

報道関係者各位

2026 年 8 月 5 日
株式会社ジーデップ・アドバンス

ソフトバンク株式会社が提供するクラウドサービス 「AI データセンター GPU サーバー」の販売パートナー契約締結のお知らせ

～データ主権を確保するオンプレミス×クラウドのハイブリッド AI 基盤構築に向けて～

株式会社ジーデップ・アドバンス(本社：東京都中央区、以下「当社」)は、ソフトバンク株式会社(本社：東京都港区、以下「ソフトバンク」)が提供する AI 計算基盤を活用したクラウドサービス「**AI データセンター GPU サーバー**」の販売パートナーとして契約を締結したことをお知らせいたします。本コンピューティング プラットフォームは、ソフトバンクが国内に構築する大規模な NVIDIA GB200 NVL72 のクラスタを基盤とし、GPU 計算リソースを国内のデータセンターで運用して提供するものです。本販売パートナーとしての活動は、当社が中長期的に目指す「**データ主権を意識したオンプレミスとクラウドのハイブリッド AI 基盤**」の実現に向けた**重要な第一歩**であり、お客様が国内で安心して大規模 AI 開発・運用に着手できる環境づくりを進めてまいります。

背景：データ主権の重要性とハイブリッド戦略

生成 AI と大規模言語モデル(LLM)の急速な普及に伴い、AI 計算需要は世界的に爆発的な拡大を続けています。経済産業省の試算をもとにした分析によれば、**2030 年に必要とされる計算能力は 1,960 エクサ FLOPS に達し**、GPU リソースの確保は企業・研究機関にとって喫緊の経営課題となっています。

一方で、多くの企業や研究機関が AI 計算環境を海外クラウドに依存している現状は、情報セキュリティ・ガバナンス・コスト・各国法規制対応の観点から大きなリスクを抱えています。学習データ・推論データ・モデルそのものに含まれる機密情報や個人情報の取り扱いを考えれば、**国内に閉じた計算資源の確保＝データ主権の確立**は、もはや選択肢ではなく前提条件となりつつあります。

当社は、「**機密性の高いデータや定常ワークロードはオンプレミスで、大規模学習やバースト的な計算需要は国内クラウド GPU で**」という、データ主権を確保しながら柔軟性も両立するハイブリッド利用が、日本企業の AI 活用における有力な解の一つであると考えています。これまでオンプレミス側のソリューション提供を強みとしてまいりましたが、**今般の「AI データセンター GPU サーバー」の取り扱いを起点として、クラウド側の選択肢を本格的にラインアップに加え**、ハイブリッド構成の提案力を段階的に拡充してまいります。

ソフトバンクの「AI データセンター GPU サーバー」の概要

ソフトバンクが提供する「AI データセンター GPU サーバー」は、**NVIDIA DGX SuperPOD および NVIDIA GB200 NVL72 上で構築され**、AI ワークロード最適化アーキテクチャーで提供されるクラウドサービスです。

アクセラレーテッド コンピューティング サーバーは高速な [NVIDIA Quantum-2 InfiniBand](#) ネットワーキングで接続されたクラスター構成で提供され、大規模 AI 学習から科学シミュレーションまで幅広く対応します。

「AI データセンター GPU サーバー」に関する Web サイトのご案内

https://www.gdep.co.jp/ai_datacenter_gpuserver/

サービスラインアップ

- **NVIDIA GB200 β 版プラン(専有)** : NVIDIA GB200 NVL72×4、ノード間 [NVIDIA NVLink Switch](#)。1 サーバーからクラスター構成で利用可能。
- **NVIDIA DGX H100 プラン(専有)** : NVIDIA DGX H100 80GB×8(GPU メモリ総計 640GB)、ノード間 NVIDIA Quantum-2 InfiniBand 400Gbps×8。1 サーバーからクラスター構成で利用可能。
- **NVIDIA DGX A100 プラン(専有)** : NVIDIA DGX A100 80GB×8(GPU メモリ総計 640GB)、ノード間 NVIDIA Quantum InfiniBand 200Gbps×8。最短 7 日間からの短期集中利用にも対応。
- **NVIDIA DGX A100 時間貸しプラン** : GPU チップ 1 枚から 1 分単位で利用可能な従量課金制。モデルのファインチューニングや検証用途に最適。

3 つの特長

- **GPU サーバーを専有して利用可能** : 1 台から複数台まで広帯域クラスター構成で専有提供。Lustre ファイルシステム+NVMe による高速ストレージ(100GB 単位、1PB 超の大容量も対応)。
- **スポット利用で柔軟な契約が可能** : 時間貸しは最短 1 分から、専有プランは最短 7 日からと、計算ニーズに合わせて柔軟に選択可能。ソフトバンクの VPN サービス「SmartVPN」による閉域接続オプションも提供。
- **開発を支援するソフトウェア環境を標準付帯** : [NVIDIA AI Enterprise](#)(50 超の AI フレームワーク+事前学習済みモデル)、Slurm ジョブスケジューラー、NGC Private Container Registry をセットアップ済みで提供。契約後すぐに計算処理を開始可能。

正式お申し込みから最短 10 営業日でご利用開始可能であり、急な大規模計算需要にも俊敏に応えられる体制が整っています。

本契約における当社の強み

1. オンプレミスの実績を基盤としたハイブリッド提案の起点

当社は、[NVIDIA DGX](#) プラットフォームをはじめとするオンプレミスアクセラレーテッド コンピューティング インフ

ラの導入実績を多数有しております。本サービスの取扱開始により、これまで対応が難しかった大規模クラウド GPU 需要にも応えられるようになり、お客様のデータ機密度・ワークロード特性に応じたハイブリッド配置設計（「定常処理はオンプレ、ピーク学習は国内クラウド GPU」）の提案を一層充実させてまいります。

2. NVIDIA 認定 Elite パートナー AI Factory Specialization としての技術的優位性

当社は NVIDIA 認定エリートパートナーとして、長年にわたり GPU コンピューティング製品の販売・サポートを国内で展開してまいりました。また、昨年には日本初となる NVIDIA AI Factory Specialization の認定を取得しています。NVIDIA GB200 NVL72/ NVIDIA DGX H100 / NVIDIA DGX A100 といった最新 GPU アーキテクチャー、NVIDIA Quantum InfiniBand による高速インターコネクト、NVIDIA AI Enterprise、Slurm 等のソフトウェアスタック全般に関する深い技術知見を活かし、お客様のワークロードに最適なリソース構成・チューニングを提案・実装いたします。

3. 研究機関・大学・エンタープライズに広がる顧客基盤

当社は、国内の主要研究機関、大学、製造業、金融機関、製薬・素材産業等の AI 研究・開発部門に幅広い顧客基盤を有しています。LLM 開発、自動運転、AI 創薬、材料開発 DX といったデータの取り扱いに厳格な要件を持つお客様にこそ、国内クラウド×オンプレのハイブリッド構成の価値を訴求できる強みがあります。

4. PoC から本番運用までの伴走支援

GPUリソースの選定にとどまらず、環境構築、ライブラリ・フレームワーク(PyTorch、NVIDIA CUDA、NVIDIA NeMo、Dynamo-Triton、NVIDIA AI Enterprise 等)の最適化、Slurm ジョブスケジューラーの運用設計までを当社の技術スタッフが一通貫でサポートします。ハイブリッド構成における「どのワークロードをどこに配置するか」の検討も、お客様と並走してまいります。

5. スケーラビリティと投資効率の両立

時間貸しプランによる 1 分単位の従量利用から、専有プランによるクラスター長期利用まで、本サービスは幅広い計算ニーズに対応します。当社のオンプレミスソリューションと組み合わせることで、お客様は自社設備への過剰投資を避けながら、必要な時に必要な規模の GPU リソースを国内で確保可能となります。設備投資 (CapEx)と運用費(OpEx)のバランスを最適化した AI インフラ戦略を実現します。

想定ユースケース

本サービスと当社のソリューションを組み合わせることで、以下のような幅広い AI ワークロードに対応いたします。

- **LLM 開発**：基盤モデル・特化型モデルの学習およびファインチューニング
- **自動運転**：自動運転システムモデルの学習・検証
- **AI 創薬**：自社シミュレーションアルゴリズムでのデータ生成検証

- **材料開発 DX**：材料探索 AI システムの動作検証
- **科学技術計算**：大規模 HPC シミュレーション、研究機関向け計算基盤

今後の展開

当社は本販売パートナーとしての活動を**ハイブリッド AI 基盤戦略の起点**と位置付け、今後さらに、運用支援サービス、AI 開発支援サービス、追加のクラウド/オンプレ連携メニュー等を順次拡充してまいります。「データを守りながら、AI 活用を加速する」というメッセージのもと、国内エンタープライズ・研究機関の AI 基盤構築を中長期にわたり支援し、日本の AI 産業基盤とデータ主権の確保に貢献することを目指してまいります。

関係者のコメント

■ ソフトバンク株式会社 法人事業統括 AI・クラウドプラットフォーム本部 執行役員本部長 竹綱洋記氏のコメント

このたび、AI インフラ領域を牽引する株式会社ジーデップ・アドバンス様を販売パートナーとしてお迎えできることを大変嬉しく思います。株式会社ジーデップ・アドバンス様が持つ高度な GPU の技術知見と、当社の大規模な NVIDIA DGX SuperPOD 環境が融合することで、データ管理に厳格な要件を持つお客様へ、安心・安全なハイブリッド AI 環境を迅速に提供することが可能となります。今後は、両社の強みを生かし、日本のイノベーション創出を共に支援してまいります。

■ 株式会社ジーデップ・アドバンス 代表取締役 CEO 執行役員 飯野匡道のコメント

生成 AI の本格活用フェーズへの移行に伴い、計算需要が爆発的に拡大する一方で、機密データの国外流出リスクや海外クラウド依存は、日本企業にとって看過できない経営課題となっています。当社はこれまでオンプレミス GPU インフラの提供を通じて、多くのお客様の AI 研究開発を支えてまいりましたが、ソフトバンク様の『AI データセンター GPU サーバー』を販売パートナーとしてお取り扱いさせていただくことで、データ主権を確保した国内クラウド GPU という新たな選択肢をお客様にお届けできるようになります。本取り組みを、当社が中長期的に目指すオンプレミスとクラウドのハイブリッド AI 基盤戦略の起点と位置付け、日本の AI 産業基盤の発展に貢献してまいります。

■ ソフトバンクのクラウドサービス「AI データセンター GPU サーバー」について

詳細はソフトバンク公式サイトをご参照ください。

<https://www.softbank.jp/biz/services/ai/ai-data-center/>

■ 株式会社ジーデップ・アドバンスについて

当社は GPU コンピューティングおよび AI インフラに特化したソリューションプロバイダーです。国内の研究機関・教育機関・企業に対し、AI/HPC 環境の設計・構築・運用支援をワンストップで提供しています。

URL : <https://www.gdep.co.jp>

■ 本件に関するお問い合わせ先

株式会社ジーデップ・アドバンス ビジネス開発 BL

TEL : 03-6803-0620 E-mail : sales@gdep.co.jp

※ NVIDIA、NVIDIA DGX、DGX SuperPOD、NVIDIA AI Enterprise、Grace、Blackwell、NGC は、NVIDIA Corporation の商標または登録商標です。SoftBank およびソフトバンクの名称は、ソフトバンクグループ株式会社の登録商標または商標です。