Geophysical Research Abstracts
Vol. 18, EGU2016-16753-1, 2016
EGU General Assembly 2016


# Datacubes as a Service Paradigm

Angelo Pio Rossi and Peter Baumann
Jacobs University, Bremen, Germany

Spatio-temporal data sets often can be represented conveniently through datacubes as a common unifying paradigm. Flexible, scalable services can be offered based on the concept of a datacube query language while hiding the technicalities, thereby allowing user-friendly visual data interaction.

One of today's most influential initiatives in Big Geo Data is EarthServer which is paving the way for flexible, scalable datacube services based on innovative NewSQL technology (Fig. 1). Researchers from Europe, the US and recently Australia have teamed up to rigorously materialize the datacube paradigm for Earth Observation, ocean, meteorological, and planetary science.

EarthServer has established client and server technology for such spatio-temporal datacubes strictly based on the open datacube standards, OGC WCS and WPCS. The underlying scalable array engine, rasdaman, enables direct interaction, including 3-D visualization, what-if scenarios, common EO data processing, and general analytics on regular and irregular grids. Integration of datacube and metadata retrieval, together with advanced visualization based on NASA WorldWind, are geared towards an effective, user-friendly access and analysis. Conversely, EarthServer is significantly shaping the ISO, OGC, and INSPIRE Big Data standards landscape by being specification editor.

Phase 1 of EarthServer has advanced scalable array data¬base technology into 100+ TB services; in phase 2, a federation of Petabyte datacubes is being built in Europe and Australia to perform ad-hoc querying and merging. Phase 1 reviewers have attested rasdaman to "significantly transform the way that scientists in different areas of Earth Science will be able to access and use data in a way that hitherto was not possible". Altogether, these large-scale deployments prove that datacubes are a convenient model for presenting users with a simple, consolidated view on the massive amount of data files gathered – "a cube tells more than a million images".