



AIME R410 - 4 GPU Rack Server

Your Deep Learning server, configurable with 4 high end deep learning GPUs which give you the fastest deep learning power available: Upto 500 Trillion Tensor FLOPS of AI performance and 64 GB high speed GPU memory.

Built to perform 24/7 at your inhouse data center or co-location with an server graded EPYC powered CPU.

— AIME R410 - Machine Learning Server

For Machine Learning a new kind of server is needed, the multi GPU AIME R410 takes on the task for delivering maximum Deep Learning training and interference performance.

With its liquid cooled CPU and high air flow cooling design it keeps operating at their highest performance levels even under full load in 24/7 scenarios.

6 high performance fans are working aligned to deliver a high air flow through the system, this setup keeps the system cooler, more performant and durable than an collection of many small fans for each component.

— EPYC CPU-Performance

The high-end AMD EPYC CPU designed for servers delivers up to 32 cores with a total of 64 threads per CPU with an unbeaten price performance ratio.

The available 128 PCI 3.0 lanes of the AMD EPYC CPU allow highest interconnect and data transfer rates between the CPU and the GPUs and ensures that all GPUs are connected with full x16 PCI 3.0 bandwidth.

A large amount of available CPU cores can improve the performance dramatically in case the CPU is used for preprocessing and delivering of data to optimal feed the GPUs with workloads.

— Well Balanced Components

All of our components have been selected for their energy efficiency, durability, compatibility and high performance. They are perfectly balanced, so there are no performance bottlenecks. We optimize our hardware in terms of cost per performance, without compromising endurance and reliability.



Definable Quad GPU Configuration

Choose the need configuration among the most powerfull NVIDIA GPUs for Deep Learning:

4 X NVIDIA RTX 2080TI

Each NVIDIA RTX 2080TI trains AI models with 544 NVIDIA Turing mixed-precision Tensor Cores delivering 107 Tensor TFLOPS of AI performance and 11 GB of ultra-fast GDDR6 memory.

With the AIME R410 you can combine the power of 4 of those adding up to more then 400 Trillion Tensor FLOPS of AI performance.

4 X Titan RTX

Powered by the award-winning Turing™ architecture, the Titan RTX is bringing 130 Tensor TFLOPs of performance, 576 tensor cores, and 24 GB of ultra-fast GDDR6 memory.

With the AIME R410 you can combine the power of 4 of those adding up to more then 500 Trillion Tensor FLOPS of AI performance.

4 X Tesla V100

With 640 Tensor Cores, Tesla V100 is the world's first GPU to break the 100 teraFLOPS (TFLOPS) barrier of deep learning performance including 16 GB of highest bandwidth HBM2 memory.

Tesla V100 is engineered to provide maximum performance in existing hyperscale server racks. With AI at its core, a single Tesla V100 GPU delivers 47X higher inference performance than a CPU server.

All NVIDIA GPUs are supported by NVIDIA's CUDA-X AI SDK, including cuDNN, TensorRT which power nearly all popular deep learning frameworks.

Up to 4TB High-Speed SSD Storage

Deep Learning is most often linked to high amount of data to be processed and stored. A high throughput and fast access times to the data are essential for fast turn around times.

The AIME R410 can be configured with two NVMe SSDs, which are connected by PCI lanes directly to CPU and main memory. We offer following 3 class types of SSD to be configured:

- **QLC-Typ:** high read rates, average write speed - best suitable for reading of static data libraries or archives
- **TLC-Typ:** highest read and high write speed - best suitable for fast read/write file access
- **MLC-Typ:** highest read and write speed - best suitable for high performance databases, data streaming and virtualization



High Connectivity and Management Interface

With the two available 10 Gbit/s LAN ports fastest connections to NAS resources and big data collections are achievable. Also for data interchange in a distributed compute cluster the highest available LAN connectivity is a must have.

The AIME R410 is equipped with a dedicated IPMI LAN interface with its advanced BMC it can be remote monitored and controlled (wake-up/reset). This features make a successful integration of the AIME R410 into a server rack cluster possible.

Developed for Machine Learning Applications

The AIME R410 was first designed for our own machine learning server needs and evolved in years of experience in deep learning frameworks and customized PC hardware building.

Our machines come with preinstalled Linux OS configured with latest drivers and frameworks like Tensorflow, Keras, PyTorch and Mxnet. Just login and start right away with your favourite machine learning framework.

Technical Details

Type	Rack Server 6HE, 45cm depth
CPU (configurable)	EPYC 7261 (8 cores, 2.5 GHz) EPYC 7232 (8 cores, 3.1 GHz) EPYC 7351 (16 cores, 2.4 GHz) EPYC 7302 (16 cores, 3.0 GHz) EPYC 7402 (24 cores, 2.8 GHz) EPYC 7502 (32 cores, 2.5 GHz)
RAM	64 / 128 / 256 GB ECC memory
GPU Options	4x NVIDIA RTX 2080TI oder 4x NVIDIA Titan RTX oder 4x NVIDIA Tesla V100
Cooling	CPU liquid cooled GPUs are cooled with an air stream high performance fans > 100000h MTBF
Storage	Upto 2 X 2TB NVMe SSD
Config Options	QLC: 1500 MB/s read, 1000 MB/s write TLC: 3500 MB/s read, 1750 MB/s, write MLC: 3500 MB/s read, 2700 MB/s write
Network	2 x 10 GBit LAN 1 x IPMI LAN
USB	2 x USB 3.0 Ports (vorne) 2 x USB 3.0 Ports (hinten)
PSU	2000 Watt power 80 PLUS Platinum certified (94% efficiency)
Noise-Level	< 50dBA
Dimensions (WxHxD)	440 x 265 x 430 mm
Price	from 7.749,00 € excl. MwSt.

