

Aufgabe 1

1 Punkt

Betrachten Sie die folgende Festplatte:

- Sektorgröße $B = 2^9\text{B}$,
- Interblock Gap Size $G = 0\text{B}$,
- Sektoren/Spur $S = 2^8$,
- Spuren pro Scheibenseite $T = 300$,
- Anzahl der **beidseitig** beschriebenen Scheiben $D = 2^3$,
- mittlerer Spurwechsel $sp = 2\text{ms}$,
- Datenrate für sequentielles Lesen, wenn von allen Seiten/Scheiben gleichzeitig gelesen wird: 150 MiB/s .

Wie hoch ist die Drehzahl (rpm)?

Lösung: _____

Name:

Matrikelnummer:

3/14

Aufgabe 2

1 Punkt

In Slotted Pages der Größe 2^{13} Bytes sollen 2^{14} Datensätze der Größe 2^4 Bytes gespeichert werden. Der Header ist dense. Es wird Byte-Adressierung verwendet. Geben Sie die minimale Zahl von Slotted Pages an, die dafür benötigt wird.

Lösung: _____

Aufgabe 3

1 Punkt

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

- $|R| = 150.000$.
- Sortierung von R nach A .
- Werte von A und B : eindeutig und zwischen 1 und 1.000.000
- Indices:
 - einstufiger sparse Index $Index_1$ (**kein** B⁺-Baum Index) auf Attribut A .
 - einstufiger sparse Index $Index_2$ (**kein** B⁺-Baum Index) auf $Index_1$.
 - dense B⁺-Baum Index auf B .
- Anzahl der Tupel pro Datenblock: 12.
- Anzahl der Einträge pro Indexblock: 60.

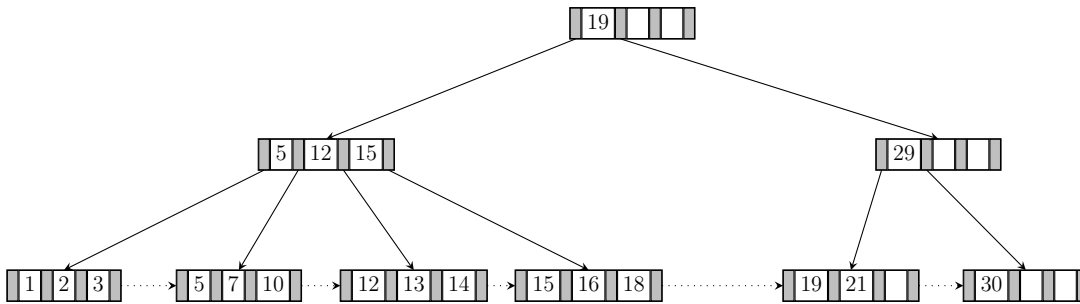
Geben Sie die **Strategie** sowie die **Anzahl der nötigen Blockzugriffe** für die effizienteste Zugriffsmethode an, um die Anfrage $\sigma_{A=50.000}(R)$ zu beantworten.

Lösung: _____

Aufgabe 4

1 Punkt

Aus folgendem B⁺-Baum mit $m = 4$ sollte der Wert 29 gelöscht werden. Der Algorithmus zum Löschen aus einem B⁺-Baum ist jedoch vorzeitig abgebrochen worden (z.B. durch einen Stromausfall). Führen Sie den Algorithmus zu Ende und zeichnen Sie den resultierenden B⁺-Baum.



Aufgabe 5**1 Punkt**

Die folgende Tabelle wird in eine statische Hash-Datei überführt. Als Suchschlüssel wird das Attribut *persNum* verwendet. Der Hashwert ist die erste Ziffer des Attributwerts. Nehmen Sie an, dass kein Bucket Overflow auftritt. **Illustrieren** Sie den Inhalt der daraus resultierenden Hash-Datei.

<i>Name</i>	<i>persNum</i>
Elizabeth	0206
Ross	6828
David	2263
Norman	0997
Jayne	5629
Gary	6927
Michael	1098
Daniel	0509
Keith	4651
Alan	1492

Aufgabe 61 Punkt

Sie wollen externes Merge-Sort auf der Relation $R[A]$ ausführen. Die Werte des Attributs A der Relation sind

$$A = (92, 28, 8, 36, 84, 16, 20, 4, 24, 80, 32, 0, 12, 40, 72, 48, 44, 68, 60, 64, 76)$$

in der angegebenen Reihenfolge. Jeder Block fasst 2 Tupel.

Geben Sie die **minimale Puffergröße in Blöcken** an, sodass genau ein Merge-Schritt nötig ist, um die Relation zu sortieren.

Lösung: _____

Aufgabe 7**1 Punkt**

Gegeben ist die Relation $R[A, B, C]$. Auf Attribut A existiert ein flacher sparse Index. Was ist die effizienteste Strategie um Bereichsanfragen von folgendem Typ zu beantworten?

$$\sigma_{a < A < b}(R)$$

Geben Sie **alle notwendigen Schritte** an.

Aufgabe 81 Punkt

Gegeben ist eine Relation $R[A, B]$. R beinhaltet 1.000.000 Tupel, wobei jeder Wert von Attribut B 2.500-mal in R vorkommt. Die Relation R ist auf 400 Blöcken gespeichert. Es gibt einen dense B^+ Baum Index auf Attribut B . Jeder Knoten im B^+ kann maximal 250 Schlüssel beinhalten. Die Duplikate werden durch Verknüpfung mit dem TID (Tuple Identifier) eindeutig gemacht. Es soll folgende Anfrage ausgeführt werden:

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

Sie können zwischen vier Strategien zur Beantwortung dieser Anfrage wählen:

- a.) Wert b im Index suchen, danach die Blätter des Indexes sequentiell scannen bis $B \neq b$ und jedem Index-Eintrag zum Datenblock folgen.
- b.) Scan der gesamten Datenblöcke nach Tupel, die $B = b$ erfüllen.
- c.) Binäre Suche auf Datenblöcken nach erstem Tupel, das $B = b$ erfüllt, danach die Datenblöcke sequentiell scannen bis $B \neq b$.
- d.) Strategie a.) mit Bitmap Index Scan.

(0.5 Punkte) Geben Sie an welche **zwei Strategien am effizientesten** sind (bzgl. der Blockzugriffe).

(0.5 Punkte) Begründen Sie Ihre Antworten (Sie müssen nicht alle Blockzugriffe ausrechnen).

Einzelheiten für Aufgaben 9 – 12.

Gegeben sind drei Relationen: R , S und T mit den folgenden Eigenschaften.

$R[\underline{A}, B, C, G]$	$S[\underline{C}, D, E]$	$T[\underline{A}, \underline{C}, F]$
$ R = 10.000$	$ S = 25.000$	$ T = 400.000$
$b_R = 500$	$b_S = 1.000$	$b_T = 20.000$
sparse Index auf A dense Index auf C	sparse Index auf C	sparse Index auf (A, C)
$V(R, A) = 10.000$	$V(S, C) = 25.000$	$V(T, A) = 400$
$V(R, B) = 2.500$	$V(S, D) = 500$	$V(T, C) = 1.000$
$V(R, C) = 1.000$	$V(S, E) = 2.500$	$V(T, F) = 100.000$
$V(R, G) = 100$		

- Die Relationen sind nach Primärschlüssel sortiert und auf hintereinander liegenden Blöcken gespeichert.
- Unterstrichene Attribute sind Primärschlüssel.
- Die Werte in den Tupeln sind gleichverteilt und unabhängig.
- $V(R, A)$ ist die Anzahl der verschiedenen Werte von Attribut A in Relation R .
- Ein Index auf (A, C) bedeutet, dass ein Schlüssel die Kombination von beiden Attributen, A und C , in dieser Reihenfolge ist.
- Alle Indizes sind B^+ -Bäume. Die Duplikate werden durch Verknüpfung mit dem TID (Tuple Identifier) eindeutig gemacht.
- Jeder Knoten in B^+ -Baum kann 100 Schlüssel speichern.
- Der Buffer hat Platz für 51 Blöcke.

Name:

Matrikelnummer:

11/14

Aufgabe 9

1 Punkt

Berechnen Sie die Kosten (Blockzugriffe) von einem Natural Join zwischen Relationen R und S mit Hilfe des Index-Nested-Loop-Join Algorithmus und geben Sie die effizientere Join-Reihenfolge ($R \bowtie S$ oder $S \bowtie R$) an.

Aufgabe 10

1 Punkt

Berechnen Sie die Kosten (Blockzugriffe) der folgenden Anfrage, wenn ein Block-Nested-Loop-Join Algorithmus für den Join benutzt wird. Die Selektion soll vor dem Join durchgeführt werden und in einer temporären Tabelle gespeichert werden.

$$\sigma_{G=g}(R) \bowtie S$$

Lösung: _____

Aufgabe 111 Punkt

Gegeben ist die folgende Anfrage:

$$\pi_E(\sigma_{C=c}(S) \bowtie \sigma_{F=f}(T))$$

(0.5 Punkte) Zeichnen Sie zuerst die algebraische Normalform (Projektionen, Selektionen, Kreuzprodukte).

(0.5 Punkte) Formulieren Sie die Anfrage so um, dass Selektionen und Projektionen so weit wie möglich nach unten geschoben werden (sodass der kleinstmögliche Join durchgeführt wird).

Aufgabe 12

1 Punkt

Schätzen Sie die Kardinalität des Ergebnisses der folgenden Anfrage ab ($\sigma_{A=a}(R) \neq \emptyset$, $\sigma_{D=d}(S) \neq \emptyset$).

$$(\sigma_{A=a}(R)) \bowtie (\sigma_{D=d}(S)) \bowtie T$$

Lösung: _____