

Datenbanksysteme II: Implementation of Database Systems

Ulf Leser

Material von

Prof. Johann Christoph Freytag

Prof. Kai-Uwe Sattler

Prof. Alfons Kemper, Dr. Eickler



-
- I apologize for Denglish

My first contact with database

- Father: “I need to track my customers”
 - Name, address, profession, prior contracts
- Little Ulf: “No problem”
 - ~1984: Turbo Pascal 4.0, Schneider CPC646, 512 KB main memory
 - Load customers from disk into memory
 - Repeat until “Q”
 - (E)dit customer
 - (D)elele customer
 - (I)nsert customer
 - (Q)uit
 - Write customers to disk
 - Bill: 250.- DM

Story part 2

- Father: “I need to track my offers”
 - Customers have projects and call for bids, father makes offers
 - Many customers have many projects **over time**
- Little Ulf: “No problem”
 - Clever plan – reuse “software architecture”
 - Load offers from disk into memory
 - Repeat until “Q”
 - (E)dit offer
 - (D)elete offer
 - (I)nsert offer
 - (Q)uit
 - Write offers to disk
 - My (winning) offer: 400.- DM

Part 3

- Disaster!
 - No dynamic arrays at this time
 - Max. 500 customers
 - Max. 40 offers
 - Max. 2 KB per offer
 - Gives 20.000 times 2.000 = 40 MEGABYTE!
 - App. size of hard disc
 - No way to load and hold all data at startup
 - Wrong architecture

Solution 1

- Repeat until "Q"
 - (E)dit offer
 - Open file
 - Read and write until record found
 - Read, edit, and write record
 - Read and write until EOF
 - (D)elele offer
 - ...
 - (I)nsert offer
 - ...
 - (Q)uit
- Very slow

Solution 2

- Some one should take care of finding, loading, storing data for me
- Buy dBase IV

Lessons learned

- **Scalability** is an issue
 - Works fine for small datasets, breaks down for large datasets
- **Data is Information is Knowledge is essential**
 - If offers-file corrupted – father goes out of business
 - Which will affect me too
- When it comes to large volumes of data – get a database and a **good database programmer**
- Think before you program
 - Project 100% over budget (dBase IV license)
 - Project 300% over time (6 months instead of 2)

Today's Database Systems

- Databases are essential part of an **enterprise's infrastructure**
 - As important as OS
 - Long-running, expensive, essential investment
 - Holds the most important business asset: information
- Database administrator is a profession
 - Many good programmers, but few good database developers/administrators
 - Always one of the skills most demanded in industry
- RDMS became an often **"invisible" piece of software**
 - "So nützlich wie fließendes Wasser" (G. Weikum, MPI Saarbrücken)
- DBAs are very complex, though apparently simple to use
- Good DBAs need to **understand their system**

History

- Prior to relational databases
 - Network databases, hierarchical databases, COBOL
- “Three waves”
 - Relational databases since 1980th
 - Object-oriented databases early 1990th
 - XML databases early 2000th
- Few „big“ vendors survived
 - Oracle, IBM DB2, Microsoft SQL Server
- Some niche players
 - Open Source: MySQL, PostgreSQL, Ingres, SAP DB
 - Data Warehousing: Teradata, Tandem
 - Mobile databases: Sybase

Importance

- Databases are behind almost every non-trivial piece of software
 - Data needs to be stored
 - Disk access is the main performance bottleneck
 - Hence: [Optimize disk access](#) in the RDBMS
 - Data is manipulated from many clients
 - Concurrent access quickly screws up data
 - Hence: [Synchronize data access](#) through the RDBMS
 - Data is used by many applications with different requirements
 - No good to design app-specific optimal data structures
 - Hence: Use [application independent](#) languages
 - Systems crash
 - Crashes cannot be avoided
 - Hence: [Secure most valuable asset](#) in a centralized manner

Complexity

- A RDBMS is a **very, very complex piece of software**
 - Oracle installs from 5+ CDs
- One code base
 - Oracle9i Database Release 2 Documentation for Microsoft Windows
 - Oracle9i Database Release 2 Documentation for 64-bit Windows
 - Oracle9i Database Release 2 Documentation for Solaris Operating System (SPARC)
 - Oracle9i Database Release 2 Documentation for AIX-Based Systems
 - Oracle9i Database Release 2 Documentation for Compaq Tru64 UNIX
 - Oracle9i Database Release 2 Documentation for HP 9000 Series HP-UX
 - Oracle9i Database Release 2 Documentation for HP-UX Itanium
 - Oracle9i Database Release 2 Documentation for hp OpenVMS Alpha
 - Oracle9i Database Release 2 Documentation for IBM OS-390 (MVS)
 - Oracle9i Database Release 2 Documentation for Linux x86-64
 - Oracle9i Database Release 2 Documentation for Linux x86
 - Oracle9i Database Release 2 Documentation for Linux Itanium
 - Oracle9i Database Release 2 Documentation for IBM Power based Linux
 - Oracle9i Database Release 2 Documentation for Linux/390
 - Oracle9i Database Release 2 Documentation for Fujitsu Siemens Computers BS2000/OSD
 - Oracle9i Database Release 2 Documentation for Stratus Continuum 600 and 1200 Series

List of Books 1

Book Title	Contents	Search	Index	PDF
Access Manager for AS/400 Installation and User's Guide	Contents	Search	Index	PDF
Administrator's Guide	Contents	Search	Index	PDF
Administrator's Guide for Windows	Contents	Search	Index	PDF
Administrator's Reference for UNIX Systems: AIX-Based Systems, Compaq Tru64 UNIX, HP 9000 Series HP-UX, Linux Intel, and Sun Solaris	Contents	Search	Index	PDF
Advanced Replication	Contents	Search	Index	PDF
Advanced Security Administrator's Guide	Contents	Search	Index	PDF
Application Developer's Guide - Advanced Queuing	Contents	Search	Index	PDF
Application Developer's Guide - Fundamentals	Contents	Search	Index	PDF
Application Developer's Guide - Large Objects (LOBs)	Contents	Search	Index	PDF
Application Developer's Guide - Object-Relational Features	Contents	Search	Index	PDF
Application Developer's Guide - Workspace Manager	Contents	Search	Index	PDF
Backup and Recovery Concepts	Contents	Search	Index	PDF
Backup and Recovery Documentation Online Roadmap	Contents	Search		
Client Installation Guide for 32-Bit Windows	Contents	Search	Index	PDF
Client Release Notes for 32-Bit Windows	Contents	Search		PDF
COM Automation Feature Developer's Guide	Contents	Search	Index	PDF
Concepts	Contents	Search	Index	PDF
Data Cartridge Developer's Guide	Contents	Search	Index	PDF
Data Guard Broker	Contents	Search	Index	PDF
Data Guard Concepts and Administration	Contents	Search	Index	PDF
Data Mining Administrator's Guide	Contents	Search	Index	PDF
Data Mining Concepts	Contents	Search	Index	PDF
Data Warehousing Guide	Contents	Search	Index	PDF
Database Tuning with the Oracle Tuning Pack	Contents	Search	Index	PDF
Directory Service Integration and Deployment Guide	Contents	Search	Index	PDF
Dynamic Services User's and Administrator's Guide	Contents	Search	Index	PDF
Enterprise Edition Installation Guide for OS/390	Contents	Search	Index	PDF
Enterprise Edition Messages Guide for OS/390	Contents	Search	Index	PDF
Enterprise Edition Release Notes for OS/390	Contents	Search		PDF
Enterprise Edition System Administration Guide for OS/390	Contents	Search	Index	PDF
Enterprise Edition User's Guide for OS/390	Contents	Search	Index	PDF
Enterprise Manager Administrator's Guide	Contents	Search	Index	PDF
Enterprise Manager Concepts Guide	Contents	Search	Index	PDF

List of Books 2

Book Title	Contents	Search	Index	PDF
Enterprise Manager Concepts Guide	Contents	Search	Index	PDF
Enterprise Manager Configuration Guide	Contents	Search	Index	PDF
Enterprise Manager Event Test Reference Manual	Contents	Search	Index	PDF
Enterprise Manager Messages Manual	Contents	Search	Index	PDF
Error Messages	Contents	Search	Index	PDF
Getting Started for Windows	Contents	Search	Index	PDF
Getting Started with Oracle Change Management Pack	Contents	Search	Index	PDF
Getting Started with Oracle Management Pack for Oracle Applications	Contents	Search	Index	PDF
Getting Started with the Oracle Diagnostics Pack	Contents	Search	Index	PDF
Getting Started with the Oracle Standard Management Pack	Contents	Search	Index	PDF
Globalization Support Guide	Contents	Search	Index	PDF
Heterogeneous Connectivity Administrator's Guide	Contents	Search	Index	PDF
Installation Guide for 32-Bit Windows	Contents	Search	Index	PDF
Installation Guide for UNIX Systems: AIX-Based Systems, Compaq Tru64 UNIX, HP 9000 Series HP-UX, Linux Intel and Sun Solaris	Contents	Search	Index	PDF
Intelligent Agent User's Guide	Contents	Search	Index	PDF
interMedia Annotator User's Guide	Contents	Search	Index	PDF
interMedia Java Classes User's Guide and Reference	Contents	Search	Index	PDF
interMedia User's Guide and Reference	Contents	Search	Index	PDF
Internet Directory Administrator's Guide	Contents	Search	Index	PDF
Internet Directory Application Developer's Guide	Contents	Search	Index	PDF
Java Developer's Guide	Contents	Search	Index	PDF
Java Stored Procedures Developer's Guide	Contents	Search	Index	PDF
JDBC Developer's Guide and Reference	Contents	Search	Index	PDF
JPublisher User's Guide	Contents	Search	Index	PDF
Label Security Administrator's Guide	Contents	Search	Index	PDF
Master Glossary	Contents	Search	Index	PDF
Master Index	Contents	Search	Index	PDF
Migration	Contents	Search	Index	PDF
Migration Workbench Frequently Asked Questions (FAQ)	Contents	Search	Index	PDF
Migration Workbench Reference Guide for IBM DB2/AS400 V4R5 Migrations	Contents	Search	Index	PDF
Migration Workbench Reference Guide for Informix Dynamic Server 7.3 Migrations	Contents	Search	Index	PDF
Migration Workbench Reference Guide for Microsoft Access 2.0, 95, 97, 2000 Migrations	Contents	Search	Index	PDF
Migration Workbench Reference Guide for Microsoft SQL Server and Sybase Adaptive Server Migrations	Contents	Search	Index	PDF
Migration Workbench Reference Guide for MySQL 3.22, 3.23 Migrations	Contents	Search	Index	PDF
Migration Workbench Release Notes	Contents	Search	Index	PDF
Net Services Administrator's Guide	Contents	Search	Index	PDF
Net Services Reference Guide	Contents	Search	Index	PDF

List of Books 3

Book Title	Contents	Search	Index	PDF
Net Services Reference Guide	Contents	Search	Index	PDF
New Features	Contents	Search	Index	PDF
OLAP Developer's Guide to the OLAP API	Contents	Search	Index	PDF
OLAP Developer's Guide to the OLAP DML	Contents	Search	Index	PDF
OLAP User's Guide	Contents	Search	Index	PDF
Oracle C++ Call Interface Programmer's Guide	Contents	Search	Index	PDF
Oracle Call Interface Developer's Guide for Windows	Contents	Search	Index	PDF
Oracle Call Interface Programmer's Guide	Contents	Search	Index	PDF
Oracle Objects for OLE	Contents	Search		
Oracle Objects for OLE C++ Class Library	Contents	Search		
Oracle Transparent Gateway for IBM DB2/400 Installation and User's Guide for IBM AS/400	Contents	Search	Index	PDF
Performance Planning	Contents	Search	Index	PDF
Performance Tuning Guide and Reference	Contents	Search	Index	PDF
PL/SQL User's Guide and Reference	Contents	Search	Index	PDF
Pro*C/C++ Precompiler Getting Started for Windows	Contents	Search	Index	PDF
Pro*C/C++ Precompiler Programmer's Guide	Contents	Search	Index	PDF
Pro*COBOL Precompiler Getting Started for Windows	Contents	Search	Index	PDF
Pro*COBOL Precompiler Programmer's Guide	Contents	Search	Index	PDF
Pro*FORTRAN Supplement to the Oracle Precompilers Guide	Contents	Search		PDF
Pro*PL/I Supplement to the Oracle Precompilers Guide	Contents	Search	Index	PDF
Procedural Gateway and Tools for IBM MQSeries for MS Windows	Contents	Search	Index	PDF
Procedural Gateway and Tools for IBM MQSeries for UNIX	Contents	Search	Index	PDF
Procedural Gateway for APPC Installation and Configuration Guide for UNIX	Contents	Search	Index	PDF
Procedural Gateway for APPC Installation and Configuration Guide for Windows	Contents	Search	Index	PDF
Procedural Gateway for APPC Messages Guide	Contents	Search	Index	PDF
Procedural Gateway for APPC User's Guide for UNIX	Contents	Search	Index	PDF
Procedural Gateway for APPC User's Guide for Windows	Contents	Search	Index	PDF
Programmer's Guide to the Oracle Precompilers	Contents	Search		PDF
Provider for OLE DB Developer's Guide	Contents	Search	Index	PDF
Real Application Clusters Administration	Contents	Search	Index	PDF
Real Application Clusters Concepts	Contents	Search	Index	PDF
Real Application Clusters Deployment and Performance	Contents	Search	Index	PDF
Real Application Clusters Documentation Online Roadmap	Contents	Search		
Real Application Clusters Real Application Clusters Guard I - Concepts and Administration	Contents	Search	Index	PDF
Real Application Clusters Real Applications Clusters Guard I Configuration Guide Release 2 (9.2.0.1.0) for UNIX Systems: AIX-Based	Contents	Search	Index	PDF

List of Books 4

Book Title	Contents	Search	Index	PDF
Real Application Clusters Real Applications Clusters Guard I Configuration Guide Release 2 (9.2.0.1.0) for UNIX Systems: AIX-Based Systems, Compaq Tru64 UNIX, HP 9000 Series HP-UX, and Sun Solaris	Contents	Search	Index	PDF
Real Application Clusters Setup and Configuration	Contents	Search	Index	PDF
Recovery Manager Quick Reference	Contents	Search	Index	PDF
Recovery Manager Reference	Contents	Search	Index	PDF
Recovery Manager User's Guide	Contents	Search	Index	PDF
Reference	Contents	Search	Index	PDF
Release Notes Addendum for Windows	Contents	Search	Index	PDF
Release Notes for 32-Bit Windows	Contents	Search	Index	PDF
Release Notes for AIX-Based 4.3.3 Systems (64-bit)	Contents	Search	Index	PDF
Release Notes for AIX-Based 5L Systems (64-bit)	Contents	Search	Index	PDF
Release Notes for Compaq Tru64 UNIX (64-bit)	Contents	Search	Index	PDF
Release Notes for HP 9000 Series HP-UX (64-bit)	Contents	Search	Index	PDF
Release Notes for Linux Intel (32-bit)	Contents	Search	Index	PDF
Release Notes for Sun Solaris (32-bit)	Contents	Search	Index	PDF
Release Notes for Sun Solaris (64-bit)	Contents	Search	Index	PDF
Replication Management API Reference	Contents	Search	Index	PDF
Sample Schemas	Contents	Search	Index	PDF
Security and Network Integration Guide for Windows	Contents	Search	Index	PDF
Security Overview	Contents	Search	Index	PDF
Services for Microsoft Transaction Server Developer's Guide	Contents	Search	Index	PDF
SNMP Support Reference Guide	Contents	Search	Index	PDF
Spatial User's Guide and Reference	Contents	Search	Index	PDF
SQL Reference	Contents	Search	Index	PDF
SQL*Module for Ada Programmer's Guide	Contents	Search	Index	PDF
SQL*Plus Getting Started for Windows	Contents	Search	Index	PDF
SQL*Plus Quick Reference	Contents	Search	Index	PDF
SQL*Plus User's Guide and Reference	Contents	Search	Index	PDF
SQLJ Developer's Guide and Reference	Contents	Search	Index	PDF
Streams	Contents	Search	Index	PDF
Supplied Java Packages Reference	Contents	Search	Index	PDF
Supplied PL/SQL Packages and Types Reference	Contents	Search	Index	PDF
Support for JavaServer Pages Reference	Contents	Search	Index	PDF
Syndication Server User's and Administrator's Guide	Contents	Search	Index	PDF
Text Application Developer's Guide	Contents	Search	Index	PDF
Text Reference	Contents	Search	Index	PDF
Transparent Gateway for DB2 Installation and User's Guide for IBM OS/390	Contents	Search	Index	PDF

List of Books

Oracle9i Database List of Books (Release 2 (9.2)) - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://otn.oracle.com/pls/db92/db92.docindex?remark=homepage

Nachsehen DBLP MyStuff Google Lehre News Suchen Buecher kaufen Reisen Paper suchen

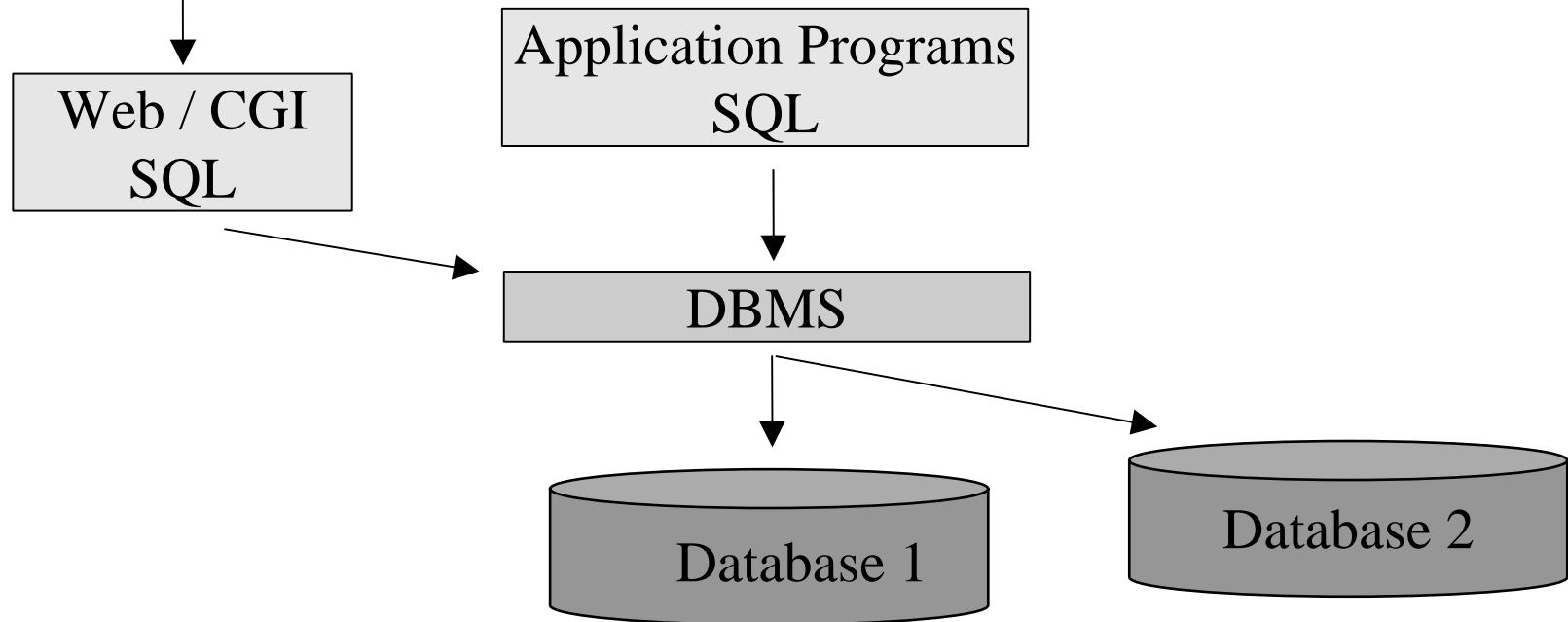
Oracle9i Database List of Books (Rel...

Transparent Gateway for DB2 Installation and User's Guide for IBM OS/390	Contents / Search / Index / PDF
Transparent Gateway for IBM DRDA Installation and User's Guide for UNIX	Contents / Search / Index / PDF
Transparent Gateway for IBM DRDA Installation and User's Guide for Windows	Contents / Search / Index / PDF
Ultra Search Online Documentation	Contents / Search / /
Universal Installer Concepts Guide	Contents / Search / Index / PDF
User-Managed Backup and Recovery Guide	Contents / Search / Index / PDF
Utilities	Contents / Search / Index / PDF
Visual Workbench for Oracle Procedural Gateways for IBM MQSeries Installation and User's Guide for Windows NT	Contents / Search / / PDF
Workflow Client Installation Notes (release 2.6.2)	Contents / Search / / PDF
Workflow Guide (release 2.6.2)	Contents / Search / / PDF
Workflow Server Installation Notes (release 2.6.2)	Contents / Search / / PDF
XML API Reference - XDK and Oracle XML DB	Contents / Search / Index / PDF
XML Database Developer's Guide - Oracle XML DB	Contents / Search / Index / PDF
XML Developer's Kits Guide - XDK	Contents / Search / Index / PDF

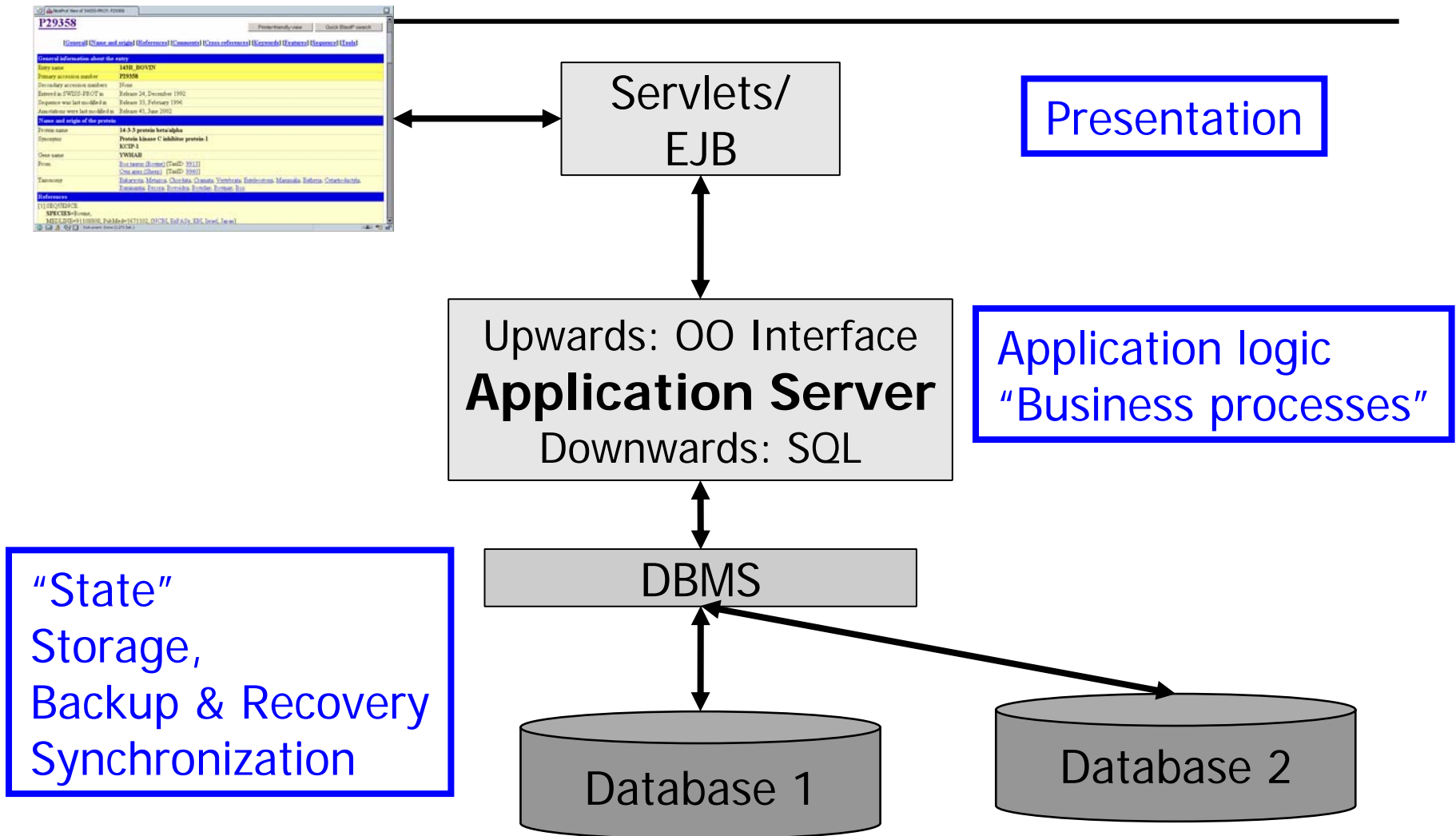
Much more than „SELECT ... FROM ... WHERE“

- Text, Multimedia, spatial data
- Object-relational features
- XML, XQuery, XPath
- Parallel and distributed databases; Cluster; GRID
- Uniform access or uniform storage
 - File system goes database
- Data Mining and Business Intelligence
- Backup & Recovery
 - Try to make your application fault tolerant ...
- Security, Auditing
- DBs are infrastructure for application server

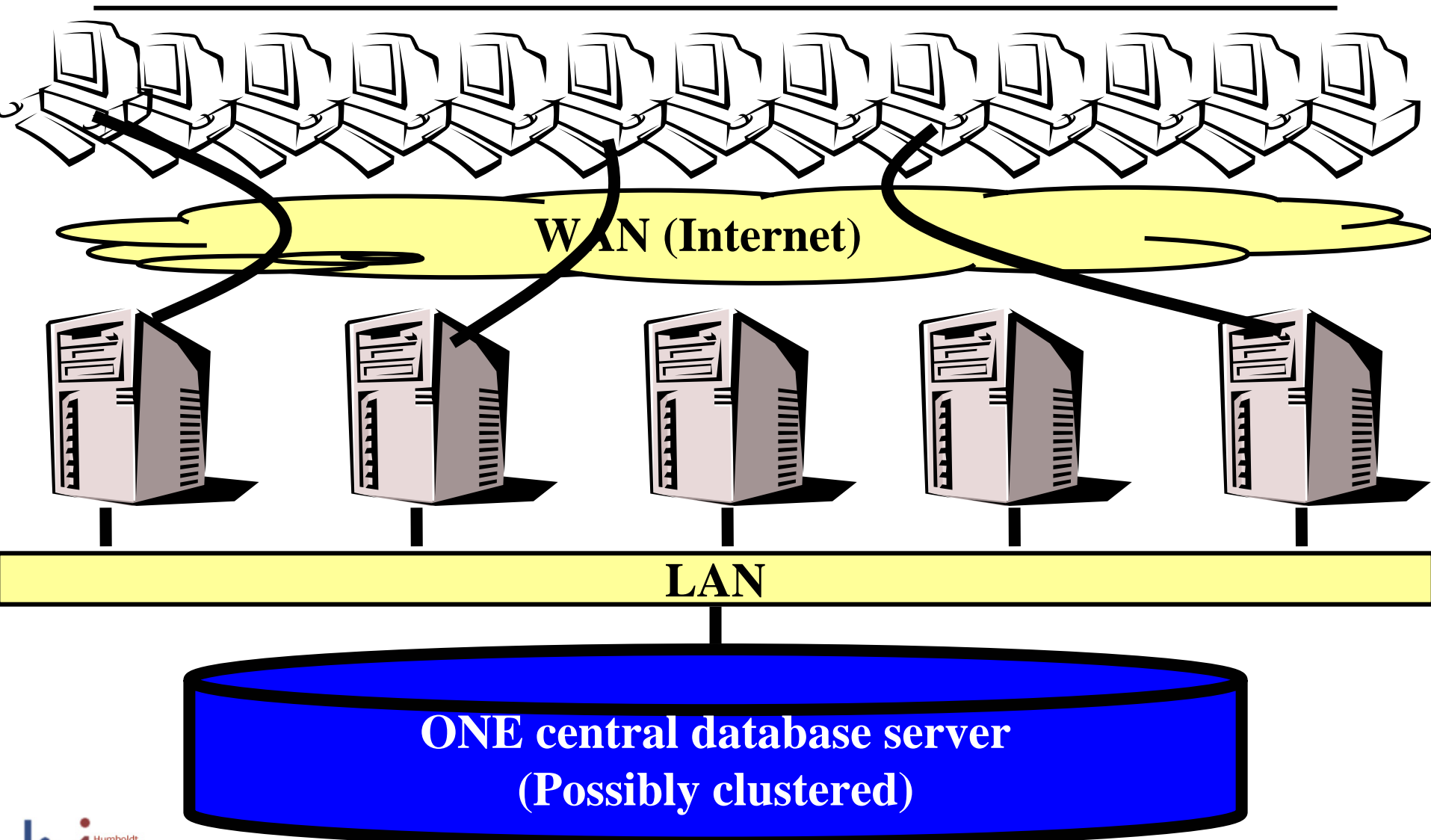
Database System – 2 Tier



3 Tier



SAP R/3: Enterprise Resource Planning



Why Databases?

- **Data is extremely valuable**
- Many programs do **different things** with the same data
 - Centralize data management
 - Avoid (uncontrolled) data redundancy
 - App-independent data model
 - App-independent access methods
 - Centralize control of data integrity
- Many users operate **concurrently** on the same data
 - Avoid blocking and inconsistencies
 - Authorization and Authentication
 - Support transactions
- Amount of data grows and grows and grows - **scalability**
 - Clients working on the same copy
 - Central query optimization, caching, connection sharing, etc.
 - Controlled, automated replication
- Data must not disappear – **fault tolerance**
 - Centralize data replication for safety
 - Centralize backup & recovery

Scalability

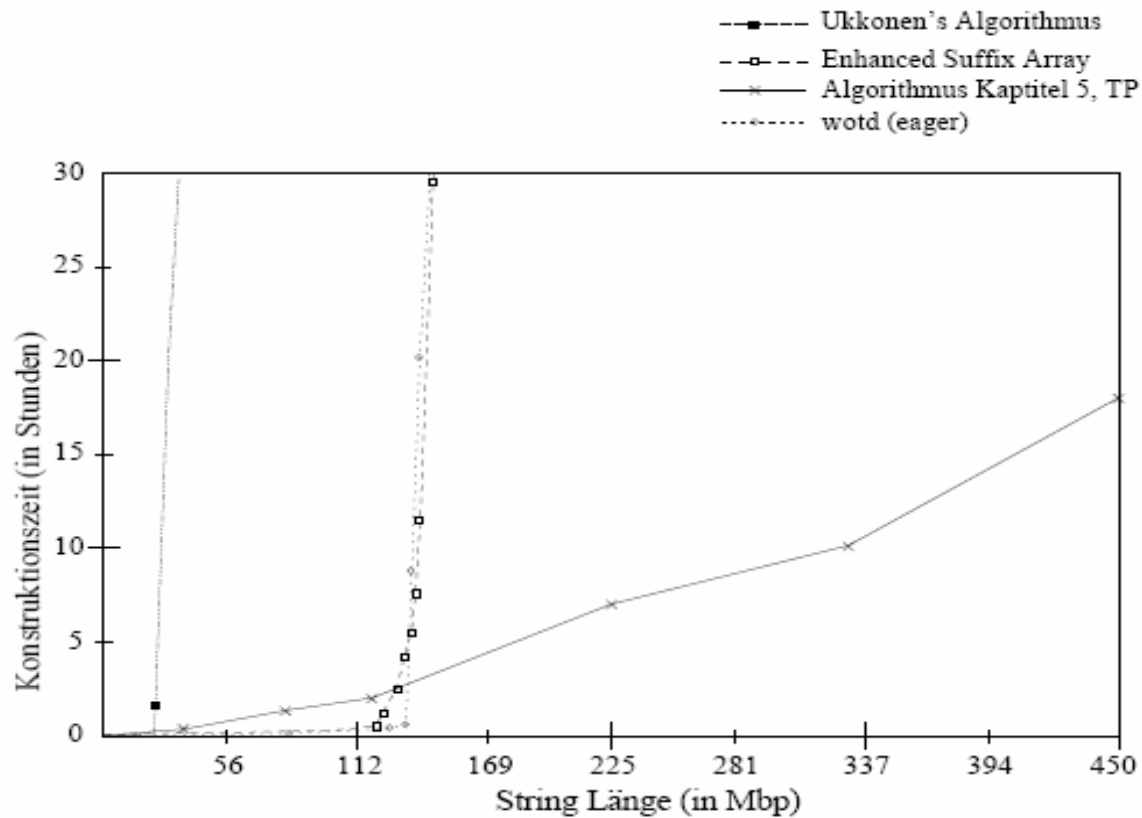


Abbildung 6.3: Entwicklung der Laufzeiten im Vergleich zu anderen Algorithmen

Quelle: G. Witterstein, Dissertation, 2005

Making the difference

- Databases are **not**
 - self-administrating
 - self-adapting
 - self-developing
 - self-healing
- Work mostly in the background – no GUI or dialogs
- Good database programs require **intensive knowledge** about the why & what & how
 - Which index to use and where to place it?
 - Which synchronization level?
 - When use parallel servers?
 - Optimization possibilities?
 - Where to look if something goes wrong?

DBS2: Implementation of Database Systems

- Lecture 4SWS
 - Tuesday, 13 – 15 , RUD26, 1'305
 - Thursday, 13 – 15 , RUD26, 1'305
 - Not : 5.5.2005, 26.5.2005

- Contact
 - Ulf Leser
 - Room: IV.103
 - Tel: (030) 2093 – 3902
 - Mail: leser (at) informatik.hu-berlin.de

Praktikum & Exams

- “Praktikum”
 - Presence & commitment are necessary
 - Implementation of file-/ buffer-/ index manager in C++
 - Two groups
 - Tuesday, 15-17, RUD 26, Room 1'308
 - Thursday, 15-17, RUD 26, Room 1'308
 - Starts next week, introduction this Thursday during lecture
 - Team:
 - Heiko Müller, Frank Huber, Jana Bauckmann, Chokri ben Necib
- Examination
 - 30 minutes, oral exam
 - **Alternative: Klausur**
 - Dates will be set mid-June
 - Requires Praktikumsschein
 - No changing between Prof. Freytag / me

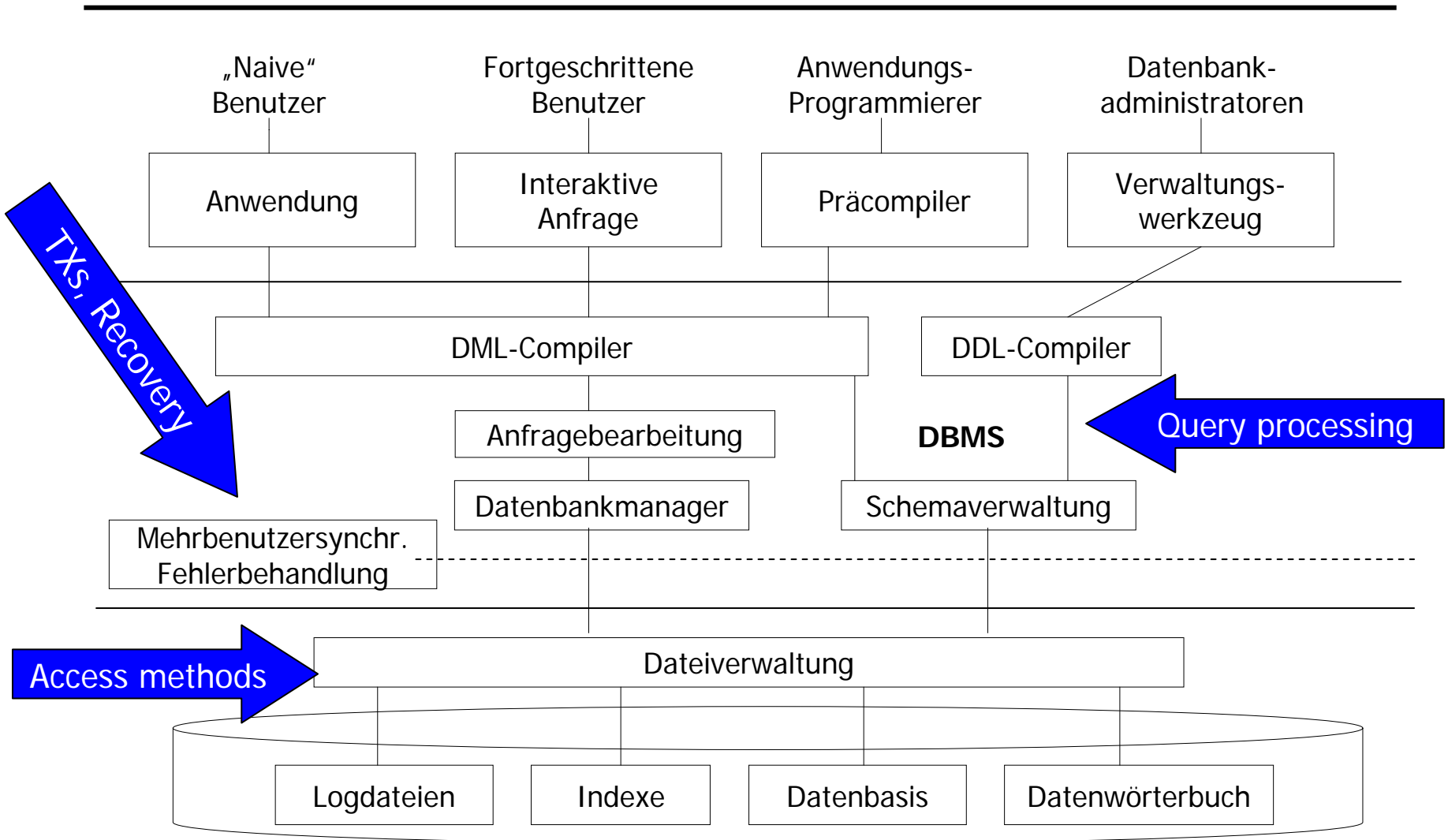
Slides

- Slides are available before / shortly after the lecture
- Please send me **any errors**
- I apologize: Mixture between English / Deutsch
- Slides are
 - No script
 - No substitution for listening to the lecture
 - No substitution for reading a book

Contents

1. Introduction
2. Basics of RDBMS
3. Overview and Architecture of DBMS
4. Access Methods
5. Query Processing and Optimization
6. Transactions and Concurrency Control
7. Recovery Concepts and Methods
8. Distributed DBMS
9. (Data Warehouse and Data Mining)

Überblick



Events

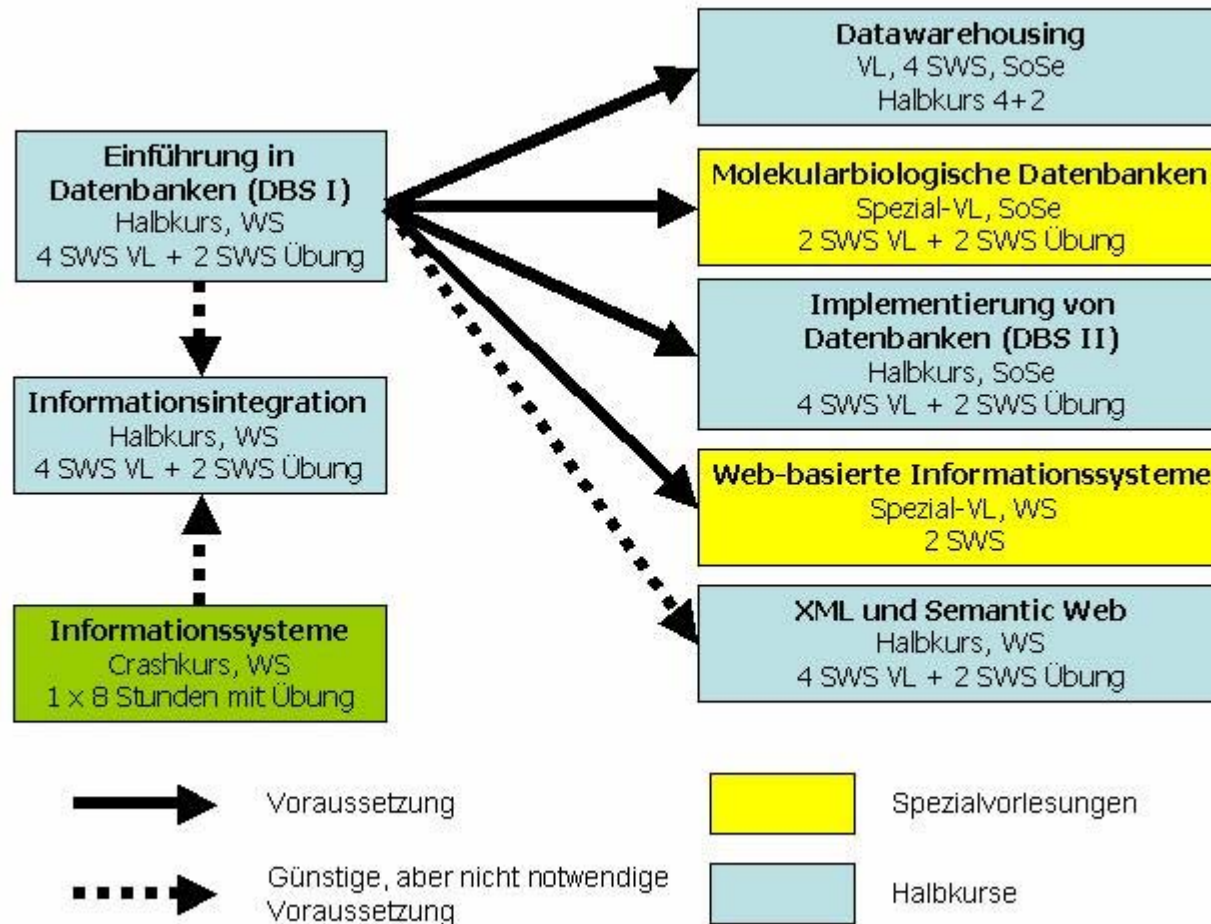
- Microsoft, SQL Server 2005
- Hopefully
 - Microstrategy
 - Oracle
 - Teradata

- Coordinated with lecture on Data Warehousing

Literature

- Primary
 - Saake, Heuer, Sattler **“Datenbanken: Implementierungstechniken”**, mitp Verlag, 2005 (2. Auflage)
 - Garcia-Molina, Ullman, Widom: **“Database System Implementation”**, Prentice Hall, 2000
 - Härder, Rahm: **“Datenbanksysteme. Konzepte und Techniken der Implementierung”**, Springer, 2. Auflage 2001
- Other
 - Kemper, Eickel: “Datenbanksysteme – Eine Einführung”, Oldenburg, 5. Auflage 2004
 - R. Elmasri und S.B. Navathe: Fundamentals of Database Systems, Benjamin Cummings
 - Deutsche Übersetzung: „Grundlagen von Datenbanksystemen“, Pearson, 2002
 - J. Gray, A. Reuter: Transaction Processing, Morgan Kaufman, 1993
 - T. Özsu und P. Valduriez: Principles of Distributed Database Systems, Prentice Hall, 1999

Datenbanken@Informatik



Wissensmanagement in der Bioinformatik

- **Schwerpunkte**
 - Management molekularbiologischer Daten
 - Datenintegration
 - Datenzentrierte Algorithmen der Bioinformatik
 - Text Mining
- **Laufend möglich**
 - Studienarbeiten
 - Diplomarbeiten
 - ... oft in Kooperation mit Bioinformatikfirmen
 - ... oft in Kooperation mit Bioinformatikinstituten

Stelle frei

- Studentische Hilfskraft
- 40 Stunden / Monat
- Thema
 - Datenintegration
 - Life Science Daten
 - Automatische Verfahren, große Datenmengen
 - Teamwork
 - JAVA, Oracle oder DB2, Datenbankkenntnisse
- Start: 1.5.2005, Laufzeit 12 Monate
- Bewerbung: Prof. Naumann

Fragen ?

Fragen meinerseits

- Diplominformatiker?
- Semester?
- Prüfung?
- DBS-I

- Warum DBS-II?