Situational Awareness

Leveraging Versant JPA for Big Data Situation Awareness Applications

Matthew Barker, Director Prof. Services
Versant Corporation

Agenda

- ► Situational Awareness Defined

- ► Short SA Video

- ► Big Data and Situational Awareness

- ► What About SQL

- ► How Versant JPA Enables SA

- ► Situational Awareness Applications

Situational Awareness "Bake Off"

What is "Situational Awareness"

Relatively New Term

- One definition: "knowing what is going on so you know what to do"
- Has it's roots in military theory
  - Sun Tzu's *Art of War*
  - Term itself first appeared during WW I from crews of military aircraft
- Some modern applications of SA
  - Smart grid systems
  - Transportation systems
  - Telco networks
  - Network attack protection, fraud protection

VERSANT

Situational Awareness Video

Watch closely, the idea is to count how many times the people in white pass the ball.

(there will be a test afterwards)

VERSANT

Big Data and Situational Awareness

Benefits of Big Data

- Enables enterprises to increase
    - Efficiency
    - Effectiveness
    - Profitability
- But what is it's "unrealized" potential?
    - To make accurate decisions *quickly*
    - Another way to describe "**Situational Awareness**"
    - Heavily under utilized, but why?
    - **"The Invisible Gorilla"!**
- Other Big Data Names for "Situational Awareness"
    - "Live Data", "Real-time Analytics", others

VERSANT

"The future of computing is not just big iron. It's big data."

Tom Kalil, deputy director of the White House Office of Science and Technology Policy

There is little utility in collecting and storing data; value is created through its analysis

VERSANT

RDB: Why So Dominant?

Standard Interface

▶ SQL (with subsequent adoption by Oracle and IBM)

Mathematical Basis

▶ Relational Calculus and Algebra

▶ Basis for formal university level courses

Data Integrity

▶ ACID transactions

Marketing

▶ *Huge marketing effort by Oracle and others*

VERSANT

RDB: Shortcomings

Actually slower then technologies it replaced

- ▶ Network (sets) and Hierarchical (trees)

Joins are expensive and don't scale (intractable)

- ▶ **ACM Article: On the optimal nesting order for computing N-relational joins (1984)**

Heuristics, widespread adoption, (and marketing) have "kept" the technology alive (even thriving)

- ▶ BUT, "Big Data" requires a "tractable" solution
- ▶ Particularly with Situational Awareness apps

VERSANT

Emergence of noSQL Technologies

Object Databases (1990's)

▶ Still in heavy use in Transportation, Fraud Detection, Simulation Software, and Telco

> (good examples of Situational Awareness apps)

▶ No joins equates to high performance and more importantly, scalability

Hadoop, Cassandra, mongoDB, XML db's, etc. (2000's)

▶ Gaining some traction in search engines, social media, and other specialized uses

▶ Hadoop has the most adoption as map-reduce problems are ubiquitous

VERSANT

Will a noSQL Technology Dominate?

So far, noSQL has been slow to be adopted

- ▶ Many of the technologies are niche
- ▶ Expensive to learn and use
- ▶ Many do not include ACID
- ▶ No Standard Interface! (*almost* all of them)
  - \> Main reason why they are expensive and difficult to adopt
  - \> Locked into one vendor

VERSANT

What Are The Qualities of a Potential noSQL Tech to Support Dominance?

▶ General purpose use with a standard interface

▶ Meets the demands of "Situational Awareness Big Data" applications

> Scales and handles complex, real-time data, preferably with ACID capabilities

▶ Includes scalable data analytics

▶ Interfaces with other technologies

▶ Expertly marketed

▶ Growing user community

VERSANT

Java Persistence API

JPA

- ▶ Originally developed as an OR-Mapping technology (Hibernate, TopLink, etc.)

- ▶ 70% of Java developers use JPA for persisting data in their applications
  - > Widely adopted standard!

Problems with JPA ORM

- ▶ Same problem as SQL/relational technologies

- ▶ Still uses joins to retrieve related data as underlying data store is a relational database

- ▶ Intractable, non scalable solution

- ▶ Completely fails for Big Data Situational Awareness

VERSANT

Introducing Versant JPA

Implements the JPA 2.0 Standard

- ▶ Leverage existing developer expertise
- ▶ But with high performance and scalability

Features

- ▶ High performing, scalable standards based API
- ▶ Scalable ingestion framework
- ▶ Scalable analytics integration (with R)
- ▶ Hadoop Integration
- ▶ Application Evolution (hot)
- ▶ *Ideal for Big Data Situational Awareness*

VERSANT
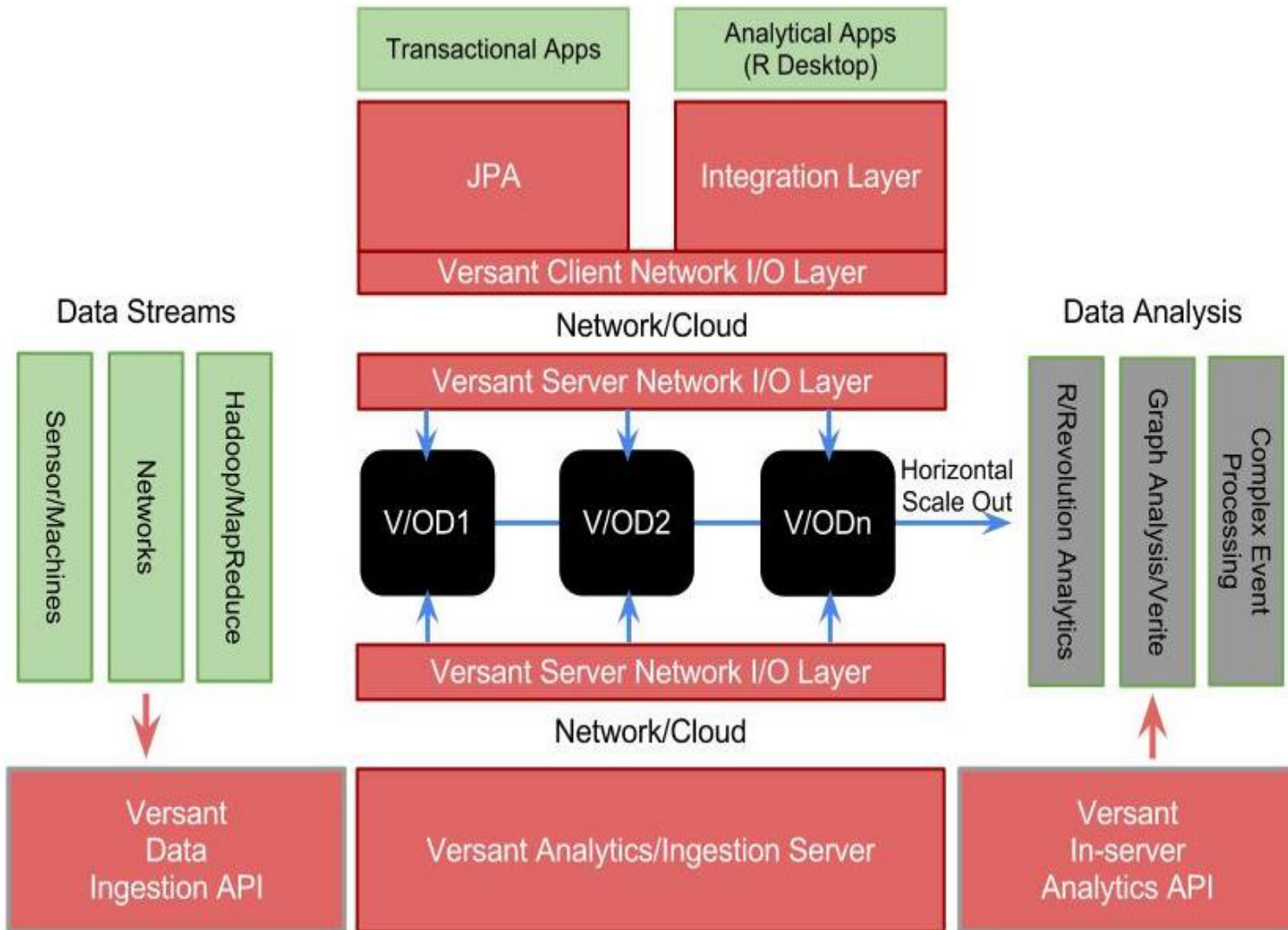
Versant JPA User Community

http://community.versant.com

▶ Hundreds of downloads per day since we introduced product

▶ If you know JPA already, then you can immediately start using Versant JPA
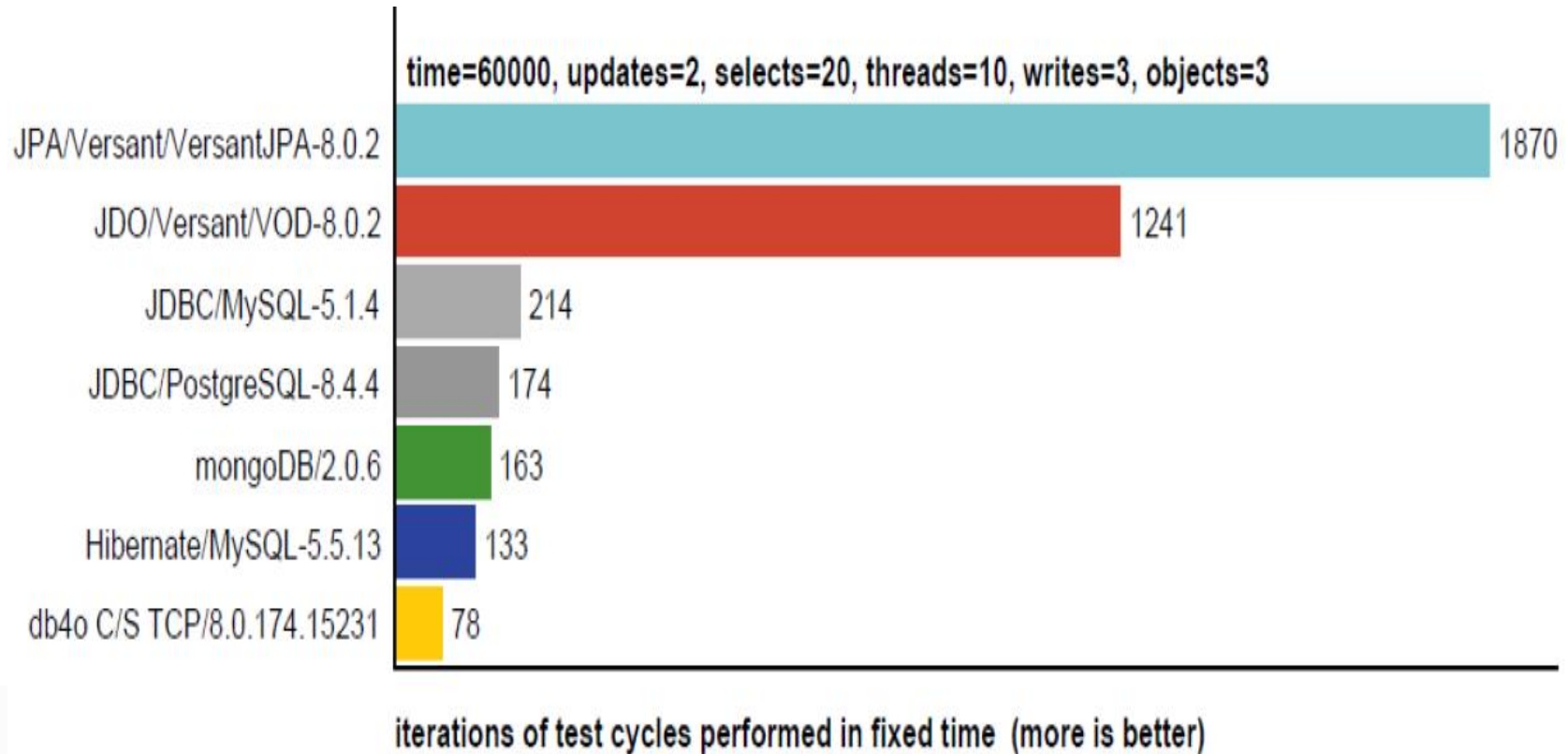
Versant Vendor Extensions

▶ Support for distributed databases

▶ Support for cursor queries

▶ Support for live schema evolution and loose schema mapping

VERSANT

Versant Situational Awareness Solution

Versant JPA Benchmark

time=60000, updates=2, selects=20, threads=10, writes=3, objects=3

JPA/Versant/VersantJPA-8.0.2	1870
JDO/Versant/VOD-8.0.2	1241
JDBC/MySQL-5.1.4	214
JDBC/PostgreSQL-8.4.4	174
mongoDB/2.0.6	163
Hibernate/MySQL-5.5.13	133
db4o C/S TCP/8.0.174.15231	78

iterations of test cycles performed in fixed time (more is better)

VERSANT

Versant Situational Awareness Apps

Fraud Detection

▶ Verizon's Telco Network fraud detection system

▶ Recently adapted for use with a Medicare fraud system developed by National Government Services at the request of the Obama administration

▶ Rules-based fraud detection system monitoring thousands of events per second and making critical real time decisions based on this data stream

VERSANT

Versant Situational Awareness Apps

Transportation

▶ **GE Transportation Rail System**

> Moves thousands of trains around North America in real time

> Benchmarks show 3X faster access for objects with medium complexity and 30X faster access for objects with high complexity

▶ **Sabre Travel Group**

> Sabre Sonic Inventory System – handles 50% of Sabre's travel business

> Travelocity site, provides real time travel schedule with pricing

> 150K availability transactions per second

> Caches the entire db (used like an "in memory" db)

VERSANT

Situational Awareness "Bakeoff"

In Smart Grid

▶ Based on CIM model, reference data model for smart grid

▶ Simulation based on real time monitoring of thousands of smart grid devices (using "end devices" in this simulation)

Same code base for simulation comparison

▶ First with Hibernate JPA ORM

▶ Then with Versant JPA

▶ System must respond to alarms and auto correct problems that arise

▶ Classic situational awareness application

VERSANT

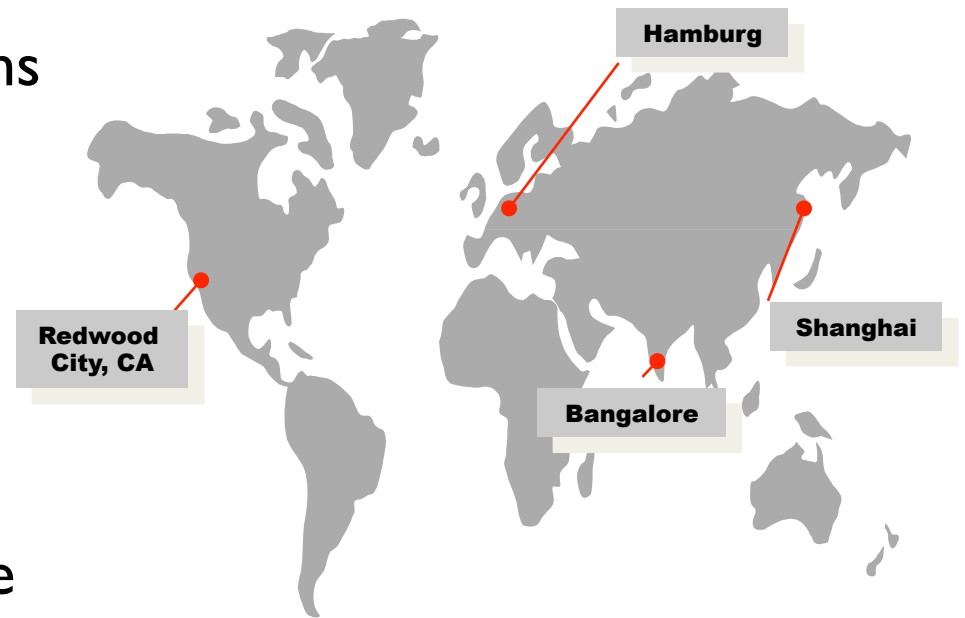# Versant JPA – More Information

http://community.versant.com

▶ Free download and free non-commercial use

▶ blog articles

▶ forum entries

jpa@versant.com

▶ asking questions

▶ Matthew Barker, mbarker@versant.com

VERSANT

About Versant

- Helping innovators to build high performance database applications
- NASDAQ:VSNT since 1996
- Telecommunications, Energy, Defense, Financial Services, Transportation, Manufacturing, other.
- Mission critical, high performance databases
- >150k installations world wide

Hamburg

Redwood City, CA

Shanghai

Bangalore

VERSANT

VERSANT