Experiments with HPC-as-a-Cloud: On-Demand ENK-Based Oil Reservoir History Matching on the Blue Gene/P

A Demonstration at the Super Computing Conference, November 2010

PI: Manish Parashar
Student: Moustafa AbdelBaky
IBM: Kirk Jordan

Cloud computing had emerged as a paradigm for providing on-demand access to resources. However, public clouds have largely proven to be ineffective for many HPC applications. Furthermore, private clouds built on grid-based solutions lack the high-speed interconnects, scalability, and manageability that supercomputers provide. Clearly, it would be desirable to provide supercomputing resources as easily accessible on-demand clouds services that can run HPC applications. In order to address this need, Rutgers and IBM are collaborating to develop and deploy a programming system that provides the abstraction of HPC-as-a-Service using supercomputers as the back-end, and can enable many-task applications to access these services easily from the desktop and even mobile systems.

As part of this effort we used an oil reservoir history matching application to demonstrate the effectiveness and utility of this programming system at the Super Computing Conference 2010, New Orleans, Louisiana, on November 13-10, 2010 (http://sc10.supercomputing.org/).

The oil reservoir history matching application workflow belongs to a strategically important class of applications that run multiple realizations with ensemble filtering to obtain near optimal solutions in situations where there is great uncertainty. In this application, multiple instances of an oil reservoir simulation are run simultaneously and the results are filtered through an Ensemble Kalman Filter (EnKF).

Specifically, for this demo, we designed a generic framework that allows users to easily deploy, run, monitor and steer applications running on IBM Blue Gene/P from a personal computer. The framework essentially converts Blue Gene/P to a private cloud, allows for public cloud integration, with no impact on performance, while increasing productivity. We provided users the capability to run their application on the private cloud, public cloud or both, thus taking advantage of both infrastructures. Finally we gave users the ability to monitor and steer applications using desktop and mobile devices such as the iPad.

Together, this integrated solution provides a more effective and convenient abstraction for using supercomputing resources for real-work problems, it can also potentially improve the productivity of the scientist and overall time-to-science.

Oil-reservoir Data Assimilation