

WHITE PAPER

Big Data: Improving Decisions thru Smart Instrumentation



Effective business decisions that drive an enterprise strategy are made through data. With the ubiquitous connectivity resulting from the “Web of Things”, Big Data is proliferating into the boardrooms of major fortune 500 companies in many different industries at an accelerating pace.^[1] It is estimated that over 10 billion devices connect to the internet via a wireless connection, with an estimate that this number increases to 30 billion by 2020.^[2] With ever-expanding surges in connectivity, more and more data becomes available for analysis. With accessible data, executives are able to make operational decisions which drive their business to accomplish its strategic goals based upon data that was not available in previous years either because of difficulty accessing it or understanding it.

BIG DATA

The term Big Data has been used in many different ways to describe many things over the last year. So what is Big Data? Big Data combines information from disparate sources inside your company from traditional sources to non-traditional sources.^[3] These sources may be databases, paper reports such as maintenance logs, preventive maintenance reports, emails, invoices, SCADA information, trending data, etc. Because of the increased processing power and more devices connected to the internet, there is a sense of data overload where understanding the data can actually hamper smart business decisions. Part of the reason for this is that decision makers do not see or fully grasp the relationships from the data. The relationships are displayed in such a way as to show correlation, and not causation; effectively defining what it is, and not why.^[1]

Nucleus Research performed case studies that show very successful results if close attention is paid to data context and connections.^[4] So how can Big Data help executives at Oil and Gas companies make more informed decisions? Analyzing “Big Data” provides decision makers with tools to make better operational decisions that impact efficiency, costs, security, and ultimately contribute to greater profits. Successful “Big Data” implementations can achieve ROI's greater than 250%.^[5] Oil and gas companies just like many other companies try to improve profitability through improving operational efficiencies. Executives at these companies work at optimizing efficiencies in many different segments of their business. For example, improving models based upon data from instruments related to flow rates, densities, pressures, gas quality and other process variables can create efficiencies and improve throughput. Process data from many different sources can intelligently let production understand how a changing well may impact the needs for artificial lift. Operationally, equipment downtime can be reduced by predictively determining remaining life of equipment based upon information collected. As instruments and instrumentation diagnostics improve, users can



leverage this information with other data in context to reduce costs and achieve goals. Improving the success rate of “Big Data” implementation can be seen in the use of contextual examination and utilization of data within enterprise parameters and corporate governance.

As data sets get larger in volume, operations on these sets require a different way of manipulating the relationships. Some techniques that have been employed consist of tensor based computational computing. According to the Future Directions in Tensor-Based Computation and Modeling, “High-dimensional modeling is becoming ubiquitous across the sciences and engineering because of advances in sensor technology and storage technology.” Although tensor computations are not new, because of the nature of the data sets being extremely large, Tensor representation is a logical choice. Tensors are a mathematical representation of vector spaces that allow the data to be represented in a matrix of multi-dimensions. Algebraic matrix operations can be performed on the data to determine specific relationships between the data.^[6] Tensor Computation is much more complex than this and beyond the scope of this article.

Advancements

Innovations in lightweight pervasive computing protocols help to bring smart sensor technology into the realm of Big Data in a significant way. With increased mobile connectivity, improved wireless technology, SQL replication services, and new parallel processing paradigms, production data becomes readily available. New lightweight protocols such as MQTT, have come into the forefront allowing devices that have low bandwidth capabilities or smaller power budgets to report data that normally would not be collected. MQTT (Message Queuing Telemetry Transport) was designed specifically with bandwidth or remote locations in mind. MQTT is well suited to a sensor network in a remote location where satellite may be the only connection mechanism. MQTT also is very effective on unreliable networks and with low powered/battery powered devices. MQTT has variants with MQTT-S (over UDP) which allows the sensor network to conserve power by sleeping for longer percentages of time. Other protocols developed to operate in this space include CoAP, and XMPP. CoAP is very similar in nature to MQTT, and XMPP is for the very large systems that consist of 100,000+ sensor nodes requiring high security.^[7] CoAP, also known as Constrained Application Protocol, is specifically designed for small, resource limited sensors or switches that rely on limited power budgets.^[8] HART, or Highway Addressable Remote Transducer protocol, is a bi-directional digital communications protocol that allows smart instrumentation and a master controller to communicate. The master controller can be a Programmable Logic Controller (PLC) or a HART Communicator. HART is unique in that it modulates a digital signal on the instrument's 4-20mA signal lines. HART has been around for many years and its benefits are not new. The HART 7 protocol released additional diagnostics that were not available in earlier releases that drastically improve analytics for smart instruments. HART allows for integrity validation, monitoring system availability, faster troubleshooting time, integrated systems with the device to provide better data for detecting problems, improved alarming to detect variations in the device which help to minimize unexpected maintenance costs, process interruptions, and plant/equipment shutdowns. HART also helps with safety by providing advanced diagnostics that improve Safety Integrity Level (SIL), facilitates automated safety shutdown testing, and enable automatic storage of compliance data for improved regulatory compliance.^[9] For a complex system, a facility may have many combinations of these various protocols all producing data that drive organizational decisions. These protocols all operate independently and concurrently populating databases with very large amounts of data.



Smart Transmitters are evolving to a point where intelligent data is being presented that allows the end user to understand the health of the device. For example, smart pressure transmitters can report remotely that a diaphragm is ruptured allowing a maintenance technician a better understanding of what is happening at a field location before he arrives as opposed to just showing low current or bad quality from the instrument. A smart Guided Wave Radar (GWR) Level Transmitter can detect if the unit is faulty, resulting in reduced level accuracy, degraded precision, or erroneous level outputs. These instruments have the ability to determine if the unit has wear due to vibration or extreme temperature changes, enhancing the maintenance on the unit for quick repair thus eliminating process upsets. With the ability of smart transmitters to have historical data in the unit, predictive analysis related to the process can be used to improve process control resulting in optimizations and efficiencies. All of this data can be made available through MQTT, HART, or other protocols that can be accessed via the boardroom to allow for better decision making capabilities. This data can be scaled into larger pieces of data that may help predict capital requirements.

Next generation technologies will help oil and gas producers not only understand how to get to resources faster, but will provide a way to comply with future government regulations. Data coming from sensor technologies has for many years been collected through SCADA systems to provide the ability to control processes. However, this data is typically not presented in an effective way outside of the HMI (Human-Machine Interface) application. Using rich visualization techniques with real-time data, operational insights can become much more apparent. Large collections of data, both historical and real-time in nature, require new approaches so that data context can be seen.

Patterns

With Big Data, much of the information is in an unstructured format. Data presented in this fashion doesn't easily translate into improvements to the bottom line for a company. Understanding trends within the business is needed in order to capitalize on the available information. In order for a company to use data that may come from intelligent sensors, legacy instruments, maintenance log sheets, preventive maintenance records, etc. it is necessary to link the structured data with the unstructured data. Patterns can be seen within this data if it is properly linked. For example, real time data coming from the pressure sensor that may indicate a ruptured diaphragm can be compared in real time to maintenance sheets, pressure relief valves, and regulators giving a maintenance employee more information as to the real cause of failure. Big Data collected from smart transmitters, legacy transmitters, and even mechanical devices can provide powerful analytics that give the user better operational control.



New computational algorithms are being developed to handle the complex patterns associated with parallel computing, non-centralized data, unreliable network connections and fault tolerance. Many large well-known companies are engaged in creating these solutions. For example, Google in its efforts to address this complexity, designed new algorithms that hide these details in a map and reduce structured operations. This allowed Google to "parallelize large computations easily", spreading out the processing needs in a decentralized fashion.^[10] Similar techniques exist at Cisco as well as

IBM. Other implementations consist of Hadoop (by Yahoo), as well as In Memory Database, and massively large parallel processing databases.

Future Enhancements

Smart transmitter technology, in the future, is envisioned to be process aware. It is not hard to envision instruments understanding what they are doing and detecting more information regarding when the process is not optimized based upon their role and location within a process. Sensors containing GPS technology to map specifically where they are installed on a pipeline, or intelligent mesh networks where transmitters actually share information as opposed to operating solely as an interpreted input to a PLC/DCS control system, may be here much more rapidly than anticipated. Managing data at this level helps companies understand future trends, predict improvements in products, improve capabilities, etc. Advances removing issues of power limitation, size/weight limitations, and complex communication structures while improving performance and reliability will continue to evolve. To create more advanced solutions, Big Data will drive the creation of smarter sensors. As instrumentation companies analyze models and simulations, sensors are becoming smaller with less power usage. Reduced power consumption and lightweight protocols for communication will be the catalyst to allow smart transmitters to fit in the world of Big Data.

For more information, please contact:

Michael Bequette, SOR® V.P. of Engineering
mbequette@sorinc.com
913-956-3040
sorinc.com

References

- [1] G. Satell, "Companies that Can't Figure Out Data Are Getting Left Behind," august 2013. [Online]. Available: <http://www.businessinsider.com/how-big-data-affects-strategy-2013-8>.
- [2] "More Than 30 Billion Devices Will Wirelessly Connect to the Internet of Everything in 2020," 9 May 2013. [Online]. Available: <https://www.abiresearch.com/press/more-than-30-billion-devices-will-wirelessly-conne>.
- [3] L. Arthur, "What is Big Data?," 15 August 2013. [Online]. Available: <http://www.forbes.com/sites/lisaarthur/2013/08/15/what-is-big-data/>.
- [4] "Informatica: Bridging the gap between traditional data and big data," [Online]. Available: <http://nucleusresearch.com/research/research/its-about-little-data-not-big-data/?page=3>.
- [5] "Where Big Data Shows Huge ROI," [Online]. Available: <http://www.information-management.com/news/big-data-ROI-Nucleus-automation-predictive-10022435-1.html>.
- [6] "Future Directions in Tensor-Based Computation and Modeling," 1 May 2009. [Online]. Available: <http://www.cs.cornell.edu/cv/tenwork/finalreport.pdf>.
- [7] "Blog about M2M, Internet of Things and the vertical applications for Smart Home, Smart City and Smart Vehicles," 5 November 2013. [Online]. Available: <http://www.iotprimer.com/2013/11/iot-protocol-wars-mqtt-vs-coap-vs-xmpp.html>.
- [8] "Constrained Application Protocol," [Online]. Available: http://en.wikipedia.org/wiki/Constrained_Application_Protocol.
- [9] "HART Communication Foundation," [Online]. Available: <http://www.hartcomm.org/>.
- [10] J. D. a. S. Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters," [Online]. Available: <http://static.googleusercontent.com/media/research.google.com/en/us/archive/mapreduce-osdi04.pdf>.