

# Cisco UCS C880A M8 ラック サーバー

---

# 目次

バリュー ステートメント .....	3
製品の概要 .....	3
優れた機能 .....	4
プラットフォームのサポート .....	4
特長と利点 .....	5
製品仕様 .....	6
発注情報 .....	7
保証情報 .....	8
シスコ サポート .....	8
製品持続可能性 .....	8
シスコおよびパートナーの提供サービス .....	8
Cisco Capital .....	9
詳細はこちら .....	9

## バリュー ステートメント

Cisco UCS C880A M8 は、次世代の NVIDIA HGX B300 NVL8 GPU により、すべてのデータセンターで高度な AI およびハイパフォーマンス コンピューティング (HPC) ワークロードを高速化します。

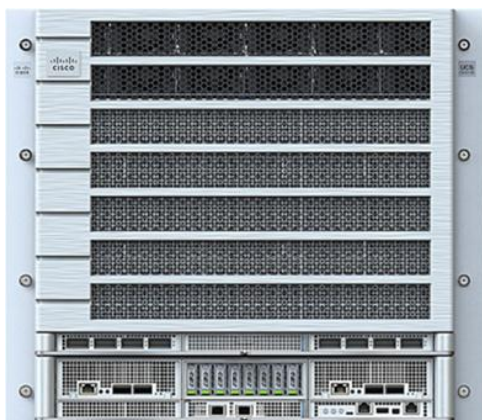


図 1.  
UCS C880A M8 前面

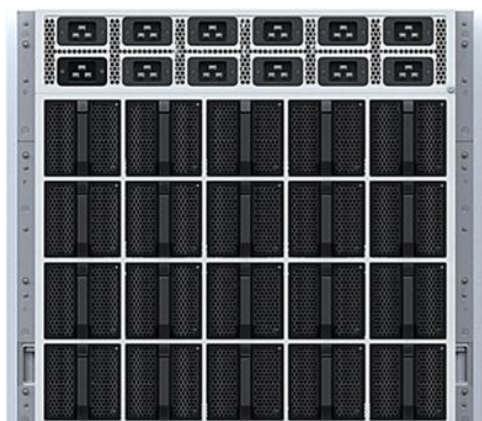


図 2.  
UCS C880A M8 背面

## 製品の概要

Cisco UCS C880A M8 は、NVIDIA HGX プラットフォームに基づいて、最も要求の厳しい人工知能 (AI) および高性能コンピューティング (HPC) ワークロードに電力を供給するように設計された、高密度の空冷ラックサーバーです。NVIDIA HGX プラットフォームで 8 つの NVIDIA HGX B300 (SXM) GPU と統合し、2 つの Intel® Xeon® 第 6 世代プロセッサを搭載しているため、リアルタイム ラージ言語モデル (LLM) の推論、次のレベルのトレーニング パフォーマンス、および大量のデータ処理をサポートします。C880A M8 は、大規模なモデルトレーニングと微調整から、リアルタイムの推論と大量のデータ処理まで、AI スタック全体のお客様をサポートします。シスコの AI 戦略にシームレスに統合し、堅牢なコンピューティング インフラストラクチャを提供することで AI の現状を接続および保護します。このサーバーは、Cisco UCS® の高密度 AI サーバー ポートフォリオを拡張し、サービスプロバイダー、金融、製造、医療、生命科学、自動車業界など、さまざまな業界の企業に強力なソリューションを提供します。高度なアーキテクチャにより、C880A M8 は比類のないパフォーマンス、拡張性、および企業管理性を保証し、大規模な AI モデルのトレーニング、微調整、推論などのコンピューティング集約型の AI のユースケースに最適です。

## 優れた機能

### NVIDIA HGX B300 で AI の可能性を最大限に引き出す

Cisco UCS C880A M8 ラック サーバーは、最新の NVIDIA HGX プラットフォームに 8 つの NVIDIA B300 (SXM) GPU を統合する点で際立っています。この強力な GPU 構成は、大規模な AI モデル トレーニング、微調整、リアルタイムの推論など、最も要求の厳しい AI ワークロードに対して次レベルのパフォーマンスを提供する機能の中核です。B300 GPU は、複雑なディープ ラーニング モデルや大規模言語モデルを加速するために不可欠な、強力な並列処理機能と高速な GPU インターコネクトを提供します。この統合により、企業はより高いトークン スループットを実現し、AI 運用の経済性を向上させ、LLM およびエージェント ワークロードの収益性を高めるスケーリングを可能にできます。

### 包括的な企業 AI の管理性

Cisco UCS C880A M8 ラック サーバーは、企業での使用に適した設計になっています。将来のリリースでは、C880A M8 で Cisco Intersight を介した管理ができるようになります。

Cisco Intersight は、サーバーのライフサイクル管理を簡素化するクラウドベースの管理プラットフォームを提供し、電源操作、広範なモニタリング メトリック、サーバー構成管理、ファームウェアバンドルリリース管理などの機能を提供します。この集中管理とオブザーバビリティにより、AI インフラストラクチャの運用を合理化し、複雑さを軽減し、データセンター全体で一貫したポリシーの適用を保証します。

### AI および HPC ワークロード向けの設計

元々のパワーを超えて、Cisco UCS C880A M8 ラック サーバーは AI と HPC の固有の要求を満たすように特別に設計されています。その設計は、リアルタイムの大規模言語モデル推論をサポートし、AI 主導のアプリケーションの迅速なデプロイメントと応答性を可能にします。また、次のレベルのトレーニング パフォーマンスにも優れ、複雑な AI モデルのトレーニングに必要な時間を大幅に削減します。さらに、大容量データ処理に対応したそのキャパシティにより、GPU アクセラレーションによる ETL プロセスを含むデータサイエンスおよびビッグデータ分析に理想的なプラットフォームになります。この特化した設計により、組織は AI モデルを効率的に構築、最適化、および利用でき、スケーラブルで高性能なソリューションでビジネスの成長を促進できます。

## プラットフォームのサポート

Cisco UCS C880A M8 は、AI および HPC ワークロードをホストし、高速化するように設計された専用のラック サーバー プラットフォームです。AI/HPC の導入用にデータセンター環境で一般的に使用されるさまざまなオペレーティング システムと仮想化プラットフォームをサポートしています。特定のソフトウェアスタックの互換性には、AI アプリケーションのデプロイメントと最適化のための NVIDIA AI Enterprise と NVIDIA の NIM (NVIDIA Inference Microservices) が含まれます。

## 特長と利点

表 1. Cisco UCS C880A M8 ラック サーバーの特徴と利点の概要

機能	利点
8つの NVIDIA B300 (SXM) GPU を搭載した NVIDIA HGX	NVIDIA のフラグメントスーパーコンピューティング GPU を活用して、AI モデル トレーニング、微調整、および推論を加速するために不可欠な、比類のない処理能力を提供
インテル Xeon第 6世代プロセッサ X 2	GPU アクセラレーションを補完するように最適化された高周波数、高スループット CPU。ボトルネックなしでトレーニングおよび推論パイプラインにデータをフィードするのに最適
ボードに統合された NVIDIA ConnectX-8 NIC (E/W トラフィック)	組み込みの East/West 800G ネットワーキング ファブリックにより、スケーラブルな AI トレーニングのための超低遅延、高帯域幅 GPU 間、およびサーバー間通信を提供
最大 8x E1.S NVMe SSD	高性能のローカル NVMe ストレージにより、超低遅延のデータ キャッシングと AI モデル トレーニングの高速チェックポイントが提供されます。
ホットスワップ可能な冗長電源構成	ダウンタイムを最小限に抑え、可用性を最大化するように設計された、容易に保守可能な冗長電源によって、エンタープライズクラスのレジリエンスを実現。
リアルの大規模言語モデル (LLM) の推論	LLM の迅速かつ効率的なデプロイメントを可能にし、即時応答を必要とするリアルタイムのアプリケーションおよびサービスをサポートします。
新たなレベルのトレーニングのパフォーマンス	大規模で複雑な AI モデルのトレーニングに必要な時間を大幅に削減し、反復と開発サイクルを迅速化
大量のデータ処理	大規模なデータセットを処理するように設計され、抽出、変換、書き込み (ETL) プロセスの高速化と GPU が加速するビッグデータ分析を促進
将来のリリース : Cisco Intersight® による企業AIの管理性	Cisco Intersight は、今後のリリースでサーバーの集中型クラウドベースの管理を提供。これにより、電源操作、モニタリング、サーバー構成管理、ファームウェアの更新などの機能が提供され、運用タスクを簡素化し、一貫した制御を確保
AI 向けの検証済みソリューション	コンピューティング、ネットワーク、ストレージ、ソフトウェアを含む検証済みの AI ソリューションを提供するシスコの戦略の一部であり、AI 工場の信頼性の高い高い最適化されたパフォーマンスを保証
拡張性に優れたインフラストラクチャ	高密度 GPU プラットフォーム向けに最適化され、AI 工場全体で予測可能なパフォーマンスを提供し、Cisco AI POD 内の回数変更可能拡張を可能にして、AI の需要の増大に対応

## 製品仕様

表 2. Cisco UCS C880A M8 ラックサーバーの主な仕様 (NVIDIA HGX B300 GPU ベースの構成)

コンポーネント	仕様
フォーム ファクタ	10RU 19 インチ ラック サーバー (NVIDIA の HGX リファレンスアーキテクチャに基づく)
プロセッサ	2x Intel Xeon 第 6 世代 6776P または 2x Intel Xeon 第 6 世代 6767P
メモリ	32x 64GB DDR5 RDIMM または 32x 96GB DDR5 RDIMM または 32x 128GB DDR5 RDIMM
GPU	8x Nvidia HGX B300 NVL8
ブート ドライブ (Boot Drive)	RAID コントローラを装備した 2x 960GB M.2 SATA SSD
内蔵ストレージ	最大 8x PCIe Gen5 x4 E1.S NVMe SSD
E-W ネットワーキング	8x GPU ボード統合 ConnectX-8
N-S ネットワーキング	2x CX-7 2x200G (暗号化) または 2x B3220 2x200G (暗号化) または 2x B3240 2x400G (暗号化) 1x OCP TFF Gen5 x8
電源モジュール	12x 80PLUS 54V 3.2kW MCRPS ホットスワップ可能な冗長 PSU (N+N)
管理	Cisco BMC
ハードウェアおよびソフトウェアの相互運用性	サポートされるオペレーティング システムと周辺機器のオプションの一覧については、『 <a href="#">シスコハードウェアおよびソフトウェアの相互運用性リスト</a> 』を参照してください。

## 発注情報

表 3. 発注情報

製品番号	製品の説明
<b>UCSC-880A-M8-B301</b>	2x Intel Xeon 6776P 2.3 GHz（最大ターボ 3.9 GHz）CPU、8x NVIDIA HGX B300 SXM GPU、32x 96GB（最大 4000 MT/s DIMM）、2x 960GB M.2 SATA ブートドライブ、2x E1.S 3.84TB NVMe SSDデータドライブ、East/West N/W 用に 8x NVIDIA ConnectX-8（GPU ボード統合型）、North/South N/W 用に 2x NVIDIA ConnectX-7（2x200G）暗号化対応、1x Intel X710-T2L OCP
<b>UCSC-880A-M8-B302</b>	2x Intel Xeon 6776P 2.3 GHz（最大ターボ 3.9 GHz）CPU、8x NVIDIA HGX B300 SXM GPU、32x 96GB（最大 4000 MT/s DIMM）、2x 960GB M.2 SATA ブートドライブ、2x E1.S 3.84TB NVMe SSDデータドライブ、East/West N/W 用に 8x NVIDIA ConnectX-8（GPU ボード統合型）、North/South N/W 用に 2x NVIDIA B3220（2x200G）暗号化対応、1x Intel X710-T2L OCP
<b>UCSC-880A-M8-B303</b>	2x Intel Xeon 6776P 2.3 GHz（最大ターボ 3.9 GHz）CPU、8x NVIDIA HGX B300 SXM GPU、32x 128GB（最大 4000 MT/s DIMM）、2x 960GB M.2 SATA ブートドライブ、2x E1.S 3.84TB NVMe SSDデータドライブ、East/West N/W 用に 8x NVIDIA ConnectX-8（GPU ボード統合型）、North/South N/W 用に 2x NVIDIA ConnectX-7（2x200G）暗号化対応、1x Intel X710-T2L OCP
<b>UCSC-880A-M8-B304</b>	2x Intel Xeon 6776P 2.3 GHz（最大ターボ 3.9 GHz）CPU、8x NVIDIA HGX B300 SXM GPU、32x 128GB（最大 4000 MT/s DIMM）、2x 960GB M.2 SATA ブートドライブ、2x E1.S 3.84TB NVMe SSDデータドライブ、East/West N/W 用に 8x NVIDIA ConnectX-8（GPU ボード統合型）、North/South N/W 用に 2x NVIDIA B3220（2x200G）暗号化対応、1x Intel X710-T2L OCP
<b>UCSC-880A-M8-B305</b>	2x Intel Xeon 6776P 2.3 GHz（最大ターボ 3.9 GHz）CPU、8x NVIDIA HGX B300 SXM GPU、32x 96GB（最大 4000 MT/s DIMM）、2x 960GB M.2 SATA ブートドライブ、2x E1.S 3.84TB NVMe SSDデータドライブ、East/West N/W 用に 8x NVIDIA ConnectX-8（GPU ボード統合型）、North/South N/W 用に 2x NVIDIA B3240（2x400G）暗号化対応、1x Intel X710-T2L OCP
<b>UCSC-880A-M8-B306</b>	2x Intel Xeon 6776P 2.3 GHz（最大ターボ 3.9 GHz）CPU、8x NVIDIA HGX B300 SXM GPU、32x 128GB（最大 4000 MT/s DIMM）、2x 960GB M.2 SATA ブートドライブ、2x E1.S 3.84TB NVMe SSDデータドライブ、East/West N/W 用に 8x NVIDIA ConnectX-8（GPU ボード統合型）、North/South N/W 用に 2x NVIDIA B3240（2x400G）暗号化対応、1x Intel X710-T2L OCP
<b>UCSC-880A-M8-B307</b>	2x Intel Xeon 6776P 2.3 GHz（最大ターボ 3.9 GHz）CPU、8x NVIDIA HGX B300 SXM GPU、32x 64GB（最大 4000 MT/s DIMM）、2x 960GB M.2 SATA ブートドライブ、2x E1.S 3.84TB NVMe SSDデータドライブ、East/West N/W 用に 8x NVIDIA ConnectX-8（GPU ボード統合型）、North/South N/W 用に 2x NVIDIA ConnectX-7（2x200G）暗号化対応、1x Intel X710-T2L OCP
<b>UCSC-880A-M8-B308</b>	2x Intel Xeon 6767P 2.3 GHz（最大ターボ 3.9 GHz）CPU、8x NVIDIA HGX B300 SXM GPU、32x 96GB（最大 4000 MT/s DIMM）、2x 960GB M.2 SATA ブートドライブ、2x E1.S 3.84TB NVMe SSDデータドライブ、East/West N/W 用に 8x NVIDIA ConnectX-8（GPU ボード統合型）、North/South N/W 用に 2x NVIDIA ConnectX-7（2x200G）暗号化対応、1x Intel X710-T2L OCP

## 保証情報

Cisco UCS C880A M8 ラック サーバには、3年間の翌営業日（NBD）ハードウェア保証と90日間のソフトウェア保証が付いています。

## シスコ サポート

Cisco UCS の保証を強化するのが Cisco Success Tracks です。Success Tracks は、デジタル インテリジェンスとヒューマン インテリジェンスのメリットをサポート エクスペリエンスに付加します。ST の成果物の詳細については、[こちら](#)の説明を参照してください。

## 製品持続可能性

Cisco の環境、社会、ガバナンス（ESG）イニシアチブおよびパフォーマンスに関する情報は、Cisco の CSR および持続可能性 [レポート](#) で提供されます。

表 4. シスコの環境保全に関する情報

持続性に関するトピック		参照先
全般	製品の材料に関する法律および規制に関する情報	<a href="#">材料</a>
	製品、バッテリー、パッケージを含む電子廃棄物法規制に関する情報	<a href="#">WEEE 適合性</a>
	製品の回収および再利用プログラムに関する情報	<a href="#">Cisco 回収および再利用プログラム</a>
	持続性に関するお問い合わせ	お問い合わせ先： <a href="mailto:csr_inquiries@cisco.com">csr_inquiries@cisco.com</a>
材料	製品パッケージの重量と材料	お問い合わせ先： <a href="mailto:environment@cisco.com">environment@cisco.com</a>

## シスコおよびパートナーの提供サービス

Cisco は、業界をリードするパートナー企業とともに、AI と高パフォーマンス コンピューティング（HPC）向けの Cisco UCS® ソリューションへの移行を加速させるサービスをお届けします。Cisco Unified Computing Services™ は、俊敏性に優れたインフラストラクチャの構築、価値実現までの時間の短縮、導入および移行時の可用性の維持をサポートします。また展開後は、ビジネス ニーズの変化に応じたパフォーマンス、可用性、および復元力の向上をサポートすることで、さらなるリスクの軽減を可能にします。

詳細については、[https://www.cisco.com/c/ja\\_jp/products/servers-unified-computing/service-listing.html](https://www.cisco.com/c/ja_jp/products/servers-unified-computing/service-listing.html) を参照してください。

## Cisco Capital

柔軟な支払いオプションにより、必要な **Cisco®** テクノロジーをこれまで以上に簡単に入手できるようになります。

**Cisco Capital®** は最先端の決済ソリューションを提供し、最も重要なこと、つまりビジネスに集中できるようにします。特定のビジネス ニーズに合わせた柔軟な支払いオプションを使用して、ビジネスの成果を促進し、イノベーションとデジタル トランスフォーメーションを加速し、市場の動きに迅速に適応できるよう支援します。総所有コストの削減、資金の節約、成長の促進に役立ちます。現在から将来に渡ってシスコのテクノロジーのメリットをすべて活用しながら、お客様のビジネス要件に最適な方法でお支払いいただくことが可能となります。従量課金モデルをお求めの場合でも、**Cisco** ハードウェア、ソフトウェア、サービス、サブスクリプション、およびサードパーティ ソリューションをバンドルする必要がある場合でも、**Cisco Capital** がどのように役立つかを [こちら](#) からご確認ください。

### 詳細はこちら

[Dense GPU - HGX and OAM](#) の Web ページ にアクセスします。

**米国本社**  
Cisco Systems, Inc.  
カリフォルニア州サンノゼ

**アジア太平洋本社**  
Cisco Systems (USA), Pte. Ltd.  
シンガポール

**ヨーロッパ本社**  
Cisco Systems International BV  
Amsterdam, The Netherlands

2023 年 11 月発行

© 2023 Cisco and/or its affiliates. All rights reserved.

Cisco および Cisco ロゴは、Cisco Systems, Inc. またはその関連会社の米国およびその他の国における商標または登録商標です。シスコの商標の一覧については、[www.cisco.com/ip/go/trademarks](http://www.cisco.com/ip/go/trademarks) をご覧ください。記載されているサードパーティの商標は、それぞれの所有者に帰属します。「パートナー」または「partner」という言葉が使用されていても、シスコと他社の間にパートナーシップ関係が存在することを意味するものではありません。1175152207 10/23

