NVIDIA A100

SUPERCHARGES CFD WITH ALTAIR AND GOOGLE CLOUD

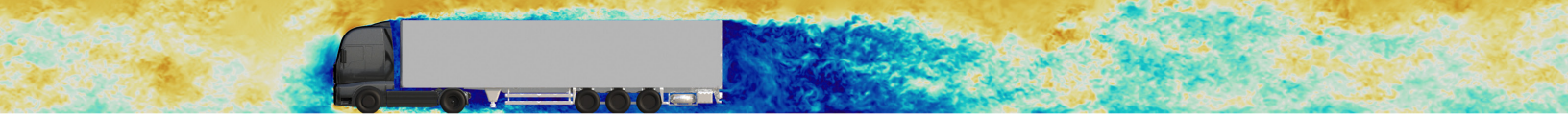


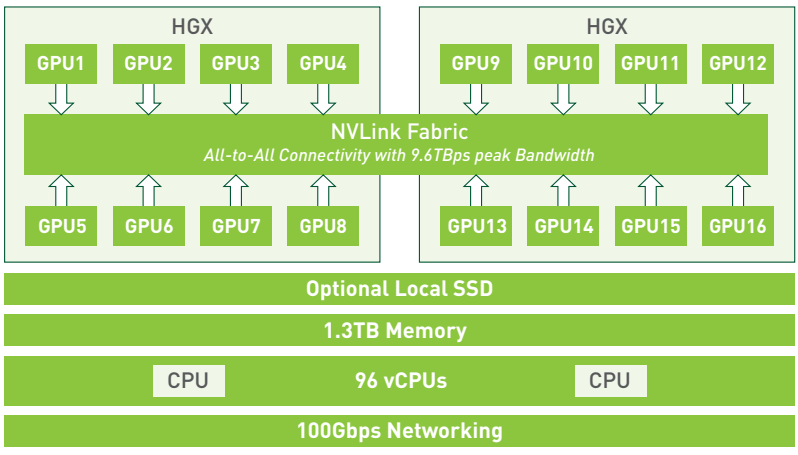
Image courtesy of Altair

Ultra-Fast CFD powered by Altair, Google Cloud, and NVIDIA

The newest A2 VM family on Google Cloud Platform was designed to meet today's most demanding applications. With up to 16 NVIDIA® A100 Tensor Core GPUs in a node, each offering up to **20X the compute performance** of the previous generation and a whopping total of up to 640GB of GPU memory, CFD simulations with Altair ultraFluidX™ were the ideal candidate to test the performance and benefits of the new normal.

With ultraFluidX, highly resolved transient aerodynamics simulations can be performed on a single server. The Lattice Boltzmann Method is a perfect fit for the massively parallel architecture of NVIDIA GPUs, and sets the stage for unprecedented turnaround times. What used to be overnight runs on single nodes now becomes possible within working hours by utilizing state-of-the-art GPU-optimized algorithms, the new NVIDIA A100, and Google Cloud's A2 VM family, while delivering the high fidelity of a transient LES aerodynamics simulation.

New A2-MegaGPU-16g instance with 16 A100 GPUs



Google Cloud

TAILORED FOR YOU

- Get exactly as many GPUs as you need, for exactly as long as you need them, and pay just for that.

FASTER OUT OF THE BOX

- The new **Ampere GPU architecture** on Google Cloud A2 instances delivers unprecedented acceleration and flexibility to power the world's highest-performing elastic data centers for AI, data analytics, and HPC applications.

THINK THE UNTHINKABLE

- With a total of up to 312 TFLOPS FP32 peak performance in the same node linked with NVLink technology, with an aggregate bandwidth of 9.6TB/s, new solutions to many tough challenges are within reach.

Google Cloud

ALTAIR

Ultra-Fast CFD powered by Altair, Google Cloud, and NVIDIA

We selected three different production-level cases to assess the new possibilities unleashed by the A2-MegaGPU-16g instance: The Altair CX1 concept design, a truck platooning case based on the generic FAT truck geometry, as analyzed by Schnepf et al.[1], and a high-fidelity platooning case with even higher resolution.

The CX1 case features 230 million voxels, a highest resolution of 1.6 millimeters, and a physical simulation time of three seconds, which used to be a typical overnight case on the previous NVIDIA Volta™ architecture. Now, the case can be run during the workday in as little as five hours when using all 16 GPUs of the instance, allowing for a completely new way of working in time-critical situations.

The truck platooning case of Schnepf et al.[1] consists of 400 million voxels at a highest resolution of four millimeters, and is run for a physical time of 40 seconds to allow for enough flow passes over the whole convoy of three trucks with a total length of 130 meters. Using the massive compute power of the whole instance, the case now just needs 15 hours to complete and can effectively be run overnight.

Finally, the high-fidelity platooning case features a resolution of down to two millimeters, leading to a total size of 640 million voxels and increasing the computational complexity compared to the other platooning case by 3.5X. Using conventional CPU-based solutions, such a heavy case would need to be run for several days on thousands of cores. In contrast, thanks to the massive compute power provided by the Ampere architecture on Google Cloud's A2 VM family, and more than 90% strong scaling efficiency on 16 vs. 8 GPUs facilitated by the extreme performance of third-generation NVIDIA NVLink® and NVIDIA NVSwitch™, ultraFluidX can turn around the case in just about 48 hours on a single A2-MegaGPU-16g instance.

With all these great new possibilities, are you ready to think the unthinkable?

RUN ULTRA-FAST

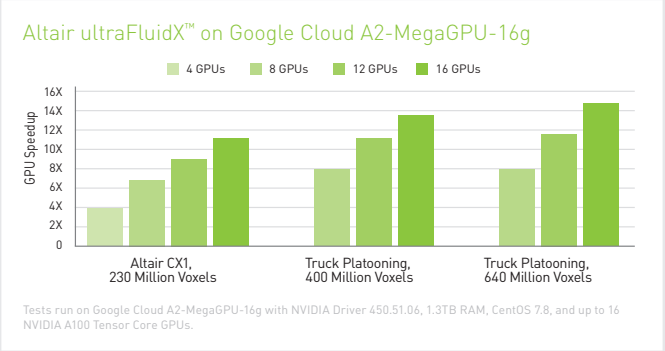
- As a native GPU-based code using a low-level CUDA C++ implementation and CUDA-aware Open MPI, ultraFluidX naturally leverages the massive compute power and memory bandwidth of the Ampere Architecture, and the extreme scalability of third-generation NVLink and NVSwitch.

RUN TRANSIENT

- Vehicle aerodynamics are highly unsteady by nature. With ultraFluidX, well-resolved transient LES simulations are now affordable. No need to press transient physics into a steady-state corset anymore.

SAVE MONEY

- Conventional simulation approaches need thousands of CPU cores to achieve the turnaround times of ultraFluidX. Our GPU-based solution massively increases throughput while reducing hardware and energy costs.



To learn more about the NVIDIA A100 Tensor Core GPU, visit **http://nv/sl**

1 Schnepf, B., Kehrer, C., and Maeurer, C., "Multidisciplinary Investigation of Truck Platooning," SAE Technical Paper 2020-37-0028, 2020, **https://doi.org/10.4271/2020-37-0028**.