

## 國科會工程處

### 115 年度「高效能晶片關鍵技術與創新應用計畫」

#### 分項一：高能效與高速運算晶片

##### 壹、 計畫背景及目的

隨著人工智慧與各式前瞻應用的開展，未來下一世代的電子產品與技術的關鍵在於高效能晶片系統，而製作與設計高效能晶片所需的相關技術為重中之重。因此，本工作項目主軸在於強化我國學術界所研發之矽基半導體晶片與系統的運算效能，結合 2D/3D 記憶體的整合設計與先進製程技術，能夠達到並進一步超越高效能運算(High Performance Computing, HPC)的指標。國內學術界雖然受到經費及人力資源的限制，因此無法使用業界半導體高階製程晶片，但在各領域均有傑出人才與充沛經驗，計畫團隊仍可發展相關技術領域，透過本計畫蓄積充沛研發能量與相關技術成果，以期結合邏輯與記憶體的整合設計，同時開發並使用先進封裝技術，建立製作高效能晶片系統所需的異質整合平台研發能量，能為國內持續提供高效能的運算晶片做出貢獻。

本工作項目將規劃三年為階段性推動目標，計畫內容包括系統架構、運算與高速傳輸關鍵設計、軟硬體技術整合、異質整合與實體驗證等領域。計畫年度目標的制定可依照計畫團隊之專長及預計執行進度以模擬、技術開發、軟硬體整合、測試載具及實體驗證漸次展現。

關於本計畫的預期關鍵成果，包括高階模擬分析技術的開發、使用測試載具實質展示異質整合平台使用於高效能運算系統所需的 2.5D 關鍵技術，以及運算及高速傳輸關鍵晶片下線與系統模擬；預期在計畫結案時能有 HPC 離形系統展示，而針對大型語言模型(>10B Model)的計算力能夠達到每秒 1 POPS 以上，滿足大型語言模型高維度與超大權重參數的運算需求。

希望透過本計畫的執行，掌握關鍵的運算及高速傳輸關鍵電路設計，與實現異質整合平台，並能夠有效的跨領域整合所需技術，完成軟硬體的協同整合，依序以模擬或是實體驗證的形式，完成 HPC 系統架構規劃、HPC 系統軟硬體的偕同驗證、高效能運算的半導體系統晶片、HPC 系統應用的開發，並展示能

處理大型語言模型的高效能運算解決方案。

## 貳、研究議題範疇

原則上以 16/7 奈米製程評估高算力晶片方案，研究主軸包含但不限於下列領域：

1. 高算力晶片架構與電路設計：採用同質(單態樣運算核心)、異質運算(多態樣核心，如 CPU+GPU+AI Engine)的整合設計，包含內建處理器(ARM, RISC-V, ...)以及客製化的硬體加速器(AI Accelerators)，並具有可延展性的實現方案，針對 8bit(定點或浮點運算)的大型語言模型(>10B Model)或小模型(<10B Model)邊緣 AI 應用需求的運算需求，達成全程 POPS 的算力指標，同時考慮前述電路與系統之先進散熱技術與封裝。
2. 整合高頻寬記憶體的高算力晶片模組：探討不同記憶體模組(DDR4/5, HBM, AIM, ...等)及軟硬體整合，有效提升整體運算效能，並以 MLPerf 的評比指標展現系統效能。
3. 超低節能的資料傳輸解決方案：透過異質整合的堆疊技術，包含晶圓層級的封裝與小晶片的可程式化先進封裝，滿足大型語言模型運算所需求的超大頻寬與節能傳輸需求。

## 參、計畫重點

隨著新興應用的蓬勃發展，高運算技術受到高度的重視，本分項的研發目標希望由系統層面考量，來提升整體的運算力、能源效率及記憶體頻寬。技術開發的項目，除了引進 FinFET 製程外，電路及架構設計的創新，以及 2D、2.5D、3D 異質整合技術等，都是發揮最大效能及算力之關鍵技術，能由不同層面切入來突破現有的瓶頸和限制，達到預期的技術指標。除此之外，也期待研發團隊能強化產學與國際合作，加速前瞻科研發展，甚至於促成創新創業、生活應用，以提升國內晶片設計產業的發展及競爭力。

## 分項二：高功率、高傳輸電路與晶片模組

### 壹、 計畫背景及目的

高效能的運算需要有巨量資料的連結及傳遞，因此高速的通訊將是未來高性能晶片驅動的重要方向之一。

而當大量的資料傳遞到雲端的高效能運算中心，高速運算晶片需要諸多的小晶片支援，如多核心 GPU 及寬頻的記憶體，運算晶片的資料吞吐量是極為巨大的，晶片間的有線連結也將由目前幾十 Gbps 進展超過 1Tbps 的傳輸量，要達到快速通訊及高效能的運算，有效率的晶片連結技術也勢必發展在追求極致的通訊，能量效率也必需同時達成，因此利用不同的半導體製程(如與矽光子的異質整合)結合創新高頻、高速電路，來完成複合式的通訊系統，也將是研發重點。

目前環境永續淨零碳排為全球相當矚目的議題，而高效率高功率及小型化的功率轉換系統在其中扮演一個非常重要的角色。面對一些新興應用，例如超大規模(hyperscale)資料中心等，全球持續投入高功率密度及高操作頻率的小型化功率轉換系統的研發。本計畫將聚焦於利用新興化合物半導體電路與模組及異質整合技術，開發下一世代的高效率及微小化的功率轉換系統。

### 貳、 研究議題範疇

探討高能效、高傳輸電路與晶片模組的研發，建議的研究主軸包含但不限於下列領域：

1. 高效能功率電路與模組：高能效功率電路與模組，應用於 AI edge；低壓、高操作頻率及高功率密度電路與模組，應用於資料中心與 AI 伺服器的電源供應；以及前述電路與系統之先進散熱技術與封裝。
2. 矽光子技術：超高速超低功耗光連結技術、異質光電整合技術、超高頻寬驅動及接收電路、光電晶片先進封裝與測試等，以應用於 AI 晶片連結、AI 高速運算、資料中心。

### 參、 計畫重點

推動高功率、高傳輸電路與模組應用之落實，以高能效功率電路與模組，

應用於 AI edge 為主。而低壓、高操作頻率及高功率密度的 AI 伺服器供電應用，應達成輸出功率、功率密度、效率分別大於 50 W、4000W/in<sup>3</sup> 與 95%為開發之技術指標。

此外，矽光子應用的開發重點，則以異質積體化先進封裝技術，以達到矽光子光收發機單通道 200Gb/s (功耗<5pJ/bit；低延遲<100ns)，且具有 WDM 技術，可在兩個波長以上進行傳輸，以突破現有門檻之應用目標。同時，規劃設計時亦需將強化 CPO 與 chiplet 系統層級的規劃、整合、驗證與模擬等完整開發能力納入整體考量。

### 分項三：邊緣運算晶片應用

#### 壹、 計畫背景及目的

現階段所有的資料驅動應用，除了強化高算力與高傳輸的晶片設計外，如何有效降低(一)運算的功耗，達成 fJ/b(每位元運算低於  $10^{-15}$ Joule)等級的節能運算，透過創新的演算模型、新型態運算架構、與電路設計的探討，提供下世代節能運算的解決方案，已成為生成式 AI 邁向應用普及化的重大挑戰議題。(二)資料感測的時間與所需求的功耗(pJ/bit)，提供高解析、高可靠的原始感測資料，滿足多元應用的需求，亦成為感測晶片能否帶動創新應用服務的關鍵議題。以具有深度資訊的影像感測晶片為例，除了是無人機、機器人、太空衛星、ADAS、自駕車等應用中不可或缺的關鍵技術外，在許多創新產業及應用中，也需要使用此類型的感測晶片與裝置。因此，在本項目的推動中，主要鼓勵研究團隊採用矽基製程，探討創新應用與服務所需求的關鍵技術與晶片，透過異質整合堆疊與小晶片封裝技術，針對各類場域所需求的超低能耗運算、高時間/空間解析度、低延遲感測時間、高節能效率的運算與感測晶片，結合軟硬體設計技術，達成關鍵晶片的技術指標與模組的離型系統，也具體呈現創新設計方案的價值與潛在落地應用的效益及影響力。

#### 貳、 研究議題範疇

本分項建議的研究主軸包含但不限於下列領域：

1. 探討超低功耗的演算法、系統架構、以及電路實現方案，達成能耗低於 fJ/b 的推論功能。
2. 高解析、低延遲感測晶片：以具有 SPAD 元件或類似的高壓製程，完成高速與低延遲的深度影像感測晶片與模組應用，達成之技術指標規格如下：(a) 影像數據量(含深度)： $> 1\text{Tb/s}$ 、(b)深度資訊： $< 1\text{cm}$ 、(c)能源效率： $< 1\text{pJ/b}$ 。
3. 以異質整合方式結合特殊應用感測晶片與高算力晶片、系統軟體，引領產業應用與服務創新者。
4. 此外因應晶片設計符合未來生活各面向之需求，以及政府施政的科技落地應用之策略，亦可提出 AI 晶片邊緣應用設計，甚至連結 IoT 聯網的完整

解決方案，可包括 edge inferencing、edge learning、驅動軟體、低耗能作業系統、前瞻通訊晶片與系統整合等技術研發。

5. 針對機器人、無人載具、衛星通訊等多元應用場域所需之專用晶片，可提出晶片設計與系統整合方案，並規劃對應之驗證方法，以確保晶片能於實際情境中落地應用，滿足百工百業之需求。

#### 參、 計畫重點

本分項將針對下世代運算的超低功耗運算與極低延遲、低能耗的智慧感測解決方案，研發關鍵晶片與軟硬體系統異質整合技術，並鼓勵學研團隊提出落地應用案例與雛形展示系統，透過具有國際競爭力的技術指標，包括有效降低運算的功耗、提升感測敏感度、降低感測延遲時間、提升感測能源效率等，展現此分項關鍵晶片的特色與價值創造。

## 分項四：先進晶片製造技術

### 壹、計畫背景及目的

傳統摩爾定律每兩年晶片電晶體數目倍增的目標，正面臨嚴峻挑戰，這主要受限於微影設備的波長與光罩大小的製程極限，以及元件二維尺度微縮的物理極限。隨著人工智慧運算需求急劇增長，國際 AI 晶片領頭公司提出了「超摩爾時代」的概念，強調晶片的運算能力必須每年倍增，以滿足未來計算需求。這項提升不僅依賴晶片設計與架構的改進，更有相當比例需仰賴半導體製造技術的不斷進步。台灣作為世界先進晶片製造技術的領導者，要維持領先地位，必須持續投入創新技術研發。其中，工業界負責五年內所需之量產技術開發，而學術界則聚焦於五到十年後的新興技術研究與評估，透過產學分工建構完整的技術研發藍圖。

目前，先進晶片製造技術已從平面轉變為三維，以有效克服二維微縮的極限。三維元件技術已成為主流趨勢，電晶體結構從平面轉變為三維的 FinFET，進而發展至兩奈米下的 NSFET 與 CFET 結構，並搭配晶片背面供電技術，讓整個電晶體結構向第三維度延伸。在電晶體後段製程中，開始整合更多不同型態的元件，例如記憶體，以提供更高的記憶體頻寬與儲存密度；此外，還包括功率元件、感測器與微機電元件等，提升晶片的整合密度並降低功耗。三維封裝技術則是「超摩爾時代」的關鍵，透過多顆晶片的垂直堆疊，提升晶片密度、性能與功能。三維元件與三維封裝的結合，我們稱之為  $3D \times 3D$  技術，這兩者的整合將成為「超摩爾時代」的核心技術。

本計畫旨在徵求學術界提案，推動  $3D \times 3D$  技術的研發，聚焦於未來五到十年間的關鍵創新，以確保台灣半導體產業持續領先。透過設定極具挑戰性的目標，引導研究團隊提出破壞性解決方案，支持邏輯、記憶體與異質整合等技術平台，達成每兩年等效性能 PPA(性能(performance)、功耗(power)與面積(area))倍增的願景。計畫目的不僅在於技術突破，更在於建構產學合作生態，加速新興技術從概念到應用的轉化，強化台灣在全球半導體供應鏈中的核心地位。

### 貳、研究議題範疇

本分項建議的研究主軸包含但不限於下列領域：

1. 關鍵材料開發：探討新型材料在  $3D \times 3D$  技術中的應用，包括新通道材料、高介電常數材料、低阻抗互連材料，以及用於三維堆疊的接合、隔離、與散熱材料，以提升元件性能並降低功耗。
2. 元件結構創新：研究先進電晶體結構，如 NSFET、CFET 與晶片背面供電技術，嵌入式記憶體（如 SRAM 替代方案）、功率元件、感測器及微機電元件，實現更高密度與多功能整合。
3. 製程技術優化：聚焦三維微縮與三維堆疊製程之模組與整合技術，克服二維物理縮極限。
4. 封裝技術推進：開發三維封裝方法，如晶片垂直堆疊、異質整合封裝，涵蓋熱管理、信號完整性與功率傳輸之優化，支援多晶片模組化。
5. 檢測分析方法：建立先進檢測工具與標準，包括奈米級成像與成分分析、電性測試及可靠性評估，用於驗證  $3D \times 3D$  技術的 PPA 指標。

以上技術開發均應建立相對應之元件或應用載具之 PPA 指標，與國際類似技術進行比較，並以每兩年倍增做為整體計畫研發主軸，提出可行路徑。

## 參、計畫重點

在本計畫中， $3D \times 3D$  技術的研發重點不再僅限於電晶體密度追求，而是轉向以 PPA 作為技術提升指標，這是包含台積電在內的主要半導體公司所採用的標準。團隊需詳細說明各自研發技術之國際最新 PPA 指標比較，以及未來五到十年間達到每兩年倍增的可行技術路徑。

PPA 指標需同時考慮性能、能耗與面積，建議指標例如 TOPS/W-mm<sup>2</sup>，同時考量處理速度、功耗倒數、元件面積倒數之乘積，乘積越大代表所發展技術之等效性能越好。

以嵌入式記憶體為例，現有先進製程中的解決方案為 SRAM，台積電兩奈米的 SRAM 單位位元面積為  $0.021 \mu\text{m}^2$ ，能耗為 sub pJ/bit，操作時間達 sub ns。所提新興嵌入式記憶體技術須以兩奈米 SRAM 為基準，在能耗與速度相當的條件下，達到單位位元面積每兩年縮小為一半的目標。