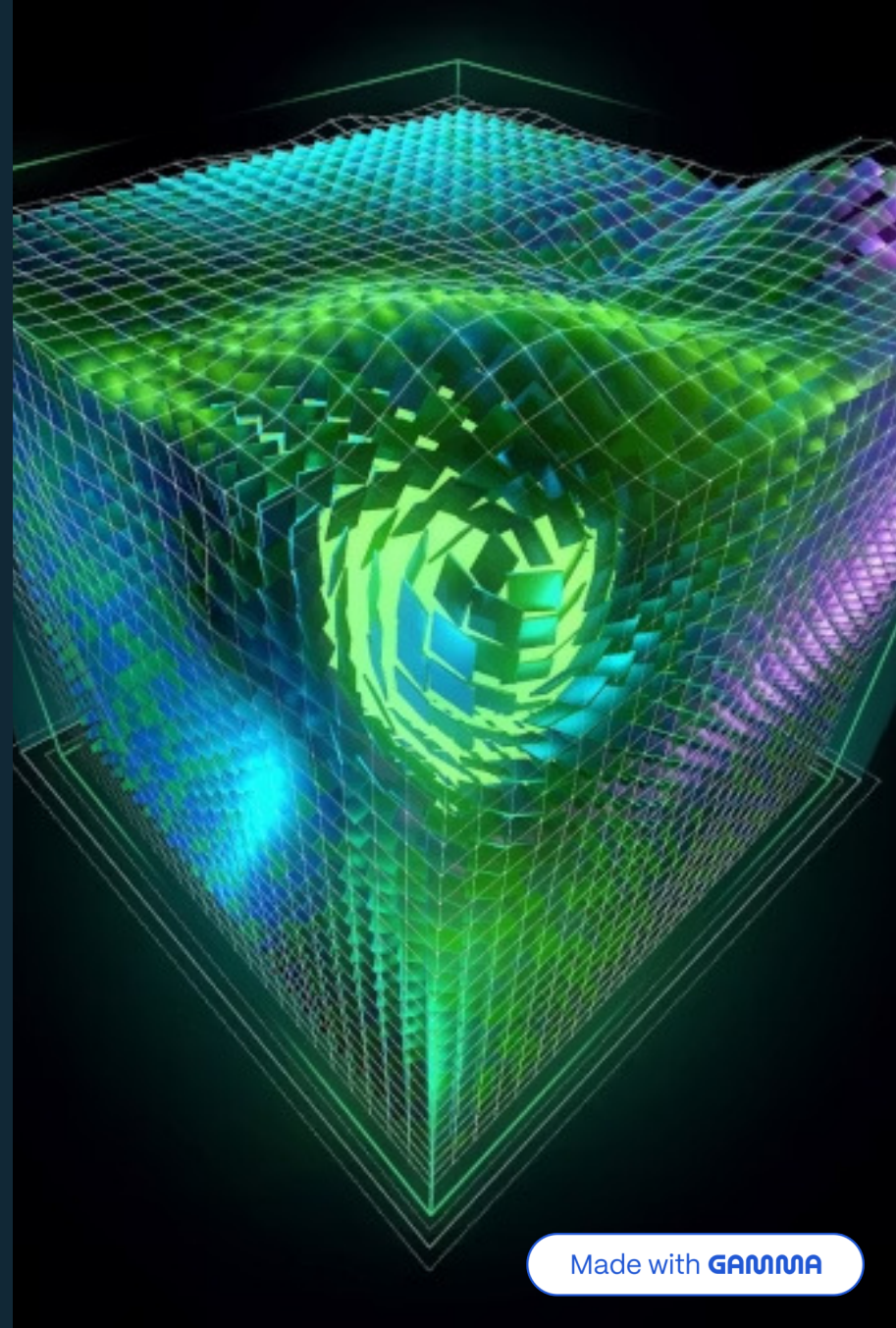
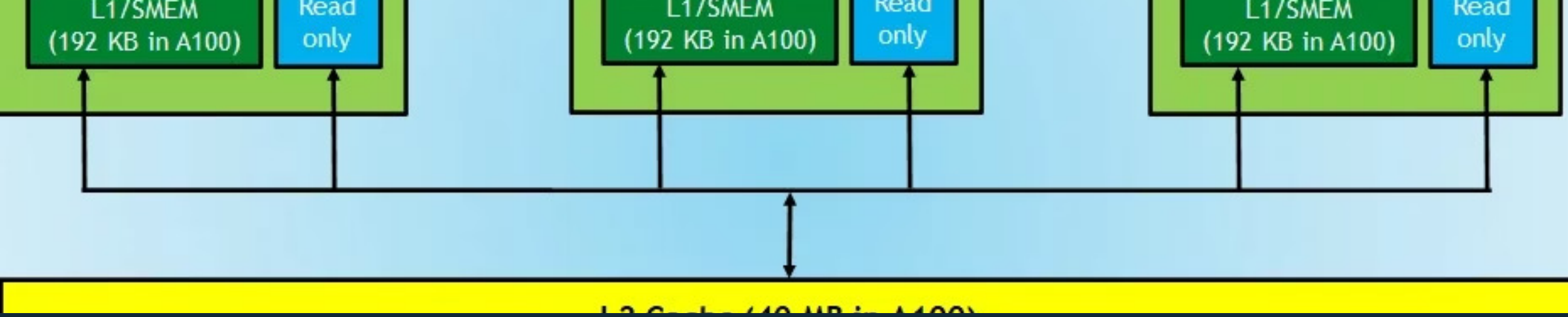
CUDA: Opening the GPU to the World

Before CUDA: GPUs were only for graphics. After CUDA: GPUs became massively parallel general-purpose computing.





Why CUDA Changed Everything

Hardware Access

Exposed hardware layer for parallel kernel execution.

Simplified Software

Unified runtime and driver simplified GPU access.

Democratized Programming

Made GPU programming accessible beyond graphics engineers.

CUDA Architecture in a Nutshell

Developer Writes Code

CUDA code (kernels) created.

Passes Through Runtime

Code goes through CUDA runtime to NVIDIA driver.

Driver Communicates

Driver directly communicates with GPU cores.

Parallel Execution

Code executes massively in parallel on hardware.

Developer Workflow with CUDA

Key Tools

- CUDA Toolkit
- NVIDIA documentation
- Developer forums

Workflow

Code → nvcc compiler → CUDA runtime/driver → GPU execution.

Accessibility

Personal machines, labs, cloud notebooks (Colab).

NVIDIA Product Ecosystem

Consumer GPUs

RTX: AI research, gaming, prosumers.

Enterprise

DGX systems: AI supercomputers.

Data Center

Tesla, A100, H100, Blackwell.

Cloud Support

Across AWS, GCP, Azure.

Open Libraries Boosted Adoption

Key Libraries

- cuBLAS: optimized linear algebra.
- cuDNN: deep learning primitives (convolutions, RNNs).

Impact

Abstracted complexity, easy integration with PyTorch, TensorFlow.

Researchers innovated without reinventing GPU programming.

Early Wins: Research & Industry

AI

OpenAI scaled deep learning training.

Healthcare

Faster drug discovery simulations.

Robotics

Real-time motion planning and vision.

Scientific Computing

Climate models, astronomy, physics.

CUDA’s Role in AI Boom

GPUs became the backbone for Machine Learning/Deep Learning.

CUDA libraries served as building blocks for AI frameworks.

Accelerated adoption led to the explosion of generative AI.

Next Steps: Hands-On with CUDA

We will now explore examples in Colab to demonstrate how to access CUDA and leverage its power for amazing applications.