NSTREAM

6 Ways Big Data Analytics Fails to Work — and How Streaming Data Applications Fix Them

The rise of big data analytics has meant that organizations could inform their decision-making by analyzing massive volumes of disparate data. It wasn't long, however, before organizations began bumping up against the limitations of this approach, especially as they tried to apply it to streaming data. Here are six examples of such limitations and how modern streaming data applications overcome them with an approach that turns traditional big data analytics on its head.

Problem 1

Big data analytics was built for data at rest.

Big data analytics was a breakthrough in handling large data sets, but only if that data was readily available in a database, data lake, or warehouse. It could then apply the massive computational power of the cloud to batch-process this data to drive insights. With reports generated every few hours, daily, or even weekly, the data driving those insights was already out of step with the real-time state of the business. Still, organizations accepted this approach as their best option, which is certainly better than having no insights.

The Fix

Streaming data applications are built for data in motion.

Modern streaming data applications are architected to process data at rest and in motion/streaming data. They apply business logic instantaneously to static data (for context) and incoming streaming data to reflect the state of a real-world object, such as a customer, asset, inventory, transaction, or IoT device. For example, a streaming data application could provide a real-time view of the state of each traffic light in a whole city or the granular location of various parts on a manufacturing factory floor.

Problem 2

Big data analytics can’t scale.

Applying a traditional big data analytics model to streaming data “breaks” the data pipeline. Incoming data must be stored before it’s processed, requiring multiple roundtrips to and from a database. The only way to accelerate this process is to increase compute power and storage, which greatly increases cost — while still failing to eliminate latency. The more you scale vertically (with higher volumes and velocity of data) or horizontally (with multiple data streams), the worse the cost and latency issues become.

The Fix

Streaming data applications scale while maintaining network-level latency.

Because streaming data applications process and push data in motion, they do not create the traffic to and from databases. What’s more, they can process multiple streams concurrently, eliminating an additional cause of latency. The more data sources — both streaming and static — that the application can process, the richer the picture it can paint of the state of the business and every real-world object within it.

The more data sources — both streaming and static — that the application can process, the richer the picture it can paint of the state of the business.

Problem 3

Big data analytics “looks down” at the data.

Because traditional models of big data analytics work on data at rest, everything it “sees” is technically in the past. Yes, it can extrapolate this historical data into broad predictions of the future, but it can’t base those predictions on what’s happening in real-time to the business. In other words, it’s not so much predictive as it is reactive.

The Fix

Streaming data applications “look up” at the business.

Streaming data applications calculate and display the state of any real-world object the instant it changes. This is possible as the data is pushed throughout the stack, right from the real-world object to the UI at near network-level latency. When you multiply that by the millions or billions of real-world objects in the enterprise, you start to get a picture of the state of the business as a whole in real-time. To take the previous example, imagine a map of the U.S. showing the state of every cell tower and every connected device in a network. That map, which is a real-time UI, is alive with state changes — for instance, devices pinging from one tower to the next or calls starting, ending, or dropping.

Streaming data applications calculate and display the state of any real-world object the instant it changes.

Problem 4

Big data analytics is bound by the "tyranny of averages."

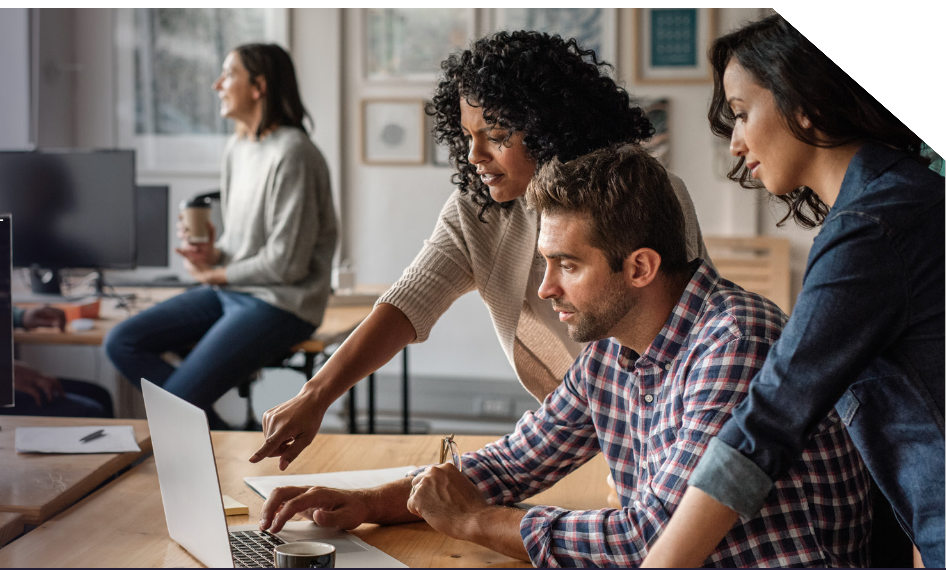
Big data analytics reveal actions to take based on past norms. Take, for example, an apartment leasing company setting rental rates for the year. When the company conducts analyses of previous years' rates to decide this year's rates, the business is acting based on what the data tells them happens "on past averages" and what has "happened" in the past, instead of also directing the business to what is actually happening currently in the rental market and what that data is communicating.

The Fix

Streaming data applications provide a clearer and more comprehensive picture.

Unlike traditional analytics, streaming data applications analyze static historical data and dynamic real-time data from various sources. In our example, they factor in additional data — real-time occupancy rates in the area, real-time prices of comparable apartments, time of year, length of lease, number of simultaneous applications at this moment, etc. — to dynamically set rates for each applicant based on real-time market conditions. Rather than rely on averages, streaming data applications help identify outliers that can help individualize and personalize decisions and actions.

Rather than rely on averages, streaming data applications help identify outliers that can help individualize and personalize decisions and actions.



Problem 5

Traditional analytics doesn't support agility.

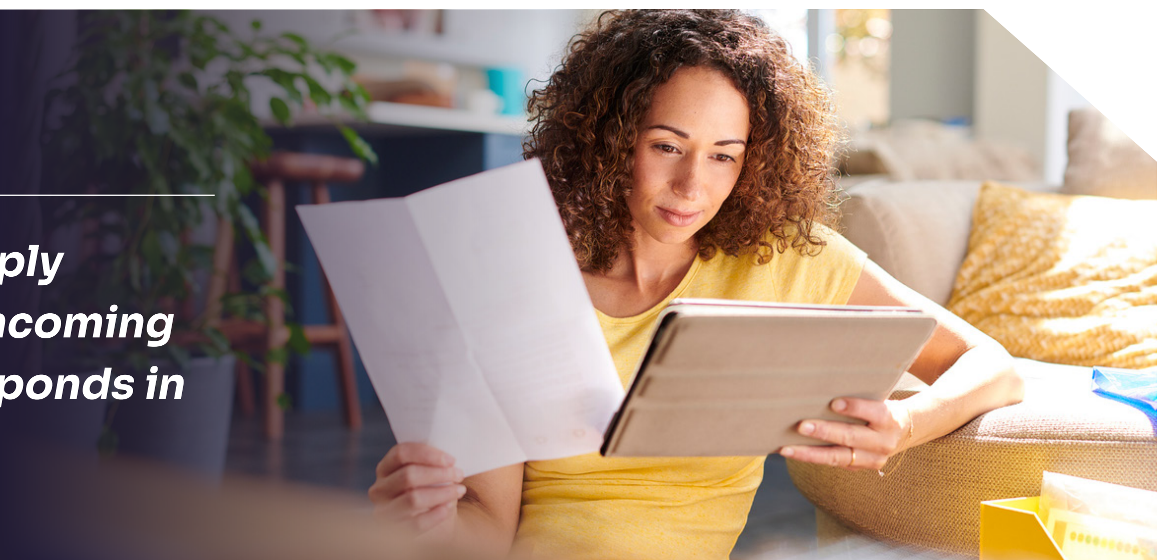
Big data analytics effectively steers the enterprise along its course based on historical data. Still, it cannot reasonably assist with the micro-adjustments — the small tacks on an hour-by-hour, minute-by-minute, or even second-by-second basis. Nor can it drive automation for exactly that reason. Given the lag between available data and current states, connecting analytics results directly to automated response via orchestration can lead to unpredictable results.

The Fix

Streaming data applications drive action in the moment.

Streaming data applications offer the real-time insights needed for a clear macro view and the fine steering in response to transient shifts in state of even a single entity, providing next-level optimization. This also means that streaming data applications can apply business logic instantaneously to incoming data, powering automation that responds in real-time to a fluid situation. For example, an ecommerce website can serve up highly personalized offers, like an in-store coupon, based not only on a customer's purchase history, but also on up-to-the-moment information such as location, sentiment, behavior, product availability, and more.

Streaming data applications can apply business logic instantaneously to incoming data, powering automation that responds in real-time to a fluid situation.



Problem 6

Traditional streaming data analytics is expensive.

The traditional model of streaming data analytics requires high compute performance and generates large volumes of traffic and storage, all of which lead to dramatic increases in infrastructure costs. But that's just the beginning. These models also require expertise in multiple applications and connections among them, which taxes resources and increases time to value. In addition, the latency inherent in these systems — which compounds at every layer with each additional source — limits the value of their output.

The Fix

Streaming data applications are cost-effective.

Streaming data applications place far less burden on infrastructure while still delivering better performance, reducing cost by several factors. They can typically replace or augment multiple applications, reducing resource needs and speeding up time to value. And because their output is truly real-time and fully contextual, they can optimize outcomes via real-time visualization and automation.

Streaming data applications place far less burden on infrastructure while still delivering better performance, reducing cost by several factors.

Build streaming data applications with an end-to-end streaming application approach

Traditional big data analytics is great for understanding what needs to be done, but it falls short when it comes to actually doing it. An end-to-end streaming application approach is required to actualize the insights that come from big data analytics.

Nstream provides an application development platform that leverages stateful objects and streaming APIs, so you can build streaming data applications that connect real-time data with real-world actions.

To learn more about Nstream, visit Nstream.io/why-nstream.