Geophysical Research Abstracts
Vol. 19, EGU2017-17574, 2017
EGU General Assembly 2017

Agile Datacube Analytics (not just) for the Earth Sciences

Dimitar Misev, Vlad Merticariu, and Peter Baumann
Jacobs University Bremen, Bremen, Germany (d.misev@jacobs-university.de)

Metadata are considered small, smart, and queryable; data, on the other hand, are known as big, clumsy, hard to analyze. Consequently, gridded data - such as images, image timeseries, and climate datacubes - are managed separately from the metadata, and with different, restricted retrieval capabilities. One reason for this silo approach is that databases, while good at tables, XML hierarchies, RDF graphs, etc., traditionally do not support multi-dimensional arrays well. This gap is being closed by Array Databases which extend the SQL paradigm of "any query, anytime" to NoSQL arrays. They introduce semantically rich modelling combined with declarative, high-level query languages on n-D arrays. On Server side, such queries can be optimized, parallelized, and distributed based on partitioned array storage. This way, they offer new vistas in flexibility, scalability, performance, and data integration. In this respect, the forthcoming ISO SQL extension MDA ("Multi-dimensional Arrays") will be a game changer in Big Data Analytics.

We introduce concepts and opportunities through the example of rasdaman ("raster data manager") which in fact has pioneered the field of Array Databases and forms the blueprint for ISO SQL/MDA and further Big Data standards, such as OGC WCPS for querying spatio-temporal Earth datacubes. With operational installations exceeding 140 TB queries have been split across more than one thousand cloud nodes, using CPUs as well as GPUs. Installations can easily be mashed up securely, enabling large-scale location-transparent query processing in federations. Federation queries have been demonstrated live at EGU 2016 spanning Europe and Australia in the context of the intercontinental EarthServer initiative, visualized through NASA WorldWind.