東京大学
THE UNIVERSITY OF TOKYO

Supercomputing & Data Science

Information Technology Center, The University of Tokyo

Welcome

Information Technology Center, The University of Tokyo

The Information Technology Center (ITC) (https://www.itc.u-tokyo.ac.jp/en/) was organized in 1999, and it consists of 4 research divisions (Campus-wide Computing, Data Science, Network, and Supercomputing). ITC is also a core organization of the "Joint Usage/Research Center for Interdisciplinary Large-Scale Information Infrastructures (JHPCN)" , and a part of HPCI (the High-Performance Computing Infrastructure) operated by the Japanese Government.

Supercomputing Research Division

The Supercomputing Division, Information Technology Center, The University of Tokyo (http://www.cc.u-tokyo.ac.jp/) was originally established as the Supercomputing Center of the University of Tokyo in 1965, making it the oldest academic supercomputer center in Japan. The Information Technology Center (ITC) was organized in 1999, and the Supercomputing Center became the Supercomputing Division (SCD) of the ITC. The three main missions of SCD/ITC are (i) Operations of Supercomputers & Services, (ii) Research & Development, and (iii) Education & Training. Currently, SCD/ITC consists of about 10 faculty members. SCD/ITC is now operating two supercomputer systems, "Massively Parallel Supercomputer System (Oakbridge-CX)" by Fujitsu with 6.61 PFLOPS, and "Integrated Supercomputer System for Simulation, Data and Learning (Wisteria/BDEC-01)" by Fujitsu with 33.1 PFLOPS. The two systems operated by SCD/ITC have 2,600+ users; 55+% of these users are from outside the university. Hands-on tutorials for parallel programming (on-line) are held 20+ times per year. Up to 10% of the total computational resources of each system are open to users from the industry.

Data Science Research Division

Nowadays, a huge volume of digital data is gathered from natural as well as artificial sources, including weather and seismic monitoring data, human and vehicle mobility data, and social activity data including business transactions, medical care, and so on. Digitally archiving historic documents and records at risk of dissipation also produces digital data. In these cases, digitization secures and enhances the value of knowledge by making it accessible regardless of physical distance. There is also a lot of new digital data being created, including web pages, social networks, academic papers, and so on. Advances in data analysis and modeling techniques, most notably machine learning, allow us to extract more meaningful and interpretable information from data and networking technology. This makes it possible to combine information from various sources. Data science is about turning raw data from a stream of digits into valuable insights and knowledge. Data science is also closely related to advances in high performance computing technologies, including high performance processors, storage and networking, big data analytics, deep learning numerical algorithms, and so on. The Data Science Research Division was established at the end of 2018, replacing the Academic Information Science Research Division. It takes a leadership role not only for the research on data science but also for designing and building a national infrastructure for the data science research community. It has been leading the development of the "mdx" platform, a national-wide academic cloud computing service in partnership with 10 universities and research institutes.

Computational Science Alliance, the University of Tokyo

At the University of Tokyo, we established the Computational Science Alliance (http://www.compsci-alliance.jp/) in 2015 by collaborating with 14 departments, including ITC. The primary purpose of this alliance is to provide an interdisciplinary education program for High-Performance Computing (HPC). The alliance started lectures in April 2017.

Computational Science Alliance The University of Tokyo

JHPCN: Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures

"JHPCN (https://jhpcn-kyoten.itc.u-tokyo.ac.jp/en/)" comprises academic supercomputer centers in Japan associated with eight national universities (Hokkaido, Tohoku, Tokyo, Tokyo Tech, Nagoya, Kyoto, Osaka, and Kyushu). This began in April 2010. The total performance of the supercomputer systems involved is 140+ PFLOPS (November 2022). JHPCN promotes collaborative research projects using the facilities and human resources of these eight centers, including the supercomputers, storage systems, and networks; interdisciplinary projects using multiple facilities are particularly encouraged. 40 or more projects have been accepted each year. New frameworks for international and industry collaborations have been initiated since 2017. Moreover, a new category focusing on data science/data analytics has been introduced in 2022, in addition to computational science. Finally, we have 63 projects (55: computational science, 8: data science/data analytics) in 2022.

JHPCN

JCAHPC: Joint Center for Advanced High Performance Computing

JCAHPC

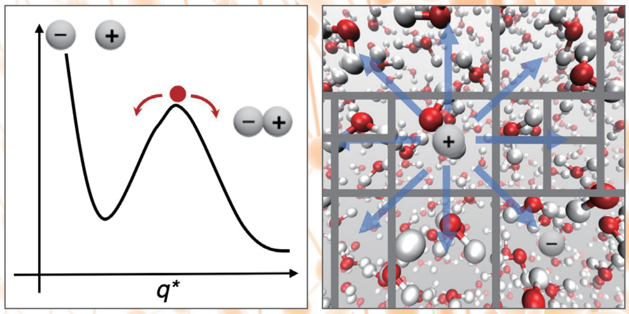
In 2013, Center for Computational Sciences, University of Tsukuba (CCS) and ITC agreed to establish the Joint Center for Advanced High-Performance Computing (JCAHPC). JCAHPC consists of more than 20 faculty and staff members of CCS and ITC. Originally, the Primary mission of JCAHPC is designing, installing, and operating the Oakforest-PACS system (OFP). In addition, CCS and ITC develop system software, numerical libraries, and large-scale applications for OFP in collaboration. JCAHPC is a new model for collaboration for research and development between multiple supercomputer centers. Oakforest-PACS (OFP) has contributed significantly to the development of computational science in Japan and around the world, especially after shutdown of the K computer in August 2019. Moreover, three proposals were adopted for the "HPCI Urgent Call for Fighting against COVID-19" in 2020. OFP retired on March 31, 2022. JCAHPC plans to introduce OFP-II with 200+ PFLOPS, the successor of OFP, which starts operation in April 2024. OFP-II has two groups of nodes: one group of heterogeneous nodes consisting of general-purpose CPUs and GPUs, and the other consisting of only CPUs.

Scientific Computing & Numerical Algorithms

Toward Establishing Compatibility Between Large-scale and Long-time Phenomena in Molecular Simulations

Molecular dynamics (MD) simulations are applied for a wide range of scientific problems, from physics and chemistry of materials & molecular biology to astrophysics. Thanks to the present status of the art where basic simulation methodology for parallel MD simulations is approaching its maturity, many open source packages are commonly available for usage on supercomputers. The problem nowadays is that MD simulations cannot be carried out for time that is long enough for describing physical processes in billion-atom-scale systems, as the increase in the clock cycles in the processors are about to peak out, and then it is now not easy to increase the number of total time steps.

We are now developing a molecular simulator that can allow computations on both large-scale and long-time. The key to realizing this is the efficient usage of machine learning for “perceiving” collective variables that appropriately describe the free-energy landscape of the system and enable unbiased long-time sampling, together with applying tree-based method for large-scale parallelism. Using the proper combinations of these methodologies, fundamental studies are ongoing regarding dynamical properties of liquid matter, including electrolytes, glasses, and other soft materials.

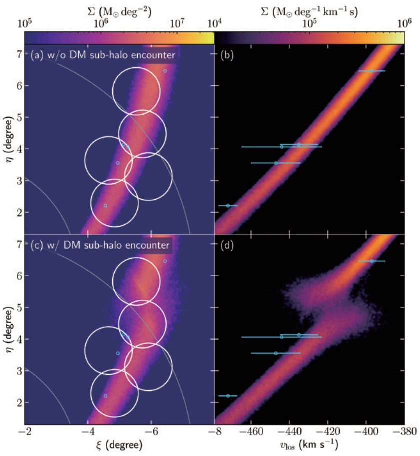


Left: Appropriate collective variables (q^*) that characterize the free energy landscape are necessary for realizing unbiased sampling of long-time trajectories.
Right: Tree methods are employed for effective computation of long-ranged electrostatics.

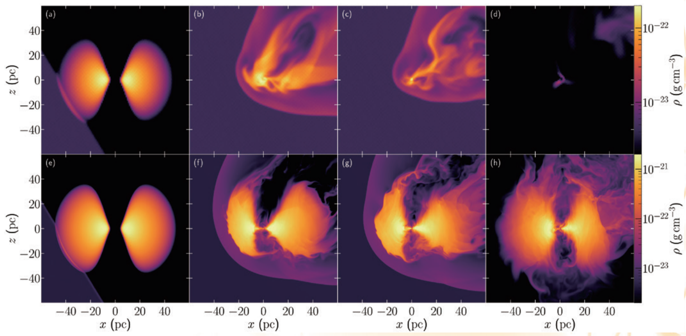
Performance Optimization of Gravitational Octree Code and Application to Galactic Archaeology

Collisionless N-body simulations are frequently employed to explore the formation and evolution of galaxies. We have developed a gravitational N-body code optimized for GPU: GOTHIC (Gravitational Oct-Tree code accelerated by HIerarchical time step Controlling). GOTHIC includes both the tree method and the hierarchical time step. The code runs entirely on GPU and is optimized for from the Fermi to the NVIDIA Ampere GPU architectures. A100, the current flagship GPU by NVIDIA, achieves a 1.3-fold acceleration compared to V100, the flagship GPU in the previous generation. The observed speed-up of 1.3 is greater than 1.2, which is the ratio of the theoretical peak performance of the two GPUs.

Stellar halos of galaxies provide fossil records of the formation and evolution of galaxies through galactic mergers. The Andromeda galaxy (M31) is an attractive laboratory for galactic archaeology due to its proximity and external perspective. N-body simulations using GOTHIC have reproduced stellar structures observed in the M31's halo. Stellar streams are promising probes to detect interactions between the streams and invisible dark matter (DM) sub-halos. The further synergy of N-body simulations on Wisteria/BDEC-01 and dedicated observations using the Subaru telescope will unveil the nature and assembly history of the DM halo.



Imprints of past DM sub-halo collision to stellar stream.

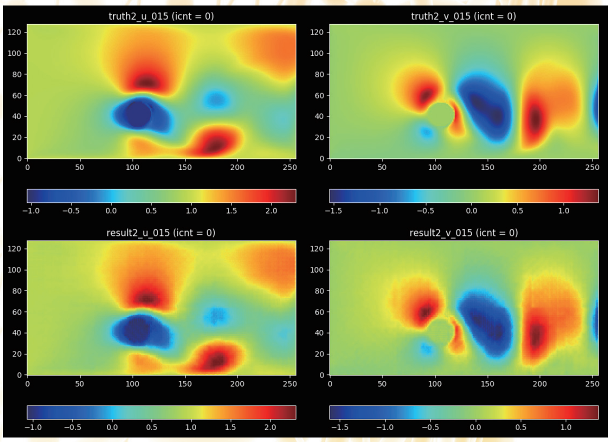


Time evolution of torus-shaped gas surrounding the massive black hole during a head-on galaxy collision. *(from Miki et al. 2021 https://doi.org/10.1038/s41550-020-01286-9)*

Development of Methods to Predict Fluid Simulation Results by Deep Learning

Computational fluid dynamics (CFD) is widely used in science and engineering. However, since CFD simulations require a large number of grid points and particles for these calculations, these kinds of simulations demand a large amount of computational resources such as supercomputers.

Recently, deep learning has attracted attention as a surrogate method for obtaining calculation results by CFD simulation approximately at high speed. We are working on a project to develop a parallelization method to make it possible to apply the surrogate method based on the deep learning to large scale geometry. Unlike the model parallel computing, the method we are currently developing predicts large-scale steady flow simulation results by dividing the input geometry into multiple parts and applying a single small neural network to each part in parallel. This method is developed based on considering the characteristics of CFD simulation and the consistency of the boundary condition of each divided subdomain. By using the physical values on the adjacent subdomains as boundary conditions, applying deep learning to each subdomain can predict simulation results consistently in the entire computational domain. It is possible to predict the simulation results in about 36.9 seconds by the developed method, compared to about 286.4 seconds by the conventional numerical method. In addition to this, we are also attempting to develop a method for fast prediction of time evolution calculations using deep learning.

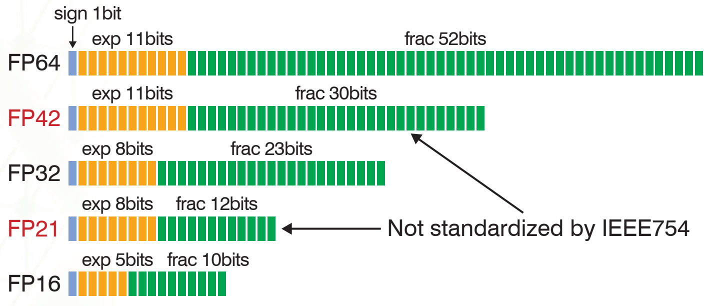


Comparison of the flow velocity results obtained by the conventional simulation (upper figure) and the prediction of these results by deep learning (lower figure).

System, Tools & Hardware

Studying Usability of Arbitrary Precisions

In recent years, the usefulness of low-precision floating-point representation has been studied in various fields such as machine learning. Low accuracy can be expected to have effects such as shortening calculation time and reducing power consumption. For example, in an application with a memory bandwidth bottleneck, the effect of reducing the calculation time by reducing the amount of memory transfer is significant. However, in fields such as iterative methods, it is common to use FP64 because the calculation accuracy strongly affects the convergence, and there are few application examples of low-precision arithmetic. This study investigates the applicability of low-precision representation to the Krylov subspace and stationary iterative methods. In this research, we focus on the FP32, FP16, and FP42, FP21, which are not standardized by IEEE754.



Comparison among each precisions

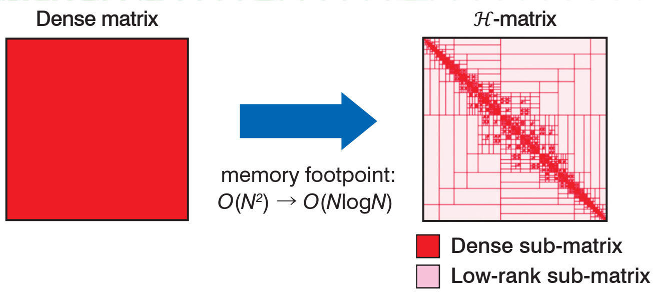
$\mathcal{H}$-matrices Library for Many-core Processors

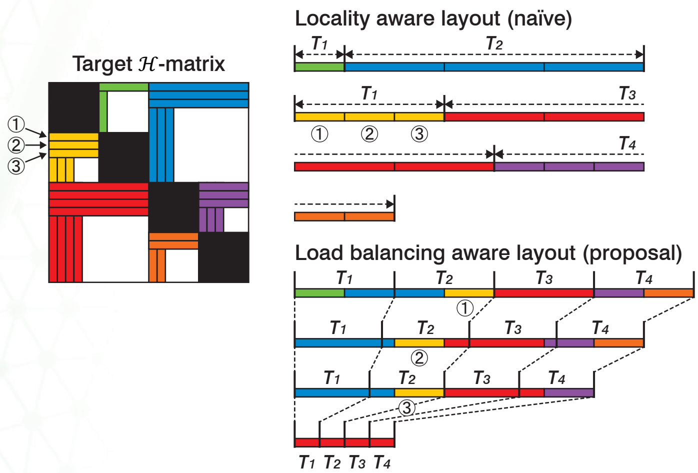
Hierarchical matrices ($\mathcal{H}$-matrices) are an approximation technique for dense matrices, such as the coefficient matrix of the boundary element method (BEM). An $\mathcal{H}$-matrix is expressed by a set of low-rank approximated and small dense sub-matrices, each of which has various ranks. The use of $\mathcal{H}$-matrices reduces the required memory footprint of dense matrices from O(N2) to O(NlogN) and is suitable for manycore processors that have relatively small memory capacities compared to traditional CPUs.

However, existing parallel adaptive cross approximation (ACA) algorithms, which are low-rank approximation algorithms used to construct $\mathcal{H}$-matrices, are not designed to exploit many-core processors in terms of load balancing.

In existing parallel algorithms, the ACA process is independently applied to each sub-matrix. The computational load of the ACA process for each sub-matrix depends on the sub-matrix's rank; however, the rank is defined after the ACA process is applied. This makes it difficult to balance the load.

We propose load-balancing-aware parallel ACA algorithms for $\mathcal{H}$-matrices that focus on many-core processors. We implemented the proposed algorithms into $\mathcal{H}$ACApK, which is an open-source $\mathcal{H}$-matrix library originally developed for CPU-based clusters.





UT-Helper: Support for HPC and Data Analysis Utilizing Unused Cores

In recent years, CPU performance improvement has been achieved by increasing the number of cores. However, using all the cores in the CPU may not be optimal for getting the best performance due to restrictions, such as memory bandwidth, power budget, and thermal dissipation.

To maximize the total efficiency, we utilize the "dark (unused) cores" and let them support the main task.

In this study, we develop the framework "UTHelper" to realize this mechanism without special modification of the existing system as the command line tools and libraries.

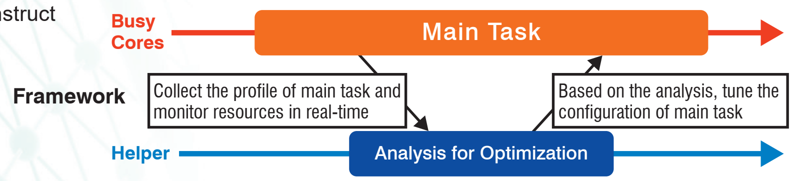
In UTHelper, to investigate optimal core assignments and to manage appropriate numbers of parallelism, we are implementing the tool based on SystemTap and OMPT in OpenMP standard that can provide the dynamic optimization for the target function without modification of the source code and recompile.

In addition, the helper threads that support properly the sub functions including communication and file IO are implemented. For example, we applied this concept to the astrophysics simulation with GPU for calculation and CPU for analysis in sequential, and successfully obtained optimal overlapping with small modification using OpenMP task construct and nested parallelism.

Work In Progress

- In-situ performance profiling
  - No additional codes is necessary for instruments
  - Measure behavior of main calculation and observe impact of helper functions
- Auto-adjust parallelism
- Cache prefetching
- Auto-adjust power budget
- Hiding communication and file IO
  - Apply to numerical library
- On-the-fly analysis during simulation and in-situ visualization
- Language extension using directive
  - Specify prefetch data, policies, QoS

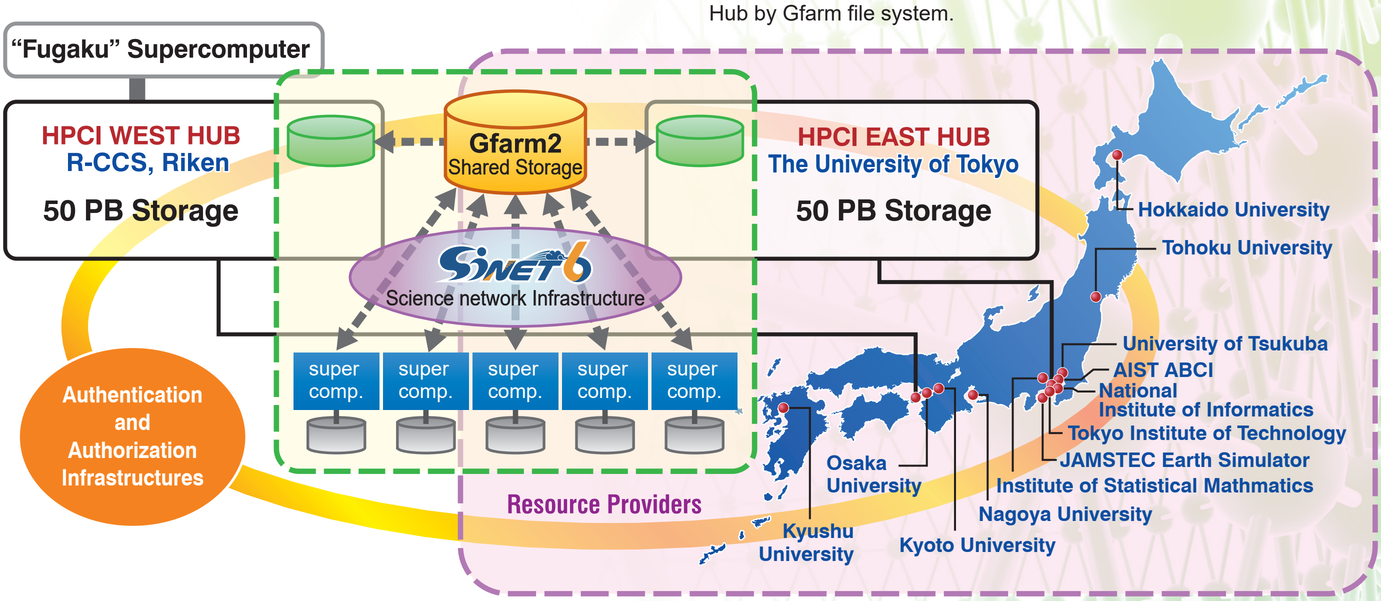
This work is supported by Japanese Government from FY.2020 to FY.2022 (JSPS Grant-in-Aid for Scientific Research (A), P.I.: Toshihiro Hanawa (ITC/U.Tokyo)).



Supercomputers at SCD/ITC

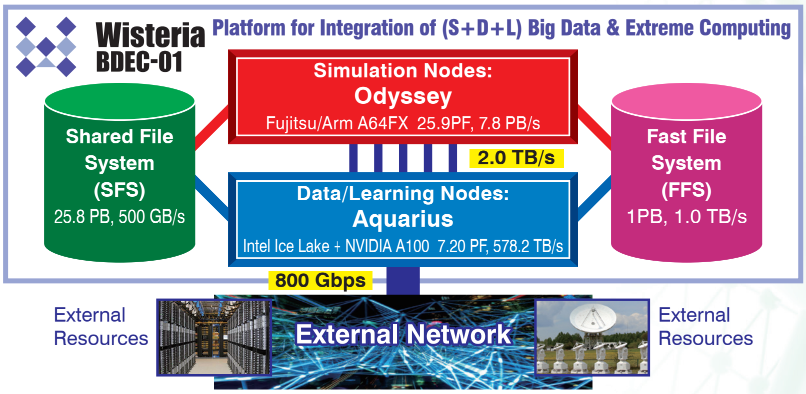
HPCI: High Performance Computing Infrastructure

High performance computing infrastructure (HPCI) is an environment that enables easy usage of flagship “Fugaku” supercomputer and other computation resources (tier-2) in Japan. In addition, HPCI is expected to match a user’s needs and computational resources to accelerate exploratory research, large-scale research, and industrial use of HPC. HPCI comprises 13 computational resource providers; nine are supercomputing centers at national universities, and four are governmental research institutes. These resource suppliers are connected via SINET6, which is a high-speed academic backbone network with 400 Gbps. SCD/ITC participates in this project as a hub resource provider in the Kanto region (the HPCI EAST Hub). The HPCI EAST Hub provides a 50-PB storage system integrated with the WEST Hub by Gfarm file system.



Wisteria/BDEC-01 (Fujitsu PRIMEHPC FX1000 & Fujitsu PRIMERGY)

We started discussions on the BDEC system (Big Data & Extreme Computing) as a platform for integration of (Simulation (S) +Data (D) +Learning (L))(S+D+L) in 2015. Wisteria/BDEC-01, which started its operation in May 2021, is the first BDEC system. Wisteria/BDEC-01 is a Hierarchical, Hybrid, Heterogeneous (h3) system, and it consists of two types of node groups for computing, Simulation Nodes (Odyssey) and Data/Learning Nodes (Aquarius), Shared File System (25.8 PB) and Fast File System (1.0 PB). The total peak performance is 33.1 PFLOPS, and aggregated memory bandwidth is 8.38 PB/sec. Simulation nodes for HPC (Odyssey) with more than 25 Peta FLOPS is based on Fujitsu’s PRIMEHPC FX 1000 with A64FX with High Bandwidth Memory. This part has the same architecture as that of the Fugaku supercomputer. Data/Learning nodes (Aquarius) are a GPU cluster consisting of Intel Xeon Ice Lake and NVIDIA A100 Tensor Core GPUs, with 7.2 Peta FLOPS for Data Analytics, AI and Machine Learning workloads. Some of Data/Learning nodes are directly connected to external resources through SINET, Japan. Odyssey and Aquarius are mutually connected through InfiniBand-EDR network with 2 TB/sec.



Wisteria/BDEC-01		Simulation Nodes: Odyssey	Data/Learning Nodes: Aquarius
Number of nodes		7680	45
Peak performance		25.9 PFlops	7.2 PFlops
Total memory Capacity		240.0 TiB	36.5 TiB
Network		Tofu-D: 6D-Mesh/Torus	InfiniBand HDR: Full-bisection Fat Tree
Compute node		FUJITSU Supercomputer PRIMEHPC FX1000	FUJITSU Server PRIMERGY GX2570 M6
CPU		Fujitsu/ARM A64FX, 48 cores, 2.2 GHz, 3.38TFLOPS, 32GB HBM2, 1,024 GB/sec	Intel Xeon Platinum 8360Y (Ice Lake), 36 cores x 2, 2.4GHz, 5.53TFLOPS, 512 GiB DDR4, 409.6 GB/sec
GPU		—	NVIDIA A100 GPUs, 108 SM’s, 19.5 TFLOPS, 40GB Memory, 1.555 TB/sec, 8 GPUs on each node
Shared File System	Name	FEFS (Fujitsu Exabyte File System)	
	OSS	DDN SFA7990XE x 16	
	Capacity & Performance	25.8 PB, 504 GB/sec	
Fast File System	Name	FEFS (Fujitsu Exabyte File System)	
	OSS	DDN SFA400NVXE x 16	
	Capacity & Performance	1.0 PB, 1.0 TB/sec	


Odyssey


Aquarius

Supercomputers at SCD/ITC

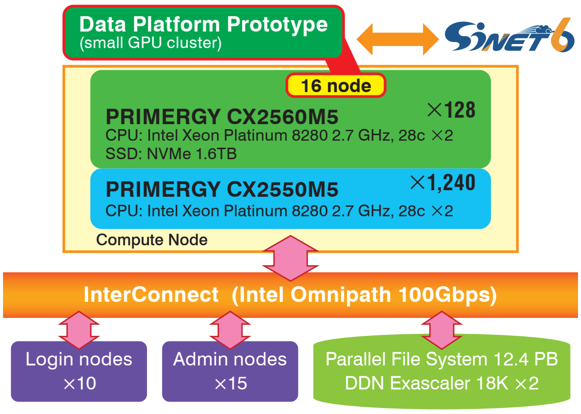
Oakbridge-CX (Fujitsu PRIMERGY)

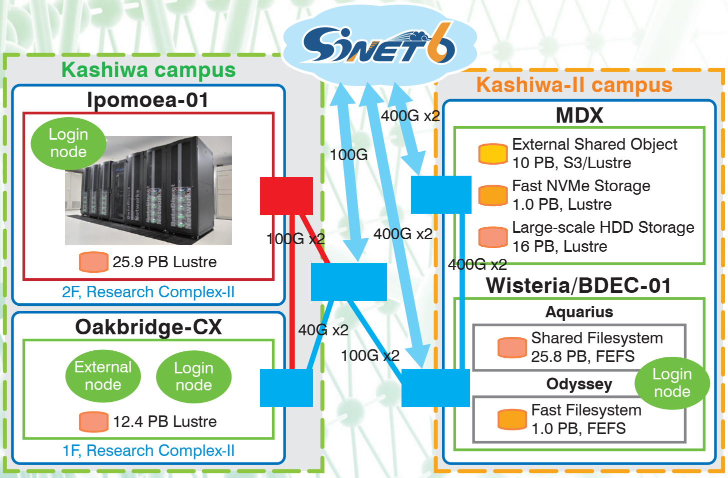
Oakbridge-CX (OBCX) is the Massively Parallel Supercomputer System using Intel Xeon Cascade Lake CPUs with the total performance of 6.61 PFLOPS. It has started operation since July 2019. The 128 nodes of compute node employ an NVMe SSD in each node for supporting staging, checkpointing, and data-intensive applications. Moreover, SSDs on designated nodes can be dynamically converged as a single shared file system using BeeGFS on Demand (BeeOND). Moreover, 16 of these 128 nodes are directly connected to external network (SINET, Japan) and can access external resources (e.g. servers, storages, sensor networks etc.). OBCX was a prototype of Wisteria/BDEC-01, which is a platform for the integration of “Simulation+Data+Learning”.

Peak performance	6.61 PFlops	
Total memory size	256.5 TByte	
Number of nodes	1240	128
Compute node	Fujitsu PRIMERGY CX2550 M5	Fujitsu PRIMERGY CX2560 M5
CPU	Intel Xeon Platinum 8280 (Cascade Lake, 28 cores, 2.7 GHz) 4.83 TFLOPS	
Memory	192 GB (DDR4)	
Interconnect	Intel Omni-Path (100 Gbps)	
Interconnect topology	Full Bisection BW Fat Tree	
SSD	—	1.6 TB(NVMe, Read: 3.20 GB/s, Write: 1.32 GB/s)
Parallel file system	Lustre Filesystem (DDN SFA18KE x2) 12.4 PB, 98 GB/s	

Ipomoea-01: Large-Scale Common Storage System

As the processing power of supercomputers has increased, the amount of data handled has also grown. SCD/ITC/UTokyo has traditionally installed storage attached to each supercomputer system, and storage for each system was independent. For this reason, the files had to be moved every time the system was replaced, which caused great inconvenience to the users. There was a strong need to introduce a common storage system that could be accessed from all systems at ITC/UTokyo. Since it is necessary to move files in 26 PB of shared storage when OFP ends its operation in March 2022, we decided to install a "Large-Scale Common Storage System (Ipomoea)" that can be accessed from each system in ITC/UTokyo including mdx. A new storage system will be installed, and replaced about every three years. Ipomoea-01 (Total Capacity: 25.9PB, Data Transfer Rate: 125 GB/sec) by Fujitsu started its operation in January 2022. The transfer of files from OFP was completed at the end of May 2022, and it has been available to the public since June.

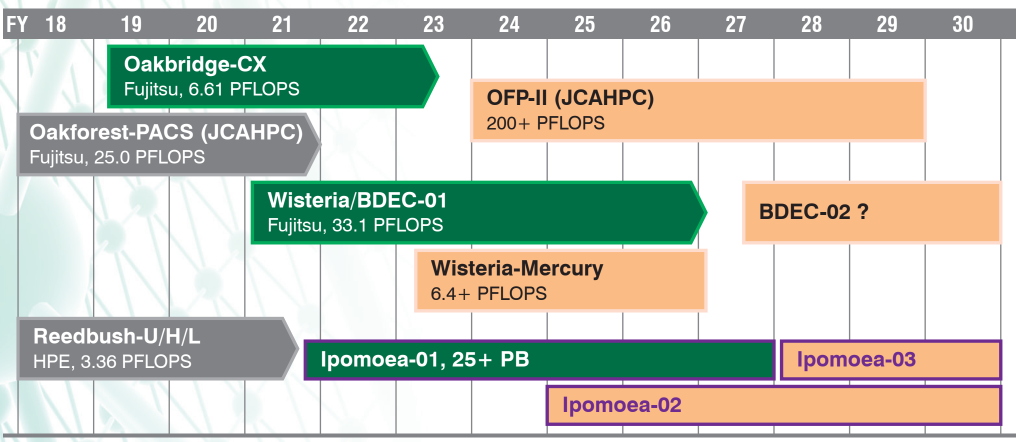




Operation	Fujitsu
Location	2F in Research Complex II,Kashiwa Campus
Filesystem	Lustre
Capacity	25.9 PB, 16.8 B i-nodes
Peak BW	125 GB/sec
Storage Servers	DDN ES7990X x5 set
Interconnect	100 GbE RoCE Ethernet

Next Supercomputer Systems in SCD/ITC

Wisteria-Mercury, an extended sub-cluster of Wisteria/BDEC-01, will be installed with 128+ GPUs and will start operation in June 2023. We expect it also plays a role as the prototype of the Oakforest-PACS-II (OFP-II) system. OFP-II, under collaboration with Center for Computational Sciences, University of Tsukuba, will start operation in Apr. 2024. It will be a heterogeneous system consisting of GPU-accelerated and CPU-only nodes with 200+ PFLOPS.



Supercomputers in SCD/ITC, University of Tokyo

Innovative Supercomputing towards Society 5.0 by Integration of "Simulation/Data/Learning (S+D+L)" using Wisteria/BDEC-01 and h3-Open-BDEC

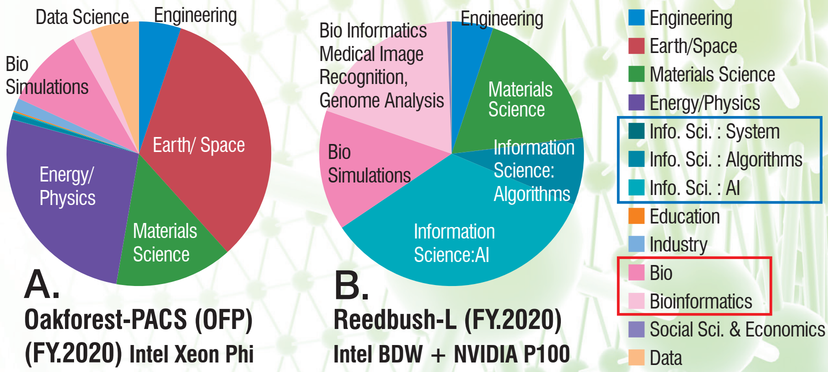
What is Society 5.0 ?

https://www8.cao.go.jp/cstp/english/society5_0/index.html

Society 5.0 was proposed in the 5th Science and Technology Basic Plan by the Cabinet Office of Japan as a future society that Japan should aspire to. It follows the hunting society (Society 1.0), agricultural society (2.0), industrial society (3.0), and information society (4.0). Society 5.0 is a human-centered society that balances economic advancement with the resolution of social problems by a system that highly integrates cyberspace and physical space, and will be achieved by Digital Innovation, such as IoT, AI, Big Data and etc.

New Directions in Supercomputing

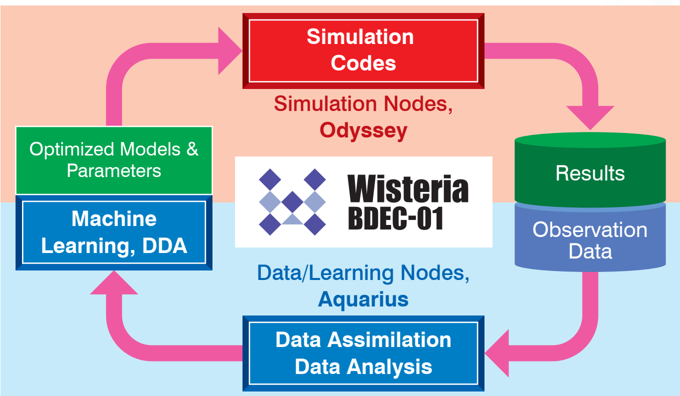
Majority of SCD/ITC supercomputer system users belong to the fields of CSE (Computational Science & Engineering), including engineering simulations (fluid dynamics, structural dynamics, and electromagnetics), earth science (atmosphere, ocean, solid earth, and earthquakes), and materials science, as shown in the pie chart A, which shows usage rate of each research area on Oakforest-PACS (OFP) with Intel Xeon Phi based on CPU hours in FY.2020. Recently, the number of users related to data science, machine learning, and artificial intelligence (AI) has been increasing, as shown in the pie chart B, which shows usage rate on Reedbush-L system with GPUs in FY.2020. Examples of new research topics are weather prediction by data assimilation, medical image recognition, and human genome analyses. Towards Society 5.0, a new type of method for solving scientific problems by integrations of "Simulation (S)" , "Data (D)" and "Learning (L)" (S+D+L) is emerging.



h3-Open-BDEC: Innovative Software Platform for Wisteria/BDEC-01

We started discussions on the BDEC system (Big Data & Extreme Computing) as a platform for integration of (S+D+L) since 2015. Wisteria/BDEC-01, which started its operation in May 2021, is the first BDEC system.

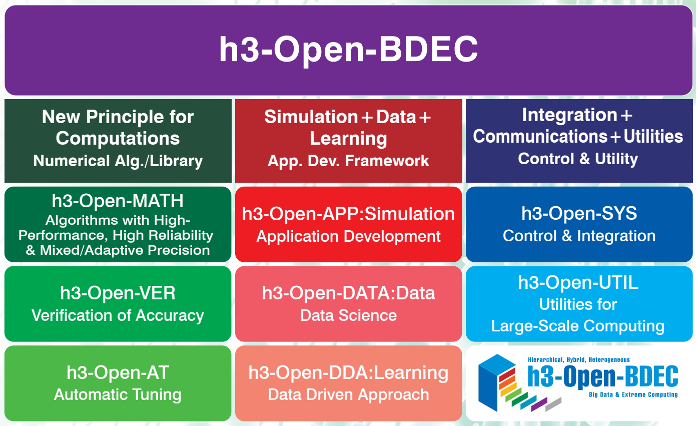
Integration of (S+D+L) on Wisteria/BDEC-01 is done by optimization of parameters for large-scale simulations on Odyssey using data analyses, data assimilation and machine learning on Aquarius.



(S+D+L) on Wisteria/BDEC-01

We develop an innovative software platform "h3-Open-BDEC" for integration of (S+D+L), and evaluate the effects of integration of (S+D+L) on Wisteria/BDEC-01. The h3-Open-BDEC is designed for extracting the maximum performance of the supercomputers with minimum energy consumption focusing on (1) innovative method for numerical analysis with high-performance/high-reliability/power-saving based on the new principle of computing by adaptive precision, accuracy verification and automatic tuning, (2) Hierarchical Data Driven Approach (*h*DDA) based on machine learning, and (3) software and utilities for heterogeneous environment, such as Wisteria/BDEC-01. Examples of (3) are described in the next section.

This work is supported by Japanese Government from FY.2019 to FY.2023 (JSPS Grant-in-Aid for Scientific Research (S), P.I.: Kengo Nakajima (ITC/U.Tokyo)). In Data Driven Approach (DDA), technique of machine learning is introduced for predicting the results of simulations with different parameters. DDA generally requires a lot of simulations for generation of teaching data. We propose the *h*DDA, where simplified models for generating teaching data are constructed automatically by machine learning with Feature Detection, MOR (Model Order Reduction), UQ (Uncertainty Quantification), Sparse Modeling and AMR (Adaptive Mesh Refinement).



Overview of h3-Open-BDEC

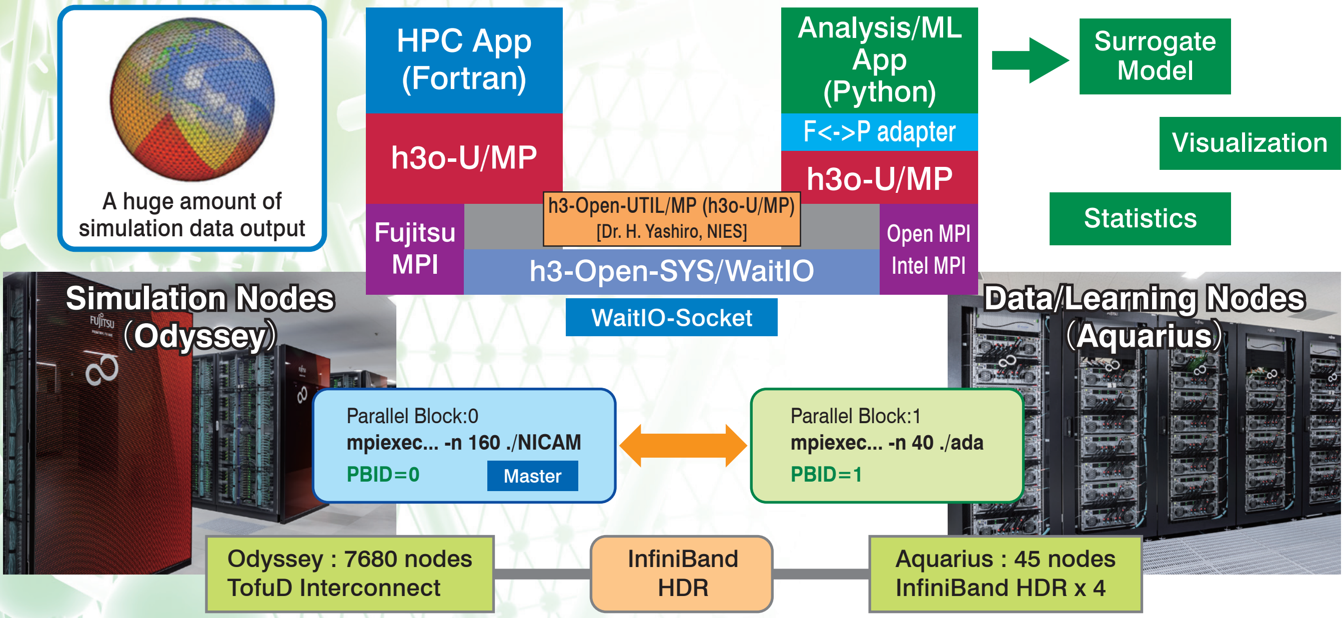
The h3-Open-BDEC is the first innovative software platform to realize integration of (S+D+L) on supercomputers in the Exascale Era, where computational scientists can achieve such integration without supports by other experts. Source codes and documents are open to public for various kinds of computational environments. This integration by h3-Open-BDEC enables significant reduction of computations and power consumptions, compared to those by conventional simulations. Possible applications using h3-Open-BDEC are combined simulations/data assimilations for climate/weather simulations and earthquake simulations, and real-time disaster simulations, such as flood, earthquake and tsunami.

Innovative Supercomputing towards Society 5.0 by Integration of "Simulation/Data/Learning (S+D+L)" using Wisteria/BDEC-01 and h3-Open-BDEC

h3-Open-SYS/WaitIO-Socket & h3-Open-UTIL/MP: System-wide Communication for Heterogeneous Coupling Computing

h3-Open-SYS/WaitIO-Socket (WaitIO-Socket) is a system-wide communication library to couple multiple MPI programs for heterogeneous environments, such as Wisteria/BDEC-01. WaitIO-Socket provides an inter-program communication environment among MPI programs and supports different MPI libraries with various interconnects and processor types. WaitIO-Socket provides an MPI like application program interfaces and can easily connect multiple MPI programs. As a higher-level software based on this library, a coupler for coupling multiple MPI programs (h3-Open-UTIL/MP) has been developed simultaneously with h3-Open-SYS/WaitIO as part of the h3-Open-BDEC project. h3-Open-UTIL/MP provides capabilities for multi-physics coupling between different grid systems. Furthermore, it is equipped with a function for combined ensemble, and with an interface for coupling codes for large-scale simulations and applications written in Python.

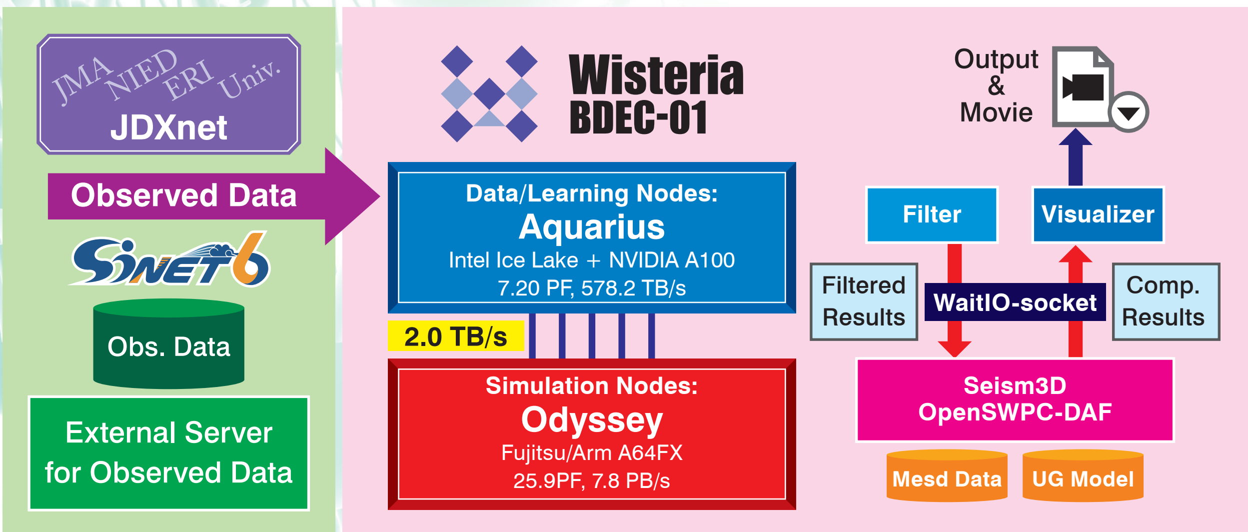
NICAM-ADA coupling utilizes h3-Open-UTIL/MP (h3o-U/MP) for converting NICAM data in a high-resolution grid on Odyssey to a low-resolution grid for machine learning by GPU's on Aquarius. While such coupled computing has been only possible by MPI on a single/homogeneous system, h3-Open-UTIL/MP with WaitIO-Socket provides more flexible interface for integration of (S+D+L) on heterogeneous system, such as Wisteria/BDEC-01.



Heterogeneous Coupling Computing by h3-Open-SYS/WaitIO and h3-Open-UTIL/MP (NICAM+ADA)

Integration of 3D Earthquake Simulations with Real-Time Data Observation & Assimilation

JDXnet is the seismic observation data exchange and distribution network in Japan, and it distributes the seismic observation data of the nine national universities, JMA, NIED, JAMSTEC, AIST etc. at more than 2,000 points. We utilize the observation data obtained from JDXnet for integration of 3D earthquake simulation in strong motion and real-time data assimilation. We developed a prototype of this integrated simulations with real-time data assimilation on Wisteria/BDEC-01 using h3-Open-SYS/WaitIO-Socket.



Integration of 3D Earthquake Simulations with Real-Time Data Observation & Assimilation on Wisteria/BDEC-01 using h3-Open-BDEC

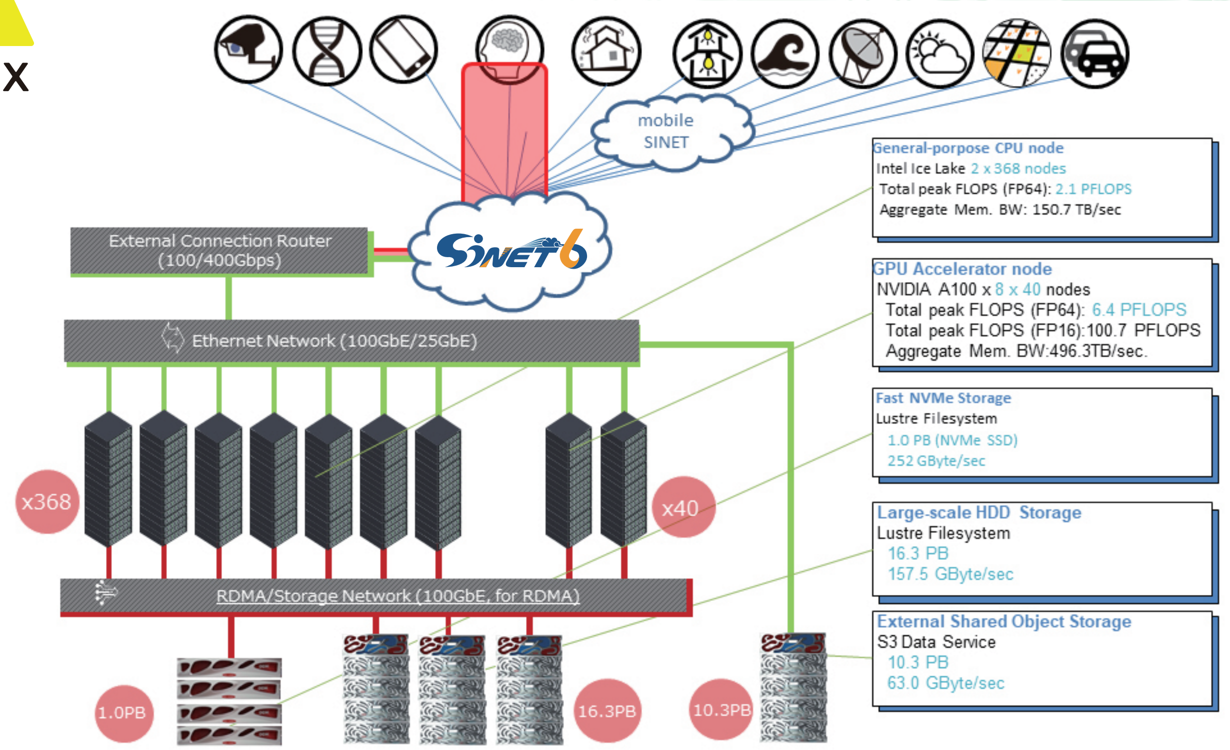
Cloud Platform and Infrastructure at ITC

mdx: A Cloud Platform for Supporting Data Science and Cross-Disciplinary Research Collaborations

The growing amount of data and advances in data science have created a need for a new kind of cloud platform that provides users with flexibility, strong security, and the ability to couple with supercomputers and edge devices through high-performance networks. We have built such a nation-wide cloud platform, called “mdx” to meet this need. The mdx platform’ s virtualization service, jointly operated by 9 national universities and 2 national research institutes in Japan, launched in 2021, and 50+ projects have been using it. Currently mdx is used by researchers in a wide variety of domains, including materials informatics, geo-spatial information science, life science, astronomical science, economics, social science, and computer science. Through the mdx platform, we anticipate more collaborations for such problems that need interdisciplinary approaches by bringing knowledge and skills in data science and application domains.

The mdx platform currently provides a VM hosting service (IaaS) and will provide data sharing and collaboration tools (PaaS) in the future. It consists of Fujitsu PRIMERGY, which is equipped with 368 CPU nodes and 40 GPU nodes as computing resources, and 1.0 Petabyte NVMe disk storage and 16.3 Petabyte hard disk for the Lustre file system, and 10.3 Petabyte S3-compliant object storage (DDN S3 Data Services) as storage resources. Each compute node employs 2 sockets of Intel Xeon Platinum 8368 with 38 cores of 2.4GHz clock, and each GPU node contains 8 GPUs of NVIDIA A100. These are virtualized by VMware vSphere virtualization platform and co-designed with SINET6, the ultra high-speed/low-latency academic backbone network in Japan. The network is virtualized via L2VPN to achieve high-performance and secure networks connecting to edge devices.

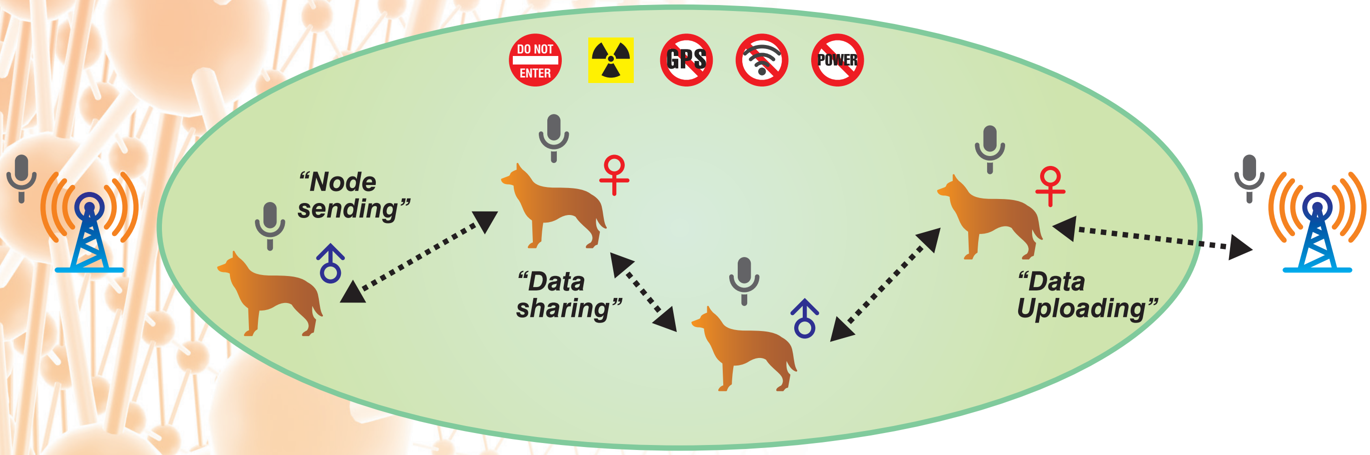
		General-purpose CPU node	GPU Accelerator node
Number of nodes		368	40
Compute node		Fujitsu PRIMERGY CX2550M6	Fujitsu PRIMERGY GX2570M6
CPU		Intel Xeon Platinum 8368 (Ice Lake SP, 38 cores, 2.4 GHz) x 2 sockets 5.83 TFLOPS	
Memory		256 GB (DDR4-3200 x 8ch x 2), 409.6 GB/s	
GPU		None	NVIDIA A100 GPUs (19.5 TFLOPS, 40 GB, 1.555 TB/s, SXM4, NVlink3) x 8
Interconnect (Front)		Ethernet (25 Gbps)	Ethernet (25 Gbps) x 2 link
Interconnect (RDMA / Storage)		Ethernet w. RoCEv2 (100 Gbps)	Ethernet w. RoCEv2 (100 Gbps) x 4 link
Storage	Fast	1.0 PB (NVMe SSD), 252 GB/s	
	Large-scale	16.3 PB (HDD), 157.5 GB/s	
	External shared object	10.3 PB (HDD), 63.0 GB/s, S3 Data service	
Virtualization		VMware vSphere, Overlay network w. EVPN-VXLAN	



Data Science Research

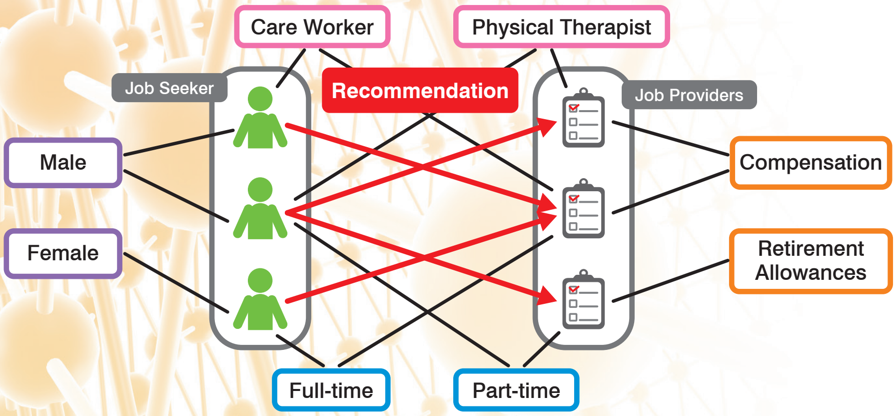
Information Design and Animal to Animal Information Network utilized outer urban infrastructure

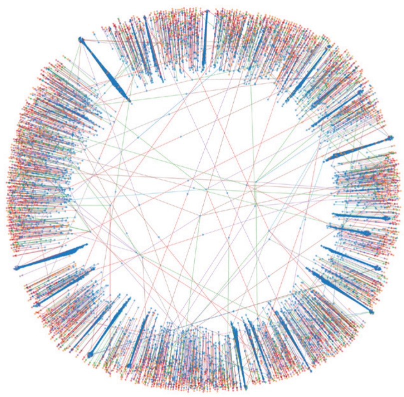
To solve environmental problems that occur in spaces where information communication technologies are difficult to be introduced, we are researching systems to realize holistic advanced information processing, in which the information space and the ecosystem are inextricably integrated. This research includes information design of animal IoT for wild animals living in the exclusion zone around the Fukushima Daiichi Nuclear Power Plant.

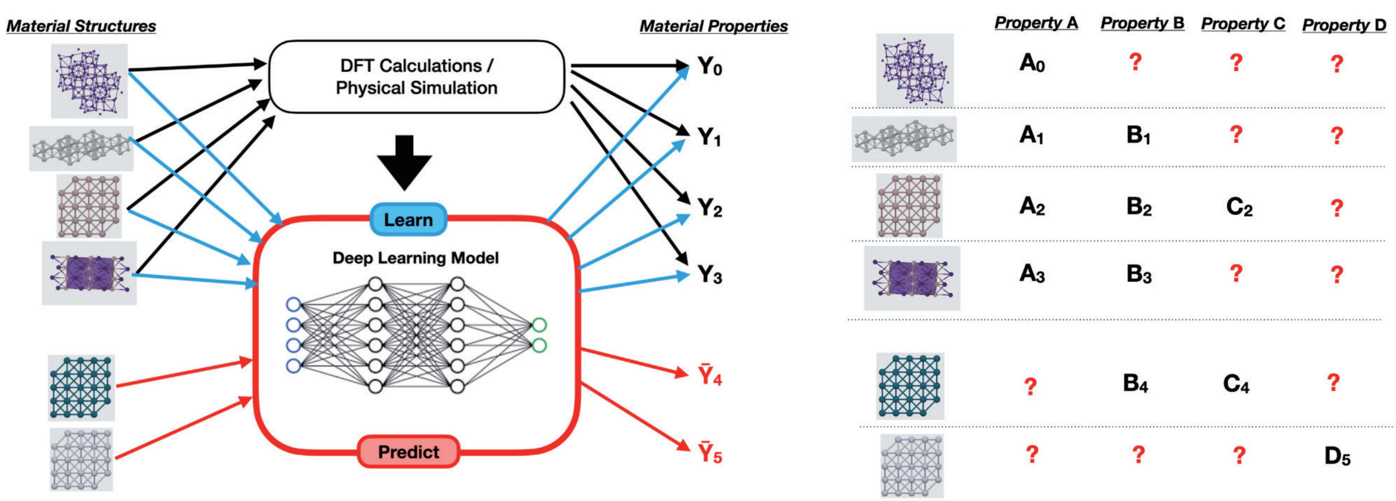


Large-Scale Graph Neural Networks

All the entities in digital space and real world - including things, facts, and human beings - and their relationships can be represented as nodes and edges - leading to large-scale dynamic graphs in graph theory. We work on graph neural networks or GNNs that can learn graph structures and roles of nodes and edges via deep learning. As GNN applications, we target various domains including recommender systems in e-commerce, news platform, and mobility, and also fraud detection in financial systems. We have been also investigating Materials Informatics, the interdisciplinary research field between data science/machine learning and materials science/engineering. We are working on the development of an effective method to store the big materials data, such as the theoretical results of the physical simulations and the real data from the experimental instruments. Moreover, by using these data, we are developing machine learning methods for the material property predictions using GNNs to effectively handle a molecular graph.



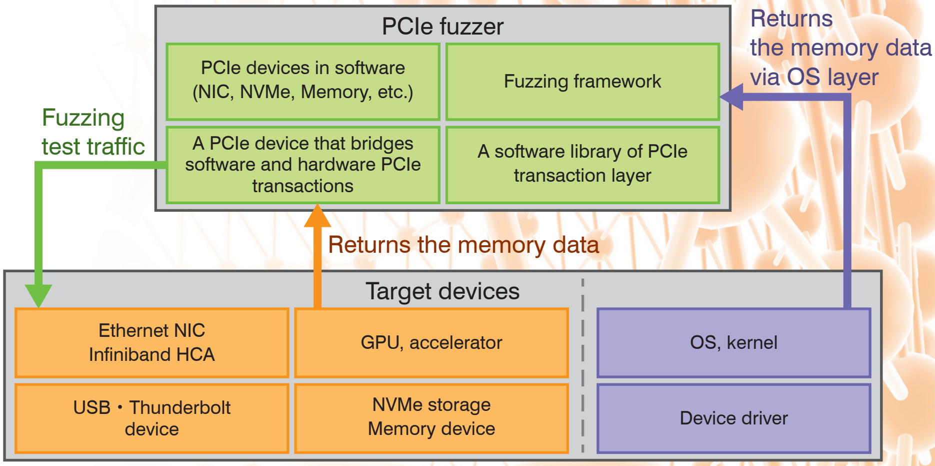




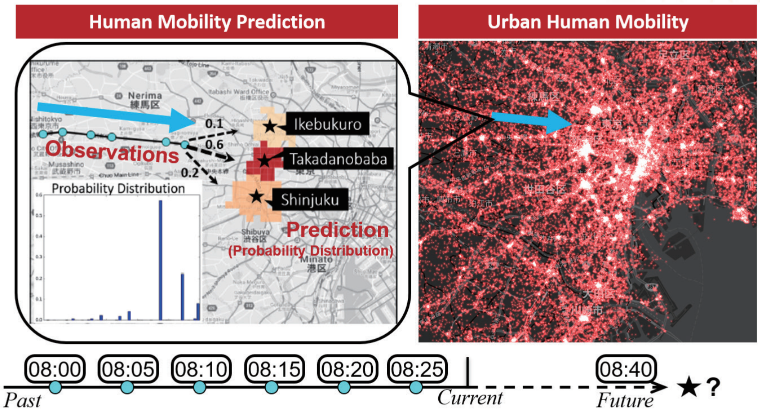
Data Science Research

Adapting software vulnerability testing for data center hardware

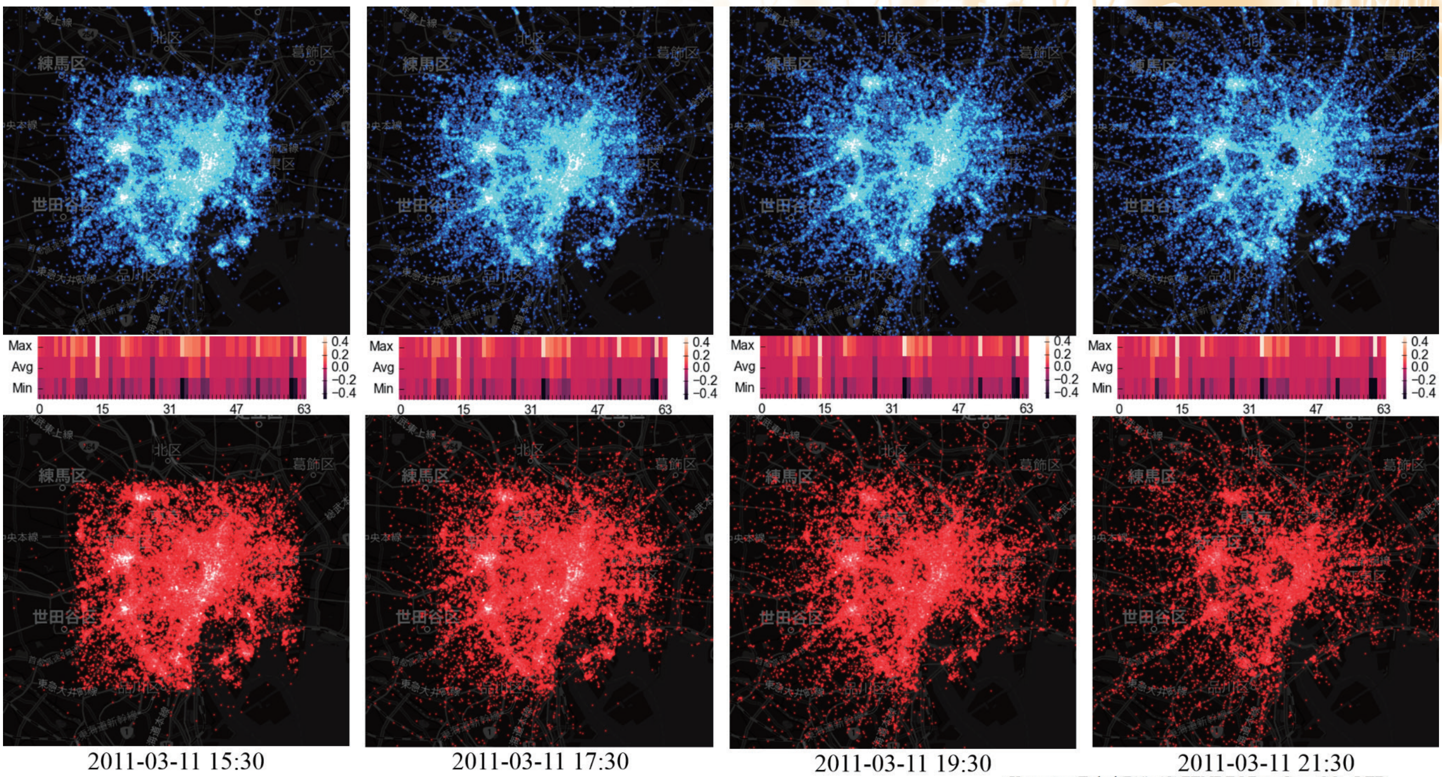
Datacenter computing environments use modern high-speed interconnects and accelerator devices to improve computing performance. We' re developing a method for early detection of hardware defects and vulnerabilities to safely use data center hardware such as GPU, SSD, and NIC.



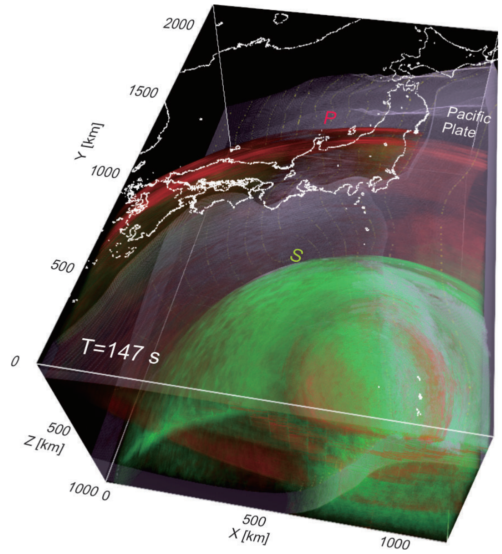
Modeling and Analysis of Urban Human Mobility with Deep Learning



Citywide human mobility prediction and simulation is one of the most important technologies for building smart cities and realizing an ultra-smart society (Society 5.0). Based on large-scale human GPS trajectory data, our research is aimed at modeling, simulation, and prediction of urban human mobility by fully utilizing state-of-the-art deep-learning technology to handle the complex spatiotemporal correlations. Through this approach, we can provide technical assistance for future urban management, such as humanitarian assistance in the event of a disaster, city emergency management, and intelligent transportation systems.

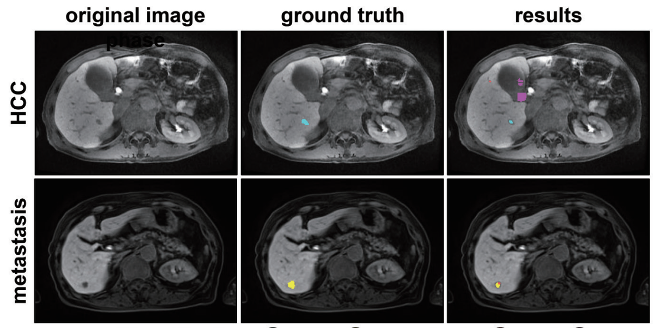


User Applications



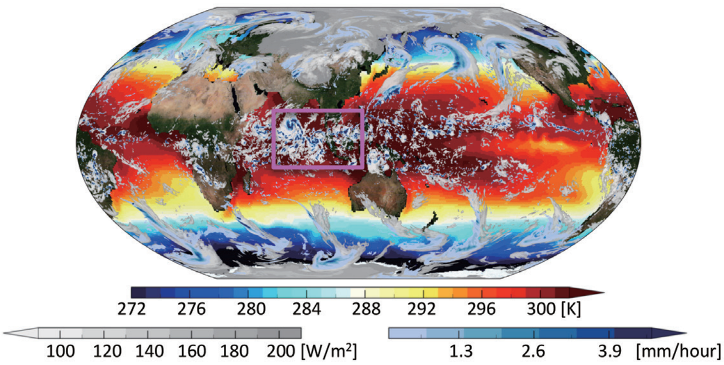
Seismic Wave Propagation Simulation
c/o T. Furumura (ERI/U.Tokyo)

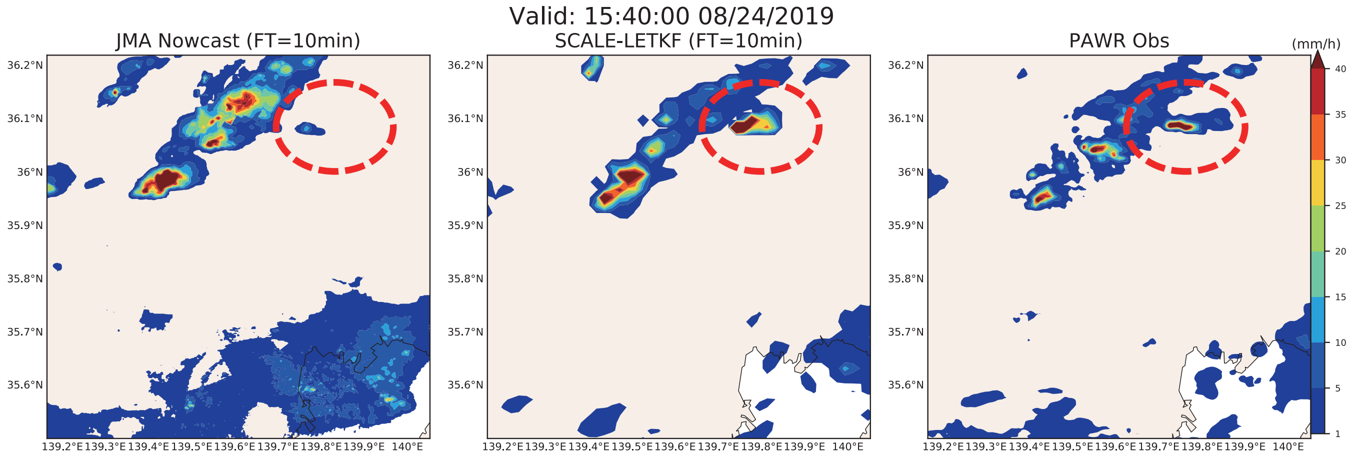
Focal Liver Lesion Detection by Deep Learning
c/o T. Takenaga (U.Tokyo Hospital)



Classification: HCC, metastasis, cyst, scar

Atmosphere-Ocean Simulations: NICOCO
c/o T. Miyakawa (AORI/U.Tokyo)





Numerical Weather Prediction System: SCALE-LETKF **c/o T. Miyoshi & T. Honda (RIKEN R-CCS)**

東京大学情報基盤センター
INFORMATION TECHNOLOGY CENTER, THE UNIVERSITY OF TOKYO
6-2-3 Kashiwanoha, Kashiwa-shi, Chiba 277-0882, JAPAN TEL: +81-4-7133-4658

Visit our website for more information
https://www.cc.u-tokyo.ac.jp/en/public/pamphlet.php

2022.10