

Prüfungsschwerpunkte „Datenbanken Implementierungstechniken“

Die Prüfung umfasst einen Fragenteil (ca. 45 Min.) sowie einen Aufgabenteil (ca. 75 Min). Im Aufgabenteil dürfen Unterlagen verwendet werden. Neben diesen allgemeinen Fragen zum Vorlesungsstoff werden Fragen zum Themenschwerpunkt des jeweiligen Referats gestellt werden.

1. Speicherverwaltung für Datenbanken

Speichermedien (Platte, SSD, Hauptspeicher): Vor- und Nachteile
Aufbau von Speichermedien: Hard Disk, Struktur eines Flash-Speichers
Ablauf eines Plattenzugriffs
Was ist RAID (allg. Prinzip)?
Puffermanagement, Ablauf eines I/O-Request
Prinzipieller Aufbau einer Datei

2. Dateiorganisation

3 Organisationsformen: Heap, Sortiert, Hash (Prinzip, keine Formeln!)
Datensätze mit fester und variabler Länge
5 Basisoperationen (in Anwendungsfällen)
Index: Prinzip und Klassifikation
- Primärindex / Sekundärindex
- geclustert / ungeclustert
- dicht / dünn
- einfach / zusammengesetzt

3. Bäume

B+ Bäume: Kenngrößen, Vor- und Nachteile
Rechnungen (vgl. Übungsblatt), z.B. mit Kapazitäten

4. Hashing

Statisches und Dynamisches Hashing: Vor- und Nachteile
Erweiterbares und lineares Hashing

5. Anfrageverarbeitung

Prinzipieller Ablauf der Verarbeitung einer Anfrage
Berechnungen von Joins (unterschiedl. Join-Implementierungen)
Berechnung von einfacher Selektion, Projektion, Scans

6. Anfrageoptimierung

grundsätzliches Problem (Optimierungsziel vs. Optimierungstechniken)
Kostenarten
Prinzip kostenbasierter Optimierung (vs. regelbasierte)
Berechnungen alternativer Pläne
Problem bei geschachtelten Anfragen
Bedeutung von Reduktionsfaktoren (Prinzip, keine Formeln!)
Anwendung algebraischer Äquivalenzen bei der logischen Optimierung

7. Physischer Datenbankentwurf

zu treffende Entscheidungen
Anwendung von Regeln bei Indexauswahl
Entscheidungen über Indexe bei konkreten Anfragen
Index-Only-Pläne: Prinzip und Anwendung

8. Tuning

Wahl der geeigneten Normalform unter Berücksichtigung von relevanten Anfragen und interrelationalen Constraints
Verbessern einzelner Anfragen (vgl. Tipps zum Tuning von Queries)

9. Transaktionen im Mehrbenutzerbetrieb

Definition und Eigenschaften von Transaktionen

Nebenläufigkeit: wozu, welche Probleme?

Erkennen von Anomalien im Mehrbenutzerbetrieb

2-Phasen-Sperrverfahren

Serialisierbarkeit

Deadlocks

Anwendung von Indexsperrern und des hierarchischen Sperrprotokolls

Andere Verfahren: Optimistisches, Zeitstempelverfahren, Mehrversionen-CC
(Bewertung, Vergleich)

10. Recovery

Fehlerarten

Pufferverwaltung Policies

Logging: Prinzip, welche Informationen

Write-Ahead-Logging

Grundsätzlicher Ablauf bei Recovery gemäß ARIES-Algorithmus: Analyse, Undo, Redo

Betrachtung eines Beispiel-Szenarios

Bedeutung von Checkpoints

11. Verteilte Datenbanken

Motivation

Fragmentierung: vertikal vs. horizontal

Begriff der Transparenz

Replikation: synchron vs. asynchron

Verteilte vs. zentralisierte DB- Komponenten: Query-Optimierung, Sperren-
verwaltung

Quantitative Bewertung verteilter DB-Anfragen (Joinkosten vs. Transportkosten)

2-Phasen-Commit-Protokoll

Abschwächung der Konsistenz: ACID vs. BASE, CAP-Theorem

12. Fragen zu den Seminarvorträgen

Dabei werden zumeist Fragen mit Bezug zum Stoff der Vorlesung gestellt, z.B.
Einordnung von Grundbegriffen oder Verwendbarkeit bestimmter Techniken