Complex Event Processing

Dr. Jürgen Krämer
VP Product Strategy IBO &
Product Management Apama

Big Techday 7
23.05.2014

software AG

Helping Organizations Transform into Digital Enterprises

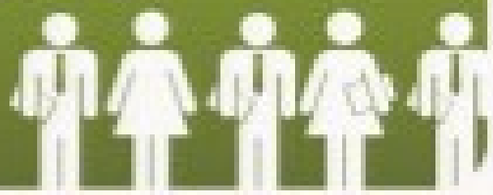
CUSTOMER BASE
More than
70%
of the Global 1000

2M+
Developers

€1B+
In Revenue

2.5M+
Users

5,000+
Employees Worldwide



Software AG dedicated to customer success for over 40 years

What is Complex Event Processing (CEP)?

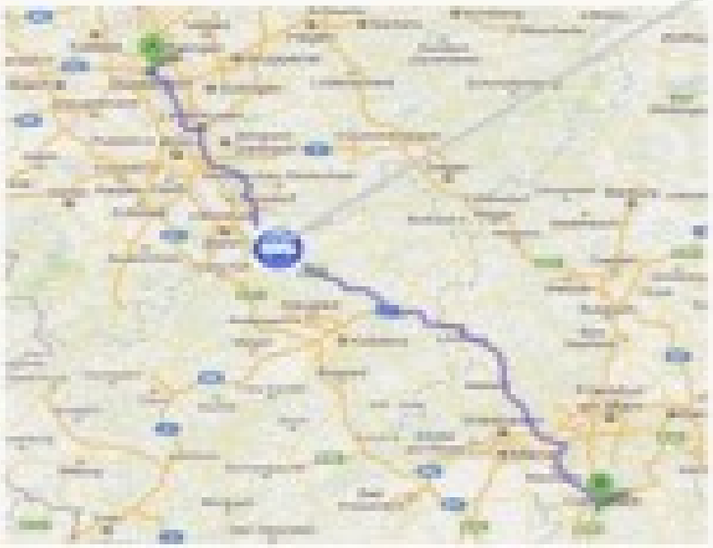
What is an Event?

What is Complex Event Processing (CEP)?

What is an Event?

Example: Vehicle

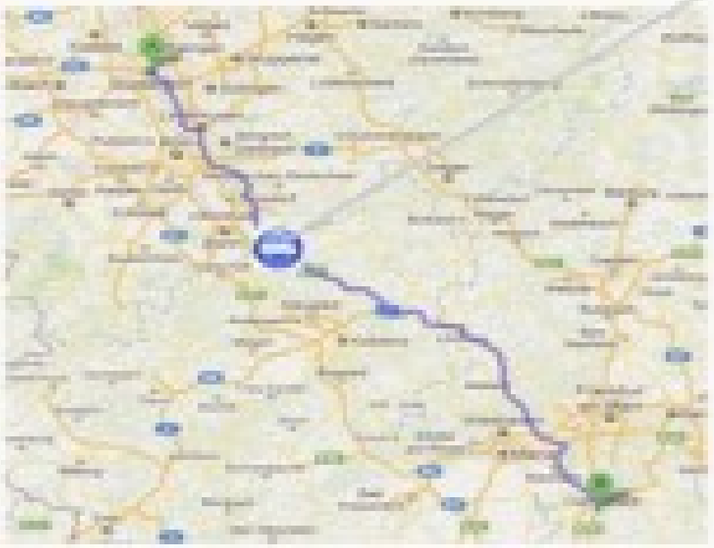
VehicleId: 23
Timestamp: 11:22
Geolocation:
+50° 43' 33.69", +7° 14' 0.89"



What is an Event?

Example: Vehicle

VehicleId: 23
Timestamp: 11:22
Geolocation:
+50° 43' 33.69", +7° 14' 0.89"



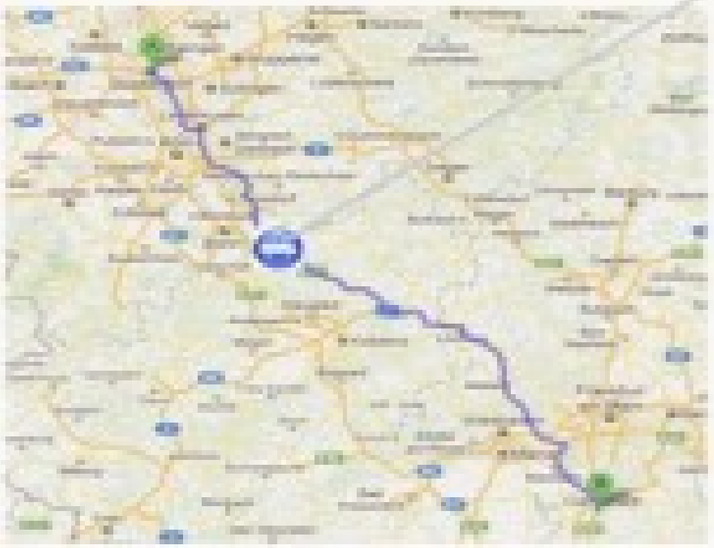
Example: SmartMeter

MeterId: 8754862
Timestamp from: 16:00
Timestamp to: 16:15
Consumption: 4.67 kWh
Max. Power: 1.12 kW

What is an Event?

Example: Vehicle

VehicleId: 23
Timestamp: 11:22
Geolocation:
+50° 43' 33.69", +7° 14' 0.89"



Example: SmartMeter

MeterId: 8754862
Timestamp from: 16:00
Timestamp to: 16:15
Consumption: 4.67 kWh
Max. Power: 1.12 kW

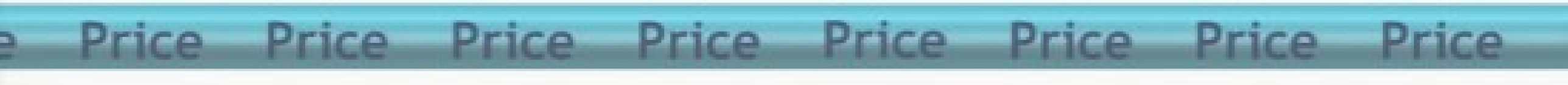
Example: Stock Price

Share: SOW
Timestamp: 14:57 Uhr
Price: €26.57

Increase in Data Velocity

Stream of position reports from a truck

Stream of stock prices



Stream of sensor readings from a smart meter

In-Memory, In-Flight Analytics

GPS GPS GPS GPS

Price Price Price Price

Value Value Value Value

Complex Event Processing (CEP)

Definition: Continuous analytics to derive meaningful business events from different event streams or other event sources like databases in real-time to gain situation awareness and trigger immediate actions.

- Event-driven, incremental processing
- High efficiency and scalability
- Enrich events with context data
- Detect patterns with time/location parameters

In-Memory, In-Flight Analytics

GPS GPS GPS GPS GPS

Price Price Price Price Price

Value Value Value Value Value

esult

Complex Event Processing (CEP)

Definition: Continuous analytics to derive meaningful business events from different event streams or other event sources like databases in real-time to gain situation awareness and trigger immediate actions.

- Event-driven, incremental processing
- High efficiency and scalability
- Enrich events with context data
- Detect patterns with time/location parameters

In-Memory, In-Flight Analytics

S GPS GPS GPS GPS

e Price Price Price Price

ue Value Value Value Value

Resul

Complex Event Processing (CEP)

Definition: Continuous analytics to derive meaningful business events from different event streams or other event sources like databases in real-time to gain situation awareness and trigger immediate actions.

- Event-driven, incremental processing
- High efficiency and scalability
- Enrich events with context data
- Detect patterns with time/location parameters

In-Memory, In-Flight Analytics

GPS GPS GPS GPS GPS

Price Price Price Price Price

Value Value Value Value Value

Complex Event Processing (CEP)

Definition: Continuous analytics to derive meaningful business events from different event streams or other event sources like databases in real-time to gain situation awareness and trigger immediate actions.

- Event-driven, incremental processing
- High efficiency and scalability
- Enrich events with context data
- Detect patterns with time/location parameters

In-Memory, In-Flight Analytics

S GPS GPS GPS GPS

ce Price Price Price Price

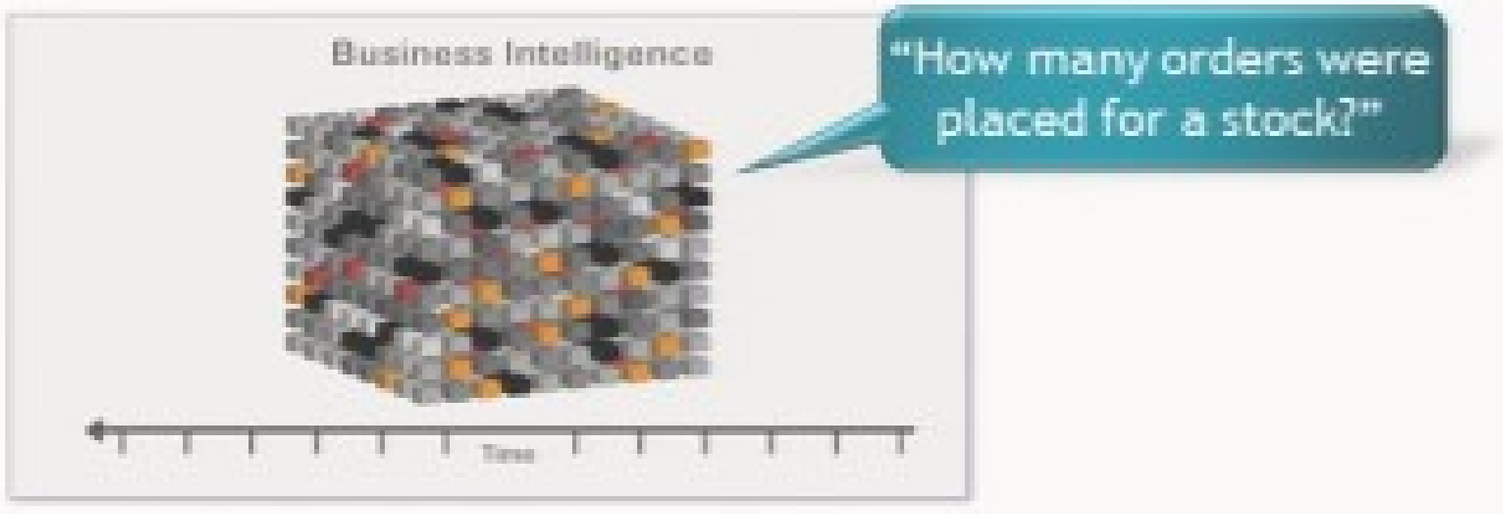
ue Value Value Value Value

Complex Event Processing (CEP)

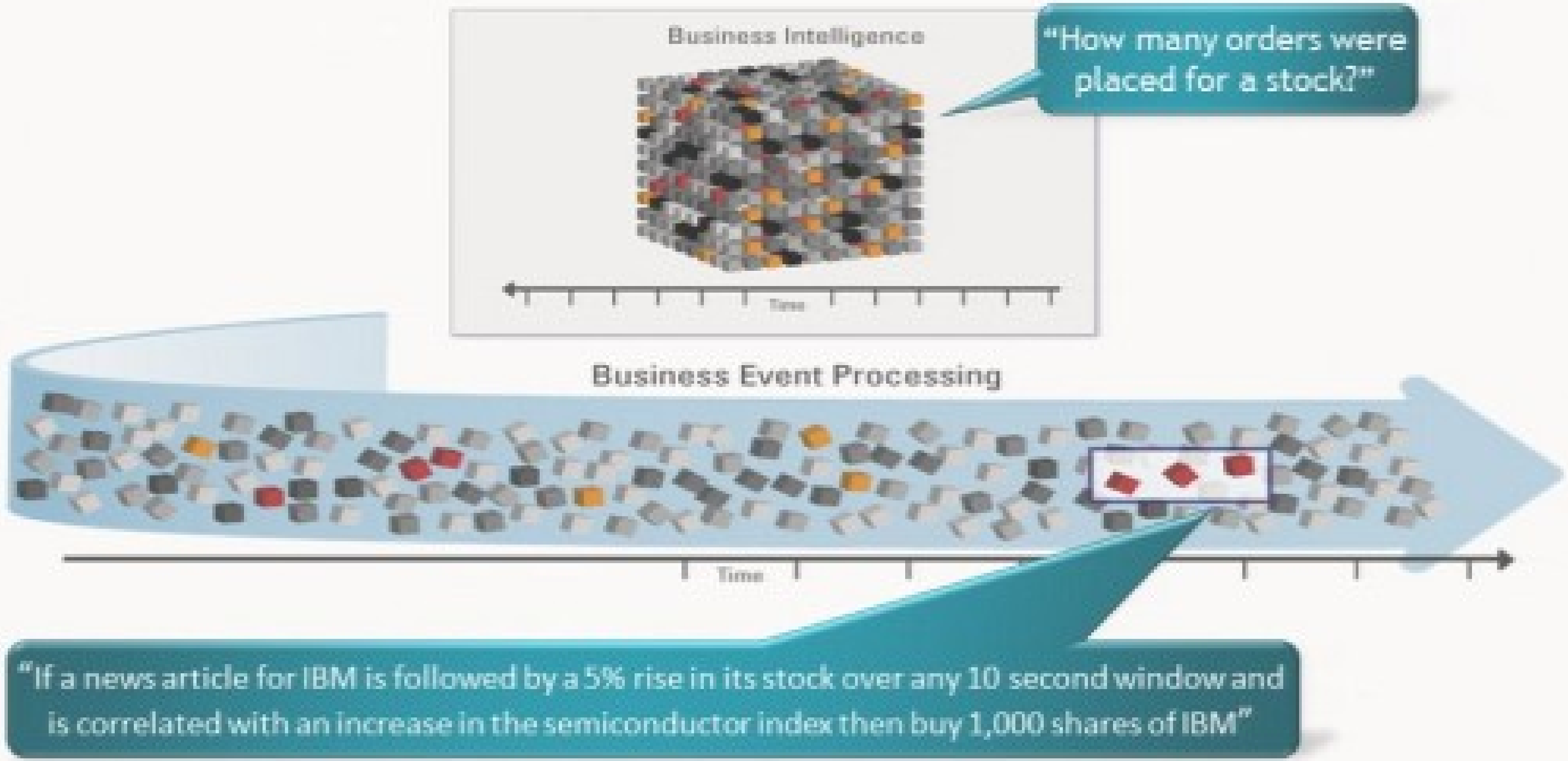
Definition: Continuous analytics to derive meaningful business events from different event streams or other event sources like databases in real-time to gain situation awareness and trigger immediate actions.

- Event-driven, incremental processing
- High efficiency and scalability
- Enrich events with context data
- Detect patterns with time/location parameters

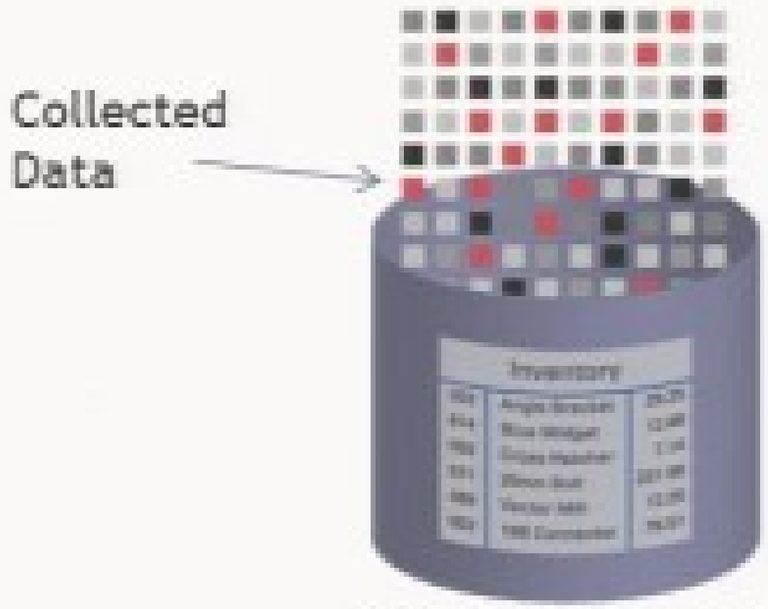
Business is Event-Driven



Business is Event-Driven

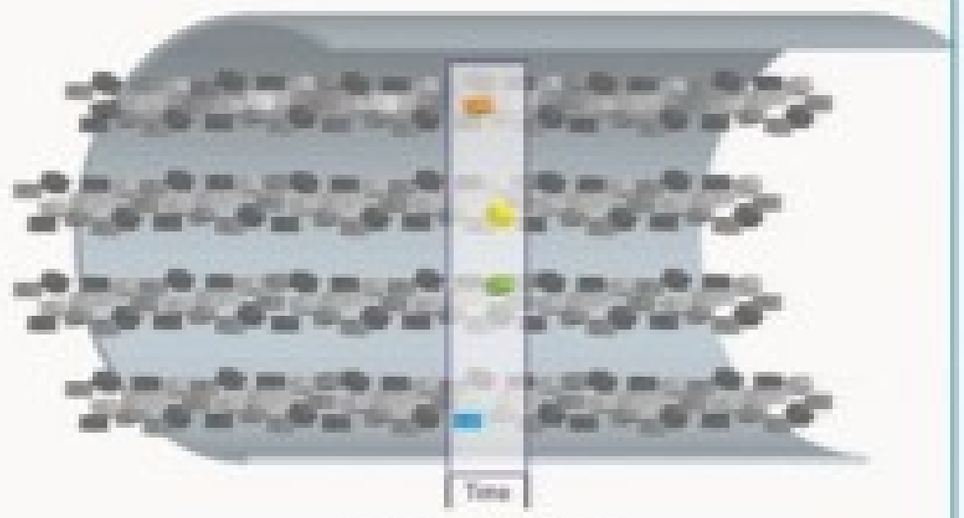


Data Processing Differences



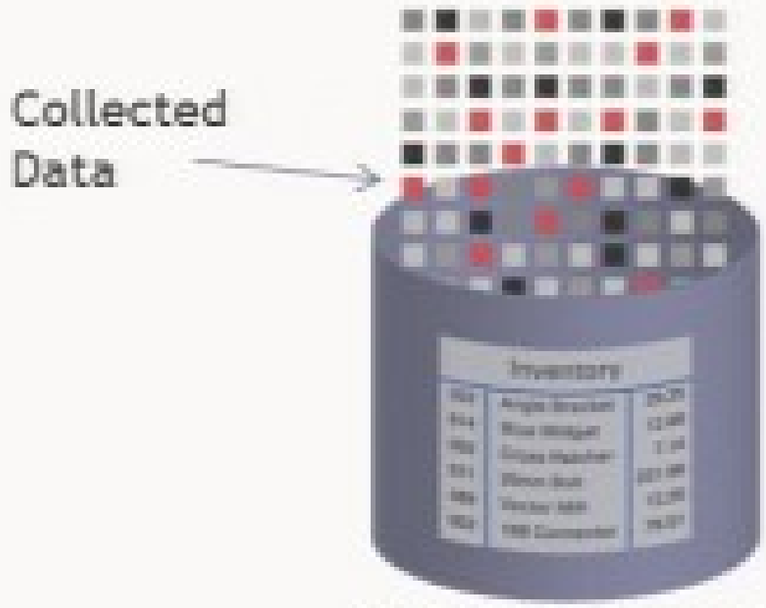
Traditional Decision Management

- Data goes **into** the database
- Analyzed **after** it is (indexed and) stored



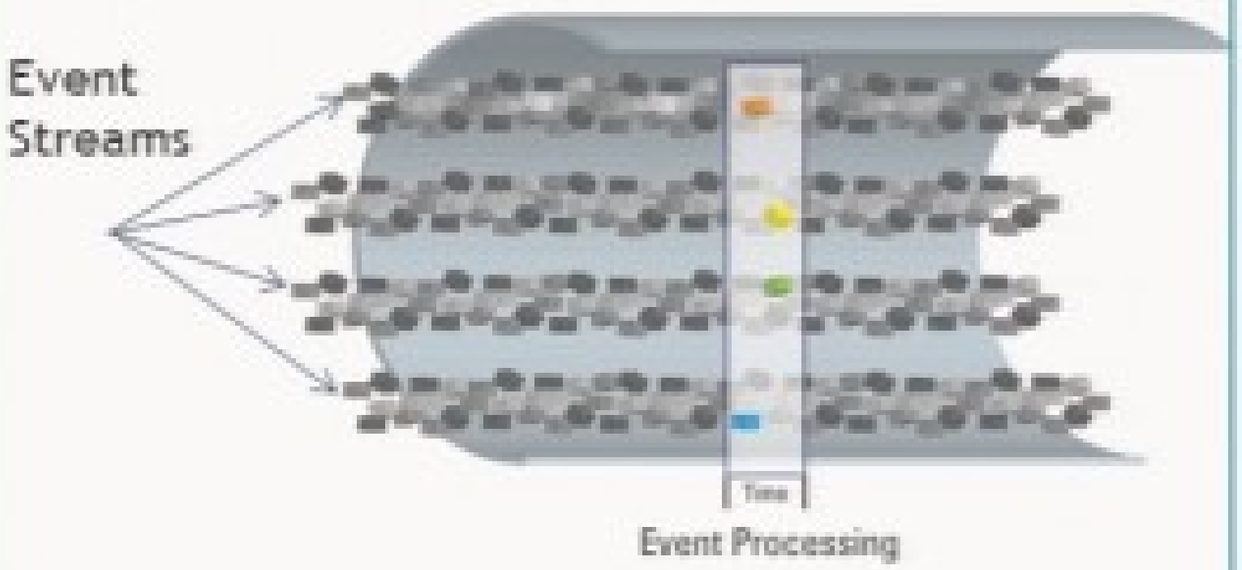
Event Processing

Data Processing Differences



Traditional Decision Management

- Data goes **into** the database
- Analyzed **after** it is (indexed and) stored

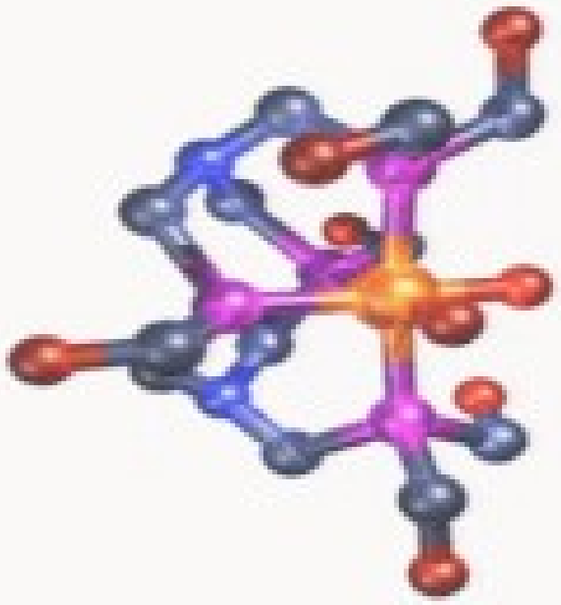


Event Processing

- Events go **through** CEP engine
- Analyzed **as they happen**
- Results **instantly** usable

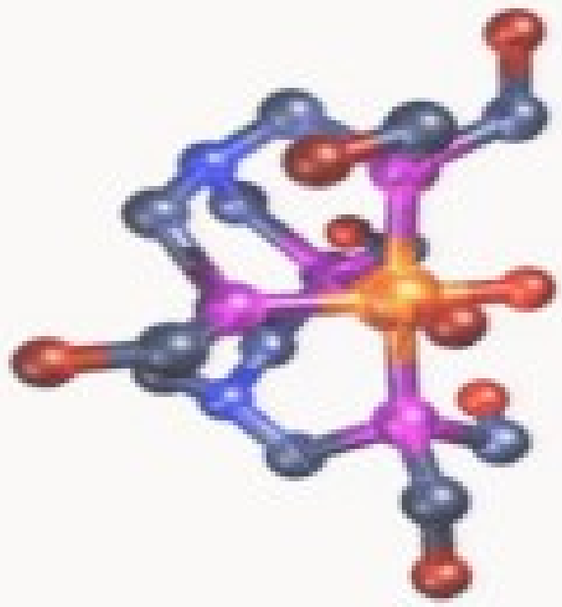
Principles of Event-based Systems

- Treat any system update as an "event"



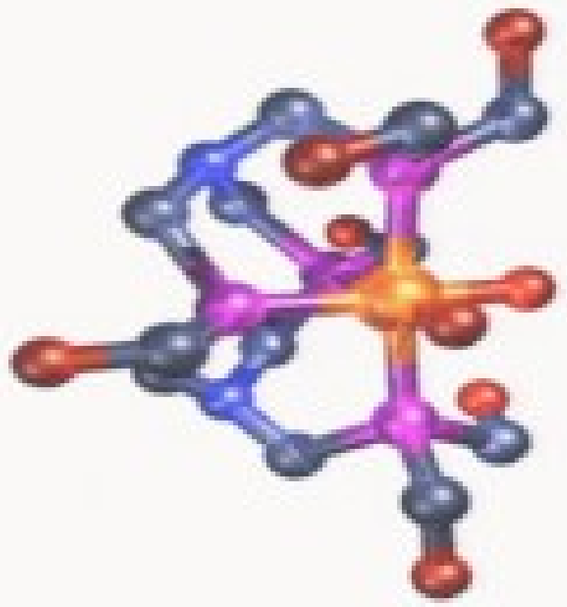
Principles of Event-based Systems

- Treat any system update as an "event"
- Enable event-based patterns to be defined to monitor, analyze and act on "event patterns"



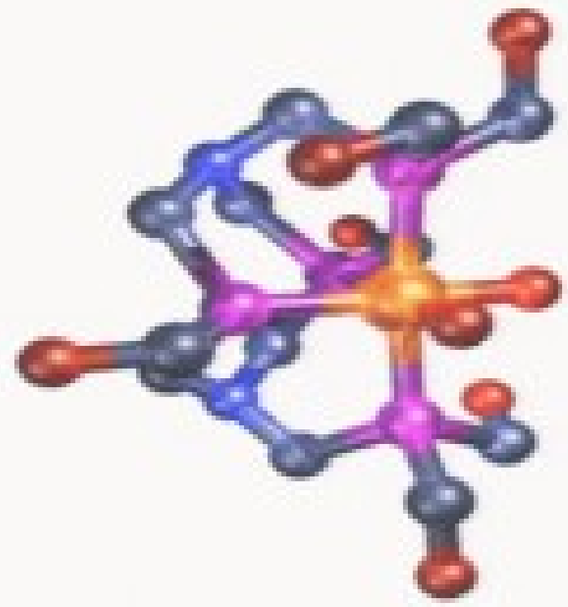
Principles of Event-based Systems

- Treat any system update as an "event"
- Enable event-based patterns to be defined to monitor, analyze and act on "event patterns"
- Patterns identifying opportunities and threats can be defined rapidly



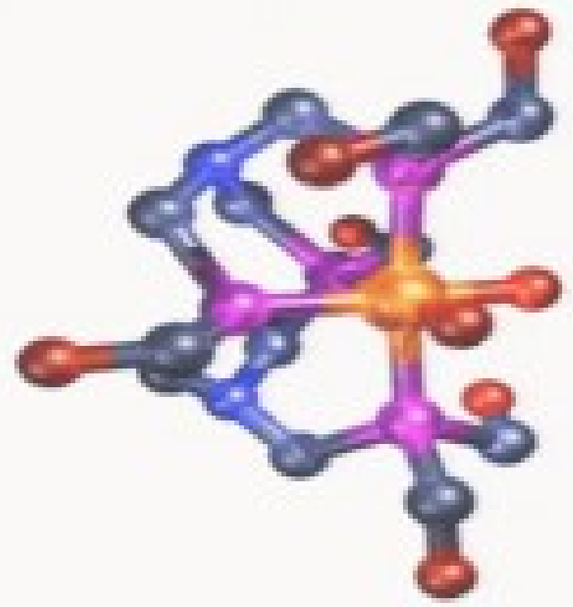
Principles of Event-based Systems

- Treat any system update as an "event"
- Enable event-based patterns to be defined to monitor, analyze and act on "event patterns"
- Patterns identifying opportunities and threats can be defined rapidly



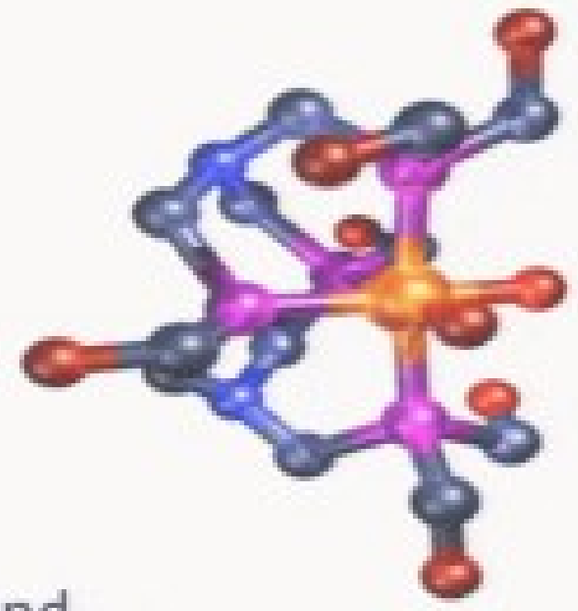
Principles of Event-based Systems

- Treat any system update as an "event"
- Enable event-based patterns to be defined to monitor, analyze and act on "event patterns"
- Patterns identifying opportunities and threats can be defined rapidly
- Patterns are loaded into a *real-time engine* that offers analysis and response with low latency



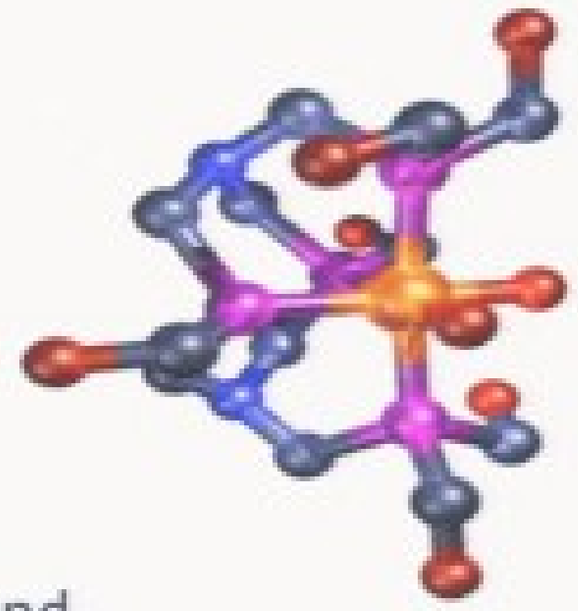
Principles of Event-based Systems

- Treat any system update as an "event"
- Enable event-based patterns to be defined to monitor, analyze and act on "event patterns"
- Patterns identifying opportunities and threats can be defined rapidly
- Patterns are loaded into a *real-time engine* that offers analysis and response with low latency
- Engine is permanently connected to multiple event sources and destinations



Principles of Event-based Systems

- Treat any system update as an "event"
- Enable event-based patterns to be defined to monitor, analyze and act on "event patterns"
- Patterns identifying opportunities and threats can be defined rapidly
- Patterns are loaded into a *real-time engine* that offers analysis and response with low latency
- Engine is permanently connected to multiple event sources and destinations
- Events can be captured and preserved in time-order for historical pattern analysis and root-cause analysis

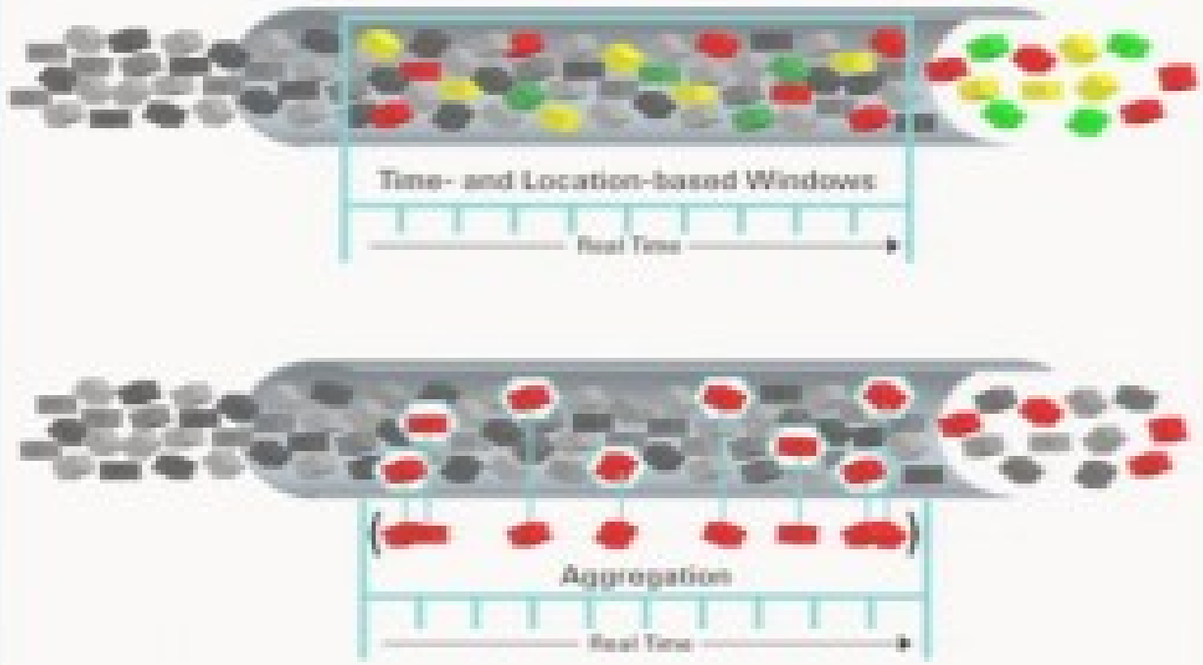


Event Processing Models

- Time- and location-based windows
  - Within, near, etc. based in real-time context

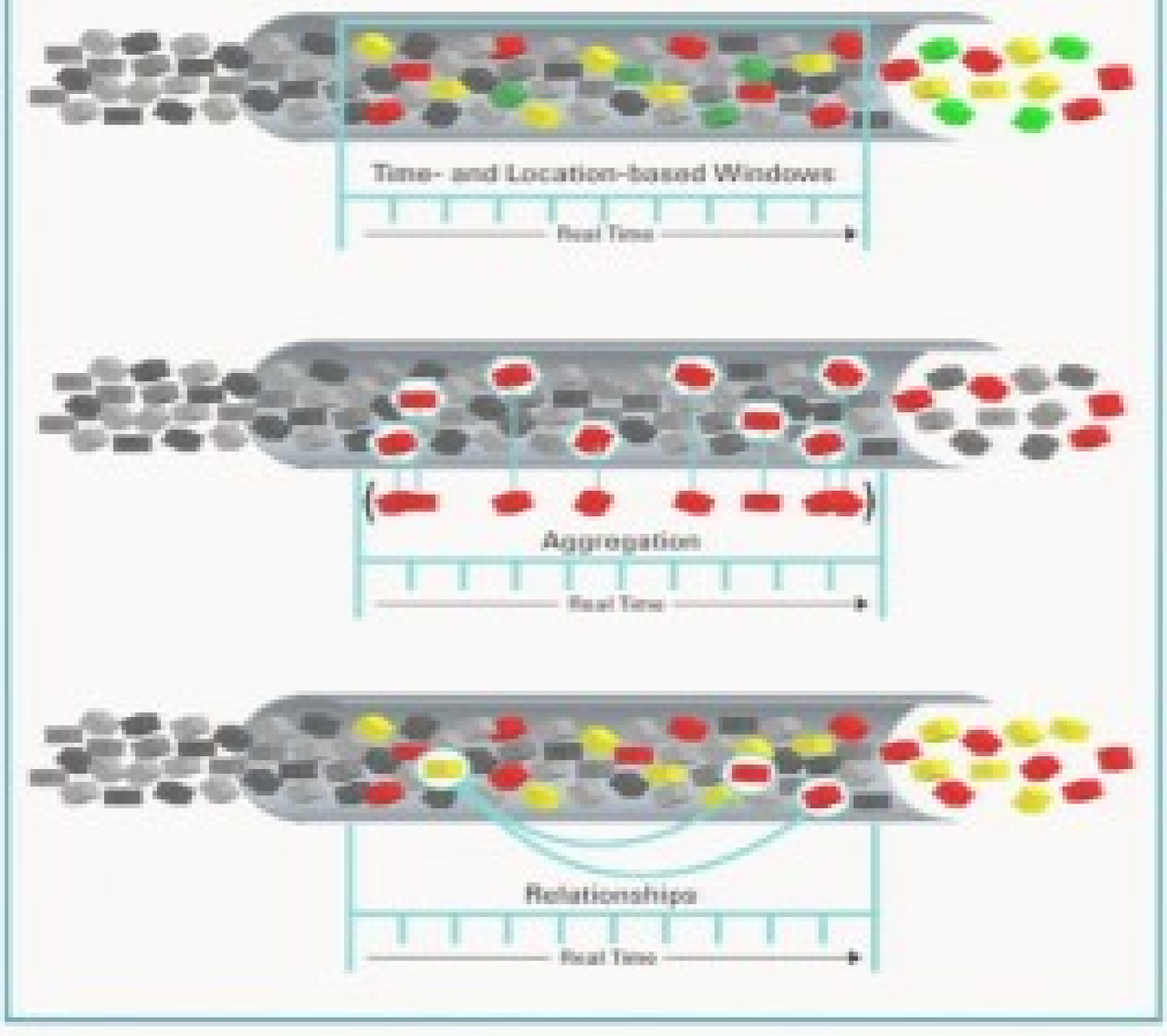
Event Processing Models

- Time- and location-based windows
  - Within, near, etc. based in real-time context
- Grouping & Aggregation
  - Accumulation of values or quantity
  - Sum, average, min, max, etc.
  - Support for custom aggregate functions



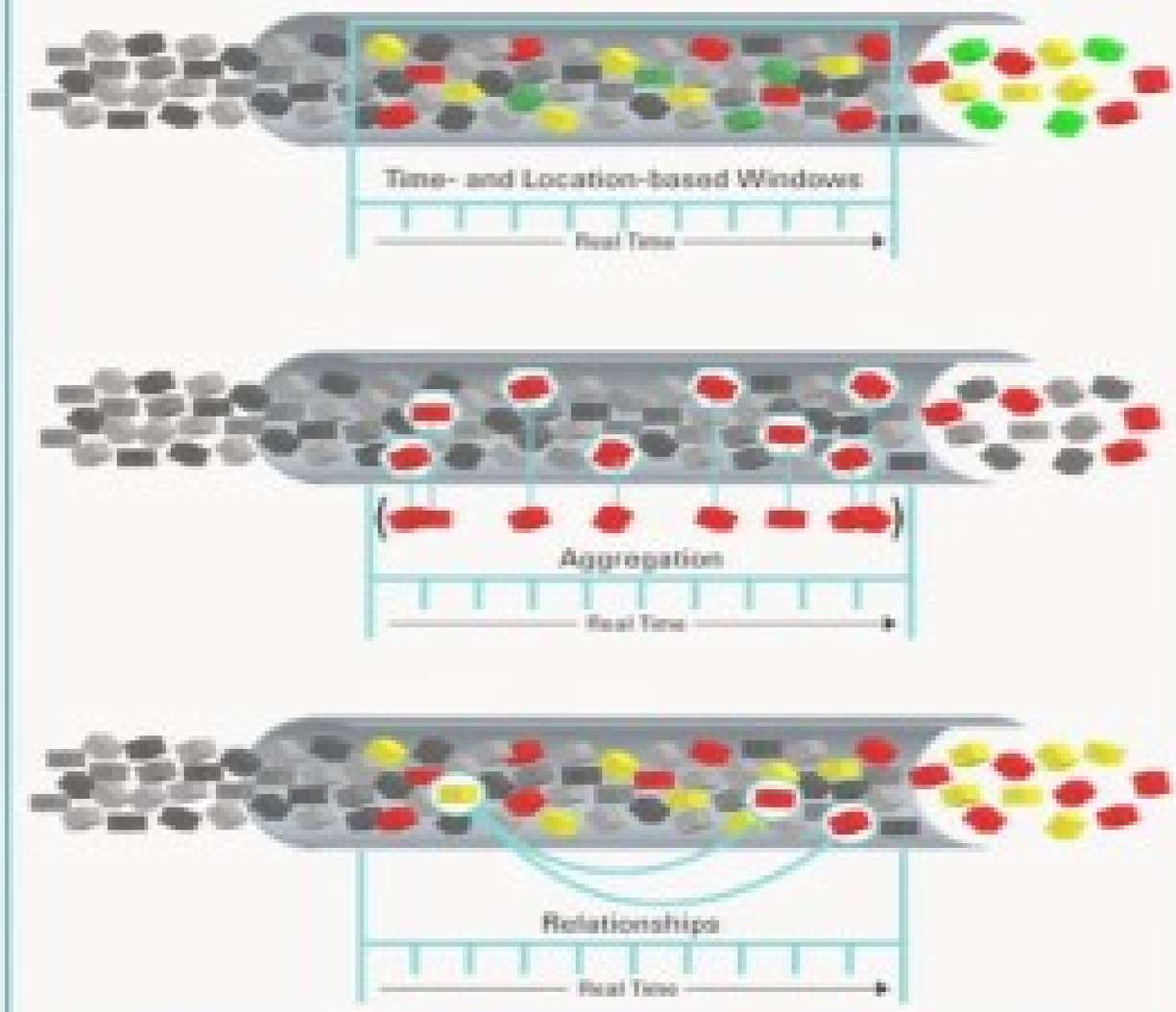
Event Processing Models

- Time- and location-based windows
  - Within, near, etc. based in real-time context
- Grouping & Aggregation
  - Accumulation of values or quantity
  - Sum, average, min, max, etc.
  - Support for custom aggregate functions
- Event Relationships
  - Event A followed by event B
  - Event A and event B
  - Event A or event B
  - The non-event
- Event Enrichment
- User-defined Functions



Event Processing Models

- Time- and location-based windows
  - Within, near, etc. based in real-time context
- Grouping & Aggregation
  - Accumulation of values or quantity
  - Sum, average, min, max, etc.
  - Support for custom aggregate functions
- Event Relationships
  - Event A followed by event B
  - Event A and event B
  - Event A or event B
  - The non-event
- Event Enrichment
- User-defined Functions
- Flexibility and ease to mix models
- Rules can be templated and parameters updated dynamically



CEP Market Landscape

Source: The Forrester Wave™: Complex Event Processing (CEP) Platforms, Q3 2009, Forrester Research, Inc., August 4, 2009

Acquisitions:

- Coral8 → Aleri → SAP
- Apama (Progress) → SoftwareAG
- StreamBase → TIBCO

Not listed:

- Microsoft StreamInsight
- Red Hat Drools Fusion
- Informatica (Agent Logic)

MARKETSANDMARKETS

"CEP market is estimated to grow from $764.5 million in 2013 to $3,322.0 million in 2018. This represents a Compounded Annual Growth Rate (CAGR) of 34.2% from 2013 to 2018."

CEP for Real-time Insights and Smarter Applications

- Capital markets trading
- Fraud detection
- Logistics management
- Customer experience mgmt
- Smart metering & smart grids
- Governance, risk & compliance
- Supply chain automation
- Industrial Internet
- Traffic management
- Transaction/log monitoring
- ...





software AG

APAMA

Real-time Analytics & Decisions

Product Overview

APAMA

Unified Productivity Tooling

Rich, intuitive development environment for developers and business analysts

The Correlator: Apama Run-Time Engine

Scalable, high-performance "Correlator"

Analytics, Patterns and Applications are deployed into the Correlator

Real-time output streams feed Apama's real-time, interactive, dashboards or any external system

Integration with Live and Static Data Systems

Cache integration

Integration with any data source/ sink or library

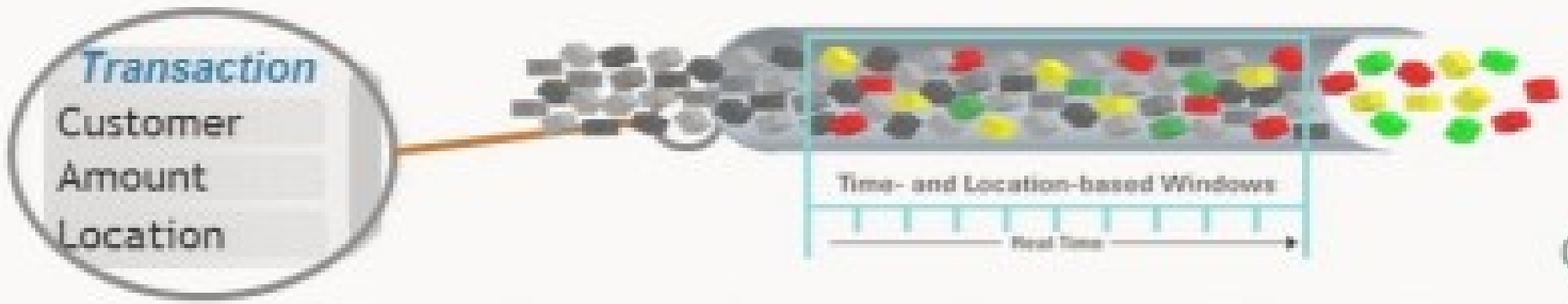
Programming the Correlator

- **The Correlator is the "container" for Apama applications**
  - Applications can be injected and removed dynamically at run-time without disrupting other running applications
- **Apama Event Processing Language (EPL)**
  - Declarative matching + Imperative processing
    - *Listeners* match, route and emit events
      - match event sequences including temporal constraints
      - dynamically set up / tear down by specifying event expressions
    - *Stream queries*
      - filter, aggregate and join temporal windows of events
    - *Actions* are EPL procedures - invoked by listeners, queries and actions
    - *Monitors* provide encapsulation
      - dynamically deployed / un-deployed
    - *Contexts* support parallel processing

- Java Plugin API to allow access to Java from EPL

Apama Uniquely Supports Complex Event Processing and Event Stream Processing

- **Complex Event Processing** - Temporal, logical and spatial attributes and relationships between events can represent business patterns, including emerging opportunities & threats

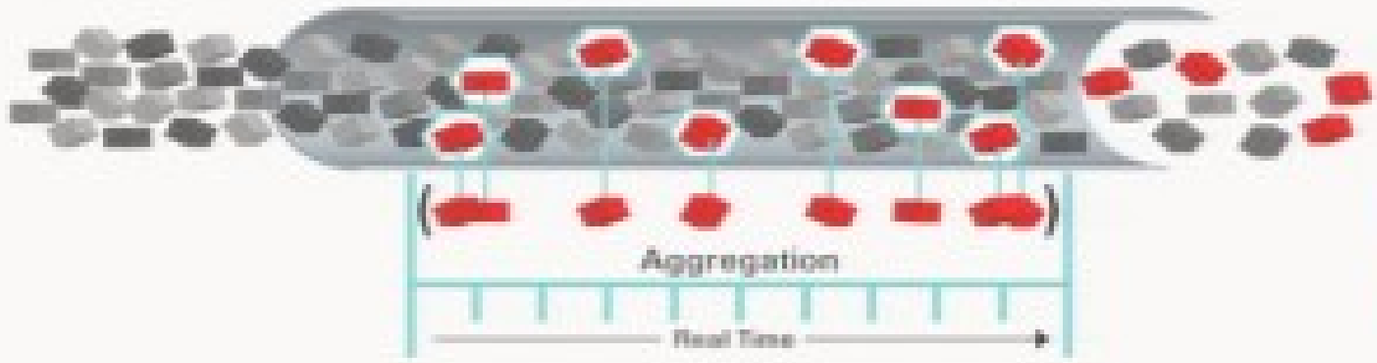


Transaction over $1,000 for a customer who then closes their account within a day


```
on Transaction (customer = ID, amount > 1000) followed-by
   Account (customer = ID, Action = 'Close') within (1 * DAY)
```


Apama Uniquely Supports Complex Event Processing and Event Stream Processing

- **Streaming Analytics** - Continuous re-calculations on a continuously moving window of events matching a particular query

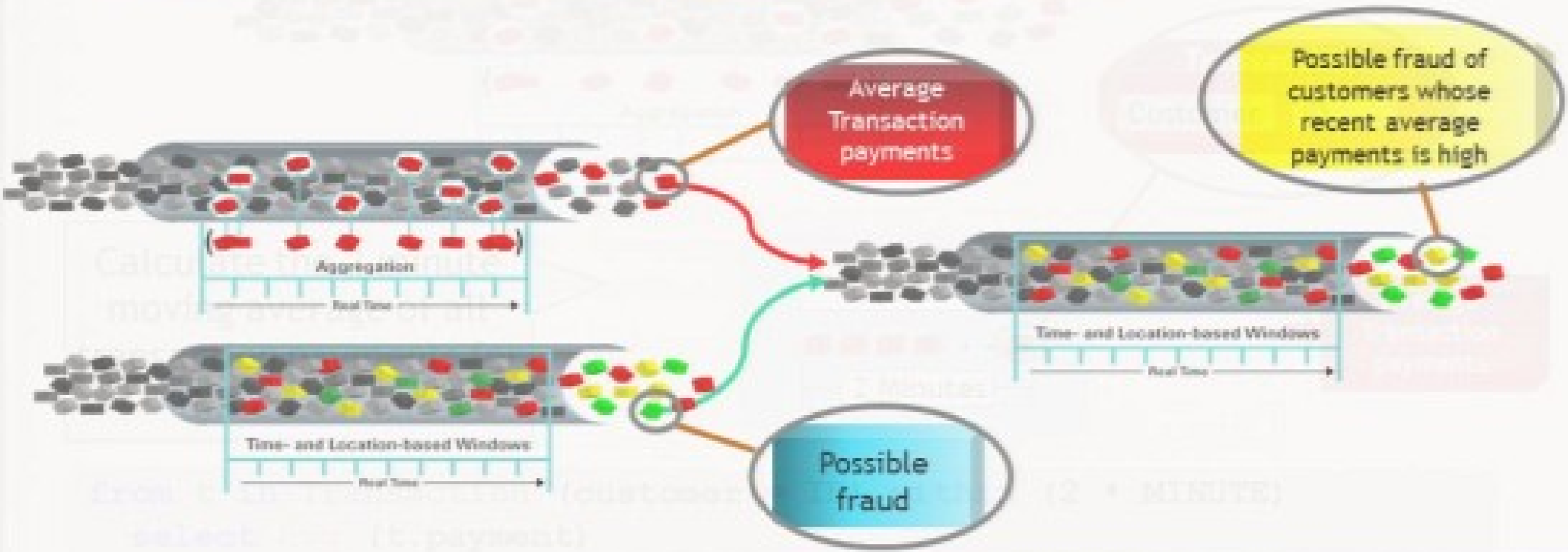


Calculate the 2 minute moving average of all transaction payments for this customer

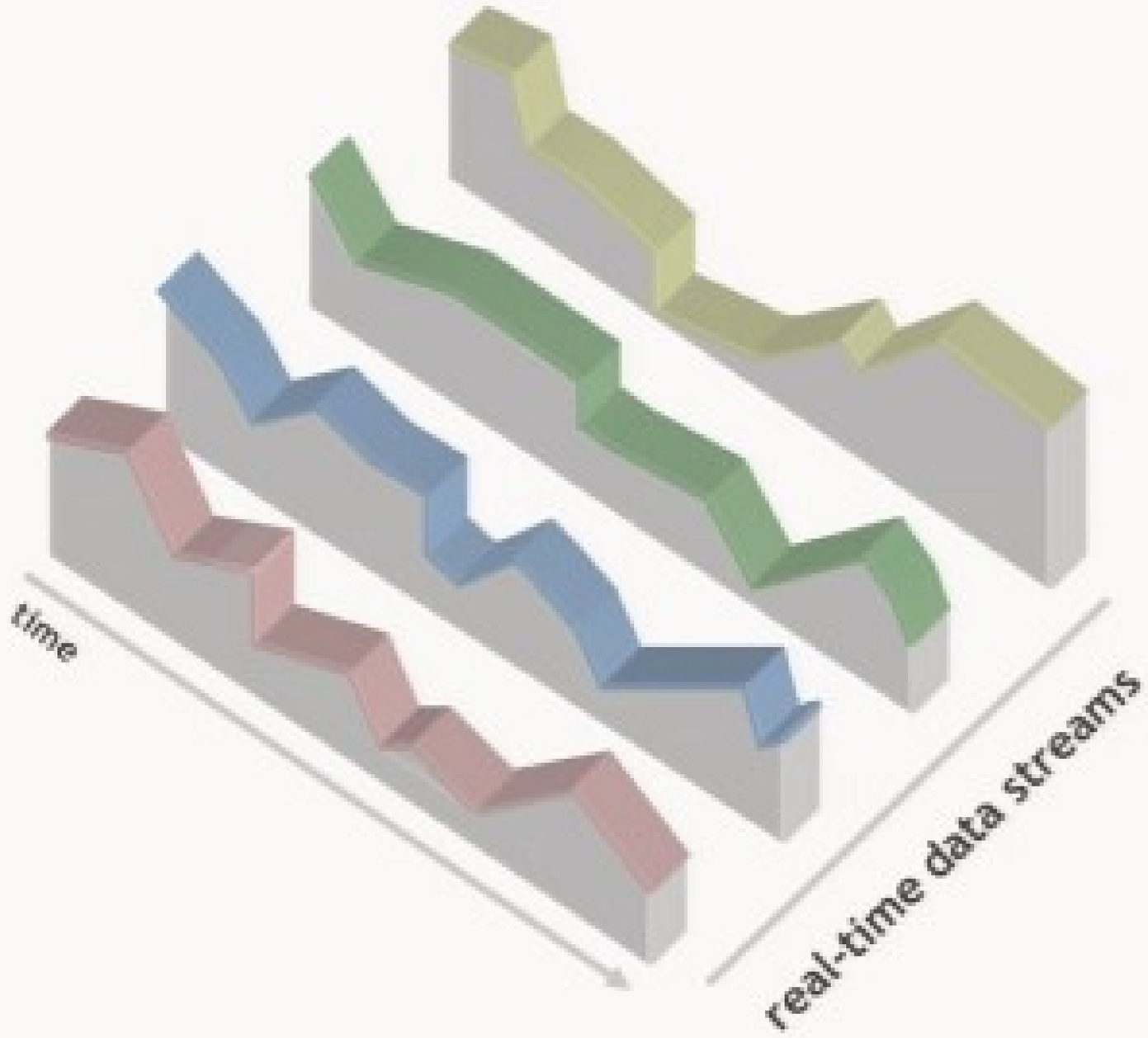

```
from t in Transaction (customer = ID) within (2 * MINUTE)
  select avg (t.payment)
```


Apama Uniquely Supports Complex Event Processing and Event Stream Processing

- **Dynamic Stream Networks** - Outputs from either Streaming Analytics or Complex Event Processing patterns can be fed into further streaming calculations or patterns.
  - The resultant network can changed dynamically: it need not be static

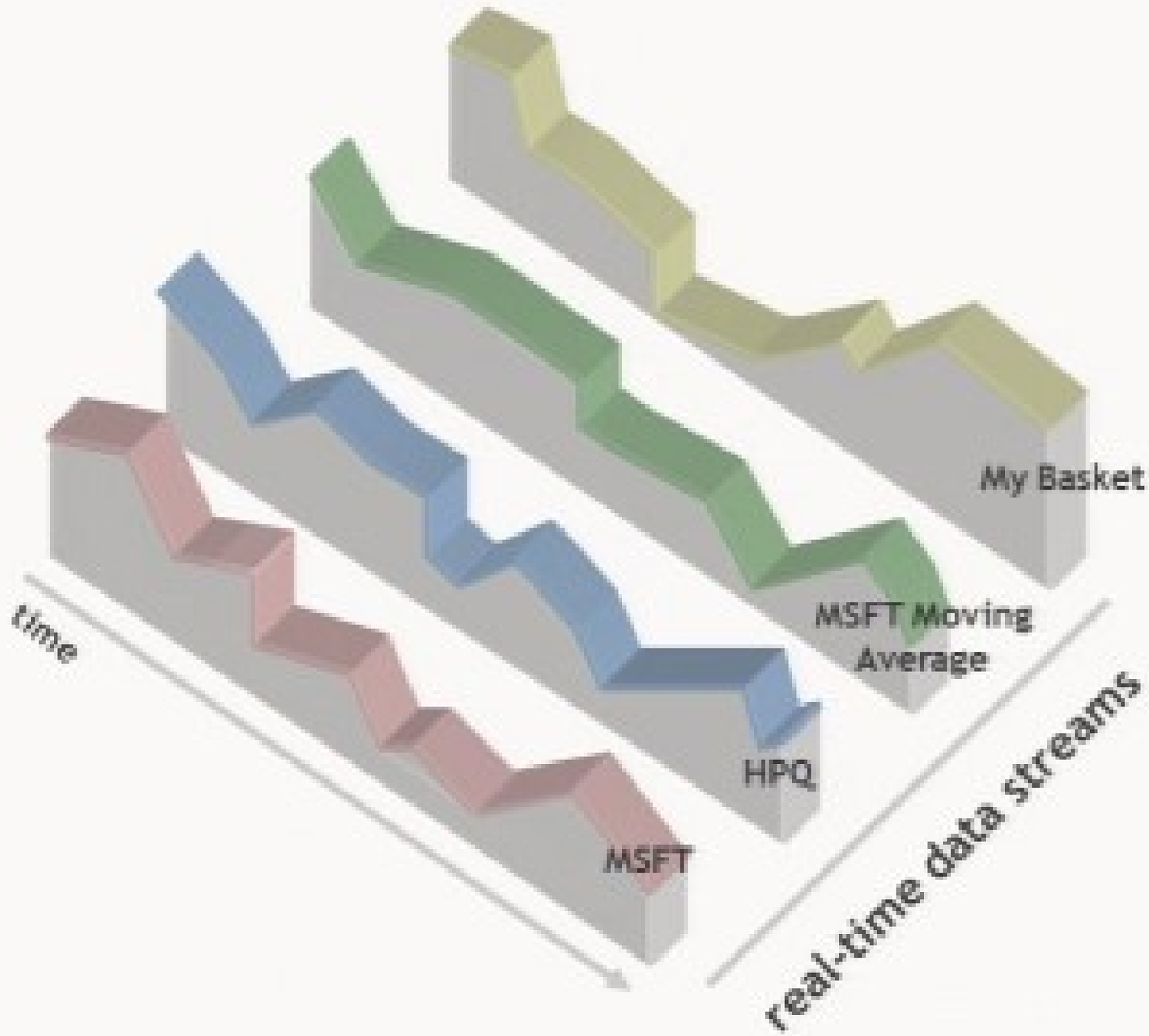


Example: Algorithmic Trading Rule



Example: Algorithmic Trading Rule

- multiple data streams

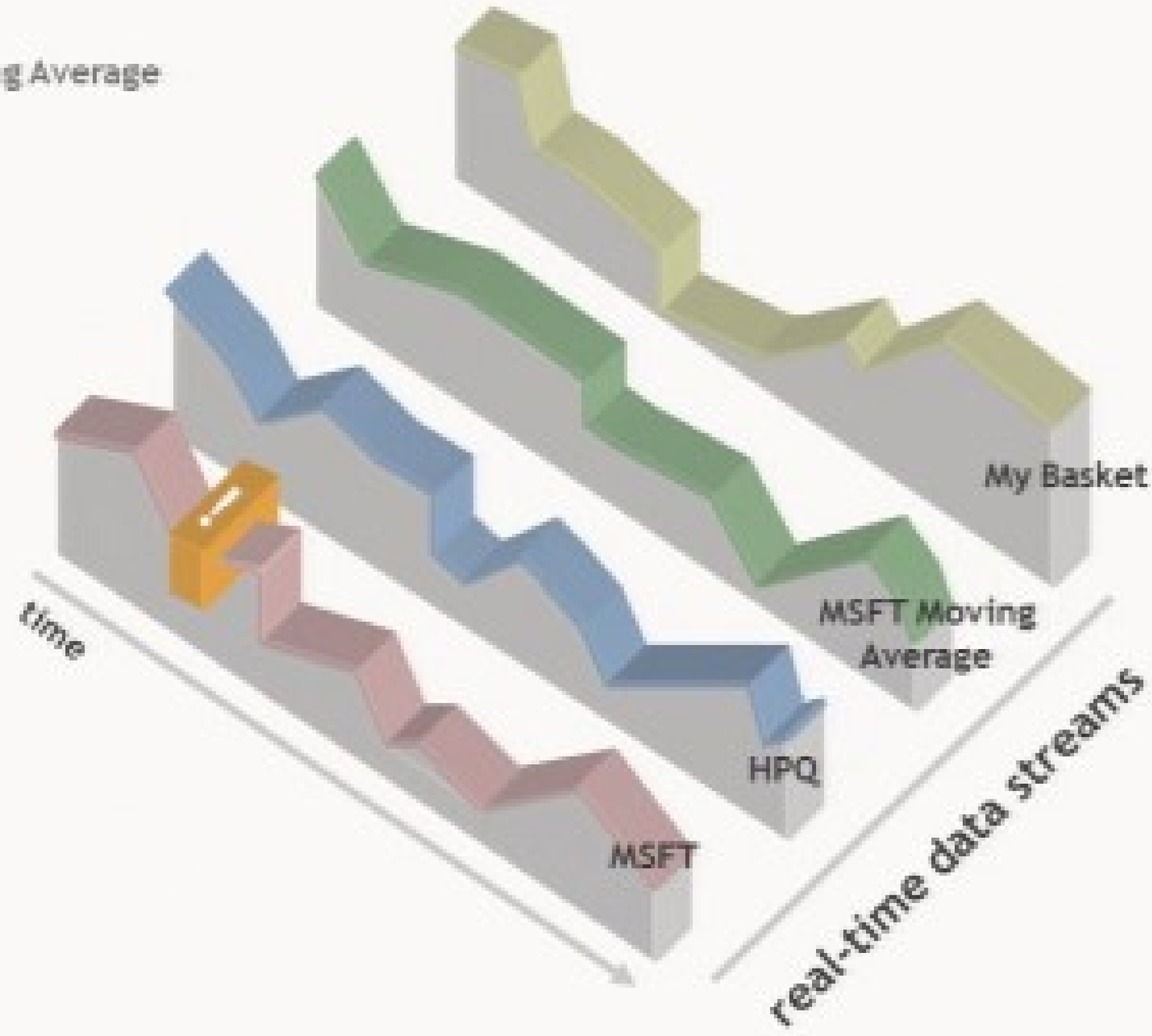


Example: Algorithmic Trading Rule

WHEN

MSFT price moves outside 2% of MSFT Moving Average

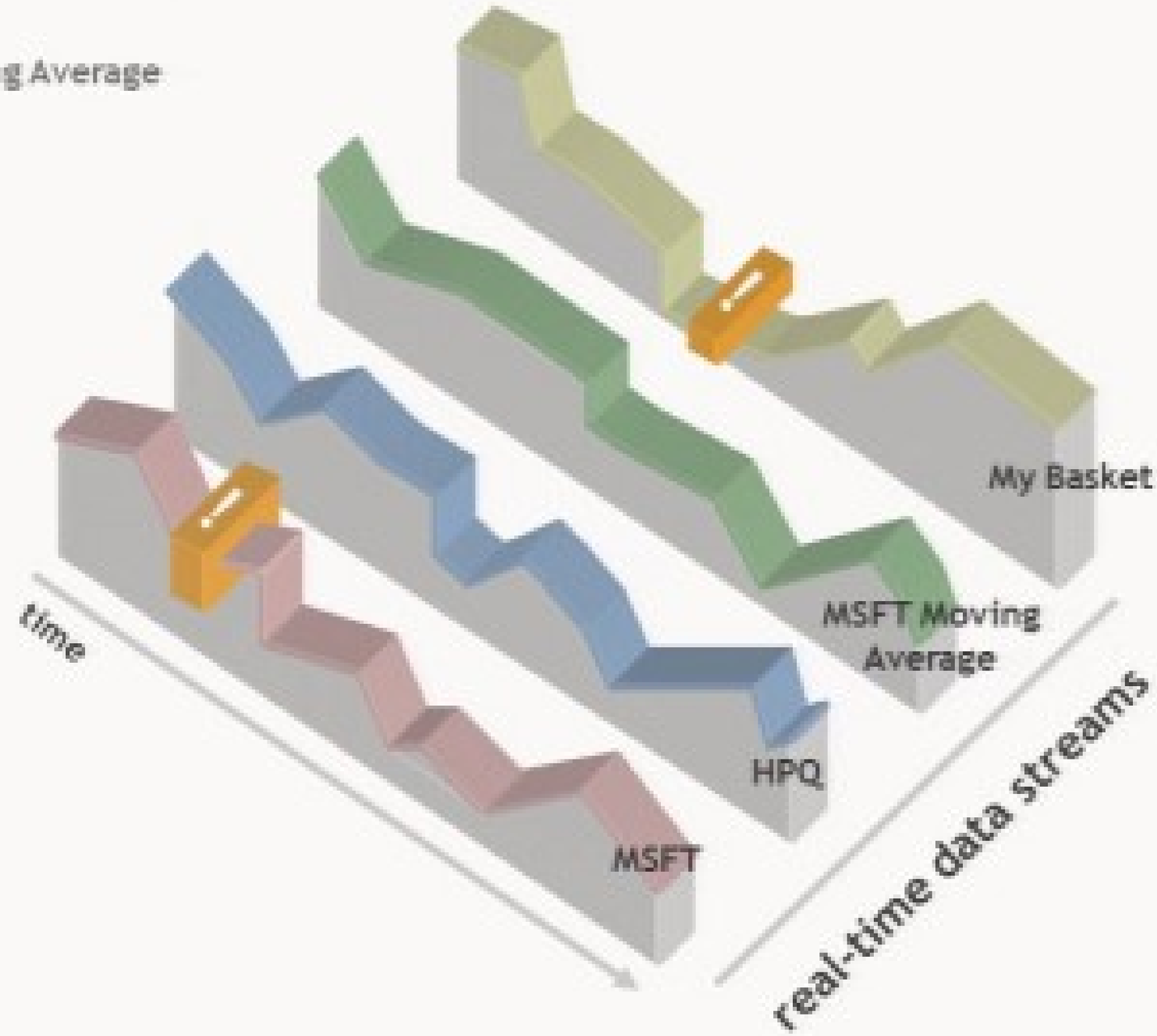
- multiple data streams



Example: Algorithmic Trading Rule

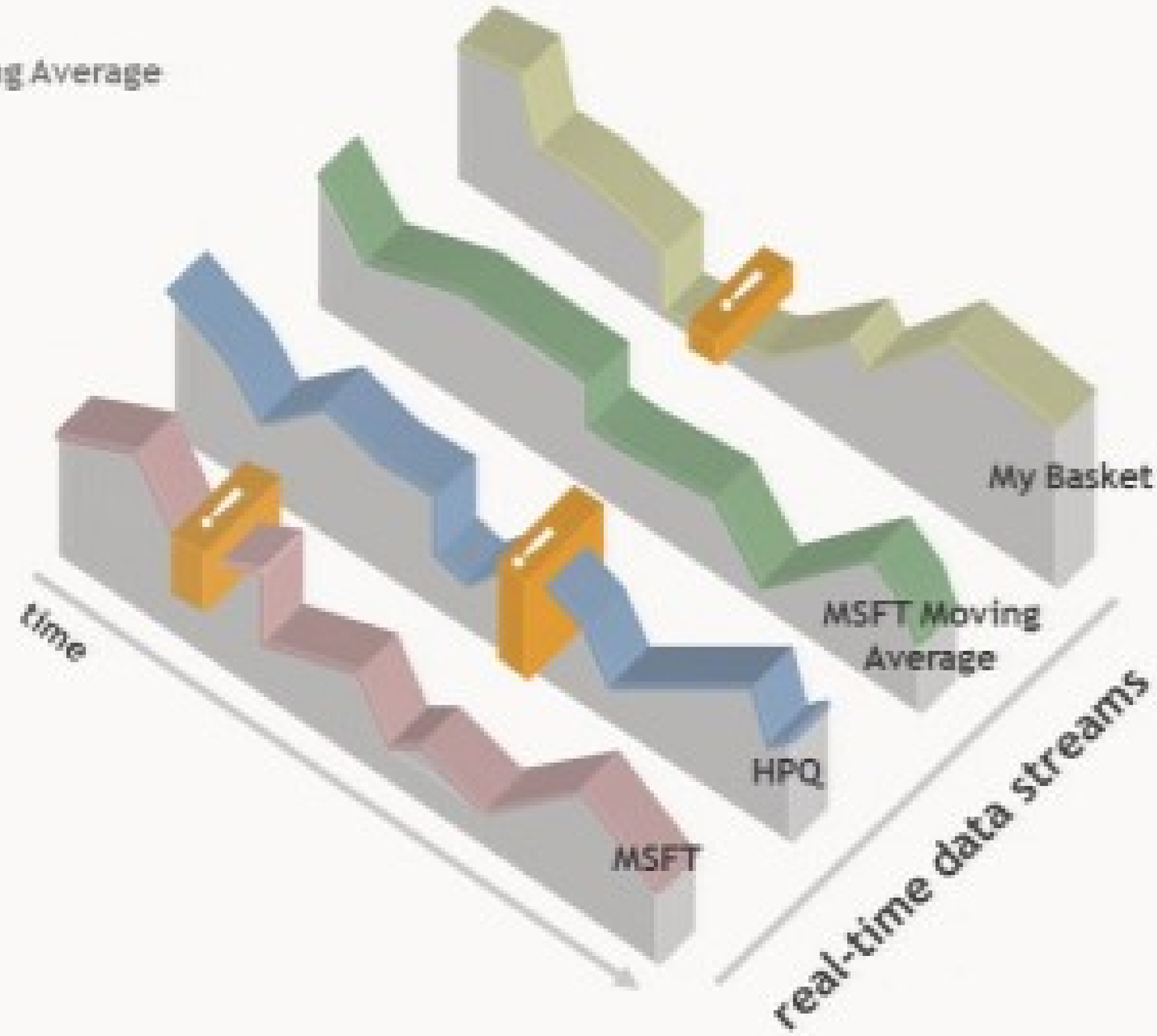
WHEN
  MSFT price moves outside 2% of MSFT Moving Average
  FOLLOWED-BY (
    My Basket moves up by 0.5%

- multiple data streams
- temporal sequencing



Example: Algorithmic Trading Rule

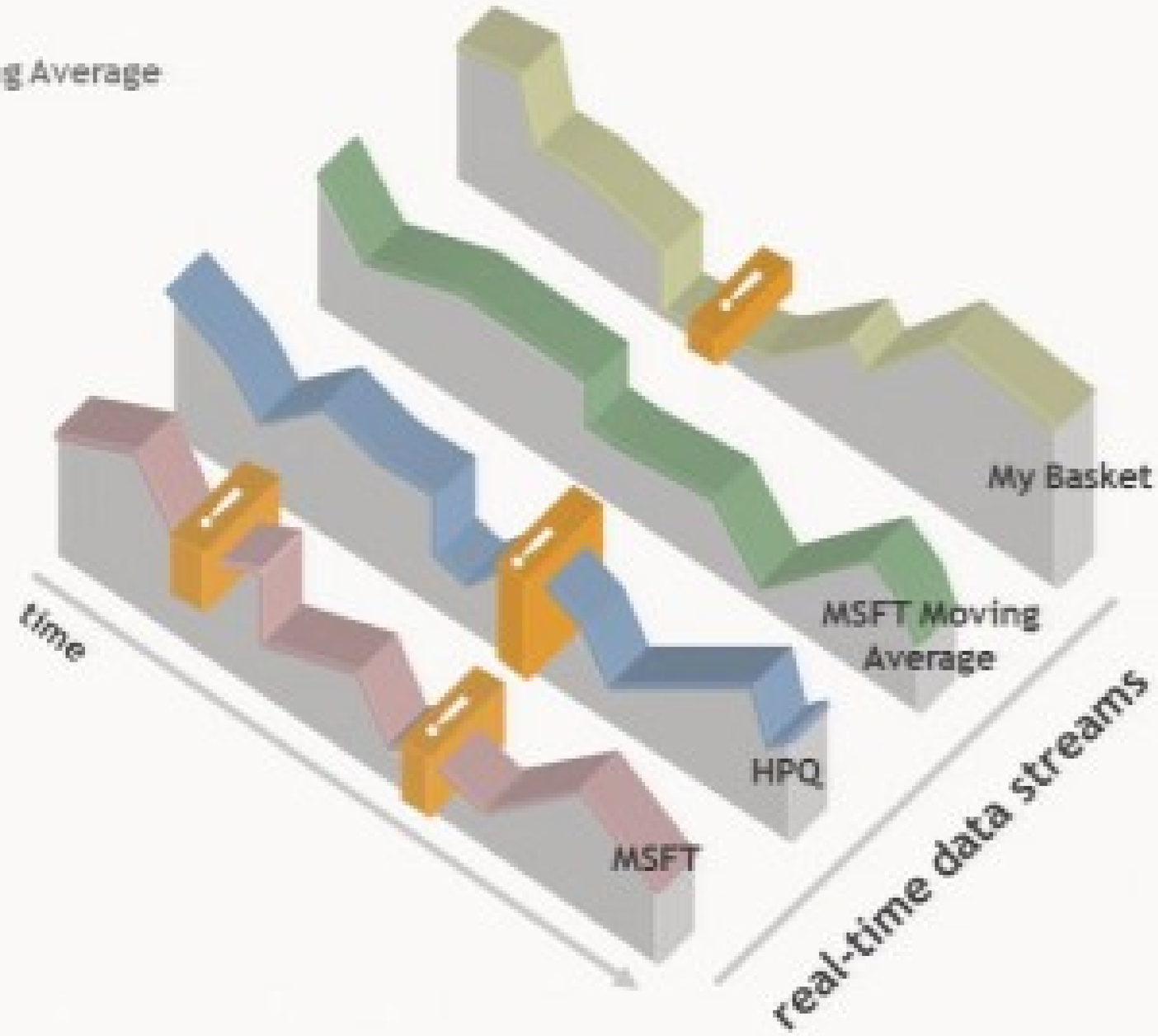
```
WHEN
  MSFT price moves outside 2% of MSFT Moving Average
  FOLLOWED-BY (
    My Basket moves up by 0.5%
    AND (
      HPQ's price moves up by 5%
```



- multiple data streams
- temporal sequencing
- complex event sequences

Example: Algorithmic Trading Rule

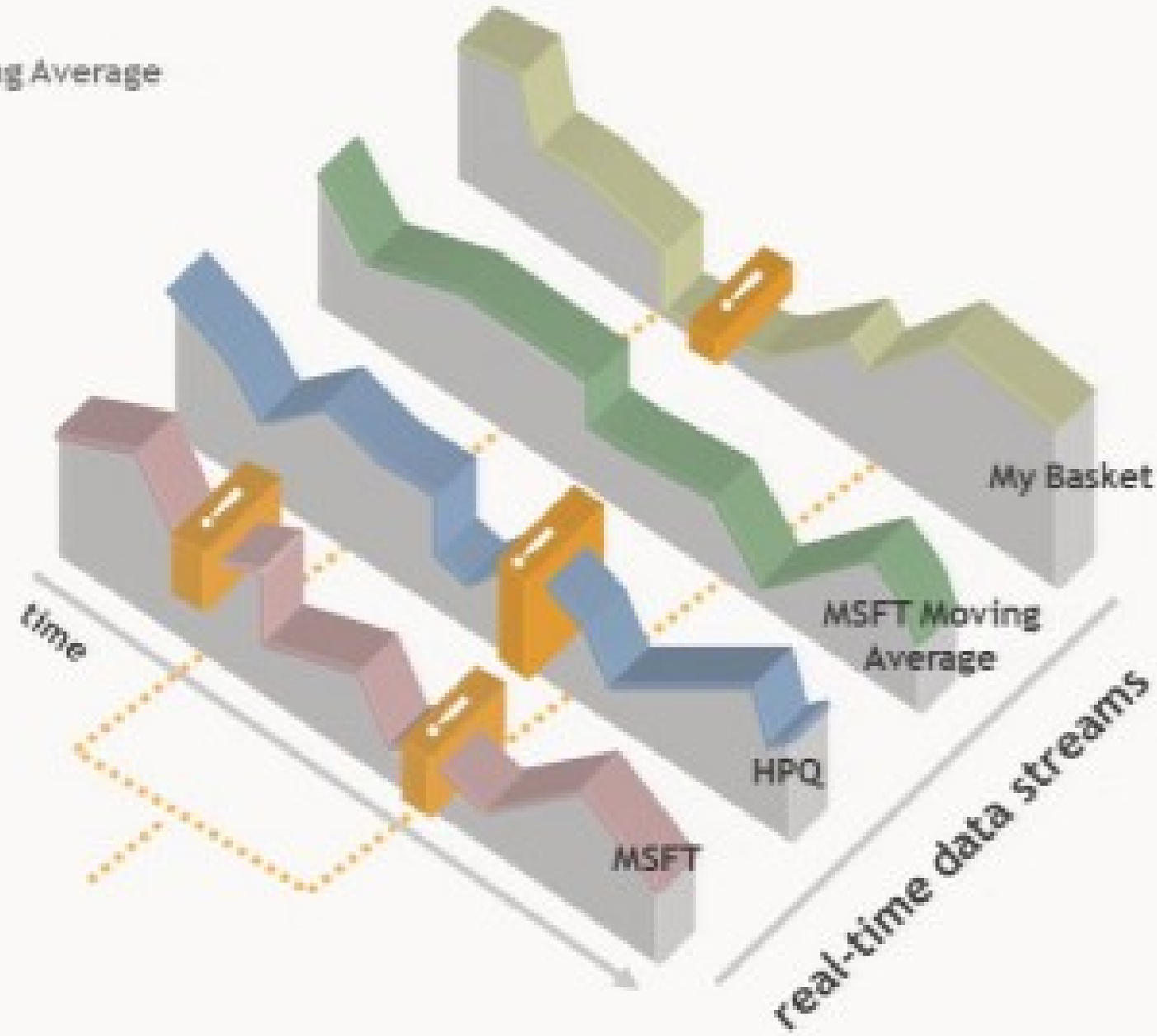
```
WHEN
  MSFT price moves outside 2% of MSFT Moving Average
  FOLLOWED-BY (
    My Basket moves up by 0.5%
    AND (
      HPQ's price moves up by 5%
      OR
      MSFT's price moves down by 2%
    )
  )
```



- multiple data streams
- temporal sequencing
- complex event sequences

Example: Algorithmic Trading Rule

```
WHEN
  MSFT price moves outside 2% of MSFT Moving Average
  FOLLOWED-BY (
    My Basket moves up by 0.5%
    AND (
      HPQ's price moves up by 5%
      OR
      MSFT's price moves down by 2%
    )
  )
ALL WITHIN
  any 2 minute time period
```



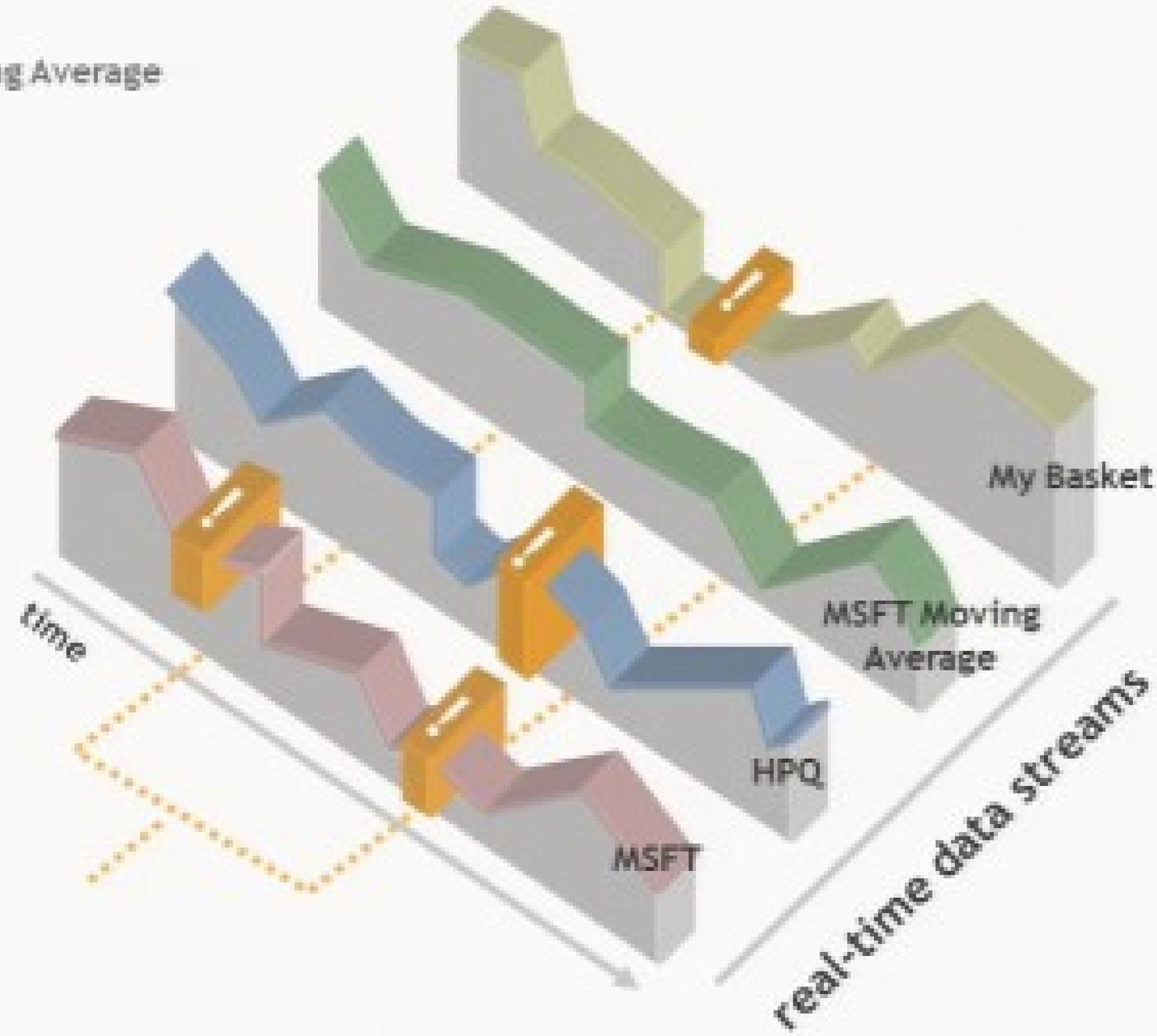
- multiple data streams
- temporal sequencing
- complex event sequences
- real-time constraints

Example: Algorithmic Trading Rule

```
WHEN
  MSFT price moves outside 2% of MSFT Moving Average
  FOLLOWED-BY (
    My Basket moves up by 0.5%
    AND (
      HPQ's price moves up by 5%
      OR
      MSFT's price moves down by 2%
    )
  )
ALL WITHIN
  any 2 minute time period

THEN
  BUY MSFT
  SELL HPQ
```

- multiple data streams
- temporal sequencing
- complex event sequences
- real-time constraints
- automated actions



Patented and Unique Apama Correlator

Logarithmic scaling of real-time event matching & **efficient processing** using partial evaluations

Micro- Scheduler maximizes use of CPU cores

Boolean, Temporal & Location logic operators

Hypertree

Temporal Sequencer

Stream Processor

State (In-memory & Cache)

Java VM

EPL (Machine Code)

In memory state for maximum performance; also **access to cache**

Dynamic windows of events for **streaming analytics**

EPL applications execute as **highly optimized machine code**

Write applications in **standard Java**

How the Correlator works

A -> B within timewindow

Simple event pattern:

Find event A followed by event B within 60 seconds

How the Correlator works

A -> B within timewindow

Account (...) -> Transaction (...) within 60 seconds

UserID = "2222"
Action = open

UserID = "2222"
Value > 1000

Using real events:

Detect when a user opens an account and then initiates a transaction of more then 1,000 within 60 seconds

How the Correlator works

A -> B within timewindow

Account (...) -> Transaction (...) within 60 seconds

UserID = "2222"
Action = open

UserID = "2222"
Value > 1000

Our simple pattern can be deconstructed into four parts and each considered separately to optimise processing

How the Correlator works

Initially we only ever look for a "open account" event – we do not need to do anything else yet

How the Correlator works

A -> B within timewindow

Account (...) -> Transaction (...) within 60 seconds

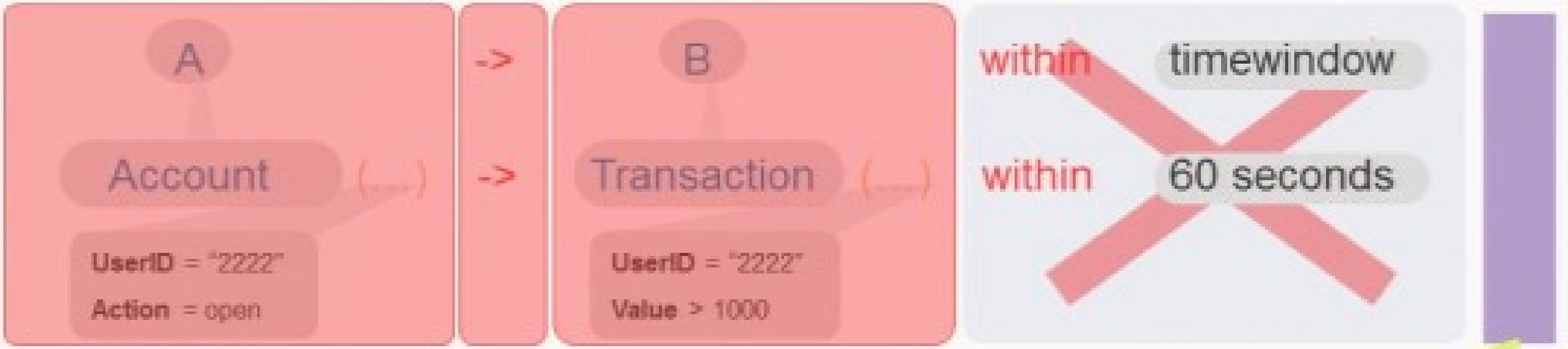
UserID = "2222"
Action = open

UserID = "2222"
Value > 1000

Now we look at the rest of the expression

We split this into two and we look for the transaction event while we start the 60 second timer

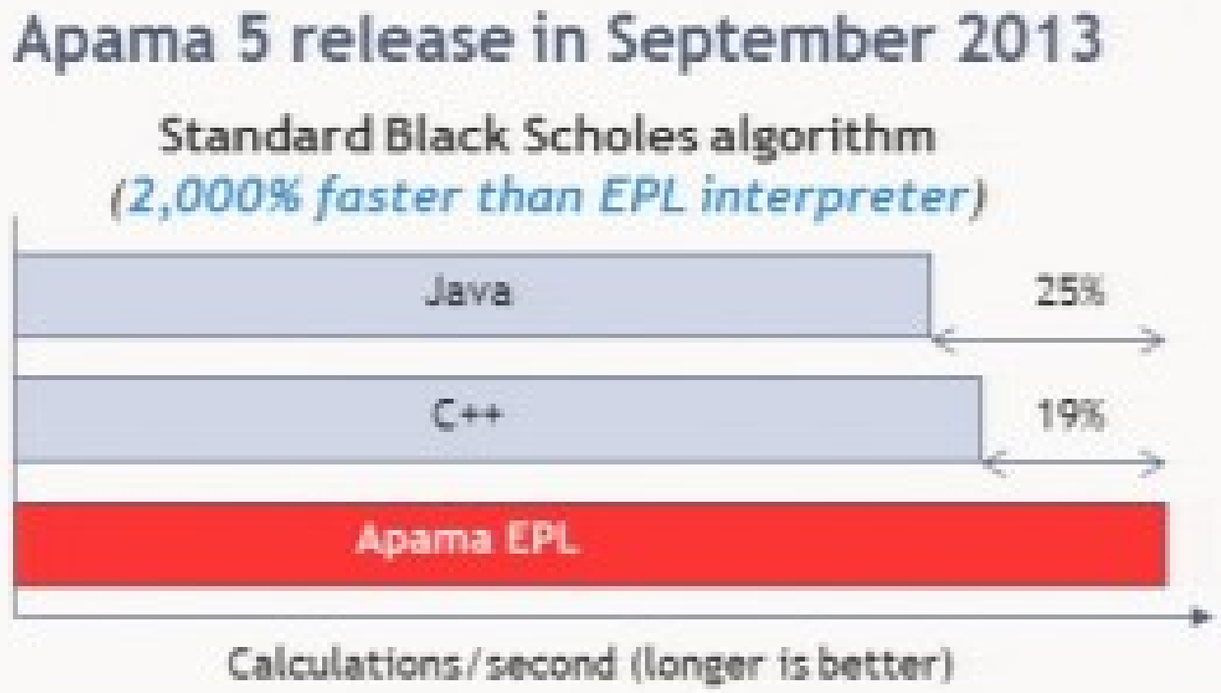
How the Correlator works



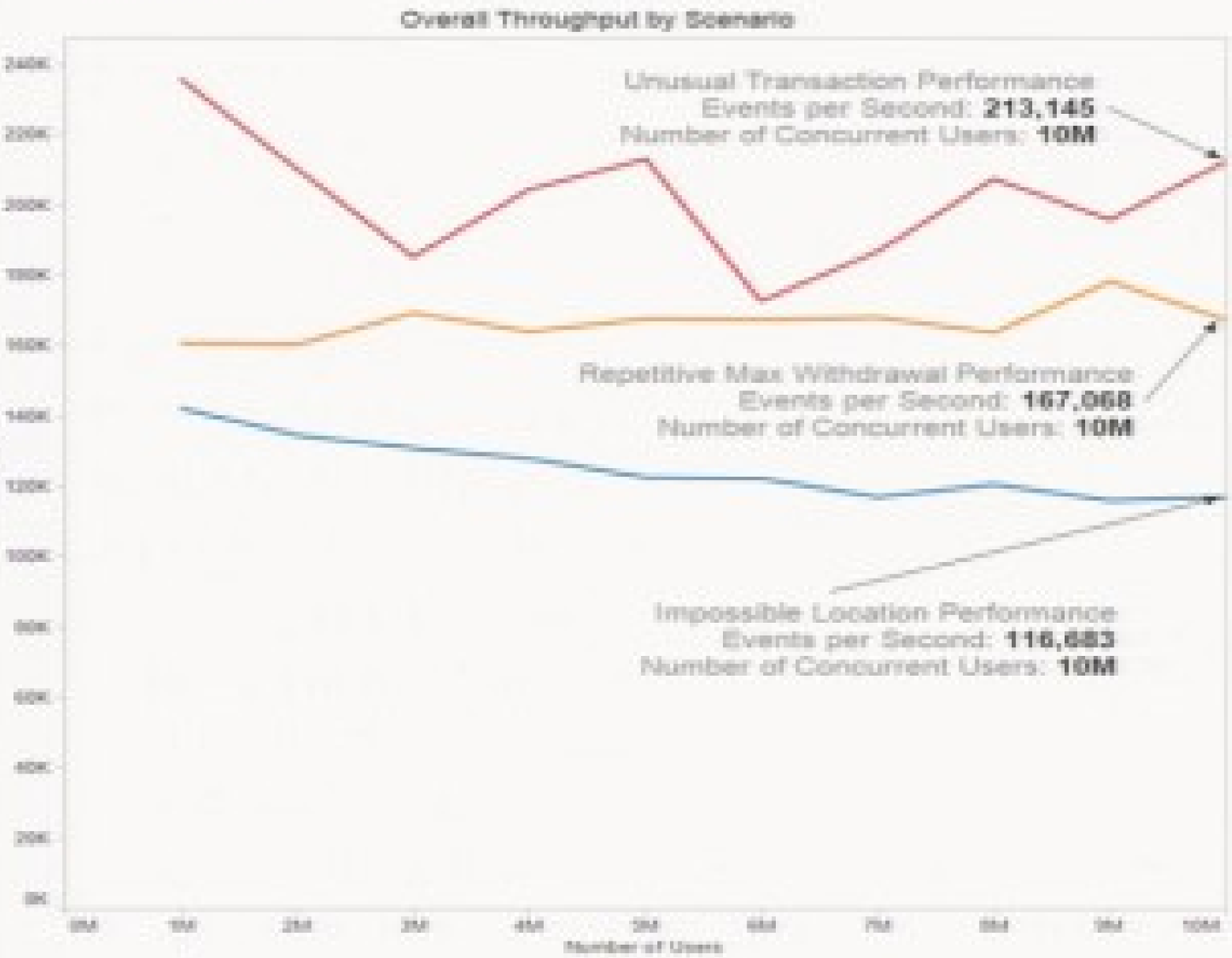
If we find the second event before we find the timer then we have found our pattern

Low-Level Virtual Machine (LLVM)

- Modern CPUs have optimized instruction sets
  - Only code complied for a specific CPU can make use of these instruction sets
- General purpose C and Java compilation is not aware of the exact CPU + Java has JVM overhead and garbage collection
  - Apama EPL is compiled at run-time to native machine code that is optimized for the exact CPU it executes on
  - Apama EPL executes more instructions per second than C or Java (and therefore any CEP engine written in C or Java)



Complex Fraud Benchmarks



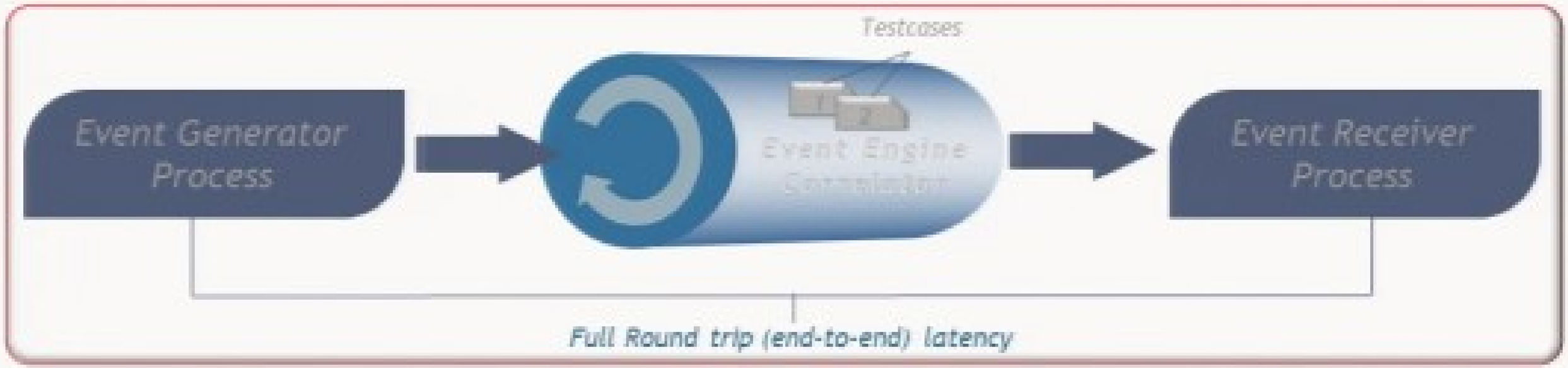
Two Intel Xeon X5570 CPUs clocked at 2.93GHz

- **Unusual Transaction**: identify unusual transactions based on the time of day and the channel
- **Repetitive Max Withdrawal**: Detect multiple cards being used on the same ATM machine each withdrawing large amounts
- **Impossible Location**: Detect when an ATM card is being used, when it was last used at a time & distance such that it could not be the same card

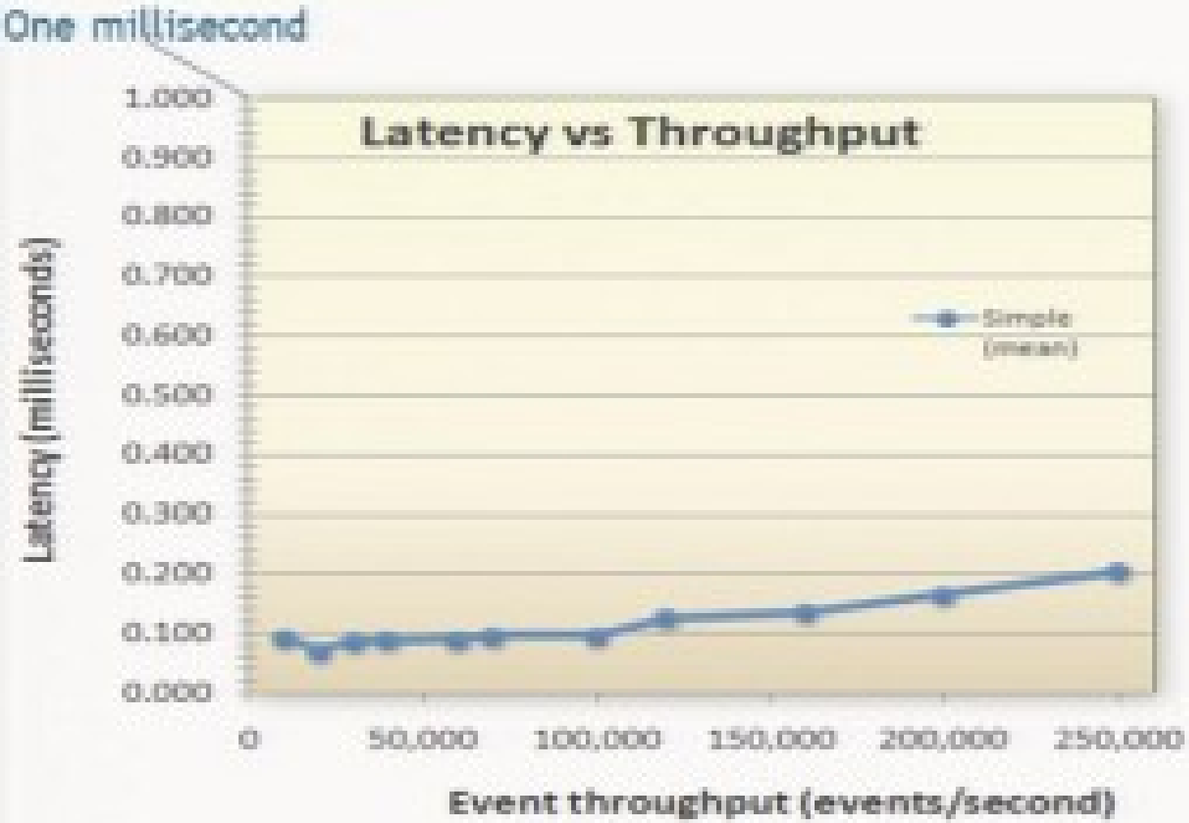
Performance test cases: Throughput & Latency

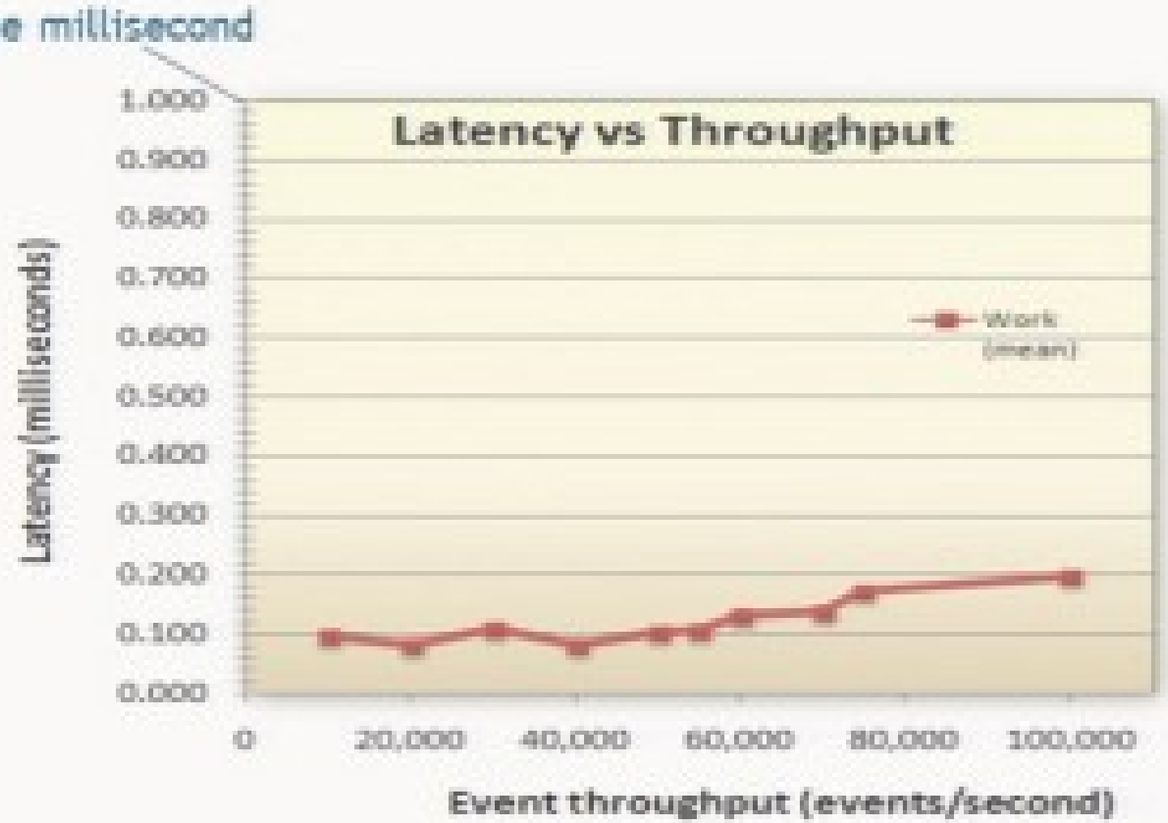
- Two test cases to demonstrate latency consistency with varying degrees of computation
  1. **Simple**: Respond to every event with little processing
  2. **Work**: Compute EWMA on price in every event, compare EWMA with price, compute weighted price and submit order
- Event rate increased until application is CPU bound

Host: Xeon X5570 (Nehalem), 2.93 GHz, Red Hat 5 64-Bit



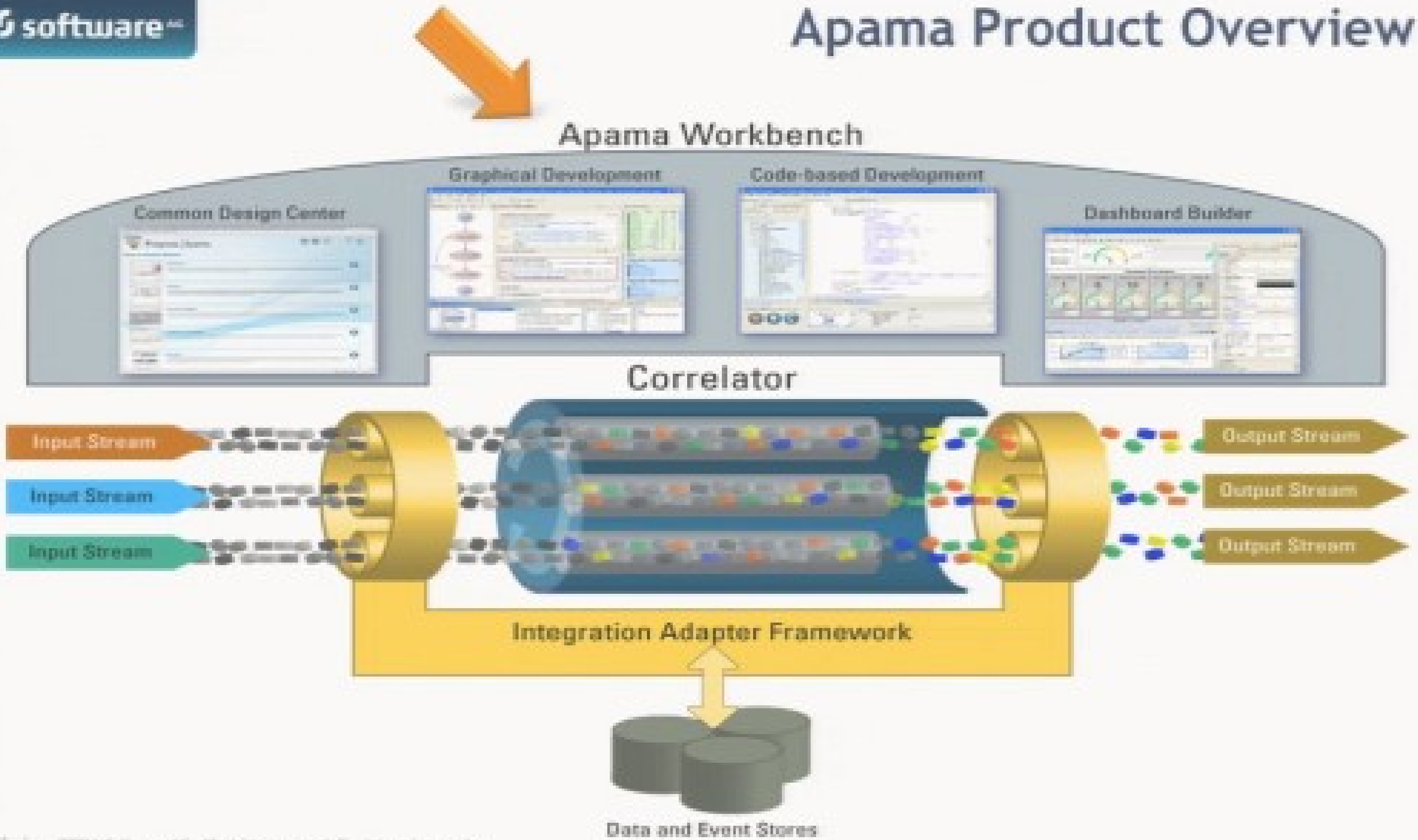
Performance results: Latency Behavior





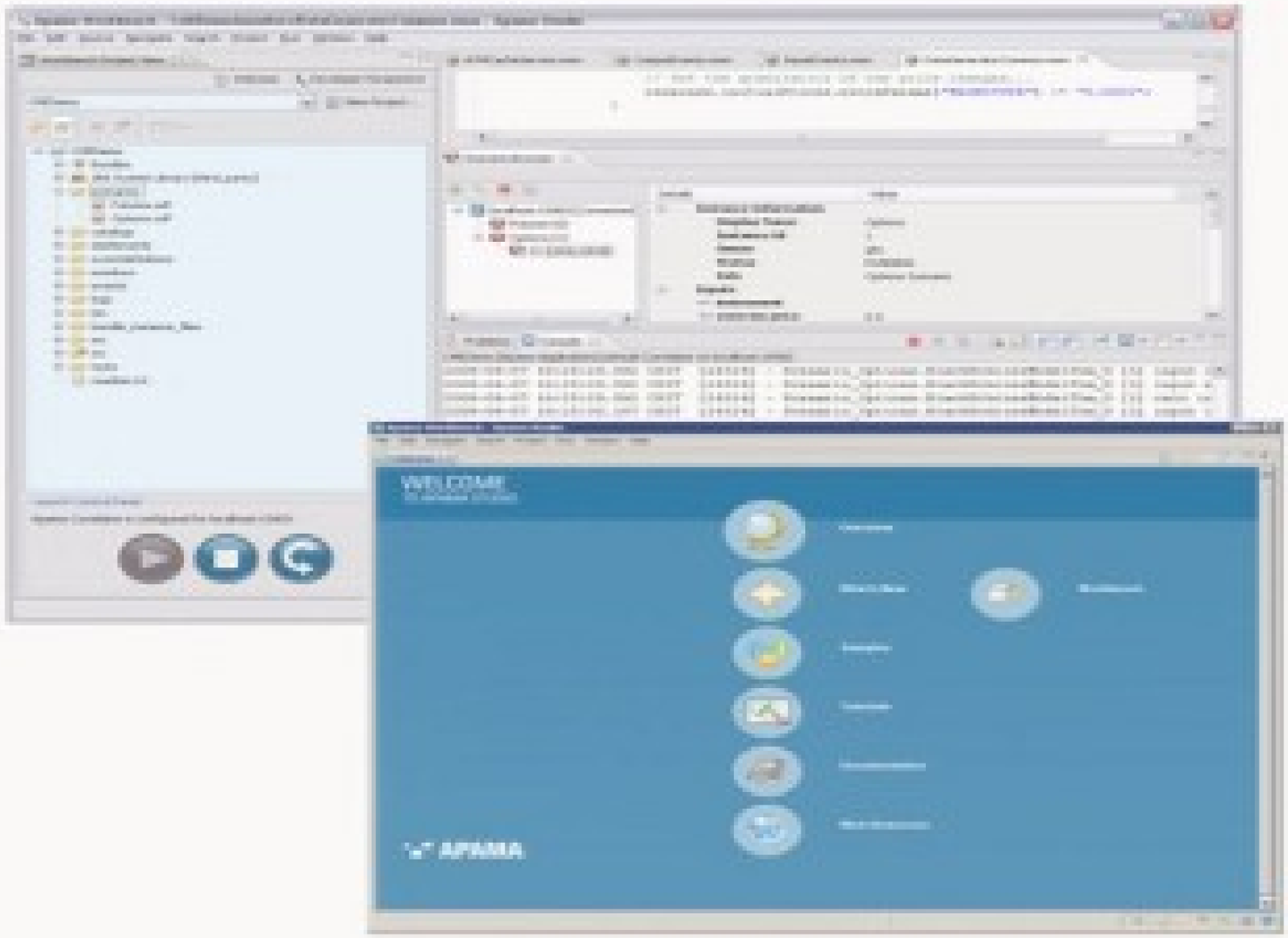
- Round-trip latency (to & from the Correlator process) is between 100 and 200 microseconds (0.1 - 0.2 milliseconds) and ***independent*** of application load
- Maintain low latency - 99th centile is also low latency

Apama Product Overview



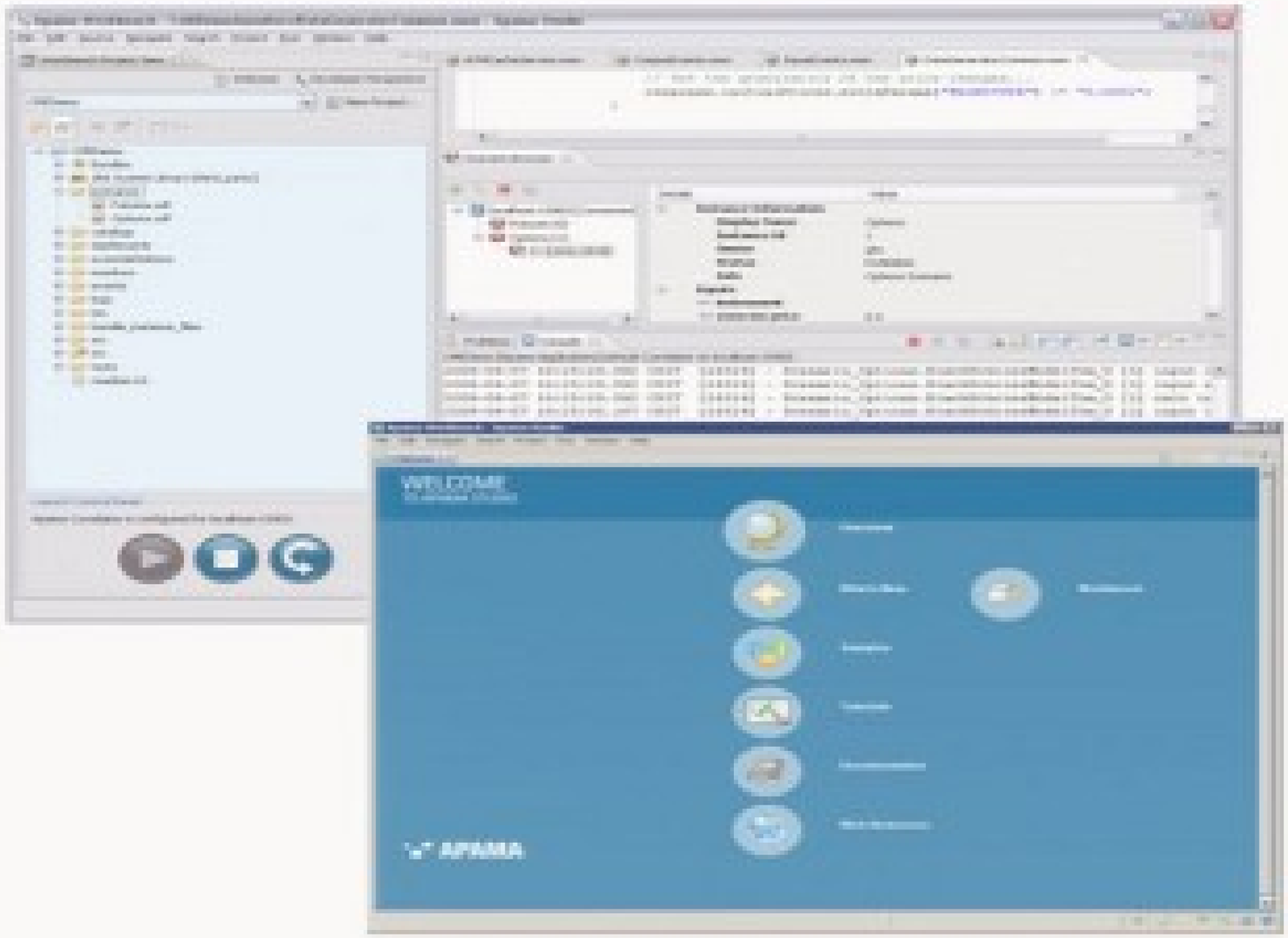
Apama Studio:
"Developer" tools - Traditional IDE

- Eclipse based
- Central development tool
  - Demos & tutorials
  - Project structure
  - Launch configurations
- MonitorScript development
  - Code assist editing
  - Source debugging
  - Application profiling
- Java development
  - Java CEP applications
  - Adapter development
  - Dashboard extension development
- Support & config tools
  - Backtesting
  - Dashboard generation & deploy
  - Adapter configuration
  - Ant export



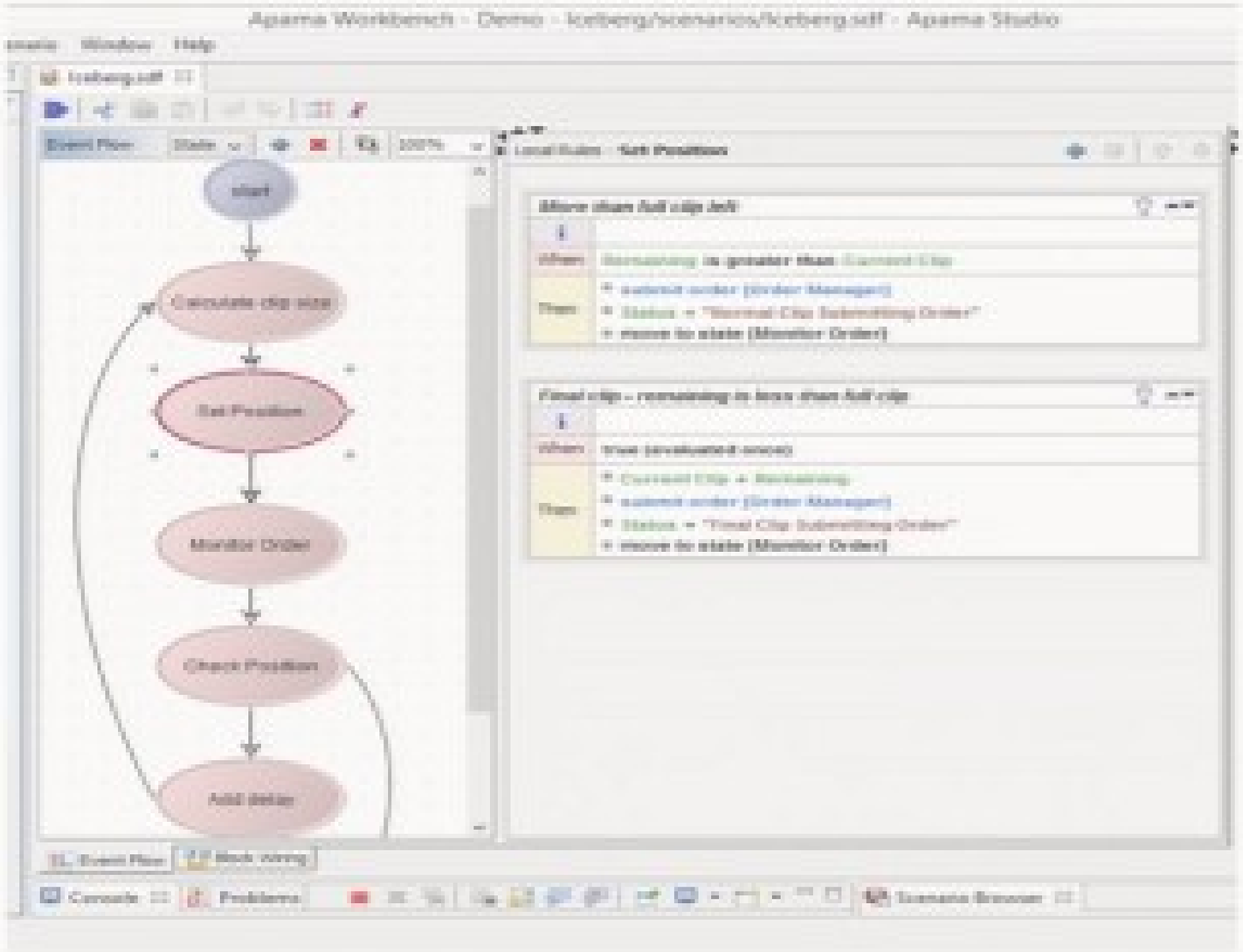
Apama Studio:
"Developer" tools - Traditional IDE

- Eclipse based
- Central development tool
  - Demos & tutorials
  - Project structure
  - Launch configurations
- MonitorScript development
  - Code assist editing
  - Source debugging
  - Application profiling
- Java development
  - Java CEP applications
  - Adapter development
  - Dashboard extension development
- Support & config tools
  - Backtesting
  - Dashboard generation & deploy
  - Adapter configuration
  - Ant export



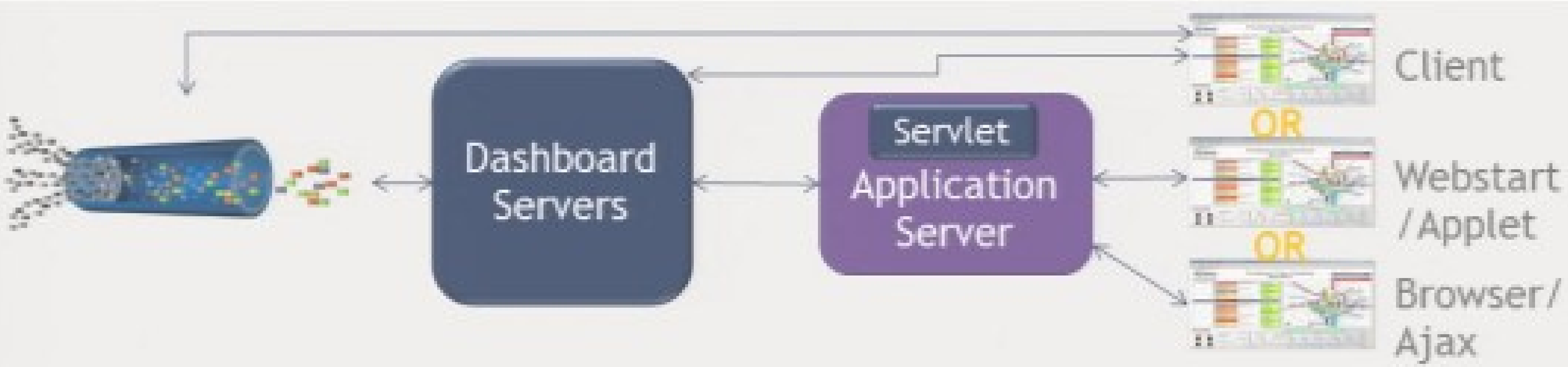
Apama Studio:
"Business Analyst" tools

- Swing based
- High-level development tool
  - Targeted at analysts, traders, …
  - All GUI driven (no programming)
- Create application "scenarios"
  - Scenarios are templates
  - Create instances at runtime
- State transition diagram
  - States of application
  - Visualize flow
- Rules
  - Conditions and actions
  - Define state transitions
- Blocks & block wiring
  - Reusable components
  - Inputs/outputs/transforms/metrics
  - Wire together to form "pipelines"



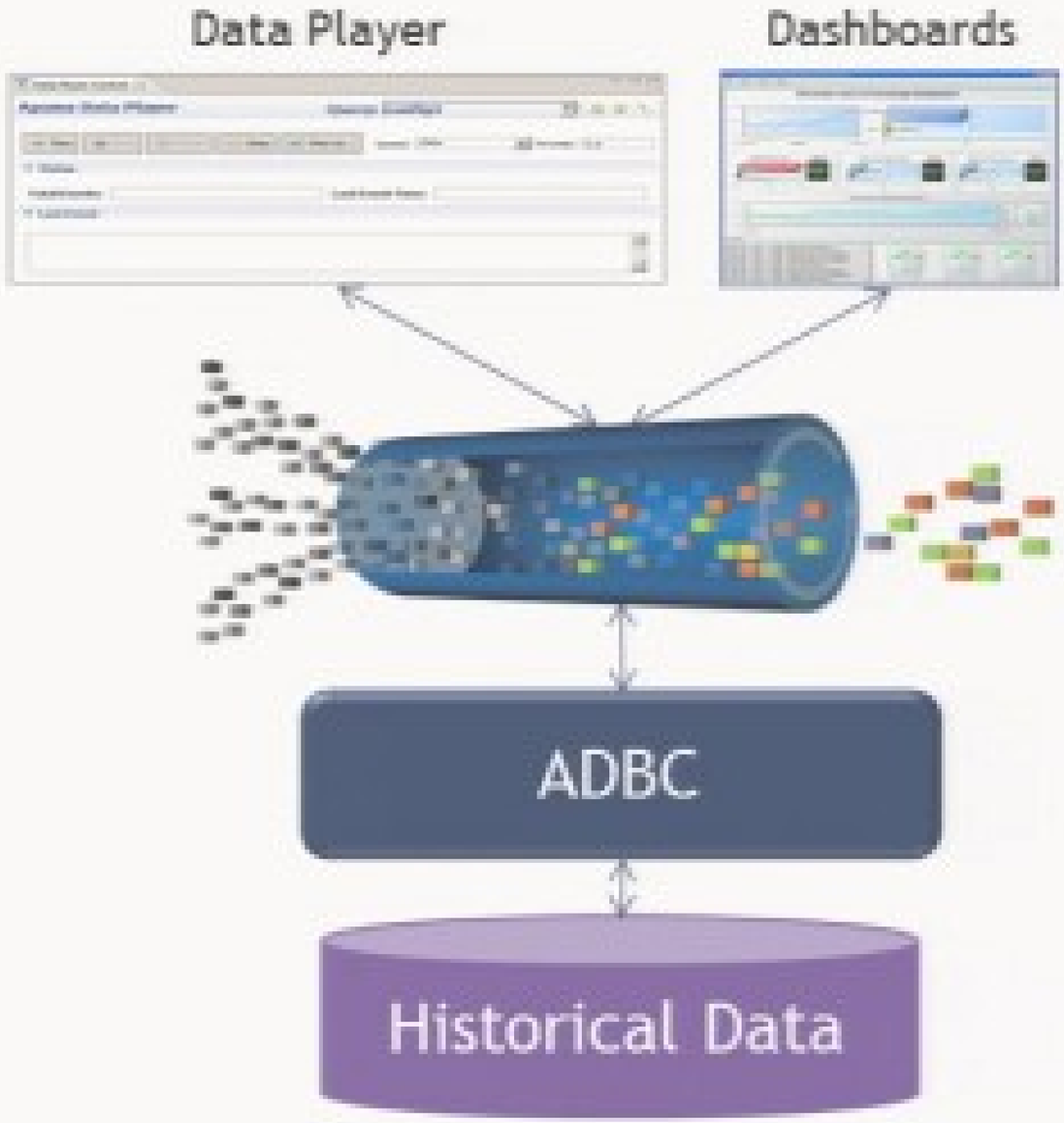
Apama Dashboards

- **Easily build UIs** to interact and monitor with applications running within Apama Correlators
  - Forms, tables, charts, dials, meters, …
- Provide **multiple deployment options** including thick client, webstart, applet, and thin client
- Multi-tier architecture for scalability and security
- Dashboards servers sit between users and Correlators

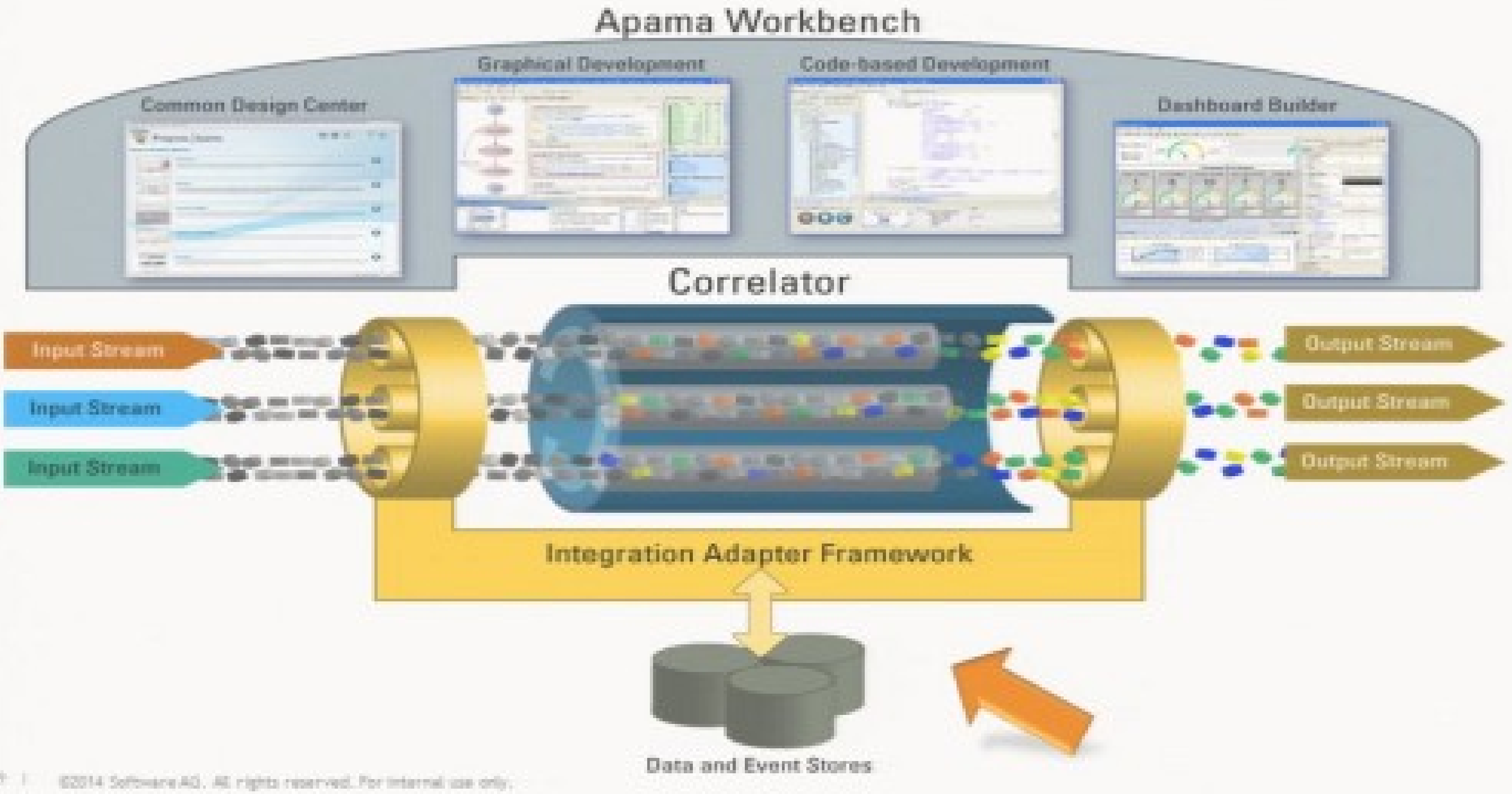


Backtesting Applications

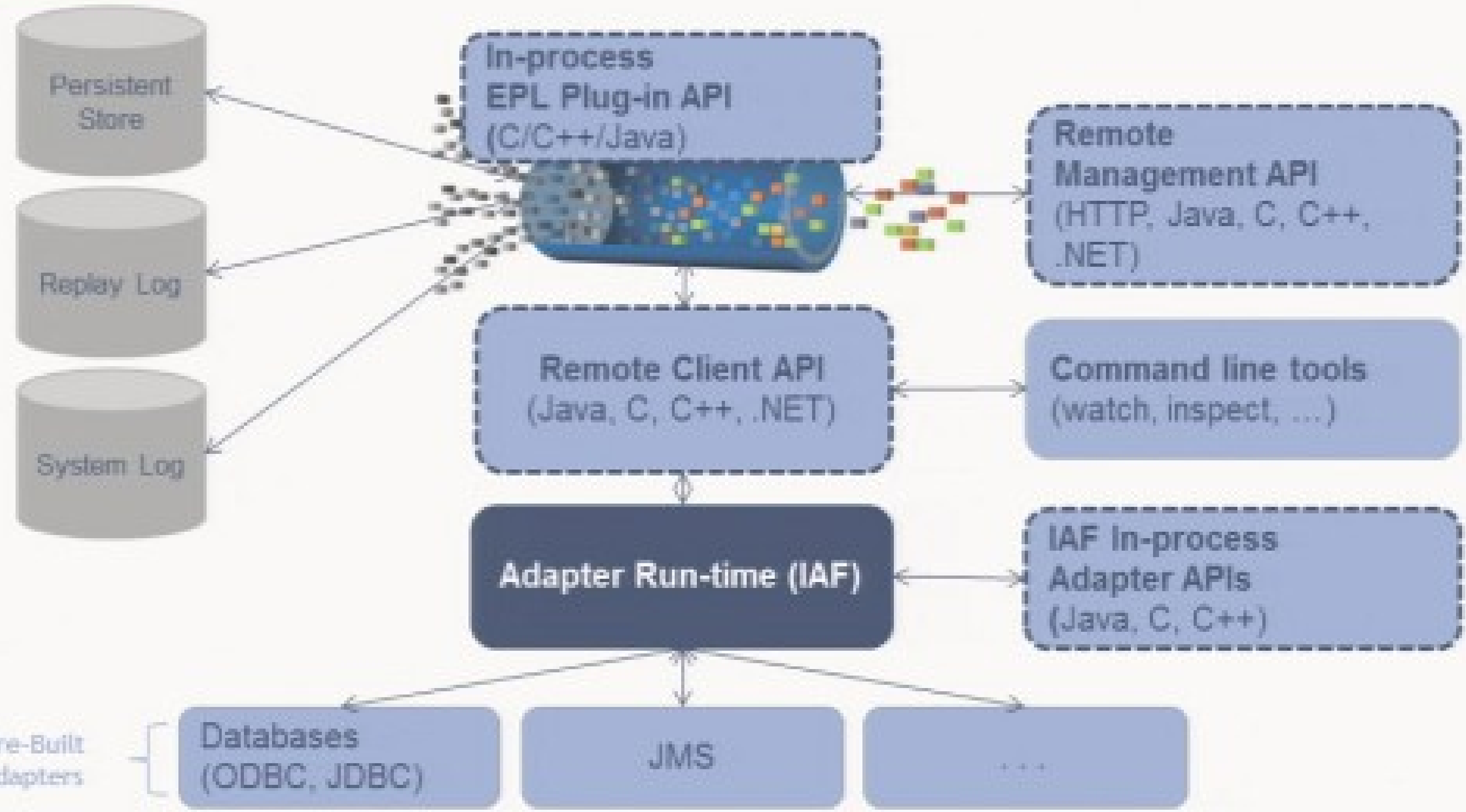
- **Backtesting** is integrated within Apama Studio
  - "What if analysis"
  - "Does an algo generate more $"
  - "Is an app responding correctly"
- **Event playback** via ADBC
  - Supports JDBC/ODBC/TXT
  - Extensible to other sources
- User controls within Data Player
  - Start/stop/pause/…
  - Playback speed
- Exploits Correlator "Clocking" capability
  - Guaranteed results regardless of replay event rate



Apama Product Overview



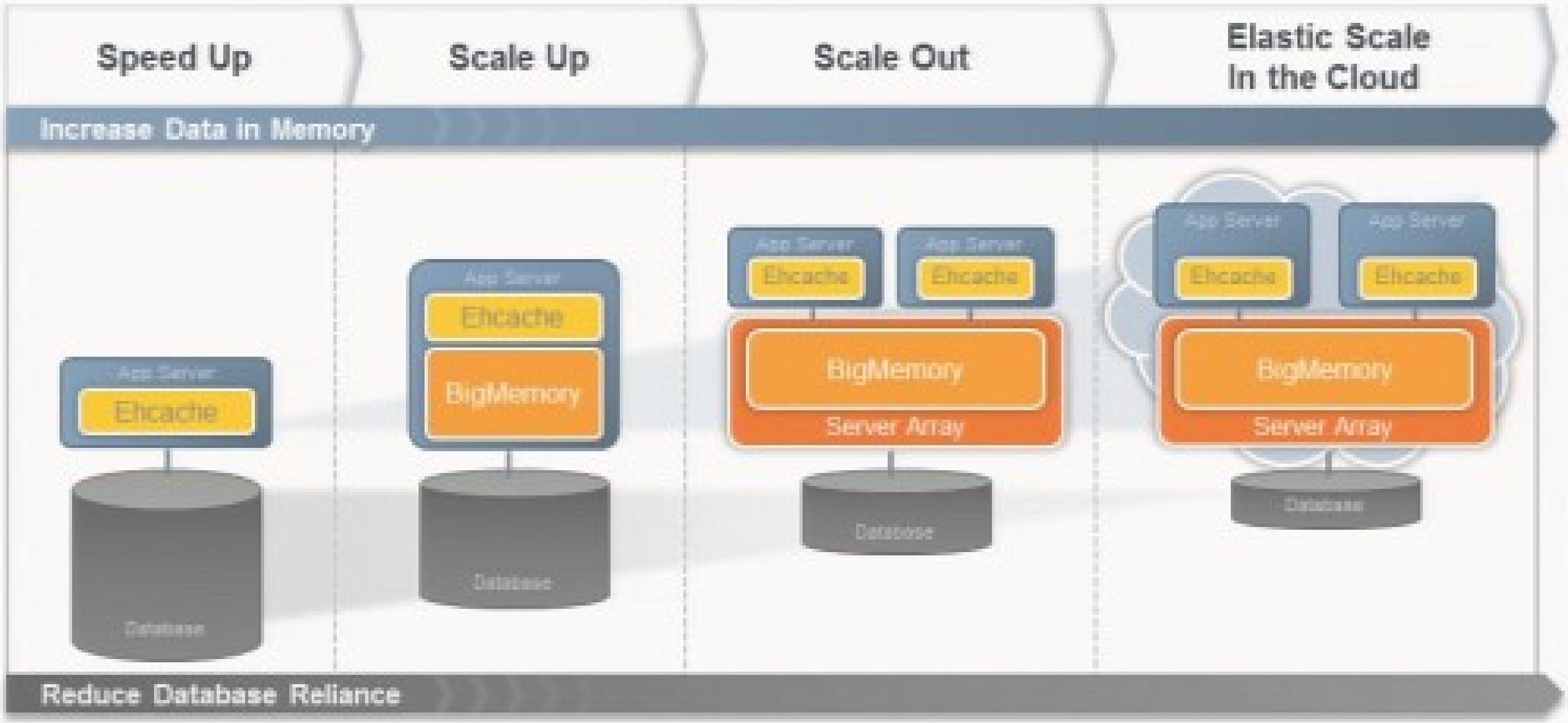
Apama Connectivity



software AG

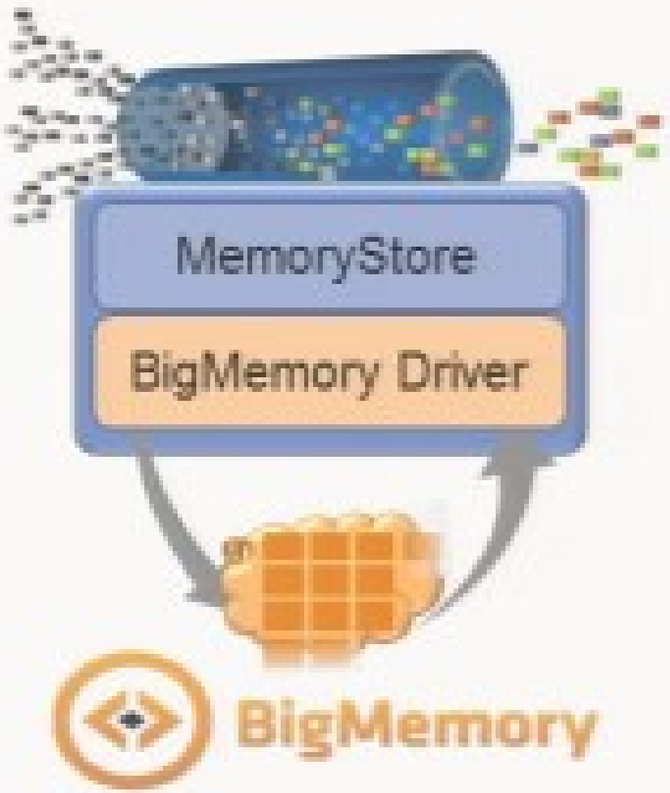
TERRACOTTA

BigMemory: In-Memory Data Management



Data in Motion & Data At Rest

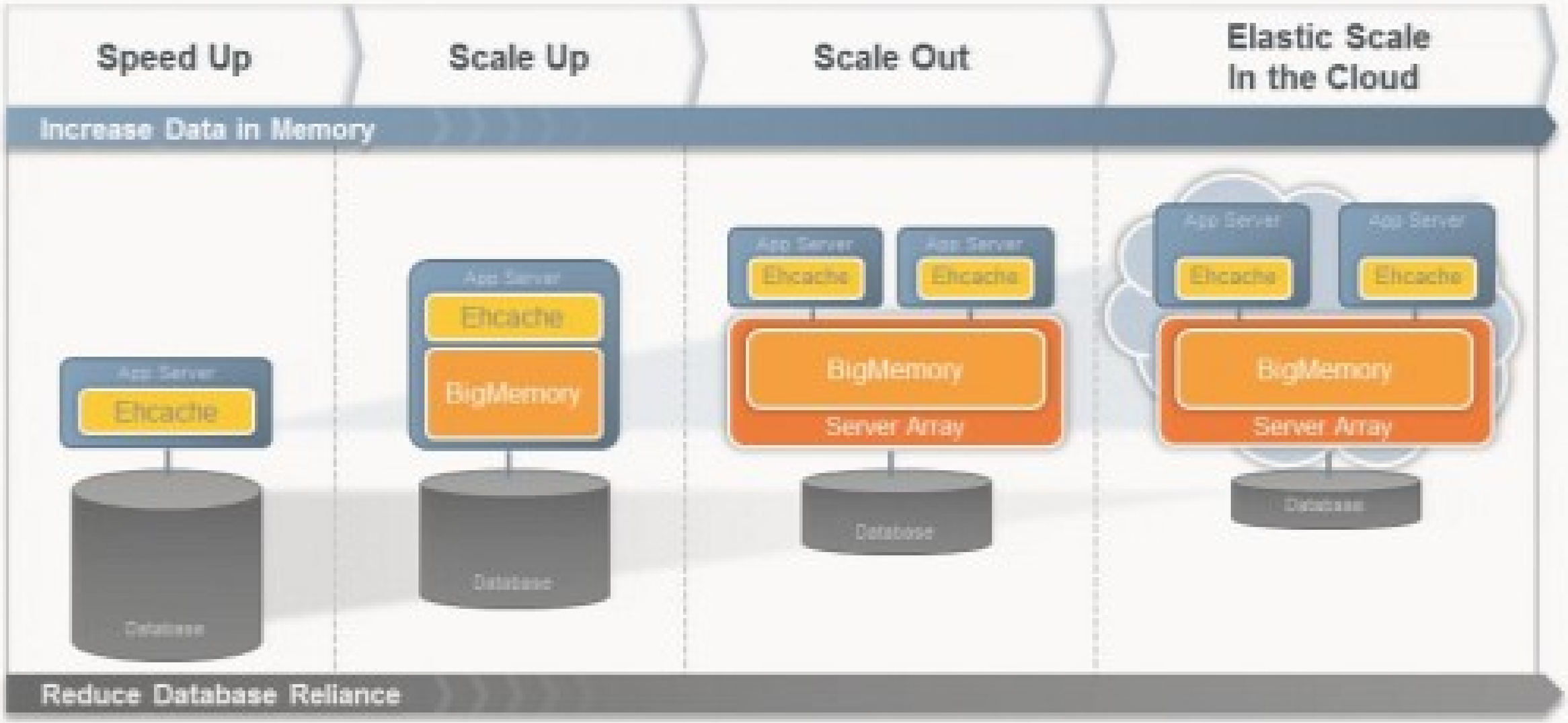
- Apama MemoryStore can be backed by BigMemory
  - In addition to databases and other apps
  - Transparent to Apama developers (simply new mode to MemoryStore)
- Provides "infinite" memory access from within Apama
  - Delivered as "Driver" to Apama "Cache" API
  - Inserts, Edits, Deletes & Notifications allow multiple clients (inc. Apama Correlators) to share data
  - Optionally exposes 'search attributes' for Apama MemoryStore cache entries to enables external clients (such as Presto) to query data



software AG

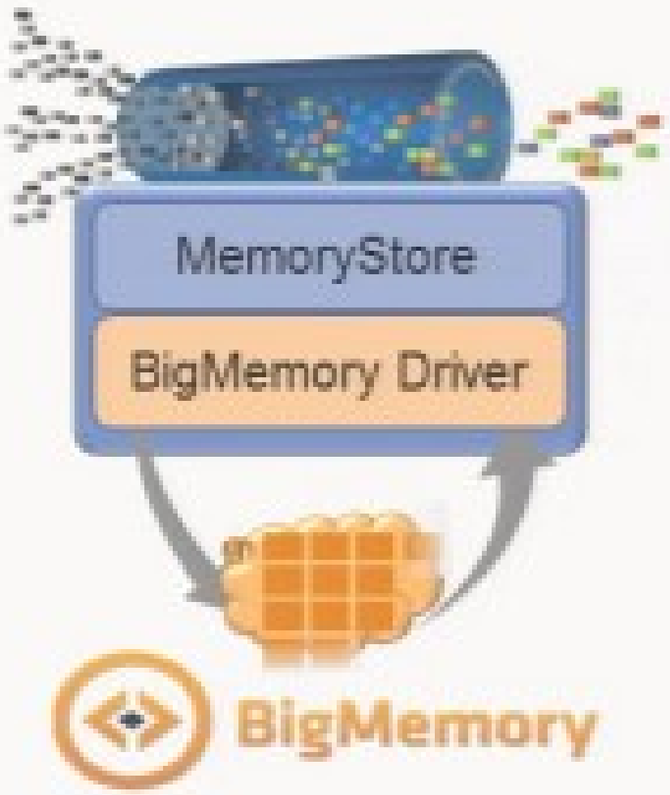
TERRACOTTA

BigMemory: In-Memory Data Management



Data in Motion & Data At Rest

- Apama MemoryStore can be backed by BigMemory
  - In addition to databases and other apps
  - Transparent to Apama developers (simply new mode to MemoryStore)
- Provides "infinite" memory access from within Apama
  - Delivered as "Driver" to Apama "Cache" API
  - Inserts, Edits, Deletes & Notifications allow multiple clients (inc. Apama Correlators) to share data
  - Optionally exposes 'search attributes' for Apama MemoryStore cache entries to enables external clients (such as Presto) to query data



Location & context-aware promotions increase revenue & loyalty

TURKCELL

Visualize campaign effectiveness in real-time; New campaigns can be deployed in 24 hours

- Customer who subscribes to the "Live Sports Results" package enters a Mall
- Partner "Sports Bar Diner" has capacity
- Push offer of 15% off any main meal this lunchtime

Location & context-aware promotions increase revenue & loyalty

TURKCELL

Visualize campaign effectiveness in real-time; New campaigns can be deployed in 24 hours

Real-time, One-to-One Marketing

- Call Records
- Program Triggers
- Profile Changes
- Location Changes

Customer location detected entering mall

- Customer who subscribes to the "Live Sports Results" package enters a Mall
- Partner "Sports Bar Diner" has capacity
- Push offer of 15% off any main meal this lunchtime

Location & context-aware promotions increase revenue & loyalty

TURKCELL

Visualize campaign effectiveness in real-time; New campaigns can be deployed in 24 hours

Real-time, One-to-One Marketing

- Call Records
- Program Triggers
- Profile Changes
- Location Changes

Customer location detected entering mall

- 35 million subscriber profiles analyzed each day
- Event streams with 40,000 events/second
- 200 business criteria checked in real-time

- Customer who subscribes to the "Live Sports Results" package enters a Mall
- Partner "Sports Bar Diner" has capacity
- Push offer of 15% off any main meal this lunchtime

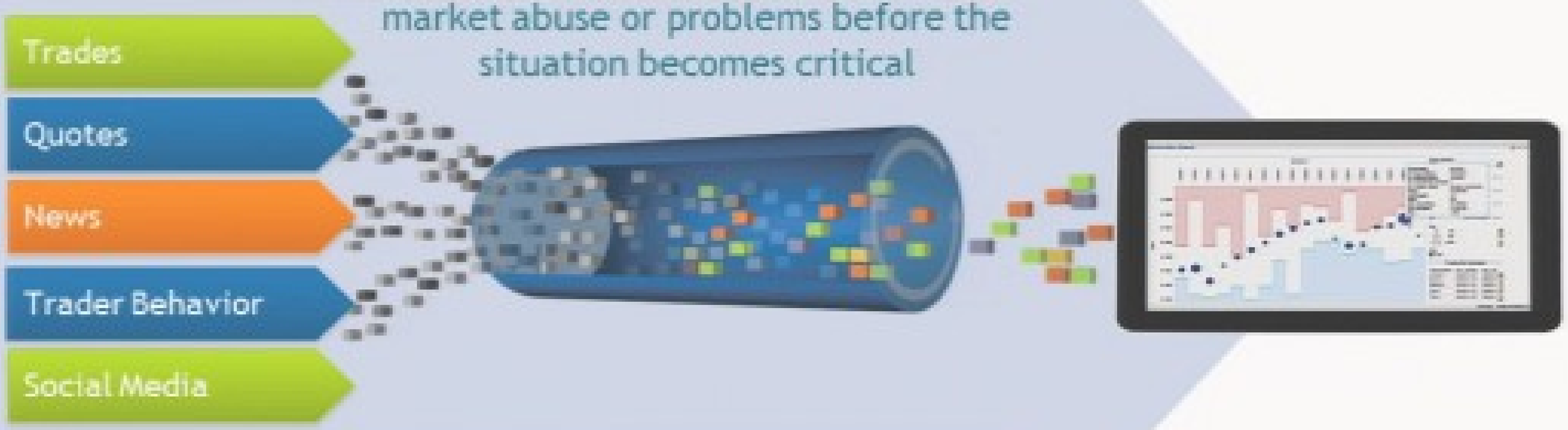
CRM Excellence Awards

- Marketing the right offer at the right time has yielded $15m revenue increase and 10x offer uptake

Mitigate potential regulatory violations before they make the headlines

citi

Large numbers of complex risk and surveillance rules that can detect market abuse or problems before the situation becomes critical



Real-time and historic market data feeds

Mitigate potential regulatory violations before they make the headlines

citi

Large numbers of complex risk and surveillance rules that can detect market abuse or problems before the situation becomes critical

Real-time and historic market data feeds

Unusual trading activity just before news release suggests insider trading

Mitigate potential regulatory violations before they make the headlines

citi

Large numbers of complex risk and surveillance rules that can detect market abuse or problems before the situation becomes critical

Real-time and historic market data feeds

Unusual trading activity just before news release suggests insider trading

Notify relevant Compliance Officer

Mitigate potential regulatory violations before they make the headlines

citi

Large numbers of complex risk and surveillance rules that can detect market abuse or problems before the situation becomes critical

Trades

Quotes

News

Trader Behavior

Social Media

Compliance staff visualize trade abuse alerts; drill in to investigate

Real-time and historic market data feeds

Unusual trading activity just before news release suggests insider trading

Notify relevant Compliance Officer

software AG

Software AG
University Relations

Monitoring Large IT Infrastructures at CERN in Real-Time: A Case for Apama

- The ATLAS experiment at CERN produces **one petabyte (10^{15}bytes) of data per second**, the equivalent of 1.5 Million CD ROMs
- Monitored by 20,000 application programs running on 2,000 server machines with 17,000 CPU cores
- Apama recognizes patterns in the monitoring data streams and initiates corrective actions in real-time

CERN

CeBIT

Philipps Universität Marburg

Thank you!

Dr. Jürgen Krämer
VP Product Strategy & Product Management
+49 (0)6421 304-800-14
+49 (0)151-22972685
juergen.kraemer@softwareag.com

For more information:
http://www.softwareag.com/corporate/products/bigdata