DATA EXPLOSION CREATES SEISMIC SHIFT IN HPC WORKLOADS FOR OIL AND GAS INDUSTRY

Michael Feldman

White paper

October 2014

MARKET DYNAMICS

Oil and gas exploration is one of the riskiest and most expensive of all commercial endeavors. With tens millions of dollars in the balance for the construction of each offshore platform and hundreds of millions of dollars devoted to leasing offshore leases, these companies rely on information that is both timely and reliable. Much of that information, and the decisions that follow, is based on results of seismic analytics workloads.

Today, finding and exploiting oil and gas deposits is enabled by high performance computing (HPC), which is tasked to provide the critical technology for processing and interpreting seismic surveys. Exploring for these oil and gas reservoirs is done by generating enormous amounts of seismic data and running that data through the computational ringer to reveal the most promising locations to drill. As is the case with many HPC applications, speed and precision are paramount – speed because time to solution is critical to business decisions in energy exploration, and precision because inaccurate information leads poor return on investment. Drilling an unproductive offshore wastes hundreds of millions of dollars, while the cost of acquiring leases can run over a billion dollars.

Technical challenges abound. Looking for oil and gas has resulted in a data deluge for Geophysicists. Geophysicists, in their quest for more accurate imagery of underground reservoirs, have successively moved from 2D, to 3D, to 4D imaging. The demand for increased resolution is also magnifying data processing needs. For example, in years past, thousands of pre-stack traces were gathered for every square mile; today as many as a million traces are used. As a result, data sets in this domain are among the largest in the industry, reaching into multiple petabytes, and which must be funneled from the storage subsystem to the compute nodes at tens of gigabytes per second.

A Seismic Workflow Challenge

Data sets of this size and the need to process them quickly compels optimal use of compute and storage resources. Acquisition, processing, interpretation and visualization of seismic images represent classic “big data” workflows, mixing compute- and data-intensive jobs, and

where the output from one job becomes the input for another. The asset team must manage this entire workflow on the underlying HPC infrastructure, and do so with the most optimal use of resources. Misused hardware resources are not only wasteful from a capital expenditure point of view, but a waste of computational opportunity from the standpoint of application throughput. Regardless of costs, there are also absolute limitations in power, cooling, and space to contend with in a given datacenter.

"Unused hardware resources are not only wasteful from a capital expenditure point of view, but a waste of computational opportunity from the standpoint of application throughput.

To complicate matters further, as new seismic data is acquired, computational priorities can shift in real time to chase down more promising prospects, making workflow management even more challenging. Even if the Geophysicists have access to elastic resources (e.g., a private or public cloud or a secondary datacenter), the complexity of orchestrating these workloads across the different compute and storage silos necessitates either a high skilled (and very busy) system administrator or a workload manager that can handle the inherent complexities of resource allocation and job scheduling across heterogeneous infrastructure. As data sets grow in size, time constraints become more demanding, and additional flexibility is desired, the manual solution to workload management becomes ever more problematic.

OPPORTUNITY FOR ADAPTIVE COMPUTING

Adaptive Computing, with its focus of intelligent workload management, has perfected advanced resource management tools for HPC users for more than 10 years. According to our latest HPC Site Census Survey, the company is the top supplier of Intelligent Workload Management solutions, with 44 percent of sites citing Adaptive as a supplier[1].

The company's Moab-branded products, which can operate across traditional clusters, private clouds, and public clouds, have been designed to provide automated job scheduling and resource management for many types of complex workflows, such as those found in oil and gas. They are geared for complex workloads in both technical and business environments, where application performance and throughput are paramount to the organization's mission.

Moab Brings Big Workflow to Workload Management

[1] HPC User Site Census: Middleware, Intersect360 Research, April 2014

Intersect360 Research
P.O. Box 60296 | Sunnyvale, CA 94088 | Tel. [888] 256-0124
www.Intersect360.com | info@Intersect360.com

Adaptive Computing has continually refined its product set, most recently incorporating the notion of "Big Workflow," a term coined by the company that provides a model for integrating data-intensive and compute-intensive workflows under a single workload management scheme[2]. Essentially, the model draws together HPC simulations with big data analytics workloads so that users can manage them in a unified manner.

In the latest Adaptive Computing version of the Moab HPC Suite (Moab 8.0) and the Moab Cloud Suite, Big Workflow services deliver dynamic scheduling, provisioning and management of multiple applications across HPC, traditional cloud and big data environments. By doing so, Adaptive Computing is able to automate much of the workflow management, which otherwise would need to be performed manually with scripts and administrator intervention. This streamlines the complexities of job management and increases application throughput, which, in turn, shortens the time to discovery of productive oil and gas. As a consequence oil and gas companies can maximize their investments in well drilling and offshore leasing.

Part of what the Big Workflow model provides is consolidating the underlying resources – bare metal clusters, virtualized datacenter servers, and private or public clouds, and optimizing their utilization. Thus Adaptive Computing is able to maximize use of hardware across multiple environments. It also streamlines analytics workflows by providing intelligent data transfers and data staging. For oil and gas, this is particularly useful given the often siloed infrastructure of compute and storage that has grown up around these evolving applications.

INTERSECT360 RESEARCH ANALYSIS

Although seismic imaging and interpretation are mature applications, the escalating costs of drilling and leasing is driving oil and gas firms to optimize their computation workflows for both speed and performance. And given the natural limitations of procuring, deploying, and operating HPC infrastructure in commercial datacenters, managing those workflows optimally is critical for the oil and gas business.

To that end, the industry looks to solutions that can deliver the highest application throughput on its HPC infrastructure in order to maximize their return on investment. The complexities and performance demands of seismic workloads, the diverse nature of HPC infrastructure that these applications run on, and the ever-growing datasets these applications must ingest, represent additional challenges in this domain.

[2] Big Workflow: More than Just Intelligent Workload Management for Big Data, Intersect360 Research, February 2014

Intersect360 Research
P.O. Box 60296 | Sunnyvale, CA 94088 | Tel. [888] 256-0124
www.Intersect360.com | info@Intersect360.com

Today only a limited number of vendors offer the kind of advanced workload manager that can meet the demands of these complex workflows. Adaptive Computing is one such vendor, and, with its Moab HPC Suite and Moab Cloud Suite, offers a unique product set for this industry. In particular, the Big Workflow model used by the Adaptive Computing software is able to integrate the seismic simulations and analytics jobs into a unified workflow that can deliver optimal utilization of the underlying compute and storage hardware. Further, it provides the flexibility to automatically scale out these workloads across all of your HPC environments.

The oil and gas industry's critical need for complex workflow management fits with Adaptive Computing's product strength. The company has a 13-year track record of delivering intelligent workload management for HPC customers with widely varying application profiles and hardware platforms, and for commercial organizations that must rely on proven technology for critical business needs, Adaptive has demonstrated itself to be a reliable partner.

Intersect360 Research
P.O. Box 60296 | Sunnyvale, CA 94088 | Tel. [888] 256-0124
www.Intersect360.com | info@Intersect360.com