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

東京大学
THE UNIVERSITY OF TOKYO

SC21 St. Louis, MO | science & beyond.

Integration of (Simulation + Data + Learning) for Innovative Scientific Computing by h3-Open-BDEC on Wisteria/BDEC-01

Wisteria BDEC-01

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

Kengo Nakajima
Information Technology Center
The University of Tokyo

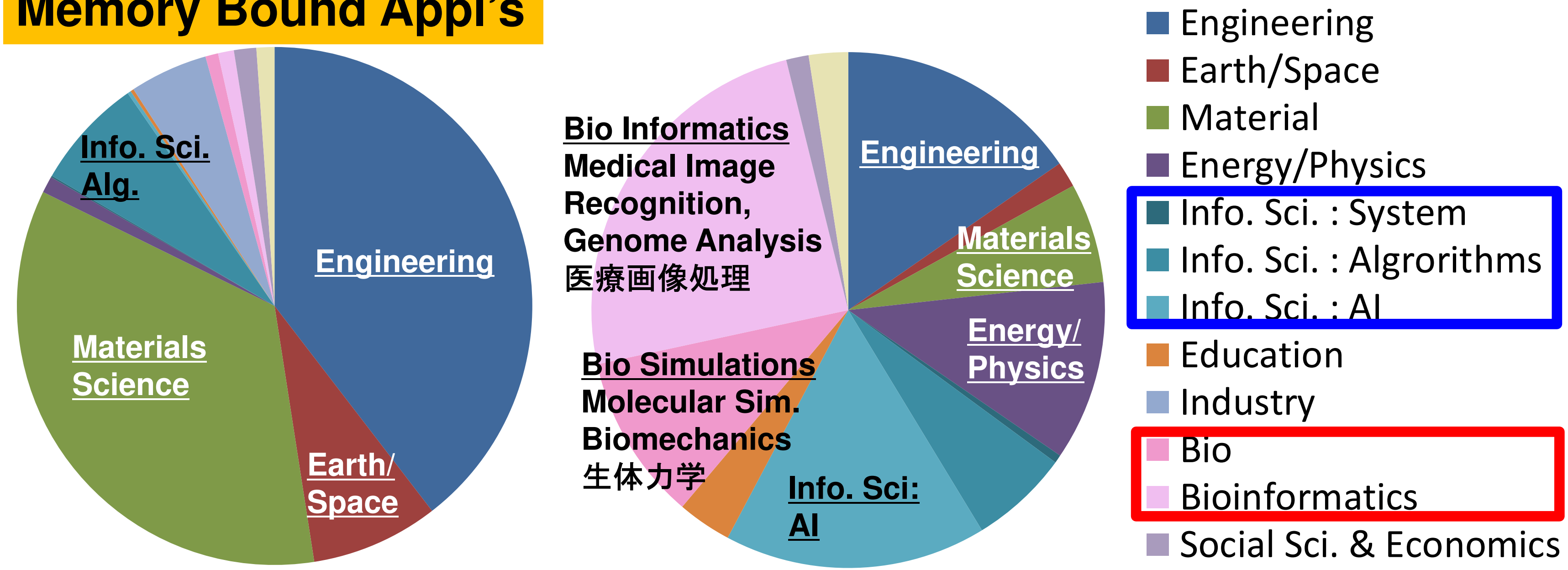
3 Systems at the end of March 2021

2,600+ users (55+% from outside of U.Tokyo)

- Reedbush (HPE, Intel BDW + NVIDIA P100 (Pascal))
  - Integrated Supercomputer Sys. for Data Analyses & Scientific Simulations
    - Jul.2016-Nov.2021 (Plan)
  - Our first GPU System, DDN IME (Burst Buffer)
  - Reedbush-U: CPU only, 420 nodes, 508 TF (Jul.2016~, retired June 2020)
  - Reedbush-H: 120 nodes, 2 GPUs/node: 1.42 PF (Mar.2017~Nov.2021)
  - Reedbush-L: 64 nodes, 4 GPUs/node: 1.43 PF (Oct.2017~Nov.2021)
- Oakforest-PACS (OFP) (Fujitsu, Intel Xeon Phi (KNL))
  - JCAHPC (U.Tsukuba & U.Tokyo)
  - 25 PF, #22 in 56th TOP 500 (November 2020) (#4 in Japan), Omni-Path Architecture, DDN IME (Burst Buffer), Sept.2016~Mar.2022
- Oakbridge-CX (OBCX) (Fujitsu, Intel Xeon Platinum 8280, CLX)
  - Massively Parallel Supercomputer System
  - 6.61 PF, #69 in 56th TOP 500, July 2019-June 2023
  - SSD's are installed to 128 nodes (out of 1,368)

Research Area based on CPU Hours (FY.2019)

Traditional CSE: Memory Bound Appl's

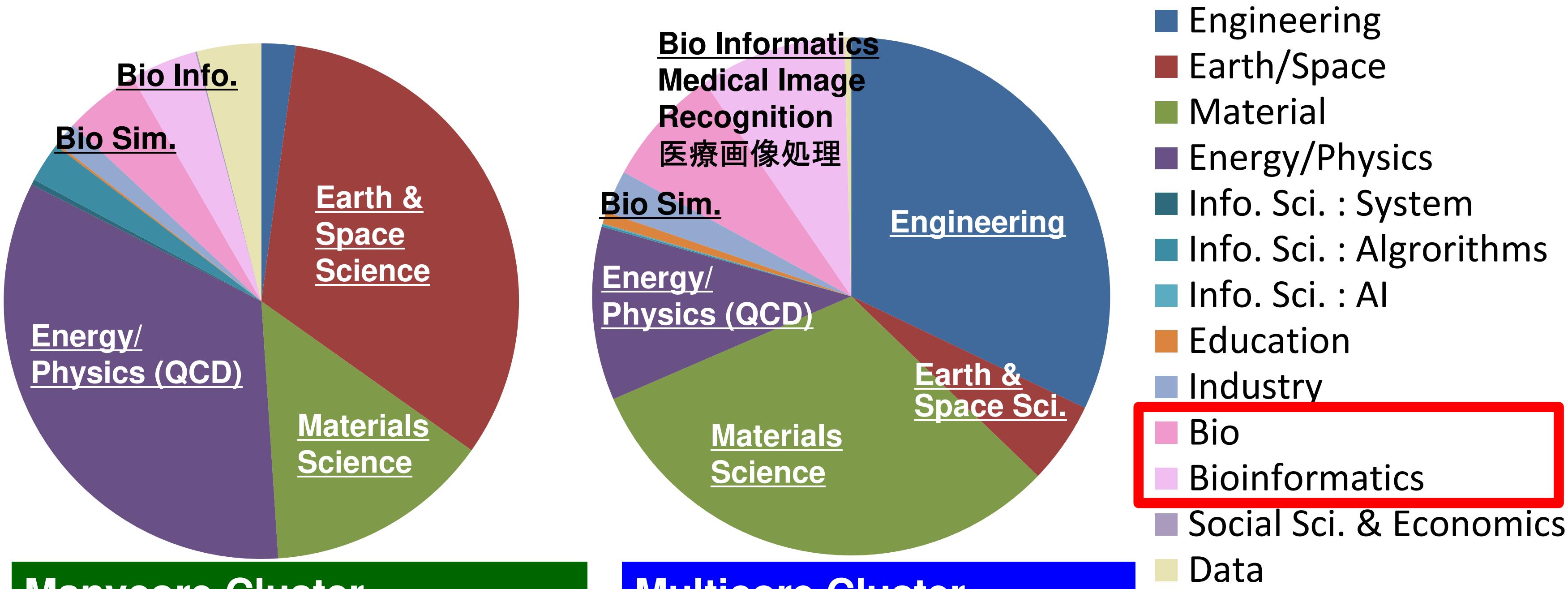


Multicore Cluster
Intel BDW Only
(Reedbush-U)

GPU Cluster
Intel BDW + NVIDIA P100
(Reedbush-H)

Research Area based on CPU Hours (FY.2019)

OBCX: October 2019~September 2020



Manycore Cluster
Intel Xeon Phi
(Oakforest-PACS) (OFP)

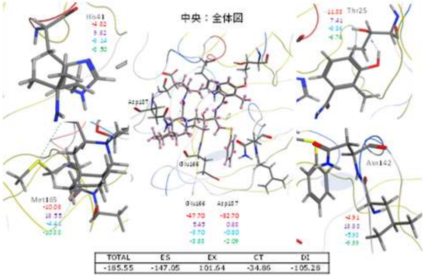
Multicore Cluster
Intel Xeon CLX
(Oakbridge-CX) (OBCX)

Traditional CSE:
Memory Bound Appl's

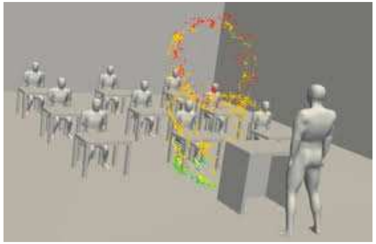
HPCI Urgent Call for Fighting against COVID-19 in Japan (FY.2020)

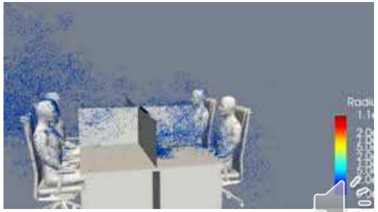
by 8 SC Centers of Natl. Univ., AIST etc.
6 of 14 accepted projects use U.Tokyo's Systems

Project Name	PI	System
Fragment molecular orbital calculations on the main protease of COVID-19	Yuji Mochizuki (Rikkyo U.)	
Study on the evaluation of arrhythmogenic risk of COVID-19 candidate drugs	Toshiaki Hisada (UT Heart)	OFP
Prediction of dynamical structure of Spike protein of SARS-COVID19	Yuji Sugita (RIKEN)	
Computer-assisted search for inhibitory agents for SARS-CoV-2	Tyuji Hoshino (Chiba U.)	
Prediction and Countermeasure for virus droplet Infection under Indoor Environment: Case studies for massively-parallel simulation on Fugaku	Makoto Tsubokura (Kobe U.)	OBCX
Spreading of polydisperse droplets in a turbulent puff of saturated exhaled air	Marco Edoardo Rosti (OIST)	



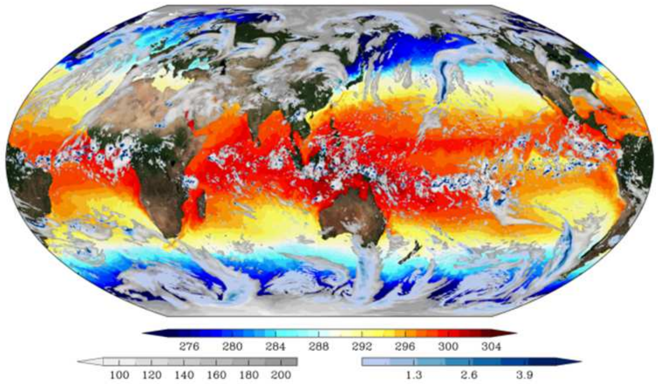
[c/o Prof. Y. Mochizuki (Rikkto U.)]

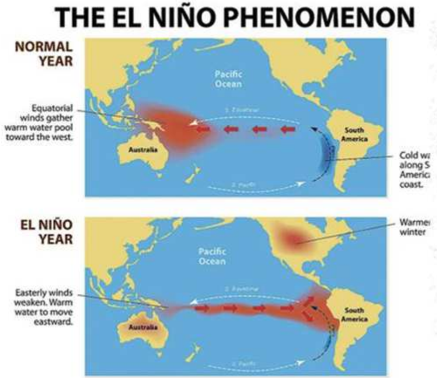


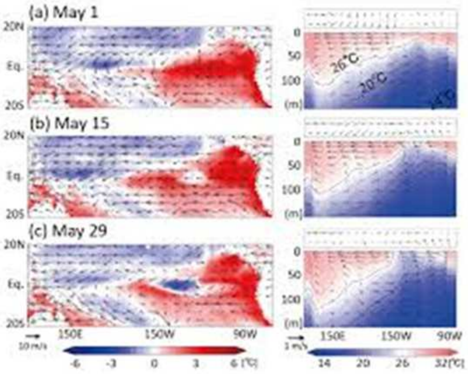


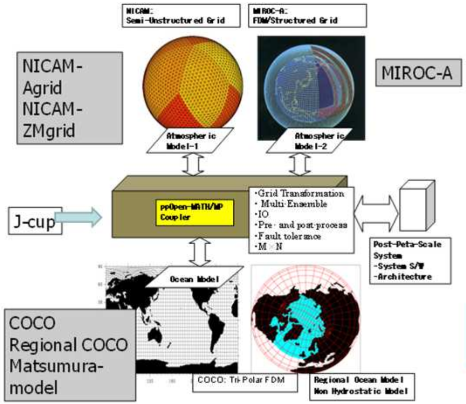
[c/o Prof. M.Tsubokura (Kobe U.)]

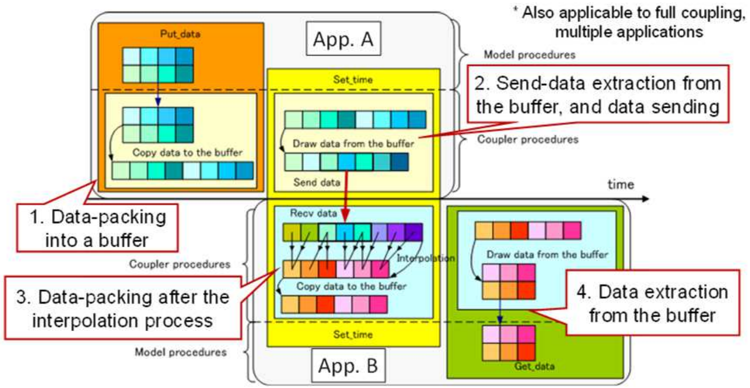
Global Atmosphere-Ocean Coupled Simulations

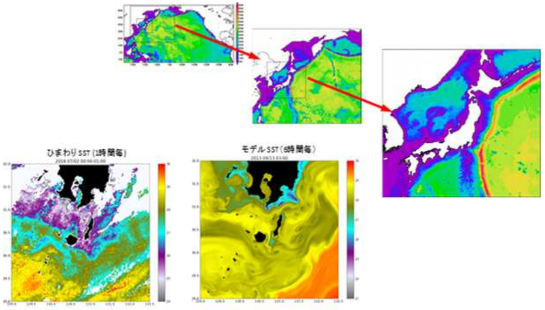








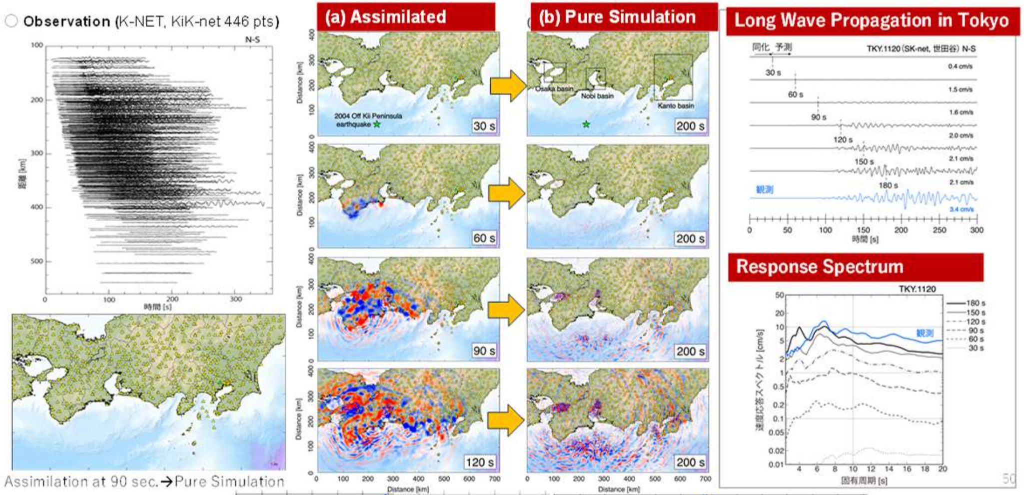


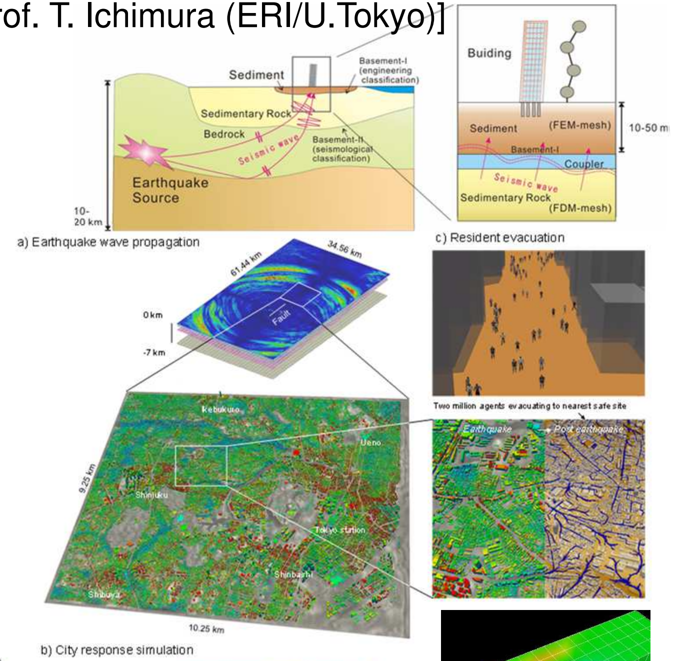


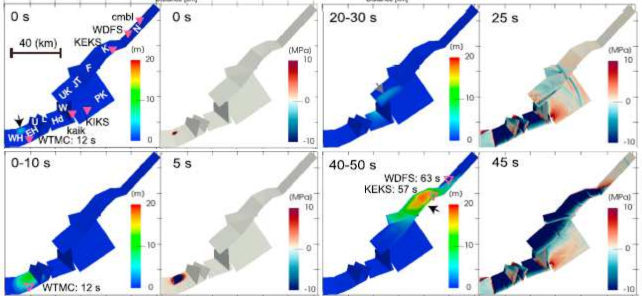
[c/o Prof. M. Sato, Prof. H. Hasumi (AORI/U.Tokyo)]

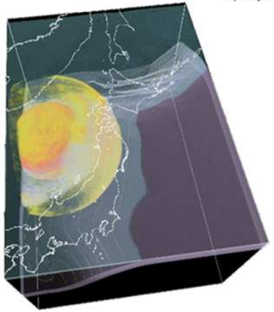
Solid Earth & Earthquake Simulations

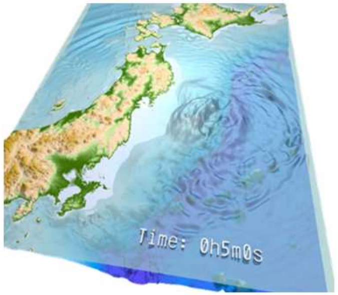
[c/o Prof. T. Furumura, Prof. T. Ichimura (ERI/U.Tokyo)]

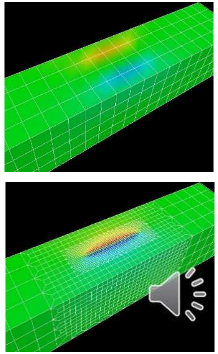








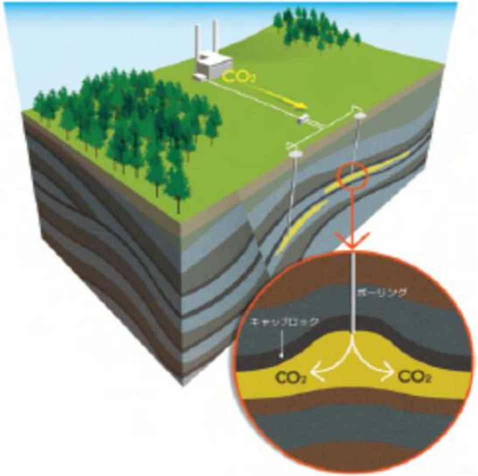




[c/o Prof. R. Ando (U.Tokyo)]

Simulation of Geologic CO_2 Storage

[c/o Dr. Hajime Yamamoto
(Taisei Corporation)]



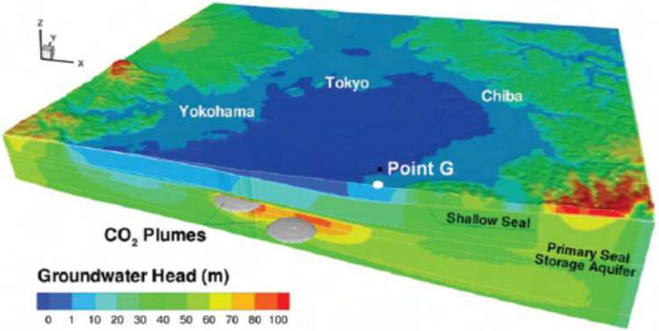
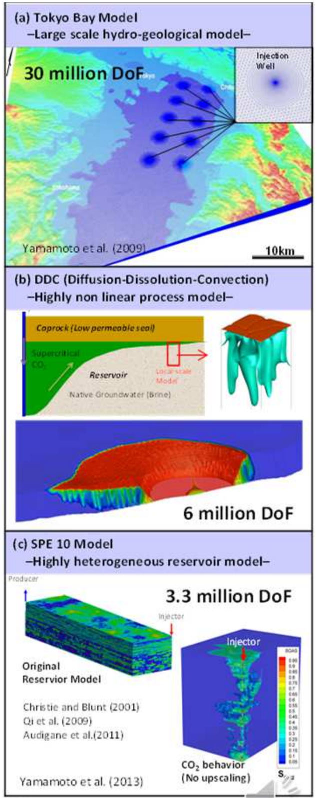
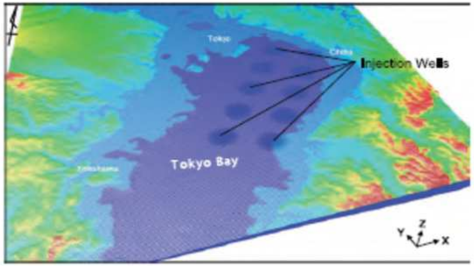
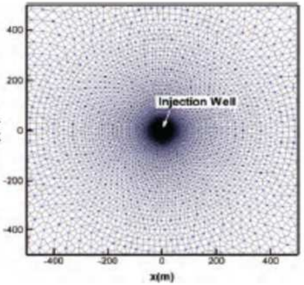


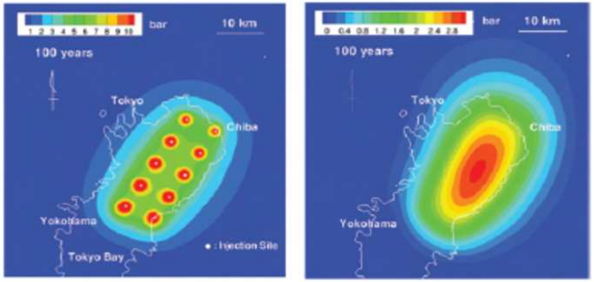
図 -4　CO_2 圧入後の地下水圧（全水頭換算）の分布（100 年後）



※DOF: degrees of freedom



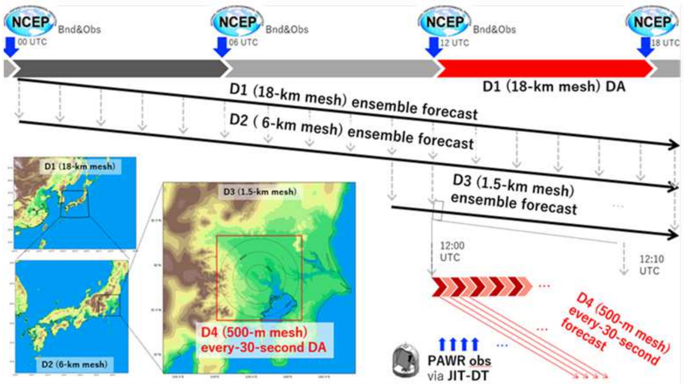




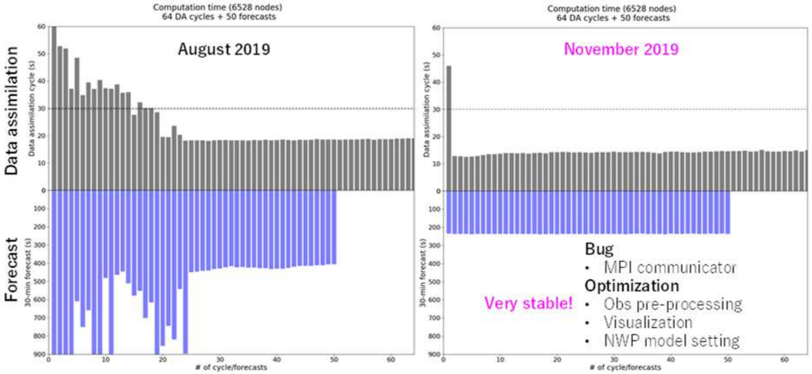
(a) 深部遮蔽層下面　　(b) 浅部遮蔽層下面

図 -5　圧力上昇量の平面分布（初期状態からの増分、圧入開始から 100 年後）

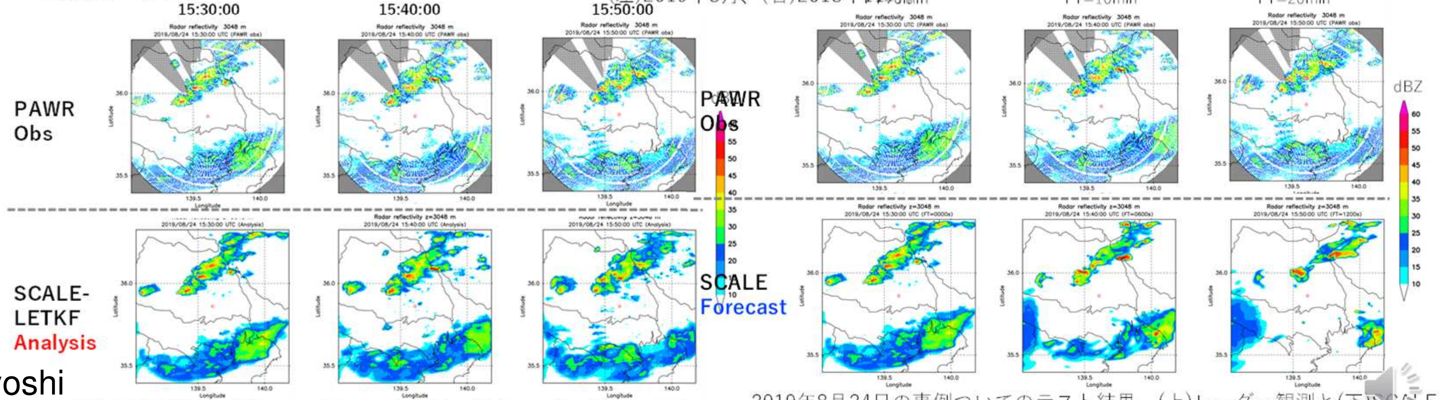
Real-Time Prediction of Severe Rainstorm by OFP



全体のワークフロー



計算性能の向上。上段はデータ同化、下段は30分予報にかかった時間(秒)。
(左)2019年8月、(右)2018年11月



2019年8月24日の事例ついてのテスト結果。(上)レーダー観測と(下)SCALE-LETKFによる解析で得られたレーダー反射強度(dBZ)を示す。

2019年8月24日の事例ついてのテスト結果。(上)レーダー観測と(下)SCALE-LETKFによる予報で得られたレーダー反射強度(dBZ)を示す。

[c/o Dr. Takemasa Miyoshi (RIKEN R-CCS)]

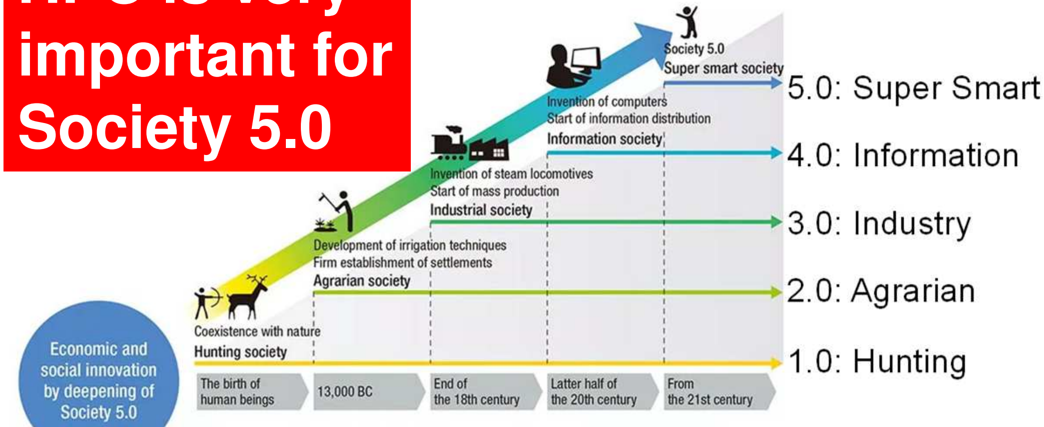
Society 5.0 & BDEC System

- We are developing an innovative method of computational science towards the Exascale Era/Society 5.0 by integration of (Simulation + Data + Learning (S+D+L)), where ideas of data science and machine learning are introduced to computational science

Society 5.0: the Cabinet Office of Japan

- Super Smart & Human-centered Society by Digital Innovation (IoT, Big Data, AI etc.) and by **Integration of Cyber Space & Physical Space**

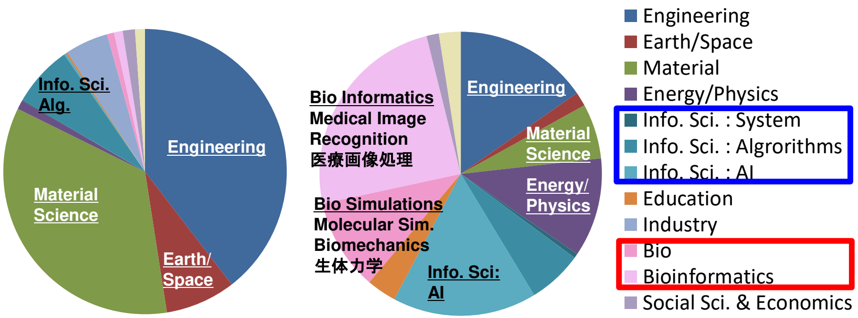
HPC is very important for Society 5.0



Source: Prepared based on materials from the Japan Business Federation (Keidanren)

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, Sophiscated Simulation**
  - **Planning started in 2015**



Multicore Cluster
Intel BDW Only
(Reedbush-U)

GPU Cluster
Intel BDW + NVIDIA P100
(Reedbush-H)

BDEC (Big Data & Extreme Computing)

S + D + L

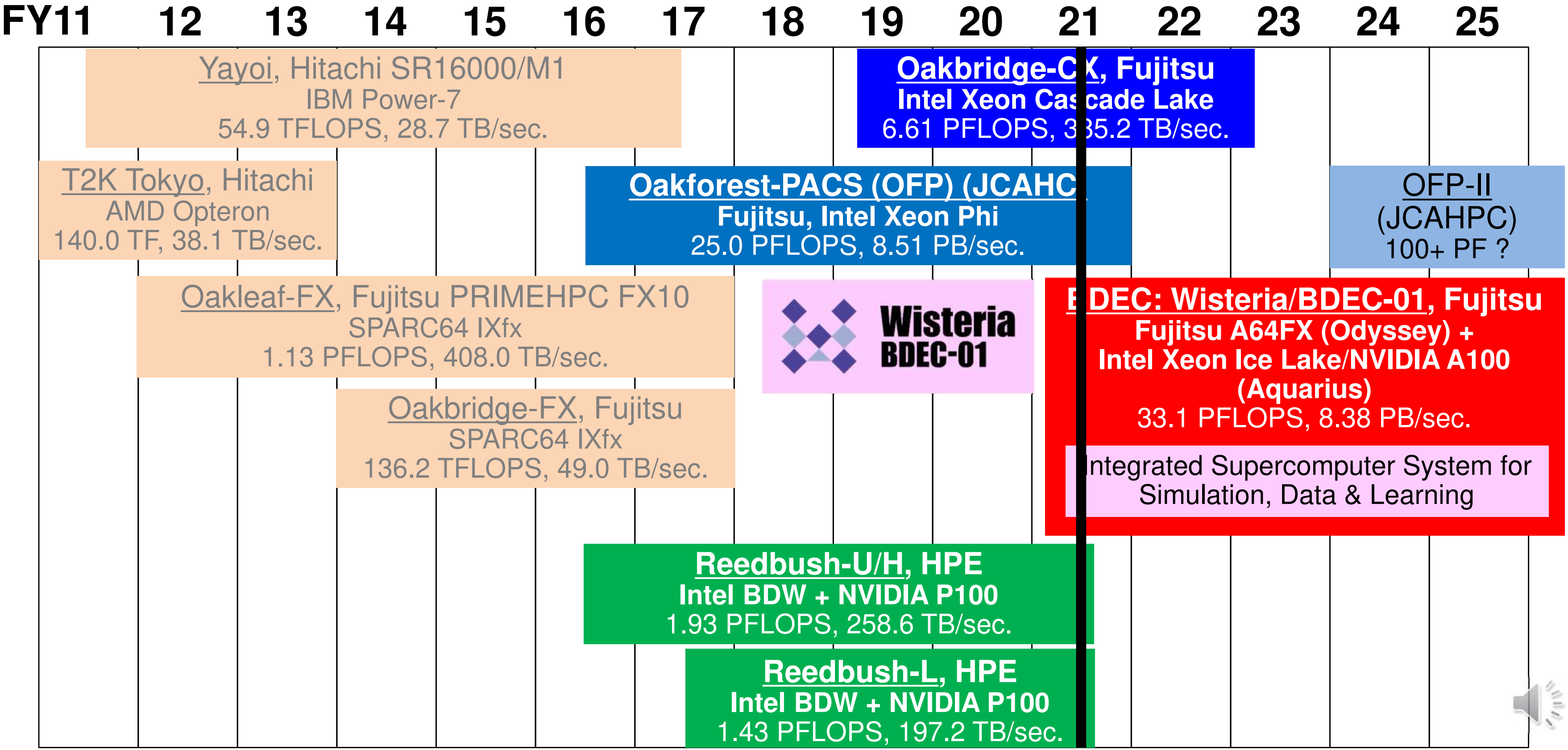
Society 5.0 & BDEC System

- We are developing an innovative method of computational science towards the Exascale Era/Society 5.0 by integration of (Simulation + Data + Learning (S+D+L)), where ideas of data science and machine learning are introduced to computational science
- **We have been operating 3 systems, and are now introducing the BDEC (Big Data & Extreme Computing) System as the Platform for Integration of (S+D+L)**
  - **Wisteria/BDEC-01 with 33.1 PF**

Wisteria
BDEC-01

Supercomputers in ITC/U.Tokyo

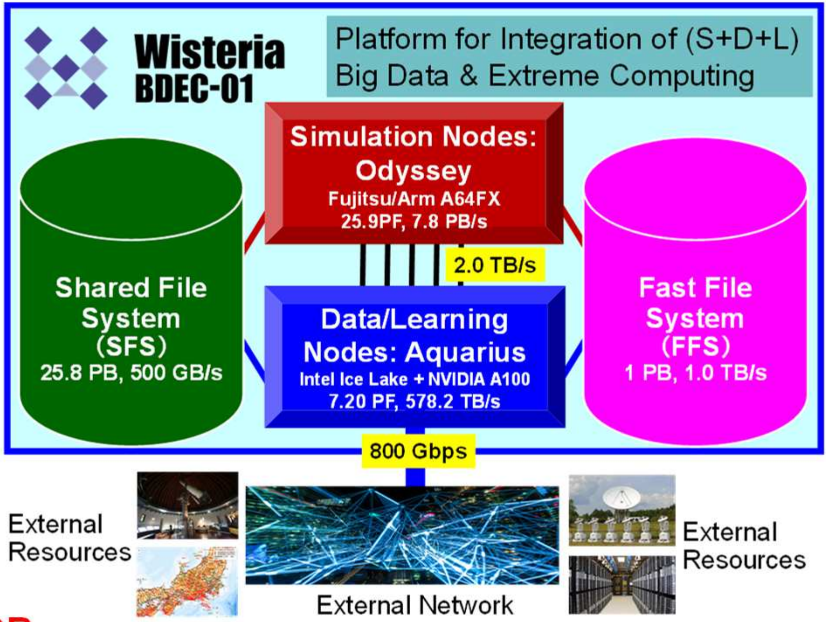
Information Technology Center, The University of Tokyo



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, ~360m²
- **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)

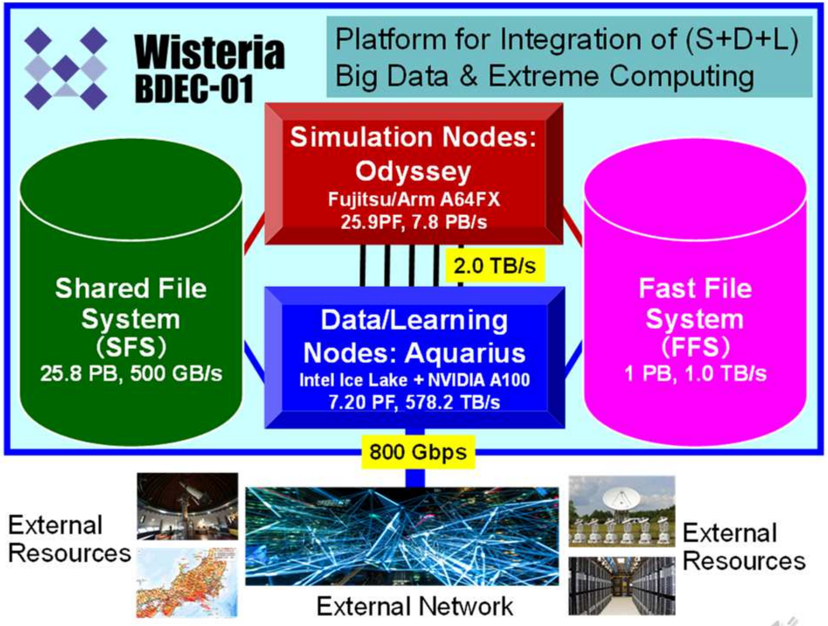


Rankings@ISC 2021 June 2021

ISC High Performance
The HPC Event.

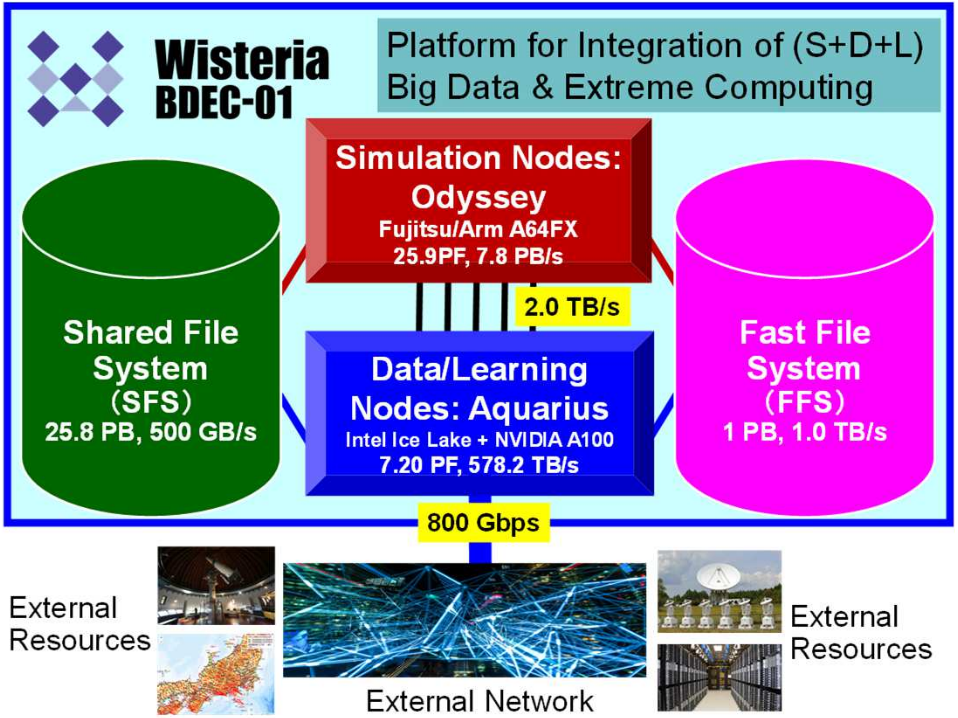
ISC HIGH PERFORMANCE 2021 DIGITAL
JUNE 24 - JULY 2, 2021
ISC-HPC.COM

	Odyssey	Aquarius
TOP 500	13	94
Green 500	21	10
HPCG	9	53
Graph 500 BFS	3	-
HPL-AI	7	-



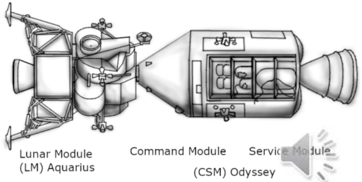
Wisteria/BDEC-01

Platform for Integration of (S+D+L)

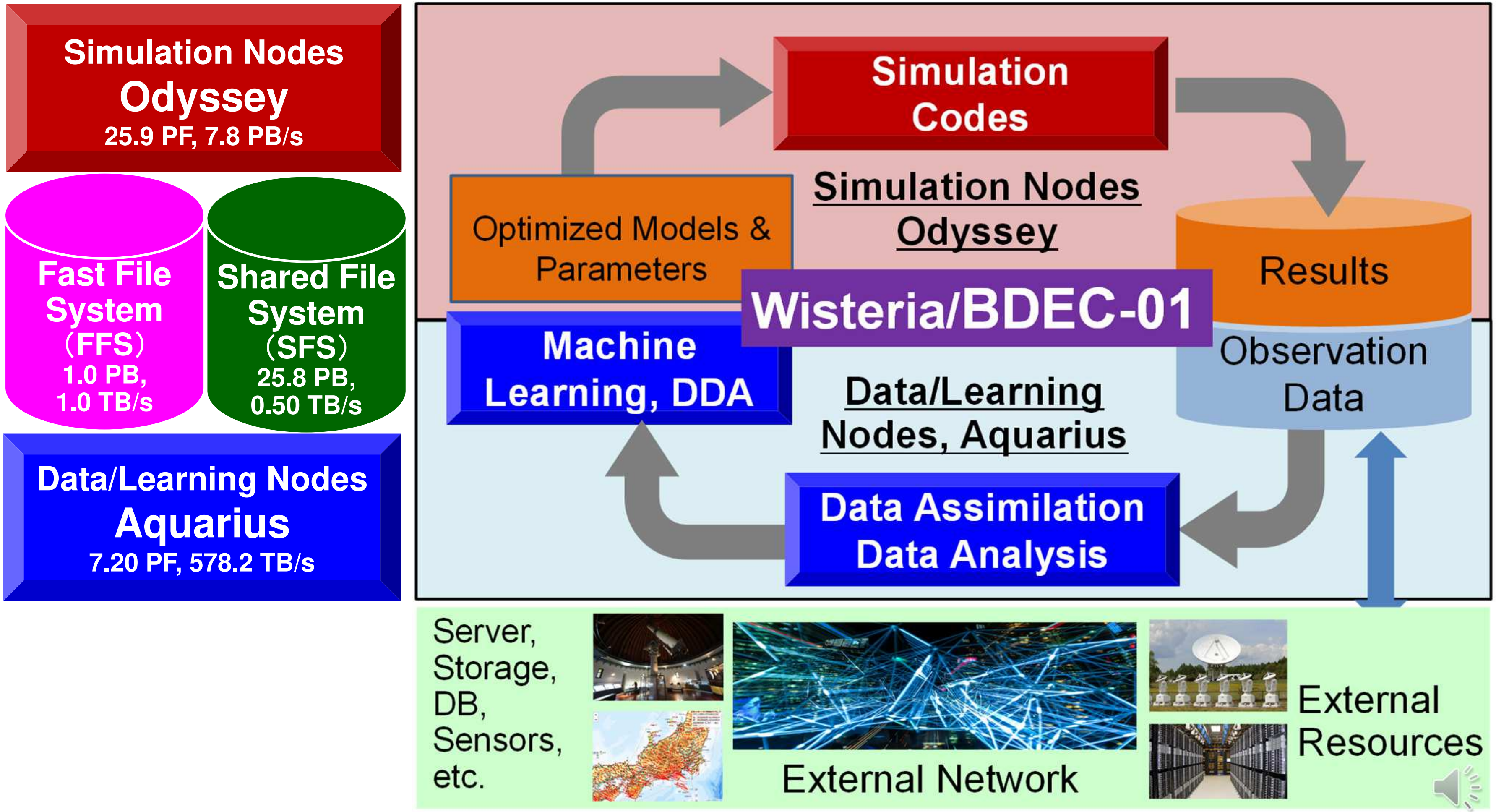


- Wisteria（紫藤）
  - "Legend of Princess Wisteria" at Lake Teganuma in Kashiwa
- Odyssey
  - Callsign of Apollo 13's Command Module (CM)
- Aquarius
  - Callsign of Apollo 13's Luna Module (LM)

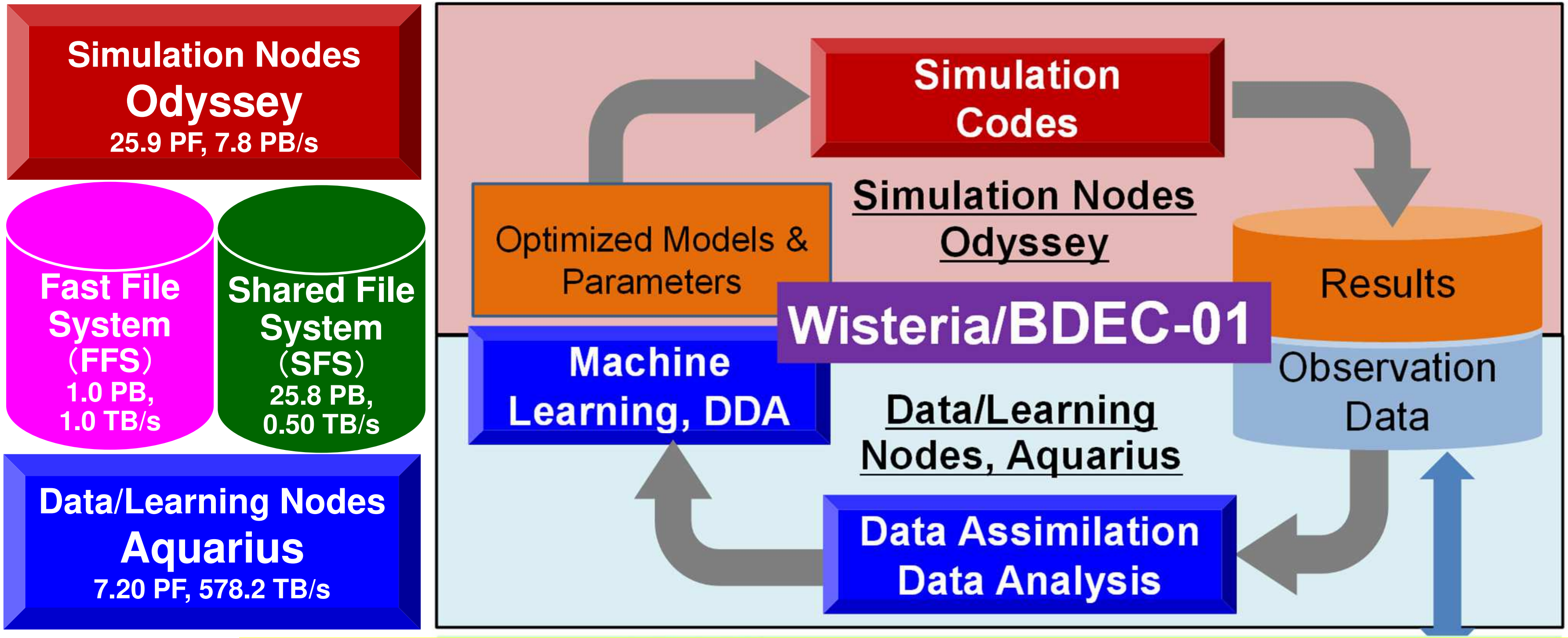
Aquarius Odyssey



https://www.cc.u-tokyo.ac.jp/public/pr/pr-wisteria.php



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)

h3-Open-BDEC on BDEC System

- We are developing an innovative method of computational science towards the Exascale Era/Society 5.0 by integration of (Simulation + Data + Learning (S+D+L)), where ideas of data science and machine learning are introduced to computational science
- We have been operating 3 systems, and are now introducing the BDEC (Big Data & Extreme Computing) System, Wisteria/BDEC-01, with 33.1 PF as the Platform for Integration of (S+D+L)
- **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 through JSPS Grant-in-Aid for Scientific Research (S) since 2019**
  - **Leading-PI: Kengo Nakajima (The University of Tokyo)**
  - **Total Budget: 152.7M JPY= 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), Co-PI, Container

h3-Open-BDEC

Innovative Software Platform for Integration of (S+D+L) on BDEC

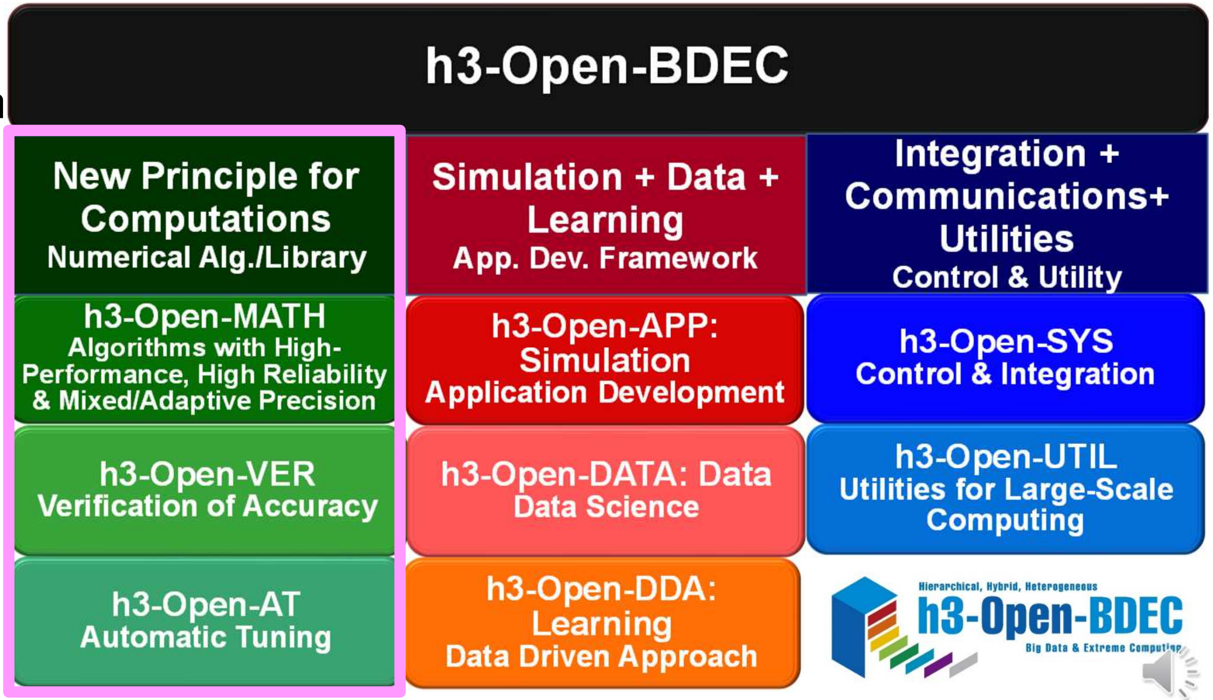
h3-Open-BDEC

New Principle for Computations Numerical Alg./Library	Simulation + Data + Learning App. Dev. Framework	Integration + Communications+ Utilities Control & Utility
h3-Open-MATH Algorithms with High-Performance, High Reliability & Mixed/Adaptive Precision	**h3-Open-APP: Simulation** Application Development	**h3-Open-SYS** Control & Integration
h3-Open-VER Verification of Accuracy	**h3-Open-DATA: Data** Data Science	**h3-Open-UTIL** Utilities for Large-Scale Computing
h3-Open-AT Automatic Tuning	**h3-Open-DDA: Learning** Data Driven Approach	Hierarchical, Hybrid, Heterogeneous h3-Open-BDEC Big Data & Extreme Computing

h3-Open-BDEC: Two Significant Innovations

① Methods for Numerical Analysis with High-Performance/High-Reliability/Power-Saving based on the New Principle of Computing by

- ✓ Adaptive Precision
- ✓ Accuracy Verification
- ✓ Automatic Tuning

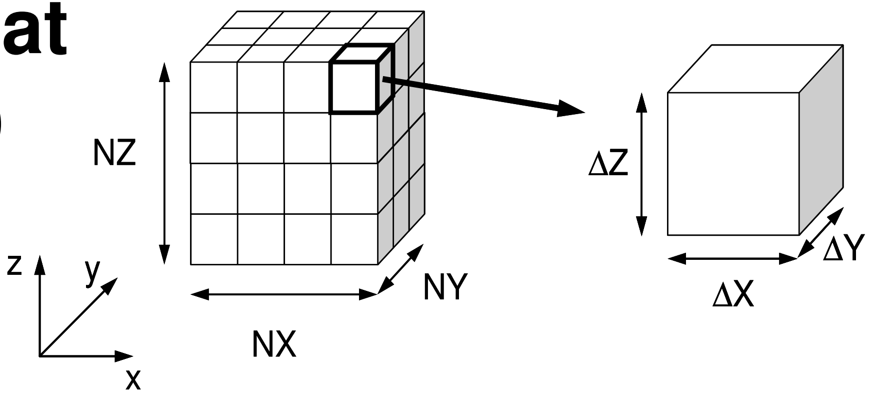


Approximate Computing with Low/Adaptive/Trans Precision

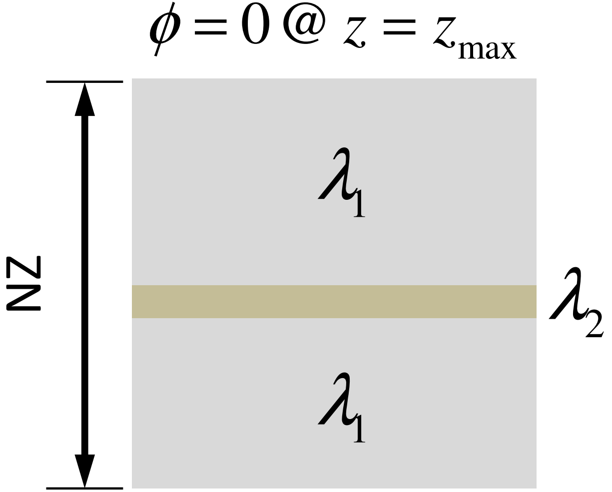
- Mostly, scientific computing has been conducted using FP64 (double precision, DP)
  - Sometimes, problems can be solved by FP32 (single precision, SP) or lower precision
- **Lower precision may save time, energy and memory**
- Approximate Computing
  - Originally for image recognition etc. where accuracy is not necessarily required
  - Also applied to numerical computations
- Computations by lower precision and by mixed precision may provide results with less accuracy

P3D: Steady State 3D Heat Conduction by FVM (1/2)

$$\nabla \cdot (\lambda \nabla \phi) + f = 0$$

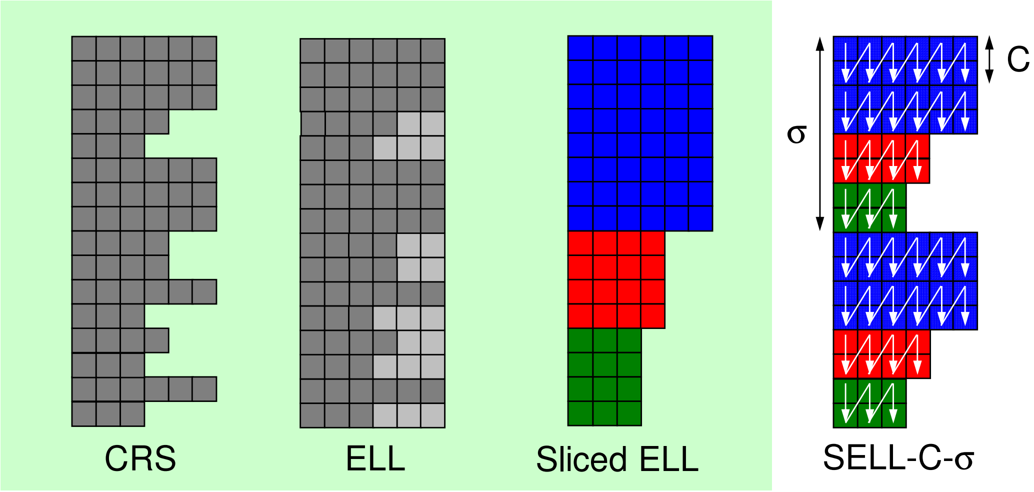


- 7-point Stencil
- Heterogenous Material Property
  - λ_1/λ_2 is proportional to the condition number of coefficient matrices
- Coefficient Matrix
  - Sparse, SPD
- ICCG Solver
- Fortran 90 + OpenMP
- CM-RCM Reordering



P3D: Steady State 3D Heat Conduction by FVM (2/2)

- Various Configurations
  - FP64 (Double), FP32 (Single), FP16 (Half) (just for preconditioning)
  - Matrix Storage Format (CRS, ELL, SELL-C-σ etc.)

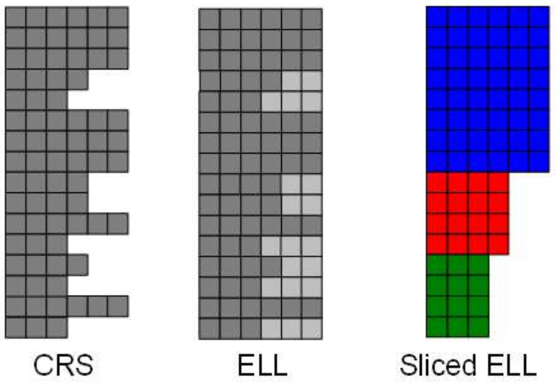


System	Oakforet-PACS	Oakbridge-CX	Oakleaf-7（FX700）
Abbreviation	OFP	OBCX	OL7
Architecture of CPU	Intel Xeon Phi 7250 (Knights Landing, KNL)	Intel Xeon Platinum 8280 (Cascade Lake, CLX)	Fujitsu A64FX（1.8GHz）
Core#/Socket	68	28	48
Socket#/Node	1	2	1
Peak Performance (DP) (GFLOPS)/Node	3,046	4,838	2,765
Memory Capacity (GB)/Node	MCDRAM: 16 DDR4: 96	192	32
Memory Bandwidth (GB/sec), Stream Triad	MCDRAM: 490 DDR4: 84.5	202	809
Compiler	Intel Parallel Studio 2019		Fujitsu FCC 4.0.0

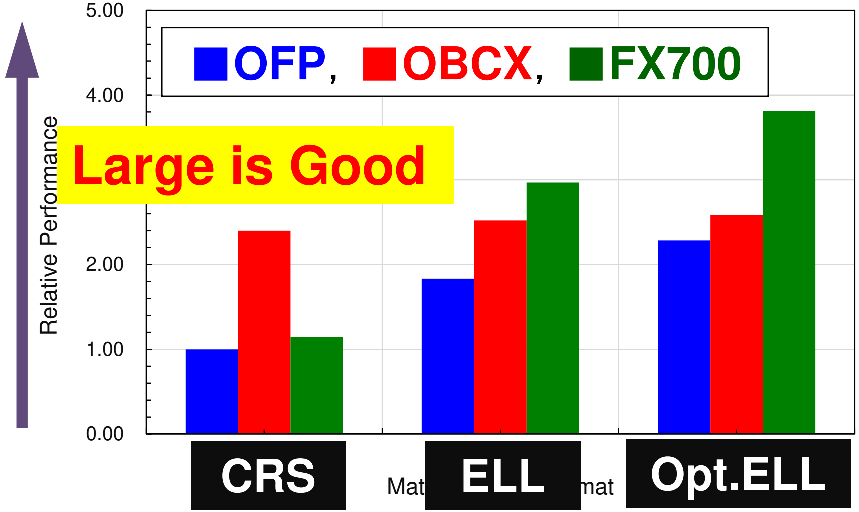
Ratio of Performance

Elapsed Computation Time for ICCG (DP), Normalized by OFP with CRS, $\lambda_1 / \lambda_2 = 1$

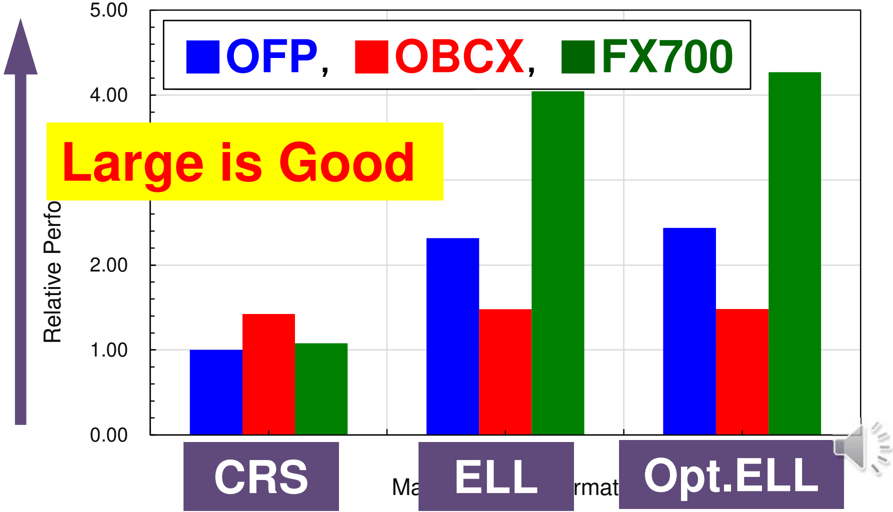
[KN et al. SWoPP 2020]



Medium : 128^3



Large : 256^3



Mixed Precision Computing of ICCG Solver for P3D on FX700

	SpMV, DAXPY, Dot Products	Preconditioning	Vectors for Preconditioning
D-D	FP64	FP64	FP64
D-S	FP64	FP32	FP32
D-H	FP64	FP16	FP32
S-S	FP32	FP32	FP32
S-H	FP32	FP16	FP32

[KN et al. SWoPP 2020]

Mixed Precision Computing for P3D on FX700 Implementation of Forward Substitution (CRS) in ICCG

[KN et al. SWoPP 2020]

FP64
FP32
FP16

D-S

```
!$omp parallel do private(ip,i)
        do ip= 1, PEsmpTOT
        do i= SMPindex((ip-1)*NCOLORtot)+1, SMPindex(ip*NCOLORtot)
          Ws(I,Z)= W(I,R)
        enddo
        enddo

!$omp parallel private(ic,ip,ip1,I,WVALs,k)
        do ic= 1, NCOLORtot
!$omp do
          do ip= 1, PEsmpTOT
            ip1= (ip-1)*NCOLORtot + ic
          do i= SMPindex(ip1-1)+1, SMPindex(ip1)
            WVALs= Ws(i,Z)
            do k= indexL(i-1)+1, indexL(i)
              WVALs= WVALs - ALs(k) * Ws(itemL(k),Z)
            enddo
            Ws(i,Z)= WVALs * Ws(i,DD)
          enddo
          enddo
        enddo
!$omp end parallel

(Backward Substitution)

!$omp parallel do private(ip,i)
        do ip= 1, PEsmpTOT
        do i= SMPindex((ip-1)*NCOLORtot)+1, SMPindex(ip*NCOLORtot)
          W(I,Z)= Ws(I,Z)
        enddo
        enddo
```

S-H

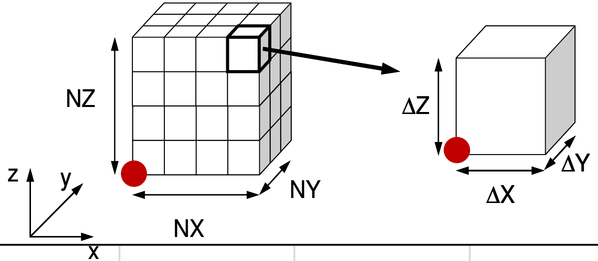
```
!$omp parallel do private(ip,i)
        do ip= 1, PEsmpTOT
        do i= SMPindex((ip-1)*NCOLORtot)+1, SMPindex(ip*NCOLORtot)
          Ws(i,Z)= Ws(i,R)
        enddo
        enddo

!$omp parallel private(ic,ip,ip1,i,WVALs,k)
        do ic= 1, NCOLORtot
!$omp do
          do ip= 1, PEsmpTOT
            ip1= (ip-1)*NCOLORtot + ic
          do i= SMPindex(ip1-1)+1, SMPindex(ip1)
            WVALs= Ws(i,Z)
            do k= indexL(i-1)+1, indexL(i)
              WVALs= WVALs - ALh(k) * Ws(itemL(k),Z)
            enddo
            Ws(i,Z)= WVALs * Wh(i,DD)
          enddo
          enddo
        enddo
!$omp end parallel
```

D-H

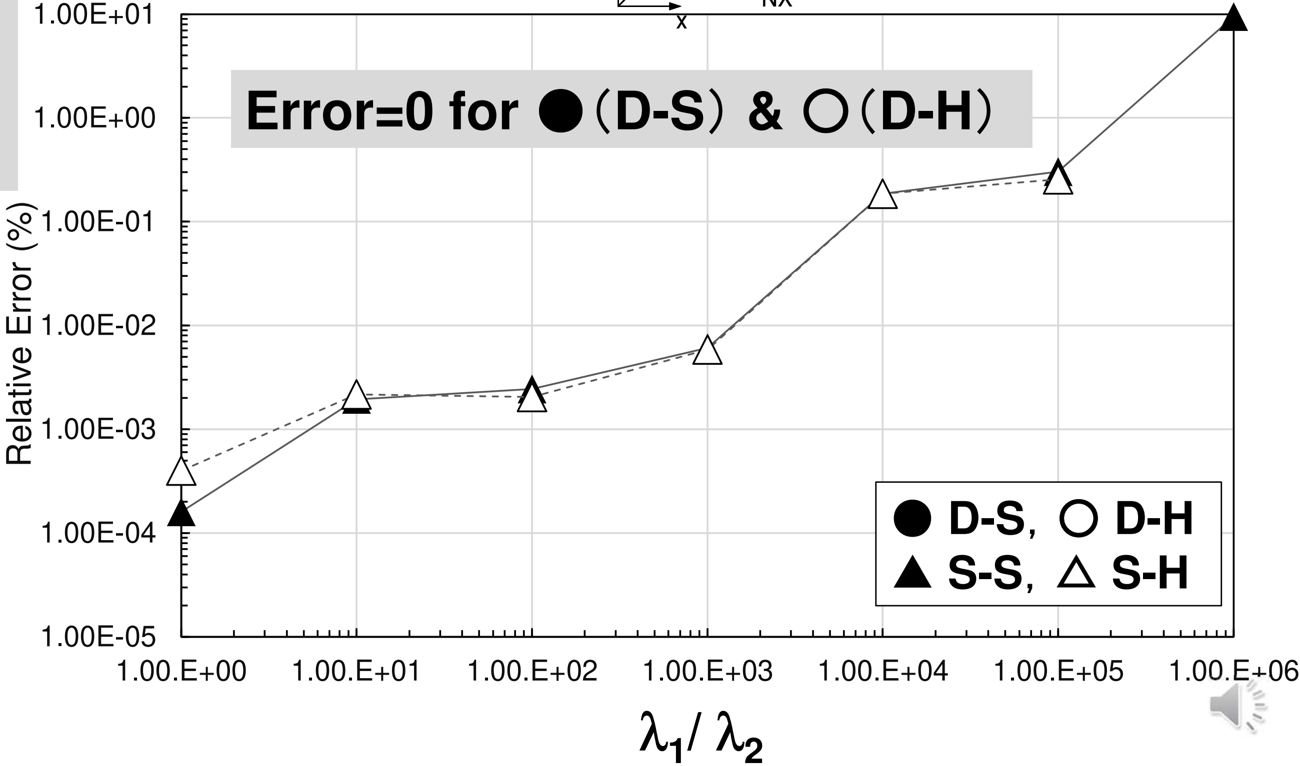
```
!$omp parallel do private(ip,i)
        do ip= 1, PEsmpTOT
        do i= SMPindex((ip-1)*NCOLORtot)+1, SMPindex(ip*NCOLORtot)
          Ws(i,Z)= W(i,R)
        enddo
        enddo

!$omp parallel private(ic,ip,ip1,i,WVALs,k)
        do ic= 1, NCOLORtot
!$omp do
          do ip= 1, PEsmpTOT
            ip1= (ip-1)*NCOLORtot + ic
          do i= SMPindex(ip1-1)+1, SMPindex(ip1)
            WVALs= Ws(i,Z)
            do k= indexL(i-1)+1, indexL(i)
              WVALs= WVALs - ALh(k) * Ws(itemL(k),Z)
            enddo
            Ws(i,Z)= WVALs * Wh(i,DD)
          enddo
          enddo
        enddo
!$omp end parallel

(Backward Substitution)

!$omp parallel do private(ip,i)
        do ip= 1, PEsmpTOT
        do i= SMPindex((ip-1)*NCOLORtot)+1, SMPindex(ip*NCOLORtot)
          W(I,Z)= Ws(I,Z)
        enddo
        enddo
```

Mixed Precision Computation

D-H/S-H do not converge at $\lambda_1/\lambda_2 = 10^6$



Relative Error (%) compared to D-D @●



[KN et al. SWoPP 2020]

Mixed Precision Computation

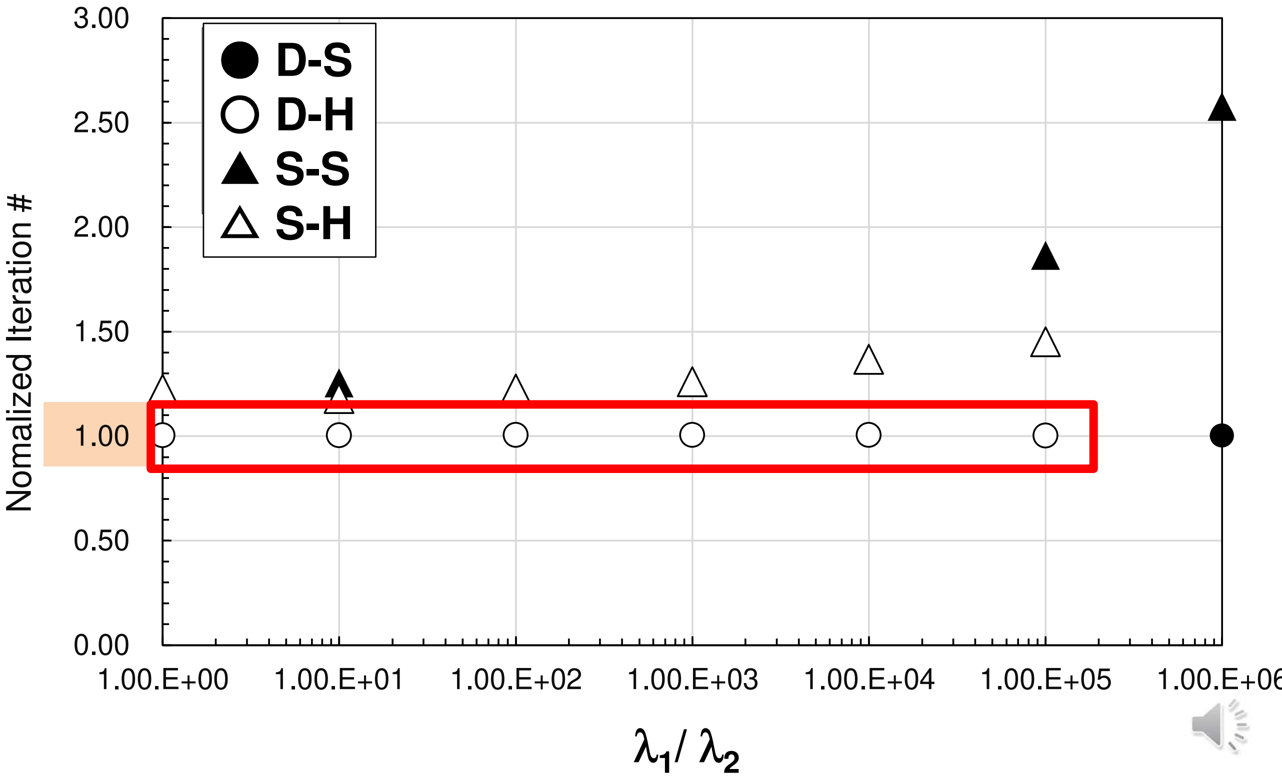
[KN et al. SWoPP 2020]

D-H/S-H do not converge at $\lambda_1/\lambda_2 = 10^6$

Number of Iterations (Normalized by that of D-D)

●～○～D-D, ▲～△

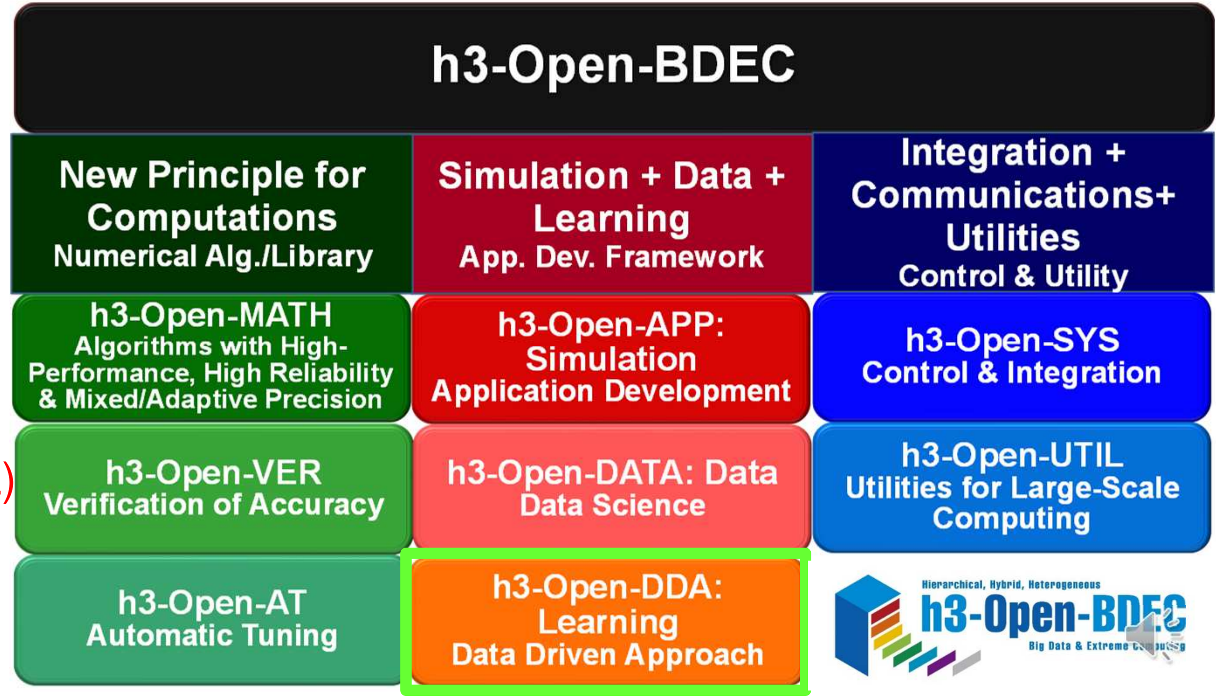
Results of (D-S, D-H) agree with those of D-D (if $\lambda_1/\lambda_2 \leqq 10^5$)



h3-Open-BDEC: Two Significant Innovations

① Methods for Numerical Analysis with High-Performance/High-Reliability/Power-Saving based on the New Principle of Computing by
- ✓ Adaptive Precision
- ✓ Accuracy Verification
- ✓ Automatic Tuning

② Hierarchical Data Driven Approach (*h*DDA) based on machine learning
- ✓ Integration of (S+D+L)
  AI for HPC

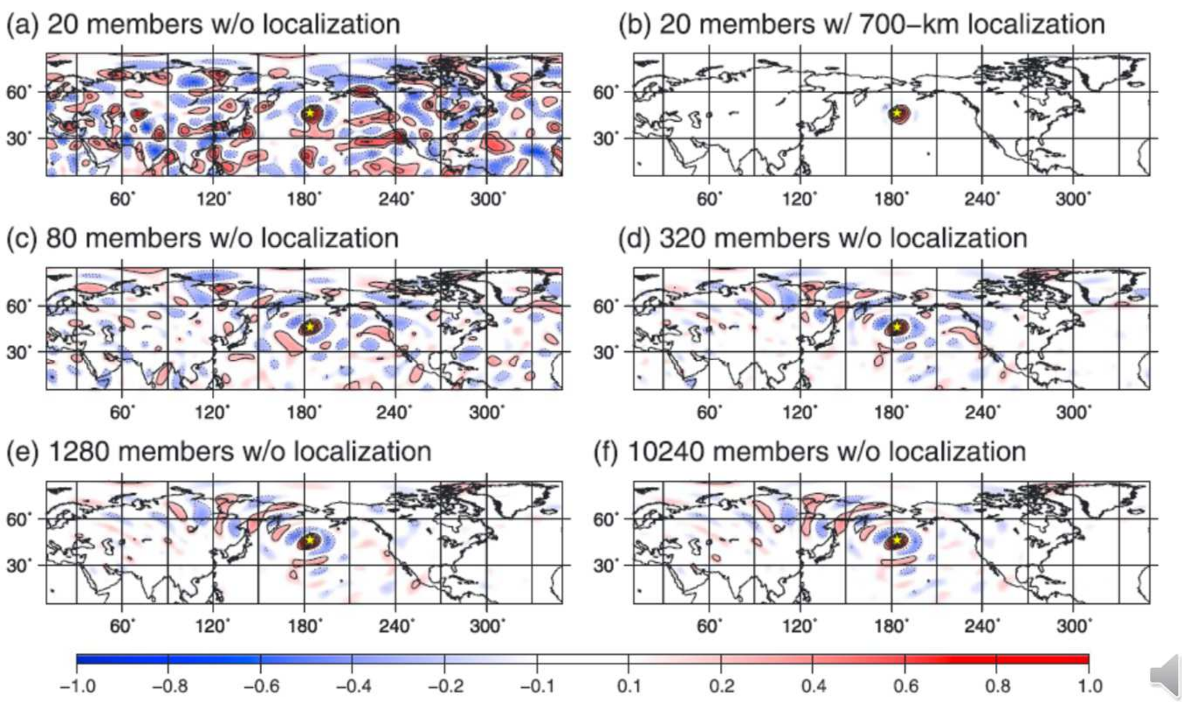


Real-World Scientific Simulations

- Non-Linear: Huge Number of Parameter Studies needed
  - ✓ Reduction of cases is very crucial

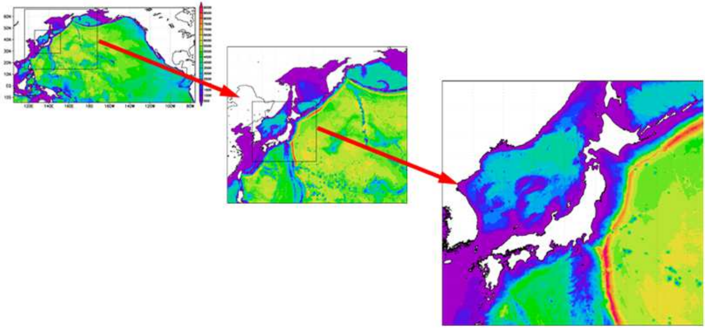
[Miyoshi et al. 2014]

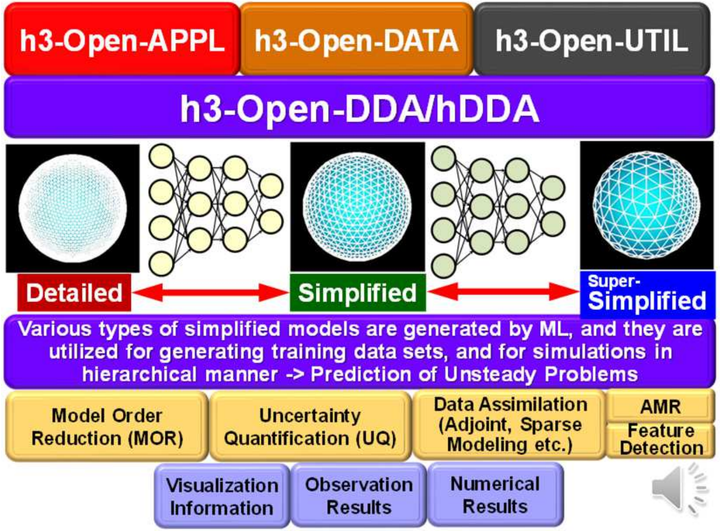
- Data Assimilation
  - ✓ Mid-Range Weather Prediction: 50-100 Ensemble Cases, 1,000 needed for accurate solution.
  - ✓ 50-100 (or fewer) may be enough for accurate solution, if opt. parameters are selected (e.g. by ML),



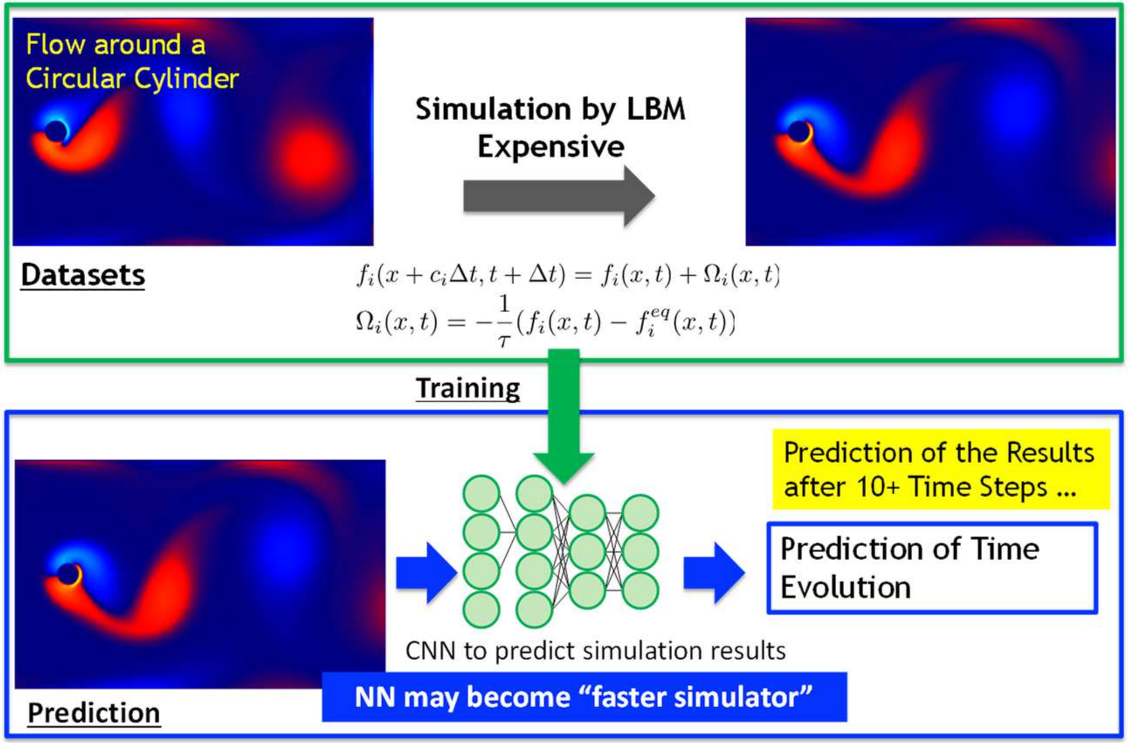
Hierarchical Data Driven Approach: *h*DDA

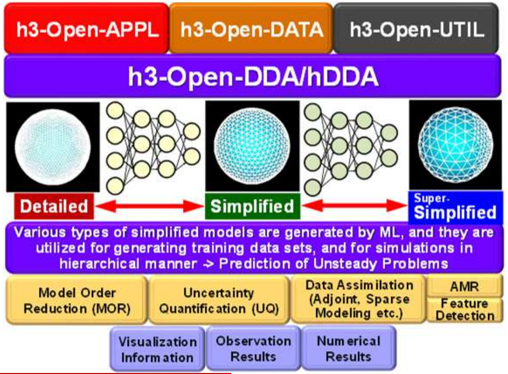
- Data Driven Approach (DDA)
  - Technique of AI/ML is introduced for predicting the results of simulations with different parameters.
  - DDA generally requires O(10^3-10^4) runs for generation of training data.
- ***h*DDA (Hierarchical DDA)**
  - Simplified models with coarser meshes (but preserving original features of physics) for efficient training are constructed automatically by Machine Learning using:
    - Feature Detection, AMR
    - MOR (Model Order Reduction)
    - UQ (Uncertainty Quantification)
    - Sparse Modeling



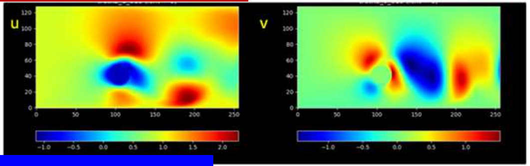


Acceleration of Transient CFD Simulations using ML/CNN Integration of (S+D+L), AI for HPC

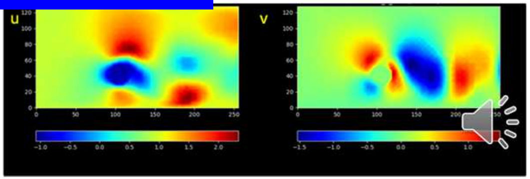




Simulations: LBM



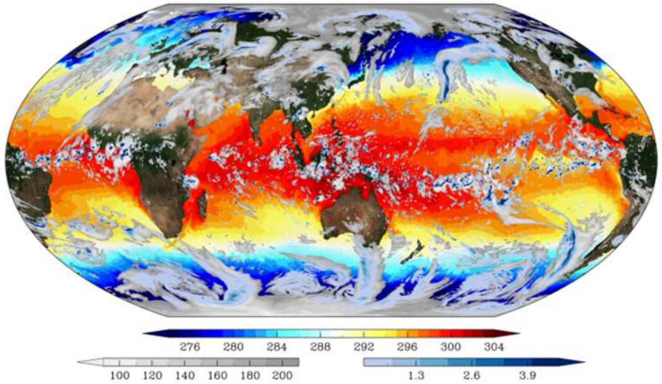
CNN Predictions

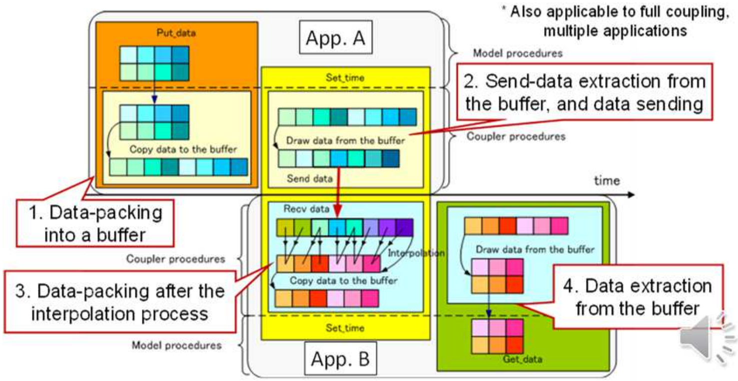


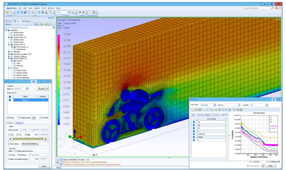
[c/o Takashi Shimokawabe (ITC/U.Tokyo)]

Possible Applications (S+D+L) on Wisteria/BDEC-01 with h3-Open-BDEC

- Simulations with Data Assimilation
  - Very Typical Example of (S+D+L)
- Atmosphere-Ocean Coupling for Weather and Climate Simulations
  - AORI/U.Tokyo, RIKEN R-CCS, NIES
- **Earthquake Simulations with Real-Time Data Assimilation**
  - **ERI/U. Tokyo**
- Real-Time Disaster Simulations
  - Flood, Tsunami
- (S+D+L) for Existing Simulation Codes (Open Source Software)
  - OpenFOAM

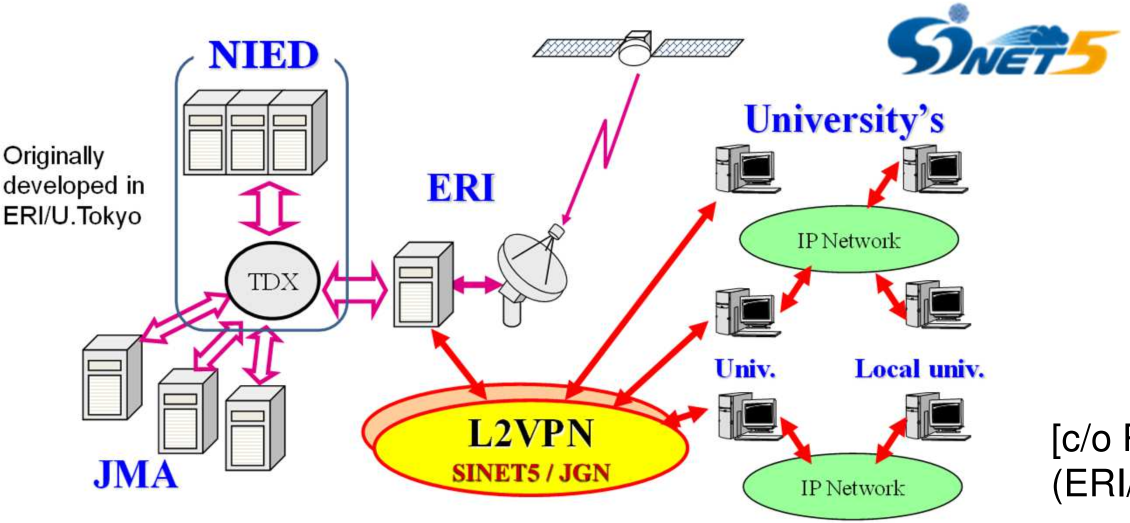






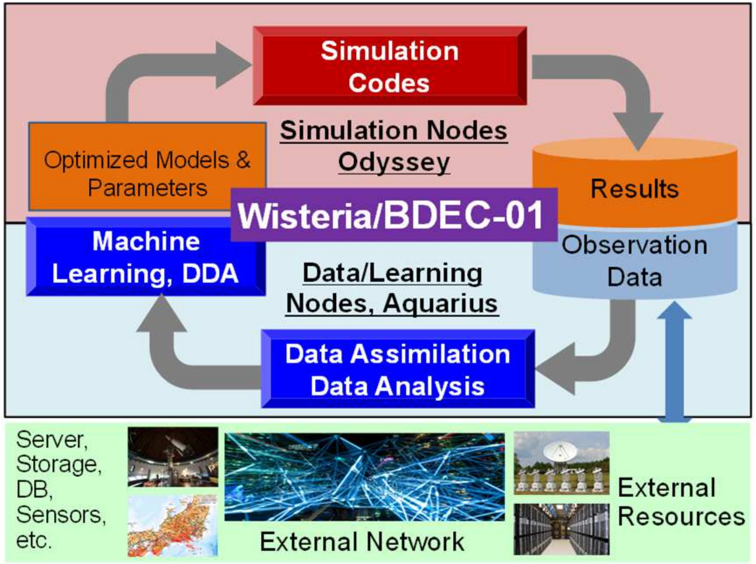
Real-Time Sharing of Seismic Observation is possible in Japan by JDXnet with SINET

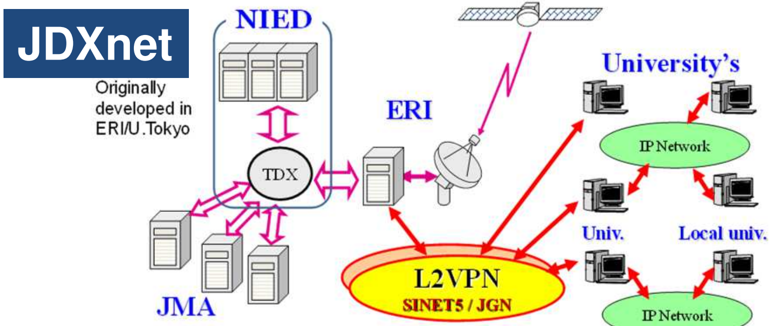
- Seismic Observation Data (100Hz/3-dir's/O(10^3) observation points) by JDXnet is available through SINET in Real Time
  - O(10^2) GB/day
  - O(10^5) pts in future including stations operated by industry

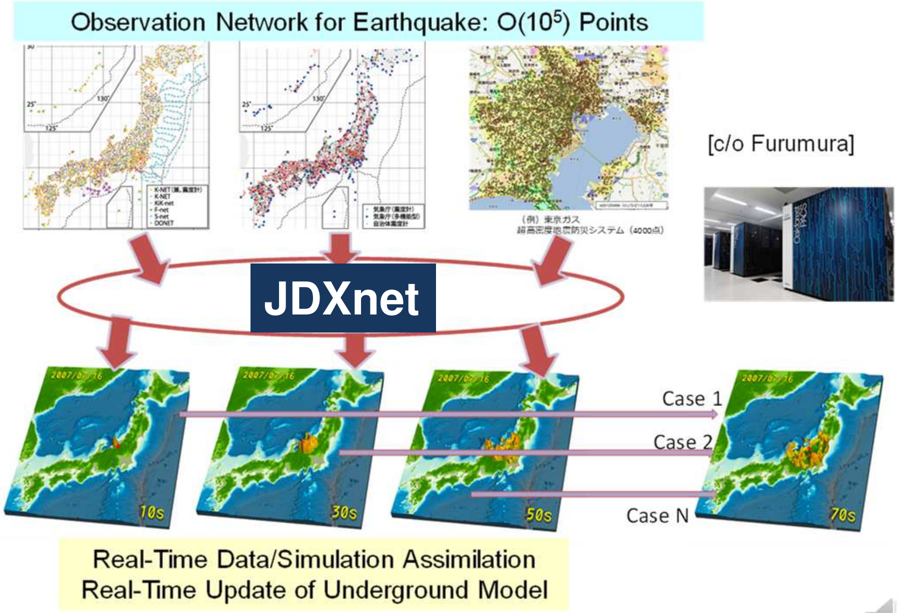


[c/o Prof. H.Tsuruoka (ERI/U.Tokyo)]

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



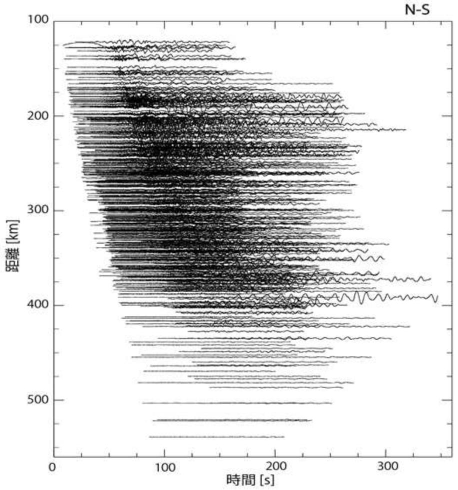


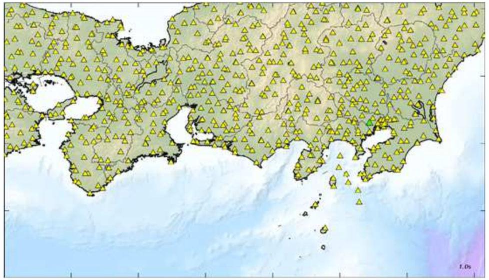


[c/o Prof. T.Furumura (ERI/U.Tokyo)]

Example of Real-Time Assimilation of (Obs.+Comp.): 2004 Kii Peninsula Earthquake (Mw 7.4) [c/o Oba & Furumura]

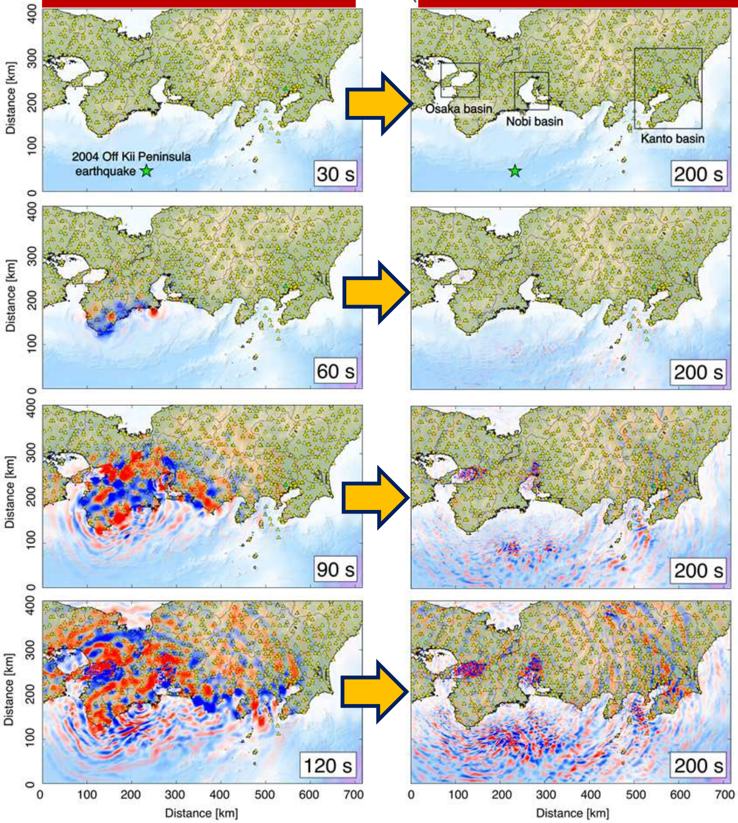
○ **Observation** (K-NET, KiK-net 446 pts)



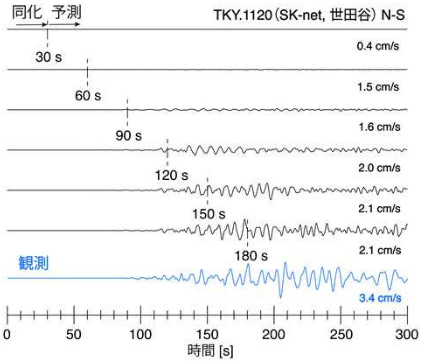


(a) Sim. + Assim.

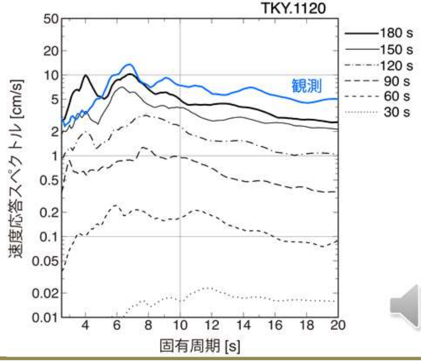
(b) Pure Simulation



Long Wave Propagation in Tokyo

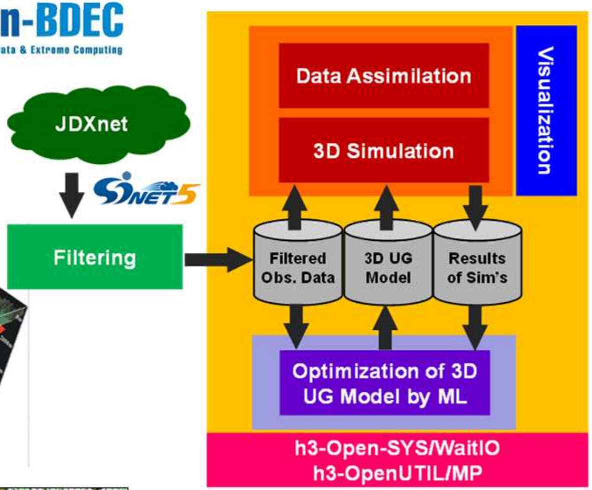


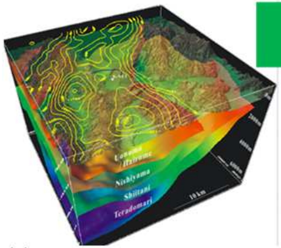
Response Spectrum

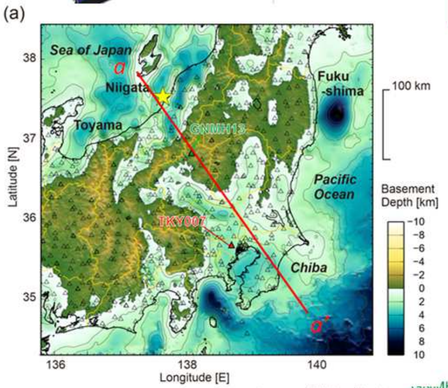


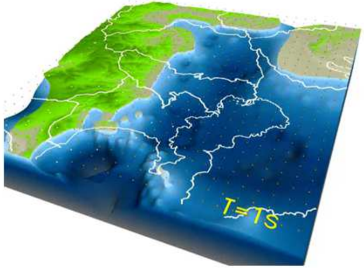
3D Earthquake Simulation with Real-Time Data Observation/Assimilation

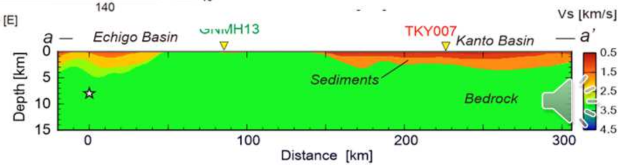
- Accurate Prediction of Seismic Wave Propagation with Real-Time Data Observation/Assimilation
  - Emergency Info. for Safer Evacuation
- 3D Underground Model
  - Heterogeneous, Observation is difficult
  - Inversion analyses of seismic waves are important for prediction of structure of underground model
  - ML may be utilized for acceleration of this prediction based on analyses of small earthquakes in normal time





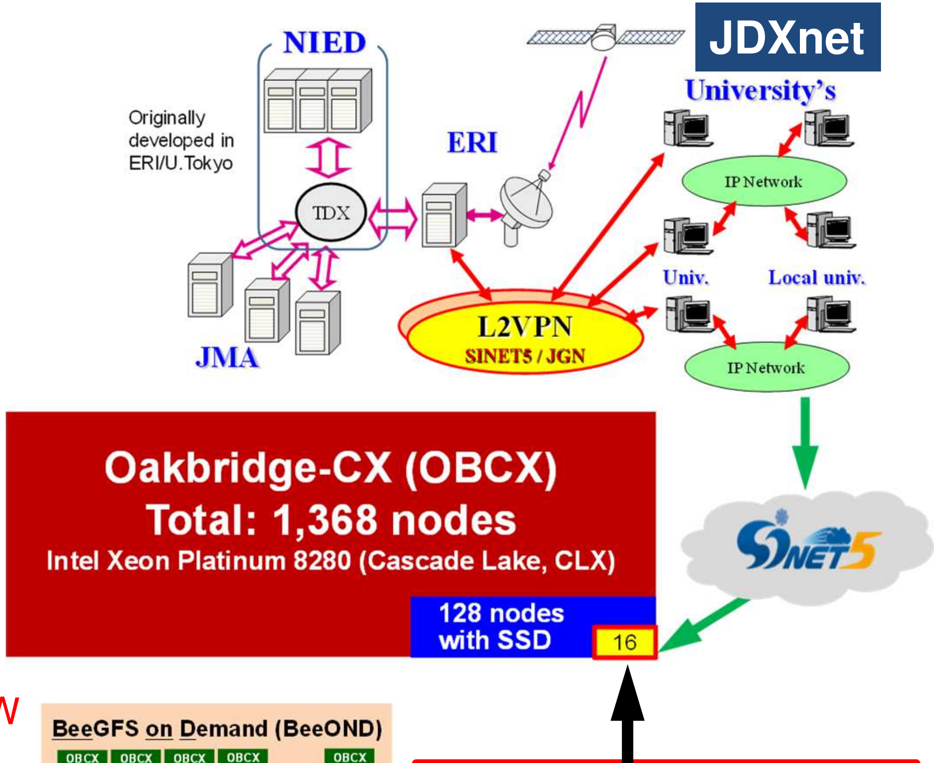


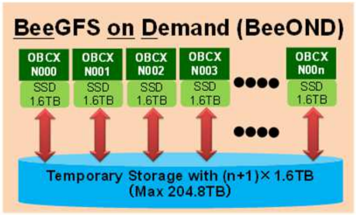




Preliminary Works on Oakbridge-CX (OBCX)

- Intel Xeon Platinum 8280 (Cascade Lake, CLX), Fujitsu
  - 1,368 nodes, 6.61 PF peak, 385.1 TB/sec, 4.2+ PF for HPL **#69 in 56th Top500 (Nov.2020)**
  - Fast Cache: SSD's for 128 nodes: Intel SSD, BeeGFS: 200+TB Fast FS
    - 1.6 TB/node, 3.20/1.32 GB/s/node for R/W
    - 16 of these nodes can directly access external resources (server, storage, sensor network etc.) through SINET
- Switching to Wisteria/BDEC-01 after May 2021

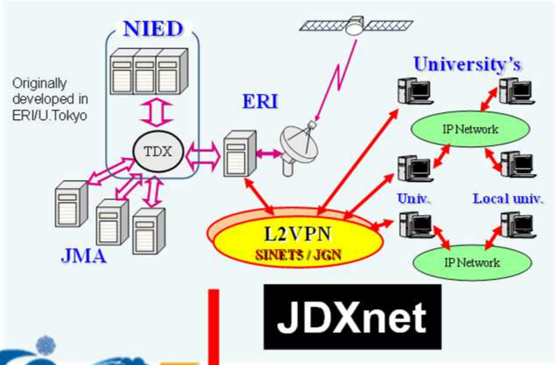




16 of 128 nodes with SSD can access external resources directly through SINET

3D Simulation + Real-Time Data Assimilation on OBCX

Seism3D/OpenSWPC-DAF: 3D FDM (Finite-Difference Method)



Filter

An External Node of OBCX

Normal Condition

Initial Velocity Distribution by Interpolation for Seism3D/OpenSWPC-DAF using O(10^1) nodes

Initial Velocity Field

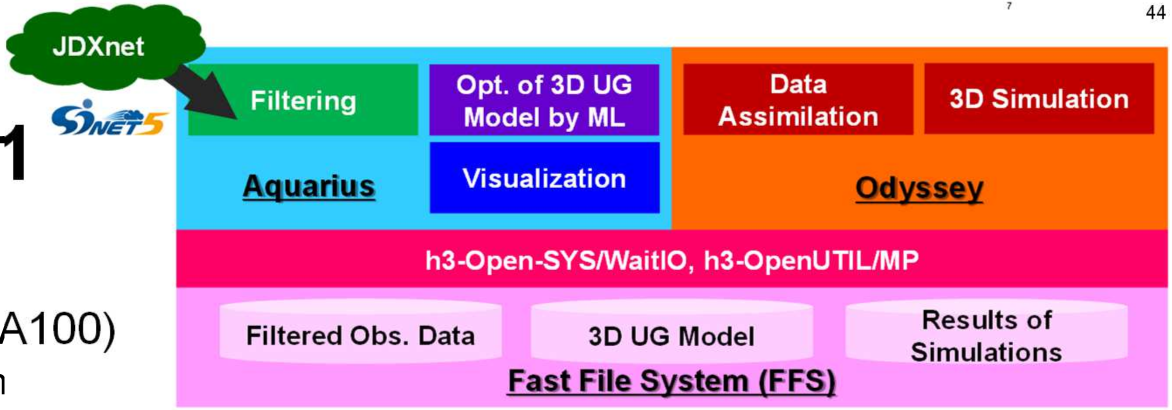
Emergency Condition

Seism3D/OpenSWPC-DAF using O(10^3+) nodes Data Assimilation + Simulation (A+S)

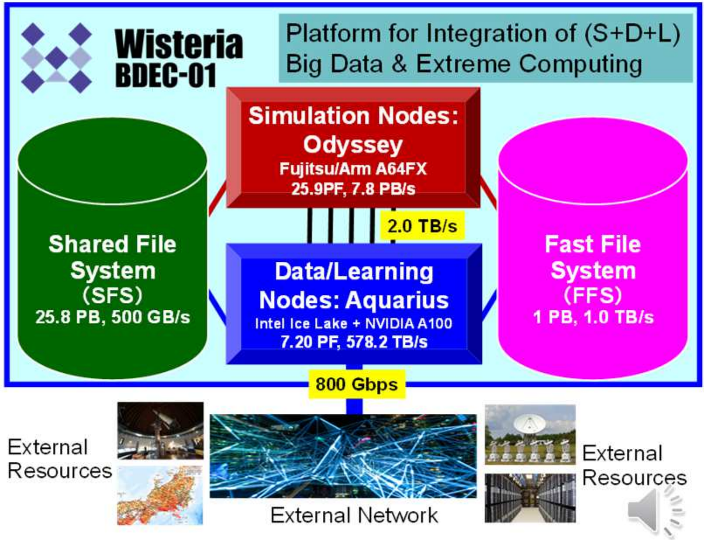
Seism3D/OpenSWPC-DAF using O(10^3+) nodes Pure Simulation (Pure S)

Oakbridge-CX (OBCX) Total: 1,368 nodes

Computing on Wisteria/BDEC-01



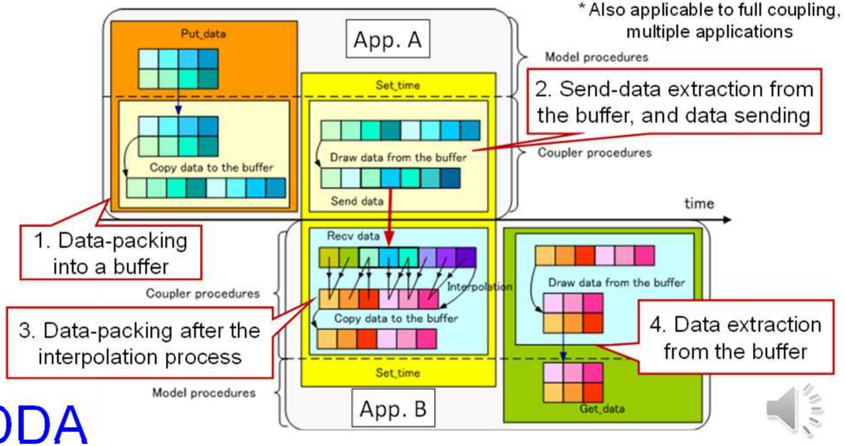
- Wisteria/BDEC-01
  - Aquarius (GPU: NVIDIA A100)
    - Filtering, ML, Visualization
  - Odyssey (CPU: A64FX)
    - Data Assimilation, Simulation
- Combining Odyssey-Aquarius (O-A)
  - Single MPI Job over O-A is impossible
  - Actually, O-A are connected through IB-EDR with 2TB/sec.
  - h3-Open-SYS/WaitIO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



h3-Open-UTIL/MP

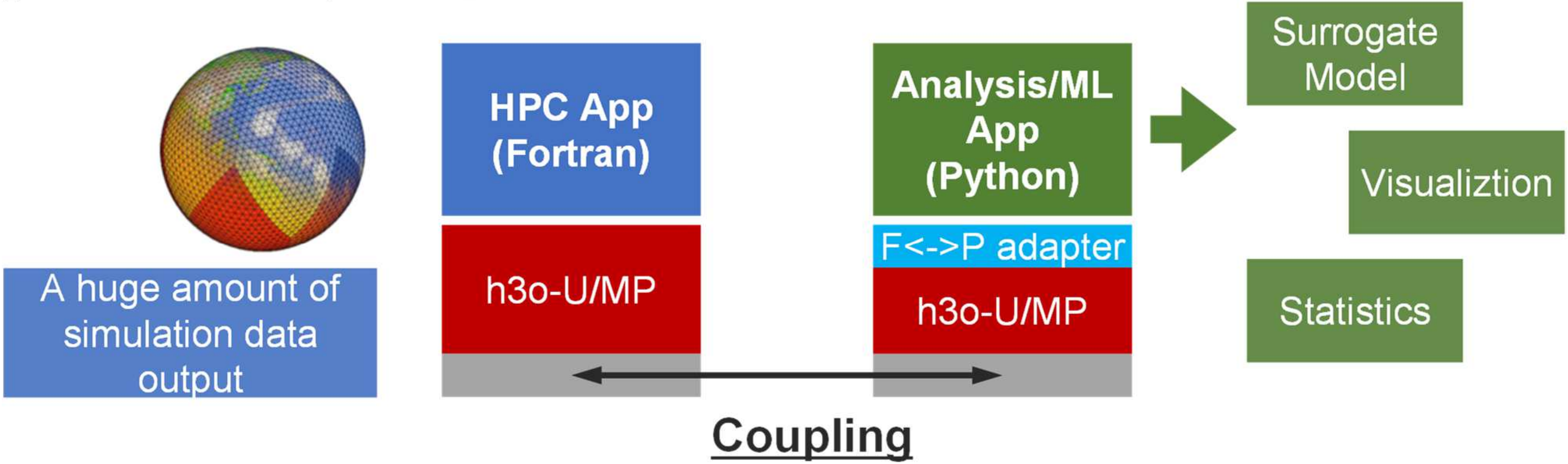
Multilevel Coupler/Data Assimilation

- Current 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



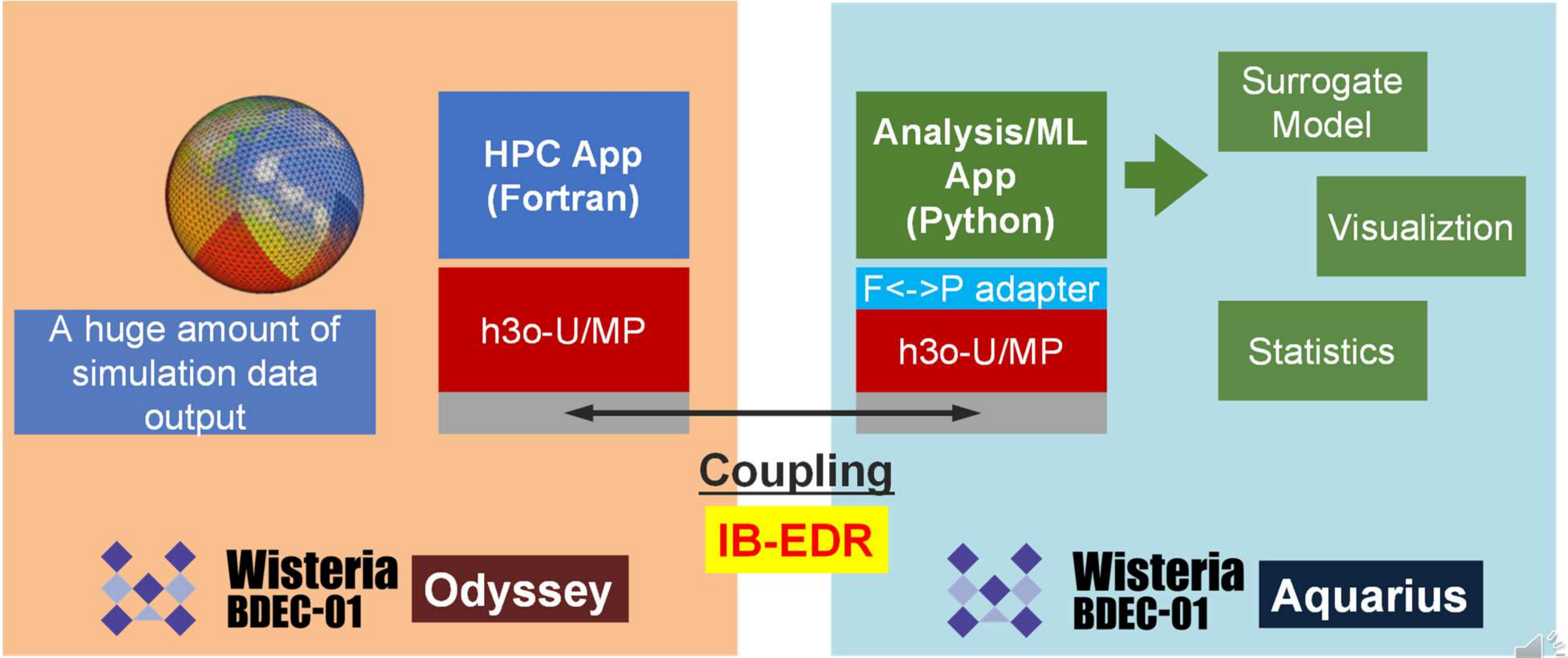
h3-Open-UTIL/MP (h3o-U/MP)

(HPC+AI) Coupling
[Dr. H. Yashiro, NIES]

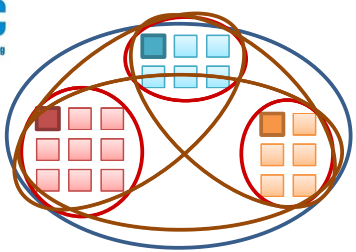


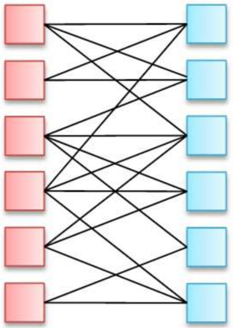
- Providing on-the-fly input/output/training data to the Analysis/ML tools
  - Easy to apply to existing HPC applications
  - Easy access to existing Python-based tools for AI/ML

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

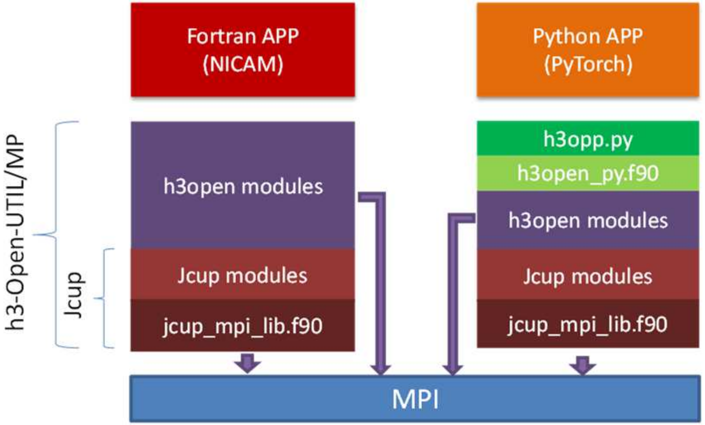


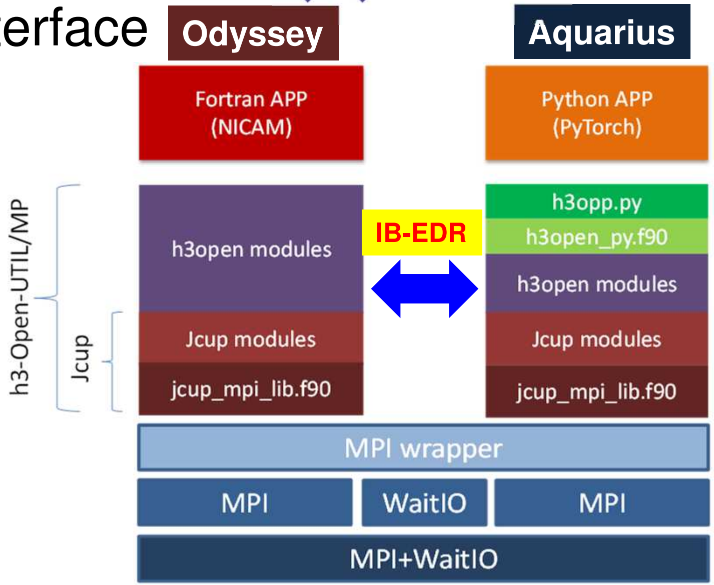
h3-Open-UTIL/MP + h3-Open-SYS/WaitIO





- Current Statue: Single MPI Job
- Direct Communication between Odyssey-Aquarius through IB-EDR by h3-Open-SYS/WaitIO, which provides MPI-like Interface



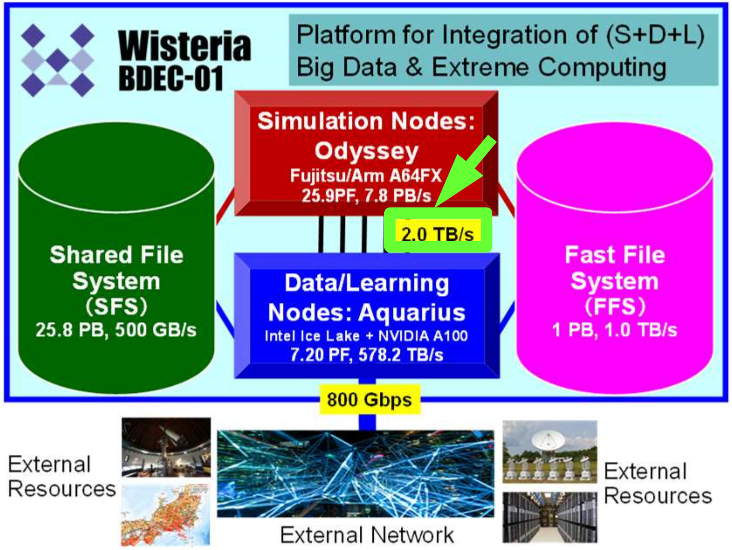


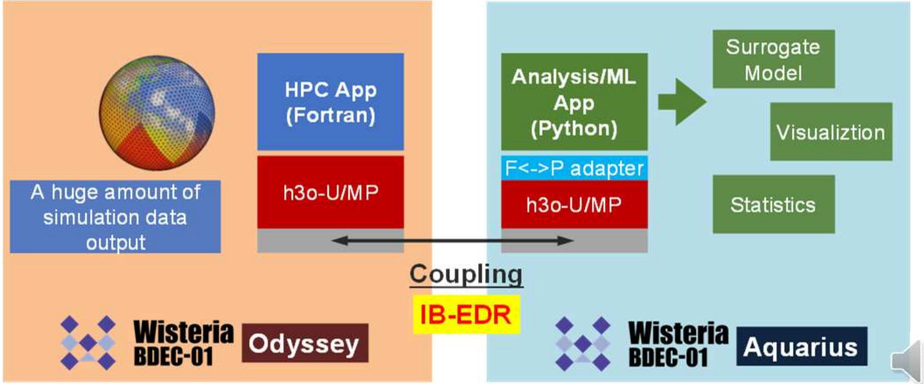
Current Status: Single MPI Job

Schedule for Public Use

Collaborations are Welcome !!

- h3-Open-SYS/WaitIO
  - October 2021, O-A Direct Communication by MPI-like Interface
- h3-Open-UTIL/MP（HPC+Python）
  - October 2021 on Odyssey only (Single MPI)
- h3-Open-UTIL/MP+h3-Open-SYS/WaitIO via IB-EDR
  - January-April 2022





Wisteria BDEC-01

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

h3-Open-BDEC: Summary

https://h3-open-bdec.cc.u-tokyo.ac.jp/

Wisteria BDEC-01

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

- By Integration of (S+D+L) using **h3-Open-BDEC (Adaptive Precision + *h*DDA)**, total energy consumption (=total computation time) for simulations will be 10% of that by the conventional methods for simulations with parameter studies
- h3-Open-BDEC is the 1st innovative software platform for integration of (S+D+L) on Exascale systems, where computational scientists can achieve such integration without supports by other experts in data analytics and AI/ML.
- Source codes and documents (in English) are open to public for various kinds of computational environments.