

1. Übungsblatt

Ausgabe: 17. Oktober 2016 · Besprechung: Ab 24. Oktober 2016

1 Storage Experiment

- Entwickle ein Programm, das die Lesegeschwindigkeit einer Festplatte für Random Access und Sequential Access misst. Bestimme Näherungen für die Parameter Seek Time t_s , Rotational Delay t_r und die sequentielle Lesebandbreite B_r anhand der gemessenen Daten und der Gerätespezifikation. Führe dazu die folgenden Experimente durch:

Ex1. Wie lange dauert es 100 MB Daten sequentiell zu lesen?

Ex2. Wie lange dauert es 1000 Segmente der Größe 1 Byte von zufälligen Lesepositionen aus einer großen Datei (2-4 mal so groß wie der Arbeitsspeicher) zu lesen?

- Vergleiche die Performance von SSD Hardware mit den Messungen zur HDD.

Lösung

Experiment 1

Um die das sequentielle lesen zu testen wird ein Ordner auf der Ziel-Festplatte als Working Directory gesetzt und die folgende Befehlsfolge ausgeführt:

```
dd if=/dev/zero of=tempfile bs=1M count=1024 conv=fdatasync,notrunc
echo 3 | sudo tee /proc/sys/vm/drop_caches
dd of=/dev/zero if=tempfile bs=1M count=1024 conv=fdatasync,notrunc
```

- Messung der Sequentiellen Lesegeschwindigkeit mit `dd if=.. of=.. bs=1M count=100 conv=notrunc,fdatasync`
- Zum generieren von Nutzlast und als Senke für Daten kann `/dev/zero/` genutzt werden
- Leeren des Filesystem Caches mit `echo 3 | sudo tee /proc/sys/vm/drop_caches`

$B_r = 75 \text{ MB/s}$

Experiment 2

Um die Random Access Lesegeschwindigkeit zu bestimmen, führen wir `fseek(..)` aus, um an zufällige Stellen in einem Großen File zu navigieren. Es ergibt sich eine Lesegeschwindigkeit von 11.9 ms pro read. Jetzt lassen sich die anderen Parameter bestimmen.

Berechnung der Parameter

Der Rotational Delay t_r ist im Durchschnitt die Dauer einer halben Rotation:

$$t_r : 5400 \text{ rpm} = 90 \text{ rps} \rightarrow 11 \text{ ms pro rotation}$$
$$t_r = 5.5 \text{ ms pro Zugriff}$$

Der Wert von t_s lässt sich nun als Rest der Random Access Zugriffszeit bestimmen:

$$t_s = 11.9 \text{ ms} - t_r = 6.4 \text{ ms}$$

HDD vs. SSD

Gerät	Lesebandbreite	Zugriffslatenz
HDD	75 MB/s	11.9 ms
SSD	300 MB/s	0.1 ms

2 Anfrageverarbeitung

- Betrachte zwei Anfragepläne p_1 und p_2 für eine Datenbankanfrage. Die Anfrage soll auf einem Datenbanksystem mit 4 KB Seiten ausgeführt werden.
- p_1 liest 100000 Pages sequentiell und liest 2500 Records mit je 512 Byte Länge von Positionen die gleichverteilt auf einer großen Anzahl von Seiten liegen.
- p_2 liest 250000 Pages sequentiell.
- Evaluiere mit Hilfe der gemessenen Hardwareeigenschaften t_s , t_r und B_r , wie lange es dauert die Anfrage mit p_1 und p_2 zu bearbeiten. Dabei genügt es die Zeit für die Festplattenzugriffe zu bestimmen. Wir gehen davon aus, dass keine Daten gepuffert werden. Diskutiere den Einfluss unterschiedlicher Zugriffsmuster und Storage-Hardware auf Anfrageverarbeitung.

Lösung

Ausführungszeit von p_1

$$t_{HDD} = \frac{(100000 + 2500) \cdot 4 \text{ KB}}{75 \text{ MB/s}} + 2500 \cdot (5.5 \text{ ms} + 6.4 \text{ ms}) = 35.2 \text{ s}$$

$$t_{SSD} = \frac{(100000 + 2500) \cdot 4 \text{ KB}}{300 \text{ MB/s}} + 2500 \cdot 0.1 \text{ ms} = 1.6 \text{ s}$$

Ausführungszeit von p_2

$$t_{HDD} = \frac{250000 \cdot 4 \text{ KB}}{75 \text{ MB/s}} = 13.3 \text{ s}$$

$$t_{SSD} = \frac{250000 \cdot 4 \text{ KB}}{300 \text{ MB/s}} = 3.3 \text{ s}$$

Diskussion

Wir beobachten, dass der optimale Plan abhängig von der zugrundeliegenden Hardware ist. Für SSD Hardware eignet sich Plan p_1 besser und für HDD Hardware Plan p_2 . Der Grund dafür ist, dass die Größe des Performanceunterschieds zwischen beiden Geräten von den eingesetzten Zugriffsmustern abhängt. Die SSD ermöglicht eine deutlichere Performance Steigerung für Random Access Zugriffe (Faktor **119x**) als für sequentielles Lesen (Faktor **4x**).

Hinweise:

- Linux stellt eine Auswahl von Befehlen bereit, die die Messung und Generierung von Daten unterstützen. Entscheide, ob sich Tools wie `dd` und `hdparm` eignen.
- Die Filestreams `/dev/null` und `/dev/rand` können bei der Datengenerierung helfen.
- Der C-Befehl `fseek(..)` eignet sich um Festplattenzugriffe auszuführen.
- Falls aus Aufgabe 1 keine aussagekräftigen Ergebnisse verfügbar sind, benutze die Werte aus Teil 2 der Vorlesung auf Folie 15.