Streaming Analytics & IoT

Peter Pugh-Jones
Streaming Analytics
insights on streaming data
“Streaming analytics typically means making analytically informed decisions in milliseconds, while examining many thousands of events per second, generated from many millions of devices which can also be enriched by many other disparate sources of data”

Jerry Baulier, Senior Director SAS R&D, Event Stream Processing

Because of this, **Streaming Analytics** is definitely a **Big Data** topic, but not necessarily restricted to being only a **Real-Time** topic.
Big data processing with SAS

I have a lot of sources of data that I want to grab from at high speed, in very large quantities

**Streaming (ESP)**

I want to keep all the data and access it really fast for as long as possible

**Storage (HADOOP & LASR)**

I want to perform predictive analytics on the data and generate models to use from it


I want to deploy and enable predictive models I have built into the streams, to help generate value as early as possible

**Streaming Analytics (ESP)**
WHAT IS THE INTERNET OF THINGS?
WHERE DOES IOT FIT?

• Big Data
  • -> Internet of Things
WHAT HAVE THESE ITEMS GOT IN COMMON?
An example of an IoT ecosystem might be a house, car or an airplane like a ……

• **Dreamliner**

• **SENSORS**
  • Many of the sensors in the Dreamliner are connected wirelessly to a central data processor. For instance, an "active gust alleviation" system uses sensors to measure turbulence at the nose, then instantly adjusts wing flaps to counter it.

• **DIAGNOSTICS**
  • A self-monitoring diagnostic system sends real-time data to technicians on the ground via a wireless broadband link. The system can predict mechanical problems, meaning less time in the hangar and fewer delays.

A good read on IoT ecosystems
HOW DOES SAS FIT?

Because we have been collecting and analysing data from devices and sensors since we started in 1976 SAS has the analytics to help with IoT already.

The differences are, that we can now do the analytics:-

- **Continuously** (as the data is in flight, or in the stream)
- **At the base level of granularity** (allowing the predictions and decisions we make to be more accurate)
- **At high speed** (as it’s streaming or real-time)
- **With machine generated data** (cleaner from the start, as it is the source)
- **From more and more devices as they become internet enabled**
WHAT IS THE IOT TO SAS?

*Streaming Analytics* is a great way of enabling what SAS already does so well, within the Internet of Things.
WHAT IS EVENT STREAM PROCESSING?
ENGINEERED FOR FAST ADAPTIVE ACTION

SAS® Event Stream Processing

Conduct low latency assessment of high-volume, high-velocity data – triggering real-time actions/alerts

Manipulate live data streams to detect, filter, aggregate & correlate signals from the noise

Apply analytics and rules to pinpoint event relevance and urgency with continuous pattern detection

Sources

Consumers
WHAT IS AN EVENT?

AN EVENT is an occurrence happening at a determinable time and place, with or without the participation of human agents, that can be recorded as a collection of fields containing DATA that we want to use to help contribute to:

• Making more informed decisions
• More accurate predictions and forecasts
• Capturing new patterns and behaviors we may have missed in the past

WORKING WITH EVENTS allows us to apply analytics while the event is, ‘IN MOTION’, before it is stored, (or lost), and may provide analytical insights you may never have had access to before and much sooner than before, generating new value from your data!
WHAT IS EVENT STREAM PROCESSING?

Event Stream Processing (ESP) is a subcategory of Complex Event Processing (CEP), focused on working with ‘Events In Motion’ called Event Streams*.

CEP encompasses methods, techniques, and tools for processing events while they occur, i.e., in a continuous and timely fashion and derives valuable, higher-level knowledge from lower-level events; this knowledge takes the form of so called Complex Events, that is, situations that can only be recognized as a combination of several events.

Complex Events can therefore be considered the output or result of performing analysis and processing on Event Streams.

* Definition provided by the Event Processing Technical Society
Simple streaming example
How many dots can you see?
(Each dot is an event)
Four
Seven
Sixteen
Let’s add some complexity
How many red dots can you see?
Fourteen
Let’s add some temporal complexity
How many yellow dots appear after the blue dots?
Twenty
WHAT ELSE?

HOW MANY RED DOTS WILL there Be next?

IF there were 20 red dots, put the next three blue dots in a specific category.

If there are ten blue dots, join to alternate data and see how many blue dots there are in total?

If there were 6 green dots, wait for an hour, then compare again.
HIGH SPEED PROCESSING

DATA IN (called Events)

SAS EVENT STREAM PROCESSING ENGINE

From huge volume of streaming data flowing at very high rate:

*Millions of records/sec*

Data (Events) are processed as soon as they arrive (happen):

*Latency: milliseconds*

DATA OUT (Events)
STREAMING DATA ANALYTICS: BENEFITS

**BUSINESS USERS**
- Descriptive, Predictive, and Optimization Analytics on continuous data streams, right now!
- Access to streaming data information
- Reduced time to decision and action
- Generation of new opportunities
- Reduced risk and cost

**IT USERS**
- Ability to deal with the increasing volume and speed of big data flow
- Reduced storage (due to data aggregation) and computational needs
- Faster to market
- Easier & quicker maintenance of analytical models
- Focus on business logic rather than production features
REAL-TIME DECISIONING & ACTIONS: BENEFITS ACROSS INDUSTRIES

- Supply Chain
- Real Time Marketing
- Industry, Energy
- Manufacturing
- Cyber Security
- Fraud Detection
- Capital Markets
- Telecom
- IT Operational
- Enterprise Decision Management
- Decision Management
Analyze data in motion with low latencies

DATA IN
(called Events)

SAS EVENT STREAM PROCESSING ENGINE

Detect Events & Patterns of Interest

Filtering
Aggregation
Pattern detection
Computations
Correlations
Procedural
Text analytics
Data quality

and much more…

DATA OUT
(Events)
Analytics and Insights ON Streaming DATA

**TECHNOLOGY INTEGRATION**
- SAS® Event Stream Processing Engine is integrated into some SAS Solutions and can be deployed at the front-end of most others.
- Complements batch and real-time capabilities of SAS solutions with streaming data analysis.

**STREAMING ANALYTICS**
- Enables SAS analytic solutions to process streaming events.
- Leverages analytical model results to provide real time insights and action on streaming data.
- Enabled the deployment of additive and incremental analytic models on streaming data.
KEY CONCEPT
Design of the rule model (called “Continuous Query”) using components (called “Windows”)

DATAFLOW CENTRIC
HIGH LEVEL DEPLOYED EXAMPLES
Continuous Data Integration & Streaming Analytics

Event Stream Processing (ESP)

Analytics Servers
In-memory data Ad-hoc Analytics

SAS Reporting & Aggregation

3rd Party Reporting, API's & Adjustment Applications

DISPARATE DATA SOURCES

RISK AGGREGATION AND REPORTING

CLOUDERA HADOOP

NEAR-CACHE

LONG TERM-CACHE(S)

(RISK) DATA LAKE

HIGH LEVEL EXAMPLE - INVESTMENT BANK SAS & HADOOP SOLUTION

In-memory data

Ad-hoc Analytics

API
INVESTMENT BANK CHALLENGES ADDRESSED BY THE SAS SOLUTION

Risk Aggregation and reporting

• Timeliness
  Performance of existing offering is slow and cumbersome
  Access to data takes a long time requiring many ad-hoc sets of data to be extracted by Risk IT

• Transparency
  Raw and standardised views of data not available
  No ability to trace information through from front-office trades/positions to VaR

• Granularity
  No ability to access granular, low level data for analysis
  Hierarchy drilldown, querying and filtering capabilities required for proper analysis are not available

• Accuracy
  Comparison of data against front-office reports highlights numerous data inconsistencies
  Current manual analysis process introduces inaccuracies

Lack of Access to Data and limited Analysis ability
Solution Overview

High level architecture (ESP-RTDM)
use case

1. Take Rate: 5x Business As Usual

2. Revenue Increase: €100k/day

3. Setup: 8 Weeks, Payback < 6 Months

* 21M Active Customers, 50K TPS, Chile Argentina Spain
INTERACTING WITH ESP
There are five interaction points with ESP.

1) Other SAS products and solutions such as Visual Scenario Designer

2) ESP Studio

3) ESP STREAMVIEWER

4) ESP XML Server engine

5) ESP C++ API Framework
STREAMVIEWER

- HTML5 interface
- Uses HTTP (RESTful) XML server
- Has 2 modes:
  - Streaming mode: display all events
  - Update: events processed with opcode
- Subscribe & Publish
Come visit me at the IoT stand!
peter.pugh-jones@sas.com

www.SAS.com