Higher Order Databases

Database Technology Everywhere

Hans-Jörg Schek
ETH Zürich

Overview of the Talk

- Introduction: Situation and Perspectives
- What are Higher Order Databases?
- Research in Higher Order Databases
  - PowerDB
  - Transactional Process Management
  - Components und Subsystems in Multimedia Information Systems

Situation and Perspectives - PITAC

Presidential Information Technology Advisory Committee (PITAC) (aus Michael Brodie, GTE, Vortrag "Que Sera Sera"):

Findings on IT

- **Produced 1/3 of US economic growth since 1992. Spectacular return on R&D investment of 20 years ago!**
- **Key driver of 21st century progress**
- **Will transform the way we live, learn, work, and play**
- **R&D funding inadequate - flat for a decade, near- term focus**
- **Recommendations (Feb. 99)**
- **2 X US Govt. R&D funding over five years (+$ 1.3 B)**
- **Focus on longer term challenges**

Situation and Perspectives - VLDB End.

Extract of my Presentation 1999 at the VLDB Endowment

THE Question: What is our (database) field?

Is it

- **"Hard core" database management system implementation? As described in DBMS textbooks?**
- **Data servers as replacement of file servers?**

Or…

- **Providing infrastructure for (distributed) information system development?**

Situation and Perspectives - VLDB End. 2

Future infrastructure for applications?

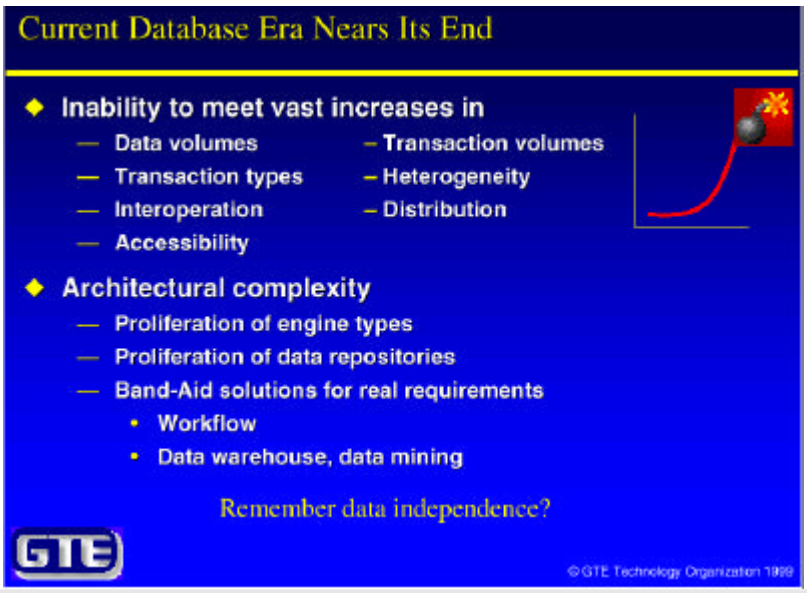
- Managing components, application services?
- Distributed client/middleware/server computing?
- Application frameworks, ERPs?
- Role in XML, e-commerce?

Role of DBMSs

- Storage manager only in all these cases.
- Our position has moved far away from the applications.
- Client application development takes place on other platforms.

Situation and Perspectives - Brodie

Brodie from his Talk:



Situation and Perspectives - Carey et al.

Carey
Hellerstein
Stonebraker

- Why we need to rethink everything!
  - All current DBMSs architected in the late 1970s
  - why the world is different now

Seminar at the University of Berkeley, 1999

- CPU, memory, disk up by 10 ** 6 in the last 20 years
- Design point of 1 Tbyte buffer pool in 2005, up from 1 Mbyte in the 1970s
- It will NOT be 250 million 4K pages!

Situation and Perspectives - Carey et al. 2

Carey
Hellerstein
Stonebraker

- Most serious applications use a TP monitor
- I.e. a three tier application architecture
  - data at the bottom in a DBMS
  - code in middle tier in TP monitor
  - user interface on the client

- DBMSs are currently “bloated”
  - stored procedures
  - object-relational features
  - warehouse features
  - triggers
  - standard benchmark hacks

Situation and Perspectives - Carey et al. 3

Carey
Hellerstein
Stonebraker

Assumptions -- Must Design for a Data and Machine Federation

- 7 X 24 operation requires wide area replication
  - understood by the DBMS
  - transactionally consistent
  - fastest mechanism is to move the log

- Incredible scalability requires more than the biggest single system

Situation and Perspectives - Carey et al. 4

Carey
Hellerstein
Stonebraker

Advantages of a Federated DBMS

- Mimics the enterprise, which is distributed
- Naturally supports mergers
- Allows “jelly bean” hardware components
- Can be incrementally built and extended

Situation and Perspectives - ETHZ/DBS

ETHZ-DBS-Answer:

Higher Order Databases (DBHO)

What is a Higher Order Database?

First rough definition:

- A Higher Order Database (DBHO) is a Database over Databases.
- A DBHO administers
  - Objects that are composed of objects
  - Transactions that are composed of transactions
- In other terms, a DBHO is a "Hyperdatabase", i.e. the primitives are again databases. (For comparison: "Hypermatrices are matrices whose elements are matrices")
- A DBHO administers distributed components in a networked environment and provides a kind of higher order "Data Independence"

"Academic game with words?

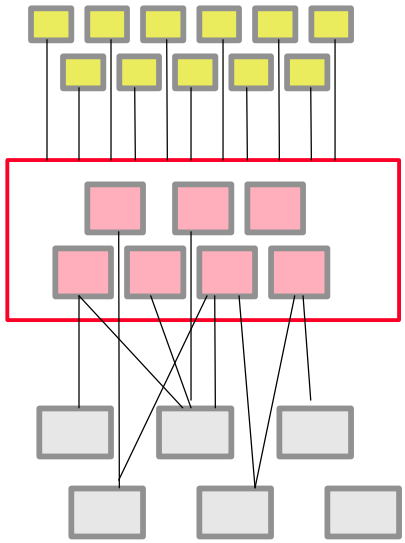
Why is a DBHO more than just a DB?

- Reminder: What is a database anyway?
  - Platform for clients concurrently accessing shared data
  - Needed: data definition, data manipulation, transactions at the interface.
  - Under the cover: query optimization, correctness for parallel access, recovery, persistency, load balancing, admission control, availability,...

Why is a DBHO more than just a DB? -2-

- Database of higher order:
  - Platform für clients, concurrently accessing shared application services
  - Needed: service definition and description, service customization, transactional processes encompassing multiple service invocations.
  - Under the cover: optimization of client requests, routing, scheduling, parallelization, correctness of concurrent accesses, flexible failure treatment, providing guaranteed termination, availability, flexible recovery, scalability,...

DBHO Simple Big Picture



Clients

access application services/components

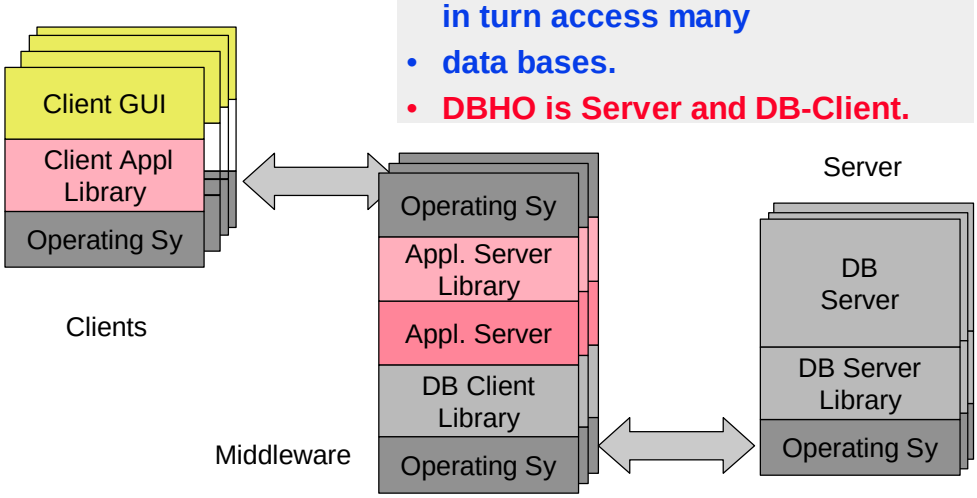
DBHO is a coordinator

managing concurrent service calls and/or distributes tasks to

Data Server
executes concurrent data access and stored procedure calls

DBHO in Middleware Presentation

- **Lots of clients access**
- **many application servers. These in turn access many**
- **data bases.**
- **DBHO is Server and DB-Client.**



DBHO - is there anything new here?

- The following examples provide the one or the other functionality of DBHO
  - TP-Monitors
  - Enterprise Resource Planning Systems
  - Extensible Databases
  - Federated DB and Multi-Databases
  - Data Warehouses
  - Multimedia DB
  - Database Cluster

Different fields?

Is there hope for a more systematic, unifying view?

Answer: DBHO

Example: TP-Monitor

In a nutshell:

TP-Monitor =

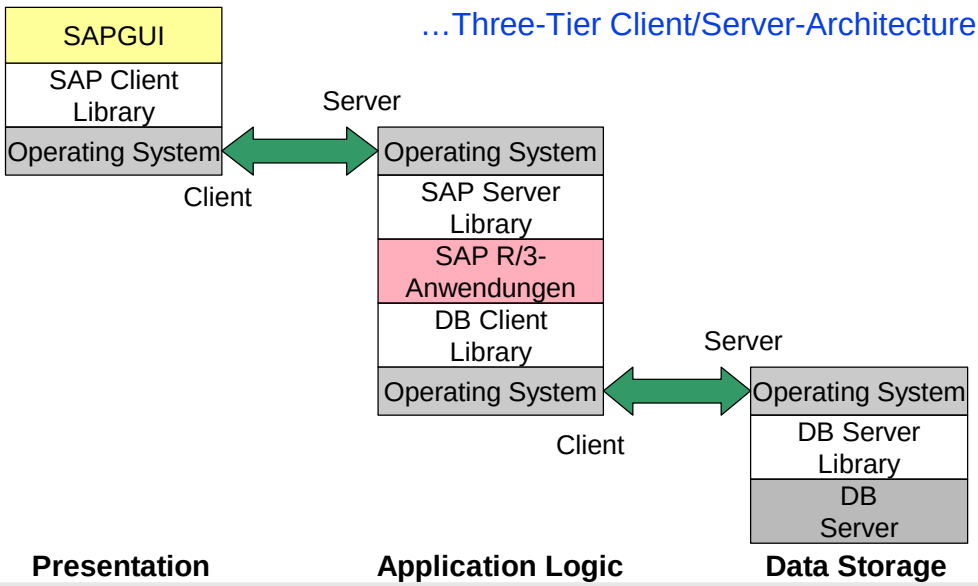
- Operating system for transaction processing and
- Infrastructure for developing and running applications, services, and components in a three-tier architecture.

A DBHO Adds

- (1) Design tools for specification of distributed applications and automatic generation of program code.
- (2) Real transaction layer above DB-transactions and transactional process management with
  - Failure treatment
  - Availability
  - Guaranteed termination and "semantic" correctness
- (3) Optimal routing to components and clever replication of whole components

Example SAP R/3

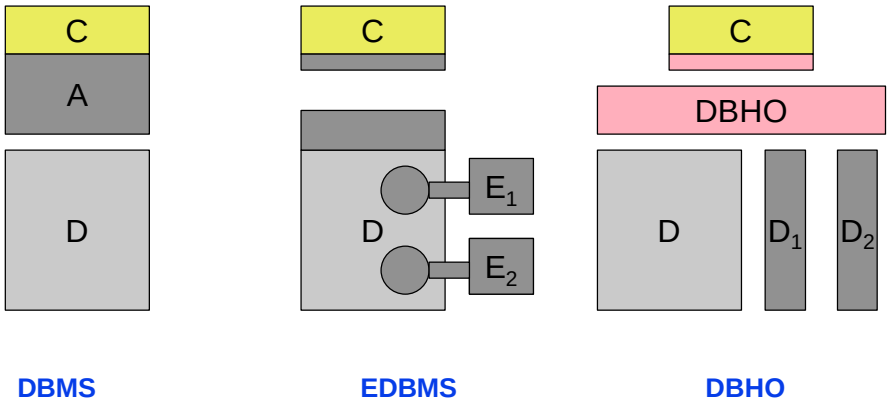
…Three-Tier Client/Server-Architecture



A DBHO Adds

- (1) Real transaction layer above DB-transactions and transactional process management with
  - Failure treatment
  - Availability
  - Guaranteed termination and “semantic” correctness
- (2) Flexible mapping to many database/storage managers
- (3) Optimal decomposition and routing of client request storage managers, clever replication of whole components

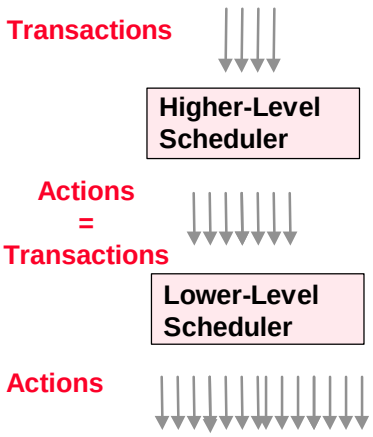
DBHO as Evolution from Extensible DB



DBHO Research necessary?

- Composite transactions
- Transactional process management
- DB Cluster
- Multimedia databases, specialized components
- Generic programming, generic agents
- Tools for specifying distributed applications
- Tools for code generation
- Tools for supporting customization
- more...

Research at ETHZ/DBS: Composite Transactions



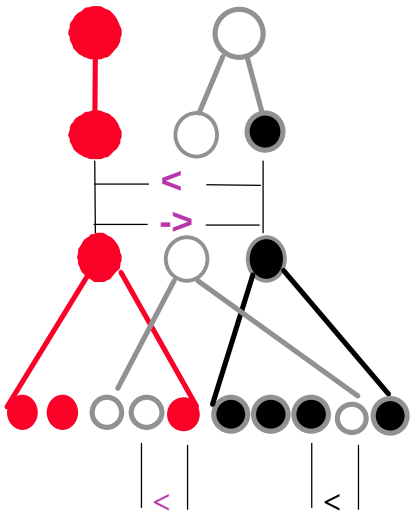
Our aim is to let several schedulers work as independently as possible

What ordering information from one scheduler is passed to the other?

What kind of orderings must be distinguished?

How can we make sure that the serialization order of one scheduler does not contradict the serialization order of the other?

Example



Cycle, because orders of conflict pairs are contradicting

Conflict Consistent Serializability (CCSR)

Redefine:

Schedule, well-formed schedule, serial schedule, equivalence to a serial schedule...

Theorem:

A schedule is CCSR if and only if the union of the serialization graph and the weak inter-transaction input order is acyclic.

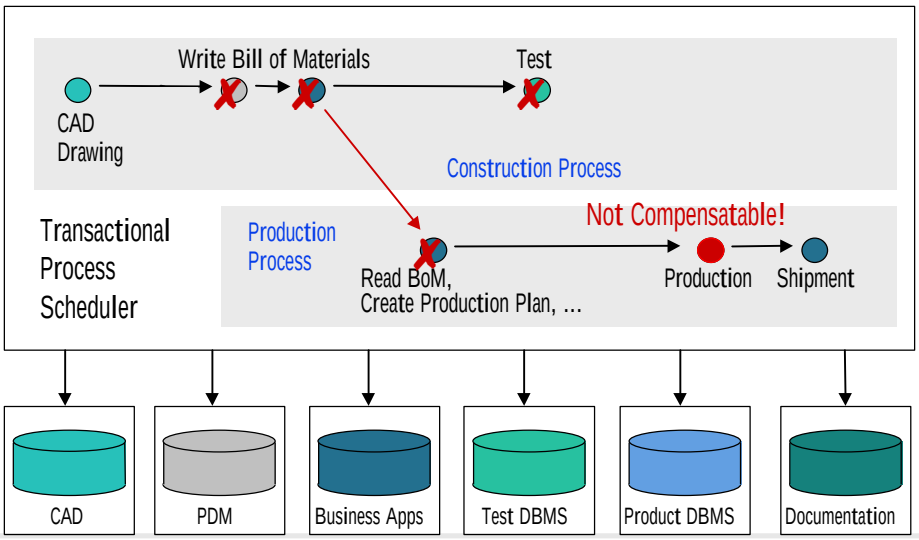
Joint work with G. Alonso

Research at ETHZ/DBS: Transactional Process Management

- Use composite transactions, add preference order for failure handling
- Distinguish different termination properties of activities
  - Compensatable:Effects can be undone (by inverse activity)
  - Retriable: Guaranteed to commit after finite number of invocations
  - Pivot: No inverse activity available

Joint work with G. Alonso

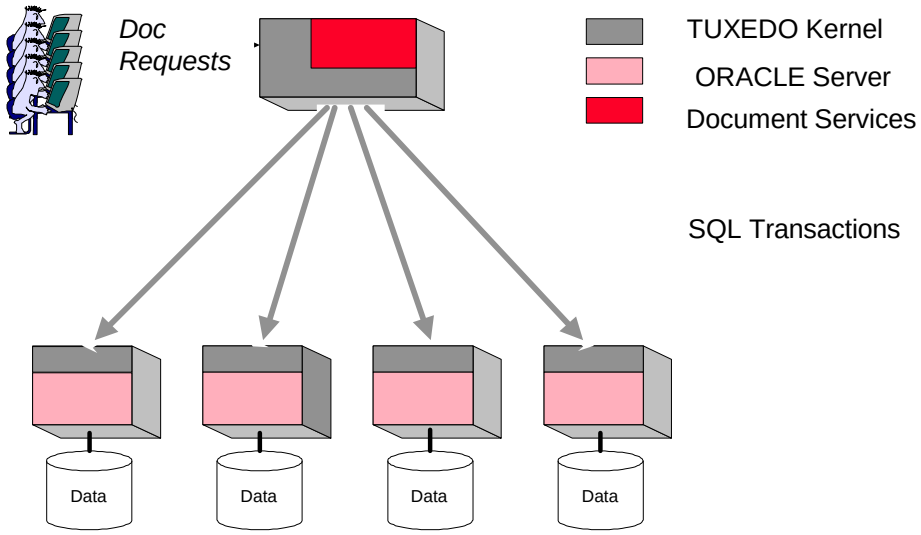
Example: CIM Scenario



Research at ETHZ/DBS: PowerDB
2^{DB}

- A powerset of databases
- A DBMS made of a PC Cluster of Databases
- A RISC approach: Build a powerful DBMS using many simple DBMSs
- A coordination DBMS of component DBMSs

Example: DocEngine Prototype System



DocEngine Measurements

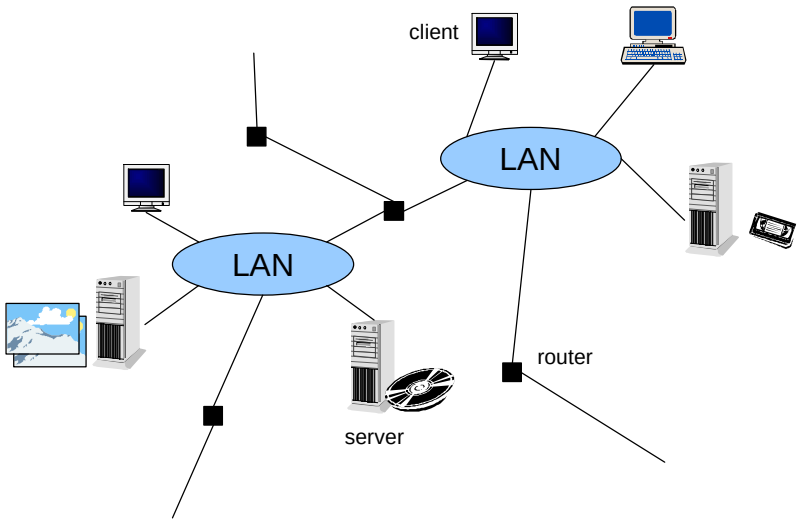
Insertion speed-up from 1 to 4 components

Placement / Workload	DISTAB	HASHLOC
(1,1)	2.8	2.5
(5,5)	4.2	6.3
(10,10)	6	12

Retrieval speed-up from 1 to 4 components

Placement / Workload	DISTAB	HASHLOC
(1,1)	2.7	2.7
(5,5)	2.5	6.4
(10,10)	2.5	15

Research at ETHZ/DBS: Multimedia Components, Image Indexing



Coordination of Specific Image Indexing Components

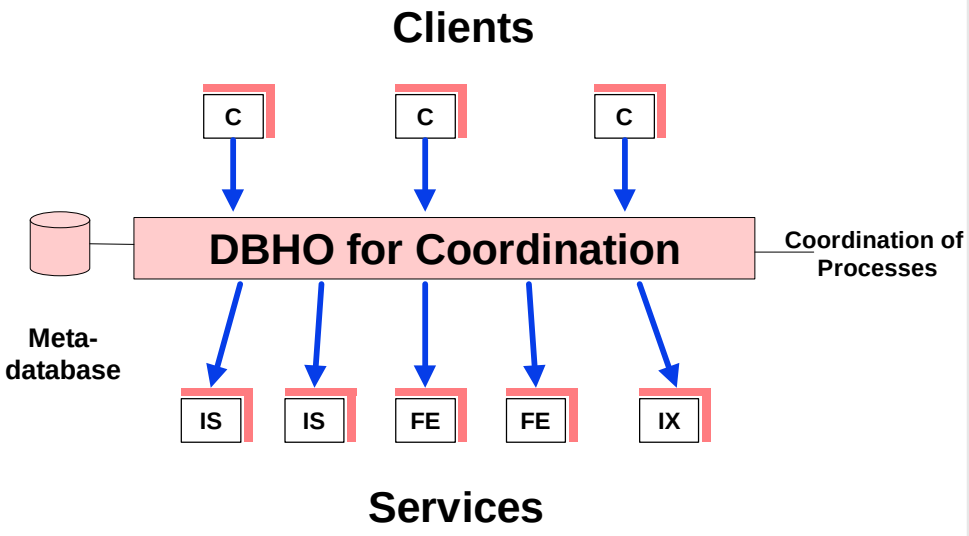


Image Similarity Search Demonstration

Demonstration will show

- Distributed Architecture
- Managing components
- VA-File, a new technique for similarity search in high-dimensional feature spaces

Acknowledgement

- Mostly: Discussions with Prof. Gustavo Alonso, ongoing joint work
- Early discussions with Prof. Gerhard Weikum, Prof. Moira Norrie, Markus Tresch
- ACM /MIT DB Working Group "Databases breaking out of the box"
- VLDB Endowment "Future Directions Group"

- Last but not least: DBS Group: Klemens Böhm, Sebastian Duda, Torsten Grabs, Evangelos Kotsakis, Uwe Röhm, Andrej Popovici, Marco Schmidt, Heiko Schuldt, Christoph Schuler, Roger Weber.

References

- Michael Brodie, Que Sera Sera, Keynote Talk at the DOM Workshop, 1999
- Michael Carey, Joseph Hellerstein, Michael Stonebraker, 1999 Seminar at the University of Berkely
- R. Orfali, D. Harkey, J. Edwards: *The Essential Client/Server Survival Guide*, John Wiley, New York, 1999
- K. Boucher, Fima Katz: *Essential Guide to Object Monitors,* John Wiley, New York, 1999
- List of Publications and Class Notes of the DBS Group at www-dbs.inf.ethz.ch