
Aufgabe 1 - Festplattenparameter.**1 Punkt**

Betrachten Sie die folgende Festplatte:

- Sektorgröße $B = 2^9$ Bytes,
- Interblock Gap $G = 0$ Bytes,
- Sektoren pro Spur $S = 2^9$,
- Spuren pro Scheibenseite $T = 300$,
- Anzahl der beidseitig beschriebenen Scheiben $D = 2^3$,
- mittlerer Spurwechsel $sp = \frac{1}{560}$ s,
- Drehzahl $dz = 15 \cdot 10^3$ rpm.

Berechnen Sie die **Zugriffszeit (in s) für 100 sequentielle Lesezugriffe.**

Lösung: _____

Aufgabe 2 - Datei-Organisation.**1 Punkt**

Die folgenden Tabelle soll in eine sequentielle Datei überführt werden. Als Schlüssel wird das Attribut **Character** verwendet. **Illustrieren Sie die sequentielle Datei.**

Actor	Character	Series
N. Reedus	Daryl	The Walking Dead
K. Harington	Jon	Game Of Thrones
B. Cranston	Walter	Breaking Bad
S. Yeun	Glenn	The Walking Dead
I. Elba	Stringer	The Wire
A. Royo	Bubbles	The Wire
M. Williams	Robb	Game Of Thrones
S. Wilson	Hershel	The Walking Dead

Aufgabe 3 - Slotted Page.**1 Punkt**

Gegeben sei eine Slotted Page mit folgenden Eigenschaften:

- Größe: $2^{13} = 8192$ Bytes,
- Adressierungstyp: **Word-Adressierung** (d.h. es kann nur jedes 2. Byte adressiert werden)

In dieser Slotted Page werden **3 Tupel** U , C , K gespeichert:

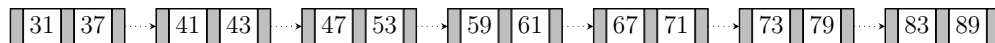
- d_1 : $|U| = 1.500$ Bytes
- d_2 : $|C| = 1.399$ Bytes
- d_3 : $|K| = 792$ Bytes

Ergänzen Sie die Slotted Page um die **fehlenden Werte/Adressen**, wobei p_i und g_i sich auf den jeweiligen Datensatz d_i beziehen. **(0.125 Punkte pro Wert/Adresse)**

a	f	g_1	p_1	g_2	p_2	g_3	p_3	\dots	d_1	d_2	d_3
									U	C	K

Aufgabe 4 - B^+ -Baum-Konstruktion.**1 Punkt**

Gegeben sind die Blätter eines B^+ -Baumes ($m = 3$). Konstruieren Sie die **darüberliegenden Ebenen** (d.h. die inneren Knoten) so, dass ein **gültiger B^+ -Baum mit minimaler Höhe** entsteht.

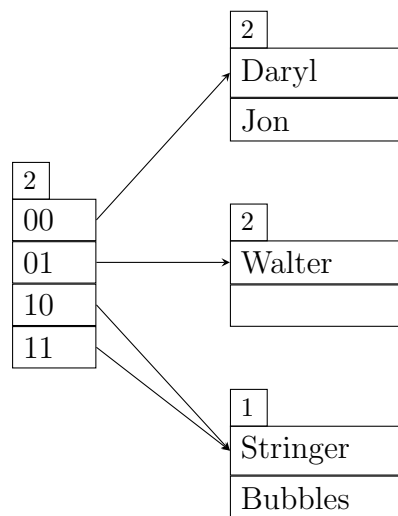


Aufgabe 5 - Erweiterbares Hashing.

1 Punkt

Die Hashfunktion $h(x)$ liefert die in der Tabelle angegebenen Binärwerte. Es soll das Tupel *Hershel* in den gegebenen Hashcontainer eingefügt werden. Ein Bucket im Hashcontainer kann bis zu 2 Tupel speichern. **Illustrieren Sie den resultierenden Hashcontainer.**

x	$h(x)$
<i>Bubbles</i>	1101
<i>Daryl</i>	0011
<i>Hershel</i>	0010
<i>Jon</i>	0001
<i>Stringer</i>	1110
<i>Walter</i>	0110



Aufgabe 6 - Externes Merge-Sort.**1 Punkt**

Gegeben sei eine Relation $R[A]$ mit folgenden Eigenschaften:

- $|R| = 180.000$ Tupel,
- Pro Datenblock werden 3.000 Tupel gespeichert.

Diese Relation soll per externem Merge-Sort sortiert werden.

Geben Sie die **minimale Puffergröße M in Blöcken** an, sodass nur ein **einzig** **Merge-Schritt nötig** ist um $R[A]$ zu sortieren.

Lösung: _____

Aufgabe 7 - Effiziente Anfragebearbeitung.**1 Punkt**

Gegeben sei eine Relation $R[A, B, C]$ mit folgenden Eigenschaften:

- $|R| = 4 \cdot 10^9$ Tupel,
- Pro Datenblock werden $8 \cdot 10^3$ Tupel gespeichert,
- Attribut A hat ganzzahlige Werte gleichverteilt im Bereich $[1; 4 \cdot 10^9]$,
- Attribut B hat ganzzahlige Werte gleichverteilt im Bereich $[1; 4 \cdot 10^6]$,
- Attribut C hat ganzzahlige Werte gleichverteilt im Bereich $[1; 4 \cdot 10^3]$,
- Duplikate werden mittels Tuple Identifier (TID) aufgelöst,
- Folgende Indizes existieren:
 - flacher sparse Index auf Attribut A mit 12.500 Einträgen pro Indexblock,
 - dense B^+ -Baum-Index auf Attribut B , $m = 2.048$, minimale Höhe,
 - dense B^+ -Baum-Index auf Attribut C , $m = 1.024$, minimale Höhe.

Es soll folgende Anfrage beantwortet werden:

$$\sigma_{B=4 \cdot 10^3 \wedge A \geq 4 \cdot 10^3} (R)$$

Geben Sie die **Strategie (0.5 Punkte)** an und berechnen Sie die **Anzahl der Blockzugriffe (0.5 Punkte)** um die Anfrage **möglichst effizient** zu beantworten (1 Knotenzugriff im B^+ -Baum entspricht 1 Blockzugriff).

Aufgabe 8 - Join-Algorithmen.**1 Punkt**

Gegeben seien folgende Relationen $R[A, B, C]$ und $S[C, D, E]$ mit folgenden Eigenschaften:

- $|R| = 10^9$ Tupel gespeichert auf $b_R = 10^6$ Blöcken,
- $|S| = 25 \cdot 10^8$ Tupel gespeichert auf $b_S = 25 \cdot 10^4$ Blöcken.

Weiters existieren folgende Indizes:

- dense Hash-Index auf $S.C$, wobei im Schnitt 3 Blöcke pro Hashwert gelesen werden müssen,
- dense B⁺-Baum-Index auf $R.C$ mit minimaler Höhe, $m = 2.048$.

Es soll der natürliche Join $R \bowtie S$ durchgeführt werden. Als Join-Algorithmus wird der **Indexed-Nested-Loop-Join** verwendet.

Berechnen Sie die **Anzahl der Blockzugriffe (0.5 Punkte)** und geben Sie die **effizienteste Join-Variante (0.5 Punkte)** ($R \bowtie S$ oder $S \bowtie R$) an.

Lösung: _____

Aufgabe 9 - Bitmap Index Scan.**1 Punkt**

Gegeben sei eine Relation $R[A, B, C]$ mit folgenden Eigenschaften:

- $|R| = 10^9$ Tupel gespeichert auf $b_R = 10^6$ Blöcken,
- Non-clustering B⁺-Baum-Index auf Attribut B , $m = 1.024$, maximale Höhe (d.h. die Knoten sind minimal befüllt),
- Duplikate werden mittels Tuple Identifier (TID) aufgelöst,
- Attribut B hat insgesamt 10^6 mal den Wert 1.500, verteilt auf 1.000 Datenblöcke.

Es soll folgende Anfrage beantwortet werden:

$$\sigma_{B=1.500}(R)$$

Berechnen Sie die **Anzahl der Blockzugriffe**, um diese Anfrage **möglichst effizient**

- a. **ohne Anwendung** eines Bitmap Index Scans (**0.5 Punkte**),
- b. **unter Anwendung** eines Bitmap Index Scans (**0.5 Punkte**)

zu beantworten.

Aufgabe 10 - Anfrageoptimierung, Join-Reihenfolge.**1 Punkt**

Gegeben seien 3 Relationen $R[A, B, C]$, $S[C, D, E]$ und $T[C, F, G]$ mit folgenden Eigenschaften:

- $|R| = 10^5$ Tupel, $V(R, A) = 1.000$, $V(R, B) = 4.000$, $V(R, C) = 7.000$
- $|S| = 10^7$ Tupel, $V(S, C) = 2.000$, $V(S, D) = 5.000$, $V(S, E) = 8.000$
- $|T| = 10^9$ Tupel, $V(T, C) = 10^9$, $V(T, F) = 6.000$, $V(T, G) = 9.000$

Weiters sei folgende SQL-Anfrage gegeben:

```
SELECT R.C, S.E, T.F
FROM   R, S, T
WHERE  R.C = S.C
       AND S.C = T.C
```

- a. Zeichnen Sie die **algebraische Normalform als Operatorbaum** für die gegebene SQL-Anfrage. **(0.5 Punkte)**
- b. Wenden Sie **heuristische Optimierung** an, um den **Operatorbaum zu optimieren**. Im resultierenden Operatorbaum soll die **Join-Reihenfolge optimal** sein (d.h. es soll zuerst der Join mit dem kleinsten Zwischenergebnis durchgeführt werden). **(0.5 Punkte)**