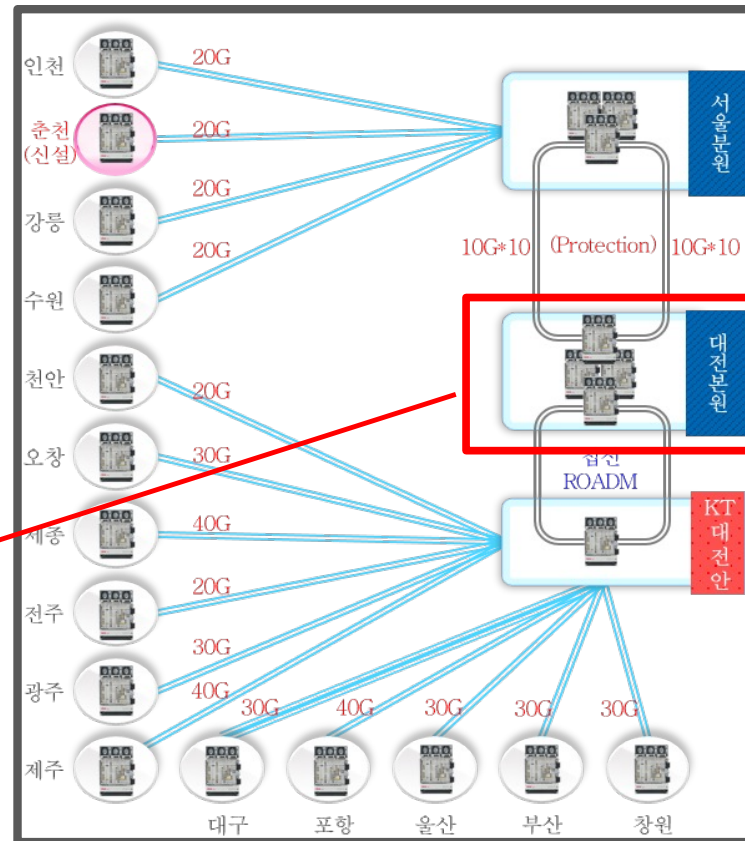
**Infrastructure pilot construction**
Seoul-Junju through high-speed connection available by KREONET

## 2. Korea Environment Institute Project grant Dec 2022

- Title : Establish an integrated impact assessment platform and build an ensemble of multi-climate scenarios based on the new climate regime (AR6)

- Participants: KEI, KISTI, Several Commercial Company

- KISTI budget : About 1M$(for 3-4years)

- Research contents (KISTI/KREONET part)
    - Building a Container based Research Platform & Ceph storage
    - Building a Distributed Research Platform based on Multi-Institutional HPC(Weather, Agriculture, Environment Institutes and statistical data)
    - Operation of various models on the Research Platform

## 3. Smart Hospital development project (Grant Sept 2022)

- Title : The Construction of patient-tailored, hyper-personalized digital medical ecosystem based on Omni-verse platform

- Participants: 20 Institutes, Hospital, and Company (Koryo Anam General Hospital, LG Electronics, Kakao, KISTI, KHNP and etc)

- Total budget: About 15M$(8 years)
- KISTI: 1.8M$ (8 years)

- Research contents (KISTI/KREONET part)
  - Big data super-highway for inter-hospital and Human & Genome data centers
  - High speed wireless communication inter/intra smart hospital
  - IT Technologies for Smart Hospital

# KREONET Giant Group Researches

[1]    Astronomy, HEP, Weather & Climate, Genome & Bio, KSTAR, TEM

## Institute/University

- SNU, Korea Univ., Sejong Univ., GSDC/KISTI, KNU, UOS, Busan Univ., Youngkwang Nuclear Generating Station, Hanyang Univ., Soosil Univ.
- CMS, ALICE

## Network Resource /Technology

- LHCOPN(10G/GSDC -to- CERN)
- LHCONE (VRF)
- Youngkwang Nuclear Generating Station Detector (10G)
- GRID Computing (KNU/10G)

## Activities

- Tier 0-Tier1-Tier2 Data Transfer
- KEK(JP)
- CMS, ALICE. Collaboration
- STAR-PROJECT (Busan Univ.)

고에너지물리 첨단연구기관

세종대학교 물리학과 김세용

고려대학교 물리학과 최수용

서울대 격자게이지이론연구단 이원종

서울대 한국중성미자연구센터 김수봉

서울시립대 슈퍼컴퓨터센터 박인규

영광원자력발전소 김수봉

고려대학교 물리학과 원은일

서울대학교 물리학과 양운기

한양대학교 물리학과 김태정

경북대 고에너지물리연구소 손동철

부산대 물리학과 유인권

Tier Center(GSDC, KNU)

Accelerator (KEK, CERN)

Youngkwang Nuclear Generating Station Detector(RENO)

## Institute/University

- **KASI, Ulsan Univ., Yeonsei Univ., KIAS, Chungbuk Univ.**
- **SDO, SDSS, KVN, KMTNET**

## Network Resource /Technology

- **Between Radio Telescope : 10-40G Carrier Ethernet**
- **Data Center(Daejeon): 100G**
- **SDO Data Center: 10G**
- **KMTNET 10G**

## Acitivies

- **eKVN Data Transfer (3 Radio Telescope, Yeonsei Univ. Tamra Univ. Ulsan Univ.)**
- **Observation Data for international : Chile, Australia, Optical Telescope remote operation/ Observation data transfer**

천체/우주 첨단연구기관

한국천문연구원 전파천문본부 김현구

연세대학교 전파망원경센터 김현구

울산대학교 전파망원경센터 김현구

탐라 전파망원경센터 김현구

한국천문연구원 광학천문본부 이충욱

한국고등과학원 물리학부 박창범

국토지리정보원 국토측량과 안기덕

한국천문연구원 SDO 이재진

충북대학교 전체우주센터 한오성

연세대학교 천문우주학과 이석영

**KVN-Korea VLBI Network(Radio Telescope/Data Center)**

**KMTNET(LSST)**

**Solar Dynamic Observatory(SDO)**

## Institute/University

- KMA, KMA Supercomputing Center(Ochang), Busan Univ., PKNU, KBS Disaster Broadcating Center, KAERI, APEC Climate Center
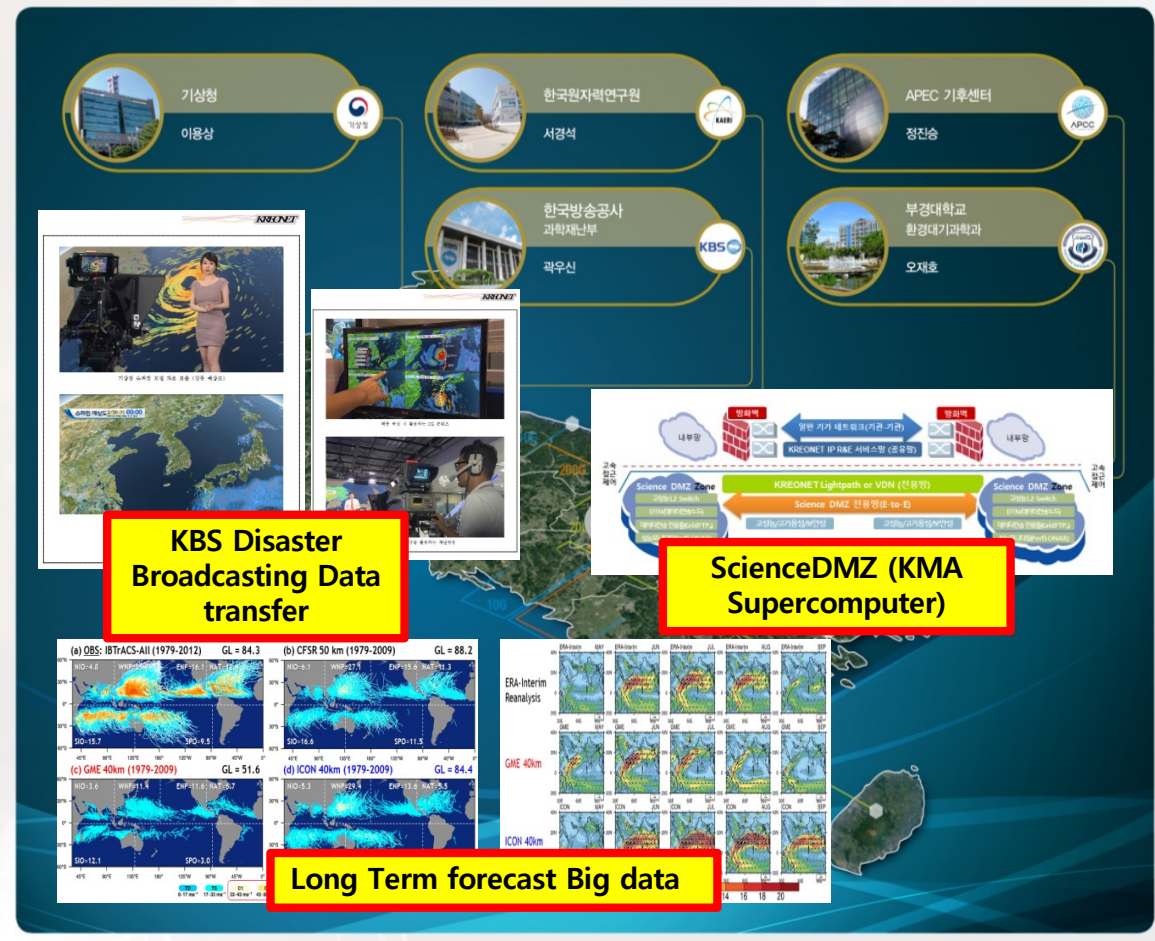
## Network Resource /Technology

- SCIENCEDMZ
 : KMA Supercomputing Center (10G)
  - APCC, Busan Univ., PKNU
- KMA(HQ)-Weather Group of R.O.K Air Force(1G)
- KBS Disaster Broadcasting(1G)

## Activities

- SCIENCEDMZ
 : KMA Supercomputing Center(10G)
  - APCC, Busan Univ., PKNU
- KMA(HQ)-Weather Group of R.O.K Air Force(1G)



기상/기후 첨단연구기관

KBS Disaster Broadcasting Data transfer

ScienceDMZ (KMA Supercomputer)

Long Term forecast Big data

## Institute/University

- SNU Hospital, Kangbook SNU Hospital, Iwa Univ. Hospital, Seoul Health College, SNU, SNU-CMI, NCC, KOBIC, CODA, Hanyang Univl, UNIST, Soongsil Univ
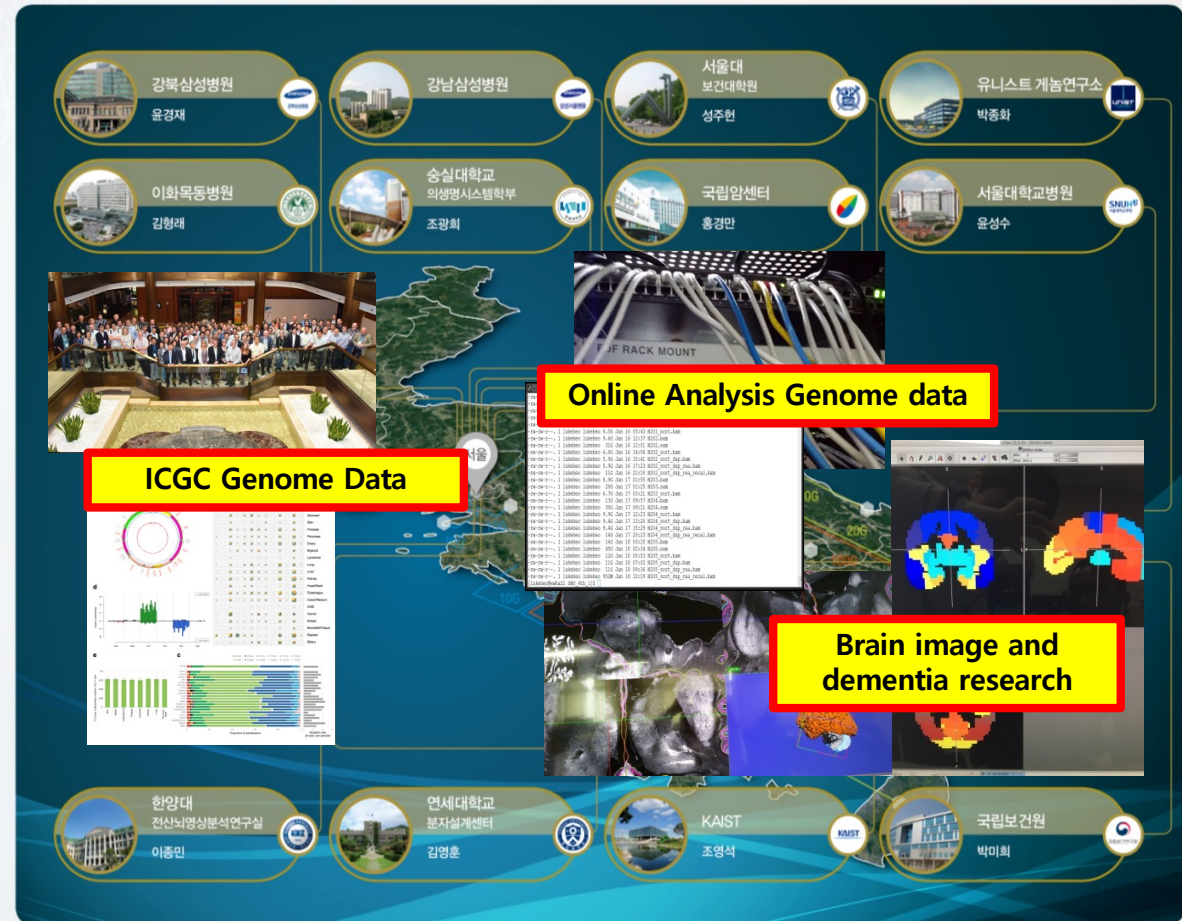
## Network Resource /Technology

- SCienceDMZ (SNU CMI/10G)
- Quantum encryption communication (CODA-KOBIC)
- ICGC Data Storage share network (6 Institutes 1G/VLAN)

## Activities

- WGS/ICGC Genome data transfer
- Patients Privacy Protect and share environment
- TB Genome data support for RDC
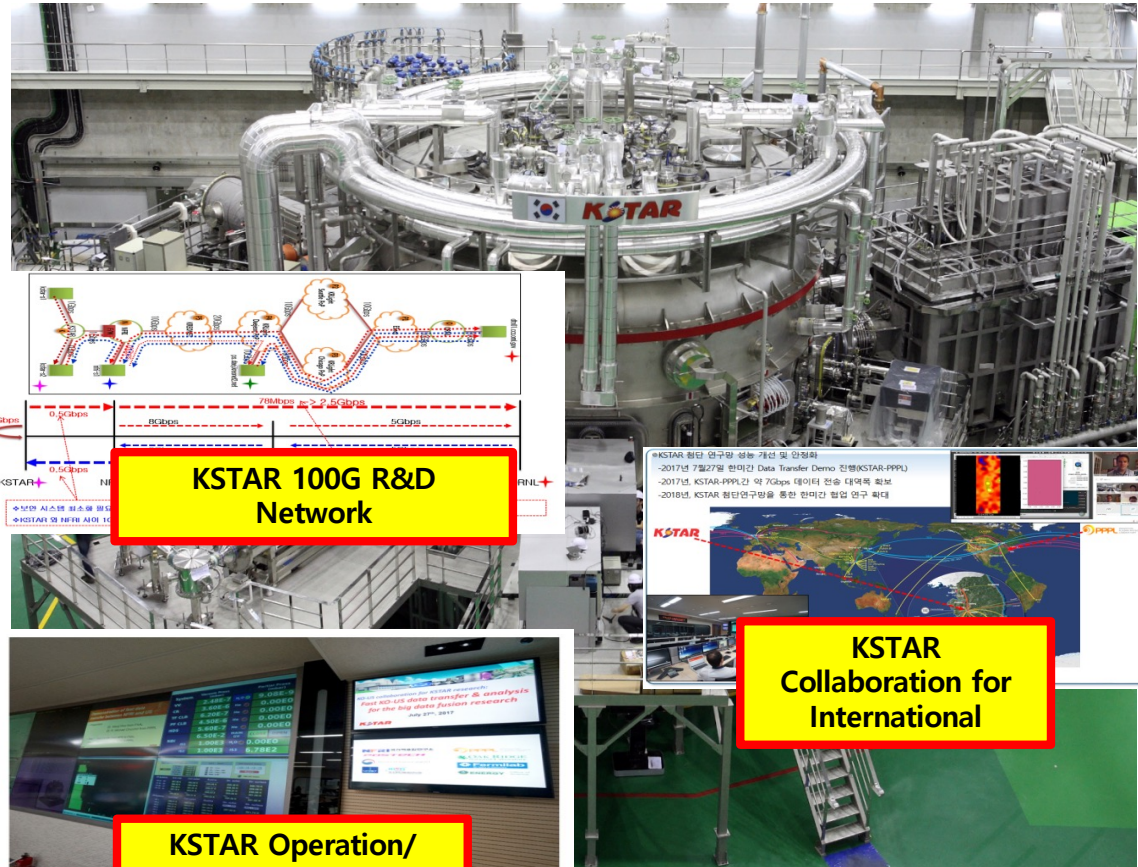
**Institute/University**

KTAR/NFRI(Nation Fusion Research Institute)

**Network Resource /Technology**

KTAR (100G)
100G Real time performance monitoring System(PerfSONAR-100G)
100G DIRECT PATH

**Activities**

KSTAR  Big data collaborative research (Kr-US
Real time remote data analysis



KSTAR 100G R&D Network

KSTAR Collaboration for International

KSTAR Operation/
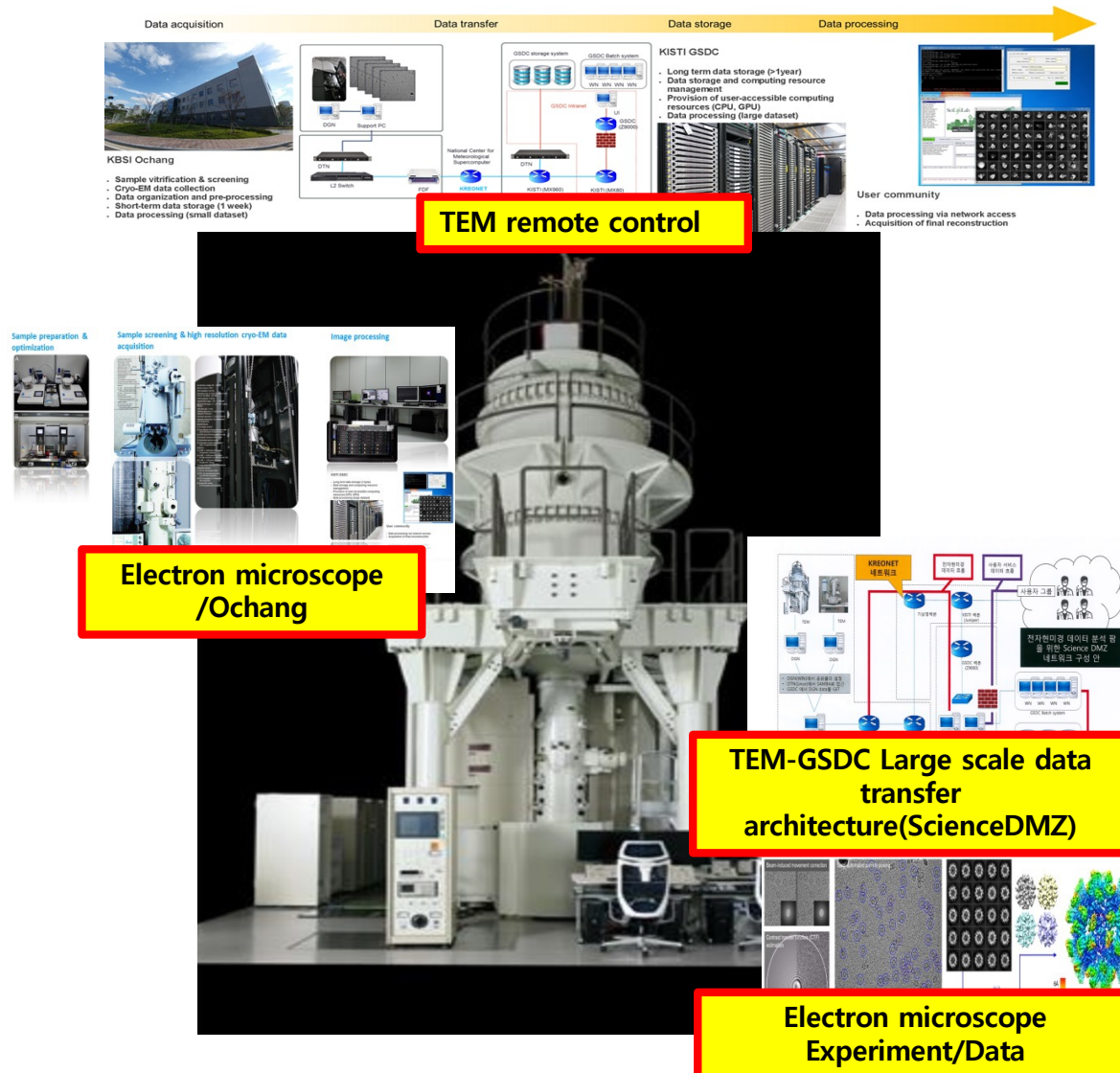
## Institute/University

- **KBSI/TEM , KISTI/GSDC**

## Network Resource /Technology

- **KBSI/TEM – GSDC 10G Light Path (LP over WDM)**
- **ScienceDMZ 10G**

## Activities

- **Experiment/Observation data**
- **Remote access/remote observation/support**



**TEM remote control**

**Electron microscope /Ochang**

**TEM-GSDC Large scale data transfer architecture(ScienceDMZ)**

**Electron microscope Experiment/Data**

# Conclusion

- **HPC over HPN**
  - Expansion and joint establishment of applied research fields through strengthening collaboration with government-funded research institutes
  - Discussion on ways to jointly utilize insufficient GPU computing resources in the field of artificial intelligence
  - Strengthening international collaboration research by expanding international big data transmission

- **Extension for 3rd party research areas**
  - Transmission and analysis of experimental data in particle physics(HEP)
  - Smart Agriculture/Environmental research on climate change
  - Cloud computing & Wireless communication

THANK YOU