



Aufgabe 1 (14 Punkte)

Sie speichern in einer Tabelle Personen, die in 1500 verschiedenen Städten wohnen:

https://hpi.de/fileadmin/user_upload/fachgebiete/plattner/teaching/EnterpriseApplications/SS2020/05_cities_stored.txt

Um unnötige Scans dieser Tabelle zu erkennen und zu vermeiden, nutzen Sie einen Bloomfilter.

Prüfen Sie für verschiedene Konfigurationen des Bloomfilters (gegeben durch die Anzahl der Bits und Hash-Funktionen), wie oft das Nichtvorhandensein der folgenden 556 Städte nicht erkannt wird:

https://hpi.de/fileadmin/user_upload/fachgebiete/plattner/teaching/EnterpriseApplications/SS2020/05_cities_queried.txt

Die Dateien für die gespeicherten und angefragten Städte speichern pro Stadt jeweils sechs Zufallszahlen aus dem Wertebereich [0; 119.999].

Der i -te Hashwert pro Stadt ergibt sich aus der i -ten Zufallszahl Modulo der Bitanzahl des Bloomfilters.

Verwenden Sie jeweils die ersten N Hash-Werte pro Zeile.

Beispiel für den Bloomfilter mit 3000 Bits und $N=4$ Hash-Funktionen:

Die zu verwendenden (ersten vier) Hash-Werte für die Zeile

'Aach (Baden-Württemberg) | 3134 | 75168 | 75713 | 15481 | 92491 | 87369'

lauten 134, 168, 713 und 481.

Die Fehlerrate für die 556 angefragten Städte beträgt ca. 57%.

Ergänzen Sie Ihre ermittelten Fehlerraten in der folgenden Tabelle.

Bits	N = 1	N = 2	N = 3	N = 4	N = 5	N = 6
2000	48%					
3000				57%		
4000						
5000				21%		
6000						24%

Aufgabe 2 (14 Punkte)

Gegeben sind drei Transaktionen mit der zeitlichen Reihenfolge ihrer Operationen.

Entscheiden Sie pro Nebenläufigkeitskontrollverfahren, ob und bei welcher Operation die einzelnen Transaktionen abbrechen. Begründen Sie Ihre Antwort.

Abgebrochene Transaktionen werden nicht neugestartet.

- Basisvariante des Zeitstempelverfahrens (5 Punkte)
- Nebenläufigkeitskontrolle durch Validierung (3 Punkte)
- MVCC auf Basis des Zeitstempelverfahrens (6 Punkte)

Hinweis: grau gefärbte Operationen sind nur für Nebenläufigkeitskontrolle durch Validierung relevant und müssen für die anderen Verfahren ignoriert werden.



Skizzieren Sie für a) und c) die Zeitstempelübergänge (zum Beispiel analog zu den Folien 10 und 26 des Foliensatzes 10_DB_Nebenlaueufigkeitskontrolle_*).

Skizzieren Sie für b), was bei den einzelnen Validierungen geprüft wird (zum Beispiel analog zu Folie 22 des Foliensatzes 10_DB_Nebenlaueufigkeitskontrolle_*).

T1	T2	T3	TS
READ(A)			1
	WRITE(B)		2
		READ(A)	3
READ(B)			4
VALIDATE			5
	WRITE(A)		6
	VALIDATE & WRITE		7
		READ(B)	8
		VALIDATE	9

Abgabeanweisung

Die Aufgaben **müssen** in Zweierteams bearbeitet werden.

Die Abgabe erfolgt über das HPI-Abgabesystem:

<https://www.osm.hpi.uni-potsdam.de/submit/dashboard/>

Alle Studenten müssen sich dafür beim Abgabesystem anmelden.

Der Name des Partners muss bei der Abgabe der Ergebnisse angegeben werden.

Reichen Sie eine txt-Datei mit ihren Lösungen ein.

Verwenden Sie dafür das folgende Template:

https://hpi.de/fileadmin/user_upload/fachgebiete/plattner/teaching/EnterpriseApplications/SS2020/submission5.txt

Abgabefrist: 16. Juli 2020 Anywhere on Earth (AoE).

Die pünktliche Abgabe und das Bestehen der Übungsblätter sind Voraussetzung für die Zulassung zur Klausur.

14 von 28 Punkte müssen zum Bestehen des Übungsblattes erreicht werden.