

## Hinweise zur Einrichtung von Postgres und TPC-H

1. Laden Sie den Sourcecode von PostgreSQL auf Ihren Rechner. Anschließend kompilieren und installieren Sie ihn. Der Download zum Postgres Quellcode wurde zur Verfügung gestellt.
2. In eine leere PostgreSQL-Datenbank laden Sie Daten des TPC-H-Benchmarks.  
TPC-H ist der wohl bekannteste Benchmark für “decision support”-Aufgaben. Auf der Webseite des TPC-H-Konsortiums finden Sie ein Paket, das Generatoren für jeweils Daten (**dbgen**) und Anfragen (**qgen**) enthält.  
In den Anfragen zum Benchmark (Verzeichnis **queries**) müssen einige Klauseln umformuliert werden, die sich auf Zeitintervalle beziehen. PostgreSQL versteht hier z. B. die Formulierung

```
l_shipdate <= date '1998-12-01' - interval '118 day'
```

Auf die PostgreSQL Syntax angepasste Query Templates für den TPC-H benchmark wurden deshalb als Download hinzugefügt.

3. Der Datengenerator **dbgen** erzeugt ein Format, in dem jede Zeile mit einem ‘|’ endet. Dieses kann mit dem **COPY**-Kommando von PostgreSQL nicht korrekt eingelesen werden. Entfernen Sie daher vor dem Laden diese ‘|’ am Zeilenende (z. B. mittels **sed -i 's/|\$/g'**).
4. Verwenden Sie bitte keine zu großen Benchmark-Instanzen. Ohne explizites Anlegen von Datenbank-Indizes können sonst sehr lange Laufzeiten entstehen und/oder die Auswertung viel Platten- und Hauptspeicher benötigen. Wir empfehlen einen Skalierungsfaktor von höchstens 1.
5. Zusätzlich werden noch einige Integritäts-Constraints importiert, die speziell für Postgres optimiert sind. Dadurch wird eine bessere Query-Optimierung möglich. Die zusätzlichen Constraints können aus die Datei **tpch-alter.sql** zur Datenbank hinzugefügt werden. (mittels **psql --file=tpch-alter.sql**)