Oracle’s Big Data Strategy

Thomas Kurian
Executive Vice President
Product Development

October 4, 2012
Safe Harbor Statement

"Safe Harbor" Statement: Statements in this presentation relating to Oracle's future plans, expectations, beliefs, intentions and prospects are "forward-looking statements" and are subject to material risks and uncertainties. Many factors could affect our current expectations and our actual results, and could cause actual results to differ materially. We presently consider the following to be among the important factors that could cause actual results to differ materially from expectations: (1) Economic, political and market conditions, including the current European debt crisis, can adversely affect our business, results of operations and financial condition, including our revenue growth and profitability, which in turn could adversely affect our stock price. (2) We may fail to achieve our financial forecasts due to such factors as delays or size reductions in transactions, fewer large transactions in a particular quarter, unanticipated fluctuations in currency exchange rates, delays in delivery of new products or releases or a decline in our renewal rates for software license updates and product support. (3) Our hardware systems business may not be successful, and we may fail to achieve our financial forecasts with respect to this business. (4) We have an active acquisition program and our acquisitions may not be successful, may involve unanticipated costs or other integration issues or may disrupt our existing operations. (5) Our international sales and operations subject us to additional risks that can adversely affect our operating results, including risks relating to foreign currency gains and losses and risks relating to compliance with international and U.S. laws that apply to our international operations. (6) Intense competitive forces demand rapid technological advances and frequent new product introductions and could require us to reduce prices or cause us to lose customers. (7) If we are unable to develop new or sufficiently differentiated products and services, or to enhance and improve our products and support services in a timely manner or to position and/or price our products and services to meet market demand, customers may not buy new software licenses or hardware systems products or purchase or renew support contracts. A detailed discussion of these factors and other risks that affect our business is contained in our SEC filings, including our most recent reports on Form 10-K and Form 10-Q, particularly under the heading "Risk Factors." Copies of these filings are available online from the SEC or by contacting Oracle Corporation's Investor Relations Department at (650) 506-4073 or by clicking on SEC Filings on Oracle’s Investor Relations website at http://www.oracle.com/investor. All information set forth in this presentation is current as of October 4, 2012. Oracle undertakes no duty to update any statement in light of new information or future events.
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 Defined

- Sources: Structured, Unstructured, and Semi-Structured Data
- “Big Data” requires new data management and analytic approaches because:
  - **Volume**: Large volume of low information-density data – requires cheap scale-out storage
  - **Velocity**: High velocity of data capture – requires rapid ingestion
  - **Variety**: Wide variety of rapidly evolving data types – requires highly flexible storage
  - **Value**: Extracting signal from noise – requires scale out query processing
Big Data Defined

- The “Big Data” Market Opportunity is an Analytics opportunity
  - Broader Data Set: Understand what Happened, Why, Predict what will happen
  - Better Analytics: Based on acquiring, capturing, organizing all/most of the relevant data
  - Deeper Analytics: New analytic methods on new kinds of data
## Big Data Defined

### Acquire
- Graph Analytics
- Decisioning
- Statistical Modeling
- Spatial Analytics
- Data Mining
- Query & Analysis
- Information Discovery
- Text Analytics

### Organize
- Text Parsing & Analysis: Various
- Script: Pig; Query: Hive
- Extract-Transform-Load: Sqoop

### Process
- Real Time: Event Processing
- Processing: Map-Reduce
- Metadata: HCatalog

### Store
- Key Value Data: NoSQL
- Columnar Data: HBase
- Raw Data: HDFS

---

**Copyright © 2012, Oracle and/or its affiliates. All rights reserved.**
“Big Data” & Hadoop

- Hadoop: A parallel computing framework and distributed file system
  - HDFS, Map/Reduce, Other Components - Hcatalog, Hbase, Hive, Pig, Oozie, Zookeeper

- Hadoop’s Value: Scalable, Fault Tolerant, Cost Efficient Data Management
  - Hadoop is a core element of Oracle’s Big Data strategy

- But Hadoop alone is not a sufficient solution for “Big Data”
  - Massively Scalable BUT Batch only . . . NO interactive analysis
  - Flexible Programming Model BUT difficult to use efficiently . . . No Optimizer
  - Highly parallel BUT unpredictable response times . . . Due to Flexible Storage Representation
  - Cost effective BUT missing enterprise class reliability . . . Due to Low Information Density data

- Hadoop is a component of a larger Warehousing & Analytics Strategy . . . not a replacement of it
  - e.g., Gartner Logical Data Warehouse model: June 21, 2012
Hardware and Software Engineered to Work Together