

Vera Rubin NVL72 整機櫃技術與供應鏈報告

Computex 2026 展示觀察、GB300 比較、ODM 供應鏈推測與投資檢核表

版本：2026-05-20

適用場景：Computex 展前研究、ODM 展示判讀、AI 伺服器供應鏈追蹤、投資觀察清單

核心資料：NVIDIA 官方 Vera Rubin NVL72 / DGX Vera Rubin NVL72 規格、NVIDIA 技術部落格、SemiAnalysis Vera Rubin Extreme Co-Design、鴻海 / Wiwynn / ASUS GTC 2026 公開展示資訊、既有 GB200/GB300、液冷、電源、連接器與光通訊供應鏈資料庫整理

目錄

1. 報告摘要
2. 為什麼 Computex 要看 Vera Rubin 整機櫃
3. Vera Rubin NVL72 核心規格
4. Vera Rubin 機櫃構造：從 rack 到 tray
5. 與 GB300 / Blackwell Ultra NVL72 的比較
6. 五個改變最多、需求最特別的地方
7. 光通訊 / CPO / AEC 擴充分析
8. Computex 展示應該怎麼看
9. ODM 廠商逐家推測
10. 既有資料庫資訊整合
11. 可能出現在整機櫃的供應鏈產品
12. 台股個股觀察名單：機會、驗證訊號與風險
13. Computex 展後投資觀察重點
14. 風險與反證
15. 來源與延伸閱讀

1. 報告摘要

一句話結論： Vera Rubin NVL72 不是單純把 GB300 換成下一代 GPU，而是 NVIDIA 進一步把 AI 伺服器機櫃標準化成「rack-scale AI factory module」。Computex 若有 ODM 展示整機櫃，重點應放在 cableless compute tray、100% 液冷、50V / 800VDC 電力架構、NVLink 6、1.6T scale-out network、L10 自動化組裝與 L11/L12 rack delivery，而不是只看外殼或 GPU 名稱。

Vera Rubin NVL72 的核心定位是把整個 rack 當作一個可部署、可維修、可監控、可量產的 AI 超級運算單元。NVIDIA 官方公開規格顯示，單一 Vera Rubin NVL72 rack 包含 72 顆 Rubin GPU、36 顆 Vera CPU、ConnectX-9 SuperNIC、BlueField-4 DPU、NVLink 6 Switch，並以 Quantum-X800 InfiniBand 或 Spectrum-X Ethernet 做 scale-out。整櫃 NVFP4 inference 最高 3,600 PFLOPS，HBM4 總容量 20.7TB，GPU HBM4 頻寬最高 1,580TB/s，CPU 端 LPDDR5X 容量 54TB。

相對 GB300，Vera Rubin 最大差異不在單點算力，而在系統工程：compute tray 從線纜式設計走向模組化 cableless；液冷從 GPU / 部分 tray 走向 100% full liquid cooled；機櫃功耗從約 120-140kW 推升到約 180-220kW；NVLink 6 讓 rack 內 scale-up bandwidth 翻倍；每 GPU scale-out bandwidth 走向 1.6T。這些改變會讓 ODM 的角色、散熱供應鏈、電源供應鏈、PCB / 連接器供應鏈重新分配價值。

Computex 觀察重點可以用一句話整理：**誰展示 full rack，誰展示可量產的 tray automation，誰展示電熱基礎設施，誰只是展示 reference mockup。**這四者的投資意義差很多。

2. 為什麼 Computex 要看 Vera Rubin 整機櫃

Computex 對台灣 AI 伺服器供應鏈的意義，不只是新產品發表，而是把平常藏在客戶認證、量產良率、資料中心 acceptance test 裡面的資訊，以展品、partner wall、模組標籤和簡報語言露出。Vera Rubin 又特別適合在 Computex 被展示，原因有三個。

2.1 台灣 ODM 是 rack-scale AI server 的主要實作端

鴻海、廣達 / QCT、緯創 / 緯穎過去在 GB200 / GB300 rack 放量中已有實際角色。SemiAnalysis 指出 Rubin compute tray L10 自動化能力目前集中在 Foxconn、Quanta、Wistron 三家；Wiwynn 則在 GTC 2026 公開表示與 Wistron 共同展示 Vera Rubin NVL72 AI factory infrastructure readiness。這代表 Computex 上若有 VR NVL72 展品，台灣供應鏈不是旁觀者，而是量產與交付端的主角。

2.2 Vera Rubin 的差異在整機櫃，不在單片 GPU

若只看 GPU，Rubin 是 3nm、HBM4、NVFP4、adaptive compression、單 GPU 高功耗的延續。但從投資與供應鏈角度，真正的增量在整櫃工程：midplane、B2B connector、高階 PCB、liquid-cooled busbar、power shelf、internal manifold、MQD / UQD、CDU、1.6T networking、L10 automation。這些都是 Computex 展場可能看到的硬體。

2.3 NVIDIA 對系統控制力更高，ODM 的差異化要重新定義

GB300 時代 hyperscaler 在前半部 chassis、networking、DPU、power delivery、local storage、management module 上有較多客製化。Rubin 因 cableless modular tray 與 midplane reference design，客製化空間被壓縮到特定模組外形內。ODM 的競爭重點會轉到：能不能拿到 L10 自動化、能不能穩定做 blind mate、能不能整合 liquid cooling 與資料中心電力、能不能完成 L11/L12 rack acceptance。

3. Vera Rubin NVL72 核心規格

項目	Vera Rubin NVL72 公開規格 / 目前推估	供應鏈意義
GPU / CPU	72 顆 Rubín GPU、36 顆 Vera CPU	每個 Vera Rubin Superchip 為 2 GPU + 1 CPU；CPU / GPU / memory / network 共同設計
AI 算力	NVFP4 inference 3,600 PFLOPS；NVFP4 training 2,520 PFLOPS	NVIDIA 主打推論成本與 MoE training efficiency，而非只賣 peak FLOPS
GPU memory	20.7TB HBM4；總頻寬最高 1,580TB/s	HBM4 pin speed、HBM4 良率、HBM4 供應商份額會影響 ramp
CPU memory	54TB LPDDR5X；每 Vera CPU 最高 1.5TB	SOCAMM / LPDDR5X 模組、CPU 旁高速記憶體連接器成為新題材
Scale-up	NVLink bandwidth 260TB/s / rack；3.6TB/s / GPU	NVLink 6 Switch、NVLink backplane、M8+ PCB、液冷 switch board 需求升級
Scale-out	ConnectX-9 SuperNIC；每 GPU 1.6Tb/s 級 scale-out	OSFP cage、1.6T optical / AEC、CPO switch、InfiniBand / Ethernet 架構都要追
Cooling	100% liquid cooled；45°C liquid cooling envelope 為官方與 OCP 論述重點	Cold plate、manifold、QD、CDU、pump、dry cooler / adiabatic tower 需求提升
Power	SemiAnalysis 推估 rack TDP 約 180–220kW；reference design 使用 4 x 110kW power shelves	50V busbar、liquid-cooled busbar、800VDC power rack、BBU / CBU 進入主線
製造	Cable-free modular tray；assembly / service time 大幅下降	L10 自動化組裝、blind mate、loading mechanism、field serviceability 成為 ODM 差異化

注意：官方性能數字多為 up to / projected specification，且 NVIDIA 可能依客戶、power profile、networking choice、DPU / storage module 配置提供不同 SKU。投資研究應把「NVIDIA 官方 reference design」和「各 hyperscaler 實際部署 BOM」分開。

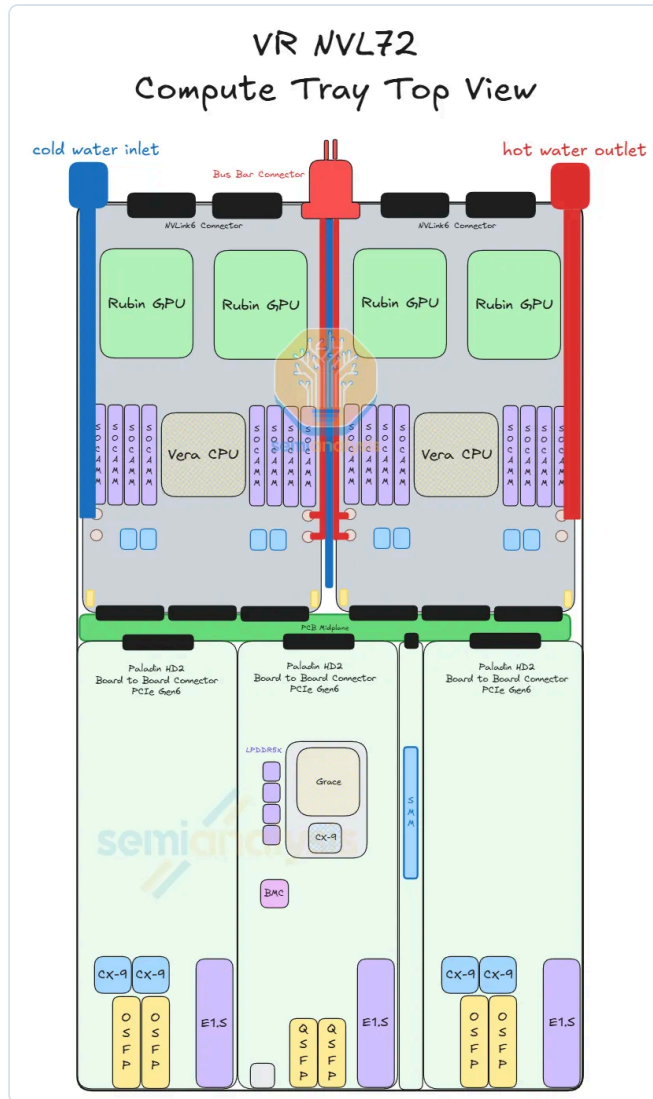
4. Vera Rubin 機櫃構造：從 rack 到 tray

4.1 Rack 層級：一櫃就是一個分散式加速器

Vera Rubin NVL72 延續 Oberon / MGX NVL72 rack-scale 路線，單一 rack 內部做 scale-up，使 72 顆 GPU 在軟體與網路上接近一個大型加速器。這是 NVIDIA 對抗 AMD Helios、AWS Trainium UltraServer、Google TPU pod 等 rack-scale solution 的核心策略：不是只提供 GPU，而是提供完整 rack 與 cluster building block。

從供應鏈角度，這意味著採購品項不再是「GPU board + server chassis」而是「rack + tray + switch + networking + cooling + power + management」。整櫃交付的難度提高，但 reference design 的標準化也讓 NVIDIA 更能壓縮 deployment time，降低客戶在不同 ODM 之間客製化造成的良率風險。

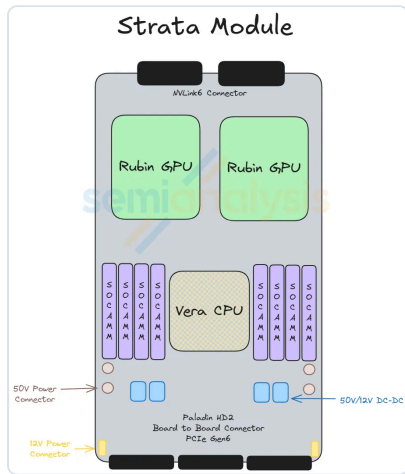
4.2 Compute tray：六類模組構成 Rubin 的最小展示單位



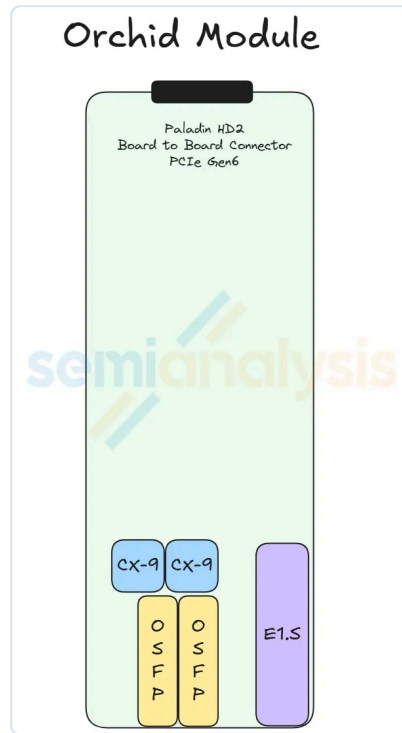
VR NVL72 compute tray 模組配置。來源：SemiAnalysis / NVIDIA VR NVL72 BoM and Power Budget Model，已保存於 wiki attachment。

一個 VR NVL72 compute tray 可拆成六類模組：

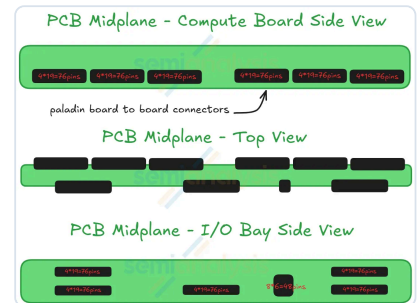
模組	位置 / 功能	Computex 觀察點
Strata module x2	後半部核心計算板；每片含 2 顆 Rubin GPU、1 顆 Vera CPU、SOCAMM 記憶體	是否展示 SOCAMM、GPU cold plate、50V 進板、NVLink spine connector
Orchid module x4	前端窄長模組；放 CX-9 NIC、OSFP cage、E1.S storage slot	是否展示前端 NIC 也被液冷、OSFP 數量、1.6T network 配置
Compute tray midplane	垂直置中的橋接板，連接 Strata 與前端模組	是否強調 cableless、B2B connector、PCIe Gen6 over PCB、M8/M9 CCL
Power delivery module	接收 50V 並配送 12V 到前端模組	是否展示 PDB、IBC、power brick、busbar、液冷 power 元件
BlueField-4 module	NVIDIA reference DPU；也可能被 hyperscaler 自研 DPU 取代	是否有 BF4 或自研 DPU slot；是否提 ICMS / KV cache / storage fabric
System management module	SMM / TPM / DC-SCM 等管理與安全功能	大型雲端客戶是否自研 management module；是否與 PDB 合併



Strata module : 2 GPU + 1 Vera CPU + SOCAMM。

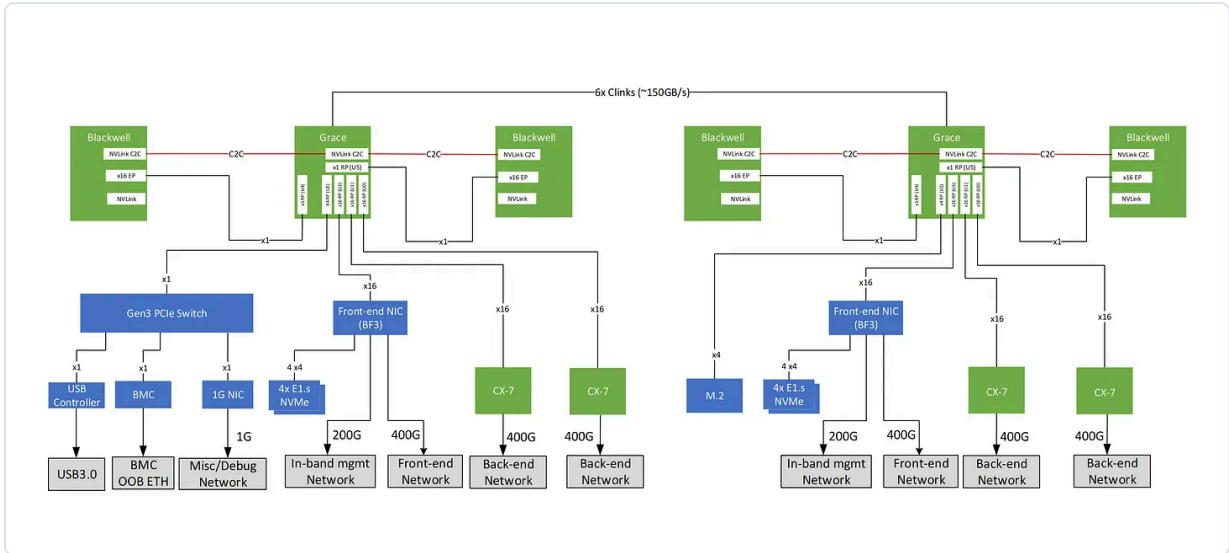


Orchid module : CX-9、OSFP、E1.S 前端化。



Midplane : cableless tray 的訊號橋。

4.3 Topology：為什麼 Rubin 可以 cableless

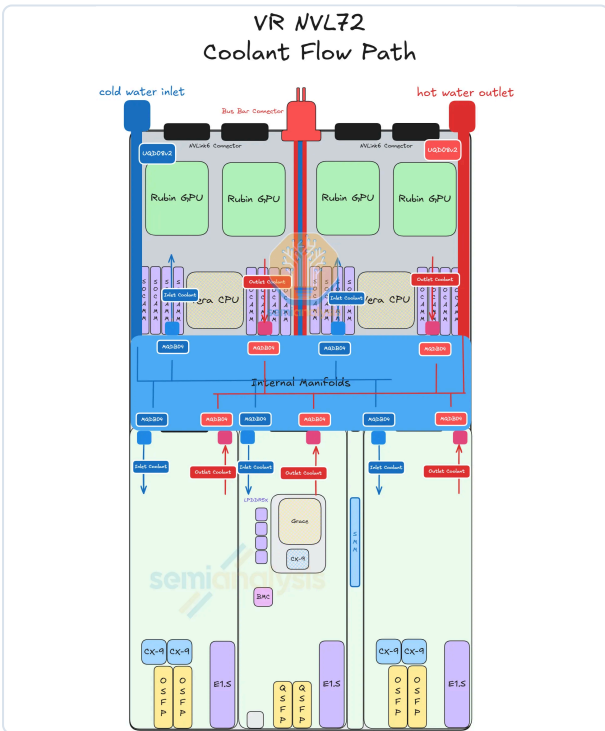


Rubin compute tray topology：CX-9 移到前端，使高速網路訊號縮短，PCIe Gen6 走較長 PCB 路徑。

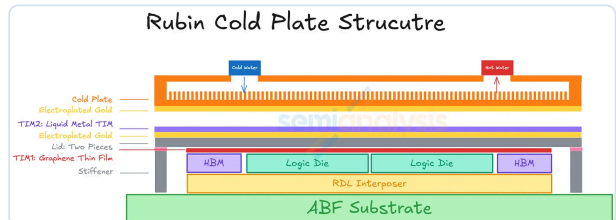
GB200 / GB300 的挑戰是 NIC 通常位於 compute tray 後半部，而 OSFP cage 在前端，高速 200G Ethernet / InfiniBand lane 走長距離 PCB 的 insertion loss 過高，因此需要 flyover cable。Rubin 的設計把 CX-9 NIC 移到前端 Orchid module，讓 200G lane 縮短；從 Strata 到 Orchid 的較長路徑則改由 PCIe Gen6 承擔。PCIe Gen6 單 lane 速率低於 200G Ethernet / InfiniBand，因此更有機會在高階 PCB 上完成長距離傳輸。

這個設計的本質不是「少了一條線」這麼簡單，而是價值轉移：原本的 flyover cable、cable routing、人工裝配風險，轉為 midplane、高階 CCL、HVLP4 銅箔、B2B connector、盲插機構、L10 automation 的需求。

4.4 Liquid cooling：整個 tray 被納入冷卻迴路



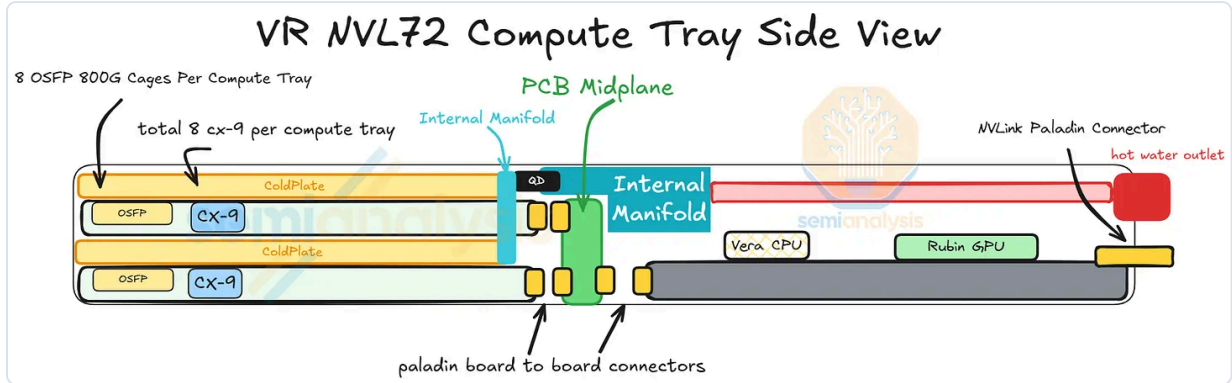
Compute tray 100% liquid cooled 架構。



冷卻液由 UQD 進出，經 internal manifold 分配。

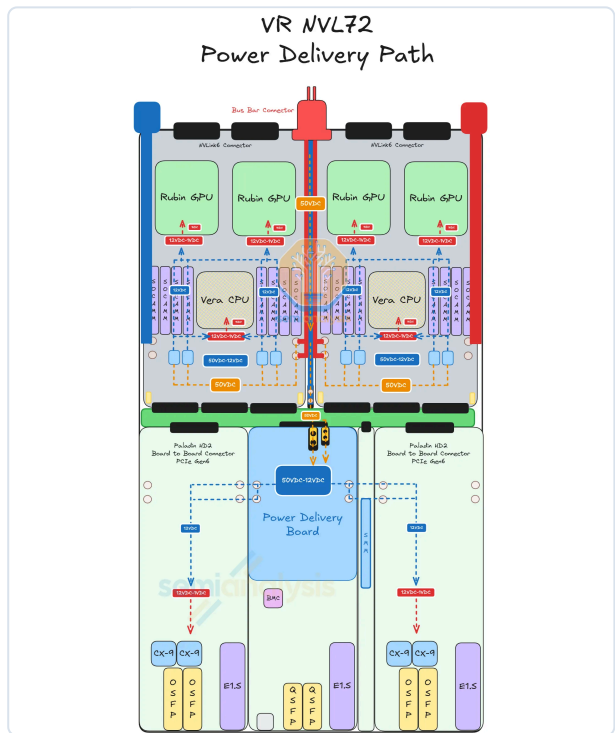
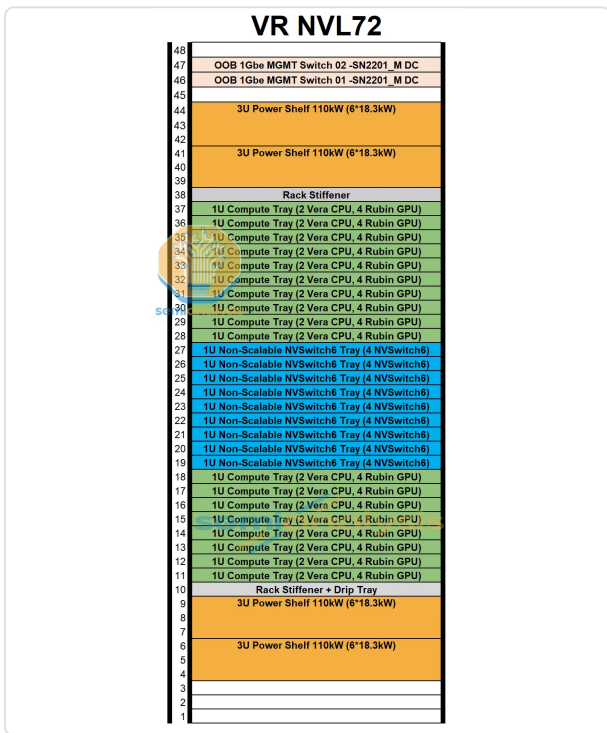
Rubin compute tray 的冷卻液由左後側 UQD 進入 tray，進到 internal manifold 後分配到不同模組，再回到 manifold，最後從右後側 UQD 流出。每個 module 上方都有 cold plate module；Strata 的冷板覆蓋 Rubin GPU、Vera CPU、SOCAMM、VRM 等元件，Orchid 也覆蓋 CX-9、E1.S SSD、OSFP cage 與 VRM。

Rubin GPU 冷板升級到 micro-channel cold plate，通道 pitch 從 Blackwell 相關世代的約 150 微米縮到約 100 微米，以提高散熱面積與熱移除能力。GPU 接觸面因液態金屬 TIM2 腐蝕風險，需要鍍金處理。這讓 cold plate 從單純銅加工品，變成微流道、表面處理、可靠度與 assembly 良率共同決定的高規格元件。



Rubin GPU micro-channel cold plate：通道變密、鍍金表面、對應液態金屬 TIM2。

4.5 Power：50V rack busbar 走向液冷化



Rack level power：四組 110kW power shelves，N+1 支援 220kW 級 rack。

Compute tray power：50V 進 Strata / PDB，再轉 12V / sub-1V。

Rubin rack power reference design 使用四組 110kW power shelves，每組 3U，內含多顆 18.3kW PSU。Power shelves 將三相 415-480VAC 轉成 50VDC，再送入 rack busbar。因整櫃功耗提高且 fanless 架構減少氣流，busbar 額定電流上升到 5,000A+，並需要 liquid-cooled busbar。

Compute tray 內，50V 直接進入 Strata board，經 IBC 轉為 12V，再由 VRM 轉成 GPU / CPU 所需的 sub-1V。這比 GB300 / Grace Blackwell 時代 12V 進 Bianca board 更有效率，原因是同樣功率下提高電壓可降低電流，電阻損耗按電流平方上升。

5. 與 GB300 / Blackwell Ultra NVL72 的比較

項目	GB300 / Blackwell Ultra NVL72	Vera Rubin NVL72	判讀
系統定位	Blackwell Ultra rack-scale AI server，客製化空間仍大	第三代 MGX NVL72 rack，NVIDIA reference control 更高	ODM 從 design house 轉向高良率製造與交付
Compute tray	Bianca board + 多處 cable / connector	Strata + Orchid + midplane，cableless modular tray	B2B connector / midplane / high-end PCB 受惠
Cooling	約 85% liquid + 15% air / hybrid	100% liquid cooled / fanless	冷板覆蓋範圍擴大到 NIC、SSD、VRM、management module
Rack power	約 120-140kW 級	約 180-220kW 級	50V busbar、liquid-cooled busbar、power rack、HVDC 議題更重要
GPU TDP	約 1,000-1,400W 級	1,800W Max-Q / 2,300W Max-P power profile	客戶可能依電力成本與資料中心容量選擇功耗 profile
CPU / memory	Grace CPU + LPDDR / platform memory	Vera CPU + SOCAMM，整櫃 54TB LPDDR5X	SOCAMM 連接器、LPDDR5X module 成為新 BOM
Scale-up	NVLink 5；18 顆 switch ASIC / rack	NVLink 6；36 顆 switch ASIC / rack	Switch tray 數量 / 冷板 / PCB / SMM 增加
Scale-out	CX-8 / 800G 級	CX-9，每 GPU 1.6T 級 scale-out	1.6T optics / AEC / CPO / OSFP cage 觀察價值提高
Assembly	多 cable route，組裝耗時且失效點多	模組盲插，官方 / 供應鏈口徑指向大幅縮短組裝與維修時間	L10 automation 是 Rubin ODM 競爭核心

GB300 到 Rubin 不是只有 GPU 世代替換，而是從「高功耗機櫃」進一步走到「電熱訊號共同設計」。這會改變投資權重：過去看 rack 出貨量與 GPU allocation，Rubin 要加看 automation、power delivery、liquid cooling、PCB / connector、scale-out networking。

6. 五個改變最多、需求最特別的地方

6.1 Cableless tray：線纜下降，但 connector / PCB / automation 上升

Rubin 最具代表性的設計是 cable-free modular tray。這不是取消互連，而是把原本依賴 flyover cable 的訊號路徑，改為 B2B connector + midplane + 高階 PCB。NVIDIA 這麼做的理由很直接：GB200 / GB300 的 cable routing 太難、失效率太高、裝配耗時太長；Rubin 若要大規模鋪進 AI factory，tray 必須像模組一樣可自動化插裝與維修。

供應鏈含義：

- 對傳統 tray 內 cable assembly 不利，但對高密度 B2B connector、有壓接技術的連接器、midplane PCB、高階 CCL、HVLP 銅箔有利。
- 對 ODM 而言，差異化不只是採購零件，而是 blind mate 良率、connector mating force、機構公差、loading mechanism 與維修流程。
- 台股映射中，SOCAMM2 / CABB 題材以優群（3217）最直接；但大型國際 connector 如 Amphenol 仍可能是主要 share。

6.2 100% liquid cooling：冷卻從 GPU 擴大到整個 tray 與 facility

Rubin 的液冷增量不只是 GPU TDP 提高，而是冷卻範圍從後端 GPU / CPU 擴大到前端 NIC、OSFP、E1.S SSD、VRM、management module。這意味著冷卻液在 tray 內必須經過更多分支，QD / MQD / UQD、manifold、leak detection、pump、cold plate module 都更重要。

供應鏈含義：

- 奇鋌 (3017) 的優勢在 total thermal solution、cold plate、inner / rack manifold、整機散熱整合。
- 高力 (8996) 的優勢在 RDHx、in-row CDU、PHE、row-level / facility-level heat exchange。
- 如果 Computex 展示是 tray 內部冷板 / manifold，奇鋌邏輯較強；若展示 CDU / facility loop / chiller-less，則高力、台達、Schneider、Vertiv 等邏輯提高。

6.3 Power：從伺服器 PSU 走向 rack power 與資料中心電力架構

Rubin rack 180–220kW 級功耗，使電源設計不能只看 PSU。Power shelf、50V busbar、liquid-cooled busbar、power distribution board、IBC、VRM、BBU、CBU、HVDC power rack 都會被放到前台。未來 megawatt-class rack 路線也會讓 800VDC / +/-400VDC 的 OCP / hyperscaler power architecture 更重要。

Computex 若有 ODM 展示 power rack、BBU、CBU、liquid-cooled busbar，這比單純說「支援 Rubin」更值得追。因為 GPU rack 的瓶頸很可能從 GPU allocation 轉向 power availability、transient load、grid interaction 與 rack-level power stability。

6.4 Network：NVLink 6 與 1.6T scale-out 使 cluster readiness 成為差異

Rubin rack 內部由 NVLink 6 scale-up；rack 外則是 ConnectX-9 / Quantum-X800 / Spectrum-X 或 hyperscaler 自有 Ethernet fabric。每 GPU 1.6T scale-out bandwidth 代表 OSFP cage、1.6T transceiver、AEC、CPO switch、multi-plane network architecture 都會成為展示亮點。

如果 ODM 展示只有單櫃外觀，資訊含量有限；如果它同時展示 rear cabling、OSFP、Spectrum-X / Quantum-X800、CPO TOR、Storage-Next / KV cache fabric，代表它在 AI factory cluster 層級更成熟。

6.5 ODM 自由度下降：L10 / L11 / L12 分工會重估估值

Rubin reference design 讓 NVIDIA 更接近交付完整系統。這對客戶有利，因為 deployment time 下降、良率與維修流程更可控；但對 ODM 毛利可能不利，因為高價值設計與 BOM 控制權被 NVIDIA 拿走更多。ODM 的價值會分成三層：

- L6 PCBA：誰能做高階 board PCBA，含 Strata / Orchid / NVSwitch board。
- L10 compute tray：誰能自動化組 tray，處理冷板、QD、B2B connector、loading mechanism。
- L11 / L12 rack delivery：誰能把 rack、cooling、network、power、storage、data center acceptance 串起來。

這也是為什麼鴻海、廣達、緯創 / 緯穎要分層看，不應只看「誰展示一台機櫃」。

7. 光通訊 / CPO / AEC 擴充分析

光通訊結論：Vera Rubin NVL72 的光通訊主軸不是「全面 CPO 取代 pluggable」，而是 1.6T scale-out 進入主戰場，同時讓 InfiniBand、Spectrum-X Ethernet、非 NVIDIA Ethernet、AEC、pluggable

optics、CPO switch 多條路線並行。Computex 看到 CPO 要拆清楚：是 Rubin NVL72 的 scale-out TOR 試行，還是 Rubin Ultra / Spectrum-X / Quantum-X 後續平台的中長期故事。

7.1 先分三層：scale-up、scale-out、scale-across

Goldman Sachs 2026-04-17 AI 光網路報告將 AI 資料中心網路分成三層：scale-up 是同一 computing unit / rack 內互連，scale-out 是 rack 與 rack / cluster 之間互連，scale-across 是跨資料中心互連。Vera Rubin NVL72 目前最重要的是前兩層。

層級	Vera Rubin NVL72 對應	主要介質	Computex 看點
Scale-up	Rack 內 72 GPU 經 NVLink 6 switch all-to-all	銅背板 / NVLink cable / high-speed PCB ; NVLink 6 雙向 SerDes	NVLink 6 switch tray、backplane、M8+ PCB、液冷 switch board
Scale-out	每 GPU 約 1.6T 對外頻寬；常見為 2 x 800G OSFP	800G / 1.6T optical、AEC、CPO switch、InfiniBand / Ethernet fabric	OSFP 數量、Quantum-X800、Spectrum-X、CPO TOR、1.6T AEC
Scale-across	跨資料中心 / 遠距傳輸	DWDM、coherent、長距光傳輸	較不會是 Rubin rack 本體展品，偏資料中心網路方案

Connection assumptions		Vera Rubin Spec A	Vera Rubin Spec B	Rubin Ultra Spec A	Rubin Ultra Spec B	Remarks
Model	GB300 NVL72	Vera Rubin NVL 72	Vera Rubin NVL 72 (25% CPO scale out)	Rubin Ultra NVL 144 (29% CPO scale out)	Rubin Ultra NVL 576 (29% CPO scale out)	
Shipment period	2H25-2026	2H26-2027	2H26-2027	2H27-2028	2H27-2028	
Scale up	Copper cable	Copper cable	Copper cable	PCB Midplane	Copper cable + CPO	
Scale out	Optical modules (1.6T)	Optical modules (1.6T)	CPO TOR Switch (1.6T) + Optical module (1.6T)	CPO TOR Switch (3.2T) + Optical module (3.2T)	CPO TOR Switch (3.2T) + Optical module (3.2T)	
# of GPU package	72	72	72	144	72 *	* For Rubin Ultra NVL576, we show per rack
# of NV Switch ASIC	18	36	36	72	108	data in the table, despite it's a 8-rack scale up
# of NIC	72	144	144	288	144	system
Scale up						
Scale up Bandwidth (unidirectional, Tb/s)	518	1,037	1,037	2,074	2,074	
Total Switching Capacity (Tb/s)	1,037	2,074	2,074	4,147	6,221	
Copper cable backplane bandwidth (Tb/s)	518	1,037	1,037	-	2,074	
Notes	100% copper cable	100% copper cable	100% copper cable	-	L1 scale up within rack	Total bandwidth = # of copper cables x Bandwidth of each cable / 2 (unidirectional)
# of copper cables (units)	5,184	5,184	5,184	-	5,184	
Bandwidth of each copper cable (Gb/s)	200	200	200	-	400	
PCB midplane bandwidth (Tb/s)	-	-	-	2,074	-	
Notes	-	-	-	PCB all to all connection	-	Total bandwidth = # of PCB Midplane x Bandwidth of each PCB
# of PCB Midplane (units)	-	-	-	2	-	
Bandwidth of each PCB (Gb/s)	-	-	-	1,036,800	-	
CPO/ NPO bandwidth (Tb/s)	-	-	-	-	1,037	
Notes	-	-	-	-	L2 scale up among racks	
# of optical fiber (units)	-	-	-	-	5,184	Total bandwidth = # of copper cables x Bandwidth of each cable / 2 (unidirectional)
Bandwidth of each optical fiber (Gb/s)	-	-	-	-	400	Total bandwidth = # of optical engine x Bandwidth of each optical engine
# of optical engine (units)	-	-	-	-	324	
Data rate of Optical engine (Tb/s)	-	-	-	-	3.2	
Scale out						
Scale out Bandwidth (unidirectional, Tb/s)	58	115	115	230	115	
Total Switching Capacity (Tb/s)	346	691	691	1,382	691	Assuming 3-layer network
GPU uplink (Tb/s)	58	115	115	230	115	GPU uplink = # of GPU per rack x GPU uplink speed (800G/ 1.6T/ 3.2T)
# of GPU (units)	72	72	72	144	72	
GPU uplink per chip (Tb/s)	0.8	1.6	1.6	1.6	1.6	
# of layer	3	3	3	3	3	
Optical module						
# of optical module (units)	216	432	324	307	153	# of optical module = Total switch capacity in scale-out network x % in optical module / Data rate of each optical module (1.6T/ 3.2T)
Optical module penetration (GSe)	100%	100%	75%	71%	71%	
Data rate of optical module (Tb/s)	1.6	1.6	1.6	3.2	3.2	
Attach rate (Optical module ports: GPU)	3.0	6.0	4.5	2.1	2.1	
CPO/ NPO						
# of CPO ports (units)	-	-	108	125	63	# of CPO ports = Total switch capacity in scale-out network x % in CPO / Data rate of each CPO port (1.6T/ 3.2T)
CPO penetration (GSe)	-	-	25%	29%	29%	
Data rate of CPO (Tb/s)	-	-	1.6	3.2	3.2	
Attach rate (CPO ports: GPU)	-	-	1.5	0.9	0.9	
Among which:						
# of optical engine (units)	-	-	108	125	63	# of CPO optical engine = Total switch capacity in scale-out network x % in CPO / Data rate each optical engine (3.2T)
Data rate of Optical engine (Tb/s)	-	-	1.6	3.2	3.2	
Attach rate (Optical engine: GPU)	-	-	1.5	0.9	0.9	
Attach rate summary						
Optical module : GPU - scale out	3.0	6.0	4.5	2.1	2.1	
Optical engine : GPU - scale up	-	-	-	-	4.5	
Optical engine : GPU - scale out	-	-	1.5	0.9	0.9	
Scale out sum	3.0	6.0	6.0	3.0	3.0	

資料庫附件：GS 依 NVIDIA roadmap 推估 GB300、Vera Rubin、Rubin Ultra 不同 scale-up / scale-out 連接方式。

7.2 Vera Rubin NVL72 的 scale-up：NVLink 6 仍以銅為主

Rubin rack 內部 scale-up 仍以 NVLink 6 銅互連為主，不是 rack 內全面光化。SemiAnalysis 指出 NVLink 6 的頻寬翻倍主要靠 simultaneous bi-directional SerDes，而不是把銅纜數量翻倍。這一點很重要：若沒有雙向 SerDes，Blackwell 世代約數千條 backplane copper cable 的複雜度可能進一步暴增；Rubin 透過雙向 SerDes 在不大幅增加 differential pair 數量的情況下提高頻寬。

供應鏈上，這代表 scale-up 受惠的不是傳統光模組，而是 NVSwitch board、M8+ CCL、低損耗 PCB、Paladin HD2 類 connector、液冷 switch tray、backplane 設計與 L10 裝配良率。Computex 若有 ODM 展示 Rubin switch tray，應記錄是否標示 NVLink 6、switch tray 冷板、SMM、backplane connector。

7.3 Vera Rubin NVL72 的 scale-out：2 x 800G 比單一 1.6T 更關鍵

SemiAnalysis 指出，VR NVL72 每個 compute tray 有 8 顆 800G CX-9 NIC，部署上有兩種可能：每 GPU 一個 1.6T OSFP cage，或每 GPU 兩個 800G OSFP cage。其 base case 偏向後者，因為兩個 800G logical ports 可以把同一 GPU 分到兩個 leaf switch / 兩個 planes，簡化多平面網路部署。

這對 Computex 判讀很有用：如果展品前面板或後端標籤看得到 OSFP cage 數量，應記錄是 4 cage / tray 還是 8 cage / tray。8 cage / tray 通常代表更適合 multi-plane scale-out network，對大型 cluster 可擴展性較佳。

配置	意義	供應鏈影響
1 x 1.6T logical port / GPU	布線較簡化，但 cluster size / plane split 較受限	1.6T transceiver 或 1.6T AEC 需求集中
2 x 800G logical ports / GPU	可拆到兩個 switch planes，擴大 cluster scaling	OSFP cage、800G optics / AEC、fiber management 複雜度提升

7.4 InfiniBand：Quantum X800 Q3400 / Q3450

Neocloud 與部分 NVIDIA full-stack 客戶仍可能偏好 InfiniBand。Vera Rubin 的 InfiniBand scale-out 可用 Quantum X800-Q3400 pluggable optical switch，也可用 Quantum X800-Q3450 CPO switch。Q3450 的特色是把多個 switch ASIC 與 optical engines 包進 switch box，等效高 radix / multi-plane 架構，並降低外部 fiber shuffle 的複雜度。

投資含義不是「Q3450 一出，所有 pluggable 都消失」。CPO switch 的優點是降低 transceiver power、減少 transceiver 成本、改善 link reliability、把 fiber shuffle 整合進 box；但相對整個 AI cluster power / capex，server rack 的功耗與成本仍佔大頭。因此 CPO 在 Rubin NVL72 可能是高階 / 大型 cluster / neocloud 的選項，並非所有部署的唯一答案。

7.5 Ethernet：Spectrum-X、SN6600 / SN6800 / SN6810 與非 NVIDIA Ethernet

NVIDIA 在 Rubin 世代也積極推 Spectrum-X Ethernet。SemiAnalysis 對 Vera Rubin 架構的整理指出，Spectrum-6 / Spectrum-X 相關 switch 有非 CPO pluggable 版本，也有 CPO 版本，例如 SN6800 以 4 ASIC / 409.6T multi-plane CPO switch 形態簡化部署，SN6810 則是 102.4T CPO switch，有不同 port configuration。

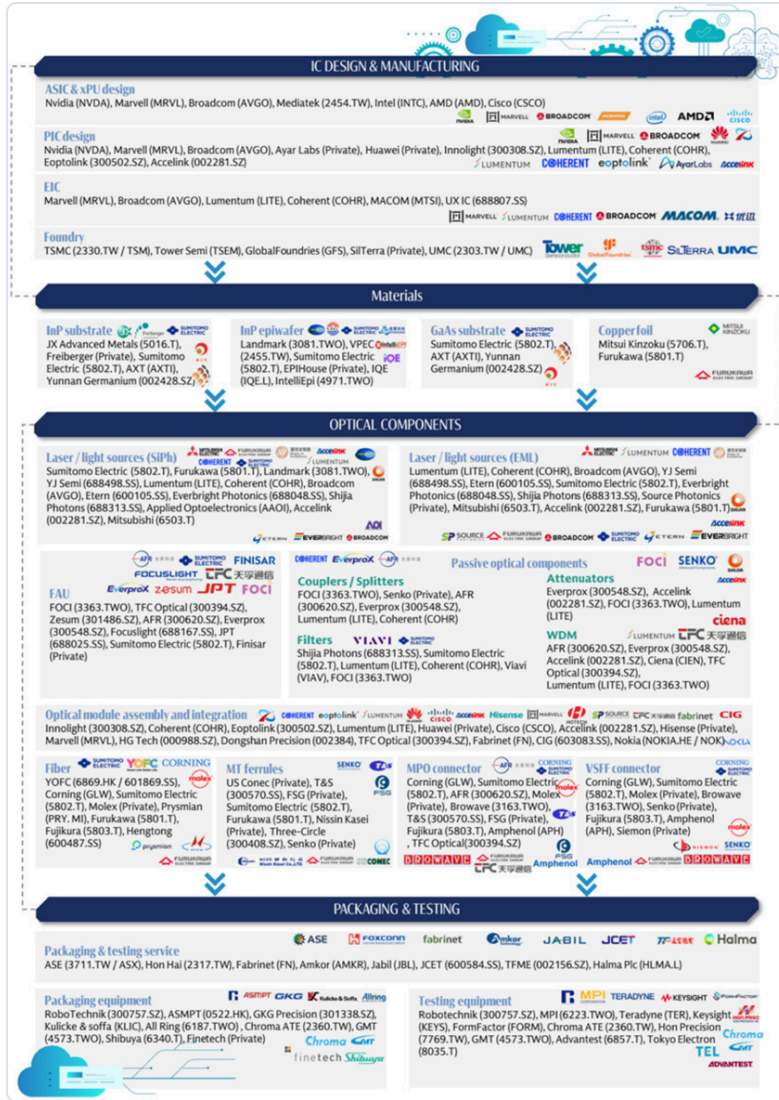
但大型 hyperscaler 未必全走 NVIDIA Ethernet。Meta 類客戶可能走 Tomahawk 6-based Minipack OCP Rack switch，並用 1.6T AEC 做 NIC-to-TOR；xAI 也可能使用 1.6T AEC 取代大量 transceiver。這使 Credo / AEC 供應鏈在 Vera Rubin 世代有重要位置。

路線	可能客戶 / 場景	Computex 展示訊號	供應鏈影響
InfiniBand + pluggable optics	Neocloud / NVIDIA full-stack cluster	Quantum X800-Q3400、800G OSFP、InfiniBand rail	800G / 1.6T 光模組、OSFP、光纖連接
InfiniBand + CPO	大型 cluster、想降低 transceiver / fiber shuffle 複雜度	Quantum X800-Q3450 CPO、OE / ELS 說明	CPO OE、ELS / CW laser、FAU / 光纖耦合
NVIDIA Ethernet + Spectrum-X	NVIDIA Ethernet cluster / enterprise AI factory	SN6600 / SN6800 / SN6810、Spectrum-X Photonics	CPO switch、pluggable optics、switch cooling
非 NVIDIA Ethernet + AEC	大型 hyperscaler 自研網路，如 Meta / xAI 類型	Tomahawk / OCP Minipack、1.6T AEC、gearbox	AEC、retimer / gearbox、Credo 類高速連接供應鏈

7.6 GS 光網路 roadmap：Vera Rubin 是 CPO 起步，Rubin Ultra 才是大放量

既有光通訊資料整理的重要結論是：GB300 → Vera Rubin → Rubin Ultra 的升級，同時推升 pluggable optical module、CPO optical engine、EML / CW Laser、PCB midplane、低損耗 PCB / CCL。GS 估算從 GB300 NVL72 到 Rubin Ultra NVL576，每 computing unit 的 scale-out / scale-up dollar content 分別提升 16x / 45x；整體 scale up + scale out value TAM 從約 150 億美元提升至約 1,540 億美元，其中 CPO 約 910 億美元、占 59%。

平台 / 期間	Scale-up	Scale-out	投資判讀
GB300 NVL72 (2H25-2026)	Copper cable	1.6T optical modules	1.6T pluggable 是主軸
Vera Rubin NVL72 (2H26-2027)	Copper cable / NVLink 6	1.6T optical modules，部分情境 CPO	CPO 開始進入 scale-out TOR，但不是全面替代
Rubin Ultra NVL144 (2H27-2028)	PCB midplane	3.2T CPO TOR + 3.2T optical modules	PCB midplane 與 CPO dollar content 明顯上升
Rubin Ultra NVL576 (2H27-2028)	Copper cable + CPO	3.2T CPO TOR + 3.2T optical modules	八 rack computing unit 使 scale-up 第二層走向光學化



GS optical supply chain 整理圖：涵蓋光模組、CPO optical engine、雷射、SiPh foundry / package、連接器、PCB / CCL。

7.7 台灣光通訊供應鏈映射

環節	台灣觀察標的	與 Vera Rubin / Rubin Ultra 的關係
1.6T pluggable optical module	華星光、光聖	Vera Rubin NVL72 scale-out 仍以 1.6T optical / OSFP 為重要主軸
CPO optical engine / SiPh 封裝	上詮、光聖、天虹	Rubin NVL72 開始試行；Rubin Ultra / Spectrum-X / Quantum-X 更重要
EML / CW Laser / Pump Laser	聯亞、全新	CPO 與外部光源 ELS 需求提升；供給緊張可能延續至 2027
FAU / D-FAU / iFAU	上詮、采鈺、Himax、合聖 / 光聖	CPO 光學最後一哩；對位容差、WLO、Meta-lens 是技術差異
PCB midplane / Paddle Card	欣興、台光電、南電	Rubin Ultra 對 PCB midplane / 低損耗小型載板需求更強；Vera Rubin 已有高階 PCB 預熱
Micro LED CPO / Intra-rack	銓創、富采、群創 / bEMC	2026-2027 偏規格 / 送樣；2028H2 後才可能形成較明顯出貨
CPO 測試 / active alignment	穎崴、旺矽	高 pin、高功耗、大封裝、光電對位使 socket / CPO insertion 設備成為量產瓶頸

7.8 Computex 光通訊現場 checklist

- OSFP cage 數量**：每 GPU 是 1 x 1.6T 還是 2 x 800G？這會影響 multi-plane network 判讀。
- Switch type**：Quantum-X800、Spectrum-X、Tomahawk / OCP switch、Cisco / Broadcom 是否出現。
- CPO 是哪一層**：是 TOR / leaf switch CPO，還是 rack 內 scale-up 光化？Vera Rubin NVL72 多數仍是 scale-out CPO。
- AEC 是否展示**：1.6T AEC 若出現，代表 hyperscaler Ethernet / cost-optimized route 很重要。
- ELS / CW Laser / optical engine 供應商**：若 partner wall 出現 Lumentum、Coherent、Broadcom、Marvell、AyarLabs 等，代表 CPO story 較完整。
- 台廠小零件**：FAU、D-FAU、micro lens、paddle card、MT ferrule、active alignment、socket / probe 是否被提到。
- 時間軸**：Vera Rubin NVL72 的 2026-2027 CPO 試行，不要和 Rubin Ultra 2027-2028 的 3.2T CPO 放量混淆。

8. Computex 展示應該怎麼看

現場看到什麼	代表意義	供應鏈推論
完整 Vera Rubin NVL72 rack	具備 NVIDIA reference rack 展示能力	仍要確認是 mockup、engineering sample、或可量產 platform
抽出 compute tray，展示 Strata / Orchid / midplane	真正切到 Rubin 核心結構	對 L10 automation、B2B connector、PCB / CCL 供應鏈較有意義
標示 blind mate、serviceable module、assembly time	展品重點在量產與維修，不只是外觀	對 Foxconn / Quanta / Wistron 類 L10 自動化能力加分
展示 100% liquid-cooled tray	前端 NIC / storage / VRM 也納入液冷	Cold plate、internal manifold、QD / MQD、leak detection 受惠
展示 CDU、rack manifold、45°C inlet、chiller-less	從 server 展示走到 facility integration	高力、台達、Schneider、Vertiv、nVent、SPX / BAC / Evapco 等觀察價值提高
展示 power rack、800VDC、BBU / CBU	AI data center 電力基礎設施升級	PSU、busbar、power shelf、power module、儲能與電力管理供應鏈受惠
展示 1.6T AEC、CPO TOR、Quantum-X800 / Spectrum-X	具備 cluster network readiness	光通訊、AEC、CPO、switch、OSFP cage 供應鏈應同步追蹤
展示 Storage-Next / ICMS / KV cache storage	切入 agentic AI 長上下文推論	Wiwynn / storage partners / BlueField-4 / DPU 生態值得追

9. ODM 廠商逐家推測

信心分級：**高** = 已有公开展示 / 官方新聞 / SemiAnalysis 指名；**中** = 有 GB200 / GB300 歷史與產品能力支持，但 Rubin full rack 尚需現場確認；**低** = 主要來自去年展場脈絡或產品線推測。

廠商	Computex 可能展示	可能強調的產品	可能使用 / 搭配供應鏈	信心
鴻海 / Foxconn	Vera Rubin NVL72 full-system rack、MDC 模組化資料中心、垂直整合零組件	L10/L11 rack、機構、散熱元件、電力傳輸、SOCAMM / connector、power / busbar	NVIDIA reference design + Foxconn 垂直整合；connector 可能含 Foxconn / Amphenol / SOCAMM；液冷可能自製加外部 TCS partner	高
廣達 / QCT	VR NVL72 或 GB300 → Rubin transition rack、CoolRack 液冷方案、QuantaGrid 系列	Rack integration、liquid cooling、customer deployment case	NVIDIA reference design + QCT rack / cooling；冷板 / CDU 可能依客戶採奇鋹、高力、台達、CoolIT 等，不宜硬判	中-高
緯創 / Wistron	L6 PCBA / L10 tray automation 能力、與緯穎共同 AI factory solution	Board-level PCBA、compute tray、automation assembly	SemiAnalysis 指 Wistron 為 L6 / L10 核心；可能和緯穎分工，Wistron 偏製造與 tray	高
緯穎 / Wiwynn	Vera Rubin NVL72 AI factory infrastructure、HGX Rubin NVL8、Storage-Next / GPU-initiated storage、液冷 rack	L11/L12 rack delivery、liquid cooling validation、storage / telemetry / leak detection	與 Wistron 共同開發；更可能強調 cloud CSP deployment、rack integration、cooling / serviceability，而非單純 L10 tray	高
ASUS	ASUS AI POD on Vera Rubin NVL72、enterprise AI factory / edge-to-data-center solution	Full liquid-cooled AI POD、storage、enterprise deployment	NVIDIA reference + ASUS rack-scale integration；偏 enterprise / regional AI factory，不一定是 hyperscaler 大量 L10 主力	中
Gigabyte	Vera Rubin NVL72 / HGX Rubin NVL8 資料中心產品、GIGAPOD / liquid cooling	NVIDIA MGX / HGX portfolio、server / rack solution	已有 Vera Rubin NVL72 文件；full rack 是否現場展示需確認。供應鏈多半走 NVIDIA reference + 自家 server platform + 外部 cooling / power partner	中
英業達 / Inventec	高密度 AI server / MGX / liquid-cooled rack concept	Rack / server integration、可能接 NVIDIA MGX ecosystem	去年已有 AI server 展示脈絡，但 Rubin full rack 公開證據較少；若展示，較可能是 reference rack 或 HGX Rubin NVL8 延伸	低-中
和碩 / Pegatron	AI server / open rack / liquid cooling concept	Rack integration、edge / enterprise AI	去年有高密度機架 / 開放架構脈絡；Rubin full rack 證據不足，需現場確認	低
Supermicro / Dell / HPE / Lenovo	海外 OEM / enterprise channel Vera Rubin rack 或 HGX Rubin platform	Turnkey enterprise AI infrastructure、service / deployment package	NVIDIA 官方 ecosystem 成員；台股鏈可從 cooling / PCB / connector / PSU 間接受惠	中

8.1 鴻海：垂直整合與 MDC 是最大敘事

鴻海已於 GTC 2026 公開展出 NVIDIA Vera Rubin NVL72 full-system AI server rack，並強調垂直整合、機構設計、散熱元件、電力傳輸與模組化資料中心 MDC。Computex 若延續這個展示，關鍵不只是 rack 本體，而是鴻海是否把 AI server、power、cooling、robotics、MDC 串成「AI factory industrialization」的故事。推測供應鏈：鴻海在 connector、機構、組裝、電力傳輸可能更多自製 / 垂直整合；但高規格連接器、液冷元件、CDU、pump、QD 仍可能有國際供應商與台廠協作。若現場有 partner wall，應優先拍下 cooling / power / connector partner。

8.2 廣達 / QCT：GB200 / GB300 交付經驗是優勢

QCT 已有 GB200 NVL72 出貨與客戶安裝案例，也有 QoolRack 液冷方案。若 Computex 展示 Vera Rubin，最值得看的是它如何從 GB300 過渡到 Rubin：是否展示相同 rack footprint 的 migration、是否有 CDU / facility cooling、是否展示 tray 可維修性。

推測供應鏈：廣達可能走 NVIDIA reference design + QCT rack / cooling integration；具體冷板、manifold、CDU 供應商需現場確認，不應直接假設全由單一台廠供應。

8.3 緯創 / 緯穎：一個看 L10，一個看 AI factory delivery

SemiAnalysis 指 Wistron 是 Rubin L6 / L10 重要角色；Wiwynn 則在 GTC 2026 公開展示 Vera Rubin NVL72 AI factory infrastructure readiness，並強調從 board-level innovation 到 rack-scale integration and validation。這組分工很值得看：緯創偏製造 / tray automation，緯穎偏 cloud infrastructure solution、L11/L12 delivery、cooling validation、storage / telemetry / leak detection。

推測供應鏈：若 Wiwynn 展示 Storage-Next / GPU-initiated storage / liquid cooling validation，代表它不是只做 rack，而是朝 CSP AI factory acceptance 與運維整合走。

8.4 ASUS / Gigabyte：企業與區域 AI factory 通路

ASUS 已公開 ASUS AI POD on NVIDIA Vera Rubin NVL72，Gigabyte 也有 Vera Rubin NVL72 資料中心文件。這兩者可能不像 hyperscaler 主力 ODM 一樣掌握最大 L10 量，但對企業 AI factory、區域 sovereign AI、edge-to-data-center 解決方案有展示價值。

推測供應鏈：更可能看到完整 solution stack、storage、network、liquid cooling、enterprise deployment，而不是大量 tray automation 細節。

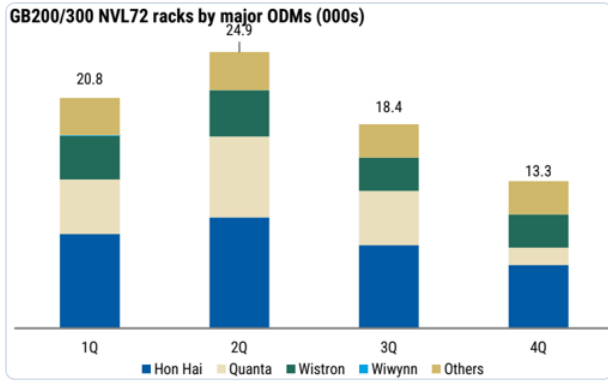
10. 既有資料庫資訊整合

本章整合 wiki / Raw_data 內已 ingest 的資料，避免只依賴外部新聞推測。核心資料包含：Morgan Stanley 2026-05-11 GB200/300 NVL72 rack 月度報告、奇鋁 2026-05-14 法說 memo、高力 2026-05-08 凱基報告、Vera Rubin 連接器供應鏈 memo、SOCAMM2 技術頁、AI 伺服器液冷供應鏈頁。

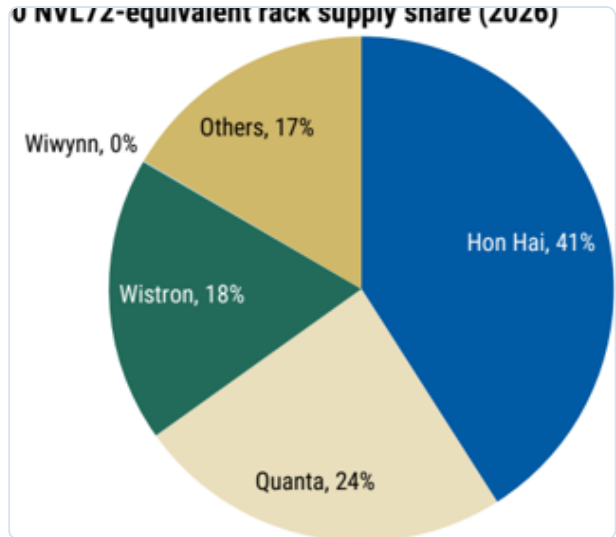
10.1 GB200/300 時代的 ODM 出貨基準

Morgan Stanley 2026-05-11 報告提供了 GB200/300 NVL72-equivalent monthly rack output。這是推測 Rubin Computex 展示與後續份額的重要基準，因為 Rubin 的供應鏈不是從零開始，而是沿著 GB200 / GB300 的 rack 量產經驗升級。

項目	資料庫數據	對 Rubin 的推論
2026 年 4 月全產業 GB200/300 rack output	約 8.3K 台，月減 3%	GB rack 已進入高量產，Rubin 將在同一批 ODM / EMS 能力上轉換
2026 年全年 rack 預估	70-80K 台，年增超過 100%；2025 約 29K 台	Rubin 不是取代小市場，而是接續已放大的 rack 裝配產能
MS GPU ODM preference	Wistron > Hon Hai > Quanta	緯創在 L10 computing tray 的重要性需高估；但 Rubin 要拆 L10 / L11 / L12 看
實際交付 vs rack equivalent	MS 提醒 Wistron computing tray rack equivalent 未必等於 L11 完整 rack 交付	Computex 看 Rubin 時，必須分辨 tray、rack、cluster acceptance 三層



MS 估算 2026 年 GB200/300 NVL72-equivalent monthly rack output。



MS 估算 2026 年 GB200/300 rack supply share。

10.2 ODM 既有出貨基準

公司	既有資訊	Computex / Rubin 判讀
鴻海 2317	MS 估 2026 年 4 月 GB rack 約 3.7K 台；2Q26 AI rack shipment 預期約 10K 台，QoQ +18%。公司公開 GTC 2026 展示 Vera Rubin NVL72 full-system rack。	若 Computex 展示 Rubin full rack，應重點看垂直整合、MDC、散熱元件、電力傳輸、connector / busbar 是否一起展示。
廣達 2382 / QCT	MS 估 2026 年 4 月 GB200/300 rack 約 2.1K 台；2Q26 仍朝 7K+ racks。QCT 有 GB200 NVL72 出貨與安裝案例。	看其是否把 GB300 / Rubin 放在同一個 rack transition story，並展示 QoolRack、CDU、serviceability。
緯創 3231	MS 估 2026 年 4 月 GB200/300 computing tray 約 1.3-1.4K rack equivalents；2Q rack shipment 估約 4.2K，QoQ +4%。SemiAnalysis 指 Wistron 是 Rubin L6 / L10 重要角色。	Rubin 看點在 L6 PCBA 與 L10 tray automation，而不只是完整 rack 外觀。
緯穎 6669	MS 報告提 2026 年 4 月營收 NT\$82.7B，MoM -16%；公司公開 GTC 2026 展示 Vera Rubin NVL72 AI factory infrastructure readiness。	緯穎更應看 L11/L12 rack delivery、cloud infrastructure solution、liquid-cooling validation、Storage-Next / telemetry。

10.3 液冷資料庫結論：奇鋌 vs 高力要分層看

既有液冷供應鏈資料把液冷拆成晶片端、tray / rack 端與 facility 端。這對 Rubin 特別重要，因為 Rubin 是 100% liquid cooled，且熱從 GPU 擴散到 NIC / SSD / VRM / management module。

公司	既有資訊	Rubin 對應
奇鋌 3017	1Q26 營收 490.38 億，YoY +110.2%；EPS 20.17；散熱產品占 63.6%，Server & Networking 占 66.4%。公司為 NVIDIA First Design Partner，VR 水冷板已小批量出貨，並布局 CX9 網卡、配電板、收發模組、Switch 周邊液冷。	若 Computex 展出 tray 內冷板、internal manifold、front module cooling，奇鋌邏輯最直接。
高力 8996	1Q26 營收 34.2 億，YoY +237.6%；EPS 6.54；液冷占營收 60%。主力為 RDHx + in-row CDU + PHE / 熱交換；凱基指出 CDU 自 2H26 成為 PHE 新動能。	若 Computex 展出 rack / row-level CDU、RDHx、facility loop、45°C liquid cooling、chiller-less，高力邏輯更強。

10.4 連接器 / SOCAMM2 資料庫結論

SOCAMM2 是 Vera CPU 記憶體連接技術，採壓縮接觸、螺絲鎖合、600+ pin，支援 LPDDR5X 9,600+ MT/s，目標是在 Vera CPU 旁提供可維修、低功耗、高頻寬的記憶體模組。優群公司資料則把 SOCAMM2 與 CABB 列為 AI 伺服器 cableless 趨勢核心產品。

產品	技術描述	Computex 看點
SOCAMM2	CPU 與記憶體模組在同一節點板，不走獨立線纜；兼具 LPDDR 低功耗與可維修性。	若 ODM 展示 Strata / Vera CPU 周邊，注意是否標示 SOCAMM、LPDDR5X module、connector vendor。
CABB / B2B connector	主板與 GPU 子板 / 模組間的壓接式板對板互連，符合 cableless 趨勢。	若展示 midplane / blind mate，記錄 connector、mating force、serviceability 相關說法。
Metal Bar	高電流傳輸與 EMI 屏蔽利基產品。	Rubin 50V / 高電流 busbar 與 power delivery 若被展示，Metal Bar 題材可追但需驗證實際用量。

10.5 光通訊 / CPO 資料庫結論

既有矽光子、COUPE 與光通訊資料整理：GB300 → Vera Rubin → Rubin Ultra 的升級，不是 pluggable 與 CPO 二選一，而是 scale-out / scale-up 連接方式同步提高 dollar content。Goldman Sachs 摘要指出 Vera Rubin NVL72 (2H26-2027) 可能開始部分情境導入 CPO；Rubin Ultra NVL144 / NVL576 才是 3.2T CPO 與 PCB midplane 更大規模導入期。

對 Computex 而言，若現場展示 Vera Rubin NVL72 搭配 1.6T optical module、AEC、CPO TOR，應把它視為「Rubin scale-out network mix」訊號，而不是直接推論所有 Rubin rack 都會用 CPO。真正大規模 CPO 主題仍需追 Rubin Ultra / Spectrum-X / Quantum-X 後續平台。

11. 可能出現在整機櫃的供應鏈產品

機櫃位置	可能產品	潛在廠商 / 觀察對象	推論基礎
Strata	Rubin GPU、Vera CPU、SOCAMM、GPU cold plate、VRM、IBC	NVIDIA、HBM / LPDDR5X 供應商、SOCAMM 連接器、冷板、DrMOS / VRM	Strata 是每 tray 核心計算板
Orchid	CX-9 NIC、OSFP cage、E1.S SSD、front-end cold plate	NVIDIA、OSFP / connector、storage、冷板 / manifold	Rubin 將 NIC 移到前端，縮短高速網路訊號
Midplane	高階 PCB、B2B connector、M8/M9 CCL、HVLP4 copper foil	Amphenol、台光電、欣興、南電、其他高階 PCB / CCL 供應商	Cableless 設計的核心增量
Cooling loop	Cold plate、internal manifold、rack manifold、MQD / UQD、leak detection	奇鋳、高力、台達、CoolIT、Boyd、AVC、CPC、Staubli、Danfoss、Parker	Rubin tray 100% liquid cooled
Rack power	110kW power shelf、PSU、50V busbar、liquid-cooled busbar、PDB	台達、power shelf / busbar / connector 供應鏈	180-220kW rack power 推動 power infrastructure 升級
Facility side	CDU、dry cooler、adiabatic tower、chiller / chiller-less design	高力、台達、Schneider、Vertiv、nVent、SPX、BAC、Evapco	45°C inlet / high-flow liquid cooling 使 facility integration 重要
Scale-out	1.6T optical、AEC、CPO TOR、Quantum-X800 / Spectrum-X	光模組、AEC、CPO、switch、fiber connector 供應鏈	每 GPU 1.6T scale-out；客戶網路架構決定 BOM
Storage / management	BlueField-4、custom DPU、Storage-Next / ICMS、BMC / DC-SCM	NVIDIA、hyperscaler 自研、storage partners、Aspeed / management ecosystem	Agentic AI 長上下文與 KV cache 需求推動 storage fabric

12. 台股個股觀察名單：機會、驗證訊號與風險

個股結論： Vera Rubin 的台股機會不只在 ODM。若 Computex 展示接近量產型整機櫃，最值得優先追的是「高功率電源線 / Whip Cable / busbar」、「SOCAMM2 / CABB / 高速連接器」、「tray 內液冷與 facility-side CDU」、「1.6T 光通訊與 CPO 被動元件」四條線。以下為研究觀察名單，屬產業鏈機會排序與驗證清單，不代表任何買賣建議。

12.1 優先觀察梯隊

梯隊	標的	Rubin 相關位置	為什麼重要	主要驗證訊號
第一梯隊	良維 6290	大功率電源線、Hi-Amp power cord、Whip Cable / 機櫃配電線材	Rubin rack 功耗升到 180–220kW，power shelf 接 415–480VAC、100A whips，外部報導指良維入列 Vera Rubin 電源線供應清單。	AI data center 線材營收比重、Hi-Amp 出貨、UL / OCP 認證、是否被 ODM 展品或客戶簡報點名。
第一梯隊	優群 3217	SOCAMM2、CABB、板對板壓接連接器、Metal Bar	Rubin cableless 把 tray 內 flyover cable 降低，價值轉向高密度 B2B connector 與 SOCAMM2；優群是 SOCAMM2 / CABB 最明確的小型台股觀察核心。	SOCAMM2 認證進度、Vera Rubin 供應商排序、單顆 ASP、CABB 是否實際進入 Rubin tray。
第一梯隊	嘉澤 3533	SOCAMM2、CPU socket、PCIe / DRAM connector	嘉澤是台灣高階 socket / connector 代表，若 SOCAMM2 從 Vera Rubin 開始放大，嘉澤是最直接的大型連接器映射之一。	伺服器 connector 出貨、SOCAMM2 是否量產、AI server 營收占比、毛利率是否受高階產品拉升。
第一梯隊	奇鈺 3017	GPU / CPU / NIC / switch tray 冷板、manifold、liquid loop	Rubin 從 GPU 局部液冷走向 100% liquid cooled，散熱從晶片延伸到 front module、CX-9、NVSwitch、PDB 周邊。	VR 水冷板放量、switch / NIC cooling 是否出貨、單櫃散熱 ASP、非中產能與良率。
第一梯隊	台達電 2308	Power shelf、PSU、800V HVDC、CDU、資料中心電力基礎設施	Rubin 的瓶頸會從單一伺服器延伸到資料中心配電；台達同時具備 power 與 thermal infrastructure 故事。	800V HVDC / power shelf 客戶導入、AI PSU 市占、CDU / rack power solution 是否與 ODM 展示綁在一起。

12.2 高功率電源線 / 配電線材：良維是這版新增重點

Rubin rack reference design 的 power shelf 需要從資料中心端接收高壓三相電，並透過 50V busbar 進入 rack。SemiAnalysis 指出 VR NVL72 reference design 採四組 110kW power shelves，每個 power shelf 接兩條 100A whips；這使外部 AC input cable、Whip Cable、high-current connector、機櫃配電線組的重要性顯著提高。市場報導指出，Vera Rubin 電源線供應清單除了鴻海，也包括貿聯-KY、佳必琪與良維；其中良維明確提到 Hi-Amp 大電流電源線與機櫃配電產品和 AI data center 高功率需求契合。

公司	機會點	Computex 要問什麼	風險
良維 6290	大功率電源線、Hi-Amp 線材、資料中心機櫃配電線組；若 Rubin rack 使用更多 100A whips，單櫃線材價值上升。	是否已進入 Vera Rubin / GB300 客戶量產清單？AI data center 營收占比？是否具備北美安規、OCP、資料中心級測試能力？	線材仍可能面臨 ASP / 毛利率競爭；實際份額需由客戶驗證與出貨數字確認。
貿聯-KY 3665	高功率線束、資料中心 power interconnect、800V HVDC 前期開發；可從單一線材往系統整合提高專案價值。	是否展示 800V HVDC、BBU / CBU、rack-level power harness？是否與 CSP 共同開發下一代高功率密度方案？	專案開發週期長，短期營收確認可能慢於題材反應。
佳必琪 6197	大電流產品、AC input connector、Whip Cable；外部報導提到 OCP 與北美 UL 認證及 CSP 量產等級訂單。	是否有 Vera Rubin / GB300 對應產品？北美雲端客戶出貨節奏？Whip Cable 與 AC input connector 量產比重？	若只是 general data center power cable，需避免直接等同於 Rubin share。
鴻海 2317	垂直整合電力傳輸、機構、connector、rack integration；可能自製或統包部分配電模組。	展品是否打開 power shelf / busbar / cable routing？是否說明自製零組件比例？	鴻海受惠明確但基期大，單一零組件對 EPS 彈性低於中小型零組件廠。

12.3 連接器：優群與嘉澤是 Rubin cableless 的核心觀察

Rubin 的 cableless design 很容易被誤解成「線材減少，所以連接器不重要」。實際上，tray 內部線纜下降後，訊號完整性、組裝公差與維修性轉到 SOCAMM2、CABB / B2B connector、midplane connector 與高階 PCB 上。也就是說，傳統 cable 的份額下降，但高階壓接連接器、socket、OSFP cage、PCIe / DRAM connector 的技術門檻上升。

公司	Rubin 連結	機會判斷	關鍵反證
優群 3217	SOCAMM2、CABB、Metal Bar	最有小型股彈性的 Rubin 連接器觀察標的。若能從第四大 SOCAMM2 供應商定位擴大到 CABB / B2B connector，題材會從記憶體插槽延伸到整個 cableless tray。	若 SOCAMM2 share 低於預期，或 CABB 沒有實際導入 Rubin，題材彈性會明顯下降。
嘉澤 3533	SOCAMM2、CPU socket、PCIe / DRAM connector	較偏大型、既有高階 socket / connector 龍頭。Rubin 若讓 SOCAMM2 與 CPU-side connector 變成標準配備，嘉澤受惠邏輯直接。	市場預期可能已反映在評價；需看到 AI server connector 出貨與毛利率同步驗證。
鴻騰 / Foxconn Interconnect	高速連接、OSFP / cable、rack 內互連	若鴻海整櫃展示強調自製 connector / cable，鴻騰可能是垂直整合受益者。	台股投資映射不如 2317 直接；需看揭露程度。
Amphenol	Paladin HD2、high-speed B2B connector、OSFP ecosystem	國際連接器龍頭仍可能是 Rubin cableless 最大贏家之一，是判斷台廠份額的 benchmark。	若 NVIDIA reference design 採國際大廠高度集中，台廠只能吃 secondary source。

12.4 液冷與電源：奇鋹、高力、台達、雙鴻、富世達要分工看

Rubin 的散熱不是單一冷板題材。展場若只看到 GPU cold plate，重點在奇鋹、雙鴻、CoolIT、Boyd；若看到 rack manifold、CDU、facility loop、45°C liquid cooling，重點轉向高力、台達、Vertiv、Schneider；若看到 quick disconnect / hinge / sliding mechanism，富世達、CPC、Staubli、Parker 等機構與流體連接件的價值會提高。

公司	主要看點	展場驗證	投資判讀
奇鋁 3017	水冷板、manifold、switch / NIC 周邊液冷、NVIDIA First Design Partner	是否展示 Rubin tray 內冷板、CX-9 / NVSwitch cooling、配電板散熱。	最直接的 tray-level thermal 受惠股，但需追毛利率與客戶議價。
高力 8996	RDHx、in-row CDU、PHE、facility-side heat exchange	若 ODM 展示整櫃 + CDU / RDHx / 45°C facility loop，高力邏輯提高。	比較像資料中心基礎設施升級股，與「chiller-less」敘事連動高。
台達電 2308	AI PSU、power shelf、CDU、800V HVDC、資料中心整體電力方案	是否展示 power rack、DC-DC shelf、BBU / CBU、CDU 與 rack-level 控制。	Rubin 若把瓶頸從伺服器推到電力基礎設施，台達是大市值核心受益者。
雙鴻 3324	冷板、CDU、快接頭相關液冷解決方案	是否被 ODM 或 CSP 展示為 GB300 / Rubin cooling partner。	若從冷板延伸到 CDU / rack solution，估值彈性提高。
富世達 6805	水冷接頭、機構件、折疊 / sliding mechanism 延伸	是否出現 blind mate liquid connector、serviceable tray、快速抽換機構。	Rubin 越強調 serviceability，機構與流體接頭價值越高。

12.5 光通訊 / CPO：短期看 1.6T，長期看 Rubin Ultra CPO

Vera Rubin NVL72 階段不應把 CPO 當成全面替代 pluggable optics 的故事。更合理的投資拆法是：2026-2027 先看 1.6T optical module、OSFP、AEC、FAU / OIN、CW laser；Rubin Ultra / Spectrum-X / Quantum-X 後續平台再看 3.2T CPO、external laser source、D-FAU / iFAU、CPO 測試與 active alignment。

公司	環節	Rubin 觀察重點	風險
波若威 3163	OIN、MPO 高密度線束、FAU / Shuffle Box	若 Vera Rubin / Rubin Ultra scale-out 光連接需求放大，FAU 與高密度光纖連接是核心位置；外部報導亦提到其打入 Vera Rubin 平台的市場說法。	MT ferrule / 上游料件短缺可能限制出貨；需驗證客戶與產能。
上詮 3363	FAU、CPO optical coupling、光纖陣列	CPO 最後一哩對位精度高，上詮若與台積電 / CPO 生態綁定，Rubin Ultra 彈性高。	CPO 大量商轉時間可能落在 Rubin Ultra，不一定立即反映 Vera Rubin NVL72。
光聖 6442	光纖連接器、SiPh 封裝 / 耦合、ELS 相關	若展場強調 CPO OE、external laser source、SiPh package，光聖是台廠映射之一。	產品 mix 複雜，需拆清楚資料中心 AI 與傳統 telecom 貢獻。
聯亞 3081 / 全新 2455	EML、CW laser、磊晶	1.6T / CPO 都需要高階雷射供給，若 external laser source 成為標準，雷射供需可能延續吃緊。	價格、良率與客戶集中度會放大波動。
穎崴 6515 / 旺矽 6223	CPO / 高速測試 socket、probe card、active alignment 測試	CPO 與高 pin 大封裝使測試設備、socket、probe card 需求增加。	驗證期長，需看實際量產節點與客戶 capex。

12.6 PCB / 材料 / 機構：不是最早反應，但會決定良率

公司	位置	Rubin 受惠邏輯	需要追的指標
台光電 2383	M8 / M9 CCL、低損耗材料	Cableless midplane 與 NVSwitch board 對低損耗材料要求提高；Rubin Ultra PCB midplane 更是長線催化。	AI server CCL 出貨、M8/M9 認證、毛利率。
欣興 3037 / 南電 8046	高階 PCB / substrate / midplane	Rubin tray 內 PCB 層數與訊號完整性要求提升；Rubin Ultra midplane 題材更強。	AI server PCB / substrate 稼動率、客戶認證、資本支出。
金像電 2368	Server motherboard PCB	GB300 / Rubin / agentic AI general server 同步推升高階伺服器板需求。	高速伺服器板占比、M8/M9 導入節奏。
川湖 2059 / 勤誠 8210 / 晟銘電 3013	滑軌、機殼、rack / sidecar 結構	高重量、高流量液冷、可維修 tray 與 sidecar cooling 會提高機構件規格。	是否與 ODM Rubin rack 展示綁定、單櫃機構件 ASP。

12.7 建議的展前持股 / 觀察分類

分類	偏短線 Computex 事件	偏中期量產驗證	適合追的問題
事件彈性	良維、優群、波若威、上詮	需確認是否真進 Rubin BOM	展場是否看到產品標籤？公司是否在法說承認 AI / Rubin 相關出貨？
基本面續航	奇鋌、台達、嘉澤、台光電	Rubin / GB300 / ASIC 共同拉動	月營收、毛利率、AI server 占比是否同步上升？
資料中心基礎設施	高力、台達、貿聯、佳必琪	Power rack、CDU、BBU / CBU、800V HVDC	客戶是否從單品採購轉成整套 rack / facility solution？
長線 CPO	上詮、光聖、聯亞、全新、穎威、旺矽	Rubin Ultra / 3.2T / CPO TOR	2026 是驗證還是量產？CPO 貢獻何時可量化？

13. Computex 展後投資觀察重點

- Full rack 還是 demo shell**：完整展示 72 GPU / 36 CPU、NVLink switch tray、power shelf、liquid cooling loop 的價值高於只放外觀機櫃。
- L10 自動化能力**：現場若提 assembly automation、blind mate、5-minute tray assembly、robotic insertion，優先追鴻海、廣達、緯創。
- L11 / L12 delivery**：若展示 cluster acceptance、cooling validation、data center integration、Storage-Next，優先追緯穎 / Wiwynn 類 cloud infrastructure provider。
- 液冷到哪一層**：只看 cold plate 不夠，要看 internal manifold、QD、rack manifold、CDU、facility loop、45°C inlet。
- 電源是否成為主題**：若出現 800VDC、BBU、CBU、liquid-cooled busbar、power rack，代表 AI data center power 已成為下一階段瓶頸。
- 連接器 / PCB 是否被點名**：SOCAMM、CABB、B2B connector、Paladin HD2、M8/M9、HVL4、quartz cloth 都是 Rubin cableless 的關鍵詞。
- 網路方案**：Quantum-X800 / Spectrum-X、1.6T OSFP、AEC、CPO TOR 的不同選擇，會影響光通訊與高速線纜供應鏈。

8. **客戶定位**：Neocloud / enterprise 可能接受 NVIDIA full stack；大型 hyperscaler 可能替換 DPU、management、network 模組。展品若可客製化，需記錄哪些模組可替換。
9. **GB300 到 Rubin 轉換節奏**：若 ODM 同時展示 GB300 NVL72、Vera Rubin NVL72、HGX Rubin NVL8，代表它在承接不同資料中心 readiness，而不是一次切換。
10. **毛利率反向風險**：NVIDIA 控制更多 BOM 與 L10 tray，可能提高出貨穩定性，但壓縮 ODM 差異化與毛利。

14. 風險與反證

14.1 展示不等於量產份額

Computex 展品可能是 reference design、engineering sample、partner demo 或 mockup。若沒有出貨時間、客戶、產能、automation、acceptance test 資訊，不應直接推成實際 share。

14.2 NVIDIA 控制力提高可能壓縮 ODM 毛利

Vera Rubin 的標準化與 L10 tray 趨勢對客戶部署有利，但可能讓 ODM 從「設計者」變成「整合 / 製造 / 交付者」。高單價 rack 不代表高毛利，尤其當核心 BOM 與 design control 在 NVIDIA 手中。

14.3 Cooling / power 供應商不一定由 ODM 決定

Hyperscaler 可能指定 cooling / power 架構，並使用其既有供應商。ODM 展出某一套方案，不代表所有客戶部署都使用同一套冷板、CDU、QD、power shelf。

14.4 HBM4 / LPDDR5X / SOCAMM ramp 仍是節奏風險

Rubin 的 performance story 仰賴 HBM4 高頻寬與 Vera CPU 大容量 LPDDR5X。若 HBM4 pin speed、SOCAMM 連接器良率、LPDDR5X module、platform qualification 延遲，rack 放量也會延後。

14.5 Chiller-less 不會一夕取代傳統 cooling

45°C liquid cooling 支援更高效率的 economizer / dry cooler / adiabatic tower，但資料中心仍需考量 mixed hall、備援、維修、氣候條件與 workload flexibility。短中期不應把 chiller 視為立即被取代。

15. 來源與延伸閱讀

本地 wiki：

- 供應鏈_Vera_Rubin_NVL72機櫃
- memo_SemiAnalysis_Vera_Rubin_NVL72_20260520
- memo_VeraRubin_機櫃連接器供應鏈_20260519
- 供應鏈_AI伺服器液冷
- 供應鏈_AI伺服器板上電源
- 技術_SOCAMM2
- 技術_SiPh
- 技術_COUPE
- 供應鏈_光通訊
- 供應鏈_CPO_D-FAU

- 報告_GS_AI光網路_20260417

外部公開來源：

- NVIDIA Vera Rubin NVL72 官方頁：<https://www.nvidia.com/en-us/data-center/vera-rubin-nvl72/>
- NVIDIA DGX Vera Rubin NVL72 官方頁：<https://www.nvidia.com/en-gb/data-center/dgx-vera-rubin-nvl72/>
- NVIDIA 技術部落格：Inside the NVIDIA Vera Rubin Platform：<https://developer.nvidia.com/blog/inside-the-nvidia-rubin-platform-six-new-chips-one-ai-supercomputer/>
- NVIDIA OCP / gigawatt AI factories blog：<https://blogs.nvidia.com/blog/gigawatt-ai-factories-ocp-vera-rubin/>
- 鴻海 GTC 2026 Vera Rubin NVL72 展示新聞：<https://www.foxconn.com/zh-tw/press-center/press-releases/latest-news/1974>
- Wiyynn GTC 2026 Vera Rubin NVL72 展示新聞：<https://www.wiyynn.com/news/wiyynn-showcases-nvidia-vera-rubin-nvl72-ai-factory-infrastructure-at-nvidia-gtc-2026>
- ASUS AI POD / Vera Rubin NVL72 新聞：<https://press.asus.com/news/press-releases/asus-ai-pod-nvidia-vera-rubin-nvl72>
- 鉅亨網 2026-03-22：輝達 Vera Rubin 電源線供應清單，良維等四台廠入列：
<https://news.cnyes.com/news/id/6393232>
- MoneyDJ 2026-05-04：良維強攻 AI 伺服器 / 資料中心，高階線材比重提升：<https://m.moneydj.com/f1a.aspx?a=62682409-63ec-4963-95ff-74993b8bf142&c=MB070100>
- 經濟日報 2026-04-27：波若威、華星光衝刺 CPO：<https://money.udn.com/money/story/5710/9465556>

註：SemiAnalysis 文章內容與圖片已整理為研究摘要與圖表引用；本報告使用摘要、轉述與圖檔引用，不重製原文長段落。