White Paper

Analytics and Big Data

Capitalize on the Untapped Potential of IoT Data

Within a few years, the "IoT" is projected to include millions, if not billions, of connected devices, generating an unfathomable amount of sensor data. With emerging, real-world use cases—such as smart metering, telemedicine, and usage-based insurance—forward-looking enterprises are reducing costs, improving customer satisfaction, and creating new business models based on the inherent value of information.

To capitalize on this emerging mega trend, your organization must do more than collect massive amounts of sensor data. Your success with these new business models hinges on your company's ability to leverage a proven and massively scalable Big Data analytics platform.

Table of Contents

With our long-standing experience and expertise as a trusted IoT data analytics partner to enterprises across every industry globally, we will guide you through these important questions so you can maximize the benefit and minimize the risk from IoT investments.

Executive Summary

In the coming years, enormous volumes of machine-generated sensor data from the Internet of Things (IoT) will fuel a wide range of data-driven products, services, and business processes. But there's a catch: to capitalize on the opportunity, your organization needs to have the right Big Data analytics platform in place.

Personal health monitors tracking your fitness, trashcans monitoring their fullness, watches telling you more than just the time, and agricultural soil monitors saying it's time to water. It seems a day doesn't go by that we don't hear about the latest "offline" thing, device, or equipment becoming "online," moving from isolation to being connected to the IoT. It's clear that integrating sensors, electronics, and network connectivity into devices can enable innovation, enhancing and extending the way we work and interact with each other and the world around us.

Gartner estimates a potential of 26 billion connected devices by 2020[1]. Morgan Stanley even estimates a potential of 75 billion connected devices by 2020[2]. More important than the device estimates, though, are the possible industry transformations and business outcomes. So, the questions we should ask ourselves as business leaders, technologists, and entrepreneurs are; where is the value? What are the challenges? What technology can make it happen?

Global sensor and device connectivity presents countless opportunities for transformation of businesses and industries, disrupting the status quo and existing marketplaces. Companies should capitalize on the opportunity, but need to do it smartly—mitigating financial and operational risks. So, understanding how value can be created from IoT, identifying the business opportunities, and understanding the challenges and the technologies needed to mitigate them are the keys to success.

With our long-standing experience and expertise as a trusted IoT data analytics partner to enterprises across every industry globally, we will guide you through these important questions so you can maximize the benefit and minimize the risk from IoT investments. In this paper, we'll share our perspectives on the value, the business opportunities, the technology challenges, and the ideal advanced analytics technology to seize the opportunity.

The Value of IoT

In order to justify the investment in R&D, infrastructure, and governance required for IoT, senior executives are demanding to understand the value to the business. In our experience, value can be expressed across three dimensions—contextual, integrated, and operational.

1 *Forecast: The Internet of Things, Worldwide, 2013, Gartner, November 2013.*
2 *Morgan Stanley: 75 Billion Devices Will Be Connected To The Internet Of Things By 2020, Business Insider, Tony Danova, October 2013.*

It should come as no surprise that the primary value of IoT is in the data, more specifically, the enhanced decision-making and automation that arise from the convergence of analytical insights, enabled by connected devices, with people. IDC predicts IoT data will account for 10 percent of the world's data by 2020, of about 44 zettabytes[3]. Clearly this is Big Data. But what can this data tell us that we didn't already know and how can we use this information to improve our well-being and optimize businesses?

With contextual data, we can start gaining insights and creating value where previously not possible.

Contextual Insights

Your health-monitoring wristband, the vending machine at the mall, aircraft engine serial number 9AB429—these are all examples of devices that are being instrumented with sensors and connected. What is important to point out about these and every sensor-enabled connected device is the notion of contextual data, the data each device creates, and contextual insights, the inference that can be made about the device and environment through analysis of the data. With contextual data, we can start gaining insights and creating value where previously not possible.

Let's take, for instance, your health-monitoring wristband. Embedded into this device is an accelerometer, which is used to sense movement. As you move, contextual data about your personal activity, such as calories burned, steps taken, and your heart rate is measured, stored, and presented to you through an app so you can monitor, maintain, and improve your personal health fitness. Should you miss a few sessions at the gym, data analytics can create contextual insights, determining this trend and having the device send you an encouraging text that it's time for a run, helping improve your well-being.

Vending machines can be embedded with pressure sensors and counters collecting contextual data on real-time inventory levels, which are made available to an operations manager or fed directly into an order management system. When inventory reaches a specified threshold, local distribution can be automatically dispatched to refill, ensuring no loss in revenue from stock outs and optimizing profit for the business.

Big Data analytics can even enable contextual insights on historical buying patterns and consumer preferences, which lead to improved demand forecasting and product planning.

Across all industries, there are billions of devices and equipment essential for operations. For instance, in the aviation industry, aircraft engines have a vital role in the operations of the airline. Airlines rely on the uptime and performance of their engines and any unplanned downtime (commonly referred to as "time off wing") causes delays to flights—which we all know can lead to a domino effect—negatively impacting customer experiences and future revenue.

By instrumenting industrial equipment such as engines, pumps, valves, and other critical assets with sensors and collecting and monitoring contextual data such as temperature, power, and pressure, enterprises can have the data they need to analyze and gain insights about the equipment (status, performance, and health) while in operation. By analyzing the data in real time, enterprises can have contextual insights

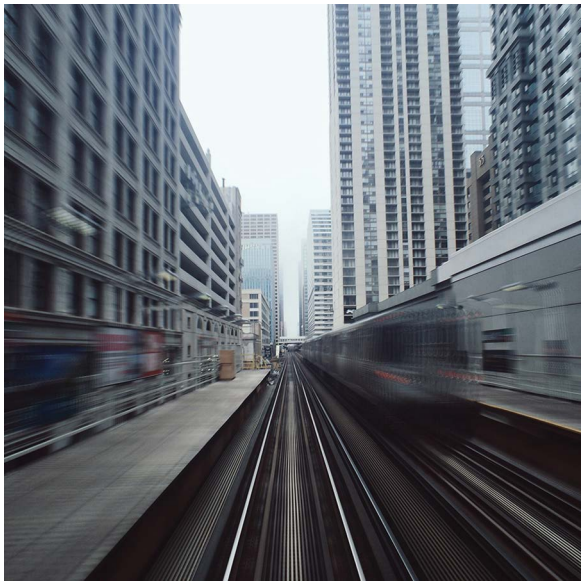
3 *"IDC Predictions 2012: Competing for 2020," Document 231720, Frank Gens, December 2011*

Contextual insights are only one part of the potential value of IoT. When we start considering how individual devices and equipment fit in an ecosystem of other connected devices within the business operations, an even larger opportunity for value creation can be possible.

delivered through advanced notification of potential problems in their equipment. So, in turn, unplanned "time off wing" for the airlines can be minimized, improving operations and revenue, and, of course, getting us to our destination safely and on time.

Contextual insights are only one part of the potential value of IoT. When we start considering how individual devices and equipment fit in an ecosystem of other connected devices within the business operations, an even larger opportunity for value creation can be possible.

Integrated Intelligence



Let's look back to the personal health-monitoring example. While there is value for each person in tracking their own health metrics, imagine the possibilities when you start combining and correlating data across disparate data sets and large populations, creating integrated intelligence. For instance, healthcare providers aspire to provide the highest quality of care for their patients (while, of course, maintaining their balance sheet). If doctors had access to their patients' personal health-monitoring data, they could have a much better understanding of the historical and real-time health and wellness of all their patients and even populations. There could be a future where treatments and prescription dosages are personalized and precise based on real-time conditions of their patients and the environment. Through a Big Data analytics platform, concerning trends in your blood pressure or other measurement could be uncovered and your doctor could adjust your treatment. This data could even be of interest to pharmaceutical companies to better understand drug interaction with environmental impacts and human activity.

Additionally, although contextual insights on each device or equipment are valuable, enterprises and industries have more than one type of equipment critical for operations. For instance, an oil field consists of compressors, pumps, heat exchangers, drilling rigs, artificial lifts, and more. Operators rely upon the uptime and performance for all equipment. They also rely upon integrated intelligence, through predictive failure notification, so that visibility into the entire production fleet will maximize operational gain.

In transportation, automotive companies are instrumenting a vast number of sensors that provide data on each vehicle and the environment around it. VTT Technical Research Centre of Finland has developed a "slipperiness detection" system for vehicles utilizing standard anti-lock brake system sensors. By analyzing the contextual data from the sensors, VTT can infer the slipperiness of the road based on friction, creating contextual insights on the road condition. As contextual insights, the system can warn the driver, avoiding a potential accident. Additionally, with respect to integrated intelligence, every vehicle across a city can report on its environment and in aggregate we can have citywide situational awareness with endless opportunities for transportation optimization.

Analyzing multiple discrete data sources can even uncover correlations, interdependencies, and patterns that lead to new insights, which is not possible by analyzing sensors independently. For instance, looking again at transportation. On-board diagnostics from a vehicle can collect data on driving behavior. Just that information alone can be used to determine an insurance rating for a driver. But by adding in information such as the car's location, the degree of traffic on the road, the weather conditions, and the driver's schedule for the day and heart rate, insurance companies might infer from the integrated intelligence that a stressed driver with a busy schedule is more likely to have an accident and provide "on the spot" incentives—such as a rebate to drive slower or arrive at his next appointment 5 minutes late.

Meanwhile, oil refineries would prefer processing discounted crude instead of premium crude for the financial savings, but doing so can potentially lead to corrosion and scaling of critical equipment, which are detrimental to the expensive assets such as cooling towers and boilers. Refineries today are instrumenting their production process to monitor water quality, flow rate, pressure, and more. A Big Data analytics platform has the potential to find patterns across the various data sets to predict equipment degradation based on historical data. So an optimal balance between processing discounted crude and equipment performance for maximum financial gain can be achieved. That's business optimization.

Operationalizing Insights

New data from IoT, the analytical insights discovered, and the technology investments would be at a loss if the insights are not actionable and delivered to the right people and systems across the enterprise at the right time and in the right format. Therefore, it is essential that actionable insights are integrated within business operations and for all stakeholders.

It is essential that actionable insights are integrated within business operations and for all stakeholders.

In the vending machine example, who should receive the stock out insights—local distribution for re-supply, fleet managers for planning delivery routes, or brand managers for understanding customer preferences? Clearly, all of them. However, given they have different roles that impact the business operations differently, each person needs a different view of the insights. Brand managers may want historical dashboards on their Web browser, local distribution may want predictive alerts summarizing total supply across their geography, and fleet managers may want real-time status on a geographic information system (GIS) map overlaid with driver GPS locations.

IoT enables us to take operations that were formerly labor intensive, requiring physical inspection or actuation, in sometimes hazardous environments and reimagine them in entirely new ways.

Let's say there is a low supply at the stadium before a big game. An application tracks the priority events, locations, and supply levels in real time. An operations manager, who is utilizing the dashboard, can make better decisions about how to manage inventory levels, then relay this information to drivers through tablets to re-route, optimizing business value. Maximizing the return on investments in a Big Data platform and IoT technologies require delivering the actionable insights to the people and systems that can best leverage them as well as having the right enabling technologies and tools to do so.

Through connectivity and increased automation, devices such as smart utility meters or streetlights can be remotely monitored and controlled. IoT enables us to take operations that were formerly labor intensive, requiring physical inspection or actuation, in sometimes hazardous environments and reimagine them in entirely new ways.

Real-time, automated control can even further increase the value of insights. In some scenarios, with the right confidence and security protocols, equipment and processes can self-adjust based on analytics. By integrating multiple data sources, integrated intelligence can enable real-time, closed feedback loops. For instance, General Electric and Siemens are manufacturers of wind turbines and by analyzing wind speed and direction in real time, their turbines can self-adjust their pitch and rotor blade angles, maximizing energy production.

Additional Data-Driven Use Cases

Across a wide range of industries, organizations are running advanced analytics on sensor data to enable new value-added products, services, and business processes. All of these use cases are driven by the ability to analyze data collected from sensors and M2M communications.

Here are some examples of the ways in which organizations are putting sensor data to work:

- **Fleet management:** Sensor data from delivery trucks is helping businesses schedule preventive maintenance before mechanical issues can disrupt fleet operations. Companies are also using sensor data to enable intelligent route optimization to reduce fuel costs and emissions harmful to the environment.
- **Healthcare sensing:** Biosensors are now used to enable better and more efficient patient care across a wide range of healthcare operations, including telemedicine, telehealth, and mobile health. Increasingly, most medical and life sciences equipment is connected for remote service.
- **Product monitoring:** Manufacturers use sensor-data analytics to monitor the health and performance of their products and to work proactively to address service and maintenance issues before they lead to product downtime.
- **Predictive maintenance:** Airlines use data from airplane sensors to proactively manage maintenance, improve reliability, reduce unplanned service work, and mitigate risk.

- **Smart meters:** Utilities are deploying digital, networked meters to systematically feed analytics tools and Web-based portals with consumption data. This enables new services, such as allowing customers to review the details of their electrical, gas, and water usage. In the back end, sensor data helps utility personnel identify losses due to theft and faulty meters.
- **Usage-based insurance:** Insurance companies use data generated from sensors in automobiles to offer drivers rates based on the amount of driving they do, their driving habits, and even where they drive and park. In addition, insurance companies can now perform predictive modeling on vehicle data to identify lower- and higher-risk customers.

Technology leaders need to be at the center of the enterprise leading the company and overcoming the hurdles for the business.

The Challenges

Unlocking the value to the Internet of Things doesn't come without challenges and requires business leaders to partner with their organizations' technology leaders.

Creating intelligent devices with sensors and processing as well as enabling secure and reliable connectivity are just the beginning. Enabling secure data, managing the data, processing insights, timely insight delivery, and cost management are some of the big challenges. Technology leaders need to be at the center of the enterprise leading the company and overcoming the hurdles for the business.

The Business Challenges and Technology Requirements for the Internet of Things

Business Challenges	Technology Requirements
Delivering business value when it's needed	■ Balancing compute at network edge and cloud to optimize latency, insights, and TCO ■ Batch, real-time, and interactive analytics ■ Flexible languages for writing queries, scripts, or machine learning ■ Flexible schemas ■ Predictive analytics Multiple data sources and aggregating data lakes ■ Enabling multiple technology stakeholders—BI analysts, programmers, data scientists ■ Access to apps when device or user is offline
Delivering business value where it's needed	■ Operationalizing the insights in real time, closed loop to operations ■ Multi-tenancy ■ Environment for agile application development ■ Enabling the various stakeholders and departments that need the insights ■ Flexible app deployment—Web and mobile platforms ■ Connectors to multiple business intelligence tools
Securing and governing the data, apps, and users end to end	■ Managing cybersecurity threats with new data from sensitive equipment ■ Securing data across the network and cloud ■ Securing control systems from intruders ■ Compliance and regulatory requirements from data ■ Regulation and authorization of applications and data ■ Managing data lifecycle ■ Privacy of sensitive data

Continued on next page

Vertica Analytics Platform offers the ideal platform for running analytics on massive volumes of sensor data.

Business Challenges	Technology Requirements
Storing and analyzing massive data at scale	■ High-velocity data ■ Accessibility, performance, and speed for multiple applications ■ Flexible, scale-out infrastructure ■ Minimizing total cost of ownership (TCO) ■ Control of IT operations, but enablement of data ■ Leveraging existing investments, but optimizing for new value potential and future requirements ■ Minimizing data center footprint and energy consumption ■ Open, standards-based architecture
Breaking down the data silos	■ Analytics across multiple data types and sources ■ Single pane view and dashboard of all critical assets for operations, consolidated across multiple vendors ■ Data ownership and access for new IoT data ■ Connectors to large ecosystem of data types ■ Data lake across multiple clouds ■ Open APIs

Figure 1. This table shows significant business challenges with corresponding technology requirements to maximize the value and minimize the financial and operational risks from IoT.

The Advanced Analytics Technology That Makes It Happen

To enable new IoT use cases, it's not enough to collect massive amounts of IoT data from across the enterprise. The real challenge is to analyze the data you collect to extract meaningful information in a timely manner.

To capitalize on data-driven business models, you need a robust advanced analytics platform that allows your organization to intelligently manage and analyze volumes of sensor data in a cost-effective manner. That's how Vertica helps you manage and derive value from Big Data.

Vertica offers the ideal platform for running analytics on massive volumes of sensor data. Among other advantages, the platform delivers the performance and economies of columnar storage and aggressive data compression; the scalability of a multimode architecture; and advanced analytics through the use of SQL extensions and in-database machine learning functions.

Columnar Storage and Execution

When it comes to analytic workloads, columnar storage offers significant gains in performance, storage footprint, and efficiency. These gains stem in part from the ability of the SQL analytics platform to read and retrieve only the columns you need, instead of reading and retrieving all the columns in the database.

Unlike traditional database vendors who struggle with retrofitting columnar storage into their legacy code for marginal gains, the columnar orientation of the Vertica Analytics Platform was deliberately designed into the core of the platform from day one.

All components of Vertica are columnar-aware for superior compression and encoding, along with better and more efficient relational join performance. The result: advanced analytical queries and reports alike run orders of magnitude faster against much larger data sets.

Aggressive Data Compression

To enable aggressive and simultaneous data compression and encoding, the Vertica Analytics Platform organizes values of similar data types contiguously in memory and on disk. This approach allows for dramatic footprint reduction and blazingly fast parallel load and query times.

Based on data cardinality and workloads, the platform's Database Designer automatically chooses from more than a dozen encoding and compression algorithms to find the top fit. The unique sorting capability of Vertica Analytics Platform further enhances the compression of the data. The result: you can achieve more with less hardware.

A Scale-Out MPP Architecture

The speed, scale, and rapid growth of today's information exceed what even the largest of supercomputers could ever handle alone. This is why massively parallel processing (MPP) systems are becoming the norm. Vertica Analytics Platform delivers a simple, yet highly robust and scalable MPP solution with linear scaling and native high availability, all on industry-standard hardware.

Unlike other MPP analytic platforms, Vertica has no leader nodes, and therefore no single point of failure. Any node in a cluster is capable of initiating loads or queries, and will evenly distribute the workload to other nodes when it makes sense to do so. Workload distribution is automatic—no administrator or user intervention is required.

Like everything else in Vertica, the MPP layer was an early and deliberate design decision, present in every release of the platform from its inception. This means every component in the advanced analytics platform is fully aware of the distributed processing capability. In contrast, many traditional relational database management system (RDBMS) vendors are now attempting to bolt on MPP elements to their legacy symmetric multiprocessing (SMP) platforms.

Unlike other MPP analytic platforms, Vertica has no leader nodes, and therefore no single point of failure.

Vertica's in-database machine learning capabilities allow users to build models and train them on historical data within Vertica, which saves time and money by building these models where the data resides.

Advanced Analytics and In-Database Machine Learning

Putting IoT data to work requires an analytics platform that gives both data scientists and business analysts the tools they need to uncover valuable insights. Vertica combines the familiarity of SQL analytics with the power of in-database machine learning so organizations can leverage a wide range IoT enabling analytical functions, such as time series, geospatial, logistic regression, k-means clustering and more.

Vertica's MPP architecture is also able to easily handle geospatial data from mobile sensors and conduct a wide variety of spatial analytics queries in fractions of a second, as compared to other solutions that would take minutes to run comparable queries. Vertica can also combine geospatial data with semi-structured sensor data and 3rd party data sources such as weather data—enabling increasingly complex use cases.

Many IoT data do not have the same units of time measure, garner data at differing intervals, or have time-gaps in measurements due to sensor failure. Time Series analytics puts data onto the same time scale and uses proprietary algorithms to interpolate missing values and get rid of outliers.

Vertica's in-database machine learning capabilities allow users to build models and train them on historical data within Vertica, which saves time and money by building these models where the data resides. The machine learning algorithms can then be run to enable IoT use cases such as predictive and preventative maintenance, which identifies potential asset faults before they happen, reducing the downtime of mission-critical assets.

If you want to go beyond the in-database functions, Vertica natively supports R for statistical computation. There are many public R libraries and functions for modeling and classification that can be easily installed on Vertica. The platform's R implementation runs in-database and allows R functions to be parallelized to take advantage of Vertica's underlying MPP platform. Vertica's SDKs allow for the development of R, Python and other functions and the integration of the wealth of public functionality. Regarding presentation of data, many different commercial and open source visualization tools can be used with Vertica.

Vertica Advanced Analytics Capabilities

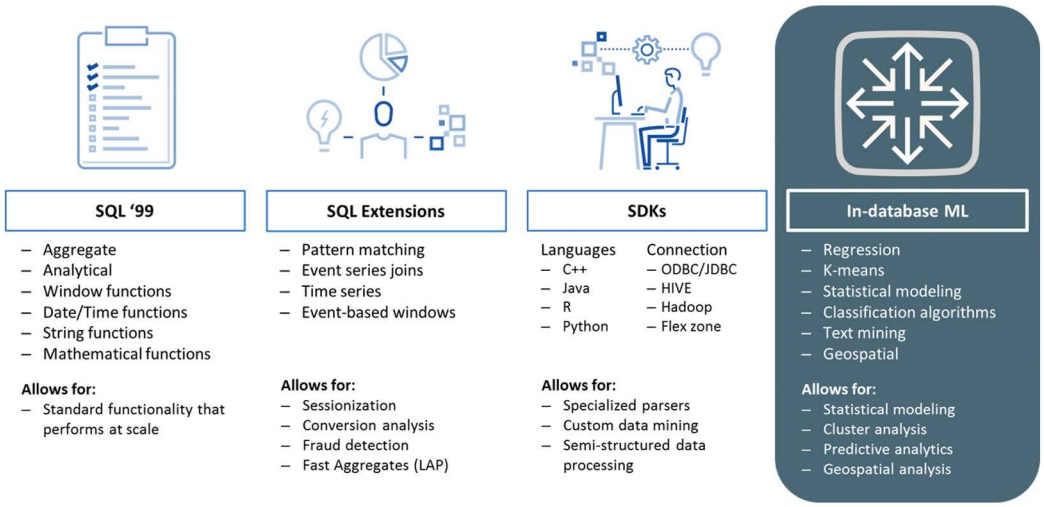


Figure 2. Putting IoT data to work requires an analytics platform that gives both data scientists and business analysts the tools they need to uncover valuable insights.

Vertica helps organizations across industries and use cases manage and analyze massive volumes of sensor data to predict and prevent operational issues and reduce service costs, improve customer satisfaction by extending operational uptime, and bring revenue-generating IoT solutions to market.

Big Data Analytics in Action: Case Studies

Vertica helps organizations across industries and use cases manage and analyze massive volumes of sensor data to predict and prevent operational issues and reduce service costs, improve customer satisfaction by extending operational uptime, and bring revenue-generating IoT solutions to market.

Suunto Unlocks Fitness Coaching Market with Big Data Project

Suunto, a subsidiary of the Amer Sports Corporation, manufactures and markets sports watches, dive computers, compasses and precision instruments helping consumers technologically connect to their active and adventurous lives.

Driven by technology improvements over the last decade, data-driven personal devices such as activity or fitness trackers have become increasingly popular amongst both professional and amateur athletes as well as sports enthusiasts. These wearables

"With the data stored in Vertica, we easily pull out current facts about how people are training and hundreds of thousands of people now benefit from the tools we provide based on this Big Data analysis."

JANNE KALLIO
Performance Business
Digital Leader
Suunto

monitor and track fitness-related information such as distances travelled, altitude climbed, dive depth, heartbeat and calories consumed. Many users frequently link their devices to smartphones for long-term data tracking.

"Over recent years, our market has changed considerably as cloud-connected devices became more prevalent," explains Janne Kallio, performance business digital leader, Suunto. "With every measurement entering the cloud, we've an opportunity to analyze this data and bring new value to our customers. We aim to create innovative products by using Big Data as a differentiator in a very competitive market."

Suunto wanted to enrich the customer experience and build brand awareness of its extreme sports watches and outdoor instruments by analyzing all the data generated by these devices. For example, analyzing fitness data in real-time allowing users to gain valuable insight into their training.

Industry: Electronics

Location: Vantaa, Finland

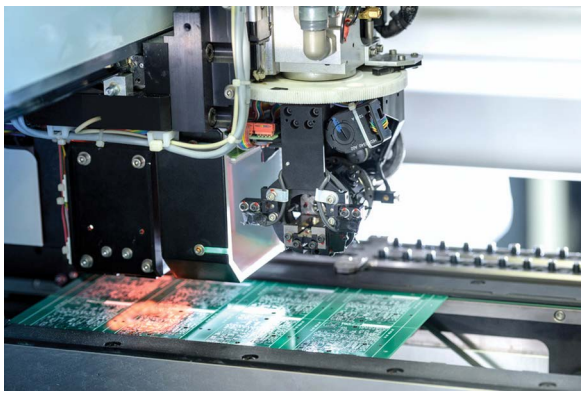
Challenge: Required a high-performance database to ingest and analyze massive data volumes created by customers' coaching solutions and enhance the user experience.

Solution: Adopted Vertica Analytics Platform to analyze customer fitness data and boost customer experience.

Results:

- Manages and analyzes massive data volumes rapidly and reliably via a highly structured database, eliminating the limitations of traditional data warehouse technology.
- Constantly updates the data pool, generating highly accurate progress and performance information to users.
- Boosts brand awareness and enhances the user experience substantially, delivering a competitive edge in a highly aggressive market.

Optimal Plus Builds a Foundation for IIoT Analytics and Actionable Intelligence—and Drives Manufacturing Excellence—with Vertica



Production issues are costly in the semiconductor and electronics industries. With its test management and manufacturing intelligence solutions, Optimal Plus helps suppliers improve product quality, yield, throughput, and performance. The company's software also enhances planning and decision making, while unmasking operational risks.

Headquartered in Holon, Israel, Optimal Plus last year processed data on 50 billion ICs and boards, which was up 15 billion year over year. Compared to homegrown data-collection alternatives, the company's holistic software enables seamless visibility across distributed supply chains. Developed for the Industrial Internet of Things (IIoT), the software connects to various plant assets, sensors, and devices to record, store, and analyze continuous streams of data.

Industry: Software Technology

Location: Holon, Israel

Challenge: Improve the delivery of time-sensitive data and analytics of semiconductor/electronics manufacturing data

Solution: Antegrate a scalable, high-performance data analytics engine with the company's solutions

Results:

- Aids collection, organization, and storage of data from 50 billion semiconductors and printed circuit boards a year
- Outpaces indexing, granularity, and query response times vs. legacy database—analytics on 2 billion data points done in less than a minute
- Runs a performance demo in 2.5 minutes against 200,000 equipment test results
- Fosters uptime of plant assets by enabling models of historical data vs. streaming data and algorithms that predict equipment faults

"One of our key buying factors is that Vertica's performance overall is excellent on huge volumes of data. This helps accelerate the speed of decisions and analytical insight."

MICHAEL SCHULDENFREI
Chief Technology Officer
Optimal Plus

Key Takeaways and Resources

IoT creates the opportunity for your organization to capture actionable information and to implement new data-driven business models that help you gain a competitive edge. However, to capitalize on this new opportunity, you need to do more than collect data. You need to arm your organization with a powerful set of tools that you can use to import, store, and analyze massive amounts of data in a timely, cost-effective manner.

With high-performance, scalable analytics, you're poised to make effective use of the enormous volumes of data from the IoT—to improve your products and services, enhance your operational processes, and bring new data-driven offerings to market. With the Vertica Analytics Platform, you have the right partner and analytics tools in place to take advantage of the massive potential of IoT data.

Learn More

Solution Page—IoT Analytics

Harbor Research Whitepaper—IoT Challenges Require a New Approach to Data Management and Analytics

Webcast—Advanced Analytics & Manufacturing Intelligence with Optimal+ & Vertica Webcast

Additional contact information and office locations:
www.vertica.com

www.vertica.com

VERTICA | MICRO FOCUS®