industrial internet CONSORTIUM

BEYOND DIGITIZATION: THE CONVERGENCE OF BIG DATA, ANALYTICS AND INTELLIGENT SYSTEMS IN OIL & GAS

October 9, 2015

TABLE OF CONTENTS

SUMMARY

By 2020, 25 billion assets will be connected[1] - working smarter, faster and more efficiently than ever before. The oil and gas industry is beginning to experience the convergence of big data, analytics and intelligent systems. This movement towards convergence challenges the current silos of disparate technologies, from legacy SCADA and asset management, to historian and decision support systems. These challenges are paving the way for new, emerging solutions that drive data and predictive analytics into the edge, the data center, and the Cloud. The Result: The industry is realizing higher levels of performance and optimization that can result in outsized return on investment for oil and gas companies leveraging these new capabilities.

This white paper will highlight how the oil and gas industry has coped with the evolution of technology from field operations to the data center, exposing the need for the industry to consider dramatic changes to leverage big data and to advance the use of analytics.

In this paper, we focus on:

- The limitations of oil and gas asset management, big data, and communications systems;
- The evolution of next generation "digitalized" industrial assets;
- The convergence vision where enterprise-wide analytics will drive all elements of industrial operations.

THE DIGITAL OILFIELD – A DECADE IN THE MAKING

Over the past decade, the oil and gas industry has experienced significant shifts in how production operations have been improved through the application of new, emerging technologies. This drive toward "digitization" has been driven primarily by the proliferation of new software applications, new data formats, and the availability of massive amounts of real-time data. New analytic-centric technologies are now revealing operational nuances that were previously captured manually, or not at all. During this period, oil and gas companies have created remote operations centers focused on optimizing operations based on these analytic insights. However, these centers have principally focused on larger assets, where such investments were cost-effective or practical. To truly achieve the Digital Oilfield or "Digital Operationalization" with embedded analytics, oil and gas executives must overcome the challenges of the inherent limitations in today's infrastructure and systems. Next-generation, real-time analytic platforms require new ways of thinking on how data processing, acquisition infrastructure, and analytics platforms are: designed, deployed, maintained, and used in a distributed and robust manner.

LAGGING DATA COMMUNICATIONS LIMITS DATA ANALYTICS

In the past, robust data connectivity in plants has been a real challenge to design, deploy, and maintain. For upstream plants (e.g. offshore rigs or mature fields distributed across large, remote areas), the fundamental telecommunications infrastructure is extremely limited or simply non-existent. Additionally, for downstream plants, such as refineries, the large amount of various metals and other materials can negatively impact the robustness and reliability of many telecommunications technologies. Historically,

[1] Gartner Symposium/ITxpo 2014, November 9-13 in Barcelona, Spain

approaches to legacy communications tend to inhibit the advantages associated with real-time connectivity. These legacy systems were never built to support high fidelity data streams and analytics. For example, leveraging new analytics and methods of insight has meant layering different telecommunications, data processing, and data storage technologies, which in turn has created many silos of data. Accordingly, these silos have only resulted in more complexity in achieving meaningful data insights, since a considerable amount of the data was unusable (or simply unavailable at the right point or in the right silo) for more advanced applications and data processing. Communication approaches that are used by oil and gas companies today include: 3rd party data transfer devices; hard-wired connected equipment; radio devices + towers for line of site communications; wireless broadband; satellite; as well as no connectivity to the central system.

Due to the remote location of most oil fields, telecommunications systems were developed and deployed based primarily on availability and cost, and not necessarily on driving business outcomes. *Why?* This occurred because, historically, oil and gas companies were simply not data-centric enterprises and largely did not consider the opportunities, advantages, and impacts of the Industrial Internet (or Industrial Internet of Things, IIoT) and the consequential demand and growth of data and data processing needs.

Additionally, the industry relied heavily on subject matter expertise (SME) of seasoned operators and their ability to react and communicate immediately and effectively regarding real-time issues. These approaches are neither scalable nor proactive nor capital efficient.

40-Year-Old Technology Still Dominates

Supervisory Control and Data Acquisition (SCADA) systems are commonly used as the central data and control system at plants, leveraging legacy telecommunications platforms facilitating remote terminal units (RTUs) and programmable logic controllers (PLCs). SCADA is typically deployed in highly distributed architectures, making it difficult to perform analytics across the entire dataset. Over the years, many SCADA systems have evolved in capabilities; yet, many still lack in some critical areas that can help enable oil and gas companies to gain and optimize data insights. Some of these limitations include:

- Underlying data structures are typically not set up optimally to handle complex statistical approaches, critical for recognizing and predicting insights;
- Most SCADA systems have relatively low fidelity of measurement and data storage ranging from a data point every 10 minutes to an hour or longer;
- In some cases, SCADA systems cannot reach the field devices due to communication issues, resulting in the inability to capture high fidelity data for that specific period of time;
- In the field, SCADA systems require manual mapping of register values resulting in a large number of improperly mapped values as well as a large number of values that are never sent back or received;
- SCADA systems also pose a huge barrier to real-time scientific monitoring of critical assets, given its current state of data availability and quality.

Established in the 1970's, Modbus is a popular and widely deployed, standardized communication protocol, predominately used across many existing PLCs and RTUs. Much like the overarching SCADA system, Modbus comes with some inherent challenges, such as its Master/Slave architecture, lack of indexed registered values, limited tracking of devices (only 250), and the limited number of data types. Newer protocols have been designed to help work around some of the legacy challenges of Modbus, but their implementation have been limited, usually specific to the vendor that created them, and generally proprietary.

Classic Historians Becomes Unfulfilling Solutions

To combat the challenges with many SCADA systems, oil and gas companies have turned to commercially-available historian platforms because they offer better user experience and visualization capabilities. Additionally, these historian systems ingest cross-functional data into a single centralized location. While historians have enabled a new level of data analysis for many companies, these systems do have several limitations, including:

- Many historian systems are built on proprietary or traditional row-based databases, requiring the partitioning of data across multiple servers, which in turn can limit complex data analysis across the entire dataset in an expedient and robust manner.
- Some historian systems archive and expunge data frequently, limiting the accessibility of historical raw data for analytics.
- Historian systems are dependent on the quality of the data provided by source systems, such as SCADA and Asset Management platforms; when provided with poor quality, inconsistent data, most historian systems will merely store that bad data along with good data.

The underlying limitations of this commonly deployed legacy architecture compromise the overall business impact of these applications, even if the analytics capabilities were present. Simply put, these legacy architectures were not designed or implemented to address the needs of complex analytics and machine learning.

The Industrial Internet - More Data, More Benefits

The widespread availability of inexpensive sensors has driven unforeseen possibilities in the Industrial Internet. Modern assets are now IIoT-enabled, shipping with imbedded smart sensors. The popularity of the Industrial Internet and the solid business outcomes realized by its implementation has pushed companies and vendors to retrofit older legacy equipment with inexpensive and robust sensor technologies. Naturally, the more IIoT-enabled assets deployed and connected to a robust telecommunications fabric, the more data is generated and available. This gives oil and gas companies, much like any other industry sector, the potential to perform greater analytic analysis and obtain business insights from such data.

One example of these new opportunities lies in remote oil and gas operations. Many remote operations centers now have rich operation capabilities that extend well beyond mere monitoring to include broad remote management and onsite reporting capabilities. This in turn allows for much more timely reaction to problems and opportunities, and enables a more proactive management approach to assets.

To add to this, greater automation of asset performance can be achieved as these assets become smarter and control logic is pushed closer to the asset. New sensors and technologies, such as video analytics and streaming data analytics, are enabling an increased number of assets in the field to demonstrate the capability of executing discrete autonomous decisions. Over time, much of the remote management of assets can be delegated directly to those smart assets that are capable of performing these edge analytics and autonomous decision-making capabilities, rather than solely relying on the remote operations centers for the decision execution.

Security Implications

IIoT edge devices will increasingly have a profound impact on the demands of security infrastructure, as the Industrial Internet becomes an integral part of the oil and gas technology ecosystem. From a security standpoint, IIoT-enabled assets present an expanded attack surface with new vectors of vulnerability across connected systems and distributed devices. Unlike traditional computing devices, IIoT edge devices are generally embedded sensors and controllers with fixed functions designed to perform specific tasks. Security concerns can shut down device-cloud-device analysis and feedback loops. In turn, this can severely impact scenarios where sub-second responses for control loops or data partitioning are required. Many of these IIoT-enabled industrial devices and sensors simply were not designed to operate in a connected ecosystem in a secure, robust manner. Many device designers simply did not anticipate the threats inherent in a connected future.

Fundamental security controls like device authentication, authorization, and encryption controls are often missing from IIoT devices. Retrofitting them with appropriate security controls (or replacing them with devices with these embedded and updatable controls) needs to be part of organizations' deployment strategies moving forward. A holistic end-to-end approach to security that spans information technology (IT) and operational technology (OT) is fundamentally required.

As companies look for new ways to gain insights from connected assets, they must ensure they can rely on the integrity of the data to make decisions. Decision-making (especially autonomous decision-making) based upon streaming telemetry from IIoT-enabled equipment will have far reaching positive impacts on all facets of the organization's business. However, decision systems must be robust enough to detect and isolate from bad, corrupted, or injected telemetry data and attacks. A cohesive strategy applying a data assurance program will be critical. Organizations will need to implement data-level security measures and implement a framework that governs data assurance across the edge infrastructure. In turn, this must instill a higher level of confidence in data-driven decisions. IoT communication protocols and data formats such as MQTT, CoAP, DDS, 6LoWPAN, ZigBee, ModBus, WirelessHart, and others offer different security capabilities and strategies, each having their own data assurance limitations. These limitations will need to be taken into account as new edge devices and management platforms are architected and deployed. Additionally, it will become increasingly critical to leverage analytics to proactively address security and integrity threats across IT and OT systems.

IT and industrial control systems (ICS) also present their own set of security challenges, especially as these platforms are increasingly converged. The traditionally long technology lifecycle of ICS can result in issues as simple as the running of outdated firmware and control software no longer supported by the vendor, or simply have well known or obvious security or reliability/integrity vulnerabilities. Physical isolation of these systems, commonly called "Air Gapping," can isolate control applications, such as PLC and SCADA platforms from connected networks. However, this is becoming increasingly difficult to maintain and justify as organizations embrace increased connectivity and its associated benefits. Attacks can originate across connected networks and onto an ICS domain by utilizing limitations in legacy networking, data protocols, and methods of security that simply were not designed for an always-connected environment. When combined with a typically cautious approach operators take towards patching and updating systems, many systems can be unpatched or updated for months, or even years, which only increases the available attack surface.

Asset Management

Today, oil and gas companies have the opportunity to implement better asset management capabilities. In the past, the low levels of data accuracy associated with incorrect or outdated information inhibited a company's ability to evolve toward being a "digital operationalization"- focused organization. Some examples of how existing asset management has limited the advancement of IIoT and analytics include:

- A lack of integration between procurement systems and asset management systems requiring duplicate entry, which typically results in inaccurate or outdated data being stored and accessed.
- Assets frequently encounter problems and require maintenance or replacement, but rarely are these maintenance processes linked to the asset management system to track all the historical interactions for a given well/asset.
- Microsoft Excel is still one of the biggest tracking mechanisms in the field. Often, various forms of Excel files are utilized across organizations to obtain critical data on Well Tests, Maintenance Logs, Failure, Tear Down Reports, and other needs. This valuable information is rarely centralized and analyzed because of the lack of accessibility within a system. In effect, many organizations have islands of Excel-based data, residing (in many cases) on individual users' computer systems.
- Vendor test, performance, and commissioning reports are commonly provided in hardcopy format. Those reports are rarely uploaded into a central system, and in some cases simply sit on a shelf awaiting later retrieval (if ever). In turn, this makes it difficult to know how current performance contrasts against commissioned performance and the specifications of the shipped asset.

Knowledge of the asset gives critical context for engineering models for determining asset performance, state of health, and proactive maintenance needs. Without accurate asset and well profile data, it is very difficult to recognize, predict, and appropriately act on issues impacting production. This results in millions of dollars in lost operational performance, misinformed decisions on well productivity resulting in increased costs, planning and maintenance costs, and more.

Additionally, with the availability of data from IIoT-enabled assets, analytics can be applied to predict and improve maintenance, enable remote monitoring of equipment, and implement planned (rather than unexpected) shutdowns. A lack of information about the reliability of a drilling rig, for example, has a significant impact on cash flow when it results in indecision, or unnecessary risk overcompensation.

Companies need to close the loop across associated processes. Asset maintenance insights need to be communicated proactively and traceably to multiple parties, while being integrated into the workflows that are implemented to ensure operationally effectiveness. For example, spare and underutilized equipment and spare service parts can be made more readily available through greater sharing of information in near real-time. Further, inventories of spare assets can be reduced when the status of all assets and their locations are well known, and the data is accurate when viewed. Asset maintenance analytics can also affect other processes, influencing outcomes not only for capital assets, but also for scheduling and utilization of these assets and associated human resources.

HUMAN CAPITAL REMAINS CRITICAL TO DECISION SUPPORT

Workflow and decision support systems are not immune to the challenges of elevating operational systems and processes to garner data analytic insights. Human intelligence and intervention remains a huge factor in interpreting real-time information, and executing appropriate decisions and actions. These scenarios can be as critical as identifying the specific root cause of equipment failure, to evaluating and resolving the entire workflow of production failure. Many scenarios include the involvement of multiple decision stakeholders, as well as process routing and operational procedures, complicating the overall effectiveness of the analysis. Without well-established decision support and event escalation processes and systems, many oil and gas companies continue to rely on operator expertise as opposed to systemized best practices for the treatment of specific events.

In the report[2], "*The Future of Work in the Oil and Gas Industry*," Accenture argues that the convergence of data, machines, and people, coupled with analytics powered in part by the emergence of the Cloud, can enable smarter decision-making by people at the edge of the organization. Accenture envisions a continuous stream of data coming from sensors that, in turn, is analyzed by a central onshore command center, leveraging the results from offshore (edge) analytics. In the immediate future, intelligent data and commands will also be fed back to the wells from onsite and offsite analysis, enabling field personnel to make wiser and more data-driven decisions based on local conditions. Enterprises may even be supplied near real-time information regarding a specific pump and its maintenance history on a wearable device that recognizes the equipment near a specific technician, operator, or engineer.

As oil and gas companies continue to digitize work and compete more on data management and analytics proficiency, this shift will transform the nature of the industry's workforce. Companies will rely less on the personal experience of specific people and rely more on science and data understood and shared by all. Digital literacy skills are becoming increasingly important across the workforce. Such literacy includes an understanding of applied analytics, use and development of new software and hardware platforms, and the ability to effectively use new tools such as live collaboration technologies. According to a benchmark study[3] by Accenture, the oil and gas industry needs to hire 11,900 new data analysts to just meet increased demand. As work becomes increasingly virtualized and analytics-based, organizations will be able to tap into new pools of workers. These new pools may include those with experienced professionals either working part time or retired, experts in different geographic locations, or those simply seeking a greater work-life balance. Such new pools of human resources could be permanent employees, or a new global digitally connected workforce of experts available on demand, or some combination of both.

THE CONVERGENCE VISION

The convergence vision is where pervasive analytics will drive all elements of industrial operations. In this vision, the gap between operations and information technologies (traditionally characterized by different systems, standards, and manufacturers) is bridged, and the integration of data from sensors and devices

[2] Accenture, "The Future of Work in the Oil and Gas Industry" 2015
[3] Accenture Digital: "Oil and Gas companies: How Digital is Transforming Work as We Know It" 2015

with control systems (SCADA), middleware (e.g., manufacturing execution systems), and back-end IT systems (e.g., ERP) is complete. Data from machines can now be captured and integrated into IT systems for analysis, frequently occurring in real-time, and tiered between distributed edge (on premise) and centralized layers. This convergence between IT and OT systems will greatly enhance the operational excellence and the subsequent profitability of oil and gas companies. This new approach will leverage new edge device and communications capabilities with data quality and assurance, as well as enhanced analytics-driven control capabilities.

This vision will be underpinned by a foundation of big data technologies, which will be necessary to handle the volumes of data with real-time and near-time analysis driven by scientific principles. The data structure necessary to enable complex scientific exploration will be the primary design focus of these next-generation architectures. It is important to note that these next generation solutions will drive the practical elimination of data silos across areas of critical information.

This information includes asset and well profile information, sensor data, well test records, maintenance and failure reports, and many other sources. Each one of these data sources is critical to recognizing and predicting real-time performance optimization opportunities.

Oil and gas companies will, in this vision, have the ability to leverage a holistic view of the enterprise, from field operations to the data center, and obtain a level of specificity in analysis and operations that will greatly improve organizational performance.

Emerging new thinking on convergence explores how complex data science and machine learning will be run across all critical data attributes in real-time. This new thinking includes leveraging derived calculated sensors, statistical classifiers for recognition and prediction of performance optimization, and maintenance and failure event analysis at a scale unimaginable given the disparate systems deployed within oil and gas today.

Built-in workflow and decision support capabilities will ensure the increasingly efficient and effective routing and handling based on specific realized events and predictive trends. Such workflow and decision support capabilities will evolve over time to include both work order and operator involvement, as well as automated control to continually adapt the systems to their most systems-level optimal settings, maximizing overall production while reducing cost. These new capabilities will have a significant and material impact on unplanned downtime and overall production performance, while improving resource utilization, efficiency, and safety.

So, why is convergence happening and why is it important to oil and gas?

Cloud computing, inexpensive sensors, progressive network availability, and big data analytics (centralized at distributed at the edge) have paved the way for this convergence. In the past, oil and gas companies, like many other industrial sectors, were able to ignore many of IT enterprise technology and process trends, relying on technology advancements within the organizations' current context that were perceived as "necessary" and cost effective. As these technology trends collide, driven by increased demand for high fidelity, higher-quality data, an environment has emerged that is pushing transformative change. Oil and gas companies can work collaboratively, using the Industrial Internet to unlock new value from device data generation, robust connectivity, and interoperability. This information, in turn, will help them to cut costs and to improve efficiency, productivity and profitability.

The Evolution of the Next Generation, Digitized Industrial Assets

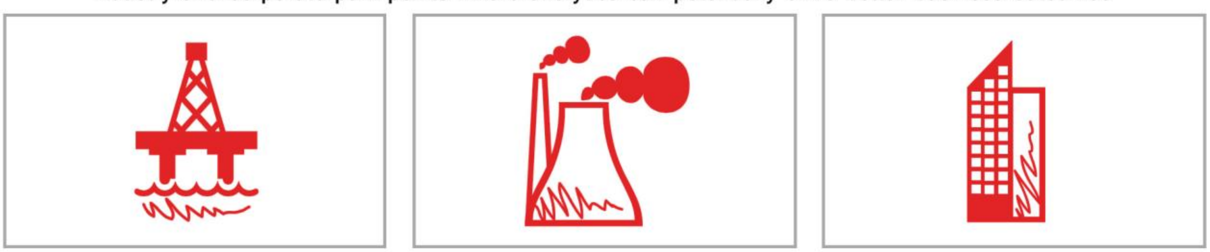
Industrial assets are characterized by very long lifecycle during which large volumes of data are produced and used by varying departments, from engineers to field forces, business project leaders, and operations directors, right through to the C-level executives. By adopting a 100% digital approach and integrating across all phases of the asset lifecycle as well as their partner ecosystem, enterprises can improve data access and gain improved insights about their assets. This could span across enterprise handling asset design to construction, operation, maintenance and disposal as well as DiALM incorporated with CAD (Computer Assisted Drawings), 3D images, timelines and other forms of non-tabular data. This will ultimately help leverage data to enhance profitability and safety.

Incremental improvements at every point of operational and business decisions. → **10% points improvement = $260 million bottom-line impact on single brownfield asset.**

Investment Opportunities for Analytics in Oil and Gas

In today's environment, where instant response is expected, oil and gas companies need advanced systems that enable them to do the things they already do but faster and with enhanced flexibility. Furthermore, advanced analytics solutions are needed to manage the volume of data (structured and unstructured) in upstream exploration and production (E&P), downstream applications and in corporate functions.

Industry and corporate pain points where analytics can potentially drive better business outcomes



Upstream Operations	Downstream Operations	Corporate Operations
Forecast and deliver production commitments	Optimize end-to-end the integrated value chain - from plant to pump	Optimize cash flow to effectively meet planned capital expenditure commitments
Efficiently deliver unconventional plays	Configure the supply chain to enable cost reduction in the manufacture of specialty lubricants	Enable or manage contingent labor
Enforce rational and appropriate working standards	Measure and manage market risk - at commercial and logistics levels	Measure and manage market risk - at a commercial level
Manage equipment supply chain	Enforce rational and appropraite working standards	Enforce rational an appropriate working standards
Execute capital projects to time, budget and specified scope		Execute capital projects to time, budget and specified scope

Source: Accenture Analytics, March 2013.

CONCLUSION

As next generation, smart assets are deployed, oil and gas companies will evolve their operations and processes to accelerate the rapid growth and volume of data. As this data is consumed, performance-based, predictive solutions will leverage current and emerging infrastructures and technologies to provide material improvements in field operations. Analyzing the data derived from assets across the value chain will not be reserved for just operators in the field. Rather, performance optimization and big data analytic solutions will automate and enhance how valuable insights are communicated throughout the enterprise, from field operations to finance, and be used to address problems as downtime, system failures, and financial projections.

Security, safety, and the interoperability of platforms interfacing or creating data and information are just some of the challenges facing the oil and gas industry as it begins to leverage these new capabilities. To accomplish this, the industry must collaborate and share best practices, in order to learn, adjust and adapt going forward. New risk scenarios need to be rigorously tested through the application of use cases and testbeds. More recently, collaborative work has been facilitated across a broad ecosystem approach through organizations such as the Industrial Internet Consortium. Using a public-private partnership model to solving the challenges in Oil & Gas, the Industrial Internet Consortium will be addressing interoperability, use cases, testbeds, and security/safety/privacy in the coming months, to provide recommendations and best practices as Oil & Gas companies begin to understand the impact of the convergence of big data, analytics and intelligent systems on their businesses.

Authors and Contributors

ABOUT THE INDUSTRIAL INTERNET CONSORTIUM

The Industrial Internet Consortium® (IIC™) is a global, member supported, organization that promotes the accelerated growth of the Industrial Internet of Things by coordinating ecosystem initiatives to securely connect, control and integrate assets and systems of assets with people, processes and data using common architectures, interoperability and open standards to deliver transformational business and societal outcomes across industries and public infrastructure. The Industrial Internet Consortium is managed by the Object Management Group (OMG). For more information, visit www.iiconsortium.org.