Open source Geospatial Business Intelligence in action with GeoMondrian and SOLAPLayers!

FOSS4G 2010

Dr. Thierry Badard, CTO
Spatialytics inc.
Quebec, Canada
tbadard@spatialytics.com

Spatialytics
Genuine Geospatial Business Intelligence

Barcelona, Spain – Sept 9th, 2010
What are GeoMondrian & SOLAPLayers?

- It is part of the geospatial BI software stack developed initially by the GeoSOA research group at Laval University in Quebec …
  - GeoKettle
  - GeoMondrian
  - SOLAPLayers

- But are now developed and supported by Spatialytics
  - [http://www.spatialytics.org](http://www.spatialytics.org) (open source community)
  - [http://www.spatialytics.com](http://www.spatialytics.com) (professional support, training)

- OK but … what is geospatial BI? ;-)

- [Spatialytics.org](http://www.spatialytics.org)
As you probably know …

- Business Intelligence applications are usually used to better understand historical, current and future aspects of business operations in a company.
- The applications typically offer ways to mine database- and spreadsheet-centric data, and produce graphical, table-based and other types of analytics regarding business operations.
- They support the decision process and allow to take more informed decision!
Data visualization to support decision ...
As you probably know …

- Business Intelligence applications are usually used to better understand historical, current and future aspects of business operations in a company.
- The applications typically offer ways to mine database- and spreadsheet-centric data, and produce graphical, table-based and other types of analytics regarding business operations.
- They support the decision process and allow to take more informed decision!
- Rely on an architecture with robust components and applications:
  - ETL tools & data warehousing (DW)
  - On-line Analytical Processing (OLAP) servers and clients
  - Reporting tools & dashboards
  - Data mining
Classical architecture of a BI infrastructure

- Transactional databases
- Web resources
- XML, flat files, proprietary file formats (Excel spreadsheets, …)
- LDAP
- …
The Data Warehouse: the crucial/central part!

- Repository of an organization’s historical data, for **analysis purposes**.
- Primarily destined to analysts and decision makers.
- Separate from operational (OLTP) systems (source data)
  - But often stored in relational DBMS: Oracle, MSSQL, PostgreSQL, MySQL, Ingres, …
- Contents are often presented in a summarized form (e.g. key performance indicators, dashboards, OLAP client applications, reports).
  - Need to define some metrics/measures
The Data Warehouse: the crucial/central part!

- Optimized for:
  - Large volumes of data (up to terabytes);
  - Fast response (<10 s) to analytical queries (vs. update speed for transactional DB):
    - de-normalized data schemas (e.g. star or snowflake schemas),
      - Introduces some redundancy to avoid time consuming JOIN queries
    - all data are stored in the DW across time (no corrections),
    - summary (aggregate) data at different levels of details and/or time scales,
    - (multi)dimensional modeling (a dimension per analysis axis).
      - All data are interrelated according to the analysis axes (OLAP datacube paradigm)
- Focus is thus more on the analysis / correlation of large amount of data than on retrieving/updating a precise set of data!
- Specific methods to propagate updates into the DW needed!
MDX query language

- MDX stands for MultiDimensional eXpressions
- Multidimensional query language
- *De facto* standard from Microsoft for SQL Server OLAP Services (now Analysis Services)
- Also implemented by other OLAP servers (Essbase, Mondrian) and clients (Proclarity, Excel PivotTables, Cognos, JPivot, …)
- MDX is for OLAP data cubes what SQL is for relational databases
- Looks like a SQL query but relies on a different model (close to the one used in spreadsheets)

```sql
SELECT
    { [Measures].[Store Sales] } ON COLUMNS,
    { [Date].[2002], [Date].[2003] } ON ROWS
FROM Sales
WHERE ( [Store].[USA].[CA] )
```
Results representation

- SELECT
  
  \{ [Product].[All Products].[Drink],
  [Product].[All Products].[Food] \} ON COLUMNS,
  
  \{ [Store].[All Stores].[USA].[WA].[Yakima].[Store 23],
  [Store].[All Stores].[USA].[CA].[Beverly Hills].[Store 6],
  [Store].[All Stores].[USA].[OR].[Portland].[Store 11] \} ON ROWS

  FROM Warehouse

  WHERE ([Time].[1997], [Measures].[Units Shipped])

- OLAP client software propose:
  - Alternate representation modes (pie charts, diagrams, etc.)
  - Different tools to refine queries/explore data
    - Drill down, roll up, pivot, …
    - Based on operators provided by MDX
Geospatial BI adds maps and spatial analysis!

Require to consistently integrate the geospatial component in all parts of the architecture!
Why merge BI and GIS software?

• Because …

“About eighty percent of all data stored in corporate databases has a spatial component” [Franklin 1992]

Why merge BI and GIS software?

- Enable the exploration of spatial relations between data
  - To take into account all aspects of data
  - And then take informed decisions

- Some phenomena can only be observed and interpreted by representing them on a map!
  - Spatial distribution,
  - Spatiotemporal evolution,
  - etc.
To implement true geo-analytical tools ...
Pentaho open source BI software stack

- [http://www.pentaho.org](http://www.pentaho.org)
Pentaho open source BI software stack

- Pentaho ([http://www.pentaho.org](http://www.pentaho.org))

+ CDF: Community Dashboard Framework
+ Other projects: olap4j, JPivot, Halogen, …
Spatialytics open source geospatial BI stack

- Spatialytics ([http://www.spatialytics.com](http://www.spatialytics.com))

& integration in various dashboard and reporting tools

- PostGIS
- Oracle Spatial

Spatialytics ETL systems

Data sources (OLTP systems)
GeoMondrian

- GeoMondrian is a "spatially-enabled" version of Pentaho Analysis Services (Mondrian)
- GeoMondrian brings to the Mondrian OLAP server what PostGIS brings to the PostgreSQL DBMS
  - i.e. a consistent and powerful support for geospatial data.
- Licensed under the EPL
- http://www.geo-mondrian.org
GeoMondrian

• As far as we know, it is the first implementation of a true Spatial OLAP (SOLAP) Server
  - And it is an open source project! ;-)

• Provides a consistent integration of spatial objects into the OLAP data cube structure
  - Instead of fetching them from an external spatial DBMS, web service or a GIS file

• Implements a native Geometry data type

• Provides first spatial extensions to the MDX language
  - Add spatial analysis capabilities to the analytical queries

• At present, it only supports PostGIS datawarehouses
  - But other DBMS will be supported in the next version!
Spatially enabled MDX

- Goal: bring to Mondrian and MDX what SQL spatial extensions do for relational DBMS (i.e. Simple Features for SQL and implementations such as PostGIS).

- Example query: filter spatial dimension members based on distance from a feature
  
  
  ```sql
  SELECT
      {[Measures].[Population]} on columns,
      Filter(
          {[Unite geographique].[Region economique].members},
          ST_Distance([Unite geographique].CurrentMember.Properties("geom"),
                       [Unite geographique].[Province].[Ontario].Properties("geom")) < 2.0
      ) on rows
  FROM [Recensements]
Spatially enabled MDX

• Many more possibilities:
  - in-line geometry constructors (from WKT)
  - member filters based on topological predicates (intersects, contains, within, …)
  - spatial calculated members and measures (e.g. aggregates of spatial features, buffers)
  - calculations based on scalar attributes derived from spatial features (area, length, distance, …)
GeoMondrian

- Demo -
SOLAPLayers

- SOLAPLayers is a lightweight cartographic component (framework) which enables navigation in geospatial (Spatial OLAP or SOLAP) data cubes, such as those handled by GeoMondrian.

- It aims to be integrated into existing dashboard frameworks in order to produce interactive geo-analytical dashboards.

- Such dashboards help in supporting the decision making process by including the geospatial dimension in the analysis of enterprise data.

- First version stems from a GSoC 2008 project performed under the umbrella of OSGeo.

- Licensed under BSD (client part) and EPL (server part).

- http://www.solaplayers.org
SOLAPLayers v1

- Version 1 was based on OpenLayers and Dojo
- It allows:
  - the connection with a Spatial OLAP server such as GeoMondrian,
  - some basic navigation capabilities in the geospatial data cubes,
  - and the cartographic representation of some measures as static or dynamic choropleth maps, maps with proportional symbols.
SOLAPLayers v1

- Demo -
SOLAPLayers v1

- Version 1 was a mostly proof of concept!
- It presents important limitations:
  - Allows only the cartographic representation (no crosstabs or charts)
  - Works only for one measure and the spatial dimension!
  - Offers limited navigation capabilities in the geospatial data cubes
  - Is able to connect to GeoMondrian only
  - Extending the framework is difficult due to the lack of flexibility and the poor documentation of Dojo,
  - Integration with other currently used geo-web and dashboard frameworks was difficult
  - ...

So, SOLAPLayers has undergone (and is still undergoing ;-) ) a deep re-engineering!

Version 2 is fully based on ExtJS/GeoExt (and hence OpenLayers)
- It will make its integration with other geo/web and BI/dashboards frameworks easier
- It provides some new ExtJS components dedicated to GeoBI!
- Based on the philosophy for the development of applications adopted by these geo-web frameworks, it allows an easier creation/maintenance of the produced geo-analytical dashboards!
- Like ExtJS, it supports internationalization!
SOLAPLayers 2.0 – Architecture

1. SOLAP Server
   - Native or XML/A

2. OLAP Server
   - Native or XML/A

- Built-in or LDAP Server

- Authentication

- MDX

- SOLAPJSON

- Client

- Geospatial data source (WFS, DBMS, ...)

- Spatialalytics.org SOLAP Server

- Spatialalytics.org Map Component

- Open Source Geospatial Business Intelligence
SOLAPLayers 2.0 – Architecture

1. GeoMondrian
   SOLAP Server
   • Native or XML/A

2. OLAP Server
   • Native or XML/A

Built-in or LDAP Server

Authentication

OLAP4J

MDX

SOLAPJSON

Client

Bridge architecture

- Maximize what is in place in organisations
- But, no Geo-MDX capabilities available!
SOLAPLayers 2.0 – Architecture

1. SOLAP Server
   - SOLAP Server
   - Built-in or LDAP Server
   - Authentication

2. OLAP Server
   - OLAP Server
   - OLAP4J
   - MDX
   - SOLAPJSON

3. Geospatial data source (WFS, DBMS, ...)
   - Geospatial data source (WFS, DBMS, ...)

Native or XML/A
SOLAPLayers 2.0 – Architecture

1. SOLAP Server
   - GeoMondrian
   - Native or XML/A

2. OLAP Server
   - Native or XML/A

3. Geospatial data source (WFS, DBMS, ...)

Built-in or LDAP Server

Authentication

MDX

SOLAPJSON

Client

- For simple geo-dashboards
- Based on transactional data
- Thematic mapping
- No Geo-MDX and drill-down or roll-up capabilities!
Define the template of the dashboard in a HTML file

```html
<!-- SOLAPLayers lib. -->
<script type="text/javascript" src="SOLAPLayers/solaplayers.js"></script>

<!-- SOLAPLayers dashboard demo -->
<script type="text/javascript" src="dashboard3.js"></script>

</head>
<body>
<div id="leftdiv" style="width: 49%; float: left;">
  <div id="querypanel"></div>
  <div id="mappanel" style="padding-top: 20px;"></div>
</div>
<div id="rightdiv" style="width: 49%; float: right;">
  <div id="tablepanel" style="padding-bottom: 20px;"></div>
  <div id="chartpanel" style="padding-bottom: 20px;"></div>
</div>
</body>
</html>
```
SOLAPLayers 2.0 – Geo-dashboard made easy!

1 Define the template of the dashboard in a HTML file

```html
<!-- SOLAPLayers lib. -->
<script type="text/javascript" src="SOLAPLayers/solaplayers.js"></script>

<!-- SOLAPLayers dashboard demo -->
<script type="text/javascript" src="dashboard3.js"></script>

</head>
<body>
<div id="leftdiv" style="width: 49%; float: left;">
  <div id="querypanel"></div>
  <div id="mappanel" style="padding-top: 20px;"></div>
</div>
</div>
<div id="rightdiv" style="width: 49%; float: right;">
  <div id="tablepanel" style="padding-bottom: 20px;"></div>
  <div id="chartpanel" style="padding-bottom: 20px;"></div>
</div>
</body>
</html>
```

2 Define your dashboard components in a JS file and map it to the div in the HTML file

```javascript
chartPanel = new SOLAPLayers.dashboard.ColumnChartComponent({
mrs: mrs,
title: 'ChartPanel',
renderTo: chartpanel,
width: 750,
height: 500
});
```
SOLAPLayers 2.0 – Geo-dashboard made easy!

Enjoy! ;-)
SOLAPLayers 2.0

- Demo -
SOLAPLayers – Sum up & roadmap

- As GeoExt which provides Geospatial extensions to ExtJS, SOLAPLayers provides GeoBI extensions to ExtJS.
- So, to make it simple: **SOLAPLayers = GeoBIExt!**
- At present, it provides the main components for creating geo-analytical dashboards
  - Map, crosstab, column chart, line chart, ...
  - But, many more to come and to develop!
    - Cube explorer, query builder, time slider/navigator, gauges, score cards, social graphs, ...
    - Advanced interaction capabilities and settings will be added to each components!
    - Additional thematic mapping capabilities are also required: multi-maps, ...
- Beta of version 2.0 to be released by the end of October 2010
- We anticipate to have a first stable version in January 2011
Questions?

- Thanks for your attention and do not hesitate to ask for more demos and to contact us for possible collaborations!

- Contact:
  
  Dr. Thierry Badard, CTO  
  Spatialytics inc.  
  Quebec, Canada  
  Email: tbadard@spatialytics.com  
  Web: http://www.spatialytics.org  
  http://www.spatialytics.com  
  Twitter: tbadard & spatialytics

- GeoKettle: http://www.geokettle.org  
  Twitter: geokettle

- GeoMondrian: http://www.geo-mondrian.org  
  Twitter: geomondrian

- SOLAPlayers: http://www.solaplayers.org  
  Twitter: solaplayers