Real-time data analytics for IoT

Yellowbrick

Analyze IoT data faster and more efficiently at every point in your network

There's no end to the list of industries that are both challenged and transformed by Internet of Things (IoT) devices and the data "exhaust" they produce every second of every day. Utilities, Healthcare, Telecommunications, Automotive, Insurance, and Manufacturing are more instrumented than they've ever been, and that trend is accelerating.

In every industry, large-scale IoT applications can generate petabytes of data in minutes, and 5G connectivity and low-latency transmission will make that volume grow exponentially. There is now a critical need for lightning-fast data analytics solutions that can support initial processing at the edge, real-time streaming data and large volumes of rapidly changing data, and complex interactive analytics in centralized private and/or public cloud environments. Those solutions also need to be sufficiently flexible to meet any need and complement any IT approach without a massive investment in technology or new skill sets.

But current options are far too limited for those tasks. NoSQL options like Cassandra, Redis, and HBase aren't designed for real-time analytics at scale; cloud-only options like Snowflake have poor and unpredictable price/performance and are unsuitable as edge systems for latency reasons; and data lakes have proven to be useful only as low-cost data stores. For most companies, those limits make IoT real-time analytics at scale a pipe dream.

Yellowbrick for modern IoT analytics
Yellowbrick Data Warehouse breaks through the limitations of legacy and cloud-only alternatives. Based on a redesign/modernization of analytics databases from the ground up to take advantage of highly optimized instances, Yellowbrick offers the best price/performance for real-time workloads available at IoT scale. At the same time, it offers the freedom to deploy data warehouses in data centers/private clouds, as SaaS in public clouds, and both (hybrid), along with near real-time replication across all instances (e.g., between edge nodes and private/public clouds).

Key facts

$1 trillion
Total that was spent on IoT worldwide in 2020
IDC

75 billion
Number of IoT devices that will be online globally by 2025
Statista

79 zettabytes
Amount of data that will be produced by IoT devices globally by 2025
IDC

IoT analytics architecture

Devices
Telco, Automotive, Logistics,
Manufacturing, Oil & Gas

- Routers
- Towers
- Phones
- Cars
- Cameras
- Sensors

Streaming
(millions of rows/sec)

Edge analytics
Filtering, aggregation, and scoring –
at the edge

Auto replication

Core analytics (EDW)
Complex workloads –
in centralized cloud

End users

BI & data science tools

Yellowbrick Data Warehouse

That unique feature set helps businesses across every industry overcome the hardest IoT challenges while unlocking new monetization opportunities, like only real-time, interactive access to vast data volumes can.

Fast & easy migrations

Yellowbrick is compliant with industry standards like ANSI SQL for plugging seamlessly into existing environments that include common data integration, BI, and identity management tools. We can set up and run a free Proof of Concept either in your data center or the cloud within a couple weeks, with production migrations of all your workloads usually completing in 30-60 days.

Next steps

Contact us to explore how Yellowbrick Data Warehouse offers the price/performance at scale and deployment flexibility that are critical for real-world IoT analytics.

Learn more at **yellowbrick.com**.

With Yellowbrick, you can:

Stream IoT data into edge nodes at millions of rows/second using Apache Kafka or Spark for filtering, aggregation, and potentially, ML model scoring

Auto-replicate data from the 5G edge to the back office, asynchronously and in near real-time
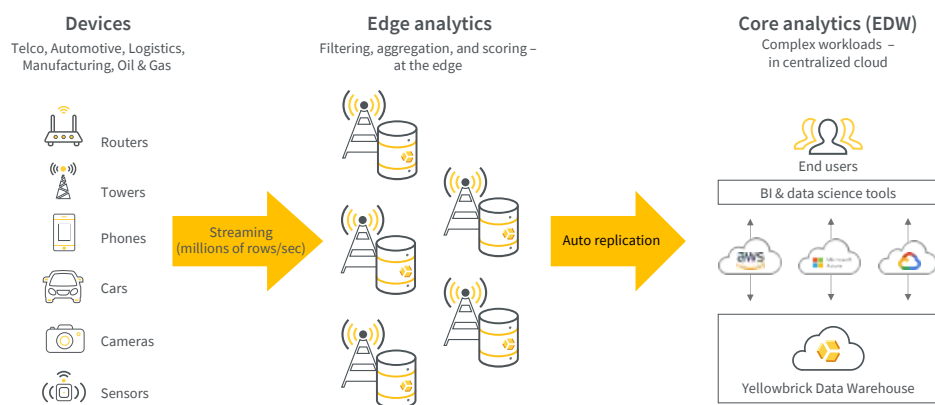
Process that data in data analytics hubs deployed in data centers/private clouds, public clouds, and both/hybrid, with pay-per-use cloud DR available

Use the same cloud console/get the same operator experience across your entire architecture

Let thousands of users analyze billions of data points at the speed of thought using their favorite BI or data science tools

Query all streamed and at rest data immediately, with no need to manually create indexes, partitions, or cubes for peak performance