

2024年6月11日

報道関係各位

GMO インターネットグループ株式会社

GMO インターネットグループ、生成 AI 向け GPU クラウドサービスに NVIDIA Spectrum-X を国内クラウド事業者として初採用 ～NVIDIA Spectrum-X 採用により、生成 AI モデルの開発期間を短縮へ～

GMO インターネットグループ株式会社（代表取締役グループ代表：熊谷 正寿 以下、GMO インターネットグループ）は、「NVIDIA H200 Tensor コア GPU(以下、NVIDIA H200 GPU)」を採用した2024年12月にリリース予定の生成 AI 向けの GPU クラウドサービスに、NVIDIA Spectrum™-X イーサネットネットワークプラットフォームを採用したことを発表いたします。^(※1)

NVIDIA Spectrum-X は AI ワークロード専用設計された世界初のイーサネットファブリック^(※2)であり、生成 AI ネットワークのパフォーマンスを飛躍的に向上させることが期待されます。大規模言語モデルの開発・研究者向けに GPU メモリ容量とメモリバス帯域幅を大幅に拡大・最適化された「NVIDIA H200 GPU」の国内最速提供に加え、Spectrum-X の採用により、お客様の大規模言語モデルの開発における期間短縮とコスト削減を実現します。NVIDIA Spectrum-X は、国内クラウド事業者として初の採用となります。^(※3)

(※1) NVIDIA、生成 AI 向けにイーサネット ネットワーキングを強化：<https://www.nvidia.com/ja-jp/about-nvidia/press-releases/2024/nvidia-supercharges-ethernet-networking-for-generative-ai/>

(※2) イーサネットファブリックとは、ネットワークデバイス間の高速かつ効率的なデータ転送を実現するために、スイッチ間の接続を最適化する技術

(※3) 2024年6月10日時点 NVIDIA 調べ



【背景】

データセンター規模で運用され、ユニークなネットワーク特性を示す生成 AI においては、従来のイーサネットネットワークは AI の高性能要求を満たすには不十分であり、NVIDIA はこれを解決するため、AI 向けに最適化されたイーサネットファブリックを開発しました。Spectrum-X は、生成 AI インフラのネットワーク性能を飛躍的に向上させ、高速かつ安定したネットワーク環境が不可欠であることから採用に至りました。

【Spectrum-X の特徴と利点】

Spectrum-X は AI 向けに構築された初のイーサネットファブリックで、従来のイーサネットファブリックと比べてネットワーク性能を 1.6 倍に向上させます。NVIDIA Spectrum SN5600 イーサネットスイッチと NVIDIA BlueField®-3 SuperNIC を特徴とする Spectrum-X は、生成 AI 向けの GPU クラウドサービス

に必要なパフォーマンスと機能を備えたエンドツーエンドの接続性で、適応ルーティングと混雑制御を活用し、AI に最適化された最高性能のイーサネットネットワークを提供します。

NVIDIA Spectrum SN5600 イーサネットスイッチは、AI 向けの世界初の毎秒 51Tb イーサネットスイッチであり、生成 AI 向けに、ネットワークの混雑を最小限に抑えながら、大規模で負荷の高いデータスループットを提供します。きめ細かなルーティングテクノロジーにより、ネットワークインフラを常に最大限に活用できます。

BlueField-3 SuperNIC は、ネットワーク集約型の超並列コンピューティング向けに設計されており、GPU サーバー間で最大毎秒 400Gb の RDMA over Converged Ethernet (RoCE) ネットワーク接続を提供し、クラスタ内のネットワークにおける 大規模言語モデルの開発パフォーマンスを向上させます。Spectrum-X を採用することにより、お客様の AI 開発期間の短縮とコスト削減を実現します。



▲NVIDIA Spectrum-X

【今後の展望】

GMO インターネットグループは、NVIDIA Spectrum-X の採用により、生成 AI 向けの GPU クラウドサービスの性能をさらに強化し、日本国内の AI 産業を支援します。これにより、AI 研究者やスタートアップは、より効率的かつ経済的に AI ソリューションの開発と展開を行うことが可能となります。GMO インターネットグループは、今後も最新技術を積極的に導入し、より高品質な生成 AI 向けの GPU クラウドサービスの提供に努めてまいります。そして、日本の AI 産業の発展、ひいては社会全体の発展に貢献してまいります。

【生成 AI 向け GPU クラウドサービスについて】

GMO インターネットグループの GPU クラウドサービスは、毎秒 4.8 テラバイト (TB/s) で 141 ギガバイト (GB) の HBM3e メモリを提供する初の GPU である「NVIDIA H200 GPU」を搭載予定です。これは、「NVIDIA H100 Tensor コア GPU」の 2 倍近い容量で、メモリ帯域幅は約 1.4 倍です。「NVIDIA H200 GPU」の大容量かつ高速なメモリは、生成 AI と LLM の開発効率を向上・開発コストを低減し、生成 AI モデルの開発力を底上げします。



▲NVIDIA H200 Tensor コア GPU

・参考：GMO インターネットグループ、NVIDIA H200 Tensor コア GPU を採用した生成 AI 向けの GPU クラウドサービスを国内最速提供へ～経済産業省による「クラウドプログラム」の供給確保計画の認定～
<https://www.gmo.jp/news/article/8933>

■提供開始時期：2024年12月(予定)

■本サービスに関するお問い合わせ先：

GMO インターネットグループ株式会社

ドメイン・ホスティング事業本部

E-mail：aicloud@gmo.jp

【GMO インターネットグループの AI 関連の取り組み】

(URL：<https://www.gmo.jp/ai-history/>)

GMO インターネットグループは、「AI で未来を創るナンバー 1 企業グループへ」を掲げ、① 時間とコストの節約、② 既存サービスの質向上、③ AI 産業への新サービス提供を軸に、グループ全パートナーを挙げて AI の活用、AI 関連サービスの提供を進めています。

①時間とコストの節約

2023 年度には賞金総額 1,000 万円の社内公募コンテスト「AI (愛) しあおうぜ！ ChatGPT 業務活用コンテスト」を実施しました。AI に関する取り組みや新サービスへつながる作品が集まり、多くが実際のサービスに実装されています。

また、全員受講必須の定期的な AI セミナーを開催。2024 年からは、非エンジニア向けのリスキリング施策として AI 人材育成プログラム「虎の穴」を実施しています。

(<https://www.gmo.jp/news/article/8802/>)

②既存サービスの質向上

生成 AI による文章や画像の生成など、既存サービスに AI を組み込むことにより、ドメイン、ホスティング、EC、広告、メディア、セキュリティなど幅広い領域でお客様にこれまで以上に利便性の高いサービスをご提供しています。

③AI 産業への新サービス提供

AI 産業をさらに拡大するための支援を進めています。

- 2023 年 5 月 24 日に、ハンズオン型 CVC「GMO Web3 株式会社」を、「GMO AI & Web3 株式会社」へと社名変更し AI スタートアップ支援を拡大しています。すでに、有望な AI スタートアップへの支援を実施しています。
- ドメイン登録サービス「お名前.com byGMO」にて、「.ai」ドメインの提供。
(<https://www.onamae.com/service/domain/ai/>)
- AI 専門家とともに「GMO 教えて AI 株式会社」を設立し、生成 AI プロンプトポータルサイト「教えて AI byGMO」を開始しました。(<https://oshiete.ai/>)
- 簡単に使える超高速 VPS『ConoHa VPS』にて、NVIDIA 社の GPU「NVIDIA H100 Tensor コア GPU」「NVIDIA L4 Tensor コア GPU」を搭載した AI 開発者向けの GPU ホスティングサービスを開始しました。(<https://www.conoha.jp/>)
- 2023 年 12 月に「NVIDIA Partner Network」に参画し、GPU の国内普及に取り組んでまいります。(<https://www.gmo.jp/news/article/8720/>)

- 2024年4月に経済安全保障推進法に基づく特定重要物資である「クラウドプログラム」の供給確保計画に関する経済産業省の認定を受けました。日本のAI産業の拡大のため、生成AIの基盤となるクラウドインフラの提供に向けて、約100億円規模のGPUサーバーへの投資をします。

(<https://www.gmo.jp/news/article/8933/>)

【GMO インターネットグループについて】

GMO インターネットグループは、ドメインからセキュリティ、決済までビジネスの基盤となるサービスをご提供するインターネットインフラ事業を軸に、インターネット広告・メディア事業、インターネット金融事業、暗号資産事業を展開する総合インターネットグループです。

お客様に喜ばれるサービスを迅速かつ低価格で提供するために、サービスは機器の選定から設置、構築、開発、運用までを内製化することを基本方針としています。そのため、グループ110社に在籍する約7,400名のパートナーのうち、ITのモノづくりを担う開発者（エンジニア・クリエイター）が約半数の49.5%を占めています。^(※4)

(※4) 2024年3月末時点

以上

【報道関係お問い合わせ先】

- GMO インターネットグループ株式会社
 本体事業管理本部 広報担当 川縁
 TEL : 03-5456-2555 E-mail : pr@gmo.jp
- GMO インターネットグループ株式会社
 グループコミュニケーション部 広報担当 新野
 TEL : 03-5456-2695 URL : <https://www.gmo.jp/contact/press-inquiries/>

【サービスに関するお問い合わせ先】

- GMO インターネットグループ株式会社
 ドメイン・ホスティング事業本部
 E-mail : aicloud@gmo.jp

【GMO インターネットグループ株式会社】 (URL : <https://www.gmo.jp/>)

会 社 名	GMO インターネットグループ株式会社 (東証プライム市場 証券コード:9449)	
所 在 地	東京都渋谷区桜丘町 26 番 1 号 セルリアンタワー	
代 表 者	代表取締役グループ代表 熊谷 正寿	
事 業 内 容	■ インターネットインフラ事業	■ インターネット広告・メディア事業
	■ インターネット金融事業	■ 暗号資産事業
資 本 金	50 億円	

Copyright (C) 2024 GMO Internet Group, Inc. All Rights Reserved.