Performance Engineering FUN3D at Scale with TAU Commander

John C. Linford, Srinath Vadlamani, Sameer Shende, Allen D. Malony
ParaTools, Inc.
Eugene, OR 97405 U.S.A.
{jlinford, srinathv, sameer, malony}@paratools.com

William Jones, William Kyle Anderson, Eric Nielsen
NASA Langley Research Center
Hampton, VA 23681 U.S.A.
{w.t.jones, william.k.anderson, eric.j.nielsen}@nasa.gov

***Abstract*—FUN3D is an unstructured-grid computational fluid dynamics suite widely used to support major national research and engineering efforts. FUN3D is being applied to analysis and design problems across all the major service branches at the Department of Defense. These applications span the speed range from subsonic to hypersonic flows and include both fixed- and rotary-wing configurations. This poster presents performance profiles of a high Reynolds number simulation of the flow over a wing-body-pylon-nacelle geometry on 14,400 cores of a Cray XC30 at the Navy DSRC. Profiles are gathered via TAU Commander, which implements a new performance engineering methodology to improve user productivity. TAU Commander highlights source code regions that limit scalability through profiling, tracing, and aggregate summary statistics with respect to computational time, memory allocation, and memory access. This analysis approach is being carefully documented to assist other DoD groups in similar performance evaluation activities.**

I. Extended Abstract

NASA Langley Research Center's FUN3D software is an unstructured-grid computational fluid dynamics suite used to tackle complex aerodynamics problems [1]. The toolset enables multidisciplinary capabilities through coupling to variable fidelity models encompassing structural effects, multibody dynamics, acoustics, radiation, optics, propulsion, and ablation. FUN3D provides the world's foremost adjoint-based design capability, enabling formal optimization of time-dependent moving-body simulations involving turbulent flows. The adjoint formulation is also used to perform mathematically-rigorous mesh adaptation and error estimation. FUN3D is widely used to support major national research and engineering efforts at NASA and among groups across U.S. industry, other government agencies, and academia. A past collaboration with the Department of Energy received the prestigious Gordon Bell Prize, which recognizes outstanding achievements in high-performance computing.

FUN3D is being applied to a broad spectrum of analysis and design problems across all the major service branches at the Department of Defense (DoD). These applications span the speed range from subsonic to hypersonic flows and include both fixed- and rotary-wing configurations as well as a diverse array of weapons systems. FUN3D is in routine use across the various DoD Distributed Supercomputing Resource Centers (DSRCs). To accommodate an ever-increasing demand for larger and more complex simulations, the FUN3D development team is partnered with computational experts from ParaTools, Inc. through the DoD/Engility Productivity Enhancement, Technology Transfer and Training Project (PETTT). Through this collaboration, the team is effectively identifying and addressing computational barriers encountered at scale.

The team is using TAU commander and related tools on the DoD DSRC systems to study FUN3D computational performance and to guide optimization efforts. TAU Commander highlights source code regions that limit scalability through profiling, tracing, and aggregate summary statistics with respect to computational time, memory allocation, and memory access patterns. TAU Commander simplifies the TAU Performance System® [2], [3], which has been used to improve the performance and stability of several DoD applications [4]–[7]. TAU Commander eliminates the troubleshooting step inherent in the traditional TAU workflow and avoids invalid TAU configurations by using a structured workflow approach that gives context to a TAU user's actions. A study of 124 workflows demonstrated that using TAU Commander reduces the number of unique steps in the performance workflow by approximately 50% and reduces the number of commands a user must know from approximately eight to exactly one [8]. TAU Commander is installed on several DoD DSRC systems. The analysis approach is being carefully documented to assist other DoD groups in similar performance evaluation activities.

To establish a baseline database of FUN3D performance metrics, the team used TAU Commander to profile a high Reynolds number simulation of the flow over a wing-body-pylon-nacelle geometry. The computation was performed using 600 nodes (14,400 Intel® Xeon® Ivy Bridge cores) on Shepard, a Cray XC30 system located at the Navy DSRC.

Load imbalance in FILL_JACOBIAN is easily visible by comparing against the lowest time in this routine and viewing the function histogram. On average, the routine took 96.6 seconds, of which 28.8 seconds are spent in MPI_Bcast. The broadcast operation takes place in LMPI_CONDITIONAL_STOP, which has dramatically different times across all MPI ranks, caused by load imbalance in solver core routines.

Analyzing data from the 14.4k core run revealed a communication bottleneck in LMPI_CONDITIONAL_STOP. Rank 0 performs a data reduction on behalf of all other ranks, so all other ranks remain idle until Rank 0 finishes the calculation. The severity of the problem grew linearly in the number of MPI ranks. The team implemented a new approach that did not require the data reduction operation. This improvement reduced runtime by up to 33% in core solver routines and is included in the FUN3D 13.0 release.

When writing the solution file to disk, there is a strong correlation between MPI rank and time spent in MPI_Bcast and MPI_Send. MPI_Bcast is called the same number of times with the same message size for all ranks, yet time spent in MPI_Bcast varies dramatically by rank. Rank 0 receives data from all others and writes it to disk, hence N-1 ranks spend the majority of their time in this operation waiting for I/O on Rank 0. This suggests implementing parallel I/O and distributed checkpoints. However, this operation is relatively infrequent once every few hours. Optimization should focus on load balancing the time-stepping PDE solution.

Acknowledgments

This work was developed in part by the User Productivity Enhancement, Technology Transfer and Training (PETTT) Project No. PETA-KY07-001 and by DoE SBIR Grant No. DE-SC0009593. The authors would like to acknowledge the computational resources and PETTT software support from the DoD High Performance Computing Modernization office under Contract No. GS04T09DBC0017.

References

[1] R. T. Biedron, J.-R. Carlson, J. M. Derlaga, P. A. Gnoffo, D. P. Hammond, W. T. Jones, B. Kleb, E. M. Lee-Rausch, E. J. Nielsen, M. A. Park, C. L. Rumsey, J. L. Thomas, and W. A. Wood, "FUN3D manual: 12.9," NASA, Tech. Rep. NASA-TM-2016-219012, 2016.

[2] S. Shende and A. Malony, "The TAU Parallel Performance System," *International Journal of High Performance Computing Applications*, vol. 20, no. 2, pp. 287–311, 2006.

[3] A. Morris, A. Malony, S. Shende, and K. Huck, "Design and Implementation of a Hybrid Parallel Performance Measurement System," in *Proceedings of the International Conference on Parallel Processing (ICPP)*, 09 2010.

[4] S. Shende, A. Malony, J. Linford, A. Wissink, and S. Adamec, "Isolating runtime faults with callstack debugging using tau," *Proc. HPEC 2012 Conference*, 2012.

[5] S. Shende, A. D. Malony, A. Morris, and A. Wissink, "Simplifying memory, I/O, and communication performance assessment using TAU," in *Proceedings of DoD HPCMP UGC Conference*. IEEE Computer Society, 2010.

[6] J. C. Linford, S. Shende, A. D. Malony, and A. Wissink, "Efficient Parallel Runtime Bounds Checking with the TAU Performance System," in *Proceedings of the 2013 IEEE High Performance Extreme Computing Conference (HPEC'13)*, Waltham, MA, 09 2013.

[7] A. Wissink and S. Shende, "Performance Evaluation of the Multi-language Helios Rotorcraft Simulation Software," in *Proceedings of DoD HPCMP UGC Conference*. IEEE Computer Society, 2008.

[8] J. C. Linford, S. Shende, and A. Malony, "Tau commander: An intuitive interface for the tau performance system," ParaTools, Inc., Chicago Operations Office, Argonne, IL, Announcement of USDOE Scientific and Technical Information OSTI ID 1252491, 04 2016.