



Dealing with Big (Geo) Data

Programme

Agile Array Analytics: why we need array databases

Peter Baumann (Jacobs University, Bremen – Germany)

Multidimensional arrays represent an information category which substantially contributes to today's "Big Data" challenge. 1D sensor timeseries, 2D satellite imagery, 3D x/y/t image timeseries and x/y/z exploration data, 4D x/y/z/t climate and ocean data constitute but some examples of array data in the geo sciences; further application domains include life, space, and social sciences as well as business and engineering. As databases do not support large, multidimensional arrays, traditionally file-based ad-hoc implementations offering limited functionality prevail in data archives and Web services holding such data. Array databases have set out to close this gap by extending standard databases with n-D array modeling and query support. A particular contribution such Array DBMSs can make for both scientific and industrial applications is their query flexibility, scalability, and information integration. In our talk we present the rasdaman ("raster data manager") Array DBMS we have developed since several years. It is fully implemented and has been evaluated in a variety of relevant application fields. According to independent experts in the field, rasdaman is the most advanced Array DBMS today. We outline conceptual model and architecture of rasdaman, highlight the query optimization potential, discuss application domains investigated, and introduce a proposal for extending the ISO SQL standard with array capabilities in a natural way.

NetCDF and NASA WorldWind viewer

Giorgio Zamboni (Politecnico di Milano)

EST-WA (Environment Space and Time Web Analyzer), a tool for managing environmental variables in NetCDF (network Common Data Form) format, is presented. The system is composed by two tools: EST-WA2D, which is the graphical 2D interface for viewing and filtering the variable; EST-WA3D, which is the graphical 4D interface implemented using the SDK Java WorldWind. The latter topic will be discussed in detail, showing the data model and the implemented functionalities.

Compression of big data with multiresolution splines

Maria Antonia Brovelli (Politecnico di Milano)

Multi-resolution bilinear spline functions to interpolate observations and to model fields are presented. The approach is discussed and compared with the traditional techniques to interpolate, store and transmit fields (raster or matrix data), considering accuracy and storage requirements.

27th September 2013

hour: 10.00-13.00

Politecnico di Milano - Campus Como

Palazzo Natta - Summer School Room

Via Natta, 12/14 (Como)

