Disclaimer

The information in this presentation is confidential and proprietary to SAP and may not be disclosed without the permission of SAP. Except for your obligation to protect confidential information, this presentation is not subject to your license agreement or any other service or subscription agreement with SAP. SAP has no obligation to pursue any course of business outlined in this presentation or any related document, or to develop or release any functionality mentioned therein.

This presentation, or any related document and SAP's strategy and possible future developments, products and or platforms directions and functionality are all subject to change and may be changed by SAP at any time for any reason without notice. The information in this presentation is not a commitment, promise or legal obligation to deliver any material, code or functionality. This presentation is provided without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. This presentation is for informational purposes and may not be incorporated into a contract. SAP assumes no responsibility for errors or omissions in this presentation, except if such damages were caused by SAP’s intentional or gross negligence.

All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, which speak only as of their dates, and they should not be relied upon in making purchasing decisions.
Continuous Intelligence with Streaming Analytics:
- insight from live event streams
- instantaneous response

Process live event streams
- Process events in real-time, as fast as they arrive
- Filter, enrich, transform, normalize
- Capture high value data in SAP HANA

Analyze data in motion
- Analyze events in the context of other events and historical data
- Watch for patterns, trends, correlations
- Apply predictive analytics to anticipate what’s coming
- Apply rules to determine what action to take

Act in Real-time
- Directly impact continuous business processes
- Alerts – while there is time to act
- Immediate response
- Continuous computation of KPI’s
- Stream key data to live dashboards
SAP rated a “Leader” in The Forrester Wave™: Streaming Analytics, Q3 2017

“Streaming analytics is essential for bringing real-time context to apps.” - Forrester

Forrester defines streaming analytics as:

Software that can filter, aggregate, enrich, and analyze a high throughput of data from multiple, disparate live data sources and in any data format to identify simple and complex patterns to provide applications with context to detect opportune situations, automate immediate actions, and dynamically adapt.
Benefits of SAP HANA streaming analytics

➢ Part of the SAP HANA Platform
  • Runs as a separate service in the HANA system
  • Managed via the HANA Cockpit
  • Access to HANA tables in streaming data models

➢ Scalable
  • Process events in real-time, as fast as they arrive
  • Hundreds of thousands to millions of events per second
  • Cluster architecture support low-cost multi-node scale out

➢ Fast – Low Latency
  • Millisecond latency from arrival to response

➢ Rapid Development
  • SQL-based event process language – familiar and concise
  • Both visual and text editors in HANA Studio
  • Rich set of testing tools in HANA studio

➢ Predictive Analytics
  • Built-in predictive models based on machine learning

➢ Range of Connectivity Options
  • Standard REST and WebSocket interfaces
  • Range of standard adapters
  • Extensible adapter toolkit
  • API’s for C++, Java and .NET

➢ Secure
  • All connectivity/actions require HANA authentication
  • Granular access control to the stream/action level

➢ Fault Tolerant
  • Multi-node clusters with auto-restart
  • Cluster managers operate as a peer network
  • Event windows can be configured as recoverable
Introduction to Streaming Analytics
Streaming data sources are everywhere

- Sensors
- Click streams
- Social media
- Transactions
- Market prices
Event stream processing uses continuous queries

**Database Queries**

- **Step 1:** Store the data
- **Step 2:** Query the data

**Continuous Queries**

- **Step 1:** Define the continuous queries and the dataflow
- **Step 2:** Wait for data to arrive. As it arrives, it flows through the continuous queries to produce immediate results
Streams and Windows

Input Streams (or windows)
Events arrive on input streams

Derived Streams, Windows
Apply continuous query operators to one or more input streams to produce a new stream

Database Analogy
- Streams are like tables
- Events are rows in the table
- Fields in an event message are columns in the table

Windows can Have State
- Retention rules define how many or how long events are kept
- Opcodes in events can indicate insert/update/delete and can be automatically applied to the window
Complex Event Processing extracts insight from raw events

Virtually no useful information in a single isolated event

Sensor readings – 10’s of thousands per second

Event window – e.g. 30 min

e.g. Compare variance of trends across multiple sensors against historical norms

Alert
Streaming analytics runs as a service of the HANA Platform

Input Streams

HANA DB

Streaming Server

SAP HANA System

adapters

Alerts

Dashboards

Applications

Hadoop
## Why Streaming Analytics?

<table>
<thead>
<tr>
<th>Situation Detection</th>
<th>Alerts</th>
<th>Immediate Response</th>
<th>Live Dashboards</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Watch for trends or patterns</td>
<td>- Example: Alert a supervisor when a machine needs adjustment before quality is affected</td>
<td>- Example: Automatically adjust prices based on market conditions</td>
<td>- Continuously compute and stream summary data to live dashboards</td>
</tr>
<tr>
<td>- Spot significant changes</td>
<td>- Example: Alert IT staff as soon as a security threat is detected</td>
<td>- Example: Dispatch a technician for urgent preventive maintenance</td>
<td></td>
</tr>
<tr>
<td>- Monitor correlations</td>
<td></td>
<td>- Example: Tailor an offer to a user based on current activity</td>
<td></td>
</tr>
<tr>
<td>- Compare current values to historical norms</td>
<td></td>
<td>- Example: Shut down a system to prevent damage</td>
<td></td>
</tr>
<tr>
<td>- Apply predictive models to anticipate what’s coming</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Receive events from streaming sources at high speeds

- Process hundreds of thousands or millions of events per second, in real-time
- REST and WebSocket interfaces
- Kafka, JMS and MQ adapters
- Extensible adapter toolkit plus other standard adapters

Filter, enrich and normalize incoming data

- Rather than capturing all the raw data in HANA, optimize HANA resources by capturing valuable information in an optimized data model

Data reduction

Options include:

- Sample high frequency data to reduce the number of data points
- Only record changes

Data Tiering:

- Store high value data in HANA in-memory tables
- Lower value data in HANA dynamic tiering or in Hadoop
- High speed HANA database loading with support for parallel writing to multiple partitions
Example Use Cases
The Internet of Things: insight from smart devices

Thousands of devices continuously sending bits of information

Insight from combining and correlating
– Combine data across related devices
– Watch for trends or patterns

Active monitoring: alerting and response
– Rather than simply storing the data and analyzing it later
– Identify where there is value in immediate awareness and response
– Let event stream processing generate alerts or initiate immediate action

- Location monitoring: alert operations staff if equipment is not where it’s supposed to be
- Predictive maintenance: detect imminent failure and take preventative action
- Smart grid: anticipate overloading and take action before it occurs
- Personalized offers – based on current activity, location, etc
Use Case – Utilities

Spatial and Graph
SAP HANA

Combine business data with geographical and streaming data.

Example: Proof of Concept with Alliander
Tracing impact of outages in electricity network (47 million records) improves from 3.5 hours to 2.5 seconds.
Use Case – Utilities

Speed layer
- Grid Monitoring Events
- Customer Issues
- Detect Voltage Irregularities
- Are they geographically related?
- Do grid issues overlap customer issues?
- Based on location evaluate impact
- Use ML scoring to predict likely cause of issue

Batch layer
- Substations
- Grid Layout
- Customer Details
- Geospatial Relationships
- Record customer incident
- Trained ML Models
- Create & update work order

Automatic response to Customer
- Generate work order with probable cause
Improving Energy Demand Response with SAP HANA® Streaming Analytics

As part of its long-term load management program, SBB worked with SAP Innovative Business Solutions to create a smart, real-time power demand management platform based on SAP HANA® streaming analytics. The solution will automatically decrease peak loads and is expected to reduce SBB’s energy costs. After a successful proof of concept, SBB now has the foundation for its ongoing energy management transformation.

Before: Challenges and Opportunities
- Manage constantly fluctuating energy demands more efficiently
- Respond to frequent, sudden peak loads during rush hours and in winter when heating is at maximum
- Implement a system that monitors power demand in real time and automatically reduces load by proactively switching off energy-consuming appliances, such as heaters

Why SAP
- Implement standard software with SAP HANA streaming analytics to form the smart core of a complex signal processing solution and provide a proven foundation for future innovation projects
- Close collaboration between SAP and SBB, working as one unified team in Switzerland from the proof of concept to development of SAP-specific system components
- Using agile methodology, the SAP Innovative Business Solutions team delivered a proof of concept in just six weeks

After: Value-Driven Results
- Postponed investments in new energy infrastructure by leveraging existing assets
- Savings in peak energy costs expected
- Increase the reliability of energy supply
- New opportunities to control other energy-consuming appliances in the future, such as train traction (engines)

“By realizing our power demand management platform, we’re building the foundation for future business cases.”
Markus Haldier, Head of Power Demand Management Program, SBB
Providing a Next-Generation Air Service with SAP® Leonardo Internet of Things

To optimize its Sigma Smart Air Service, Kaeser worked with SAP Digital Business Services to deploy SAP Leonardo IoT capabilities as its innovation foundation together with SAP Asset Intelligence Network and SAP Predictive Maintenance and Service. Kaeser’s new solution connects its compressors smartly in the cloud, allowing it to offer a next-generation service at a lower price.

Before: Challenges and Opportunities
- Service team unable to access calibration data and other equipment-specific information, which was stored in on-premise systems
- No solution to meet the needs of dealers and companies’ service providers
- Need for track-and-trace capabilities with selected suppliers to scale up potential

Why SAP
- Strategic partnership and co-innovation with SAP
- SAP Asset Intelligence Network based on SAP S/4HANA® Cloud for use in maintenance of equipment information and greater transparency over usage
- Standard integration between SAP S/4HANA, SAP Predictive Maintenance and Service, the SAP Manufacturing Execution application, and third-party solutions
- Ability to establish a “digital twin” to build entirely new data-driven business models

After: Value-Driven Results
- Increased service efficiency with one single source of product, component, spare-part, and service information
- Higher customer and dealer satisfaction with an up-to-date digital replica and improved supplier collaboration
- Greater transparency across the company’s installed base, enabling it to offer additional digital services and new business models
- Enhanced flexibility of a cloud-based solution

“SAP Leonardo technologies provide the framework for different applications that can be easily combined to orchestrate your own business model.”
Falko Lameter, CIO, Kaeser Kompressoren SE

New
Business models based on real-time data

Higher
Customer lifetime value and service efficiency

>1 million
Measurements per day to avoid unplanned downtime

See how Kaeser boosted its air-service business with SAP software.
Continuous Manufacturing - Paper Breakage Use Case

Preventing paper breakage, reducing waste production and time-to-quality are key.

Fusion of sensor measurements and quality parameters to derive predictive models for real-time implementation to reduce paper breakage.

Leads to:
- Decrease production losses
- Increase plant utilization
- Reduce time-to-quality, in particular for rarely and never-before produced materials
- Monitor, guarantee and possibly improve continuous quality standards
Industrial Worker Safety – Gas Leak & Fall Detected

Scenario 1

Wearable Sensors
Joe is an electrician working in hazardous conditions. He is equipped with a number of wearable IoT sensors that can detect dangerous gas levels, his heart rate and his posture.

Dangerous Gas and a Fall!
The sensors monitor in real-time for unsafe conditions. As the sensor detects dangerous levels of gas Joe is alerted and attempts to make his way to safety, on his way he falls and is left lying on the ground.

Alert Manager
Data from Joe’s sensors stream real-time to his manager’s dashboard. The safety manager is first alerted of the gas and then a second alert to indicate the fall. An emergency response team is sent to aid Joe.

Report Incident
With Joe now safe, his manager can complete the necessary Incident Reporting, this task is sped up due to an automatically created incident report which has prefilled much of the incident information based on the alert.

Analyze Data
Once the incident is submitted, analysts at head office can use relevant IoT sensor data automatically included with the report to look for trends or patterns and to try to make the workplace safer.
The magnitude of the problem:
• The most important logs in an SAP system generate altogether between 100k and 1M records/hour
• 300 systems generate up to 300M records/hour, which is about 90GB/hour

Types of attack patterns:
• Related to human behavior
• Of technical origin

Browse and explore events
• Overview by system, timeframe, context, …
• Filter and search
• Create alerts

View and work on alerts
Real-time click stream analysis

Monitor user behavior in aggregate
• Respond to trends in real-time
• Jump on surges before they subside
• Give more prominence to topics attracting viewers now

Customize offers based on individual patterns of behavior
• Don’t limit customization to historical behavior, but include current behavior

Prevent fraud
• Watch for patterns that indicate likelihood of fraud; act before the fraud takes place
Real-Time Financial Risk Consolidation and Monitoring

- Real-time consolidation across multiple trading platforms, multiple asset classes
- Real-time valuation using live market prices
- Custom pricing algorithms for non-liquid securities
- Incoming trades immediately update all sensitivity calculations
- Real-time P&L calculations
- Continuous limit monitoring and alerting; set limits at any aggregation levels
Deployment
What does this look like to a HANA application developer?

- For production systems, it’s normally recommended to run streaming on dedicate machines – but can be low-cost commodity Linux servers (doesn’t require HANA certified hardware)
- Design time tools are available as a plugin to SAP HANA Studio or SAP Web IDE
- Installed via HANA Installer
- Streaming server(s) runs on dedicated host(s)
- Administration via SAP HANA Studio and Cockpit
- Streaming projects are managed in the HANA repository and can be included in, and deployed as DU’s
- Optional – licensed separately from base
SAP HANA system with a one or more streaming nodes

- Streaming capacity scales independently of SAP HANA database
- Add streaming nodes for additional stream processing capacity
- Streaming projects connect to SAP HANA database via SAP HANA ODBC driver
- Streaming projects can send events to XS application using http output adapter
Hardware Layout View

Streaming hosts can be, but do not have to be, HANA-certified servers

Certd. HW Box

Certd. HW Box

Certd. HW Box

HANA Scale-Out

Node 1

SAP HANA core

Node 2

SAP HANA core

Standby Node

SAP HANA core

Host

DT node

Dynamic Tiering Node

Host

Streaming Cluster

Streaming Node

Streaming Node

Certd. Storage for data and redo logs of HANA core and DT

Certd. Storage for data and redo logs of HANA core and DT

Certd. Storage for data and redo logs of HANA core and DT

Certd. Storage for data and redo logs of HANA core and DT

Client Network

Intra-node Network

Storage Network for core and DT

Streaming Persistency Network

HANA System (One SID)

© 2016 SAP SE or an SAP affiliate company. All rights reserved.
Hardware considerations

- Although streaming does not require HANA-certified configuration, hardware should be from an SAP hardware partner: [http://scn.sap.com/docs/DOC-8760](http://scn.sap.com/docs/DOC-8760)

- See the streaming analytics sizing guide for indicative capacity for each of the T-shirt sizes. Some systems can go much larger (e.g. 200 cores)

- System configuration must consider network bandwidth and HANA DB workload for stream capture

- [SAP HANA Smart Data Streaming Sizing and Configuration Guidelines](#)
Streaming Scalability

- Streaming projects are multi-threaded and will scale across all available cores
- Streaming hardware scales independently of HANA core; can run one or more streaming nodes in a single HANA system
- Large projects can be broken into smaller inter-connected projects running on different nodes
- Partitioning feature eliminates bottlenecks within a project by providing multiple parallel processing paths

SAP HANA streaming analytics provides extreme scalability. Throughput capacity can always be increased by adding hardware.
Installation of HANA streaming analytics

1. **Download from Service Marketplace**
   - Separate download from SAP HANA core (since it’s optional)

2. **Install via SAP HANA installer**
   - Add to an existing SAP HANA system
   - Install as part of a new SAP HANA system install

3. **Additional download packages:**
   - Streaming studio plugin (for HANA Studio)
   - Streaming analytics client package – adapters, ODBC driver, SDK’s
Streaming server(s) in the SAP HANA Administration Console

- Streaming nodes will be listed in the landscape view of the SAP HANA Administration Console in the HANA Studio.
- Will show status and provide ability to stop/start nodes. Use HANA Cockpit for project-level control.

![Streaming server(s) in the SAP HANA Administration Console](image)
**HANA Cockpit: Centralized Monitoring and Administration**

- Options for managing nodes, workspaces, projects, adapters, etc.

- Drilling down provides additional detail. For example, clicking on the “Manage Streaming Projects" option will open up the page “Project Details".
Streaming Lite: HANA HSA Remote Deployment Option

Deployment option included with SAP HANA streaming analytics
- CCL projects can run at the edge
- Small footprint package
- Supported on ARM and Intel processors
- Certified on Red Hat and Raspbian

* Streaming lite can exist by itself, but may also be deployed in conjunction with other gateway components

**Diagram:** Devices → Gateway (Streaming Lite*, SQL Anywhere) → HTTPS → Data Center (Applications, Streaming Analytics, HANA Database) → SAP HANA Platform

- Alerts, KPI’s, Action events
Streaming Lite

- Light-weight HSA server process designed for gateway deployment in IoT scenarios
  - Smaller install image - reduced disk and memory footprint

- Compiled CCL project runs as a stand-alone Linux process (no HANA system, no SDS cluster)

- Supported on Linux: 32 bit ARM and 64 bit Intel x86 (or compatible) processors
  - Certified on Raspberry Pi with Raspbian (for 32 bit ARM)

- Streams output via WebSocket to SDS
  - New internal output adapter to stream output from Streaming Lite (at the gateway) to HANA smart data streaming in the cloud, via WebSocket

- **Filter close to source**
  - Reduce bandwidth and need for excessive central capacity
  - Not limited to simple filtering: can apply complex, dynamic filter criteria

- **Aggregate**
  - Examples include: reduce sampling rate; reporting trending rather than raw values

- **Local alerting/response**
  - Generate local alerts or response without relying on round-trip to center (or even connectivity to center)
Design Time
Design Time

• Streaming “projects” define the input streams, continuous queries, and outputs

• Projects are defined in CCL – the stream processing language used
  • Derived from SQL for familiarity and ease of use
  • Models can be defined in just a few lines of CCL
  • Maximizes user productivity

• Streaming plugin for Eclipse provides both a CCL editor and visual editor, along with testing tools
Design time tools for Eclipse / SAP HANA Studio

Streaming plug-in for Eclipse includes a visual editor for defining continuous queries and directing stream flow, plus run/test tools:

- Event processing language: CCL (like SQL for streams)
- CCL Script (like event driven stored procedures for streams)
**HANA Streaming Analytics Design Time Tools in SAP Web IDE for SAP HANA**

Web based tooling is now viable as an SDS design time environment independent of SAP HANA Studio

- Add streaming modules to multi-target applications
- **Enhancements to CCL editor:** Parsing and Syntax highlighting, editor auto-completion, semantic error checking, improved compilation error handling, ability to add CCL snippets for ease of use.
- **Visual project viewer:** view the diagram of a CCL project (read only – no visual editing)
- **Run-Test Tools:** project start/stop, stream viewer, record/playback, manual input, file upload, manage data services

_Note: Not yet a full functional replacement for SAP HANA Studio. There are still advanced features available in Studio that are not yet available in the browser-based tools._
CCL: Continuous Computation Language

SQL-based Event Processing Language

Leverage familiarity and simplicity of SQL

Instead of “snapshot” queries, continuous queries
  Filter (WHERE), Aggregate (GROUP BY), Join, Compute

Extensions for event streams
  Windows, Patterns

<table>
<thead>
<tr>
<th>Example</th>
<th>CCL Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculation of 10 minute moving average</td>
<td>CREATE WINDOW MovAvg10 PRIMARY KEY DEDUCED AS SELECT instream.ID AS ID , avg (instream.VALUE) AS Avg FROM instream KEEP 10 MINUTES GROUP BY instream.ID ;</td>
</tr>
</tbody>
</table>
Predictive Analytics for Streams
Machine Learning Algorithms

Native Algorithms
• Adaptive Hoeffding Tree
• Denstream Clustering

Imported Scoring Algorithms
• Leverage HANA Predictive Analytics Library ("PAL")
• CART
• C4.5
• CHAID

Tooling Integration
• SAP Business Objects Predictive Analytics
• Automated Analyst Persona
• Scoring algorithms generated in CCL
Machine Learning Algorithms: predictive analytics for streams

Adaptive Hoeffding Tree
- Train against live data streams, continuously updating the model, discovering predictive relationships
- Score incoming data using the trained model

Denstream (Clustering)
- Incremental clustering algorithm continuously adapts to incoming data
- Assign incoming events to clusters
- Identify outliers

Imported Decision Trees
- Train models using standard PAL decision tree algorithms against historical data in SAP HANA
- Import models into HSA project and use them for live scoring of incoming events

Identify outliers

Decide when action is needed – and what action to take
Streaming projects connecting to HANA tables and views

- High speed event stream capture into HANA tables

- Update HANA tables based on analysis of incoming event streams
  - Streaming projects can apply inserts, updates, and deletes to HANA tables – not limited to event logging
  - This allows HANA tables to “mirror” event windows in streaming projects

- Join event streams to HANA tables/Views
  - Reference HANA tables/Views directly in CCL projects

- Pre-load reference data from HANA DB into streaming projects
  - With ability to set refresh interval

- Run event-driven analytics on HANA DB
Connectivity
The HSA server provides connectivity to HSA projects via a “streaming web service”
Publishers have a choice of http/REST or WebSockets
Subscribers can browse meta-data via REST API and subscribe to output streams via WebSockets
Streaming projects can connect as Kafka consumers, producers, or both.

HANA database

HSA Project

Producer

Consumer

Producer

Consumer

Alerts

Applications

Kafka Support
Overview of other HSA integration options

Data sources

Adapters

HSA Project

input

output

HSA pub/sub library

Alerts

Applications

Application with HSA connectivity

Application with HSA connectivity

Input adapters receive/fetch messages using a protocol supported by the source, parse the message, and publish it to HSA as an event. Output adapters do the same – in reverse.

The HSA SDK (C++, Java, .NET) provides a library that can be embedded in adapters or applications to connect to SDS and publish/subscribe.
SAP HANA streaming analytics provides a range of connectivity options

**Standard:**
- http: REST, WebSockets
- Messaging: Kafka, JMS, MQ, TIBCO
- SOAP
- Databases
- Files
- OSIsoft
- TCP Sockets
- SAP RFC
- SAP Replication Server (in)
- Logfile (in)
- Microsoft Excel (out)
- Email (out)

**Parsing/Formatting**
- JSON, Avro
- XML events
- XML docs
- CSV
- FIX
- JMS Object Arrays

**Custom/Extensible:**
- Adapter Toolkit supports pluggable transport/parse/format modules written in Java
- API’s for C/C++, Java, .NET
Operational Dashboards and UI options for streaming applications

- **Design Studio 1.4 and up** - includes support for visualization of streaming data, with connectivity to HANA streaming projects

- **HTML 5 (inc. SAP UI5)** - Standard REST interface and websocket support enables custom html5 applications (including SAP UI5) to receive/display streaming data
Resources

SAP HANA streaming analytics Developer Center

https://developers.sap.com/topics/hana-streaming-analytics.html
Thank you

Contact information:

Rob Waywell
SAP HANA Product Management

https://developers.sap.com/topics/hana-streaming-analytics.html