

AIとスパコンが生み出す次世代 の科学技術シミュレーション

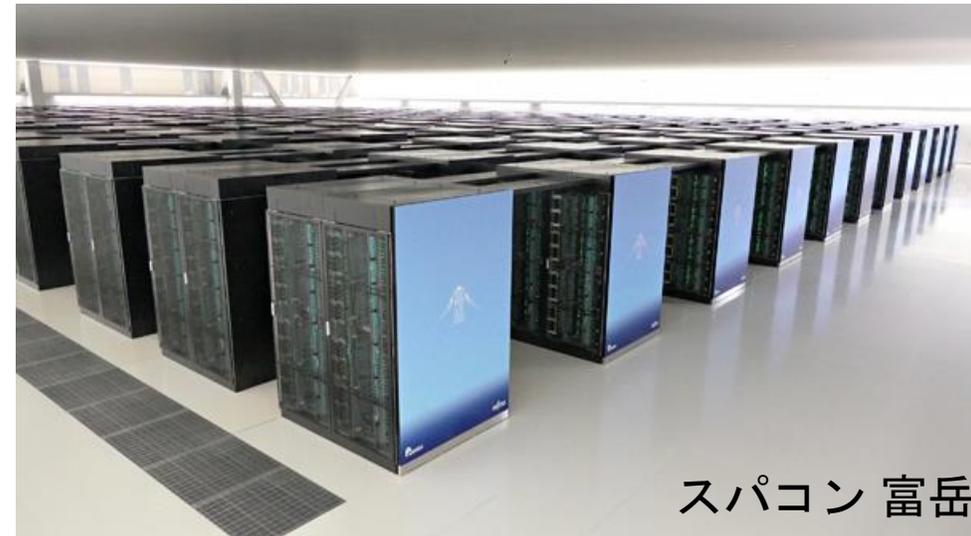
東京大学 情報基盤センター
スーパーコンピューティング研究部門
塙 敏博

スーパーコンピュータ（スパコン）とは？

- （本来は）科学技術計算を主目的とする大規模なコンピュータ
 - 一般的なコンピュータに比べて、高性能かつ大規模
 - ざっくり**数千倍以上の規模/性能/容量**
- 大量の複雑な計算を高速かつ**正確に**解くことが求められてきた



2024年10月26日(土)

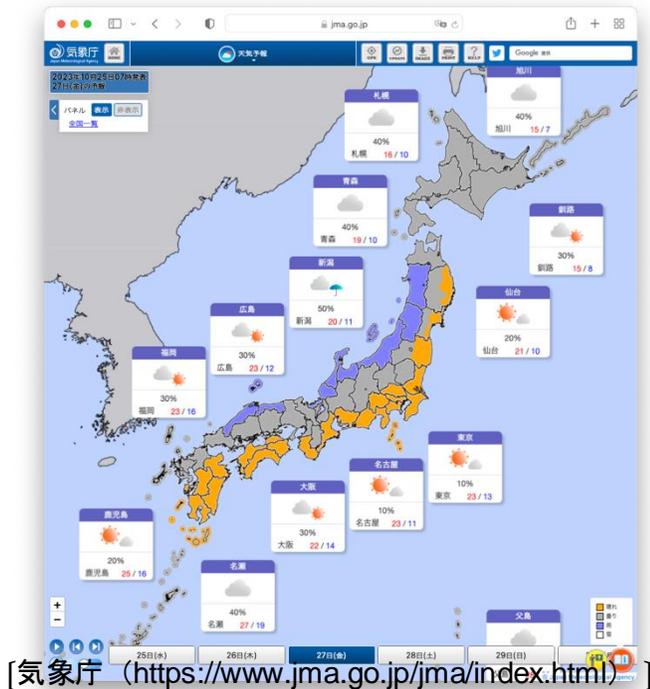


柏キャンパス一般公開

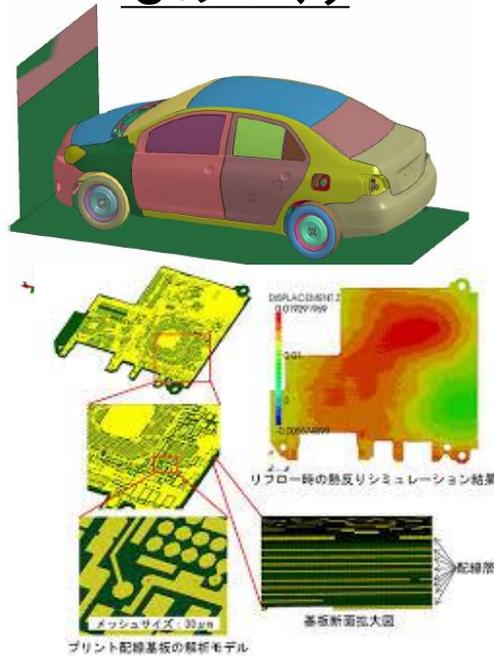
スパコンの活用事例

- 大規模/精密、かつ高速に数値シミュレーションを実行できることが求められてきた

気象庁の「天気予報」

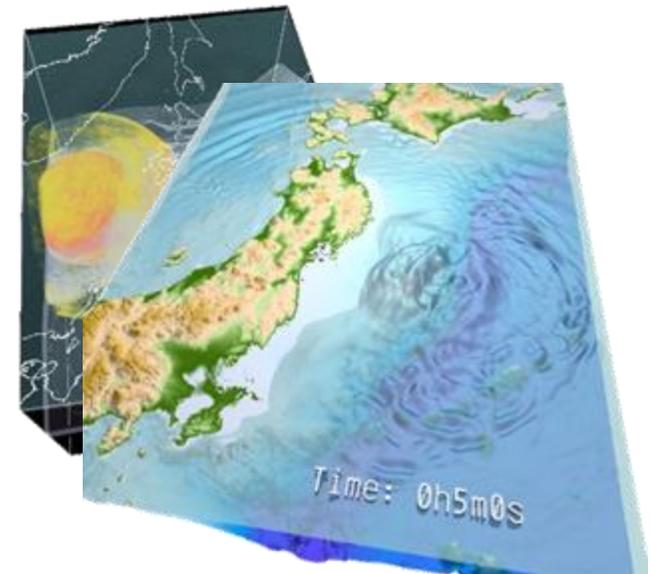


ものづくり



〔画像提供: 日本自動車工業会 (JAMA)〕

地震シミュレーション



〔画像提供: 古村孝志教授・市村強教授 (東大・地震研)〕

スパコンの速さはどこから来るの？

- プロセッサ (CPU: Central Processing Unit, 中央演算装置)

- 計算・処理の**中枢**
- コア数を増やして性能を上げる
 - 1個ずつのコアを速くするのはほぼ限界(20年前)
 - **電気を食うので、コア性能・コア数をやたら上げられない**



昔



現在: 96コアの製品も使われている

- メモリ

- 処理中のデータやプログラムの置き場 (電源を切ると消える)
- CPUとの間で高速に繋がらないと性能は出ない
 - 例えば96コアにスムーズにデータを届けられないといけない

各コンピュータ

- ネットワーク (インターコネクト)

- 全体 (数千~10数万のコンピュータ群) を束ねて一体のシステムとして動かす
- 超高速ネットワーク

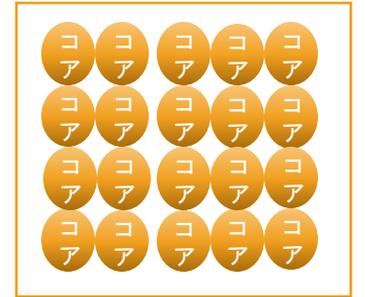
GPUのスパコンへの導入 (世界的には2009年ごろ~)

- プロセッサ (CPU: Central Processing Unit, 中央演算装置)

- 計算・処理の**中枢**
- コア数を増やして性能を上げる
 - 1個ずつのコアを速くするのはほぼ限界(20年前)
 - **電気を食うので、コア性能・コア数をやたら上げられない**



昔



現在: 96コアの製品も使われている

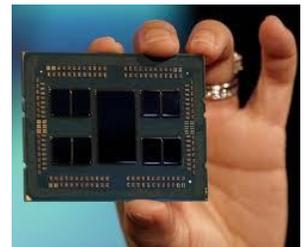
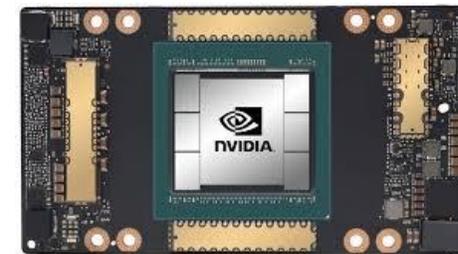
- **GPU**: Graphics Processing Unit、CPUに対してadd-onする**演算装置: 演算加速装置、アクセラレータ**

- **単純化したコアをたくさん並べる**
 - 例えば64個の演算器をまとめて制御、制御単位(CPUコアのような)=132
→ 演算器は $64 \times 132 = 8,448$ 個
 - これをフルに使えないと性能が出たことにならない
 - **電力効率が良い**

x GPU単独では動作できない、プログラミングが複雑

- 最新のスパコンランキング(Top500)の上位**10位のうち9台**はGPU搭載

- 4位の富岳だけがGPU搭載なし



一方AIは、

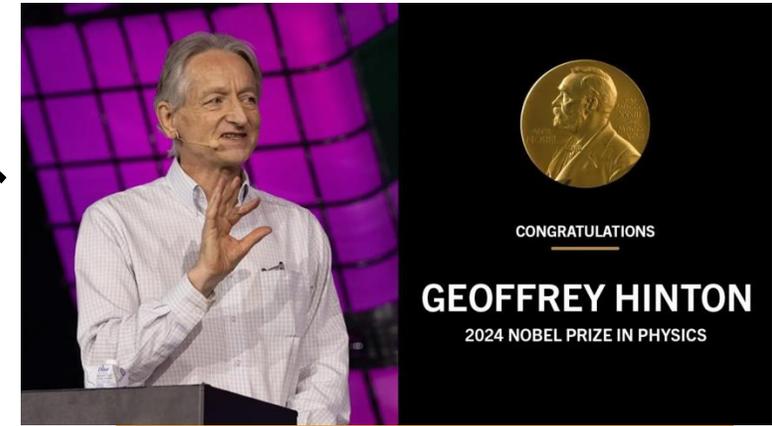
- 2012年：画像認識の精度を競う大会 (ILSVRC) で、トロント大学が**深層学習**を使って圧倒的優勝

→ GPUを使って大量の計算を高速化



ここから爆発的なAIブームに突入

- GPUの進化：AI技術が牽引
 - 最初は単なる画像処理エンジン
 - 実数の演算ができる高速な演算加速装置としてスパコンに導入
 - AI向けの機能が次々に追加



ジェフリー・ヒントン
(当時トロント大教授)
2024年ノーベル物理学賞受賞!!

人工知能(AI)とは

現代は機械学習を指すことが多い

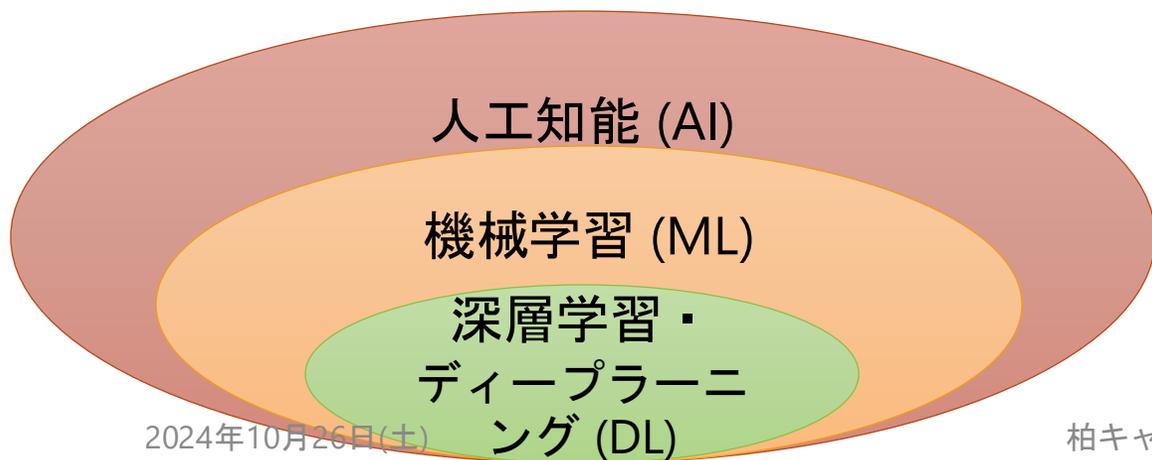
- 機械学習 (ML: Machine Learning)
 - コンピュータで人間の学習に相当する仕組みを実現したもの
- 本来のAIはもっと広い意味

- 深層学習・ディープラーニング (DL: Deep Learning)

- 多数の層から成るニューラルネットワークを用いて行う機械学習の手法
- DL技術の進歩が現在の人工知能ブームを支えている



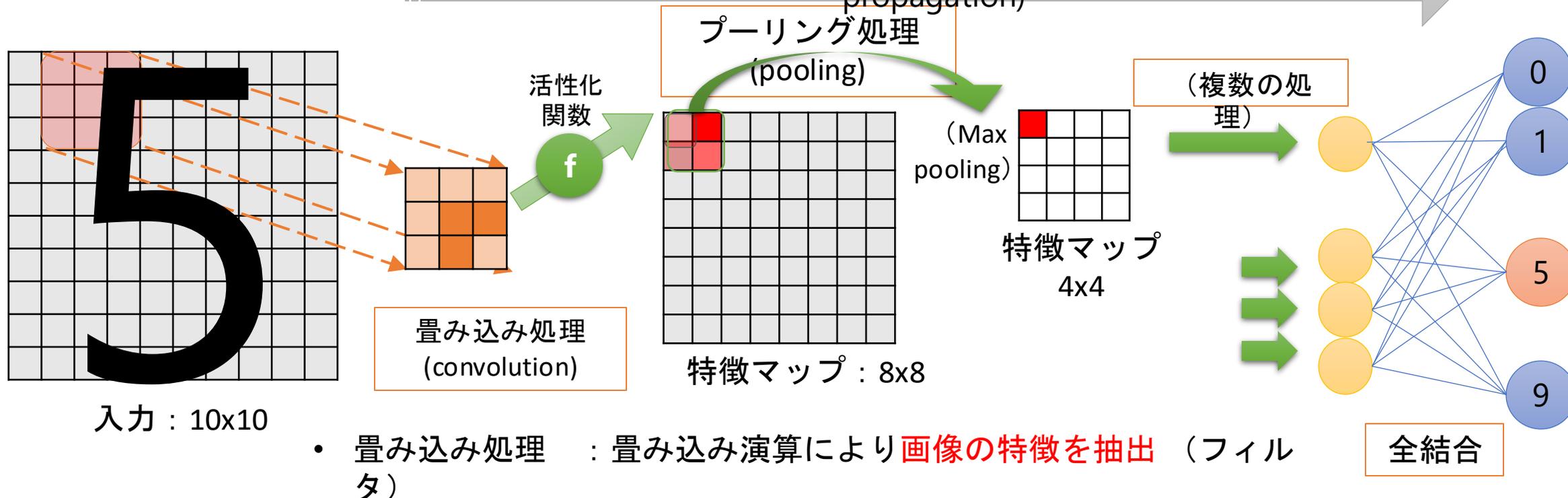
- 生成AI, 大規模言語モデル(LLM)



CNN (Convolutional Neural Network)の例

学習：誤差逆伝搬によりパラメータ決定 (back propagation)

推論：学習で得たパラメータを使って順方向伝搬 (forward propagation)



- 畳み込み処理 : 畳み込み演算により画像の特徴を抽出 (フィルタ)

AI処理の特徴 :

- 計算は多少不正確でも正しい結果が得られる
- 決まった形の計算 (小さな行列・テンソル演算) が大量に必要

生成AI・大規模言語モデル(LLM)

• 鍵になる技術

• Transformer

- 注目する入力を決める仕組みによって**独立して同時に多数の処理**が可能に

→ GPUを効率よく使えるようになった！

• 自己教師あり学習

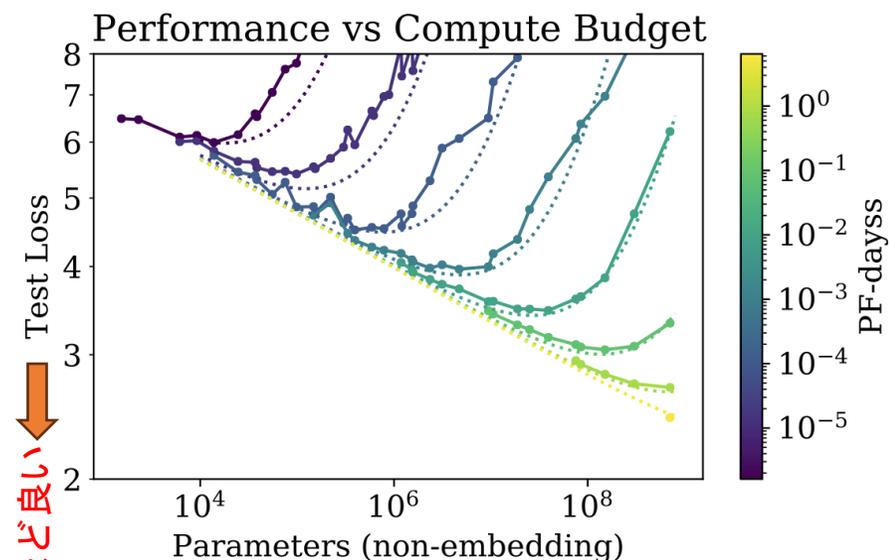
- 従来は教師データを別に用意する必要があった
- データの中から生成モデルによって教師データを作成
 - 例えば「穴埋め問題」を自分で作れば、正解（教師データ）は実はわかっている

→ 大量のデータセットを扱うことが可能

• スケーリング則

- 言語モデルのパラメータ数（=モデルのサイズ）、データセットのサイズ、トレーニングに使用される計算量が増えるにつれて、誤差 = 性能がべき乗則に従って減少する

→ AIにもスパコンが必要に！！



処理できる計算性能(右のカラーバー)とパラメータ数が多いほど、モデルの性能がよくなる
(合わせてデータセットサイズも増やさないとイケない)

J. Kaplan 他(Open AI), Scaling Laws for Neural Language Models, <https://arxiv.org/pdf/2001.08361>

AIを支えるGPU

$$1.34567 + 5.34723 = 6.69290$$

- 科学技術計算：正確さが命

⇔ AI処理：多少雑に計算しても構わない=低精度演算、演算が簡単な分
高速

$$1 + 5 = 6$$

- AI処理を支援するハードウェア

- テンソルコア、Transformerエンジン、様々な種類の低精度演算

- OpenAIがChatGPTに使った計算資源

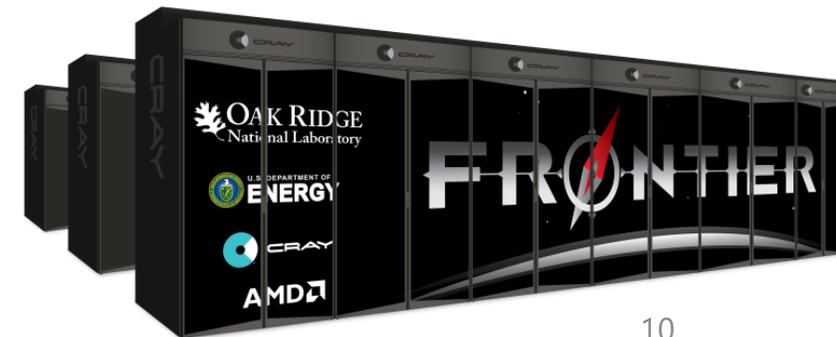
- GPU数：約10,000
- 数週間の学習 (training)

- 今後のモデル開発には30,000GPUは必要とか

→まさにスパコンそのもの、
GPU数さえあればいい
ということものではない、
構築・利用技術も重要

2024年6月26日(土)

柏キ



例：タンパク質の立体構造予測 Alpha Fold2

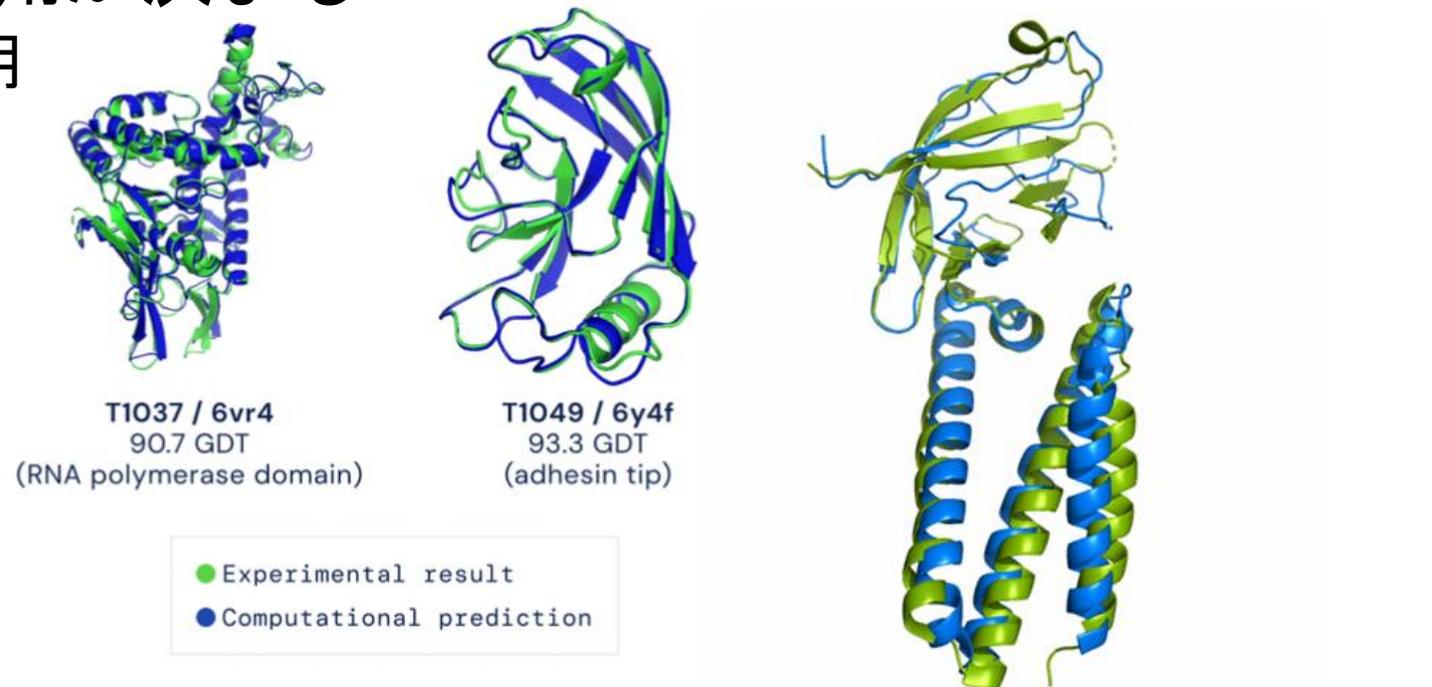
- タンパク質は、どのような配列できているか、に加えて、立体構造によって、機能や相互作用が決まる
 - 創薬、洗剤、生命の起源解明

従来は

- X線結晶解析、電子顕微鏡等で測定
- 時間をかけてシミュレーションをする

→ AIによる予測

(スパコンでなくてもある程度可能)



新型コロナウイルス中のタンパク質ORF3aの構造を推定、実測とほぼ一致

青: AlphaFoldによる推定

緑: UC Berkeley Brohawn Labによる実測値

AI for Science, AIによるシミュレーション支援

1. 複雑なシミュレーションは本質的に時間がかかる

→時間を短縮できないか？

• 時間がかかる理由

- トライ&エラー：形状や構造、割合などを少しずつ変えてシミュレーション、最適な条件を探す

- 例：自動車や航空機などの空力解析

- アンサンブル計算：様々な条件の下でシミュレーション、結果を統計処理

- 例：天気予報

→効率の良いパラメータをAIで推定、全体で必要なシミュレーション回数を減らす

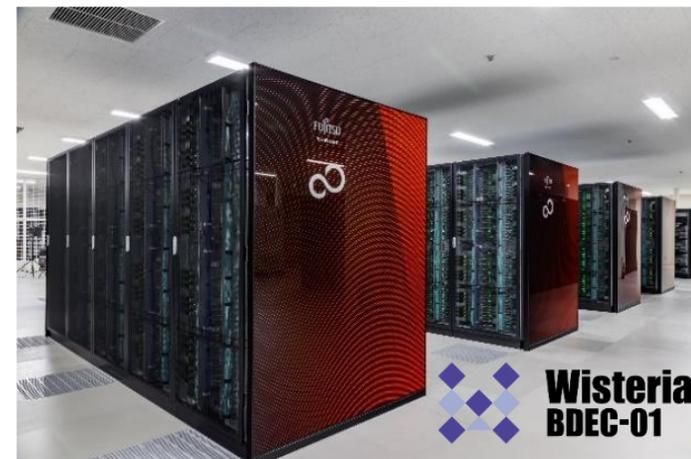
2. 実験をせずに画期的なものを発見できないか？

- 過去の実験データなどから全く新しい条件を推定

- 例：物質材料の探索

(シミュレーション (計算) + データ + 学習) 融合

- 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」
 - 今で言う AI for Science
- 2021年5月に運用を開始した「Wisteria/BDEC-01」は「BDEC計画」の1号機
 - 「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
 - Odyssey (シミュレーションノード) : 富岳同型機 (A64FX)
 - 25.9 PF, 7,680ノード
 - Aquarius(データ学習ノード): GPUクラスタ
 - 7.2 PF, 45ノード



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

観測データ

Aquarius

データ同化
データ解析



Wisteria
BDEC-01

2024年10月26日(土)

サーバー
ストレージ
DB
センサー群
他



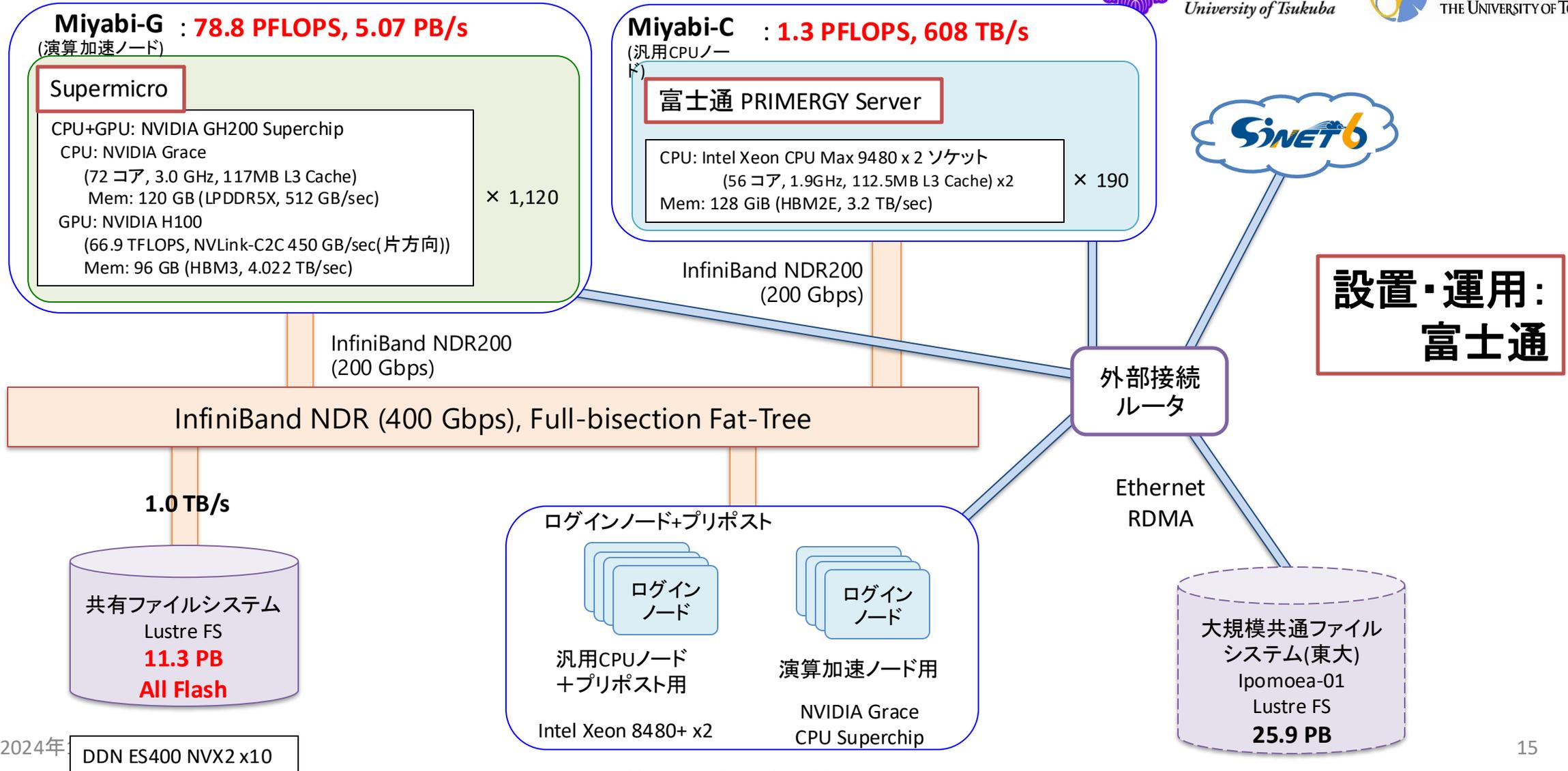
外部ネットワーク



外部
リソース

Miyabi (OFP-II)の概要 (1/2)

2025年1月運用開始



Miyabi(OFP-II)の概要 (2/2)



• Miyabi-G: 演算加速ノード: NVIDIA GH200

– 計算ノード: NVIDIA GH200 Grace-Hopper Superchip

- Grace: 72c, 3.45 TF, 120 GB, 512 GB/sec (LPDDR5X)

- H100: 66.9 TF DP-Tensor Core, 96 GB, 4,022 GB/sec (HBM3)

- CPU-GPU間はキャッシュコヒーレント

- NVMe SSD for each GPU: 1.9TB, 8.0GB/sec, GPUDirect Storage

– **合計 (CPU+GPUの合計値)**

- **1,120 ノード, 78.8 PF, 5.07 PB/sec, IB-NDR 200**

• Miyabi-C: 演算加速ノード: NVIDIA GH200 (SPR)

– 計算ノード

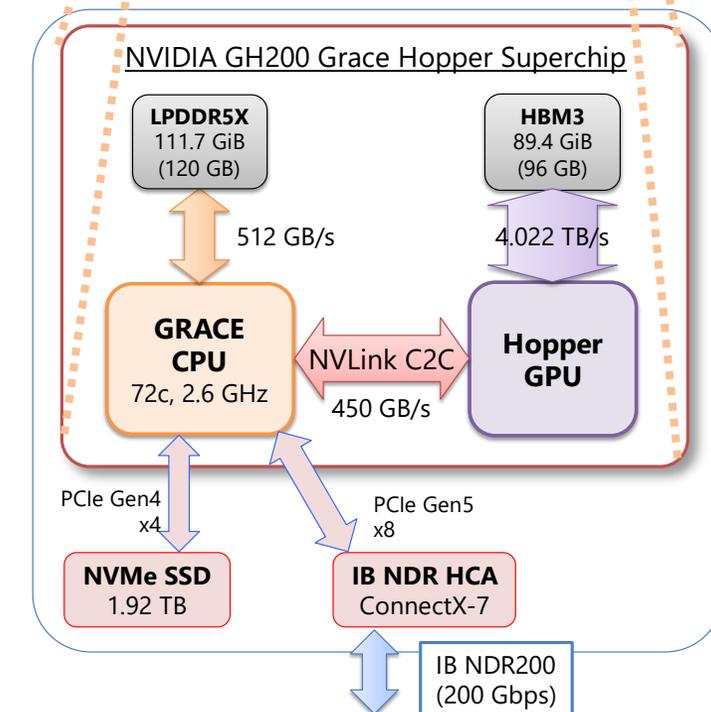
- 6.8 TF, 120 GB, 512 GB/sec

– **合計**

- **190 ノード**

- **372 TB/sec for STREAM Triad (Peak: 608 TB/sec)**

来年の一般公開を
お楽しみに！

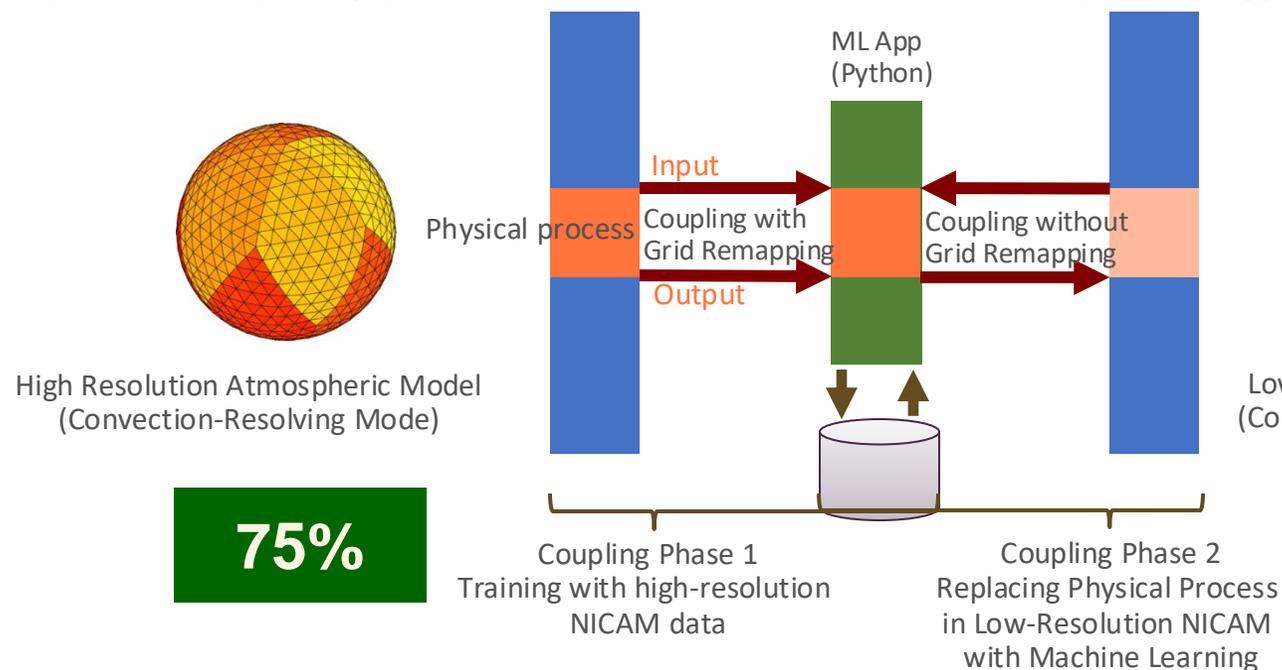


「計算+データ+学習」融合を支援する ライブラリ・実行環境の開発



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

- 異なるモデルのシミュレーションを結合
- シミュレーションと機械学習を結合
- 異なる種類のスパコン同士で通信を可能にする: Odyssey ⇔ Aquarius



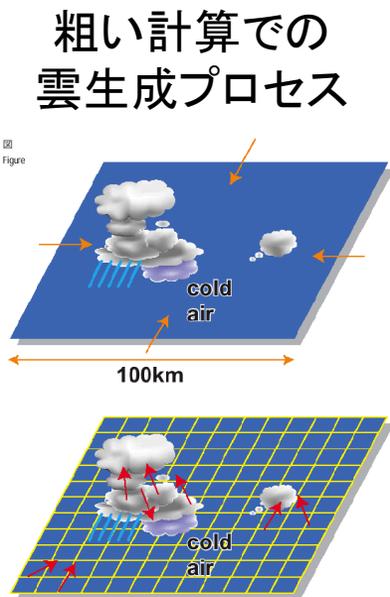
75%

Low Resolution Atmospheric Model (Convection-Parameterization Mode)

この部分を AIによる推論に

25%

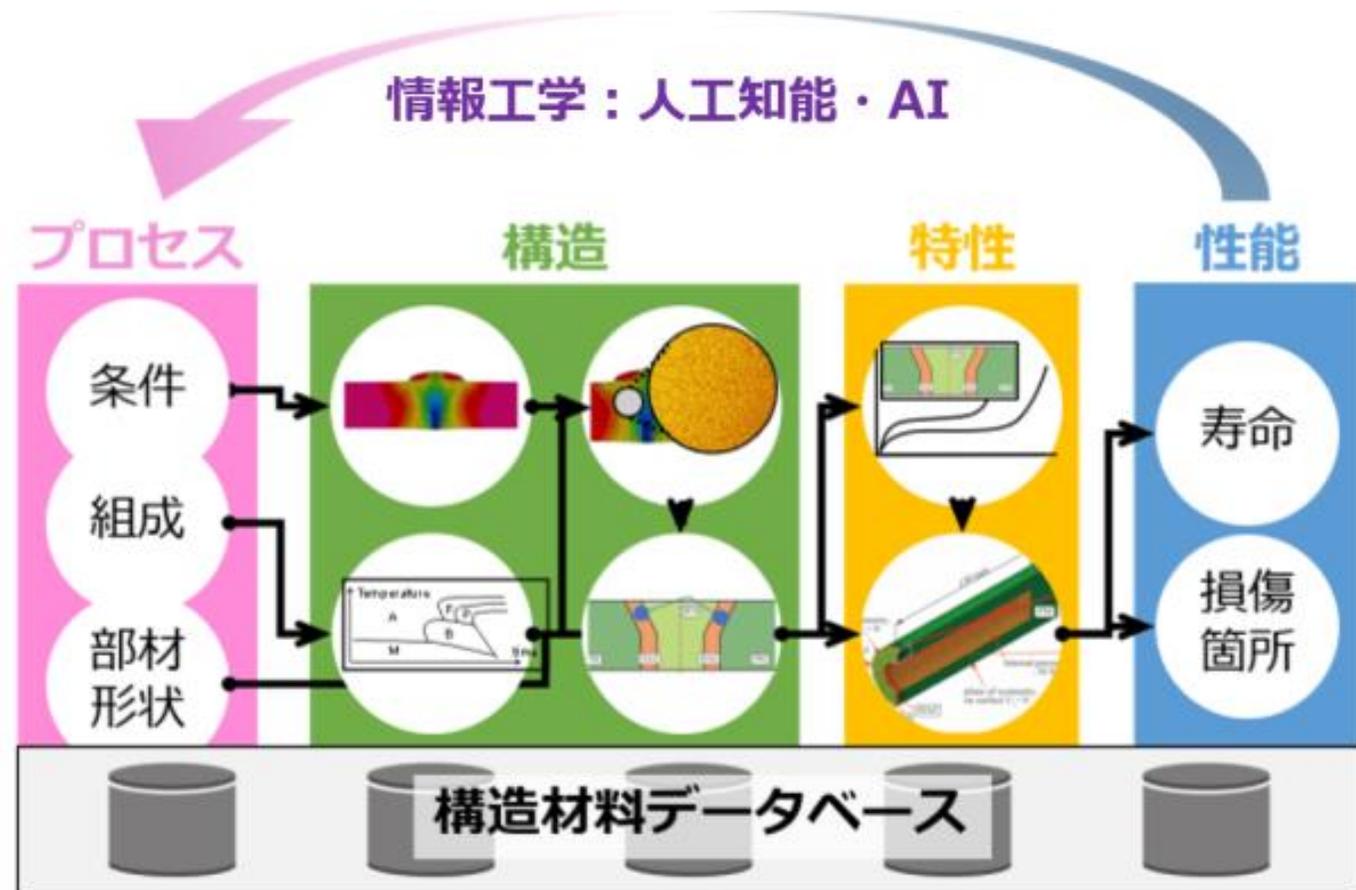
~0%



気象コードにおける機械学習との連携[八代・荒川 2020]

材料開発へのAI活用 (国立物質材料研究機構：NIMS)

- 所望の特性に応じた、
 - 材料の組成、
 - 形状、
 - 製造の条件など、をAIによって推定し候補を提示
- シミュレーション時間の低減
- 実験の効率化



まとめ

- **AI for Science**によって
 - これまでは簡単に発見できなかったことが見つかる（かもしれない）
 - これまでの全シミュレーション実行時間を**1/10（あるいはもっと）**に短縮できる
- Wisteria/BDEC-01は、先進的なAI for Science, AI for HPCを実現する基盤として提供
 - ソフトウェア: h3-Open-BDEC
- 今後も「計算＋データ＋学習」融合 = BDEC プロジェクトを推進
 - Wisteria/BDEC-01 → BDEC-02
- **2025年1月**導入予定の **Miyabi (OFP-II)**ではGPUが主力、AI for Scienceを加速
 - JCAHPC: 筑波大学計算科学研究センターとの共同運用



筑波大学
計算科学研究センター
Center for Computational Sciences

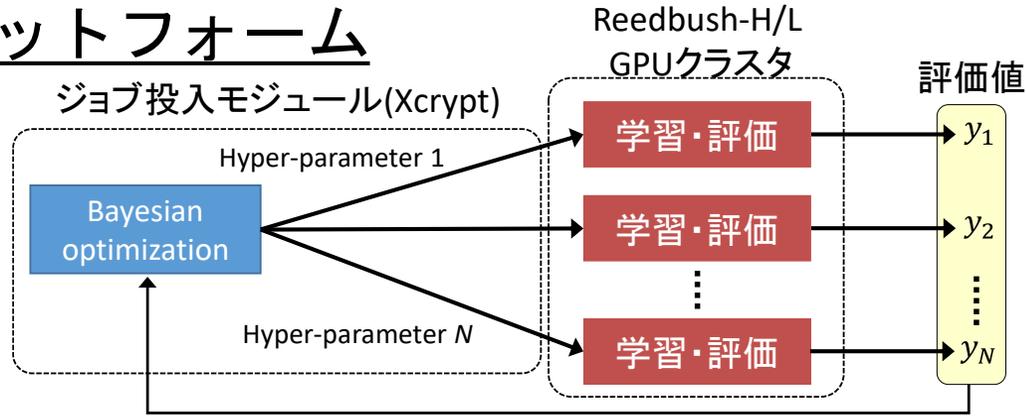


東京大学情報基盤センター
INFORMATION TECHNOLOGY CENTER, THE UNIVERSITY OF TOKYO

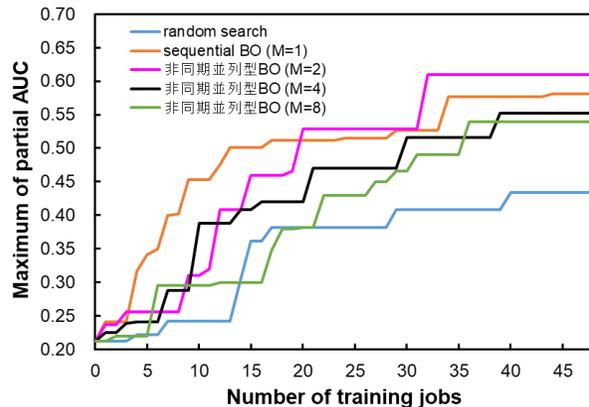
バイオインフォマティクス：医療画像処理

東大病院等

深層学習自動チューニングプラットフォーム

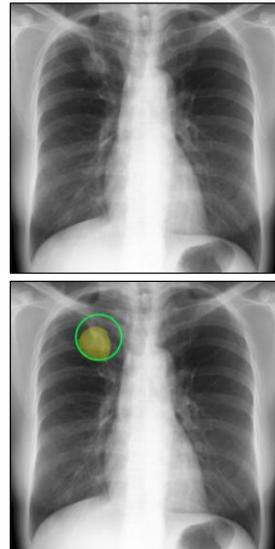


胸部X線写真の肺腫瘍検出



学習ジョブ数と評価値(partial AUC)の最大値との関係

2024年10月20日(土)

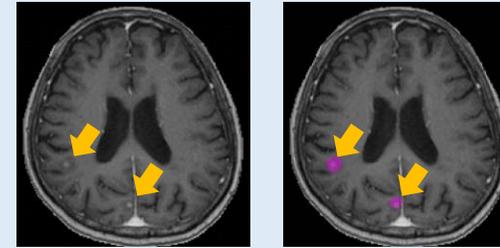


上：元画像、下：検出結果
(黄、緑丸：病変領域)

開発中のソフトウェア

頭部造影MR画像の転移性脳腫瘍検出

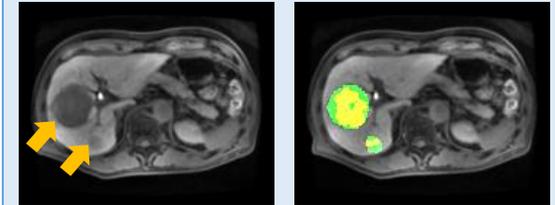
村田, JAMIT2018



検出結果例
左：元画像、右：検出結果(マゼンダ)

造影MR画像の肝結節性病変検出

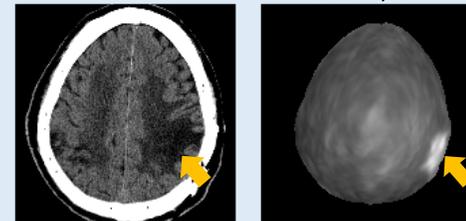
Takenaga T, CARS 2018



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

頭部救急CT画像の異常検知

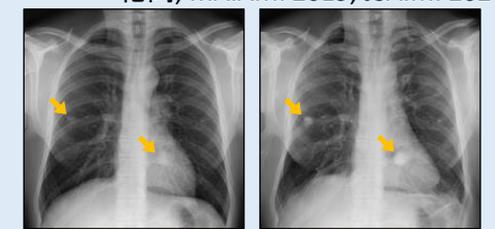
Sato D, SPIE MI 2018



脳梗塞症例
(左：元画像、右：異常度マップ)

胸部X線画像の異常強調

花岡, MAIAMI 2019, JSAIMI 2020



強調画像例
(左：元画像、右：強調結果、矢印：肺腫瘍)

Nomura Y, J Supercomput. 20 Jan 2020 (Epub ahead)

Wisteria/BDEC-01

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

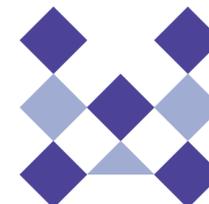
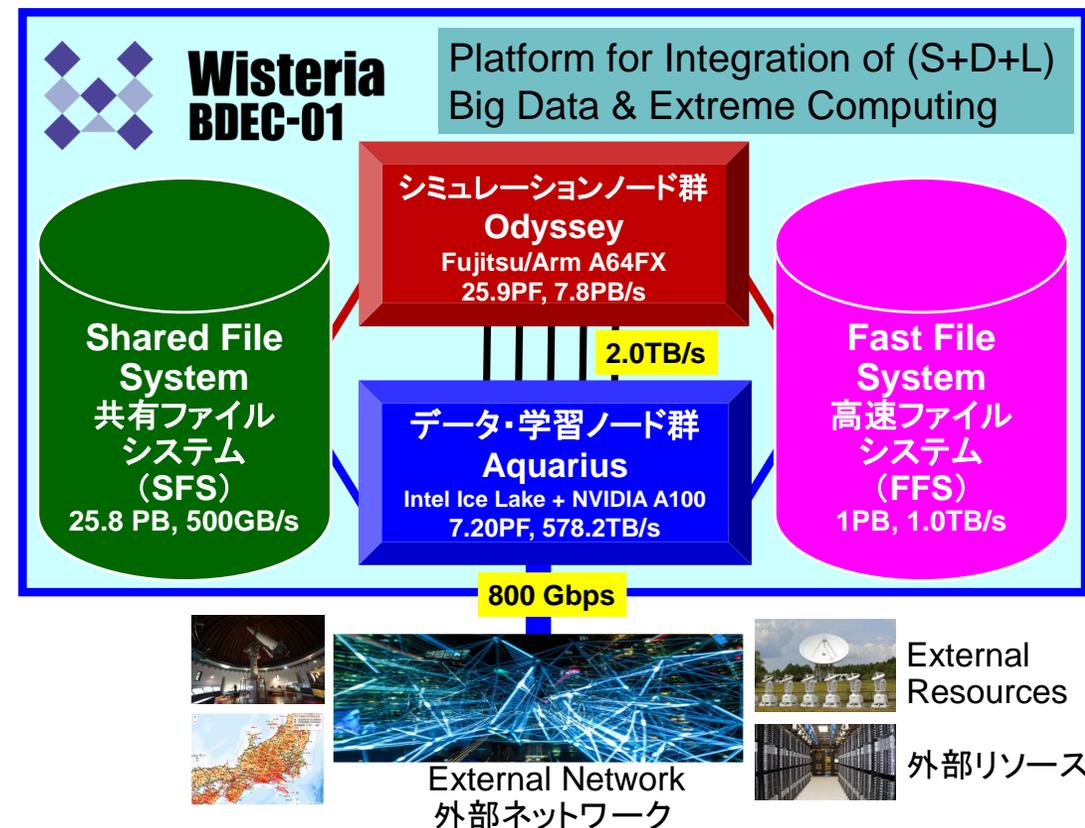
• 2種類のノード群

- シミュレーションノード群 (S, SIM) : **Odyssey**
 - 従来のスパコン
 - Fujitsu PRIMEHPC FX1000 (A64FX), 25.9 PF
- データ・学習ノード群 (DL, 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

2024年10月26日(土)

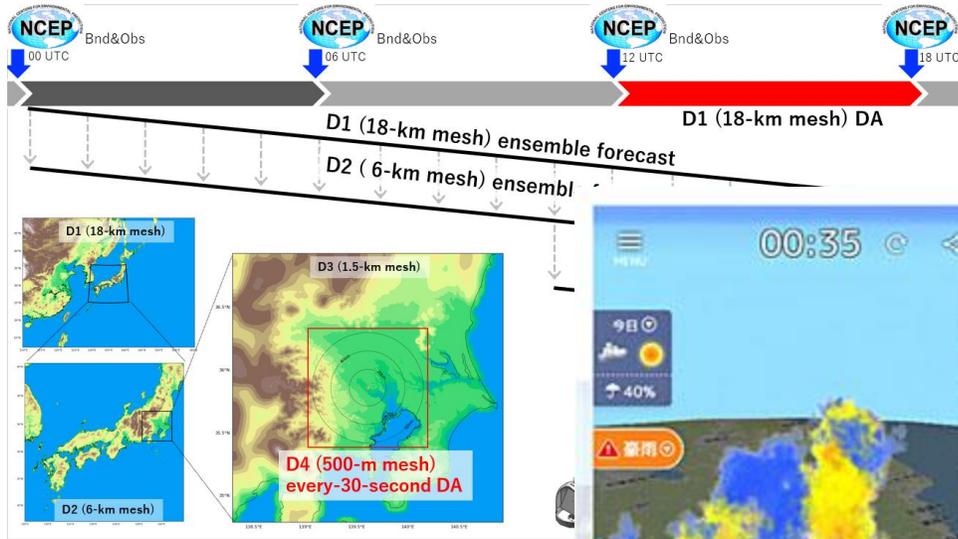
柏キャンパス一般公開

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)

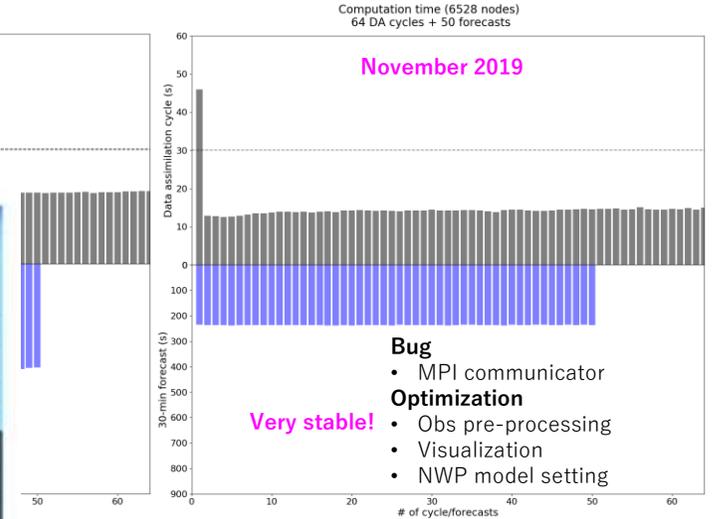
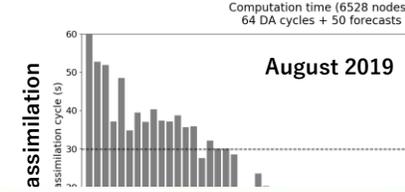


Wisteria
BDEC-01

ゲリラ豪雨予測のリアルタイム実証実験 (理化学研究所)



全体のワークフロー

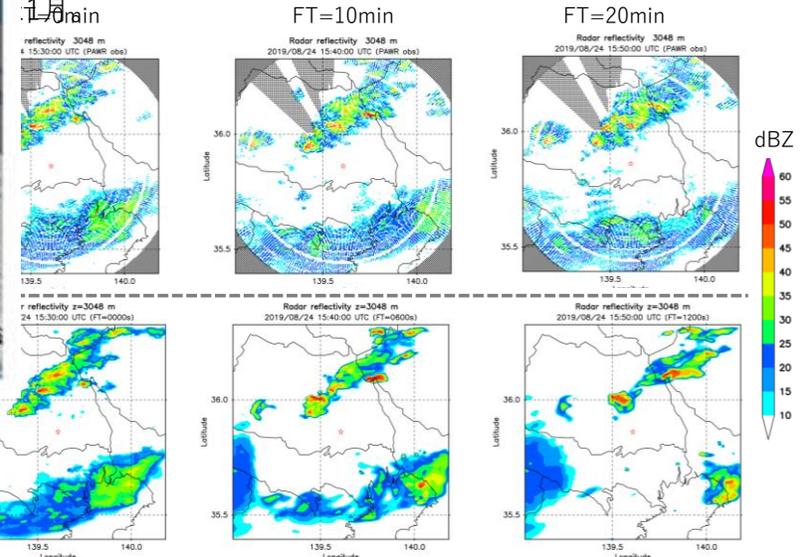


夕同化、下段は30分予報にかかった時間(秒)。



PAWR Obs

SCALE-LETKF Analysis



〔画像提供: 三好建正博士 (理化学研究所)〕

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

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

スパコンの速さの源泉： プロセッサ・メモリ・ネットワーク

- CPU (Central Processing Unit, 中央演算装置)
 - 並列計算が基本：PC, スマホも同様
 - 最新スパコンでは 40~100コア程度、それを1~2個
 - マルチコア、メニーコア
- GPU (Graphic Processing Unit)
 - 描画用途⇒高いデータ処理能力を計算に利用
 - 数百のコアを有する「メニーコア」
- メモリ：三次元積層、高速なデータアクセス
- **マルチコア・メニーコアCPU (+GPU) を超高速ネットワークで接続した並列計算機 (Parallel Computer) = 「スパコン」**
 - 家庭用LANの数百倍, 100Gbps~800Gbps

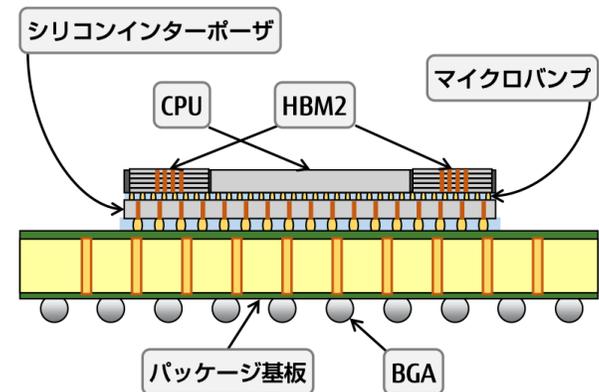
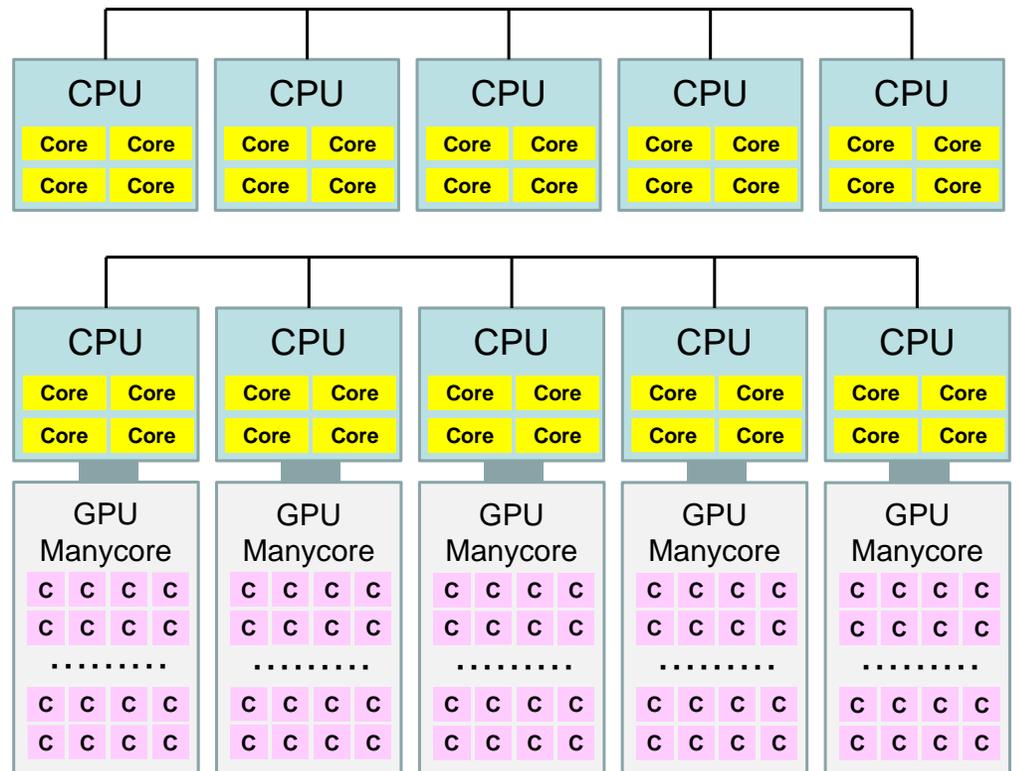
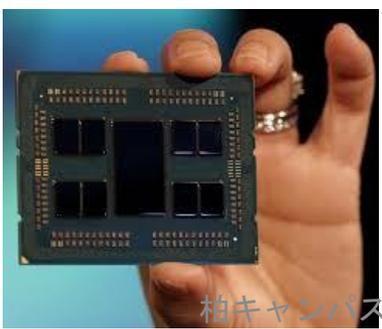


図 12 2.5次元パッケージの構成 (断面模式図)

「whitepaper
富士通
PRIMEHPC
FX1000」
<https://www.fujitsu.com/downloads/JP/jsuper/primehpc-fx1000-hard-ja.pdf>

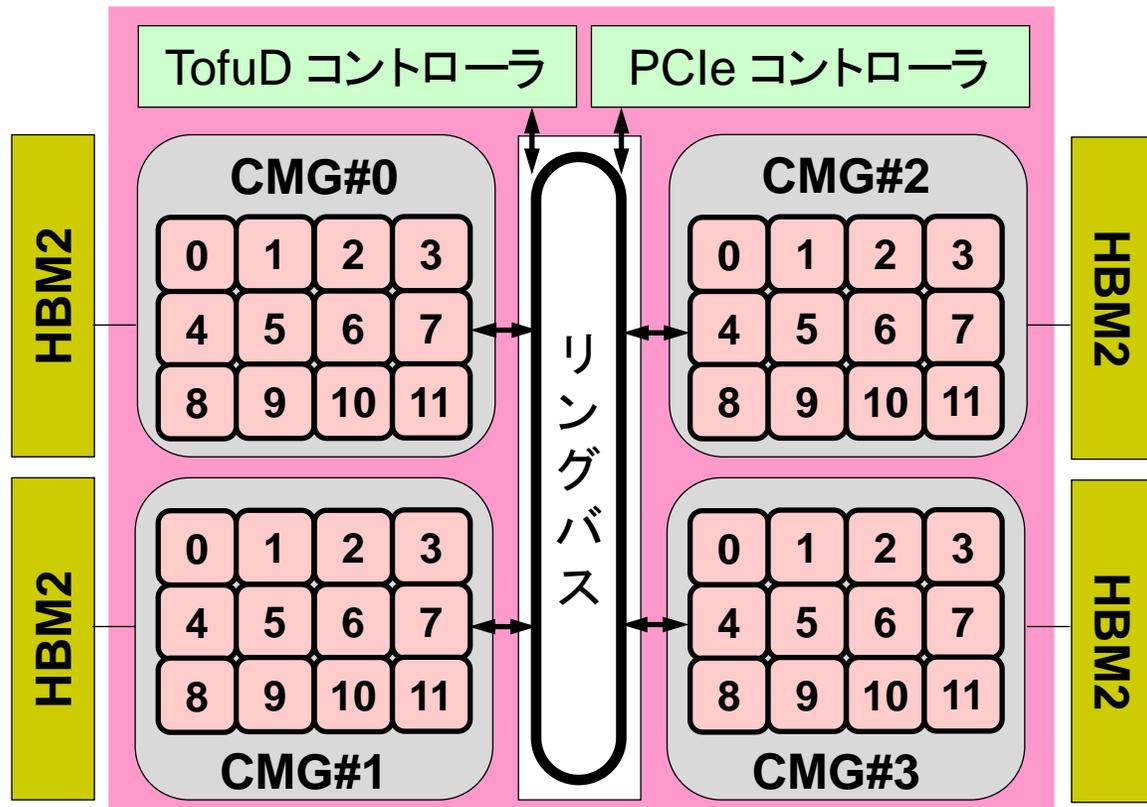


2024年10月26日(土)



柏キャンパス一般公開

富士、Wisteria/BDEC-01に使われている技術



A64FX プロセッサ

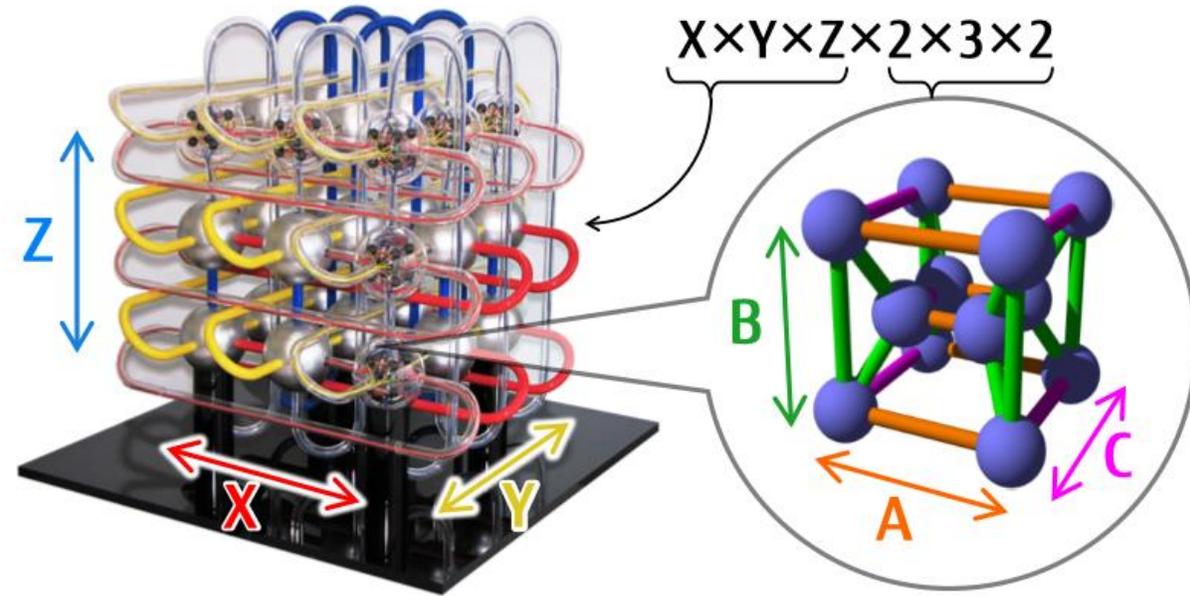
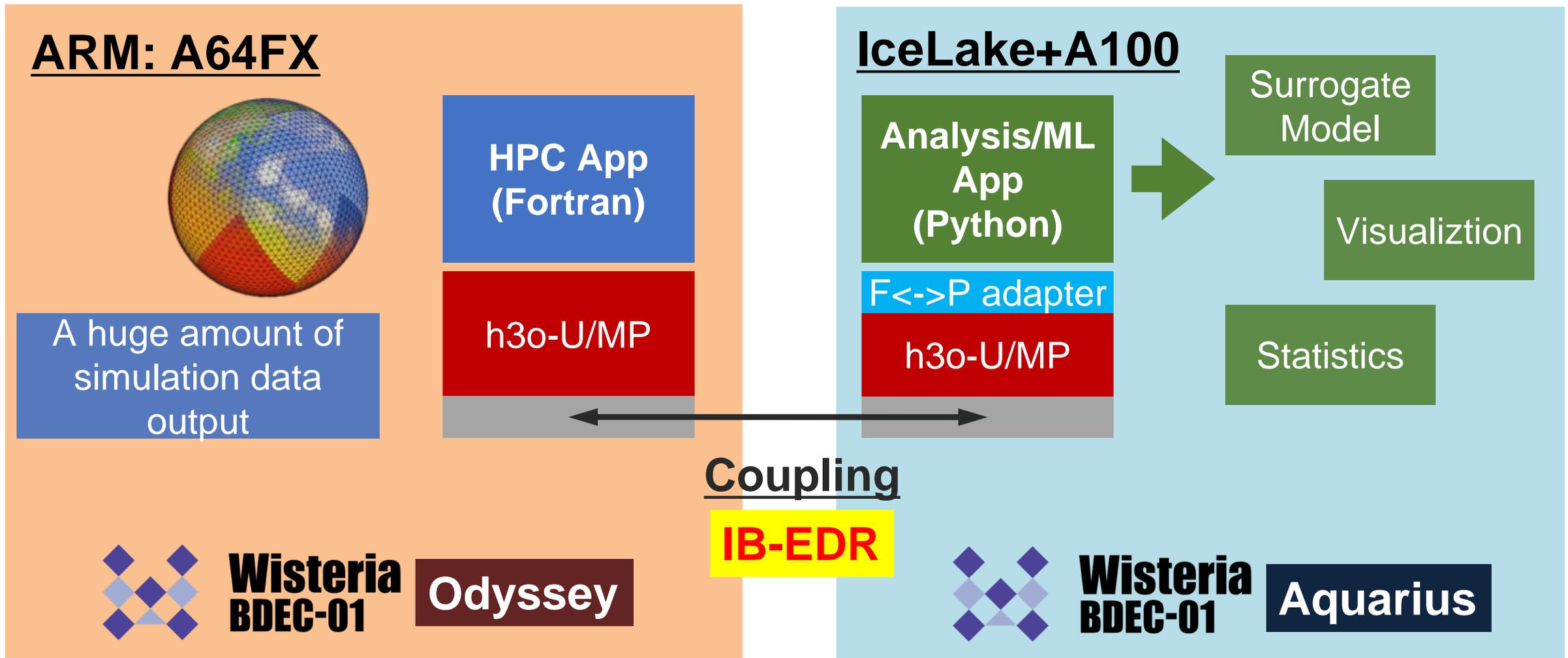


図 13 6次元メッシュ/トーラスのトポロジーモデル

Tofu-D ネットワーク

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



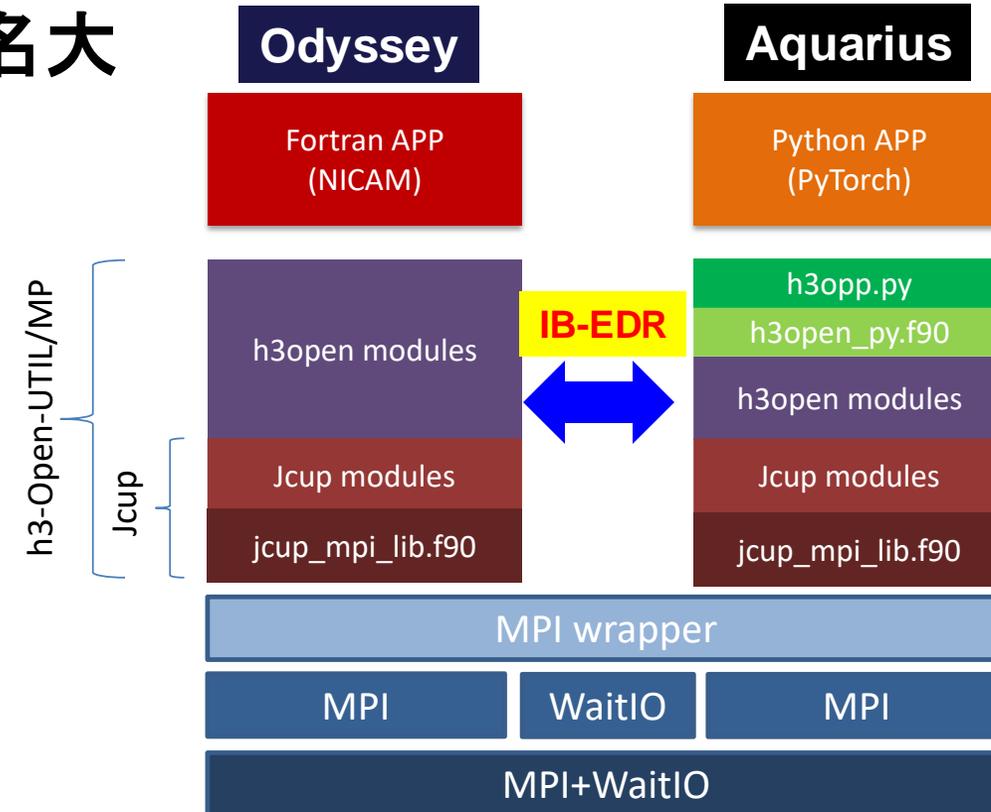
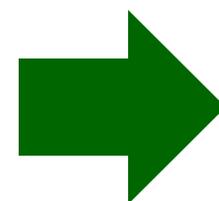
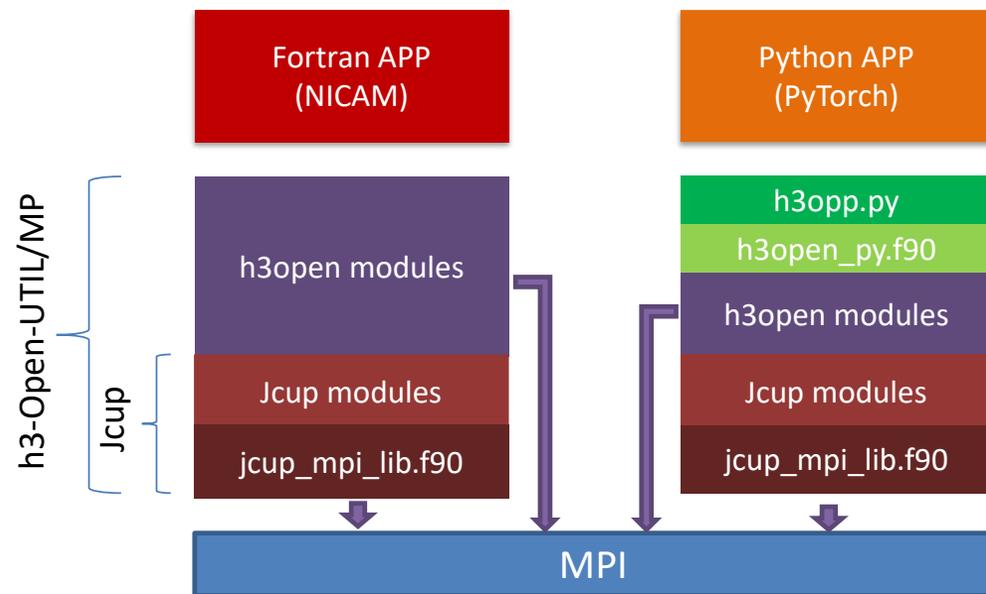


h3-Open-UTIL/MP・h3-Open-SYS/WaitIO-Socket連携 2022年6月から利用可能

2022年度はFS経由のWaitIO-File整備：名大



Wisteria BDEC-01



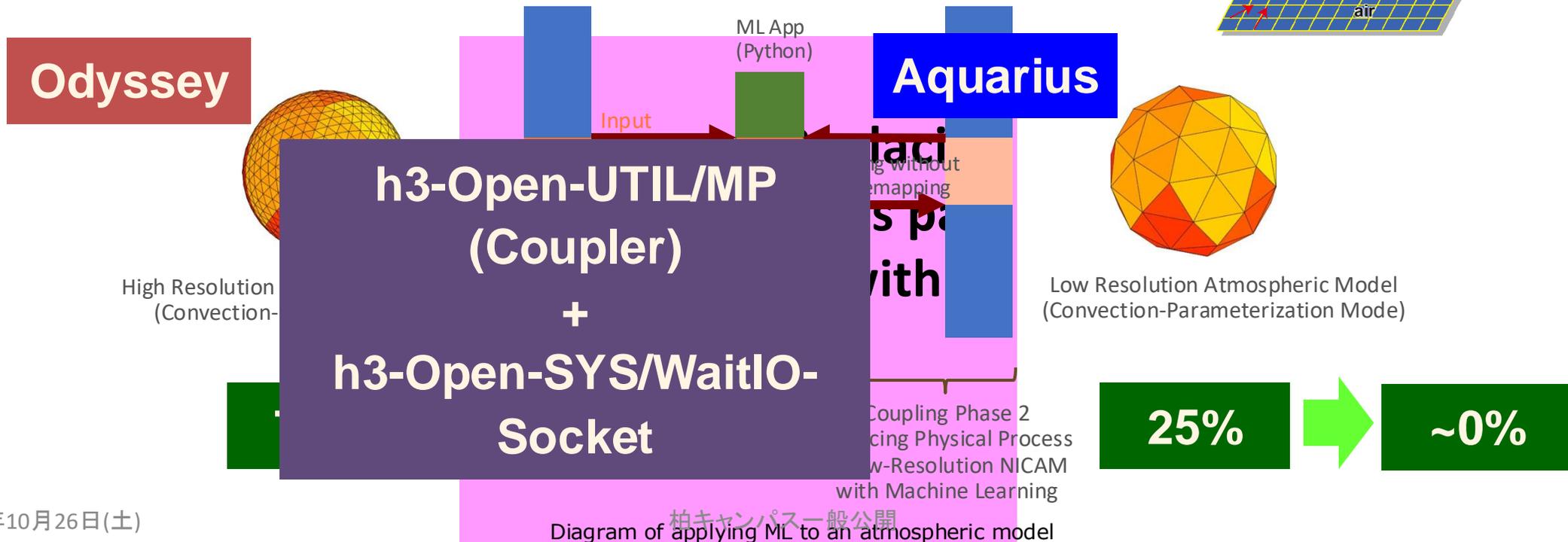
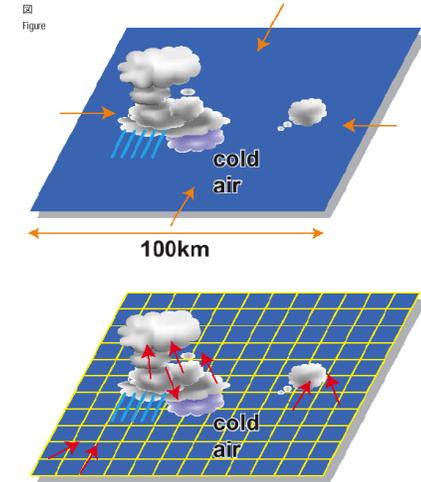
2021年4月：MPI通信可能な環境を前提

2022年6月：Coupler + WaitIO

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!



おまけ：ChatGPTに聞いてみた

- なかなかいいことを言っている
 - でも15分じゃ無理だね
- ➔お主やるな...

The screenshot shows a chat interface with two messages from a user (HA) and one response from ChatGPT. The user asks for a 15-minute summary of a talk. ChatGPT provides a structured summary with a title and five main points, each with sub-points.

HA AIと... 考え

HA 15分で話せるぐらいの内容にしてください

「AI

1. イン

-
-

2. スパ

-
-

3. AIと

-
-

4. スパ

-
-
-

5. 次世

-
-
-
-

もちろん、15分の講演に合わせて簡略化した内容にしましょう。

タイトル: "AIとスパコン：科学技術シミュレーションの未来"

1. イントロダクション (1分)

- スパコンとAIが科学技術をどのように変えているかの概要

2. スパコンとAIの基礎 (2分)

- スパコンとは何か？性能と進化
- AIの基本原理と役割

3. スパコンとAIの連携 (2分)

- ハードウェアとソフトウェアの統合
- GPUの役割と分散処理

4. 次世代の科学技術シミュレーション (3分)

- 原子・分子レベルのシミュレーション
- 気象予測と気候モデリング
- 宇宙探査のシミュレーション

柏キャンパス 一般公開