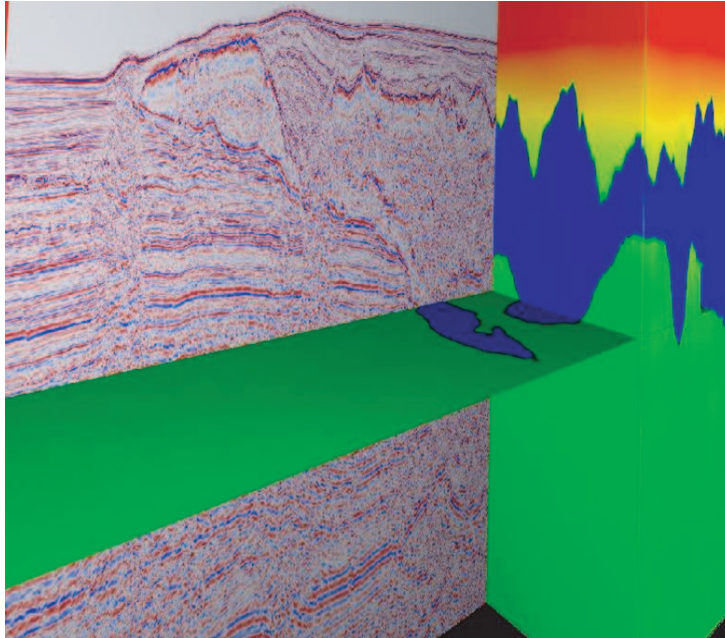IBM

Deep Computing Capacity on Demand: Upstream Petroleum Industry

Highlights

■ *Addresses peak workload demands for high performance computing*

■ *Offers flexible pricing, allowing customers to pay only for the capacity reserved and helps avoid up-front capital investment*

■ *Helps create a flexible, scalable and responsive IT infrastructure*

IBM Deep Computing

IBM Deep Computing can address the needs of customers who require powerful solutions to address their most complex and challenging business needs. These customers include scientists, engineers, knowledge workers, and other professionals whose core business or research relies on high performance computing (HPC) for deep levels of computation, data access, visualization, and communications. Deep Computing is more

than traditional HPC hardware, software and services. It also embraces emerging strategic technologies, research initiatives, open standards, industry and customer partnerships, and industry expertise. IBM Deep Computing takes a comprehensive approach to helping our customers gain insight and transform their business and organizations. Industries and business segments include Life Sciences, Digital Media, Financial Services, Petroleum, Electronics, Automotive, Aerospace, Government and Higher Education.

Using HPC in the Upstream Petroleum industry

The upstream petroleum industry is ideally positioned to take advantage of an on demand computing model. To support processes such as seismic interpretation and reservoir modeling, petroleum companies often deploy high performance computing (HPC) infrastructures. HPC capability and availability can help companies perform petroleum exploration, appraisal and production. However, because fluctuating oil prices can

drive project-and asset-based activity in this industry, building HPC capacity to meet peak demand has not been an economical option.

IBM's offerings for high performance computing on demand can help upstream petroleum companies meet peak demands for HPC capacity, while providing a better way to manage capital and operating costs. Petroleum companies can achieve attractive price/performance for compute-intensive applications and processing of numerous large-data files. These organizations also can reduce the administrative costs associated with managing distributed systems and data, allowing them to focus on their core business activities rather than on IT cost recovery or infrastructure complexity.

Meeting upstream petroleum business challenges

Oil companies making field investment decisions face increasing requirements for HPC capacity in order to process, manage and use ever-expanding quantities of data. These requirements are exacerbated by the need to improve productivity (by reviewing and analyzing a larger portfolio of assets in a shorter time in response to business needs); to build better, more accurate reservoir models (through, for example, parametric studies and compute-intensive algorithms that yield finer granularity and hence more accurate results); to reduce time spent on value-adding decisions (such as how to optimally extract oil and gas from producing fields) in a near real-time environment; and to support new fields and developments in remote and often environmentally hostile regions. Such activities can be critical to the continued success of oil companies as they endeavor to use scarce and valuable resources while containing costs and mitigating risks, with the sole aim of producing oil and gas efficiently, effectively, safely and responsibly.

For service and supply organizations that support oil companies in their quest to efficiently and cost-effectively use oil and gas reserves, the upstream petroleum industry has become increasingly competitive. Short- and long-term fluctuations in oil prices, geopolitical influences, ongoing licensing rounds and other factors can directly affect the demands for seismic acquisition, data processing and accurate reservoir modeling and simulation, and the ability to perform these activities can be influenced by HPC capacity. Planning against such fluctuating demand becomes inherently difficult. Companies must provide services quickly, efficiently and cost-effectively. The successful services companies are those that meet the changing needs of the oil companies while helping to control costs and improve profit margins. Using IBM cluster services to obtain HPC capacity on demand can contribute to this success.

IBM can help oil companies and service and supply organizations meet these business challenges by helping to maximize capacity flexibility while minimizing capital investment. IBM offerings for HPC on

demand are based on state-of-the-art cluster technologies and services, and leverage IBM's industry and technology expertise.

Examining HPC requirements for the upstream petroleum industry

Providing HPC resources on demand requires an infrastructure of high-speed processors and storage that can provide measurable compute capacity in a flexible, scalable and highly secure manner. Users require seamless access to these resources, regardless of where the physical systems may reside. An HPC infrastructure for an upstream petroleum company has the following objectives:

Scalability: Applications and databases often exceed the capacity of even the largest single systems and companies quickly outgrow their systems.

Flexibility: While they focus on controlling costs, oil companies and service and supply organizations need increased flexibility in a rapidly changing environment.

Significant processing capacity: As processing requirements continue to grow, more and more companies are finding that they can benefit from using clustered servers and Grid computing. Clusters link separate nodes to create a single, powerful, scalable system, and Grid computing increases overall compute resource sharing and utilization.

Flexible financial and delivery models: Petroleum companies can benefit from balancing their fixed and variable IT infrastructure and operations costs and choosing between onsite owned/operated versus offsite hosted IT solutions.

Clustering and Grid computing, combined with a wide range of server technologies, are designed to provide:

- *High availability with failover protection*
- *Scalability with minimal downtime*
- *Large-capacity data volumes*
- *Capability to handle peaks in workload, permitting flexible responses to changing business requirements*

Providing high performance computing on demand

Fundamentally, providing HPC capacity on demand is a simple concept. Instead of purchasing computers to meet peak demands, companies can procure sufficient hardware for average demand and then contract remote processing power to help meet peak loads, or offload all processing to a remote facility. In practice this process can be more complex. However, IBM can help shield customers from much of this complexity— one of the primary advantages of a hosted on demand environment.

High performance computing can have many different requirements, and purchasing sufficient hardware can become expensive. Some problems are computationally parallel, and a large number of servers working in parallel can solve the problem quickly and efficiently. Other problems require ultrafast interprocess communication or a single large memory model. A hosted environment can supply

access to these resources, and instead of paying for equipment that may be underutilized, customers simply pay for the cycles and compute or storage resources that they reserve for a specified period of time.

IBM provides a scalable, highly secure and extensible environment designed to handle peak workloads. This solution combines the best of IBM technologies and open standards with on demand computing, providing an infrastructure that enables customers to help meet their goals of increased scalability, flexibility, and processing capacity.

Building a world-class infrastructure to support upstream petroleum companies

IBM has a comprehensive approach to deliver on demand computing for petroleum companies. IBM's advanced clustering capabilities and services can help reduce the time, effort and expense required to

manage and analyze seismic and reservoir data, which can accelerate decision making.

The HPC infrastructure is based on the IBM @server™ Cluster 1350, an innovative Linux-based cluster that provides a robust, highly scalable compute facility centrally managed and controlled by IBM. The Cluster 1350 combines the power of IBM @server xSeries® Intel Xeon™ processor 32-bit and IBM @server AMD Opteron™ processor 32-bit/64-bit compute servers with IBM Cluster Systems Management (CSM) software, IBM TotalStorage® products, and leading third-party networking components to enable powerful, flexible solutions for high performance computing. IBM also offers clusters of IBM @server pSeries® POWER™ 64-bit compute servers running AIX 5L™ or Linux. IBM plans to offer other models, platforms and technologies over time.

Customers can access virtual dedicated Cluster 1350 cluster capacity based on the xSeries 335 2-way Intel Xeon processor-based server or IBM @server 325 2-way AMD Opteron processor-based server. The xSeries 345 or IBM @server 325 is used as a management node for centralized control of the system. Each virtual cluster has its own dedicated management node. Customers can also access virtual cluster capacity based on the pSeries 655 4-way server with a pSeries management node. Optional Fibre Channel or SCSI external disk storage capacity can be provided. Customers also have access to an IBM 3590, LTO, or DLT serially shared tape server for loading data and performing backups.

Customers have full control of the compute and storage resources within their assigned environment, and each virtual cluster has a dedicated cluster and management network. A robust networking infrastructure is designed to allow customer

data and applications to be highly available and secure, and a virtual private network (VPN) service provides remote access to the facility. Customers pay for their assigned compute and storage capacity for the length of time that they use it.

Advanced systems management and Grid-enabling software

IBM Cluster Systems Management (CSM) for Linux software provides robust capabilities for resource monitoring, automated operations, remote hardware control, distributed command execution, configuration file management and parallel network installation. Access to the system is accomplished through a remote shell interface. Customers may use CSM directly to perform remote management tasks, install and use their own management tools, or request comprehensive management and monitoring services from IBM.

The IBM General Parallel File System (GPFS) for Linux and AIX® is a high performance scalable file system—ideal for petroleum companies that manage extremely large data files. GPFS can provide an enhanced layer of scalability, availability and performance. Optional dynamic job scheduling and workload management tools are also planned to be available to optimize cluster resources and help increase job throughput.

Customers supply the additional software required to run their workloads; this software may include in-house, open source, or IBM and third party applications and tools. Customers can use Grid software and services, acquired separately from IBM or a third party, to develop new applications or run existing applications in a Grid environment. The open source Globus Toolkit™ is designed to provide the underlying elements for Grid security, communication, information infrastructure, resource management and portability. The IBM Grid Toolbox

for Linux and AIX is an integrated set of tools and software based on the Globus Toolkit that facilitates the building of Grids and development of Grid applications. In addition to running production workloads, customers may run benchmarks or perform other types of analysis.

Services for delivering HPC on demand

IBM offers a variety of services to support IBM-supplied hardware, software, and networking components and infrastructure. Base services include the facility, provisioning and deployment of resources, security, virtual private network, monitoring, maintenance, and help desk support. Customers can benefit from a range of fee-based custom services, such as enhanced VPN bandwidth, software deployment and customization, advanced monitoring and management, storage management and backup, Grid-related services, and more.

High Performance Computing

Delivering dynamic HPC resources

IBM offers dynamic HPC resources on demand that are designed to match customer requirements for scalability and capacity. The on demand computing model is economical, enabling customers to enjoy flexible pricing and terms and helping them to realize lower operational and infrastructure fixed costs. This computing model also can help reduce infrastructure complexity, IT training and staffing and ongoing maintenance and upgrades. Customers can focus on their core engineering issues and be insulated from rapid information technology shifts.

IBM is extending its leadership in high performance computing to the on demand environment. By combining leading IBM @server technologies, advanced cluster and systems management capabilities, growth and innovation in Grid and autonomic computing, and demonstrated industry expertise, IBM is delivering solutions designed to enable petroleum customers to conduct their businesses more efficiently and effectively in a highly competitive market.

Figure A. Possible technical components: Deep Computing Capacity on Demand for the Upstream Petroleum Industry

Component	Features/Benefits
HARDWARE	
IBM @server Cluster 1350	• Combines IBM @server servers, IBM TotalStorage storage and leading third-party networking components to create powerful, flexible solutions for HPC and commercial application environments • Uses xSeries 335 (1U) and xSeries 345 Intel Xeon 32-bit dual processor nodes • Uses BladeCenter Intel Xeon 32-bit dual processor nodes • Uses IBM @server 325 (1U) AMD Opteron 32-bit/64-bit dual processor nodes • Runs the Linux operating system standard • Runs the Windows operating system by exception • Uses CSM for central management
AMD Opteron Technology	• Runs 32-bit and 64-bit applications simultaneously • Easy migration to 64-bit computing capability with software investment protection • Extreme price/performance design—high performance at affordable pricing • Runs Linux and Windows
IBM pSeries Technology	• Features POWER 64-bit processors in an ultra-dense packaged server for high performance computing • Includes the AIX 5L operating system and a rich suite of cluster software
IBM TotalStorage	• FAStT Fibre Channel and SCSI disk options • 3590, LTO, and DLT tape server and cartridges options
SOFTWARE	
Operating Systems	• Linux and Windows provided by customer • AIX provided by IBM
IBM Cluster Systems Management (CSM)	• Provides robust, powerful management from a central point of control • Simplifies administrative tasks and may reduce life-cycle costs • Offers a highly reliable infrastructure and event monitoring • Provides software installation and updates, remote hardware control, distributed command execution, configuration file management, and diagnostics
IBM General Parallel File System (GPFS) (optional)	• Provides shared access to files across multiple disk drives on multiple nodes • Provides a common file system abstraction for data shared among multiple nodes • Allows applications to easily access files using standard POSIX (Portable Operating System Interface for UNIX®) file system interfaces • Enables parallel applications to simultaneously access either the same or different files • Provides high availability through automatic recovery from node and disk failures
Job scheduling (optional) (planned availability)	• Provides dynamic job scheduling and workload management • Designed to optimize cluster resources and increase job throughput • Works with CSM to facilitate management of cluster resources
SERVICES	
IBM Global Services	• Support Deep Computing Capacity on Demand with a full range of services for solution customization • Provide Grid enablement services, including Grid Innovation Workshops, Grid Pilot Implementation Services, and Grid Rollout Implementation Services

For more information

To learn more about IBM Deep
Computing Capacity on Demand for
the Upstream Petroleum industry,
contact your IBM representative.

AMERICAS

Alan McCarter

e-mail: armccar@us.ibm.com

ASIA PACIFIC

Sinisa Nikolic

e-mail: 9sinisa@my.ibm.com

EUROPE/MIDDLE EAST/AFRICA

Ian Green

e-mail: ian_green@uk.ibm.com