



マイクロン、データセンター・インフラストラクチャ向け大容量256GB LPDRAM SOCAMM2で世界初の新たな基準を確立

ニュースのハイライト:

- 業界初のモノリシック32Gb LPDDR5Xダイにより、標準的なRDIMMと比較して消費電力と実装面積を1/3に削減
- 長文コンテキストのLLM推論でTime-to-First-Token (TTFT: 最初のトークン生成までの時間)を2.3倍高速化、スタンドアロンCPUアプリケーションではワットあたりの性能を3倍向上
- モジュールあたりの容量を1.33倍に拡張。AIと高性能コンピューティング(HPC)の両方に対応する8チャンネルサーバーCPUあたり2TB LPDRAMを実現

2026年3月3日、米国アイダホ州ボイシ — Micron Technology, Inc. (Nasdaq:MU)は本日、業界最高容量のLPDRAMモジュールとなる256GB SOCAMM2の顧客向けサンプル出荷を開始し、低消費電力サーバーメモリにおけるリーダーシップを強化しました。業界初のモノリシック32Gb LPDDR5X ダイの採用により実現したこの画期的な成果は、AIデータセンターの変革を進展させ、新しいシステムアーキテクチャーの可能性を拓く大容量の低消費電力メモリを提供します。

AIの学習、推論、エージェントAI、そして汎用コンピューティングの融合により、メモリにはより厳しい要件が求められ、データセンターのシステムアーキテクチャーの再構築を促しています。最新のAIワークロードは、大規模なモデルパラメータ、拡張し続けるコンテキストウィンドウ、永続的なキー・バリュー(KV)キャッシュを必要とし、コアコンピューティングでは、データの集約度、並行処理、メモリの実装面積の向上・改善が重要とされています。

これらのワークロードのすべてで、メモリの容量、帯域幅効率、レイテンシ、電力効率がシステムレベルでの主要な制約事項とされ、性能、拡張性、総所有コスト(TCO)に直接的な影響を及ぼします。LPDRAMは、これらの特性を独自に組み合わせ、消費電力と熱の制約が厳しいデータセンター環境において、AIサーバーとコアコンピューティング・サーバーの双方の基盤となるソリューションです。マイクロンは、高度なAIインフラのニーズに対応する先進的なメモリの共同設計においてNVIDIAと協業しています。

マイクロンのクラウドメモリビジネスユニットのシニアバイスプレジデント兼ゼネラルマネージャー、ラジ・ナラシンハン氏は「マイクロンの256GB SOCAMM2製品は、AIとHPCの両方に最適な電力効率に優れたCPU接続のメモリソリューションです。本日の発表は、マイクロンのテクノロジーとパッケージングの進歩の証であり、業界最小の実装面積で大容量、最小電力のモジュール式メモリソリューションを実現しました。マイクロンは、データセンター向け低消費電力メモリソリューションにおける継続的なリーダーシップのもと、32GbモノリシックLPDRAMダイを初めて提供し、独自の地位を確立しました。マイクロンは、より電力効率に優れ、大容量のシステムアーキテクチャーの採用を促進します」



容量、電力効率、ワークロード性能の最適化

マイクロンの256GB SOCAMM2は、さまざまなAIや汎用コンピューティング・ワークロードに対し、より大きいメモリ容量、大幅に低い消費電力、より高速な性能を提供します。

- **AIサーバーのメモリ容量の拡張：**
256GB SOCAMM2は、従来の最大容量である192GB SOCAMM2から容量を3分の1以上向上させ、8チャンネルCPUあたり2TBのLPDRAM搭載を可能にします。これにより、大きなコンテキストウィンドウや複雑な推論ワークロードに対応します。
- **低消費電力と省スペース化：**
SOCAMM2は、同等のRDIMMと比較して消費電力が1/3、実装面積も1/3に削減されるため、ラック密度が向上し、TCOを削減します*1。
- **推論とコアコンピューティング性能の向上：**
統合メモリアーキテクチャーにおいて、KVキャッシュオフロードに256GB SOCAMM2を使用した場合、現在利用可能なソリューションと比較して*2、長文コンテキストのリアルタイムLLM推論におけるTTFTの時間を2.3倍以上高速化します。またスタンドアロンCPUアプリケーションでは、LPDRAMは、HPCワークロード向けの標準的なメモリモジュールと比較して、ワットあたりの性能を3倍以上向上させます*3。
- **保守性と拡張性を実現するモジュール設計：**
モジュール式のSOCAMM2の設計により、保守性が向上し、液冷サーバーアーキテクチャーにも対応し、AIとコアコンピューティングのメモリ要件が厳しくなる中、将来的な容量拡張を可能にします。

NVIDIA データセンターCPU部門 製品統括、イアン・ファインダー氏は「高度なAIインフラは、負荷の高いAI推論ワークロードにおいて最大限のパフォーマンスと効率性を実現するため、あらゆるレイヤーで卓越した最適化を必要とします。マイクロンの256GB SOCAMM2は、従来のサーバーメモリと比べ消費電力を抑えながら、大容量かつ高帯域幅のメモリを実現しており、次世代AI CPUの実現を支えています」と述べています。

業界標準を推進し、低消費電力メモリの採用を加速

マイクロンは、JEDEC SOCAMM2仕様の策定を主導し続けるとともに、システム設計者との緊密な技術協業を維持し、次世代データセンタープラットフォームでの業界全体の電力効率と性能の向上を推進しています。

マイクロンは現在、256GB SOCAMM2の顧客向けサンプルを出荷しています。8GB～64GBコンポーネントから、48GB～256GB SOCAMM2モジュールまで、業界で最も幅広いデータセンターLPDRAMポートフォリオを提供しています。

その他のリソース

- [LPDDR at Scale: 大容量メモリによる効率的なLLM推論の実現](#) (英文)
- [すべてのワットが重要: データセンターを変革する低消費電力メモリの役割](#) (英文)
- [SOCAMM2 ウェブページ](#)
- [データセンターメモリウェブページ](#)



*1: 消費電力の3分の1は、128GB、128ビットのバス幅のSOCAMM2モジュール1つが使用する電力ワット数に基づいて計算され、64GB、64ビットのバス幅のDDR5 RDIMM2つと比較されます。フットプリントの3分の1の計算では、SOCAMM2面積(14x90mm)と標準的なサーバーRDIMMを比較しています。

*2: 結果は、500Kコンテキスト長と16人の同時ユーザーを使用したLlama3 70Bモデル(FP16量子化)によるリアルタイムの推論に関するマイクロンの内部テストに基づいています。70B TTFTレイテンシの改善は、CPUあたり2TB LPDRAMで0.12秒、CPUあたり1.5TB LPDRAMで0.28秒のレイテンシに基づいています CPU。テスト条件の詳細については、今月初めに発表されたホワイトペーパー([LPDDR at Scale: 大容量メモリによる効率的なLLM推論の実現](#)) (英文)をご覧ください。

*3: LPDDR5XとDDR5の同一容量でPot3DソーラーフィジックスHPCコードパフォーマンスを測定するマイクロンの内部テスト。

Micron Technology, Inc.について

マイクロンは、情報のあり方を変革し、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及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、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(AI)や計算集約型アプリケーション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をご覧ください。

© 2025 Micron Technology, Inc. All rights reserved. 情報、製品、仕様は予告なく変更されることがあります。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権利者に帰属します。