

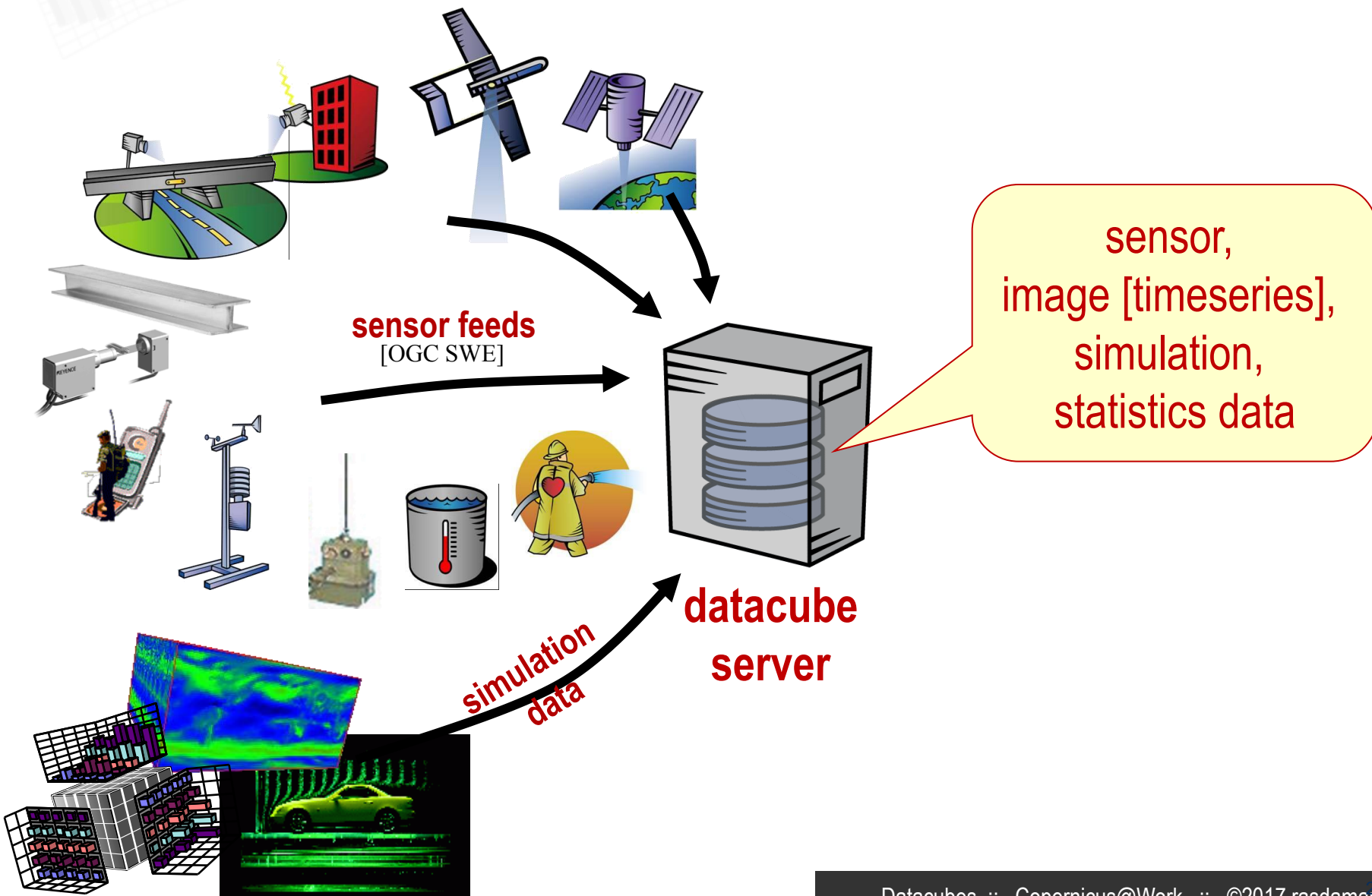
Agile Analytics & Fusion with Spatio-Temporal Datacubes

Copernicus@Work, Berlin, 2017-mar-14

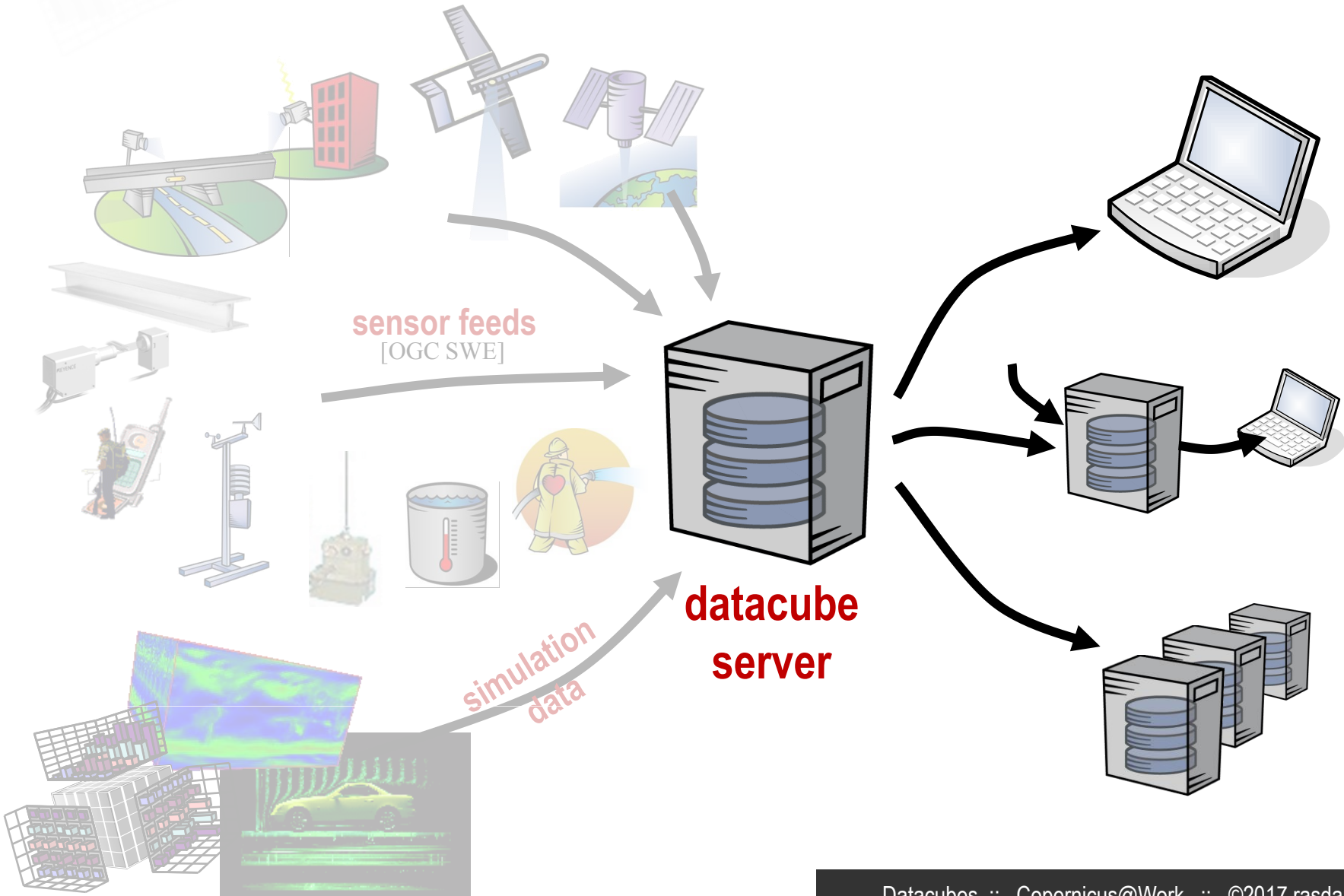
Peter Baumann

Jacobs University | rasdaman GmbH

Homogenizing Data Services



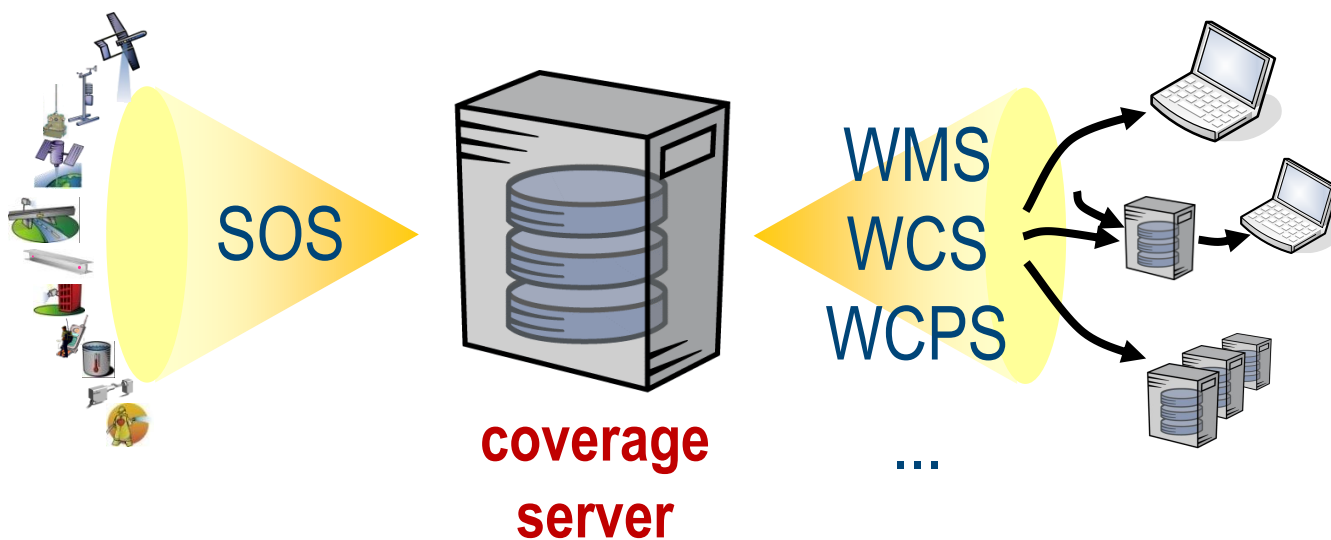
Homogenizing Data Services



In OGC/ISO/INSPIRE: Coverages

SOS: upstream
sensor data capturing

W*S: downstream
download, processing, visualization



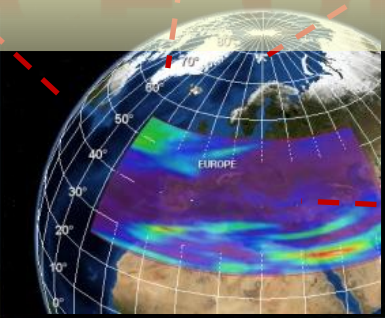
“coverages naturally are served through WCS”
- G. Escriu, INSPIRE facilitator, 2017

SERVICE QUALITY

SERVICE QUALITY

```
rwX--x--- 1 rasdata users 1485 Oct 13 2004 4251NW.ASC
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251NWGR.tfw
rwX--x--- 1 rasdata users 640432 Oct 13 2004 4251NWGR.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251NWGW.tfw
rwX--x--- 1 rasdata users 779368 Oct 13 2004 4251NWGW.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251NWRL.tfw
rwX--x--- 1 rasdata users 712492 Oct 13 2004 4251NWRL.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251NWWL.tfw
rwX--x--- 1 rasdata users 62830 Oct 13 2004 4251NWWL.tif
rwX--x--- 1 rasdata users 1498 Oct 13 2004 4251S0.ASC
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251S0GR.tfw
rwX--x--- 1 rasdata users 1076750 Oct 13 2004 4251S0GR.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251S0GW.tfw
rwX--x--- 1 rasdata users 197142 Oct 13 2004 4251S0GW.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251S0RL.tfw
rwX--x--- 1 rasdata users 936348 Oct 13 2004 4251S0RL.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251S0WL.tfw
rwX--x--- 1 rasdata users 119990 Oct 13 2004 4251S0WL.tif
rwX--x--- 1 rasdata users 1485 Oct 13 2004 4251SW.ASC
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251SWGR.tfw
rwX--x--- 1 rasdata users 577868 Oct 13 2004 4251SWGR.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251SWGW.tfw
rwX--x--- 1 rasdata users 352188 Oct 13 2004 4251SWGW.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251SWR.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251SWR.tif
rwX--x--- 1 rasdata users 1485 Oct 13 2004 4252N0.ASC
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252N0GR.tfw
rwX--x--- 1 rasdata users 355774 Oct 13 2004 4252N0GR.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252N0GW.tfw
rwX--x--- 1 rasdata users 49046 Oct 13 2004 4252N0GW.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252N0RL.tfw
rwX--x--- 1 rasdata users 600964 Oct 13 2004 4252N0RL.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252N0WL.tfw
rwX--x--- 1 rasdata users 46714 Oct 13 2004 4252N0WL.tif
rwX--x--- 1 rasdata users 1485 Oct 13 2004 4252NW.ASC
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252NWGR.tfw
rwX--x--- 1 rasdata users 1445064 Oct 13 2004 4252NWGR.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252NWGW.tfw
rwX--x--- 1 rasdata users 410426 Oct 13 2004 4252NWGW.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252NWRL.tfw
rwX--x--- 1 rasdata users 655374 Oct 13 2004 4252NWRL.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252NWWL.tfw
rwX--x--- 1 rasdata users 108612 Oct 13 2004 4252NWWL.tif
rwX--x--- 1 rasdata users 1485 Oct 13 2004 4252S0.ASC
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252S0GR.tfw
rwX--x--- 1 rasdata users 607646 Oct 13 2004 4252S0GR.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252S0GW.tfw
rwX--x--- 1 rasdata users 685092 Oct 13 2004 4252S0GW.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252S0RL.tfw
rwX--x--- 1 rasdata users 632172 Oct 13 2004 4252S0RL.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252S0WL.tfw
```

SERVICE QUALITY

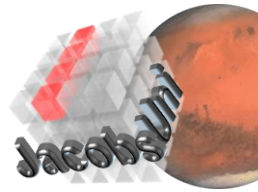


```
rwX--x--- 1 rasdata users 1485 Oct 13 2004 4251NWGR.ASC
rwX--x--- 1 rasdata users 640432 Oct 13 2004 4251NWGR.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251NWGW.tfw
rwX--x--- 1 rasdata users 779368 Oct 13 2004 4251NWGW.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251NWRL.tfw
rwX--x--- 1 rasdata users 712492 Oct 13 2004 4251NWRL.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251NWWL.tfw
rwX--x--- 1 rasdata users 62830 Oct 13 2004 4251NWWL.tif
rwX--x--- 1 rasdata users 1498 Oct 13 2004 4251S0.ASC
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251S0GR.tfw
rwX--x--- 1 rasdata users 1076750 Oct 13 2004 4251S0GR.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4251S0GW.tfw
rwX--x--- 1 rasdata users 197700 Oct 13 2004 4251S0GW.tif
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251S0RL.tfw
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251S0RL.tif
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251S0WL.tfw
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251S0WL.tif
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251SW.ASC
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251SWGR.tfw
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251SWGR.tif
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251SWGW.tfw
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251SWGW.tif
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251SWWL.tfw
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251SWWL.tif
rwX--x--- 1 rasdata users 95200 Oct 13 2004 4251S0.ASC
rwX--x--- 1 rasdata users 355774 Oct 13 2004 4252S0.ASC
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252S0GR.tfw
rwX--x--- 1 rasdata users 49046 Oct 13 2004 4252S0GR.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252S0GW.tfw
rwX--x--- 1 rasdata users 600964 Oct 13 2004 4252S0GW.tif
rwX--x--- 1 rasdata users 600964 Oct 13 2004 4252S0RL.tfw
rwX--x--- 1 rasdata users 600964 Oct 13 2004 4252S0RL.tif
rwX--x--- 1 rasdata users 600964 Oct 13 2004 4252S0WL.tfw
rwX--x--- 1 rasdata users 600964 Oct 13 2004 4252S0WL.tif
rwX--x--- 1 rasdata users 600964 Oct 13 2004 4252S0.ASC
rwX--x--- 1 rasdata users 607646 Oct 13 2004 4252S0GR.tfw
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252S0GR.tif
rwX--x--- 1 rasdata users 685092 Oct 13 2004 4252S0GW.tfw
rwX--x--- 1 rasdata users 685092 Oct 13 2004 4252S0GW.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252S0RL.tfw
rwX--x--- 1 rasdata users 632172 Oct 13 2004 4252S0RL.tif
rwX--x--- 1 rasdata users 216 Oct 13 2004 4252S0WL.tfw
rwX--x--- 1 rasdata users 632172 Oct 13 2004 4252S0WL.tif
```



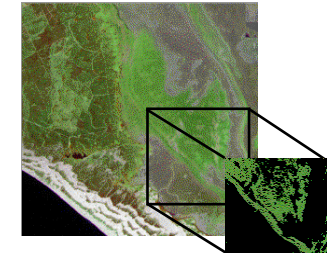
EarthServer: Datacubes At Your Fingertips

- **Agile Analytics** on x/y/t + x/y/z/t Earth & Planetary **datacubes**
 - EU rasdaman + NASA WorldWind
 - 250+ TB now, next: **1+ PB cube**
- Intercontinental initiative, started 2011:
EU + US + AUS
- Towards global
datacube federation



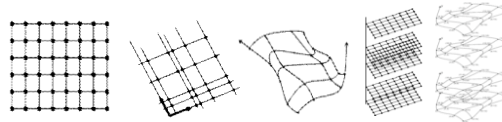
rasdaman: Agile Array Analytics

= „raster data manager“: SQL + n-D arrays



- Scalable parallel “tile streaming” architecture
[VLDB 1994, VLDB 1997, SIGMOD 1998, VLDB 2003, ..., VLDB 2016]

- Regular & irregular grids



■ Mature, in operational use, on OSGeo Live

- blueprint for Big Datacube standards: ISO, OGC, INSPIRE

“world leading environment”, “standard working horse for OGC standardisation on these innovative data access interfaces” - G. Landgraf, ESA, 2017

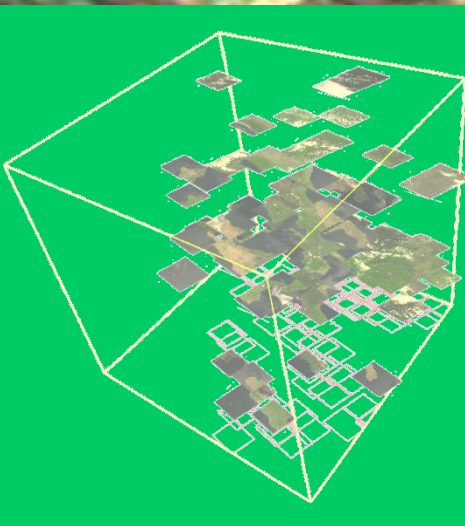


Direct Data Visualization

```

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

```



Linear Algebra Ops

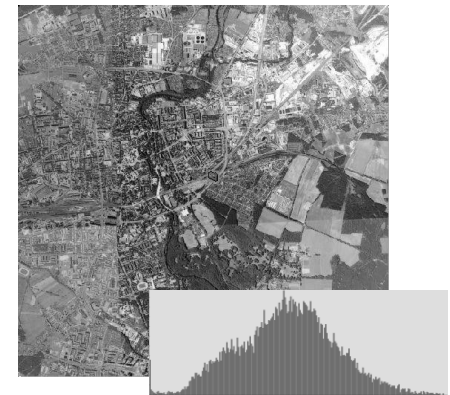
- Matrix multiplication

$$(\mathbf{AB})_{ij} = \sum_{k=1}^m A_{ik} B_{kj}$$

```
select marray i in [0:m], j in [0:p]
       values condense +
           over      k in [0:n]
           using     a [ i, k ] * b [ k, j ]
from   matrix as a, matrix as b
```

- Histogram

```
select marray bucket in [0:255]
       values count_cells( img = bucket )
from   img
```

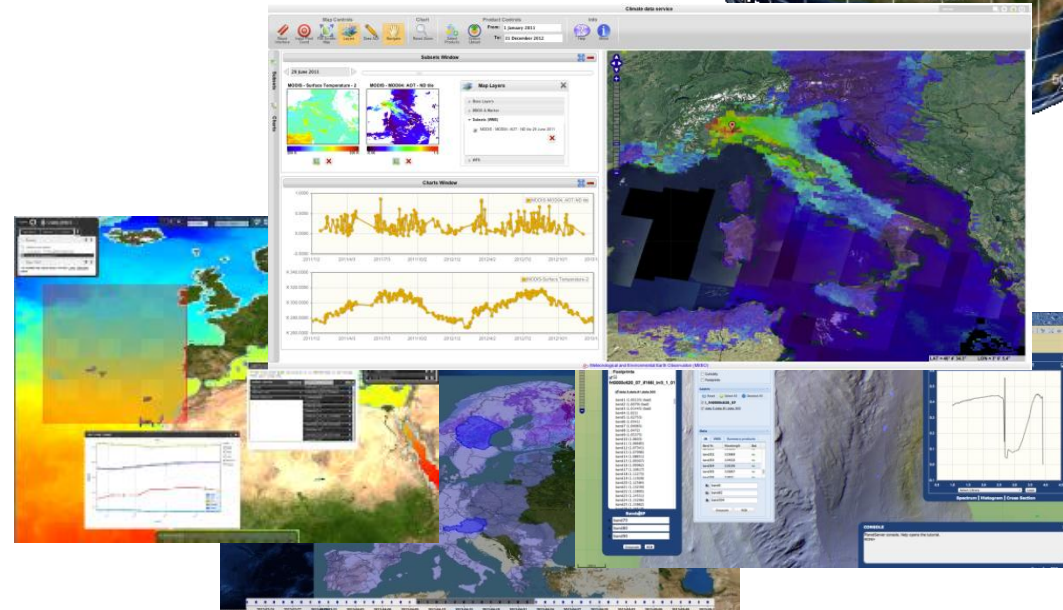
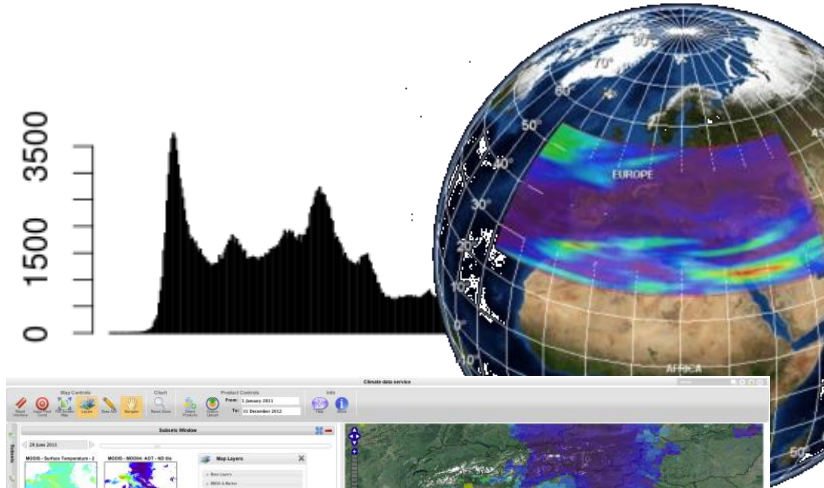


Science & GIS Tool Interfacing

- Geo tools
 - MapServer, GDAL, QGIS, OpenLayers, Leaflet, NASA WorldWind, Cesium, ...

- General-purpose scientist tools
 - Java, C++
 - python
 - R (under work)

- ...any OGC/INSPIRE compliant client
 - rasdaman
Reference Implementation



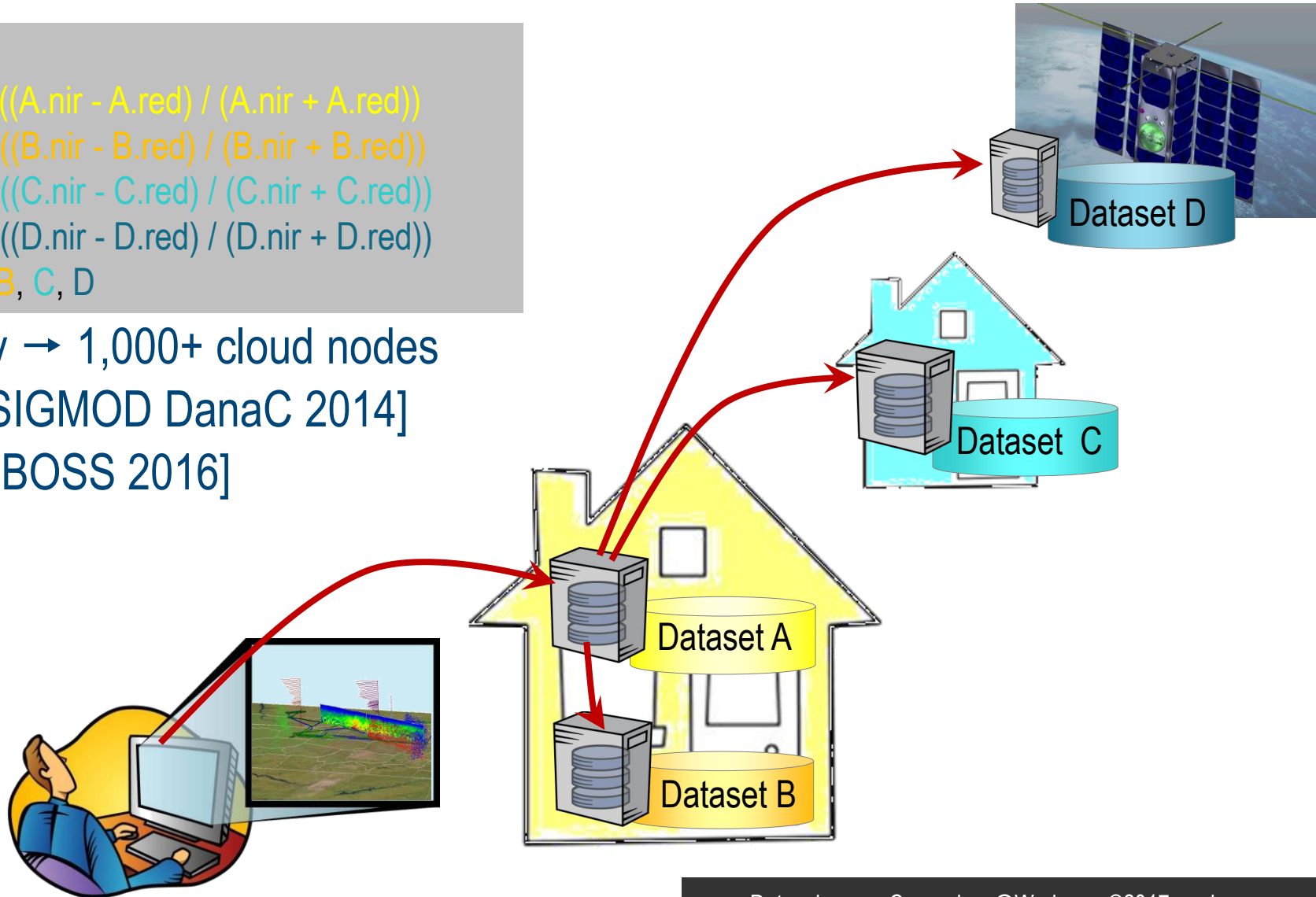
Parallel, Distributed Processing

select

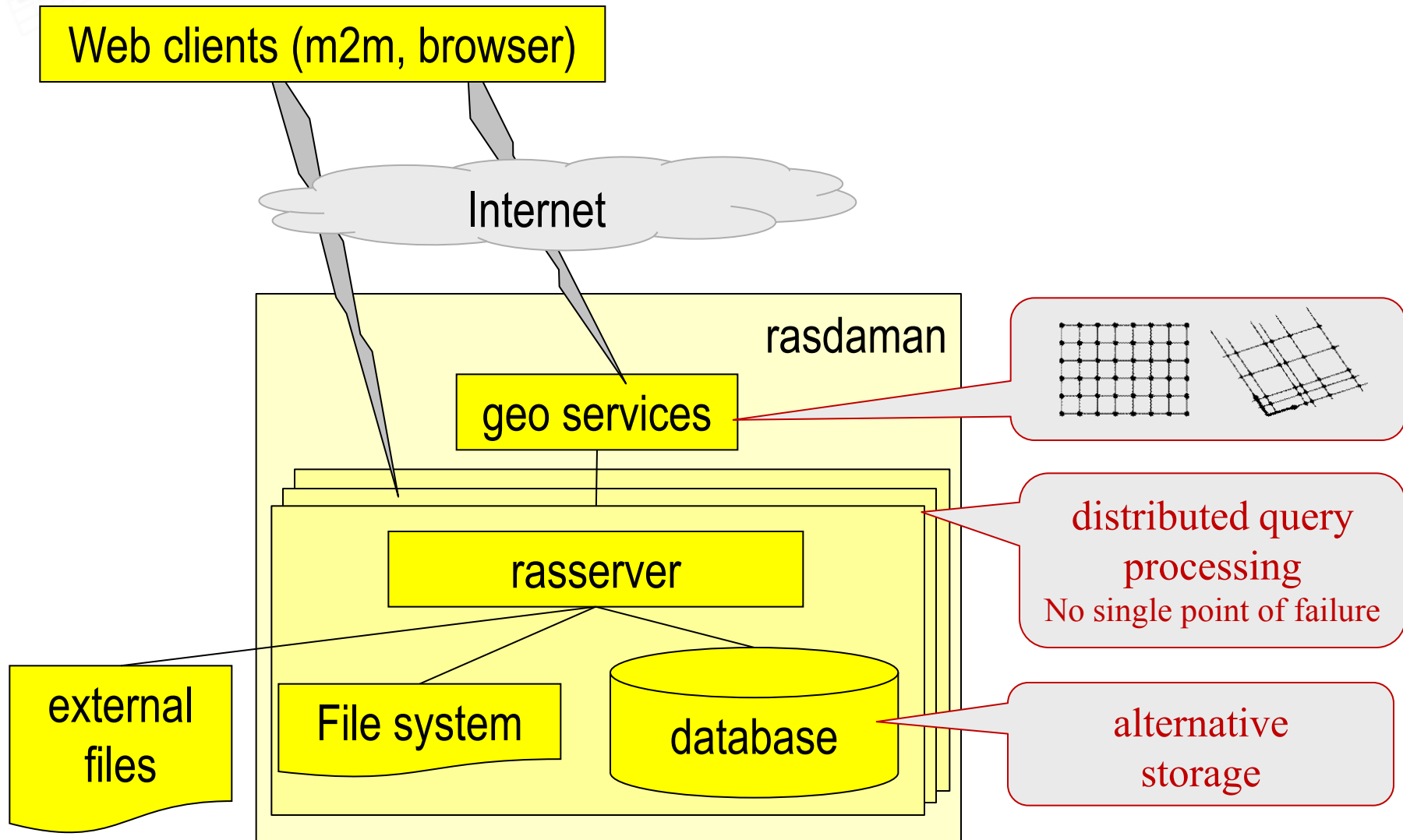
$\max((A.nir - A.red) / (A.nir + A.red))$
 - $\max((B.nir - B.red) / (B.nir + B.red))$
 - $\max((C.nir - C.red) / (C.nir + C.red))$
 - $\max((D.nir - D.red) / (D.nir + D.red))$

from A, B, C, D

1 query → 1,000+ cloud nodes
 [ACM SIGMOD DanaC 2014]
 [VLDB BOSS 2016]



rasdaman Architecture



ESA Tech Harm Strat Meetg, 2017-feb-21

German recommendations for Roadmap

EO Big Data

- Establish „European Datacube Federation“ as a strategic goal
 - Advancing user service quality, specifically: timeseries analytics
 - Enable flexible framework for value-adding services, through open standards
 - Maintain European technology lead
 - EarthServer as reference example to be enlarged
- Foster and Organize joint activities with platform developers and operators in member states
 - Analyse existing EO Exploitation Platforms with respect to common functions and interoperable interfaces
 - Define federation requirements including IT security aspects on appropriate levels
 - Support harmonization and standardization on interfaces of federated platforms
 - Develop components and workflows supporting harmonized scenarios and interfaces

Wrap-Up

*„A cube says more
than a million images“*

- Arrays, aka spatio-temporal **Datacubes**, major „Big Data“ category
 - = *sensor, image (timeseries), simulation, statistics datacubes*
 - Wide **consensus** by bodies (OGC, ISO, INSPIRE) + implementers
 - *SQL/MDA game changer for science, engineering, and beyond*
- **EarthServer**: towards planetary-scale datacube federation
- **rasdaman**: flexible, scalable array engine
 - 250+ TB; 1,000x query parallelization
 - blueprint for „Big Data“ standards, often reference implementation
 - See us: www.earthserver.eu, www.rasdaman.org, www.jacobs-university.de/isis