

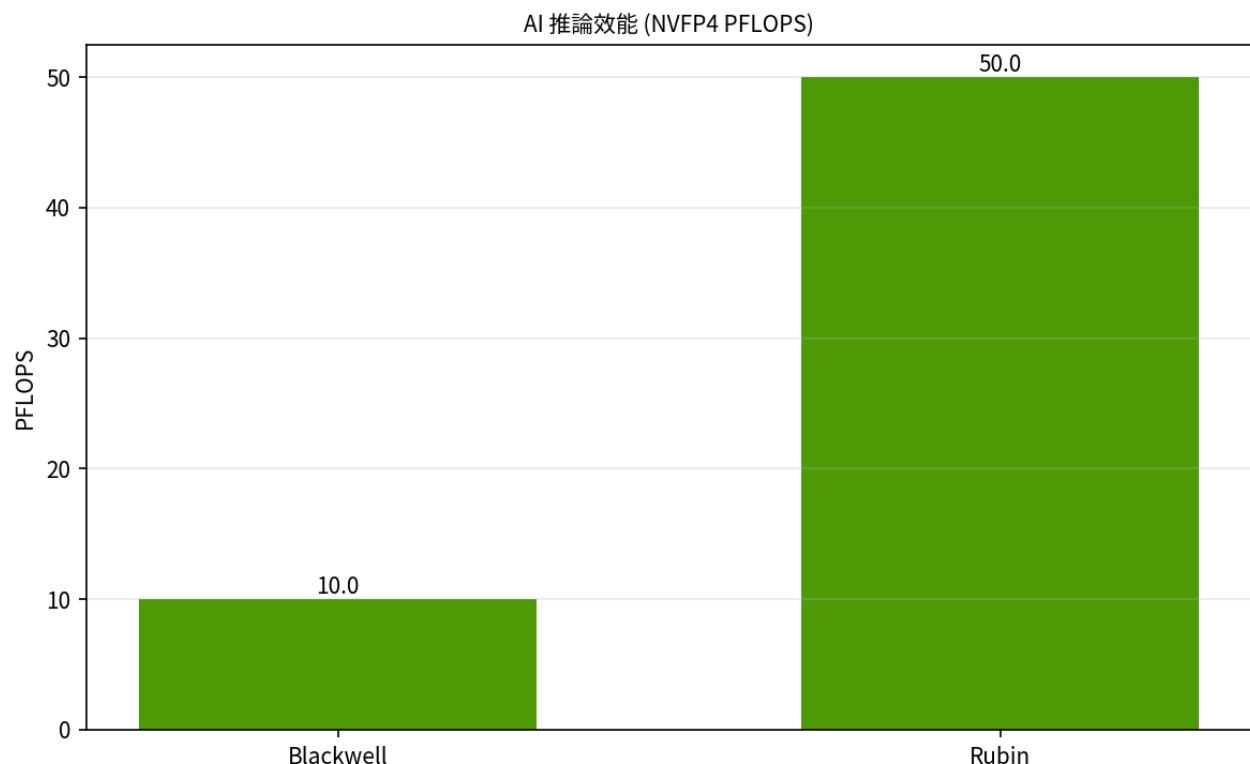
NVIDIA CES 2026 : Vera Rubin 平台深度解析

定義 AI 工廠 (Defining the AI Factory)

執行摘要：運算單元的典範轉移

NVIDIA 2026 年推出的 Vera Rubin 平台重新定義了資料中心的基本單位：從「晶片」轉變為「機架」。Vera Rubin NVL72 不僅僅是一台伺服器，它是一台重達 2 公噸的超級電腦，運作起來就像一個單一的「運算單元 (Computing Unit)」。憑藉每顆 GPU 50 PFLOPS 的算力以及每秒傳輸 260 TB 數據的神經系統（相當於全球網際網路總流量的 2 倍以上），它完美解決了下一代推理模型面臨的「記憶體牆 (Memory Wall)」問題。

1. 運算層：Vera Rubin Compute Tray (運算盤)



基礎建構模組

「運算盤 (Compute Tray)」是機架的核心建構模組。每個 Tray 都是一個獨立的高性能節點，專為處理 AI 工廠的大規模資料編排與吞吐而設計。

運算盤規格參數

- 2 顆 Vera CPU：採用 NVIDIA 自研 Olympus 核心 (Arm 架構)。效能是 Grace CPU 的 2 倍。
 - 專為資料編排、程式碼編譯與壓縮而設計。
- 4 顆 Rubin GPU：Blackwell 的繼任者。
 - 單顆提供 50 Petaflops (NVFP4) 推論效能 (是 Blackwell 的 5 倍)。
- 1 顆 BlueField-4 DPU：整合 64 個 Grace 核心。
 - 負責 100% 的安全、存儲與管理工作，確保 GPU 算力專注於 AI。
- 8 顆 ConnectX-9 NIC：單埠 800G (每顆 GPU 總出口頻寬達 1.6 Tb/s)。
 - 為每顆 GPU 提供專屬的高速 Scale-out 通道，解決吞吐瓶頸。

2. 機架層：Vera Rubin NVLink72 Rack

巨型 GPU (The Giant GPU)

當 18 個運算盤與 9 個交換器盤組合在一起，便形成了 NVL72 Rack —— 這是一個由 72 顆 Rubin 晶片互連組成的「巨型 GPU」。

機架規格與「神經系統」

- 規模：18 個運算盤 (總計 72 顆 Rubin GPU) + 9 個 NVLink Switch Trays。
- 交換架構：NVLink 6 交換器盤 (封包交換、無阻塞架構)。
- 總頻寬：約 260 TB/s (All-to-All 雙向頻寬)。
- 對比：這比全球網際網路的總流量還要大 2 倍以上。
- 複雜度：220 兆個電晶體 (整台機架半導體重量接近 2 公噸)。
- 神經系統：這 9 個交換器盤就像機架的「大腦」，讓任一顆 GPU 都能以 3.6 TB/s 的速度與其他 GPU 直接對話。

3. 網路層：Spectrum-X Ethernet Photonics

Spectrum-X 與光子技術

為了將成千上萬個機架連接成一座「AI 工廠」，NVIDIA 引入了突破性的光子技術：

- 512 Lanes & 200 Gb SerDes：全球首款支援 200G 電氣信號的乙太網路交換器。
- CPO (Co-packaged Optics)：將光學引擎直接與交換器晶片封裝在一起。
- 捨棄傳統的可插拔式光模組。
- 功耗降低 3.5 倍。
- 可靠性提升 10 倍 (徹底解決 Link Flap 連線不穩問題)。

Rubin GPU Connectivity Triangulation

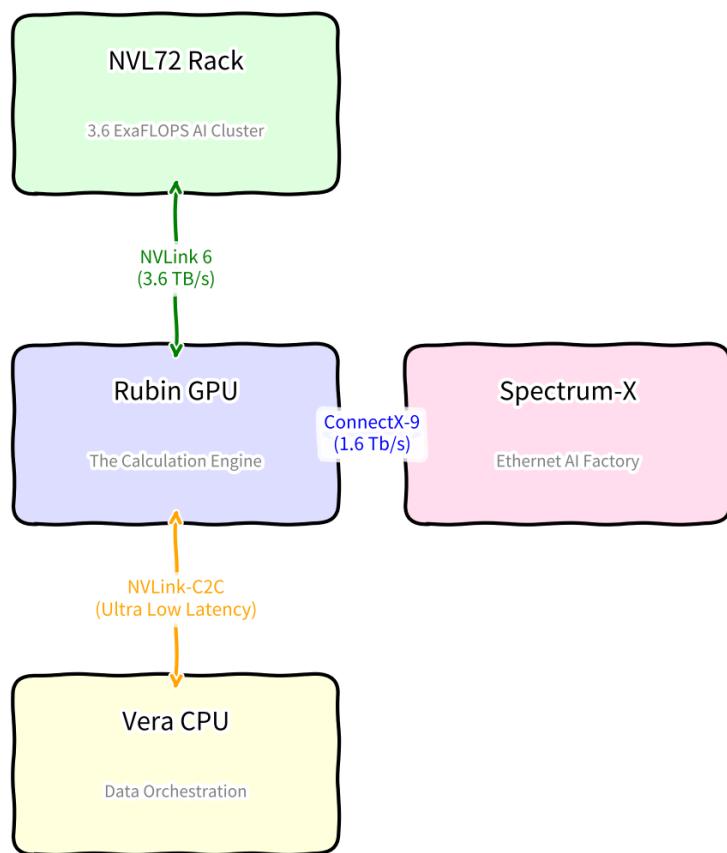


圖: Rubin GPU 連接三部曲 (Scale-Up, Scale-Out, Host)

4. 技術深度解析：為什麼需要 8 張網卡？

在一個 Tray 中配置 8 顆 ConnectX-9 NIC 對應 4 顆 Rubin GPU，絕非偶然，而是為了維持「運算」與「通訊」完美對稱的精密架構設計。

工程邏輯解析

1. 實現 1:1 的頻寬平衡：

- 每顆 Rubin GPU 需要 1.6 Tb/s 的對外頻寬。
- 透過 8 張網卡，每顆 GPU 分配到 2 個 800G 通道。
- 結果：每顆 GPU 都有專屬的高速出口，確保大規模擴充時不塞車。

2. 匹配 PCIe 6.0 的吞吐量：

- PCIe 6.0 x16 提供約 128 GB/s (接近 1024 Gb/s) 的頻寬。
- 必須使用 800G/1.6T 等級的網卡才能餵飽這種「線速 (Wire Speed)」。

3. 可靠性與多租戶隔離 (Multi-tenancy)：

- 冗餘設計：單張網卡故障不會導致整個 Tray 癱瘓。
- 流量隔離：物理上將「模型並行」與「數據並行」的流量分開處理。

5. 技術深度解析：Spectrum-X 的 512 通道與 32 光收發器之謎

要理解如何透過僅僅 32 個光收發器 (Optical Transceivers) 實現 512 條通道 (Lanes) 的吞吐量，關鍵在於「單一模組的通道密度」以及 NVIDIA 在「共封裝光學 (CPO)」技術上的突破。

表 1: Spectrum-X 核心頻寬數學 (The Core Math)

參數 (Parameter)	數值 (Value)	邏輯說明 (Logic)
總通道數 (Total Lanes)	512	機架對外的總吞吐通道
光收發器數量 (Transceivers)	32	透過 CPO 繫湊排列於 ASIC 四周
單模組通道數 (Lanes/Mod)	16	$512 \text{ Lanes} \div 32 \text{ Mods} = 16 \text{ Lanes}$
單通道速率 (Rate/Lane)	200 Gb/s	採用 PAM4 高速訊號技術
單模組頻寬 (BW/Mod)	3.2 Tb/s	$16 \text{ Lanes} \times 200 \text{ Gb/s} = 3.2 \text{ Tb/s}$
總交換容量 (Total Capacity)	102.4 Tb/s	$32 \text{ Mods} \times 3.2 \text{ Tb/s} = 102.4 \text{ Tb/s}$

Spectrum-X Architecture: 32 CPO Engines

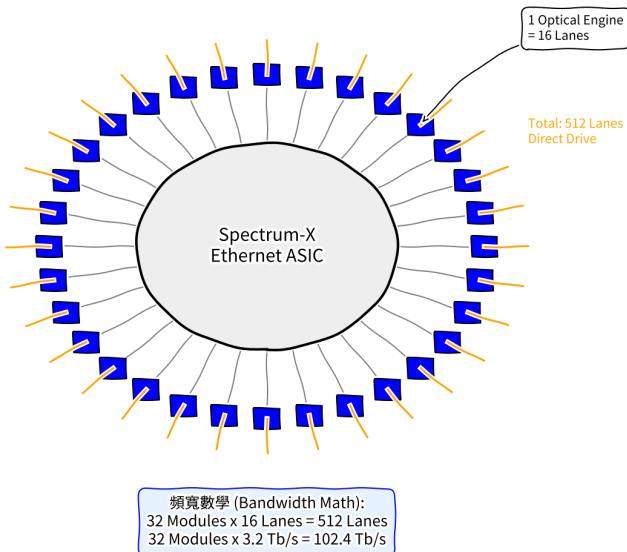


圖 2. Spectrum-X 32 CPO Engines 架構示意圖

技術解析：CPO 與架構優勢

1. 物理實現：從 Pluggable 到 CPO (Co-packaged Optics)

- 在 200G SerDes 的極高速率下，訊號在 PCB 上傳輸超過幾公分就會產生劇烈損耗。
- NVIDIA 將 32 個光引擎直接封裝在交換器晶片 (ASIC) 基板上，取代了傳統面板上密密麻麻的插拔式模組。
- 每個 3.2T 收發器直接對接 MPO/MTP 高密度光纖，實現光訊號直出。

2. 戰略優勢：為什麼要濃縮到 32 個？

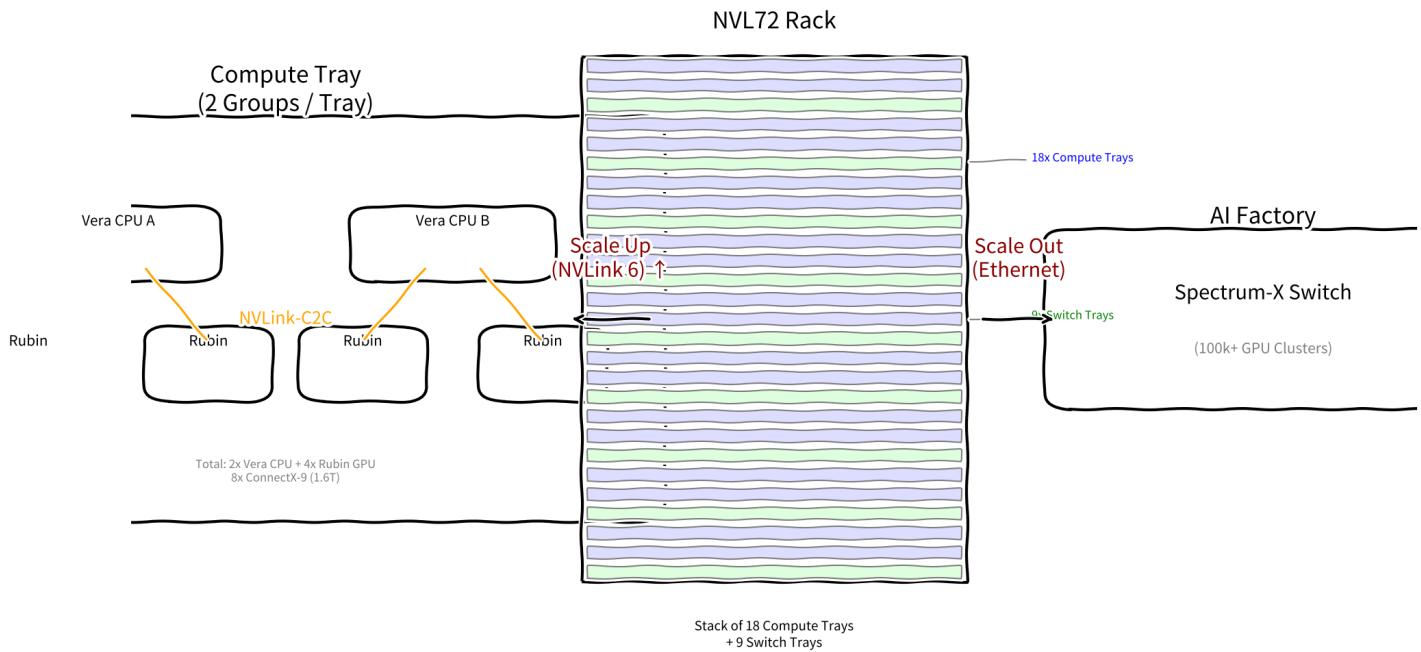
- 功耗降低：減少了電氣重驅動晶片 (Retimers)，節省約 40% 能源。
- 散熱空間：騰出空間給液冷散熱塊 (Cold Plates)，實現「無風扇」設計。
- 可靠性：大幅減少故障點 (Points of Failure)，對萬卡叢集的穩定性至關重要。

3. 與 ConnectX-9 的完美對接

- 每個 3.2T 光模組透過分支線纜 (Breakout Cable) 分成 2 個 1.6T 埠。
- 正好對接 2 張 ConnectX-9 網卡 (單張 1.6T)。
- 一台 Spectrum-X 交換器可穩穩連接 64 個 CX9 節點。

6. 從 Tray 到 AI Factory 的擴展 (Scaling from Tray to Factory)

Vera Rubin 擴展架構：從 Tray 到 Factory (Rack View)

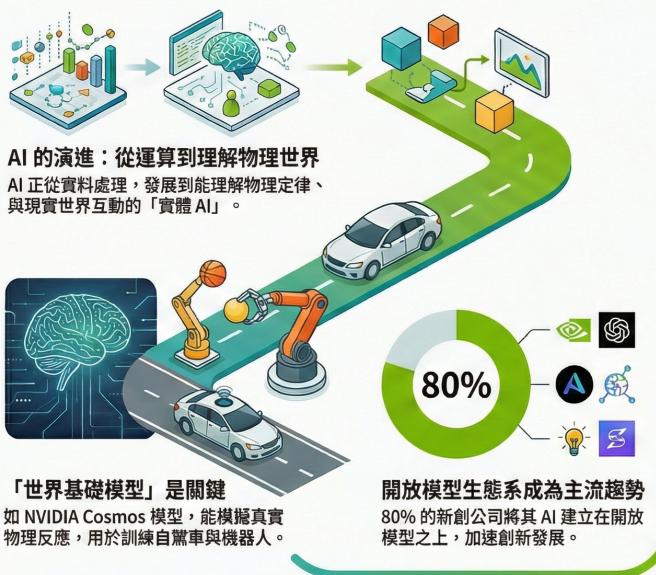


這張圖展示了 Vera Rubin 平台如何從單一個「運算盤 (Tray)」透過 NVLink 6 向上擴展 (Scale Up) 成一個機架，再透過 Spectrum-X 向外擴展 (Scale Out) 成擁數十萬顆 GPU 的 AI 工廠。

7. 系統架構圖

輝達 CES 2026 重磅發布：Vera Rubin 平台開啟實體 AI 新紀元

AI 的下一波浪潮：實體 AI



動力核心：全新 Vera Rubin 平台

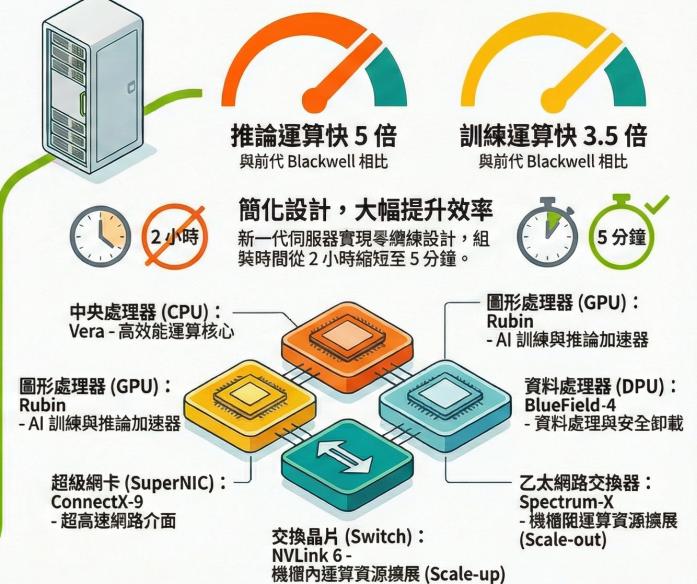


圖 1. NVIDIA AI Flow & Ecosystem Visualization (來源: CES 2026)

8. 世代演進：Vera Rubin (R100) vs Blackwell (B200)

從 Blackwell 到 Vera Rubin, NVIDIA 不僅僅是提升了數字，而是徹底改變了資料中心的物理型態。以下是兩代旗艦平台的終極規格對決：

表 2: 世代對決 (Generation Comparison)

指標 (Metric)	Blackwell (B200)	Vera Rubin (R100)	提升幅度 (Improvement)
AI 推論 (Inference)	10 PFLOPS	50 PFLOPS	5x (更強的思考能力)
HBM 頻寬 (Bandwidth)	8 TB/s	22 TB/s	2.8x (解決記憶體牆)
NVLink 頻寬 (Scale-Up)	1.8 TB/s	3.6 TB/s	2x (更快的叢集互連)
網路架構 (Network)	Plug. Optics	CPO (512 Lanes)	S-X 專用光子引擎
機架規模 (Rack Scale)	NVL72 Gen1	NVL72 Gen2	18 Trays 超高密度

架構質變總結

架構質變分析：

1. 從能耗到光子：Blackwell 仍依賴傳統電氣互連，而 Rubin 透過 Spectrum-X CPO 進入了「光子運算」時代。
2. 從訓練到思考：50 PFLOPS 的推論性能顯示 Rubin 專為 OpenAI o1/o3 等具備「長鏈思考 (Chain of Thought)」能力的模型而生。
3. 從單卡到機架：NVL72 Gen2 的出現，標誌著「機架即 GPU (Rack is the GPU)」的理念已完全成熟。

9. 訊號傳輸革命：SerDes 從 Blackwell 到 Rubin

AI 運算速度的瓶頸，最終回到了物理層。

當單通道速率從 Blackwell 的 100G (112G PAM4) 翻倍至 Rubin 的 200G (224G PAM4) 時，傳統銅線 (Copper) 與電路板 (PCB) 已觸及物理極限 (Skin Effect)，這就是為什麼 CPO 是唯一的出路。

表 3: SerDes 物理層對決 (The Physics Battle)

技術指標 (Spec)	Blackwell Era	Vera Rubin Era	物理影響 (Impact)
SerDes 速率	100G (112G PAM4)	200G (224G PAM4)	訊號衰減劇增 (Signal Loss)
傳輸介質 (Medium)	PCB / DAC 銅線	Glass / Fiber (CPO)	銅線無法支撐 200G 長距離
單埠速率 (Port)	800 Gb/s (8x100)	1.6 Tb/s (8x200)	介面密度翻倍
互連架構 (Interconnect)	Scale-Out (電氣)	Scale-Up (光子)	光學引擎取代電氣 Retimer
交換器晶片 (ASIC)	51.2T (Spectrum-4)	102.4T (Spectrum-X)	單晶片吞吐量翻倍

物理極限解析

為什麼 200G 是分水嶺？

- 趨膚效應 (Skin Effect)：在 224G 頻率下，電流僅在銅線極表層流動，阻抗急劇上升。
- 介電質損耗 (Dielectric Loss)：PCB 板材本身會吸收高頻訊號，導致訊號在幾吋內就衰減殆盡。
- 結論：Blackwell 是「電氣互連」的巔峰，而 Vera Rubin 是「光子互連」的元年。

10. 巔峰對決：NVIDIA Spectrum-X vs Broadcom Tomahawk 6

在 102.4 Tb/s 的頂級交換機領域，目前只有兩位真正的選手：NVIDIA Spectrum-X 與 Broadcom Tomahawk 6。這不僅是規格的較量，更是兩種不同世界觀的碰撞：「AI 垂直整合」對抗「開放通用晶片 (Merchant Silicon)」。

表 4: 王者之爭 (The Clash of Titans)

比較項目 (Item)	Broadcom Tomahawk 6	NVIDIA Spectrum-X	勝負關鍵 (Key)
交換容量 (Capacity)	102.4 Tb/s	102.4 Tb/s	平手 (Tie)
物理層 (SerDes)	224G PAM4	224G PAM4	平手 (均需 CPO)
設計哲學 (Philosophy)	開放生態 (Open)	垂直整合 (Vertical)	通用 vs 專用
目標市場 (Target)	雲端 / 白牌交換機	AI Factory / NVL72	廣度 vs 深度
軟體堆疊 (Software)	SAI / SONiC	NCCL / CUDA	相容性 vs 效能

戰略總結 (Strategic Conclusion)

戰略局勢分析：

1. Broadcom (AVGO)：是「軍火商」，提供最強的通用引擎 (Tomahawk) 級 Cisco, Arista 等各大品牌，主宰了 90% 的傳統資料中心與雲端網路。
2. NVIDIA (NVDA)：是「特種部隊」，Spectrum-X 是為了讓 GPU 跑得更快而存在的。它不在此乎通用性，只在乎 NCCL 通訊效率與 AI 叢集效能。
3. 結論：如果您要蓋通用雲端，選 Broadcom；如果您要蓋 AI 超級電腦 (AI Factory)，Spectrum-X 是唯一解。

11. 光子戰爭：Broadcom CPO vs NVIDIA CPO

CPO (Co-Packaged Optics) 是 200G 時代的關鍵技術，但 Broadcom 與 NVIDIA 選擇了截然不同的物理實作路徑。這是一場關於「標準化維修」與「極致整合密度」的哲學之戰。

表 5: 光子引擎對決 (The CPO Battle)

比較項目 (Item)	Broadcom (Davisson)	NVIDIA (Spectrum-X)	技術差異 (Diff)
核心引擎 (Engine)	Davisson (基於 Bailly)	Custom Spectrum CPO	迭代 vs 原生
光學技術 (Tech)	TSMC COUPE (SiPh)	TSMC COUPE + MRM	微環調變 (密度)
維護策略 (Service)	可插拔雷射 (Field Replaceable)	高度整合 (Integrated)	易修性 vs 密度
目標優化 (Goal)	良率與互通性 (Yield)	功耗與空間 (Power)	通用 vs 極致
生態系 (Ecosystem)	OIF 標準化 (Open)	NVLink 封閉優化	標準 vs 效能

光子架構總結 (Architectural Conclusion)

光子技術解析：

1. Broadcom 的 Davisson (Tomahawk 6 CPO)

強調「可維修性」，雷射模組可以像傳統光模組一樣更換，這對於傳統資料中心運維至關重要。

2. NVIDIA 則採用更激進的微環調變器 (MRMs) 技術，追求極致的「密度」與「每瓦效能」，因為在 NVL72 機架中，空間與電力比維修便利性更珍貴。

3. 這是「通用伺服器思維」與「超級電腦思維」的典型差異。

11.1 分析師觀點：SemiAnalysis 評析 (Analyst View)

知名半導體分析機構 SemiAnalysis 對這場「後端網路之戰 (Backend Network Battle)」提出了關鍵洞察：

SemiAnalysis 關鍵洞察

1. 成本結構 (Cost)：

- Merchant Silicon (Broadcom) 通常比 NVIDIA 的整體解決方案便宜約 50% (Per Port)。
- 這對大規模雲端業者 (Hyperscalers) 具有極大吸引力。

2. 效能與鎖定 (Performance vs Lock-in)：

- NVIDIA 宣稱 Spectrum-X 在 AI 負載下有 1.6x 的效能優勢 (透過 RoCE 優化)。
- 但 Broadcom 擁有更開放的生態系，且 Tomahawk 往往在上市時間 (Time-to-Market) 領先 6-9 個月。

3. 市場區隔：

- Broadcom 嬴在「標準化」與「多供應商」的 AI 叢集。
- NVIDIA 嬴在追求「極致單一效能」的垂直整合 AI Factory。

12. 記憶體革命：HBM4 vs HBM3e (The Memory Revolution)

Rubin 平台最大的躍進之一在於採用了第六代高頻寬記憶體 (HBM4)。這不僅僅是速度的提升，更是架構的徹底重構。

特徵 (Feature)	Blackwell (HBM3e)	Vera Rubin (HBM4)
介面頻寬 (Bus Width)	1024-bit	2048-bit (雙倍通道)
堆疊高度 (Stack Height)	8-Hi / 12-Hi	12-Hi / 16-Hi (更高密度)
邏輯底座 (Logic Die)	標準製程 (Standard)	客製化製程 (Custom 12nm/5nm)
單顆容量 (Capacity)	24GB / 36GB	48GB / 64GB+
功能整合 (Integration)	單純記憶體 (Storage)	記憶體內運算 (Near-Memory Compute)

HBM4 架構優勢

戰略意義 (Strategic Impact)：

1. 頻寬牆的突破：2048-bit 介面讓單顆 HBM 的傳輸速度翻倍，解決了 GPU 運算雖快但資料餵不進去的瓶頸。

2. 客製化邏輯層：HBM4 允許 NVIDIA

在記憶體底部加入自己的邏輯電路，這意味著簡單的運算可以直接在記憶體內完成，大幅減少資料搬移的功耗 (Data Movement Costs)。

13. AI 儲存關鍵：PCIe Gen 6 與 SSD (The Storage Backbone)

雖然 GPU 與 HBM Rubin 吸引了大部分目光，但將資料「餵進」這些巨獸的管道同樣關鍵。Vera 平台將見證 PCIe Gen 6 SSD 的全面導入。

特徵 (Feature)	Hopper/Blackwell (Gen 5)	Vera Rubin (Gen 6)
傳輸速率 (Transfer Rate)	32 GT/s	64 GT/s (兩倍速)
x4 通道頻寬 (Bandwidth)	~16 GB/s	~32 GB/s per SSD
Checkpointing	慢 (可能導致 GPU 閑置)	快 (降低訓練中斷時間)
資料載入 (Data Loading)	由 CPU 解壓	GPUDirect Storage (直通顯卡)
主要瓶頸 (Bottleneck)	IOPS 與 頻寬	散熱與功耗 (High Power)

PCIe Gen 6 SSD 的戰略價值

SSD 在 AI 中的兩大關鍵角色 (Critical Roles)：

1. Checkpointing (存檔)：訓練大模型每小時需將 TB 級的參數寫回 SSD 以防當機。PCIe Gen 6 將存檔時間減半，直接提升昂貴 GPU 的有效運算時間 (Effective Training Time)。
2. Data Caching (快取)：Vera CPU 透過 Gen 6 匯流排從 SSD 預取資料，確保 HBM4 永遠有資料可算，避免「GPU 等待 SSD」的 Data Starvation 慘劇。

14. 供應鏈版圖變遷 (Supply Chain Shifts)

Vera Rubin 的技術大躍進將重塑 AI
硬體供應鏈。從晶圓代工到散熱模組，舊有的平衡將被打破，新的贏家將會誕生。

領域 (Sector)	關鍵變革 (Change)	潛在受惠 (Beneficiaries)	風險/挑戰 (Risks)
晶圓代工 (Foundry)	N4 -> N3 (3nm)	TSMC (獨家)	產能與良率 (Capacity)
封裝 (Packaging)	CoWoS-L 面積倍增	TSMC, Amkor	基板短缺 (Substrate)
記憶體 (HBM)	HBM4 邏輯底座整合	TSMC (Logic Die), SK Hynix	標準化後的價格競爭
光通訊 (Optics)	可插拔 -> CPO 共封裝	Broadcom, TSMC (COUPE)	傳統模組廠 (Transceiver)
散熱 (Thermal)	氣冷 -> 液冷/浸沒式	Vertiv, CoolIT, 鴻海	機發建設 (Infrastructure)
載板 (PCB/CCL)	M8/M9 等級材料	Unimicron, 台光電 (EMC)	良率與翹曲 (Warpage)

CPO 帶來的產業衝擊

CPO 的破壞式創新 (The CPO Disruption) :

最值得關注的變革在於 CPO (Co-packaged Optics)。隨著 SerDes 速度達到物理極限，將光引擎直接封裝在 GPU 或 Switch 重心旁成為必然。這將導致傳統光模組廠 (Pluggable Optics) 面臨長期且巨大的轉型壓力，而掌握先進封裝 (TSMC) 與光晶片 (Broadcom/NVIDIA) 的巨頭將進一步整碗端走高階市場。

14.1 分析師觀點：SemiAnalysis 供應鏈評析 (Analyst View)

SemiAnalysis 對於 Vera Rubin 世代的供應鏈提出了兩大警訊與機會：

SemiAnalysis: 供應鏈權力重組

1. CoWoS 產能牆 (The CoWoS Wall) :

- 儘管 TSMC 積極擴產，但 Rubin 的封裝複雜度 (Reticle limit 突破) 將消耗更多 CoWoS 產能。
- 預測：CoWoS 仍將是 2026 年前的最大出貨瓶頸，而非 HBM 或 GPU 本身。

2. HBM4 的權力轉移 (Power Shift to Foundry) :

- HBM4 引入客製化邏輯底座 (Base Die) 意味著晶圓代工廠 (TSMC) 的角色將比記憶體原廠 (SK Hynix/Micron) 更關鍵。
- 價值鏈將從「純記憶體製造」轉向「先進封裝與邏輯整合」，TSMC 將成為最大贏家。

15. 關鍵規格與單位總表 (Key Specifications & Units)

名詞 (Term)	定義與換算 (Definition)	意義/場景 (Context)
PFLOPS PFLOPS	Peta-FLOPS (\$10^{15}\$ 次/秒) 1 PFLOPS = 1,000 TFLOPS	AI 算力的基本單位 (FP4/FP8/FP16)
ExaFLOPS ExaFLOPS	Exa-FLOPS (\$10^{18}\$ 次/秒) 1 ExaFLOPS = 1,000 PFLOPS	下一代 AI Factory 或超級電腦算力等級
Tb/s vs TB/s Tb/s vs TB/s	1 Byte (B) = 8 bits (b) 1 TB/s = 8 Tb/s	Tb/s: 網路頻寬 (Network) TB/s: 記憶體頻寬 (Memory)
HBM (3e/4) HBM (3e/4)	High Bandwidth Memory (高頻寬記憶體) 堆疊式 DRAM 架構	決定 AI 模型的「推論速度」與「參數量」
Radix Radix	Port Count (交換器埠數) 例如: Radix 512 = 512 個連接埠	決定單層交換器能連接多少 GPU (擴展性)
SerDes SerDes	Serializer/Deserializer (串列/解串器) 例如: 224G SerDes	物理層訊號速度，銅纜傳輸的極限與瓶頸
CPO CPO	Co-packaged Optics (光電共封裝) 將光引擎封裝在切換晶片旁	NVL72 骨幹技術，解決長距離高耗電問題
NVLink vs Ethernet NVLink vs Ethernet	私有協議 (Scale-Up) vs 標準協議 (Scale-Out)	NVLink: 像多核 CPU 內部的 Bus Ethernet: 像電腦之間的 LAN
Scale-Up Scale-Up	垂直擴展 (增加單一節點算力)	NVL72 機架內，GPU 像在同一個主機板上
Scale-Out Scale-Out	水平擴展 (增加節點數量)	透過 Spectrum-X 交換器連接數萬張 GPU