

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不同的政府官方刊物、可得的公開市場研究資料來源及獨立供應商的其他資料來源，以及灼識行業諮詢有限公司（「灼識諮詢」）編製的獨立行業報告。我們委聘灼識諮詢編製有關[編纂]的獨立行業報告（「灼識諮詢報告」）。我們認為，該等資料的來源為該等資料的適當來源，且我們在提取及重現該等資料時已採取合理審慎措施。我們並無理由相信該等資料屬虛假或具有誤導性，亦無理由相信遺漏任何事實會導致該等資料屬虛假或具有誤導性。來自政府官方來源的資料並無經我們、獨家保薦人、[編纂]、任何[編纂]、彼等各自的任何董事、監事及顧問或參與[編纂]的任何其他人士或各方獨立核實，且概無就其準確性發表任何聲明。因此，本文件所載來自政府官方來源的資料未必準確。

中國AI存儲行業概覽

AI基礎設施概覽

隨著新一波AI發展浪潮來臨，大模型的訓練和推理需要進行海量數據處理和高併發計算，在過去3年內對包括計算、存儲和網絡在內的AI基礎設施的需求顯著增加，並將在未來保持這樣的增長趨勢。

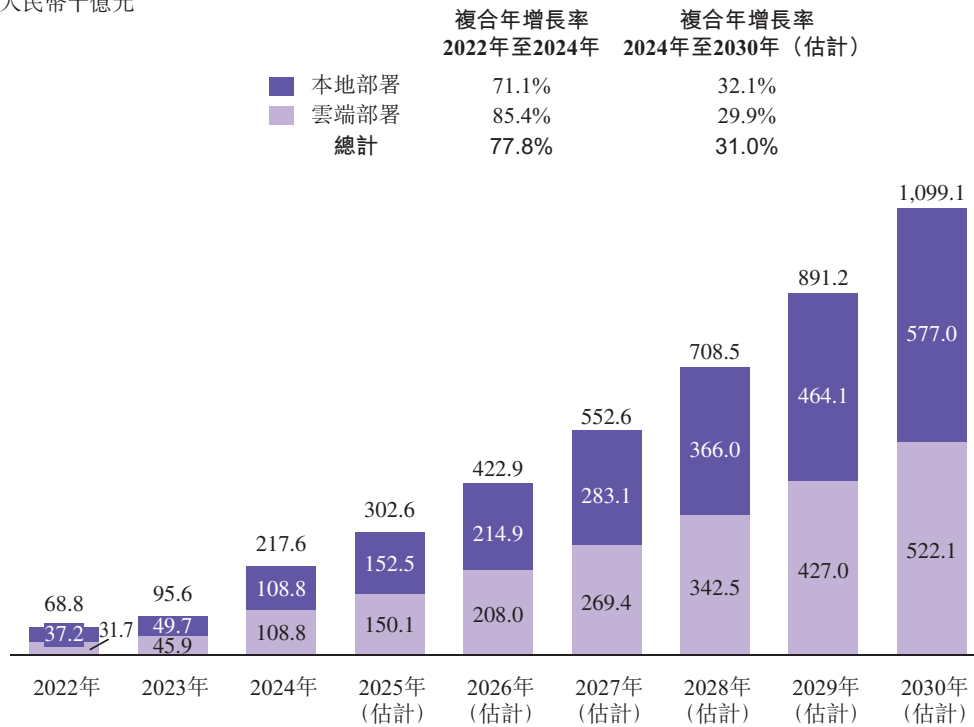
根據部署模式的不同，AI基礎設施可分為雲端部署和本地部署兩類。雲端部署的AI基礎設施由第三方雲服務商運營，用戶在雲端按需租用計算、存儲與網絡資源，具有彈性高、部署快、前期投入低等特點，尤其在互聯網、中小企業、初創公司中廣泛應用。而本地部署的AI基礎設施部署在企業自有或受控環境中，資源完全由企業掌控，更注重數據主權與系統可控性。隨著AI在智能製造、金融機構、電信運營商及技術企業等領域的深入應用，有越來越多需要構建可控、高性能的AI基礎設施的客戶開始選擇本地部署。

根據灼識諮詢的資料，中國AI基礎設施市場規模在2024年達到人民幣2,176億元，佔全球AI基礎設施市場約15%。預計市場規模將在2030年擴大至人民幣10,991億元，複合年增長率為31.0%，預測屆時中國在全球市場的佔比將上升至約25%。其中，預計本地部署AI基礎設施市場規模將從2024年的人民幣1,088億元增長至2030年的人民幣5,770億元，複合年增長率為32.1%。

行業概覽

中國AI基礎設施市場規模，按部署形式劃分，2022年至2030年（估計）

人民幣十億元



資料來源：IDC、行業專家訪談、灼識諮詢

附註：歷史數據自2022年起呈列，是由於隨著大型語言模型於2022年前後在全球興起，中國AI基礎設施和AI存儲的實質性商用部署加速，該時間段的數據更具相關性和代表性。

行業概覽

AI存儲概覽

AI的工作流程通常包括數據採集與清洗、模型訓練以及推理應用等環節。在這一過程中，大量非結構化數據需要被持續收集、處理、計算與再存儲，對數據流的穩定性與吞吐能力要求極高。存儲系統貫穿了AI全生命周期的數據讀寫作業，其性能、可擴展性及安全性直接影響各類計算任務的效率、成本和數據治理水平。隨著數據體量的迅速增長，高效、安全的存儲系統已成為極其關鍵的AI基礎設施組成。

此外，存儲的客戶黏性在整個IT系統中都是極高的，一旦部署特定存儲系統，數據遷移成本包括數據搬遷、系統兼容性調整、業務中斷風險以及運維流程重構等均非常高。這使得企業在選擇存儲系統時極為謹慎。

雲計算時代的存儲系統主要面向事務型與歸檔型負載，優化方向集中在容量擴展、成本控制和訪問可靠性，在設計上偏向靜態數據的長久保存與低頻次訪問。但在面對有高併發、多節點並行訪問和低延遲特徵的AI任務時，這樣的傳統架構易出現數據讀寫的瓶頸，並最終限制算力發揮。因此，AI的普及應用也要求存儲系統的協同進化，對存儲的需求從容量管理升級為智能數據底座，其重要性隨AI應用規模的擴大而持續提升。

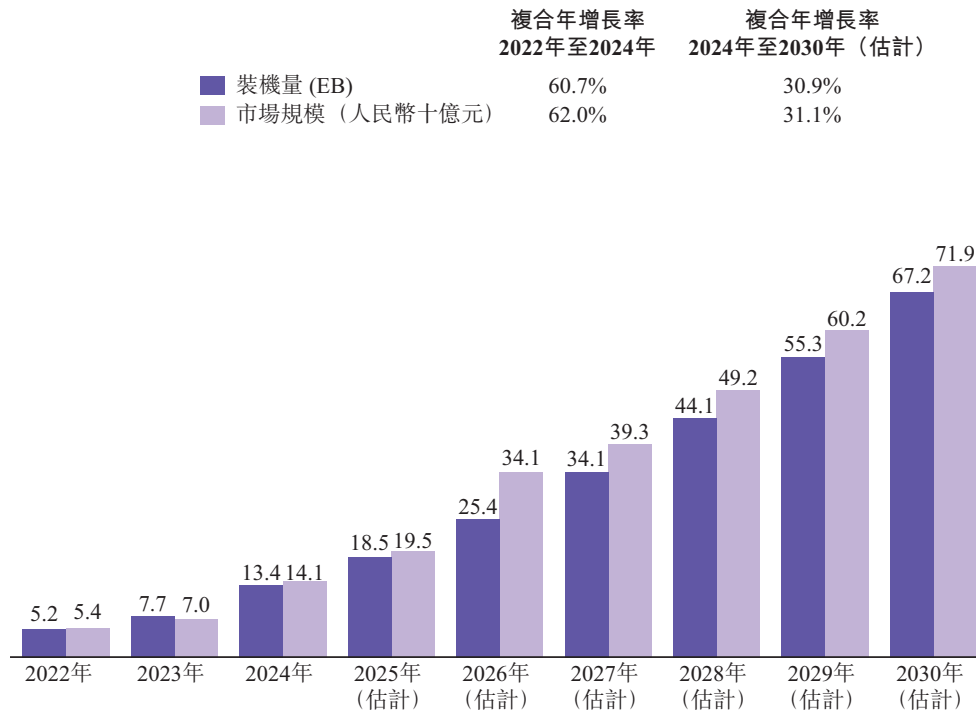
AI存儲是指專為AI任務設計並優化的數據存儲系統，旨在滿足AI任務中對超大容量數據保存和高性能數據讀寫的要求。未來，企業內部長期積累的、高價值的專有數據，將成為其在AI浪潮中的核心競爭力。本地部署的AI存儲能讓企業在安全可控地管理數據，防止洩露與外部使用。此外，本地部署支持數據的持續回流與迭代，使企業能夠構建自主的數據飛輪，形成自己的智能壁壘。這意味着，本地部署的AI存儲不僅是數據主權的保障，更是長期競爭力的基石。

根據灼識諮詢的資料，中國本地部署AI存儲裝機量在2024年達到13.4EB，並預計在2030年擴大至67.2EB，複合年增長率為30.9%。同時，中國本地部署AI存儲市場規模在2024年達到人民幣141億元，並預計在2030年增長至人民幣719億元，複合年增長率為31.1%。¹

¹ 1 EB=1,000 PB，1 PB = 1,000 TB。

行業概覽

中國本地部署AI存儲裝機量與市場規模，2022年至2030年（估計）



資料來源：IDC、灼識諮詢

AI存儲產業鏈

AI存儲產業鏈主要包括上游基礎硬件供應商、中游軟件及軟硬一體解決方案提供商和下游應用終端客戶。

上游：AI存儲所需的核心硬件，包括存儲介質（HDD、SSD等）、網絡設備、服務器機櫃等。

中游：AI存儲解決方案提供商，通過配套分布式文件系統、數據調度與緩存管理等底層系統軟件，整合軟硬件資源，提供解決方案，是產業鏈的價值核心。

下游：AI存儲終端客戶，涵蓋智能製造、金融、電信、互聯網與科技、機器人及自動駕駛等領域的企業。該等用戶通常對數據主權、存儲性能及可擴展性均具有較高要求。

行業概覽

AI存儲產業鏈

★ 本公司所在環節



資料來源：灼識諮詢

行業概覽

中國分布式AI存儲行業概覽

存儲系統面向AI任務的挑戰

訓練側的「I/O牆」。目前主流的大模型參數規模普遍達到十億以上，訓練集群通常有上千乃至上萬塊AI芯片在同時工作。訓練過程需要AI芯片不斷讀取海量數據、並定期保存模型快照(checkpoint)，產生短時間內的巨大讀寫壓力。同時，訓練期間會有大量小文件被隨機訪問，顯著增加存儲系統的管理和響應負擔。傳統存儲難以支撐這種高強度併發訪問，容易導致AI芯片長時間閒置等待數據，整體訓練效率大幅下降。

推理側的「內存牆」。更長的上下文和更複雜的交互，是大模型的一大發展趨勢。為了支持這些特性，大模型需要保留大量中間計算結果作為緩存，而緩存通常存放在AI芯片的顯存中，來保證其能被高速讀寫。未來，可以預見這部分緩存的規模將迅速膨脹，遠超AI芯片的顯存容量，而傳統存儲系統不具備媲美緩存的高速讀寫性能，因此企業不得不為顯存而繼續採購昂貴的AI芯片，導致推理成本仍有極大的優化空間。

數據側的「重力牆」。企業的數據往往分散在不同系統中，原始數據可能在大容量的存儲系統中里，而訓練又需要高讀寫性能的存儲支撐。數據在這些系統之間遷移耗時長、成本高，還容易造成數據孤島和版本混亂的問題。

分布式存儲是AI存儲的最佳解決方案

根據架構不同，AI存儲可分為集中式存儲與分布式存儲兩類。

集中式存儲以單一存儲節點為核心，數據管理與訪問均在該節點上完成，架構相對簡單，適用於數據規模較小、併發需求有限的AI場景，但難以滿足日益增長的性能和擴展性要求。分布式AI存儲則通過多個存儲節點的協同工作，將數據分片存儲在不同節點上，實現數據的並行讀寫與負載均衡。這樣的設計既能提供大容量和高吞吐性能，也能通過副本冗餘與容錯機制，確保數據在高負載條件下的穩定性與可靠性。

行業概覽

因此，對於AI訓練和推理任務，分布式存儲已成為主流架構選擇。下表總結了集中式與分布式存儲的特點。

	集中式存儲	分布式存儲
架構形態.....	由少量控制器承擔全部I/O，I/O能力無法線性擴展	可以多節點橫向擴展，I/O能力隨節點數量線性增長
I/O性能.....	受限於控制器數量和帶寬，面對大型計算集群時容易成為瓶頸	支持多節點並行處理的大帶寬讀寫，I/O能力隨節點數量擴大而提升，可支持大型計算集群
推理緩存支持....	控制器的緩存容量與帶寬有限、且無法橫向擴展，僅能依靠AI芯片顯存，難以承載TB~PB級的推理緩存	作為外存，可與AI芯片顯存、內存共同形成多級緩存架構，可隨集群規模擴展緩存容量與帶寬，適配推理的高併發小文件讀寫
可靠性與容錯性..	系統故障影響大，尤其是對於恢復緩慢的高容量系統	多節點分布式管理，可實現數據冗餘，並支持各類數據透明遷移和動態調度，降低數據移動成本
適用場景.....	中小規模AI實驗室、單機推理任務	大規模AI訓練、多集群推理、RAG數據湖、向量數據庫、KV Cache、企業級AI平台

行業概覽

分布式AI存儲架構以軟件為中心，核心軟件層負責元數據管理、數據分片、負載均衡與冗餘保護，硬件資源則被統一抽象與調度，形成靈活可擴展的資源池，從而實現系統的橫向擴展與高效交付。

分布式AI存儲的架構



資料來源：灼識諮詢

對於本地部署存儲而言，實體部署架構直接可見，且與系統設計緊密耦合，因此區分集中式存儲和分佈式存儲具有意義。相比之下，雲端存儲將底層基礎設施對用戶進行抽象化處理，且預設採用分佈式架構，因此從用戶或服務的角度來看，這種區分基本上無關重要。因此，本文中對分佈式存儲的討論僅指本地部署存儲。

行業概覽

分布式AI存儲的產品分類

分布式AI存儲的產品根據任務需求可分為兩類：AI數據湖存儲和AI訓推存儲。

AI數據湖存儲

AI數據湖存儲主要面向AI生命週期中長期留存的溫、冷數據，如訓練樣本歸檔、歷史模型版本、日誌數據及中間特徵文件等。此類系統強調高容量密度、成本效率與數據可靠性，通過分層存儲、壓縮編碼與副本冗餘技術實現低成本的數據持久化。

其部署規模通常與整體數據資產規模直接掛鉤。對於擁有海量數據資源的客戶，AI數據湖存儲構成其數據湖與知識庫體系的底座，用於支撐長期的數據積累與治理。系統性能指標以容量利用率、存儲成本、數據恢復可靠性等為核心。

AI訓推存儲

AI訓推存儲聚焦AI計算階段對熱數據的高併發訪問需求，是模型訓練與推理性能釋放的關鍵環節。其系統架構通常採用高帶寬互聯、分布式文件系統與數據分片機制，以實現高吞吐、低延遲、強一致性的訪問體驗。

行業概覽

與AI數據湖存儲不同，AI訓推存儲是AI基礎設施中與算力最緊密耦合的系統，其部署規模與AI算力規模高度相關，強調數據局部性與算存協同優化，常與AI集群的高速網絡深度綁定，以減少數據傳輸瓶頸。

根據灼識諮詢，中國分布式AI存儲裝機量在2024年達到10.3EB，並預計在2030年擴大至60.5EB，複合年增長率為34.3%。於2024年，AI數據湖存儲佔市場的約80%，而AI訓推存儲佔20%。預計以上佔比將於2030年分別轉為約72%及28%。

中國分布式AI存儲裝機量，按產品劃分，2022年至2030年（估計）

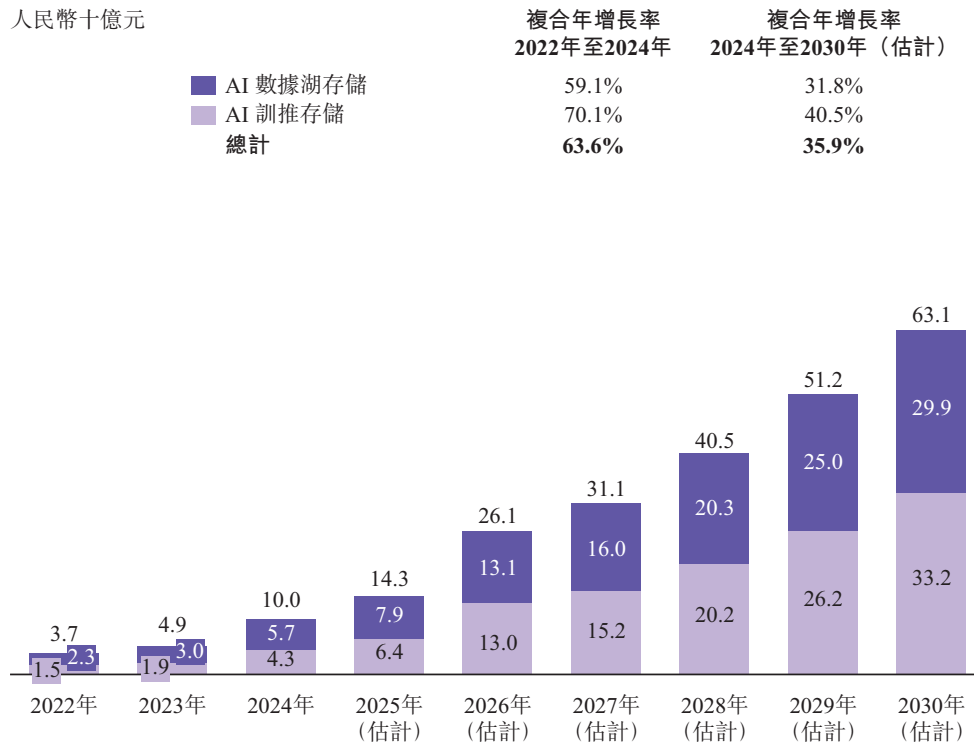


資料來源：IDC、灼識諮詢

同時，根據同一資料來源，中國分布式AI存儲市場按收入計在2024年達到人民幣100億元，並預計在2030年增長至人民幣631億元。

行業概覽

中國分布式AI存儲市場規模，按產品劃分，2022年至2030年（估計）



資料來源：IDC、灼識諮詢

分布式AI存儲的交付模式

分布式AI存儲的交付模式分為兩類：以軟件為核心的一站式交付與軟件交付。

以軟件為核心的一站式交付通常交付軟硬一體機，依托於軟件的調度算法，配套提供經過預配置的硬件和網絡，直接交付給客戶。在分布式AI存儲系統的質量控制和兼容性驗證過程中，軟硬件一體的交付模式憑藉其軟硬件的深度耦合及系統級優化能力，擁有更強的系統兼容性和穩定性，同時也減低了運維售後風險。

軟件交付只提供存儲系統的軟件層（文件系統、對象存儲、分布式調度等），客戶根據存儲系統要求採購硬件設備、並完成軟硬件集成。IT集成能力和硬件維護能力較強、或對特定硬件供應商有較強需求的客戶，通常會選擇這種模式。在中國市場，許多大型互聯網企業、科研機構以及國有企業出於合規性和國產化的考慮，通常需要單獨採購特定的服務器、網絡和算力等硬件資源，其後在其現有基礎設施上部署分布式AI存儲軟件。

行業概覽

市場驅動因素

數據體量的爆炸式增長。隨著AI模型參數規模和訓練數據集的迅速擴大，全球數據總量呈指數級上升，尤其是圖像、視頻、語音等非結構化數據的佔比顯著提升。這一趨勢對存儲系統的容量、擴展性及成本控制提出更高要求，推動企業持續加大在高密度與彈性存儲方面的投入。

AI場景對存儲容量及性能要求逐漸提高。在大模型訓練、推理及生成類應用中，存儲系統不僅要具備海量數據承載能力，更需支撐高併發訪問與低延遲傳輸。分布式AI存儲通過多節點並行與高速互聯，實現數據的高效流動與性能線性擴展，成為滿足AI場景計算密集型需求的關鍵基礎設施。

企業用戶對跨平台兼容與靈活架構的需求不斷增加。隨著多雲與混合雲環境普及，企業希望在不同計算平台之間靈活調用與遷移數據，分布式存儲憑藉開放式架構與模塊化設計，可實現資源的彈性擴展與異構系統的無縫協同。

合規與數據主權推動本地部署存儲的佔比增加。在數據安全、隱私保護及行業監管要求日益嚴格的背景下，越來越多企業選擇自建可控的分布式存儲系統，以確保核心數據的安全可控與本地合規存儲，從而進一步推動分布式AI存儲市場的滲透與發展。

行業概覽

中國分布式AI存儲市場競爭格局

中國分布式AI存儲市場解決方案供應商大致上可分為獨立廠商和非獨立廠商。獨立廠商是專注於提供存儲系統解決方案的廠商，不依附於其既有的硬件組件、服務器或其他業務體系，因此具備很強的生態合作開放性與上游硬件供應鏈的靈活性。獨立廠商可以避免單一硬件選擇帶來的供應鏈風險，能為客戶選擇最適配其需求與場景的組合。下表展示了分布式AI存儲獨立與非獨立廠商的對比。

	獨立廠商	非獨立廠商
供應鏈	多供給，可替換	單一供給，策略受限
協議與生態	支持多種協議，為不同的應用場景和生態提供高效通路和兼容	廠商自有協議優先，專注自身生態搭建
性能	以存儲系統自身的帶寬密度和容量效率為目標，追隨主流的硬件迭代路線	以超出存儲系統自身的端到端、一體化業務為目標，鎖定自有路線
交付	可分離採購，靈活度較高	不可分離採購，靈活度低

鑒於部分供應商（特別是獨立供應商）的一大部分收入來自軟件解決方案，而其他供應商的收入主要來自一站式解決方案，因此市場參與者的收入不能直接比較。然而，各供應商軟件配置的裝機量則有統一基準，可供比較。根據灼識諮詢的資料，按2024年裝機量計，中國五大分佈式AI存儲解決方案供應商合計持有市場份額的52.3%。按2024年裝機量計，我們是中國第二大分佈式AI存儲解決方案供應商，也是最大獨立解決方案供應商，市場份額為10.4%。

公司	分布式AI存儲裝機量， PB，2024年	市佔率，2024年	類型
公司A	1,519	14.7%	非獨立
本公司	1,067	10.4%	獨立
公司B	956	9.3%	非獨立
公司C	945	9.2%	非獨立
公司D	907	8.8%	非獨立
總計	5,395	52.3%	

行業概覽

資料來源：上市公司年報、行業專家訪談、灼識諮詢

附註：

1. 公司A是一家信息通信技術解決方案提供商，總部位於中國廣東省，成立於1987年。該公司主要提供信息通信技術基礎設施、雲端運算、數字能源、消費性電子和智能汽車解決方案產品及服務。該公司並非上市公司。
2. 公司B是IT基礎設施解決方案提供商，總部位於中國山東省，成立於1998年。該公司主要提供IT基礎設施產品、解決方案及相關服務，包括雲端運算、大數據和人工智能解決方案。該公司是上海證券交易所上市公司。
3. 公司C是一家信息通信技術解決方案提供商，總部位於中國浙江省，成立於2003年。該公司主要提供涵蓋雲端運算、大數據、人工智能、工業互聯網、信息安全、智慧互聯、邊緣運算和統一運維的一站式數字化解決方案，以及端到端的技術服務。該公司並非上市公司。
4. 公司D是一家高性能計算服務器製造商，總部位於中國北京，成立於1996年。該公司主要提供高端計算、儲存、安全和數據中心產品。該公司是上海證券交易所上市公司。

行業概覽

同時，本公司在分布式AI存儲平台的多個關鍵維度上展現出行業領先的能力。在同業當中，我們提供最全面的協議兼容性，是首個支持單個系統異構芯片的參與者，並推出首個在單一集群內支持塊、對象及文件的分佈式存儲平台。

行業概覽

競爭壁壘

軟件核心能力。分布式AI存儲的核心競爭在於底層軟件算法的成熟度，包括調度算法、元數據管理、數據布局與資源優化等，這些能力決定了系統在複雜AI任務下的數據流效率與資源利用率。

性能優化能力。在AI訓練等高併發、大吞吐場景下，系統需保持穩定的讀寫性能與低延遲響應。廠商需具備深厚的系統優化能力，以保障算力集群的持續高效運行。

交付與服務能力。AI存儲系統往往涉及數百個節點，需要跨機架、多集群甚至多數據中心部署、集成、測試驗證，同時不同企業對數據主權和合規的要求以及部署環境也存在較大差異。分布式AI存儲服務商必須具備端到端交付能力、可以實現方案的快速部署及一站式交付方案，從而提高客戶滿意度。

生態兼容性。分布式AI存儲需與主流AI框架及多種計算平台深度集成，從而減少數據傳輸與等待的時間，簡化開發者的對接流程，增強存儲架構的拓展性，形成廣泛的生態適配性，以提升客戶使用的便利性與擴展性。

客戶黏性與長期支持。行業客戶對系統穩定性與長期運維依賴度高，廠商需具備規模化運維能力與持續技術支持能力，通過軟件更新、服務延展與生態綁定形成高客戶黏性。

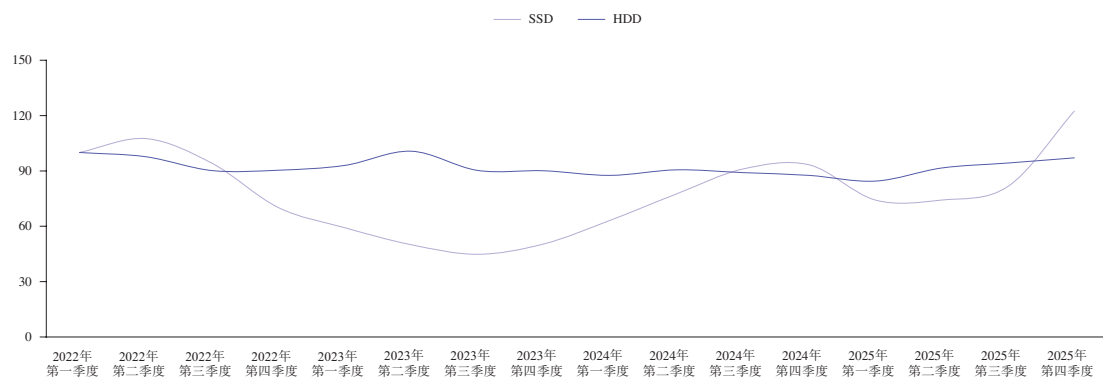
行業概覽

價格與成本趨勢分析

分布式AI存儲系統的成本結構主要包括儲存媒介（SSD及HDD）、計算服務器、網絡設備和系統軟件。其中，SSD及HDD是對系統性能最關鍵的組件，也是存儲硬件的主要部分，此乃由於其直接決定存儲效能、可拓展性及整體資本支出。

在歷史期間，SSD的單價呈現明顯的週期性波動，其特徵是價格快速下跌後出現部分回升，導致其價格波動幅度相對較大。相較之下，HDD的單價走勢則更為穩定，在較窄的範圍內進行平緩且有限的調整，這反映其作為一種成熟且具有成本錨定作用的存儲媒介的地位。特別是，SSD及HDD的價格在2025年下半年均出現反彈，其中SSD的漲幅更為顯著。在供應收緊和需求復甦的推動下，預計該上升趨勢將於2026年持續。下圖展示SSD及HDD的歷史價格走勢。

SSD及HDD平均價格指標，2022年第一季度至2025年第四季度



資料來源：福布斯、灼識諮詢

附註：上圖中的數據乃基於2022年第一季度，為統計及計算目的，2022年第一季度SSD及HDD組件的平均單價已調整為100。

資料來源

灼識諮詢獲委託對中國AI存儲行業及其他相關經濟數據進行研究、提供分析並作出報告，費用為人民幣550,000元。委託報告已由灼識諮詢在不受本公司或其他利益相關方影響的情況下獨立編製。灼識諮詢提供行業諮詢服務、商業盡職審查及戰略諮詢。在顧問團隊積極追蹤TMT、汽車、消費品及服務、農業、化工、營銷及廣告、文化及娛樂、能源及工業、金融及服務、醫療保健及交通運輸等各行業的最新市場趨勢下，灼識諮詢擁有該等行業高度相關及極具洞察力的市場情報。

行業概覽

灼識諮詢利用各種資源採用一手及二手研究方法。一手研究包括與主要行業專家及領先參與者的訪談，而二手研究則涉及分析來自公開可得來源的數據。委託報告中的市場預測基於預測以下主要假設：(i) 全球整體社會、經濟及政治環境預期於預測期間維持穩定趨勢；(ii) 若干關鍵行業驅動因素預計將在預測期內繼續推動市場增長；及(iii) 於預測期內，並無任何極端不可抗力或不可預見的行業法規致使市場可能受到重大或根本性影響。