



## 美國科技產業

# Nvidia GTC 2024：新 GPU 平台、軟體服務與機器人

### 焦點內容

1. Nvidia 舉辦 GTC 2024 大會，推出：(1) 如期推出 Blackwell 平台；(2) 推出 Nvidia NIM、NeMo 微服務；(3) 用於人型機器人的 Project GR00T，與 Jetson Thor 晶片；(4) 以 API 形式提供 NVIDIA Omniverse Cloud。
2. GB200 NVL72(水冷機櫃，含 36 個 GB200 超級晶片)，在推論、訓練、能耗與資料處理表現上均較 HGX H100 大幅提升。多家雲端 CSP (Microsoft、AWS、Google Cloud 和 Oracle Cloud Infrastructure)均表示將會部屬 GB200。
3. 我們預期廣達、鴻海、緯創與緯穎都將是主要受惠者。我們預期 2025-26 年液冷滲透率將上升，進而嘉惠酷碼、奇鋳與雙鴻。

### Nasdaq 指數

Nasdaq 指數, 點



資料來源：Bloomberg

### SOX 指數

SOX 指數, 點



資料來源：Bloomberg

### 重要訊息

Nvidia 舉辦 GTC 2024 大會，推出次世代資料中心 GPU 平台 Blackwell，以及 AI Foundry 服務。會議重點：(1) 如期推出 Blackwell 平台，包含 B100 晶片、GB200 超級晶片、GB200 NVL72 機櫃，以及用於 GPU 互聯的 NVlink Switch 晶片；(2) 推出 Nvidia NeMo 與 NIM 微服務，搭配 DGX Cloud，皆為 AI Foundry 服務一環，簡化企業開發、部署、客製化與運行生成式 AI 模型；(3) 推出 Project GR00T 人型機器人通用基礎模型，以及 Jetson Thor 晶片；(4) 以 API 形式提供 NVIDIA Omniverse Cloud，讓軟體業者便於創造數位孿生應用。

**如期推出 Blackwell GPU 平台。**Blackwell GPU 使用兩個晶粒封裝(台積電 N4P，單一晶粒電晶體較 Hopper 提升約 30%)，並包含第二代 Transformer 模型與第五代 NVLINK，訓練上較 Hopper 提升 2.5 倍 TFLOPS FP8 算力，推論上帶來 5 倍 FP4 算力提升，Blackwell HGX 與 Hopper HGX 系統兼容，利於快速部署。Blackwell HGX 與 Hopper HGX 相容，因此用戶可輕易升級。Blackwell 將於今年晚些時候推出，Nvidia 指出美國主要 CSP，包含：AWS、Google Cloud、Microsoft Azure 和 Oracle Cloud Infrastructure)、以及部分二線 CSP 廠商，例如：Applied Digital、CoreWeave、Crusoe、IBM Cloud 和 Lambda，以及主權 AI 雲服務商均會部屬 Blackwell。

**GB200 有望為下一代資料中心部屬標準。**GB200 超級晶片包含 2 個 B200 GPU 和 1 個 Grace CPU。隨著多模態 AI 模型參數達兆級，算力需求持續提升，Nvidia 也推出 GB200 NVL72(含 36 個 GB200 超級晶片及液冷機櫃的 DGX GB200 系統)，在推論、訓練、能耗與資料處理表現上均較 HGX H100 大幅提升。GB200 亦預計於今年晚些時候推出，相較於 GH200 佔 Hopper 系列比重較低(定位為讓客戶熟悉 Grace 系統)，多家雲端 CSP (Microsoft、AWS、Google Cloud 和 Oracle Cloud Infrastructure)均表示將會部屬 GB200。

**GB200 強化 AI 機櫃系統與液冷機會。**GB200 強化 AI 機櫃系統與液冷機會。我們預期廣達 (2382 TT, NT\$257, 增加持股)、鴻海 (2317 TT, NT\$136, 增加持股)、緯創 (3231 TT, NT\$122.5, 增加持股) 與緯穎 (6669 TT, NT\$2,175, 增加持股) 都將是 GB200 主要受惠者。Nvidia 的主要美系客戶 AWS、Microsoft、Oracle、Meta、Open AI 與 Tesla 都將是採用 GB200 超級晶片的 DGX 電腦的買家，而其在台的 ODM 業者將從中受惠。儘管 GB200 設計上沒有基板，但緯創將提供多項板材(如 NVLink 交換板與其他)，因此供貨之內含價值可能高於目前的 H100。然而，GB200 預計今年稍後才會上市，故凱基預期系統/機櫃生產時間可能落在 2Q25 之後。另外，根據 Nvidia 設計，配備 GB200 超級晶片的 DGX 電腦及 GB200 伺服器機櫃將使用 MGX 液冷設計，因而我們預期 2025-26 年液冷滲透率將上升，進而嘉惠酷碼(未上市)、奇鋳 (3017 TT, NT\$569, 增加持股) 與雙鴻 (3324 TT, NT\$655, 增加持股)；冷卻液監控主機 (CDU)、歧管與快接頭需求亦可望隨之升溫，增加市場商機。

### 投資建議

凱基維持 Nvidia 「增加持股」評等，與目標價 950 美元。台股相關供應鏈看好廣達、鴻海、緯創、緯穎、奇鋳與雙鴻，將成為 AI 趨勢與 GB200 新產品週期的主要受惠者。

### 投資風險

全球消費及 ICT 需求不如預期；消費者對新產品的接受度不高。

圖 1：主要指數表現

Index	Recent close (pts)	1W (change, %)	2W	1M	3M	6M	12M	YTD
Dow Jones	38,790	0.1	12.0	0.4	4.0	12.0	21.7	2.9
Nasdaq	16,103	0.5	17.5	2.1	8.0	17.5	38.5	7.3
SOX	4,758	(2.7)	36.2	5.1	15.9	36.2	54.3	13.9

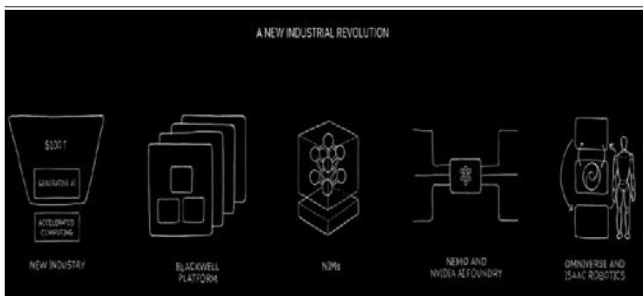
資料來源：Bloomberg

圖 2：同業評價比較表

Company	Ticker	Market cap (US\$bn)	Turnover (30 days moving avg.) (US\$mn)	Rating	Target price (US\$)	Share price (US\$)	Upsides (%)	PE (x)			EPS (US\$)			EPS YoY (%)		
								2023	2024F	2025F	2023	2024F	2025F	2023F	2024F	2025F
NVIDIA	NVDA	2,211	46,790	OP	950.0	884.6	7.4	68.2	28.7	22.9	12.96	30.78	38.61	283.8	137.5	25.4
Broadcom	AVGO	573	4,265	OP	1,435.0	1,237.2	16.0	29.3	22.2	21.9	42.25	55.78	56.39	12.3	32.0	1.1
AMD	AMD	308	12,911	OP	200.0	190.7	4.9	72.0	46.9	36.1	2.65	4.06	5.27	(25.1)	53.4	29.8
Intel	INTC	181	1,913	OP	53.0	42.7	24.1	40.9	30.8	20.3	1.04	1.39	2.10	(43.4)	32.7	51.6
Marvell	MRVL	58	1,114	OP	85.0	67.2	26.5	44.7	29.3	24.0	1.50	2.29	2.80	(29.1)	52.3	22.4
Dell	DELL	76	850	OP	140.0	106.6	31.3	15.0	13.4	10.8	7.12	7.97	9.88	(6.4)	11.9	24.0
HPE	HPE	22	291	N	18.0	17.1	5.5	7.9	8.6	8.3	2.15	1.97	2.06	6.8	(8.3)	4.4

資料來源：Bloomberg、凱基

圖 3：Nvidia GTC 五大宣布



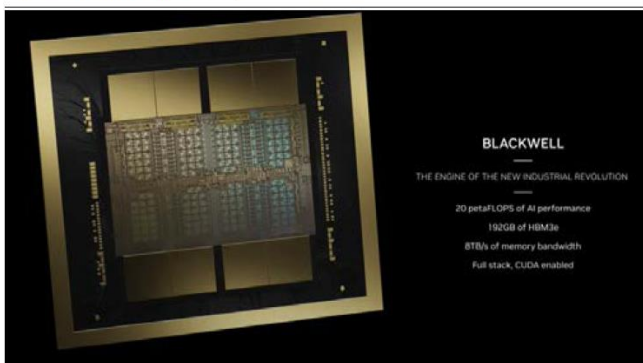
資料來源：公司資料

圖 4：Nvidia Blackwell 平台



資料來源：公司資料

圖 5：Blackwell GPU 特色



資料來源：公司資料

圖 6：Blackwell 對比 Hopper 提高 5 倍 AI 表現



資料來源：公司資料

圖 7: GB200 超級晶片



資料來源：公司資料

圖 8: 由兩個 Blackwell GPU 和 1 個 Grace CPU 組成



資料來源：公司資料

圖 9: Blackwell 計算節點包含 2 個 CPU、4 個 GPU



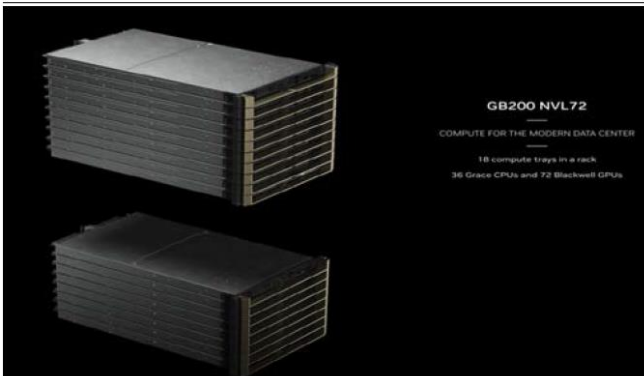
資料來源：公司資料

圖 10: GB200 NVL72 單一機櫃算力達 1.4 exaFLOPS



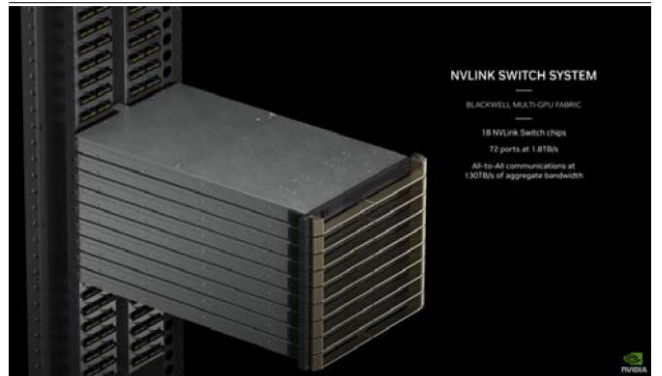
資料來源：公司資料

圖 11: GB200 NVL72 由 18 個 Blackwell 計算節點組成



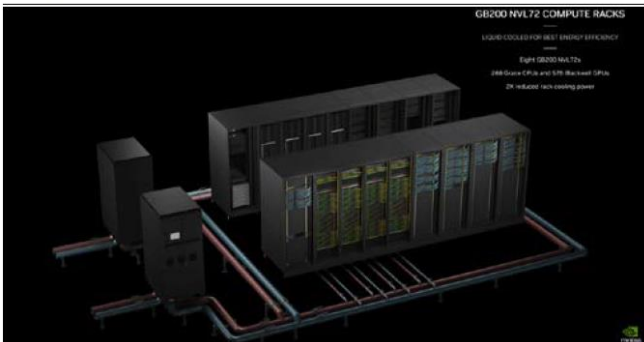
資料來源：公司資料

圖 12: GB200 NVL72 由 9 個交換器節點組成(18 個交換晶片)



資料來源：公司資料

圖 13: 8 個 GB200 NVL72 組成 DGX SuperPod，水冷系統將提供最優效率



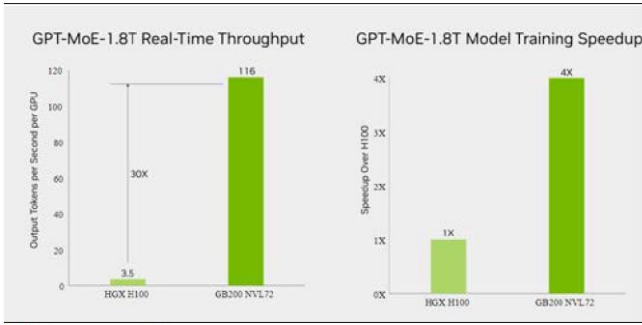
資料來源：公司資料

圖 14: 整個資料中心完整部屬可達 32000 個 GPU



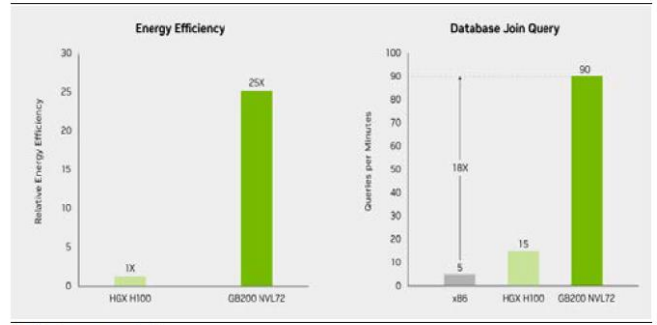
資料來源：公司資料

圖 15: GB200 NVL72 在推論與訓練表現上均大幅提升



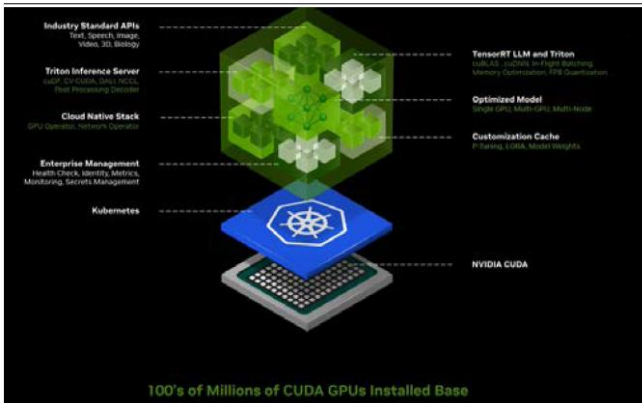
資料來源：公司資料

圖 16: GB200 NVL72 在能耗與資料處理表現上均大幅提升



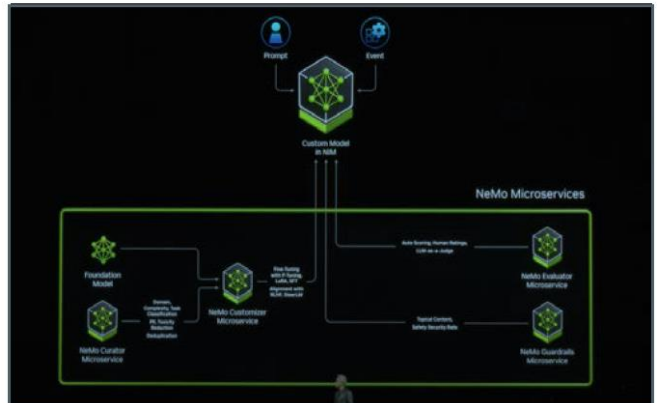
資料來源：公司資料

圖 17: Nvidia NIM (Nvidia Inference Microservice)



資料來源：公司資料

圖 18: Nvidia NeMo 微服務



資料來源：公司資料

圖 19 : GB200 與 H100 系列規格比較

	GB200 Superchip	GH200 Superchip	H200 SXM	H100 SXM	H100 PCIe
Configuration	1 Grace CPU : 2 Blackwell GPUs	1 Grace CPU : 1 Hopper GPU	1 Hopper GPU	1 Hopper GPU	1 Hopper GPU
FP64 Tensor Core	90 teraFLOPS	67 teraFLOPS	67 teraFLOPS	67 teraFLOPS	51 teraFLOPS
TF32 Tensor Core	5,000 teraFLOPS	989 teraFLOPS	989 teraFLOPS	989 teraFLOPS	756 teraFLOPS
BFLOAT16 Tensor Core	10,000 teraFLOPS	1,979 teraFLOPS	1,979 teraFLOPS	1,979 teraFLOPS	1,513 teraFLOPS
FP16 Tensor Core	10,000 teraFLOPS	1,979 teraFLOPS	1,979 teraFLOPS	1,979 teraFLOPS	1,513 teraFLOPS
FP8 Tensor Core	20,000 teraFLOPS	3,958 teraFLOPS	3,958 teraFLOPS	3,958 teraFLOPS	3,026 teraFLOPS
INT8 Tensor Core	20,000 teraFLOPS	3,958 TOPS	3,958 TOPS	3,958 TOPS	3,026 TOPS
FP4 Tensor Core	40,000 teraFLOPS				
GPU Memory	Up to 384 GB HBM3e	Up to 96GB HBM3 Up to 144 GB HBM3e	HBM3e 141GB	HBM3 80GB	HBM3 80GB
GPU Memory Bandwidth	16 TB/s	Up to 4TB/s Up to 4.9TB/s HBM3e	4.8TB/s	3.35TB/s	2TB/s
Max Thermal Design Power (TDP)		Programmable from 450W to 1000W (CPU+GPU+memory)	Up to 700W (configurable)	Up to 700W (configurable)	300-350W (configurable)
Form Factor	Superchip module	Superchip module	SXM	SXM	PCIe dual-slot air-cooled
Interconnect	NVLink: 3,600GB/s	NVIDIA NVLink: 900GB/s	NVIDIA NVLink: 900GB/s PCIe Gen5: 128GB/s	NVIDIA NVLink: 900GB/s PCIe Gen5: 128GB/s	NVLink: 600GB/s PCIe Gen5: 128GB/s
Server Options			NVIDIA HGX™ H200 partner and NVIDIA-Certified Systems™ with 4 or 8 GPUs	NVIDIA HGX H100 Partner and NVIDIA-Certified Systems™ with 4 or 8 GPUs NVIDIA DGX H100 with 8 GPUs	Partner and NVIDIA-Certified Systems with 1-8 GPUs

資料來源：公司資料、凱基



圖 20: H100/GB200 供應商比較

GPU型號	GB200	H100
平台	MGX	HGX
主要零組件		
主要供應商		
GPU 模組 (OAM)	鴻海 (工業富聯)	鴻海 (工業富聯)
GPU基板 (UBB)	X	緯創 (主要) 鴻海 (工業富聯)
NVLink 交換板	緯創	
主板	英業達	英業達
	緯創	緯創 神達 廣達
散熱		鴻海 (工業富聯)
	- 水冷 Cooler Master CoolIT 奇鋹 雙鴻 Vertiv (CDU) CPC, Staubli (manifold)	- 氣冷 (3D VC) Cooler Master 奇鋹
伺服器 (L10)	廣達	廣達
機櫃 (L11)	鴻海 (工業富聯)	鴻海 (工業富聯)
	緯創集團 (緯穎)	緯創集團 (緯穎)
	ZT System	ZT System
		Supemicro 技嘉 神達 Dell 華碩 Oracle

資料來源：公司資料、凱基

上述為證監會持牌人，隸屬凱基證券亞洲有限公司從事相關受規管活動，其及 / 或其有聯繫者並無擁有上述有關建議股份，發行人及 / 或新上市申請人之財務權益。

**免責聲明** 部份凱基亞洲有限公司股票研究報告及盈利預測可透過 [www.kgi.com.hk](http://www.kgi.com.hk) 取閱。詳情請聯絡凱基客戶服務代表。本報告的資料及意見乃源於凱基證券亞洲有限公司的內部研究活動。本報告內的資料及意見，凱基證券亞洲有限公司不會就其公正性、準確性、完整性及正確性作出任何申述或保證。本報告所載的資料及意見如有任何更改，本行并不另行通知。本行概不就因任何使用本報告或其內容而產生的任何損失承擔任何責任。本報告亦不存有招攬或邀約購買或出售證券及 / 或參與任何投資活動的意圖。本報告只供備閱，並不能在未經凱基證券亞洲有限公司書面同意下，擅自複印或發佈全部或部份內容。凱基集團成員公司或其聯屬人可提供服務予本文所提及之任何公司及該等公司之聯屬人。凱基集團成員公司、其聯屬人及其董事、高級職員及雇員可不時就本報告所涉及的任何證券持倉。