



Wisteria/BDEC-01

「計算・データ・学習」融合 スーパーコンピュータシステム



東京大学情報基盤センター

<https://www.cc.u-tokyo.ac.jp/supercomputer/wisteria/service/>

東京大学情報基盤センター



- 東京大学大型計算機センター（1965年）
- 東京大学情報基盤センター（1999年～）
 - 全国共同利用施設
 - 学際大規模情報基盤共同利用・共同研究拠点 中核拠点（2010年～）
 - 革新的ハイパフォーマンス・コンピューティング・インフラ(HPCI) 構成機関（2010年～）
 - 最先端共同HPC基盤施設（JCAHPC）（2013年～）
 - 筑波大学計算科学研究センター・東大情報基盤センター：OFP
- 2023年10月現在
 - 1式のシステムを運用
 - Oakforest-PACS（OFP）：2022年3月末に運用終了
 - Oakbridge-CX（OBCX）：2023年9月末に運用終了
 - Wisteria/BDEC-01（「計算・データ・学習」融合スーパーコンピュータシステム）：2021年5月運用開始
 - ^{2023/10/19}データ活用社会創成プラットフォーム（mdx）：2021年3月設置



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

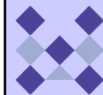
OBCX
(Fujitsu)
6.61 PF

Hitachi HA8000
T2K Today
140 TF

Oakforest-
PACS (Fujitsu)
25.0 PF

OFP-II
150+ PF

Fujitsu FX10
Oakleaf-FX
1.13 PF

 Wisteria
BDEC-01 Fujitsu
33.1 PF

BDEC-
02
250+ PF

東京大学情報基盤
センターのスパコン
利用者2,600+名
55%は学外

Reedbush-
U/H/L (SGI-HPE)
3.36 PF

Ipomoea-01 25PB

Ipomoea-
03

Ipomoea-02

2001-2005

2006-2010

2011-2015

2016-2020

2021-2025

2026-2030

Hitachi SR8000

1.024 PF

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

Oakforest-

PACS (Fujitsu)

1.1 PF

OFP-II

150+ PF

Accelerators

AMD Opteron

Intel Xeon Phi

Fujitsu FX10

Oakleaf-FX

1.13 PF

Wisteria BDEC-01 Fujitsu

33.1 PF

BDEC-

02

250+ PF

疑似ベクトル

汎用CPU

加速装置付

SPACR64 IXfx

Reedbush-

U/H/L (SGI-HPE)

3.36 PF

A64FX,

Intel Icelake+

NVIDIA A100

Accelerators

Intel BDW +

NVIDIA P100

Ipomoea-01 25PB

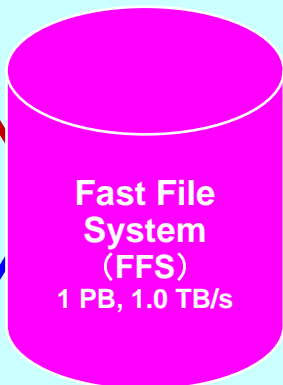
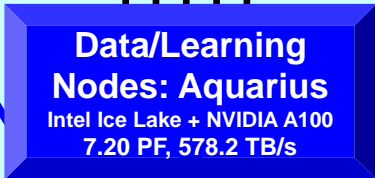
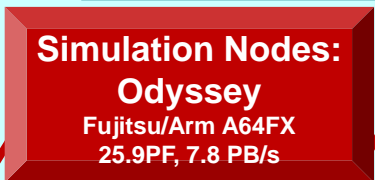
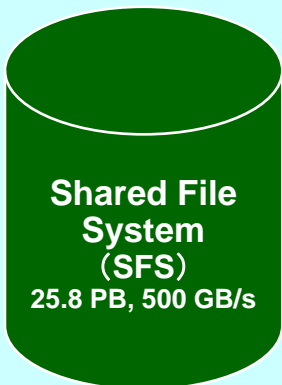
Ipomoea-

03

Ipomoea-02



Platform for Integration of (S+D+L)
Big Data & Extreme Computing



800 Gbps

External Resources



External Network



External Resources



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

Reedbush (HPE, Intel BDW + NVIDIA P100 (Pascal))

- データ解析・シミュレーション融合スーパーコンピュータ
- 2016年7月～2021年11月末(引退)
- 東大ITC初のGPUクラスタ, ピーク性能3.36 PF

Oakforest-PACS (OFP) (Fujitsu, Intel Xeon Phi (KNL))

- JCAHPC (筑波大CCS・東大ITC), 2016年10月～2022年3月末(予定)
- 25 PF, #39 in 58th TOP 500 (November 2021)

Oakbridge-CX (OBCX) (Fujitsu, Intel Xeon CLX)

- 2019年7月～2023年9月末(予定)
- 6.61 PF, #110 in 58th TOP500-Sep 2023 (Plan)



Wisteria/BDEC-01 (Fujitsu)

- シミュレーションノード群 (Odyssey) : A64FX (#17)**
- データ・学習ノード群 (Aquarius) : Intel Icelake+NVIDIA A100 (#106)**
- 33.1 PF, #13 in 57th TOP 500, 2021年5月14日運用開始
- 「計算・データ・学習(S+D+L)」融合のためのプラットフォーム
- 革新的ソフトウェア基盤「h3-Open-BDEC」
(科研費基盤(S) 2019年度～2023年度)



Simulation Nodes
(Odyssey)

2023/10/19



Data/Learning Nodes
(Aquarius)



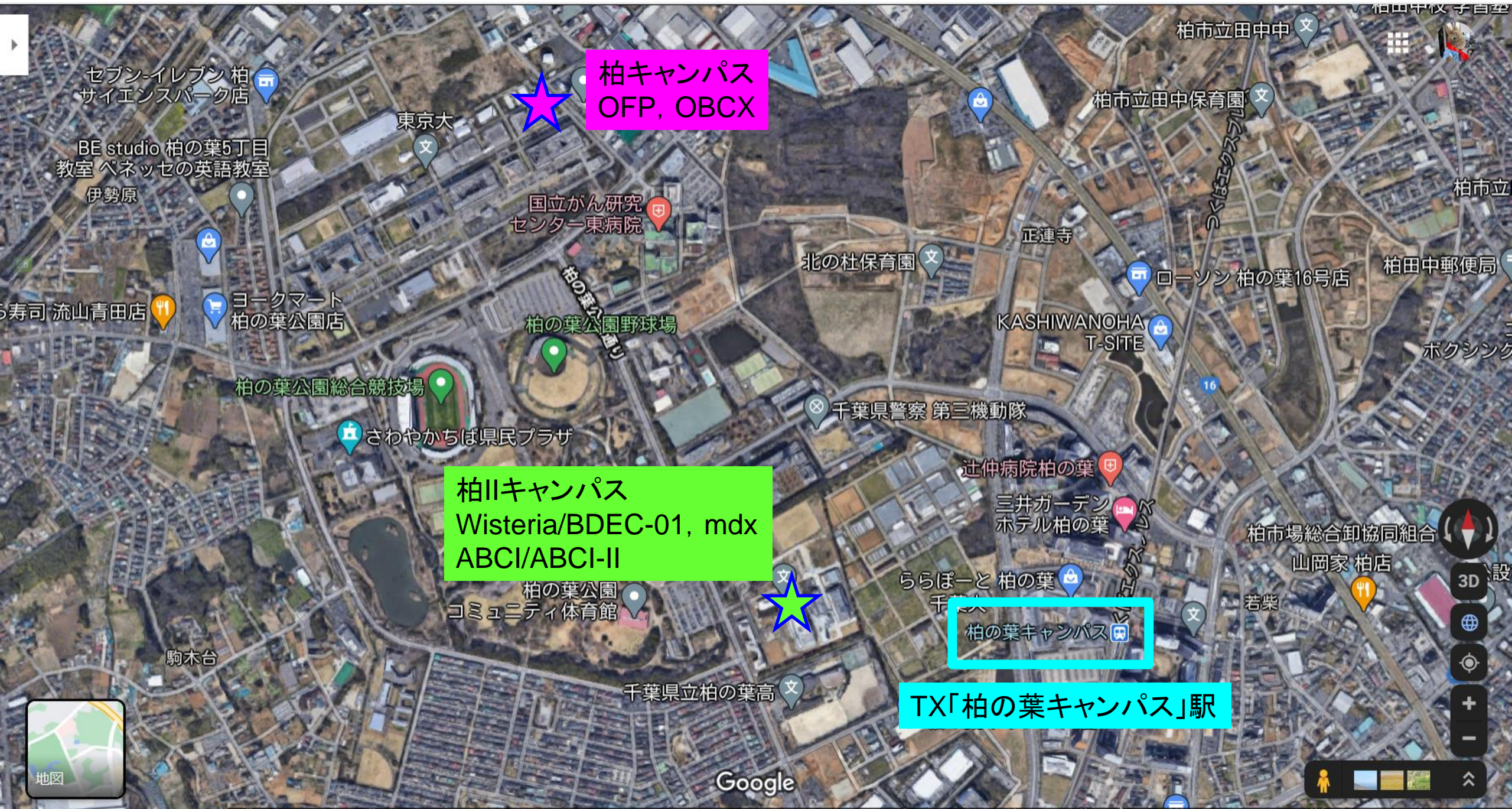
Reedbush



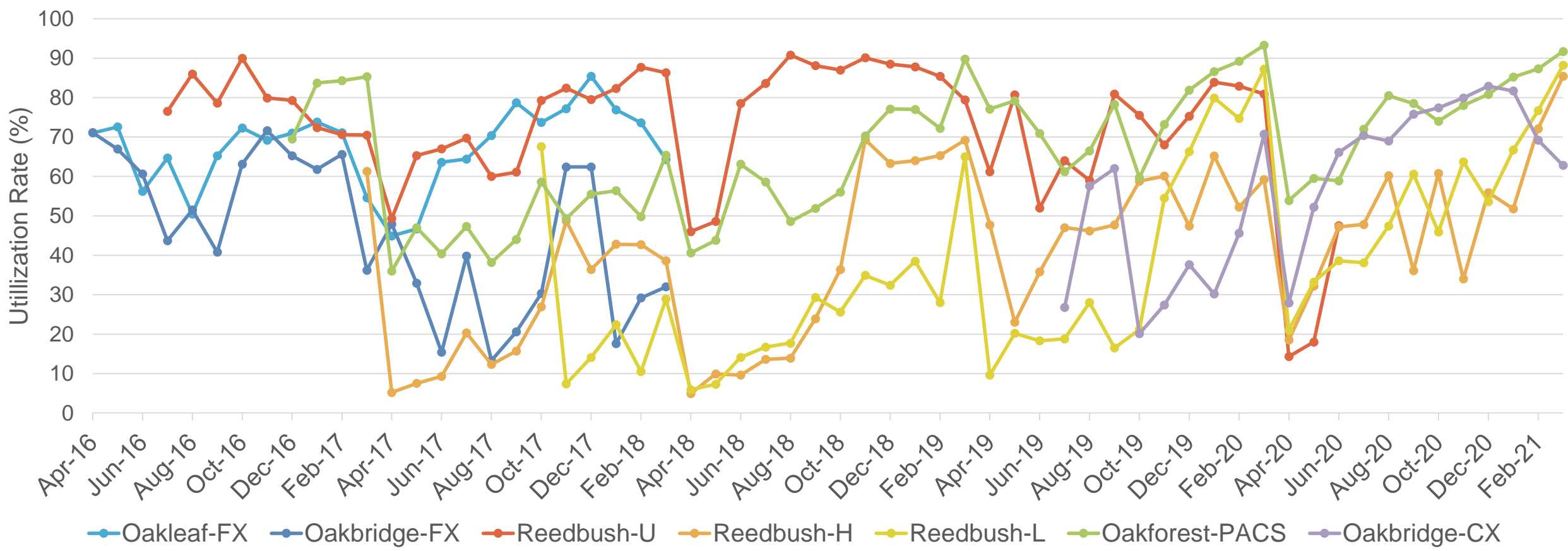
Oakforest-PACS



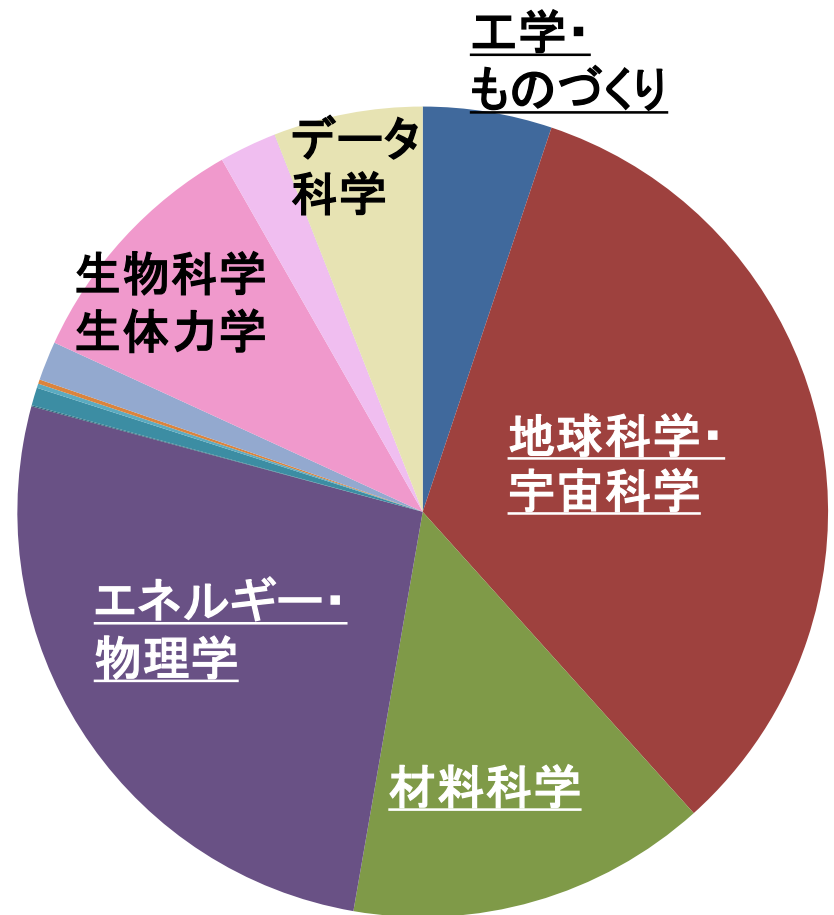
Oakbridge-CX



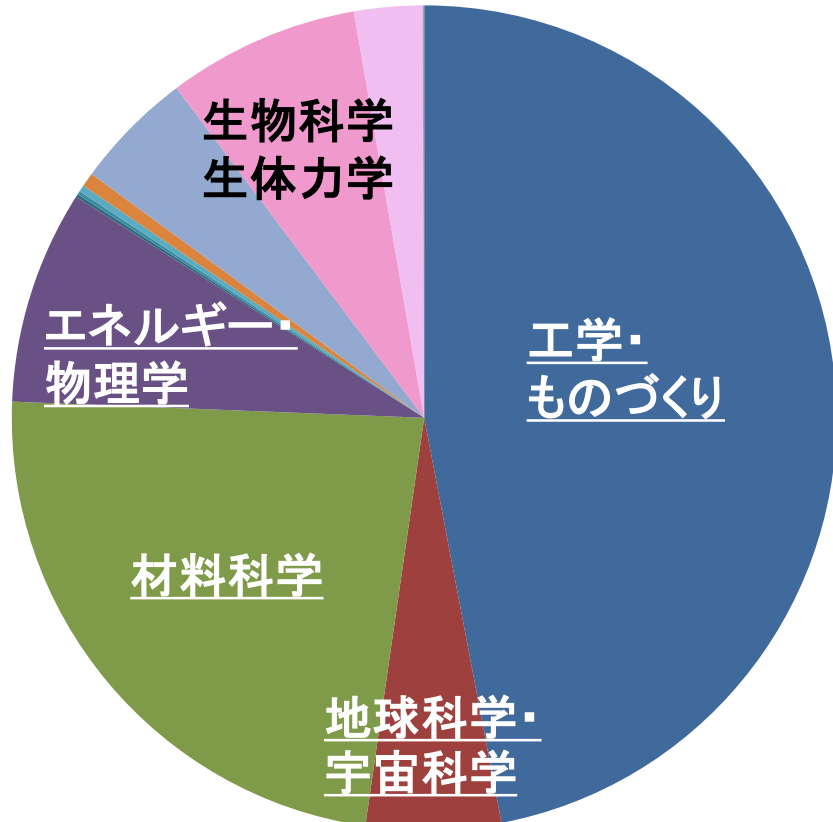
各システムの月平均利用率履歴



研究分野別利用CPU時間割合(2020年度)



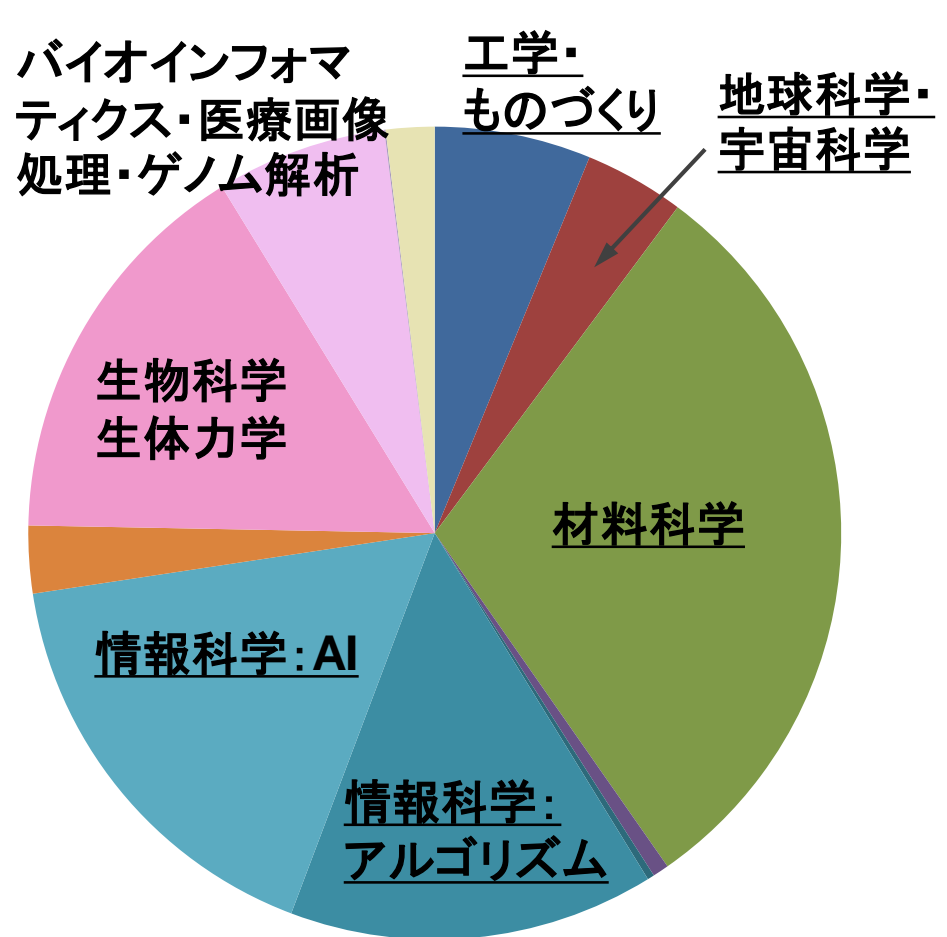
メニィコアクラスタ
Intel Xeon Phi
(Oakforest-PACS)



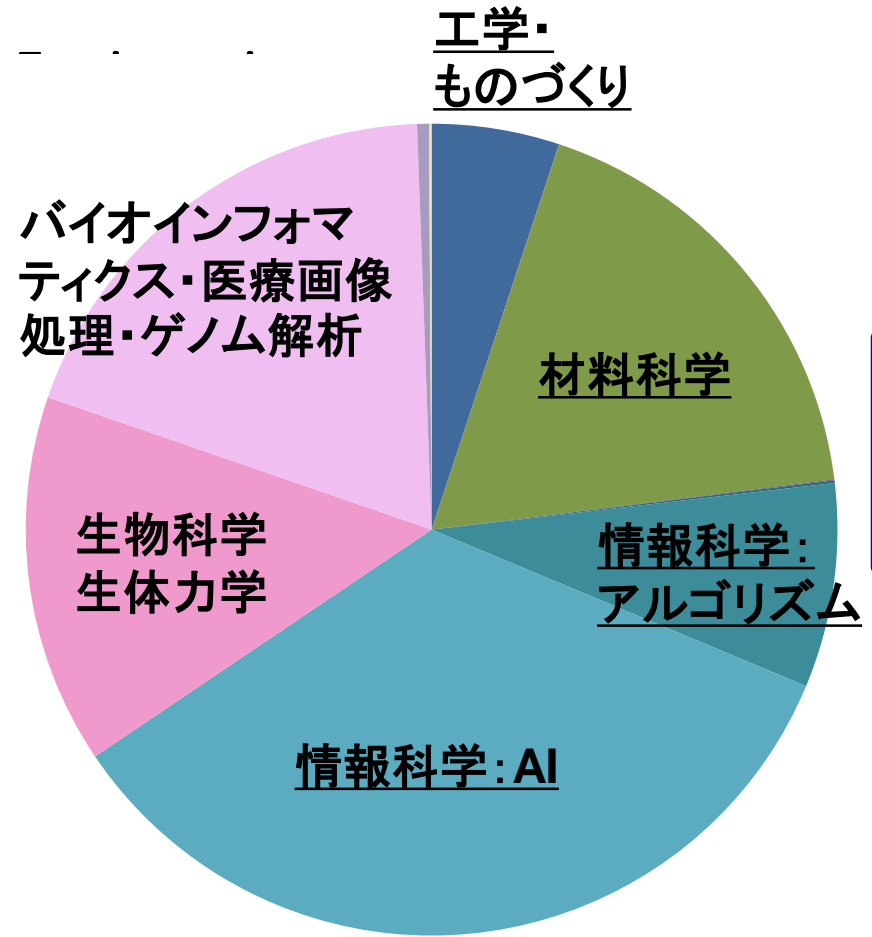
マルチコアクラスタ
Intel CLX
(Oakbridge-CX)

- 工学・ものづくり
- 地球科学・宇宙科学
- 材料科学
- エネルギー・物理学
- 情報科学: システム
- 情報科学: アルゴリズム
- 情報科学: AI
- 教育
- 産業利用
- 生物科学・生体力学
- バイオインフォマティクス
- 社会科学・経済学
- データ科学・データ同化

研究分野別利用CPU時間割合 (2020年度)



GPUクラスタ (2GPUs/node)
Intel BDW + NVIDIA P100
(Reedbush-H)

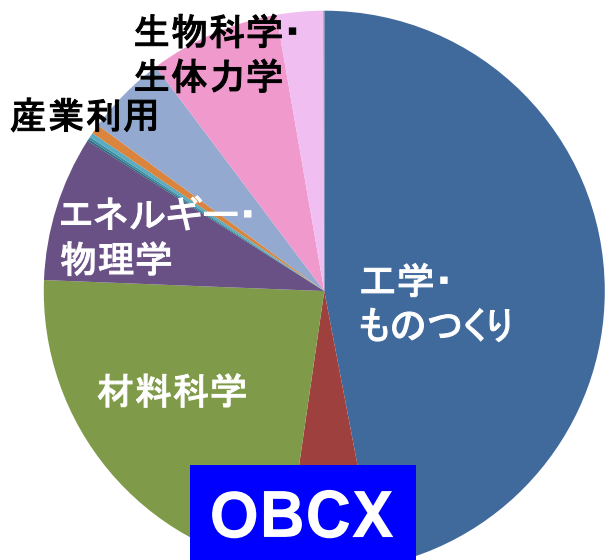
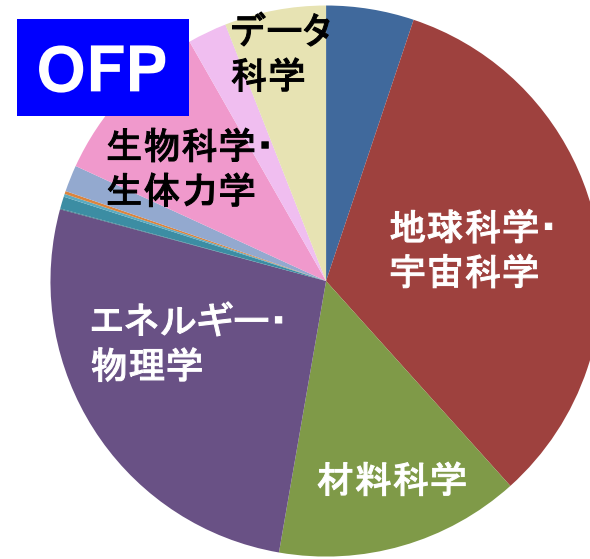
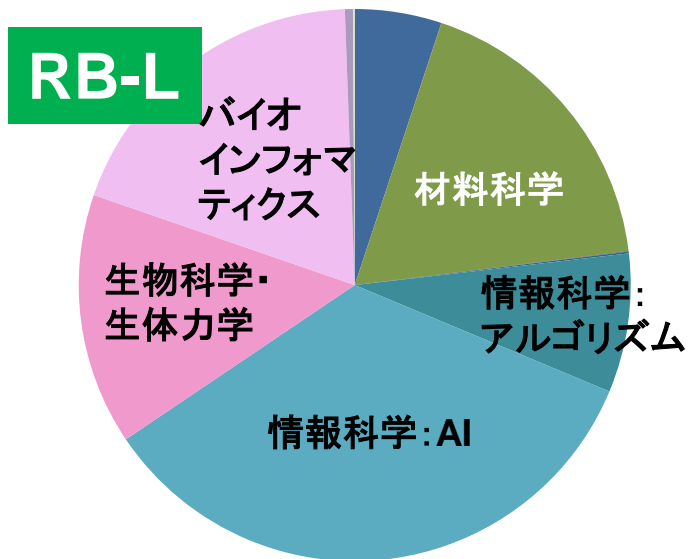
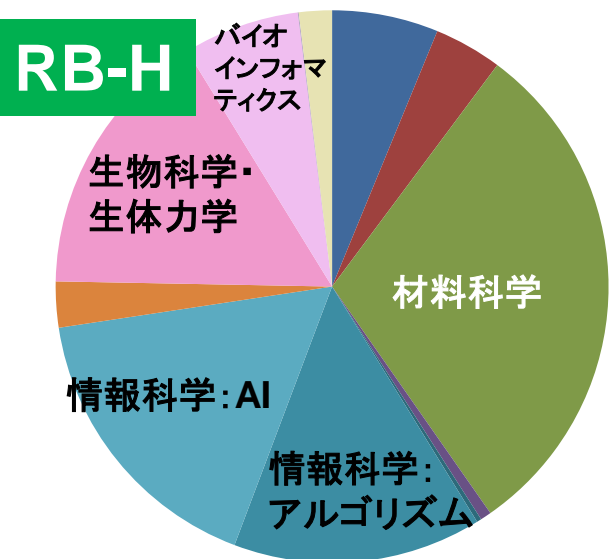


GPUクラスタ (4GPUs/node)
Intel BDW + NVIDIA P100
(Reedbush-L)

- 工学・ものづくり
- 地球科学・宇宙科学
- 材料科学
- エネルギー・物理学
- 情報科学: システム
- 情報科学: アルゴリズム
- 情報科学: AI
- 教育
- 産業利用
- 生物科学・生体力学
- バイオインフォマティクス
- 社会科学・経済学
- データ科学・データ同化

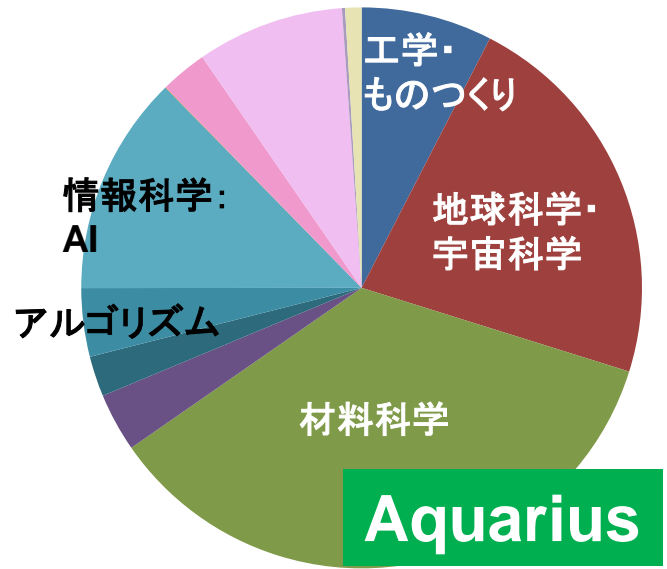
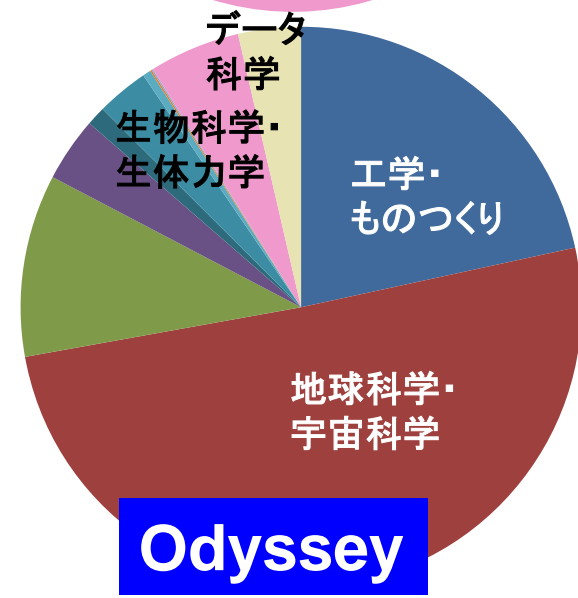
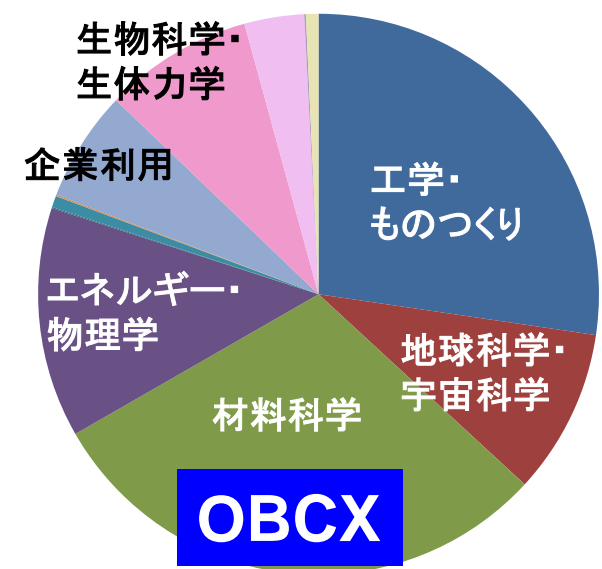
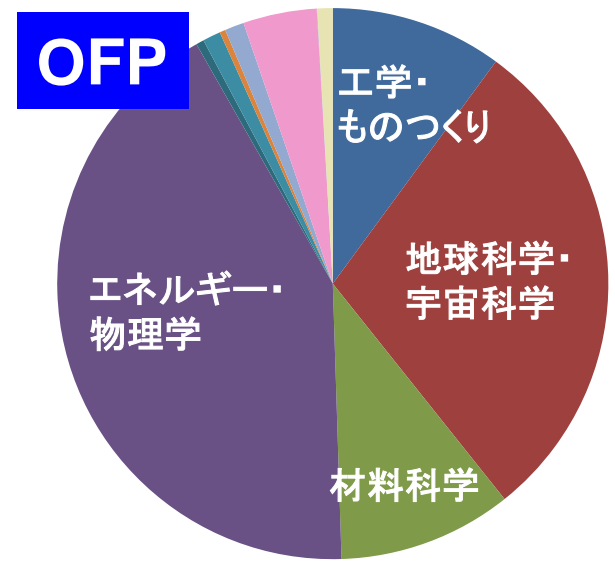
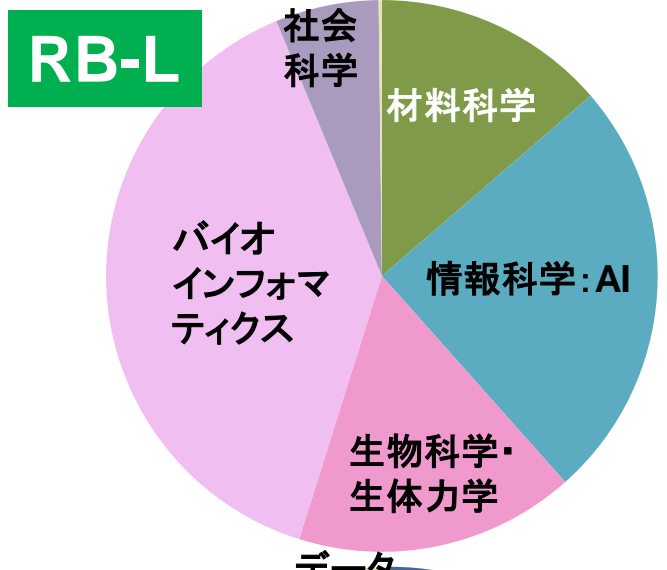
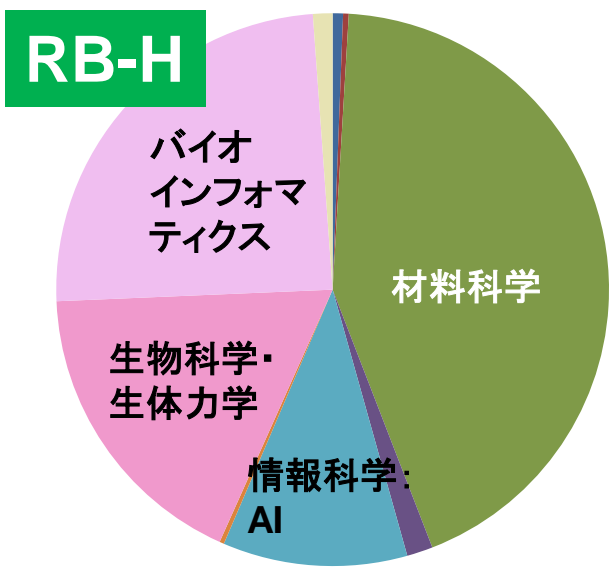
2020年度分野別 ■汎用CPU, ■GPU

- 工学・ものづくり
- 地球科学・宇宙科学
- 材料科学
- エネルギー・物理学
- 情報科学:システム
- 情報科学:アルゴリズム
- 情報科学:AI
- 教育
- 産業利用
- 生物科学・生体力学
- バイオインフォマティクス
- 社会科学・経済学
- データ科学・データ同化



2021年度分野別 ■汎用CPU, ■GPU

Odyssey, Aquariusは8月以降, RB-H, RB-Lは11月末時点

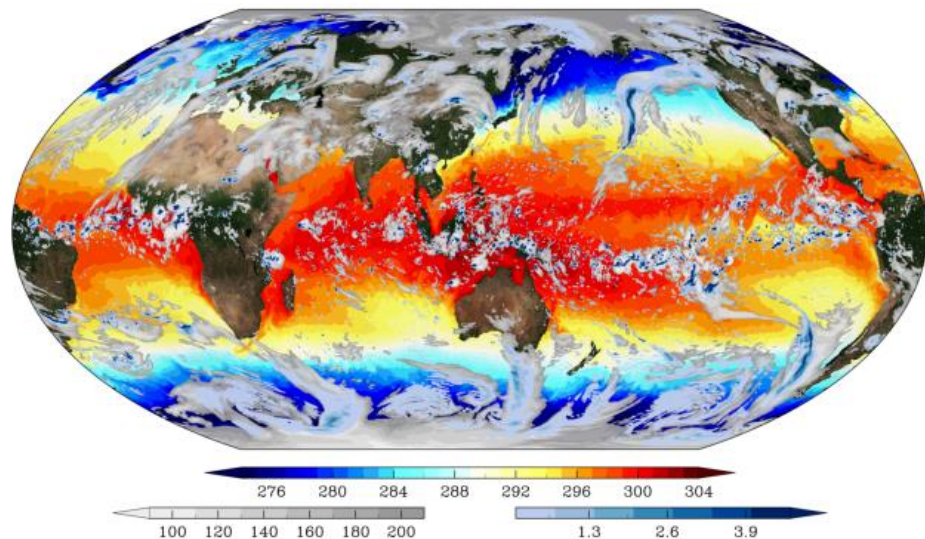


- 工学・ものづくり
- 地球科学・宇宙科学
- 材料科学
- エネルギー・物理学
- 情報科学: システム
- 情報科学: アルゴリズム
- 情報科学: AI
- 教育
- 産業利用
- 生物科学・生体力学
- バイオインフォマティクス
- 社会科学・経済学
- データ科学・データ同化

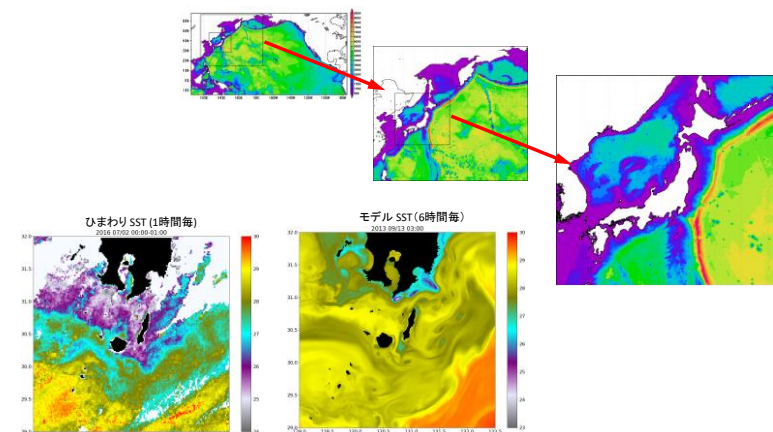
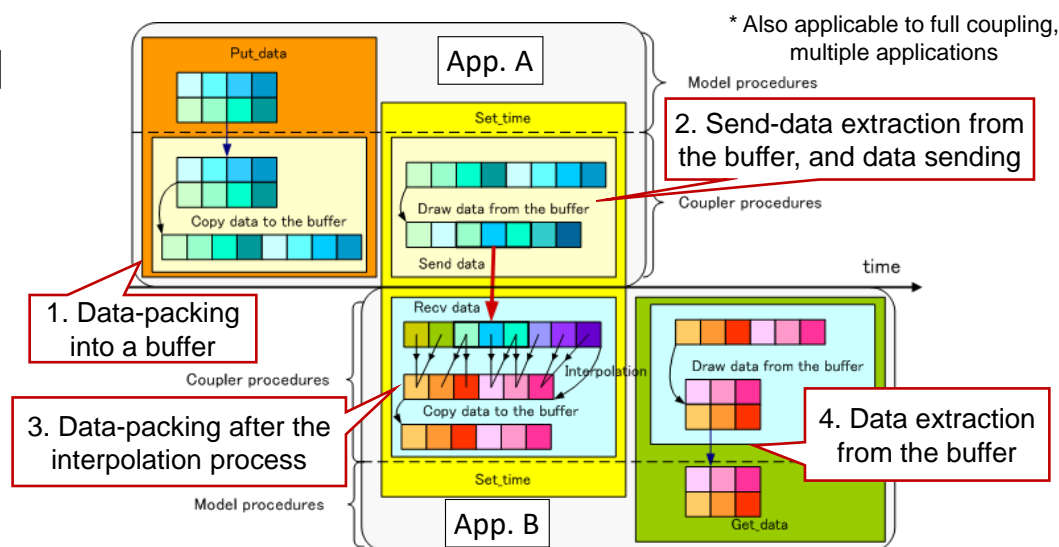
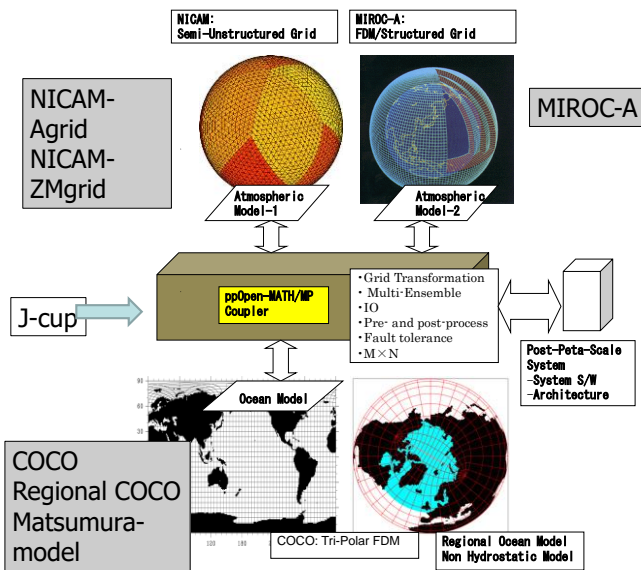
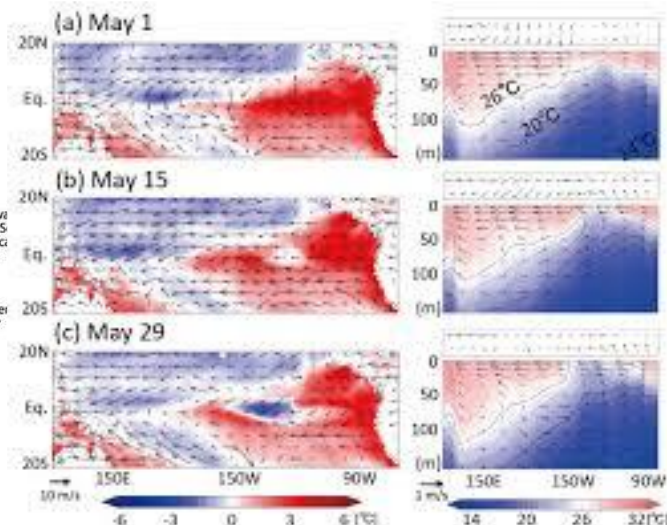
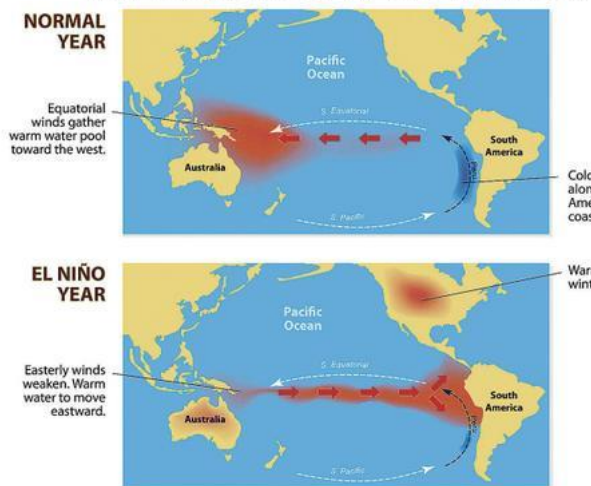
地球科学・宇宙科学分野ではOFP ⇒ Wisteria/BDEC-01への移行が順調に進んでいる

全地球大気環境シミュレーション

東大大気海洋研究所, 東大理学系研究科等



THE EL NIÑO PHENOMENON



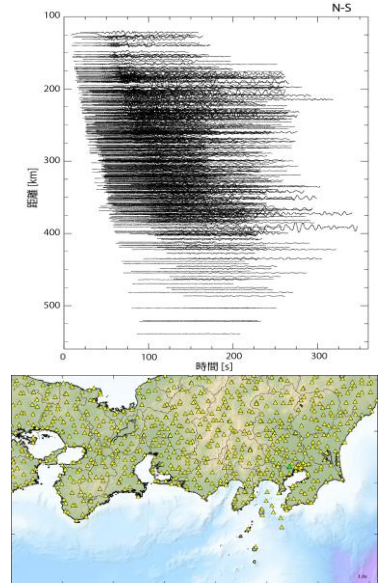
[画像提供: 佐藤正樹教授・羽角博康教授(東大・大気海洋研)]

地震シミュレーション・地殻変動

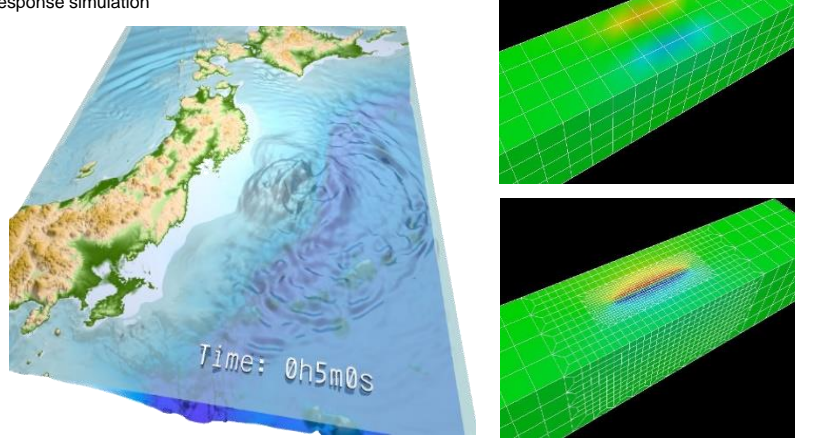
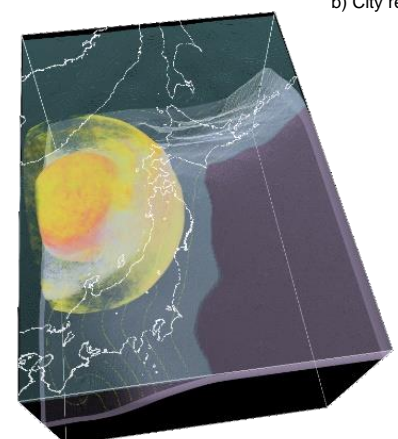
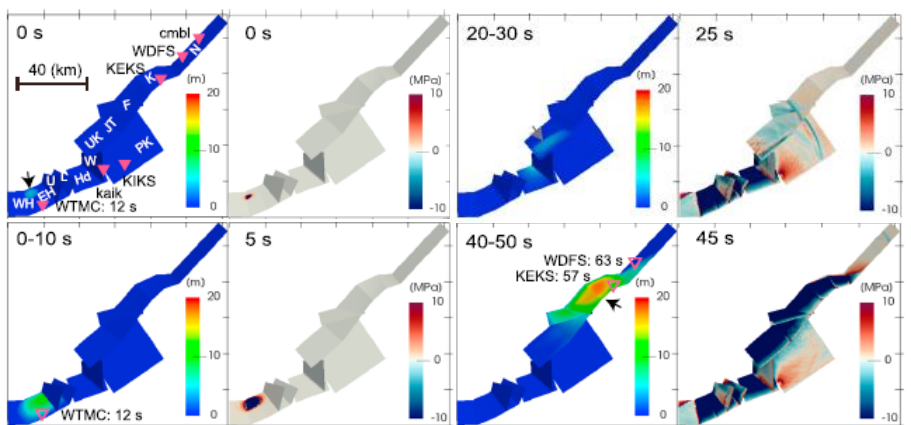
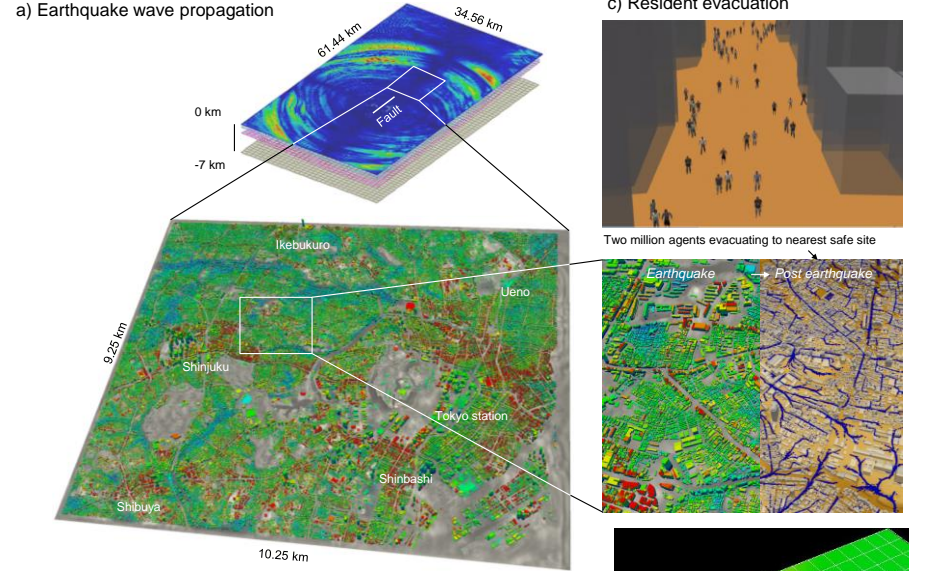
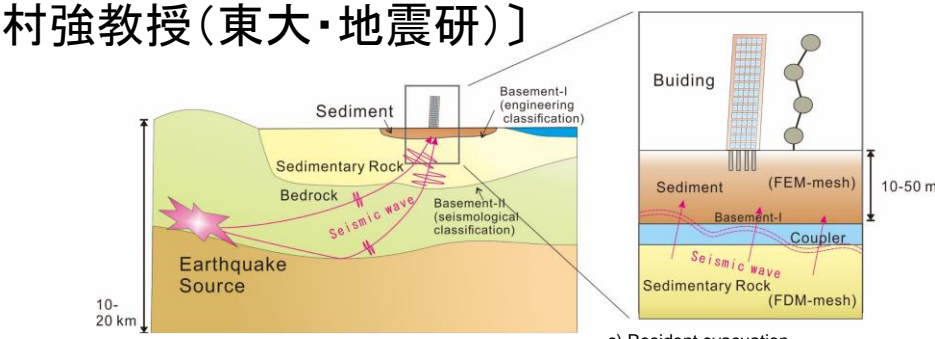
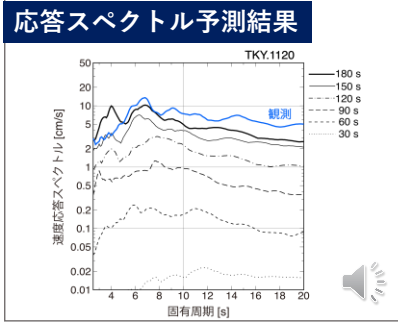
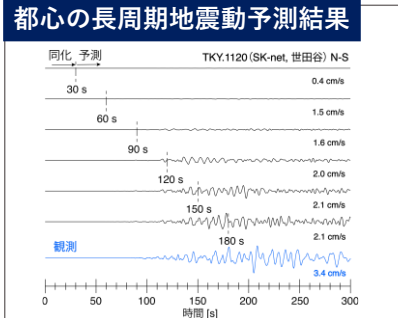
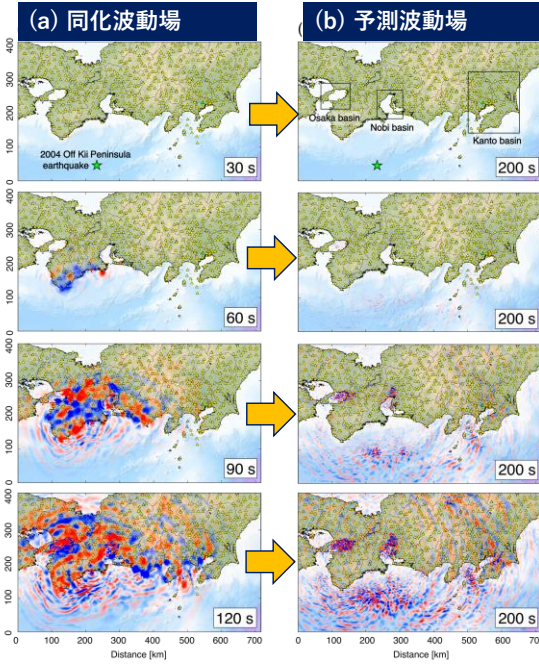
東大地震研究所, 東大理学系研究科等

[画像提供: 古村孝志教授・市村強教授(東大・地震研)]

○ 使用データ(K-NET, KiK-net 446点)



90秒間の同化→予測計算



[画像提供: 安藤亮輔准教授 (東大・理学系)]

二酸化炭素地下貯留シミュレーション

大成建設, 理化学研究所等

[画像提供:
山本肇博士(大成建設)]

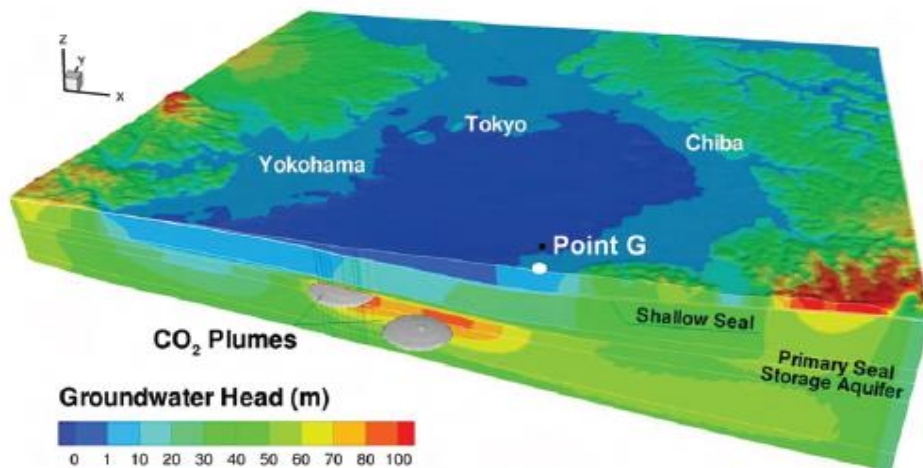
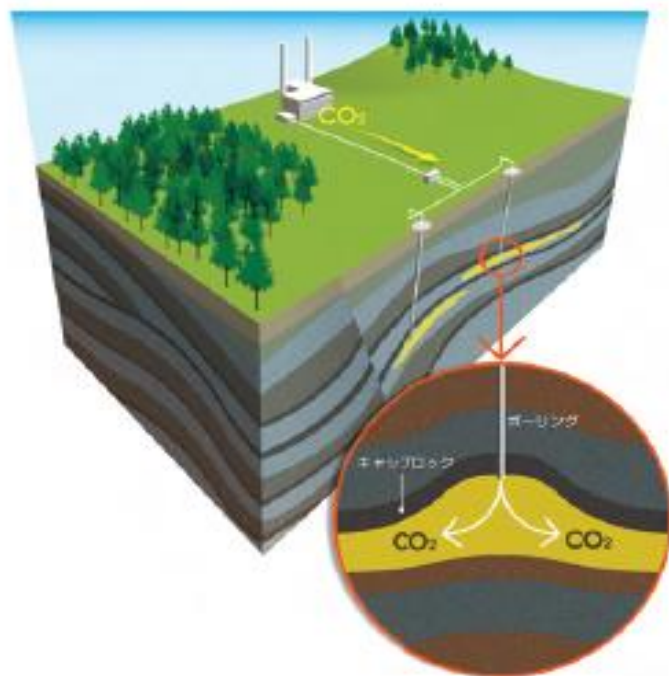
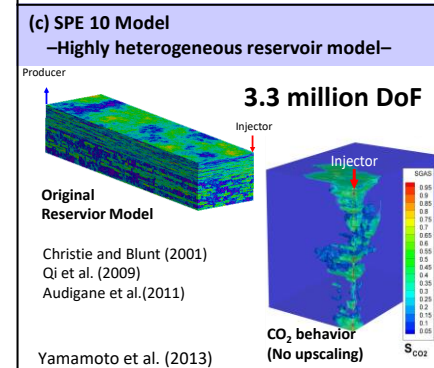
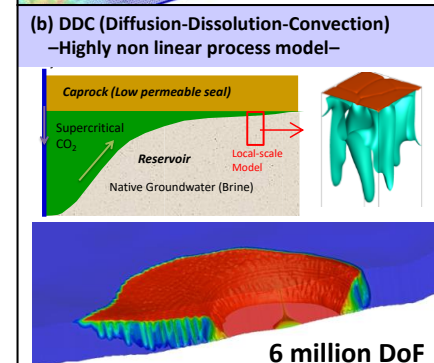
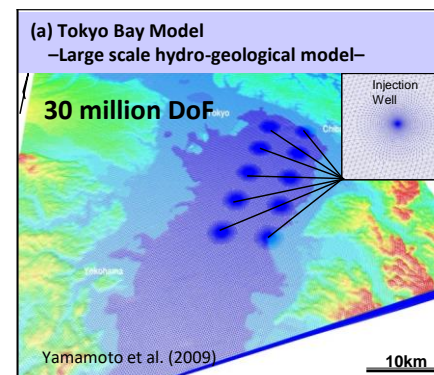
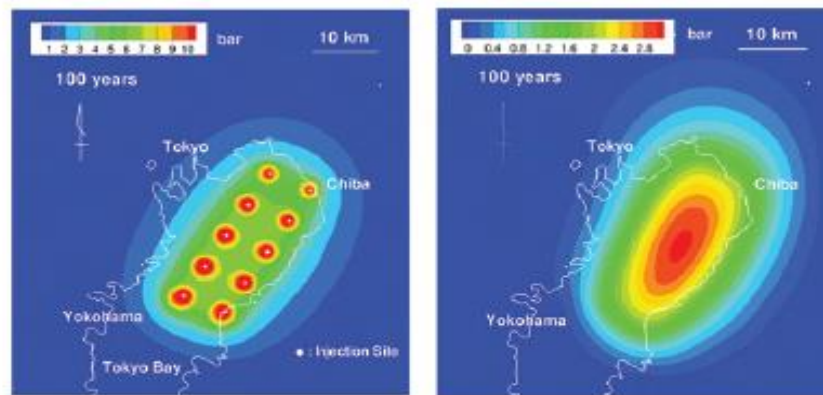


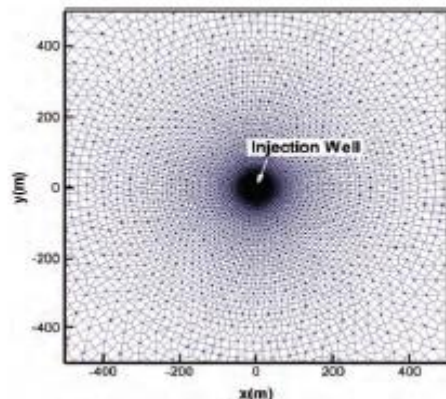
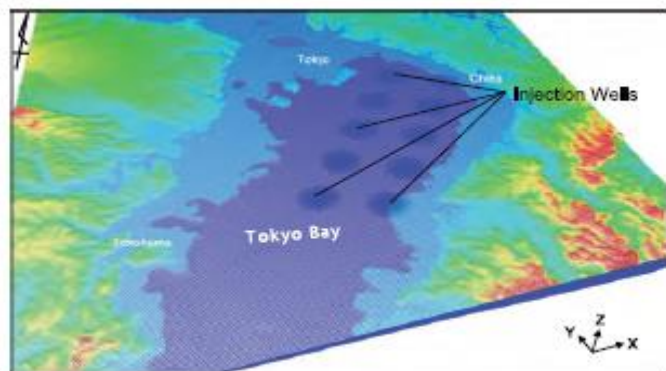
図-4 CO₂ 圧入後の地下水圧 (全水頭換算) の分布 (100 年後)



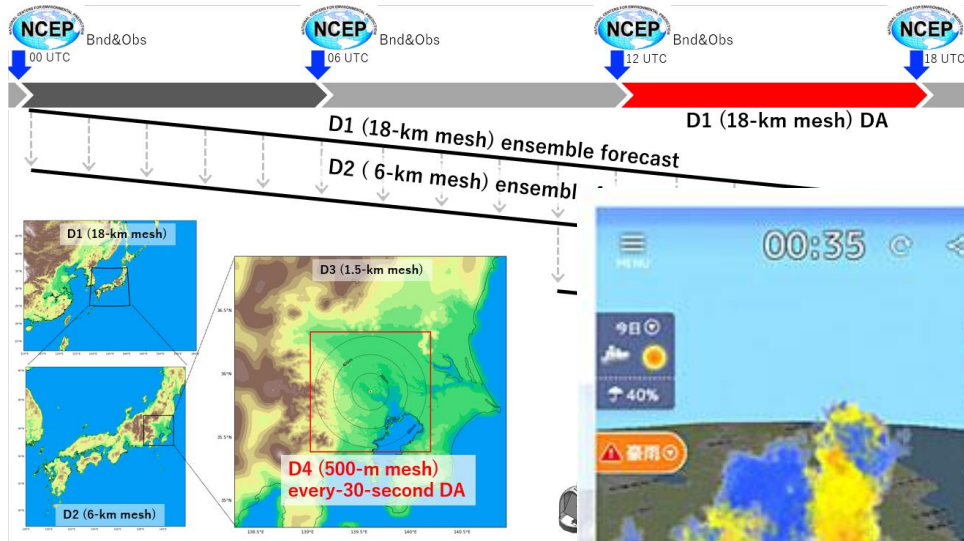
※DOF: degrees of freedom



(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)



ゲリラ豪雨予測のリアルタイム実証実験 (理化学研究所)

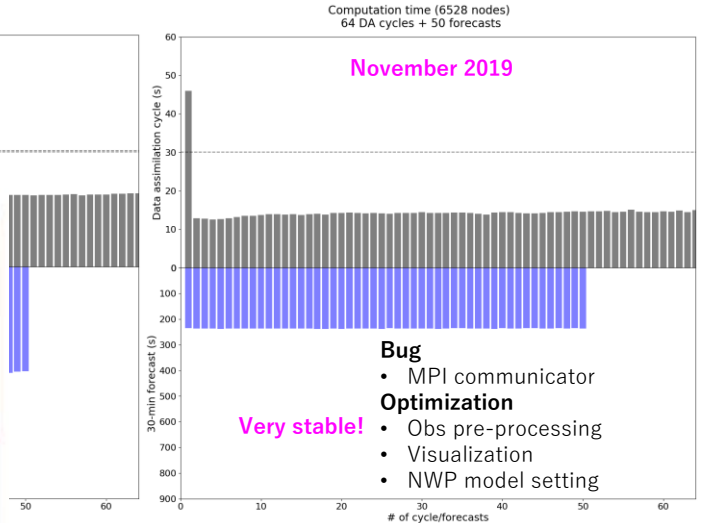
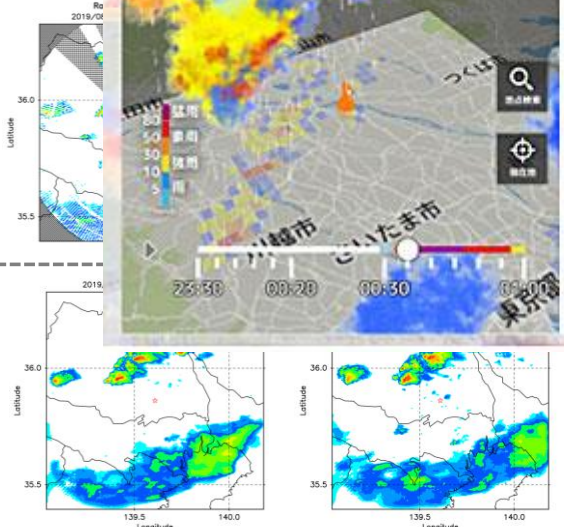


全体のワークフロー

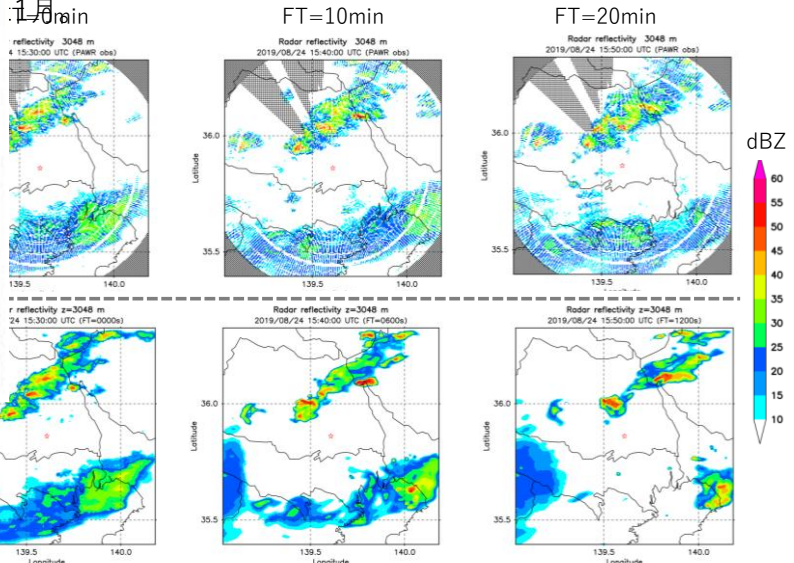


PAWR Obs

SCALE-LETKF Analysis



夕同化、下段は30分予報にかかった時間(秒)。



〔画像提供: 三好建正博士 (理化学研究所)〕

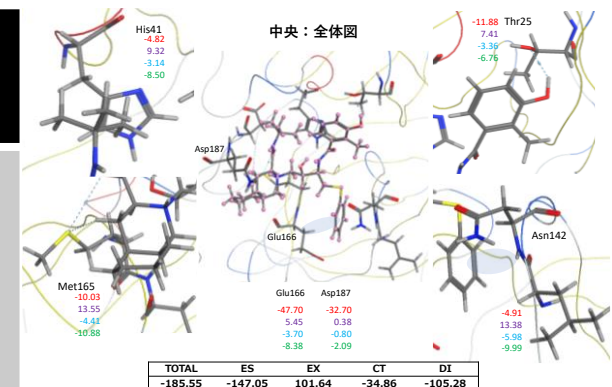
2019年8月24日の事例についてのテスト結果。(上)レーダー観測と(下)SCALE-LETKFによる解析で得られたレーダー反射強度(dBZ)を示す。

「COVID-19」対応HPCI臨時公募課題

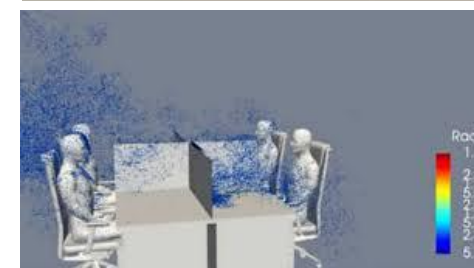
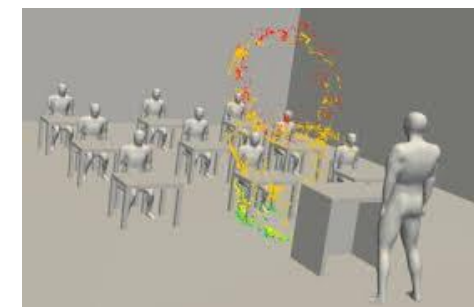
全14のうち6課題が東大システムを利用(2020年度)



課題名	代表者(所属)	使用システム
新型コロナウイルスの主要プロテアーゼに関するフラグメント分子軌道計算	望月 祐志 (立教大学)	Oakforest PACS
COVID-19治療の候補薬: chloroquine、hydroxychloroquine、azithromycinの催不整脈リスクの評価ならびにその低減策に関する研究	久田 俊明(株式会社UT-Heart研究所 / 東大)	
新型コロナウイルス表面のタンパク質動的構造予測	杉田 有治 (理化学研究所)	
計算機解析によるSARS-CoV-2増殖阻害化合物の探索	星野 忠次 (千葉大学)	Oakbridge CX
室内環境におけるウイルス飛沫感染の予測とその対策: 富岳大規模解析に向けたケーススタディ	坪倉 誠 (神戸大学)	
Spreading of polydisperse droplets in a turbulent puff of saturated exhaled air	Marco Edoardo Rosti (OIST)	



資料提供: 望月祐志教授(立教大学)



資料提供: 坪倉誠教授(神戸大学)

YouTubeチャンネルのご紹介

研究事例紹介や、セミナー・講習会の録画などをご覧になれます。

- 「東京大学情報基盤センター」チャンネル

<https://www.youtube.com/channel/UC2CHaGp1AO-vqRIV7wmU0-w>

- Wisteria/BDEC-01システム紹介

https://www.youtube.com/watch?v=SXjYtatzo-4&list=PLobjSv_ny85IW03OAPUJ9DWJoHhNiQgvY&index=3&t=104s

- 第10回JCAHPCセミナー

https://www.youtube.com/playlist?list=PLobjSv_ny85mfPTuCC2i7r_sPQYKZvy2e

- 柏キャンパス一般公開

https://www.youtube.com/playlist?list=PLobjSv_ny85kr1lg2m-bUiMC2a9W6k53u
<https://www.youtube.com/watch?v=q-0QtU7Ops4&t=116s>

- JCAHPCセミナー:「人類と地球を護るスーパーコンピューティング」

https://www.youtube.com/playlist?list=PLobjSv_ny85l-z-VJCy690ZjIAA04xCRA

- お試しアカウントつき講習会

https://www.youtube.com/playlist?list=PLobjSv_ny85kXY2Mtnhn1k7pM-epQaD2y

スーパーコンピューティング の今後

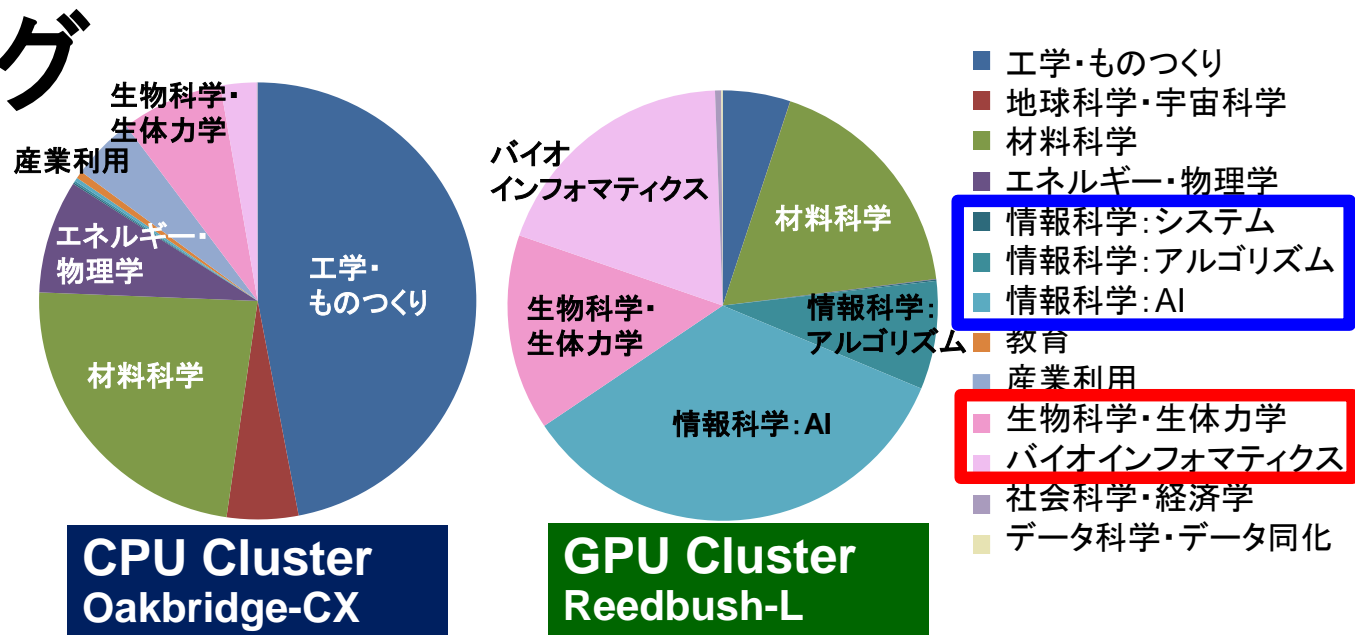
- ワークロードの多様化
 - 計算科学, 計算工学: Simulations
 - 大規模データ解析
 - AI, 機械学習

- (シミュレーション(計算) + データ + 学習) 融合 ⇒ Society 5.0 実現に有効, 2015年頃から取り組み

– フィジカル空間とサイバー空間の融合

- S: シミュレーション(計算) (Simulation)
- D: データ(Data)
- L: 学習(Learning)

– Simulation + Data + Learning =
S+D+L



(シミュレーション(計算)+データ+学習)融合(S+D+L)

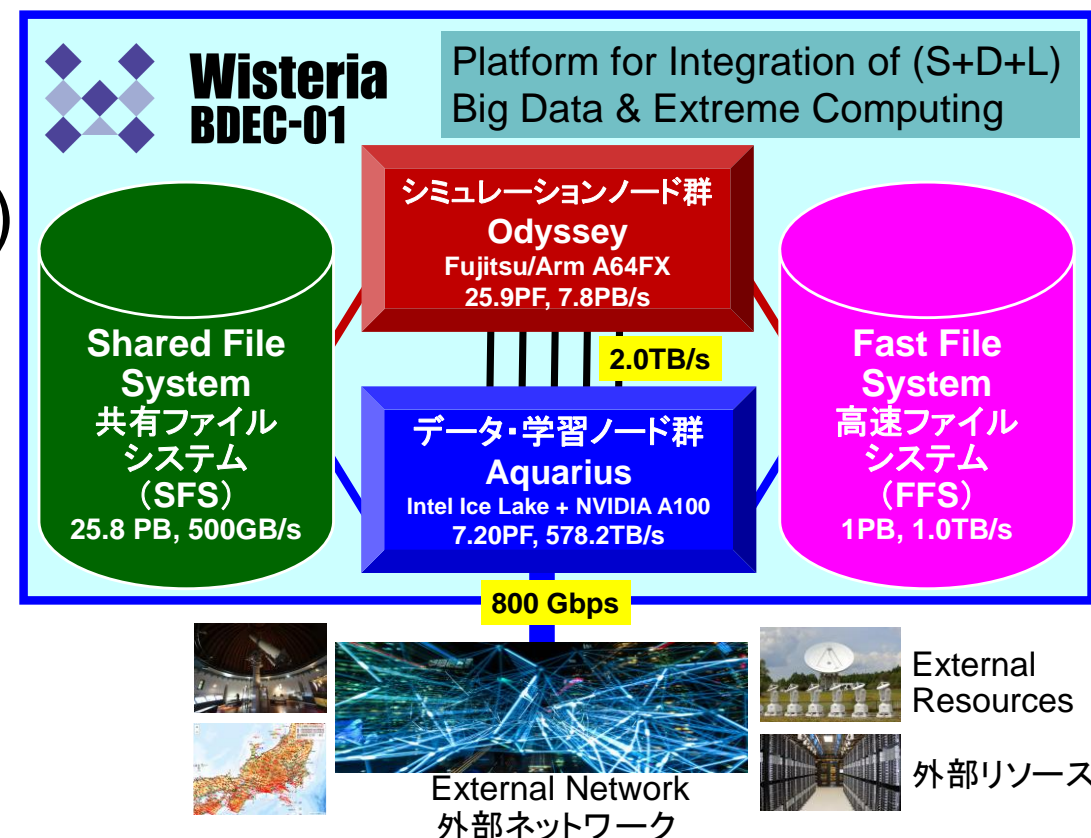
- 東大情報基盤センターでは、2015年頃から「(S+D+L)融合」の重要性に注目し、それを実現するためのハードウェア、ソフトウェア、アプリケーション、アルゴリズムに関する研究開発を開始
 - BDEC計画(Big Data & Extreme Computing)
 - 「データ+学習」による、より高度な「シミュレーション」
 - AI for HPC
 - 地球科学関連では自然な発想(すでに実施されている)
- 2021年5月に運用を開始した「Wisteria/BDEC-01」は「BDEC計画」の1号機
 - Reedbush, Oakbridge-CXは「BDEC」のプロトタイプと位置づけられる
 - 「計算・データ・学習(S+D+L)」融合を実現する、世界でも初めてのプラットフォーム



Wisteria/BDEC-01

- 2021年5月14日運用開始
 - 東京大学柏Ⅱキャンパス
- 33.1 PF, 8.38 PB/sec., **富士通製**
 - ~4.5 MVA(空調込み), ~360m²
- Hierarchical, Hybrid, Heterogeneous (h3)
- **2種類のノード群**
 - シミュレーションノード群(S, SIM) : **Odyssey**
 - 従来のスパコン
 - **Fujitsu PRIMEHPC FX1000 (A64FX), 25.9 PF**
 - 7,680ノード(368,640 コア), 20ラック, Tofu-D
 - データ・学習ノード群(D/L, DL) : **Aquarius**
 - データ解析, 機械学習
 - **Intel Xeon Ice Lake + NVIDIA A100, 7.2 PF**
 - 45ノード(Ice Lake:90基, A100:360基), IB-HDR
 - 一部は外部リソース(ストレージ, サーバー, センサーネットワーク他)に直接接続
 - ファイルシステム: 共有(大容量) + 高速

BDEC:「計算・データ・学習(S+D+L)」
融合のためのプラットフォーム
(Big Data & Extreme Computing)



Wisteria
BDEC-01

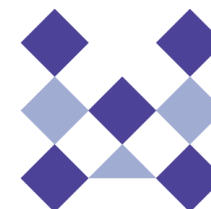
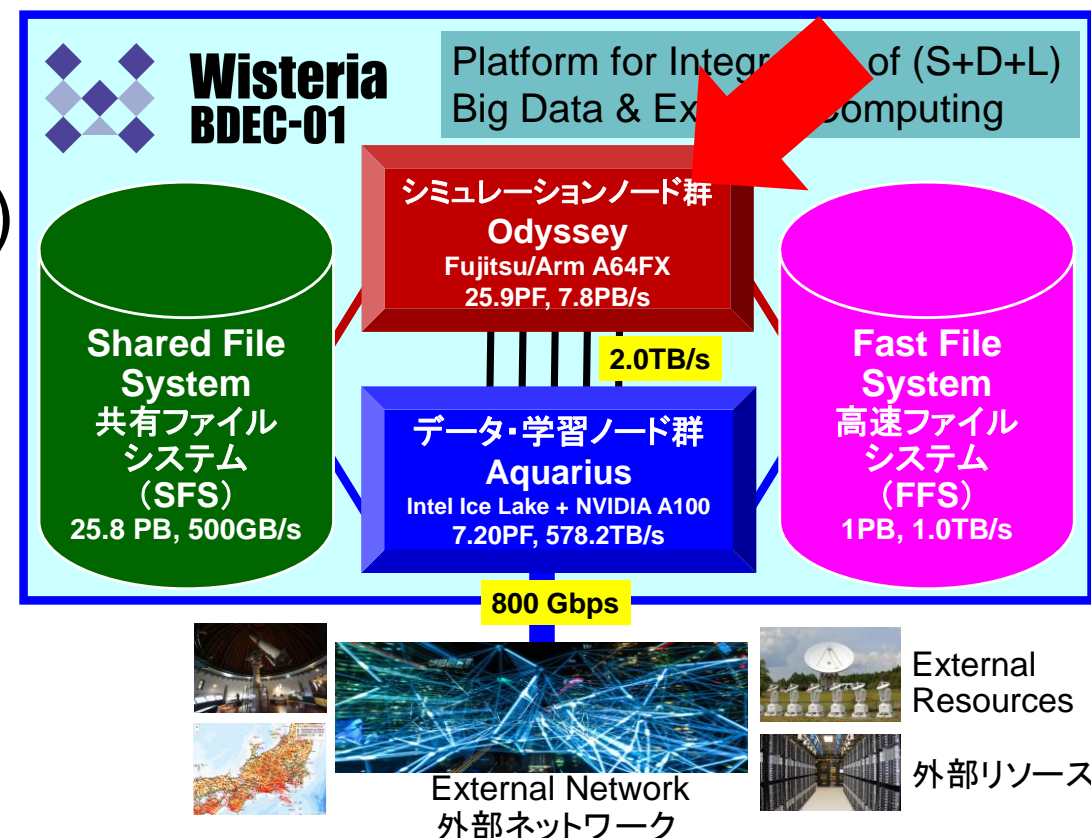
Wisteria/BDEC-01

- 2021年5月14日運用開始
 - 東京大学柏Ⅱキャンパス
- 33.1 PF, 8.38 PB/sec., **富士通製**
 - ~4.5 MVA(空調込み), ~360m²
- Hierarchical, Hybrid, Heterogeneous (h3)
- **2種類のノード群**

- **シミュレーションノード群 (S, SIM) : Odyssey**
 - 従来のスパコン
 - **Fujitsu PRIMEHPC FX1000 (A64FX), 25.9 PF**
 - 7,680ノード(368,640 コア), 20ラック, Tofu-D

- **データ・学習ノード群 (D/L, DL) : Aquarius**
 - データ解析, 機械学習
 - **Intel Xeon Ice Lake + NVIDIA A100, 7.2 PF**
 - 45ノード(Ice Lake:90基, A100:360基), IB-HDR
 - 一部は外部リソース(ストレージ, サーバー, センサーネットワーク他)に直接接続
- ファイルシステム: 共有(大容量) + 高速

BDEC:「計算・データ・学習 (S+D+L)」
融合のためのプラットフォーム
(Big Data & Extreme Computing)

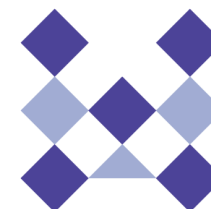
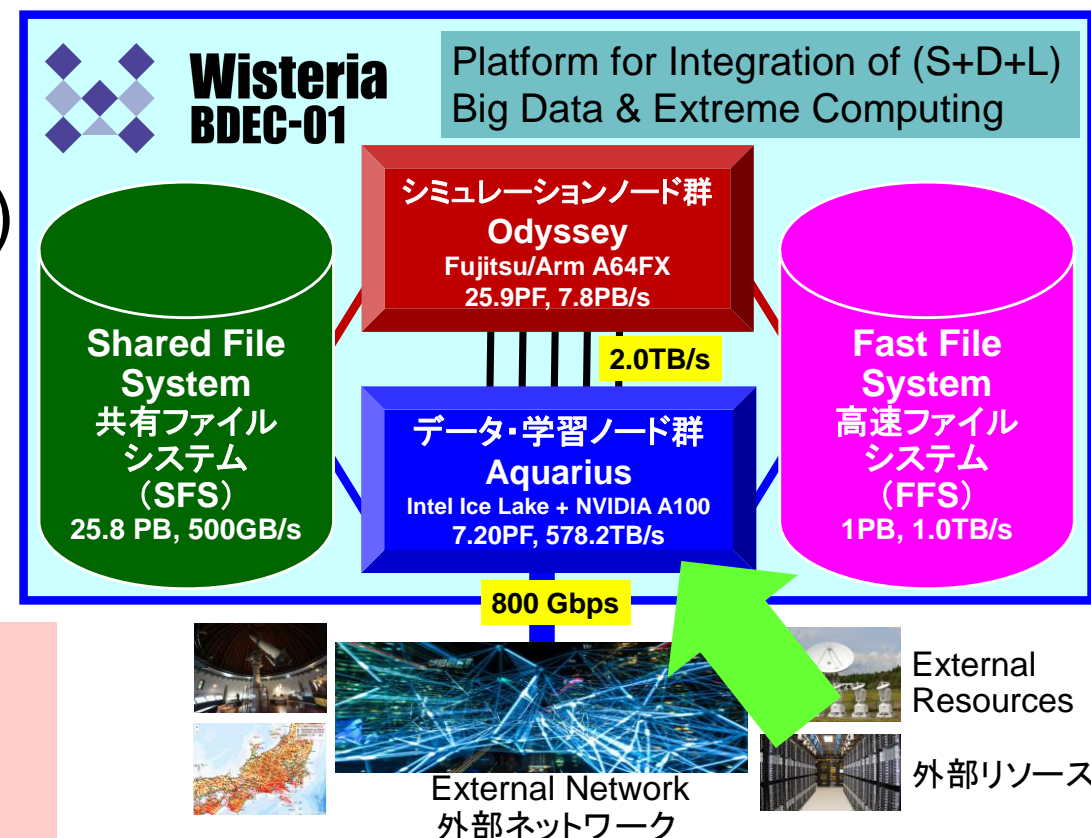


Wisteria
BDEC-01

Wisteria/BDEC-01

- 2021年5月14日運用開始
 - 東京大学柏Ⅱキャンパス
- 33.1 PF, 8.38 PB/sec., **富士通製**
 - ~4.5 MVA(空調込み), ~360m²
- Hierarchical, Hybrid, Heterogeneous (h3)
- **2種類のノード群**
 - シミュレーションノード群(S, SIM) : **Odyssey**
 - 従来のスパコン
 - **Fujitsu PRIMEHPC FX1000 (A64FX), 25.9 PF**
 - 7,680ノード(368,640コア), 20ラック, Tofu-D
 - **データ・学習ノード群(D/L, DL) : Aquarius**
 - データ解析, 機械学習
 - **Intel Xeon Ice Lake + NVIDIA A100, 7.2 PF**
 - 45ノード(Ice Lake:90基, A100:360基), IB-HDR
 - 一部は外部リソース(ストレージ, サーバー, センサーネットワーク他)に直接接続
 - ファイルシステム: 共有(大容量) + 高速

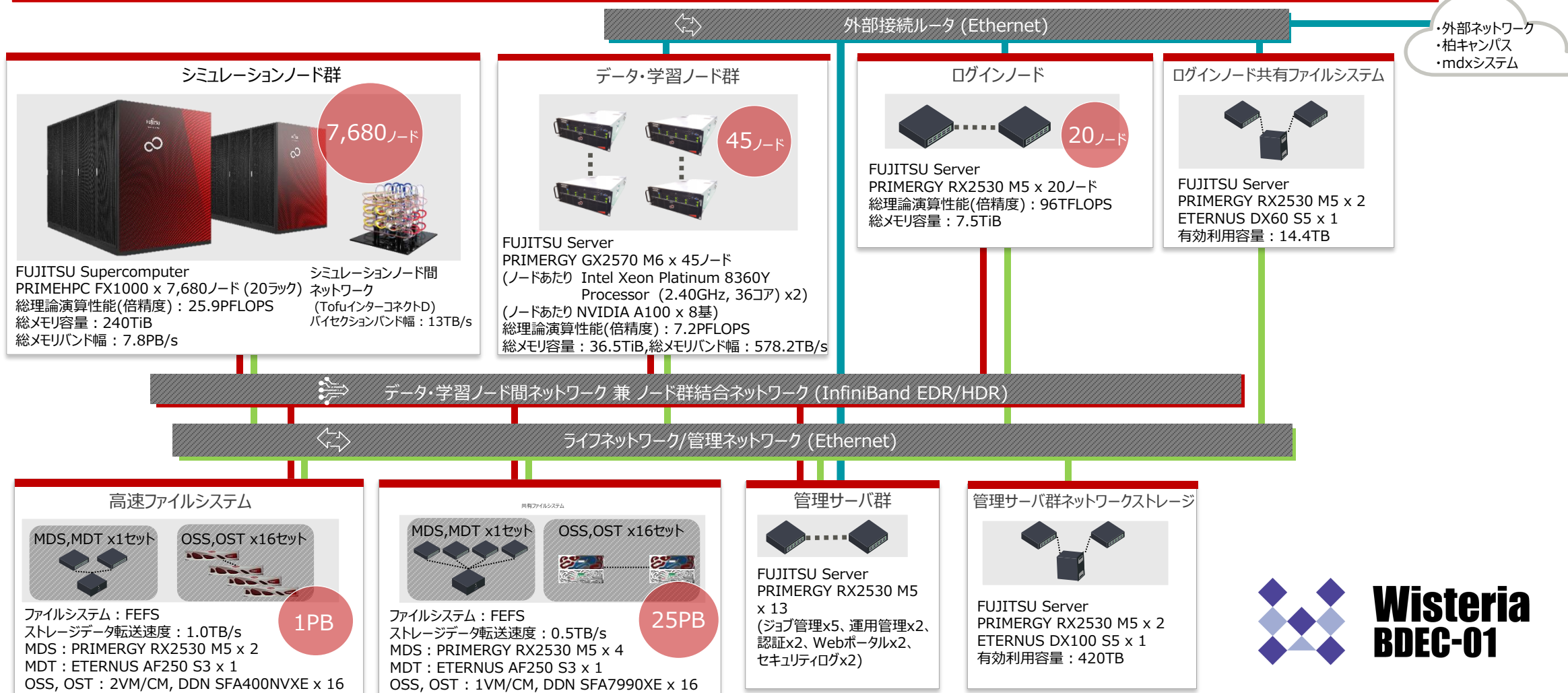
BDEC:「計算・データ・学習(S+D+L)」
融合のためのプラットフォーム
(Big Data & Extreme Computing)



**Wisteria
BDEC-01**

システム構成図

シミュレーションノード：7,680ノード (総理論演算性能 25.9 PFLOPS、総メモリバンド幅 7.8 PB/s)
 データ・学習ノード：45ノード (総理論演算性能 7.2 PFLOPS、総メモリバンド幅 578.2 TB/s)



項目		Wisteria-O (Odyssey)	Wisteria-A (Aquarius)
総理論演算性能		25.9 PFLOPS	7.2 PFLOPS
総ノード数		7,680	45
総主記憶容量		240.0 TiB	36.5 TiB
ネットワークトポロジー		6次元メッシュ / トーラス	Full-bisection Fat Tree
共有ファイルシステム	システム名	FEFS (Fujitsu Exabyte File System)	
	サーバ(OSS)	DDN SFA7990XE	
	サーバ(OSS)数	16	
	ストレージ容量	25.8 PB	
	ストレージデータ転送速度	504 GB/s	
高速ファイルシステム	システム名	FEFS (Fujitsu Exabyte File System)	
	サーバ(OSS)	DDN SFA400NVXE	
	サーバ(OSS)数	16	
	ストレージ容量	1.0 PB	
	ストレージデータ転送速度	1.0 TB/s	

項目		Wisteria-O (Odyssey)	Wisteria-A (Aquarius)
マシン名		FUJITSU Supercomputer PRIMEHPC FX1000	FUJITSU Server PRIMERGY GX2570 M6
CPU	プロセッサ名	A64FX	Intel Xeon Platinum 8360Y (開発コード名: Ice Lake)
	プロセッサ数 (コア数)	1 (48+アシスタントコア2 or 4)	2 (36+36)
	周波数	2.2 GHz	2.4 GHz
	理論演算性能	3.3792 TFLOPS	5.53 TFLOPS
	メモリ容量	32 GB	512 GiB
	メモリ帯域幅	1,024 GB/s	409.6 GB/s
GPU	プロセッサ名	-	NVIDIA A100
	SM数(単体)		108
	メモリ容量(単体)		40 GB
	メモリ帯域幅(単体)		1,555 GB/s
	理論演算性能(単体)		19.5 TFLOPS
	搭載数		8
	CPU-GPU間接続		PCI Express Gen4 x 16レーン (1レーンあたり片方向32 GB/s)
	GPU間接続		NVLink x 12本 (1本あたり片方向25GB/s)
インターコネク		Tofuインターコネク	InfiniBand HDR(200Gbps) x 4

ソフトウェア群

項目	Wisteria-O (Odyssey)	Wisteria-A (Aquarius)
OS	Red Hat Enterprise Linux 8 (aarch64)	Red Hat Enterprise Linux 8 (x86_64)
コンパイラ	GNU コンパイラ	GNU コンパイラ
	富士通社製 コンパイラ (Fortran77/90/95/2003/2008、C、C++)	Intel コンパイラ(Fortran77/90/95/2003/2008、C、C++) NVIDIA HPC SDK (Fortran77/90/95/2003/2008、C、C++、OpenACC 2.7) NVIDIA CUDA SDK (CUDA C、CUDA C++)
メッセージ通信 ライブラリ	富士通社製MPI	Intel MPI、Open MPI

項目	Wisteria-O (Odyssey)	Wisteria-A (Aquarius)
ライブラリ	SuperLU、SuperLU MT、SuperLU DIST、METIS、MT-METIS、ParMETIS、Scotch、PT-Scotch、PETSc、Trillinos、FFTW、GNU Scientific Library、NetCDF、Parallel netCDF、HDF5、Parallel HDF5、CMake、Miniconda、Xabclib、ppOpen-HPC、MassiveThreads、Boost C++、mpiJava	
	富士通社製ライブラリ(BLAS、CBLAS、LAPACK、ScaLAPACK)	Intel社製ライブラリ(MKL)(BLAS、CBLAS、LAPACK、ScaLAPACK)、NVIDIA社製ライブラリ(cuBLAS、cuSPARSE、cuFFT、cuDNN、NCCL)、MAGMA
アプリケーション	OpenFOAM、ABINIT-MP、PHASE、FrontFlow/blue、FrontISTR、REVOCAP-Coupler、REVOCAP-Refiner、OpenMX、MODYLAS、GROMACS、BLAST、R packages、bioconductor、BioPerl、BioRuby、BWA、GATK、SAMtools、Quantum ESPRESSO、Xcrypt、ROOT、Geant4、LAMMPS、CP2K、NWChem、DeepVariant、Paraview、VisIt、POV-Ray、TensorFlow、Chainer、PyTorch、Keras、Horovod、MXNet	
		Theano
フリーソフトウェア	autoconf、automake、bash、bzip2、cvs、emacs、findutils、gawk、gdb、make、grep、gnuplot、gzip、less、m4、python、perl、ruby、screen、sed、subversion、tar、tcsh、tcl、vim、zsh、git など	
		Globus Toolkit、Gfarm、FUSE
コンテナ仮想化	Singularity Community Edition	
商用アプリ		MATLAB (New!)

Jun. 2023 (ISC23)の諸ランキング

Wisteria/BDEC-01のシミュレーションノード群 (Odyssey) とデータ・学習ノード群 (Aquarius) は別々に測定・申請

System	TOP500	Green500	HPCG	Graph500	HPL-MxP (HPL-AI)
	連立一次方程式 (密行列)	TOP500:消費電力 当たり計算性能	連立一次方程式 (疎行列)	グラフ処理	連立一次方程式 (密行列・混合精度)
Oakbridge-CX	140	92	85	-	-
Wisteria/BDEC-01 (Odyssey)	25	50	12	6	14
Wisteria/BDEC-01 (Aquarius)	136	32	68	-	-

61st TOP500 List (Jun., 2023)

R_{max} : Performance of Linpack (TFLOPS)

<http://www.top500.org/>

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

	Site	Computer/Year Vendor	Cores	R_{max} (PFLOPS)	R_{peak} (PFLOPS)	Power (kW)
1	Frontier, 2022, USA DOE/SC/Oak Ridge National Laboratory	HPE Cray EX235a, AMD Optimized 3 rd Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	8,730,112	1,102.00 (=1.102 EF)	1,685.65	21,100
2	Fugaku, 2020, Japan R-CCS, RIKEN	Fujitsu PRIMEHPC FX1000, Fujitsu A64FX 48C 2.2GHz, Tofu-D	7,630,848	442,010 (= 442.0 PF)	537,212.0	29,899
3	LUMI, 2022, Finland EuroHPC/CSC	HPE Cray EX235a, AMD Optimized 3 rd Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	2,220,288	309.10	428.70	6,016
4	Leonard, 2022, Italy EuroHPC/Cineca	BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64GB, Quad-rail NVIDIA HDR100	1,824,768	238.70	304.46	7,404
5	Summit, 2018, USA DOE/SC/Oak Ridge National Laboratory	IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR InfiniBand	2,414,592	148.60	200.79	10,096
6	Sierra, 2018, USA DOE/NNSA/LLNL	IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR InfiniBand	1,572,480	94.64	125.71	7,438
7	Sunway TaihuLight, 2016, China National Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	10,649,600	93.01	125.44	15,371
8	Perlmutter, 2021, USA DOE/NERSC/LBNL	HPE Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-10	761,856	70.87	93.75	2,528
9	Selene, 2020, USA NVIDIA	NVIDIA DGX A100 SuperPOD, AMD EPYC 7742 64C 2.25GHz, NVIDIA GA100, Mellanox Infiniband HDR	555,520	63.46	79.22	2,646
10	Tianhe-2A, 2018, China National Super Computer Center in Guangzhou	TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000	4,981,760	61.44	100.68	18,482
24	ABCI 2.0, 2021, Japan AIST	PRIMERGY GX2570 M6, Xeon Platinum 8360Y 36C 2.4GHz, NVIDIA A100 SXM4 40 GB, InfiniBand HDR	504,000	22.21	54.34	1,600
25	Wisteria/BDEC-01 (Odyssey), 2021, Japan ITC, University of Tokyo	PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu interconnect D	368,640	22.12	25.95	1,468

Green 500 Ranking (Jun., 2023)

	TOP 500 Rank	System	Accelerator	Cores	HPL Rmax (Pflop/s)	Power (kW)	GFLOPS/W
1	255	Henri, Flatiron Institute, USA	NVIDIA H100 80GB	8,288	2.88	44	65.396
2	34	Frontier TDS, ORNL, USA	AMD Instinct MI250X	120,832	19.20	309	62.684
3	12	Adastra, GENCI-CINES, France	AMD Instinct MI250X	319,072	46.10	921	58.021
4	17	Setonix-GPU, Paswey SCC, Australia	AMD Instinct MI250X	181,248	27.16	477	56.983
5	77	Dardel GPU, KTH, Sweden	AMD Instinct MI250X	52,864	8.26	146	56.491
6	1	Frontier, ORNL, USA	AMD Instinct MI250X	8,699,904	1,194.00	22,703	52.592
7	3	LUMI, EuroHPC/CSC, Finland	AMD Instinct MI250X	2,220,288	309.10	6,016	51.382
8	483	amplitUDE (GPU Partition), University of Duisburg-Essen, Germany	NVIDIA H100 80GB	6,768	1.95	38	51.343
9	70	Goete-NHR, Universitaet Frankfurt, Germany	AMD Instinct MI210	96,768	9.09	195	46.543
10	187	ATOS THX.A.B, Atos, France	NVIDIA A100 SXM 64GB	25,056	3.50	86	41.411
32	136	Wisteria/BDEC-01 (Aquarius), Fujitsu, Japan	NVIDIA A100 40GB	42,120	4.425	183	24.06

HPCG Ranking (Jun., 2023)

	Computer	Cores	HPL Rmax (Pflop/s)	TOP500 Rank	HPCG (Pflop/s)
1	Fugaku	7,630,848	442.01	2	16.004
2	Frontier	8,730,112	1,102.00	1	14.054
3	LUMI	2,220,288	309.10	3	3.408
4	Leonard	1,824,768	238.70	4	3.114
5	Summit	2,414,592	148.600	5	2.926
6	Perlmutter	761,856	70.870	8	1.905
7	Sierra	1,572,480	94.640	6	1.796
8	Selene	555,520	63.460	9	1.622
9	JUWELS Booster Module	449,280	44.120	12	1.275
10	Dammam-7	672,520	22.400	21	0.881
12	Wisteria/BDEC-01 (Odyssey)	368,640	22.121	25	0.817

Graph500 BFS, Jun. 2023

	Site	Computer/ Vendor	Cores	Scale	GTEPS
1	<u>Fugaku, 2020, Japan</u> R-CCS, RIKEN	Fujitsu PRIMEHPC FX1000, Fujitsu A64FX 48C 2.2GHz, Tofu-D	7,299,072	42	137096
2	<u>Pengcheng Cloudbrain-II, 2022, China</u> HUST-Pengcheng Lab-HUAWEI	HUST, Kunpeng 920+Tesla A100, Custom	6,999,552	40	121804.3
3	<u>Frontier, 2022,USA</u> DOE/SC/ORNL	HPE, Cray EX235a, Slingshot-11	8,730,112	40	29654.6
4	<u>Pengcheng Cloudbrain-II, 2022, China</u> HUST-Pengcheng Lab-HUAWEI	Huawei Atlas 900 AI Cluster	93,696	40	25242.9
5	<u>Sunway TaihuLight, 2016, China</u> National Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	10,599,680	40	23755.7
6	<u>Wisteria/BDEC-01 (Odyssey), 2021, Japan</u> ITC, University of Tokyo	Fujitsu PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu interconnect D	368,640	37	16118
7	<u>Toki-Sora, 2021, Japan</u> JAXA	PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu interconnect D	276,480	36	10813
8	<u>NAPS-FX1000, 2023, Japan</u> JMA	PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu interconnect D	221,184	36	10158
9	<u>LUMI-C, 2021, Finland</u> EuroHPC/CSC	Cray EX, SlingShot-10	190,976	38	8467.7
10	<u>Summit (CPU Only), 2018, USA</u> DOE/SC/Oak Ridge National Laboratory	IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR InfiniBand	2,414,592	40	7665.7

HPL-MxP (Jun. 2023)低精度演算利用による性能向上

R_{\max} : Performance of Linpack (TFLOPS)

	Site	Computer/Vendor	Cores	HPL-AI (EFLOPS)	Top500	HPL R_{\max} (EFLOPS)	Speedup
1	<u>Frontier, 2022, USA</u> DOE/ORNL	HPE Cray EX235a, AMD Optimized 3 rd Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	8,699,904	9.950	1	1.194	8.3
2	<u>LUMI, 2022, Finland</u> EuroHPC/CSC	HPE Cray EX235a, AMD Optimized 3 rd Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	2,174,976	2.168	3	0.3091	7.0
3	<u>Fugaku, 2020, Japan</u> R-CCS, RIKEN	Fujitsu PRIMEHPC FX1000, Fujitsu A64FX 48C 2.2GHz, Tofu-D	7,630,848	2.0	2	0.4420	4.5
4	<u>Leonard, 2022, Italy</u> EuroHPC/Cineca	BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64GB, Quad-rail NVIDIA HDR100	1,824,768	1.842	4	0.2387	7.7
5	<u>Summit, 2018, USA</u> DOE/SC/Oak Ridge National Laboratory	IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR InfiniBand	2,414,592	1.411	5	0.1486	9.5
6	<u>Selene, 2020, USA</u> NVIDIA	NVIDIA DGX A100 SuperPOD, AMD EPYC 7742 64C 2.25GHz, NVIDIA GA100, Mellanox Infiniband HDR	555,520	0.630	9	0.0630	9.9
7	<u>Perlmutter, 2021, USA</u> DOE/NERSC/LBNL	HPE Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-10	761,856	0.590	8	0.0709	8.3
8	<u>JUWELS Booster Module, 2020, Germany</u> Julich (FZJ)	Bull Sequana XH2000, AMD EPYC 7402 24c 2.8GHz, NVIDIA A100, Mellanox InfiniBand HDR	449,280	0.47	12	0.0440	10
9	<u>Adastral, 2022, France</u> GENCI-CINES	HPE Cray EX235a, AMD Optimized 3 rd Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	319,072	0.303	11	0.0461	6.6
10	<u>Setonix-GPU, 2022, Austraria</u> Pawsey Supercomputing Centre	HPE Cray EX235a, AMD Optimized 3 rd Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	181,248	0.175	15	0.0272	6.4
14	<u>Wisteria/BDEC-01 (Odyssey), 2021, Japan</u> ITC, University of Tokyo	Fujitsu PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu interconnect D	368,640	0.100	23	0.0220	4.5

利用制度の紹介

- 一般利用
 - 大学・公共機関に在籍の方（大学院生は代表者としては申し込めません）
 - 電気代相当料金の利用負担金支払いが必要
- 企業利用
 - 企業に在籍の方
 - 利用負担金は一般利用の約1.2倍
 - 書面・ヒアリング審査あり，成果報告（公開）義務あり
- 若手・女性利用
 - 大学・公共機関に在籍の方
 - 4月1日現在40歳以下の若手，または女性，または学生
 - 利用負担金なし
 - 書類審査あり，成果報告義務あり
- 学際大規模情報基盤共同利用・共同研究拠点（JHPCN）への課題申請
- HPCI課題への課題申請

一般利用コース

<https://www.cc.u-tokyo.ac.jp/supercomputer/wisteria/service/course.php>

- グループコース(1人またはそれ以上から構成されるグループ)

- Wisteria/BDEC-01から「パーソナルコース」は廃止
- 代表者は大学・公共機関所属者



- トークン(ノード時間)を購入

- Odyssey, Aquariusを利用できる
- O/Aでそれぞれ消費係数が異なる

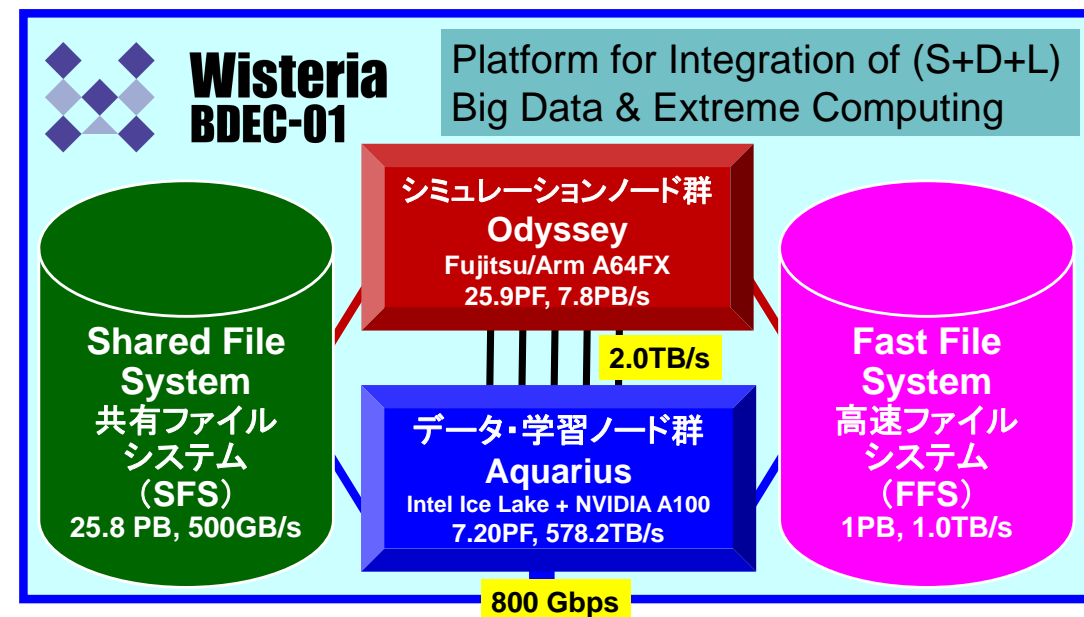
- 一般利用

- ノード固定

- Aquariusの1ノード(8GPU)を占有して利用
- 審査有り(ヒアリング)

- GPU占有

- Aquariusの1・2・4GPUを占有して利用
- 審査無し



External Network
外部ネットワーク



External Resources

外部リソース

ジョブクラス

<https://www.cc.u-tokyo.ac.jp/supercomputer/wisteria/service/job.php>

- インタラクティブ
- バッチジョブ
- プリポスト
- 現状ではOdyssey, Aquariusは同時利用はできない

- Wisteria-O (Odyssey): シミュレーションノード群
 - XXX-o
 - priority-o: 優先キュー, トークン消費量1.5倍
- Wisteria-A (Aquarius): データ・学習ノード群
 - XXX-a ノード単位
 - share-XXX GPU単位
 - MIG (Multi-Instance GPU)により, GPU内を更に分割可能だが, 本システムでは採用せず

スーパーコンピュータシステムの詳細

- <https://www.cc.u-tokyo.ac.jp/guide> をご覧ください
 - 利用申請方法
 - 運営体系
 - 料金体系
 - 利用の手引き
- 利用制度説明会（10/8）の資料
 - <https://www.cc.u-tokyo.ac.jp/events/seminar/20211008.php>

- 東大情報基盤センターのスパコン概要
- Wisteria/BDEC-01
- **Ipomoea-01**
- h3-Open-BDEC

大規模共通ストレージシステム「Ipomoea」

- スーパーコンピュータの処理能力の向上に伴い、扱うデータ量も増加の一途
- 東大センターでは従来ストレージは各システムに附属して導入され、各システムのストレージは独立
- このような状況（注：ストレージがシステム毎に独立）は利用者に多大な不便を強いることになり、東大センターの全システムからアクセス可能な共通ストレージの導入が強く求められていた
- 各システムからアクセスできる「大規模共通ストレージ（Ipomoea）」導入決定
 - OFP運用終了が契機
 - 1システムを約5–6年使用し、約3年ごとに新しいストレージシステム（25+PB）を導入し、入れ替えることを想定している



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

**東京大学情報基盤
センターのスパコン**
利用者2,600+名
55%は学外

Reedbush-U/H/L (SGI-HPE)
3.36 PF

Ipomoea-01 25PB

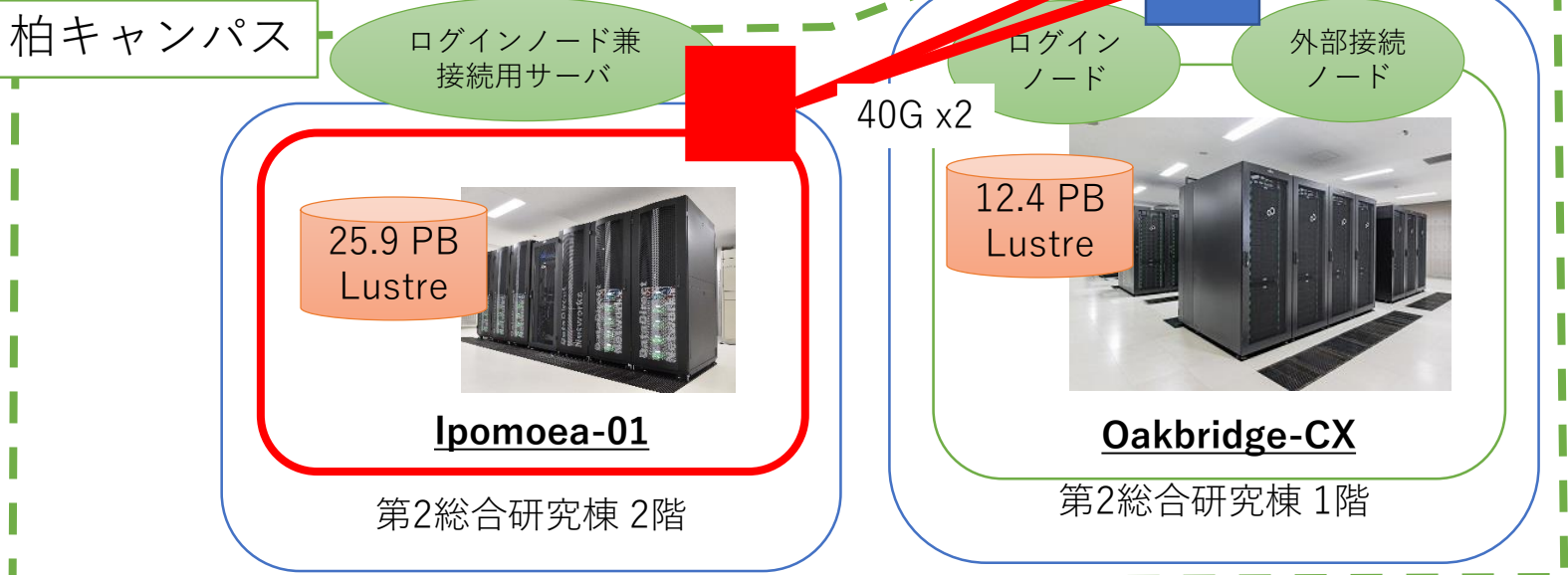
Ipomoea-03

Ipomoea-02

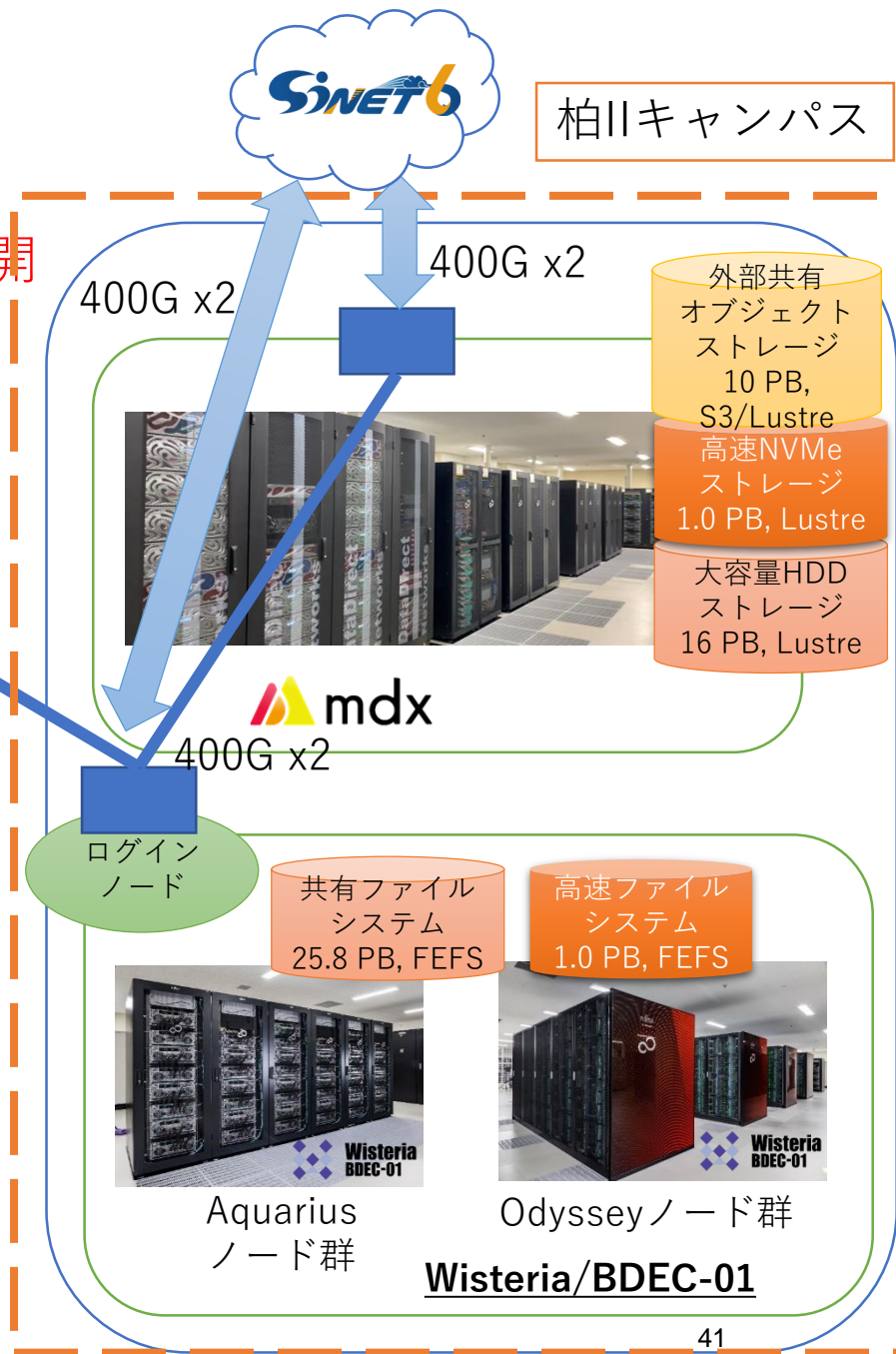
Ipomoea-01

- 2022年1月運用開始, 25+PB
 - 富士通製
- OFPのLustre領域の必要ファイルの移行完了, 2022年6月から公開
- 割当容量
 - 東大センターのシステムのいずれかに利用者番号(教育利用, 講習会除く)を有する場合
 - 各利用者ごとに5TB
 - 各グループごとに登録システムで付与されている容量の15%を無償で付与
 - 追加負担金(企業はこの2割増し)
 - 7,200円/TB/年, 2,100,000円/PB/年
 - Ipomoea-01のみの利用申し込みも可能

柏キャンパス



柏IIキャンパス



- 東大情報基盤センターのスパコン概要
- Wisteria/BDEC-01
- Ipomoea-01
- **h3-Open-BDEC**



(計算+データ+学習)融合によるエクサスケール時代の革新的シミュレーション手法

- エクサスケール(富岳+クラス)のスパコンによる科学的発見の持続的促進のため、計算科学にデータ科学、機械学習のアイデアを導入した(計算+データ+学習(S+D+L))融合による革新的シミュレーション手法を提案
 - (計算+データ+学習)融合によるエクサスケール時代の革新的シミュレーション手法(科研費基盤S, 代表: 中島研吾(東大情基セ), 2019年度~2023年度)
- 革新的ソフトウェア基盤「h3-Open-BDEC」の開発: 東大BDECシステム(Wisteria/BDEC-01), 「富岳」等を「S+D+L」融合プラットフォームと位置づけ、スパコンの能力を最大限引き出し、最小の計算量・消費電力での計算実行を実現するために、下記2項目を中心に研究
 - 変動精度演算・精度保証・自動チューニングによる新計算原理に基づく革新的数値解法
 - 階層型データ駆動アプローチ(hDDA: Hierarchical Data Driven Approach)等に基づく革新的機械学習手法
 - Hierarchical, Hybrid, Heterogeneous \Rightarrow h3



**Wisteria
BDEC-01**



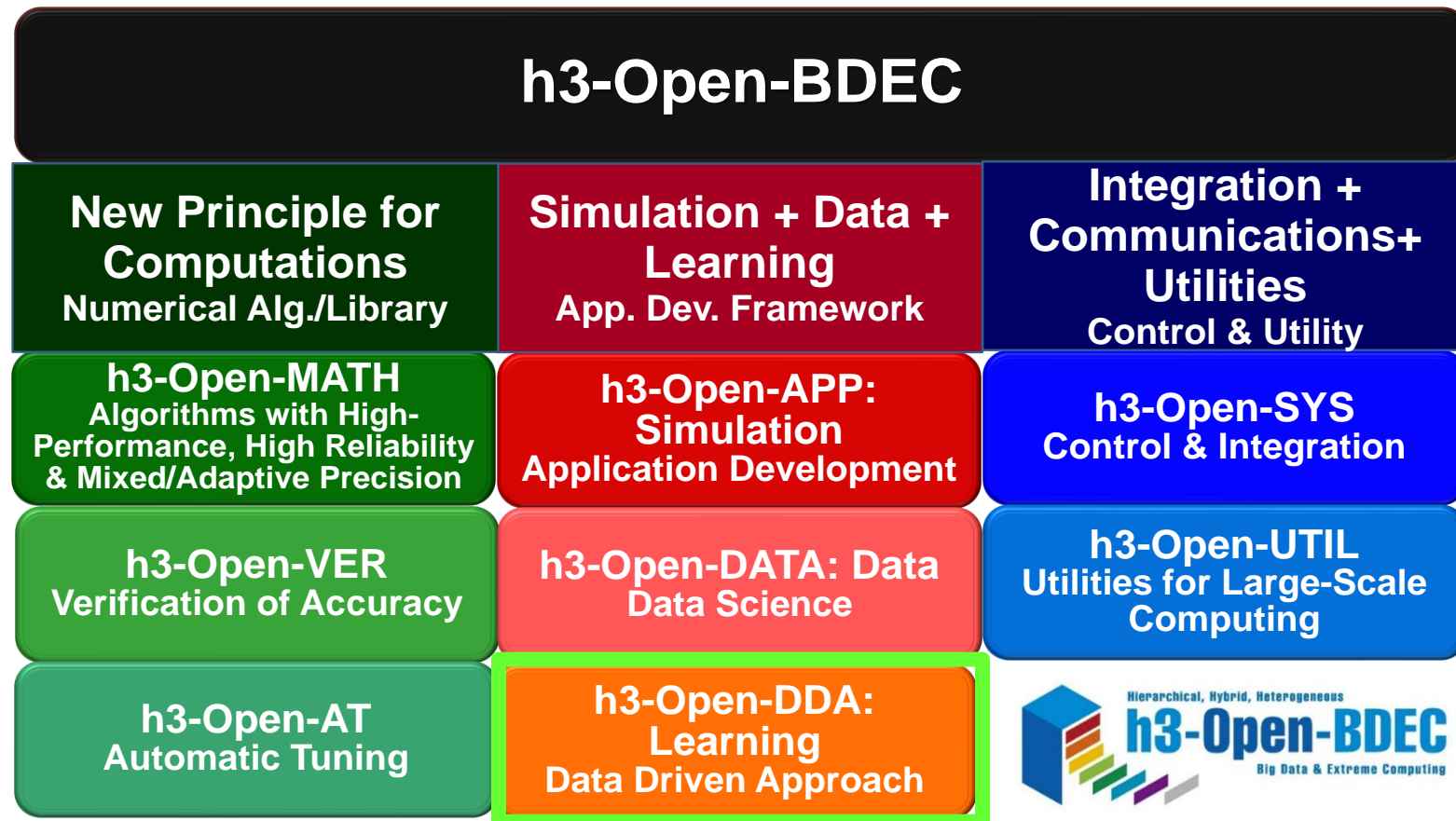
h3-Open-BDEC

「計算＋データ＋学習」融合を実現する革新的ソフトウェア基盤

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

- ① 変動精度演算・精度保証・自動チューニングによる新計算原理に基づく革新的数値解法
- ② 階層型データ駆動アプローチ (hDDA: Hierarchical Data Driven Approach) 等に基づく革新的機械学習手法

✓ Hierarchical, Hybrid, Heterogeneous ⇒ h3



期待される成果と意義

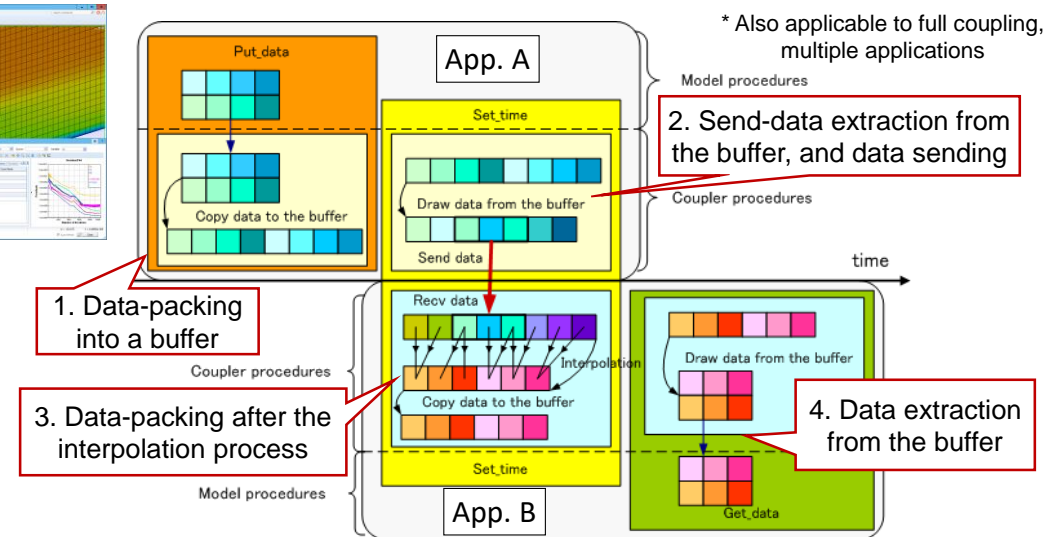
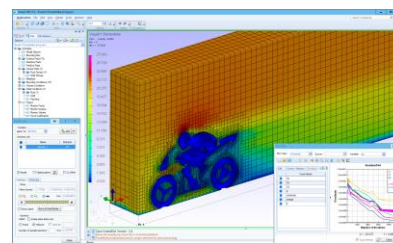
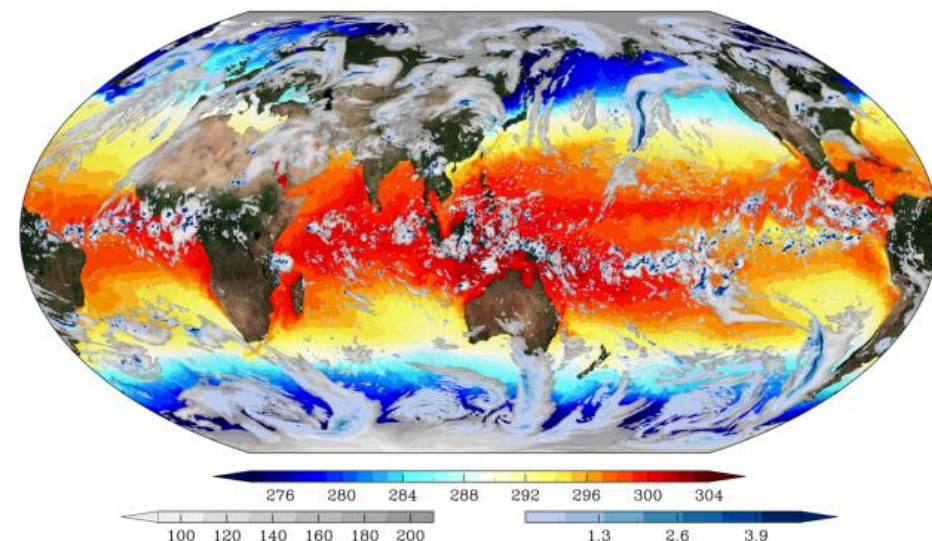
<http://nkl.cc.u-tokyo.ac.jp/h3-Open-BDEC/>



- 計算科学の専門家のみで(S+D+L)融合を容易に実現
 - 機械学習の専門家のサポートを必要としない
- ソースコード, マニュアル類も含めて一般に公開, 様々なエクサスケールシステムでの普及を目指す
 - ポスト富岳も含めたポストムーア時代への展開
- h3-Open-BDEC利用による(S+D+L)融合シミュレーションにより従来手法と同等の正確さを保ちつつ, 大幅な計算量・消費電力削減を目指す(10分の1が目標)。
- シミュレーション高度化: パラメータスタディのケース数を削減できる
- リアルタイム災害シミュレーション等への適用

Wisteria/BDEC-01上における h3-Open-BDECを使用した(S+D+L)融合

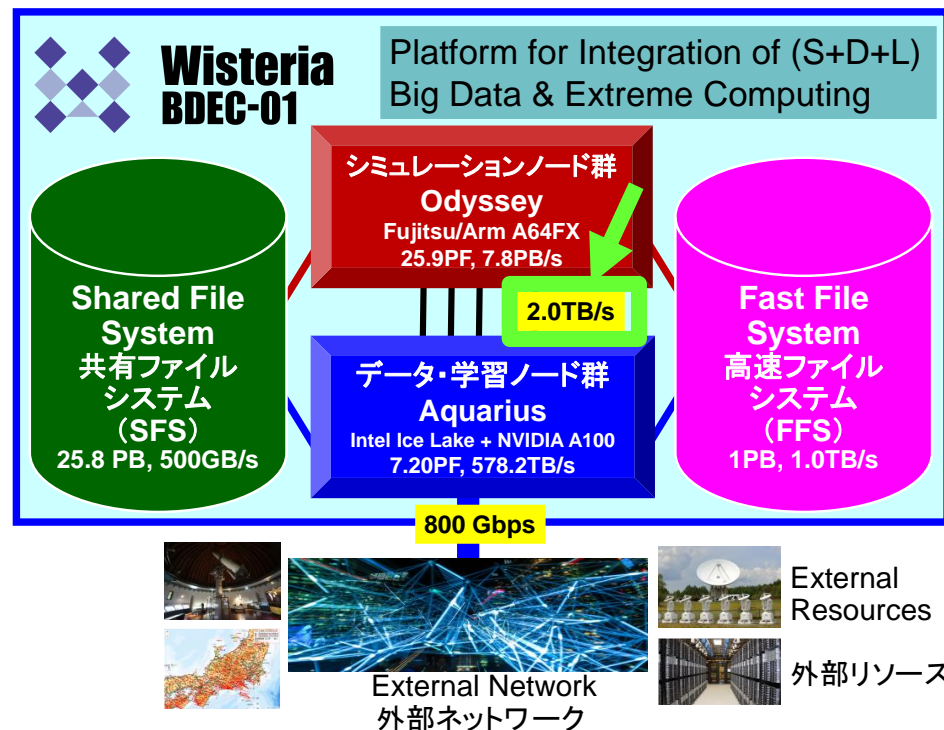
- シミュレーションとデータ同化の融合
 - 典型的・伝統的な(S+D+L)融合
- 気候・気象のための大気海洋連成シミュレーション
 - 東大大気海洋研, 理研, 国立環境研他
- **リアルタイム同化+三次元強震動シミュレーション**
 - **東大地震研**
- リアルタイム災害シミュレーション
 - 洪水, 津波
- 既存シミュレーションコードの(S+D+L)融合による高度化
 - OpenFOAM




AI for HPC の実現



- Odyssey-Aquarius連携
 - MPIによる通信は不可
 - O-Aを跨いでMPIプログラムは動かない
 - Odyssey-Aquarius間はInfiniband-EDR (2TB/sec)で結合されている
- ソフトウェア開発
 - O-A間通信: h3-Open-SYS/WaitIO
 - IB-EDR経由 (WaitIO-Socket)
 - 高速ファイルシステム (FFS) 経由連携 (WaitIO-File)
 - 高機能カプラー: h3-Open-UTIL/MP



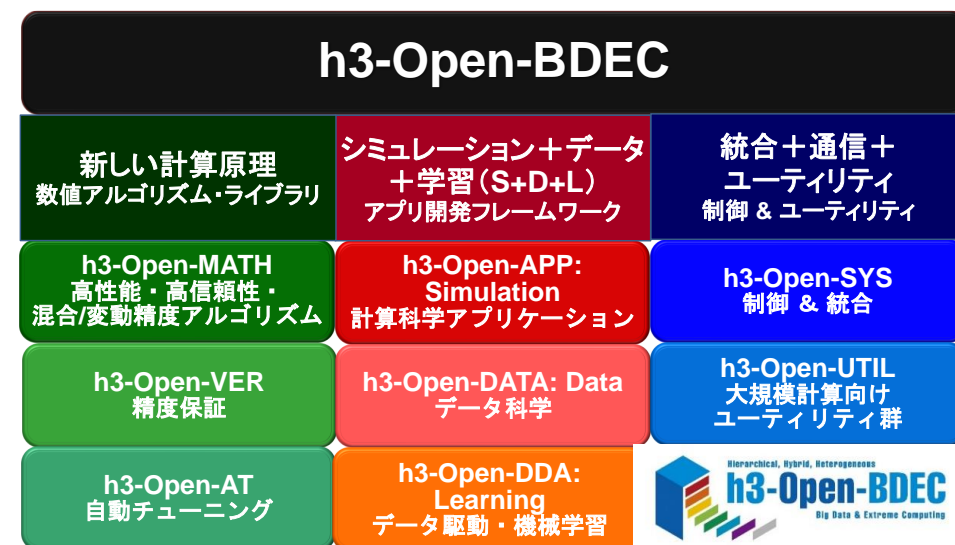
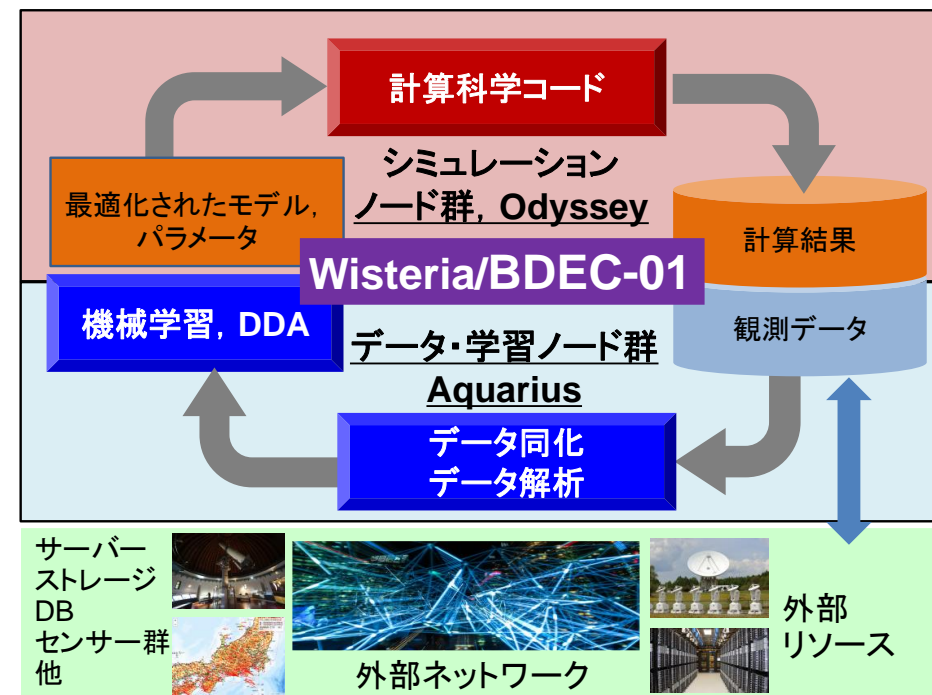
h3-Open-BDEC		
新しい計算原理 数値アルゴリズム・ライブラリ	シミュレーション+データ +学習 (S+D+L) アプリ開発フレームワーク	統合+通信+ ユーティリティ 制御 & ユーティリティ
h3-Open-MATH 高性能・高信頼性・ 混合/変動精度アルゴリズム	h3-Open-APP: Simulation 計算科学アプリケーション	h3-Open-SYS 制御 & 統合
h3-Open-VER 精度保証	h3-Open-DATA: Data データ科学	h3-Open-UTIL 大規模計算向け ユーティリティ群
h3-Open-AT 自動チューニング	h3-Open-DDA: Learning データ駆動・機械学習	



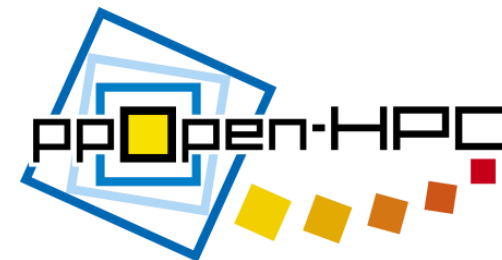
h3-Open-SYS/WaitIO

データ受け渡しライブラリ[松葉, 2020]

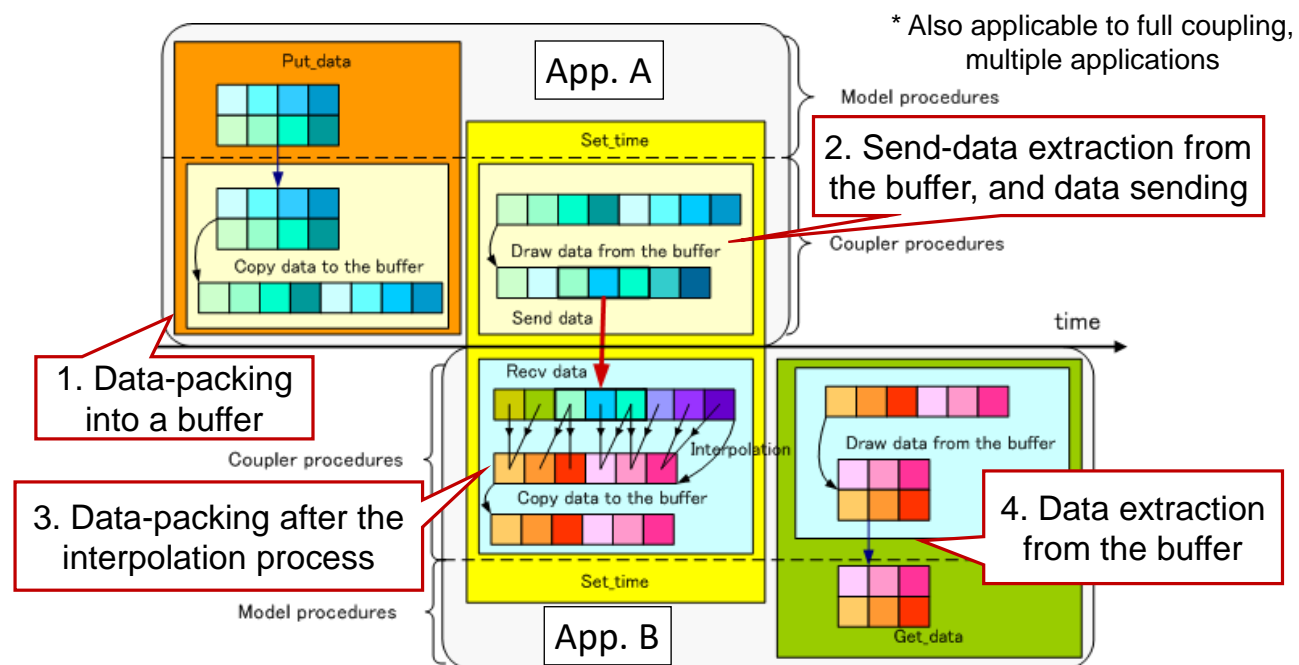
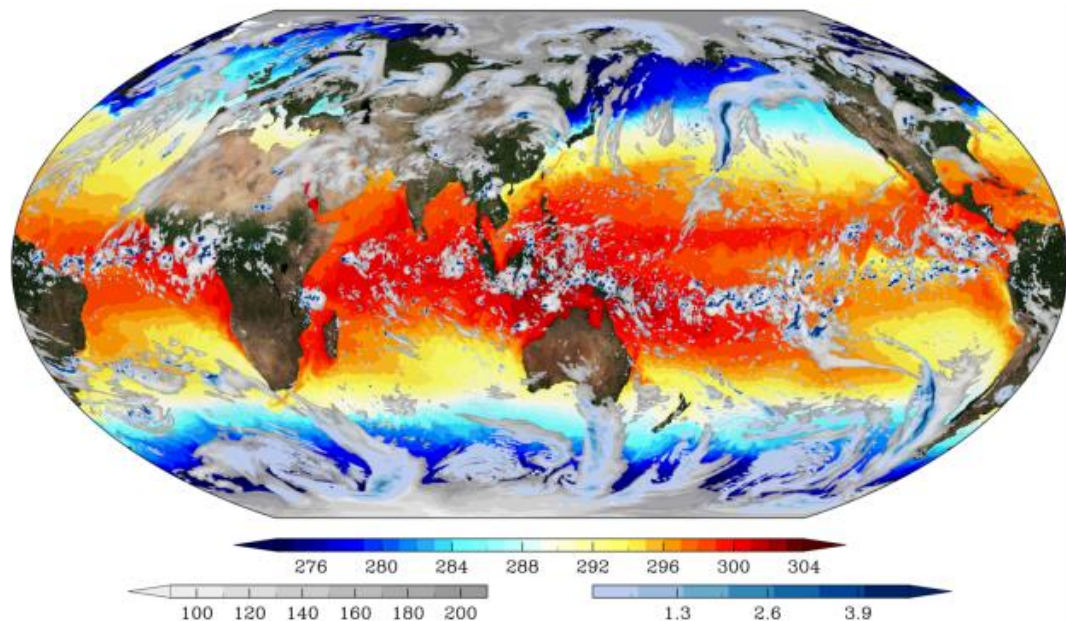
- ヘテロジニアス環境下での異なるコンポーネント間ファイル経由連携ライブラリとして考案
- 機能
 - ✓ Odyssey～Aquarius間連携
 - IB-EDR経由通信 (WaitIO-Socket)
 - ファイル経由 (WaitIO-File)
 - ✓ 外部からのデータ取得 (観測データ等)
 - ✓ 読み込み・書き出しの同期
- API: C/C++, Fortranから呼び出し可能
 - ✓ MPIライクなインタフェースを提供
- 多機能カプラー(h3-Open-UTIL/MP)との連携



連成シミュレーションのためのカプラー 〔荒川, 八代〕



- 従来のカプラー (Coupler) : ppOpen-MATH/MP
 - 複数 (通常2つ: 大気 (NICAM) + 海洋 (COCO)) のアプリケーションの弱連成 (Weak Coupling) をサポート
 - 各アプリケーションは1種類の計算をやる



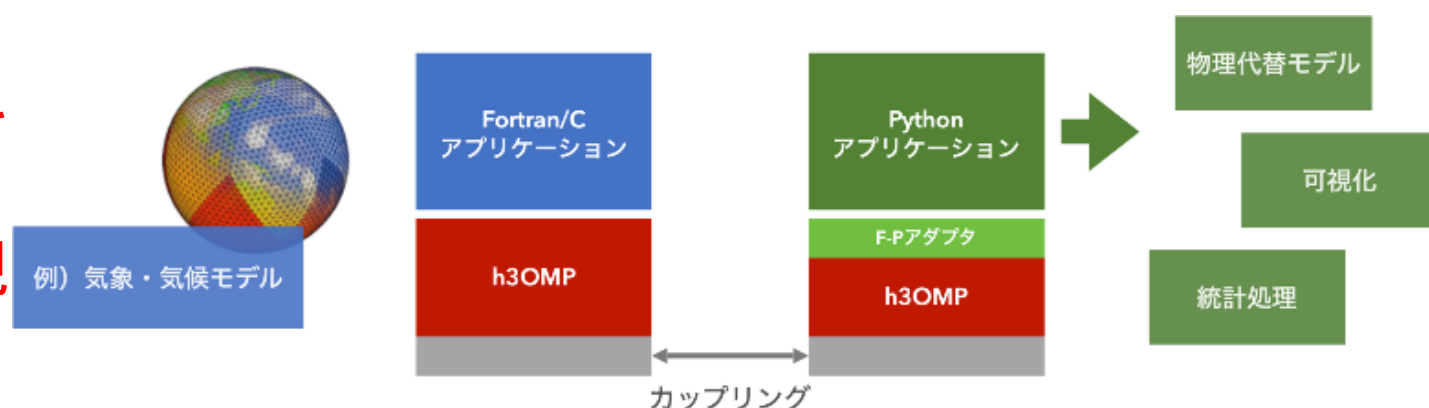
「計算＋データ＋学習」融合を支援する 多機能カプラーh3-Open-UTIL/MP



- 異なる物理モデル連成のアンサンブル実行を支援・統合するための機能
 - MPI通信、時刻同期、格子系間マッピング等の管理機能の他、従来のカプラーには無い、複数の弱連成結合シミュレーションのアンサンブル実行、片側のモデルのみをアンサンブル実行する多対1の弱連成結合が可能
 - スパコン上で、全地球大気海洋連成シミュレーションによって動作検証済み

Fortran/Cコード(物理モデル)とPythonコードの弱連成を実現する機能

- FortranやCで記述されたプログラム同士の連成計算に限って開発を行ってきたカプラーを、Pythonによって記述されたAI・機械学習、可視化処理系のワークロードから活用できるように機能拡充。



Fortran/CアプリとPythonアプリの連成計算の模式図
〔八代・荒川 2020〕

- O-A利用: WaitIOとの連携

更に詳細な情報

- A64FX (富士通)
 - <https://www.fujitsu.com/jp/products/computing/servers/supercomputer/a64fx/>
 - https://old.hotchips.org/hc30/2conf/2.13_Fujitsu_HC30.Fujitsu.Yoshida.rev1.2.pdf
- FUJITSU PRIMEHPC FX1000
 - <https://www.fujitsu.com/jp/products/computing/servers/supercomputer/>
- 3rd Gen Intel Xeon Scalable
 - <https://www.intel.com/content/www/us/en/newsroom/news/3rd-gen-intel-xeon-scalable-video.html#gs.zb3u0m>
 - <https://www.intel.com/content/www/us/en/newsroom/news/3rd-gen-xeon-scalable-processors.html#gs.zb4d00>
 - https://www.hotchips.org/assets/program/conference/day1/HotChips2020_Server_Processors_Intel_Irm_a_ICX-CPU-final3.pdf
- NVIDIA A100 TENSORコア GPU
 - <https://www.nvidia.com/ja-jp/data-center/a100/>
 - https://www.hotchips.org/assets/program/conference/day1/HotChips2020_GPU_NVIDIA_Choquette_v01.pdf

参考リンク(ビデオ)

- Wisteria/BDEC-01 利用説明会
 - <https://www.youtube.com/watch?v=1bbZVO6-UQg>
- h3-Open-BDEC: プロジェクトHP (工事中)
 - <http://nkl.cc.u-tokyo.ac.jp/h3-Open-BDEC/>
- Wisteria/BDEC-01 & h3-Open-BDEC 紹介講演 (日本語)
 - https://www.youtube.com/watch?v=CsJ_9aGNXCg
 - <https://www.pccluster.org/ja/event/pccc20/exhibition/itc-u-tokyo.html>
- Wisteria/BDEC-01 & h3-Open-BDEC 紹介講演 (英語)
 - <https://www.youtube.com/watch?v=jX51NF2LniE>



**Wisteria
BDEC-01**

