

Einführung

Dr. Angelika Reiser
Lehrstuhl für Informatik III:
Datenbanksysteme
www-db.in.tum.de

TU München / Garching

reiser@in.tum.de

Vorlesung

- Vorlesungswebseite siehe TUMonline-Eintrag

www-db.in.tum.de/teaching/ws1516/DBSandere

- IN4714:
 - Teil des Pflichtmoduls Geodatenbanken und Visualisierung
 - Umfang: 2V SWS und 2 ECTS

Modalitäten Vorlesung

- 2 Stunden Vorlesung wöchentlich
- Mittwochs, 16.30 – 18.00 Uhr
- Klausur: Termin steht noch nicht fest
- IN4714: 40 Minuten
- Bestehensgrenze: 50%
- Fragen gerne während der Vorlesung
- nach der Vorlesung individuelle Fragen

Modalitäten Vorlesung (cont.)

- Lesestoff zur Vorbereitung
- Eingebettete Übungen
- Vorrechnen an der Tafel

→ interaktive Veranstaltung

Vorlesungsinhalt

- Datenbankentwurf
 - E/R-Modellierung
 - UML-Modellierung
- Relationales Datenmodell
- Relationale Anfragesprache SQL
- Datenintegrität

Vorlesungsinhalt, cont.

- Physische Organisation der Daten
 - B-Bäume
 - Hashing
- Anfragebearbeitung
- Transaktionsverwaltung
- (Hauptspeicherdatenbanken, NoSQL Datenbanken, Data Warehouses, ...)

→ Vorbereitung für Geodatenbanken,
Andreas Donaubaue, ab Dezember 2015

Verwendete Materialien

Folien von Prof. Kemper:

www-db.in.tum.de/teaching/bookDBMSeinf

und Prof. Neumann:

www-db.in.tum.de/teaching/ws1415/grundlagen

Danke 😊 - Fehler gehen auf mich ☹️

Literatur

Alfons Kemper und André Eickler
Datenbanksysteme: Eine Einführung
10. Auflage (2015)
(ältere Auflagen sind auch ok)
Oldenbourg Verlag, München
(ca 50 Euro)

www-db.in.tum.de/teaching/bookDBMSeinf

Übungsbuch dazu ...

Alfons Kemper und Martin Wimmer
Übungsbuch Datenbanksysteme
3. Auflage (2011)
(ältere Auflagen sind auch ok)
Oldenbourg Verlag, München
(ca 30 Euro)

Zusatzmaterial

www-db.in.tum.de/teaching/bookDBMSeinf

- Folien
- Videos von Vorlesungen
- Daten zum Aufbau von eigenen Datenbanken
- SQL-Schnittstelle
- Programmbeispiele für
 - IBM DB2
 - Oracle
 - MS SQL Server

Weitere Literatur

A. Silberschatz, H. F. Korth und S. Sudarshan
Database System Concepts, 6. Auflage,
McGraw-Hill, 2010.

codex.cs.yale.edu/avi/db-book/db6/slide-dir/

R. Elmasri, S.B. Navathe

Fundamentals of Database Systems, 6. Auflage,
Addison-Wesley, 2010. (auch in Deutsch
erhältlich)

R. Ramakrishnan, J. Gehrke

Database Management Systems, 3. Auflage,
2003.

<http://pages.cs.wisc.edu/~dbbook/>

Weitere Literatur, cont.

J.D. Ullmann, J. Widom

A First Course in Database Systems, Prentice Hall, 3. Auflage, 2007.

infolab.stanford.edu/~ullman/fcdb.html

MOOCS

- Self paced mini courses, Stanford
class.stanford.edu/courses/DB/2014/SelfPaced/about
- Datenmanagement mit SQL, HPI
open.hpi.de/courses/sql

Weitere Literatur, cont.

- Informationssysteme/
Einführung in Datenbanksysteme, Uni Saarland
infosys.uni-saarland.de/datenbankenlernen/

Begriffsklärung

- Was ist ein Datenbanksystem (DBS)?

Ein System zum Speichern und Verwalten von Daten.

- Warum kein herkömmliches Dateisystem verwenden?

Ausfallsicherheit und Skalierbarkeit nur mit hohem Aufwand erreichbar.

Beispiele

Traditionelle Anwendungsgebiete:

- Geschäftsdaten
- Buchhaltung
- Verwaltung

...

Heute sehr viel breiter:

- Wissenschaftliche/Medizinische Daten
- Data Mining
- Geoinformationssysteme
- Websuche

...

Beispiele, cont.

Indirekt benutzen wir ständig Datenbanken:

- Websuche bei Google, Yahoo, ...
- Anfragen bei Amazon, EBay, ...
- Backend vieler großer Webseiten

Viele Spielarten (DB/IR, zentralisiert, dezentralisiert etc.)

Datenbanken werden fast immer eingesetzt wenn

- die Datenmengen groß sind
- die Daten wertvoll sind

Beispiele, cont.

Die großen kommerziellen Datenbanksysteme:

- Oracle
- IBM DB2
- Microsoft SQL Server
- Sybase

Einige freie Datenbanksysteme:

- PostgreSQL
- MySQL
- MonetDB

Noch viele weitere, teils stark spezialisierte Systeme

Motivation für den Einsatz eines Datenbanksystems

Typische Probleme bei Informationsverarbeitung ohne DBMS

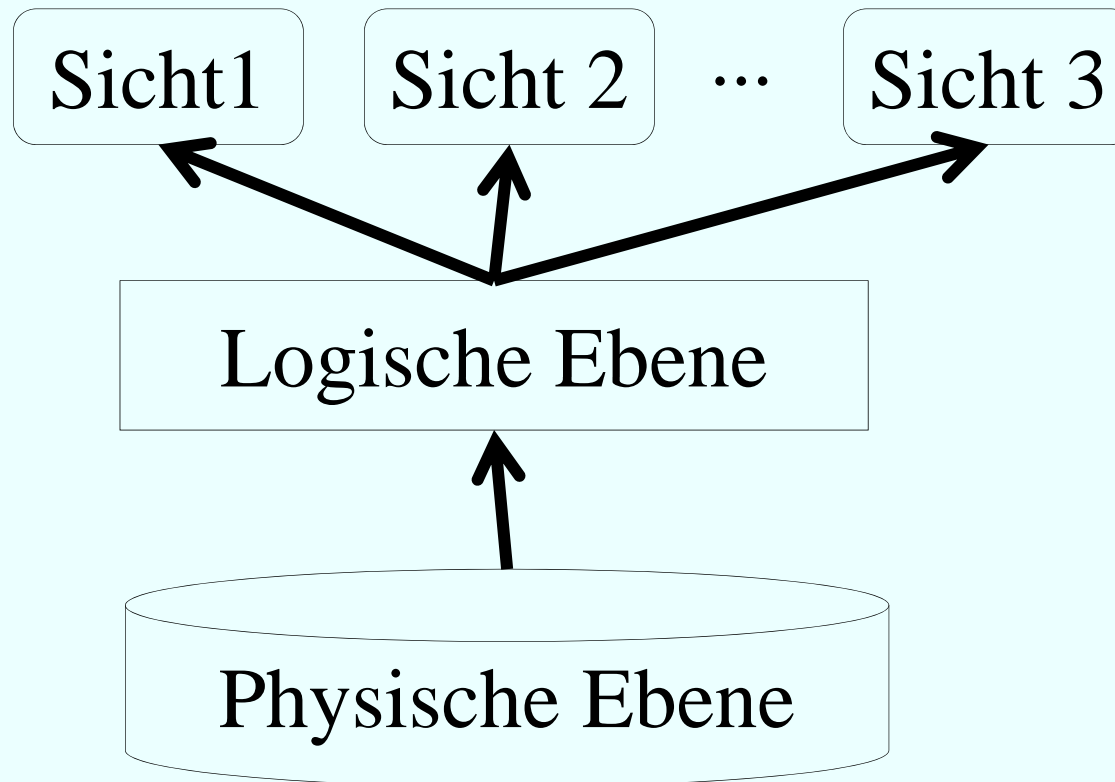
- Redundanz und Inkonsistenz
- Beschränkte Zugriffsmöglichkeiten
- Probleme beim Mehrbenutzerbetrieb
- Verlust von Daten
- Integritätsverletzung
- Sicherheitsprobleme
- hohe Entwicklungskosten für Anwendungsprogramme

Motivation für den Einsatz eines Datenbanksystems, cont.

Gründe für den DBS-Einsatz

- Deklarative Anfragesprachen
- Mehrbenutzersynchronisation
- Fehlerbehandlung
- Sicherstellung der Datenintegrität
- Effizienz und Skalierbarkeit

Abstraktionsebenen eines Datenbanksystems



Abstraktionsebenen eines Datenbanksystems, cont.

Sicht:

beschreibt wie ein Benutzer die Daten sieht

Logische Ebene:

beschreibt wie die Daten strukturiert sind

Physische Ebene:

beschreibt wie die Daten gespeichert werden

Abstraktionsebenen eines Datenbanksystems, cont.

DBS entkoppelt Anwendungen von der Struktur und Speicherung der Daten:

- Logische Datenunabhängigkeit
(einfache) Änderungen auf der logischen Ebene haben keinen Einfluß auf Anwendungen
- Physische Datenunabhängigkeit
Änderungen auf der physischen Ebene haben keinen Einfluß auf Anwendungen

Wird in fast allen modernen DBS durchgesetzt

Eigenschaften von Datenbanksystemen

Deklarative Anfragesprache

- Benutzer sagt DBS **was** für Daten geholt werden sollen . . .
. . . und **nicht wie** die Daten geholt werden sollen
- Weniger fehleranfällig (beim Formulieren von Anfragen/Entwickeln von Anwendungen), da kein Wissen über die tieferen Schichten des DBS nötig

Eigenschaften von Datenbanksystemen, cont.

Mehrbenutzersynchronisation

- Wenn mehrere Benutzer ohne jegliche Kontrolle gleichzeitig Daten ändern können, gibt es große Probleme
- DBS erlaubt gleichzeitigen Zugriff und verhindert schlimme Seiteneffekte

Eigenschaften von Datenbanksystemen, cont.

Fehlerbehandlung

- DBS kann Zustand zum Zeitpunkt eines Absturzes rekonstruieren
- Dafür werden Logdateien vom DBS angelegt und verwaltet

Eigenschaften von Datenbanksystemen, cont.

Datenintegrität (Konsistenz)

- Datenverarbeitung in einer Anwendung läuft nicht völlig zufällig ab, sondern folgt gewissen Regeln

→ DBS befolgt (angegebene) Regeln und schützt so (automatisch) vor:

- Benutzerfehlern
- Programmfehlern

Eigenschaften von Datenbanksystemen, cont.

Effizienz und Skalierbarkeit

- DBSe sind für groß angelegte Anwendungen konzipiert
- In DBSen sind Techniken integriert, die mit großen Datenvolumen umgehen können

typisch: 100 GB (Gigabyte) – transaktionale Daten (auch Expressversionen)
bis zu EB (Exabyte) maximale Größe

Eigenschaften von Datenbanksystemen

- Deklarative Anfragesprache
- Mehrbenutzersynchronisation
- Fehlerbehandlung
- Datenintegrität
- Effizienz und Skalierbarkeit

Architekturübersicht eines DBMS

