

# JHPCNから「計算・データ・学習」融合を目指す 学際大規模情報基盤共同利用・共同研究拠点公募説明会 @東大



**Wisteria  
BDEC-01**

東京大学情報基盤センター  
スーパーコンピューティング研究部門  
<http://www.cc.u-tokyo.ac.jp/>  
問合せ先: [uketsuke@cc.u-tokyo.ac.jp](mailto:uketsuke@cc.u-tokyo.ac.jp)

# 本日の趣旨



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

- 東京大学情報基盤センターのスーパーコンピュータ群
- 学際大規模情報基盤共同利用共同研究拠点 (JHPCN) 公募課題
- 質疑
  
- 現在運用中のシステム
  - Oakbridge-CX (2023年9月末運用終了)
  - Wisteria/BDEC-01 (2021年5月運用開始)

# 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
- 2022年11月現在
  - 2式のスパコンシステムを運用
    - Oakbridge-CX(OBCX): 2023年9月末運用終了
    - 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
  - 大規模共通ストレージ「Ipomoea-01」: 2022年1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



2001-2005

2006-2010

2011-2015

2016-2020

2021-2025

2026-2030

Hitachi SR8000  
1,024 GFHitachi SR11000  
J1, J2  
5.35 TF, 18.8 TFHitachi SR16K/M1  
Yayoi  
54.9 TFHitachi  
SR2201  
307.2GFHitachi  
SR8000/MPP  
2,073.6 GFOBCX  
(Fujitsu)  
6.61 PFHitachi HA8000  
T2K Today  
140 TFOakforest-  
PACS (Fujitsu)  
25.0 PFOFP-II  
200+ PFFujitsu FX10  
Oakleaf-FX  
1.13 PFWisteria  
BDEC-01 Fujitsu  
33.1 PFBDEC-  
02  
250+ PFReedbush-  
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  
1.024 PF  
SR8000

Hitachi SR11000  
J1, J2  
IBM Power5+

Hitachi SR16K/M1  
Yayoi  
IBM Power7

Hitachi SR2201  
HARP-1E

Hitachi SR8000/MPP  
SR8000

Intel CLX  
OBCX (Fujitsu)  
6.61 PF

Hitachi HA8000  
T2K Today  
140 TF

Oakforest-PACS (Fujitsu)  
35.0 PF

OFP-II  
200+ 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

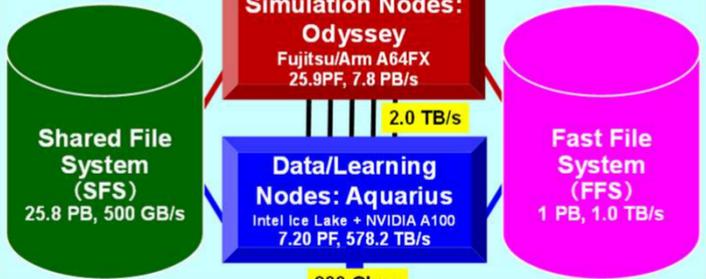
Ipomoea-01 25PB

Ipomoea-03

Ipomoea-02



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



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 (Reedbush-H/L)

### 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 (#23)**
- データ・学習ノード群 (Aquarius) : Intel Icelake + NVIDIA A100 (#125)**
- 33.1 PF, 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

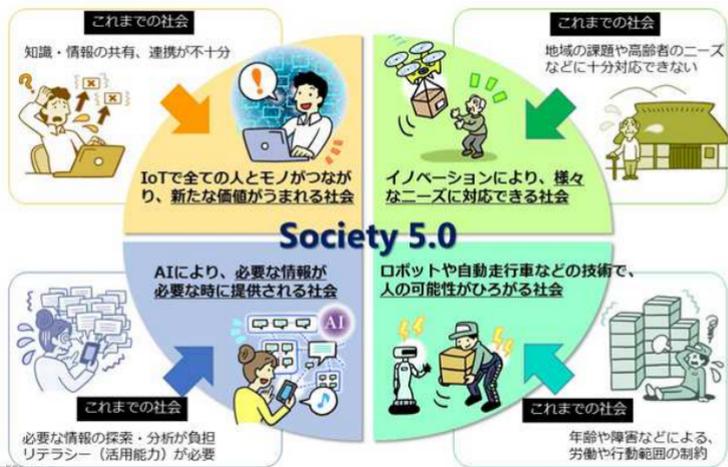


Oakbridge-CX

# mdx データ活用社会創成プラットフォーム

超スマート社会 Society 5.0 : AI, IoT, ビッグデータなどの革新技术を社会全体に活用し、サイバー空間（仮想）とフィジカル空間（現実）を高度に融合させた社会

持続可能な開発目標(SDGs) : 経済発展と社会的課題解決の両立



用途に応じてオンデマンドで短時間に**構築・拡張・融合**できるデータ収集・集積・解析機能を提供するプラットフォーム



## 3本柱

### 1 高性能計算環境によるデータ科学と計算科学の融合

データ科学、計算科学の手法を融合し、さらに国内最高の計算環境を用いて他に無い高精度の予測を行えるようにする

### 2 SINETを活かしたリアルタイム収集・集積・解析環境の動的な構築

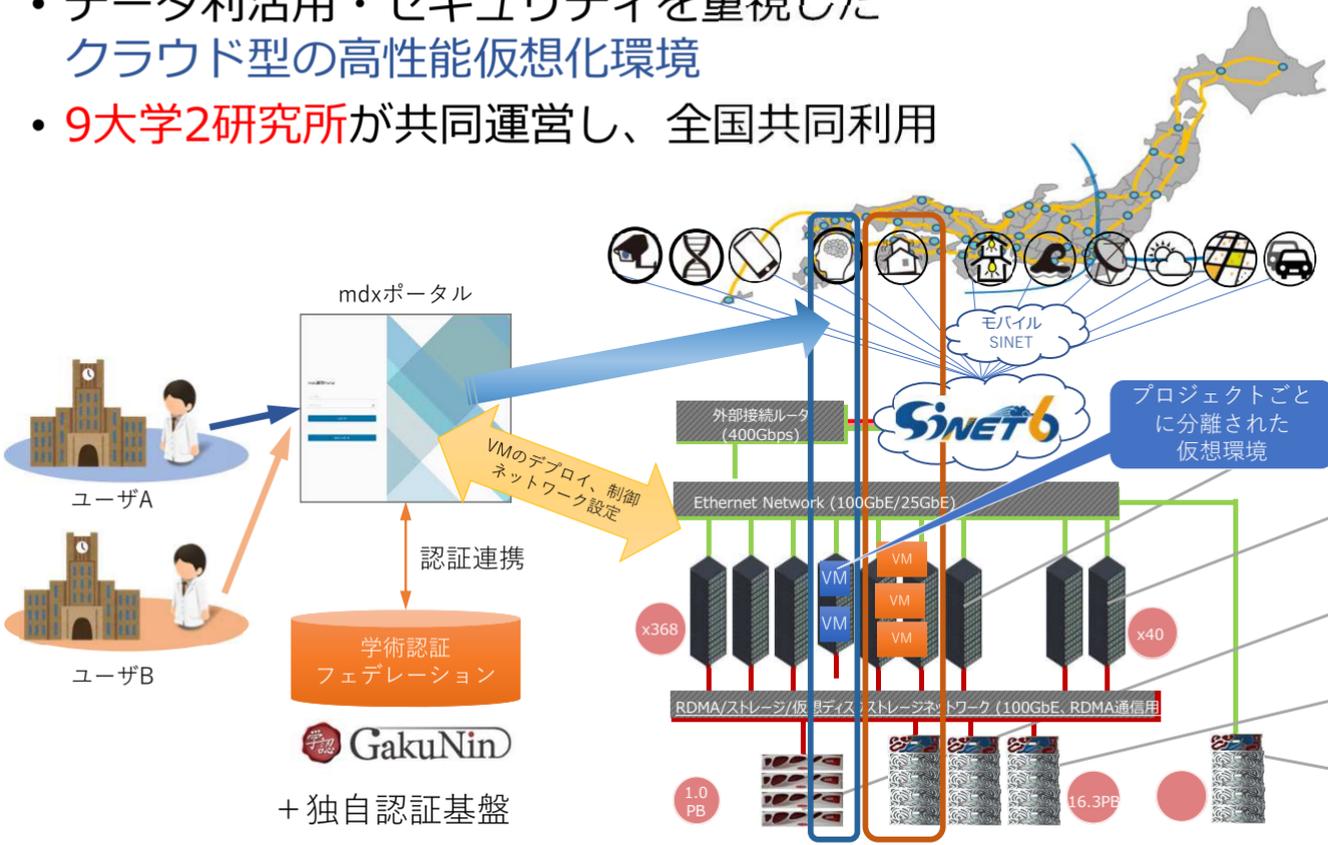
遠隔地のセンサーやストレージ、データプラットフォームの計算資源、ストレージをつないで、リアルタイムに入力から出力を得られるアプリケーションごとの収集・集積・解析環境を、使いたいときに即時に構築するSINETモバイル基盤によりセンサー等のデータを安定してセキュアにつなぐ

### 3 異種データ・異種知識の融合活用の推進

様々な分野のデータ保持者、解析者、利用者が産学にまたがって連携するコミュニティーを形成し、新たな価値創造につなげる。

# mdx データ活用社会創成プラットフォーム

- データ利活用・セキュリティを重視したクラウド型の高性能仮想化環境
- 9大学2研究所が共同運営し、全国共同利用



**ネットワーク**  
 12つのネットワーク 外部接続ネットワーク  
 SINET6と400G x2で接続  
 内部高速ネットワーク RDMA  
 ストレージ

**汎用CPUノード**  
 Intel IceLake x2ソケット x368ノード  
 理論ピーク性能 (FP64): 2.1PFLOPS  
 総メモリバンド幅: 150.7 TB/s

**GPU 演算加速ノード**  
 Intel IceLake x2ソケット+NVIDIA A100 x8 x40ノード  
 理論ピーク性能 (FP64): 6.4PFLOPS  
 理論ピーク性能 (FP16): 100.7PFLOPS  
 総メモリバンド幅: 496.3 TB/s

**高速 NVMe ストレージ**  
 Lustre Filesystem  
 1.0 PB (NVMe SSD)  
 252 GByte/sec

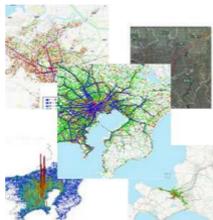
**大容量HDDストレージ**  
 Lustre Filesystem  
 16.3 PB (HDD)  
 157.5 GByte/sec

**外部共有オブジェクトストレージ**  
 S3 Data Service  
 10.3 PB (HDD)  
 63.0 GByte/sec

GakuNin  
 + 独自認証基盤

# mdx データ活用社会創成プラットフォーム

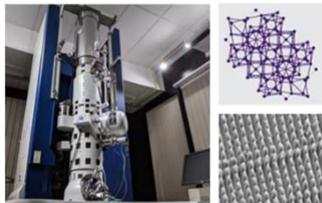
- 2021年秋より運用を開始し、70以上のプロジェクトがmdxを利用
- プロジェクトの分野は、地理空間情報、材料科学、カーボンニュートラル、経済、地球科学、自然言語処理、機械学習、深層学習など多岐にわたる



地理空間情報



データ科学技術



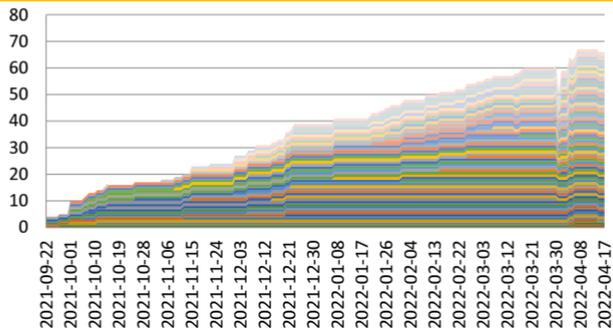
ARIM Japan

材料科学 x 情報科学



脱炭素社会に向けた情報基盤

## mdx におけるプロジェクト数の推移



## mdx を活用しているJHPCN採択課題の一例

- 財務ビッグデータの可視化と統計モデリング
- 医療・介護領域の人材マッチングに最適化された大規模グラフニューラルネットワーク
- エージェントモデルと統計データを用いた全国規模の疑似人流データの開発
- 大規模な日本語モデル構築・共有のためのプラットフォームの形成
- グラフニューラルネットワークとマルチタスク学習による汎用的物性予測モデルの構築
- ビヨンド・"ゼロカーボン"を目指し地域と技術をつなぐ情報基盤の構築
- 単語間に区切りのない書写言語における係り受け解析エンジンの開発
- 多次元高精度地表情報（MHESD）の地球科学・歴史考古学における高度利活用

# スパコン利用諸制度 (○:代表者, △:参加者)

制度名	種別	大学等	企業	学生	個人	審査	無料	報告書	A	B	C	D	備考	募集
通常利用	一般	○	△	○					✓	✓	✓			随時
	トライアル	○	△	○				✓			✓	✓	年度内	随時
お試し利用		○	○	○	✓		✓				✓	✓	1ヶ月限定	随時
JHPCN		○	○	△		書類	✓	✓		✓				年1回(1月)
HPCI	一般・若手	○	△	△		書類	✓	✓		✓				年1回(10-11月)
	産業		○			書類	✓	✓		✓				
若手女性	一般	○	○	○	✓	書類	✓	✓		✓	✓			年2回(8・2月)
	インターン			○	✓	書類	✓	✓			✓			年1回(夏季)
AI for HPC		○	○	△		書類	✓	✓		✓	✓			随時, 年4回審査
HPCチャレンジ		○	○	○		書類	✓	✓						年数回
講習会		△	△	△	✓		✓						1ヶ月有効UID	年20回程度
教育利用		○	○	○		書類	✓	✓					企業研修等可	随時
企業利用	一般	△	○	△		+面接		✓		✓				年2回(8・2月)
	トライアル	△	○	△		+面接	一部	✓			✓	✓	3ヶ月無料 年度内	随時, 年4回審査

(A:トークン移行, B:ノード固定, C:Odyssey⇔Aquarius移行可能, D:1システム1回限り応募可能)

- JHPCNの概要
- 東大情報基盤センターのスパコン概要
- Wisteria/BDEC-01
- Oakbridge-CX
- Ipomoea-01
- h3-Open-BDEC
  
- 利用事例
- JHPCN利用申込み

# 学際大規模情報基盤共同利用共同研究拠点(JHPCN)

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

- 大規模スーパーコンピュータを有する8国立大学センターから構成されるネットワーク型拠点
  - 北海道, 東北, 東大(中核拠点), 東工大, 名古屋, 京都, 大阪, 九州
- 文部科学省「共同利用・共同研究拠点」として認可され, 2010年4月から活動開始(6年に一回見直し(+3年))
  - 東大:地震研, 大気海洋研, 物性研など
- 学際的な共同研究課題の推進
  - 計算科学+計算機科学
  - 各センタースパコン及び関連設備の利用(無料)
- 2016年度以降は一般課題に加えて, 国際, 産業, 萌芽の3カテゴリー
  - 萌芽は各センター独立:本学「若手・女性」, 「AI-for-HPC」は「萌芽」の一つ
- 2022年度から第3フェーズ
  - 枠組みが変化(計算科学, データ利活用)(+ mdx)



# JHPCNの目的

- グランドチャレンジ的な問題について、学際的な共同利用・共同研究を実施することにより、我が国の学術・研究基盤の更なる高度化と恒常的な発展に資する
  - 超大規模計算機, 超大容量ストレージ, 超大容量ネットワークなどの情報基盤
  - 地球環境, エネルギー, 物質材料, ゲノム情報, Webデータ, 学術情報, センサーネットワークからの時系列データ, 映像データ, プログラム解析, その他
- JHPCNには、上記の各分野における多数の先導的研究者が在籍し、これらの研究者との共同研究によって、研究テーマの一層の発展が期待できる。



# 共同研究分野(2021年度まで)

- 超大規模数値計算系応用分野
  - 地球環境, エネルギー, 物質材料等の科学・工学分野における科学技術シミュレーション, 関連するモデリング, 数値解析アルゴリズム, 可視化手法, 情報基盤等
- 超大規模データ処理系応用分野
  - ゲノム情報, Webデータ, 学術情報コンテンツ, センサーネットワークからの時系列データ, 映像を始めとするストリームデータなどに対する高度なメディア情報処理, プログラム解析, アクセスや検索, 情報抽出, 統計的・意味的分析、データマイニング
- 超大容量ネットワーク技術分野
  - 超大規模データ共有のためのネットワーク品質の確保や制御, 超大規模ネットワーク自体の構築と運用に必要な監視や管理, そのようなネットワークの安全性の評価と維持確保、およびこれら諸技術の研究開発の支援のための技術等
- 上記研究分野を統合した超大規模情報システム関連研究分野
  - ペタスケール／エクサスケールコンピュータのアーキテクチャ, 高性能計算基盤ソフトウェア, グリッド, 仮想化技術, クラウドコンピューティング等

# 共同研究分野(2022年度から)

- 分野1:大規模計算科学分野:2021年度までのJHPCN
  - 本日は(主として)これについて説明する
  - 超大規模数値計算系応用分野
  - 超大規模データ処理系応用分野
  - 超大容量ネットワーク技術分野
  - 超大規模情報システム関連研究分野
- 分野2:データ科学・データ利活用分野:2022年度からの新分野
  - mdxを中心に利用するような分野
- 分野1にも「大規模データ処理系応用分野」があるが、分野2は人文・社会系も含むより広い範囲が対象、分野1の「データ処理系」は科学技術計算に関わるデータ処理、データ解析、その手法の開発、が中心

# 共同研究のポイント: 全部を満たす, ということではない

- 学際的研究体制
  - 情報科学分野(計算機科学・データ科学)と応用分野による, 学際共同研究の推進
  - 審査においては学際的な研究体制の構築がなされていることを重視
- ソフトウェアおよびデータ活用推進
  - 単なる公開だけでなく, 成果の幅広い利活用を目指す取り組みを高く評価
- IT 基盤技術開発
  - アーキテクチャー, システムソフトウェア, セキュリティ等, IT 技術の基盤的研究につながる課題を高く評価: 各構成研究拠点の研究者との共同研究
- 拠点連携
  - 複数構成拠点の資源を活用, 複数拠点の研究者との連携体制
  - 広域分散型の大規模情報システム, アプリケーションのマルチプラットフォーム実装
- 大規模データ・大容量ネットワーク利用
- 計算機資源の利用のみではなく, 研究開発的要素が大きい課題を採択

# 共同研究の分類

- 一般共同研究課題(採択課題数全体の8割程度)
- 国際共同研究課題(採択課題数全体の1割程度)
  - 国際共同研究課題は, 国内の研究者のみでは解決や解明が困難である問題について, 国外の研究者と学際的な共同研究を実施する
  - 国外の共同研究者との打ち合わせ等の出張旅費を助成する制度がある
    - 1名・2週間程度の滞在可能
- 企業共同研究課題(採択課題数全体の1割程度)
  - 産業応用を重視した学際的な共同研究
- 各課題とも, 大学・研究機関, 民間企業からの参加が可能
- (分野1, 分野2), (HPCI(スパコン利用), 非HPCI)という組み合わせがあるが, ここでは(分野1・HPCI)に限定

# 応募資格

- 課題代表者
  - 日本国内の機関(大学・研究機関, 民間企業等)に所属する者
- 学生参加
  - 大学院生のみ, 課題代表者, 副代表者にはなれない
- 非居住者の参加(国際共同研究課題含む)
  - 利用する計算機を運用している構成拠点に所属している研究者が共同研究者として参加する必要がある⇒東大のスパコンを使う場合は必ず東大センターのメンバーを含む(ご相談ください)(規定★)
- 国際共同研究課題
  - 副代表者として日本国外機関所属研究者を1名以上含む必要がある
  - 英文申込書による応募を必須, 規定★を満たす
- 企業共同研究課題
  - 課題代表者は必ず民間企業に所属する研究者
  - 副代表者の1名以上は規定★を満たす

# スパコンの「利用資格」(1/3) : 「通常利用」

- 東京大学情報基盤センタースーパーコンピュータシステム利用規程
  - <https://www.cc.u-tokyo.ac.jp/guide/files/riyoukitei.pdf>
- 原則, 国内の教育機関(高専, 大学, 大学院)の教職員・学生, 公的研究機関の研究者を対象: 専ら研究開発に従事する人々(企業の一部の人々も含む)
- 企業の技術者, 研究者, 国外の研究者(非居住者)等については一部制限がある場合があるが, 一定の条件を満たせば使用可能
  - 利用制度によって異なる(ここでは「通常利用」を中心に簡単に説明)
  - 特に非居住者の利用については外国為替及び外国貿易法で厳密に規定されているが, 本センターの非居住者利用ルールは比較的フレキシブル⇒本学安全保障輸出管理支援室, 経済産業省と密接に協力して制定
- 企業勤務者(科研費等の交付を受けて専ら研究開発に従事する人々を除く)
  - 「企業利用」: 申請ベース, 利用者資格審査委員会の審査(後述)
  - HPCI, JHPCN等の研究グループ所属者は利用可能
  - 「通常利用」グループのメンバーと共同研究, 役務契約の実体があれば当該グループのメンバーとなることは可能(別途書類提出必要あり), グループ代表にはなれない

# スパコンの「利用資格」(2/3):「通常利用」

参考

- 非居住者・外国籍利用者
  - 非居住者(海外の機関に所属(日本国籍の場合も含む)する研究者, 来日6ヶ月未満の留学生)には利用制限(マニュアル閲覧制限等)あり
    - 非居住者は利用グループ(一般利用, JHPCN, 企業利用等)の代表者にはなれない
    - ホワイト国と非ホワイト国で区別あり
    - 来日6ヶ月未満の留学生:安全保障関連手続き等がクリアされれば講義等でのアクセスは可能
  - 国内機関に雇用されている場合は, 外国籍でも居住者扱い(制限無し)
  - 留学生も来日6ヶ月経過していれば原則居住者扱い(制限無し)
    - 安全保障関連手続きは必要
  - HPCI, JHPCN等の研究グループ所属者は非居住者も利用可能(一部制限あり)
  - 「通常利用」グループのメンバーと共同研究, 役務契約の実体があれば当該グループのメンバーとなることは可能(別途書類提出必要あり, 一部制限あり)
  - 機関同士(できれば部局同士)の共同研究協約・MOUが結ばれていればOK
- 企業勤務者, 非居住者は通常利用グループの代表者にはなれないが, 一定の条件を満たせば通常利用グループのメンバーになることは可能

# スパコンの「利用資格」(3/3)

- ご不明な点をご遠慮なくお問い合わせください
  - uketsuke(at)cc.u-tokyo.ac.jp
- 特に「通常利用」については、原則として、手続きさえ踏めば、「利用できない」ということはまずありません。
- 「JHPCN」の場合は更にゆるい
  - 非居住者
  - 企業の方
  - 提案して採択されれば、共同研究の実態がある、と本学では判断している([利用規程](#))
    - 第3条(8)項
    - 追加の書類, 審査等は不要

# みなし輸出管理(2022年5月より)

<https://www.cc.u-tokyo.ac.jp/guide/application/non-resident.php>

- 外国為替及び外国貿易法の関連法令(以下「外為法等」)が改正され, 2022年5月1日からみなし輸出管理の対象が拡大
  - 外為法等が対象とする貨物の輸出及び技術の提供
  - 国境を超えた貨物輸出, 技術提供の他, 国内であっても, 居住者から非居住者への技術の提供が「輸出とみなさ」れ, 外為法による輸出管理の対象となる(みなし輸出管理)
- 改正外為法等では, 「居住者⇒居住者」の技術提供についても, 技術提供を受ける居住者が外為法等に規定される外国政府・外国法人等から強い影響を受けている状態に該当(以下「類型該当」)する場合, みなし輸出管理の対象
  - 改正外為法等を遵守し, みなし輸出管理を徹底するために, 個々のスーパーコンピュータ利用者に類型に該当するか否かについて大学に対して自己申告することを求める
  - 当センターのスーパーコンピュータをご利用の全利用者の方が自己申告を行う必要があることから, UID発行のたびに自己申告フォームより回答(5分程度)をお願いしている

# その他

- 共同研究期間
  - 2023年4月1日～2024年3月31日
- 利用負担金：無料
- 課題審査：書類審査
- 資源量
  - 後述
- 応募方法(分野1・HPCI)
  - <https://jhpcn-kyoten.itc.u-tokyo.ac.jp/ja/cfp>
  - 様式記入, Web経由で登録, 所属機関長(部局長)の連絡先を登録
  - 利用分野(数値計算, データ処理, ネットワーク, 情報システム)指定

# 留意事項

- 採択研究課題の目的にのみ研究資源等を利用すること。
- 平和利用目的の提案であること。
- 人権および利益保護への配慮を行うこと。
- 文部科学省「生命倫理・安全に対する取組」に適合すること。
- 厚生労働省「研究に関する指針について」に適合すること。
- 経済産業省「安全保障貿易管理について」に適合すること。
- 実質的に同一の研究課題と思われる課題(ほぼ同等の研究体制・研究テーマの課題や, 研究対象のみが異なる課題など)は採択しません。

# スケジュール

- **募集要項公開・募集開始**
  - 2022年11月末(以降)
  - <https://jhpcn-kyoten.itc.u-tokyo.ac.jp/>
- 応募
  - 申込み締切:2023年1月6日(金)17:00(厳守)
  - 2023年3月中旬までに結果通知
- 研究実施
  - 開始:2023年4月1日
  - 第15回JHPCN シンポジウム(研究内容紹介):2023年7月上旬
  - 中間報告書:2023年10月中旬
  - 共同研究期間終了:2024年3月31日
  - 最終報告:2024年5月中旬
  - 第16回JHPCN シンポジウム(研究成果報告):2024年7月上旬
- 中間報告書(2ページ程度), 最終報告書提出は必須
- シンポジウム発表必須(2020・2021:完全オンライン, 2022:ハイブリッド)

- JHPCNの概要
- **東大情報基盤センターのスパコン概要**
- Wisteria/BDEC-01
- Oakbridge-CX
- Ipomoea-01
- h3-Open-BDEC
  
- 利用事例
- JHPCN利用申込み

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  
200+ PF

Fujitsu FX10  
Oakleaf-FX  
1.13 PF

Wisteria Fujitsu  
BDEC-01  
33.1 PF

BDEC-02  
250+ PF

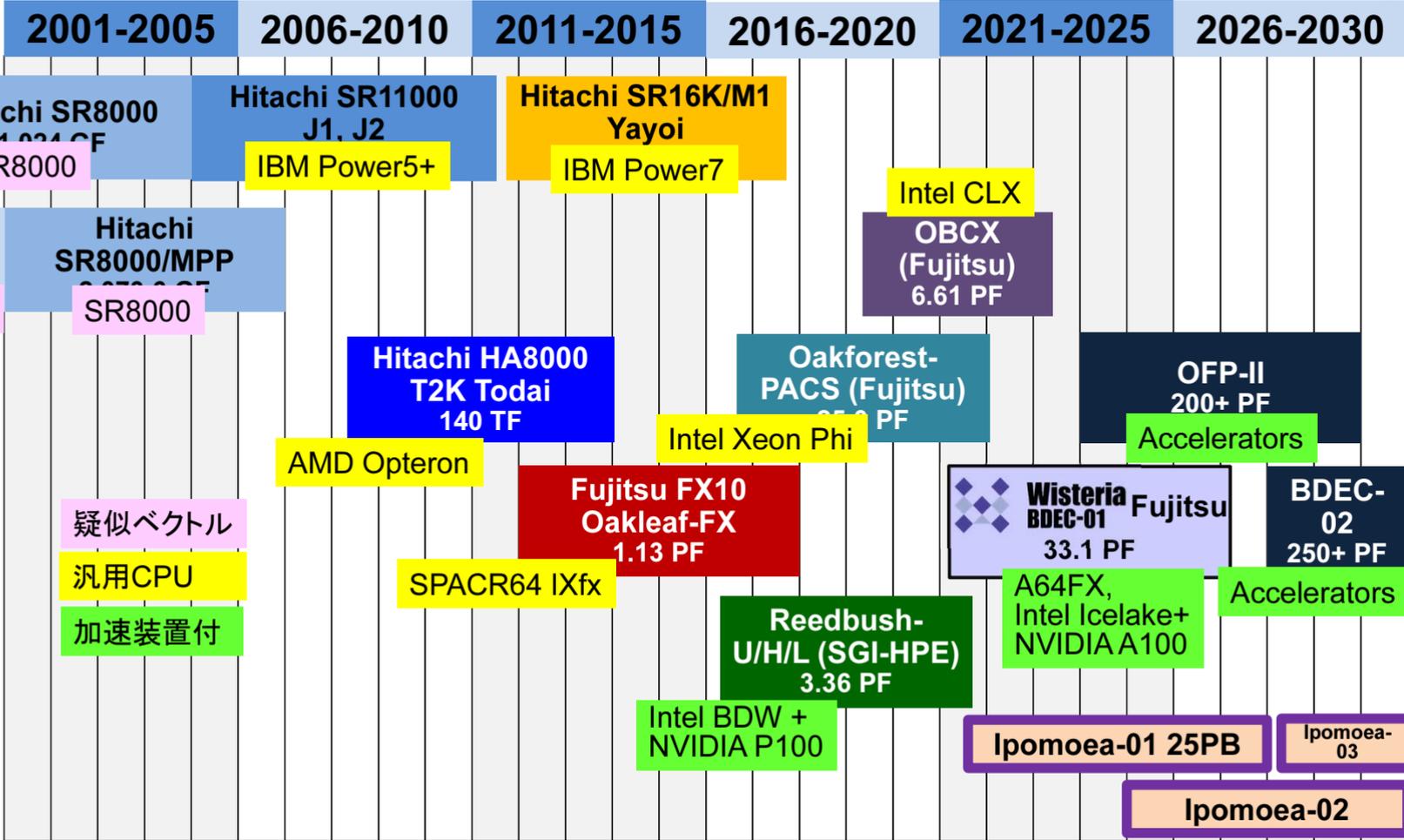
東京大学情報基盤  
センターのスパコン  
利用者2,600+名  
55%は学外

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

Ipomoea-01 25PB

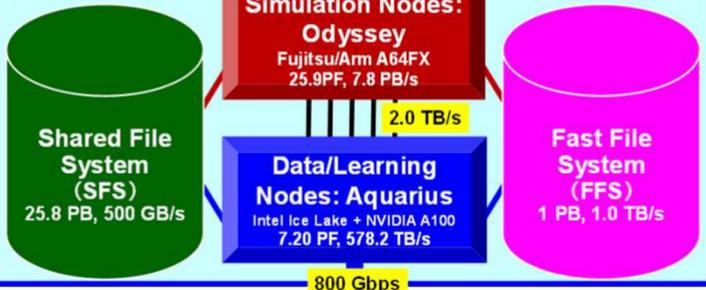
Ipomoea-03

Ipomoea-02





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



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 (Reedbush-H/L)

### 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 (#23)**
- データ・学習ノード群 (Aquarius) : Intel Icelake + NVIDIA A100 (#125)**
- 33.1 PF, 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



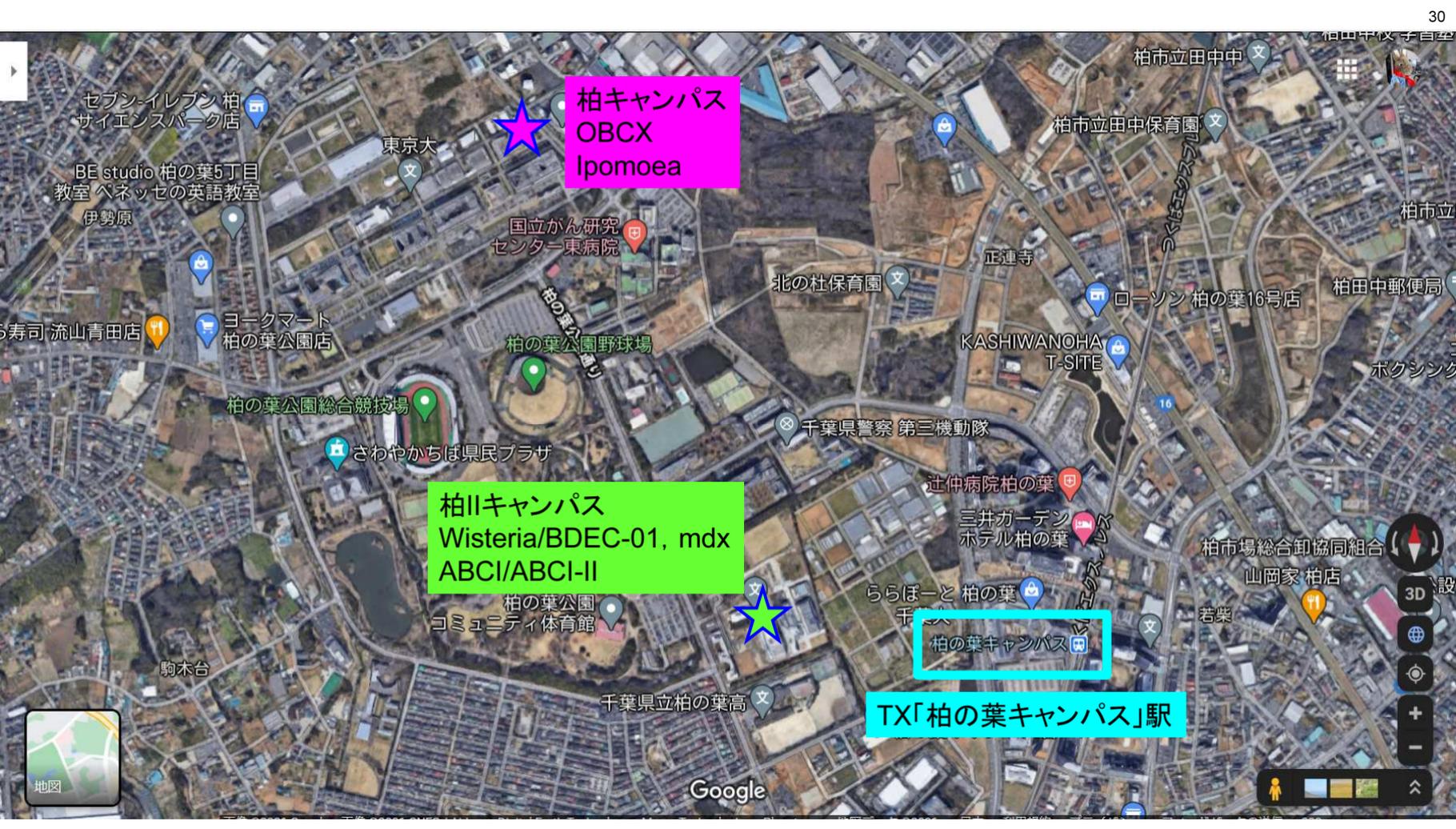
Oakbridge-CX

★ 柏キャンパス  
OBCX  
Ipomoea

★ 柏IIキャンパス  
Wisteria/BDEC-01, mdx  
ABCI/ABCI-II

★ 柏の葉キャンパス

TX「柏の葉キャンパス」駅



# 各種ベンチマーク(1/2)

- TOP 500 (Linpack, HPL)
  - 連立一次方程式ソルバー(直接法), 計算速度(FLOPS値)
  - 規則的な密行列: 連続メモリアクセス
  - 計算性能
- HPCG
  - 連立一次方程式ソルバー(反復法), 計算速度(FLOPS値)
  - 有限要素法から得られる疎行列(ゼロが多い)
    - 不連続メモリアクセス
    - 実アプリケーションに近い
  - メモリアクセス性能, 通信性能
- Green 500
  - HPL (TOP500) 実行時のFLOPS/W値

## 各種ベンチマーク(2/2)

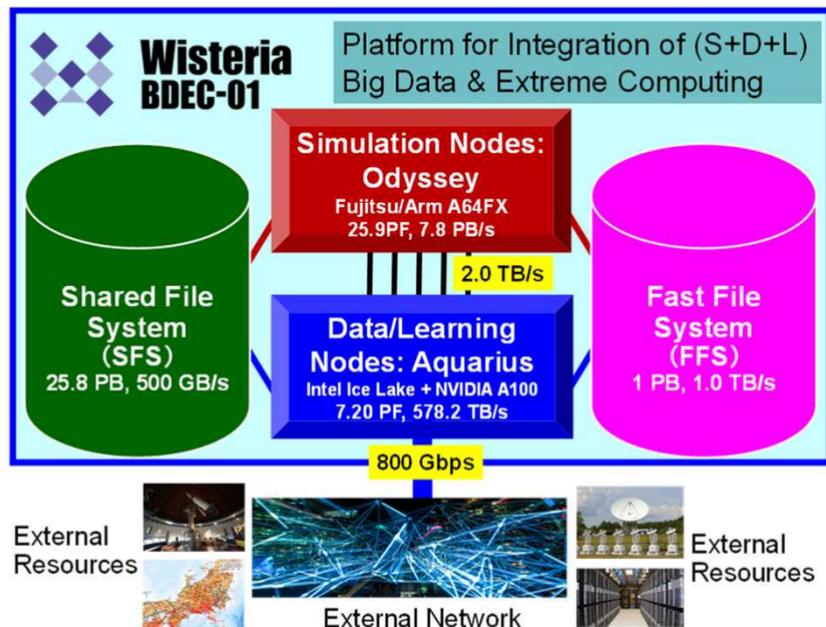
- Graph 500 BFS
  - グラフ処理におけるデータ処理効率
  - BFS(Breadth-First-Search, 幅優先探索)
- HPL-MxP (HPL-AI)
  - AI・機械学習向けHPL(for TOP500)
  - 科学技術計算(HPL)⇒倍精度演算(FP64)
  - HPL-MxP⇒低精度演算を含む混合精度演算

# SC22における諸ランキング (2022年11月)



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

\*) ISC 2022 (June 2022)



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

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

<http://www.top500.org/>
 $R_{\text{peak}}$ : Peak Performance (TFLOPS), Power: kW

	Site	Computer/Year Vendor	Cores	$R_{\max}$ (PFLOPS)	$R_{\text{peak}}$ (PFLOPS)	Power (kW)	
1	<b><u>Frontier, 2022, USA</u></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><u>Fugaku, 2020, Japan</u></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><u>LUMI, 2022, Finland</u></b> EuroHPC/CSC	HPE Cray, EX235a, AMD Optimized 3 <sup>rd</sup> Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	1,110,144	151.90	214.35	2,942	
4	<b><u>Leonardo, 2022, Italy</u></b> EuroHPC/CINECA	Atos, BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Infiniband HDR	2,414,592	148.60	200.79	10,096	
5	<b><u>Summit, 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	148.60	200.79	10,096	
6	<b><u>Sierra, 2018, USA</u></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><u>Sunway TaihuLight, 2016, China</u></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><u>Perlmutter, 2021, USA</u></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	2,528	
9	<b><u>Selene, 2020, USA</u></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><u>Tianhe-2A, 2018, China</u></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><u>ABC1 2.0, 2021, Japan</u></b> AIST	Fujitsu, 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><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	22/22/6/6	22.12	25.95	1,468

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

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

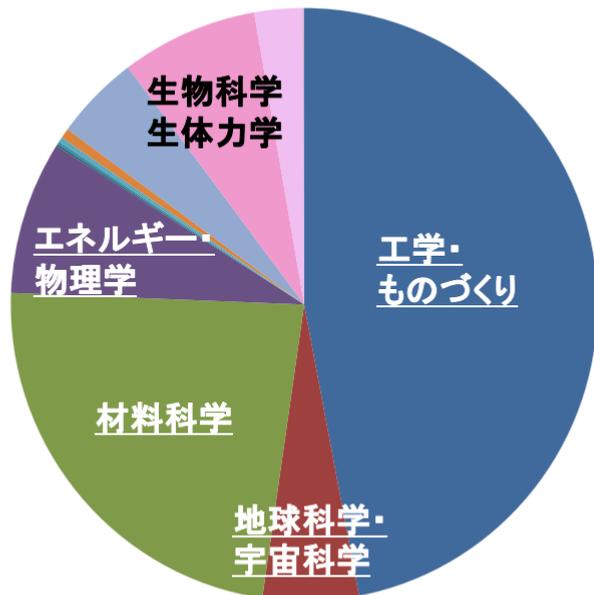
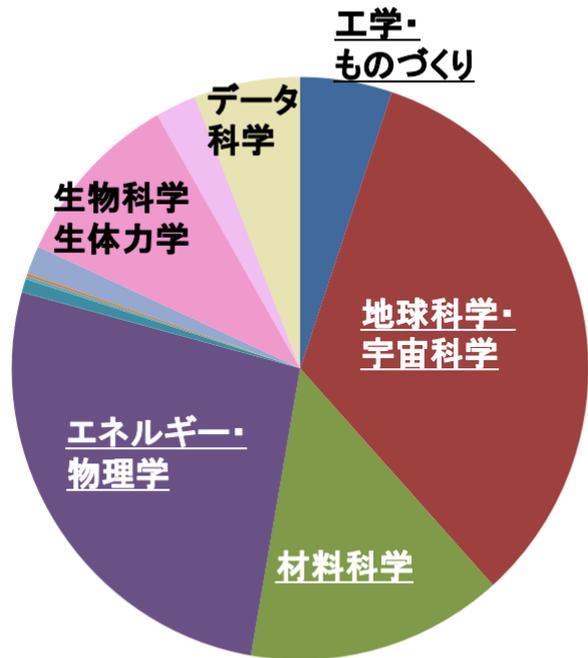
<http://www.top500.org/>
 $R_{\text{peak}}$ : Peak Performance (TFLOPS), Power: kW

	Site	Computer/Year Vendor	Cores	$R_{\max}$ (PFLOPS)	$R_{\text{peak}}$ (PFLOPS)	Power (kW)
1	<b><u>Frontier, 2022, USA</u></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><u>Fugaku, 2020, Japan</u></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><u>LUMI, 2022, Finland</u></b> EuroHPC/CSC	HPE Cray, EX235a, AMD Optimized 3 <sup>rd</sup> Gen. EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	1,110,144	151.90	214.35	2,942
4	<b><u>Leonardo, 2022, Italy</u></b> EuroHPC/CINECA	Atos, BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Infiniband HDR	2,414,592	148.60	200.79	10,096
5	<b><u>Summit, 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	148.60	200.79	10,096
6	<b><u>Sierra, 2018, USA</u></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><u>Sunway TaihuLight, 2016, China</u></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><u>Perlmutter, 2021, USA</u></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	2,528
9	<b><u>Selene, 2020, USA</u></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><u>Tianhe-2A, 2018, China</u></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><u>ABCI 2.0, 2021, Japan</u></b> AIST	<b>日本2位・柏1位 (産業技術総合研究所, 柏IIキャンパス)</b>		22.21	54.34	1,600
23	<b><u>Wisteria/BDEC-01 (Odyssey), 2021, Japan</u></b> ITC, University of Tokyo	<b>日本3位・柏2位</b> FX1000, A64FX 48C 2.2GHz, Tofu	368,640	22.12	25.95	1,468

# GFLOPS (ピーク性能) 当たり利用負担 (円) : 電気代 GFLOPS/W (Green 500) (2023年度から値上げ予定)

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)	20.7	5.076
<b>Wisteria-Odyssey (Fujitsu/Arm A64FX)</b>	17.8	15.069
<b>Wisteria-Aquarius (Intel Xeon Ice Lake + NVIDIA A100x8)</b>	9.00	24.058

# 研究分野別利用CPU時間割合(2020年度)

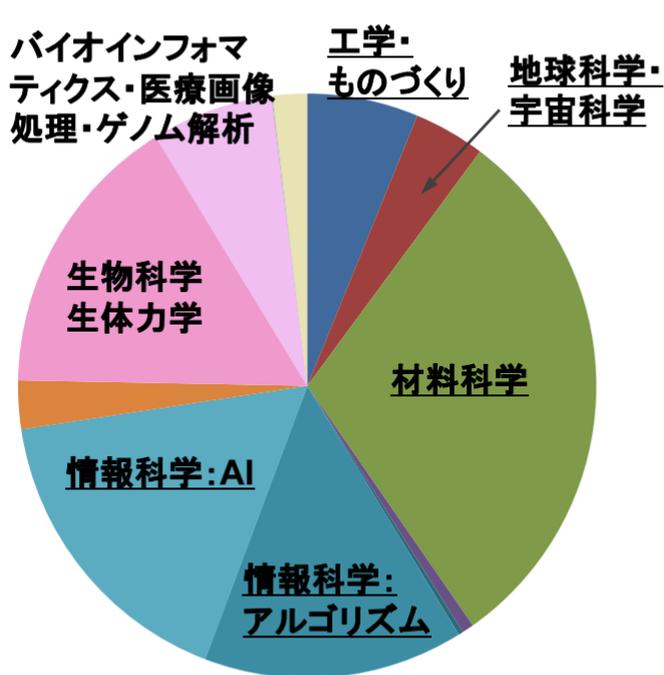


- 工学・ものづくり
- 地球科学・宇宙科学
- 材料科学
- エネルギー・物理学
- 情報科学:システム
- 情報科学:アルゴリズム
- 情報科学:AI
- 教育
- 産業利用
- 生物科学・生体力学
- バイオインフォマティクス
- 社会科学・経済学
- データ科学・データ同化

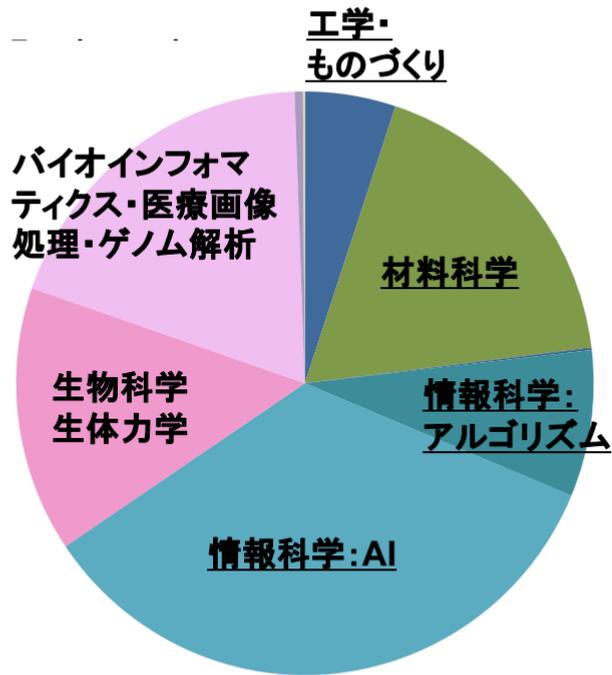
メニコアクラスタ  
Intel Xeon Phi  
(Oakforest-PACS)

マルチコアクラスタ  
Intel CLX  
(Oakbridge-CX)

# 研究分野別利用CPU時間割合(2020年度)



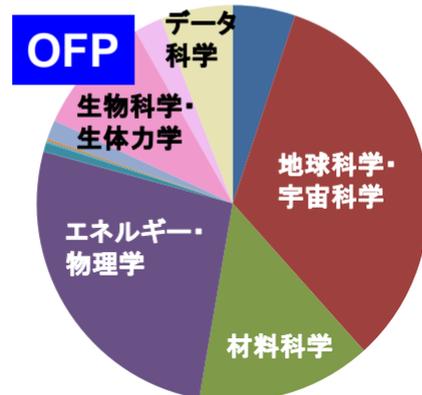
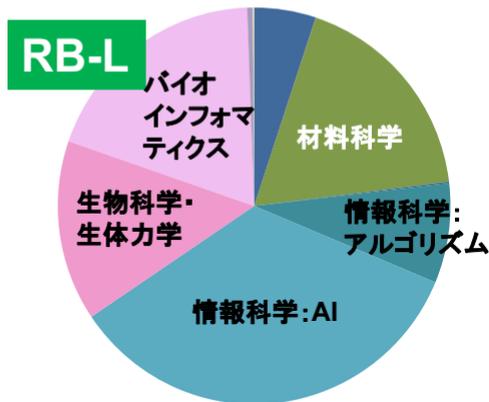
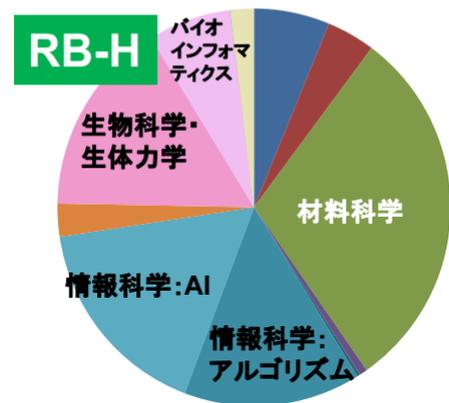
GPUクラスタ (2GPUs/node)  
Intel BDW + NVIDIA P100  
(Reedbush-H)



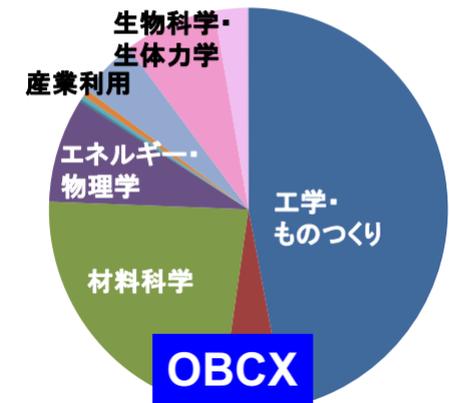
GPUクラスタ (4GPUs/node)  
Intel BDW + NVIDIA P100  
(Reedbush-L)

- 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

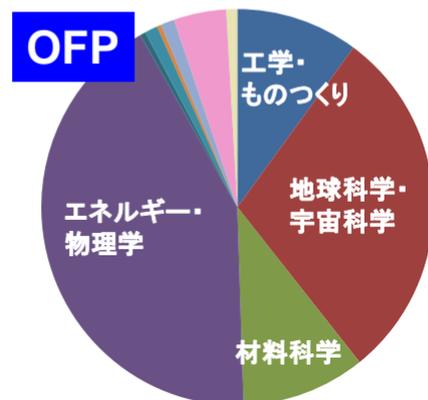
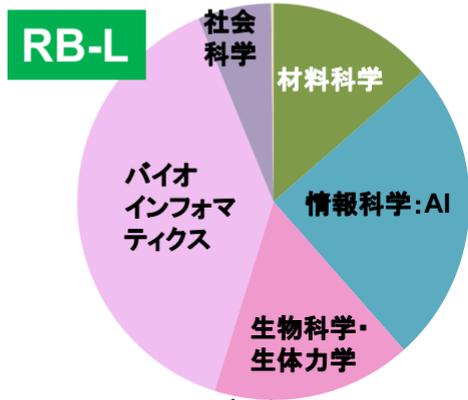
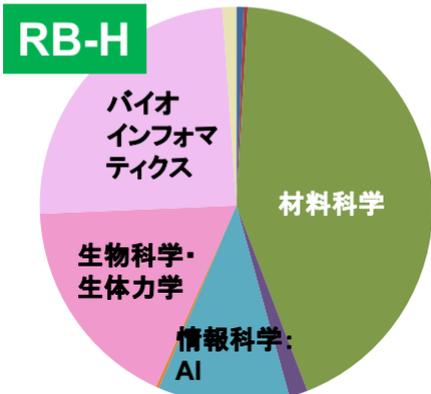


- 工学・ものづくり
- 地球科学・宇宙科学
- 材料科学
- エネルギー・物理学
- 情報科学: システム
- 情報科学: アルゴリズム
- 情報科学: AI
- 教育
- 産業利用
- 生物科学・生体力学
- バイオインフォマティクス
- 社会科学・経済学
- データ科学・データ同化

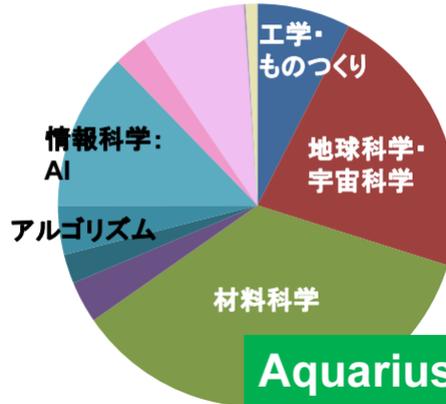
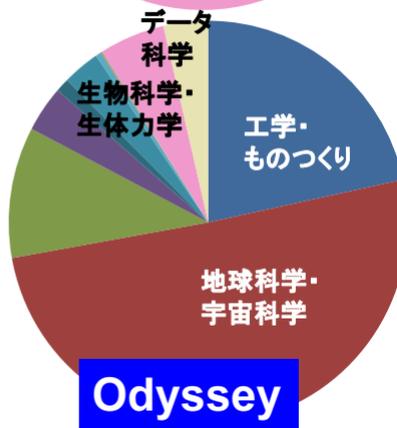
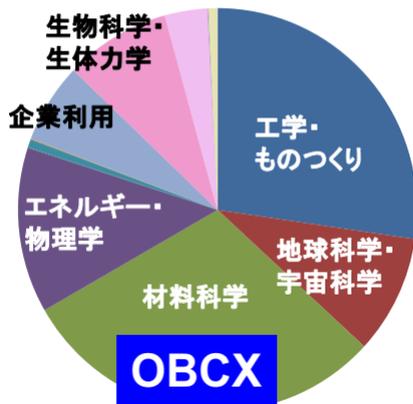


# 2021年度分野別 ■ 汎用CPU, ■ GPU

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



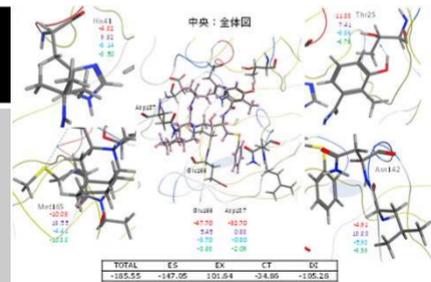
地球科学・宇宙科学分野ではOFP ⇒ Wisteria/BDEC-01への移行が順調に進んでいる

# 「COVID-19」対応HPCI臨時公募課題

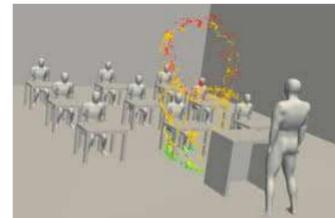
## 全14のうち6課題が東大システムを利用(2020年度)



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

- 柏キャンパス一般公開(2022年度)

[https://www.youtube.com/playlist?list=PLobjSv\\_ny85IZVF5GsMq1aNeltledMgpZ](https://www.youtube.com/playlist?list=PLobjSv_ny85IZVF5GsMq1aNeltledMgpZ)

- 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\\_ny85I-z-VJCy690ZjIAA04xCRA](https://www.youtube.com/playlist?list=PLobjSv_ny85I-z-VJCy690ZjIAA04xCRA)

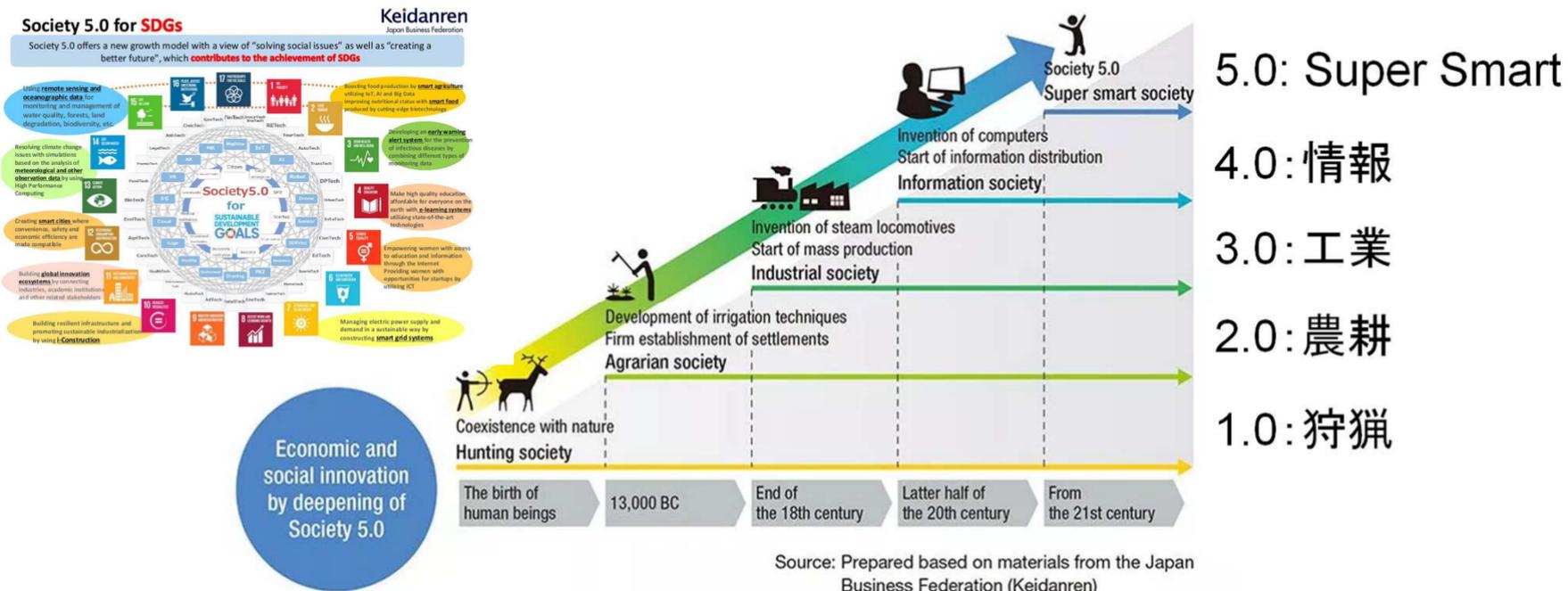
- お試しアカウントつき講習会

[https://www.youtube.com/playlist?list=PLobjSv\\_ny85kXY2Mtnhn1k7pM-epQaD2y](https://www.youtube.com/playlist?list=PLobjSv_ny85kXY2Mtnhn1k7pM-epQaD2y)

- JHPCNの概要
- 東大情報基盤センターのスパコン概要
- **Wisteria/BDEC-01**
- Oakbridge-CX
- Ipomoea-01
- h3-Open-BDEC
  
- 利用事例
- JHPCN利用申込み

# Society 5.0: 日本が提唱する未来社会のコンセプト

デジタル革新・イノベーション (IoT, AI, ビッグデータ等) により, サイバー空間 (仮想) と フィジカル空間 (現実) を高度に融合させたシステムを構築し, 経済発展と社会的課題の解決を両立する, 超スマートな人間中心の社会



# 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

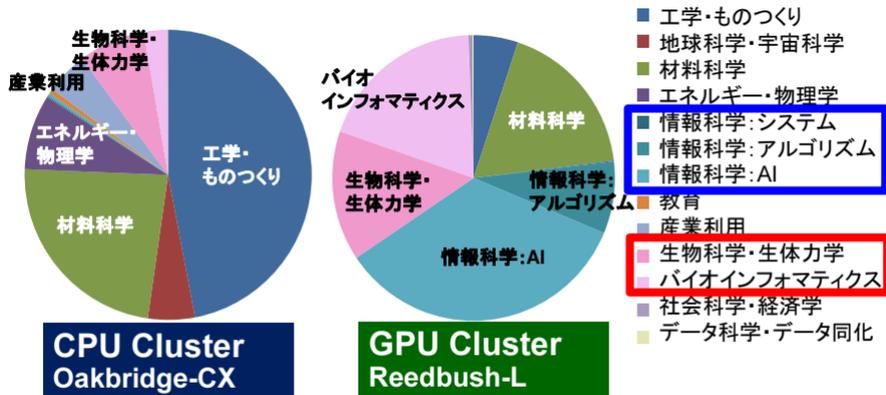
### フィジカル空間とサイバー空間の融合

- 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

## 2021年春に柏IIキャンパスで始動

- BDEC(Wisteria/BDEC-01): 賢いスパコン
- Data Platform(mdx): Cloud的, よりフレキシブル



**BDEC: S + D + L**

**mdx: s + D + L**

# (シミュレーション(計算)+データ+学習)融合(S+D+L)

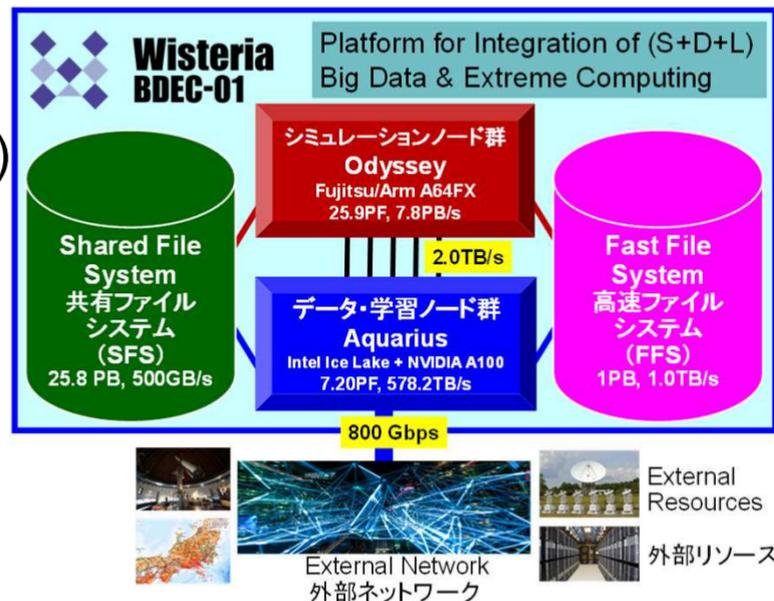
- 東大情報基盤センターでは、2015年頃から「(S+D+L)融合」の重要性に注目し、それを実現するためのハードウェア、ソフトウェア、アプリケーション、アルゴリズムに関する研究開発を開始
  - BDEC計画(Big Data & Extreme Computing)
  - 「データ+学習」による、より高度な「シミュレーション」
    - AI for HPC
  - 地球科学関連では自然な発想(すでに実施されている)
- 2021年5月に運用を開始した「Wisteria/BDEC-01」は「BDEC計画」の1号機
  - Reedbush, Oakbridge-CXは「BDEC」のプロト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), ~360m<sup>2</sup>
- 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**

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>
- Hierarchical, Hybrid, Heterogeneous (h3)
- 2種類のノード群**

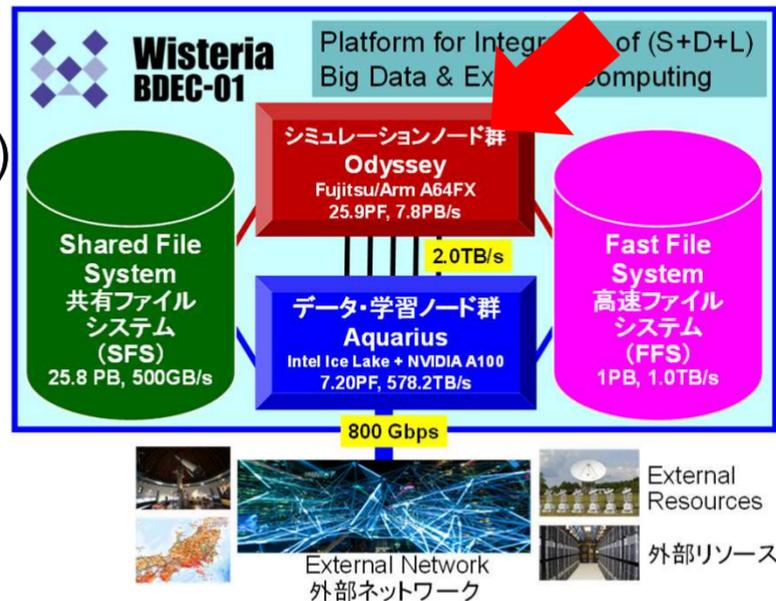
## シミュレーションノード群(S, SIM) : Odyssey

- 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)

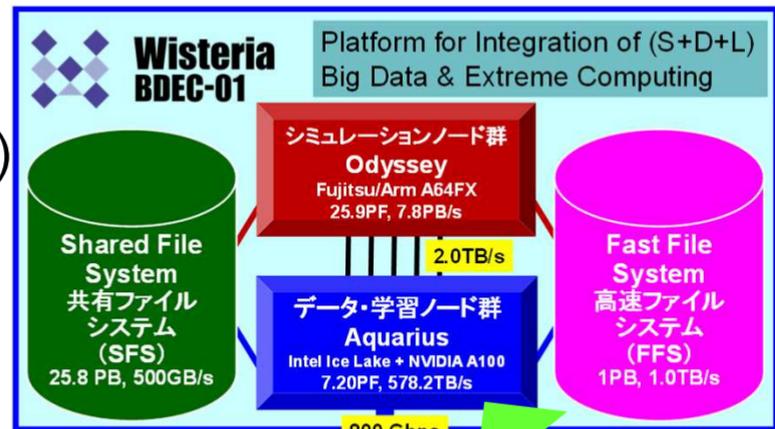


**Wisteria**  
**BDEC-01**

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

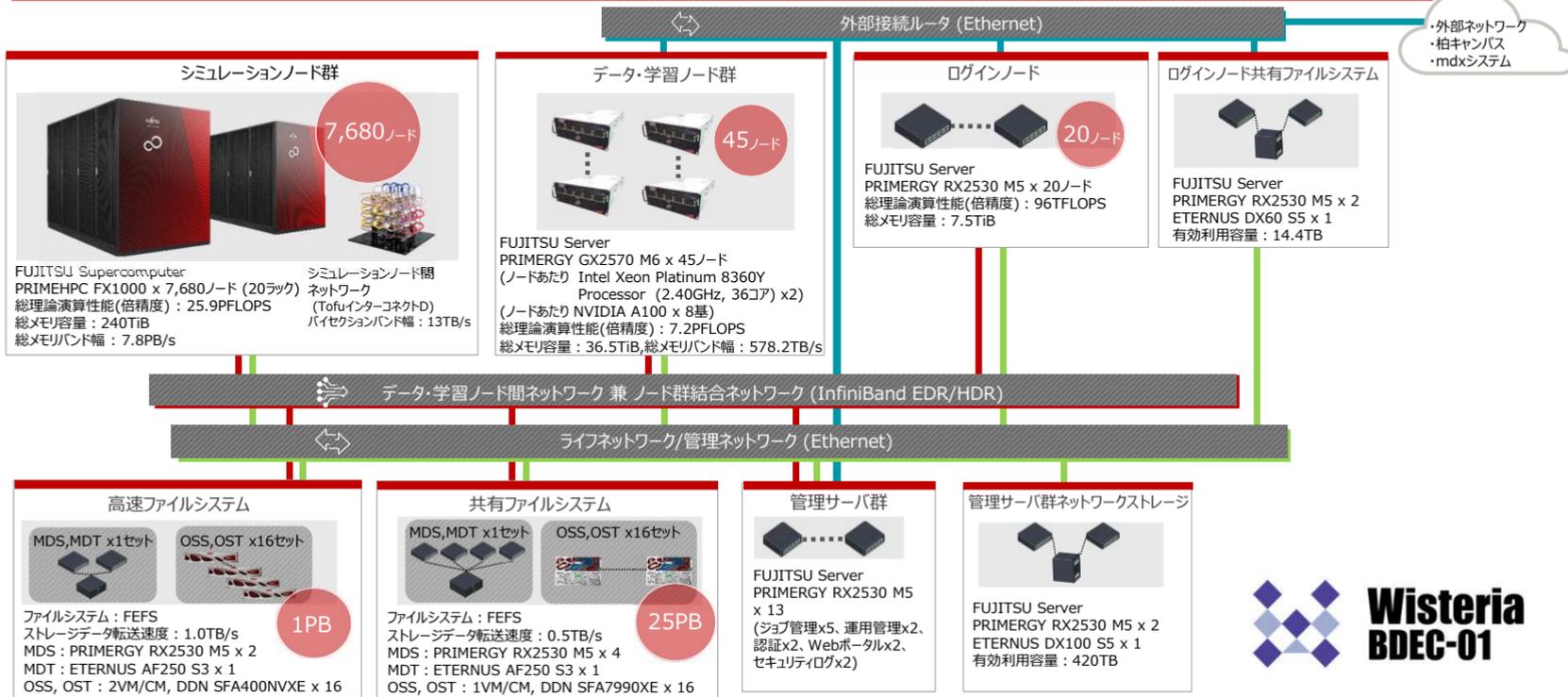
BDEC:「計算・データ・学習(S+D+L)」  
融合のためのプラットフォーム  
(Big Data & Extreme Computing)



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

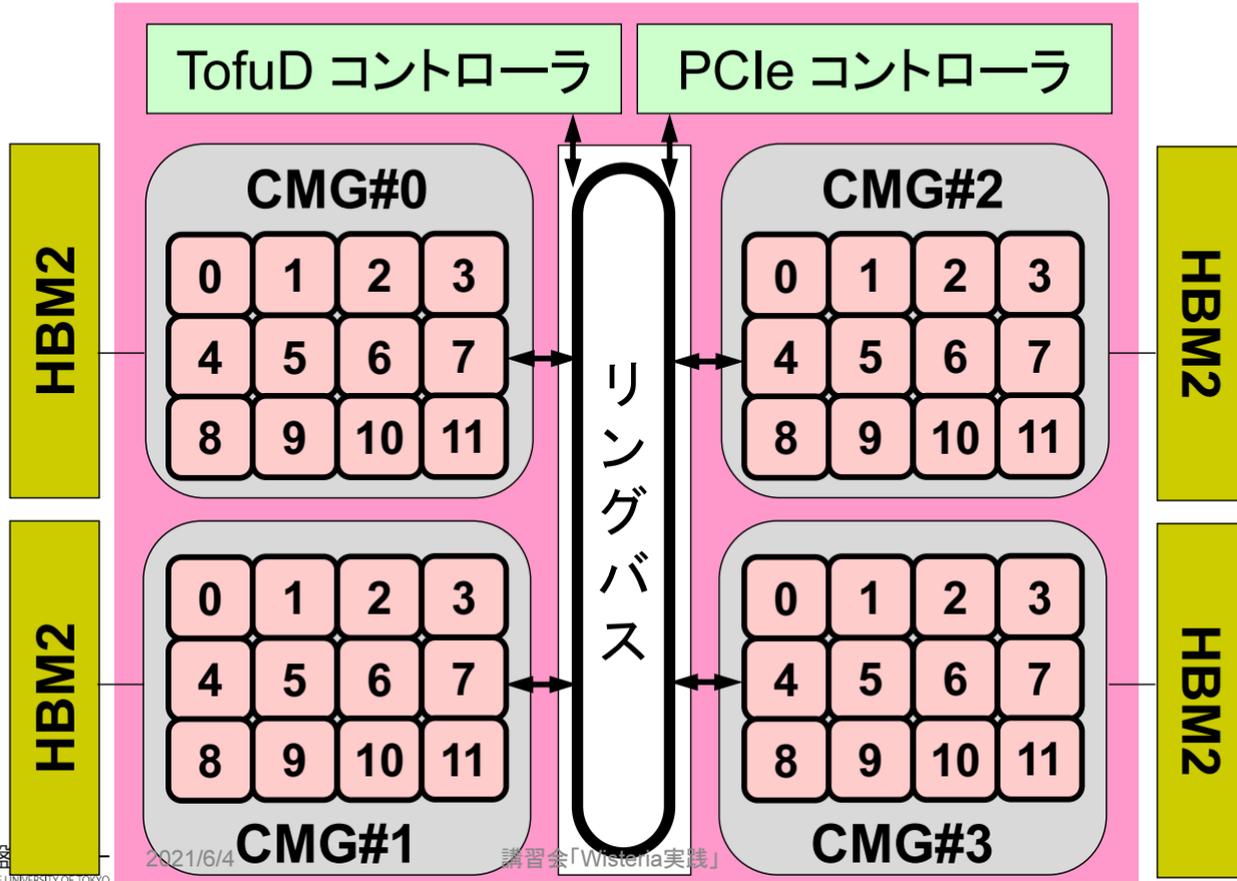
# システム構成図

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)



*The Wisteria/BDEC-01 is a supercomputer system  
operated by the Information Technology Center,  
The University of Tokyo.*

# A64FX : CMG (Core Memory Group)



# Wisteria-OdysseyのTofu-D

- Tofu-D 形状: 全20ラック
    - $(X, Y, Z, a, b, c) = (10, 8, 8, 2, 3, 2)$
  - ➔ 3次元、2次元、1次元に展開して指定可能
- 典型的な例 (全系)

- 3次元:  $20 \times 24 \times 16$ 
  - $X^*a, Y^*b, Z^*c$
- 2次元:  $60 \times 128$ 
  - $X^*a*b, Y^*Z^*c$
- 1次元: 7680

- 参考：
  - ジョブ運用ソフトウェア  
エンドユーザー向けガイド
  - MPI使用手引書

シェルフ内に  
12枚搭載  
シェルフあたりは  
(1, 1, 2, 2, 3, 2)

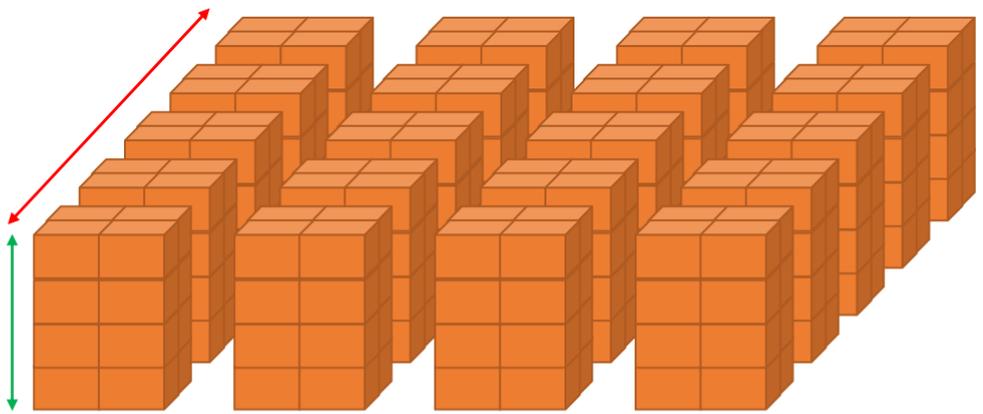


c軸

X軸 (2シェルフ x 5ラック = 10) トーラス

Z軸  
(2x4シェルフ=8)  
トーラス

Y軸 (2シェルフ x 4ラック = 8) メッシュ





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
ネットワークポロジ		6次元メッシュ / トーラス	Full-bisection Fat Tree
インターコネク		TofuインターコネクD	InfiniBand HDR(200Gbps) x 4
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)

# 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)、cuBLAS、cuSPARSE、cuFFT、MAGMA、cuDNN、NCCL
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、tclsh、tcl、vim、zsh、git など	
		Globus Toolkit、Gfarm、FUSE
コンテナ仮想化	Singularity Community Edition	

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

計算科学コード

シミュレーション  
ノード群, Odyssey

最適化されたモデル,  
パラメータ

計算結果

Wisteria/BDEC-01

機械学習, DDA

データ・学習ノード群  
Aquarius

観測データ

データ同化  
データ解析



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

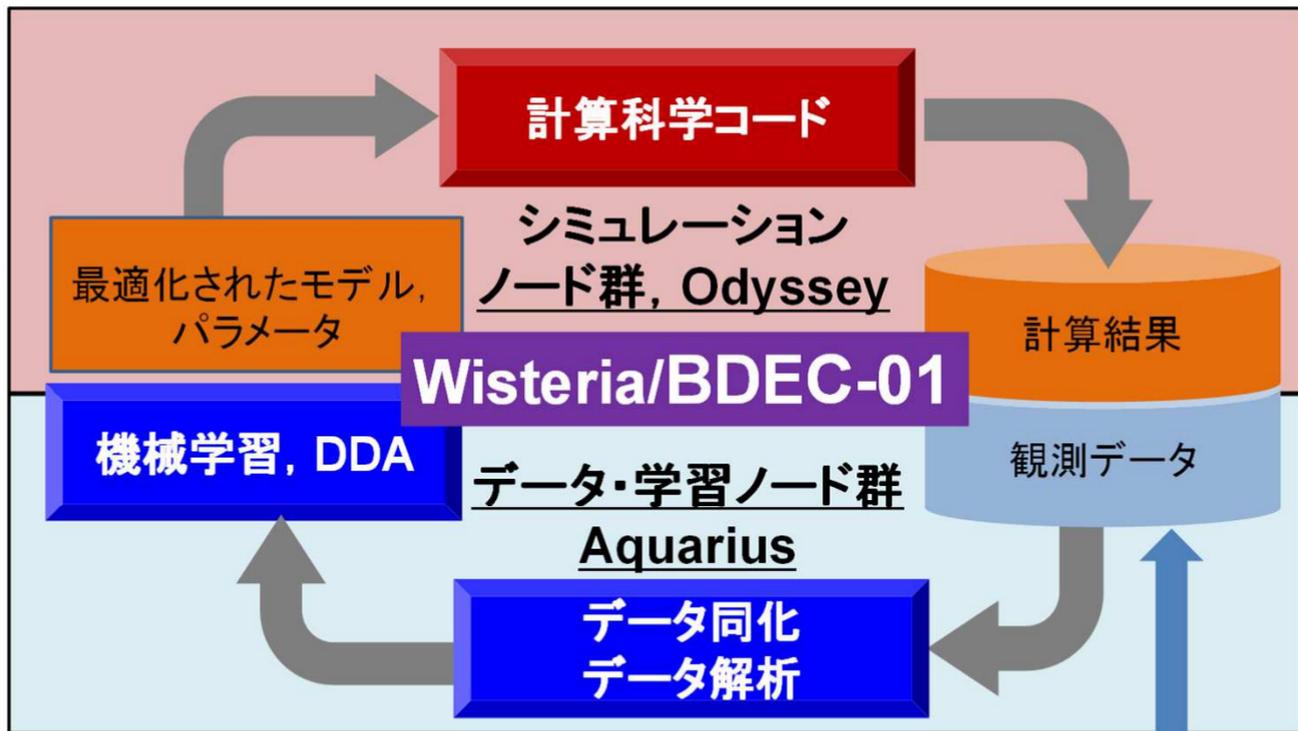
Shared File  
System  
(SFS)  
25.8 PB,  
0.50 TB/s

## Data/Learning Nodes Aquarius

7.20 PF, 578.2 TB/s



**Wisteria  
BDEC-01**



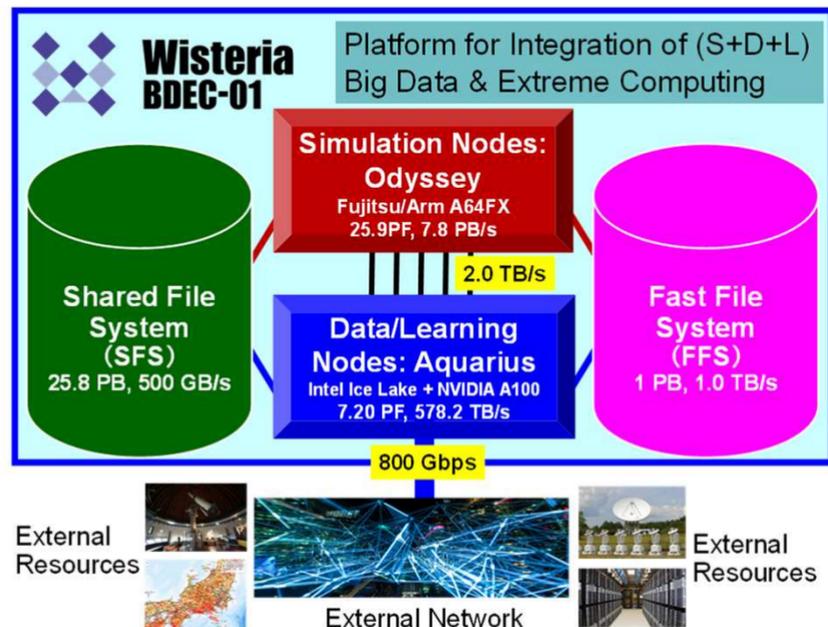
**シミュレーションのためのモデル・パラメータのデータ解析, AI/機械学習による最適化 (S+D+L)**

# SC22における諸ランキング (2022年11月)



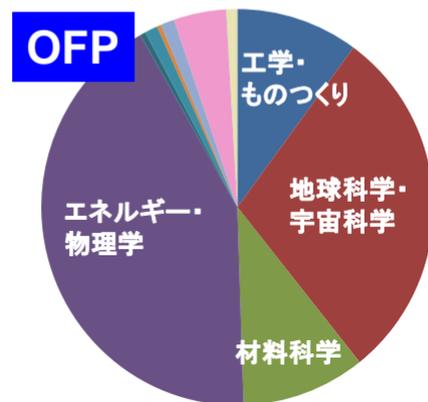
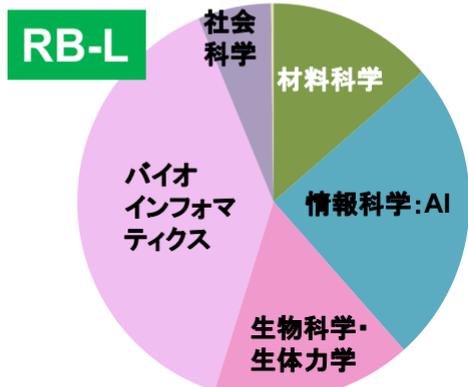
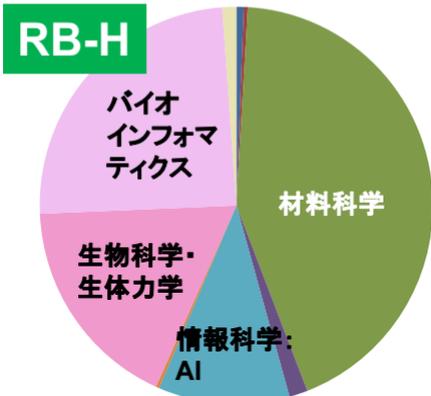
	Odyssey	Aquarius
TOP 500	23	125
Green 500	45	28
HPCG	12	68
Graph 500 BFS	4	-
HPL-MxP (HPL-AI)	10*	-

\*) ISC 2022 (June 2022)

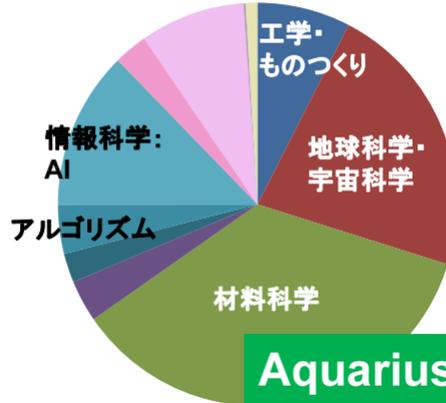
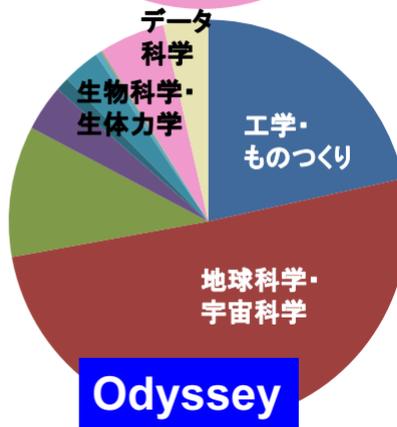
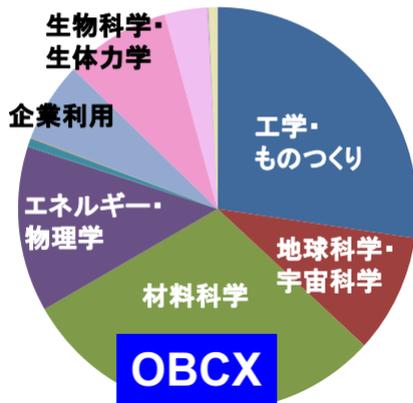


# 2021年度分野別 ■ 汎用CPU, ■ GPU

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

# 技術的な特徴など



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

- Odyssey
  - SVE (Scalable Vector Extension)
    - Armv8-A命令セットアーキテクチャをスーパーコンピュータ向けに拡張
  - FP16
  - 機械学習・AIワークロードへの適用
- Aquarius
  - HPC・計算科学への適用
  - CPU: Intel Xeon Ice Lake
    - 3<sup>rd</sup> Generation Intel Xeon Scalable Processors
    - 推論, 単独での利用は難しいが
  - GPU: NVIDIA A100 Tensor Core
    - Tensor Core + Tensor Float [TF32]
- Odyssey-Aquarius
  - InfiniBand-EDR

# 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](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/>
- 3<sup>rd</sup> 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](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](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)
  - <https://h3-open-bdec.cc.u-tokyo.ac.jp/>
- Wisteria/BDEC-01 & h3-Open-BDEC紹介講演(日本語)
  - [https://www.youtube.com/watch?v=CsJ\\_9aGNXCg](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>



- JHPCNの概要
- 東大情報基盤センターのスパコン概要
- Wisteria/BDEC-01
- **Oakbridge-CX(2023年9月末運用終了)**
- **Ipomoea-01**
- h3-Open-BDEC
  
- 利用事例
- JHPCN利用申込み

# 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: PRIMERGY CX400 M4 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  
手記憶容量: 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

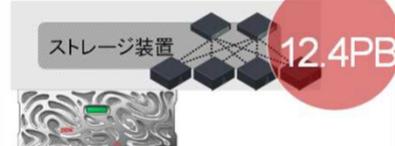
### 管理サーバ群



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)

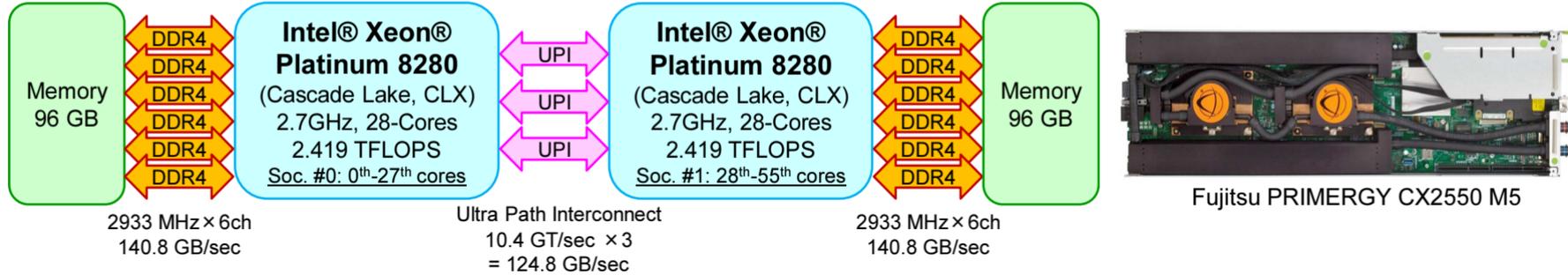
### 並列ファイルシステム



12.4PB

ストレージ装置: DDN ES18KE 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)

# Intel Xeon Platinum 8280 (Cascade Lake-SP, CLX)



- Intel Xeon Platinum 8280 (Cascade Lake, CLX), 富士通
  - 1,368 nodes, 6.61 PF peak, 385.1 TB/sec,
  - **4.2+ PF for HPL**
- Network: Intel Omni-Path, 100 Gbps, Full Bi-Section
- Storage: DDN EXAScaler (Lustre)
  - 12.4 PB, 193.9 GB/sec
- Power Consumption:
  - 950.5 kVA

# 全体構成

項目		仕様
総理論演算性能		6.61 PFLOPS
総ノード数		1,368=1,240+112+16
総主記憶容量		256.5 TiB
ネットワークポロジ		Full-bisection Fat Tree
並列ファイルシステム	システム名	Lustreファイルシステム
	サーバ(OSS)	DDN ES18K
	サーバ(OSS)数	8
	ストレージ容量	12.4 PB
	ストレージデータ転送速度	193.9 GB/s

# ノードの構成

項目		仕様	
製品名		Fujitsu PRIMERGY CX2550 M5	Fujitsu PRIMERGY CX2560 M5
ノード数		1240	112+16
CPU	プロセッサ名	Intel® Xeon® Platinum 8280 (開発コード名 : CascadeLake)	
	プロセッサ数(コア数)	2 (28+28)	
	周波数	2.7 GHz	
	理論演算性能	4.8384 TFLOPS	
Memory		192 GiB(DDR4)	
インターコネク		Intel® Omni-Path ネットワーク (100 Gbps)	
SSD	容量	-	1.6 TB(NVMe)
	読み出し性能		3.20 GB/s
	書き込み性能		1.32 GB/s

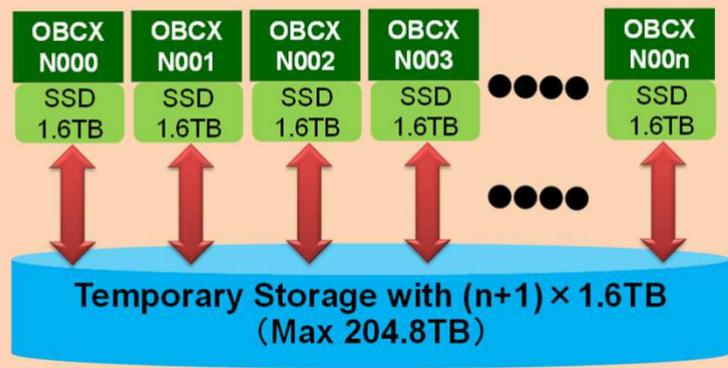
# ソフトウェア構成

項目	構成
OS	Red Hat Enterprise Linux 7, CentOS 7
コンパイラ	GNU コンパイラ Intel コンパイラ(Fortran77/90/95/2003/2008, C, C++)
メッセージ通信ライブラリ	Intel MPI, Open MPI, Intel Omni-Path Fabric Software
ライブラリ	Intel社製ライブラリ(MKL)(BLAS, CBLAS), その他(LAPACK, ScaLAPACK, SuperLU, SuperLU MT, SuperLU DIST, METIS, MT-METIS, ParMETIS, Scotch, PT-Scotch, PETSc, Trillinos, FFTW, GNU Scientific Library, NetCDF, Parallel netCDF, HDF5, Cmake, Anaconda, Xabclib, ppOpen-HPC, ppOpen-AT, MassiveThreads
アプリケーション	Mpjava, OpenFOAM, ABINIT-MP, PHASE, FrontFlow/blue, FrontISTR, REVOCAP-Coupler, REVOCAP-Refiner, OpenMX, xTAPP, AkaiKKR, MODYLAS, ALPS, feram, GROMACS, BLAST, R packages, Bioconductor, BioPerl, BioRuby, BWA, GATK, SAMtools, Quantum ESPRESSO, Xcrypt, Paraview, VisIt, POV-Ray
フリーソフトウェア	Autoconf, automake, bash, bzip2, cvs, emacs, nndutils, gawk, gdb, make, grep, gnuplot, gzip, less, m4, perl, ruby, sed, ubversion, tar, tcsh, tcl, zsh, FUSE, git 等
コンテナ仮想化	singularity (dockerイメージ利用可)

# 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

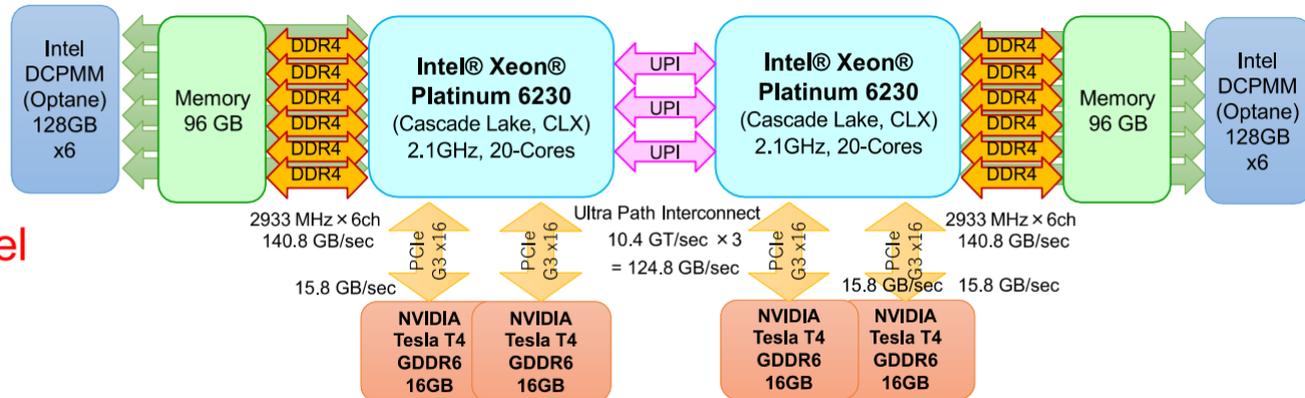


# Mini-DP (1/2)

- GPUサーバ1 × 1台
  - Intel Xeon Gold 6230 (Cascade Lake, ) (20コア) × 2
  - NVIDIA Tesla V100 32GB × 2
  - 192 GBメモリ, 100 G bit Ethernet

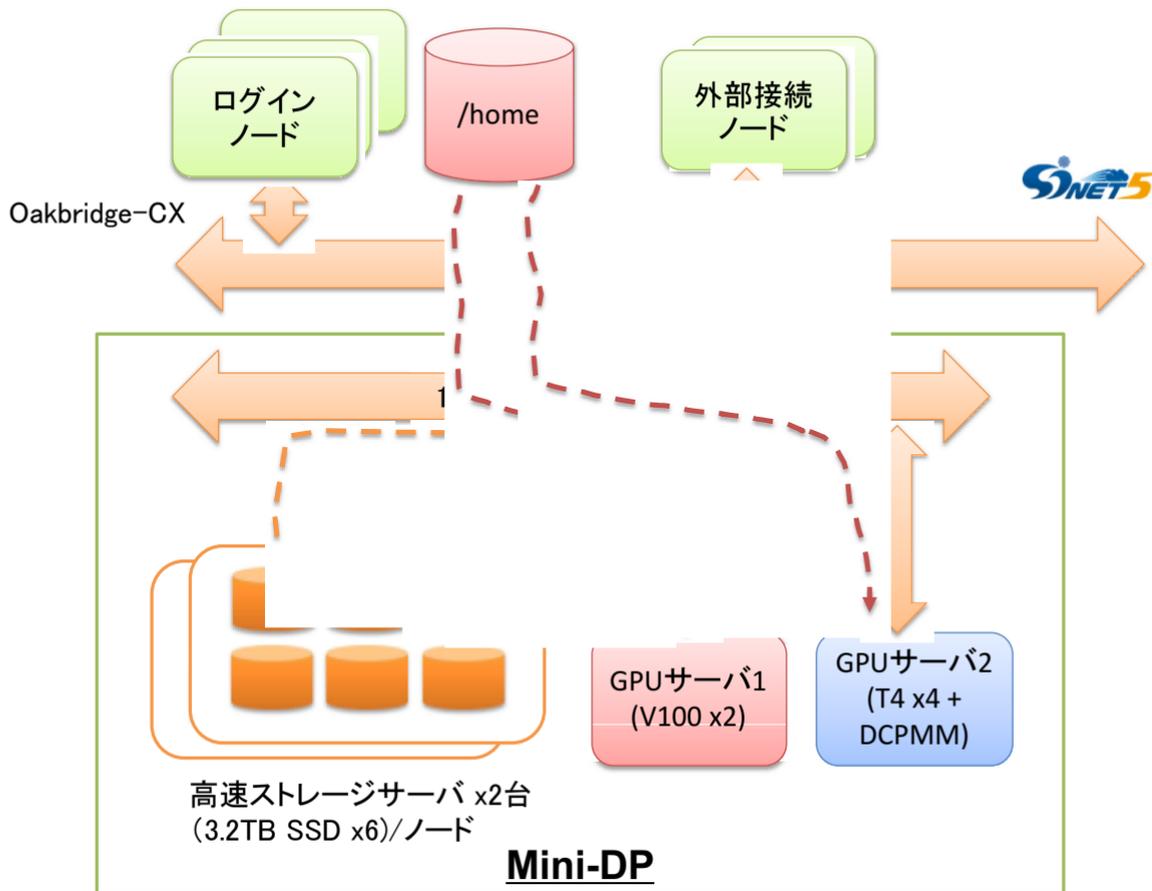
- GPUサーバ2 × 1台

- Intel Xeon Gold 6230 (CLX) (20コア) × 2
- NVIDIA Tesla T4 16GB × 4
- 192 GBメモリ + 1.5 TB不揮発メモリ (Intel DCPMM, Optane DIMM)
- 100 G bit Ethernet



# Mini-DP (2/2)

- 高速ストレージサーバ  
×2台 (3.2TB NVMe  
SSD ×6枚搭載)
- GPUサーバ1, 2は  
OBCXのログインノード  
ファイルシステムをマウ  
ント、高速ストレージサー  
バのNVMe SSDを  
NVMe over Fabricsによ  
りブロックデバイスとして  
attachして利用可能(構  
成変更も応相談)



# Mini-DP (Data Platform)

Total: 1,368 nodes

128 nodes  
with SSD

16

SINET5

External  
Resources

## OBCXの16ノード

SINET経由で外部計算機  
資源に直接接続  
BDECにおける外部ノード  
(EXN)と融合ノード(ITN)  
の中間的役割

## Mini-DP

GPUサーバー with NVMe  
SSD, BDECのデータ学習  
ノードとmdxの中間的役割

# Oakbridge-CXに関する情報

- 全般
  - <https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/>
- 利用コース
  - <https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/course.php>
- ジョブクラス
  - <https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/job.php>
- 利用申込・利用負担金
  - <https://www.cc.u-tokyo.ac.jp/guide/application/>
- FAQ
  - <https://www.cc.u-tokyo.ac.jp/faq/obcx.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



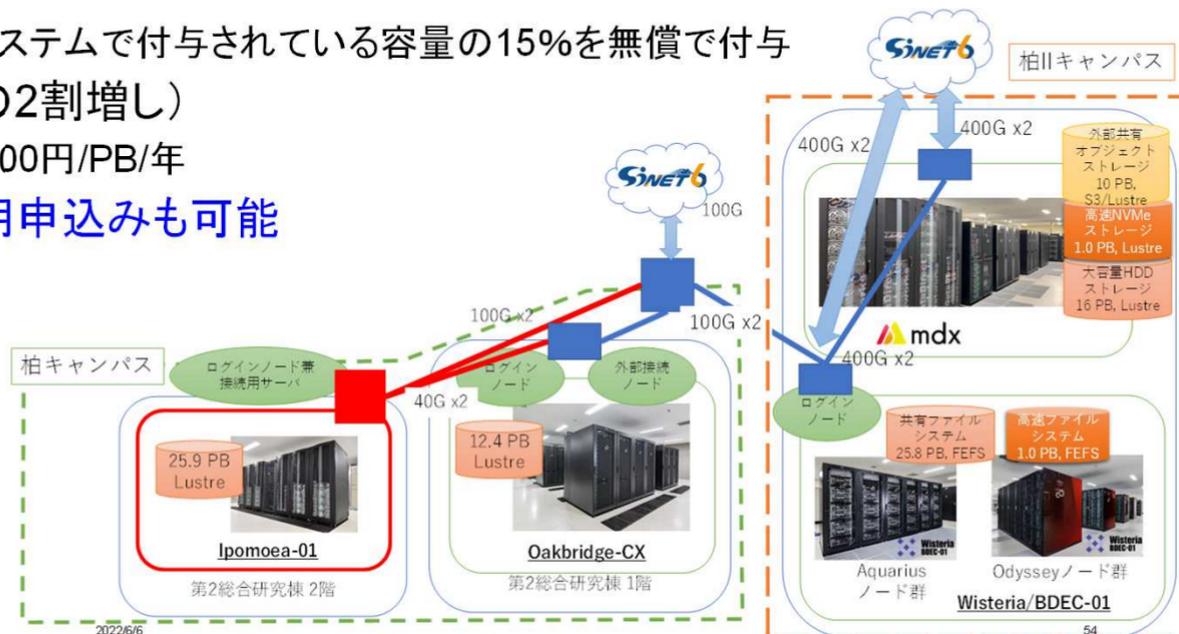
# 大規模共通ストレージシステム「Ipomoea」

- 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



- 2022年1月運用開始・6月より一般に公開, 25+PB, 富士通製
  - 2022年5月末までにOFPのLustre領域の必要ファイルの移行完了
- 割当容量
  - 東大センターのシステムに利用者番号(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のみの利用申込みも可能

# Ipomoea-01



- JHPCNの概要
- 東大情報基盤センターのスパコン概要
- Wisteria/BDEC-01
- Oakbridge-CX
- Ipomoea-01
- **h3-Open-BDEC**
- 利用事例
- JHPCN利用申込み

# h3-Open-BDEC

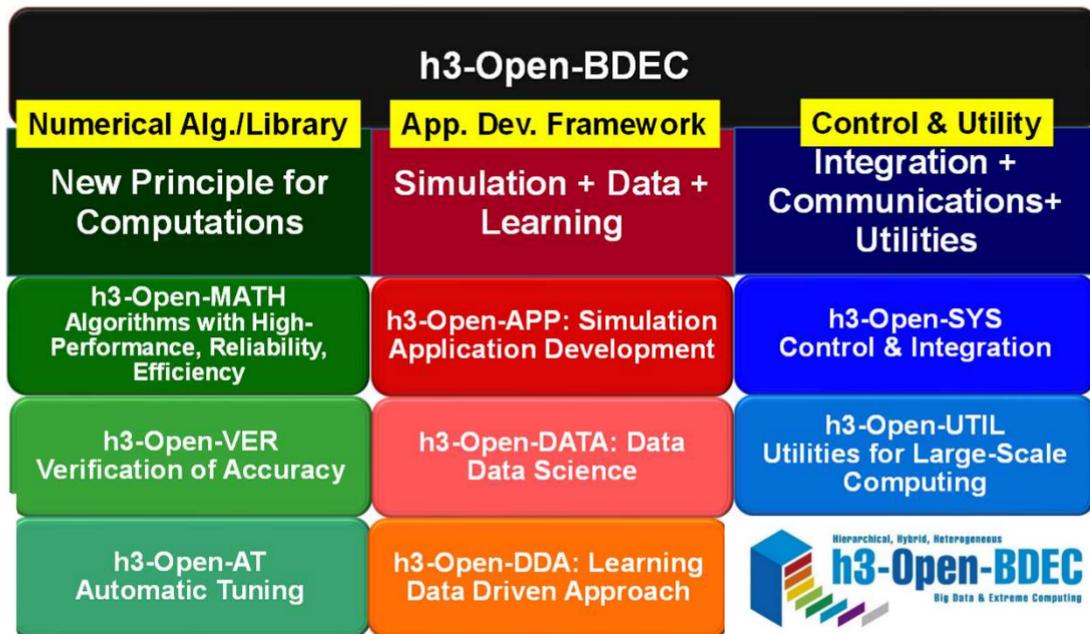
「計算+データ+学習」融合を実現する革新的ソフトウェア基盤  
科研費基盤研究(S)(2019年度~23年度, 代表: 中島研吾)

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

Hierarchical,  
Hybrid,  
Heterogeneous

Big Data &  
Extreme  
Computing

- 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等に基づく革新的機械学習手法
- ③ ヘテロジニアス環境 (e.g. Wisteria/BDEC-01) におけるソフトウェア, ユーティリティ群



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

計算科学コード

シミュレーション  
ノード群, Odyssey

最適化されたモデル,  
パラメータ

計算結果

Wisteria/BDEC-01

機械学習, DDA

データ・学習ノード群  
Aquarius

観測データ

データ同化  
データ解析



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

# AI for HPC, AI for Science の実現へ向けて



## 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

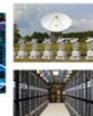
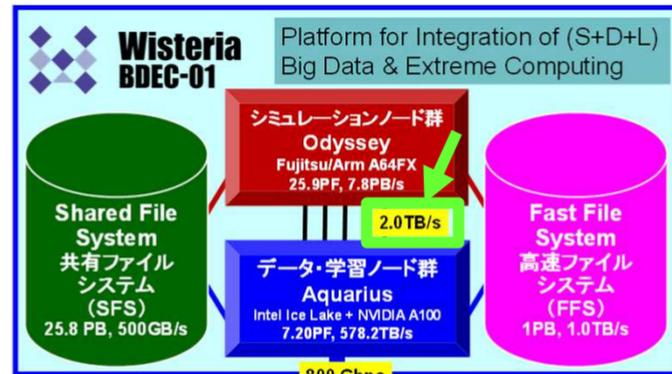
## ソフトウェア開発

– 高機能カプラー: h3-Open-UTIL/MP

– 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)



External Resources

外部リソース

External Network  
外部ネットワーク

## 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]

[住元他, HPC-181, 2021]

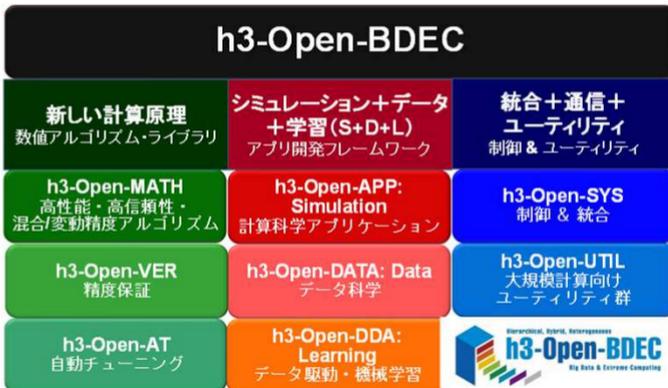
- ヘテロジニアス環境下での異なるコンポーネント間ファイル経由連携ライブラリとして考案

## 機能

- ✓ Odysseus～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

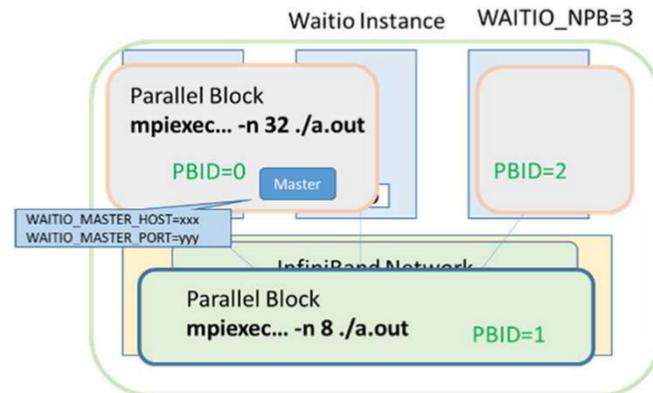
- API: C/C++, Fortranから呼び出し可能

- ✓ MPIライクなインタフェースを提供



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

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

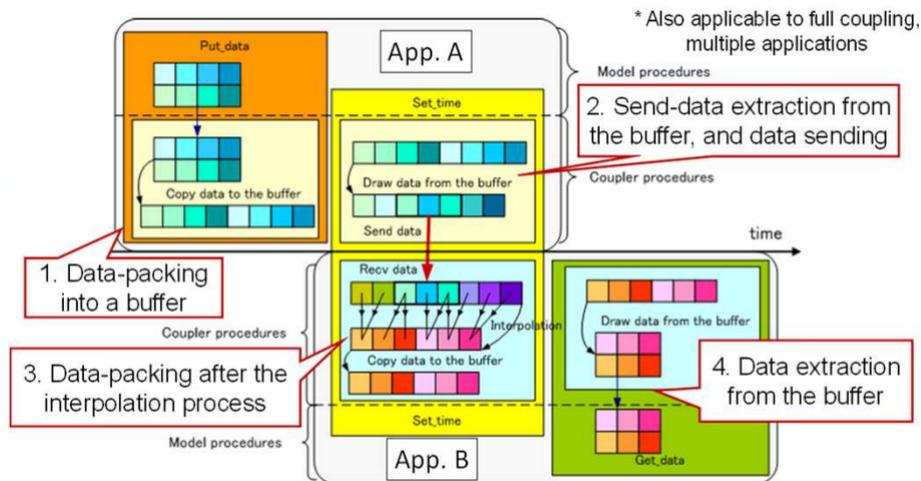
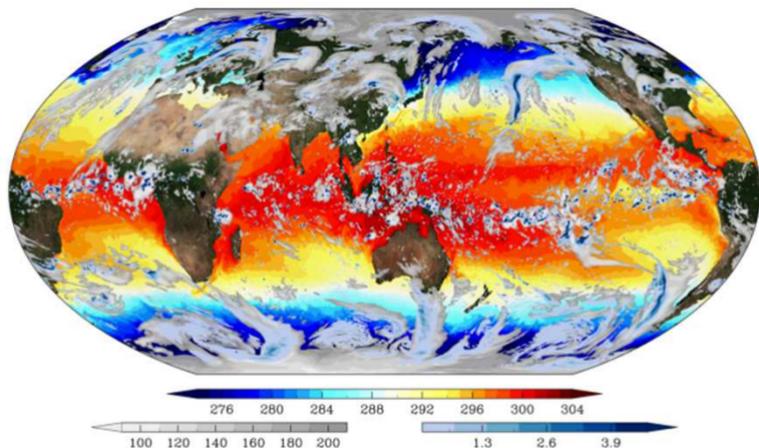


[Sumimoto et al. 2021]

# 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



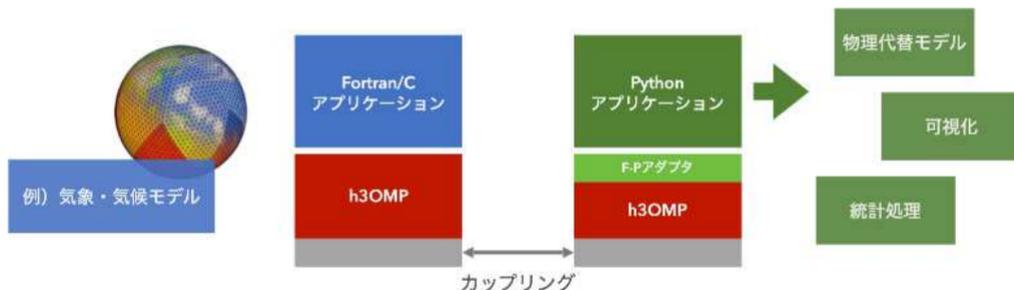
# 「計算＋データ＋学習」融合を支援する 多機能カプラーh3-Open-UTIL/MP



- 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

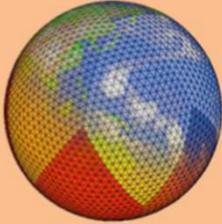
- FortranやCで記述されたプログラム同士の連成計算に限って開発を行ってきたカプラーを、Pythonによって記述されたAI・機械学習、可視化処理系のワークロードから活用できるように機能拡充。



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

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

## ARM: A64FX



A huge amount of  
simulation data  
output

HPC App  
(Fortran)

h3o-U/MP

## IceLake+A100

Analysis/ML  
App  
(Python)

F<->P adapter

h3o-U/MP

Surrogate  
Model

Visualization

Statistics

Coupling

IB-EDR



**Wisteria  
BDEC-01**

**Odyssey**



**Wisteria  
BDEC-01**

**Aquarius**

# h3-Open-UTIL/MP

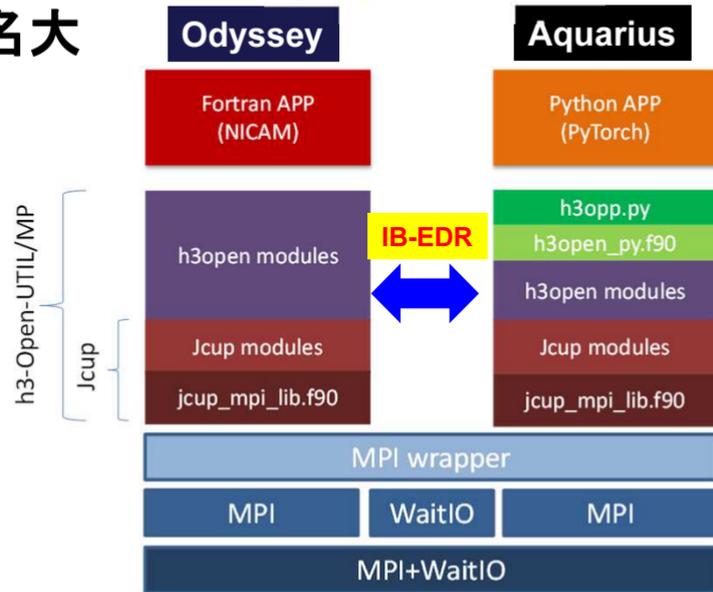
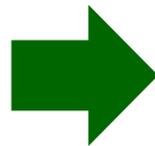
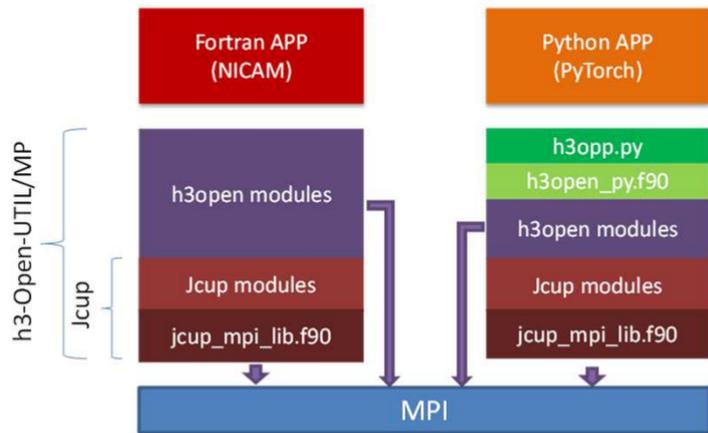
## h3-Open-SYS/WaitIO-Socket連携

2022年6月から利用可能

2022年度はFS経由のWaitIO-File整備:名大



**Wisteria  
BDEC-01**



2021年4月: MPI通信可能な環境を前提

2022年6月: Coupler+WaitIO

# 解説記事 : h3-Open-UTIL/MP・ h3-Open-SYS/WaitIO-Socket



- h3-Open-UTIL/MP

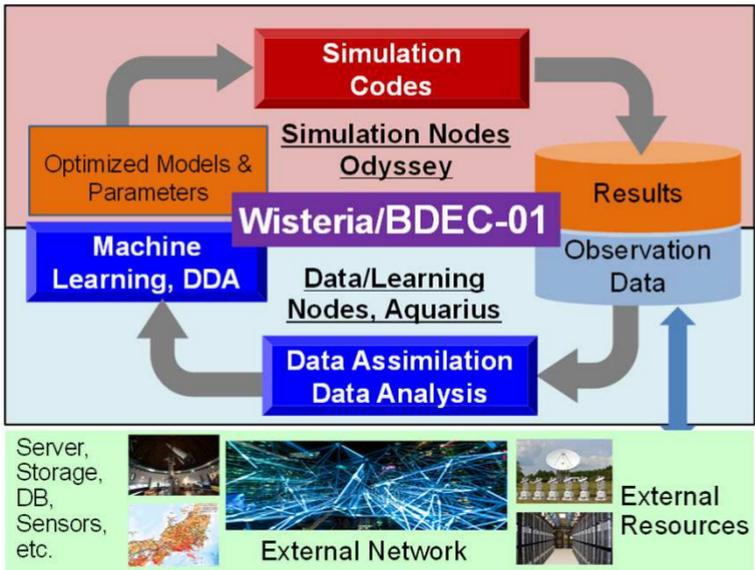
- [https://www.cc.u-tokyo.ac.jp/public/VOL24/No3/13\\_202205-Wisteria-2.pdf](https://www.cc.u-tokyo.ac.jp/public/VOL24/No3/13_202205-Wisteria-2.pdf)
- <http://nkl.cc.u-tokyo.ac.jp/files/202207UtilMPfinal.pdf>



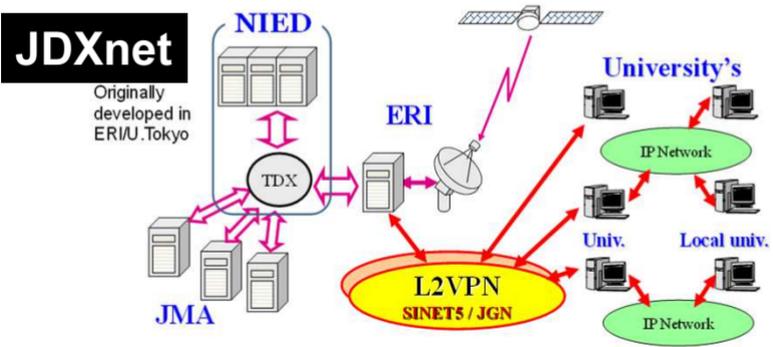
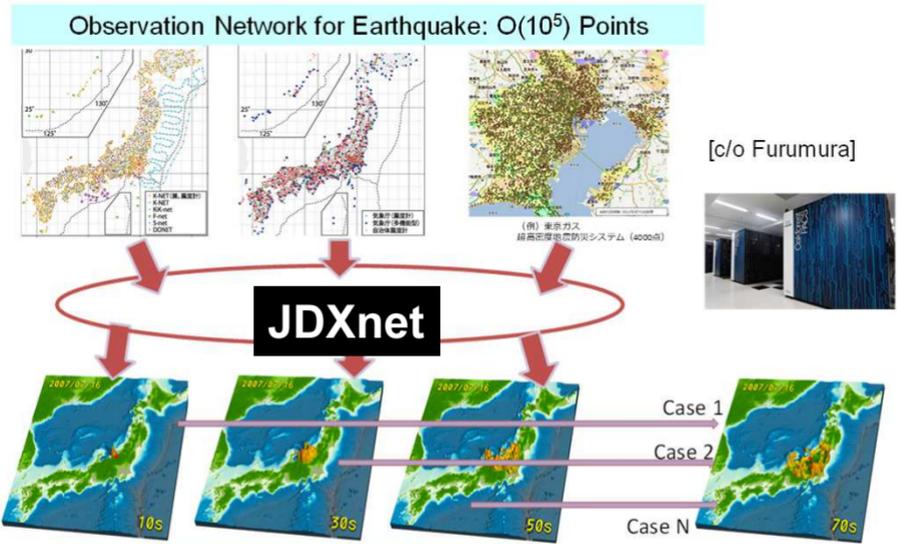
- h3-Open-SYS/WaitIO-Socket

- [https://www.cc.u-tokyo.ac.jp/public/VOL24/No2/10\\_202203Wisteria-1.pdf](https://www.cc.u-tokyo.ac.jp/public/VOL24/No2/10_202203Wisteria-1.pdf)
- [https://www.cc.u-tokyo.ac.jp/public/VOL24/No3/12\\_202205-Wisteria-1.pdf](https://www.cc.u-tokyo.ac.jp/public/VOL24/No3/12_202205-Wisteria-1.pdf)





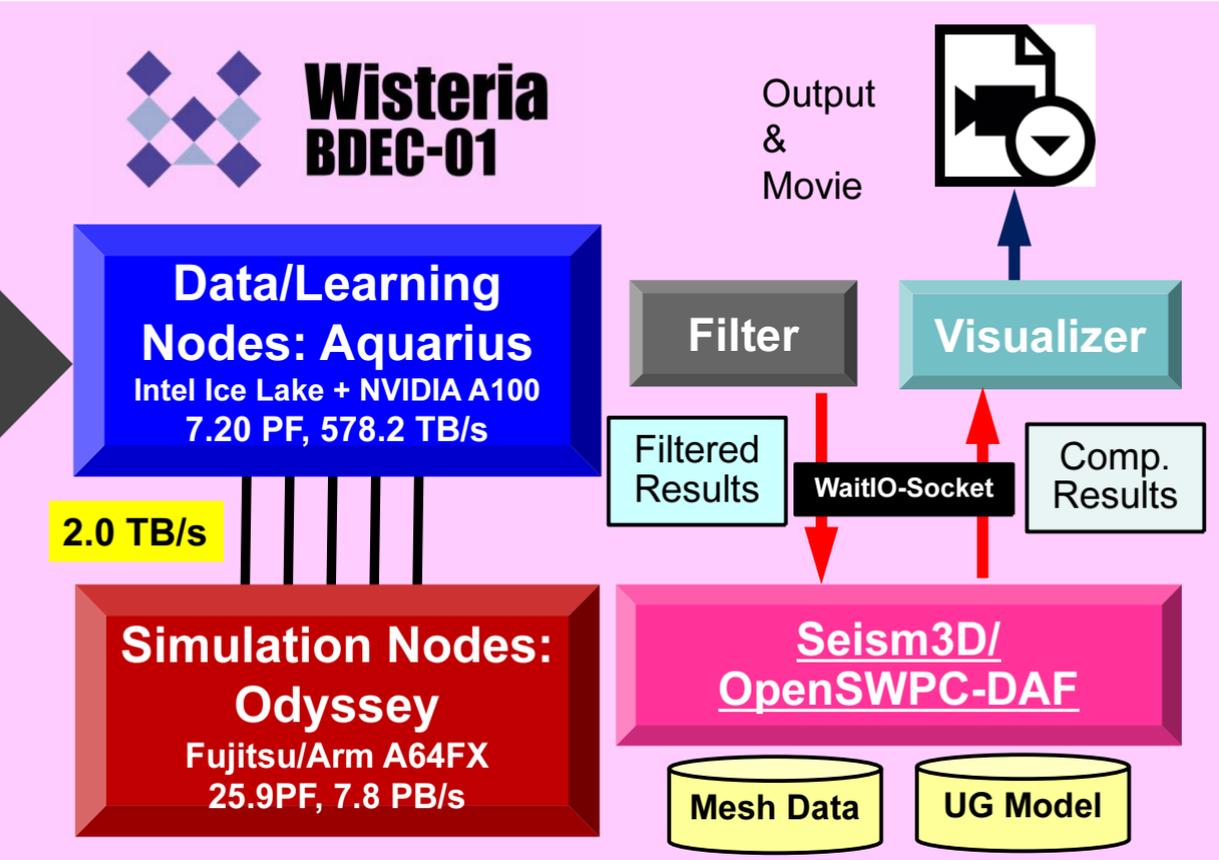
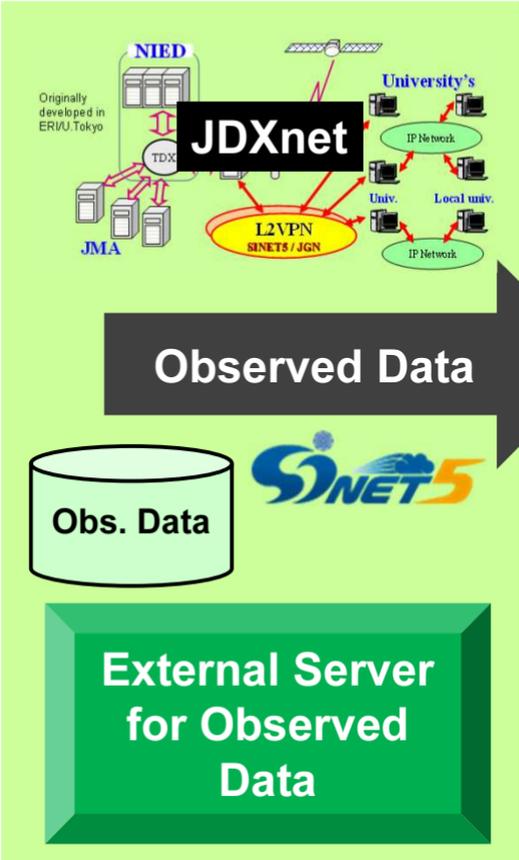
# リアルタイムデータ同化+ 3D強震動シミュレーション融合 JDXnetによるリアルタイム観測データ活用



Real-Time Data/Simulation Assimilation  
Real-Time Update of Underground Model

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

# 長周期地震動シミュレーション+観測データ同化



# Communications by WaitIO-Socket

[Kasai et al. 2021]

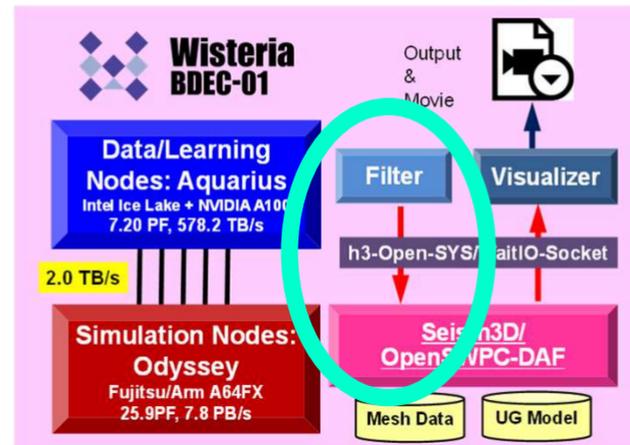
## Aquarius: SEND

```
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_INTEGER, 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_INTEGER, 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

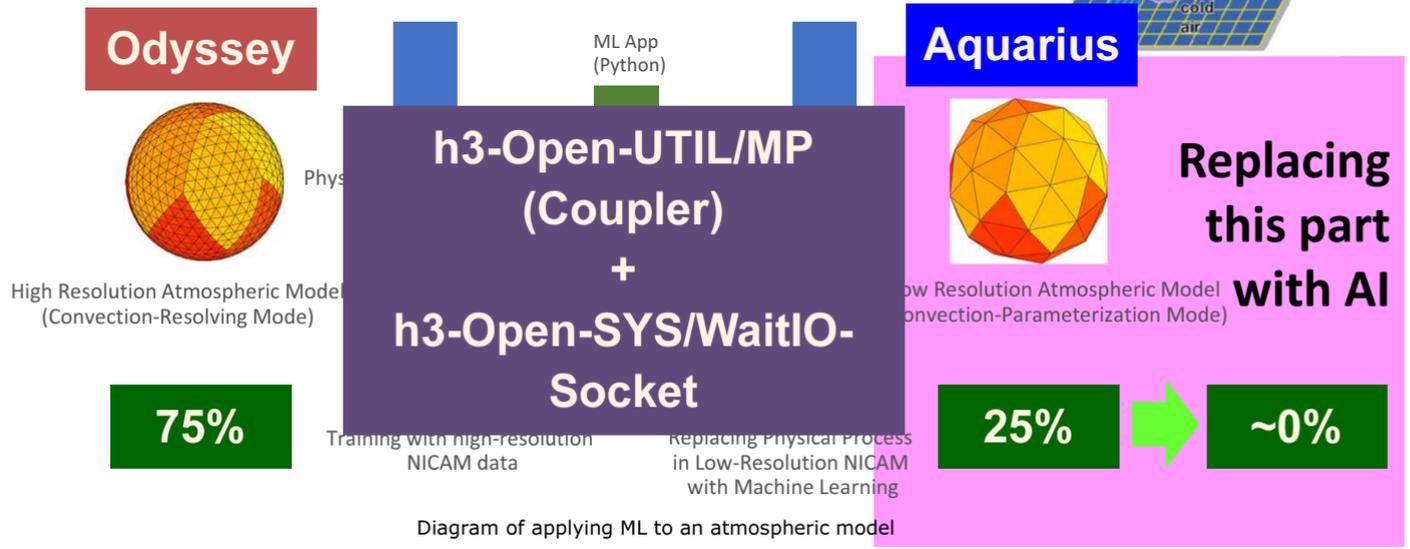
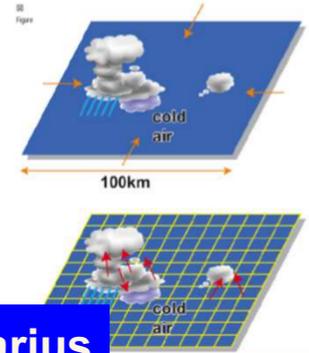
```
call WAITIO_MPI_RECV (NTMAX1_o, 1, WAITIO_MPI_INTEGER, 0,1, WAITIO_COMM_UNIVERSE, ...)
call WAITIO_MPI_RECV (DT_o, 1, WAITIO_MPI_FLOAT, 0,2, WAITIO_COMM_UNIVERSE, ...)
call WAITIO_MPI_RECV (NST_o, 1, WAITIO_MPI_INTEGER, 0,3, WAITIO_COMM_UNIVERSE, ...)
call WAITIO_MPI_RECV (AT_o, 1, WAITIO_MPI_FLOAT,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



# Atmosphere-ML Coupling

[Yashiro (NIES), Arakawa (ClimTech/U.Tokyo)]

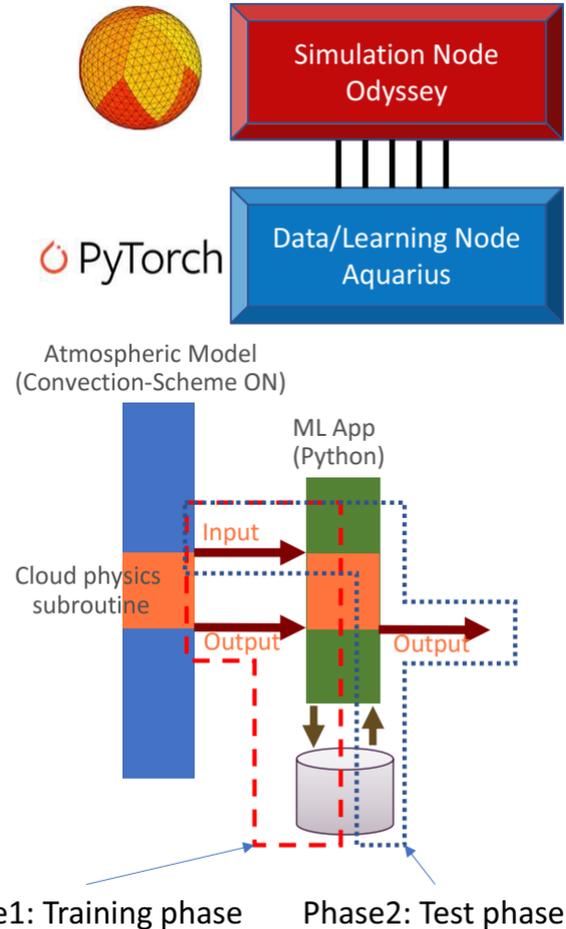
- Motivation of this experiment
  - Two types of Atmospheric models: Cloud resolving VS Cloud parameterizing
  - Cloud resolving model is difficult to use for climate simulation
  - Parameterized model has many assumptions
  - Replacing low-resolution cloud processes calculation with ML!



# Experimental Design

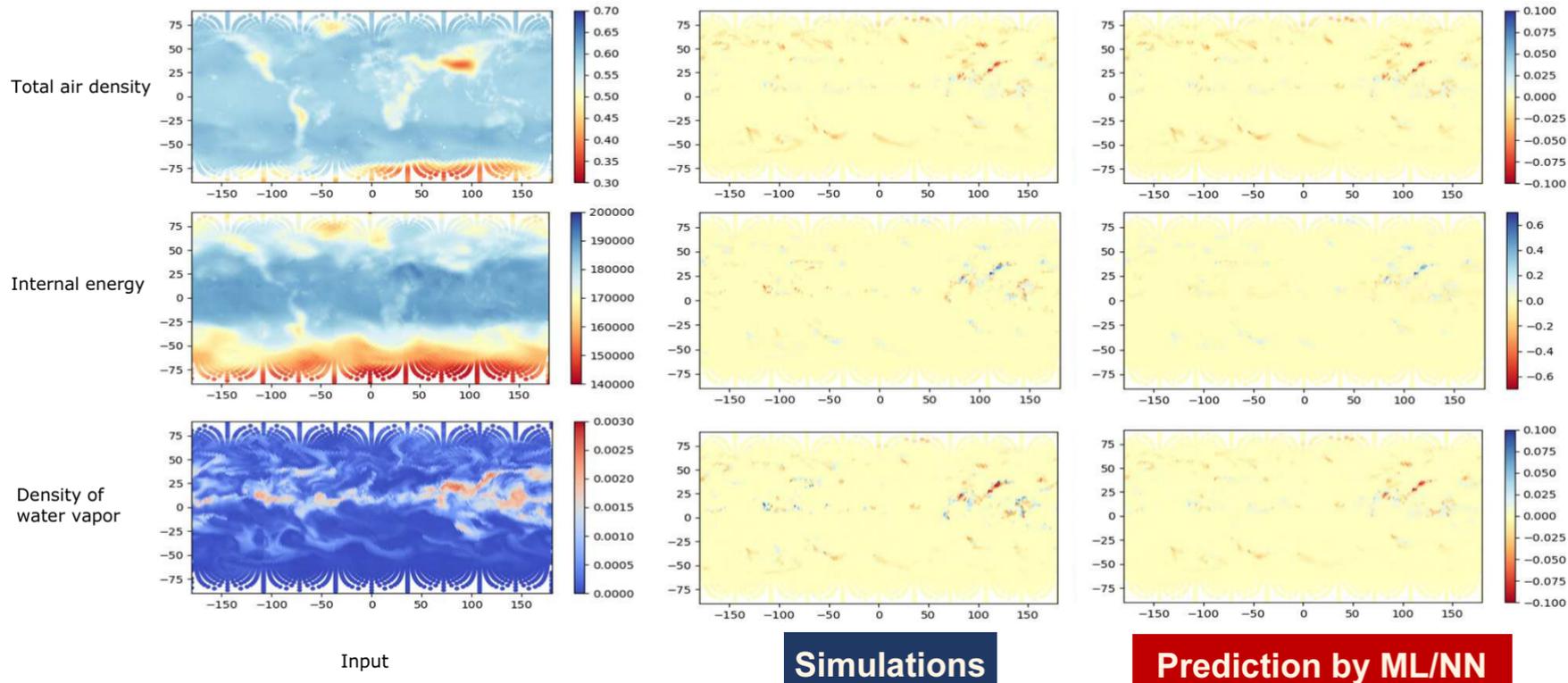
- Atmospheric model on Odyssey
  - NICAM : global non-hydrostatic model with an icosahedral grid
  - Resolution : horizontal : 10240, vertical : 78
- ML on Aquarius
  - Framework : PyTorch
  - Method : Three-Layer MLP
  - Resolution : horizontal : 10240, vertical : 78
- Experimental design
  - Phase1: PyTorch is trained to reproduce output variables from input variables of cloud physics subroutine.
  - Phase2: Reproduce the output variables from Input variables and training results
- Training data
  - Input : total air density ( $\rho$ ), internal energy ( $e_{in}$ ), density of water vapor ( $\rho_q$ )
  - Output : tendencies of input variables computed within the cloud physics subroutine

$\frac{\Delta \rho}{\Delta T}$	$\frac{\Delta e_{in}}{\Delta T}$	$\frac{\Delta \rho_q}{\Delta T}$
--------------------------------	----------------------------------	----------------------------------



# Test calculation

- Compute output variables from input variables and PyTorch
  - The rough distribution of all variables is well reproduced
  - The reproduction of extreme values is no good



# 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 fjmp
module load waitio

export WAITIO_MASTER_HOST=`hostname`
export WAITIO_MASTER_PORT=7100
export WAITIO_PPID=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 omp
module load intel
module load impi
module load waitio

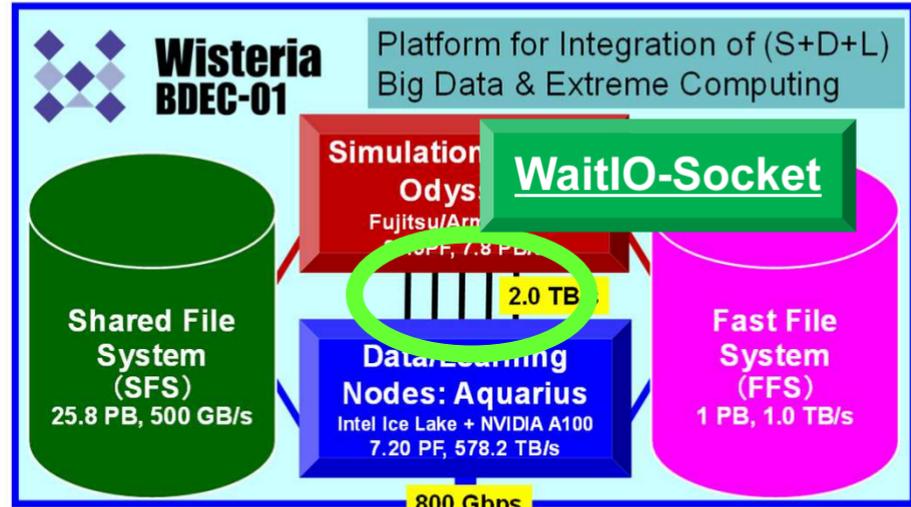
export WAITIO_MASTER_HOST=`waitio-serv -c`
export WAITIO_MASTER_PORT=7100
export WAITIO_PPID=1
export WAITIO_NPB=2

module unload intel
module unload impi
module load gcc omp

mpiexec -n 10 ./ada
```

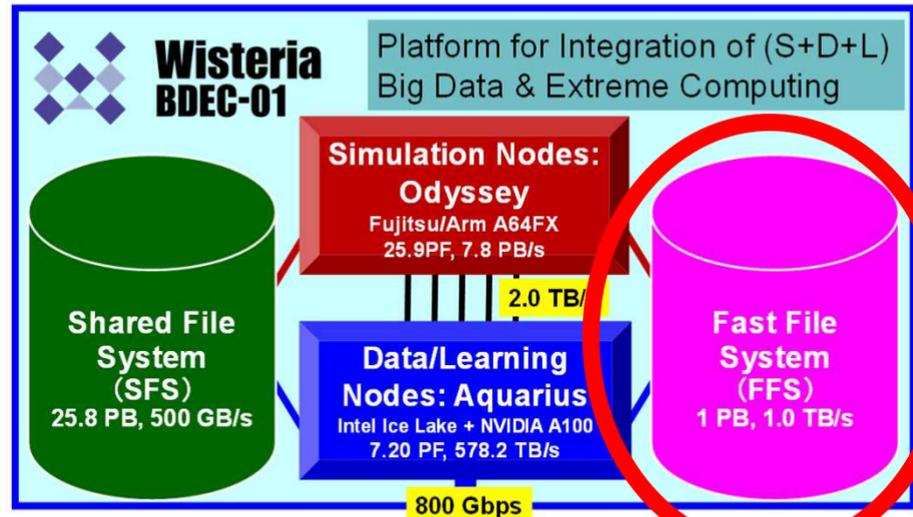
# h3-Open-SYS/WaiIO-Socket

- 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/WaiIO-Socket**
    - Library for Inter-Process Communication through IB-EDR with MPI-like interface



# h3-Open-SYS/WaiIO-File

- 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/WaiIO-File**
    - **Library for Inter-Process Communication through FFS with MPI-like interface**



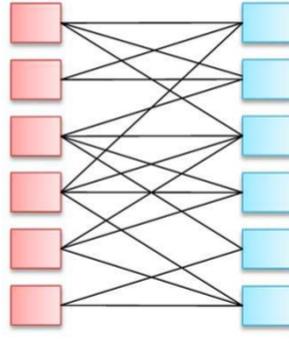
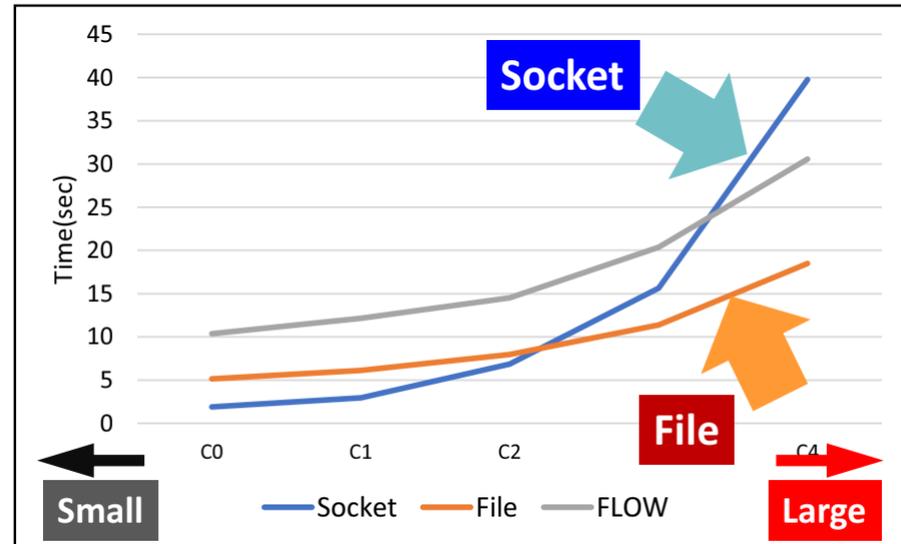
# Preliminary Evaluation: WaitIO-Socket, -File

- Odyssey-Aquarius, Flow (Nagoya University)
- Test Cases
  - Model A: 160 proc's/40 nodes: A64FX (Odyssey, Flow I)
  - Model B: 20 proc's/1 node: Intel Xeon (Aquarius, Flow II)
    - Waitio\_isend/irecv/wait

## Results

- “Socket” is better for small cases, “File” is better for large ones
- “Flow” with WaitIO-Files is 2x slower than Odyssey-Aquarius
  - Reasonable Number

Case	Grid #	Total Size
C0	81,920	26,214,400
C1	163,840	52,428,800
C2	327,680	104,857,600
C3	655,360	209,715,200
C4	1,310,720	419,430,400

- JHPCNの概要
- 東大情報基盤センターのスパコン概要
- Wisteria/BDEC-01
- Oakbridge-CX
- Ipomoea-01
- h3-Open-BDEC
  
- **利用事例**
- JHPCN利用申込み

# 2020年度東大利用課題(OB外:OBCX外部接続)(1/2)

計算科学, 計算機科学・システム, データ・学習, 融合

課題番号	課題名・代表者(所属)	RB	OPF	OB	OB外	FENNEL
jh200002-NAH	電磁流体力学乱流の高精度・高並列LESシミュレーションコード開発研究 三浦英昭(核融合科学研究所)		○			
jh200005-NAH	大規模並列計算による格子の最短ベクトル探索の効率化に関する研究 柏原賢二(東京大学)	○				
jh200008-NAH	Developing Accuracy Assured High Performance Numerical Libraries for Eigenproblems 片桐孝洋(名古屋大学)	○		○		
jh200016-NAH	データサイエンスに基づく高分子材料の構造物性相関 天本義史(九州大学)	○				
jh200020-MDH	機械学習に基づくマクロ経済変動の数理モデリング 齊木吉隆(一橋大学)	○				
jh200022-DAH	社会の分析とシミュレーションのための合成人口データ提供システム 村田忠彦(関西大学)					○
jh200023-NAHI	Hierarchical low-rank approximation methods on distributed memory and GPUs 横田理央(東京工業大学)	○				
jh200027-ISH	超巨大ニューラルネットワークのための分散深層学習フレームワークの開発とスケーラビリティの評価 田中正弘(情報通信研究機構)	○		○		
jh200029-NAH	時空間領域境界積分方程式法の高速解法の開発と巨大地震シミュレーションへの応用 安藤亮輔(東京大学)		○			
jh200036-MDHI	High resolution simulation of cardiac electrophysiology on realistic whole-heart geometries 中島研吾(東京大学)		○	○		
jh200037-NAH	高性能・変動精度・高信頼性数値解析手法とその応用 中島研吾(東京大学)	○	○	○		

# 2020年度東大利用課題(OB外:OBCX外部接続)(2/2)

## 計算科学, 計算機科学・システム, データ・学習, 融合

課題番号	課題名・代表者(所属)	RB	OFFP	OB	OB外	FENNEL
jh200038-MDH	HPCと高速通信技術の融合による大規模データの拠点間転送技術開発と実データを用いたシステム実証試験 村田健史(情報通信研究機構)			○		
jh200041-NAH	Innovative Multigrid Methods II 藤井昭宏(工学院大学)		○	○		
jh200042-DAH	Deep Learningを用いた医用画像診断支援に関する研究 佐藤一誠(東京大学)	○	○	○		
jh200043-MDHI	Development of Fast Surrogate for Approximating Large-scale 3D Blood Flow Simulation 下川辺隆史(東京大学)	○	○	○		
jh200045-NAH	エクサスケール時代の数値計算手法に対する性能予測技術 深谷猛(北海道大学)		○	○		
jh200046-DAH	分散機械学習技術を用いた大規模医用画像処理の実現に向けた研究 大島聡史(名古屋大学)	○				
jh200047-NWH	ハイブリッドクラウド構築とゲノム情報解析の効率的な運用に関する研究 長崎正朗(京都大学)			○	○	
jh200051-NAHI	Scalable Multigrid Poisson solver for AMR-based CFD applications in Nuclear Engineering 小野寺直幸(日本原子力研究開発機構)			○		
jh200062-NAH	ゲリラ豪雨予測のリアルタイム実証実験 三好建正(理化学研究所)		○			
jh200064-NAH	機械学習を用いた風環境予測精度の向上と防災技術への応用 高木洋平(横浜国立大学)		○	○		
jh210003-NWJ	財務ビッグデータの可視化と統計モデリング 地道 正行(関西学院大学)					○

# 2021年度東大利用課題(OB外:OBCX外部接続)(1/2)

計算科学, 計算機科学・システム, データ・学習, 融合

課題番号	課題名・代表者(所属)	RB	OFFP	OB	OB外	W-O	W-A
jh210002-NAHI	Developing Accuracy Assured High Performance Numerical Libraries for Eigenproblems 片桐孝洋(名古屋大学)			○		○	
jh210003-NAH	GPU・CPU・ARMプロセッサに対する原子力CFDアプリケーション用の混合精度ポアソン解法 小野寺直幸(日本原子力研究開発機構)			○			○
jh210004-NAH	電磁流体力学乱流の高精度・高並列LESシミュレーションコード開発研究 三浦英昭(核融合科学研究所)		○			○	
jh210009-MDH	HPCと高速通信技術の融合による大規模データの拠点間転送技術開発と実データを用いたシステム実証試験 村田健史(情報通信研究機構)				○		
h210011-DAH	Deep Learningを用いた医用画像診断支援に関する研究 佐藤一誠(東京大学)	○					○
jh210015-NAH	高性能かつ高信頼な数値計算手法とその応用 萩田武史(東京女子大学)		○	○		○	○
jh210017-MDH	Development of physics informed machine learning for soft matter: polymer flows and beyond John Molina(京大)			○			○
jh210018-NWH	ハイブリッドクラウド構築とゲノム情報解析の効率的な運用に関する研究 長崎正朗(京都大学)			○	○		
jh210021-MDHI	High resolution simulation of cardiac electrophysiology on realistic whole-heart geometries 中島研吾(東京大学)		○	○		○	
jh210022-MDH	三次元強震動シミュレーションとリアルタイムデータ同化の融合 中島研吾(東京大学)		○	○	○	○	○
jh210023-NAH	時空間領域境界積分方程式法の高速解法の開発と巨大地震シミュレーションへの応用 安藤亮輔(東京大学)		○				

# 2021年度東大利用課題(OB外:OBCX外部接続)(2/2)<sup>06</sup>

## 計算科学, 計算機科学・システム, データ・学習, 融合

課題番号	課題名・代表者(所属)	RB	OFF	OB	OB外	W-O	W-A
jh210024-NAHI	Hierarchical low-rank approximation methods on distributed memory and GPUs 横田理央(東京工業大学)					○	○
jh210026-NAH	Innovative Multigrid Methods II 藤井昭宏(工学院大学)		○	○		○	
jh210027-NAH	機械学習モデルのリアプノフ指数ならびにリアプノフベクトルの解析 齊木吉隆(一橋大学)		○				
jh210030-DAH	大規模分散医用画像処理に向けた医用画像処理アプリケーションの最適化 大島聡史(名古屋大学)						○
jh210034-NAH	高レイノルズ数乱流のデータ科学プラットフォームの構築 石原卓(岡山大学)		○				
jh210035-NAH	GPUの高速並列計算で実現する交差禁止制御可能な高分子シミュレータの開発 萩田克美(防衛大学校)						○
jh210040-MDH	合成人口プロジェクト: 従業地・通学地属性の確率的割当てと深層学習による空中写真からの住宅判別 村田忠彦(関西大学)(HPCI共有ストレージ使用)						
jh210044-NAH	エクサスケール時代の数値計算手法に対する性能予測技術 深谷猛(北海道大学)		○	○		○	
jh210046-NAH	グラフ構造で一般化された動的負荷分散フレームワークの構築と重合メッシュ法への適用 森田直樹(筑波大学)			○			
jh210049-MDH	Developing data driven analysis methods for extreme scale numerical simulations 朝比祐一(日本原子力研究開発機構)						○
jh210051-MDH	Development of Fast Surrogate for Approximating Large-scale 3D Blood Flow Simulation 下川辺隆史(東京大学)	○	○				○
jh211011-NWJ	財務ビッグデータの可視化と統計モデリング 地道 正行(関西学院大学)(FENNEL使用)						

# 2022年度東大利用課題(OB外:OBCX外部接続)(1/3)

計算科学, 計算機科学・システム, データ・学習, 融合

課題番号	課題名	代表者(所属)	OB	OB外	W-O	W-A
jh220005	電磁流体力学乱流の高精度・高並列LESシミュレーションコード開発研究	三浦英昭(核融合研究所)			○	
jh220007	偏った訓練データに基づく力学系の機械学習モデリング	齊木吉隆(一橋大学)			○	
jh220009 国際共同	Hierarchical low-rank approximation methods on distributed memory and GPUs	横田理央(東工大)				○
jh220011	大規模分散医用画像処理アプリケーションの実用化に向けた研究	大島聡史(名古屋大)				○
jh220012	ノードを跨ぐ多数GPU環境下でのマルチフィジックス粒子法の高速度化	浅井光輝(九州大)				○
jh220013	時空間領域境界積分方程式法の高速度解法の開発と巨大地震シミュレーションへの応用	安藤亮輔(東大)			○	
jh220014	ハイブリッドクラウドを用いたゲノム情報に基づく構造多型パネルの構築とアノテーション	長崎正朗(京大)	○	○		
jh220019	MPMとFEMによる未解明な大規模土砂災害の数値シミュレーション	寺田賢二郎(東北大)			○	
jh220020	極端気象現象予測における不確実性の起源の解明(AI-for-HPC発展課題)	澤田洋平(東大)			○	○
jh220022	高性能かつ高信頼な数値計算手法とその応用	荻田武史(東京女子大)	○		○	○

# 2022年度東大利用課題(OB外:OBCX外部接続)(2/3)

計算科学, 計算機科学・システム, データ・学習, 融合

課題番号	課題名	代表者(所属)	OB	OB外	W-O	W-A
jh220029	三次元強震動シミュレーションとリアルタイムデータ同化の融合	中島研吾(東大)	○	○	○	○
jh220030	超高解像度の即時予測の実現に向けた都市街区内部風況データベースの構築	小野寺直幸(原研)			○	○
jh220031	Targeting exa-scale systems: performance portability and scalable data analyses	朝比祐一(原研)			○	○
jh220036	大規模アプリケーションの高性能な実用的アクセラレータ対応手法	下川辺隆史(東大)				○
jh220038	GPU並列計算による高分子材料系シミュレーションの高速化技法の検討	萩田克美(防大)				○
<a href="#">jh220041</a> <a href="#">国際共同</a>	High resolution simulation of cardiac electrophysiology on realistic whole-heart geometries	中島研吾(東大)	○		○	○
jh220042	人と衣服と気流の連成相互作用シミュレーション・フレームワークの構築	青木尊之(東工大)				○
jh220046	次世代演算加速装置とそのファイルIOに関する研究	埴敏博(東大)				○
jh220047	グラフ構造で一般化された静的負荷分散フレームワークの高度化とメッシュフリー法への適用	森田直樹(筑波大)	○			
jh220048	HPCと高速通信技術の融合による大規模データの拠点間転送技術開発と実データを用いたシステム実証試験	村田健史(NICT)		○		

# 2022年度東大利用課題(OB外:OBCX外部接続)(3/3)

計算科学, 計算機科学・システム, データ・学習, 融合

課題番号	課題名	代表者(所属)	OB	OB外	W-O	W-A
jh220049 国際共同	Innovative Multigrid Methods II	藤井昭宏(工学院大)	○		○	
jh220052	時空間発展するシミュレーションを予測する代理モデルの開発	下川辺隆史(東大)				○
jh220054	ソフトマター流動の機械学習	John Molina(京大)				○
jh220055	数値シミュレーションと機械学習との融合による水圏生態系予測 (AI-for-HPC発展課題)	菊地淳(理研)			○	○
jh220057	統合機械学習分子動力学システムの構築	奥村雅彦(JAEA)	○		○	○
jh220058	次世代学術情報基盤に向けた基盤ソフトウェアの実践的な研究・ 開発・評価	杉木章義(北大)				○

- JHPCNの概要
- 東大情報基盤センターのスパコン概要
- Wisteria/BDEC-01
- Oakbridge-CX
- Ipomoea-01
- h3-Open-BDEC
  
- 利用事例
- **JHPCN利用申込み**

# 提案書(2023年度は若干変わる可能性あり)

- 3頁以内
  - 研究目的
  - JHPCN公募型共同研究として実施する必要性, 研究の意義, 研究計画
  - これまでの研究成果(継続の場合)
  - 拠点連携課題, 大規模データ・大容量ネットワークにおける具体的計画
  - 主な関連発表成果
- 資源利用計画等(別ファイル)
- 執筆のポイント
  - 明確な目標, 学際的な研究, 学術的意義, JHPCN共同研究としての意義(次頁(再掲))
  - 新規課題は準備が十分整っている(と思わせる)ことが重要
    - **これから並列化を始める, というのはやや難しいかも知れない: 個別にご相談ください**
    - 継続課題は前年度成果(特に中間報告書段階)をある程度重視
  - 審査は毎年実施されるが, 複数年計画として「研究計画」を記述することは可能
    - 特に新規課題は, 将来へ向けての全体像がより明確になる
  - 明確な資源利用計画(とは言え, なかなか事前に書くのは難しい)
    - ○ノード×○時間×○ケース

# 共同研究のポイント: 全部を満たす, ということではない

- 学際的研究体制
  - 情報科学分野(計算機科学・データ科学)と応用分野による, 学際共同研究の推進
  - 審査においては学際的な研究体制の構築がなされていることを重視
- ソフトウェアおよびデータ活用推進
  - 単なる公開だけでなく, 成果の幅広い利活用を目指す取り組みを高く評価
- IT 基盤技術開発
  - アーキテクチャー, システムソフトウェア, セキュリティ等, IT 技術の基盤的研究につながる課題を高く評価: 各構成研究拠点の研究者との共同研究
- 拠点連携
  - 複数構成拠点の資源を活用, 複数拠点の研究者との連携体制
  - 広域分散型の大規模情報システム, アプリケーションのマルチプラットフォーム実装
- 大規模データ・大容量ネットワーク利用
- 計算機資源の利用のみではなく, 研究開発的要素が大きい課題を採択

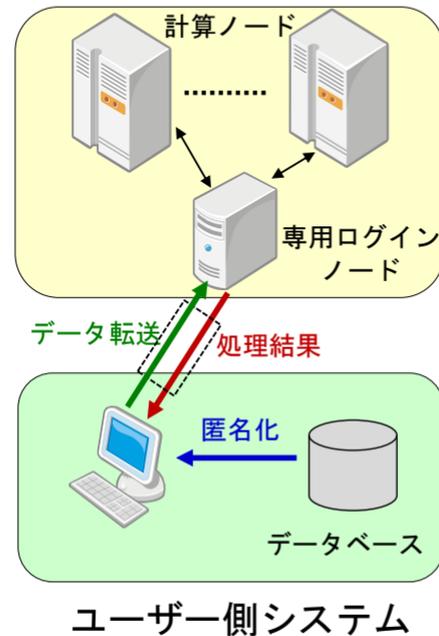
# 利用可能な資源

- スパコン等: 他大学(他拠点)の資源との併願も可能(推奨)
  - Oakbridge-CX(OBCX)
    - <https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/>
    - (要相談)外部接続ノード, Mini-DP利用
  - Wisteria/BDEC-01(Odyssey, Aquarius)
    - <https://www.cc.u-tokyo.ac.jp/supercomputer/wisteria/service/>
    - (要相談)Odyssey-Aquarius連携, 外部接続(Aquarius), Wisteria-Messenger利用
  - Ipomoea-01
    - 規定に応じて各ユーザーに配分(予定)
- 基本的に資源の追加はできない
  - 自前の予算で追加することは可能(審査有り)
- 利用ノード数の上限があるが, 1ヶ月に一回「大規模HPCチャレンジ(後述)」で, より大規模な資源を利用可能(後述)
  - 提案書に利用計画として記述することは可能
- Aquariusは「ノード固定」可能

# ノード固定

[https://www.cc.u-tokyo.ac.jp/supercomputer/reedbush/service/fixed\\_node.php](https://www.cc.u-tokyo.ac.jp/supercomputer/reedbush/service/fixed_node.php)

- ライセンスサーバが必要な商用プログラムを利用したい
- データに対するセキュリティを確保するために、専用HDを使用したい
- データ転送用に専用通信回線を敷きたい
- 一般バッチジョブキューでは対応できないような長時間ジョブを実行したい
- **Aquariusは「ノード固定」可能: 事前に必ずご相談ください**



# Oakbridge-CX (1/2)

- 概要
  - Intel Platinum 8280 (Cascade Lake) 搭載
  - 全1,368ノードのうち128ノードに高速SSD搭載
    - うち16ノードは外部ネットワークに直接接続する「外部接続ノード」
  - ネットワーク経由でssh, scp/sftpを利用したログインノードアクセス, ファイル転送
  - 高速なNVMe-SSD搭載ノード, BeeGFS on Demand (BeeOND)によるオンデマンドな単一の共有ファイルシステムを用いたビッグデータ解析
- 特殊な利用形態 (要相談)
  - 外部接続ノードを利用して, 外部リソース (ストレージ, サーバー, データベース, センサーネットワーク等) に直接接続, リアルタイムデータ取得
  - ノード固定, ログインノード持込・カスタマイズ
  - SINET L2VPN経由でログインノード, 外部接続ノードへの直接接続
  - 東大センター担当窓口 ([uketsuke@cc.u-tokyo.ac.jp](mailto:uketsuke@cc.u-tokyo.ac.jp)) 宛てご相談ください).

# Oakbridge-CX (2/2)

- ソフトウェア

- 【言語コンパイラ】

- Fortran, C, C++

- 【ライブラリ】

- MPI, BLAS, LAPACK/ScaLAPACK, FFTW, PETSc, METIS/ParMETIS

- 【アプリケーションソフトウェア】

- OpenFOAM, ABINIT-MP, PHASE, FrontFlow/Blue, FrontISTR, REVOCAP, ppOpen-HPC, **h3-Open-BDEC, HyperWorks, MATLAB**

- 【コンテナ仮想化】

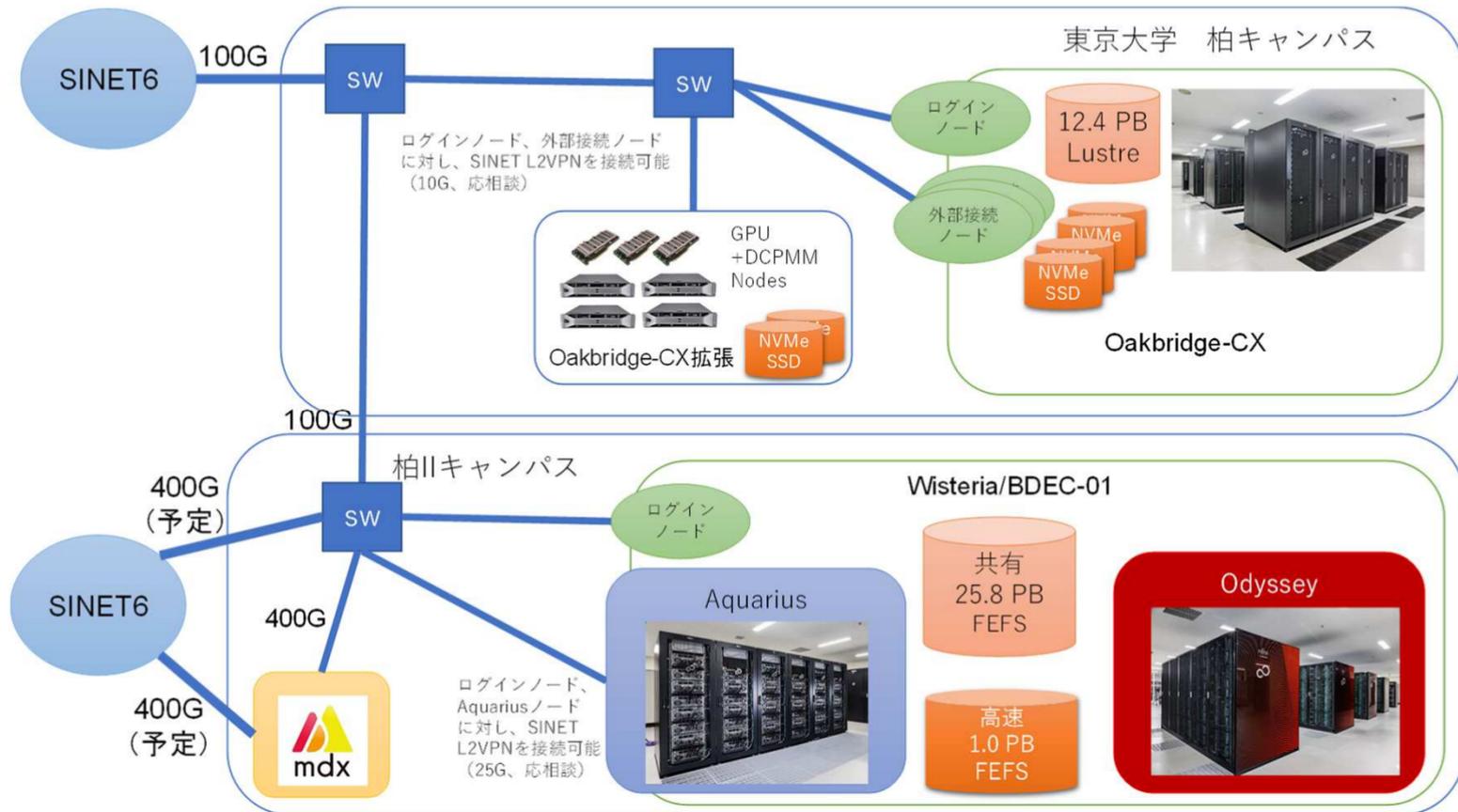
- singularity (docker image可能)

# 大規模HPCチャレンジ

<https://www.cc.u-tokyo.ac.jp/guide/hpc/>

- 月1回(月末)実施, 公募制, 無料
  - グループまたは個人
  - 基本的に居住者をグループ代表とすることを前提
  - 現ユーザーには限定せず, 大学・研究機関・企業から幅広く募集
  - 外部専門家を含む審査委員会で審査
- **全系を24時間占有利用, 年3回募集(3~4ヶ月分をまとめて募集)**
  - **2022年度は8時間(09:00-17:00)で実施**
- Oakbridge-CX
  - 上限: 1,280ノード
- Wisteria/BDEC-01
  - Odyssey(6,144ノード)とAquarius(36ノード)は(当面は)別々に実施
- 成果公開: ニュースレターへの寄稿, セミナー等での発表が義務

# 利用可能な資源とSINETとの接続



# Wisteria/BDEC-01 (1/3)

- 概要

- Odyssey: シミュレーションノード群, Fujitsu A64FX 7,680ノード搭載
- Aquarius: データ・学習ノード群, Intel Xeon Platinum 8360Y・NVIDIA A100 Tensorcore 45ノード搭載(ノード当たり8 GPU, 合計360 GPU)
- ネットワーク経由でssh, scp/sftpを利用したログインノードアクセス, ファイル転送
- 高速なファイルシステムを用いたビッグデータ解析が可能
- OdysseyとAquariusを両方使用したい場合は, それぞれ別々に申し込む

- 特殊な利用形態(要相談)

- Aquariusの各ノードは, 外部リソース(ストレージ, サーバー, データベース, センサーネットワーク等)に直接接続可能
- Odyssey・Aquarius連携による「計算・データ・学習」融合(h3-Open-BDEC)
- SINET L2VPN経由でログインノード, 外部接続ノードへの直接接続
- 東大センター担当窓口(uketsuke@cc.u-tokyo.ac.jp)宛てご相談ください).

# Wisteria/BDEC-01 (2/2)

- ソフトウェア

- 【言語コンパイラ】

- Fortran, C, C++

- 【ライブラリ】

- MPI, BLAS, LAPACK/ScaLAPACK, FFTW, PETSc, METIS/ParMETIS

- 【アプリケーションソフトウェア】(一部GPU化非対応)

- OpenFOAM, ABINIT-MP, PHASE, FrontFlow/Blue, FrontISTR, REVOCAP, ppOpen-HPC, **h3-Open-BDEC, HyperWorks (A), MATLAB (A)**

- 【コンテナ仮想化】

- singularity (docker image可能)

# 利用カテゴリ

- ①OBCX
  - ②OBCX(外部接続ノード)
  - ③Odyssey
  - ④Aquarius(一般利用)
  - ⑤Aquarius(GPU専有)
- Odyssey, Aquariusを両方利用したい場合は, ③及び④または⑤をそれぞれ申し込む, ③・④・⑤を全て申し込むことも可能

# 2023年度計算資源(東大分)

ストレージはノード時間・GPU時間に比例して配分

Wisteria/BDEC-01, Ipomoeaは各システムの合計が配分される

**Ipomoeaはグループ毎配分の外, 各ユーザーも5TBまで無料で利用可能**

NH: ノード時間, GH: GPU時間

	上限	ストレージ	Ipomoea (グループ)	単価	実行ジョブ上限
OBCX	43,200 NH (10ノード・半年)	4.00TB (4,320NH当たり)	0.60TB (4,320NH当たり)	17.36円/NH	256ノード
OBCX (外部接続ノード)	4,320 NH (1ノード・半年)			26.04円/NH	1ノード
Odyssey	276,480 NH (32ノード年)	2.00TB (1ノード年当たり)	0.30TB (1ノード年当たり)	10.41円/NH	2,304ノード
Aquarius (一般利用)	95,040 GH (11GPU年)	6.00TB (1GPU年当たり)	0.90TB (1GPU年当たり)	31.25円/GH	8ノード(64GPU)
Aquarius (GPU専有)	69,120 GH (8GPU年) 1・2・4・8 必ず8,640の倍数			42.19円/GH	GPU専有数分

# 2023年度計算資源

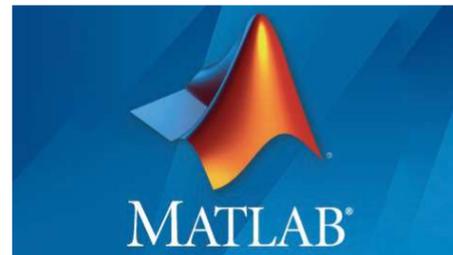
Excel

- 要求資源量の上限値が明文化された
  - 複数センター計算資源を利用の場合は負担金360万円相当
  - 各センターについては300万円相当
  - mdxの利用は, 基本的に一律100万円相当
    - スパコンとmdxを両方利用する場合には, min(360万-スパコン合計値, 100万)
- 東大情報基盤センターのスパコンの負担金
  - 2023年度は2022年度の50%増し(一律), 全体の予算が変わらないために, 利用できる資源量は従来の1.5分の1程度となる
  - Oakbridge-CX(1ノード年, 8,640ノード時間): 10万円⇒15万円
  - Wisteria/BDEC-01(Odyssey)(1ノード年): 6万円⇒9万円
  - Wisteria/BDEC-01(Aquarius)(1GPU年): 18万円⇒27万円
    - 8GPU専有: 216万円⇒291.6万円: ほぼ予算を使い切ってしまう
- **Oakbridge-CX(OBCX)は2023年9月末運用停止**

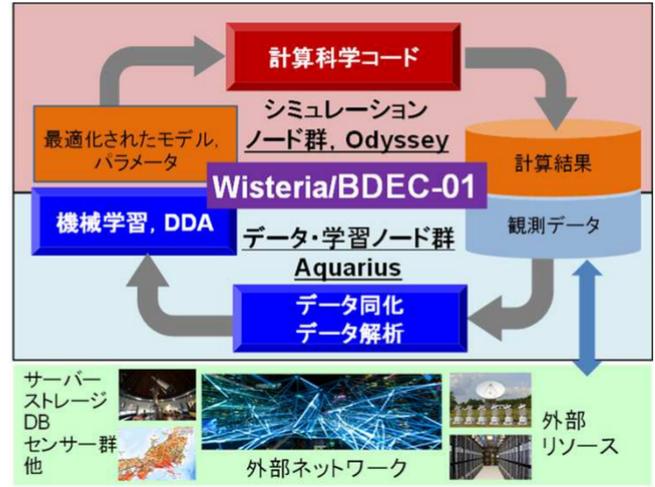
# スパコン利用にあたっての指針(1/3)

## 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
    - OBCX, Aquarius
    - 国内大学教職員・学生のみ利用可能



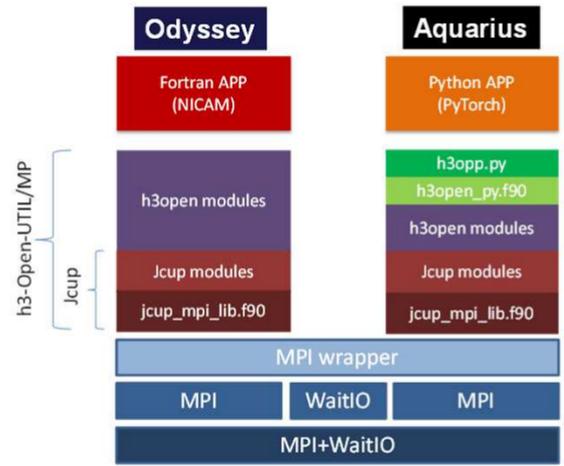
# MATLABの導入 「S+D+L」融合, AI for HPCの実現



- 2022年3月からOBCX, Aquariusで利用可能
- MATLAB

- ✓ 多様な機能
- ✓ ユーザーのプログラムからの関数呼び出し重視⇒データ解析, 機械学習系の豊富な機能⇒高度化
- ✓ MATLABはAquarius(データ・学習ノード群)でのみ稼働するが, h3-Open-BDECと連携させて, Odyssey(シミュレーションノード群)上で実施する大規模シミュレーションのパラメータ最適化に適用する⇒「S+D+L」融合, AI for HPC

- h3-Open-BDECは様々な環境で動作 ⇒MATLABと組み合わせた使用による普及



# スパコン利用にあたっての指針(2/3)

## OBCX, Odyssey, Aquarius

- 計算科学・大規模シミュレーション(S)
  - データ科学(D)
  - 機械学習・AI(L)
  - 「S+D+L」融合
- 
- 全てのシステム(OBCX, Odyssey, Aquarius)がそれぞれの項目に対応可能
    - Aquarius(データ・学習ノード)でもシミュレーションはできる
  - データ科学(D), 機械学習・AI(L)
    - コンテナ仮想化(Singularity)により対応

# スパコン利用にあたっての指針(3/3)

## OBCX, Odyssey, Aquarius

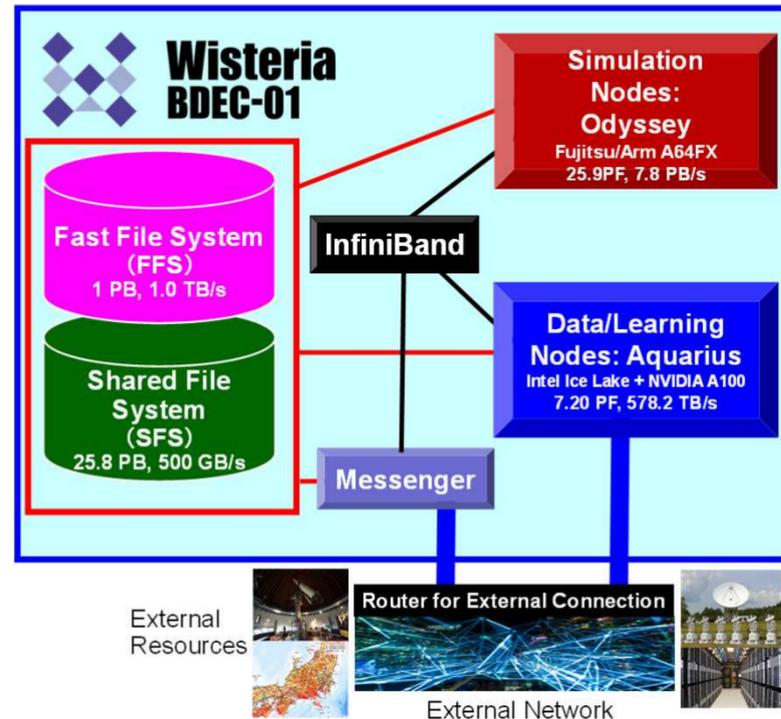
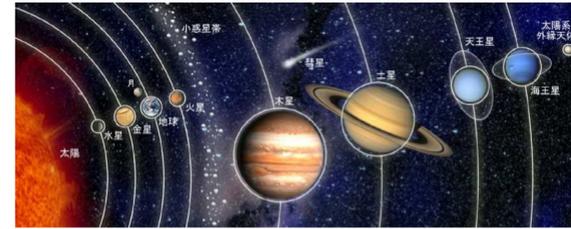
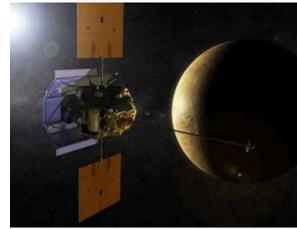
	OBCX	Odyssey	Aquarius	O+A
計算科学	◎	◎	◎	-
データ科学	◎	◎	◎	-
機械学習・AI	○	○	◎	-
「S+D+L」融合	○	○	◎	◎
商用コード・ MATLAB等	○	×～△	○	-
その他の特徴	<ul style="list-style-type: none"> <li>Intel Xeon CPUのため特殊なチューニングは必要なく、そこそこの性能が出る</li> <li>外部接続ノード, SSD搭載ノード</li> </ul>	<ul style="list-style-type: none"> <li>A64FX(Arm)</li> <li>チューニング必須</li> <li>FP16</li> <li>商用コードへの対応がやや遅れている</li> </ul>	<ul style="list-style-type: none"> <li>CPU(Ice Lake): 高い推論性能</li> <li>GPU(A100): Tensor Core + Tensor Float [TF32]</li> <li>超大規模シミュレーションには不向き</li> </ul>	<ul style="list-style-type: none"> <li>O-A連携(h3-Open-BDEC, WaitIO, Coupler)</li> <li>2022年6月から利用可能</li> </ul>

# JHPCN共同研究課題への応募

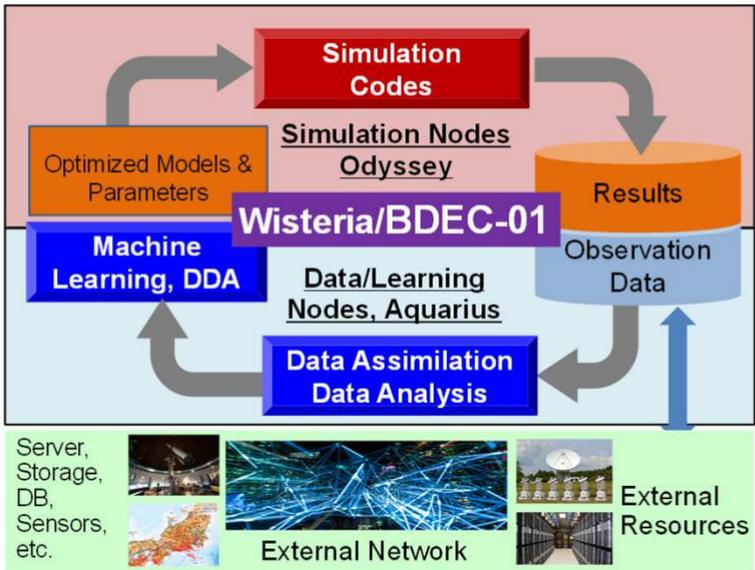
- 計算科学アプリケーションの並列化
- 計算科学アプリケーションの最適化
  - 「GPU化」はテーマとなりうるが、MPI・OpenMPについては要相談
  - Odyssey向け最適化は十分それだけでテーマとなりうる
  - 得られた知見を利用者へ積極的に公開(Odyssey:名古屋大とも情報を共有)
- マルチフィジックス連成
- 「S+D+L」融合(S+D, S+L)
  - h3-Open-BDEC, h3-Open-UTIL/MP, WaitIO-Socket/File活用:検証, 改良に貢献
- 特に, リアルタイムデータ処理を含む場合
  - OBCX(外部接続ノード), Aquarius(またはMessenger)
  - 大規模並列処理が必要となるような場合を想定
  - 単にデータを取得, 貯蔵する場合は「分野2」へ応募されるのが良いかも知れない
- 困ったことがありましたら, お気軽にご相談ください。(募集要項が)威圧的な書き方ですみません。

# Wisteria-Messenger

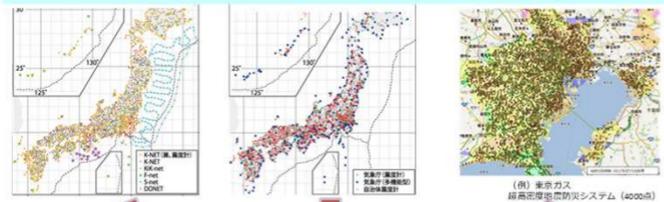
- Intel Xeon Gold 6348 (IceLake)
  - 2.6 GHz, 28 cores/socket
  - 2 sockets × 6 nodes
  - IB-HDR
- Directly working with “Odyssey” for supporting “Aquarius”
  - can access external resources directly
  - e.g. “Filtering” in the Earthquake Simulation
  - h3-Open-SYS/WaitIO-Socket
- Operation has started
  - JHPCN users can access
  - Plan for public use is pending
- 本当は「Mercury」という計画があって、そのプロトタイプとしてMessengerがある。



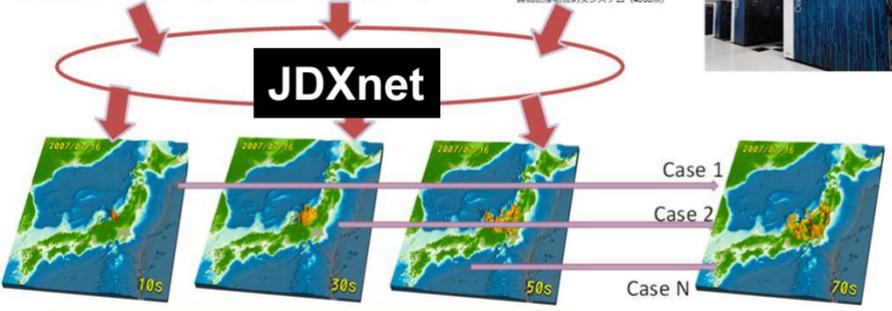
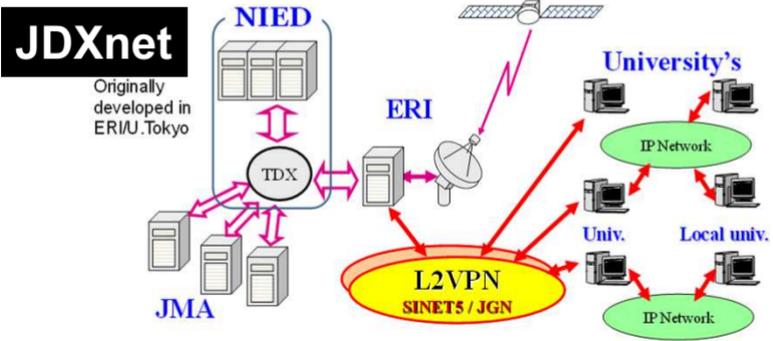
# リアルタイムデータ同化+ 3D強震動シミュレーション融合 JDXnetによるリアルタイム観測データ活用



Observation Network for Earthquake:  $O(10^5)$  Points



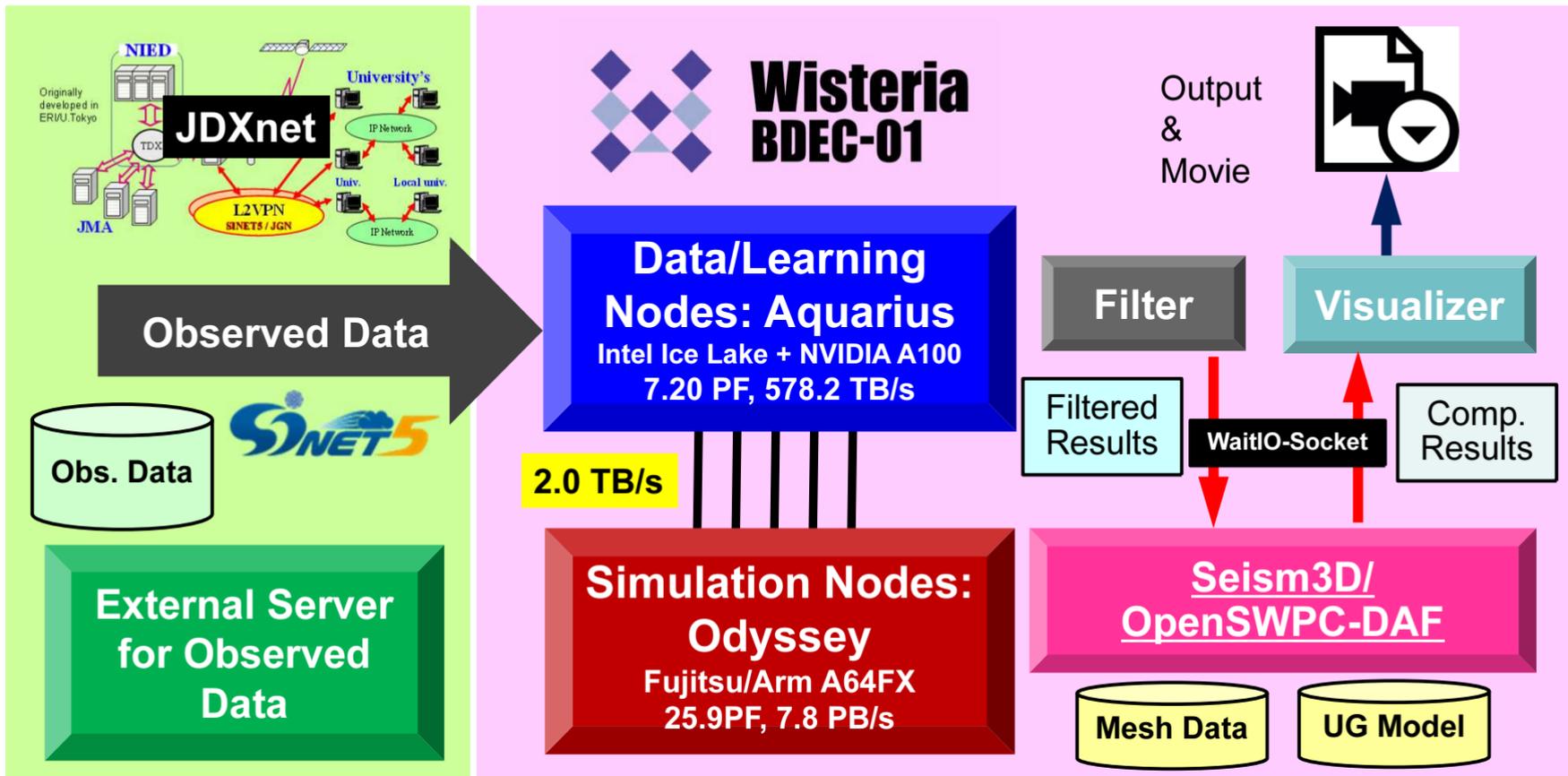
[c/o Furumura]



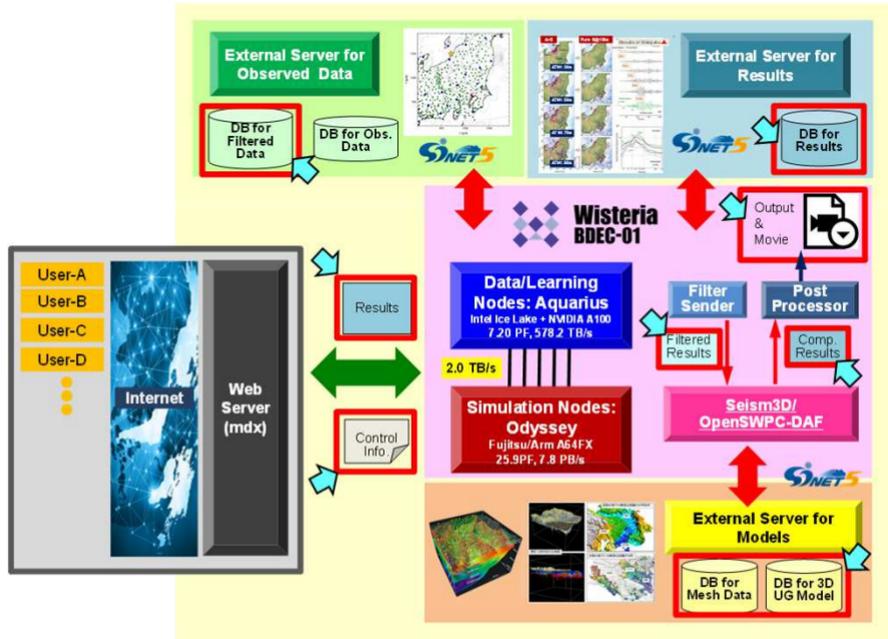
Real-Time Data/Simulation Assimilation  
Real-Time Update of Underground Model

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

# 長周期地震動シミュレーション+観測データ同化



# Webベース シミュレーション体験・ データ利活用システム mdxとの連携事例



• 「3D長周期地震動+リアルタイムデータ同化」融合シミュレーションシステムの「防災・減災」啓蒙・教育へ向けた利用・展開を図るため、Webベースのシミュレーション体験・データ利活用環境を構築(2022年度)

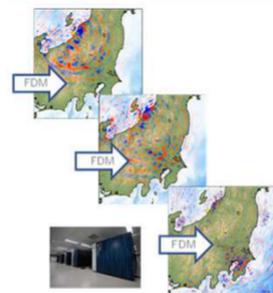
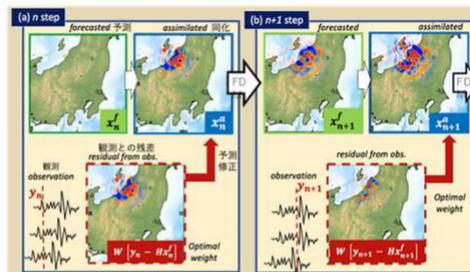
• 利用者はWeb Server(mdx上)にアクセスし、スパコン(Wisteria/BDEC-01)上でのシミュレーションの実施、計算結果、観測結果の可視化処理、表示等を行う。

• Web経由でデータ群をスパコン上で処理するフレームワークは様々なアプリケーションへの転用が可能

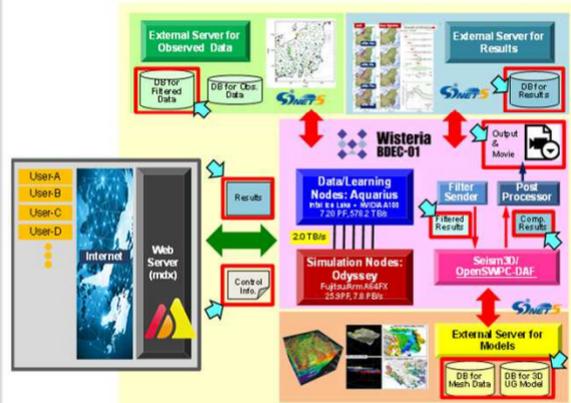
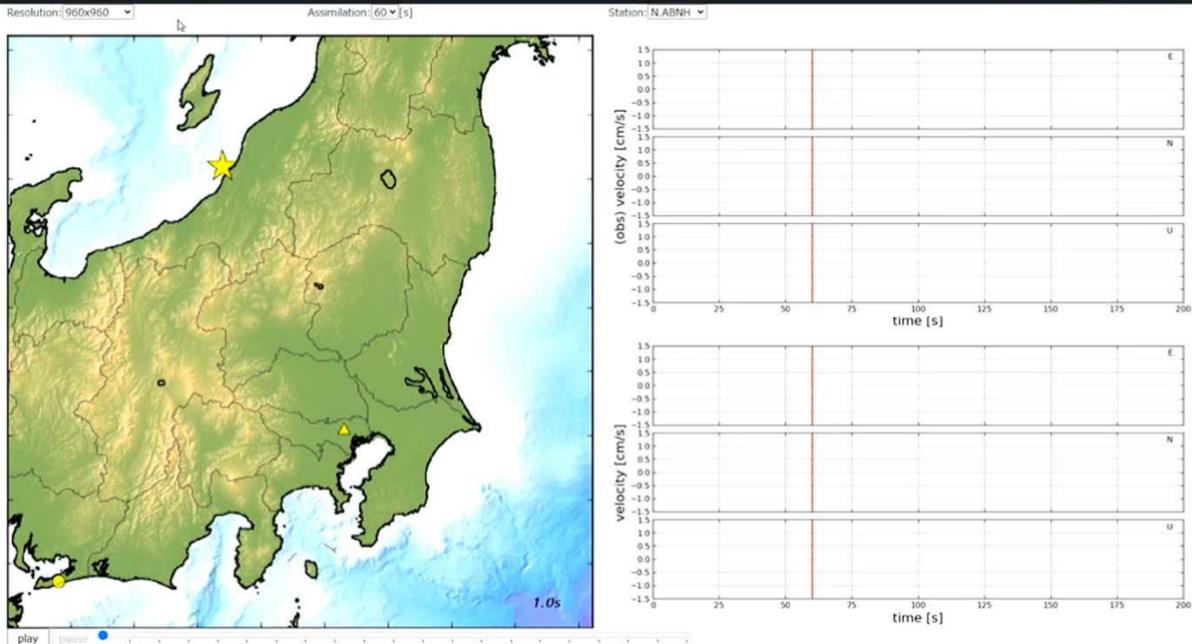
# mdxからWeb経由で大規模シミュレーション・データ同化をインタラクティブに実行

(A+S) Assimilation+Simulation

(Pure S) Pure Simulation/Forecast

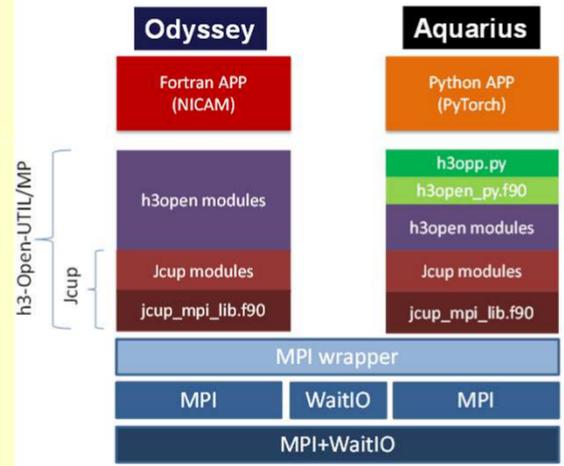
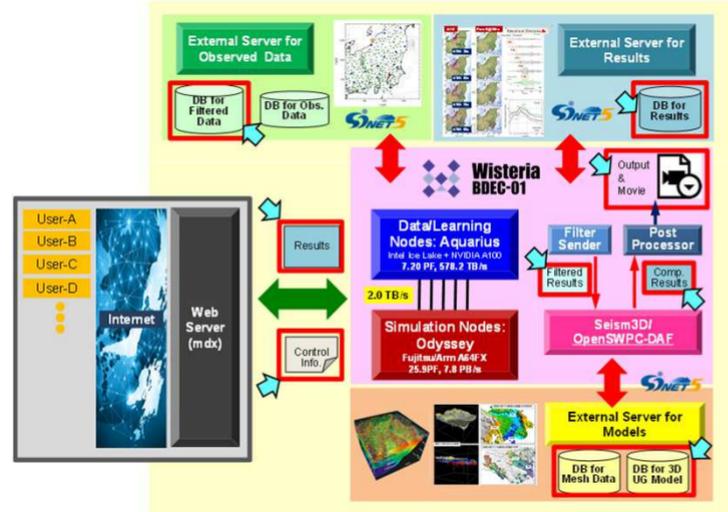
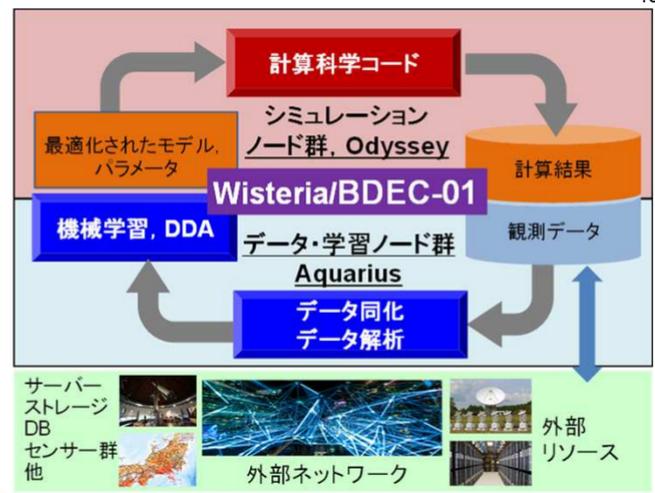


Chuetsu offshore earthquake, 2007



# Odyssey-Aquarius連携, Wisteria-mdx連携

- 検討されておられる方は是非お気軽にご相談ください。
  - h3-Open-SYS/WaitIO, h3-Open-UTIL/MP
- uketsuke@cc.u-tokyo.ac.jp



- JHPCNの概要
- 東大情報基盤センターのスパコン概要
- Wisteria/BDEC-01
- Oakbridge-CX
- Ipomoea-01
- h3-Open-BDEC
  
- 利用事例
- JHPCN利用申込み
- **番外編:代替制度**

# JHPCNは少しハードルが高いと感じられる場合

- 無料・審査無し
  - お試しスパコン利用(無料体験)
  - お試しアカウント付き並列プログラミング講習会
- 無料・審査あり: JHPCNの一步手前の位置づけ
  - 若手女性利用
  - AI for HPC: 特にこれから「計算+データ+学習」融合にとりかかろうとしている場合
- 有料・報告義務あり
  - トライアル利用(通常の30%の負担金)
  - 企業の場合は事前審査あり, 最初の3ヶ月間は無料
- 「企業利用」も含めて2023年1月下旬に来年度公募説明会を予定している

# 若手女性・AI-for-HPC (○:代表者, △:参加者)

制度名	種別	大学等	企業	学生	個人	審査	無料	報告書	A	B	C	D	備考	募集
通常利用	一般	○	△	○					✓	✓	✓			随時
	トライアル	○	△	○				✓			✓	✓	年度内	随時
お試し利用		○	○	○	✓		✓				✓	✓	1ヶ月限定	随時
JHPCN		○	○	△		書類	✓	✓		✓				年1回(1月)
HPCI	一般・若手	○	△	△		書類	✓	✓		✓				年1回(10-11月)
	産業		○			書類	✓	✓		✓				
若手女性	一般	○	○	○	✓	書類	✓	✓		✓	✓			年2回(8・2月)
	インターン			○	✓	書類	✓	✓			✓			年1回(夏季)
AI for HPC		○	○	△		書類	✓	✓		✓	✓			随時, 年4回審査
HPCチャレンジ		○	○	○		書類	✓	✓						年数回
講習会		△	△	△	✓		✓						1ヶ月有効UID	年20回程度
教育利用		○	○	○		書類	✓	✓					企業研修等可	随時
企業利用	一般	△	○	△		+面接		✓		✓				年2回(8・2月)
	トライアル	△	○	△		+面接	一部	✓			✓	✓	3ヶ月無料 年度内	随時, 年4回審査

(A:トークン移行, B:ノード固定, C:Odyssey⇔Aquarius移行可能, D:1システム1回限り応募可能)

# 若手・女性利用

<https://www.cc.u-tokyo.ac.jp/guide/young/>

- 対象
  - 若手研究者(各年度4月1日現在において40歳以下), 女性研究者, 学生
- 一般枠(1月, 7月の年2回募集): 2022年からmdx利用可能
  - 個人研究, 企業もOK
  - 非居住者は原則認めないが, 来日6ヶ月未満の留学生@国内は可能(閲覧制限等有)
  - 4月開始: 1年または半年, 10月開始: 半年
- インターン制度
  - 学部学生・大学院生を対象(個人研究)
  - 期間中1週間程度の東大センター滞在を想定していたが, オンラインへ移行の予定
  - グループ利用は2022年度から廃止(「教育利用」へ移行)
- 成果公開
  - 報告書(ニューズレターへの寄稿), 報告会
- 特に優れた一般枠課題は, JHPCN萌芽型共同研究課題として推薦
  - 7月のJHPCNシンポジウムでポスター発表できる

# 若手・女性利用(応募可能な計算機資源・負担金未定)

<https://www.cc.u-tokyo.ac.jp/guide/young/>

Wisteria/BDEC-01 Odyssey	Wisteria/BDEC-01 Aquarius	Oakbridge-CX	mdx <a href="https://mdx.jp/mdx/">https://mdx.jp/mdx/</a>
<b>6ヶ月(90,000円相当)</b> ① 8,640(2ノード×6ヶ月) ② 2,304ノード ③ 1.00(1ノード) ④ 4TB	<b>6ヶ月(135,000円相当)</b> ① 12,960(1GPU×6ヶ月) ② 8ノード(64GPU) ③ 3.00(GPU) ④ 6TB	<b>6ヶ月(150,000円相当)</b> ① 8,640(2ノード×6ヶ月) ② 256ノード ③ 1.00(1ノード) ④ 8TB	✓ 608vCPU(4ノード相当) ✓ 1GPU ✓ 仮想ディスク 100GB ✓ 高速内部ストレージ 1TiB ✓ 大容量内部ストレージ 2TiB ✓ オブジェクトストレージ 2TiB ✓ グローバルIPアドレス 1個
<b>12ヵ月(180,000円相当)</b> ① 17,280(2ノード年) ② 2,304ノード ③ 1.00(1ノード) ④ 4TB	<b>12ヶ月(270,000円相当)</b> ① 25,920(1GPU年) ② 8ノード(64GPU) ③ 3.00(GPU) ④ 6TB	<b>12ヵ月(300,000円相当)</b> ① 17,280(2ノード年) ② 256ノード ③ 1.00(1ノード) ④ 8TB	
<b>ディスク容量追加(1TB)</b> 6ヶ月:3,240円 12ヶ月:6,480円	<b>ディスク容量追加(1TB)</b> 6ヶ月:3,240円 12ヶ月:6,480円	<b>ディスク容量追加</b> 6ヶ月:3,240円 12ヶ月:6,480円	<b>ディスク容量追加料金は検討中 (最大50%値上げ)</b>

①トークン量上限, ②並列実行ノード数(最大), ③トークン消費係数, ④ディスク容量

# AI for HPC: Society 5.0実現へ向けた人工知能・データ科学による計算科学の高度化(試行)(1/2)

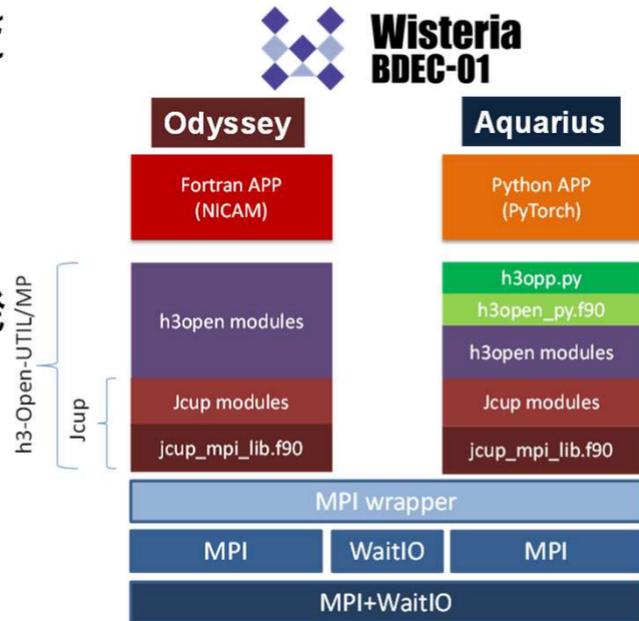
<https://www.cc.u-tokyo.ac.jp/guide/exploratory/AIforHPC/>

- (計算+データ+学習(S+D+L))融合実現, データ科学, 機械学習, 人工知能による計算科学の高度化を目指す提案を募集
- 原則として, 計算科学シミュレーション(自作またはオープンソース)を, データ科学, 人工知能, 機械学習等によって高度化, 効率化することを目的とする
  - 大規模データ同化と人工知能を融合するような研究も受け付ける。
  - プログラム本体のチューニング, アルゴリズム高度化などは対象外ですが, 自動チューニングによって最適アルゴリズムを選択するような提案は歓迎いたします。
- 応募者グループ・センター教員の共同研究として実施
  - 代表者: 居住者(大学・研究機関・企業), メンバー: 非居住者参加も可能
  - 次年度以降JHPCN共同研究課題応募を目指す

# AI for HPC: Society 5.0実現へ向けた人工知能・データ科学による計算科学の高度化(試行)(2/2)

<https://www.cc.u-tokyo.ac.jp/guid/e/exploratory/AIforHPC>

- 随時募集, 年4回審査(変更)
- 計算機資源を無償で提供(負担金50万円相当(変更の可能性あり))
  - OBCX, Wisteria/BDEC-01
  - Wisteria/BDEC-01 (Odyssey+Aquarius) 利用推進
- Wisteria/BDEC-01向けソフトウェア群の共同開発
  - h3-Open-BDEC
- 成果公開
  - 報告書(ニューズレターへの寄稿), 報告会(若手・女性と合同, JHPCNに採択の場合は免除)



# これまでの採択課題

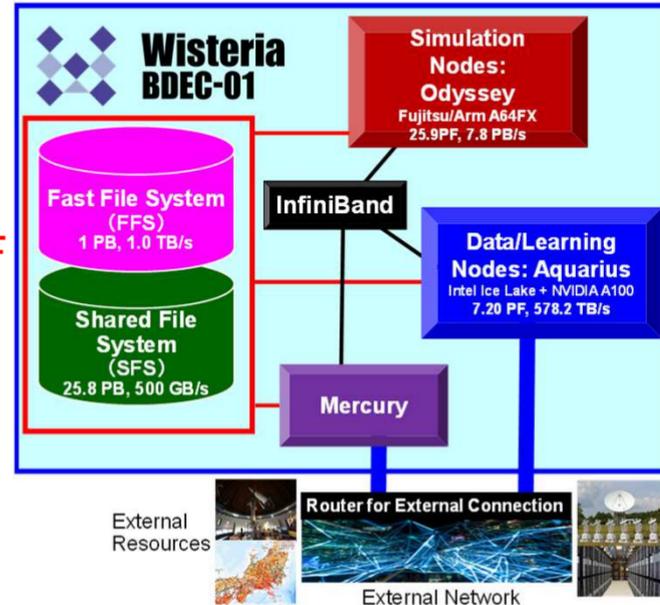
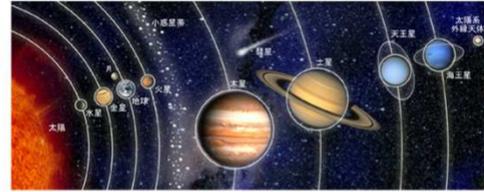
## 計算＋データ同化＋機械学習融合

年度	代表者	課題名	使用計算機	備考
2020	澤田洋平 (東京大学工学系 研究科・准教授)	地球科学シミュレーションの 不確実性定量化の新展開	Oakforest-PACS	
2021	澤田洋平 (東京大学工学系 研究科・准教授)	超巨大アンサンブル 計算と機械学習の協調によ る地球科学シミュレーション の不確実性定量化	Wisteria/BDEC-01 (Odyssey) Wisteria/BDEC-01 (Aquarius)	h3-Open-BDECの提供する Odyssey-Aquarius連携ライ ブラリの開発にも貢献しても らう予定
2021	菊地淳(理化学研 究所環境資源科学 研究センター・チー ムリーダー)	数値シミュレーションと機械 学習との融合による東京 湾の赤潮予測	Wisteria/BDEC-01 (Aquarius) Oakbridge-CX	
2022	川崎猛史(名古屋 大学 大学院理学 研究科・講師)	機械学習を用いた非晶質固 体系の破壊予測システムの 開発	Wisteria/BDEC-01 (Aquarius) Oakbridge-CX	

- JHPCNの概要
- 東大情報基盤センターのスパコン概要
- Wisteria/BDEC-01
- Oakbridge-CX
- Ipomoea-01
- h3-Open-BDEC
  
- 利用事例
- JHPCN利用申込み
- **番外編:GPU移植への取り組み(OFP-II)**

# 将来構想: OFP-II, Mercury

- スパコンへの性能要求, 省電力, 脱炭素化⇒演算加速器搭載は不可避
- **Wisteria-Mercury (2023年10月～)**
  - GPUクラスタ
  - OFP-IIプロトタイプ
- **OFP-II (2024年4月)**
  - OFP後継機 (JCAHPC: 筑波大学と共同), 200+PF
  - GPUクラスタ
    - Mercuryと同じ, もしくはその後継GPU
  - **CPU Onlyのノード群もあり**
- **アプリケーションの移植が必要 (3,000人)**
  - 講習会, GPUミニキャンプによる対応



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  
200+ PF

Fujitsu FX10  
Oakleaf-FX  
1.13 PF

Wisteria Fujitsu  
BDEC-01  
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

Mercury

Ipomoea-01 25PB

Ipomoea-03

Ipomoea-02

# スケジュール概要

## Mercury & OFP-II



筑波大学  
University of Tsukuba



東京大学  
THE UNIVERSITY OF TOKYO

- GPU移行のための諸作業は遅くとも、2022年秋に始める必要がある
- それ以前にMercury・OFP-IIIに搭載するGPU(両者は同じ)を決める必要あり
- 2022年2月～3月
  - プリベンチマークを各社に依頼(計算科学系7種類, Fortran, C)
- 2022年6月
  - GPUベンダーを決定
  - ポイント:性能, ポーティングのしやすさ, サポート体制, Fortranへの対応
- 2022年秋～
  - ポーティング開始, 当初はAquariusをプラットフォームとして使用(多分12月頃)
- 2023年秋～
  - Mercuryを使用した最適化, 評価
- 2024年4月:OFP-II運用開始

# お試しアカウント付き並列プログラミング講習会

<https://www.cc.u-tokyo.ac.jp/events/lectures/>

[https://www.youtube.com/playlist?list=PLobjSv\\_ny85kXY2Mtnhn1k7pM-epQaD2y](https://www.youtube.com/playlist?list=PLobjSv_ny85kXY2Mtnhn1k7pM-epQaD2y)

- Wisteria/BDEC-01, Oakbridge-CXを利用した講習会
- 2020年度から完全オンラインで実施: 年20回程度, 概して好評(移動不要)
- 既存利用者に限定せず, 企業の技術者・研究者も受講可能
  - ✓ 受講者の1/2~2/3は企業から受講: 裾野拡大に大きな貢献
  - ✓ PCクラスタコンソーシアム(実用アプリケーション部会)と共催
- 1~2日間の講習, 1ヶ月有効な「お試しアカウント」
  - ✓ スパコン超入門: Linuxの使い方
  - ✓ MPI基礎, MPI応用(並列有限要素法), マルチコアプログラミング
  - ✓ GPUプログラミング, GPUミニキャンプ(ハッカソン)
  - ✓ ライブラリ利用(センター教員開発のライブラリ普及)(H行列, WaitIO/Coupler)
  - ✓ OpenFOAM(初級, 中級)
  - ✓ Altair HyperWorks
  - ✓ 利用ノード数, 実行時間に制限あり, スパコンを使用しない講義もある

# 講習会 (○:代表者, △:参加者)

制度名	種別	大学等	企業	学生	個人	審査	無料	報告書	A	B	C	D	備考	募集
通常利用	一般	○	△	○					✓	✓	✓			随時
	トライアル	○	△	○				✓			✓	✓	年度内	随時
お試し利用		○	○	○	✓		✓				✓	✓	1ヶ月限定	随時
JHPCN		○	○	△		書類	✓	✓		✓				年1回(1月)
HPCI	一般・若手	○	△	△		書類	✓	✓		✓				年1回(10-11月)
	産業		○			書類	✓	✓		✓				
若手女性	一般	○	○	○	✓	書類	✓	✓		✓	✓			年2回(8・2月)
	インターン			○	✓	書類	✓	✓			✓			年1回(夏季)
AI for HPC		○	○	△		書類	✓	✓		✓	✓			随時, 年4回審査
HPCチャレンジ		○	○	○		書類	✓	✓						年数回
<b>講習会</b>		<b>△</b>	<b>△</b>	<b>△</b>	<b>✓</b>		<b>✓</b>						<b>1ヶ月有効UID</b>	<b>年20回程度</b>
教育利用		○	○	○		書類	✓	✓					企業研修等可	随時
企業利用	一般	△	○	△		+面接		✓		✓				年2回(8・2月)
	トライアル	△	○	△		+面接	一部	✓			✓	✓	3ヶ月無料 年度内	随時, 年4回審査

(A:トークン移行, B:ノード固定, C:Odyssey⇔Aquarius移行可能, D:1システム1回限り応募可能)

# MPI+OpenMPで並列化されたFortranプログラムのGPUへの移行手法(オンライン)

<https://www.cc.u-tokyo.ac.jp/events/lectures/196/>

## 趣旨

- 本講習会ではOFP-IIへのプログラム移植に向け、既存のMPI+OpenMPでCPU向けに並列化されたFortranプログラムのGPU環境への移植手法を学ぶ
- 受講料は無料
- GPUへの移行は基本的にOpenMP(CPU向け)で並列化されたループを対象として実施
- GPU向けのループ並列化の手段としてはOpenACC, OpenMP 5.x, do concurrent, CUDA Fortranなど様々であるため、それぞれの特徴や使い方を学ぶ
- Wisteria/BDEC-01(Aquarius)を活用し、OpenMPでCPU向けに並列化された有限要素法のプログラムを題材とした、GPUへの移植の演習を行う

**12月7日午後開催(オンライン)(申込締切:11月30日!!)**

# GPUミニキャンプ(オンライン)(1/2)

<https://www.cc.u-tokyo.ac.jp/events/lectures/197/>

## • ミニキャンプとは？

- 参加者がコードやデータセットを持ち込み、CUDA、OpenACC、Deep Learning など、GPUに関連した課題に対して、メンターからの助言を受けながら、その課題解決に取り組みます。

## • メンター

- エヌビディア合同会社、東大情報基盤センターなどから参加し、参加者の課題解決にご協力します。



## • 次回は12月12日～19日に実施(申込締切:11月30日!!)

- 各チームで実践中は、ベストエフォートでメンターがQ&A対応
- 年度内にもう1回、来年度は4回以上実施の予定

# GPUミニキャンプ(オンライン)(2/2)

<https://www.cc.u-tokyo.ac.jp/events/lectures/197/>

- 実施形式
  - Zoom と Slack を使ったオンライン形式
- 参加資格
  - 国公立大学・高専の教員・学生・研究生
  - 研究機関研究員
  - 企業に所属する研究者・技術者(非営利目的に限る)
  - 作業に必要なコードおよびデータセット等をセンターに持ち込める方。
  - コマンドラインによるLinux上での作業やエディタ利用に支障のない方
- 2022年度は、企業利用に発展した事例もあり

# 2023年度向け利用説明会(オンライン)(予定)

- 2022年11月29日(JHPCN)
- 2023年1月下旬(企業, 若手・女性, AI-for-HPC)
- 2023年7月下旬(企業, 若手・女性, AI-for-HPC)



**Wisteria  
BDEC-01**



- 諸制度

- 若手・女性 <https://www.cc.u-tokyo.ac.jp/guide/young/>
- 企業利用 <https://www.cc.u-tokyo.ac.jp/guide/company/>
- AI-for-HPC <https://www.cc.u-tokyo.ac.jp/guide/exploratory/AlforHPC/>
- 講習会 <https://www.cc.u-tokyo.ac.jp/events/lectures/>