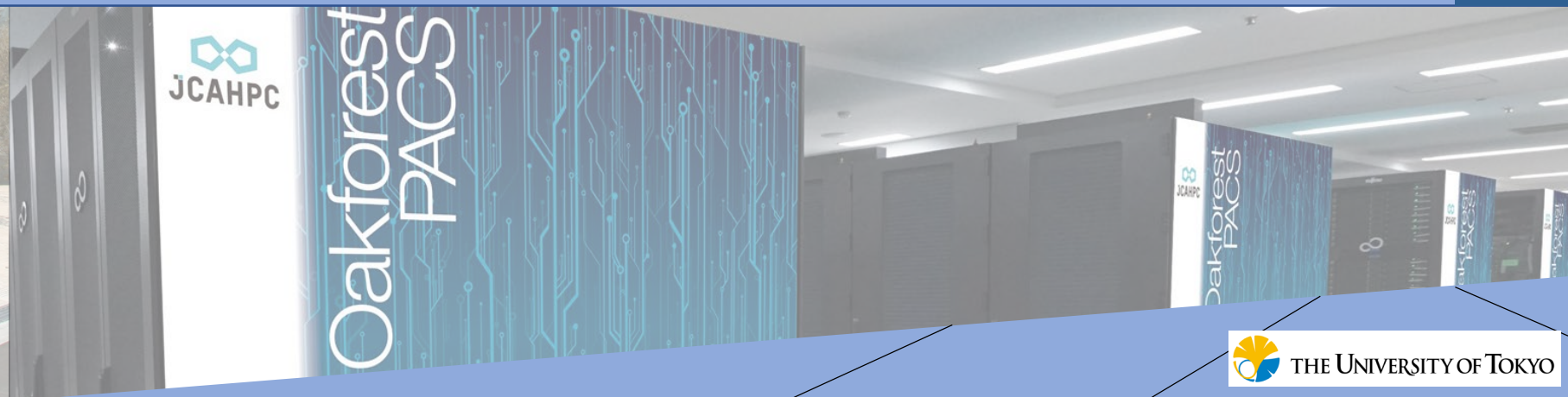
INFORMATION TECHNOLOGY CENTER
THE UNIVERSITY OF TOKYO

JCAHPC

mdx:
Infrastructure for leveraging data



THE UNIVERSITY OF TOKYO

mdx: Data Platform project

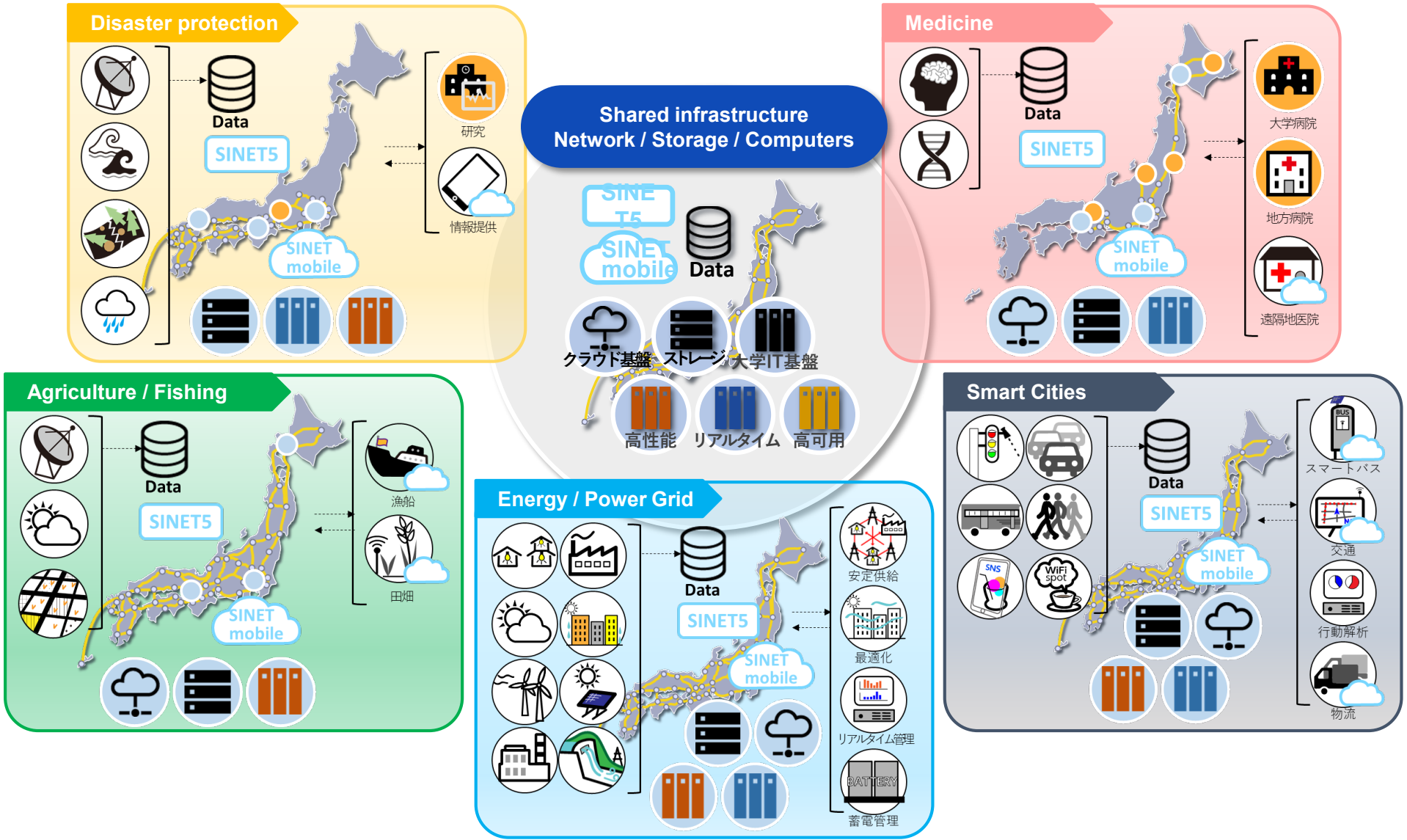
- **Target is to leverage data utilization at all over Japan making full use of high performance R&E network "SINET"**
  - SINET is an R&E network of Japan operated by NII (National Institute of Informatics)
- **Project supported by Japanese Government**
- **Currently jointly being designed by:**
  - 8 National Universities (Tokyo, Hokkaido, Tohoku, Tokyo Tech, Nagoya, Kyoto, Osaka, Kyushu)
  - NII (National Institute of Informatics)
  - AIST (National Institute of Advanced Industrial Science and Technology)
- **Will invite universities and public research institutes of all over Japan to use the platform for industry-academia and local government-academia collaboration activities.**
- **Operation starts from Mar. 2021.**

mdx: Data Platform project (2)

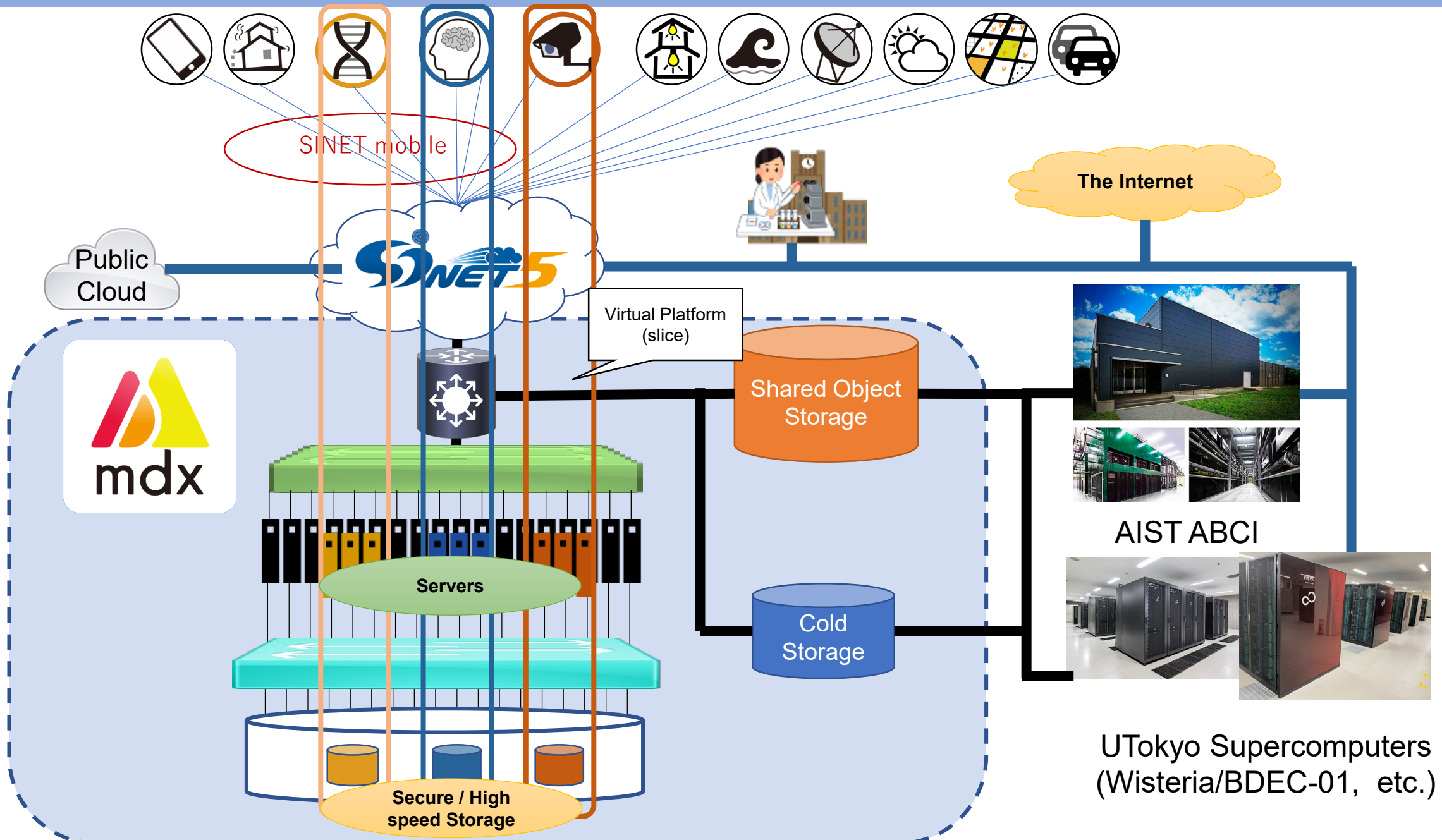
- **a rapid PoC environment for R&D data utilization activities including industry-academia collaboration projects.**
  - Shared platform for various data utilization activities
  - Combine SINET and high performance computing and storage infrastructure
- **Users can use wide bandwidth low latency “slices”**
  - Wide-area virtual infrastructure isolated from “the internet”
  - Connect edge devices with high performance computing and storage infrastructure and supports real-time data processing
- **Will host:**
  - Various data exploiting activities, especially in SMEs, local governments and agriculture / fishing
  - Key to solve the regional disparity problems

On-demand platform

- A **real-time** data processing environment.
- **Geographically distributed IaaS** including network



Infrastructure of the Data Platform

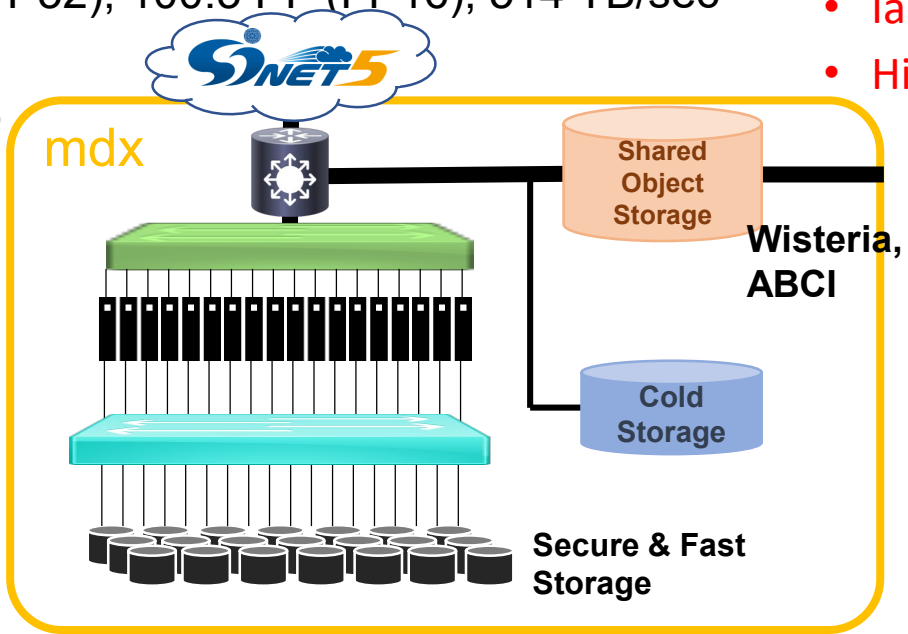


Infrastructure of the Data Platform

- **The infrastructure of the Data Platform is more like a cloud (IaaS) spreading over wide area**
  - Network connecting data and IoT devices can be provisioned with compute and storage resources
- **The platform provides virtual infrastructure (slices) to users**
  - Users can use the provided infrastructure (slice) as if it is a dedicated infrastructure for the user.
  - A slice is isolated from the internet or other slices.
    - User can provision gateway(s) to outside on a slice

Overview of infrastructure

- **Facility**
  - < 2.0 MW including Cooling, <170 m^2
  - Same location with Wisteria/BDEC-01, same campus with ABCI
- **Compute nodes**
  - the general purpose nodes:
    - 368 nodes, Intel Xeon Platinum 8368 (38c, 2.4GHz, IceLake-SP) x2 CPU sockets/node
    - 2.14 PF (DP), 150.7 TB/sec
  - the computation accelerator nodes:
    - 40 nodes, Intel Xeon Platinum 8368 (IceLake-SP) x2 socket + NVIDIA A100 x8 GPUs/node (19.5 TF, 1.555 TB/s)
    - 6.47 PF (FP64), 50.38 PF (FP32), 100.3 PF (FP16), 514 TB/sec
- **Storage**
  - Fast Storage with NVMe SSD
    - 1 PB, 250 GB/sec
  - Internal Storage
    - 16 PB. 157 GB/sec
  - Shared Object Storage (S3 compatible)
    - 10 PB, 63 GB/sec
  - Cold Storage
    - Optical disc drive

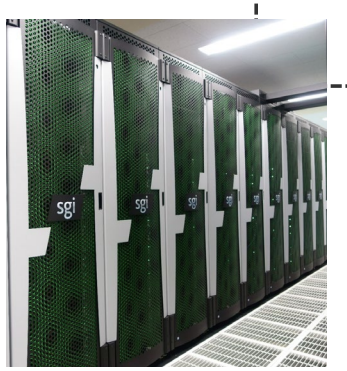


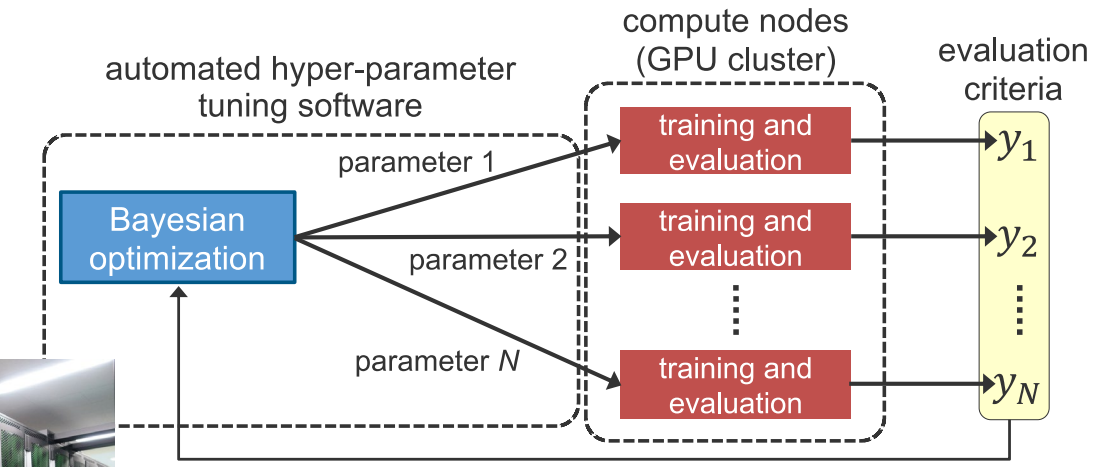
- **Network**
  - 25G Ethernet for frontend
    - 100G to SINET
    - 400G to BDEC
  - 100G Ethernet with RoCEv2 for RDMA and Storage as backend
  - Overlay with EVPN-VXLAN
- **Software, etc.**
  - VM & Container
    - VMware vSphere
  - IaaS like management
  - High security, high availability

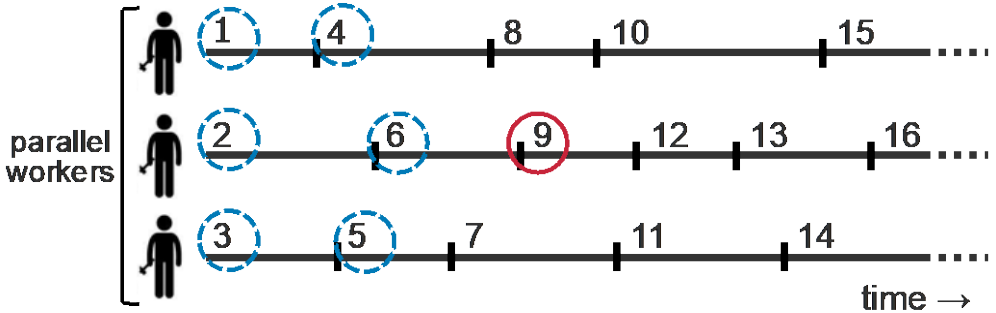
New building in Kashiwa II Campus

Prototype for mdx: Medical image recognition by UTokyo hospital

- Hyper-parameter auto-tuning platform on Reedbush

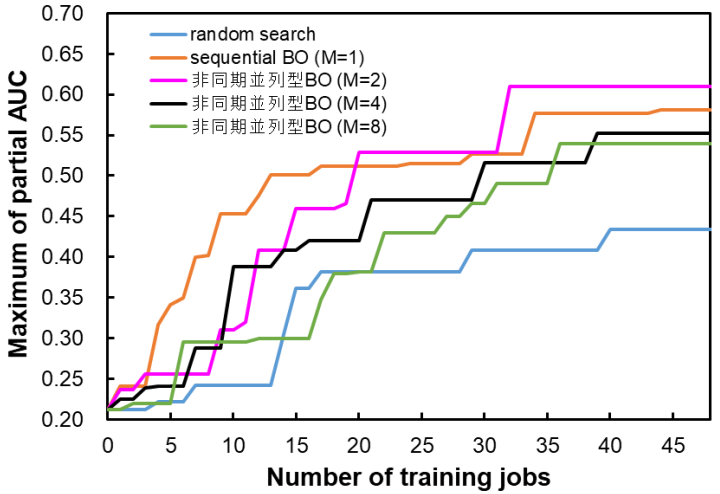




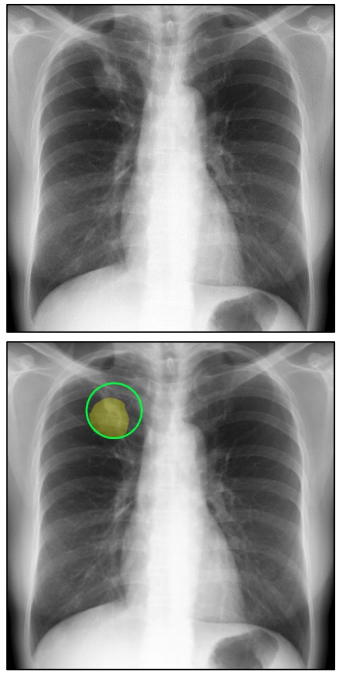


Asynchronous parallel BO

Lung mass detection in chest radiographs



Changes in partial AUC values of validation data, where each value is the maximum value in past evaluations

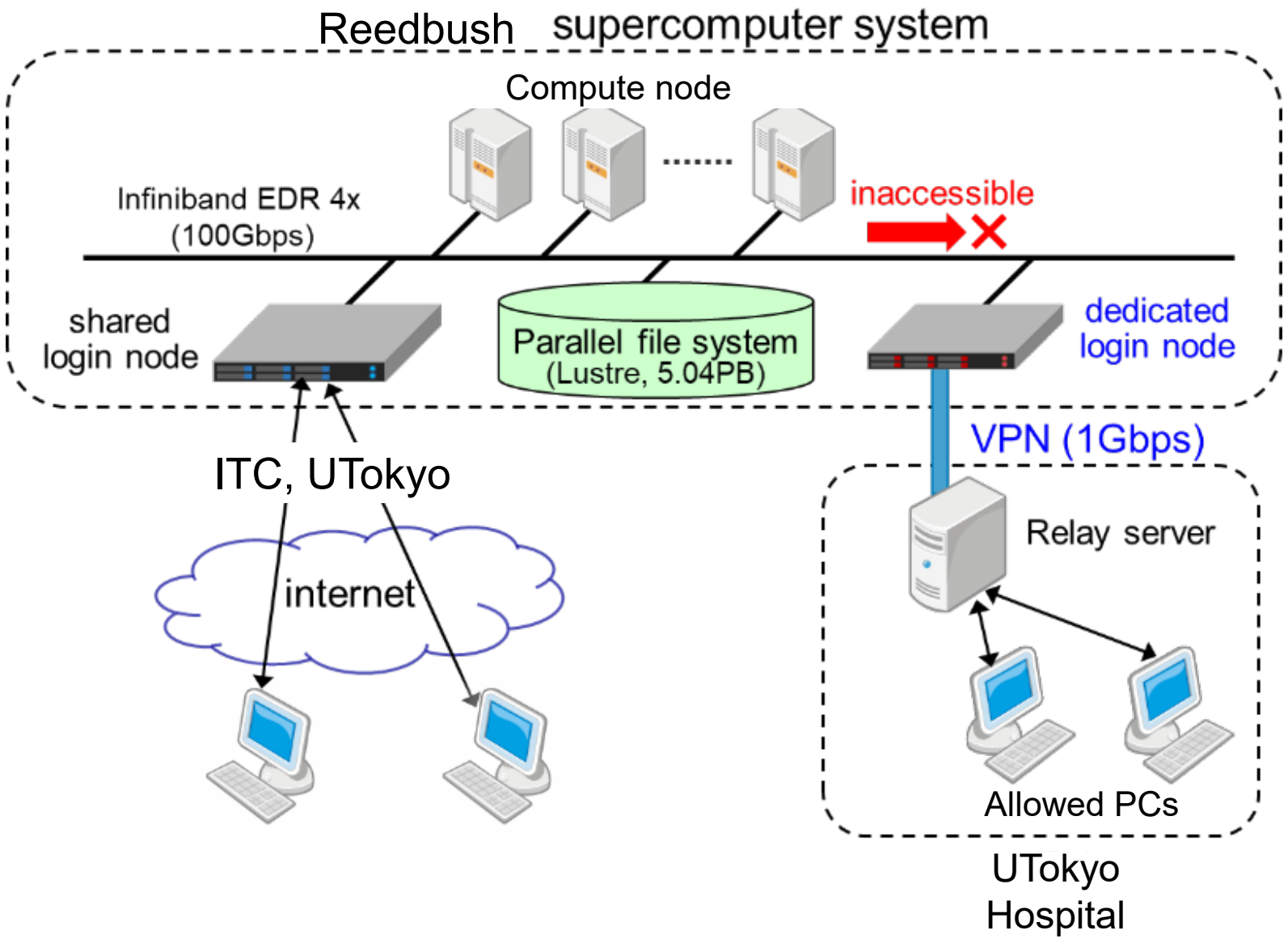


Upper: original image, Lower: detection result (Green circle and yellow filled region: lesion area)

Nomura Y, J Supercomput. 20 Jan 2020 (Epub ahead)

Coordination between UTokyo Hospital and Reedbush System

- **Security is crucial !!**
  - Anonymized personal data is transferred to RB
  - Dedicated login node and VPN are introduced for isolation with other projects
  - Only least amount of data required for calculation is placed on RB



Summary

- **The Data Platform (mdx) is “more for every day applications than big sciences”**
  - Data utilization for everyone: SMEs, local governments, agricultures, fishing etc.
  - Provide PoC environment for commercial applications
- **A real-time data processing environment.**
  - It is a geographically distributed IaaS, directly connectable to edge devices.
- **Infrastructure of mdx**
  - Virtual infrastructure (slices)
  - 368 CPU nodes: 2.1 PF (DP)
  - 40 GPU nodes: 8 GPUs/node, 6.4 PF (FP64), 6.7 PF (FP32), 100 PF (FP16)
  - 1 PB High-speed Storage with NVMe SSD, 16 PB Internal Storage, 10 PB Shared Object Storage
- **The platform can work as a “streaming data gathering infrastructure” for super computers such as ABCI or Wisteria/BDEC-01**
  - Leveraging the SINET mobile infrastructure

Thank you for watching