

# 理化学研究所を中核とした国際連携による「富岳NEXT」開発体制とその意義

~ 計算による課題解決を支える次世代AI-HPCプラットフォームの構築へ ~



理化学研究所 計算科学研究センター(R-CCS) センター長 松岡 聡 次世代計算基盤開発部門 部門長 近藤 正章

2025年8月22日 HPCI計画推進委員会



### 「富岳」の系譜を継ぎつつ新たな次世代計算基盤が目指すもの

### • フラッグシップシステムが持つ価値と国家的な重要性

- ゼタ(Zetta)スケールの大規模計算環境による計算科学領域での世界的優位性の継続的な確保
- スケールメリットによる高い導入・運用効率、およびコミュニティ形成の力を生かした裾野の拡大
- 計算能力が勝負となるAI開発でも国内最高計算能力を持つ研究プラットフォームとして活用

#### 計算可能領域の拡張を通じ、価値創造を支援するゼタ(Zetta)スケール計算資源の提供

#### • 産業競争力の強化に資する「富岳NEXT」の開発

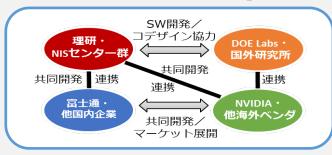
- 「富岳」のARMアーキテクチャを受け継ぎながら、さらなる発展を目指したシステム構築
- 「AI for Science」や量子コンピュータとの連携を含め、計算による課題解決の先駆的な取り組みを推進
- 加速部としてGPUを導入し、日米共同開発体制のもと、AIの積極的利用も含めて我が国のアプリケーションを現代化することで成果の社会実装を加速
- 理研を中核とした「Made with Japan」の開発体制構築と半導体戦略も踏まえたプロジェクト推進

世界的に訴求力のある国産技術の高度化や技術継承、情報産業での戦略的不可欠性の確保とグローバルマーケットへの展開を推進



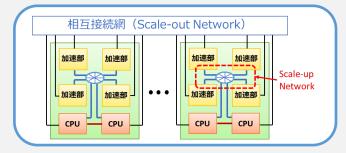
### 科学技術・産業競争力強化に向けた「富岳NEXT」開発戦略

### Made with Japan



- 基本設計ベンダー・他国内外ベンダーとの 協調による、訴求力のあるシステム構築と 国産技術高度化、グローバルマーケットへ の展開、国内半導体産業の振興
- 国際連携でのプロジェクト推進による高度人材の育成、技術革新の加速、情報技術における主権の確保

### 技術革新



- 高性能CPU-GPUの密結合による広帯 域・ヘテロジニアスなノードアーキテクチャ、 先進的なメモリ技術等の共同開発
- AI・HPC技術のさらなる発展と、その高度利用による最大100倍程度のアプリケーション実効性能の向上

#### 持続性/継続性



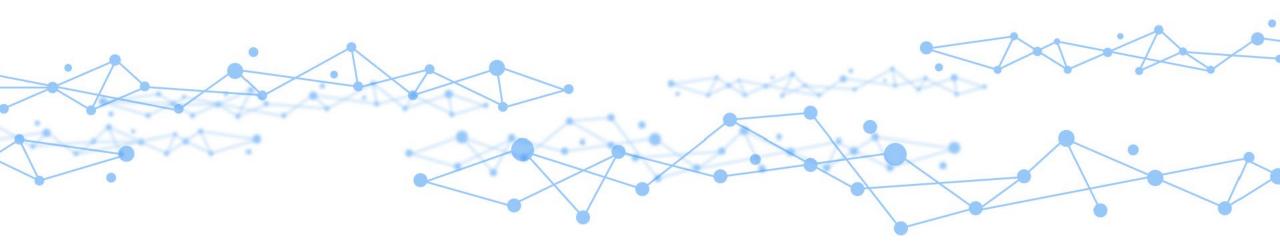
- ●標準規格や既存のエコシステムとの親和性が高いシステムの構築(ソフトウェア環境の継続整備)
- 今後の先端システムに合わせたアプリケーションの現代化とそのサポート体制構築
- ●「富岳」での取組みを進化させた運用技術 の高度化による省エネルギー化の実現

#### 「富岳NEXT」エコシステム構築とそれを利用した我が国の半導体・情報基盤の強化

- 次世代AI-HPCプラットフォーム開発による計算可能領域の拡張と「AI for Science」による科学の推進
- 先端的AI技術や計算基盤の開発における日本の主権の確保
- 継続的な先端半導体開発、計算資源確保のロードマップ構築とそれに基づく持続的な研究開発の実施



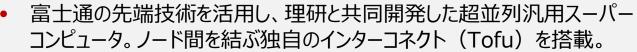
# スーパーコンピュータ 「京」から「富岳」そして「富岳NEXT」へ





### スーパーコンピュータ「京」から「富岳」そして「富岳NEXT」へ

### 「京」





- 世界に先駆け10ペタフロップス (PF) の演算性能を達成。2016年6月、 11月、2期連続で計算性能 (TOP500) 世界一を達成。
- 2012年共用開始。2019年運用終了。



### 理化学研究所

次世代スーパーコンピュータ 開発実施本部

> 計算科学研究機構 (AICS)

計算科学研究センター

(R-CCS)

### 「富岳」

理研と富士通とで共同開発。CPUのA64FXは、ARMベースで構築され、 スマートフォンなどに使われている命令セットを世界で初めてスーパーコン ピュータに拡張(SVE: Scalable Vector Engineを搭載)。



- コデザインを推進し、稼働と同時に成果を創出。特に、試験運用時には、 コロナ渦における飛沫シミュレーション等で成果を挙げ、その後もさまざまな 社会課題の解決に貢献するキーインフラストラクチャーとして活躍中。
- 2020年6月から4期連続でTOP500などのベンチ―マークテストで4冠を 達成。2021年から共用開始。世界トップレベルの省エネ性能を実現。



#### 「富岳NEXT」

- 理研が中核となり、高性能ARMベースCPUで世界をリードする「富士通」と、AI/HPC 向けGPUで世界トップシェアを誇る「NVIDIA」とで共同開発。AI性能(FP8)で世界初のゼタ(Zetta)スケールの性能を目指す。
- システムソフトウェアを含めたソフトウェア開発は、国際連携によるオープンな体制で実施。ハードウェアとソフトウェア両面において「Made with Japan」のコンセプトで日本の優位性を確保。我が国の情報基盤における重要インフラストラクチャーとして整備。
- 2030年頃、理研神戸地区において運用開始予定。

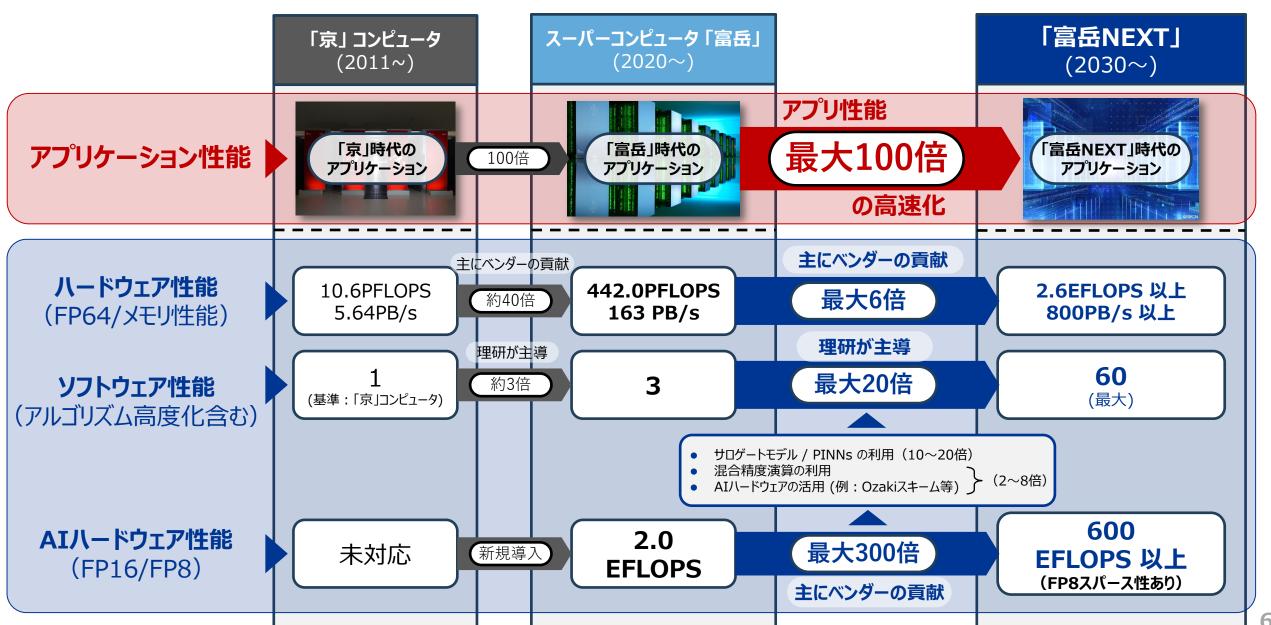


理研本部が主導するプロジェクト 推進組織から、「京」の稼働に合 わせてAICSを設置。AICSを発 展させたR-CCSは、計算機科学 /計算科学の世界的な拠点 (COE) として欧米の機関と比 肩する組織へ。

AICS発足時の8チーム、1ユニット、1部門、約150人体制から、R-CCSでは19チーム、5部門、約400名を超える体制へと拡大。



### アプリケーション性能 最大100倍への道





### 低精度演算器を活用したアプリの高度化・高速化へのアプローチ

#### 高度化·高速化手法

#### アプリへの効果

(1) Ozakiスキーム

- 高精度行列を分割→低精度で乗算
- GPUテンソルコア等を活用
- 結果は高精度で累積

(2) 混合精度演算

- 重要部分のみ高精度、他は低精度
- 反復計算で低精度主+高精度補正
- メモリ/通信量を抑制

(3) サロゲートモデル (PINN含む)

- 高コスト計算をMLで近似し高速化
- PINN (Physics-Informed Neural Network) を取り入れることでさらに物 理制約を考慮

- FP64級の精度を確保
- 高精度専用HWなしで高速化
- 消費電力削減
- メモリ帯域の効率化によるスルー プット向上
- 学習/シミュレーションの時間短縮
- 大幅な計算時間短縮によるリアルタイム予測・探索の実現
- 大規模なパラメータ走査が容易に

上記の主要な技法に加え、さまざまな高速低精度演算器を活用した技術や アルゴリズムの高度化を併用し活用することで高精度を保ちながら高速化・省電力化を理研主導で実現



### 「Made with Japan」による開発体制の意義

#### • 最高技術の組み合わせによる世界最高水準のシステム開発

- 高性能ARMベースCPUで世界をリードする富士通が次世代の後継CPUを開発。
- 世界初となるゼタ(Zetta)スケールAI/HPCシステムの実現
- AI/HPC向けのGPUと開発環境で世界トップのシェアを誇るNVIDIAとも共同開発を実施。
- その他の関連する企業とも密に連携し、国際連携体制で先端技術の粋を集めたシステムを構築。

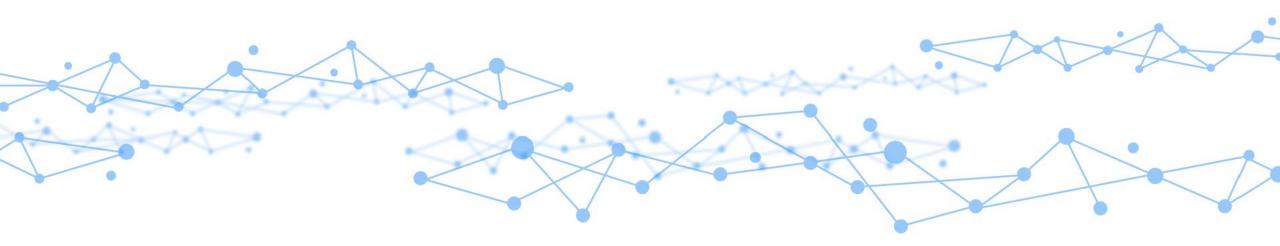
#### • 国際連携によるオープンなソフトウェア開発を「富岳NEXT」プロジェクトの一環で実施

- 米国DOEの研究所をはじめとして各国の研究所と協力したソフトウェア・アプリケーション開発を推進。
- 開発ソフトウェアを世界的なAI/HPCエコシステムの一部として普及させることで、ハードウェアとソフトウェア含む「富岳NEXTシステム」全体を高付加価値。

- 互いの強みを生かすことで、世界的に訴求力のある国産技術の高度化と技術継承
- AI/HPC融合時代における世界標準となる次世代計算基盤の構築
- 三者が協力しハイパースケーラー等への訴求などグローバルなマーケットへ展開を目指す



# 「富岳NEXT」システム概要





### 「富岳NEXT」に求められるシステム

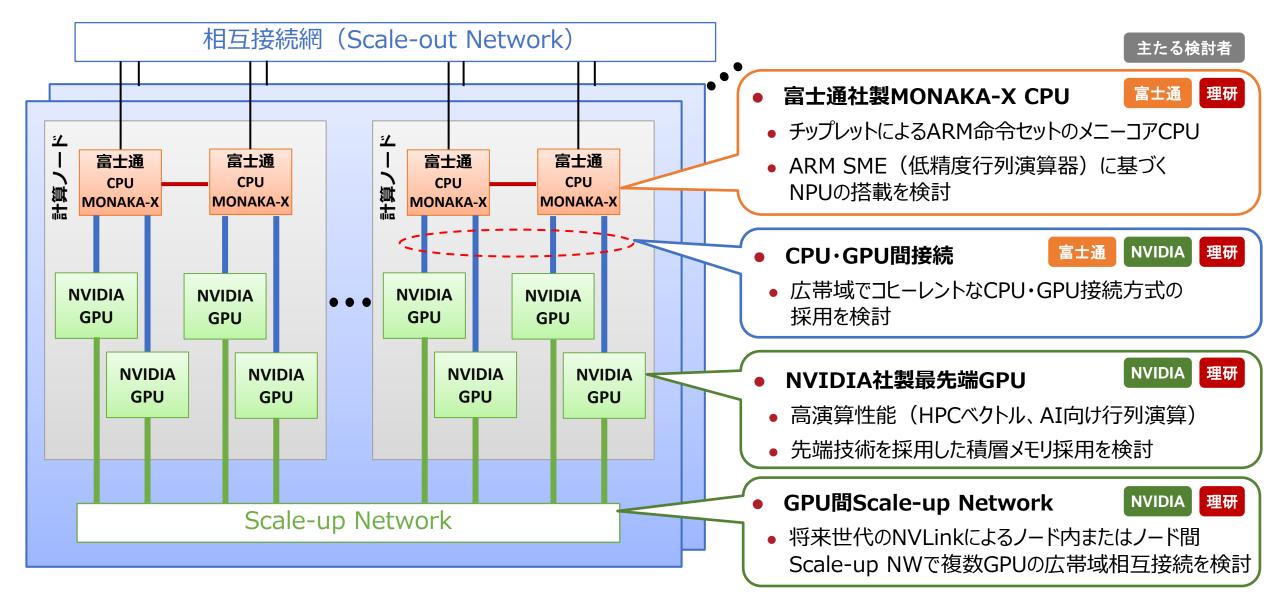
- 既存HPCアプリケーションで現行の5~10倍以上の実効計算性能
- AI処理でゼタ(Zetta) FLOPSスケールのピー ク性能を念頭に50EFLOPS以上の実効性能
- シミュレーションとAIの融合により、総合的に数十倍のアプリケーション高速化を目標

- アプリケーションの実効性能を最優先とする「アプリケーションファースト」の理念
- 電力制約下で目標を達成するため、「富岳」で培った アプリケーションソフトなどの資産を有効活用できる 電力効率の高いCPU部と帯域重視の演算処理加速部 を組み合わせた、高帯域およびヘテロジニアスなノード アーキテクチャを基本構成としたシステム

項目	CPU	加速部	「富岳」	「富岳」比
合計ノード数	3400ノー	-ド以上	158,976	
理論 FP64ベクトル性能	48 PFLOPS以上	2.6 EFLOPS以上	537 PFLOPS	x5.7 以上
理論 FP16/BF16行列演算性能	1.5 EFLOPS以上	150 EFLOPS以上	2.15 EFLOPS	x70.5 以上
理論 FP8行列演算性能	3.0 EFLOPS以上	300 EFLOPS以上	_	
sparsity考慮の 同理論性能		600 EFLOPS以上	_	
メインメモリサイズ	10 PiB以上	10 PiB以上	4.85 PiB	x4.1 以上
メインメモリバンド幅	7 PB/s以上	800 PB/s以上	163 PB/s	x4.9 以上
合計消費電力	40 MW以下(計算ノードおよびストレージ)		約30 MW	



### 「富岳NEXT」システム・アーキテクチャ概要





### 富士通 MONAKA-X 概要

### **FUJITSU-MONAKA-X(1.4nm)**

FUJITSU

国産の世界最先端プロセッサー for Al

#### NPUによるAI高速化処理

- ・サーバ用CPUで世界初、低精度行 列演算器 Arm SME を実装、 低レイテンシAI処理を実現
- ・AI高速化フレームワーク・ライブラリの 拡充

#### HPCに最適な高いスケーラビリティ

- 次世代3Dメニーコアアーキテクチャ による超メニーコア化
- SIMD拡張による更なる高速計算 処理の実現。
- ・ HPC向け高速コンパイラ・ライブラリの 拡充

#### GPUとの密結合

・GPUとの広帯域データ転送の採用 により、AI学習やGPUに最適化 されたアプリの高速処理を実現



#### 高い省電力性とセキュリティ

- ・先端半導体プロセス採用
- •超低電圧動作制御
- Confidential Computing



#### **HPC**

Big data processing

- 気象変動モデリング
- 新たな治療薬・方法開発
- 金融サービスの高度化 etc.



#### データセンタ

Scalability for Cloud Computing

- 省電力・省スペース
- AI学習・推論インフラ最適化
- 高度なセキュリティ性 etc.



#### エッジコンピューティング

Real-Time and Edge AI

- 安全保障
- 通信インフラ
- ロボティクス etc.

C 2025 Fujitsu Limited



### NVIDIA GPU 概要

## NVIDIA GPU 概要

富岳NEXT (2030年頃) に向けて、

アプリケーション性能を最大化するハードウェア・ソフトウェア・システムを

コデザインするとともに、GPUポーティングをはじめアプリケーションユーザーを支援。

CPUとの広帯域なコヒーレント接続、新たなメモリ技術や電力効率向上を目指した

最新技術など、幅広く検証し推進する。

<システム全体性能目標(加速部)>

FP64 Vector : 2.6EF 以上

FP16/BF16行列演算性能 :150EF 以上

FP8行列演算性能(疎性有り) :600EF 以上

: 10PiB 以上 メモリサイズ

メモリバンド幅 :800PB/s 以上



# 富士通 MONAKA-X+先端NVIDIA GPUが実現するHPC・AI融合 ~ 理研が主導するシステムソフトウェア開発戦略 ~

#### • アプリ開発者が容易に利用できるためのプログラミング環境

- OpenMP, OpenACC/Kokkos (Fortranを含む)等、幅広い開発フレームワーク環境に対応
- HPCとAIワークロードの両方を意識した最適化コンパイラ/ランタイム開発
- CPU/GPUハイブリッド最適化のためのデバッグイング・プロファイリングツール整備

#### ● 性能を最大限に引き出すための数値計算ライブラリ・ミドルウェア

- Ozakiスキームに基づく高精度演算エミュレーションでCPU/GPUの低精度演算器を活用
- 混合精度演算(FP8/FP16/FP64)を組み合わせた高速・高精度数値計算ライブラリ開発
- BLAS, FFT, Sparse Solver 等をCPU/GPU環境向けに最適化

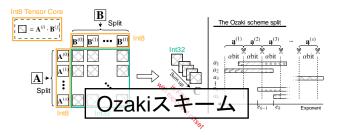
#### ● 大規模HPC・AI アプリのスケーラブルな実行を可能にする通信ライブラリ

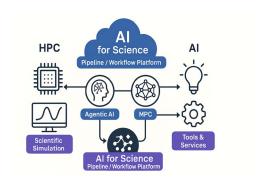
- NVLink等を想定したScale-up/Scale-out NWを最大限活用
- MPI, UCX, UCCなどOSS通信ライブラリの最適化と拡張
- 富岳NEXT向け超低レイテンシ・高スループット集団通信の実装

#### HPC・AIの有機的な融合を可能にするAI関連ソフトウェア

- Agentic AI, MPCを活用した科学シミュレーションと各種ツール・サービスの融合
- AI for Science向けPipeline/Workflowプラットフォームを構築し、
   個々のAIソフトウェアを有機的に結合して、自動化・統合運用可能な研究開発環境を提供



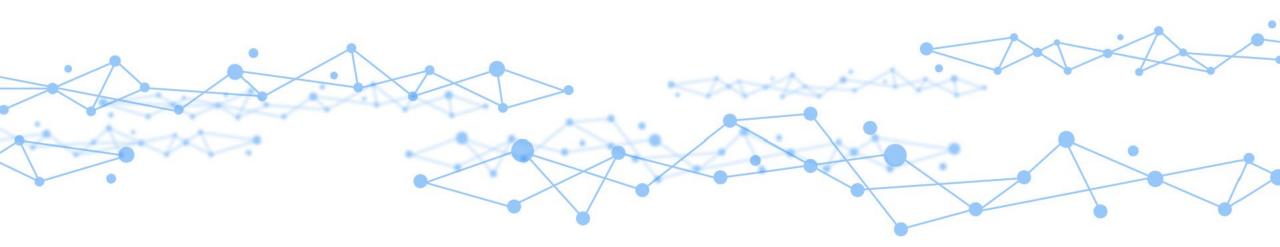




富士通 MONAKA-X と最先端のNVIDIA GPUのハイブリッド性能を最大限に引き出すため、 理研の主導によるHPC・AIの両面で最適化された統合システムソフトウェアを開発



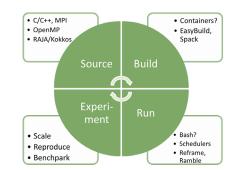
# コデザインの推進およびエコシステムの構築

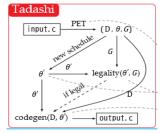


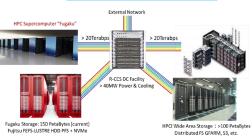


### コデザイン戦略および「富岳NEXT」エコシステム形成に向けて

- 幅広いアプリケーションでのコデザインの実施
  - CI/CD/CBフレームワークを用いた多様なアプリでのベンチマーキング
    - DOE/MEXTの枠組みで共同開発したBenchparkも活用
    - 国外のHPC関連研究所にも積極的に情報提供しコデザインを連携
- AI駆動のアプリ開発(コーディング/最適化)
  - R-CCS開発中のコーディング/最適化AIの利用やChatbotを活用
  - エージェントAI対応のHPCアプリ開発
- 「富岳NEXT」テストベッドシステムの構築・提供
  - 「富岳NEXT」に近い構成のシステムをテストベッドとしてユーザに提供
  - NISや「AI for Science」システムを含むGPUシステム/クラウド環境を活用
  - 潜在的な「富岳NEXT」ユーザコミュニティを形成
- バーチャル富岳への開発ソフトウェアの早期展開
  - ソフトウェア検証・高度化と「富岳NEXT」エコシステム拡大へ









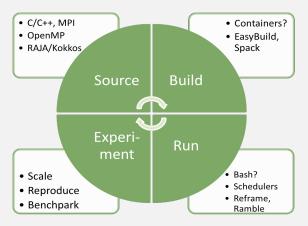
ツール等の展開について:今年度は富岳NEXTプロジェクトのアプリチーム内でテストを実施しつつ来年度以降に幅広く展開していくための体制構築と制度設計を行う



### 理研による先進的アプリ開発支援の設計・開発の具体例

### ① 移植・性能評価支援

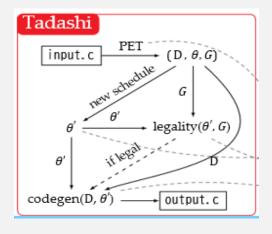
CI/CD/CB技術へ取り入れた移植・ 性能評価の自動化



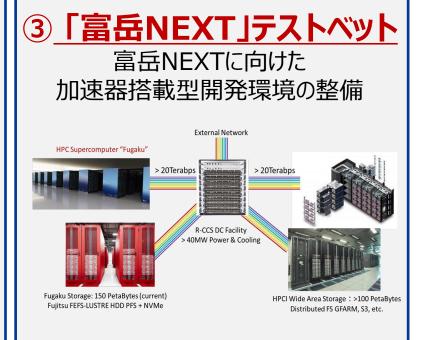
- DOEとの共同研究による自動ベンチマーキング環境(Benchpark)利活用による継続的な性能評価支援
- 幅広い分野のアプリケーション・コミュニティ とのコデザインを実施

### ② 高度化支援

AIチャットボットによる開発支援・AIによる自動プログラム最適化



- AIチャットボットによる開発・利用支援の 運用を「富岳」で運用開始
- 加速部向けの移植・最適化のためのAI によるコーディング技術 (Tadashi) を 開発中

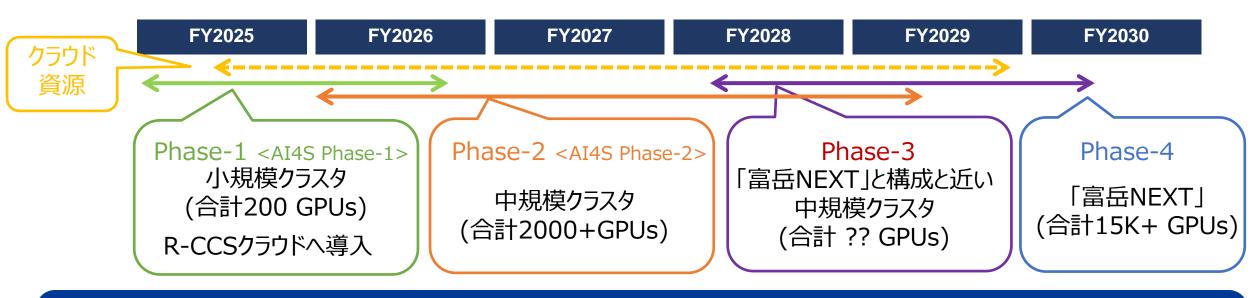


- 加速器を搭載したテストベッド環境を導入・整備 (「AI for Science」システム、 量子-HPC システムやクラウド環境等)
- 幅広いアプリケーションユーザへ提供しコデ ザインを推進

「富岳NEXT」の運用開始直後からの科学的成果創出に向け、各分野のロードマップも踏まえつつ、 オープンなアプリケーション・コミュニティとの密な連携体制を構築し、アプリケーション・ファーストによるコデザインを推進



### 「富岳NEXT」テストベッドの整備計画(案)



- Phase1 (2025年4月):さまざまなGPUからなる小規模クラスタシステム、「AI for Science」スーパーコンピュータ (AI4Sシステム)のPhase1として導入。バーチャル「富岳」向けのソフトウェア開発ベースとしても活用(合計約200GPU)
- Phase2 (2025年度末): AI4SシステムのPhase2と「量子HPC連携プラットフォーム」のGPUシステムを合わせた中規模クラスタ(合計約2000GPU)、CI/CD/CBプラットフォームとしても利用を想定
- Phase 3 (2027年度): 「富岳NEXT」に近いシステム構成を持つ中規模GPUクラスタ、開発システムソフトウェアやアプリケーションの評価・検証、CI/CD/CBプラットフォームや運用準備マシンとして利用
- Phase 4 (2030年度): 「富岳NEXT」システム自身
- (必要に応じて):最新GPUやネットワークテストのためNISのシステムやクラウド資源を活用



### 「AI for Science」開発用スーパーコンピュータ 概要

#### ■ 背景·目的

- AIとシミュレーションを密連携させた「AI for Science」推進のための新計算基盤
- 世界最高水準のAI性能と「富岳」等との協調で科学研究を加速
- 科学研究基盤モデルの開発・活用による研究サイクルの飛躍的加速・探索空間拡大

#### ■ システム構成

- NVIDIA Grace Blackwellスーパーチップ搭載ノード400台超(GPU 1,600基以上)
- ノード間接続: NVIDIA InfiniBand XDR (最大3.2Tbps)
- 性能: 64.16PFLOPS以上(FP64)、FP8 15.539EFLOPS以上(FP8)
- 温水冷却対応のSuper Micro製サーバを採用し高性能と省エネを両立

#### ■「富岳」との比較

- ◆ FP8演算性能は「富岳」の約7.23倍(「富岳」で代替可能なFP16の約2.15EFLOPSとの比)
- FP64を得意とする「富岳」との連携でAI×Science融合研究を実現

#### ■ 今後の予定

- 2025年度中に構築完了、2026年度初頭から本格運用開始
- 米国アルゴンヌ国立研究所(2024年MOU締結済)や他の機関との連携の発展・加速も期待

#### AI推論性能

15.539 Exaflops以上(FP8性能)





### 理研主導による持続可能なソフトウェア開発と標準化に向けた取組み

#### • 「富岳」から「富岳NEXT」へのスムーズな移行と高い互換性

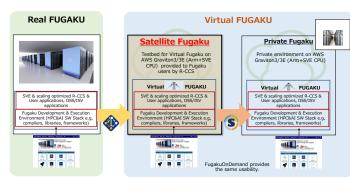
- 「富岳」の技術・経験を基に互換性と先進性を両立させた進化型のソフトウェア基盤を実現
- 「富岳」で動作する全アプリを富岳NEXTで即利用可にするためのシステムソフトウェア整備
- Open OnDemand (「富岳」のクラウド化) により使い慣れた環境を富岳NEXTでも継承
- 科学シミュレーションだけでなく、AI for Science、量子-HPC計算や融合計算にも対応

#### OSS利活用と持続的な発展

- システムソフトウェアは可能な限りOSSを採用し、開発したものも原則OSSとして公開
- 改変や拡張を通じて富岳NEXTや国内外コミュニティの発展に貢献
- 理研が既に参画しているHPSF (High Performance Software Foundation) など国際HPCコミュニ ティと連携し持続的な開発体制を確立(例: E4S、Spack、OpenChami、Kokkos)

#### クラウド・国内外基盤への展開

- SpackによるHPC環境のパッケージ化 により標準利用ソフト環境を提供し国内外へ共有
- バーチャル富岳(クラウドの「富岳」化)の実績を発展させ、富岳NEXTでもクラウド上に 次世代フラッグシップシステム相当の環境を展開















富岳NEXTでは、移行性・互換性・先進性を兼ね備えたシステムソフトウェアを核に 国内外のOSSエコシステムと標準化活動を通じて、理研主導による持続可能な開発・利用基盤を構築



### 「バーチャル富岳」によるエコシステム構築

- ●「バーチャル富岳」を基盤としたエコシステム構築
  - 「富岳NEXT」などさまざまな開発プロジェクトの成果 (ソフトウェア) を「バーチャル富岳」へ同梱、世界中へ提供
  - テストベッド・利用環境としてクラウドなど多様な環境へ「バーチャル富岳」を展開⇒「様々なプロジェクト成果」を高度に精錬するとともに、ユーザーコミュニティ構築へも貢献
  - 「富岳」などスパコンシステムの**運用の一貫として、「バーチャル富岳」の持続的なエンハンスを保証**
- ソフトウェアの開発・評価・利用による価値創出のサイクルを持続する仕組み=「バーチャル富岳」 により、スパコン向けソフトウェアのエコシステムを構築

#### 持続可能なソフトウェア環境による新たな価値創造

- 世界トップ・レベルのサイエンス・プラット フォーム「富岳」や「富岳NEXT」で修練され た高度なソフトウェア環境を容易に利用可能に
- ⇒ スパコン活用による新たな価値創造へ貢献

#### 多様なテストベッド・利用環境での 評価実績により利用者を拡大

- 「富岳NEXT」テストベッドへの適用
- AWS (展開済) などクラウド環境への展開
- 同梱ソフトウェアを修練するとともに利用者を拡大⇒ 各ソフトウェアの価値を向上



⇒ 運用業務の延長で共通基盤であるソフトウェ アの維持管理を持続的に実現

#### 開発成果の集約・提供

「富岳NEXT」等ナショナルプロジェクト (コデザインやAI活用開発、等)の成果を集約

- 多様な開発コミュニティや標準化団体と協働
- SPACK・コンテナ技術を活用し集約・公開





### 「富岳NEXT」に向けた運用技術・施設整備の方向性

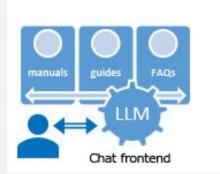
### 持続性/継続性



- ●「端境期」を極力生じさせない利用環境を整備し、世界最高水準の**計算性能と計算 資源量を継続的かつ安定的に提供**
- バーチャル「富岳」を発展させ、アプリケーションの継続的利用環境を確保
- 計算資源のみならず、研究データについても持続的な提供やデータ共有・連携環境の整備を目指す

### 運用の高度化





- HPCI整備計画調査研究事業と連携し、 NISとの利用環境共通化や異種計算基盤 の一体型運用を推進
- R-CCS内の量子コンピュータやAI-for-Scienceマシンとは密接な連携により、次 世代計算基盤を確立
- AI技術の積極的な導入により、運用の省力化と最適化を実現

### <u>低炭素化</u>



- 最新の冷却技術の導入や「富岳」で実証した省電力機構の高度化により、高エネルギー効率と低カーボン排出を実現
- スーパーコンピュータ本体だけでなく、データセンター全体としても運用コストの効率化・ 最適化を推進
- 蓄電技術・再生可能エネルギー・廃熱再利用による環境負荷の低減を検討



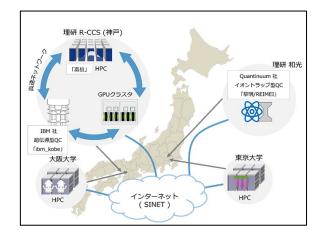
### 「富岳」と量子コンピュータとの連携(量子古典ハイブリッド)

#### 「富岳」と最先端量子コンピュータ(QC)による世界有数の量子・HPC連携の計算研究環境を構築

- 「富岳」とIBM Quantum System Two 「ibm\_kobe」を同じ建物に設置し、 低遅延の高速ネットワークで密に結合した計算環境。
- 高速で極めて低いゲートエラー率等の優れた性能を備え、大規模なワークロードを実行可能な最新プロセッサ「IBM Heron (156量子ビット)」搭載。
- 超伝導型の「ibm\_kobe」やイオントラップ型の「黎明」と、「富岳」を用いた量子・HPC連携プラットフォームの構築・運用(ソフトバンク・大学等の国内機関や米国企業(IBM、Quantinuum社)等と連携・協力)。



IBM Quantum System Two "ibm\_kobe"



量子HPC連携プラットフォーム

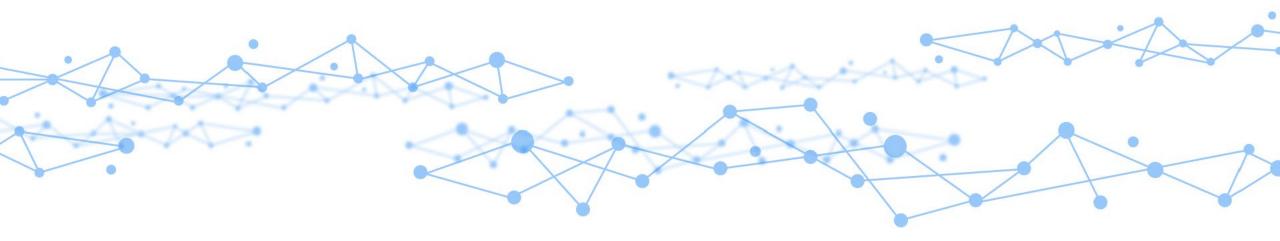
これまでのスーパーコンピュータだけでは実現できなかった 計算可能領域の拡大

量子古典ハイブリッドによる "Quantum Utility"量子計算の有用性を実証へ

「富岳NEXT」を核とした次世代の「HPC-AI-QC」プラットフォームの構築を実現し 科学技術と産業発展、さらには地球規模の課題解決に貢献



# アプリケーション高度化への道





### アプリケーション性能向上に向けた取組み

- アプリケーション開発者から見たテーマ
  - 開発: CPU→GPU化、GPUテンソルコア/低精度演算を活用するアルゴリズム、AI/学習による加速、チューニング
  - 開発環境活用: CD/CI/CBシステム、テストベッド(「富岳NEXT」に近いシステム)、AI駆動のアプリ開発環境
- 開発者サポート タウンホール/情報提供
  - 「富岳NEXT」アプリケーションセミナー開催 1~2回/月 (第4回まで開催済み)
  - 環境、情報活用のための各種ハッカソン開催
  - 例:次世代GPUシステム GPUの帯域を最大活用するための手法 ベンダーと連携
- モデルケース構築/活用
  - **モデル先行ケース**として 10数件程度のアプリ (コミュニティ推薦)
  - 開発ヒストリのAI学習; 得られた情報の公開、活用 → 幅広いアプリ開発に生かす
  - 開発アプリ → コデザイン による次世代計算基盤設計の最適化
  - 「富岳NEXT」に留まらず永続的な計算機開発に生かしていくフレームワークとしての開発ループ設計

#### 連携

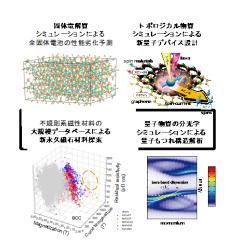
- ・ベンダーや
- 「次世代HPC·AI開発支援拠点」(文部科学省 委託事業)との連携
- 国際連携 (検討中): 「富岳NEXT」開発連携プログラム(仮称) 海外のスパコンセンターと連携
- オープンシステム だれもが参加できる ←→ 先端開発機密保持-の両立



### 次世代計算基盤で期待されるブレークスルー(例)

#### 新物質・エネルギー

- ・「富岳」の5-10倍の規模の数値計算が可能となることで、大規模磁性材料データベース構築、全固体電池・太陽電池材料のAI加速第一原理計算、トポロジカル量子物質の超高速応答およびレーザーと量子物質の相互作用のシミュレーション、さらには量子物質内の量子縺れ制御のための数値実験の実現を目指す。
- ・高性能モーターや高効率太陽電池、高性能・高信頼 二次電池、省電カスピンデバイス、量子デバイス材料 の探索・実装に必要な時間・資源の削減が期待される。 またデバイス材料の性能劣化予測が可能となることが 見込まれ、次世代計算基盤が安定的なエネルギー網の 構築に貢献することが期待される。

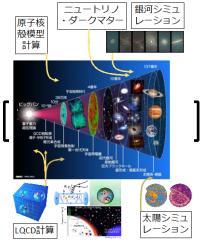


#### 基礎科学分野

- 素粒子・原子核・宇宙・惑星の各分野において、 「富岳」の約10倍の計算能力によって可能になる:
- ・LQCDから標準模型検証とそれを超える物理の探索、 宇宙や物質のなりたちの解明に迫る第一原理計算群例: カイラルフェルミオンによる u,d,s,c,~b クォーク 上計算(現行アルゴリズムで「富岳」x50倍程度)を

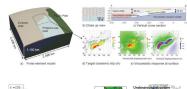
#### 低精度演算加速/AI加速による実現を目指す

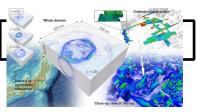
- ・中性子過剰核の存在限界やクラスター構造の解明
- ・銀河全体→分子雲内部構造:一貫したシミュレーション
- ・宇宙に存在する多様な銀河の形成進化史に迫る
- ・太陽内部:磁場生成から黒点形成まで一貫して追う
- ・銀河団・銀河群スケールのダークマターハローを 分解できる(ニュートリノDM)数値シミュレーション



#### 地震·津波防災

- 広域のマクロ地殻変動シミュレータと特定領域のミクロな地 殻変動・地震動シミュレータを組み合わせたマルチスケール シミュレータを開発することで、地殻変動・地震動の観測 データに整合するモデルを構築。これにより、一定規模の地 震が発生した後、周囲での地震発生の推移を予測できる か検証できるようになると期待
- 上記の実現に向けて、**富岳の5-10倍程度**のシミュレーション性能の実現を目標とする。そのため、次世代計算基盤のシステム特性に合わせた物理シミュレーション、及び、AI等のデータ駆動型のアルゴリズム・実装を開発

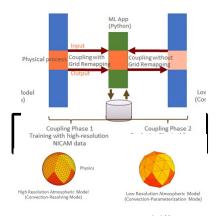




地殻変動と地震動のマルチ スケールシミュレータ

#### デジタルツイン·Society5.0

- ・高解像度シミュレーションとAIの連携により、 データ駆動型のパラメタリゼーションモデルを構築 し、これを活用した超高解像度メッシュを用いずと も高精度な全球気候モデルの計算を実現。富岳の5 倍程度の計算性能とAIを併用することで、現行の 高空間解像度計算よりも100倍の高速化と数百年規 模の長期気候計算の実現。
- ・富岳の10倍程度の計算が可能となることで、東京都心部の10km四方に対して、IoT機器のデータをリアルタイムに同化をともなった25cm解像度のリアルタイム風況デジタルツインを実現。これは現在のシミュレーションで達成できる典型的な解像度の約8倍の高精細計算。



シミュレーションとAIの連携による気象計算の高速化

│ ポスト富岳FS - アプリケーション調査研究グループ 2024年度まとめより(一部修正)



### 次世代計算基盤で期待されるブレークスルー(例)

2011∼ 「K computer」







2020∼ 「Fugaku」

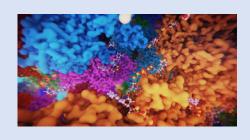




2030 [FugakuNEXT]



#### 細胞内分子動態シミュレーション

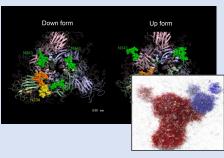


「京」では1億原子系の短い時間の

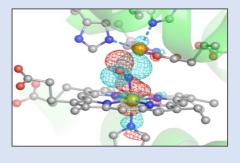
全原子分子動力学 計算の高速化 (100倍以上)



長時間ダイナミクスと細胞機能 マルチスケールモデルの利用



「富岳」ではさらに大きな系の 長時間ダイナミクスが可能に



電子状態を考慮したダイナミクスが可能に 生体デジタルツイン抗体創薬など

### 自動車空力

ダイナミクスが実現



高解像度LESによる風洞代替など 基礎研究



デジタルツイン(上) AI支援多目的最適化(下) による自動車設計の期間短縮



生成AIによる最適形状の提案 などによる自動車設計の自動化

AIによる自動運転技術の確立



### 次世代計算基盤で期待されるブレークスルー(例)

2011∼ 「K computer」









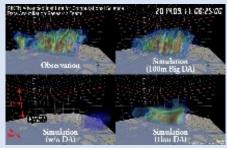




2030 [FugakuNEXT]



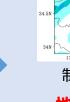
#### 気象·気候



「京」を用いたゲリラ豪雨予報 手法の開発



2021年東京オリンピック・パラリンピック期間中「富岳」によるリアルタイムゲリラ豪雨予報(世界初)



制御なし \*制御なし

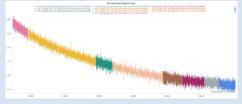
地球規模の気候危機の解決

社会や都市のデジタルツインやAI との融合により、政策の仮想試行 および提言へ

#### Fugaku LLM (130億パラメタ)

対象モデル	学習可能 トークン数
13B Transformer モデル	230B トークン

Fugaku LLMの事前学習に、 およそ "10~15年" 要する



Fugaku LLMの事前学習を、 富岳の 1/11 の規模の利用で **"1ヶ月強"** で完了



Fujitsu Research Portalで 無償利用可能 米国SambaNova社も商用基盤で提供

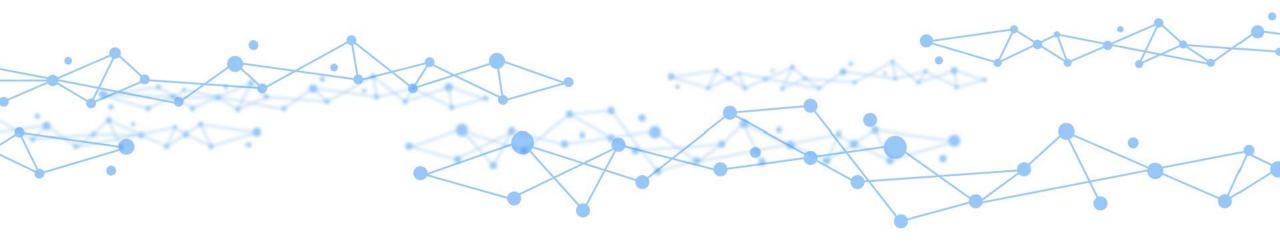
https://portal.research.global.fujitsu.com/

最先端の〜**兆級パラメタ**の 基盤モデルの 事前学習を**2か月以下**で完了

AI for Science加速によるイ ノベーションサイクルの劇的進化



# 参考資料





### 尾崎スキームIおよびIIの動向

	RTX4090		GH200		Rubin 300 Ultra (estimate by S. Matsuoka)	
DGEMM (cuBLAS)	0.62	TFLOPS	60.9	TFLOPS	??? 100+ TFLOPS per 4 GPU diextrapolation from I	
Ozaki Scheme I	5.84	TFLOPS	34.5	TFLOPS		Rubin Ultra
Ozaki Scheme II (new)	7.4–9.8	TFLOPS	66.9	-80.2 TFLOPS	800~1300 TFLOPS per 4 GPU di	4 Reticle-Sized GPUs 100PF FP4   1TB HBM4e ie package

With Ozaki Scheme II, emulation outperforms native FP64 even on data center GPUs. Consumer-grade GPU  $\rightarrow$  RTX4090: INT8TC: FP64 = 512:1 DATA Center GPU  $\rightarrow$  GH200: INT8TC: FP64TC = 29.5:1 DATA Center GPU  $\rightarrow$  B200: INT8TC: FP64TC = 112.5:1

Code is available at RIKEN's GitHub presented by Dr. Uchino

https://github.com/RIKEN-RCCS/accelerator\_for\_ozIMMUhttps://github.com/RIKEN-RCCS/GEMMul8



### 「富岳NEXT」の開発推進体制

● 次世代のフラッグシップシステムの開発・整備や、その先のフラッグシップシステムに向けた最先端の要素技術の開発・ 調査研究、TRIPとの協業や、国際連携等を進めるためのハブ機能を果たす組織として、計算科学研究センター (R-CCS) に新たに「次世代計算基盤開発部門(部門長:近藤正章)」を2025年4月に設置。

● オープンで開かれた開発環境を構築するとともに、ユーザー・コミュニティとの対話を定期的に進め、「富岳NEXT」プロ

ジェクトを推進。

#### 文科省-DOE連携のPAの署名(2024年4月9日)



大学•研究機関

理化学研究所 計算科学研究センター



タウンホールミーティング の積極的な開催

次世代計算基盤 開発部門

ベンダー群

米国・エネルギー省(DOE) Labs.等

情報基盤センター群

ユーザー・コミュニティ

文科省-DOE連携ワークショップ (2023年2月8日)



















センター長 松岡 聡



副センター長



副センター長 杉田 有治 (兼務) (計算による科学)



プロセッサ研究チーム 佐野 健太郎



大規模並列数値計算 技術研究チーム 今村 俊幸



次世代高性能 アーキテクチャ研究チーム 近藤 正章 (兼務)



高性能ビッグデータ 研究チーム 佐藤 賢斗

#### 2025年8月1日 組織図



高性能人丁知能 システム 研究チーム Mohamed WAHIB



高性能計算モデリング 研究チーム Jens DOMKE



大規模デジタルツイン 研究チーム



高性能クラウドシステム: セキュアソフトウェア 研究チーム 竹房 あつ子

19チーム 5部門 (1T、22U、1室)



山口 弘純





計算

の 科学



連続系場の理 論研究チーム 青木 保道



離散事象シミュレー ション研究チーム 伊藤 伸泰



量子系分子科学 研究チーム 中嶋 隆人



量子系物質科学 研究チーム 柚木 清司



粒子系生物物理 研究チーム 杉田 有治



複合系気候科学 研究チーム 富田 浩文



複雑現象統一的解法 研究チーム 坪倉 誠



データ同化研究チーム 三好 建正



計算構造生物学 研究チーム Florence TAMA



総合防災・減災 研究チーム※ 富田 浩文 (兼務)



デジタル材料科学 研究チーム 常行 真司

※専任PIを今後公募予定。

「富岳」Society 5.0 推進拠点



拠点長 松岡 聡 (兼務)



副センター長& 拠点長代理 村山 綾介



コーディネーター 白井 宏樹

HPC/AI駆動型医薬 プラットフォーム部門



部門長& バイオメディカル 計算知能ユニット 奥野 恭史



副部門長 本間 光貴



創薬化学AI アプリケーション ユニット 池田 和由



分子デザイン 計算知能ユニット 池口 満徳



AI創薬連携基盤 ユニット 奥野 恭史 (兼務)

量子HPC連携 プラットフォーム部門



部門長 佐藤 三久



副部門長 児玉 祐悦



副部門長 小野寺 民也



量子HPCソフトウェア 環境開発ユニット 辻 美和子



量子計算 シミュレーション 技術開発ユニット 伊藤 伸泰 (兼務)



量子HPC プラットフォーム 運用技術ユニット 三浦 信一(兼務)

AI for Science プラットフォーム部門



部門長 松岡 聡 (兼務)



AI開発計算環境 運用技術ユニット 三浦 信一 (兼務)



次世代AIデバイス 開発研究ユニット 佐野 健太郎 (兼務)



AI学習最適化 基盤開発ユニット Mohamed WAHIB (兼務)





AI for Science基盤 モデル研究チーム 横田 理央



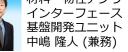
AI学習・推論データ 管理基盤開発ユニット 佐藤 賢斗 (兼務)



生命・医科学アプリ インターフェース 基盤開発ユニット 松永 康佑



材料・物性アプリ インターフェース



次世代計算基盤開発部門



部門長 近藤 正章



副部門長 庄司 文由 (兼務)

次世代計算基盤 システム開発 ユニット 佐野 健太郎 (兼務)

次世代計算基盤 アプリケーション 開発ユニット 青木 保道 (兼務)



次世代計算 基盤運用技術 ユニット 山本 啓二 (兼務)



先進的計算基盤技 術開発ユニット 佐藤 賢斗 (兼務)



嶋田 庸嗣 (兼務)

運用技術部門



部門長 庄司 文由



副部門長& システム運転 技術ユニット 井口 裕次



施設運転 技術ユニット 三浦 信一



ソフトウェア開発 技術ユニット 村井 均



データ連携 技術ユニット 甲斐 俊彦



先端運用 技術ユニット 山本 啓二