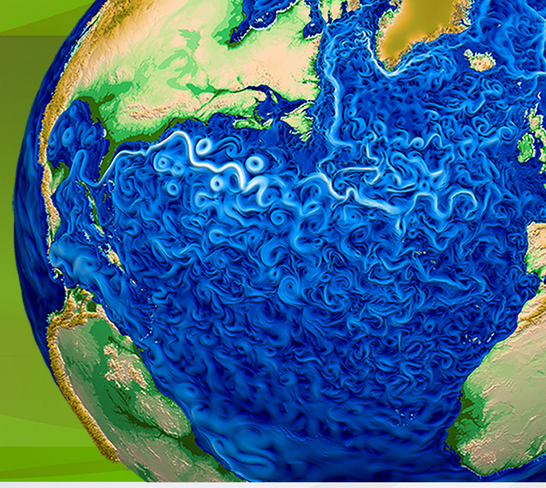
R: Progress on porting the Community Atmosphere Model - Spectral Element (CAM-SE) to the GPU-CPU hybrid architectures

Matthew Norman (ORNL) ; Irina Demeshko (SNL)



Objective

Port ACME to Efficiently Utilize Current and Emerging Accelerated Computing

The tracer transport routines of CAM-SE have been ported to use GPUs. During ACME's first year, this port has been improved substantially for performance and a wider range of science targets. To achieve ACME's ambitious scientific goals, it is crucial that the codebase efficiently utilize modern architectures so that it can compete for large allocations on LCF resources.

Improvements to the Existing Port

Description of the Exising Port

- The CAM-SE tracer transport routines were ported during Oak Ridge Leadership Computing Facility's (OLCF's) previous Center for Accelerated Application Readiness (CAAR) effort;
- Using CUDA FORTRAN, kernels were ported to the GPU, targeting Titan's K20x GPUs;
- A speed-up of 2x over a single AMD Interlagos CPU was achieved;
- A new vertical remap algorithm was coded, and element packing / unpacking was overlapped with MPI and PCI-express data transfers.

Improvements during ACME's First Year

- The ability to sub-cycle over vertical remaps was added;
- Number of vertical levels were previously limited. Now, any number of vertical levels can be used, and CUDA blocksizes are determined by the developer;
- Several crucial variables were placed into fast GPU cache memory;
- The efficiency of overlapping element edge packing and unpacking was improved by removing barriers and synchronizations where unnecessary;
- The pre-computations for tracer transport routines were moved to the GPU;
- Monotone limiter has been ported to GPU with CUDA FORTRAN and shows up to 3x speedup over a single CPU implementation.

Current Performance

- Performance depends on the work per node because MPI begins to dominate.
- Assuming 50 tracers and 30 vertical levels, **CAM-SE dycore only,** GPU speed-up is now 1.69x and 1.34x at 128 and 64 elements per node, respectively.

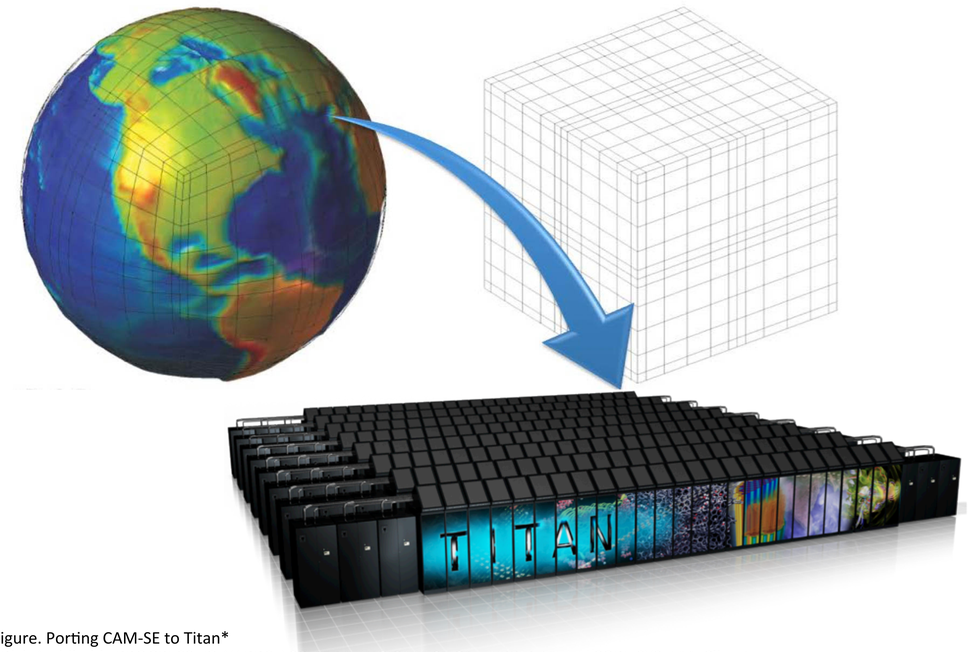


Figure. Porting CAM-SE to Titan*
*Rick Archibald , PROGRESS TOWARDS ACCELERATING CAM-SE ON HYBRID MULTI-CORE SYSTEM, Climate and Earth System Modeling PI Meeting Model Interoperability, 2011

Future Work

Completing the Monotone Limiter Porting

- The monotone limiter itself is ported with machine-precision similarity;
- The supporting routines are being ported with OpenACC.

Improving Inter-comparison Capabilities

- Other ACME work is developing ensemble-based testing that looks at globally, annually averaged statistics as well as transient statistics such as extremes;
- Need to ensure the port is statistically indistinguishable from CPU results within the range given by CPU compiler flag changes.

OpenACC Porting

- OpenACC is a directives-based approach, which requires significantly less work for porting, resulting in very similar code;
- The existing port will be moved to OpenACC for enhanced portability and code maintainability.

Expanding the Breadth of the Port

- The next targets for porting are CAM-SE dynamics, then MPAS-O tracers.

This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.
Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO SAND2015-3387 C

ACME Accelerated Climate Modeling for Energy

For additional information, contact:
Matthew Norman
Computational Climate Scientist

Oak Ridge National Laboratory
(865) 576-1757
normanmr@ornl.gov

climatemodeling.science.energy.gov/acme

U.S. DEPARTMENT OF ENERGY