



重塑AI應用邊界： AMD與MiTAC技術賦能AI Appliance開發

林祺政博士 AI研發中心
神通資訊科技股份有限公司



豐富的系統整合經驗



MiTAC 是台灣少數擁有多樣客戶、並具備大型專案整合輸出能力的系統整合商，在資通訊市場扮演重要的腳色，服務涵蓋政府機構、國/民營企業、教育單位、金融機構、交通、工廠等

MiTAC-AI落地戰略核心



專注於AI解決方案

發展以生成式AI機器人為主軸的企業級解決方案。



提供企業級的落地型AI

服務對象涵蓋金融、教育和各類企業客戶。



AMD深度合作

Leverages AMD的AI硬體與軟體平台，實現高效能、成本效益與未來就緒的AI基礎設施。

MiSBOT一體機提供端點服務的AI體驗

- 為企業量身打造的聊天機器人服務，部署於搭配AMD Instinct™ 系列的 AI Server，確保在地化部署效能，或運用 AMD Ryzen AI 輕量化架構提供神奇的AI體驗

AMD戰略夥伴

- 合理的價格以及優秀的效能
- 完善的產品Roadmap為未來AI藍圖做好準備



AMD AI技術 x 平台賦能

AMD Instinct™ GPU：世界級加速技術

包括MI200、MI300和MI325系列加速器

開放軟體生態系統：ROCm™

企業級的開放AI軟體，與PyTorch、Triton、Huggingface等開源工具兼容，從「移植」到「開發」，提供卓越價值

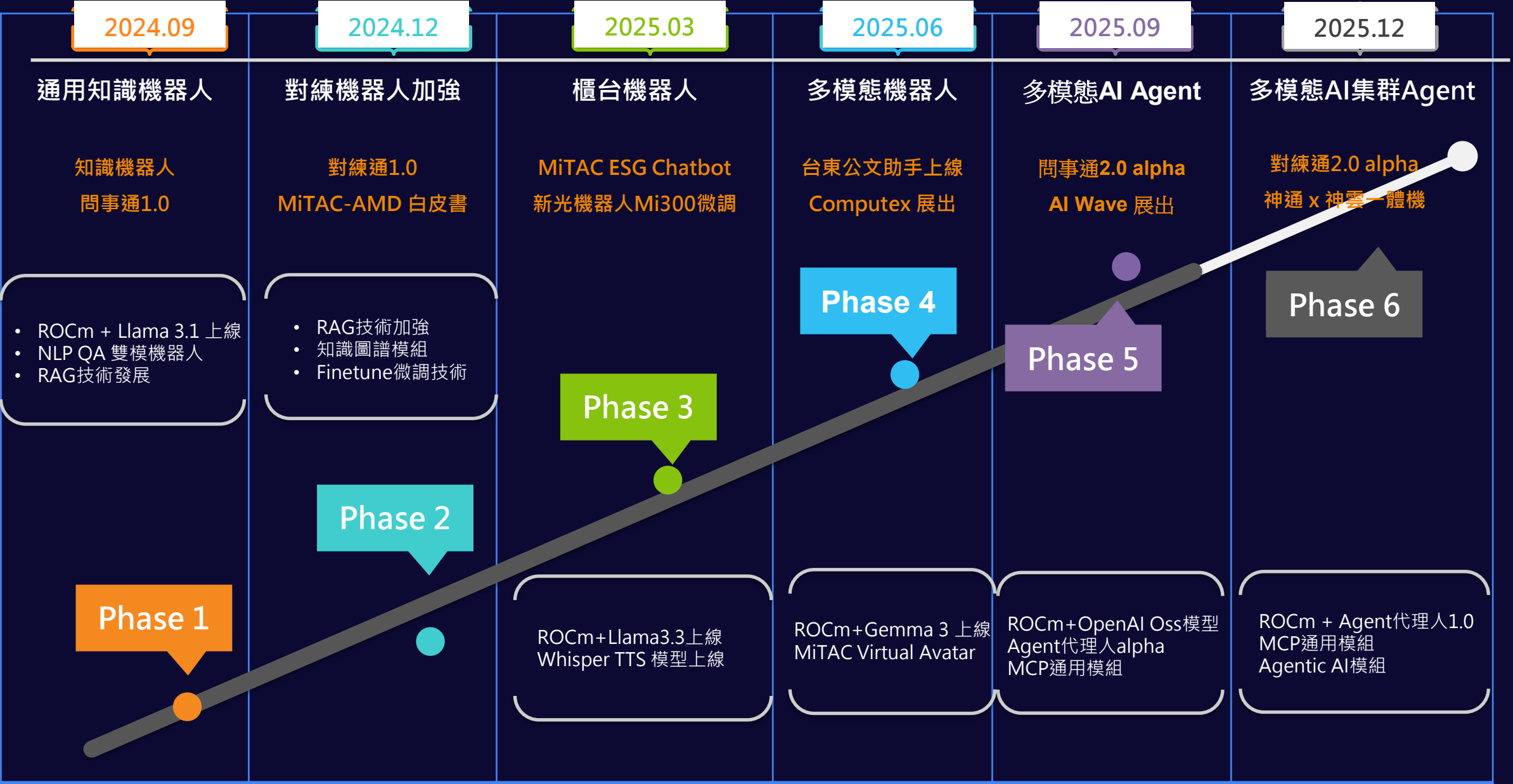
AMD Instinct™ MI300X 的領導優勢

- 192GB HBM3記憶體，峰值理論頻寬約5.3 TB/s。
- 支援高達2.4倍更大的模型尺寸，減少CPU卸載，提升約2倍效能。
- 在記憶體容量、記憶體頻寬、HPC及AI效能上均具領先地位。

AMD Ryzen AI PC：協同運算中樞

- 整合NPU、CPU、GPU，提供協同處理，達到最佳效率與能效。
- NPU是產生式/內容創作模型的理想選擇。

AI產品歷程



神通生成式AI落地案例



整合多功能 AI 機器人技術的平台，包括文字、語音機器人和虛擬 AI 助理

ChatBot

- 法務諮詢專家
- 庶務全能通
- 人資諮詢助理
- 業務小助理
- 公文生成系統

VoiceBot

- ESG服務諮詢
- 優秀理專系統
- 業務招募助手
- 智能客服外撥

Virtual Assistant

- 智能教練
- AI Coach
- 理專對練
- 業務培訓

AI Training Course

- AI 素養 • 數據匯流
- AI 工具
- AI 應用
- AI Lab





神通MiSBOT一體機

- 企業級客製化聊天機器人系統，部署於搭載AMD Instinct™ MI210 AI Server
- 採用Meta-Llama-3-8B-Instruct/Gemma 3 27B模型進行調校

效能實測驗證

- 搭載單一Mi210 AI Server 在維持服務品質下 (TTFT約2秒，TPOT約0.2秒)，同時服務80位併發用戶
- 一般企業使用可服務100人以上

MiSBot.ai 大模型方案



整合多功能AI機器人技術平台

包括文字、語音機器人和虛擬AI助理



GPT管理功能

設定啟用、模型資源查詢、使用統計、紀錄稽核



Chatbot應用



Voicebot應用

金融保險領域 x AMD 應用

金融模型在地開發與優化

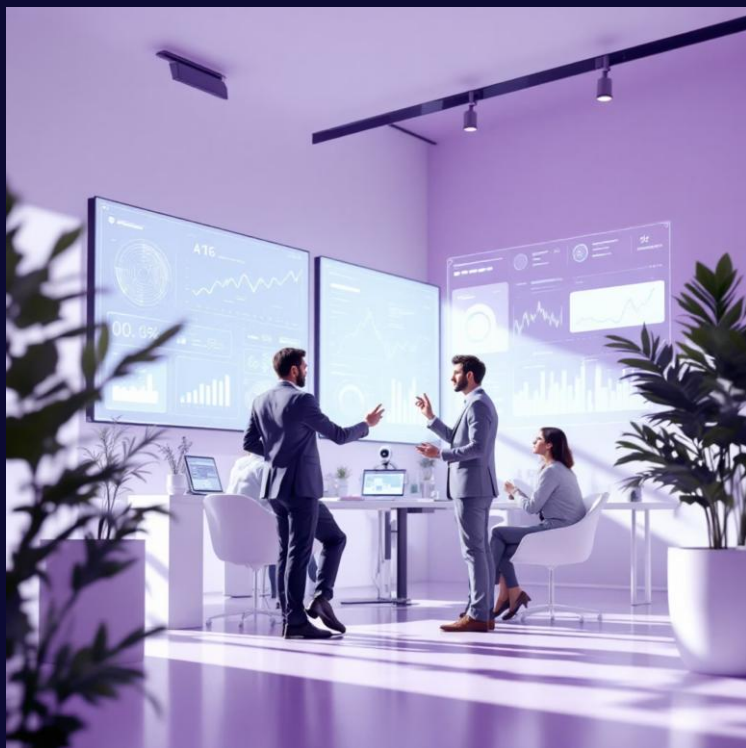
MiTAC協助金融機構開發在地模型，並提供技術優化顧問服務
特別針對AMD GPU賦能的金融應用

MI300系列GPU微調金融模型

使用 AMD Instinct™ MI300系列GPU微調金融模型，
發揮其大記憶體與高效能優勢

金融領域MiSBOT 3.0

提供ESG服務諮詢、優秀理專系統、業務招募助手等功能，支援金融業務。



語音機器人



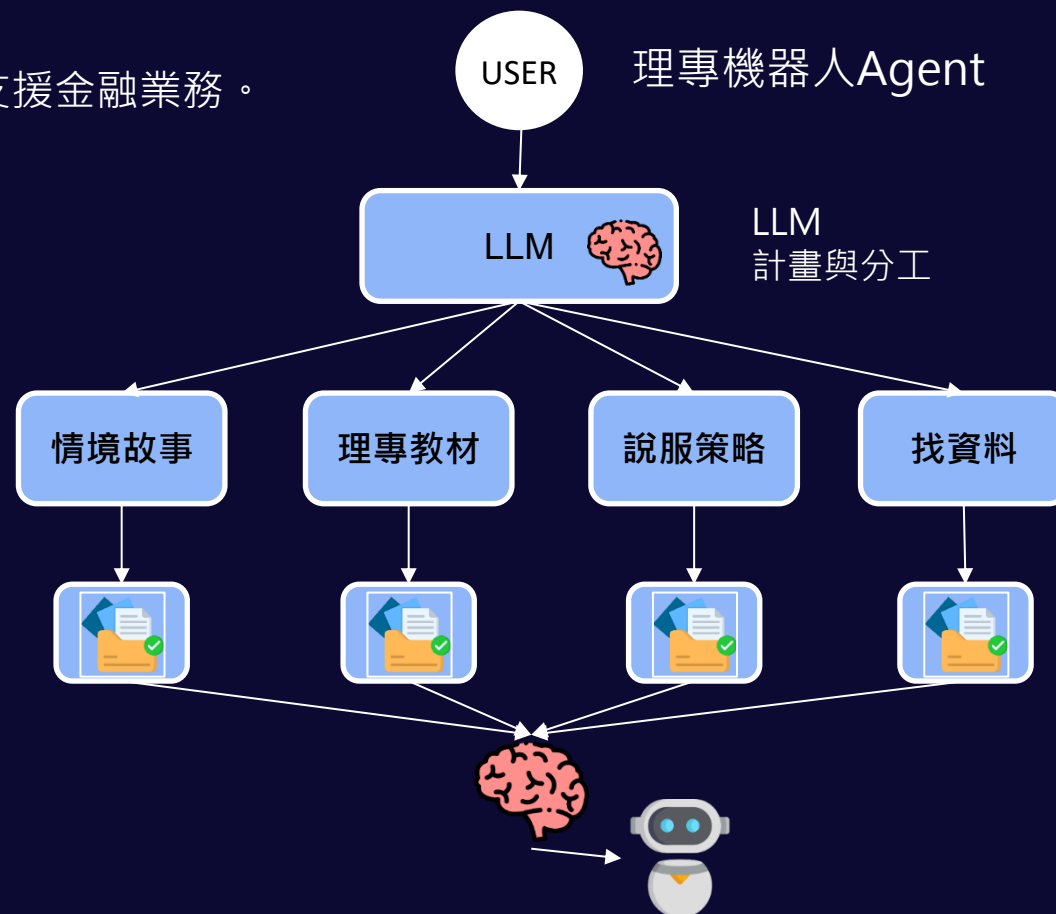
ESG服務諮詢



優秀理專系統



業務招募助手



科技執法與智慧城市 x AMD 應用

環保稽查科技執法系統

自動偵測機車與行人違規丟棄的行為，實現環保稽查的智能化

科技執法 AMD Instinct™ MI210 Video Service

2025年3月推出，利用 AMD Instinct™ MI210的運算能力進行執法相關的影像分析

新北市環保局、台北市衛生局、宜蘭縣衛生局、新竹縣衛生局、屏東縣衛生局上線



台北市衛生局
抽菸偵測、嫌疑姿態偵測

新北市環保局
垃圾丟棄，每月1000+案件

新北市環保局案例 實際效益

New Taipei City Environmental Protection
Bureau Case Study



BEFORE

人工辨識負擔大

新北市環保局在違規熱點建置1029支移動式監視器，過去要透過人工逐一辨識影片畫面，縱使運用縮時攝影方式審查，人工辨識也大約需要**2-3小時**，耗費大量人力時間成本，加重稽查負擔。

科技稽查 效率提升8倍以上

經由本系統輔助，每日24小時錄影只需**3-5分鐘**即可判讀完成，效率較過去人工辨識提升 **8 倍以上**，大幅提升執法效率，減輕勤務負擔。



AFTER

神通AI教育領域 x AMD 應用

神通高中AI教育理念

致力於打造涵蓋AI素養、AI工具、AI應用和AI實驗室的完整高中AI教育體系，培育AI教師與學生，建立完整AI人才生態系統

AI實驗室建置與ROCm實作經驗

透過AMD Instinct GPU效能與開源軟體ROCm實作，賦能大語言模型在教育領域的應用，為教學實作提供強大支持



神通AI教育領域 x AMD 應用

型號	規格	數量
MiTAC AX2 Chatbot	CPU	AMD EPYC Milan 7C13, 64 Core 2.45GHz
	GPU	AMD Mi210 64Gb GDDR
		AMD Infinity Fabric link 2-port
	RAM	DDR4-3200 64GB
	Storage NVMe	2.5" U.2 SSD 1.0TB
	Storage SATA	2.5" SATA SSD 2.0TB
	SFP+	Intel X710-DA2 Dual Port 10GbE/1GbE
	Power	2200W AC 200~240V/12.6A
	Cooling	8cm Air Cooling Fans
	Dimension (D x W x H)	32.71" x 17.25" x 3.43" (831 x 438.4 x 87mm)
		2U Rackmount



型號	規格	數量
MiTAC AV1 AutoML	CPU	AMD EPYC 7C13 64Core 128 Thread 2.25GHz
	RAM	DDR4-2999 32GB
	Storage SATA	2.5" SATA SSD 1.92TB
	MegaRAID	SAS 9361-8i
	LAN	Broadcom BCM5720 1GbE Port
	Power	800Watts (200-240V AC input)
	Cooling	8cm Air Cooling Fans
	Dimension (D x W x H)	25.59" x 17.32" x 3.43" (650 x 440 x 87mm)
		2U Rackmount



種子老師課程

圖像生成



馬公高中

教材規劃



羅東高中

AI素養



石門實中

機器學習



課程構想



生成式AI



佈建高中AI運算資源

運算門檻降低，AI專題不再只是夢想

可即時使用高效能GPU

支援LLM訓練與多輪對話測試

不需自建昂貴設備也能啟動AI環境



落地架構

Application Server

AI Core Platform



Dell / HP / ASUS AI PC
GPU : AMD Ryzen AI PRO 340
總算力 : 50 AI TOPs

30~100人以下 AI 推論主機
GPU : AMD MI210 4片
總算力 : 724 AI TOPs

30人以下 AI 推論主機
GPU : AMD Radeon PRO W7900 DS
總算力 : 123 TFLOPs

AI 預訓練主機 / AI 模型微調
GPU : AMD MI3xx 系列 8顆
最高總算力 : 28,856 AI TOPs

OneMiTAC x AMD-完整的整合與協同



神通資料扮演專業系統整合商，
結合了神雲科技高效能的硬體
規格，完整實現AI多領域落地的
戰略佈局

- ROCm提供企業級的開放式AI軟體，
與PyTorch及Huggingface開源相容
- 賦能MiTAC在AMD Platform上快速部署和
優化AI模型，提供強大支援
- AMD Instinct MI300X等領先GPU加速器
為在地AI部署提供世界級的強大效能
- 共同發布「MiTAC-AMD White Paper」
提升彼此雙方協同目標





重塑 AI 應用邊界： AMD 與 MiTAC 技術賦能 AI Appliance 開發

陳玟錡 Stephanie | 亞太區業務總監

神雲科技股份有限公司

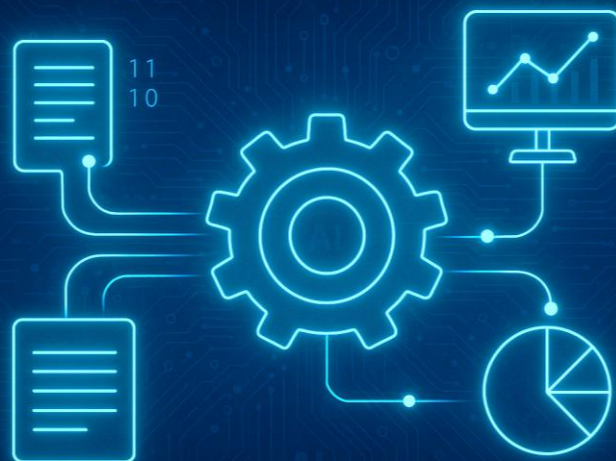
MiTAC Computing x AMD AI 落地策略 | 驅動企業轉型

即時提供客戶服務



AI 能透過智慧客服與對話式系統，快速回應客戶需求

自動化工作流程



AI 助手可以自動蒐集、整合並分析數據，協助企業達到更高效能

數據驅動決策

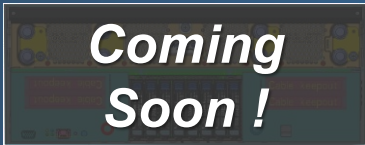


AI 能協助管理層從中萃取洞察，優化決策過程

AI 模型選擇策略

應用場景	模型大小	範例模型	同時使用人數	Token/s 需求	精度	散熱需求	備註
簡單問答 (FAQ、客服)	< 70B	8B、70B	10-50	~10-15	FP16 / INT8 (可壓縮)	氣冷即可	成本低，反應快，適合客服、內部知識查詢
複雜問答 (專業分析、文件摘要)	≥ 70B	70B、235B	30-100	~15-20	FP16	氣冷即可	更高的上下文需求，適合文件分析、商務應用
高精度決策支援 (金融、供應鏈優化)	≥ 235B	235B	50-100	~20-30	FP16 / BF16	氣冷或水冷	大模型推理能力佳，適合策略分析、數據洞察
超大規模 AI 推理 (大語言模型部署、AI 工廠)	≥ 685B	685B	30-100+	>30	FP16 / BF16 / FP8	氣冷或水冷	超大模型，需高並行算力與高效能冷卻系統

MiTAC Computing x AMD 伺服器解決方案

模型大小	使用人數	Token/s	精度	散熱	對應伺服器	GPU 支援	記憶體	儲存/插槽	特點
70B	50-100	15	FP16	氣冷	TYAN TN85-B8261 	4× AMD Instinct MI210	24 DDR5-6400, up to 6TB	6× PCIe 5.0 x16, 8x NVMe U.2	相容多樣 PCIe，可以根據需求去擴充
235B	100	15	FP16	氣冷	G8825Z5 	8× AMD Instinct 325X	24 DDR5-6400, up to 6TB	12× PCIe 5.0 x16, 8x NVMe U.2	OAM的 GPU模組設計，擁有 GPU/NIC1 對1，可以滿足GPU叢集擴充
685B	30	15	FP16	水冷	G4826Z5 	8× AMD Instinct MI355X	24 DDR5-6400, up to 6TB	10× PCIe 5.0 x16, 8x NVMe E1.S	液冷設計，使高功耗 GPU穩定運行。



Thank you!