The New Generation of Data Management

In-Memory Technologies in Action

Stephan Grotz
Terracotta Solution Architect

Dr. Jürgen Krämer
VP CEP & Analytics

10.01.2013

Agenda

1) Motivation
2) Introduction to Terracotta
3) Introduction to CEP
4) Introduction to Nirvana
5) In-Memory Data Management Platform
6) Q&A

Motivation

Data explosion compounds challenges

software AG

Data is getting bigger

Rapid Growth of Global Data from 2009-2020[1]

From 1 to 35 Zetabytes

70% of the data generated by individuals[1]

RFID Market to see some serious growth[2]

125 Billion
RFID tags in 2020

The number of mobile-connected devices will exceed the world's population in 2012[3]

7 Billion

Every minute in the Internet[4]

100.000 Twitter tweets

240.000 shared Facebook content

Global mobile data traffic will surpass[3]

10 exabytes in 2016

(1)   CSC Report „big data growth infographic", (2) itechex report, (3) Cisco Visual Networking Index 2011-2016, (4) Intel

Get There Faster.™

Big Data can generate significant financial value across sectors
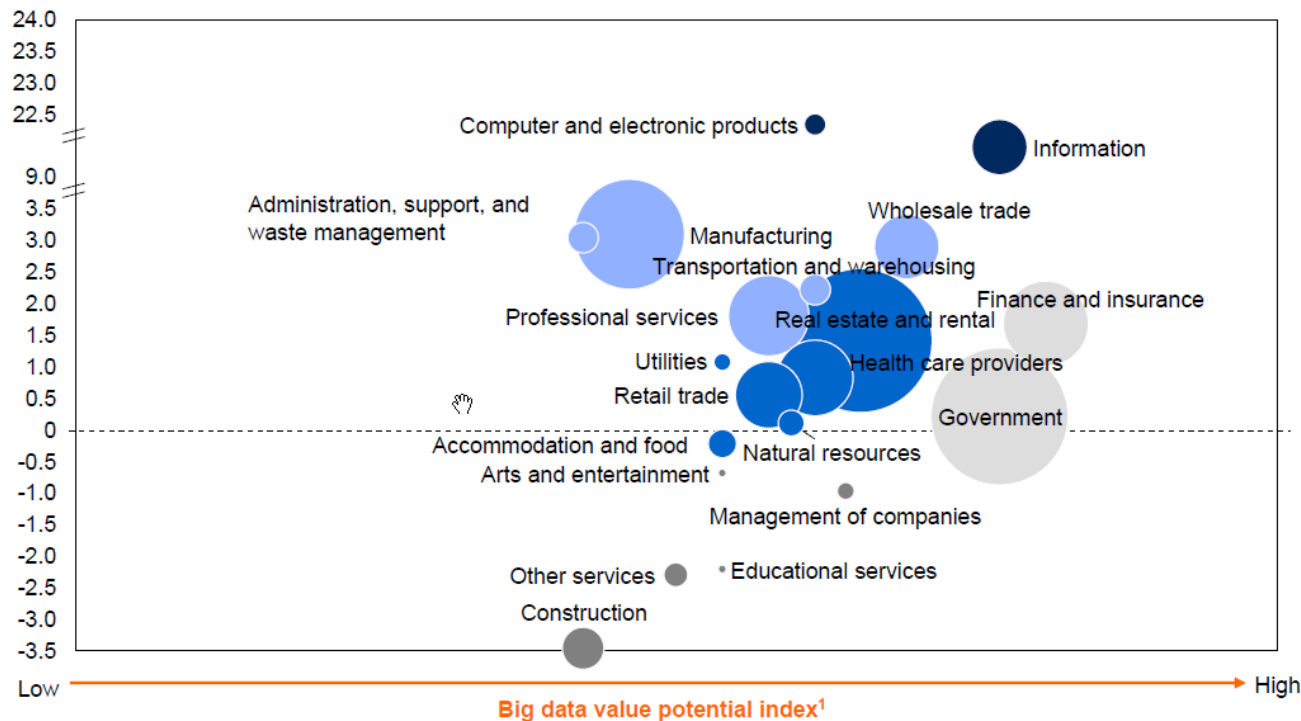
Verticals (Gartner)

Securities Trading

Banking

Travel and Transportation

Telecom

Media and Entertainment

Online Gaming

High Tech Manufacturing

eCommerce/Retail

...

McKinsey: Potential increase with Big Data in US retailers' operating margin[1] **60%**

Get There Faster.™

Classical data management architecture

- Operational and analytical environments are separated

- Mainly based on structured data

- Data exchange and transformation with latency
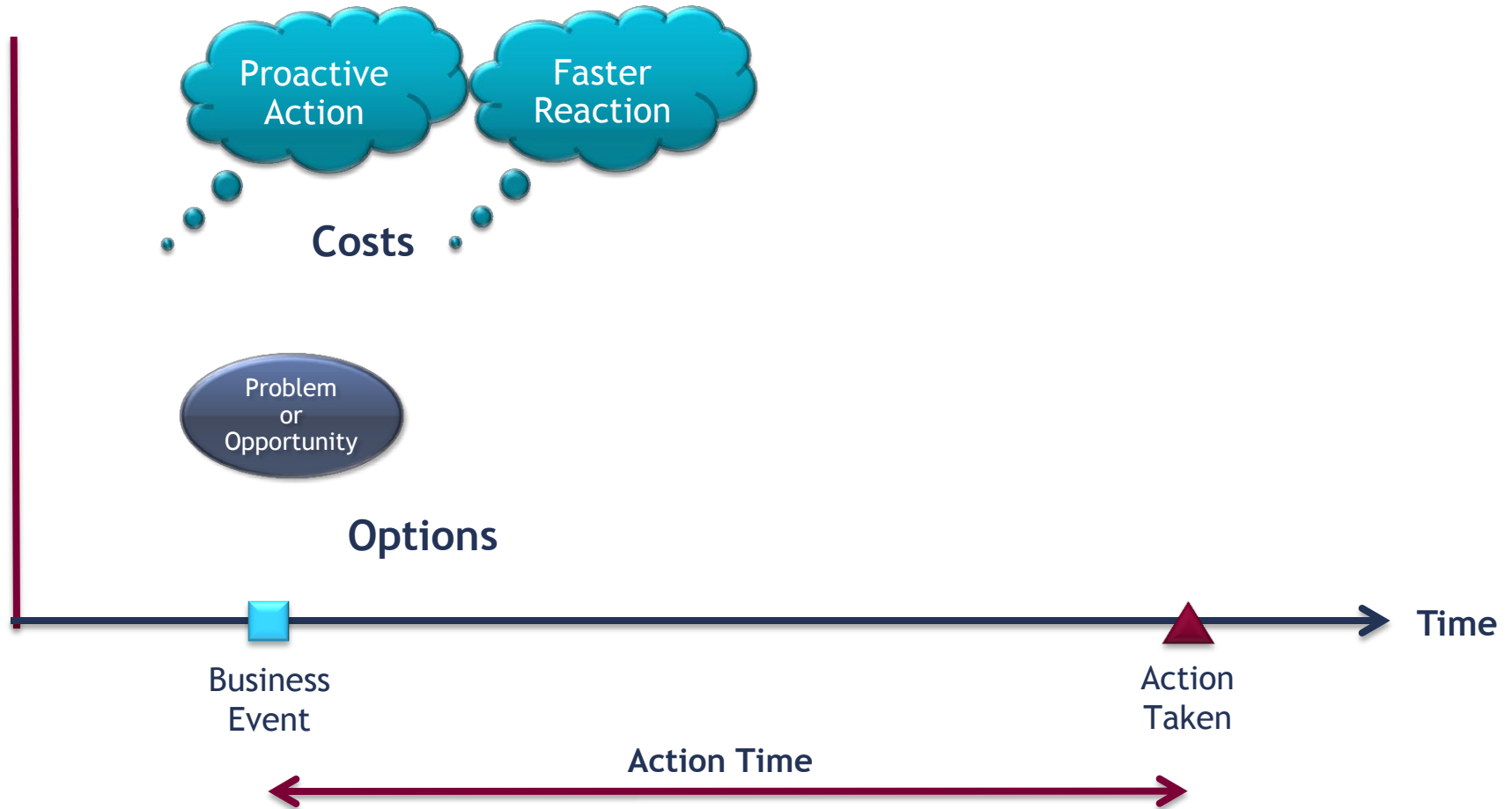
- Different ways end-user interacts with data

Applications

Business Analytics

OLTP DBMS

ETL

Data Warehouse
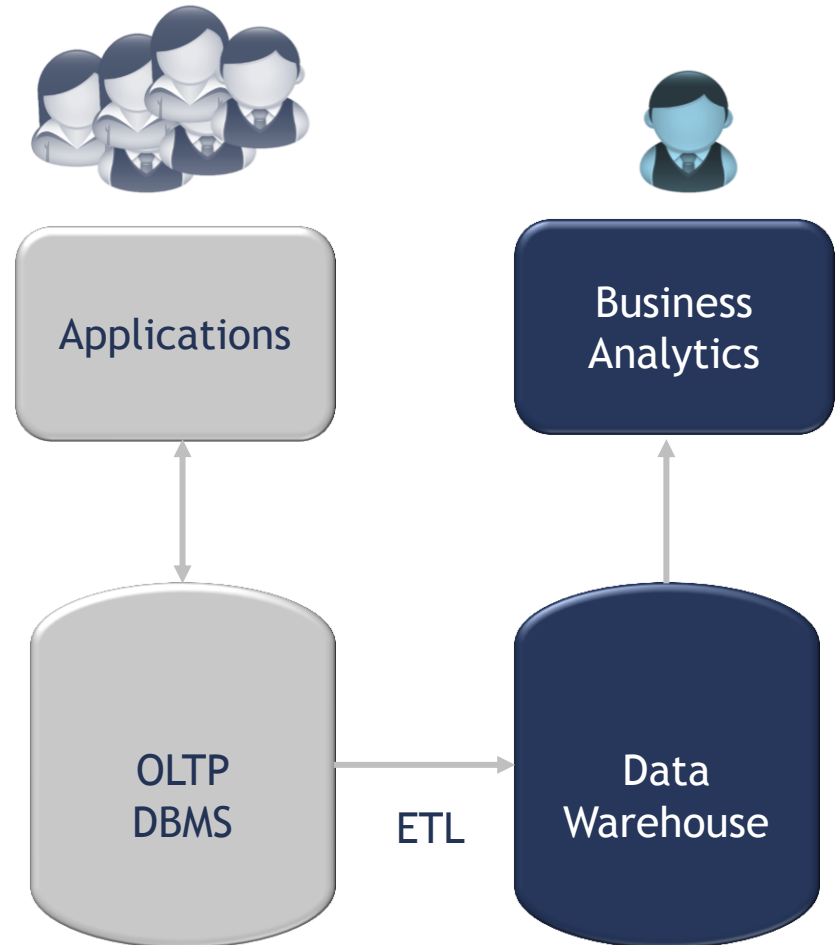
Pace of Business Drives Need for Real-time Analytics

Business is Moving to Real-Time Transactions

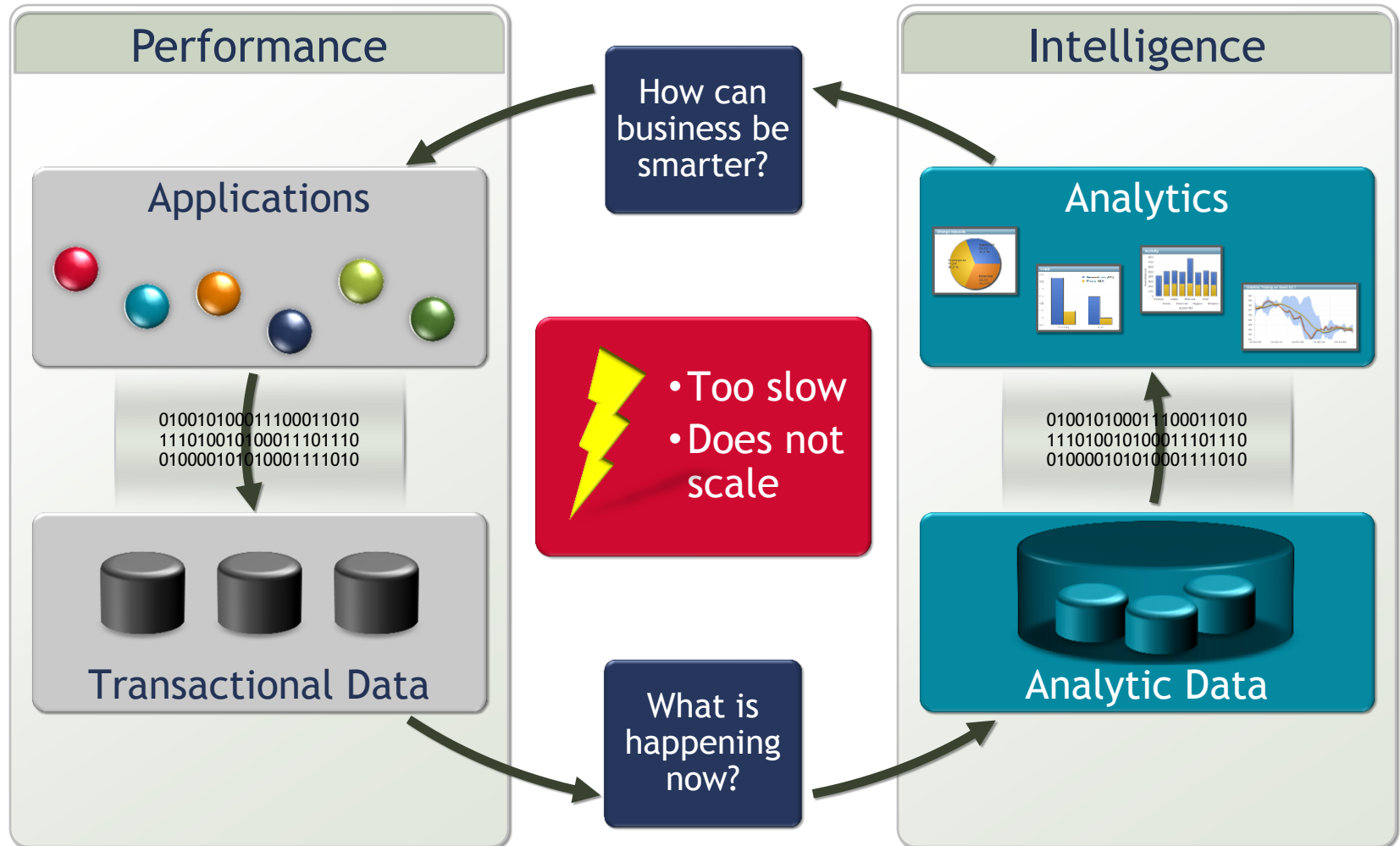|	months	weeks	days	hours	minutes	seconds
Trading Analytics					30 minutes →	5 seconds
Airline Operations					20 minutes →	30 seconds
Call Center Inquiries				8 hours →		10 seconds
Track Financial Position			1 day →		5 minutes	
Supply Chain Updates			1 day →		15 minutes	
Document Transfer			3 days → *Mail... express... fax... e-mail*			45 seconds
Refresh Data Warehouse	1 month →			1 hour		
Trade settlement		5 days →	1 day			
Build-to-Order PC	6 weeks →		1 day			

Source: **Gartner**

Get There Faster.™

Business Value of Faster Actions

Real-time data management goals

In-Memory data management for the real-time business
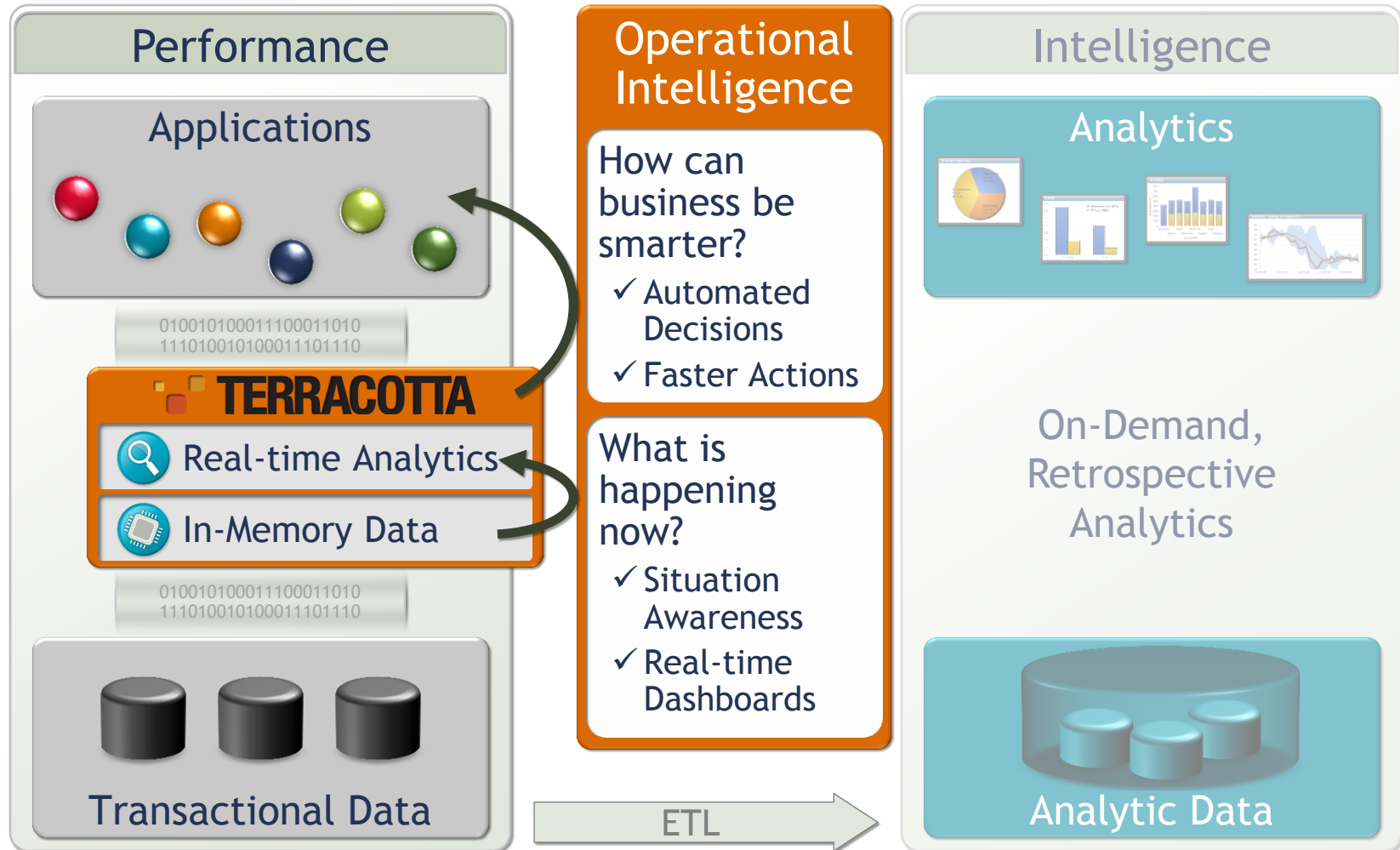
Introduction to Terracotta

Get There Faster.™

TERRACOTTA

- Provider of in-memory data management solutions for the enterprise

- Behind some of the most widely deployed software for scalability and performance

- 70% of Global 10000 use Terracotta

- Wholly-owned subsidiary of Software AG

500,000+ Enterprise Deployments
Majority of Global 2000

Bank of America VISA JPMorgan

verizon at&t BT 中国电信 CHINA TELECOM

Comcast ESPN CABLEVISION

PayPal Google salesforce.com

hp CISCO Adobe Sabre Holdings UPS

KAISER PERMANENTE STAPLES LANDS' END

LOCKHEED MARTIN SOCIAL SECURITY ADMINISTRATION USA NATIONAL INSTITUTES OF HEALTH FEDERAL AVIATION ADMINISTRATION

Get There Faster.™

Big Memory: Deliver Real-Time, High Volume Data Access
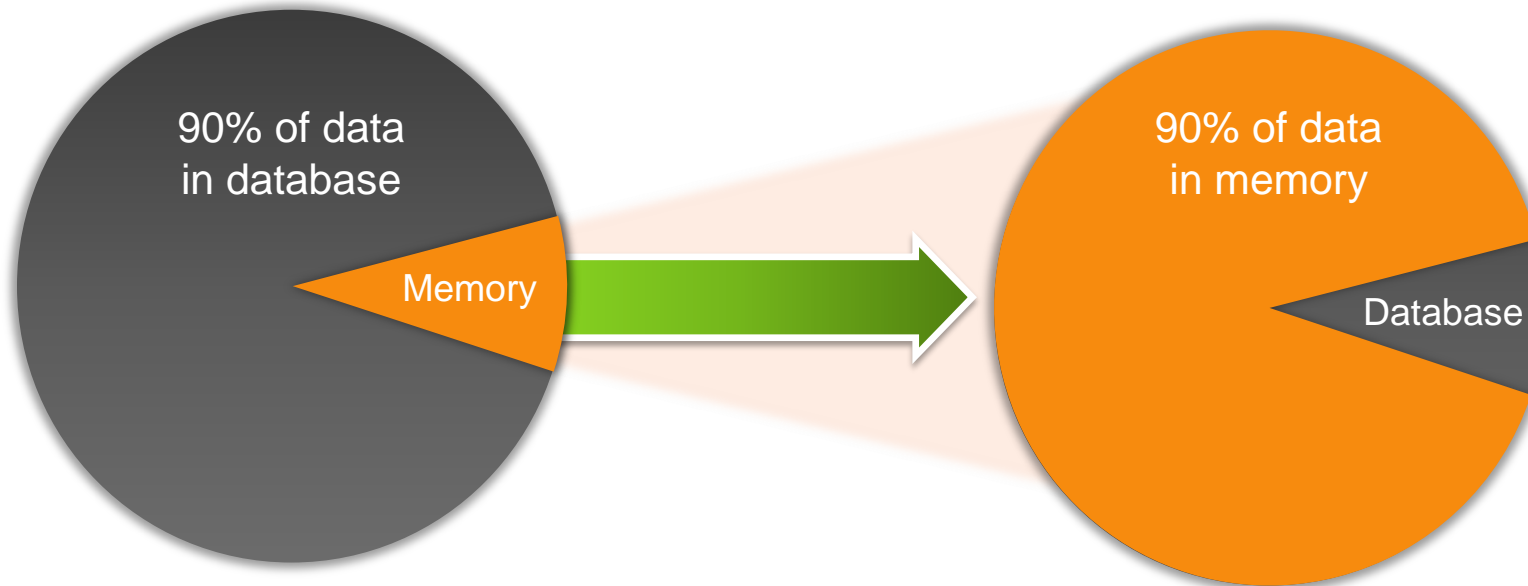
In-Memory

How do we maximize our use of cheap memory?

Steep drop in price of RAM

Big Data

How do we unlock the all value in our data?

Explosion in volume of business data

Get There Faster.™

BigMemory's Innovation

90% of data in database

Memory

90% of data in memory

Database

Milliseconds
App Response Time

Microseconds
App Response Time

Terracotta BigMemory Delivers Scale & Speed

Real-time access to massive amounts of business data

GO BIG.

MORE

Data

Users

Customers

Transactions

GO FAST.

QUICKER

Processing

Analysis

Services

Decisions

Get There Faster.™

Introduction to Complex Event Processing (CEP)

Get There Faster.™

What is an Event?

Example: Vehicle

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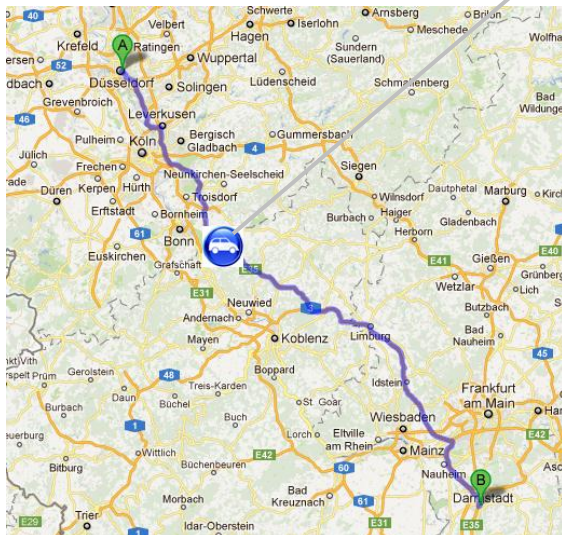
Example: SmartMeter

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

Event Streams

Stream of position reports from a truck

PS GPS GPS GPS GPS GPS GPS GPS GPS

Stream of stock prices

ce Price Price Price Price Price Price Price Price

Stream of sensor readings from a smart meter

ue Value Value Value Value Value Value Value Value

Increase in Data Velocity

Stream of position reports from a truck
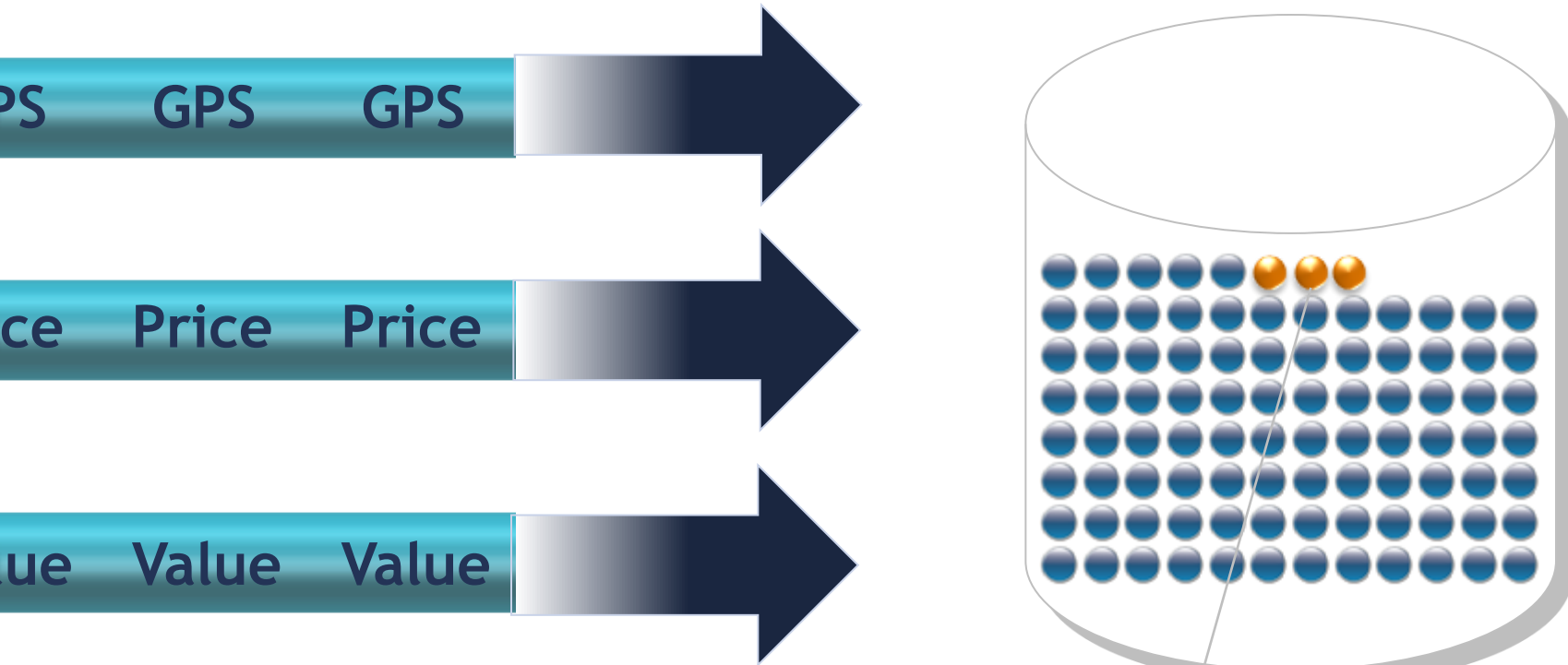
PS **GPS** **GPS** **GPS** **GPS** **GPS** **GPS** **GPS** **GPS**

Stream of stock prices

ce **Price** **Price** **Price** **Price** **Price** **Price** **Price** **Price**

Stream of sensor readings from a smart meter

lue **Value** **Value** **Value** **Value** **Value** **Value** **Value** **Value**

Increase of Data Volumes

Stream of position reports from a truck

PS GPS GPS

Stream of stock prices

ce Price Price

Stream of sensor readings from a smart meter

ue Value Value

Need for Real-time Analytics

PS   GPS   GPS

ce   Price   Price

ue   Value   Value
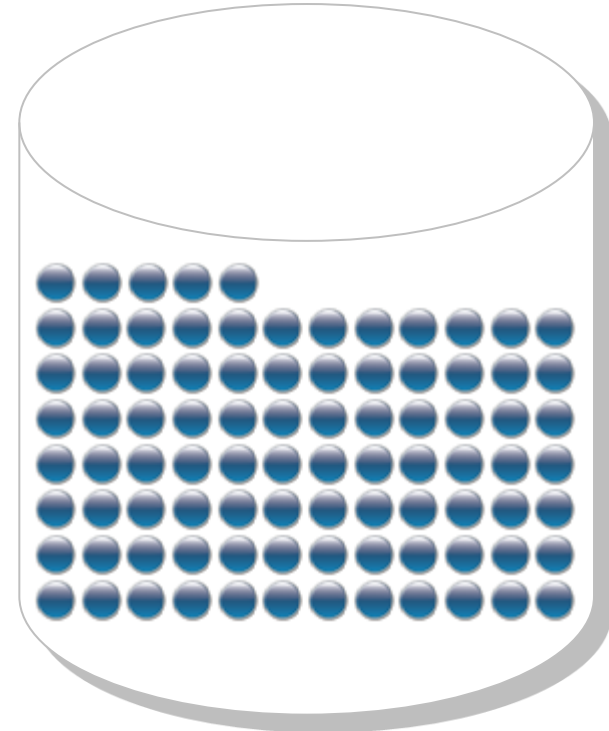
➔Newest data is more relevant!
➔Database becomes bottleneck!

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

Traditional „Store-and-Analyze"

Request

Response
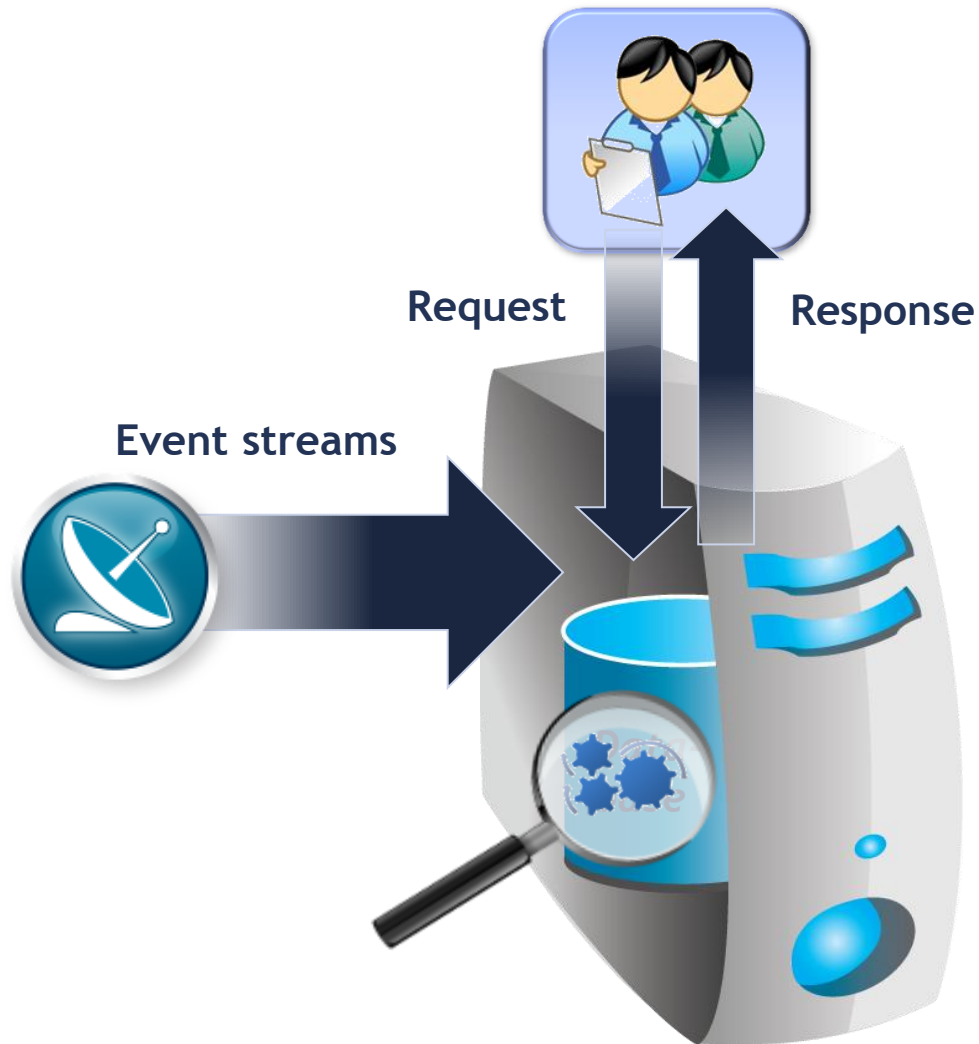
Event streams

Two phases
1. **Store data**
2. **Pull-based analysis**

Problems

- **Data store grows permanently**
- ➔ **Expensive search & analysis**
- **Not designed for continuous query evaluation**
- ➔ **Workarounds entail high load**

Event Stream Processing

Two phases
1. **Store business logic**
2. **Push-based analysis**
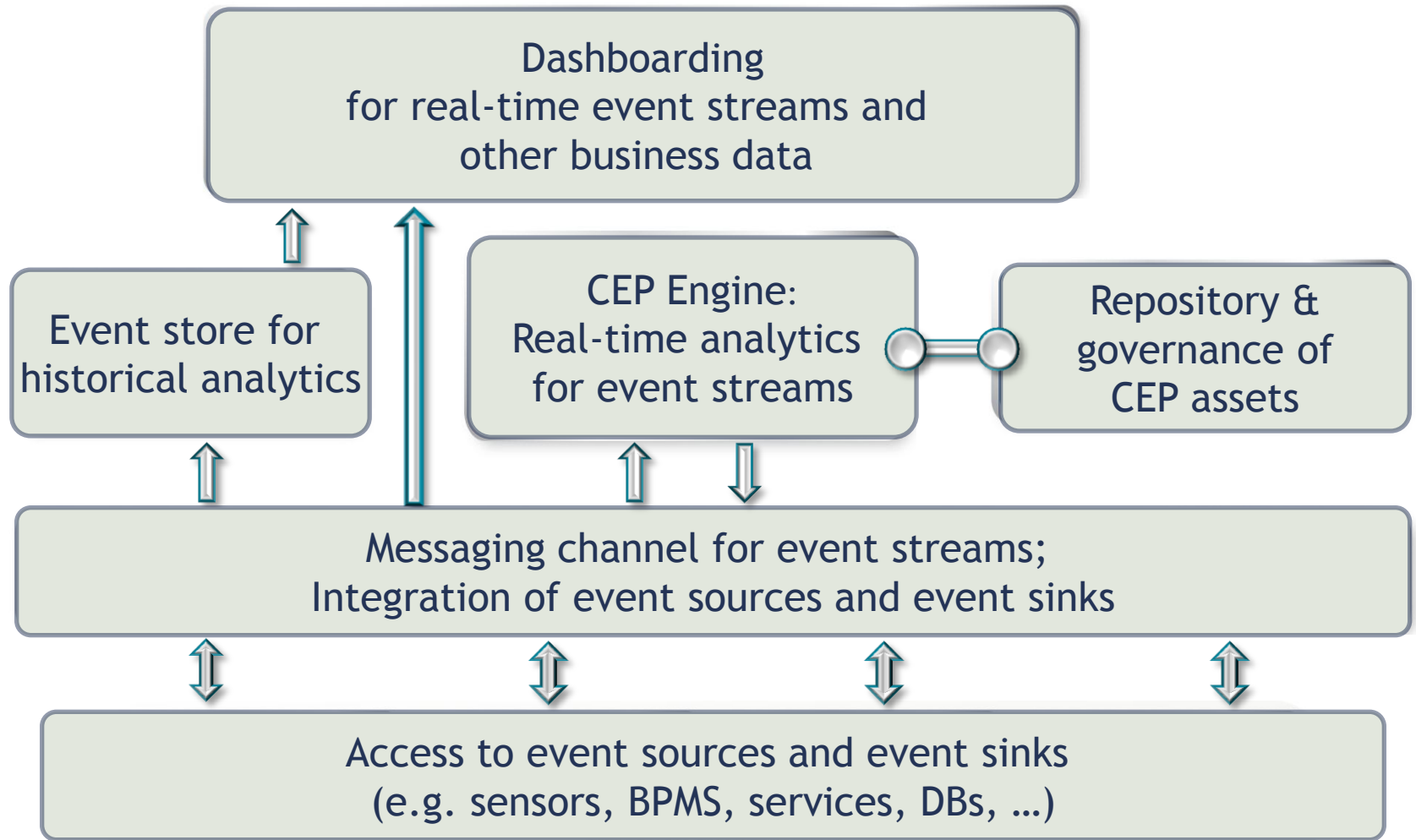
**Register business logic
(patterns, KPIs, ...)**

Input streams

Complex Event Processing Engine

Output streams

In-memory, on-the-fly analysis!

Software AG's CEP Platform: webMethods Business Events

Real-time analytics for Big Data in motion

External streams
(news feeds, click streams)

Event Bus

Continuous Analytics

CEP Engine

Pattern and trend detection

Input streams (cache changes) or context information

Output streams (KPIs, business events)

Real-time Dashboards

Alarms & Actions

Event Bus

Enterprise Applications

In-Memory Data **Terracotta**

Transactional Data

Get There Faster.™
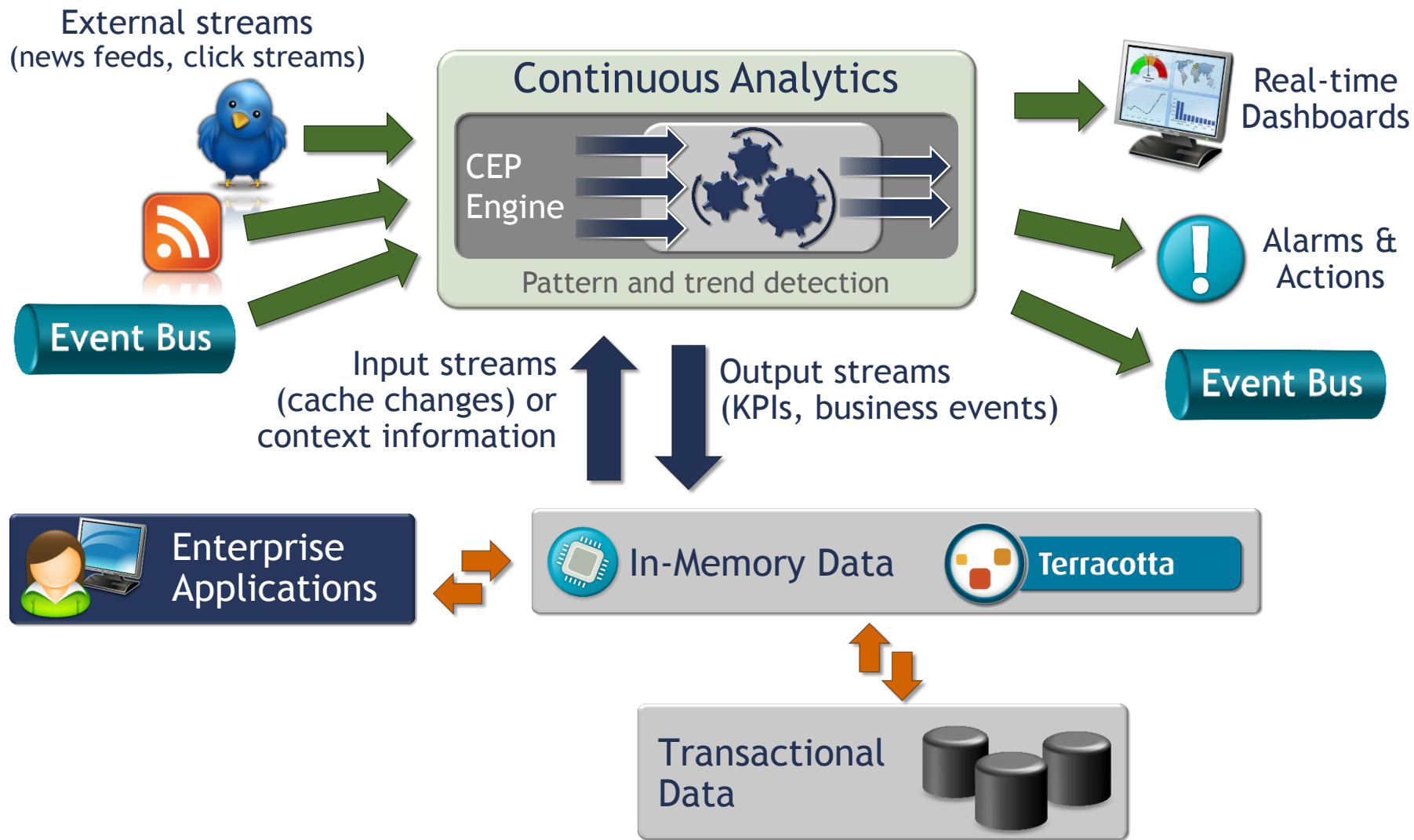
Basic idea of continuous cache queries

EHCache

key	value		
id	customer	region	revenue
1	AB AG	EMEA	1.000.000 €
4	CD GmbH	NA	2.000.000 €
6	EF Ltd	EMEA	30.000 €
8	AB AG	EMEA	3.770.000 €

Continuous queries:
- `revenue ≥ 1.000.000 €`
- `sum(revenue), avg(revenue) group by region`
- `sum(revenue), avg(revenue)`

CEP Engine

key	value		
id	customer	region	revenue
1	AB AG	EMEA	1.000.000 €
4	CD GmbH	NA	2.000.000 €
8	AB AG	EMEA	3.770.000 €

key	value	
region	sum	avg
EMEA	4.800.000 €	1.600.000 €
NA	2.000.000 €	2.000.000 €

key	value	
agg	sum	avg
AVG	6.800.000 €	1.700.000 €
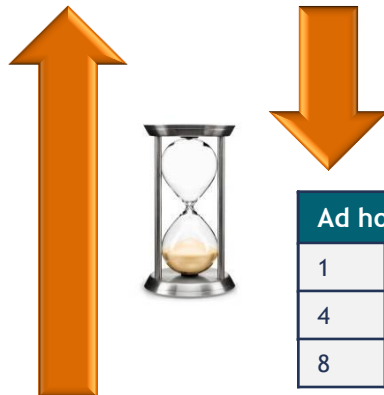
Ad hoc query result			
1	AB AG	EMEA	1.000.000 €
4	CD GmbH	NA	2.000.000 €
8	AB AG	EMEA	3.770.000 €

Ad hoc query:
`revenue ≥ 1.000.000 €`

- Continuous Cache Queries
  - Use Cache change events
  - Maintain result cache
    - Always up to date
    - Access without additional costs/delay

Technical demo

CEP Projects in Software AG Designer

<u>Single </u>Eclipse-based development environment for all webMethods products

Project definition

Database integration

Test data

Event Type definition

Event Streams

Business logic

Editor for SQL-style queries

Testing & Debugging
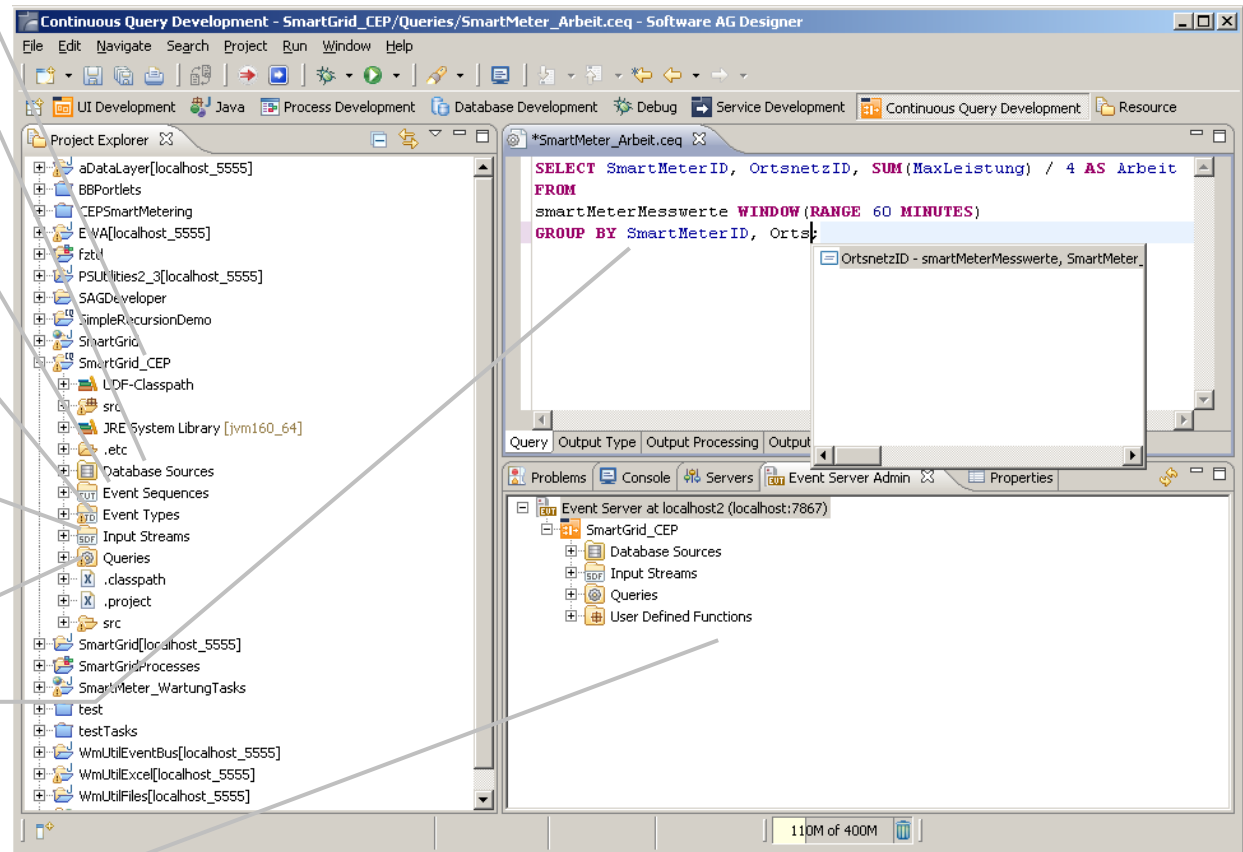
Business logic via SQL

Ease of use

- Standardized, well-established query language
- Same query language for event streams and historic data (relations)
- → Lots of developers are familiar with SQL
- No custom coding, no proprietary languages
- Implicit handling of time concepts
- Update of business logic at runtime
- → Higher reusability, flexibility, and quality
- → Reduced development times

Deterministic and sound semantics (based on application time)

- Snapshot-compliant with databases
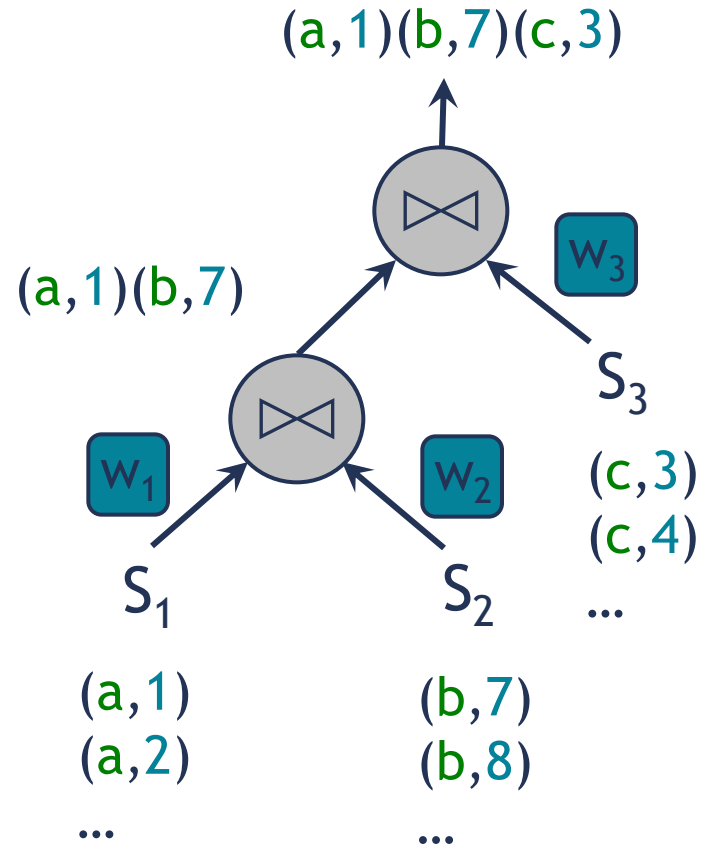- Enables query optimization for continuous queries

Get There Faster.™

Reproducible and traceable results

Issues for well-defined query semantics

- Notion of time
- Windowing constructs
- Schema stability
- Continuous output
- Stream ordering
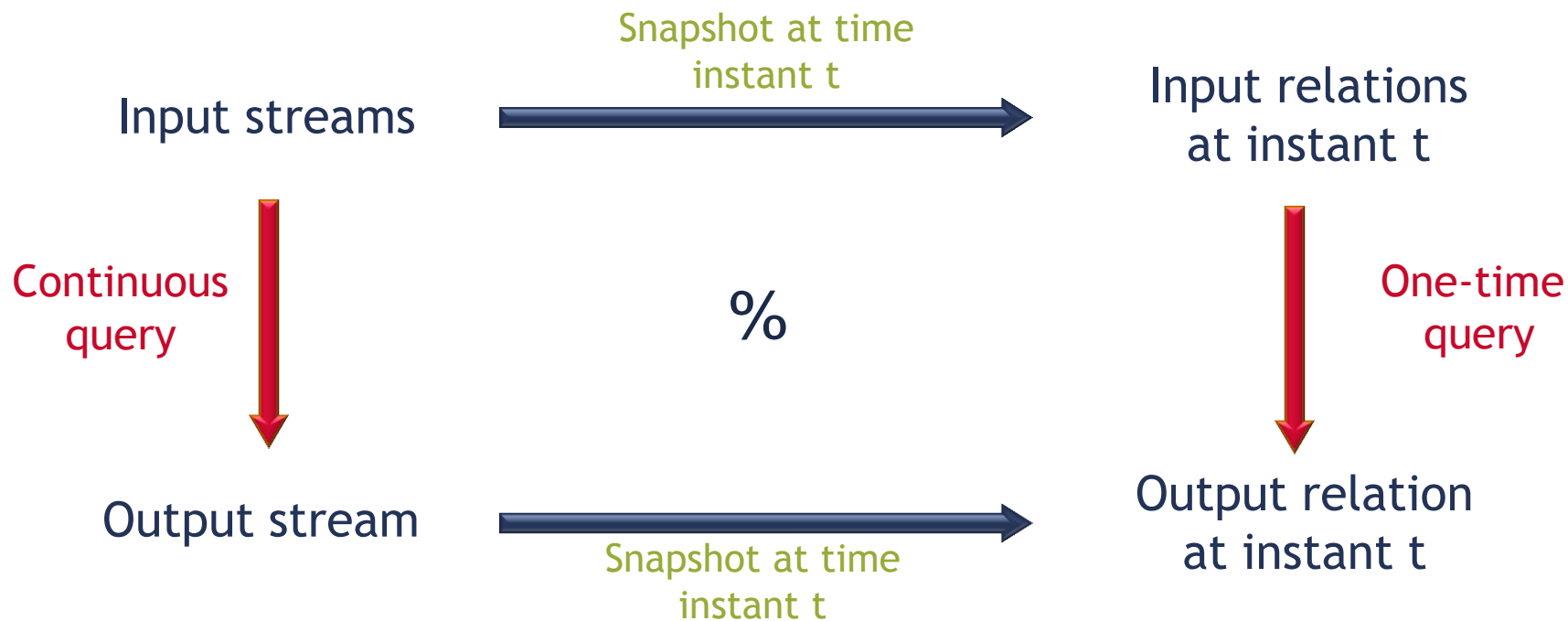- Determinism
- Optimizations

Semantics must not depend on system internals!

$(a,1)(b,7)(c,3)$

$\bowtie$ W_3

$(a,1)(b,7)$

S_3

W_1 $\bowtie$ W_2 $(c,3)$
$(c,4)$

S_1 S_2 ...

$(a,1)$ $(b,7)$
$(a,2)$ $(b,8)$

...   ...

Semantic compliance with databases

Input streams

Snapshot at time instant t →

Input relations at instant t

Continuous query ↓

%

One-time query ↓

Output stream

Snapshot at time instant t →

Output relation at instant t

Exact specification of query results for any point in time
- DBS would produce identical results if applied to every single time instant

All conventional transformation rules applicable due to snapshot reducibility
➜ Powerful query optimizations applicable

Event processing – design patterns

Basic

- Filtering
- Transformation
- Merge
- Correlation
- Aggregation
- Grouping
- Enrichment

EXACT ANSWERS

Advanced

- Duplicate elimination
- Difference
- Pattern matching
- State machines
- Non-event detection
- Set memberships
- User-defined functions
- Recursion
- Database lookups & writes

Temporal

- Time-based windows
- Count-based windows
- Partitioned windows
- Control over window advance
- Control over time granularity
- Temporal patterns

Stream Mining

APPROXIMATE ANSWERS

- Dynamic forecasting
- Outlier detection
- Density estimation
- Cluster detection

Get There Faster.™

Event engine – design patterns

Operator-level

- ✓ Push-based implementations
  - • Time-interval algebra
  - • Positive-negative algebra
  - • Punctuations
- ✓ Heartbeats
- ✓ In-memory caching & indexing
- ✓ Pub/Sub mechanism
- ✓ Pull-based implementations for request-respond queries
- ✓ User-defined functions & operators

Query-level

- ✓ Dynamic registration
- ✓ Semantic subscriptions
- ✓ Authentication & authorization

System-level

- ✓ Application time
- ✓ Simulation & replay
- ✓ Query batches
- ✓ Metadata management
  - • Schema information
  - • Runtime statistics
- ✓ Multi-threading
- ✓ High availability
- ✓ Query optimization
  - • Subquery sharing
  - • Cost model
- ✓ Distributed execution
- ✓ Fine-granular testing & debugging
- ✓ webMethods Deployer
- ✓ SAG Designer

Get There Faster.™

Graphical Query Modeler in next release
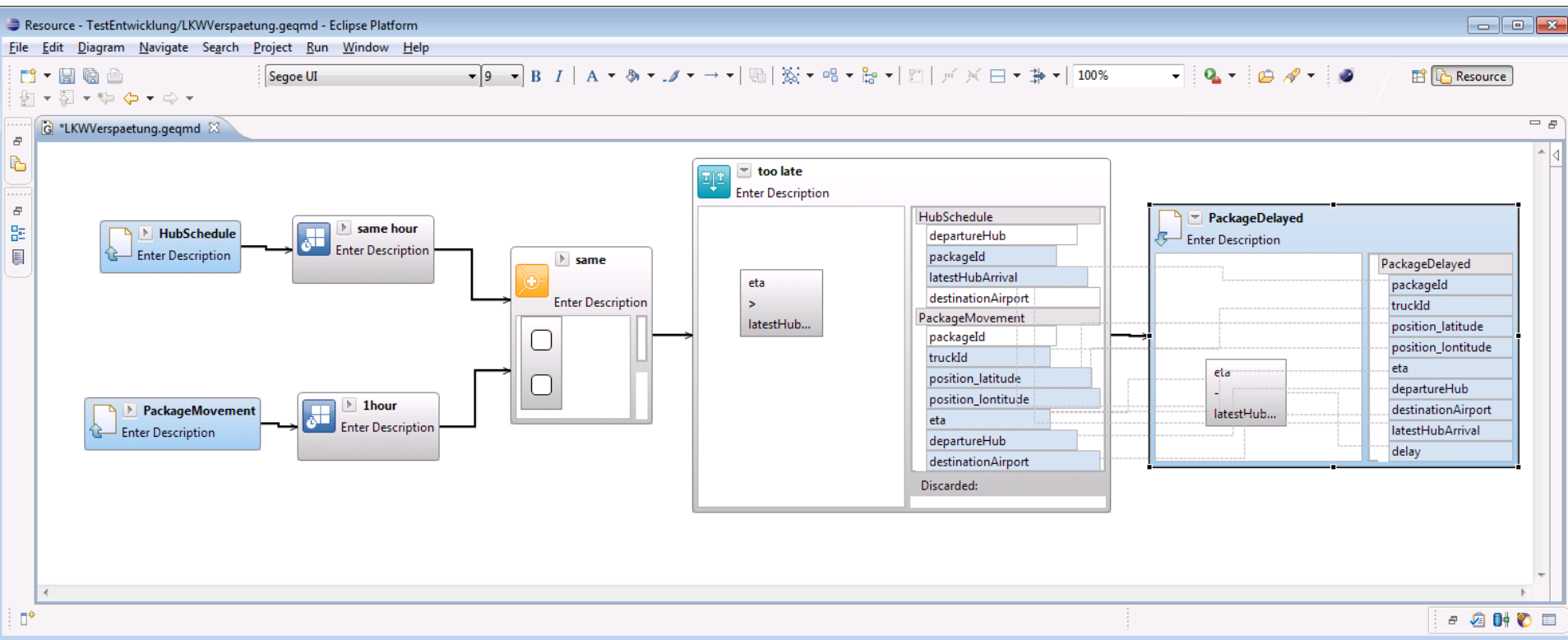
Mashup Dashboards for Live Insights

ARIS MashZone Creates Transparency

MashApp

Feeds

Data sources

| Event Streams | CSV | XML | ARIS PPM | Excel | RSS | JDBC |
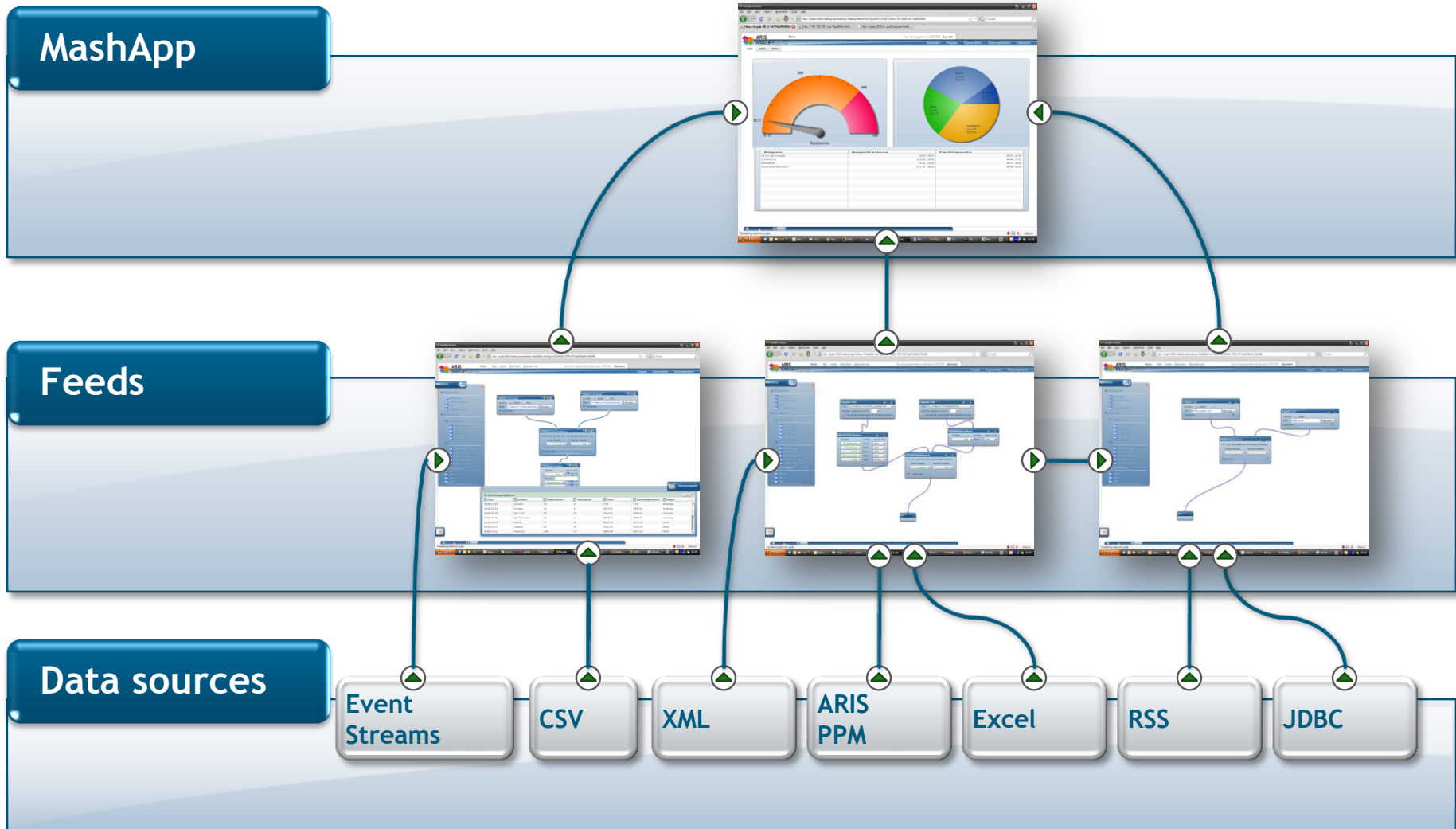
Event Driven Architecture (EDA)

Get There Faster.™
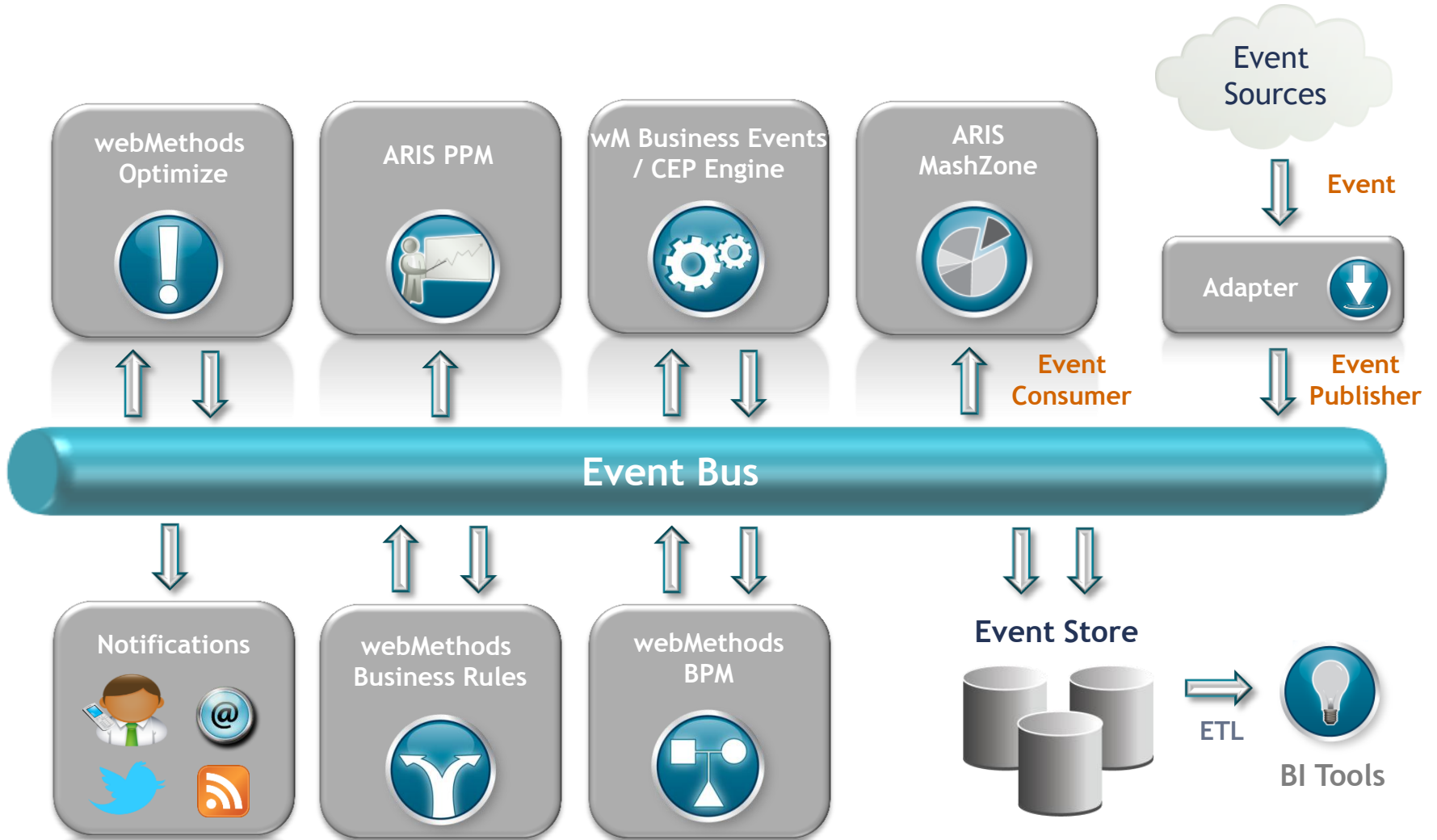
Transform Real-time Insights into Immediate Actions

webMethods Business Events

ARIS MashZone

Meter Sensor reading data

Event

Adapter

Meter outage detected

Event Consumer

Event Publisher

Event Bus

Taking action on meter outage

Services

Meter reset by field engineer

webMethods BPM

Intelligent Business Operations: Next Best Action

Intelligent Business Operations (IBO)

Decision process

| Event management to enhance situation awareness | Decision management to make faster and better decisions | BPM flow management to coordinate more-flexible processes |

Observe, orient → **Decide** → **Act**

My business operations

A → C → D

Other departments, customers, partners and the environment
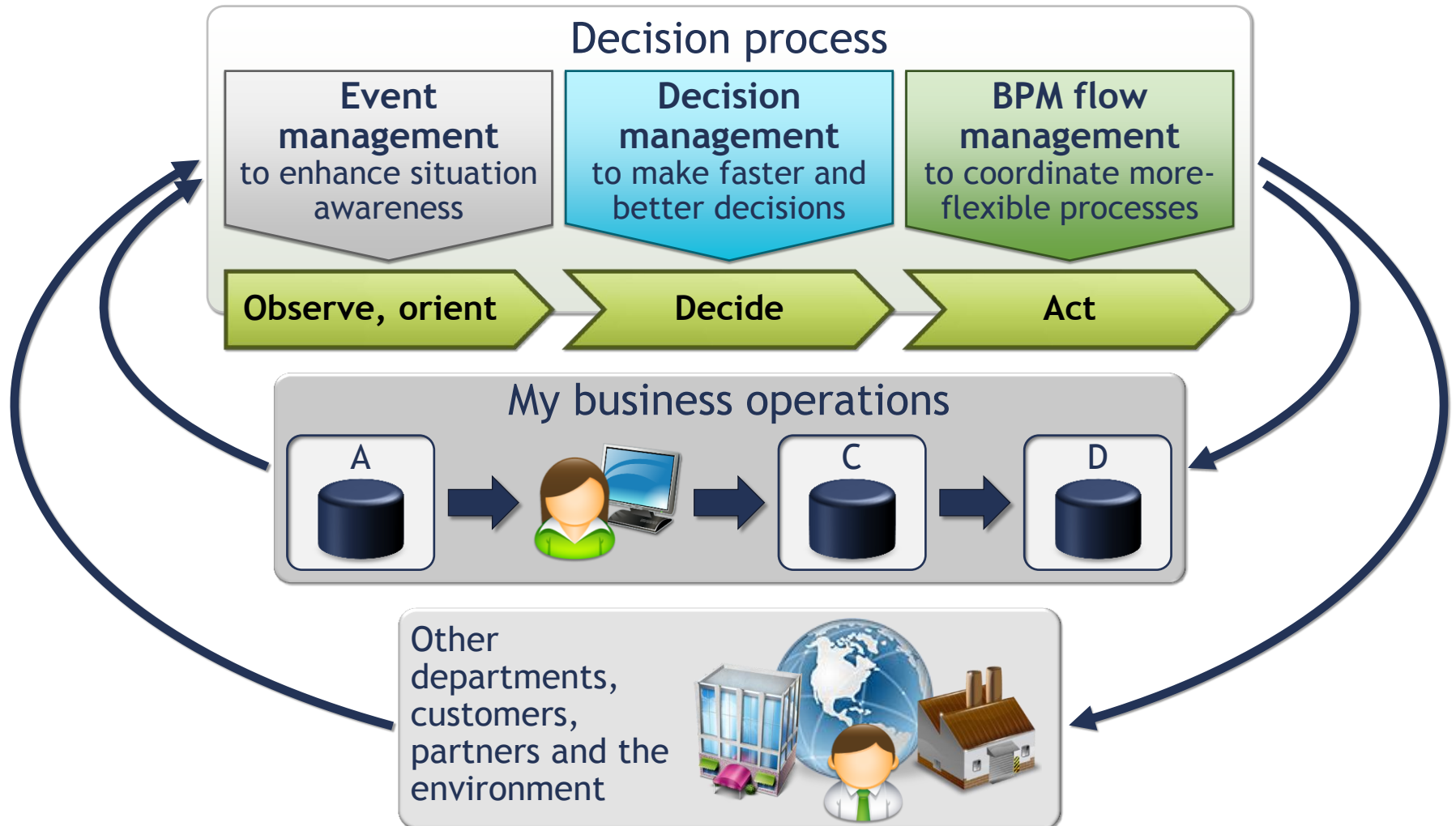
Source: W. Roy Schulte: Apply Three Disciplines to Make Business Operations More Intelligent, Gartner Inc., March 2012

Get There Faster.™

Business Impact of CEP

➢ Improves the quality of decision making by presenting information that would otherwise be overlooked.

➢ Enables faster response to threats and opportunities.

➢ Helps shield business people from data overload by implementing a management-by exception policy.

➢ Reduces the cost of manually processing the growing volume of event data in business.

CEP is one of the pillars of the emerging "big data" movement.

Gartner Hype Cycle for Data Management 2011

Source: Gartner Inc.: W. Roy Schulte; Zarko Sumic: „Hype Cycle for Data Management", 2011.

High-Speed Messaging

webMethods Nirvana Messaging

High Speed Universal Messaging

- A single end-to-end solution for the delivery of real-time data
    - Across multiple delivery channels (Enterprise, Web, Mobile)
    - Between multiple languages and platforms

- Supporting an unrivalled selection of languages, messaging standards and protocols using common APIs

- Proven track record in the most demanding markets
    - Our customers deliver real-time data to tens of thousands of connected users every day

webMethods Nirvana Messaging

Enterprise

Web

Mobile

Get There Faster.™
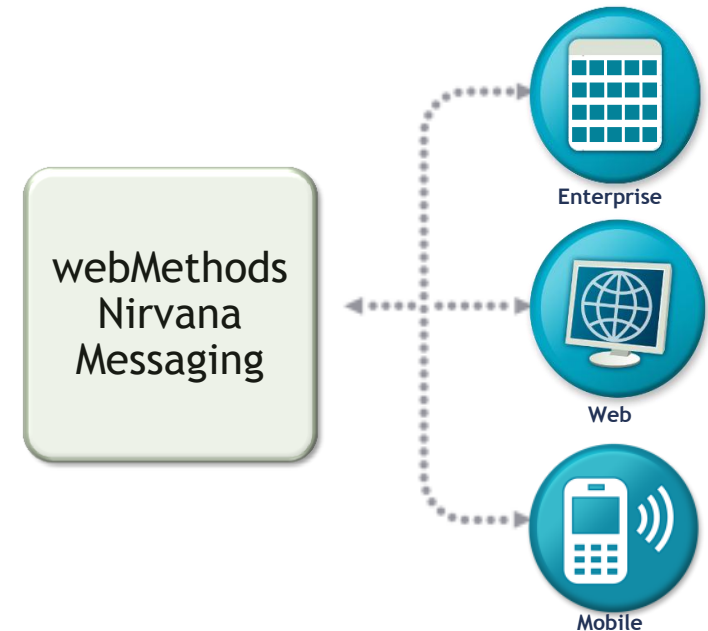
Benefits of Nirvana Universal Messaging

SIMPLICITY

Common APIs for all delivery channels
One product – multiple use cases

SCALABILITY & PERFORMANCE

No bridging between components

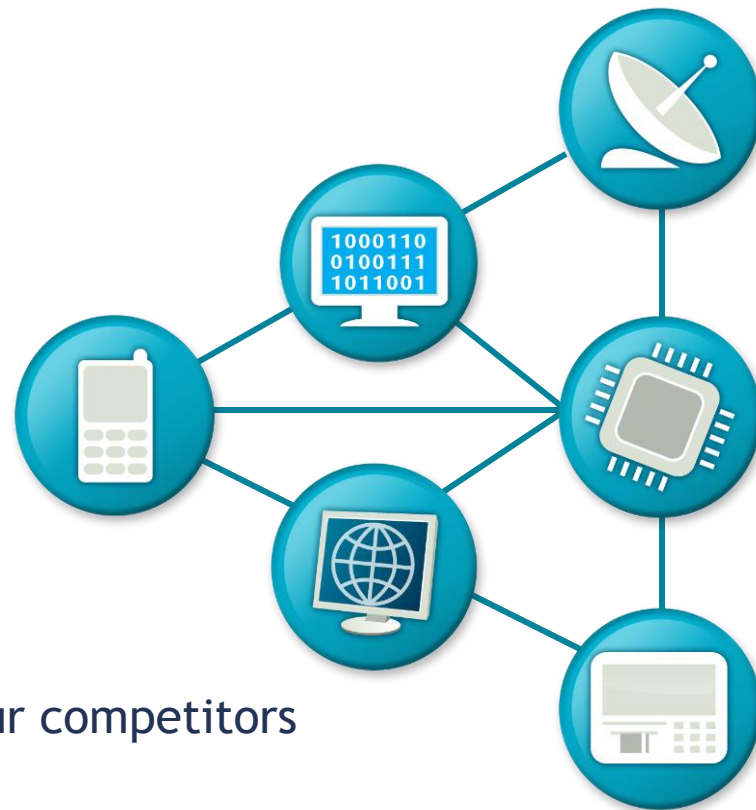Data travelling over fewer hops

FLEXIBILITY

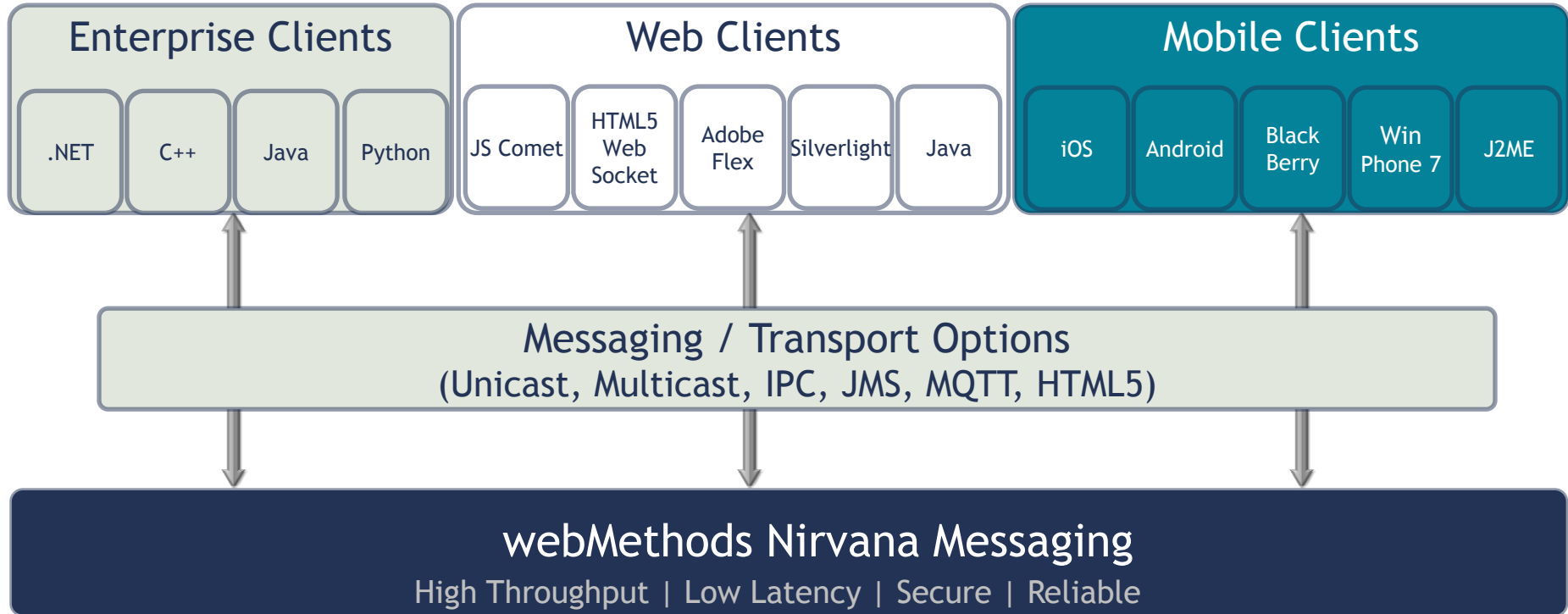Support for all your delivery channels

FUTURE PROOFING

Ready for new delivery channels before your competitors

VALUE

Less vendors, less cost – hardware, software, operational savings

Universal Messaging Today

Enterprise Clients

- .NET
- C++
- Java
- Python

Web Clients

- JS Comet
- HTML5 Web Socket
- Adobe Flex
- Silverlight
- Java

Mobile Clients

- iOS
- Android
- Black Berry
- Win Phone 7
- J2ME

Messaging / Transport Options
(Unicast, Multicast, IPC, JMS, MQTT, HTML5)

webMethods Nirvana Messaging
High Throughput | Low Latency | Secure | Reliable

webMethods Nirvana Messaging Customers

Customer Successes

- Global customer base

- Powers 2 of the top 3 market-leading Foreign Exchange platforms
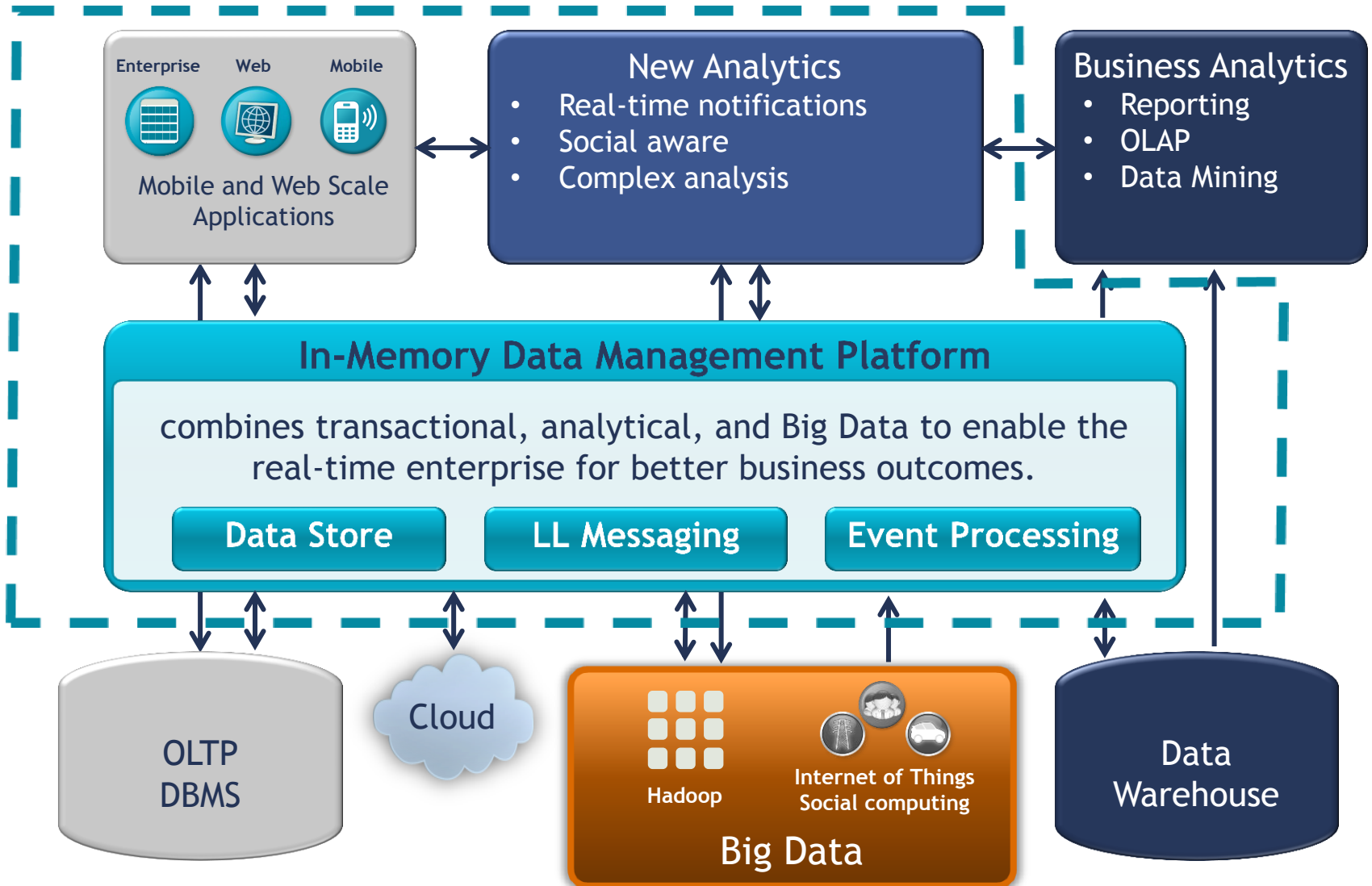
- Increasing reach beyond financial services

Global Customer Base

In-Memory Data Management Platform

Get There Faster.™

Key Advantages

➢ **Leverage enterprise-proven and high-performance in-memory technology (scale-up & -out)**

➢ **Process data in real-time (move, manage, analyze)**

➢ **Use commodity hardware and software**

➢ **Preserve existing business data and data stores**

➢ **Implement in modular fashion**

Get There Faster.™

Thank you!

Get There Faster.™

Contact: juergen.kraemer@softwareag.com