

Globale Statistiken schnell erzeugen

Das Erstellen von globalen Statistiken ist vor allem für große Faktentabellen in einem Data Warehouse sehr aufwändig und nimmt deshalb häufig viel Zeit in Anspruch. Insbesondere wenn das Datenvolumen der Faktentabellen stetig wächst, benötigt die Aktualisierung der globalen Statistiken immer mehr Zeit.

Faktentabellen sind jedoch in der Regel partitioniert, und es ändern sich meist nur die Daten einer oder weniger Partitionen. In diesem Szenario kann seit Oracle 11g der Einsatz inkrementeller Statistiken in Betracht gezogen werden. Die Voraussetzung: Die lokalen Statistiken der einzelnen Partitionen müssen aktuell sein und die inkrementellen Statistiken für die betroffenen Tabellen aktiviert werden.

Inkrementelle Statistiken aktivieren

Inkrementelle Statistiken werden mit Hilfe der Prozedur `DBMS_STATS.SET_TABLE_PREFS` für einzelne Tabellen aktiviert. Dazu ist die Preference `INCREMENTAL` auf `TRUE` zu setzen. (siehe Listing 1)

Durch das Aktivieren der inkrementellen Statistiken speichert die Oracle-Datenbank für jede Partition der Tabelle ein sogenanntes Synopsis-Objekt im `SYSAUX`-Tablespace. Die Synopsis-Objekte enthalten statistische Metadaten für die einzelnen Partitionen und die Spalten in den Partitionen. Im Vergleich zu den vollständigen Partitionsdaten sind die Synopsis-Objekte sehr klein.

Globale Statistiken inkrementell aktualisieren

Wird eine neue Partition hinzugefügt oder eine bestehende Partition geändert, müssen für diese zunächst die lokalen Statistiken aktualisiert werden. Anschließend kann die Oracle-Datenbank anhand der Synopsis-Daten die globalen Statistiken erzeugen, ohne die gesamte Tabelle lesen zu müssen. Der Ablauf könnte z. B. wie folgt aussehen. (siehe Listing 2)

Reduzierte Laufzeiten für die Erzeugung globaler Statistiken

Durch dieses Verfahren wird die Laufzeit für die Erzeugung der globalen Statistiken erheblich reduziert. Insbesondere steigt die Laufzeit auch

bei stetigem Wachstum der Faktentabelle im Vergleich zur nicht inkrementellen Erzeugung der Statistiken deutlich geringer.

Die folgende Abbildung zeigt die schnellere Laufzeit inkrementell erzeugter gegenüber nicht inkrementell erzeugten Statistiken am Beispiel einer partitionierten Faktentabelle. Betrachtet wird ein Zeitraum von 31 Tagen, in dem täglich 10 Millionen neue Datensätze in die Faktentabelle eingefügt werden. (siehe Abbildung 1)



Globale Statistiken schnell erzeugen

Listing 1:

```
begin
  dbms_stats.set_table_prefs
    (ownname => 'MART'
    , tabname => 'FAKT_VERKAUF'
    , pname => 'INCREMENTAL'
    , pvalue => 'TRUE');
end;
```

Listing 2:

```
begin
  -- lokale Statistiken für Partition mit neuen Daten aktualisieren
  -- (hier Partition D20150302)
  dbms_stats.gather_table_stats(
    (ownname => 'MART'
    , tabname => 'FAKT_VERKAUF'
    , partname => 'D20150302');
    , granularity => 'PARTITION');

  -- globale Statistiken inkrementell aktualisieren
  dbms_stats.gather_table_stats
    (ownname => 'MART'
    , tabname => 'FAKT_VERKAUF');
end;
```

Abbildung 1:

