Geophysical Research Abstracts
Vol. 17, EGU2015-8550, 2015
EGU General Assembly 2015

Toward a more efficient and scalable checkpoint/restart mechanism in the Community Atmosphere Model

Valentine Anantharaj

Oak Ridge National Laboratory, National Center for Computational Sciences, Oak Ridge, United States (anantharajvg@ornl.gov)

The number of cores (both CPU as well as accelerator) in large-scale systems has been increasing rapidly over the past several years. In 2008, there were only 5 systems in the Top500 list that had over 100,000 total cores (including accelerator cores) whereas the number of system with such capability has jumped to 31 in Nov 2014. This growth however has also increased the risk of hardware failure rates, necessitating the implementation of fault tolerance mechanism in applications. The checkpoint and restart (C/R) approach is commonly used to save the state of the application and restart at a later time either after failure or to continue execution of experiments.

The implementation of an efficient C/R mechanism will make it more affordable to output the necessary C/R files more frequently. The availability of larger systems (more nodes, memory and cores) has also facilitated the scaling of applications. Nowadays, it is more common to conduct coupled global climate simulation experiments at 1 deg horizontal resolution (atmosphere), often requiring about 10^3 cores. At the same time, a few climate modeling teams that have access to a dedicated cluster and/or large scale systems are involved in modeling experiments at 0.25 deg horizontal resolution (atmosphere) and 0.1 deg resolution for the ocean. These ultrascale configurations require the order of 10^4 to 10^5 cores. It is not only necessary for the numerical algorithms to scale efficiently but the input/output (IO) mechanism must also scale accordingly.

An ongoing series of ultrascale climate simulations, using the Titan supercomputer at the Oak Ridge Leadership Computing Facility (ORNL), is based on the spectral element dynamical core of the Community Atmosphere Model (CAM-SE), which is a component of the Community Earth System Model and the DOE Accelerated Climate Model for Energy (ACME). The CAM-SE dynamical core for a 0.25 deg configuration has been shown to scale efficiently across 100,000 cpu cores. At this scale, there is an increased risk that the simulation could be terminated due to hardware failures, resulting in a loss that could be as high as 10^5 - 10^6 titan core hours. Increasing the frequency of the output of C/R files could mitigate this loss but at the cost of additional C/R overhead. We are testing a more efficient C/R mechanism in CAM-SE. Our early implementation has demonstrated a nearly 3X performance improvement for a 1 deg CAM-SE (with CAM5 physics and MOZART chemistry) configuration using nearly 10^3 cores. We are in the process of scaling our implementation to 10^5 cores. This would allow us to run ultra scale simulations with more sophisticated physics and chemistry options while making better utilization of resources.