CASE STUDY

Improving Gas Turbine Performance with HiPSTAR

> "In aviation, the urgent need to reduce emissions is paramount, and this work will be an important milestone in developing the next generation of high-efficiency aircraft."
>
> *–Professor Richard Sandberg, University of Southampton*

Organizations

University of Southampton
Aerodynamics and Flight Mechanics Research Group

The University of Southampton is one of the top 15 research universities in the U.K. Engineering at Southampton is a world-class centre for education, research and enterprise.
www.southampton.ac.uk

GE Global Research

GE Global Research has served as the cornerstone of GE innovation for more than a century. More than 3,000 of the world's brightest scientists, engineers and researchers are working together to deliver technical breakthroughs for GE customers.
www.geglobalresearch.com

Cray® XC30™ Supercomputer "ARCHER"

ARCHER (Advanced Research Computing High End Resource) is the U.K.'s National Supercomputing Service for academic research. The November 2014 Top500 list clocked ARCHER at 2.55 petaflops.

Cray Inc.
901 Fifth Avenue, Suite 1000
Seattle, WA 98164
Tel: 206.701.2000
Fax: 206.701.2500
www.cray.com

Gas turbines (GTs) are the backbone of propulsion- and power-generation systems. Given the very large installed base worldwide, any GT efficiency increase has significant potential to reduce fuel burn and environmental impact. For example, General Electric's installed GT base alone burns $150 billion per year in oil and gas. In power generation, every percentage point increase in combined cycle efficiency would reduce GE's turbine fuel costs by $1.5 billion per year and cut CO_2 emissions per megawatt by 1.5 percent.

Although GT performance has improved considerably, it is becoming increasingly difficult to make further advances with current design tools. High-fidelity computational fluid dynamics (CFD) promises some advancement. Still, resolving all scales of turbulence present in GTs constitutes a formidable computational challenge that can be met only by algorithms that exploit the latest HPC architectures.

Research by GE and the University of Southampton is leveraging recent advances in high-fidelity CFD and ongoing increases in computing power to improve understanding of the unsteady physics that occur in the high-turbulence environment of a GT. The team, led by Richard Sandberg, professor of fluid dynamics and aeroacoustics, in partnership with Vittorio Michelassi, chief engineer of aerodynamics at GE Global Research, is conducting DNS of low- and high-pressure turbines using HiPSTAR (High-Performance Solver for Turbulence and Aeroacoustics Research), a highly accurate structured multiblock compressible fluid dynamics code in curvilinear/cylindrical coordinates, written in Fortran.

HiPSTAR was developed to conduct cutting-edge direct numerical simulations (DNS), the solution of fully nonlinear, time-dependent and three-dimensional Navier-Stokes equations, with no empirical closure assumptions, on today's high performance computing (HPC) systems. More recently, subgrid-scale models have been implemented that also provide large eddy simulation capability, greatly reducing the cost of computation at high Reynolds numbers.

To take full advantage of massively parallel HPC systems, the code was initially parallelized using the Message-Passing Interface (MPI) system and was then extended to hybrid OpenMP/MPI (OMP/MPI) parallelisation through a Cray Centre of Excellence project in Edinburgh. In this project, HiPSTAR was adapted to use hybrid parallelism using the OpenMP application programming interface (API). The project implemented a novel approach to efficiently parallelising compact difference schemes to avoid traditionally used global data transposes or MPI all-to-all calls. GPU parallelisation using OpenACC was implemented recently by John Levesque at the Cray Centre of Excellence at Oak Ridge National Laboratory and Richard Pichler at the University of Southampton.

On "ARCHER" (a Cray® XC30™ supercomputer) most production runs — typically DNS of flows involving low-pressure turbine blades — use between 4,000 and 8,000 cores for grid sizes of up to 500 million nodes. Simulations of high-pressure turbine flows, which require greater grid sizes (approximately 10^9 nodes) have recently been conducted on about 16,000 cores of a Cray® XC40™ system ("Hornet" at HLRS, Stuttgart).

Scaling tests have been performed on ARCHER on a single-block test problem with a total of 1.3×10^9 collocation points. Figure 1 shows the good scaling of the code up to 36,864 cores, for which a real speedup factor over wall clock time on 288 cores (corresponding to 12 nodes of ARCHER, the minimum required to avoid paging) of 102 is achieved versus the ideal factor of 128.

Weak scaling tests were also conducted on ARCHER. In these tests, the number of operations performed by every core, and the number of MPI messages every core had to send/receive, were kept constant with core count. Figure 2 shows the results. The

Figure 1 (left): Strong scaling on "ARCHER." Grid points per unit wall clock time – the computation rate – is used here and in the other figures as a performance metric. A minimum of 12 nodes (288 cores) is required to run this case.

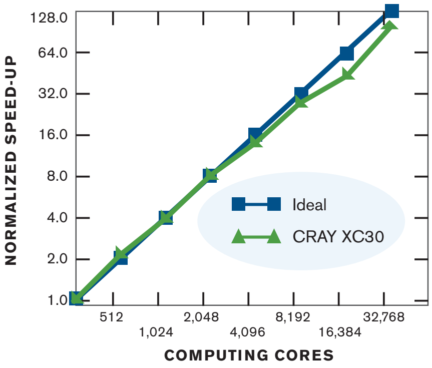
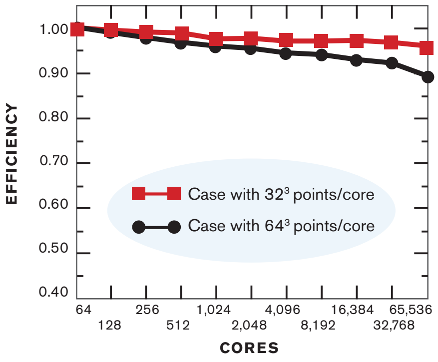


Figure 2 (right): Weak scaling on "ARCHER." Grid points per core kept constant at 270,336.



lowest core count considered was 24, corresponding to a single node of ARCHER (the smallest economic allocation unit); normalising using this result to obtain an efficiency yields a figure of approximately 90 percent at 49,152 cores (corresponding to a simulation with 13 x 10^9 collocation points). In addition to the code's excellent parallel scaling, the performance of the underlying algorithm also appears to be good. A sustained performance of approximately 12.5 percent of peak performance in HiPSTAR's core computation kernels has been measured on a Cray system.

The code was also tested on "Titan," a hybrid GPU/CPU Cray system at the Oak Ridge Leadership Computing Facility. This system is executing production simulations, typically using 512-plus GPUs. Scaling to much greater GPU counts can be demonstrated for more challenging problems, however. Figure 3 shows strong scaling for three different problem sizes using grid points per unit wall clock time as a performance metric. The results are presented with GPU count as the independent variable, but the number of MPI ranks per GPU was also varied. The line showing ideal (linear) scaling is defined separately for each grid size (extrapolating from the result obtained at the lowest GPU count in each case); nonetheless, data points relating to the performance of the larger simulations lie on or near the extrapolated ideal scaling line of the smaller cases. This shows that weak scaling — which is particularly relevant, as the cases to be considered are very ambitious — is excellent. This is further demonstrated in Figure 4, which shows that if the ratio of MPI ranks to GPUs is kept constant, in addition to the grid size to GPU ratio, wall clock time per time step remains constant to over 8,000 GPUs.

Figure 3 (left): Strong scaling on "Titan," in terms of computation rate.

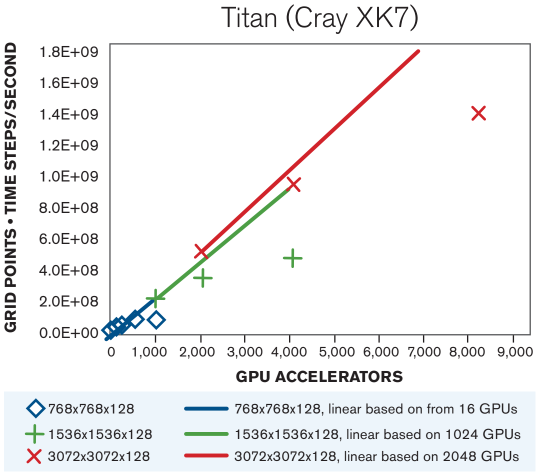
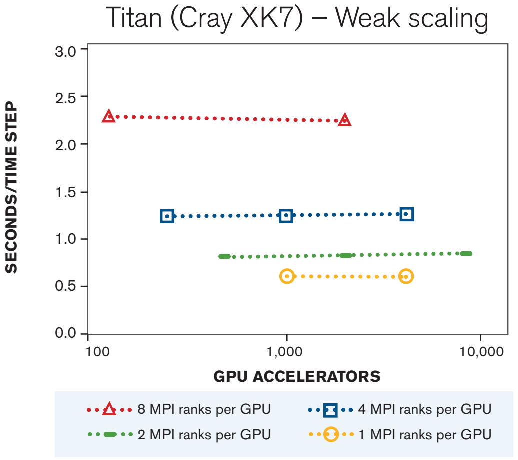


Figure 4 (right): Grid points per GPU kept constant at 73,728 in all cases.



Figures 5 and 6 show the results. The simulation of the high-pressure turbine was conducted by Dr. Andrew Wheeler on a grid with more than 1 billion points.

Figure 5 (left): DNS of low-pressure turbine, conducted on "Titan" and "ARCHER."

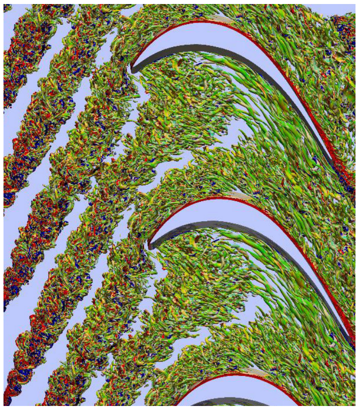
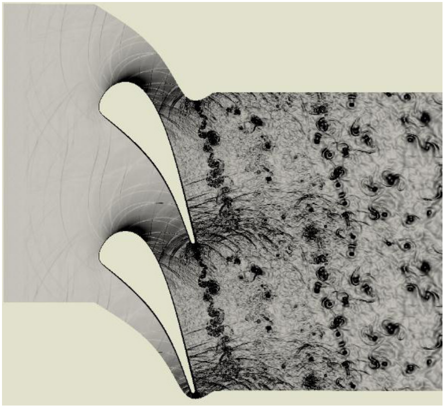


Figure 6 (right): DNS of high-pressure turbine, conducted on "Hermit," "Hornet" and "ARCHER."



Research partner GE will be able to advance gas turbine technology, with particular focus on aviation engines, by analyzing the results of several unprecedented first-principle calculations performed in engine-relevant operating conditions.