rasdaman:
Array Databases Boost Spatio-Temporal Analytics

Peter Baumann
School of Engineering
Jacobs University
Bremen, Germany
p.baumann@jacobs-university.de

Rasdaman ("raster data manager") is the pioneer in Array Database Systems, the next generation in scalable scientific data services: it provides agile analytics on massive multidimensional raster data ("arrays"), such as regular and irregular spatio-temporal grids. An SQL-style query language allows users to flexibly build their own product in a "mix and match" style. The underlying engine boosts performance through strong optimizations, large-scale parallelization, and use of new hardware. Key distinguishing features of rasdaman are:

- **Flexibility** – any query, any time, from 1D through 4D spatio-temporal data and beyond
- **Scalability** - individual dynamic optimization and parallelization for each incoming query, tested by distributing single queries to 1000+ cloud nodes and hundreds of Terabytes
- **Performance** – real-time access, processing, mixing, and filtering of any-size spatio-temporal data

On such sensor, image, simulation, and statistics data appearing, e.g., in earth, space, and life science applications rasdaman allows to quickly set up array-intensive services which are distinguished by their flexibility, speed, and scalability. Rasdaman embeds itself smoothly into PostgreSQL or other standard databases, or simply uses any file system.

The transatlantic EarthServer initiative utilizes rasdaman as its enabling platform technology for hundreds of Terabytes of satellite, atmosphere, ocean, and geology data.

The rasdaman concepts are heavily impacting international standardization, having shaped the OGC raster query language, Web Coverage Processing Service (WCPS) and currently leading the ISO Array SQL initiative.