行業概覽

*本節及本文件其他章節所載資料及統計數據摘錄自弗若斯特沙利文編製的獨立行業報告，以及來自各種政府官方刊物及其他公開可得刊物。我們委聘弗若斯特沙利文編製有關***[編纂]***的獨立行業報告。我們認為本資料摘錄自適當的來源，且已合理審慎摘錄及轉載有關資料。然而，我們、聯席保薦人、***[編纂]***、彼等各自的任何董事及顧問或參與***[編纂]***的任何其他人士或各方並無獨立核實來自政府官方來源的資料，且概不就其準確性發表任何聲明。*

資料來源

就**[編纂]**而言，我們委聘獨立市場研究顧問弗若斯特沙利文對GPGPU市場進行分析，及編製弗若斯特沙利文報告。弗若斯特沙利文為於1961年在紐約成立的獨立全球諮詢公司，提供行業諮詢、市場策略諮詢及企業培訓等服務。就其提供的市場研究服務，我們已向弗若斯特沙利文支付費用人民幣850,000元，我們認為該費用符合現行市場價格。

編纂及編製弗若斯特沙利文報告時，弗若斯特沙利文進行：(i)一手研究，包括訪談行業參與者、競爭對手、下游客戶及公認的第三方行業協會；及(ii)二手研究，包括審閱企業年報、相關官方機構的數據庫，以及弗若斯特沙利文過去數十年建立的獨家數據庫。弗若斯特沙利文報告中的市場預測基於下列關鍵假設作出：(i)所論述的全球市場之社會、經濟及政治環境於預測期內將保持穩定；(ii)預測期內全球及中國GPGPU市場的政府政策將保持一致；及(iii)GPGPU市場將繼續受弗若斯特沙利文報告所述因素驅動。

全球GPU行業概覽

大模型與AI智能體的定義

大模型主要可分為通用大模型與行業大模型兩類。通用大模型通過大量跨領域數據集訓練，具備自然語言理解、邏輯推理等泛化能力。相反，行業大模型則依託特定領域數據與知識進行優化，能夠精準解決專業問題，但跨行業適應能力相對較弱。兩種大模型可相輔相成：通用大模型為各行業大模型的開發提供底層技術基礎，降低行業大模型的開發門檻並提升訓練效率，而行業模型則提供垂直知識，加強通用模型的泛化能力。

AI智能體是一種基於人工智能技術的智能系統，通常以大模型(尤其是行業大模型)作為核心基礎構建而成。在人類用戶設定目標並提供算力、數據等必要資源後，AI智能體能夠通過感知、動態決策與持續學習，自主規劃任務流程、調度資源，執行任務並適應複雜環境。

行業大模型的發展前景與優勢

行業大模型正成為GPGPU領域的藍海市場。鑑於通用大模型持續成熟，正在各行各業迅速普及，推動其核心基礎設施GPGPU的同步成長。

行業概覽

行業大模型建基於通用大模型，可微調以適配垂直領域需求。該方式具有高度專業性、適配性與成本效益，能降低算力／數據需求，縮短開發週期及解決傳統企業AI限制，成為推動AI技術在行業大規模落地的關鍵力量。

行業大模型為AI智能體的發展提供核心支撐(其可加速滲透)，其配備通用認知能力與領域專業知識。對內，以私有方式部署的智能體可通過實現即時監控及智能決策加強風險合規與運營自動化。對外，其能夠刺激營銷／客服，以提升用戶體驗並推動收入增長。通過多智能體協作，企業可實現複雜任務的自主協同調度並提升整體智能化水平。

行業大模型與AI智能體相互賦能可彰顯前者作為通用能力及領域專業知識融合載體的戰略價值，在各行各業需求的上升的同時亦凸顯其強勁的增長潛力。

AI計算芯片比較

AI計算芯片大致可分為GPGPU及非GPU芯片。依託強大的可編程性、高算力及成熟的開發者生態系統，GPGPU歷來都是全球及中國市場大規模AI訓練及推理的主流芯片類型。非GPU芯片主要包括ASIC、FPGA、可重構計算芯片、光計算芯片及其他新興技術路線，在全球市場(尤其是中國)的重要性正日益凸顯。在中國AI計算芯片市場，按收入計，2025年GPGPU市場達到人民幣1,236億元，佔市場的59.2%，而非GPU市場達到人民幣852億元，佔市場的40.8%。按出貨量計，2025年中國AI計算芯片市場達到約4.0百萬顆，其中GPGPU與非GPU芯片各佔約50%的市場份額。

AI計算所用主要芯片比較

		GPGPU	FPGA	ASIC	其他
功能與性能	人工智能應用範圍	• 人工智能訓練✓ • 人工智能推理✓	• 人工智能推理✓	• 人工智能訓練✓ • 人工智能推理✓	• 包括可重構計算芯片、光計算芯片、量子計算芯片及其他新興技術路線。
	靈活性	●* 通用性。功能可通過編程定義，無需永久性物理改動	◑ 半定製化。功能可通過編程定義，無需永久性物理改動，但亦存在限制	○ 完全定製化。硬件設計決定其功能且不可更改	
	算力	● 高於FPGA及ASIC	◔ 落後於GPU及ASIC	◑ 高於FPGA但低於GPU	
功能與性能	開發者生態系統	● 具備成熟的開發環境以及龐大且成熟的開發者生態系統	◕ 開發工作具挑戰性，且與主流成熟的GPU開發環境不兼容	◕ 開發工作具挑戰性，且與主流成熟的GPU開發環境不兼容	
CTU(使用成本)	能耗	◔ 能耗最高	◑ 能耗成本介於GPU及ASIC之間	● 能耗性能具有顯著優勢	
	價格	● 相同用途下最具成本效益	◑ 成本相對較高	◕ 初期開發成本高，但在規模擴大及商業化後，成本將低於FPGA	

*附註：** 哈維球代表不同芯片在統一標準下的相對性能水平。實心圓表示性能優異，空心圓則表示性能欠佳。

資料來源： 弗若斯特沙利文、半導體行業協會

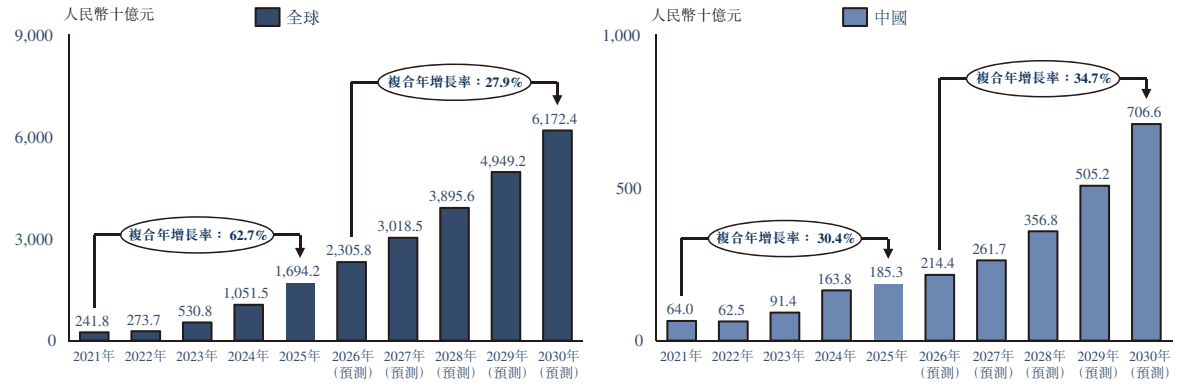
全球及中國GPU市場規模

過去五年，隨著ChatGPT與其他大模型技術等應用取得突破，全球對AI算力的需求激增，使GPU市場規模呈現爆炸性增長，於2025年達到約人民幣1.7萬億元。展望將來，隨著中國大模型技術(如DeepSeek)、人形機器人及智能駕駛技術等技術進步，中國AI技術水平逐步達到國際領先水準。在此背景下，全球經濟正踏入AI普及加速的新時代，全球GPU市場規模預計於2030年將達到人民幣6.2萬億元。

中國GPU市場亦經歷快速增長，由2021年的人民幣640億元增長至2025年的人民幣1,853億元。鑑於AI應用持續多元化及算力基礎設施需求旺盛，預期GPU在中國將加速普及。到2030年，中國GPU市場規模預計將增長至約人民幣7,066億元。

全球GPU市場呈現由國際巨頭主導的頂端集中競爭格局，而中國GPU市場亦有類似的情況，通過政策支持和特定場景的突破，積極推進國產替代品。例如，中華人民共和國工業和信息化部等部門於2023年聯合印發的《算力基礎設施高質量發展行動計劃》，鼓勵採用安全可信的基礎軟硬件，並形成自主可控的解決方案。此外，國務院於2025年發佈的《關於深入實施「人工智能+」行動的意見》支持AI芯片攻堅創新與使能軟件生態培育，從而進一步推動中國GPGPU市場的國產替代。受市場發展階段、政策扶持、需求擴張及供給側改善等綜合因素的推動，預測期內，中國GPU市場的增長速度預計將快於全球GPU市場。首先，全球GPU市場已建立在相對龐大及成熟的市場基礎上，而中國GPU市場仍處於相對早期的發展階段，具備更大的滲透空間。其次，在政策層面，中國的扶持政策持續推動AI芯片、智算基礎設施及自主可控計算解決方案的發展。第三，在需求層面，大模型訓練及推理的快速發展，加上企業AI及智能終端在智慧城市、車載計算、工業應用等場景的加速普及，正推動GPU算力需求持續增長。第四，在供應層面，先進製程產能的逐步提升、國產GPGPU產品持續通過客戶認證及國產硬件與軟件生態系統的持續發展，有望緩解供應限制，並支持GPU在中國更廣泛的商業化及部署。

2021年至2030年(預測)全球及中國GPU市場規模(按收入計)



資料來源： 弗若斯特沙利文、CSIA

全球及中國GPU行業的主要發展趨勢及驅動因素

- ***下游市場旺盛推動需求增長***

GPU在多個下游市場發揮著重要作用，例如大模型訓練與推理、邊緣計算、元宇宙、數字孿生、遊戲及科學計算領域。隨著該等市場繼續快速發展，GPU產品的需求已大幅增加。在當今AI普及加速時代，GPU(尤其是GPGPU)憑藉其強大並行計算能力加速模型訓練與推理，已成為向大模型應用轉型的關鍵基礎設施。

- ***企業AI推理應用需求激增***

隨著行業大模型廣泛普及，企業AI推理需求呈現爆發式增長。企業AI應用呈現出顯著特徵：行業大模型的本地化部署要求具備即時推理能力，如智慧城市的實時影像分析、工業質檢的精準缺陷判斷等。企業AI計算通過將算力下沉至企業網絡私有節點降低數據傳輸延遲，直接滿足即時推理需求，推動近年的急速增長。與此同時，企業AI的爆發式增長對部署於私有節點的GPU提出了極高的能效要求，推動了兼具高算力與低能耗的異構GPU的發展，該等GPU已成為支撐行業大模型在企業端高效落地的重要硬件基石。

中國GPGPU市場

GPGPU的定義及核心組成

GPGPU是一種結合高度並行架構設計及通用計算能力的處理器。面對人工智能對多元高效算力日益增長的需求，GPGPU融合GPU的並行計算優勢，針對AI應用場景進行系統性優化。這使其在持續提升能效比等核心性能指標的同時，維持強大算力輸出並利用成熟的開發生態。憑藉高度並行架構與大量計算單元，GPGPU能夠高效執行科學計算、數據分析及深度學習模型訓練與推理等非圖形化的大規模並行計算任務，並廣泛應用於深度學習、科學計算、大數據分析、AI等需要高性能計算的領域。GPGPU產品的性能通常根據客觀且可量化的技術指標綜合評估，包括計算性能、支持的計算精度、內存容量及內存帶寬。特別是，高性能GPGPU產品一般指專為企業AI及數據中心應用設計的GPGPU加速器產品，其FP16計算性能約為50 TFLOPS或以上，或INT8計算性能約為200 TOPS或以上，內存容量約為32GB或以上，內存帶寬約為200 GB/s或以上，並支持包括FP16/BF16及INT8在內的多種計算精度。

GPGPU包含計算單元、存儲單元及控制單元等核心組件。計算單元由眾多流處理器構成，能夠運用並行計算高效處理各類複雜運算。存儲單元由片內緩存及顯存組成，其中緩存能快速存取常用數據，顯存則用於儲存大量數據。控制單元則管理與調度計算過程，並調配資源讓計算單元從存儲單元中獲取數據，執行運算後回傳結果。該等組件緊密合作，確保高性能與高效計算。

異構計算及其在GPGPU中的應用

異構計算

異構計算是一種專門的並行分佈式計算方式，利用不同類型指令集及體系架構的處理單元組成單一系統。其核心應用場景涵蓋公有雲AI及企業AI。異構計算的價值在於協調多種不同計算單元，以滿足多樣化計算需求，並提升整體性能。例如，典型的CPU+GPU組合可分別處理不同特性的數據結構與存取模式，從而提升效率及可擴展性並克服單一架構計算的局限性。

異構計算硬件架構

物理架構上，異構計算架構可分為兩大類別：

片外異構：一種集成多個架構、功能或工藝各異的芯片(包括GPU、ASIC及CPU)至單一系統的技術解決方案。該等芯片相互作用並協作，以實現異構集成。

片內異構：一種將具有獨立架構與接口的功能單元集成至同一芯片上的系統集成。通信通過片內總線進行，實現跨功能分區協同以支持高性能計算及數據中心。

張量引擎與GPGPU引擎的片內異構集成、調度管理、內存管理、數據通信與功耗管理，實現了動態資源分配與協同加速，提升了芯片的整體性能與適配性。調度管理單元作為性能調度中樞，將任務分配至適配調度器，並通過資源衝突檢測避免算力與內存爭用及數據壅塞，從而提升效率並降低延遲。圖形內存管理單元作為圖形內存調控中樞，依託定製DDR控制器及稀疏矩陣壓縮引擎實現硬件級顯存壓縮，並依靠斷點續存，在斷電後10秒內即可恢復任務。數據通信單元作為數據傳輸的中樞，採用動態帶寬分配器與硬件數據隊列優化，促進數據流高效傳輸。功耗管理單元作為能效管理中樞，利用動態頻率調整算法與溫度響應模型進行負載預測與電壓調控，並在過熱時通過梯度降頻穩定算力，實現能效控制。此外，為適配多樣化軟件生態，生態兼容層作為異構芯片軟件適配基礎，通過CUDA/OpenCL轉換器轉換代碼指令，以動態補丁引擎修復運行兼容性問題，並通過虛擬化接口支持多租戶隔離，確保生態的適配性與資源的靈活分配。

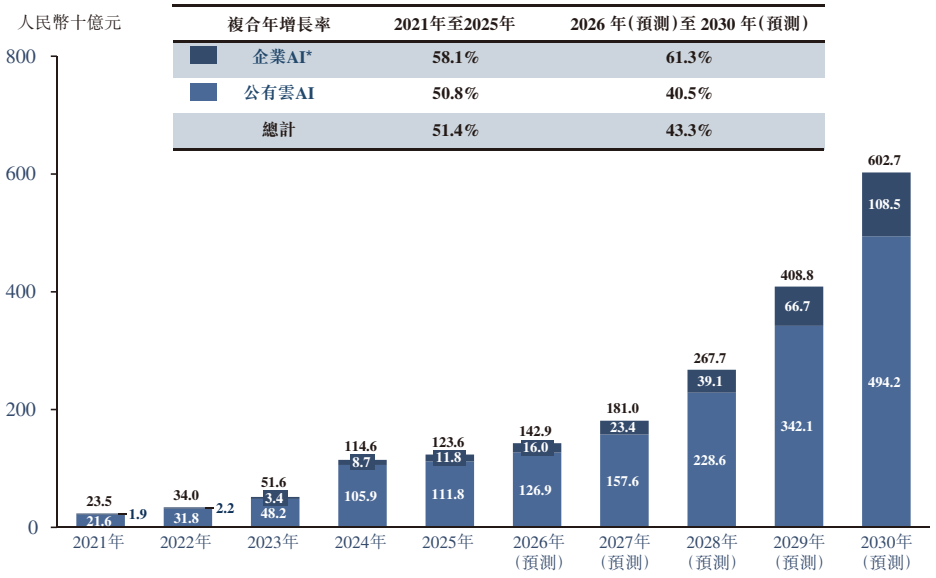
中國GPGPU市場規模及應用分佈

隨著對支持AI訓練推理、科學計算加速等先進應用的通用高性能計算需求不斷攀升，GPGPU需求亦同步激增。中國GPGPU市場呈現快速增長態勢，由2021年的人民幣235億元增至2025年的人民幣1,236億元。未來，隨著AI應用深度拓展及各行業數字化轉型加速，對GPGPU等核心計算基礎設施的需求預計將迎來爆發式增長。預計到2030年，中國GPGPU市場將增至人民幣6,027億元。

就市場結構而言，GPGPU市場可細分為企業與公有雲AI兩大場景。中國的企業AI GPGPU市場按58.1%的複合年增長率由2021年的人民幣19億元增加至2025年的人民幣118億元。同期，中國的公有雲AI GPGPU市場按50.8%的複合年增長率由2021年的人民幣216億元增加至2025年的人民幣1,118億元。於2026年至2030年，企業AI GPGPU市場預期會維持61.3%的高複合年增長率，並於

2030年達人民幣1,085億元，而公有雲AI GPGPU市場預期將按40.5%的複合年增長率增長，並於2030年達人民幣4,942億元。受兩個分部共同推動，中國的整體GPGPU市場於2026年至2030年間預期將實現43.3%的複合年增長率。

2021年至2030年(預測)按應用領域細分的中國GPGPU市場規模(按收入計)



附註：

(1) 公有雲AI指雲服務提供商在公有雲數據中心集中部署計算資源，並通過共享雲平台提供AI計算服務的AI計算場景。

(2) 企業AI指企業在私有雲、本地數據中心、專用計算集群或邊緣節點部署或使用AI計算資源，以支持行業特定應用的AI計算場景。

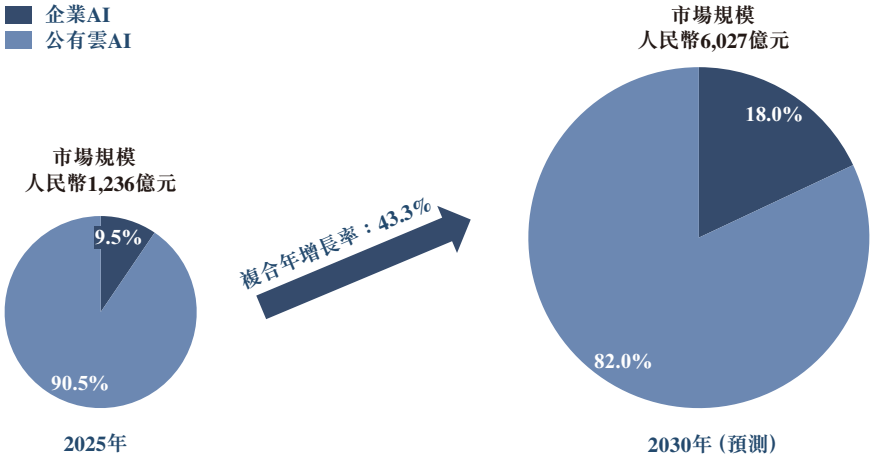
資料來源： 弗若斯特沙利文、CSIA

中國公有雲AI GPGPU市場增長放緩主要歸因於供給側制約。於2025年，NVIDIA及AMD的市場份額合共達90%以上。受出口限制的影響，若干NVIDIA及AMD產品在中國長期供應受限。特別是，公有雲AI數據中心通常需要高性能GPGPU以支持大規模模型訓練及推理，因此該領域對高端產品的可用性相對更為敏感。與此同時，國內GPGPU供應商由於先進製程產能受限、產品認證及導入週期較長，無法快速填補由此產生的供應缺口。因此，公有雲AI GPGPU市場的複合年增長率預計將由2021年至2025年的50.8%放緩至2026年(預測)至2030年(預測)的40.5%，但其市場規模預計將由2025年的人民幣1,118億元增至2030年的人民幣4,942億元。隨著先進製程產能逐步提升及更多國產產品完成客戶認證，供應短缺有望大幅緩解。

相比之下，NVIDIA及AMD就企業AI應用供應的GPGPU產品通常有別於公有雲AI數據中心所用的GB200及H200等高端產品，故受相關制約的影響較輕。因此，企業AI市場供應受限程度遠輕於公有雲AI市場。此外，企業AI GPGPU市場於2025年的市場規模為人民幣118億元，遠低於公有雲AI GPGPU市場的人民幣1,118億元，因此增長率較高。企業AI市場亦處於滲透的相對早期階段，私有雲、本地數據中心、專用計算集群、邊緣節點及其他行業特定應用的需求繼續擴大。因此，企業AI GPGPU市場的複合年增長率預計將由2021年至2025年的58.1%增至2026年(預測)至2030年(預測)的61.3%，到2030年其市場規模將達人民幣1,085億元。

於中國GPGPU市場的應用結構中，呈現出公有雲AI與企業AI協同發展的格局。於2025年，中國GPGPU市場總規模為人民幣1,236億元，其中公有雲AI應用以90.5%的市場份額佔據主導市場，而企業AI應用份額為9.5%。到2030年，市場規模預計將達人民幣6,027億元，2026年至2030年的複合年增長率為43.3%。雖然公有雲AI應用將繼續佔據大部分市場份額，企業AI應用的佔比預計將大幅提升至18.0%，反映在算力提升、算法優化等技術賦能下，企業AI計算場景的需求正逐步釋放。

2025年至2030年(預測)中國GPGPU下游應用分佈



資料來源：弗若斯特沙利文、半導體行業協會、中國信息通信研究院

中國GPGPU市場的驅動因素及發展趨勢

中國GPGPU市場正由多項因素推動其快速發展：

- ***對大模型訓練與推理的需求呈現爆炸式增長***

近年來，大模型呈現爆炸式增長，其參數規模呈指數級上升，已逐步從數十億參數攀升至數千億甚至更高。該等模型的訓練需處理海量數據並執行梯度計算、反向傳播等複雜運算，這對算力提出了極高要求。GPGPU憑藉其強大的並行計算能力、高吞吐量及處理大規模矩陣運算的能力，已成為滿足此類需求的核心硬件。隨著AI技術的持續普及，AI推理功能的應用場景亦不斷拓展至金融、工業、教育等更多行業領域，進一步擴大了GPGPU的需求。

- ***智能終端的廣泛佈局帶動企業AI算力需求的增長***

數字化轉型步伐加速推動智慧城市、車載計算、工業質檢等眾多應用場景蓬勃發展，該等場景均需要強大高效的計算能力。例如，在智慧城市建設方面，需在城市各個角落部署傳感器及監控設備，實時採集並處理數據，以便支持交通管理、環境監測、及安全預警等功能。自動駕駛車輛的車載計算需要快速準確應對複雜路況以保證安全性。該等場景要求GPGPU具備低延遲、高性能及高能效特性。GPGPU能夠將強大的算力部署至靠近數據源的邊緣端，避免數據傳輸至雲端造成延遲問題，顯著提升響應速度及處理效率，推動企業AI快速普及。

- ***拓展業務邊界，發展全棧算力基礎設施***

在中國，以創新為驅動的GPGPU企業正突破傳統芯片製造商的邊界，加速從硬件提供商向全棧算力基礎設施建設者轉型。該等企業以核心計算技術突破為支點，建立涵蓋硬件、軟件及服務的結構化生態系統。在硬件層面，圍繞自主研發GPU延伸產業鏈；在軟件層面，確保與操作系統及優化算法框架的兼容性；在服務層面，提供面向多樣化應用場景的增值解決方案。這集成模式為中國AI產業實現自主可控及規模化落地奠定基礎。

中國GPGPU行業競爭格局

GPGPU行業競爭格局概覽

中國GPGPU市場呈現國際巨頭主導高端領域，而國內參與者尋求差異化突破的格局。在多方利益相關者的支持下，國內企業發展前景廣闊，預計市場份額逐步提升。市場競爭呈現分層格局：領先的國際企業憑藉先發優勢主導通用計算領域，豐富的產品組合橫跨多元應用場景。其軟件平台與解決方案樹立了行業標竿。相反，國內企業則透過差異化創新加速發展，尤其是企業AI計算與行業專用應用。透過聚焦架構設計與能效及依託本土供應鏈優勢，該等公司正縮小與國際競爭對手的差距並開發契合國內市場的解決方案。按2025年中國GPGPU市場收入計，本公司排名前十，市場佔有率低於1.0%。

2025年中國GPGPU公司排名及市場份額(按收入計)

排名	公司	收入(人民幣十億元)	市場份額(%)
1	公司A	101.2	81.9%
2	公司B	12.0	9.7%
3	公司E	3.1	2.5%
4	公司F	1.6	1.3%
5	公司G	1.5	1.2%
…	其他	4.2	3.4%
	總計	123.6	100.0%

資料來源：弗若斯特沙利文、專家訪問、公司年度報告

企業AI場景競爭格局

按2025年收入計，中國GPGPU市場在企業AI場景競爭格局中，公司A以人民幣98億元領先，公司B以人民幣14億元緊隨其後。本公司在企業AI場景的技術適應與市場拓展中位居前列，2025年收入為人民幣115.9百萬元。

2025年中國企業AI GPGPU市場排名(按收入計)

排名	公司	收入(人民幣十億元)	市場份額(%)
1	公司A	9.8	83.1%
2	公司B	1.4	12.3%
3	本公司	0.12	1.0%
4	公司C	0.10	0.9%
5	公司D	0.09	0.8%

附註：

1. 公司A是一家在納斯達克上市的半導體公司，成立於1993年，總部位於美國加利福尼亞州。其主要專注於GPU、AI計算平台、加速計算解決方案及相關軟硬件生態系統，業務覆蓋全球主要市場。公司A於2025財政年度錄得總收入約1,305億美元。公司A被視為全球領先的GPU及AI計算芯片公司。按收入計，公司A於2025年在中國整體AI計算芯片市場的市場份額約為48.5%，於2026年上半年進一步下跌至約37.7%。有關市場發展整體上與公司A在中國的市場份額減少趨勢一致。誠如上文所披露，AI計算芯片可大致分為GPGPU及非GPU芯片。雖然公司A及本公司主要提供GPGPU產品，部分頂尖中國國產AI計算芯片供應商主要提供非GPU芯片。於2025年，按收入計，GPGPU佔中國整體AI計算芯片市場約59.2%。在中國AI計算芯片市場的GPGPU分部，公司A保持其主導地位，市場份額約為81.9%，相當於整體AI計算芯片市場約48.5%。與此同時，中國國產供應商已在非GPU分部建立重大市場地位。因此，公司A在中國GPGPU市場的81.9%市場份額反映其在中國整體AI計算芯片市場的GPGPU分部中的市場地位。同樣地，上文所披露的中國企業AI GPGPU市場排名具體涵蓋在企業AI場景部署的GPGPU產品，而非在該場景中所用的所有類型的AI計算芯片，因此公司A在企業AI GPGPU市場的83.1%市場份額應理解為僅限於該特定市場分部。

2. 公司B是一家在納斯達克上市的半導體公司，成立於1969年，總部位於美國加利福尼亞州。其主要專注於高性能CPU、GPU、數據中心加速器、自適應計算產品及嵌入式半導體解決方案，業務覆蓋全球主要市場。公司B於2025財政年度錄得總收入約346億美元。公司B被視為全球領先的高性能計算及AI加速器芯片公司。

3. 公司C是一家成立於2018年的私營半導體公司，總部位於中國上海。其主要專注於用於AI計算、圖形渲染、視頻處理及內容生成的高端GPU芯片及全棧芯片解決方案，業務主要在中國。公司C被視為中國服務於數據中心、邊緣計算及圖形相關應用場景的關鍵GPU芯片公司。

4. 公司D是一家在香港聯交所主板上市的半導體公司，成立於2015年，總部位於中國上海。其主要專注於GPGPU產品及AI計算解決方案，業務主要在中國。公司D於2025年錄得收入約人民幣10億元。公司D被視為中國領先的GPGPU及AI計算解決方案公司。

5. 公司E是一家在上海證券交易所科創板上市的半導體公司，成立於2014年，總部位於中國天津。其主要專注於高端處理器(包括CPU及DCU)，業務主要在中國。公司E於2025年錄得營業收入約人民幣144億元。公司E被視為中國領先的高端處理器及AI計算芯片公司。

6. 公司F是一家在上海證券交易所科創板上市的半導體公司，成立於2020年，總部位於中國上海。其主要專注於用於AI訓練及推理、通用計算及圖形渲染的全棧高性能GPU產品及計算平台，業務主要在中國。公司F於2025年錄得營業收入約人民幣16億元。公司F被視為中國關鍵的全棧高性能GPU公司。

7. 公司G是一家在上海證券交易所科創板上市的半導體公司，成立於2020年，總部位於中國北京。其主要專注於用於AI計算加速、圖形渲染、物理仿真、科學計算及超高清視頻處理的全功能GPU芯片及相關計算加速平台，業務主要在中國。公司G於2025年錄得營業收入約人民幣15億元。公司G被視為中國領先的全功能GPU公司。

資料來源：弗若斯特沙利文、專家訪問、公司年度報告

平均售價分析

GPGPU市場的一般價格區間為每單位人民幣10,000元至人民幣200,000元，依據效能、配置和品牌而有所不同。過往，在AI需求激增及全球半導體供應緊張的推動下，2020年至2022年間價格出現顯著波動。儘管同樣產品型號的價格在其成熟後會逐年下降，但產品迭代會推動價格上調，而新推出的型號通常會在性能升級和先進技術配置的支持下大幅漲價。

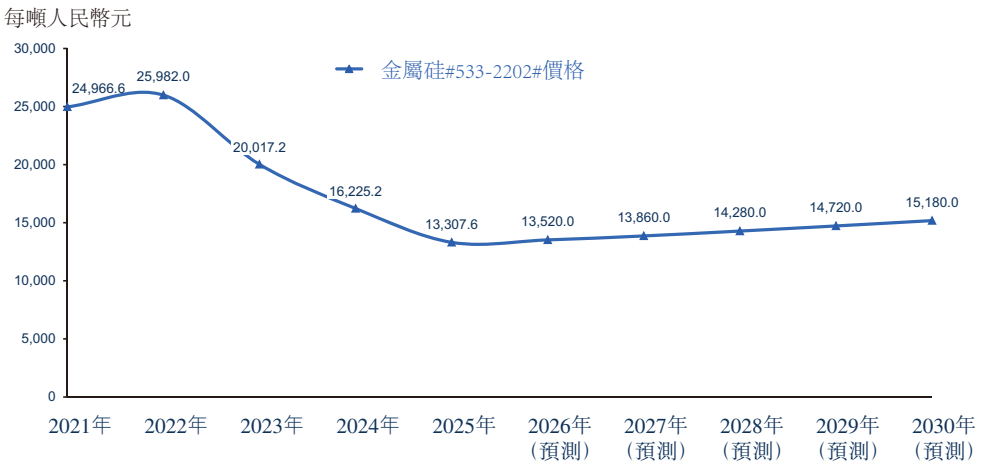
原材料分析

本公司採用fabless商業模式，並不會直接購買或處理原材料。因此，於評估原材料價格趨勢時，本公司主要參考晶圓代工廠使用的主要原材料之定價。

若干主要原材料的價格於2020年至2022年間顯著波動，其後逐步轉穩。例如，533級冶金硅為晶圓製造使用的主要原材料，為一種工業級硅及半導體供應鏈中的上游硅基原料，可進一步加工成用於晶圓製造更高純度硅材料及硅晶圓。533級冶金硅的價格由2020年的每噸人民幣13,240.8元增加至2021年的每噸人民幣約25,000元，並進一步增加至2022年的高峰每噸人民幣25,982.0元。價格自此穩步下降，跌至2025年的人民幣13,307.6元。展望將來，由於供應條件逐漸改善及需求增長放緩，供需趨於平衡，而價格亦恢復至更正常的水平。於2021年至2022年觀察到的波幅主要由暫時性的供需錯配所推動。於需求側，下游領域的快速增長導致整體半導體需求飆升，繼而增加晶圓代工廠的訂單量及原材料採購數量。於此期間，大規模採用電動車，連同車輛電子化及智能化的加速趨勢，均大幅提升了每輛車輛的半導體搭載量，進一步刺激了總需求。此外，地緣政治愈來愈多的不確定因素令部分下游客戶預先下達訂單及囤積即時生產需要以外的存貨，加劇了半導體價值鏈的短期需求壓力。於供應側，半導體生產能力於結構上集中於有限數目的領先晶圓代工廠，而疫情相關的中斷亦限制了額外產能的及時釋放。強勁需求增長與供應受限結合，導致晶圓製造階段的生產條件緊張，此情況向上游傳導，令原材料價格上升。

該等原材料價格波動透過晶圓代工廠收取的晶圓製造費用的變動，間接且通常存在時間延遲地影響本集團產品的歷史售價。該等費用不僅受上游原材料成本的影響，亦受製程節點、產能利用率、製造良率及商業條款等因素的影響。於2021年及2022年，原材料價格上升連同晶圓製造產能緊張，增加了晶圓代工廠的成本壓力。倘該等額外成本反映於晶圓製造費用中，則會增加本集團的生產成本，並在產品定價時予以考量，從而對其產品售價產生了上行壓力。隨後原材料價格的下降緩解了該等成本壓力，但並未導致本集團產品售價出現即時或相應的減少。這是由於原材料成本僅為晶圓製造費用的組成部分之一，且該等成本的變動僅沿供應鏈部分且逐步地傳導。因此，本集團產品的歷史售價並未與533級冶金硅的價格直接成比例變動，而是主要受上述產品成熟度與迭代動態、產品性能與配置、產品組合以及現行市場供需狀況的影響。

2021年至2030年(預測)中國的金屬硅#533-2202#價格

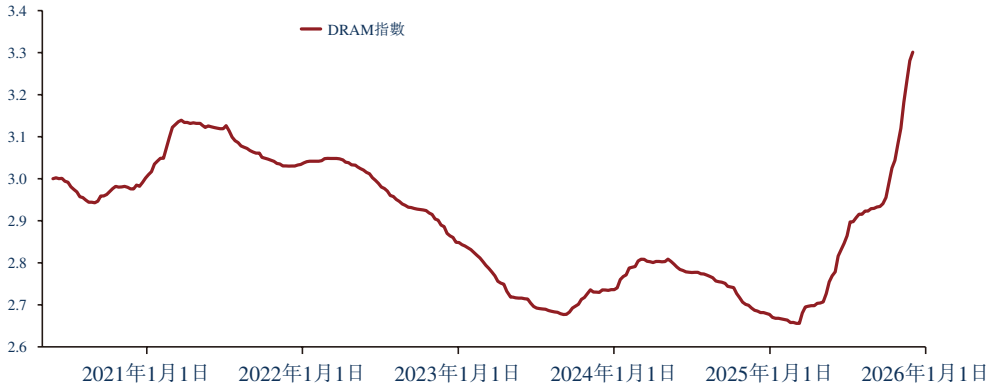


資料來源：弗若斯特沙利文、半導體行業協會、中國有色金屬工業協會

DRAM閃存價格分析

近年來，DRAM價格經歷了階段性波動。於2021年，受全球供應鏈限制及遠程工作及相關終端市場的強勁需求影響，DRAM價格維持在高位。於2022年，由於消費電子需求疲軟及全行業去庫存，價格大幅下跌。於2023年及2024年波動後，2025年DRAM價格大幅上漲，主要受AI應用快速部署、AI服務器需求上升、存儲廠商產能向HBM及DDR5等高端產品轉移，以及供應持續受限等因素所推動。

DRAM閃存指數



資料來源： CFM、弗若斯特沙利文

GPGPU行業准入壁壘

- ***技術壁壘***

在GPGPU行業，技術壁壘極為顯著。高精度浮點運算作為行業的核心技術之一，需在芯片架構設計、電路優化及算法協同等多個層面達到高標準。所涉及的技術專業知識的複雜度與深度不容小覷。此外，隨著企業AI計算領域等應用場景的擴展，企業不僅需確保強大的算力輸出，亦須優化能效這一關鍵性能指標。這意味著企業須精細打磨諸多環節，包括選擇先進製造工藝、

電源管理模組設計及計算單元能效優化。該等工作需投入大量時間、人力及資本進行研發，同時面臨諸多技術挑戰。從基礎理論突破到工程實踐應用，這一過程充滿困難，使得新入行者突破該等壁壘變得極為困難。

- ***供應鏈壁壘***

供應鏈挑戰是GPGPU設計公司面對的另一項重大准入壁壘。在上游，獲取晶圓代工廠資源需建立深度穩定的合作夥伴關係。精準掌握產能與工藝能力至關重要，同時需應對複雜的商業談判以確保高質量晶圓的穩定供應。下游獲取核心客戶資源同樣充滿挑戰，須滿足客戶多維度的嚴苛要求，並具備快速應對需求變化的能力。此外，整合生態系統資源並確保產業鏈無縫協作，需長期積累供應鏈管理專業知識。該等因素共同構成了顯著的供應鏈壁壘。