A Demonstration of HadoopViz: An Extensible MapReduce System for Visualizing Big Spatial Data

Ahmed Eldawy  Mohamed F. Mokbel  Christopher Jonathan
Department of Computer Science and Engineering  University of Minnesota

Motivation
- Tera-bytes of satellite data
- Billions of geotagged tweets
- Peta-bytes of medical images
- Millions of space images

Challenges
- Visualization is an international language for communication
- Different visualizations for different datasets and applications
- Spatial data needs to be smoothed
- Giga-pixel images to catch all the details
- Limited resources of a single machine

Visualization Abstraction

1. smooth
Fuses nearby records together before drawing

2. create-raster
Initializes a canvas on which records are plotted

3. rasterize
Visualizes a record on a raster layer created by the create-raster function

4. merge
Merges two partial raster layers

5. write
Writes the final raster layer to the output as an image

Single Level Plot
Implemented in MapReduce using the five abstract functions

Multilevel Plot
Multilevel Image

- Millions of tiles
- Supports zoom in/out
- Giga-pixels resolution

Flat Partitioning
- Partition to tiles in levels that are multiple of k
- Each partition generates all tiles in up-to k levels
- No merging step is required

Pyramid Partitioning

Case Studies

I. Scatter plot
II. Road network
III. Frequency heat map
IV. Satellite heat map

Interface

Performance

This work is supported in part by NSF under Grants IIS-0952977 and IIS-1218168