MAPR INDUSTRY GUIDE FOR MANUFACTURING

MAPR INDUSTRY GUIDE FOR MANUFACTURING

INTRODUCTION

To say that digital technologies are changing the face of manufacturing is an understatement. Led by big data analytics, miniaturization of sensors enabling the Internet of Things (IoT), easy-to-deploy machine learning toolkits, and the digitalization of logistics and the entire supply chain, a century-old manufacturing industry is transforming for the better. The challenge for both business and IT leaders is to skillfully manage this transformation, being mindful of existing technology investments, requirements of future investments, and the impact of digital transformation on the current workforce. The days of competing in manufacturing largely on the basis of labor costs are forever fading in the rearview mirror. Success is coming to those who are the most adept at newer technologies and opportunistic in the face of highly dynamic markets and global economic forces.

As shown by the table below, the smart factories of tomorrow are more about data than they are about assembly lines, inexpensive labor, and manual processes. Going forward, driverless finished products to customers and end users. Predictive maintenance technologies will mostly eliminate production glitches owing to failed equipment – at least for those manufacturers with a resilient, robust, underlying data platform to support such data-intensive operations in real time.

Manufacturers have been successfully using data to increase efficiency and quality but are finding that lean production and cost cutting are no longer enough to remain competitive. The goal today is to integrate and gain insights from data across their complex, global, and often fragmented supply chains.

DATA-DRIVEN TRENDS: FROM YESTERDAY TO TODAY TO TOMORROW

YESTERDAY Factories with Monolithic Apps	TODAY Connected Factories with Connected Apps	TOMORROW Smart Factories with Data-Driven Apps
Manual Assembly	Semi-Automated Assembly	Robot-Based, Fully Automated Assembly
Inadequate Visibility into the Supply Chain	Partially Connected via Supply Chain Management Software (SCMS)	Automated Supply Ordering Based on Depletion Monitoring
Unconnected Equipment; Manual Prognosis	IoT-ized Equipment	Predictive Maintenance and Proactive Health Monitoring
Manual Business Processes for Selling Goods and Services	E-Commerce and E-Tailers Enabled via the Internet	Quicker and Seamless Selling, Using Bots & Voice-Services
Handover to the "Delivery Guys"	Connected Delivery Vehicles with Accurate Delivery Alerts to Consumers	Drones for Delivery; Autonomous Delivery Vehicles

A recent global report by PwC clearly demonstrates how digital transformation has descended on the manufacturing community. In this report, 91% of industrial companies are investing in creating digital factories in the geographies surveyed. But only 6% report being "fully digitized," meaning that a great deal of energy and effort will be dedicated to digital transformation in the next few years. Of those investing in digitalization, 98% fully expect to increase efficiency with various digital technologies, including predictive maintenance. And perhaps most encouraging, 90% of those surveyed said digitalization of manufacturing offers more opportunities than risks, which should help with securing proper funding for digital transformation efforts.

What the digital transformation pioneers and early leaders have discovered is the relative inability of many existing digital platforms to 'get the job done' in digital manufacturing. As this guide will show, the IoT is being widely adapted and leveraged in manufacturing in a number of use cases, such as predictive maintenance and logistics. Most of the data generated 24/7 by IoT devices is unstructured data. Conventional platforms and databases simply cannot aggregate, store, process, and 'normalize' this data in an efficient manner. Newer and more data-agnostic platforms are needed.

The same holds true with the multitude of manufacturing applications for advanced AI and machine learning technologies. Again, these technologies demand platforms capable of working fluidly with enormous and continuous data volumes. Thus, bridging the gap between existing IT investments and a fully modernized infrastructure is usually Job #1 with manufacturing CIOs today.

The role of digital technologies and of IT in manufacturing is perhaps best summed up by Mattias Ulbrich, the long-time CIO of AUDI and winner of many major digital transformation awards: "The prevailing wisdom that IT follows business no longer fits. The advent of digital means we can use new approaches and technologies to establish new business cases, models, and services more quickly. The job of IT is to proactively support the innovative transformation of the company as a whole."

This ebook will clearly demonstrate that, going forward, global manufacturing is all about data-driven growth, not simply competing on labor costs or relying on decades-old models to execute certain business processes. Without the proper platforms for analyzing the highly complex data sets in huge volumes, a predictable manufacturing business in an environment of market volatility, business uncertainty, and constant change is just not possible.

MANUFACTURING INDUSTRY TRENDS

Deloitte minces few words in a major report on information technology in manufacturing: "The world is undergoing a fourth industrial revolution, one fueled by smart, intelligent automation and marked by an unprecedented, *exponential* pace of change." Addressing the current crop of leading edge manufacturers, the report notes that, "Many leading 21st century manufacturers are converging digital and physical worlds in which sophisticated hardware combined with innovative software, sensors, and massive amounts of data and analytics is expected to produce smarter products, more efficient processes, and more closely connected customers, suppliers, and manufacturers."

Throughout the report, Deloitte frequently refers to the recurring theme of the role of big data analytics as manufacturers seek to transform their businesses. The biggest manufacturing spenders, Deloitte notes, are in the U.S. and Western Europe, the regions stung the most by the emergence of China as the world's leading manufacturing country. Much of China's emergence

has owed to its abundant source of relatively inexpensive labor and state-controlled capital markets. The West is fighting back with big data.

Another recent report from PwC reinforces the notion that big data analytics lies at the heart of digital transformation in manufacturing. This report states boldly, "Artificial intelligence and data analytics are driving the digital factory." The report, based on extensive interviews with global manufacturers, shows that more than half of them are already using smart algorithms to boost operational performance. PwC maintains that manufacturers' adoption of machine learning analytics that improves predictive maintenance will jump 38% in the next few years. Further, machine learning will reduce supply chain forecasting errors by half and lost sales by 65% with better and more reliable product availability.

As shown in study after study from leading global consultancies and major think tanks, big data analytics along with machine learning algorithms, AI techniques and advanced hardware-agnostic platforms capable of processing any data format are helping manufacturers blaze a pathway to new business models. Along the way, they are optimizing long-standing manufacturing operations at the factory floor level and through the entire manufacturing process.

But these new business models and advanced technologies have also been shown to put excessive strains on existing ERP, CRM, and PLM systems. Manufacturers are rightfully loath to undertake massive rip and replace IT strategies, and instead find themselves searching for newer platforms capable of integrating as seamlessly as possible with existing investments.

Meanwhile, manufacturers are very mindful about cybersecurity ramifications as they go full bore into the creation of digital factories. One potential source of data leakage and, therefore, point of attack are the thousands of IoT devices deployed or planned for deployment. So, not surprisingly, IT leaders are giving closest consideration these days to emerging big data platforms that can deliver the following:

- Capability to lock down data security across clouds, on-premises as well as at the edge, with management and encryption
- Ability to detect anomalies in sensor data
- Ability to maintain audit trails and data lineage
- Creation of a dashboard that leverages data science and machine learning approaches to stay ahead of attackers

Thus, in manufacturing the shift is on to move from mostly product-centric manufacturing models to instead creating other sources of value, largely leveraging big data to do so. However, one warning sign emerges from the Deloitte report: most manufacturing executives interviewed said their organizations are either not preparing or not moving fast enough to address the many disrupters in their environment. For them, time is not on their side.

KEY STAKEHOLDERS

In talking about the key stakeholders in manufacturing, it is helpful to draw a distinction between the two principal types of manufacturing. Discrete manufacturing generally refers to the production of distinct items, such as cars, trucks, planes, power equipment, medical hardware, electronics, toys, furniture, and so on. By contrast, one can think of process manufacturing as production of end products by combining supplies of raw substances or 'ingredients' using a formula or, in the case of food manufacturing, a recipe. Thus, examples include food, beverages, plastics, chemicals, and so on.

The Leaders

Global manufacturers come in all sizes and span all discrete manufacturing categories. Getting accurate data on the largest of these is dependent upon the availability of public information about them, which in the case of, say, Chinese manufacturers is often state-controlled. And these manufacturers can rank among the largest. In fact, China displaced the U.S. as the largest manufacturing nation in 2010 and by all accounts has widened its lead since then. Those dramatic gains in China were owing to a number of key factors, not the least of which is China's relatively low labor costs. Thus, the hopes for manufacturers in the West to compete against the Chinese rest with the digital transformation of manufacturing, which is very much underway in the U.S. and Western Europe.

These considerations aside, the top ten global leaders in manufacturing from the venerable Fortune 500 list (based on reliable reports of revenues) include seven companies engaged primarily in auto manufacturing, two in electronics, and one in diverse manufacturing. Three are U.S.-based; two each are based in Germany and Japan; and one each in China, Korea, and Taiwan. In descending order, they include Toyota, Samsung, Daimler, General Motors, Ford, Hon Hai, Honda, General Electric, SAIC Motor, and BMW Group. They are also among the most IT- and data-intensive organizations in the world.

Given this global diversity in ownership of the leading manufacturers (which does manifest itself pretty much throughout the top 100 global manufacturers), it is not surprising that global manufacturing today is a potent mix of economics, geopolitics, environmental battles, climate change battles, and the usual business challenges of a highly competitive sector. Each of these various factors is a data-rich trove of potentially competitive advantage, if the data can be consumed, stored, analyzed, and then used for key decision-making. For that reason, manufacturing will continue to evolve into a richer digital-first business, largely unrecognizable by 2020 from manufacturing at the turn of the century.

Distributors/Logistics Firms

Manufacturers must move their products from factories to warehouses and eventually to consumers and other businesses, all in the same timely fashion with which they expect raw materials from suppliers to create these products. While major manufacturers tend to manage their own logistics and distribution, a great deal of this business involves distribution partners. But regardless of who is doing the actual movement of product, the same data-intensive rules apply to maximizing efficiency and profit, just as it does with any other logistics operations.

Third-Party Suppliers and Partners

All midsize and large manufacturers source materials globally in a never-ending search for highest quality at lowest possible cost. Real-time communications between a manufacturer and supplier are table stakes to ensure production inventories closely match production schedules (which, in turn, are matched to a broad host of economic, marketing, and sales-related data). However, third-party suppliers and partners also include a broad range of pure digital plays, including all manner of cloud providers, notably SaaS and PaaS but also IaaS providers as well.

University/Government/Research Providers

These stakeholders are important providers of data from historic global market and sales data to real-time weather data affecting logistics to geopolitical indicators on global economic trends – among other data. Increasingly, much of this data, like sensor data, is unstructured, including social media. Thus, it is key that platforms are able to consume and process data of all types efficiently, which most traditional systems were never designed to do.

Regulators/Government

Like any other vertical sector, manufacturing must comply with a number of state, national, and international regulations and compliance edicts. In particular these days, this often means the ability to show compliance with the rising tide of data privacy regulations, including the General Data Privacy Regulation (GDPR) for any organization doing business with citizens in the European Union and the related APPI in Japan and PIPA in Korea. In addition, various governmental regulatory agencies aggregate large volumes of digital data from manufacturers in specified formats on employment, worker safety, output, supplier payments, cybersecurity efforts, and in the case of defense and associated manufacturing, detailed manufacturing and customer reports.

Moreover, in order to ensure compliance to these regulations, manufacturers require a robust data governance and security policies built into their business processes. A data platform that understands these regulations is now a necessity for manufacturers.

New Disrupters

Clearly, the most digitally advanced and most watched new automobile company in the last decade is Tesla, which also happens to be among the most highly digitized – from its factory floors (perhaps the most automated) to marketing models (no dealers per se, rather showrooms located often in malls) to sales techniques (more people buy Teslas via the internet than buy any other brand). Tesla is a manufacturing company, through and through.

Amazon is also testing the manufacturing waters. Last year the company was awarded a patent for an on-demand manufacturing system designed to very quickly produce custom clothing, and other products, once a customer order is placed. This move was widely seen as a first step in a process of Amazon leaping into the manufacturing-as-a-service business, just as it has come to dominate the IaaS business. Chinese giant Tencent is seen as another potential manufacturing disrupter. The key point here is that digital technologies can effectively pose insurmountable barriers to new companies entering manufacturing sectors dominated today by massively large firms. This reality significantly raises the bar, and anxiety levels, of entrenched firms when it comes to leveraging data competitively.

KEY DATA SOURCES AND STANDARDS

One of the most compelling factors of the data associated with the contemporary manufacturing industry is its enormous and fast-growing volume. Much of that volume and most of its growth going forward is semi-structured and unstructured data – precisely the data formats on which traditional database platforms choke and often grind to a halt. In addition, the data most useful to the business is a combination of historic data and very large volumes of real-time data generated and sent to big data platforms, 24/7 year-round.

Research aided by study after study shows a clear link between big data analytics and profitability as well as growth in the manufacturing sector. In fact, one major study over a ten-year period of some 250 global manufacturers shows manufacturers leveraging digital technologies to effectively become faster, cheaper, and more flexible – precisely the characteristics needed to prosper in a world of continuous change.

Here below are the key data sources that will pump enormous and growing volumes of data into manufacturers' data platforms. The ability of these platforms to aggregate, store, process, and analyze this data will be intrinsically linked to overall success – or the lack thereof.

Factory Floor Operations/Machine Logs/Factory Operator Notes/Equipment Log/ Operational Applications

This includes data from programmable logic controllers (PLCs), manufacturing execution systems (MES), and human machine interface (HMI) terminals, to name a few examples. Given the highly distributed nature of manufacturing, much of this data originates from remote or distant sites, the so-called 'edge.' Locally processing this data and then uploading it to central sites is key to speedy overall data analysis, decision-making, and action. Moreover, sources such as machine logs and equipment logs tend to follow a certain format but are examples of unstructured data.

IoT Devices

As with so many other vertical markets such as oil and gas, the proliferation of IoT devices continues apace in manufacturing, where they are proving particularly useful in a predictive maintenance role for all manner of machinery. They are used further to monitor performance of heating and cooling systems (HVAC), sending large volumes of semi-structured and unstructured data to big data platforms for real-time analysis.

Marketing and Sales Systems

This data for the most part is structured today, though increasingly marketing teams are tapping into scores of new data sources, such as social media sentiment data, which is generally unstructured.

Distributors/Logistics Partners

These organizations ingest and generate business-critical data that generally is structured and designed for instant analysis, report generation, and decision-making. A vital feature of any manufacturing big data analytics platform is the ability to work seamlessly with GPS data for better fleet management and to integrate this data into other solutions to drive faster decisions.

Third-Party Suppliers/Partners as well as Retailers and E-Tailers

Most data sent back and forth with third-party suppliers is structured and tailored for various inventory and just-in-time manufacturing systems.

Real-Time Cybersecurity Threat Data

This absolutely essential data arrives 24/7 and often requires immediate action. Existing security tools require logs to be massaged into a specific format and cannot analyze or act on the information in real time. While these tools have their place, they are far more effective when used in combination with a highly robust data platform with flexible machine learning and support for high-speed ingestion.

University and Other Research Organizations

While less dependent on pure research than sectors like pharmaceuticals, manufacturers do maintain relationships with various research outlets. This data can find its way into marketing and sales systems as well as into factory operational systems.

Crowdsourcing

Virtually all this data is acquired from both major and secondary social media sites and thus is both semi- and unstructured for the most part. Both marketing and sales in manufacturing organizations are the biggest consumers of this data.

Trade Organizations and Associations

These organizations generate and distribute various reports on various aspects of manufacturing (economics, geopolitical dynamics, supply chain data, raw materials forecasts, labor statistics,

etc.). Too numerous to mention them all, some major ones include the National Association of Manufacturers, the Alliance for American Manufacturing, various sector-specific associations such as the Aluminum Association, and horizontally-focused organizations such as the Production and Operations Management Society.

USE CASES

As Tata Consultancy Services says, "The range of big data use cases in the manufacturing industry is limited only by available data and imagination." This global consultant asked manufacturers to rate a list of perceived big data benefits on a scale from 1 (lowest) to 5 (highest):

- Product quality and defects tracking – 3.37
- Supply planning – 3.34
- Manufacturing process defect tracking – 3.32
- Supplier, components, and parts defect tracking – 3.11
- Supplier performance data to inform contract negotiations – 3.08
- Output forecasting – 3.03
- Increasing energy efficiency – 2.97
- Testing and simulation of new manufacturing processes - 2.88
- Support for mass-customization of manufacturing – 2.75

Below is a deeper dive into individual manufacturing use cases for big data and advanced analytics.

Preventative and Predictive Maintenance

From assembly lines to warehouses to trucks and other delivery vehicles, modern manufacturing is characterized by huge amounts of industrial hardware. It can take surprisingly little, such as a seized motor or an electrical fault in a power backup system, to bring an entire assembly line to a sudden and costly halt. Virtually every piece of this equipment spews out a constant stream of data related to the overall health of this vital gear. The volumes and complexity of this data are far beyond the computing and analysis capabilities of traditional database systems. Such systems must not only be able to properly integrate this data into current systems, but also be able to leverage advanced AI and machine learning solutions to accurately and consistently *predict* equipment failures before they occur, then trigger timely corrective actions. In fact, the capacity of deep learning to analyze enormous quantities of high-dimensional data is taking preventative maintenance to entirely new levels of excellence. For example, AI can extend the life of a large delivery truck beyond what is traditionally possible by combining truck model data, maintenance histories, IoT sensor data such as anomaly detection on engine vibration data, and images and video of engine condition. Millions of dollars of cost from preventable downtime are at stake here.

Assembly-Line Quality Assurance and Real-Time Parts Flow Monitoring

Using an array of sophisticated cameras, vibration sensors, thermal imaging sensors, and light sensors, manufacturers can take measurements on the assembly line, identify defects as early as possible, identify any potential process or design flaws, and suggest corrective actions in real time. Since defects are typically the result of many factors, analyzing long histories of assembly line sensor data can find subtle anomalies that signify product flaws. A robust data platform is capable of storing long histories of sensor data while also enabling high-speed, real-time, early-warning analytics that correlate real-time measurements with other disparate data, then compare them to quality models. Manufacturers can have a real-time view into their production process.

Supply Chain and Logistics

Perhaps nothing is more irksome and troubling to manufacturers than to create finished products with maximum efficiency and speed, only to have those products stuck in delivery bottlenecks. A capable data platform accurately tracks the movement of vehicles and products to identify the costs of various transportation and process options. Such platforms must be capable of analyzing historical, time-stamped location data, enabling manufacturers to calculate optimal delivery routes and enable dynamic rerouting to minimize the impact of arbitrary obstacles like traffic, energy prices, and weather. Businesses can also leverage the optimal delivery system as a revenue-generating basis for premium/expedited delivery services to consumers as these services become increasingly fashionable. Essentially, big data allows supply chains to boost both service and efficiency by syncing product and external data with business decisions. Despite this promise of big data-driven logistics, McKinsey found that most organizations have yet to unlock this potential, so much remains to be done.

Security

Manufacturers of all sizes need to identify events or patterns that could indicate an imminent security threat or cyberterrorist attack in order to keep their personnel, property, and equipment safe. Predictive analytics plays a central part in identifying patterns and detecting these threats in advance. Manufacturers often use Security Information Event and Management (SIEM) software as point solutions to help identify threats, develop anomaly detection techniques, and reduce the likelihood of such incidents. AI/machine learning solutions are particularly adept at just this sort of early detection and early warning. But such solutions often don't leverage all the data silos (across clouds, on-premises, and edge environments), resulting in missed or delayed threat detection. A data platform capable of handling large data volumes associated with AI/ML as well as the wide variety of data types and data sources involved has now become a necessity for manufacturers.

Market Pricing and Planning

Business leaders in manufacturing firms know full well that this step is the key to maximizing profitability. Simply stated, a stable and reliable data platform in manufacturing can analyze the dynamics of global demand, integrate supply factors, even integrate social media and geopolitical data (such as today's rapidly shifting tariff environment) to bring products to market quickly and then price them accordingly. In fact, customer analytics, operational analytics, and to a lesser extent fraud and compliance, along with new product innovation, have emerged as the most popular big data use cases for manufacturing sales and marketing groups.

Smart Meter Analysis

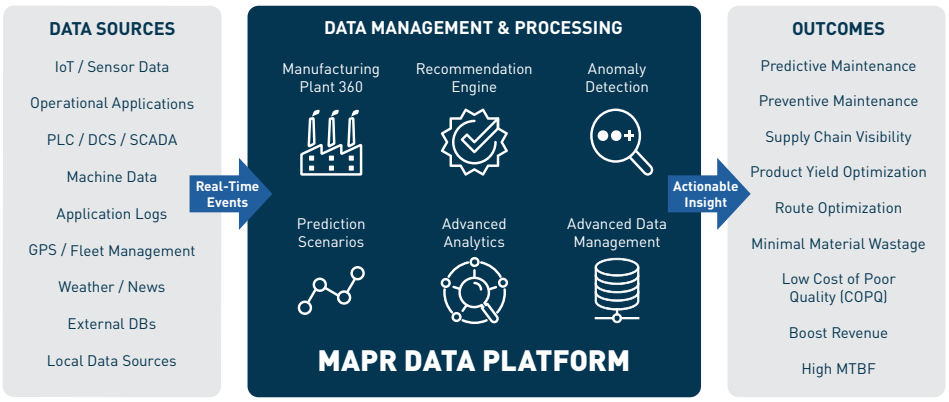
A capable data platform allows manufacturers to extract very granular information from so-called smart meters that aggregate data on energy utilization at a per-site or per-installation basis in order to better identify both pricing and utilization recommendations.

Product Configuration Planning

Related to market pricing and planning, product configuration planning can help accelerate production by enabling faster delivery times of countless product configurations. Advanced pattern analysis, available in the better big data platforms, essentially predicts the most popular product configurations, based on numerous data feeds, including unstructured social media data. Customization is arguably associated with higher factory floor costs. Leveraging data analytics helps manufacturers with a snapshot of the effect customization will have on the factory floor as well as on production schedules, machinery availability, and inventory.

MAPR INDUSTRY ARCHITECTURE

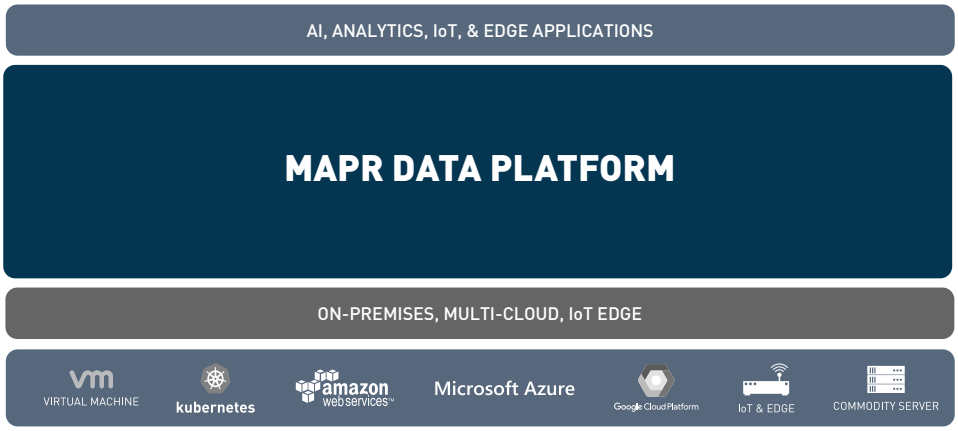
As has been shown, IT efforts in the manufacturing sector are driven today by a need to harness formidable big data initiatives in the face of highly complex and dynamic global market conditions. It is specifically to resolve the many issues and unique challenges facing the manufacturing sector that MapR has tailored its data platform, enabling firms in this sector to offer truly data-centric applications very cost-effectively. The MapR Data Platform allows tapping into a highly complex myriad of manufacturing-related data, regardless of where the data is located.



MapR Data Platform for Manufacturing

By pursuing a data-centric vision for a new generation of applications, MapR has created a data platform that converges the management of data of any size, speed, and format. It was for this work that we were awarded a patent for the MapR Data Platform.

The MapR Data Platform is designed to deliver utility-grade data services and commercially supported, open source innovations to development teams, IT operations, business analysts, and data scientists. Open source technology provides a fantastic creative force when looking to tackle the sophisticated new challenges that big data – and especially new data – can uncover.



Without a converged data platform, critical information can get stuck in "data silos" and an inefficient use of hardware resources can result in a costly "cluster sprawl" of underutilized servers and storage. With the MapR Data Platform, businesses can enjoy real-time insights based on secure, protected, high-fidelity data.

The MapR-XD Cloud-Scale Data Store is the only exabyte-scale data store for building intelligent applications with the MapR Data Platform. MapR-XD was built in response to rising user demand for intelligent applications, capable of automating real-time operational decisions based on application of deep analytics insights. MapR-XD includes the MapR multi-temperature Global Namespace and data management in the form of security, compression, snapshots, multi-tenancy, and self-healing. It is delivered either via disk or flash.

With a consistent focus on the integrity of data, MapR has created a hardened, clustered platform that can withstand multiple hardware failures, data center outages, and malicious attacks and intrusions from cybercriminals. Many proven methods of data protection – such as failover, redundancy, and access controls – are built into the MapR Data Platform.

Game-changing big data applications and analytics will continue to rely on open source software. As a company founded in and contributing to the open source world, MapR continues to define enterprise requirements and best practices for successfully using the latest open source innovations. We deliver monthly updates to open source software packages to ensure you have the latest innovations.

MAPR DIFFERENTIATORS IN MANUFACTURING

In the face of the myriad challenges facing manufacturers as they attempt to build smart factories, the MapR Data Platform and the MapR architecture for manufacturing enables manufacturers to improve production profitably, grow revenues, and build better operational predictability by tapping into all datasets and sources. This includes sensor data from thousands of IoT devices, factory floor data, data from the edge from remote production sites, text files, social media data, and other data, regardless of its format and structure. The MapR Data Platform then converges this data onto one hardware-agnostic platform for processing and analysis, regardless of where the data is located.

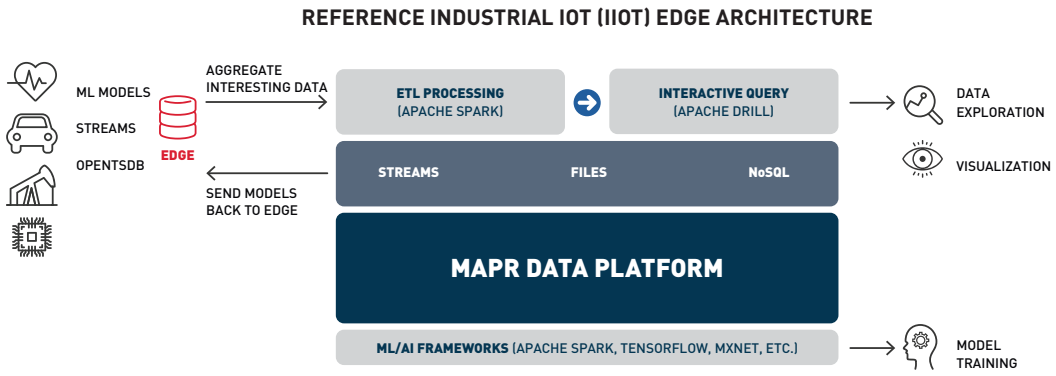
With MapR, manufacturers have already boosted productivity, increased revenues, and better managed the highly dynamic supply-demand curves, characteristic of today's industry. Specifically, the MapR architecture for manufacturing delivers the capability to:

- **Boost quality assurance efforts.** With the MapR Data Platform, manufacturers can take measurements of work-in-progress to find defects as early as possible while identifying process or design flaws. They can also store long histories of sensor data, while enabling faster, early warning analytics to correlate real-time measurements with other isolated data to improve product quality and yield. And with a high-performing NoSQL database management system and an event streaming system built into the platform, test data is made available and analyzed in real time to reveal previously hidden insights.
- **Mitigate supply chain risk.** Manufacturers are constantly required to analyze historical and time-stamped location, map out potential delays, and dynamically optimize delivery routes to ensure production isn't interrupted by traffic, energy prices, and weather. The MapR Data Platform allows them to ingest GPS data, refrigeration data, and truck route information, process it, and analyze this data, thereby helping them reduce transportation costs and minimize waste.
- **Make easy work of data access.** Loading data into MapR is as simple as copying data to a standard file system. Direct Access NFS lets business analysts access data without special tools, so they can read and write files in MapR with their existing file system-based analytical applications. The integrated security in MapR ensures that users can access only the data they are authorized to access.

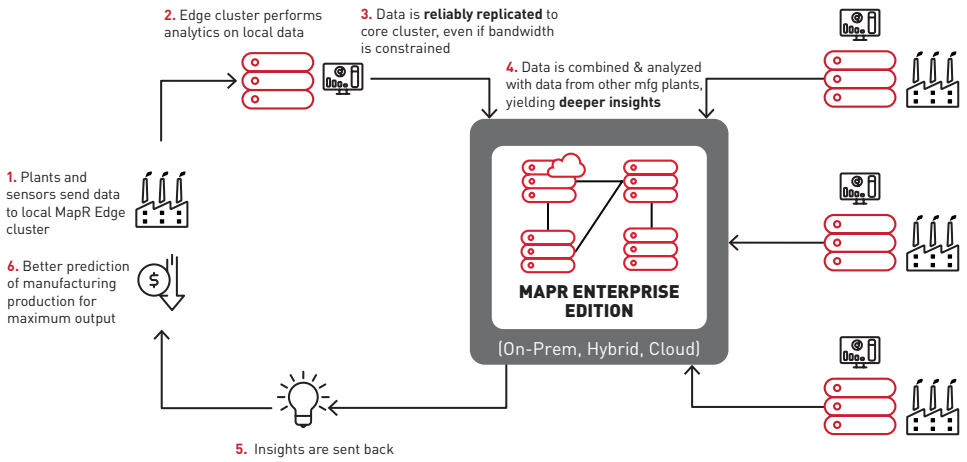
- **Enable business continuity.** Because manufacturing environments demand ongoing analysis of sensor data, MapR ensures continuity with its proven track record of reliable production deployments. MapR provides integrated high availability (HA), data protection, and disaster recovery (DR) capabilities to protect against hardware failure as well as site-wide failure.
- **Optimize small file support.** Most distributions of Apache Hadoop are intended only for large files, while the MapR Data Platform is ideal for handling files of all sizes, including many small files. This is critical in an operational intelligence system in which, say, IoT sensor data can be stored and analyzed as many independent small files coming from many different sources.
- **Deliver massive scale and hardware efficiency.** MapR was designed for high performance, with respect to both high throughput and low latency. This ensures MapR can keep up with the high volume and high velocity of incoming sensor data from thousands of sources. In addition, a fraction of servers are required for running the MapR versus other Hadoop distributions, leading to architectural simplicity and lower capital and operational expenses. Massive scalability ensures that adding new nodes to the cluster can effectively accommodate rapidly growing data volumes.

MapR Edge and IoT Data

In today's increasingly distributed manufacturing environment, IT faces tough challenges when it comes to bringing data from remote assets, spread across a large manufacturing plant or sites. With a small footprint and reliable replication capabilities, MapR Edge is ideally suited for space- and bandwidth-constrained manufacturing environments, hence addressing issues related to local data collection, very common in manufacturing facilities. Used in combination with a core MapR Enterprise deployment (on-premises or in the cloud), MapR Edge empowers organizations to securely process data locally, quickly aggregate insights on a global basis, and ultimately push intelligence back to the edge for faster and more significant business impact.



MapR Edge is already proving itself as an ideal solution to bringing high-level, real-time analysis of IoT data to local sites traditionally constrained by bandwidth and storage capabilities. Essentially a small-footprint edition of the MapR Data Platform, MapR Edge pushes the benefits of convergence to remote computing sites, thus overcoming the IoT data analysis obstacles that otherwise hamstring local decision-makers. MapR Edge is the answer to the question, "Can we do limited analytics on-board the machines or equipment themselves instead of transferring all data to a central cluster with a flaky internet connection?" The answer is "Yes."



As noted by Jason Stamper, lead analyst for Data Platforms and Analytics at 451 Research, "The use cases for IoT continue to grow, and in many situations the volume of data generated at the edge requires bandwidth levels that overwhelm the available resources. MapR is pushing the computation and analysis of IoT data close to the sources, allowing more efficient and faster decision-making locally, while also allowing subsets of the data to be reliably transported to a central analytics deployment."

MapR Edge helps manufacturers make sense of the massive quantities of data generated by IoT sensors. MapR Edge also provides much-needed unified security, so vital in an era of cyber threats. This security approach protects data stored at the edge and in motion, while reducing IT complexity with a consistent security framework. MapR Edge also simplifies code development, owing to its standards-based interfaces.

Finally, the MapR Edge-to-Cloud File Migration service, deployed to each edge site, is on continuous alert for new files, which are automatically transferred to the cloud. The intelligent use of MapR metadata services ensures performance and reliability.

PRODUCTION EXAMPLES

Leading Chip Manufacturer – Gaining High Availability While Lowering Costs

A U.S.-based semiconductor firm with 200 global locations found itself in need of a solution to store, access, and quickly process vast amounts of data. Most of it originated from archives of all test data from test houses and foundries worldwide – often in a wide variety of formats. This data was available for advanced analytics, but the manufacturer's existing systems were creating major bottlenecks in both importing and exporting the data. The system also had zero high-availability features, necessary to complete reporting of the analytics to customers.

After examining competing major big data platform providers, the manufacturer elected to deploy the MapR Data Platform. They did so, owing to its lightning-fast data import/export capabilities, high availability, network file system (NFS) capabilities, and the potential to lower infrastructure costs. As one senior manufacturing manager noted, "Because MapR works with Hadoop, we get an entire ecosystem with machine learning tools and packages. By using MapR and its excellent NFS, we are able to give our customers access to the data directly from the source through NFS for machine-to-machine programmatic access."

To date, the manufacturer has achieved data import/export speeds an order of magnitude faster than that of their old system. Given the real costs of test time, the vastly speedier availability of data allows engineers to make faster decisions that can more quickly be fed to the factory test floor, saving money. In addition, the manufacturer has estimated that its ability to run MapR on commodity hardware is costing a third of what they had paid out formerly. Those savings doubled when the manufacturer added high availability, which since the deployment has delivered zero downtime.

MAG45 – Data Lake Slashes Customer On-Boarding Time by 50%

This 70-year-old Netherlands-based industrial integrator provides maintenance and repair services to help customers improve both productivity and efficiency of production lines and reduce TOC for production operations. MAG45 embarked on a project to automate the manual processes of managing and accessing the product data of more than five million products from more than 3,000 suppliers. However, its initial efforts were beset by various challenges, not the least of which were:

- Its traditional SQL databases could not easily handle non-standardized product data.
- Storage and retrieval of increasingly large volumes of data became very time-consuming.
- The process of matching customer data to the MAG45 database was highly manual and therefore unscalable.
- It could take several months for MAG45 to set up a supply chain for a new customer.

MAG45 worked with a MapR partner, Anchormen, to develop a Cisco UCS-based big data solution, designed to consolidate all of MAG45's semi-structured data into a data lake. Anchormen was able to demonstrate MapR stability, scalability, embedded security, and ease of integration of new data sources using NFS.

The resulting business benefits have been substantial. MAG45 has reduced the time needed to implement new supply chains by 50%. Because MAG45 now has more insight into its own data, it expects savings in both supply chain and procurement costs, which ultimately will be reflected in improved services and lower TCO for its customers.

AUDI – More Efficient Design and Production

With a world-renowned auto design and manufacturing legacy going back 85 years to 1932, AUDI, now part of the Volkswagen Group, has always sought greater efficiencies in all aspects of its business. For that reason, AUDI has deployed the MapR Data Platform as a significant part of its automotive development process.

Working with consultant NorCom IT, AUDI engineers deployed a large Hadoop cluster, based on the MapR Data Platform with an initial goal of analyzing vehicle development data. NorCom and AUDI deployed several pilot projects, testing the suitability of big data technologies for data analysis. It was following this extensive testing phase of competing technologies that AUDI chose MapR for its development analysis platform. As noted by Thomas Kriegel, Head of Virtual Vehicle Management, Processes, & Methods for AUDI AG, "The MapR Data Platform provides by far the best technology for the integration of existing applications. In the growing field of data processing, this is an absolute must."

Incitec Pivot Ltd. – Gains 'Explosive' Results

In its illustrious 99-year history, Incitec Pivot Ltd. (IPL) has served many industries with its fertilizers, explosive chemicals, and mining services. Recently, this venerable manufacturer sought to improve upon its explosives' yield and quality, while offering new services to explosives customers that evaluate blasts at mining sites.

But challenges were everywhere. IPL's existing historical systems could not scale sufficiently to handle full-fidelity data. Historian sprawl resulted in up to 6-week data collection delays for the company's data science team. Chronically clogged batch data pipelines hampered meaningful data insights. And analysts were unable to gain the visibility needed into remote mining sites.

The solution chosen to these pressing challenges is the MapR Data Platform. Once deployed at the main factory, the platform began collecting data from programmer logic controllers (PLCs) and sensors to optimize operations and perform predictive maintenance. And at remote mine sites, the manufacturer deployed MapR Edge to essentially bring the extensive benefits of the MapR Data Platform directly to a local site with limited bandwidth and compute resources. As a direct result of the MapR Data Platform installation, IPL has leveraged preventive maintenance features to improve manufacturing efficiency, while developing a new revenue stream through 'explosions as a service' analytics at actual mine sites.

RECOMMENDATIONS

In manufacturing around the globe, the race is on to build smart factories while fully digitizing virtually every aspect of manufacturing. At stake, quite literally, is survival. Note the alarmingly rapid descent of General Electric as it struggles today to recover from the impact of more digitally savvy manufacturing disruptors, among other things. Surely, the era of big data is upon virtually all industries, with manufacturing being at a particular inflection point. But experts at Harvard and elsewhere are quick to issue this caveat: companies are enamored with the concept of big data and its potential to transform the organization for the better but very often lack the proper strategy to make that happen. Look for prudently aggressive manufacturers to do the following as they navigate the risky waters ahead:

- Embrace and promote the essential concept of data-centric strategy and data-driven decision-making, based in large measure on making more complete use of the voluminous data available.
- Deploy edge solutions in order to drive real-time decision-making to where the decisions are going to have the most impact, such as on assembly lines or in equipment.
- Look for a big data solution that enables easy ingestion of any kind of data – including temperature, industrial IoT sensors, weather, equipment logs, process software, and logistics – and fast processing of this data in a single platform. Such a platform solution should count real-time processing capabilities as a major feature, enabling early intervention to avoid supply chain disruptions. Further, the platform should seamlessly combine the latest AI/ML computing and security techniques with existing tools investments.
- Don't think 'rip and replace' when it comes to developing a data-centric strategy. Rather, look to new solutions that leverage the massive investments already made, a solution that requires minimal incremental investment to enable inclusion of all data to deliver the competitive advantage that comes with improved real-time decision-making.