Bringing NSK to the Big Data Ecosystem

Sami Akbay
Web Action
Big Data and NonStop

• Introduction
• What is the big deal about big data
• NonStop in the Big Data World
• Real-time and Big Data
  – Introduction / architecture
• From Big Data to Big “value” data
• Q+A
Introductions – WebAction Team
(some familiar NonStop names)

Ali Kutay – president & ceo
• founder/investor/ceo: GoldenGate (Oracle),
  WebLogic (Oracle), AltoWeb(Borland) Formtek
  (Lockheed)

Sami Akbay – evp
• ceo: altibase / vp marketing + pm: GoldenGate,
  Embarcadero, AltoWeb / technical roles: Hearst,
  Rabobank, American Stock Exchange,
  MediaMetrix

Alok Pareek – evp
• cto: GoldenGate, vp: Oracle development, Data
  Integration & Data Replication: database kernel
developer (10 yrs):Oracle

Steve Wilkes – cto
• head of atg: GoldenGate / cloud DI: Oracle /
  architect: Middleware Company / cto: AltoWeb
  /atg: Cap Gemini (uk)

Simon Whitworth
• Director of Engineering - Barclays, Tandem,
  HP, GoldenGate, Oracle

Jim Knudsen
• Director of Application Engineering – Xypro,
  Electracard, Oracle, GoldenGate, ACI

Dale VanStratten
• Head of Application Sales – RS Software,
  Electracard, ACI, Insession, Data Processors
  International, First National Bank of Omaha,
  First Data Corp

Keith Dick
• Sr. Engineer – HP, Tandem, Compaq, Chi
Big Data Craze!

• According to Gartner 64% of enterprises surveyed indicate they’re deploying or planning Big Data projects.

• Yet even more acknowledge they still don’t know what to do with Big Data.
What is Big Data Anyway?

• 3Vs of Big data
  – Variety: structured, semi-structured, unstructured
  – Velocity: frequency at which records are created
  – Volume: size

• The 4\textsuperscript{th} V – “Value”
The Signal and The Noise

If the quantity of information is increasing by \textbf{2.5 quintillion} bytes per day, the amount of useful information almost certainly isn't. \textbf{Most of it is just noise}, and the noise is increasing faster than the signal. There are so many hypotheses to test, so many data sets to mine--but a relatively constant amount of objective truth.

\textbf{Separate signal from the noise:}
Pick the events of interest from a flood of information that’s flowing through the enterprise (and there’s more noise than ever!)
Everyone’s doing IT

• Big Data Adoption:
  – 27% of enterprises launched projects in 2012
  – 31% launching in 2013
  – 64% will have deployed solutions by the end of the year

• Top Drivers:
  – Enhance customer experience (54%)
  – Improve process efficiency (42%)
  – Launch new products/services (39%)
Big Data is here to Stay

Industries are using big data to transform business models and improve performance in many areas

<table>
<thead>
<tr>
<th>Retail</th>
<th>Manufacturing</th>
<th>Media and telecommunications</th>
<th>Advertising and public relations</th>
<th>Energy</th>
<th>Healthcare and life sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Customer relationship management</td>
<td>• Fraud detection and prevention</td>
<td>• Network optimization</td>
<td>• Demand signaling</td>
<td>• Smart grid</td>
<td>• Pharmacogenomics</td>
</tr>
<tr>
<td>• Store location and layout</td>
<td>• Supply chain optimization</td>
<td>• Customer scoring</td>
<td>• Targeted advertising</td>
<td>• Exploration</td>
<td>• Pharmaceutical research</td>
</tr>
<tr>
<td></td>
<td>• Dynamic pricing</td>
<td></td>
<td></td>
<td>• Operational modeling</td>
<td>• Clinical outcomes research</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Power-line sensors</td>
<td></td>
</tr>
<tr>
<td>Financial services</td>
<td>Financial services</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Algorithmic trading</td>
<td>• Fraud detection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Risk analysis</td>
<td>• Portfolio analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: A.T. Kearney analysis
Big Data Data-Stores

• Data Stores:
  – Hadoop / Cloudera / HortonWorks
  – Cassandra / DataStax
  – MongoDB
  – Couchbase

• Data Warehouses
  – HP Vertica
  – EMC Greenplum / Pivotal
  – Teradata / Asterdata
  – Oracle Exadata
  – IBM Netezza
  – Sybase – ASE

Store and process massive amounts of information that spans a long period of time.

AKA: Institutional knowledge, corporate intelligence, strategic intelligence, customer value, click stream, historic information …
Combining Real-time and History

- Identify and separate signals during data ingestion
- Accumulate all data in the big data store
- Send signals to NonStop
- Act on signals to build new generation of real-time Big Data apps
NonStop for the Events of Interest

BIG Data

WebAction Platform

All Events
(signal + noise)

Historic Information
institutional Knowledge

Big Data
Data Stores

Events of Interest
(signal)

Alerts, Big Data Apps
Operational Intelligence

Email
Texts
Logs
Clickstream Data
Social Media
Transactional Data

HP Integrity NonStop System
Comprehensive Approach to Big Data

WebAction

Real-Time (In Memory) Processing

Real-Time Enterprise

Batch Processing

Hadoop, MongoDB, Cassandra

Data Warehouse

Immediate, Real-Time Decisions (Seconds after Data is produced)

Historic data to support real-time decisions
Acquire Big Data

- Data from transactional sources is acquired via redo or transaction logs
- No Production Impact
- No Application changes
Process

- Fuse Data from current and historic sources
- Scale data using partitioned streams, and caches
- Build Fast In-memory Aggregates
- Identify Anomalies
- Reduce Storage
Deliver Value

- Continuous Big Data Records
- Real-Time Dashboards
- Business Trends
- Data Patterns
- Outliers
Create a **WAction** to track POS authorizations

**DATA**
ACI BASE24 on HP NSK and/or SQUARE generated request

**PROCESSING**
- GROUP BY merchant_id
- Window based Computation
  - Type, frequency
  - Time, 5 minutes
  - Operator = “SUM”
- Historical Cache
- Compare 10%

**WACTION**

<table>
<thead>
<tr>
<th>Merchant ID</th>
<th>3AB22A11NE44122SGCV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date Time</td>
<td>12-Dec-2012 11:10:10 PM</td>
</tr>
<tr>
<td>Location</td>
<td>123, High St, MN, 55118</td>
</tr>
<tr>
<td>Authorization Count</td>
<td>100</td>
</tr>
<tr>
<td>Historic Count</td>
<td>60</td>
</tr>
</tbody>
</table>

If authorization requests for a merchant within a 5 minute interval deviates from the historical average by 10%, notify the merchant and POS administrator.
This following is for information only and represents WebAction Inc.'s current view of its product development cycle.

Features and release dates are best estimates and should be considered provisional and subject to change without notice.

There can be no guarantee that the release dates will be met or that the product or enhancements will be released at all.

WEBACTION, INC. MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS DOCUMENT.