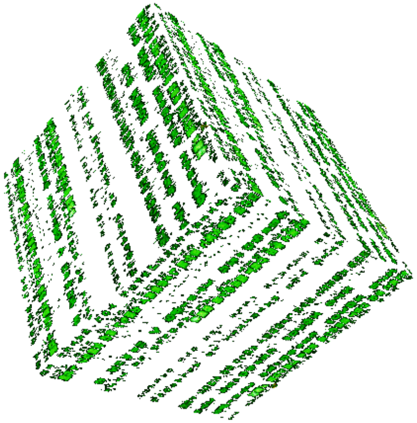
[co-funded by EU through H2020 EarthServer -2]



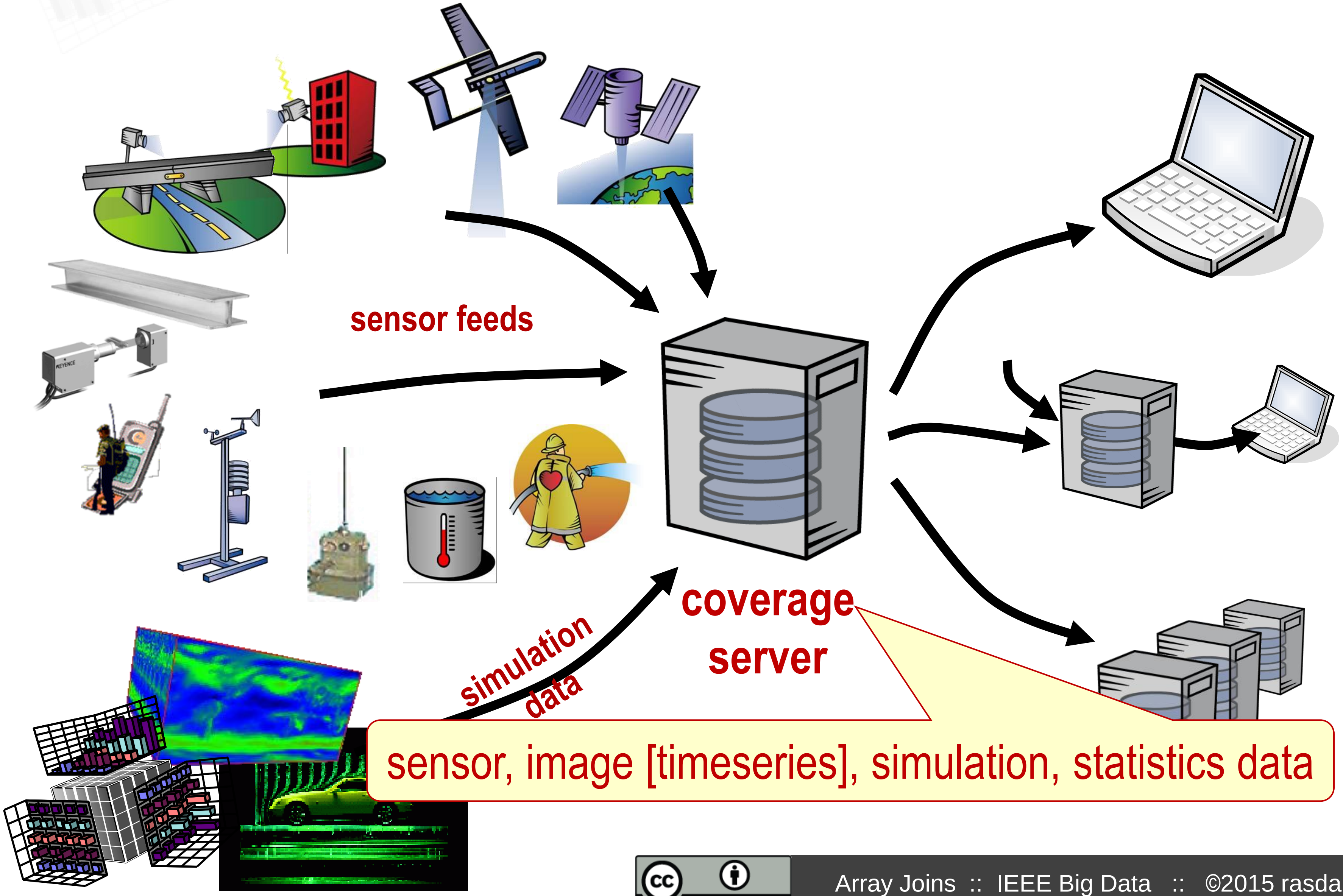
[gamingfeeds.com]

On the Efficient Evaluation of Array Joins

Big Data in the Geosciences Workshop
IEEE Big Data, Santa Clara, US, 2015-oct-29

Peter Baumann, Vlad Merticariu
Jacobs University | rasdaman GmbH
{p.baumann,v.merticariu}@jacobs-university.de

Data Homogenization With OGC Standards



sensor feeds

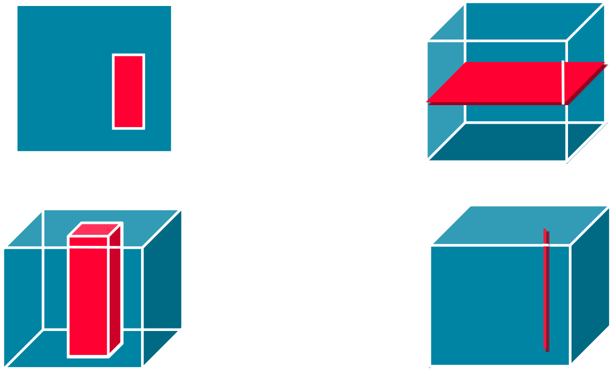
coverage server

simulation data

sensor, image [timeseries], simulation, statistics data

Web Coverage Service (WCS)

- OGC Coverages unifying regular & irregular grids, point clouds, meshes
  - OGC Coverage Implementation Schema
- WCS Core: access to spatio-temporal coverages & subsets
  - subset = trim | slice

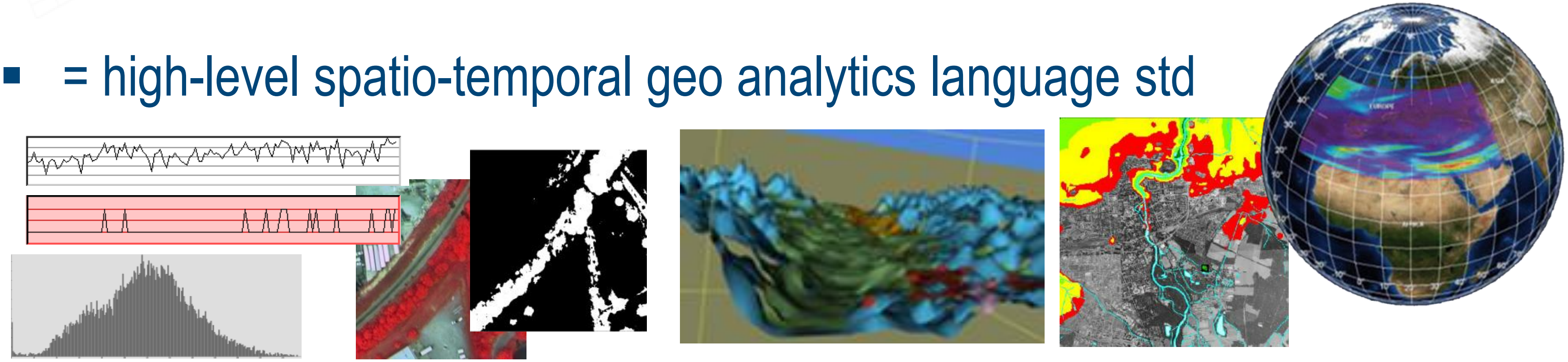


- WCS Extensions: optional functionality facets
  - Scaling, CRS transformation, …

Large, growing implementation basis: rasdaman, GDAL, QGIS, OpenLayers, OPeNDAP, MapServer, GeoServer, NASA WorldWind, EOx-Server; Pyxis, ERDAS, ArcGIS, ...

OGC Web Coverage Processing Service (WCPS)

- = high-level spatio-temporal geo analytics language std



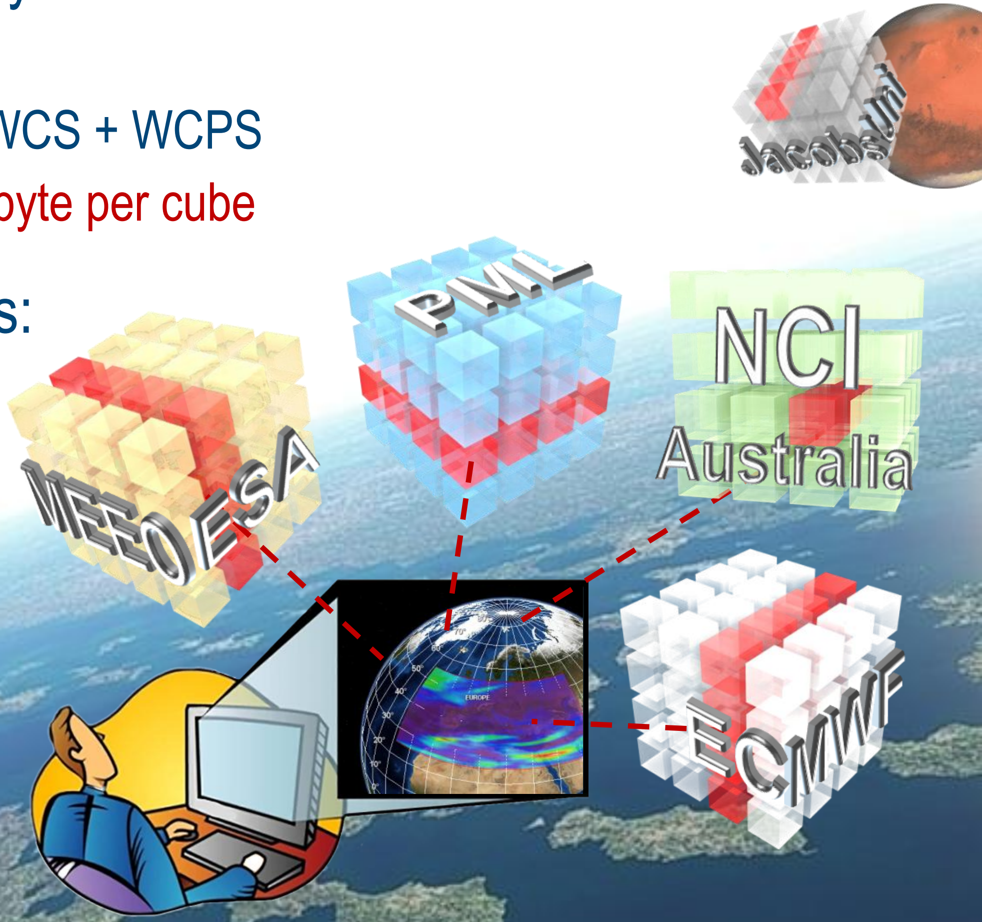
[JacobsU, FhG; NASA; data courtesy BGS, ESA]

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

```
for $c in ( M1, M2, M3 ),
    $lm in ( LandMask )
where
    some( $c.nir > 127 and $lm )
return
    encode($c.red - $c.nir, “image/tiff“ )
```

EarthServer: Datacubes At Your Fingertips

- Agile Analytics on Earth & Planetary datacubes
  - rasdaman + NASA WorldWind
  - Rigorously standards: OGC WMS + WCS + WCPS
  - 100s of TB online now, next: 1+ Petabyte per cube
- Intercontinental initiative, 3+3 years:
  EU + US + AUS

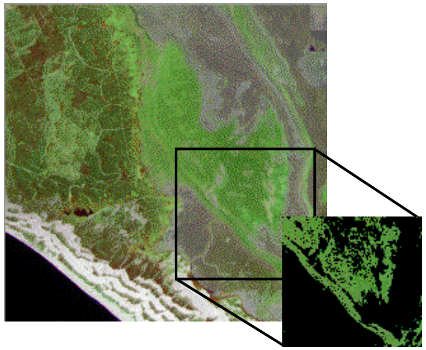


Co-funded by the European Union

www.earthserver.eu

Agile Array Analytics: rasdaman

- „raster data manager“: SQL + n-D arrays
- Scalable parallel “tile streaming” architecture
  - Joins!
- Supports R, QGIS, OpenLayers, MapServer, GDAL, EOxServer, Pyxis, ERDAS, ArcGIS, ...
- Blueprint for OGC WCPS, ISO Array SQL stds

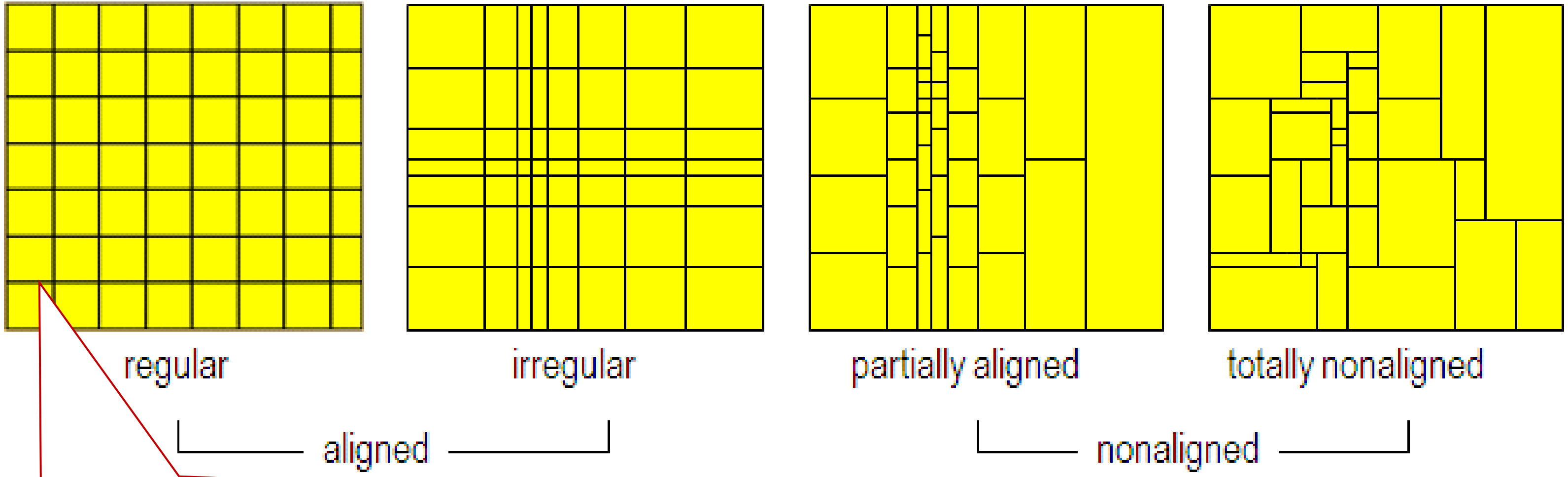


rasdaman visitors

The Earth Monitoring Competition
Copernicus masters
WINNER
T-SYSTEMS BIG DATA CHALLENGE
2014
T · ·Systems·

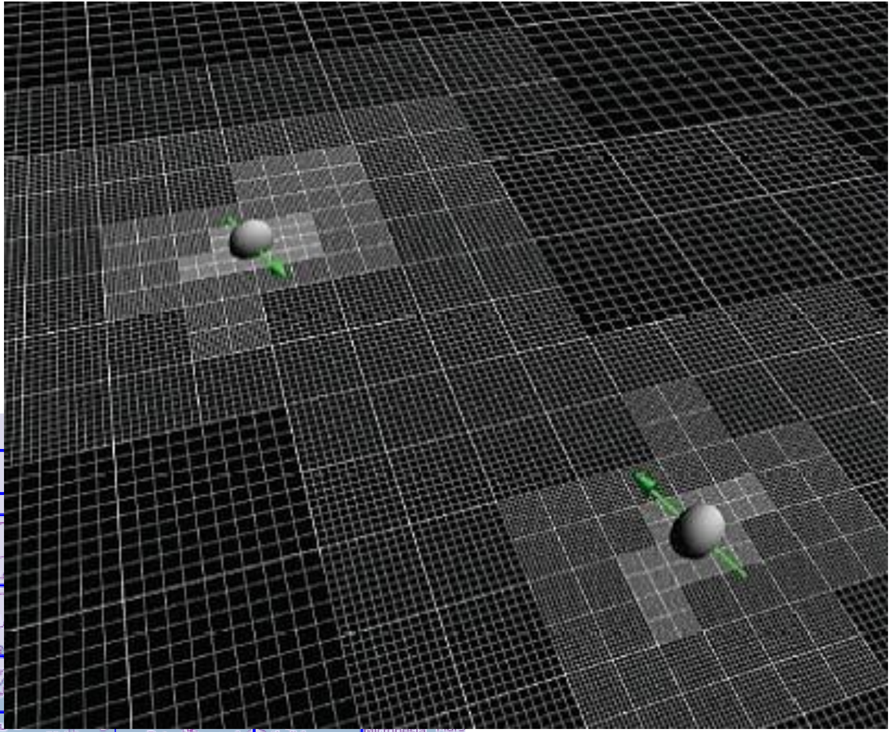
Array Partitioning

- Goal: faster loading by adapting storage units to access patterns
- Approach: split n-D array into n-D partitions („tiles“)
- Tiling classification based on degree of alignment [ICDE 1999]
    - all implemented in rasdaman

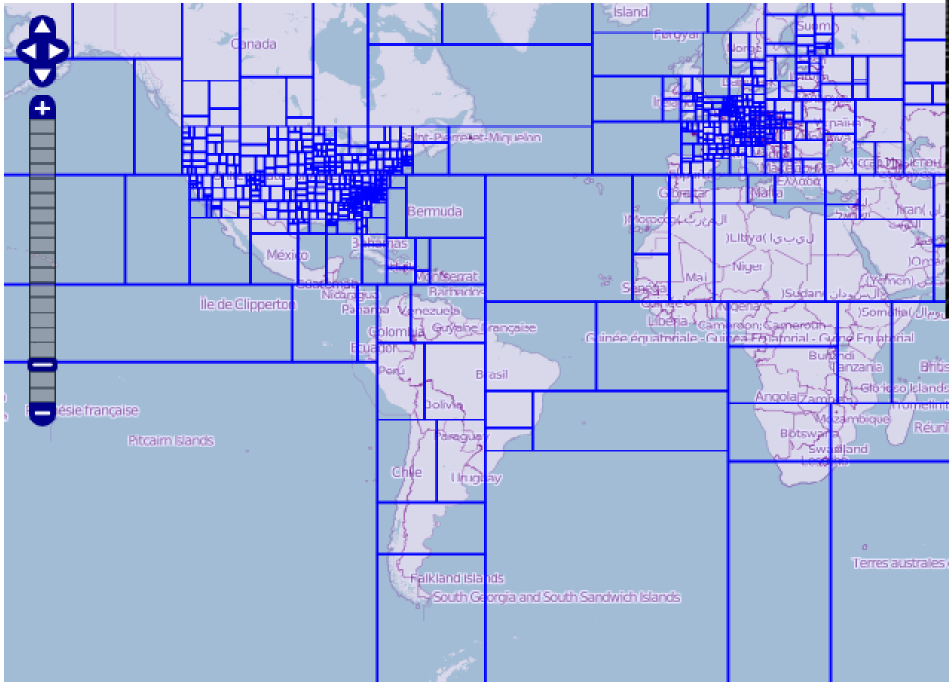


chunking [Sarawagi, Stonebraker, DeWitt, ...]

Why Irregular Partitioning?



[Centrella et al: scidacreviews.org]



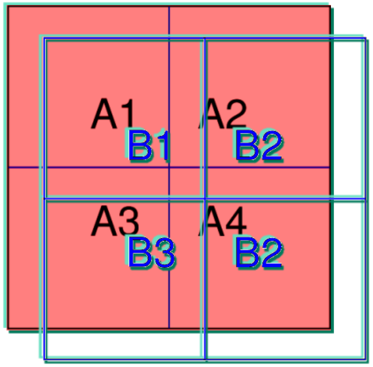
[OpenStreetMap]

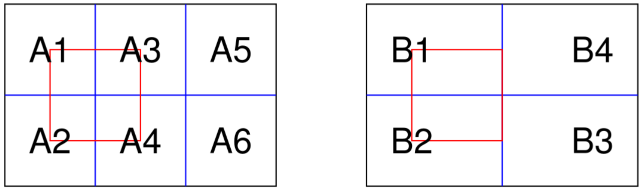
Array Join: What Happens?

- „MODIS red band with cloud mask applied“:

```
$Modis.red * $CloudMask
```

- Case 1: same tile shapes, same position
  - Easy
- Case 2: same tile shapes, different position
  - Overlaps, not so easy
- Case 3: different tile shapes
  - Worse
- Case 4: different dimensions
  - Gimme a break!

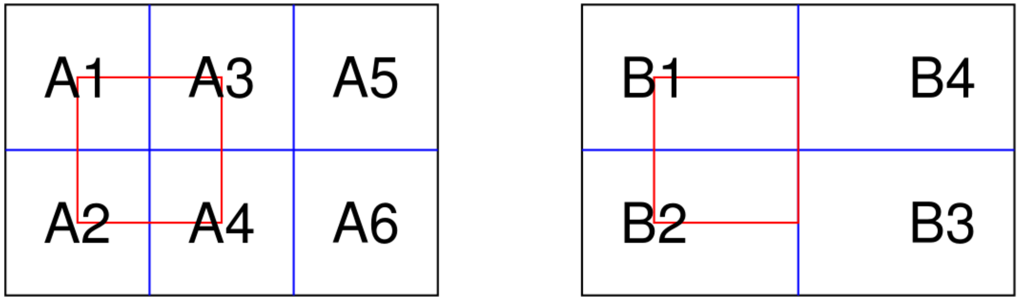




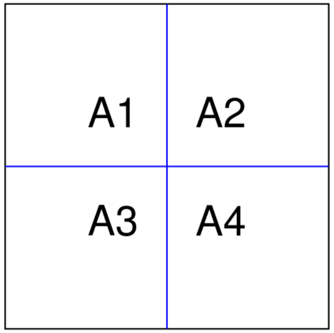
Array Join: Problem Statement

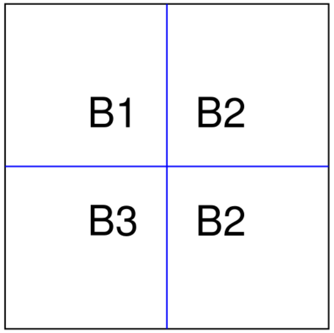
- Goal: minimize tile reads when evaluating „A op B“ in face of some arbitrary, independent partitioning of A and B

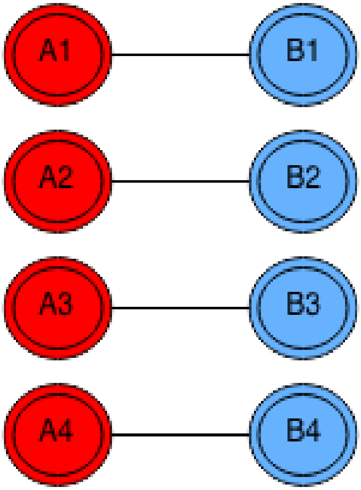
```
$Modis.red * $CloudMask
```

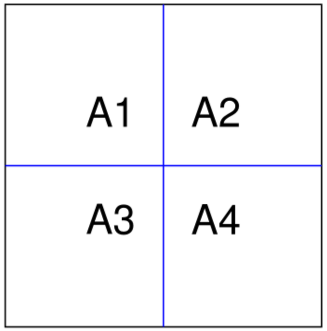


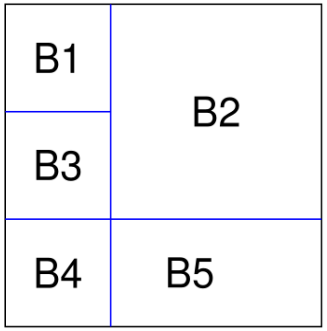
Bipartite Traversal Graphs

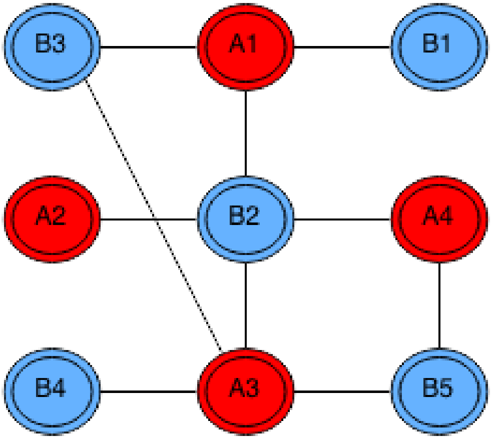






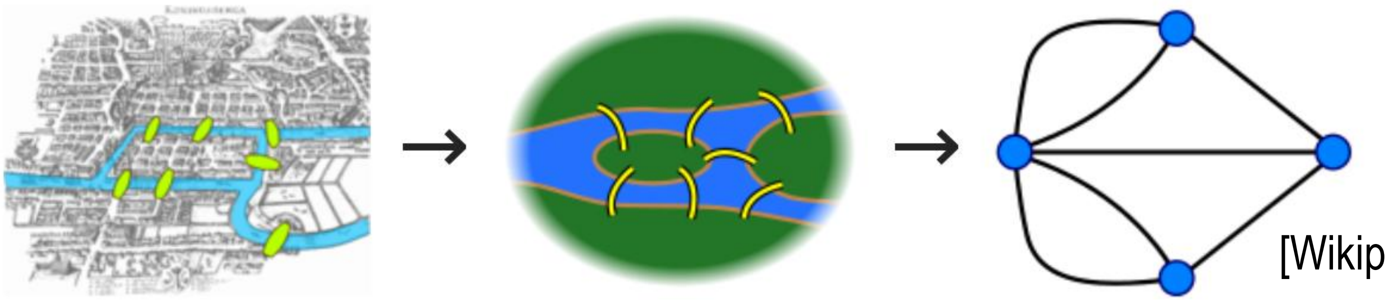






Finding Complete Paths

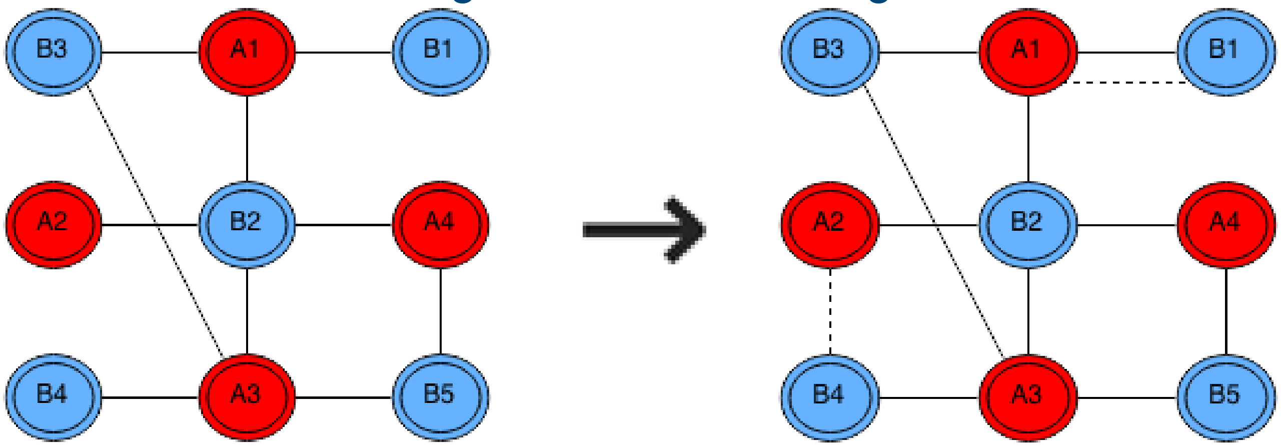
- Leonhard Euler: „complete, minimal path for even-degree nodes“



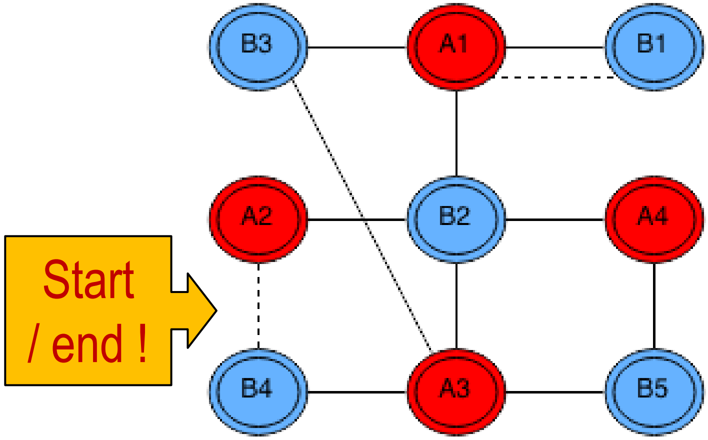
Pregel river, Königsberg

[Wikipedia]

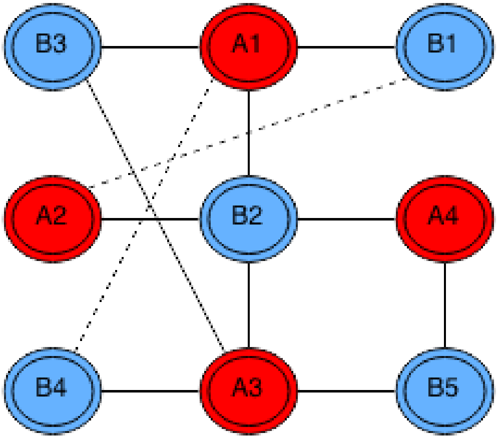
- Carl Hierholzer: „not even degree? Add aux edges!“



Finding Shortest Paths



Start / end !



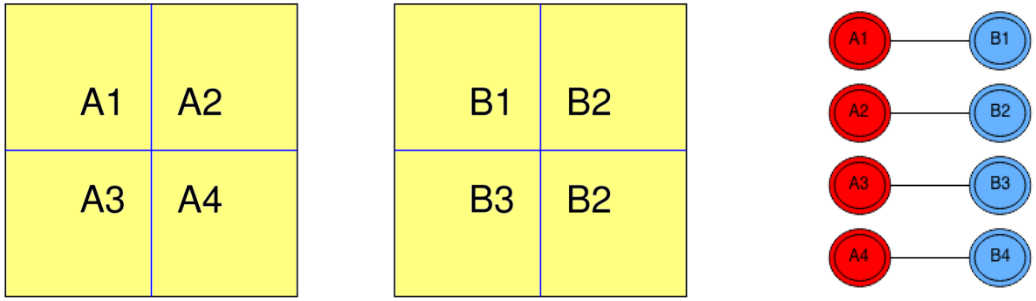
<B4,A3,B2,A1,B1,B3,
A3,B5,A4,B2,A2>

<B4,A3,B3,A1,B1,A2,
B2,A4,B5,A3,B2,A1>

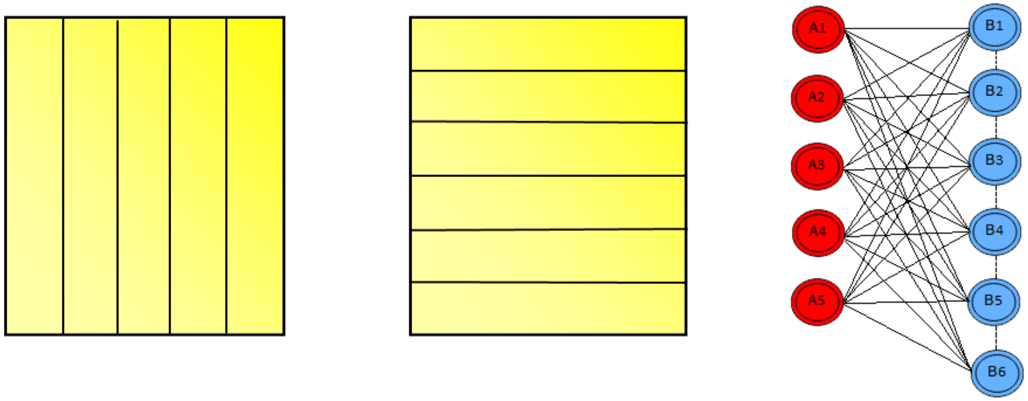
- Assumption: can hold 1 red, 1 blue tile
- How to find shortest path? See paper!

Complexity?

- Best case: full alignment → $|E_A|+|E_B|$

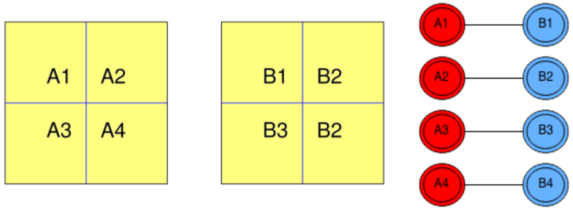


- Worst case: All tiles of A x all tiles of B → $|E_A|*|E_B|$

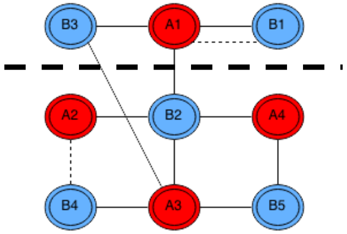


What else?

- We have*: tile traversal with minimum duplicate tile reads*
- Variation 1: *buffer size to avoid dup reads?*
  - query cost estimation
- Variation 2: *for buffer size given, how much improvement?*
- Variation 3: *parallelize disconnected subgraph*

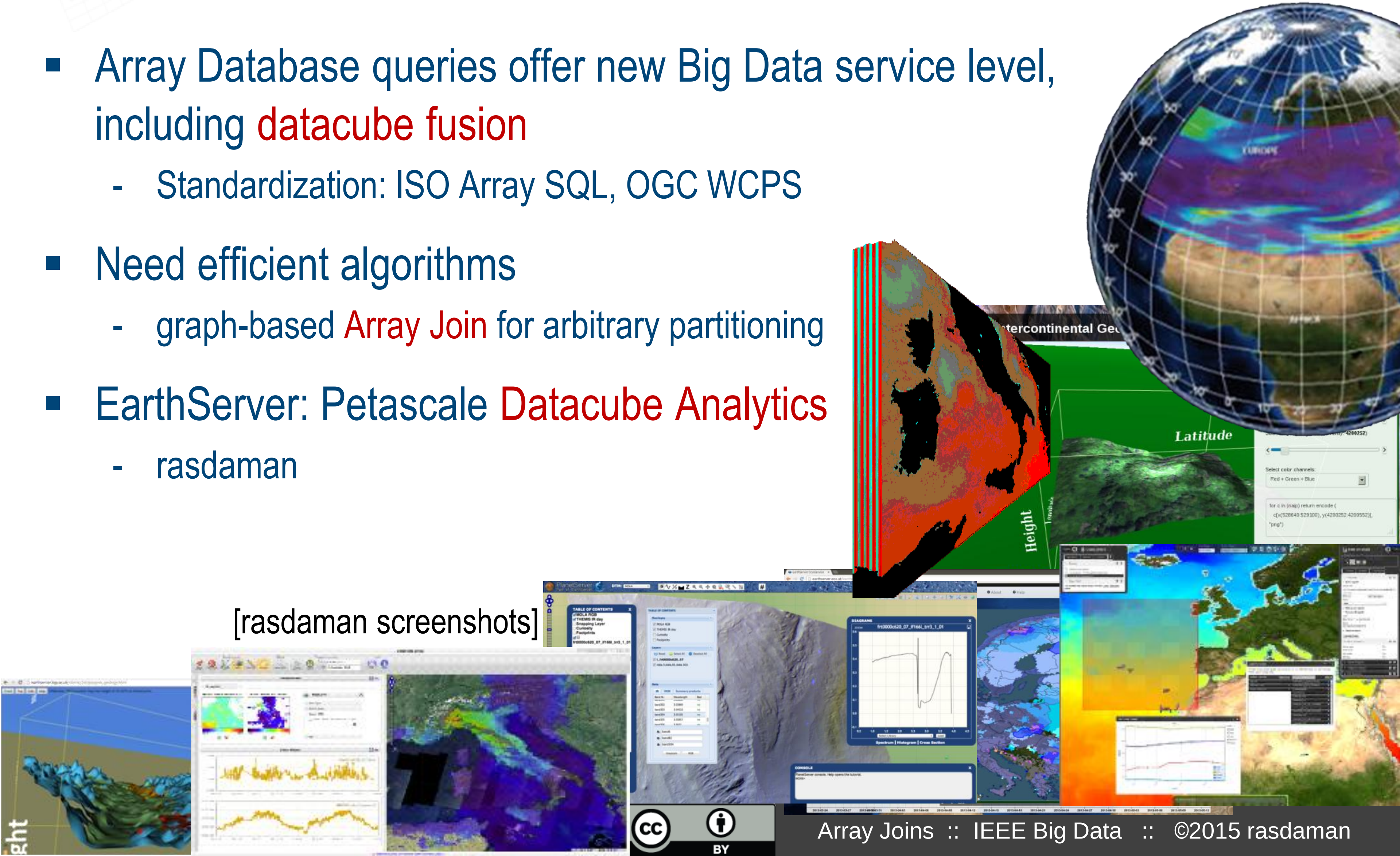


- Variation 4: *how many tiles to ship between nodes for optimal parallelization?*



Wrap-Up

- Array Database queries offer new Big Data service level, including datacube fusion
  - Standardization: ISO Array SQL, OGC WCPS
- Need efficient algorithms
  - graph-based Array Join for arbitrary partitioning
- EarthServer: Petascale Datacube Analytics
  - rasdaman



```
for c in (naip) return encode (
   c[x(528640:529100), y(4200252:4200552)],
"png")
```

[rasdaman screenshots]

Datacube Research @ Jacobs U

- Large-Scale Scientific Information Systems research group
  - Flexible, scalable n-D array services
  - www.jacobs-university.de/lsis
- Main results:
  - pioneer Array DBMS, rasdaman
  - standardization:
    - OGC Big Geo Data (also ISO, INSPIRE, W3C)
    - ISO Array SQL

Hiring PhD students, PostDocs