Science SQL

OGC TC Barcelona, 2015-Mar 10

Peter Baumann
Jacobs University | rasdaman GmbH
baumann@rasdaman.com

Research supported by EU EarthServer, PublicaMundi
Agile Big Data Analytics for Earth & Planetary Sciences

- 20 ... 132 TB spatio-temporal databases
  - running rasdaman

- EarthServer-2 starting May 2015

www.earthserver.eu
What Users Really Want

- "Given me all of the images in this geographic area in this this time span that are at least 80% cloud free have been radiometrically corrected and are from these satellites and then pass those images into a workflow to perform functions x,y,z"
  - Carl Reed

- “Find images taken by the SEVIRI satellite on August 25, 2007 which contain fire hotspots in areas which have been classified as forests according to CORINE Land Cover, and are located within 2km from an archaeological site in the Peloponnese.”
  - INSPIRE related

- „probability that the combined wave and swell height in area X, for a daylight time window of 12 hours, will be less than 1.5m, is 90%“
  - WMS BP draft

complexity + flexibility
...aka „Variety“
Sample Application: Database Visualization

```sql
select
  encode(
    struct {
      red: (char) s.image.b7[x0:x1,x0:x1],
      green: (char) s.image.b5[x0:x1,x0:x1],
      blue: (char) s.image.b0[x0:x1,x0:x1],
      alpha: (char) scale( d.elev, 20 )
    },
    "image/png"
  )
from SatImage as s, DEM as d
```

[JacobsU, Fraunhofer; data courtesy]
Human Brain Imaging

- **Research goal:** to understand structural-functional relations in human brain

- **Experiments capture activity patterns** (PET, fMRI)
  - Temperature, electrical, oxygen consumption, ...
  - → lots of computations → „activation maps“

- **Example:** “*a parasagittal view of all scans containing critical Hippocampus activations, TIFF-coded.*“

```sql
select tiff( ht[ $1, *:**, *:* ] )
from HeadTomograms as ht,
     Hippocampus as mask
where count_cells( ht > $2 and mask )
     / count_cells( mask )
     > $3
```

$1 =$ slicing position, $2 = $ intensity threshold value, $3 = $ confidence
Gene Expression Analysis

- Gene expression = reading out genes for reproduction
- Research goal: capture spatio-temporal expression patterns in Drosophila

http://urchin.spbcas.ru/Mooshka/ [Samsonova et al]

```
select jpeg( scale( {1c,0c,0c}*e[0,*:*,*:*] +{0c,1c,0c}*e[1,*:*,*:*] +{0c,0c,1c}*e[2,*:*,*:*], 0.2 ) )
from EmbryoImages as e
where oid(e)=193537
```
ISO 9075 Part 15: SQL/MDA
- resolved by ISO SQL WG in June 2014

n-D arrays as attributes

```
create table LandsatScenes(
    id: integer not null, acquired: date,
    scene: row( band1: integer, ..., band7: integer ) array [ 0:4999,0:4999]
)
```

declarative array operations

```
select id, encode(scene.band1-scene.band2)/(scene.nband1+scene.band2), "image/tiff"
from LandsatScenes
where acquired between "1990-06-01" and "1990-06-30" and
    avg( scene.band3-scene.band4)/(scene.band3+scene.band4)) > 0
```
OGC WCPS

- **OGC Web Coverage Processing Service (WCPS)**
  = high-level geo raster query language; adopted 2008

- WCPS 2: all grid types:

- "From MODIS scenes M1, M2, M3: difference between red & nir, as TIFF"
  - ...but only those where nir exceeds 127 somewhere

```python
for $c$ in ( M1, M2, M3 )
  where some( $c$.nir > 127 )
  return encode( $c$.red - $c$.nir, "image/tiff" )
```
**rasdaman: Agile Array Analytics**

- **“raster data manager“**: SQL + n-D arrays

```sql
select img.green[x0:x1,y0:y1] > 130
from LandsatArchive as img
where avg_cells( img.nir ) < 17
```

- Scalable parallel “tile streaming” architecture

- Implemented & in use: ESA, NASA, PML, DWD, ...
  - OGC WCS Core Reference Implementation
  - Copernicus Masters Big Data Challenge 2014 winner
Adaptive Tiling

- Sample tiling strategies [Furtado]:
  - regular
  - directional
  - area of interest

- rasdaman storage layout language

```
insert into MyCollection
values ...
tiling area of interest [0:20,0:40], [45:80,80:85]
tile size 1000000
index d_index storage array compression zlib
```
Optimization 1: Query Rewriting

- **understood:**
  - heuristic optimization
  - 150 rules in rasdaman [Ritsch 2002]

- **partially understood:**
  - cost-based optimization
Optimization 2: Just-In-Time Compilation

- **Observation:** interpreted mode slows down

- **Approach:**
  - cluster suitable operations
  - compile & dynamically bind

- **Benefit:**
  - Speed up complex, repeated operations

- **Variation:**
  - compile code for GPU (later)

[Jucovschi, Stancu-Mara 2008]

```sql
select x*x*...*x
from float_matrix as x
```

Times [ms] for $512^2 \times n$ ops
Parallel / Distributed Query Processing

\[
\text{select} \quad \max((A.nir - A.red) / (A.nir + A.red)) \\
- \max((B.nir - B.red) / (B.nir + B.red)) \\
- \max((C.nir - C.red) / (C.nir + C.red)) \\
- \max((D.nir - D.red) / (D.nir + D.red)) \\
\text{from} \; A, B, C, D
\]

1 query → 1,000+ cloud nodes
Inset: Hadoop is not the Answer to All

- **no built-in knowledge** about structured data types
  - “Since it was not originally designed to leverage the structure [...] its performance [...] is therefore suboptimal” [Daniel Abadi]
- M. Stonebraker (XLDB 2012): „will hit a scalability wall“
A Brief History of Array Databases
Conclusion

- **n-D Arrays** a major datatype, central to science, engineering, business
  - Massive spatio-temporal sensor, image, simulation, statistics data

- **Query language** = flexibility + scalability + information integration
  - 130+ TB databases, 2D...5D datacubes & timeseries
  - 1 query → 1,000+ cloud nodes

- **ISO SQL/MDA** a game changer
  - Any question, any time
  - No data/metadata divide

- **Visit us:**
  - [www.rasdaman.org](http://www.rasdaman.org)
  - [www.earthserver.eu](http://www.earthserver.eu)