AI Transportables

Buyers Guide

What is AI Transportables?

The past decade of technological advances in High Performance Computing (HPC), such as GPU's expanded role in compute, has led to a surge in adoption of machine-learning and artificial intelligence applications. This workflow has seen applications expand to the most extreme edge environments. While compute technologies have advanced significantly in datacenters, the same capabilities have seen slower adoption in edge/transportable environments.

AI Transportable systems are no-compromise compute platforms with the capabilities of the datacenter that can function in the same edge environments as embedded solutions. Through PCIe expansion, GPUs running the compute engine can have direct high-speed bus connections to equally fast NVMe storage, NIC add-in cards, and FPGA data capture. This PCIe design supports the latest AI system architectures for direct memory transactions in a disaggregated workflow.

AI Transportable systems meet stringent MIL-STD requirements for shock and vibration, broad operating temperature ranges, altitudes, humidity, and redundant power.

The AI Transportables industry represents the corner market between the latest advances in high-performance computing and edge (transportable) applications. According to MarketsandMarkets, the edge computing market is expected to grow at an average annual growth rate of 34 percent to $15.7 billion by 2025.
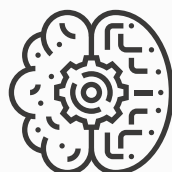
Things to know before buying AI Transportables systems:

- ☐ Applications
- ☐ Use Cases
- ☐ First-to-Market Advantages
- ☐ PCIe Bandwidth and Form Factors
- ☐ Certifications and Quality
- ☐ AI Transportable Systems in Comparison
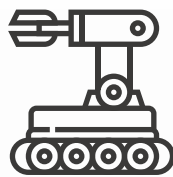
Transportability

AI Transportable Platform:
AI Ready
Compact
Modular

AI Ready

Realtime Capability

Transportable AI systems for faster inferencing

Transportable AI is necessary in industries with large incoming datasets including autonomous driving, military object identification, and oil and gas drilling operations. Two off-the-shelf options include datacenter rackmount systems or low-power embedded systems, both of which have severe limiting factors. Datacenter grade systems cannot meet the rigors these transportable industries when it comes to ruggedization. As datacenter systems are designed for stationary air-conditioned racks , they do not address the road-related vibration or cooling limitations encountered in the autonomous vehicle industry.

While embedded solutions integrate low-powered GPUs with low SWaP footprints, these products significantly underperform when compared to datacenter focused GPUs. Software developers for applications using embedded products must compromise on the sensor capture, data throughput, and the overall capabilities of their deployment. AI Transportable applications require both low SWaP footprints

and high performance. This is where AI Transportable systems come into play. The focus of AI Transportable systems is to bring the latest GPU technologies to the transportable edge. Integrating the latest GPUs into vehicles enables greater data capture and inference capability for AI applications.
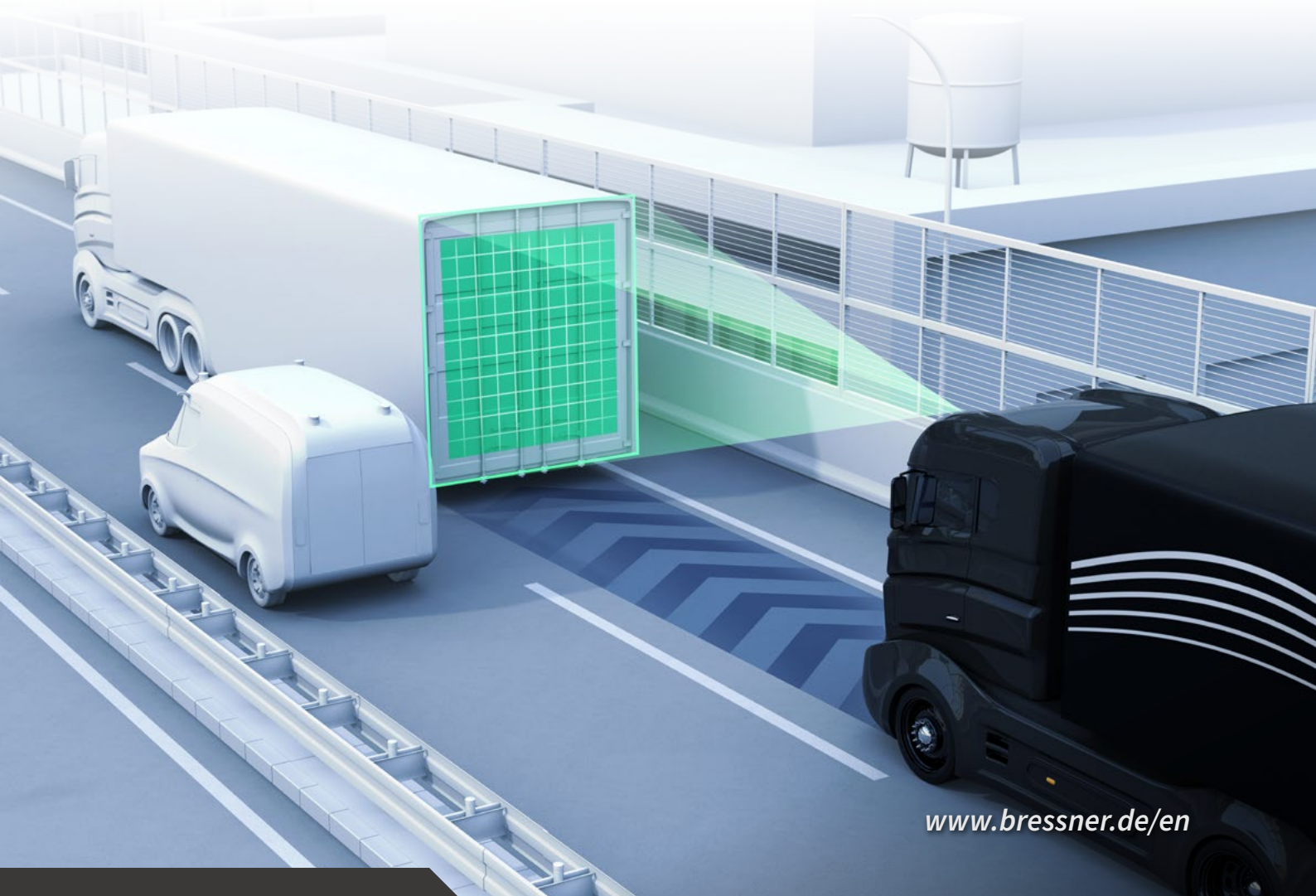
Market Share Edge Computing Global:

- ◆ **2019:** 3.5 billion USD
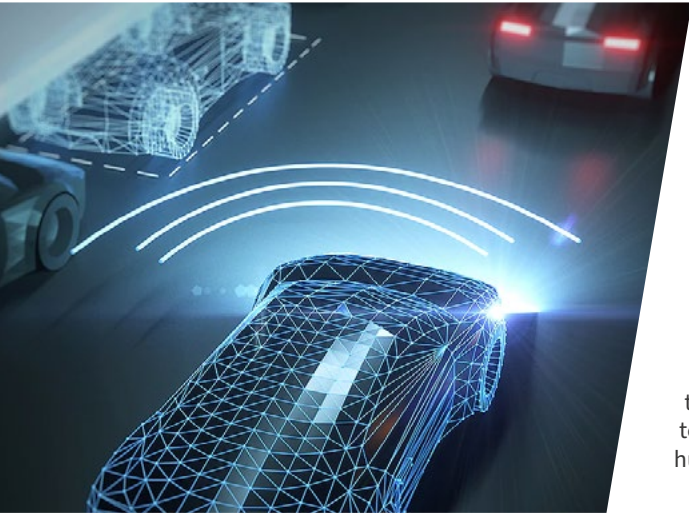- ◆ **2027:** 43.4 billion USD

Source: Grand View Research

Market Share AI Transportables:

- ◆ **2019:** 200 to 400 million USD
- ◆ **2027:** 3 to 5 billion USD

Source: One Stop Systems

Applications

Autonomous Vehicles

Artificial Intelligence is a core component of deploying truly autonomous vehicles in commercial, industrial and defense applications. These AI Transportable applications require no-compromise performance deployed in harsh operating environments supporting real time decision making. Autonomous vehicles are outfitted with specialized high performance edge computing equipment with high bandwidth data ingest components tied to a myriad of video, radar and LIDAR sensors, high capacity and low latency storage subsystems and high-performance compute engines that can perform the AI machine learning and inference tasks needed to enable the vehicle to see, hear, think and make decisions just like human operators.

Aerospace

Modern aircraft generate masses of data during flight and also during their stay at the airport. The incoming information must be seamlessly collected, filtered, and forwarded through various communication paths in order to be processed and archived at the end. Telemetry, entertainment and cabin systems require specially certified hardware that can withstand not only the pressure drop at high altitudes, but also extreme temperature fluctuations and mechanical stresses. OSS brings two decades of experience in specification, development, manufacturing and repair of aerospace components and systems. Development practices patterned after DO-254 and DO-178-C are combined with an AS9100-D Quality System. Flight safety and certification according to DO-160G and related Boeing and Airbus requirements are offered according to customer need.

Industrial Automation

Artificial Intelligence is being deployed to address a wide range of applications across numerous industrial markets including Oil and Gas, Mining, Construction and Agriculture, and Life Sciences and Healthcare. In each of the industries data is growing exponentially and it is being used to develop sophisticated algorithms to provide predictive intelligence to enable real time actions in the field. The latency and security issues involved in transferring data to remote data centers is unacceptable. These AI Transportable applications also have the unique requirement to operate in often harsh or extreme environmental conditions or have severe constraints in terms of size, weight, and power.

Media and Entertainment

Artificial Intelligence is augmenting high-end video production workflows in Media and Entertainment applications driving the need for no compromise high-performance compute and storage systems designed specifically for rugged operation on-set and in the field. Rugged transportable AI systems are required at venues to process very high-resolution video in real time and manage complex event video special effects and light shows. In addition, AI transportable systems are required on set to provide extremely large amounts of memory and storage bandwidth to capture the high-speed data flows generated with 8K raw camera output. Ruggedized removeable storage is needed to transfer the massive data files created on set for transport to post-production systems without the unacceptable latency and cost of network connections.

Marine and Sea Transportation

Sea-based AI transportable applications include advance threat detection systems, autonomous or semi-autonomous surface ships and undersea vessels, and automated maintenance alert systems. AI capabilities are embedded in many shipboard defensive and offensive mission systems designed for rapid awareness and reaction to threats. Unmanned surface ships have been demonstrated to travel thousands of miles without any crew paving the way to ocean patrols for months at much reduced costs. Unmanned Underwater Vehicles (UUVs) using AI technology are deployed in anti-mine missions. AI technology is also used to automate maintenance, identifying weak systems before they break down at sea.

Defense and Security

Artificial Intelligence is rapidly becoming a core element of defense systems across all mission environments including on land, in the air and on the sea. These applications require the highest possible performance deployable in the harshest conditions. AI will be critical in coping with the huge data flows being acquired from sophisticated sensor networks and converting this to actionable intelligence in real time. The latency involved in relying on remote datacenters is unacceptable in these situations. These systems are located in harsh environments but require the highest levels of reliability and security along with the latest in CPU, GPU, data ingest and storage technology. The military theater of tomorrow is one that will demand "AI transportability" at every node.

Use Cases

Transportable AI hardware enables organizations to deploy off-the-shelf systems with AI capabilities into the field to solve a range of critical problems in real time. Capabilities such as machine learning, deep learning, and deep neural networks are already being used in very different market segments, both commercial and military:

◆ Truck manufacturers are deploying an AI-based autonomous driving system that enables their trucks to respond in real time to highway conditions, including road construction, that are missed by GPS mapping systems

◆ Entertainment/media applications are deploying robust, transportable AI systems at venues to process high-definition video and manage complex video and light shows for events

◆ Ride-sharing companies are deploying AI and high-performance hardware in prototype vehicle fleets

◆ Companies are developing autonomous civilian vehicles, aside from cars. These include long-haul commercial trucks, heavy construction equipment, mining and agricultural machinery, buses, subways, as well as freight and passenger trains

◆ Military and civilian drone companies are enabling cooperative behavior between aerial and land drones

◆ Offshore oil and gas facilities use AI to monitor complex surface and subsea operations and analyze sensor data to detect suboptimal operations and impending problems, including equipment failures, before they occur

◆ AI-based maritime monitoring and analysis systems are used onboard ships to automate the detection of faulty systems for maintenance purposes

◆ Aerospace prime contractors are developing an AI-based threat detection system aboard U.S. Navy aircraft

◆ AI cybersecurity applications monitor real-time access to industrial assets at manufacturers and utilities, track authorized access, and detect patterns indicative of cyberattacks

◆ Civilian connected aircraft are using on-board AI and SATCOM or 5G wireless links to collect data on aircraft operational performance and process it on the spot

◆ Medical imaging and diagnostic devices utilize AI to process and analyze scans for faster, more accurate and detailed diagnoses

◆ The military is using AI in mobile command centers at the edge to quickly process a flood of tactical information into a comprehensive and intelligible picture of the battlefield

◆ Automated targeting systems are using advanced sensors, machine-learning algorithms and touchscreen displays to enable army tank crews to detect and respond to incoming threats faster than ever before

First-to-Market Advantages

There are three key considerations to assess when choosing the right technology supplier to achieve a first-to-market advantage in your AI Transportable industry: time to market, constraints and trade-offs, and execution.

Time to Market

You can speed up your time to market by using a modular, scalable hardware platform. A commercial off-the-shelf platform that is modular and scalable is critical to speeding time to market for high performance, real-time, AI-enabled applications in the field.

Benefits at a glance:

- ◆ Modular and scalable products to serve the AI transportable space which vary based on the environmental parameters and the amount of compute performance needed

- ◆ Datacenter-oriented GPU integration and packaging in thermally and structurally optimized enclosures via design simulation and qualification testing

- ◆ Project lifecycle management: defining requirements, milestone project planning, qualification testing, material obsolesecence mitigations, and establishing close, ongoing customer support

BECOME A MARKET LEADER IN YOUR AI TRANSPORTABLE INDUSTRY

Constraints and Trade-Offs

You can optimize your tradeoffs by leveraging vendor expertise and experience. Constraints for AI Transportables fall into three areas: computing power, AI capacity and survivability.
Companies should assess potential technology suppliers for their demonstrated expertise in breaking through these constraints to enable reliable, powerful AI processing in rugged, edge solutions.

Execution

You can establish a multi-level technical partnership to execute projects on-time and on-budget.
Companies should evaluate systems suppliers that adopt a customer success business model. Such a model, at the outset, establishes a mutual relationship that has a shared focus: the specific business objectives and deliverables for a given program.

PCIe Bandwidth and Form Factors

| PCIe	Bandwidth per lane	PCIe x1	PCIe x4	PCIe x8	PCIe x16	Encoding
1.0 / 1.1	2.5 GT/s	0.25 GByte/s	1.0 GByte/s	2.0 GByte/s	4.0 GByte/s	8b10b
2.0 / 2.1	5 GT/s	0.50 GByte/s	2.0 GByte/s	4.0 GByte/s	8.0 GByte/s	8b10b
3.0 / 3.1	8 GT/s	0.97 GByte/s	3.9 GByte/s	7.8 GByte/s	15.8 GByte/s	128b/130b
4.0	16 GT/s	1.90 GByte/s	7.8 GByte/s	15.8 GByte/s	31.5 GByte/s	128b/130b
5.0	32 GT/s	3.90 GByte/s	15.8 GByte/s	31.5 GByte/s	63.0 GByte/s	128b/130b
6.0	64 GT/s	7.50 GByte/s	30.3 GByte/s	60.5 GByte/s	121.0 GByte/s	PAM-4

The importance of PCIe for AI Transportables

The latest version of PCI Express (PCIe) is Gen 5, (PCIe 5.0) which was released in July 2017. PCIe 5.0 supports a data rate of up to 256 Gbps per lane, transmitting twice as fast as PCIe 4.0. PCIe 5.0 is backward compatible with all previous versions of PCIe and can be deployed into existing systems.

PCIe 5.0 enables AI Transportable devices to use twice the bandwidth of the previous generation, which is critical for this industry where sensor technology continues to increase the size of data captured for inference emodels. In addition, PCIe 5.0 supports lower latency, which is important for real-time applications where every millisecond counts.

PCIe form factors

U.2

100 mm

69.9 mm

M.2

80 mm

42 mm

22 mm

22 mm

CEM add-in card

67.3 mm

167.6 mm
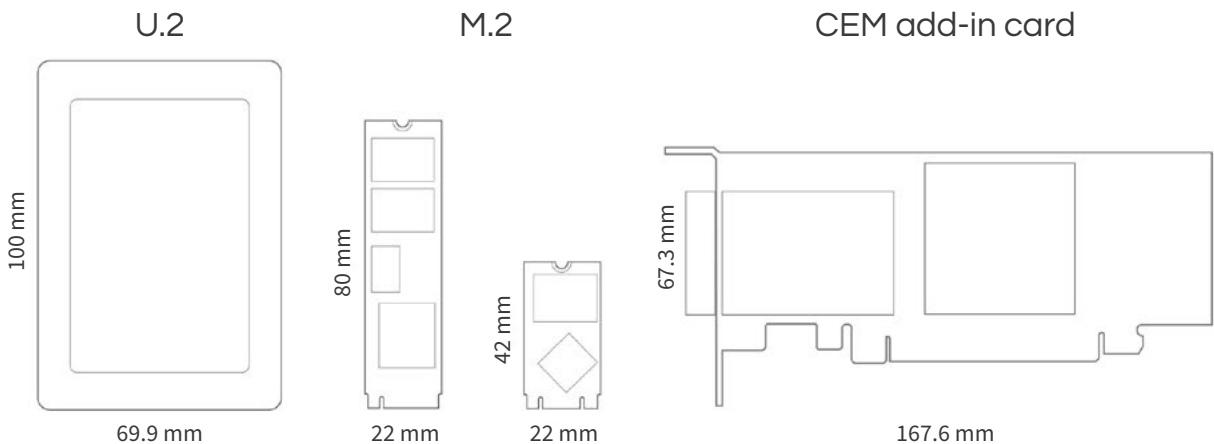
Certifications and Quality

What do the MIL-STD standards mean?

Standard	Meaning and scope
MIL-STD-810	A set of standardized test methods for the U.S. military, defining, among others, the compatibility of equipment in dealing with high temperature fluctuations, atmospheric pressure, humidity, vibration, or solar radiation.
MIL-STD-461	Specification of electromagnetic compatibility requirements for military products and solutions.
MIL-STD-464	Establish environmental electromagnetic interface requirements and test criteria for air, sea, space, and ground systems, including associated ordnance.
MIL-STD-704	Ensure compatibility between the aircraft electrical system, external power supply, and airborne equipment.

AS9100 quality management

One Stop Systems has both ISO 9001:2015 and AS9100D certifications. This scope of certification includes: design, manufacture and supply of industrial computers for the AI Transportable industry.

AS9100 is a quality management system developed to fill the gaps that aerospace companies have identified in the ISO 9001 standard, in terms of how they conduct business for their customers. It was originally approved in 1999 and has since undergone four revisions, culminating in our current version, AS9100D.

A quality system is a formal system that documents procedures and processes to meet customer requirements. When you have a quality system, your customers know that you have a complete system in place to produce the product or service you provide in a high-quality, repeatable manner while meeting all regulatory requirements. Through the risk and opportunity assessment conducted as part of the AS9100 process, there are always opportunities to improve your business or enhance your customers' experience. When companies like OSS have AS9100 in place, the world knows

that we are looking at those risks and opportunities and proactively address them.

The idea of continuous improvement is difficult to integrate into a corporate culture, but AS9100 helps accelerate the process and engage team members around these issues. For OSS, the journey began many years ago with the ISO-9001 standard and the company's commitment to this quality management system. In 2020, OSS began the process of certifying to AS9100 and received accreditation in early 2022. In a short period of time, the company has made significant progress in becoming compliant with the standards throughout the organization. The OSS company culture has adopted continuous improvement as a way of life.

AI Transportable Systems in Comparison

Rugged GPU Servers

Model	3U Short Depth Server	Rigel Edge Supercomputer	GAS-R
CPU	Intel® Xeon® Ice Lake or AMD EPYC™ 7003 Dual or Single	AMD EPYC™ 7003 processor Single socket SP3 (LGA4094)	AMD Epyc™ 7002 up to 64 cores or Dual Intel® scalable processors
GPU	Full-Height, Full-Length	4x NVIDIA® A100 SXM GPUs 320 GB GPU storage 3. Generation NVLink 2.4 TB/s aggregated total bandwidth	8x NVIDIA® A100 SXM4 GPUs (NVLink 3.0)
RAM	DDR4 RAM up to max. 4TB	8x ECC DDR4 RDIMM/LRDIMM slots (support modules up to 128 GB)	Up to 4TB DDR4 ECC RAM
Storage Capacity	12GB SAS-3 or 6GB SATA III 32GB NVMe x4	PCIe 4.0 x4 NVMe M.2 slot	Up to 200TB U.2 PCIe 4.0 NVMe flash
PCIe Expansion	4x PCIe 4.0 x16 FH 1x PCIe 4.0 x16 FH, Half-Length	4x PCIe 4.0 x16 HHHL slots	4x PCIe 3.0 x16 HH/FL Double-Width 2x PCIe 3.0 x16 HH/HL Single-Width 1x PCIe 3.0 x4 HH/HL with physical x8 connector 32x PCIe 3.0 x8 HH/HL (removable)
I/O	2x USB 3.2 1x USB 3.1 Gen 2 (Type C)	4x RJ-45 10/100/1000 (rear) 3x RJ-45 Unified Baseboard Management (rear) 2x USB-C 1x VGA	2x RJ-45 (10-Gigabit Ethernet) 4x Ethernet & Infiniband up to 200GB/s 4x FibreChannel up to 32GB 5x USB 3.0, 2x rear, 2x front, 1x internal 4x USB 2.0 (2x rear, 2x internal)
Chassis	Lightweight aluminium frame (3U) Black anodized outer case	Lightweight aluminium frame (4U) Black anodized outer case	Black anodized aluminium chassis (8U)
Cooling	92mm fan, 40mm fan	2x 80 x 80mm high CFM PWM controlled GPU fans 4x 80 x 28mm PWM controlled server fans	80mm fan, 92mm fan, 80dB(A) at 1-meter distance
Power Supply	Single/Dual AC 1600W, 110-240VAC Single/Dual DC 1600W, 48VDC Single/Dual AC 2600W, 110-240VAC	90-264VAC (47-63Hz) 110VAC (400-800Hz) 3Ø 180-300 VDC 48 VDC	100-250VAC (47-63Hz or 400Hz) 200-370VDC 2+1, hot swap power supply up to 1600W
Temperature	Operation: 0° ~ 35°C Storage: -40° ~ 71°C	Operation: 0° ~ 35°C Storage: -40° ~ 71°C	Operation: 0° ~ 35°C Storage: -40° ~ 71°C
Humidity	10% ~ 90%	0% ~ 95%	5% ~ 95%
Weight	20.4kg (45.0 lbs)	22.7kg (50.0 lbs)	60.3kg (132.9 lbs)
Dimensions	482.6 x 177.8 x 469.9mm	177.8 x 218.4 x 678.1mm	368.3 x 431.8 x 622.3mm
Regulations	FCC, CE, CISPR 22, MIL-STD-810G	FCC, ICES-003, CE, CISPR 22, RoHS 6 of 6, WEEE, MIL-STD-461E	MIL-STD-810G, MIL-STD-461E, MIL-STD464A, MIL-STD-704E, IEC 61000-4-2

GPU Accelerators and Storage Systems

EB4400	Centauri	FSAn-4	
-	-	Dual Intel® Xeon® Skylake, Cascade Lake, Cascade Lake-X, LGA 3647	**CPU**
-	-	Half-Height, Full-Length, GPU connectors on rear side	**GPU**
-	-	Up to 4TB DDR4 ECC RAM	**RAM**
-	Up to 8x U.2/U.3 PCIe 3.0 or 4.0 NVMe SSDs with 128TB per 3U system (15.36TB NVMe drives)	Supports 32x PCIe 3.0 x8 HH/HL PCIe NVMe flash, up to 400TB usable	**Storage Capacity**
Backplane Dual OSS-538: 1x Single-Width PCIe 4.0 x16 FHFL 4x Dual-Width PCIe 4.0 x16 FHFL Backplane Dual OSS-521: 1x SIngle-Width PCIe 4.0 x16 FHFL 6x Single-Width PCIe 4.0 FHFL 1x Dual-Width PCIe 4.0 x16 FHFL	1x PCIe 4.0 x16 SFF-8644 host up-link for connecting several Centauri systems	4x PCIe 3.0 x16 HH/FL Double-Width, 2x PCIe 3.0 x16 HH/HL Single-Width, 1x PCIe 3.0 x4 HH/HL with physical x8 connector 32x PCIe 3.0 x8 HH/HL (removable)	**PCIe Expansion**
1x PCIe 4.0 x16 host-to-target uplink kits (32GBit/s) 2x PCIe 4.0 x16 host-to-target uplink kits (64GBit/s) SmartNIC host	2x RJ-45 Ethernet ports (for web server GUI access)	2x RJ-45 (Intel® X550) 10-GbE slots 5x USB 3.0, 2x rear, 2x front, 1x internal	**I/O**
Rackmount steel chassis (4U)	Aluminium chassis (3U)	Rackmount aluminium chassis (4U)	**Chassis**
3x 92mm (180CFM) fan standard PWM controlled Optional IPMI system monitoring Optional Quadrafoam 45 PPI removable fan filters	BMC and fan control	4x 70 x 38mm PWM hot swap fan	**Cooling**
Single / Dual AC 2600W power supply Single / Dual AC 1600W power supply Single / Dual DC 1600W power supply	Dual 1600W redundant power supplies: 100-240VAC full range 48VDC 50/60Hz	Power input rear: 100-250VAC (47-63 or 400Hz) or 200-370VDC 2+1, hot swap power supply up to 1600W	**Power Supply**
Operation: 0° ~ 35°C; short-term: -5° ~ 40°C Storage: -40° ~ 71°C	Operation: 10° ~ 50°C Storage: -40° ~ 71°C	Operation: 0° ~ 35°C Storage: -40° ~ 71°C	**Temperature**
5% ~ 95%	5% ~ 95%	5% ~ 95%	**Humidity**
11.3kg (24.9 lbs) - w/o GPUs	5.9kg (13.0 lbs)	31.8kg (70.1 lbs) to 38.6kg (85.1 lbs)	**Weight**
271.8 x 177.8 x 470mm	215.9 x 133.4 x 508mm	431.8 x 177.8 x 609.6mm	**Dimensions**
FCC Class A, CE, UL, cUL, RoHS3	FCC Class A, ICES-003, UL/IEC 60950-1, CE, CISPR 22, WEEE & RoHS 3	MIL-STD-810G, MIL-STD-461E, MIL-STD-464A, MIL-STD-704E	**Regulations**

BRESSNER
A ONE STOP SYSTEMS COMPANY

Contact

EMEA

BRESSNER Technology GmbH
Industriestrasse 51
82194 Groebenzell
Germany

Phone: +49 8142 47284-70
E-Mail: vertrieb@bressner.de
Web: www.bressner.de/en

USA

One Stop Systems Inc.
2235 Enterprise Street #110
Escondido, CA 92029, USA

Phone: +1 (877) 438-2724
Fax: +1 (760) 745-9824
E-Mail: sales@onestopsystems.com