DataTorrent
Powered by Apache Apex

Apache Apex enables GE to transform industrial internet of things (IIoT) data into opportunity

GE is a leading global digital industrial company providing advanced IoT analytics solutions and software-defined machines to millions of customers across an array of industry verticals. With multiple lines of business in more than 150 countries, it is one of the world's largest and most innovative companies.

KEY OUTCOME METRICS

Apache Apex enabled GE to meet the challenge of ingesting and analyzing machine data from thousands of disparate sources and achieve:

Real-time ingestion and analytics, zero data loss

Unified monitoring of all connected devices

Sub-millisecond latency—preserving data's fleeting value

CHALLENGE

Ingesting data from millions of connected machines and sensors in real-time with zero data loss

The industrial internet revolution is transforming every industry, from energy and transportation to healthcare and aviation. Connected devices and sensor-enabled machines are proliferating rapidly. These devices include medical equipment, drill rigs, wind turbines, assembly line components, and aircraft engines, among more—and all generate data in real time. As a result, organizations are inundated with time-series machine data with significant volume, variety, and velocity. This data promises dramatic upside: businesses can use it to reduce unplanned downtime, increase productivity, reduce human error, and improve operational efficiency.

To unlock this potential, GE built a cloud-based analytics platform, Predix, to help organizations leverage data from industrial machines and use it to build industrial internet applications. These applications transform machine data into actionable insights—ultimately to build smart, agile businesses.

But harnessing the value of this massive volume of machine data presents a huge challenge. Short latency requirements add complexity: most of the data generated from the connected machines loses value in seconds or even milliseconds. To power Predix's Time Series and Data Ingestion services, GE needed a fast, reliable turnkey solution with the ability to ingest massive amounts of data with sub-millisecond latency and zero data loss. At the same time, the solution needed to be easily and efficiently scalable in an enterprise environment to meet current as well as future data needs.

Some of the key challenges GE faced were:

- Ingest high-volume, high speed data from thousands of devices and end-customer sensors with ultra-low millisecond latency—to enable predictive maintenance and support
- Prevent data loss to provide timely alerts based on sensor measurements
- Reduce time and expertise required to build industrial internet applications
- Create easy scalability, without incurring additional hardware costs,

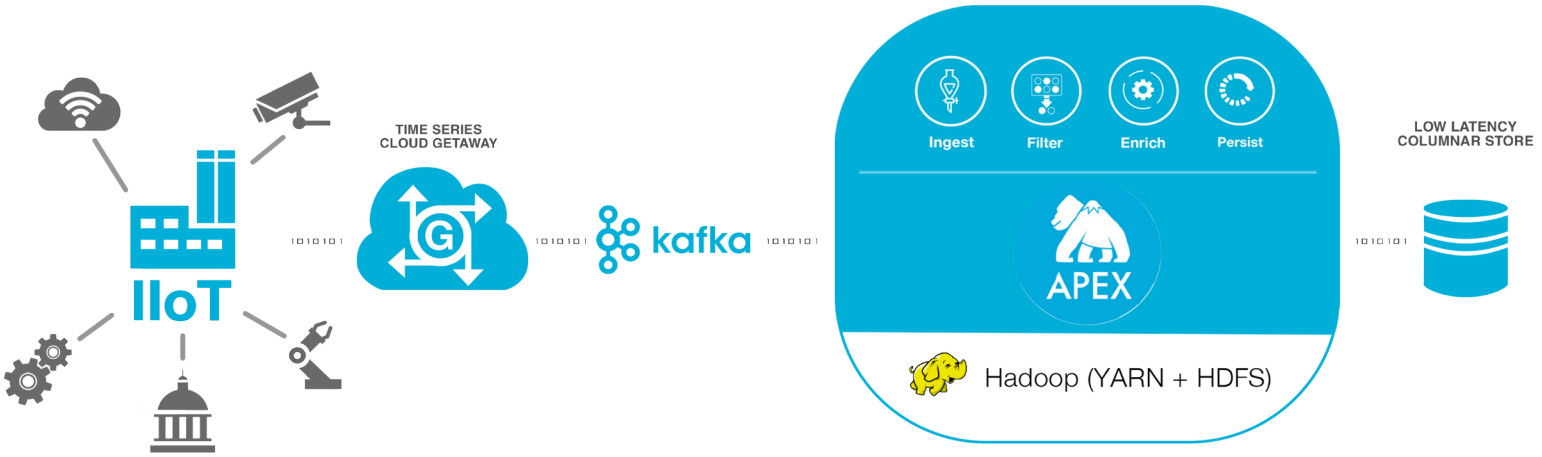
SOLUTION

GE found everything it required—and more—in Apache Apex.

Key features of Apache Apex that benefitted the digital industrial company included:

- Fast time-to-production: with a comprehensive library of pre-built operators, pre-built application templates, management and monitoring UI, and widgets, Apache Apex accelerates time-to-market
- Linear scalability: automatic scaling with the capacity to process billions of events per second with millisecond latency
- Guaranteed data processing: support for exactly-once semantics
- High availability: built-in fault tolerance enables applications to self-heal with no data loss, no state loss, and no human intervention
- State management: highly efficient and distributed state management enables automatic check-pointing with minimal impact on latency—without writing code
- Ease of use: easy integration with current technology stack—databases, message buses, and data sinkss

DATATORRENT
Powered by Apache Apex



OUTCOME

Apache Apex enabled GE to:

- Ingest and analyze high-volume, high-speed data from thousands of real-time data sources, including connected machines and sensors, with zero data loss
- Build industrial internet applications quickly and easily to lower barriers to entry, accelerate time-to-market, and drive innovation faster
- Provide timely sensor alerts
- Scale automatically to accommodate any data processing needs with single-digit millisecond latency and high throughput
- Easily manage and monitor application performance and debug applications
- Power predictive analytics with the right data at the right time to reduce costly maintenance and unplanned downtime, improve asset performance, and lower costs
- Unify monitoring of all connected sensors and devices to minimize disruptions
- Reduce total cost of ownership on development and operations

Apache Apex delivered the data ingestion and analytics capabilities GE needed to power Predix. With Apache Apex, GE is positioned to unlock the vast potential of the high-volume, high-speed data in thousands of data sources, all in real time. And, by harnessing this vast, continually renewing well of data, GE transforms the IIoT into a key advantage.

Imagine what your business could achieve with real-time data-in-motion ingestion and analytics.

Download the Sandbox, Enterprise, Community, or Cloud version of the DataTorrent platform, and explore our difference today. Learn more about our products and solutions at www.datatorrent.com