Azure Stream Analytics

Azure Stream Analytics is a real-time analytics and complex event-processing engine that is designed to analyze and process high volumes of fast streaming data from multiple sources simultaneously. Patterns and relationships can be identified in information extracted from a number of input sources including devices, sensors, clickstreams, social media feeds, and applications. These patterns can be used to trigger actions and initiate workflows such as creating alerts, feeding information to a reporting tool, or storing transformed data for later use. Also, Stream Analytics is available on Azure IoT Edge runtime, enabling to process data on IoT devices.

The following scenarios are examples of when you can use Azure Stream Analytics:

Analyze real-time telemetry streams from IoT devices

Web logs/clickstream analytics

Remote monitoring and predictive maintenance of high value assets

Real-time analytics on Point of Sale data for inventory control and anomaly detection

How does Stream Analytics work?

An Azure Stream Analytics job consists of an input, query, and an output. Stream Analytics ingests data from Azure Event Hubs (including Azure Event Hubs from Apache Kafka), Azure IoT Hub, or Azure Blob Storage. The query, which is based on SQL query language, can be used to easily filter, sort, aggregate, and join streaming data over a period of time. You can also extend this SQL language with JavaScript and C# user-defined functions (UDFs). You can easily adjust the event ordering options and duration of time windows when performing aggregation operations through simple language constructs and/or configurations.

Each job has one or several outputs for the transformed data, and you can control what happens in response to the information you've analyzed. For example, you can:

Send data to services such as Azure Functions, Service Bus Topics or Queues to trigger communications or custom workflows downstream.

Send data to a Power BI dashboard for real-time dashboarding.

Store data in other Azure storage services (for example, Azure Data Lake, Azure Synapse Analytics, etc.) to train a machine learning model based on historical data or perform batch analytics.

Key capabilities and benefits

Azure Stream Analytics is designed to be easy to use, flexible, reliable, and scalable to any job size. It is available across multiple Azure regions, and runs on IoT Edge or Azure Stack.

Programmer productivity

Azure Stream Analytics uses a SQL query language that has been augmented with powerful temporal constraints to analyze data in motion. You can also create jobs by using developer tools like Azure PowerShell, Azure CLI, Stream Analytics Visual Studio tools, the Stream Analytics Visual Studio Code extension, or Azure Resource Manager templates. Using developer tools allows you to develop transformation queries offline and use the CI/CD pipeline to submit jobs to Azure.

The Stream Analytics query language allows to perform CEP (Complex Event Processing) by offering a wide array of functions for analyzing streaming data. This query language supports simple data manipulation, aggregation and analytics functions, geospatial functions, pattern matching and anomaly detection. You can edit queries in the portal or using our development tools, and test them using sample data that is extracted from a live stream.

You can extend the capabilities of the query language by defining and invoking additional functions.

Fully managed

Azure Stream Analytics is a fully managed (PaaS) offering on Azure. You don't have to provision any hardware or infrastructure, update OS or software. Azure Stream Analytics fully manages your job, so you can focus on your business logic and not on the infrastructure.

Run in the cloud or on the intelligent edge

Azure Stream Analytics can run in the cloud, for large-scale analytics, or run on IoT Edge or Azure Stack for ultra-low latency analytics. Azure Stream Analytics uses the same tools and query language on both cloud and the edge, enabling developers to build truly hybrid architectures for stream processing.

Reliability

Azure Stream Analytics guarantees exactly once event processing and at-least-once delivery of events, so events are never lost. Exactly once processing is guaranteed with selected output as described in Event Delivery Guarantees.

Azure Stream Analytics has built-in recovery capabilities in case the delivery of an event fails. Stream Analytics also provides built-in checkpoints to maintain the state of your job and provides repeatable results.

As a managed service, Stream Analytics guarantees event processing with a 99.9% availability at a minute level of granularity.

Security

In terms of security, Azure Stream Analytics encrypts all incoming and outgoing communications and supports TLS 1.2. Built-in checkpoints are also encrypted. Stream Analytics doesn't store the incoming data since all processing is done in-memory. Stream Analytics also supports Azure Virtual Networks (VNET) when running a job in a Stream Analytics Cluster.