Innovative Supercomputing by Integration of Simulation/Data/Learning

Kengo Nakajima
Information Technology Center
The University of Tokyo
RIKEN R-CCS

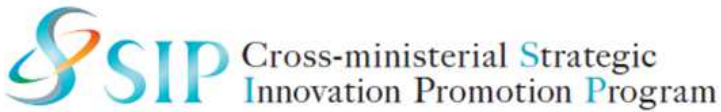
Wisteria
BDEC-01

Hierarchical, Hybrid, Heterogeneous
h3-Open-BDEC
Big Data & Extreme Computing

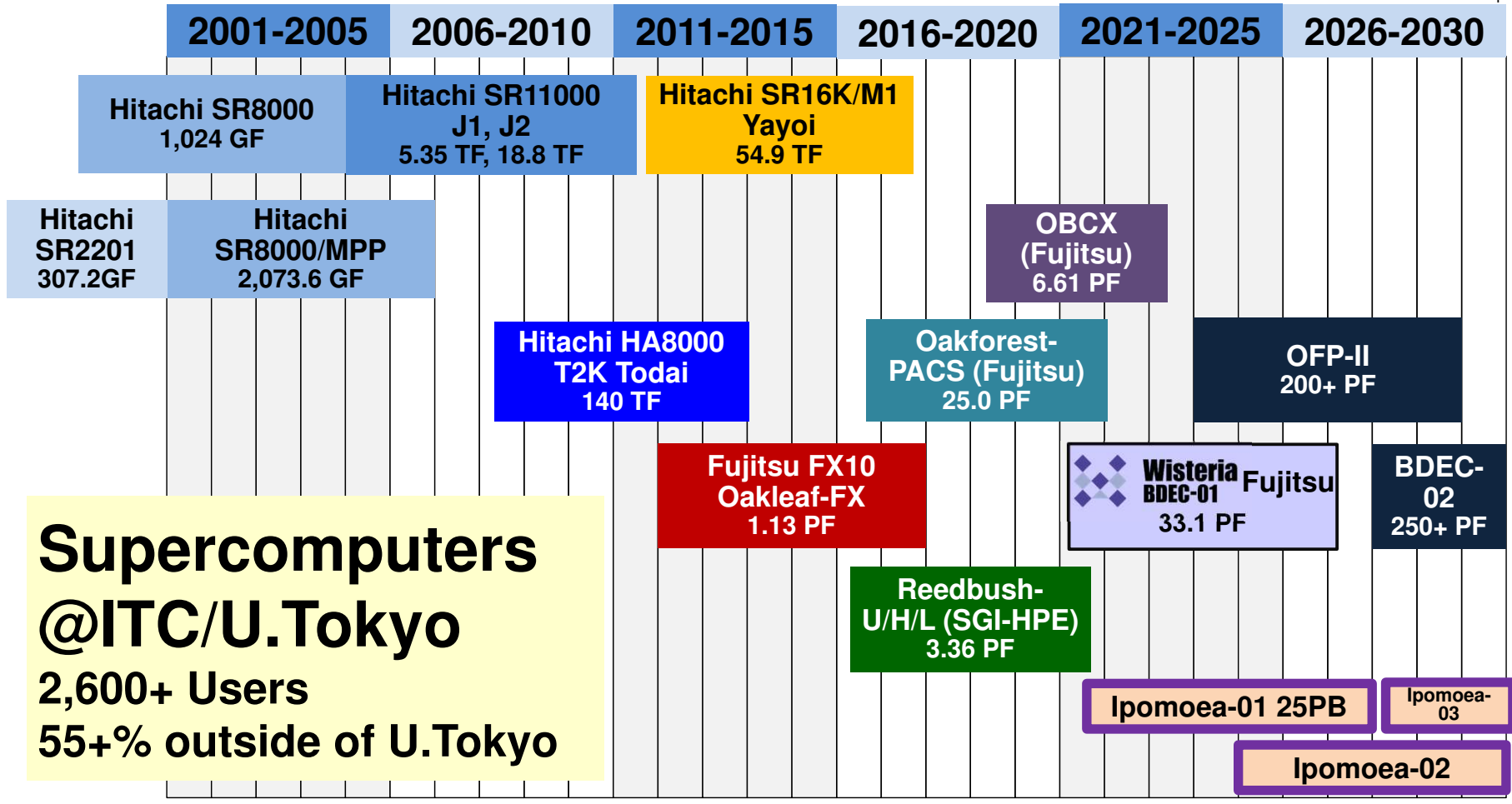
HPC Challenges for New Extreme Scale Applications
Paris, France, March 6-7, 2023

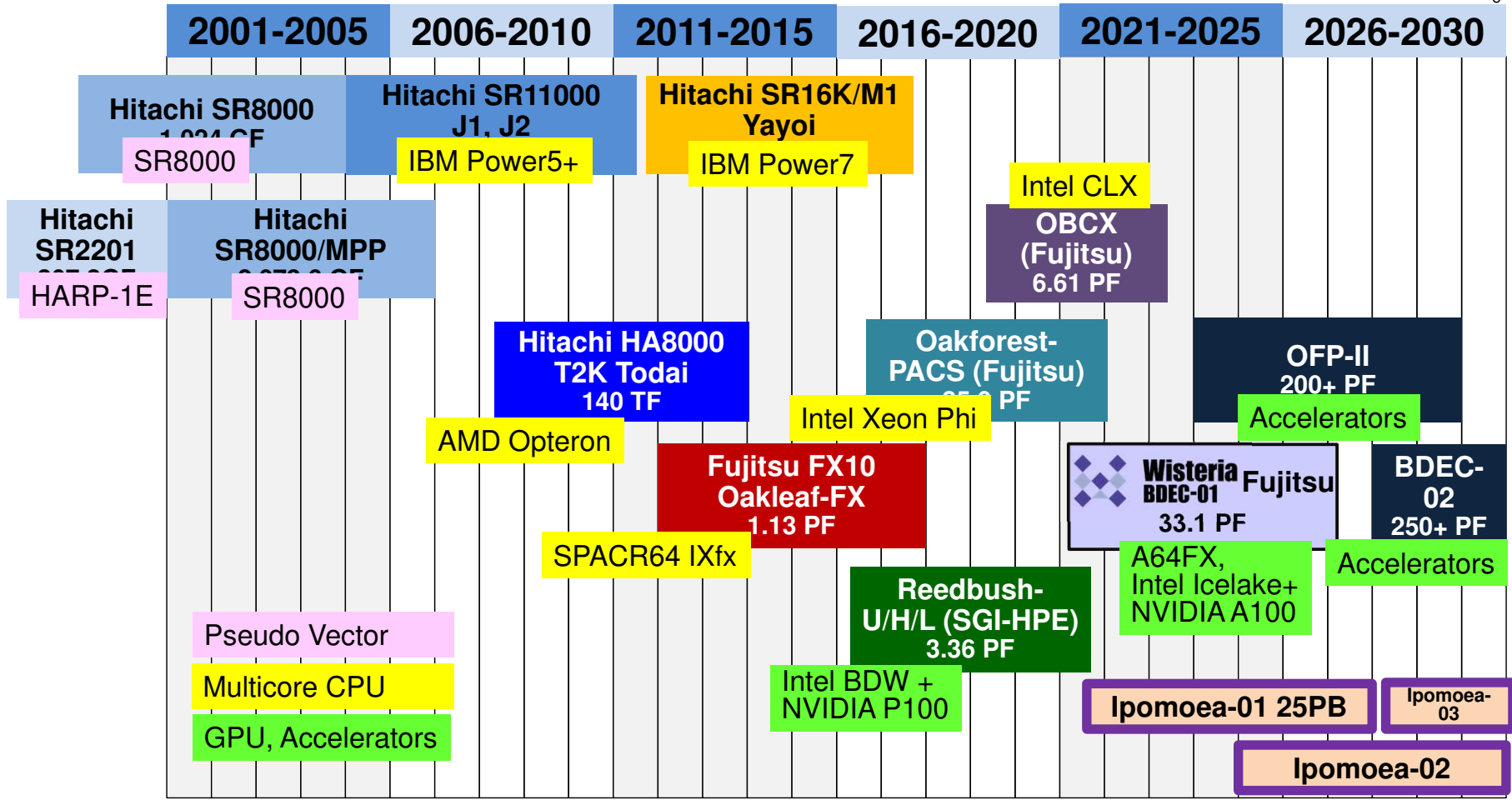
Acknowledgements

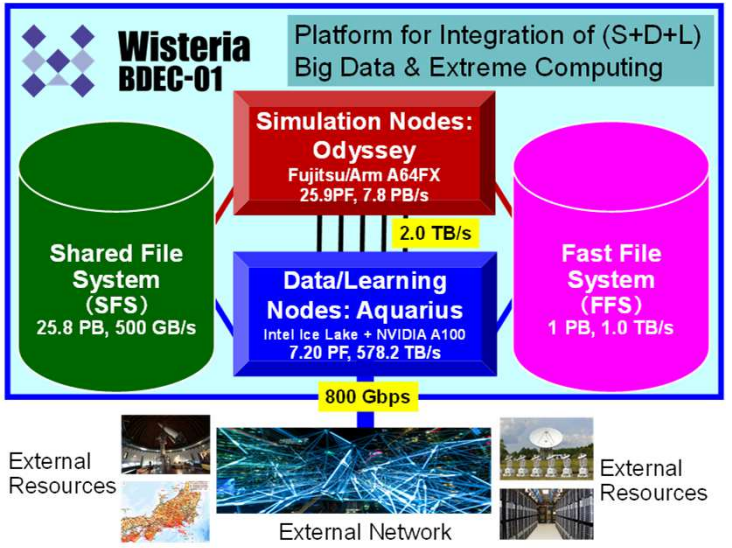
- JSPS Grant-in-Aid for Scientific Research (S) (19H05662)
- New Energy & Industrial Technology Development Organization (NEDO): Cross-ministerial Strategic Innovation Promotion Program (SIP): Big-Data and AI-Enabled Cyberspace Technologies
- Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures (JHPCN)
  - jh210022-MDH, jh220029



- Wisteria/BDEC-01
- h3-Open-BDEC
- Earthquake Simulations combined with Real-Time Data Assimilation
  - h3-Open-SYS/WaitIO
- HPC-AI Workload: Global Cloud Simulations
  - h3-Open-UTIL/MP
- How to run such (a little bit complicated) workflows
  - Wisteria/BDEC-01
  - Wisteria/BDEC-01 + mdx
- Future/Ongoing Works
  - International Collaborations
  - OFP-II


2001-2005
2006-2010
2011-2015
2016-2020
2021-2025
2026-2030
Hitachi SR8000
1,024 GF
Hitachi SR11000
J1, J2
5.35 TF, 18.8 TF
Hitachi SR16K/M1
Yayoi
54.9 TF
Hitachi
SR2201
307.2GF
Hitachi
SR8000/MPP
2,073.6 GF
OBCX
(Fujitsu)
6.61 PF
Hitachi HA8000
T2K Todai
140 TF
Oakforest-
PACS (Fujitsu)
25.0 PF
OFP-II
200+ PF
Fujitsu FX10
Oakleaf-FX
1.13 PF
Wisteria
BDEC-01
Fujitsu
33.1 PF
BDEC-
02
250+ PF
Reedbush-
U/H/L (SGI-HPE)
3.36 PF
Ipomoea-01 25PB
Ipomoea-
03
Ipomoea-02
Supercomputers
@ITC/U.Tokyo
2,600+ Users
55+% outside of U.Tokyo


2001-2005
2006-2010
2011-2015
2016-2020
2021-2025
2026-2030
Hitachi SR8000
SR8000
Hitachi SR11000 J1, J2
IBM Power5+
Hitachi SR16K/M1 Yayoi
IBM Power7
Intel CLX
OBCX (Fujitsu) 6.61 PF
Hitachi SR2201
HARP-1E
Hitachi SR8000/MPP
SR8000
Hitachi HA8000 T2K Todai 140 TF
AMD Opteron
Oakforest-PACS (Fujitsu)
Intel Xeon Phi
OFP-II 200+ PF
Accelerators
Fujitsu FX10 Oakleaf-FX 1.13 PF
SPACR64 IXfx
Wisteria BDEC-01 Fujitsu 33.1 PF
A64FX, Intel Icelake+ NVIDIA A100
BDEC-02 250+ PF
Accelerators
Reedbush-U/H/L (SGI-HPE) 3.36 PF
Intel BDW + NVIDIA P100
Pseudo Vector
Multicore CPU
GPU, Accelerators
Ipomoea-01 25PB
Ipomoea-03
Ipomoea-02



Simulation Nodes
(Odyssey)

Data/Learning Nodes
(Aquarius)

東京大学 THE UNIVERSITY OF TOKYO
東京大学情報基盤センター INFORMATION TECHNOLOGY CENTER, THE UNIVERSITY OF TOKYO

Oakbridge-CX (OBCX) (Fujitsu)

- Intel Xeon Cascade Lake
- July 2019-September 2023
- 6.61 PF, #129 in 69th TOP500

Oakbridge-CX

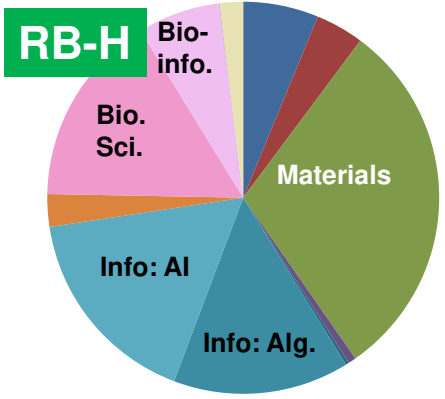
Wisteria/BDEC-01 (Fujitsu)

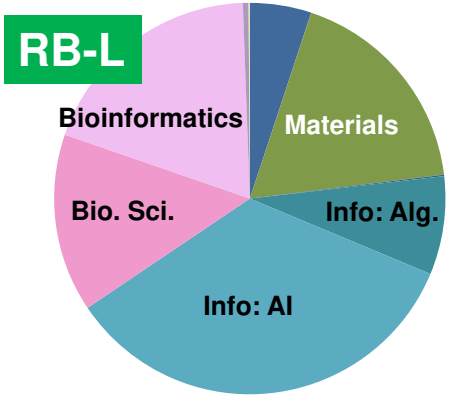
- **Simulation Nodes (Oddysey): A64FX (#23)**
- **Data/Learning Nodes (Aquarius) : Icelake + A100 (#125)**
- 33.1 PF, May 2021-April 2027
- Platform for Integration of "Simulation+Data+Learning (S+D+L)"
- Innovative Software Platform "h3-Open-BDEC" supported by Japanese Government (JSPS Grant-in-Aid for Scientific Res. (S) FY.2019-2023)

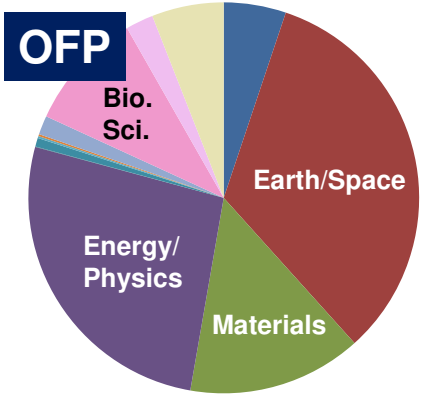
Wisteria BDEC-01

h3-Open-BDEC
Hierarchical, Hybrid, Heterogeneous
Big Data & Extreme Computing

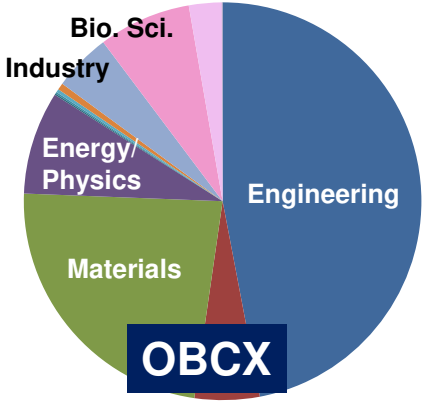
Research Area based on CPU Hours (FY.2020)







- Engineering
- Earth/Space
- Material
- Energy/Physics
- Info. Sci. : System
- Info. Sci. : Algrorithms
- Info. Sci. : AI
- Education
- Industry
- Bio
- Bioinformatics
- Social Sci. & Economics
- Data

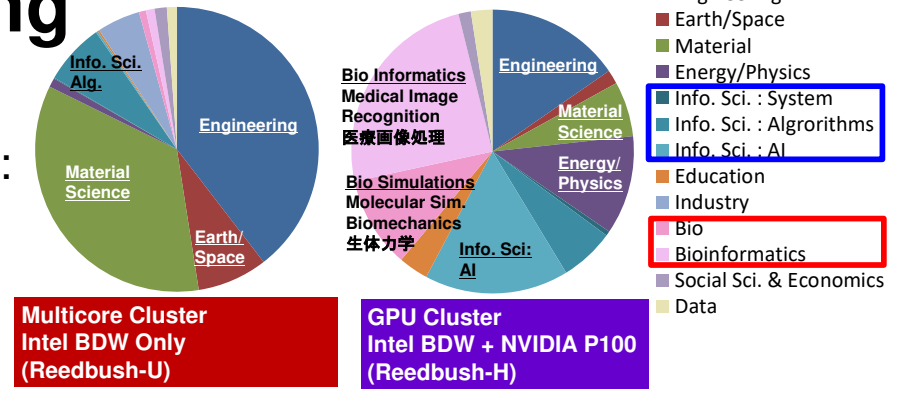


CPU

GPU

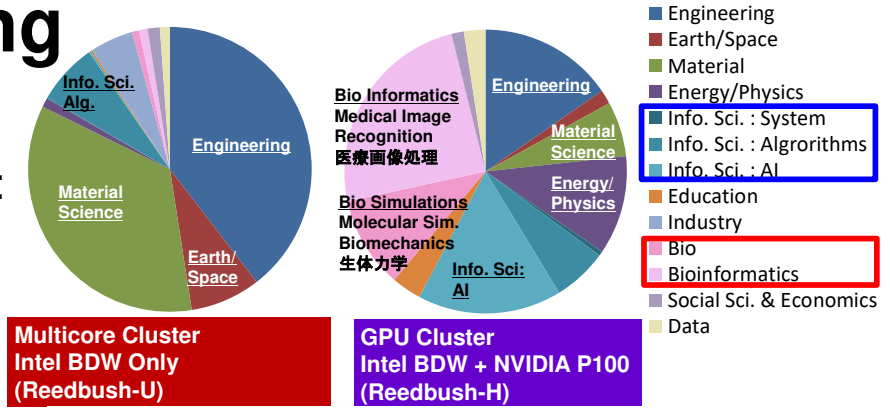
Future of Supercomputing

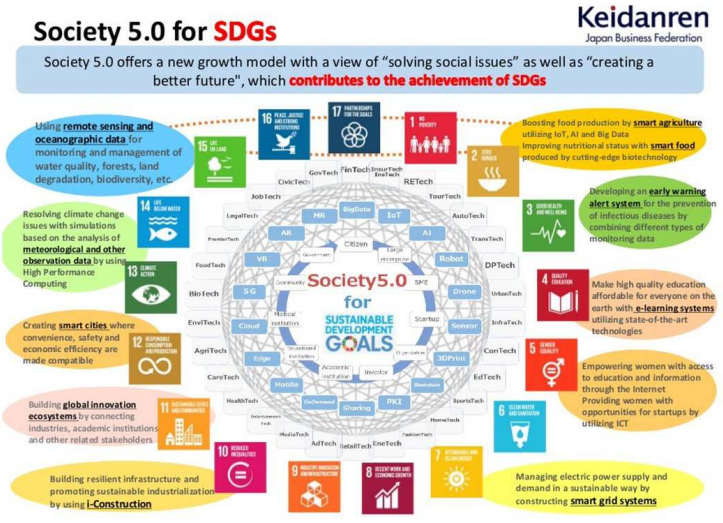
- Various Types of Workloads
  - Computational Science & Engineering: Simulations
  - Big Data Analytics
  - AI, Machine Learning …



Future of Supercomputing

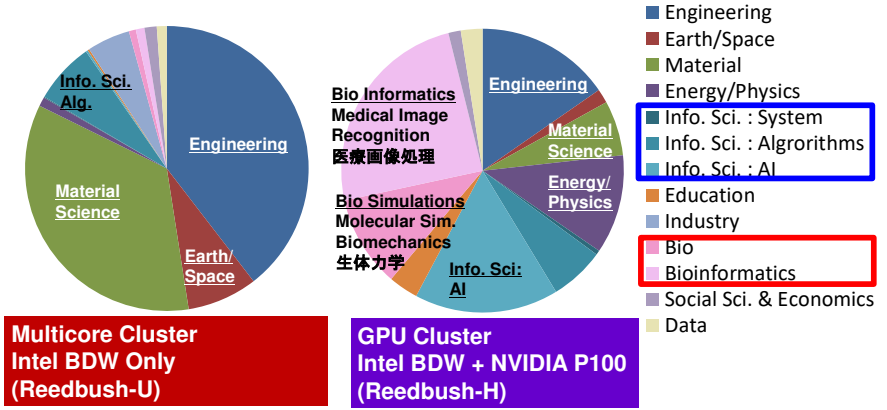
- Various Types of Workloads
  - Computational Science & Engineering: Simulations
  - Big Data Analytics
  - AI, Machine Learning …
- Integration/Convergence of (Simulation + Data + Learning) (S+D+L) is important towards Society 5.0 proposed by Japanese Government
  - Super Smart & Human-centered Society by Digital Innovation (IoT, Big Data, AI etc.) and by Integration of Cyber Space & Physical Space





Future of Supercomputing

- Various Types of Workloads
  - Computational Science & Engineering: Simulations
  - Big Data Analytics
  - AI, Machine Learning …
- **Integration/Convergence of (Simulation + Data + Learning) (S+D+L) is important towards Society 5.0**
- **BDEC (Big Data & Extreme Computing)**
  - **Platform for Integration of (S+D+L)**
  - **Focusing on S (Simulation)**
    - **AI for HPC, AI for Science, Digital Twins**
  - **Planning started in 2015**



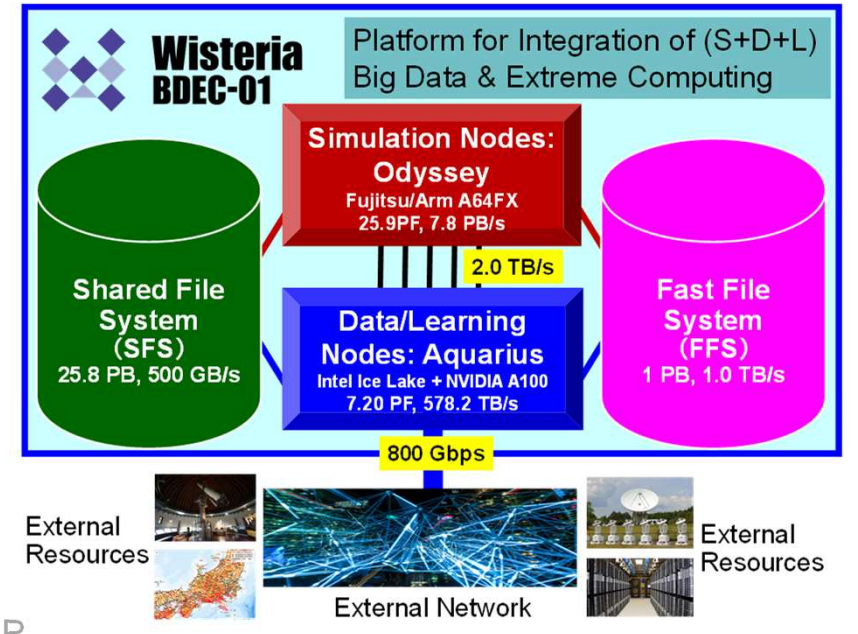
BDEC (Big Data & Extreme Computing)

S + D + L

Wisteria/BDEC-01

The 1st BDEC System (Big Data & Extreme Computing) Platform for Integration of (S+D+L)

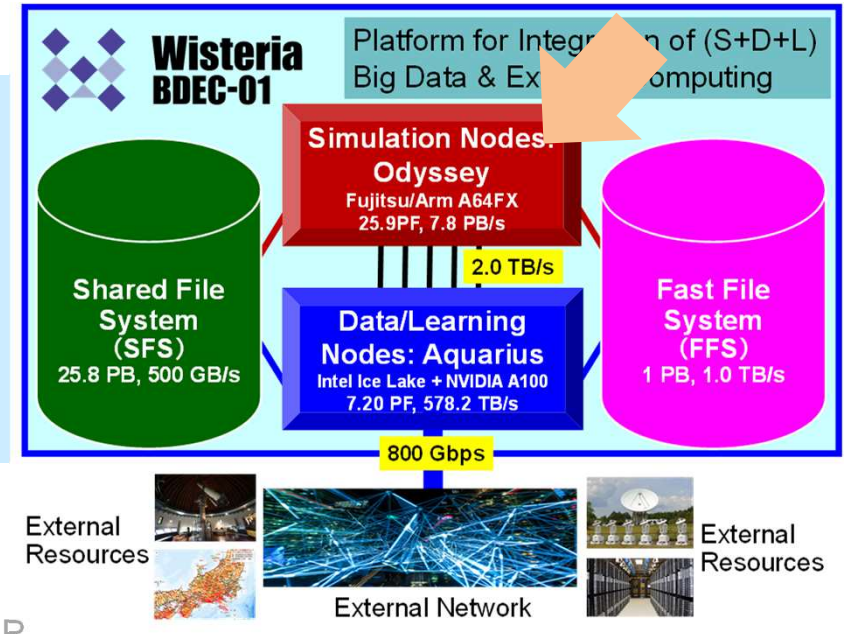
- Operation starts on May 14, 2021
- 33.1 PF, 8.38 PB/sec by **Fujitsu**
  - ~4.5 MVA with Cooling, ~360m^2
- 2 Types of Node Groups
  - Hierarchical, Hybrid, Heterogeneous (h3)
  - Simulation Nodes: Odyssey
    - Fujitsu PRIMEHPC FX1000 (A64FX), 25.9 PF
      - 7,680 nodes (368,640 cores), Tofu-D
      - General Purpose CPU + HBM
      - Commercial Version of "Fugaku"
  - Data/Learning Nodes: Aquarius
    - Data Analytics & AI/Machine Learning
    - Intel Xeon Ice Lake + NVIDIA A100, 7.2PF
      - 45 nodes (90x Ice Lake, 360x A100), IB-HDR
    - Some of the DL nodes are connected to external resources directly
- File Systems: SFS (Shared/Large) + FFS (Fast/Small)



Wisteria/BDEC-01

The 1st BDEC System (Big Data & Extreme Computing) Platform for Integration of (S+D+L)

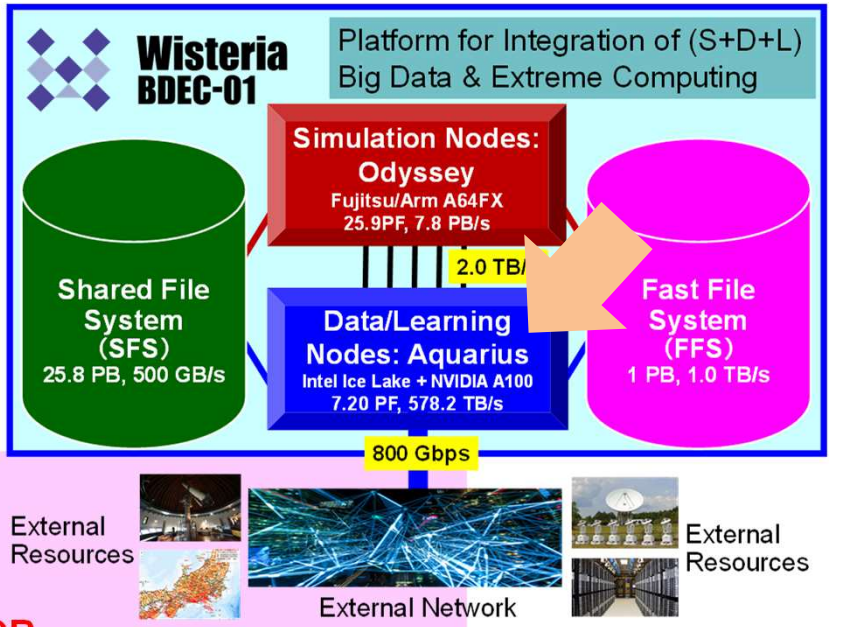
- Operation starts on May 14, 2021
- 33.1 PF, 8.38 PB/sec by **Fujitsu**
  - ~4.5 MVA with Cooling, ~360m^2
- **2 Types of Node Groups**
  - Hierarchical, Hybrid, Heterogeneous (h3)
  - **Simulation Nodes: Odyssey**
    - **Fujitsu PRIMEHPC FX1000 (A64FX), 25.9 PF**
      - **7,680 nodes (368,640 cores), Tofu-D**
      - **General Purpose CPU + HBM**
      - **Commercial Version of "Fugaku"**
  - Data/Learning Nodes: Aquarius
    - Data Analytics & AI/Machine Learning
    - Intel Xeon Ice Lake + NVIDIA A100, 7.2PF
      - 45 nodes (90x Ice Lake, 360x A100), IB-HDR
    - Some of the DL nodes are connected to external resources directly
- File Systems: SFS (Shared/Large) + FFS (Fast/Small)



Wisteria/BDEC-01

The 1st BDEC System (Big Data & Extreme Computing) Platform for Integration of (S+D+L)

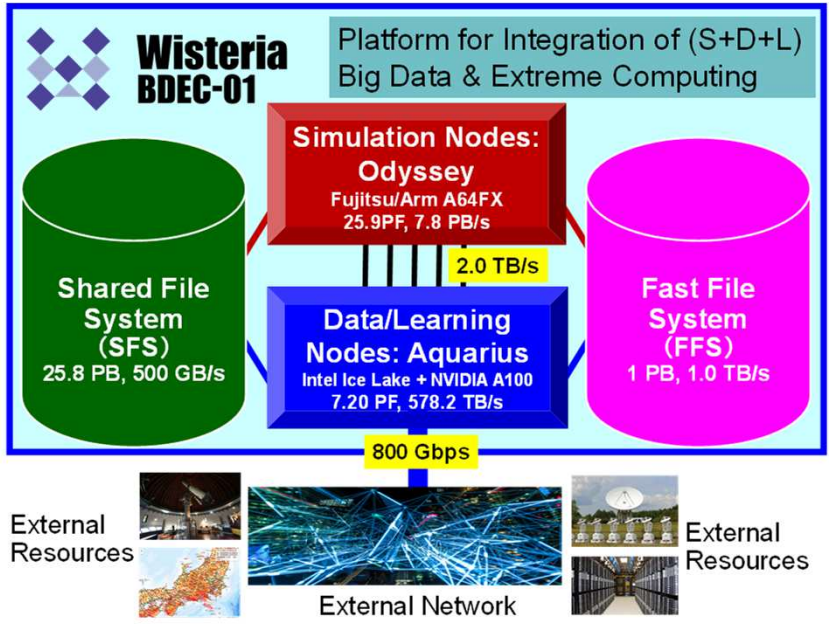
- Operation starts on May 14, 2021
- 33.1 PF, 8.38 PB/sec by **Fujitsu**
  - ~4.5 MVA with Cooling, ~360m^2
- **2 Types of Node Groups**
  - Hierarchical, Hybrid, Heterogeneous (h3)
  - **Simulation Nodes: Odyssey**
    - **Fujitsu PRIMEHPC FX1000 (A64FX), 25.9 PF**
      - **7,680 nodes (368,640 cores), Tofu-D**
      - **General Purpose CPU + HBM**
      - **Commercial Version of "Fugaku"**
  - **Data/Learning Nodes: Aquarius**
    - **Data Analytics & AI/Machine Learning**
    - **Intel Xeon Ice Lake + NVIDIA A100, 7.2PF**
      - **45 nodes (90x Ice Lake, 360x A100), IB-HDR**
    - **Some of the DL nodes are connected to external resources directly**
- File Systems: SFS (Shared/Large) + FFS (Fast/Small)



Rankings@SC22
November 2022

	Odyssey	Aquarius
TOP 500	23	125
Green 500	45	28
HPCG	12	68
Graph 500 BFS	4	-
HPL-MxP (HPL-AI)	10*	-

*) ISC 2022 (June 2022)



Simulation Nodes
Odyssey
25.9 PF, 7.8 PB/s

Fast File System (FFS)
1.0 PB, 1.0 TB/s

Shared File System (SFS)
25.8 PB, 0.50 TB/s

Data/Learning Nodes
Aquarius
7.20 PF, 578.2 TB/s

Simulation Codes

Optimized Models & Parameters

Simulation Nodes Odyssey

Results

Wisteria/BDEC-01

Machine Learning, DDA

Observation Data

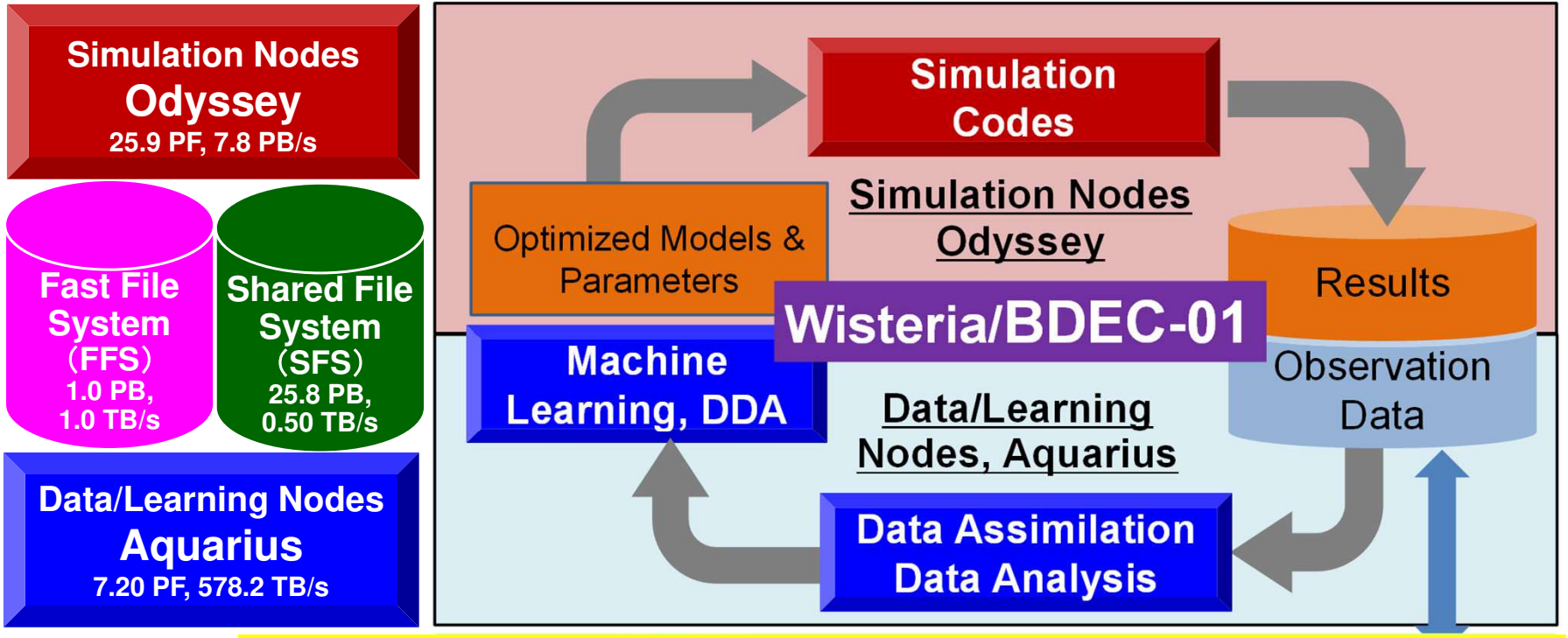
Data/Learning Nodes, Aquarius

Data Assimilation Data Analysis

Server, Storage, DB, Sensors, etc.

External Network

External Resources

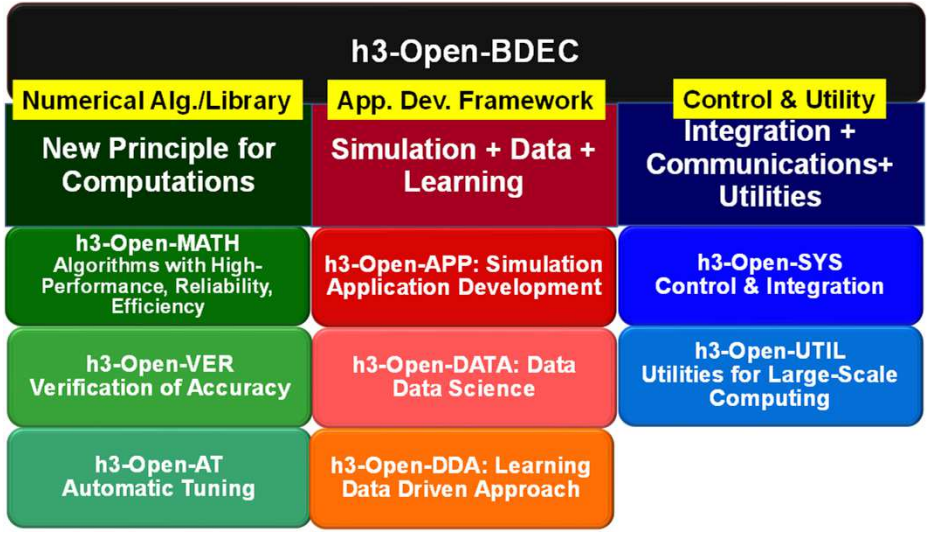


Optimization of Models/Parameters for Simulations by Data Analytics & Machine Learning (S+D+L)

- Wisteria/BDEC-01
- **h3-Open-BDEC**
- Earthquake Simulations combined with Real-Time Data Assimilation
  - h3-Open-SYS/WaitIO
- HPC-AI Workload: Global Cloud Simulations
  - h3-Open-UTIL/MP
- How to run such (a little bit complicated) workflows
  - Wisteria/BDEC-01
  - Wisteria/BDEC-01 + mdx
- Future/Ongoing Works
  - International Collaborations
  - OFP-II

h3-Open-BDEC Innovative Software Platform for Integration of (S+D+L) on the BDEC System, such as Wisteria/BDEC-01

- 5-year project supported by Japanese Government (JSPS) since 2019
- Leading-PI: Kengo Nakajima (The University of Tokyo)
- Total Budget: 1.41M USD



Members (Co-PI's) of h3-Open-BDEC Project

Computer Science, Computational Science, Numerical Algorithms, Data Science, Machine Learning

- Kengo Nakajima (ITC/U.Tokyo, RIKEN), Leading-PI
- Takeshi Iwashita (Hokkaido U), Co-PI, Algorithms
- Hisashi Yashiro (NIES), Co-PI, Coupling, Utility
- Hiromichi Nagao (ERI/U.Tokyo), Co-PI, Data Assimilatio
- Takashi Shimokawabe (ITC/U.Tokyo), Co-PI, ML/hDDA
- Takeshi Ogita (TWCU), Co-PI, Accuracy Verification
- Takahiro Katagiri (Nagoya U), Co-PI, Appropriate Computing
- Hiroya Matsuba (ITC/U.Tokyo, Hitachi), Co-PI, Container

Hierarchical, Hybrid, Heterogeneous
h3-Open-BDEC
Big Data & Extreme Computing

東京大学
THE UNIVERSITY OF TOKYO

北海道大学
HOKKAIDO UNIVERSITY

国立研究開発法人
国立環境研究所
National Institute for Environmental Studies

Tokyo Woman's Christian University
東京女子大学

名古屋大学
NAGOYA UNIVERSITY

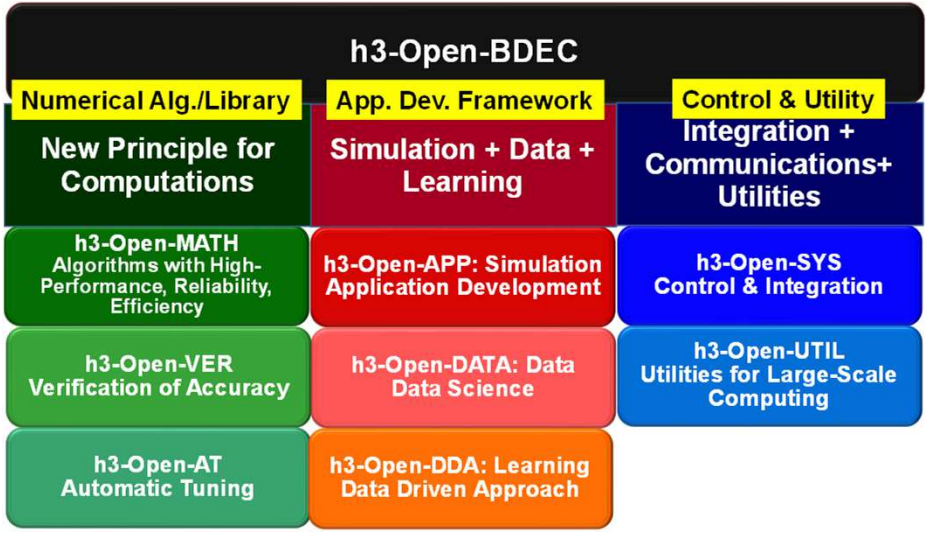
HITACHI

RIKEN

R-CCS

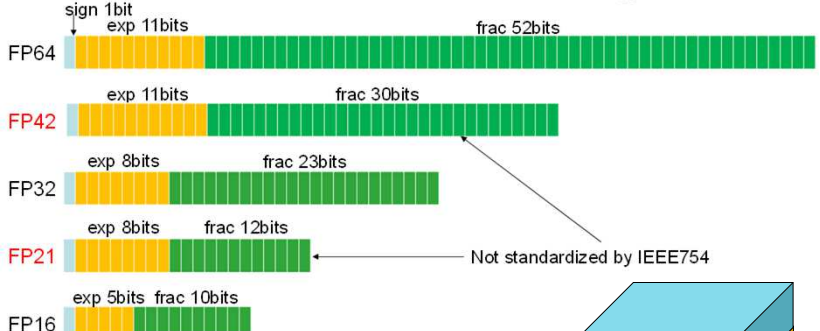
h3-Open-BDEC Innovative Software Platform for Integration of (S+D+L) on the BDEC System, such as Wisteria/BDEC-01

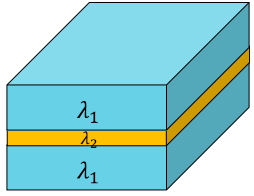
- "Three" Innovations
  - New Principles for Numerical Analysis by Adaptive Precision, Automatic Tuning & Accuracy Verification
  - Integration of (S+D+L) by Hierarchical Data Driven Approach (*h*DDA)
  - Software & Utilities for Heterogenous Environment, such as Wisteria/BDEC-01



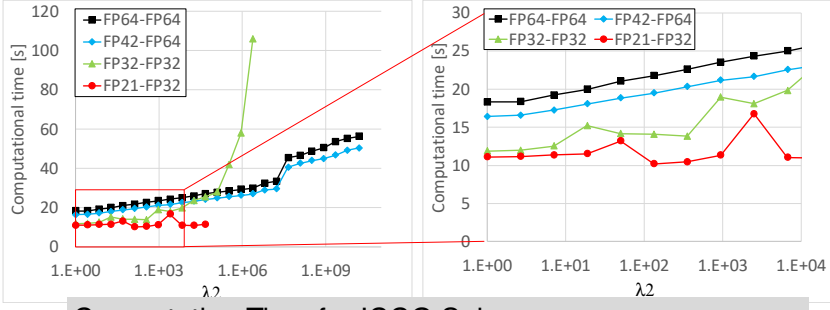
Adaptive Precision Computing with FP21/FP42

Masatoshi Kawai (kawai@cc.u-tokyo.ac.jp)





Heat Conduction with
Heterogeneous Material Property

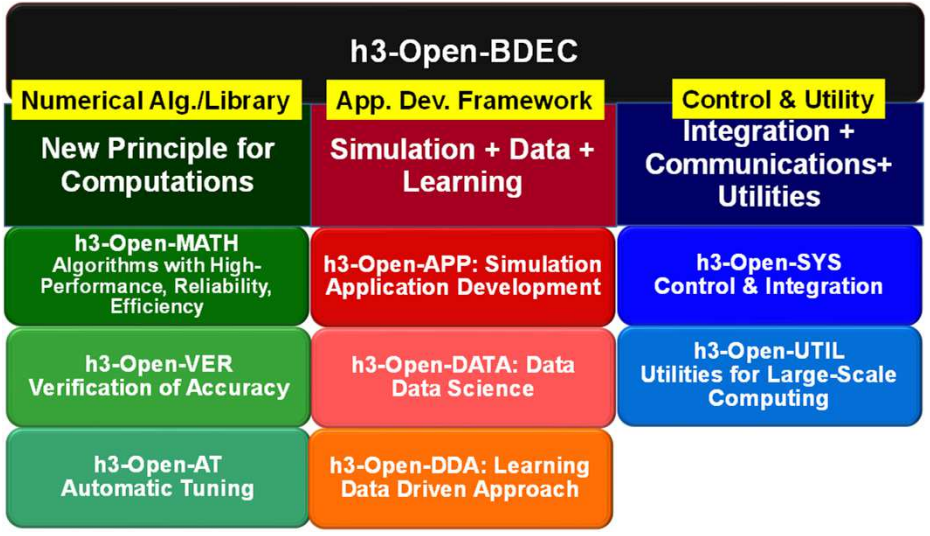


Computation Time for ICCG Solver
Various Types of Precisions on Intel Xeon Cascadelake

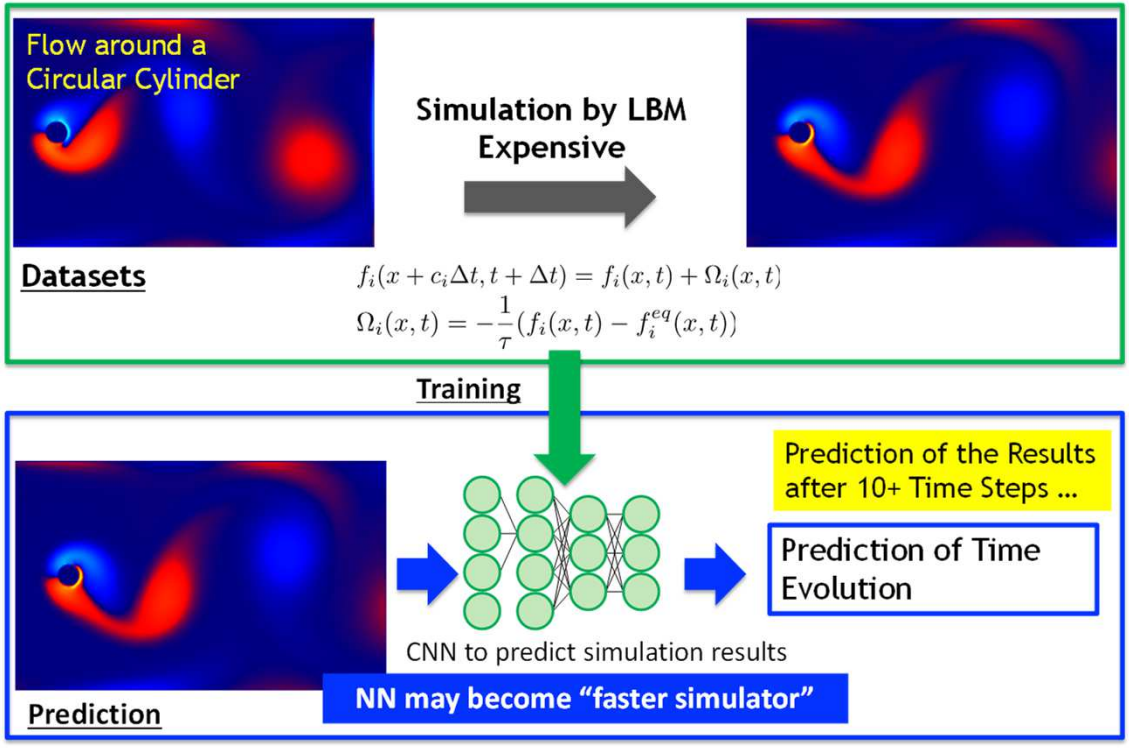
In recent years, the usefulness of low-precision floating-point representation has been studied in various fields such as machine learning. Low accuracy can be expected to have effects such as shortening calculation time and reducing power consumption. For example, in an application with a memory bandwidth 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. Developed method has been evaluated for ICCG solver, which solves linear equations derived from 3D FVM code for steady-state head conduction with heterogeneous material property ($\lambda_1=10^0$, $\lambda_2=10^0$~10^9). Generally, computation with lower precision (e.g. FP32-FP32, FP21-FP32) becomes unstable, if condition number of the coefficient matrix is larger (λ_2 is larger), FP21-FP32 provides the best performance if λ_2 is up to 10^4. (“FP21-FP32” means “matrices are in FP21, and vectors are in FP32)

h3-Open-BDEC Innovative Software Platform for Integration of (S+D+L) on the BDEC System, such as Wisteria/BDEC-01

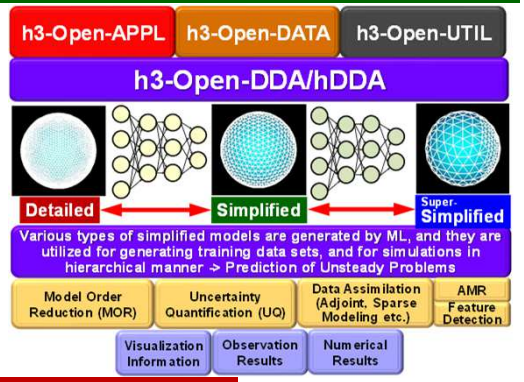
- "Three" Innovations
  - New Principles for Numerical Analysis by Adaptive Precision, Automatic Tuning & Accuracy Verification
  - Integration of (S+D+L) by Hierarchical Data Driven Approach (*h*DDA)
  - Software & Utilities for Heterogenous Environment, such as Wisteria/BDEC-01



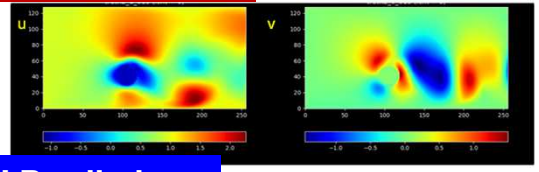
Acceleration of Transient CFD Simulations using ML/CNN Integration of (S+D+L), AI for HPC/AI for Science



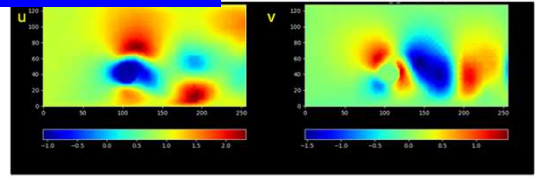
[c/o Takashi Shimokawabe (ITC/U.Tokyo)]



Simulations: LBM

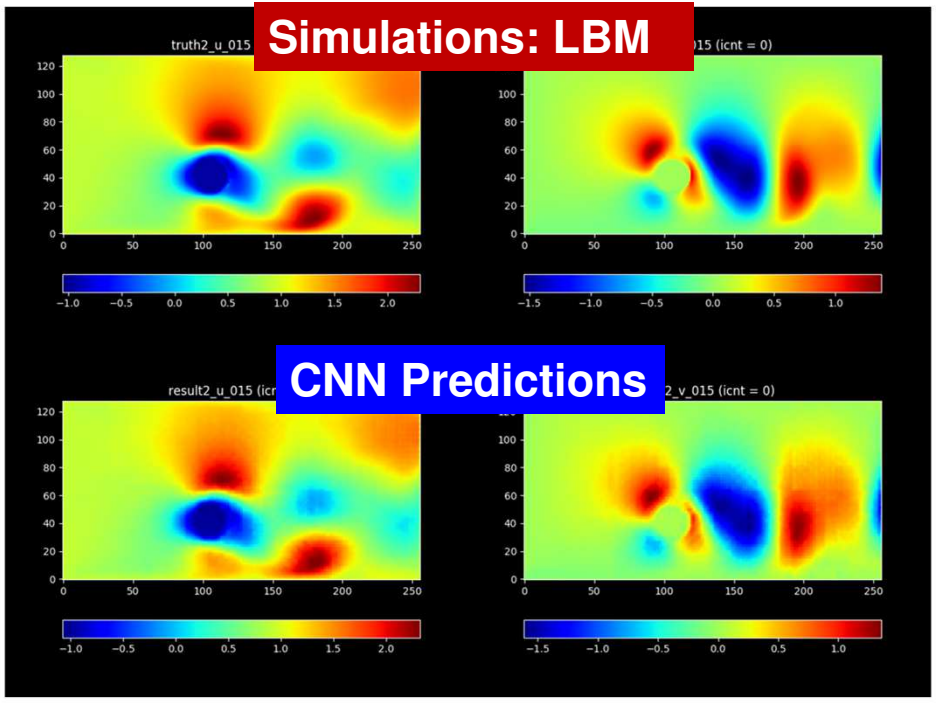


CNN Predictions



Prediction of CFD Simulation by ML/CNN

Takashi Shimokawabe (shimokawabe@cc.u-tokyo.ac.jp)

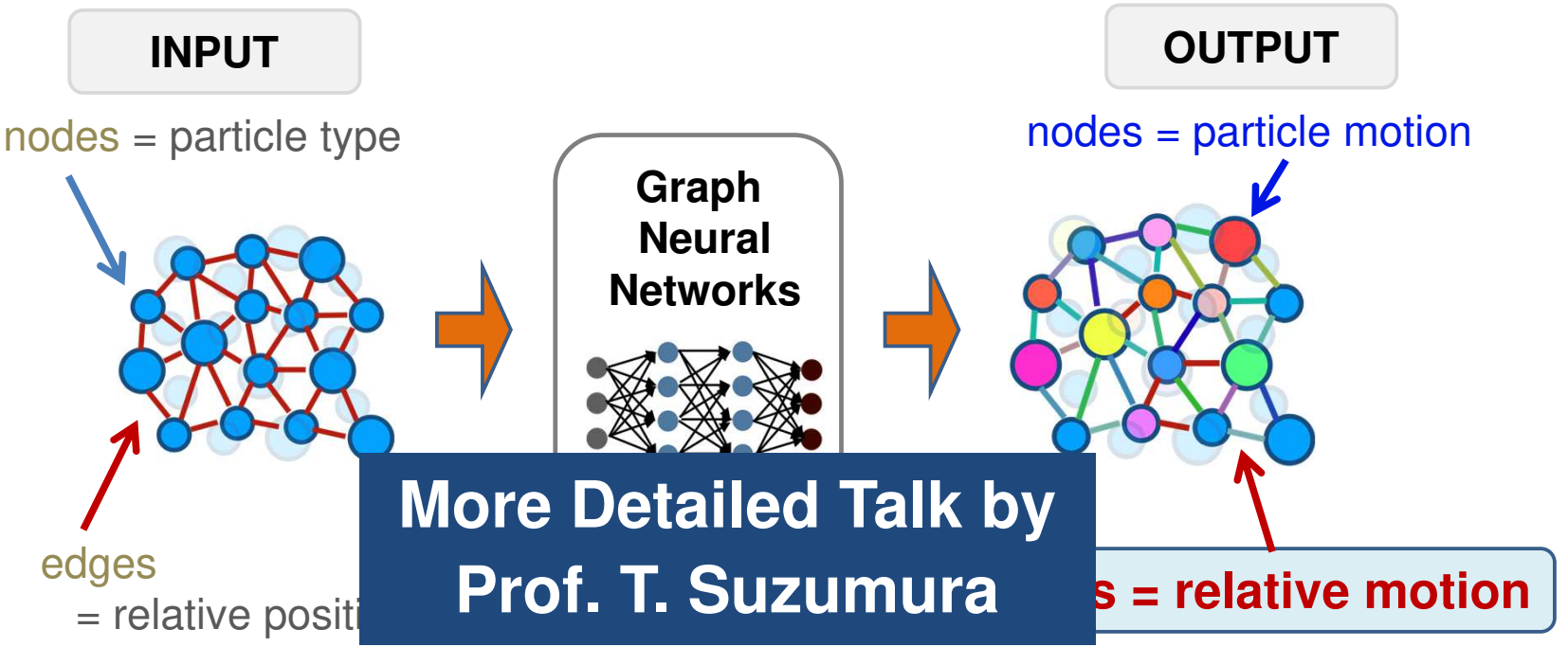


Comparison of the flow velocity results obtained by the conventional simulation (upper) and the prediction of these results by deep learning (lower)

Computational fluid dynamics (CFD) is widely used in science and engineering. However, since CFD simulations requires 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.

Machine learning slow molecular dynamics

Our proposal — **BOnd Targeting Network (BOTAN)**

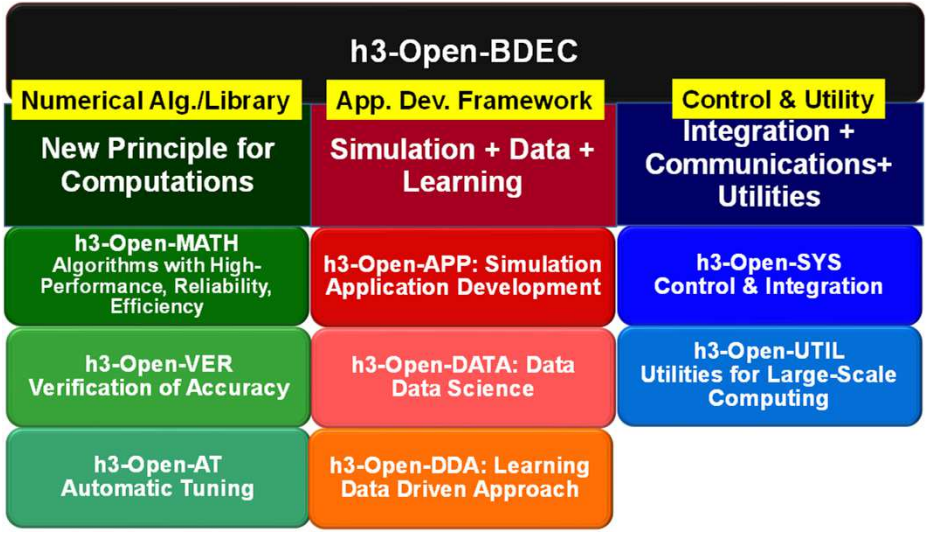


H. Shiba, M. Hanai, T. Suzumura, and T. Shimokawabe, arXiv:2206.14024 (2022)

- Wisteria/BDEC-01
- h3-Open-BDEC
- **Earthquake Simulations combined with Real-Time Data Assimilation**
  - **h3-Open-SYS/WaitIO**
- HPC-AI Workload: Global Cloud Simulations
  - h3-Open-UTIL/MP
- How to run such (a little bit complicated) workflows
  - Wisteria/BDEC-01
  - Wisteria/BDEC-01 + mdx
- Future/Ongoing Works
  - International Collaborations
  - OFP-II

h3-Open-BDEC Innovative Software Platform for Integration of (S+D+L) on the BDEC System, such as Wisteria/BDEC-01

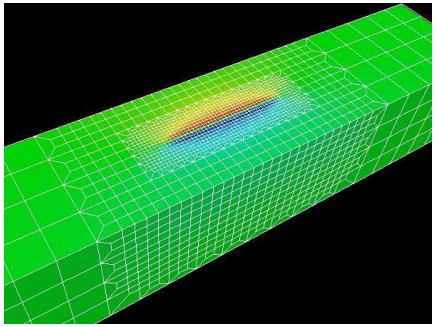
- "Three" Innovations
  - New Principles for Numerical Analysis by Adaptive Precision, Automatic Tuning & Accuracy Verification
  - Integration of (S+D+L) by Hierarchical Data Driven Approach (*h*DDA)
  - Software & Utilities for Heterogenous Environment, such as Wisteria/BDEC-01

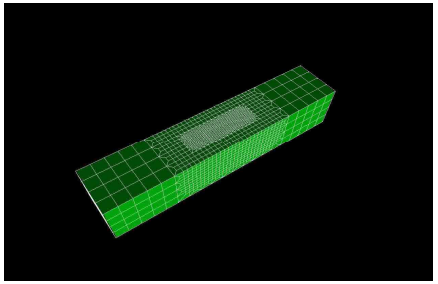


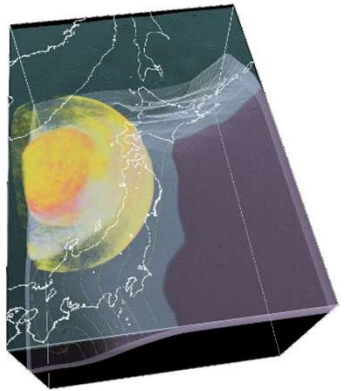
Earthquake Simulations

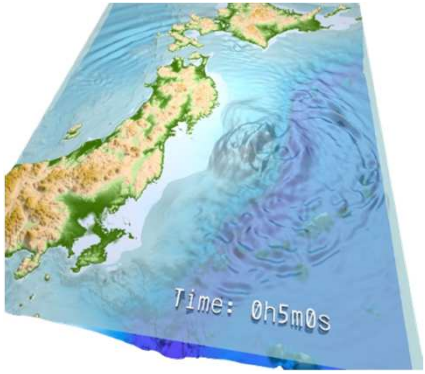
Three Categories, Experts in Each

- Earthquake Generation Cycle
  - Stress accumulation at plate boundaries/faults
  - $O(10^1)$~$O(10^2)$ years
  - BEM, BIM
- Dynamic Rupture
  - Stress Accumulation⇒Rupture
    - Estimation of source terms, very local
  - $O(10^1)$~$O(10^2)$ seconds
  - BEM, BIM
- **Seismic Wave Propagation**
  - **Strong Ground Motion**
  - **$O(10^1)$~$O(10^2)$ seconds**
  - **FEM, FDM**







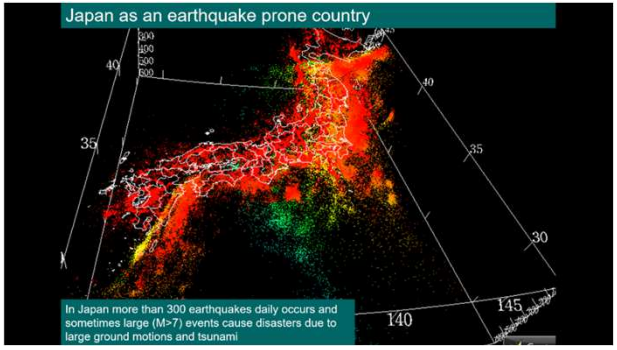


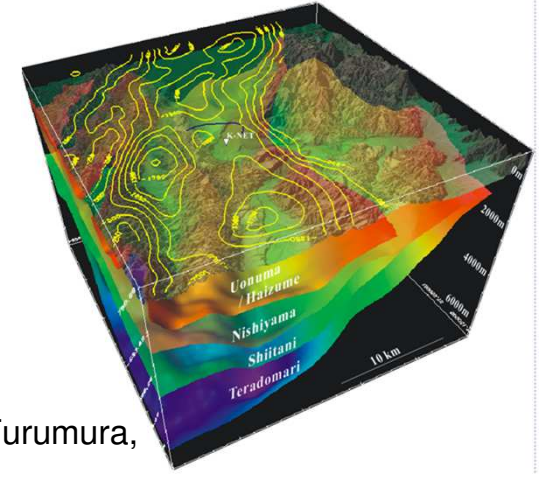
Early Forecast of Long-Period Ground Motions via Data Assimilation of Observation and Simulations [Furumura et al. 2017]

- New method for the early forecast of long-period (> 3–10 s) ground motions generated by large earthquakes based on the data assimilation of observed ground motions and FDM simulations of seismic wave propagation in a 3-D heterogeneous structure (**Seism3D/OpenSWPC-DAF（Data-Assimilation-Based Forecast）**).
- **This approach uses the dense nationwide network in Japan and supercomputers to perform forecasts using the assimilated wavefields at speeds much faster than the actual wave propagation speed.**
- **An early alert can be issued prior to the occurrence of strong motions due to large, distant earthquakes.**
- **This research inspired me to develop a system like Wisteria/BDEC-01, where (Simulation, Data, Learning) are integrated on a single system.**

Earthquake simulation is always with uncertainty

- Subsurface/Underground Structure
  - Heterogenous, Random, Stochastic
  - Fluctuations
- Traditional Simulations
  - Forward Simulations
- **Integration of Simulation/Observation is essential**
- **New Types of Methods for Simulations combined with Data Assimilation/Real-Time Observation is under development**
  - **Forecast by Simulations, Correction by Data Assimilation**

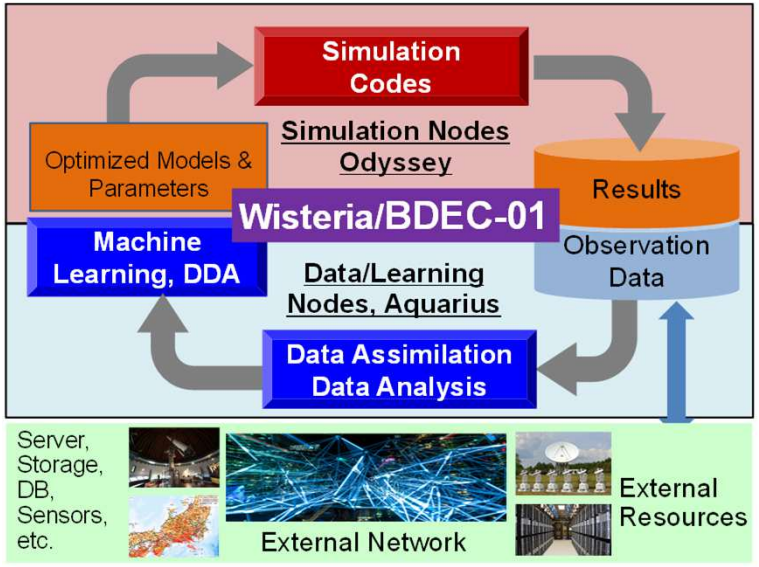


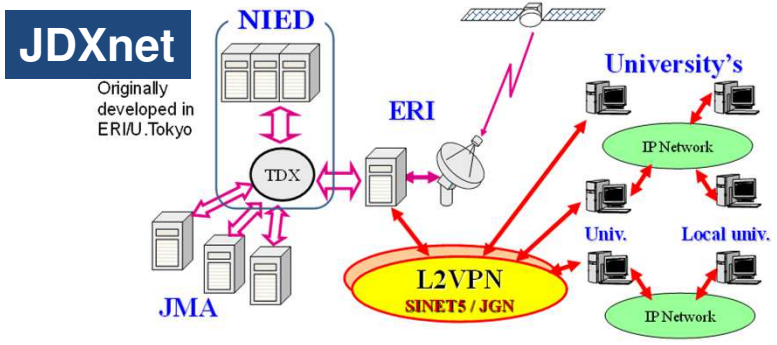


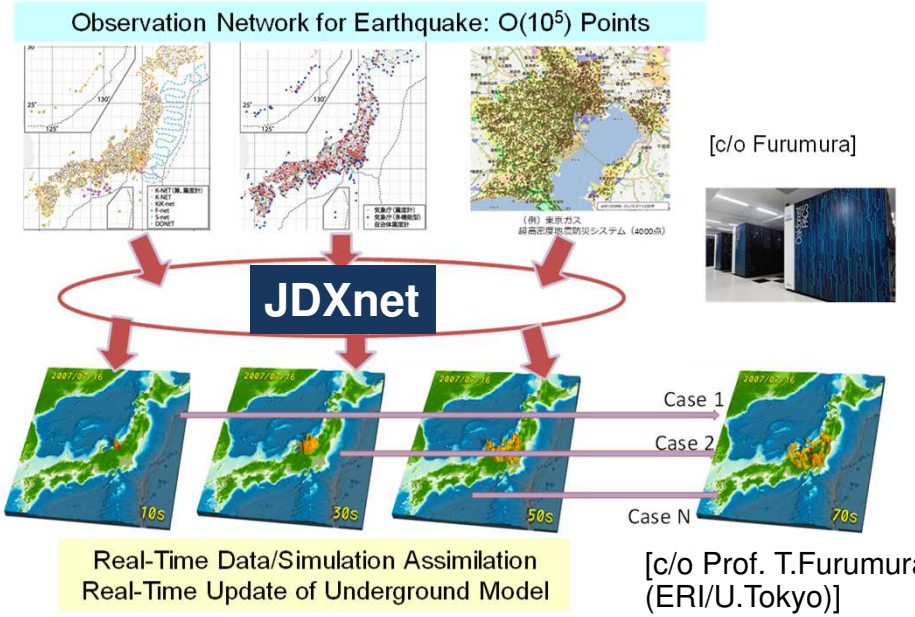
[c/o Prof. T. Furumura, ERI/U.Tokyo]

3D Earthquake Simulation with Real-Time Data Observation/Assimilation

Simulation of Strong Motion (Wave Propagation) by 3D FDM





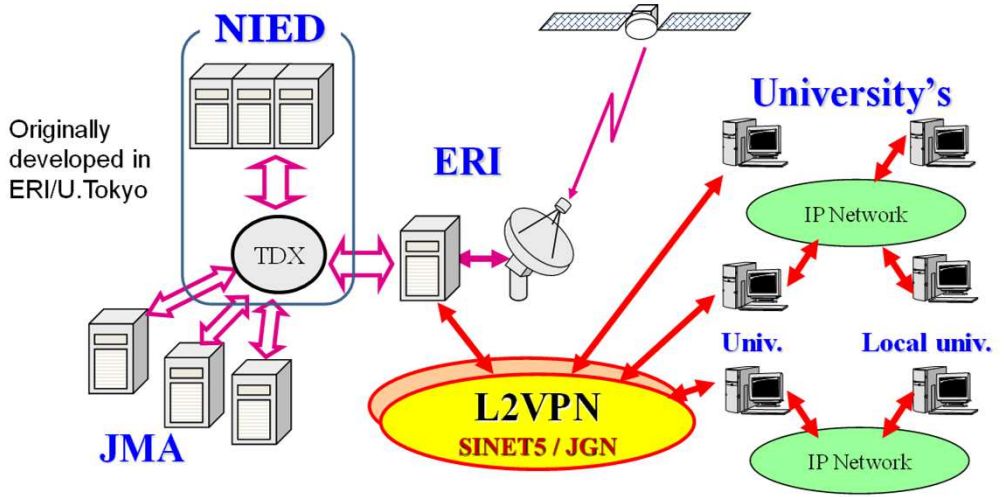


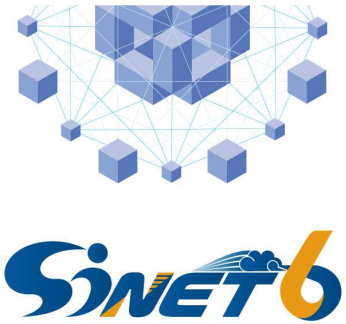
[c/o Prof. T.Furumura (ERI/U.Tokyo)]

Real-Time Sharing of Seismic Observation is possible in Japan by JDXnet with SINET

Japan Data eXchange network

- Seismic Observation Data (100Hz/3-dir's/O(10^3) observation points) by JDXnet is available through SINET in Real Time
  - O(10^2) GB/day: available at Website of NIED
  - O(10^5) pts in future including stations operated by industry





[c/o Prof. H.Tsuruoka (ERI/U.Tokyo)]

Real-Time Assimilation of "Observation+Computation" in Seismic Wave Propagation [c/o Oba & Furumura]

- **Data Assimilation of Wave Propagation by "Optimal Interpolation Technique"**

Assim. Comp. Residual Obs. Comp.

$$x_n^a = x_n^f + W(y_n - Hx_n^f)$$

Comp. Assim.

$$x_{n+1}^f = Fx_n^a$$

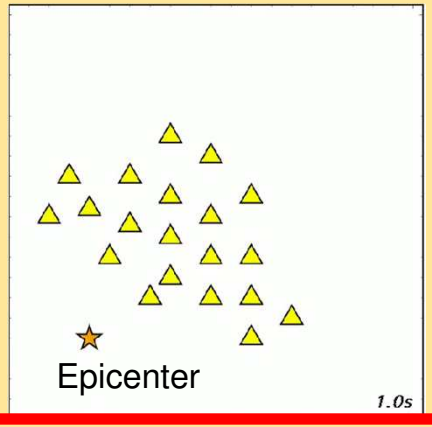
n: Time Step

W: **Weighting Matrix**

F: Wave Propagation simulation

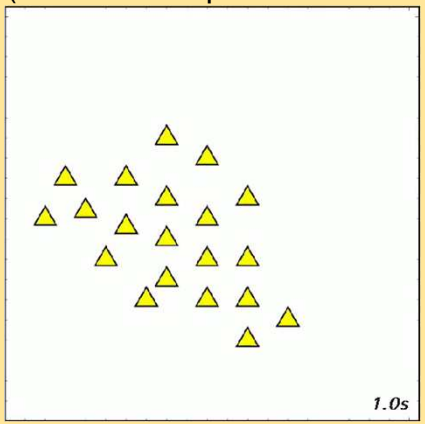
(A) Pure Simulation

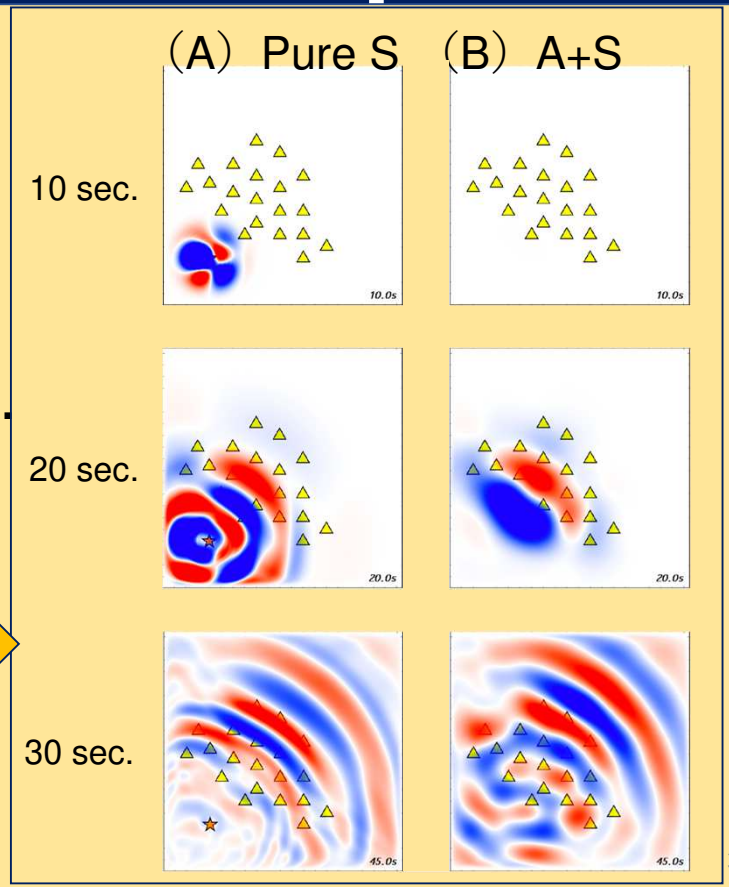
△: Obs. Pts.



(B) Assimilation+Sim.

(No info for Epicenter needed)





Real-Time Assimilation of "Observation+Computation" in Seismic Wave Propagation [c/o Oba & Furumura]

- **Data Assimilation of Wave Propagation by "Optimal Interpolation Technique"**

$$x_n^a = x_n^f + W(y_n - Hx_n^f)$$

Assim. Comp. Residual Obs. Comp.

$$x_{n+1}^f = Fx_n^a$$

Comp. Assim.

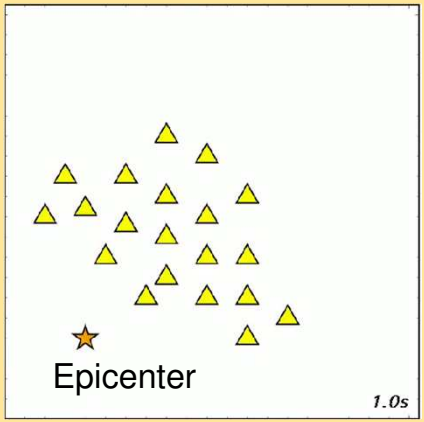
n: Time Step

W: **Weighting Matrix**

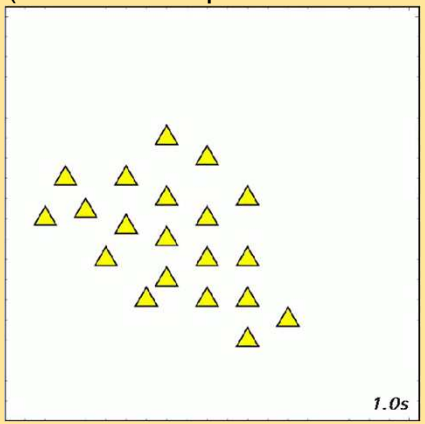
F: Wave Propagation simulation

(A) Pure Simulation

△: Obs. P s.



(B) Assimilation+Sim. (No info for Epicenter needed)

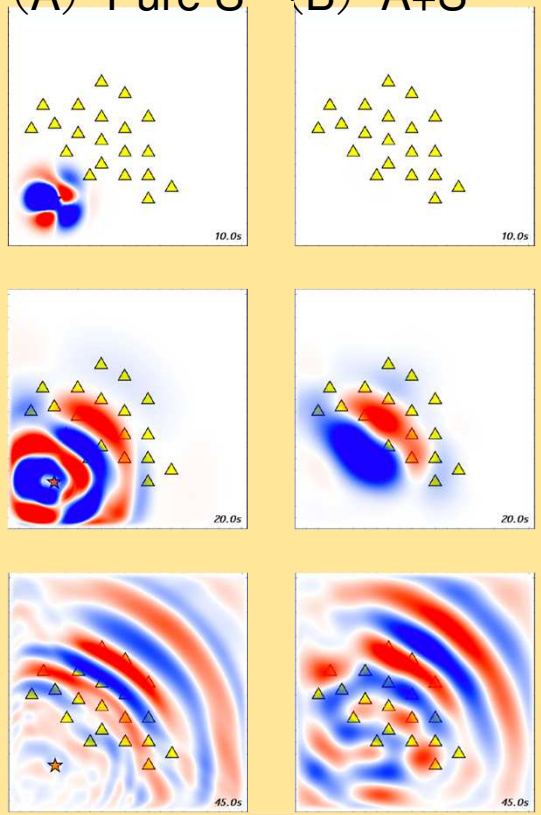


(A) Pure S (B) A+S

10 sec.

20 sec.

30 sec.



Starting from (A+S: Assim+Sim.) to (Pure S: Pure Simulation)

Residual

Assim. Comp. Obs. Comp.

$$x_n^a = x_n^f + W(y_n - Hx_n^f)$$

Comp. Assim.

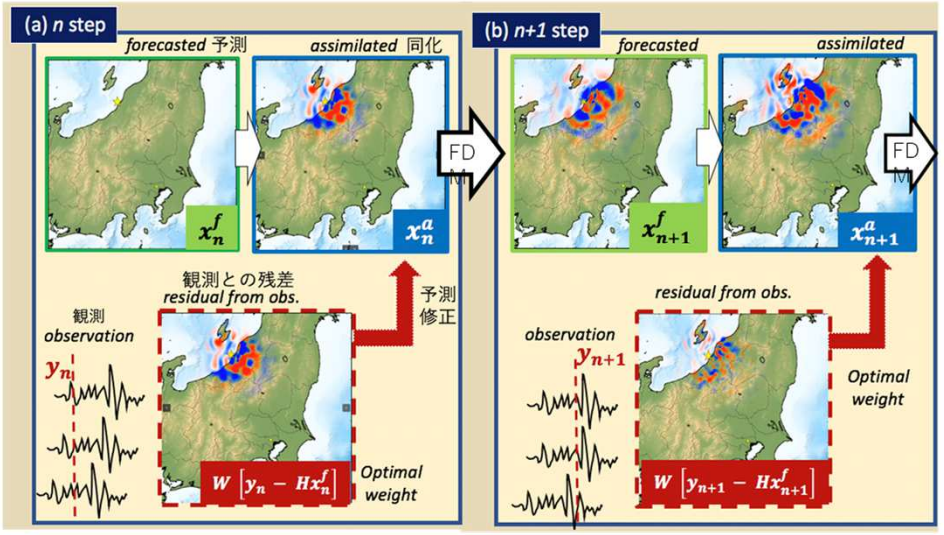
$$x_{n+1}^f = Fx_n^a$$

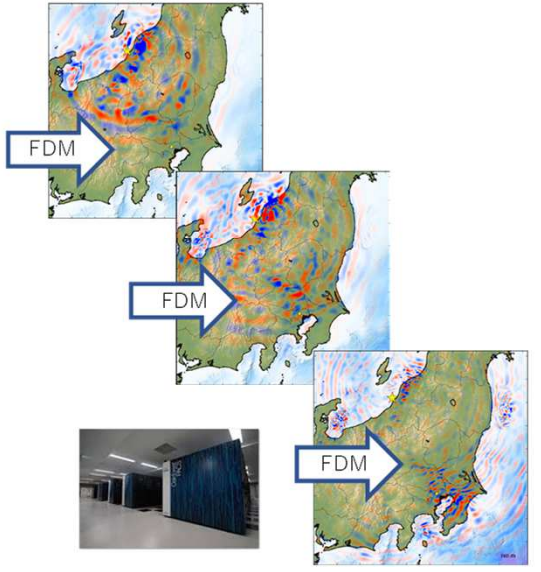
n: Time Step

W: **Weighting Matrix**

F: Wave Propagation simulation

(A+S) Assimilation+Simulation → **(Pure S) Pure Simulation/Forecast**

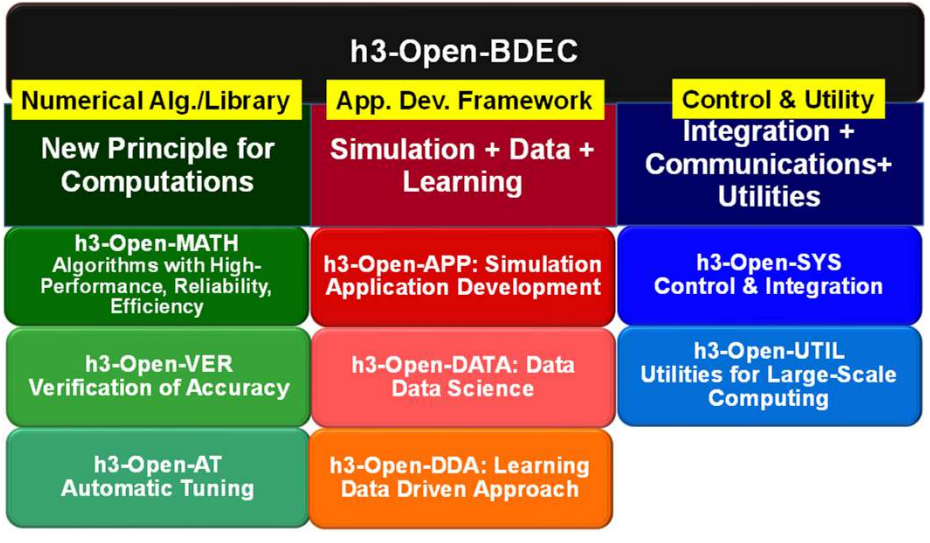




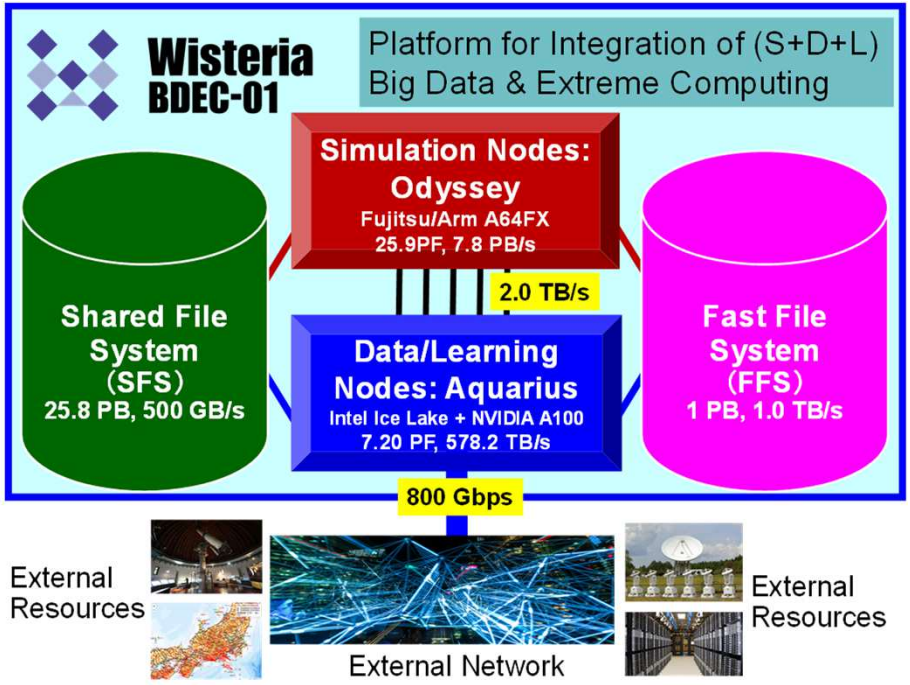
[c/o Prof. T. Furumura, ERI/U.Tokyo]

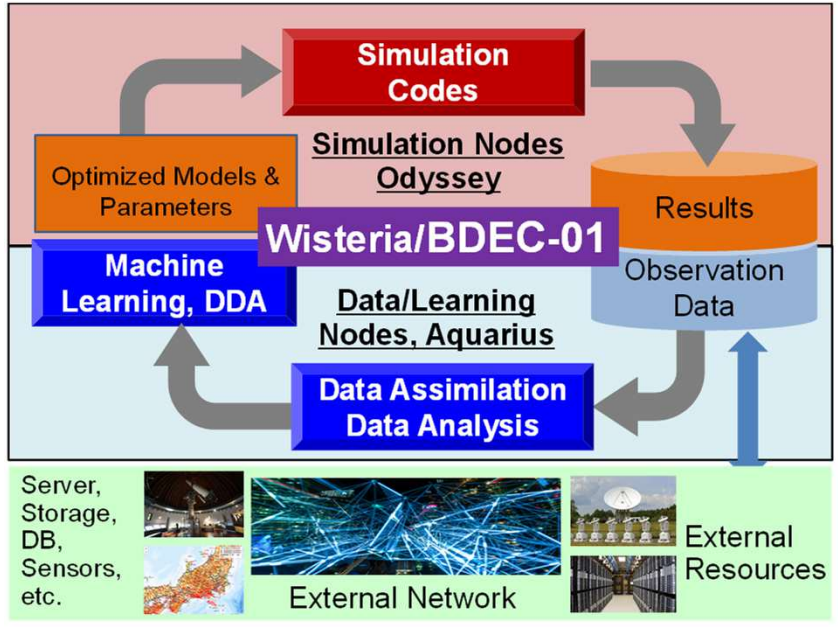
h3-Open-BDEC Innovative Software Platform for Integration of (S+D+L) on the BDEC System, such as Wisteria/BDEC-01

- "Three" Innovations
  - New Principles for Numerical Analysis by Adaptive Precision, Automatic Tuning & Accuracy Verification
  - Integration of (S+D+L) by Hierarchical Data Driven Approach (*h*DDA)
  - Software & Utilities for Heterogenous Environment, such as Wisteria/BDEC-01



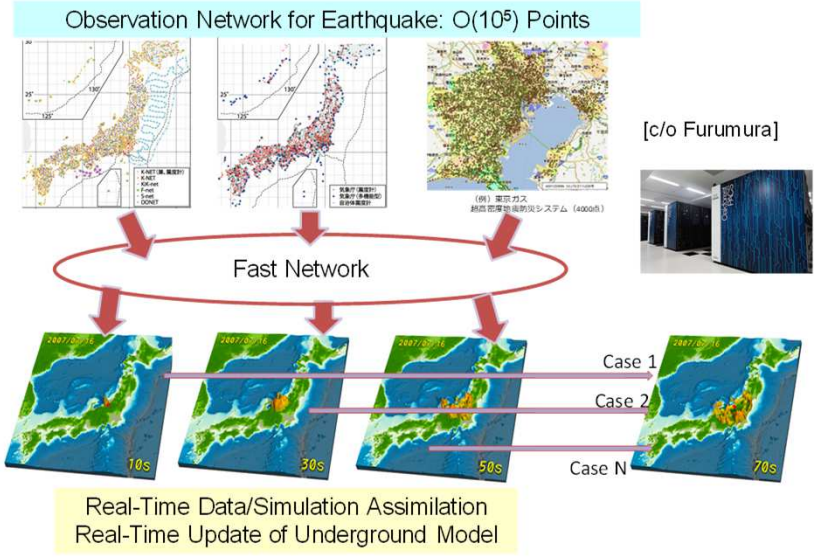
Wisteria/BDEC-01: The First "Really Heterogenous" System in the World



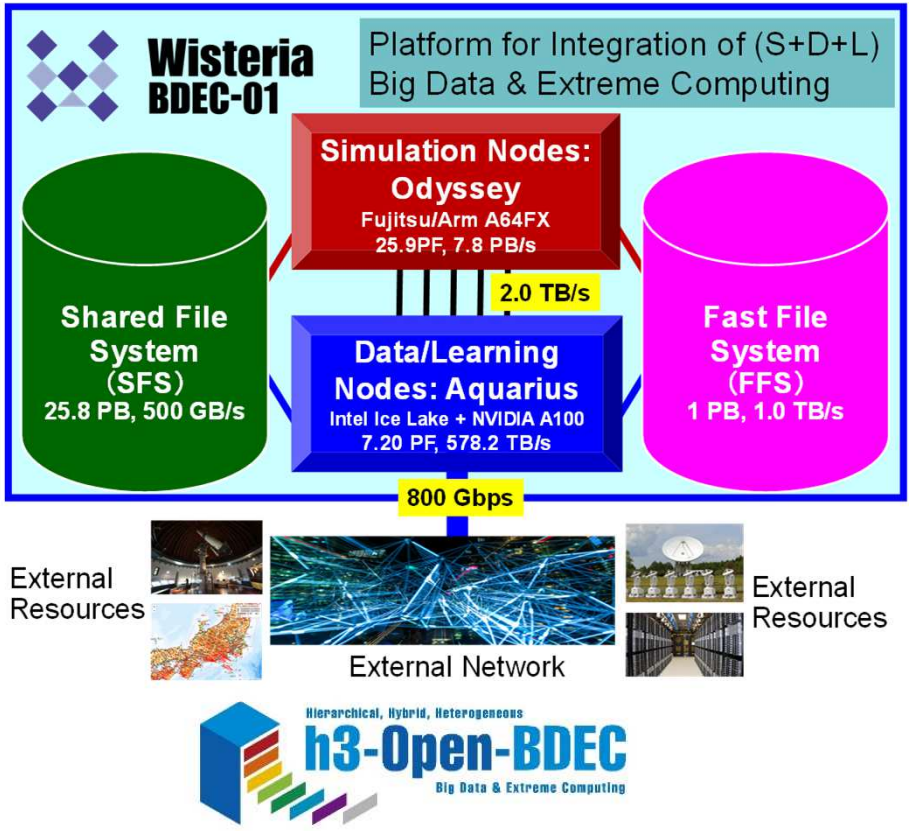


Computing on Wisteria/BDEC-01

- **Wisteria/BDEC-01**
  - **Aquarius (GPU: NVIDIA A100)**
    - **Filtering, Visualization**
  - **Odyssey (CPU: A64FX)**
    - **Data Assimilation, Simulation**

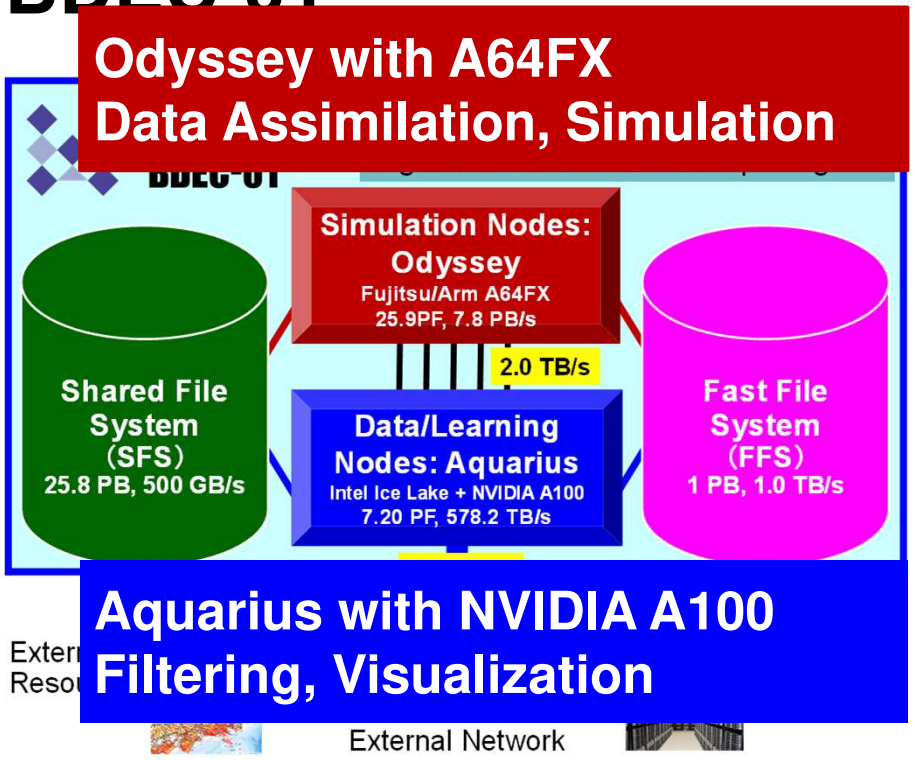


[c/o Furumura]



Computing on Wisteria/BDEC-01

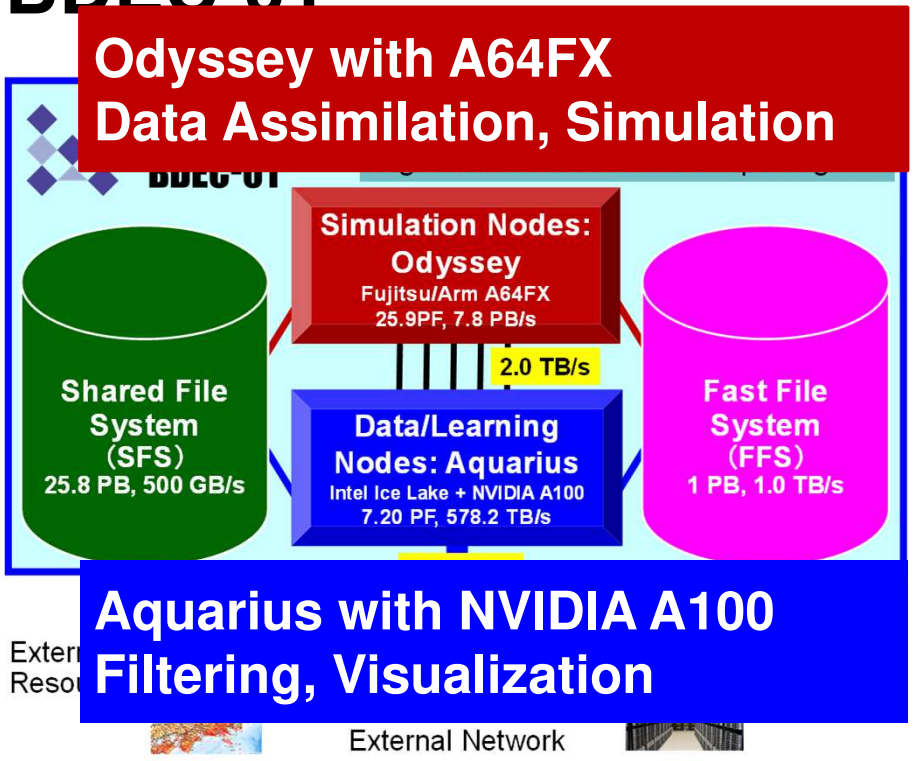
- **Wisteria/BDEC-01**
  - **Aquarius (GPU: NVIDIA A100)**
    - **Filtering, Visualization**
  - **Odyssey (CPU: A64FX)**
    - **Data Assimilation, Simulation**
- Combining Odyssey-Aquarius
  - Single MPI Job over O-A is impossible



Hierarchical, Hybrid, Heterogeneous
h3-Open-BDEC
Big Data & Extreme Computing

Computing on Wisteria/BDEC-01

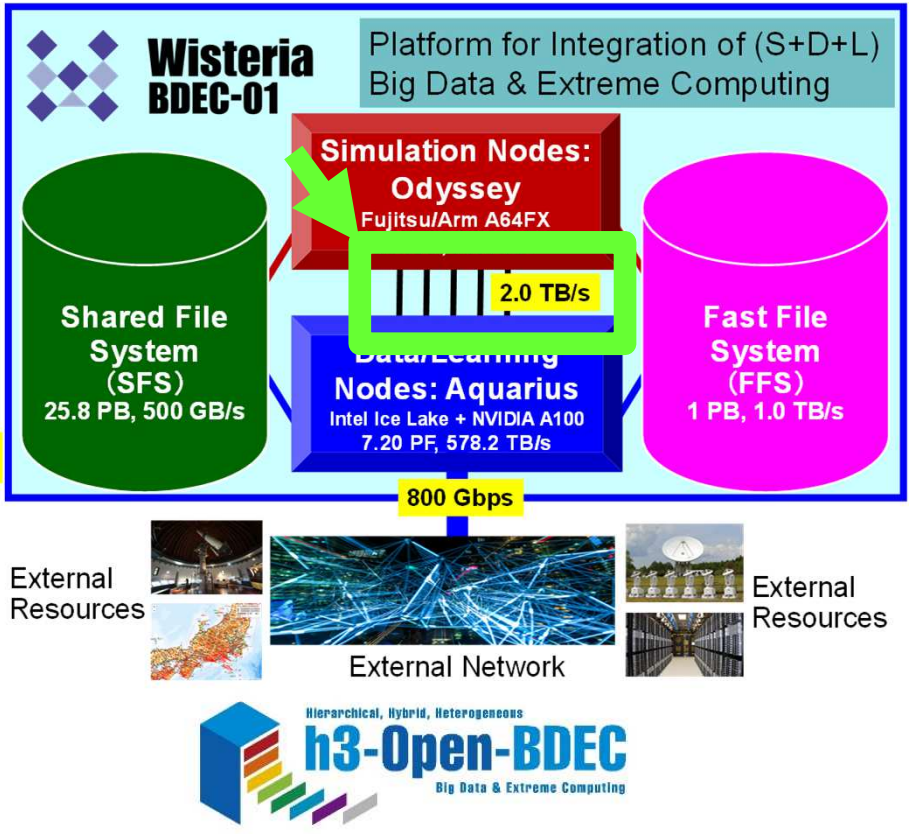
- **Wisteria/BDEC-01**
  - **Aquarius (GPU: NVIDIA A100)**
    - **Filtering, Visualization**
  - **Odyssey (CPU: A64FX)**
    - **Data Assimilation, Simulation**
- Combining Odyssey-Aquarius
  - Single MPI Job over O-A is impossible



Hierarchical, Hybrid, Heterogeneous
h3-Open-BDEC
Big Data & Extreme Computing

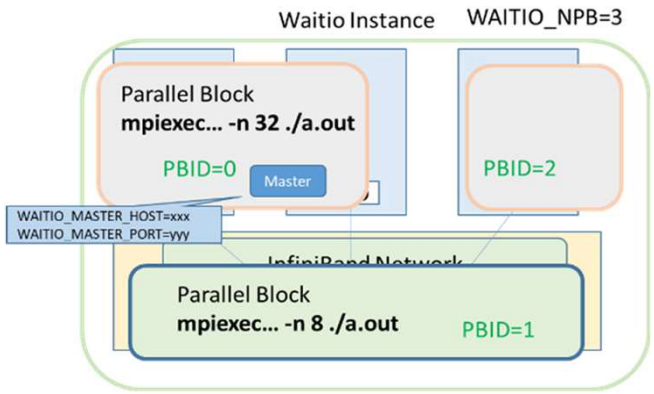
Computing on Wisteria/BDEC-01

- Wisteria/BDEC-01
  - Aquarius (GPU: NVIDIA A100)
    - Filtering, Visualization
  - Odyssey (CPU: A64FX)
    - Data Assimilation, Simulation
- Combining Odyssey-Aquarius
  - Single MPI Job over O-A is impossible
  - Actually, O-A are connected through IB-EDR with 2TB/sec.
  - h3-Open-SYS/WaitIO-Socket
    - Library for Inter-Process Communication through IB-EDR with MPI-like interface
  - h3-Open-UTIL/MP
    - Multiphysics Coupler



API of h3-Open-SYS/WaitIO-Socket
PB (Parallel Block): Each Application
More Detailed Talk by Prof. S. Sumimoto

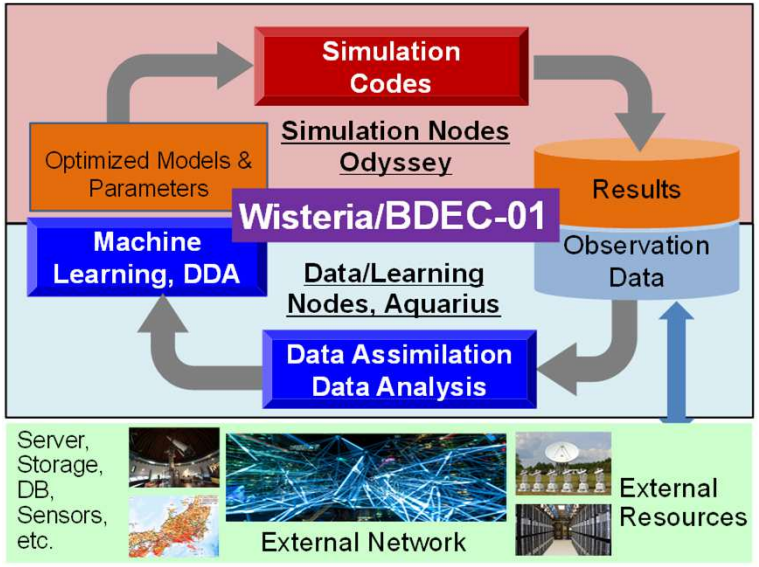
WaitIO API	Description
waitio_isend	Non-Blocking Send
waitio_irecv	Non-Blocking Receive
waitio_wait	Termination of waitio_isend/irecv
waitio_init	Initialization of WaitIO
waitio_get_nprocs	Process # for each PB (Parallel Block)
waitio_create_group waitio_create_group_wranks	Creating communication groups among PB's
waitio_group_rank	Rank ID in the Group
waitio_group_size	Size of Each Group
waitio_pb_size	Size of the Entire PB
waitio_pb_rank	Rank ID of the Entire PB

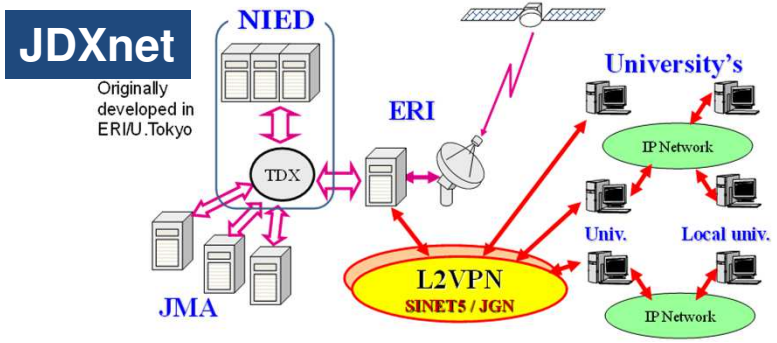


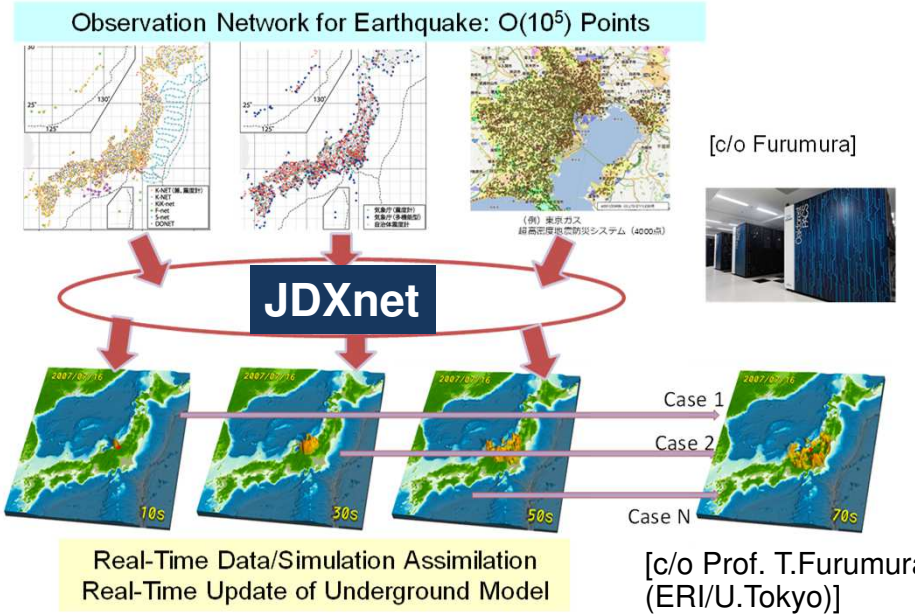
[Sumimoto et al. 2021]

3D Earthquake Simulation with Real-Time Data Observation/Assimilation

Simulation of Strong Motion (Wave Propagation) by 3D FDM

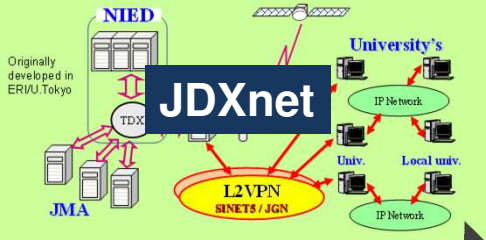






[c/o Prof. T.Furumura (ERI/U.Tokyo)]

System on Wisteria/BDEC-01 using WaitIO



Observed Data

Obs. Data

SINET5

External Server for Observed Data

Wisteria BDEC-01

Data/Learning Nodes: Aquarius
Intel Ice Lake + NVIDIA A100
7.20 PF, 578.2 TB/s

2.0 TB/s

Simulation Nodes: Odyssey
Fujitsu/Arm A64FX
25.9PF, 7.8 PB/s

Output & Movie

Filter

Visualizer

h3-Open-SYS/WaitIO-Socket

Seism3D/ OpenSWPC-DAF

Mesh Data

UG Model

Communications by WaitIO-Socket

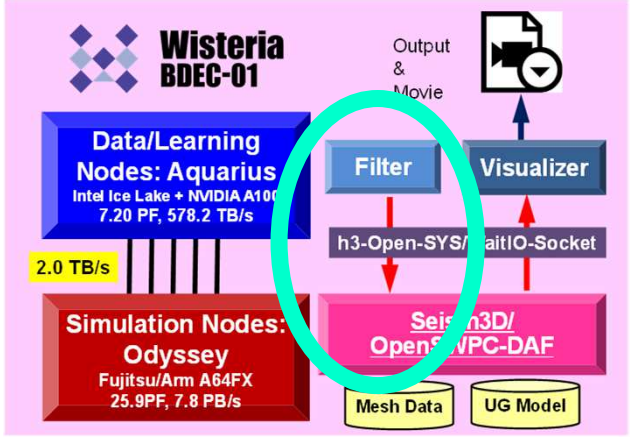
[Kasai et al. 2021]

Aquarius: SEND

```
program dmy_filter
<省略: 型宣言等>
call mpi_init (ierr)
call mpi_comm_size (MPI_COMM_WORLD, nprocs, ierr)
call mpi_comm_rank (MPI_COMM_WORLD, myrank,  ierr)
call WAITIO_CREATE_UNIVERSE (WAITIO_COMM_UNIVERSE, ierr)

if (myrank==0) then
  open(100,file='./obsfile_list.txt', form='formatted', status='old', iostat=ierr)
  do i=1,300
    <省略: obsデータ読み込み処理>
    print *,"Send obs data ..... "
    call WAITIO_MPI_ISEND (NTMAX1_o, 1,           WAITIO_MPI_INTEGER,  2,1, WAITIO_COMM_UNIVERSE,req(1,1), ierr)
    call WAITIO_MPI_ISEND (DT_o,     1,           WAITIO_MPI_FLOAT,    2,2, WAITIO_COMM_UNIVERSE,req(1,2), ierr)
    call WAITIO_MPI_ISEND (NST_o,    1,           WAITIO_MPI_INTEGER,  2,3, WAITIO_COMM_UNIVERSE,req(1,3), ierr)
    call WAITIO_MPI_ISEND (AT_o,     1,           WAITIO_MPI_FLOAT,    2,4, WAITIO_COMM_UNIVERSE,req(1,4), ierr)
    call WAITIO_MPI_ISEND (T0_o,     1,           WAITIO_MPI_FLOAT,    2,5, WAITIO_COMM_UNIVERSE,req(1,5), ierr)
    call WAITIO_MPI_ISEND (ISO_X_o,  NSMAX,       WAITIO_MPI_INTEGER,  2,6, WAITIO_COMM_UNIVERSE,req(1,6), ierr)
    call WAITIO_MPI_ISEND (ISO_Y_o,  NSMAX,       WAITIO_MPI_INTEGER,  2,7, WAITIO_COMM_UNIVERSE,req(1,7), ierr)
    call WAITIO_MPI_ISEND (ISO_Z_o,  NSMAX,       WAITIO_MPI_INTEGER,  2,8, WAITIO_COMM_UNIVERSE,req(1,8), ierr)
    call WAITIO_MPI_ISEND (ISTX_o,   NST,         WAITIO_MPI_INTEGER,  2,9, WAITIO_COMM_UNIVERSE,req(1,9), ierr)
    call WAITIO_MPI_ISEND (ISTY_o,   NST,         WAITIO_MPI_INTEGER,  2,10,WAITIO_COMM_UNIVERSE,req(1,10),ierr)
    call WAITIO_MPI_ISEND (ISTZ_o,   NST,         WAITIO_MPI_INTEGER,  2,11,WAITIO_COMM_UNIVERSE,req(1,11),ierr)
    call WAITIO_MPI_ISEND (STC_o,    6*NST,       WAITIO_MPI_CHAR,     2,12,WAITIO_COMM_UNIVERSE,req(1,12),ierr)
    call WAITIO_MPI_ISEND (VxAll_obs,NST*NOBS_LEN,WAITIO_MPI_FLOAT,    2,13,WAITIO_COMM_UNIVERSE,req(1,13),ierr)
    call WAITIO_MPI_ISEND (VyAll_obs,NST*NOBS_LEN,WAITIO_MPI_FLOAT,    2,14,WAITIO_COMM_UNIVERSE,req(1,14),ierr)
    call WAITIO_MPI_ISEND (VzAll_obs,NST*NOBS_LEN,WAITIO_MPI_FLOAT,    2,15,WAITIO_COMM_UNIVERSE,req(1,15),ierr)
    call WAITIO_MPI_WAITALL (15,req, status, ierr)
    call sleep(1)
  enddo
  close (100)
endif
call WAITIO_FINALIZE (ierr)
call mpi_finalize (ierr)
end
```

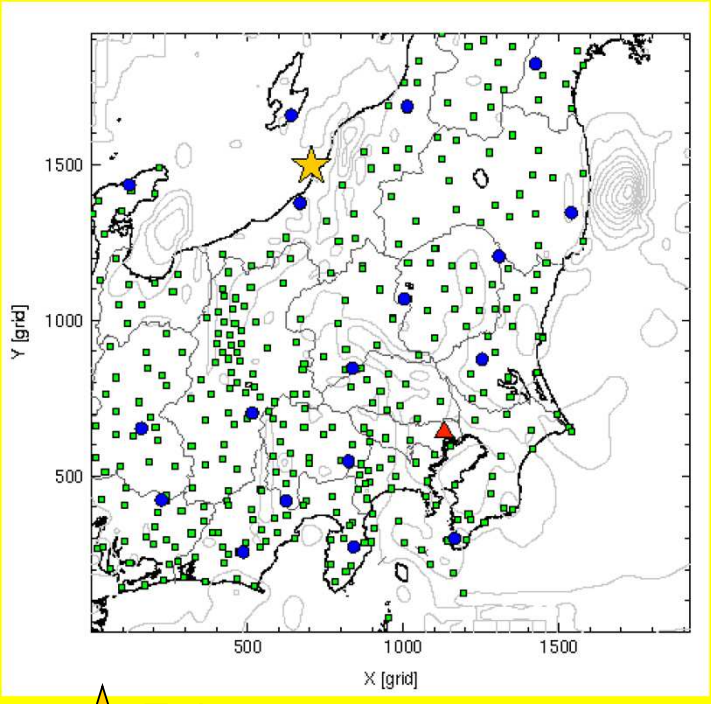
Odyssey: RECV

```
call WAITIO_MPI_IRECV (NTMAX1_o, 1,           WAITIO_MPI_INTEGER, 0,1, WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (DT_o,     1,           WAITIO_MPI_FLOAT,   0,2, WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (NST_o,    1,           WAITIO_MPI_INTEGER, 0,3, WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (AT_o,     1,           WAITIO_MPI_FLOAT,   0,4, WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (T0_o,     1,           WAITIO_MPI_FLOAT,   0,5, WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (ISO_X_o,  NSMAX,       WAITIO_MPI_INTEGER, 0,6, WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (ISO_Y_o,  NSMAX,       WAITIO_MPI_INTEGER, 0,7, WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (ISO_Z_o,  NSMAX,       WAITIO_MPI_INTEGER, 0,8, WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (ISTX_o,   NST,         WAITIO_MPI_INTEGER, 0,9, WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (ISTY_o,   NST,         WAITIO_MPI_INTEGER, 0,10,WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (ISTZ_o,   NST,         WAITIO_MPI_INTEGER, 0,11,WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (STC_o,    6*NST,       WAITIO_MPI_CHAR,    0,12,WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (VxAll_obs,NST*NOBS_LEN,WAITIO_MPI_FLOAT,   0,13,WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (VyAll_obs,NST*NOBS_LEN,WAITIO_MPI_FLOAT,   0,14,WAITIO_COMM_UNIVERSE,...)
call WAITIO_MPI_IRECV (VzAll_obs,NST*NOBS_LEN,WAITIO_MPI_FLOAT,   0,15,WAITIO_COMM_UNIVERSE,...)
```



Off Niigata 2007 Mw6.6 Earthquake

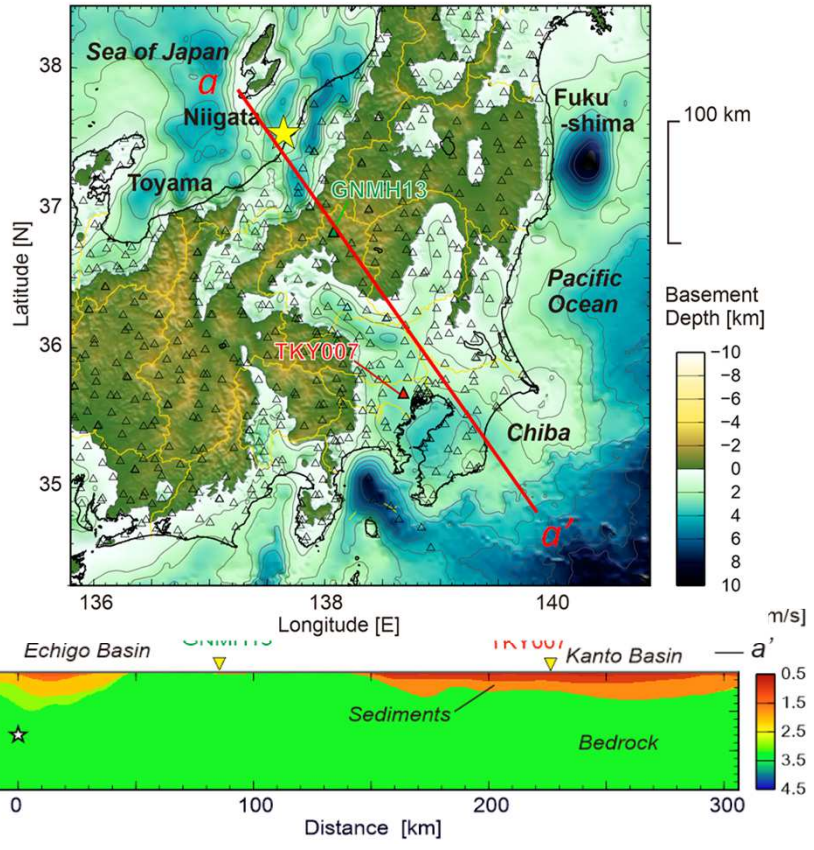
[c/o Prof. T. Furumura, ERI/U.Tokyo]



Epicenter

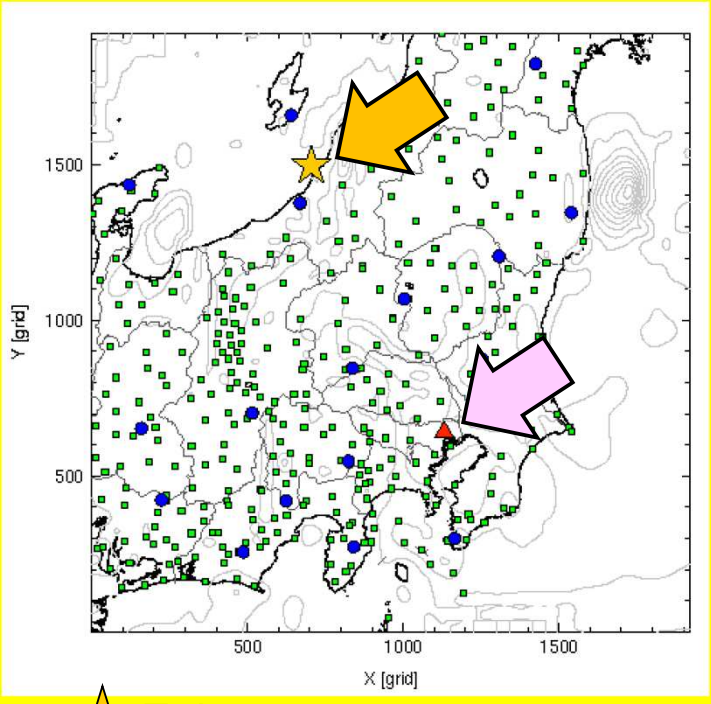
Hi-net (Short Period) 349 pts

F-net (Broadband) 18 pts



Off Niigata 2007 Mw6.6 Earthquake

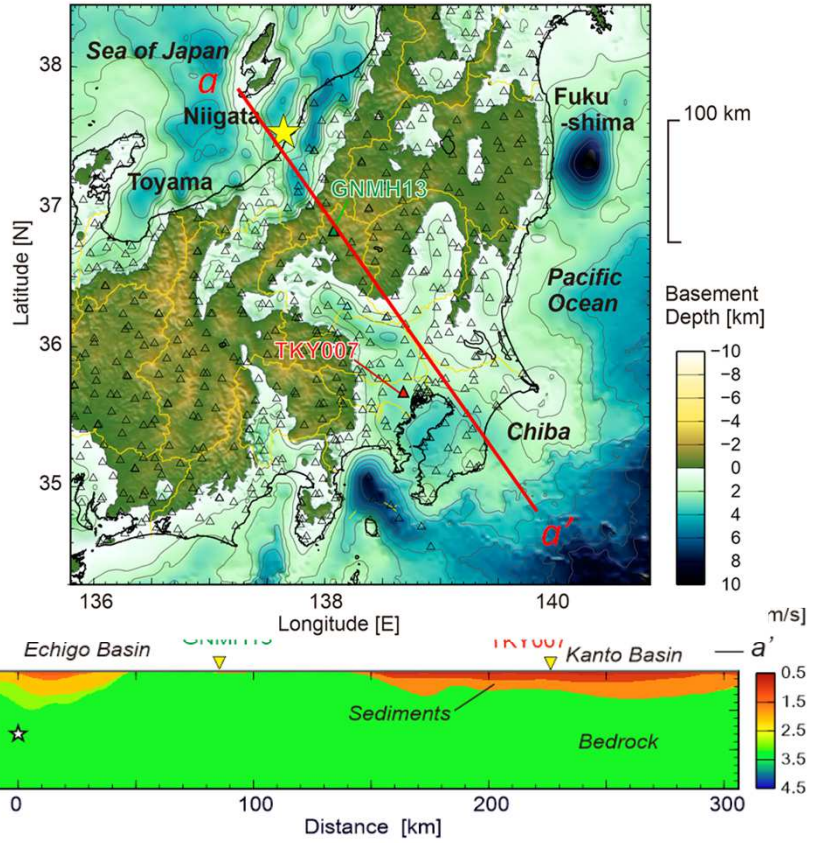
[c/o Prof. T. Furumura, ERI/U.Tokyo]



☆ Epicenter

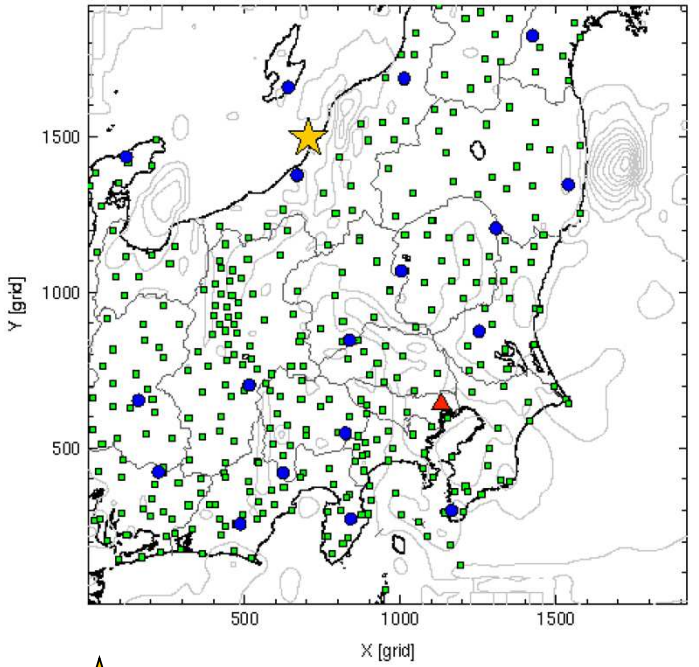
■ Hi-net (Short Period) 349 pts

● F-net (Broadband) 18 pts



Off Niigata 2007 Mw6.6 Earthquake

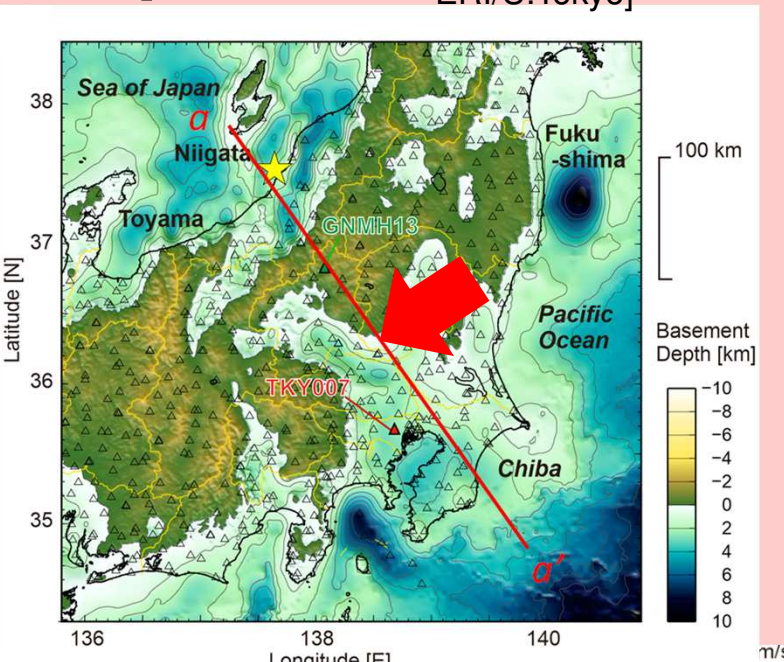
[c/o Prof. T. Furumura, ERI/U.Tokyo]

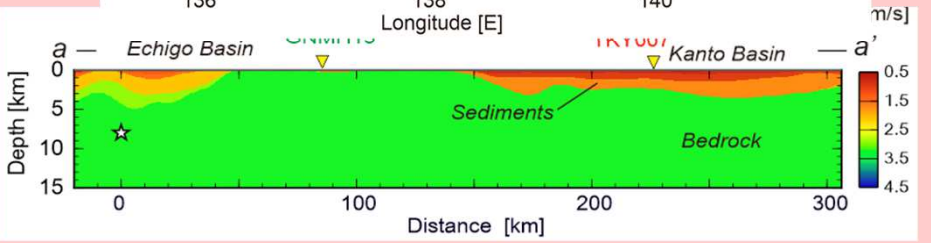


Epicenter

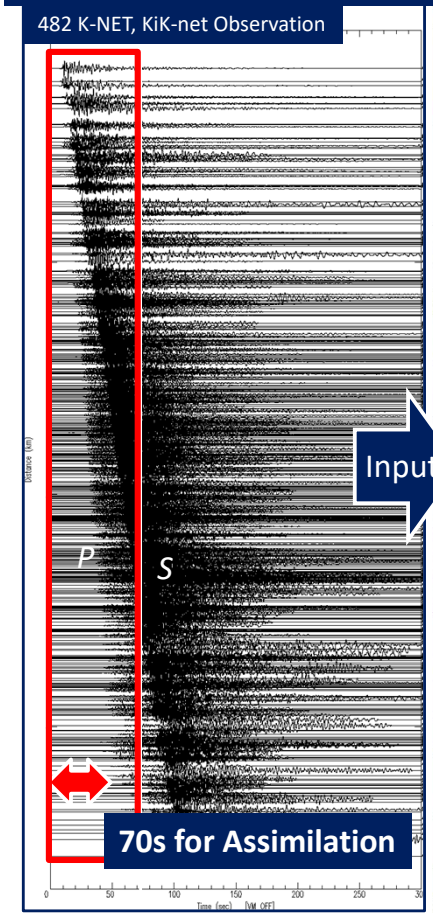
Hi-net (Short Period) 349 pts

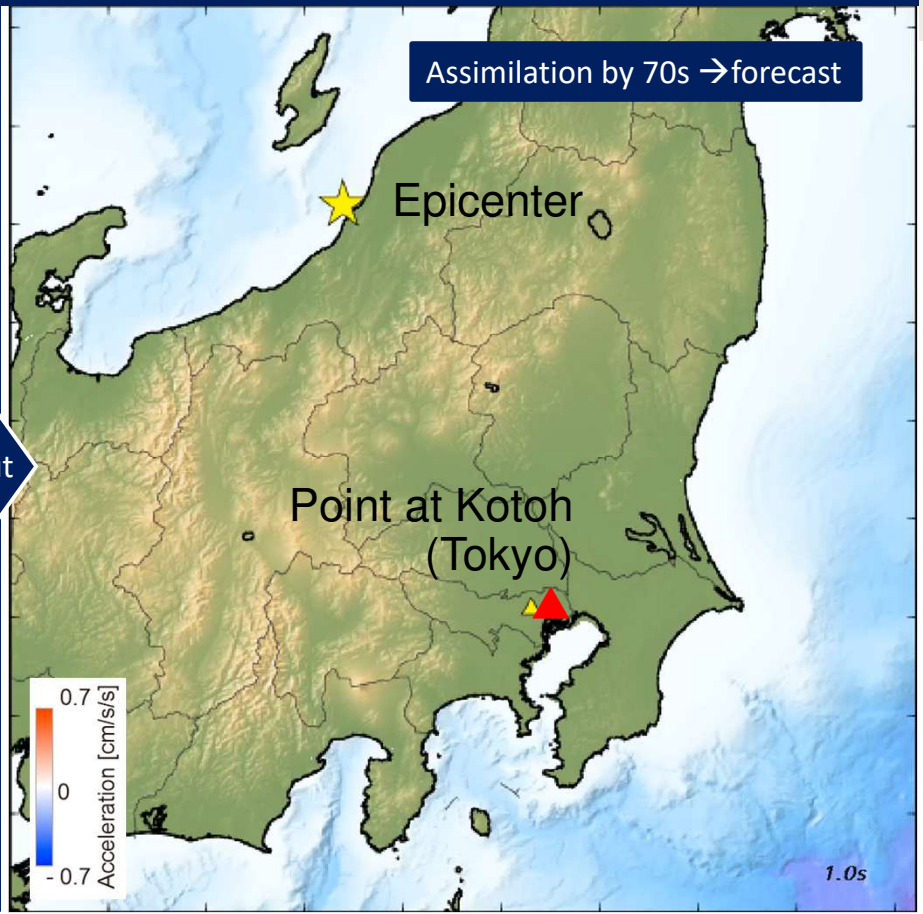
F-net (Broadband) 18 pts



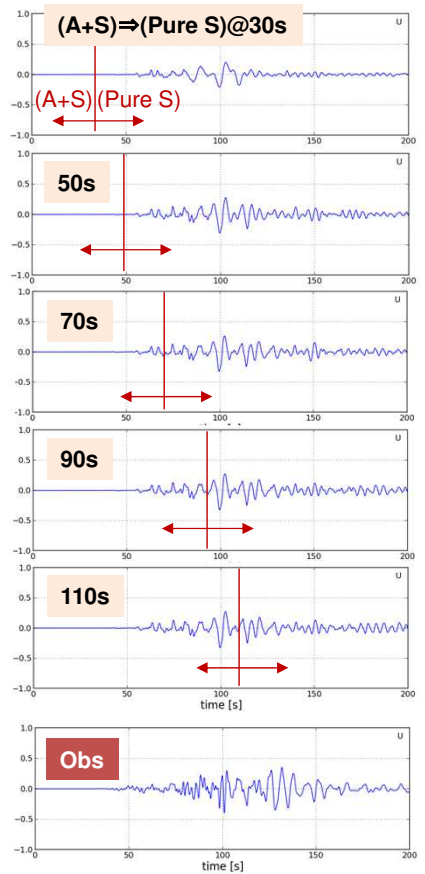


Data Assimilation + Pure Simulation/Forecast



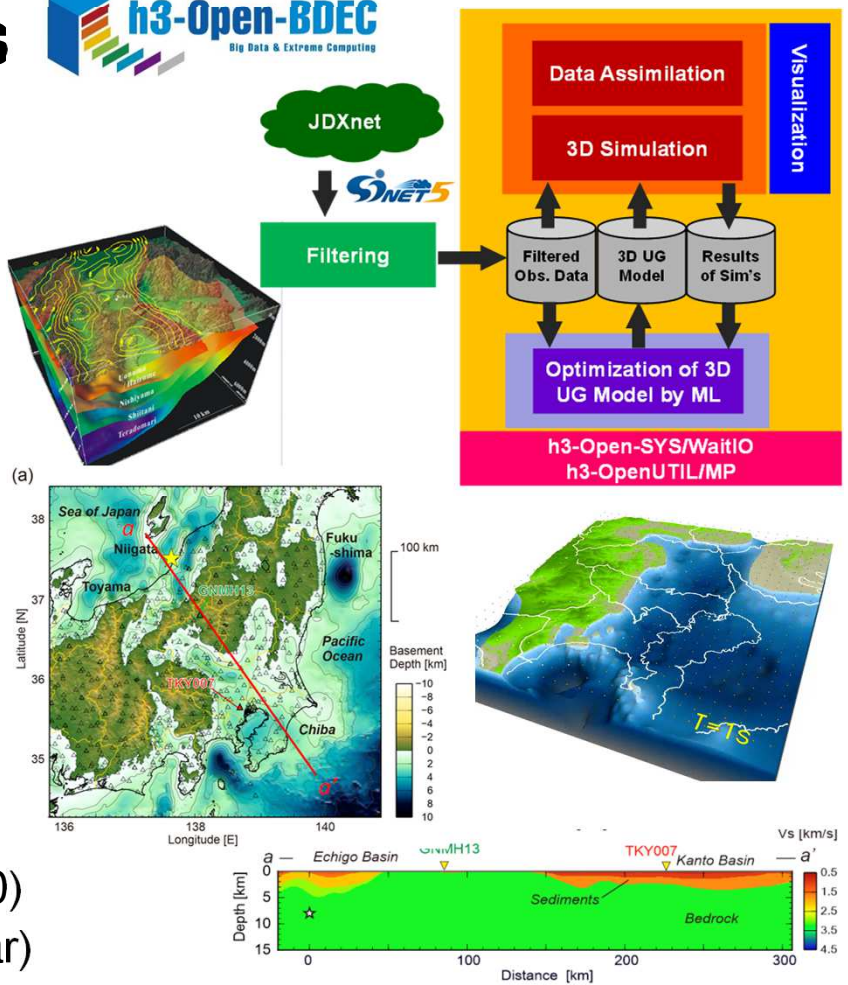


Results at Kotoh ▲ (N.KOTH)
N 35° 37.0'
E 139° 46.9'



Future Directions towards Integration of (S+D+L)

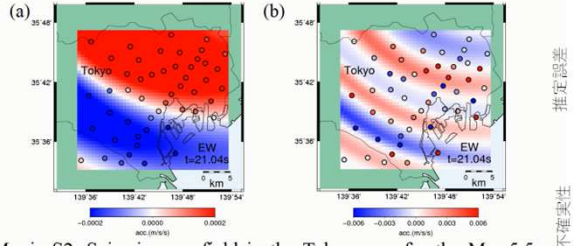
- Accurate Prediction of Seismic Wave Propagation with Real-Time Data Observation/Assimilation
  - Emergency Info. for Safer Evacuation
- 3D Underground Model
  - Heterogeneous, Observation is difficult
  - Inversion analyses of seismic waves are important for prediction of structure of underground model
  - ML may be utilized for acceleration of this prediction based on analyses of small earthquakes in normal time (e.q. Mw < 3.0)
  - More sophisticated DA method (e.g. 4DVar)



Actually, construction of 3D Underground Model by this Model for Long-Period Seismic Wave Propagation is not realistic

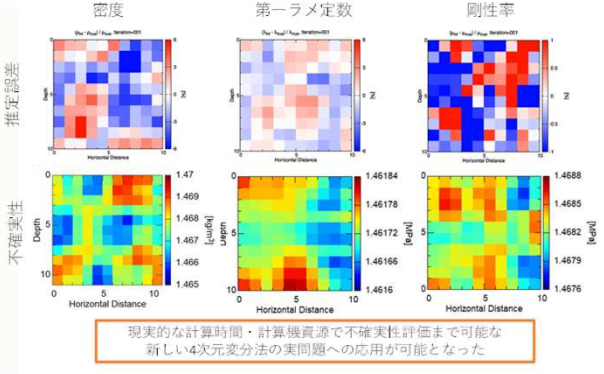
- Local models with smaller meshes should be used

Replica Exchange Monte Carlo Nagao et al.

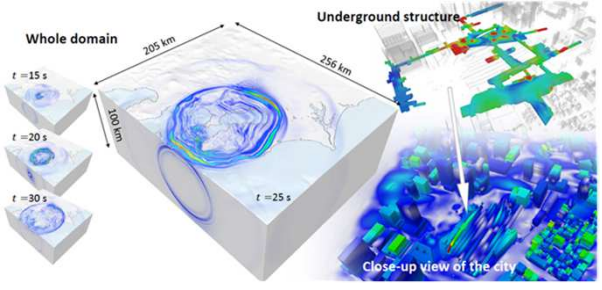


Movie S2. Seismic wavefield in the Tokyo area for the Mw 5.5 earthquake of 16 September 2014 in the northern Kanto area, in the frequency band (a) 0.10–0.20 Hz and (b) 0.10–1.0 Hz, computed with the optimum model parameters, compared to the observations (circles).

2nd Order Adjoint Nagao et al.



Large-Scale ML Ichimura, Fujita SC22 GB Finalists

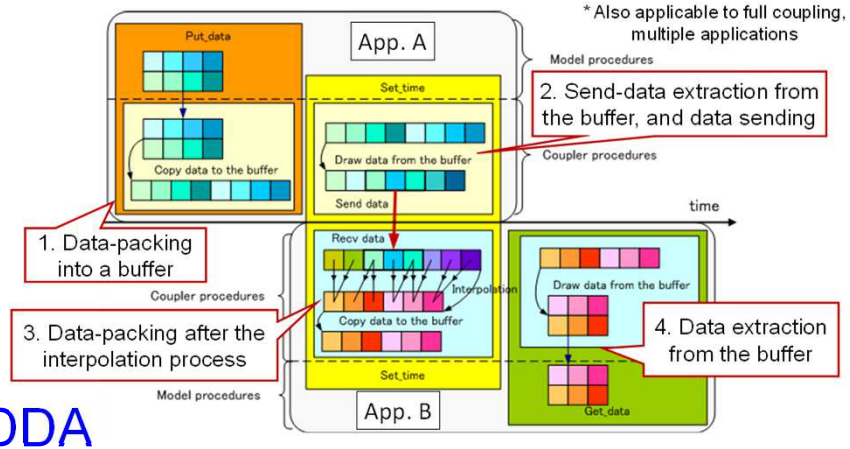


- Wisteria/BDEC-01
- h3-Open-BDEC
- Earthquake Simulations combined with Real-Time Data Assimilation
  - h3-Open-SYS/WaitIO
- **HPC-AI Workload: Global Cloud Simulations**
  - **h3-Open-UTIL/MP**
- How to run such (a little bit complicated) workflows
  - Wisteria/BDEC-01
  - Wisteria/BDEC-01 + mdx
- Future/Ongoing Works
  - International Collaborations
  - OFP-II

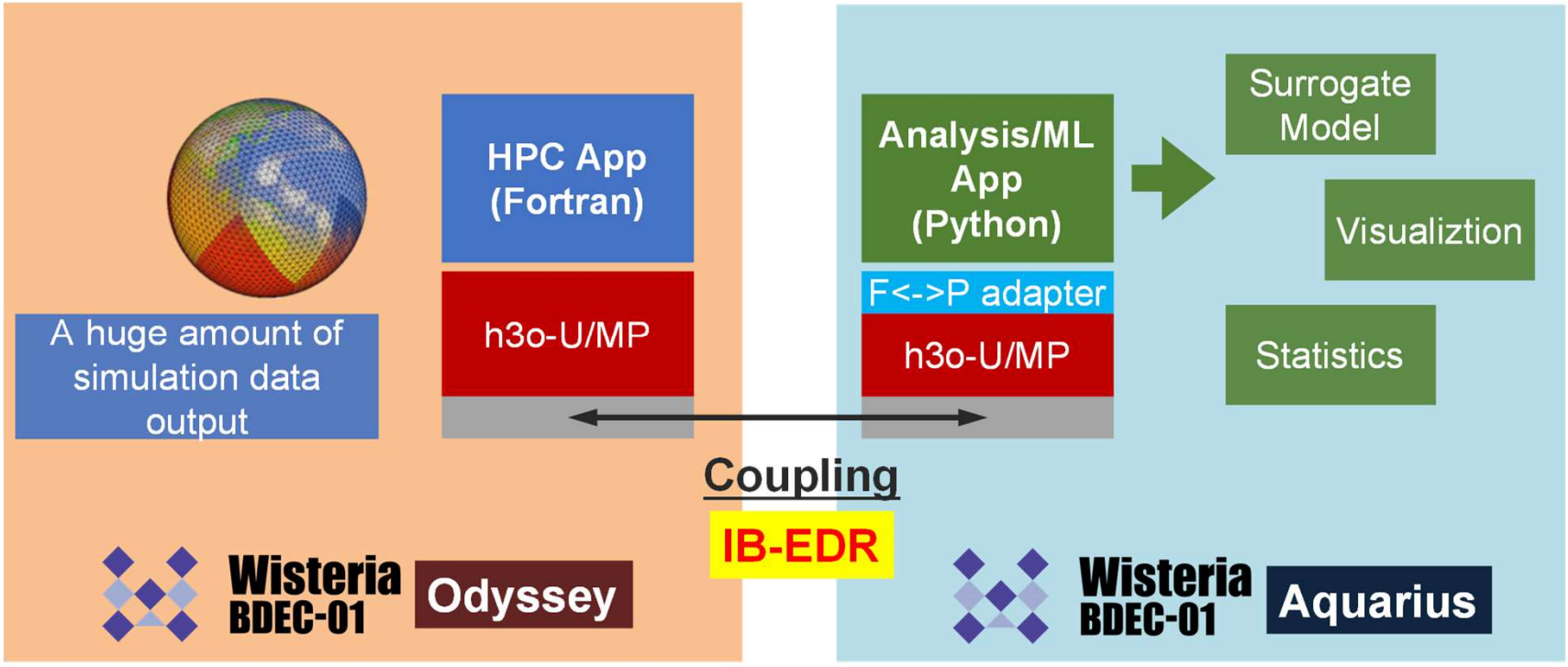
h3-Open-UTIL/MP

Multilevel Coupler/Data Assimilation

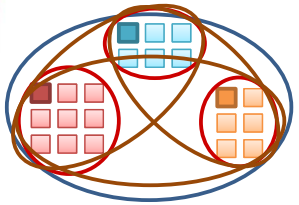
- Traditional Coupler: ppOpen-MATH/MP
  - Weak-Coupling of Multiple (usually two) Applications
    - Each application does a single computation
- h3-Open-UTIL/MP
  - Data Assimilation (Multiple Computations: Ensemble)
  - Assimilation of Computations with Different Resolutions
    - h3-Open-DATA, h3-Open-APP
  - Data Assimilation by Coupled Codes
    - e.g. Atmosphere-Ocean
- Data Assimilation: h3-Open-DATA
  - Karman Filter, Particle Karman Filter
  - LETKF
  - Adjoint Method
- Generation of Simplified Models in hDDA

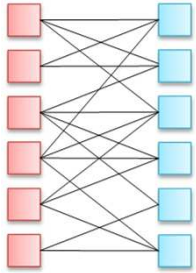


h3-Open-UTIL/MP (h3o-U/MP) + h3-Open-SYS/WaitIO-Socket



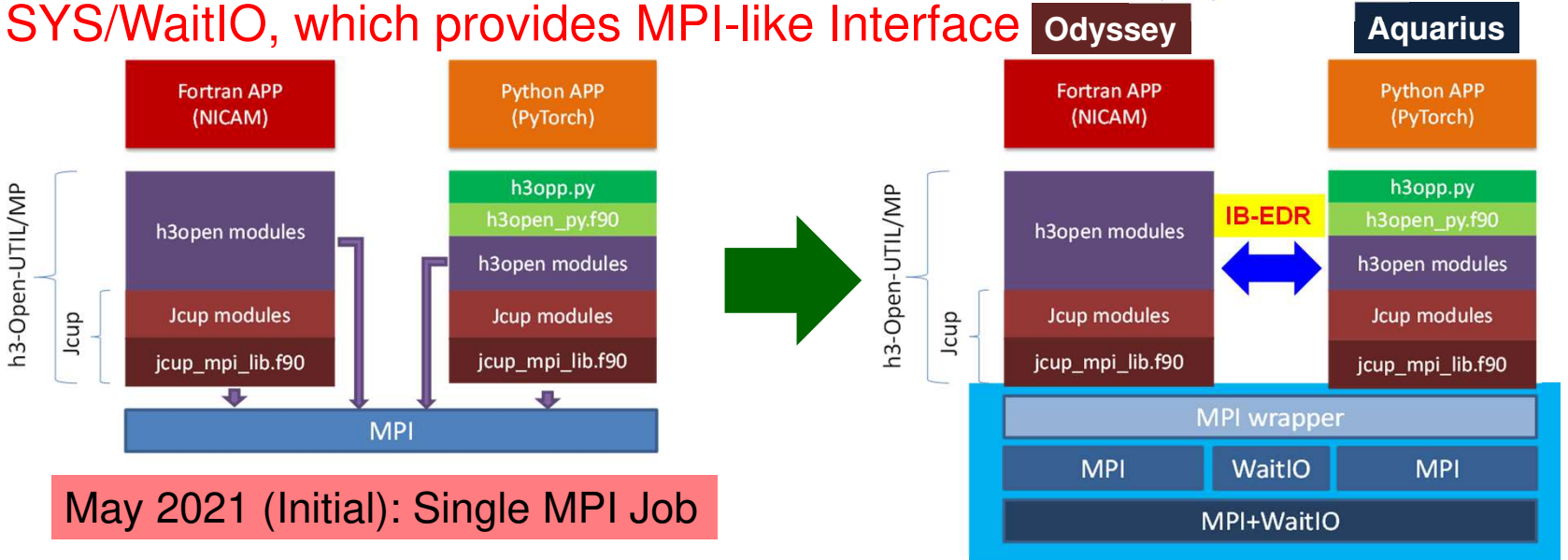
h3-Open-UTIL/MP + h3-Open-SYS/WaitIO-Socket





- Single MPI Job (May 2021)
- Direct Communication between Odyssey-Aquarius through IB-EDR by h3-Open-SYS/WaitIO, which provides MPI-like Interface

Wisteria BDEC-01



May 2021 (Initial): Single MPI Job

Atmosphere-ML Coupling
[Yashiro (NIES), Arakawa (ClimTech/U.Tokyo)]

- Motivation of this experiment
  - Tow types of Atmospheric models: Cloud resolving VS Cloud parameterizing
  - Could resolving model is difficult to use for climate simulation
  - Parameterized model has many assumptions
  - Replacing low-resolution cloud processes calculation with ML!

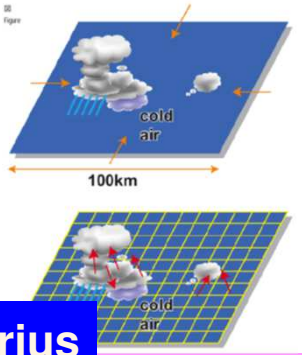
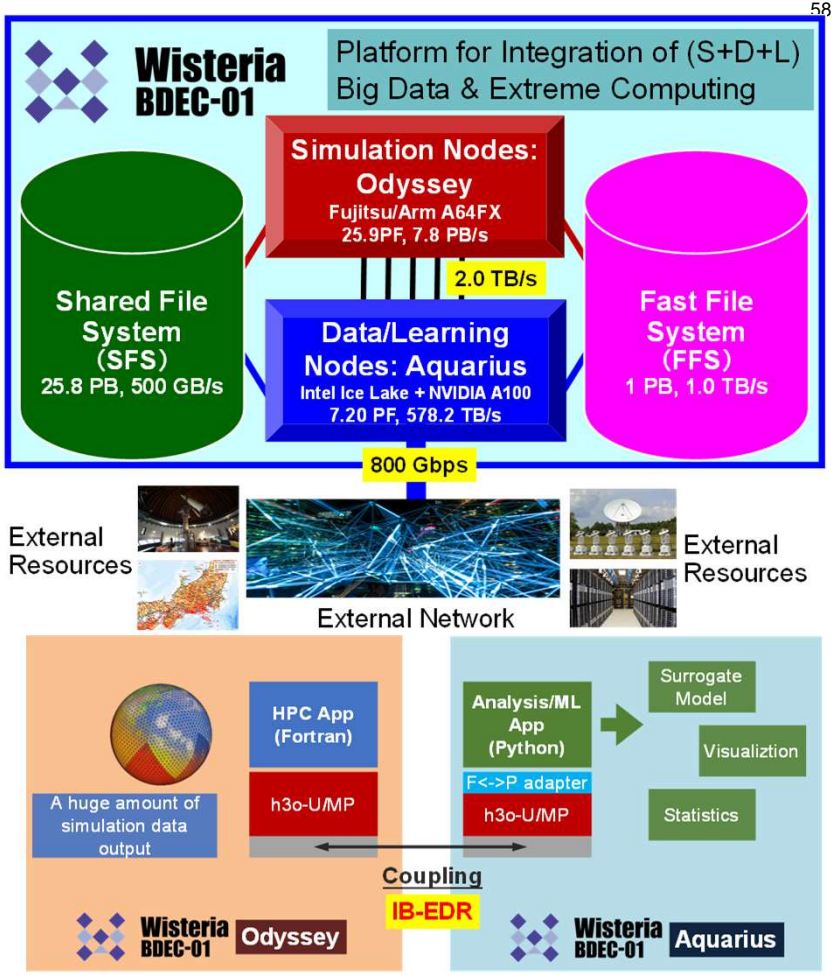


Diagram of applying ML to an atmospheric model

- Wisteria/BDEC-01
- h3-Open-BDEC
- Earthquake Simulations combined with Real-Time Data Assimilation
  - h3-Open-SYS/WaitIO
- HPC-AI Workload: Global Cloud Simulations
  - h3-Open-UTIL/MP
- **How to run such (a little bit complicated) workflows**
  - **Wisteria/BDEC-01**
  - **Wisteria/BDEC-01 + mdx**
- Future/Ongoing Works
  - International Collaborations
  - OFP-II

How to run the workloads

- Total Number of Nodes
  - Odyssey: 7,680 nodes: not so crowded
  - Aquarius: 45 nodes, 360 GPUs, very crowded
- One node of Aquarius is reserved for this type of workload on the integration of (S+D+L)
- 2 separate jobs (Odyssey, Aquarius) should be submitted
- If both jobs “grab” resources, execution starts.



Examples of Scripts [Sumimoto, Arakawa]

Odyssey for Simulation

```
#!/bin/bash
#PJM -N "test_waitio"
#PJM -L rscgrp=coupler-lec-o
#PJM -L node=10:noncont
#PJM --mpi proc=80
#PJM -L elapse=00:10:00
#PJM -g gt00
#PJM -j
#PJM -e err

module load fj
module load fjmpi
module load waitio

export WAITIO_MASTER_HOST=`hostname`
export WAITIO_MASTER_PORT=7100
export WAITIO_PBID=0
export WAITIO_NPB=2

hostname
waitio-serv-a64fx -d -m $WAITIO_MASTER_HOST

#mpiexec -oferr-proc errnicam -np 160 ./nicam
mpiexec -np 80 ./nicam
```

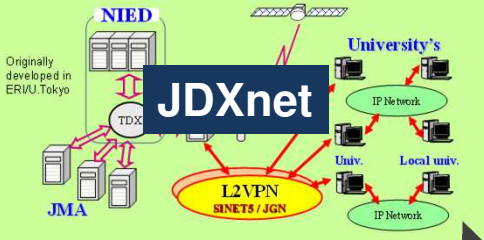
Aquarius for AI

```
#!/bin/bash
#PJM -N "test_waitio"
#PJM -L rscgrp=coupler-lec-a
#PJM -L node=1
#PJM --mpi proc=10
#PJM -L elapse=00:10:00
#PJM -g gt00
#PJM -j
#PJM -e err

module unload aquarius
module unload gcc ompi
module load intel
module load impi
module load waitio

export WAITIO_MASTER_HOST=`waitio-serv -c`
export WAITIO_MASTER_PORT=7100
export WAITIO_PBID=1
export WAITIO_NPB=2

module unload intel
module unload impi
module load gcc ompi

mpiexec -n 10 ./ada
```

Lessons Learned & Future

- **Software (h3-Open-BDEC: WaitIO, Coupler) enabled integration of (S+D+L) on Odyssey-Aquarius**
  - **WatioIO-Socket/File**
  - **Job Submission System**
- Policy for Operation
  - Current status is preliminary, Very few workloads for (S+D+L)
  - More flexible (& complicated) policy needed
- Publication
  - Shinji Sumimoto, et al. PDCAT'22, Dec.7-9, 2022 (Best Paper Award)
    - https://www.hpc.is.tohoku.ac.jp/pdcat2022/

System on Wisteria/BDEC-01 using WaitIO



Observed Data

Obs. Data

SINET5

External Server for Observed Data

Wisteria BDEC-01

Data/Learning Nodes: Aquarius
Intel Ice Lake + NVIDIA A100
7.20 PF, 578.2 TB/s

2.0 TB/s

Simulation Nodes: Odyssey
Fujitsu/Arm A64FX
25.9PF, 7.8 PB/s

Output & Movie

Filter

Visualizer

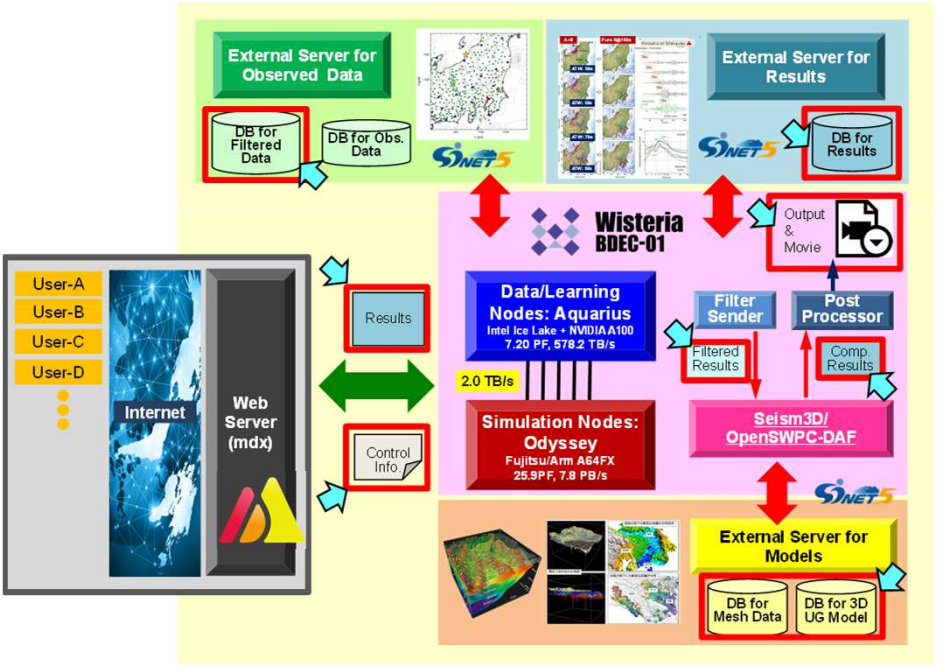
h3-Open-SYS/WaitIO-Socket

Seism3D/ OpenSWPC-DAF

Mesh Data

UG Model

Web-based Simulation System for Outreach Activities

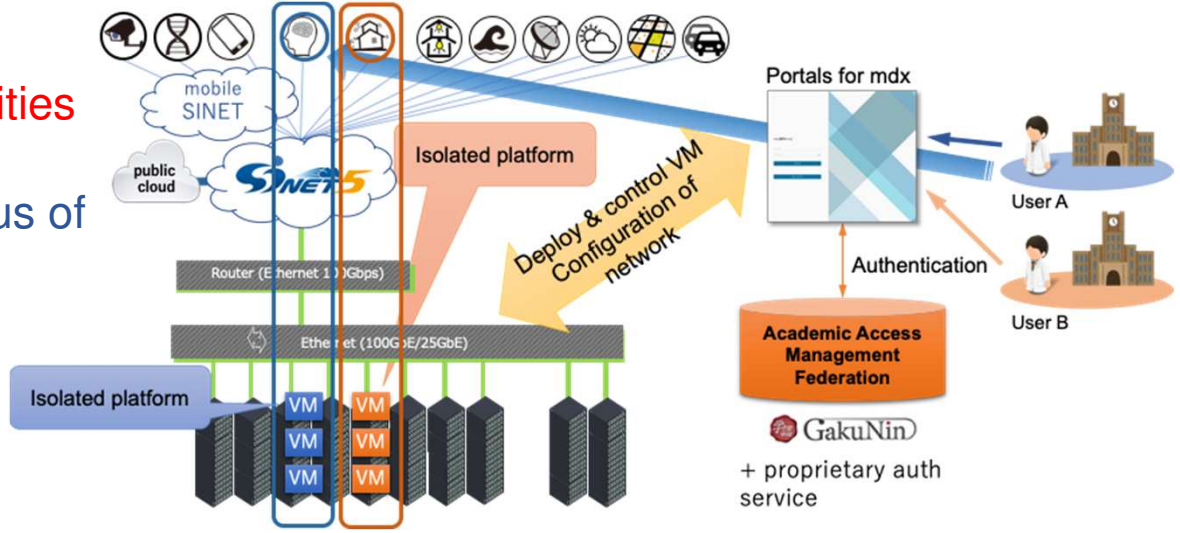


- Web-based simulation system for enlightenment of disaster prevention/mitigation awareness using “3D Earthquake Simulation with Real-Time Data Observation/Assimilation”
- Users including general citizens and high-school/junior-high-school students, access the web-server on the mdx system, and manipulate simulations on the Wistera/BDEC-01.
- The framework can be utilized in various types of applications.

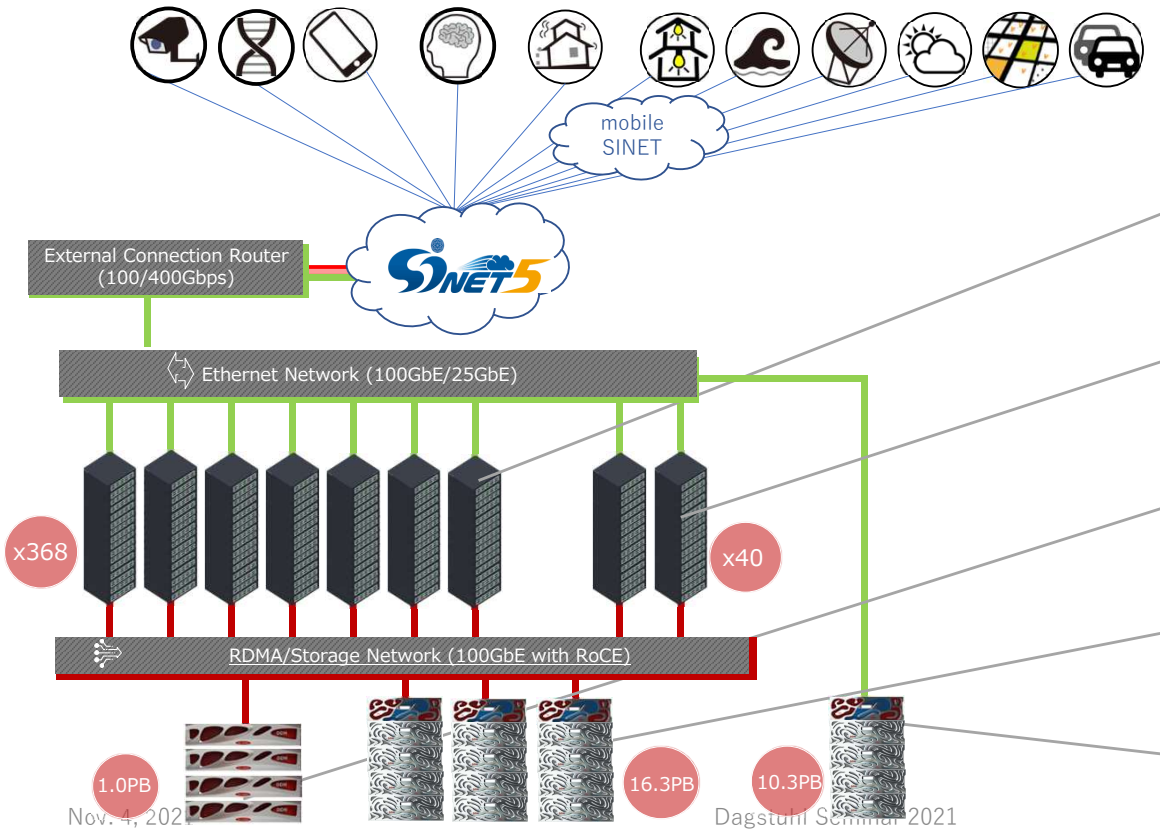
“Data Platform”

- High-performance virtualization environment focusing on leveraging data with security like “Cloud” using same HW as supercomputers.

- National joint usage system jointly operated by 9 universities and 2 research institutes
- Located at Kashiwa II campus of U.Tokyo



mdx Overview



Two kinds of networks

- External Network
  - ✓ 400Gbps x2 connection to SINET6
- Internal High Speed Interconnect
  - ✓ RDMA, Storage

General-purpose CPU node
Intel Ice Lake 2 x 368 nodes
Total peak FLOPS (FP64): 2.1 PFLOPS
Aggregate Mem. BW: 150.7 TB/sec

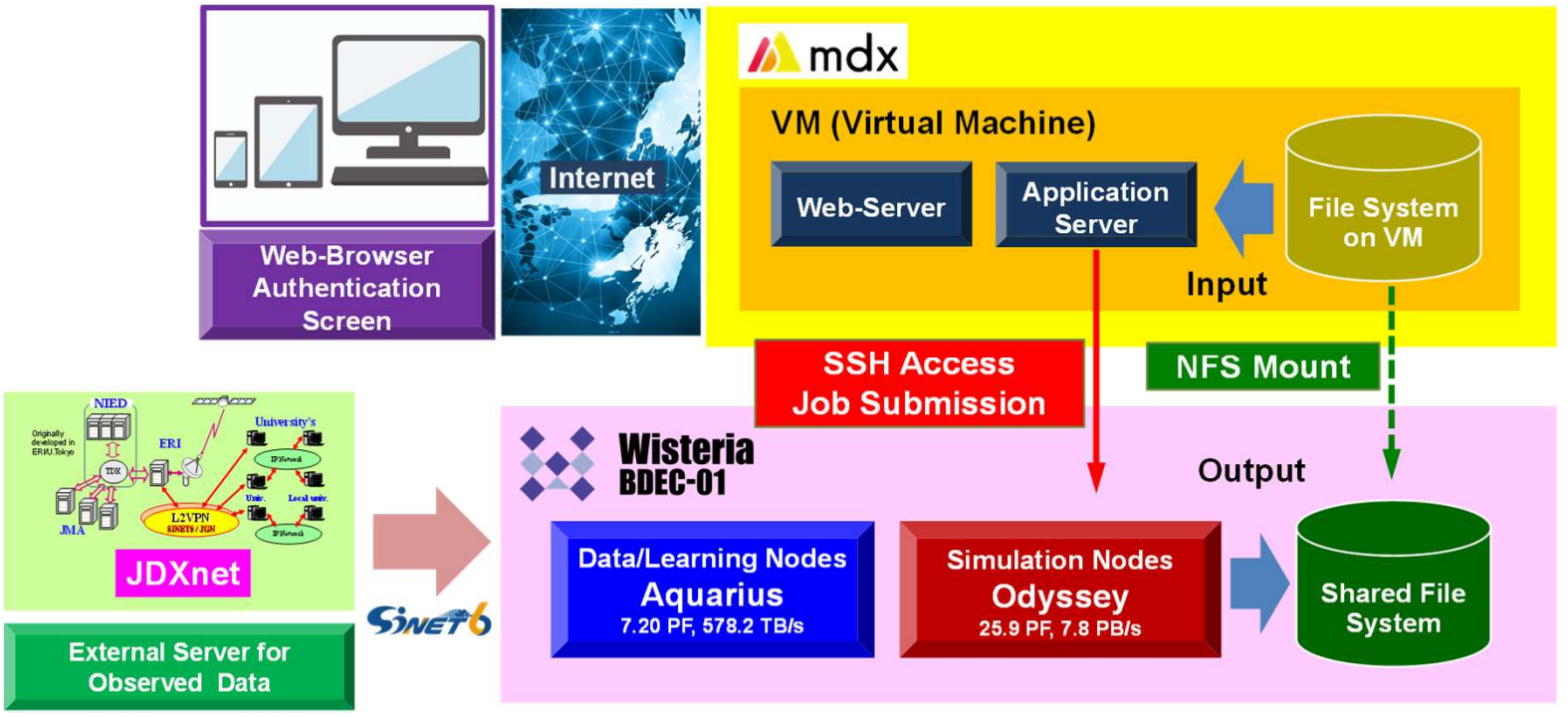
GPU Accelerator node
NVIDIA A100 x 8 x 40 nodes
Total peak FLOPS (FP64): 6.4 PFLOPS
Total peak FLOPS (FP16):100.7 PFLOPS
Aggregate Mem. BW:496.3TB/sec.

Fast NVMe Storage
Lustre Filesystem
1.0 PB (NVMe SSD)
252 GByte/sec

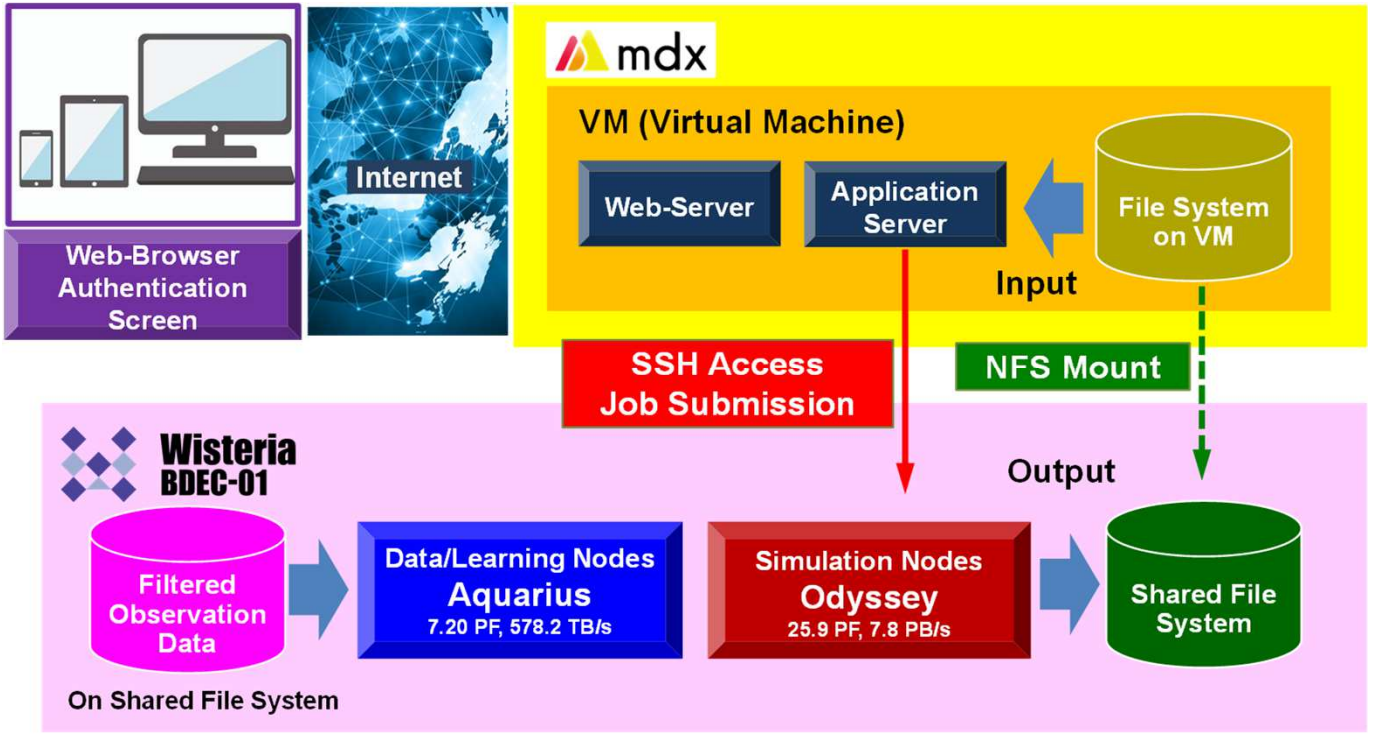
Large-scale HDD Storage
Lustre Filesystem
16.3 PB (HDD)
157.5 GByte/sec

External Shared Object Storage
S3 Data Service
10.3 PB (HDD)
63.0 GByte/sec

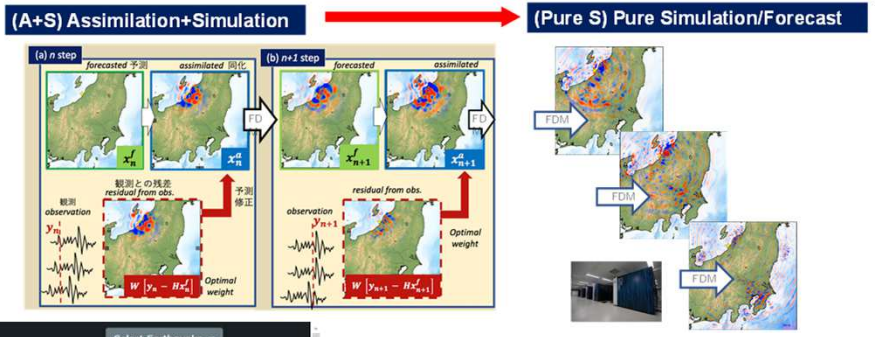
Web-based Simulation System for Outreach Activities

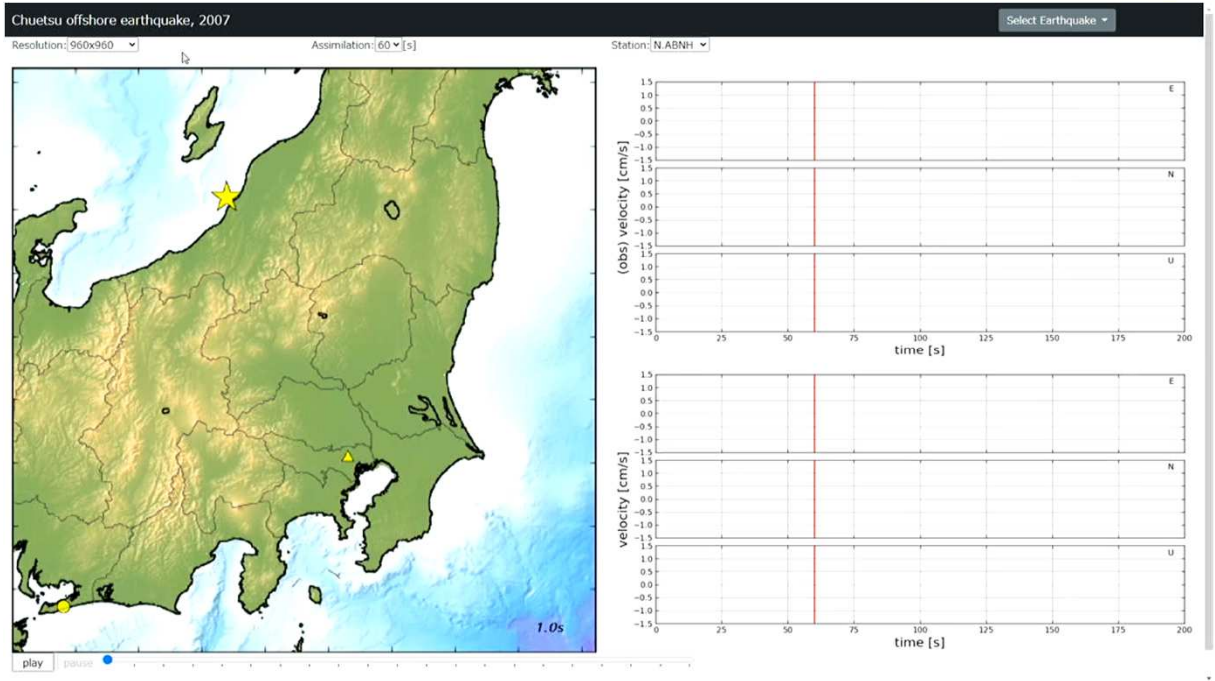


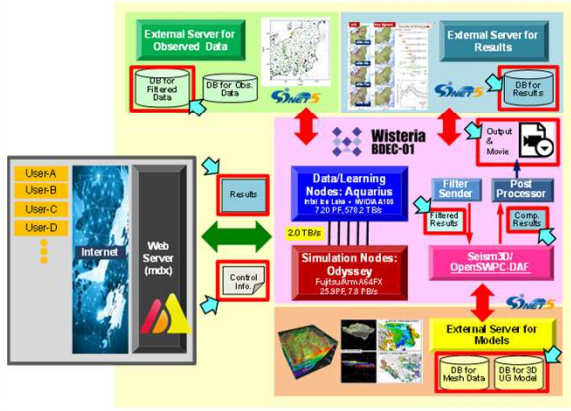
Web-based Simulation System for Outreach Activities (Prototype for Demonstration)



Web-based Simulation System for Outreach Activities (Prototype)





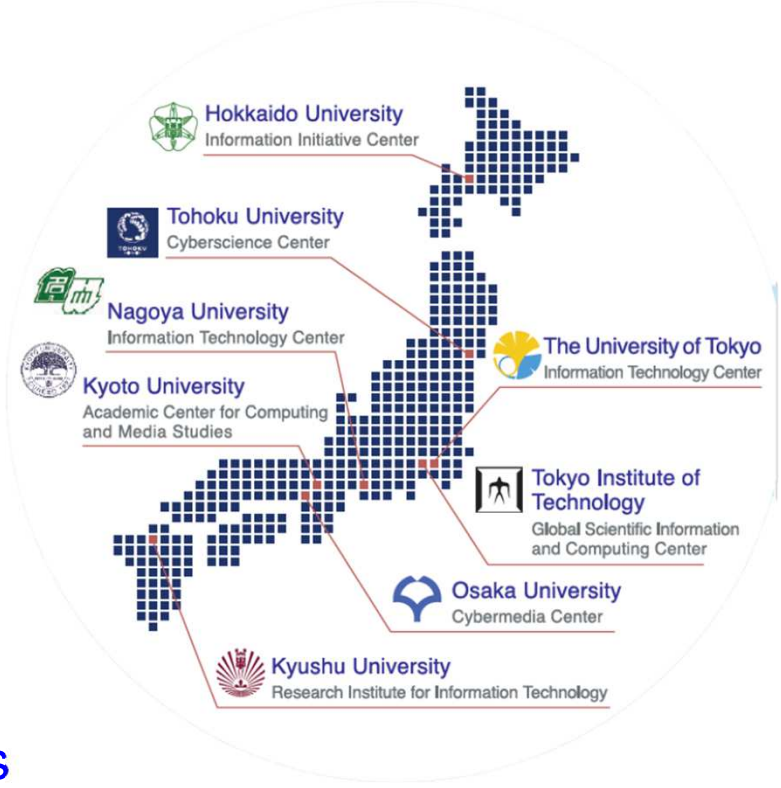


- Wisteria/BDEC-01
- h3-Open-BDEC
- Earthquake Simulations combined with Real-Time Data Assimilation
  - h3-Open-SYS/WaitIO
- HPC-AI Workload: Global Cloud Simulations
  - h3-Open-UTIL/MP
- How to run such (a little bit complicated) workflows
  - Wisteria/BDEC-01
  - Wisteria/BDEC-01 + mdx
- **Future/Ongoing Works**
  - **International Collaborations**
  - **OFP-II**

JHPCN

https://jhpcn-kyoten.itc.u-tokyo.ac.jp/en/

- Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures (2010-)
- Alliance of SC Centers of 8 National Universities (corresponding to NHR in Germany)
  - 7 “Imperial” Universities + Tokyo Tech
  - Core Institute: ITC/U.Tokyo
  - Total 140+PFLOPS (May 2022)
- Promotion of collaborative (fundamental, interdisciplinary) research projects using facilities & human resources in 8 Centers



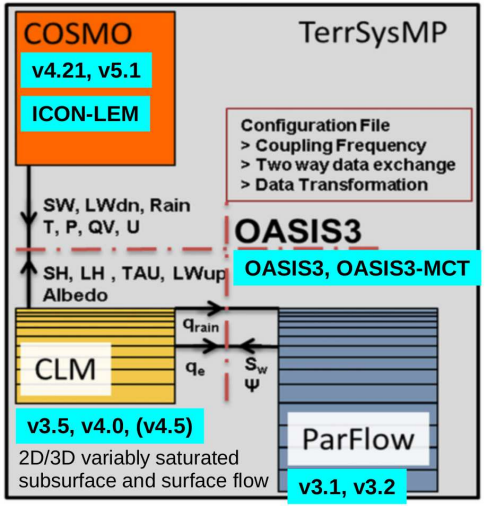
Joint Proposal for FY.2023 JHPCN under Eval.

https://jhpcn-kyoten.itc.u-tokyo.ac.jp/en/

- Innovative Computational Science by Integration of Simulation/Data/Learning under Heterogeneous Computing Environment
  - FY.2021 & 2022: Focused on Earthquake Simulations
    - Univ. Tokyo (ITC, ERI), Nagoya U., Kyushu U., NIES, Fujitsu
  - FY.2023-2025 (plan): Other applications and International Collaborations
    - Jülich Supercomputing Centre（JSC）：Modular Supercomputing
    - Rudjer Boskovic Institute, Centre for Informatics and Computing, Croatia
    - Friedrich-Alexander-Universität Erlangen-Nürnberg（FAU）
    - French Atomic Energy Commission (CEA)
- Target Systems in Japan
  - Wistreia/BDEC-01，Flow@Nagoya U., mdx

Applications from JSC & CEA

- JSC
  - Terrestrial Systems Modeling Platform (TSMP)
    - Coupling: Groundwater Flow & Atmosphere
    - https://www.terrsysmp.org/
  - Chebyshev Accelerated Subspace Eigensolver (ChASE)
    - Quantum Chemistry, Heterogeneous Environment
    - https://github.com/ChASE-library



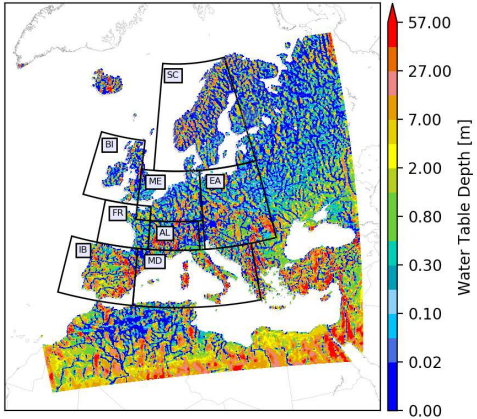
- CEA
  - Selection of inhibitors of the SARS-CoV-2 Main Protease
  - based on BigDFT
    - https://bigdft.org/
    - Already optimized for Fugaku/A64FX by CEA-R-CCS collaboration
    - France Boillod-Cerneux

ALGORITHM 1: The ChASE algorithm: SI plus our original contributions

Require: Hermitian matrix A, number of desired eigenpairs nev, threshold tolerance for residuals tol, initial polynomial degree deg, search space increment nex, approx and optim flags, vector matrix $\hat{V} \equiv [\hat{v}_1 \ \dots \ \hat{v}_{nev+nex}]$ and estimates μ_1 and $\mu_{nev+nex}$.

Ensure: nev extremal eigenpairs $(\Lambda, \hat{Y})$, with $\Lambda = [\lambda_1 \ \dots \ \lambda_{nev}]$ and $\hat{Y} \equiv [\hat{y}_1 \ \dots \ \hat{y}_{nev}]$, and their residuals $[\mathrm{Res}(\hat{y}_1, \lambda_1) \ \dots \ \mathrm{Res}(\hat{y}_{nev}, \lambda_{nev})]$

1: $m_{1:nev+nex} \leftarrow$ deg ▷ INITIAL CONSTANT DEGREE
2: $(b_{sup}, \mu_1, \mu_{nev+nex}, \hat{V}) \leftarrow$ LANCZOS(A, approx) ▷ BOUND AND OPTIONAL INPUT
3: **while** size($\hat{Y}$) < nev **do**
4: $\hat{V} \leftarrow$ FILTER(A, b_{sup}, μ_1, $\mu_{nev+nex}$, $\hat{V}$, m, optim) ▷ USES ARRAY OF DEGREES
5: $\hat{Q} \leftarrow$ ORTHOGONALIZE($[\hat{Y}\ \hat{V}]$) ▷ QR FACTORIZATION
6: $\hat{Q} \leftarrow \left[\hat{Q}_{:,\mathrm{size}(\hat{Y})+1} \ \dots \ \hat{Q}_{:,nev+nex}\right]$ ▷ REDUCE TO ACTIVE SUBSPACE
7: $(\hat{V}, \tilde{\Lambda}) \leftarrow$ RAYLEIGH-RITZ(A, $\hat{Q}$)
8: Compute the residuals $\mathrm{Res}(\hat{V}, \tilde{\Lambda})$
9: $(\hat{V}, \Lambda, \hat{Y}) \leftarrow$ DEFLATION & LOCKING($\hat{V}, \tilde{\Lambda}, \mathrm{Res}(\hat{V}, \tilde{\Lambda}), \hat{Y}$)
10: $\mu_1 \leftarrow \min([\Lambda\ \tilde{\Lambda}])$; $\mu_{nev+nex} \leftarrow \max([\Lambda\ \tilde{\Lambda}])$
11: $c \leftarrow \frac{b_{sup}+\mu_{nev+nex}}{2}$; $e \leftarrow \frac{b_{sup}-\mu_{nev+nex}}{2}$
12: **for** $a = 1 \rightarrow$ size($\hat{V}$) **do**
13: $m_a \leftarrow$ DEGREES(tol, Res ($\hat{V}_{:,a}, \tilde{\lambda}_a$), λ_a, c, e) ▷ COMPUTE POLYNOMIAL DEGREE
14: **end for**
15: Sort $\mathrm{Res}(\hat{V}, \tilde{\Lambda})$, $\hat{V}$, $\tilde{\Lambda}$, m according to m
16: **end while**



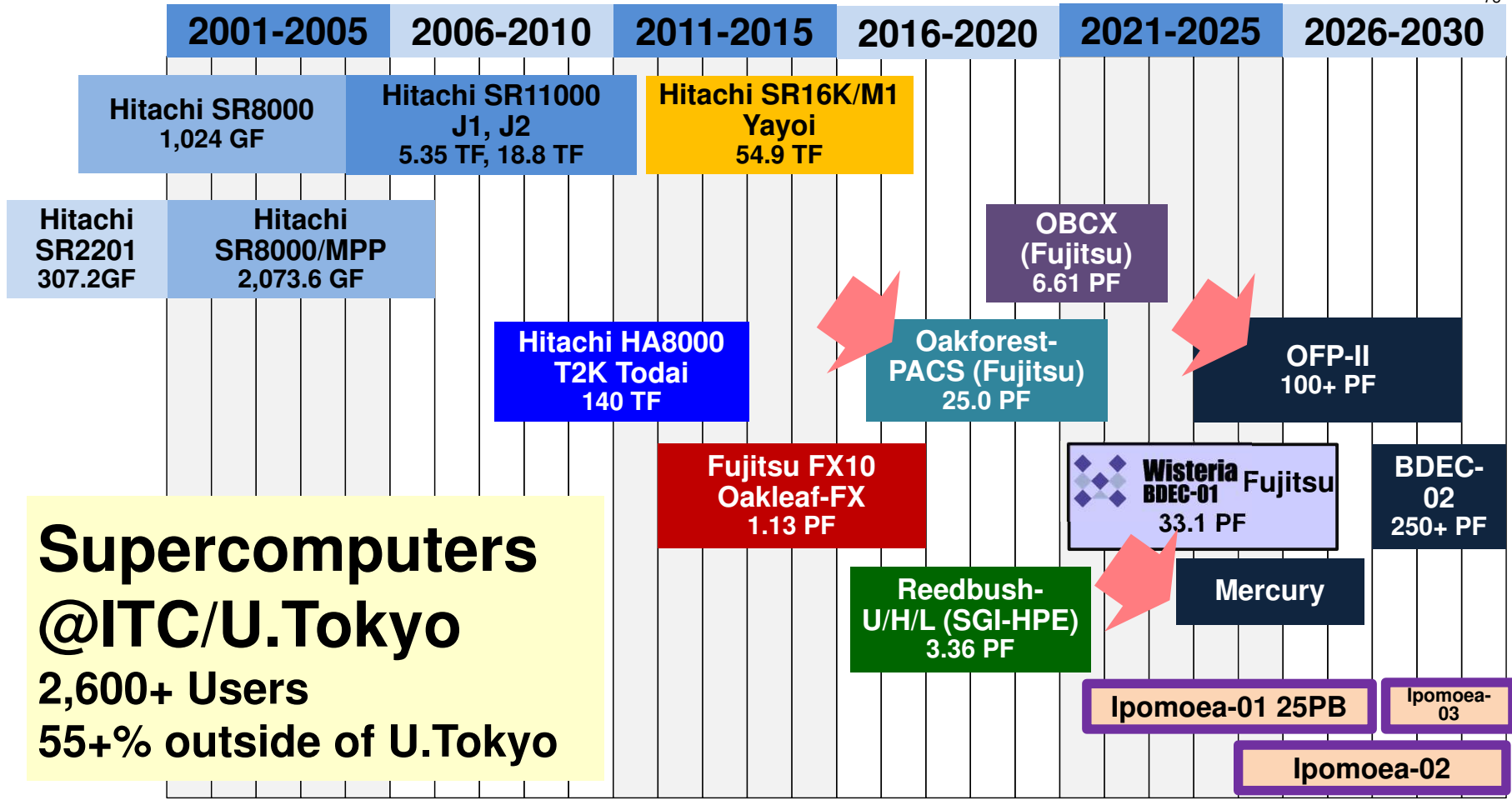
JCAHPC

http://jcahpc.jp/eng/index.html

筑波大学 University of Tsukuba　東京大学 THE UNIVERSITY OF TOKYO

JCAHPC

- Joint Center for Advanced High Performance Computing, since 2013
  - University of Tsukuba & University of Tokyo
    - Budgets of 2 Centers are combined
  - Promotion on Computational Science, Design/Procurement/Operation of Large-scale Systems
- Oakforest-PACS (OFP), 1st System of JCAHPC
  - 8,208 Intel Xeon Phi, 25PF, Fujitsu
  - Top500 (#6 (Nov.2016), #1 in Japan)
  - Retired in the end of March 2022 (#39 (Nov.2021))
- National Flagship System in FY.2019/2020
  - Between K and Fugaku


2001-2005
2006-2010
2011-2015
2016-2020
2021-2025
2026-2030
Hitachi SR8000
1,024 GF
Hitachi SR11000
J1, J2
5.35 TF, 18.8 TF
Hitachi SR16K/M1
Yayoi
54.9 TF
Hitachi
SR2201
307.2GF
Hitachi
SR8000/MPP
2,073.6 GF
OBCX
(Fujitsu)
6.61 PF
Hitachi HA8000
T2K Todai
140 TF
Oakforest-
PACS (Fujitsu)
25.0 PF
OFP-II
100+ PF
Fujitsu FX10
Oakleaf-FX
1.13 PF
Wisteria
BDEC-01
Fujitsu
33.1 PF
BDEC-
02
250+ PF
Reedbush-
U/H/L (SGI-HPE)
3.36 PF
Mercury
Ipomoea-01 25PB
Ipomoea-
03
Ipomoea-02
Supercomputers
@ITC/U.Tokyo
2,600+ Users
55+% outside of U.Tokyo

Future Perspective: OFP-II

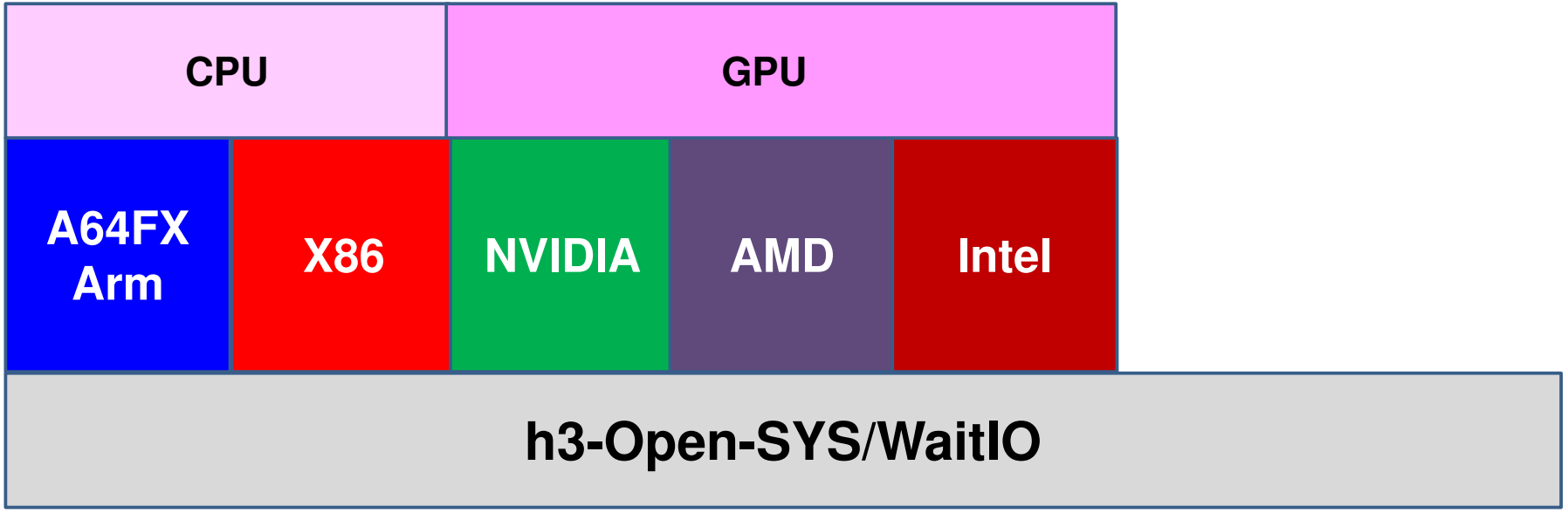
- U.Tokyo is shifting to GPUs/Accelerators in next 10 years
  - Maximum performance under constraint of power consumption
- OFP-II (April 2024 or later)
  - Successor of OFP
  - GPU Cluster (same
  - Integration of (Simu
  - Group-A (CPU+GPU), Group-B (Only CPU): CPUs in Group-A and Group-B could be different
- Wisteria-Mercury (October-November 2023)
  - Prototype of OFP-II

More Detailed Talk by Prof. T. Hanawa

Summary

- Wisteria/BDEC-01
- h3-Open-BDEC
  - h3-Open-SYS/WaitIO
  - h3-Open-UTIL/MP
- Examples of a little bit complex workflows
  - Earthquake Simulations + Real-Time Data Assimilation
  - Global Cloud Simulations
- Some Future Perspectives
  - International Collaborations
  - OFP-II

Anything is possible with WaitIO
WaitIO over Internet/cloud is possible



Anything is possible with WaitIO
WaitIO over Internet/cloud is possible