1 Generelles

Erläutern Sie die Motivation hinter der Datenbanktheorie. Gehen Sie insbesondere auf die Probleme ein, und beschreiben Sie was es heißt wenn ein Datum integer ist.

fhwedel
UNIVERSITY OF APPLIED SCIENCES

Motivation

- Daten sollen dauerhaft gespeichert sein und jederzeit schnell verfügbar sein
- Integrität der Daten muss zu jedem Zeitpunkt gewährleistet sein (Richtigkeit d. Daten)
- Mehrbenutzerbetrieb mit stark unterschiedlichen Nutzungsszenarien (Endanwender, Programmierer, Administrator)
- Probleme hierbei:
  - Konflikte um Ressourcen -> Anfrageoptimierung notwendig
  - Inkonsistente Zustände -> Deadlocks
  - Transaktionsverwaltung -> 2 Phasen-Sperrprotokoll
  - Datenschutz/Autorisierung -> Grant/Revoke

Um diese Aufgaben zu erleichtern und einzelne Bereiche eines DBMS unabhängig vom anderen zu machen, entstanden mehrstufige Architekturmodelle

Datenbanktheorie und Implementierung - FH-Wedel SS 2019 - Michael Predeschly 3

Erläutern Sie das ANSI-SPARC Schema. Welches Ziel wird hierbei verfolgt.

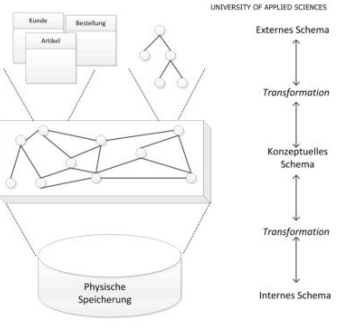
fhwedel
UNIVERSITY OF APPLIED SCIENCES

Architektur von Datenbanken

ANSI/SPARC 3 Schema-Architektur (1975)

Unterscheidung von drei Ebenen

- **Externe Schema** (Benutzersicht)
- **Konzeptuelles Schema** (Referenzschema)
- **Internes Schema** (physische Realisierung)



Ziel:
Trennung der einzelnen Schichten, um Modularität und Austauschbarkeit zu gewährleisten (einzelne Schichten müssen nur Schnittstellen kennen).

Datenbanktheorie und Implementierung - FH-Wedel SS 2019 - Michael Predeschly 4

Wie werden Daten auf einer Magnetplatte adressiert? Nennen Sie sowohl den Überbegriff als auch die einzelnen Komponenten aus denen sich dieser zusammensetzt.

Wägen Sie ab warum es sinnvoll sein kann bei der Speicherung von Daten auf einer Magnetplatte auf größere Datenblöcke zu setzen.

Bei zu kleinen Datenblöcken müssen die zu speichernden Daten öfter aufgeteilt werden. Da diese im Zweifelsfall über die gesamte Platte verteilt sein können wird viel Zeit für diverse Schreib- und Lesevorgänge benötigt.

Bei zu großen Datenblöcken kann es vorkommen, dass eine kleines Datum (z.B. eine Zahl) nur einen geringen Teil der verfügbaren Blockgröße verwendet.

Erläutern Sie grob wie ein Raid-System funktioniert. Was versteht man unter dem sog. Striping sowie dem Mirroring ?

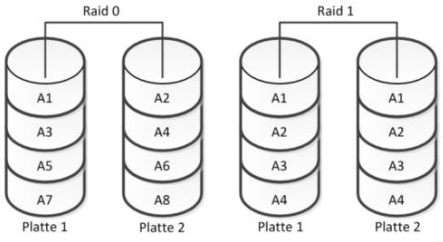
Physische Speicherung - Raid

Festplatten können zu sogenannten RAIDs (**R**edundant **A**rray of **I**ndependent **D**isks) zusammengeschalten werden, um die Ausfallsicherheit oder die Zugriffsgeschwindigkeit zu erhöhen.

Wie funktioniert ein Raid?

RAID 0 - Striping (max. Performance)
↳ Daten werden auf zwei Platten aufgeteilt

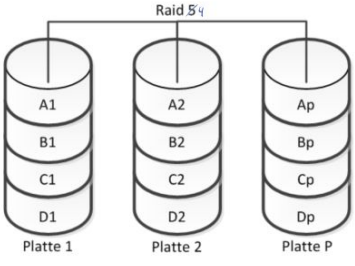
RAID 1 - Mirroring
↳ Daten werden kopiert & auf zwei Platten gespeichert



Erläutern Sie was man unter einem Raid-4 System versteht. Wie wird hierbei die Parität ermittelt?

RAID 5 - Parität auf mehreren Platten

- Vorteile von RAID 0 und RAID 1 kombiniert
- Parity Information auf Platte P erlaubt bei Ausfall einer Platte die Wiederherstellung der Daten der ausgefallenen Platte.
- RAID 5 ersetzt kein Backup!!
- Werden Daten gelöscht, können diese auch mit der Parity-Information nicht wiederhergestellt werden



Beispiel Paritätsb

Platte 1: 0101 11

Platte 2: 1110 00

Paritätsplatte: 1011 11

Wie vorgehen bei mehr als

¡img src=paste-18150531792897.jpg

a) Berechnen Sie die Parity Information der Platte P.
b) Nach erfolgreicher Parity-Berechnung fällt Platte 2 aus. Rekonstruieren Sie diese auf einer Platte 2'. Wichtig ist hierbei der Rechenweg.

¡/div¿¡missing WORD¿

fhwedel
UNIVERSITY OF APPLIED SCIENCES

Übungsaufgabe

Gegeben sind folgende 4 Platten:

Platte 1: 1111 0111
Platte 2: 1011 1100
Platte 3: 0111 0011

Pl.1	1 1 1 1	0 1 1 1
Pl.3	0 1 1 1	0 0 1 1
Par.	0 0 1 1	1 0 0 0
Pl.2	1 0 1 1	1 1 0 0

Platte P: 0011 1000

a) Berechnen Sie die Parity Information der Platte P.
b) Nach erfolgreicher Parity-Berechnung fällt Platte 2 aus. Rekonstruieren Sie diese auf einer Platte 2'. Wichtig ist hierbei der Rechenweg.

Datenbanktheorie und Implementierung - FH-Wedel SS 2019 - Michael Predeschly 9

Warum erhöht ein Raid 4 die Ausfallsicherheit, ersetzt aber kein Backup?

Manuelles Löschen führt weiterhin zum Datenverlust.

Weswegen wird in der Datenbanktheorie eine möglichst effiziente Pufferverwaltung angestrebt? Beschreiben Sie grob wie so eine aussehen könnte.

fhwedel
UNIVERSITY OF APPLIED SCIENCES

Motivation

Wenn Daten von Festplatte gelesen werden ist dies sehr viel langsamer als wenn diese im Arbeitsspeicher liegen. Die Zugriffszeit gibt dabei an, wie schnell Daten bereitgestellt werden können. Typische Zugriffszeiten heutiger Speichermedien:

- Festplatten: 9 ms
- SSD: 250 µs
- Arbeitsspeicher 60-70 ns
  (Quelle: https://de.wikipedia.org/wiki/Zugriffszeit)

-> Daten sollten komplett im Arbeitsspeicher gehalten werden. Dies ist aufgrund hoher Kosten meist nicht komplett möglich.

-> Intelligente Auslagerung der am häufigsten/dringendsten benötigten Daten in den Arbeitsspeicher

Datenbanktheorie und Implementierung - FH-Wedel SS 2019 - Michael Predeschly 12

Was versteht man unter dem Mapping von Speicheradressen?

Zuordnung von Speicheradressen auf Scheiben, Spuren und Sektoren.

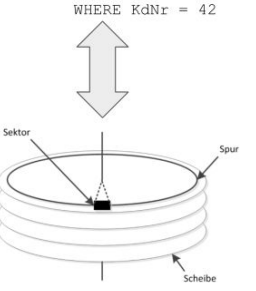
fhwedel
UNIVERSITY OF APPLIED SCIENCES

Speicherverwaltung

- Verbindung der Daten mit dem Festplatten-Speicher (Mapping von Speicheradressen auf Scheiben, Spuren und Sektoren)
- Abstraktion von konkreter Hardware auf der Speicherverwaltung realisiert
  -> Technologieunabhängigkeit
- Effizientes Auslesen der angeforderten Daten in den Arbeitsspeicher
- Schnelles Auffinden freier Datenblöcke

```
SELECT *
FROM Kunde

UPDATE Kunde
SET Name = Bilbo
WHERE KdNr = 42
```

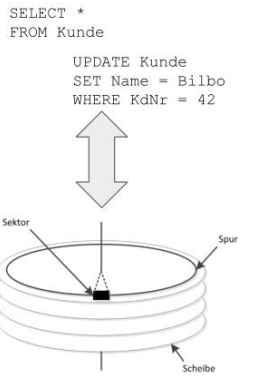


Datenbanktheorie und Implementierung - FH-Wedel SS 2019 - Michael Predeschly 13

Beschreiben Sie generell den Begriff der Speicherverwaltung.

Speicherverwaltung

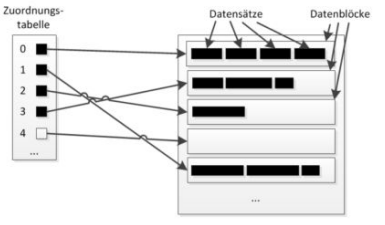
- Verbindung der Daten mit dem Festplatten-Speicher (Mapping von Speicheradressen auf Scheiben, Spuren und Sektoren)
- Abstraktion von konkreter Hardware auf der Speicherverwaltung realisiert
  -> Technologieunabhängigkeit
- Effizientes Auslesen der angeforderten Daten in den Arbeitsspeicher
- Schnelles Auffinden freier Datenblöcke



Beschreiben Sie grob den Begriff der Speicherzuordnung. Mit welchem Hilfsmittel ist es hierbei möglich die Datenblöcke zu lokalisieren? Müssen die Datenblöcke stets die gleich Länge aufweisen?

Speicherzuordnung

- **Zuordnungstabelle** als Inhaltsverzeichnis für Datenblöcke
- Datenblöcke enthalten mehrere **Datensätze** unterschiedlicher Länge
- Eine Datei kann mehrere Einträge in der Zuordnungstabelle beanspruchen
- Freie Datenblöcke können leicht über Zuordnungstabelle gefunden werden



Was sind die Ziele der Pufferverwaltung ?

Pufferverwaltung

Ziel Pufferverwaltung:

- Reduktion von physischer Ein-/Ausgabe
- Häufig genutzte Daten werden in **Seiten** im Hauptspeicher vorgehalten
- **Verdrängung** alter Daten im Hauptspeicher durch intelligente Verfahren

-> Weniger Zugriffe auf die Festplatte und damit schnellere Verarbeitung der Anfragen möglich

Pufferverwaltung in Datenbanken hat viele Ähnlichkeiten zum Vorgehen von Cache-mechanismen bei Betriebssystemen.

Wo werden die Pufferdaten zwischengelagert?

In sog. Seiten im Hauptspeicher (RAM).

Pufferverwaltung

Ziel Pufferverwaltung:

- Reduktion von physischer Ein-/Ausgabe
- Häufig genutzte Daten werden in **Seiten** im Hauptspeicher vorgehalten
- **Verdrängung** alter Daten im Hauptspeicher durch intelligente Verfahren

-> Weniger Zugriffe auf die Festplatte und damit schnellere Verarbeitung der Anfragen möglich

Pufferverwaltung in Datenbanken hat viele Ähnlichkeiten zum Vorgehen von Cache-mechanismen bei Betriebssystemen.

Datenbanktheorie und Implementierung - FH-Wedel SS 2019 - Michael Predeschly 15

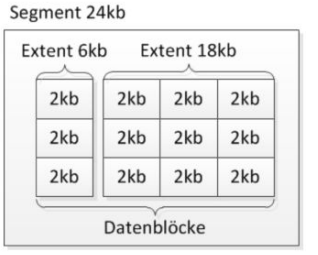
front

Beispiel Oracle

- **Segmente**: logische Einheit
- **Extent**: kleinste logische Speichereinheit in Oracle

Segmente besteht aus einem oder mehreren Extents, die die Informationen zu einem Objekt gespeichert haben (z.B. Relation).

Reicht Speicherplatz für Segment nicht aus, werden weitere Speicherbereiche im Rahmen des Tablespaces allokiert.



Datenbanktheorie und Implementierung - FH-Wedel SS 2019 - Michael Predeschly 17

Da eine Datenbank i.d.R. ein eigenes Filesystem anlegt ist es komplett losgelöst von sämtlichen Restriktionen von Festspeicher Betriebssystem.