Analytics for the Internet of Things

Just as big data was becoming the new normal, technology ushered in the Internet of Things with its exponentially greater data volumes. So, we introduced in-database and edge analytics to change the way data is collected and analyzed.

The Internet of Things, commonly called IoT, could be the most significant technology trend of our era. By enabling equipment and sensors with internet and M2M connectivity, the IoT greatly expands by orders of magnitude the volume of data available for organizations to optimize operations, provide innovative products and services, and manage resources like never before.

However, while the IoT process begins with capturing and aggregating high-velocity data, it is the analysis of this data that provides its real value. Analysis enables organizations to make decisions and take actions in real or near-real time. Unfortunately, not all traditional analytic models and techniques are easily adapted to high-dimensional, streaming data. The very nature and delivery of IoT data will require new analytic workflows and processing architecture to ensure a

modern data system remains practical and useful. So, unless a company does something different, the ease of collecting huge quantities of IoT data will quickly outstrip its ability to analyze it.

NDAA is the IoT differentiator.

Statistica advanced analytics platform addresses this challenge with its Native Distributed Analytics Architecture (NDAA), a unique in-database methodology designed specifically to execute platform-specific analytic operations on aggregates computed in a database, edge gateway or device/sensor. It then returns results—often at the size of an email attachment—to the desktop, whose role would be limited to building models and displaying results, rather than crunching burdensome data streams.

With NDAA, any one can conduct in-database processing with no knowledge of SQL or databases

"Statistica deploys Analytic Atoms onto any edge device or IoT gateway anywhere in the world to run analytic workflows right where data is created. Only the most meaningful, pre-processed, aggregate information is delivered upstream."

By moving the math to the data, Statistica's in-database and edge analytics will:

- **further empower Citizen Data Scientists,** line-of-business users adept at meeting their own analytics needs independently of statisticians.

- **provide reliability** in the case of slow/intermittent data connections between data and end consumers of results.

- **improve security** by reducing the amount of raw data (v. aggregate data) in network traffic.

- **enhance performance and operations management**, with a relatively small setup time required to move to a new database platform.

- **improve processing time and integrity** by leaving data in the native systems where it resides.

- **reduce duplication of infrastructure** that would otherwise be required to store, manage, and protect critical data sets.

NDAA and Edge Scoring
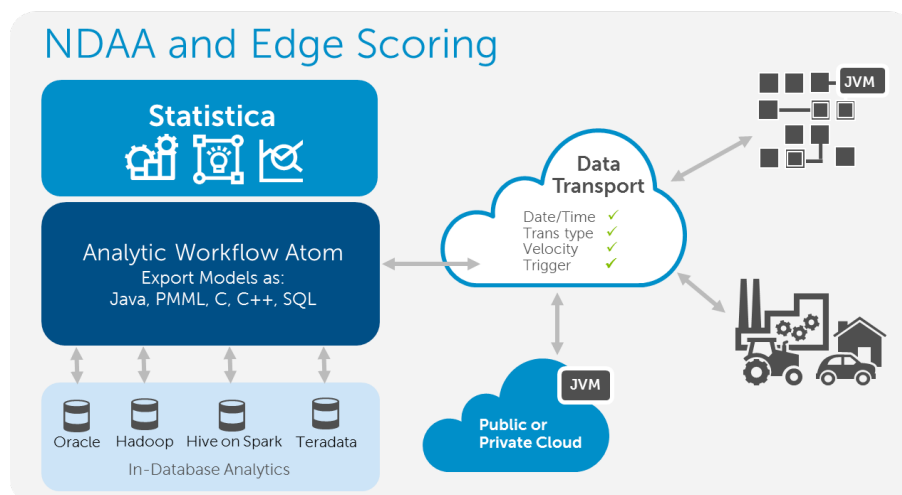
In a typical NDAA scenario, a user builds a data prep or analytic workflow, such as with a Random Forest. He can export it in the language (e.g., Java, C, C++, PMML, SQL) that is most appropriate to the target platform (e.g., Hadoop, Teradata, Netezza, Exadata) on which the data resides. Once there, the analytic workflow executes platform-specific operations.

required, because all performance-intensive computations are carried out on the database side. And Statistica's NDAA solution is database-agnostic, supporting Apache Hive (on Spark), MS SQL Server, Oracle, Teradata and more.

Push analytics to the data

NDAA models may also run as "edge analytics" close to the physical locations where IoT sensors and controllers are located: on the leading edge of the data system. Such edge analytics dispense with the need to collect huge amounts of data from the source. Instead, the math goes to the data no matter where it resides.

Essentially, with Statistica you can deploy Analytic Atoms onto any

edge device (even a non-computing device) or IoT gateway—anything with an IP address anywhere in the world—enabling you to run analytic workflows at the edge of the network where data is created. Not only does this functionality enable local and/or automatic decision-making and action-taking, but it eliminates the need to stream all IoT data to centralized data centers. Only the most meaningful, pre-processed, aggregate information needs to be delivered upstream to another gateway, to the datacenter or to the cloud for further core analytics.

This means your entire network becomes "smart" at any scale, making decisions in real time or to the millisecond. Edge analytics could easily be the key to making truly smart cities and super-smart factories a reality.

Benefits of NDAA and edge analytics

- **Analyze and respond** to real-time data in critical situations where network latency or periods of lost connectivity pose risks to safety, productivity, or customer experience.

- **Ensure decisions and actions** are made quickly based on the most relevant data.

- **Minimizes the consumption** of network bandwidth.
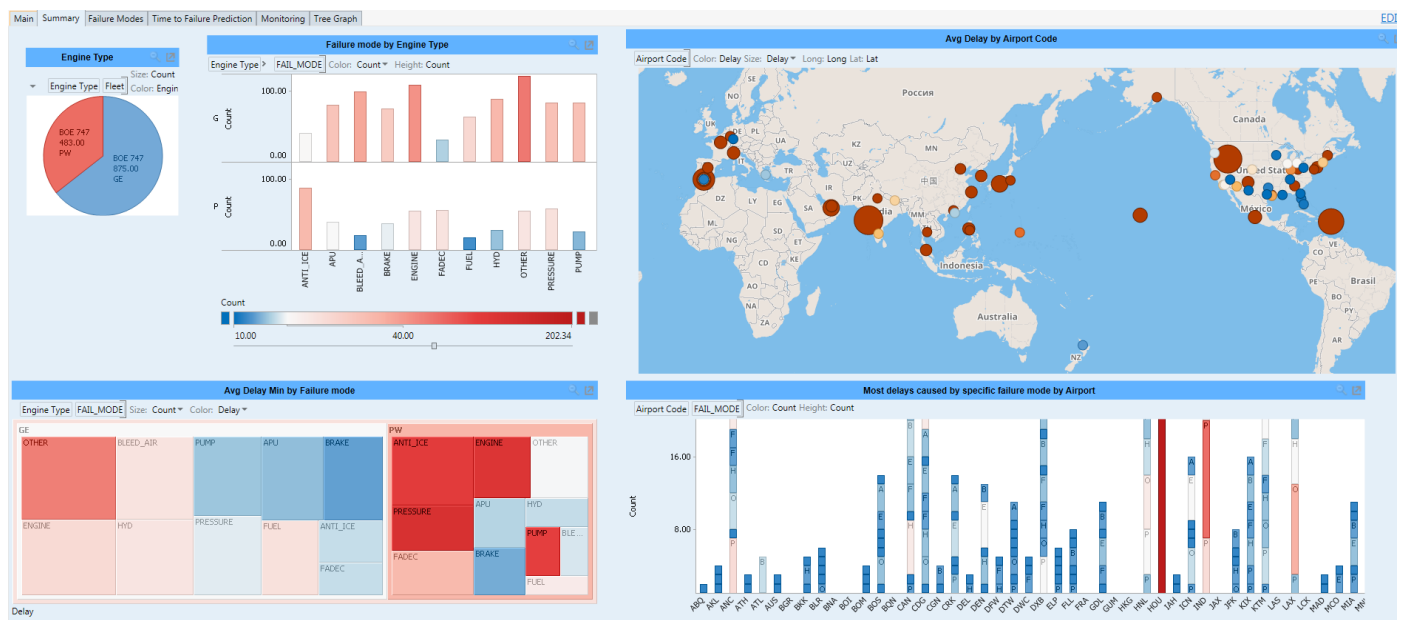
Why Statistica?

Deployed on virtually any platform, Statistica NDAA and edge analytics can make the difference between success and mediocrity in a diverse IoT landscape:

- **Vertical- and domain-enabled** — Deploy industry specific models and algorithms to create faster value

- **Rich partner ecosystem** — Industry-leading and vertical market partners

- **Digital Strategy Consulting** — Engage in collaborative system design, development and delivery.

Powerful results can be achieved when Statistica's NDAA, edge and core analytics are combined. Integrating these analytics in a distributed system enables algorithms created at the core to run locally at the edge and helps deliver the full power of IoT. At Statistica, we combine our advanced analytics capabilities with technology and domain expertise from our partners to deliver the right analytics for your unique business.

About Dell Statistica

Since 1984, Dell Statistica has been delivering scalable, affordable and simple-to-use analytics solutions that enhance decision-making and mitigate risk while enabling its customers to embed analytics everywhere, empower more people, and innovate faster. www.dell.com/statistica

Statistica's dynamic visualization and dashboard tools enable citizen data scientists to research and explain edge data effectively.