WHITE PAPER
MAY 2018

IT@INTEL

Driving Improvement in Manufacturing through Advanced Data Analytics

Embracing a world of devices that connect and interact can fundamentally change how we use information technology in business.

Dr. Steve Chadwick
Senior Principal Engineer, Intel IT

Executive Overview

As technology evolves, Intel's worldwide factories have experienced growing complexity in products and processes. We use data analysis and sensors to improve product quality, reduce capital costs, and speed time to market. Often a single tool is equipped with multiple sensors, each collecting specific, unique data. We also use Internet of Things (IoT) sensors to collect additional data not associated with tools. These sensors collect tens and sometimes hundreds of data points per second, and each factory contains thousands of sensors, resulting in massive amounts of data. Understanding the data life cycle (store, mine, integrate, notify, report) is necessary for solving complex business problems while protecting and managing the IT environment.

While data from any source can provide greater business insights, utilizing it to its fullest potential can present challenges:

- **Analyzing large amounts of data.** Traditional analysis methods become inadequate as data volumes explode, and data must be converted from multiple sources.
- **Identifying useful data.** Once integrated, data must be filtered for the most meaningful information, and correlations between seemingly unrelated data must be made in order to derive the highest value from that data.
- **Delivering information.** Useful information must be delivered to users in formats that are easy to act upon.

Intel has decades of experience investing in and using advanced data analytics, and we are sharing our best practices to help enable other enterprises to solve complex problems. Along with optimizing the data life cycle, we have integrated all manufacturing data into a single database to allow faster, agile reporting and data investigation.

Our integrated manufacturing data and advanced data analytics enable us to perform calculations never before possible. We also recently upgraded our servers to a newer generation of Intel® Xeon® processor, which has reduced the time it takes to process extremely large data sets (four billion records) from 28 hours to just 15 hours. Together, advanced analytics and hardware upgrades help improve yield and time to market.

Contents

Acronym

IoT Internet of Things

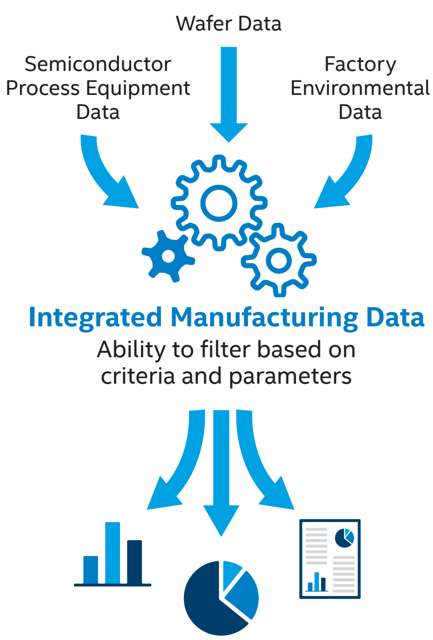


Figure 1. Delivering useful information to people who need it, in a format that is easy to act on, requires using integrated manufacturing data from many sources, combined with advanced data analytics and criteria- and parameters-based reporting.

The Ever-Increasing Complexity of Data Analytics

Intel operates factories around the world that have experienced growing complexity in products and manufacturing processes. This complexity generates massive amounts of data. For example, each silicon wafer is associated with roughly one gigabyte of data, and there are thousands of wafers sorted per day. This creates a large volume of data that must be analyzed with the resultant learnings applied to the manufacturing process. Intel® factories also use a variety of data-collection sources including sensors to collect data on individual tools in factories across the globe. This data helps us reduce capital costs, increase product quality, and improve time to market.

Many organizations, including Intel, are beginning to use the Internet of Things (IoT) to collect data, in addition to tool data, at its source. This brings the ability to leverage more data and use technologies such as 4G and Wi-Fi* for faster analysis. The IoT environment consists of diverse sensors, devices, and technologies connecting, interacting, and negotiating with one another to help meet business challenges. We use sensors in Intel's manufacturing process equipment, and often one tool is equipped with multiple sensors, each collecting specific, unique data. For example, a tool that deposits metal (sputtering) has sensors that monitor pressure, temperature, and gas composition. These sensors provide feedback that enables technicians to monitor product quality and act on that information in real time.

Massive amounts of data can lead to greater business insights and more actionable information, but utilizing the data to its fullest potential requires a different approach to data management and analysis than in the past.

Thousands of sensors collecting tens to hundreds of data points per second generate enormous amounts of data that must be stored and analyzed before it is useful. Traditional analysis methods become inadequate as the volume of data explodes. In addition, integrating multiple data sources requires conversion and contextualization before data can be meaningfully analyzed or acted upon. Once data is integrated, the most meaningful information must be identified, and correlations between seemingly unrelated data must be made to derive the highest value from that data. For example, factory personnel may need to explore data associated with all factories in the system or specific types of tools in one or more factories. Even then, this information must meet specific criteria and parameters and be delivered to decision makers and other users in a format that is easy to act on. In addition, HTML5's touch-enabled features and rapid evolution presents challenges for developers to deliver touch-enabled information, which is what employees have come to expect.

For example, as shown in Figure 1, Intel IT collects, integrates, and analyzes process and equipment data, wafer data, and factory environmental data.

This data can be used to create criteria- and parameters-based reporting to deliver useful information previously hidden in huge volumes of data to the people who need it. Our approach to managing Intel's integrated manufacturing data is enabling us to remain competitive, reduce costs, and increase throughput despite the rise in data complexity.

Managing the Data Life Cycle for Improved Insights

To remain competitive in a complex data environment, Intel continues to shift toward real-time decision making whenever possible. With the availability of Wi-Fi and 4G networking, analyzing data at the edge and streaming data from factory sensors is possible. In addition to real-time analytics, large amounts of data are captured and stored for after-the-fact analysis, which can provide historic trends and new opportunities for revealing previously undiscovered information.

Step 1: Understanding the Data Life Cycle

Understanding the data life cycle (shown in Figure 2) and managing it more efficiently allows Intel IT to protect information, solve complex problems, and manage the IT environment effectively. Each stage of the data life cycle includes activities and requirements to optimize the value of the data.

The data life cycle consists of the following stages:

- **Data storage.** We use storage methods that can handle hundreds of data points per second, terabytes of summary data, rapid read/write capabilities, and archiving.
- **Data mining.** Mining the data in accordance with Intel® Privacy Principles helps discover correlations in data from a single source and reveals new insights. We choose Intel® Xeon® processors, technologies, and networks that can meet high-performance computing needs. (See "Intel® Technology for Advanced Data Analytics" for more information.)
- **Data integration.** To solve complex problems and create meaningful insights, we perform data integration across disparate sources. This integration provides multiple, distinct viewpoints that, when combined, reveal a holistic view.
- **Notification.** Effective notification must be prioritized, be delivered to appropriate devices, highlight what is most important, and prompt a response.
- **Reporting.** Effective reporting must deliver actionable information tailored to the device and user receiving it.

Intel® Technology for Advanced Data Analytics

Intel® technology enables advanced data analytics to bring real-world insights to business. The benefits include:

- **Performance and reliability.** Intel® Xeon® processors deliver high performance, agility, and reliability across a full range of workloads. Built-in security features are equally applicable to micro-servers or high-end data centers.
- **Scalability.** Intel® Omni-Path Architecture scales to thousands of nodes, increasing message throughput and overall fabric efficiency with minimal latency.
- **Memory management.** Intel® Xeon® Scalable processors provide innovative memory management necessary for big data analytics.

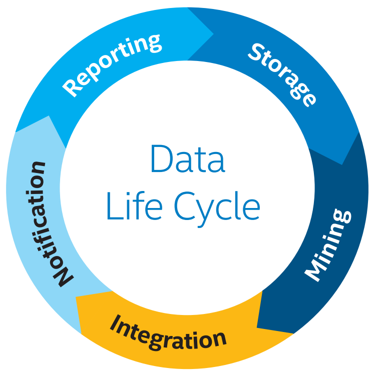


Figure 2. Each stage of the data life cycle includes specific requirements, such as correlating information in the mining stage and delivering actionable information in the notification stage.

Step 2: Analyzing, Identifying, and Delivering the Most Useful Data

Intel maintains several powerful state-of-the-art data centers built with Intel® Architecture that perform ad-hoc, real-time, and after-the-fact data analysis. To balance computing requirements, jobs are distributed across these data centers, yet remain close to the data source for faster processing. From this analysis, Intel IT generates reports based on the desired outcome—knowing what information is important, what devices it will be delivered to, and how it will be acted upon. To achieve this goal, we filter data based on importance, traverse data from multiple sources, and use HTML5 to create easily useable reports.

The results, as described in the sidebar, "Increasing Uptime, Accelerating Throughput, and Improving Quality," have been significant.

Filtering data for importance

Knowing what information is most important is a critical factor in data analytics. As individuals, we are often overwhelmed with data. From too much incoming email to social media and news sources, filtering information to include only what is important to us is an ongoing challenge. Incoming data from factory tools, sensors, and other sources is no different. It is impractical to process it all—there must be balance along with prioritization.

Traversing multiple sources to solve problems

Solving problems with data requires traversing the hierarchy of the data across multiple sources and making correlations between them. For example, locating a car in the United States based solely on its license plate involves analyzing data from multiple sources, beginning with a search of the Department of Motor Vehicles' records in all fifty states. Once a match is found, the vehicle is correlated to its owner, and the owner's address is retrieved. The address must then be mapped using a different system to locate the car. Similarly, we traverse data from multiple sensors, sources, and factories to find information needed to solve Intel's business challenges.

Delivering useful information

With HTML5, content can be delivered to any platform—smartphone, tablet, laptop, or desktop. It can also include interactive "point and click" functionality for ease of use. Smartphone functionality, such as touch capabilities, has become the standard baseline of user expectations, and HTML5 allows us to create touch-enabled reports that users find easy to use. Along with its unique features and rapid evolution, HTML5 often presents challenges that we continue to work through to provide the best user experience possible.

Increasing Uptime, Accelerating Throughput, and Improving Quality

To monitor yield, cycle time, uptime, process health, and maintenance schedules, Intel's process engineers collect data and transform it into actionable information with Intel's advanced data analytics system. Engineers can use advanced data analytics to correlate parts-replacement frequency to end-of-line yield. For example, a sensor that is monitoring the pressure inside a manufacturing tool may detect a change in pumping efficiency and alert the technician. This type of notification can help technicians proactively address the issue and minimize unscheduled downtime.

Technicians conduct quality checks on tools when running production materials. Quality checks increase the time it takes a wafer to get through the factory, so an engineer wants to check only as often as absolutely necessary. Intel uses advanced data analytics and modeling to improve quality across manufacturing by dynamically adjusting the monitor frequency and optimizing throughput.

Real-time sensor data from equipment in each factory represents huge "snapshots" of tens to hundreds of data points per second per sensor. Manufacturing IT data analysis tools perform real-time, as well as end-of-line, correlation. Instead of looking at thousands of graphs, now engineers can prioritize issues in their area to increase efficiency by focusing on the specific issues. Processing more than 5 billion points of sensory data per day per factory results in significant, measurable improvement in equipment availability and yield.

Step 3: Start Small and Grow

Making use of these large and ever-growing data sets requires complex time series analysis; processing this data in real time is not a trivial task. But realizing the potential of data does not require a vast network of sensors in the beginning; it starts with understanding how to manage and optimize the sensor data available. Intel IT started small and experienced a learning curve in mining and integrating data across multiple sources. We then were able to identify what was most important in the vast sea of information and we soon discovered actionable insights. Over time, our advanced data analytics process has grown to support increasing amounts data in addition to more traditional sources.

Solution Architecture

We have developed an integrated manufacturing data system that enables us to combine data from various manufacturing sources (Figure 3). These include quality data, data from the Manufacturing Execution System, and control systems data in individual factories, as well as IoT sensor data. We use extract, transform, and load (ETL) processes to compile that data and cross-reference different data domains. We physically separate the servers that perform IoT data pre-processing from the servers that process the factory data for added efficiency. The entire system is based on powerful Intel Xeon processors that provide the memory and compute capacity to process databases that can contain billions of rows.

The unique combination of distributed computing (at the fabrication site level) and centralized computing (at the integrated data level) enables us to minimize computing impact at the facilities but also crunch exponentially growing amounts of data. It also enables us to standardize calculations, which is an important part of Intel's Copy Exactly (CE!) manufacturing processes.

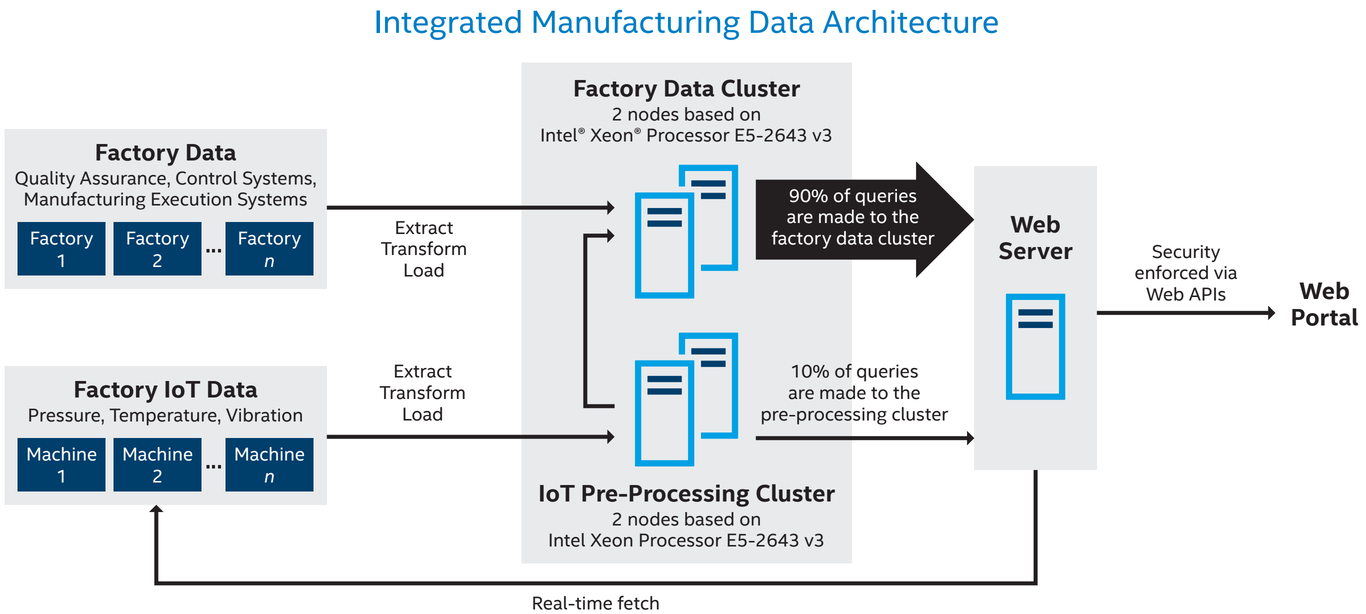


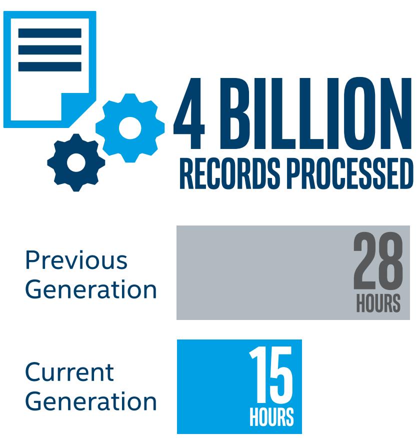
Figure 3. Our integrated manufacturing data system, running on high-performance Intel® Xeon® processors, pulls data from many sources and makes that data available through a security-enabled web portal. Users can also drill down into the IoT data as necessary.

We use governance and access controls to protect data to maintain a high level of security and consistency between factories, throughout the life cycle of the data and all subsequent containers. Our approach provides flexibility by allowing users to directly query site-level data while simultaneously protecting data integrity. For the integrated multi-site data, we provide users with a web portal and modern user interface. They can perform a wide variety of data visualization tasks using the portal, and we continue to protect the data by allowing access to the integrated data only through the user interface.

Results

We are realizing benefits from our integrated manufacturing data system on many fronts. These results can be divided into two primary categories: benefits resulting from upgrading technology and benefits resulting from centralized, integrated data. The system is now used across all of Intel's front-end factories by thousands of engineers daily. Usage of the system continues to grow year-over-year.

We use Agile methodology to promote continuous improvement at a high velocity. The adoption of Agile practices enables us to deliver improvements to production systems within hours of a user request. We integrate testing into the build process to maintain a high level of quality.



Technology Upgrade Speeds Insights

Near the end of 2017 we upgraded our IoT pre-processing cluster servers from the Intel® Xeon® processor X5650 (2 GHz) to the Intel® Xeon® processor E5-2643 v3. The additional memory and compute power has enabled significant improvements in our ability to process vast amounts of data in less time. For example, we can now process one 4-billion record data set in 15 hours instead of 28 hours, nearly cutting our processing time in half. And put simply, faster results can save millions of dollars and increase time to market.

Integrated Data Provides New Capabilities

Although we are realizing tremendous value from our upgraded servers, a purpose-built repository of integrated manufacturing data provides additional value by enabling us to accomplish things we simply could not do before. We are using business intelligence to integrate data from different domains, and the more data we can correlate, the more insights we can uncover. The centralization and reorganization of data allows us to perform comparisons and calculations that were previously impossible. It also creates an eminently scalable architecture that can keep up with data growth.

A Connected Future

Intel IT sees transformational possibilities for sensors in the evolution of enterprise IT systems. IoT applications are becoming more prevalent in our corporate facilities and factories in a variety of use cases (see the sidebar, "The Increasing Influence of Sensor Networks"). The IoT environment is changing the way we operate today and into the future, including the following applications:

- **Smart buildings.** IoT sensors are improving energy efficiency, security, asset tracking and management, and comfort for occupants in smart buildings. For example, occupancy sensors automatically adjust lighting and temperature depending upon whether a location, such as a conference room, is in use or vacant.
- **Data centers.** IoT sensors are improving energy efficiency and environmental control in the data center, as well as helping IT organizations manage physical equipment and its utilization. Sensors that regulate temperature automatically adjust heating and cooling equipment to remain within acceptable ranges.
- **Factories.** Predictive maintenance, intra-factory data transport, remote monitoring, and maintenance are increasingly managed by IoT sensors with real-time adjustment capabilities to meet specifications. Sensors can help engineers predict impending equipment failure and proactively schedule maintenance to help avoid unexpected downtime.
- **Supply chains.** Inbound tracking, warehousing, supplier security management, and capacity utilization can be optimized through the use of the IoT. Sensors can help maintain real-time inventory counts and proactively schedule orders before inventory gets too low.
- **Worker productivity.** IoT applications can improve information assistance, mobile productivity, and group collaboration for mobile workers. Employees can move between locations and automatically connect to networks and equipment, such as conference room controls, without additional setup.

Embracing a world of smarter devices that connect and interact with one another and the data it provides to IT organizations paves the way for IT-based design flexibility and agility. It can fundamentally change how we use information technology in business.

Putting the Compute Where it Makes Most Sense

Intel IT has built a robust private enterprise cloud. We also have a vast manufacturing Internet of Things (IoT)—hundreds of thousands of sensors deployed throughout Intel's factories worldwide, collecting a wide variety of data such as location, temperature, humidity, and vibration levels. If we sent every data point collected by every sensor through the network to our cloud, we would quickly overwhelm the network as well as drive up data transmission costs. On the other hand, if we kept all the data at the edge, it would be impossible to integrate different sources of data to reveal important insights.

The combination of a centralized computing capability in the cloud and smart sensors and gateways at the edge gives us the best of both worlds. Sensors located on manufacturing equipment can perform edge analytics and respond in real time, such as shutting off the machine or alerting a technician. These smart sensors can also filter data and send only the most important data points to the cloud. Then our cloud-based applications can perform additional analytics to correlate seemingly unrelated data, identify patterns, and perform data visualization.

The edge-cloud relationship is dynamic—insights found at the cloud level can be translated into algorithms, which can then be deployed to the edge. This iterative process enables us to gradually improve the decisions made by smart devices.

We have found that edge computing is best for situations where data is needed in near real time, whereas the cloud is best for large data sets that are less time-sensitive, such as inventory management and manufacturing execution systems.

By using both edge and cloud computing, we can reduce the number of device-to-cloud data round trips, thereby controlling connectivity and computing costs without compromising responsiveness.

Conclusion

We look forward to a future for Intel factories where most sensors are wirelessly connected to the network, edge analytics proactively locate impending equipment failures, and sensor systems keep track of equipment inventory and automatically place orders to suppliers.

Data integration and analysis is helping to lower capital costs, increase yield, and improve cycle time. We are taking advantage of powerful Intel Xeon processor-based technology to analyze oceans of data and to convert that data to actionable business insights. Integrated manufacturing data and an optimized data life cycle that includes the IoT are important factors in converting information into revelations that can help us solve Intel's business challenges.

For more information on Intel IT best practices, visit **intel.com/IT**.

IT@Intel

We connect IT professionals with their IT peers inside Intel. Our IT department solves some of today's most demanding and complex technology issues, and we want to share these lessons directly with our fellow IT professionals in an open peer-to-peer forum.

Our goal is simple: improve efficiency throughout the organization and enhance the business value of IT investments.

Follow us and join the conversation:

- Twitter
- #IntelIT
- LinkedIn
- IT Center Community

Visit us today at **intel.com/IT** or contact your local Intel representative if you would like to learn more.

Related Content

Visit **intel.com/IT** to find content on related topics:

- Horizontal IoT Platform Paves the Way to Enterprise IoT Success paper
- Joining IoT with Advanced Data Analytics to Improve Manufacturing Results paper
- Transforming Intel's Supply Chain with Real-Time Analytics paper
- Using Big Data in Manufacturing at Intel's Smart Factories paper
- Extending Enterprise Business Intelligence and Big Data to the Cloud paper
- Using Apache Hadoop* for Context-Aware Recommender Systems paper
- Reducing Client Incidents through Big Data Predictive Analytics paper