



Integrated Array/Metadata Analytics

Dimitar Misev and Peter Baumann

Jacobs University Bremen, Germany (d.misev@jacobs-university.de)

Data comes in various forms and types, and integration usually presents a problem that is often simply ignored and solved with ad-hoc solutions. Multidimensional arrays are an ubiquitous data type, that we find at the core of virtually all science and engineering domains, as sensor, model, image, statistics data. Naturally, arrays are richly described by and intertwined with additional metadata (alphanumeric relational data, XML, JSON, etc). Database systems, however, a fundamental building block of what we call "Big Data", lack adequate support for modelling and expressing these array data/metadata relationships. Array analytics is hence quite primitive or non-existent at all in modern relational DBMS.

Recognizing this, we extended SQL with a new SQL/MDA part seamlessly integrating multidimensional array analytics into the standard database query language. We demonstrate the benefits of SQL/MDA with real-world examples executed in ASQLDB, an open-source mediator system based on HSQLDB and rasdaman, that already implements SQL/MDA.