



# AI晶片發展與半導體市場趨勢分析

劉智文 產業分析師

產業情報研究所(MIC)

財團法人資訊工業策進會

2020.12.28

[wenliu@iii.org.tw](mailto:wenliu@iii.org.tw)  
[mic.iii.org.tw](http://mic.iii.org.tw)

**MIC**<sup>®</sup>



# 簡報大綱

---

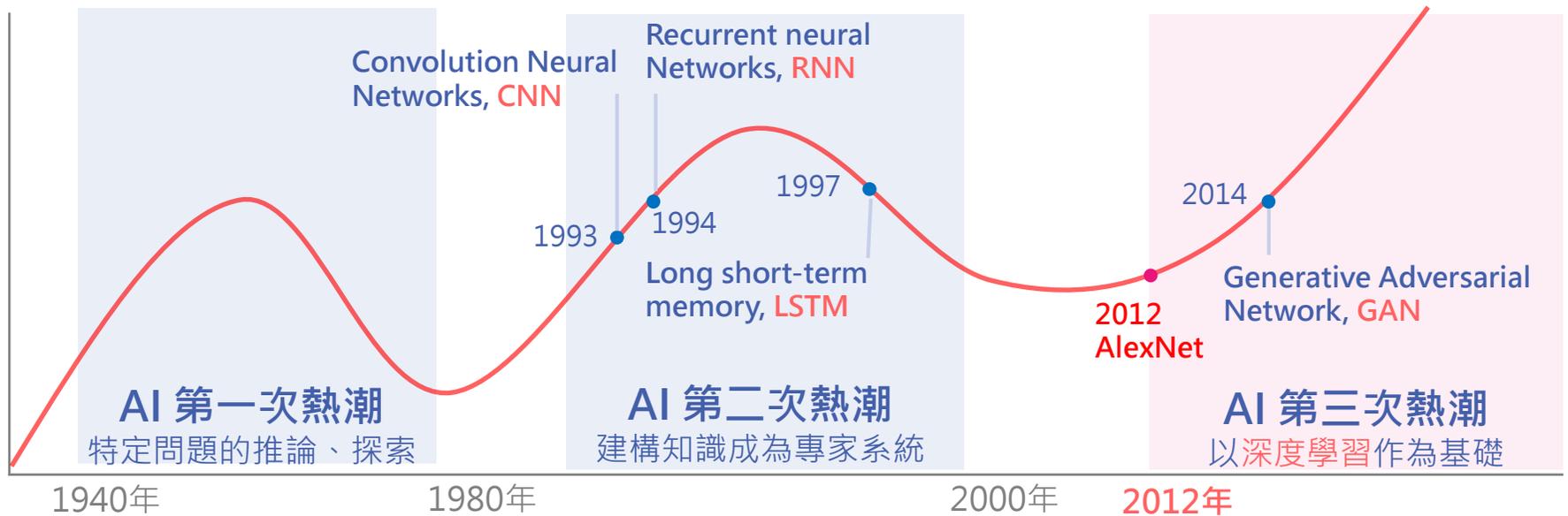
- ❖ AI晶片發展現況
- ❖ AI晶片市場現況
- ❖ AI晶片對半導體產業的影響
- ❖ 結論



# AI晶片發展現況



# Alexnet打破影像辨識瓶頸 2012開啟AI元年



資料來源：MIC，2020年12月

## AI影像視覺技術 發展重要里程碑



**LeNet (1993-1998)**  
手寫郵遞區號辨識，  
第一個能夠以良好的  
速度和準確度識別手  
寫數字卷積網路



**Nvidia(2007)**  
進入AI領域，提供  
硬體加速

## AlexNet

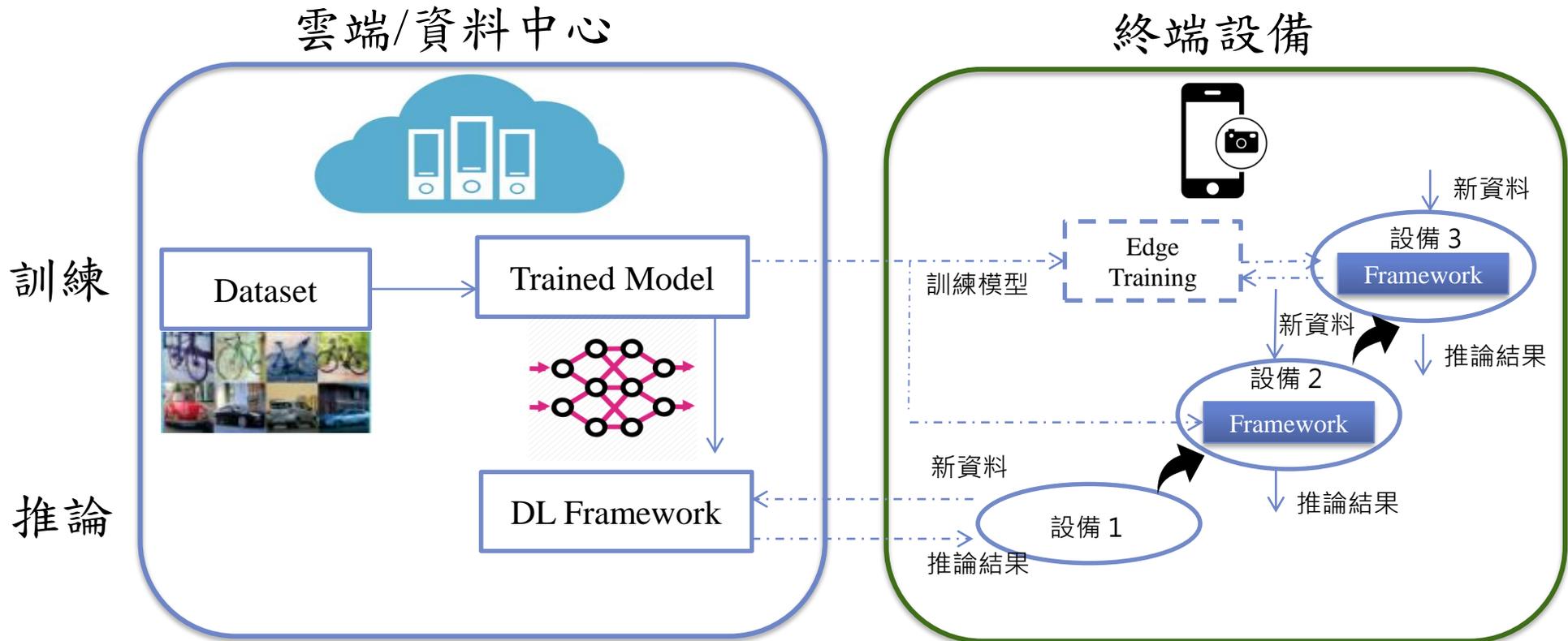
**AlexNet(2012)**  
以顯著的優勢贏得  
ILSVRC 2012影像識  
別比賽，錯誤率降低  
至16.4% (8層)

## Microsoft ResNet

**ResNet(2015)**  
微軟亞洲研究院提出  
ResNet 取得ILSVRC  
2015第一，錯誤率  
3.57%打敗人類  
(152層)



# 應用情境定義AI硬體系統架構



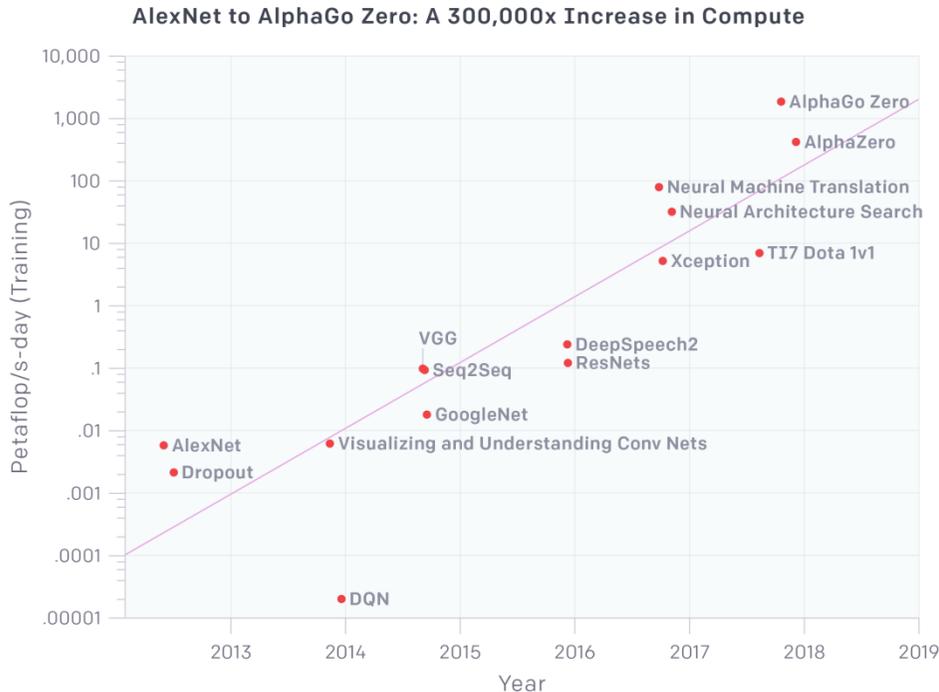
資料來源：MIC，2020年12月

- ❖ 主流為在雲端進行資料訓練，推論部分則分別在雲端或終端進行
- ❖ 終端學習因個人化需求、隱私權等因素興起需求，但目前仍缺少解決方案

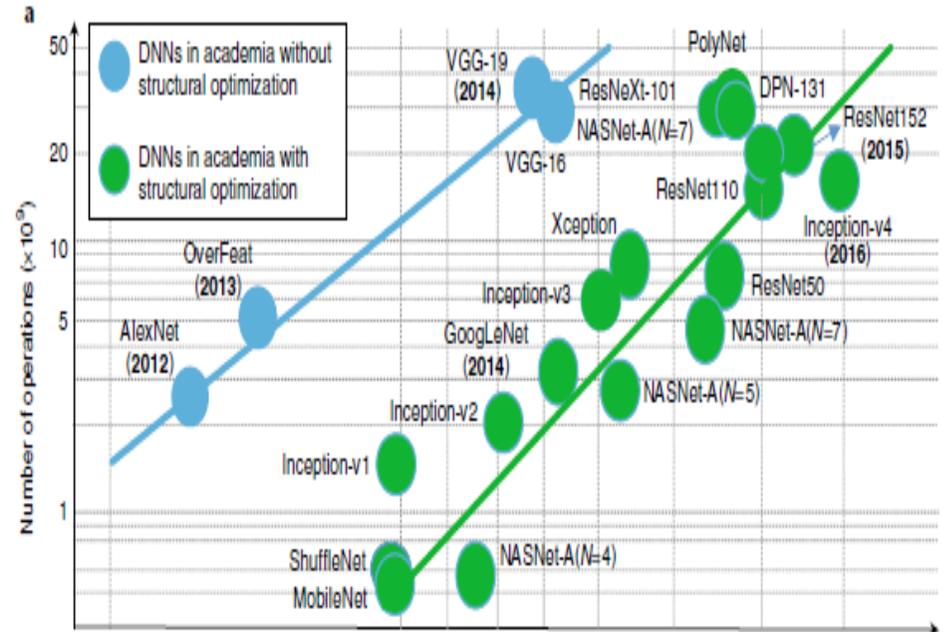


# AI運力需求指數增加 硬體資源吃緊

## AI訓練運算需求成長



## AI推論運算需求成長



資料來源：OpenAI，Nature Electronics(2018)，2020年12月

- ❖ AI模型不斷變大，造成運算需求、記憶體存取次數越不斷膨脹
- ❖ AI模型採用演進，CNN->RNN->RL->GANs，也使運算量指數成長(訓練及推論運算)
- ❖ 實際應用情境將可能為多模型組合



# 多元創新增加AI系統運算能力

## AI晶片三大創新方向

### 軟體設計

目標：演算法/模型精準度  
與模型大小之平衡

作法：

- Pruning
- Quantization
- BNNs
- Dynamic-precision
- Compression
- Compiler design
- ..

### 系統設計

目標：減少運算量及記憶  
體存取或增加頻寬

作法：

- Systolic array
- HBM(High Bandwidth Memory)
- 3D封裝(記憶體、感測器..)
- SRAM
- ..

### 新興架構

目標：半導體製程改善  
運算能耗效率

作法：

- In memory computing
- Analog computing
- MRAM/RRAM/PCRAM
- Neuromorphic
- Spiking neural network
- ..



# 大廠搶占雲端 AI晶片新創輩出

雲端(訓練/推論)

終端(推論)

CPU



GPU



FPGA



ASIC (XPU)



TESLA



groq

KNUEdge

Cambricon



Kneron



GRAPHCORE



brainchip



BITMAIN



新創  
企業

資料來源：MIC，2020年12月



# Intel強化架構整合提升競爭力



- ❖ AI運算除了強調高效能外，低延遲也成為今年各廠商主打的重點，尤其在物聯網的相關應用上，低延遲更成為不可或缺的功能，預期在5G滲透率提高的帶動下，具備通訊功能的AI裝置將會持續增加
- ❖ 多元處理架構具備高效能、低功耗等優點，滿足人工智慧運算的需求，成為各大廠的標準配置，也使得廠商在開發平台的設計上，能提供更多裝置種類的支援
- ❖ 多元處理架構使得晶片市場競爭加劇，吸引新創公司投入，也帶動併購的風潮



# 多元創新增加AI系統運算能力

## AI晶片三大創新方向

### 軟體設計

目標：以精準度換取模型大小

作法：

- Pruning
- Quantization
- BNNs
- Dynamic-precision
- Compression
- Compiler design
- ..

### 系統設計

目標：減少運算量及記憶體存取或增加頻寬

作法：

- Systolic array
- HBM(High Bandwidth Memory)
- 3D封裝(記憶體、感測器..)
- SRAM
- ..

### 新興架構

目標：半導體製程改善  
運算能耗效率

作法：

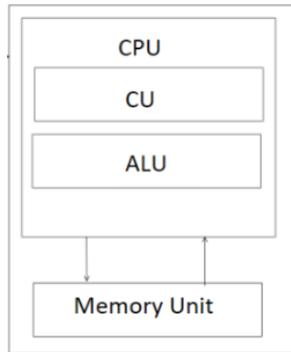
- In memory computing
- Analog computing
- MRAM/RRAM/PCRAM
- Neuromorphic
- Spiking neural network
- ..



# 新興架構未來將可降低能耗兩個數量級以上

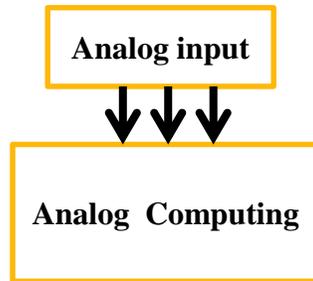
100pJ/MAC

馮紐曼架構



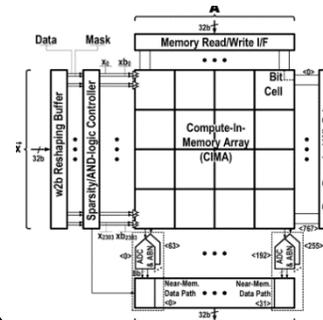
1pJ/MAC

類比運算架構



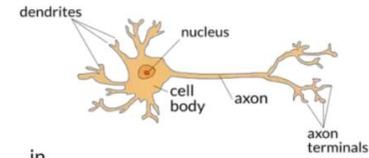
0.1 pJ/MAC

記憶體內運算  
架構



<0.1 pJ/MAC

神經型態運算



Today

Future

以類比電路解決數位乘加運算電晶體數，降低能耗及晶片面積

以記憶體內加入運算方式降低資料搬移，降低能耗及增加算力

模仿人類神經傳導，以脈衝方式降低運算能耗

資料來源：經濟部AI領航推動計畫，MIC，2020年12月



# AI晶片市場現況



# 2019年半導體市場規模小幅衰退

(單位：十億美元)

### 2011年~2021年全球半導體市場規模



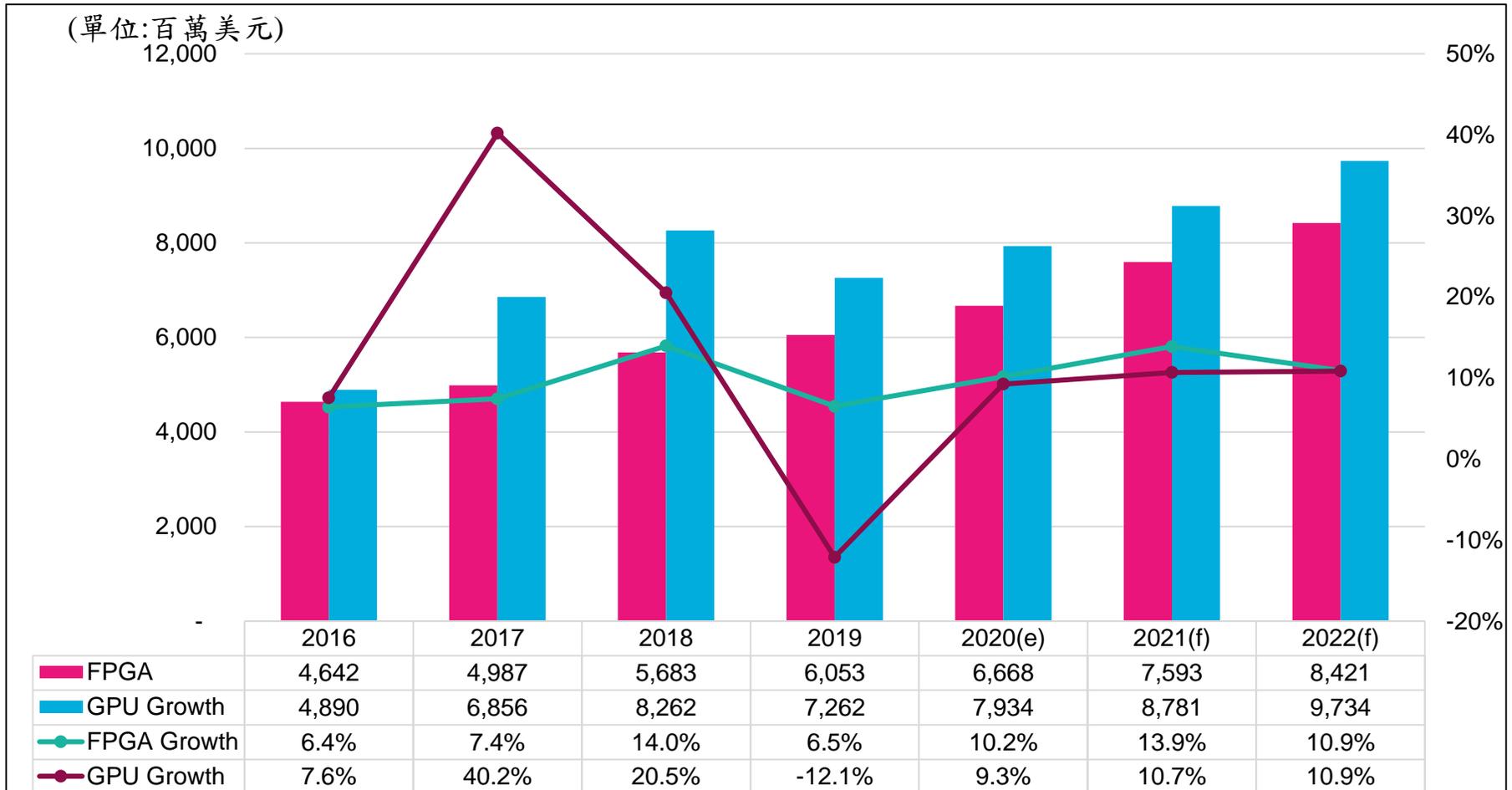
資料來源：WSTS，MIC預測，2020年12月

- ❖ 2019年記憶體價格下跌，導致全球半導體市場衰退(若不計入記憶體因素，則呈現微幅成長)，2020年則受新冠肺炎影響，整體成長將不如預期，相關需求動能將向後遞延1~2季





# 全球GPU及FPGA市場現況



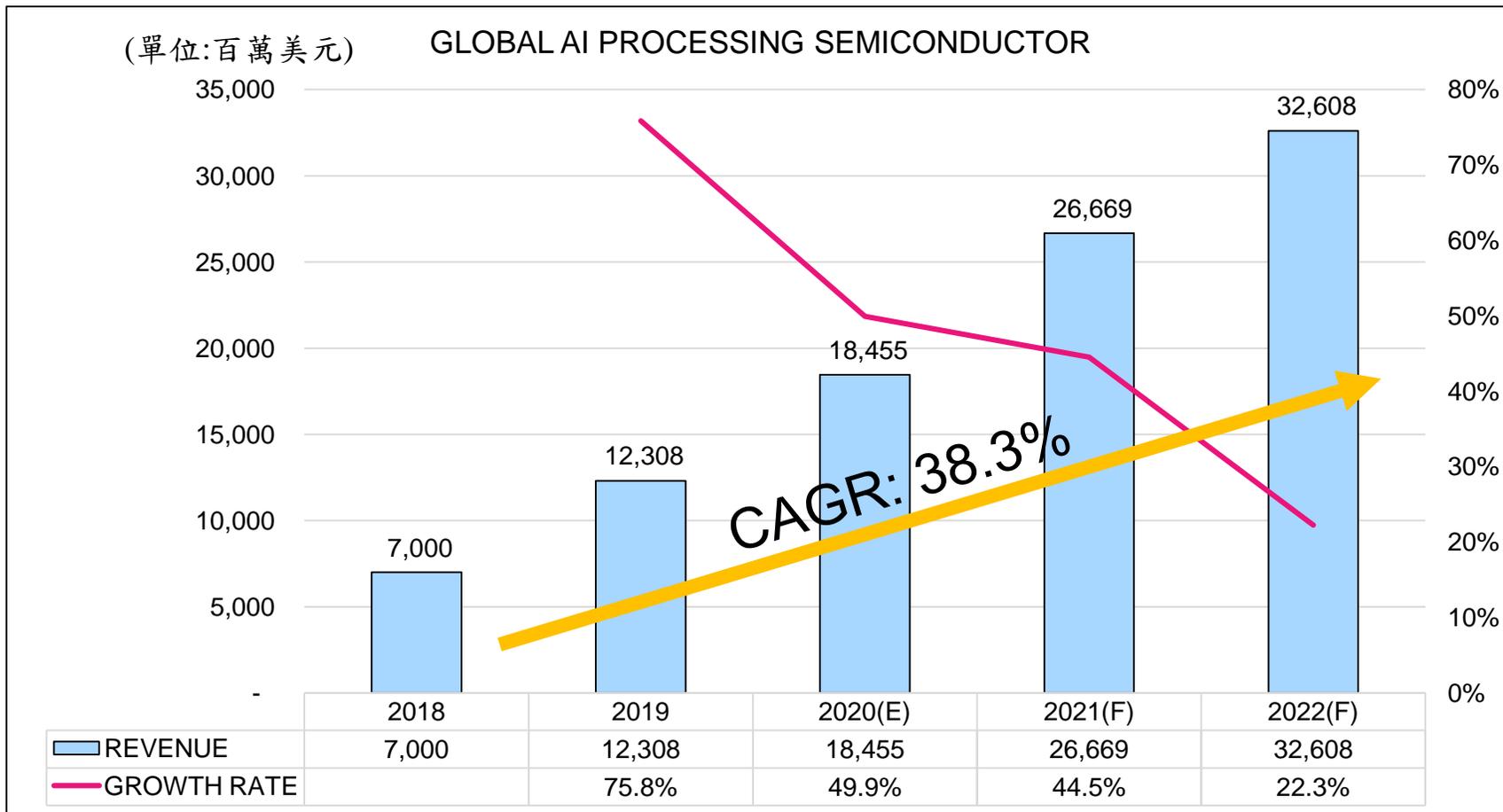
資料來源：Gartner，MIC整理，2020年12月

❖ 2019全球GPU及FPGA市場產值分別為72.6億美金以及60億美金123億美金，預計至2022年每年將保持38.3%的高速成長





# 全球AI 處理器市場現況

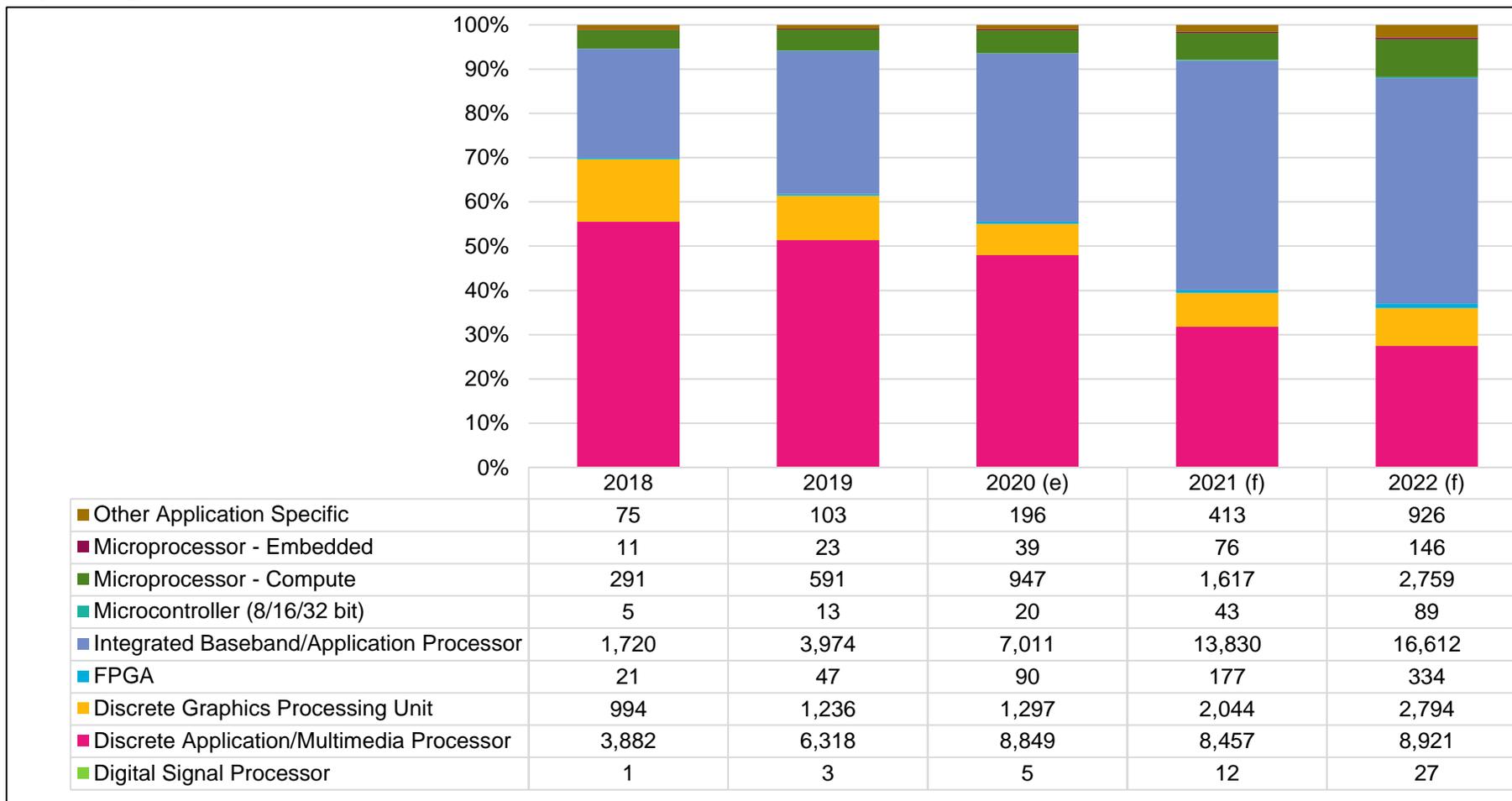


資料來源：Gartner，MIC整理，2020年12月

- ❖ 2019全球AI晶片市場產值約為123億美金，預計至2022年每年將保持38.3%的高速成長



# 全球AI晶片市場晶片類型分布

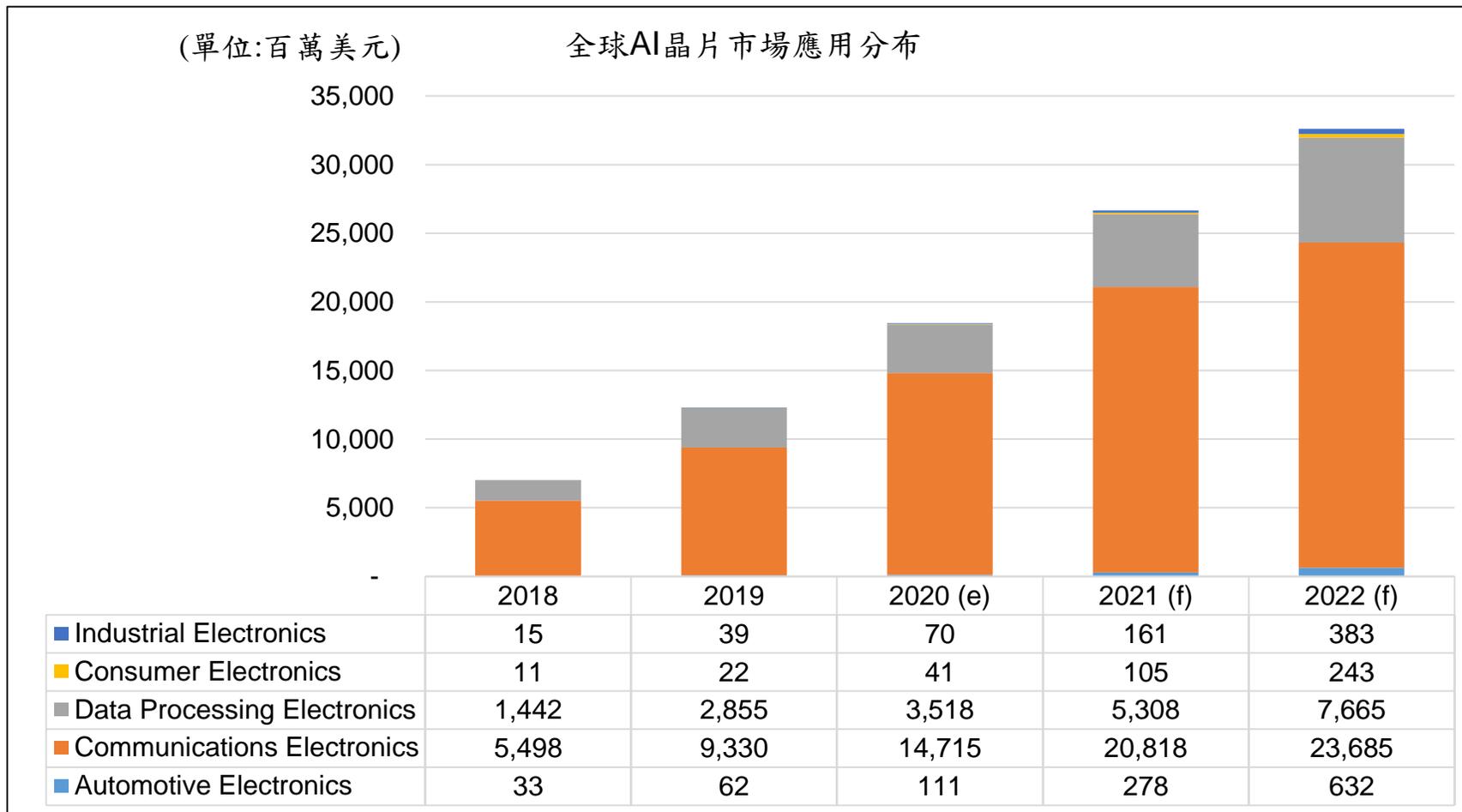


資料來源：Gartner，MIC整理，2020年12月

❖ 隨著網通技術的進步，未來具備聯網功能的AI晶片將會隨著邊緣運算進入各項應用角落，大幅提升AI晶片的滲透率



# 全球AI晶片市場應用分布



資料來源：Gartner，MIC整理，2020年12月

- ❖ 全球AI晶片市場滲透率仍以電腦及通訊應用為大宗，汽車、消費性電子、以及工業應用較少，但在開發環境逐漸成熟下，未來幾年將保持高度成長



# AI晶片對半導體產業的影響



# 3C結合AI技術創造新價值與新應用場域



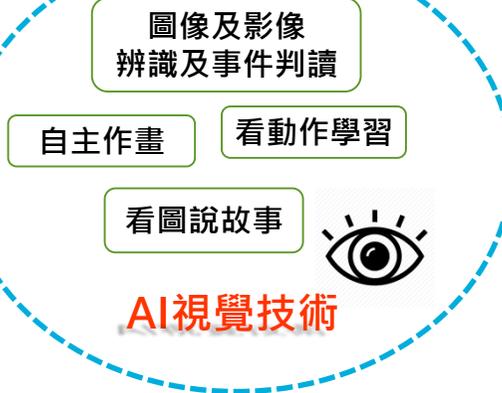
Tencent 腾讯



Manufacture

Car

Medical



Computing

Communications

Consumer

資料來源：MIC，2020年12月

- ❖ 資通訊硬體終端結合AI視覺、語音等技術與內容服務業者，創造新應用市場與新服務，矩陣式思考創造新商機



# 新興應用領域帶動不同半導體需求

著重因素：

體積

延遲

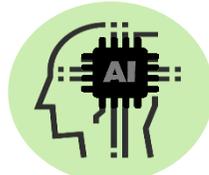
成本

效能

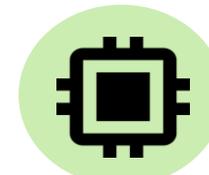
功耗



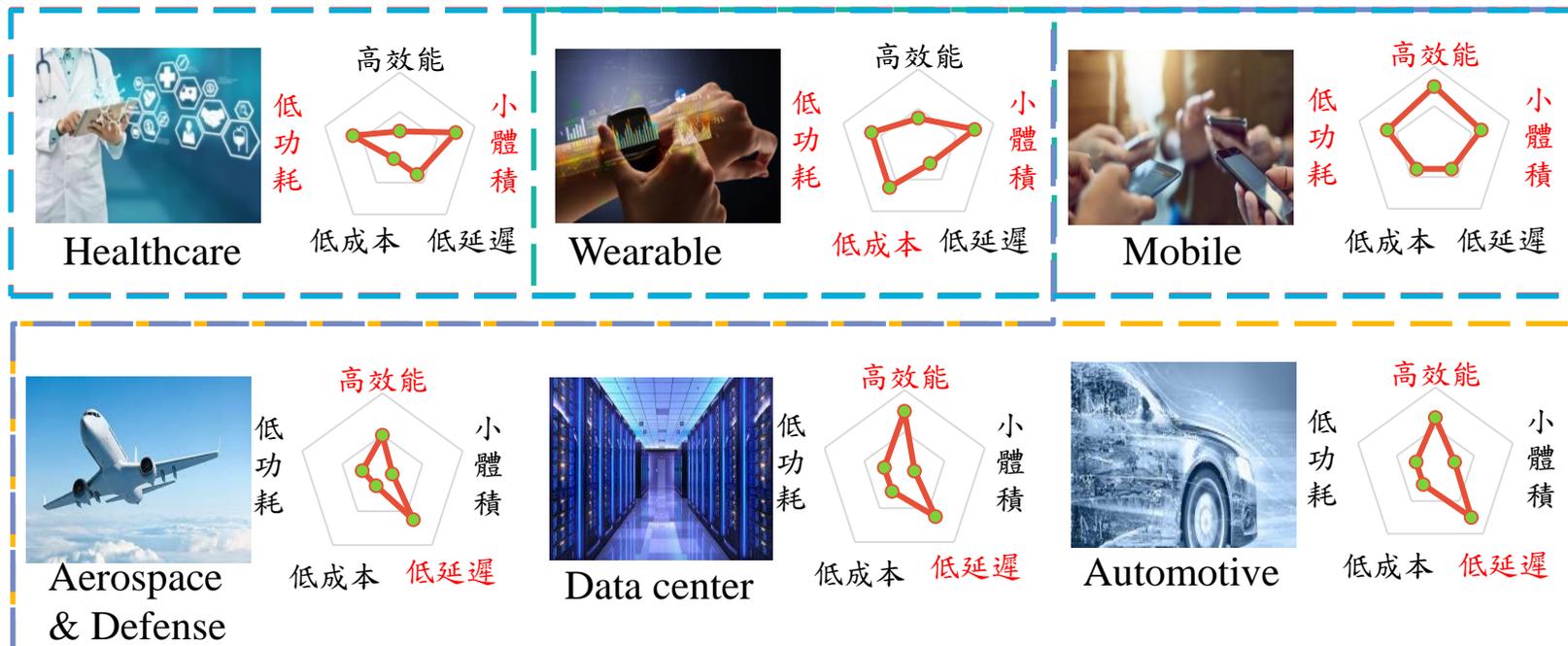
HI



AI



IC



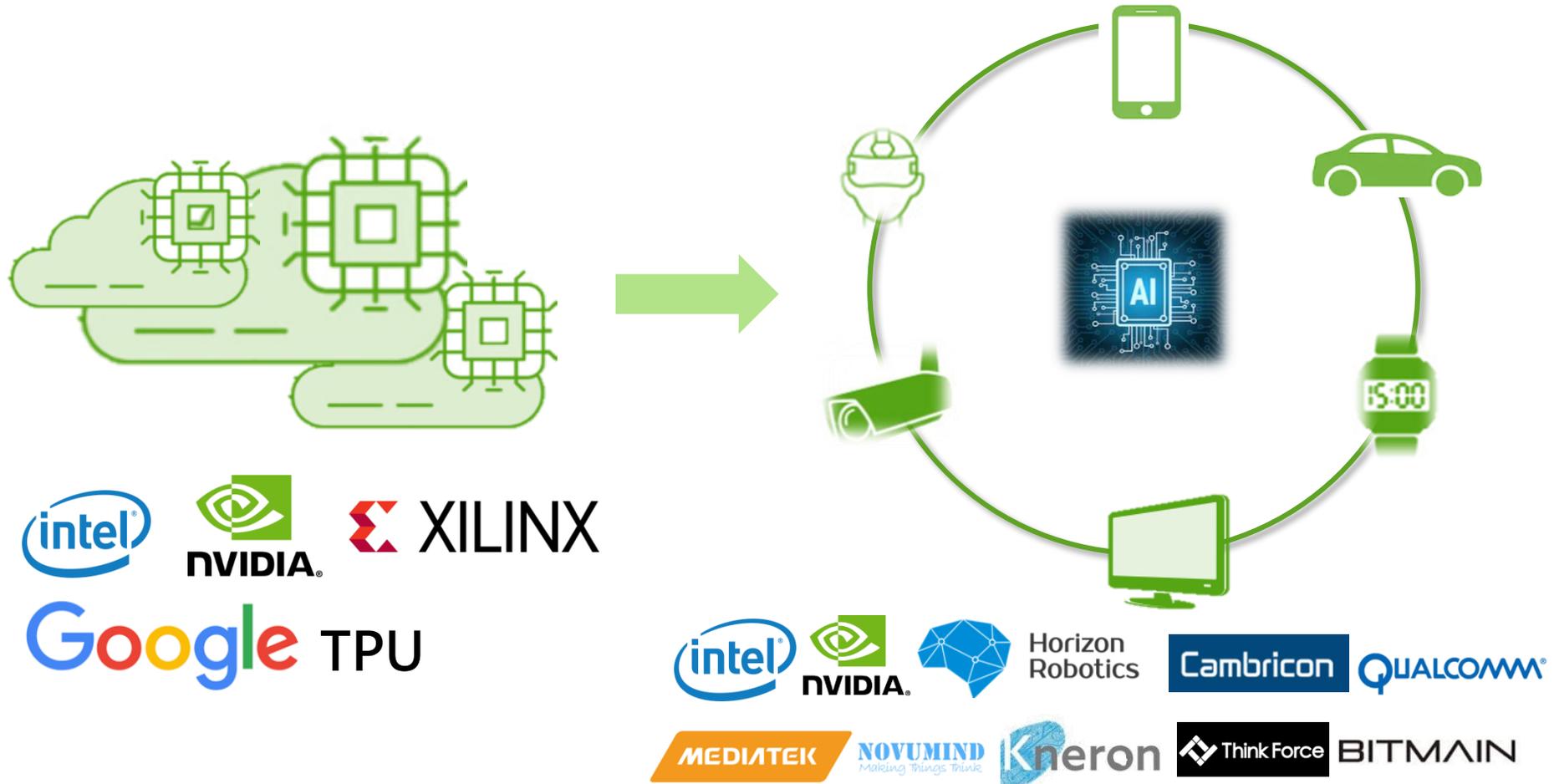
備註：紅字為前兩大重點需求因素

資料來源：MIC，2020年12月

❖ 人類智慧、人工智慧及半導體晶片的加成，使得越來越多的應用領域相應而生，多元新興應用市場對於成本、效能、體積等晶片要求各有差異



# 一、AI晶片在Edge端走向分化

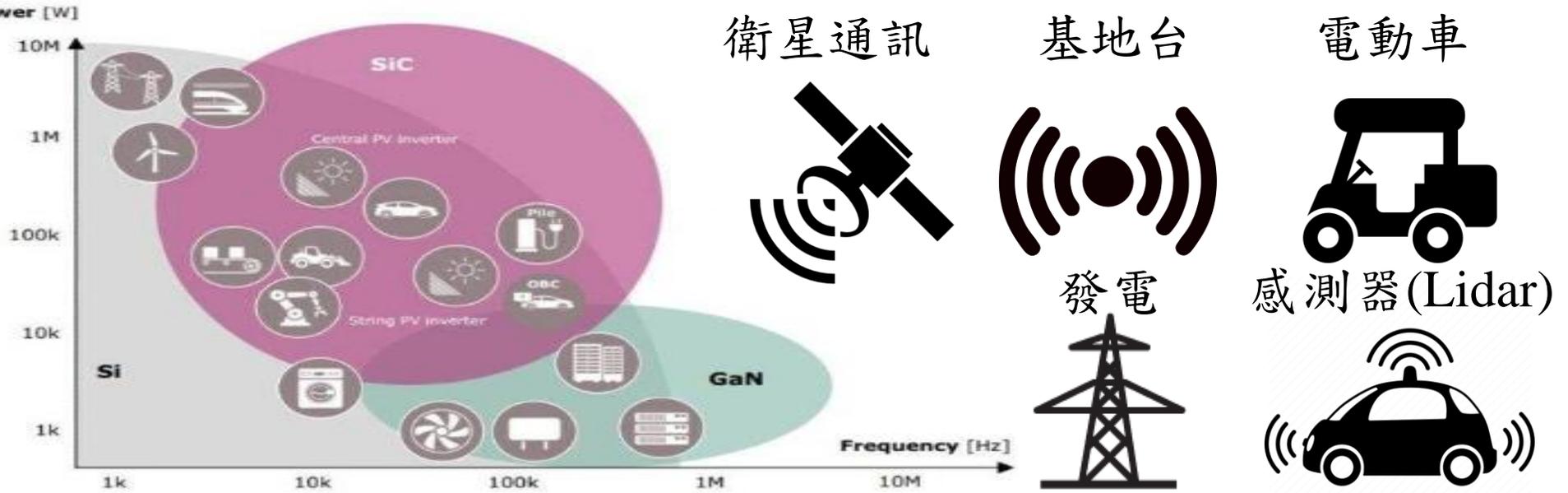


資料來源：MIC，2020年12月

- ❖ 由於Edge及Device裝置對深度學習網路的運算需求增加，也因應用情境不同，因此促成許多大廠及新創投入開發

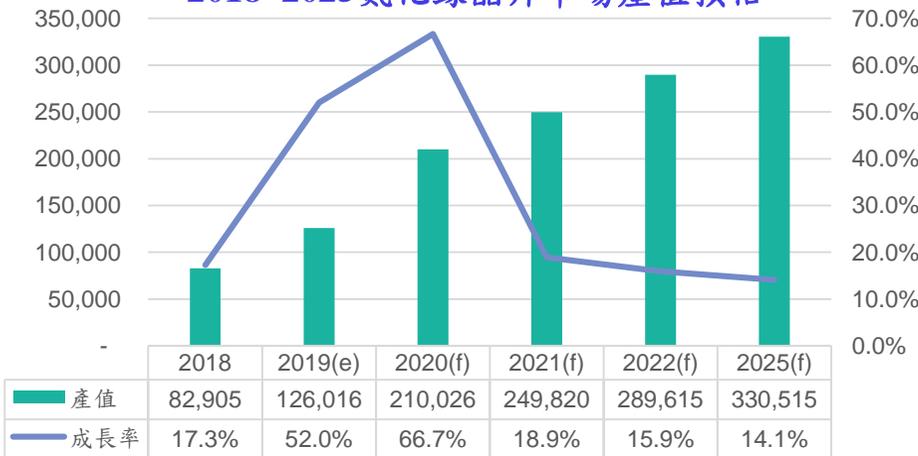


## 二、應用需求強化新興半導體需求成長



單位:百萬美元

2018~2025氮化鎵晶片市場產值預估



❖ 化合物半導體具備高頻、高功率、能矽寬等優勢，能有效提升能源轉換效率及散熱效果。

❖ 台廠如台積電、世界先進、穩懋、漢磊、茂矽、聯穎等，皆積極投入開發



• 與STMicroelectronics合作開發氮化鎵技術

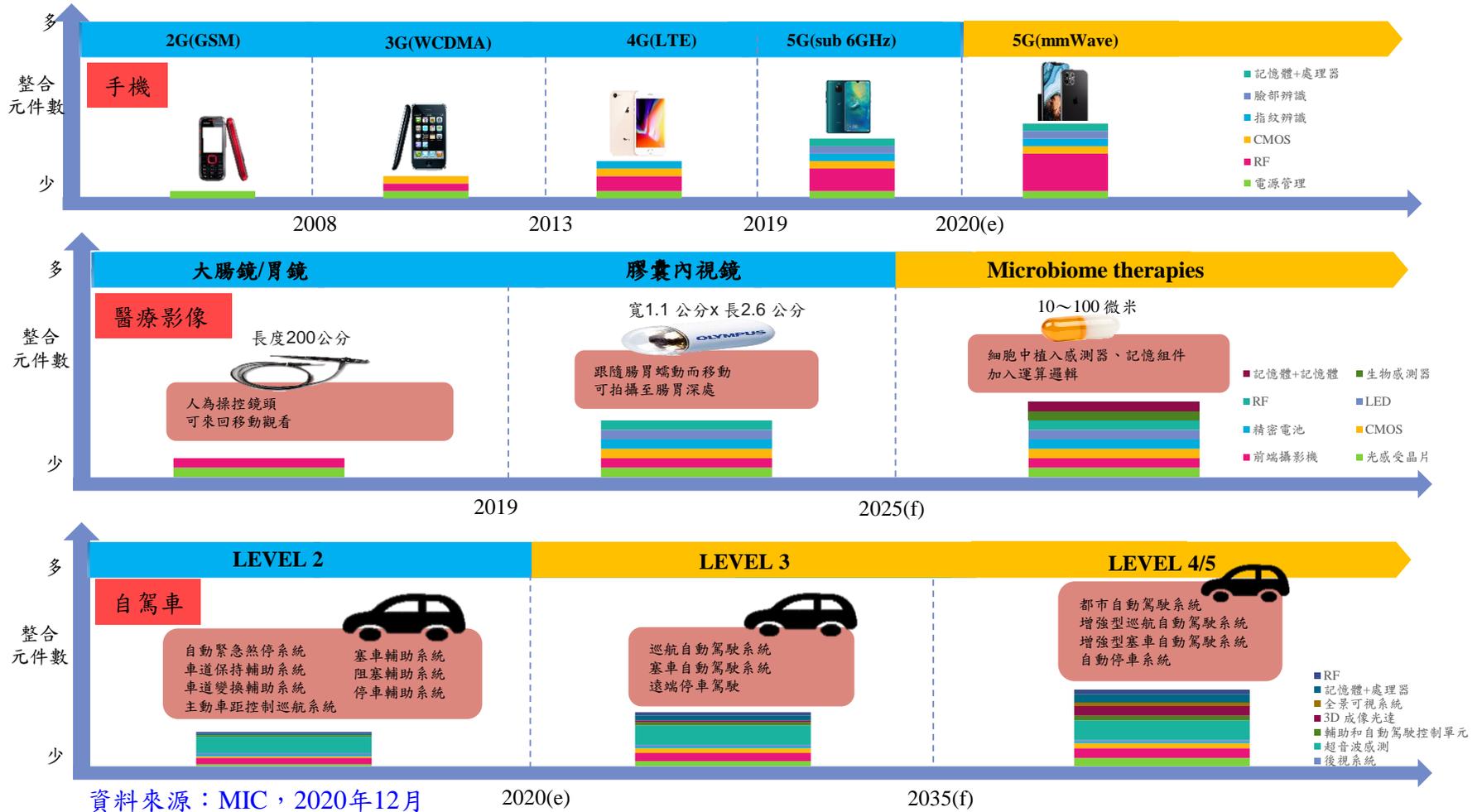


• 和Qromis共同合作開發8吋GaN-on-QST

資料來源：富士總研，MIC整理，2020年12月



# 三、市場需求走向多晶片、客製化發展

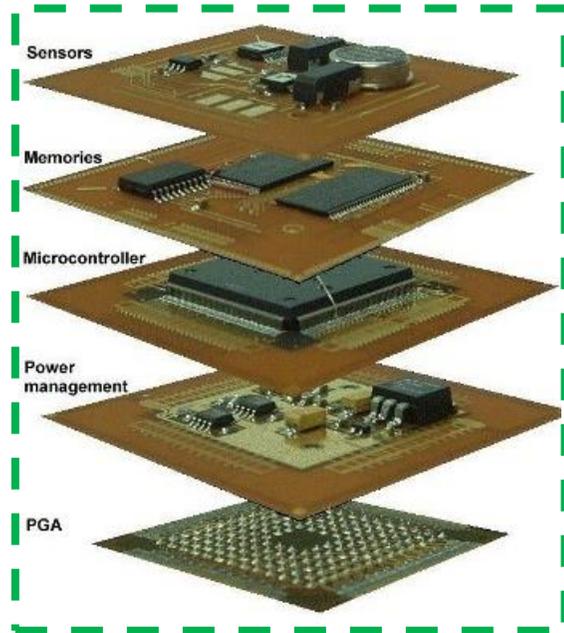


❖ 多元新興應用市場不僅將使未來晶片的發展朝向更多元件的整合，且隨著產業需求的不同，未來晶片將朝向客製化的走向

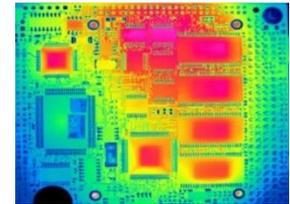
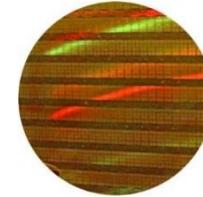
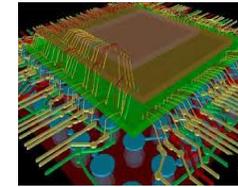
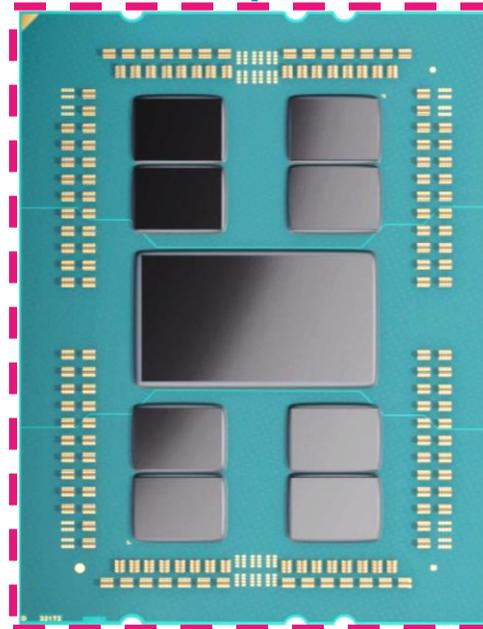


# 異質整合及Chiplet成效能發展關鍵及方向

## 異質整合



## Chiplet



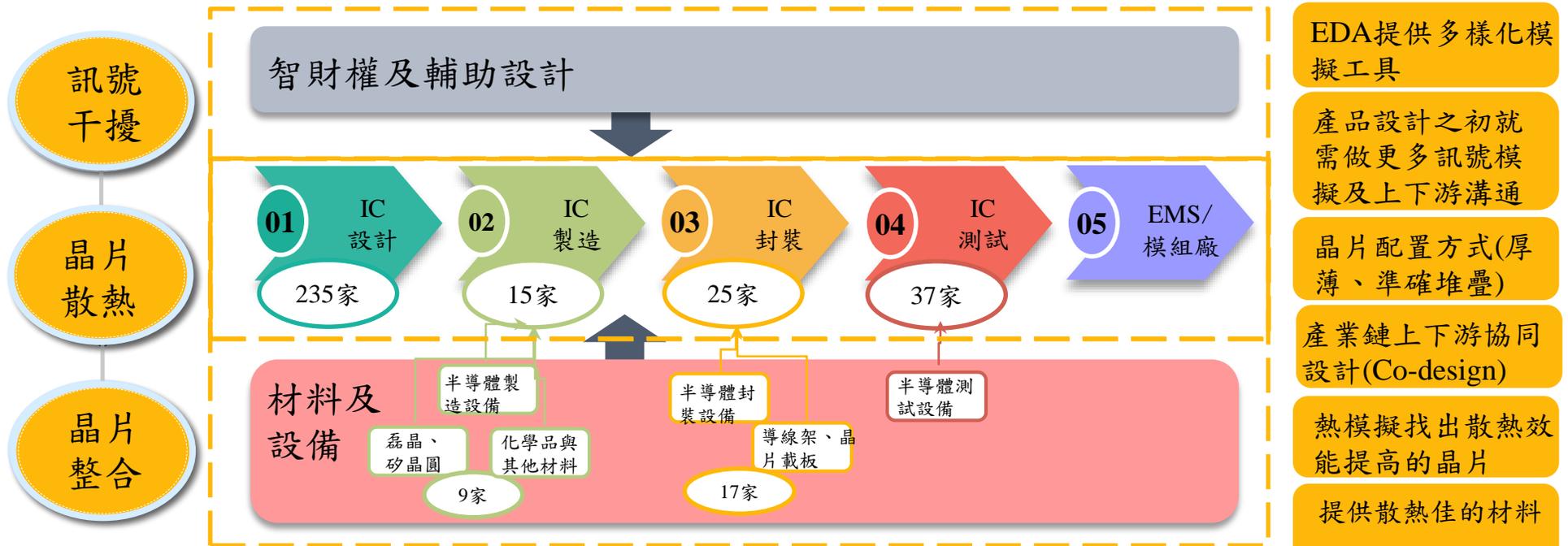
驗證設計流程優化  
(EDA)  
產品可靠度驗證

資料來源：各公司，MIC整理，2020年12月

- ❖ SiP 模組不著重於單一晶片效能，而是整體零組件搭配後的發揮。在初始設計廠商便需考量各類零組件在功耗、散熱的差異以及晶片相互的干擾及影響，並將體積、晶片佈線等因素加以規劃，以發揮模組化最大效能
- ❖ 目前上下游廠商皆依靠EDA軟體工具加快設計流程並模擬及驗證產品效能，採用標準化設計流程，幫助廠商提升良率、降低流片失敗的可能性
- ❖ 由於SiP使用零組件眾多，其產品的設計與驗證流程將變得更為重要，如何提升產品可靠度並縮小最終產品效能與設計之初的差異，將是未來發展最重要的課題



# SiP供應鍊複雜度高 需半導體產業鏈更緊密合作



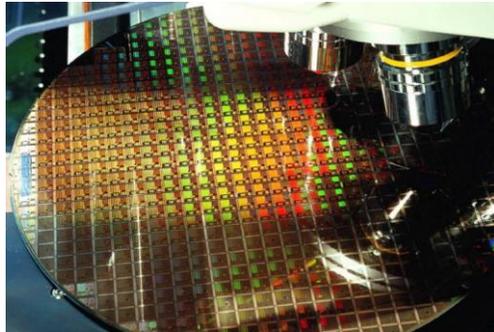
備註：○ 為台灣廠商家數  
資料來源：MIC，2020年12月

- ❖ 多晶片整合的SiP模組為發揮最大效能，在晶片設計、製造過程，必須考量訊號干擾、晶片散熱及晶片整合度等問題
- ❖ 除了上下游廠商對於晶片配置、熱模擬等需進行協同合作外，EDA廠商也須提供多樣化模擬工具以利進行多類型訊號模擬，材料及設備廠也應提供相對應的先進材料，幫助幫助廠商提升良率、降低流片失敗的可能性
- ❖ 隨著SiP模組日趨複雜化，如何提升產品可靠度並縮小最終產品效能與設計之初的差異，將是未來發展最重要的課題。



# 封裝業者與晶圓製造業者合作愈趨緊密

## IC製造



## IC封裝



強化合作  
關係

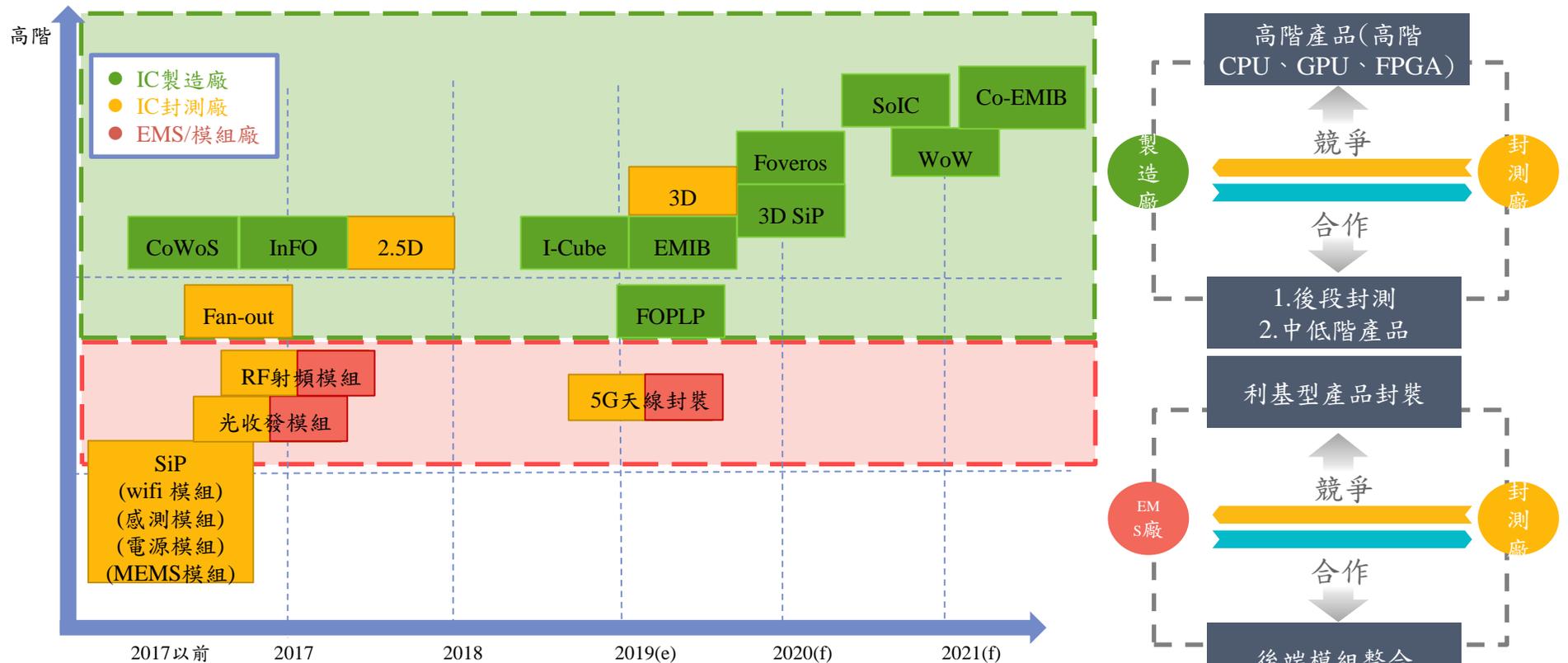
資料來源：MIC，2020年12月

晶圓廠	封測廠	封測技術
台積電	自行開發	CoWoS、InFo SoIC、WoW
Samsung	自行開發	PoP、I-Cube
Intel	自行開發	Foveros
中芯	江蘇長電	3D SiP
南亞科	福懋科	--
華邦電	華東	--
長江存儲	宏茂	Xtacking

- ❖ 摩爾定律趨緩，晶片效能成長仰賴封裝技術，吸引晶圓廠商開發封裝技術
- ❖ 由於晶圓廠資本額較大，且位居產業上端，因此目前相關合作案大都由晶圓製造業者出資入股封裝廠商
- ❖ 製造與封測結合可提高廠商對製造流程的掌握程度，並簡化產線驗證的流程，有助於廠商搶奪可靠度要求較高的市場(如車用、能源以及航太)



# 半導體產業競合關係改變



備註：僅列出異質整合相關封測技術  
資料來源：MIC，2020年12月

- ❖ IC製造廠及下游的EMS/模組廠相繼以各自優勢的方式跨入封測產業
- ❖ 半導體廠商在進行合作的同時，彼此間競合關係也將隨技術發展而有所改變
- ❖ IC封測廠與IC製造廠在高階技術重疊性高；與EMS廠則將在5G天線封裝、RF模組等利基型產品競爭



# 結論



# 結論

- ❖ 全球AI晶片市場成長動能從雲端滲透至終端，AIoT、車聯網、醫療用電子等三大應用將成為半導體產業未來五年主要的成長動能
- ❖ 多元處理架構具備高效能、低功耗等優點，滿足人工智慧運算的需求，CPU、GPU、FPGA等通用處理器結合ASIC晶片的架構，將成為AI晶片市場主流
- ❖ 多元運算架構以及多晶片的整合技術使得封測影響力提升，但也使得半導體產業的垂直競合產生變化，封測產業市場競爭加劇



# 智慧財產權暨引用聲明

- ❖ 本活動所提供之講義內容或其他文件資料，均受著作權法之保護，非經資策會或其他相關權利人之事前書面同意，任何人不得以任何形式為重製、轉載、傳輸或其他任何商業用途之行為
- ❖ 本講義內容所引用之各公司名稱、商標與產品示意照片之所有權皆屬各公司所有
- ❖ 本講義全部或部分內容為資策會產業情報研究所整理及分析所得，由於產業變動快速，資策會並不保證本活動所使用之研究方法及研究成果於未來或其他狀況下仍具備正確性與完整性，請台端於引用時，務必注意發布日期、立論之假設及當時情境