電子硬體產業

TPU v8 放量出貨，推升 2027 年 AI 成長動能

焦點內容

1. Google (美) TPU v8 首次分為 8t 與 8i 兩套針對不同工作負荷量最佳化的架構，預計 4Q26 於硬體供應鏈進入量產階段，將對 2027 年增添 AI 伺服器市場硬體供應鏈動能
2. TPU v8 機櫃設計雖分化為 8t/8i，但多半為 v7 之延伸，放量嘉惠規格升即之散熱、電源、PCB/CCL、載板與光通 OCS 業者。
3. 2027 年將是 Google TPU 明顯放量年，從晶片量的角度，預期 TPU v8 出貨量將由 2026 年之約 402 萬顆增至 2027 年之約 910 萬顆 (年增 126%)，v8 機櫃總出貨量可望翻倍，將同步拉動 TPU 晶片、ODM、散熱、電源、PCB/CCL 與光通訊需求。主要受惠廠商且營收貢獻度達 10%以上之公司包括: 晶片設計商聯發科（2454 TT）、主要 L10/L11 系統組裝商 Celestica（CLS US）、電源廠商台達電（2308 TT）、ABF 載板供應商欣興（3037 TT）、CCL 供應商台光電（2383 TT）、PCB 業者金像電（2368 TT）、水冷板新進廠商奇鋐（3017 TT）、HVLP4 銅箔供應商金居（8358 TT）。

重要訊息

Google (美) TPU v8 預計 4Q26 於硬體供應鏈進入量產階段，將增添 2027 年 AI 伺服器市場硬體供應鏈動能。

評論及分析

Google TPU v8 首次分為 8t 與 8i 兩套針對不同工作負荷量最佳化的架構，預計 4Q26 於硬體供應鏈進入量產階段。Google TPU 已進入第八代 (v8)，與前一代 TPU v7 (Ironwood) 同時承擔訓練與推論不同，TPU v8 首次明確分化為 TPU 8t (Zebrafish) 與 TPU 8i (Sunfish)兩套針對不同工作負荷量 (workload) 最佳化的架構。其中 TPU 8t（Training）定位大型模型 pre-training 與 embedding-heavy workload，延續 3D Torus scale-up 架構，單一 Superpod 可擴展至 9,600 顆 TPU；而 TPU 8i（Inference）則鎖定 post-training、sampling、real-time inference 與 reasoning，導入全新的晶片間互連 Boardfly inter-chip interconnect (ICI) topology，單一 pod 最多可配置 1,152 顆、其中最高 1,024 顆 active TPU。TPU 的需求來源可分成 Google 內部需求與外部雲端客戶兩大部分。內部需求主要來自 Google DeepMind 與 Google 自有 AI 服務，包括 Gemini 等 frontier models 的 training、post-training 與 inference；至於外部需求中，目前最重要且公開規模最大的客戶之一為 Anthropic (US)，2025 年 Anthropic 宣布擴大 Google Cloud TPU 使用規模至最高 100 萬顆 TPU、超過 1GW 產能，於 2026 年陸續上線；而於 2026 年 4 月更進一步宣布與 Google、Broadcom 簽訂新的合作協議，取得數 GW (multiple gigawatts) 的下一世代 TPU 產能 (TPU v8)，預計 2027 年開始上線，主要支援 Claude frontier 模型。

v8 機櫃設計雖分化為 8t/8i，但多半為 v7 之延伸。8t 延續適合大型 training 的 3D Torus 架構，但晶片間互連頻寬 (Inter-chip interconnect (ICI) bandwidth) 較前代提高 2 倍，並新增 Virgo scale-out network，8i 則改採全新的 Boardfly topology，每個 tray 以 4 顆 TPU 形成 building block，8 個 board 組成 32-chip group，再由 36 個 group 透過 OCS 互連，形成最多 1,152 顆 physical TPU、最高 1,024 顆 active TPU 的 pod。依目前供應鏈資訊，v8 每層 L6 板子上除了 TPU 之外，還進一步整合 Power 與高速 Transceiver/Optics，並導入 Google 第四代（Gen4）液冷；因此 v8 機櫃升級的核心已不只是 TPU 晶片效能提升，而是從運算、記憶體、網通、電源、散熱與光通訊全面進入系統級之設計。目前主要 L10/L11 組裝廠商仍為 Celestica (美)，鴻海(2317 TT, NT$251.5, 增加持股)繼 v7 組裝 CPU 層與機櫃後，於 v8 增加 TPU 層與 L11 機櫃之組裝份額，而英業達(2356 TT, NT$59.5, 持有)則主要供應 TPU 之 L6 為主，亦將有機會擴及系統。

v8 放量嘉惠規格升即之散熱、電源、PCB/CCL、載板與光通 OCS 業者。

水冷設計: TPU v8 由晶片液冷升級至系統級液冷，CDU 容量與液冷零組件同步升級，市場資訊顯示由 v7 約 1.4MW 升至 v8t 約 2MW，並擴大 Manifold、QD、Tubing 等液冷零組件需求，我們預估單櫃水冷價值（未

含 CDU）估約 4.5–5 萬美元；8t 水冷板主要供應商為奇鋐 (3017 TT, NT$3,435, 增加持股)、古河(日)，8i 則以 Cooler Master(台)、Boyd(美)為主。

電源方面: v8 進入 10kW PSU，並開始導入 HVDC，單櫃電源價值明顯提高，v8 在 50V 架構下預估採 3 排 Power shelf，每個含 6 個 10kW PSU，相較 v7 的 5.5k 升規，PSU 單櫃價值由約 2 萬美金增至 2.8-3 萬美金 (提升 40–50%)，並開始相容±400V DC，主要受惠供應商為台達電 (2308 TT, NT$1,890, 增加持股)，其次為 Flex (美)、光寶科 (2301 TT, NT$279.5, 增加持股)。

PCB／CCL／載板方面: v8 同時受惠面積、層數、材料三重升級，ABF 載板由 v7 約 80–85mm／16L 放大至 v8 約 90–100mm／18–22L，估 ASP 提升 40–50%，主要受惠欣興（載板）(3037 TT, NT$1,175, 增加持股)；新一代 v8t、v8i 主板 34-40 層，CCL 材料升級至 M8+ 等級並搭配 HVLP4 銅箔，使 CCL/PCB 平均售價增加 70%-80%，主要受惠台光電（CCL)(2383 TT, NT$4,935, 增加持股)、金像電（PCB）(2368 TT, NT$1,120, 增加持股) 與金居 (8358 TT, NT$495, 未評等)（HVLP 銅箔）。

光通訊方面: 8t 為 Scale、8i 為 Latency 的網路分流，帶動 Optics／Fiber／OCS 需求。8t 維持 3D Torus 並新增 Virgo scale-out，我們預估 Optical Attach Rate 最高可由 v7 的 1.5 提升至 2.0／TPU（+33%），8i 則以 Boardfly 將最大 network diameter 由 16 降至 7 hops、attach rate 約 1.25；TPU 放量仍將擴大光模組、MPO/高密度光纖與 OCS 需求，主要供應鏈包括 Lumentum（OCS）、光聖 (6442 TT, NT$1,490 未評等)（光纖/配線、Scale-out）、聯亞 (3081 TT, NT$2,695, 未評等)／華星光 (4979 TT, NT$569, 未評等)／光環（ 3234 TT, NT$173, 未評等)／聯鈞 (3450 TT, NT$523, 未評等)（LD 相關）及全新 (2455 TT, NT$552, 未評等)／環宇-KY (4911 TT, NT$488, 未評等)（PD 相關）。

投資建議

2027 年將是 Google TPU 明顯放量年，從晶片量的角度，預期 TPU v8 出貨量將由 2026 年之約 402 萬顆增至 2027 年之約 910 萬顆 (年增 126%)，其中由聯發科設開發之 v8t 晶片量預計約為 360 萬顆，Broadcom 開發之 v8i 晶片量預期約為 550 萬顆，v8 機櫃總出貨量可望翻倍成長，將同步拉動 TPU 晶片、ODM、散熱、電源、PCB/CCL 與光通訊需求。相關受惠廠商包括: 晶片設計商聯發科（2454 TT, NT$4,920 增加持股，TPU v8t ASIC，2027 年營收貢獻約 40%）、Broadcom (美, TPU v8i ASIC，約 45%)、主要 L10/L11 系統組裝商 Celestica（CLS US，25–30%）、電源廠商 (Power Shelf／PSU／±400V HVDC) 台達電(2308 TT）(20–25%)、CCL 供應商台光電（2383 TT，市佔率 70%）(20-25%)、PCB 業者金像電（2368 TT，15–20%）、水冷板新進廠商奇鋐（3017 TT，10–15%）、 HVLP4 銅箔供應商金居（8358 TT，5–10%），其他營收貢獻較不顯著之供應商 (小於 10%)尚有鴻海（2317 TT，TPU 層與 L11 機櫃 ODM）、英業達（2356 TT，L6 組裝）、健策（3653 TT，stiffener）、富世達（6805 TT, NT$2,455, 增加持股，液冷零組件 QD）、ABF 載板供應欣興（3037 TT）與光寶（2301 TT，PSU）。此外，光通訊相關業者應屬 TPU 需求上揚之最大受惠族群，TPU 相關營收貢獻 2027 年皆有雙位數以上，包括聯亞（3081 TT，LD）、光聖（光纖/配線與 Scale-out）及 Lumentum（LITE US，OCS）皆為主要受益廠商。

投資風險

晶片供應順暢度；組裝與零組件瓶頸。

TPU v8 晶片設計與迭代時程

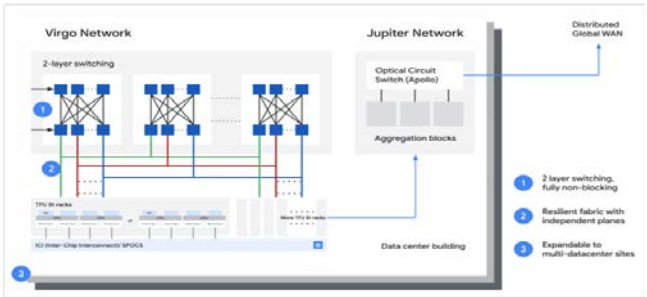
Google TPU 已進入第八代，與前一代 TPU v7 (Ironwood) 同時承擔訓練與推論不同，TPU v8 首次明確分化為 TPU 8t (Zebrafish) 與 TPU 8i (Sunfish) 兩套針對不同 workload 最佳化的架構。其中 TPU 8t（Training）定位大型模型 pre-training 與 embedding-heavy workload，延續 3D Torus scale-up 架構，單一 Superpod 可擴展至 9,600 顆 TPU，並加入原生 FP4、SparseCore、Virgo scale-out network 等設計；而 TPU 8i（Inference）則鎖定 post-training、sampling、real-time inference 與 reasoning，導入 384MB on-chip SRAM、Collectives Acceleration Engine（CAE）及全新的 Boardfly ICI topology，單一 pod 最多可配置 1,152 顆、其中最高 1,024 顆 active TPU。Google 於 2026 年 4 月 Cloud Next 正式公布 TPU 8t/8i，兩者預計於 2026 年陸續提供 Cloud 客戶使用；這代表 Google TPU 的發展已從單純追求算力提升，進一步走向訓練 / 推論架構分流與工作負載特定最佳化(workload-specific optimization)。

圖 1：TPU v8 系列受 memory wall 限制，將其分為訓練(v8t)與推論(v8i)系列

Year	2023	2023	2024	2025	2026	2026	2027F	2028F	2028F	2029F	2029F
	v5e	v5p	v6e	v7x	v8t	v8i	v8px	v9	v9	v10	v10
晶片版本	Viperlite	Viperfish	Ghostlite	Iroodwood	Zebrafish (A5921)	Sunfish	Whalefish (Upgraded Sunfish)	Humufish	Triggerfish (A5922)	Merope	Icefish
ASIC 合作夥伴	Broadcom	Broadcom	Broadcom	Broadcom	聯發科	Broadcom	Broadcom	聯發科	聯發科	Broadcom/Marvell?	聯發科
製程節點	台積電 N5	台積電 N5	台積電 N4	台積電 N3	台積電 N3	台積電 N3	台積電 N3	台積電 N3/N2	台積電 N3/N2	台積電 N2	台積電 N2
記憶體世代	HBM2E	HBM3	HBM3	HBM3E	HBM3E	HBM3E	HBM4E	HBM4	HBM4E	HBM4E/HBM5?	HBM4E/HBM5?
記憶體容量	16GB	95GB	32GB	192GB	216GB	288GB	576GB?	288GB	?	?	768GB?
2.5D 封裝	CoWoS-S	CoWoS-S	CoWoS-S	CoWoS-S	CoWoS-S	CoWoS-S	CoWoS-S	EMIB-T	EMIB-T	CoWoS-L	EMIB-T
混和鍵合/SoIC	N	N	N	N	N	N	N	N	N	Y?	?

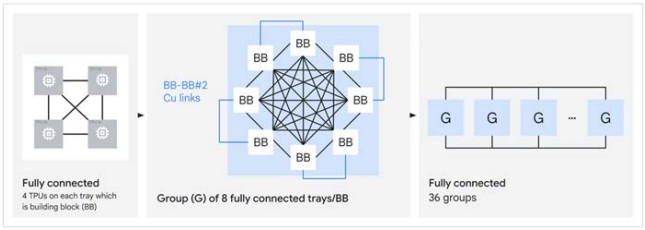
資料來源：凱基

圖 2：TPU 8t 機架級至 Virgo Fabric 的連接



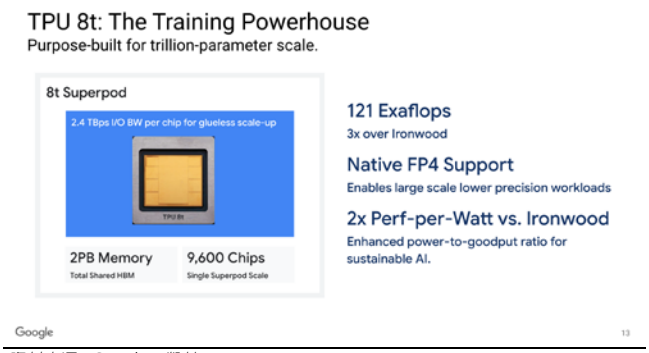
資料來源：Google；凱基

圖 3：TPU 8i 採用分層 Boardfly 拓撲結構：以四個全互連晶片構成的單元為基礎，組成全互連的八板集群；再將 36 個此類集群全互連，構成一個 TPU 8i Pod



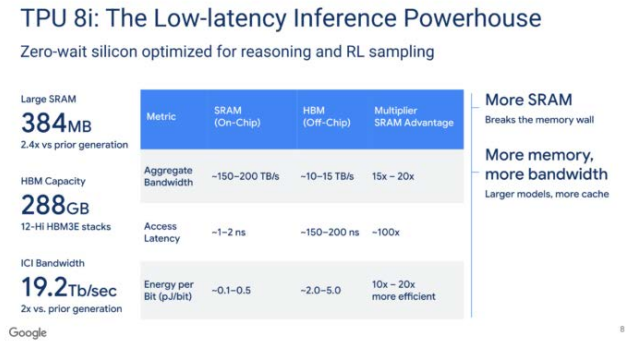
資料來源：Google; 凱基

圖 4：TPU 8t 聚焦大規模訓練



資料來源：Google；凱基

圖 5：TPU 8i 聚焦低延遲即時推論

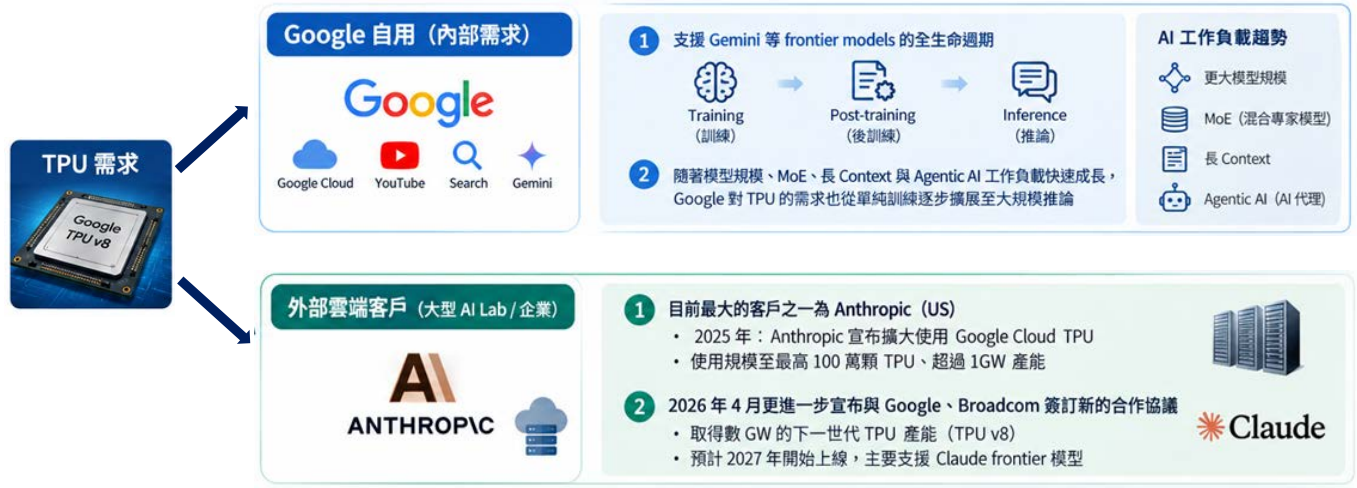


資料來源：Google; 凱基

TPU 之主要需求包括 Google 自用與外部客戶

TPU 的需求來源可分成 Google 內部需求與外部雲端客戶兩大部分。內部需求主要來自 Google DeepMind 與 Google 自有 AI 服務，包括 Gemini 等 frontier models 的 training、post-training 與 inference；隨著模型規模、MoE、長 Context 與 Agentic AI 工作負載快速成長，Google 對 TPU 的需求也從單純訓練逐步擴展至大規模推論。至於外部需求中，目前最重要且公開規模最大的客戶之一為 Anthropic (US)，2025 年 Anthropic 宣布擴大 Google Cloud TPU 使用規模至最高 100 萬顆 TPU、超過 1GW 產能，於 2026 年陸續上線；而於 2026 年 4 月更進一步宣布與 Google、Broadcom 簽訂新的合作協議，取得數 GW (multiple gigawatts) 的下一世代 TPU 產能 (TPU v8)，預計 2027 年開始上線，主要支援 Claude frontier 模型。因此 TPU 的需求結構正在由過去以 Google 自用為主，逐步轉變成「Google 自用 + 大型外部 AI Lab」雙引擎驅動，也使 TPU 開始具備更明顯的人工智慧加速器之商業化特性。

圖 6：TPU 主要需求包括 Google 自用與外部客戶



資料來源：凱基

TPU v8 與 v7 晶片之主要差異

第七代 Ironwood（TPU v7）仍採一套架構同時涵蓋訓練 (training) 與推論 (inference)，每顆 TPU 配備 192GB HBM、7.38TB/s HBM 頻寬，並透過晶片間互連頻寬 (ICI) 3D Torus 架構進行大規模 scale-up；到了 v8，Google 則將不同工作負荷 拆分為 8t 與 8i。TPU 8t 擅長處理大規模、運算密集的訓練工作負載，其設計旨在提供更大的運算吞吐量和更強的可擴展頻寬，其配置 216GB HBM、6.53TB/s HBM bandwidth、128MB SRAM、12.6 PFLOPS FP4，並加入 native FP4、SparseCore 與 Virgo Network，晶片由聯發科設計；TPU 8i 則具有更高的記憶體頻寬，旨在服務於對延遲最為敏感的推理工作負載，強調記憶體容量(memory capacity) 與通訊延遲性 (communication latency)，將 HBM 提高至 288GB、8.60TB/s bandwidth、SRAM 提高至 384MB，並加入 CAE 與 Boardfly topology，FP4 算力為 10.1 PFLOPS，晶片則由 Broadcom 設計。相較 v7 (Ironwood)，8t 大型 training 的 performance-per-dollar 最高提升 2.7 倍，8i 針對大型 MoE 低延遲 inference 的 performance-per-dollar 最高提升 80%，兩者 performance-per-watt 最高均提升約 2 倍；因此 v7 至 v8 最重要的改變並非單純提高 FLOPS，而是透過運算、記憶體、網路及工作負載專業化的共同升級，提高整體系統效率。

圖 7：TPU v8 系列透過運算、記憶體、Compute、Memory、網路及工作負載專業化的共同升級，提高整體系統效率

	TPU v7 Ironwood	TPU v8t	TPU v8i
設計合作夥伴	Broadcom	MediaTek	Broadcom
主要工作負載	大規模訓練、推論與服務	大規模預訓練	後訓練、推理
網路拓樸	3D torus	3D torus	Boardfly
每顆晶片HBM 容量	192 GB	216 GB	288 GB
片上 SRAM（VMEM）	TBA	128 MB	384 MB
峰值 FP4／晶片	TBA	12.6 PFLOPS	10.1 PFLOPS
每顆晶片 HBM 頻寬	7.38 TB/s	6.528 TB/s	8.601 TB/s
HBM 世代	HBM3E	TBA	TBA
CPU host／header	TBA	Arm-based Google Axion	Arm-based Google Axion
單一 Pod／rack	9,216 顆 TPU	9,600 顆 TPU	1,152 顆（ 1,024 顆在線）
Pod／機架 HBM 總容量	1.77 PB	約 2 PB	331.8 TB
單一晶片功耗	1000w	<800w	1300w
散熱	液冷+氣冷	液冷	液冷
載板	80 x 85 mm /16L	80~90 x 80~90 mm/18-20L(TBC)	80~90 x 80~90 mm/18-20L(TBC)
CCL/PCB	24L/M7 ; HVLP2	34L-40L/M8+ ; HVLP4	34-40L/M8+ ; HVLP4
光收發模組	800G (Scale Up)	1.6T (Scale Up)	1.6T (Scale Up)
電源	5.5kW	10kW	10kW

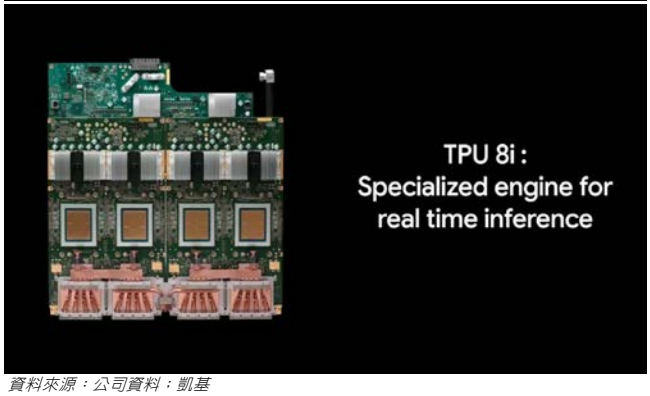
資料來源：凱基

TPU v8 機櫃設計多半為 v7 之延伸

TPU v7 Ironwood 的基本運算單元以 4 顆 TPU 為一個實體主機 (physical host)，單一機櫃 (rack) 採左右對稱設計，共配置 64 顆 TPU，也就是由 16 組 4-TPU host 組成。 每一層包括 4 顆 TPU 與 4 顆 Axion CPU 組成一個計算單元，左右各 8 個單元，顯示機櫃內包含 64 顆 TPU，透過高速 ICI 連接並形成一個 3D Torus Cube。當運算規模超越單一機櫃時，Google 再利用 Optical Circuit Switch（OCS）將多個 64-chip Cube 動態連接，例如 4 個 Cube 即可形成 256-chip pod，而最大 Ironwood Superpod 則由 144 個 Cube × 64 顆 TPU = 9,216 顆 TPU 組成，使數千顆晶片可以共同執行大型模型的訓練與推論。OCS 除了負責 scale-up，也能在 Cube 或 link 發生故障時重新配置光路並繞過故障節點，提高大型叢集的可用性。

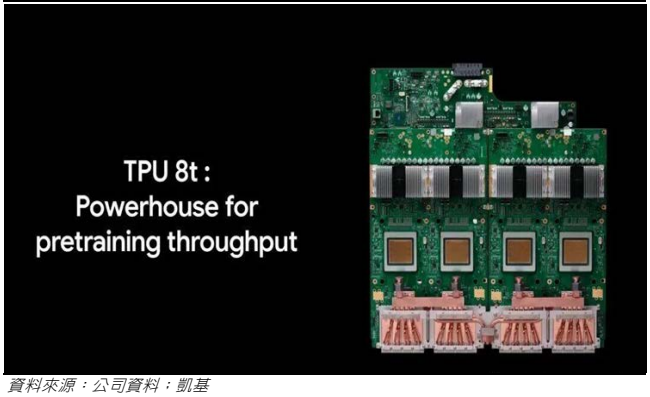
至於 v8 機櫃設計，則從單一架構進一步分化為 8t / 8i。TPU v8 在機櫃與系統設計上最大的改變，是 8t 與 8i 在機櫃架構上不完全相同。8t 延續適合大型 training 的 3D Torus 架構，但晶片間互連頻寬 (ICI bandwidth) 較前代提高 2 倍，並新增 Virgo scale-out network，單一 Virgo fabric 可連結超過 134,000 顆 8t、提供最高 47Pb/s non-blocking bisection bandwidth，再搭配 JAX 與 Pathways，可進一步形成超過 100 萬顆 TPU 的 logical training cluster；8i 則改採全新的 Boardfly topology，每個 tray 以 4 顆 TPU 形成 building block，8 個 board 組成 32-chip group，再由 36 個 group 透過 OCS 互連，形成最多 1,152 顆 physical TPU、最高 1,024 顆 active TPU 的 pod，將 1,024-chip 系統最大 network diameter 由 3D Torus 的 16 hops 降至僅 7 hops，以降低 MoE 與 reasoning workload 的 all-to-all communication latency。依目前供應鏈資訊，v8 每層 L6 板子上除了 TPU 之外，還進一步整合 Power 與高速 Transceiver/Optics，並導入 Google 第四代（Gen4）液冷；因此 v8 機櫃升級的核心已不只是 TPU 晶片效能提升，而是從運算、記憶體、網通、電源、散熱與光通訊全面進入系統級之設計。目前主要 L10/L11 組裝廠商仍為 Celestica (美)，鴻海繼 v7 組裝 CPU 層與機櫃後，於 v8 增加 TPU 層與 L11 機櫃之組裝份額，而英業達(2356 TT, NT$59.5, 持有)則主要供應 TPU 之 L6 為主，亦將有機會擴及系統。

圖 8：TPU v8i tray 示意圖



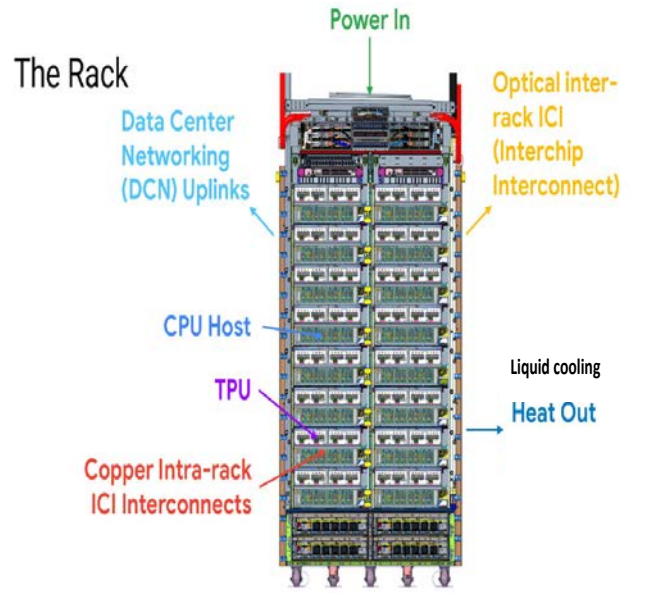
資料來源：公司資料；凱基

圖 9：TPU v8t tray 示意圖



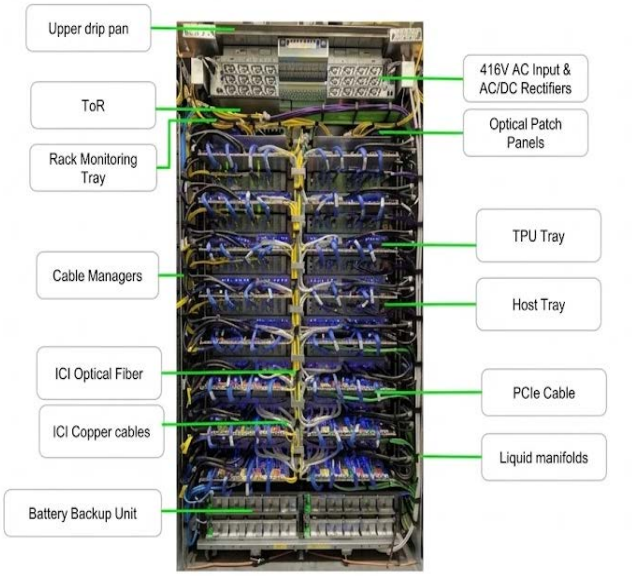
資料來源：公司資料；凱基

圖 10：TPU v7 機櫃示意圖



資料來源：公司資料；凱基

圖 11：TPU v7 基本運算單元以 4 顆 TPU 為一個實體主機，單一機櫃採左右對稱設計，共配置 64 顆 TPU

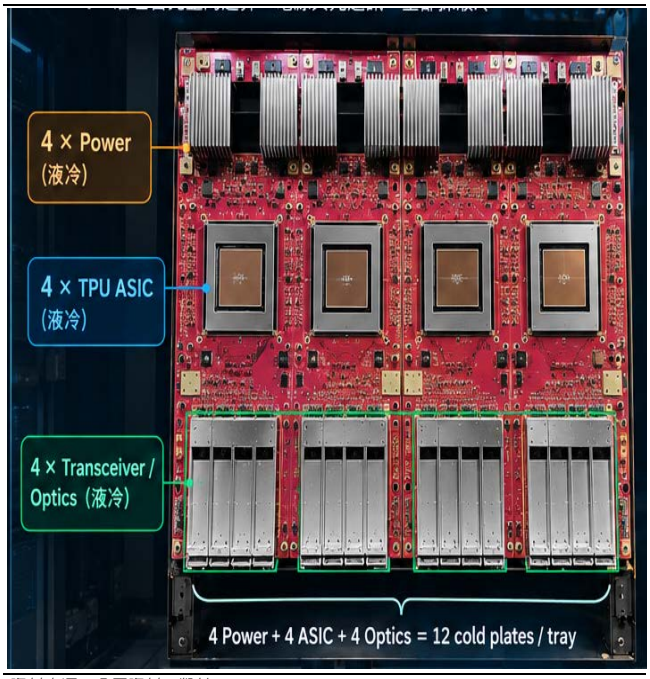


資料來源：公司資料；凱基

TPU v8 散熱從「晶片液冷」走向「系統級液冷」

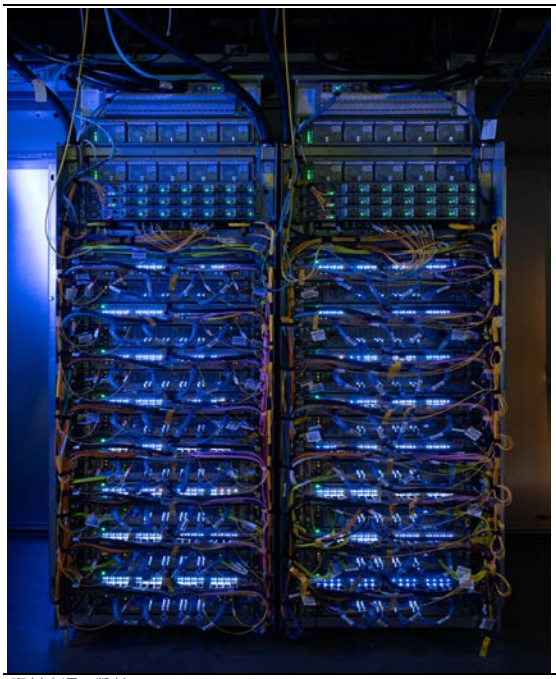
散熱架構是 TPU v7 Ironwood 升級至 v8 的重要變化之一。Ironwood 採用 Google 第三代液冷系統（Gen3），以 CDU（Cooling Distribution Unit）將冷卻液送至機櫃，再透過 manifold、管路與水冷板 (cold plate) 直接帶走高功耗晶片產生的熱量；Google 官方也展示了 Ironwood Superpod 所使用的第三代 CDU，到了 TPU v8，Google 進一步升級至第四代液冷（Gen4），主要原因是 8t／8i 的運算密度與系統功耗持續提高，散熱需求已不再只集中於 TPU 本身，而逐步延伸至整個高密度運算系統。依目前供應鏈拆解資訊，v8 每層水冷板需要 12 片，包含 4 顆 TPU、4 組 Power 及 4 組 Transceiver，若以一機櫃 16 層來計算，即相當於一櫃約 192 片水冷板，由於單層水冷板片數較多，加上快接頭 (QD)，預期單層水冷價值可望達 2,500-3,000 美元，單櫃價值可望達 4.5-5 萬美金。此水冷價值尚未包括 CDU，Google 同時已提出下一代 (第五代) Project Deschutes CDU 規格，散熱能力提升至約 2MW、流量 500 GPM、壓力達 80–90 psi 等級，為未來更高功率密度 AI 機櫃預留散熱能力。由於 v8t 與 v8i 存在晶片與記憶體設計上之不同，因此液冷散熱之供應商也有所區分，聯發科設計之 v8t 之水冷板供應商以奇鋐與古河（日)為主，而 Broadcom 設計之 v8i 版本之水冷板供應商則以 Cooler Master (台灣)與 Boyd (美)為主。市場資訊表示，v7 之 CDU 之設計有採用 1.4MW CDU，而 v8t 開始導入約 2MW CDU，目前常見架構為一台 CDU 搭配約 8 個 機櫃，雖然 CDU 額定容量不能直接視為機櫃實際功耗，但由 1.4MW 提升至 2MW 仍顯示 Google 必須為下一代 TPU 預留更大的散熱能力，此外，v8 的液冷範圍亦由水冷板逐漸擴展至分歧管 (Manifold)、快接頭、Tubing、Rack distribution 及 Optics cooling。

圖 12：TPU v8 系列在每一層 tray 將會有 12 片水冷板



資料來源：公司資料；凱基

圖 13：TPU v8 系列在 16 層 tray 將採用類似設計，並採用第四代 CDU



資料來源：凱基

圖 14：TPU v8 升規搭配第四代液冷（Gen4）CDU

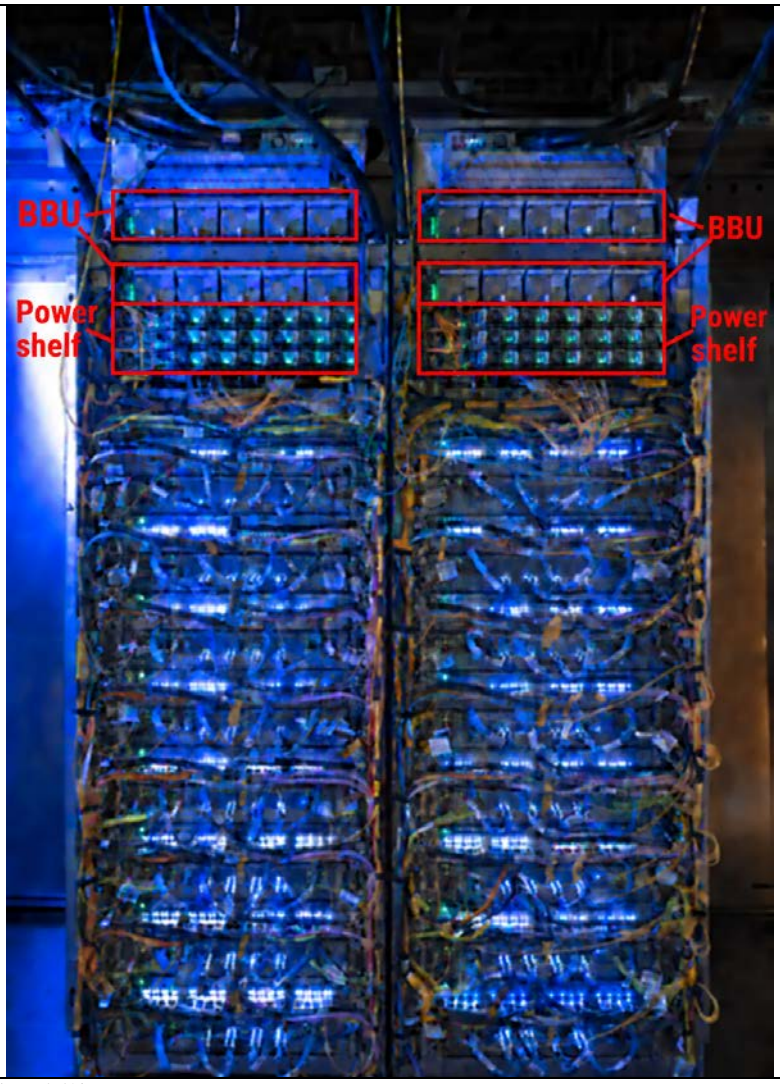


資料來源：Google; 凱基整理

v8 版本開始採用高壓直流（HVDC）電源架構

隨著 TPU 的迭代升級，TDP（熱設計功耗）趨勢向上，v8 版本的電源系統規格也持續提升。最顯著的變化在於相容±400V DC 配置，惟現行 50V DC 方案在本代產品中仍將佔據主流地位。在 50V DC 架構下，預計 8i 機櫃將採用 3 排 Power shelf，內含 6 個 10kW 的 PSU；相較之下，v7 機架則採用 6 排 22kW 的 Power shelf（每個包含 4 個 5.5kW 的 PSU）。我們估算，單櫃的 PSU 總價值將從 v7 版本的約 2 萬美元升至 8i 版本的 2.8 萬至 3 萬美元，增幅達 40-50%。此外，BBU（備用電池單元）和 DC-DC 電源模組預計將繼續作為 TPU 機櫃電源系統的核心組件。我們認為台達電（2308 TT）是±400V DC 電源櫃的主要供應商，目前已開始出貨，並在電源市場享有高市佔率；Flex（美）被視為第二供應商，而光寶科（2301 TT）據信也已於今年成功切入 TPU 供應鏈。

圖 15：電源架構隨 TPU 迭代持續升級，HVDC 初步導入

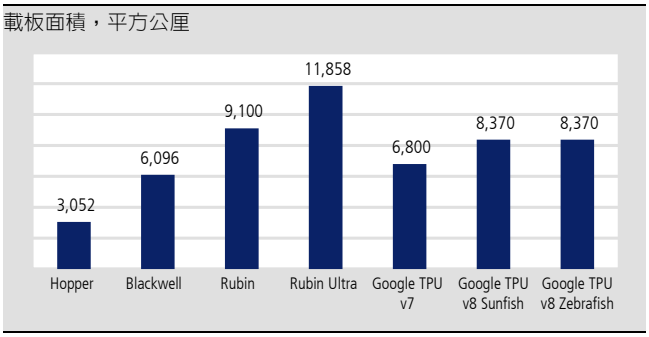


資料來源：凱基整理

TPU v8 載板大尺寸化，PCB 展現質、量雙軌成長

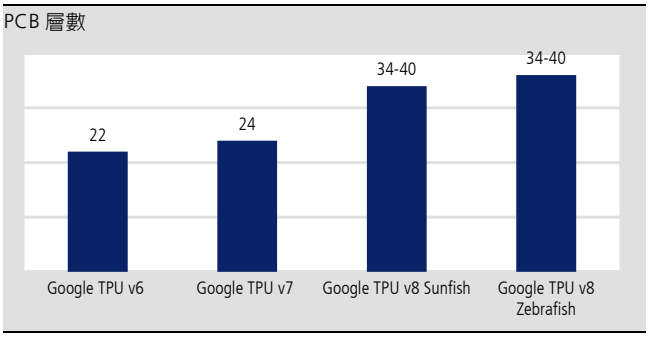
相較於 TPU v7 Ironwood， v8 同時受惠面積、層數、材料三重升級，ABF 載板由 v7 約 80–85mm／16L 放大至 v8 約 90–100mm／18–22L，估 ASP 提升 40–50%；新一代 v8t、v8i 主板 34-40 層，CCL 材料升級至 M8+ 等級並搭配 HVLP4 銅箔，使 CCL/PCB 平均售價增加 70%-80%，我們看好市占率高達 70% CCL 之 台光電、 首度進入供應鏈的金像電(20-25% PCB 市佔率)、金居、ABF 供應商欣興(3037 TT, NT$1,175, 增加持股) 將受惠此趨勢。

圖 16：TPU v8 載板大尺寸化



資料來源：公司資料；凱基預估

圖 17：TPU V8 PCB 走高層數與多層 HDI



資料來源：公司資料；凱基預估

Scale-Up 每顆 TPU 的光學用量正在上升

TPU v7 (Ironwood)為訓練與推論通用設計，採 3D torus 拓樸，每 64 顆 ASIC 以 4×4×4 組成一個 Cube，對應一個機櫃，每個 Cube 對外拉出 96 條光纖；總計一個 pod 有 144 個 Cube，合計 9,216 顆 TPU，故一個 pod 需要 13,824 條光纖、13,824 支光收發模組，以及 48 台 OCS，每顆 TPU 平均光學搭載率(Optical Attach Rate)為 1.5 條光纖與 1.5 支光收發模組。

TPU v8 世代把訓練與推論拆開，兩種晶片採不同拓樸設計。推論用的 v8i，沿用每個 Tray 4 顆 ASIC，但拓樸從 3D torus 換成 Boardfly，可使任兩顆距離最遠的晶片，最多轉 7 次(hops)就可溝通，相較 3D torus 的 16 次，可讓延遲少掉一半以上。

拓樸布建上，每個 Tray 對外使用 5 條光纖，一個 pod 則有 1,152 顆晶片，故一個 pod 需要 1,440 條光纖、1,440 支光收發模組，OCS 預估使用 5 台，Optical Attach Rate 為 1.25。

訓練用的 v8t 則保留 3D torus 架構，每個 Tray 亦是 4 顆 ASIC，惟 Google 尚未公開單一機櫃的晶片組態，此將會決定 3D 幾何如何切分。

研究部產業調研，認為機櫃內以 4×4×2、32 顆 ASIC 為一個互連單元機率高，因此每個機櫃對外將使用 64 條光纖；依此計算，一個 pod 有 9,600 顆 TPU 的狀況下，將對應 300 個機櫃、19,200 條光纖、19,200 支光收發模組、64 台 OCS，Optical attach rate 上升到 2，比 v7 高 33%。

Scale Out 拓樸資訊仍少；v8i 尚未揭露互聯架構，v8t 則在既有的 Jupiter 架構之外，新增 Virgo Network Fabric，可使交換器部署層級數由三層壓縮至兩層。整體而言，TPU 總部署量成長，仍將使 Scale Out 端的光收發模組、光纖線纜與配線盒用量同步放大。

圖 18：各 TPU Pod 的拓樸

	TPU v7 (Ironwood)	TPU v8i (推論)	TPU v8t (訓練)
【Scale-Up 拓樸】			
拓樸架構	3D torus	Boardfly	3D torus
每Tray晶片數	4	4	4
單元幾何	4×4×4	—	4×4×2/4×4×4*
單元內晶片數 (顆)	64	4	32/64
單元對外光纖數 (條)	96	5	64/96*
【Pod 規模】			
Pod實體晶片數 (顆)	9,216	1,152	9,600
可用晶片數 (顆)	9,216	1,024	9,600
備援晶片數 (顆)	0	128	0
【每Pod光學用量】			
光纖數 (條)	13,824	1,440	19,200/14,400*
光收發模組數 (支)	13,824	1,440	19,200/14,400*
OCS 台數 (台)	48	5(估)	64/48*
假設單台OCS埠數 (埠)	300	300	300
【Optical Attach Rate】			
每顆TPU光纖數 (條)	1.50	1.25	2.00/1.50
每顆TPU光模組數 (支)	1.50	1.25	2.00/1.50
Attach Rate vs v7 (%)	-	-16.7%	33.3%/--

** 註：凱基調查 TPU v8t 目前有兩類拓樸方案，最終拓樸方式尚未明確定案*
資料來源：凱基整理

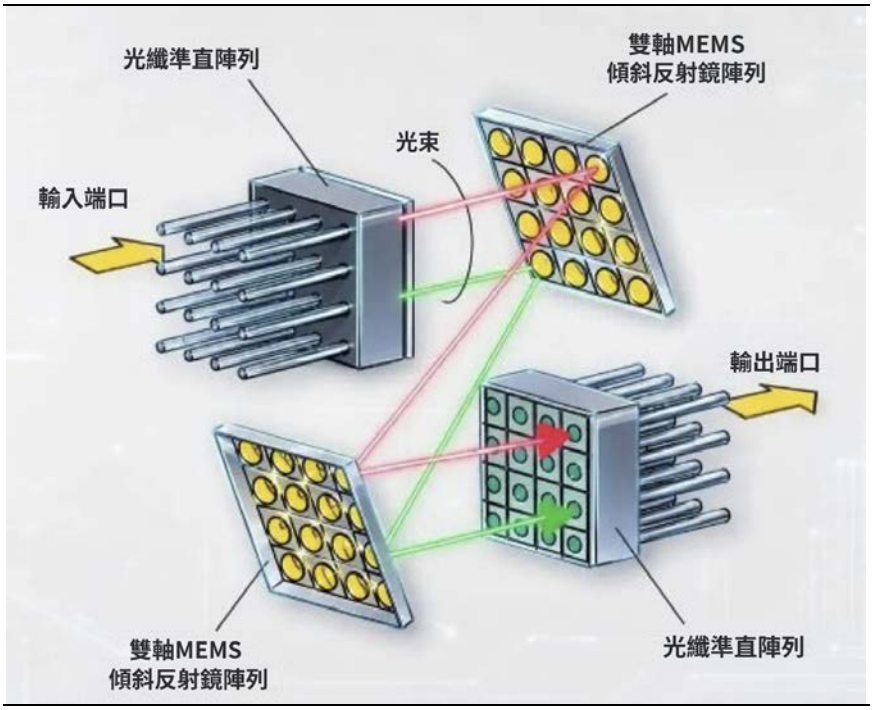
低延遲、低功耗 Google 大量使用 OCS

Google 採用 OCS (Optical Circuit Switch)的原因，是因為 OCS 可直接在光纖埠之間切換光束，不做光-電-光(OEO)轉換，因此可(1)降低傳輸訊號延遲、(2)光收發模組從 400G、800G、升級到 1.6T，不須替換 OCS 本體、(3)省去 SerDes 與 DSP，單埠功耗遠低於電交換器。

Google 使用的 OCS 採 MEMS 技術，主要由輸入光纖陣列、準直透鏡陣列、兩組對向配置的 2D MEMS 微反射鏡陣列、輸出透鏡陣列、輸出光纖陣列等元件構成。MEMS 本身是用半導體製程在矽晶圓上蝕刻出可雙軸偏轉的微反射鏡，施加電壓控制偏轉角度，就能改變光束路徑，完成輸入埠到輸出埠的切換。我們推測 Google 的 OCS 走自研與跟 Lumentum 合作兩條路線，台系廠商在 OCS 本體上著墨有限。

整體而言，我們認為，隨著 TPU 出貨顆數放大，將帶動光收發模組、高密度光纖陣列、MPO 跳線與配線的用量，將有利台灣光纖集線、配線廠商光聖 (主責 Scale Out)、光收發模組內 PD 供應鏈全新(2455 TT)、環宇-KY (4991 TT)、LD 供應鏈聯亞 (3081 TT)、華星光 (4979 TT)、光環 (3234 TT)、聯鈞 (3450 TT)。

圖 19：MEMS 式 OCS 的光路結構



資料來源：Lumentum、凱基預估

圖 20：Lumentum R300 光路交換器

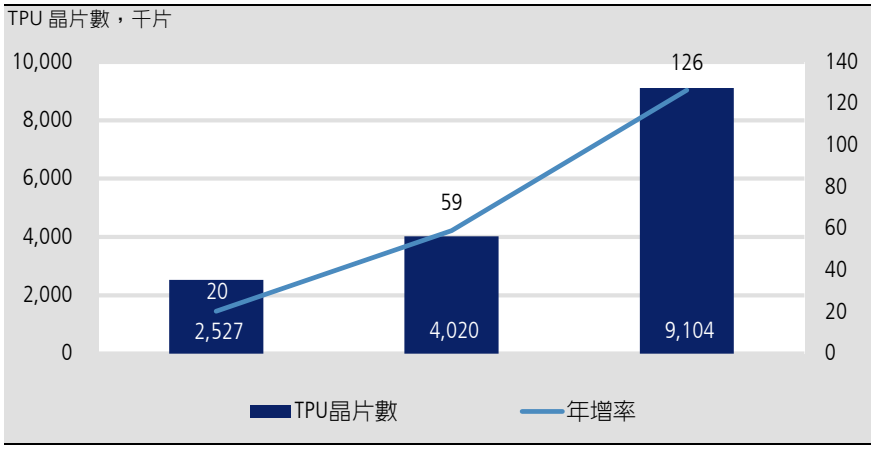


資料來源：Lumentum、凱基整理

2027 年 TPU 晶片翻倍，預計機櫃量亦將翻倍成長

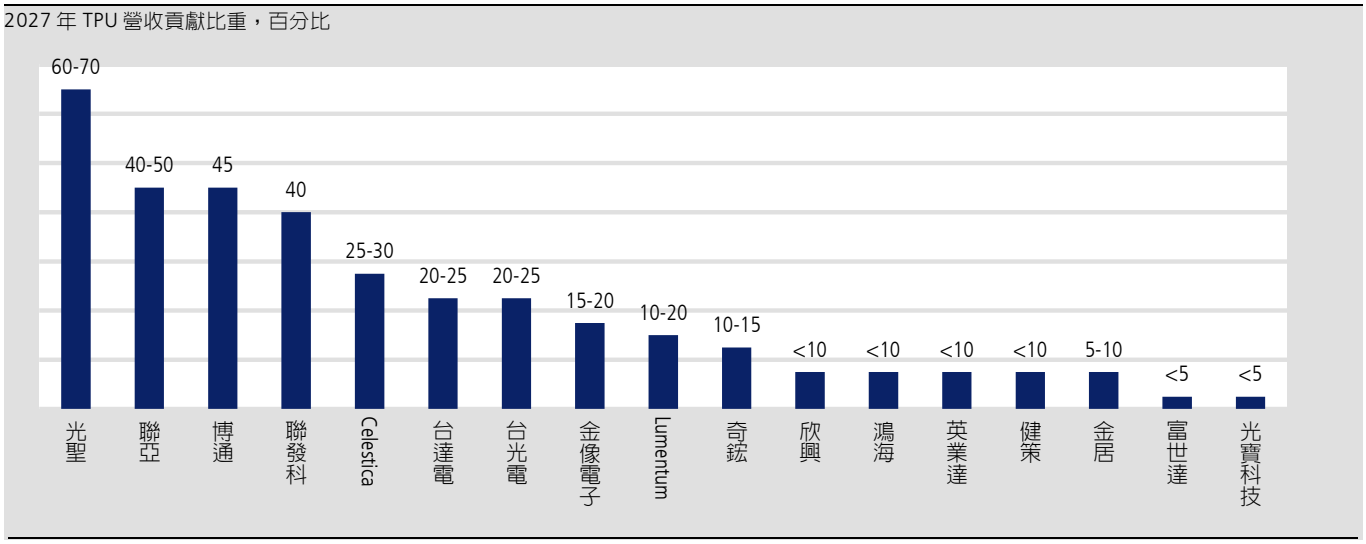
2027 年將是 Google TPU 明顯放量年，從晶片量的角度，預期 TPU v8 出貨量將由 2026 年之約 402 萬顆增至 2027 年之約 910 萬顆 (年增 126%)，其中由聯發科設開發之 v8t 晶片量預計約為 360 萬顆，Broadcom 開發之 v8i 晶片量預期約為 550 萬顆，因而預期 2027 年之 v8 機櫃總出貨量可望由 2026 年翻倍，將同步拉動 TPU 晶片、ODM、散熱、電源、PCB/CCL 與光通訊需求。相關受惠廠商包括: 晶片設計商聯發科（2454 TT，TPU v8t ASIC，2027 年營收貢獻約 40%）與 Broadcom (美, TPU v8i ASIC，2027 年營收貢獻約 45%)、主要 L10/L11 系統組裝商 Celestica（CLS US， 25–30%）、電源廠商台達電（2308 TT，Power Shelf／PSU／±400V HVDC，20–25%）、CCL 供應商台光電（2383 TT，20–25%）、PCB 業者金像電（2368 TT，15–20%）、水冷板新進廠商奇鋐（3017 TT，10–15%）、HVLP4 銅箔供應商金居（8358 TT，5–10%），其他營收貢獻較不顯著之供應商 (小於 10%) 尚有鴻海（2317 TT，機櫃 ODM）、英業達（2356 TT，L6 組裝）、健策（3653 TT， stiffener）、富世達（6805 TT，液冷零組件 QD）、欣興（3037 TT，ABF 載板供應商）與光寶（2301 TT，PSU）。此外，光通訊相關業者應屬 TPU 需求上揚之最大受惠族群，TPU 相關營收貢獻 2027 年皆有雙位數以上，包括聯亞（3081 TT，LD）、光聖（6442 TT，光纖/配線與 Scale-out）及 Lumentum（LITE US，OCS）皆為主要受益廠商。

圖 21：TPU ASIC 出貨顆數預計將於 2027 年翻倍成長



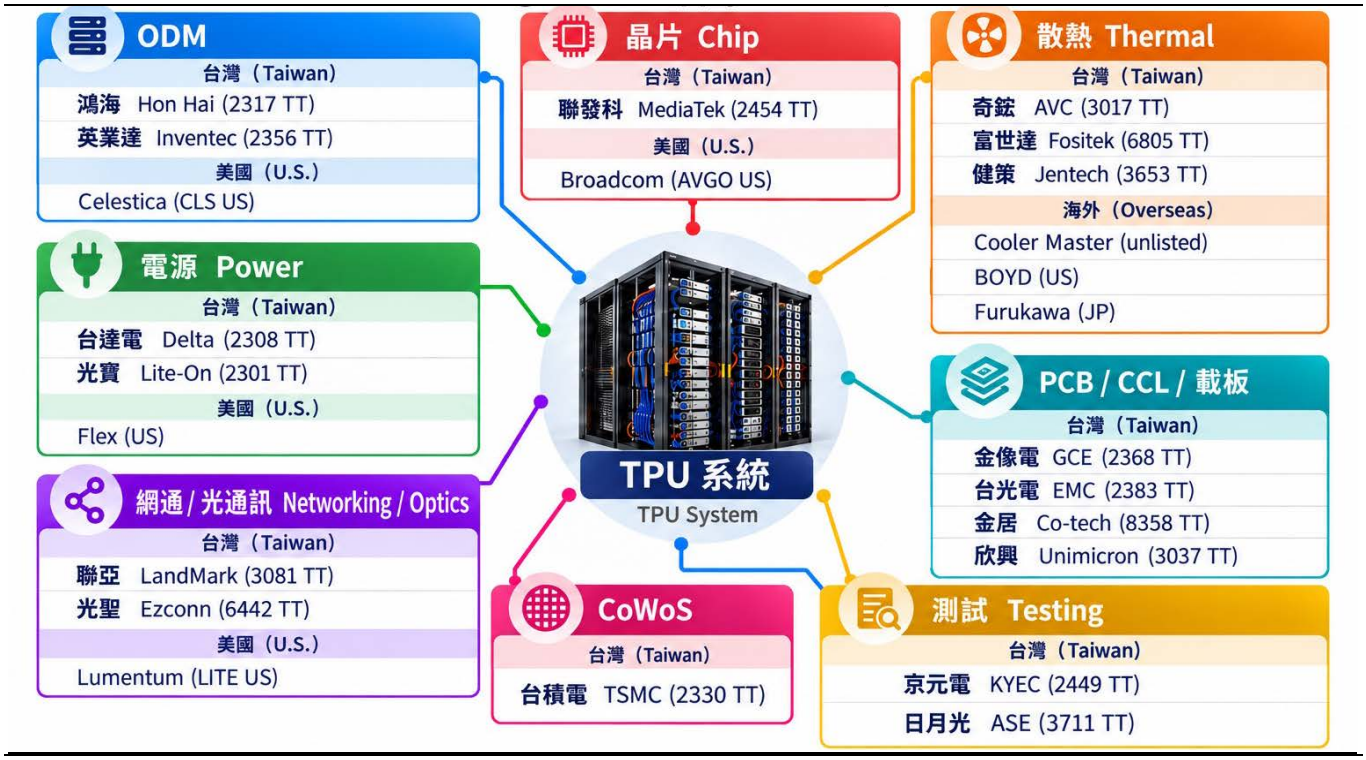
資料來源：公司資料；凱基預估

圖 22：TPU 供應鏈中光聖、聯亞、博通、聯發科、Celestica、台達電、台光電、金像電子、Lumentum、奇鋐等公司占其 2027 年營收比重達雙位數



資料來源：凱基預估

圖 23：Google TPU 供應鏈關鍵零組件與台廠布局



資料來源：凱基預估

凱基證券集團據點

中國	上海	上海市靜安區南京西路 1601 號越洋國際廣場 1507 室 郵政編號：200040
台灣	台北	104 台北市明水路 700 號 電話 886.2.2181.8888·傳真 886.2.8501.1691
香港		香港灣仔港灣道十八號中環廣場四十一樓 電話 852.2878.6888·傳真 852.2878.6800
泰國	曼谷	8th - 11th floors, Asia Centre Building 173 South Sathorn Road, Bangkok 10120, Thailand 電話 66.2658.8888·傳真 66.2658.8014
新加坡		珊頓大道 4 號#13-01 新交所第二大廈 郵政編號：068807 電話 65.6202.1188·傳真 65.6534.4826
印尼		Sona Topas Tower Fl.11 Jl. Jend. Sudirman kav.26 JakartaSelatan 12920 Indonesia 電話 62 21 250 6337

股價說明

等級	定義
增加持股 (OP)	對個股持正面看法，預期個股未來十二個月的表現超越凱基證券集團所追蹤的相關市場的總報酬。
持有 (N)	對個股持中性看法，預期個股未來十二個月的表現符合凱基證券集團所追蹤的相關市場的總報酬。
降低持股 (U)	對個股持負面看法，預期個股未來十二個月的表現低於凱基證券集團所追蹤的相關市場的總報酬。
未評等 (NR)	凱基證券未對該個股加以評等。
受法規限制未評等 (R)	受凱基證券集團內部政策和/或相關法令限制使凱基證券集團無法進行某些形式的資訊交流，其中包括提供評等給投資人參考。
	*總報酬 = (十二個月目標價-現價)/現價

免責聲明

上述凱基分析員為證監會持牌人，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。

於本報告內所載的所有資料，並不擬提供予置身或居住於任何法律上限制凱基證券亞洲有限公司 （「凱基」） 或其關聯成員派發此等資料之司法管轄區的人士或實體使用。此等資料不構成向任何司法管轄區的任何人士或實體作出的任何投資意見、或發售的要約、或認購或投資任何證券或其他投資產品或服務的邀請、招攬或建議，亦不構成於任何司法管轄區用作任何上述的目的之資料派發。請特別留意，本報告所載的資料，不得在美國、或向美國人士（即美國居民或按照美國或其任何州、屬土或領土之法律成立的合夥企業或公司）或為美國人士之利益，而用作派發資料、發售或邀請認購任何證券。於本報告內的所有資料只作一般資料及參考用途，而沒有考慮到任何投資者的特定目的、財務狀況或需要。該等資料不擬提供作法律、財務、稅務或其他專業意見，因此不應將該等資料賴以作為投資專業意見。

部份凱基股票研究報告及盈利預測可透過 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的內部研究活動。本報告內的資料及意見，凱基不會就其公正性、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，凱基並不另行通知。凱基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，並不能在未經凱基書面同意下，擅自以任何方式轉發、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、高級人員及僱員可不時就本報告所涉及的任何證券持倉。