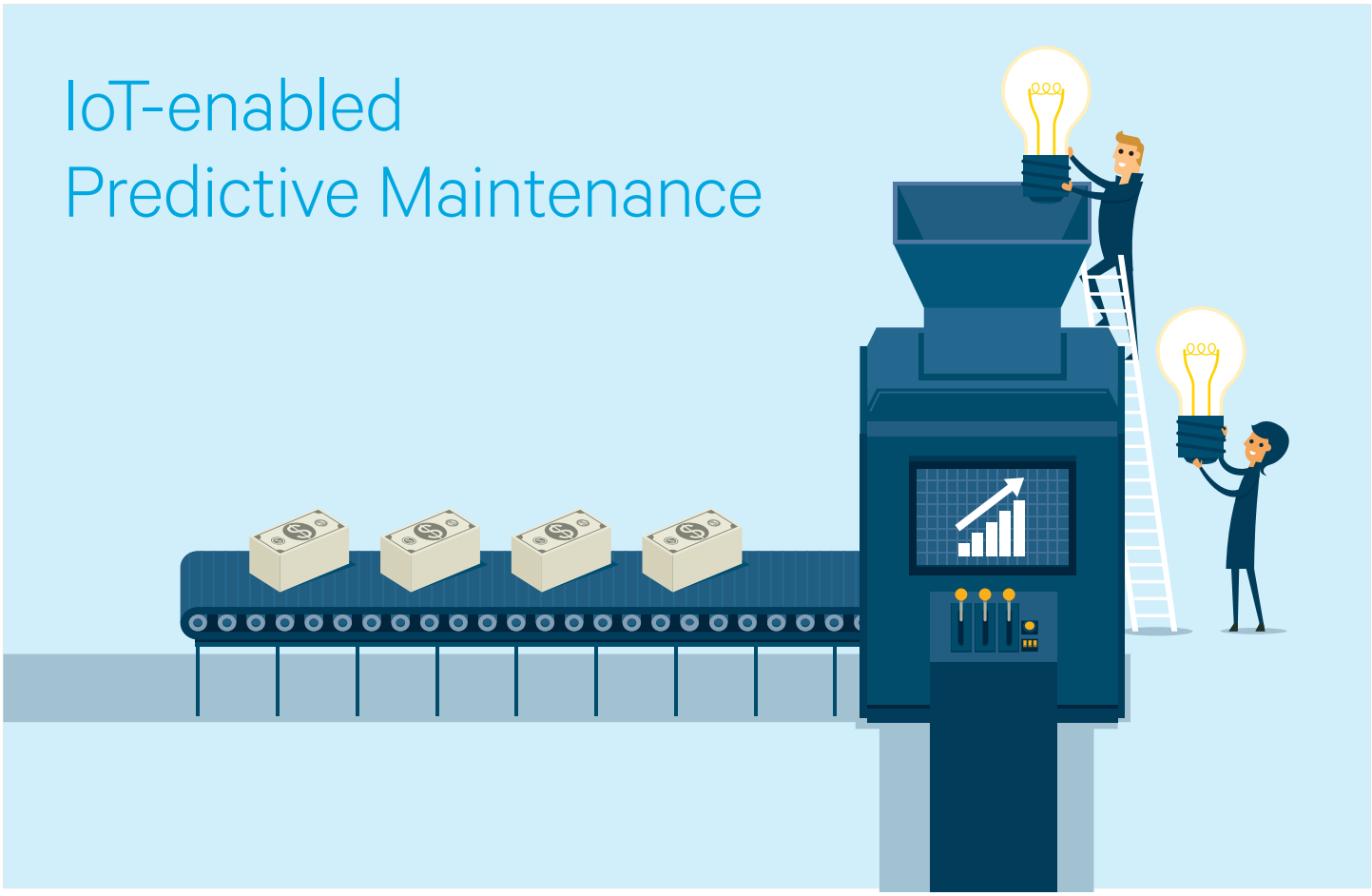
cloudera

WHITE PAPER

Driving Predictive Maintenance in a Connected World



Version: 105

cloudera

Table of Contents

cloudera

Predictive Maintenance in the industrial world

Powered by sensors, connectivity and smart machines, the Internet of Things (IoT) is reshaping the manufacturing and industrial processes, effectively changing the paradigm from one of 'repair and replace' to more of 'predict and prevent'. In an industrial scenario, unplanned equipment downtime can be extremely costly to business. Today, manufacturers and organizations in other asset heavy verticals cannot afford to wait till a machine or equipment breaks down in order to figure out what went wrong. On the other hand, enterprises also don't want to spend costly time and resources doing unwanted maintenance to all of their equipment and machinery without really needing to do so.

Predictive Maintenance, in a nutshell, is all about figuring out when an asset should be maintained, and what specific maintenance activities need to be performed, based on an asset's actual condition or state, rather than on a fixed schedule, so that you can maximize uptime and productivity. It is all about predicting & preventing failures and performing the right maintenance routines in order to reduce costly equipment downtimes.

So why is Predictive Maintenance gaining increasing importance today, especially among asset intensive verticals? It is driven primarily by three key factors:

- **Unplanned outages equals lost revenues** – For asset heavy industries, unplanned equipment outages can mean big losses in revenues and productivity. For example some of the leading automotive manufacturers estimate that every minute of unplanned downtime can cost them as much as $15,000 - $20,000 and a single downtime event can cost approximately $2 Million. Given the business impact, it is not surprising that these industries have been focusing on driving predictive maintenance to minimize downtime and losses
- **Unnecessary costs and work** – It costs valuable resources, time, and material to do ongoing maintenance activities and any unnecessary maintenance done leads to unwanted costs and redundant work. Leading organizations want to ensure that they are doing the right level of maintenance, and replacing the right parts for the right machinery at the right time.
- **Moving from reactive to proactive** – Currently, most organizations employ a combination of corrective and preventative maintenance on their machinery to minimize the impact of unplanned downtime. However most of the current methodologies are still reactive and after-the-fact, or leads to redundant and unnecessary work. With predictive maintenance organizations are moving towards effectively predicting and preventing issues before they impact their business.

For some of the leading automotive manufacturers, unplanned downtime can cost them as much as $15,000 - $20,000/ minute and a single downtime event can cost approximately $2 Million.

Data generated per day by a typical offshore oil platform

1-2TB

Source: Cisco, Whitepaper - A New Reality for Oil & Gas

cloudera

According to McKinsey research, implementing predictive maintenance using IoT can reduce equipment downtime by as much as 50 percent and reduce maintenance costs of factory equipment by 10 to 40 percent.

Industrial IoT and Predictive Maintenance

Today, IoT and connected devices are essentially re-shaping the predictive maintenance landscape. By capturing and utilizing data streaming from sensors and connected devices, businesses can now gain visibility into the condition of their valuable assets and specific components in real time. According to research from McKinsey, implementing predictive maintenance using IoT can reduce equipment downtime by as much as 50 percent and reduce maintenance costs of factory equipment by 10 to 40 percent.

One of the key enablers and pillars for driving predictive maintenance is **Condition Based Monitoring (CBM)**.

The goal of condition based monitoring is to spot upcoming equipment failure so maintenance can be proactively scheduled when it is needed – not before and not after. Equipment failures, usually, doesn't happen in a split second. As indicated in Figure 1 below, the asset condition continuously degrades over time before it breaks down. So you really have a window of opportunity in the early stages, when failure begins, to identify and detect the signs of potential breakdown. The goal is to detect the onset of the failure right when the asset condition starts to deteriorate, in the early stages, in order to identify and prevent an outage or failure. The earlier you can detect the signs of failure, the easier it is to fix and the cheaper it is.

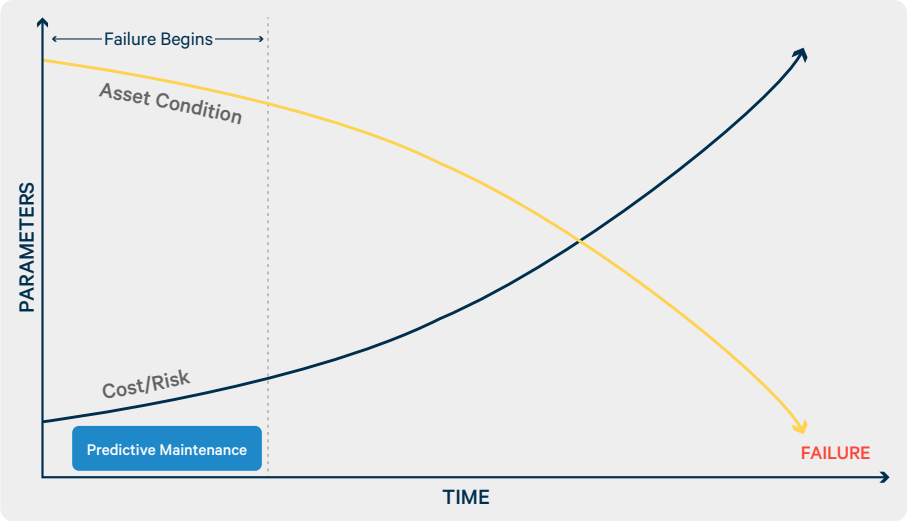


Figure 1: Condition-Based Monitoring

So, how do you effectively detect the onset of failure?

This is where IoT, sensors, and connected devices come into play. With IoT, you can now have sensors that continuously monitor key attributes or performance indicators of machines, such as temperature, pressure, vibrations/second, noise levels etc. Utilizing IoT and sensor data from connected equipment and machinery, organizations can now monitor the health of their assets, effectively predict when an asset might fail, detect variances, understand warning signals, and quickly identify any patterns that might indicate a potential breakdown. Data centric organizations are using machine learning to detect anomalies or patterns that are indicative of failure and intervene as soon as initial signs of failure are detected to perform the right maintenance activities. Early detection is the key to reducing failures and costs.

Today, organizations can effectively carry out advanced analytics and machine learning on sensor data to accurately predict the chances of a machine being down — including when and how an asset will fail. This enables early and corrective measures to be planned and introduced in the most effective way, thereby avoiding unplanned downtimes and costly staff and resources.

Thus, IoT enabled predictive maintenance enables organizations to continuously monitor the health of their assets and enable corrective measures to be introduced early, thereby avoiding unplanned downtime and reducing the burden of costly staff and resources. IoT enabled predictive maintenance can have a profound impact, especially on asset-heavy industries such as manufacturing, automotive, mining, energy and utilities, etc., delivering lowered costs associated with maintenance, lowered equipment downtimes, and improved productivity and quality.

cloudera

IoT Data Management platform

An effective data management platform for IoT should:

- Drive cost efficiency - Low cost/ TB
- Handle the volume & variety of IoT data
- Enable predictive analytics and machine learning
- Enable analytics in real-time
- Be fundamentally secure
- Support both Cloud & On -prem deployments

Key Data Management Challenges for IoT

Even though there is significant value to be gained from the continuous monitoring of assets and enabling predictive maintenance, organizations often struggle to keep up with the volume, variety, and velocity of data that these connected machines generates. Data streaming from sensors quickly adds up and managing petabytes of data from millions of sensors, and driving analytics on data in real-time, is one of the biggest challenges organizations have to account for when it comes to successfully implementing IoT use cases.

Some of the critical challenges and key considerations that organizations face with respect to data management for IoT enabled predictive maintenance include:

- **Cost of Data Management:** Traditional data management mechanisms tend to be notoriously expensive, do not scale easily, and were not built for capturing and processing the petabytes of IoT data streaming from connected devices. Today, organizations need a more flexible and scalable data management and analytics platform that can easily ingest, store, manage, and process all of these streaming data from IoT sources at a lower cost per terabyte.
- **Ability to handle the volume and variety of IoT Data:** To enable condition based monitoring and emerging IoT use cases, information architects need a platform that can handle not just traditional structured data, but all types of diverse data structures and schemas, including everything from intermittent readings of temperature, pressure, and vibrations per second to handling fully unstructured data (e.g., images, video, text, spectral data) or other forms such as thermographic or sound/noise from machines.
- **Managing the complexity of real-time data:** In order to drive continuous monitoring and predictive maintenance, the data management platform needs to enable real-time analytics on streaming data. What that means is that the platform needs to have capabilities to effectively ingest, store, and process the data streaming in from sensors in real time or near-real time in order to instantly deliver insights and action.
- **Diverse Analytical Capabilities:** Existing platforms offer limited ability to provide insights and analytics into product usage and performance. For real-time predictive maintenance, you need a platform that can provide a wide range of analytical options — including everything from SQL analytics and search capabilities to tools to support machine learning and modelling, along with tight integration with leading business intelligence (BI) solutions.
- **Predictive Modeling Capabilities:** Predictive modelling capabilities are key to delivering insights and current platforms provide limited to no modeling or machine learning capabilities to predict and prevent issues before they impact the operations.

Thus, in order to effectively drive condition-based monitoring and predictive maintenance capabilities, organizations need a scalable, elastic, and cost-effective data management platform that can handle the volume, variety and velocity of data presented by IoT. Furthermore, the platform should be able to handle the complexity of both data at rest as well as data in motion, offer enterprise-grade security and management tools, and deliver a range of analytical options including proven machine learning and predictive modelling capabilities.

Cloudera Enterprise — The Data Management Platform for IoT

More and more organizations functioning in asset-heavy industries are utilizing Cloudera's Enterprise Data Hub (EDH), powered by Apache Hadoop and leading open source technologies, to effectively drive predictive maintenance and other key IoT use cases in order to improve performance, reduce costs, and enable insights in real time.

With Cloudera Enterprise, organizations can easily bring together sensor data along with data from diverse manufacturing systems and multiple sources into a single, unified platform at considerably lower cost. Since Hadoop is built on a highly scalable and flexible file system, any type of data—including structured data from manufacturing systems or historians or streaming data from sensors—can be loaded into the platform without altering its format, in order to perform machine learning and analytics to predict failures. Data generated by machines and sensors, including log files, can be collected in real time and streamed directly into an EDH instead of being staged in a temporary file system or data mart. By enabling both real-time stream processing and streaming data integration, Cloudera helps users ensure all decisions and processes can be informed by data.

And because Hadoop uses industry-standard hardware, the cost per terabyte of storage is, on average, 50-85% cheaper than a traditional data management mechanisms.

cloudera

Cloudera for IoT

- Effectively handles both data in motion and data at rest
- Easily ingests millions of events/sec
- Industry leadership in Spark
- Real-time processing and analytics
- Supports hybrid Cloud deployments
- Data Science for the Enterprise
- Data security beyond compromise
- Proven success across diverse IoT use cases

Cloudera for IoT and Predictive Maintenance

Some key attributes of Hadoop and Cloudera Enterprise that make them perfect for driving IoT-enabled predictive maintenance include:

- **Real-Time Data Ingest:** Easily ingest data from multiple data sources, including both batch and real-time data ingest from connected devices and sensors, using tools such as Apache Kafka and Apache Flume.
- **Manage Data Variety:** A flexible and scalable platform that can handle all types of structured and unstructured data sources. Based on specific use cases, the platform can easily manage multiple data formats, structures, and schemas—from structured data from diverse manufacturing systems to intermittent sensor readings of temperature and pressure or live video and audio feeds.

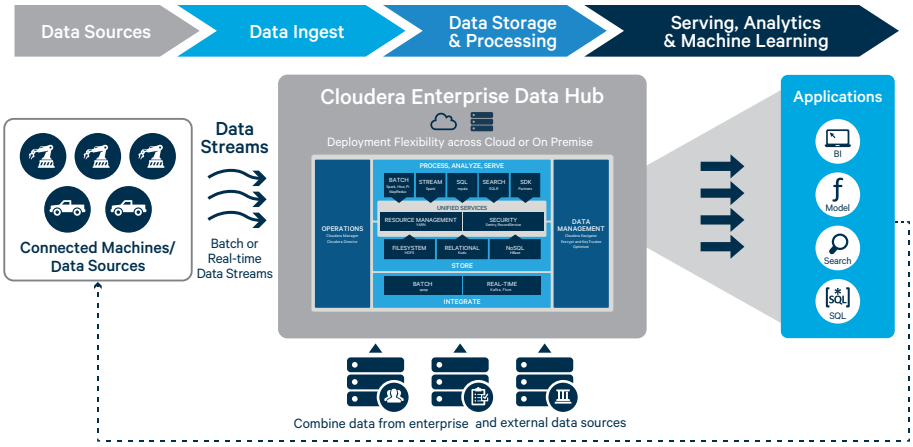


Figure 2: IoT Data Management Value Chain Using Cloudera Enterprise

- **Fast Analytics on Fast Data:** Apache Kudu provides the storage substrate for fast analytics on streaming data—providing a combination of fast inserts and updates alongside efficient columnar scans to enable IoT type real-time analytic workloads across a single storage layer.
- **Real-Time Analytics:** Enables real-time data processing on streaming data using Apache Spark and Spark Streaming, supported by storage options like Apache HBase and Apache Kudu.
- **Machine Learning Capabilities:** Provides out-of the-box machine learning libraries with Apache Spark that enable organizations to easily build predictive models and continuously iterate on them.
- **Enable Data Science for the Enterprise:** Accelerate exploratory data science at scale and build machine learning models using Cloudera Data Science Workbench by taking advantage of massively parallel compute and expanded data streams.
- **Diverse Analytical Options:** Offers diverse analytical engines, including search and SQL analytics, with tools such as Apache Impala (incubating), statistical analysis, and machine learning libraries, to suit the diverse needs within your organization. Moreover, Cloudera is already integrated with industry-leading BI and visualization tools, so that organizations can continue to leverage their existing BI tools and assets.
- **End-to-End Data Security:** Delivers complete protection of sensitive data, including total encryption, key management, access policy enforcement, audit, lineage, and complete data governance capabilities using Cloudera Navigator.

cloudera

Key IoT Predictive Maintenance Use Cases Supported:

- Connected Vehicles
- Connected Factories
- Connected Machinery
- Industrial IoT
- Heavy Equipment
- Smart Buildings
- Smart Ports
- Smart Airports

Predictive Maintenance Solution — Delivery Models

When it comes to putting all of this together, Cloudera offers a variety of options for building and deploying a predictive maintenance solution. Organizations have a choice of **a)** building a custom solution, **b)** assembling a solution using some of our leading certified technology partners, or **c)** buying a prepackaged solution from systems integrators to deploy the use case.

- **Build:** Organizations can work with Cloudera's professional services team to build a customized predictive maintenance solution that fits their specific industry and business needs. Our service professionals will work hand in hand with client teams to engineer a solution that can be implemented on-premise, in the cloud, or in a hybrid environment, leveraging the appropriate components of data ingestion, storage, processing, and analytics mentioned earlier. A typical implementation time for a build model ranges from eight to nine weeks.
- **Assemble:** Cloudera can also work with clients to assemble a solution based on the Cloudera Enterprise platform along with some of the leading technology partners from our partner ecosystem. These partners have created various components and accelerators for simplifying and providing faster time-to-value for data ingestion, data processing, machine learning and analytics. Cloudera certifies that these technologies and solution providers go through a rigorous certification process. Depending on the partner solutions deployed, a typical engagement will take around nine weeks for implementation with flexible deployment options.
- **Buy:** A third option is to buy a pre-built and proven joint solution from our partners that is deployed on top of Cloudera's enterprise data management platform. Some of the leading Systems Integrators (SIs) have created end-to-end solutions for managing and implementing remote monitoring and predictive maintenance, using Cloudera Enterprise, that offer pre-built integrations, configurations, and compatibility to get started quickly.
  - An example of one such solution is the TCS Sensor Data Analytics Framework, which is a highly customizable big data framework, powered by Cloudera, to collect, store and analyze a variety of sensor and log data from IoT deployments. Examples of its applications are predictive maintenance, machine diagnostics, telematics processing, remote monitoring, and early warning systems.
  - Another example of a proven joint solution for predictive maintenance and monitoring is the Kogentix Asset Maintenance Prediction (AMP) solution. Kogentix AMP makes it possible to build, manage, monitor, and deploy AI applications leveraging cost effective, scalable open source software. Business leaders and data scientists can focus on using their data for superior results, not scripting and managing a wide variety of ever changing tools. For data to have an impact, the entire data science process must be automated and operationalized. Kogentix AMP delivers.

cloudera

Case Study 1 - Navistar

Using sensor data from 300,000 connected vehicles, Navistar is able to lower maintenance costs by up to 40 percent.

Predictive Maintenance — Customer Use Cases

Leading organizations across the globe are adopting Cloudera's Enterprise Data Hub, as the data management and analytics platform for driving predictive maintenance and IoT. Provided here is a summary of some of the key customer use cases, across different verticals, that highlights how some of the leading organizations are utilizing Cloudera and the power of IoT analytics to drive predictive maintenance.

Customer Case Study - 1
Industry: Automotive Manufacturing
Driving Predictive Maintenance for Connected Vehicles

Customer Overview:

Navistar is a leading manufacturer of commercial trucks, buses, defense vehicles, and engines, widely known for its International® Truck and IC Bus® brands.

Challenge

Typically, vehicle manufacturers schedule vehicle maintenance based on miles traveled or time since last appointment. But these are very rudimentary approaches and needed a much more data driven model to minimize truck downtime and unplanned maintenance.

To help fleet and vehicle owners move from a reactive approach to a more predictive model, Navistar needed to analyze a wider range of data in real time, including vehicle sensor data. However, its traditional data warehouses couldn't support the growing volume of fast-moving, high-volume telematics data. "As we collected more data, the analytic process slowed to a near halt on our legacy systems," said Ashish Bayas, chief technology officer, Navistar.

Solution

Navistar built an IoT-enabled remote diagnostics platform, called OnCommand® Connection, on Cloudera Enterprise. The platform brings in over 70 telematics and sensor data feeds from more than **300,000** connected vehicles—including engine performance, truck speed, acceleration, coolant temperature, and brake wear. This data is then correlated with other Navistar and third party data sources, including meteorological, geolocation, vehicle usage, traffic, historical warranty, and parts inventory information. The platform currently stores over 60 terabytes (TB) of data and uses machine learning and advanced analytics to automatically detect engine problems early and predict maintenance requirements.

Fleet and vehicle owners can now monitor truck health and performance from smartphones or tablets, prioritize needed repairs, and quickly identify the nearest dealer service locations that have the relevant parts in stock, available technicians, and available service bays.

Results

With OnCommand Connection, Navistar has helped fleet and vehicle owners reduce maintenance costs by up to **40 percent**. One Navistar customer reduced the maintenance cost-per-mile for its vehicles, which previously was 12 to 15 cents, to less than three cents.

Early detection also reduces vehicle downtime by up to 40 percent and minimizes towing costs. In addition, when downtime occurs, vehicle owners typically lose US$1,000 in revenue per vehicle daily. With over 300,000 vehicles across 2,300 customers, the total impact can be significant.

cloudera

Case Study 2

IoT solution based on acoustics helps plant operators continuously monitor the health of their assets and potentially predict turbine, generator, and other equipment issues

Customer Case Study - 2
Industry: Energy & Utilities
IoT based Acoustic Monitoring Solution for Hydroelectric Power plants

Customer Overview:

Our Client is a leader in building industrial turbines, generators, machinery, and solutions for hydroelectric power plants. They also offer industrial automation solutions to help companies improve operations efficiencies and productivity.

Challenge

Our Client's tremendous footprint in the hydropower market gives its staff unique insights into industry challenges. Even a few minutes of downtime for these massive turbines can mean significant costs due to lost productivity. And that meant hydropower plants were still dependent on sending experienced inspection personnel every week to check on the condition and health of equipment. The challenge our Client was trying to solve was: How could we automate inspections and monitor the health of these assets so they could learn about potential equipment issues earlier and avoid costly downtimes?

Solution

Their answer: An innovative IoT based predictive maintenance solution for hydroelectric power plants based on acoustic monitoring. The solution captures, analyzes, and interprets sounds in hydropower plants to diagnose the health of the turbines and generators. Just like an auto mechanic diagnoses problems based on sounds coming from an engine, they are able to analyze the noises coming out of the machinery in the power plants and can provide early warnings based on the acoustic signals.

In order to create this innovative solution, they teamed with Cloudera to implement an advanced analytics and machine learning platform for its IoT data. The solution uses acoustic sensors to capture turbine, generator, and equipment noises. Each plant has multiple power units, which each generate approximately 60 gigabytes (GB) of data per month. Just connecting 1,000 power units, generates roughly 60 terabytes of data a month.

Cloudera Enterprise Data Hub, enables them to capture, store, analyze, and detect anomalies in the audio, and apply machine learning algorithms. They are using machine learning algorithms to analyze sound signals, understand what is "normal" and detect anomalies and potential issues.Based on the noise level and type of noise, the platform can detect if there is a problem with the equipment and will eventually be able to predict when this issue might impact plant operations. Operators, who might be stationed hundreds of miles away, can easily view any alerts in the platform dashboard and take the appropriate action—such as dispatching maintenance or scheduling service activity.

Results

The innovative industrial IoT solution helps hydropower plant operators continuously monitor the health of these massive machines and potentially predict turbine, generator, and other equipment issues much earlier and more cost effectively. Earlier detection can help dramatically reduce equipment downtime and maintenance costs, enabling increased energy production and increased profit margins for the asset owner.

Cloudera is helping transform vast amounts of complex data such as noise levels into clear and actionable insights, thereby enabling our Client bring new, innovative solutions to their customers. Ultimately, the goal is moving from anomaly detection to predictive maintenance and to use the insights gained to provide better equipment for future plants.

cloudera

Case Study 3

Zero Down Time (ZDT) Application that helps monitor and prevent issues for 10,000 connected robots across manufacturing facilities in real-time.

Customer Case Study - 3
Industry: Industrial Robotics
Driving towards Zero Downtime in Industrial Robotics

Customer Overview:

Our Client is one of the world's leading suppliers of robotics equipment and factory automation systems, helping manufacturers maximize their efficiency, reliability, quality, and profitability. In this context, they are providing industrial robotics solutions and services for some of the leading automotive manufacturers.

Challenge:

For auto manufacturers, operational uptime is of paramount importance. Some auto manufacturers estimate that unplanned downtime in a factory can cost them as much as $15,000 - $20,000 per minute and that a single downtime can cost them approx. $2 MN in lost revenues. Given the business impact of unplanned downtimes, our Client needed to ensure that the thousands of their robotic equipments and machinery that were installed in the plant were always up and running and in good operational health. They needed the ability to continuously monitor the performance of the robotics equipment in real-time, while being able to predict and detect issues before they impacted the operations.

Solution:

Our Client has built Zero Down Time (ZDT) Application and robotics monitoring solution, using Cloudera Enterprise Data Hub as the data management and advanced analytics platform. They are using Cloudera Enterprise as the data management engine to gather, store, process, and analyze sensor data files from **10,000** robots across manufacturing facilities in real-time. The ZDT Application analyzes data from coming from robots throughout its factories to monitor the health of these critical machines and detect potential issues that could lead to failures in the production line. Based on the real-time data collected, if a potential failure is detected, ZDT alerts the customer's service center. Parts and support can then be delivered to tackle the issue before any downtime occurs.

Results

Using the ZDT application, and powering it with real time data analytics, has enabled our Client to transform their business and operations, as they edge towards a zero-down time goal for their robotic equipment. This has tremendous potential to impact both operational efficiencies and their customer satisfaction metrics. Apart from lowering down time, using ZDT, our customers is able to use the data generated from their robotics equipment to determine how to optimize the manufacturing environment including things such as reducing energy consumption, extending the equipment life and also improving cycle time and product quality.

cloudera

Case Study 4

Enabling predictive maintenance and driving operational efficiencies across millions of connected cargo handling equipments.

Customer Case Study - 4
Industry: Cargo Handling & Shipping
Using IoT to Transform the Cargo Handling Industry

Customer Overview:

Having delivered half a million loader cranes to customers, and servicing a quarter of the world's 600 million containers, our Client is one of the leading provider of cargo and load handling solutions globally.

Challenge:

In order to improve their efficiencies in cargo-handling operations at scale, our Client wanted to collect, process, and analyze data from **100,000+** connected equipments worldwide to derive insights and build predictive models to serve their end users. They wanted to bring together data from connected machinery in order to provide a comprehensive view of issues related to cargo handling to ensure their customers have continuous, reliable and sustainable performance. More than just managing data from connected machinery, they also needed to add context to sensor data by being able to combine and correlate it with other internal and external data sources. Also, they needed to open up their data to the data scientists to do exploratory analytics and model building as well.

Solution:

Cloudera together with TCS architected a cloud-based IoT solution, built on Cloudera Enterprise Data Hub, that uses advanced analytics and machine learning to drive insights from the streams of data across the millions of cargo handling equipment to enable remote monitoring and predictive maintenance. The end goal is to improve utilization, reduce unplanned equipment downtime and further provide a rich, comprehensive view of issues related to cargo handling to ensure their customers have continuous, reliable and sustainable performance.

The IoT-as-a-Service solution is based on TCS Sensor Data Analytics Framework (SDAF) that will help collect, store, analyze and correlate sensor data streams with data from both internal, external and third-party data sources. The platform ingests data from millions of connected equipment such as — tail lifts in trucks to loader cranes, offshore cranes and loading equipments in some of the biggest cargo container vessels, pulls in weather patterns and forecasts, and contrasts geography — to perform key analysis for remote monitoring, predictive equipment maintenance and anomaly detection all in real-time. Sensor data is also being merged with other enterprise data sources and data from the backend systems for equipment performance analysis.

Additionally, the Client data science team utilizes Cloudera Data Science Workbench, a collaborative hub and integrated development environment capable of running Python, R or Scala with support for Apache Spark to build machine learning solutions.

Results:

Before having a big data analytics platform, multiple systems and geographies were functioning in silos and there was no comprehensive view or insights into the operational performance of machines. All of the tracking was done manually, was inefficient, inflexible and error-prone. Today Cloudera is powering their Internet of Things (IoT) journey to enable predictive maintenance and drive operational efficiencies across millions of connected equipments. They now have a real-time view of all the connected equipments and machinery and use analytics to improve their operations. The data stored is now used as the data lake for building insights and predictive models and the platform will support our Client's strategic goal to connect all equipments across 35 countries by 2020.

cloudera

To learn more about Cloudera for IoT and key use cases, please visit: https://www.cloudera.com/solutions/iot.html

Conclusion

IoT and the power of advanced analytics and machine learning is shifting the paradigm in predictive maintenance today. Utilizing the power of Cloudera Enterprise Data Hub, organizations can now easily ingest, store, process, and analyze unlimited volumes and varieties of sensor data; use powerful processing and analytics tools across data in motion as well as data at rest; and leverage machine learning capabilities across petabytes of data to drive real-time predictive maintenance. With Cloudera Enterprise, organizations are able to benefit from the power of open source technologies while leveraging Cloudera's industry-leading management and data security tools that are critical to IoT production deployments. Today Cloudera is powering some of the most compelling IoT use cases across diverse verticals, and, together with our industry-leading partners, can accelerate the time-to-value for your IoT investments.

About Cloudera

Cloudera delivers the modern platform for machine learning and advanced analytics built on the latest open source technologies. The world's leading organizations trust Cloudera to help solve their most challenging business problems by efficiently capturing, storing, processing, and analyzing vast amounts of data. Learn more at cloudera.com.

cloudera.com

1-888-789-1488 or 1-650-362-0488
Cloudera, Inc. 395 Page Mill Road, Palo Alto, CA 94306, USA