



マイクロン、業界最速・最高容量のHBMを提供し、生成AIのイノベーションを促進

July 26, 2023 at 10:00 AM JST

先端1 β プロセスノードにより、1.2TB/秒の帯域幅、優れた電力効率、 業界初の8層での24GBを実現したHBM3 Gen2を提供

2023年7月26日 - アイダホ州ボイシ発 —Micron Technology, Inc. (NASDAQ:MU) は、全体の帯域幅で毎秒1.2テラバイト(1.2TB/s)、ピンあたりの帯域幅で毎秒9.2ギガビット(9.2Gbps)を超え、現在出荷中のHBM3ソリューションと比較して50%の高速化が図られただけでなく、業界初となる8層(8-High)で24ギガバイト(GB)の大容量化を実現したHBM3 Gen2メモリのサンプル出荷を開始したと発表しました。このHBM3 Gen2は、前世代製品と比較してワット当たりの性能が2.5倍向上し、人工知能(AI)向けデータセンターに求められるメモリ指標(性能、容量、電力効率)で新たな基準を達成します。これらの改良により、GPT-4やこれに続く大規模言語モデル(LLM)の学習時間を短縮できるだけでなく、AI推論向けにも高効率なインフラを構築でき、優れた総所有コスト(TCO)を実現できます。

マイクロンの広帯域メモリ(HBM)ソリューションの基盤は、業界をリードする1 β (1ベータ)DRAM プロセスノードです。このノードにより、業界標準のパッケージ・サイズ内でダイ当たり24ギガビット(Gb)のDRAMを8層キューブ(立体構造)で組み立てることができます。さらに、マイクロンでは12層積層の36GB製品を2024年第1四半期にサンプル出荷する予定です。マイクロンは、積層の高さが同一の既存の競合ソリューションと比較して、容量を50%以上増大させます。マイクロンのHBM3 Gen2で実現された電力性能比とピン速度の向上は、電力に対する要求が高い現在のAIデータセンターに不可欠です。この電力効率の改善は、競合のHBM3製品の倍となるシリコン貫通電極(Through-Silicon-Vias: TSV)の数、金属密度の5倍化による熱抵抗の改善、エネルギー効率の高いデータパス設計などのマイクロンの先進技術により実現しました。

メモリでの2.5D/3D 積層および先進パッケージング技術での実績を有するマイクロンは、TSMCの3DFabric Allianceのパートナーであり、半導体とシステム・イノベーションの未来の形成に貢献できることを誇りに思っています。HBM3 Gen2製品開発の一環としてマイクロンとTSMCが進めている協業は、AIおよび高性能コンピューティング(HPC)用途のコンピューティング・システムの円滑な導入と統合の基盤になります。TSMCはすでに、マイクロンのHBM3 Gen2メモリのサンプルを入手しており、顧客のイノベーションを支援する次世代HPCアプリケーション用途としてその評価とテストをマイクロンと緊密に連携しながら進めています。

マイクロンのHBM3 Gen2メモリは、数兆のパラメータに達するマルチモーダルAIモデルなどの生成AIの需要に対応します。単一キューブで24GBに達する大容量化とピン速度9.2Gbps超の高速化により、LLMの学習時間を30%以上短縮し、TCOを低減します。加えて、今回のメモリ製品は、一日当たりのクエリ実行数を大幅に増大させ、学習済みモデルを効率的に活用できるようにします。さらに、クラス最良のワット当たり性能により、最新のAIデータセンターでのコストを目に見える形で削減します。1,000万個のGPUが利用される例として、HBMキューブ当たり5ワットの節電が実現した場合、運用コストは5年間で最大5億5,000万ドル削減できると見積もられます。

マイクロンのコンピュータ事業部門 バイスプレジデント 兼 ゼネラル・マネージャーのプラヴィーン・ヴァイディヤナサン

は「マイクロンのHBM3 Gen2技術は、顧客と業界に向けて優れたAI/HPCのコンピューティング・ソリューションの実現を目指し開発されました。HBM3 Gen2製品を顧客のプラットフォームに容易に統合できる点もマイクロンにとって重要な基準の一つです。フルスペックのピン速度で動作する完全プログラマブルなメモリBIST (Memory Built-In Self-Test)により、顧客と実施できるテスト能力も向上し、より効率的なコラボレーション機会も創出され、顧客は製品化に要する時間を短縮できるようになります」と述べました。

NVIDIA ハイパースケールおよびHPCコンピューティング担当バイスプレジデントのイアン・バック (Ian Buck) 氏は「生成AIの中核は、エネルギー効率に優れたHBMの広帯域幅を活用したアクセラレーテッド・コンピューティングです。NVIDIAは、マイクロンと広範な製品で協業してきた長い歴史があり、HBM3 Gen2でマイクロンと協力を通じて、AIのイノベーションを加速させることに期待しています」と述べました。

マイクロンは、米国での設計とプロセス開発、日本でのメモリ製造、台湾での先進パッケージングによる世界規模のエンジニアリング組織力を生かし、今回の画期的な製品を開発しました。本日の発表は、マイクロンの技術リーダーシップを示す重要な成果になります。マイクロンではすでに[1α\(1アルファ\)ノードによる24GbのモノリシックダイのDRAMを実装した大容量サーバー・ソリューション向けの96GB DDR5モジュール](#)を発表しており、本日、1βノードによるダイ当たり24Gbの24GB HBM3製品を発表しました)。さらに、1βノードによるモノリシックダイ当たり32GbのDRAMを実装した128GB DDR5モジュールの提供を2024年上半期に予定しています。これらの製品を通じて、マイクロンはAIサーバー向けの先進技術イノベーションを実証します。

マイクロンのHBM3は、下記をご参照ください。

- HBM3のサイト(英文) : <https://www.micron.com/hbm3>
- Six Five Insider Podcast Interview with Girish Cherussery(英文) :

<https://www.youtube.com/watch?v= I1vEmlnp8I>

Micron Technology, Inc.について

マイクロンは、情報活用のあり方を変革し、すべての人々の生活を豊かにするために、革新的なメモリおよびストレージソリューションを提供するリーディングカンパニーです。顧客第一主義を貫き、テクノロジーの最前線でリーダーシップを発揮し続け、洗練された製造技術と事業運営を妥協なく追求するマイクロンの製品ポートフォリオは、DRAM、NAND、NORの各種メモリからストレージ製品まで多岐にわたり、Micron®またはCrucial®のブランドを冠した高性能な製品を多数展開しています。マイクロンで生まれた数々のイノベーションは、データの活用を加速すると同時に、人工知能や5Gといった最先端分野の進歩の基盤として、データセンターからインテリジェントエッジ、さらにはクライアントコンピューターとモバイルをまたいだユーザーエクスペリエンスまで、さまざまな事業機会を新たに生み出し続けています。Micron Technology, Inc. (Nasdaq: MU)に関する詳細は、[micron.com](https://www.micron.com)をご覧ください。

©2023 Micron Technology, Inc. All rights reserved. 情報、製品および仕様は予告なく変更される場合があります。マイクロン、マイクロンのロゴ、およびその他のすべてのマイクロンの商標はMicron Technology, Inc.に帰属します。他のすべての商標はそれぞれの所有者に帰属します。