Spatial data in NoSQL databases

Wilko Quak, Peter van Oosterom, Martijn Meijers, Irene de Vreede, Oscar Martinez Rubi
Delft University of Technology

Abstract :

In 2005 database guru Stonebraker states that the time for one size fits all databases has come and gone and that there will be a collection of independent database engines (Stonebraker, 2005). With the upcoming popularity of NoSQL databases this has become reality. Each NoSQL databases differs from the traditional (object-)relational DBMS in a different way depending on the specific problem it is designed to solve. Also spatial support differs widely between the products. In recent years we did experiments with several NoSQL databases, testing the spatial functionality and performance:

- Graph/network data: Storage of the geometry and topology of the waterways network of the Netherlands in the graph database Neo4j with spatial extension.
- Trajectory/moving objects data: Storage and analysis of the current and historic AIS shipslogs of ships in the Netherlands in MongoDB. MongoDB claims to be highly scalable, allows ad-hod querying and has support for point objects making it a suitable candidate. We also compared the MongoDB implementation with a PostgreSQL implementation.
- Point cloud data: Benchmarking the storage of the 640 billion points of the AHN2 dataset in various databases. In this research the column store MonetDB was tested against file based solutions and Oracle and PostgreSQL storage.
- Point cloud (and spatial other) data: Review of potential use of Spatial Hadoop for point cloud data as the Hadoop systems are very common nowadays. Its main features are presented, and specifically Spatial Hadoop is introduced (with similar developments in Spark).
- nD-array data: Benchmarking storage and retrieval of large hydrologic (in form multidimensional array/ 3D, 4D, 5D grids) in the SciDB database and comparing NetCDF file storage and retrieval.

Bases on the results of our various experiments and literature research we establish some conclusions on the state-of-play in the field spatial support in NoSQL databases and developments in this field.

References:

Stonebraker, M. and Cetintemel, U." One size fits all": an idea whose time has come and gone *21st International Conference on Data Engineering (ICDE'05),* **2005**, 2-1