Simplifying Big Data
Best Practices for On-Premises and Cloud Architectures
CON8746

Paul Kent
Vice President Big Data, SAS

Jean-Pierre Dijcks
Master Product Manager Big Data, Oracle

Presented 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.
Big Data Platform options from Oracle

• **Oracle Big Data Appliance**
  – Simple to Deploy
  – Fast out of the Box
  – Lower Cost than DIY Clusters

• **Oracle Big Data SQL**
  – Oracle SQL across ALL your data
  – Secure access to data in NoSQL, Hadoop and Oracle Database
  – High performance on un-modeled data

• **Oracle Big Data Cloud Service**
  – Same architecture as on-premises
  – High performance data processing
  – Enables Big Data SQL in Oracle Public Cloud

• **Oracle Exadata**
  – Best platform for all Oracle Database workloads
  – Unique software that maximizes Oracle Database
  – Standardize, optimized, hardened end-to-end
100% Upward Compatibility with On-Premises Enables Coexistence and Migration
Program Agenda

1. Right Engine – Right Job
2. Right Platform for the Engine
3. Configuring for Mixed Workloads
4. Real-World Examples with SAS
Right Engine – Right Job
A simple set of criteria

Performance

Security

Cost
## High-level Technology Comparison

<table>
<thead>
<tr>
<th>Ingest</th>
<th>HDFS</th>
<th>NoSQL</th>
<th>RDBMS</th>
</tr>
</thead>
<tbody>
<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>
</tbody>
</table>
# High-level Technology Comparison

<table>
<thead>
<tr>
<th></th>
<th>HDFS</th>
<th>NoSQL</th>
<th>RDBMS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Type</strong></td>
<td>Chunk</td>
<td>Record</td>
<td>Transaction</td>
</tr>
<tr>
<td><strong>Write Type</strong></td>
<td>Synchronous</td>
<td>Eventually Consistent</td>
<td>ACID Compliant</td>
</tr>
<tr>
<td><strong>Data Preparation</strong></td>
<td>No Parsing</td>
<td>No Parsing</td>
<td>Parsing and Validation</td>
</tr>
<tr>
<td><strong>DR Type</strong></td>
<td>Second Cluster</td>
<td>Node Replica</td>
<td>Second RDBMS</td>
</tr>
<tr>
<td><strong>DR Unit</strong></td>
<td>File</td>
<td>Record</td>
<td>Transaction</td>
</tr>
<tr>
<td><strong>DR Timing</strong></td>
<td>Batch</td>
<td>Record</td>
<td>Transaction</td>
</tr>
</tbody>
</table>
# High-level Technology Comparison

<table>
<thead>
<tr>
<th></th>
<th>HDFS</th>
<th>NoSQL</th>
<th>RDBMS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Type</strong></td>
<td>Chunk</td>
<td>Record</td>
<td>Transaction</td>
</tr>
<tr>
<td><strong>Write Type</strong></td>
<td>Synchronous</td>
<td>Eventually Consistent</td>
<td>ACID Compliant</td>
</tr>
<tr>
<td><strong>Data Preparation</strong></td>
<td>No Parsing</td>
<td>No Parsing</td>
<td>Parsing and Validation</td>
</tr>
<tr>
<td><strong>DR Type</strong></td>
<td>Second Cluster</td>
<td>Node Replica</td>
<td>Second RDBMS</td>
</tr>
<tr>
<td><strong>DR Unit</strong></td>
<td>File</td>
<td>Record</td>
<td>Transaction</td>
</tr>
<tr>
<td><strong>DR Timing</strong></td>
<td>Batch</td>
<td>Record</td>
<td>Transaction</td>
</tr>
<tr>
<td><strong>Complex Analytics?</strong></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Query Speed</strong></td>
<td>Slow</td>
<td>Fast for simple questions</td>
<td>Fast</td>
</tr>
<tr>
<td><strong># of Data Access Methods</strong></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**
Right Platform for the Engine
Big Data Appliance X5-2 Hardware

• Sun Oracle X5-2L Servers, with each
  – 2 * 18 core Intel Xeon E5-2699 v3
  – 128GB DDR4 Memory
  – 96TB SAS Disk Space

• 3 InfiniBand switches for all internal Hadoop and Spark traffic

• 1 Cisco Management Switch

Note: Support for Cloudera CDH Data Hub included
Hadoop Hardware – Pricing Trends with BDA (List)
Hardware Profiles for Hadoop

Storage
- Dense drives provide flexibility as well as low cost per TB
- HDFS supports dense drives with more scalability enhancements in CDH 5.5
- **Recommendation:** Storage is cheap, ignore it for sizing

CPU
- “Hadoop” nodes are increasingly running mixed workloads and “external processes”:
  - HDFS
  - ETL & Stream Processing
  - Machine Learning
- **Recommendation:** Size CPUs per node towards the high-end core counts

Memory
- Mixed workloads and external processes do need memory
- Large DIMMs do still command a premium in 2-socket servers
- **Recommendation:** Size memory to the mid-spectrum due to hockey stick effects. This is changing!
Hardware Profiles for Hadoop – Conclusion

• Stop worrying about disk space. It is cheap and available.
  – SSD? Once real data tiering is available, or for NoSQL Databases

• Worry instead about:
  – Network
    • Expensive and often the forgotten bottleneck
    • Think about data ingest and the required capacity
  – CPU
    • Mixed workloads will require more CPUs within the node
    • Everyone wants to run compute at the same time
  – Memory
    • Be conscious of actual needs and restrictions
    • Not all SW can make use of lots of memory

Be wary of myths and lore. The hardware world has changed...
Instead, do the math and calculate the bottlenecks and usages you expect
Virtualization Profiles for Hadoop

CPU

Disk
Virtualization Profiles for Hadoop

Performance / Data Storage

Retain “direct attached disks” for Hadoop workloads

Goal: Data permanently lives in the “Hadoop nodes”

Hypervisor overhead is typically low, but the IO overhead has impact, which is why SR-IOV is applied to ensure high throughput.
Virtualization Profiles for Hadoop

Flexibility / Application Hosting

Goal: Spin up environments and the data quickly
Drive towards flexibility at the expense of performance
Recommendations and Future?

• If you don’t need them, don’t use them
  – Remember you will need to apply OS updates to all VMs, and when they have data, this may not be desirable

• Don’t create many small VMs, rather create large VMs
  – Pay attention to the number of VMs on a physical node, and consider what happens if the physical node goes offline (data loss?)

• Futures
  – Docker?
Configuring for Mixed Workloads
Sharing and Managing Resources

Private Cloud

• Note: this is not about “security / isolation” but instead about multiple workloads (tenants) on a single cluster
  – Hence we are ignoring VMs

• Two main ways (outside of VM/Docker) of preventing resource grabs:
  – Linux Control Groups (cgroups)
  – Yet Another Resource Negotiator (YARN)
YARN

• **CapacityScheduler**
  – Designed for multiple tenants
  – Based on resource queues
  – Enables administrators to distribute resources to apps and users

• **FairScheduler**
  – Designed to average out resources per application
  – Does support queues, but as a minimum guarantee
  – Lower maintenance, but less control

The ResourceManager has two main components: Scheduler and ApplicationsManager.
Why cgroups?

• YARN does require developers to write applications in a specific way
  – Sometimes it is quicker and simpler to not adhere to YARN

• Sometimes the requirements are not overly complex
  – cgroups enables “guarantees” by “hard partitioning” of the nodes at the Linux level
  – cgroups are static but create a simple to administer and robust separation of resources
Real World Examples with SAS
4 Segments

- SAS and Oracle Partnership
- Data Lake Dream
- Customer #1 – Telco (Asia)
- Customer #2 – Banking

- Things you might not know...
Big Data – Dream IT. Build IT. Realize IT

Paul Kent, Vice President, Big Data, SAS
Andy Mendelsohn, Senior Vice President, Database Server Technologies, Oracle Corporation
Maureen Chew, Oracle Corporation
Gary Granito, Oracle Corporation
488-2013
<table>
<thead>
<tr>
<th><strong>SAS AND ORACLE</strong></th>
<th>WORKING TOGETHER TO CREATE CUSTOMER VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Joint R &amp; D development and Product Management teams in Cary and Redwood Shores</td>
<td></td>
</tr>
<tr>
<td>• Focus on driving SAS technology components to run natively in Oracle database</td>
<td></td>
</tr>
<tr>
<td>• Joint performance engineering optimizations</td>
<td></td>
</tr>
<tr>
<td>• Template physical architectures developed based on use-cases</td>
<td></td>
</tr>
<tr>
<td>• Physically tested and benchmarked together</td>
<td></td>
</tr>
<tr>
<td>• Reduction in physical effort</td>
<td></td>
</tr>
<tr>
<td>• Overall reduction in lifecycle costs</td>
<td></td>
</tr>
<tr>
<td>• Best Practice papers</td>
<td></td>
</tr>
<tr>
<td>• SAS and Oracle Engineers provide joint &quot;Sizing and Architecture Analysis and Design&quot;</td>
<td></td>
</tr>
</tbody>
</table>
Data Lake Dream
Future State: Data Lake Pattern

- Credit Card Transactions
- Demand Deposits Data
- Credit Applications Data
- Mortgage Data

ETL

- Finance
  - Forecasting
  - Profitability
  - Asset Tracking
  - Others
    - Federal

- Regulatory
  - SEC
  - Compliance Reporting
  - Others

- Risk
  - Client Scanning
  - Operational Exposure
  - Commercial
  - Consumer
  - Money Laundering
  - Check Fraud
  - Loan Fraud

- Fraud
  - Other

- Regulatory
  - Federal
  - SEC
  - Compliance Reporting
  - Others

- Fraud
  - Other

- DataMart

- SAS
- Cognos
- QlikView
- Query Tool

Math Rack

Analysts
3 Stage Refinement

/source
/card
/checking
/mortgage

/transform
/customer
/product

/data
SAS BIG DATA ON BIG DATA APPLIANCE

- Flexible Architectural options for SAS deployments
  - Can run on Starter, Half and Full configurations

- Optionally select nodes “N, N-1, N-2, …” for additional SAS Services such as SAS Compute Tier, SAS MidTier

- Optionally select node subset “N, N-1, N-2, N-3, …) for more dedicated resources for SAS Analytic Compute Environment by shifting Big Data Appliance roles

- Option to selectively add more memory on a per node basis depending on specific workload distribution
STARTER BDA

SAS VISUAL ANALYTICS, HIGH-PERFORMANCE ANALYTIC COMPUTE ENVIRONMENT CO-LOCATED WITH HADOOP
FULL RACK BDA

SAS VISUAL ANALYTICS, HIGH-PERFORMANCE ANALYTIC COMPUTE ENVIRONMENT CO-LOCATED WITH HADOOP
Customer #1
Telco Provider (Asia)
**TELCO (ASIA) CHALLENGES**

**THE SITUATION...**

- Increasing need for Data
- More data held at more granular levels
- Expansion of Analytics users in the enterprise
TELCO (ASIA) BIG DATA ARCHITECTURE IMPLEMENTATION APPROACH

Phase 0
ADM & ABT DI Jobs (ETL SAS Target)

Phase 1
ADM & ABT DI Jobs (ETL Hadoop Target)

In Hadoop Processing
ADM DI Jobs using Sqoop (Hive)
ABT DI Jobs (ELT Hadoop Target) (Pass Through + Explicit SQL + Hive Optimization)

Phase 2
SAN Storage

EDW
FDM
Net Cube
EXADATA

BDA - CLOUDERA HADOOP

ADM (Hive)
ABTs (Hive)
**TELCO (ASIA) DATA MANAGEMENT ON HADOOP**

**SAMPLE JOB FLOW - ADM**

**Phase 0**
- Source Table ➔ Extract Data (Implicit SQL to partitioned table) ➔ Table Loader (Proc Append) ➔ Target Table (SAS)

**Phase 1**
- Source Table ➔ Extract Data (Implicit SQL to partitioned table) ➔ Table Loader (Proc Append) ➔ Target Table (Hive)

**Phase 2**
- Source Table ➔ Sqoop Data (Oozie & Sqoop Import) ➔ Target Table (Hive)
## TELCO (ASIA) DATA MANAGEMENT ON HADOOP

OPTIMIZED USING HADOOP SQOOP+OOZIE VIA CUSTOM SAS DI SQOOP TRANSFORM

<table>
<thead>
<tr>
<th>Subject Area</th>
<th>Table</th>
<th>Load Duration in hours (pre-optimization)</th>
<th>Load Duration, in hours (post-optimization)</th>
<th>Improvement (in hours)</th>
<th>% Improvement</th>
<th>Delta Load Aggregation Level</th>
<th>Delta Load Data Size (in GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytic Data Mart (ADM)</td>
<td>Fact Table 1</td>
<td>0.7000</td>
<td>0.2383</td>
<td>0.4617</td>
<td>65.95%</td>
<td>DAILY</td>
<td>5.60</td>
</tr>
<tr>
<td></td>
<td>Fact Table 2</td>
<td>1.9500</td>
<td>0.5031</td>
<td>1.4469</td>
<td>74.20%</td>
<td>WEEKLY</td>
<td>18.50</td>
</tr>
<tr>
<td></td>
<td>Fact Table 3</td>
<td>3.8167</td>
<td>1.1969</td>
<td>2.6197</td>
<td>68.64%</td>
<td>MONTHLY</td>
<td>37.20</td>
</tr>
<tr>
<td></td>
<td>Fact Table 4</td>
<td>6.0667</td>
<td>0.6214</td>
<td>5.4453</td>
<td>89.76%</td>
<td>DAILY</td>
<td>7.00</td>
</tr>
<tr>
<td></td>
<td>Fact Table 5</td>
<td>0.5333</td>
<td>0.0617</td>
<td>0.4717</td>
<td>88.44%</td>
<td>DAILY</td>
<td>1.70</td>
</tr>
<tr>
<td></td>
<td>Fact Table 6</td>
<td>0.2500</td>
<td>0.0528</td>
<td>0.1972</td>
<td>78.89%</td>
<td>DAILY</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Fact Table 7</td>
<td>0.4833</td>
<td>0.1011</td>
<td>0.3822</td>
<td>79.08%</td>
<td>WEEKLY</td>
<td>2.20</td>
</tr>
<tr>
<td></td>
<td>Fact Table 8</td>
<td>0.9667</td>
<td>0.2597</td>
<td>0.7069</td>
<td>73.13%</td>
<td>MONTHLY</td>
<td>6.30</td>
</tr>
<tr>
<td></td>
<td>Fact Table 9</td>
<td>5.7000</td>
<td>0.0697</td>
<td>5.6303</td>
<td>98.78%</td>
<td>DAILY</td>
<td>1.70</td>
</tr>
<tr>
<td></td>
<td>Fact Table 10</td>
<td>0.9333</td>
<td>0.1094</td>
<td>0.8239</td>
<td>88.27%</td>
<td>WEEKLY</td>
<td>4.10</td>
</tr>
<tr>
<td></td>
<td>Fact Table 11</td>
<td>1.7333</td>
<td>0.2517</td>
<td>1.4817</td>
<td>85.48%</td>
<td>MONTHLY</td>
<td>8.10</td>
</tr>
<tr>
<td></td>
<td>Fact Table 12</td>
<td>2.0833</td>
<td>0.2033</td>
<td>1.8800</td>
<td>90.24%</td>
<td>DAILY</td>
<td>10.10</td>
</tr>
<tr>
<td></td>
<td>Fact Table 13</td>
<td>4.4000</td>
<td>0.8367</td>
<td>3.5633</td>
<td>80.98%</td>
<td>WEEKLY</td>
<td>39.80</td>
</tr>
<tr>
<td></td>
<td>Fact Table 14</td>
<td>8.1333</td>
<td>2.6169</td>
<td>5.5164</td>
<td>67.82%</td>
<td>MONTHLY</td>
<td>104.80</td>
</tr>
<tr>
<td></td>
<td>Fact Table 15</td>
<td>0.4833</td>
<td>0.0447</td>
<td>0.4386</td>
<td>90.75%</td>
<td>DAILY</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Fact Table 16</td>
<td>1.1000</td>
<td>0.0842</td>
<td>1.0158</td>
<td>92.35%</td>
<td>WEEKLY</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>Fact Table 17</td>
<td>1.3500</td>
<td>0.1008</td>
<td>1.2492</td>
<td>92.53%</td>
<td>MONTHLY</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td>Fact Table 18</td>
<td>0.3833</td>
<td>0.0950</td>
<td>0.2883</td>
<td>75.22%</td>
<td>DAILY</td>
<td>2.40</td>
</tr>
<tr>
<td></td>
<td>Fact Table 19</td>
<td>0.6833</td>
<td>0.3014</td>
<td>0.3819</td>
<td>55.89%</td>
<td>WEEKLY</td>
<td>14.70</td>
</tr>
<tr>
<td></td>
<td>Fact Table 20</td>
<td>4.8333</td>
<td>1.1042</td>
<td>3.7292</td>
<td>77.16%</td>
<td>MONTHLY</td>
<td>50.90</td>
</tr>
<tr>
<td></td>
<td>Dimension Table 1</td>
<td>8.1000</td>
<td>2.9511</td>
<td>5.1489</td>
<td>63.57%</td>
<td>N/A</td>
<td>213.70</td>
</tr>
</tbody>
</table>
• Reuse existing workflows
• Retarget outputs to Hadoop Friendly Formats
• Selectively upgrade processing to Hadoop Optimal

• Batch Window Dragon Slayed!
• Infrastructure no longer challenged by explosive growth
Customer #2
Bank (Europe)
Data exploration at massive scale
Intuitive visual analytics
Descriptive and Predictive Modeling

Model comparison

Dynamic group-by processing
In-Memory Statistics for Hadoop:

Programming interface for SAS model development
THINGS OF NOTE

1. SAS was not licensed for all nodes in the BDA

   - SAS EP jobs scheduled by YARN will have full visibility to data across the cluster
   - SASHDAT (SAS High Performance Data Binary Format) needs data sets to stay on the nodes licensed for SAS; some attention with Hadoop Balancer required.

2. SAS was not licensed for all cores on the nodes it was licensed for

   - SAS Licensing Posture has improved – you can license say 200 cores on a 1000 core cluster and control limits with CGROUPS and YARN
Things you might not know
**SPD Engine** with Hadoop

- Support for running on MapR 4.0.2
- Support for Code Accelerator
- Enhanced WHERE pushdown: AND, OR, NOT, parenthesis, range operators and in-lists
- Parallel write support can improve write performance up to 40%
- Optionally uses Apache Curator/Zookeeper as a distributed lock server. No more physical lock files.

```
libname spdat spde '/user/dodeca' hdfshost=default;
```
Leverage the processing power of the cluster

- SAS runs completely in Hadoop cluster
- YARN with Oozie prioritizes and schedules SAS jobs in cluster
- All existing SAS Grid Manager integrated (e.g. EM) capabilities
- Reduction of SAN storage
- No separate SAS compute tier
ORACLE ENGINEERED SYSTEMS FOR

ExaData
ExaLogic
SuperCluster
Database (Backup, Recovery, Logging Appliance)
ZFS Storage Appliance
Virtual Compute Appliance
Big Data Appliance
Integrated Cloud
Applications & Platform Services