Simplifying Big Data
SAS Big Data on Oracle Big Data Appliance

April 2016
SAS Global Forum

Mathew Steinberg
Product Management
Oracle Database Development

In association with:
Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle’s products remains at the sole discretion of Oracle.
SAS Institute

• Analytics
  – Big Data Analytics
  – Data Mining
  – Machine Learning
  – Predictive Analytics
  – Statistical Analysis

• Visualization – Explore relevant data, identify drivers
  – SAS Visual Statistics – interactive predictive modeling

• Discovery
  – Point and click access
## Data Management Technology Comparison

**Select the Right Tool for the Job**

<table>
<thead>
<tr>
<th></th>
<th>HDFS</th>
<th>NoSQL</th>
<th>RDBMS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ingest</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Type</td>
<td>Chunk</td>
<td>Record</td>
<td>Transaction</td>
</tr>
<tr>
<td>Write Type</td>
<td>Synchronous</td>
<td>Eventually Consistent</td>
<td>ACID Compliant</td>
</tr>
<tr>
<td>Data Preparation</td>
<td>No Parsing</td>
<td>No Parsing</td>
<td>Parsing and Validation</td>
</tr>
<tr>
<td><strong>DR</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DR Type</td>
<td>Second Cluster</td>
<td>Node Replica</td>
<td>Second RDBMS</td>
</tr>
<tr>
<td>DR Unit</td>
<td>File</td>
<td>Record</td>
<td>Transaction</td>
</tr>
<tr>
<td>DR Timing</td>
<td>Batch</td>
<td>Record</td>
<td>Transaction</td>
</tr>
<tr>
<td><strong>Access</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complex Analytics?</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Query Speed</td>
<td>Slow</td>
<td>Fast for simple questions</td>
<td>Fast</td>
</tr>
<tr>
<td># of Data Access Methods</td>
<td>One (full table scan)</td>
<td>One (index lookup)</td>
<td>Many (Optimized)</td>
</tr>
</tbody>
</table>

- **Affordable Scale**
- **Low Predictable Latency**
- **Flexible Performance**

Copyright © 2015, Oracle and/or its affiliates. All rights reserved.
Oracle Engineered Systems for Data Management

• Big Data Appliance
  Operational Simplicity
  Simplify Access to ALL Data
  Open Analytics Platform
  Full Stack Integration
  – Cloudera CDH Data Hub included

• Exadata
  Ideal Database Hardware
  Smart System Software
  Full-Stack Integration
  Highest Performance, Lowest TCO
SAS Big Data on Big Data Appliance

On Premise or Hosted

- Deploy SAS Analytics on Hadoop nodes
  - SAS Grid: Visual analytics, Scoring, etc
- Reduce Total Cost of Ownership
- Flexible Architectural options for SAS deployments
  - Can run on Starter, Half and Full configurations

SAS Visual Analytics High-Performance Analytic Compute Environment co-located with Hadoop
SAS on Exadata
On Premise or Hosted

• Real Time Enterprise – Faster Applications

• For the lowest TCO in small installations, deploy SAS Analytics in Virtual Machines on Exadata for ‘light consolidation’
SAS Enterprise Analytics “Data Lake” with Oracle Big Data Appliance and Exadata

Big Data Appliance
Hadoop & NoSQL

Exadata
Relational

Oracle Big Data Connectors
Oracle Big Data SQL
Oracle Governance & Security

Execution

Innovation

Data Management
- Data Integration
- Data Quality
- Data Preparation for Hadoop
- Data Governance & MDM
- Event Stream Processing

SAS Analytics
- Data Mining
- Statistical Analysis
- Forecasting
- Text Analytics
- Optimization & Simulation

SAS Visualization
- Visual Data Exploration
- Easy Analytics
- Interactive Reporting & Dashboards
- Collaboration
- Mobile BI

SAS Discovery
- Data Exploration
- Highly visual and/or interactive
- “you don’t know the questions to ask”

Enterprise Analytics Platform
Telco (Asia) – SAS on Oracle Big Data Appliance with Exadata

Challenges

• Increasing need for data
• More data held at more granular levels
• Expansion of Analytics users in the enterprise

Solution

• SAS on Big Data Appliance
• Exadata Database Machine
• Reuse existing workflows
• Retarget output to Hadoop friendly formats
• Removed batch window
• Infrastructure no longer challenged by explosive growth
Major European Bank – SAS on Oracle Big Data Appliance

Challenge
- Dramatic increase in amount of data collected
- Data exploration at massive scale
- No ability for predictive analysis

Solution
- SAS on Oracle Big Data Appliance
- SAS Visual Analytics
  - Intuitive visual analytics
- SAS Visual Statistics
  - Descriptive and Predictive Modeling
  - Model Comparison
  - Dynamic group processing
- SAS in-memory Statistics for Hadoop
  - In-Memory Statistics for Hadoop
  - Programming interface for SAS model development
SAS Enterprise Analytics on Oracle Cloud

- Analytics
- Visualization
- Discovery

VPN / Fast Connect

IaaS / Paas

ORACLE CLOUD

Exadata Cloud Service

ORACLE CLOUD

Big Data Cloud

ORACLE CLOUD
On Premise or in the Oracle Cloud

100% Upward Compatibility with On-Premise
Enables Coexistence and Migration

- Private Cloud
- CoExistence and Migration
- Same Architecture
- Same Standards
- Same Products
- Oracle Cloud
SAS and Oracle **Better** Together

- Analytics + Data Management = Analyze Your Business
- Visualization + Big Appliance = Faster
- Discovery + Exadata = More Nimbly
- + = Lowest TCO
See Oracle at Booth #
Integrated Cloud
Applications & Platform Services