

11. Übungsblatt

Ausgabe: 18. Januar 2016 · Besprechung: 25. Januar 2016

1 Star Schema Benchmark

Wir haben in der Übung bereits den TPC-H Benchmark kennengelernt, der in Industrie und Forschung zum quantitativen Vergleich von Datenbankmanagementsystemen eingesetzt wird. Der Star Schema Benchmark [2] ist eine abgewandelte Version des TPC-H Benchmarks.

1. Welche Gründe sprechen für ein Star Schema? In welchen Anwendungsszenarien kommen Star Schemata zum Einsatz?
2. Erläutern Sie Schema und Anfragen des Star Schema Benchmarks! Welche Muster sind in den Anfragen erkennbar?

2 Star Join Anfragen

1. Was verstehen Sie unter einem Star Join?
2. Welche Techniken haben Sie kennengelernt, um Star Join Anfragen effizient zu berechnen? Diskutieren Sie Vorteile und Nachteile der Techniken und in welchen Situationen sie eingesetzt werden sollten.
3. Was ist ein Semi Join? Wie könnte man Semi Joins zur Optimierung von Star Joins einsetzen?
4. Wenden Sie alle Techniken beispielhaft auf Anfrage Q2.3 des Star Schema Benchmarks an!
5. Eine fortgeschrittene Technik zur effizienten Berechnung von Star Joins ist der sogenannte „Invisible Join“ von Abadi et. al. [1]. Erklären Sie die Funktionsweise und wenden Sie die Technik auf die Beispielanfrage an!

Literatur

- [1] Daniel J. Abadi, Samuel R. Madden, and Nabil Hachem. Column-stores vs. row-stores: How different are they really? In *SIGMOD*, pages 967--980. ACM, 2008.
- [2] Pat O'Neil, Elizabeth J. O'Neil, and Xuedong Chen. The star schema benchmark (ssb), 2009. Revision 3, <http://www.cs.umb.edu/~poneil/StarSchemaB.PDF>.