



Halle 18. Stand EG E20

Internet-Zugriff auf multidimensionale Terabyte-Datenbanken

EU-Standard für Supercomputing-Datenbanken

(München) Eine Initiative großer Supercomputing-Zentren hat jetzt begonnen, einen europäischen Standard für den performanten, flexiblen Zugriff auf Multi-Terabyte-Datenbanken zu entwickeln. Von vornherein wird dabei nicht nur konzeptuelle Arbeit geleistet, sondern unter konsequenter Einbeziehung der Nutzer streng auf die Praxistauglichkeit geachtet. Eine Familie unterschiedlicher Client-Applikationen, basierend auf einer gemeinsamen Datenbank-Plattform, dient der Evaluierung. Die Realisierung nutzt das multidimensionale Datenbanksystem RasDaMan, dessen neue Java-Schnittstelle RasJ zur E-Commerce-Kopplung auf der Hannover-Messe vorgestellt wird.

Die "Speicherschweine" unter den Informationstypen sind längst nicht mehr die klassischen betriebswirtschaftlichen Daten, Stücklisten etc., sondern Sensor- und Bilddaten (z.B. Satellitendaten und Strömungssimulationen) sowie Statistikdaten (z.B. Data Warehouses). Gigabyte- (GB-) Datenbanken sind heute Alltag, Terabyte-Datenbanken (1 TB = 1.000 GB) die aktuelle Herausforderung und Petabyte-Datenbanken (1 PB = 1.000.000 GB) ein wichtiges Forschungsgebiet. Zum Vergleich: bei heutiger Plattentechnologie (18 GB) belegt ein PB über 50.000 Platten. Der wesentliche Engpaß entsteht nicht einmal bei der Speicherung an sich, sondern vielmehr beim Zugriff durch den Anwender. Insbesondere die Netzwerke werden durch die großen Datenmengen völlig überlastet; die Folge: Der Anwender wartet und gibt irgendwann auf - und viele wichtige Analysen finden nicht statt.

In dem dreijährigen EU-Projekt ESTEDI (*European Spatio-Temporal Data Infrastructure for High-Performance Computing*, URL: www.estedi.org), das in diesem Februar gestartet ist, haben sich daher bedeutende Supercomputing-Zentren in Europa, USA und Rußland zusammengefunden, um mit Europas führenden Spezialisten in multidimensionaler Datenbanktechnologie dieses Problem zu lösen. Im ersten Schritt wird unter der Leitung von ERCOFTAC (*European Research Community on Flow, Turbulence and Combustion*) eine umfassende Anforderungsanalyse durchgeführt; als deutsche Partner sind daran die DLR als Betreiber von Europas größtem Satellitenbildarchiv und das Deutsche Klimarechenzentrum beteiligt. In enger Zusammenarbeit mit den *High-Performance Computing*- (HPC-) Zentren sowie industriellen Partnern erstellen daraufhin die Datenbank-Experten von FORWISS (Bayerisches Forschungszentrum für wissenschaftliche Systeme) und Active Knowledge GmbH die Schnittstellenspezifikation, welche vollständig offengelegt wird, so daß jeder Datenbank- und Werkzeughersteller darauf Bezug nehmen kann.

Die prototypische Realisierung der ESTEDI-Plattform erfolgt durch die beiden Datenbank-Partner auf Basis des multidimensionalen Datenbanksystems RasDaMan (siehe Kasten "RasDaMan"). Zur Evaluierung entwickelt jeder HPC-Partner für sein Spezialgebiet eine Applikation auf der generischen Plattform, die sich im Alltagseinsatz bei seinen Kunden bewähren muß.

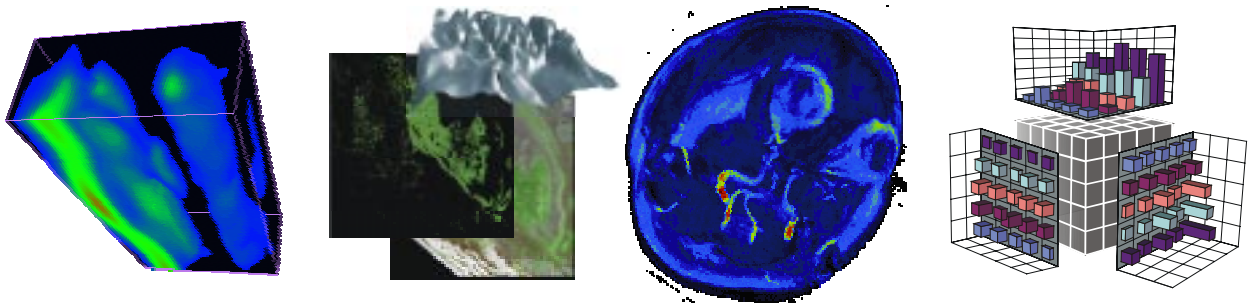
In der *User Interest Group* (UIG) werden nicht nur die initialen Anforderungen aufgenommen, sondern auch die Ergebnisse laufend von Praktikern bewertet. Neben so bekannten Namen wie CERN

und Parsytec befindet sich auch ein US-Partner in der UIG. Vielleicht ist ESTEDI mit seiner Ausdehnung von USA bis Rußland damit das weitgespannteste EU-Projekt aller Zeiten.

Alle wesentlichen Aspekte des Datenmanagements werden beleuchtet, so z.B. interaktiver E-Commerce mit Satellitendaten, Data Mining auf Klima-Modellen und genetischen Daten sowie insbesondere einfacher, flexibler Nutzerzugang zu extremen Datenvolumina. Die Anwendungsgebiete reichen von Strömungsdynamik, Erdbeobachtung, Chemical Process Engineering und Genetik bis zu Kosmologie, wo die weltweit exakteste Simulation des Universums als Testfall dienen wird.

Naturgemäß sind Performanceaspekte ein zentrales Thema, das z.B. mit intelligenter Optimierung und parallelen Datenbank-Engines angegangen wird. Aber auch der nutzerorientierte, flexible Zugang zu solchen Terabyte-Datenbanken bildet ein wichtiges Thema; Java-Schnittstellen und Visualisierungstechniken bilden hier wichtige Eckpunkte. Durch die Anbindung von Bandarchiven an die Server-Engine lassen sich praktisch unlimitierte Datenbankgrößen realisieren, welche jedoch durch neu zu entwickelnde Optimierungstechniken gesteuert werden müssen.

Letztendlich versprechen sich Experten von diesem interdisziplinären, weltweit umfassendsten Ansatz einen pan-europäischen Information-Highway für multidimensionale Daten, der diesen Namen auch verdient.



RasDaMan

RasDaMan ist ein Client/Server-Datenbanksystem, das auf Meß-, Bild- und Statistikdaten beliebiger Dimension spezialisiert ist. Die Software ist optimiert für Inhouse-Entwickler und Unternehmen, die nutzerspezifische Web-Informationssysteme erstellen. Schlüsselvorteile von RasDaMan sind schnelle Anwendungsentwicklung, höhere Flexibilität und Performance sowie deutliche Kosteneinsparungen an Netz- und Client-Hardware durch das innovative Architekturprinzip "what you get is what you need". Die serverbasierte Anfrageauswertung benutzt neuartige Techniken zur intelligenten Anfrage- und Speicheroptimierung sowie transparente Kompression. RasDaMan ist interoperabel mit konventionellen relationalen und objektorientierten Datenbanksystemen und ermöglicht damit erstmalig die integrierte Verwaltung alphanumerischer und multidimensionaler Daten. Unterstützte Plattformen sind mehrere Unix-Derivate - darunter Linux - sowie Windows NT.

Die international patentierte Datenbank-Innovation "made in Germany" hat hochrangige Innovationspreise erhalten, darunter im Jahr 1998 den begehrten European IT Prize. Entstanden ist diese Datenbank-Innovation bei FORWISS, dessen Datenbankspezialisten schon seit vielen Jahren internationale Reputation besitzen; seit zwei Jahren wird RasDaMan von der FORWISS-Ausgründung Active Knowledge GmbH erfolgreich vertrieben und weiterentwickelt. In ESTEDI kombinieren beide Partner ihr Know-How, um die Herausforderungen komplexer Anfragen auf extremen Volumina anzugehen.

Weitere Informationen:

Active Knowledge GmbH
Dr. Peter Baumann

Tel. 089 / 48095-207, Fax -195
Email: baumann@active-knowledge.de



Diesen Text können Sie von www.active-knowledge.de abrufen, indem Sie auf der Leitseite dem Link "+++ News +++" folgen.