



《天下雜誌》總主筆
陳良榕

雲與端的戰爭 黃仁勳能笑多久？

剛結束的輝達GTC大會，盛況超越蘋果發表會，生成式AI引爆的硬體需求，幾乎都由它吃下。

輝達拿太多了嗎？當算力從雲落到端，局勢會怎麼變？

「這在幾年前難以想像，之前都是終端裝置消費多數矽晶片，」張曉強說，但「感謝AI爆發，局勢已驟變。」

自從ChatGPT在二二年底問世，過去一年間生成式AI引爆的硬體需求，規模之大、速度之迅猛，都是我們這輩科技記者生平所未見。更誇張的是，買氣幾乎集中在輝達一家公司。

因此出現以下驚人數字。輝達的資料中心部門二二年全年營收，大概跟聯發科整家公司的規模差不多，都是四千多億台幣。

結果今年一月結束的最新季度，該部門單季營收超過五千億，較去年同期成長四〇九%。

對此類千億規模的大公司而言，這是史無前

不

久前，我在《天下》辦公室附近

一家人氣零食店大吃一驚，店員竟指著金黃色芒果乾告訴我，「Lori最喜歡吃這個。」

他親熱直呼其名的老主顧，正是輝達台裔創辦人黃仁勳的太太。夫妻倆曾在保鏢陪同下，多次來此採購。對面的豆花攤，也貼了大張的黃仁勳合影。

這是很奇特的感覺。時常在台北街頭趴趴走的黃仁勳，也是每年固定出現在台北國際電腦

展的老面孔。十多年來，他招牌的黑皮衣與結實體態不變，一頭濃黑頭髮逐漸轉為灰白。

這位台灣科技記者熟悉的「鄰家大叔」，竟在年過六十之後，一夕躍成超級巨星。

輝達股價在ChatGPT推出後的一年多時間，漲了四倍多，市值一舉突破兩兆美元，成為世界第三大企業。

在矽谷舉辦，剛結束的輝達GTC年度產品發表大會，竟然湧進一萬多名觀眾，火紅程度

勝過蘋果發表會。這位全球科技業在位最久的CEO，因此被美國媒體封為「新賈伯斯」。

新賈伯斯的魔幻歷程

要理解這個魔幻歷程，可從台積電資深總經理張曉強，三月初在國際固態電路研討會（ISSCC）主題演講說起。他預期二〇三〇年全球半導體市場，以伺服器為主的高性能運算市佔會達到四〇%，大幅超越手機、筆電等終端裝置的三〇%。

例的成長速度。即便是

特斯拉大舉開出平價電動車Model 3產能，改寫汽車歷史的二一年，也不過成長七四%。

這樣的規模和高速成長，能持續下去嗎？

「雲」仍搶購輝達產品

質疑聲四起。許多人擔心微軟、亞馬遜會不會一年買光未來幾年的需求？也就是半導體業常見的暴起暴落循環。前車之鑑是疫情間的恐慌性搶購，之後業界花了快兩年來消化庫存。

但目前看來，買氣仍旺。三大雲端業者亞馬遜、微軟、Google都還在大幅擴張雲端建設——微軟最新會計年度的資本支出高達四六〇億美元，比台積電還高。而且，每一家都表

示來年還要加碼。

這些錢都將用在刀口上——搶購一塊高達百萬台幣的輝達GPU。

為此，亞馬遜還將舊有伺服器的折舊年限，從四年逐步延長為六年，以挪出更多銀彈。

黃仁勳在法說會得意表示，這代表CPU已經過氣，再也不需定期更新、升級。業界大勢「已從（CPU代表的）通用運算，轉到（GPU代表的）加速運算。」這是要搶下英特爾的市佔嗎？法人追問，黃仁勳笑瞇瞇說出金句，「輝達從不爭奪市場，我們只創造市場。」輝達最新一季毛利率高達七六·三%，這不但是半導體業前所未聞的超高利潤，甚至比微軟的六八·四%還高。

雖說創新者的獎賞就是贏家全拿，不過看在客戶眼裡，輝達也拿太多了吧！

「端」開始醞釀反攻

Google DeepMind傑出電腦科學家紀懷新，在參加「天下經濟論壇」時告訴我，現在所有雲端大廠，都希望將AI的算力下放到「端」上面，也就是近來業界喊得響亮的AI手機、AIPC，就不需用輝達GPU，「因為在雲上執行運算實在太貴。」

他解釋其背後邏輯，「任何新科技發展，一開始問的都是What's possible?（有什麼可能性?）」當新科技潛能被充分挖掘，進入成熟期後，「接下來問的就會是What's the lowest cost?（最低成本是多少?）」

生成式AI已進入追求性價比階段了嗎？至少某些領域還沒。眾所期待的生成式AI下一波大浪，是包含視覺、語言及語音的多模態基礎模型，這會消耗遠比現在更大的算力、需要更多強大的GPU。

一位輝達主管也告訴我，AI創新仍源源不絕，「好像西部拓荒一樣，還看不到疆界。」

看來「雲的王者」黃仁勳還能笑上一段時間。

但「端」的陣營，也在醞釀新一波反攻。

根據彭博報導，蘋果已聯手Google，打算將Google的大型語言模型Gemini整合進iPhone。如果順利，二四年底推出的iPhone 16，將成為史上最強AI手機。■