

搭載 Arm AGI CPUs 的 Supermicro 解決方案

最佳化架構，解鎖突破性的效能、效率與機架密度，支援 24/7 全時運作的 Agentic AI



AI 協調調度的新一代運算解決方案

- 新型 2U、5U 與多節點架構，提供液冷與氣冷選項
- 機架規模設計，應對現代 Agentic AI 日益增長的需求
- 透過 Supermicro 完整的建構區塊解決方案 (DCBBS) 模組化基礎設施，實現更快速的大規模部署
- 基於高效能、低功耗且高核心密度的 Arm AGI CPU，每顆 CPU 最高 136 核心，每核心記憶體頻寬達 6GB/s

專為大規模 AI 設計

Supermicro 搭載 AI 導向 Arm AGI CPU 的解決方案，正是為大規模 Agentic AI 協調調度時代所打造，提供高效能、高效率與高密度的特性，最大化機架規模部署的經濟效益。Supermicro 憑藉領先市場的實績，結合大規模建構區塊解決方案 (DCBBS) 的部署能力，以及自研熱管理技術，打造出專為現代 Agentic AI 優化的全新架構類別。

Supermicro 更快部署 AI 基礎設施

這些全新且創新的、針對特定工作負載的架構，旨在最大化 Arm AGI CPU 的效能與密度，同時確保電力與散熱效率，降低 TCO 與 TCE (對環境的總成本)。Supermicro 的模組化建構區塊理念、成熟的自有液冷技術，以及機架整合專業能力，意味著這一系列新架構能夠更快進入市場，讓組織得以更早開始實現 AI 基礎設施投資的回報。憑藉全球每月最高達 6,000 個機架的製造規模，Supermicro 已準備好在任何規模下交付完整整合的 AI 叢集，這些叢集均在出貨前於廠內完成整合、測試與驗證，可於交付當日立即上線運作。

DCBBS 完整 Agentic AI 解決方案

Supermicro 的資料中心建構區塊解決方案 (DCBBS) 是現代資料中心的藍圖，提供完整、模組化的 AI 基礎設施，由經驗證的元件與子系統所組成，具備端到端的部署靈活性。透過利用 Arm AGI CPU，Supermicro 解決方案能比傳統架構提供每機架超過 2 倍的效能，並協助企業在每 Gigawatt 的 AI 資料中心容量上節省高達 100 億美元的資本支出 (CAPEX)。基於 Supermicro 業界領先的機架密度與每瓦效能，這些解決方案能確保資料中心空間與電力資源獲得最大化利用。

新一代 Agentic AI 運算解決方案

透過充分利用 Arm AGI CPU 的高核心密度與每瓦效能，Supermicro 系統在相同電力範圍內，能比傳統平台提供高達 2 倍的每機架效能與 2 倍的核心密度，大幅節省電力消耗與機房空間。136 核心的 Arm AGI 微架構能減少傳統負擔，在每個週期完成更多工作，實現持續且不節流的效能。其他特色還包括每核心 6GB/s 記憶體頻寬與延遲最佳化的記憶體存取，以支援線性擴展；同時具備擴展的記憶體容量與靈活的 I/O，能打造能源效率高且分散式的 Agentic AI 基礎設施，讓 CPU 得以協調調度數千個平行任務。



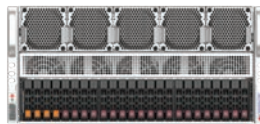
2U 4 節點 ORV3 機架規模解決方案
超高密度多節點機架規模液冷架構



1U 4 節點 ORW 機架規模解決方案
最高運算密度的多節點液冷 ORW 機架架構

機架規模解決方案

主要應用	大規模 AI 推論、雲端運算、超大規模	大規模 AI 推論、Agentic AI 協調調度、雲端運算、超大規模 CPU 服務、高密度 Scale-out 工作負載
機架配置	44-OU/48-OU OCP ORV3 每機架最高支援 76 個節點	44-OU/48-OU OCP ORW 每機架最高支援 168 個節點
每機架 CPU/核心數	最高支援 152 顆 CPU / 20,672 核心	最高支援 336 顆 CPU / 45,696 核心
冷卻解決方案	CPU - 直接晶片液冷 列內式或機架內 CDU	CPU - 直接晶片液冷 列內式或機架內 CDU
每機架電力預算	最高 95kW	最高 210kW



5U 8-GPU
新一代效能與效率的旗艦式
機架安裝解決方案



2U Hyper
熱最佳化架構，具備靈活的 GPU 支援



2U Hyper-E
單 CPU 前置 I/O 架構，
針對邊緣部署最佳化

高密度機架式

型號	ARS-522GP-NR	ARS-222H-NR	ARS-212HE-FNR
主要應用	GPU 加速推論、AI 訓練、模型開發、HPC、模擬、渲染	AI 協調調度、CPU 驅動推論、記憶體密集型工作負載、雲端運算、虛擬化、靈活的 GPU 加速推論	AI 推論、邊緣運算、雲端運算、CDN
CPU	雙 Arm AGI CPU Neoverse V3，每顆 CPU 支援 64/128/136 核心	雙 Arm AGI CPU Neoverse V3，每顆 CPU 支援 64/128/136 核心	單 Arm AGI CPU Neoverse V3，每顆 CPU 支援 64/128/136 核心
記憶體	24 個 DIMM 插槽；1DPC 最高支援 6TB DDR5-8800MT/s	24 個 DIMM 插槽；1DPC 最高支援 6TB DDR5-8800MT/s	12 個 DIMM 插槽；1DPC 最高支援 6TB DDR5-8800MT/s
PCIe 擴充插槽	8 PCIe 5.0 x16 FHFL double-width slots 4 PCIe 5.0 x16 FHHL slots 1 PCIe 6.0 x16 FHFL slot 1 PCIe 5.0 x8 AIOM slot (OCP 3.0 compatible)	5 PCIe 6.0 x16 FHHL slots 1 PCIe 6.0 x8 FHHL slot 1 PCIe 6.0 x8 AIOM slot (OCP 3.0 compatible)	Default 3 PCIe 6.0 x16 FHHL slots 1 PCIe 6.0 x8 FHHL slot 1 PCIe 6.0 x16 AIOM slot (OCP 3.0 compatible) Option A 4 PCIe 6.0 x16 FHHL slots 1 PCIe 6.0 x16 AIOM slot (OCP 3.0 compatible) Option B 1 PCIe 6.0 x16 FHHL slot 4 PCIe 6.0 x8 FHHL slots 1 PCIe 6.0 x16 AIOM slot (OCP 3.0 compatible)
GPU 支援	最高支援 8 張雙寬度	最高支援 2 張雙寬度	最高支援 2 張雙寬度
儲存	8 個前置熱抽換 2.5" NVMe	8 個前置熱抽換 2.5" NVMe	預設 6 個前置熱抽換 E1.S NVMe 選項 A 4 個前置熱抽換 E1.S NVMe 選項 B 6 個後置熱抽換 2.5" NVMe
電源	6 組 2700W 冗餘 (3+3) 鈦金級 (96%) 熱抽換電源供應器	2 組 2700W 冗餘 (1+1) 鈦金級 (96%) 熱抽換電源供應器	2 組 2000W 冗餘 (1+1) 鈦金級 (96%) 電源供應器