

1. Übungsblatt

Ausgabe: 19. Oktober 2015 · Besprechung: 26. Oktober 2015

1 Storage Disks

In der Vorlesung wurden Hard-Disk Drives und Solid-State Drives behandelt.

1. Erklären Sie kurz die Funktionsweise beider Speichermedien!
2. Was sind jeweils die Vor- und Nachteile?
3. Welche Konsequenzen ergeben sich daraus für ein Datenbankmanagementsystem?
4. Gegeben sei eine Festplatte mit 512 Byte großen Blöcken mit folgenden Geschwindigkeiten: Lesen mit 50 MB/s; Schreiben mit 40 MB/s; Zugriffszeit (Seek): 10 ms; Rotationsgeschwindigkeit: 7200 rpm

Gegeben sei zudem ein NAND-Flash-Medium, das in 128 KB große Blöcke mit je 64 Seiten unterteilt ist, und folgende Geschwindigkeiten aufweist: 25 μ s um einen Lesevorgang zu beginnen; 100 μ s um eine Seite zu lesen; 2 ms zum Löschen eines Blocks; 200 μ s zum Schreiben einer Seite.

Wie lange dauern folgende Operationen:

- (a) Sequentielles Lesen von 1 GB
- (b) Lesen von 1 GB, bei dem alle Daten in 4 KB großen Datenblöcken zufällig verteilt sind
- (c) Schreiben von 10 MB (sequentiell)
- (d) Schreiben von 1 Byte

2 PostgreSQL

Wir werden uns später noch das Verhalten von PostgreSQL anschauen, dessen Implementation oft ein gutes Beispiel ist für die in der Vorlesung diskutierten Techniken.

1. Laden Sie den Sourcecode der aktuellen PostgreSQL-Version auf Ihren Rechner. Anschließend kompilieren und installieren Sie ihn.
2. In eine leere PostgreSQL-Datenbank laden Sie Daten des TPC-H-Benchmarks.
TPC-H ist der wohl bekannteste Benchmark für "decision support"-Aufgaben. Auf der Webseite des TPC-H-Konsortiums finden Sie ein Paket, das Generatoren für jeweils Daten (**dbgen**) und Anfragen (**qgen**) enthält.
3. Führen Sie die TPC-H-Anfragen aus und messen Sie jeweils die Laufzeit.