
ORACLE®

What's New with Spatial and Graph?

Technologies to Better Understand Complex Relationships

Jim Steiner
Vice President, Product Management
Server Technologies
January, 2016

ORACLE

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.

Spatial and Graph Technology

Graph for Analysis of:

- Social Media relationships
- “Linked data” for large scale integration and reasoning
- Internet of Things interactions

Spatial for Analysis of:

- Understanding behavior and performance as it relates to location
- Data Enrichment
- Proximity and containment analysis

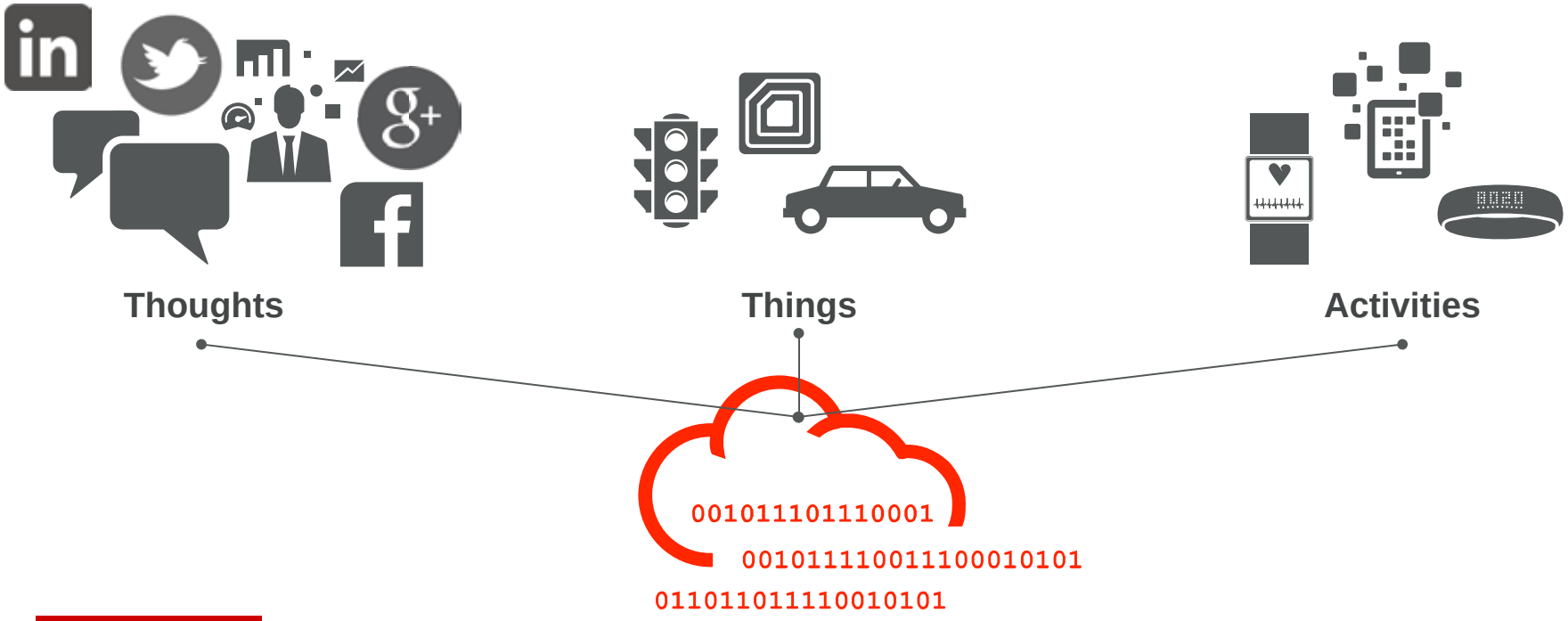
ORACLE

Graph and Spatial Analysis – It is about relationships

- Are things in the same location? Who is the nearest? What tax zone is this in? Where can deliver in 35 minutes? What is in my sales territory? Is this built in a flood zone?
- Which supplier am I most dependent upon? Who is the most influential customer? Do my products appeal to certain communities? What patterns are there in fraudulent behavior?

New Data, New Platforms

The Datafication Of Every“thing”



New Technologies Enable New Architectures

Big Data

- Highly Parallel and In-Memory Database Architectures
- Hadoop
- NoSQL Database Models
  - Key Value, Columnar, Graph Stores

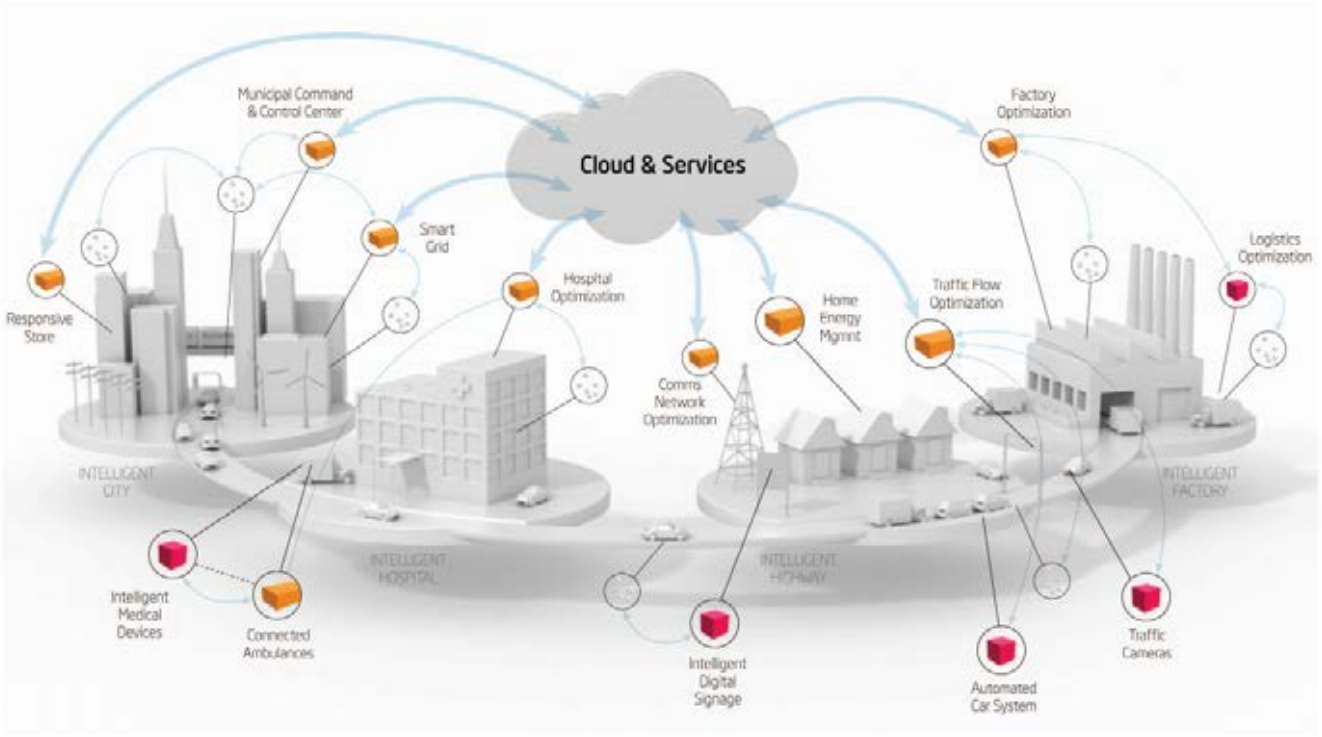
Fast Data

- Stream Processing
- In-memory clusters
  - Apache Spark
- Real time Discovery
- Real time Analytics
  - Hybrid in-memory search / analytic engine

ORACLE

Cloud Architecture

Connecting a System of Systems



The Internet of Things

And changes possibilities (and expectations)

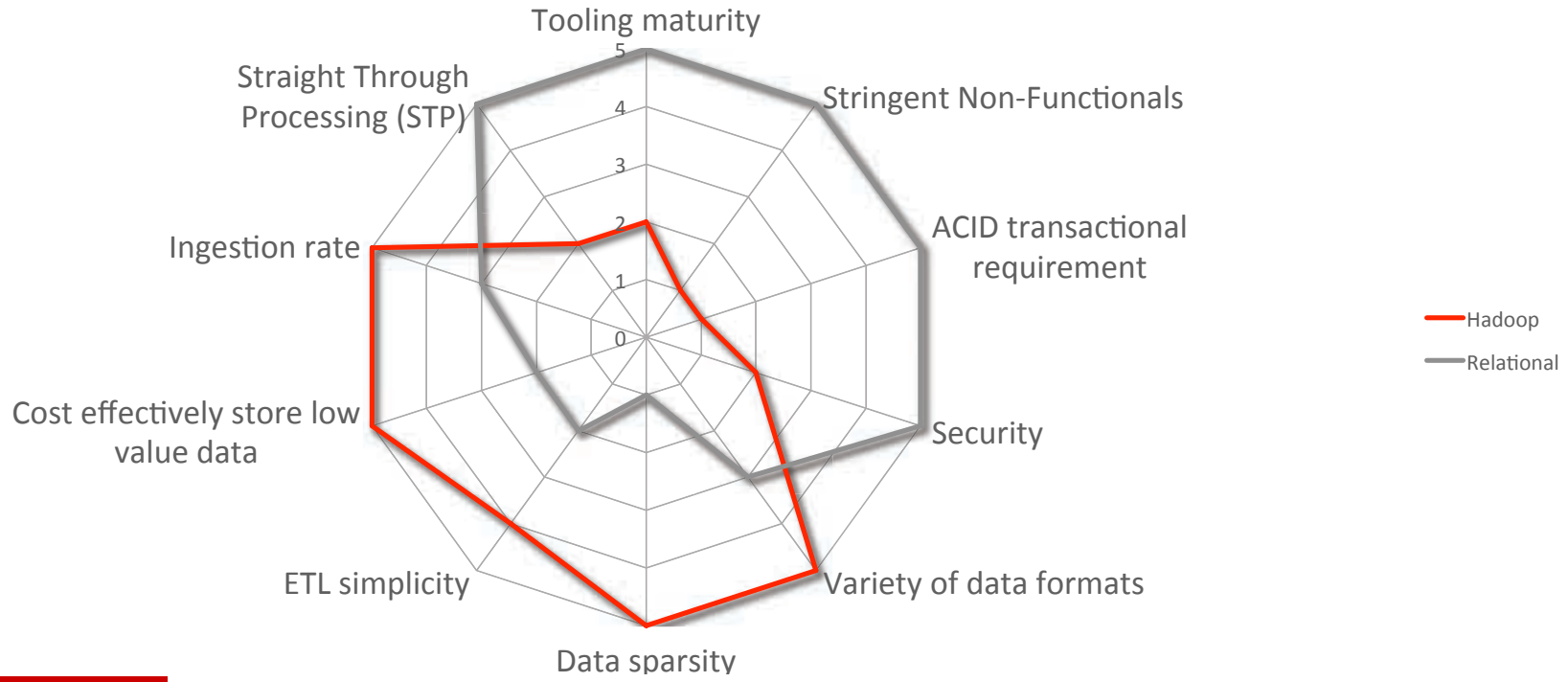
Thing	+	Technology *SW/HW/NW*	=	What the thing always did	+	Value Added Service

New Offerings for Spatial and Graph

Oracle Big Data Spatial And Graph

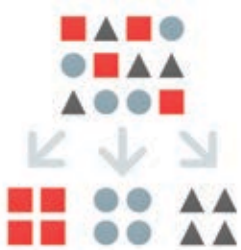
Database or Big Data technologies

Typical technical decision criteria



Big Data Spatial and Graph

Big Data often needs an Organizing principle – Data Harmonization



Big Data analysis is often about relationships not aggregation

Big Data platform is economically compelling for working with massive data sets found in spatial and graph workflows

ORACLE

Oracle Big Data Spatial and Graph (BDSG)

Spatial Analysis Features

Property Graph Database

Who is most important? There Are Lots of Answers.

- Answers from **Aggregation**
  - Who spends the most?
  - Who buys the highest margin goods?
  - Who is most consistently a top contributor?

Tabular questions:
Well-suited to SQL-like tools

- Answers from **Connectivity**
  - Who's most influential?
  - Which supplier do I depend on the most?
  - What is the right product mix for millennials?

Graph questions:
We need something different!

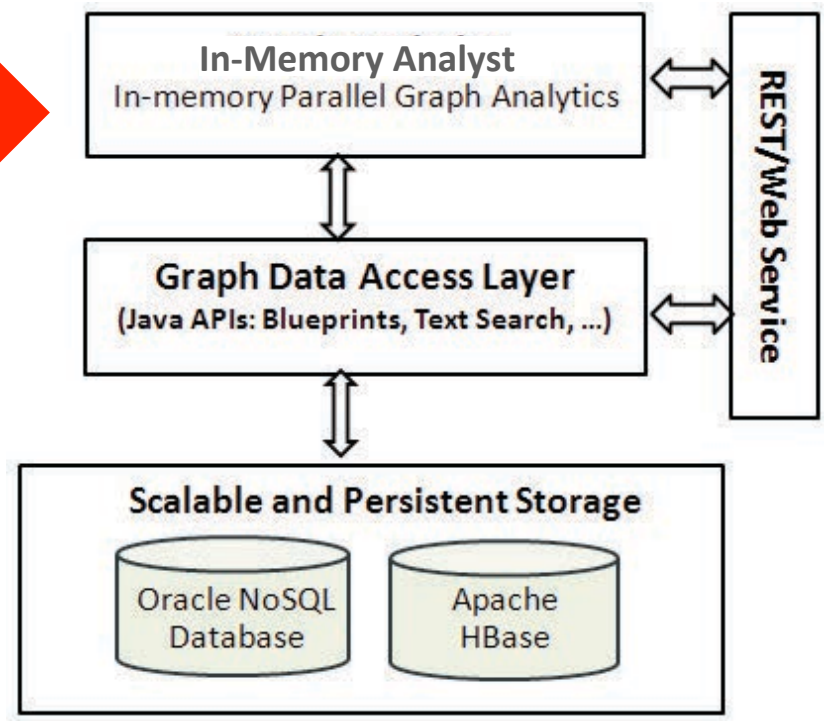
Big Data Graph Architecture

Lightning-Fast In-Memory Analytics

- YARN Container
- Standalone Server
- Embedded

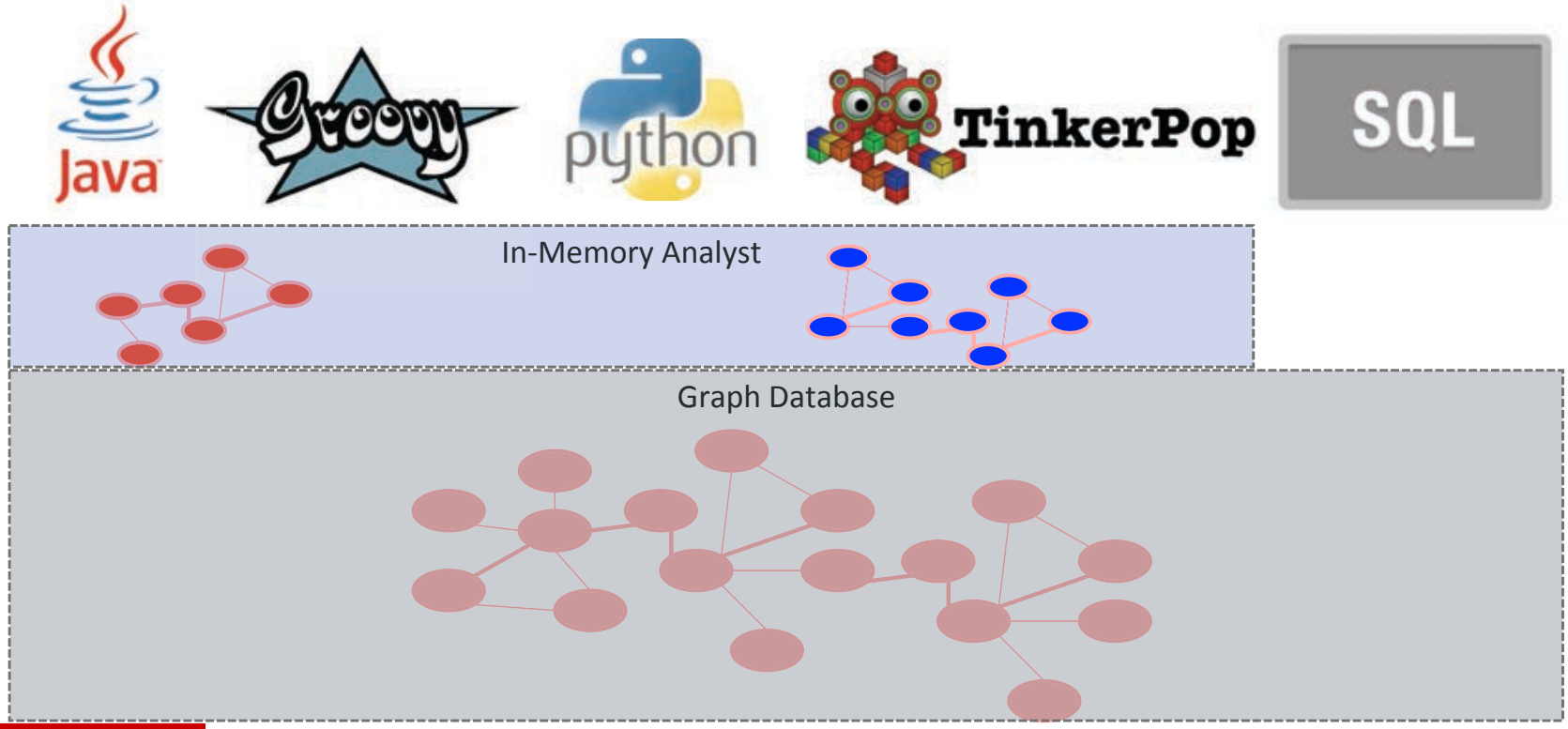
Massively Scalable Graph Store

- Oracle NoSQL
- HBase



ORACLE

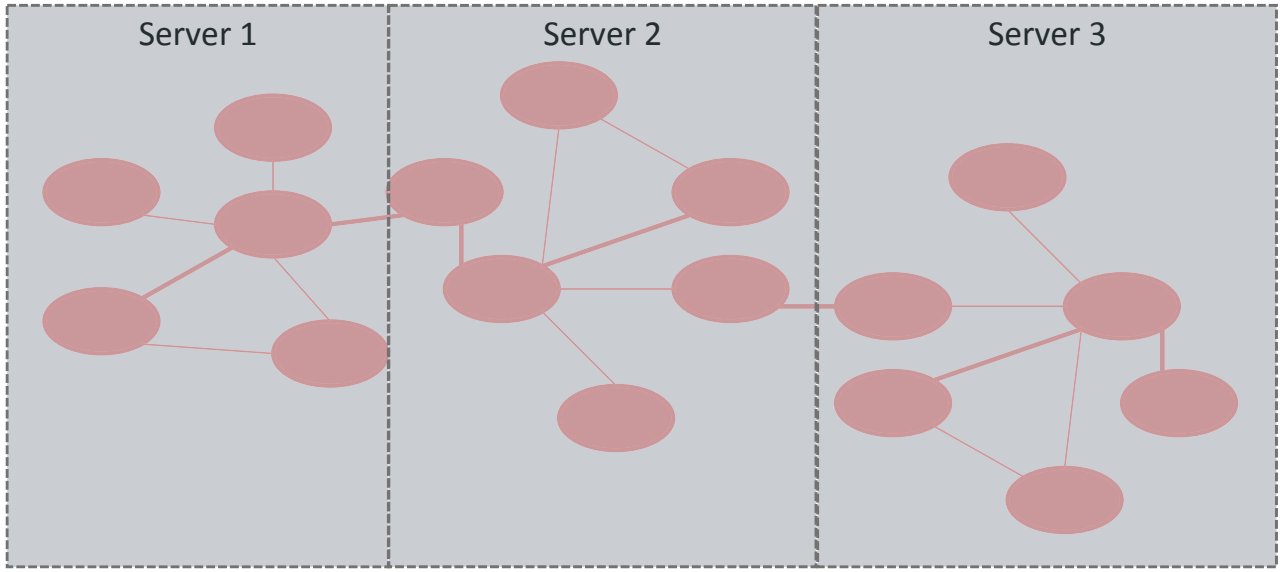
Multiple Interfaces for Many Kinds of Users



ORACLE

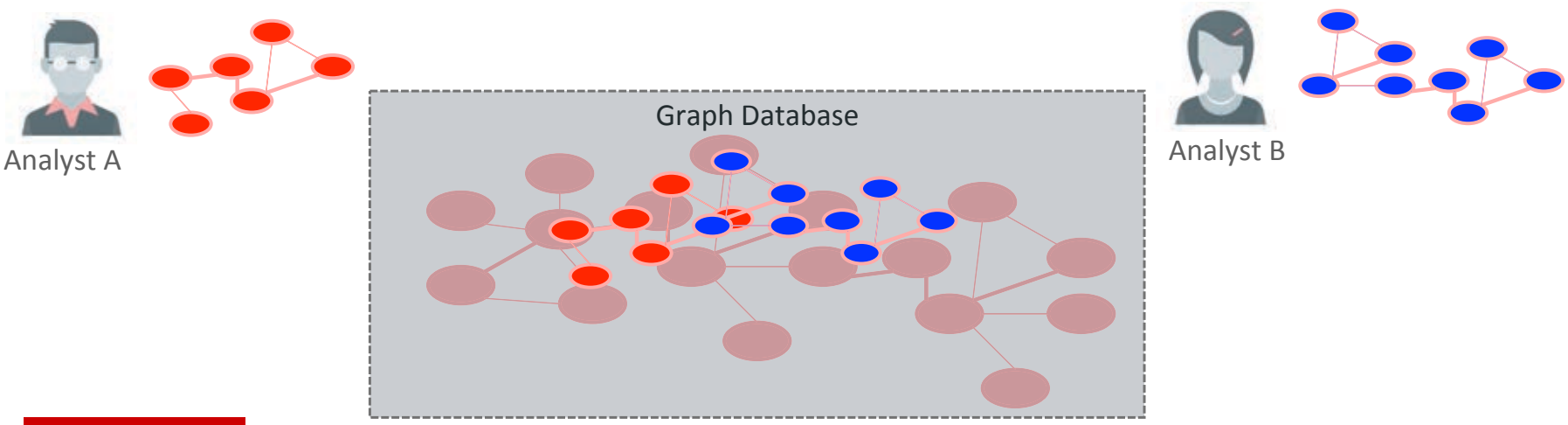
Scales to Trillions of Edges

- Leverage HBase or Oracle NoSQL for massive storage scalability
- Use underlying engine for node/edge access and graph filtering



In-Memory Analyst

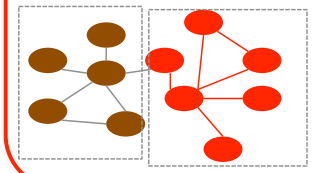
1. Analyst engine runs on one or more servers
2. Users request loads subgraph of interest into engine
3. Users call simple functions for complex analyses
4. Write results back or share on-the-fly



ORACLE

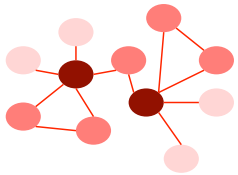
35+ Graph Functions

Detecting Components and Communities



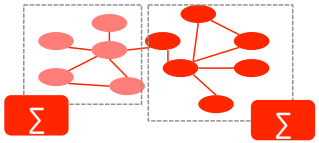
Tarjan's, Kosaraju's,
Weakly Connected Components, Label Propagation (w/ variants), Soman and Narang's

Ranking and Walking



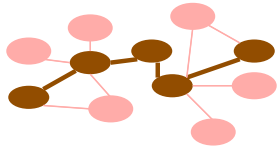
Pagerank, Personalized Pagerank,
Betweenness Centrality (w/ variants),
Closeness Centrality, Degree Centrality,
Eigenvector Centrality, HITS,
Random walking and sampling (w/ variants)

Evaluating Community Structures



Conductance, Modularity
Clustering Coefficient (Triangle Counting)
Adamic-Adar

Path-Finding



Hop-Distance (BFS)
Dijkstra's,
Bi-directional Dijkstra's
Bellman-Ford's

Link Prediction

SALSA
(Twitter's Who-to-follow)

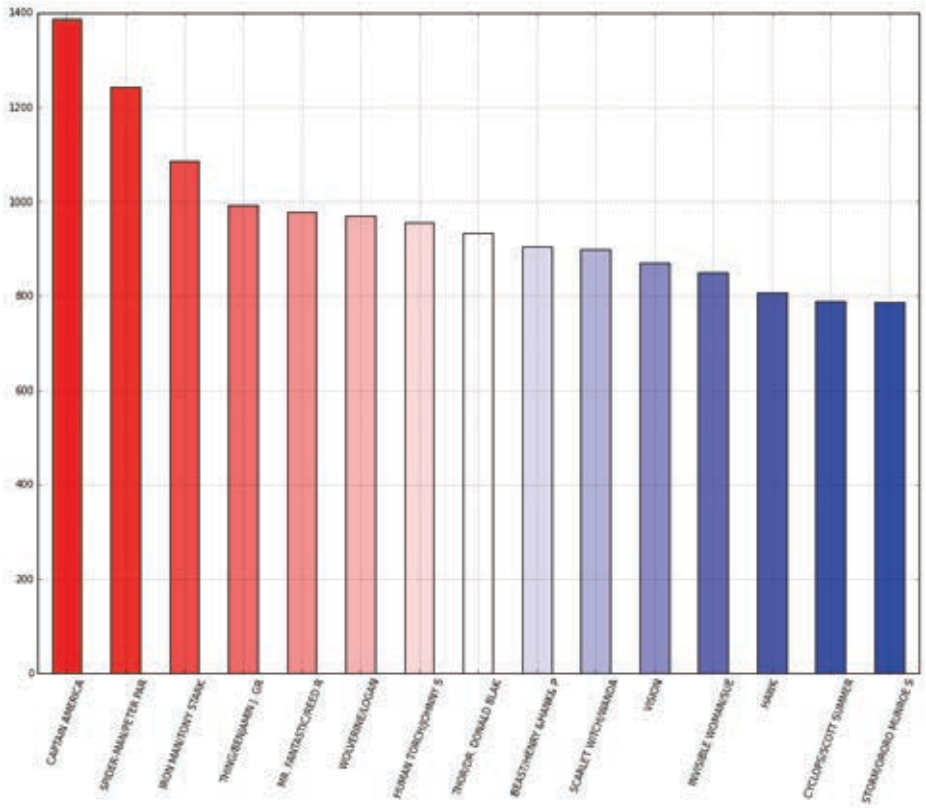
Other Classics

Vertex Cover
Minimum Spanning-Tree(Prim's)

Degree Centrality in Big Data Graph

Code

```
heroInfluence =
analyst.inDegreeCentrality()
```



Differentiators -- Graph

- Commercial, supported software
- "Best of Both Worlds" Graph DB
  - In-memory graph algorithms execution – Like Neo4J
  - Distributed graph database model – Like Datastax Titan
- Dozens of pre-built parallel in-memory graph algorithms
- 10-50x faster analytics than competitors' offerings
- Analyze 20-30 Billion edge graph in memory on a single BDA node

Oracle Big Data Spatial and Graph (BDSG)

Spatial Analysis Features

Property Graph Database

Motivation for Oracle Big Data Spatial

Emergence of Hadoop for spatial analysis in business and spatial workflows

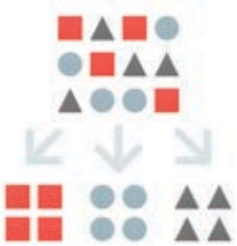
Existing Hadoop-based Spatial technologies are GeoSpatial-centric not Application-centric

No significant commercial offerings

ORACLE

What problems can Big Data Spatial analysis address?

Data Harmonization using any location attribute (address, postal code, lat/ long, placename, etc).



Categorization and filtering based on location and proximity

Preparation, validation and cleansing of Spatial and Raster data

Visualizing and displaying results on a map

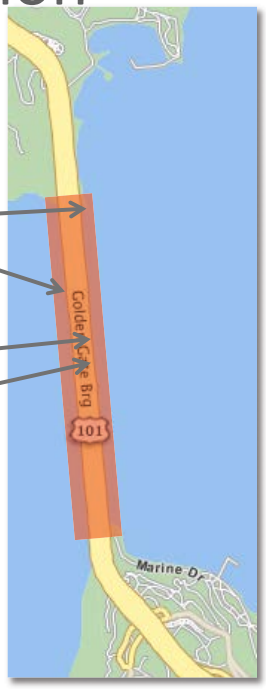
ORACLE

Data Harmonization: Linking information by location

Are these data points related?

- Tweet: sailing by #goldengate
- Instagram image subtitle: 골든게이트 교*
- Text message: Driving on 101 North , just reached border between Marin County and San Francisco County
- GPS Sensor: N 37°49’11” W 122°28’44”

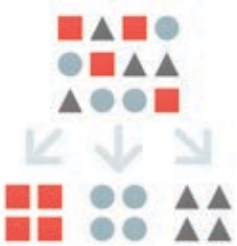
- Now find all data points around Golden Gate Bridge ...



* Golden Gate Bridge (in Korean)

What features does Big Data Spatial have?

Data enrichment service API using GeoNames and geometry hierarchy data



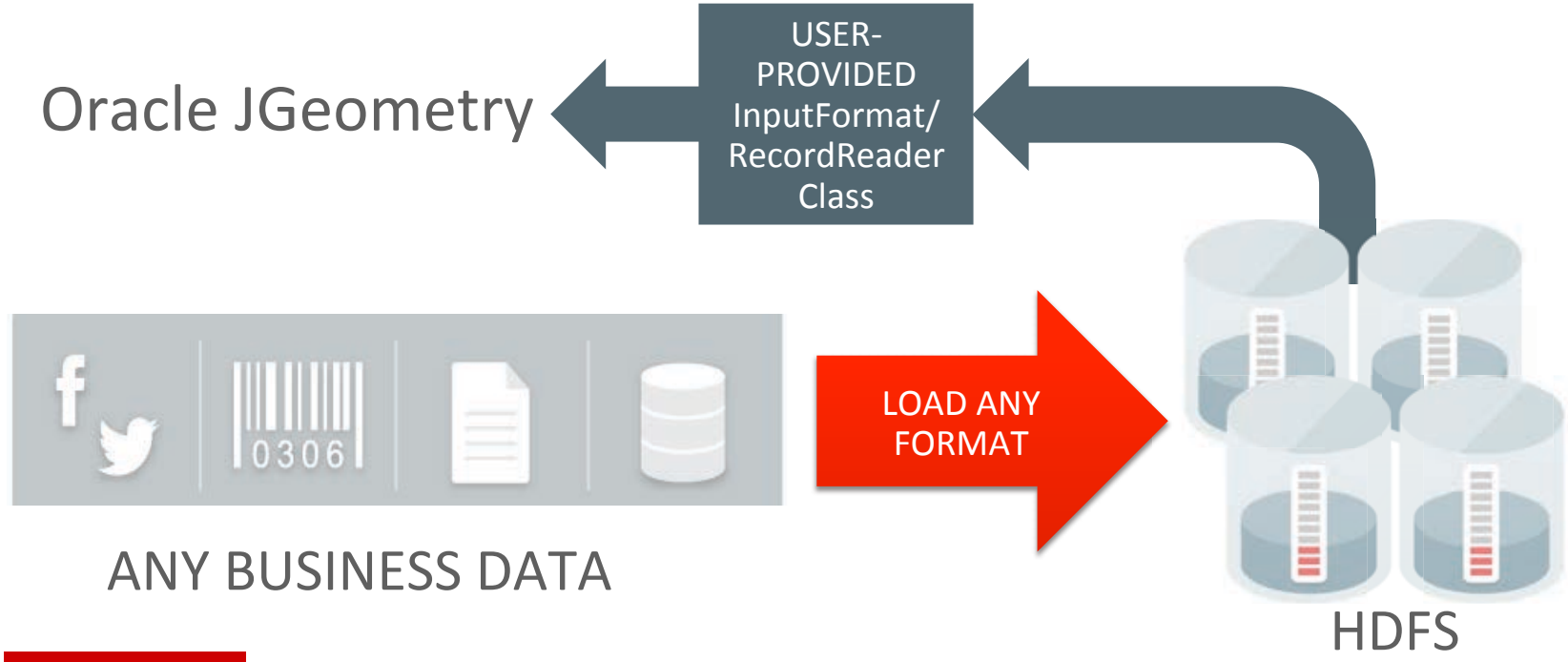
MapReduce routines for distance calculations, PointInPolygon, buffer creation, Categorization, KMeansClustering, Binning, etc.

Spatial processing of data stored in HDFS. Raster processing operations: Mosaic and sub-set operations. Geodetic and Cartesian data

HTML5 Map Visualization API

ORACLE

Store any business data with spatial information in HDFS



ORACLE

Vector Data Processing API Functions

Single Geometry

- Length
- Area
- Buffer
- Simplify

Geometry Pairs

- Range Queries
  - Point in Polygon
  - Touch, Overlap, Intersect, Contains, Any Interaction
- Join Queries
  - Interactions on sets of data
  - E.g.: Find all the dropped cell calls in all coverage areas

Categorization and Enrichment

- Associate a data set with a known geometry or named hierarchy
  - Process all Tweets for a period of time and count how many are associated with each city, county, state, etc.

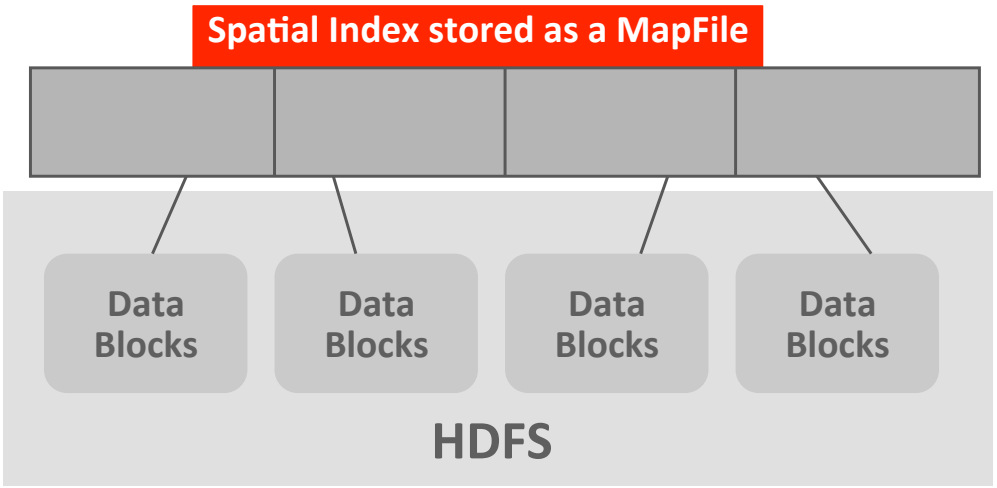
Data Categorization Services

Any hierarchical geometry data set for reference

Customers choose a set of layers For example, they can select (continents, countries, cities) or (countries, states, counties) as the hierarchy

Big Data Spatial map-reduce job processes the customer data and produces a result file

Spatial Index for Spatial Queries



MapReduce Job with Index

Copy the index to distributed cache

Mapper reads the index data for the corresponding HDFS block

Process only those records that return hits from the index search

Big Data Spatial and Graph

Spatial Vector Processing Framework

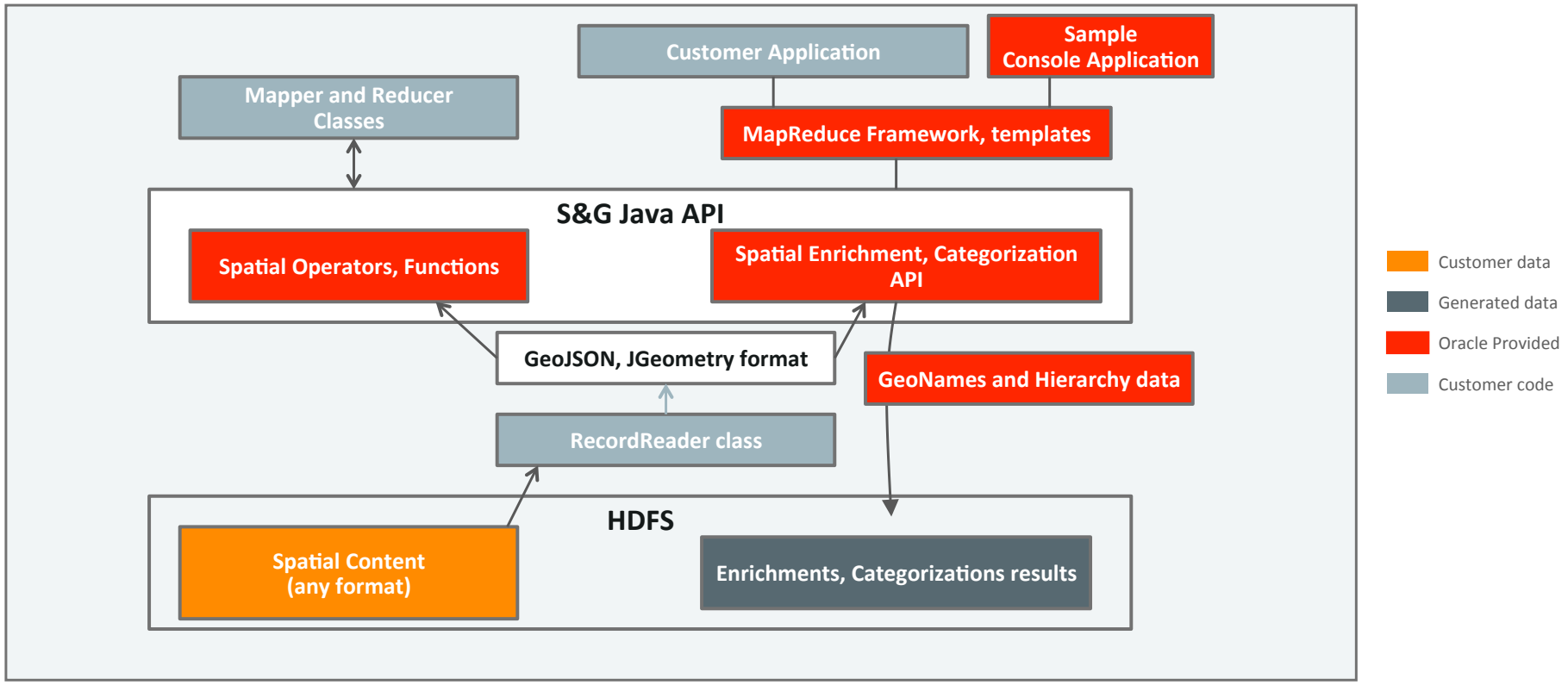
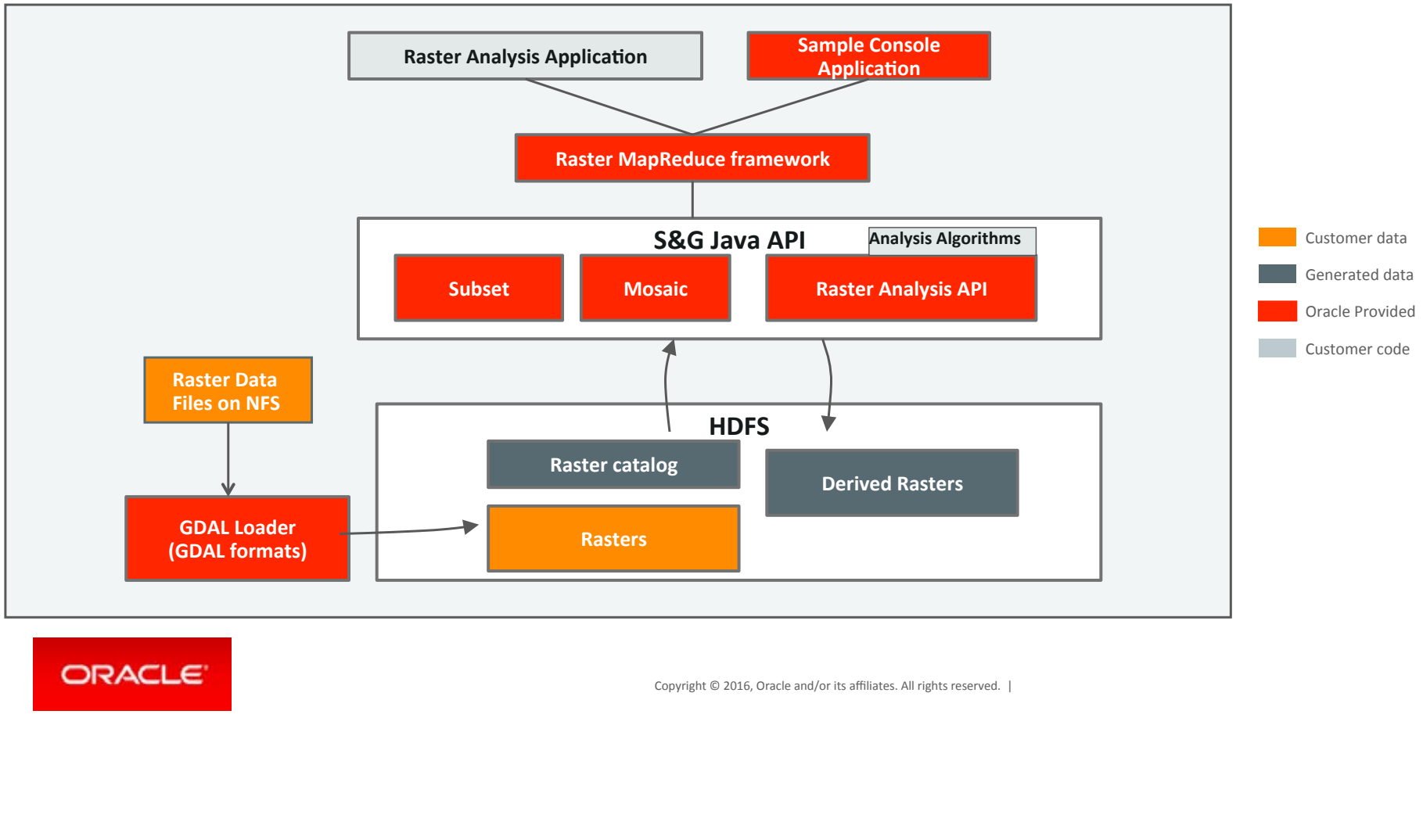


Image Server

- HDFS storage for the image or raster files
  - We can support dozens of file formats (GDAL supported formats)
  - Images are geo-referenced
  - Images can be in different coordinate systems and resolutions
- Three main capabilities
  - Loader to load raster data from NFS to HDFS
  - Mosaic and subset operations based on a virtual mosaic
  - Image processing framework for raster analysis

Big Data Spatial and Graph

Spatial Raster Processing Framework



Differentiators -- Spatial

- Commercial, supported software
- Application-centric approach vs. GIS-centric
  - Works on any data that includes location info, datatype and file format
- GeoEnrichment services including global geographic hierarchy
- Supports both spatial processing and spatial analytics
- Very rich set (~50) of spatial operators and functions
- Both Vector and Raster services
- Includes map visualization

Automated Database Cloud Service – Full Instance Editions

Enterprise Edition (EE)

adds…

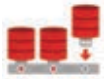
- Transparent Data Encryption (TDE)
- Data Guard
- All standard EE features

Standard Edition

- Full database instance
- Up to 16 OCPUs

High Performance

adds…


Multitenant


Partitioning


Advanced Compression


Advanced Security, Label Security, Database Vault

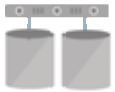

Real Application Testing


OLAP, Analytics, Spatial and Graph


Management Packs

Extreme Performance

adds…


Real Application Clusters (RAC)

In Memory

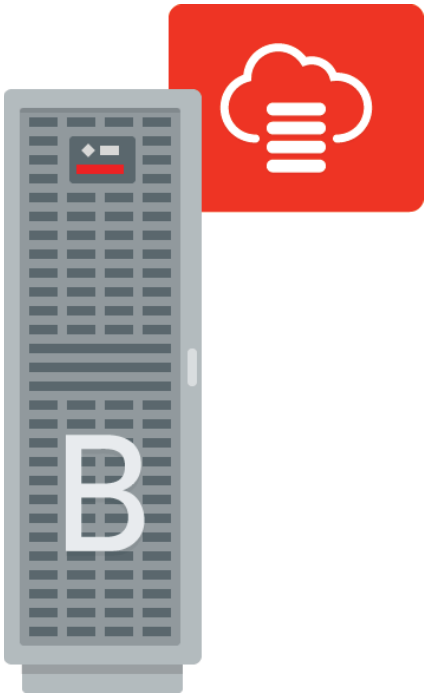

Active Data Guard

Reference: http://www.oracle.com/us/products/database/enterprise-edition/comparisons/index.html

Greater Capabilities

ORACLE

Oracle Big Data Cloud Service – Now Production!



- Fast, reliable and secure service on Oracle Big Data Appliance
- Fully automated service for Hadoop and Spark
- Available in scalable units of:
  - 216 cores
  - ¾ TB memory
  - 96 TB disk storage

Oracle Big Data Cloud Service

Key Features

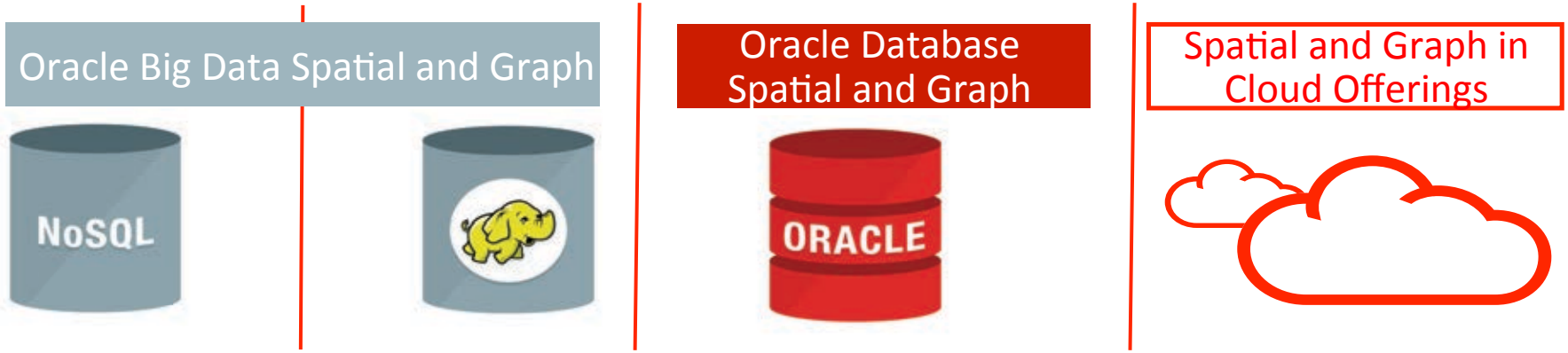
- Big Data (Hadoop, Spark) as an Automated Service
  - Cloudera Enterprise – Data Hub Edition 5.4
  - Oracle Big Data Connectors
  - Oracle Big Data Spatial and Graph
  - Oracle Data Integrator with Advanced Big Data Option
  - Database Cloud Service integration (via Connectors)
- Big Data SQL service add-on [Coming Soon]
  - Unified query across Big Data and Exadata Cloud Services
- Platform for new Big Data Services
  - Big Data Discovery [Coming Soon]

Benefits

- Consistently high performance
- Integration with Exadata Cloud Service means one fast SQL query on all your data

Oracle's Spatial and Graph Strategy

Enable Spatial and Graph use cases on every platform



The Spatial and Graph SIG

- The SIG promotes interaction and communication that can drive the market for spatial technology and data
- Members connect and exchange knowledge via online communities and at annual conferences and events

- Meet us at the Summit
  - Morning Receptions
    - Tuesday and Wednesday / 7:45 to 8:30 a.m. / Registration Area
  - Birds of a Feather Session
    - Wednesday / 12 to 1 p.m. / Auditorium – Look for "Spatial and Graph SIG" table
- Join us online
  - LinkedIn (search for "LinkedIn Oracle Spatial")
  - Google+ (search for "Google+ Oracle Spatial")
  - IOUG SIG (sign up for free membership through www.ioug.org)
  - OTN Spatial – Communities (search for "Oracle Spatial and Graph Community")
- Contact the Board
  - oraclespatialsig@gmail.com

ORACLE


ORACLE®