
業 務

概覽

我們是誰

我們是中國領先的智算服務商，專為支持涵蓋大規模模型訓練及推理的算力密集型AI工作負載而構建。我們構建、經營及變現AI算力。我們為智算中心（「AIDC」）項目提供一站式智算基礎設施解決方案，並提供由我們的專有計算操作系統驅動的按需雲計算服務。根據灼識諮詢的資料，按2025年收入計，我們在中國智算服務商中排名第九，並為前十名中唯一的智算新雲（「智算新雲」），使我們成為中國最大的智算新雲。作為AI原生基礎設施提供商，我們有別於超大規模智算服務商，從底層開始圍繞高密度計算部署、異構資源編排及模型中心型計算需求所需的運營穩定性構建我們的架構。

在生成式AI（涵蓋大規模模型訓練及高併發推理）的推動下，整個計算行業正在經歷一場根本性轉型。通用雲架構過往針對事務型及擴展Web工作負載而設計，該等複合型工作負載催生了一套截然不同的基礎設施需求，而傳統以CPU為核心的雲端環境無法滿足該等需求。因此，根據灼識諮詢的資料，運算優化型AI原生基礎設施對於維持持續的AI進展至關重要。我們透過AI基礎軟件（「AIFS」）服務建立了自身能力，其後於2023年年初戰略性將資源集中於智算基礎設施及雲服務，隨著企業應用開始加速，將我們在數據處理、訓練編排及部署方面的模型工程專業能力延伸至基礎設施層。

我們的戰略性升級使我們站在中國AI基礎設施建設的前沿，並以普惠算力為核心營運理念。我們透過異構調度、serverless編排及按使用量計費，降低智算的准入門檻。我們的專有「一度算力」標準以標準化算力計量單位（「DCU」）計量算力，並使收費與實際算力消耗（而非預留的物理容量）相匹配。根據灼識諮詢的資料，其為中國智算行業首個此類可量化計費單位。我們亦已開始將此理念延伸至模型推理，將我們的計算集群及選定的第三方基礎模型封裝為標準化推理輸出，客戶透過我們的應用程序編程接口（「API」）獲取該等輸出，並按詞元消耗量付費。

如今，我們經營兩大主要業務線，其將AI算力的創造與其變現相連接。透過我們的智算基礎設施解決方案，我們作為一站式解決方案提供商，負責AIDC的架構設計、硬件採購及集成部署，整合我們的專有計算操作系統，為地方政府、國有企業及民營企業提供完全同步、符合生產級標準的計算設施，同時亦確保長期獲取算力。透過我們的智算雲服務，我們經由專有AI原生平台提供池化計算、網絡及存儲資源，主要通過訂閱、按量付費及專用計算集群服務企業客戶及開發者，將我們的算力池變現。

業 務

我們主要透過長期安排獲取算力，而非自行建設及擁有設施。我們的合作夥伴（包括設施擁有人及計算基礎設施基金）為設施建設提供資金，而我們則提供涵蓋硬件採購、架構設計、系統部署及持續運營管理的全套技術及運營能力。此安排所需的前期資本較自行建設為少，並使我們能夠快速擴展區域產能。隨著我們規模的擴大，我們計劃以選擇性自建基礎設施補充該安排，以加強長期資產控制、支持部署我們的專有計算操作系統及提升單位經濟效益。

我們的業務在運營及財務層面均實現快速擴張。為確保運營一致性，我們在計算資源的整個運營週期內對其實施全面管控，涵蓋設施架構規劃、系統部署及由我們專有計算操作系統驅動的日常動態調度。截至最後實際可行日期，在我們的「中央調度＋區域節點」框架下，我們的可控算力在中國七個省份已超過29,000 PFLOPS，根據灼識諮詢的資料，按簽約生命週期管控容量計，躋身中國領先的智算新雲之列。截至最後實際可行日期，我們已透過我們的智算基礎設施解決方案及智算雲服務為超過150名企業客戶提供服務。隨著企業對智能基礎設施的需求加速，以及兩大業務線的商業化程度不斷深化，我們的收入由2023年的人民幣106.2百萬元增長至2024年的人民幣395.5百萬元，再增長至2025年的人民幣1,098.2百萬元，兩年內增長近十倍，並由截至2025年6月30日止六個月的人民幣306.4百萬元增長至截至2026年6月30日止六個月的人民幣588.4百萬元。

中國最大智算新雲
按2025年收入排名⁽¹⁾

人民幣10.98億元
2025年收入

221.52%
2023-2025年收入
年複合增長率⁽²⁾

「一度算力」
中國首個此類可量化
計費單位⁽¹⁾

> 29,000 PFLOPS
可控算力⁽³⁾

國內首個AI操作系統
通過中國信通院四項認證⁽⁴⁾

54.1%
30天MFU

大規模集群
≥ 85.9%
彈性伸縮速度提升

> 99.97%
即時彈性交付率

附註：

- (1) 根據灼識諮詢的資料
- (2) 按我們2025年的總收入除以我們2023年的總收入，將所得商值開平方，再減去一計算
- (3) 截至最後實際可行日期，根據灼識諮詢的資料
- (4) 根據灼識諮詢的資料，涵蓋四個關鍵領域，即算力調度、模型訓練、模型推理及數據處理

業 務

我們的市場機會

AI的快速規模化已為計算基礎設施創造了結構性需求拐點。根據灼識諮詢的資料，2025年中國AI基礎設施行業的市場規模已超過人民幣3,711億元，預計到2030年將突破人民幣16,720億元，2025年至2030年的年複合增長率約為35.1%。在此更廣闊的市場中，中國的智算服務市場規模於2025年達到約人民幣684億元，預計到2030年將超過人民幣2,775億元，同期年複合增長率為32.3%。於2025年中國智算服務市場中，智算基礎設施解決方案貢獻約人民幣148億元，而智算雲服務則貢獻超過人民幣536億元。

此增長反映智算需求的快速擴張持續重新定義行業的基礎設施需求。模型訓練引發了第一波需求，需要持續的高密度計算能力，其規模是通用基礎設施從未被設計用於承載的。隨著模型日趨成熟及部署範圍擴大，推理已超越訓練成為主要的算力消耗來源，將AI處理以大規模、持續的方式嵌入企業及個人工作流程中。這一轉變日益受到智能體AI、具身AI及物理AI融合的推動，在此情況下，自主系統、機器人智能及具備物理感知能力的認知模型可全天候持續執行複雜的多步驟任務。Token工廠其後進一步深化此趨勢，將推理從偶發性需求轉化為持續的工業級生產模式。貫穿該等轉變的共同核心要求是：AI工作負載需要專用基礎設施，提供通用雲架構並非被設計用於提供的計算密度、資源彈性及運營穩定性。

中國的智算服務市場包括兩類主要參與者：超大規模智算服務商及智算新雲。超大規模智算服務商受限於兩項制約因素：通用基礎設施對AI工作負載存在顯著的適配摩擦，以及專有模型開發利益與其所服務的獨立開發者及企業之間產生結構性衝突。智算新雲提供立足於業務中立性及技術專業化的獨特價值主張，結合AI原生架構、完全的商業中立性、異構硬件兼容性及使基礎設施成本與實際計算利用率相匹配的按用量計費商業模式。智算新雲的結構性優勢體現在市場份額軌跡上。根據灼識諮詢的資料，智算新雲在中國智算服務市場的份額預計將從2025年的15.7%增長至2030年的27.3%，增速超越整體市場增長，體現AI原生中立基礎設施模式的商業驗證持續深化。

我們的技術實力

我們的服務基於AI原生架構構建，該架構從底層開始圍繞模型中心型工作負載的特定需求而設計，其中計算、網絡、存儲及系統韌性被設計為完全同步運作。此架構協調一致性正是AI原生基礎設施有別於為AI需求而改造的通用系統的關鍵所在。

業 務

為實現此專有性能，我們已優化基礎設施的每一層：

- **計算層。**我們對我們的算力池進行編排，其截至最後實際可行日期由超過29,000 PFLOPS的可控算力組成。我們基於拓撲結構的智能調度能感知硬件之間的物理距離，從而將任務分配至最高效的計算集群，最大限度地提高通信速度，同時最大限度地減少跨節點延遲。通過將加速器級增強與專用通信庫及軟硬件協同加速相結合，我們顯著降低了交互延遲，並在持續生產條件下提升了整體計算吞吐量。
- **網絡層。**我們部署了一個經過優化的「軌道優化」互連拓撲結構，從而實現了節點間的高帶寬、低時延通信。此架構通過為同級硬件建立直接連接，消除了交換機之間的流量衝突。此外，我們採用網絡內計算，即數據在傳輸過程中由交換機進行處理，這可將集合通訊任務的性能提升一倍，並實現超過95%的網絡帶寬利用率。此網絡設計可防止通信開銷制約計算吞吐量，這對大規模分佈式訓練而言是一項關鍵要求。
- **存儲層。**我們開發了一套針對AI數據模式進行優化的高性能混合存儲系統。該系統支持對大模型的模型進度（「檢查點」）進行快速存儲，並通過區分不同文件類型，從而最大限度地提高數據吞吐量。該系統旨在管理和訪問艾字節級數據，使計算資源能夠在訓練及推理管道中以其設計利用率上限運行。
- **系統層。**我們亦在系統層嵌入系統級韌性機制，包括自動硬件故障檢測、快速節點隔離以及重啟功能。為緩解傳統基礎設施中常見的長時間資源配置瓶頸，我們的專有自動化集群部署方法以虛擬化磁盤及網絡介面架構取代傳統網絡啟動流程。在千節點集群的實際企業部署中，該方法將整體擴展時間縮短逾85%，並將節點配置速度提升逾7倍。

除物理編排外，我們正透過AI工廠（一項持續開發中的技術架構）將我們的架構能力延伸至模型執行及輸出層。AI工廠包含支持大規模模型預訓練及強化學習的訓練工廠，以及旨在透過結合自動化模型部署、動態存儲器分配及分佈式快取以推動高吞吐量模型推理，從而最大限度提高詞元生成效率的Token工廠。為優化持續推理工作負載下的集群性能，我們正在開發Alaya Atom Pod參考架構，以在模型執行的預填充和解碼階段分離計算資源。Alaya Atom Pod採用分佈式鍵值緩存層，旨在儲存、傳輸和復用中間推理狀態，預期可最大限度地減少長上下文和多輪智能體工作流程中的冗餘矩陣計算，並降低每詞元的計算開銷。

業 務

補足我們的生產引擎，我們的ModelOps工具鏈涵蓋AI開發的全生命週期，從數據處理、訓練編排到部署及迭代優化，均在統一的運行環境中完成。在此基礎上，我們的AgentiCTRL強化學習環境通過將底層基礎設施複雜性抽象化為可編程接口，簡化了後訓練及對齊工作流程。此架構縮短策略調校週期，並使應用團隊能夠將算法目標優化與集群管理分離，從而加快自主工作流程的部署。

我們的架構的效率及穩定性反映於性能基準測試中。我們一體化架構的商業效益已獲高標準的營運基準測試所驗證。根據灼識諮詢的資料，於滿載條件下持續生產運行期間，我們的基礎設施維持了54.1%的模型FLOP利用率（「MFU」），日間波動率僅為1.60%，完全處於2%的嚴格行業穩定性閾值以內，並超越30%至50%的行業平均MFU範圍。該等卓越指標證明我們的平台有能力在大型分佈式任務期間，透過盡量減少通信開銷及減少由硬件引致的訓練中斷，以維持高效算力利用率及一致的營運穩定性。

為持續針對真實技術邊界完善我們的服務，我們的研究團隊透過在我們的基礎設施上調校原型架構，積極追蹤生成式人工智能的前沿趨勢。DART-GUI-7B模型完全在我們的AgentiCTRL基礎設施上訓練而成，並於2026年在30次迭代設置下的全球OS World 7B基準測試賽道中榮獲第一名，充分證明了我們所提供系統環境的編排效率。

我們的服務

我們的技術通過兩大主要服務線進行商業化。我們將底層的調度、編排和基礎設施能力轉化為可供客戶使用的產品。

智算基礎設施解決方案

透過我們的智算基礎設施解決方案，我們作為一站式解決方案提供商，負責AIDC的架構設計、硬件採購及集成部署，服務於需要專為AI工作負載而建構的生產級計算設施的地方政府、國有企業及民營企業。我們的解決方案涵蓋AIDC開發的全生命週期，從早期需求評估及架構設計，經硬件採購及系統集成，到各個AIDC的「點亮」及持續容量優化。我們並非交付各自獨立的組件，而是將分散的硬件組件轉化為整合型的、AI即用的，並就各客戶的特定工作負載特徵作出校準的設施。

我們基礎設施解決方案的技術基礎為Alaya NeW OS，此乃我們專有的計算操作系統，其於所有AIDC部署中作為平台控制層發揮作用。Alaya NeW OS提供全棧式優化，在單一框架內抽象化異構硬件並同步計算、網絡及存儲資源，使開發人員能夠像使用傳統的桌面調度程序一樣，以同樣直觀便捷的方式管理高性能計算資源。根據灼識諮詢的資

業 務

料，Alaya NeW OS為中國首個在四個關鍵領域(即算力調度、模型訓練、模型推理及數據處理)獲得中國信息通信研究院(「中國信通院」)認證的AI操作系統。該認證反映其符合系統設計、運營可靠性及技術穩健性方面的國家級標準，並為我們的政府、國有企業及受監管行業客戶提供經獨立驗證的基礎設施。

智算雲服務

透過我們的智算雲服務，我們經由Alaya NeW Cloud(我們的專有AI原生開發及生產平台)提供標準化、池化計算、網絡及存儲資源。Alaya NeW Cloud作為serverless交付接口，提供了一個統一的环境，能夠自動分配資源，並消除傳統雲服務中的操作碎片化現象。

我們的雲產品組合分為兩個主要服務層級，以滿足不同規模的需求：

- **虛擬容器計算服務(「VKS」)**旨在為運行需要靈活、彈性計算資源的推理及微調工作負載的客戶而設計。利用serverless架構，VKS根據實時需求提供自動擴展及資源分配，讓開發團隊無需預留冗餘的物理容量，亦可承接突發密集型或週期性工作負載。VKS服務根據我們的「一度算力」標準以DCU計量，該標準使成本與實際算力消耗相匹配。
- **專屬容器計算服務(「DKS」)**旨在為進行需要持續性能隔離及營運穩定性的規模、超大規模模型訓練的企業客戶而設計。DKS提供對專用計算集群的專屬存取，提供最密集的分佈式訓練工作負載所要求的一致性及可靠性。客戶按協定的訂閱期訂閱DKS，為我們在訂閱期內帶來經常性收入及可預見性。

我們亦推出Alaya Token，此乃我們的託管式模型推理服務，讓客戶透過統一API接口訪問選定的第三方大語言模型，主要按詞元消耗量計費。Alaya Token由我們的Token工廠技術架構支持，並於2026年通過中國信通院開展的高質量Token雲服務能力評估。截至最後實際可行日期，我們已就Alaya Token服務訂立約十份合約，而於往績記錄期間並無就該安排確認任何收入。

我們的財務表現

於往績記錄期間，我們錄得穩健且持續的增長。截至2023年、2024年及2025年12月31日止年度以及截至2025年及2026年6月30日止六個月，我們的收入分別為人民幣106.2百萬元、人民幣395.5百萬元、人民幣1,098.2百萬元、人民幣306.4百萬元及人民幣588.4百萬元；毛利分別為人民幣43.2百萬元、人民幣46.3百萬元、人民幣191.4百萬元、人民幣56.3百萬元及人民幣182.4百萬元。截至2023年、2024年及2025年12月31日止年度以及截至2025年及2026年6月30日止期間，我們的經調整年/期內淨虧損(非國際財務報告準則計量)分別為人民幣194.3百萬元、人民幣206.2百萬元、人民幣91.3百萬元、人民幣59.9百萬元及人民幣14.9百萬元。

業 務

我們的優勢

立足開放生態的領先智算新雲

我們是國內首批專注於人工智能算力需求的智算服務商之一。根據灼識諮詢的資料，以2025年收入計，我們在國內全部智算服務商中位列第九，並為中國前十大智算服務商中唯一的智算新雲。我們不商業化專有基礎模型，亦不構建AI應用程序。我們對智算雲服務的堅定信念，使我們成為中立的首選基礎設施合作夥伴，因為我們不會在任何下游市場與任何客戶競爭，同時獨立模型開發者、企業、政府部門及國有企業可在我們的基礎設施上構建而無需作出戰略性妥協。根據灼識諮詢的資料，智算新雲在中國智算服務市場的份額預計將從2025年的15.7%增長至2030年的27.3%，反映市場對中立、專用AI基礎設施的偏好持續深化。

我們的市場領先地位體現在客戶群中。我們的智算基礎設施解決方案及智算雲服務的企業客戶數量於2023年、2024年、2025年以及截至2025年及2026年6月30日止六個月分別為三名、七名、82名、45名及87名。Forrester將我們評為全球AI基礎設施雲平台領域的「表現強勁者」，並將我們列為亞太地區普惠算力利用效率的領先者之一。根據IDC發佈的《2025 IDC MarketScape：中國算力中心運營及服務廠商評估報告》，我們被歸入「主要廠商」梯隊。我們亦入選華興資本旗下Rongo Group發佈的2024至2025年度「中國獨角獸TOP50」及「AI產業創新領軍企業TOP50」榜單。我們已獲得多項國家級榮譽。具體而言，我們於2025年獲中華人民共和國工業和信息化部（「工信部」）認定為國家級重點「小巨人」企業，此為表彰專業領域內新穎及表現優秀中小企業的國家級「小巨人」計劃中的一個精英類別。該等榮譽體現了我們的技術實力及持續發展的往績記錄。身處產業生態前沿，我們亦參與了國家及行業標準的制定工作。作為核心起草參與者，我們參與制定了中國首個大語言模型行業標準《大規模預訓練模型技術和應用評估方法》，並參與制定了《人工智能研發運營一體化》、《大模型推理平台技術要求》、《大模型訓練平台技術要求》以及《大模型數據處理平台技術要求》的技術要求。該等標準構成了中國大語言模型生態系統的基礎框架。

AI原生架構與持續的研發能力

我們已構建專有的AI原生架構，專門為大規模模型訓練、推理及強化學習工作負載的嚴苛需求而設計，並以持續及深化的研發投入為支撐。截至2026年6月30日，我們擁有超過150名研發人員，佔員工總數逾50%。截至最後實際可行日期，我們已獲得125項已註冊專利，共同構成了前瞻性的技術組合。

業 務

我們研發能力的深度體現在獨立技術驗證中。Alaya NeW OS是我們的專有計算操作系統，為中國首個獲得中國信通院在四個關鍵領域(算力調度、模型訓練、模型推理及數據處理)認證的AI操作系統。於2026年，我們的Alaya NeW Cloud成功通過中國信通院開展的智算雲平台能力評估。我們的技術架構已在標準工程條件下經過嚴格測試。在504個最新一代智算加速器以63個節點規模運行的為期30天的持續生產運行中，我們的基礎設施實現了54.1%的MFU，相當於每個加速器約535 TFLOPS，在Qwen2.5-72B-Instruct預訓練工作負載上，日間MFU波動為1.60%。根據灼識諮詢的資料，該持續效率在可比參數規模下顯著高於行業平均水平，為我們先進的工程設計提供了清晰的明證。

我們的專業研發團隊積極探索最先進的模型工程方法，以確保我們的平台針對下一代架構轉變持續保持優化。作為衡量我們基礎設施支持高度複雜強化學習工作負載能力的基準，於2026年，在最多30次迭代的設定下，於我們專有平台上訓練的DART-GUI-7B模型在OS World 7B賽道中取得第一名。此項評估凸顯了我們實現快速算法迭代的能力，並鞏固了我們作為純基礎設施及平台軟件供應商的核心市場地位。我們的應用研究成果已多次在ICLR、NeurIPS、ACL及AAAI等領先國際AI會議上發表，體現了我們在應用硬件工程及基礎軟件優化兩個層面的參與。

我們的研發投資專注於促進我們平台內工程經驗的系統性積累。我們並非追求孤立的開發項目，而是藉助我們的全生命週期工程框架，確保從反覆模型部署中獲得的技術洞察可直接整合至我們的計算操作系統及平台軟件的持續完善中。此持續優化過程使我們能夠隨著每次後續部署，逐步提升軟硬件協同效應、提高我們平台的算力利用效率，並增強我們核心架構的長期技術能力。

創新重新定義智算耗用

智算雲服務傳統上按時間定價，透過按小時、按月或按年租賃指定硬件規格。按時間計費適合持續、可預測的工作負載，例如大規模預訓練。對於波動的工作負載，包括推理、微調及突發密集型訓練，按時間計費會導致客戶為閒置容量付費，或在需求激增時等待容量。根據灼識諮詢的資料，中國智算雲服務市場的計費正從供應端的按時間定價轉向按使用量計量。我們憑藉兩項創新處於此轉變的前沿：提供彈性資源編排的AI原生serverless架構，以及使商業邏輯與實際消耗相匹配的專有算力計量標準。

我們的serverless架構將資源調度及產能擴張邏輯直接嵌入系統中。算力資源根據任務需求自動分配，並在數秒內完成擴縮容。透過我們的專有兼容層，工作負載可在單一計算環境中跨異構加速器架構進行編排，抽象化底層硬件複雜性，並在不同資源配置下提供穩定、自適應的算力解決方案。此架構為具有週期性或突發密集型需求的客戶消除了冗餘的產能預留，並提升我們集群的利用效率。

業 務

根據灼識諮詢的資料，我們的「一度算力」標準為中國智算行業首個此類可量化計費單位。對標能源計量原則，該標準將算力能力抽象化為標準化DCU，並受我們的專有專利保護，其中包括於2025年獲授的普惠算力技術及算力資源包框架。由能夠實時追蹤異構資源消耗的專有軟件插件支持，此框架將商業邏輯從基於時間的服務器佔用轉變為實際算力消耗，使客戶能夠根據可量化的算力輸出付費，而非為預留的物理容量付費。此舉降低中型企業、小型企業及個人開發者的准入門檻，並擴闊我們的客戶基礎，使其不再局限於大型機構客戶。

Alaya Token為我們的託管式推理服務，處於商業化早期階段，將此按消耗量計費的模式從算力單位延伸至詞元消耗。透過將底層基礎設施部署抽象為標準API接口，Alaya Token提供按實際輸入和輸出詞元消耗量計費的託管式模型推理服務。依託我們的Token工廠架構，該服務將我們的普惠算力理念從原始計算單元擴展至顆粒度的詞元消耗，使客戶能夠利用大語言模型，同時無需承擔獨立集群管理帶來的資本或營運開銷。

該等創新共同降低高性能AI基礎設施的准入門檻，提升集群利用率，並使我們處於行業向按消耗量計費計算轉型的前沿。

長期管控容量及全生命週期交付能力

我們透過長期安排獲取算力，該等安排使我們在每個設施的整個運營週期內對其擁有運營控制權。我們的智算基礎設施解決方案項目為我們獲取該算力的主要渠道，並輔以長期租賃第三方AIDC及我們自有的服務器。於此安排下，行業合作夥伴出資建設主要實體設施，而我們則直接負責架構設計、系統集成及持續的營運管理。此模式所需的前期資本較自行建設為少，並使我們能夠快速擴展至不同區域。為配合此速度並確保對基礎設施的長期絕對控制，我們正在尋求選擇性自建策略。該均衡的資產組合確保我們不斷擴張的產能在結構上保持穩定，並免受供應端波動影響，使我們能夠在維持跨區域快速擴張的同時，獲取自有基礎設施的成本效益。

我們對底層計算資源的長期控制，乃我們技術及商業創新的不可或缺基礎。Alaya NeW OS、我們的軟件棧及我們的「一度算力」計費模式，依賴穩定且深度整合的基礎設施，以優化異構調度。彼等無法在零散、短期的第三方租賃基礎上有效運行。這種獨特的軟硬件協同效應，使我們有別於缺乏我們的一體化技術棧及資產結構化能力的傳統基礎設施轉售商。截至最後實際可行日期，在我們的中央調度及區域節點架構下可控算力已超過29,000 PFLOPS，根據灼識諮詢的資料，按生命週期可控算力規模計，我們躋身中國領先的智算新雲之列。

業 務

我們提供涵蓋智能計算中心建設全生命週期的綜合服務。我們的流程涵蓋早期階段的需求評估及架構設計、中期階段的系統集成及部署，以及持續的運營優化。在集成階段，我們會進行嚴格的兼容性驗證及壓力測試，以將多供應商硬件整合到我們統一的編排框架中。在運營階段，我們持續優化資源配置及工作負載分配，以適應不斷演進的模型規模及應用需求，從而在我們整個跨區域雲網絡實現自動化及高效率的資源部署。

具適應性的基礎設施部署戰略與標準化的生命週期交付流程相結合，使我們能夠在擴大業務規模的同時，保持運營規範並控制風險。這種多元化的方式為區域擴張和持續產能增長建立了可複製的路徑。

遠見卓識的領導力推動持續創新

我們的管理團隊兼具深厚學術素養與全球行業經驗，為我們引領智算服務行業發展築牢技術根基、賦予前瞻戰略視野。董事會主席方磊博士曾任職於微軟總部，擔任項目負責人，參與微軟Azure孵化及數據中心管理系統研發工作。此後，其負責必應搜索引擎的大數據索引與機器學習架構設計，助力我們形成計算基礎設施與人工智能融合領域的核心技術優勢。董事會副主席尚明棟先生擁有微軟總部多年軟件工程從業經驗，主攻大型系統的穩定性與可擴展性方向。其豐富的工程管理經驗，能夠推動前沿研究成果轉化為具備商業價值的AI原生產品，保障產品高效交付與穩定運行。

在此背景下，我們積極把握智算雲服務快速擴張的機遇。我們為業內首批將普惠算力及強化學習雲服務概念作為切實可行的基礎設施戰略而非理論構想予以闡述並付諸實施的企業之一。

在這一戰略方向的指引下，我們已組建一支既具備工程實施能力，又擁有先進研究實力的研發團隊。我們的研究成果已於ICLR、NeurIPS、ACL及AAAI等國際頂尖AI會議上發表，體現了我們在應用工程與基礎AI研究領域均有所涉獵。

為持續推進計算平台及相關技術的迭代與商業化，我們一直在擴充研發團隊。截至2026年6月30日，我們僱用超過150名研發人員，佔員工總數的比例超過50%。

業 務

我們的戰略

我們致力於創造智能、探索未知，憧憬一個透過智能轉型賦能全球企業的未來。基於此使命，我們提供由專有計算操作系統驅動的一站式智算基礎設施解決方案及AI原生雲服務。透過消除計算的技術及經濟障礙，我們使各行各業的機構能夠部署先進模型、優化運營效率，並在AI時代實現持續的商業價值。

通過持續的研發和架構創新，執行我們的AI優先技術路線圖

我們擬通過在整個產品服務和研發框架中貫徹AI優先原則來鞏固我們的市場領導地位。我們計劃系統性地增加研發投入，重點關注計算操作系統、工具鏈以及智能體AI能力的演進。

我們的主要工作重點將為優化我們專有的異構計算資源的統一編排引擎。通過增強彈性資源管理和分佈式訓練效率，我們旨在滿足全球對超大規模(10,000-PFLOPS級)模型訓練和高併發推理日益增長的需求。

與此同時，我們將推進Token工廠戰略，順應模型規模化應用的結構性行業趨勢。我們擬將Token工廠打造為核心引擎，全面提升端到端推理吞吐量。這套架構搭載自動化調度與內存管理模塊，可在個人、企業及科研各類應用場景下實現大批量詞元生成、壓降單位成本，助力客戶規模化部署高性價比AI解決方案。

我們計劃透過推進強化學習、多智能體協作及節能計算方面的先進舉措，加強我們的核心技術能力，以提升我們底層基礎設施的性能。我們將繼續吸引系統軟件和雲架構領域的世界級人才，同時深化與全球研究機構的合作，以彌補基礎AI研究與工業級工程之間的差距。

擴展我們的服務矩陣及開源生態系統

為滿足更廣泛的企業需求，我們將繼續完善分級服務。我們計劃優化計算套餐產品和彈性容器集群，為通用AI工作負載提供一個具有成本效益且高效能的環境。對於有嚴格安全和合規要求的企業客戶，我們提供隔離的資源池和網絡環境，以確保工業級別的穩定性和安全性。

此外，我們致力於構建和優化我們的開源生態系統。通過投資生態系統運營，我們旨在提升AI工具鏈的可複用性，降低開發人員的准入門檻，並最終實現源自我們核心平台架構的長期商業價值及複合軟件權益最大化。

業 務

擴展計算能力，並強化全棧基礎設施能力

我們計劃系統性地擴大我們的生命週期管控算力容量。我們將繼續透過與設施擁有人的長期安排獲取算力，深化與電信營運商、基礎設施投資者及政府部門的戰略聯盟，以共同營運下一代計算中心，並以選擇性自建基礎設施補充該等安排。

作為區域計算互聯計劃的積極參與者，我們的平台採用以中央調度+區域節點為特徵的分佈式架構。該架構基於我們專有的Alaya NeW OS內核，支持在不同地區及硬件集群之間實現任務遷移。我們旨在利用我們於算力中心規劃、技術建設及軟件定義編排方面的端到端專業知識，為受監管行業提供定製的基礎設施解決方案。我們利用Alaya NeW OS作為詞元自動化生產的基礎操作系統，旨在促進計算的標準化及廣泛流動性。

深化生態系統協同效應，拓展客戶覆蓋範圍和商業疆界

我們計劃與國內外硬件製造商、獨立軟件供應商以及研究機構建立一個強大的合作網絡。這種以生態系統為導向的方法使我們能夠提供全棧式、本地化的解決方案，將底層硬件與高級應用軟件有機融合。

我們正通過與合作夥伴共同推進旗艦項目，系統性地提升市場拓展效率，以此展示我們解決方案帶來的切實投資回報；同時，我們利用合作夥伴成熟的銷售網絡，開拓新的垂直領域，並在現有客戶群中加速實施「立足並拓展」戰略。透過將我們標準化的智算基礎設施及專業的AI工具鏈與我們積累的特定行業專門知識相結合，我們建立了高度可複製的交付模式，使本集團能夠在平行或相鄰行業中迅速複製我們的成功經驗，並把握不斷擴大的市場機遇。

加快全球擴張步伐，把握AI基礎設施帶來的機遇

我們致力於系統性的國際擴展，將我們的Alaya NeW OS定位為一個中立性的跨平台層，以滿足全球對主權AI的需求。我們的技術專為異構環境量身打造，既支持跨供應商的資源調度，又能嚴格遵守當地的數據本地化和安全法規。

依託我們在東南亞主要增長市場的既有佈局（體現於我們與新加坡研究機構的技術合作，以及我們在印尼與一家領先電訊營運商附屬公司積極部署並已開始產生初步收益的項目），我們擬進一步投資於海外本地化算力樞紐的開發及營運優化。作為尋求構建自主算力架構的地區政府及企業值得信賴的合作夥伴，我們旨在使收益來源多元化，並把握本地化的結構性增長。為拓寬我們的長期戰略視野，我們計劃逐步將我們的本地化合作夥伴模式複製至其他戰略區域，瞄準亞洲、中東及歐洲（包括日本、沙特阿拉伯、阿拉伯聯合酋長國及西班牙）的高增長擴張機遇。

業 務

我們的業務模式

概覽

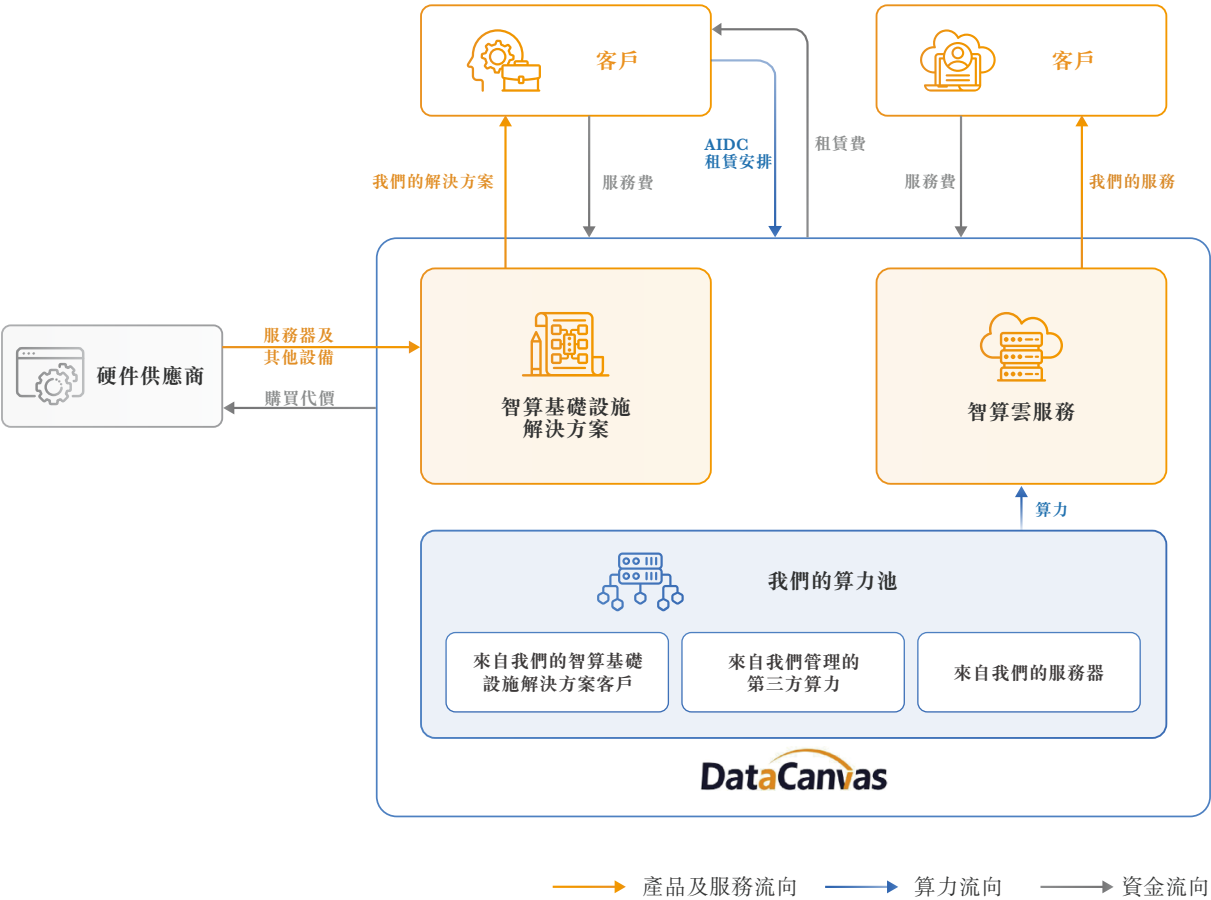
我們採用雙引擎業務模式，以智算基礎設施解決方案和智算雲服務為核心，構建智算供給和消耗閉環，以便我們能夠快速擴展物理智算基礎設施，同時降低AI開發的准入門檻。

- **智算基礎設施解決方案。**我們通過快速部署大規模AI基礎設施，解決算力短缺的根本問題。我們提供一站式解決方案，負責AIDC的架構設計、硬件採購及集成部署，結合我們的工程專業知識及Alaya NeW OS的能力。
- **智算雲服務。**我們通過我們的Alaya NeW Cloud平台提升算力的可及性和效率。通過將物理資源虛擬化為細粒度的按需服務，我們最大限度地提高了基礎設施的利用率，並促進了計算的廣泛應用。

我們的智算基礎設施解決方案及智算雲服務作為我們業務模式中兩個相互協同的組成部分運作。透過我們的智算基礎設施解決方案，我們將向第三方硬件供應商採購的硬件設備與我們的工程及軟件能力相結合，以向客戶交付AIDC。於交付AIDC後，在商業上合理的情況下，我們可根據另行訂立的協議向客戶租賃AIDC的指定部分以產生算力。同時，我們根據合作安排管理第三方AIDC所產生的算力以獲取算力，並據此對租賃設施進行技術升級及標準化。我們亦從自有的服務器獲取算力。我們將該等計算資源整合至我們的算力池。憑藉該算力池，我們的智算雲服務向客戶提供算力資源。此模式將計算基礎設施建設與其後的算力運營及商業化相連接，在我們兩個業務分部之間產生協同效應，並支持算力供應、聚合及消費的閉環。為支持我們算力池的持續擴張，同時維持資本效率，我們亦透過項目層面的安排，與基礎設施投資者(包括計算基礎設施基金)合作。我們擬繼續與合適的合作夥伴探索具資本效益的合作模式，以支持我們計算能力的規模化擴展。有關更多詳情，請參閱「—智算雲服務—我們的算力池」及「—客戶與供應商的重疊」。

業 務

下圖展示我們雙引擎業務模式的協同作用：



除了我們的核心雙引擎服務外，我們亦提供AIFS服務。AIFS服務包括專為專業機構量身定製的AI軟件解決方案。此類解決方案提供了相關工具，可協助客戶開發、訓練和部署AI模型，從而優化內部工作流程。

下表概述按業務分部分類的我們的服務、業務模式以及相應的定價和收入確認政策：

業務分部	服務類型	業務模式	定價	收入確認
智算基礎設施解決方案	基於項目的解決方案	我們為AIDC的建立提供一站式服務，包括AIDC的架構設計、硬件採購及集成部署。我們向地方政府及其部門、國有企業及民營企業交付解決方案。我們通常通過招投標程序獲得項目。	通常為通過招標確定的固定價格合同，已考慮採購成本、軟件許可、集成複雜性及計算集群的規模。	主要於客戶驗收時點確認。

業 務

業務分部	服務類型	業務模式	定價	收入確認
智算雲服務	虛擬容器計算服務 (「VKS」)	面向AI企業、研究機構及開發人員等客戶的靈活的serverless計算資源，利用我們的DCU標準。我們提供兩種模式的VKS：滿足穩定但具彈性需求的訂閱模式，以及滿足靈活需求的按用量付費模式。	(i) 訂閱 ：根據預付費資源包的訂閱收費。 (ii) 按量付費 ：根據DCU的實際消耗量收費。	對於基於訂閱的安排，在合同期內按比例確認；或對於基於使用量的安排，則根據實際消耗確認。
	專屬容器計算服務 (「DKS」)	為AI企業、研究機構及開發人員等客戶提供對計算集群的獨家訪問，確保穩定性和性能隔離。客戶訂閱我們的DKS服務，以訪問由我們平台管理的專屬硬件資源。	基於時間的訂閱，根據專屬資源的規模和訂閱期收費。	
	Alaya Token	託管式模型推理服務，讓用戶透過API進入選定的第三方大語言模型。客戶無需採購或直接管理底層計算資源或模型部署環境，即可使用模型推理的輸出結果。	主要基於詞元實際輸入與輸出的消耗量，採用按用量計費的定價模式。單價可能因所選模型及服務規格而異。	根據詞元實際消耗量，在提供相關推理服務時予以確認。於往績記錄期間，概無任何收入就Alaya Token服務確認。
AIFS服務	基於項目的解決方案	我們為專業機構提供量身定製的AI軟件解決方案。我們以項目為單位開展工作，提供一個集成工具集的綜合軟件平台。這些解決方案通常以完整產品的形式交付予客戶。	通常是通過商業談判或招標確定的固定價格合同，其中考量了軟件定製的複雜性、開發工時以及具體的功能要求。	主要於客戶驗收完全定製化產品時點確認。

業 務

於往績記錄期間，我們的收入主要來自智算基礎設施解決方案、智算雲服務和AIFS服務。下表載列我們於往績記錄期間按業務分部劃分的收入明細：

	截至12月31日止年度						截至6月30日止六個月			
	2023年		2024年		2025年		2025年		2026年	
	(人民幣千元)		(人民幣千元)		(人民幣千元)		(人民幣千元)		(人民幣千元)	
	%		%		%		%		%	
	(未經審核)									
智算基礎設施解決方案	24,938	23.5	333,963	84.4	1,001,958	91.2	269,449	87.9	434,046	73.8
智算雲服務	-	-	196	0.1	77,614	7.1	29,605	9.7	153,839	26.1
AIFS服務	81,294	76.5	61,346	15.5	18,583	1.7	7,363	2.4	498	0.1
總計	<u>106,232</u>	<u>100.0</u>	<u>395,505</u>	<u>100.0</u>	<u>1,098,155</u>	<u>100.0</u>	<u>306,417</u>	<u>100.0</u>	<u>588,383</u>	<u>100.0</u>

有關我們收入的詳細分析，請參閱「財務資料」。

我們的戰略性業務升級

市場演進及技術理據

我們的戰略性升級應對中國AI產業由以軟件為中心的發展向計算密集型商業化的結構性轉變。基礎模型需要大規模參數擴展及高併發推理，推動計算需求呈指數級增長。根據灼識諮詢的資料，中國AI總算力於2025年達1,590 EFLOPS，並預期於2030年達8,050 EFLOPS，年複合增長率為38.3%。同時，下游消費已進入由詞元經濟驅動的推理時代，於此時代中，Token工廠的高容量處理及基於智能體的強化學習，已使數據中心建設成為一門複雜的工程學科。由於傳統IT基礎設施及企業自建集群面臨較高的技術及資本門檻，市場對專業服務提供商的需求日益增加。

為把握該等機遇，我們主動且有序地從提供一次性交付獨立軟件的AIFS服務轉型，以專注於我們的核心服務。我們於2023年開始交付智算基礎設施解決方案，並於2024年9月推出智算雲服務。此調整應對市場向軟件定義架構設計及按使用量計算的算力消費轉型。

我們於AIFS營運期間建立的編譯軟件、自動化代碼部署及系統編排方面的既有專業知識，提供了營運高性能基礎設施所需的技術底蘊。具體而言，我們在管理分佈式工作流程方面的經驗，使我們能夠設計專有的計算操作系統Alaya NeW OS。為維持營運延續性，我們的傳統AIFS服務的功能模組已整合至我們的AI開發平台Alaya Lab，使傳統客戶能夠在不中斷其工作流程的情況下遷移至我們的雲端環境。

業 務

營運轉變及財務軌跡

於往績記錄期間，我們的業務升級以有序及受控的逐步退出方式進行。截至2023年、2024年及2025年12月31日止年度，我們來自智算基礎設施解決方案及智算雲服務的收入合計分別佔我們總收入的23.5%、84.5%及98.3%，而截至2025年及2026年6月30日止六個月則分別佔我們總收入的97.6%及99.9%，並反映此審慎的資源分配。同時，我們的AIFS服務的收入貢獻由2023年的76.5%下降至2024年的15.5%，及下降至2025年的1.7%，並由截至2025年6月30日止六個月的2.4%下降至截至2026年6月30日止六個月的0.1%。

為支持我們升級的業務模式的可擴充性，我們大幅擴展了我們管理的技術版圖。截至最後實際可行日期，我們在橫跨中國七個省的分佈式網絡中交付了超過31,000 PFLOPS的算力，並持有超過29,000 PFLOPS的可控算力。截至同日，我們透過智算基礎設施解決方案及智算雲服務為超過150名企業客戶提供服務。

此項戰略調整從根本上重組了我們的營運成本結構及利潤表現。在我們業務升級之前，我們的成本結構主要為勞工密集型，主要受分配予定制軟件開發的員工成本及技術服務費驅動。隨著我們擴展智算基礎設施解決方案規模並將智算雲服務商業化，我們的成本構成狀況轉向資本密集型模式，其特徵越來越多地表現為初始硬件採購成本及使用權資產折舊。

我們於往績記錄期間的毛利率狀況反映了此過渡的不同執行及爬坡階段。在我們推出智算基礎設施解決方案及智算服務時，我們的毛利率受壓。該等業務於早期階段的毛利率較低，主要由於硬件的採購及部署成本較高。具體而言，我們的智算雲服務於2024年及2025年在其初期爬坡階段錄得毛損，原因是租賃算力的成本於利用率及收益達到規模前已產生，而於截至2026年6月30日止六個月轉為錄得毛利。我們的智算基礎設施解決方案的毛利率於2025年及截至2026年6月30日止六個月有所增加，主要是因為於該等期間交付的若干項目涉及較大比例的軟件組件，而軟件組件通常具有較高的毛利率。有關我們於往績記錄期間的利潤率、銷售成本及分部業績的詳細討論，請參閱「財務資料－選定綜合損益及其他全面收益表組成部分的說明」。

戰略性升級亦改變了我們的營運風險狀況。有關此等不斷變化的營運變量的全面描述，請參閱「風險因素－我們的智算基礎設施解決方案及智算雲服務的商業化歷史有限」。

業 務

主要營運數據

下表載列截至所示日期及期間的主要營運數據，我們認為該等數據有助於了解我們於往績記錄期間的業務表現、營運規模及增長趨勢。該等指標應與本文件中包含的我們的經營業績及其他財務資料的討論一併閱讀：

	截至12月31日止年度/截至12月31日 ⁽¹⁾			截至 6月30日止 六個月/截至 6月30日 ⁽¹⁾
	2023年	2024年	2025年	2026年
智算基礎設施解決方案				
已交付的AIDC項目數量	3	5	6	3
我們項目覆蓋的省份數量	1	4	5	3
已交付的算力(PFLOPS)	–	1,346	17,690	12,672
智算雲服務				
可控算力(PFLOPS) ⁽²⁾	–	2,284	10,696	26,824
計算資源利用率(%) ⁽³⁾	–	–	56.0	79.6
企業客戶數量	–	2	75	84

附註：

- (1) 除累計可控算力外，表內所有其他數字均指相關期間的數據。
- (2) 可控算力指截至相關期末我們直接控制的總算力，包括我們自有的服務器算力及根據合約安排取得的算力，並由我們負責其架構規劃、系統部署及持續調度。
- (3) 計算資源利用率按已利用DCU除以產生的理論最大DCU總數計算。就DKS而言，集群產生的算力將轉換為DCU，並被視為已充分利用；就VKS而言，已利用DCU乃根據實際使用情況計算。

我們的服務

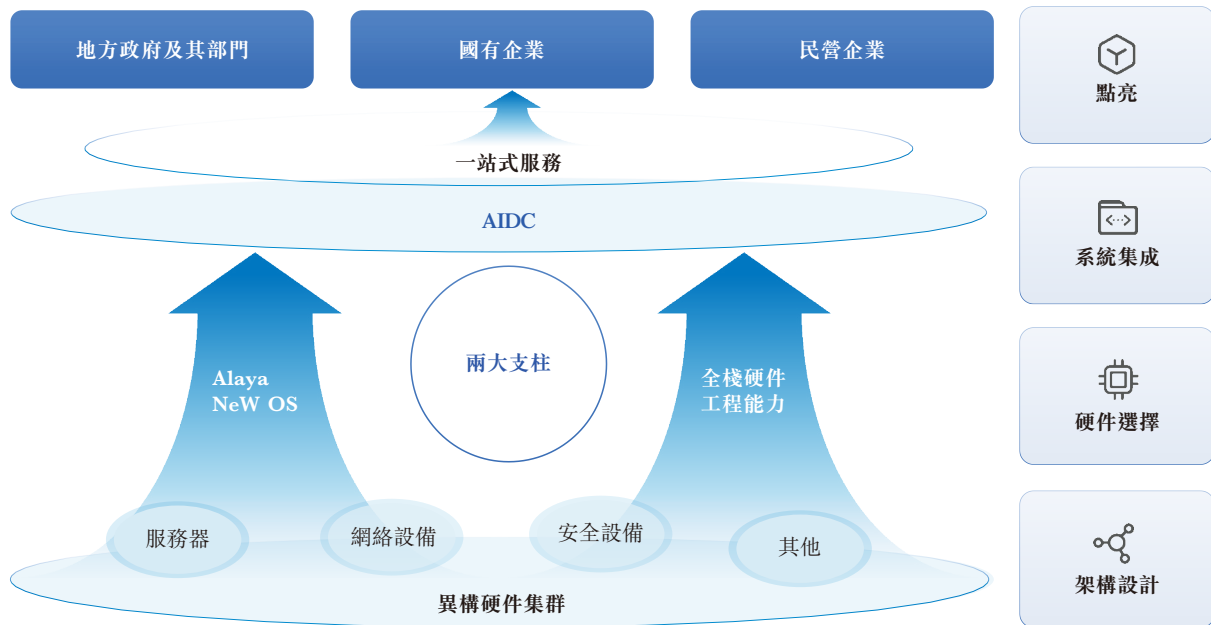
我們的服務包括：(i) 智算基礎設施解決方案，旨在協助地方政府、國有企業及民營企業建設和啟動大型AIDC；(ii) 智算雲服務，旨在為AI企業、研究機構及開發人員等客戶提供靈活、具有成本效益的算力；及(iii) AIFS服務，旨在為銀行及金融企業等專業機構提供量身定製的AI軟件解決方案。

業 務

智算基礎設施解決方案

我們的智算基礎設施解決方案為建立AIDC提供一站式解決方案。我們主要服務於需要專門的技術專長進行系統設計和集成的地方政府及其轄下部門、國有企業及民營企業。我們的服務範圍包括從硬件選擇和架構設計到系統集成和設施最終「點亮」的整個生命週期。

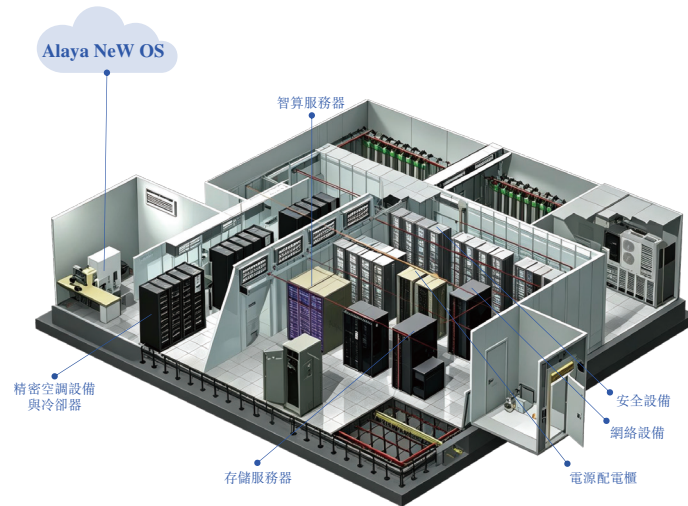
下圖展示我們提供的智算基礎設施解決方案的架構：



為提升所交付AIDC的性能和穩定性，我們的解決方案依託兩大支柱：部署我們專有的軟件基礎設施Alaya NeW OS及我們的全棧硬件工程能力。通過將Alaya NeW OS作為平台控制層嵌入到我們所集成的異構硬件集群中，我們將裸金屬服務器轉變為一個統一的、可用於生產的環境，針對AI工作負載進行優化。

業 務

下圖展示我們所建設的AIDC的結構佈局：



下圖展示我們於中國各地所建設的AIDC的分佈情況：



我們解決方案的工作流程

我們智算基礎設施解決方案下的項目通常遵循約六個月的生命週期。該過程從接洽和設計開始，我們與客戶合作定義架構要求並選擇最佳硬件配置。在主要通過公開招標獲得合同後，我們將啟動訂單驅動的採購。實施階段涉及現場硬件集成和Alaya NeW OS的部署，最終在設施「點亮」並作為生產就緒環境交付之前進行嚴格的測試和驗收程序。

業 務

我們工作流程的關鍵階段如下：

關鍵階段	已執行工作及時間跨度	資金流
接洽、設計及招標	我們於早期規劃階段與客戶接觸，以了解對規模及應用場景的具體要求。我們於招標過程中提供架構設計及硬件選擇建議。此過程通常耗時約三個月。	無資金流
合約簽訂及預付款	我們與客戶訂立合約，並通常就整個項目收取固定價格，該價格透過招標過程釐定。此過程通常耗時約一至二週。	客戶於合約簽訂時支付前期預付款
硬件採購及軟件部署	我們可為客戶採購硬件，並將 Alaya NeW OS 部署到硬件集群上。我們的工程師會進行系統調試及網絡優化。此過程通常耗時約兩至三個月。	客戶於硬件交付時支付進度款
點亮及驗收	該項目以 AIDC 的「點亮」結束，隨後進行系統檢查、性能測試及由客戶正式驗收。此過程通常耗時約兩週。	客戶於系統驗收時支付尾款；視質保完成情況而定的可選留存款項

面向 AIDC 的 AI 系統：Alaya NeW OS

Alaya NeW OS 為我們專有的雲原生操作系統，專為 AI 工作負載的獨特需求而設計。作為我們智算基礎設施解決方案的核心軟件基礎設施，Alaya NeW OS 為異構硬件集群與各種 AI 應用之間的關鍵層。部署 Alaya NeW OS 可將一組物理硬件轉換為統一的智算環境。

業 務

Alaya NeW OS集資源管理、工作調度及系統監控功能為一體，可有效管理具有一定複雜性的現代AI基礎設施。通過將物理資源虛擬化到標準池，Alaya NeW OS使用戶免受底層硬件配置的複雜性的困擾。這種虛擬化使AI任務能夠在單個集群內跨不同類型的服務器執行。

不同於主要關注服務器穩定性及基本虛擬化的傳統數據中心管理軟件，Alaya NeW OS專門為AI訓練和推理的高吞吐量、低時延要求而構建。通過其複雜的內核和調度算法，Alaya NeW OS管理整個集群的數據和指令流，確保算力高效地傳遞到最需要的地方。這種方法最大限度地利用了我們管理的物理基礎設施。根據灼識諮詢的資料，Alaya NeW OS為中國首個在四個關鍵領域(即算力調度、模型訓練、模型推理及數據處理)獲得中國信通院認證的AI操作系統。

硬件工程能力

我們擁有整合複雜硬件和軟件系統的深度工程能力。我們的服務完成了部件採購到運營準備之間的環節，使AIDC達到AI應用所需的嚴格性能標準。

我們已實施一項全面的戰略，包括計算資源、存儲系統及網絡設備的配置。具體而言，我們可構建採用「軌道優化」互連技術的先進網絡拓撲，以最大限度地提高帶寬利用率。這種架構允許我們的客戶利用適合特定需求的最佳硬件解決方案，同時保持智算基礎設施的穩定性和持續性。

關鍵能力

我們的智算基礎設施解決方案以一系列高級功能為基礎：

- **異構資源調度。** Alaya NeW OS提供對各種計算資源的統一管理。我們的軟件可抽象化底層硬件的差異，允許異構服務器能被匯聚成一個標準化的資源池。
- **Serverless架構。** Alaya NeW OS採用serverless架構，實現了硬件資源與軟件任務之間的完全異步解耦。該設計可實現算力更加靈活地分配，而無需硬件鎖定。
- **端到端性能優化。** Alaya NeW OS協調存儲、網絡和計算資源，以最大限度地減少延遲和消除瓶頸。對該等資源的優化可提高利用率並減少閑置計算時間。
- **跨AIDC分佈式網絡。** Alaya NeW OS促進跨多個AIDC的資源互連和統一調度。這種能力使得客戶能夠將算力擴展至單個數據中心的物理限制以外。

業 務

用例

案例A: 一家國有集團

我們賦能中國西部一家國有集團建立一個區域公共計算中心，以推動文化旅遊及相關產業的創新。為支持這一倡議，我們部署了我們的全棧解決方案，用於建設其大模型智算中心。我們定製了國際先進的異構計算架構，並規劃了一個容量為超過4,500 PFLOPS的高性能智算集群。我們的解決方案利用了當地的自然冷卻資源和清潔能源，結合了高效的冷卻系統和綠色運營方案。此外，我們亦開發了一個統一的智能雲服務平台，以實現計算資源的按需彈性調度，提供從基本計算租賃到AI模型訓練和行業應用集成的端到端工具支持。該項目已成為該地區綠色智能計算的示範項目，使當地企業能夠以低成本獲得高性能算力，並為沉浸式文化旅遊、製造業優化及精準農業等多種場景提供支持。

案例B: 一家大型國有投資集團

我們服務於一家大型國有投資集團，致力於鞏固數字經濟基礎，助力當地產業高質量升級。客戶選擇我們的定製解決方案來構建一個集成的智算中心。我們設計了一個配備智算設備的超過2,000 PFLOPS主流智算架構。至關重要的是，我們嚴格按照國家信息安全分級保護標準，構建了涵蓋硬件、網絡和數據層的多維保護體系。此外，我們為客戶建立了多元化的服務模塊，提供計算租賃、聯合模型研發、行業解決方案實施等定製化服務，同時與國內領先的AI軟件企業培育合作生態系統。自運營以來，該中心提供了穩定的智算輸出，成功支持了智能製造檢測、金融風險控制建模及生物製藥研發等關鍵場景，從而促進了傳統產業向智能製造的轉型。

智算雲服務

我們的智算雲服務通過Alaya NeW Cloud平台提供，為企業和開發人員提供雲原生AI開發和生產平台。儘管我們的智算基礎設施解決方案擴展了算力的物理供應，但Alaya NeW Cloud專注於該等資源的可及性及商業化。

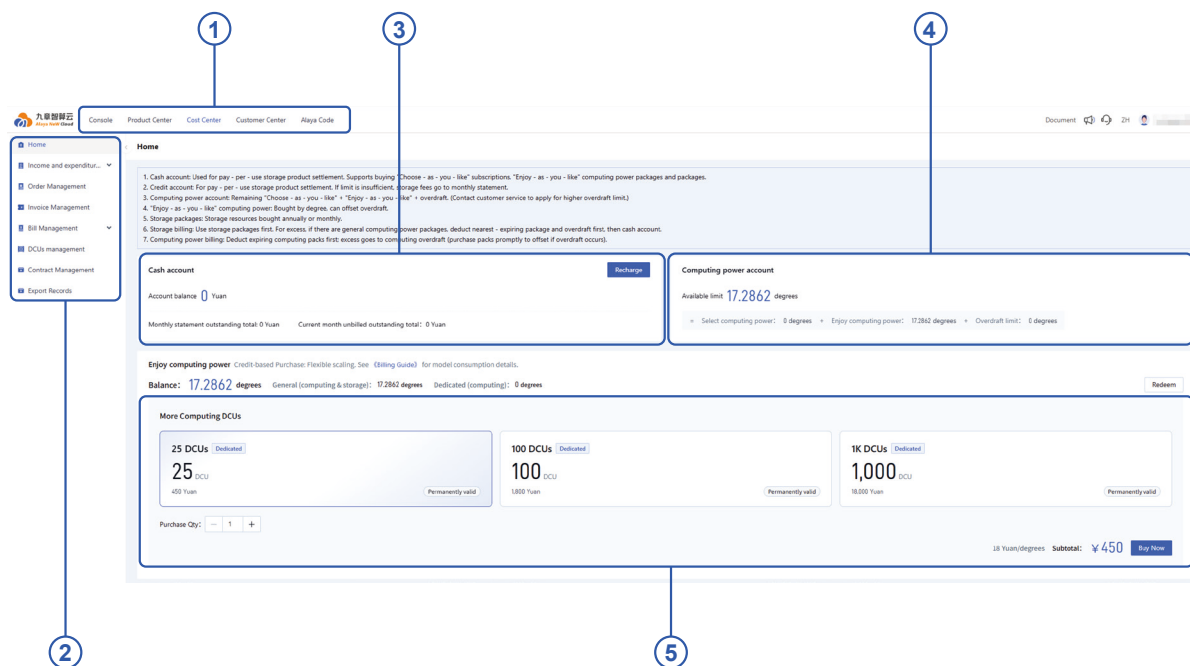
該分部體現了我們的普惠算力理念。通過抽象底層物理基礎設施的複雜性，從智算服務器到網絡設備，我們將算力轉化為標準化公用資源，使其能夠無縫、按需使用。這種方法顯著降低了AI開發人員的准入門檻，使他們能夠專注於模型創新而非基礎設施管理。Alaya NeW Cloud保護用戶免受底層硬件碎片化的影響。該平台將異構資源聚合到一個資源池中，將硬件虛擬化為標準化公用資源。於2026年，我們的Alaya NeW Cloud成功通過中國信通院進行的智算雲平台能力評測。

業 務

下圖展示我們提供的智算雲服務的架構：



我們 Alaya NeW Cloud 平台的用戶界面展示如下：



附註：

- (1) 平台導航欄。顯示Alaya NeW Cloud平台的主要功能模組，使用戶能夠於該平台的服務及集成工具鏈之間進行導覽。
- (2) 賬戶管理側邊欄。提供賬戶管理功能的集中入口，支持平台透明且用戶友好的計費操作。
- (3) 現金賬戶面板。顯示用戶的實時現金結餘及未結清的月結單金額，並提供充值功能，以便順暢完成預付費結算。

業 務

- (4) 算力賬戶面板。實時顯示用戶可用的DCU配額，並細分列示已訂閱算力及透支限額。
- (5) DCU包購買面板。展示分級規模的標準化專用DCU包，採用固定單價，使用戶能夠靈活獲取算力資源。

我們的產品組合

我們提供一個分層的產品組合，包括VKS、DKS以及Alaya Token。VKS和DKS讓客戶能夠以不同程度的靈活性與性能隔離來存取計算資源，而Alaya Token則透過提供基於詞元消耗的託管式模型推理，進一步抽象化底層基礎架構與模型部署層。

VKS

我們的VKS專為需要高度靈活性的場景而設計，例如推理及微調。VKS利用我們的serverless架構來提供可在數秒鐘內進行擴展或縮減的彈性計算資源。我們於VKS下提供兩種主要服務模式：

- **訂閱**。針對需求穩定但有彈性的客戶，用戶可以自由組合資源，優先使用旗艦算力服務器，同時訂閱一個根據實際使用情況扣費的靈活配額。
- **按量付費**。該模式專為具有波動需求的用戶而設計，其特點是無需預付費用，並嚴格根據所消耗的DCU計費。

DKS

DKS專為需要持續高性能的大規模模型訓練任務而設計，可為客戶提供對資源的獨家訪問，以避免資源爭用，並滿足了大規模模型開發與訓練的嚴格要求。

Alaya Token

Alaya Token是我們的託管式模型推理服務，可讓企業客戶與開發者透過統一的API接口調用彼等選定的第三方大語言模型。在該服務模式下，模型廠商提供多元模型，而我們提供底層算力以部署及運行該等模型，從而生成詞元，而該等詞元其後由終端用戶按需使用。使用者無需自行採購計算資源，亦無需獨立部署與運營相關模型，只需透過我們的平台向指定模型提交請求，即可按需取得模型輸出結果。Alaya Token旨在支持主流API規格，讓客戶僅需進行有限的調整適應，即可將服務整合至現有的開發工作流程中。

業 務

下圖載列Alaya Token下如何生成詞元。



我們管理底層的計算資源調度、模型部署及推理操作，並致力於透過我們的Token工廠技術架構提升服務效率。於2026年，Alaya Token通過中國信通院進行的高質量Token雲服務能力評估。有關更多詳情，請參閱「我們的技術—AI工廠—Token工廠」。截至最後實際可行日期，我們已就Alaya Token服務訂立約十份合約。於往績記錄期間，並無就該安排確認任何收入。

AI開發工具鏈

為促進我們計算資源的有效利用，Alaya NeW Cloud集成了一套功能工具，旨在簡化AI開發生命週期。

Alaya Lab

Alaya Lab為我們專為AI研究人員和工程師打造的一站式AI開發平台，旨在提供一個貫穿AI項目全生命週期（從數據準備、模型構建到調試與部署）的統一開發環境。Alaya Lab解決了傳統AI開發中團隊被迫在工作流程的各個階段使用數十個互不連通的工具的固有碎片化問題。

在實際應用中，這種碎片化問題給開發效率帶來了巨大的隱性成本：在多步並行流水線、實驗跟蹤工具、訓練系統和部署平台之間進行上下文切換，會耗費本可投入到模型創新本身的時間與認知資源。Alaya Lab通過將數據管理、代碼調試、模型開發和一鍵提交訓練等功能深度整合到統一工作空間，解決了這一結構性低效問題。開發人員無需再處理本地環境與雲資源的適配問題，也無需解決依賴衝突。

業 務

在團隊層面，Alaya Lab支持協作工作流，使成員能夠共享數據集與模型版本、統一流程，並累積機構知識。值得注意的是，Alaya Lab與極核訓練實現原生集成：用戶在Alaya Lab內完成數據準備與模型調試後，可無縫將訓練任務提交至極核訓練進行分佈式計算與資源編排。這種集成在單一平台內實現了從開發到訓練的閉環，使研究人員無需離開主要工作空間，即可獲得企業級分佈式訓練能力。

極核訓練

極核訓練為我們基於Kubernetes的serverless分佈式訓練服務，旨在解決企業級AI訓練中固有的計算和管理複雜性。其將基礎設施管理、資源調度和分佈式框架適配封裝到一個統一的核心任務引擎中。這種設計允許用戶一鍵啟動高性能的AI模型訓練任務，有效地將訓練過程與底層操作和維護解耦。通過提供資源隔離和成本可控的訓練環境，極核訓練顯著提高了客戶的訓練效率。

為確保無縫的用戶體驗，極核訓練提供了一個預先集成的、無需配置的標準化計算環境，支持主流AI框架，同時亦兼容自定義鏡像。該服務採用「佔用計算，完成後釋放」模型，這有助於集群資源的高效周轉，並最大限度地提高智算服務器利用率。

此外，極核訓練採用智能的多維調度策略，在集群效率與用戶公平性之間取得了平衡。其識別任務特徵以優化排序，確保較短的任務不會被資源密集型任務所阻礙。極核訓練建立在具有故障自愈功能的成熟集群架構之上，能夠自動緩解操作故障，為大規模分佈式訓練任務提供了穩健可靠的基礎。

AgentiCTRL

AgentiCTRL是我們專為先進AI智能體訓練打造的強化學習平台，深耕基於Transformer的大模型訓練場景。該平台旨在為智能體AI的研究與部署提供全面的工程基礎設施。傳統的強化學習工作流程給研究團隊帶來了巨大的工程負擔：從零開始構建訓練流水線、管理高成本的環境模擬，以及彌合大語言模型與強化學習框架之間的兼容性差距，一直是長期存在的挑戰。AgentiCTRL旨在系統性地解決這些瓶頸。

該平台基於統一的訓練引擎構建，將工作流程編排、環境模擬、模型交互和策略優化深度整合到單一操作層中。無論任務涉及單智能體決策還是多智能體協同，AgentiCTRL都能提供一個穩定且可擴展的訓練基礎，從而顯著增強模型在複雜環境中的自主控制和自適應任務執行能力。平台提供開箱即用的仿真環境和完備的訓練工具包，原生兼容主流強化學習框架，並針對性地進行優化。

業 務

AgentiCTRL還與Alaya Lab原生集成：用戶可在Alaya Lab中開發和調試智能體，而後調用AgentiCTRL完成強化學習閉環。兩個平台共同構成了端到端的基礎設施，賦予智能體在真實場景下自主推理和行動的能力。對於深耕智能體算法前沿領域的強化學習研究人員以及大規模部署智能體AI的工程團隊而言，AgentiCTRL提供了核心訓練基礎設施，助力新一代智能體應用落地。

關鍵能力

我們的智算雲服務以一系列高級功能為基礎：

- **創新定價和度量。**我們引入了DCU標準，並於市場上將其推廣為「一度算力」，其定義為等同於312 TFLOPS的FP16算力持續運行一小時。這種標準化保護用戶免受異構硬件複雜性的影響，允許將算力作為公用資源進行計量和交易。我們基於DCU的模型提供了靈活的資源配置和可視化監控，使用戶能夠精確地為其消耗的算力付費。通過消除與閒置資源相關的成本，我們大大降低了AI開發的准入門檻，為客戶實現了高成本效益。
- **資源彈性和簡化運營。**我們的平台採用serverless架構，提供即時可擴展性和簡化操作。其抽象化基礎設施管理層，允許用戶專注於模型開發，而無需配置或管理物理服務器。資源可立即獲取和釋放，以滿足實時工作負載需求，確保了動態配置並最大限度地提高了利用率。
- **集成工具鏈與直觀體驗。**我們的平台配備了一套全面的預配置AI開發工具和框架，用戶可通過高度直觀的用戶界面進行訪問。通過簡化複雜的環境配置並提供用戶友好的交互式控制台，我們顯著降低了AI開發的技術准入門檻。藉助該平台，開發人員及研究人員能夠無縫啟動、管理及監控其模型訓練及推理項目，而無需具備底層基礎設施運維的深厚專業知識。
- **智能體AI強化學習。**憑藉我們專有的AgentiCTRL功能，我們通過強化學習來訓練基於Transformer的智能體。該功能增強了AI智能體的決策和控制能力，使平台能夠支持複雜、動態的AI任務，而不僅僅是簡單的模式識別。
- **託管式模型推理。**透過Alaya Token，我們的平台將模型部署與推理操作抽象化為可透過API存取的服務。客戶無需管理底層服務器或部署環境，即可調用指定的第三方模型，而我們的平台則負責管理底層資源分配、推理操作、吞吐量及服務可用性。

業 務

我們的算力池

我們的算力池是我們智算雲服務的基礎。我們匯聚來自三大來源的算力，並在我們的「中央調度＋區域節點」框架下，透過 Alaya NeW OS 將其作為單一資源池進行調度。Alaya NeW OS 抽象化硬件類型與設施位置之間的差異，使我們能夠基於工作負載特徵、資源可用性及網絡就近原則，跨區域節點靈活分配工作負載。

- **來自我們的智算基礎設施解決方案客戶。**於交付 AIDC 後，我們可根據另行訂立的協議向客戶租賃 AIDC 的指定部分。我們承擔租賃部分內服務器及相關設備的營運管理。由於該等設施由我們設計及集成，其原生運行 Alaya NeW OS，並能整合進我們的算力池。將部分設施租賃予我們，使彼等得以提升其資產的利用率，同時使我們得以在毋須為設施建設出資的情況下確保長期獲得算力。有關更多詳情，請參閱「－客戶與供應商的重疊」。
- **來自我們管理的第三方算力。**我們根據合作安排自第三方營運的 AIDC 租賃算力。我們對所租賃的設施進行技術升級及標準化改造，使其可與我們的其他資源一併透過 Alaya NeW OS 進行調度。
- **來自我們的服務器。**我們亦透過自有的服務器產生算力，該等服務器部署於向第三方租賃的數據中心設施內。

進一步詳情請參閱「－主要運營數據」。

用例

案例C: 一家具身AI企業

我們為一家位於北京的企業提供服務，該企業專門從事具身AI和大模型研究，專注於視覺語言動作（「VLA」）技術和機器人應用。為解決與傳統固定資源採購相關的高成本問題，客戶採用 VKS 服務下我們的按量付費服務。利用彈性擴展模型，計算資源可根據實際工作負載需求動態擴展或釋放。利用我們獨特的 DCU 計量標準，客戶可精確地支付其訓練任務所消耗的實際計算時間。該模型為客戶提供了強大的性能，同時通過消除與閒置資源相關的成本，顯著降低了成本，幫助客戶將總體成本降低了 73.8%。

案例D: 一家領先的服裝行業互聯網企業

我們賦能一家服裝行業企業利用多模式 AIGC 技術來提高時裝設計和營銷效率。在傳統的裸金屬租賃模式下，客戶的資源利用率不到 30%，導致閒置機器的大量預算浪費。通過遷移到我們的雲平台並以按量付費方式配置 VKS，客戶現在可根據實時工作負載需求動態擴展資源。這種基於使用量的計費模式使客戶能夠避免閒置資源的成本，從而提升其資源效率。

業 務

案例E:一家領先的在線社交網絡平台

我們為一家於中國擁有重要市場份額的頂級在線關係型社交網絡平台提供計算支持。該客戶的計算需求不時波動。傳統的每月租賃裸金屬服務器導致在低流量期間產生過多浪費。通過採用我們的VKS服務(該模型支持最大並發性，同時僅對實際使用量進行計費)，客戶的月度成本顯著降低。這意味著與之前的模型相比節省88.2%的預算，證明了我們的彈性計算服務在工作負載波動的企業中具有成本效益。

AI基礎軟件服務

我們亦提供AIFS服務，該服務包含專為銀行和金融企業等專業機構量身定製的AI軟件解決方案。AIFS旨在優化內部運營，作為一款軟件平台，為客戶提供數據處理及AI模型管理所需的工具集。我們以項目為基礎運營AIFS業務。此類解決方案通常在開發後作為完整的軟件產品交付。

隨著我們戰略性地將重心轉向智算基礎設施解決方案和智算雲服務，AIFS業務正逐步退出市場。進一步詳情請參閱「我們的業務模式—我們的戰略性業務升級」。

我們的技術

技術創新是我們提供普惠、高效、可靠的智算服務的基石。我們已建立全面的全棧技術架構，融合基礎設施與先進的軟件平台及以生產為導向的架構。憑藉強大的計算、網絡、存儲及自動化運維能力，我們的技術棧支持涵蓋數據處理、分佈式模型訓練及高吞吐量推理的AI工作流程，從而提供AI原生基礎，最大限度地提升資源利用率及系統可靠性。

智算基礎設施及系統

智算集群

我們的計算層旨在支持AI模型所需的大型規模。我們已部署能夠為單個集群提供大規模計算資源的計算池，滿足具有數千億參數的模型的嚴格訓練需求。這一層面的核心是我們的拓撲感知智能調度系統。該系統感知計算服務器的物理拓撲，確保任務分配予物理距離最近的計算服務器，從而最大限度地提高通信效率。此外，我們已採用聯合軟硬件加速技術，將服務器性能增強與軟件優化(如專用通信庫及RDMA)相結合。這種協同作用顯著降低了交互延遲，提高了整體計算性能。

業 務

我們的計算架構旨在支持大規模AI工作負載的縱向擴展及橫向擴展部署模型。於縱向擴展層面，針對需要頻繁數據交換的工作負載，我們尋求優化高密度運算領域內加速器之間的通訊。在橫向擴展層面，我們透過高速網絡連接跨多服務器及集群的計算資源，並透過我們的統一編排平台對其進行管理。我們的拓撲感知調度系統旨在分配任務時，考量計算資源在實體及網絡上的接近程度、工作負載特徵及資源可用性。該方法旨在為合適的工作負載減少不必要的跨服務器通訊，並支持跨多集群及區域服務器協調使用計算資源。

效率導向型智算網絡

為解決大規模訓練中數據傳輸的瓶頸問題，我們已構建一個具有「軌道優化」互聯性的高性能網絡層。這種基於虛擬化的軌道優化拓撲實現了同一平面上相同等級服務器的直接連接，消除了交換機之間的流量衝突。我們亦利用網絡內計算技術，在傳輸過程中，數據由交換機處理，結果直接返回，使得All-Reduce性能提高兩倍，數據流量減少50%。此外，我們的網絡擁有極高的帶寬利用率，能夠動態地感知鏈路擁塞並自動重新路由數據包。這使得網絡帶寬利用率超過95%，有效地解決了由網絡延遲引起的服務器等待瓶頸。

新一代存儲系統

我們的存儲層由DingoFS提供支持，這是一個完全開源的混合存儲系統，由我們為AI的大規模並行讀寫需求而專門設計。我們實現了分鐘級檢查點優化，將數百GB規模的模型的歸檔時間從30多分鐘縮短至不到60秒。這確保了服務器始終專注於計算，而非等待數據讀寫。我們亦實現了優化的混合讀/寫加速，對小文件採用高速緩存，對大文件採用流式讀取，支持EB級數據的高效存儲和訪問。

自動運維與故障管理系統

我們已建立一個專有的自動運維與故障管理系統，以確保長期AI訓練任務的極端穩定性。該系統具有硬件故障自動監測和隔離功能。其實時監控我們的集群的狀態，自動探測異常，並在數秒鐘內隔離故障。至關重要的是，該系統支持從斷點進行不可察覺的恢復，允許訓練任務在並無人為干預的情況下從最近的檢查點自動恢復。

業 務

資源與工具深度融合的AI原生雲架構

我們的Alaya NeW Cloud體現了有別於傳統雲平台的轉變。這是一個深度集成的智能雲，將龐大的計算資源與複雜的工具鏈相結合。我們的全棧AI原生架構優化了端到端模型和AI原生應用的整個工作流程。通過重建計算、網絡和存儲，我們提供了一個無縫的環境，使客戶能夠自主創建智能。我們的平台覆蓋了AI開發的整個生命週期，具有數據處理、模型訓練和推理服務等專業模塊，以及AgentiCTRL強化學習平台，確保了對現代AI工程的全面覆蓋。

AI工廠

AI工廠是我們用於模型訓練、後訓練及大規模推理的技術架構，目前正持續進行開發與優化。旨在將我們底層的運算、網絡、儲存及調度能力，轉化為以生產為導向的系統，其涵蓋訓練工廠及Token工廠。訓練工廠支持異構運算資源上的大規模預訓練、微調及強化學習工作負載，通過整合資源調度、分佈式訓練、故障恢復及環境模擬，提高集群利用率並縮短模型迭代週期。Token工廠通過自動化模型部署、彈性資源調度、內存管理及分佈式快取，降低單位輸出的成本，並將吞吐量及資源利用率最大化，為Alaya Token及Alaya NeW Cloud上的其他推理工作負載提供支撐。

Alaya Atom Pod: 我們用於大規模AI推理的參考架構

為支持對大規模AI推理及詞元生成日益增長的需求，我們正在開發Alaya Atom Pod（一種可擴展AI推理集群的參考架構），整合高密度計算資源、標準計算服務器、高速網絡、分佈式快取及我們的統一調度軟件。Alaya Atom Pod旨在根據預填充和解碼階段的不同需求動態分配資源，其配備了分佈式鍵值緩存層，用於儲存、傳輸及復用中間推理數據，從而減少重複計算，並提升長上下文、多輪對話及智能體AI應用的效率。我們計劃以模塊化、可擴展的單元形式部署該架構，以管理不斷增長的工作負載所涉及的異構計算資源；同時，該架構的持續開發和分階段部署仍能適應不斷演變的模型架構和硬件配置。

研發

我們對研發的高度關注是我們運營的基礎，亦為我們競爭優勢的主要來源。我們對研發的持續投入使我們能夠保持在技術創新的前沿，並滿足客戶不斷變化的需求。我們已經並將繼續在研發活動中投入大量資源。於2023年、2024年及2025年以及截至2025年及2026年6月30日止六個月，我們的研發開支分別為人民幣111.6百萬元、人民幣103.4百萬元、人民幣121.6百萬元、人民幣53.1百萬元及人民幣66.8百萬元，佔各期間我們總收入的105.0%、26.2%、11.1%、17.3%及11.4%。我們擁有一支高素質、經驗豐富、專業背景多樣的研發團隊。截至2026年6月30日，我們的研發團隊由156名成員組成，佔員工總數的51.5%。

業 務

我們的研發流程

為有效地將市場需求轉化為高質量、商業上可行的產品，我們已建立一個系統化、結構化的研發流程。該流程旨在確保在我們研發活動的所有階段進行嚴格的質量控制和有效的項目管理，其中通常包括以下關鍵階段：

需求洞察與架構定義。我們的研發週期始於由產品設計管理團隊主導的結構化探索流程，將客戶實際使用中的智算痛點轉化為明確的產品需求與功能定義。隨後進行嚴格的全棧技術審查，確保每項產品決策皆兼顧市場相關性與技術可行性。

模塊化及協同工程。我們的工程架構圍繞擁有貫穿整個技術棧的特定領域專業知識的專門團隊構建。我們的研發人員跨關鍵職能領域展開協作，涵蓋底層內核優化、存儲性能調優及虛擬化，以及智能資源調度、編排及包括IDE、Web 控制台及統一API在內的面向用戶界面。關鍵在於，所有代碼在合併至主代碼庫前，均須經過強制性的靜態分析與安全漏洞掃描，確保從源頭對憑證洩漏或高風險安全漏洞零容忍。

智算仿真與全棧壓測。任何功能在部署到客戶環境前，均須通過我們專有雙層質量保證流程。首先，我們在單個集群內模擬並發計算資源調度，在模擬我們最嚴苛企業部署場景的條件下，對調度器的吞吐量及排隊延遲進行壓力測試。其次，我們進行涵蓋存儲、智算和網絡的端到端基礎設施壓力測試，重點關注大模型訓練場景，通過模擬海量檢查點寫入操作，以驗證在極端負載下的高吞吐量、網絡穩定性及最佳MFU。這項在發佈前模擬測試的投入，直接保障了客戶的系統運行時間，並維護了我們自身的可靠性聲譽。

灰度發佈與自動化部署。我們採用灰度發佈策略，系統性地降低每次產品更新的風險。新版本會先在內部測試環境完成驗證，接著逐步推送給部分客戶，最後再進行全面生產部署。整套升級流程由我們自主研發的命令控制平面統一調度，可實現終端用戶的無感升級。

可觀測性與閉環反饋。我們的營運基於全景監控系統，以實時跟蹤利用率及RDMA網絡丟包率等關鍵指標。生產中一旦偵測到任何異常均會自動觸發結構化研發回溯工作流程，形成「發現、修復、測試、歸檔」完整且可追溯的閉環機制。這一自我優化的反饋機制使得每次故障都會隨著時間推移系統性地提升系統穩定性，並隨著部署規模的擴大，進一步鞏固我們平台的可靠性優勢。

業 務

我們的研發重點

我們的研發策略從根本上植根於解決AI時代的關鍵挑戰：算力稀缺、效率瓶頸及開發複雜性。目前，我們的研發工作集中於三個層次上：鞏固底層基礎設施底座、突破性能瓶頸及完善AI工具鏈。我們亦計劃透過我們的Token工廠降低推理服務的每個詞元成本，並開創能源定義的算力協同調度。我們的目標是構建一個與全球標準相媲美的高性能、企業級的智算雲平台。

基礎設施層：極致性能與硬件卸載

在我們平台的基礎設施層，我們的工程重點在於充分發揮現代硬件的極致性能，同時有效控制成本。其中一項關鍵措施是我們的硬件卸載架構，通過此架構將網絡與存儲處理的負擔從主CPU轉移至專用協同處理器。最終結果是，我們的客戶不僅能獲得接近裸金屬的智算性能，同時仍能享有強大的多租戶安全隔離。

在網絡方面，我們正分階段遷移至專為AI工作負載設計的高性能網絡協議。傳統互聯網協議從未針對大規模模型訓練的全互聯通信模式進行設計，這會造成吞吐量瓶頸，直接拖慢訓練任務的速度。我們的路線圖首先將存儲流量轉移至新協議，隨後擴展至智算流量，此舉亦具備降低對硬件供應商的依賴及削減基礎設施成本的附加戰略優勢。最後階段將導入進階流量管理機制，以消除影響AI訓練效率的網絡不穩定性。這一切的基礎是我們專有的DingoStack網絡平台，我們持續對其進行優化，以消除高併發工作負載下的性能瓶頸。

容器層：大規模異構資源調度

容器層體現了我們大規模及高靈活性管理智算資源的能力。我們的調度引擎是旗艦功能，單一調度系統可同時協調數千台服務器，且不會降低穩定性。這直接使客戶能在我們的平台上運行最大規模的AI訓練任務，無需自行處理複雜的多集群解決方案。

除了規模之外，我們正投資於發展硬件中立的資源管理。我們的平台不會受限於單一服務器製造商，而是能夠靈活地匯集並調度來自不同硬件供應商的智算資源。這使我們的客戶免受供應鏈中斷的影響。

最後，我們的命令控制平面正進行優化，以實現近乎即時的資源調配，將啟動時間縮短至數秒之內。這是我們serverless服務的技術核心，確保閒置智算能力能被迅速回收並重新部署，從而直接提升整體資源利用率的經濟效益。

業 務

平台層：全棧AI開發體驗

於平台層，我們專注於降低AI開發者的門檻。我們自主研發的DingoFS存儲系統，旨在取代長期以來為智算雲供應商帶來主要基礎設施開支的高成本並行文件系統。我們的工程專注於AI工作負載中的兩項性能特徵：數據集加載時對大量小文件的快速存取，以及模型檢查點保存時的高吞吐量寫入。我們輔以我們的DingoSpeed緩存層，使數據集能夠在數秒內(而非數分鐘)分佈於不同數據中心。

在面向開發者的層面，我們正將Kubernetes託管服務從一項管理的服務，轉型為具備服務水平承諾的企業級平台。最重要的是，我們的Alaya Lab平台正解決AI開發中一個長期存在的痛點：開發者本地環境與雲基智算集群之間的脫節。通過我們內部研發的高速隧道技術，開發者可在個人筆記本電腦上編寫及調試代碼，同時將其直接在我們的雲集群上實時執行，從而消除傳統雲開發中繁瑣的上傳—配置—重啟流程。該流程導致開發體驗緩慢且造成阻滯。該體驗的提升，是促進開發者採用及留存的重要動力。

Token工廠：降低每個詞元成本

我們專門設立了一項研究方向，致力於系統性地降低透過Token工廠交付的高質量輸出的單位成本。在營運方面，我們正實現模型部署自動化，以支持標準化、大規模的多模型發佈，從而減少人力需求，同時構建在秒級內擴展的彈性推理能力，在需求週期內回收閒置資源，並嚴格按消耗計費。我們的很大一部分工作旨在提高模型執行層面的計算效率。結合DingoFS緩存，我們透過重複使用先前計算的鍵值狀態來提高KVCache命中率，減少首個詞元時間並避免冗餘的預填充計算；我們亦正分別優化DingoFS Connector緩存路徑，以進一步提高命中率及吞吐量。我們亦推行異構預填充與解碼分離，在不同類型的加速器上分離預填充及解碼階段，以增加每單位算力的有效輸出。同時，我們正在開發推理工作負載所需的底層通信庫及穩定性支持。綜合而言，該等舉措旨在隨著時間大幅降低每個詞元的推理成本，與我們專注於降低每單位產出成本的方向一致。

能源定義架構：算力協同調度

該舉措將能源從被動的成本投入重新定義為主動的調度變量，以決定計算工作負載部署及執行的地點與時間。該架構由我們與大學及技術合作夥伴合作開發，包含一個理論框架、兩項調度技術及一個作為閉環運作的模擬環境。理論基礎對計算需求與電力供應之間的耦合進行建模，為工作負載動態響應峰谷電價提供基礎。第一項調度技術為綠電優先遷移，其按碳彈性對工作負載進行分類，並跨越時間及地域重新分配工作負載，以將可再生能源消耗最大化。第二項為多因素感知調度，對電價、碳價及可再生能源供應等輸入應用強化學習，以共同優化跨地區的首個詞元時間、碳排放、能源消耗及資源利用率。結合新能源、電網、計算及存儲領域的模擬環境，可用於在投入生產前驗證及調整部署，並在實時環境中實現協同控制，最終目標是減少每個詞元的能源消耗。

業 務

我們的銷售網絡

於往績記錄期間，我們的解決方案及服務主要在中國銷售。我們已建立一個依靠直銷服務我們的核心客戶的銷售網絡。我們亦已在銷售網絡下發展渠道代理，以補充我們直銷團隊的能力，擴大我們的地理覆蓋範圍，及提高我們的品牌知名度。我們於往績記錄期間的業務受季節性因素影響，詳情請參閱「風險因素－我們的收入受季節性波動影響」。

我們對客戶採用直銷模式。在這種模式下，我們的內部銷售團隊直接與客戶溝通，了解彼等的具體計算需求和業務場景。這種方法使我們能夠提供高度定製的解決方案，如定製的DKS或VKS服務、Alaya Token服務以及集成的智算基礎設施解決方案。對於我們的智算基礎設施解決方案項目，通常涉及國有企業的複雜要求，我們的直銷團隊管理整個項目生命週期，從最初的諮詢和投標到交付和持續維護，確保了高水平的服務質量和客戶粘性。

我們的項目制業務

於往績記錄期間，我們的部分業務(具體而言為智算基礎設施解決方案及AIFS服務)乃按項目制模式開展。智算基礎設施解決方案項目通常涉及AIDC的交鑰匙建設、集成與交付。鑒於此類基礎設施開發的規模及性質，我們的客戶通常委託我們開展獨立的、一次性項目，而非基於週期性合作，儘管客戶隨後可能會委託我們進行擴建階段或運營支持。AIFS服務項目交付一次性獨立軟件，隨著我們於2023年初進行戰略轉型，將資源集中於專用智算基礎設施及雲服務，我們正逐步淘汰該等軟件。更多詳情請參閱「－我們的業務模式－我們的戰略性業務升級」。截至2023年、2024年及2025年12月31日以及2026年6月30日，我們分別有零個、一個、一個及一個正在進行的智算基礎設施解決方案項目，以及46個、31個、十個及七個AIFS服務項目，未結合同總價值分別為人民幣45.8百萬元、人民幣254.6百萬元、人民幣113.7百萬元及人民幣155.6百萬元。截至2023年、2024年及2025年12月31日止年度以及截至2026年6月30日止六個月，我們自項目制業務產生的收入分別為人民幣106.2百萬元、人民幣395.3百萬元、人民幣1,020.5百萬元及人民幣434.5百萬元，分別佔我們總收入的100.0%、99.95%、92.9%及73.9%。

業 務

就我們的項目制業務部分而言，下表列示所示期間內按業務線劃分的項目數目變動情況：

	截至12月31日止年度			截至6月30日 止六個月
	2023年	2024年	2025年	2026年
年／期初在建項目數量	47	46	32	11
－智算基礎設施解決方案	－	－	1	1
－AIFS服務	47	46	31	10
新承接項目數量	109	76	23	4
－智算基礎設施解決方案	3	6	6	3
－AIFS服務	106	70	17	1
已完成項目數量	110	90	44	7
－智算基礎設施解決方案	3	5	6	3
－AIFS服務	107	85	38	4
年／期末在建項目數量	46	32	11	8
－智算基礎設施解決方案	－	1	1	1
－AIFS服務	46	31	10	7

就我們的項目制業務部分而言，下表列示了所示期間內按未結合同金額及業務線劃分的各項目未完成訂單變動情況：

	截至12月31日止年度			截至6月30日 止六個月
	2023年	2024年	2025年	2026年
(人民幣千元)				
年初／期初未結合同價值(不含增值稅)	53,935	45,752	254,586	113,737
－智算基礎設施解決方案	－	－	230,892	105,853
－AIFS服務	53,935	45,752	23,694	7,885
加：合同金額(不含增值稅) ⁽¹⁾	98,048	604,144	879,691	476,365
－智算基礎設施解決方案	24,938	564,855	876,918	476,098
－AIFS服務	73,111	39,288	2,773	268
減：合同金額(不含增值稅) ⁽²⁾	106,232	395,309	1,020,541	434,544
－智算基礎設施解決方案	24,938	333,963	1,001,958	434,046
－AIFS服務	81,294	61,346	18,583	498
年末／期末未結合同價值(不含增值稅)	45,752	254,586	113,737	155,559
－智算基礎設施解決方案	－	230,892	105,852	147,905
－AIFS服務	45,752	23,694	7,885	7,655

業 務

附註：

- (1) 合同價值的增加乃指因年內或期內新承接的項目而導致的合同價值增加。
- (2) 合同金額的扣減乃指於年內或期內已確認收入的合同金額。

品牌及營銷

我們的營銷策略以我們的品牌定位為中心，即作為普惠算力提供商。為了有效觸達我們的目標企業客戶並塑造行業話語權，我們採用多層次的營銷方法來提高我們的品牌知名度，並接觸到我們的目標受眾。我們在主要機場等高流量交通樞紐戰略性地投放品牌廣告，以瞄準企業客戶。我們積極參加世界移動通信大會、全球雲網寬帶產業協會、GITEX Global及一帶一路軟件館等頂級行業會議和論壇。通過該等平台，我們展示了我們的技術能力，發佈了新產品更新，並與行業領導者合作，共同塑造AI基礎設施標準的行業話語權。我們亦利用社交媒體平台傳播有關我們產品的內容。

作為企業客戶拓展的補充，我們大力推動產品的可及性，以觸達更廣泛受眾(包括個人開發者和學生)。在營銷活動中，我們重點突出Alaya NeW Cloud高度直觀的界面和豐富的內置工具鏈，將複雜的AI開發轉化為易用、友好的體驗。為進一步降低使用門檻，我們針對個人用戶開展了廣泛的教育推廣活動，賦能沒有專業AI知識或編程背景的用戶無縫採用和使用我們的計算服務。通過在各社交媒體平台進行精準數字營銷和內容傳播，我們持續與開發者社區互動，以培育普惠的計算生態系統。

定價

我們的定價策略旨在靈活、以市場為導向，滿足客戶的多樣化需求。我們已針對不同的業務部門建立不同的定價模式：

對於我們的智算基礎設施解決方案及AIFS服務，我們一般採用按單個項目基準確定的固定價格模式。合同價格通常通過招投標程序或商業談判確定。我們對智算基礎設施解決方案項目的定價會考慮：(i)硬件採購成本；(ii)我們專有軟件的許可費；(iii)系統集成和架構設計的複雜性；及(iv)需交付的計算集群的規模。我們對AIFS服務項目的定價，會綜合考慮軟件定製的複雜度、開發工時以及具體的功能要求。我們的智算基礎設施解決方案及AIFS服務的價格因規模及服務範圍而異。

業 務

對於我們的智算雲服務，我們採用訂閱制定價模式及按使用量計費的定價模式。

- **DKS**。我們採用訂閱制模式。客戶支付定期費用，以在約定的訂閱期內獨佔特定配置的計算資源。費用通常按每台服務器、每月收取，主要根據硬件資源的類型和性能、集群規模、資源稀缺程度、當前市場狀況以及訂閱期限釐定，與訂閱期內的實際使用率無關。
- **VKS**。我們在VKS下採用兩種定價模式：
 - **訂閱制**。客戶在兩種主要安排中選擇其一，支付月度定期費用：(i)算力配額訂閱，客戶就協定數量的算力額度支付固定費用；或(ii)預留計算資源訂閱，指定計算資源專門分配給客戶，惟須受約定的使用上限所規限。兩種安排的費用通常按月計算，主要根據所需服務器的類型、性能、數量以及訂閱期限釐定。
 - **按使用量計費**。客戶根據其實際的DCU消耗量，按適用的單價付費，無需提前承諾訂閱。我們以「一度算力」標準，推廣此種按量計費模式。適用費率主要根據資源可用性、硬件性能和市場狀況釐定。
- **Alaya Token**。我們採用按使用量計費的定價模式，客戶主要按照詞元實際的輸入與輸出消耗量付費。詞元輸入與輸出量分開計量，且適用的單位費率主要根據所選模型及相關服務規格(包括上下文視窗大小、吞吐量及服務等級安排)釐定。詞元服務費一般按實際消耗量計算，並定期結算。

我們定期審查和調整我們的定價，以反映我們的運營成本(尤其是服務器的採購成本)、市場狀況的變化以及我們促進算力獲取的戰略目標。對於智算雲服務，我們可能會向教育機構和學生開發人員提供促銷價格或折扣，以培育AI人才生態系統。

知識產權

保護我們的知識產權為我們業務的一個關鍵方面。我們未來的商業成功部分取決於我們在數個關鍵領域的能力：為我們的核心技術取得和維持知識產權保護；捍衛和執行我們的專利組合；為我們的商業秘密保密；及確保我們的運營不會侵犯他人的知識產權。

業 務

截至最後實際可行日期，我們於中國擁有125項已註冊專利。截至同日，我們於中國擁有117項軟件著作權及於中國擁有174項已註冊商標。我們通過自主開發獲得專利。截至最後實際可行日期，我們擁有大部分專利以及專利申請，並與第三方就我們的專利及專利申請存在少量共同擁有或共同分享安排。

數據安全

我們高度重視數據安全和保護。我們提供算力和平台，同時數據的所有權和控制權仍由我們的客戶保留。我們已採取一套全面的保護措施，包括保密分類、嚴格的訪問控制、數據加密和脫敏，以防止數據未經授權的訪問、洩露、不當使用、修改、損毀或遺失。

我們已成立一個專門的數據合規部門，並已委任專門的信息安全專業人員來監督我們的數據安全實踐。我們已獲得公認的信息安全認證，包括國家網絡安全等級保護制度三級、信息技術服務標準二級（運維）、能力成熟度模型集成五級以及ISO27001、ISO20000、ISO45001、ISO14001及ISO9001認證，證明我們符合嚴格的行業標準。

數據收集

我們的業務主要涉及向企業客戶提供智算算力和相關雲服務。於我們的正常服務過程中，我們一般不會訪問、收集或存儲我們企業客戶的專有數據集、算法模型或訓練結果。我們的客戶保留對其駐留在我們的計算集群或雲平台上的數據的完全控制權。我們僅收集必要的業務信息，如賬戶註冊詳情、使用日誌及計費記錄，用於賬戶管理和服務提供。

數據安全管理系統

我們已建立一個穩健的數據安全管理框架，涵蓋從生成和傳輸到存儲和銷毀的整個數據生命週期。我們已實施《數據分類管理程序》，將我們的數據分為不同的類別。對高級別信息的訪問嚴格限制在確需知道的基礎上，並須經過嚴格的審批程序。

我們為數據實施生命週期治理。我們利用先進的加密技術在傳輸過程中保護數據，防止未經授權的攔截或篡改。敏感數據存儲在具有嚴格訪問控制的安全加密環境中。我們維持完善的操作日誌，以跟蹤任何數據訪問或修改。此外，我們對過期或已刪除的數據採用不可逆的數據銷毀方法，以確保其無法恢復。

我們要求所有員工（尤其是擔任核心技術和研發職務的員工）在入職時簽訂保密協議。我們亦定期進行數據安全培訓，並為員工入職和離職期間處理機密信息制定了嚴格的協議。

業 務

事故響應機制

我們已制定《網絡和信息安全事件應急預案》，以有效管理潛在的安全漏洞。我們根據安全事件的嚴重程度和影響將其分為四個級別。每個級別對應一組特定的應急響應協議，涵蓋從部門處理至全公司應急動員。我們遵循「早發現、早報告、早處置」的原則，盡量減輕任何安全事件的影響。我們的安全團隊持續監控系統狀態，並須立即報告任何異常情況。如果發生數據洩露或安全漏洞，我們將立即啟動應急響應計劃，以控制事件、評估損害並恢復受影響的系統。我們亦將及時通知受影響的各方，並依法向公安局等相關監管部門報告。

安全架構及技術措施

我們已建立一個穩健的信息安全架構，將我們的專有協議與戰略合作夥伴的安全能力相結合。我們高度重視底層基礎設施的物理和網絡安全。我們已實施多層次防禦機制，包括防火牆、入侵探測系統及抗DDoS保護。

此外，我們對內部運營保持嚴格的數據訪問政策。我們已部署一個安全的內部雲磁盤系統用於文檔管理。訪問我們的內部網絡和敏感數據需要通過多因素身份驗證。我們嚴格將我們的內部開發環境與客戶生產環境隔離開來，以防止交叉污染或未經授權的訪問。敏感數據（包括客戶憑據和關鍵業務數據）在傳輸和存儲過程中會使用行業標準加密協議進行加密。

使用開源模型及軟件

我們就智算雲服務使用若干開源基礎模型及開源軟件組件，涵蓋模型訓練、部署與推理等環節。尤其是，我們的平台設計支持多種第三方模型與軟件框架，而Alaya Token為用戶提供訪問選定的第三方大語言模型（包括開源模型）的權限。我們制定了程序以管理該等開源模型及軟件組件的使用，包括訪問控制、安全審查、漏洞掃描，以及知識產權與授許可合規措施。我們的專有技術、系統架構以及透過平台處理的客戶數據，均受我們的內部信息安全政策及技術保障措施（包括訪問控制、數據隔離及加密）規限。有關更多詳情，請參閱「風險因素—與我們的業務及行業相關的風險—我們使用開源技術可能會對我們的業務營運造成限制。」

於往績記錄期間及直至最後實際可行日期，我們並未經歷任何重大數據洩露、丟失或安全漏洞，亦未受到監管機構有關數據隱私或安全的任何重大罰款或處罰。

業 務

我們的供應商

我們的主要供應商

我們的主要供應商為中國的硬件供應商、技術服務提供商及AIDC租賃服務提供商。於2023年、2024年及2025年以及截至2026年6月30日止六個月，各年度／期間對我們五大供應商的採購額分別為人民幣98.0百萬元、人民幣584.6百萬元、人民幣1,385.9百萬元及人民幣794.9百萬元。於往績記錄期間各期，對我們五大供應商的採購額分別佔我們總採購額的47.1%、74.8%、84.8%及74.5%。於往績記錄期間各年度／期間，我們的五大供應商均為獨立第三方。於往績記錄期間，我們並無遭遇任何硬件設備或算力租賃服務供應的重大短缺或延誤。

下表載列於往績記錄期間各年度／期間我們五大供應商的詳細資料：

截至2023年12月31日止年度

序號	供應商	提供的 產品／服務	採購額	佔總採購額 的百分比	信貸條款	與我們開始業務 關係的年份
			(人民幣千元)			
1	供應商集團A ⁽¹⁾	設備	33,519.2	16.1	就服務器而言，預先付款；就其他設備而言，兩期款項分別於合約簽訂及驗收後15個工作日內支付	2023年
2	供應商B ⁽²⁾	設備	26,400.0	12.7	預付款	2023年
3	供應商C ⁽³⁾	設備	15,520.0	7.5	預付款	2023年
4	供應商D ⁽⁴⁾	物業租賃	15,117.9	7.3	預付款	2023年
5	供應商E ⁽⁵⁾	技術服務	7,395.8	3.5	總款項的60%須於合約簽署後20個日曆日內支付，餘下40%須於服務完成並經驗收後20個日曆日內支付	2022年
總計			97,952.9	47.1		

附註：

- (1) 一家於2003年在中國註冊成立並於深圳證券交易所上市的股份有限公司，主要從事柔性電子技術的研發及應用。
- (2) 一家於2008年在中國註冊成立的有限責任公司，主要從事計算機硬件、軟件及網絡產品的銷售及開發。

業 務

- (3) 一家於1998年在中國註冊成立並於深圳證券交易所上市的股份有限公司，主要從事提供IT基礎設施產品及解決方案。
- (4) 一家於2020年在中國註冊成立的有限責任公司，主要從事科技推廣及應用服務。
- (5) 一家於2020年在中國註冊成立的有限責任公司，主要從事數據服務。

截至2024年12月31日止年度

序號	供應商	提供的 產品/服務	採購額	佔總採購額 的百分比	信貸條款	與我們開始業務 關係的年份
(人民幣千元)						
1	供應商F ⁽¹⁾	AIDC租賃服務	178,256.8	22.8	預付款	2024年
2	供應商G ⁽²⁾	設備	153,090.0	19.6	於驗收當日支付	2024年
3	供應商H ⁽³⁾	AIDC租賃服務	105,886.5	13.6	於發票開具後的10個工作日內支付	2024年
4	供應商I ⁽⁴⁾	AIDC租賃服務； 設備	93,325.1	11.9	兩種付款安排： 預付款；或於驗收後60個工作日 內悉數支付	2024年
5	供應商J ⁽⁵⁾	設備	54,000.0	6.9	預付款	2023年
總計			584,558.4	74.8		

附註：

- (1) 一家於2001年在中國註冊成立的有限責任公司，主要從事廣播電視傳輸設備的製造及廣播電視節目的運營。
- (2) 一家於2017年在中國註冊成立的有限責任公司，主要從事技術推廣及增值電信服務。
- (3) 一家於2024年在中國註冊成立的有限責任公司，主要從事空間智能技術的研發及應用。
- (4) 一家於2015年在中國註冊成立的有限責任公司，主要從事提供算力及數據處理服務。
- (5) 一家於2022年在中國註冊成立的有限責任公司，主要從事計算機、通信設備及其他電子設備的製造。

業 務

截至2025年12月31日止年度

序號	供應商	提供的 產品/服務	採購額	佔總採購額 的百分比	信貸條款	與我們開始業務 關係的年份
			(人民幣千元)			
1	供應商集團K ⁽¹⁾	設備	514,673.7	31.5	預付款於合約簽訂後10個工作日內支付；剩餘款項於驗收後5個工作日內支付	2024年
2	供應商L ⁽²⁾	AIDC租賃服務	293,593.9	18.0	預付款	2024年
3	供應商集團M ⁽³⁾	AIDC租賃服務	234,934.1	14.4	於每個結算期結束後7個工作日內支付	2024年
4	供應商N ⁽⁴⁾	AIDC租賃服務	210,652.3	12.9	於服務開始後按月後付費	2025年
5	供應商H	AIDC租賃服務	132,073.6	8.0	於發票開具後10個工作日內支付	2024年
總計			<u>1,385,927.6</u>	<u>84.8</u>		

附註：

- (1) 一家於2021年在中國註冊成立的有限責任公司，主要從事科技推廣及應用服務。
- (2) 一家於2023年在中國註冊成立的有限責任公司，主要從事在山東省運營大型科學儀器及設備共享平台。
- (3) 一家於1999年在中國註冊成立的有限責任公司，主要從事文化旅遊產業投資及運營平台的運營。
- (4) 一家於1997年在中國註冊成立並於聯交所及深圳證券交易所上市的股份有限公司，主要從事提供全面的信息及通信技術解決方案。

業 務

截至2026年6月30日止六個月

序號	供應商	提供的 產品/服務	採購額	佔總採購額 的百分比	信貸條款	與我們開始業務 關係的年份
			(人民幣千元)			
1	供應商O ⁽¹⁾	AIDC租賃服務	538,425.1	50.5	季度預付，各計費週期的服務費於收到服務費確認通知及發票後，於該計費週期首月的第五日前支付	2025年
2	供應商P ⁽²⁾	AIDC租賃服務	112,984.2	10.6	次季度支付，上一季度的算力使用費應於各季度的首月支付，並於收到相應的增值稅專用發票後七個工作日內支付	2025年
3	供應商Q ⁽³⁾	設備	56,837.9	5.3	36%於合約簽署後五個工作日內支付；22%於設備交付及交付回執簽署後30日內支付；42%於設備上架及驗收單簽署後60日內支付	2026年
4	供應商R ⁽⁴⁾	設備	43,621.0	4.1	30%於合約生效及收到發票後五日內支付；餘下70%於設備集群部署及上架完成並驗收，且收到發票後20日內支付	2023年
5	供應商S ⁽⁵⁾	設備	42,997.9	4.0	35%於合約簽署及收到發票後三日內支付；30%於設備交付及交付回執簽署時支付；餘下35%於設備集群部署及上架完成並驗收，且收到發票後60日內支付	2025年
總計			<u>794,866.1</u>	<u>74.5</u>		

業 務

附註：

- (1) 一家於2003年在中國註冊成立的股份有限公司，主要從事電信營運支持系統、垂直領域大模型的研發及企業智能化建設。
- (2) 一家於2002年在中國註冊成立的有限責任公司，主要從事大數據服務、互聯網數據服務及信息技術諮詢服務。
- (3) 一家於2023年在中國註冊成立的有限責任公司，主要從事技術推廣及應用服務。
- (4) 一家於2015年在中國註冊成立的有限責任公司，主要從事軟件開發及信息系統集成。
- (5) 一家於2018年在中國註冊成立的有限責任公司，主要從事電子設備銷售。

於往績記錄期間各年度／期間，我們的董事及其各自的聯繫人或持有5%或以上已發行股份總數的股東，概無於我們的任何五大供應商中擁有任何權益。

重大供應協議的主要條款

我們一般與供應商訂立框架協議或獨立採購訂單。我們與主要供應商訂立的協議的主要條款通常包括：

- **期限。**我們協議的期限因採購性質而異。對於硬件採購，協議的期限一般為約三至六個月。對於軟件採購，期限根據具體的業務需求確定，通常介乎一至五年。對於算力租賃服務，我們一般訂立為期五年的協議。
- **付款協議。**我們通常通過銀行轉賬與供應商結算。
- **價格。**價格於採購訂單中指定。
- **交付和驗收。**供應商負責將產品交付至我們指定的數據中心或倉庫。損失風險和所有權通常於我們經檢驗驗收產品後轉移予我們。
- **質保和維護。**供應商一般為硬件產品提供三至五年的標準質保期。我們要求供應商維持國內維修基地並提供備件，以確保對硬件故障做出快速反應。
- **產品退換。**我們有權以各種原因拒收、更換或退回產品，包括產品質量、產品規格或數量與所下訂單不符。
- **終止。**當供應商未能履約且於收到我們的書面通知後未能及時糾正時，我們有權終止採購訂單。

業 務

於往績記錄期間及直至最後實際可行日期，我們與主要供應商訂立的協議具有法律約束力，並無發生可能對我們的業務、財務狀況及經營業績產生不利影響的任何重大違約行為。

供應商篩選及管理

我們通常會委聘信譽良好的供應商以確保產品的質量。我們對每種產品的質量和性能實行嚴格控制。我們已對供應商的選擇和持續管理實施嚴格的多層面流程，以確保我們供應鏈的質量和可靠性。

- **供應商准入。**我們對新供應商遵循嚴格的准入程序。我們的採購部門對潛在供應商進行全面的盡職審查。我們根據候選人的資本安全性、歷史交付表現及維護能力對其進行評估。必要時，我們亦對彼等的倉庫和設施進行現場檢查。
- **表現管理。**我們對現有供應商實施全面的評估制度。我們根據質量、交付及時性、成本競爭力及服務支持等關鍵維度進行定期評估和評級。我們根據供應商的表現得分將其分為不同的等級。戰略合作夥伴在採購中獲賦予優先權，而表現不佳的供應商則須接受整改計劃或被淘汰。
- **供應鏈韌性。**為降低與供應鏈波動及地緣政治因素相關的風險，我們通過為關鍵部件委聘多家授權分銷商，積極實現供應商基礎的多樣化。我們亦維持核心部件的戰略庫存並探索替代解決方案，以確保業務持續性。

供應商集中度

於往績記錄期間，我們主要向我們的主要供應商採購設備及AIDC租賃服務。根據灼識諮詢的資料，供應商集中度符合行業慣例。我們已與主要供應商建立穩定的合作關係，以確保供應鏈的質量及可靠性以及營運的連續性。由於我們從主要供應商採購大部分設備及算力租賃服務，倘我們與主要供應商的關係終止、中斷或以任何對我們不利的方式變更，我們的營運及業務可能會中斷。更多詳情請參閱「風險因素－與我們的業務及行業相關的風險－我們面臨供應商集中風險」。

我們已實施多項措施以減輕與供應商集中度相關的潛在風險。就設備採購而言，我們已擴大供應商網絡以引入多家硬件提供商，並持續加強與現有設備供應商的合作關係。就AIDC租賃服務而言，我們從主要供應商採購的部分AIDC租賃服務來自我們先前交付智算基礎設施解決方案的客戶，根據有關安排，我們回租AIDC以產出算力。有關更多詳情，請參閱本節「我們的業務模式」及「客戶與供應商的重疊」。我們亦從自有的服

業 務

務器獲取算力。除與智算基礎設施解決方案客戶的安排及我們自有的計算資源外，我們根據合作安排管理第三方AIDC所產生的算力以獲取算力，並據此對租賃設施進行技術升級及標準化，從而維持多元化的算力池，並非獨家依賴任何單一來源。我們在兩大類別的主要供應商之間動態分配採購量，以避免過度依賴。

於2023年，我們的採購以設備為主，當時我們正處於智算基礎設施解決方案的早期階段，向五大供應商的採購額佔我們總採購額的47.1%。隨著我們的業務於2024年及2025年快速擴張，我們的採購日益包括大量算力租賃服務。向五大供應商的採購額於2024年及2025年分別增至佔我們總採購額的74.8%及84.8%。此日益增加的集中度反映我們業務規模的增長以及來自客戶的租賃安排。截至2026年6月30日止六個月，向我們五大供應商的採購額下降至佔我們總採購額的74.5%。此外，我們採購的設備及算力租賃服務均由多個市場參與者提供，市場上亦存在具有可比服務能力的替代供應商，這支持了我們維持供應連續性的能力。

於往績記錄期間，我們最大供應商的身份因年份而異，反映我們採購的演變性質。於2023年，供應商集團A（一家設備供應商）為我們的最大供應商，佔該年度總採購額的16.1%。於2024年，供應商F（一家算力租賃服務提供商）為我們的最大供應商，佔該年度總採購額的22.8%。於2025年，供應商集團K（一家設備供應商）為我們的最大供應商，佔該年度總採購額的31.5%。截至2026年6月30日止六個月，供應商O（一家AIDC租賃服務提供商）為我們的最大供應商，佔該期間總採購額的50.5%。我們選定供應商O時乃依據對其技術能力、交付表現及質量標準的全面評估。於往績記錄期間各年度/期間，最大的供應商並非我們的獨家供應商。我們與供應商O保持建設性的合作關係，且於往績記錄期間及直至最後實際可行日期，我們與供應商O並無發生任何重大糾紛。基於前述因素，我們的董事認為，我們與主要供應商的關係不太可能發生任何重大不利變動或被終止。

物流及存貨管理

物流

對於智算服務器和設備的採購，我們通常安排從供應商直接交付至我們或客戶指定的數據中心或項目地點。

存貨管理

我們一般採用訂單驅動的採購策略來最大限度地降低存貨風險。我們通常不會維持大量的服務器或計算設備存貨。採購主要由確認的客戶訂單或特定項目需求觸發。我們的存貨主要包括服務器及其他硬件設備。此外，我們通過智能管理平台持續監控計算資源的使用情況和狀態。這使得我們能夠優化資源配置和採購規劃，從而確保我們有足夠的能力來滿足客戶需求，而不會產生過高的存貨持有成本。

業 務

我們的客戶

於往績記錄期間，我們的客戶主要包括科技公司、研究機構和大學、地方政府及其部門、國有企業及民營企業。於2023年、2024年及2025年以及截至2026年6月30日止六個月，於往績記錄期間各年度／期間我們的五大客戶分別共同產生人民幣58.3百萬元、人民幣339.0百萬元、人民幣966.5百萬元及人民幣475.6百萬元的收入，分別佔我們總收入的54.8%、85.7%、88.0%及80.8%。於往績記錄期間各年度／期間我們的五大客戶均為獨立第三方。

下表載列於往績記錄期間各期我們五大客戶的詳細資料：

截至2023年12月31日止年度

序號	客戶	銷售的產品	收入	佔總收入的百分比	信貸條款	與我們開始業務關係的年份
			(人民幣千元)			
1	客戶A ⁽¹⁾	智算基礎設施解決方案	22,796.5	21.5	於合約簽訂後4個工作日內支付	2023年
2	客戶B ⁽²⁾	AIFS服務	13,003.6	12.2	大部分服務費於工作量確認後30個工作日內支付，剩餘服務費於最終驗收後30個工作日內支付	2020年
3	客戶C ⁽³⁾	AIFS服務	9,499.7	8.9	分四期付款：生效日期、驗收、確認及服務完成，每期款項於發票開具後60個工作日內支付	2019年
4	客戶集團D ⁽⁴⁾	AIFS服務	6,631.3	6.2	大部分款項於首次驗收後30個工作日內支付，剩餘款項於最終驗收後30個工作日內支付	2022年
5	客戶E ⁽⁵⁾	AIFS服務	6,330.2	6.0	分三期付款：合約簽訂、初步驗收及最終驗收，每期款項於15個工作日內支付	2023年
總計			<u>58,261.3</u>	<u>54.8</u>		

附註：

- (1) 一家於2021年在中國註冊成立的有限責任公司，主要從事技術推廣、服務及開發。
- (2) 一家於2015年在中國註冊成立的有限責任公司，主要從事數字金融。
- (3) 一家於1992年在中國註冊成立並於上海證券交易所上市的國有企業，主要從事商業銀行服務。

業 務

- (4) 一家於2004年在中國註冊成立並於聯交所上市的股份有限公司，主要從事移動電信服務。
- (5) 一家於2019年在中國註冊成立的有限責任公司，主要從事商用車智能自動駕駛技術及高級駕駛輔助系統的研發、生產及推廣。

截至2024年12月31日止年度

序號	客戶	銷售的產品	收入	佔總收入的百分比	信貸條款	與我們開始業務關係的年份
			(人民幣千元)			
1	客戶F ⁽¹⁾	智算基礎設施解決方案	172,331.5	43.6	就服務器而言，兩期款項分別於合約簽訂及驗收後15個工作日內支付；就服務而言，分四期付款：生效日期、初步驗收、最終驗收及質保期結束，每期款項於15個工作日內支付	2024年
2	客戶集團D	智算基礎設施解決方案、AIFS服務	83,896.0	21.2	分四期付款，各期款項須於15個工作日內支付，且均以預先收取相應終端用戶付款為先決條件：分別於累計交付、全部安裝及驗收以及3年質保期屆滿時支付	2022年
3	客戶G ⁽²⁾	智算基礎設施解決方案	36,708.0	9.3	於合約簽訂後1個工作日內支付	2024年
4	客戶H ⁽³⁾	智算基礎設施解決方案	30,619.5	7.7	於合約簽訂後4個工作日內支付	2024年
5	客戶C	AIFS服務	15,447.6	3.9	兩種付款模式：大部分款項於提供服務三個月後支付，剩餘款項於最終驗收後支付，兩者均於發票開具後60個曆日內支付；或分四期付款：生效日期、驗收、確認及服務完成，每期款項於發票開具後60個工作日內支付	2019年
總計			<u>339,002.6</u>	<u>85.7</u>		

附註：

- (1) 一家於2023年在中國註冊成立的有限責任公司，主要從事在山東省運營大型科學儀器及設備共享平台。
- (2) 一家於2012年在中國註冊成立的有限責任公司，主要從事互聯網信息技術服務。
- (3) 一家於1993年在中國註冊成立並於上海證券交易所上市的股份有限公司，主要從事數字產品生產、產業數字化及新型數字基礎設施。

業 務

截至2025年12月31日止年度

序號	客戶	銷售的產品	收入	佔總收入的百分比	信貸條款	與我們開始業務關係的年份
			(人民幣千元)			
1	客戶I ⁽¹⁾	智算基礎設施解決方案	527,441.9	48.0	75%預付款於合約簽訂後15個工作日內支付；10%於交付驗收後10個工作日內支付；5%於集群驗收後10個工作日內支付；餘款(不含質保金)於集群穩定運行滿3個月後，10個工作日內支付；質保金分4期等額按年支付，於每年穩定運行期滿後10個工作日內支付。	2025年
2	客戶集團J ⁽²⁾	智算基礎設施解決方案	230,892.1	21.0	分五期付款：硬件交付至倉庫、硬件驗收、初步驗收及點亮、最終驗收及質保期屆滿，每期款項於發票開具後10個工作日內支付	2024年
3	客戶集團K ⁽³⁾	智算基礎設施解決方案	121,067.4	11.0	兩種付款模式：預付款項於生效後15個工作日內支付，交付款項於5個工作日內支付，驗收款項於10個工作日內支付，及遞延付款項分五期按年支付，每期於各確認日期後10個工作日內支付；或於合約簽訂後3個工作日後悉數支付	2025年
4	客戶L ⁽⁴⁾	智算基礎設施解決方案	66,590.8	6.1	30%於合約簽訂後10個工作日內支付，60%於設備交付並簽署交付驗收單後10個工作日內支付，10%於項目驗收後20個工作日內支付	2025年
5	客戶集團M ⁽⁵⁾	智算基礎設施解決方案	20,470.9	1.9	設備款項分三期支付：30%於合約簽訂後10個工作日內支付，20%於全部交付時支付，50%於質量驗收後5個工作日內支付；集成、運行及維護服務費分5期等額年度支付，首期款項應於質量驗收滿1個月後5個工作日內支付，其後每年按期支付	2025年
總計			<u>966,463.1</u>	<u>88.0</u>		

附註：

- (1) 一家於2001年在中國註冊成立的有限責任公司，主要從事廣播電視傳輸設備的製造及廣播電視節目的運營。
- (2) 一家於1999年在中國註冊成立的有限責任公司，主要從事文化旅遊產業投資及運營平台的運營。
- (3) 一家於2021年在中國註冊成立的有限責任公司，主要從事科技推廣及應用服務。

業 務

- (4) 一家於2025年在中國註冊成立的有限責任公司，主要從事互聯網相關服務。
- (5) 一家於2003年在中國註冊成立並於深圳證券交易所上市的股份有限公司，主要從事柔性電子技術的研發及應用。

截至2026年6月30日止六個月

序號	客戶	銷售的產品	收入	佔總收入的百分比	信貸條款	與我們開始業務關係的年份
			(人民幣千元)			
1	客戶N ⁽¹⁾	AIDC建設	172,741.7	29.4	20%於合約簽署及收到發票後五個工作日內支付；53%於設備交付及驗收後滿180日後，並於收到發票後五個工作日內支付；餘下27%於系統驗收後滿180日後，並於收到發票後五個工作日內支付	2026年
2	客戶O ⁽²⁾	AIDC建設	153,451.7	26.4	30%於合約簽署後三個工作日內支付；60%於設備交付及驗收後三個工作日內支付；餘下10%於系統驗收後三個工作日內支付	2026年
3	客戶P	AIDC建設	105,852.6	18.0	兩種付款安排：(i) 30%於合約簽署及收到發票後七個工作日內支付，60%於設備交付及驗收後七個工作日內支付，餘下10%於最終驗收及簽署驗收報告後七個工作日內支付；或(ii) 40%於系統驗收及收到發票後10個工作日內支付，餘下60%分三期按年支付，每期20%，於驗收及收到發票後每滿12個月後10個工作日內支付	2025年
4	客戶Q ⁽³⁾	智算雲	21,315.6	3.6	次月支付，服務費於各月服務期結束後及收到發票及結算單後三個工作日內支付	2026年
5	客戶R ⁽⁴⁾	智算雲	20,248.8	3.4	履約保證金及平台服務費於服務開始後10個工作日內支付；履約保證金用於抵扣第二及第三個結算期的服務費；首個結算期的服務費於該期間結束及收到發票後15個工作日內支付	2026年
總計			475,610.4	80.8		

業 務

附註：

- (1) 一家於2022年在中國註冊成立的有限責任公司，主要從事技術推廣、信息技術諮詢服務及新材料技術推廣服務。
- (2) 一家於2015年在中國註冊成立的有限責任公司，主要從事計算機軟件、安全產品及電子產品的研發及銷售。
- (3) 一家於2018年在中國註冊成立的有限責任公司，主要從事工程技術研發、信息系統集成服務及電信業務營運。
- (4) 一家於2023年在中國註冊成立的有限責任公司，主要從事技術推廣及應用服務。

於往績記錄期間各年度／期間，我們的董事及其各自的聯繫人或持有5%或以上已發行股份總數的股東，概無於我們的任何五大客戶中擁有任何權益。

重大服務協議的主要條款

我們一般與客戶訂立框架協議或服務訂單。我們與主要客戶的協議的主要條款通常包括：

- **期限。**我們與客戶的協議通常為長期框架協議。
- **服務範圍。**協議規定有計算服務的類型、資源的配置以及任何額外的運行或維護服務。
- **付款協議。**對於標準計算服務，我們一般要求客戶預付服務費。對於大型建設或集成項目，付款通常根據項目里程碑分期進行。我們通常通過銀行轉賬與客戶結算。
- **知識產權。**我們保留對在與客戶接觸之前開發的知識產權的所有權和使用權。當該等知識產權應用於向客戶提供的產品或解決方案時，客戶被授予對該等知識產權的使用權。
- **風險轉移。**風險於客戶完成檢驗並確認收到解決方案後轉移予客戶。
- **終止。**協議可因重大違約而終止，或在若干情況下，經事先通知並結算未付費後無因終止。

於往績記錄期間及直至最後實際可行日期，我們與五大客戶訂立的協議具有法律約束力，且未曾發生任何可能對我們的業務、財務狀況及經營業績產生不利影響的任何重大違約行為。

業 務

客戶集中度

於往績記錄期間，各年度／期間我們五大客戶合計產生的收入分別佔我們2023年、2024年、2025年及截至2026年6月30日止六個月總收入的54.8%、85.7%、88.0%及80.8%。於往績記錄期間，客戶集中程度相對較高且持續上升，主要由於我們的智算基礎設施解決方案業務於往績記錄期間處於相對早期及快速增長階段。大型項目本身價值較高，每個項目單獨來看均能帶來可觀的收入，導致其收入結構本身就具有高度集中性。詳情請參閱「風險因素－與我們的業務及行業相關的風險－於往績記錄期間，我們依賴有限數量的客戶，任何主要客戶的流失均可能對我們的業務、經營業績及財務狀況產生不利影響」。

我們相信，隨著業務的不斷成熟及多元化發展，我們的客戶集中度將隨著時間推移而降低。隨著我們持續擴大我們的智算雲服務的規模及覆蓋範圍，該服務向各行各業的廣泛客戶群提供經常性、按需使用的計算資源，我們預期收入基礎將會擴大。此外，隨著我們於智算基礎設施解決方案方面的聲譽及往績不斷提升，我們預期將服務更多來自不同行業的客戶，從而進一步降低集中度。此外，於往績記錄期間，我們最大客戶於各年度／期間產生的收入分別佔我們2023年、2024年、2025年及截至2026年6月30日止六個月總收入的21.5%、43.6%、48.0%及29.4%。我們最大客戶的身份於各年度／期間均有所不同，反映出我們的智算基礎設施解決方案以項目為基礎的性質。根據灼識諮詢的資料，此與行業慣例一致。基於上述情況，董事認為，我們的客戶集中度反映了智算基礎設施解決方案的項目驅動性質，以及我們的業務轉型及發展進程，並預期將隨著我們的智算雲服務規模擴大及客戶群多元化而逐步下降。我們認為客戶關係發生重大不利變動的風險有限。

客戶與供應商的重疊

於往績記錄期間，我們存在以下主要客戶與供應商重疊的情況。通常情況下，我們向該等客戶／供應商提供智算基礎設施解決方案，並與彼等訂立租賃安排，以向其租賃AIDC。根據該等安排，我們向客戶租用AIDC的指定部分，承擔該租用部分內服務器及其他設備的營運管理責任以產生算力，而我們的客戶則保留AIDC的其餘部分供其自行使用。該等安排具有戰略互補性，讓我們能夠確保長期獲取原生運行Alaya NeW OS的算力，同時使我們的客戶（其中許多客戶將AIDC開發為區域計算基礎設施）能夠提升其設施的使用率。有關更多詳情，請參閱「－智算雲服務－我們的算力池」。該安排符合我們的業務戰略，即通過智算基礎設施解決方案項目以確保獲得對我們營運至關重要的算力能力，並體現我們創收活動與採購需求之間的商業互補關係。根據灼識諮詢的資料，該安排符合行業慣例。

業 務

客戶集團D為我們於往績記錄期間2023年及2024年的五大客戶之一。於往績記錄期間，我們主要向客戶集團D銷售智算基礎設施解決方案及AIFS服務，並向其採購寬帶服務。於2023年、2024年、2025年及截至2026年6月30日止六個月，客戶集團D貢獻(i)收入分別為人民幣6.6百萬元、人民幣83.9百萬元、人民幣6.5百萬元及零，分別佔我們同期總收入的6.2%、21.2%、0.6%及零；及(ii)總採購金額分別為人民幣39.1千元、人民幣39.1千元、零及人民幣0.2百萬元，分別佔我們同期總採購金額的0.02%、0.01%、零及0.02%。

於往績記錄期間，客戶F/供應商L為我們於2024年的五大客戶之一及於2025年的五大供應商之一。於往績記錄期間，我們主要向客戶F/供應商L銷售智算基礎設施解決方案，並自其採購AIDC租賃服務。於2023年、2024年及2025年以及截至2026年6月30日止六個月，客戶F/供應商L貢獻(i)收入分別為零、人民幣172.3百萬元、零及零，分別佔我們同期總收入的零、43.6%、零及零；及(ii)總採購金額分別為零、零、人民幣293.6百萬元及人民幣0.2百萬元，佔我們同期總採購金額的零、零、18.0%及0.01%。

於往績記錄期間，客戶I/供應商F為我們於2024年的五大供應商之一及於2025年的五大客戶之一。於往績記錄期間，我們主要向客戶I/供應商F銷售智算基礎設施解決方案，並自其採購AIDC租賃服務。於2023年、2024年及2025年以及截至2026年6月30日止六個月，客戶I/供應商F貢獻(i)收入分別為零、零、人民幣527.4百萬元及零，分別佔我們同期總收入的零、零、48.0%及零；及(ii)總採購金額分別為零、人民幣178.3百萬元、零及人民幣6.8百萬元，佔我們同期總採購金額的零、22.8%、零及0.6%。

於往績記錄期間，客戶集團J/供應商集團M為我們於2025年的五大客戶之一及五大供應商之一。於往績記錄期間，我們主要向客戶集團J/供應商集團M銷售智算基礎設施解決方案，並自其採購AIDC租賃服務。於2023年、2024年及2025年以及截至2026年6月30日止六個月，客戶集團J/供應商集團M貢獻(i)收入分別為零、零、人民幣230.9百萬元及零，分別佔我們同期總收入的零、零、21.0%及零；及(ii)總採購金額分別為零、零、人民幣234.9百萬元及人民幣21.0百萬元，佔我們同期總採購金額的零、零、14.4%及2.0%。

於往績記錄期間，客戶集團K/供應商集團K為我們於2025年的五大客戶之一及五大供應商之一。於往績記錄期間，我們主要向客戶集團K/供應商集團K銷售智算基礎設施解決方案，並自其採購服務器。於2023年、2024年及2025年以及截至2026年6月30日止六個月，客戶集團K/供應商集團K貢獻(i)收入分別為零、零、人民幣121.1百萬元及零，分別佔我們同期總收入的零、零、11.0%及零；及(ii)總採購金額分別為零、人民幣1.2百萬元、人民幣514.7百萬元及零，佔我們同期總採購金額的零、0.2%、31.5%及零。

業 務

於往績記錄期間，客戶集團M／供應商集團A為我們於2025年的五大客戶之一及於2023年的五大供應商之一。於往績記錄期間，我們主要向客戶集團M／供應商集團A銷售智算基礎設施解決方案，並自其採購服務器。於2023年、2024年及2025年以及截至2026年6月30日止六個月，客戶集團M／供應商集團A貢獻(i)收入分別為零、零、人民幣20.5百萬元及零，分別佔我們同期總收入的零、零、1.9%及零；及(ii)總採購金額分別為人民幣33.5百萬元、人民幣0.2百萬元、零及零，佔我們同期總採購金額的16.1%、0.02%、零及零。

於往績記錄期間，供應商H為我們於2024年及2025年的五大供應商之一。於往績記錄期間，我們主要向供應商H採購AIDC租賃服務，亦向其銷售智算雲服務。於2023年、2024年及2025年以及截至2026年6月30日止六個月，供應商H貢獻(i)收入分別為零、零、人民幣0.6百萬元及人民幣1.3百萬元，分別佔我們同期總收入的零、零、0.1%及0.2%；及(ii)總採購金額分別為零、人民幣105.9百萬元、人民幣132.1百萬元及人民幣13.1百萬元，佔我們同期總採購金額的零、13.6%、8.0%及1.2%。

於往績記錄期間，客戶P／供應商P為我們截至2026年6月30日止六個月的五大客戶及供應商之一。於往績記錄期間，我們主要向客戶P／供應商P銷售服務器、系統軟件及集成服務，並自其採購AIDC租賃服務。於2023年、2024年及2025年以及截至2026年6月30日止六個月，客戶P／供應商P貢獻(i)收入分別為零、零、零及人民幣105.9百萬元，分別佔我們同期總收入的零、零、零及18.0%；及(ii)總採購金額分別為零、零、零及人民幣113.0百萬元，佔我們同期總採購金額的零、零、零及10.6%。

於往績記錄期間，我們就與重疊客戶與供應商相關的銷售及採購協議條款的談判乃按個案獨立進行，且彼此之間不存在互為條件的關係。於往績記錄期間，不存在捆綁銷售或背對背交易安排，且該等重疊客戶與供應商的關鍵條款通常與我們其他客戶及供應商的條款並無重大差異。根據灼識諮詢的資料，該等重疊關係符合行業慣例。我們的董事認為，該等安排屬於我們在日常業務過程中按正常商業條款進行的交易。

質量管理

我們致力於為客戶提供穩定且可靠及的計算服務。鑒於我們的業務性質，我們的質量管理體系側重於雲服務的穩定性和底層硬件基礎設施的可靠性。

服務質量及服務水平協議

我們已建立全面的服務質量保證體系。對於我們的雲服務，我們與客戶遵守嚴格的服務水平協議（「SLA」），承諾高標準的可用性和正常運行時間。我們已在平台內實施一個自動化測試和監控框架，使我們能夠持續監控計算資源的健康狀態，實時探測異常，並自動調度資源以盡量減少服務中斷。

業 務

供應鏈質量控制

由於我們的算力依賴自第三方供應商採購的硬件，因此我們對供應鏈的質量實施嚴格控制。我們根據嚴格的標準選擇供應商，包括其資金安全、歷史交付記錄及維護能力。我們要求供應商維持國內維修中心和備件存貨，以確保對硬件故障做出快速反應。我們的供應商有合同義務為有缺陷的硬件提供現場維護和更換服務，確保任何硬件問題得到及時解決，而不會對我們的服務交付產生重大影響。

客戶投訴與退款政策

我們已建立客戶投訴處理程序，以確保及時解決任何服務問題。對於我們的在線訂閱服務，一旦提供服務，我們通常不會進行退款，除非出現重大服務故障或法律另有要求。然而，我們有退款政策，在某些情況下，經我們批准，客戶可終止其訂閱並獲得預付費用中未使用部分的退款。於往績記錄期間及直至最後實際可行日期，我們並未收到對我們的業務、運營或財務狀況產生重大不利影響的與服務質量有關的任何重大訴訟。截至同日，我們並無遭遇任何重大產品召回或客戶投訴。

競爭

AIDC已成為AI發展與商業化的先決條件。市場需要能夠提供涵蓋規劃、營運及開發在內的全生命週期整合服務的專業智算服務商。根據灼識諮詢的資料，中國智算服務市場的市場規模，按收入計，於2025年達約人民幣684億元，預期於2030年將超過人民幣2,775億元，2025年至2030年的年複合增長率為32.3%。

中國智算服務市場正處於快速發展階段，其特點為參與者多樣化及競爭格局動態變化。按2025年中國智算服務收入計，本公司在中國智算服務商中排名前十。按2025年中國智算服務收入計，本公司為中國最大的智算新雲。更多詳情請參閱「行業概覽－中國智算服務市場的競爭格局」。

員工

截至2026年6月30日，我們擁有303名全職員工。下表載列按職能劃分的我們的員工人數：

員工職能	員工人數	佔總數百分比(%)
研發	156	51.5
銷售及營銷	80	26.4
行政	67	22.1
總計	303	100

業 務

為精簡人力資源管理，我們已制定一套全面的內部管理措施，概述了招聘、培訓、內部推薦等程序和標準。有關更多詳情，請參閱「—環境、社會及治理—社會責任—員工管理」。

根據中國法律及法規的要求，我們參加由適用的地方及省級政府組織的各種員工社會保障計劃，包括住房、養老、醫療、工傷、生育及失業福利計劃，根據該等計劃，我們按員工薪金的特定百分比做出供款。於往績記錄期間，曾出現我們未有為若干員工足額繳存社會保險或住房公積金的情況。有關更多詳情，請參閱「風險因素—未能根據法規規定向各項僱員福利計劃作出足夠供款可能使我們面臨處罰」。

我們認為，我們與員工保持著良好的工作關係，於往績記錄期間及直至最後實際可行日期，我們並無經歷任何重大勞資糾紛或在為業務運營招聘員工方面遭遇任何困難。

環境、社會及治理

概覽

我們將環境、社會及治理（「**ESG**」）考量融入我們的戰略規劃、技術研發、運營及商業決策全過程，在持續提升智算服務的核心技術能力和商業價值的同時，積極回應員工、客戶、業務夥伴及社會各界的期望。

作為一家智算新雲，我們高度重視我們業務的可持續發展。我們的創始人兼董事會主席方磊先生入選2025年福布斯中國ESG影響力人物榜單，以表彰其在推動綠色計算基礎設施建設、促進普惠算力及踐行全球化ESG標準方面的引領性貢獻。在其帶領下，我們在日常營運中積極推動普惠算力服務，提升資源利用效率，並持續優化我們的營運，以減少能源消耗及對環境的影響。

ESG 治理

ESG 組織架構

我們將ESG風險融入我們的整體風險管理與內部控制框架。為進一步提升ESG治理水平，本公司計劃建立由董事會牽頭、管理層統籌推進、全體員工共同參與的ESG治理體系。

董事會全面負責ESG機制與政策的制定，核心職責包括：評估關鍵ESG議題優先級、審批ESG戰略、監督執行情況，並在重大戰略決策中充分納入ESG考量，同時每年一次審議並審批年度ESG報告。高管層負責建立ESG委員會以統籌事務執行並定期向董事會匯報而員工執行層則按ESG三大維度設立協調小組，負責日常落地工作。

業 務

風險及合規管控

本公司正積極識別與ESG相關的重大議題及短期、中期和長期的潛在ESG風險，評估各項風險發生的可能性與潛在影響。作為一家智算新雲，本公司不運營任何生產設施，而是主要在辦公室、智算中心和線上開展業務。因此，本公司在日常運營中不會面臨重大的健康、安全或環境風險。

潛在物理與氣候風險

強降水、極熱天氣、颶風等極端天氣發生頻率增加，對我們的運營設施(如辦公樓、智算中心等)的穩定性造成影響。同時，極端天氣可能影響員工通勤與商務出行。氣候的變化將影響公司決策及運營。本公司將持續監察環境風險，以實現業務的長期可持續發展。

監管與聲譽風險

隨著相關監管政策及行業規範逐步完善，未來可能對智算服務市場產生一定影響。與此同時，客戶及利益相關方對環境因素的關注日益增強，在選擇合作夥伴時，愈發傾向於將綠色低碳作為重要考量標準。為此，公司將密切關注監管動態與市場環境變化。

社會與人才風險

本公司高度重視人才管理及勞工權益保障，嚴格遵守適用的勞工法律及法規並及時響應監管更新，通過建立完善的員工權益保障與職業發展機制，減少用工風險、人才流失及社會輿論壓力，保障企業品牌形象和運營穩定性。同時公司注重職業健康與心理健康建設，通過建立相關機制，力求預防職業健康或壓力相關問題，提升員工滿意度與忠誠度。

合規與商業道德

我們在業務運營過程中，嚴格遵守適用的數據安全、隱私保護、反壟斷及反不正當競爭等相關法律法規，並通過制度建設、內控流程、培訓宣導及監督機制，推動合規文化在本公司內部的有效落地。我們亦致力於與同樣重視誠信經營、數據與隱私保護以及可持續行業發展的合作夥伴開展合作。於往績記錄期間，並無對公司或員工提出的已審結的貪污訴訟案件。

業 務

環境責任

我們嚴格遵守《中華人民共和國環境保護法》《中華人民共和國水污染防治法》《中華人民共和國固體廢物污染環境防治法》等在我們所有經營所在地的適用環境法律法規。於往績記錄期間，本公司未發生因環境違規而受到行政處罰或引發負面輿論的情況。

資源使用

我們嚴格遵守《中華人民共和國節約能源法》《中華人民共和國水法》等相關法律法規，建立並執行內部節能管理制度，規範能源與水資源使用流程，不斷提升資源利用效率。本公司在日常運營中積極踐行節能減排理念，通過多項措施降低資源消耗並減少環境影響。倘可行及適用，公司在辦公區域設置節約用紙、節約用電等提示標識，倡導員工在日常工作中養成節約資源的意識及良好習慣。公司將根據業務發展情況持續關注節能減排相關管理措施，並不斷完善相關管理機制：

- 於往績記錄期間，我們大部分的電力消耗來自智算中心的運營消耗、以及辦公照明和空調系統。我們持續開展節能宣傳，包括但不限於在辦公區域張貼節能倡導標語、在員工手冊中強調節能等舉措，從而引導員工在日常辦公中養成人走燈滅、設備閒置及時斷電的良好習慣。我們已設定明確的節能目標，每年將用電量減少約5%。
- 於往績記錄期間，本公司用水主要集中在智算中心運營以及日常辦公環節。我們通過行政宣傳，張貼節水倡導標語等方式持續提升員工的節水意識。我們倡導員工養成隨手關閉水龍頭、防止滴漏的良好用水習慣。我們已設定明確的節水目標，並正在實施及持續監察各項措施，目標是每年減少約5%的用水量。
- 於往績記錄期間，我們資源消耗的情況如下：

指 標	單 位	截至12月31日止年度		
		2023年	2024年	2025年
耗電量	MWh	150.1	1,093.3	14,167.2
耗水量	噸	1,102	1,783	1,937

業 務

於往績記錄期間，受智算中心業務擴張及其高能耗特徵影響，本公司的總耗電量呈現上升趨勢。本公司在用電規劃上將會全面推行精細化能源管理，未來將通過引入高效供配電、持續提升綠電使用比例，以逐步降低用電密度。水資源管理方面，儘管業務擴張致使總耗水量有所增加，但得益於規模經濟效應及節水措施的有效落實，本公司已實現耗水量同比增長率逐漸趨緩。本公司將通過完善用水政策，推進節水改造及提升水資源利用效率，在業務擴張的同時進一步降低耗水量。

智算中心能源管理

我們致力於將綠色低碳原則貫穿智算中心規劃、建設與運營全生命週期。截至2026年6月30日止六個月，我們在運營的智算中心平均算力利用率已超過79.6%，而2026年6月的月度利用率達到89.1%。我們的能效比(PUE)降至1.26，近60%的機房達到2級能效等級，有兩間機房達到1級能效等級。截至2026年6月30日，我們其中一間機房已開始使用可再生能源發電，且我們將逐步在各智算中心擴大可再生能源的應用範圍，同時持續優化我們的能源結構及提高能源效率，旨在以負責任的方式為數字經濟提供更清潔、更高效的計算基礎設施。

於2026年，本公司通過於能源協調、綠電利用及低碳數據中心開發方面與行業合作夥伴開展合作，進一步推進了綠色計算舉措，重點提升了智算運營過程中能源消耗的效率及可追溯性。

應對氣候變化

我們持續將氣候應對措施融入經營管理。氣候變化導致的極端天氣可能擾亂供應鏈、影響生產並增加成本，為此我們將建立定期監測與氣象預警機制，降低物理風險。此外，本公司在數據中心運營中積極實施節能實踐，持續優化運營效率，以達到有效減少碳排放。

廢棄物管理

我們高度重視廢棄物管理，致力於通過系統化的管理流程減少廢棄物產生量，從而降低對環境的影響。於往績記錄期間，我們在日常運營中不涉及大量的有害固體廢物排放，我們所產生的固體廢棄物主要為日常行政辦公產生的無害廢棄物。2023年至2025年，我們所產生的無害固體廢物產生量分別為0.46噸、0.48噸及0.71噸，無害固體廢物強度由每億元人民幣收入0.43噸下降至0.12噸及0.06噸。公司業務運營過程中不涉及廢氣或污染物排放，為進一步保護環境，我們制定了廢棄物100%合規處置的目標。

業 務

溫室氣體排放

我們正持續完善溫室氣體（「溫室氣體」）監督與管控機制。我們將繼續推行減排行動方案，包括但不限於使用綠電智算中心、推廣新能源運輸車輛的應用、提高能源使用效率、選用更環保的包裝材料、建立監測與評估機制。我們設立了碳排放減排目標，計劃在2028年前將溫室氣體排放減少約10%。

於往績記錄期間，我們的溫室氣體排放歷史量化數據如下：

指標	單位	截至12月31日止年度		
		2023年	2024年	2025年
溫室氣體排放量(範圍1) ⁽¹⁾	噸二氧化碳當量	不適用	不適用	不適用
溫室氣體排放量(範圍2) ⁽²⁾	噸二氧化碳當量	105.1	765.3	9,917.1
溫室氣體排放量(範圍3) ⁽³⁾	噸二氧化碳當量	2.7	3.0	4.2

附註：

- (1) 由於我們的運營活動不涉及化石燃料的直接使用，因此於往績記錄期間未產生範圍1直接溫室氣體排放。
- (2) 溫室氣體排放計算範圍(範圍2)包括我們所購電力產生的排放。
- (3) 溫室氣體排放計算範圍(範圍3)包括廢紙的處理以及處理淡水和污水過程中所消耗的電力對應的排放。

於往績記錄期間，我們的營運不涉及直接燃燒化石燃料，本公司的核心碳排放來自於外購電力折算的間接排放。隨著業務規模的擴大，於往績記錄期間，整體外購電量的增加促使間接排放的溫室氣體上升。展望未來，我們將會持續控制能耗的同時，通過建立完善的碳排放計量機制、參與市場化綠電交易等方式，系統性地降低因業務增長帶來的溫室氣體排放壓力。

綠色產品與責任

作為普惠算力的行業領導者之一，本公司致力於在踐行自身運營綠色低碳轉型的同時，將普惠算力與綠色產品責任理念融入業務，為社會提供更先進、智能且環境友好的綠色算力服務。公司通過Alaya NeW Cloud，創新構建低碳算力調度網絡，通過智能調度綠電資源，實現了清潔能源的高效轉化與穩定供給，為行業提供了可驗證的綠色算力解決方案。本公司亦尋求通過行業合作推動綠色計算，包括與能源及環境合作夥伴開展合作，以促進綠電利用、能源計算協調及低碳計算基礎設施的建設。

業 務

作為綠色算力的先驅，我們已在智算中心廣泛部署節能設備，並透過技術創新系統性地提升算力資源的使用效率。我們的Serverless+RL技術架構能夠實現算力資源池化與智能調度，向下管理硬件資源的協同，從基礎設施層面降低能耗與碳足跡。同時，我們通過以算力計量單位(DCU)為核心的普惠算力服務模式，通過結合按需彈性供給機制，從服務層面有效促進社會算力資源的集約利用，降低閑置與浪費風險，體現綠色產品與服務責任。

社會責任

社區貢獻

我們堅信，算力鴻溝是制約實現人工智能普惠發展的關鍵。為此，我們秉持「讓算力回歸生產工具屬性」的理念並致力於推動實現「人人易得、用得起」的普惠算力目標。本公司亦尋求通過基於詞元的AI服務，使AI能力更易於獲取，使得客戶及開發者能夠以靈活、按使用量付費的方式使用智算及模型能力。此方法有助於降低採用AI相關的預付費用及技術門檻，並支持更廣泛地參與AI創新。我們構建並運營的開發者生態，不僅是技術交流與協作的樞紐，更是將「軟件定義算力」能力與「按需計量、按量計費」的彈性服務模式向廣泛創新者延伸的重要渠道。於往績記錄期間，我們有效降低了千行百業獲取和利用計算的門檻與技術成本，使算力真正成為可靈活調度的數字化生產工具。

本公司持續深化產學研深度融合與知識共享，深度參與WAIC 2026、WAIC 2025及VALSE 2025等頂尖學術與行業盛會，並與多家頂尖高校達成定製化智算合作，推動前沿AI技術普惠與人才培養。此外，我們通過與大學、研究機構及學術組織開展合作，積極推動普惠算力於科學研究中的應用。例如，我們通過向研究人員提供易於獲取的智算資源及AI驅動的研究能力，支持了AI for Science計劃，有助於降低科學計算的門檻及培育一個開放協作的研究生態系統。社區持續的反饋與共創，亦反向驅動著我們自研的智算操作系統及服務模式持續優化，從而為跨越算力鴻溝、激發產業新動能構築了堅實的技術與生態基礎。我們堅信，與行業夥伴及社區開發者合作共同推動普惠算力的落地，是履行企業社會責任、促進數字經濟增長的重要路徑。

知識產權保護

知識產權是我們的核心資產之一。本公司成立了知識產權部，負責公司知識產權申請、管理，建立了系統的創新激勵和保護體系，包括《知識產權管理辦法》、《九章雲極知識產權及相關技術成果獎勵辦法》、《九章雲極商標管理辦法》、《九章雲極軟件著作權登

業 務

記管理辦法》等，從而打造適合本公司的專屬的專利組合。有關我們所取得知識產權的更多詳情，請參閱「一知識產權」。未來公司將持續在智算服務領域構築技術壁壘，強化知識產權保護能力。

信息安全治理及數據保護

本公司已設立專門的信息安全管理部門，以應對日益複雜的網絡安全環境及監管要求，並保障公司業務的穩健運行。對於公司內部，本公司已制定《網絡安全管理制度》、《網絡與信息安全事件應急預案》及《信息安全管理規範》，形成較為完善的內部管理要求。在數據安全與隱私保護方面，我們不僅通過協議明確責任，更以技術踐行承諾，構建嚴密的多維防護體系。本公司通過簽署具體的數據保護協議從制度層面明確雙方的權責與隱私邊界，持續通過定期開展的第三方安全審計與內部合規檢查，嚴密監控並評估現有數據安全措施的有效性。

此外，我們在技術與運營管理上不斷深化信息安全防護建設，已制定完善的網絡安全應急預案。Alaya NeW智算運營管理平台已順利通過國家《信息系統安全等級保護3級》認證，同時公司層面已獲得ISO/IEC 27001信息安全管理體系認證。除了加強內部信息安全治理外，我們亦透過安全監控、訪問控制、數據保護及治理以及營運風險管理等措施，持續提升智算服務的安全性。我們致力於提供安全、穩定且可靠的智算服務，同時在服務生命週期內保障客戶信息及數據的機密性、完整性及可用性。

雲服務質量與責任

作為一家智算新雲，我們將產品安全、數據責任與創新保護深度融入運營，在計算機應用軟件設計開發及技術服務領域取得ISO 9001質量管理體系認證。此外，我們已獲得ISO/IEC 20000信息技術服務管理體系認證，該認證透過標準化服務管理流程及持續營運改善，有助於我們提供穩定、可靠且高質量的智算服務。於往績記錄期間，得益於定期的安全審核、系統巡檢與嚴格的質量控制流程，我們的平台未發生因數據及隱私安全或其他系統漏洞缺陷導致的客戶投訴。我們高度重視客戶反饋，所有通過技術服務協議渠道收到的產品與技術服務相關諮詢與改進建議均已得到及時響應與閉環處理。

業 務

供應鏈管理

我們充分認識到供應鏈對產品和服務質量的關鍵重要性，我們將供應商視為重要的合作夥伴，共同為客戶提供創新、高效的智算服務。針對供應商選擇，我們已建立明確的產品質量和環境與社會風險篩選標準。我們要求供應商提供質量管理體系與環境管理體系證書、產品相關檢測報告、相關行業案例和經驗、及其他資質證明文件。

我們優先選擇在質量、交付、成本及售後服務方面綜合表現優異的供應商，並將環境與社會風險評估作為核心篩選標準。供應商須具備環境合規相關的合法合規資質。我們優先考慮已獲得綠色認證、具備完善環保設施且資源利用效率較高的企業。在社會責任方面，供應商需嚴格遵守有關勞工常規、工作場所安全及職業健康的適用法律及法規，保障員工合法權益。對存在重大環境違規、安全事故或不良信用記錄的企業，我們將審慎評估，暫不納入合作名錄。未來，我們將持續完善供應商ESG評估機制，積極推動供應鏈向低碳化與可持續發展轉型。

員工管理

我們致力於營造公平、尊重且包容的工作環境，並嚴格遵守中華人民共和國各項勞動法律法規，包括《中華人民共和國勞動法》《中華人民共和國勞動合同法》《中華人民共和國未成年人保護法》及《禁止使用童工規定》。我們禁止基於任何受法律保護特徵的歧視、騷擾、霸凌、未經授權披露保密信息等行為。

作為一家有社會責任感的智算新雲，我們積極推動包容性就業，為女性、殘疾人士等群體創造平等就業機會。公司主動面向殘疾人開放適配崗位，改造辦公環境、提供輔助設備與個性化支持。我們高度重視女性員工的成長與發展，努力消除職場隱性壁壘，營造公平溫暖的工作氛圍。我們堅信，多元化的人才隊伍是公司創新活力的重要源泉。截至2026年6月30日，本公司共有303名員工，其中女性員工在員工總數中佔有相當比例，且有殘疾員工4名。本公司根據適用的法律及法規保障女性員工在孕期、產假等階段的合法權益，並堅決禁止任何形式的性騷擾與職場虐待。我們將持續在公司的各個層級推動性別多元化與機會平等，同時力爭殘疾員工人數每年保持淨增加以確保公司踐行多元化、公平性與包容性承諾。於往績記錄期間，公司並無發生任何童工、強迫勞動或歧視事件。

業 務

我們高度重視每位員工的貢獻，並為其提供公平的薪酬和職業發展機會。本公司建立了多層次的人才培養與晉升機制。通過設立管培生項目，面向具備潛力的應屆畢業生及早期職業人，加速其成長為具備綜合管理能力的核心成員。為了打通多元化晉升通道，公司特別設立了「開明訓練營」，通過提供專業的系統性指導與培訓，助力高潛力員工快速成長與能力躍升，從而為本公司的長遠發展持續注入新的活力與動能。員工亦可以得到多元化的學習與成長機會，使其能夠提升專業技能與知識水平，推動其在本公司內的長期職業發展。

職業健康及安全 (OHS)

我們嚴格遵守《中華人民共和國安全生產法》《中華人民共和國消防法》及其他適用的相關法律法規，已建立完善的內部管理制度，並且取得ISO14001環境管理體系認證和ISO45001職業健康安全管理体系認證。於往績記錄期間，本公司在所有重大方面均遵守中國有關健康與職業安全的適用法律法規，運營過程中未發生經裁定的工傷事故或死亡事件。

物業

我們的主要行政辦公室分別位於中國北京。根據《公司(豁免公司及招股章程遵從條文)公告》第6(2)條，本文件獲豁免遵守《公司(清盤及雜項條文)條例》第342(1)(b)條有關《公司(清盤及雜項條文)條例》附表三第34(2)段的規定(該條款要求就我們於土地或建築物中的所有權益提供估值報告)，原因是截至最後實際可行日期，我們租賃的任何物業的賬面金額均未超過我們合併資產總額的15%。

截至最後實際可行日期，我們主要於中國租賃8處物業，總建築面積為4,729.38平方米，主要作為我們的辦公室。我們認為中國內地有足夠的物業供應，且我們的業務運營並不依賴現有的租賃。我們認為，我們目前的設施足以滿足我們目前的需求。

於往績記錄期間及直至最後實際可行日期，我們並未收到有關部門要求我們完成某些租賃物業租賃登記備案的任何通知或行政處罰決定。我們的中國法律顧問表示，未進行租賃登記備案不會影響租賃協議的有效性，亦不會對本公司的整體業務產生重大不利影響。

業 務

保 險

我們為我們的員工提供補充醫療保險。補充醫療保險為我們的員工提供了超出法定社會保險要求的更強保障。根據灼識諮詢的資料，我們的保險覆蓋範圍符合市場慣例。我們及時審查我們的保單，以確保其符合中國法定法律及法規。我們認為，我們現有的保險覆蓋範圍對於我們的業務運營已屬足夠，並且符合一般市場慣例。

於往績記錄期間，我們並未遭受任何重大保險索賠。更多詳情請參閱「風險因素－我們有限的保險範圍可能使我們面臨巨額成本及業務中斷」。

牌照、批准及許可

據我們的中國法律顧問告知，截至最後實際可行日期，我們已從相關政府部門取得對我們的業務營運屬重要的所有必要執照、許可、批准及證書，且於該等執照、許可及批准屆滿時續期並無重大法律障礙。下表載列截至最後實際可行日期我們業務運營所需的重要牌照及許可的詳情。

牌照/許可	持有人	授予日期	屆滿日期
增值電信業務經營許可證	本公司	2024年12月27日	2028年11月14日
增值電信業務經營許可證	上海數中科技有限公司	2025年7月17日	2030年7月17日
信息系統安全等級保護備案證明	本公司	2025年7月29日	不適用
增值電信業務經營許可證	北京一章	2026年9月17日	2031年9月17日

有關詳情，請參閱「風險因素－與我們的業務及行業相關的風險－倘我們未能取得及維持監管環境下適用於我們業務的必需執照及批准，我們的業務、經營業績及財務狀況可能會受到重大不利影響。」

法律訴訟及合規

於往績記錄期間及直至最後實際可行日期，我們並無涉及我們認為會對我們的業務、經營業績、財務狀況或聲譽和合規產生重大不利影響的任何實際或未決的法律、仲裁或行政訴訟(包括任何破產或接管程序)。於往績記錄期間及直至最後實際可行日期，根據我們中國法律顧問的意見，除本文件所披露者外，我們在所有重大方面均已遵守相關適用的中國法律及法規。

業 務

風險管理及內部控制

我們已建立並正在維持風險管理和內部控制系統，其中包括我們認為適合我們業務運營的政策和程序。我們致力於不斷改進該等系統。我們已於業務運營的各個方面採取並實施風險管理政策。我們的董事會負責建立和更新我們的內部控制系統，而我們的高級管理層則監督各附屬公司和職能部門的內部控制程序和措施的日常實施。

營運風險管理

為管理該等風險，我們已為我們的業務實施不同的內部控制程序。該等風險主要源於內部流程和系統的缺陷、人員錯誤、IT基礎設施的中斷或外部事件。我們認為該等風險為我們業務的主要風險，但認為可通過適當的營運程序有效地控制和減輕該等風險。我們的政策規定各部門的主要營運風險和相應的控制措施。對於我們的智算基礎設施解決方案，我們已成立項目決策委員會以監督項目生命週期。項目決策委員會對項目的成本、風險及潛在盈利能力進行全面評估。我們至少每年進行一次風險管理評估，我們的內部審計職能部門將評估結果報告予我們的董事，以進行持續改進。

財務報告風險管理

我們已制定一套與財務報告風險管理相關的會計政策，包括預算管理政策、資金管理政策及費用管理政策。我們設有各種程序和IT系統來實施我們的會計政策，我們的財務部門根據該等程序審查我們的管理賬目。我們的財務部門於集中管理模式下運作，附屬公司的資金由本集團集中管理，以確保資金安全和效率。我們亦為財務部門員工提供定期培訓，以確保彼等了解我們的財務管理和會計政策，並在日常運營中貫徹執行。

合規風險管理

為管理我們對中國不斷變化的監管環境的合規性，我們已實施一系列內部控制措施。這點至關重要，因為我們在不同地區的業務運營均受到獲取和更新執照、許可及批准的各種要求的約束。我們有指定人員定期監控相關政府部門頒佈的法律、法規及政策的變化，以確保我們獲得經營業務所需的牌照。我們亦設有專門的數據合規部門以監督我們的數據安全和隱私合規情況，從而確保我們的運營符合適用的法律。我們的法務部門審查所有重要合同和業務安排，以識別並減輕潛在的法律風險。

知識產權風險管理

我們已建立內部知識產權管理制度，並已實施各種政策，以妥善管理我們的知識產權組合，降低侵權訴訟的風險。作為一家技術密集型公司，我們在運營過程中可能會面臨持有專利或其他知識產權的第三方的索賠，指控我們侵權或主張其權利。我們已實施嚴格的保密措施，並要求員工簽署保密協議，以保護我們的專有技術和商業秘密。我們亦

業 務

在推出新產品之前進行知識產權檢索和分析，以盡量減少侵犯第三方知識產權的風險。有關更多詳情，請參閱「風險因素－我們可能會面臨知識產權侵權索償，這可能會導致我們的財務和管理資源分散，或對我們的業務、經營業績及財務狀況產生不利影響。」

獎項和認可

於往績記錄期間及直至最後實際可行日期，我們在產品、技術及創新方面收獲眾多獎項和認可，其中部分重要的獎項和認可載列如下：

獎項／認可	頒獎機構	頒授年份
入圍《高質量數字化轉型產品及服務全景圖(2022)》圖譜「人工智能」板塊	中國信通院	2023年
2022年度AI金雁獎－人工智能領軍企業大獎	中國電子商會人工智能委員會	2023年
2024大模型典型示範應用案例集－DataCanvas Alaya NeW智算操作系統	中國信通院華東分院	2024年
首批通過中國信通院四大領域評測的AI系統	中國信通院	2025年
中國獨角獸企業TOP50	融中榜	2025年及2026年
2025年算力互聯領域優秀案例－「算力計量計價」技術創新實踐項目	北京市通信與互聯網協會	2025年
國家級專精特新「小巨人」企業	工業和信息化部(MIIT)	2025年
2025年度獨立智算雲獨角獸企業(金獎)	德本諮詢	2025年
「AI中國」機器之心2025年度評選人工智能領軍企業TOP 20	機器之心	2025年
Alaya NeW Cloud智算雲平台能力評測	中國信通院	2026年
Alaya Token高質量Token雲服務能力評估	中國信通院	2026年
入選亞太經合組織數智賦能案例集 (Alaya NeW Cloud 賦能印尼NeutraDC算力中心項目)	亞太經合組織(APEC)	2026年
入選2025年度新型信息基礎設施協調發展典型案例名單(算力計量計價產業應用)	工業和信息化部	2026年
2026中國AI新雲市場領導獎	弗若斯特沙利文	2026年