Data Sheet

Gryf

The First Ever Suitcase-sized AI Supercomputer

HIGHLIGHTS

EASILY TRANSPORTABLE

Compact, 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 datacenter 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
FP64 Performance	NA	30 TFLOPS
FP32 Performance	91.6 TFLOPS	60 TFLOPS
FP8 Performance *	1,466 TFLOPS	3,341 TFLOPS
Memory Size	48GB	94GB
Memory Bandwidth	864GB/s	3.9TB/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

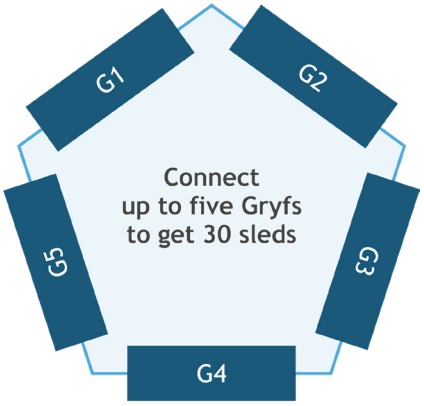
By bringing powerful AI and HPC capabilities directly to the edge of the network, Gryf enables unprecedented levels of real-time intelligence and decision-making in field deployments. This is enabled by GigaIO's AI fabric, which unifies the edge-to-core infrastructure and allows for the dynamic deployment 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 uses, from media & entertainment and sports analytics to military & defense and energy / oil & gas. 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, Ipsotek
- Open Gryf sled architecture

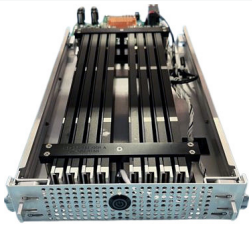


The platform's modular design enables customization for specific applications through its open sled architecture, supporting diverse computing needs from AI/ML workloads to high-performance computing applications.

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
Compute Sled	2	2	1
Accelerator Sled	1	0	0
Storage Sled	2	3	5
Network Sled	1	1	0
Capabilities			
System Memory	1TB	1TB	512GB
Storage	984TB	1.5PB	2.5PB
GPU Memory	48GB (L40S), 94GB (H100 NVL)	-	-

Compute Sled	Storage Sled	Network Sled	Accelerator Sled
			

Gryf Specifications

Compute Sled	
CPU	1x AMD EPYC™ 7003 family, 7713P 64-core or 7313P 16-core
System Memory	4x 128GB (512GB total) or 4x 64GB (256GB total)
OS Storage	1x 960GB NVMe-M2 SSD
OS Support	Linux Rocky 8/9 or Ubuntu 20/24, planned support for Windows Server 2019 & 2022
Networking	2x QSFP56-100GbE
BMC / IPMI	Via platform rear panel ports
Accelerator Sled	
Accelerator Slot	Single/Double-wide PCIe-FHFL form factor, up to 350W
Accelerator	1x NVIDIA L40S 48GB, or 1x H100 NVL 94GB
Storage Sled	
Storage	8x 61.44TB NVMe-E1.L SSD (492TB total) or 8x 30.72TB NVMe-E1.L SSD (246TB total)
Network Sled	
Ports	2x QSFP28-100GbE 6x SFP28-25GbE
Platform	
Sled Slots	6x for Compute Sleds, Accelerator Sleds, Storage Sleds, Network Sleds
GigaIO AI Fabric	Internal: PCIe 128Gb/s board-to-board Expansion & data offload: 8x fabric mini-SAS-HD-32G (256Gb/s total)
Rear Panel Ports	4x SFP+ 10GbE and 1x SFP+ 1GbE for remote management and provisioning 2x USB 2.0 1x HDMI display
Software	GigaIO AI Fabric Manager
Power	Dual 2,500W 1+1 power supplies IEC-320-C20 power inlet, 100-240 VAC @ 50 to 60Hz
Fans	6x 60mm fans dynamically optimized for system workloads
Fan Filters	Removable 45 PPI filters
Dimensions	With top & bottom covers: 9.00" x 14.00" x 24.75" (228.6mm x 355.6mm x 628.7mm) Without top & bottom covers: 9.00” x 14.00” x 21.25” (228.6mm x 355.6mm x 539.8mm)
Weight	Configuration dependent
Enclosure	Ruggedized carbon fiber, detachable top with folding handle, detachable bottom with wheels
Environment	Operating temperature: 10°C to 32°C (50°F to 90°F)
Serviceability	Field replaceable units (FRUs): sleds, NVMe SSDs, power supplies, fan tray with fan filtration, case top and bottom covers
Compliance (Pending)	FCC Class A, CE, Trade Agreements Act (TAA) / Made in the USA, IP55 in transport