Executive Summary
Manufacturers have been collecting and storing data for years, but now big data technologies enable more constructive use of this information, such as to increase throughput, boost yields, improve efficiency, and reduce downtime. Big data is characterized by huge data sets and varied data types (e.g., images, text, and machine log files), which the production line is producing at a much faster rate than ever before. When this data is analyzed using new tools available on the market, manufacturers can gain valuable insights derived from finding patterns, extracting meaning, and ultimately making decisions that lead to greater efficiency.

However, many machine tools operate in relative silos, so it is a major challenge to collect, analyze, and act on data generated across the factory floor. This is why Dell* and Intel have assembled various Internet of Things (IoT) and big data technologies that provide the connectivity, security, interoperability, and analytics capabilities that enable higher performance manufacturing.
This paper outlines an Internet of Things (IoT) solution very similar to one implemented in an Intel manufacturing facility to show how data analytics applied to factory equipment and sensors can bring operational efficiencies and cost savings to manufacturing processes. The Intel pilot is forecasted to save millions of dollars annually along with additional return on investment business value.1 Likewise, this paper presents a solution that can help manufacturers benefit from the revolutionary advances in computer processing power, database technology, and tools for big data.

### Emergence of Operational Technology

According to Gartner*,2 operational technology (OT) is hardware and software that detects or causes a change through the direct monitoring and/or control of physical devices, processes, and events in the enterprise. Information technology (IT) is the common term for the entire spectrum of technologies for information processing, including software, hardware, communications technologies, and related services. In general, IT will only include embedded technologies that generate data for enterprise use.

A clear distinction of IT and OT can be best described by referencing the ISA 95 standard, as summarized in Figure 1. The ISA 95 is the international standard Level 4 and 5 are typical IT layers, while Level 1, 2, and 3 are operational technology devices and processes of the organization.3

### Key Business Objectives

Increasing production line performance and improving the bottom line using data that is generated throughout the production cycle to help make decisions, solve business problems, and identify opportunities.

### Business Challenge

Data throughout the production cycle contains clues that ultimately explain variation in yield, quality, and output.

---

**Figure 1. ISA with Added Business Management**
However, getting this information into the right hands and at the right time is made more difficult by large numbers of factory automation networks that do not talk to each other.

Moreover, some manufacturing environments generate massive data files (gigabytes in a few days per tool type), limiting the ability to store, analyze, and extract useful information from them using conventional methods. Without the use of big data technologies, it is extremely hard to even visualize the information in large data sets from various sources.

**Solution Benefits**

Through the use of IoT technologies and big data analytics in its factories, Intel demonstrated benefits in many areas, including:

- **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 are used to identify good units from defective units roughly ten times faster than the 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**

The following IoT big data analytics solution, a collaborative effort between Dell, Intel and other industry ecosystem partners, can be applied to factory equipment and sensors to increase operational efficiency and manufacturing performance.

**High-Level Architecture**

Figure 2 shows a high-level IoT manufacturing architecture for small to large data sets spanning one or many factories. It supports data acquisition, aggregation, and analytics workloads for various types of data from the

---

**Find the Value in the Internet of Things**

Dell® believes that the Internet of Things represents a remarkable opportunity to find new efficiencies and innovations by instrumenting your physical world. To overcome the concerns with emerging technologies, Dell recommends that organizations Start Small, Build Fast and Connect what Matters.

**Start Small**

- Everything is more feasible and manageable when you start small. By keeping to a ‘walled garden,’ you can ensure security, make things interoperable, and establish a new flow of data.

**Build Fast**

- New data are like potato chips, you inevitably want more after you’ve had a first taste. We’ve seen high demand for both increased amounts of data and new data streams once people start imagining the possibilities. Dell recommends building in the flexibility and products to scale as the organization demands.

**Connect what Matters**

- ROI matters, even on small projects. Show immediate value by focusing even early activities on improving the core metrics that matter to your organization. The ability to move the needle on key KPIs should support the business case for follow-on investments of resources.
manufacturing shop floor and the manufacturing network, which opens up data mining and visualizing possibilities to create new business intelligence.

- **Data Acquisition**: Tools and sensor networks send factory data to industrial-grade IoT gateways that filter and secure the information before sending it to the data store platform.

- **Data Collection and Aggregation**: The data store platform based on the Hadoop* core collects data from throughout the factory, including structured data from existing databases and unstructured data from tool sensors, log files, and SMS text messages. It then cleans, extracts, transforms, and consolidates the data.

- **Data Analytics**: The data is analyzed by analytics software and high-level factory applications running on the data store platform.

- **Data Visualization**: The results of the analysis are presented to users via intuitive visualization capabilities in the business intelligence layer of the network, which may reside outside of the factory, thereby requiring firewalls and other measures to thwart hackers.

The **Data Collection and Aggregation** and **Data Analytics** functions can run in the cloud, the enterprise data center, or in a new category of computing infrastructure, called an industrial data center (IDC). The IDC is a platform situated on the factory floor – between factory equipment and the enterprise data center – designed for analyzing manufacturing data in motion. It supports a centralized manufacturing data store, manufacturing analytics, and workload consolidation. For manufacturers, the IDC allows multiple decision points throughout the business to deliver improved operational insight and real-time reporting that helps them become more efficient, productive, and competitive.

Figure 3. Solder Balls Attached to the Backside of a Processor

**Usage Model Example**

The following details some of Intel's groundbreaking work and what they discovered through the integration of big data analytics and technologies for the IoT in manufacturing.

**Use Case**: Minimizing yield losses by reducing incorrect ball assembly in ball-attach equipment

**Background**: Intel's ball-attach manufacturing module places solder balls on the underside of packaged chips (Figure 3) so they can be mounted on printed circuit boards. The solder balls are placed into the ball attach lands of a packaging substrate using solder paste to hold them in place. The entire package then goes through a reflow oven that melts the paste and balls to the substrate lands.

A key piece of equipment in the manufacturing process is the placement head, which holds the solder balls using vacuum pressure before they are attached to the substrate. After the head is inspected for excessive or missing balls, it is aligned to the substrate and the balls are placed in the solder paste. After releasing the balls, the placement head is inspected for any remaining balls, and a vision system (camera) checks the substrate for any missing or shifted balls.

**Problem Statement**: Units with missing balls are faulty material and contribute to yield loss. One cause is a placement head with inadequate vacuum pressure.
**Big Data Solution:** By visualizing and correlating sensor readings with data from various machines and the execution system, Intel was able to reduce yield losses, optimize maintenance cost, and avoid sudden equipment downtime. As a result, technicians can proactively address the problem using this predictive maintenance capability.

Figure 4 illustrates the five major steps of the IoT big data analysis process:

1. Data sources (e.g., ball attach module and data logs) – continuously send information to a gateway.
2. The gateway filters the data and securely sends it to the data store platform.
3. The data store platform ingests time series data and writes to the database.
4. Analytics software supports interactive, iterative, and graphical data experimentation.
5. Visualization software supports custom business intelligence dashboards.

**Results:** This solution, along with similar improvements to associated processes, allows approximately 25 percent of the yield loss to be recovered, thus saving millions of dollars annually. For more information about this usage model, please read the white paper, “Optimizing Manufacturing with the Internet of Things.”

**Technology**

This section describes an example of technology ingredients available to deploy a manufacturing solution like the one previously described.

**Gateways for Data Acquisition**

Intel® Gateway Solutions for the Internet of Things (Intel® Gateway Solutions for IoT) offer companies a key building block to enable connectivity in both legacy and new building systems. They integrate technologies and protocols for networking, embedded control, enterprise-grade security, and easy manageability on which application-specific software can run.

**Computing and Storage Platform**

The Dell In-Memory Appliance for Cloudera Enterprise* is designed to enable users to ingest, store, analyze, and build insights from big data. To deliver fast analytics and stream processing, the appliance is bundled with Cloudera Enterprise*, which includes Apache* Spark. Cloudera Enterprise allows manufacturers to...
implement powerful end-to-end analytic workflows, comprising batch data processing, interactive query, navigated search, deep data mining, and stream processing, all from a single common platform.

With a highly-available, common platform, there is no need to maintain separate systems that drive up complexity and cost by having separate data, metadata, security, and management. The solution delivers performance-packed computing with the latest Intel® Xeon® processor E5-2600 v2 product family and up to 24 dual in-line memory modules (DIMMs). Built with 22-nanometer process technology and up to 12 cores per processor, the Intel® Xeon® processor enables super-fast processing for compute-intensive tasks.

A particularly relevant aspect of the Intel® Xeon® processor E5 family architecture in the context of Hadoop, Intel® Data Direct I/O Technology (Intel® DDIO) makes a substantial contribution to the platform's overall I/O performance. Intel DDIO intelligently directs I/O packets to the processor cache, skipping main system memory. This action can dramatically reduce latency and improve overall system bandwidth and power utilization by eliminating unneeded trips to memory.5 Read more about platform considerations when running Hadoop.

Table 1. Dell® In-Memory Appliance for Cloudera Enterprise® Configurations

Dell In-Memory Appliance for Cloudera Enterprise, built on the Dell PowerEdge® two-socket Intel-based servers and 10G network connectivity, scales economically to support 8, 16, and 24 node clusters, as shown in Table 1. Any of these configurations can be used to support the big data software stack represented in Figure 5 and described in the following.

Red Hat® Enterprise Linux® operating system for servers is designed to support scalable and fully virtualized data centers. It runs on high-performance, multi-core systems that support the most demanding workloads. The operating system also manages the underlying system complexity, allowing users to get high performance with minimal management overhead.

Data Store and Management

Cloudera Enterprise offers a unified platform for big data by providing one place to store, process, and analyze all data, thus enabling fundamental new ways to derive value insights. The software is 100 percent Apache* licensed open source and is unique

Red Hat® Enterprise Linux®

Figure 5. Software Stack Running on the Big Data Analytics Server
in offering unified batch processing, interactive SQL, and interactive search and role-based access controls.

Apache Hadoop* is an open source software platform for scalable distributed computing. Written in Java, it runs on a cluster of industry-standard servers configured with direct-attached storage and cost-effectively scales performance by adding economical nodes to the cluster.

MonetDB* is an open source column-oriented database management system designed to provide high performance on complex queries against large databases, such as combining tables with hundreds of columns and multimillion rows. It has been applied in high-performance applications for data mining, online analytical processing (OLAP), geographic information systems, and streaming data processing.

PostgreSQL* is a powerful, open source object-relational database system used for OLTP.

Master Data Management
Dell Boomi* and Dell Toad* can be used to normalize data across disparate data sets to facilitate accurate data analysis. Using the Intel manufacturing example, these Dell tools could ensure

“fab name” is same type of data (text) across the different data sources: vision system, ball attach module, data logs, and yield data.

Extract-Transform-Load (ETL) Tools
Dell Boomi AtomSphere* makes it easier to sync data across applications located on-premises or in the cloud – without using software or appliances. The solution helps eliminate the costs associated with integrating and maintaining integrations with legacy middleware, appliances, or custom code.

Data Analytics
Dell Toad Data Point* is a set of data analysis tools that simplify data access, integration, reporting, and sharing of analytics results.

Dell Statistica* facilitates data mining, predictive analytics, machine learning, and analysis of structured and unstructured data.

Business Intelligence
Dell Toad Intelligence Central* is a set of data reporting tools that provides a centralized repository for up-to-date information, making data provisioning more efficient.

Dell Kitenga* is an analytics suite providing integrated information modeling and visualization capabilities in a big data search and business analytics platform. Combining technologies such as Hadoop for scalability and performance, Lucene/SOLR search, Mahout machine learning, 3D information modeling, and advanced Natural Language Processing, Kitenga is a fully integrated, configurable, cloud-enabled software platform that can be deployed in minutes.

Security
Security solutions from Intel, Dell, and Sonicwall* can protect manufacturing networks, as well as data centers, while sustaining performance and simplifying management. The solution shown in Figure 6 applies to the case where a factory is connected to a data center over the Internet. It enables secure access from virtually any endpoint and protects against viruses, spam, phishing, and other attacks with multiple threat-detection techniques.

Trusted Zone
McAfee Embedded Control maintains the integrity of devices, gateways, and servers by allowing only authorized code to run and only authorized changes to be made. It automatically creates a dynamic whitelist of the
"authorized code" on the system. Once the whitelist is created and enabled, the system is locked down to the known good baseline and no program or code outside the authorized set can run.

**Demilitarized Zone**

*Dell Sonicwall Aventail Advanced Reporting* provides powerful analysis of remote access to network resources. A robust hierarchical log analysis tool, Advance Reporting tracks and evaluates all remote user access to enterprise resources over a secure remote access (SRA) solution.

*Dell SecureWorks* is a security service provider offering advanced endpoint threat detection, penetration testing, and incident response, among other services.

*Dell SecureWorks Security Information Management (SIM)* is a security-as-a-service solution that collects, filters, and categorizes security events from logs and files across virtually any device, around the clock.

*Dell KACE* performs software distribution, inventory, patch management, and more.

**Cloud and Enterprise Data Center**

*Dell Sonicwall SuperMassive 9000 Network Security Appliance Series* is a next-generation firewall (NGFW) that detects and blocks the most sophisticated threats with minimal latency for every network connection, delivering deep security for the enterprise at multi-gigabit speeds.

**IoT Tenets**

The IoT big data solution developed by Dell and Intel is designed to provide security and interoperability from machines to data center or Internet cloud 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 flexible language programming of Intel Gateway Solutions for the IoT simplifies the development of Plug and Play drivers used to connect to sensors and other end devices.

- **Data normalization** through protocol abstraction to improve interoperability
  - Intel Gateway Solutions for the IoT can maintain various device and communication protocols.

- **Broad analytics infrastructure** from edge to cloud to realize customer value
  - Revolution Analytics and Dell supply a comprehensive set of data analytics and business intelligence software that can be used to turn data into actionable insights.

- **Infrastructure** to monetize hardware, software, and data management from edge to cloud
  - This IoT big data solution provides the tools needed to realize substantial improvements in manufacturing performance.
Increasing Manufacturing Performance with the Internet of Things (IoT)

Summary
Big data analytics and Internet of Things technologies provide the critical backbone to enable higher levels of manufacturing performance. The readily-available ingredients presented in this paper map out a smart manufacturing solution capable of making manufacturers more profitable and competitive. Dell and Intel are working to drive a data revolution in manufacturing that will yield new productivity and efficiency gains.

Moving Forward
Dell believes that any organization can start implementing IoT technologies in its own manufacturing facilities. By starting with relatively small projects—perhaps just by analyzing a single process—organizations can build capabilities that result in a competitive edge. For those looking for external guidance, Dell is able to integrate technologies from across Intel, Dell, and other partners. Dell Services is available to provide support and leadership as desired throughout a project. Dell Services can help develop an overall technology strategy and provide program management. They can concentrate on specific pieces of a solution, such as business intelligence and analytics, to help aggregate, consolidate, integrate, and derive insights from data. They can help design a cloud solution—and deliver the right elastic platform from partner or Dell components that are secure and optimized for any environment. Dell Services has a team dedicated to developing custom mobile applications to deliver insights to a company’s team members, who are best positioned to act on the information.

Dell OEM Solutions is available to help bring IoT intellectual property to market. Accelerate time to market through:
- Dedicated OEM team
- OEM product and solution design
- World class supply chain
- Global services and support
- End to end product portfolio

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® Gateway Solutions for the Internet of Things Development Kits
Intel Gateway Solutions for the Internet of Things 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 Dell and IoT, visit dell.com/oem.


3 Source: http://www.isa.org
4 Results might vary depending on package size, process, and equipment used in the manufacturing process.

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL’S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Intel may change specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked “reserved” or “undefined.” Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request. Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order. Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4726, or by visiting Intel’s Web site at www.intel.com.

Copyright © 2014 Intel Corporation. All rights reserved. Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.

* Other names and brands may be claimed as the property of others.