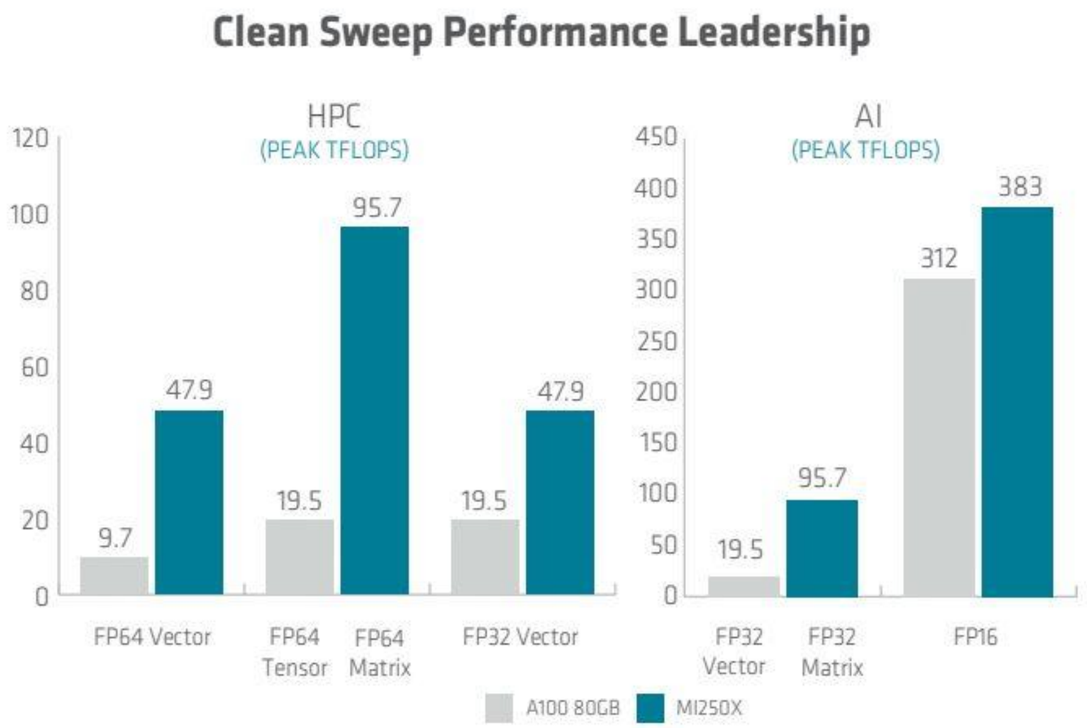
AMD INSTINCT™ MI200 SERIES ACCELERATOR

World's Fastest HPC and AI Accelerator[1]

The era of exascale is here. Immense computational power coupled with the fusion of HPC and AI is enabling researchers and scientists to tackle our most pressing challenges from climate change to vaccine research. With the AMD Instinct™ MI200 accelerators and ROCm™ 5.0 software ecosystem, innovators can tap the power of the world's most powerful HPC and AI data center GPUs to accelerate their time to science and discovery.[1]

Based on the 2nd Gen AMD CDNA™ architecture, AMD Instinct™ MI200 accelerators deliver a quantum leap in HPC and AI performance over competitive data center GPUs today. With an up to 4x advantage in HPC performance compared to competitive GPUs, the MI200 accelerator is the first data center GPU to deliver 383 teraflops of theoretical mixed precision FP16 performance for deep learning training, offering users a powerful platform to fuel the convergence of HPC and AI.[1]



Graph 1: Peak TFLOPS across range of mixed-precision Compute[1]

Innovations Delivering Performance Leadership

AMD innovations in architecture, packaging and integration are pushing the boundaries of computing by unifying the most important processors in the data center, the CPU and the GPU accelerator. With industry-first multi-chip GPU modules along with 3rd Gen AMD Infinity Architecture, AMD is delivering performance, efficiency and overall system throughput for HPC and AI using AMD EPYC™ CPUs and AMD Instinct™ MI200 series accelerators

Key Features

PERFORMANCE	MI250	MI250X
Compute Units	208CU	220CU
Stream Processors	13,312	14,080
Peak FP64/FP32 Vector	45.3 TFLOPS	47.9 TFLOPS
Peak FP64/FP32 Matrix	90.5 TFLOPS	95.7 TFLOPS
Peak FP16/BF16	362.1 TFLOPS	383.0 TFLOPS
Peak INT4/INT8	362.1 TOPS	383.0 TOPS

MEMORY		
Memory Size	128GB HBM2e	128GB HBM2e
Memory Interface	8,192 bits	8,192 bits
Memory Clock	1.6GHz	1.6GHz
Memory Bandwidth	up to 3.2TB/sec[2]	up to 3.2TB/sec[2]

RELIABILITY	MI250	MI250X
ECC (Full-chip)	Yes	Yes
RAS Support	Yes	Yes

SCALABILITY		
Infinity Fabric™ Links	up to 6	up to 8
Coherency Enabled	No	Yes
OS Support	Linux™ 64 Bit	Linux 64 Bit
AMD ROCm™ Compatible	Yes	Yes

BOARD DESIGN		
Form Factor	OAM	OAM
Thermal	Passive & Liquid	Passive & Liquid
Max Power	500W & 560W TDP	500W & 560W TDP
Bus Interface	PCIe® Gen 4 Support	
Warranty	3 Year Limited[4]	