



北海道大学

150<sup>th</sup>  
HOKKAIDO UNIVERSITY



北海道大学  
情報基盤センター  
HOKKAIDO UNIVERSITY  
INFORMATION INITIATIVE CENTER

資料4-3-3

2026年、北海道大学は創基150周年を迎えます。

AI for Scienceに不可欠な計算資源の戦略的増強  
(ii) 既存の計算資源等による共用の効率化を図る取組

# 「AI4S対応GPUテストベッド」の整備

2026年 6月30日

**事業代表者：棟朝 雅晴**

国立大学法人 北海道大学 副理事（情報・DX）  
情報環境推進本部 副本部長・情報化推進室長  
情報基盤センター長・教授

※数量やスケジュール等は計画段階の内容であり、今後変更となる可能性があります。

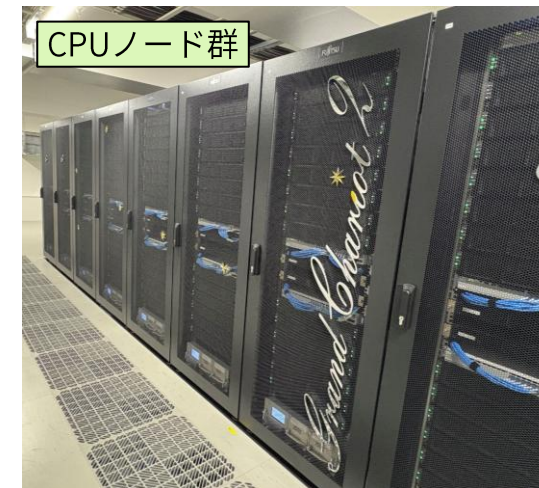
# 既存設備：学際大規模計算機システム

## ● 概要

- ✓ 2025年4月導入、7月正式サービス開始 コンテナ(Kubernetes)ベース
- ✓ スーパーコンピュータシステムと研究クラウドシステムが中核

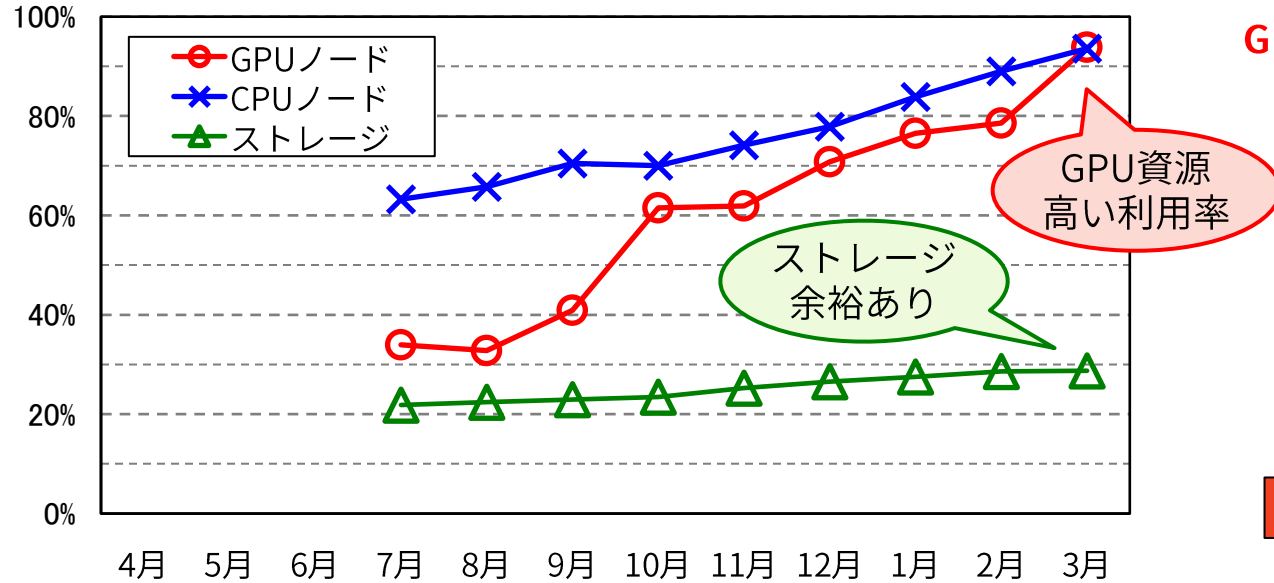
## ● スーパーコンピュータ「Grand Chariot 2」

- ✓ 演算資源：
  - CPUノード群 (480ノード)
  - **GPUノード群 (24ノード) ※北大スパコン初のGPUシステム**
- ✓ スパコンストレージ：
  - 17PBの**オールフラッシュ**構成 (国内大学スパコンで最大規模)
  - テープ装置 (@北見工業大学) への**遠隔バックアップ**機能付き
- ✓ 提供サービス：
  - **占有利用**コース (CPUソケット/GPUカード単位、年間定額)
  - **共用利用**コース



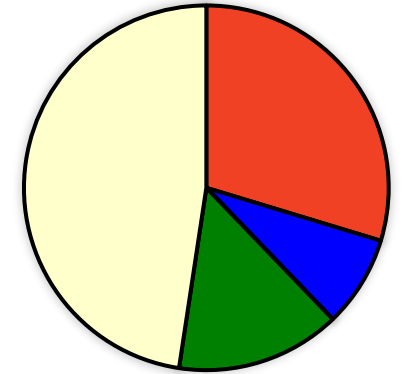
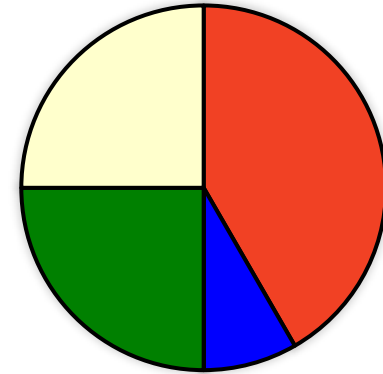
# Grand Chariot 2の運用状況

### 2025年度 システム使用率



### 2026年度 利用申請状況 (4/1 現在)

GPUノード (総量: 96GPU年)    CPUノード (総量: 960CPU年)



■ 占有 (一般)    ■ 占有 (民間)    ■ 共用 (HPCI等)    ■ 残り

※占有利用は事前受付分のみ、共用利用の受付はこれから。

- ✓ 演算資源 (GPUノード・CPUノード) は運用開始初年度 (2025年度) で高い利用率に到達
- ✓ 2026年度開始時点で既に多くの利用申請あり (GPUは残り25%、占有利用への高い需要)
- ✓ 現時点では高性能な (オールフラッシュの) ストレージシステムに十分な余裕あり

**計算資源 (GPU) の重点的強化**によりサービス (AI4S向け) の拡充を図る！

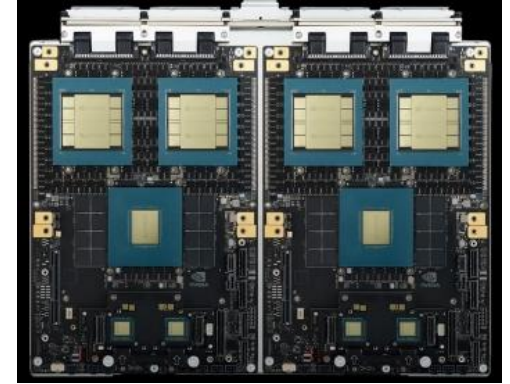
# 整備計画の概要

## ● 構成

- ✓ 全体 : **AI向けハイエンドGPU32枚** + 必要最小限の周辺機器
- ✓ ノード : GPU 4基、総メモリ1.5TB程度
- ✓ 総性能 : 約1.2 PFLOPS (FP64) 、 **約300PFLOPS (FP8)**

## ● ポイント

- ✓ ストレージ、ログインノード、管理ノード、NWスイッチ、冷却設備等は**既存設備を最大限利用**
- ✓ 既存システムで需要が高い**占有利用**を想定して**GPU数の最大化**を優先
  - 現有冷却設備の余剰能力と予算制約の双方を同時に満たす実行可能な最大構成
- ✓ 運用中のスパコンのジョブスケジューラ (**独自カスタマイズ済み**) との互換性
  - 異なる利用タイプ (占有利用・共用利用) の柔軟かつ効率的な運用が可能
- ✓ 現スパコンのGPUノード群のAI性能 (FP8 : 190FLOPS) の**1.5倍以上**の性能を実現



<https://www.nvidia.com/ja-jp/data-center/grace-cpu/>

# 提供予定サービスと狙い

## ● 占有利用型サービス

- ✓ 現スパコン（および前スパコン）で**十分な運用経験**あり
- ✓ **バッチジョブ** + （グループ毎の）**専用ジョブクラス**（ジョブキュー）で実現
- ✓ 利用負担金は**定額**
- ✓ カスタマイズ済みのジョブスケジューラによる**遊休資源の有効活用**も可能  
（短時間ジョブ用の共用利用や占有状態へ戻す機能などを組み込み済み）

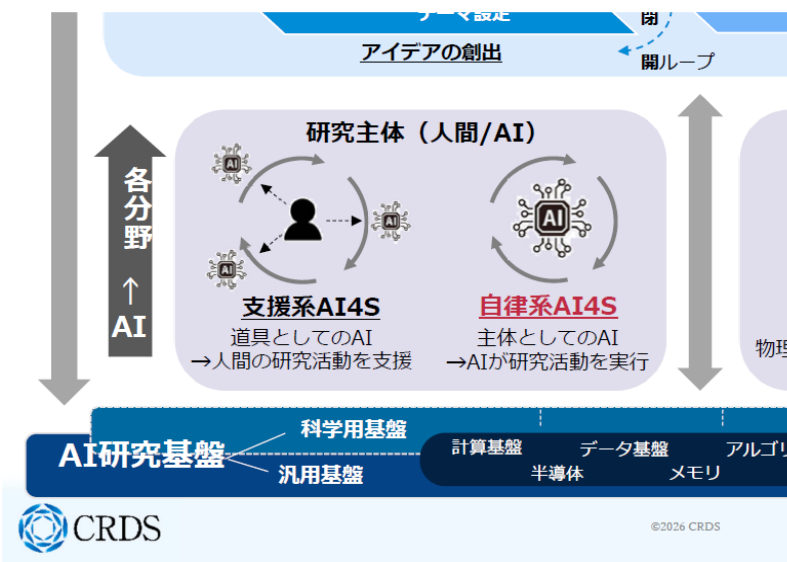
## ● 狙い

- ✓ 一般的な共用利用形態に合わない需要への対応
- ✓ 研究室のサーバ等からの**ステップアップが容易** → HPCIに馴染みのない研究者の参入を促進
- ✓ 定額制 → 利用者・運用者の双方で**研究費・運用費の予見性・計画性が向上**
- ✓ バッチジョブベース → 大規模実行を目的とした**他のHPCIシステムへのスムーズな移行**が可能

# 期待される効果・AI4S推進の展望

## ● 主たるターゲット：支援系AI4S・自律系AI4S

- ✓ 従来の一般的な**共用利用形態**では対応が難しいAI4S
  - ・ 研究活動と併走する形での**継続的**かつ**反復的**なAI利用
  - ・ 実験、観測、シミュレーション等と**連携**したワークロード
  - ※ **実行タイミングが保証**されない共用利用型とは相性が悪い
- ✓ 北大J-PEAKSでの事例：AI駆動による農林水産分野の研究DX（大規模画像解析、シミュレーションとの密な連携等）
- ✓ **知識生成基盤研究部門**（2025年度新設）による利用者支援



第1回 AI for Science推進委員会 参考資料5 (P19)

## ● HPCIへの貢献：新しい利用形態の実証

- ✓ 従来の共用利用型：単発の大規模シミュレーション・学習には最適 → **利用者を限定**
- ✓ 占有利用の利点「使いたいときに使える」は大きな魅力 → HPCIの**新規利用者獲得**への期待

**期待：支援系/自律系を含む多様なAI4S研究の支援 + HPCI利用者層の拡大**

# HPCIに関する取り組み

## ● HPCIへの資源提供計画

- ✓ 資源量：28~32GPU年程度（運用上必要な最小限の資源を除いた全て）
- ✓ 期間：3年（2027年度～2029年度） + 1～2年（現学際システムのリース期間延長に応じて）
- ※ 現システムでHPCIへ資源提供済みであり、**運用開始後速やかに資源提供を開始可能**

## ● 利用制度・利用環境面の取り組み

- ✓ HPCIにおける**占有利用形式の有効性の検証**と得られたノウハウ等の共有・フィードバック
- ✓ 共通のインターフェース（例：Open OnDemand）やGPU対応ソフトウェアを積極的に導入

## ● 他機関との連携

- ✓ 本事業採択機関（筑波大・東大・九大）との連協・協力（6/8付で**覚書締結**）
- ✓ NVIDIAのCPU-GPU密結合型システムの運用機関（JCAHPC、名大、理研R-CCS）との連携