



Bringing Ad-Hoc Analytics to Big Earth Data: the EarthServer Experience

Peter Baumann

Jacobs University, CASE (Center for Advanced Systems Engineering), Bremen, Germany (p.baumann@jacobs-university.de)

From the commonly accepted Vs defining the Big Data challenge - volume, velocity, variety - we more and more learn that the sheer volume is not the only, and often not even the decisive factor inhibiting access and analytics. In particular variety of data is a frequent core issue, posing manifold issues. Based on this observation we claim that a key aspect to analytics is the freedom to ask any questions, simple or complex, anytime and combining any choice of data structures, whatever diverging they may be.

Actually, techniques for such "ad-hoc queries" we can learn from classical databases. Their concept of high-level query languages brings along several benefits: a uniform semantic, allowing machine-to-machine communication, including automatic generation of queries; massive server-side optimization and parallelization; and building attractive client interfaces hiding the query syntax from casual users while allowing power users to utilize it.

However, these benefits used to be available only on tabular and set oriented data, text, and - more recently - graph data. With the advent of Array Databases, they become available on large multidimensional raster data assets as well, getting one step closer to the Holy Grail of integrated, uniform retrieval for users.

EarthServer is a transatlantic initiative setting up operational infrastructures based on this paradigm. In our talk, we present core EarthServer technology concepts as well as a spectrum of Earth Science applications utilizing the EarthServer platform for versatile, visualisation supported analytics services. Further, we discuss the substantial impact EarthServer is having on Big Geo Data standardization in OGC and ISO. Time and Internet connection permitting a live demo can be presented.