# Research towards Next-Gen Supercomputing Systems

Fujitsu Next Generation Computing Infrastructure Collaborative Research Cluster

In October 2022, Tokyo Institute of Technology and Fujitsu have established the "Fujitsu Next Generation Computing Infrastructure Collaborative Research Cluster" at Tokyo Tech's Suzukakedai Campus. The purpose is to realize a next-generation computing infrastructure capable of extremely large-scale data processing and simulations based on AI and high-performance computing (HPC) technologies.
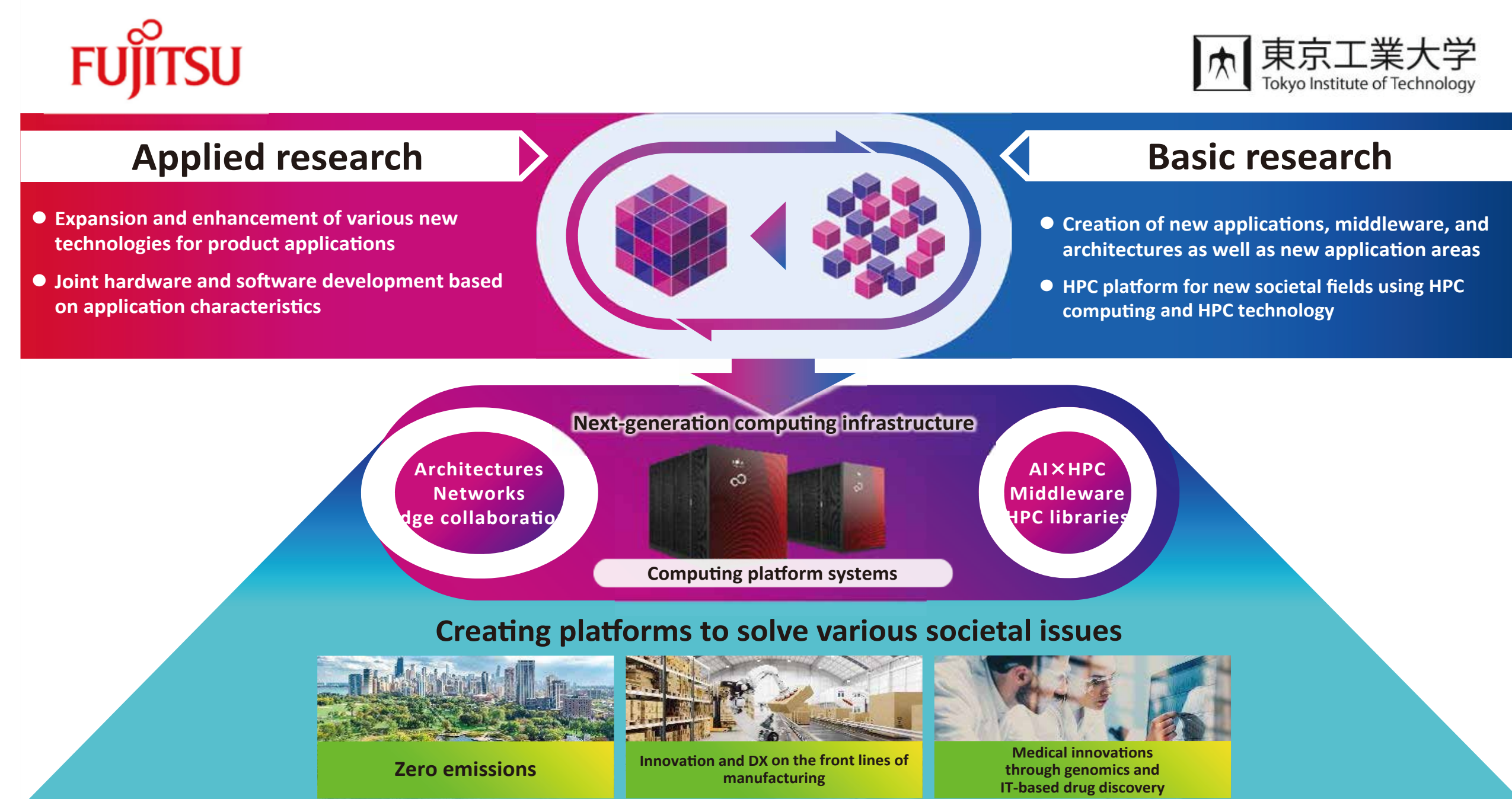
Leader: Hidehiko Masuhara (Tokyo Tech)

Sub-leaders: Naoki Akaboshi (Fujitsu),   Toshio Endo (Tokyo Tech)

Members (partial):

|	
Masahiro Miwa (Fujitsu)	Isao Ono (Tokyo Tech)
Miwa Ueki (Fujitsu)	Ryo Onishi (Tokyo Tech)
Takafumi Kanamori (Tokyo Tech)	Ryuichi Sakamoto (Tokyo Tech)
Rio Yokota (Tokyo Tech)	Akihiro Nomura (Tokyo Tech)

System Infrastructure Accommodating Interactive HPC

Motivation:

Usage of HPC system is wide-spreading, including interactive or realtime jobs. In order fulfill requirements, next-gen scheduling method is required.
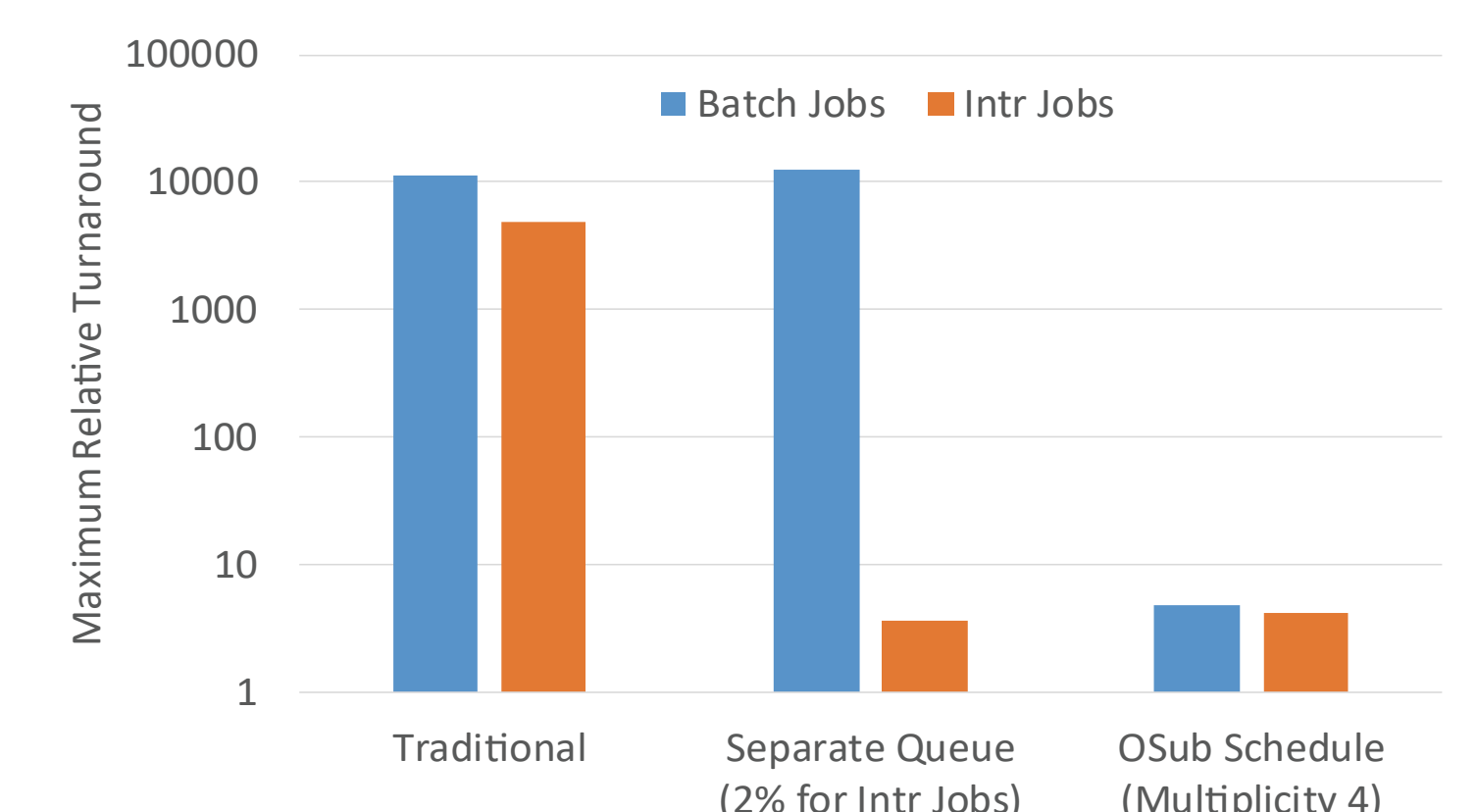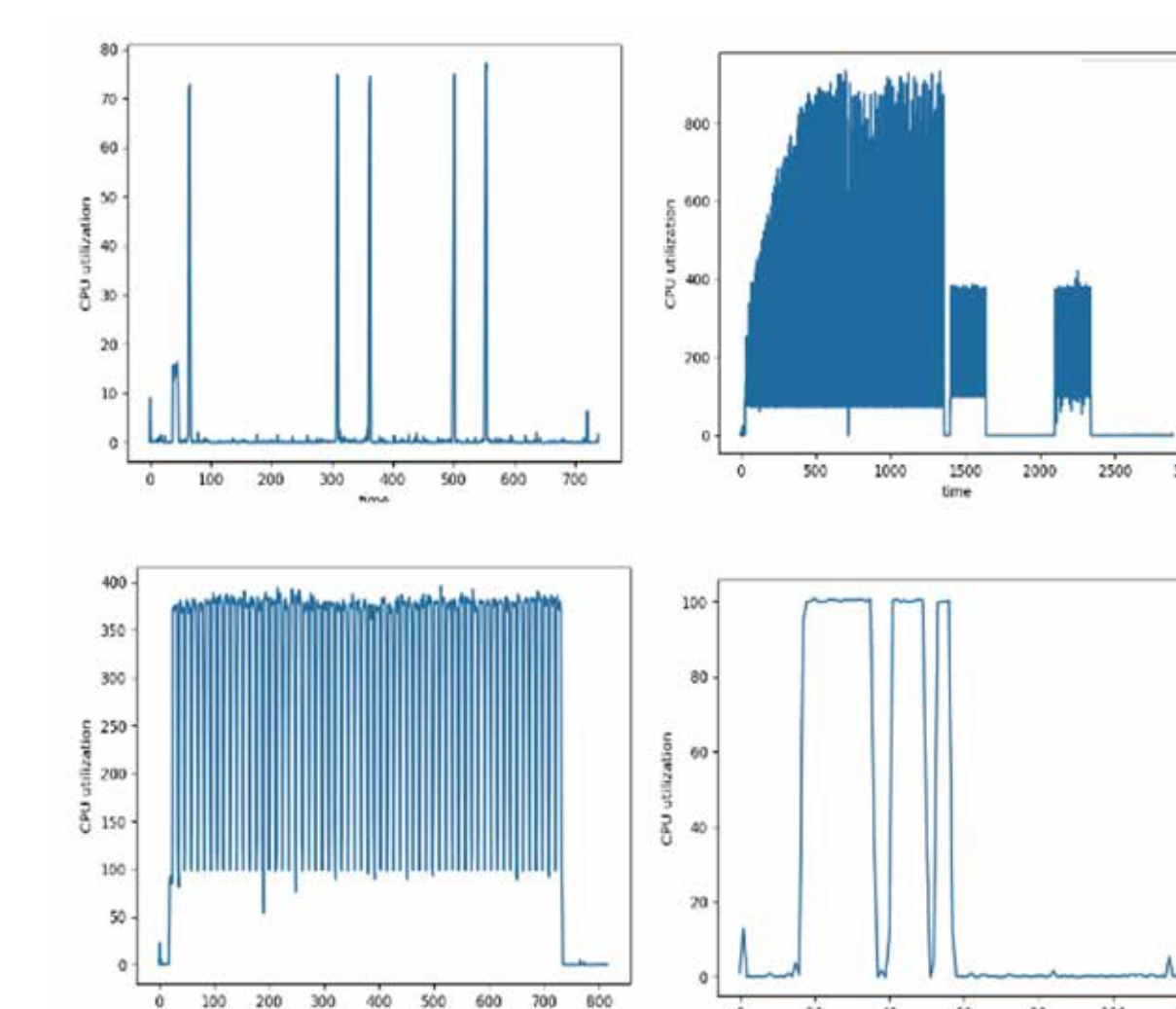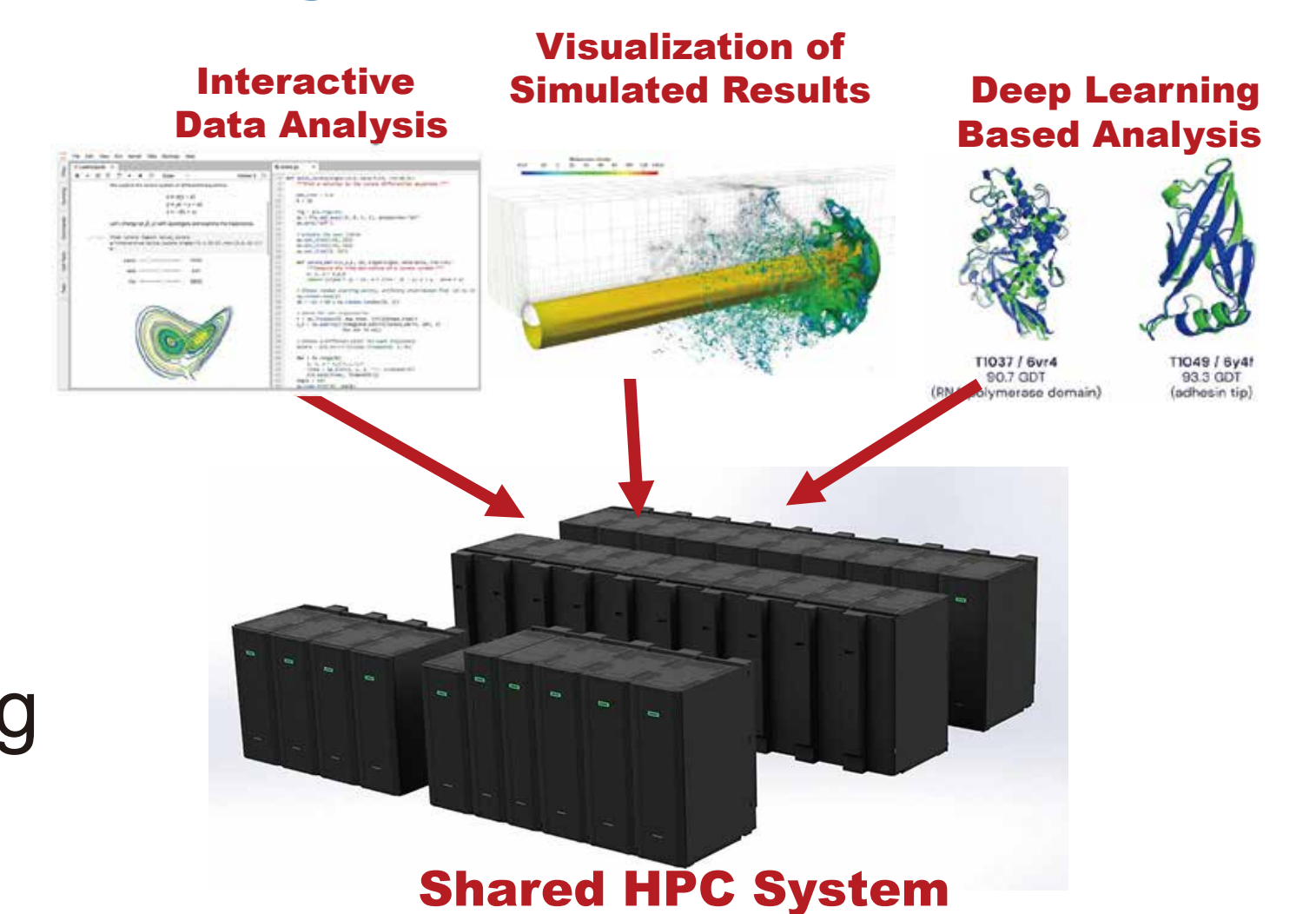
Interactive jobs have divergent CPU usage pattern. Data are from TSUBAME3

Scheduling simulation shows Oversubscribing (OSub) scheduling drastically improves responsiveness of jobs.
[Minami et al. IEEE HPEC 2023]

Research and Study of Operation Technologies

Collabolative research with U. Tokyo, RIKEN R-CCS, NII, and other institutes.

The Feasibility Study (FS) for next-generation computing infrastructure by MEXT, Japan was started on August 2022, and we are studying the field of operation technology of HPC. We aim to offer a next-generation system that can serve as a research DX infrastructure with an advanced digital twin to provide a platform for solving issues to realize broad SDGs.

Carbon Neutrality

- Procurement of green (low carbon emission) electricity
- Energy conservation of the entire facility including HPC systems
- Promotion of reuse of waste heat from HPC systems
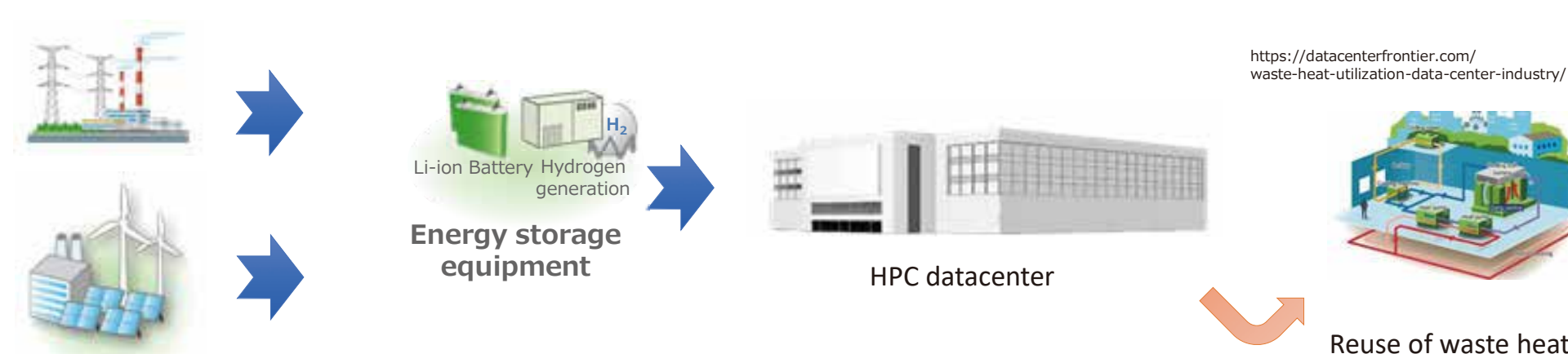
Facility and Equipment

- Investigation of hot-water cooling technology
- Analysis and demonstration of power-saving operation by real-time power consumption monitoring

Resource Management

- Coordinated resource management technology among major supercomputers
- Integration with the cloud (cloud bursting)
- Networking technologies to enable coordination
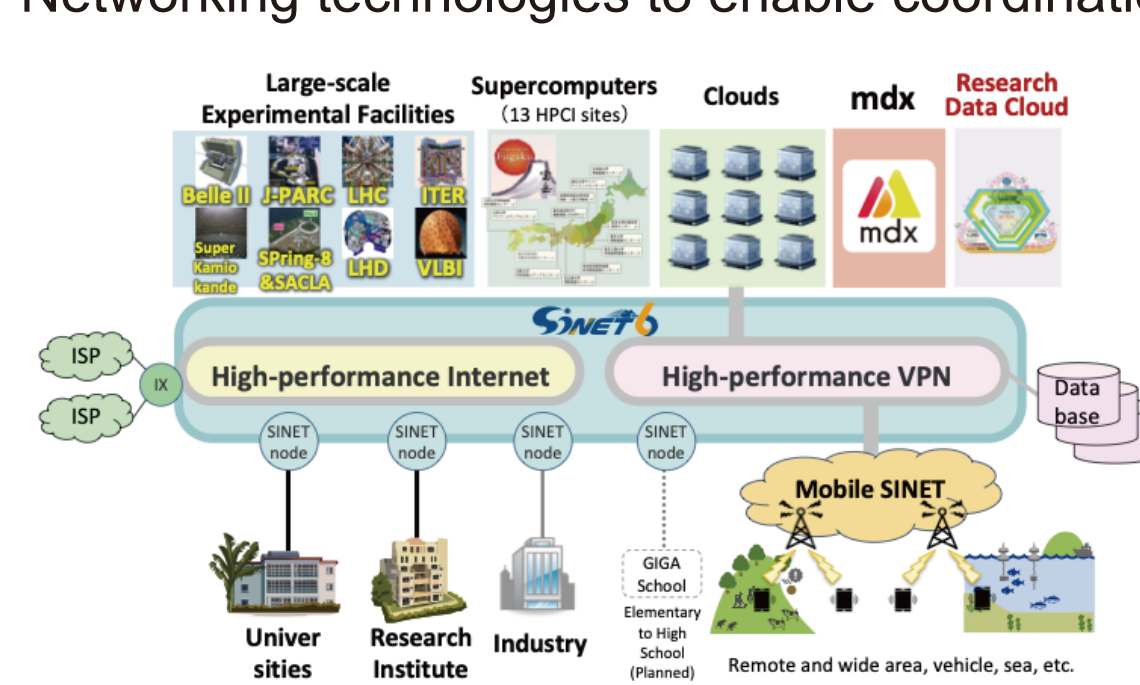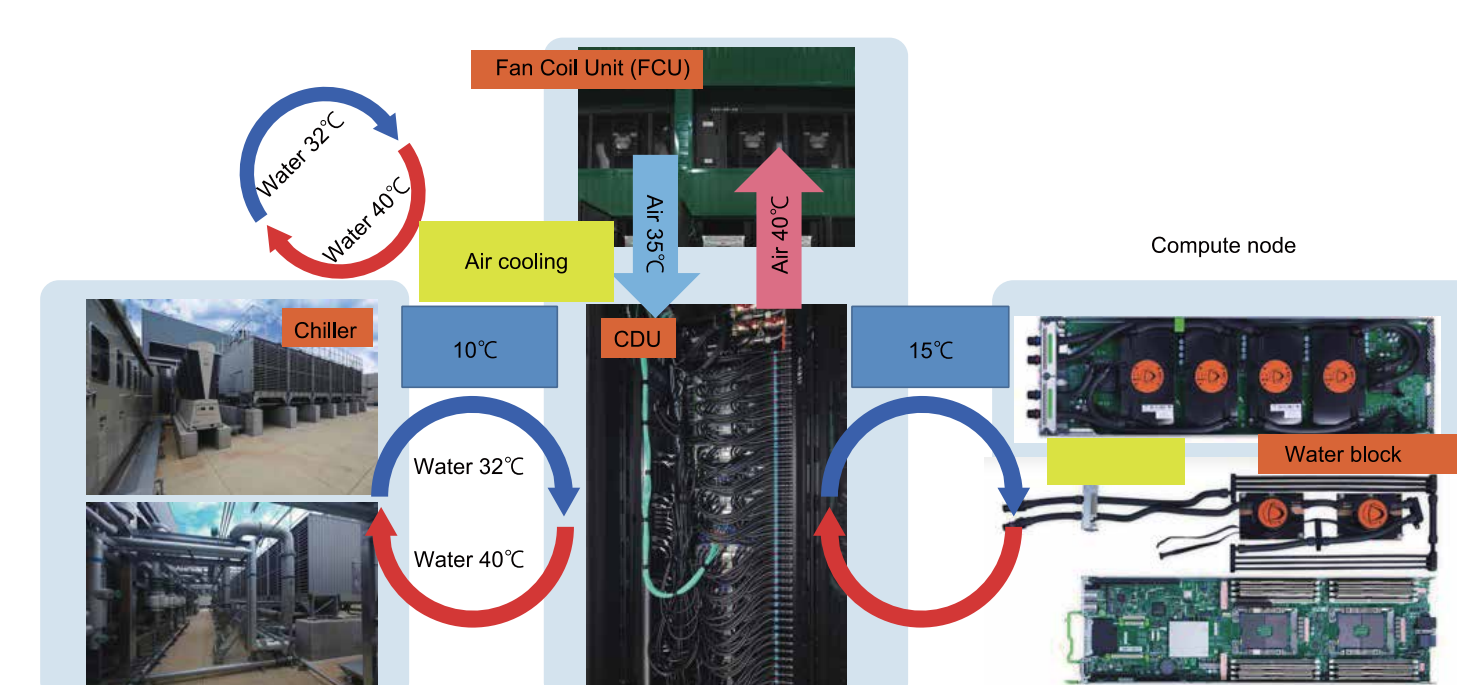
Data Leverage

Society 5.0, SDGs, Research DX

HPCI federated operation

Scalable Tracing for Post-5G Distributed System

Collabolative research with AIST, and other institutes.

In Post-5G era, data analysis systems are more distributed and heterogeneous including edge devices. In order to monitor and manage such systems, scalable distributed tracer is required.

Approaches:

- Collection methods of tracing data optimized for each system layer
- Cooperation among multiple layers and hierarchical distibuted storage
- Data compression and efficient storage using techniques

Initial Evaluation:

Towards scalable tracing, initial evaluation on response time with several system configurations has been done to investigate bottleneck.
- The Bookinfo microservice has been deployed and measured.
- Multiple connections using wrk2 are created and varying numbers of requests are sent in parallel.
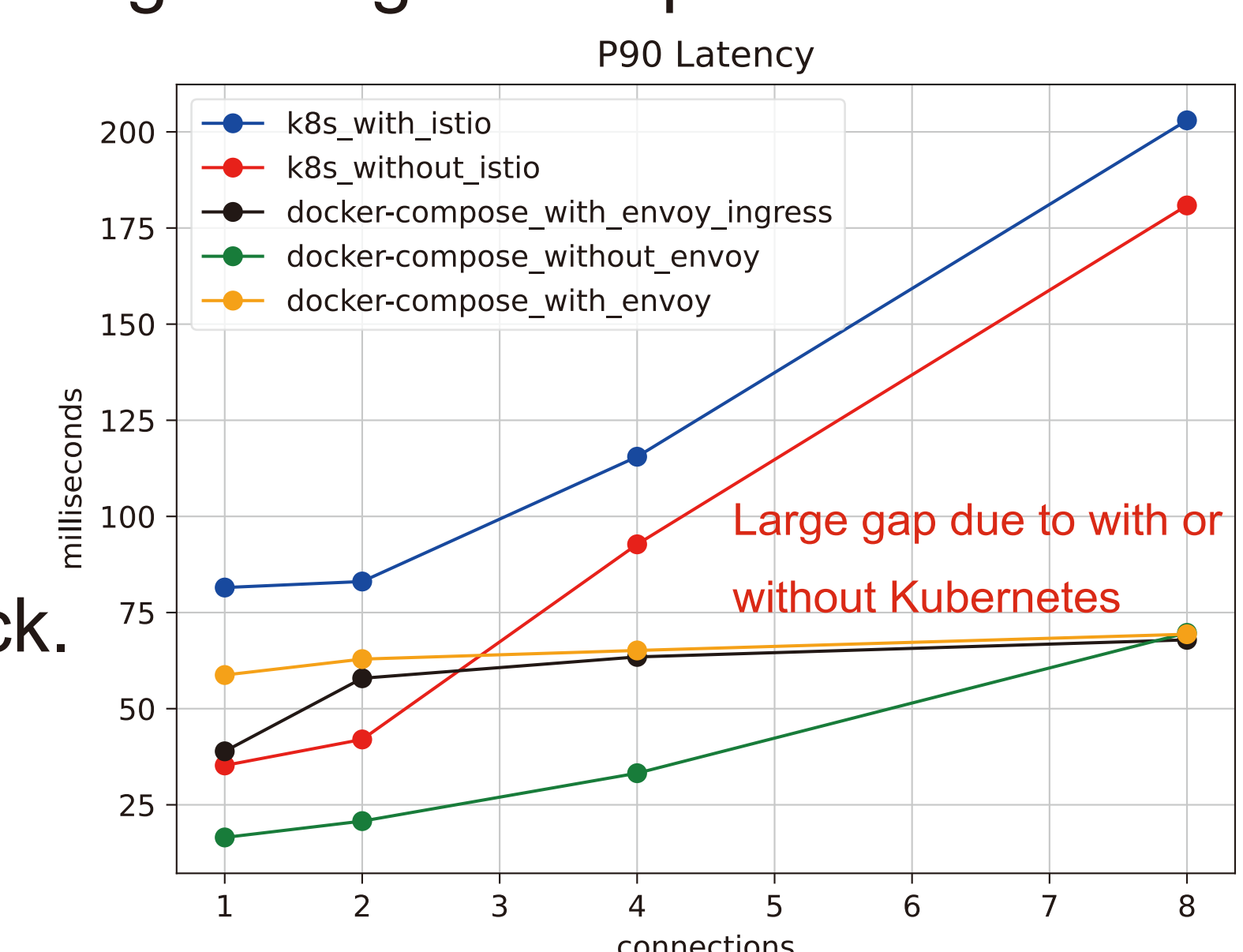
Costs introduced by Kubernetes (K8s) are observed.

http://www.gsic.titech.ac.jp/sc23