

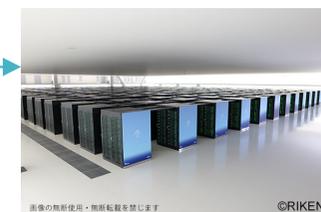
NFS及びNISのGPU化対応と 次世代HPC・AI開発支援拠点事業について

朴泰祐

次世代HPC・AI研究開発支援センター・センター長
(筑波大学計算科学研究センター)

国内フラッグシップスーパーコンピュータNFS (National Flagship System) の変遷とNISシステムのGPU化

System	Year	Processor	Architecture	Peak Performance
(Earth Simulator)	2002	NEC Vector (SX-5 base)	Vector	40TF
K	2012	Fujitsu SPARC64 VIIIfx	Multicore CPU (8c)	10PF
Fugaku	2020	Fujitsu A64FX	Manycore CPU (48c)	488PF
Fugaku-NEXT	2030?	Fujitsu MONAKA + NVIDIA GPU	CPU+GPU (multi-socket/node)	??? (for HPC + AI)

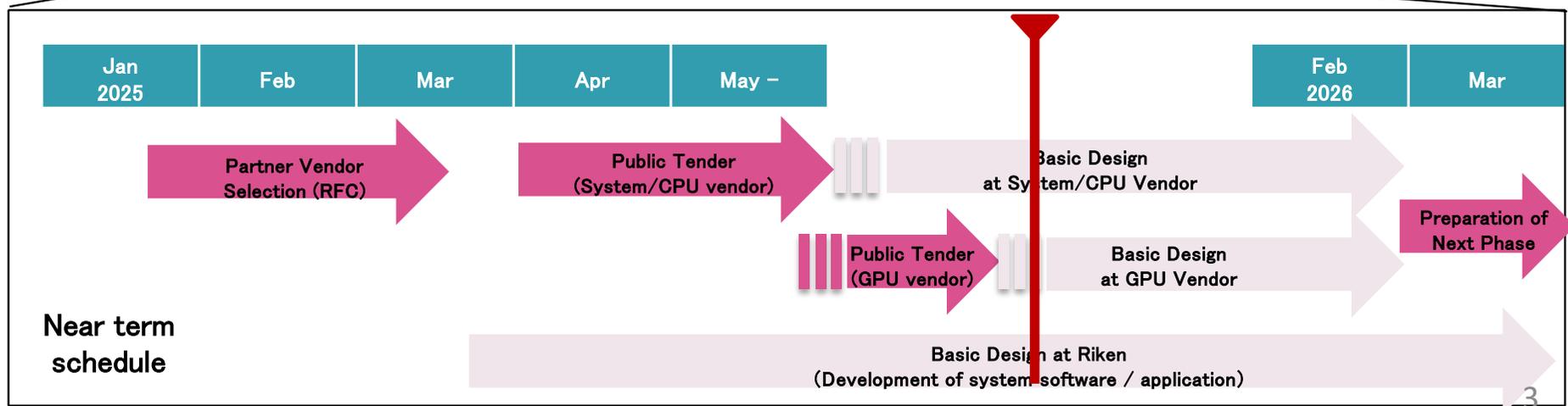
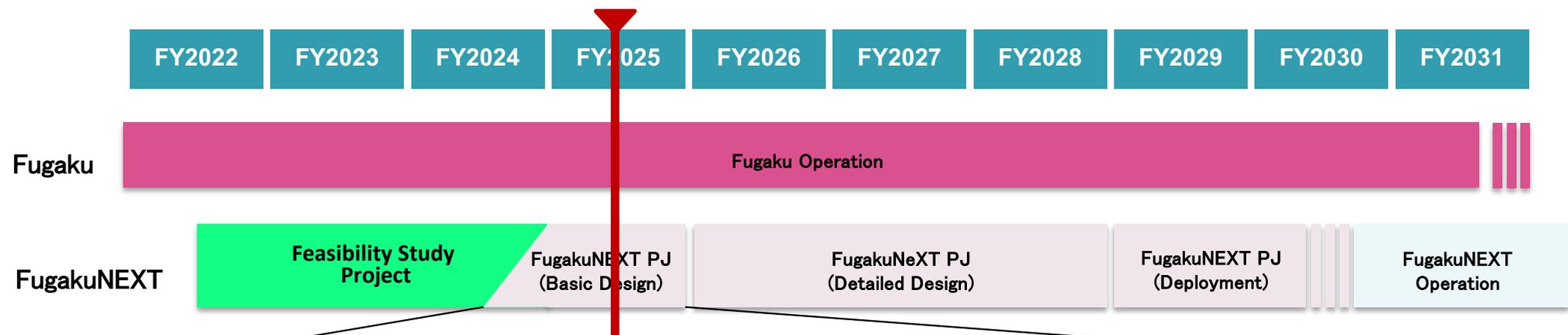


- HPCIにおけるNFS (National Flagship System)となる富岳NEXTとNIS (National Infrastructure Systems) のGPU化に対応するためのGPUコード開発が急務
 - 各センターは独自のGPU教育資料の作成やハッカソンの開催を実施
 - ユーザ向けの最先端GPUコンピューティングに対する統一的教育プログラム・コースが必要, またCPU・GPU統合のような先進的システムへの対応も課題 (例: GH200, MI300A)
- 世界的に進められているHPC/AI/AI4S (AI for Science) にキャッチアップするためにGPU教育とコード開発を支援する中心的組織が必要

(Slide courtesy by M. Kondo@RIKEN)

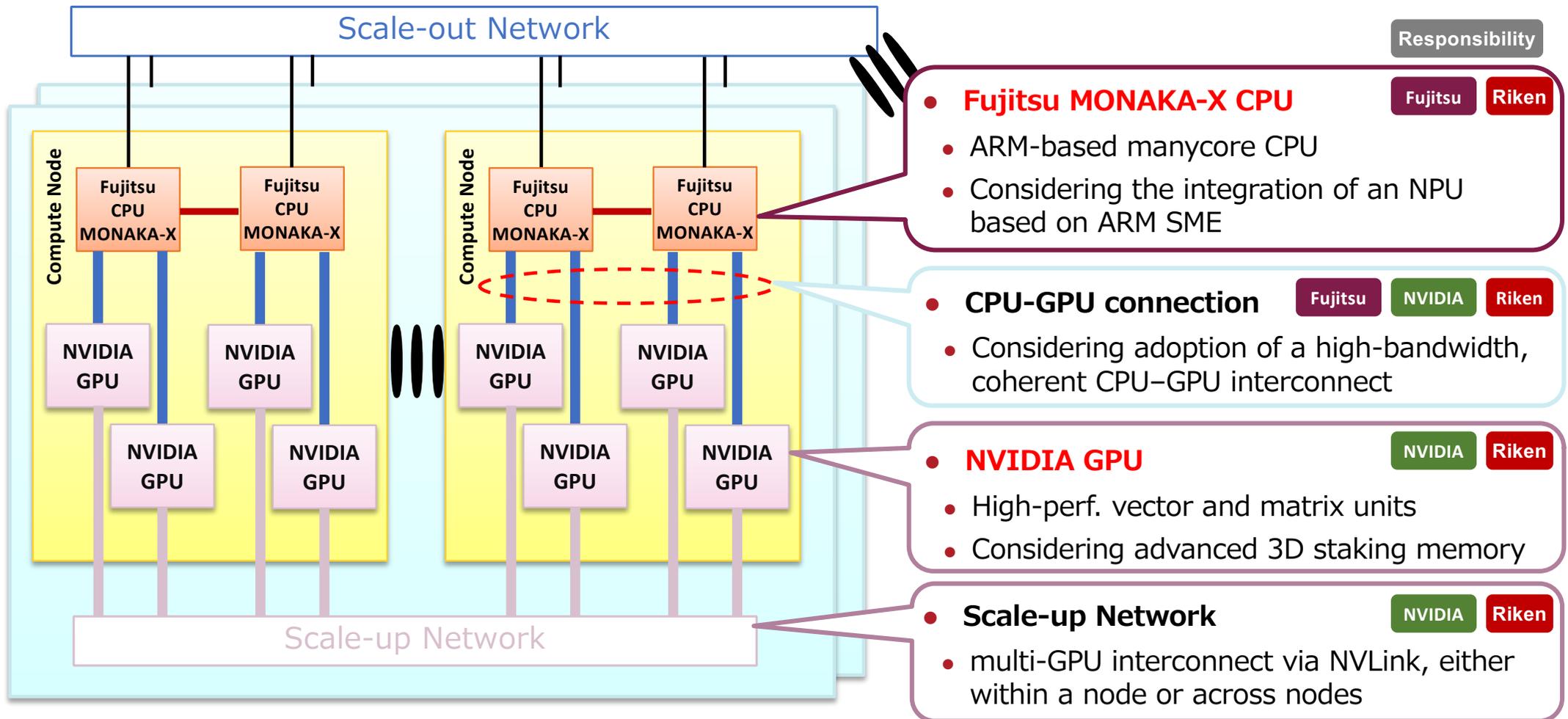
Expected Timeline of FugakuNEXT R&D

- FugakuNEXT development and deployment schedule



Overview of FugakuNEXT System Architecture

(Slide courtesy by
M. Kondo@RIKEN)



Detailed system architecture configuration will be determined in basic design phase 4

NIS (National Infrastructure Systems)の状況

- GPU化を先進的に進めているセンター(GPUが基本)
 - JCAHPC(筑波大学+東京大学):Miyabi-G
 - 筑波大学:Pegasus, Sirius (2026/2~)
 - 東京大学:Wisteria/BDEC-01 Aquarius, mdx
 - 東京科学大学:Tsubame-4.0
- GPU化を進めているセンター(CPUが基本だが一部GPUを導入)
 - 北海道大学: Grand Chariot 2 (GPU部)
 - 京都大学: Gardenia
 - 名古屋大学:不老 Type II
 - 九州大学:玄界ノードグループB
- GPU以外の演算加速アーキテクチャが主のセンター(ベクトルプロセッサ)
 - 東北大学:AOBA
 - 大阪大学:Squid
- 大規模GPUシステム導入を予定しているセンター
 - 理化学研究所計算科学研究センター:AI for Science system(仮称)

電力当たり性能の向上, 絶対的演算性能, メモリバンド幅重視等の理由でGPU化が進んでいる

文科省「次世代HPC・AI開発支援拠点形成事業」

- 「次世代計算基盤のユーザビリティに関する提言」、「次世代計算基盤に関する調査研究(Post富岳 Feasibility Study)」報告より、演算加速機構を持つ次世代NFSの開発とこれを支えるソフトウェア技術、人材育成、ユーザビリティの向上が不可欠という結論
→ 演算加速機構を持つシステムへのアプリケーション移行と人材育成等を目的とする事業が発足
- 2025年7月に公募開始
 - 8月4日提案締め切り
 - 9月上旬, 書類審査を経てヒアリング
- RIST(代表機関) + 筑波大・東大・東京科学大(中核機関)による提案を申請
→ 9月下旬, 採択(一部条件付き) → 「次世代HPC・AI研究開発支援センター」発足へ
(2025年10月~2030年3月)
- 提案の骨子
 - 本事業では、これまで「京」・「富岳」を中核とする HPCI において開発が進められてきた**既存アプリケーション**のほか、ポスト「富岳」時代を見据えた新たな科学研究分野(AI for Science等)における**新規アプリケーション**について、新たなフラッグシップシステムにおいて導入が予定されている **加速部への対応等**をはじめとした**次世代のHPC・AI開発の技術支援を行うための拠点を形成する**。本事業においては、ポスト「富岳」及び現在また今後 HPCI システム提供機関において運用される**スーパーコンピュータにおける演算加速部は主としてGPUによって構成されるものと想定する**。
 - 技術支援においては、**CPU・GPU・インターコネクト各部の開発ベンダー**等とも連携して知見の蓄積を図るとともに、国内の**計算科学分野の研究者や民間企業のユーザ**に対する**普及啓発活動**等を実施する。また、文部科学省にて同時平行的に実施されている**HPCI整備計画調査研究事業の各プログラム**とも連携して**分野横断型のコミュニティ**を形成し、効果的かつ効率的に**技術支援を進めるとともに、人材育成に資する取組を実施・推進する**。

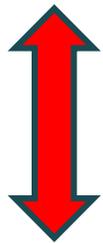
次世代HPC・AI研究開発支援センター (HAIRDESC) 及び中核機関・協力機関

代表機関 (RIST)
 - HAIRDESC: Advanced HPC-AI Research & Development Support Center
 (センター長: 朴泰祐
 副センター長: 下司雅章)
 神戸・ポートアイランドに設置



協力機関

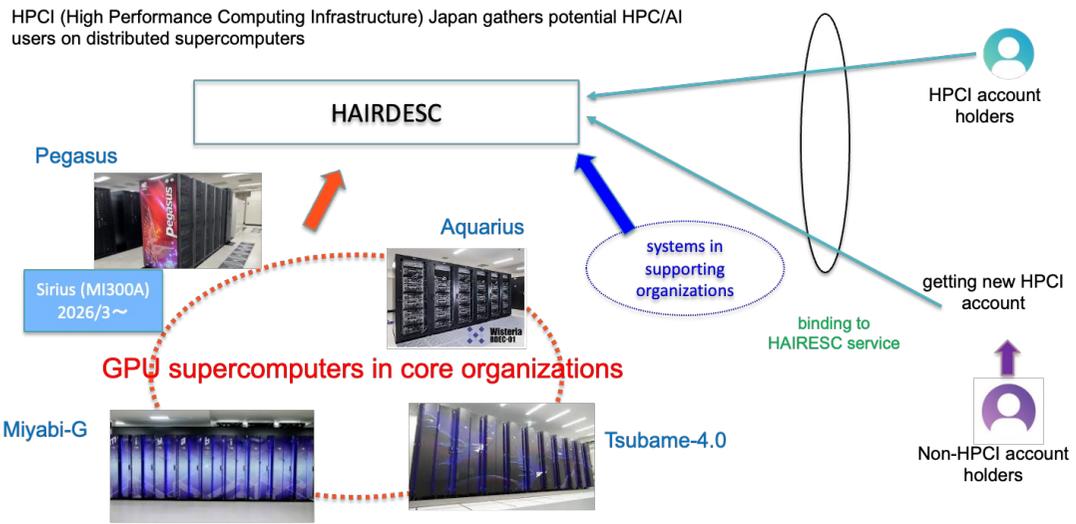
- 国立大学スーパーコンピュータセンター
 北海道大学, 東北大学, 名古屋大学, 京都大学, 大阪大学, 九州大学
- 理化学研究所計算科学研究センター (R-CCS)
- GPUベンダー: AMD社, NVIDIA社



GPU技術・GPU教育・人材の全ての活動について協力し合い, 教育教材・教育用GPUリソースを共有, ハッカソン等のイベントを実施

中核機関

- 筑波大学(計算科学研究センター)
 PI: 額田彰
- 東京大学(情報基盤センター)
 PI: 下川辺隆史
- 東京科学大学(総合研究院スーパーコンピューティング研究センター)
 PI: 横田理央



事業実施組織と支援体制

- **代表機関: 高度情報科学技術研究機構 (RIST): HAIRDESC**
 - 拠点事業全体の取りまとめ、予算配分、中核機関を中心とする各種運営会議体の運営
 - GPU教育コース資料(プレゼン、ドキュメント、ビデオ等)の一元管理とGPUシステム利用ユーザ向け配布
 - オンデマンド方式でのGPU教育コース資料の閲覧システム構築
 - 研究員等の雇用による特定アプリケーションのGPU化支援
- **中核機関: 筑波大学、東京大学、東京科学大学:**
計算科学研究センター, 情報基盤センター, 総合研究院スーパーコンピューティング研究センター
 - GPU教育コース資料の体系化: 既存資料の整理と補填資料作成とアーカイブ
 - GPU化コーディングのチュートリアル、ハッカソンの運営
 - GPUスーパーコンピュータ資源の提供
 - 特定アプリケーションのGPU化支援: 外部業者委託によるGPUコード作成の監督と技術的支援
 - 先進的GPUコード作成のための研究開発

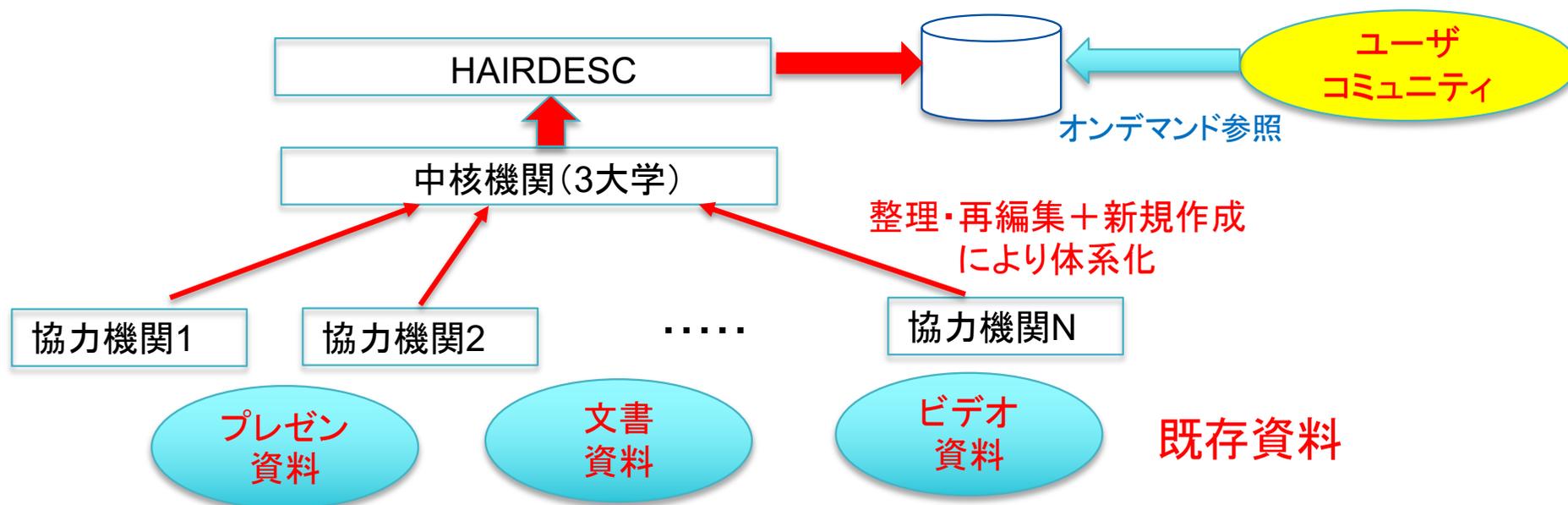
- **協力機関:**
 - **大学:** 北海道大学、東北大学、名古屋大学、京都大学、大阪大学、九州大学
 - GPU教育コース教材の提供: 既存、新規作成
 - **GPUスーパーコンピュータ資源の提供:** 各大学のプログラムにより有償
 - 研究員等の人材の斡旋
 - **アプリケーションユーザの本事業参加の斡旋**
 - **理化学研究所計算科学研究センター(R-CCS)**
 - **ポスト富岳の設計・開発における技術連携**
 - **特定アプリケーションのGPU化の相互協力**
 - 研究員等の斡旋
 - **企業:** AMD、NVIDIA
 - GPUコード開発のための**技術・資料の提供**
 - **最新GPUの動向についての技術情報提供**
 - ハッカソン等における**講師の供出**
 - **特定アプリケーションのGPU化における開発業者の斡旋**
- 協力機関についてはR8年度以降、必要に応じて拡大する可能性あり
 - 計算科学分野コミュニティ
 - 国立研究所
 - 企業(例: Intel)

アプリケーション開発者支援・分野横断コミュニティ形成

- 既存(主としてCPU向け)コードのGPU化のための教育とコード開発支援
 - 国内における「標準的GPU化」コースの策定と啓蒙・教育
 - 特定ベンダーのGPUに限らないアーキテクチャ想定: NVIDIA & AMD
 - レベルに応じたチュートリアル／ハッカソンコース設定
 - 初級: GPUとは何でありどのようにプログラムするか
 - システムの理解: 動作原理、アーキテクチャ、典型的GPUシステム紹介
 - コーディングの基本: OpenACC, OpenMP, CUDA, HIP等
 - MPIを代表とする並列通信とGPUの組み合わせ
 - 中級: GPUコードの高速化のノウハウ
 - パラメータ設定、ハードウェアリソース制約、効率的なループ、データ構造
 - 通信と演算のオーバーラップ、データ共有方法等
 - 上級: 先進的GPUテクノロジーを活かした高度なチューニング
 - MPS, MIGのようなGPUリソース分割共有手法の活用
 - GPU・CPU一体モジュール(GH200, MI300A等)の高度利用
 - 実アプリケーション例を用いたチューニングのノウハウ、bad example等
 - 数ヶ月毎のハッカソンを実施しオンラインを含むハンズオンセミナーによりユーザとのコミュニケーションを取りつつ相談に乗る(GPUミニキャンプのイメージ)
 - Slack等を利用したセミオンラインでのQ&A対応
 - 全てのイベントをオンライン化し、ビデオ記録してアーカイブ

GPU教育のためのマテリアルの集約・整理・アーカイブ

- 中核機関・協力機関におけるこれまでのGPU教育マテリアルを集約し、体系化された教育プログラムのために整理する
- GPU教育マテリアルを代表機関に設置したサーバにアーカイブし、ビデオ教材を含めオンデマンドで取得・視聴可能にする
- シンポジウム、GPUミニキャンプ等のあらゆるビデオ録画アーカイブも



- **分野横断的GPUユーザコミュニティの形成と人材育成**

- 年2回程度のシンポジウムを開催し、本事業におけるGPU化の成功例や問題となった点を一般公開して議論
 - ⇒ ハイブリッド開催と録画・アーカイブ
- シンポジウムで分野横断的な議論の場を設け、アプリケーションは異なるがGPU化手法が共通するプログラム等を通じて分野横断的な連携コミュニティを構築
- GPUハッカソン等を通じ、各分野のアプリケーションをGPU化してだけでなく、GPU化自体を技術的に体得し、今後のGPU化に資する人材を育成
 - ⇒ 日本における「GPU使い」を育成
- 国内だけでなく国際的な連携も積極的に進める
 - GPU先進国の大学・研究機関との個別交流の推奨（交流旅費の援助等も検討）
 - シンポジウムでの招待講演等
 - GPU化された国産コードの海外利用の斡旋
 - 海外のGPU利用技術の紹介

• GPUプログラミング教育教材の策定

- GPUプログラミングに関する教材は従来, それぞれのセンターや特定コミュニティ, ベンダー等によってバラバラに管理・配布され, ユーザから見ると情報が重複していたり不足していて, 教育効率が悪かった
- HAIRDESC及び3中核機関で散在しているマテリアル(スライド, ドキュメント, ビデオ等)を集約・再編し, 過不足のチェックと必要に応じた補填を実施
→ 教育マテリアルのスリム化・ワンストップ化・効率化

• 直交する様々な技術内容

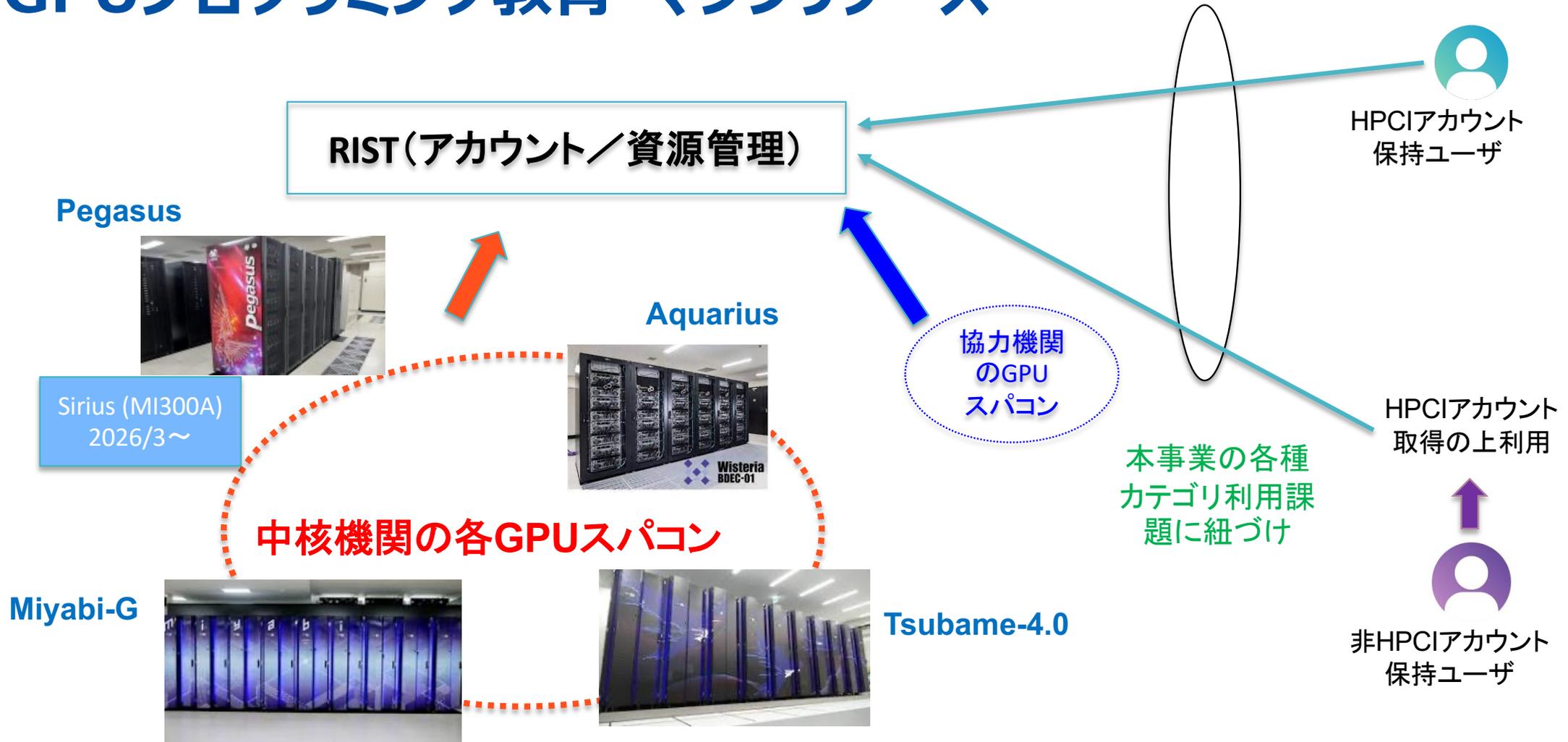
- 言語の抽象度
 - トップレベル: OpenACC, OpenMP, Python等
 - 低レベル: CUDA, HIP等
- GPU並列度
 - 単一GPU
 - ノード内複数GPU
 - ノード間(大規模)GPU
- ベンダー依存
 - NVIDIA
 - AMD
- CPUとの関係等の高度な技術, 性能チューニング
 - CPU/GPU共有メモリ(GH200, MI300A等)
 - GPUレベル, MPI通信を含む高並列化
性能プロファイル/モニタリング方法

いかに簡潔かつ過
不足なしに教材を
準備していくか?
→ R7年度内に基
本部分の策定

GPUアプリケーションに関する国際連携

- CPU-centricシステムからGPU-centricシステムへの移行
 - 従来の国策としての計算科学アプリケーション開発はNFSである「富岳」及びNISの各スパコンを対象としていたが、特に「京」「富岳」の系譜はCPUのみ(=演算加速部なし)のシステムであり、GPU化が急速に進む欧米での利用が大きくは期待されなかった。
 - 逆にNFSがCPUシステムであることから、欧米の先進的GPUアプリケーションやフレームワークが国内で利用しづらかった。
- 現在NISはGPU化が着々と進み、ポスト富岳(Fugaku-NEXT)もGPUを全面採用することから、欧米と互換性のあるアプリケーション開発を行う機運が熟成しつつある。
- GPUアプリケーションの国際的相互乗り入れ
 - 欧米の先進的GPUアプリケーション／利用技術⇒国内での利用
 - 国産の先進的GPUアプリケーション／利用技術⇒海外での利用
- 計算科学の国際連携・共同研究の一層の活性化が期待される。
- アプリケーションのコンテナ化サポートをHAIRDESCとして行う。

GPUプログラミング教育・マシンリソース



加速部対応等に係る研究開発及び技術支援

• 特定アプリケーションコードのGPU化支援(R8年度～)

- 代表機関及び中核機関の合議により、非常に重要な中核的アプリケーションでGPU化が行われていない、あるいは不十分なものを選定し、予算とのバランスで年度単位でGPU化・大規模並列化を進める
 - 各機関の研究者だけでは十分な対応ができないため、外部委託業者を中心として行い、各機関に配置した担当者(複数、必要に応じて多機関に渡る)による技術支援を通してGPU化を行う
 - Slack等を用いたオンラインでのセミリアルタイム議論の他、一定期間毎の進捗や問題点共有のミーティングも行う
 - 結果と評価については年度内のシンポジウム等で一般ユーザとも共有
- アプリケーション選定(R8年度より実施の予定)
 - 例1) 文科省で実施すると想定されるアプリケーション開発プロジェクトや分野振興活動と連携
⇒ 一部のアプリケーションのGPU化協力
 - 例2) 各中核機関及びそのユーザが開発しているコードのGPU化を高度化するため予算と人材を集中させる
 - 例3) 本拠点で一般から公募し、有識者会議等を経て選定？ → おそらく難しい
- 公的機関における研究を対象とし、個別企業対応は行わないが、民間企業におけるGPUコンピューティングとAI for Science利用はプロモーションする

今後の予定

- 2025/12～
 - 富岳NEXTチームとの理研・HAIRDESC連携
- 2026/01末
 - web page, portal siteの開設
- 2026/03中旬
 - GPU教育マテリアル ver.1 公開
 - HPCI-IDを利用した教材配布・ビデオ配信
- 2026/03下旬
 - GPU教育向けテストアカウント発行:HPCIアカウントの利用
- **2026/03/30**
 - **第1回HAIRDESCシンポジウム(オンライン)**
- 2026/04以降の早い時期
 - GPUプログラミング・チュートリアルビデオと教材の公開
 - GPUプログラミング・ハッカソン開始
- 2026/??
 - 各種文科省プログラムとの連携開始？

まとめ

- RIST内に「次世代HPC・AI研究開発支援センター(HAIRDESC)」を設立
+ 筑波大・東大・東京科学大の3中核機関で事業を推進
- GPU対応プログラミング支援⇒ 大学・国研・企業を含め all Japan で支えていく
- 一部の重要な計算科学・AI研究のためのコードの集中的なGPU化支援
- GPU対応プログラミングを支援するGPUスパコンリソースの提供
- GPUアプリケーションの国際化の支援
- 先進的GPUシステム利用技術の研究開発
- 理化学研究所との協力関係の下でポスト富岳(富岳NEXT)に向けたGPUコード開発を推進
- ベンダー、海外機関との連携
- Home page公開予定 2026/01/26 <https://www.hairdesc.jp>
- **人材募集！**
 - HAIRDESC及び中核機関大学ではGPUプログラミング教育、GPUコーディング、先進的GPU研究を積極的に進める研究者を募集します。
 - 技術に強い人だけでなく、運営サポートができる方も歓迎します。