

2025 年 2 月 20 日

プレスリリース

報道関係者各位

HPC システムズ株式会社

代表取締役 小野 鉄平

(コード番号：6597 東証グロース)

問合せ先 取締役管理部長 下川 健司

(電話番号：03-5446-5530)

HPC システムズ、NVIDIA 最新 Blackwell 搭載 GPU サーバーとソフトウェアパッケージで ローカル LLM 導入を加速するターンキーソリューションを提供開始 ～最大級モデルも動作、研究開発を強力にサポート～

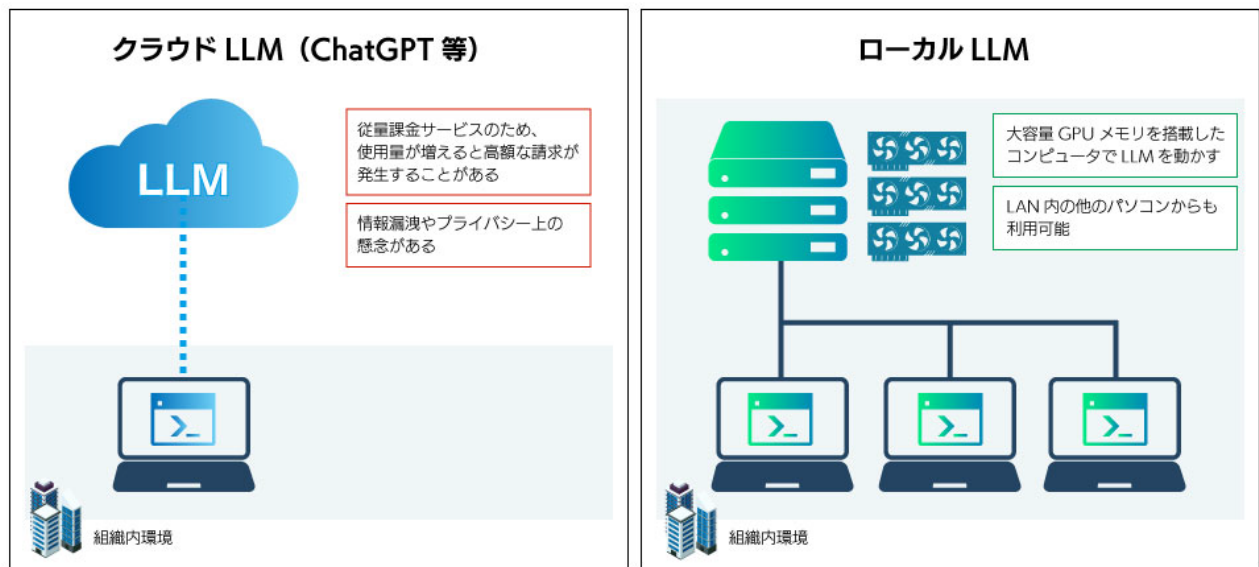
HPC システムズ株式会社（本社：東京都港区、代表取締役 小野 鉄平、以下 HPC システムズ）は、NVIDIA HGX B200（最新 Blackwell GPU アーキテクチャ）を 8 基搭載したハイエンド GPU サーバー「HPC5000-XGRAGPU8R10S-NVL」を中核とする、ローカル LLM（大規模言語モデル）環境の導入・運用をワンストップで実現するソリューションの提供を開始しました。

【背景：高まるローカル LLM 需要、大規模・高性能モデル、そしてセキュリティへの意識】

ChatGPT の登場以来、企業や研究機関において、大規模言語モデル（LLM）を自社環境（ローカル）で運用する「ローカル LLM」の需要が急増しています。

ローカル LLM は、ローカル環境で動作するため、クラウド利用に伴う従量課金が発生しません。さらに、クラウド型のチャット AI で懸念されている情報漏洩やプライバシー上の問題もなく、企業や研究機関が機密性の高い情報を扱う場合に最適です。

特に近年、より高度なタスクに対応できる、パラメータ数が数百億～数千億規模の「高性能 LLM」へのニーズが高まっています。これらの高性能 LLM は、より自然な文章生成、高度な質問応答、複雑な推論などが可能であり、研究開発、コンテンツ生成、顧客対応など、幅広い分野での活用が期待されています。しかし、このような高性能 LLM を動作させるには、大容量の GPU メモリを搭載した特殊なコンピュータが必要となり、導入・運用のハードルが高いという課題がありました。



【本ソリューションの概要】

1. 最新鋭のハードウェア:

- NVIDIA の最新 Blackwell アーキテクチャに基づく HGX B200 GPU を 8 基搭載した GPU サーバー「HPC5000-XGRAGPU8R10S-NVL」を提供。
- 180GB の GPU メモリ×8 基からなる合計 1.4TB 超の大容量メモリにより、現在最大級の LLM も動作可能。
 - 例：Meta が開発した Llama 3 の最大モデル（405B パラメータ）の推論に必要なメモリは約 810GB とされていますが、本製品はそれを大幅に上回る 1.4TB 超のメモリを搭載しており、余裕をもって対応可能。
 - より大規模なモデルや、複数のモデルを同時に動かすことも可能であり、将来的な LLM の進化にも対応可能。
- サーバー本体だけでなく、電源ユニット（PDU）やラックもセットで提供。

2. ソフトウェアも完備:

- 主要な LLM 関連ソフトウェア（Dify、OpenWebUI、Ollama、Xinference など）をインストール済み。
- オプションで、AI 開発を加速する「NVIDIA AI Enterprise（有償）」のセットアップにも対応。

3. 複数チームでの利用も簡単:

- 複数チームでのリソース共有やジョブ管理を効率化する「Slurm Workload Manager」の構築・運用をサポート。

4. 安心のサポート体制:

- 導入時の環境構築（ネットワーク設定、ストレージ設定など）から運用後の保守・技術サポートまで、HPC システムズがトータルでサポート。
- システム構成のコンサルティングや現地調査も実施。
- HPC システムズは、NVIDIA のパートナー認定制度「NPN」で最上位の「ELITE PARTNER」認定を取得しており、NVIDIA 製品を熟知した専門家がサポート。

【前世代からの進化：圧倒的な性能向上】

今回発表の製品は、前世代の HGX H200 と比較して、以下の点で飛躍的な性能向上を達成しています。

- GPU 性能: 2 倍以上 (FP8 精度で 72PFLOPS)
- メモリ容量: 約 1.3 倍 (1.44TB の HBM3e メモリ)
- メモリ帯域幅: 約 1.7 倍 (8TB/s)
- NV Switch GPU 間帯域: 2 倍 (1.8TB/s)

【本ソリューションのメリット】

- 導入期間の短縮: 環境構築の手間を省き、研究者はすぐに LLM 研究を開始可能。
- 専門知識不要: 複雑なシステム構築・運用は HPC システムズが代行。
- コスト削減: 外部委託に比べ、導入・運用コストを抑制。従量課金の心配も不要。
- 安定稼働: 専門家によるサポートで、システムを安定運用。
- 最新技術の活用: NVIDIA の最新 GPU で、最先端の研究開発を支援。
- 高性能 LLM に対応: 最大級の LLM を動作させ、より高度な研究開発を実現。
- セキュリティ強化: 情報漏洩リスクを排除し、機密情報を安全に保護。

【今後の展開】

HPC システム株式会社は、本ソリューションを皮切りに、LLM 分野の研究開発を支援する様々な製品・サービスを提供していく予定です。

【ハードウェア構成例（抜粋）】

コンポーネント	仕様
GPU	NVIDIA B200 SXM6 180GB ×8
GPU メモリ	合計 1.44TB
CPU	Intel Xeon 6-6960P (72core 2.7GHz 500W) ×2
システムメモリ	3TB (128GB DDR5-6400 ECC Registered ×24)
ストレージ	OS : 1.92TB U.2 NVMe ×2 内蔵ストレージ : 3.84TB U.2 NVMe ×2
電源	5250W 80PLUS TITANIUM 電源 ×6 (C19-C20 電源ケーブル 12 本同梱)



製品ページ : <https://www.hpc.co.jp/product/hardware/hpc5000-xgragpu8r10s-nvl/>

【HPC システムズについて】

HPC システムズは、ハイパフォーマンスコンピューティング（HPC）分野のニッチトップ企業です。

HPC 事業では、科学技術計算用高性能コンピュータとシミュレーションソフトウェア販売、科学技術計算やディープラーニング（深層学習）環境を構築するシステムインテグレーションサービス、シミュレーションソフトウェアプログラムの並列化・高速化サービス、計算化学ソフトウェア、マテリアルズ・インフォマティクスプログラム開発・販売、受託計算サービス・科学技術研究開発支援、創薬研究開発や素材・材料研究開発分野向けサイエンスクラウドサービスをワンストップで提供しています。また、CTO 事業では、顧客の用途、課題をヒアリングしながら、価格・性能・品質・高低温・防塵・防水・静電対策・過酷な環境に対する高耐久性など多種多様の対応が求められる、工場生産設備・製造装置・検査装置、制御機器や交通インフラ、自動運転、リテール店舗などのコントローラーとしての産業用コンピュータやエッジコンピュータの仕様提案から開発、生産、保守サポート、長期安定供給を実現しています。

社名 HPC システムズ株式会社

所在地 東京都港区海岸 3 丁目 9 番 15 号 LOOP-X 8 階

設立 2006 年 7 月 3 日

資本金 2 億 3,024 万円 (2024 年 12 月 31 日時点)

代表者 代表取締役 小野 鉄平

プレスリリースに関するお問い合わせ

https://www.hpc.co.jp/contact/company_form/