Information Systems

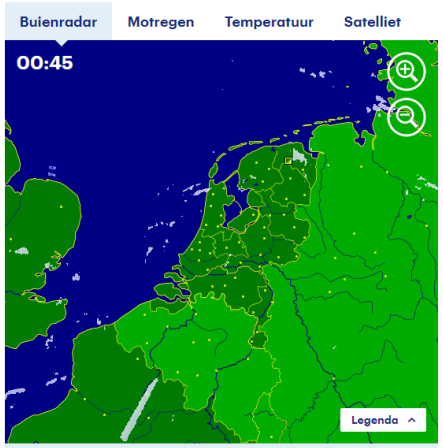
Workshop 1/9: Applied Information Systems & NoSQL

An introduction by TNO

Prerequisits

- Work in pairs
- Your own laptop with:
  - Test datasets
  - Your favorite programming language
  - One or more data stores:
    - RDBMS: postgres/mysql
    - Column store: Cassandra/InfluxDB
    - KeyValue store: Redis
    - FileSystem

Setup



- Goal: Fastest way to serve buienradar information, both:
  - Images for the user
  - Expected Rain for x/y location
- Context:
  - Frequency: New radar image each 5 minutes ~= 100.000 images per year
  - Resolution = 550x500 = 275.000 values per image
  - Image is sparsely filled, i.e. it doesn’t rain everywhere (usually).
  - Generate Random data! (or simulate a cloud moving across, etc)
- Make things testable
  - Create a test set where the value for a given x,y coordinate is checkable. For example value = width*y+ x (index) that way you will be able to check what you’re doing is correct

Tasks

- Define your own schema to get images and rain per x/y coordinate in your DB
- Write 100mln values (if possible) to your solution and calculate performance over time (what happens?)
- After different orders of magnitude (100.000, 1mln, 10mln, 100mln) of writes do a read query and calculate performance (what happens?)
- Which query (image / rainperxy) performs better (why?)
- Open question: what could you do to improve things?
- Create a table with your performance scores and share with the group
- Change schema/db/solution and experiment, try to improve your scores

Example result (minimally 2 solutions) more = bonus

Your schema/DB	100.000	1mln	10mln	100mln
Write performance	???? inserts/sec	???? inserts/sec	???? inserts/sec	???? inserts/sec
Read performance Image	???? Reads/sec	???? Reads/sec	???? Reads/sec	???? Reads/sec
Read performance xy rain	???? Reads/sec	???? Reads/sec	???? Reads/sec	???? Reads/sec

Explain/show your schema

…….

Zipfile of your code

Contribution student 1:

……

Performance deliberations: (Why)

…….

What could you do to improve things?:

……

New & Improved scores thanks to:

……

Contribution student 2:

……..