Real-time Streaming Insight & Time Series Data Analytic

For Smart Retail

Sudip Majumder
Senior Director – Development
Industry IoT & Big Data
## Economic Characteristics of Data

*“Data is the New Oil” …..then “Analytics is the new fuel”*

<table>
<thead>
<tr>
<th>Oil ----&gt; Fuel</th>
<th>Data ----&gt; Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oil is raw and is of little direct use</strong></td>
<td><strong>Data is raw and is of little direct use</strong></td>
</tr>
<tr>
<td><strong>Oil has Potential Energy</strong></td>
<td><strong>Data has Potential Value</strong></td>
</tr>
<tr>
<td><strong>Gas (refined Oil) has more 5x to 10x the Potential Energy than oil</strong></td>
<td><strong>Analytics (refined Data) has more Potential Value than Data</strong></td>
</tr>
<tr>
<td><strong>Burning Gas to create motion converts Potential Energy to Kinetic Energy</strong></td>
<td><strong>Applying data sciences to optimize decisions converts Potential Value to Kinetic Value</strong></td>
</tr>
<tr>
<td><strong>Oil is converted into fuel that powers the economy</strong></td>
<td><strong>Data is converted into Analytics that powers the business</strong></td>
</tr>
</tbody>
</table>
4Vs Meeting 4Cs

**Big Data**
- Volume
- Velocity
- Variety
- Veracity

**IoT**
- Connectivity
- Collection
- Context
- Cognition
## Mapping

<table>
<thead>
<tr>
<th>Connectivity</th>
<th>Collection</th>
<th>Context</th>
<th>Cognition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Volume</strong></td>
<td>Bandwidth to handle data pipeline, embedded capacity</td>
<td>Sheer amount of data generated and collected in sensors, mobile devices, wearable etc.</td>
<td>Sort out data, cleansing, Ingest, extract, transform, load</td>
</tr>
<tr>
<td><strong>Velocity</strong></td>
<td>Network latency, high speed Wi-Fi, Internet connection, HaLow</td>
<td>Fast request/responses or one-way traffic between devices and backend systems</td>
<td>Timely processing, streaming, near real time, seamless scale-out</td>
</tr>
<tr>
<td><strong>Variety</strong></td>
<td>Flexibility on protocols e.g. Http, CoAP, MQTT, Websocket</td>
<td>Different formats of data, diverse devices, proprietary apps</td>
<td>Location data, clickstream, user behavior, motion</td>
</tr>
<tr>
<td><strong>Veracity</strong></td>
<td>Shut down devices or disconnect the network as needed</td>
<td>Shield the noise or bad data</td>
<td>Filter dirty/false data and correlating applicable data for semantics</td>
</tr>
</tbody>
</table>
Some Bold Imperatives:

- Without IoT – Companies React to Failure
- From Sensors to Make Sense
- Radical Departure From Past to Digital Fabric
Complexity of the Consumer Experience Demands Integrated Analytics

How well can you anticipate and serve the needs of your consumer at each step of the shopping process?

**Consumer Actions**
- **Pre Shop**
  - Need
  - Plan and Research
  - Pick Location
- **Shop**
  - Find It in Store
  - Select Product
  - Buy Product
- **Post Shop**
  - Receive Product
  - Use / Consume Product
  - Tell Others

**Retailer Actions**
- Advertising, Digital/Web Marketing, Sunday Circular, Direct Mail, Mobile, Social, Email
- Distribution, Inventory, Assortment, Pricing, Plan-O-Gram, Promotion / Mobile, Store Layout, Brick/Mortar and Web Store Operations, Transactions
- Returns, Customer Service (call center, web, social), Loyalty Program, Digital Engagement

**Supplier Actions**
- Brand Marketing, Shopper Marketing, Advertising, Digital/Web Marketing, Direct Mail, Mobile, Social, Email
- Demand Planning, Pricing, Shipments, DS0, Shelf Position, Trade Promotion, Shopper Marketing
- Consumer Affairs, Social Sentiment Monitoring, Digital Engagement

**Consumer Location**
- At Home
- On the Go
- In the Store

Copyright © 2014 Oracle
Realities of Retail – Inefficiencies and Opportunities

- No integration of data or analytics across functions.
- Lack of measurement standards and common metrics.
- No alignment to the Consumer Experience.
- Pervasive Silos (Online-In Store Merchandising).

- No single view of the customer or inventory.
- Unresponsive supply chain.
- A Consumer Experience held back by legacy systems.
Opportunity: Integrate and Enable New Analytics

- Each function understands its role in serving the consumer.
- Insights and actions directed by a common set of metrics.
- Discovery and collaboration becomes possible across multiple business dimensions.
- IT becomes a strategic enabler.

KEY QUESTION: How do you get started?
High Level Architecture

Oracle IoT Cloud Service

SENSORS
WEATHER
POS
Local Storage

Real Time Streaming

Oracle
DB/Exadata
Cloud Service

HTTP Receiver
SMS Alert
Data Loader

Industry Data Model

Oracle
Integration
Cloud Service
High Level Architecture

Oracle IoT Cloud Service

Real Time Streaming

Local Storage

SENSORS

WEATHER

POS

Real time Capture & Streaming Analytics

Collection
Decoding
Analytics
Publishing

In Memory Oracle No SQL
Big Data

Oracle DB/Exadata Cloud Service

HTTP Receiver
Java
SMS Alert

Data Loader

Industry Data Model

Oracle Integration Cloud Service

Oracle IoT Cloud Service

Data Access & Presentation

Big Data Discovery & Mining
Big Data Analytics
Big Data Discovery & Mining

Web Sockets
Responsive & Material Design

Responsive & Material Design

REST Services

Java

Reporting

Data mining

Oracle Integration Cloud Service

Copyright © 2016, Oracle and/or its affiliates. All rights reserved.
Our Solution – Real Time Analytics Architecture

Production View

**Hybrid Data Store**

- **Raw and Long Term Data Store**
- **Enriched Data Store**
- **In Memory Data Store**

**Data Acquisition**

**Event Processor**

**Smart Query Processor**

**Structured System of Records**

**ORDM (Retail DW)**

**In Database Analytics**

**In Memory Data Store**

**JAVA API to Load DB with Exception and/or Aggr data for DW & Deep Learning**

**Data could be forked to other processes here.**

**Properties Graph Map**

**POS**

**INV**

**LOYALTY**

**CHANNEL**

**CRM**

**SENSORS**

**WEATHER**

**RFID**

**Property Graph & Spatial Reports**

**Real Time Capture, Streaming Analytics & Big Data Discovery**

- Dashboard, Reports
- Metadata in NoSQL
- Web Socket for "Push"
- REST API for “Pull” in-memory and Historical
- Micro Services Example: SMS Alert to Marketing Ops
- Real Time Utilization Reporting
- Customer Purchase Behavior, Service History, CEI Information

**Communication with Customer**

**Customer Experience Data**

**Property Graph Map**

**Structured System of Records**

**Real Time Utilization Reporting**

**REST API for “Pull” in-memory and Historical Data Acquisition**

**Event Processor**

**Smart Query Processor**

**Structured System of Records**

**Real Time Capture, Streaming Analytics & Big Data Discovery**

- Dashboard, Reports
- Metadata in NoSQL
- Web Socket for "Push"
- REST API for “Pull” in-memory and Historical
- Micro Services Example: SMS Alert to Marketing Ops
- Real Time Utilization Reporting
- Customer Purchase Behavior, Service History, CEI Information

**Communication with Customer**

**Customer Experience Data**

**Property Graph Map**

**Structured System of Records**

**Real Time Capture, Streaming Analytics & Big Data Discovery**

- Dashboard, Reports
- Metadata in NoSQL
- Web Socket for "Push"
- REST API for “Pull” in-memory and Historical
- Micro Services Example: SMS Alert to Marketing Ops
- Real Time Utilization Reporting
- Customer Purchase Behavior, Service History, CEI Information

**Communication with Customer**

**Customer Experience Data**

**Property Graph Map**

**Structured System of Records**
Oracle Solution – Real Time Analytics Architecture
Architectural Components Description

- Feed Orchestrator
- Hybrid Data Store
- Smart Query Processor
- Actionable Intelligence Platform
Oracle Solution – Real Time Analytics Architecture

Feed Orchestrator

The Feed Orchestration Layer provides capture and distribution of records received from the data source to downstream processes.

- Designed to support horizontal scale by adding / removing capacity to meet processing demands
- Provides **Lambda Architecture** capability to the solution by enabling writing on multiple types of data stores
- Provides feeds to:
  - Real-time Message Queue/Stream Processing Modules
  - Relational SQL Database Writers for Quick Response Data Storage
  - NoSQL Database Writers for Intermediate Storage and Processing
  - Hadoop Data Store for Cold Storage and Batch Processing
  - External sources such as Wholesale partners
Oracle Solution – Real Time Analytics Architecture

Hybrid Data Store

The Hybrid Data Store consists of:

• Hadoop Data Store
• NoSQL/Real Time Data Store
• Relational Data Store
Oracle Solution – Real Time Analytics Architecture

Hybrid Data Store

The Hybrid Data Store provides Peta scale capability to handle:

• Real Time / Near Real Time Streams
• SQL Data store for Relational data store and self-serve reporting data model
• NoSQL Data store for Schema less data store and transaction level events processing
• Hadoop Data store for Long Term Storage of Raw data / transactions
• Provides flexible linear horizontal and vertical scaling capability - expansion of individual data stores
• Serves both the Batch Layer and Speed Layer in support of a Lambda Architecture
Oracle Solution – Real Time Analytics Architecture

Smart Query Processor

The Smart Query Layer provides the capability to query multi-platform back end data stores using standards-based SQL tools

• Provides singular interface for querying the Hybrid Data Store
• Enables retrieval and merging of query results from Near Real Time feeds and SQL/No-SQL storage
• Provides JDBC Connectivity to downstream applications
• Provides load balancing capabilities to query servers
• Acts as the Serving Layer of the Lambda Architecture
The Actionable Intelligence Platform is composed of 3 modules

- **Real Time Insights**
  - Based on open source / Angular JS
  - Provides capability to query and report on real time feeds, create dynamic dashboards and canned reports
  - Provides flexibility to internal development teams to create rich visualization for specific users or use cases

- **OBIEE / VA (Optional)**
  - Provides Self-Service, Dash boarding and Reporting capabilities
  - Provides pre-built reports and dashboards for use by various Stake Holders, such as Store Ops, Marketing, Finance, Category Management, Loyalty Program, etc.

- **Big Data Discovery (Optional)**
  - Provides the ability to explore, hypothesize and validate the scenarios for critical business problems and potential business opportunities
  - Provides test bed interface to analyze data without impacting production reports
Workflow

Hi speed data extraction

Analytics and Reporting Engine

Presentation

Predictions and decisions

- Logs are collected from POS Machine, RFID tags, IOT devices etc. with utmost efficiency not to hamper the device performance.
- As the data capture size if huge, it needs significant amount of storage. There comes filtering that helps you extract selected data rather than all and save you from running out of storage space.

- Streaming analysis is much faster and near real time.
- Streaming analytics based on Esper CEP engine to generate continuous KPI stream
- Data analytics and decision engine, persistence storage and data mining capabilities built using Apache Hadoop components for horizontal expandability and faster query execution

- Data modeling and reporting built using Hive and R.
- Visualization of outcome through multi-channel presentation layer relevant to individual stakeholders from dashboard view to drill-down capability of correlated transaction tracking
- Responsive UI design

- Identify business prospects
- Foresee challenges
- Logging and surveillance
- Generate innovative business plans and pricing scheme
Key Take-a-Ways

- Like the Internet, **IoT Monetization will occur**
- **Value** delivery will dictate early winners
- IoT Monetization silver bullet is **packaging and pricing flexibility**
- **Real time visibility and agility** are essential to monetization
- Scale goes beyond volume, infrastructure needs to support **business scale**