Internet of Things (IoT) and Big Data Capabilities

Can you help your organization position itself to capitalize on the opportunities presented by IoT and Big Data?

In the business world, it is common wisdom to know that if you are standing still, you're moving backwards. The "Internet of Things" (IoT) and "Big Data" are current buzzwords and reflect a trend in enterprise computing world towards capturing and analyzing large amounts of data from a variety of disparate data sources.

While the term "Big Data" is relatively new, the process of collecting and storing large amounts of information is old; the terms reflect an evolution of technological advances rather than a revolution. The ability to collect more data from different places has resulted in an increase in the volume, velocity, and variety of data.

Table of Contents

How Can Emerson Help?

The Data Management consultants specialize in designing, implementing, and supporting real-time, historian based, reporting and integration solutions for companies. We are a vendor agnostic group with extensive experience with all facets of the data management life cycle and have worked with a plethora of different tools for data collection and storage, analytics and visualization including many of those listed in this article. We can help your company take advantage of the new advances in system connectivity and analytical competency. Emerson is the leader in the collection, integration, and analysis of time-series data. By understanding your environment, goals and needs we will come up with a recommendation that's right for you. We have already helped several clients achieve many of the key benefits of IoT and big data listed in this article, such as improved performance management, real-time operational visibility, reduced maintenance costs, and overall business intelligence. Your organization can lean on our expertise; together we will come up with a road map to empower you to turn your data into insight and achieve both qualitative and quantitative benefits.

Background

IoT is an environment in which objects can be assigned unique identifiers and the ability to transfer data over a network. The objects can be anything, for example – animals with biochips, people with heart monitor implants, or automobile vehicles with sensors on tires to communicate pressure values. One of the key factors permitting the IoT trend is the large range of addresses made possible with IPv6, which has a 128-bit address space. This means that there are 2128 or approximately 3.4 x 1038 addresses available (IPv6, 2016). Cisco estimates that by 2020, 50 billion devices will be connected to the Internet. Yet today, more than 99 percent of things in the physical world are not connected (New Cisco Internet of Things (IoT) System Provides a Foundation for the Transformation of Industries, 2015). Newer wireless network connectivity specifications are starting to offer advantages over traditional options in cost and simplicity (11 Internet of Things (IoT) Protocols You Need to Know About, 2015). For example, the ZigBee wireless standard is intended to be simpler and less expensive than other wireless personal area networks such as Bluetooth or Wi-Fi. ZigBee devices also require less power and can transmit data over mesh networks of intermediate devices to achieve better reliability (ZigBee, 2016).

Industrial IoT

In the industrial world, some view IoT as something which they have been using already with the functionality provided as part of supervisory control and data acquisition systems (SCADA) (Dingman, 2015). However, more recently, IoT wave has encompassed advances in sensor technology, connectivity, and analytics which will expand the impact of data and improve decision support systems. In the upstream and midstream oil and gas industry, the IoT evolution has led the way for C1D1 certified wireless I/O devices to accommodate multiple analogue and digital inputs and communicate at the 900MHz or 2.4MHz frequencies. These devices can then transmit data to remote terminal units (RTU), distributed control systems (DCS) or SCADA systems. Innovations such as these offer a significant cost savings opportunity as companies can deploy connectivity devices efficiently and inexpensively while not sacrificing data collection and safety (Longley, 2015). The availability of drone technology also offers other data application possibilities such as monitoring and surveying pipelines in disconnected environments (Williams, 2015).

The Benefits

Advances in connectivity protocols and technology in the industrial sector, along with upsurge in mobile digital devices such as wearable technology and the rapid global increase of mobile technology will serve to increase the number and types of data sources in an organization's information landscape. Here are some of the areas in which IoT can improve operational visibility (Challenges, Opportunities and Strategies for Integrating IIoT Sensors with Industrial Data Ecosystems, 2015):

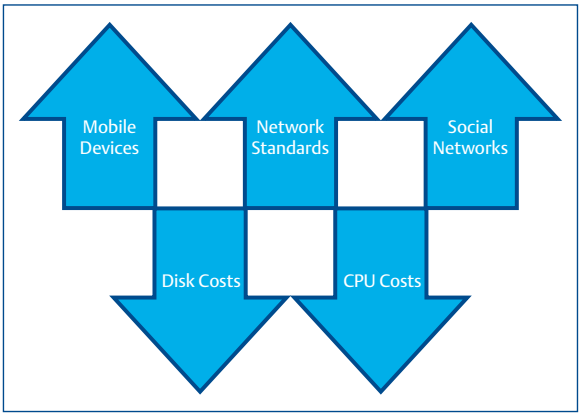
- Performance management – Sensor data can enhance real-time situational awareness and mobilize assets that were previously monitored manually. Information collected can be used to cut operational costs, improve process efficiency and asset availability.
- Improved planning – IoT enables industries to collect data from geographically disperse assets, and when this data is combined with process data, new insights can be achieved which better position organizations from a planning perspective.
- Collaborative benefits – secure data sharing will enable organizations to share data more easily and establish data driven collaboration and partnerships.

What Is Big Data?

In the broader enterprise computing infrastructure view, both IoT and big data should be discussed together since they share a close relationship. Big data is a term that describes large volumes of structured and unstructured data that is captured by organizations daily. The analysis of this data can lead operational insights which enable better business decisions. IoT will be one of the key trends driving the adoption of big data analytics tools.

Trends Impacting Big Data

Despite all the hype, IoT isn't the single driving factor behind the big data push. Over the last 20 years several different trends have affected the ability to create and store more data. These trends have promoted the need for scalable data collection and analytics architectures. Below ISA lists of some of the key items. Mobile revolution – smart phones, tablets, wearable technology. In 2013 there were approximately 6 billion mobile subscriptions, which is approximately 87% of the world's population (Knudsen, 2013). Improved network connectivity – advanced wireless communication standards, fiber optics, and IoT enabling more devices to connect at faster speeds. Social networking – billions of users now on social networks such as Facebook and Twitter. Cheap storage – the average cost of hard drive storage has decreased exponentially in the last 30 years making it more affordable to store more types and quantities of data (Average Cost of Hard Drive Storage, 2015). Inexpensive computing – Like storage costs, the average size of computer chips has decreased significantly in the last 30 years. At the same time, processors have become more powerful and less inexpensive (Markoff, 2015).



Breaking Down Big Data

While IoT plays a key role in the data collection process, the big data processing functionality can be broken down into three broad sections – storage, analytics, and visualization.

Storage

With the massive amounts of data coming in from a variety of sources, the first challenge with getting value from big data is storage. The huge, continuous stream of data will increase the need for more robust and scalable data storage and warehousing. Whether the data centers are managed by the organization itself or through a cloud or platform as a service provider, there are a variety of data storage options available. Data governance and security, as well as data virtualization frameworks can broadly be included in this category of big data management. The advance of in-memory data processing has allowed traditional database and ERP vendors such as Microsoft, Oracle, SAP, and IBM to provide high performance data storage solutions (Mullins, 2015). Whereas these vendors have proven solutions for structured data, other firms and solution providers are positioning themselves for better handling of unstructured or semi-structured data. Apache Hadoop is an open source framework for distributed storage and processing large datasets and provides a suite of capabilities founded on top of its Hadoop Distributed File System (HDFS) (What is Apache Hadoop?, 2015). NoSQL databases such as Apache CouchDB and MongoDB excel at document storage and unstructured data (Pal, 2015), while historian systems such as OSIsoft PI are best positioned for time-series data.

Analytics

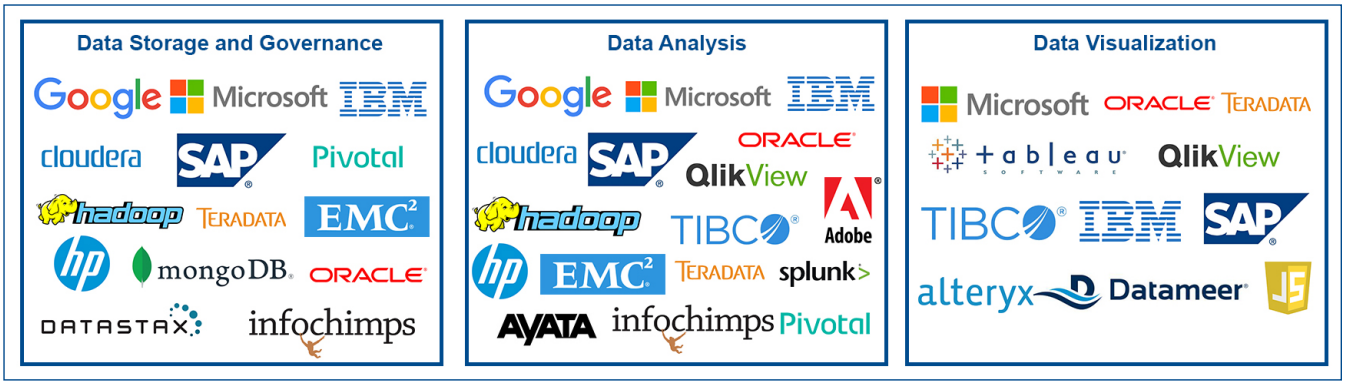
Once the data is stored, the data must be analyzed to understand it and produce usable information. This is the second layer of big data management. Analytics can correlate events, calculate metrics, prepare statistics and recognize patterns within the data. However, before the analytical goals can be achieved there are several challenges that need to be typically addressed. Real-time data is not always be correct and there may be a requirement for cleansing the data. Additionally, real-time data may need to be merged with other structured or unstructured data to provide context-specific information. These are some of the challenges that are addressed in addition to providing the analytical capabilities as part of this layer. As with data storage, there are several options available when it comes to big data analytics. Apache's Hadoop platform provides tools such as Pig (a high-level language for data analysis programs), Spark (suited for in-memory processing algorithms), and MapReduce (framework for writing applications that process large amounts of data) (What is Apache Hadoop?, 2015). Traditional big players such as Microsoft, SAP and Oracle also provide data analytics capabilities while there are several newer names on the list of popular service providers, such as Cloudera and Pivotal (16 Top Big Data Analytics Platforms, 2014). Vendors such as Microsoft aim to create frameworks open to leveraging capabilities of other big data analytics tools such as Hortonworks, while still providing existing BI solutions to cleanse (SSIS), analyze (SSAS) and report data (SSRS) (Understanding Microsoft big data solutions, 2016).

Visualization

The final step of the value chain for consuming and analyzing large amounts of data is visualization. This stage isn't always required since some tools provide a native interface for accessing data. However, when analytical data needs to be shared there is often the need for visualization tools. The suite of tools in this layer may augment functionality of analytical tools and by providing basic statistical analysis capabilities. Tools in this layer include Microsoft Power BI, Tibco Spotfire, Tableau, Qlikview, SSRS, etc. Versions of Microsoft Excel 2013 and higher include PowerPivot, which provides on-demand analytical capabilities as well (Russo, 2013).

Vendor Summary

The figure below broadly classifies some of the key players in the industry based on the capabilities offered. While this diagram is by no means comprehensive it highlights some of the prominent vendors. In the data visualization space, JavaScript is noted since there are several freely available visualization frameworks based off JavaScript.



What Can Big Data And IoT Do For You?

Several case studies have been performed and there are innumerable benefits specific to organizations that can be achieved from harnessing large amounts of raw data. Here are some examples.

CenterPoint Energy (Power and Utilities)

CenterPoint Energy operates in the power domain, supplying electricity and natural gas to several states in central and southern United States. CenterPoint implemented a smart grid program by using IoT and smart meters to automate and increase data collection from smart meters. With its SAP HANA big data solution, the company has dramatically reduced time to diagnosis for events like power outages. It can quickly determine if the outage is due to something inside a customer's home or a broader power outage. Furthermore, the company has saved millions of dollars by reducing fleet fuel consumption and thousands of tons of fuel emissions by automating data collection. With its IoT and big data framework, the company can analyze complex problems and visualize information to troubleshoot problems faster and improve safety and customer service (Mullich, 2013).

Royal Dutch Shell (Upstream Oil and Gas)

Royal Dutch Shell leverages advances with the IoT and big data solutions to realize several benefits in its upstream drilling and exploration program. The solution uses fiber optic cables which transmit high frequency sensor data to private servers maintained by Amazon Web Services. This allows Shell to accurately determine the characteristics of prospective oil fields. The firm also uses its big data solution to ensure machines are working properly as machines are fitted with sensors collecting performance data to support predictive maintenance. Finally, Shell brings several different data streams such as economic indicators and weather patterns into its big data solution and uses complex algorithms to determine demand, allocate resources, and set prices at pumps in the most optimal manner considering all these external inputs (Marr, 2015).

Carnegie Mellon University (Education)

Carnegie Mellon University, based in Pittsburgh, Pennsylvania, is a research institution with seven globally recognized schools and colleges including the Carnegie Institute of Technology. As part of its big data architecture, Carnegie Mellon University uses the OSIsoft PI system to store and analyze data from sensors in its technical infrastructure where large volumes of data are collected. The university also uses Microsoft Azure in a cloud configuration to gain real-time insights through predictive analytics and machine learning. Using its big data solution, the organization has been able to significantly reduce building maintenance costs and energy costs. By leveraging Microsoft Azure machine learning it has been able to better predict and diagnose system faults which results in more efficient operations (Carnegie Mellon Sees a Way to Cut Energy Use by 20 Percent with Cloud Machine Learning Solution, 2014).

City of Barcelona (Administration)

The city of Barcelona, Spain achieved significant benefits in civic planning and infrastructure with its implementation of a big data solution. By implementing Microsoft Azure, Azure HDInsight Service, and Microsoft SQL Server, the city was able to consume and analyze the variety of information sources and create a portal depicting key performance indicators for the city services such as bike usage, number of people using bus routes, etc. The city was also able to gain insight into the optimal location to situate emergency centers. By gaining insights from users' Twitter and Facebook comments around organized events, the city was able to study the entertainment and food venues, citizen interest and satisfaction, people mobility, and detect incidents (City of Barcelona – City Deploys Big Data BI Solution to Improve Lives and Create a Smart-City Template, 2015).

Contact us today to learn more about our case examples and how we can help your organization.

Emerson
Operational Certainty Consulting
1100 Louis Henna Blvd
Round Rock, TX 78681 USA
+1 888 889 9170

www.emerson.com/opcertcon