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

Integration of Simulation/Data/Learning

Kengo Nakajima
Information Technology Center
The University of Tokyo

Wisteria BDEC-01

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

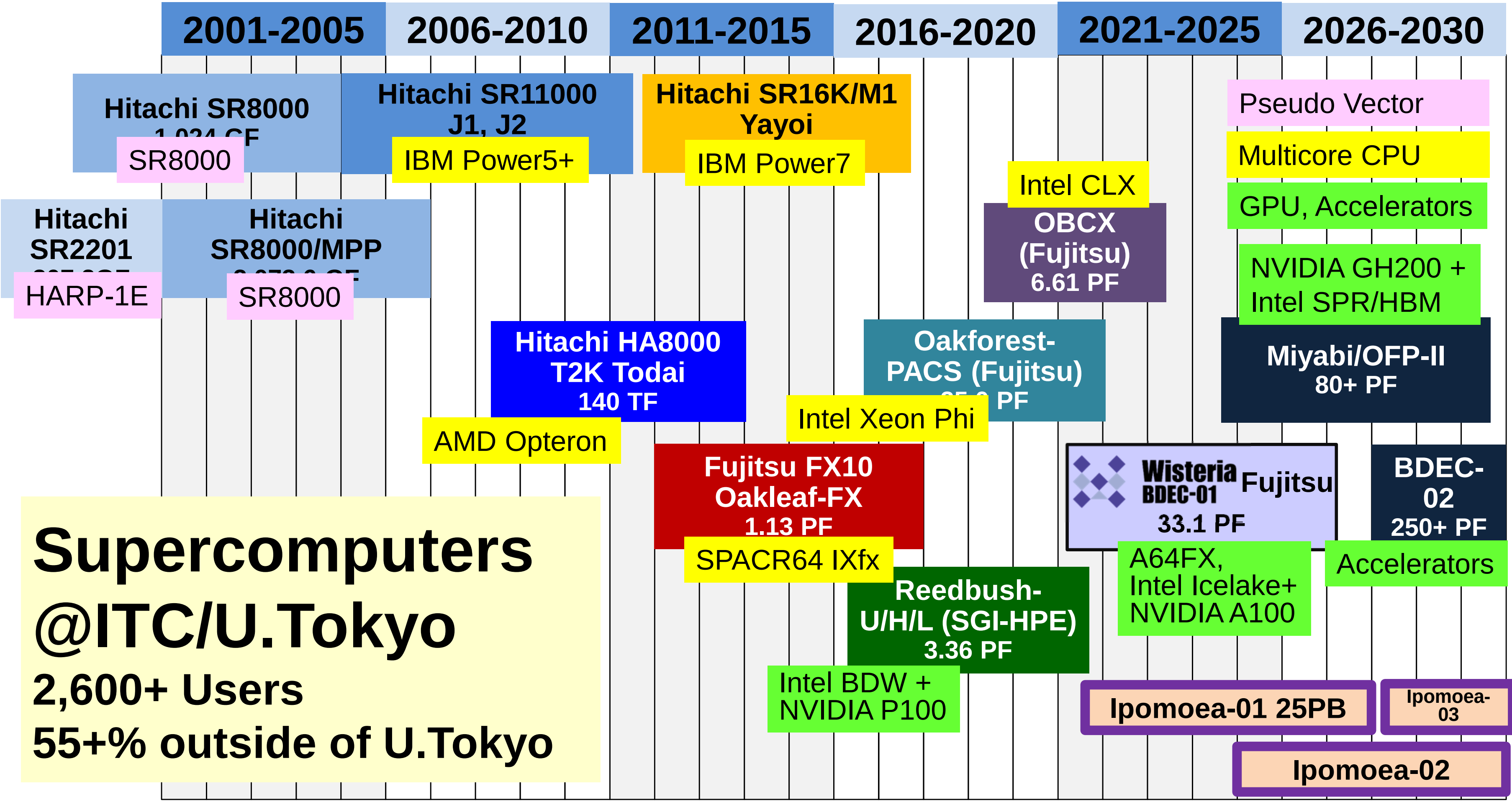
- Integration of (Simulation/Data/Learning)
  - Wisteria/BDEC-01
  - h3-Open-BDEC
- Applications on Wisteria/BDEC-01 with h3-Open-BDEC

65th TOP500 List (June, 2025)

R_{max}: Performance of Linpack (TFLOPS)
R_{peak}: Peak Performance (TFLOPS), Power: kW

http://www.top500.org/

	Site	Computer/Year Vendor	Cores	R_{max} (PFLOPS)	R_{peak} (PFLOPS)	GFLOPS/W	Power (kW)
1	**El Capitan, 2024, USA** DOE/NNSA/LLNL	HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A, Slingshot-11, TOSS	11,039,616	1,742.00 (=1.742 EF)	2,746.38 63.4 %	58.99	29,581
2	**Frontier, 2021, USA** DOE/SC/Oak Ridge National Laboratory	HPE Cray EX235a, AMD Optimized 3rd Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	9,066,176	1.353.00	2,055.72 65.8 %	54.98	24,607
3	**Aurora, 2023, USA** DOE/SC/Argonne National Laboratory	HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11, Intel	9,264,128	1,012.00	1,980.01 51.1 %	26.15	38,698
4	**JUPITER Booster, 2025, USA** EuroHPC/FZJ	EVIDEN, BullSequana XH3000, GH Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail NVIDIA InfiniBand NDR200, RedHat Enterprise Linux	4,801,344	794,40	930.00 85.3 %	60.62	13,088
5	**Eagle, 2023, USA** Microsoft	Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR	2,073,600	561.20	846.84 66.3 %		
6	**HPC 6, 2024, Italy** Eni S.p.A.	HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, RHEL 8.9	3,143,520	477.90	606.97 66.3 %	56.48	8,461
7	**Fugaku, 2020, Japan** R-CCS, RIKEN	Fujitsu PRIMEHPC FX1000, Fujitsu A64FX 48C 2.2GHz, Tofu-D	7,630,848	442.01	537.21 82.3 %	14.78	29,899
8	**Alps, 2024, Switzerland** Swiss Natl. SC Centre (CSCS)	HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11	2,121,600	434.90	574.84 75.7 %	61.05	7,124
9	**LUMI, 2023, Finland** EuroHPC/CSC	HPE Cray EX235a, AMD Optimized 3rd Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	2,752,704	379.70	531.51 71.4 %	53.43	7,107
10	**Leonard, 2023, Italy** EuroHPC/Cineca	EVIDEN, BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64GB, Quad-rail NVIDIA HDR100	1,824,768	241.20	181.49 78.7 %	32.19	7,494
15	**ABCI 3.0, 2025, Japan** AIST	HPE Cray XD670, Xeon Platinum 8558 48C 2.1GHz, NVIDIA H200 SXM5 141 GB, Infiniband NDR200, Rocky Linux 9	479,232	145.10	130.44 80.0 %	59.29	3,596
22	**CHIE-3, 2024, Japan** SoftBank, Corp.	NVIDIA DGX H100, Xeon Platinum 8480C 56C 2GHz, NVIDIA H100, Infiniband NDR400, Ubuntu 22.04.4 LTS	297,840	91.94	138.32 66.5 %		
24	**CHIE-2, 2024, Japan** SoftBank, Corp.	NVIDIA DGX H100, Xeon Platinum 8480C 56C 2GHz, NVIDIA H100, Infiniband NDR400, Ubuntu 22.04.4 LTS	297,840	89.78	138.32 64.9 %		
27	**ABCI-Q, 2025, Japan** AIST	Fujitsu, Supermicro SYS-221GE-TNHT-LCC, Intel Xeon Platinum 8558 48C 2.1GHz, NVIDIA H100 SXM5 80GB, Infiniband NDR, Rocky Linux 9.4	315,210	74.58	99.35 75.1 %	40.67	1,834
36	**FPT, 2025, JAPAN** FPT AI Factory	HPE, HGX H200, Xeon Platinum 8558 48C 2.1GHz, NVIDIA H200 SXM5 141 GB, Infiniband NDR400, Ubuntu 22.04.5 LTS	146,304	49.85	67.44 73.9 %		
37	**Miyabi-G, 2024, Japan** JCAHPC	Fujitsu, Supermicro ARS 111GL DNHR LCC, Grace Hopper Superchip 72C 3GHz, Infiniband NDR200, Rocky Linux	80,640	46.80	72.80 64.3 %	47.59	983
46	**TSUBAME 4.0, 2024, Japan** Institute of Science Tokyo	HPE Cray XD665, AMD EPYC 9654 96C 2.4GHz, NVIDIA H100 SXM5 94 GB, Infiniband NDR200	172,800	39.62	61.60 64.3 %	48.55	816
73	**Wisteria/BDEC-01 (Odyssey), 2021, Japan** U.Tokyo	Fujitsu PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu D	368,640	22.12	25.95 85.2 %	15.07	1,468


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
Pseudo Vector
Multicore CPU
GPU, Accelerators
Intel CLX
OBCX
(Fujitsu)
6.61 PF
Hitachi
SR2201
HARP-1E
Hitachi
SR8000/MPP
SR8000
NVIDIA GH200 +
Intel SPR/HBM
Hitachi HA8000
T2K Todai
140 TF
AMD Opteron
Oakforest-
PACS (Fujitsu)
PF
Intel Xeon Phi
Miyabi/OFP-II
80+ PF
Fujitsu FX10
Oakleaf-FX
1.13 PF
SPACR64 IXfx
Wisteria
BDEC-01
Fujitsu
33.1 PF
BDEC-
02
250+ PF
A64FX,
Intel Icelake+
NVIDIA A100
Accelerators
Reedbush-
U/H/L (SGI-HPE)
3.36 PF
Intel BDW +
NVIDIA P100
Ipomoea-01 25PB
Ipomoea-
03
Ipomoea-02
Supercomputers
@ITC/U.Tokyo
2,600+ Users
55+% outside of U.Tokyo

Integration of (S+D+L) has been our main strategy in recent 10 years

- Various Types of Workloads
  - Computational Science & Engineering: Simulations
  - Big Data Analytics +AI, Machine Learning …
- **Integration of (<u>Simulation+Data+ Learning</u>) (S+D+L) is important towards Society 5.0, Human-Centered Society proposed by Japanese Gov.**
  - **By Integration of Cyber & Physical Space**
- **BDEC (Big Data & Extreme Computing)**
  - **Platform for Integration of (S+D+L)**
  - **Focusing on S (Simulation)**
    - **AI for HPC, (Classical) AI for Science**
  - **Planning started in 2015**

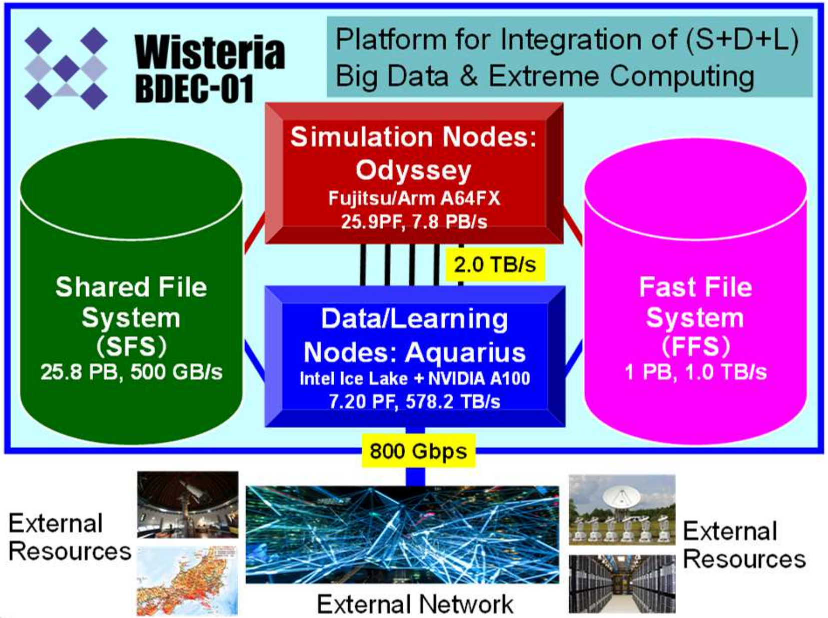
BDEC (Big Data & Extreme Computing)

S + D + L

Wisteria/BDEC-01

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

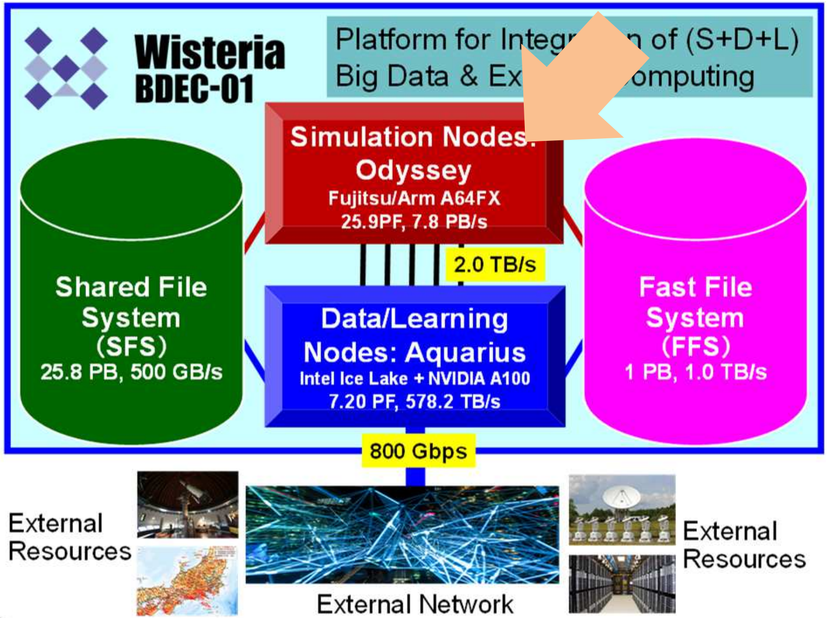
- Operation starts on May 14, 2021
- 33.1 PF, 8.38 PB/sec by **Fujitsu**
  - ~4.5 MVA with Cooling, ~360m^2
- 2 Types of Node Groups
  - Hierarchical, Hybrid, Heterogeneous (h3)
  - Simulation Node Group: Odyssey
    - Fujitsu PRIMEHPC FX1000 (A64FX), 25.9 PF
      - 7,680 nodes (368,640 cores), Tofu-D
      - General Purpose CPU + HBM
      - Commercial Version of “Fugaku”
  - Data/Learning Node Group: Aquarius
    - Data Analytics & AI/Machine Learning
    - Intel Xeon Ice Lake + NVIDIA A100, 7.2PF
      - 45 nodes (90x Ice Lake, 360x A100), IB-HDR
    - Each DL node is 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)

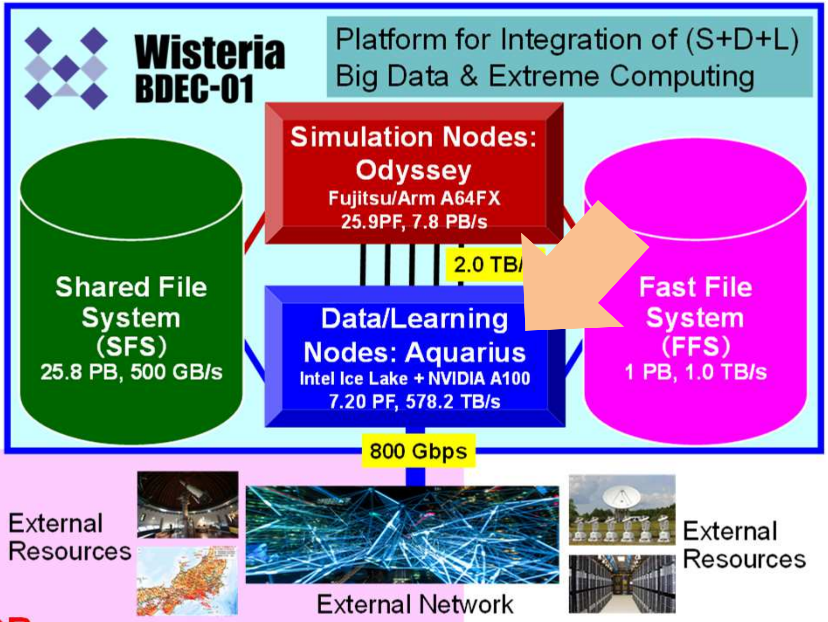
- Operation starts on May 14, 2021
- 33.1 PF, 8.38 PB/sec by **Fujitsu**
  - ~4.5 MVA with Cooling, ~360m^2
- **2 Types of Node Groups**
  - Hierarchical, Hybrid, Heterogeneous (h3)
  - **Simulation Node Group: Odyssey**
    - **Fujitsu PRIMEHPC FX1000 (A64FX), 25.9 PF**
      - **7,680 nodes (368,640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

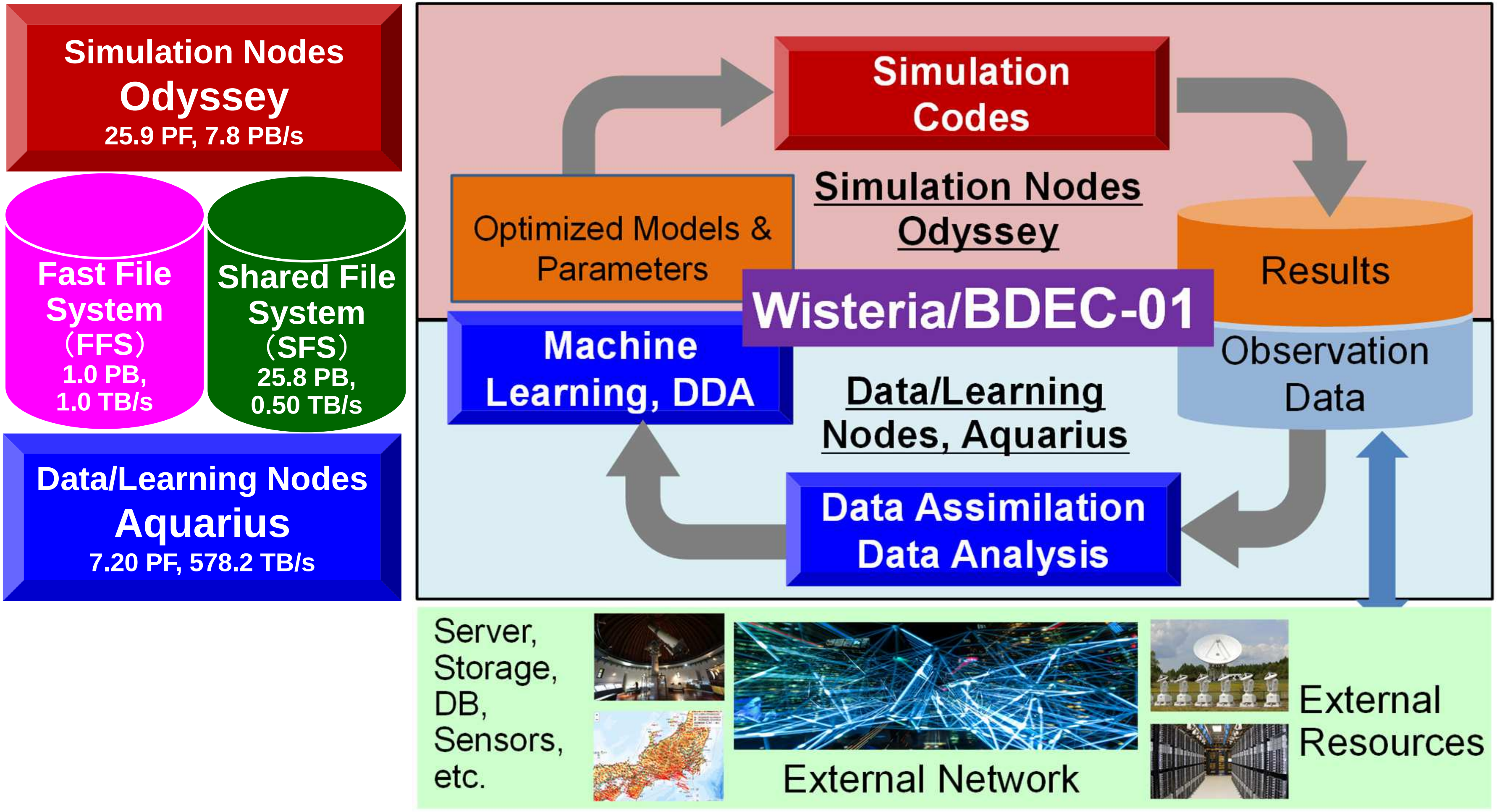


Wisteria/BDEC-01

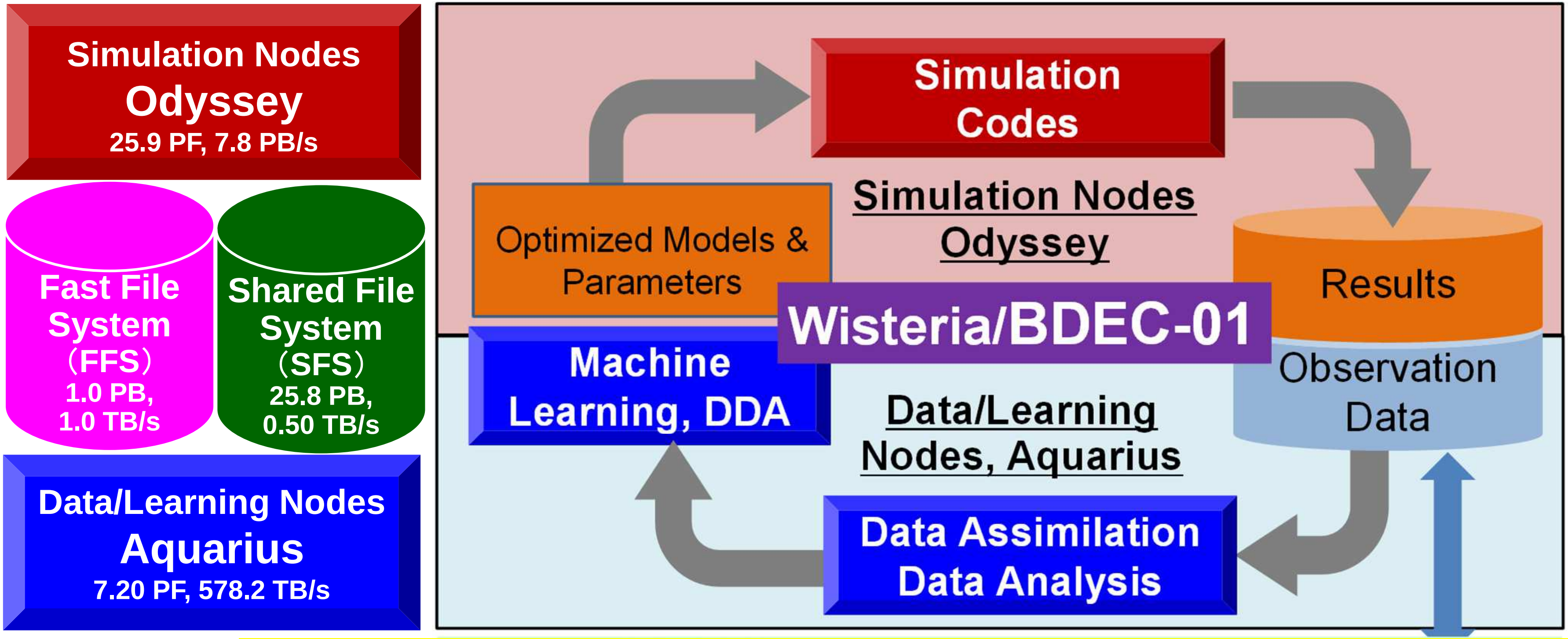
The 1st BDEC System (Big Data & Extreme Computing)
Platform for Integration of (S+D+L)

- 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 Node Group: 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 Node Group: Aquarius**
    - **Data Analytics & AI/Machine Learning**
    - **Intel Xeon Ice Lake + NVIDIA A100, 7.2PF**
      - **45 nodes (90x Ice Lake, 360x A100), IB-HDR**
    - **Each DL node is connected to external resources directly**
- File Systems: SFS (Shared/Large) + FFS (Fast/Small)



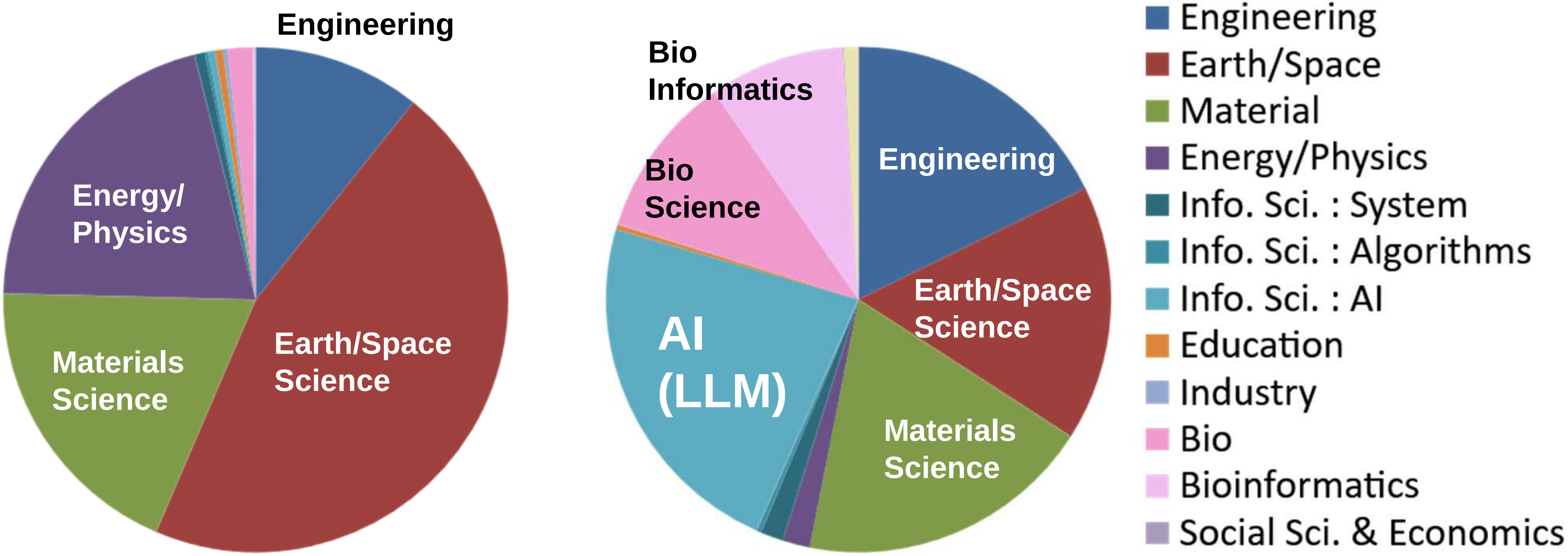


Simulation Nodes
Odyssey
25.9 PF, 7.8 PB/s
Fast File System (FFS)
1.0 PB, 1.0 TB/s
Shared File System (SFS)
25.8 PB, 0.50 TB/s
Data/Learning Nodes
Aquarius
7.20 PF, 578.2 TB/s
Simulation Codes
Optimized Models & Parameters
Simulation Nodes Odyssey
Results
Wisteria/BDEC-01
Machine Learning, DDA
Observation 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)

Research Area based on Machine Hours CPU, GPU (April 2024-March 2025)



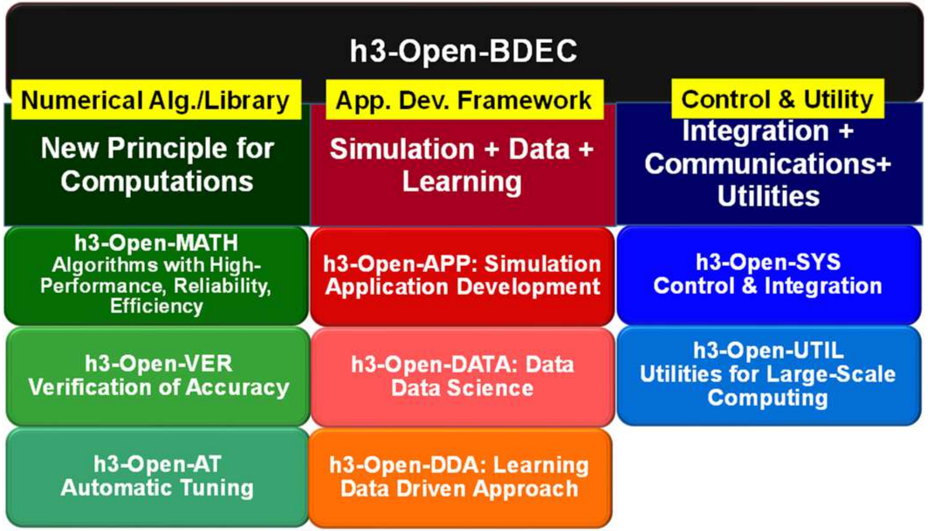
Odyssey A64FX

Aquarius A100

- Integration of (Simulation/Data/Learning)
  - Wisteria/BDEC-01
  - **h3-Open-BDEC**
- Applications on Wisteria/BDEC-01 with h3-Open-BDEC

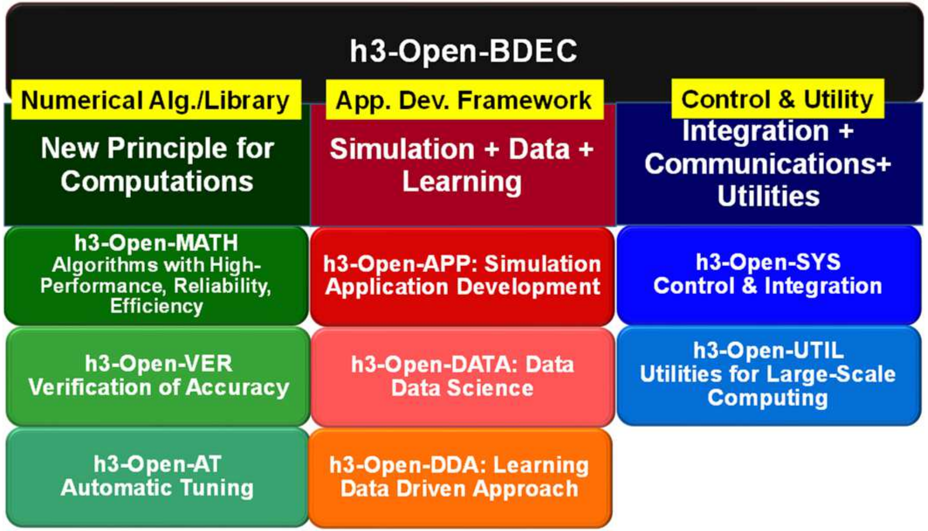
h3-Open-BDEC: Innovative Software Platform for Integration of (S+D+L) on the BDEC System, such as Wisteria/BDEC-01

- 5-year project supported by Japanese Government (JSPS) since 2019
  - FY.2023 is the final year
    - Until then end of March 2024
- Leading-PI: Kengo Nakajima (The University of Tokyo)
- Total Budget: 1.41M USD



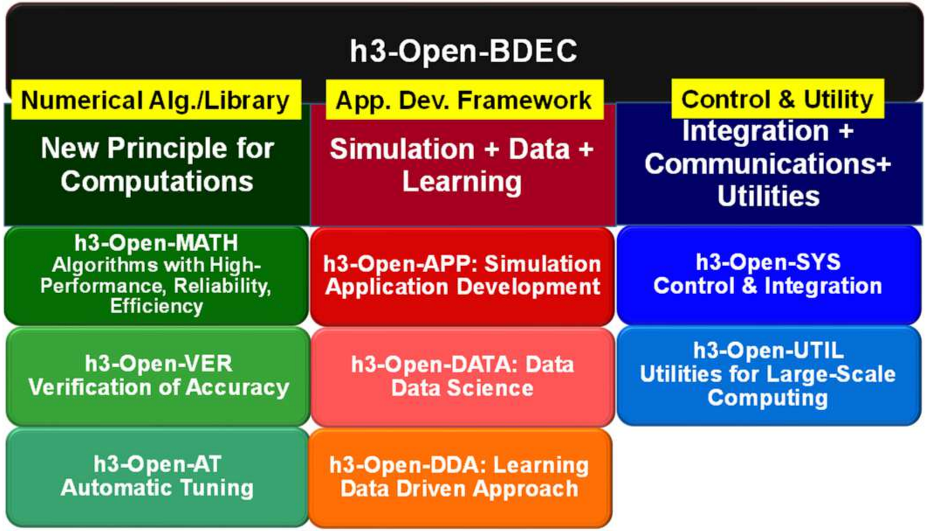
h3-Open-BDEC Innovative Software Platform for Integration of (S+D+L) on the BDEC System, such as Wisteria/BDEC-01

- "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

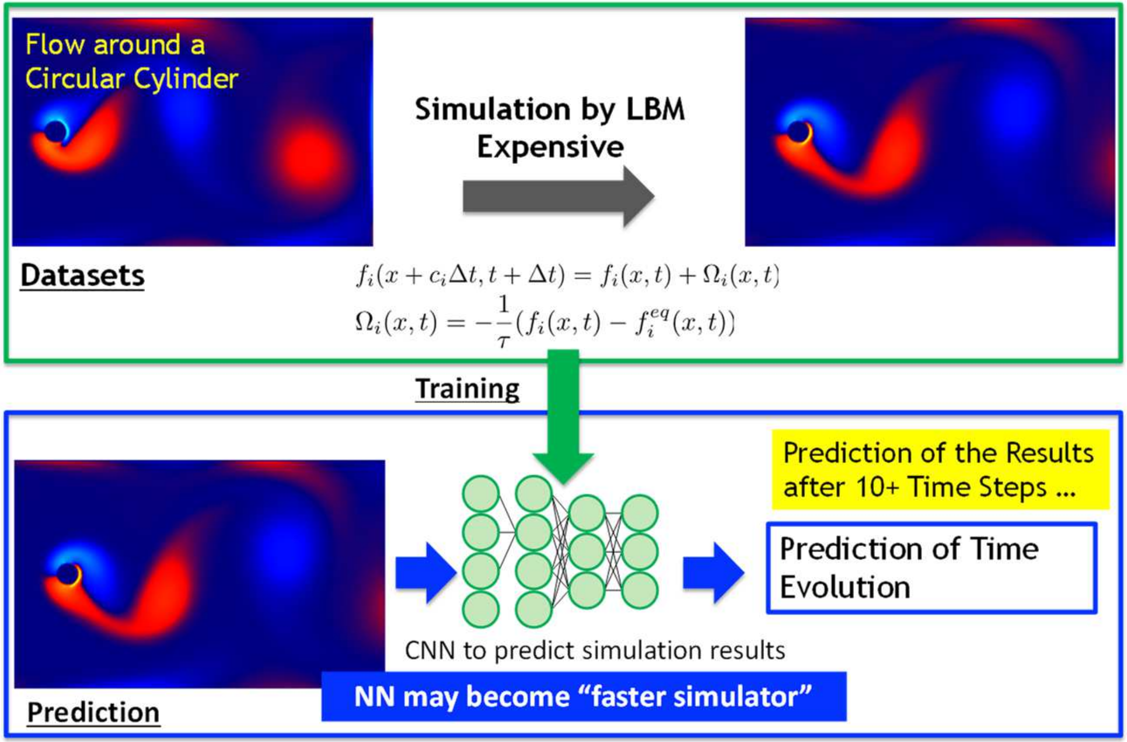


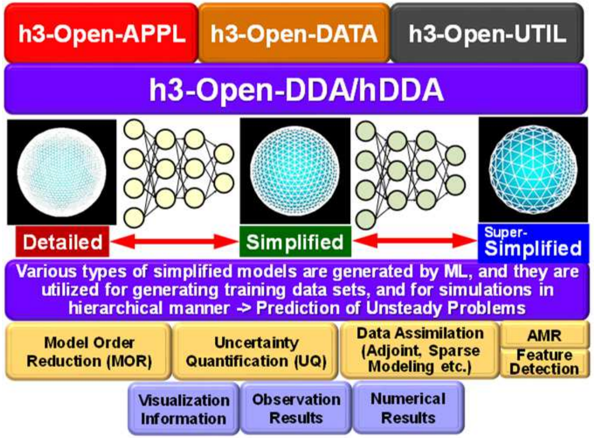
h3-Open-BDEC Innovative Software Platform for Integration of (S+D+L) on the BDEC System, such as Wisteria/BDEC-01

- "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





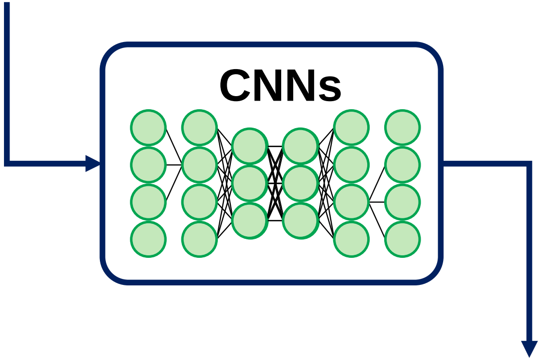
Initial Target: Estimating O(10) time steps ahead in transient CFD simulations

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

Prediction of steady flows using convolutional neural networks (CNNs)

Input

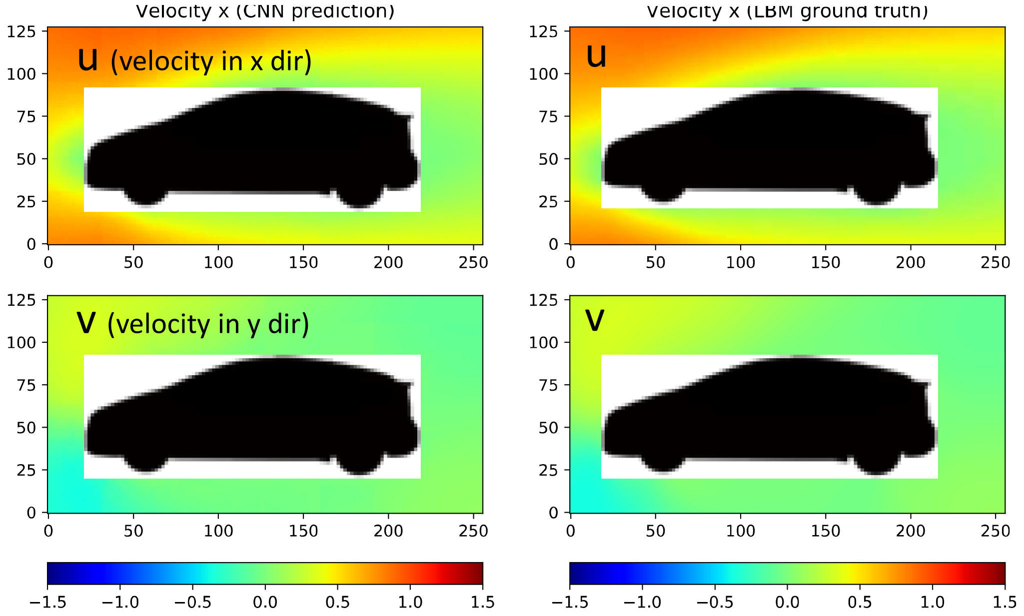
- Signed distance function (Geometry)
- Boundary conditions of velocity (u, v)



Output

- Velocity (u, v)

CNN Prediction **LBM Ground truth**



Computation time

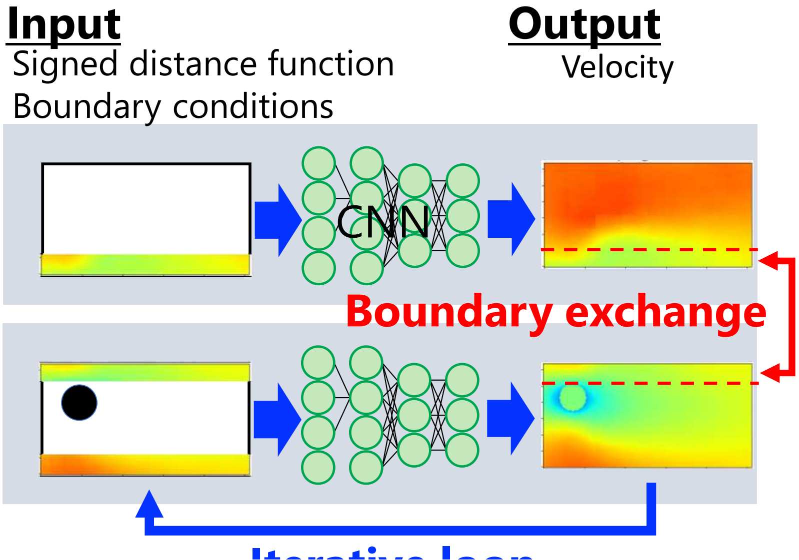
LBM (82,000steps) : 41.1 sec
CNN prediction: 0.6 sec

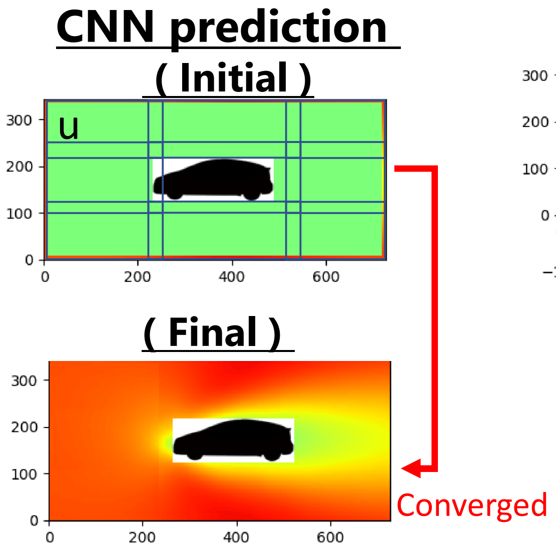
[c/o Takashi Shimokawabe (ITC/U.Tokyo)]

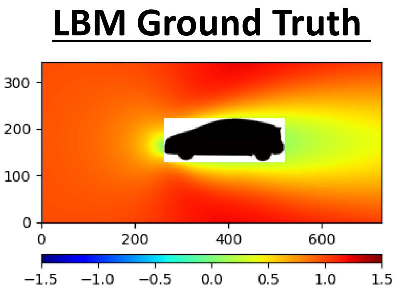
CNN prediction has achieved high accuracy with significant reduction in calculation time.

Prediction by CNN with boundary exchange

- Predicting simulation results on large domain using CNN with boundary exchange.
- The network model trained for a single domain is applied to the decomposed subdomains to predict the simulation results in each subdomain.
- In order to maintain consistency between values in the subdomains, boundary exchange between neighbor subdomains is performed.
- CNN and boundary exchange are performed iteratively until values converge.



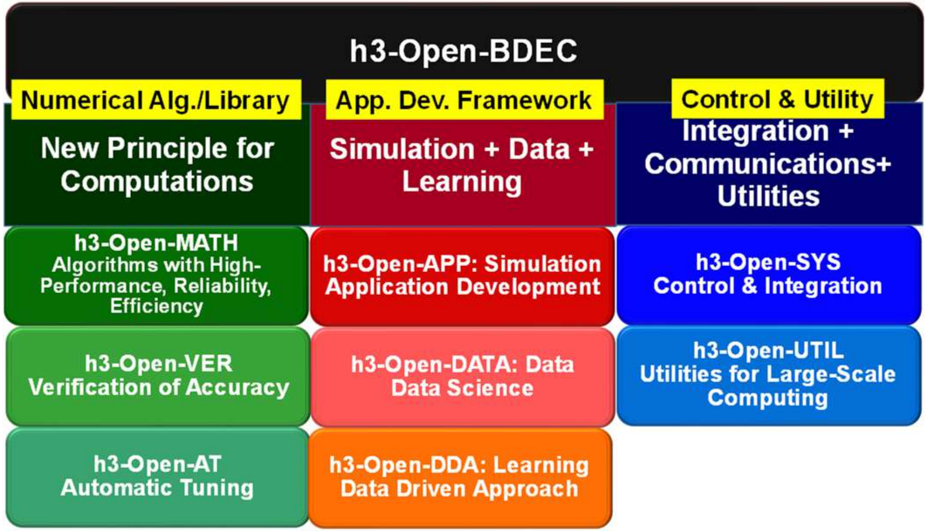




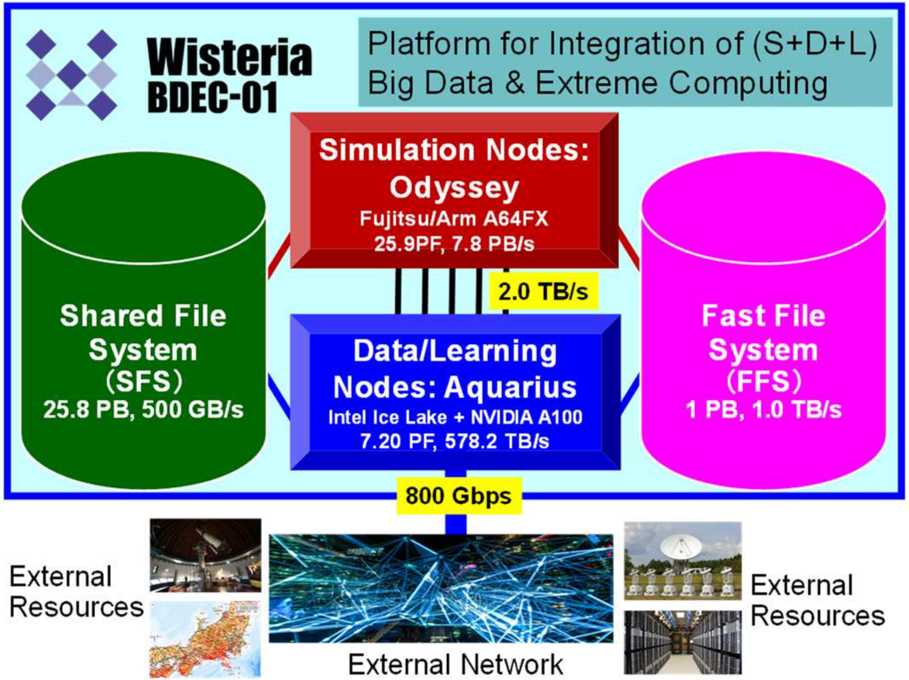
Domain size：748 x 364
(9 decomposed subdomains)
Mean error ：3.89％
Comp. time ：3.82 s

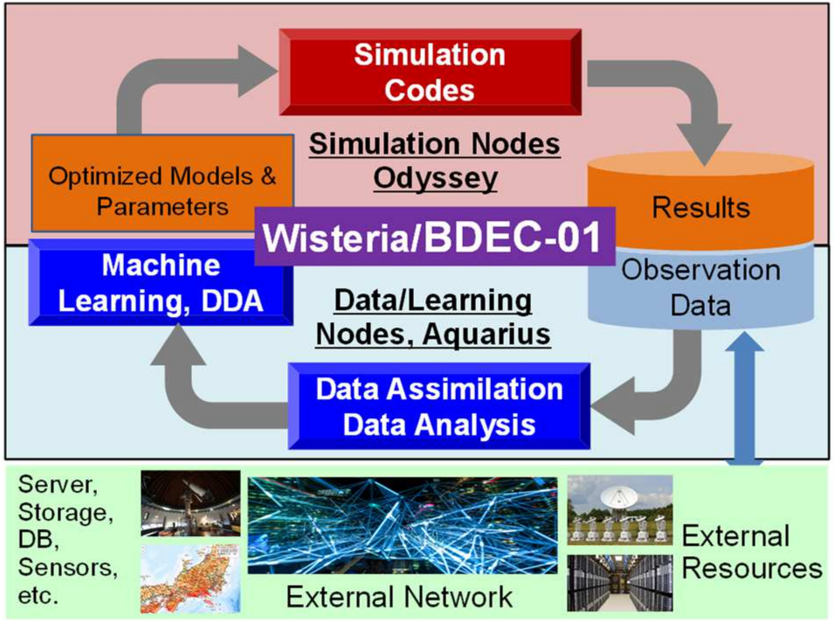
h3-Open-BDEC Innovative Software Platform for Integration of (S+D+L) on the BDEC System, such as Wisteria/BDEC-01

- "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



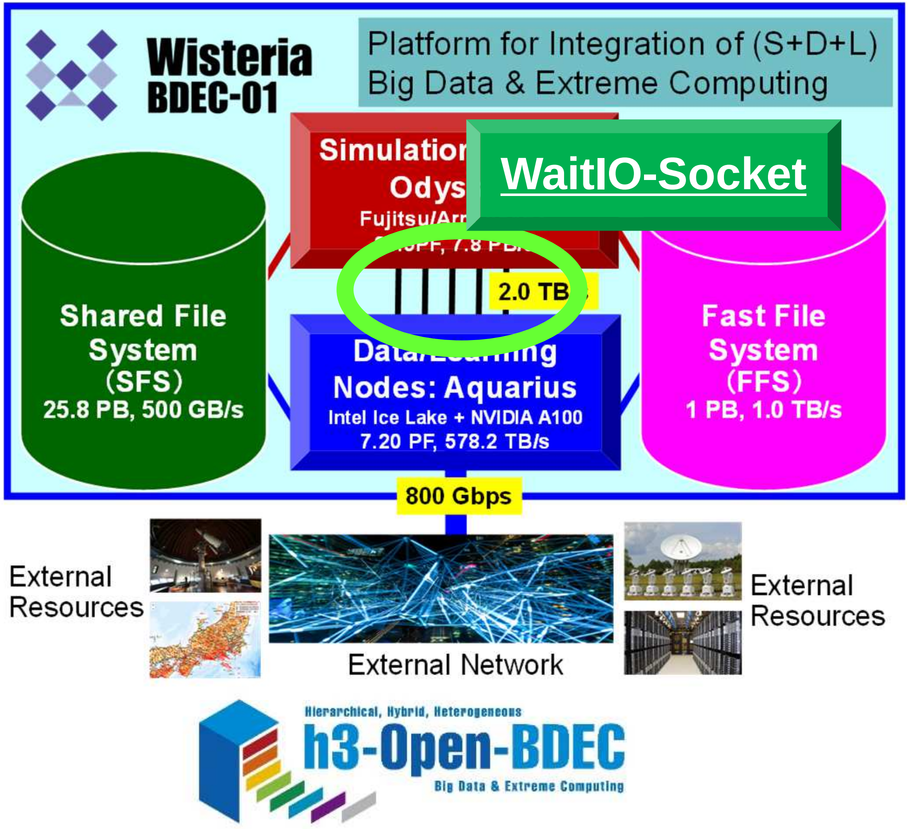
Wisteria/BDEC-01: The First "Really Heterogenous" System in the World





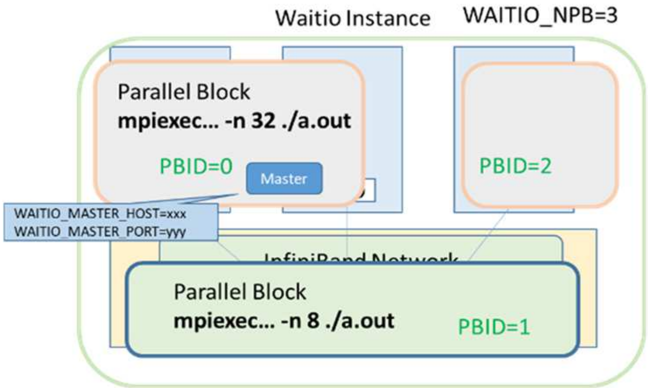
h3-Open-SYS/WaiIO-Socket

- Wisteria/BDEC-01
  - Aquarius (GPU: NVIDIA A100)
  - Odyssey (CPU: A64FX)
- Combining Odyssey-Aquarius
  - Single MPI Job over O-A is impossible
- **Connection between Odyssey-Aquarius**
  - **IB-EDR with 2TB/sec.**
  - **Fast File System**
  - **h3-Open-SYS/WaitIO-Socket**
    - **Library for Inter-Process Communication through IB-EDR with MPI-like interface**



API of h3-Open-SYS/WaitIO-Socket
PB (Parallel Block): Each Application

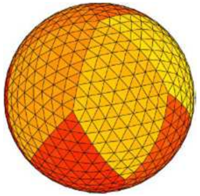
WaitIO API	Description
waitio_isend	Non-Blocking Send
waitio_irecv	Non-Blocking Receive
waitio_wait	Termination of waitio_isend/irecv
waitio_init	Initialization of WaitIO
waitio_get_nprocs	Process # for each PB (Parallel Block)
waitio_create_group waitio_create_group_wranks	Creating communication groups 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

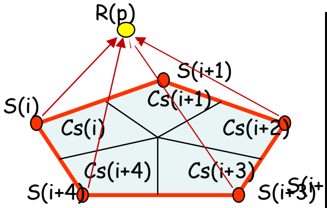


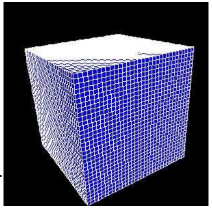
[Sumimoto et al. 2021]

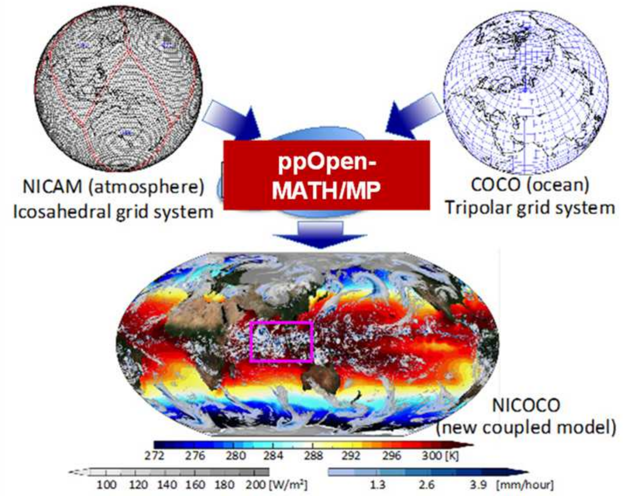
Coupler/Multiphysics Coupler

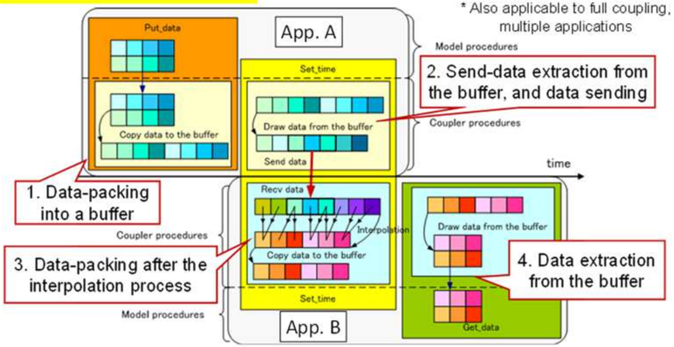
- *Coupler* is a tool for coupling multiple scientific applications loosely
  - *Multiphysics Coupler: Ocean-Atmosphere, Fluid-Structure etc.*
  - Connecting multiple applications using various grid/mesh systems: *e.g. FEM-FDM*
  - Interpolation/Extrapolation
  - Non-Centralized/Fully Distributed Operations, No Master Process





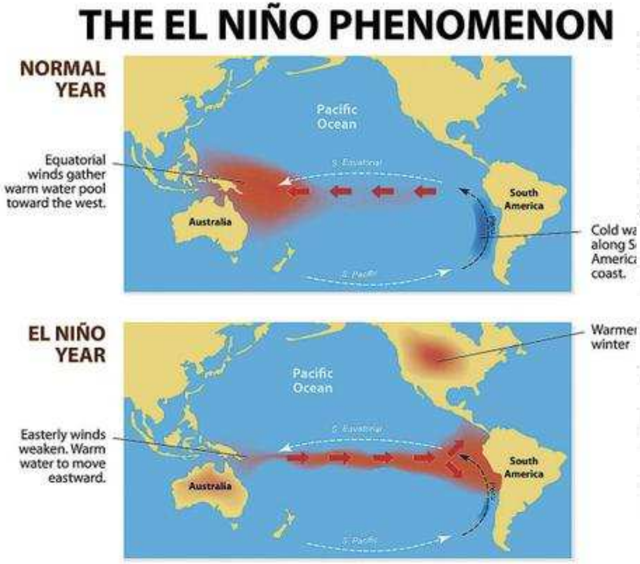


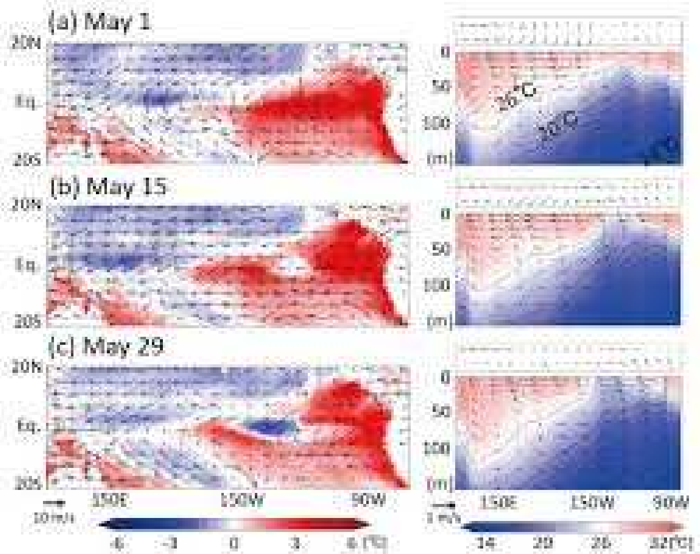




Simulations of Super El Niño

Mechanism of the Abrupt Terminate of Super El Niño in 1997/1998 has been revealed by Atmosphere-Ocean Coupling Simulations for the Entire Earth using ppOpen-HPC (Multiphysics Coupler) on the K computer [Sato et al. 2017]






2017.9.20 日経(夕)

「スーパーエルニーニョ」終息
巨大雲 引き金に
東大など解明

東京大学大気海洋研究所と理化学研究所、海洋研究開発機構は、1997～98年に発生して世界に異常気象をもたらした「スーパーエルニーニョ」現象が急に終息した原因を突き止めた。熱帯の巨大雲が引き金になっていた。台風の発生予測などの精度向上にもつながる成果という。

エルニーニョは太平洋の東側の海面温度が平年より高くなる現象。特に強まった状態はスーパーエルニーニョと呼ばれ、97～98年のほか2015～16年にも発生した。世界各地で豪雨など異常気象の一因になったと考えられている。

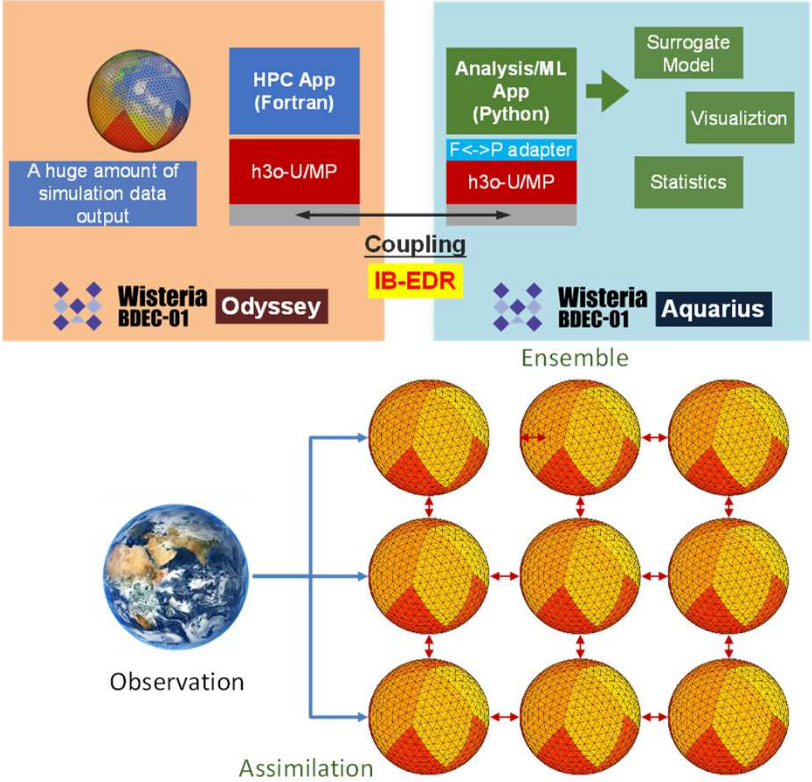
研究チームは雲の発生や動きを計算する手法を改良し、海洋との相互作用を精度良く予測する手法を開発した。理研のスーパーコンピューター「京」を使った数値計算で、赤道付近に巨大な雲ができて東に進む「マッデン・ジュリアン振動」と呼ぶ現象がスーパーエルニーニョ終息の引き金になったと突き止めた。

東西方向の数千キロに広がる巨大な雲から吹き出す東風が強まって海面付近の温かい海水が動かされ、入れ替わりに冷たい海水が下から上がってくるようになったという。

h3-Open-UTIL/MP

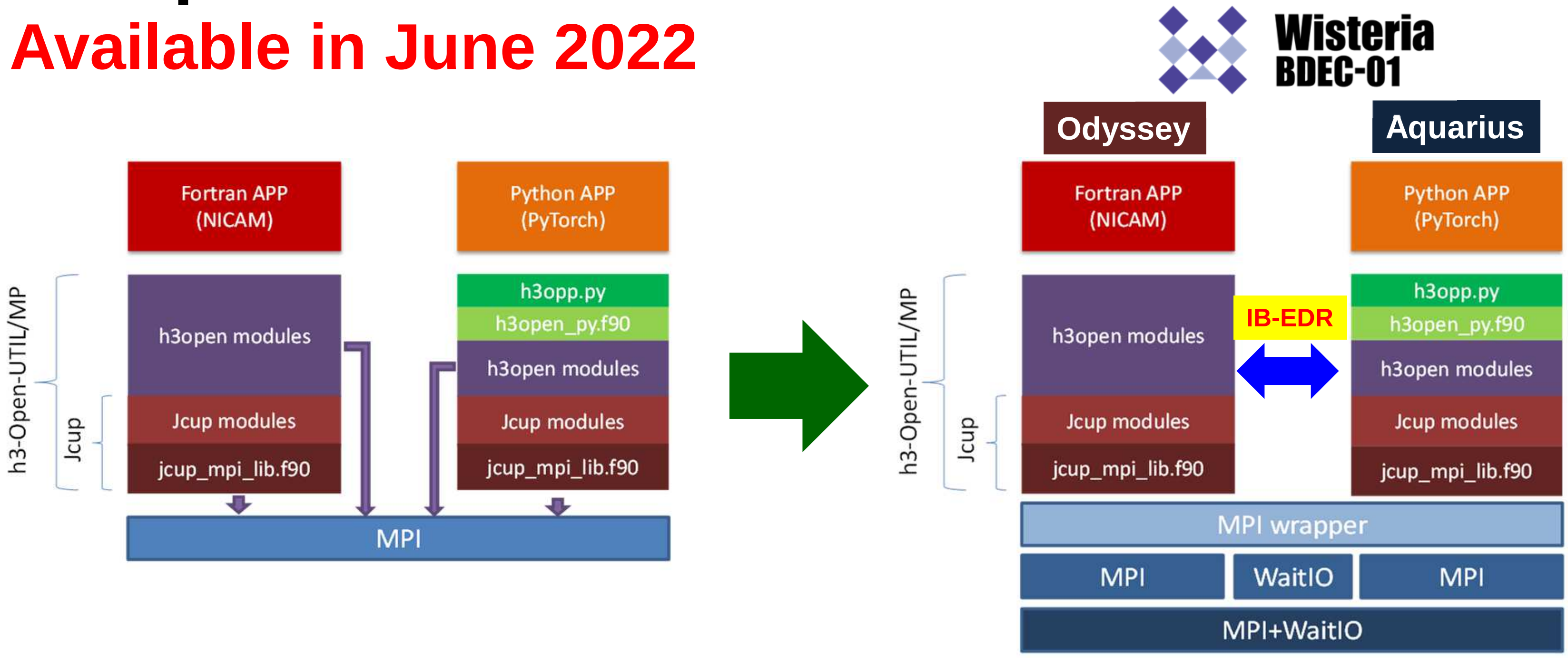
Integration of (S+D+L)
Multilevel Coupler/Data Assimilation

- Extended Version of Multi-Physics Coupler
- Coupling of Simulations on Odyssey and AI on Aquarius
  - Combined with WaitIO
- Special Capabilities for Data Assimilation and Ensemble Simulations
  - Data Assimilation by Coupled Codes
    - e.g. Atmosphere-Ocean
  - Ensemble Coupling



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

Available in June 2022



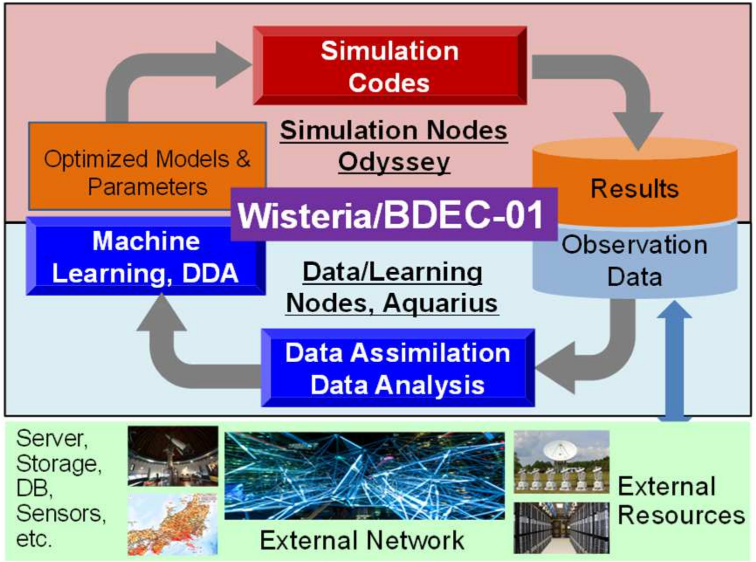
May 2021: MPI Only

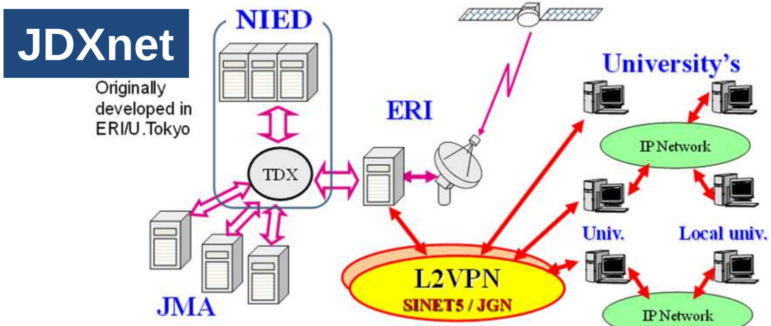
June 2022: Coupler＋WaitIO

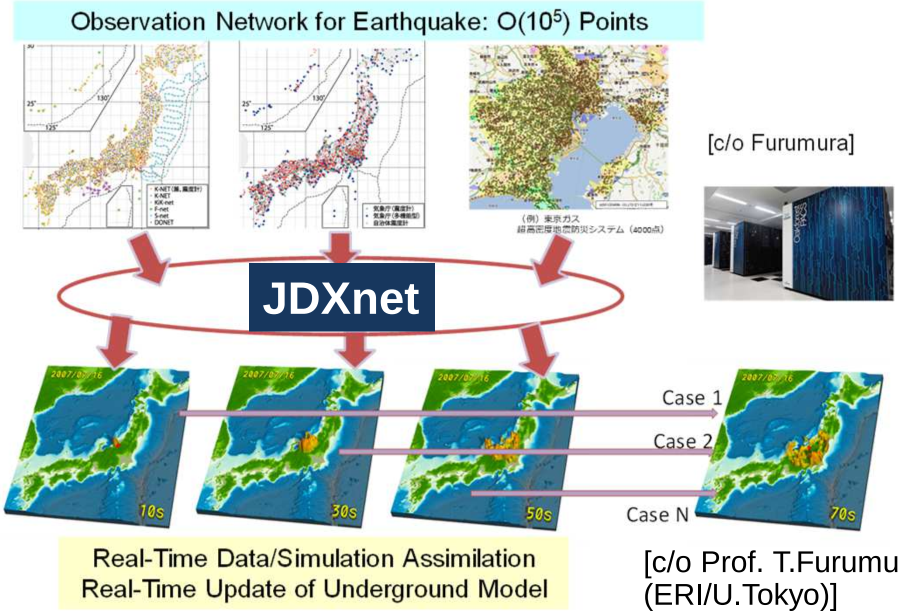
- Integration of (Simulation/Data/Learning)
  - Wisteria/BDEC-01
  - h3-Open-BDEC
- **Applications on Wisteria/BDEC-01 with h3-Open-BDEC**
  - **Earthquake Simulations**
  - (Global Cloud Simulation+AI) Coupling
  - Ensemble Coupling

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

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





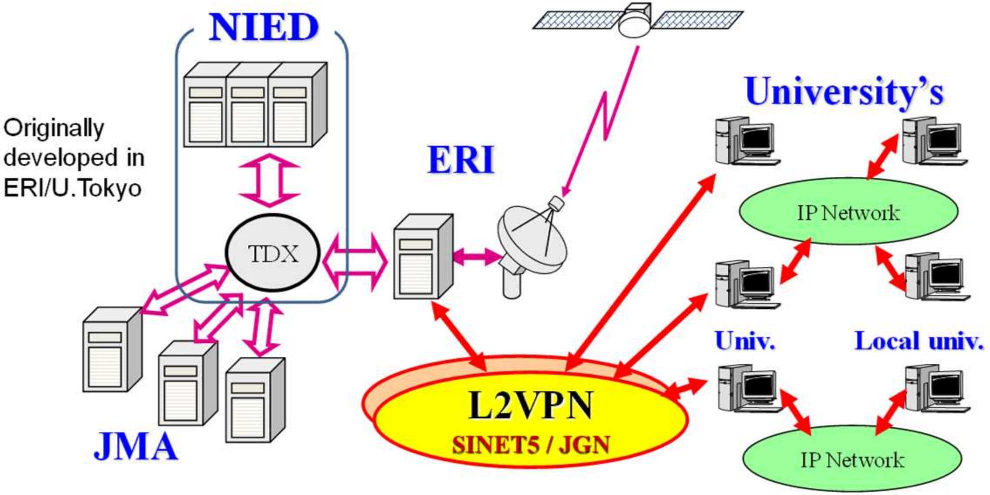


[c/o Prof. T.Furumura (ERI/U.Tokyo)]

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)]

Seism3D/OpenSWPC-DAF

Starting from (A+S: Assim.+Sim.) to (Pure S: Pure Simulation), 3D FDM [Prof. T. Furumura, ERI/U.Tokyo]

Residual

$$\underbrace{x_n^a}_{\text{Assim.}} = \underbrace{x_n^f}_{\text{Comp.}} + W(\underbrace{y_n}_{\text{Obs.}} - H\underbrace{x_n^f}_{\text{Comp.}})$$

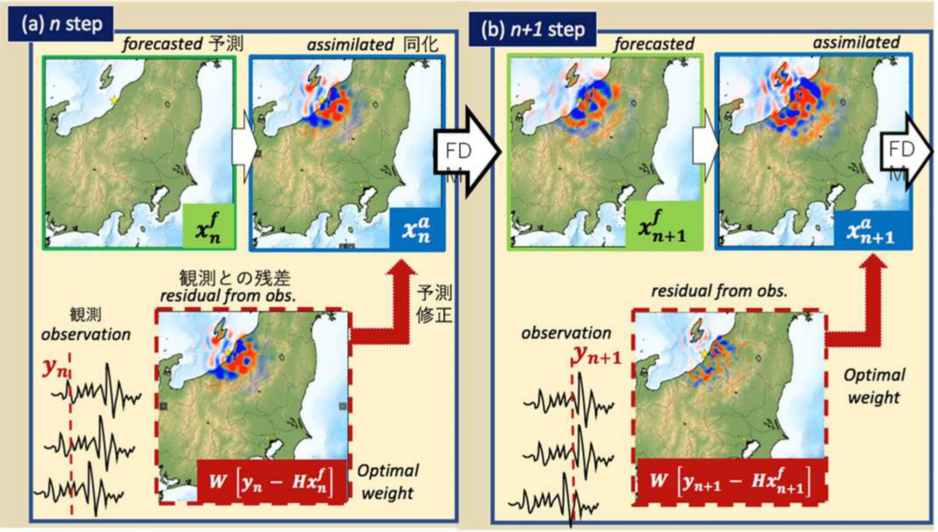
$$\underbrace{x_{n+1}^f}_{\text{Comp.}} = F\underbrace{x_n^a}_{\text{Assim.}}$$

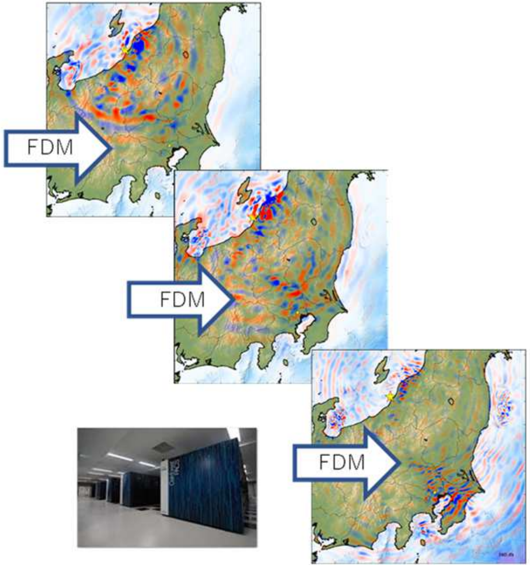
n: Time Step

W: **Weighting Matrix**

F: Wave Propagation simulation

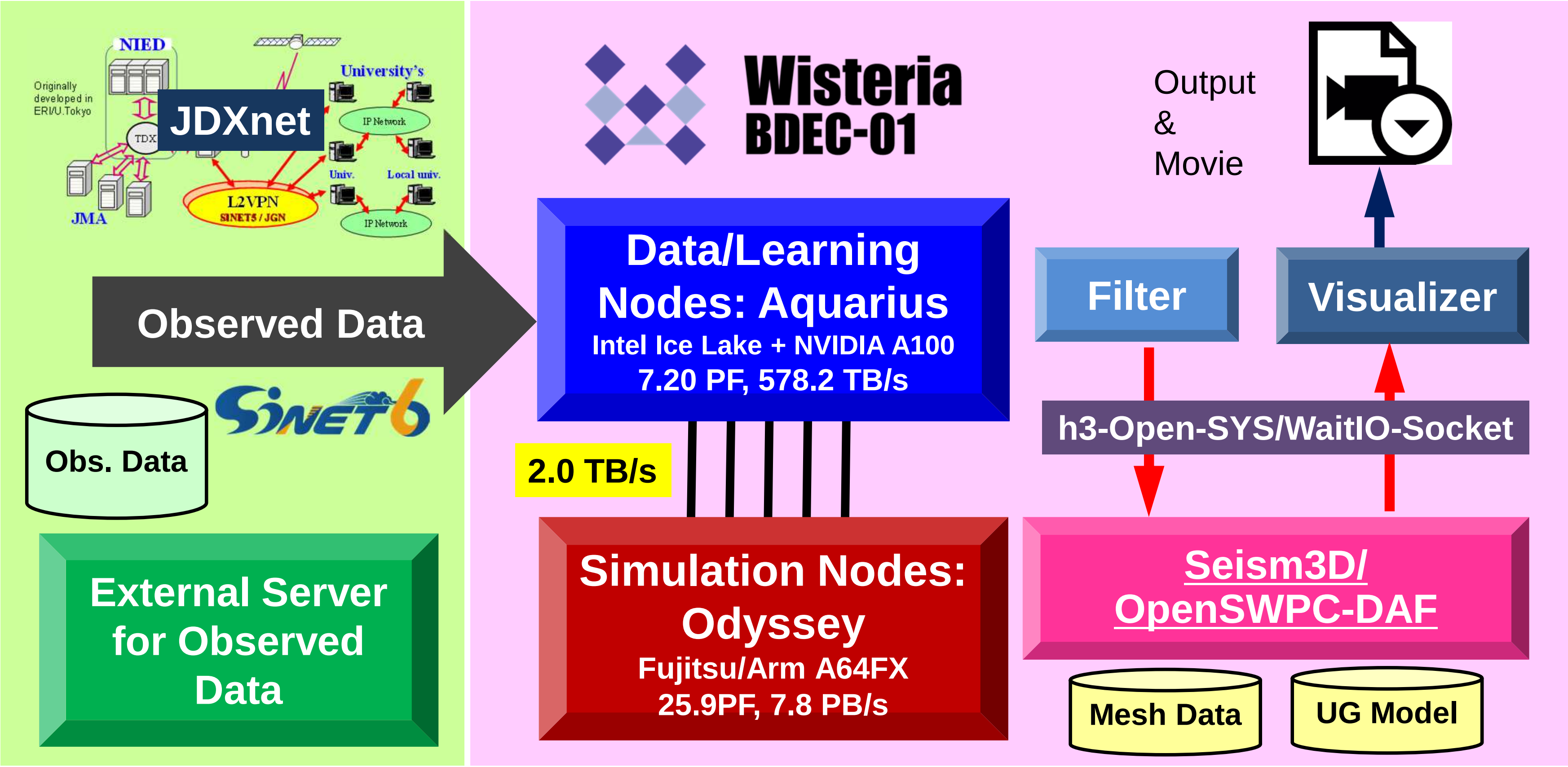
(A+S) Assimilation+Simulation → **(Pure S) Pure Simulation/Forecast**





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

3D Earthquake Simulation with Real-Time Data Observation/Assimilation on Wisteria/BDEC-01



Communications by WaitIO-Socket

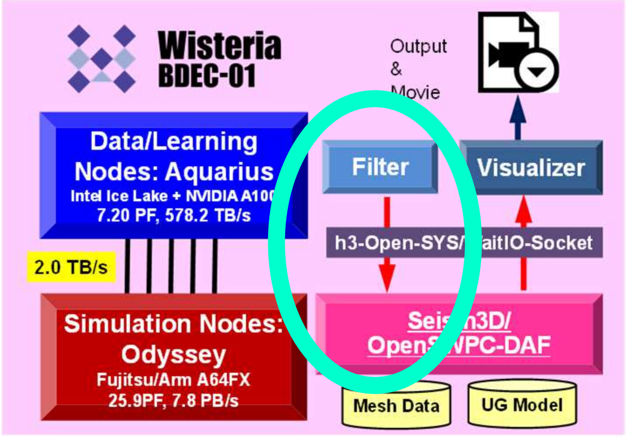
[Kasai et al. 2021]

Aquarius: SEND

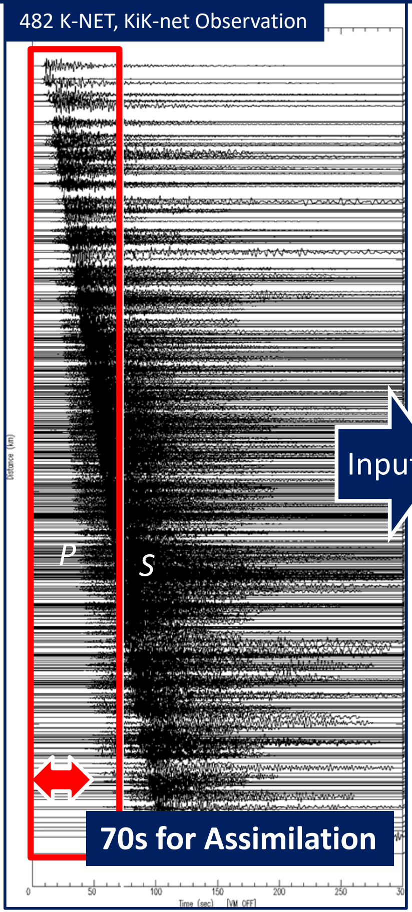
```
program dmy_filter
<省略: 型宣言等>
call mpi_init (ierr)
call mpi_comm_size (MPI_COMM_WORLD, nprocs, ierr)
call mpi_comm_rank (MPI_COMM_WORLD, myrank,  ierr)
call WAITIO_CREATE_UNIVERSE (WAITIO_COMM_UNIVERSE, ierr)

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
```

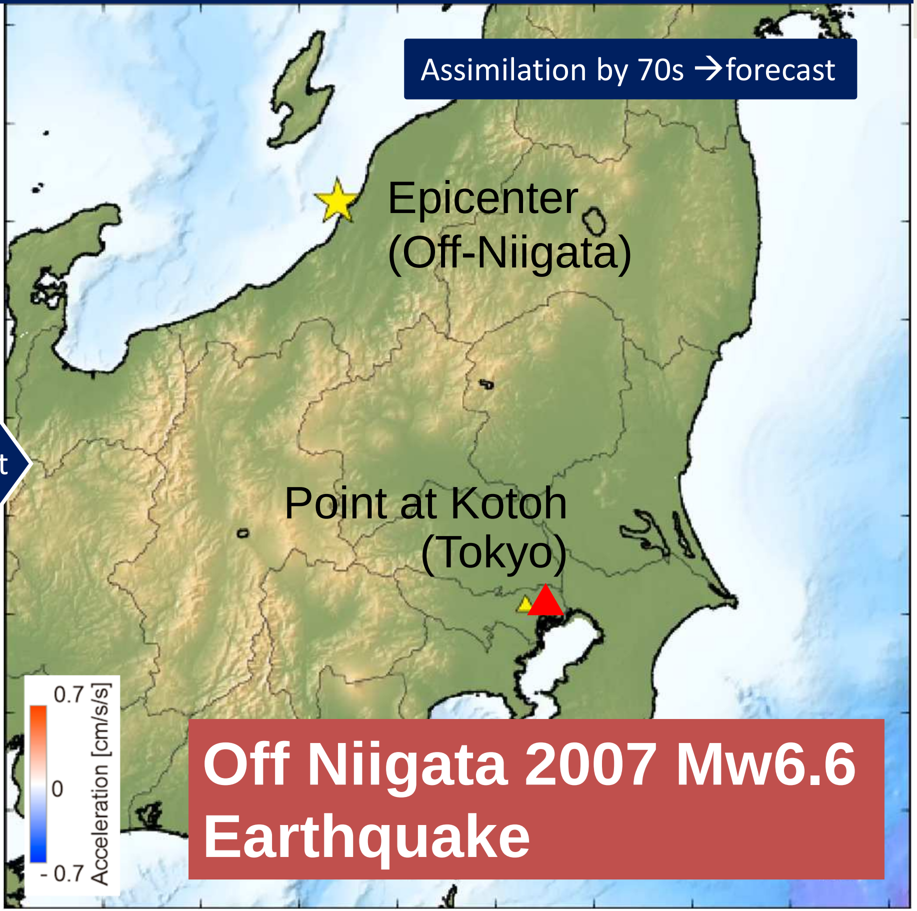
Odyssey: RECV

```
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,...)
```

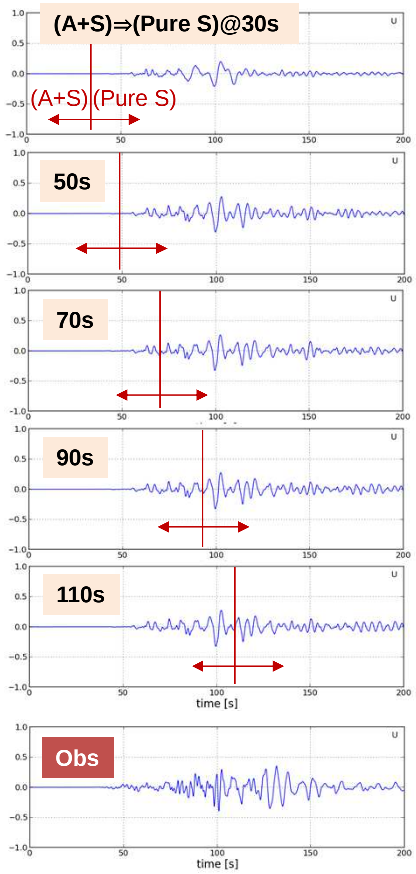


Data Assimilation + Pure Simulation/Forecast



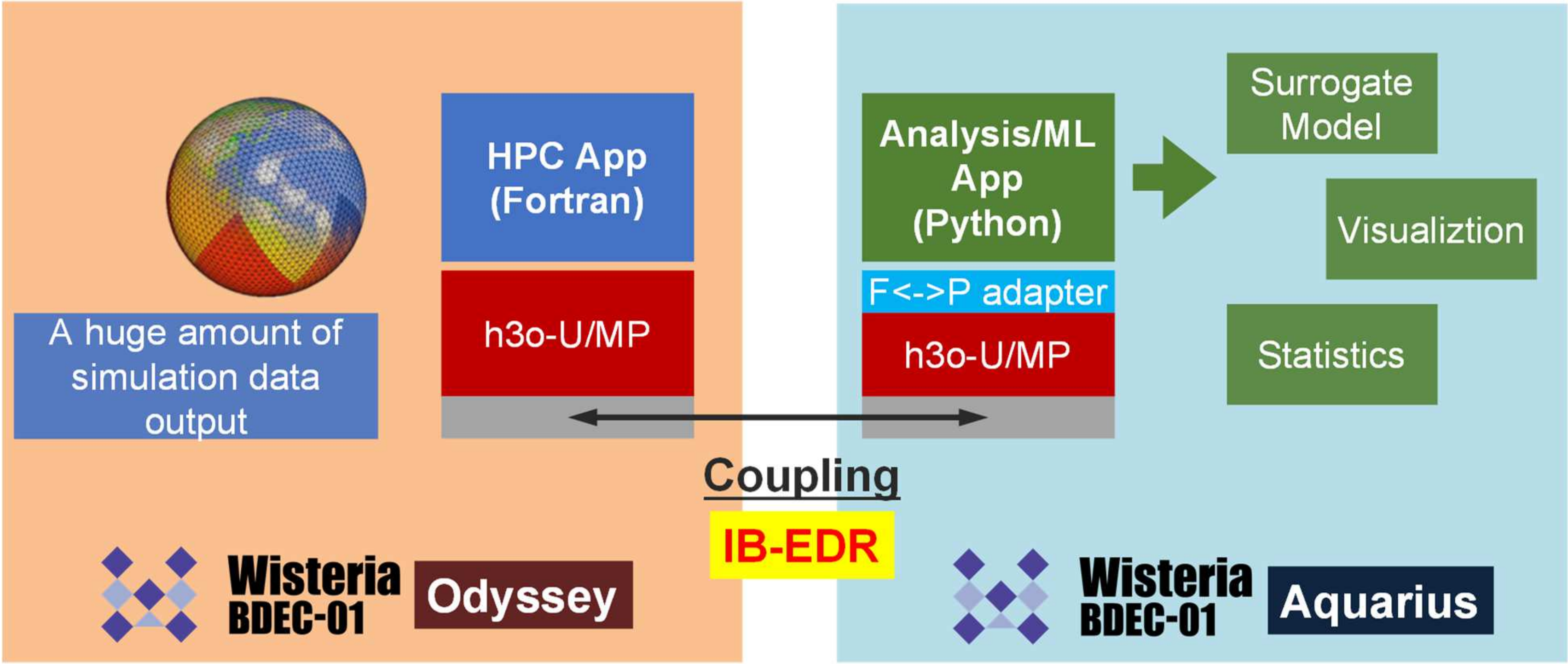


Results at Kotoh ▲ (N.KOTH)
N 35° 37.0'
E 139° 46.9'



- Integration of (Simulation/Data/Learning)
  - Wisteria/BDEC-01
  - h3-Open-BDEC
- **Applications on Wisteria/BDEC-01 with h3-Open-BDEC**
  - Earthquake Simulations
  - **(Global Cloud Simulation+AI) Coupling**
  - Ensemble Coupling

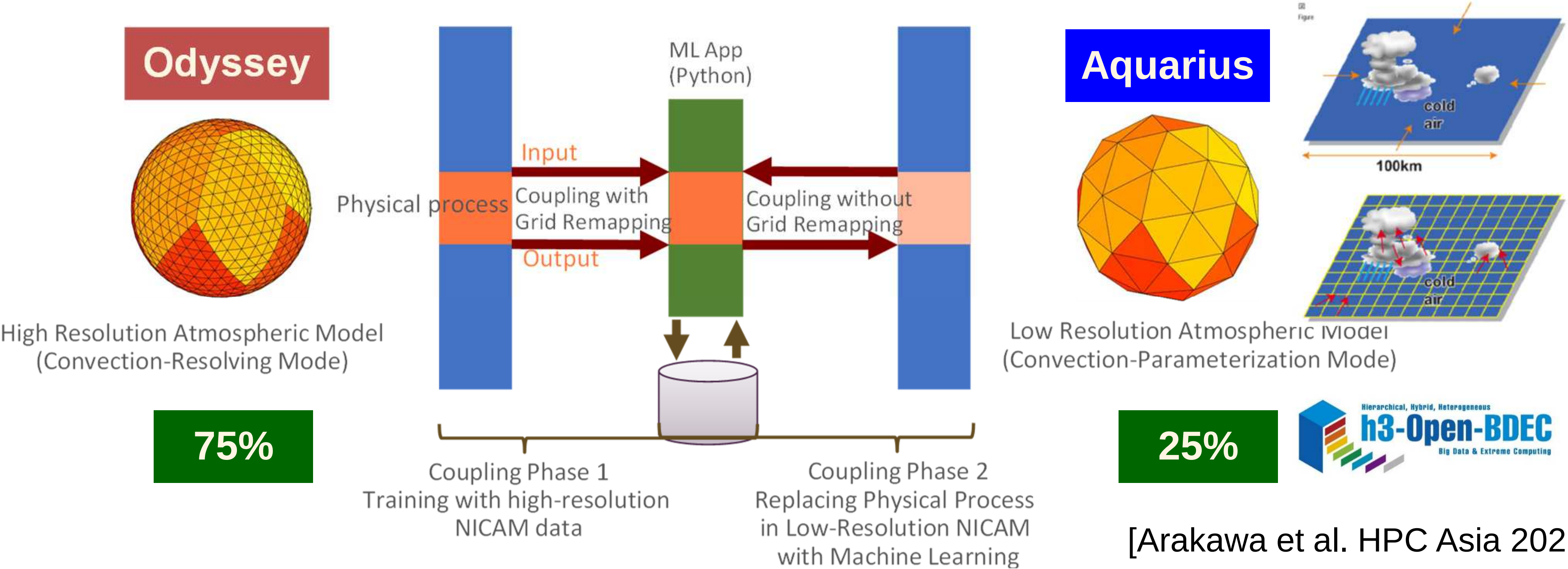
h3-Open-UTIL/MP (h3o-U/MP) Extended Multiphysics Coupler



Global Cloud Physics (1/2)

High-Resolution Model/Low-Resolution Model for Parameterization

Replace Low-Resolution Model Simulations (25%) with AI (trained by results of High-Resolution Model)



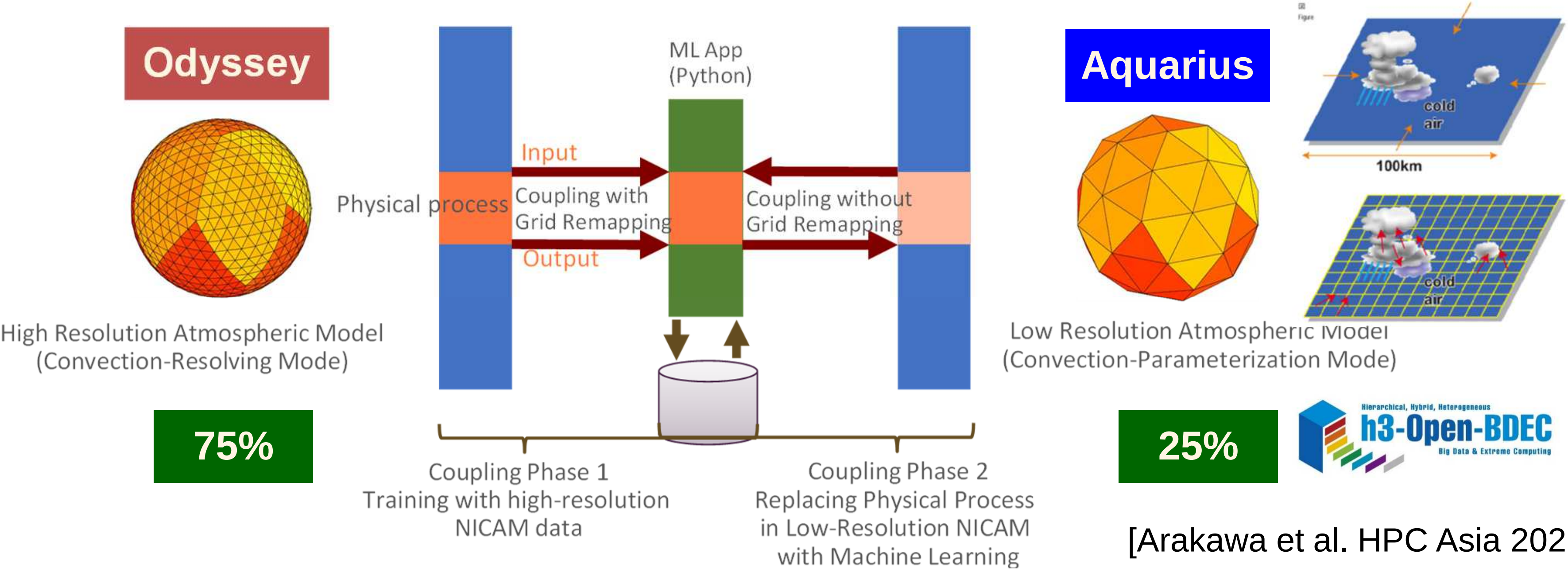
[Arakawa et al. HPC Asia 2022]

Global Cloud Physics (2/2)

High-Resolution Model: NICAM in Fortran on Odyssey

Training/Inferencing: (PyTorch: 3-Layer Perceptron) on Aquarius

Seamless Operations on Odyssey/Aquarius by h3-Open-BDEC: Training of AI by Simulation, and Integration of (Simulation+Inferencing)



[Arakawa et al. HPC Asia 2022]

Comparison: Simulations and AI

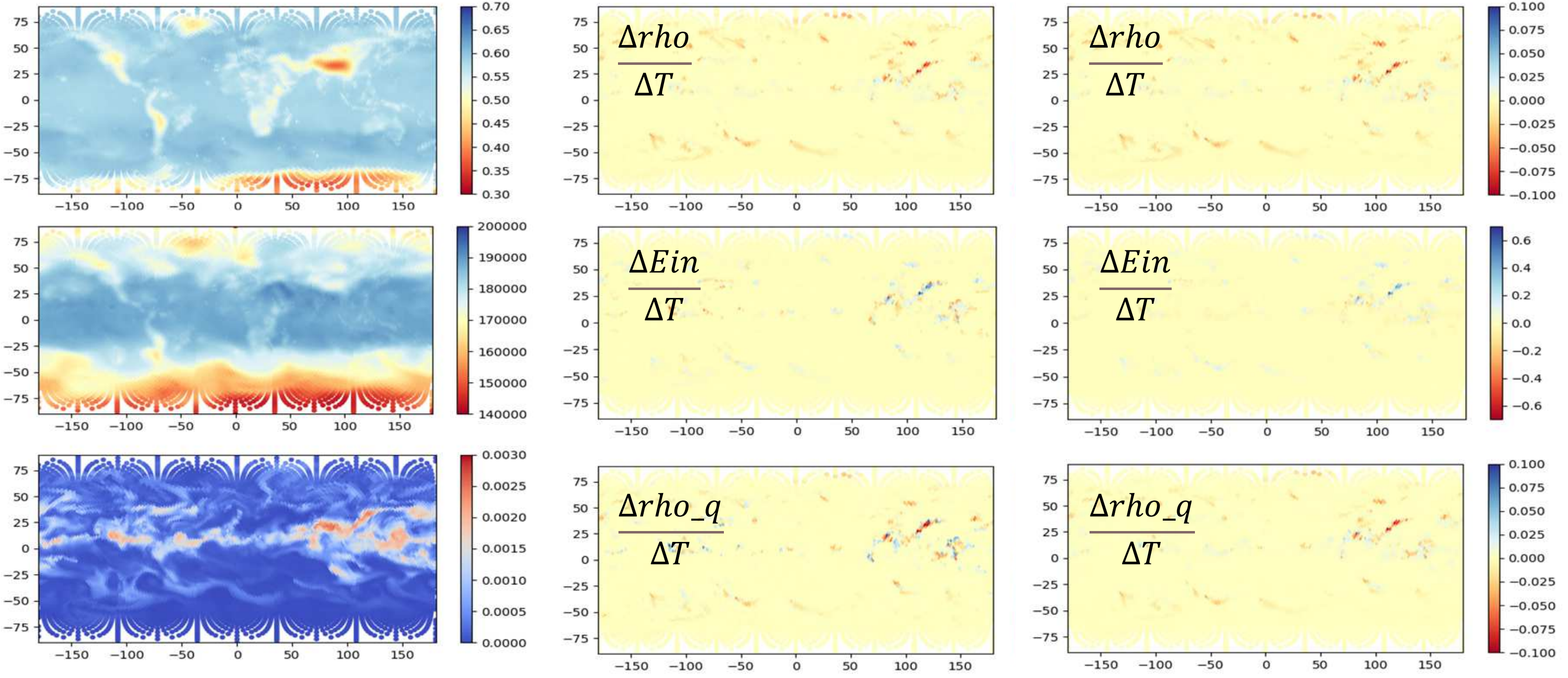
Input: Air Density, Internal Energy of Air, Vapor Density

Output: Gradient in Time Direction of 3 Parameters [Arakawa et al. HPC Asia 2022]

Air Density *rho*

Internal Energy *Ein*

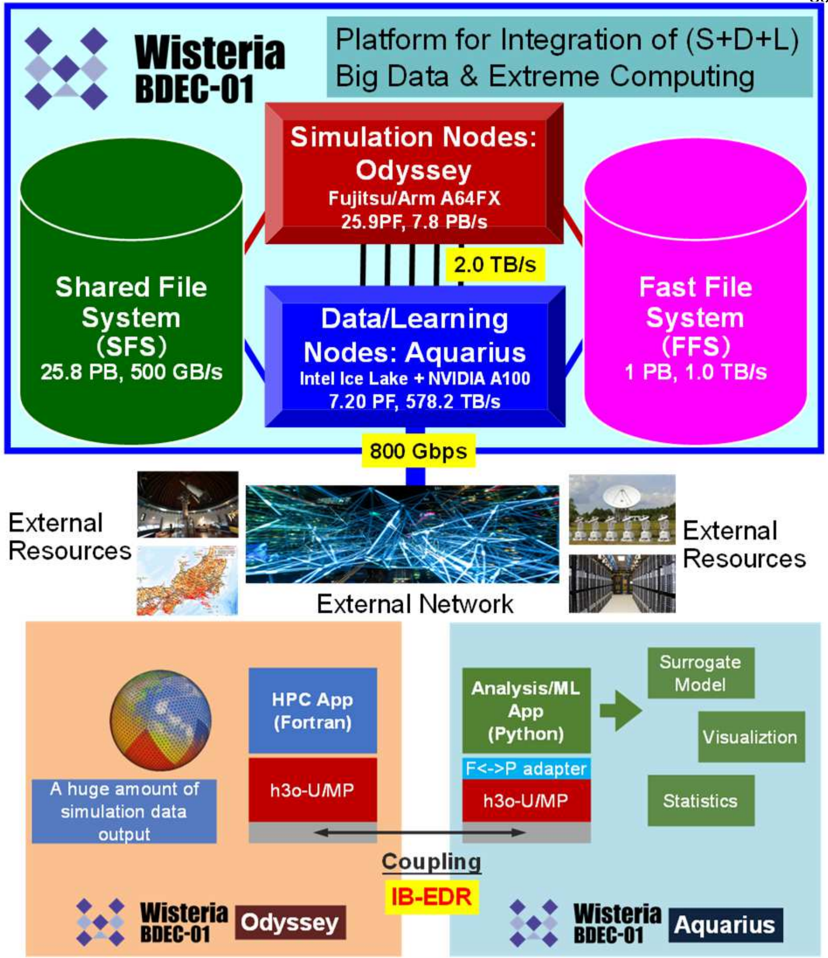
Vapor Density *rho_q*



Input **Simulations** **Prediction by ML/NN**

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.
- More flexible (& complicated) policy needed



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
```

Summary

- Integration of (Simulation/Data/Learning) at ITC/U.Tokyo
- Wisteria/BDEC-01
- h3-Open-BDEC
- Applications

Wisteria BDEC-01

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