FUN3D v14.1 Training
FUN3D Performance at NAS

Eric Nielsen

December 13, 2024

(Updated January 2026 to include AMD Turin CPU and NAS GH200 GPU data)

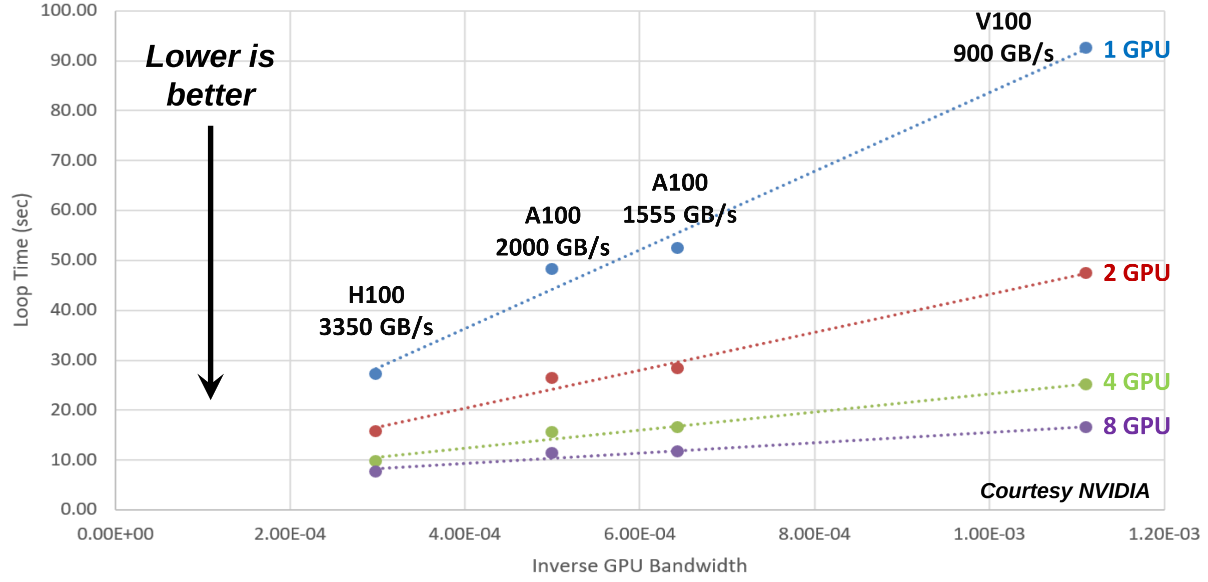
Goal and Scope

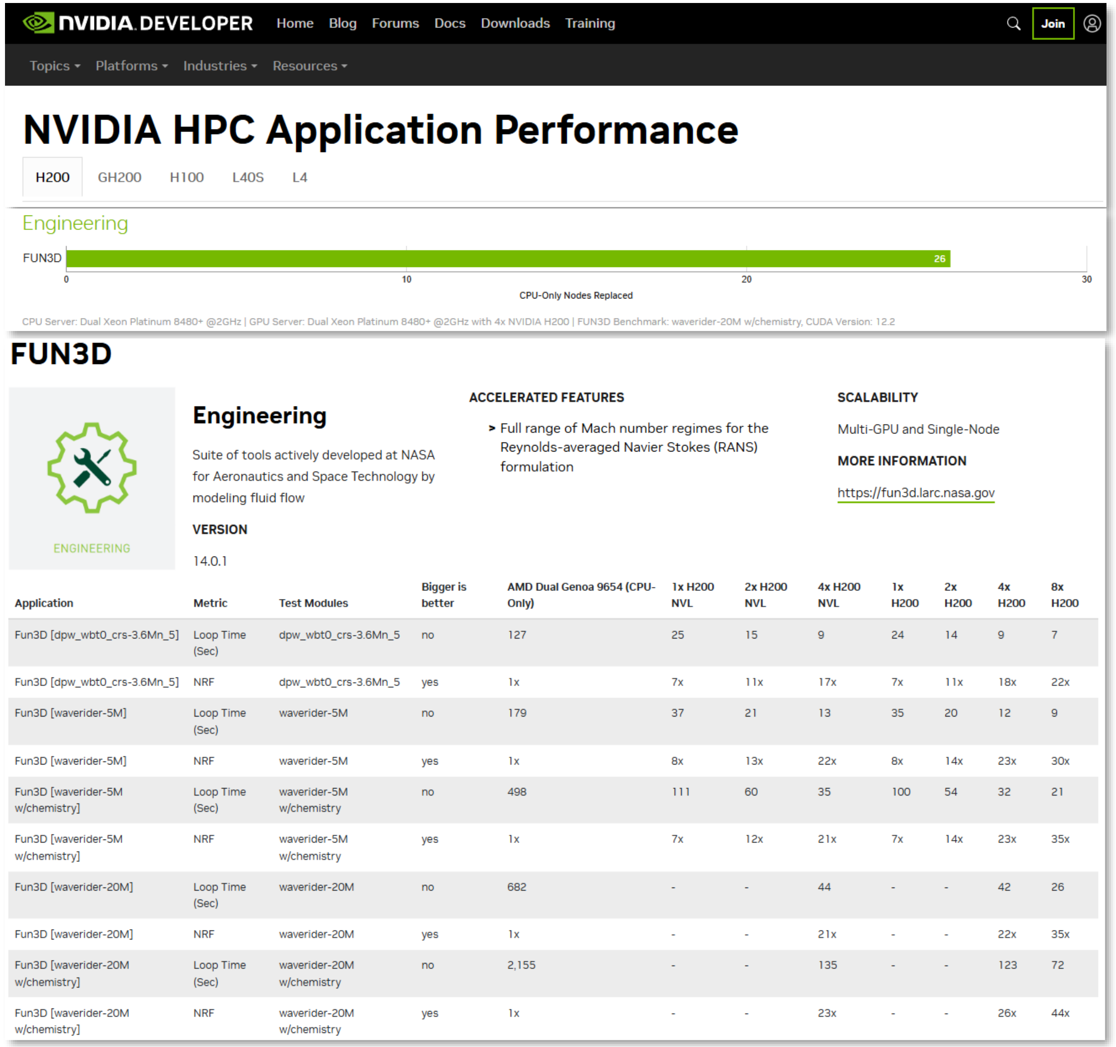
- Goal is to provide device-level performance data for FUN3D finite-volume execution on recent CPU- and GPU-based hardware options available at NAS using simple, yet relevant simulations
- Hopefully such data will be useful in scoping future allocation requests and campaigns on the NAS systems
- What is not covered here:
  - How to run a case
  - Scaling performance
  - What options are / are not available for GPU execution
  - Execution on AMD or Intel GPUs
  - User workflows vary widely in practice; potential impacts are only briefly touched on here
- For such topics, please see the FUN3D User Manual, FUN3D publications, and/or the broad range of training content archived on https://fun3d.larc.nasa.gov , as well as the documentation available on the NAS website
- If questions remain, please reach out to fun3d-support@lists.nasa.gov

Background

FUN3D is memory bound:
In general, performance scales with memory bandwidth

FUN3D Performance on Recent NVIDIA Architectures





NVIDIA DEVELOPER Home Blog Forums Docs Downloads Training Join

Topics Platforms Industries Resources

NVIDIA HPC Application Performance

H200 GH200 H100 L40S L4

Engineering

FUN3D 26

0 10 20 30

CPU-Only Nodes Replaced

CPU Server: Dual Xeon Platinum 8480+ @2GHz | GPU Server: Dual Xeon Platinum 8480+ @2GHz with 4x NVIDIA H200 | FUN3D Benchmark: waverider-20M w/chemistry, CUDA Version: 12.2

FUN3D

ENGINEERING

Engineering

Suite of tools actively developed at NASA for Aeronautics and Space Technology by modeling fluid flow

VERSION

14.0.1

ACCELERATED FEATURES

> Full range of Mach number regimes for the Reynolds-averaged Navier Stokes (RANS) formulation

SCALABILITY

Multi-GPU and Single-Node

MORE INFORMATION

https://fun3d.larc.nasa.gov

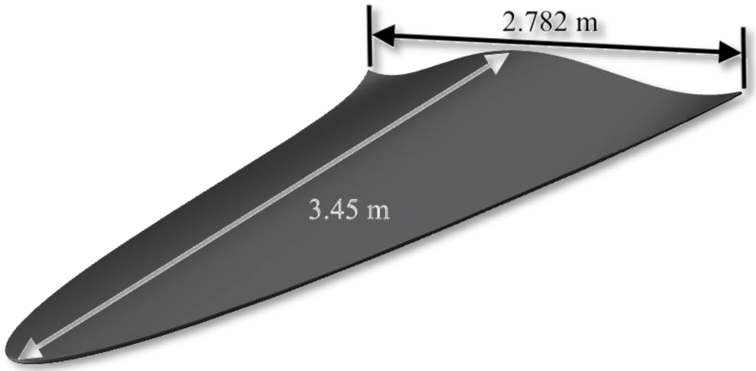
Application	Metric	Test Modules	Bigger is better	AMD Dual Genoa 9654 (CPU-Only)	1x H200 NVL	2x H200 NVL	4x H200 NVL	1x H200	2x H200	4x H200	8x H200
Fun3D [dpw_wbt0_crs-3.6Mn_5]	Loop Time (Sec)	dpw_wbt0_crs-3.6Mn_5	no	127	25	15	9	24	14	9	7
Fun3D [dpw_wbt0_crs-3.6Mn_5]	NRF	dpw_wbt0_crs-3.6Mn_5	yes	1x	7x	11x	17x	7x	11x	18x	22x
Fun3D [waverider-5M]	Loop Time (Sec)	waverider-5M	no	179	37	21	13	35	20	12	9
Fun3D [waverider-5M]	NRF	waverider-5M	yes	1x	8x	13x	22x	8x	14x	23x	30x
Fun3D [waverider-5M w/chemistry]	Loop Time (Sec)	waverider-5M w/chemistry	no	498	111	60	35	100	54	32	21
Fun3D [waverider-5M w/chemistry]	NRF	waverider-5M w/chemistry	yes	1x	7x	12x	21x	7x	14x	23x	35x
Fun3D [waverider-20M]	Loop Time (Sec)	waverider-20M	no	682	-	-	44	-	-	42	26
Fun3D [waverider-20M]	NRF	waverider-20M	yes	1x	-	-	21x	-	-	22x	35x
Fun3D [waverider-20M w/chemistry]	Loop Time (Sec)	waverider-20M w/chemistry	no	2,155	-	-	135	-	-	123	72
Fun3D [waverider-20M w/chemistry]	NRF	waverider-20M w/chemistry	yes	1x	-	-	23x	-	-	26x	44x

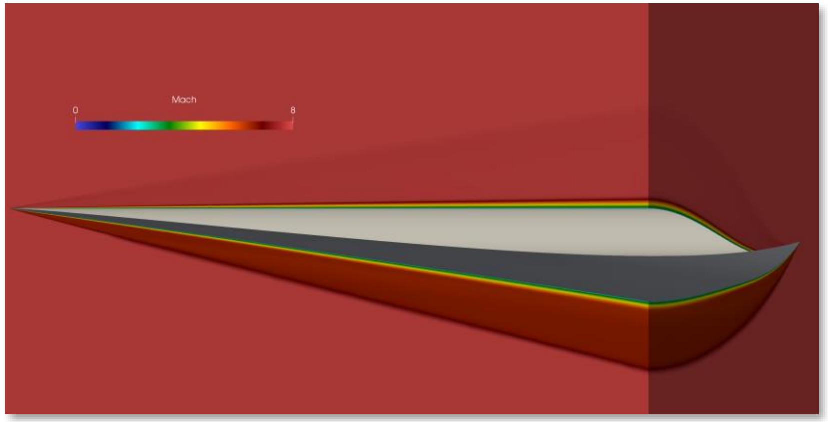
https://developer.nvidia.com/hpc-application-performance

Test Cases

- Tests are performed using the FUN3D v14.1 modules deployed on NAS
- Timings reflect solely the cost of solving the governing equations
- Two test cases are used here, both based on a conceptual hypersonic waverider vehicle
- Freestream conditions are M_∞ = 8.0, T_∞ = 227 K, and Re_L = 9.5 million (30-kilometer altitude)
- Grid consists of 5 million points, 5.3 million prisms, 13.5 million tetrahedra, and 5 thousand pyramids
    - Sized to fit within available memory on the target hardware
- Cases are run using a Reynolds-averaged Navier-Stokes formulation with the single-equation Spalart-Allmaras turbulence model and implicit time integration
- All other inputs taken as FUN3D default values

Case 1 uses a perfect-gas formulation
Case 2 uses a one-temperature, 5-species gas model





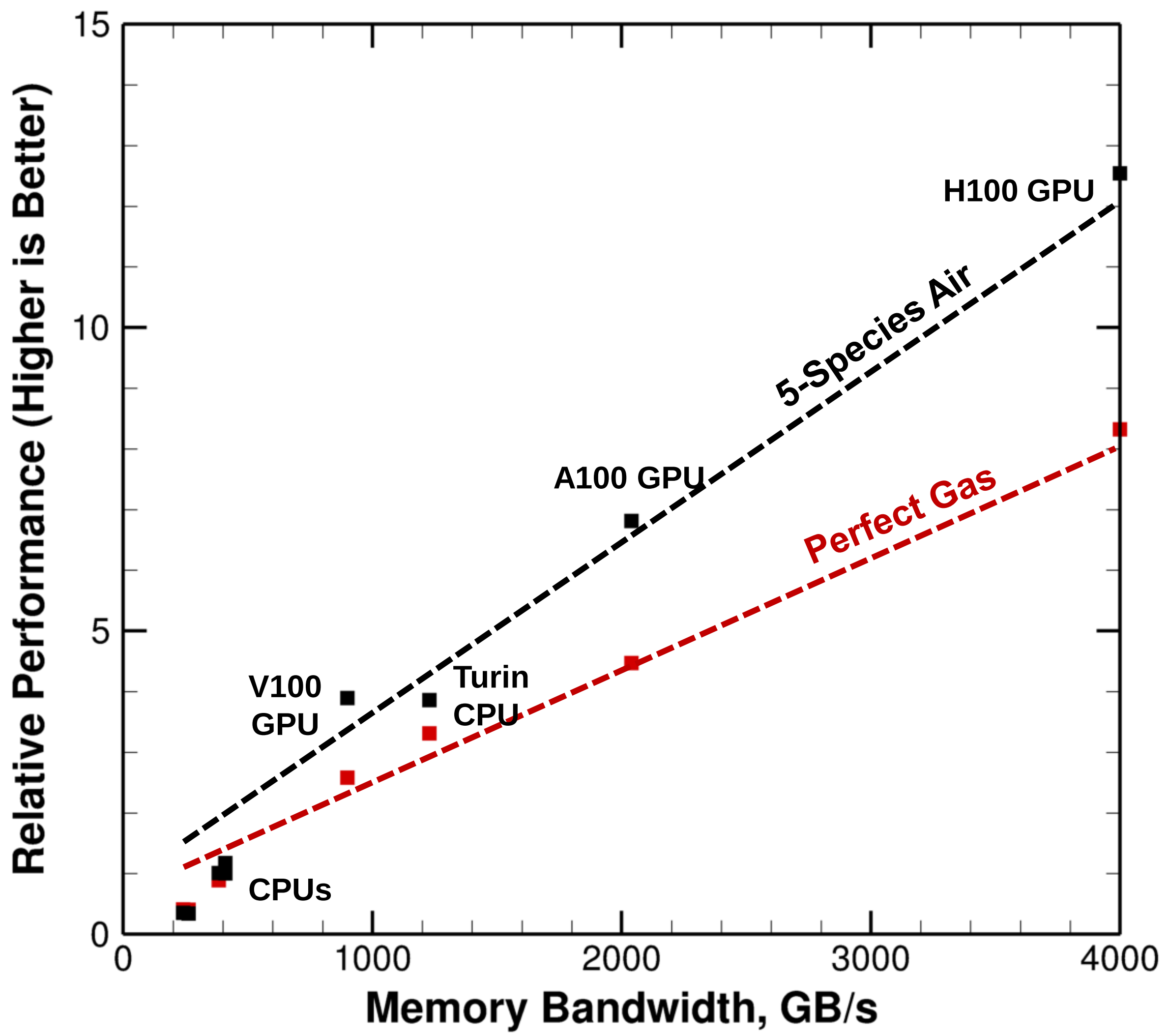
NAS Hardware Specs and Relative FUN3D Performance

	Intel Skylake CPU*	Intel Cascade Lake CPU†	AMD Rome CPU‡	AMD Milan CPU¶	AMD Turin CPU‡‡	NVIDIA V100 GPU§	NVIDIA A100 GPU**	NVIDIA GH200††	
								CPU	GPU
Hardware Details	2 x 20c Xeon 6148	2 x 20c Xeon 6248	2 x 64c 7742	2 x 64c 7763	2 x 128c 9745	32 GB SXM	80 GB SXM	72c Grace	96 GB SXM H100
Peak Memory Bandwidth, GB/s	240	262	410	410	1,228	900	2,039	384	4,000
FUN3D Performance: Perfect Gas	0.41	0.40	1.00	1.14	3.31	2.58	4.47	0.89	8.32
FUN3D Performance: 5-Species Air	0.35	0.34	1.00	1.17	3.86	3.89	6.81	1.01	12.54

- All simulations are performed using a single CPU node (all available cores) or a single GPU
- Performance is normalized by the Rome CPU result and **independently for each gas model**; higher numbers are better

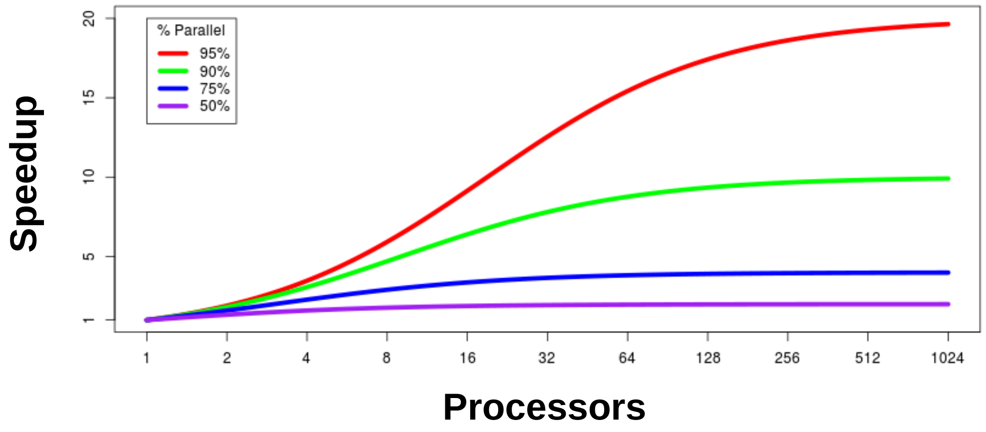
* *https://en.wikichip.org/wiki/intel/xeon_gold/6148*
† *https://en.wikichip.org/wiki/intel/xeon_gold/6248*
‡ *https://www.amd.com/content/dam/amd/en/documents/products/epyc/amd-epyc-7002-series-datasheet.pdf*
¶ *https://www.amd.com/en/products/processors/server/epyc/7003-series/amd-epyc-7763.html*
‡‡ *https://www.amd.com/en/products/processors/server/epyc/9005-series/amd-epyc-9745.html*
§ *https://images.nvidia.com/content/technologies/volta/pdf/tesla-volta-v100-datasheet-letter-fnl-web.pdf*
** *https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/a100/pdf/nvidia-a100-datasheet-nvidia-us-2188504-web.pdf*
†† *https://resources.nvidia.com/en-us-grace-cpu/grace-hopper-superchip*

Relative FUN3D Performance on NAS Hardware



GPU Computing and Amdahl's Law

- GPU computing can be game-changing for costly CFD applications, e.g.,
  - Scale-resolving simulations
  - Aeroelastic applications
  - Trajectory analysis / 6-DOF simulations
  - Finite-rate chemistry with complex gas models
- Similar benefits can certainly be realized for engineering-class applications (e.g., steady-state RANS using perfect gas), but one should be aware of some implications related to Amdahl's Law



Amdahl's Law at the Solver and Workflow Levels

Solver Level

- All steps needed to solve the governing equations are performed on the GPU
- However, some auxiliary operations take place on the CPU; some are optional but some may not be
  - Some of these kernels do not map well to a GPU, while the team has simply lacked bandwidth to address others
  - Long-duration simulations easily amortize many of these operations
  - Engineering-class simulations that run in minutes call for potential mitigation strategies outlined in the User Manual
- We are working hard on new approaches and paradigms here – hopefully to appear in future releases

Workflow Level

- Most engineering workflows are far more complicated than simply solving Navier-Stokes on the latest gold shiny hardware, e.g.,
  - Operations / applications / motifs that do not scale or perhaps do not map to certain architectures
  - COTS applications and components
  - File manipulations and file system interactions
  - CAD applications, mesh generation / adaptation
  - Distributed / remote computing
  - Multidisciplinary / multiphysics concerns
- Tough to offer general solutions here, but always happy to chat

- **Get a small allocation and run some tests with your actual workflow or a close surrogate**
- **Please reach out to us – your real-world challenges and feedback are extremely valuable**

Summary

- Performance and scaling will vary with specific user workflows, but the general trends shown here should hold
- Explore some representative tests ahead of time if possible
- See User Manual, publications, and prior tutorials regarding performance and suggestions when computing in GPU environments
- Please feel free to contact us at fun3d-support@lists.nasa.gov to discuss specific workflows or questions