Big Event Data Benchmarking

Hans-Arno Jacobsen

Big Events -
for low latency analytics & BPM,
application performance management,
smart grids, and participatory sensing

Hans-Arno Jacobsen

Middleware Systems Research Group / MSRG (msrg.org)
Low Latency Analytics & BPM

Vendor

Sale

Manufactory

Warehouse

Finance

Marketing

Packaging

FedEx

Pick up

Process

Pay

Check

Signature

Print receipt

Payment

Feedback

Goods delivery

Fill dispatch bill

Fill out-inventory bill

Credit card

Target price

Prototype

Out Take Control

Assign

Confirm

Determinate plan

Check stock Raw materials

Audit Raw materials

Determinate plan

Execute plan

Process control

Monitor

Strategy Design Marketing Order

Manufactory Payment

Requirement collection

Feature selection

Goods selection

Confirm features

Material Make plan

Monitoring

Statistic

Chart
Low Latency Analytics & BPM
Application Performance Management (APM)

Diagram showing the relationships between different components such as Identity Manager, Application server, Web server, SAP, Message Queue, Database, Web Service, Message Broker, Main Frame, 3rd Party, and Agent.
E-Energy, Smart Grids & Green Middleware

Need middleware to:
• Analyze
• Control
• Monitor

Changes in traditional energy power grids:
• Distributed energy resources
• Fine-grained monitoring of production and consumption
• Fine-grained control of consumption
Participatory Sensing

LTE 4G

Internet

DSRC

Traffic Database

Filter

Notification

Analytical Queries

Geographic Information
Towards Showers & Streams

- Scenarios that require **filtering** of events at large scales

- Scenarios that require **filtering & storage** of events at large scales

- Filtering & storage of “**event streams**”

- Filtering & storage of “**event showers**”
Showers vs. Streams

Event Processing¹
- Limited, no single schema
- Events vary in shape and size from one to the next
- Processing of many event expressions
- Tends to require support for aggregation
- Broader model & paradigm (dissemination, matching, coordination)

Event Stream Processing
- Schema-based, single schema per stream
- Stream tuples follow schema
- More single-expression processing-based
- Aggregation is a key requirement
- Focused on processing queries/expressions over event streams

¹ publish/subscribe-style processing
Call to Arms: Big Event Data

- No standards
- No benchmarks
- No future

We need to develop standards & benchmarks for event processing & Big Events!