Senao AI Server

White Paper

Building Intelligent Infrastructure

from Edge to Enterprise

A comprehensive guide to building scalable AI, enterprise, and edge computing infrastructure with the Senao AI Server family.

Building Tomorrow's Compute Power

The Evolution of Compute Infrastructure

In an era where data is both a driver and a byproduct of every business process, computing infrastructure has evolved into a key enabler of digital transformation. Historically, enterprise servers were primarily deployed to support well-defined workloads such as file sharing, database hosting, email services, and centralized business applications. These environments were relatively predictable, and scaling infrastructure often meant simply adding more servers with similar specifications.

Today, the computing landscape has fundamentally changed. Organizations are embracing artificial intelligence (AI), virtualization, cloud-native applications, real-time analytics, and edge computing to accelerate innovation and improve operational efficiency. Modern infrastructure must therefore support a much broader range of workloads—from large-scale AI model training and enterprise virtualization to intelligent edge inference and distributed computing environments.

Rather than relying on a single server platform, enterprises increasingly require a flexible server portfolio that enables the right compute resources to be deployed for each workload. High-performance AI servers power model development and data-intensive computing, enterprise servers support virtualization and business-critical applications, while compact edge platforms bring AI intelligence closer to where data is generated.

The rapid growth of AI applications has also introduced new infrastructure challenges. Organizations must balance compute performance, memory capacity, storage bandwidth, accelerator support, power efficiency, and deployment flexibility. As AI workloads continue to evolve, selecting the appropriate server architecture becomes a strategic decision rather than simply a hardware purchase.

The Senao AI Server family is designed around this modern infrastructure philosophy. By providing platforms optimized for high-performance computing, enterprise applications, and edge AI deployment, organizations can build scalable infrastructure that adapts to both current requirements and future technological advancements.

The Senao AI Server Portfolio

To address the diverse requirements of modern AI and enterprise computing, the Senao AI Server portfolio offers five platforms optimized for different deployment scenarios.

Model	Primary Positioning	Typical Workloads
SR810 Series	High-Performance AI Computing	AI Training, HPC, Large Language Models, Data Analytics
SR710 Series	Enterprise AI & Virtualization	Virtualization, Private AI, Database, Enterprise Applications
SR610 Series	General-Purpose Enterprise Server	Business Applications, File Services, General Computing
SE210 Series	Edge AI Computing	Machine Vision, Industrial AI, Video Analytics, Edge Inference
SE110 Series	Compact Edge AI Platform	Remote Sites, AI Gateway, Smart Retail, Branch Deployment

Each platform shares a common design philosophy while being optimized for its target workloads, enabling organizations to deploy the right level of compute resources without unnecessary complexity or overprovisioning.

Compute as a Strategic Investment

Modern organizations no longer view computing infrastructure as simply an IT expense. Instead, server platforms have become strategic assets that directly influence innovation, operational efficiency, and business competitiveness. Whether supporting enterprise applications, accelerating AI development, enabling industrial automation, or powering intelligent edge services, the right computing platform enables organizations to respond faster to changing business demands.

As AI adoption continues to expand, infrastructure decisions must also consider long-term flexibility. Selecting the appropriate server architecture today can significantly reduce future deployment costs while simplifying infrastructure expansion as workloads evolve.

Enterprise server platforms are therefore evolving beyond traditional CPU performance. Modern systems combine high-performance processors, large memory capacity, high-speed storage, accelerator support, and flexible networking to create balanced computing environments capable of supporting diverse enterprise workloads.

Performance Must Match the Workload

Raw computing power alone does not guarantee better performance. Every workload has different infrastructure requirements, and the optimal server configuration depends on how computing, memory, storage, networking, and accelerator resources work together.

For example, AI model training demands high-core-count processors, multiple GPU accelerators, large memory capacity, and extremely high storage throughput. In contrast, enterprise virtualization requires

balanced CPU performance, scalable memory, and reliable storage, while edge AI deployments prioritize compact design, low power consumption, and real-time inference performance.

Rather than deploying identical servers across every environment, organizations increasingly benefit from matching infrastructure capabilities to specific application requirements.

Matching Infrastructure to Business Requirements

The Senao AI Server family is designed around this workload-oriented approach.

Instead of a one-size-fits-all platform, each server model is optimized for a specific deployment scenario:

Model	Optimized For
SR810 Series	AI Training, High Performance Computing, Large-scale Analytics
SR710 Series	Enterprise AI, Virtualization, Database Services
SR610 Series	General Enterprise Applications, Collaboration, File Services
SE210 Series	Edge AI, Industrial Automation, Machine Vision
SE110 Series	Remote AI Deployment, Smart Retail, IoT Gateway

This workload-driven architecture allows organizations to optimize performance, improve infrastructure utilization, and reduce total cost of ownership by selecting the most appropriate platform for each application rather than overprovisioning every deployment.

As enterprise computing continues to evolve, scalability becomes equally important as performance. Infrastructure must not only meet today's workload requirements but also provide sufficient flexibility to accommodate future AI applications, increasing data volumes, and evolving business demands. This balance between performance, scalability, and operational efficiency forms the foundation of modern AI infrastructure.

Looking Ahead

Building the Future with the Senao AI Server Family

As artificial intelligence becomes increasingly integrated into everyday business operations, organizations require more than high-performance hardware. They need infrastructure that is scalable, adaptable, and capable of supporting diverse workloads across data centers, enterprise environments, and distributed edge locations.

Future-ready infrastructure is no longer defined by a single server platform. Instead, it is built upon a portfolio of purpose-designed systems that work together to deliver the right balance of compute performance, power efficiency, storage capacity, and deployment flexibility.

The Senao AI Server family has been developed with this philosophy in mind—providing optimized platforms for AI computing, enterprise applications, virtualization, and edge intelligence while maintaining a unified architecture that simplifies deployment and long-term management.

Why Senao AI Servers?

Organizations adopting AI are often faced with a difficult question:

Which server should be deployed for each workload?

Rather than requiring customers to compromise with a single hardware platform, Senao provides purpose-built server platforms optimized for different deployment requirements.

Platform	Primary Role
SR810 Series	High-performance AI training and data-intensive computing
SR710 Series	Enterprise AI, virtualization, database, and mission-critical workloads
SR610 Series	General-purpose enterprise computing and business applications
SE210 Series	Industrial AI, machine vision, intelligent edge computing
SE110 Series	Compact AI gateway, remote offices, smart retail, and branch deployments

This portfolio-based approach allows organizations to build infrastructure that scales naturally as business requirements evolve, eliminating unnecessary hardware complexity while maximizing resource utilization.

One Architecture, Multiple Deployment Scenarios

Although each server platform is optimized for different workloads, they all share the same design philosophy:

- Enterprise-class reliability
- Modern processor architecture
- High-speed storage technologies
- Flexible expansion capabilities
- AI-ready infrastructure
- Long-term scalability

This unified platform strategy enables IT teams to standardize deployment, simplify lifecycle management, and reduce operational complexity while maintaining the flexibility required for modern AI applications.

Preparing for the Next Generation of AI

Artificial intelligence continues to evolve rapidly. New AI models, larger datasets, and increasingly distributed applications demand infrastructure capable of adapting without requiring complete hardware replacement.

By selecting server platforms designed around scalability, workload optimization, and architectural flexibility, organizations can confidently support today's enterprise requirements while preparing for tomorrow's AI-driven innovations.

The Senao AI Server family provides the foundation for this journey—enabling businesses to build intelligent infrastructure from enterprise data centers to distributed edge environments with confidence, efficiency, and long-term investment protection.

Demystifying AI Infrastructure:

Preparing for the Next Wave of Intelligence

Introduction

Artificial Intelligence (AI) has rapidly evolved from a research topic into a foundational technology that is transforming industries worldwide. Organizations are adopting AI to improve operational efficiency, automate decision-making, enhance customer experiences, and create new business opportunities. As AI applications continue to expand, infrastructure has become a critical factor in determining whether these initiatives can be deployed successfully and scaled efficiently.

However, AI infrastructure is often misunderstood. Many organizations assume that deploying AI simply requires purchasing the most powerful server available. In reality, successful AI deployment depends on selecting infrastructure that aligns with the intended workload, deployment location, and long-term operational requirements.

What Do We Mean by AI?

The term "Artificial Intelligence" encompasses a wide range of technologies, including machine learning (ML), deep learning (DL), computer vision, natural language processing (NLP), generative AI, and intelligent automation. Although these applications differ significantly, they all rely on one common foundation—computing infrastructure capable of processing large volumes of data efficiently.

Modern AI workloads generally fall into two categories:

- **AI Training** – Building and optimizing AI models using massive datasets and high-performance computing resources.
- **AI Inference** – Running trained AI models to generate real-time predictions, decisions, or insights closer to where data is generated.

Understanding this distinction is essential because each workload requires different infrastructure characteristics. While AI training prioritizes maximum computing performance and accelerator scalability, AI inference often emphasizes low latency, energy efficiency, and deployment flexibility.

AI Infrastructure Is Not One-Size-Fits-All

As organizations expand their AI adoption, they rarely rely on a single type of server. Instead, AI infrastructure typically consists of multiple computing layers working together.

Centralized enterprise servers provide the computing resources required for virtualization, enterprise applications, and AI model development, while edge AI platforms process data locally to reduce latency, improve responsiveness, and minimize bandwidth consumption.

This distributed computing model enables organizations to deploy AI where it delivers the greatest business value while maintaining centralized management and operational consistency.

Senao Perspective

The Senao AI Server family is designed to support every stage of this AI journey.

Rather than offering a single hardware platform for all workloads, Senao provides purpose-built server platforms optimized for different deployment scenarios—from centralized AI computing and enterprise virtualization to distributed edge AI applications.

By selecting the right platform for each workload, organizations can improve infrastructure efficiency, simplify future expansion, and maximize long-term return on investment.

Training, Inference, and the Compute Divide

Unlike traditional software, which follows predefined instructions, AI models are created through a training process in which large volumes of data are analyzed to identify patterns, optimize model parameters, and generate predictive capabilities. The effectiveness of AI depends not only on sophisticated algorithms, but also on the computing infrastructure that supports data processing, model development, and deployment.

AI workloads are generally divided into two major stages: **training** and **inference**.

Training is a computationally intensive process that requires powerful processors, large memory capacity, high-speed storage, and often multiple GPU accelerators. During this stage, AI models continuously process massive datasets to improve their accuracy and performance.

Inference, on the other hand, is the operational phase in which trained models generate predictions or make decisions using live data. Because inference frequently occurs close to users, devices, or production systems, infrastructure priorities shift toward low latency, deployment flexibility, and operational efficiency rather than maximum computing performance.

Although both stages are essential to AI, they place very different demands on computing infrastructure.

Selecting the Right Platform for Each AI Stage

Understanding the distinction between training and inference helps organizations avoid both under-provisioning and over-investment.

Rather than deploying identical infrastructure for every AI workload, organizations achieve greater efficiency by matching server capabilities to their operational requirements.

For example:

AI Stage	Infrastructure Priority	Recommended Senao Platform
AI Training	Maximum compute performance, GPU scalability, memory bandwidth	**SR810 Series**
Enterprise AI Services	Virtualization, databases, private AI applications	**SR710 Series**
General Enterprise Computing	Business applications and collaboration services	**SR610 Series**
Edge AI Inference	Machine vision, industrial automation, real-time analytics	**SE210 Series**
Compact Edge Deployment	Smart retail, branch offices, remote AI gateways	**SE110 Series**

This workload-oriented deployment strategy allows organizations to optimize infrastructure investments while maintaining a unified architecture across centralized and distributed computing environments.

Why This Matters

As AI adoption continues to grow, enterprises are increasingly deploying both centralized AI infrastructure and distributed edge intelligence simultaneously.

A well-designed AI infrastructure therefore requires more than powerful hardware—it requires selecting the right platform for each stage of the AI lifecycle. By aligning computing resources with workload characteristics, organizations can improve performance, simplify future expansion, and reduce total cost of ownership.

Choosing the Right Compute Platform

Artificial Intelligence places significantly different demands on computing infrastructure than traditional enterprise applications. Selecting the appropriate processor architecture is therefore one of the most important decisions when designing an AI-ready data center.

While CPU performance remains the foundation of every server platform, today's AI workloads also require scalable memory bandwidth, high-speed storage, accelerator expansion, and flexible networking.

The optimal platform is determined not only by processor performance, but by how the entire system architecture supports the intended workload.

Processor Architecture Matters

Different AI deployment scenarios require different processor capabilities.

For large-scale AI model development, high core counts, large memory capacity, and extensive PCIe expansion are essential to support multiple GPU accelerators and high-throughput storage.

For enterprise computing environments, balanced CPU performance provides the flexibility needed for virtualization, databases, collaboration platforms, and private AI services.

At the network edge, processor efficiency, integrated I/O, and compact system design become equally important for delivering real-time AI inference while minimizing power consumption and deployment complexity.

The Senao Compute Portfolio

Rather than relying on a single server architecture, the Senao AI Server family provides purpose-built platforms optimized for different compute requirements.

Platform	Compute Architecture	Best Suited For
SR810 Series	Intel® Xeon® 6 platform with extensive PCIe Gen5 expansion and support for multiple accelerator cards	AI Training, HPC, Large Language Models
SR710 Series	Dual Intel® Xeon® 6 platform with high memory capacity and up to 24 NVMe Gen5 drive bays	Enterprise AI, Virtualization, Database Services
SR610 Series	Dual Intel® Xeon® 6 platform in a compact 1U design	Business Applications, Cloud Infrastructure
SE210 Series	Intel® Xeon® Granite Rapids-D SoC with integrated high-speed networking	Edge AI, Machine Vision, Industrial Computing
SE110 Series	Intel® Atom® Grand Ridge SoC optimized for compact and energy-efficient deployments	Smart Retail, Remote Sites, AI Gateway

Beyond the Processor

Although the processor is often the first specification considered, overall system architecture ultimately determines AI performance.

For example, the **SR810 Series** provides extensive PCIe Gen5 expansion for accelerator deployment, making it well suited for GPU-intensive AI model training. The **SR710 Series** combines high memory capacity with flexible storage expansion to support enterprise-scale virtualization and AI services.

Meanwhile, **SE210 Series** and **SE110 Series** emphasize compact integration, efficient power consumption, and high-speed networking for distributed edge AI deployments.

By aligning processor architecture with real-world workload requirements, organizations can build infrastructure that delivers higher efficiency, better scalability, and lower total cost of ownership.

Accelerated Computing for Modern AI

As AI models become increasingly sophisticated, processor performance alone is no longer sufficient for many workloads. Modern AI infrastructure combines CPUs with hardware accelerators to dramatically improve training speed, inference performance, and overall system efficiency.

Rather than replacing the CPU, accelerators work alongside it. The CPU manages operating systems, data preparation, orchestration, and application logic, while GPUs and other accelerator devices execute highly parallel AI computations. This collaborative architecture enables organizations to process larger datasets, train more complex models, and deliver AI services more efficiently.

PCIe: The Foundation of AI Expansion

The ability to deploy AI accelerators depends largely on the server's expansion architecture.

PCI Express (PCIe) provides the high-bandwidth connectivity required for GPUs, AI accelerator cards, high-speed networking, and NVMe storage. As AI workloads continue to grow, PCIe Gen5 significantly increases data throughput, reducing communication bottlenecks between processors, accelerators, and storage devices.

A scalable PCIe architecture also gives organizations the flexibility to expand AI infrastructure as business requirements evolve, without replacing the entire server platform.

Accelerator Flexibility Across the Senao AI Server Family

Different AI workloads require different levels of accelerator capability.

Platform	Expansion Focus	Typical AI Deployment
SR810 Series	Up to eight PCIe Gen5 x16 expansion slots for high-density accelerator deployment	Large-scale AI training, LLM development, HPC
SR710 Series	Flexible PCIe Gen5 expansion supporting enterprise AI and virtualization	Private AI, databases, enterprise GPU computing
SR610 Series	PCIe Gen5 expansion for application-specific acceleration	Business applications and cloud infrastructure
SE210 Series	PCIe Gen5 expansion options for edge AI acceleration	Machine vision, industrial automation, intelligent analytics
SE110 Series	Integrated, power-efficient architecture optimized for compact edge deployments	AI gateways, smart retail, remote branch AI

Matching Accelerator Resources to Workloads

Not every AI application requires the same level of hardware acceleration.

Large language model (LLM) training and scientific computing typically benefit from multiple high-performance GPU accelerators, making platforms such as the **SR810 Series** an ideal choice. Enterprise AI services often require a balanced combination of CPU performance, storage, and selective accelerator expansion, where the **SR710 Series** provides an optimal foundation.

At the edge, workloads frequently prioritize low latency, compact system design, and efficient power consumption over maximum GPU density. The **SE210 Series** and **SE110 Series** therefore enable organizations to deploy AI inference closer to where data is generated, improving responsiveness while reducing network bandwidth requirements.

Building Infrastructure That Can Grow

A well-designed AI platform should support today's workloads while remaining adaptable for future expansion.

By combining modern processor architectures with PCIe Gen5 expansion, the Senao AI Server family allows organizations to start with the resources they need today and scale compute, networking, storage, and accelerator capabilities as AI adoption continues to evolve.

Memory and Storage

Unlocking the Full Potential of AI Workloads

High-performance processors and accelerators are only part of the AI computing equation. As AI models grow in size and enterprise datasets continue to expand, memory capacity and storage performance become equally critical to overall system efficiency.

Insufficient memory can limit model execution and virtualization density, while slow storage creates bottlenecks when loading datasets, checkpoints, and enterprise applications. A balanced infrastructure therefore requires compute, memory, and storage to work together as a unified architecture.

High-Bandwidth Memory for Modern Computing

Modern AI, virtualization, and enterprise workloads demand large memory capacity with high bandwidth to efficiently process multiple applications simultaneously.

The Senao AI Server family is designed to support these evolving requirements through DDR5 memory technology, delivering increased bandwidth, improved power efficiency, and greater scalability compared to previous generations.

Depending on deployment requirements, each platform provides an optimized memory architecture:

Platform	Memory Architecture	Best Suited For
SR810 Series	Up to **32 DDR5 RDIMM slots** for large-capacity AI and HPC workloads	AI model training, large datasets
SR710 Series	Up to **32 DDR5 ECC RDIMM slots**, supporting capacities up to **8 TB**	Enterprise AI, virtualization, databases
SR610 Series	Up to **32 DDR5 ECC RDIMM slots**	Enterprise applications and cloud services
SE210 Series	**8 DDR5 RDIMM slots** optimized for edge computing	Industrial AI, machine vision
SE110 Series	**2 DDR5 SO-DIMM slots** for compact edge deployments	Smart retail, branch offices, AI gateways

This tiered memory architecture enables organizations to select the most appropriate platform based on workload characteristics rather than deploying excessive resources across every environment.

High-Speed NVMe Storage

AI workflows frequently involve moving and processing extremely large volumes of data. From training datasets and AI model checkpoints to enterprise databases and virtual machine images, storage performance has a direct impact on application responsiveness and overall productivity.

To meet these requirements, the Senao AI Server family incorporates NVMe technology across the product portfolio.

- **SR810 Series** provides hot-swappable NVMe storage for high-performance AI and HPC environments.
- **SR710 Series** supports **up to 24 hot-swappable PCIe Gen5 NVMe drive bays**, enabling exceptional throughput for enterprise AI, virtualization, and large-scale database applications.
- **SR610 Series** supports **up to 12 PCIe Gen5 NVMe drive bays**, providing an excellent balance between storage density and system footprint.
- **SE210 Series** combines U.2 NVMe hot-swappable storage with M.2 options, offering flexible storage configurations for edge deployments.
- **SE110 Series** integrates M.2 NVMe storage in a compact platform optimized for embedded and distributed AI applications.

Balancing Capacity with Performance

The optimal storage architecture depends on application requirements.

Large AI training environments benefit from maximum throughput and rapid access to training datasets, while enterprise virtualization emphasizes storage density and reliability. Edge AI deployments often prioritize compact system integration and simplified maintenance without compromising responsiveness.

By offering multiple storage architectures across the Senao AI Server family, organizations can deploy the right level of performance for each workload while maintaining a consistent infrastructure strategy.

A Foundation for Scalable AI Infrastructure

As AI adoption accelerates, infrastructure requirements continue to evolve.

The combination of DDR5 memory technology and NVMe storage enables the Senao AI Server family to support increasingly demanding AI models, enterprise applications, and distributed edge services. Whether deployed in centralized data centers or remote edge locations, these platforms provide the scalability and responsiveness required for future-ready computing.

Building Connected AI Infrastructure

Modern AI infrastructure extends far beyond individual servers. As organizations deploy AI across data centers, enterprise campuses, and edge locations, seamless connectivity, centralized management, and scalable networking become essential components of a successful deployment.

An effective AI platform is therefore built not only on compute performance, but also on how efficiently systems communicate, exchange data, and can be managed throughout their operational lifecycle.

High-Speed Networking for AI Workloads

AI applications generate significantly more data than traditional enterprise workloads.

Large training datasets, distributed storage, virtualization clusters, and real-time inference all depend on high-bandwidth networking capable of moving information quickly and reliably between compute nodes, storage systems, and users.

The Senao AI Server family supports modern high-speed networking technologies that enable organizations to build scalable AI infrastructures while reducing communication bottlenecks.

Depending on deployment requirements, networking capabilities are optimized across different platforms:

Platform	Networking Focus	Typical Deployment
SR810 Series	High-speed networking with flexible PCIe Gen5 and OCP NIC expansion	AI clusters, HPC environments
SR710 Series	Enterprise networking with scalable expansion for virtualization and private AI	Enterprise data centers
SR610 Series	Standard enterprise networking for business-critical applications	Cloud infrastructure and business services
SE210 Series	Integrated **25GbE SFP28** and **10GbE** networking designed for industrial edge AI	Smart manufacturing, intelligent transportation
SE110 Series	High-density **2.5GbE**, **10GbE**, and **25GbE** connectivity for compact edge deployments	Retail, branch offices, remote AI sites

By selecting the appropriate networking architecture for each deployment scenario, organizations can improve system responsiveness while simplifying future infrastructure expansion.

Intelligent Remote Management

Managing AI infrastructure efficiently is just as important as delivering computing performance.

As server deployments continue to grow, IT administrators require centralized tools that enable remote monitoring, health diagnostics, firmware maintenance, and system recovery without requiring physical access to each device.

The Senao AI Server family incorporates enterprise-class server management capabilities, supporting industry-standard technologies including:

- IPMI 2.0
- Redfish API
- Dedicated BMC management
- Remote health monitoring
- Remote firmware updates
- Power and thermal monitoring

These capabilities help reduce operational complexity while improving service availability across enterprise and edge deployments.

Designed for Long-Term Operations

AI infrastructure is expected to remain in service for many years.

Organizations therefore require platforms that not only deliver high performance today but also simplify maintenance, hardware upgrades, and future expansion.

Through standardized management interfaces, flexible networking options, and modular expansion architecture, the Senao AI Server family enables IT teams to manage diverse server platforms using a consistent operational model, reducing administrative overhead while improving deployment efficiency.

Connecting Compute, Storage, and Intelligence

Successful AI infrastructure depends on the integration of computing, networking, storage, and management into a unified platform.

By combining enterprise-class networking with intelligent management capabilities, the Senao AI Server family enables organizations to build AI environments that are scalable, reliable, and ready to support future business growth.

Deploying AI Across Enterprise and Edge

Artificial Intelligence is no longer confined to centralized data centers. Modern organizations increasingly deploy AI across multiple environments, combining centralized computing resources with distributed edge intelligence to support a wide range of business applications.

Rather than relying on a single server platform, successful AI infrastructure is built using purpose-designed systems that work together to deliver the appropriate level of compute performance, storage capacity, networking, and operational efficiency for each workload.

A Layered AI Infrastructure

Most enterprise AI deployments can be viewed as three interconnected layers:

AI Development Layer

This layer focuses on model development, AI training, and large-scale data processing. These workloads require maximum compute performance, extensive accelerator expansion, and high memory bandwidth.

The **Senao SR810 Series** is designed to address these requirements with its Intel® Xeon® 6 platform, high-density PCIe Gen5 expansion, and enterprise-class architecture for AI and HPC environments.

Enterprise Services Layer

Once AI models are trained, organizations must integrate them into daily business operations.

This layer supports virtualization, enterprise databases, private AI services, ERP systems, collaboration platforms, and application servers that require balanced compute performance, high availability, and scalable storage.

The **Senao SR710 Series** provides an ideal platform for enterprise AI infrastructure, while the **SR610 Series** delivers an efficient solution for general-purpose enterprise workloads and cloud services.

Edge Intelligence Layer

Many AI applications require decisions to be made where data is generated rather than transmitting all information back to a centralized data center.

Industrial automation, machine vision, smart transportation, retail analytics, and IoT deployments all benefit from localized AI inference, reducing latency while improving responsiveness and bandwidth efficiency.

The **Senao SE210 Series** is optimized for industrial edge AI environments with integrated high-speed networking and flexible expansion capabilities. For compact deployments such as branch offices, smart retail, and remote AI gateways, the **SE110 Series** provides an efficient, space-saving platform while maintaining enterprise-class reliability.

A Unified AI Infrastructure Strategy

Although each platform serves a different purpose, they are designed to complement one another as part of a unified AI infrastructure.

Infrastructure Layer	Senao Platform	Primary Purpose
AI Development	**SR810 Series**	AI Training, HPC, Large Language Models
Enterprise AI	**SR710 Series**	Private AI, Virtualization, Enterprise Databases
Enterprise Services	**SR610 Series**	Business Applications, Cloud Infrastructure
Edge AI	**SE210 Series**	Machine Vision, Industrial Automation
Remote Edge	**SE110 Series**	Smart Retail, Branch AI, IoT Gateway

This layered deployment model enables organizations to place computing resources where they create the greatest business value while maintaining a consistent operational framework across the entire infrastructure.

Building an AI Journey That Can Scale

Few organizations deploy AI at full scale from day one.

Most begin with a limited number of AI projects before expanding into enterprise-wide deployments. As adoption grows, infrastructure must scale without requiring a complete redesign.

The Senao AI Server family enables organizations to expand AI capabilities incrementally—from centralized AI development to enterprise services and distributed edge intelligence—using platforms

specifically optimized for each stage of growth.

Accelerating AI Transformation

Building AI infrastructure is only the first step toward digital transformation. The true value of AI is realized when organizations successfully integrate intelligent computing into daily operations, enabling faster decision-making, improved operational efficiency, and new business opportunities.

Rather than deploying AI as a standalone technology, leading organizations are incorporating AI across multiple business functions—from enterprise operations and manufacturing to retail, logistics, healthcare, and smart city applications.

AI Adoption Is a Journey

Most organizations do not deploy AI infrastructure all at once. Instead, adoption typically progresses through several stages, allowing IT teams to validate workloads, optimize infrastructure, and expand AI capabilities as business requirements evolve.

Phase 1 – Establish the Foundation

Organizations begin by deploying reliable enterprise infrastructure capable of supporting virtualization, business applications, and early AI initiatives.

Recommended Platforms

- **SR710 Series** – Enterprise AI, virtualization, and private AI services
- **SR610 Series** – Business applications, cloud infrastructure, collaboration services

Phase 2 – Scale AI Computing

As AI adoption increases, organizations expand computing resources to support model development, advanced analytics, and high-performance AI workloads.

Recommended Platform

- **SR810 Series** – AI model training, large language models (LLMs), high-performance computing (HPC), and data-intensive analytics

Phase 3 – Extend Intelligence to the Edge

Once AI models are deployed, inference increasingly moves closer to where data is generated. This reduces latency, improves responsiveness, and minimizes bandwidth consumption.

Recommended Platforms

- **SE210 Series** – Industrial automation, machine vision, intelligent transportation
- **SE110 Series** – Smart retail, branch offices, IoT gateways, distributed edge AI

One Platform Family, Multiple AI Journeys

Although every organization follows a different path toward AI adoption, the underlying infrastructure should remain consistent.

The Senao AI Server family enables customers to build AI environments incrementally, selecting the most appropriate platform for each stage of growth while maintaining a unified operational framework.

This modular deployment strategy helps organizations:

- Accelerate AI implementation
- Optimize infrastructure investment
- Reduce operational complexity
- Simplify future expansion
- Protect long-term technology investments

From Infrastructure to Business Value

Successful AI initiatives are measured not only by computing performance, but by measurable business outcomes.

By aligning infrastructure with business objectives, organizations can deploy AI more efficiently, improve resource utilization, and create a scalable foundation capable of supporting future innovation.

Whether supporting enterprise applications, developing AI models, or enabling intelligent edge services, the Senao AI Server family provides a flexible platform designed to grow alongside business requirements.

Designing AI Infrastructure for Long-Term Success

As organizations expand their AI initiatives, infrastructure decisions must consider more than immediate performance requirements. A successful AI platform should be designed to support future growth, simplify ongoing operations, and protect long-term technology investments.

Rather than replacing infrastructure as workloads evolve, organizations increasingly seek platforms that can scale incrementally while maintaining operational consistency across data centers and distributed edge environments.

Scalability by Design

AI workloads rarely remain static.

As datasets become larger, AI models grow more sophisticated, and business requirements continue to evolve, infrastructure must be capable of expanding without disrupting existing operations.

The Senao AI Server family supports this approach through a portfolio of platforms optimized for different deployment scales—from compact edge computing to enterprise AI and high-performance computing.

Organizations can begin with the platform that best fits current requirements and expand their AI infrastructure over time while maintaining a consistent architecture.

Simplified Infrastructure Management

Managing AI infrastructure becomes increasingly challenging as deployments grow across multiple locations.

Centralized management capabilities help IT teams monitor system health, perform remote maintenance, deploy firmware updates, and maintain operational visibility throughout the infrastructure lifecycle.

With enterprise-class management features such as dedicated BMC, IPMI 2.0, and Redfish support, the Senao AI Server family enables consistent management across server platforms, reducing operational complexity and improving service availability.

Protecting Technology Investments

Infrastructure investments should continue delivering value as technology evolves.

Modern AI platforms should support future processor generations, high-speed networking, accelerator expansion, and increasing storage requirements without requiring complete infrastructure replacement.

By adopting standardized architectures based on DDR5 memory, PCIe Gen5 expansion, NVMe storage, and enterprise management technologies, organizations gain the flexibility to adapt to future AI workloads while maximizing return on investment.

A Unified Platform for Future Growth

Although AI applications continue to evolve, the underlying infrastructure principles remain consistent.

The Senao AI Server family is designed to provide a unified platform that supports organizations throughout every stage of their AI journey—from enterprise computing and AI development to distributed edge intelligence.

This architecture enables businesses to scale confidently while maintaining operational efficiency, reducing complexity, and preparing for the next generation of AI innovation.

AI Infrastructure in Action

Artificial Intelligence delivers the greatest value when it is integrated into real business operations. From enterprise IT and manufacturing to retail and smart cities, organizations are increasingly adopting AI to automate processes, improve operational efficiency, and generate actionable insights.

While every industry has unique requirements, successful AI deployments share one common characteristic—they rely on infrastructure that is optimized for the intended workload and deployment environment.

Enterprise AI

Many organizations are deploying AI to improve productivity across daily business operations.

Common workloads include:

- Private AI assistants
- Intelligent document search (RAG)
- Enterprise knowledge management
- Database analytics
- Virtualization
- Business intelligence

These applications require balanced compute performance, scalable memory, and reliable storage.

Recommended Platforms

- **SR710 Series** – Enterprise AI, virtualization, database services
- **SR610 Series** – General business applications and cloud infrastructure

AI Development and Data Analytics

Organizations developing proprietary AI models require significantly greater computing resources.

Typical workloads include:

- Large Language Model (LLM) development
- AI model training
- High-performance computing (HPC)
- Scientific simulation
- Data analytics

These environments demand maximum compute density, accelerator expansion, and high-bandwidth memory.

Recommended Platform

- **SR810 Series** – AI training and large-scale compute environments

Industrial Edge AI

Industrial AI applications require decisions to be made immediately where data is generated.

Examples include:

- Machine vision inspection
- Predictive maintenance
- Robotics
- Factory automation
- Quality assurance

These workloads benefit from localized AI inference, reducing latency while improving operational efficiency.

Recommended Platform

- **SE210 Series** – Industrial edge AI and machine vision

Smart Retail and Distributed Edge

Many organizations operate across hundreds or thousands of distributed locations.

Examples include:

- Smart retail analytics
- Self-service kiosks
- Digital signage
- Branch office AI
- IoT gateway services

Compact, power-efficient platforms simplify deployment while enabling real-time AI processing.

Recommended Platform

- **SE110 Series** – Compact edge AI deployment

One Infrastructure, Many Industries

Although AI applications continue to diversify, the infrastructure principles remain remarkably consistent.

By selecting the appropriate platform for each deployment scenario, organizations can create a unified AI environment that supports centralized computing, enterprise services, and distributed intelligence without introducing unnecessary operational complexity.

The Senao AI Server family enables customers to build scalable AI infrastructure that adapts to evolving business requirements across multiple industries and deployment models.

Why the Senao AI Server Family?

Artificial Intelligence continues to evolve rapidly, but the principles of successful infrastructure remain consistent. Organizations require platforms that are scalable, manageable, and capable of supporting a wide range of workloads without introducing unnecessary operational complexity.

Rather than offering a single server designed for every application, the Senao AI Server family provides a portfolio of purpose-built platforms optimized for AI computing, enterprise services, and distributed edge intelligence.

This architectural approach enables organizations to deploy the right level of computing resources where they deliver the greatest business value.

Designed Around Workloads

Different workloads require different infrastructure characteristics.

Instead of optimizing every platform for maximum performance, Senao designs each server to address a specific operational requirement.

Platform	Primary Value
SR810 Series	Maximum compute performance for AI training, HPC, and large-scale analytics
SR710 Series	Balanced enterprise platform for virtualization, private AI, and mission-critical applications
SR610 Series	Cost-effective infrastructure for business applications and cloud services
SE210 Series	High-performance edge AI platform for industrial environments and intelligent automation
SE110 Series	Compact, energy-efficient platform for distributed AI and remote deployments

This workload-centric strategy enables organizations to optimize both infrastructure performance and investment efficiency.

Built for Enterprise Operations

Modern AI infrastructure must remain operational, secure, and manageable throughout its lifecycle.

The Senao AI Server family incorporates enterprise-class capabilities including:

- Intel® Xeon® 6 and Intel® Xeon® D processor platforms
- DDR5 memory architecture
- PCIe Gen5 expansion
- NVMe storage technologies
- High-speed Ethernet connectivity
- Dedicated BMC management
- IPMI 2.0 and Redfish support

Together, these technologies provide a robust foundation for enterprise AI deployments while simplifying long-term operations and system management.

Scalable from Enterprise to Edge

AI infrastructure is no longer limited to centralized data centers.

Organizations increasingly deploy AI across headquarters, regional offices, manufacturing facilities, retail stores, and remote locations.

The Senao AI Server family enables this distributed architecture through complementary platforms that support:

- Centralized AI development
- Enterprise applications
- Private AI services

- Industrial AI
- Smart retail
- Remote edge computing

By maintaining a unified architecture across these environments, organizations can simplify deployment, reduce operational complexity, and accelerate AI adoption.

Ready for the Future

As AI technologies continue to evolve, infrastructure must remain flexible enough to support new applications, increasing data volumes, and future hardware expansion.

The Senao AI Server family provides organizations with a scalable foundation that supports growth from initial AI projects to enterprise-wide deployments—helping businesses adopt AI with confidence while protecting long-term technology investments.

Conclusion

Building the Future of AI with Senao

Artificial Intelligence is transforming the way organizations operate, innovate, and compete. As AI applications continue to expand across enterprise data centers, manufacturing facilities, retail environments, and distributed edge locations, infrastructure has become the foundation that enables successful AI adoption.

However, no single server platform can efficiently address every workload. Organizations require an infrastructure strategy that balances performance, scalability, operational efficiency, and long-term investment protection.

The Senao AI Server family has been designed with this philosophy at its core.

A Complete AI Infrastructure Portfolio

Rather than offering one platform for every application, Senao provides purpose-built servers optimized for each stage of the AI journey.

Platform	Primary Role
SR810 Series	AI Training, High-Performance Computing, Large Language Models
SR710 Series	Enterprise AI, Virtualization, Private AI
SR610 Series	Business Applications and Enterprise Services
SE210 Series	Industrial AI, Machine Vision, Intelligent Edge
SE110 Series	Smart Retail, Remote Office, IoT Gateway

Together, these platforms enable organizations to build AI infrastructure that aligns with business objectives while maintaining a unified operational framework.

Designed for Sustainable Growth

The future of AI will continue to evolve.

New workloads, larger datasets, and increasingly distributed applications will require infrastructure that is not only powerful, but also flexible enough to adapt without unnecessary complexity.

By combining modern processor architectures, high-speed memory, scalable storage, enterprise networking, and intelligent management, the Senao AI Server family provides a reliable foundation for organizations planning long-term AI adoption.

Why Senao

The Senao AI Server family is built around four guiding principles:

Performance

Purpose-built platforms optimized for AI, enterprise computing, and edge intelligence.

Scalability

A comprehensive product portfolio supporting growth from initial AI projects to enterprise-wide deployments.

Simplicity

Unified management capabilities and standardized enterprise technologies simplify deployment and lifecycle management.

Investment Protection

Future-ready system architectures help organizations maximize infrastructure value while preparing for evolving AI requirements.

Looking Forward

Artificial Intelligence is no longer a future technology—it is rapidly becoming an essential part of modern business.

Whether developing AI models, modernizing enterprise infrastructure, or deploying intelligent edge applications, organizations need a trusted technology foundation that can evolve alongside changing business demands.

The Senao AI Server family delivers that foundation—providing scalable, reliable, and purpose-built infrastructure for every stage of the AI journey.

From Enterprise to Edge.From Today's Workloads to Tomorrow's Innovation.