半導體產業學習與七大巨鯨共游

* 2019年10月4日
* Rick Merritt，EE Times矽谷採訪中心主任

那些曾經顛覆商業、改變文化的企業巨鯨，也就是超大規模企業(Hyperscalers)，正開始重塑半導體產業。他們到底做了什麼史無前例的改變？由此產生的後果是什麼？

全球七大資料中心(又稱超大規模業者，Hyperscalers)創造了的龐大晶片市場，推動半導體產業達到了新的性能高度和成本底線。一些人擔心巨頭們分散了晶片廠商的注意力，使他們不再關注更小、更多樣化的用戶的需求。

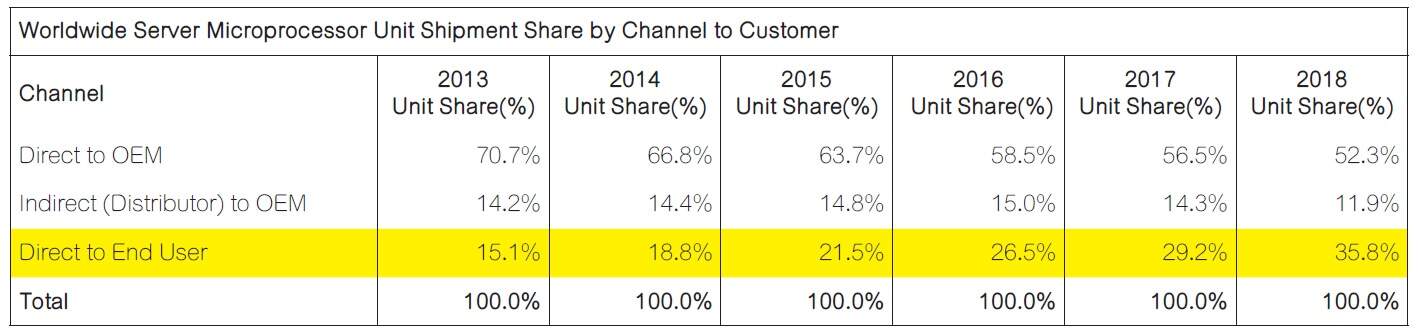
更令人擔憂的是，大多數Hyperscalers都建立了自己的世界級設計團隊，宣稱要成為半導體產業的角逐者。他們已經在設計領先的人工智慧(AI)加速器和智慧型網路控制器，而且剛剛開始發佈開放原始碼RTL，並聲稱這將是一個不斷增長的趨勢。

與這些巨鯨競爭是艱難的，但完全失去這些巨鯨的訂單更加可怕。這極有可能發生，因為歐洲和美國的監管機構現在已瞄準了少數幾個頂級Hyperscalers，威脅要瓦解它們。

然而最有可能的情況是，亞馬遜(Amazon)、Google、微軟(Microsoft)、Facebook、阿里巴巴、騰訊和百度作為晶片客戶同時也是設計師將繼續成長，促使半導體產業得開始慢慢學會如何與這些巨鯨共游。

據思科(Cisco)的資料顯示，2016年Hyperscalers營運了338個資料中心；到2021年，這個數量將達到628個，佔據全球所有資料中心流量的一半以上。據說最大的Hyperscalers安裝的伺服器至少達到300萬台，甚至可能超過800萬台。

根據IDC的資料，透過新的銷售管道，Hyperscalers在2018年直接購買了(主要來自英特爾和AMD)全球35%的伺服器處理器，高於2013年的15%。「例如，選擇x86處理器已使他們成為業界王者，」IDC研究副總裁Shane Rau說。

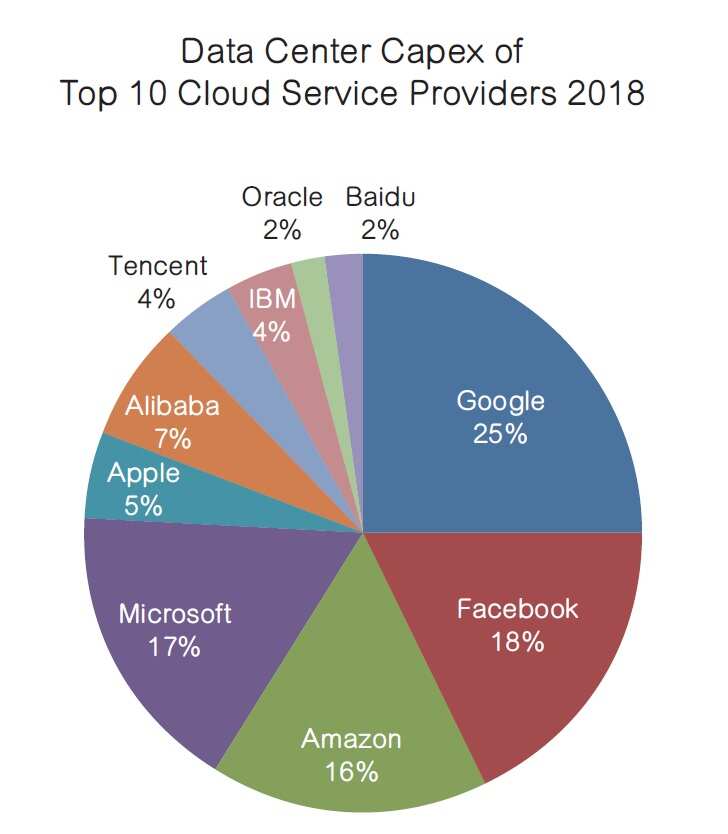
*Hyperscalers為伺服器處理器創建了一個寬廣的新航道。(資料來源：IDC)*

IHS Markit分析師Vlad Galabov表示，總體而言，Hyperscalers管理著所有公共雲端服務約70%的業務，他們購買了大約三分之一的全球資料中心設備。他並提到，亞馬遜、Google和微軟是英特爾最新Xeon可擴展晶片的首批客戶。

這些晶片提供上一代英特爾主記憶體兩倍的容量。因此，Dell'Oro集團分析師Baron Fung認為，Hyperscalers在過去兩年內購買量大增的情況下幾乎壟斷了DRAM市場，導致了記憶體晶片價格的歷史性飆升。

Fung提到，去年，包括蘋果(Apple)、IBM和甲骨文(Oracle)在內的十大Hyperscalers花費了近1,000億美元用於購買資料中心設備，比2017年增加了約45%。650 Group創始人Alan Weckel表示，包括蘋果在內的前八大Hyperscalers去年購買了所有雲端運算設備的71%，總價值達到1,090億美元，而亞馬遜獨佔其中的12%。

「亞馬遜的伺服器安裝基數大於整個電信產業，」同時關注這兩大市場的Weckel表示。相比之下，中國的三家Hyperscalers合起來仍小於亞馬遜，而最小規模的百度大約只佔亞馬遜的十分之一。



*部分人士認為蘋果、IBM和甲骨文也屬於Hyperscalers。(資料來源：Dell'Oro)*

支出看起來很龐大，但收入也很可觀。去年僅亞馬遜、Google和Facebook三家的營收就超過4,000億美元。

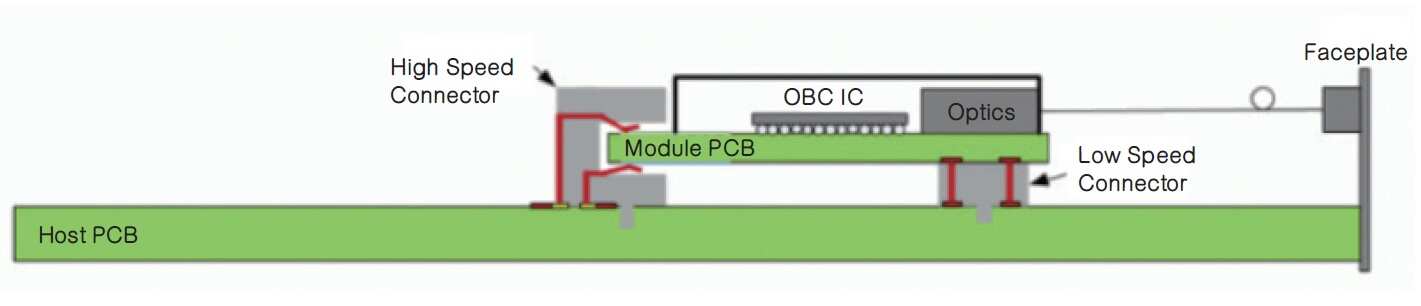
定義銅和光纖網路發展藍圖

由於Hyperscalers的資料中心需要高速連接數萬台伺服器，他們的需求推動了乙太網路路控制器、交換機和光模組產品的採購，以及網路標準的建立。早在2010年，Facebook工程師就要求其供應商開始研究Terabit乙太網路路標準，即便在今天這種標準仍不實際。

微軟Azure雲端服務網路硬體工程經理Brad Booth表示，「十年前，網路巨頭是Cisco和Juniper，他們站在世界之巔。現在，每個客戶都會先找到我們和其他Hyperscalers。」在保持雲端服務領先地位的競爭中，「我們需要不斷評估新技術並在適當的時候介入。」

這就是為什麼在2015年，Booth協助創立了Consortium for On-Board Optics(COBO)。COBO定義了一種方法，可將ASIC和光學元件壓縮到建構資料中心網路主幹的架頂交換機中，這種光學模組降低了交換機功耗與成本，使之輕鬆過渡到當今網路最前瞻的400Gbps交換機。

微軟目前正在評估它的第一個COBO模組，希望於2021年部署它們。但首先，它必須在一個「灰度集群(canary cluster，也譯作金絲雀集群)」中對它們進行測試，該集群使用神經網路應用來驅動高水準的測試流量。

*Hyperscalers定義了一種將ASIC和光學元件結合在一起的方法。(資料來源：COBO)*

追求更高速率的競賽仍在持續進行中。今年稍早，微軟和Facebook宣佈了一項合作，為交換機的下一個重要發展步驟定義元件——將ASIC和光學元件放在同一個封裝內。

「巨鯨們推動了巨大的銷量…但只有少數幾家Hyperscalers可以證明在晶片上進行功能設置和投資是合理的。」博通(Broadcom)產品經理Pete Del Vecchio說道，該公司是網路交換晶片的長期霸主。

Linley Group網路分析師Bob Wheeler提到，近年來，少數幾家擁有創新交換機晶片的新創公司對Broadcom提出了挑戰，但「到目前為止，沒有一個真正能對Broadcom的市佔率造成重大影響。」

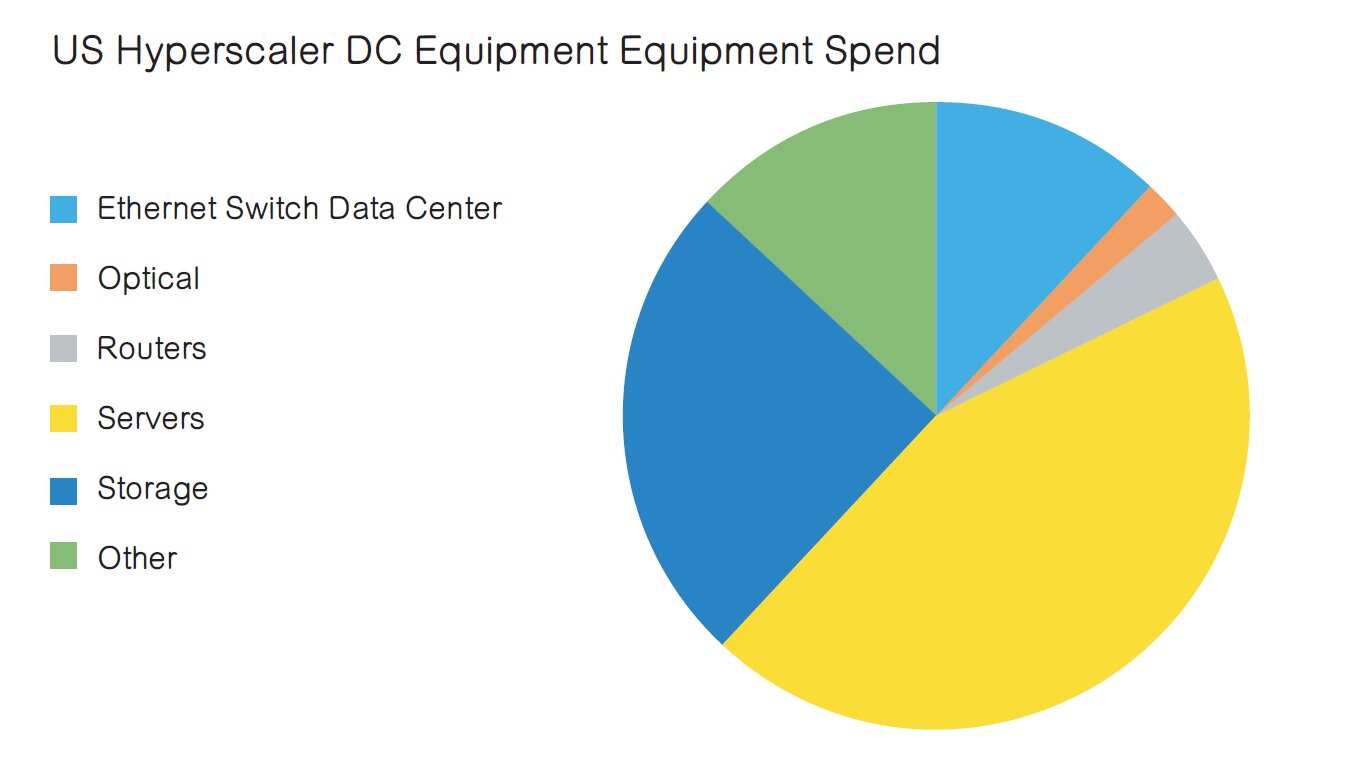
Wheeler解釋，Hyperscalers的「數量驚人地快速增長。十年前，當營運商主導市場時，他們通常每季購買10,000個交換機埠，但現在Hyperscalers每季需要100,000個埠。規模和速度是前所未有的，這對新創公司來說的確是一個挑戰。」

然而，巨鯨們切斷採購的速度也很快。「當我們從10G切換到40G產品時就停止購買10G產品，不拖泥帶水，」Booth表示，而即將演進到400G交換機可能又會讓今天尖端的100G產品沒剩多少市場空間。

Hyperscalers還推進了構成網路端點的伺服器乙太網路路控制器的發展。早在2014年，Booth和其他人就一起創建了一個專門小組，定義了25G和50G規範，以使當時的10G控制器得到更快地發展，而這些在一定程度上破壞IEEE流程的東西恰恰為電信營運商定義了部分40G和100G標準。

如今，微軟已經將其伺服器升級到50G乙太網路路鏈路，其規範幾年前就已不存在，而且微軟已經計畫在幾年內轉向100G控制器。

有些人認為Hyperscalers劫持了乙太網路路標準流程，其推動的需求可能無法滿足傳統業務交換機和伺服器。而且，Hyperscalers更想要快速達到下一個最高速率，這遠遠超過了他們對互用性的需求，因為他們只專注於少數供應商。

*Hyperscalers也是網路和儲存領域極具影響力的買家。(資料來源：650 Group)*

「他們絕對推動了這個產業，」一位不願透露姓名的資深乙太網路路標準負責人表示。「獲得專案批准的最簡單方法就是有Hyperscaler需要它。」

「這種形勢的缺點是Hyperscalers不願意花太多時間去深度開發一個乙太網路路標準。他們引領轉向下一個互連標準的速度比其他公司要快…但是，並非每個人都是大猩猩，大猩猩和小公司需求之間的差距越來越大。」

當大客戶成為小競爭對手時

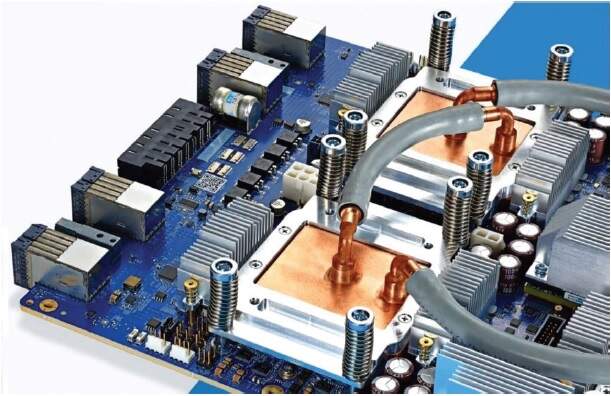
為進一步提升網路性能，亞馬遜則採用了不同的策略，其收購了新創公司Annapurna，從而得到用於處理網路通訊協定的智慧控制器。Galabov說，「亞馬遜在英特爾或其他公司那裡沒有找到滿意的乙太網路路元件，所以購買了擁有基於Arm核心晶片的以色列公司Annapurna。現在，亞馬遜已經成為排名前五的乙太網路路適配器供應商，」僅次於英特爾、Mellanox、Broadcom和Marvell。

Annapurna被亞馬遜收購時約有200名工程師，這個數字今天應該已多了一倍。「四大網卡供應商的地位一直都沒有變過，亞馬遜迅速成為第五大乙太網路路適配器供應商，」Galabov說，估計亞馬遜內部設計現在已佔全球乙太網路路控制器埠和收益的8%。

為了與Hyperscalers建立更深厚的聯繫，Nvidia也出價69億美元意欲收購網路專家Mellanox，這是其迄今為止最努力的一次收購。Fung表示，Mellanox三分之一的銷售額都來自大中型雲端服務供應商，其產品已成功應用於阿里巴巴、百度、Facebook和騰訊等公司。

Gabalov說，我們期待更快的轉變。當運算工作負載發生變化時，Hyperscalers是最先感受到的，為了與對手競爭，他們需要迅速做出反應。而這正是隨著深度學習的興起衍生而來，2012年，新型AI應運而生。

跟隨著形勢的發展，亞馬遜、阿里巴巴和百度已經發佈了用於深度學習的獨立加速器，部分將於年內部署。而Google擁有三代TPU且已在其網路中運作，處於領先地位。



*Google最新的TPU運行速度非常快，需要液體冷卻。(圖片來源：Google)*

「憑藉TPU，Google帶動了整個AI加速器產業，」Rau說。「當Hyperscalers找不到他們想要的東西時，同時又財力雄厚且有高性能需求，他們會自己來設計，」他說。「對於半導體產業來說，以更低的成本提供標準元件是一項很艱巨的挑戰。」

對於許多新創公司和已建立的晶片製造商而言，上述緊張局勢正引發他們對推出第一批AI加速器的焦慮。他們擔心巨鯨級企業內部開發的晶片可能就已滿足市場預期的最大需求。

長久來看，深度學習會得到廣泛應用，私營企業也將轉向雲端服務巨頭來訓練其神經網路，因為世界上最大的處理器群要完成這項工作也可能需要幾周時間。「如果發生這種情況，我們會看到運算將更加不均衡地湧向Hyperscalers，」Weckel表示。

Weckel擔心巨鯨們會在某種情況下(由於中美貿易戰而變得更加複雜)打造各自的半導體供應鏈，在這種情況下，任何一家半導體公司的產量，都不能產生可觀的收益。「如果Hyperscalers繼續投資自研晶片，他們可以直接與三星和台積電這樣的代工廠合作，」Galabov說。

Hyperscalers「將在一段時間內生產自己的ASIC，但他們終究會感受壓力，而回來採用更節省成本的解決方案，」Rau說。

「情況並不像看起來那麼糟糕，」Linley Group總裁暨分析師Linley Gwennap說。亞馬遜和其他公司目前都只是討論過推理，而沒有開始訓練ASIC，許多項目可能根本不會見諸天日。「更多可能是他們沒有得到他們想要的東西，而不是他們真正想要設計自己的晶片。」

來自Facebook硬體團隊的觀點

最糟糕的情況是，Hyperscalers將繞過晶片供應商，就像十年前他們拋棄戴爾(Dell)和惠普(HP)等伺服器製造商。Google是首批上市伺服器的廠商之一，其伺服器規格毫無花俏之處，該公司直接與廣達和緯創等原始設計製造商(ODM)合作，跳過了品牌原始設備製造商(OEM)這些不必要的盈利中間商。

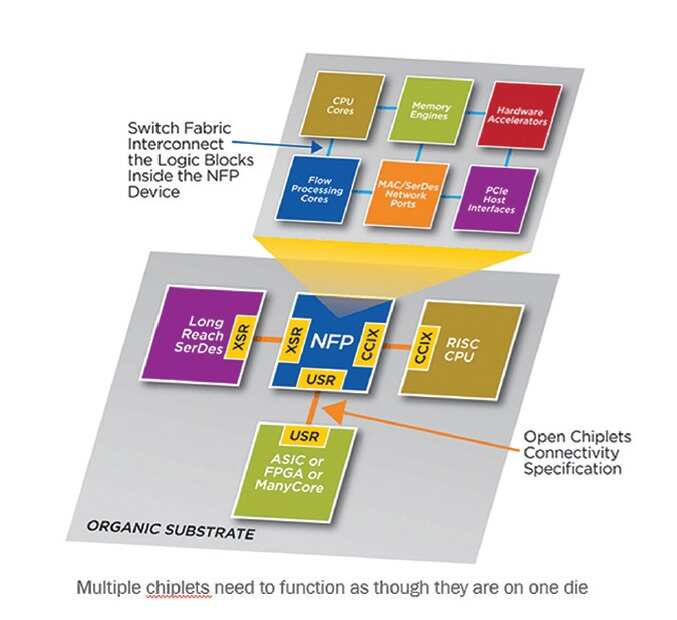
結果是伺服器市場「變得更加殘酷」，而中國的華為和浪潮已經躋身前六位，Weckel表示。

從一定程度上講，這種情況已經發生了。「我們已經開始直接購買光學模組，而不是從交換機供應商處購買，因為這中間支付的差價太高了，」Booth表示。

微軟邁出了更大的一步，於今年3月發佈了一個新的開放原始碼RTL資料壓縮標準，為Hyperscalers定義了系統規範。這個資料壓縮標準是基於2011年Facebook創建的開放運算計畫(OCP)。

微軟透過Facebook於2011年創建的OCP發佈了一個新的開放原始碼RTL資料壓縮標準，為Hyperscalers定義了系統規範。微軟伺服器工程總經理Kushagra Vaid說：「在雲端服務供應商中，我們首先…開創了推動RTL的新先例，希望其他企業也會跟隨我們的腳步。」

OCP的另一項努力是為chiplets定義開放標準，許多人認為隨著摩爾定律的放緩，這將是半導體設計的未來。OCP成員之一，晶片供應商Netronome，將致力於為集團提供其多核心網路處理器中使用的800Gbps結構RTL。

*Open Compute Group旨在為chiplets定義開放規範。(資料來源：OCP)*

另外，Google是RISC-V革新的早期支持者，其發佈了許多新處理器核心指令集的開放原始碼。

長遠來看，開放原始碼IP模組和工具的不斷積累可能會改變晶片產業，但這一天不會很快到來。

開放原始碼硬體「與開放原始碼軟體一樣具有潛力，關鍵區別在於軟體可以從零開始而硬體不行，」Facebook新硬體部門硬體工程總監Aaron Sullivan說。「對於硬體，你必須讓其上運行的東西工作，這不像將開放原始碼編譯成二進位那麼簡單。」

「我們不具有構建EDA工具和IP的核心競爭力，」Facebook基礎架構總監Vijay Rao表示，在加入社群網路之前的15年時間，Rao一直在AMD和英特爾設計處理器和EDA流程。「在Facebook，我們著眼於創建更好的基礎架構，這是我們最關注的。當然如果需要構建更好的EDA工具或晶片，我們也會去做。」

「對伺服器OEM的去仲介化並不需要太多前期投資，反而設計資料中心晶片是一項巨大的投資，而且隨著向7奈米及以下製程的推進，投資會越來越大，因此Facebook更願意宣導而不是設計自己的晶片。」

例如，「我和20家公司進行了交流，可市場還沒能為資料中心提供良好的視訊轉碼器，因此我們向市場提供了創意並協助公司來製造它們，Broadcom和Verisilicon就是我們合作的兩家公司，」Rao說。

同樣，Facebook與Broadcom、高通(Qualcomm)，以及其他公司合作，定義了推理處理器。他補充，Facebook積極幫助設計其編譯器而不是晶片本身。「如果由於數量太少廠家不願意生產，那我們可能會自己做一些晶片。我們有AR/VR產品、Telco Infra專案，可能會需要一些EDA流程。未來有太多的事情我們可能選擇去做，包括內部製造或外部製造的東西。」

Booth對此表示贊同：「我們有工程師團隊定義了(系統級)開關、光學零件和元件。思科有數千名工程師負責交換機設計，但微軟並沒有聘請這麼多工程師，對我們來說，採用OEM產品或白盒(White box)系統更划算。」

七條巨鯨航線各不相同

儘管本文將Hyperscalers組合在一起，但他們在社交網路、電子商務和媒體中的業務和工作量各不相同。「他們彼此競爭激烈，因此需要創新和差異化，」Galabov說道。

例如，「Hyperscalers推動了與英特爾合作的新方式，」他說。「英特爾過去並不習慣定制CPU的模式，但到2018年，它向雲端服務客戶銷售的一半CPU都是定制的，高於2013年的20%，英特爾必須學會這項全新的能力。」

光通訊也出現了這種趨勢，比如Google採用CDWDM4規範連接其資料中心。而相對後來者Facebook和微軟則採用更廉價的單模光纖。「微軟快把供應商逼瘋了，他們希望我們所有人都有共同的願景和方向，」Booth表示。

在努力為每個巨頭製造獨特的產品時，半導體公司們已經忘記了傳統的行銷規則。根據Booth的說法，他們停止「整合多個市場的需求，並為所有市場打造同一個產品；相反，他們在拆分自己的市場。」這是一個奇怪的轉變。

另一些人則指出，超過12種乙太網路規範已經出爐，這是標準化工作中出現碎片化的一個例子。

AI估計也會和Hyperscalers的不同需求一樣，被分成不同的目標行業。「託管的加速與Facebook的加速非常不同，Facebook目前還沒有託管的計畫，」Rao說。「我們對我們想要的東西有深入瞭解，希望構建解決不同問題的加速器。我們的模型規模與其他Hyperscalers有很大不同，模型類型和參數也大不相同。」

Rao的團隊已經與大多數AI晶片初創企業進行了交流，「他們並非都主攻同一領域，有些開發邊緣訓練，有些開發類似於Alexa的設備，這解決了問題的不同部分。」

Booth贊同了這一點。「微軟想要做的機器學習可能與Google或亞馬遜想要做的有很大不同，現在，其中一些更是區分我們的AI即服務(AI-as-a-service)的秘密武器，所以，不太可能製造一台設備賣給所有人。」

有一件事情是所有Hyperscalers從始至終都專注的，那就是想盡辦法節省每一分錢。「晶片供應商很難真正理解我們所致力的營運規模，」Booth說，6年前他加入微軟之前，曾在英特爾、Applied Micro和Dell工作過。「那些看起來不起眼的、功耗只有半瓦特的東西，正呈爆炸式增長，聚集在一起消耗著兆瓦級的功率。」

*Google的數十個資料中心之一。(資料來源：Google)*

監管機構會讓巨鯨滅絕嗎？

資料中心巨頭們如今還面臨著許多威脅，在2016年美國總統大選中被披露隱私洩露及其作為「馬前卒」(編按：指為他人奔走效力的人)的醜聞之後，公眾情緒已經轉向他們的對立面。此外，歐盟委員會、美國司法部和聯邦貿易委員會也正在進行或考慮反壟斷調查。

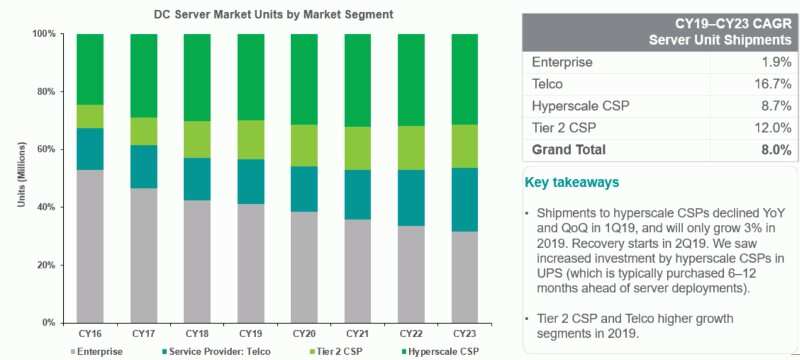
作為審查的重中之重，Facebook執行長馬克·祖克柏(Mark Elliot Zuckerberg)已經做出了許多承諾。該公司正朝著保護隱私和防止濫用社交網路進行營運的新方向發展，但其龐大且高度自動化的網路仍然容易洩密和受到攻擊。

有傳聞指出，一些跡象顯示Hyperscalers已經跌落神壇不再是媒體的寵兒。一位分析師指出，那些有前途的畢業生越來越多地認為他們和矽谷一樣，不再是那麼有吸引力的雇主了。

但是，儘管草根階層發起了關閉Facebook帳戶的運動，這家社交網路巨頭和其他hyperscalers仍在繼續成長。人們無法預測未來的選舉、安全性漏洞和監管制裁可能會帶來什麼，但有經驗的投資者仍押寶巨鯨將繼續發展。

晶片製造商今年已初嘗沒有Hyperscalers的生活會是什麼樣子，當時巨頭們在經歷了兩年的強勁增長之後縮減了開支。Rau預測從今年秋季開始，對hyperscalers的伺服器處理器整體銷售將回升至約3.5%。

Galabov表示，在接下來的5年中，Hyperscalers的伺服器消費量在單位和收入上會達到7.5%的增長。他補充，規模小得多的二線雲端服務供應商和電信營運商的伺服器支出增長速度將更快。

*Galabov認為電信營運商和二線雲端供應商的增長速度超過了Hyperscalers。(資料來源：IHS Markit)*

Fung更加樂觀，他預測Hyperscalers將在其十幾歲的青少年期恢復增長。很多跡象說明，許多公司正持續將更多的工作轉移到AWS、Azure和Google Cloud等公共雲端服務上。

巨鯨們掀起的下一次巨浪可能最快在明年到來。因為他們會發掘新的伺服器平台、更快的網路和第一批AI加速器。建議晶片製造商們準備好他們的游泳裝備迎接更大的浪潮。

在美國西部半導體展(Semicon West)上，Google的一位工程師說，這家搜尋巨頭為了追求更高的AI性能，已經徹底改變了一切，從加速器到互連和浮點格式，他並邀請晶片製造商加入對驅動第四代TPU(或許是一種全新的電晶體)研究的行列中來。

這只是Hyperscalers風頭正勁的眾多邀請之一，半導體產業需要更積極參與。

(參考原文：[Semiconductors Swim with Seven Whales](https://www.eetimes.com/document.asp?doc_id=1334931&page_number=1)，by Rick Merritt)

**本文同步刊登於EE Times Taiwan 10月號雜誌**