

# Gryf

## The First Ever Suitcase-sized AI Supercomputer

### HIGHLIGHTS

#### EASILY TRANSPORTABLE

Compact, lightweight, airline cabin-friendly carry-on form factor

#### ADAPTABLE & SCALABLE

Scale to five interconnected units, each with six sleds for GPU, compute, storage, and networking

#### DATACENTER PERFORMANCE

AI fabric technology powers high-density GPU compute and extensive storage capabilities

#### COMPREHENSIVE ECOSYSTEM

Leading storage vendors, AI/HPC software providers, and accelerator manufacturers

Gryf is a highly flexible and reconfigurable mobile data center designed to bring powerful AI and machine learning capabilities directly to the edge.

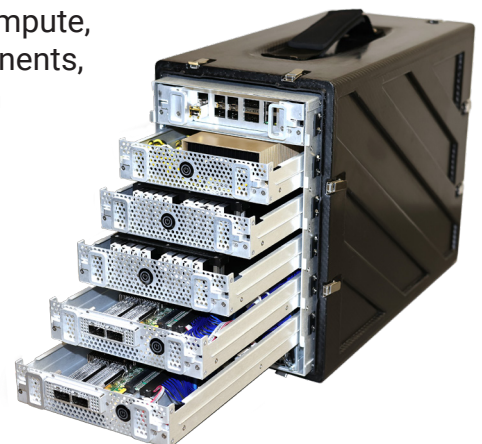
Co-designed by GigaIO and SourceCode for portability and scalability, Gryf allows users to quickly transform the vast amounts of sensor data collected at the edge into actionable solutions. It sets a new standard for on-demand configurability in the field and is unique in its ability to scale to the performance of a field supercomputer, all in the airline cabin-friendly form factor of carry-on luggage.

This scaling is made possible by GigaIO's groundbreaking AI fabric, which unifies edge-to-core infrastructures to dynamically deploy any mission application for actionable real-time intelligence – anywhere.

### Datacenter Class Performance

Gryf combines high-performance GPUs with substantial storage capacity to execute complex AI tasks at the edge. Gryf can scale up to five interconnected units, each housing six customizable sleds for GPU, compute, storage, and networking components, enabling massive capabilities in a single server and dynamic reconfiguration at the edge.

The following table shows the performance characteristics of Gryf's different GPU offerings.



## Gryf Accelerator Specifications

	NVIDIA L40S	NVIDIA H100 NVL	NVIDIA H200 NVL
<b>FP64 Performance</b>	NA	30 TFLOPS	30 TFLOPS
<b>FP32 Performance</b>	91.6 TFLOPS	60 TFLOPS	60 TFLOPS
<b>FP8 Performance *</b>	1,466 TFLOPS	3,341 TFLOPS	3,341 TFLOPS
<b>Memory Size</b>	48GB	94GB	141GB
<b>Memory Bandwidth</b>	864GB/s	3.9TB/s	4.8TB/s

\* tensor cores with sparsity

### Easily Transportable

Gryf is designed for maximum mobility and ease of deployment. The compact platform features a detachable top with a folding handle and a detachable bottom with wheels, allowing the entire unit to be easily transported and rolled into the field. These features make Gryf ideal for rapid deployment in remote or challenging environments.

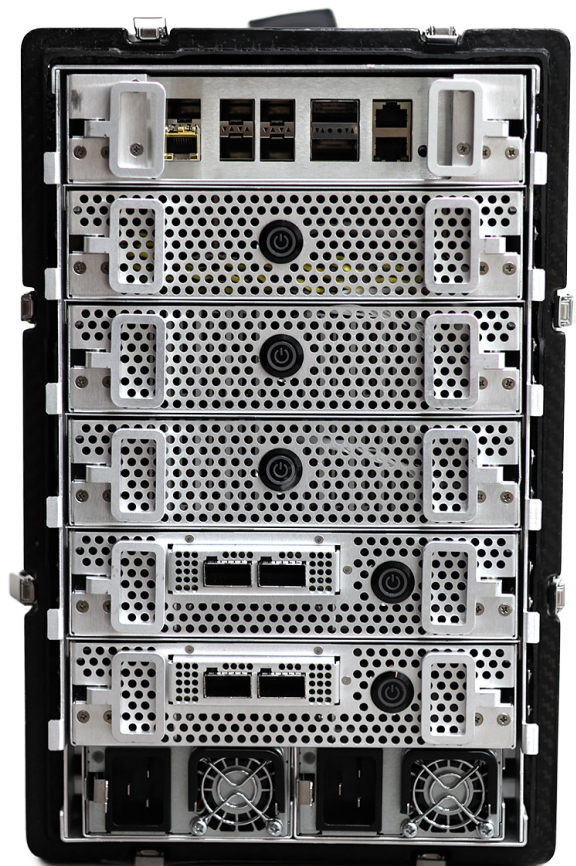
### Adaptable & Scalable

The modular architecture of Gryf allows it to be tailored to specific mission requirements. The platform features multiple slots that can be populated with a variety of compute, storage, accelerator, and network sleds. Users can mix and match these components to create the optimal configuration for their needs, whether that's focused on AI/ML, high-performance computing, or data-intensive storage.

### Edge Intelligence

By bringing powerful AI and HPC capabilities directly to the edge of the network, Gryf enables unprecedented levels of real-time intelligence and decision-making in field deployments. This is enabled by GigaIO's AI fabric, which unifies the edge-to-core infrastructure and allows for the dynamic deployment of mission-critical applications.

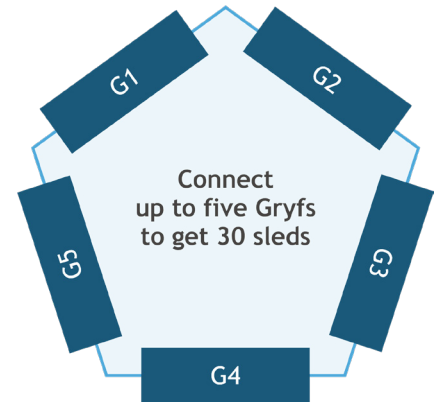
Gryf is designed to provide a highly versatile and capable edge computing platform that can be rapidly deployed and reconfigured to address a wide range of mission requirements, from combat operations to disaster response to scientific research in remote locations. Its combination of performance, agility, and modularity makes it a valuable asset for organizations that need to push the boundaries of what is possible at the tactical edge.



## Comprehensive Ecosystem

Gryf is supported by a growing ecosystem, including:

- Storage vendors
- AI/HPC software providers
- Accelerator manufacturers
- Key partners: VAST Data, WEKA, Hammerspace, Peak:AIO, Edge Runner, Catapult
- Open Gryf sled architecture

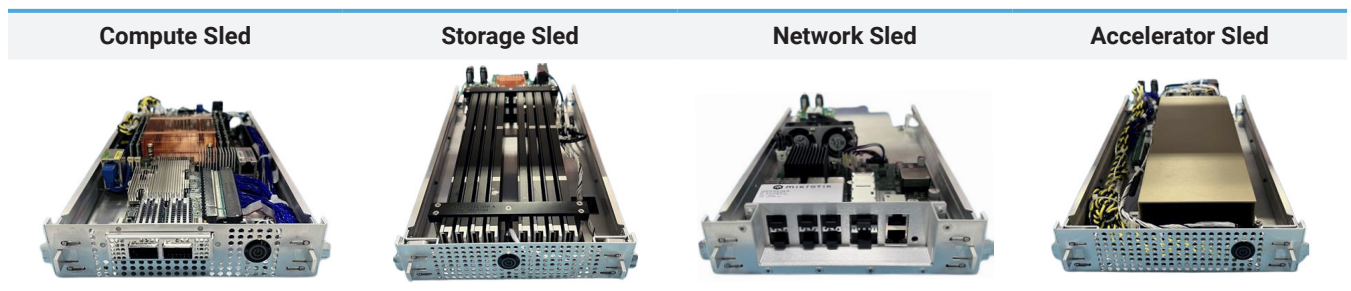


The platform’s modular design enables customization for specific applications through its open sled architecture, supporting diverse computing needs from AI/ML workloads to high-performance computing applications.

## Sample Configurations

Gryf is easily configured and reconfigured to meet your deployment needs by populating the platform with sleds of your choice. Initial fixed configuration options at launch are as follows:

	AI/ML	Compute	Storage
<b>Compute Sled</b>	1	2	1
<b>Accelerator Sled</b>	2	0	0
<b>Storage Sled</b>	2	3	5
<b>Network Sled</b>	1	1	0
<b>Capabilities</b>			
<b>System Memory</b>	512GB	512GB per sled (1TB total)	512GB
<b>Storage</b>	492TB per sled (0.9PB total)	492TB per sled (1.5PB total)	495TB per sled (2.5PB total)
<b>GPU Memory</b>	96GB (L40S), 188GB (H100 NVL), 282GB (H200 NVL)	-	-



## Gryf Specifications

Compute Sled	
<b>CPU</b>	(1) AMD EPYC™ 7003 family, 16-, 32-, or 64-core options
<b>System Memory</b>	(4) DDR4 DIMMs: 4x32GB (128GB), 4x64GB (256GB), 4x128GB (512GB) options
<b>OS Storage</b>	(1) 512GB NVMe-M.2 SSD
<b>OS Support</b>	Linux Rocky 8/9 or Ubuntu 20/24, planned support for Windows Server 2019 & 2022
<b>Networking</b>	(2) QSFP56-100GbE
<b>BMC / IPMI</b>	Via platform RJ45
Accelerator Sled	
<b>Accelerator Slot</b>	Single/Double-wide PCIe-FHFL form factor, up to 350W
<b>Accelerator</b>	(1) Nvidia L40S 48GB, or H100 NVL 94GB, or H200 NVL 141GB
Storage Sled	
<b>Storage</b>	(8) NVMe-E1.L SSD, options for 8x15.36TB (123TB total), 8x30.72TB (246TB total), 8x61.44TB (492TB total)
Network Sled	
<b>Ports</b>	(2) QSFP28-100GbE (6) SFP28-25GbE
Platform	
<b>Sled Slots</b>	(6) for Compute Sleds, Accelerator Sleds, Storage Sleds, Network Sleds in any location
<b>GigaIO AI Fabric</b>	Internal: PCIe 128Gb/s board-to-board Expansion & data offload: (8) fabric mini-SAS-HD-32G (256Gb/s total)
<b>Rear Panel Ports</b>	(5) SFP+ 10GbE for remote management (2) USB 2.0 (1) HDMI display
<b>Software</b>	GigaIO AI Fabric Manager
<b>Power</b>	Dual AC/DC 2,500W 1+1 power supplies IEC-320-C20 power inlet, 100-240 VAC @ 50 to 60Hz
<b>Fans</b>	(6) 60mm fans dynamically optimized for system workloads
<b>Fan Filters</b>	Removable 45 PPI filters
<b>Dimensions</b>	With wheels: 9.00" x 14.00" x 24.50" (228.6mm x 355.6mm x 622.3mm) Without wheels: 9.00" x 14.00" x 22.00" (228.6mm x 355.6mm x 558.8mm)
<b>Weight</b>	Less than 55lbs (24.95 kgs) max
<b>Enclosure</b>	Ruggedized carbon fiber, detachable top with folding handle, detachable bottom with wheels
<b>Environment</b>	Operating temperature: 10°C to 32°C (50°F to 90°F)
<b>Serviceability</b>	Field replaceable units (FRUs): sleds, NVMe SSDs, power supplies, fan tray with fan filtration, case top and bottom covers
<b>Compliance (Pending)</b>	FCC Class A, CE, Trade Agreements Act (TAA) / Made in the USA, IP55 in transport

