HRG Assessment

Vblock™ Unified Data Solution for Oil & Gas

It is 2014 and oil and gas industry leaders are engaged in making the digital oil field a reality. While this is a good first step in improving Up Stream production and efficiencies HRG sees this as only the necessary first step towards a fully integrated digital oil and gas supply chain comprising Up Stream, Mid-Stream, and Down Stream operations. HRG's view of the future for the energy sector (oil, gas and electricity) is one based on the establishment of a fully integrated, manageable and controllable supply chain that spans from development to delivery. This future view is in keeping the current work being done by General Electric in the area of manufacturing process digitization, control, management, and automation. GE uses the term the Industrial Internet to describe this multi-year initiative – they are currently in year three of this initiative focused on supply / value chain automation and facilitation to deliver customer desired outcomes. VCE's Vblock™ Unified Data Solution for Oil and Gas represents a key enabling technology platform for the evolving and fully integrated digital oil and gas supply chain.

The oil and gas industry is extremely competitive and highly regulated environment. Characterized by limited reserves of natural resources, fluctuating demand, and price volatility oil and gas companies need to better match production to demand, optimize costs, and reduce the impact of environmental risks. Finding and developing new oil and gas resources requires collecting and analyzing growing volumes of data generated by increasingly sophisticated engineering technology such as downhole multiphase sensors, and measurement while drilling (MWD) applications. Additional real-time drilling data is delivered by oil field technologies like fiber-optic pressure and temperature gauges that provide sustained downhole profiling data captured through MWD and remotely steerable downhole tools. With remote data collecting sensors and programmable logic controllers (PLCs) that enable real-time monitoring, analysis, management and control using distributed sensors high-speed communications and data mining techniques oil and gas companies can monitor and fine-tune remote drilling operations.

Unified Oil and Gas Big Data brings together data that has not previously been analyzed in this aggregated form. Analytics applied to this data drives new insights and provides personnel with access to searchable institutional knowledge that helps compensate for limited expert staffing and the growing shortage of experienced personnel.

Unified oil and gas supply chain data analytics provides insights into logistics, supply chain, marketing, and trading to improve and enable management of the entire energy production and delivery supply chain from first oil to delivery. Predictive analytics combining current and real-time data with historical data helps oil and gas suppliers predict and eliminate potential equipment and well failures and eliminate human error. The supply chain wide collection and archiving of vibration, strain, pressure, flow, temperature, and other critical data helps improve the accuracy of predictive analytics results.

Outcomes

By putting the right data, at the right time in the hands of decision-makers oil and gas suppliers will realize the following outcomes:

- Improve new well delivery
- Improve drilling accuracy and safety
- Enhance oil recovery from unconventional sources
- Improve safety and prevent risks
- Increase production and efficiency
- Reduce time to first oil
- Prevent catastrophic equipment failure
- Benefit from more accurate forecasting

Oil and gas companies use thousands of sensors installed in wells and throughout surface facilities to provide continuous data collecting and real-time monitoring of equipment, wells, and environmental conditions. Aside from real-time streaming data from well heads, drilling equipment, and sensors much of the data that providers collect is unstructured such as emails, documents, spreadsheets, images, live multi-day multimedia (such as: video feeds, audio and voice recordings). The BMMsoft EDMT component (described in detail later in this document) of VCE's VCAP Unified Big Data Solution was specifically designed to ingest this type of data

The increased use of remote monitoring devices to collect and transmit data (regardless of location) contributes significantly to the growth in Oil and Gas related Big Data. This telemetric data enables engineers to monitor and track wells, reservoirs, and facilities and enables early intervention for alarm conditions as they are detected in order to avoid catastrophic system failures.

VCE's VCAP (Vblock™ Converged Analytics Platform) Oil and Gas solution can be used to improve the interoperability of dissimilar geographically dispersed systems as a result of EDMT's unique ingest and metadata index capabilities. Industry requirements, as well as, production, management and delivery mechanisms continue to evolve keeping pace with new remote monitoring and management technologies and the ability to collect, analyze and report on real-time data. In HRG's opinion EDMT as implemented in VCE's VCAP Solution is fully scalable and able to meet much of the oil and gas sectors evolving unified data and analytics requirements for the foreseeable future.

Unified Data Strategy

What is Unified Data? Unified Data is the combination of Big Data and Enterprise data (sometimes called "Little Data") that has been ingested, indexed, and stored in a searchable and query-able enterprise-class database. Unified Data can then be can cross-searched and cross-analyzed using currently available data analysis tools and techniques, as well as, SQL, "R" and other in-database analytics tools and methods.

Things to consider when developing a Unified Big Data Strategy are:

- ➢ Data Variety – structured, unstructured, and real time
- ➢ Data Volume – trillions of transactional records, "Internet of Things", and files
- ➢ Data Velocity – the ability to ingest and store the volumes of data at the rate required
- ➢ Choice of database –data should be stored in an ACID-compliant repository
- ➢ Running analytics and searches on (and across) structured and unstructured data from multiple sources

Data Sources

Oil and gas industry professionals are becoming increasingly aware of the benefits of having access to Unified Big Data as they strive to improve outcomes, decrease cost, increase efficiency, and deliver a reasonable return on investment (profit). With the introduction of VCE's VCAP Solution no data, regardless of type, location, or format is beyond integration. As Oil and Gas suppliers continue to collect increasing volumes of data at increasing velocities from their operations due to the proliferation of real time sensors, PLCs, and other sources keeping track of that data and analyzing it to provide meaningful and actionable information will drive reduced costs and increased production. Some of the sources of Big Data in the Oil and Gas sector are:

➢ The continually increasing volumes of data and transactions originating from the remote sensors, PLCs, telemetric devices, smart phones, and other constituents of the oil and gas sector specific "Internet of Things".
➢ Social media sites used for sharing ideas, opinions, images, documents, related to Oil and Gas suppliers.
➢ Emerging technology developments that facilitate organizational change, drive productivity, increase creativity, speed new development, and create data
➢ Web portal, e-mail, SMS based interactions, as well as, device-to-device transactions all create data
➢ Sources of unstructured data include drawings, maintenance records, specifications, seismic imaging, documents, e-mails, audio and video.
➢ The changes driving recent increases in the volume and velocity of data in the oil and gas industry include the increasing use of sensors, larger numbers of unconventional wells, and the growth in remotely managed and automated conventional sources.
➢ Increasing numbers of sensors in wells and on drilling equipment provide a constant stream of data that can be used to determine when a piece of equipment might fail so that a preemptive solution can be implemented.

VCE's VCAP Unified Data Solution for Oil and Gas can be used to help inform oil field personnel of the most effective solutions and fixes for maintenance issues and impending equipment problems thus avoiding unnecessary down time and cost. Leading suppliers need to become actively involved and encouraged to engage and participate in pilot project programs designed to enable this type of next generation use of Unified Big Data.

Meeting The Challenge

The current challenge for Oil and Gas Information Technology executives and engineers is to ingest, integrate, analyze, and administer all available and relevant data from sensors, PLCs, and archives regardless of form, format, or location of data. Using Unified Big Data and data analytics tools engineers and other Oil and gas professionals can compare current conditions and data against all known and similar situations. This enables professionals to more effectively and rapidly make correct decisions. VCE's VCAP Unified Data Solution represents a key enabling technology for the fully integrated (first oil to delivery) digital oil and gas supply chain of the future.

Looking to the future of the oil and gas sector ideally a Unified Big Data Oil and Gas Solution like VCE's VCAP Solution will be used to capture and make available all known and available traditional and non-traditional data. More specifically such a solution will:

➢ Identify all appropriate and relevant data sources and determine the best way to capture that data.
➢ Require an information technology architecture that will scale to fit future workload requirements as demand increases.
➢ Identify and make available, either directly or through APIs, all best in class and most appropriate data analytics and assessment tools
➢ Incorporate the collection of data from Smart Phones used by field personnel and telemetric devices.
➢ Ensure total security and with no risk of sensitive data being accessed and misused – read as completely hacker proofed
➢ Ensure that the data and affiliated applications and tools are available 24x7.

➢ Capture and analyze current processes and combine with historic data for total process and outcome analysis, optimization, and improvement.

➢ Monitor oil and gas wells-downhole and at the wellhead for fluid rate, fluid phase, and reservoir pressure.

➢ Optimize oil well pump jacks (artificial lift) production by assessing down hole pump efficiency based on analysis of data on rod position and rod load and drawing comparisons against similar historic data.

➢ Optimize pipeline compressor efficiency by collecting data on current net flow, maximum capacity, and potential maximum capacity to analyze current single and parallel compressor efficiency and to draw comparisons against similar historic data to anticipate compressor failure and determine the presence of leaks in the system.

➢ Reduce nonproductive time and increase the production capacity of current reserves through enhanced oil recovery, tightly controlled maintenance costs, and improved health, safety and environmental regulatory compliance.

➢ Apply advanced analytics such as pattern recognition on unified oil and gas data to identify new potentially productive seismic trace signatures that previously had been over looked.

Each of these workloads and more (see graphic titled "Oil and Gas Data Sources") benefit significantly from the scalability and flexibility of VCE's VCAP Solution. VCAP brings extreme data ingest speed and capacity for mixed workload multiple source and format Big Data analytics. Oil and Gas suppliers will realize significant benefit through faster ROI, reduced TCO, and improved time to result. Right sized scalable compute and data intensive solutions like VCE's VCAP Solution address the computational, data handling, and analytical requirements needed to make the digital oil and gas supply chain a reality.

Unified Data Solution for Oil & Gas

VCE with their Vblock™ Systems provide a solution that meets all of the big data challenges head on. The Converged Analytics Platform solution consists of the Vblock™ System 300 and 700 families combined with BMMsoft EDMT and SAP IQ. The Unified Big Data challenge is to ingest, index, archive, manage, analyze, and unify high volume, high velocity, structured, unstructured, static, and streaming data from a continually increasing variety of sources and locations.

Workloads and applications that benefit as a result of implementing a VCAP Unified Big Data Oil and Gas solution (*see diagram – Oil & Gas Data Sources*) include: Asset Maintenance, Decision Support, HMI – Drill down GUI, IoT Data Analysis and control, Cost & Profitability Analysis, Predictive Analytics, Regulatory Compliance: health, safety, & Environmental, Demand, cost, pricing, & utilization Analysis, Production Metrics and Forecast, and Supply Chain & Logistics.
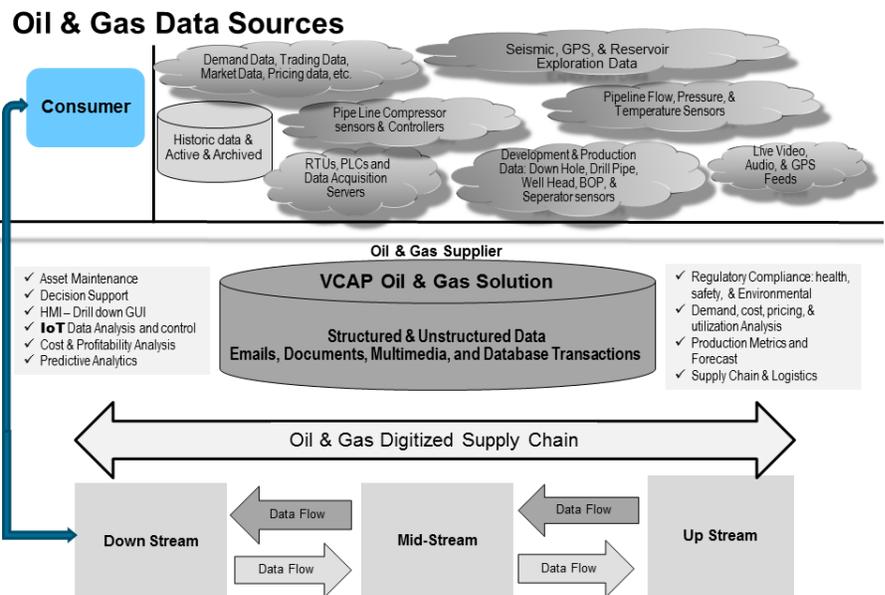
Vblock™ Converged Analytics Platform for Unified Data

The Vblock™ Converged Analytics Platform Solution is an integrated platform of compute, storage and connectivity infrastructure for ingesting, indexing, and storing Unified Data. This solution easily scales to keep pace with changing oil and gas supply chain requirements. With the Vblock™ Solution customers can run searches and queries against the indexed unified columnar data store.

Vblock™ Systems deliver the efficiency and business agility of virtualization and cloud computing, integrating industry leading compute, network, and storage technologies.

Vblock™ Systems are designed, engineered, manufactured and certified to leverage storage technologies from EMC, networking and compute from Cisco, and, optionally, VMware virtualization software. Cisco UCS servers and Nexus switches support innovations like Unified Fabric, embedded management, and policy based computing.

The Vblock™ Converged Analytics Platform solution, a services based solution, uses EMC's VNX storage for the Vblock™ 300 series platforms, and EMC's VMAX product for the Vblock™ 700 series platforms. EMC's VNX provides Fully Automated Storage Tiering (FAST) for the efficient and intelligent use of various data storage sub-systems.

The Vblock™ System with the Converged Analytics Platform solution is designed to handle mixed enterprise class application workloads and benefits significantly from Vblock's integrated operational, management, and administrative simplicity.

VCE Vision™ intelligent operations provides the systems management for Vblock™ Systems, optimizes services for a converged infrastructure, integrates directly with VMware technologies, and provides an extensible API for leading management tools. VCE Vision's features include virtualization optimization, converged operation and an open management API.

VCE Vblock™ Systems are delivered fully configured 45 days from the time of a customer order and they are fully operational within 48 hours of delivery. All Vblock™ physical and logical builds are completed and tested to ISO quality standards at the factory prior to shipment to the customer.

VCE Single Point of Contact Support includes a built in problem escalation process that seamlessly assigns any unresolved service tickets to the appropriate Cisco, EMC or VMware engineering support personnel. VCE told HRG that their first line support staff typically resolves 96% of calls without having to engage personnel from Cisco, EMC or VMware.
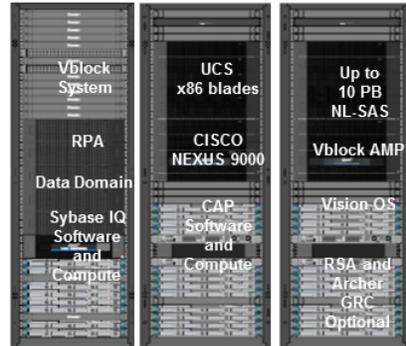
BMMsoft EDMT® Unified Data Solution

BMMsoft's EDMT® Solution ingests and indexes structured and unstructured data using up to 3 million independent data processing channels that ingest this data into a single analytic SAP IQ data repository. EDMT® automatically creates and stores metadata for the ingested content. The EDMT Universe connector for BusinessObjects, lets BusinessObjects users access, view, and analyze all of the structured and unstructured data in an EDMT archive.

EDMT stands for Emails, Documents, Multimedia, and database Transactions. EDMT® in combination with SAP IQ 16 is at the heart of the Vblock™ Converged Analytics Platform (VCAP). This converged solution delivers the performance, reliability and scalability needed for real-time ingesting, indexing, accessing, and cross analysis of extreme volumes of heterogeneous Big Data. In recent customer test, VCAP loaded 23 PB (23,000 TB) of mixed data using only 42 ETL channels at speed of 3 PB per day. During the same test, over 100 billion documents, files and emails were loaded – in addition to over 350 Trillion database and sensor records. EDMT was configured to use external storage to store actual files/images of those 100 billion files (whose metadata were loaded in EDMT/IQ) in external file systems i.e. EMC Isilon or WORM – or over 100 PB of external objects fully managed, versioned and searchable by VCAP.

EDMT supports enterprise-wide searches across disparate mixed data types. Virtually any type of electronic data regardless of format or location can be ingested, indexed, and queried in close to real time. When ingesting streaming data EDMT does not perform the extract and transform operations which means that even higher rates of data ingest can be realized making ingested data available for search, query, and analysis in close to real-time with ingest latency as low as 20 milliseconds. Using EDMT normal business level queries typically return results in the sub second range while highly complex queries can return a result in from 1 to 300 seconds.

Both EDMT and SAP IQ use column data stores providing significant benefits in terms of data compression, improved query/analytics speed, and better data security. EDMT's use of Multiplexing is an excellent fit with the SAP IQ technology as both benefit from the full SAN connectivity between EDMT, SAP IQ and storage.

EDMT and SAP IQ comply with the ACID rules for database architectures providing high-quality data that is critical for enterprise applications and reliable verifiable analytic results based on clean and complete data. EDMT uses unadulterated or "accent-free" SQL that ensures precise data definition, ACID compliant data relationships, and precise query results.

Using EDMT businesses can run unified real-time text analysis, SQL analysis, and cross-analysis of mixed data types. With the EDMT Solution real-time monitoring, instant cross-analysis of new and historical data and real-time response to market changes, product problems, customer dissatisfaction, litigation, audit, fraud threats, and competitive threats can all be handled from a single system with no impact on production systems. EDMT can enforce data retention policies to meet EIS and regulatory compliance requirements while leaving the source data in its original state. Furthermore, EDMT is unique in that it is compatible with all other enterprise class SQL based applications.

EDMT's high availability features support the implementation of multi-site active-active disaster recovery sites. Remote replication to multiple sites ensures that problems with one or more replication targets or channels will not impact replication. According to BMMsoft the latency between the primary site and the secondary site(s) can be configured to be as low as 2 seconds.

SAP IQ 16

SAP IQ, a column oriented database, provides significant storage compression, query speed and performance advantages when compared to traditional row oriented databases. Column oriented DBMS outperform traditional row-based database management systems running analytic workloads on average by a factor of 100 times.

With SAP IQ 16, SAP introduced a native MapReduce API, Hadoop integration, Predictive Model Markup Language (PMML) support, and an expanded library of statistical and data mining algorithms that leverage the power of distributed query processing across a Massively Parallel Processing (MPP) grid based on SAP IQ's Multiplex technology.

New SAP IQ APIs enable the implementation of proprietary algorithms that run in-database. SAP claims that running proprietary algorithms in database delivers greater than 10 times the performance acceleration as compared to existing approaches. Additional improvements have been made for text data compression and bulk data loading interfaces.

With the release of SAP IQ 16 SAP introduced In-Memory Row-Level Versioning Store extended to IQ Multiplex for scale out IQ deployments. Point-in-Time Recovery restores an IQ backup plus all committed database operations to a specified point in time. An OData (Open Data Protocol) Server that lets web clients communicate with an SAP IQ database server using the OData interface. SAP IQ Cockpit, a new graphical administration tool for on-board management and monitoring of SAP IQ, is the first phase in the evolution of SAP Control Center towards enterprise scale administration and monitoring of all SAP database technology.

The SAP IQ and BMMsoft EDMT Multi-node designs are very tightly coupled. The SAP IQ shared disk / shared database is made possible because SAP IQ is not a partitioned database. With SAP IQ you can add storage without being required to add additional servers just as you can add servers without having to add storage. (*It is worth reiterating that with SAP IQ 16 storage can be added independently from servers and vice versa which is NOT possible with Hadoop.*) This architecture allows the addition of servers of any size to the SAP IQ grid.

EDMT and SAP IQ both use MPP "shared disk" architecture. Adding or removing servers ("nodes") with SAP IQ is straight forward not requiring data reorganization or re-partitioning. A single query in either EDMT or SAP IQ grid can be spread across all nodes, or a predefined subset of all nodes, or use just one server. This functionality provides substantial flexibility in terms of workload management and user isolation. All available nodes can be regrouped within 1 second (according to BMMsoft) to provide support for a big query.

The nonstop functionality of both SAP and EDMT is closely tied to their multiplex capability. In the event of a node failure or replacement there is no requirement for the admin to take any action to reestablish access to "lost data" because all of the nodes in the SAP IQ MPP environment "see" and access all of the shared data directly. For disaster recovery EDMT's Multi-site Replication feature is used to replicate all shared EDMT data to the disaster recovery site.

Conclusion

VCE's Vblock™ Unified Data Solution for Oil and Gas provides a key enabling technology platform for the evolving and fully integrated (first oil to delivery) digital oil and gas supply chain. Oil and gas suppliers will achieve faster ROI, reduced TCO, and improved time to result using Unified Big Data analytics which take advantage of VCAP's data ingest speed and capacity for multi-source and multi-format data. In addition a decision support system based on VCAP's ability to unify and analyze data from disparate sources will help field engineers effectively deal unexpected situations as they arise.

Unified oil and gas Big Data that brings together data that has not previously been analyzed in this aggregated form provides personnel with access to searchable institutional knowledge helping to compensate for limited expert staffing and the growing shortage of experienced personnel.

Increased use of remote monitoring devices to collect and transmit data (regardless of location) contributes significantly to the growth in Oil and Gas related Big Data. VCE's VCAP Solution is fully scalable and able to meet most of the oil and gas sectors unified data analytics requirements for the foreseeable future. VCAP brings extreme data ingest speed and capacity for mixed workload multiple source and format Big Data analytics.

Harvard Research Group is an information technology market research and consulting company. The company provides highly focused market research and consulting services to vendors and users of computer hardware, software, and services. For more information please contact Harvard Research Group:

Harvard Research Group
PO. Box 297
Harvard, MA 01451 USA

Tel. (978) 456-3939

e-mail: hrg@hrgresearch.com

http://www.hrgresearch.com