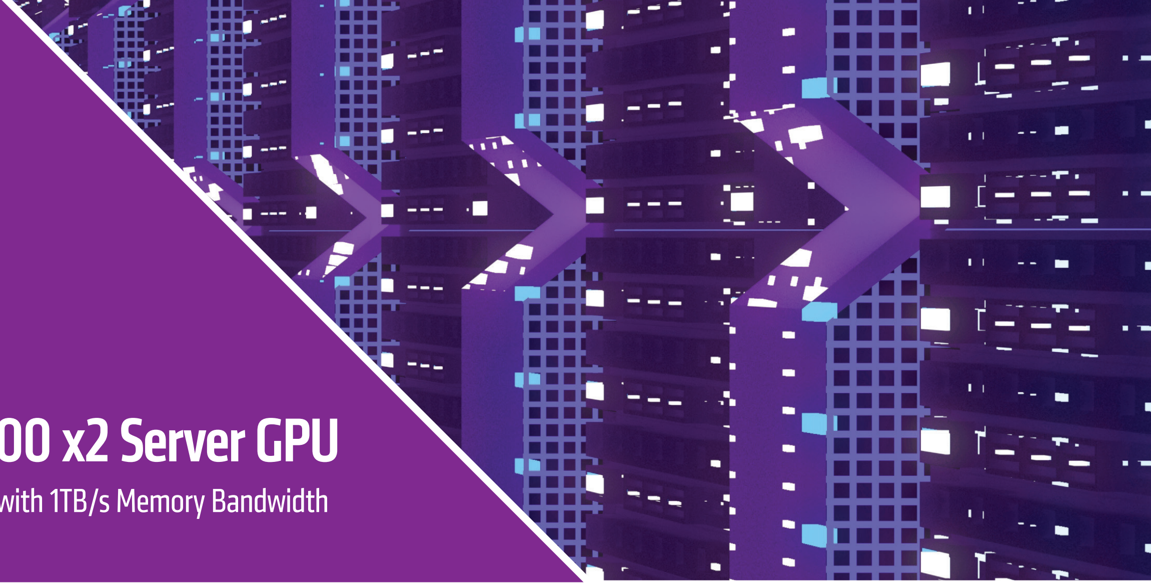


AMD FirePro™ S9300 x2 Server GPU

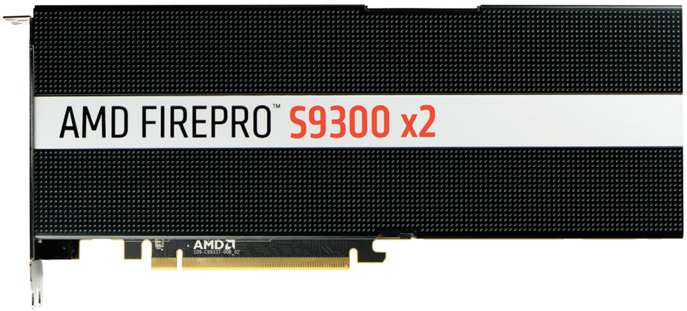
The World's First GPU Accelerator with 1TB/s Memory Bandwidth

Speed up your most complex HPC workloads in data analytics, or signal processing using C++ or OpenCL™, and witness the power of the world's first GPU Accelerator with HBM memory, the AMD FirePro™ S9300 x2. Experience the fastest 32-bit compute GPU accelerator, with up to 13.9 TFLOPS of peak single-precision compute accelerating your HPC workflows.[1,3]

Heterogeneous-compute Interface for Portability (HIP) Tool – Already have code written in CUDA? Don't want to be tied to a single vendor? Easily convert your code to C++ with this free and open source tool, while maintaining compatibility with CUDA compilers. The HIP tool allows developers to port the majority of their CUDA code over to C++ in a snap. Get started on the AMD FirePro S9300 x2 GPU, an open-source friendly accelerator from AMD, today.

Key Features

- *Max Power: 300W*
- *Bus Interface: PCIe® Gen 3 x16*
- *Form Factor: Dual Slot, Full Length, Full Height*
- *Cooling: Passive*
- *Memory Size: 8GB HBM*
- *OS Support: Linux® 64-bit*



Unparalleled Single Precision Performance

With 13.9 TFLOPS of peak single precision performance, the AMD FirePro S9300 x2 is the fastest GPU accelerator available, delivering up to 2x peak single precision performance over NVIDIA's Tesla M40 and 1.6x peak single precision performance over NVIDIA's Tesla K80.[2,3] The latest 3rd generation GCN architecture brings a number of enhancements to the GPU including a new 16-bit floating point and integer instruction set for more efficient use of memory bandwidth and reduced memory footprint.

World's First GPU Accelerator with HBM Memory

The AMD FirePro™ S9300 x2 Server GPU is the industry's first server GPU with 8GB of ultra-fast HBM memory and features 1 TB/s of memory bandwidth (512GB/s per GPU). HBM is a new type of memory design with low power consumption and ultra-wide communication lanes. It uses vertically stacked memory chips interconnected by microscopic wires called "through-silicon vias," or TSVs, placed directly onto the interposer, shortening the distance information has to travel between memory and processor.

Open Developer Ecosystem

AMD FirePro™ S9300 x2 Server GPU is fully compatible with the new AMD GPUOpen software stack, which includes a new open source Linux driver, built from the ground up and optimized for compute. This driver includes native support for GPU-to-GPU communication both within and between nodes.

Developers now have the option of using OpenCL™ or C++ to accelerate code on the GPU. The free, open source Heterogeneous Compute Compiler (HCC) leverages popular CLANG/LLVM technology and supports compilation of C++ code to x86 or GPU targets. Since a majority of developers already use C++ programming, there's no need to learn a new language to start accelerating on the GPU. The open source HCC compiler will allow developers to have single source code development for the new AMD FirePro™ S9300 x2 GPU.

FEATURES

13.9 TFLOPS of Peak Single Precision

Helps speed up time required to complete single precision floating point operations used within Simulations, Video Enhancement, Signal Processing, Video Transcoding and Digital Rendering applications where high performance takes precedence over accuracy. With the AMD FirePro™ S9300 x2 delivering 13.9 TFLOPS of peak single precision compute performance, one can configure a 2P server with 8 GPUs to achieve over 111 TFLOPS of peak single precision compute performance. In a standard 42U rack with 10x 4U servers, that's potentially over 1 PFLOP of single precision compute performance!

870 GFLOPS of Peak Double Precision

Helps speed up time required to complete double precision floating point operations used within Computational Fluid Dynamics, Structural Mechanics, Reservoir Simulation and Aerodynamics applications, where numerical precision is mission critical.

Half Precision (FP16) Support

Developers who do not need the accuracy of 32-bit mathematical operations can now use 16-bit operations to help achieve high performance through a more efficient use of memory bandwidth and reduced memory footprint.

8GB HBM Memory

HBM is a new type of memory design with low power consumption and ultra-wide communication lanes. It uses vertically stacked memory chips interconnected by microscopic wires called "through-silicon vias," or TSVs, placed directly onto the interposer, shortening the distance information has to travel between memory and processor.

GPUOpen Professional Compute

Comprised of an open-source Linux driver optimized for compute, support for GPU acceleration using a new compiler to process code written in the C++ programming language, and other developer tools such as the Heterogeneous-compute Interface for Portability (HIP) Tool to port code written for CUDA to C++.

OpenCL™ 1.2 Support

Helps professionals tap into the parallel computing power of modern GPUs and multicore CPUs to accelerate compute-intensive tasks in leading CAD/CAM/CAE and Media & Entertainment applications that support OpenCL. The AMD FirePro S9300 x2 Server GPU supports OpenCL™ 1.2, allowing developers to take advantage of new features that give GPUs more freedom to do the work they are designed to do.

AMD PowerTune

AMD PowerTune Technology is an intelligent power management system that monitors both GPU activity and power draw. AMD PowerTune optimizes the GPU to deliver low power draw when GPU workloads do not demand full activity and delivers the optimal clock speed to ensure the highest possible performance within the GPU's power budget for high intensity workloads.[4]

For more information, please visit **AMD.com/firepro/hpc**

[1] The AMD FirePro™ S9300 x2 GPU is the first professional dual GPU accelerator card with HBM memory available as of March 31st, 2016. FP-186

[2] AMD internal testing as of March 2016. System configurations may vary, yielding different results. The AMD FirePro™ S9300 x2 delivers up to 13.9 TFLOPS peak single-precision compute performance, while the NVIDIA Tesla M40 delivers up to 7 TFLOPS peak single-precision performance. For AMD, calculation based on a formula using a combination of shader units and clock speeds. For Nvidia, source of data as of March 18, 2016: http://www.nvidia.com/object/tesla-m40.html FP-187.

[3] AMD internal testing as of March 2016. System configurations may vary, yielding different results. The AMD FirePro™ S9300 x2 delivers 13.9 TFLOPS of peak single-precision compute performance. As of March 18th, 2016, NVIDIA's fastest Tesla board is the Tesla K80, which delivers up to 8.73 TFLOPS of single-precision compute performance. For AMD, calculation based on a formula using a combination of shader units and clock speeds. For Nvidia, source of data as of March 18, 2016 - http://www.nvidia.com/object/tesla-k80.html FP-189

[4] AMD PowerTune technology is offered by certain AMD FirePro™ products, which is designed to intelligently manage GPU power consumption in response to certain GPU load conditions. Not all products feature all technologies – check with your component or system manufacturer for specific model capabilities.