OVERVIEW

Who Are We

We are a general-purpose graphics processing unit ("**GPGPU**") provider that develops and sells GPGPU products to meet the AI computing needs of China's diverse industries. Anchored in our proprietary GPU+ (Minsky™) architecture — a domestically developed, software-defined on-chip heterogeneous computing framework — we focus on enterprise AI as our core target market, while also seeking to expand into the public cloud market and emerging fields such as robotics. Our goal is to deliver energy-efficient and secure computing platforms along with scenario-optimized solutions, positioning us as a foundational partner in China's industrial AI upgrade. China enterprise AI GPGPU market grew at a CAGR of 58.1%, from RMB1.9 billion in 2021 to RMB11.8 billion in 2025. We were the third largest company in China enterprise AI GPGPU market in terms of revenue in 2025 with a market share of 1.0%, according to Frost & Sullivan. This position reflects our deep alignment with the core needs of enterprise AI, including power efficiency, cost-effective design, software compatibility, and precision-tailored hardware and services.

Our GPU+ (Minsky™) architecture integrates GPGPU cores, AI acceleration units, and a high-bandwidth on-chip memory subsystem into a single chip. Through tight software-hardware co-design, it enables dynamic resource scheduling and energy-efficiency optimization. This architecture is powered by three proprietary technology engines, each addressing a critical bottleneck in real-world deployment:

- ***Dynamic Heterogeneous Scheduling Engine***

  In enterprise edge scenarios such as industrial inspection lines and retail analytics, computing resources are often constrained, and workload fluctuations lead to inefficient usage. Robotics applications, in particular, require architectures that are both low-power and capable of real-time response. Our dynamic heterogeneous scheduling engine intelligently dispatches workloads to either GPGPU cores or AI acceleration units, depending on real-time task characteristics such as image recognition, time-series analysis, or decision inference. This engine also supports scheduling strategies optimized for robotics. It enables maximized resource utilization, significantly improved energy efficiency, and reliable low-latency performance in time-critical domains, such as manufacturing inspection, smart monitoring, and embedded automotive systems.

- ***High-Density Multi-Precision Computing Engine***

  AI applications present diverse requirements for computational precision, as different scenarios demand varying trade-offs among accuracy, storage, and efficiency. For example, smart city often relies on INT8 precision for large-scale video analytics, while industrial inspection typically requires FP16 precision to ensure accurate defect detection. In advanced manufacturing sectors, ultra-high precision such as FP64 is needed for scientific modeling and engineering simulation. In generative AI, low-precision formats such as FP4 and FP8 are increasingly adopted to alleviate computing and memory constraints of LLMs, striking a balance between efficiency and performance. These multi-level precision needs make it challenging for general-purpose hardware to simultaneously deliver flexibility and computing efficiency. Our solution integrates specialized AI computing units that feature multi-precision cores and dynamic sparsity acceleration. These units deliver up to three times the computing density of conventional architectures and include built-in scalability for large-system deployment, such as thousand-card clusters. This engine enables a flexible, cost-effective platform that accommodates diverse inference workloads, including vision tasks, scientific simulations, and quantized LLM inference, helping enterprises optimize total cost of ownership through precision-aligned performance.

- ***High-Bandwidth, Low-Latency On-Chip System***

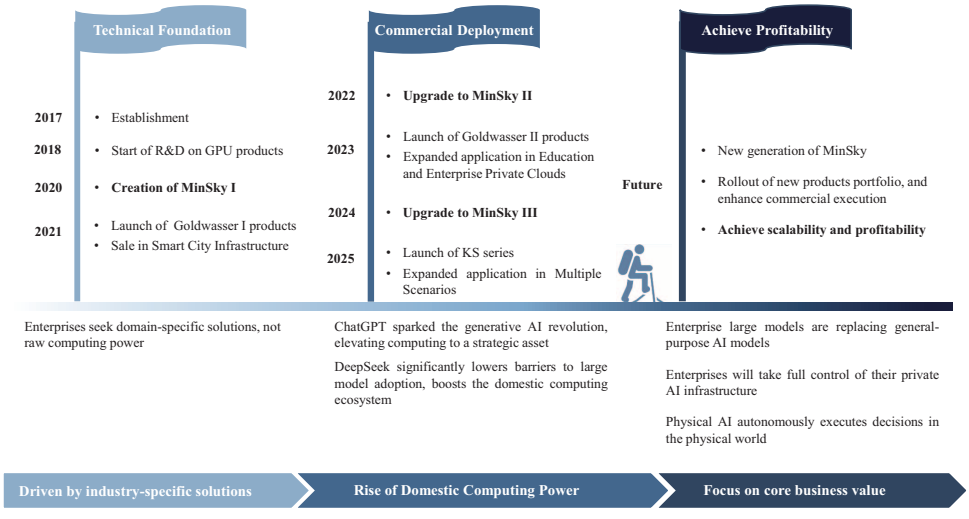
  In high-concurrency, real-time workloads, such as large-scale inference clusters, autonomous driving sensor fusion, or millisecond-level financial decisioning, performance is often bottlenecked by data movement across memory subsystems. Off-chip bandwidth costs and latency make conventional solutions impractical at scale. We address this with a proprietary

on-chip interconnect architecture that delivers high bandwidth and low latency, while significantly reducing the number and distance of data movement paths. This design reduces reliance on costly external high-bandwidth memory ("**HBM**") by achieving comparable performance and power efficiency at a lower cost, while ensuring sustained and deterministic system performance. It supports stringent latency and throughput requirements across scenarios like second-level alerts in smart cities, real-time control in industrial automation, and sub-millisecond analytics in high-frequency trading environments.

Compared to traditional GPU architectures, our GPU+ (Minsky™) architecture maintains full compatibility with CUDA and OpenCL while overcoming key inefficiencies associated with homogeneous computing pipelines and reliance on external bandwidth. It achieves higher computing efficiency and density under the same process node, enabling scalable, scenario-adaptive AI performance across applications such as smart cities, industrial inspection, and LLM inference. At the same time, our GPU+ (Minsky™) architecture offers the energy efficiency of fixed-function ASICs, while retaining the programmability and flexibility of GPGPUs. Its reconfigurable AI cores and software-defined control stack allow for rapid algorithm iteration and deployment, mitigating the obsolescence risks associated with static hardware. This hybrid capability is especially valuable in dynamic environments like industrial inspection, where continuous adaptation and workload evolution are the norm.

We have advanced through successive generations of our proprietary GPU+ (Minsky™) architecture, each marking a milestone in our product evolution. Minsky™ I, launched in 2020, powered our first-generation GPGPU products, Goldwasser I, and established the foundation for integrating heterogeneous computing units within a single chip. In 2022, Minsky™ II was deployed in Goldwasser II products, delivering significant improvements in computing density, power efficiency, and ecosystem compatibility. Minsky™ III, introduced in 2024 and adopted in our KS Series, further enhanced adaptability for enterprise AI, particularly in edge and enterprise private cloud scenarios. Looking ahead, the first and second phases of Minsky™ IV products are expected to be launched in December 2026 and the second half of 2027, respectively, extending our architecture into new high-performance domains. While maintaining the efficiency of our on-chip memory system, Minsky™ IV may also incorporate optional HBM or other advanced memory technologies for certain scenarios in the public cloud market, such as large-scale intelligent computing centers, where extreme throughput and memory capacity are critical. This generational roadmap demonstrates how we continuously scale the versatility of our GPU+ (Minsky™) architecture to address the evolving needs of enterprise AI, public cloud, and robotics.

The following diagram illustrates the milestones in the evolution of our architecture and products.



We focus on enterprise AI as our primary market while also seeking to expand into the public cloud market and next-generation robotics. Powered by our proprietary GPU+ (Minsky™) on-chip heterogeneous architecture, we have built a structured product matrix that addresses the needs of three

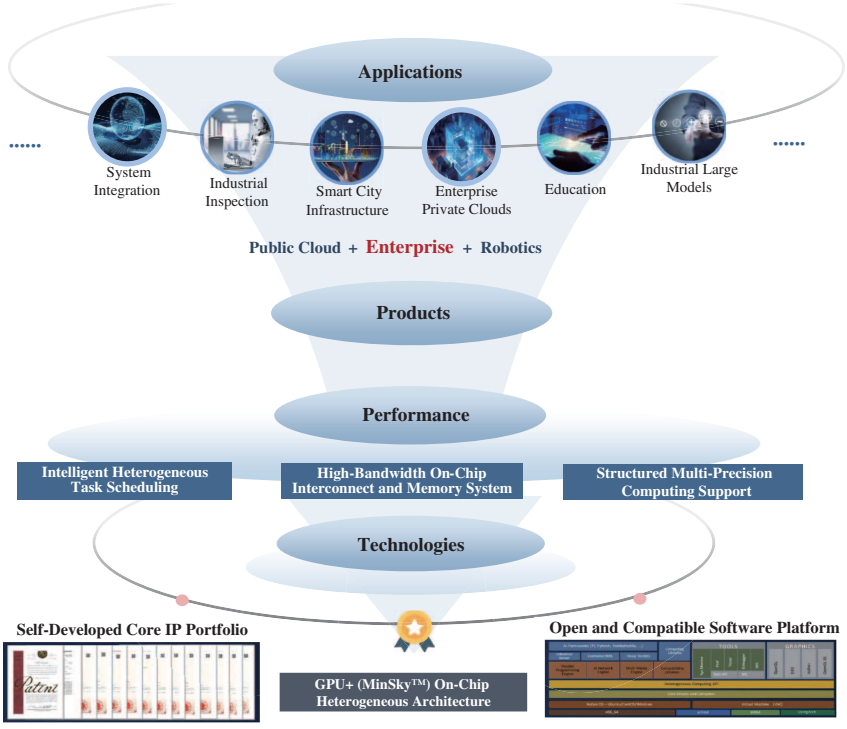
strategic market segments: enterprise, public cloud, and robotics. Our platform delivers, energy-efficient computing for a wide range of real-world scenarios — from industrial automation and smart city to regional AI inference nodes and embedded robotic systems.

Our GPGPU product portfolio mainly consists of our Goldwasser and KS series, designed to balance performance, energy efficiency, and ecosystem compatibility. Our products have been deployed at scale in high-density enterprise scenarios such as industrial inspection, smart city, and education. In addition to our data center-oriented products, we offer a range of edge inference accelerators that deliver low-power AI computing in constrained environments, supporting applications such as robotics, intelligent terminals, and embedded medical systems. These offerings reflect a roadmap designed to support AI computing across cloud, edge, and scientific domains. See "— Our Products."

Our integrated technology stack supports various applications in enterprise, public cloud, and robotics markets. Our position has been acknowledged externally. In the 2023 AIGC Enterprise Ranking by Cyzone, we were named one of the "Most Notable AIGC Companies" among over 500 candidates. In the smart city sector, our products stand out for their performance, energy efficiency, and ecosystem adaptability, and have been widely deployed in collaboration with leading vendors. In intelligent computing centers, our accelerators have been adopted for LLM training clusters, powering inference and training workloads across municipal, telecom, and internet environments.

Looking ahead, as AI inference workloads continue to shift from cloud to enterprise environments, and computing demand expands across emerging fields such as industrial simulation and the metaverse, we expect market momentum to further accelerate. Our scenario-optimized solutions, spanning hardware adaptation to software optimization, strengthen our position as a central enabler within the enterprise AI ecosystem. By capitalizing on the growth opportunities from large-model deployment and AI agent collaboration, and benefiting from the government's efforts to promote domestic GPU production, we are well-positioned to strengthen our market presence and achieve ongoing market share growth.

The diagram below provides an overview of our core capabilities, spanning foundational technologies, integrated solution components, products, and application domains.



Our Strategic Markets

Our GPGPU products are designed to serve three strategic markets, enterprise, public cloud, and robotics, each comprising differentiated AI application scenarios. While our current business focus is concentrated in the enterprise market, we are actively expanding our product portfolio and technology roadmap to address the emerging needs of the public cloud and robotics markets. For each market, we pursue a scenario-driven strategy, tailoring our products and solutions through technologies to precisely match industry-specific requirements.

Enterprise Market

The enterprise market is our current core business domain and the principal focus of our product deployment. The driving force behind ongoing industrial transformation is the pressing demand from enterprises across industries to improve operational efficiency and decision-making accuracy. From real-time defect detection on factory production lines and intelligent interpretation of medical images to inventory optimization in retail and warehousing and risk monitoring in financial transactions, enterprises are embedding AI deeply into their core business processes. This transformation creates fundamental demand for AI computing that can efficiently process massive datasets, while maintaining energy efficiency and deployment flexibility. With the rapid adoption of generative AI, enterprises are accelerating the shift from data analytics to real-time business decision making, positioning AI as the foundational infrastructure for business innovation. In this process, we provide enterprise customers with end-to-end support from product development to scenario deployment.

Public Cloud Market

China's public cloud market is experiencing rapid growth, driven primarily by surging demand from large-model training and intelligent computing centers. Unlike the enterprise market, which emphasizes private deployment and data sovereignty, or the robotics market, which relies on specialized hardware and real-time control, the public cloud market leverages platform-based, shared resources to lower customers' upfront investment through elastic leasing and pay-as-you-go models. The market is currently dominated by several leading providers, while other providers are seeking breakthroughs through vertical industry cloud offerings and high-performance computing scenarios. Going forward, we plan to collaborate with public cloud service providers to integrate our products into public cloud platforms and to provide customized, streamlined solutions for industries such as finance and healthcare that have distinctive computing requirements. With our strong energy efficiency and performance-to-cost advantages, we intend to gradually establish a solid position within the public cloud ecosystem.

Robotics Market

China's robotics market is expanding rapidly, propelled by the advancement of intelligent manufacturing and the proliferation of AI-driven real-world applications. In contrast to the enterprise market, which emphasizes private deployment and data processing, and the public cloud market, which provides platform-based resources, the robotics market demands tight integration of hardware and software as well as real-time perception, decision-making, and execution in physical environments. Robotics applications are now deeply embedded in industries such as manufacturing, medical rehabilitation, and logistics. Leveraging our accumulated expertise in high-performance computing, we plan to collaborate with robot manufacturers and system integrators to deliver AI computing solutions that combine high performance with low power consumption, thereby enabling stronger real-time perception and decision-making capabilities in robots. In addition, we are developing optimized embedded solutions tailored for scenarios emphasizing physical interaction and specialized tasks, such as industrial inspection, warehouse sorting, and interactive companionship.

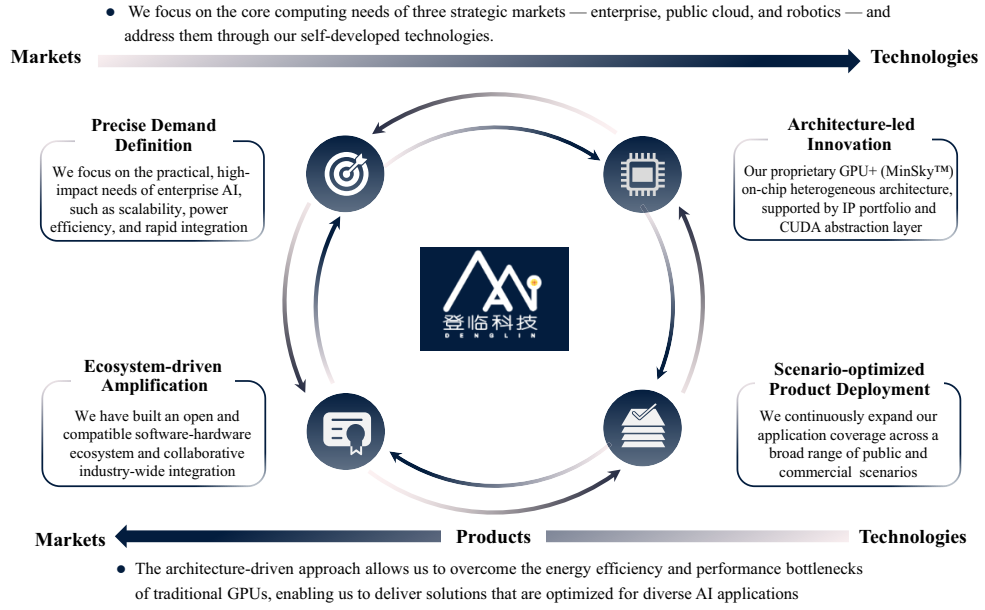
Application of Our Products

Our GPGPU products have evolved alongside the development of the AI industry, progressing from early deployments in computer vision-centric workloads to supporting generative AI models and, more recently, powering enterprise-level AI Agents. In response to these trends, we have continuously

expanded our application coverage and deepened integration with downstream ecosystems. Today, our products are deployed across a broad range of public and commercial scenarios, where they deliver energy-efficient and cost-effective computing. We work closely with ecosystem partners and end customers to address real-world needs, and our GPGPU solutions have been adopted in six key application areas: smart city, enterprise private clouds, education, industrial inspection, industrial large models, and system integration. See "— Our Products — Applications of Our Products."

OUR ECOSYSTEM

Our business model is driven by precise demand definition, architecture-led innovation, scenario-optimized product deployment, and ecosystem-driven amplification, as illustrated in the diagram below.



We focus on the core computing needs of three strategic markets, enterprise, public cloud, and robotics, and address them through our proprietary GPU+ (Minsky™) on-chip heterogeneous architecture, supported by a fully self-developed IP portfolio and a deeply compatible CUDA abstraction layer. This architecture-driven approach allows us to overcome the energy efficiency and performance bottlenecks of traditional GPUs while enabling flexible, scenario-specific hardware configurations. By aligning computing architecture with real-world deployment requirements, we deliver solutions that are optimized for scalability, power efficiency, and rapid integration across diverse AI applications.

Our ecosystem is structured around a three-layered technology stack encompassing hardware platforms, scenario-optimized software libraries, and algorithm toolchains. In collaboration with ecosystem partners, we have built a vertically integrated value chain that spans from the chip layer to system-level integration and application deployment. This structure forms a self-reinforcing growth loop: real-world demands drive architectural iteration, scenario-based deployments validate and refine our product matrix, and ecosystem expansion accelerates domestic substitution.

Our Market Opportunities

Enterprise AI is becoming the primary growth engine of China's GPGPU market. According to Frost & Sullivan, the size of China enterprise AI GPGPU market reached RMB11.8 billion in 2025 and is projected to surge to RMB108.5 billion by 2030, representing a CAGR of 61.3% from 2026 to 2030. This rapid expansion is driven by the widespread deployment of intelligent terminals, such as smart city analytics systems and industrial inspection devices, and the increasing alignment between edge computing and the demands of real-time inference. We expect to benefit from the expansion of this

high-growth market. The rise of industry large models is further amplifying demand for GPGPUs as foundational infrastructure for AI deployment. In verticals such as financial risk control and intelligent transportation, the need for high-throughput, low-latency inference is accelerating. With broad downstream coverage across sectors including smart cities and energy management, we believe that we are well positioned to capture market share through rapid and targeted product deployment in diversified application environments.

We believe that our market opportunity is supported by three structural drivers. First, AI inference workloads are shifting from public cloud to enterprise endpoints, where the need for low-latency, efficient GPGPUs, particularly for applications such as real-time video analysis and industrial defect detection, is growing rapidly. These requirements are closely aligned with our technical strengths in heterogeneous architecture optimization, computing density, and power control. Second, emerging fields such as industrial simulation and the metaverse are releasing new computing demand at edge nodes, further expanding the addressable market for GPGPU-based solutions. Third, national policy is providing strong support for domestic GPGPU development and adoption, including initiatives such as "Eastern Data, Western Computing", which promotes intelligent computing infrastructure, and procurement guidelines under the trusted computing initiative, which accelerate the scale-up of domestic GPUs in critical industries. Through our solutions spanning hardware and software integration, we aim to capture long-term opportunities arising from the convergence of large models and AI agent collaboration.

COMPETITIVE STRENGTHS

We Are an Early Developer in On-Chip Heterogeneous GPGPU Architecture, Enabling Breakthroughs in Performance and Power Efficiency

Through architectural breakthroughs, proprietary IP development, and deep hardware-software co-design, we have built a computing platform that combines high performance with low power consumption. Our GPU+ (Minsky™) architecture integrates AI accelerators, GPGPU computing units, and large-bandwidth on-chip memory subsystems into a single chip. It overcomes the efficiency limitations of traditional GPU designs, while ensuring full compatibility with CUDA and OpenCL, thereby supporting model migration and significantly reducing integration costs for customers.

Our GPU+ (Minsky™) architecture delivers three key performance breakthroughs:

- ***Intelligent Heterogeneous Task Scheduling***

  Based on real-time workload characteristics, computing tasks are dynamically routed to either GPGPU computing units or AI acceleration units. This approach greatly improves computing efficiency and raises effective throughput per watt compared to traditional homogeneous GPU designs.

- ***High-Bandwidth On-Chip Interconnect and Memory System***

  We developed a proprietary high-bandwidth interconnect fabric paired with a large-scale distributed shared cache system. Most data movement is handled within the chip, substantially reducing dependency on external memory bandwidth and resolving latency and throughput bottlenecks.

- ***Multi-Precision Computing Support***

  We support a full range of numeric precisions, from FP4 to FP64, through integrated fixed-point and floating-point computing units. This enables the same architecture to support low-precision edge inference and high-precision scientific computing.

We have built a fully self-developed portfolio of core IPs that spans instruction sets, computing engines, and memory systems. This enables us to maintain full architectural autonomy, optimize performance for real-world applications, and reduce reliance on external suppliers:

- ***Instruction Sets Enabling Real-Time Heterogeneous Scheduling and Flexible Precision***

  Our dynamic scheduling instruction set allows ultra-low-latency coordination across GPU and AI acceleration units. It serves as the core of our software-defined architecture, enabling microsecond-level task assignment that adapts dynamically to workload demands. In parallel, our mixed-precision instruction set supports transitions between FP4, INT8, FP16, BF16, and FP64 formats. This level of flexibility allows a single architecture to support both lightweight edge inference and precision-critical scientific computing, maximizing resource efficiency and simplifying system design.

- ***Computing Engines for AI Acceleration, General-Purpose Computing, and High-Precision Floating Point***

  Our AI acceleration units are built around in-house developed tensor engines that support multi-precision operations, dynamic sparsity, and structured operator acceleration. These features enable high-throughput inference for both dense and sparse models with strong energy efficiency. Our GPGPU computing units are designed with self-developed microarchitectures that retain full compatibility with CUDA and OpenCL, while being deeply optimized for our hardware stack. Additionally, our architecture includes FP64 double-precision computing units tailored for scientific modeling, engineering simulations, and other HPC-grade applications that demand high numerical accuracy.

- ***Memory System Designed for High Bandwidth, Low Latency, and Efficient Data Management***

  We have developed a distributed shared cache system with large-capacity, high-bandwidth on-chip memory, coordinated by compiler and driver-level controls to improve data reuse and execution efficiency. This system reduces dependence on external memory bandwidth to just one-fifth of that required by comparable international solutions, according to our internal testing results. Complementing this is our proprietary on-chip interconnect, capable of terabyte-scale data transfer between computing engines and memory under high-concurrency conditions. An intelligent memory access controller further enhances performance by dynamically optimizing scheduling and access patterns — helping to overcome the "memory wall" and support real-time, power-efficient computing at scale.

These intellectual property assets cover core architectural modules such as instruction sets, computing engines, and memory systems, and form the technological foundation for our long-term scalability, ecosystem compatibility, and architectural independence.

We Focus on Core Enterprise AI Scenarios and Serve as a Scalable Engine for Intelligent Solution Delivery

We have a deep understanding that the core value of enterprise AI lies in driving business growth and optimizing operational efficiency. In fragmented, high-frequency scenarios such as smart city and industrial inspection applications, our proprietary GPU+ (Minsky™) heterogeneous on-chip architecture significantly enhances effective computing power per unit of energy consumption. This "precision efficiency" addresses the shortcomings of traditional GPUs, which are low utilization, high energy consumption, and coarse-grained computing, helping enterprises substantially reduce total cost of ownership and meet the rigid demand for cost reduction, efficiency improvement, and green operations.

To address enterprise concerns over ecosystem migration and supply chain uncertainties, we provide practical and reliable safeguards. Our CUDA/OpenCL compatibility layer and automated migration toolchain enable customers' core business algorithms to transition smoothly, lowering both

the technical threshold and time cost of migration. At the same time, we have built a stable and reliable supply chain that ensures business continuity for enterprise customers and supports long-term, stable product delivery, offering a secure and credible option for domestic substitution.

Our open architecture design grants enterprise customers valuable technological autonomy. Our modular heterogeneous on-chip architecture allows key computing units to evolve independently. This gives enterprises the flexibility to choose their technical roadmap and upgrade pace according to their own needs, effectively avoiding risks of technology lock-in or mandatory system-wide overhauls, and safeguarding the long-term value of their investments.

To better support the diverse applications of enterprise AI, we have built a structured product portfolio covering a wide range of scenarios. In core fields such as smart city and industrial inspection, our products, coupled with deeply optimized algorithms, provide mature solutions proven in large-scale projects. In response to rising demand for LLM inference, our innovative memory solutions offer a highly cost-effective option. At the same time, in areas such as scientific computing, critical infrastructure, and intelligent terminals, we provide high-precision computing and edge intelligence platforms, demonstrating strong adaptability and a structured layout capable of meeting enterprises' diverse AI needs.

We Have Built an Open and Compatible Software-Hardware Ecosystem to Facilitate Domestic GPU Migration and Adoption

We have contributed to the development of a domestic GPU ecosystem through a systematic strategy grounded in application programming interface ("**API**")-level compatibility, deep software stack optimization, and collaborative industry-wide integration. This approach significantly lowers the migration threshold for customers transitioning from international GPU platforms and accelerates the large-scale deployment of domestic AI solutions across key industries. At the core of this effort is our compatibility with CUDA and OpenCL. We provide standardized APIs for drivers, computing libraries, and inference engines, while implementing deep hardware-specific optimizations under the hood to ensure stable, efficient execution on our GPU+ (Minsky™) architecture.

To support developer adoption and reduce learning costs, we offer a suite of tools, spanning device monitoring, performance profiling, and debugging. Developers familiar with international toolchains can quickly become productive on our platform without having to overhaul their workflows. In addition, we provide model conversion and optimization utilities that support rapid migration of AI models originally built for international GPU platforms. These tools retain model fidelity while optimizing execution on our hardware, effectively shortening integration cycles and accelerating deployment. This migration toolchain has already proven itself in large-scale commercial use: we have completed the replacement of over 10,000 Company N-based accelerator cards at a large smart city service provider and have validated our ecosystem integration with about 300 algorithms and server partners across multiple AI verticals.

Beyond software compatibility, we have built an ecosystem spanning hardware, operating systems, frameworks, and application-layer support. Our GPU+ (Minsky™) architecture supports collaboration with both international and domestic CPU platforms, including x86, ARM, and RISC-V (an open-source instruction set architecture), enabling flexible heterogeneous deployment across diverse infrastructure environments. We offer native drivers for both global operating systems such as Windows and Linux, and Chinese mainstream operating system (OS) platforms including Kylin, UnionTech (Tongxin), and openEuler, ensuring stable system-level integration in both domestic and international settings.

We have also optimized leading AI frameworks such as PyTorch and TensorFlow for our platform and completed performance tuning for domestic frameworks, delivering a consistent and efficient development experience across software stacks. For application-level enablement, we provide a pre-integrated software development kit (SDK) and reference implementations for priority scenarios such as smart city and industrial inspection. Working with partners, we have also introduced integrated AI devices, including edge servers and terminal systems, that incorporate our computing capabilities into turnkey and deployment-ready products. This open, compatible, and vertically integrated ecosystem facilitates domestic GPU adoption.

We Are Led by a Forward-Looking, Technically Grounded Management Team and Sustained by a Pragmatic Culture

We are guided by a management team with deep domain expertise, global perspective, and a strong track record of execution. The core management team comprises experts who have spent over two decades in the GPU industry and previously held senior technical and commercial roles at well-established domestic and international GPU companies. They bring with them end-to-end capabilities that span from architectural design and R&D to business strategy and go-to-market execution. Under their leadership, we have achieved rapid product iteration and established a differentiated competitive position. The continuity of our strategic direction is reinforced by zero turnover among key technical and managerial positions, which ensures alignment across product development, roadmap evolution, and organizational execution.

Our management has consistently demonstrated a forward-looking mindset. Drawing from deep insights into global AI infrastructure trends, they anticipated the breakout growth of AI inference workloads, the rise of edge computing, and the increasing demand for intelligent terminals. This foresight enabled us to position our architecture early around low latency and efficient edge computing and to align product planning with the real-world needs of applications such as smart city governance and embedded AI. Their early strategic decisions have allowed us to capture critical market opportunities ahead of the curve and build lasting competitive advantages in targeted verticals.

Talent has been a key driver of our growth. We have built a highly capable and relatively young team, with nearly 70% of our workforce engaged in R&D as of April 30, 2026, most of whom have over ten years of industry experience, including approximately 20% with backgrounds at leading international GPU companies. This combination of technical depth, global experience and a dynamic workforce supports efficient execution of our product roadmap. Our culture emphasizes delivering practical, customer-oriented innovation through continuous investment in core technologies and independent IP development, supported by a co-design approach that aligns product development with customer needs and accelerates deployment. We also adopt an open and compatible ecosystem strategy to facilitate adoption, and incentivize talent through competitive compensation and equity-based incentives, fostering long-term commitment and operational agility in a rapidly evolving market.

DEVELOPMENT STRATEGIES

Expanding Our Product Portfolio to Serve Public Cloud and Robotics Markets through Scenario-Driven Development

We plan to extend our core GPU capabilities beyond the enterprise market into the strategically important and fast-growing public cloud and robotics sectors. Our GPGPU products are currently deployed primarily in enterprise AI scenarios, and we are actively broadening our product roadmap and capabilities to address emerging use cases in cloud-based AI inference and robotic computing. These expansion efforts are designed to deepen our presence in technology-intensive verticals such as smart cities, internet services, LLM inference, healthcare AI, and intelligent mobility.

For each strategic market, enterprise, public cloud, and robotics, we pursue a scenario-driven product strategy that emphasizes product customization and ecosystem integration. Our objective is to establish flagship use cases that enable a full-cycle value chain from chip-level innovation to end-market application. This entails in-depth research into domain-specific technical requirements, including deployment environments, performance constraints, and compliance standards. Building on our standardized foundation architecture, we tailor our solutions accordingly through hardware-software co-design, algorithm adaptation, and systems integration.

To accelerate adoption, we are partnering with leading enterprises in each vertical to launch pilot projects that validate our solutions under real-world conditions. These collaborations enable us to collect actionable feedback, optimize our product offerings, and iterate. Over time, we aim to build industry-ready and replicable solution frameworks that support scaled commercialization across diverse application scenarios. This strategy will not only drive our technology deployment across a broader market base but also enhance the resilience and sustainability of our long-term growth.

Continued Investment in GPU R&D to Drive Architectural Upgrades and Enhance Product Competitiveness

We will continue to increase our investment in GPU research and development, guided by frontier trends in architecture design, computing performance, and energy efficiency. By allocating additional capital, talent, and organizational focus to core GPU innovation, we aim to maintain a sustained technological lead and strengthen our position in high-performance computing markets.

To support this objective, we plan to build a high-caliber R&D team and establish industry-academic collaboration platforms with leading universities and research institutions. These partnerships will focus on critical areas such as GPU architecture optimization, computing throughput scaling, and breakthroughs in performance-per-watt. Our roadmap includes continuous iteration of our core GPU product lines to support next-generation AI and high-performance computing workloads.

We will implement a fast iteration mechanism that incorporates both real-time market feedback and internal technology evolution. By accelerating the release of updated GPU products with enhanced computing capabilities, lower power consumption, and improved ecosystem compatibility, we aim to deliver differentiated value in terms of performance, cost-effectiveness, and developer experience. Looking ahead, we have completed IP development for a high-performance architecture targeting AI for Science, and are actively developing a GPU architecture optimized for graphics rendering and visualization. This will ensure that our products remain competitive across performance tiers and reinforce our position as a core player in China's GPU industry.

Enhancing Our Software Toolchain to Improve Product Compatibility and Development Efficiency

We will continue to strengthen our software toolchain to further enhance the compatibility, accessibility, and ease of use of our GPU+ (Minsky™) architecture. Our architecture is natively compatible with CUDA and OpenCL, enabling integration into existing development ecosystems and significantly reducing customer migration costs. Building on this foundation, we plan to enrich our software offerings through cross-functional collaboration among market researchers, system architects, and user experience specialists, with a focus on emerging domains such as AI, cloud computing, and data center infrastructure.

We have released the Hamming® software toolchain, an integrated suite that supports major domestic and international AI frameworks, operating systems, CPU architectures, container environments, virtual machines, and inference servers. This toolchain is designed to help customers rapidly deploy their algorithms in real business scenarios and to maximize hardware utilization across cloud and data center environments. It also provides robust support for development, testing, and monitoring across diverse workloads and runtime environments.

As we continue to evolve the platform, we will deepen collaboration with upstream and downstream partners, including hardware vendors, software developers, and industry solution providers. Through open technical interfaces and the provision of development kits, we aim to attract more ecosystem participants and jointly build a complete product system that spans hardware, software, algorithms, and application layers. These efforts will increase the breadth of our market coverage, strengthen developer engagement, and further enhance the stickiness and scalability of our platform in enterprise and infrastructure environments.

Expanding Our Customer Base and Deepening Strategic Partnerships to Drive Commercial Adoption

We will continue to expand our customer base by reinforcing our presence in core industries and cultivating long-term partnerships with leading enterprises. We adhere to a sector-focused strategy that emphasizes deep engagement in specific verticals — an approach that allows us to better understand industry-specific application requirements and build tailored solutions with strong product-market fit. This focus has enabled us to develop competitive advantages in key markets such as smart cities and large model computing, where we aim to consolidate our market position.

Given the globally leading demand scale of Chinese enterprise customers, we intend to leverage our geographic proximity to respond to downstream needs with greater speed and precision. Early access to customer feedback will enable us to accelerate product iteration and solution expansion, ensuring that our offerings remain relevant and aligned with evolving industry needs.

To support this expansion, we will build a more structured sales and customer engagement network. Our go-to-market strategy will be supported by a "professional + vertical" team matrix, combining deep technical knowledge with domain-specific insight. This structure will enable us to identify and serve new demand scenarios across industries, improve customer conversion and retention rates, and ultimately enhance the commercial deployment and scalability of our AI computing solutions.

Attracting and Developing Top Industry Talent to Sustain Long-Term Innovation and Commercial Growth

We are committed to building a structured talent strategy that aligns with our long-term goals in both technology development and market expansion. To enhance our market position in the GPU field, we will continue to attract top-tier R&D experts with proven experience in architecture design, compiler optimization, and AI computing systems. These professionals will strengthen our ability to drive innovation across our core product lines and support ongoing breakthroughs in performance, energy efficiency, and ecosystem integration.

At the same time, we will expand our commercial and customer-facing teams, including solution engineering, sales, and marketing, to better capture opportunities in high-growth verticals. By cultivating a well-rounded team that combines technical excellence with market insight, we aim to improve customer acquisition and retention, accelerate the rollout of new products, and enhance overall commercial execution. This dual-track approach to talent development will ensure we maintain strong momentum as we scale our platform and solutions across broader markets.

OUR BUSINESS MODEL

GPGPU Products

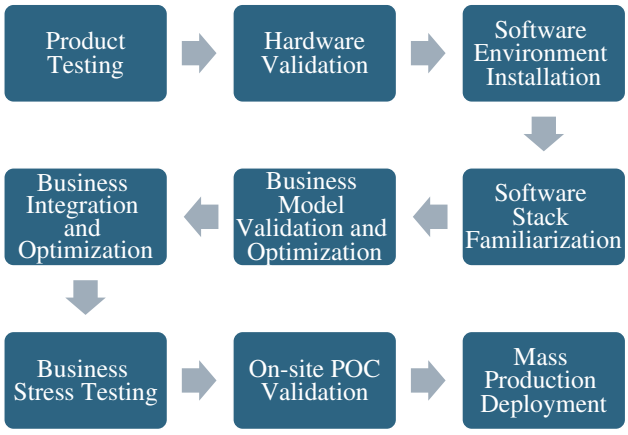
We focus on the development and sale of hardware products, especially board-level GPGPUs, which are supplied to original equipment manufacturers (OEMs), system integrators, and AI solution providers. We operate under a fabless business model, which is the prevailing model in the semiconductor industry for companies focused on chip innovation and architectural advancement. This model allows us to concentrate our resources on the design and development of energy-efficient GPGPUs, while leveraging the specialized capabilities of third-party partners for manufacturing. Specifically, we retain full control over the development of our GPU+ (Minsky™) architecture, system software stack, and proprietary IPs, while wafer fabrication, chip assembling and testing are outsourced to foundries and OSAT providers. Our GPGPU products are deployed across a broad range of enterprise AI application scenarios, including smart city, enterprise private clouds, education, industrial inspection, industrial large model deployment, and system integration. This model enables us to remain capital-efficient while scaling our product delivery across diverse and growing AI markets.

Our GPGPU products are primarily standardized hardware products offered under defined specifications and performance tiers. Customization, where applicable, is generally limited to software configuration, system integration parameters or deployment optimization, and does not involve fundamental changes to the underlying hardware architecture.

Pricing for GPGPU products is transaction-based and determined by product model, performance specifications, order volume and prevailing market conditions. Variations in pricing among different product lines primarily reflect differences in computing performance, memory capacity and application positioning.

Revenue from sales of GPGPU products is recognized when control of the products is transferred to customers, which typically occurs upon delivery of the products and customer acceptance in accordance with the relevant contractual terms. Cost of sales is recognized concurrently with revenue recognition and primarily comprises wafer fabrication costs, packaging and testing fees, materials and other directly attributable production costs.

Our business process reflects our standard procedure for onboarding a customer and deploying our GPU acceleration solutions in real-world scenarios. This step-by-step flow ensures smooth technical integration, efficient validation, and rapid productization, helping customers adopt our platform with minimal friction and maximum performance.



The key steps in the process are as follows:

- *Product Testing* — Initial shipment of the product to the customer for testing purpose.
- *Hardware Validation* — Confirmation by the customer of basic hardware functionality.
- *Software Environment Installation* — Deployment of required drivers, toolkits, and middleware.
- *Software Stack Familiarization* — Familiarize customer with our unified software platform.
- *Business Model Validation and Optimization* — Benchmarking and tuning of target AI models and applications.
- *Business Integration and Optimization* — Integration of the solution into customer workflows and adaptation to specific system needs.
- *Business Stress Testing* — Full workload simulation to evaluate stability and scalability.
- *On-site POC Validation* — Customer conducts proof-of-concept at deployment site to validate end-to-end performance.
- *Mass Production Deployment* — Final confirmation and ramp-up of volume shipments into production environments.

Technical Services

We also provide technical services relating to GPGPUs, such as chip design services to customer-specific projects covering hardware-software co-design, customization of our architecture for specific workloads, and optimization of models and algorithms for deployment in customer environments. These services are typically project-based and enable us to address the specific technical and deployment requirements of large enterprise and institutional customers. In addition, as part of such engagements, we may license selected proprietary technologies, software components or computing

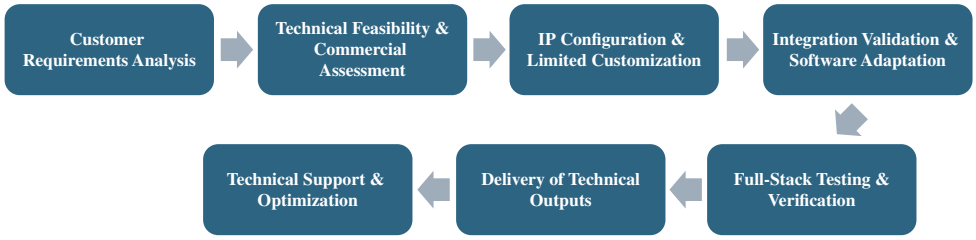
modules to customers for incorporation into their systems or chip designs. Our chip design services and IP licensing offerings are primarily based on standardized architectures, reusable IP modules and software platforms, with limited project-specific configuration or customization to meet customer requirements, allowing us to deliver services efficiently while maintaining consistency and scalability across projects.

IP licensing arrangements are historically offered on a standalone basis, and may also be provided in connection with technical service engagements. Such arrangements are typically structured as non-exclusive, non-transferable and non-sublicensable licenses limited to the customer's internal use within the scope of the relevant project or agreed application scenario. License consideration is usually one-off, determined as part of the overall fixed-fee or milestone-based service pricing, and we do not currently generate material recurring royalty income from such arrangements. Under these arrangements, our pre-existing and background intellectual property, including core GPU architectures, reusable IP blocks, software platforms and general-purpose technologies, remains owned by us. While customers may own or be granted rights to project-specific deliverables developed for their commissioned projects, we retain ownership of, and the right to reuse, underlying technologies, methodologies and non-customer-specific know-how. Customers are generally restricted from reverse engineering, sublicensing or using our proprietary IP beyond the agreed scope and are subject to confidentiality and IP protection obligations, enabling us to support customer-specific implementation while preserving control over our core technologies and their continued application across multiple projects and customers.

Chip design services and IP licensing are priced on a transaction-based basis, typically under fixed-fee or milestone-based arrangements determined by the agreed scope of services, delivery complexity and project duration.

Depending on the nature of the engagement, revenue from our chip design services is recognized on a point-in-time basis, i.e., recognized in full upon completion of the service and formal customer acceptance. For contracts that include design services and the production and delivery of samples, revenue is recognized based on the actual completion of distinct performance obligations upon delivery of the relevant services or samples and customer acceptance. For IP licensing services, revenue is recognized upon delivery of the licensed materials and customer confirmation. Cost of sales is recognized in line with revenue recognition, with costs associated with sample production and delivery recognized concurrently with the related revenue, and costs for other technical services recognized as incurred. Costs primarily comprise personnel costs, engineering resources and project-specific expenses.

Below is a simplified flow chart for our chip design services business line.



- *Customer Requirements Analysis* — Identification of customer application scenarios, technical objectives and system requirements.
- *Technical Feasibility and Commercial Assessment* — Assessment of technical feasibility, delivery scope, resources and commercial terms.
- *IP Configuration and Limited Customization* — Configuration of existing IP modules and software platforms, with minor customization if required.
- *Integration Validation and Software Adaptation* — Validation of system integration and adaptation of software and toolchains.

- *Full-Stack Testing and Verification* — Regression testing across hardware and software layers against agreed specifications.
- *Delivery of Technical Outputs* — Delivery of agreed software, documentation and related technical deliverables.
- *Technical Support and Optimization* — Ongoing technical support during deployment and post-delivery optimization.

Overall, our business model for both business lines is transaction-based, with revenue generated from discrete sales of GPGPU products, as well as project-based technical services, rather than subscription-based models. Cash receipts from sales of products and technical services are generally settled through customer bank transfers.

OUR PRODUCTS

We have developed a structured, multi-generational GPGPU product portfolio to address the growing and evolving demands of enterprise AI, which has become a key driver of the deployment of AI computing power across China. Our products are designed to serve enterprise AI across a range of deployment environments, including edge nodes, public cloud, private clouds, and terminal devices and are built on an original GPU+ (Minsky™) architectural foundation supported by a set of unified Hamming® Software Stack. This enables consistent performance scaling and efficient hardware-software co-optimization.

Since 2021, we have launched multiple product generations across two core categories: Enterprise Data Center Products and Edge Inference Products, each targeting distinct deployment environments and technical requirements in enterprise AI infrastructure. These products are all built on successive generations of our proprietary GPU+ (Minsky™) architecture, which provides the common foundation for computing density, power efficiency, and ecosystem compatibility across our portfolio.

- Our Enterprise Data Center Product line has evolved through three generations. The Goldwasser I L Series (L128, L256), launched in 2021 and powered by Minsky™ I, focused on efficient computing for foundational AI workloads. The Goldwasser II GS Series (GS30, GS40), introduced in 2023 and based on Minsky™ II, delivered enhanced throughput and memory capacity for high-concurrency inference. The KS Series (KS28, KS38), launched in 2025 and incorporating Minsky™ III, combines strong computing performance with structured domestic compatibility to meet the growing demand for secure and localized infrastructure in sectors such as smart city and enterprise private clouds.
- Our Edge Inference Product line has also advanced in tandem with the evolution of the Minsky™ architecture. The Goldwasser I UL/MXM Series (UL32, UL64, MXM32, MXM64), launched in 2021 and built on Minsky™ I, was designed for low-power inference in space-constrained environments such as industrial inspection and education. The KS Series (KS19, KS20), introduced in 2025 and powered by Minsky™ III, provides enhanced performance and ecosystem compatibility under sub-50W power budgets. The KS Series also reflects our strategic response to rising demand for fully domestically produced and deployed AI solutions across both data center and edge environments.

Looking ahead, we expect to launch the first and second phases of Minsky™ IV products in December 2026 and the second half of 2027, respectively, further extending our GPU+ platform into high-performance domains. The first phase will continue to use our LPDDR-based on-chip memory subsystem, while the second phase, which we expect to launch in the second half of 2027, will incorporate HBM3e or other advanced memory technologies, to address the extreme throughput and memory needs of public cloud and intelligent computing center scenarios.

The principal performance specifications of our major product series are summarized below.

- Our Minsky™ I-based Goldwasser I products deliver computing power of up to 256 TOPS (INT8), 64 TFLOPS (FP16) and 16 TFLOPS (FP32), with 8GB to 64GB of memory, memory bandwidth of up to 239GB/s and power consumption of up to 55W, and are primarily used for inference in smart city scenarios.
- Our Minsky™ II-based Goldwasser II products deliver computing power of up to 512 TOPS (INT8) and 256 TFLOPS (FP16), with 32GB to 128GB of memory, memory bandwidth of up to 192 GB/s and power consumption of up to 95W, and support both inference and training across smart city, intelligent computing center and other scenarios.
- Our Minsky™ III-based KS Series products deliver computing power of up to 560 TOPS (INT8) and 280 TFLOPS (FP16), with 8GB to 128GB of memory, memory bandwidth of up to 819.2 GB/s and power consumption of up to 130W, and support both inference and training.

Each successive series reflects the progressive iteration of our GPU+ (Minsky™) architecture and offers enhanced overall performance; the series are positioned across different power and performance tiers rather than being simple substitutes for one another, and we promote our latest series to new customers while continuing to support and migrate existing customers in line with their own product iteration cycles.

The following table sets forth our major products and their features.

Category	Generation	Series	Model	Illustration	Key Specifications
Enterprise Data Center Products.	I	Goldwasser I L Series	L128		• Typical Power Consumption: 25W • Memory Capacity: 16/32/64GB • Peak Memory Bandwidth: 239 GB/s
			L256		• Typical Power Consumption: 35W • Memory Capacity: 16/32/64GB • Peak Memory Bandwidth: 239 GB/s
	II	Goldwasser II GS Series	GS30		• Typical Power Consumption: 40W • Memory Capacity: 32/64/128GB • Peak Memory Bandwidth: 192 GB/s
			GS40		• Typical Power Consumption: 60W • Memory Capacity: 32/64/128GB • Peak Memory Bandwidth: 192 GB/s

Category	Generation	Series	Model	Illustration	Key Specifications
Edge Inference Products	III	KS Series	KS28		• Typical Power Consumption: 50W • Memory Capacity: 32/64GB • Peak Memory Bandwidth: 204.8 GB/s
			KS38		• Typical Power Consumption: 90W • Memory Capacity: 64/128GB • Peak Memory Bandwidth: 409.6 GB/s
	I	Goldwasser I UL/MXM Series	UL32		• Typical Power Consumption: 15W • Memory Capacity: 4/8/16GB • Peak Memory Bandwidth: 58.325 GB/s
			UL64		• Typical Power Consumption: 20W • Memory Capacity: 8/16/32GB • Peak Memory Bandwidth: 116.65 GB/s
			MXM32		• Typical Power Consumption: 10W • Memory Capacity: 4/8GB • Peak Memory Bandwidth: 30 GB/s

Category	Generation	Series	Model	Illustration	Key Specifications
			MXM64		• Typical Power Consumption: 15W • Memory Capacity: 8/16GB • Peak Memory Bandwidth: 60 GB/s
	II	KS Series	KS19		• Typical Power Consumption: 25W • Memory Capacity: 8GB • Peak Memory Bandwidth: 102.4 GB/s
			KS20		• Typical Power Consumption: 25W • Memory Capacity: 8/16/32/64GB • Peak Memory Bandwidth: 102.4 GB/s

Enterprise Data Center Products

Generation I — Goldwasser I L Series (L128, L256)

The Goldwasser I L Series, comprising L128 and L256, marked our initial product generation for enterprise data centers. These cards support mainstream AI model inference tasks with power consumption ranging from 25W to 35W. Designed to meet the requirements of smart city, enterprise private clouds, and industrial inspection, the L Series offers stable performance and is fully compatible with our Hamming® Software Stack.

Generation II — Goldwasser II GS Series (GS30, GS40)

Our Goldwasser II GS Series, featuring GS30 and GS40, offers significant advancements in computing density, memory bandwidth, and energy efficiency. With support for 128GB of memory per card, these products are optimized for general-purpose large models and high-concurrency inference in settings such as education, smart city, and enterprise private clouds. They are widely deployed in video-based analytics, science assessments, and LLM inference, and are fully integrated with both domestic and international AI frameworks through our GPU+ (Minsky™) architecture.

Generation III — KS Series (KS28, KS38)

The KS Series represents our latest data center products to date. KS28 and KS38 are designed for high-density deployment in AI server environments and are particularly well suited for industrial large models and system integration. These cards support FP32 to INT4 precision and offer efficient scaling under multi-card scheduling scenarios. They are widely adopted by customers seeking cost-effective, high-concurrency computing platforms compatible with both domestic and global ecosystems.

Edge Inference Products

Generation I — Goldwasser I UL/MXM Series (UL32, UL64, MXM32, MXM64)

The Goldwasser I UL/MXM Series includes UL32, UL64, MXM32, and MXM64 — low-power inference modules tailored for edge-side deployment. With power consumption under 20W, these modules are ideal for education and smart city scenarios requiring compact design and responsive inference. Typical applications include AI-assisted invigilation, science lab assessments, and real-time video analytics at the edge. All models are compatible with the Hamming® Software Stack and support localized AI workloads with minimal integration overhead.

Generation II — KS Series (KS19, KS20)

KS19 and KS20 are the second-generation edge inference products developed under the KS Series. These cards are suitable for edge deployment and are optimized for high-concurrency inference in education, smart city, and system integration. KS20 supports up to 100 concurrent video streams per card, making it suitable for multi-classroom environments and compact AI server integration. The cards are compatible with mainstream domestic OS platforms and deliver strong performance under low-latency, real-time inference conditions.

Applications of Our Products

Our GPGPU products have evolved alongside the development of the AI industry, progressing from early deployments in computer vision-centric workloads to supporting generative AI models and, more recently, powering enterprise-level AI Agents. In response to these trends, we have continuously expanded our application coverage and deepened integration with downstream ecosystems. Today, our products are deployed across a broad range of public and commercial scenarios, where they deliver energy-efficient, and cost-effective computing. We work closely with ecosystem partners and end customers to address real-world needs, and our GPGPU solutions have been adopted in six key application areas: smart city, enterprise private clouds, education, industrial inspection, industrial large models, and system integration.

Smart City

Challenges: Smart city places stringent demands on AI computing platforms due to the scale and complexity of video analytics and public safety applications. Key challenges include insufficient GPU memory for large facial recognition libraries, limited processing capacity for structured video analytics across high-density streams, and the inability of legacy solutions to support frequent model iteration or deploy across compact, energy-constrained systems. At the same time, a wide variety of AI models — spanning facial comparison, object detection, pedestrian re-identification, and behavior recognition — must be executed concurrently and cost-efficiently, often under national cybersecurity and localization mandates.

Our Solution: Our Goldwasser Series provides a scalable, efficient computing platform optimized for smart city workloads. With support for high memory capacity and computing density, the platform enables real-time analysis of massive video streams and facilitates retrieval across million-level facial databases and billion-scale pedestrian libraries. It accommodates diverse matching algorithms by supporting multiple similarity metrics, including cosine, Hamming, and hash distance. Goldwasser products can be flexibly deployed in AI boxes and 2U to 4U servers, and are engineered for dense integration with power envelopes as low as 25W. Our solution has been widely adopted in smart city applications, including public safety monitoring, video structuring, emergency and environmental monitoring, and food safety supervision. It ensures full compatibility with domestic software ecosystems while meeting the precision and flexibility requirements of large-scale AI governance systems.

Case Study: Our products have been deployed in a wide range of smart city projects to support AI-based traffic analysis, urban mobility, and vehicle-road collaboration. In a smart transportation project implemented in Yantai, Shandong Province, our accelerator cards were used to run AI algorithms for real-time processing and analysis of large-scale traffic data. The system enables accurate traffic flow prediction, route optimization, and congestion mitigation. Our products demonstrated strong usability, reliability, and stability during deployment, and delivered a compelling price-performance advantage for the customer.

In another smart city project in Xiangyang, Hubei Province focused on autonomous driving and V2X applications, our AI accelerator cards were embedded in roadside AI boxes installed by municipal traffic authorities. These devices enabled vehicle perception and video structuring on the roadside, with data synchronized to autonomous vehicles in real time. The integrated system allowed the vehicles to make driving decisions based on both onboard and road-level sensing inputs. The low power consumption and high computing performance of our products were critical to the successful deployment of this collaborative perception framework.

Enterprise Private Clouds

Challenges: As enterprise adoption of AI continues to deepen, private cloud environments must support increasingly complex and computation-intensive workloads, including LLM inference, training-finetuning cycles, and data-intensive analytics. These intelligent computing centers face persistent challenges such as rising energy consumption, suboptimal system scalability, and weak compatibility with evolving software ecosystems. Enterprises also require flexible computing infrastructure that can accommodate a diverse and rapidly changing AI workload mix — ranging from NLP and computer vision to recommendation systems — while ensuring high utilization rates and compliance with security and localization standards. Traditional data center GPUs often struggle to deliver the necessary performance-per-watt or to support containerized, cross-framework deployment efficiently.

Our Solution: Our Goldwasser Series products are built for large-scale, enterprise-grade AI workloads, offering competitive performance density, energy efficiency and deployment flexibility. With support for FP32, TF32, FP16, BF16, and INT8 computing precision, our cards enable parallel execution of multi-modal tasks such as model training, inference, image rendering, and analytics within a unified infrastructure. The architecture is fully compatible with CUDA and mainstream AI development frameworks, and supports virtualization, container orchestration, and secure cloud-native

deployment. These features make our platform ideal for use in enterprise AI computing clusters and data centers. By improving workload throughput and reducing overall total cost of ownership (i.e. the total cost of a product or service over its entire lifecycle, including direct costs (such as the initial purchase price) and indirect costs (such as time spent adjusting to new systems), as well as for short and long-term costs and cost savings), our solution empowers customers to build scalable and future-ready private cloud infrastructure that meets the evolving demands of the AI era.

Case Study: A leading state-owned enterprise deployed our accelerator cards within its internal intelligent computing center to support end-to-end industrial large model applications. These included RAG-based models for internal knowledge retrieval and task automation, which significantly enhanced workflow efficiency. In partnership with the customer, we co-developed a large model featuring both static and dynamic sparsity, and customized our hardware-software stack to optimize for their unique sparsity algorithm, substantially accelerating inference while maintaining model accuracy. Our solutions were further adapted to support the customer's customized multimodal and Stable Diffusion models, along with tailored vLLM implementations, enabling flexible deployment across diverse user groups and model types within the enterprise.

Education

Challenges: The adoption of AI in K12 education is accelerating, particularly in scenarios such as science lab assessments (physics, chemistry, and biology), AI-assisted invigilation, physical education scoring, and campus video monitoring. Regional education platforms and middle schools face several practical constraints: limited power and space at the edge, the need for high concurrency in multi-classroom environments, and growing requirements for domestic software compatibility. Additionally, schools increasingly demand support for both computer vision-based scoring systems and lightweight generative models that assist teachers with student evaluation and real-time feedback. Traditional GPU solutions often fail to meet these needs in terms of cost, energy efficiency, and system-level integration.

Our Solution: We offer a range of GPGPU products optimized for education-sector AI scenarios. The KS20 and KS28 deliver efficient inference under 50W power budgets and support up to 50–100 concurrent video streams per card, ideal for deployments in AI-assisted science assessments and invigilation. For more demanding applications such as LLM inference and multimodal interaction in smart classrooms, the KS38 offers high memory bandwidth in a compact, edge-ready form factor. All models are fully compatible with the Hamming® Software Stack and domestic OS platforms such as Kylin and UnionTech, supporting integration and long-term ecosystem support. Compared to legacy solutions, our education-specific offerings provide enhanced price-performance and localization, making them ideal for scaled deployment in regional teaching platforms.

Case Study: A leading domestic provider of AI-powered scoring equipment for middle school science exams has adopted our solutions to support large-scale deployment of intelligent grading systems for physics, chemistry, and biology practical exams. Since 2019, provinces across China have begun incorporating lab performance assessments into high school entrance examinations, creating strong demand for scalable, objective, and real-time evaluation methods. Traditional manual grading methods suffer from low efficiency, high labor costs, and inconsistent scoring. By integrating our cards into their scoring terminals, the customer has enabled fast and accurate image-based recognition of students' experimental actions, such as correct use of lab equipment and procedural steps. The system applies AI-driven analysis aligned with preset grading rubrics to generate impartial, real-time scoring outcomes across thousands of candidates.

The solution has already been rolled out in multiple regions and has significantly improved both the fairness and efficiency of exam evaluation. Our hardware platform not only ensures stable and secure operation, but also meets rising localization requirements in China's education equipment sector. Compared to imported alternatives, our products offer higher adaptability, cost-effectiveness, and better alignment with domestic educational needs. We also provide customized technical support and algorithm optimization tailored to science exam scenarios, helping our customer reduce deployment

costs and accelerate adoption. This deployment showcases the transformative role of AI in modern education assessment and demonstrates the value of domestically developed AI infrastructure in delivering scalable, real-world solutions.

Industrial Inspection

Challenges: AI deployment in industrial inspection remains technically complex and highly application-specific. Manufacturers in sectors such as automotive, electronics, and energy require real-time, high-precision defect detection on dynamic production lines. Key pain points include the high cost of traditional AI deployment, fragmented and specialized application scenarios, and difficulty adapting to different product types or defect forms. Vision-based inspection systems must process massive volumes of image data under tight latency constraints while minimizing false positives and supporting continuous model retraining. For safety-critical products — such as lithium-ion batteries — requirements are even stricter: high-speed, image-based inspection must precisely identify surface defects and maintain tight tolerances for electrode alignment, as quality deviations may lead to safety risks or reduced yield.

Our Solution: Our Goldwasser Series offers an edge-to-cloud solution tailored for pixel-level visual inspection across diverse industrial workflows. With support for up to 128GB of memory and high parallel computing density, our products enable real-time processing of high-resolution images captured by line-scan or area-scan cameras. This allows the system to detect micro-defects, such as bubbles, scratches, or edge misalignments, under complex lighting and background conditions. Our solution supports classification of over 10 types of surface defects, along with positional anomaly detection, multi-camera concurrency, and rapid model switching across production lines. Goldwasser accelerators are deployed in smart inspection cabinets and integrated with robotic arms and optical systems to perform continuous, automated inspection with inference-level consistency that exceeds human operators. The system supports a variety of AI frameworks and inference engines, and integrates smoothly with MES and quality feedback systems for yield improvement.

Case Study: A leading manufacturer of lithium battery production equipment has adopted our Goldwasser Series solution in its high-speed vision inspection system for the winding process. During battery cell winding, multiple layers of electrode sheets and separator films are rapidly aligned and spooled under continuous motion. Our solution is deployed at the critical winding stage, where a high-frame-rate line-scan camera captures ultra-fine edge features of each electrode layer in real time. The images are streamed to a GPU-accelerated inference module powered by our KS38 and GS40 cards, where neural networks measure the spacing between electrode and separator layers down to sub-millimeter precision. The system detects alignment drift, burrs, and separator overlap with millisecond response time and feeds real-time feedback into the control system to dynamically adjust mechanical tolerances, significantly improving first-pass yield and reducing defect-related rework.

In a downstream process, the same manufacturer has also implemented our visual inspection platform for the stacking and cutting stages. These stages are critical to the dimensional integrity of pouch cells, where cutting errors or electrode misalignment can lead to edge short circuits or thermal events. Our solution uses structured illumination and multi-angle imaging to detect slicing defects, uneven edges, and contamination particles. Each camera channel is managed through a multi-stream inference pipeline that enables concurrent detection across several inspection points. The system is fully integrated with the customer's MES, allowing for defect traceability and real-time quality reporting at the lot level. By replacing manual inspection with AI-driven automation, the customer has improved defect classification accuracy, enhanced safety compliance, and scaled production capacity without proportional increases in headcount or floor space.

Industrial Large Models

Challenges: As the focus of AI deployment shifts from general-purpose large models to domain-specific industrial models, including both LLMs and large vision models ("**LVMs**"), enterprise and government users increasingly demand computing systems that can support high-concurrency inference, multi-modal workloads, and flexible fine-tuning. These industrial large models are often built on top of general-purpose foundation models (e.g., LLaMA, ChatGLM) and fine-tuned with proprietary

industry data using techniques such as mixture-of-experts (MoE). While these models offer higher task alignment and lower inference cost, they also place significant pressure on computing infrastructure in terms of throughput, memory bandwidth, and deployment flexibility.

Our Solution: We offer a line of integrated appliances based on our Goldwasser Series, designed to support inference and lightweight training of large models in enterprise environments. Each single-node system can be configured with up to 8 Goldwasser accelerator cards and up to 128GB of memory per card. The systems support mixed-precision computation from FP32 to INT4 and are optimized for multi-batch inference. Deployment options range from 1U to 4U servers, compatible with domestic and international CPU platforms. Our appliances are validated against over 30 mainstream models and support MoE scheduling, multi-card coordination, and mainstream fine-tuning toolchains. These turnkey solutions help enterprise and public sector users rapidly deploy industrial large models in areas such as finance, media, education, and public services, accelerating the transition to AI-native operations with improved efficiency and reduced cost.

Case Study: A leading domestic server and AI infrastructure provider has adopted our Goldwasser-based LLM inference system as the core engine for its general-purpose AI appliance offering. This system supports mainstream LLMs, and is designed to serve government, education, and enterprise customers seeking to deploy LLM capabilities across private cloud environments.

The customer uses this platform to offer fine-tuning and inference-as-a-service capabilities for institutional users, enabling downstream clients to run retrieval-augmented generation (RAG), document Q&A, summarization, and text generation workloads. The system supports multiple inference backends and toolchains, such as vLLM, and offers flexible memory partitioning and multi-card task scheduling. Compared to international GPU-based systems, the Goldwasser-powered platform offers enhanced performance-per-watt and competitive total cost of ownership, enabling scalable deployment of LLMs in scenarios where data privacy, local adaptation, and ecosystem compatibility are critical.

System Integration

Challenges: As demand for AI computing accelerates across industries, system integrators and OEMs face growing pressure to deliver AI servers that meet diverse performance, power, and ecosystem compatibility requirements. Key pain points include limitations in GPU form factor flexibility, lack of compatibility with domestic CPU and OS platforms, challenges in system-level thermal and power optimization, and high total cost of ownership associated with mainstream international GPU solutions. Integrators also need to support a range of deployment environments, from compact AI edge boxes to data center-grade 4U rackmount servers, while ensuring long product lifecycle, stable supply, and streamlined software migration.

Our Solution: We offer a highly modular GPGPU product portfolio, optimized for integration across edge-to-cloud AI server environments. Our products are available in multiple form factors, including half-height, single-slot, and full-height dual-slot, supporting diverse chassis and thermal design profiles. All series are fully compatible with domestic CPUs and operating systems (including Kylin, UnionTech, and openEuler), in addition to mainstream international platforms. We provide structured driver, runtime, and toolchain support via the Hamming® Software Stack, ensuring integration with existing CUDA-based development environments. With memory capacities of up to 128GB and power options ranging from 15W to 90W, our products meet a broad range of AI server requirements, from low-power edge inference systems to high-density intelligent computing nodes. Through close collaboration with leading server OEMs and industrial computer providers, we help accelerate system-level deployment, reduce BOM cost, and improve delivery reliability across verticals.

Multi-Scenario Deployments

While our products are designed to address specific enterprise AI application scenarios, namely smart city, enterprise private clouds, education, industrial inspection, industrial large models, and system integration, some customers deploy our solutions across multiple domains. These customers seek unified and flexible platforms capable of supporting diverse use cases within a common architecture, enabling integrated deployment, performance optimization, and long-term ecosystem scalability.

Case Study: We began our collaboration with Company H, a leading industry player in the domain of smart city, where our accelerated AI appliances were deployed for real-time video stream analysis in municipal governance and street-level enforcement. These systems accurately identified events such as illegal street vending and unauthorized parking, significantly improving response efficiency and urban management outcomes. Building on the success of this deployment, Company H rapidly expanded the use of our products into four additional domains:

- *Industrial Inspection* — In wind energy operations, Company H deployed our all-in-one training-inference systems to monitor turbine conditions in real time using vision-based analytics, while performing lightweight fine-tuning on edge devices to improve fault prediction accuracy.
- *Financial Security* — We jointly delivered integrated hardware-software servers for banking scenarios, enabling AI functions such as perimeter protection, facial recognition, and employee behavior monitoring, alongside business process automation.
- *Industrial Large Models* — We helped Company H deploy lightweight multimodal models to reduce inference costs and improve accuracy in complex industrial scenarios previously limited by long-tail distribution issues.
- *System Integration* — Together with Company H, we built an open AI ecosystem platform combining our hardware and their vertical-specific algorithms. Applications include large-model-based API generation tools that significantly enhance development efficiency across their enterprise solutions.

This case illustrates how a leading industry player can benefit from our integrated technology portfolio across multiple sectors, demonstrating the scalability, adaptability, and ecosystem integration capabilities of our platform.

OUR TECHNOLOGIES

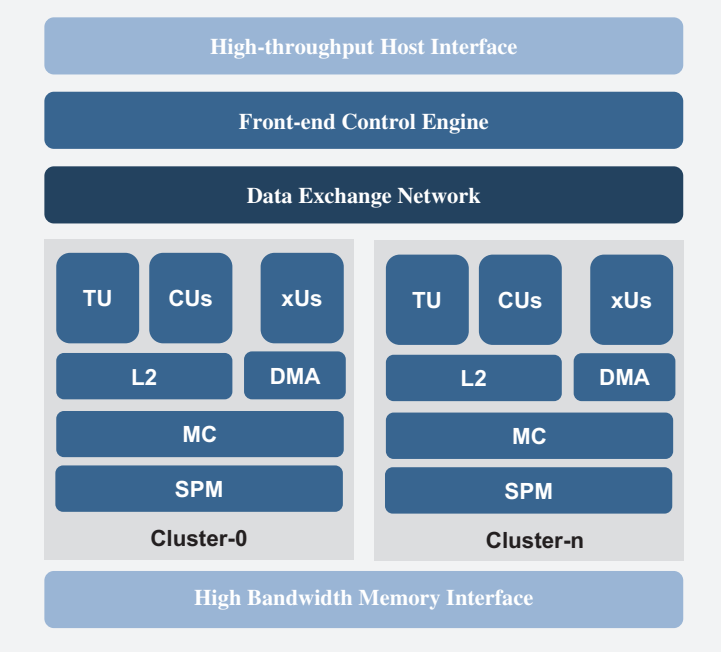
Industry Background — Limitations of Conventional AI Computing Architectures

GPUs have emerged as one of the core enablers of modern artificial intelligence development. While the edge AI domain has seen a proliferation of application-specific integrated circuit (ASIC)-based accelerators, the cloud and data center segments remain dominated by Company N's GPU products, driven largely by their general-purpose programmability and a mature CUDA-centric software ecosystem. However, conventional GPU architectures, originally designed for graphics rendering and high-performance computing, face inherent challenges in AI workloads, including relatively low computing density, limited tensor computation efficiency, and high dependence on external memory bandwidth.

In an attempt to address these limitations, Company N introduced Tensor Cores in its Volta architecture in 2017, boosting performance for tensor operations. However, this did not fundamentally resolve the inefficiencies embedded in its homogeneous design. Concurrently, AI accelerators based on ASIC architectures offer excellent performance for specific operations such as convolutions and matrix multiplications, but suffer from limited programmability. Their reliance on host CPUs or lightweight coprocessors often restricts performance, while the absence of hardware-supported data coherence further complicates system integration.

Traditional heterogeneous computing systems, comprising distinct computing units such as CPUs, GPUs, DSPs, and FPGAs, also encounter significant challenges. These include high data transfer overhead, inefficient resource scheduling, and a lack of a unified memory model, all of which reduce overall system efficiency. The demand for next-generation architectures that integrate high-performance, low-power, and programmable computing engines has never been more pressing.

Our Foundational Architecture — The GPU+ (Minsky™) On-Chip Heterogeneous Architecture



To address the core demands of AI acceleration, namely high computing performance, low power consumption, and software ecosystem compatibility, we have independently developed our proprietary GPU+ architecture (Minsky™), a software-defined, on-chip heterogeneous computing framework. Unlike conventional discrete heterogeneous systems that rely on external components (e.g., CPUs, GPUs, and FPGAs), our GPU+ (Minsky™) architecture tightly integrates programmable general-purpose computing cores, tensor-centric AI accelerators, and a high-bandwidth on-chip memory subsystem within a single chip, enabling fine-grained co-optimization of hardware and software.

At the heart of this architecture is a dual-engine computing design built around two complementary components:

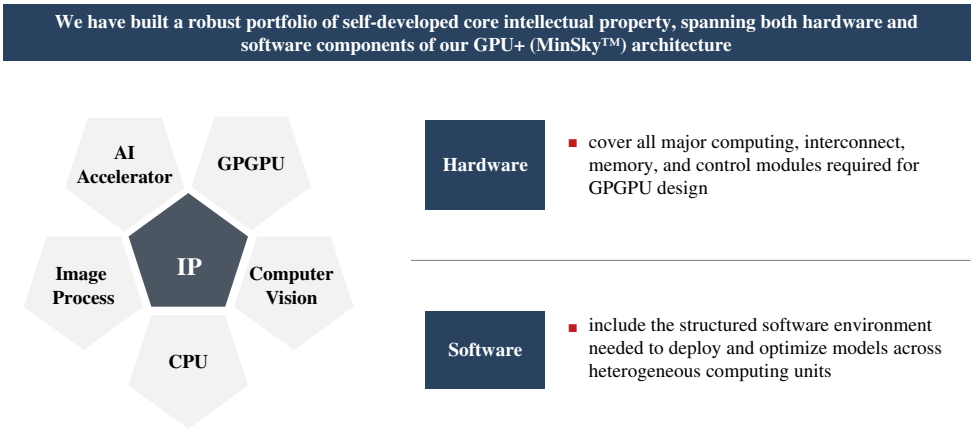
- The first is a tensor computing engine optimized for the structured operations that dominate AI workloads (e.g., matrix multiplication and convolution). This specialized hardware directly maps to the multi-dimensional structure of tensor data and the deterministic logic of AI operators, avoiding inefficiencies common in general-purpose cores and significantly improving computational efficiency.
- The second is our proprietary CU, which inherits and develops the Single Instruction, Multiple Thread (SIMT) parallelism model found in modern mainstream GPGPU. The CU natively supports mainstream programming languages such as CUDA and OpenCL, allowing developers to migrate existing algorithms with minimal code change. This ensures that AI workloads, whether based on standard operators or custom logic, can be deployed with ease and performance predictability.

All computing engines are interconnected through a high-speed on-chip network fabric and communicate via a large-capacity on-chip memory system. Unlike traditional GPU designs that rely heavily on off-chip HBM, our architecture features a software-managed memory hierarchy that drastically reduces data movement latency and power overhead. This integrated memory fabric addresses the well-known "memory wall" problem and enables efficient parallel execution across high-throughput AI tasks.

The GPU+ (Minsky™) architecture also supports modular extension of fixed-function accelerators, including image signal processing (ISP), video encoding/decoding, and lossless compression units. These modules are managed through unified scheduling and integrated via the same on-chip interconnect. As a result, our platform delivers the efficiency of accelerators alongside the flexibility and ecosystem compatibility of general-purpose GPUs, allowing developers to target a broad spectrum of computing-intensive tasks ranging from AI training and inference to image processing, signal analysis, and scientific computing.

Together, this unified architecture allows our GPU+ (Minsky™) architecture to adapt to a wide range of deployment scenarios, from LLM inference in cloud data centers to ultra-low-latency edge computing at the terminal. It delivers competitive performance per watt, excellent scalability, and full compatibility with existing software ecosystems, enabling customers to deploy AI with lower total cost of ownership and faster time-to-solution across evolving AI application domains.

Self-Developed and Continuously Evolving Core IP Portfolio



Built upon our proprietary GPU+ (Minsky™) architecture, we have developed a structured and fully self-owned portfolio of core intellectual property (IP), covering a broad range of computing and acceleration functionalities. Unlike IP providers such as ARM or Imagination Technologies, whose GPU IPs are typically designed for mobile applications and lack the performance necessary for high-end cloud workloads, we focus on cloud-scale, high-performance applications that demand architectural integration and optimization at scale.

We believe there is no shortcut to delivering scalable AI computing through IP aggregation alone. Instead, we have adopted a vertically integrated approach to IP development, enabling us to tightly couple computing units, memory subsystems, and task scheduling within a unified framework. This strategy allows us to align R&D priorities with market needs, maintain technology advantage, and reduce dependencies on external suppliers. Moreover, as a domestic company in China, our self-developed IPs help shield us from potential disruptions arising from changes in the international trade environment.

Through successive generations of product development, we have built a robust portfolio of self-developed core intellectual property, spanning both hardware and software components of our GPU+ (Minsky™) architecture.

On the hardware side, our IP covers all major computing, interconnect, memory, and control modules required for GPGPU design. Key in-house developments include:

- Tensor computing engines comprising over 1,000 parallel MAC units;
- Precision-adaptive low-bit accelerators supporting INT4 and FP8 with sub-2ns switching latency;

- Sparse computation modules with >95% zero-skipping efficiency under 70% sparsity conditions;
- A unified HBM2E memory subsystem supporting up to 128GB and 1.2TB/s bandwidth;
- A hierarchical on-chip network with 256GB/s bandwidth and sub-50ns latency;
- Hardware fault tolerance units supporting physical fuse and 8-path redundancy;
- RISC-V dual-core task schedulers with 1.5GHz frequency and 128-bit vector extensions.

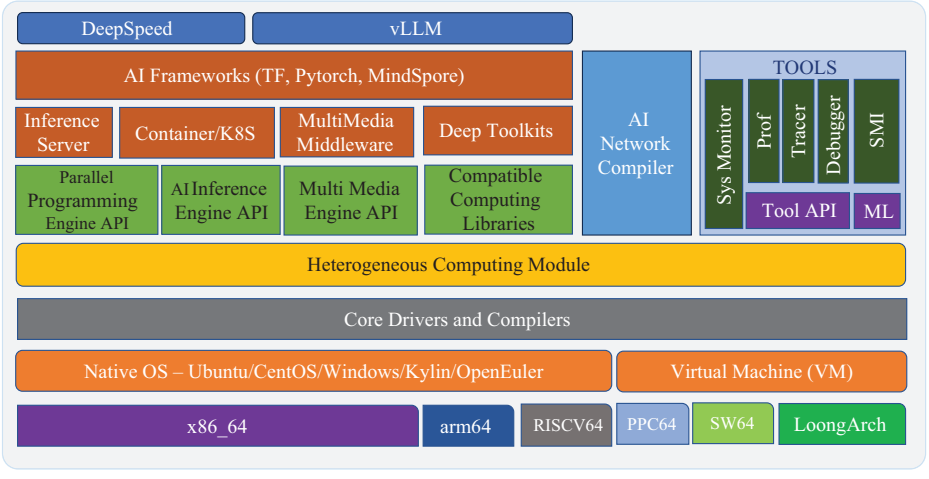
On the software side, our IP portfolio includes the structured software environment needed to deploy and optimize models across heterogeneous computing units. These include:

- A CUDA/OpenCL compatibility layer based on LLVM IR, covering over 90% of standard APIs;
- The DLNNE inference engine with support for trillion-parameter model inference under 100ms latency;
- A sparse model compression engine with 60% memory footprint reduction, coordinated with hardware;
- Operator fusion optimizers reducing kernel launch overhead by over 60%;
- A dynamic task scheduler with <10ns thread-switching latency, integrated with hardware time-stamping;
- A virtualization interface supporting up to 8 vGPU slices per physical card;
- LLVM toolchain plugins for automatic TensorCore code generation.

Together, these IP components form the basis of our independently developed GPGPU computing platform, supporting our roadmap across training, inference, and edge computing scenarios. This modular and extensible IP platform enables us to rapidly assemble computing products tailored for diverse use cases, ranging from edge inference cards to high-throughput training accelerators. By leveraging internally developed IPs and tightly integrating hardware and software, we can deliver energy-efficient solutions optimized for real-world AI workloads.

Unified Software Platform

Overview of Our Software Stacks



To complement our self-developed hardware, we have built a fully integrated software platform, the Hamming® Software Stack, that mirrors the structure of Company N's CUDA ecosystem while being specifically optimized for our architecture. Hamming® is designed to ensure maximum compatibility with widely adopted open APIs, enabling migration of AI applications and accelerating customer adoption.

The major software modules in the Hamming® Software Stack include:

- ***Open-source application frameworks*:** We provide optimized versions of widely used CUDA-ecosystem open-source AI frameworks such as DeepSpeed, vLLM, PyTorch, MindSpore, and TensorFlow. These retain functional parity with their community versions, but are optimized for our hardware.
- ***Inference and container system layer*:** We offer a networked inference system based on inference servers and Kubernetes container orchestration, functionally aligned with Company N's Triton inference server. We also support mainstream multimedia middleware such as GStreamer and FFmpeg, both enhanced to utilize our hardware acceleration capabilities.
- ***Deep Toolkits*:** We provide a fully custom lightweight software framework — Deep Toolkits — which integrates video decoding, preprocessing, and AI inference functions, and is released as open source to assist customers with application system integration.
- ***Drivers and basic CUDA-equivalent libraries*:** We supply our own implementation of CUDA-compatible drivers and libraries, including cuDNN, cuBLAS, cuFFT, and others, with API-level compatibility but optimized backend implementations tailored for our hardware. Our OpenCL drivers also comply with industry standards.
- ***Network inference engine*:** Our proprietary inference engine supports the optimization and accelerated execution of AI inference workloads on our products. Its APIs are similar to those of Company N's TensorRT, but implementations are optimized specifically for our GPU+ (Minsky™) architecture.
- ***Video and image codec drivers*:** These include video and JPEG codec APIs that are self-defined and directly interface with our onboard hardware accelerators for fast encoding and decoding.
- ***Developer tools*:** We provide a toolchain, including a device management library (compatible with Company N's NVML APIs), a graphical and CLI management tool, and profiling/debugging tools.
- ***Low-level drivers and runtime environment*:** These modules are not directly exposed to users but serve as the foundation of our heterogeneous computing support. They comprise a custom heterogeneous computing runtime with unified APIs for GPU and tensor processing unit (TPU) management, a hardware abstraction layer and kernel drivers tailored for our architecture, and support for a broad range of domestic and international CPU architectures, including x86, ARM, and RISC-V, and operating systems, including Kylin, UnionTech, Euler, Red Hat, and Windows.

Our unified software platform is designed around a software-defined architecture that ensures integration between software and hardware at every stage — from design to deployment. Starting from real-world AI application requirements, we break down each use case into functional modules and quantify their computing demands. These demands are then mapped to chip-level architectural configurations, including the mix and scale of Computing Units (CUs), Tensor Units (TUs), and codec engines, as well as deeper design factors such as the ratio of matrix to vector engines inside TUs and the allocation of computing precision formats (e.g., FP32, FP16, INT8).

Once the system is deployed, the Hamming® software stack orchestrates runtime execution through a tightly coupled four-step process. First, the unified driver automatically loads and identifies the hardware configuration, detecting available computing engines and memory structures. Second,

incoming AI workloads are decomposed into sub-tasks, distinguishing between tensor-heavy and control-intensive operations. Third, the platform dispatches these sub-tasks to the appropriate computing engines (e.g., convolutional layers in CNNs are routed to TUs), while recurrent structures in RNNs or scalar-heavy tasks are directed to CUs. Finally, the system fuses the results, aligning outputs to the desired precision (typically FP32) to ensure accuracy and consistency. This software-hardware co-design enables efficient execution with minimal developer intervention and allows our platform to scale across diverse deployment environments, from edge inference to large-scale AI training clusters.

SPECIALIST INDUSTRY AND ACCEPTABLE SECTORS

We are primarily engaged in the research, design, and development of GPGPUs, and have developed a proprietary on-chip heterogeneous chip architecture, GPU+ (Minsky™), together with a fully self-developed and vertically integrated software stack, the Hamming® Software Stack. Our business spans the entire logic and physical design process of semiconductor products, encompassing architecture definition, front-end and back-end development, verification, and system-level performance validation. In addition, we have built a structured suite of low-level drivers, CUDA-equivalent libraries, and AI inference toolchains that enable real-time, high-throughput AI workloads.

Based on the nature of our technology and business model, we believe we fall within the following Specialist Technology and Acceptable Sectors:

Specialist Technology Products	Specialist Technologies	Acceptable Sectors
GPGPU and related technical services	Semiconductors	Advanced hardware and software — Semiconductors — Design: logic and physical design, and validation and verification

Our Directors are of the view that our business falls within the acceptable sector of "advanced hardware and software — Semiconductors — Design: logic and physical design, and validation and verification," because our core activities encompass the full-cycle design and development of GPGPUs, including architecture definition, front-end logic design, back-end physical design, and multi-level functional and system validation, together with software that supports and enables the operation and verification of our chips. From an industry perspective, Frost & Sullivan considers that developers of GPGPUs are classified within the semiconductor design segment, because the development of such products requires advanced logic and physical design processes and extensive verification consistent with industry practice. Having considered the nature of our technology and development workflow, the Joint Sponsors concur with the views of our Directors and Frost & Sullivan and are of the view that our business appropriately falls within this acceptable sector under Chapter 18C of the Listing Rules.

COMMERCIAL DEVELOPMENT OF OUR PRODUCTS

We have achieved initial commercial traction through the deployment of our Minsky™ architecture-based GPGPU products across a wide range of AI computing use cases spanning cloud, edge and terminal environments. These deployments highlight our ability to meet rising demand for domestically developed, high-performance computing solutions nowadays. Our solutions offer both power efficiency and cost-effectiveness. Our commercialization strategy emphasizes mainstream ecosystem compatibility and differentiated product performance, targeting high-growth AI verticals such as smart cities, intelligent vehicles, enterprise data centers, industrial automation, healthcare, and education.

Leveraging our proprietary GPU+ (Minsky™) architecture and the Hamming® software stack, we have established commercial relationships with leading enterprises, system integrators, and solution providers in China. We are scaling the adoption of our major products through a phased rollout approach, while concurrently advancing commercialization efforts for next-generation products in our Minsky™ roadmap.

Sales Volume

The following table sets forth the sales volume of our major GPGPU products for the periods indicated:

	For the year ended December 31,			For the four months ended April 30,	
Product	**2023**	**2024**	**2025**	**2025**	**2026**
	Units	*Units*	*Units*	*Units*	*Units*
Goldwasser I	3,480	1,998	2,759	713	91
Goldwasser II	719	3,297	2,985	100	7
KS Series	—	—	6,400	—	10,246
Total	4,199	5,295	12,144	813	10,344

The overall sales volume of our GPGPU products increased during the Track Record Period, driven primarily by an expanding customer base and growing demand from enterprise AI application scenarios. The sales volume of Minsky™ I-based Goldwasser I products declined in 2024 compared to 2023, mainly because customers shifted to purchasing the upgraded Minsky™ II-based Goldwasser II products. In 2025, sales volume of the earlier-generation Goldwasser series remained relatively stable overall despite the introduction of the KS Series. Goldwasser I sales volume in 2025 increased compared to 2024, primarily because (i) certain smart city customers experienced increased demand for image-comparison applications which the Goldwasser I products were well-suited to meet, and (ii) following the launch of the KS Series, certain customers placed one-off orders for Goldwasser I products for a total of 1,077 units, with an aggregate value of approximately RMB6.1 million, to complete their then-existing project deployments before migrating to the KS Series for subsequent projects. As of the Latest Practicable Date, the delivery of Goldwasser I products to such customers had been completed. Subsequent to the Track Record Period and up to the Latest Practicable Date, we have secured new orders for over 2,000 units of the Goldwasser I products, with a contract value exceeding RMB6.0 million. Goldwasser II sales volume in 2025 declined slightly compared to 2024, reflecting the gradual transition of certain customers to the KS Series. The KS Series, our first commercial products based on the Minsky™ III architecture, recorded sales volume of 6,400 units in 2025, becoming a key contributor to total sales volume during the year. In the four months ended April 30, 2026, the sales volume of the Goldwasser I and Goldwasser II products declined compared to the same period in 2025, while the KS Series sales volume increased in the four months ended April 30, 2026 as compared with full year 2025, reflecting the gradual transition of certain customers to the KS Series.

The KS Series is intended to progressively replace both the Goldwasser I and Goldwasser II Series, and we currently expect to discontinue mass production of the Goldwasser I (launched in April 2021) and Goldwasser II (launched in June 2023) Series by the end of 2026, resulting in a co-existence period of approximately 21 months with the KS Series. We currently expect the KS Series' commercial life cycle to extend to approximately 2030, and plan to launch the first and second phases of our Minsky™ IV products in December 2026 and the second half of 2027, respectively, each with an expected commercial life cycle of approximately five years. To manage product transition risk, we generally notify customers of upcoming discontinuations at least nine months in advance, coordinate implementation timelines with each customer's product life cycle, and continue to provide after-sales support to existing customers during their standard warranty periods. Having regard to the foregoing, our Directors are of the view that our Group is not exposed to any significant product obsolescence risk in connection with the transition between product generations.

Average Selling Prices

The following table sets forth a breakdown of the average selling prices of our major GPGPU products for the periods indicated:

	For the year ended December 31,			For the four months ended April 30,	
Product	**2023**	**2024**	**2025**	**2025**	**2026**
	RMB'000	*RMB'000*	*RMB'000*	*RMB'000*	*RMB'000*
Goldwasser I	5.7	5.2	5.1	5.2	5.6
Goldwasser II	9.9	16.9	9.0	20.8	17.8
KS Series	—	—	11.7	—	16.2

The price range of our major GPGPU products largely remained stable during the Track Record Period, reflecting consistent product positioning across different power and performance tiers. The average selling price of our Goldwasser II products increased in 2024 compared to 2023, primarily due to a higher proportion of customized configurations delivered to certain customers. These configurations primarily involved customization of our Hamming® software stack and toolchain (including dynamic and static sparsity support, operator-level performance optimization and AI framework adaptation) on top of our standard hardware platform, undertaken principally for one enterprise customer to enable enhanced performance in emerging large-model inference workloads. Pricing was determined on a cost-plus basis with reference to the incremental R&D investment and the value of the customization to the customer, resulting in an approximately 80% mark-up over comparable standard-configuration products.

In 2025, the average selling price of Goldwasser II products declined compared to 2024, primarily due to (i) the entry of competing products of a leading PRC peer into the qualified supplier list of certain of our key customers, which increased customers' bargaining power on pricing in respect of these products; (ii) our strategic decision to offer pricing concessions to certain customers who placed sizeable, multi-product orders of strategic importance to our market position; and (iii) the gradual transition of customer demand from the Goldwasser II series towards the KS Series as part of our normal product roadmap progression. Overall, our pricing strategy is to benchmark our products against comparable offerings from leading domestic and international GPGPU suppliers, while granting selective pricing concessions to support large or strategically important orders. Further, we expect to continue to face pricing pressure on our products. If we are unable to develop, launch and commercialize new or upgraded products in a timely manner to keep pace with leading market participants, our path to commercialization and profitability may be materially and adversely affected. See "Risk Factors — Risks Relating to the Commercialization of Our Products — We face substantial competition and continuing pricing pressure, which may materially and adversely affect our path to commercialization and profitability" for further details.

In the four months ended April 30, 2026, the average selling price of Goldwasser II products declined compared to the same period in 2025, primarily due to a higher proportion of bare dies sold to our customers with lower selling prices compared to the same period in 2025, when a higher proportion of higher-priced products were sold. The increase in average selling price of Goldwasser II from RMB9.0 thousand in 2025 to RMB17.8 thousand in the four months ended April 30, 2026 was primarily influenced by certain individual orders with higher specifications and does not reflect long-term pricing trends. However, as we only shipped a small amount of Goldwasser II in the four months ended April 30, 2026, its average selling price was susceptible to those high-specification orders. The average selling price of KS Series products in the four months ended April 30, 2026 increased compared to that in 2025, primarily due to a higher proportion of our more advanced KS Series products such as KS38 sold to our customers, which have higher selling prices.

At the same time, the introduction and scaling of the KS Series products in 2025 contributed to diversification of our overall average selling price profile across different performance tiers. According to Frost & Sullivan, the ranges of selling prices of our products during the Track Record Period are in line with the industry.

PATH TO COMMERCIALIZATION

We are a Pre-Commercial Company as defined under Chapter 18C of the Listing Rules at the forefront of developing GPGPU products and providing related technical services for AI and high-performance computing applications. Our commercialization strategy is designed to capture the substantial growth opportunities in China's rapidly expanding GPGPU market while laying a foundation for sustainable revenue generation and margin improvement. Our business strategies and development plans are aligned with prevailing and expected trends in enterprise AI computing, including increasing demand for localized, cost-efficient and energy-efficient computing infrastructure, wider adoption of heterogeneous computing architectures, and expanding AI use cases across enterprise, industrial intelligence and edge computing scenarios. Consistent with sector practices among emerging GPGPU providers, we adopt a staged commercialization approach that progresses from product validation and early enterprise adoption to broader commercial deployment and scale-up, focusing on application-driven demand and ecosystem integration. According to Frost & Sullivan, the PRC AI infrastructure market is expected to grow at a double-digit CAGR over the next five years, driven by surging demand for training and inference of large language models, accelerated enterprise digital transformation, and increasing deployment of AI applications in sectors such as finance, energy, and robotics. Positioned with our advanced architecture, fabless model, and growing customer base, we believe we are well placed to commercialize our technologies and products.

Scaling GPGPU Products

Our commercialization roadmap has been driven by continuous upgrades of our proprietary GPU+ (Minsky™) architecture. Since 2020, we have advanced from Minsky™ I, which powered our Goldwasser I products, to Minsky™ II, underpinning Goldwasser II products launched in 2023 with enhanced throughput and memory capacity. In 2025, we launched the KS Series, the first wave of Minsky™ III-based products, which delivered meaningful improvements in computing density, energy efficiency and ecosystem compatibility. The KS Series addresses the growing demand for localized and secure enterprise AI infrastructure. In 2025, we completed our supply chain transition by optimizing outsourced manufacturing partners and switching to a local supplier to enhance production efficiency and supply chain security, thereby streamlining wafer fabrication and the corresponding packaging and testing processes, and consequently expanding our supply chain capabilities. This transition has reduced our exposure to supply uncertainties and shortened our logistics cycles by nearly 50%, contributing to a more stable and predictable lead time for our products. Production yield levels following the transition have been broadly in line with our prior experience, with no material impact on our cost structure. With supply chain transition now completed, shipment volumes of the KS Series ramped up in the second half of 2025 and are expected to continue throughout 2026.

Minsky™ III represents a broader product generation beyond the KS Series. While we have prioritized bringing the KS Series to volume production as the initial Minsky™ III products, additional offerings based on the Minsky™ III architecture are under development and will further expand our product coverage across enterprise AI verticals. Looking ahead, we are developing Minsky™ IV products, which we plan to commercialize in two phases. The first phase of Minsky™ IV products, which we expect to launch in December 2026, will utilize LPDDR memory and is designed to deliver approximately twice the computing performance of the KS Series under comparable energy consumption, providing enterprise AI customers with improved total cost of ownership. The second phase of Minsky™ IV products, expected to be launched in the second half of 2027, will incorporate PRC domestically produced advanced HBM3e memory together with die-to-die (D2D) interconnect and advanced packaging technologies, and is targeted to deliver approximately four times the peak computing performance of the LPDDR-based first-phase products. Architecturally, Minsky™ IV represents an iterative extension of our proprietary GPU+ (Minsky™) architecture, building on the mature designs of the prior three generations to deliver enhanced single-card computing power, broader native support for additional data formats, and differentiated memory and interconnect configurations across the first- and second-phase products. The first-phase Minsky™ IV products, configured with LPDDR memory, are expected to deliver computing power of up to 1,200 TOPS (INT8) at the board level, as compared with a maximum of 560 TOPS (INT8) for the KS Series, enabling us to address enterprise AI applications with higher computing-density and performance requirements within our

existing enterprise AI verticals. The second-phase products are intended to serve large-model training and inference workloads at enterprise AI customers and intelligent computing centers, and are expected to deliver FP8 computing power of up to 2,000 TFLOPS, with target customers including operators of large-scale intelligent computing centers and public cloud providers. Based on our preliminary design specifications and according to Frost & Sullivan after reviewing publicly available information about such peer companies, including but not limited to their annual reports, product manuals and official websites, the second-phase Minsky™ IV products are expected to deliver comparable FP8 computing performance and memory capacity, together with approximately 1.7 times the memory bandwidth, relative to the corresponding flagship product of a leading international GPU manufacturer, and approximately 2.0 times the FP8 computing performance and memory bandwidth, together with approximately 1.3 times the memory capacity, relative to a corresponding product offered by a leading PRC cloud accelerator card provider. The phased commercialization approach is designed to manage potential volatility in pricing and availability of advanced PRC domestic memory components. Like the KS Series, the Minsky™ IV products will continue to rely on our PRC domestic supply chain. We do not expect the launch of the Minsky™ IV products to replace, or materially reduce demand for, the KS Series. Rather than serving as a substitute, Minsky™ IV extends our product portfolio into higher power and performance tiers above those served by the KS Series, thereby expanding our addressable market, while the KS Series is expected to continue to serve mainstream enterprise AI scenarios throughout its expected commercial life cycle to approximately 2030. The two product lines are positioned to serve distinct application scenarios: the KS Series addresses mainstream enterprise AI deployments such as smart city, industrial inspection and enterprise private cloud, while the Minsky™ IV products target higher computing-density workloads such as large-model inference and training at intelligent computing centers and, over time, public cloud providers. We accordingly expect the two product lines to be complementary, with the Minsky™ IV products capturing incremental demand in market segments that the KS Series does not address, rather than cannibalizing sales of the KS Series. This staged progression of the Minsky™ roadmap allows us to continuously optimize performance and efficiency, deepen customer adoption across enterprise AI and position ourselves for longer-term opportunities in the public cloud market.

Our scale-up of GPGPU products is reflected in measurable commercialization metrics. Total GPGPU sales volume increased from 4,199 units in 2023 to 5,295 units in 2024, reflecting the transition from the Goldwasser I series to the Goldwasser II series. In 2025, total sales volume further increased significantly to 12,144 units, primarily driven by the KS Series, the first wave of Minsky™ III-based products, which recorded sales volume of 6,400 units and became a key contributor to overall sales volume. During the same year, sales volume of the earlier-generation Goldwasser series remained relatively stable overall despite product iteration, with sales volume of Goldwasser I products increasing to 2,759 units and sales volume of Goldwasser II products declining slightly to 2,985 units year-on-year, resulting in an overall increase in total sales volume of the Goldwasser series compared to 2024. The total sales volume continued to grow from 813 units in the four months ended April 30, 2025 to 10,344 units in the four months ended April 30, 2026, primarily driven by the KS Series. Compared to the four months ended April 30, 2025, the sales volume of Goldwasser series continued to decline in the four months ended April 30, 2026 with sales volume of Goldwasser I products decreasing from 713 units to 91 units and sales volume of Goldwasser II products decreasing from 100 units to seven units, while the KS Series sales volume increased from nil in the four months ended April 30, 2025 to 10,246 units in the same period in 2026, reflecting the gradual transition of certain customers to the KS Series. During the Track Record Period, the majority of customers for our GPGPU products were private companies, with the balance comprising public companies. To support continued scale-up, we have placed orders to secure core wafer capacity resources equivalent to approximately 1,400 wafers, and are expanding packaging and testing capacity as well as PCB assembly capabilities through third-party partners.

Driving Commercialization Through Enterprise Adoption

Our enterprise AI customers procure our GPGPU products and related computing solutions primarily for the deployment of artificial intelligence applications within their own businesses, rather than for the operation of public computing infrastructure services to be leased to third parties (which is

the typical use case for public cloud AI providers). Enterprise AI applications span across various industry verticals, including: (i) smart city, where AI is used for traffic management, public safety and environmental monitoring applications such as image recognition, multi-modal recognition and behavioral analysis; (ii) enterprise private cloud, where AI is deployed within an enterprise's internally-managed private infrastructure to support workflow automation, internal knowledge management, intelligent meeting and document workflows; (iii) industrial inspection, where AI is used for high-precision quality testing, production flow control and intelligent equipment monitoring; (iv) education; and (v) industry-specific large model deployment. Compared with public cloud AI, enterprise AI customers typically place a higher emphasis on data security, real-time responsiveness, energy efficiency and total cost of ownership, and often have stricter requirements on the configuration of computing infrastructure within their organizations.

Our path to commercialization is underpinned by favorable industry dynamics. The adoption of GPGPUs in China is accelerating as enterprises seek domestic solutions to mitigate supply chain risks associated with foreign suppliers, a trend further reinforced by government policies that promote indigenous innovation in semiconductors. As an early mover in GPGPU technologies, with differentiated architecture and proven design service capabilities, we are well positioned to capture this demand shift and strengthen our role as a key domestic provider.

Our commercialization efforts are currently centered on the enterprise market, where our products have been deployed across multiple industry verticals, including smart city, enterprise private clouds, education, and industrial inspection. Each of these verticals requires a distinct balance of computing density, energy efficiency, and ecosystem compatibility, which our GPU+ (Minsky™) architecture is suited to provide. A representative example is our collaboration with Company H, where our products have been deployed across multiple scenarios ranging from large-model RAG-based enterprise knowledge management to multimodal AI applications. This multi-scenario adoption showcases the adaptability of our platform and serves as a strong reference case for other industry players.

Looking ahead, we plan to further deepen our penetration into enterprise verticals by strengthening collaboration with leading players and scaling adoption across their downstream ecosystems. As industrial large models increasingly permeate enterprise workflows, bringing about a future of "LLM everywhere," we believe our technology will serve as a scalable, cost-effective, and secure computing backbone, driving broader commercialization across the enterprise AI landscape.

Over the next 12 to 36 months, we plan to expand our product coverage in these verticals through the staged rollout of our Minsky™ IV products, which are designed to address both edge-oriented scenarios, meaning deployments at or near the locations where data is generated, which typically require low power consumption and compact form factors (such as industrial inspection, logistics and part of smart city scenarios) and higher-performance scenarios, meaning deployments in data centers or dedicated computing clusters that require substantially higher computing throughput and memory capacity (such as large-model inference and training), while maintaining a consistent software stack across product generations. To drive adoption of this expanded coverage at scale, we collaborate with market-representative enterprise customers and ecosystem partners through three complementary mechanisms: (i) joint solution development, under which we integrate our GPGPU products with the customer's domain-specific software, datasets and workflows to form a deployable end-to-end solution, supported, where appropriate, by limited intellectual property licensing or limited customization services based on our mass-production products; (ii) co-validation, under which we and the customer jointly benchmark the solution against the customer's reference metrics for performance, energy efficiency and total cost of ownership, generating validation results that support subsequent procurement decisions within the customer's organization and across the same vertical; and (iii) application-level optimization, under which we tune our hardware configuration, compiler stack and software tool chains to the customer's workload characteristics so that evaluation-phase performance can be reproduced in production.

As a representative example, we have collaborated since 2022 with the end customers of Customer A, which are subsidiaries of a leading provider of vision-based solutions, across more than ten rounds of hardware and software iteration, which has enhanced the suitability and software maturity of the first

three generations of our Minsky™ products in vision-centric smart city scenarios. Customer A, which served as our sole distributor during the Track Record Period, acted as the procurement and distribution channel through which such end customers' orders, deliveries and settlements were effected in accordance with their internal procurement requirements, and the technical collaboration on hardware and software iteration described above was conducted directly between our engineering team and such end customers. During the Track Record Period, this collaboration generated revenue of approximately RMB29.1 million for our Group recognized through our sales to Customer A as the distribution channel. More broadly, by partnering with enterprises that have leading technology capabilities and market influence in their respective verticals and by incorporating their guidance on board-card design, software functionality and future optimization priorities, we develop reference solutions that we expect to translate into repeat deployments within the same customer, broader adoption across the relevant vertical and increased demand for our Minsky™ IV and other GPGPU products. We further believe that demand for enterprise private cloud AI deployments is increasing, as enterprise customers increasingly seek to retain their sensitive data and proprietary models within their own infrastructure for data security and regulatory compliance reasons, as well as for latency, customization and total cost of ownership considerations. As our GPGPU products are well-suited to such private cloud and on-premise AI deployments, we expect this trend to further support demand for our products across our target enterprise AI verticals.

Expanding into Strategic Markets of Public Cloud and Robotics

In addition to our enterprise AI focus, we intend to selectively explore opportunities in the public cloud and robotics markets, which we believe represent substantial longer-term growth opportunities. As of the Latest Practicable Date, we have not entered into any signed contracts or binding cooperation arrangements with public cloud customers, and our exploratory activities consist primarily of early-stage technical discussions and product roadmap development. China's public cloud market has been growing rapidly, driven by surging demand from large-model training and intelligent computing centers, and is expected to continue its robust expansion in the coming years. At the same time, the public cloud market is currently served by established participants and is highly concentrated. According to Frost & Sullivan, NVIDIA, T-Head, AMD and Hygon together accounted for over 90% of China's public cloud AI computing market in 2025 by relevant product revenue, supported by mature software ecosystems, established customer bases and large-scale cluster deployment experience. New entrants also face high technical, ecosystem and customer-validation barriers, as public cloud deployments require large-scale cluster interconnection and scalability as well as compatibility with mainstream operating systems, AI frameworks and cloud platforms, and cloud service providers typically conduct lengthy qualification and validation of new suppliers, which may lengthen our customer-acquisition cycle in this market. While our commercialization efforts have historically focused on enterprise AI, we already possess cost-efficient products with the performance and scalability required for public cloud deployments. Our entry into the public cloud market is based on incremental product adaptation rather than a fundamental shift in strategy. We are enhancing product compatibility and certification, improving virtualization support and optimizing computing resource scheduling efficiency to meet the elastic computing requirements of public cloud providers, while leveraging our existing architecture and software stack.

In parallel, we intend to selectively explore opportunities in the robotics market, which we believe is experiencing favorable momentum in China underpinned by intelligent manufacturing, logistics automation and AI-driven physical interaction. Our exploratory focus in the robotics market is on selected sub-segments, including industrial robots, service robots and embodied intelligence robots, with an initial emphasis on industrial inspection scenarios. We similarly expect to face competition in the robotics market, where, according to Frost & Sullivan, overseas participants such as NVIDIA, Tesla and Intel, together with certain PRC SoC providers, hold leading market positions by virtue of their first-mover advantages in chip performance, software ecosystems and application adaptation. Robotics computing chips must also balance computing performance, real-time responsiveness and energy efficiency within constrained size, power and thermal envelopes, and typically require customized adaptation to, and lengthy validation with, different robot platforms and application scenarios, which we expect to lengthen our development and customer-engagement cycles in this market. Our roadmap

includes developing new architectures and system-on-chip (SoC) designs tailored for robotics applications. By integrating selected IP modules with our existing architecture platform, we aim to provide efficient, low-power computing engines that can support real-time perception, decision-making, and control in diverse robotics scenarios. If we are successful in expanding into the public cloud and robotics markets, we expect to diversify our commercialization channels and capture additional revenue streams over time, although there is no assurance that such expansion will be achieved within the timeframe we currently contemplate or at all.

Technical Services Relating to GPGPU Products

In addition to product sales, we generate revenue from technical services relating to GPGPU products, which draw on our design expertise and proprietary intellectual property. This business line primarily consists of chip design services, where we work with customers to customize architectures, optimize performance and adapt our products to their specific deployment scenarios, as well as selective IP licensing. Since 2023, we have started delivering such services, including a back-end chip design services project for Customer H in 2025 that contributed meaningfully to revenue during the period. Customer H, incorporated in July 2020 with a registered capital of RMB10 million and headquartered and operating in Beijing, is a national-level new-type scientific research institution established under the auspices of a leading PRC university with expertise in frontier, strategic and disruptive research in intelligent technologies. We commenced our engagement with Customer H through a non-public procurement process. The project was accepted and completed in May 2025, and accordingly no renewal is contemplated. The salient terms of our service agreement with Customer H, being our largest technical services customer during the Track Record Period, are summarized below:

- *Tenure:* the contract was effective from signing in November 2024 through December 31, 2025, followed by a quality warranty period effective from the date of service acceptance till June 30, 2026.
- *Rights and obligations:* we were required to deliver in accordance with the agreed technical specifications and to perform remedial work during the warranty period; Customer H was required to conduct acceptance within the contractually prescribed period.
- *Credit terms and payment milestones:* the contract fee was payable in two installments: a prepayment shortly after signing and the balance shortly after acceptance.
- *Ownership of intellectual property:* the intellectual property rights in the deliverables, and in any new technical achievements developed by us in performing the contract, vested in Customer H.
- *Termination:* Customer H was entitled to terminate and require refund of amounts paid, together with liquidated damages, on specified defaults by us; either party was entitled to terminate on force majeure rendering the purpose of the contract incapable of being achieved.

Customization is particularly valuable because different industries and application scenarios have distinct requirements for computing performance, power efficiency, cost, and software ecosystem compatibility. Through customized design, we can optimize hardware architectures for specific computational tasks, such as by adjusting the configuration of computing units, or integrating acceleration modules IPs, thereby improving performance efficiency while reducing overall power consumption and cost. This approach allows customers to overcome the limitations of general-purpose chips, accelerate time-to-market for solutions aligned with their product roadmaps, and enhance their competitiveness.

Customer H was not our sole customer under the technical services segment during the Track Record Period, and we also entered into other technical or technology-related service or licensing arrangements with additional customers during the same period, albeit with smaller contract sizes. We have been actively expanding our customer base for technical services by engaging with additional enterprise, industrial and institutional customers across different application scenarios. During the Track

Record Period, we entered into nine chip design services projects and IP licensing agreements with our customers, with an aggregate contract amount of approximately RMB150.6 million, of which approximately RMB106.5 million had been recognized as revenue during the Track Record Period. Of the above, (i) six chip design services projects had an aggregate contract amount of RMB60.8 million, among which RMB1.0 million, RMB8.7 million, RMB43.3 million and RMB0.8 million was recognized as revenue in 2023, 2024, 2025 and the four months ended April 30, 2026, respectively; and (ii) three IP licensing agreements had an aggregate contract amount of RMB89.8 million, among which nil, RMB7.7 million, RMB12.1 million and RMB32.8 million was recognized as revenue in 2023, 2024, 2025 and the four months ended April 30, 2026, respectively. The remaining unrecognized contract amount of approximately RMB32.8 million as of April 30, 2026 had subsequently been recognized in full prior to the Latest Practicable Date. The majority of these customers for both business lines were private-sector enterprises, with the balance comprising state-owned or other public-sector customers. The majority of such customers were new customers, and we did not receive repeat purchases of pure technical services from any single customer during the Track Record Period.

Our existing IP licensing engagements have each been one-off arrangements reflecting that each such arrangement grants the customer a one-time right to use a specified item of our intellectual property, or an earlier-generation architecture, for integration into its own products, and therefore does not by its nature give rise to recurring fees. We accordingly expect our IP licensing arrangements to be predominantly one-off in nature going forward, although we do not exclude the possibility of generating recurring royalty income from future engagements. As of May 31, 2026, we had acquired approximately 21 new customers across our business as a whole since January 1, 2025 (comprising both new customers of our GPGPU products and new technical services customers, rather than technical services customers only), covering verticals such as smart city, industrial, robotics, intelligent computing centers, transportation, electricity and oil, and distributed across Beijing, Shanghai, Guangdong, Jiangsu, Zhejiang, Hunan, Hubei and Sichuan, and we target to expand our cumulative new customer base by an additional 150 to 200 customers over the next 12 to 24 months on a nationwide basis. We accordingly regard the technical services segment as complementary to our GPGPU product sales, and expect this segment to be supported over time by a combination of selected flagship projects and a more diversified customer base as our commercialization progresses.

We have been actively expanding our technical services customer base by engaging with additional chip design services customers and pursuing selective IP licensing arrangements. In particular, in November 2025 we entered into a GPU250-ARCH architecture licensing contract with a new enterprise customer, pursuant to which we granted the enterprise customer a permanent, non-exclusive and non-transferable license to use the GPU250-ARCH architecture, with the licensed use restricted to the territory of the PRC. The GPU250-ARCH architecture is derived from the base architecture underlying our products that are currently in mass production, and was delivered together with only the corresponding basic standard development tools available at the relevant time rather than our latest software tool chain. The enterprise customer is an Independent Third Party with its registered office and principal place of business in Wuxi, Jiangsu Province, whose principal business is integrated circuit design and sales, and intends to integrate the licensed architecture into its existing SoC product line for the purpose of adding AI processing capabilities to such SoC chips. We do not expect this licensing arrangement to result in the enterprise customer becoming a competitor of our Group: the enterprise customer's SoC product line targets low-power and low-cost applications, whereas our GPGPU products target enterprise AI infrastructure scenarios that require significantly higher computing performance, and the licensed technology, considered together with the limited supporting development tools, is not, in our view, capable of being further developed into products that would compete with our existing or pipeline GPGPU products in our target markets. As our commercialization progresses, we expect the technical services segment to be supported by a more diversified set of customers alongside larger flagship projects.

The salient terms of our IP licensing arrangements are illustrated by the GPU250-ARCH architecture licensing contract, the largest IP licensing arrangement entered into during the Track Record Period, which are summarized below:

- *Scope and tenure:* a perpetual, non-exclusive and non-transferable license to use the licensed architecture for the design, manufacture and sale of the licensee's own chip products, restricted to the territory of the PRC; our subsequent upgrades and derivative versions are excluded.
- *Material rights and obligations:* we were required to deliver the agreed technical materials of the licensed architecture and to provide technical support for two years from delivery, included in the license fee; the licensee may not use the architecture beyond the licensed scope, sublicense, transfer or grant third parties access to it (other than its permitted service providers and manufacturers), or reverse engineer it.
- *Credit terms and payment milestones:* the license fee is payable in four installments: 50% upon signing, 30% following delivery and acceptance, and 10% each within one year and two years of delivery, respectively.
- *Ownership of intellectual property:* the intellectual property rights in the licensed architecture remain vested in us, and improvements independently developed by a party belong to that party.
- *Termination:* neither party may terminate the contract at will; either party may terminate for the other party's material breach that remains uncured for 30 days after notice or for force majeure.

Our revenue from technical services is expected to remain relatively stable in the near term, representing a recurring and complementary income stream that strengthens the resilience of our business model. While individual technical service engagements are typically project-based and completed upon delivery for a specific stage of a customer's requirements, demand for such services arises on a recurring basis across our broader customer base, as different customers engage us for similar technical support at various stages of product adoption, including initial deployment, customization and performance optimization. Leveraging our experience in GPU architecture and IP development, we have entered into architecture licensing arrangements that demonstrate our ability to monetize our technology, as illustrated by the contract entered into in November 2025. See "Business — Path to Commercialization — Achieving Revenue Requirement for a Commercial Company — Remaining Contracted Revenue and Basis of Forecast." More importantly, this business line deepens customer engagement by embedding us more closely into their development cycles, reinforces the adoption of our GPGPU products in downstream applications, and creates opportunities to monetize our IP portfolio in a targeted manner. Over time, we expect technical services to play a strategic role in supporting customer stickiness, diversifying our revenue mix, and strengthening the long-term commercialization of our technology platform.

Strengthening Sales and Marketing Efforts

We are pursuing a focused sales and marketing strategy designed to expand adoption of our products across priority industries and customer segments. Our approach is to first establish deep relationships with leading customers in core verticals, such as smart city, enterprise private clouds, education, and industrial inspection, and then leverage these partnerships to broaden penetration into mid- and long-tail customers within the same industries. To maximize effectiveness, we have organized our sales efforts into regional industry clusters, combining direct engagement by frontline sales personnel with coordinated support from ecosystem partners and technical teams. This integrated model enables us to discover and capture new opportunities while providing customers with flexible product configurations and optimized server solutions that deliver strong economic benefits. It supports our commercialization by shortening customer evaluation cycles, improving conversion from testing to scaled procurement and increasing repeat purchases from existing customers. Please see "Business — Sales and Marketing of Our Products."

Looking ahead, we plan to further expand our sales and marketing team to improve efficiency and deepen customer coverage. In particular, we aim to broaden our customer base by securing new customers and to drive incremental purchases from existing customers. Some of our existing customers who previously purchased a small quantity of our GPGPU products for testing may increase their purchases as they validate performance, while others with significant strategic value may require longer testing and evaluation cycles before scaling up to bulk orders. By maintaining close engagement and supporting customers throughout these processes, we expect to gradually convert trial users into long-term, large-volume purchasers, thereby accelerating commercialization across multiple application scenarios.

Financial Outlook and Operating Leverage

Our financial performance during the Track Record Period reflects the progression of our commercialization efforts and the operating leverage inherent in our business model. Our revenue increased from RMB27.9 million in 2023 to RMB82.6 million in 2024 and RMB171.3 million in 2025, and significantly increased from RMB5.8 million in the four months ended April 30, 2025 to RMB200.4 million in the four months ended April 30, 2026, primarily driven by the scaling of GPGPU product shipments and the expansion of technical services supporting product deployment and customer adoption. Gross profit also improved significantly over the same period, reflecting optimization of our product mix toward higher-value models, increased shipment volumes and the contribution of technical service revenue. Although we remained loss-making during the Track Record Period, our losses narrowed period-on-period as revenue scaled and operating efficiency improved, demonstrating the operating leverage inherent in our business model.

For the four months ended April 30, 2026, our revenue was RMB200.4 million, as compared to RMB5.8 million for the four months ended April 30, 2025, representing a substantial increase. The increase was primarily attributable to the increase in our total sales volume from 813 units in the four months ended April 30, 2025 to 10,344 units in the same period in 2026, driven by the ramp-up of shipments of our KS Series products following their launch in March 2025 and the completion of our supply chain transition to a local supplier to enhance production efficiency and supply chain security, measured against a low revenue base in the corresponding period of 2025 prior to such ramp-up. Our gross profit for the four months ended April 30, 2026 was RMB102.0 million, as compared to RMB3.4 million for the four months ended April 30, 2025, primarily due to the significant increase in our total sales volume and revenue.

As of the Latest Practicable Date, the aggregate value of confirmed orders from our customers (comprising legally binding contracts) was approximately RMB604.3 million, which our Directors consider sufficient to support our revenue for 2026, and the aggregate value of indicative orders (comprising non-binding purchase intentions and ongoing customer engagements that have not yet been confirmed by purchase orders) was approximately RMB680.5 million, providing further support for our revenue in 2026 and subsequent years. Of the foregoing, at least RMB35.4 million of the confirmed orders are expected to be recognized as revenue in 2026, with the remainder expected to be recognized in 2027 and beyond.

We expect our operating leverage and financial performance to continue to improve in 2026 as we scale our business under our existing business model. In particular, we expect our revenue to increase significantly, driven primarily by the continued ramp-up of our KS Series shipments and growing enterprise AI adoption, and we expect our gross profit to grow at a faster rate than our revenue and our gross profit margin to improve, reflecting our optimized product mix, increasing economies of scale as shipment volumes rise and the benefits of our completed supply chain transition. We also expect our selling and distribution expenses and administrative expenses to grow more slowly than our revenue, and therefore to decline as a percentage of our revenue, while we continue to invest in research and development to support our product roadmap and long-term competitiveness, funded principally by the net **[REDACTED]** from the **[REDACTED]**.

Our commercialization plan is designed to translate product deployment, ecosystem collaboration and market expansion into sustained improvements in operating metrics over time, including revenue scale, gross profit and operating leverage. The nature of our commercial arrangements typically progresses from initial deployment and system integration to scaled procurement as customer usage deepens, which supports increased demand for our GPGPU products and related services and contributes to recurring demand over time. As commercialization advances across comparable enterprise and industry application scenarios, a greater proportion of our fixed cost base is expected to be absorbed, supporting margin improvement and operating leverage.

Looking ahead, we intend to continue executing our commercialization initiatives in a disciplined manner, focusing on scaling adoption in core enterprise AI application scenarios while selectively expanding into public cloud and robotics markets. While we will continue to invest in research and development to sustain product competitiveness, we expect that continued execution of our commercialization strategy and disciplined market positioning will strengthen our competitive position in our target segments and support progress toward sustainable growth and profitability.

Achieving Revenue Requirement for a Commercial Company

We expect to achieve the revenue requirement applicable to a Commercial Company under Chapter 18C in 2026. This expectation is based on (i) execution and completion of existing signed contracts with scheduled revenue recognition milestones, (ii) conversion of advanced-stage projects currently under negotiation, and (iii) continued scaling of sales of our GPGPU products and technology licensing arrangements.

Remaining Contracted Revenue and Basis of Forecast

We have entered into several contracts with customers that provide a concrete contractual basis for our forecast revenue. In January 2026, we entered into a framework procurement agreement with an enterprise customer engaged in information technology and communication services, with an aggregate contract amount of approximately RMB300.0 million. Under this agreement, we will supply GPGPU accelerator boards based on our KS Series, which form part of our Specialist Technology Products, with specific purchase orders to be placed and executed on a batch basis. The contract sets out the general commercial terms, including product specifications, delivery arrangements, acceptance procedures and payment terms. Consideration is payable by bank transfer and is typically structured based on delivery and customer acceptance. We have already received an initial payment of RMB5.0 million as well as an additional payment of RMB86.0 million upon delivery of products in the first quarter of 2026. The contract does not contain unusual termination rights, contingent pricing mechanisms or material conditions precedent that would impede performance, other than customary delivery and acceptance requirements.

In March 2026, we entered into a procurement contract with an enterprise customer engaged in intelligent computing and system integration, with an aggregate contract amount of approximately RMB58.3 million, pursuant to which we will supply GPGPU accelerator boards based on our KS Series, which form part of our Specialist Technology Products, together with the related server, networking and security equipment required to deploy them as an AI computing cluster. The contract sets out agreed product configurations, quantities and pricing, with consideration payable by bank transfer based on delivery and customer acceptance. Revenue is expected to be recognized upon delivery and acceptance in accordance with applicable accounting standards. The contract includes standard provisions on inspection and acceptance, delivery timelines and payment arrangements, and does not include unusual termination rights or material impediments to completion. Delivery of products under this contract was completed in April 2026.

In addition, in November 2025, we entered into an architecture licensing contract with an enterprise customer engaged in integrated circuit system solutions, pursuant to which we granted a permanent, non-exclusive and non-transferable license for the use of our GPU architecture in China, with a total contract value of RMB69.5 million. The consideration is payable in installments tied to contract execution and delivery milestones, including an upfront prepayment and subsequent payments following delivery and acceptance. The customer has already paid the initial installment, and customer

acceptance was subsequently completed in January and May 2026, upon which the remaining contracted revenue has been recognized. The contract specifies the licensed scope, delivery standards, acceptance procedures and payment timetable, and does not contain unusual termination rights or material conditions that would impede completion.

Taken together, these contracts provide a diversified and contractually secured basis for our forecast revenue, with the remaining contracted revenue expected to be progressively recognized in 2026 and contributing a meaningful portion of our projected revenue for that year, thereby supporting our expected trajectory to meet the Commercial Company revenue requirement.

Customers Confirming Interest through Binding Orders or Letters of Intent

In addition to signed contracts, we have obtained procurement intention agreements from a number of enterprise customers confirming their interest in our GPGPU products. As of the Latest Practicable Date, we had entered into ten procurement intention agreements with enterprise customers, with an aggregate intended contract value of approximately RMB680.5 million.

These customers are primarily enterprise and industrial users with identified AI computing deployment plans across our Group's six key application areas, namely smart city, enterprise private clouds, education, industrial inspection, industrial large models, and system integration. The procurement intention agreements typically specify the intended product models, indicative procurement quantities, pricing framework and expected deployment scale, and reflect customers' internal budget planning and medium-term procurement roadmaps.

While such procurement intention agreements are generally subject to the execution of definitive purchase contracts and satisfaction of customary internal approval procedures, we believe that they provide meaningful quantitative visibility on near-term commercial demand and form an important basis for our production planning and revenue forecasting. Based on the aggregate intended contract value of the signed procurement intention agreements, together with our historical conversion experience and expected delivery schedules, we expect a substantial portion of such intended procurement to be converted into firm purchase orders within 2026.

Contracts under Negotiation and Contract Pipeline

In addition to the procurement intention agreements described above, we are currently in active negotiations with several potential enterprise customers regarding the supply of GPGPU products and related solution deployment. These potential projects primarily relate to applications in intelligent computing centers and AI infrastructure. Among these, we are in discussions with a technology company regarding a potential cooperation project with an indicative contract value of approximately RMB600 million, which remains subject to further technical validation, commercial negotiations and internal approval processes of the relevant parties. In addition, we are in discussions with other potential customers with an aggregate indicative contract value of approximately RMB200 million.

These discussions are at early to mid-stage, involving evaluation of product configurations, deployment scale, delivery timelines and commercial terms. Subject to satisfactory completion of the relevant evaluation and approval procedures, we expect that, if successfully concluded, such negotiations could result in the execution of definitive contracts within approximately six to twelve months. While there is no assurance that these negotiations will result in binding contracts, such ongoing discussions reflect continued market interest in our GPGPU products and solutions and complement our existing pipeline of signed contracts and procurement intention agreements.

Conclusion on Revenue Requirement

Taken together with existing signed contracts, customers' purchase intentions and our active contract pipeline, our Directors are of the view that (i) our business is sustainable; and (ii) our commercialization plan provides a reasonable and sustainable basis for achieving the revenue requirement under Rules 18C.03(4) and 18C.19 within the expected timeframe. The Joint Sponsors concur with the views of our Directors.

OUR RESEARCH AND DEVELOPMENT

We have established a structured, system-level R&D framework that encompasses every phase of our chip development and commercialization cycle, from early-stage product definition and architecture design to backend verification, mass production, and ongoing optimization. Our development model is built around a tightly integrated hardware-software co-design process, enabling us to iterate and respond efficiently to evolving market needs and technological advancements.

We adopt a modular and reusable development strategy, allowing IP components, such as computing cores, interconnects, memory subsystems, and codecs, to be incrementally refined and reused across multiple chip generations. This approach significantly shortens development cycles of each generation of products and ensures functional consistency across our GPGPU product family. To enforce discipline and quality in execution, we implement a formal development process based on the Integrated Product Development methodology, supported by detailed engineering checklists and performance evaluation benchmarks.

R&D Process

The typical lifecycle of our chip development follows a structured, multi-step process as illustrated below:

Step	Description
Product Definition	Define overall product strategy, core application scenarios, performance and power targets, chip configuration, process node, and cost constraints.
Architecture Design	Design core system architecture, including computing core clusters, cache structures, memory hierarchies, peripheral interfaces, and IP integration plan.
Front-End Design	Perform RTL coding, module-level simulation, system integration, and functional verification; adopt reusable IP structure for consistency.
Functional Verification	Utilize simulation platforms and formal tools for full-chip logic verification, focusing on functional correctness, interface integrity, and stability.
Back-End Design	Conduct floorplanning, placement and routing, power and timing optimization; complete design rule checks and prepare for tape-out.
Tape-Out and Fabrication	Submit final GDSII to the foundry for mask generation and wafer fabrication; conduct sample production and initiate yield optimization.
Testing and Validation	Carry out bring-up tests, electrical and signal integrity validation, functional correctness checks, and reliability evaluation under working conditions.
Mass Production and Iteration	Implement design for manufacturability (DFM) and mass production transfer; continuously improve performance and yield through silicon feedback.

This disciplined and repeatable process supports our ability to deliver reliable products with rapid iteration cycles.

R&D Team

Our R&D efforts are led by a seasoned core team with deep domain expertise across GPU architecture, high-performance computing, and embedded systems. The team brings together veterans from global semiconductor leaders and top-tier domestic institutions. As of the Latest Practicable Date, we operated R&D centers in Suzhou (headquarters), Shanghai, Beijing, Hangzhou, and Chengdu. We established R&D centers in these cities principally to access the specialized talent pools and industry resources available in each location and to stay close to our customers and partners. Suzhou, our headquarters, anchors our operations, supply chain and mass-production functions and benefits from its position in the Yangtze River Delta semiconductor cluster; Shanghai provides deep chip-architecture and software talent and proximity to the design ecosystem; Beijing supports frontier technology research and engagement with government, industry and research institutions; Hangzhou provides software and AI talent and close access to internet and cloud customers; and Chengdu, our largest software and chip back-end base, offers a sizable and stable engineering talent pool together with lower operating costs and supportive local policies. Nearly 70% of our total workforce are R&D personnel as of April 30, 2026, most of whom have industry experience of over ten years. A majority of our R&D professionals, particularly those involved in hardware architecture, hold postgraduate degrees, primarily in electronic engineering, computer science, and related disciplines.

Member of our core R&D Team	Profile
Mr. Wang Chenhui	Mr. Wang Chenhui is a co-founder and the chief executive officer of our Company. He graduated from Shanghai University in July 2000 with a bachelor's degree in engineering, majoring in computer science and technology. Mr. Wang began his career at 上海思源笙軟件科技有限公司, an EDA software company, where he worked on the front-end of compiler development for hardware description languages such as LEF/DEF and VHDL. He later joined a domestic operating system development company, where he led the OS kernel development team. From 2007 to 2017, he served as the software team leader at Vivante Corporation (Shanghai) Co., Ltd., where he was responsible for GPU driver development, system integration, and technical support. Since joining our Company in November 2017, Mr. Wang has led the development of our software platform, various compilers, runtime systems, and software toolchains designed for heterogeneous computing. He currently serves as Chairman and General Manager. Mr. Wang was recognized as a core member of both the Suzhou Industrial Park Major Talent Program (2024) and the Suzhou Major Innovation Team (2025).
Mr. Wang Ping	Mr. Wang Ping is a co-founder and the chief innovation officer of our Company. He holds a bachelor's degree in engineering, majoring in microelectronics, from Tsinghua University (清華大學). Mr. Wang has over 20 years of experience in chip architecture design and engineering management, with particular expertise in high-performance processors and AI accelerators. Prior to co-founding our Company in 2017, he held senior technical and managerial roles at several leading semiconductor companies, where he led the development of heterogeneous SoCs and AI computing units. At our Company, Mr. Wang is responsible for implementing product and technology-related innovation and overall chip architecture planning, including the design and evolution of our proprietary GPU+ (Minsky™) on-chip heterogeneous architecture. He has played a key role in defining computing unit structure, precision configurations, interconnect architecture, and memory systems across the Goldwasser product lines.

Member of our core R&D Team	Profile
Mr. Wang Gang	Mr. Wang Gang serves as Director of Chip Development at our Company. He holds a bachelor's degree in Computer Science and Technology and a master's degree in Computer Architecture, both from Northwestern Polytechnical University. Mr. Wang began his career at Vivante Corporation (Shanghai) Co., Ltd., where he held roles from Chip Design Engineer to Senior Chip Design Engineer. He later joined VeriSilicon Microelectronics (Shanghai) Co., Ltd. as staff Chip Design Engineer and Chip Design Manager. Since joining our Company in December 2017, Mr. Wang has held successive roles including design manager, IP Design Manager, and currently oversees our full-chip development process as Director of Chip Development.
Mr. Li Xiaoquan	Mr. Li Xiaoquan is Software R&D Manager at our Company. He holds a master's degree in Computer Architecture from Huazhong University of Science and Technology. Mr. Li has nearly two decades of experience in GPU software engineering, having served as Engineer at OmniVision Technology (Shanghai) Co., Ltd., Senior Software Engineer at Vivante Corporation (Shanghai) Co., Ltd., and Principal Engineer at VeriSilicon Microelectronics (Shanghai) Co., Ltd. He joined our Company in December 2017 and currently leads software development for key driver and middleware components.
Mr. Ma Liang	Mr. Ma Liang is Chief Design Engineer at our Company. He holds a master's degree in Microelectronics and Solid-State Electronics from Hefei University of Technology and has over 15 years of experience in chip design. He previously served as Engineer at Hangzhou Guoxin Technology Co., Ltd., Sr. Staff Engineer at Vivante Corporation (Shanghai) Co., Ltd., and Senior Engineer at HiSilicon (Shanghai) Co., Ltd. Since joining our Company in November 2017, Mr. Ma has taken charge of key module design and currently leads our overall front-end and back-end integration as Chief Design Engineer.

Retention of Core R&D Personnel

We place significant emphasis on retaining our core research and development personnel, who are critical to our technology development and long-term competitiveness. To incentivize long-term commitment and align the interests of our core R&D members with ours, we have adopted equity-based incentive arrangements under which eligible core R&D personnel are granted share options subject to multi-year vesting schedules. The share options granted to our core R&D personnel generally vest over a four-year period, which we believe is effective in promoting retention and continuity of our R&D capabilities. In addition, we have entered into service agreements and related arrangements with our key management and technical staff that include confidentiality obligations, intellectual property ownership provisions and post-termination restrictive covenants. These arrangements are designed to protect our proprietary technologies and mitigate the potential impact of any departure of key personnel.

The salient terms of our standard labor contract, confidentiality agreement and non-competition agreement entered into with our key management and technical staff include:

- *Term of Employment*: Fixed-term employment contracts, typically with a term of four years, including a probation period at the commencement of employment.
- *Duties and Work Arrangements*: Employees agree to perform their designated roles and duties as assigned by us, and we may adjust their responsibilities and work location in accordance with our business needs and applicable PRC laws.

- *Remuneration and Benefits*: Remuneration is primarily in the form of monthly salary, together with statutory social insurance and other employee benefits provided in compliance with PRC laws and regulations.
- *Confidentiality Obligations*: Employees are subject to various confidentiality obligations covering technical, commercial, customer, supplier and other proprietary information, which continue to apply after termination of employment.
- *Intellectual Property Ownership*: Intellectual property created by employees in the course of their employment, including inventions, software, source code, designs and other work products, is owned exclusively by us, with employees required to assign all related rights to us.
- *Non-competition and Non-solicitation*: Certain key management and technical staff are subject to post-termination non-competition obligations, typically for a period of up to one year, during which we pay statutory non-competition compensation, as well as non-solicitation restrictions in respect of our employees, customers and business partners. Such non-solicitation terms are expressly included in the standard labor contract, confidentiality agreement and non-competition agreement entered into between us and our key management and technical staff.
- *Termination*: Termination rights and procedures are set out in accordance with PRC labor laws, including notice requirements, grounds for termination and statutory compensation where applicable.

OUR SUPPLY CHAIN

Our Fabless Production Model

We adopt a fabless production model, focusing on the design and development of GPGPUs while outsourcing the manufacturing, packaging, and part of the testing processes to third-party wafer foundries and OSAT providers. This model, widely used across the integrated circuit (IC) design industry, allows us to maximize the efficiency of our R&D resources and maintain a capital-light structure without the need to invest in large-scale manufacturing infrastructure. Since our inception, there has been no material change to our production model, which is consistent with industry norms and best practices.

Our production process consists of three major stages:

- *Wafer Fabrication*: Once our chip design is finalized, we tape out the design and transfer it to a designated wafer foundry for fabrication. This step converts our GPU+ (Minsky™) architecture into physical silicon wafers.
- *Packaging and Testing*: Upon completion of wafer fabrication, we conduct a chip probing test to verify initial die-level performance. The wafers are then sent to packaging partners, where they are assembled into complete chips. Post-packaging, these chips are returned for final functional and system-level testing before being qualified for production release.
- *Board-Level Integration*: The final step involves producing complete accelerator cards at electronic manufacturing services (EMS) factories. This stage involves multiple materials, including chemical compounds, mechanical components, and various types of electronic materials. To improve cost and management efficiency, we implement a composite procurement model: we directly procure and manage high-value raw materials (e.g., DDR memory), while delegating the purchasing of lower-value components to the EMS factories.

Supply Chain Management

To support the transition from chip design to product delivery, we have established an integrated supply chain managed by four specialized divisions: (i) the Planning and Fulfillment Division, responsible for supplier engagement, procurement and overall supply planning; (ii) the Silicon Operation Division, overseeing wafer production, packaging and testing; (iii) the EMS Operation Division, supervising board-level manufacturing at EMS factories; and (iv) the Quality Assurance Division, ensuring quality control across all stages of the supply chain from upstream silicon to final assembly. We generally secure our access to production capacity from key wafer foundries through contractual capacity reservation arrangements by placing orders in advance. In particular, we have secured our access to future wafer production capacity till the end of 2027 through a wafer capacity reservation agreement. We generally make prepayments and pay deposits to reserve wafer production capacity under agreed capacity plans. Our cash and cash equivalents of RMB449.3 million as of April 30, 2026 and committed unutilized banking facilities of RMB1,870.4 million as of August 31, 2026 provide us with flexibility to settle such payments. We also closely monitor our cash flow and liquidity position to manage the related working capital risks.

We emphasize proactive quality management throughout the production process by closely monitoring OEM partners and integrating real-time operational and testing data into our internal systems for centralized oversight. We typically enter into quality assurance agreements with key manufacturing partners, requiring them to meet specified yield targets and assume responsibility for process-related losses. In addition, we have implemented an end-to-end traceability system, under which each chip and board is assigned a unique serial number linked to detailed production and testing records, enabling effective quality control, customer support and timely response to any defects or field issues.

Selection of Suppliers and Cost Control Measures

During the Track Record Period, we primarily procured wafer fabrication services from third-party foundries, packaging and testing services from OSAT providers, and EDA tools for chip design and verification, while maintaining two to three suppliers for each major production step to reduce concentration risk and enhance flexibility. We select and manage suppliers based on the "QCDS + T" framework (quality, cost, delivery, service and technical development), under which a cross-functional team conducts structured evaluations, including potential on-site audits, before admitting qualified suppliers into our pool. We control procurement costs through competitive bidding, requiring detailed cost breakdowns and benchmarking submissions against internal and industry standards, and for components with high price volatility, we monitor market trends and obtain real-time quotations to secure competitive pricing. Where appropriate, we may increase safety stock in advance to mitigate pricing fluctuations and ensure supply continuity, supporting efficient scaling while maintaining product quality and cost discipline.

Product Storage and Inventory Management

Our finished products and in-process materials are primarily stored at the packaging facilities and EMS factories we engage, which operate established warehouse management systems enabling real-time tracking of inventory movements. We maintain a systemized inventory management approach to ensure alignment between physical stock and internal records, with all items verified for accuracy and proper storage conditions. To further strengthen control, our supply chain and finance teams conduct joint monthly audits of raw materials, work-in-progress and finished goods to reconcile records and promptly identify any discrepancies, supporting effective inventory management, risk mitigation and efficient order fulfillment.

Logistics

We manage logistics selectively based on the value and criticality of materials, taking responsibility for transporting wafers from foundries due to their high value and strategic importance, while subsequent logistics to packaging and testing facilities are typically handled by our manufacturing partners. For board-level materials, transportation is generally arranged through couriers,

except for high-value components such as DDR memory modules, for which we oversee delivery to ensure security and traceability. In terms of insurance, we previously arranged coverage on a per-shipment basis but transitioned in 2025 to an annual transportation insurance policy based on projected shipment volumes, improving risk management and cost efficiency across our supply chain operations. See "— Insurance."

OUR SUPPLIERS

Our suppliers primarily consist of outsourced foundries for wafer fabrication, OSAT providers for chip packaging and testing, EMS providers for PCBA and board assembly, software and design tool providers, and suppliers of key hardware materials such as memory and PCB. In 2023, 2024, 2025 and the four months ended April 30, 2026, our purchases from our five largest suppliers accounted for 83.1%, 65.2%, 56.9% and 76.3% of our total purchases, respectively, while our purchase from the largest supplier accounted for 59.5%, 31.1%, 31.2% and 21.0% of our total purchases, respectively.

We usually settle with most raw material suppliers and OEMs on a monthly basis with a 30 days' credit period. However, wafer suppliers often require full advance payment in line with industry practice, so we usually make full payment in advance during the early stage of our business relationship with such suppliers.

During the Track Record Period, the research and development support services provided by our suppliers principally comprised (i) the licensing of EDA software and third-party design IP, which we use as standard inputs in our chip design process, and (ii) wafer manufacturing (including tape-out as a design validation step) and chip packaging and testing services for the downstream production of our designs provided by our outsourced foundries and OSAT providers. We did not outsource any of our core research and development processes, namely architecture design, front-end design and design verification, to any third party at any time during the Track Record Period. During the early stage of our operations, when our in-house engineering headcount was more limited, we engaged a third party to assist with a portion of back-end physical implementation work, which is a common collaboration arrangement among fabless chip design companies and did not involve any transfer of our core intellectual property. The performance of full-mask tape-out by third-party foundries is the industry-standard arrangement between fabless chip design companies and wafer fabrication facilities and likewise does not involve any transfer or change of ownership of intellectual property.

The following tables set out the details of our five largest suppliers based on purchases from them in each period during the Track Record Period:

For the year ended December 31, 2023

Ranking	Supplier	Background of supplier	Nature of Purchase	Year of commencement of business relationship with us	Purchase amounts	Percentage of total purchase	Credit Terms
					(RMB in millions)		
1	**Supplier A**. .	A Shanghai-based company primarily engaged in the manufacturing of computers, communications and other electronic devices, established in 2015 with the registered capital of RMB229.3 million and listed on the Shanghai Stock Exchange with the net assets of RMB5.5 billion as of December 31, 2025.	R&D support services (such as engineering trial production and software and hardware personnel support)	2019	96.2	59.5%	Prepaid
2	**Supplier D** .	A Suzhou-based company primarily engaged in the manufacturing of cultural education, arts and crafts, sports and entertainment products, established in 2001 with the registered capital of RMB29.7 million.	Packaging, testing	2018	13.3	8.2%	Net 30 days

Ranking	Supplier	Background of supplier	Nature of Purchase	Year of commencement of business relationship with us	Purchase amounts	Percentage of total purchase	Credit Terms
					(RMB in millions)		
3	**Supplier B**. .	A California-based company established in 2018.	R&D support services (such as engineering trial production and software and hardware personnel support)	2018	12.0	7.4%	Net 30 days
4	**Suzhou Etron Technologies Co., Ltd. (蘇州易德龍科技股份有限公司)** .	A Suzhou-based company primarily engaged in the manufacturing of instruments and meters, established in 2001 with the registered capital of RMB160.4 million.	Manufacturing	2021	7.5	4.6%	Prepaid or net 30 days
5	**Supplier C** .	A Hong Kong-based company primarily engaged in the business of other electronic components and equipment, established in 2009.	Material	2020	5.4	3.4%	Prepaid
Total					134.4	83.1%	

For the year ended December 31, 2024

Ranking	Supplier	Background of supplier	Nature of Purchase	Year of commencement of business relationship with us	Purchase amounts	Percentage of total purchase	Credit Terms
					(RMB in millions)		
1	**Supplier F** . .	A Zhuhai-based company primarily engaged in software and information technology services, established in 2020 with the registered capital of RMB48.8 million.	Customized chip-design services (such as IP integration, chip design, verification, design-for-testability enhancement and electrical performance analysis)	2022	30.6	31.1%	Prepaid
2	**Supplier G** . .	A Shanghai-based company primarily engaged in technology promotion and application services, established in 2018 with the registered capital of RMB1.0 million.	R&D support services (such as engineering trial production and software and hardware personnel support)	2023	10.3	10.5%	Net 30 days
3	**Suzhou Etron Technologies Co., Ltd. (蘇州易德龍科技股份有限公司)** . .	A Suzhou-based company primarily engaged in the manufacturing of instruments and meters, established in 2001 with the registered capital of RMB160.4 million.	Manufacturing	2021	8.9	9.0%	Prepaid or net 30 days
4	**Supplier H** . .	A Chengdu-based company primarily engaged in technology promotion and application services, established in 2012 with the registered capital of RMB80.0 million.	Software	2024	7.5	7.6%	Installment

Ranking	Supplier	Background of supplier	Nature of Purchase	Year of commencement of business relationship with us	Purchase amounts	Percentage of total purchase	Credit Terms
					(RMB in millions)		
5	**Supplier D** . .	A Suzhou-based company primarily engaged in the manufacturing of cultural education, arts and crafts, sports and entertainment products, established in 2001 with the registered capital of RMB29.7 million.	Packaging, testing	2018	6.8	6.9%	Net 30 days
Total					64.1	65.2%	

For the year ended December 31, 2025

Ranking	Supplier	Background of supplier	Nature of Purchase	Year of commencement of business relationship with us	Purchase amounts	Percentage of total purchase	Credit Terms
					(RMB in millions)		
1	**Supplier I** . .	A Shanghai-based company mainly engaged in the semiconductor industry, established in 2020 with the registered capital of RMB10.0 million.	Tape-out services and procurement of wafers	2024	78.0	31.2%	Prepaid
2	**Supplier F** . .	A Zhuhai-based company primarily engaged in software and information technology services, established in 2020 with the registered capital of RMB48.8 million.	Tape-out services	2022	19.3	7.7%	Prepaid
3	**Supplier P** . .	A Shanghai-based technology company primarily engaged in advanced semiconductor IP R&D and services, integrated circuit design services, with the founding capital of RMB152.9 million.	IP	2021	17.7	7.1%	Installment
4	**Supplier Q** . .	A Hong Kong-based enterprise primarily engaged in semiconductor component-related business and technical services, with the founding capital of HKD677.0 million.	Material	2025	14.9	5.9%	Prepaid
5	**Supplier A** . .	A Shanghai-based company primarily engaged in the manufacturing of computers, communications and other electronic devices, established in 2015 with the registered capital of RMB229.3 million and listed on the Shanghai Stock Exchange with the net assets of RMB5.5 billion as of December 31, 2025.	Tape-out services and procurement of wafers	2019	12.3	4.9%	Prepaid
Total					142.2	56.9%	

For the four months ended April 30, 2026

Ranking	Supplier	Background of supplier	Nature of Purchase	Year of commencement of business relationship with us	Purchase amounts	Percentage of total purchase	Credit Terms
					(RMB in millions)		
1	Supplier R . .	A Suzhou-based technology company primarily engaged in computer system integration, weak current engineering and IT equipment sales, with the registered capital of RMB10.1 million.	Material	2025	37.3	21.0%	Net 30 days
2	Supplier I. . .	A Shanghai-based company mainly engaged in the semiconductor industry, established in 2020 with the registered capital of RMB10.0 million.	Tape-out services and procurement of wafers	2024	32.5	19.4%	Prepaid
3	Shenzhen Zhenqi Electronics Co., Ltd. (深圳振啟電子有限公司). .	A Shenzhen-based technology company primarily engaged in R&D and sales of electronic products and computer hardware & software, with the registered capital of HKD35.0 million.	Material	2023	26.9	17.0%	Net 30 days
4	Supplier S . .	A Shenzhen-based technology company primarily engaged in R&D, packaging, testing and sales of semiconductor memory chips, with the registered capital of RMB470.8 million and listed on the SSE STAR Market with the net assets of RMB5.5 billion as of December 31, 2025.	Material	2025	20.5	12.9%	Prepaid
5	Supplier Q . .	A Hong Kong-based enterprise primarily engaged in semiconductor component-related business and technical services, with the founding capital of HKD677.0 million.	Material	2025	10.8	6.0%	Prepaid
Total					128.0	76.3%	

We acquainted our five largest suppliers in each period during the Track Record Period through direct marketing efforts or friends' referrals, which mainly include referrals from our business partners and friends from peer companies.

Supplier B was a company incorporated in California, the United States in 2018, which provided us with R&D support services (principally engineering trial-production and software and hardware personnel support) on normal commercial terms. Supplier B was controlled by Mr. Li Jianwen Jeff, one of our Controlling Shareholders, and we engaged it during the early stage of our development to access overseas-based engineering and technical resources. Our purchases from Supplier B amounted to approximately RMB12.0 million in 2023 and nil in each of 2024 and 2025. Supplier B was deregistered in 2024, and we have not had, and do not expect to have, any further transactions with it.

To the best of our knowledge, during the Track Record Period and up to the Latest Practicable Date, except for Supplier B, all of our five largest suppliers in each period during the Track Record Period were Independent Third Parties. To the best of our Directors' knowledge, save for Supplier B which was controlled by one of the Controlling Shareholders, none of our Directors or their respective close associates or any person who, to the knowledge of our Directors, owned more than 5% of our issued share capital, had any interest in any of our five largest suppliers, except for Supplier B, in each period during the Track Record Period and as of the Latest Practicable Date. There was no past or

present relationship (including business, employment, financing, family, trust or otherwise) between the five largest suppliers, except for Supplier B, during the Track Record Period and us, our Directors, shareholders or senior management, or any of their respective associates.

The market price of dynamic random-access memory (DRAM), one important component of our GPGPU products, fluctuated in 2024 and increased significantly in 2025. According to Frost & Sullivan, the average procurement price of DRAM increased by approximately 320% over the period from January 1, 2024 to December 31, 2025, primarily driven by strong demand from AI infrastructure deployment and supply-side constraints. This upward trend further accelerated in recent periods, with DRAM prices increasing by approximately 150% from September 30, 2025 to the end of 2025 and by a further approximately 70% from the end of 2025 to March 31, 2026, reflecting continued supply tightness. We primarily utilize higher-specification memory components, for which supply demand conditions are comparatively more constrained and shortages more pronounced, and accordingly the price increases we experienced have been more significant than the overall DRAM market trend during the same periods. Memory components (primarily DRAM) account for a substantial proportion of our cost of sales. As a result, such sustained price increases have exerted upward pressure on our cost of sales and may adversely affect our gross profit margin if not effectively mitigated.

To address the impact of rising memory prices, we have adopted a combination of pricing, cost control and operational optimization measures. We have implemented selective price adjustments for certain products, including reducing discount levels and passing through part of the cost increases to customers in a gradual and commercially acceptable manner. At the same time, we have strengthened supply chain management and internal cost control, including centralized procurement and price-locking arrangements where feasible, improving production efficiency and strictly controlling non-essential expenditures, which have partially offset cost pressures. In addition, we have optimized our product and customer mix, focusing on higher-margin products and prioritizing key customers to maintain stable order volumes, thereby supporting overall profitability. See "Risk Factors — Risks Relating to Our Operations — Dependence on foundries and raw material price volatility may affect our operations and profitability" for further details. Except as disclosed above, during the Track Record Period and up to the Latest Practicable Date, we did not experience any significant fluctuation in prices set by our suppliers, material breach of contracts on the part of our suppliers, or delay in delivery of our orders from our suppliers. During the Track Record Period and up to the Latest Practicable Date, we did not have any material disputes with the aforementioned suppliers.

Supplier Management

We procure certain research and development support services from third-party suppliers, primarily to support engineering trial production and to supplement our internal software and hardware engineering resources during specific development stages. To ensure supplier compliance with applicable laws and regulations and to protect our proprietary technology, know-how, trade secrets and other intellectual property rights, we have implemented a structured supplier management and control framework.

This framework includes a stringent supplier onboarding and evaluation process covering legal compliance, technical capability and information security standards. We enter into written agreements with such suppliers that contain clear and robust provisions on confidentiality, intellectual property ownership and usage restrictions, under which all foreground intellectual property arising from the engagement is owned by us or otherwise exclusively licensed to us. Suppliers are contractually prohibited from unauthorized use, disclosure or transfer of our proprietary information and are subject to ongoing compliance obligations. In addition, access to sensitive technical information is limited to a need-to-know basis, and critical development activities and core intellectual property remain under our direct control.

Key Terms of Procurement Agreements with Suppliers

We enter into procurement agreements with backend suppliers, such as packaging and testing service providers, to process semiconductor wafers and deliver finished integrated circuit products in accordance with our design and quality specifications. These arrangements typically consist of a

framework or master agreement that sets out general commercial and technical terms, together with individual purchase orders or work orders that specify the relevant quantities, pricing and delivery schedules. The salient terms of a typical arrangement are summarized below:

- *Scope of Services*: The supplier performs packaging and/or testing services for semiconductor wafers and delivers finished integrated circuit products as instructed by us. The framework agreement governs the overall relationship, while individual purchase orders or work orders constitute binding contracts for each batch.
- *Materials Supply*: We provide wafers and certain materials free of charge, while the supplier is responsible for procuring other auxiliary or customized materials based on mutually agreed forecasts. Unused customer-specific materials may be subject to buyback obligations.
- *Order Placement and Acceptance*: Orders are placed through written work orders with three days' advance notice. If the supplier does not respond within 24 hours, the order is deemed accepted.
- *Yield Requirements*: Assembly yield must meet specified minimum levels (e.g., not less than 98% for single-die products, with defined deductions for multi-die products). If yields fall short due to supplier fault, we may claim compensation.
- *Testing and Inspection*: The supplier performs final product testing and accepts our "golden samples" as reference for calibration. Additional charges may apply for special inspection requirements.
- *Process Supervision and Reports*: We may dispatch personnel to inspect production, and the supplier must provide periodic process reports upon request.
- *Logistics and Delivery Terms*: We bear the cost of inbound wafer delivery, and the supplier is responsible for outbound shipment of finished goods. Shipping terms generally follow agreed trade terms.
- *Payment Terms*: Fees are determined in individual quotations or work orders. Payment is due within 30 days of invoice issuance and after acceptance of deliverables.
- *Intellectual Property and Confidentiality*: All wafers, test data, process specifications, and resulting products remain our property. Both parties are subject to strict confidentiality obligations.
- *Term and Termination*: Framework agreements are typically entered into for an initial term of one year and may be automatically renewed. Either party may terminate the agreement in accordance with its terms, including upon material breach not remedied within a specified period. These framework agreements do not contain minimum purchase commitments, exclusivity arrangements or fixed multi-year pricing obligations, and actual purchases are made on an order-by-order basis.

SALES AND MARKETING OF OUR PRODUCTS

We acquire new customers primarily through referrals, industry events and exhibitions, and direct marketing. Our sales and marketing team, comprising sales, ecosystem and technical solution sales functions, is responsible for identifying opportunities, engaging customers and providing pre-sales technical support, supported by 46 personnel as of April 30, 2026 with strong familiarity with our architecture and product roadmap. We focus on economically developed regions with high concentrations of AI-driven enterprises and leverage local sales entities for regional coverage and service delivery.

Our strategy focuses on enhancing customer acquisition and retention by optimizing existing products for core application scenarios, initially targeting leading customers and then scaling to broader markets, while also developing new products to address emerging demands such as supply chain

security and domestic substitution. We further strengthen market reach through strategic collaborations with ecosystem partners in enterprise AI verticals, including joint development, solution integration and coordinated go-to-market efforts, and have entered into framework agreements with certain key customers to facilitate recurring business and streamline order placement.

We adopt a combination of a direct sales model and a distributorship model. Direct sales to large enterprise customers form the core of our customer engagement during the Track Record Period, reflecting our strategy of building long-term relationships with key industry players. Distributorship is a common practice in the semiconductor industry, particularly during the early stages of market development when it is difficult to be included in the approved supplier lists of certain companies. Distributors may also provide value-added functions such as supply chain management and faster settlement of trade receivables, enabling us to accelerate market penetration and improve working capital efficiency.

Relationship with Customer A and Distribution Arrangement

During the Track Record Period, given our primary focus on direct sales to large customers we engaged only one distributor, Customer A. Customer A, which is an Independent Third Party, is a Shenzhen-headquartered public company primarily engaged in technical development and consulting of electronic products and development of information technology, established in 2014 with the registered capital of RMB760.0 million and listed on the Shenzhen Stock Exchange with the net assets of RMB5.4 billion as of December 31, 2025. As our business continues to expand, we may adopt a broader mix of direct sales and distributor arrangements to capture additional opportunities and enhance sales coverage. In 2023, 2024, 2025 and the four months ended April 30, 2026, our sales to the distributor Customer A amounted to RMB10.6 million, RMB11.4 million, RMB6.8 million and RMB0.4 million, accounting for 37.9%, 13.8%, 4.0% and 0.2% of our total revenue for the corresponding periods. The decrease of our revenue derived from Customer A in 2025 and the four months ended April 30, 2026 was temporary, primarily as the end customers were in the process of testing new products due to our supply chain transition to a local supplier to enhance production efficiency and supply chain security. Our supply chain transition, which was completed in 2025, involved a change of our outsourced manufacturing partner, subsequent to which the chips are produced in a different manufacturing process and become, from the perspective of the end customers, new products which shall undergo the end customers' testing and qualification before commercial deployment and before further orders are placed through Customer A. We entered into a distribution agreement with Customer A for 2026 on substantially the same terms as the previous agreements. From January 1, 2026 to August 31, 2026, Customer A placed 11 purchase orders with us with an aggregate contract value of approximately RMB1.2 million.

We adopted a distributor arrangement with Customer A primarily because the end customers behind Customer A have internal procurement and compliance requirements under which their relevant purchases are required to be centrally managed and executed through Customer A. The end customers behind Customer A are subsidiaries of a Hangzhou-headquartered public company with its shares listed on the Shenzhen Stock Exchange, which is primarily engaged in the manufacturing and sales of security equipment, network equipment and smart devices. To the best knowledge of our Directors, the end customers and their ultimate beneficial owners are Independent Third Parties. Under the distributor arrangement, we sell our products to Customer A, which in turn resells them to the end customers. To the best of our knowledge, such distribution relationship among end customers, Customer A and us was formed primarily as Customer A is a qualified distributor of the end customers and a large-scale, reputable and leading distributor in the industry to ensure a stable distribution arrangement among the parties. Our distribution arrangement with Customer A is not exclusive, and, to the best of our knowledge, the distribution arrangement between Customer A and the end customers is not exclusive. Customer A places purchase orders with us based on the underlying end-customer demand, specifying the relevant product models, quantities, pricing and delivery schedules, while we continue to engage directly with the end customers in relation to technical support, product deployment and validation. The distribution agreement operates as a framework arrangement, under which binding obligations arise only when a purchase order placed by Customer A is accepted by us, rather than through advance shipments or inventory or consignment arrangements. While the distribution agreement contains a

best-efforts sales target and minimum order quantity requirements at the purchase-order level (as further described below), it does not contain take-or-pay obligations, exclusivity arrangements or fixed multi-year pricing commitments.

The salient terms of our distribution agreements with Customer A are set out below:

- *Appointment and Scope*: We appointed Customer A as a non-exclusive distributor of our products within the agreed authorized territory. The arrangement does not restrict us from selling our products directly or appointing other distributors.
- *Term and Renewal*: The distribution agreement has an initial term of one year and is automatically renewable for successive one-year periods unless terminated or not renewed in accordance with its terms.
- *Purchase Orders*: Customer A places separate purchase orders specifying product models, quantities, pricing and delivery schedules. Each purchase order constitutes a binding contract for the relevant transaction.
- *Pricing and Payment Terms*: Pricing is determined on a purchase-order basis. Purchase orders during the Track Record Period typically provided for monthly settlement with 30-day payment terms.
- *Delivery and Inventory*: Products are delivered pursuant to confirmed purchase orders. We do not maintain inventory, consignment stock or advance shipments with Customer A.
- *Acceptance and Inspection*: Products are subject to inspection and acceptance procedures in accordance with the agreement and applicable purchase orders and are deemed accepted if no objection is raised within the agreed inspection period.
- *Warranty and After-sales Support*: Products are generally subject to a one-year warranty from receipt, with remedies including repair, replacement or refund in accordance with the agreement.
- *Intellectual Property and Confidentiality*: All intellectual property rights in our products remain vested in us, and Customer A is subject to confidentiality and related contractual obligations.
- *Termination*: Either party may terminate the agreement in accordance with its terms, including upon material breach not remedied within the agreed cure period.
- *Minimum Order Quantity*: Each monthly purchase order is required to meet a minimum order quantity per product item, with samples, trial-production units and evaluation boards exempt from this requirement. The minimum order quantity may be modified by us upon one month's prior written notice.
- *Sales Targets and Consequences for Non-Performance*: Customer A is required to use its best efforts to purchase a targeted amount of our products on a monthly or quarterly basis. If Customer A fails to achieve a specified percentage of the agreed target, we are entitled to cancel its distributorship and terminate the distribution agreement.
- *Returns*: Once a purchase order has been accepted by us in writing, it is non-cancellable and non-refundable, and the purchase price is not subject to set-off, claim, counterclaim or adjustment. Returns are accordingly limited to the warranty remedies described above.
- *Sub-distribution Arrangements*: Customer A may appoint its affiliates as sub-agents or sub-distributors within the authorized territory, subject to the terms of the distribution agreement. The term of any sub-distribution arrangement may not exceed the term of the distribution agreement, and Customer A remains jointly and severally liable for any breach of the distribution agreement by its affiliates.

Given the above commercial arrangements and the fact that Customer A was our sole distributor during the Track Record Period, we did not face cannibalization risk among multiple distributors. In addition, we believe the risk of channel stuffing in relation to Customer A was low, as Customer A's purchases from us were generally aligned with the underlying end-customer demand, we did not stock products with Customer A, and shipments were made based on confirmed purchase orders with defined delivery and payment terms. Furthermore, we have implemented internal control measures to prevent channel stuffing, including prompt communications with the end-customer subsequent to Customer A's placement of purchase orders to verify the actual demand, as well as regular communications with Customer A to check its inventory level generally conducted every two weeks. We have not discovered channel stuffing issues since the commencement of our business relationship with Customer A. For the same reasons, we did not identify material risks relating to improper distributor marketing activities during the Track Record Period.

If we engage additional distributors in the future, we plan to implement appropriate internal control measures to manage potential channel-related risks, including cannibalization, channel stuffing and compliance of distributor marketing activities. These measures are expected to include aligning distributor purchases with underlying end-customer demand, monitoring distributor order and shipment patterns, implementing internal approval and documentation requirements for distributor transactions, and conducting periodic reviews of distributor sales and marketing activities to ensure compliance with our contractual arrangements and internal policies.

Direct Sales

The salient terms of our standard direct sales agreements during the Track Record Period are set out below:

- *Product and Pricing:* The agreement specifies the model number, quantity, and unit price (inclusive of tax) for the GPGPU products being purchased. The order details are typically listed in an annexed table.
- *Delivery Arrangements:* The seller (our Company) is responsible for shipping the goods to the customer's designated address, as agreed in the contract.
- *Inspection and Acceptance:* The buyer is required to inspect the goods upon receipt and notify the seller in writing of any discrepancies. If no objections are raised within three working days, the goods are deemed accepted.
- *Payment Terms:* Common payment terms include credit terms of 30 days except for certain customers, where prepayment is normally required. For major customers, the credit term may be extended up to 90 days or staged payment. Bank account and invoicing details are included in the contract. The seller issues a formal invoice upon full payment.
- *Confidentiality:* Both parties are required to keep all confidential technical and commercial information strictly confidential. This includes product specifications, pricing, and customer use cases.
- *Liability and Breach:* Either party may terminate the agreement or seek damages in the event of a material breach not remedied after notice. Penalties for late payment or non-performance may apply.
- *Force Majeure:* The contract includes a standard force majeure clause excusing liability for performance failures due to unforeseen events such as natural disasters or government restrictions.

After-sale Services

Our PRC Legal Advisors are of the view that, under the PRC Civil Code, (a) where product defects cause personal injury or damage to property of others, the producer may bear tort liability, (b) where product issues relate only to product quality or performance and do not result in personal injury or third-party property damage, liabilities are generally addressed through contractual remedies in accordance with the relevant contractual arrangements and applicable PRC laws, including repair, replacement, return or compensation for losses. We provide after-sale services for our GPGPU products in accordance with our warranty policy and internal service procedures. We offer a two-year warranty for products delivered to certain customers. For customers with extended warranty requirements, we may, following necessary commercial discussions and agreement with such customers, provide an extended warranty period ranging from two to five years. Our after-sale services are coordinated through our business team, which serves as the primary interface with customers for matters such as product complaints, return merchandise authorizations (RMA) and returns. Upon receiving customer feedback, our quality department works with the relevant internal teams to investigate and handle the issues, and coordinates the subsequent resolution process. We provide timely updates on the handling progress to customers through our business team.

OUR CUSTOMERS

Our customers are primarily corporate and institutional users especially in the key application scenarios such as smart city, enterprise private clouds, and education. These customers include solution providers, OEMs, and integrators that apply our products in real-world environments and business scenarios such as intelligent invigilation, industrial inspection, LLM inference, and edge-to-cloud AI computing. Our customers are mainly located in Beijing, Hangzhou, Shanghai, Suzhou, Shenzhen and Chengdu. In 2023, 2024, 2025 and the four months ended April 30, 2026, we had nine, 28, 67 and 41 customers, and eight, 22, 51 and 21 new customers. Our business model for both business lines is transaction-based, and we count a counterparty in a transaction as a customer only after revenue is recognized under such transaction. Therefore, the number of our customers for the four months ended April 30, 2026 is smaller than that for the full year 2025. During the Track Record Period and up to the Latest Practicable Date, we have accumulatively serviced 125 customers in total, and our customer retention rate is approximately 100%, 67%, 57% and 30% in 2023, 2024, 2025 and the four months ended April 30, 2026. The decrease in our customer retention rate was primarily driven by the rapid expansion of our customer base. Specifically, the customer retention rate was higher in 2023 primarily due to a smaller customer base in 2022. In 2024 and 2025, although the number of our repeat customers continuously increased, the total number of our customers increased at a faster pace as we acquired a large number of new customers in each year, which lowered our customer retention rate in percentage terms. The relatively lower customer retention rate in the four months ended April 30, 2026 was mainly due to a relatively shorter observation period, as certain customers had not yet placed orders with us again in the four months ended April 30, 2026.

In 2023, 2024, 2025 and the four months ended April 30, 2026, revenue contributed from our five largest customers accounted for 98.4%, 85.1%, 83.0% and 90.4% of our total revenue, respectively, while revenue from our largest customer accounted for 37.9%, 25.7%, 29.4% and 38.0%, respectively, for the same periods. During the Track Record Period, we generally granted customers credit terms of 30 days except for certain customers, where prepayment is normally required. For major customers, we may grant credit terms of up to 90 days or staged payment, depending on customer profile and risk assessment.

The following tables set out the details of our five largest customers based on their revenue contribution in each period during the Track Record Period:

For the year ended December 31, 2023

Ranking	Customer	Background of customer	Products/service provided by us	Year of commencement of business relationship with us	Revenue	Percentage of total revenue	Credit Terms
					(RMB in millions)		
1	**Customer A** . . .	A Shenzhen-headquartered public company primarily engaged in technical development and consulting of electronic products and development of information technology, established in 2014 with the registered capital of RMB760.0 million and listed on the Shenzhen Stock Exchange with the net assets of RMB5.4 billion as of December 31, 2025.	GPGPU products used in smart city scenarios	2022	10.6	37.9%	Net 30 days
2	**Customer B** . . .	A Beijing-based private company primarily engaged in the promotion and application services of technology, established in 2003 with the registered capital of RMB20.0 million.	GPGPU products used in smart city scenarios	2023	8.2	29.5%	Net 60 days
3	**Customer C** . . .	A Beijing-based private company primarily engaged in the manufacturing of computers, communications, and other electronic devices, established in 2022 with the registered capital of RMB5.0 million.	GPGPU products used in system integration scenarios	2023	5.3	19.0%	Staged payment
4	**Customer D** . . .	A Tianjin-based company primarily engaged in technology promotion and application services, established in 2022 with the registered capital of RMB1.0 million.	GPGPU products used in smart city scenarios	2023	2.3	8.3%	Prepayment
5	**Customer E** . . .	A Chongqing-based private company primarily engaged in software and information technology services, established in 2019 with the registered capital of RMB748.5 million.	GPU accelerator card design and software adaptation services	2023	1.0	3.7%	Staged payment
Total					27.4	98.4%	

For the year ended December 31, 2024

Ranking	Customer	Background of customer	Products/service provided by us	Year of commencement of business relationship with us	Revenue	Percentage of total revenue	Credit Terms
					(RMB in millions)		
1	**Customer F** . . .	A Chengdu-based private company primarily engaged in software and information technology services, established in 2021 with the registered capital of RMB59.3 million.	GPGPU products used in enterprise private cloud scenarios	2024	21.2	25.7%	Prepayment
2	**Customer B** . . .	A Beijing-based private company primarily engaged in the promotion and application services of technology, established in 2003 with the registered capital of RMB20.0 million.	GPGPU products used in smart city scenarios and IP licensing	2023	15.2	18.4%	Net 60 days

Ranking	Customer	Background of customer	Products/service provided by us	Year of commencement of business relationship with us	Revenue	Percentage of total revenue	Credit Terms
					(RMB in millions)		
3	**Customer G**	A Beijing-based private company primarily engaged in technology development, technical services and sales of electronic products, established in 2019 with the registered capital of RMB2.2 million.	GPGPU products used in system integration scenarios	2024	13.8	16.7%	Net 90 days
4	**Customer A**	A Shenzhen-headquartered public company primarily engaged in technical development and consulting of electronic products and development of information technology, established in 2014 with the registered capital of RMB760.0 million and listed on the Shenzhen Stock Exchange with the net assets of RMB5.4 billion as of December 31, 2025.	GPGPU products used in smart city scenarios	2022	11.4	13.8%	Net 30 days
5	**Customer H**	A Beijing-based research institute primarily engaged in R&D, established in 2020 with the founding capital of RMB10.0 million.	Back-end chip design services	2024	8.7	10.5%	Prepayment
Total					70.3	85.1%	

For the year ended December 31, 2025

Ranking	Customer	Background of customer	Products/service provided by us	Year of commencement of business relationship with us	Revenue	Percentage of total revenue	Credit Terms
					(RMB in millions)		
1	**Customer K**	A Hangzhou-based private technology company primarily engaged in information transmission, software and information technology services, established in 2023 with the founding capital of RMB10.0 million.	GPGPU products used in enterprise private cloud scenarios	2025	50.3	29.4%	Net 90 days
2	**Customer H**	A Beijing-based research institute primarily engaged in R&D, established in 2020 with the founding capital of RMB10.0 million.	Back-end chip design services	2024	39.3	23.0%	Prepayment
3	**Customer L**	A Suzhou-based private technology company primarily engaged in R&D, sales of communication equipment, computer software and hardware, and related technical services, established in 2015 with the founding capital of RMB10.0 million.	GPGPU products used in enterprise private cloud scenarios	2025	30.6	17.8%	Net 90 days
4	**Customer M**	A Singapore-based private enterprise primarily engaged in research and experimental development on electronics (excluding medical science), established in 2024 with the founding capital of SGD10,000.	IP licensing	2025	12.1	7.0%	Prepayment
5	**Customer N**	A Suzhou-based private intelligent technology company primarily engaged in information transmission, software and information technology services, established in 2018 with the founding capital of RMB385.0 million.	GPGPU products used in enterprise private cloud scenarios	2025	9.9	5.8%	Net 90 days
Total					142.1	83.0%	

For the four months ended April 30, 2026

Ranking	Customer	Background of customer	Products/service provided by us	Year of commencement of business relationship with us	Revenue	Percentage of total revenue	Credit Terms
					(RMB in millions)		
1	**Customer S** . . .	A Guangzhou-based technology company primarily engaged in new generation information technology, computing infrastructure construction and intelligent computing services, with a registered capital of RMB157.6 million and listed on the Shanghai Stock Exchange with the net assets of RMB250.9 million as of December 31, 2025.	GPGPU products used in enterprise private cloud scenarios	2026	76.1	38.0%	Staged payment
2	**Customer T** . . .	A Shanghai-based technology company primarily engaged in software development, artificial intelligence application development and technology promotion services, with a registered capital of RMB10.0 million.	GPGPU products used in enterprise private cloud scenarios	2026	51.6	25.8%	Staged payment
3	**Customer Q** . . .	A Wuxi-based technology company primarily engaged in integrated circuit design, manufacturing, testing and sales, with a registered capital of RMB500.0 million.	IP licensing and GPGPU products used in system integration scenarios	2025	36.2	18.1%	Net 10 days
4	**Customer L** . . .	A Suzhou-based technology company primarily engaged in R&D, sales of communication equipment, computer software and hardware, and related technical services, with the founding capital of RMB10.0 million.	GPGPU products used in enterprise private cloud scenarios	2025	8.5	4.3%	Net 30 days
5	**Customer U** . . .	A Suzhou-based technology company primarily engaged in intelligent weak current system solutions, software development and information system integration services, with a registered capital of RMB20.0 million.	GPGPU products used in smart city scenarios	2026	8.5	4.3%	Net 30 days
Total					181.1	90.4%	

We acquainted our five largest customers in each period through direct marketing efforts or customers and friends' referrals, which mainly include referrals from our downstream customers, industry seminar participants and business partners. In addition, we plan to broaden our customer base and strengthen our relationship with customers by developing sales centers. Setting up new sales centers provides convenience for customer maintenance and our business trip to nearby cities, and allows us to have an in-depth understanding of local market conditions, visit potential customers more frequently and penetrate into the markets effectively through the sales centers.

To the best of our knowledge, during the Track Record Period and up to the Latest Practicable Date, all of our five largest customers in each period were Independent Third Parties. To the best of our Directors' knowledge, none of our Directors or their respective close associates or any person who, to the knowledge of our Directors, owned more than 5% of our issued share capital, had any interest in any of our five largest customers in each period as of the Latest Practicable Date. There was no past or present relationship (including business, employment, financing, family, trust or otherwise) between the five largest customers during the Track Record Period and us, our Directors, shareholders or senior management, or any of their respective associates.

During the Track Record Period and up to the Latest Practicable Date, we did not have any material disputes with the aforementioned customers, nor did we receive any material complaints from such customers. We did not receive any material product returns from our customers during the Track Record Period and up to the Latest Practicable Date, and to the best knowledge of our Directors and senior management, there were no potential material product returns as of the Latest Practicable Date.

OVERLAPPING CUSTOMER AND SUPPLIER

During the Track Record Period, certain counterparties acted as both our customers and suppliers in the ordinary course of business. All transactions with such counterparties were conducted on an arm's length basis, on normal commercial terms, and did not give rise to any conflict of interest.

We purchased customized chip-design services from Supplier F, amounting to nil, RMB30.6 million, RMB19.3 million and nil in 2023, 2024, 2025 and the four months ended April 30, 2026, respectively, representing approximately nil, 31.1%, 7.7% and nil of our total purchase for the corresponding periods. The services provided by Supplier F principally comprised customized chip-design solutions, including IP integration, chip design, verification, design-for-testability (DFT) enhancement and electrical performance analysis. These services are priced on a workload basis with reference to prevailing market rates for comparable customized design services, which are on normal commercial terms. Supplier F has long-standing relationships with certain overseas wafer foundries. We engaged Supplier F in connection with certain of our earlier-generation product designs that targeted overseas process nodes, in order to leverage its established interfaces with the relevant foundries. We did not purchase customized chip-design services from Supplier F in the four months ended April 30, 2026, following the completion of our supply chain transition to a local supplier to enhance production efficiency and supply chain security. We also sold sample GPGPU products to Supplier F for its internal testing purposes, generating revenue of nil, RMB17.7 thousand, RMB17.7 thousand and nil in 2023, 2024, 2025 and the four months ended April 30, 2026, respectively, which accounted for approximately nil, 0.02%, 0.01% and nil of our total revenue for the corresponding periods. The corresponding gross profit was nil, approximately RMB12.0 thousand, RMB14.0 thousand and nil, representing a gross profit margin of nil, approximately 67.8%, 79.3% and nil, and contributing approximately nil, 0.02%, 0.02% and nil of our total gross profit for the same periods. The relatively high gross profit margin is primarily attributable to the limited volume and the sample nature of these sales.

We purchased AI edge computing devices from Customer D, amounting to RMB29.0 thousand, RMB355.0 thousand, RMB10.4 thousand and nil in 2023, 2024, 2025 and the four months ended April 30, 2026, respectively, representing approximately 0.02%, 0.4%, 0.004% and nil of our total purchase for the corresponding periods. We also sold GPGPU products to Customer D, generating revenue of RMB2.3 million, RMB1.3 million, RMB4.4 million and RMB1.6 million in 2023, 2024, 2025 and the four months ended April 30, 2026, respectively, which accounted for approximately 8.3%, 1.5%, 2.5% and 0.8% of our total revenue for the corresponding periods. The corresponding gross profit was approximately RMB1.2 million, RMB0.2 million, RMB1.5 million and RMB0.8 million, representing a gross profit margin of approximately 50.5%, 14.7%, 35.2% and 48.2%, and contributing approximately 10.7%, 0.4%, 1.7% and 0.7% of our total gross profit for the same periods. The fluctuation in gross profit margin across the years reflected differences in product mix and order specifications under the relevant purchase orders.

We purchased server equipment from Customer L, amounting to nil, nil, RMB312.0 thousand and nil in 2023, 2024, 2025 and the four months ended April 30, 2026, respectively, representing approximately nil, nil, 0.1% and nil of our total purchase for the corresponding periods. We also sold GPGPU products to Customer L, generating revenue of nil, nil, RMB30.6 million and RMB8.5 million in 2023, 2024, 2025 and the four months ended April 30, 2026, respectively, which accounted for approximately nil, nil, 17.8% and 4.3% of our total revenue for the corresponding periods. The corresponding gross profit was nil, nil, approximately RMB20.7 million and RMB2.4 million, representing a gross profit margin of nil, nil, approximately 67.9% and 28.0%, and contributing approximately nil, nil, 23.0% and 2.3% of our total gross profit for the same periods. Our gross profit margin on sales to Customer L was broadly consistent with our gross profit margin for GPGPU product sales generally, reflecting standard product pricing on arm's length terms.

Except as disclosed above, none of our major customers during the Track Record Period was also a supplier, and vice versa.

SEASONALITY

Our business is subject to certain seasonal fluctuations, primarily due to the procurement patterns of downstream customers such as large enterprises and government agencies. These customers often concentrate purchases in specific quarters, leading to uneven revenue distribution throughout the year. The first quarter is typically a low season, largely due to the Spring Festival holiday and the tendency for urgent procurement projects to be completed by the end of the previous year. In contrast, the third and fourth quarters are generally peak periods, consistent with typical seasonal trends observed across the semiconductor industry. As we continue to expand our market presence and diversify our customer base and application scenarios, we believe that we are increasingly well-positioned to mitigate seasonal fluctuations in revenue and operating performance.

INTELLECTUAL PROPERTY

We regard our patents, trademarks, trade secrets and other intellectual property rights as critical to our business operations. As of the Latest Practicable Date, we hold 95 granted patents and 12 patent applications, and 43 registered copyrights, which provide us with IP protection for our Specialist Technology Products during the Track Record Period and up to the Latest Practicable Date. We have self-developed and solely own all these intellectual properties for each of our Specialist Technology Products and there were no co-ownership or co-sharing arrangements with other third parties during the Track Record Period and up to the Latest Practicable Date. We have not in-licensed any material intellectual property rights or outsourced any core R&D processes to third parties. We expect to continue to incur expenses to maintain, protect and expand our intellectual property portfolio.

The following table sets out the information of our material patents granted in China as of the Latest Practicable Date:

Field of Application	Name of Patent	Expiry Date	Ownership
Computer Technology, Convolutional Computation	Convolution Computation Method, Data Processing Method, Chip, and Electronic Device	February 10, 2043	Solely Owned
Processor Technology	Convolution Computing Unit, AI Computing Array, and Related Devices	December 14, 2042	Solely Owned
Memory Management in High-Performance Parallel Processors	Method and Corresponding Processor for Dynamically Shared Memory Space in Parallel Processors	September 2, 2042	Solely Owned
Chip Technology	Method, Device, and Computer Equipment for Task Processing	May 9, 2042	Solely Owned
Computer Technology	Scheduling Method and Scheduling System for Heterogeneous Hardware Architecture	January 25, 2039	Solely Owned
Processor Technology	Method for Managing Register File Units	January 21, 2039	Solely Owned

The following table sets out the information of major proprietary rights of IC layout design filed and granted in China as of the Latest Practicable Date, which are directly related to our major products. We consider the following proprietary rights of IC layout design as our major and typical ones.

Field of Application	Name of Proprietary Right of IC Layout Design	Expiry Year
Cloud Inference	Cloud Inference Chip	2032
General-Purpose High-Performance Computing	AI High-Performance Computing Chip DL-hum	2032
Convolutional Computation	AI Convolutional Computing Chip DL-te_sub_base	2032

Field of Application	Name of Proprietary Right of IC Layout Design	Expiry Year
Compilation	Compiler Chip (V1)	2032
General-Purpose High-Performance Computing	AI Chip	2032
General-Purpose High-Performance Computing	Goldwasser-V1 Parallel Processor	2030
General-Purpose High-Performance Computing	Kehan-M0 Parallel Processor	2030

For detailed information about our material intellectual property, see "Appendix IV — Statutory and General Information — B. Further Information about our Business — 2. Our Intellectual Property Rights."

We protect our intellectual property through a combination of patents, trademarks, trade secrets, unfair competition laws and contractual arrangements, including confidentiality agreements. We typically define ownership and protection of intellectual property in employment and commercial agreements, and implement supporting measures such as internal policies and designated personnel for intellectual property management, timely registration and monitoring of rights, prompt action against potential infringements and engagement of professional service providers.

We have entered into agreements with confidentiality clause with all of our employees, and agreements with non-competition clause with our senior management and other employees who have access to trade secrets or confidential information about our business. As of the Latest Practicable Date, we had not been subject to any material disputes or claims for infringement upon third parties' intellectual property rights and, to the best knowledge of our Directors and senior management, they are not aware of any such disputes. As advised by our Intellectual Property Legal Advisors, as of the Latest Practicable Date, we did not encounter any material impediments to our intellectual property applications in the PRC.

DATA PRIVACY AND INFORMATION SECURITY RISK MANAGEMENT

In the course of our business, we collect, store and process business data and transaction data. As we only make transactions with enterprises, we do not collect or process personal data of any external parties. During the Track Record Period and up to the Latest Practicable Date, we did not experience any material cybersecurity incidents, data breaches, data leakage or material data loss. We maintain a financial system, a human resource management system and a business management system. See "Risk Factors — Risks Relating to Our Operations — Security breaches and other disruptions could compromise our confidential and proprietary information, which could cause our business and reputation to suffer" and "Regulatory Overview — Regulations on Internet Information Security and Privacy Protection."

The principal types of data processed by us include (i) employee personal information collected for human resources management purposes, and (ii) business data generated in our research and development and related business activities.

- Employee personal information, such as name and contact details is collected and processed by our human resources department strictly on a notice or consent basis and used solely for internal management. Such personal information is stored for the minimum period required by applicable laws or business needs and is securely deleted or anonymized upon employee departure. We do not conduct any cross-border transfer of such personal information.

- Business data, primarily including production and operation data, supply chain data, and R&D and testing data (including software test reports, runtime logs, internal R&D documents, code commit records and other materials), are generated and processed in the course of our research and development and related business activities. Such data is accessed only by authorized personnel within their scope of duties through internal systems. In

accordance with our data classification and protection policies, business data are stored on dedicated servers located within the PRC with enhanced security protection measures, including encryption and role-based access controls. We do not provide such data outside the PRC.

We do not process important or core data as defined under applicable PRC data laws and regulations in our business operations, nor do we engage in any cross-border transfer of data. All data collected by us in our business operations are stored on servers located within the territory of the Chinese Mainland.

We pay close attention to risk management relating to our IT system, as storage and protection of corporate data and related information is critical to us. To ensure data security, we have adopted a rigorous encrypted algorithm to store sensitive data and strictly execute a data accessing and transmitting policy to ensure the confidentiality of our data. We have also developed strict internal control and data accessing mechanisms and detailed approval and operation procedures regarding data storage and processing, including data classification standards that distinguish data by sensitivity, role-based access controls linked to job functions, and internal approval workflows for the creation, modification and access of sensitive data. We have established a set of internal protocols on data security, which set forth detailed, strict requirements in relation to the use, disclosure and protection of confidential information. Among other things, such internal protocols provide limited authorization to our employees holding specific positions at specific levels to access and process corporate data on a need-to-know basis, who shall use such data only for the purposes of performing their work assignment.

All of our employees are required to sign a confidentiality agreement with us, which prohibits them from disclosing any confidential information relating to their work without our consent. We have a structured data backup system to encrypt and store data on servers in different locations in order to minimize the risk of data loss. We also conduct data restoration tests to examine the status of the backup system on a regular basis.

In addition, we have established a remote disaster recovery system for our server by setting up multiple storage for the same information and data of long time dimension on the cloud, local and remote locations. Even if the server is damaged due to the highest level of disasters such as earthquakes, mudslides and other irresistible natural disasters, we believe that it can safeguard and guarantee that the service and data can be completely restored promptly.

As advised by our PRC Legal Advisors, during the Track Record Period and up to the Latest Practicable Date, our Group has not been subject to any administrative penalties due to non-compliance with the data privacy and information security. Given that (i) we have established a structured data protection and risk mitigation framework in accordance with the Personal Information Protection Law, the Cybersecurity Law, the Data Security Law, and other applicable laws and regulations, covering institutional, organizational, and technical aspects; (ii) the volume of personal information processed by us covered fewer than one million individuals, and we are therefore not subject to the statutory requirements to designate a personal information protection officer or to make the related regulatory filings; (iii) the number of individuals whose personal information we process is significantly below the 10 million threshold, and we are therefore not subject to the statutory obligation to conduct a personal information protection compliance audit; and (iv) we had no material violations or non-compliance records during the Track Record Period and up to the Latest Practicable Date, our PRC Legal Advisors are of the view that we had complied with applicable PRC laws and regulations on data privacy, cybersecurity and personal information protection in all material respects during the Track Record Period and up to the Latest Practicable Date.

COMPETITION

We operate in a competitive and rapidly evolving GPGPU market, where success depends on product performance, energy efficiency, ecosystem compatibility, cost-effectiveness, service quality and brand credibility, as well as our ability to meet the technical and deployment needs of customers. At the same time, the industry presents significant barriers to entry, including high technical complexity, long development cycles, close coordination with manufacturing partners and the need to build customer

trust through reliable, long-term delivery. Enterprise AI applications further require integrated hardware and software optimization, favoring providers with full-stack capabilities and proven execution. For more information on the industry landscape, please see "Industry Overview."

EMPLOYEES

As of the Latest Practicable Date, we had 278 full-time employees, all of whom were based in Chinese mainland. The following table sets forth the number of our employees by function as of the Latest Practicable Date:

Function	Number of employees
R&D	189
Administration	25
Sales and marketing	42
Supply chain	22
Total	**278**

We recognize that attracting and retaining qualified personnel is critical to our operations and continued development. We recruit primarily through campus hiring, online channels and employee referrals, and support new hires, particularly in R&D roles, through training, mentoring and periodic technical exchanges. Our compensation framework includes base salary, performance-based bonuses and equity incentives, and we typically enter into non-compete agreements with senior management and key technical staff while conducting annual performance reviews. We have not established a labor union and have maintained stable employee relations, with no material labor disputes or strikes during the Track Record Period and up to the Latest Practicable Date.

Social Insurance and Housing Provident Fund Contributions

As required by PRC laws and regulations, we participate in social security schemes organized by municipal and provincial governments, including pension, medical insurance, work-related injury insurance, unemployment insurance, maternity insurance, and housing provident fund. We are required to make contributions to these schemes at specified percentages of the salaries, bonuses, and certain allowances of our employees, up to a maximum amount specified by the local governments from time to time. During the Track Record Period and up to the Latest Practicable Date, we have made such contributions in accordance with applicable PRC laws and regulations in all material respects. However, during the Track Record Period, certain subsidiaries engaged third-party human resources agencies to make contributions on their behalf for a limited number of employees. The amounts of social insurance contributions made through third party human resources agencies were approximately RMB0.8 million, RMB0.7 million, RMB0.5 million and RMB0.06 million as of December 31, 2023, 2024, 2025 and April 30, 2026, respectively. The amounts of housing provident fund contributions made through third party human resources agencies were approximately RMB0.5 million, RMB0.4 million, RMB0.3 million and RMB0.04 million as of December 31, 2023, 2024, 2025 and April 30, 2026, respectively. As of the Latest Practicable Date, only two employees remain under such arrangements, and we expect to rectify the remaining arrangements by the first quarter of 2027. As advised by our PRC Legal Advisors, the engagement of a third-party human resources agency to pay social insurance premium and housing provident funds for certain employees is not in strict compliance with relevant PRC laws and regulations, and according to the Social Insurance Law of the PRC (《中華人民共和國社會保險法》), if an employer fails to pay social insurance contributions on time and in full and fails to make the payment or supplement the amount within a specified time limit ordered by the government authority, the government authority may impose a fine ranging from one to three times the amount of the unpaid premiums. With respect to housing provident funds, the relevant government authorities may order us to make up the outstanding contributions within a prescribed period, and where such payment is not made within the prescribed period, the relevant authorities may apply to the competent court for compulsory enforcement, and a fine ranging from RMB10,000 to RMB50,000 may be imposed. The third-party

human resources agency and our Directors have confirmed that, during the Track Record Period and up to the Latest Practicable Date, there had been no outstanding, underpaid or deficient contributions in respect of any payments made by us through third-party human resources agencies.

According to the Special Credit Report of Business Entity (Proof of No Illegal or Non-compliant Record) (《經營主體專項信用報告(有無違法違規紀錄證明版)》) issued by the Credit China (Chengdu, Sichuan) platform (信用中國(四川成都)平台) on February 25, 2026 and June 1, 2026 and the Special Credit Report (for Substitution of Certificate of No Violation Record) (《專用信用報告(替代有無違法紀錄證明專用版)》) issued by the Shanghai Public Credit Information Service Center (上海市公共信用信息服務中心) on June 1, 2026, which consolidate regulatory compliance information maintained by the relevant competent governmental authorities, Chengdu Denglin and Shanghai Suanzhen had no record of violations in the areas of human resources and social security, medical insurance or housing provident fund administration during the Track Record Period. Our PRC Legal Advisors are of the view that the Credit China (Chengdu, Sichuan) platform and Shanghai Public Credit Information Service Center are the competent authorities to issue the credit report for Chengdu Denglin and Shanghai Suanzhen, respectively. As advised by our PRC Legal Advisors, if (i) there is no material change in the currently effective PRC laws and regulations, (ii) no employee complaint or dispute arises, and (iii) our Group is able to rectify the non-compliance as instructed by the relevant authorities within the specified period, the likelihood of our Group being fined by government authorities and thereby suffering a material adverse effect on business operations will be relatively low.

Pursuant to Article 19(1) of the Supreme People's Court's Interpretation (II) on Several Issues Concerning the Application of Law in Labor Dispute Cases (《最高人民法院關於審理勞動爭議案件適用法律問題的解釋(二)》) (the "**New Judicial Interpretation**"), promulgated on July 31, 2025, and effective as of September 1, 2025, if an employer and an employee agree or the employee commits that social insurance contributions are not required to be paid, the People's Court shall deem such agreement or commitment invalid, and where an employer fails to pay social insurance contributions, and the employee requests to terminate the labor contract and claims economic compensation from the employer in accordance with the PRC Labor Contract Law, the People's Court shall support such claim. In light of our PRC Legal Advisors' opinion that the New Judicial Interpretation does not expand our penalty exposure or repeal the social insurance laws and regulations currently in force of the PRC, we believe the New Judicial Interpretation would not have a material adverse effect on our business or financial results.

LAND AND PROPERTIES

We are headquartered in Suzhou, and maintain certain operation functions in Shanghai, Beijing, Hangzhou, Shenzhen, Xi'an and Chengdu. As of the Latest Practicable Date, we did not own any property, and leased 12 properties in the PRC with an aggregate GFA of approximately 7,150 sq.m. from third parties. These properties were used primarily as premises of office spaces, R&D activities and daily operations. Our lease agreements in respect of the abovementioned leased properties generally have lease terms ranging from one to five years.

Under applicable PRC laws, lease agreements for properties must be registered with the local housing authorities. As of the Latest Practicable Date, registration had not been completed for two of our leased properties in China, both of which are located in Hangzhou. One of these leased properties is used for office purposes and has a gross floor area of 984.56 sq.m. According to the confirmation letter issued by the lessor of this leased property in June 2026, the property owner and other relevant parties have not yet provided the complete materials required for the lease registration, and therefore, the lease registration has not yet been completed as of the Latest Practicable Date. The other leased property is also used for office purposes and has a gross floor area of 260.25 sq.m. This property was recently leased, and the registration of this property has been completed on October 21, 2026. According to our PRC Legal Advisors, failure to register a lease may result in administrative orders to complete registration within a specified time limit and shall impose a fine between RMB1,000 and RMB10,000 per unregistered lease which fail to make corrections within the specified time limit, but does not affect the legal validity of the lease itself and will not result in relocation. We have not received any such orders or penalties to date. Our Directors believe that the lack of registration does not have any

material adverse impact on our operations or financial condition, and we will continue to engage with lessors and local authorities to complete the registration process where practicable. For further details, see "Risk Factors — Risks Relating to Our Operations — We may be liable for failure to register and file our lease agreements, which may subject us to penalties." We have established the leased asset management policy that covers job responsibilities, detailed procedures for leasing, and the filing and ongoing management of leased assets. As advised by the internal control consultant engaged by our Group, this policy is adequate and effective to prevent future occurrence of the similar incidents.

INSURANCE

During the Track Record Period, we provided mandatory social insurance for our employees as required by PRC social insurance regulations, such as pension insurance, unemployment insurance, work injury insurance and medical insurance. In addition, we maintain insurance coverage that is typical for a fabless semiconductor design company of our scale, including transportation insurance arranged either on a per-shipment basis or under an annual policy. Our Industry Consultant considers that the scope of our coverage is generally consistent with industry practice among comparable semiconductor design companies. Based on the above, our Directors are of the view that our insurance policy aligns with industry practice and complies with applicable requirements. During the Track Record Period, we had not been the subject of any product liability claims. See "Risk Factors — Risks Relating to Our Operations — We may not have sufficient insurance coverage to cover our potential liability or losses and as a result, our business, financial conditions, results of operations and prospects may be materially and adversely affected should any such liability or losses arise."

AWARDS AND RECOGNITIONS

During the Track Record Period, we received awards and recognition in respect of our Company and our chip products, significant ones of which are set forth below:

Year	Award/Title	Granting Authority
2021. . .	SAIL Top 30 Award — Highest Honor at World Artificial Intelligence Conference	SAIL Award Committee, composed of WAIC Presidium Experts, Shanghai AI Strategic Advisory Experts, and other industry, investment, and media experts
2021. . .	Outstanding AI Chip Application Case	China Electronics Standardization Institute
2022. . .	2022 Shanghai Specialized and New SMEs List	Shanghai Municipal Commission of Economy and Informatization
2022. . .	Top 10 Most Investable Incubation-stage Enterprises	Science and Technology and Economy Commission of Pudong New Area, Shanghai
2023. . .	National "Little Giant" Enterprise (Specialized and New)	Ministry of Industry and Information Technology of the PRC
2023. . .	High and New Technology Enterprise	Jiangsu Provincial Department of Science and Technology, Department of Finance of Jiangsu Province, and Jiangsu Provincial Tax Service of the State Taxation Administration
2025. . .	Major Innovation Team	Suzhou Municipal People's Government
2025. . .	Leading Scientific and Technological Talent	Suzhou Industrial Park Administrative Committee
2025. . .	Leading Scientific and Technological Talent	Suzhou Municipal People's Government

LICENSES, PERMITS AND APPROVALS

During the Track Record Period and up to the Latest Practicable Date, as advised by our PRC Legal Advisors, we had obtained all material licenses and permits required for our business operations (i.e. business license) in the PRC, and such business license had remained in full effect. The following table sets forth the material licenses and permits in connection with our business operations as of the Latest Practicable Date:

No.	Certificate/Permit Name	Holder	Scope	Issue Date	Expiry Date
1	Business License (營業執照)	Company	computer technology, etc.	December 31, 2025	/
2	Business License (營業執照)	Chengdu Denglin	technical services, etc.	January 21, 2026	/
3	Business License (營業執照)	Beijing Denglin	integrated circuit design, etc.	January 16, 2026	/
4	Business License (營業執照)	Suzhou Denglin Hanhai	software development, etc.	January 19, 2026	/
5	Business License (營業執照)	Suzhou Suanzhen	software development, etc.	January 20, 2026	/
6	Business License (營業執照)	Wuhu Denglin	information system integration services, etc.	January 20, 2026	/
7	Business License (營業執照)	Hangzhou Denglin Hanhai	information system integration services, etc.	November 6, 2024	/
8	Business License (營業執照)	Shanghai Suanzhen	technical services, etc.	January 21, 2026	/
9	High and New Technology Enterprise Certificate (高新技術企業證書).	Company	/	December 12, 2023	December 11, 2026
10	High and New Technology Enterprise Certificate (高新技術企業證書).	Chengdu Denglin	/	December 8, 2025	December 7, 2028
11	Specialized, Sophisticated, Distinctive and Innovative "Little Giant" Enterprise (專精特新"小巨人"企業)	Company	/	July 1, 2026	June 30, 2029
12	Registration of Consignee and Consignor of Import and Export Goods (進出口貨物收發貨人備案) . .	Company	/	May 20, 2020	December 31, 2099
13	Registration of Consignee and Consignor of Import and Export Goods (進出口貨物收發貨人備案) . .	Beijing Denglin	/	April 16, 2025	December 31, 2099

We intend to apply for renewal of our key licenses and permits as required by applicable laws and regulations, the procedures for which are expected to be initiated timely prior to their respective expiration date as applicable. We are actively following up for the renewal of any license, permit, approval or certificate which were due to expire soon as of the Latest Practicable Date and the successful renewal of which will be subject to our fulfillment of relevant requirements. As of the Latest Practicable Date, our Directors were not aware of any reason that would cause or lead to the non-renewal of our existing licenses, permits, certificates and approvals. Our PRC Legal Advisors are of the view that, as of the Latest Practicable Date, there was no substantial legal impediment for us to renew our existing licenses, permits, approvals and certificates in the PRC as long as we would comply with the relevant legal requirements.

ENVIRONMENTAL, SOCIAL AND GOVERNANCE

ESG Governance

ESG Management and Targets

We have integrated environmental, social and governance principles into our strategic planning, daily operations and risk management processes, and incorporated ESG-related matters into the Group's risk management and internal control system. To support the implementation of ESG management, we have established a structured governance framework comprising the Board of Directors, the ESG Strategy Committee and the ESG Working Group, with defined responsibilities and coordination mechanisms to facilitate the management of ESG-related risks and implementation of relevant measures. The Board of Directors is responsible for overseeing ESG matters, including reviewing ESG strategies and targets, overseeing ESG-related risks and ensuring alignment with the Group's overall strategy and risk management framework. The ESG Strategy Committee serves as the coordinating management body, responsible for formulating work plans and targets, coordinating ESG initiatives across departments, monitoring ESG risks, reviewing key performance data and implementation progress, and reporting to the Board. The ESG Working Group is responsible for data collection, management and internal verification, supporting the implementation of ESG measures across departments, and ensuring the accuracy and completeness of ESG disclosures. Under the above framework, our ESG-related risk management and internal control process covers risk identification, assessment, implementation of control measures, performance monitoring and follow-up improvement. Relevant functional departments identify ESG-related risks and control matters in the course of daily operations, including those relating to product responsibility, employee health and safety, supply chain compliance, anti-corruption, environmental management and climate-related matters. The ESG Working Group consolidates relevant information and performance data, conducts internal checks on the completeness and accuracy of such information, and follows up on implementation issues identified in the process. The ESG Strategy Committee reviews key ESG risks, implementation progress and performance data on a regular basis, coordinates corrective measures across departments where necessary, and reports material ESG matters and progress to our Board where appropriate.

Win-Win Collaboration and Co-creation

Product Responsibility

We have established a quality management system covering the entire process of research and development, procurement, production and after-sales support in compliance with applicable laws and regulations, including the *Law of the People's Republic of China on Product Quality* and the *Law of the People's Republic of China on the Protection of Consumer Rights and Interests*, to ensure product safety, stability and compliance. We implement a full-process quality control mechanism covering key stages including product design and development, verification and testing, supply chain management and delivery, with defined responsibilities and traceable processes, and continuously improve product quality through customer feedback, root cause analysis and corrective measures. For details of our product quality management and customer service, please see "Business — Sales and Marketing of Our Products." As of April 30, 2026, we had not experienced any product recall incidents due to quality reasons, nor had we received any customer complaints arising from quality issues.

Employee Development

We regard our employees as a fundamental driver of sustained innovation and steady development. Through standardized recruitment processes, protection of employees' lawful rights and interests, and a structured training framework, we enhance organizational vitality and employees' sense of belonging, providing strong talent support for long-term sustainable development. In terms of recruitment and diversity, we strictly comply with the *Labor Law of the People's Republic of China* and other relevant laws and regulations, and have established the *Human Resources Control Procedure* to ensure that recruitment processes are standardized and transparent. We respect diversity in gender, age, ethnicity and professional background, ensure equal opportunities in recruitment, promotion and remuneration, and cultivate an inclusive and collaborative working environment.

In terms of employee protection and well-being, in accordance with the *Employee Welfare Policy*, we provide statutory social insurance and benefits in accordance with applicable laws, implement standardized labor contract management, and maintain stable and harmonious labor relations. In respect of dismissal and termination of employment, we follow our internal human resources management procedures, under which termination during the probation period generally requires three days' written notice where the relevant employee is found not to meet the employment conditions, while termination after probation generally requires 30 days' written notice and internal review and approval by the relevant department supervisor, the human resources department and the chairman, as applicable. We also handle handover arrangements and salary and benefit settlement in accordance with applicable requirements. Through established communication and feedback channels, we respond to employees' reasonable requests in a timely manner. We also promote employees' physical and mental well-being and strengthen team cohesion through cultural and sports activities and team-building initiatives. Benefiting from a stable organizational culture and talent development mechanism, our core team turnover rate was 0 from 2021 to the four months ended April 30, 2026.

In terms of employee training and development, we attach importance to the continuous enhancement of employees' professional capabilities and overall competencies, establish a multi-level training mechanism to continuously enhance professional and technical capabilities. We organize on-site technical training by industry professionals and support employees in participating in external professional training and technical exchange programs based on role requirements. As of April 30, 2026, our employee training coverage rate was 100%, with an average of 5.2 training hours per employee.

Occupational Health and Safety

We attach great importance to employees' occupational health and workplace safety, regard safety management as an important foundation for steady operations, continuously improve relevant systems and management mechanisms, and foster a safe, standardized and orderly working environment. We comply with the *Law of the People's Republic of China on the Prevention and Control of Occupational Diseases*, the *Law of the People's Republic of China on Work Safety* and other relevant laws and regulations, and incorporate occupational health and safety management into our daily operational system.

In terms of employee health and safety initiatives, we have established an occupational health management mechanism to monitor employees' physical and mental well-being and organize regular health examinations in accordance with applicable laws and regulations. Taking into account the specific characteristics of different job roles, we pay close attention to ventilation, lighting, and equipment safety conditions within the office environment, and conduct regular inspections and maintenance to mitigate potential health risks. In addition, we enhance employees' awareness of health protection through internal communications and training, fostering a health management culture that combines self-protection with mutual supervision.

With respect to emergency management and safety training, we have formulated contingency plans covering fire incidents, power outages, natural disasters, and other emergencies, clearly defining responsibilities and response procedures. Through regular drills and safety training sessions, we strengthen employees' emergency response capabilities. Following each drill, we conduct reviews and

implement improvements to continuously optimize our emergency management mechanisms and enhance overall organizational resilience. Benefiting from our systematic and tiered health and safety management system, as of April 30, 2026, we had not recorded any work-related fatality incidents.

Sustainable Supply Chain

We attach great importance to the stability and sustainable development of the supply chain, continuously improve our supply chain management system, and adhere to the principles of quality first, compliant operations and controllable risk in supplier selection and cooperation. In line with the ISO 9001 quality management system concept and our internal management requirements, we carry out standardized supply chain management and promote the gradual implementation of environmental, social and governance requirements throughout the supply chain. For details of our supply chain management and supplier management, please see "Business — Our Supply Chain." As of April 30, 2026, we had not experienced any material non-compliance incidents in relation to our supply chain management.

Community Contribution

We fulfil our corporate social responsibilities by engaging in industry collaboration and supporting social inclusion, and maintain communication with academic institutions and industry participants to facilitate technology exchange and industry cooperation. In particular, we have engaged in technical exchange and industry discussions to support collaboration within the industry. In July 2022, we conducted technical exchange activities with the Department of Electronic Engineering and the Institute of Integrated Circuits of Tsinghua University, engaging in discussions on chip design optimization and advanced packaging technology pathways. In 2023, we participated in an industry roundtable forum organized by Alibaba Cloud, where we discussed industrial collaboration, supply chain development and green manufacturing. We also promote equal employment opportunities and social inclusion. As of April 30, 2026, we employed two employees with disabilities and provided corresponding work arrangements and support based on position requirements.

Anti-corruption

We attach great importance to integrity in business conduct and anti-corruption management and regard them as an important foundation of corporate governance and sustainable operations. We strictly comply with the *Law Against Unfair Competition of the People's Republic of China* and other relevant laws and regulations, and continuously improve our integrity management system to prevent risks relating to bribery, fraud, conflicts of interest and other improper conduct, ensuring that our operations are carried out in accordance with the law and in compliance with applicable requirements.

In terms of system development, we have formulated and implemented the *Internal Management Policy on Integrity in Business Conduct*, which expressly prohibits commercial bribery, kickbacks, improper benefit transfers and other improper transactions, and sets out arrangements for conflict of interest declarations, gifts and hospitality management and supplier engagement standards. At the same time, we have established a reporting and investigation mechanism for violations, with clearly defined reporting procedures, investigation processes and allocation of responsibilities, to ensure that relevant issues are addressed in a timely and impartial manner. For violations confirmed through investigation, we impose accountability measures in accordance with our internal policies and implement corrective actions, forming a feedback-based management mechanism covering risk identification, investigation and follow-up rectification. In 2025, 100% of our Directors received training on integrity in business conduct. In terms of governance effectiveness, as of April 30, 2026, we had not been involved in any legal cases relating to corruption, bribery or other major corrupt practices, and had not identified any material violations of business ethics.

Green Operations

Climate Change Response

Climate change has become a significant global environmental and development issue, exerting profound impacts on economic operations, industrial structures and the business environment. We actively respond to the national "dual carbon" goals and green low-carbon development strategy, strictly comply with relevant laws and regulations, incorporate climate-related risks and opportunities into our governance and risk management framework, continuously improve our greenhouse gas emissions management mechanisms, and promote the steady transition of our operations towards a low-carbon model.

Taking into account our industry characteristics and business model, we have conducted a systematic identification and assessment of climate-related risks and opportunities. In terms of risks, the principal risks we have identified include:

Risk type	Risk event	Risk level	Time horizon	Potential impact	Response measures
Physical risks . .	Typhoons, rainstorms and other extreme weather events;	Low	Short term	May affect the normal operation of our office and research and development facilities, potentially causing power interruptions or logistics delays.	We strengthen maintenance of office facilities, establish emergency response mechanisms and improve business continuity management arrangements.
Transition risks .	Policy and regulatory risks	Medium	Long term	Stricter carbon emission regulation, information disclosure requirements and other climate-related policies in China and globally may increase compliance costs.	We continuously monitor regulatory developments, improve greenhouse gas management and information disclosure mechanisms and strengthen internal ESG training.
	Specialized technology product risks	Low	Long term	Failure of our specialized technology products to meet increasingly stringent energy efficiency and environmental standards, regulatory requirements or customer expectations may adversely affect our reputation, customer relationships and business expansion.	We strengthen product energy efficiency optimization, monitor international environmental and energy efficiency standards and enhance environmental management across the product life cycle.

In terms of opportunities, we believe that climate change also brings opportunities for green technology upgrading and industrial transformation, including:

Opportunities	Opportunity overview	Potential impact	Response measures
Technology and research & development.	Global green and low-carbon transition drives increasing demand for high-performance and low-power chips.	Contributes to enhancing product value-added and market competitiveness.	We continue to optimize chip energy efficiency design and strengthen research and development investment.
Market and brand.	Growing customer focus on green supply chains and ESG performance.	Contributes to enhancing brand image and customer trust.	We strengthen ESG management across the supply chain and improve carbon emissions information disclosure.

In accordance with the *Guidelines for Accounting and Reporting of Greenhouse Gas Emissions by Enterprises (2022 Revision)*, the *2006 IPCC Guidelines for National Greenhouse Gas Inventories* and the *Greenhouse Gas Protocol: Corporate Accounting and Reporting Standard (2021)*, we have established a greenhouse gas emissions management framework covering the entire process of emission source identification, data collection, accounting and measurement, and target setting. From the perspective of emission source identification, as a fabless semiconductor design enterprise, our greenhouse gas emissions primarily arise from non-production operational activities and value chain-related activities. Scope 1 mainly includes direct greenhouse gas emissions from the use of company vehicles; Scope 2 mainly includes indirect greenhouse gas emissions from purchased electricity and water consumed at office and research and development premises; and Scope 3 mainly includes other indirect emissions generated from our business activities, including employee business travel and commuting, purchased goods and services, emissions arising from upstream raw materials and outsourced processing activities, emissions related to outsourced packaging and testing service providers and electronic manufacturing service providers, emissions from logistics and transportation activities, and other value chain-related activities.

Based on the identified emission sources, we conducted a thorough carbon inventory and statistical assessment, and calculated our total greenhouse gas emissions and emission intensity. As of 2025, our total greenhouse gas emissions amounted to 2,306.8 tons of carbon dioxide equivalent, of which Scope 1 emissions were 0.4 tons, Scope 2 emissions were 1,604.8 tons and Scope 3 emissions were 701.6 tons. Our greenhouse gas emission intensity was 13.3 tons of carbon dioxide equivalent per RMB million of revenue. To minimize greenhouse gas emissions, we continue to improve our greenhouse gas emissions management system and promote green office and low-carbon operational measures. In particular, we improve the energy efficiency of our office and research and development facilities by promoting energy-saving lighting, reasonably managing air-conditioning temperature settings and reducing standby energy consumption. We also optimize travel management arrangements, encourage low-carbon travel practices where practicable, and strengthen ESG management across our supply chain. In addition, we continue to optimize chip energy efficiency design and monitor relevant environmental and energy efficiency standards, so as to enhance environmental management across the product life cycle.

	For the year ended December 31,			For the four months ended April 30,
	2023	2024	2025	2026
Total greenhouse gas emissions (tons of CO2 equivalent)	1,596.2	1,823.3	2,306.8	971.4
Scope 1 — direct greenhouse gas emissions (tons of CO2 equivalent)	1.2	1.3	0.4	0.9
Scope 2 — indirect greenhouse gas emissions (tons of CO2 equivalent)	1,249.4	1,404.0	1,604.8	482.6
Scope 3 — other indirect greenhouse gas emissions (tons of CO2 equivalent)	345.6	418.0	701.6	487.9
Total greenhouse gas emission intensity (tons of CO2 equivalent/million of revenue)	57.1	22.1	13.3	4.8

Based on the above carbon inventory results, we conducted targeted data analysis and, taking 2025 as the base year, set the following phased emission reduction targets:

- Short-term target: By 2027, reduce greenhouse gas emission intensity by 2% compared with 2025;
- Medium- to long-term target: By 2030, reduce greenhouse gas emission intensity by 5% compared with 2025.

Waste Management

We comply with the *Law of the People's Republic of China on the Prevention and Control of Environmental Pollution by Solid Waste* and other relevant national and local regulations, and have established the *Waste Management Policy* to set out clear requirements for the classification, collection, storage and disposal of waste, forming a systematic and standardized full-process waste management mechanism. As a fabless semiconductor design enterprise, we primarily generate office waste, electronic waste and a small amount of hazardous waste in the course of daily operations. All such waste is subject to standardized management and full-process supervision to ensure that handling procedures are safe, compliant and traceable.

In terms of supervision and performance management, we incorporate waste management into our internal control system and conduct regular process inspections and ledger reviews to ensure that data is accurate, records are complete and disposal is compliant. As of 31 December 2025, our compliant waste disposal rate was 100%. On this basis, taking 2025 as the base year, we have set a phased management target: by 2027, maintain a 100% compliant disposal rate for both general waste and hazardous waste.

Resource Management

We adhere to a green and low-carbon operational philosophy, incorporate energy conservation and efficient resource utilization into our daily management requirements, continuously improve our energy and water resource management mechanisms, and enhance resource efficiency in office and research and development activities. We have established an energy and resource management system covering all departments, clarified the allocation of responsibilities, and incorporated electricity-saving and water-saving measures into routine inspection procedures. Administrative and logistics personnel conduct regular inspections of air-conditioning, lighting and water supply systems to ensure that resource use remains safe, standardized and efficient. Taking 2025 as the base year, we aim to improve energy and water use efficiency in our operations, reduce overall energy consumption intensity and water consumption intensity per unit by 2% by 2027, and maintain 100% compliance in wastewater discharge.

In terms of energy management, our energy consumption primarily arises from electricity use in office and research and development activities. In 2025, our total energy consumption amounted to 732.2 tons of standard coal equivalent, and our energy consumption intensity was 4.2 tons of standard coal equivalent per RMB million of revenue. Through optimizing lighting system layouts, promoting LED energy-saving lighting, reasonably setting air-conditioning temperature ranges and conducting energy-saving awareness initiatives, we continuously reduce energy consumption in our operations.

In terms of water resource management, our daily water use mainly arises from office and research and development activities. In 2025, our total annual water consumption amounted to 1,208.0 cubic meters. We reduce water waste and improve water use efficiency by using sensor-operated taps, installing water-saving reminder signage, conducting regular inspections of water supply systems and strengthening water conservation awareness initiatives.

	For the year ended December 31,			For the four months ended April 30,
	2023	2024	2025	2026
Total water consumption (cubic meters)	3,270.6	1,795.0	1,208.0	885.0
Total energy consumption (tons of standard coal)	565.3	637.1	732.2	220.7
Energy consumption intensity (tons of standard coal per RMB million of revenue)	20.2	7.7	4.2	1.1

LEGAL PROCEEDINGS AND COMPLIANCE

During the Track Record Period and up to the Latest Practicable Date, to the best knowledge of our Directors, we had not been and were not a party to any material legal, arbitral or administrative proceedings, and we were not aware of any material pending or threatened legal, arbitral or administrative proceedings against us or our Directors. To the best knowledge of our Directors and as advised by our PRC Legal Advisors, our business operations had been carried out in compliance with applicable laws and regulations in all material aspects during the Track Record Period and up to the Latest Practicable Date, except as disclosed elsewhere in this document.

During the Track Record Period and up to the Latest Practicable Date, our business activities did not involve any Sanctioned Countries or business partners subject to SDN sanctions or sanctions imposed by other Relevant Jurisdictions. Our business activities thus should not have implicated the applicable international sanctions and such sanctions did not have any material impact on our business.

During the Track Record Period and up to the Latest Practicable Date, our products incorporated certain U.S.-origin electronics and we also utilized certain U.S. technology in our GPU design process. That said, our International Sanctions Legal Advisors have advised that our products sold should not be subject to the Export Administration Regulations ("**EAR**"). This is because, judging by the Export Control Classification Numbers of the U.S.-origin electronics incorporated in our products, they are not deemed "controlled" under U.S. export controls and thus not counted in the *de minimis* calculation. Our products sold thus contained 0% of controlled U.S.-origin items, which is below the 25% *de minimis* threshold for items destined to our customers' countries. In addition, our products sold did not satisfy both the product scope and country and end-use/end-user scope of any Foreign Direct Product Rules. Moreover, none of our business partners would cause us to violate the EAR because we did not procure, export or otherwise transfer any items subject to the EAR in violation of applicable U.S. export control restrictions.

With respect to international tariffs, since we only sold our products directly to mainland China and Hong Kong during the Track Record Period and up to the Latest Practicable Date, our sales activities were not subject to international tariffs. Moreover, to the best knowledge of our Directors, our products were not indirectly offered to any overseas jurisdictions (including the U.S.). Therefore, our Directors believe that international tariffs have had no impact on us.

In relation to the Outbound Investment Rule, our International Sanctions Legal Advisors have advised that we are a covered foreign person due to our business activities, specifically those related to the design of GPGPUs, which is a covered activity under the Outbound Investment Rule, subject to different view of the Treasury. U.S.-person investments in us, including acquiring our non-public shares to facilitate the **[REDACTED]**, thus would constitute notifiable transactions and require notifications to the Treasury. That said, our International Sanctions Legal Advisors have advised that such U.S.-person investments should not be prohibited under the Outbound Investment Rule, because we do not engage in any covered activities described in the definition of "prohibited transactions."

Specifically,

- we do not develop or produce any electronic design automation software for the design of integrated circuits or advanced packaging, any semiconductor fabrication equipment, equipment for performing volume advanced packaging, or commodity, material, software, or technology designed exclusively for use in or with extreme ultraviolet lithography fabrication equipment;
- we do not engage in integrated circuits fabrication or packaging; and
- we do not design any integrated circuit that meets or exceeds the performance parameters in Export Control Classification Number 3A090.a in supplement No. 1 to 15 CFR part 774, or integrated circuits designed for operation at or below 4.5 Kelvin.

Our International Sanctions Legal Advisors have advised that 31 CFR 850.501 of the Outbound Investment Rule provides an exception to the notification requirement allowing U.S. persons to **[REDACTED]** in our Company's **[REDACTED]**, so long as the **[REDACTED]** do not afford the U.S. person rights beyond standard minority shareholder protections with respect to us.

Our Directors are of the view that the Outbound Investment Rule is not expected to have a material adverse impact on our business operations, financial performance or investment prospects, or the **[REDACTED]**, as none of our major shareholders, the Joint Sponsors and the Overall Coordinators is a U.S. person, and an exception allows U.S. persons to invest in our **[REDACTED]**.

RISK MANAGEMENT AND INTERNAL CONTROL

It is the responsibility of our Board to ensure that we maintain sound and effective internal controls and risk management system to safeguard our Shareholders' investment and our assets at all times. We maintain internal manuals setting out operating procedures, internal control procedures and other policies and guidelines. We have adopted and implemented structured risk management policies in various aspects of our business operations, such as financial reporting and compliance.

Our Board of Directors and our general manager are responsible for the establishment, updating and implementation of our internal control policies and systems, while our management team monitors the daily implementation of the internal control procedures and measures with respect to our functional teams.

Financial Reporting Risk Management

We have in place a set of accounting policies in connection with our financial reporting risk management. We have various procedures in place to implement accounting policies, and our financial team reviews our management accounts based on such procedures. We also provide regular training to our finance team members to ensure that they understand our financial management and accounting policies and implement them in our daily operations.

Compliance Risk Management

In order to effectively manage our compliance and legal risk exposures, we have adopted strict internal procedures to ensure the compliance of our business operations with the applicable rules and regulations. In accordance with these procedures, our in-house legal team performs the basic function of reviewing and updating the form of contracts we enter into with our customers, partners, and suppliers. Our legal team examines the contract terms and reviews all relevant documents for our business operations, including licenses and permits obtained by the counterparties to perform their obligations of business contracts and all the necessary underlying due diligence materials, before we enter into any contract or business arrangements.

Our in-house legal team is responsible for obtaining any requisite governmental pre-approvals or consents, including preparing and submitting all necessary documents for filing with relevant government authorities, within the prescribed regulatory timelines. We continuously improve our internal policies according to changes in laws, regulations and industry standards, and update internal templates for legal documents. We undertake compliance management over various aspects of our operations and employee activities. We have also established an accountability system in respect of employees' violations of laws, regulations and internal policies. In addition, we continually review the implementation of our risk management policies and measures to ensure our policies and implementation are effective and sufficient.