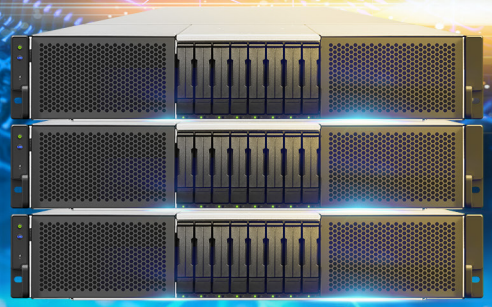
Infortrend®

Explore More



Enterprise Cloud

Highlights

High Performance

- One KS provides up to 1M IOPS, 30 GB/s read, and 25 GB/s write throughput
- Selected KS models feature dual Intel® Xeon® Scalable 64-core or AMD EPYC™ 9004/9005 Series 160-core CPUs, along with up to 4 Nvidia GPUs

Comprehensive Data Protection

- Multi-node data replication ensures no data loss
- Remote backup of VMs and application related data to cloud or storage

User-Friendly Management

- EonOne for managing compute, GPU, and storage resources
- EonKube for managing VMs and all deployed applications

Scalability

- Expand KS cluster up to 20 nodes
- Expand storage with EonStor GS and DS solutions

Flexible Hardware Options

- Five models with configurable GPU options for various application needs

Support for Different Application Types

- Run cloud-native applications, traditional applications, and virtual machines

Introduction

Enterprise Cloud is a private cloud platform that integrates compute, storage, and Kubernetes management into a single solution. It offers cloud capabilities such as scalability and reliability and enables enterprises to run both cloud-native and legacy applications. Deployment starts with three identical compute nodes, each equipped with high-performance CPUs, GPUs, and NVMe SSDs, making the platform well suited for AI and other high-performance workloads. The platform supports node protection and allows for scale-out to enhance performance. Additionally, it provides high-performance storage to meet demanding data storage needs.

High Performance

IEC compute nodes, KS, deliver outstanding storage performance, achieving 1M IOPS, 30 GB/s read, and 25 GB/s write throughput with high-speed U.2 NVMe SSDs. Offering flexible and powerful computing options, KS includes two processor configurations: KSi models with dual Intel Xeon® Scalable processors, capable of scaling up to 1280-core computing power in a cluster, and KSa models with dual AMD EPYC™ 9004 or 9005 Series processors, supporting up to 320 cores per appliance and 6400 cores per cluster.

Furthermore, the system supports GPU computing, making it ideal for AI, deep learning, HPC, and complex data analysis that demand extensive parallel processing capabilities.

Comprehensive Data Protection

KS supports multi-node data replication (Replica x2 and x3), ensuring no data loss in the event of node or hard disk failure. To protect applications, KS can regularly back up data to remote locations, offering flexible backup options, including backup of VMs and application-related data. This ensures fault tolerance and comprehensive disaster recovery capabilities of KS.

User-Friendly Management

KS offers a user-friendly solution for efficient management with EonOne and EonKube graphical user interfaces. EonOne provides comprehensive hardware management for compute, GPU, and storage resources. Users can effortlessly monitor the resource status of each node and the overall cluster, making it simple to identify hardware issues. This enhances management efficiency and ensures optimal performance across the system.

EonKube is designed for managing software, including VMs and all deployed applications. It provides a simplified Kubernetes application deployment, making it more accessible for enterprises by eliminating the need for complex configuration and coding. The EonKube dashboard offers a clear view of system resources, such as CPU and GPU usage for each application, allowing administrators to monitor and adjust resource allocation as needed. Users can also easily check for the latest versions of built-in applications maintained and verified by Infortrend. Updates can be applied with a single click, streamlining the process and eliminating manual version checks, update package searches, and testing.

BMC (Baseboard Management Controller) support allows IT teams to remotely manage and monitor KS systems in both data center and edge environments.

Scalability

KS features a scalable architecture that starts with a three-node cluster and enables seamless expansion by non-disruptively adding one KS node at a time, up to a maximum of twenty nodes. As enterprises' compute and storage requirements grow, they can easily add KS nodes to expand the cluster in line with business and application needs. This flexible scale-out design helps optimize resource utilization while supporting growing workloads and performance demands.

To expand storage capacity, KS supports various storage expansion options, allowing connection to EonStor GS unified storage or EonStor DS SAN storage.

Flexible Hardware Options

KS offers five compute node models with configurable options for GPUs and U.2 NVMe SSDs, allowing enterprises to select the model that best fits their application needs. All KS models come in a space-efficient 2U form factor and support 25GbE and 100GbE network interfaces. Designed for standard enterprise data centers and high-performance workloads, the KS 5000U series is ideal for AI training, deep learning, HPC, and other compute-intensive applications. The series includes the following models:

- **KSi 5008U / KSa 5008U** is designed as a GPU-intensive computing node, supporting up to four high-performance Nvidia GPU cards.
- **KSi 5016U / KSa 5016U** is a general-purpose computing node with two Nvidia GPU cards, delivering efficient performance for diverse workloads.
- **KSi 5024U i**s a storage-oriented computing node supporting 24 SSDs.

Support for Cloud-Native and Traditional Applications

KS features a built-in Kubernetes platform for running cloud-native containerized applications. With a built-in hypervisor, KS provides support for traditional applications by running VMs. This facilitates enterprises to run all workloads on a single platform. Furthermore, KS offers a built-in Application Marketplace with a diverse selection of applications for multiple domains and industries. The Marketplace includes commonly used enterprise applications, such as database systems, big data analytics platforms, and software development tools like CI/CD, as well as domain-specific applications for HPC, AI, and machine learning (ML), along with tools for managing related tasks. To minimize maintenance efforts, KS supports seamless one-click deployment and rolling updates, ensuring that applications stay up to date.

High Availability

The high-availability design of KS features the advanced failover function. The system can automatically detect node failures and quickly perform a failover, transferring applications to healthy nodes. Faulty node can be completely removed from and rejoined to the cluster system after repair, without requiring system downtime. This greatly enhances availability and ensures the continuity of enterprise applications.

Automatic Scaling

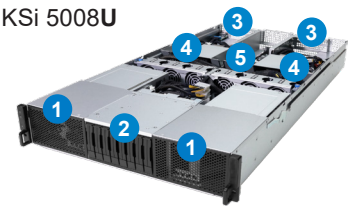
KS can dynamically scale cloud-native applications across nodes according to real-time node load conditions. It achieves this by continuously monitoring the hardware load of computing nodes, including metrics such as CPU, GPU, memory, storage space, and network usage. In response to heightened computing demands on an application, the system reallocates resources by automatically scaling the application to underutilized computing nodes. Conversely, when applications no longer require additional system resources, KS seamlessly downscales the allocated resources to optimize efficiency. By dynamically adjusting to changing demands, the automated scaling mechanism helps to ensure consistent application performance and optimized resource utilization.

Easy Deployment

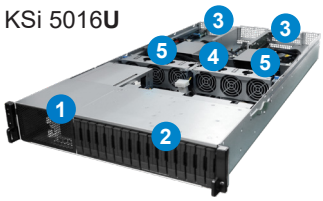
KS combines computing, storage, and Kubernetes platform within a single node appliance. This greatly simplifies deployment and management as KS eliminates the need for users to purchase these components separately and navigate through complex configurations, which can otherwise require substantial time, finances, and technical expertise. The platform features an intuitive installation wizard that guides users through the entire setup process, which is completed within 30 minutes. After setup, enterprises can immediately start installing and running applications. The initial cluster setup can start with three KS or more computing nodes, and expansion is easy – users can simply add KS nodes one by one, with each node requiring just a few simple steps through the management interface. Alternatively, users can start with one KS node in a setup mode that does not support expansion.

KSi 5000U Series

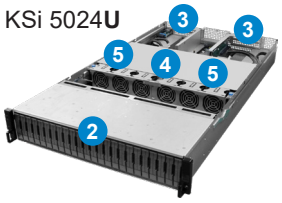
KSi 5008**U**



KSi 5016**U**



KSi 5024**U**



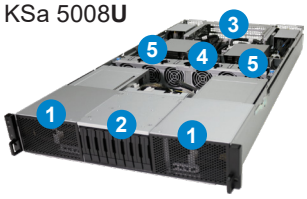
1. FHFL dual-width GPU slots
2. NVMe SSD slots
3. Network adapter slots
4. DDR5 DIMM slots
5. Intel® Xeon® Scalable Processor (4th or 5th Gen)

PHYSICAL SPECIFICATIONS

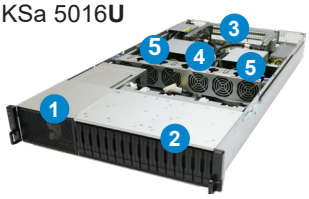
Product Series		KSi 5000U			
Model		KSi 5008**U**		KSi 5016**U**	KSi 5024**U**
Form Factor		2U 8-bay		2U 16-bay	2U 24-bay
Mounting Options		Rackmount			
Processor		Dual 4th Gen Intel® Xeon® Scalable Processors	Dual 5th Gen Intel® Xeon® Scalable Processors	Dual 4th Gen Intel® Xeon® Scalable Processors	
Processor Core Available		12 to 32 cores	64 cores	12 to 32 cores	12 to 32 cores
Cache Memory		• 32 x DDR5 DIMM slots • Default 128GB, up to 2TB DDR5 RDIMM			
Supported Drives		2.5" U.2 NVMe SSD			
		Note: 1. A minimum of four drives is required. 2. For the latest Compatibility Guide, refer to our official website.			
PCIe Expansion Slots		• 4 x FHFL PCIe Gen5 x16 slots for GPUs • 2 x FHHL PCIe Gen5 x16 slots • 2 x OCP NIC 3.0 PCIe Gen5 x16 slots		• 2 x FHFL PCIe Gen5 x16 slots for GPUs • 2 x FHHL PCIe Gen5 x16 slots • 2 x OCP NIC 3.0 PCIe Gen5 x16 slots	• 2 x FHHL PCIe Gen5 x16 slots • 2 x OCP NIC 3.0 PCIe Gen5 x16 slots
		Note: The number of supported network adapters per node via PCIe expansion slots depends on the deployment mode (standalone or cluster mode).			
Network Adapter Options		• Dual-port 25GbE SFP28 OCP NIC 3.0 adapter • Dual-port 100GbE QSFP56 adapter			
		Note: Network adapters must be purchased from Infortrend.			
Max. FHFL Dual-Width GPUs		4	4	2	0
System Fans		6			
Rear IO Ports		• 1 x RJ45 dedicated BMC/IPMI LAN port • 1 x RJ45 1GbE management port • 1 x RJ45 serial port (COM) • 1 x Mini-DP port (supports VGA output) • 2 x USB 3.2 Gen1 ports (Type A)			
Dimensions (W x H x D) (Without Chassis Ears and Protrusions)		449 x 88 x 922 mm			449 x 88 x 852 mm
Package Dimensions (W x H x D)		596 x 262 x 1188 mm			596 x 262 x 1118 mm
Power Supply	Power Supplies	2 x 2700W 80 PLUS Titanium redundant hot-plug	2 x 3200W 80 PLUS Titanium redundant hot-plug	2 x 2700W 80 PLUS Titanium redundant hot-plug	2 x 1600W 80 PLUS Titanium redundant hot-plug
	AC Input	200-240VAC/16-13.5A, 50/60Hz	100-127VAC/18A, 200-240VAC/16A, 50/60Hz	200-240VAC/16-13.5A, 50/60Hz	100-127VAC/12A, 200-240VAC/10A, 50/60Hz
		Note: Please use 200-240VAC for all KS model			
Environmental (Without GPU)	Temperature	Operating: 10 to 35°C; non-operating: -40 to 60°C			
	Altitude	Operating: Sea level to 3,660m (12,000ft); non-operating: sea level to 12,192m (40,000ft)			
	Relative Humidity	5 to 95% non-condensing, operating and non-operating			
Safety Standards		• Electromagnetic compatibility: CE, BSMI, FCC • Safety: UL, BSMI, CB			

KSa 5000U Series

KSa 5008**U**



KSa 5016**U**



1. FHFL dual-width GPU slots
2. NVMe SSD slots
3. Network adapter slots
4. DDR5 DIMM slots
5. AMD EPYC™ 9004 / 9005 Series Processor

PHYSICAL SPECIFICATIONS

Product Series		KSa 5000U	
Model		KSa 5008**U**	KSa 5016**U**
Form Factor		2U 8-bay	2U 16-bay
Mounting Options		Rackmount	
Processor		Dual AMD EPYC™ 9004 / 9005 Series Processors	
Processor Core Available		16 to 160 cores	
Cache Memory		• 24 x DDR5 DIMM slots • Default 192GB, up to 1.5TB DDR5 RDIMM	
Supported Drives		2.5" U.2 NVMe SSD	
		Note: 1. A minimum of four drives is required. 2. For the latest Compatibility Guide, refer to our official website.	
PCIe Expansion Slots		• 4 x FHFL PCIe Gen5 x16 slots for GPUs • 3 x FHHL PCIe Gen5 x16 slots • 1 x OCP NIC 3.0 PCIe Gen5 x16 slot	• 2 x FHFL PCIe Gen5 x16 slots for GPUs • 3 x FHHL PCIe Gen5 x16 slots • 1 x OCP NIC 3.0 PCIe Gen5 x16 slot
Network Adapter Options		• Dual-port 25GbE SFP28 OCP NIC 3.0 adapter • Dual-port 100GbE QSFP56 adapter	
		Note: Network adapters must be purchased from Infortrend.	
Max. FHFL Dual-Width GPUs		4	2
System Fans		6	
Rear IO Ports		• 1 x RJ45 dedicated BMC/IPMI LAN port • 1 x RJ45 1GbE management port • 1 x RJ45 serial port (COM) • 1 x DB15 VGA port • 2 x USB 3.2 Gen1 ports (Type A)	
Dimensions (W x H x D) (Without Chassis Ears and Protrusions)		449 x 88 x 922 mm	
Package Dimensions (W x H x D)		596 x 262 x 1188 mm	
Power Supply	Power Supplies	• Models with 160-core CPU: 2 x 3200W 80 PLUS Titanium redundant hot-plug • Other models: 2 x 2700W 80 PLUS Titanium redundant hot-plug	
	AC Input	200-240VAC/16-13.5A, 50/60Hz	
Environmental (Without GPU)	Temperature	Operating: 10 to 35°C; non-operating: -40 to 60°C	
	Altitude	Operating: Sea level to 3,660m (12,000ft); non-operating: sea level to 12,192m (40,000ft)	
	Relative Humidity	5 to 95% non-condensing, operating and non-operating	
Safety Standards		• Electromagnetic compatibility: CE, BSMI, FCC	• Safety: UL, BSMI, CB

SOFTWARE SPECIFICATIONS

Management	EonOne, EonKube

SOFTWARE LICENSE

EonSight		Optional	EonSight enables AI-driven system monitoring and analytics, providing anomaly detection, performance insights, and capacity forecasting to help optimize cluster stability and operational efficiency.
Application Services	AI	Optional	AI License enables one-click deployment of built-in AI applications via the App Marketplace, accelerating AI environment setup and supporting advanced data analytics and machine learning workloads.
	HPC	Optional	HPC License enables deployment and management of built-in HPC applications through the App Marketplace, supporting efficient job scheduling, workload monitoring, and centralized management of HPC data and tasks.
	Manufacture	Optional	Manufacturing License enables one-click deployment of built-in manufacturing applications via the App Marketplace, providing a streamlined way to support and manage manufacturing operations and workflows.

WARRANTY AND SERVICE

Service and Support	Standard Service	3-year limited hardware warranty and 8 x 5 phone, web, and email support (batteries are covered under warranty for 2 years)
	Upgrade or Extension Options	Warranty extension: Standard service can be extended up to 5 years. The following services can be upgraded to 5 years. • Upgrade: Replacement part dispatch on the next business day • Advanced service: Phone, web, and email support + onsite diagnostics on the next business day • Premium service: Phone, web, and email support + onsite diagnostics within 4 hours
		Note: Options may vary by region. For more details, please contact our sales representatives.
	Technical Support	Get information on system installation and maintenance, download technical documents and software, or issue a support ticket
	Product Services	Register products, download firmware, apply for licensing services, create product repair tickets, or check product repair status

Asia Pacific (New Taipei, Taiwan)
Infortrend Technology, Inc.
Tel : +886-2-2226-0126
E-mail : sales.ap@infortrend.com

China (Beijing, China)
Infortrend Technology, Ltd.
Tel : +86-10-6310-6168
E-mail : sales.cn@infortrend.com

Japan (Tokyo, Japan)
Infortrend Japan, Inc.
Tel : +81-3-5730-6551
E-mail : sales.jp@infortrend.com

Americas (Sunnyvale, CA, USA)
Infortrend Corporation
Tel : +1-408-988-5088
E-mail : sales.us@infortrend.com

EMEA (Düsseldorf, Germany)
Infortrend Technology, Inc.
E-mail: sales.de@infortrend.com

Contact Sales

Visit Our Website

• Any information provided herein is without warranties of any kind of and is subject to change without prior notice.
• Copyright © 1999-2026 Infortrend Technology, Inc. Copyright to the documents and programs on the Site(s) is owned and/or performed by Infortrend Technology, Inc. All rights reserved.
• Infortrend, SANWatch, EonOne, EonStor and EonServ are registered trademarks or trademarks of Infortrend Technology, Inc. Other names prefixed with "IFT", "DS", "CS", "GS", "GSe", "GSe Pro", "GSx", and "KS" are trademarks or brand names of Infortrend Technology, Inc. All other names, brands, products or services are trademarks or registered trademarks of their respective owners.