Using OGC Services to Interoperate Spatial Data Stored in SQL and NoSQL Databases

Claudio de Souza Baptista, Odilon Francisco de Lima Junior, Maxwell Guimaraes de Oliveira, Fabio Gomes de Andrade, Tiago Eduardo da Silva, Carlos Eduardo Santos Pires
Introduction
Due to the large volume of data generated on the Internet today, new forms of data storage and data processing are required.

Traditional architectures of DBMS for storing structured data have proven inadequate to deal with this enormous volume of data.
Introduction
Introduction

The popularity of spatial applications and also the supporting devices for geo-referenced data gathering have all contributed to an increase in this information volume.
For these applications, NoSQL databases provide distributed storage and indexing techniques using map/reduce functions.
Introduction

- For these applications, NoSQL databases provide distributed storage and indexing techniques using map/reduce functions.
Problem:
- combine this volume of georeferenced information that emerges from social networks with the traditional geo-referenced information, stored, for example, in SQL spatial databases or in spatial data infrastructures.
Solution:

- implement OGC interoperability services among spatial data on the NoSQL spatial database so as to integrate it to the SQL spatial DBMS.
The SANY project [Havlik et al. 2009] designed a service to provide a single point of access to data spread across the various nodes of a network of sensors.
The ORCHESTRA project [Usländer 2007] describes a spatial data infrastructure for risk management applications.
Related Work

- Map/Reduce functions have been used to carry out a number of tasks in the geographical domain:
  - Generation of spatial indexes [Akdogan et al. 2010][Cary et al. 2009]
  - Query processing [Jardak et al. 2010]
  - Prediction of natural disasters [Hasenkamp et al. 2010]
The architecture of our system was developed on three layers:

- Application
- Services
- Persistence
Architecture

Application Layer

Service Layer (WMS/WFS)

Persistence Layer

Client

1. NoSQL Map Server

2. SQL Map Server

NoSQL Spatial Database

SQL Spatial Database
Application Layer

- Responsible for the interaction between users and services.
- In our prototype, we used GEO-STAT (Geographic Spatio-Temporal Analysis Tool), a Web map viewer that we have developed.
Service Layer

- Makes possible to submit queries to a NoSQL database with the same syntax used to query a SQL database in a simple and transparent manner for the user’s application.
Service Layer
Persistence Layer

- We used CouchDB-GeoCouch for NoSQL data.
- CouchDB
  - Document-oriented schema free database
  - A database is stored as a collection of documents JSON (JavaScript Object Notation)
  - All interaction is performed entirely by using the HTTP protocol through a RESTful interface
Case Study - Setup

- We configured two servers (for SQL and NoSQL databases); each providing WMS and WFS services.

- Both servers stored spatial records about the Brazilian state of Paraíba, including all its 223 municipalities, the highways that cross the state, and all fire outbreaks detected in the state in 2010.
Case Study - Experiment 1

- Goal: Checking OGC requests placed to WMS and WFS services

- Functional tests were performed on the implemented NoSQL Services accessing the NoSQL database server.

- The result set was compared to GeoServer in order to validate the accuracy of our implementation.
Case Study - Experiment 1
Case Study - Experiment 1
Goal: Evaluation of interoperability between the NoSQL and the SQL databases
Case Study - Experiment 2

- In the first server, based on SQL, we left available only data related to municipalities and highways in the state of Paraiba.
- In the second server, based on NoSQL, we left available only data about fire outbreaks.
Case Study - Experiment 2

Show all fires detected in the city of Monteiro in 2010.

1\textsuperscript{st} step: retrieves along with the SQL Map Server the geometry and the corresponding identifier of the city of Monteiro in the GML format.
2nd step: the geometry that comes from the first step is used to filter all fires (geometries in NoSQL spatial database) that are inside the area represented by that geometry.
Case Study - Experiment 2
Experiments have demonstrated that it is possible to submit queries using the same syntax for SQL and NoSQL spatial databases in a simple and transparent manner for the user’s application.
Future Work

- Conduct experiments on the performance and scalability of services;
- Extend the server to allow the use of other NoSQL databases
Many Thanks!

Questions?
Using OGC Services to Interoperate Spatial Data Stored in SQL and NoSQL Databases

Cláudio de Souza Baptista
baptista@dsc.ufcg.edu.br

Carlos Eduardo Santos Pires
cesp@dsc.ufcg.edu.br

Odilon Francisco de Lima Junior
odilonflj@gmail.com