NASA STTR 2021 Phase I Solicitation

T9.02   Rapid Development of Advanced High-Speed Aerosciences Simulation Capability

Lead Center: ARC

Participating Center(s): JSC, LaRC

Scope Title:

Aerothermal Simulation on Advanced Computer Architectures

Scope Description:

Aerothermodynamic simulations of planetary entry vehicles such as Orion and Dragonfly are complex and time consuming. These simulations, which solve the multispecies, multitemperature Navier-Stokes equations, require detailed models of the chemical and thermal nonequilibrium processes that take place in high-temperature shock layers. Numerical solution of these models results in a large system of highly nonlinear equations that are exceptionally stiff and difficult to solve efficiently. As a result, aerothermal simulations routinely consume 20 to 50 times the compute resources required by more conventional supersonic computational fluid dynamics (CFD) analysis, limiting the number of simulations delivered in a typical engineering design cycle to only a few dozen. Moreover, entry system designs are rapidly increasing in complexity, and unsteady flow phenomena such as supersonic retropropulsion are becoming critical considerations in their design. This increases the compute resources required for aerothermal simulation by an additional one to two orders of magnitude, which precludes the delivery of such simulations in engineering-relevant timescales.

In order to deliver the aerothermal simulations required for NASA's next generation of entry systems, access to greatly expanded compute resources is required. However, scaling up conventional central processing unit (CPU-) based supercomputers is problematic due to cost and power constraints. Many-core accelerators, such as Nvidia's general-purpose graphical processing units (GPGPUs), offer increased compute capability with reduced cost and power requirements and are seeing rapid adoption in top-end supercomputers. As of June 2020, 144 of the top 500 fastest supercomputers leveraged accelerators or co-processors, including 6 of the top 10 [1]. All three of the U.S. Department of Energy's upcoming exascale supercomputers will be accelerated using GPGPUs [2]. NASA has deployed Nvidia v100 GPGPUs to the Pleiades supercomputer [3]. Critically, NASA's aerothermal simulation tools are fundamentally unable to run on many-core accelerators, and must be reengineered from the ground up to efficiently exploit such devices.

This scope seeks to revolutionize NASA's aerothermal analysis capability by developing novel simulation tools capable of efficiently targeting the advanced computational accelerators that are rapidly becoming standard in the world's fastest supercomputers. A successful solution within this scope would demonstrate efficient simulation of a large-scale aerothermal problem of relevance on an advanced many-core architecture, for example, the Nvidia

Volta GPGPU, using a prototype software package.

Expected TRL or TRL Range at completion of the Project: 2 to 5
Primary Technology Taxonomy:
Level 1: TX 09 Entry, Descent, and Landing
Level 2: TX 09.1 Aeroassist and Atmospheric Entry
Desired Deliverables of Phase I and Phase II:

- Software

Desired Deliverables Description:

The desired deliverable at the conclusion of Phase I is a prototype software package capable of solving the multispecies, multitemperature, reacting Euler equations on an advanced many-core accelerator such as an Nvidia v100 GPGPU. Parallelization across multiple accelerators and across nodes is not required. The solver shall demonstrate offloading of all primary compute kernels to the accelerator, but may do so in a nonoptimal fashion, for example, using managed memory, serializing communication and computation, etc. Some noncritical kernels such as boundary condition evaluation may still be performed on a CPU. The solver shall demonstrate kernel speedups (excluding memory transfer time) when comparing a single accelerator to a modern CPU-based, dual-socket compute node. However, overall application speedup is not expected at this stage. The solver shall be demonstrated for a relevant planetary entry vehicle such as FIRE-II, Apollo, Orion, or the Mars Science Laboratory.

A successful Phase II deliverable will mature the Phase I prototype into a product ready for mission use and commercialization. Kernels for evaluating viscous fluxes shall be added, enabling computation of laminar convective heat transfer to the vehicle. Parallelization across multiple accelerators and multiple compute nodes shall be added. Good weak scaling shall be demonstrated for large 3D simulations (>10M grid cells). The implementation shall be sufficiently optimized to achieve an ~5x reduction in time-to-solution compared to NASA's Data-Parallel Line Relaxation (DPLR) aerothermal simulation code, assuming each dual-socket compute node is replaced by a single accelerator (i.e., delivered software running on eight GPGPUs shall be 5 times faster than DPLR running on eight modern, dual-socket compute nodes). Finally, the accuracy of the delivered software shall be verified by comparing to the DPLR and/or LAURA codes. The verification study shall consider flight conditions from at least two of the following planetary destinations: Earth, Mars, Titan, Venus, and Uranus/Neptune.

State of the Art and Critical Gaps:

NASA's existing aerothermal analysis codes (LAURA, DPLR, US3D, etc.) all utilize domain-decomposition strategies to implement coarse-grained, distributed-memory parallelization across hundreds or thousands of conventional CPU cores. These codes are fundamentally unable to efficiently exploit many-core accelerators, which require the use of fine-grained, shared-memory parallelism over hundreds of thousands of compute elements. Addressing this gap requires reengineering our tools from the ground up and developing new algorithms that expose more parallelism and scale well to small grain sizes.

Many-core accelerated CFD solvers exist in academia, industry, and government. Notable examples are PyFR from Imperial College London [4], the Ansys Fluent commercial solver [5], and NASA Langley's FUN3D code, which recently demonstrated a 30x improvement in node-level performance using Nvidia v100 GPUs [6]. However, nearly all previous work has focused on perfect gas flow models, which have different algorithmic and resource requirements compared to real gas models. The Sandia Parallel Aerodynamics and Reentry Code (SPARC) solver is the only project of note to have demonstrated efficient real-gas capability at scale using many-core accelerators [7].

Relevance / Science Traceability:

This scope is directly relevant to NASA space missions in both HEOMD and SMD with an entry, descent, and

landing (EDL) segment. These missions depend on aerothermal CFD to define critical flight environments and would derive large, recurring benefits from a more responsive and scalable simulation capability. This scope also has potential cross-cutting benefits for tools used by ARMD to simulate airbreathing hypersonic vehicles. Furthermore, this scope directly supports NASA's CFD Vision 2030 Study, which calls for sustained investment to ensure that NASA's computational aeroscience capabilities can effectively utilize the massively parallel, heterogeneous (i.e., GPU-accelerated) supercomputers expected to be the norm in 2030.

References:

1. "Japan Captures TOP500 Crown with Arm-Powered Supercomputer," June 22, 2020.
2. R. Smith: "El Capitan Supercomputer Detailed: AMD CPUs & GPUs To Drive 2 Exaflops of Compute," March 4, 2020.
3. NASA HECC Knowledge Base: "New NVIDIA V100 Nodes Available," 21 June, 2019.
4. F. Witherden, et al.: "PyFR: An Open Source Framework for Solving Advection–Diffusion Type Problems on Streaming Architectures Using the Flux Reconstruction Approach," Computer Physics Comm., vol. 185, no. 11, pp. 3028-3040, 2014.
5. V. Sellappan and B. Desam: "Accelerating ANSYS Fluent Simulations with NVIDIA GPUs," 2015.
6. E. Neilsen, et al.: "Unstructured Grid CFD Algorithms for NVIDIA GPUs," March 2019.
7. M. Howard, et al.: "Employing Multiple Levels of Parallelism for CFD at Large Scales on Next Generation High-Performance Computing Platforms," ICCFD10, Barcelona, Spain, 2018.
8. J. Slotnick, et al.: "CFD Vision 2030 Study: A Path to Revolutionary Computational Aerosciences," NASA/CR–2014-218178, March 2014.

Scope Title:

Robust Aerothermal Simulation of Complex Geometries

Scope Description:

NASA's production aerothermodynamic flow solvers all share a common characteristic: they utilize second-order accurate finite volume schemes to spatially discretize the governing flow equations. Schemes of this type are ubiquitous in modern compressible CFD solvers. They are simple to implement, perform well on current computer architectures, and provide reasonable accuracy for a wide range of problems. Unfortunately, one area where these schemes struggle to deliver high accuracy is at hypersonic speeds when a strong shock wave forms ahead of the vehicle. In such cases, the computed surface heat flux exhibits extreme sensitivity to the design of the computational grid near the shock [1], which must be constructed from cell faces that are either parallel or perpendicular to the shock to minimize error.

This stringent requirement for shock-aligned grids precludes the use of fully unstructured tetrahedral meshes in aerothermal simulation. While this restriction is manageable for simple or idealized entry systems [2], unstructured grids have significant accuracy and efficiency benefits for complex vehicle geometries, for example, ADEPT, and flow fields, for example, Mars 2020 reaction control system (RCS) firings, where large disparities in length scales must be resolved accurately. Moreover, unstructured grids can be developed much more rapidly and with a much higher degree of automation than traditional structured grid topologies [3]. As such, they are widely used in most

other CFD subdisciplines.

Fortunately, recent research has demonstrated that high-order, finite-element schemes such as the Discontinuous Galerkin (DG) method can achieve high-quality solutions for shock-dominated flows on unstructured grids when appropriate stabilization mechanisms are employed [4][5]. This research also suggests high-order methods are largely insensitive to the choice of the upwind flux function, potentially resolving a long-standing deficiency of second-order finite volume schemes at high speeds. However, while DG methods are robust and commonly applied in subsonic regimes, their continued development for aerothermal applications is hampered by ad hoc implementations in research-level codes.

This scope seeks to revolutionize NASA's aerothermal analysis capability by enabling rapid, robust, and highly automated analysis of complex entry systems using fully unstructured tetrahedral grids. A successful solution within this scope would demonstrate accurate simulation of a 3D capsule geometry at conditions relevant to planetary entry using DG or an equivalent numerical scheme in a prototype software package.

Expected TRL or TRL Range at completion of the Project: 2 to 5
Primary Technology Taxonomy:
Level 1: TX 09 Entry, Descent, and Landing
Level 2: TX 09.1 Aeroassist and Atmospheric Entry
Desired Deliverables of Phase I and Phase II:

- Software
- Prototype

Desired Deliverables Description:

The desired deliverable at the conclusion of Phase I is a prototype software package capable of solving the two-dimensional, multispecies, multitemperature, reacting Euler equations on unstructured triangular grids at planetary entry velocities (>7 km/s). The software shall demonstrate robust capturing of the bow shock ahead of a simple cylinder at a variety of flight conditions without requiring adjustment of algorithm parameters, for example, artificial viscosity scale factors. The postshock flow field shall be free of the numerical noise in the entropy field, which is typical of conventional second-order finite volume schemes on triangular grids. Convergence to machine precision shall be demonstrated for all calculations.

A successful Phase II deliverable will mature the Phase I prototype into a product ready for use on mission-relevant engineering problems. Extension to the laminar, multispecies, multitemperature Navier-Stokes equations shall be implemented. Extension to three spatial dimensions using unstructured tetrahedral grids shall be implemented, with efficient multinode parallelization targeting modern high-performance computing (HPC) platforms such as the NASA Pleiades supercomputer. The software shall be demonstrated on a range of planetary entry problems that include at least two of the following destinations: Earth, Mars, Titan, Venus, and Uranus/Neptune. Surface heat flux predictions shall be verified by comparison with NASA's DPLR and/or LAURA simulation codes, and must be free of numerical noise typically observed when using second-order finite volume codes on unstructured tetrahedral grids. Computational

performance, as measured by total time-to-solution for a given heat flux accuracy, shall be characterized and compared to DPLR/LAURA, but no specific performance targets are required.

State of the Art and Critical Gaps:

Multiple academic [4][5][6][7] and NASA [8] groups have demonstrated promising results when using high-order DG/finite element methods (FEMs) to perform steady-state aerothermodynamic analysis at conditions relevant to planetary entry. The bulk of these studies were conducted using structured grids with some degree of shock alignment (though not sufficient alignment to support a second-order finite volume scheme). However, [4] and [5] demonstrate equally accurate results on fully unstructured grids, suggesting that their technologies are capable of meeting the objectives of this scope. An additional shortcoming of current research is that all efforts examine the same 5 km/s flight condition (relatively slow for planetary entry) with simplistic, nonionized flow models. An infusion of resources is needed to mature these promising algorithms into scalable, production-ready software that can be tested across a full entry trajectory with best-practice thermochemical models.

Relevance / Science Traceability:

This scope is directly relevant to NASA space missions in both Human Exploration and Operations Mission Directorate (HEOMD) and Science Mission Directorate (SMD) with an EDL segment. These missions depend on aerothermal CFD to define critical flight environments and would see significant, sustained reductions in cost and time-to-first-solution if an effective unstructured simulation capability is deployed. This scope also has strong crosscutting benefits for tools used by ARMD to simulate airbreathing hypersonic vehicles, which have stringent accuracy requirements similar to those in aerothermodynamics. Finally, this scope aligns with NASA's CFD Vision 2030 Study, which calls for a "much higher degree of automation in all steps of the analysis process" with the ultimate goal of making "mesh generation and adaptation less burdensome and, ultimately, invisible to the CFD process." In order for the aerothermal community to realize these goals, we must eliminate our dependence on manually designed, carefully tailored, block structured grids. This scope is an enabling technology for that transition.

References:

1. Candler, et al.: "Unstructured Grid Approaches for Accurate Aeroheating Simulations." AIAA-2007-3959, 2007.
2. Saunders, et al.: "An Approach to Shock Envelope Grid Tailoring and Its Effect on Reentry Vehicle Solutions." AIAA 2007-0207, 2007.
3. Kleb, et al.: "Sketch-to-Solution: A Case Study in RCS Aerodynamic Interaction." AIAA-2020-067, 2020.
4. Ching, et al.: "Shock Capturing for Discontinuous Galerkin Methods With Application to Predicting Heat Transfer in Hypersonic Flows." Journal of Computational Physics, Issue 376, pp. 54-75, 2019.
5. Gao, et al.: "A Finite Element Solver for Hypersonic Flows in Thermochemical Non-equilibrium, Part II." International Journal of Numerical Methods for Heat & Fluid Flow, Vol. 30 No. 2, pp. 575-606, 2020.

Scope Title:

Efficient Grid Adaption for Unsteady, Multiscale Problems

Scope Description:

The current state of the art for production CFD simulation in EDL is the solution of steady-state problems on fixed computational grids. However, most of the current challenge problems in the discipline are unsteady. Examples include supersonic retropropulsion, where engine plumes exhibit unsteady behavior across a wide range of

timescales [1]; capsule dynamic stability, where the vehicle pitch motion is amplified by the unsteady wake dynamics [2]; and single-event drag modulation, where a high-drag decelerator is separated from the main vehicle at hypersonic speeds [3]. Successful analysis of these phenomena require simulating many seconds of physical time while simultaneously resolving all features of the flow field with high accuracy. Since critical features, for example, shocks, shear layers, etc., will evolve and move through the computational domain over time, current practice requires large, globally refined grids and stringent limitations on simulation time step. This makes these problems computationally infeasible without dedicated access to leadership-class supercomputers.

One promising method to reduce the cost of these simulations is to employ feature-based grid adaption such that the computational grid is only refined in the vicinity of critical flow features. Adaptive techniques, particularly metric-aligned anisotropic adaption [4], have been shown to dramatically reduce computational cost for a wide range of steady-state flow problems, often by as much as an order of magnitude. These techniques have been successfully used to solve large-scale, EDL-relevant problems with high Reynolds number boundary layers by incorporating prismatic near-wall layers [5]. Application of efficient adaptive techniques to unsteady problems is less established, but recent advancements have demonstrated a nearly 100x reduction in compute time required to achieve an equivalent level of space-time accuracy relative to globally refined grids [6].

This scope seeks to accelerate the infusion of cutting-edge algorithms for unsteady grid adaption that promise to radically reduce the time required to simulate unsteady fluid phenomena. A successful solution within this scope would demonstrate an order of magnitude reduction in computational cost without compromising solution accuracy for an unsteady supersonic or hypersonic flow problem relevant to EDL.

Expected TRL or TRL Range at completion of the Project: 2 to 4
Primary Technology Taxonomy:
Level 1: TX 09 Entry, Descent, and Landing
Level 2: TX 09.4 Vehicle Systems
Desired Deliverables of Phase I and Phase II:

- Prototype
- Software

Desired Deliverables Description:

The desired deliverable at the conclusion of Phase I is a prototype software package employing adaptive grid refinement algorithms for the simulation of unsteady, shocked flows in at least two spatial dimensions. An inviscid, perfect gas model is acceptable for Phase I efforts. The prototype software shall be demonstrated on a suitable challenge problem. Suggested challenge problems are prescribed motion of a cylinder relative to the computation domain subject to Mach 6+ supersonic flow or 2D axisymmetric simulation of a shock tube with an initial pressure ratio >50. Other challenge problems of similar complexity are acceptable. The prototype software is not expected to be scalable or performant at this stage.

A successful Phase II deliverable will mature the Phase I prototype into a product ready for use on mission-relevant engineering problems. The code shall be extended to solve the unsteady laminar Navier-Stokes equations in three spatial dimensions with appropriate controls to manage adaption in the boundary layer and the far field, if needed. Extension to reacting, multitemperature gas physics is desired, but not required. The software shall be parallelized to enable simulation of large-scale problems using

modern HPC platforms such as the NASA Pleiades supercomputer. The software shall be demonstrated on a 3D challenge problem such as a single jet supersonic retropropulsion configuration at zero angle of attack; free-to-pitch simulation of the Orion entry capsule at supersonic free-stream conditions; or aerodynamic interaction and separation of multiple spheres in a supersonic free stream. The software shall demonstrate a 10x speedup relative to a nonadaptive, time-marched calculation without significantly degrading simulation accuracy as measured by an appropriate solution metric (average reflectance measurement system (RMS) pressure fluctuation, final capsule pitch angle, etc.).

State of the Art and Critical Gaps:

Multiple academic, government, and commercial software packages exist that implement some form of solution-adaptive mesh refinement. NASA's LAURA and DPLR codes offer simplistic clustering algorithms for structured grids that solve the limited problem of resolving strong bow shocks [7][8]. NASA's FUN3D code implements an advanced metric-based, anisotropic refinement capability that has been demonstrated on large-scale aerospace calculations [7]. However, unsteady solution-adaptive algorithms have yet to be demonstrated for EDL-relevant problems outside of academic research codes. Significant investment is required to implement these algorithms into a production-quality flow solver with the performance and scaling characteristics required to address NASA's requirements for unsteady flow simulation.

Relevance / Science Traceability:

This scope has extremely broad applicability across multiple NASA mission directorates. In particular, ARMD, HEOMD, SMD, and STMD each contend with complex, unsteady flow phenomena that could be more readily analyzed with the aid of the proposed technology: flutter analysis, parachute inflation, fluid slosh, and atmospheric modeling are just a few examples. In EDL specifically, a robust time-space adaption capability would enable simulation of supersonic retropropulsion at Mars using NASA's existing supercomputing assets. Capsule stability could be analyzed in the preliminary design phase, allowing mission designers to utilize low-heritage capsule shapes without adding significant cost or risk to the project. Drag skirt separation could be modeled in detail to reduce risk prior to a technology demonstration mission. The potential benefits of this technology are widespread, making this a critical investment area for the Agency.

References:

1. Korzun, et al.: "Effects of Spatial Resolution on Retropropulsion Aerodynamics in an Atmospheric Environment." AIAA-2020-1749, 2020.
2. Hergert, et al.: "Free Flight Trajectory Simulation of the ADEPT Sounding Rocket Test Using US3D." AIAA-2017-446, 2017.
3. Rollock, et al.: "Analysis of Hypersonic Dynamics During Discrete-Event Drag Modulation for Venus Aerocapture." AIAA-2020-1739.
4. Alauzet, et al.: "A decade of progress on anisotropic mesh adaptation for computational fluid dynamics." Computer Aided Design, Issue 72, pp. 13-39, 2016.
5. Sahni, et al.: "Parallel Anisotropic Mesh Adaptation with Boundary Layers for Automated Viscous Flow Simulations." Engineering With Computers, Issue 33, pp. 767–795, 2016.
6. Alauzet, et al.: "Time-accurate multi-scale anisotropic mesh adaptation for unsteady flows in CFD." J. of Computational Physics, Issue 373, pp. 28-63, 2018.
7. Saunders, et al.: "An Approach to Shock Envelope Grid Tailoring and Its Effect on

Reentry Vehicle Solutions." AIAA 2007-020, 2007.

8. Gnoffo: "A finite-volume, adaptive grid algorithm applied to planetary entry flowfields." AIAA Journal, Volume 21, No. 9, 1983.

9. Bartels, et al.: "FUN3D Grid Refinement and Adaptation Studies for the Ares Launch Vehicle." AIAA-2010-4372, 2010.