

第203回 お試しアカウント付き  
並列プログラミング講習会  
「MPI基礎：並列プログラミング入門」

東京大学 情報基盤センター  
三木 洋平

# 東京大学情報基盤センター



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

- 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4月現在
  - 2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）
    - 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
  - データ活用社会創成プラットフォーム（mdx）：2021年3月設置



**JCAHPC**

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  
100+ PF

Fujitsu FX10  
Oakleaf-FX  
1.13 PF

 Wisteria  
BDEC-01 Fujitsu  
33.1 PF

BDEC-  
02  
250+ PF

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

Ipomoea-01 25PB

Ipomoea-  
03

Ipomoea-02

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

2001-2005

2006-2010

2011-2015

2016-2020

2021-2025

2026-2030

Hitachi SR8000

SR8000

Hitachi SR11000

IBM Power5+

Hitachi SR16K/M1

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)

100 PF

OFP-II  
100+ PF

AMD Opteron

Intel Xeon Phi

Accelerators

疑似ベクトル

汎用CPU

加速装置付

Fujitsu FX10

Oakleaf-FX  
1.13 PF

Wisteria BDEC-01  
Fujitsu  
33.1 PF

BDEC-02  
250+ PF

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

Shared File System (SFS)  
25.8 PB, 500 GB/s

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

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

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

2.0 TB/s

800 Gbps

External Resources



External Network



External Resources



Simulation Nodes  
(Odyssey)



Data/Learning Nodes  
(Aquarius)



東京大学  
THE UNIVERSITY OF TOKYO



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

### 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
- 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
- 25 PF, #39 in 58<sup>th</sup> 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, #129 in 60<sup>th</sup> TOP500 (November 2022)



### Wisteria/BDEC-01 (Fujitsu)

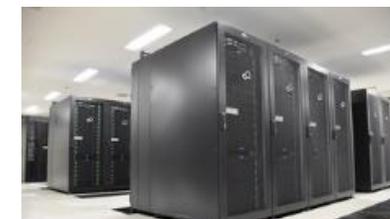
- シミュレーションノード群 (Odyssey) : A64FX (#17)
- データ・学習ノード群 (Aquarius) : Intel Icelake+NVIDIA A100 (#106)
- 33.1 PF, #17 in 58<sup>th</sup> TOP 500, 2021年5月14日運用開始
- 「計算・データ・学習 (S+D+L)」融合のためのプラットフォーム
- 革新的ソフトウェア基盤「h3-Open-BDEC」  
(科研費基盤 (S) 2019年度～2023年度)



Reedbush



Oakforest-PACS

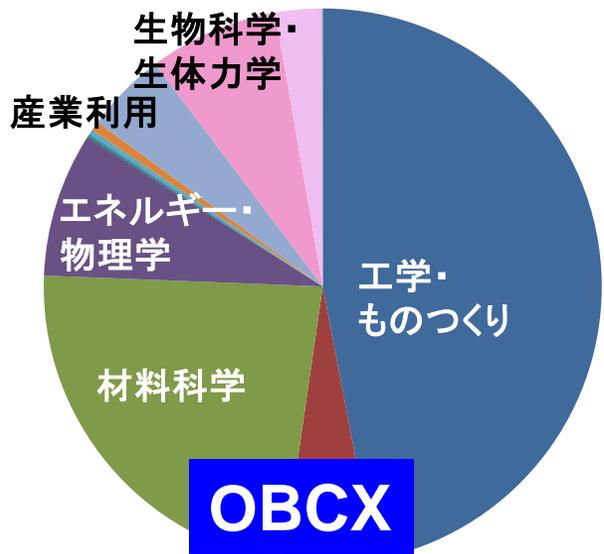
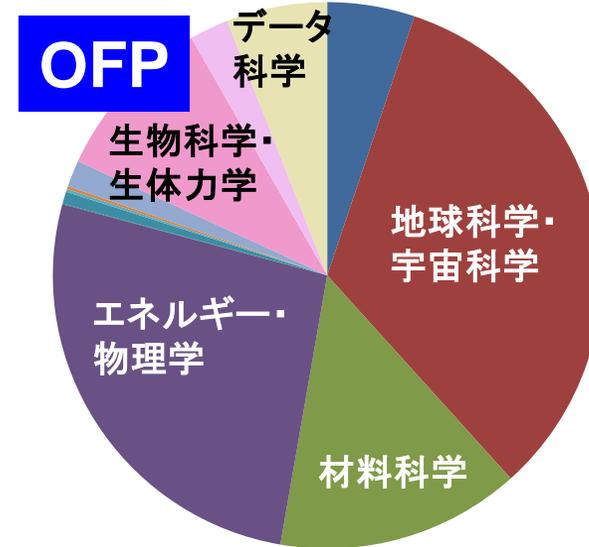
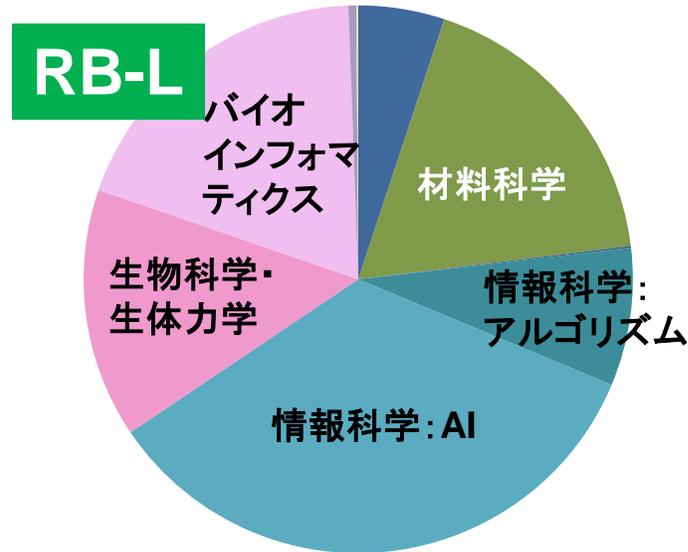
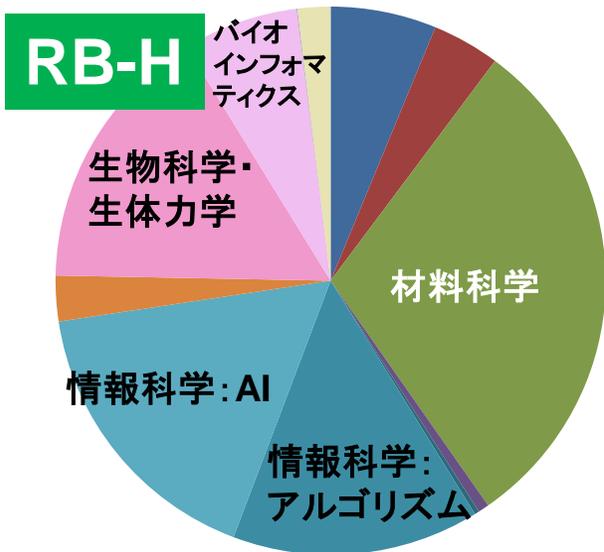


Oakbridge-CX

# 2020年度分野別 ■汎用CPU, ■GPU

RB: Reedbush, OFP: Oakforest-PACS, OBCX: Oakbridge-CX

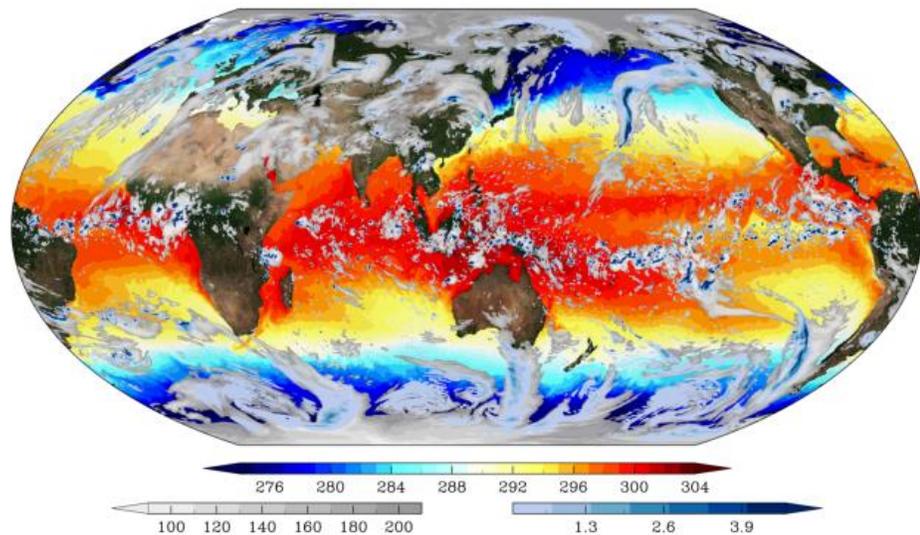
- 工学・ものづくり
- 地球科学・宇宙科学
- 材料科学
- エネルギー・物理学
- 情報科学:システム
- 情報科学:アルゴリズム
- 情報科学:AI
- 教育
- 産業利用
- 生物科学・生体力学
- バイオインフォマティクス
- 社会科学・経済学
- データ科学・データ同化



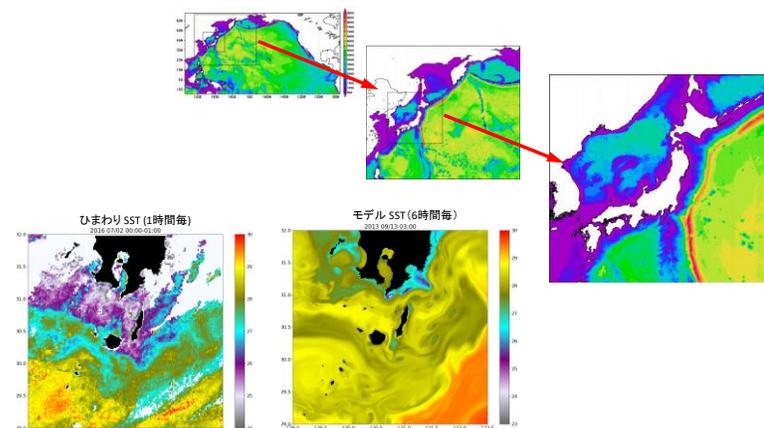
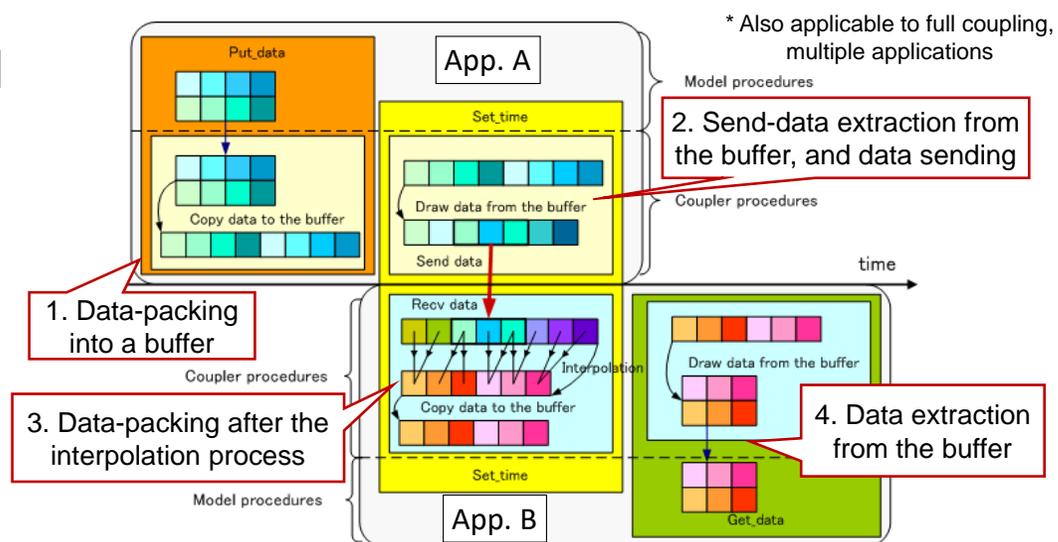
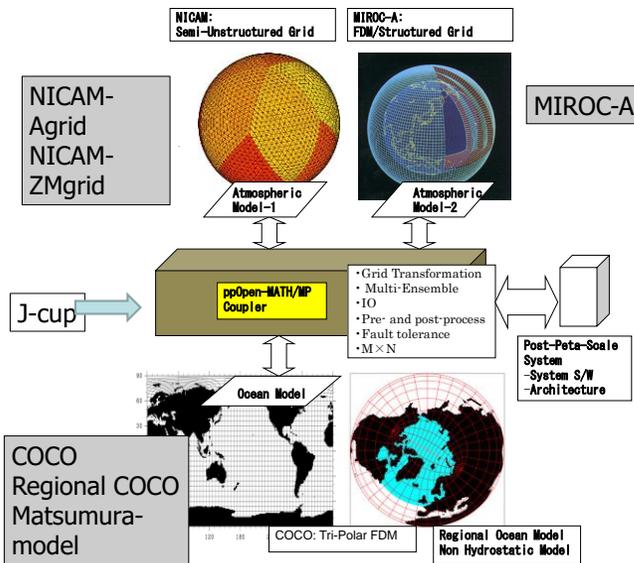
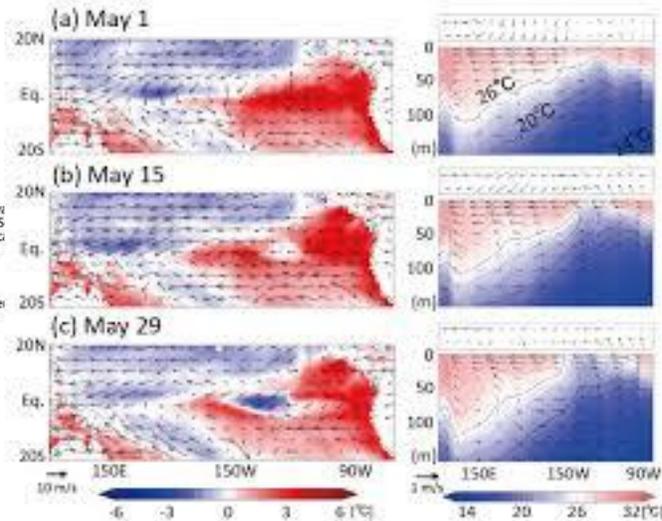
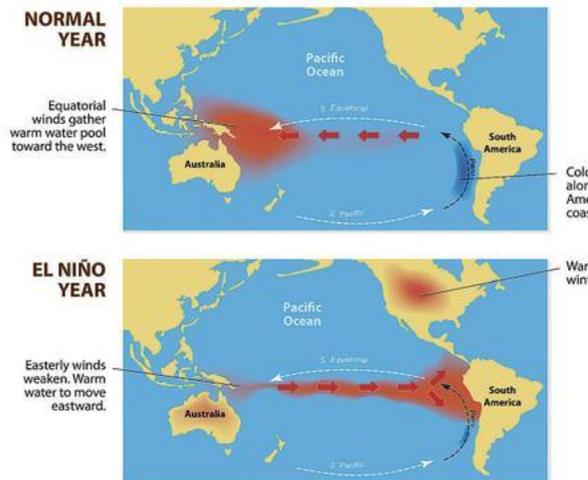
- 工学・ものづくり
- 地球科学・宇宙科学
- 材料科学

# 全地球大気・海洋シミュレーション(気候・気象)

東大大気海洋研究所, 東大理学系研究科等



## THE EL NIÑO PHENOMENON

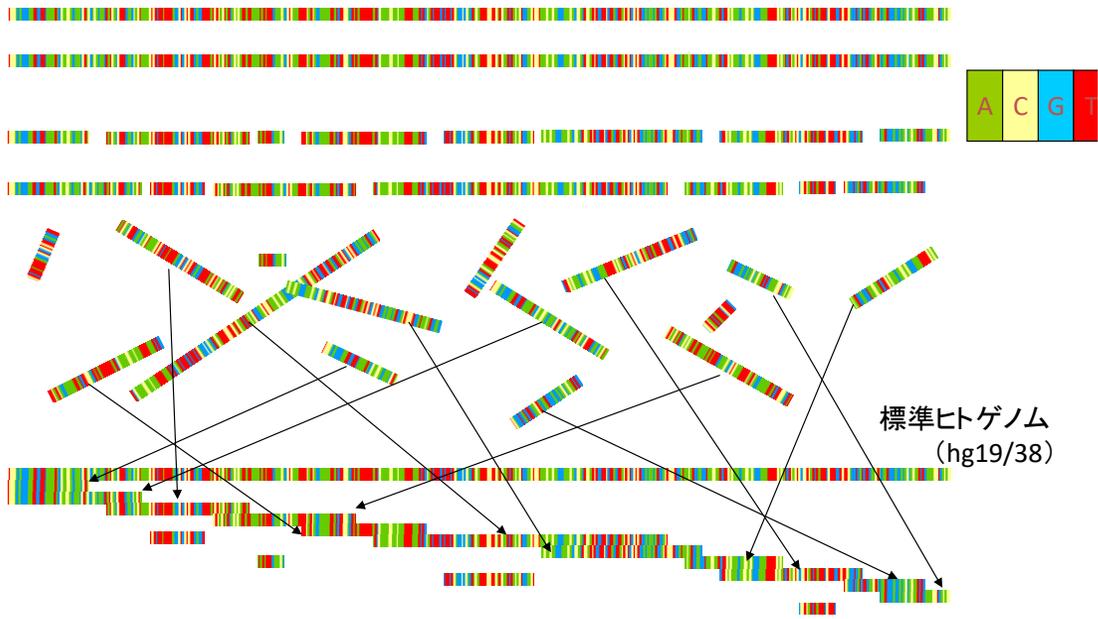


[画像提供: 佐藤正樹教授・羽角博康教授(東大・大気海洋研)]

# バイオインフォマティクス：ゲノム解析

東大新領域創成科学研究科 等

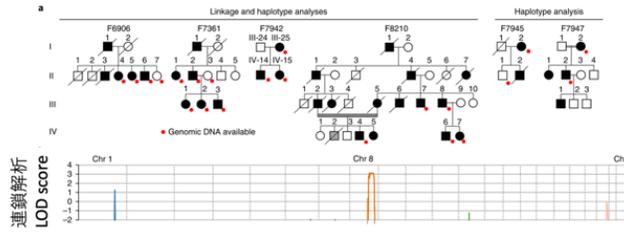
## ヒト個人ゲノムはどのように再解読するか？



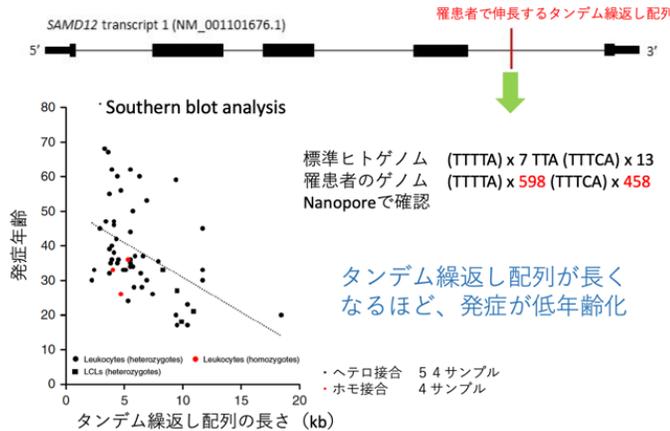
長さ 100~150 塩基のDNA 断片を 10~20億本収集 (ヒト1人当たり)  
接尾辞配列 (suffix array), Burrows-Wheeler 変換等が活用される

[画像提供：  
森下真一教授 (東京大学新領域創成科学研究科)]

良性成人型家族性ミオクローヌスてんかんの家系



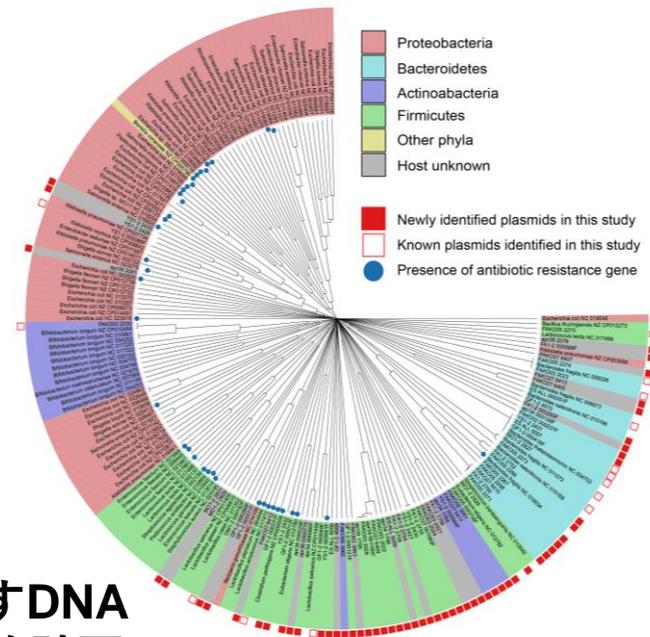
家系の連鎖解析により絞り込んだ8番染色体の領域に存在する38個遺伝子のコード領域には、原因となる1塩基変異が見つからなかった。しかしこの領域のSAMD12のイントロンに、罹患者で伸長するタンデム繰返し配列を発見。



Nature Genetics 2018

新たに発見された疾患を引き起こすDNAの繰返し配列伸長異常：様々な動的計画法, de Bruijn グラフ探索等が活用

## ヒト腸内細菌叢から発見された多様なプラスミド・ファージ配列の全貌

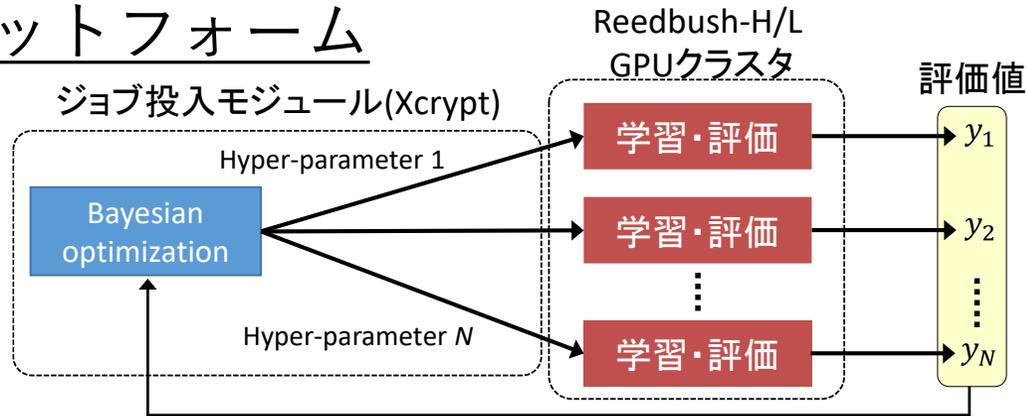


Microbiome 2019

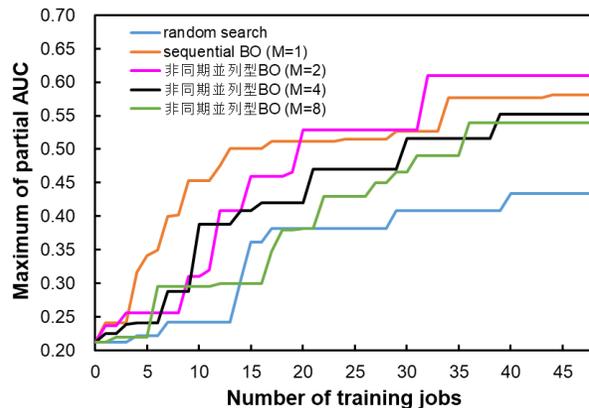
# バイオインフォマティクス：医療画像処理

東大病院等

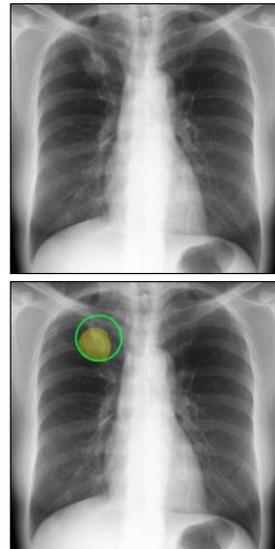
## 深層学習自動チューニングプラットフォーム



## 胸部X線写真の肺腫瘍検出



学習ジョブ数と評価値(partial AUC)の最大値との関係

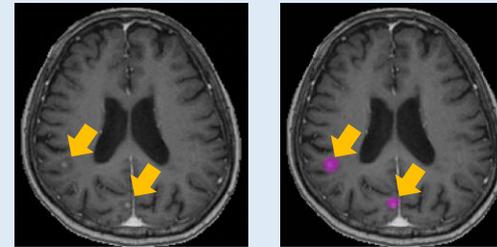


上：元画像、下：検出結果  
(黄、緑丸：病変領域)

## 開発中のソフトウェア

### 頭部造影MR画像の転移性脳腫瘍検出

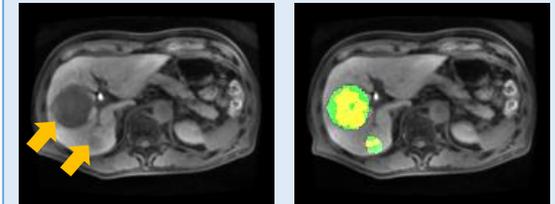
村田, JAMIT2018



検出結果例  
左：元画像、右：検出結果(マゼンダ)

### 造影MR画像の肝結節性病変検出

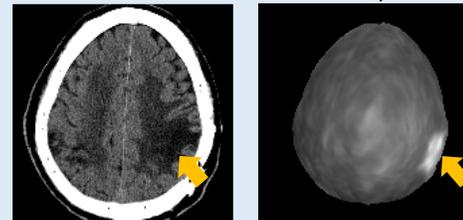
Takenaga T, CARS 2018



検出結果例  
(肝細胞がん、左：元画像、右：検出結果)  
●：検出、●：過検出、●：未検出

### 頭部救急CT画像の異常検知

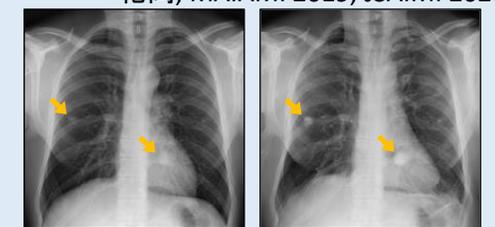
Sato D, SPIE MI 2018



脳梗塞症例  
(左：元画像、右：異常度マップ)

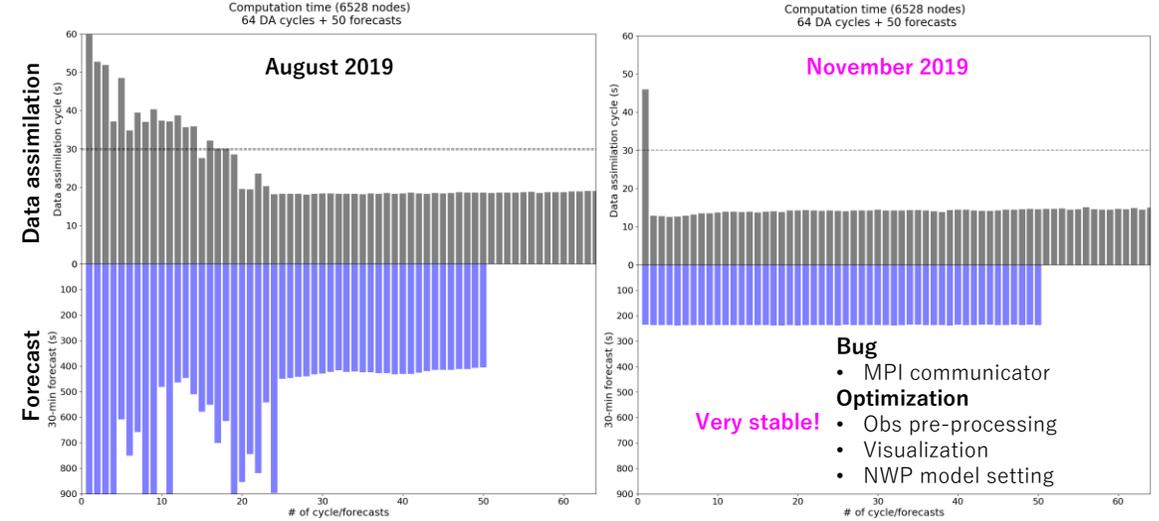
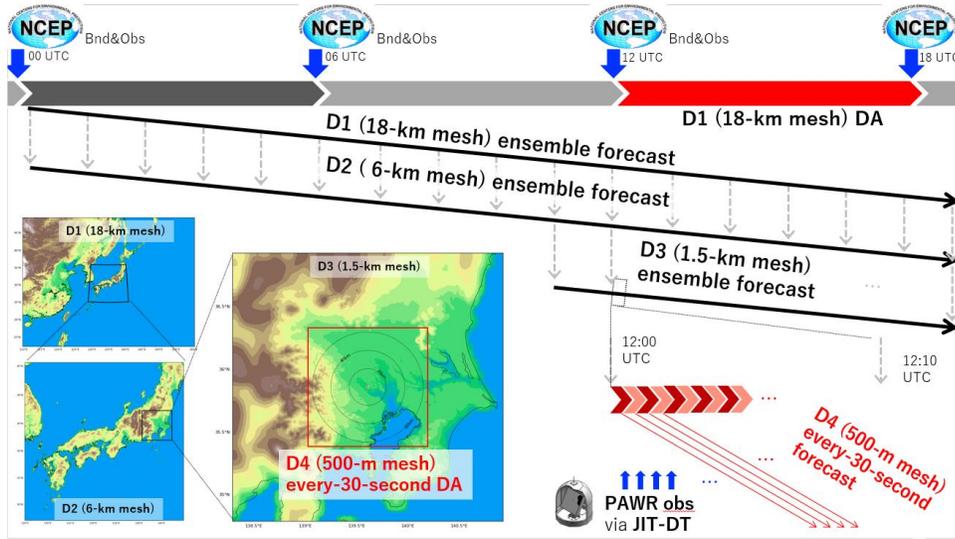
### 胸部X線画像の異常強調

花岡, MAIAMI 2019, JSAIMI 2020



強調画像例  
(左：元画像、右：強調結果、矢印：肺腫瘍)

# ゲリラ豪雨予測のリアルタイム実証実験 (理化学研究所)



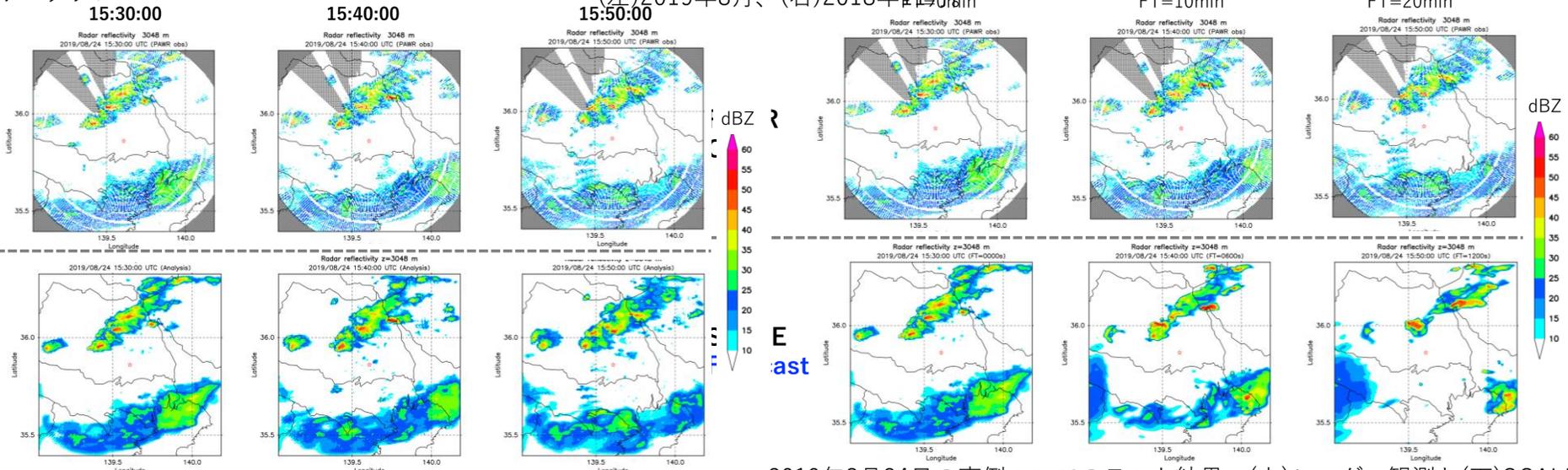
全体のワークフロー

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



PAWR Obs

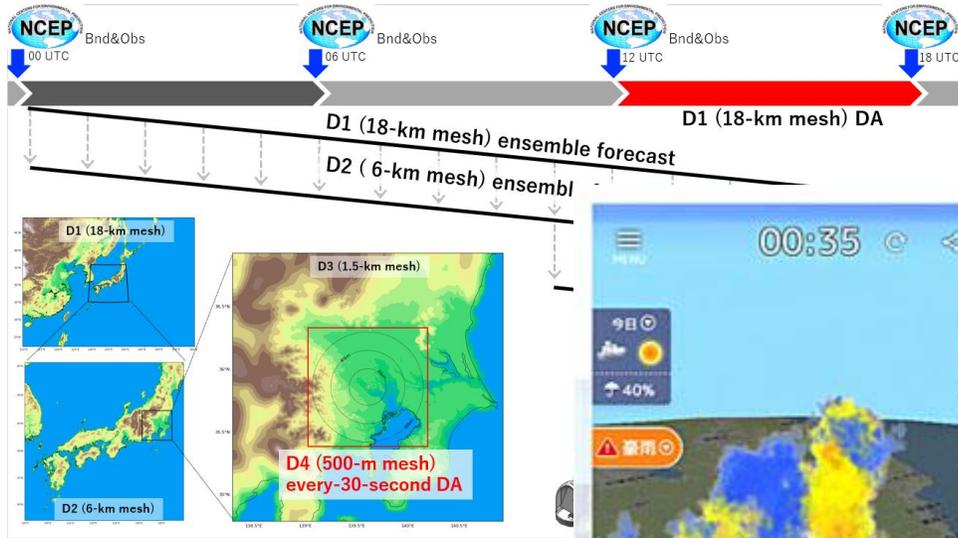
SCALE-LETKF Analysis



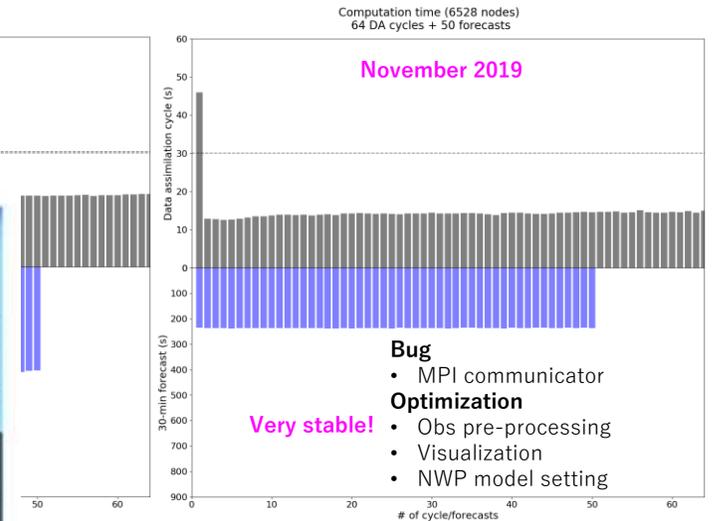
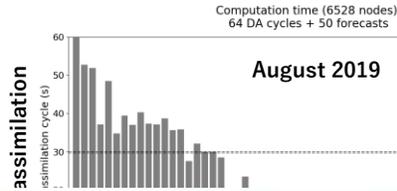
〔画像提供: 三好建正博士 (理化学研究所)〕

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。

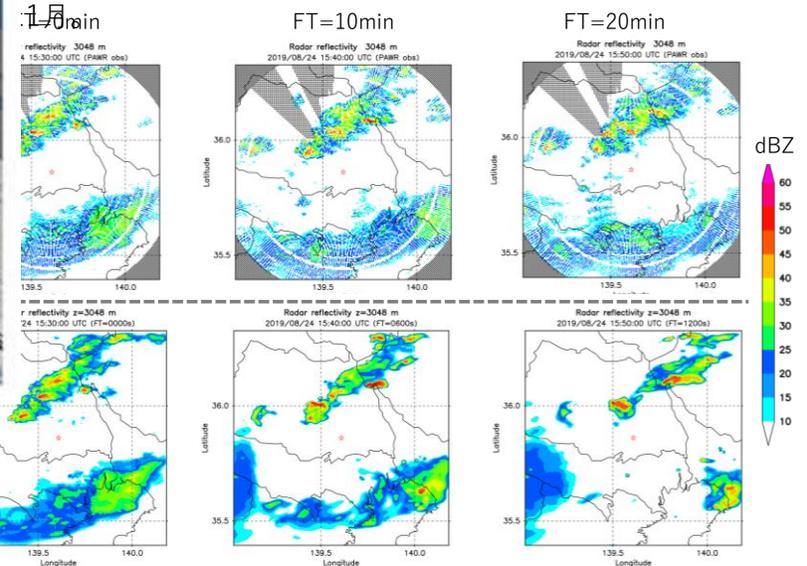
# ゲリラ豪雨予測のリアルタイム実証実験 (理化学研究所)



全体のワークフロー

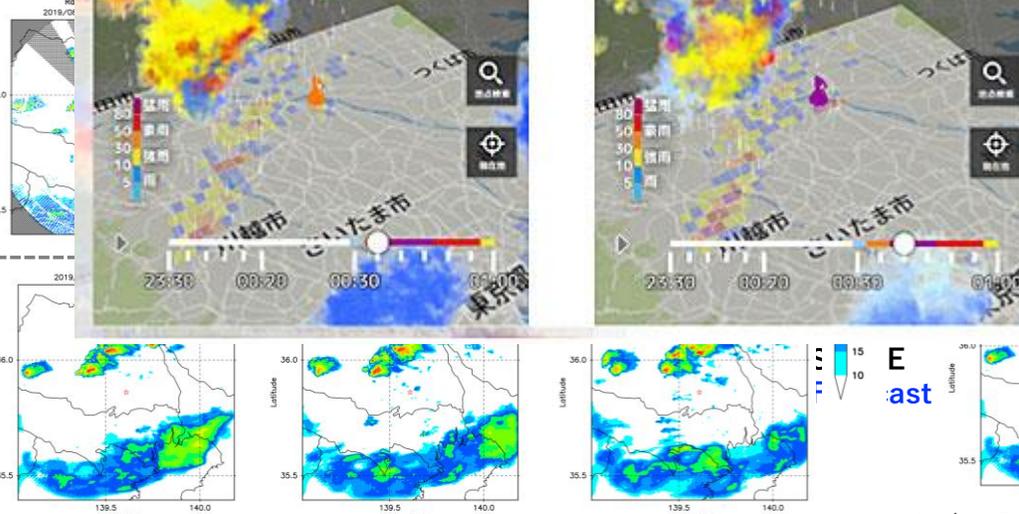


夕同化、下段は30分予報にかかった時間(秒)。



PAWR Obs

SCALE-LETKF Analysis



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。

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による予報で得られたレーダー反射強度(dBZ)を示す。

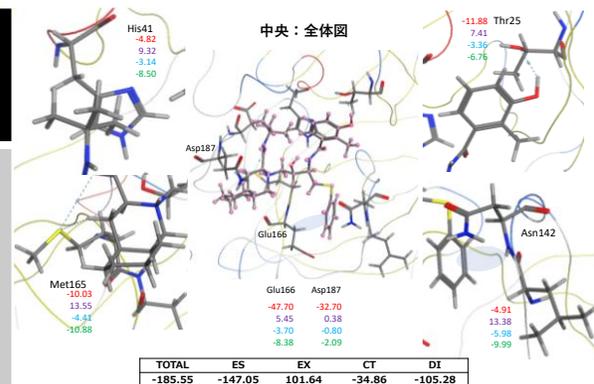
〔画像提供: 三好建正博士 (理化学研究所)〕

# 「COVID-19」対応HPCI臨時公募課題

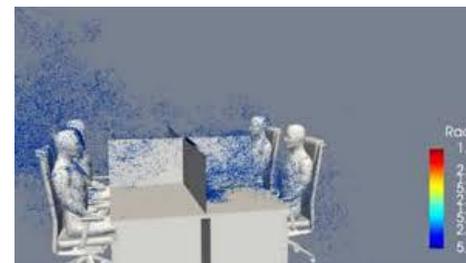
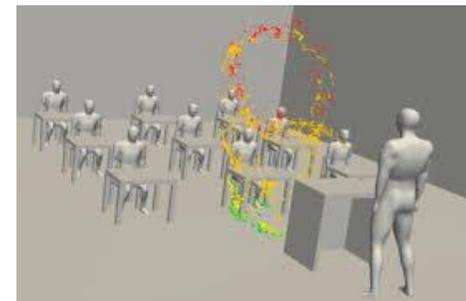
全14のうち6課題が東大システムを利用（2020年度）



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/UC2CHaGp1A0-vqR1V7wmU0-w>

- Wisteria/BDEC-01システム紹介

[https://www.youtube.com/watch?v=SXjYtatz0-4&list=PLobjSv\\_ny85lW030APUJ9DWJoHhNiQgvY&index=3&t=104s](https://www.youtube.com/watch?v=SXjYtatz0-4&list=PLobjSv_ny85lW030APUJ9DWJoHhNiQgvY&index=3&t=104s)

- 第10回JCAHPCセミナー

[https://www.youtube.com/playlist?list=PLobjSv\\_ny85mfPTuCC2i7r\\_sPQYKZvy2e](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\\_ny85kr1Ig2m-bUiMC2a9W6k53u](https://www.youtube.com/playlist?list=PLobjSv_ny85kr1Ig2m-bUiMC2a9W6k53u)  
<https://www.youtube.com/watch?v=q-0QtU70ps4&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」

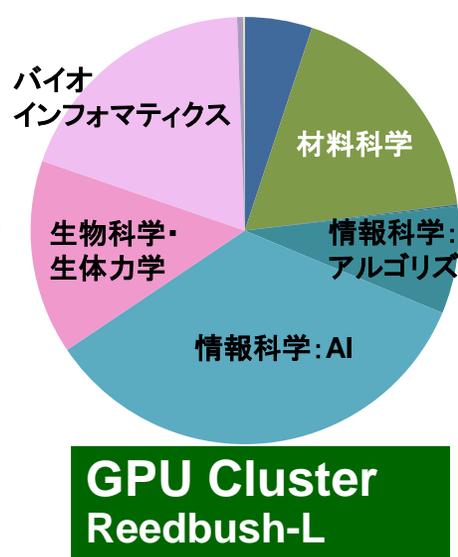
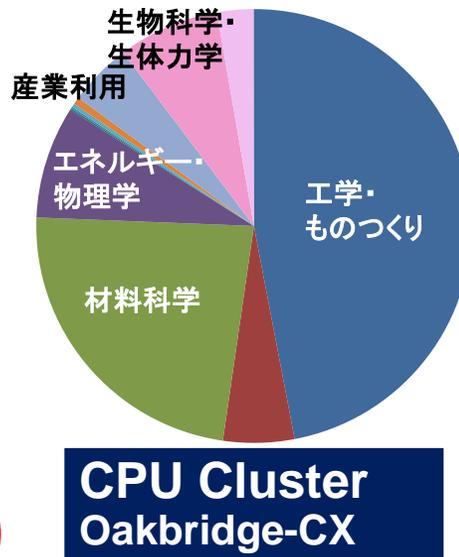
[https://www.youtube.com/playlist?list=PLobjSv\\_ny85l-z-VJCy690Zj1AA04xCRA](https://www.youtube.com/playlist?list=PLobjSv_ny85l-z-VJCy690Zj1AA04xCRA)

- お試しアカウントつき講習会

[https://www.youtube.com/playlist?list=PLobjSv\\_ny85kXY2Mtnhn1k7pM-epQaD2y](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**



- 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



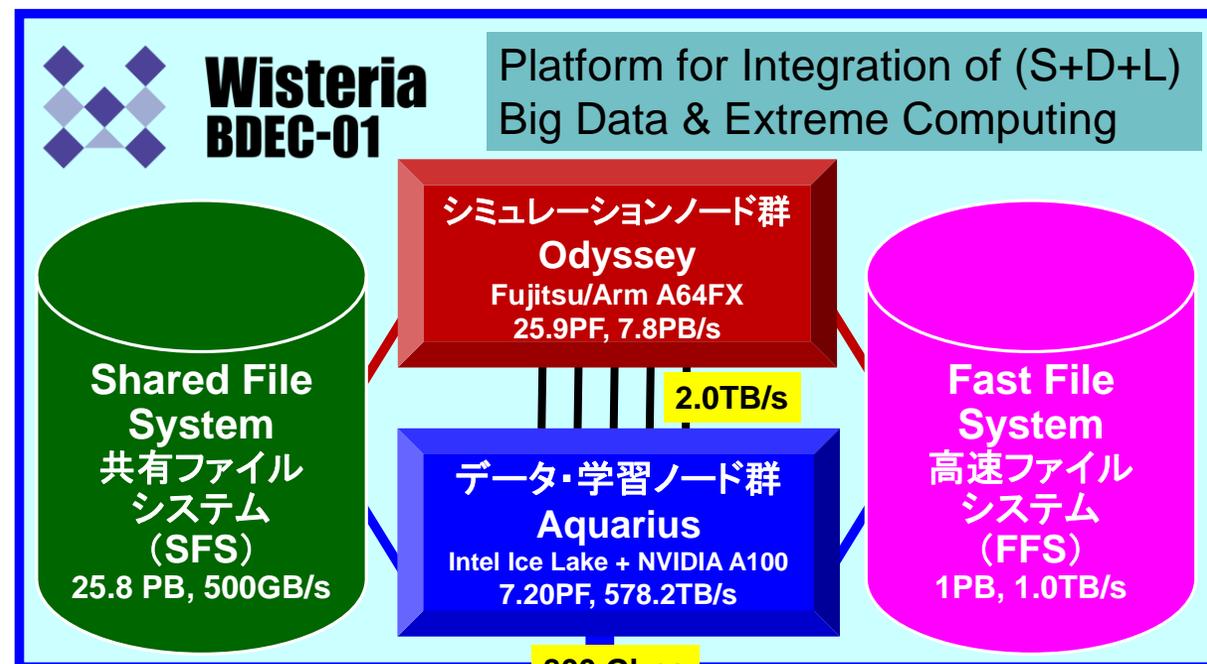
# Wisteria/BDEC-01



# Wisteria BDEC-01

2021年5月14日から柏Ⅱキャンパスで運用を開始した国内最大級のシステム

「Wisteria/BDEC-01」は、「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



External Network  
外部ネットワーク



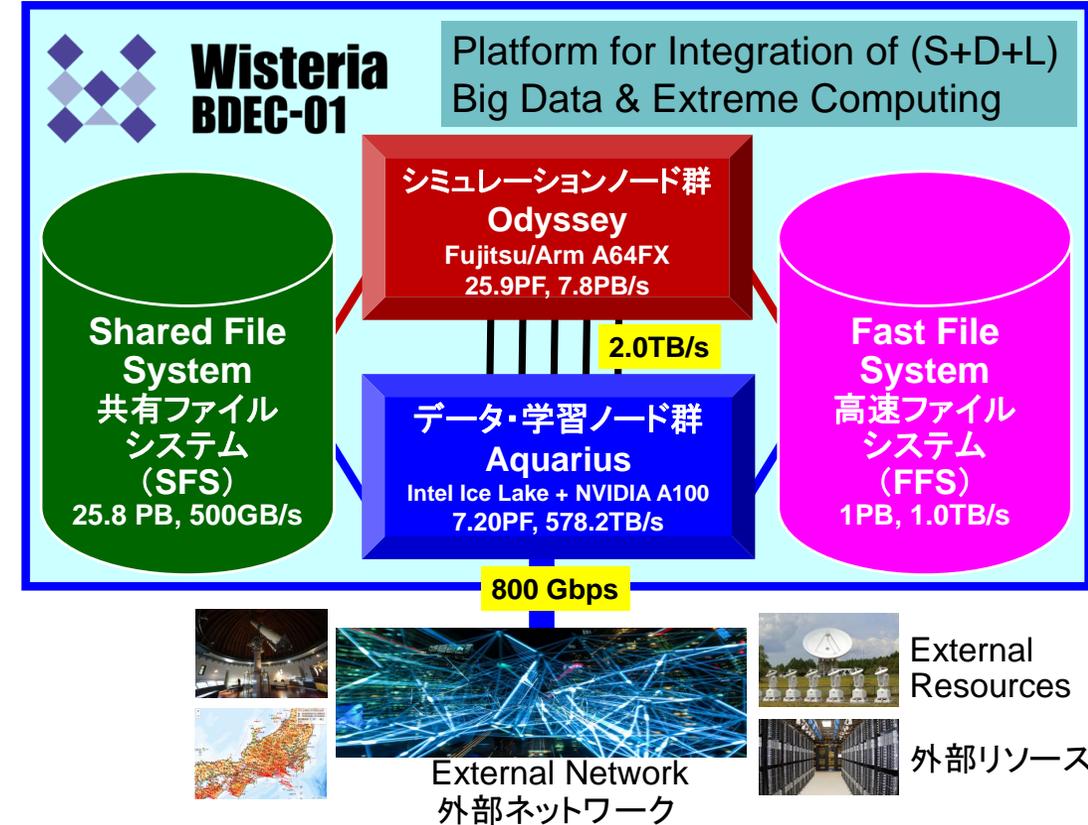
External Resources

外部リソース

# 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) , ~360m<sup>2</sup>
- **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

- 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) , ~360m<sup>2</sup>
- **2種類のノード群**

- シミュレーションノード群 (S, SIM) :

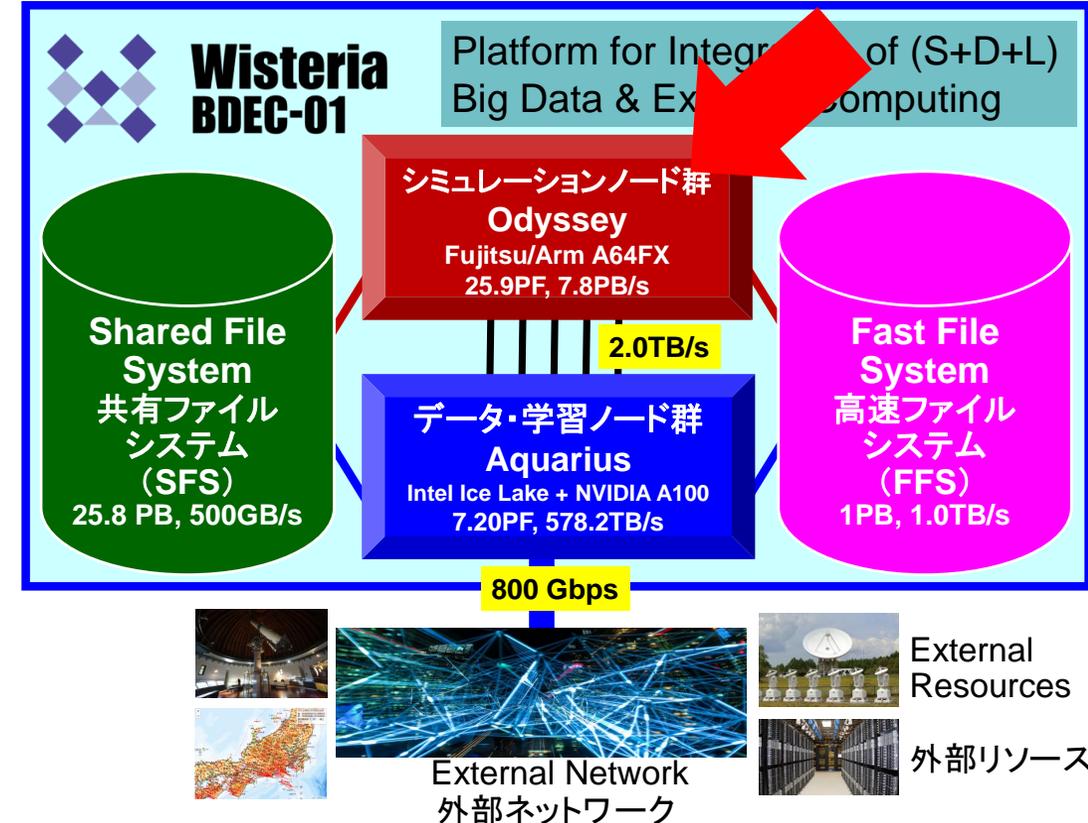
## Odyssey

- 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)



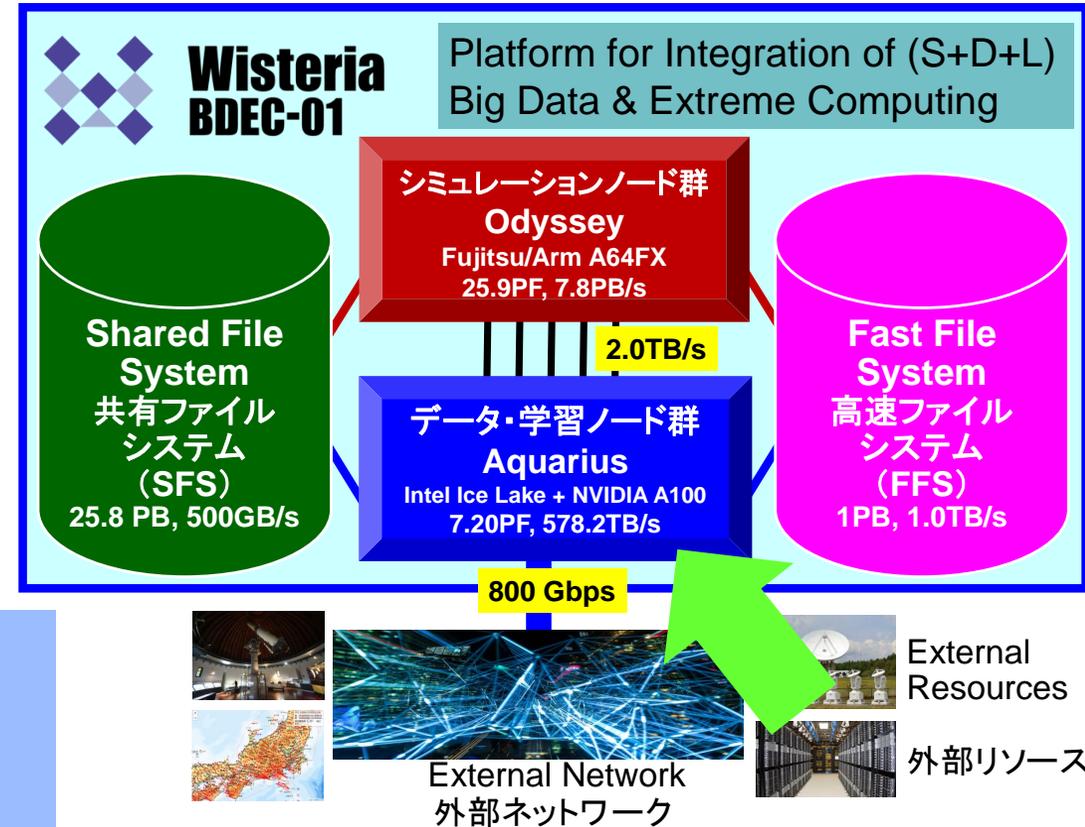
**Wisteria  
BDEC-01**

# 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) , ~360m<sup>2</sup>
- **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

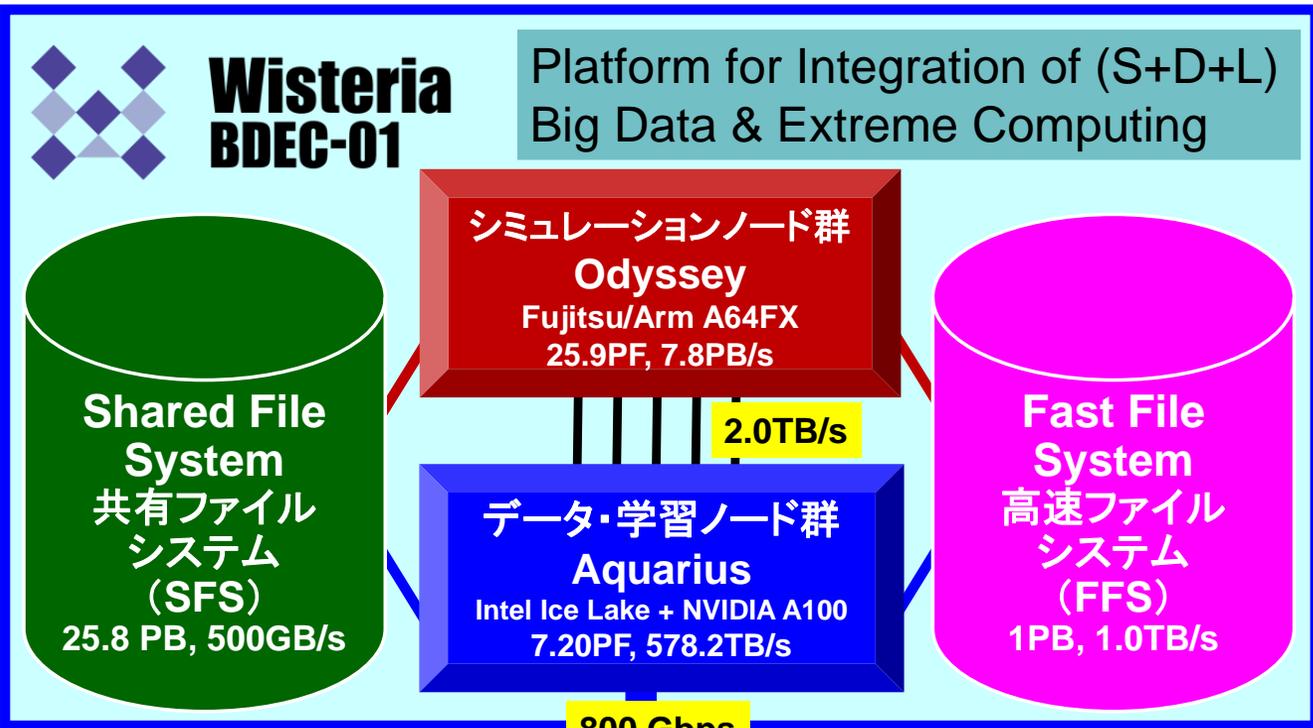
BDEC: 「計算・データ・学習 (S+D+L)」  
融合のためのプラットフォーム  
(Big Data & Extreme Computing)



**Wisteria**  
**BDEC-01**

# Wisteria/BDEC-01

## (S+D+L) 融合プラットフォーム



- **Wisteria (紫藤)**
  - 手賀沼 (柏市) に伝わる「藤姫伝説」
- **Odyssey**
  - アポロ13号・司令船 (Command Module, CM) のコールサイン
- **Aquarius**
  - アポロ13号・月着陸船 (Lunar Module, LM) のコールサイン



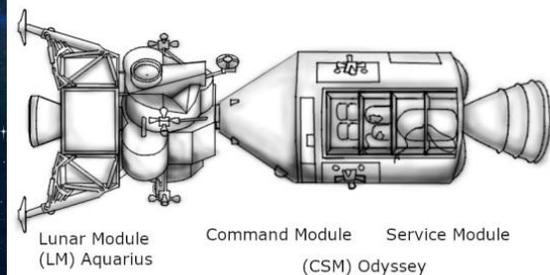
External Network  
外部ネットワーク



External Resources

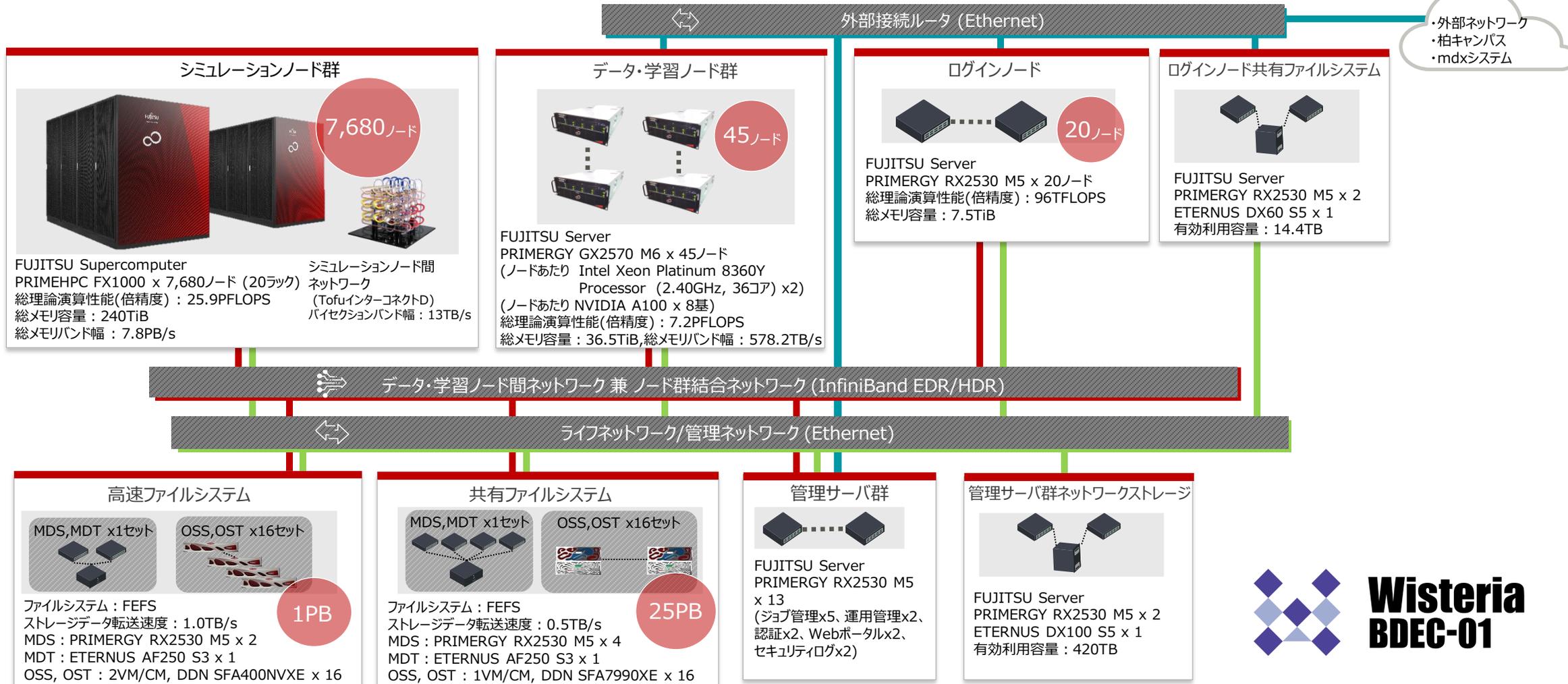


外部リソース



# 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)



Simulation Nodes

**Odyssey**

25.9 PF, 7.8 PB/s

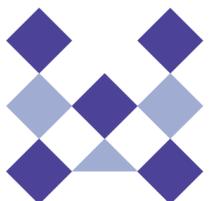
Fast File System (FFS)  
1.0 PB, 1.0 TB/s

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

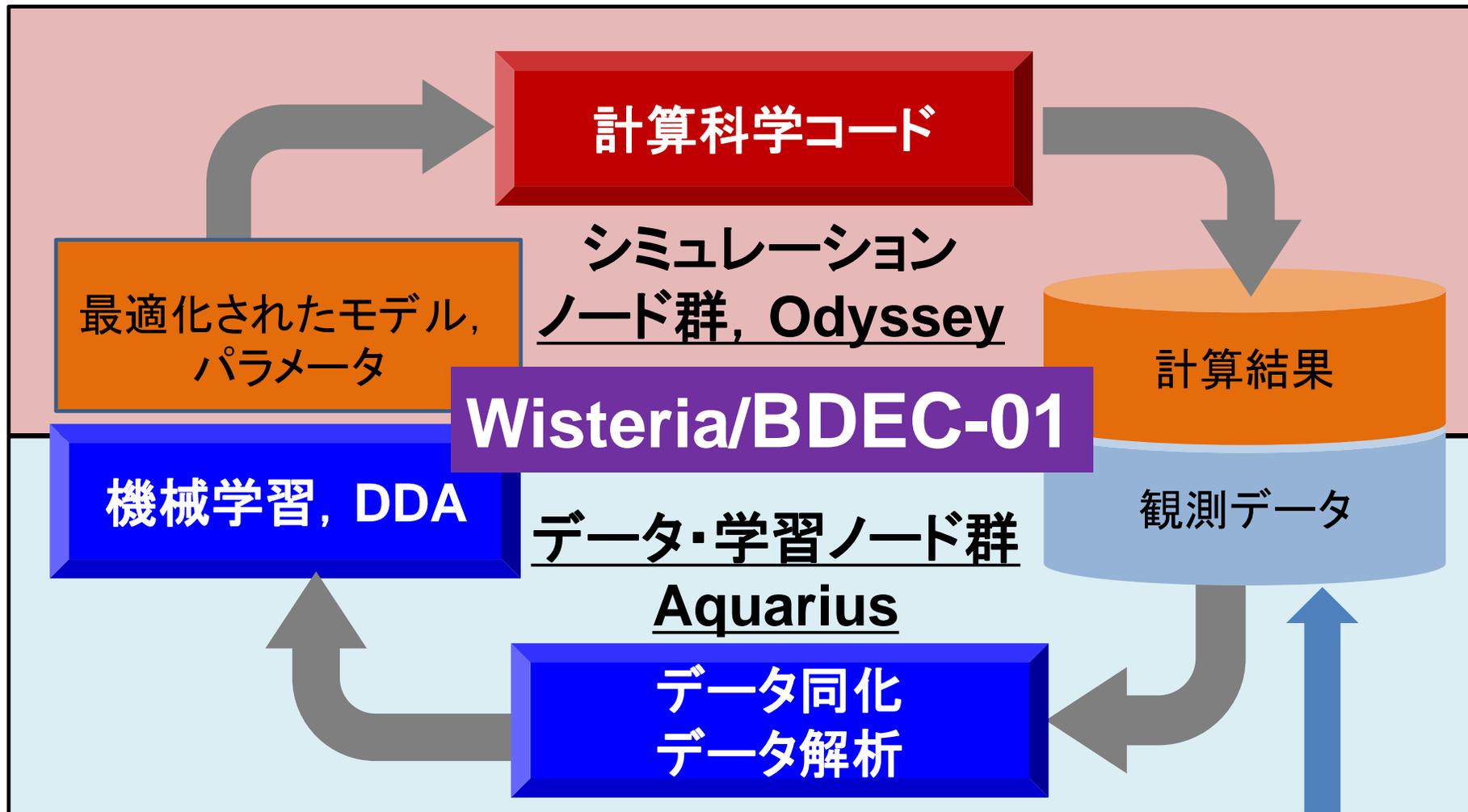
Data/Learning Nodes

**Aquarius**

7.20 PF, 578.2 TB/s



**Wisteria BDEC-01**



サーバー  
ストレージ  
DB  
センサー群  
他



外部ネットワーク



外部  
リソース

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

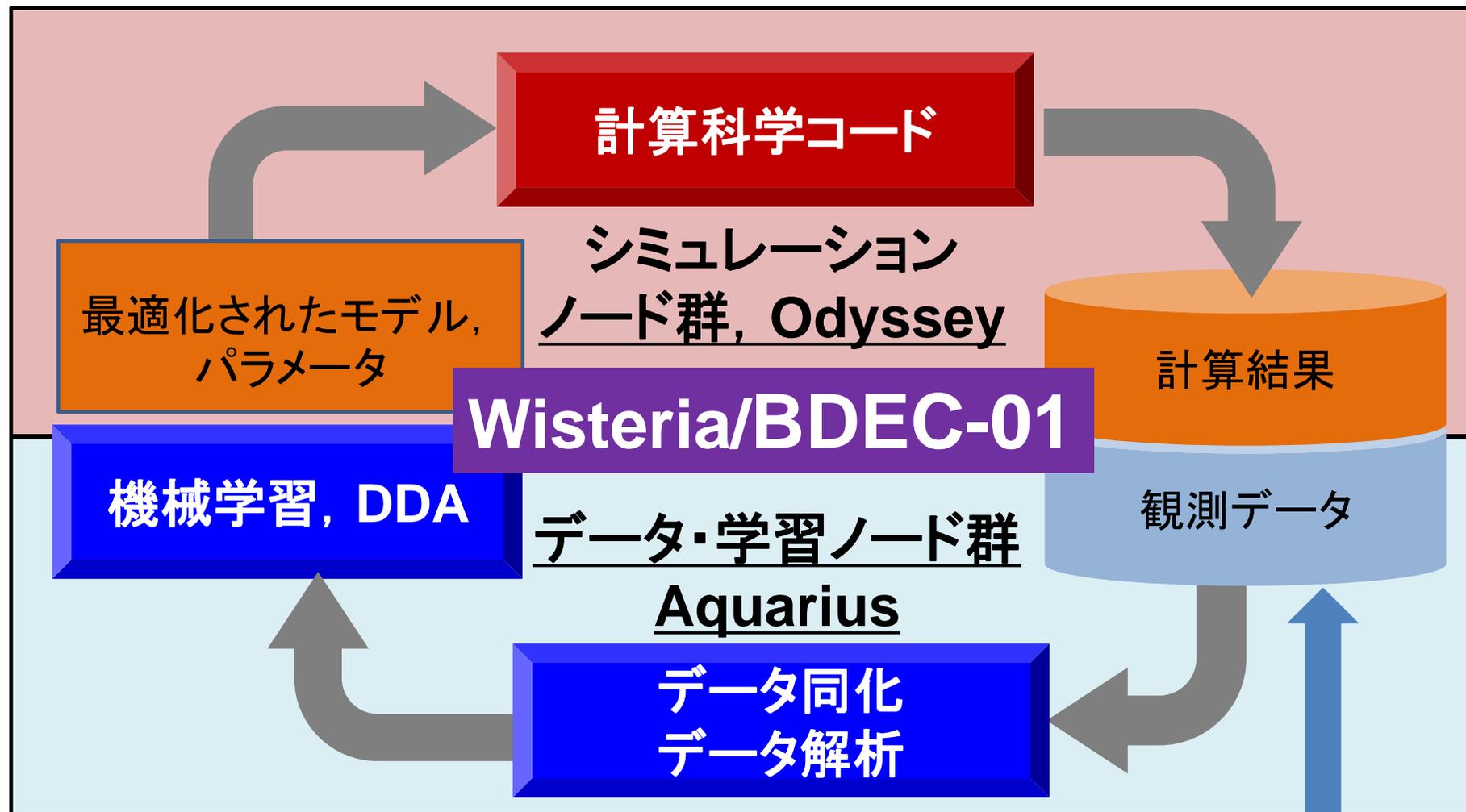
Data/Learning Nodes

**Aquarius**

7.20 PF, 578.2 TB/s



**Wisteria  
BDEC-01**



シミュレーションのためのモデル・パラメータのデータ解析, AI/機械学習による最適化 (S+D+L) = AI for HPC

# November 2022 (SC22)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
	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	129	85	83	-	-
Wisteria/BDEC- 01 (Odyssey)	23	45	12	4	14
Wisteria/BDEC- 01 (Aquarius)	125	28	68	-	-

# 60<sup>th</sup> TOP500 List (Nov., 2022)

R<sub>max</sub>: Performance of Linpack (TFLOPS)

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

R<sub>peak</sub>: Peak Performance (TFLOPS), Power: kW

	Site	Computer/Year Vendor	Cores	R <sub>max</sub> (PFLOPS)	R <sub>peak</sub> (PFLOPS)	Power (kW)
1	<b>Frontier, 2022, USA</b> DOE/SC/Oak Ridge National Laboratory	HPE Cray EX235a, AMD Optimized 3 <sup>rd</sup> Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	8,730,112	1,102.00 (=1.102 EF)	1,685.65	21,100
2	<b>Fugaku, 2020, Japan</b> 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	<b>LUMI, 2022, Finland</b> EuroHPC/CSC	HPE Cray EX235a, AMD Optimized 3 <sup>rd</sup> Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	2,220,288	309.10	428.70	6,016
4	<b>Leonardo, 2022, Italy</b> EuroHPC/CINECA	BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64GB, Quad-rail NVIDIA HDR Infiniband	1,436,616	174.70	255.75	5,610
5	<b>Summit, 2018, USA</b> 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	<b>Sierra, 2018, USA</b> 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	<b>Sunway TaihuLight, 2016, China</b> National Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	10,649,600	93.01	125.44	15,371
8	<b>Perlmutter, 2021, USA</b> 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	2,589
9	<b>Selene, 2020, USA</b> 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	<b>Tianhe-2A, 2018, China</b> 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
22	<b>ABCI 2.0, 2021, Japan</b> 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
23	<b>Wisteria/BDEC-01 (Odyssey), 2021, Japan</b> ITC, University of Tokyo	PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu interconnect D	368,640	22.12	25.95 <sup>24</sup>	1,468

# Green 500 Ranking (Jun., 2022)

<https://www.top500.org/lists/green500>

	TOP 500 Rank	System	Accelerator	Cores	HPL Rmax (P flop/s)	Power (kW)	GFLOPS/W
1	405	Henri, Flatiron Institute, USA	NVIDIA H100 80GB PCIe	5,920	2.04	31	65.091
2	32	Frontier TDS, ORNL, USA	AMD Instinct MI250X	120,832	19.20	309	62.684
3	11	Adastral, GENCI-CINES, France	AMD Instinct MI250X	319,072	46.10	921	58.021
4	15	Satonix--GPU, Pawsey Supercomputing Centre, Australia	AMD Instinct MI250X	181,248	27.16	477	56.983
5	68	Dardel GPU, KTH - Royal Institute of Technology, Sweden	AMD Instinct MI250X	52,864	8.26	146	56.491
6	1	Frontier, ORNL, USA	AMD Instinct MI250X	8,730,112	1,102.00	21,100	52.227
7	3	LUMI, EuroHPC/CSC, Finland	AMD Instinct MI250X	2,220,288	309.10	6,016	51.382
8	159	ATOS THX.A.B, Atos, France	NVIDIA A100 SXM4 64GB	25,056	3.50	86	41.411
9	359	MN-3, Preferred Networks, Japan	MN-Core	1,664	2.181	53	40.901
10	331	Champollion, HPE, France	NVIDIA A100 SXM4 80GB	19,840	2.32	60	38.555
28	125	Wisteria/BDEC-01 (Aquarius), Fujitsu, Japan	NVIDIA A100	42,120	4.425	183.93	24.06

# HPCG Ranking (November 2022)

	Computer	Cores	HPL Rmax (Pflop/s)	TOP500 Rank	HPCG (Pflop/s)
1	Fugaku	7,630,848	442.010	2	16.004
2	Frontier	8,730,112	1102.00	1	14.054
3	LUMI	2,220,288	309.10	3	3.408
4	Summit	2,414,592	148.600	5	2.926
5	Leonardo	1,463,616	174.70	4	2.567
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	23	0.817

# Graph500 BFS, November 2022

	Site	Computer/ Vendor	Cores	Scale	GTEPS
1	<b><u>Fugaku, 2020, Japan</u></b> R-CCS, RIKEN	Fujitsu PRIMEHPC FX1000, Fujitsu A64FX 48C 2.2GHz, Tofu-D	7,630,848	41	102956
2	<b><u>Pengcheng Cloudbrain-II, China</u></b> Pengcheng Lab	Kunpeng 920 + Ascend 910	93,696	40	25242.9
3	<b><u>Sunway TaihuLight, 2016, China</u></b> National Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	10,599,680	40	23755.7
4	<b><u>Wisteria/BDEC-01 (Odyssey), 2021, Japan</u></b> ITC, University of Tokyo	Fujitsu PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu interconnect D	368,640	37	16118
5	<b><u>Toki-Sora, 2021, Japan</u></b> JAXA	PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu interconnect D	276,480	36	10813
6	<b><u>LUMI-C, 2021, Finland</u></b> EuroHPC/CSC	Cray EX, SlingShot-10	190,976	38	8467.7
7	<b><u>Summit (CPU Only), 2018, USA</u></b> 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
8	<b><u>SuperMUC, 2018, Germany</u></b> Leibniz	ThinkSystem SD650, Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path	196,608	39	6279.47
9	<b><u>Lise, 2021, Germany</u></b> ZIB	Bull Intel Cluster Intel Xeon Platinum 9242 48C 2.3GHz Intel Omni-Path	121,920	38	5423.94
10	<b><u>DepGraph Supernode, China</u></b> National Engineering Research Center for Big Data Technology and System	DepGraph +GPU Tesla A100	128	33	4623.379

# HPL-MxP (November 2022)

$R_{\max}$ : Performance of Linpack (EFLOPS)

<https://hpl-mxp.org/>

	Site	Computer/Vendor	Cores	HPL-AI (EFLOPS)	Top500	HPL $R_{\max}$ (EFLOPS)	Speedup
1	<a href="#"><u>Frontier, 2022, USA</u></a> DOE/ORNL	HPE Cray EX235a, AMD Optimized 3 <sup>rd</sup> Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	8,730,112	7.942	1	1.1020	7.2
2	<a href="#"><u>LUMI, 2022, Finland</u></a> EuroHPC/CSC	HPE Cray EX235a, AMD Optimized 3 <sup>rd</sup> Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	2,174,976	2.168	3	0.3091	7.0
3	<a href="#"><u>Fugaku, 2020, Japan</u></a> R-CCS, RIKEN	Fujitsu PRIMEHPC FX1000, Fujitsu A64FX 48C 2.2GHz, Tofu-D	7,630,848	2.0	2	0.4420	4.5
4	<a href="#"><u>Leonardo, 2022, Italy</u></a> EuroHPC/CINECA	BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100, NVIDIA HDR100 InfiniBand	1,463,616	1.842	4	0.1682	11.0
5	<a href="#"><u>Summit, 2018, USA</u></a> 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	<a href="#"><u>Selene, 2020, USA</u></a> 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	<a href="#"><u>Perlmutter, 2021, USA</u></a> 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	<a href="#"><u>JUWELS Booster Module, 2020, Germany</u></a> 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	<a href="#"><u>Adastra, 2022, France</u></a> GENCI-CINES	HPE Cray EX235a, AMD Optimized 3 <sup>rd</sup> Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	319,072	0.303	11	0.0461	6.6
10	<a href="#"><u>Satonix -- GPU, 2022, Australia</u></a> Pawsey Supercomputing Centre	HPE Cray EX235a, AMD Optimized 3 <sup>rd</sup> Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	181,248	0.175	15	0.0272	6.4
14	<a href="#"><u>Wisteria/BDEC-01 (Odyssey), 2021, Japan</u></a> ITC, University of Tokyo	Fujitsu PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu interconnect D	368,640	0.10	23	22,121	4.5



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

Shared File System (SFS)  
25.8 PB, 500 GB/s

Simulation Nodes:

**Odyssey**

Fujitsu/Arm A64FX  
25.9PF, 7.8 PB/s

2.0 TB/s

Data/Learning Nodes:

**Aquarius**

Intel Ice Lake + NVIDIA A100  
7.20 PF, 578.2 TB/s

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

800 Gbps

External Resources



External Network



External Resources



Simulation Nodes (Odyssey)



Data/Learning Nodes (Aquarius)



東京大学  
THE UNIVERSITY OF TOKYO



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

### 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
- 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
- 25 PF, #39 in 58<sup>th</sup> 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, #129 in 60<sup>th</sup> TOP500 (November 2022)



### 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, #17 in 58<sup>th</sup> 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)



Reedbush



Oakforest-PACS



Oakbridge-CX

# Oakbridge-CX (OBCX)

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

- 富士通製
- Intel Xeon Platinum 8280 (Cascade Lake, CLX) (28コア) × 2
  - 合計1,368ノード, Omni-Path Architecture (OPA)
- 共有ファイルシステム (Lustre)
- 128ノードはSSD搭載, 総容量200TBの高速ファイルシステムとして運用可
  - SSD搭載128ノードのうち16ノードは外部ネットワークに直接接続しており (外部接続ノード), 外部リソース (サーバ, ストレージ, センターネットワーク等) との通信可能
- ピーク性能6.61 PFLOPS, TOP500で69位 (2020年11月)
- 更に外部ネットワークに直接接続し, OBCX各計算ノードと連携するGPUサーバ (通称Mini-DP, Wisteria/BDEC-01の「データ・学習ノード」に相当) を利用可能

# Oakbridge-CX (OBCX) 世界第129位 2022年11月現在



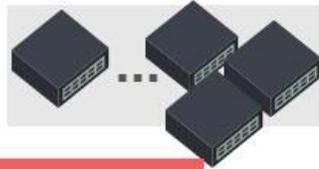
Fujitsu PRIMERGY CX2550 M5



Fujitsu PRIMERGY CX400 M1  
シャーシ当たりCX2550 M5 × 4搭載

## 計算ノード

Chassis: PRIEMRGY CX400 M1 x342 <4node / Chassis>  
Node: PRIMERGY CX2550 M5 x1,240, CX2560 M5 x128



x1,368 node



## 全体性能

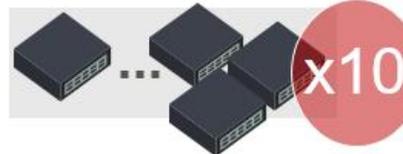
理論演算性能: 6.61PF  
主記憶容量: 256.5TiB  
メモリバンド幅: 385.1TB/s  
ラック数: 21ラック  
SSD搭載: 128ノード

## ノード単体

理論演算性能: 4.8384 TF  
主記憶容量: 192GiB  
メモリバンド幅: 281.6GB/s

計算ノード間ネットワーク (Omni-Path Architecture)  
通信性能 100Gbps

## ログインノード



x10

FUJITSU Server  
PRIMERGY CX2560 M5 x 10

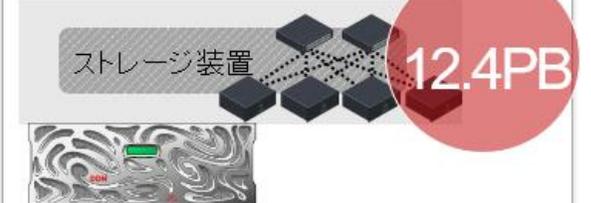
## 管理サーバ群



x15

FUJITSU Server  
PRIMERGY RX2530 M4 x 15  
(ジョブ、運用、認証、Web、  
セキュリティログ保存)

## 並列ファイルシステム



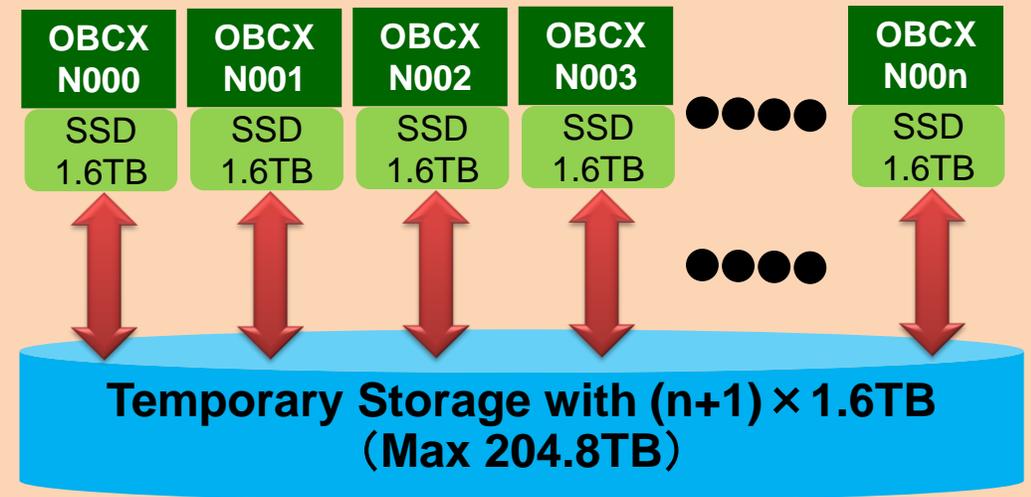
12.4PB

ストレージ装置: DDN ES18K x2セット  
ファイルシステム: DDN ExaScaler  
(Lustreベースファイルシステム)

# Oakbridge-CX (OBCX) : BDECに向けた実験システム

- 全1,368ノードのうち128ノードにSSD (Solid State Drive) 搭載
  - Intel SSD + BeeGFS
    - 容量 : 1.6 TB/node
    - 読み書き性能 : 3.20/1.32 GB/s/node
    - BeeOND (BeeGFS-on-Demand) によって合計 200+TB (128 × 1.6) の高速ファイルシステムとして使用可能
  - データ科学アプリケーション
    - ソフトウェア類も充実
  - ステージング, チェックポイント
  - 128ノードのうち16ノードはSINET経由で外部リソース (サーバー, ストレージ, センサーネットワーク) に直接接続⇒外部接続ノード

## BeeGFS on Demand (BeeOND)



Total: 1,368 nodes

128 nodes  
with SSD

16

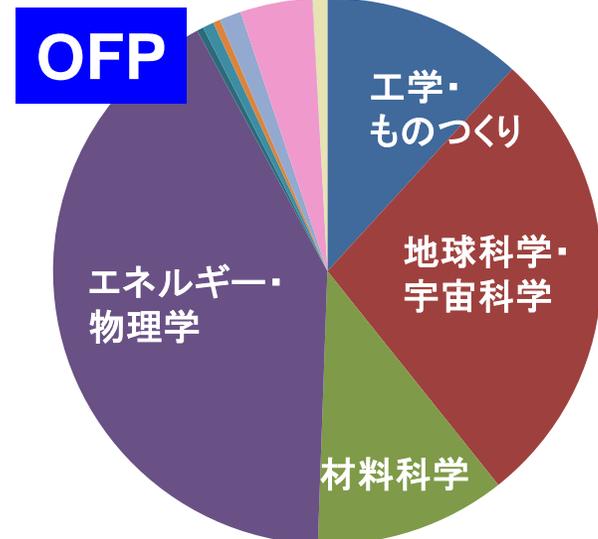
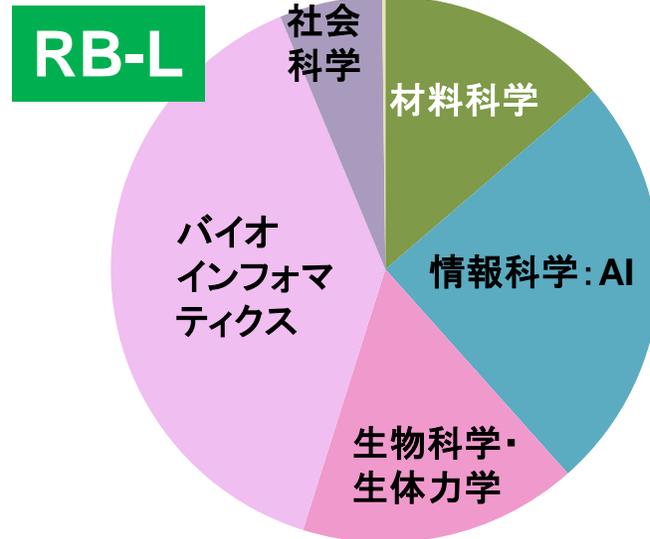
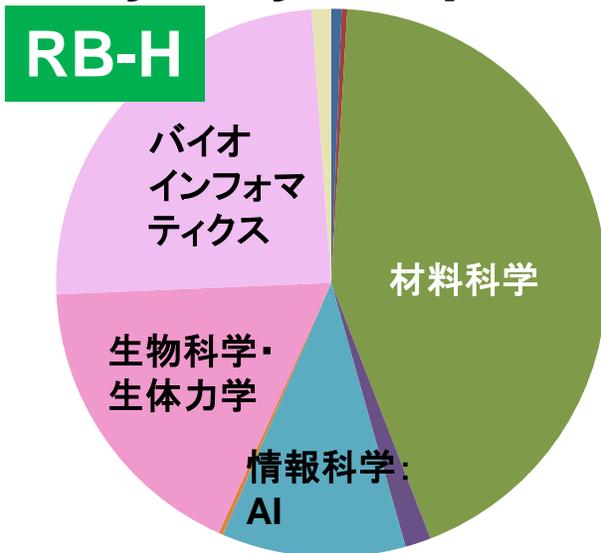
**OBCXの16ノード(外部接続ノード)**  
SINET経由で外部計算機資源に直接接続,  
BDECにおけるデータ・学習ノード群と同様の  
役割



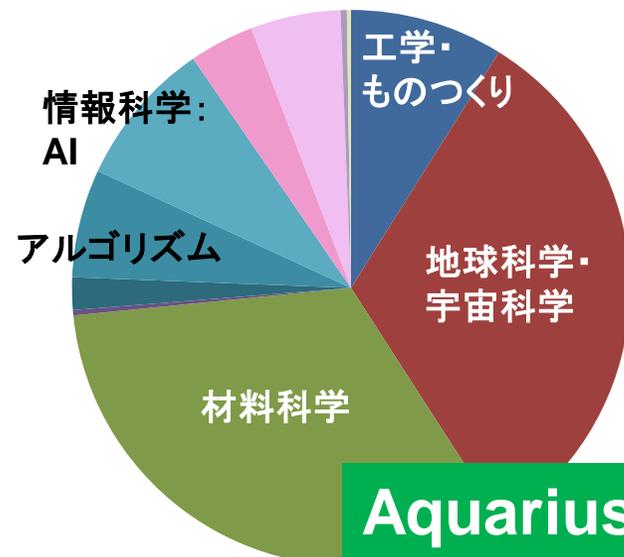
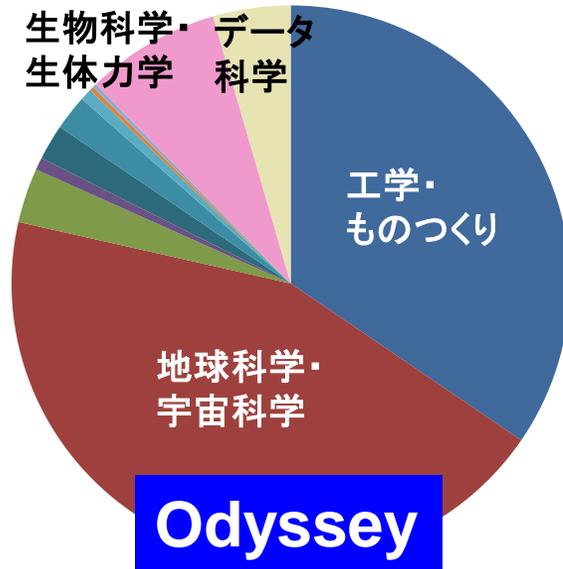
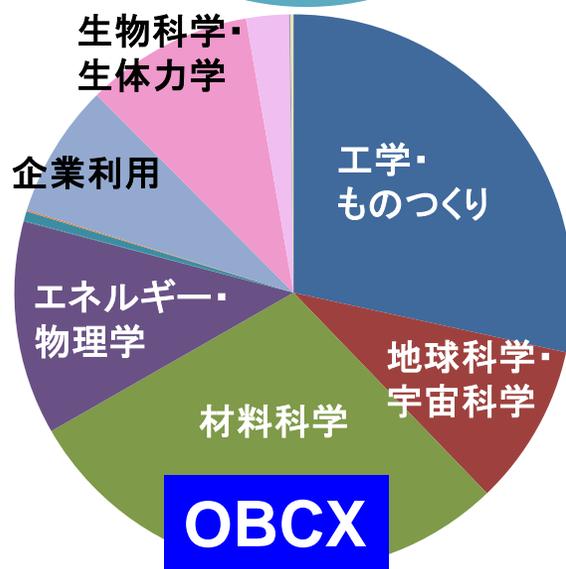
# 2021年度分野別 (12月末時点) ■汎用CPU,

## ■GPU

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



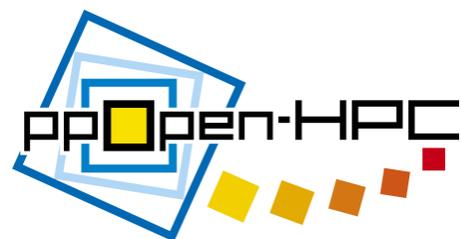
- 工学・ものづくり
- 地球科学・宇宙科学
- 材料科学
- エネルギー・物理学
- 情報科学: システム
- 情報科学: アルゴリズム
- 情報科学: AI
- 教育
- 産業利用
- 生物科学・生体力学
- バイオインフォマティクス
- 社会科学・経済学
- データ科学・データ同化



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

# スパコン利用にあたっての指針 (1/2)

## OBCX, Odyssey, Aquarius



- 基本的には、自作コード、オープンソースの利用を前提

- OpenFOAM (流体)

- OBCX, Odyssey

- 今野雅博士 (客員研究員) : OpenFOAM関連チュートリアル

- FrontISTR, FrontFlow, ABINIT (東大生研)

- ppOpen-HPC, h3-Open-BDEC (東大センター)



- 商用コード

- Altair HyperWorks (汎用CAEコード)

- <https://www.altairjp.co.jp/hyperworks/>

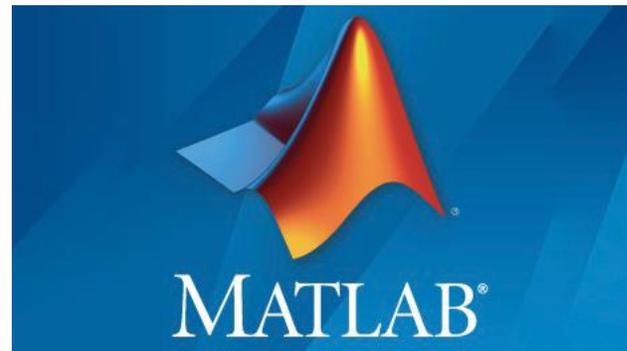
- OBCX, Aquarius (一部)

- 国内大学教職員・学生のみ利用可能

- 研究機関, 企業の場合は別途ライセンス取得が必要

- MATLAB (2022年3月から利用可能)

- OBCX, Aquarius



# スパコン利用にあたっての指針 (2/2)

## OBCX, Odyssey, Aquarius

	OBCX 1,368ノード	Odyssey 7,680ノード	Aquarius 45ノード, 360GPU	O+A
計算科学 (S)	◎	◎	◎	-
データ科学 (D)	◎	◎	◎	-
機械学習・AI (L)	○	○	◎	-
大規模計算	○	◎	○	◎
「S+D+L」融合	○	○	◎	◎
商用コード	◎	△	○	-
特徴	<ul style="list-style-type: none"><li>Intel Xeon CPUのため特殊なチューニングは不要</li><li>外部接続ノード (128) SSD搭載ノード (16)</li></ul>	<ul style="list-style-type: none"><li>富岳と同じA64FX搭載, 高いメモリ性能</li><li>超大規模シミュレーション可</li><li>FP16容易に利用可</li><li>チューニング必須</li></ul>	<ul style="list-style-type: none"><li>高性能GPU (NVIDIA A100) 搭載</li><li>外部接続可能</li></ul>	<ul style="list-style-type: none"><li>O-A連携ソフトウェア開発中 (h3-Open-BDEC, WaitIO)</li></ul>
こんな用途に	<ul style="list-style-type: none"><li>研究室内のマシン等での計算の大規模化</li></ul>	<ul style="list-style-type: none"><li>大規模計算, FEM・差分法等によるシミュレーション</li></ul>	<ul style="list-style-type: none"><li>機械学習, 深層学習</li><li>粒子法シミュレーション</li></ul>	<ul style="list-style-type: none"><li>(S+D+L) 融合</li></ul>

# GFLOPS (ピーク性能) 当たり利用負担 (円) : 電気代 GFLOPS/W (Green 500)

System	JPY/GFLOPS Small is Good	GFLOPS/W Large is Good
Oakleaf-FX/Oakbridge-FX (Fujitsu) (Fujitsu SPARC64 IXfx)	125	0.866
Reedbush-U (HPE) (Intel Xeon Broadwell (BDW))	61.9	2.310
Reedbush-H (HPE) (Intel BDW+NVIDIA P100x2/node)	15.9	8.575
Reedbush-L (HPE) (Intel BDW+NVIDIA P100x4/node)	13.4	10.167
Oakforest-PACS (Fujitsu) (Intel Xeon Phi/KNL)	16.5	4.986
Oakbridge-CX (Fujitsu) (Intel Xeon Cascade Lake)	<del>20.7</del> → 31.05	5.076
<b>Wisteria-Odyssey (Fujitsu/Arm A64FX)</b>	<del>17.8</del> → <b>26.7</b>	<b>15.07</b>
<b>Wisteria-Aquarius (Intel Xeon Ice Lake + NVIDIA A100x8)</b>	<del>9.00</del> → <b>13.5</b>	<b>24.06</b>

# 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

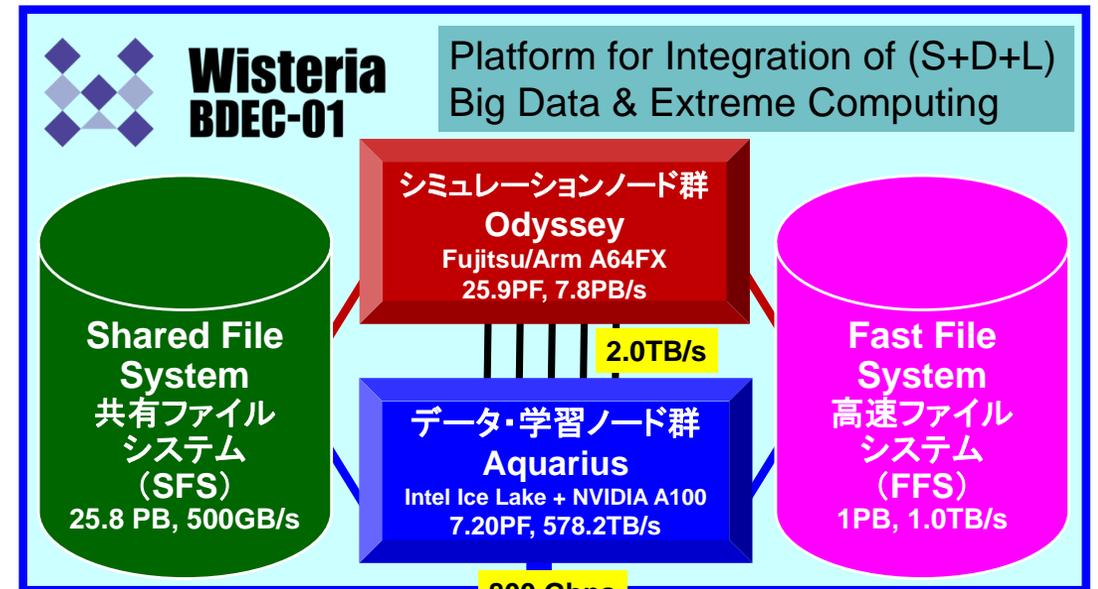
- 一般利用

- 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
  
- 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（1/19）の資料
  - <https://www.cc.u-tokyo.ac.jp/events/seminar/20230119.php>

# 大規模共通ストレージシステム「Ipomoea」

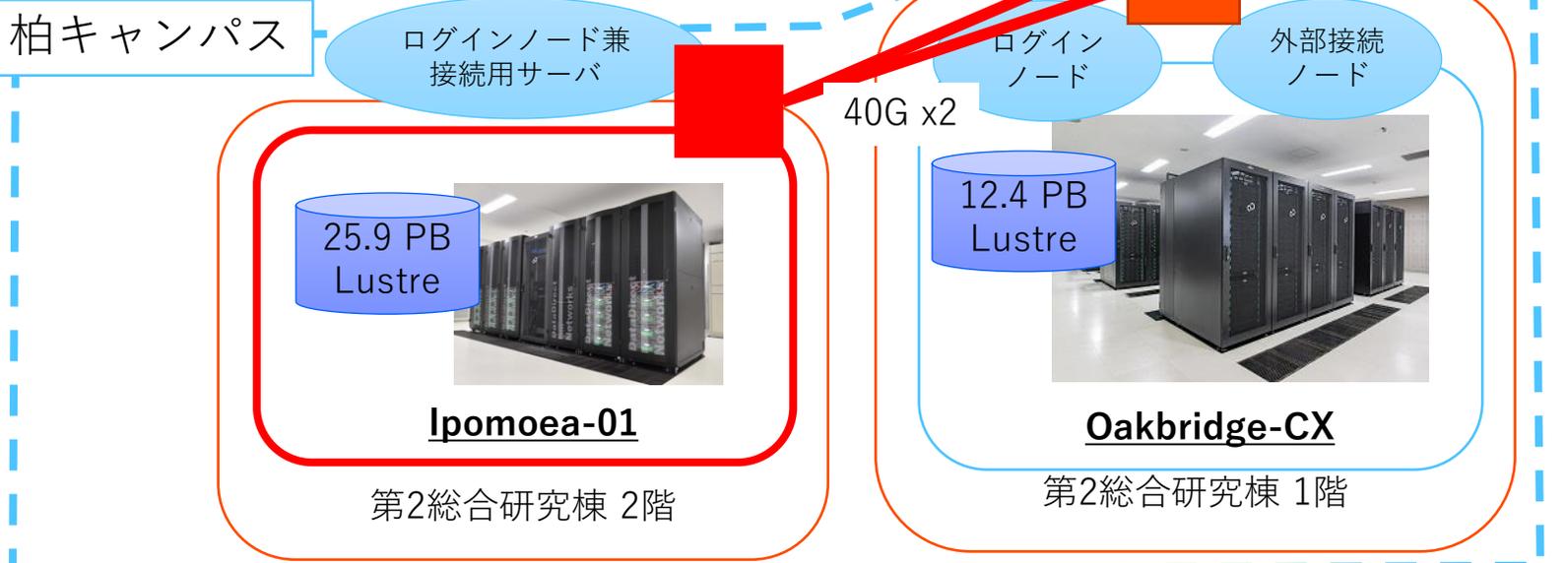
- スーパーコンピュータの処理能力の向上に伴い，扱うデータ量も増加の一途
- 東大センターでは従来ストレージは各システムに附属して導入され，各システムのストレージは独立
- このような状況（注：ストレージがシステム毎に独立）は利用者に多大な不便を強いることになり，東大センターの全システムからアクセス可能な共通ストレージの導入が強く求めら
- **各システムからアクセスできる「大規模共通ストレージ（Ipomoea）」導入決定**
  - OFP運用終了が契機
  - 1システムを約5–6年使用し，約3年ごとに新しいストレージシステム（25+PB）を導入し，入れ替えることを想定している



# Ipomoea-01

- 2022年1月運用開始, 25+PB
  - 富士通製
- FP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

