
Aufgabe 1

Aufgabe 1 a)

Im folgenden Listing wird die Implementierung des in der Vorlesung vorgestellten Skippy-Benchmarks gezeigt. Erklären Sie die Funktionsweise des Benchmarks.

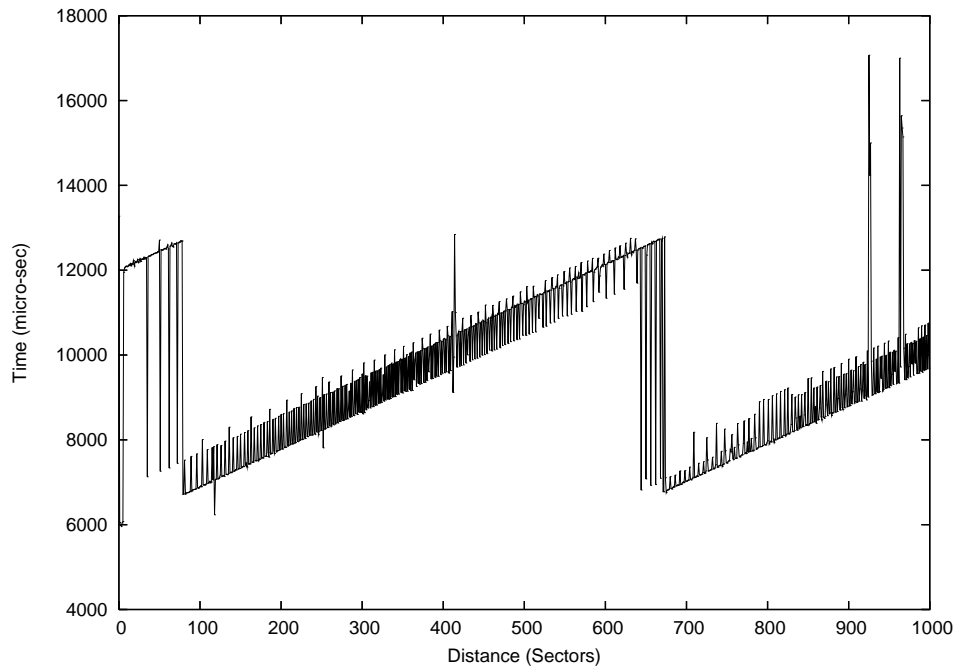
```
#include <stdlib.h>
#include <stdio.h>
#include <sys/types.h>
#include <sys/stat.h>
#include <sys/time.h>
#include <fcntl.h>
#include <unistd.h>

#define SINGLE_SECTOR 512

int
main() {
    int lNoMeasure = 1000;
    char lBuffer[SINGLE_SECTOR];
    int i;
    int lTime;

    struct timeval lTimeValStart;
    struct timeval lTimeValEnd;
    struct timezone lTimeZone;
    for(i = 0; i < SINGLE_SECTOR; ++i) {
        lBuffer[i] = 'x';
    }

    int lFd = open("/dev/sda3", O_RDWR | O_SYNC);
    if(lFd < 0) {
        fprintf(stderr, "Error occurred while opening raw device\n");
        exit(-1);
    }
}
```



```

for(i = 0; i < lNoMeasure; ++i) {
    /* time the following sequence, and output <i, time> */
    gettimeofday(&lTimeValStart, &lTimeZone);
    lseek(lFd, i * SINGLE_SECTOR, SEEK_CUR);
    write(lFd, lBuffer, SINGLE_SECTOR);
    gettimeofday(&lTimeValEnd, &lTimeZone);

    lTime = (lTimeValEnd.tv_sec - lTimeValStart.tv_sec) * 1000000;
    lTime += (lTimeValEnd.tv_usec - lTimeValStart.tv_usec);

    printf("%d %d\n", i, lTime);
}
close(lFd);
}

```

Aufgabe 1 b)

Interpretieren Sie die mit dem Benchmark ermittelten Beispielmessungen! Welche Parameter lassen sich ablesen?

Aufgabe 2

Bestimmen Sie den break-even point, bis wann sich der Einsatz eines Index gegenüber einem Scan lohnt.

Aufgabe 3

Angenommen der Optimierer hat die Kosten für einen scan um einen Faktor $a, a > 1$

unterschätzt und die Kosten für den Index um einen Faktor $b, b < 1$ überschätzt. Wie groß ist der maximale Fehler durch diese Fehleinschätzung?

Aufgabe 4

Aufgabe 4 a)

Untersuchen Sie den Operator `Indexscan` in `tinydb`. Wie lassen sich Prädikate auswerten? Welches Zugriffsverhalten hat der Operator?

Aufgabe 4 b)

Schreiben Sie ein Java-Programm, das z.B. analog zu `JoinSample.java` unter Verwendung des `Tablescan` bzw. `Indexscan` die folgende Anfrage auswertet.

```
SELECT *
FROM CUSTOMER, NATION
WHERE c_nationkey = n_nationkey
      AND n_name = "GERMANY"
      AND c_custkey = 300
```

Vergleichen Sie die Laufzeit der beiden Programme.

Aufgabe 4 c)

Erläutern Sie wie der `Indexscan` zusätzlich zu `Tablescans` in einem der join ordering Algorithmen, z.B. in `GOO`, als Planalternativen berücksichtigt werden können.