Bring Location Analysis to Your Big Data Workloads Using Oracle Big Data Spatial and Graph

Big Data Showcase Theater

Jean Ihm        David Lapp

Oracle Spatial and Graph Technologies
September 2016
Agenda

1. Why Location Analysis for Big Data
2. Use Cases
3. Features
4. Demo
Spatial Analysis – It is about relationships

• Are things in the same location? Who is the nearest? What tax zone is this in? Where can we deliver in 35 minutes? What is in my sales territory? Is this built in a flood zone?
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
Big Data Challenges

Requires more development resources and data scientists

Build your own environment from commodity hardware and open source software.
Big Data Solution

Make developers and data scientists more productive

Pre-built, parallel, MapReduce and Spark Spatial Algorithms

Raster and Vector processing Frameworks

Optimized, pre-configured Big Data Appliance and Cloud platforms
A Complete Big Data Portfolio

Oracle Data Integrator
Oracle GoldenGate
Oracle Big Data Preparation
Oracle Big Data SQL
Oracle Big Data Connectors
Oracle NoSQL Database

Oracle BI EE + Data Visualization
Oracle Big Data Discovery
Oracle Big Data Spatial and Graph
Oracle Advanced Analytics
Oracle Data-as-a-Service
Oracle Applications
Oracle Big Data Appliance
Oracle Big Data Cloud
Oracle Public Cloud
Oracle Big Data Spatial and Graph
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)
Insurance Industry

Use Case: Linking Information by Location

86%
Of Insurance companies agree that analyzing multiple data sources together is crucial to making accurate predictions

88%
Agree that linking information by location is key to combining disparate sources of Big Data

Source: “The big data: How data analytics can yield underwriting gold.” Survey conducted by Ordnance Survey and Chartered Insurance Institute, 25 April 2013.
GeoSocial Analysis Use Cases: “Nearest Friends”

“It’s 11:30. Want to meet Jon, Melli, and Albert for lunch @Milano’s today at Noon?”

URGENCY WITH LIMITED AVAILABILITY: Lunch promotion to targeted potential “table of 4” who know each other within 1 km.

“We know you and your old college buddies love Elton John. Get together with Tom, Dick and Hari and the rest of the frat next month.”

BROADER SOCIAL REACH; WIDER DISTANCE. NO TIME CONSTRAINT: Concert promotion.

“Great seats @ Cinema 18 for 7:30 show of new Avengers movie tonite. Free popcorn and soft drink for you and Mary. Text her at 555-1234.”

TIME CONSTRAINED: Target people within 20 minute drive 30 minutes before the show.
What problems can Big Data Spatial analysis address?

- **Data Harmonization** using any location attribute (address, postal code, lat/long, placename, etc).
- **Preparation, validation and cleansing** of Spatial and Raster data.
- **Visualization and displaying results on a map**.
- **Categorization and filtering** based on location and proximity.
- **Spatial querying and analysis** of Hadoop data with SQL.
### What features does Big Data Spatial have?

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data enrichment service API using GeoNames and geometry hierarchy data</td>
<td>MapReduce routines for distance calculations, PointInPolygon, buffer creation, Categorization, KMeansClustering, Binning</td>
</tr>
<tr>
<td>Spatial processing of data stored in HDFS or NoSQL. Raster processing operations: Mosaic and sub-set operations. Geodetic and Cartesian data</td>
<td>HTML5 Map Visualization API</td>
</tr>
<tr>
<td></td>
<td>Hive SQL API Query from Oracle DB with Big Data SQL &amp; Oracle SQL Connectors for Hadoop</td>
</tr>
</tbody>
</table>
Demo

Discovering location patterns in social media and global incidents

Using Big Data Spatial Analytics and Big Data Discovery
Summary
Oracle Big Data Spatial and Graph

• Commercial, supported software

• Componentry to boost efficiency of data scientists and developers – save time on custom development

• Bring location context to big data – harmonize and offer new insight into customers, assets, organizations

• Extend existing Big Data applications, complement relational environments

• Oracle Big Data Spatial and Graph offers
  – Dozens of built-in algorithms/enrichment services, and map visualization
  – Scaleable storage and parallel processing on Hadoop /Oracle NoSQL
  – Runs on commodity hardware or BDA, both on-premise or in the Cloud
<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monday, Sept 19</strong></td>
<td><strong>8:30 a.m. - 10:30 a.m.</strong> Building a Java Recommender System in 15 Minutes with Graph Technologies [TUT5289]</td>
<td>Parc 55 - Cyril Magnin II/III</td>
</tr>
<tr>
<td></td>
<td><strong>4:15 p.m. - 5:00 p.m.</strong> Graph and Link Analysis: Discovering Network Relationships in Big Data [CON6445]</td>
<td>Park Central - Olympic</td>
</tr>
<tr>
<td></td>
<td><strong>4:00 p.m. - 6:00 p.m.</strong> Apply Location Intelligence and Spatial Analysis to Big Data with Java -- Hands On Lab [HOL6255]</td>
<td>Hilton - Franciscan Room C/D</td>
</tr>
<tr>
<td><strong>Tuesday, Sept 20</strong></td>
<td><strong>1:00 p.m. – 1:20 p.m.</strong> Bring Graph Analysis to Relational and Hadoop Data [THT7821]</td>
<td>Moscone South Exhibition Hall - Big Data Theater</td>
</tr>
<tr>
<td></td>
<td><strong>4:00 p.m. – 5:00 p.m.</strong> High-Speed Video and Image Processing with Java and Hadoop [CON5543]</td>
<td>Parc 55 - Mission</td>
</tr>
<tr>
<td></td>
<td><strong>6:15 p.m. - 7:00 p.m.</strong> Meet the Experts: Oracle’s Big Data Management System [MTE7224]</td>
<td>Moscone South - 306</td>
</tr>
<tr>
<td></td>
<td><strong>7:15 p.m. - 8:00 p.m.</strong> Meet the Experts: Spatial and Graph Technologies for Database, Big Data, and the Cloud</td>
<td>Moscone South - 306</td>
</tr>
</tbody>
</table>
## Big Data Spatial and Graph at OOW 2016


<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thursday, Sept 22</td>
<td>Build Applications on Spark, Hadoop, and NoSQL with Oracle Big Data Spatial and Graph [CON6585]</td>
<td>Park Central - Olympic</td>
</tr>
<tr>
<td>10:45 a.m. - 11:30 a.m</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Showcase Theater Presentations

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuesday, Sep 20</td>
<td>Bring Graph Analysis to Relational and Hadoop Data</td>
<td>Big Data Showcase Theater</td>
</tr>
<tr>
<td>1:00 - 1:20 p.m.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wednesday, Sep 21</td>
<td>High Speed Video and Image Processing Using Oracle Big Data Platform</td>
<td>Big Data Showcase Theater</td>
</tr>
<tr>
<td>1:00 - 1:20 p.m.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wednesday, Sep 21</td>
<td>Bring Location Analysis to Your Big Data Workloads Using Spatial</td>
<td>Big Data Showcase Theater</td>
</tr>
<tr>
<td>2:30 - 2:50 p.m.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Big Data Spatial and Graph at OOW 2016


### Demos

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday - Wednesday</td>
<td>Oracle’s Spatial Technologies for Database, Big Data, and the Cloud</td>
<td>Moscone South Exhibition Hall – 101 Oracle Database Showcase (left side)</td>
</tr>
<tr>
<td>Monday - Wednesday</td>
<td>Oracle’s Graph Database and Analytics for Database, Big Data, and the Cloud</td>
<td>Moscone South Exhibition Hall – 101 Oracle Database Showcase (left side)</td>
</tr>
<tr>
<td>Monday - Wednesday</td>
<td>Discover and Analyze Relationships with Oracle Big Data Spatial and Graph</td>
<td>Moscone South Exhibition Hall Big Data Showcase (SBD-016)</td>
</tr>
</tbody>
</table>

### Partners

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday – Wednesday</td>
<td>Tom Sawyer Software</td>
<td>Moscone South Exhibition Hall - Booth #436</td>
</tr>
<tr>
<td>Monday – Wednesday</td>
<td>DataNinja (DocomoInnovations)</td>
<td>Moscone South Exhibition Hall - Booth #1741</td>
</tr>
<tr>
<td>Monday – Wednesday</td>
<td>interRel Consulting</td>
<td>Moscone South Exhibition Hall - Booth #3208</td>
</tr>
</tbody>
</table>
Resources

• Oracle Big Data Spatial and Graph OTN product page: www.oracle.com/technetwork/database/database-technologies/bigdata-spatialandgraph
  – White papers, software downloads, documentation and videos

• Oracle Big Data Lite Virtual Machine - a free sandbox to get started: www.oracle.com/technetwork/database/bigdata-appliance/oracle-bigdatalite-2104726.html

• Hands On Lab for Big Data Spatial: tinyurl.com/BDSG-HOL

• Blog – examples, tips & tricks: blogs.oracle.com/bigdataspatialgraph

• @OracleBigData, @SpatialHannes, @JeanIhm

• LinkedIn: Oracle Spatial and Graph Group