



Aufgabe 1 (32 Punkte)

Lesen Sie Ausschnitte aus Hasso Plattners Veröffentlichungen

- **„A Common Database Approach for OLTP and OLAP Using an In-Memory Column Database“**
Abschnitt 1 – 3
<http://www.sigmod09.org/images/sigmod1ktp-plattner.pdf>
- **„The Impact of Columnar In-Memory Databases on Enterprise Systems“**
Abschnitt 1 – 5.2
<http://www.vldb.org/pvldb/vol7/p1722-plattner.pdf>

und fassen Sie die Inhalte unter Beantwortung folgender Fragen zusammen.

- a) Was sind Gründe für die historische Trennung von OLTP und OLAP in verschiedene Systeme? (6 Punkte)
- analytische Anfragen verlangsamen das transaktionale System
 - verschiedene Datenbankschemata notwendig
 - inkompatible SQL-Dialekte
 - kein Bedarf an Vereinigung
 - performante und flexible analytische Anfragen möglich
 - unterschiedliche Workload-Eigenschaften
- b) Warum und wie ist es heutzutage möglich OLTP- und OLAP-Anfragen in einem System effizient zu bearbeiten? (7 Punkte)
- Schnellere Hardware (Moore's Law)
 - Entwicklung spaltenbasierter Hauptspeicher-DBs
 - Entwicklung von NV-RAM
 - Entwicklung von Streaming-Systemen
 - größerer und günstigerer Hauptspeicher
 - Entwicklung der Blockchain
 - Multicore-CPU's
- c) Warum werden Inserts in Unternehmensanwendungen durch den Umstieg auf spaltenbasierte Hauptspeicherdatenbanken nicht unbedingt langsamer als in einer Zeilendatenbank? (3 Punkte)
- Inserts werden unter keinen Umständen schneller
 - Weil redundante materialisierte Aggregate weggelassen und nicht aktualisiert werden müssen
 - Weil ein spaltenbasiertes Speicherlayout Inserts besser unterstützt



d)

Erklären Sie warum die Vereinigung von OLTP und OLAP die Entwicklung neuer Unternehmensanwendungen vereinfacht. (2 Punkte)

Einfacheres Schema mit konsistenten Daten.

Alle transaktionalen Daten können für Analysen genutzt werden (keine Auswahl von materialisierten Aggregaten).

e) Welche Vorteile haben spaltenorientierte gegenüber zeilenorientierten Datenbanken? (3 Punkte)

Komprimierung und Verarbeitung auf komprimierten Daten gut möglich.

Ein spaltenbasiertes Speicherlayout ist für alle Anfragetypen besser geeignet

Wenn SQL-Anfragen viele Werte weniger Attribute lesen, erlaubt eine Spaltenorientierung besseres Caching.

f) Was ist der TPC-C Benchmark? (2 Punkte)

Ein standardisierter OLTP Benchmark

Ein standardisierter OLAP Benchmark

g) Welches sind analytische Benchmarks des Transaction Processing Performance Councils? (5 Punkte)

TPC-A

TPC-C

TPC-DS

TPC-E

TPC-H

h) Welche zwei Themen (Konzepte) aus den gelesenen Abschnitten fanden Sie am spannendsten? (2 Punkte)

i) Welche zwei Themen (Konzepte) aus den gelesenen Abschnitten fanden Sie nicht so interessant? (2 Punkte)



Aufgabe 2 (14 Punkte)

Für die Bearbeitung der zweiten Übung benötigen Sie Zugang zu folgendem SAP HANA-Datenbanksystem:

Host Name: vm-syene.eaalab.hpi.uni-potsdam.de (192.168.30.23)
Instance Nummer: 02
User Name: STUDENT2020
Password: Student2020

Der Zugriff auf die Datenbank ist nur aus dem HPI-Netz möglich. Ihr Laptop muss außerdem unter <https://byod.hpi.de/> für die Nutzung interner HPI-Ressourcen registriert sein.

Auf welche SAP-Tabellen des Schemas „SAPHPB“ haben Sie Zugriff? Sie können dafür folgende SQL-Anfrage nutzen: (7 Punkte)

```
select table_name from tables where schema_name = 'SAPHPB';
```

Welche Informationen sind in den einzelnen Tabellen gespeichert? (7 Punkte)

Für Datenbankanfragen können Sie beispielsweise den (von uns empfohlenen) Python SAP HANA Connector (<https://github.com/SAP/pyhdb/>) verwenden.

Die SAP HANA Tools (<https://tools.hana.ondemand.com/#hanatools>) bieten zusätzlich eine grafische Oberfläche zur Eingabe von SQL-Anfragen und Datenanzeige.

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/submission1.txt

Abgabefrist: 11. Mai Anywhere on Earth (AoE).

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

23 von 46 Punkte müssen zum Bestehen des Übungsblattes erreicht werden.