

Diss. ETH Nr. 15021

A Scalable Architecture for Large Scientific Databases

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH

for the degree of
Doctor of Technical Sciences

presented by
Etzard Stolte
Diplom-Informatiker Universität Bielefeld
born September 30, 1963
citizen of Germany

accepted on the recommendation of
Prof. Dr. G. Alonso, examiner
Prof. Dr. G. Weikum, co-examiner

2003

Abstract

Modern scientific data repositories impose unique requirements on the underlying databases, such as very large data volumes, complex integration with external processing subsystems, and often non-predictable queries. Such challenges are commonly encountered in many applications from various disciplines, such as physics (e.g., astrophysics, geophysics), biology (e.g., ecology, genomics, proteomics), and medicine (e.g., biomedical imaging, clinical knowledge management). This thesis presents variety of novel techniques and a client/server architecture to optimize storage and processing of such data. Most importantly, we present a method to stream approximated numeric data to specialized database clients for processing and visualization. Our system architecture supports the handling and query processing of very large data volumes at any level of approximation needed. The practicability and effectiveness of these innovations has been demonstrated with the HEDC (RHESSI Experimental Data Center), a multi-terabyte repository built as part of this thesis for the NASA RHESSI satellite. The experimental validation of HEDC has proven that these techniques significantly accelerate data analysis and visualization, reduce server load, and keep system overhead (I/O, transfer across network, processing) low. Our methodology is generic in the sense that processing and visualization of any numeric attribute may be optimized, if approximated values make *sense* in that context.

The main contributions of this thesis are

- an extension mechanism for generic database systems to build *approximated materialized views*
- a method to provide multi-resolution *trial-and-error analysis* over very large data spaces
- a method for truly *interactive database visualization* of millions of tuples
- a *generic 3-tier architecture* to provide such services transparent to external programs and users, while at the same time coping with the typical changes in processing algorithms, computing environments, database schemas, and usage patterns
- a production system based on this architecture, the *HEDC*

Abstract

- the *StreamCorder*, a HEDC client written in Java that offers progressive data processing and interactive database visualization

Kurzfassung

Wissenschaftliche Informationssysteme stellen besondere Anforderungen an die ihnen zugrunde liegende Datenbank, wie zum Beispiel ein sehr grosses Datenvolumen, eine komplexe Integration externer Analysesysteme und oft schwer vorherzusagende Suchen. Diese Art von Problemen sind typisch für Informationssysteme in einer Vielzahl von Fachdisziplinen, wie der Physik (Astrophysik, GeoPhysik), der Biologie (Ökologie, Genetik) und der Medizin (bildgebende Verfahren, klinische Datenverwaltung). In dieser Doktorarbeit werden eine Reihe von neuen Methoden und eine Client/Server Architektur vorgestellt, um die Verwaltung und Nutzung solcher Systeme zu optimieren. Insbesondere werden Techniken zur approximativen Verarbeitung von numerischen Attributen innerhalb der Datenbank vorgestellt. Die Effektivität der Methodik konnte beim HEDC (RHESSI Experimental Data Center) unter Beweis gestellt werden, dem im Rahmen dieser Dissertation entwickelten, mehrere TeraByte grosse Datenzentrum des NASA RHESSI Satelliten. Experimente belegen, dass die in dieser Arbeit vorgestellten Methoden nicht nur Analyse und Visualisierung entscheidend beschleunigen, sondern auch die Serverlast (Suchen, I/O, Netzwerkübertragung, Datenverarbeitung) deutlich verringert. Die Methodik ist verallgemeinerbar in dem Sinne, dass die Verarbeitung und Visualisierung jedweder numerischer Daten optimiert werden kann, so weit approximative Werte *Sinn* machen. Die wichtigsten Beiträge dieser Arbeit sind:

- eine Methode handelsübliche Datenbanken mit *approximativen Materialisierten Sichten* zu erweitern
- eine Methode zur progressiven Analyse sehr grosser Datenmengen
- eine Methode zur *interaktiven Visualisierung* von Millionen Datensätzen
- eine 3-Tier *System Architektur* um solche Methoden für beliebige Datenbanken innerhalb der üblicherweise sehr dynamischen Umgebung von Analysealgorithmen, Rechenumgebungen, Datenformaten und Benutzermuster einzusetzen
- ein Produktionssystem, das *HEDC*, welches mit dieser Architektur entwickelt wurde

Kurzfassung

- der *StreamCorder*, ein in Java entwickelter HEDC Klient, der alle oben erwähnten Methoden zur interaktiven Analyse und Visualisierung einsetzt