Streaming Data from the Network Edge

How In-Memory Technologies and Machine Learning Catalyze Innovation

RTInsights.com

hazelcast

Companies are increasingly generating huge volumes of data at the network edge. Massive volumes of data are flowing from smart meters, Internet of Things (IoT) sensors, autonomous vehicles, health monitors, and industrial automation devices, to name just a few sources.

Many teams want to use data streaming from the edge to better understand their business. Imagine an oil exploration company that aggregates and filters sensor data at the edge, extracting the most critical time-sensitive data points for subsequent processing at a centralized data center. Or a manufacturer that uses a machine learning application to diagnose a potentially failing component or spot a quality issue in a remote location.

However, it's not practical to transmit these growing data volumes from the edge into a centralized data center for processing or analysis. Some companies manually collect data from the edge onto storage devices and physically transport it to data centers. But this process is time-consuming and cannot yield real-time insights.

Nor is it easy to process data near the location where it is created. Often these edge environments lack space for computer hardware, or using the available space comes with a big opportunity cost. What's more, processing power at the edge is typically insufficient for huge data volumes, since small-footprint hardware designed for edge computing tends to be smaller than hardware running in data centers. And poor latency caused by network bottlenecks prevents transactions from being processed in a timely manner.

When companies cannot effectively process streaming data, many critical events go unnoticed or are spotted too late for timely action. Only when event-streaming IoT data is processed using modern technologies can companies gain real value from their data assets. By deploying a technology foundation with the appropriate speed, scalability, stability, and security, you can process data from the edge—and gain valuable insights that can catalyze innovation (see sidebar).

INNOVATIVE APPLICATIONS FOR STREAMING DATA

Predictive maintenance

Oil and gas companies using IoT devices on remote drilling equipment need real-time data to understand operational issues such as drill head stability. The data helps analysts predict the need for maintenance before an equipment break creates a need for costly, time-consuming repairs or downtime. At one company, this insight reduced the time needed to achieve true vertical depth of the well by 10 percent, creating savings of millions of dollars per week per rig.

Healthcare insight

Medical professionals use telematics to instantly see and treat remote patients. One hospital uses the technology to conduct robotically assisted surgery on patients, where the surgeon instantly receives data and insight from cameras, robotic devices, and patient monitors. This rapid feedback makes it easy to quickly adjust to any change in the patient's condition or address surgical challenges as they arise.

Customer service

A media company utilizes data from consumers' set-top boxes to better understand viewing habits, purchase history, customer service interactions, and billing issues. Combining this customer experience data with an AI-powered application enables chatbot conversations that give service reps the information necessary to optimize every customer contact or service call. In addition, the company can quickly address customer complaints—such as those resulting from system failure—because the AI-enabled, human-like bots can handle high call volumes while personalizing customer communications.

Real-Time Event Streaming Accelerates Business Insight

A modern, real-time streaming solution that supports edge computing and incorporates machine learning technology offers valuable benefits to your organization.

From a business perspective, the technology delivers powerful analysis that can help you detect potential and actual failures in real time, allowing instant responses. For example, a mechanical issue with an autonomous vehicle could be identified and addressed before the vehicle loses power or control.

Streaming solutions can also streamline more complex analysis, such as linear or multiple regression testing of goods being manufactured. For instance, the technology can help a manufacturer better understand the likelihood of failure due to multiple environmental conditions, machinery problems, or other manufacturing flaws—before a costly production run is complete.

Real-time processing speeds up responsiveness and enables faster time to results. The technology ingests, categorizes, and processes vast amounts of data with ultra-low latency, supporting continuous intelligence practices. It can also process time-sensitive data from numerous sources.

Performance and optimization at such a fine level enables businesses to shrink the time spent on a project, which can result in potential savings of millions of dollars. Faster insights also help enterprises reduce the amount of time needed to perform tasks or make adjustments, increasing efficiency and improving the ability to make faster, smarter decisions.

Using artificial intelligence and machine learning with event streaming also helps enterprises enhance their insight about customers and operations. One energy company uses the technology to generate huge volumes of high-frequency data that can derive insights from machine learning models. Another firm uses machine learning to power a chatbot, which delivers analysis that improves customer knowledge.

A modern event-streaming solution also simplifies technology deployment and operation. A single, lightweight system can address even the most challenging architectural requirements. By choosing an integration-friendly solution, you can avoid adding further complexity or massive administrative overhead to your environment.

Choose a Technology Built for Edge Computing

To gain maximum business value from IoT data, companies should look for a real-time event-streaming solution that enables edge computing. Rather than solely relying on a centralized data center, companies should explore ways to push the compute layer closer to the data sources. This lets them reduce the bottleneck of slow/limited networks, and derive actionable insights in real time. An ideal IoT processing platform includes the following capabilities and features.

Processing speeds at extreme scale. Considering the volumes of data created at the edge, an operational, in-memory computing platform that manages data using in-memory storage is best to deliver microsecond-level speeds at scale. It also performs parallel execution for fast application speed and supports ultra-low latency. High-speed processing can help companies take action faster than their competitors and accomplish more in the same amount of time—catalyzing new opportunities, including those yet to be imagined.

Run-anywhere, portable edge computing. To handle edge processing applications, the solution should use specialized IoT platforms that are compact and durable. These platforms, which require less RAM and lower CPU power, are perfect for running lightweight software applications.

Elastic, seamless scalability. A solution that can easily scale up and down without interrupting jobs helps you keep pace with changing workloads. You should be able to add or remove hardware to allocate the right amount of resources to the workloads without disrupting operations.

Rapid job submission. The solution should support the ability to submit new jobs quickly and easily. You should also be able to replace long-running existing jobs with enhanced versions. This capability is especially important for deploying machine learning models, since most users repeatedly refine their models and need to get them into production quickly.

Stability. Reliable solutions are always available for processing. IoT solutions require data replication to provide a robust yet performant means of fault tolerance. In the case of a temporary outage or power loss to the processing cluster, a hot restart feature helps you get the node up and running again quickly.

Enterprise-grade security. As event streaming introduces many more access points to the processing network, the need for security rises. Authentication and role-based access controls protect data from unauthorized viewers, while encrypted data transmission helps ensure data privacy.

Intuitive manageability. By providing tools to monitor your system, a management console can help you ensure service-level agreements (SLAs) are met. It also provides the tools required to troubleshoot problems when they arise, reducing the cost of downtime.

Modern technology components. Look for a solution that is open source, based on standards, and offers a cloud-native design. It should include:

- An operational, in-memory computing platform that manages data using in-memory storage and performs parallel execution for fast application speed and scale
- Scalable storage for caching data from third-party systems
- An ultra-fast, application-embeddable stream processing engine that supports low-latency batch and stream processing

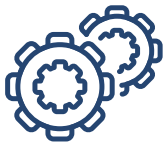
USE CASE:

Real-Time Event Stream Processing for Safer, More Efficient Operations

A leading oil and gas system integrator wanted to improve customer operations using data-driven insight.

CHALLENGE:

The company wanted an event-based technology solution that could analyze and display high-frequency data, helping drill operators process, store, and analyze data from the edge. It also needed a new way to detect low-latency issues earlier. Another priority was the ability to automate the remediation process, so the company could avoid lost production time and optimize well productivity.



SOLUTION:

The firm deployed a streaming solution with an in-memory distributed computing platform that acts as a processing backbone. It supports the application monitoring of well sensors with varying formats and frequency. The solution also adjusts rig settings in real time.

RESULTS:

Now the company can record or sample high-frequency data or events that occur, applying proprietary algorithms to make fine-tuned adjustments to the drilling process. For example, it can make real-time adjustments to the RPM of the drilling string and bit, which can prevent equipment failure and costly delays to the drilling process.

The integrator can gather, monitor, and analyze rig data, tweaking drilling parameters in near-real time. The company's clients can manage physical resources better through high-frequency feedback on a per-well basis. They are often able to reduce the drilling time by as much as 20 percent, from a typical 15 days to 12 days—a benefit that saves millions of dollars per week per rig.

With the new solution, the integrator also can create streaming key performance indicators. These measures help clients reduce the amount of time to drill a well, creating millions of dollars in savings. The technology also provides clients with dashboard capabilities that enable real-time data visualization and interpretation.

Instant Insight Spurs Action and Innovation

As the use of event streaming data expands, the right solution represents a critical global business opportunity for insight and innovation. Companies should look for the following core enablers:

- **Speed** at a rate that can only be delivered by in-memory systems
- **Scalability** using cloud-based solutions that allow businesses to scale up and down instantly in response to changing market conditions
- **Stability** with a distributed architecture that protects against downtime and related business impacts
- **Security** so companies can protect sensitive data from unauthorized users

The Hazelcast In-Memory Computing Platform

Hazelcast delivers the System of Now™, an industry-leading in-memory computing platform that provides Global 2000 enterprises with ultra-high performance for time-sensitive, cloud-native applications.

The Hazelcast In-Memory Computing Platform comprises **Hazelcast IMDG**, the most widely deployed in-memory data grid, and **Hazelcast Jet**, the industry's most advanced in-memory stream processing solution. This technology is uniquely designed to allow you to gain computing insights faster, enable actions within shorter durations, and engage new data at the speed with which it is arriving. In addition, a distributed caching architecture allows you to scale up to hundreds of terabytes and scale out for maximum efficiency when dealing with remote data or edge processing.

Built for ultra-fast processing at extreme scale, Hazelcast's cloud-native in-memory data grid and event stream processing technologies are trusted by leading companies such as JPMorgan Chase, Charter Communications, Ellie Mae, UBS, and National Australia Bank to accelerate business-critical applications. The world's largest e-commerce sites rely on the Hazelcast Platform for sub-millisecond response times to support massive volume spikes associated with Black Friday, Cyber Monday, or Singles' Day.

Intel® Optane™ DC Persistent Memory

Since many of the event-streaming requirements around performance are dependent on in-memory processing, the one big hurdle that emerges is the cost of random-access memory (RAM). In many cases, the investment in more RAM-heavy hardware servers is justifiable, and as RAM prices continue to decrease, the use of in-memory processing becomes more accessible.

Recent innovations make the adoption of in-memory processing even more practical. The Intel Optane DC Persistent Memory technology offers two ways in which in-memory processing can be more cost-effective. The first way is in volatile memory mode, in which Optane chips act as an alternative to RAM, and run at nearly the same speed but at a much lower cost and much higher capacities. This lets businesses more easily justify in-memory technologies and thus take advantage of the performance benefits that in-memory processing offers.

The second way in which Optane supports in-memory technologies is in the persistent mode. In this mode, Optane can be used as a faster alternative to solid state drives (SSDs). For example, Hazelcast provides a hot restart capability in which in-memory data is persisted in non-volatile memory so that if a node goes down temporarily, it can be restored quickly by reading data from the hot restart store. If the hot restart data is stored in Optane in persistence mode, recovery of that node can be up to 3.5x faster than using SSDs.

When generating and processing data at the edge, the speed delivered by a modern event-streaming solution helps you innovate and stay ahead of the competition. To learn how in-memory processing can optimize streaming data, visit https://hazelcast.com/.

RTInsights.com
Accelerate Your Business with Real-Time Insights

RTInsights is an independent, expert-driven web resource for senior business and IT enterprise professionals in vertical industries. We help our readers understand how they can transform their businesses to higher-value outcomes and new business models with AI, real-time analytics, and IoT. We provide clarity and direction amid the often confusing array of approaches and vendor solutions. We provide our partners with a unique combination of services and deep domain expertise to improve their product marketing, lead generation, and thought leadership activity.

hazelcast

Hazelcast delivers the System of Now™, an in-memory computing platform that empowers Global 2000 enterprises to deliver innovative, low-latency, data-centric applications. Built for ultra-fast processing at extreme scale, Hazelcast's cloud-native in-memory data grid and event stream processing technologies are trusted by leading companies such as J.P. Morgan Chase, Charter Communications, Ellie Mae, UBS and National Australia Bank to accelerate business-critical applications. Hazelcast is headquartered in San Mateo, CA, with offices across the globe. To learn more about Hazelcast, visit https://hazelcast.com/.