(intel)

Internet of Things (IoT) Delivers Business Value to Manufacturing

Intel demonstrates IoT success using big data analytics to bring cost savings, predictive maintenance, and higher product yields to its own discrete manufacturing processes, enabled by commercially available and open-source building blocks.

Driving a manufacturing evolution using real-time big data analytics

Executive Summary

The volume, variety, and velocity of data[1] being produced in manufacturing is growing exponentially, creating opportunities to gain a competitive edge, respond to changing market dynamics, and increase manufacturing margins, productivity, and efficiency. Equipment across the factory floor is generating thousands of different data types, such as unit production data at multiple levels, equipment operational data, process data, and human operator data, which can be stored and made available for analysis.

Although large manufacturers have been using proven modeling, statistical process control, and data analysis to optimize production for years, the extreme composition and availability of today's data provides opportunities to deploy new approaches, infrastructure, and methods to further improve bottom line P&L. Manufacturing industries are ready to embrace the use of big data, supported by higher compute performance, open standards, the availability of industry know-how, and the growing availability of highly skilled data statisticians fueled by funded academic research. With access to a new level of industrial intelligence, manufacturers can improve quality, increase throughput, have better insights into the root cause of issues, reduce downtime, and optimize the day-to-day maintenance of manufacturing equipment. With these new business values and technology capabilities, manufacturers can further improve supply chain management and introduce the use of customized services to shorten time-to-market for consumer products across geographies.

This blueprint outlines the various building blocks used in an end-to-end Internet of Things (IoT) big data analytics solution pilot in one of Intel's manufacturing facilities. This big data implementation demonstrates how analytics can be applied to factory equipment data, bringing operational cost saving efficiencies to manufacturing processes. With industry collaboration from Cloudera, Dell, Fusionex, Mitsubishi Electric, and Revolution Analytics, this end-to-end IoT big data analytics solution is forecast to save millions of dollars annually for Intel manufacturing along with additional return-on-investment business values. The solutions from these leading vendors are available today for manufacturers to adopt and integrate end-to-end across the factory automation architecture. The solution described is not specific to any manufacturing vertical or business problem and outlines examples of use cases that show how the solutions from these leading vendors work together.

Table of Contents

Key Business Objectives

Manufacturers are continuously driven to increase operational efficiency and reduce maintenance costs to improve the bottom line while driving increased competitive differentiation across the product portfolio. Manufacturing shop floor improvements and efficiencies enable manufacturers to react to dynamic economic and market conditions, improve product yield and quality, and fundamentally allow manufacturers to competitively produce in a more cost-optimized manner.

The solution blueprint addresses the objectives of various stakeholders as follows:

- **Manufacturing Automation Senior Officers** (CTO/ VPs)—Work to identify smart manufacturing strategies and roadmaps that implement data driven improvements across: implementation models, return-on-investment methods, cost reduction targets, production improvement opportunities, advanced monitoring techniques and better control through predictive modeling capabilities, with long-term goals focused on the journey of achieving more autonomous operations (e.g., closed-loop analysis and control).

- **Manufacturing IT Professionals**—Focused on selecting the right infrastructure and building blocks to enable, sustain, and differentiate the business with real-time, predictive operations. The solution needs to close the gap between IT and business by enabling actionable insights, uncovering hidden influencing factors from the wealth of data that are being produced or are already available. To empower business intelligence, the solution needs to have clear ease of use by all levels of people from operators, to shift managers, to business decision makers.

- **System Integrators (Factory Automation and IT/OT Integrators)**—Take advantage of a set of proven, end-to-end implementation recipes and use case examples, which clarify the business value and provide a framework to discuss how the solution can be implemented.

Business Challenge

Automated processes are generating an ever-increasing volume of data, and in many cases, too much data to extract information from it. Existing data may also be sitting in disparate locations on the network and shop floor. Yet, this data may hold the key that enables manufacturers to improve shop floor efficiency, react to changing economic and market conditions, increase product yields and quality, and fundamentally produce more competitive and cost-optimized products.

Manufacturers desire new ways of extending the life and value of their manufacturing assets by using data to prevent costly downtime and move toward better maintenance strategies supported by improved diagnosis and prognosis capabilities.

Clear return on investment (ROI) on technology and solutions requires infrastructure and capabilities that can scale as manufacturing operations become increasingly complex. A key consideration is the ability to statistically use manufacturing data from disparate sources to differentiate existing operations. This is where big data is a game changer.

Big Data in Manufacturing

Big data is characterized by huge data sets with varied data types, which can be classified as structured, semi-structured, or unstructured, as shown in Table 1. Structured data has a formal structure, defined model, and can be formatted into tables, making it relatively easy to manage and process. Structured data has the advantage of being easily entered, stored, queried, and analyzed. Examples of structured data include manufacturing data stored in relational databases, and data from manufacturing execution systems and enterprise systems.

On the other hand, unstructured data, such as images, text, machine log files, human-operator-generated shift reports, and manufacturing social collaboration platform texts, may be in a raw format that requires decoding before data values can be extracted. Semi-structured data is a form of structured data that does not conform to the formal structure of data models associated with relational databases or other forms of data tables, but nonetheless contains tags or other markers to separate semantic elements and enforce hierarchies of records and fields within the data.

In manufacturing, process variability stemming from various factors, like the type of material used, different process

MANUFACTURING INDUSTRIES EXAMPLES	REAL-TIME, SEMI-STRUCTURED DATA	UNSTRUCTURED DATA	STRUCTURED DATA
Semiconductor	Machine builder standards like SECS/GEM, EDA or custom-based on COM, XML	Operator shift reports	RDBMS databases
Electronics		Machine logs	NoSQL
Solar	Sensors (vibration, pressure, valve, and acoustics), Relays	Error logs	Enterprise data warehouse
Machinery		Texts	
Energy	RFID	Vision images	Spreadsheets
Automobiles	Direct from PLCs, Motor and drives	Audio/Video	
Aerospace			
Chemical and Pharma	Direct from motion controllers, robotic arm	Manufacturing collaboration social platforms	
Metalworking		Maintenance Logs	
Food and Beverage	Manufacturing historians (time series data structures)		
Pulp and Paper			
Clothing and Textiles			
Furniture			

Table 1. Manufacturing Data Examples

recipes and methods, and equipment and machine operator differences, creates a high volume of many types of data. This drives a real business need for manufacturers to turn to big data solutions deployed on a scalable infrastructure that can grow with their business and manufacturing require-ments. Manufacturing equipment gen-erates massive data sets (i.e., gigabytes in a week per tool type across multiple tools) as shown in Table 2, making it difficult to store, analyze, and extract useful information from them using conventional methods.

DATA TYPES	DATA SIZE (PER WEEK)	EXAMPLES
Machine parameters and error logs	~5 GB per machine	Used to monitor machine performance: dispense height, placement (x, y, z), belt speed, flow rate, oven temperature, laser power, etc.
Machine events	~10 GB per machine	Used to measure process time: start dispense, end dispense, start setup, and end setup
Defect images from vision equipment	~50 MB per unit or 750 GB per lot	Used to identify root cause of failure modes, defect commonality, defect mapping

Table 2. Data Size Examples

Big Data Analytics in Action

Over the past few years, Intel manufacturing has developed more than a dozen big data projects that have bolstered both operational efficiency and the bottom line. Here are a couple of examples:

Reduced Product Test Times

Every Intel® chip produced undergoes a thorough quality check involving a complex series of tests. Intel found that by using historical information gathered during manufacturing, the number of tests required could be reduced, resulting in decreased test time. Implemented as a proof of concept, this solution avoided test costs of $3 million in 2012 for one series of Intel® Core™ processors. Extending this solution to more products, Intel expects to realize an additional $30 million in cost avoidance.[2]

Improved Manufacturing Monitoring

Data-intensive processes also help Intel detect failures in its manufacturing line, which is a highly automated environment. Intel pulls log files out of manufacturing tools and testers across the entire factory network—up to five terabytes an hour. By capturing and analyzing this information, Intel can determine when a specific step in one of its manufacturing processes starts to deviate from normal tolerances.[3]

The power of big data solutions available today stems from the ability to connect to data sources from disparate locations, and aggregate, store, manage, and centralize various types of data in a cost-effective, scalable manner. Analytics can be performed to merge and correlate these data set types to create business value through newfound insights. In addition to visualizing data to find root causes and describe the relationships between large data sets (descriptive and root cause analytics), analytics and machine learning create the opportunity to examine how machine data can be correlated to yield, quality, and output to predict their behavior, and classify and preempt problems earlier in the overall manufacturing process, thereby providing valuable information to proactively detect processes that are getting out of control.

Solution Benefits

In general, the solution should enable manufacturers to improve decision making, leading to lower product and operating costs, higher yields, less time to resolve quality issues, and the ability to implement IoT-connected and analytics capabilities into legacy and new equipment. For IT organizations supporting manufacturing, the solution provides forward-looking, smart manufacturing operational technology in areas such as scalable data store management, end-to-end security, and business intelligence that generates a competitive advantage.

Supply chains will benefit from the manufacturer's ability to dynamically react to before-and-after sales events via secured business-to-business communication that enables the exposure of APIs that support the creation of new services. Retailers of manufactured goods will see faster time to product availability, better selling costs from manufacturers, quality improvements and, where it matters most, the ability for manufacturers to adjust supply based on rapidly changing market demands.

Consumers will ultimately enjoy lower product prices resulting from manufacturing cost reductions and will be able to continue to demand high-quality, affordable solutions and services.

For the specific implementation described here, the Intel factory pilot program employed IoT technologies and big data analytics, which enabled the following benefits:

- **Increased Manufacturing Throughput**
  The production line can run for longer periods of time due to shortened planned maintenance, resulting from preventive maintenance measures that reduce the number of routine part replacements.

- **Higher Yields**
  A test tool that previously rejected good units when one of its parts malfunctioned is now repaired in advance, using real-time predictive maintenance methods to trigger a response before the process control system can detect errors.

- **Improved Efficiency**
  Image analytics identify good units from defective units in about one-tenth the time of the equivalent manual method.

- **Reduced Downtime**
  Tool failures are avoided through the use of preventive maintenance to identify worn parts that need to be replaced in advance of planned maintenance.

Solution Overview

In collaboration with Cloudera, Dell, Fusionex, Mitsubishi Electric, and Revolution Analytics, the solution has successfully pioneered advanced capabilities that have made tremendous headway in the use of data mining science to solve practical issues common to discrete manufacturing environments, saving Intel millions of dollars through cost avoidance and improved decision making. The overall solutions matrix of end-to-end building block suppliers that support equipment integration, gateway device options, manufacturing network infrastructure components, and localized industrial data center components, with links to enterprise data centers and business-to-business (B2B) cloud services, are increasingly being embraced for operational success.

The building blocks for end-to-end manufacturing that enable manufacturing intelligence from the factory floor to the industrial data center are composed of specific architectural components, as shown in Figure 1. A private cloud infrastructure is common where the data is stored and ingested, and analytics tools are made available to enable both offline modeling analysis and online recommendations to optimize specific tool performance. Linked to the manufacturing subnet, gateway devices are connected to the equipment to extract data using the required protocols. Business intelligence (BI) software solutions are connected through the corporate intranet via VPN for users to access BI results and visualization tools in support of manufacturing operations.

Figure 1 shows the IoT and big data analytics end-to-end solution implementation, including the software and hardware building blocks deployed. On the shop floor, automation assets are distributed over the manufacturing network, including the control-room-focused industrial data center that is resident within the factory floor design. This high-level IoT manufacturing deployment supports small to large, architecturally distributed, and varied data set ingestion requirements for the Hadoop*-based platform. The implementation further supports the data acquisition and aggregation of various types of data from the manufacturing shop floor and the manufacturing network, which opens up the possibilities of visualizing, monitoring, and data analytics for creating new business intelligence.
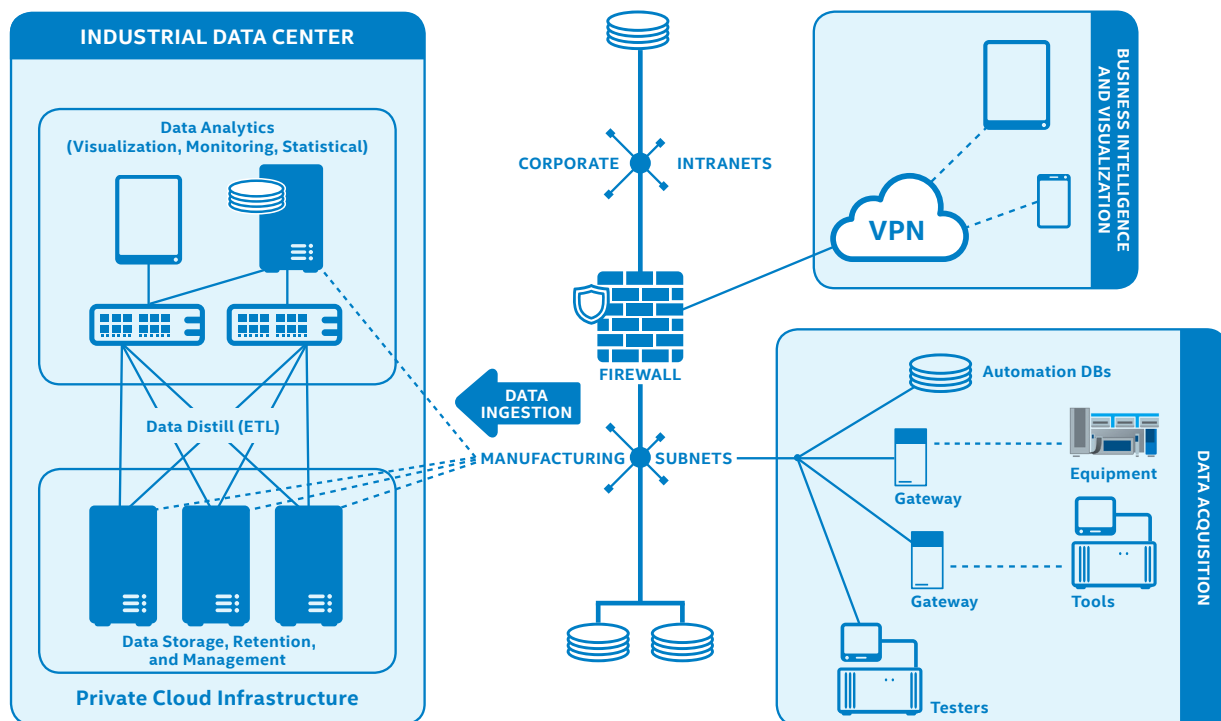

Figure 1. Building Blocks for End-to-End Infrastructure Enabling Manufacturing Intelligence from the Factory Floor to the Industrial Data Center

USE CASES

The following use case summaries describe some of the groundbreaking work and discoveries Intel accomplished in three very distinct problem areas across different types of data and production processes. These use cases provide context, problem-solution examples, and ROI discussion points that could be helpful when planning an IoT and big data analytics project.
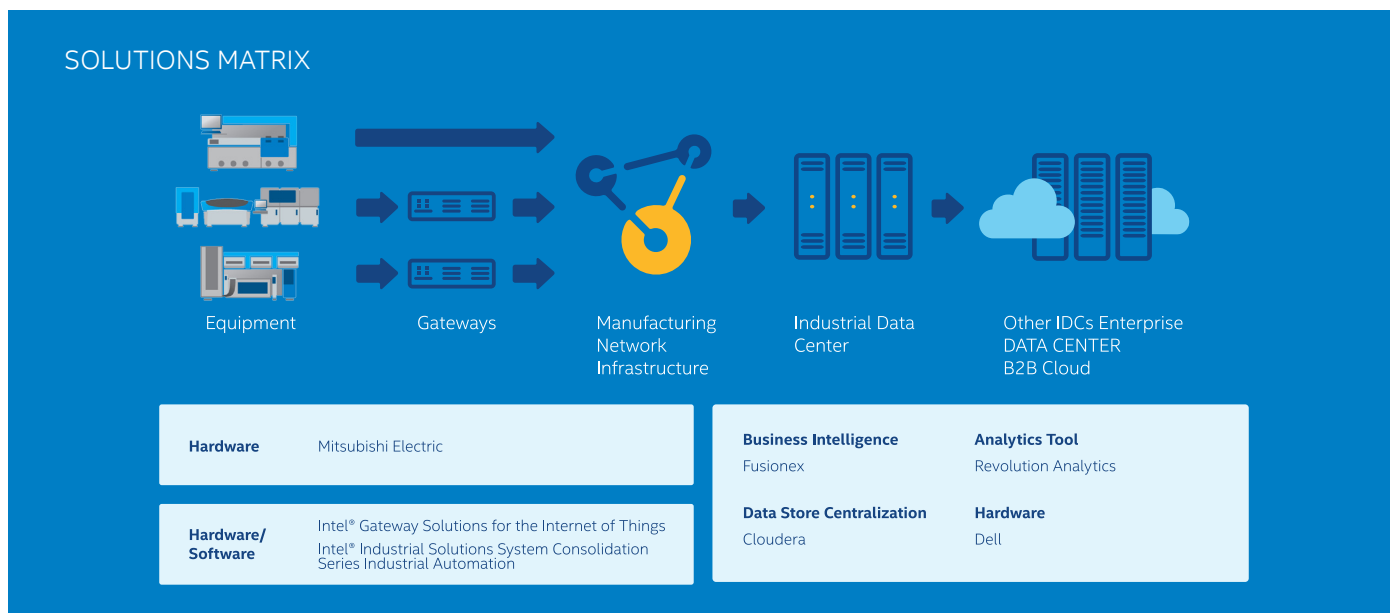
SOLUTIONS MATRIX

Equipment   Gateways   Manufacturing Network Infrastructure   Industrial Data Center   Other IDCs Enterprise DATA CENTER B2B Cloud

Hardware	Mitsubishi Electric
Hardware/ Software	Intel® Gateway Solutions for the Internet of Things Intel® Industrial Solutions System Consolidation Series Industrial Automation

Business Intelligence	Analytics Tool
Fusionex	Revolution Analytics
Data Store Centralization	**Hardware**
Cloudera	Dell

Figure 2. End-to-End Building Block Suppliers

Use Case 1

Reducing non-genuine production yield loss through the monitoring and analysis of machine parametric values and timely replacement of parts before they fail.

Background

The Automated Test Equipment (ATE) used within the Intel factory is a machine designed to perform tests on different devices. When under test, these devices are referred to as devices under test (DUTs). An ATE uses control systems and automated information technology to rapidly perform tests that measure and evaluate a DUT.[4] The ATE is interfaced to an automated placement tool called a handler that physically places the DUT on a Test Interface Unit (TIU) so it can be tested by the equipment.

Problem Statement

Defective TIUs will wrongly categorize good units as bad, which negatively impacts Intel manufacturing operation costs. Defective TIUs are due to mechanical failures of test components used to execute testing scenarios of the DUT. Defective TIUs can cause DUTs to be wrongly categorized, including the rejection of good units. Intel's manufacturing objective is to detect underperforming TIU components prior to failure and repair or replace them before they wrongly categorize units. When a faulty TIU wrongly categorizes good units as bad, the units are scrapped. To avoid such problems, some equipment components were replaced with spare parts during regular preventive maintenance, even if they were still operating properly.

Results and Benefits Yield

The analytics capability predicted up to 90 percent of potential TIU failures before being triggered by the existing factory's online process control system, as shown in Figure 3. This early notification prompted the replacement of defective TIUs before they began over-rejecting good units, thus reducing yield losses by up to 25 percent.[5] In addition, Intel reduced the need to replace spare parts before they failed during preventive maintenance, resulting in an estimated 20 percent reduction in spare parts cost.

Figure 3. Analyzed Test Interface Unit Performance Indicates an Imminent TIU Component Failure

All data and parametric values reflected here are for illustrative purposes only.

Use Case 2
Reducing yield losses by eliminating and minimizing incorrect ball assembly in ball attach equipment

Background

The ball attach module is where solder paste is printed onto the lands of the substrates. Solder balls are placed into the ball attach lands, and the paste holds them in place. The entire package goes through a reflow oven, which melts the paste and balls to the substrate lands.

Solder balls are vacuumed onto tiny holes of the placement head. The head is inspected for excess or missing balls. Once the head is aligned to the substrates, the balls are placed in the solder paste on the substrates. After releasing the balls, the placement head is inspected for any remaining balls. Finally, the substrates are inspected by a camera vision system for any missing or shifted balls.

Problem Statement

Units with missing balls are considered faulty material and classified as a yield loss. There are numerous scenarios in which balls are missing from the units, including inadequate vacuum pressure.

Results and Benefits Yield

By visualizing and correlating sensor readings (Figure 4) with various machine and execution system data, Intel was able to reduce yield loss, optimize maintenance cost, and avoid sudden equipment downtime.[5] These real-time monitoring and predictive maintenance capabilities enabled the technicians to proactively fix the problem before it occurred.
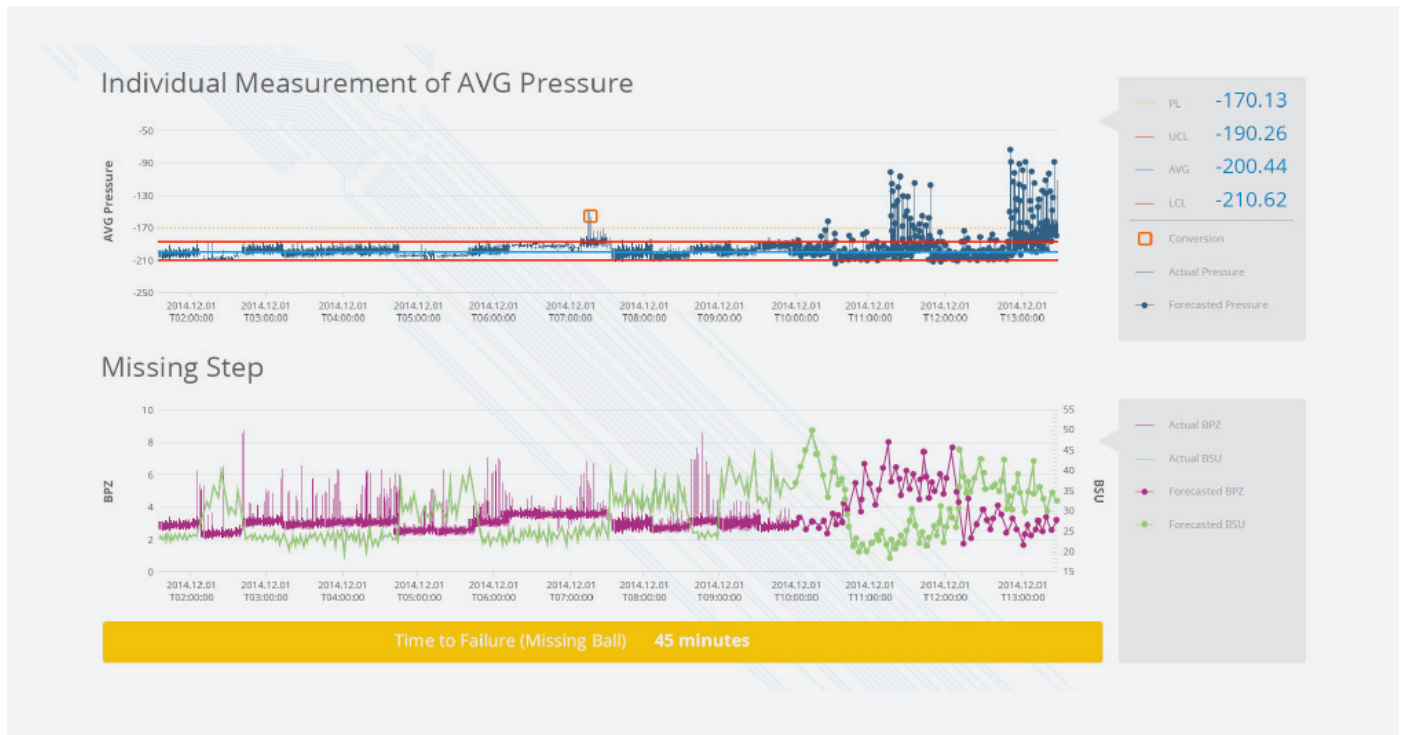
Use Case 3
Using image classification to identify good or defective units

Background

Machine vision equipment is used in a module that screens units and categorizes them into good and marginal units. The good units are sent forward to be processed while the marginal ones are inspected and determined by a manufacturing specialist to be either good or defective. This manual process takes a significant amount of time.

Problem Statement

The manual process to inspect and categorize marginal units is cumbersome and can take up to eight hours to successfully segregate a host of true reject units from marginal ones. This includes the time it takes for the units to reach the operator, then flow to a segregation module, and lastly be segregated into bins. In contrast, an inspection module using image analytics can identify reject units just moments after being scanned.

Results and Benefits Yield

The marginal images recorded at the machine vision equipment module are preprocessed. Each image, which is unstructured data, is first resized, cropped, and converted into grayscale, and then each pixel is converted into a binary format. The next stage of the process involves feature selection, where the unstructured image is defined by a set of distinct values. These values are then fed into various machine learning algorithms (Figure 5) to differentiate true and marginal rejects. The image analytics shorten the time to save true rejects from a pool of marginal units without sacrificing quality requirements. Image analytics identifies defects roughly ten times faster than the manual method.[5]

Figure 4. Analysis of Predicted Average Pressure Readings and Machine Motor Missteps Predicts When Missing Ball Occurrences Will Increase

All data and parametric values reflected here are for illustrative purposes only.

Technologies

This section describes the technology ingredients that are built into the end-to-end solution.

Industrial Data Center in the Manufacturing Shop Floor Control Room

Figure 6 shows the building blocks of an Industrial Data Center (IDC) that hosts the required data store and management, data analytics, security, and business intelligence software stack. The IDC (see sidebar) caters to the needs of the local factory operations and resides in the local manufacturing shop floor control room. It connects to the enterprise network and other IDCs deployed at other factories of the manufacturer. The key hardware and software components of the IDC are described here.

Figure 5. Machine Learning Algorithms Used to Improve Quality Defects Classifications on Each Unit of the Lot at a Much Faster Rate

All data and parametric values reflected here are for illustrative purposes only.

Figure 6. Industrial Data Center
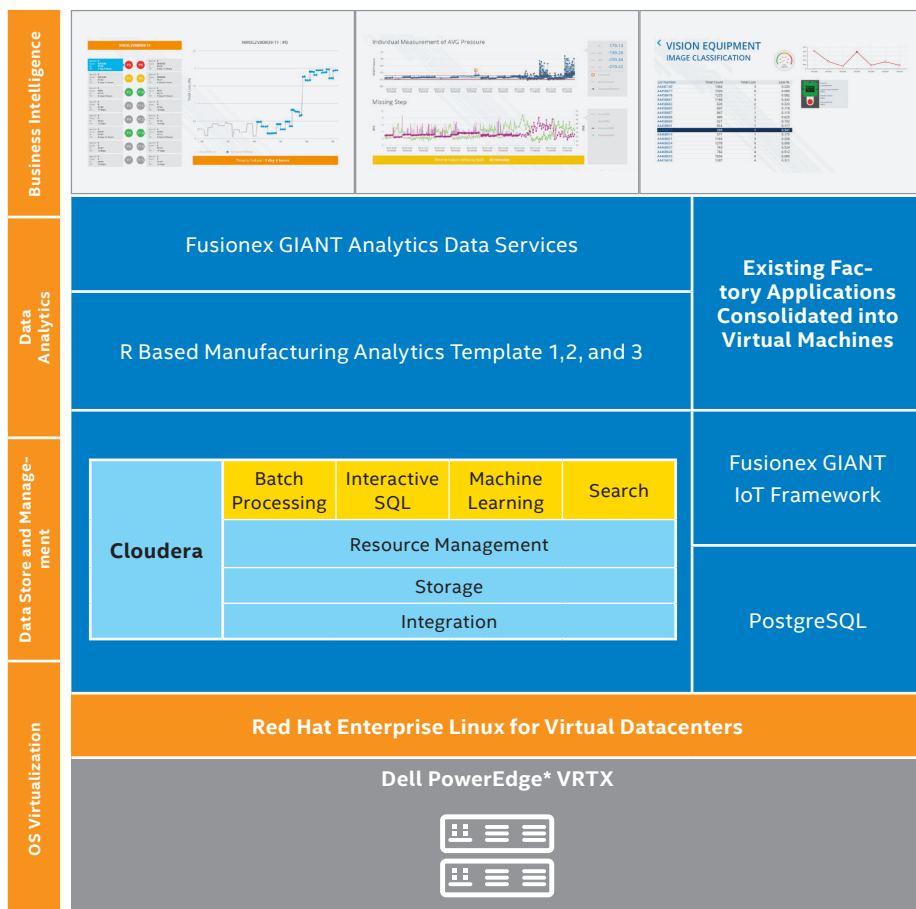
Hardware Components

The hardware platform is a Dell PowerEdge* VRTX, a compact system that hosts the big data analytics software in a private-cloud setting as an on-premise system of servers. The system consists of a Dell PowerEdge VRTX chassis with twenty-five 900GB hard disks and two Dell PowerEdge* M820 blade servers, each equipped with four Intel® Xeon® processors from the E5-4600 product family. The Intel Xeon processor E5-4600 product family provides a dense, cost-optimized, four-socket processor solution with up to eight cores, up to 20MB of last-level (L3) cache, and up to 1.5TB maximum memory capacity, along with communication pathways to move data more quickly.

Industrial Data Center (IDC) Overview

An Industrial Data Center (IDC) is an on-premise server platform implementing operational technology on the manufacturing shop floor. It is purpose-built and specially configured to be an "IoT cloud" for manufacturing data store, management, and analytics. IDC hardware must meet stringent requirements for high availability, redundancy, and harsh manufacturing environments (e.g., extended temperature, vibration). IDC software capabilities may include:

- Data center virtualization and orchestration
- Data store (e.g., Hadoop, RDBMS SQL)
- Secure data store and processing
- Manufacturing analytics
- Equipment edge IoT gateway management
- Business intelligence
- Data gateway between the shop floor and external data center (enterprise or cloud)

- API exposure and new services creation for internal company and B2B external communications
- Support for existing factory applications, such as process control and shop floor manufacturing execution systems

The IDC connects to a bounded, secure manufacturing network, enabling closed-loop control operations that are typically managed by manufacturing IT.
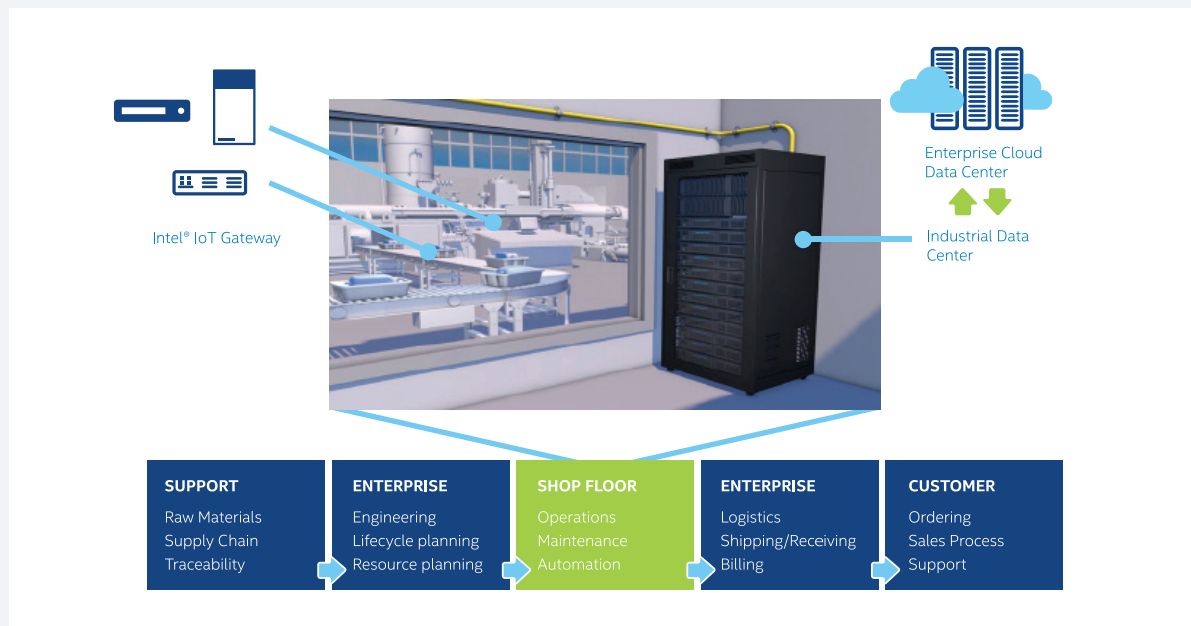

Intel® IoT Gateway

Enterprise Cloud Data Center

Industrial Data Center

SUPPORT	ENTERPRISE	SHOP FLOOR	ENTERPRISE	CUSTOMER
Raw Materials Supply Chain Traceability	Engineering Lifecycle planning Resource planning	Operations Maintenance Automation	Logistics Shipping/Receiving Billing	Ordering Sales Process Support

Figure 7. Industrial Data Center and Intel IoT Gateway

Software Components

Figure 7 shows two Dell PowerEdge* M820 servers hosting data analytics and application software, and Hadoop nodes, which run in multiple virtual machines (VMs). Red Hat Enterprise Linux for Virtual Datacenters operating system provides a complete, scalable virtualization software solution for the servers.

Analytics and Application Node

The node hosts multiple VMs that run the various data analytics and application software workloads. This includes:

- Revolution R Enterprise* from Revolution Analytics is an analytics software tool built upon the powerful open source R statistics language. There are numerous packages and libraries, which support different analytical models available in the R support community. Manufacturing analytics models are built by data scientists or statisticians, who will typically consult with manufacturing automation equipment owners and IT to determine the problems and analytics requirements, and how data models can be deployed. The Revolution R Enterprise software supports a variety of big data statistics, predictive modeling, and machine learning capabilities and provides users with cost-effective and fast big data analytics that are fully compatible with the R language. Offering high-performance, scalable, enterprise-capable analytics, Revolution R Enterprise supports a variety of analytical capabilities, including exploratory data analysis, model building, and model deployment.[6]
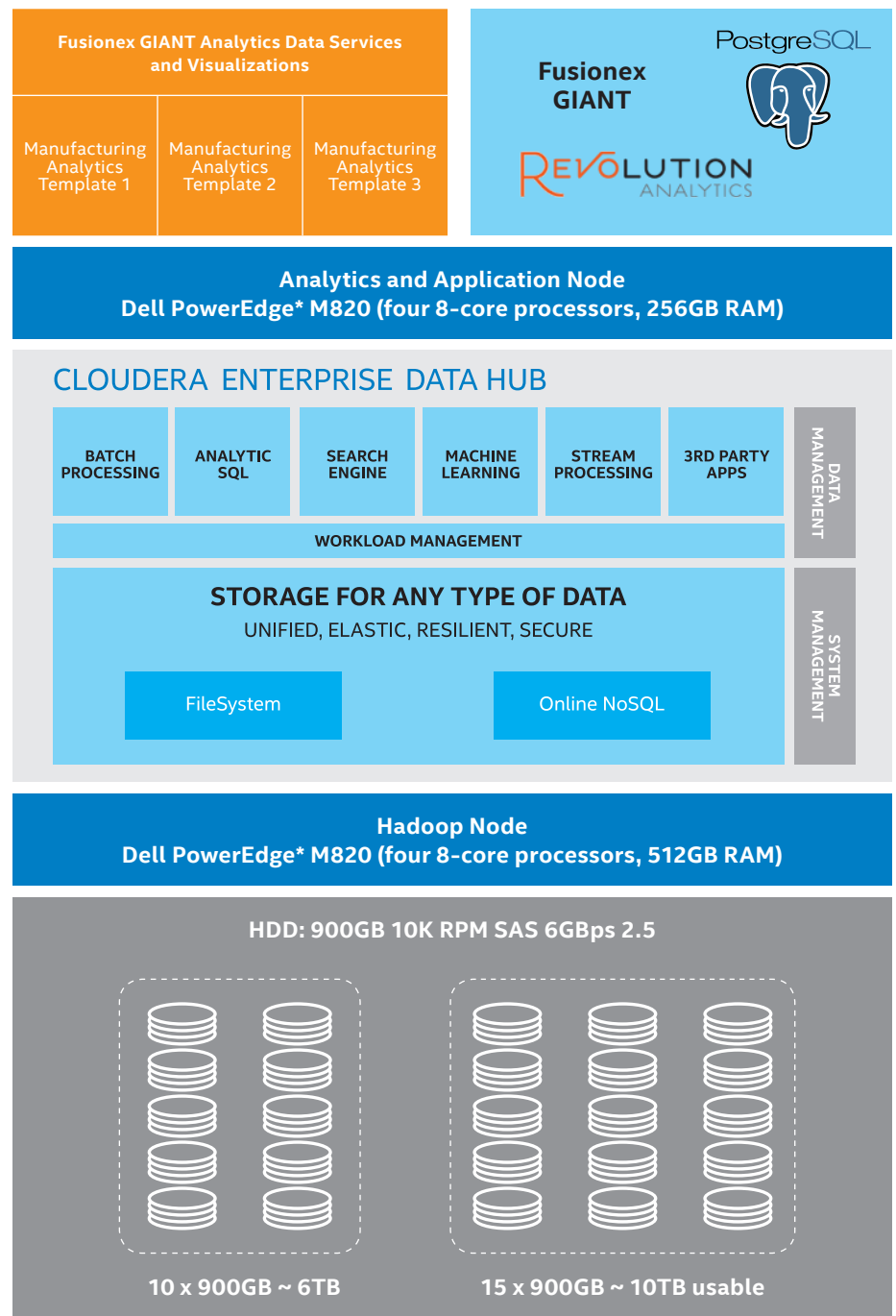
Fusionex GIANT Analytics Data Services and Visualizations

Manufacturing Analytics Template 1	Manufacturing Analytics Template 2	Manufacturing Analytics Template 3

Fusionex GIANT — PostgreSQL — REVOLUTION ANALYTICS

Analytics and Application Node
Dell PowerEdge* M820 (four 8-core processors, 256GB RAM)

CLOUDERA ENTERPRISE DATA HUB

| BATCH PROCESSING | ANALYTIC SQL | SEARCH ENGINE | MACHINE LEARNING | STREAM PROCESSING | 3RD PARTY APPS | DATA MANAGEMENT |

WORKLOAD MANAGEMENT

STORAGE FOR ANY TYPE OF DATA
UNIFIED, ELASTIC, RESILIENT, SECURE

FileSystem — Online NoSQL

SYSTEM MANAGEMENT

Hadoop Node
Dell PowerEdge* M820 (four 8-core processors, 512GB RAM)

HDD: 900GB 10K RPM SAS 6GBps 2.5

10 x 900GB ~ 6TB   15 x 900GB ~ 10TB usable

Figure 8. Software Runs in Virtual Machines on the Industrial Data Center

Revolution R Enterprise from Revolution Analytics

By linking with high-performance Intel® Math Kernel Library (Intel® MKL) multi-threaded math libraries, Revolution R Enterprise uses the power of multiple processors to accelerate common matrix calculations. In addition, Revolution R Enterprise parallel, distributed algorithms enable statistical analysis of big data to scale linearly with multi-core servers, high-performance computing clusters, big data appliances, and Hadoop.

- PostgreSQL is a powerful, open source object-relational database for most SQL:2008 data types, including INTEGER, NUMERIC, BOOLEAN, CHAR, VARCHAR, DATE, INTERVAL, and TIMESTAMP. It supports storage of binary large objects, including pictures, sounds, or video. In addition, there are native programming interfaces for C/C++, Java, .Net, Perl, Python, Ruby, Tcl, and ODBC.[7]

- Fusionex GIANT is a big data business intelligence software solution with streaming capabilities for real-time monitoring and analytics. The technology allows for encapsulation of R-based and other predictive models as "RESTful" web services. FusionexGIANT provides integration capabilities such as connecting and synchronizing multiple data sources, performing data warehousing ETL tasks, and ultimately, empowering end users to discover, drill down, and drill through in exploration of the underlying data via rich, stunning visualization. Rendered through HTML 5, Fusionex GIANT visuals are device-agnostic, supporting Android*, iOS, Blackberry*, etc. The solution supports communication to various production databases in the network, like PostgreSQL, Microsoft SQL, Oracle, and IBM DB2* with ODBC and JDBC connections.[8]

Hadoop node

Four virtual machines are provisioned to run a basic, four-node Cloudera Hadoop cluster, consisting of one head node and three worker nodes.

- Cloudera Enterprise Data Hub Edition offers a unified platform for big data by providing one place to store, process, and analyze all of the collected data, resulting in an extension of existing investment value while enabling fundamental new

ways to derive value from data. It includes CDH – Cloudera Distribution for Hadoop, which is 100 percent Apache licensed and open source. It also supports unified batch processing, interactive SQL, and interactive search and role-based access controls.[9]

CDH delivers the core elements of Hadoop, scalable storage and distributed computing, along with additional components, such as a user interface; necessary enterprise capabilities, such as security; and integration with a broad range of hardware and software solutions. The solution can help businesses transform and accelerate the way big data is used. When combined with data center architecture based on Intel® Xeon® processors, Cloudera and Intel provide a comprehensive approach for industrial and manufacturing businesses seeking to analyze data to make factories run more efficiently.

Internet of Things (IoT) Gateway Connecting to the Shop Floor Equipment

An IoT gateway is a new breed of network and connectivity infrastructure that connects to machines and acquires their data. When factory personnel want more data about how a machine or a sub-component within a machine is operating, they can add new sensors (e.g., vibration, temperature, pressure) to the machine and collect additional real-time data. Commercially available gateways have different compute, storage, and memory configurations to support a range of data acquisition requirements.

The major function of an IoT gateway is to transmit manufacturing data from factory equipment (source) to the industrial data center (IDC), thereby supporting interoperability between equipment and the IDC (Figure 9). This requires the IoT gateway to support various fieldbus protocols so it can

connect to legacy equipment and sensors. It then collects, normalizes, and secures the manufacturing data before sending it to the IDC.

The Intel big data analytics pilot program implemented the Mitsubishi Electric C Language Controller of MELSEC-Q Series*,[10] an Intel® Atom™ processor-based gateway used to aggregate and securely send data to the industrial data center.

The Mitsubishi Electric C language controllers of MELSEC-Q Series are embedded solutions equipped with numerous features characteristic of intelligent systems, including robust network connectivity and the high computational performance needed to process large amounts of data collected from sensors or via the network when supporting sophisticated system control and operations. At the core of this controller is a hardware platform based on Intel® architecture and the Wind River VxWorks* real-time operating system.

The C language controllers of MELSEC-Q Series satisfy the diverse requirements of factory automation, including excellent reliability, tolerance of harsh environments, and long-term availability. It is a robust and reliable product that requires little maintenance for IoT manufacturing applications.

In place of ladder logic used in conventional programmable logic controllers, the C language controllers of MELSEC-Q Series use the international standard C languages (C and C++) for greater programming flexibility. This allows users to take full advantage of their existing C language software and development.

CIMSNIPER* is a data acquisition and processing software package for Mitsubishi Electric C Language Controller of MELSEC-Q Series. It can collect process data (including machine builder standard like SECS messages) and manufacturing equipment errors without modifications of existing systems.
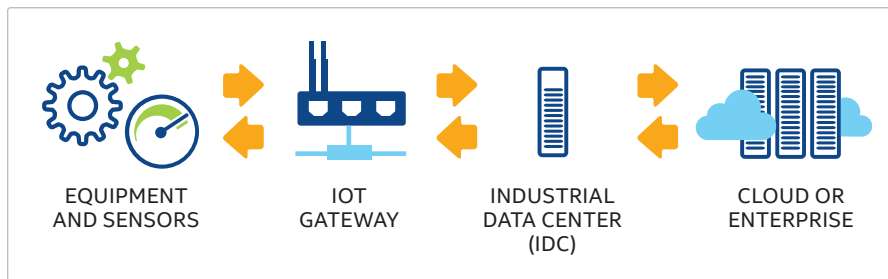
Figure 9. IoT Gateway Supports Interoperability between Factory Equipment, the IDC, and the Cloud or Enterprise Infrastructure

EQUIPMENT AND SENSORS — IOT GATEWAY — INDUSTRIAL DATA CENTER (IDC) — CLOUD OR ENTERPRISE

Data Flow

Figure 10 shows how data flows in the previously discussed use cases.

- The IoT gateway acquires machine and sensor data in real time, sending the data to the IDC after performing protocol abstraction and data filtering to clean and format the data per IDC requirements. For instance, machine data from the ball attach module is acquired over machine interface ports and from analog sensors mounted in the robotic arm of the ball attach module.

- All incoming factory data is stored in Hadoop. The following non-exhaustive list shows what can be achieved using readily available capabilities in Hadoop:

  – HTTP: The big data analytics server exposes an authenticated REST HTTP endpoint that supports operations into HDFS.

  – Apache Sqoop* provides the connectivity tool for moving data from non-Hadoop data stores – such as relational databases and data warehouses into Hadoop.

  – Apache Flume* can receive a stream of continuous data.

- The data specific to the previously discussed use cases is in the form of Comma-Separated Value (CSV) files or raw images. While the Hadoop ecosystem includes plenty of ways to ingest data, some factory machines have limited network transfer capabilities, which requires custom engineering to deliver data into HDFS:

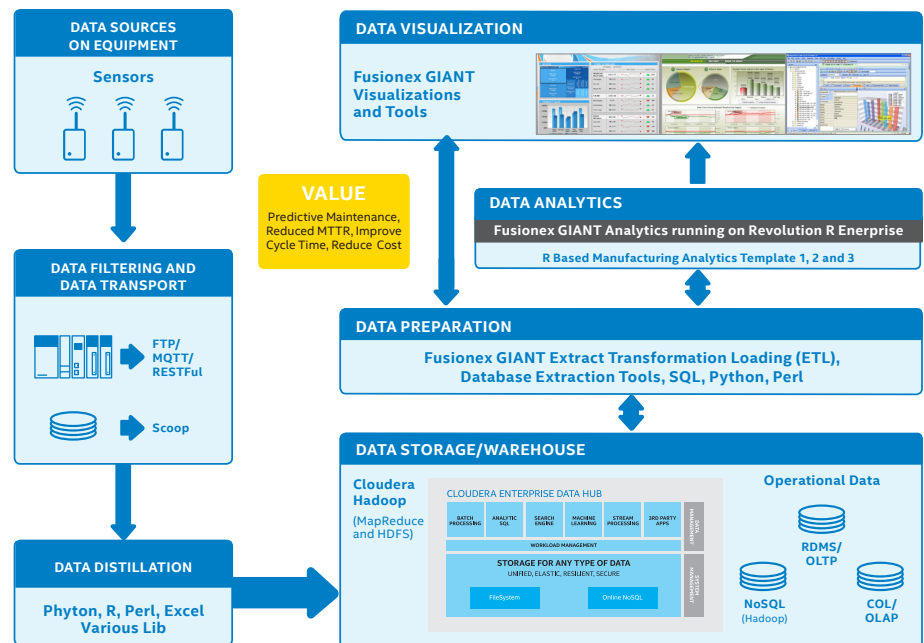 – FTP: The IoT gateway has an FTP client that periodically connects to the big data analytics server

Figure 10. High-Level Data Flow

and transfers the most recently acquired data directly into HDFS. Other streaming protocols, like MQTT and REST, can be used, depending on real-time streaming and analytic requirements.

  – Common Internet File System (CIFS)/Windows* share: The IoT gateway can copy files into a CIFS/ Windows share directory provided by the IDC.

- The CSV files are directly imported into HDFS, while the raw images are pre-processed using computer vision techniques to produce a textual data representation of the image.

- Depending on the operational and data schema requirements, data is stored in one of two databases: NoSQL (Hadoop) or RDMS/SQL.

- At this point, various data preparation tasks, such as data cleansing and data distillation, can be performed.

- Data can also be accessed and processed for tasks such as ad hoc reporting, workflow scheduling, and database integration.

- At the same time, the Cloudera Distribution for Hadoop performs various operations on the data that is typically too big for data management platforms.

- The distilled data will then be used for model building, training, and scoring, using Revolution R Enterprise analytics software.

- Depending on the analytics use cases, analyzed data is presented to users in easy-to-understand dashboards and alerts can be generated to enable actions on monitored and predictive conditions.

Implementation Considerations

Manufacturers planning an IoT and big data analytics project are likely to have different workload requirements than the ones presented in this blueprint; hence, their hardware and requirements for the industrial data center and IoT gateway may differ. This is especially critical when moving from pilot to production (full factory deployment). For instance, infrastructure requirements are driven by the nature of the data sets, networking capabilities of the existing manufacturing network, network configuration, processing speed (i.e., manufacturing analytics turnaround time), data retention strategies, and most importantly, compute performance of the infrastructure.

From this blueprint perspective, it is best to consult with the referenced hardware vendors to determine the most effective infrastructure solution deployment option. The software stack on the industrial data center and the gateway could scale to various types of problems, use cases, equipment, and manufacturing environments. Provisioning or deployment of the various software capabilities requires adjusting the number of nodes and virtual machines on the Industrial Data Center hardware, along with data connectivity requirements of the IoT gateway.

The implementation of this end-to-end solution required a project team made up of members from three key groups: the shop floor automation engineering group, the IT/OT integration team, and data scientists.

A more detailed system integration implementation guide for this Intel big data analytics pilot program featuring IoT and big data analytics solutions can be made available. The guide documents implementation learnings, possibilities, and challenges of Intel's own implementation using the solution building blocks discussed. The guide can be used as a quick start reference material for a project implementation team.

IoT Tenets

The Intel big data analytics pilot program was designed to provide security and interoperability from machines to the IDC in keeping with five key tenets defined by Intel:

- **World-class security as the foundation:**
  The solution protects the entire manufacturing environment with state-of-the-art security solutions.

- **Automated discovery and provisioning of edge devices** to ease deployment:
  The IoT Gateway interfaces seamlessly to machines and sensors, and connects securely to the IDC.

- **Data normalization** through protocol abstraction to improve interoperability:
  The IoT gateway can maintain various device and communication protocols.

- **Broad analytics infrastructure** from edge to cloud to realize customer value:
  The solution includes capabilities to support advanced analytics for factory operations and manufacturing processes.

- **Infrastructure** to monetize hardware, software, and data management from edge to cloud:
  This IoT big data solution enables substantial improvements in manufacturing performance, leading to a fast ROI.

Summary

An IoT and big data analytics solution can help manufacturers address various current and future manufacturing challenges while producing a clear ROI. The building blocks used in this end-to-end solution can be applied to existing equipment, assets, and data irrespective of equipment types, processes, issues, and deployment requirements of a manufacturer. The solution is therefore scalable to future business needs. If set up, managed, and used correctly, the solution allows manufacturers to differentiate their smart manufacturing capabilities with new business intelligence.

This blueprint can be applied and implemented by manufacturers who have not started to take advantage of the intelligence contained in manufacturing data. Manufacturers who have already been using data to improve their efficiency can take their data mining and analytics capability to the next level.

Intel integrated and validated this IoT and big data analytics solution with data extracted from its own manufacturing network and from IoT gateways connected to equipment and sensors. The Mitsubishi Electric C language controllers of MELSEC-Q Series were used as IoT gateways. The pilot program benefited from close collaboration between the factory engineers, the IT department, and industry experts from Cloudera, Dell, Fusionex, Mitsubishi Electric, and Revolution Analytics. The team started with existing machine performance and monitoring data, then proceeded to using big data analytics and modeling to ingest additional data used to predict potential excursions and failures. Being able to predict machine component failures enables advanced alerting and allows operators, engineers, and managers to repair and prevent the excursion, hence reaping tremendous savings from improving yield, reducing time to repair, and using fewer spares.

For Intel, this pilot is forecast to save millions of dollars annually, along with other business value that Intel is still realizing. Benefits include improving equipment component uptime, minimizing wrong classification of good units as bad (thereby increasing yield and productivity), enabling predictive maintenance, and reducing component failures. There are many other types of parametric, product, and equipment data—both structured and unstructured—within Intel manufacturing environment and machines, which could be mined and analyzed to extract new business values. Capitalizing on this opportunity will enable further efficiency and productivity enhancements to the factory and ultimately create a competitive advantage. These business benefits can be realized by other manufacturers adopting a similar blueprint implementation.

Resources

Intel® Internet of Things Solutions Alliance

Members of the Intel® Internet of Things Solutions Alliance provide the hardware, software, firmware, tools, and systems integration that developers need to take a leading role in IoT.

Intel® IoT Gateway Development Kits

Intel® IoT Gateway development kits enable solution providers to quickly develop, prototype, and deploy intelligent gateways. Available for purchase from several vendors, the kits also maintain interoperability between new intelligent infrastructure and legacy systems, including sensors and data center servers.

For more information about Intel® solutions for IoT, **visit www.intel.com/iot.**

[1] Source: Gartner analyst Doug Laney introduced the 3Vs concept in a 2001 MetaGroup research publication, 3D data management: Controlling data volume, variety and velocity. blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.

[2] Source: "2012-13 Intel IT Annual Report," www.intel.com/content/www/us/en/it-management/intel-it-best-practices/intel-it-annual-performance-report-2012-13.html.

[3] Source: "Intel Cuts Manufacturing Costs With Big Data," InformationWeek, March 18, 2013, www.informationweek.com/software/information-management/intel-cuts-manufacturing-costs-with-big-data/d/d-id/1109111.

[4] Source: http://www.techopedia.com/definition/2148/automatic-test-equipment-ate.

[5] Results might vary depending on package size, process, and equipment used in the manufacturing process. The TCO or other cost reduction scenarios described in this document are intended to provide a better understanding of how the purchase of a given Intel product, combined with a number of situation-specific variables, might affect future cost and savings. Nothing in this document should be interpreted as either a promise of or contract for a given level of costs.

[6] Source: www.revolutionanalytics.com/revolution-r-enterprise.

[7] Source: www.postgresql.org/about.

[8] Source: http://www.fusionex-international.com/Products/Analytics-Big-Data.

[9] Source: www.cloudera.com/content/cloudera/en/products-and-services/cloudera-enterprise.html.

[10] Source: http://www.mitsubishielectric.com/fa/products/cnt/plcq/items/cpu.