

# 爭搶製程商機 八仙過海各憑本事

■文：任莖萍



照片人物：力積電 (PSMC) 董事長黃崇仁

同樣基於整合理念，也有人將念頭動到了記憶體身上。力積電 (PSMC) 董事長黃崇仁提及，從 1G 到 5G 通訊約莫走了四十個年頭，資料速度提升了 800 萬倍，半導體要飛速追趕，熱量與頻寬成了最大挑戰，加之運算又浪費太多時間在資料進出記憶體上；反之，即使擁有大容量記憶體、運算速度卻跟不上，就會需要大量緩衝區 (buffer)。為翻轉此困境，力晶突破製程壁壘將邏輯 IC 和記憶體整合在一個基板上，採用 25nm DRAM 製程做成名為「AIM」的單晶片，特別適用於 5G 和 AI 等特定用途的高速邊緣運算，具備以下

三大特點。

一是省略 I/O、沒有頻寬顧慮，二是可內嵌四個 CPU 或 MPU 並允許混搭 ARM 和 RISC 架構，四通道合計最高傳輸率可達 4096 bit/s，三是功耗和成本遠低於張量處理器 (TPU) 或智能處理器 (IPU) 等新型 AI 加速處理器。黃崇仁主張，神經網路涵蓋輸入、運算和儲存，DRAM 若沒有成功連接處理器就沒用，就像人腦有千億個神經元、需要千兆連接；如果前段 DRAM 故障就會罹患阿茲海默症，若是後段硬碟損壞就是帕金森氏症。他堅信，再好的應用都需要半導體支持，否則都是空談，並透露正在開發 AI 感測器，可就地運算數據再丟出去，降低負載。

## 明日製程重點：MEMS 感測

日月光半導體執行長吳田玉接話，製程關注的不外乎就是密度和功耗，摩爾定律 (Moore's Law) 講述的只是處理器 + 記憶體的大腦運作，缺少對眼、耳、口、鼻、手的論調，但諸如微機電 (MEMS) 等感測技術才是明日重點。然而，摩

爾與泛摩爾沒有衝突、而是相輔相成的，因為腦子好，所有感官會更靈光；未來六十年，大腦會繼續循摩爾軌跡前行，感測器則將從穿戴外置，轉成嵌入在感官或大腦中。他坦言，雖然半導體產業在過去三十年形勢大好，幾呈同步線性成長，但它其實是很殘忍的行業，餅很大、卻未必吃得到，只適合「極端樂觀者」投入。



照片人物：日月光半導體執行長吳田玉

吳田玉的看法是：製程升級確會帶動營收節奏，但嚴格來說並非由摩爾定律直接牽引，而是先帶動系統發明、為社會創造經濟效益和價值並產生大規模量化效益使然；何況，今天的摩爾定律已發生質變！從業者必須先弄清楚日後生

意打哪兒來？要在什麼時間點切入？今後三十年，有哪些應用可維持十年成長榮景？找對題目、設立願景，才有機會大啖商機。他認為，將 CPU 與記憶體綁在一起只是摩爾定律的延伸，不算「異質整合」(Heterogeneous Integration, HI)，不須過於執著此道；真正的異質整合必須是「摩爾完全沒考慮到的東西」，且能發揮槓桿作用。

## 3D 堆疊起飛，「量子運算」備受期待



照片人物：旺宏電子總經理盧志遠

3D 堆疊封裝是另一個製程焦點。旺宏電子總經理盧志遠拆解，2D 平面因閘口相距太遠，難以控制電流洩漏路徑，所以往 3D 鰭式電晶體 (FinFET) 發展，其優勢在於：邏輯閘控制較佳，實作可高達 192 層，2014 年後的多層單元 (MLC) NAND 封裝多採 3D 立體封裝。小晶片 (Chiplet) 模組化是另一值得留意的趨勢，不過也有人反其道而行，將整個 300mm 的晶圓做成一個超大晶片，成功率不是

零、就是一百，根本沒有良率可言；無論尺寸如何，都少不了錯誤修正 (ECC) 技能相助。此外，盧志遠相當看好「量子運算」(Quantum Computing) 前景。

IBM 早在 1970 ~ 1980 年間，即對超導電性著手研究，隨著近日 53 量子位元 (Qbit) 的電腦問世，量子運算再度引發關注，台灣大學也正積極籌組「IBM 量子電腦中心」。統整專家意見，多只在光電和通訊兩類半導體器件實施量子點 (Quantum Dot) 超導電路，而建構量子演算法須先弄清楚代碼，因為每個設備都必須將量子電路映射到量子設備系統，使所有電子設備更接近量子晶片；結合特殊 CMOS 在極低溫環境工作可降低傳統晶片的嚴重電阻，對於啓用量子電腦的電晶體至關重要，亦是梗阻所在，須仰賴最大數量的門控精確度 (Fidelity) 警告。

## 「世界工廠」成過去式，供應鏈轉趨陣營化

半導體創新風向也吹拂到系統商。為打通電子業上、下游經脈做垂直整合，鴻海集團三年前開始進軍半導體領域，以實際行動展現轉型「解決方案提供者」的決心。S 次集團副總經理陳偉銘表示，受限於製程，IC 設計須與晶圓廠緊密合作、無法「為所欲為」，設計規則已從最初的數百增至上萬，現在更向下延伸到封裝和系統並尋求跨域突破；有些系統商更進一步發展專用晶片 (ASIC) 以強化自己



照片人物：鴻海 S 次集團副總經理陳偉銘

的生態系統，而軟體可能成為創造差異化的關鍵：相較於硬體，演算法對改善運算效能的幫助更大，甚至出現軟體定義晶片 (Software Defined Everything, SDX)。

陳偉銘指出，在保護主義、貿易戰盛行下，每個地區都希望發展屬於自己的產業，不太可能再出現世界工廠，供應鏈亦將一改通力合作之勢、轉趨陣營化，產品規格也將涇渭分明，只有資源豐沛的大集團有能力同時投入多種研發，而半導體和 AI 的進展將決定國家優勢。他剖析，早年 PC 時代，台灣有許多系統 ODM 廠、建立完整產業鏈，是促進半導體蓬勃發展的成功因素之一，才得以造就今天台灣在晶圓代工和委外封裝測試廠 (OSAT) 穩座全球第一、IC 設計位居第二的榮景；面對全球產業轉型，延伸完整供應鏈是最佳策略。

## Facebook : AI + VR/ AR 穿戴式裝置將繼起， 無線技術能耗驚人



照片人物：Facebook 副總裁兼直升機與技術工程負責人 Shahriar Rabii

今年在國際半導體展 (SEMICON Taiwan 2019) 系列活動中，有一場《智慧數據國際高峰論壇》難得請到臉書 (Facebook) 副總裁兼直升機與技術工程負責人 Shahriar Rabii 分享所謂系統商的觀點。他表示，此刻的我們正以史無前例的便捷方式進行交流並獲得數位資訊，過去十年來，我們見證了智慧手機和個人電腦的革命；未來，提供虛擬實境／擴增實境 (VR/AR) 的智慧眼鏡等穿戴式裝置將繼之而起。例如，記不起某人名字時，智慧眼鏡可用來增強人們的記憶力和認知能力；在境外有飲食需求，翻譯眼鏡可在共享白板上做出健康的飲食選擇。

以智慧手機 AR 所做的所有事情都會附加基於位置的訊息，可與世界各地的朋友在互動中共享私人便箋。Rabii 認為，AR 眼鏡可增強我們的感知力，路由器將

連接虛擬與現實，讓人們無縫、高效地與世界接軌。這些功能需要獲得足夠高、在日光下也能清楚分辨的高解析度顯示器，且須用圖像識別感興趣的個體後、針對註解 (annotation) 和舞動加以渲染 (render)，融會貫通任務的上、下文意識，以提供最佳用戶體驗。我們當然不希望每次進行查詢或命名時都遭遇漫長等待，想要順理成章、泰然自若地回應，需要始終處於上、下文感知狀態的 AI 協助。

### 減少數據傳輸量 vs. 功耗的兩難

AR 可看作是具有完美記憶力的小助手，透過互聯網連接直達眼、耳感官，目標是在設計點、乃至整個過程中以可接受的功耗提供堪用性能並優化嚴苛的外形規格限制，以打造輕巧而時尚的眼鏡，同時將皮膚溫升降至最低以確保用戶舒適度。為快速連接圖形架構和顯示技術，需為佩戴者在其他任務暗置燈光、音訊以追蹤手勢，接著將用戶位置在地圖上定位、計算 3D 對象並分析工作量。若在傳統系統單晶片 (SOC) 架構運行乃基於不同操作的能耗實現：高功率可增加數據傳輸速度，但最終無線技術能耗將提高 12,000 倍、甚至 300,000 倍以上。

因此，「減少數據傳輸量」對節能有明顯助益，需要使用閾值電壓的混合來啓動 1V 電壓，避免時間漏電功率下降、並迅速增加設備數量以保持泄漏螺栓 (bolts of

leakage) 供應數量的電晶體遷移；但弔詭的是，這會導致更高的功率、設計難度更大。由於定義動態功耗為主的工作負載經常變化、且執行接近閾值的電壓，AI 設備的槓桿效應非常大、屬於數據傳輸密集型架構，加速器可為優化 AI 工作負載提供巨大機會，且有助於獲得重要聲音；接收到語音後會進入喚醒詞檢測和自然語言等子任務，這是一個永遠在線 (always-on) 的模組，需以超低功率運行。

假如相關資源也可用於管理其他應用，更有利於訓練 AI 模型；又此類模型的工作精度可低至只有兩位，能顯著降低功耗。另一方面，執行、配置文件對電腦和記憶體的要求亦截然不同，此時，加速器可發揮作用——藉由在數據處理附近使用高密度的記憶體選項找到平衡，讓運行卷積神經網路 (CNN) 減少能量消耗並提高錯誤率，節省約 26% 的 SRAM。Rabii 總結，在技術定義的過程中可能會採取不同途徑並提供不同功率和性能；此前，總能量或因製程技術設計出現差異，這可能是經由材料和物理方法導入革命性的新佈線而產生，效果極佳。 **CTA**