



AMD ROCm™ 6 OPEN SOFTWARE FOR AI AND HPC WORKLOADS



FREE YOUR WORKLOADS WITH THE ROCm™ 6 PLATFORM

The AMD ROCm™ 6 open-source software platform is optimized to extract the best HPC and AI workload performance from AMD Instinct™ MI300 accelerators while maintaining compatibility with industry software frameworks. ROCm consists of a collection of drivers, development tools, and APIs that enable GPU programming from low-level kernel to end-user applications and is customizable to meet your specific needs. Develop, collaborate, test, and deploy applications in a free, open source, and integrated software ecosystem. Once designed the software is then portable, allowing movement between accelerators from different vendors or between different inter-GPU connectivity architectures—all in a device-independent manner. ROCm is particularly well-suited to GPU-accelerated high-performance computing (HPC), artificial intelligence (AI), scientific computing, and computer aided design (CAD). A collection of advanced GPU software containers and deployment guides for HPC, AI & Machine Learning applications are now available on the [Infinity Hub](#) to help speed-up your system deployments and time to insights.

PROPEL YOUR AI APPLICATIONS

Upstream support for the most popular AI and ML frameworks—TensorFlow, JAX, and PyTorch—make it easy to adopt the ROCm platform. A broad range of supporting libraries to help improve your models and run-time environments including DeepSpeed, ONNX-RT, Jax, and CuPy all are supported as part of the ROCm platform. These features and optimizations provide an 8x gen on gen performance improvement for ROCm 6 + MI300X over ROCm 5 + MI250. The [AMD Infinity Hub](#) provides ready-to-run containerized frameworks, and our [GPU Accelerated Applications Catalog](#) lists the broad set of supporting libraries that are available to AI and ML practitioners.

ACCELERATE YOUR HIGH-PERFORMANCE COMPUTING WORKLOADS

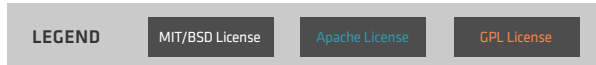
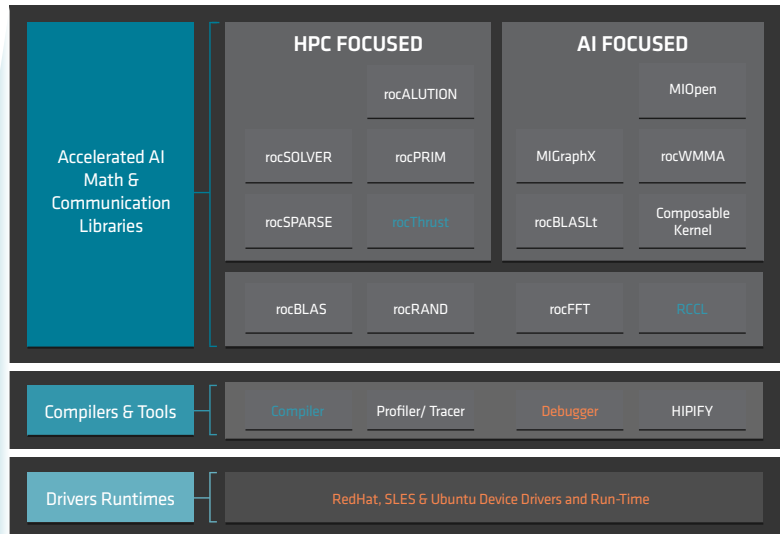
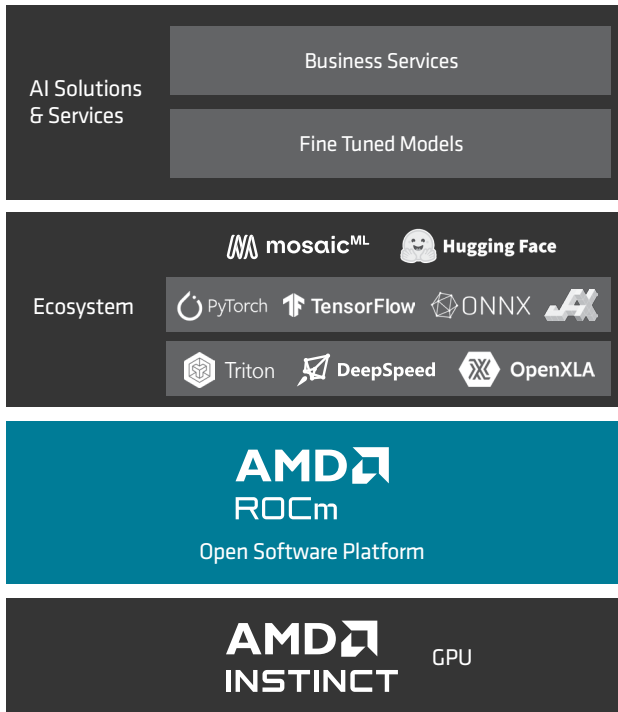
Some of the most popular HPC frameworks are part of the ROCm platform, including those to help parallelize operations across multiple GPUs and servers, handle memory hierarchies, and solve linear systems. Our [GPU Accelerated Applications Catalog](#) includes a vast set of platform-compatible HPC applications, including those in astrophysics, climate & weather, computational chemistry, computational fluid dynamics, earth science, genomics, geophysics, molecular dynamics, and physics. Many of these are available through the [AMD Infinity Hub](#), ready to download and run on servers with AMD Instinct™ accelerators.

WHAT'S NEW IN ROCm 6?

- 1. Expanded support:** For AMD Instinct™ MI300A and MI300X
- 2. Key AI supported features:** ROCm transformer engine, highly optimized attention algorithms, and proven collective communications libraries
- 3. Optimized performance:** Dynamic FP16, BF16, FP8 utilization, optimized HIPGraph
- 4. Enabling developer success:** Support for structured sparsity and quantization libraries
- 5. Expanded ecosystem support:** with latest frameworks, models, and ML pipelines

“Using Lamini software, ROCm has achieved software parity with CUDA for LLMs. We chose the AMD Instinct MI250 as the foundation for Lamini because it runs the biggest models that our customers demand and integrates finetuning optimizations. We use the large HBM capacity (128GB) on MI250 to run bigger models with lower software complexity than clusters of A100s.”

Greg Diamos
CTO at Lamini



COMBINING NEXT-GEN HARDWARE & SOFTWARE
LLAMA 2 70B INFERENCE



LEARN MORE

We've made the ROCm 6 platform even easier to learn about, install, and run:

- [AMD ROCm Docs](#) site features all the latest ROCm documentation including release notes, how to guides, tutorials, examples, and other resources for developers.
- [AMD Infinity Hub](#) gives you access to HPC applications and ML frameworks packaged as containers and ready to run. We've expanded our already big catalog to include weather, QFD, quantum chemistry, and more.
- [ROCm Application Catalog](#) gives the most full and up-to-date list of applications supported by the ROCm platform.
- [ROCm Developer Hub](#) is the new home for all developer resources, including training webinars, videos, blogs, and more.

FOR MORE INFORMATION VISIT AMD.COM/ROCM

© 2024 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow, AMD Instinct, ROCm, and combinations thereof are trademarks of Advanced Micro Devices, Inc. Other trademarks used in this publication are for identification purposes only and may be trademarks of their respective companies. PID#: 232383225