

Analyzing the Internet of Things

Just as big data was becoming the new normal, technology ushered in the next level of challenges.

BY MOVING ANALYTICS TO THE DATA, TIBCO STATISTICA™:

- Further empowers citizen data scientists, line-of-business users adept at meeting analytics needs independently of statisticians.
- Provides reliability in the case of slow/intermittent data connections between data and consumers of results.
- Improves security by reducing the amount of raw data (aggregate data) in network traffic.
- Enhances performance and operations management with a relatively small setup time required to move to a new database platform.
- Improves processing time and integrity by leaving data in the native systems where it resides.
- Reduces needs for infrastructure that would normally be required to store, manage, and protect critical data sets.
- Interactive dashboards and visualizations.

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

TIBCO 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, anyone can conduct in-database processing with no knowledge of SQL or databases required, because all performance-intensive computations are carried out on the database side. And the Statistica NDAA solution is database-agnostic, supporting Apache Hive (on Spark), MS SQL Server, Oracle, Teradata, and more.

“Statistica deploys analytic workflows onto any edge device or IoT gateway anywhere in the world, right where data is created.”

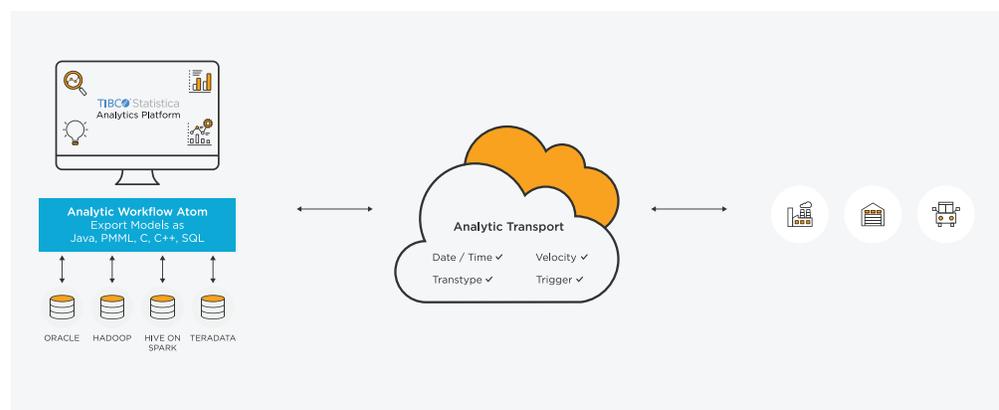


Figure 1: In a typical NDAA scenario, a user builds a data prep or analytic workflow, such as with a Random Forest. They can export it in the language (for example, Java, C, C++, PMML, SQL) that is most appropriate to the target platform (Hadoop, Teradata, Netezza, Exadata, etc.) on which the data resides. Once there, the analytic workflow executes platform-specific operations.

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. Edge analytics dispense with the need to transfer 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 workflows 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 them 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, 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 up-stream to another gateway, the datacenter, or 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

- 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.
- Minimize 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: Industry-specific models and algorithms create faster value
- Rich partner ecosystem: Industry-leading and vertical market partners work together
- Digital strategy consulting: Collaborative system design, development, and delivery

Smart things can happen when the Statistica NDAA, edge, and core analytics unite in a distributed system, where algorithms created at the core run locally at the edge to deliver the IoT’s full power. At TIBCO, we combine our advanced analytics capabilities with domain expertise from our technology partners to deliver the right analytics for your business.



Figure 2: Statistica dynamic visualization and dashboard tools enable citizen data scientists to research and explain edge data effectively.