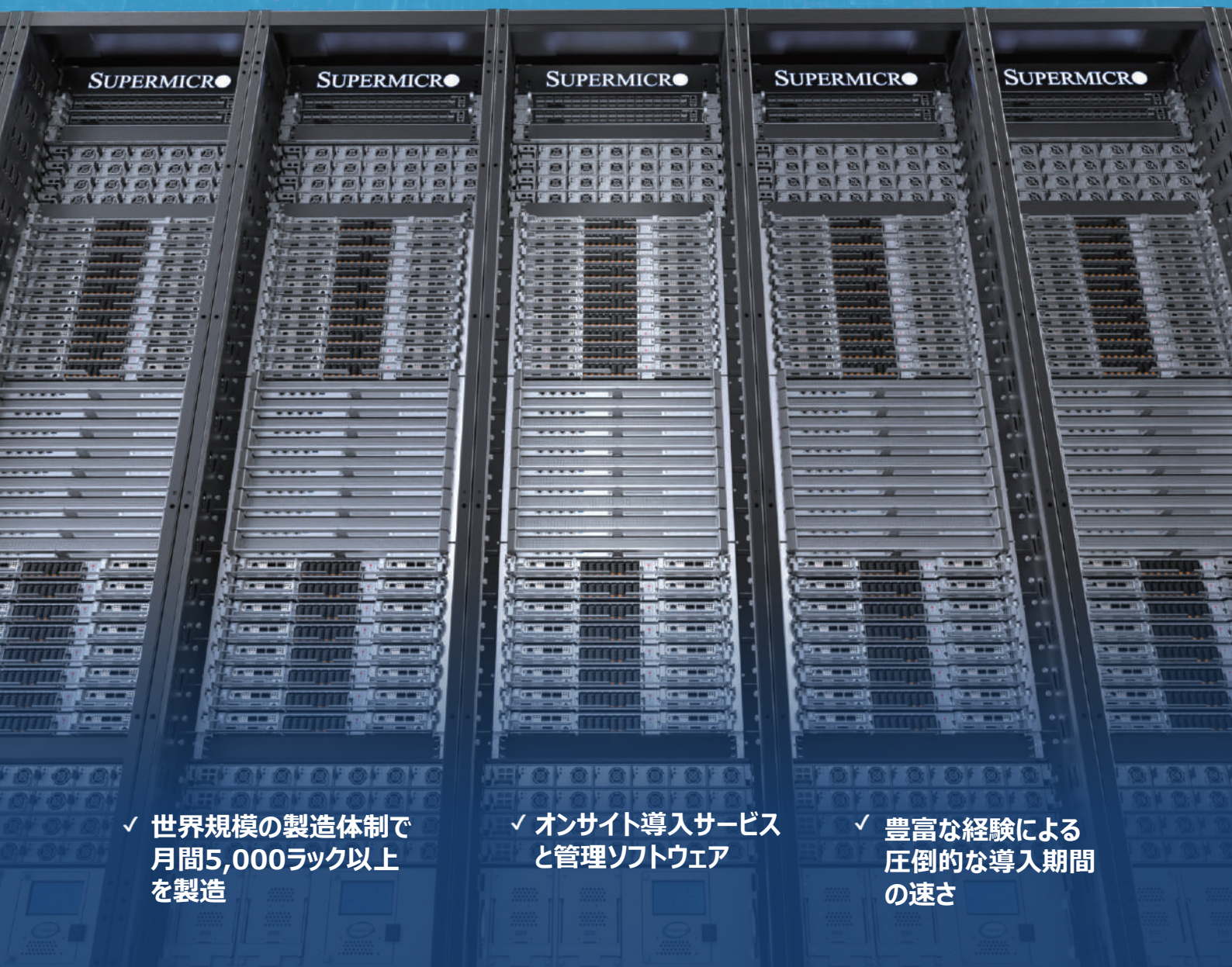




NVIDIA GB200 NVL72 SuperCluster

AIデータセンター向けエンドツーエンド液冷ソリューション



✓ 世界規模の製造体制で
月間5,000ラック以上
を製造

✓ オンサイト導入サービス
と管理ソフトウェア

✓ 豊富な経験による
圧倒的な導入期間
の速さ

1ラックでエクサスケール・スーパーコンピューターを実現



兆パラメータAIモデル用に構築された72GPUスケラブル・コンピューティング・ユニット

最も強力で効率的なNVIDIA Blackwellプラットフォーム

Supermicroは、NVIDIA Blackwell の導入を機に、液冷データセンターへの移行を加速し、AIインフラストラクチャーの急速に高まる需要に対してエネルギー効率が高い新しいソリューションを提供します。大規模なDLC（Direct-to-Chip）液冷AIシステムの豊富な導入経験を持つSupermicroの最先端の技術で、NVIDIA GB200 NVL72を強化し、シングルラックでエクサスケールコンピューティングを実現し、前世代と比べ、同じ電力で最大25倍のパフォーマンスを実現します。



- 72 NVIDIA Blackwell GPU：ラック内で最も効率的なエクサスケールコンピューティングを実現するために、HBM3eメモリの巨大なプールを備えた1つのGPUとして動作します。
- 液冷のパイオニア：データセンターの電力コストを最大40%削減するエンドツーエンドの液冷ソリューション
- 圧倒的な製造規模：最大の液冷ラックレベルの製造能力を持つSupermicroは、カリフォルニア州サンノゼ、ヨーロッパ、アジアの製造施設を持ち、GB200 NVL72のタイムリーで高品質な展開を実現します。
- 包括的なサービスの提供：Supermicroは、概念実証(PoC)から本番運用に至るまで、必要な部品、ネットワークソリューション、オンサイトの設置サービスをワンストップで提供します。
- 高度なネットワーキングへの対応：Supermicroは、NVIDIA BlueField®-3 SuperNIC、Spectrum™-X、Quantum-2、および次世代800Gb/秒ネットワーキングプラットフォームをいち早く採用しています。

Supermicroのエンドツーエンド液冷

施設側のクーリングタワーを含め、直接液体冷却インフラを簡単に導入・維持できます。

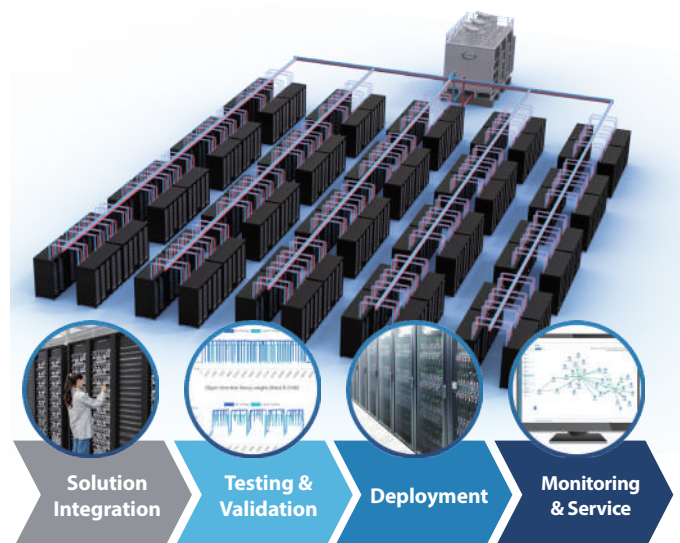


液冷データセンター管理のためのSuperCloud Composer (SCC)

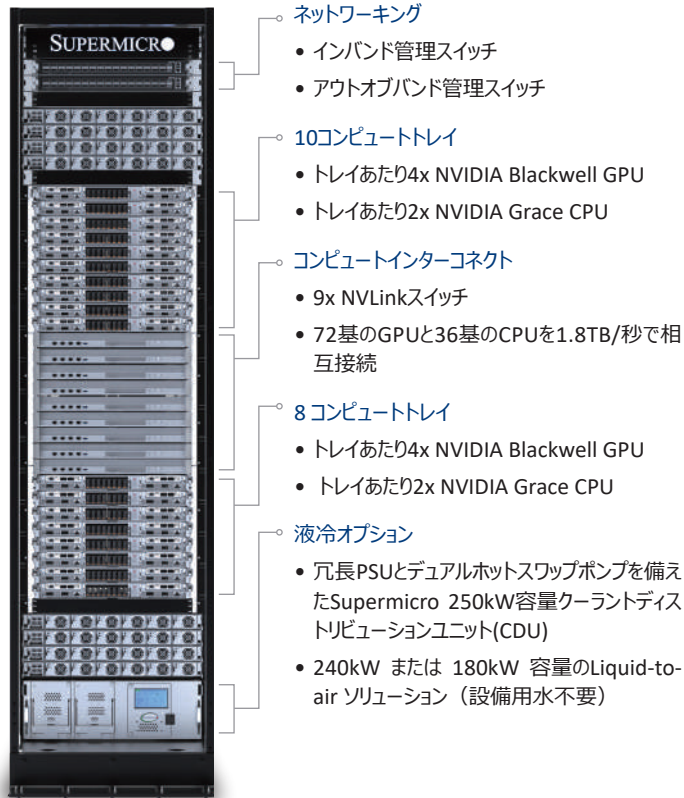
Supermicroの包括的なデータセンター管理プラットフォームであるSuperCloud Composerソフトウェアは、圧力、湿度、ポンプやバルブの状態など、液冷システムとラック、冷却水分配ユニット、冷却塔に関する重要な情報を監視する強力なツールを提供します。SuperCloud ComposerのLCCM (Liquid-Cooling Consulting Module) は、運用コストと管理コストを最適化します。

Supermicroのエンドツーエンド液冷ソリューションを搭載

Supermicro NVIDIA GB200 NVL72 SuperClusterは、NVIDIA GB200 Grace™ Blackwell Superchipsを収容するコンピュータトレイ用に設計された、最先端のインラック冷却液分配ユニット (CDU) とカスタムコールドプレートを搭載しています。NVIDIA GB200 NVL72は、完全に統合された液冷能力を備えたシングルラックで、エクサスケールコンピューティング能力を提供します。72基の NVIDIA Blackwell GPU と 36 基の Grace CPU が搭載され、NVIDIA の過去最大の NVLink™ ネットワークによって相互接続されています。NVLinkスイッチシステムは、低レイテンシーで毎秒130テラバイト (TB/s) のGPU通信を可能にし、AIやハイパフォーマンス・コンピューティング (HPC) ワークロードのパフォーマンスを向上させます。



NVIDIA GB200 NVL72ラックスケール構成



エンドツーエンドのオンサイト導入サービス

概念実証（PoC）から本番運用まで、Supermicroはワンストップショップとして、必要な部品、液冷、ネットワークソリューション、管理ソフトウェア、オンサイト設置サービスをすべて提供します。ワンストップショップとして、Supermicroは、様々なGPU、CPU、メモリモジュールに最適化されたカスタム設計のコールドプレート、複数の冷却液分配ユニットのフォームファクターと容量、マニホールド、ホース、コネクタ、冷却塔、監視・管理ソフトウェアを含む、包括的な社内液冷エコシステムを提供します。このエンドツーエンドのソリューションは、ラックレベルの構成にシームレスに統合され、システム効率を大幅に高め、サーマルスロットリングを緩和し、同時にAI時代のデータセンター運用の総所有コスト（TCO）と環境への影響の両方を低減します。



72-GPU Scalable Unit		SRS-GB200-NVL72-M1
GPUs	72x NVIDIA Blackwell B200 GPUs	
CPUs	36x NVIDIA 72-core Grace Arm Neoverse V2	
Compute Trays	18x 1U ARS-121GL-NB0	
NVLink Switch Trays	9x NVLink Switch, 4-ports per compute tray connecting 72 GPUs to provide 1.8TB/s GPU-to-GPU interconnect	
Power Shelves	8x 1U 33kW (6x 5.5kW PSUs), total power 132kW	
Rack Dimensions (mm)	W 600 x D 1068 x H 2236	
Liquid Cooling Options	<ul style="list-style-type: none"> 1x in-rack Supermicro 250kW capacity CDU with redundant PSU and dual hot-swap pumps 1.3MW capacity in-row CDU 180kW/240kW capacity liquid-to-air solutions for facilities without cooling tower and water supply 	

Compute Tray		ARS-121GL-NB0
Overview	1U Liquid-cooled System with 2x NVIDIA GB200 Grace Blackwell Superchips	
CPU and GPU	<ul style="list-style-type: none"> 2x 72-core NVIDIA Grace Arm Neoverse V2 CPU 4x NVIDIA Blackwell B200 per Superchip 	
GPU Memory	Up to 384GB HBM3e per Superchip (768GB per tray)	
CPU Memory	Up to 480GB LPDDR5X per Superchip (960GB per tray)	
Networking	4x NVIDIA NVLink Switch ports	
Storage	8x E1.S PCIe 5.0 drives	
Power Supply	Shared power through 4+4 rack power shelves	

Subject to change

Subject to change

AIデータセンターのエンドツーエンド液冷

幅広いAIデータセンター環境に対応する総合液冷ソリューション



冷却水分配ユニット

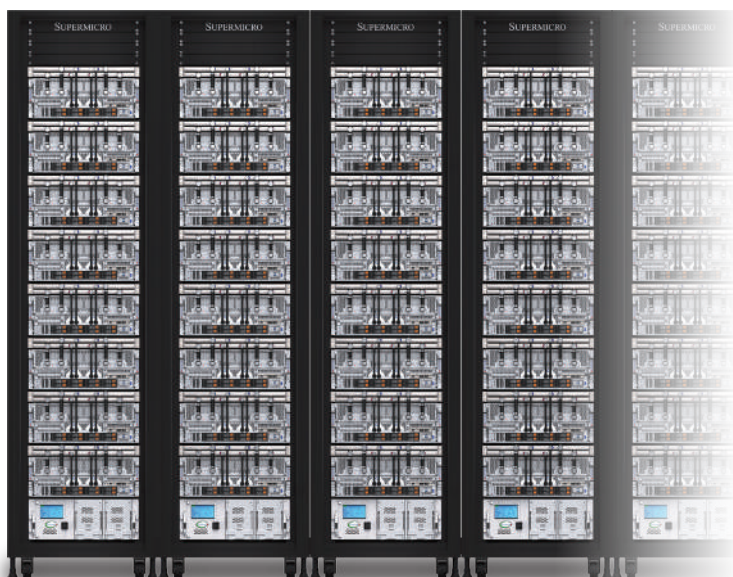


Liquid-to-Air クーリングラック
(設備用水不要)



冷却塔

NVIDIA HGX B200 8-GPU SuperCluster



NVIDIA HGX B200 8GPU用の新しい4U液冷および10U空冷システム

Supermicroの主要な4U液冷システムと新しい10U空冷システムは、NVIDIA HGX™ B200 8-GPUをサポートし、現在生産準備が整っています。新しく開発されたコールドプレートと250kW容量のラック内冷却水分配ユニットは、8xシステムの性能と効率を最大化し、48Uラックで64x 1000W Blackwell GPUと16x 500W CPUを提供します。新しい10U空冷システムは、ラックに完全に統合された最大4倍のシステムをサポートし、従来と同じ密度で最大15倍の推論性能と3倍のトレーニング性能を提供します。

4U液冷ラック構成



ネットワーキング

- インバンド管理スイッチ
- アウトオブバンドIPMI管理スイッチ
- ノンブロッキングネットワーク
- ネットワーク専用ラックまたは各コンピュータ・ラック内のリーフ・スイッチ

コンピュータとストレージ

- ラックあたり8台のSYS-422GA-NBRT-LCCまたはAS-4126GS-NBR-LCC
- 8x NVIDIA HGX B200 8-GPU/ラック
- 64x NVIDIA B200 Tensor Core GPU
- ラックあたり8x 1440GB HBM3e
- NVIDIA GPUDirect RDMAをフルサポートするローカルまたは専用ストレージファブリックによる柔軟なストレージオプション

液冷オプション

- 冗長PSUおよびデュアルホットスワップポンプを備えたSupermicro 250kW容量冷却液分配ユニット
- Supermicro冷却液分配マニホールド(水平または垂直)
- 240kW または 180kW容量のLiquid-to-air ソリューション (設備用水不要)

次世代スケーラブルユニット

数兆パラメータのAIモデル向けにインフラストラクチャーをスケーリングするSupermicroは、NVIDIA BlueField®-3 SuperNICと400Gb/秒のConnectX-7、さらには、NVIDIA Blackwell プラットフォームにおいて将来的に800Gb/秒ネットワーキングを可能にするConnectX-8、Spectrum-4、Quantum-3など、InfiniBandとイーサネットの両方にネットワーキングのイノベーションをもたらしています。NVIDIA Quantum-2 InfiniBandおよびSpectrum-X Ethernetと、Supermicroの4U液冷および8U空冷NVIDIA HGX H100およびH200システムクラスターは、現在、これまでで最大規模のAI配備を支えています。



64-GPU Scalable Unit

GPUs	8x NVIDIA HGX B200 8-GPU (64 GPUs)
CPUs	16x Intel® Xeon® 6 or AMD EPYC™ 9005 Series Processors
GPU Systems	8x SYS-422GA-NBRT-LCC / AS-4126GS-NBR-LCC
NVLink	5th Generation NVIDIA NVLink at 1.8TB/s
Networking*	<ul style="list-style-type: none"> • NVIDIA Quantum-2 InfiniBand 400G NDR • NVIDIA Spectrum-X Ethernet 400Gb/s • 2x Ethernet ToR management switch
Rack Dimension*	48U 800mm x 1200mm
Liquid Cooling*	<ul style="list-style-type: none"> • 1x in-rack Supermicro 250kW capacity CDU with redundant PSU and dual hot-swap pumps • 1.3MW capacity in-row CDU • 180kW/240kW capacity liquid-to-air solutions for facilities without cooling tower and water supply

*Subject to change

4U 8-GPU System

SYS-422GA-NBRT-LCC / AS-4126GS-NBR-LCC

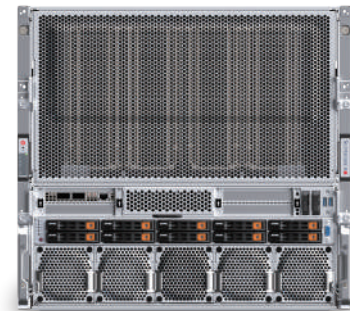
Overview	4U Liquid-cooled System with NVIDIA HGX B200 8-GPU
CPU	Dual Intel® Xeon® 6 or AMD EPYC™ 9005 Series Processors
Memory	<ul style="list-style-type: none"> • Up to 6TB DDR5-6400 or MRDIMM 8800MT/s (Intel) • Up to 9TB DDR5-6000 (AMD)
GPU	NVIDIA HGX B200 8-GPU (180GB HBM3e per GPU) 1.8TB/s NVLink GPU-GPU interconnect with NVSwitch
Networking*	<ul style="list-style-type: none"> • 8x NVIDIA ConnectX®-7 or BlueField®-3 SuperNICs • 2x NVIDIA ConnectX®-7 Dual-port 200Gbps/NDR200 QSFP112 NICs • 1:1 networking to each GPU to enable NVIDIA GPUDirect RDMA and Storage
Storage	8 front hot-swap 2.5" PCIe 5.0 x4 NVMe drive bays
Power Supply	4x 6.6kW redundant Titanium Level power supplies

*Subject to change.

10U空冷ラック構成

空冷または液冷

データセンター固有の冷却要件とインフラに応じて、複数のシステム構成とフォームファクターが利用可能です。最新の10Uフォームファクターは、空冷環境においてNVIDIA HGX B200 8-GPUをサポートし、モジュール式GPUトレイは、強化された熱性能と容易なメンテナンスのための大型ヒートシンクを収容することができます。GPU密度を最大化するために、液冷4UアーキテクチャをSupermicroのエンドツーエンド液冷ソリューションに統合することができます。標準48Uラックに最大8システム、合計64基のGPUを搭載することができます。



10U 8-GPU System

SYS-A22GA-NBRT / AS-A126GS-TNBR

Overview	10U Air-cooled System with NVIDIA HGX B200 8-GPU
CPU	Dual Intel® Xeon® 6 or AMD EPYC™ 9005 Series Processors
Memory	<ul style="list-style-type: none"> Up to 6TB DDR5-6400 or MRDIMM 8800MT/s (Intel) Up to 9TB DDR5-6000 (AMD)
GPU	<ul style="list-style-type: none"> NVIDIA HGX B200 8-GPU (180GB HBM3e per GPU) 1.8TB/s NVLink GPU-GPU interconnect with NVSwitch
Networking*	<ul style="list-style-type: none"> 8x NVIDIA ConnectX®-7 or BlueField®-3 SuperNICs 2x NVIDIA ConnectX®-7 Dual-port 200Gbps/NDR200 QSFP112 NICs 1:1 networking to each GPU to enable NVIDIA GPUDirect RDMA and Storage
Storage	10 front hot-swap 2.5" PCIe 5.0 x4 NVMe drive bays
Power Supply	6x 5250W redundant Titanium Level power supplies

*Subject to change



ネットワーキング

- インバンド管理スイッチ
- アウトオブバンドIPMI管理スイッチ
- ノンブロッキングネットワーク
- ネットワーク専用ラックまたは各コンピュータ・ラック内のリーフ・スイッチ

コンピュータおよびストレージ

- ラックあたり4x SYS-A22GA-NBRT または AS-A126GS-TNBR
- 4x NVIDIA HGX B200 8-GPU/ラック
- 32x NVIDIA B200 Tensor Core GPU
- ラックあたり4x 1440GB HBM3e
- NVIDIA GPUDirect RDMAをフルサポートするローカルまたは専用ストレージファブリックによる柔軟なストレージオプション

32-GPU Scalable Unit

GPUs	4x NVIDIA HGX B200 8-GPU (32 GPUs)
CPUs	8x Intel® Xeon® 6 or AMD EPYC™ 9005 Series Processors
GPU Systems	SYS-A22GA-NBRT or AS-A126GS-TNBR
NVLink	5th Generation NVIDIA NVLink at 1.8TB/s
Networking*	<ul style="list-style-type: none"> NVIDIA Quantum-2 InfiniBand 400G NDR NVIDIA Spectrum-X Ethernet 400Gb/s 2x Ethernet ToR management switch
Rack Dimension	48U 750mm x 1200mm

*Subject to change

AIデータセンター向けエンドツーエンド液冷ソリューション



Worldwide Headquarters

Super Micro Computer, Inc.
980 Rock Ave.
San Jose, CA 95131, USA
Tel: +1-408-503-8000

EMEA Headquarters

Super Micro Computer, B.V.
Het Sterrenbeeld 12, 5215 ML,
's-Hertogenbosch, The Netherlands
Tel: +31-73-640-0390

APAC Headquarters

Super Micro Computer, Taiwan Inc.
3F, No. 150, Jian 1st Rd., Zhonghe Dist.,
New Taipei City 235, Taiwan
Tel: +886-2-8226-3990



お問い合わせ：スーパーマイクロ株式会社

〒150-0031 東京都渋谷区桜丘町 20-1 渋谷インフォスター 21 階

電話：03-5728-5196 FAX：03-5728-5197 Email：Sales_Inquiry_JP@Supermicro.com

www.supermicro.com

© Super Micro Computer, Inc. Specifications subject to change without notice. All other brands and names are the property of their respective owners.



Please Recycle

MKT-0002_2024-GB200-NVL72_R08