Update Statistics for the Oracle Cost-Based Optimizer

White Paper: Oracle Database Administration

February 2012
TABLE OF CONTENTS

OVERVIEW ........................................................................................................................................... 3
UPDATE STATISTICS METHODS ........................................................................................................... 4
CONFIGURATION..................................................................................................................................... 5
Object Information................................................................................................................................. 5
Default Settings.................................................................................................................................... 5
TODO Settings....................................................................................................................................... 7
INTERNAL RULES FOR UPDATE STATISTICS ...................................................................................... 8
RECOMMENDATIONS FOR PRODUCTION OPERATION........................................................................... 10
ORACLE TABLE MONITORING ............................................................................................................. 11
PARALLEL UPDATE STATISTICS ......................................................................................................... 12
THE DBMS_STATS PACKAGE ................................................................................................................ 13
BRCONNECT PARAMETER STATS_DBMS_STATS .................................................................................. 14
UPDATE STATISTICS FOR INFOCUBES WITH BRCONNECT ............................................................... 15
UPDATE STATISTICS FOR INDIVIDUAL PARTITIONS ....................................................................... 16
VALIDATING THE STRUCTURE OF ORACLE DICTIONARY TABLES ...................................................... 17
CREATING SPACE STATISTICS WITH BRCONNECT ............................................................................. 18
BRCONNECT OPTION -FORCE WITH UPDATE STATISTICS ............................................................... 19
ADDITIONAL INFORMATION ................................................................................................................ 20
SAP Library ........................................................................................................................................... 20
SAP Notes ............................................................................................................................................ 20
OVERVIEW

The cost-based optimizer (CBO) is an essential component for improving Oracle database performance. It uses statistics describing database objects and the distribution of data within them. Statistics on data distribution help identify which data is the most selective.

Using these statistics, the CBO determines the best plan for data access. Cost-based optimization gives better performance, but only if the statistics are up-to-date. Therefore, it is important that you update optimizer statistics regularly to maintain the best possible Oracle database performance. Out-of-date statistics can cause the CBO to choose poor access paths, which reduces performance.

The SAP functionality to update statistics is fully up-to-date in BRCONNECT. This functionality is also available in the DBA Planning Calendar in transaction DBACOCKPIT or DB13 of the Computing Center Management System (CCMS), enabling you to schedule update statistics regularly without manual intervention.

The main benefit in update statistics with BCONNECT is parallel processing, which results in significantly faster run times when activated. For more information, see “Parallel Update Statistics” [Page 12].

This document answers the following questions:

- What methods are used to update statistics?
- How is update statistics configured?
- What are the internal rules for update statistics?
- What are the recommendations for production operation?
- How can Oracle table monitoring improve run times for update statistics?
- How can parallel update statistics improve run times?
- What are the BRCONNECT -force options for update statistics?
UPDATE STATISTICS METHODS

To update statistics, BRCONNECT normally:

1. Deletes statistics for certain tables, where the statistics are misleading, including pool and cluster tables (up to Oracle 9i) and objects where the ACTIVE field in the DBSTATC table is set to N.
2. Checks each table to see if the statistics are out-of-date or non-existent.
3. If required, updates the statistics on the table immediately after the check

BRCONNECT uses the DBSTATC table as a control table. It performs the check and update (the final two steps above) in a single phase. The section "BRCONNECT Option -force with Update Statistics" [Page 19] describes how you can influence update statistics.
The **DBSTATC** control table controls update statistics. In a new installation of the SAP System, the **DBSTATC** table includes entries for the most important tables, that is, large and dynamic tables. The **DBSTATC** table influences how update statistics is executed, but no longer contains the intermediate results (as in earlier releases).

- You can create entries in the **DBSTATC** table for either of the following reasons:
- To specify non-standard analysis method and sample size

To create temporary statistics for the SAP application monitor for tables that normally have no statistics, such as pool and cluster tables (up to Oracle 9i)

You can customize update statistics by editing this table using transaction DB21, in which you can change the following (the field names are given in brackets):

**Object Information**
- **Database object** (**DBOBJ**)  
  Table name (for example, APQD)
- **Object type** (**DOTYP**)  
  Database object type (standard setting for tables 01)
- **Object owner** (**OBJOW**)  
  Table owner (standard setting SAPS3/SAPSR3 or SAP<SAPSID>)
- **Database** (**DBTYP**)  
  Database system used (for example, ORACLE, or blank if database-independent)

**Default Settings**
- **Use** (**VWTYP**)  
  This identifies the service or application for which the table entry is used:
  - **A** (application monitor)  
    For the SAP application monitor (ST07) and the optimizer
  - **O** (optimizer)  
    Only for the optimizer (standard setting)
- **Active** (**ACTIV**)  
  This indicates if and how the statistics will be updated and in which context they will be used. Possible settings are:
  - **A** (Active)  
    Statistics are checked and updated if needed.
  - **N** (No) (up to Oracle 9i, later only in exceptional circumstances)  
    Statistics cannot be created. You can only set this if Use is O. BRCOMPDELETE deletes existing statistics during next run of brconnect -f stats.
  - **P** (Priority)  
    Priority Statistics (P) are checked and updated if needed before statistics with lower priority (A).
  - **R** (Restrictive)  
    Certain statistics values are externally determined. Other values follow the standard check and update procedure (as with value A, see above). For more information, see SAP Note 1020260.
  - **U** (Unconditional)
Statistics are updated without checking during every run of brconnect -f stats.

- **I** (Ignore)
  
  BRCONNECT does **not** check whether create or update statistics is required. If statistics are present, they are **not** updated or deleted.

- **Changed on** (AEDAT)
  
  Date when one of the settings was changed. This field is automatically set by transaction DB21.

- **Analysis method** (AMETH)
  
  This is the analysis method used to create or update statistics. The possible analysis methods are:

  - **E**
    
    Estimate table statistics including indexes
  
  - **EH**
    
    Estimate table statistics and create histograms
  
  - **EI**
    
    Estimate table statistics and verify index structure
    
    Caution: this locks indexes.
  
  - **EX**
    
    Estimate table statistics, create histograms, and verify index structure
    
    Caution: this locks indexes.
  
  - **C**
    
    Compute table statistics including indexes
  
  - **CH**
    
    Compute table statistics and create histograms
  
  - **CI**
    
    Compute table statistics and verify index structure
    
    Caution: this locks indexes.
  
  - **CX**
    
    Compute table statistics, create histograms, and verify index structure
    
    Caution: this locks indexes.
  
  - **A**
    
    Collect statistics using auto-sampling, including indexes
  
  - **AH**
    
    Collect statistics using auto-sampling, including indexes, and create histograms
  
  - **AI**
    
    Collect statistics using auto-sampling, including indexes, and verify index structure
    
    Caution: this locks indexes.
  
  - **AX**
    
    Collect statistics using auto-sampling, including indexes, create histograms, and verify index structure
    
    Caution: this locks indexes.

- **Analysis Option** (OPTIO)

  Specification of sample size (only relevant to **E**, **EH**, **EI**, and **EX**)

  - **P<\textit{p}>**
    
    \textit{p} percent of the table rows are used for the analysis (1 - 100, .001 - .999)
- $r<\text{thousand table rows are used for the analysis (}1\text{ – 9999)}$

- **History (HISTO)**
  - If this is selected, the statistic results are archived to the `DBSTATHORA` and `DBSTAIHORA` tables.

- **Customer (PLAND)**
  - Use this field to suppress warnings issued by BRCONNECT if pool or cluster tables were specified in the `DBSTATC` table to always have statistics (up to Oracle 9i). Only use this field in exceptional circumstances.

**TODO Settings**

- **Check (DURAT)**
  - If this is selected, BRCONNECT checks once whether new statistics are required and updates them if necessary, even when `Analysis method` is `I` or `N`. If required, you can manually change the entries in the `DBSTATC` control table with transaction DB21. When it has finished, `brconnect -f stats` resets the `Check` field.

- **Update (TOBDO)**
  - If this is selected, BRCONNECT updates statistics once without check even when `Analysis method` is `I` or `N`. When it has finished, `brconnect -f stats` resets the `Update` field.

- **Changed on (TDDAT)**
  - Date on which the `Check` or `Update` fields were set. This field is automatically set by transaction DB21. BRCONNECT does not add or delete rows to or from the `DBSTATC` table and does not change the analysis method and sample size. The `DBSTATC` table is used purely as a control table.
INTERNAL RULES FOR UPDATE STATISTICS

This algorithm is used by BRCONNECT to update statistics.

1. BRCONNECT determines the working set of tables and indexes to be checked and updated. To do this, it uses:
   - Options `-t|-table` and `-e|-exclude` set with `brconnect -f stats` (these options take priority)
   - `stats_table` and `stats_exclude` parameters from `init<DBSID>.sap`

2. If the working set contains pool, cluster (up to Oracle 9i) or other tables that have the `ACTIVE` flag in the `DBSTATC` table set to `N`, BRCONNECT immediately deletes the statistics for these tables if available, because they can negatively affect database performance.

3. BRCONNECT checks statistics for the remaining tables in the working set, including tables that have the `ACTIVE` flag in the `DBSTATC` table set to `A`, `E`, or `R` as follows:
   - If the table has the MONITORING attribute set (standard since Oracle 10g), BRCONNECT reads the number of inserted, deleted, and updated rows from the `DBA_TAB_MODIFICATIONS` table.
   - Otherwise, BRCONNECT uses the standard method (see table below) to update statistics on a unique index.

BRCONNECT uses the following standard method to check and update a table's statistics:
- Method and sample defined for the table in the `DBSTATC` table (has highest priority)
- Method and sample from the options `-m|-method` or `-s|-sample_size` of `-f stats` (takes priority) or the `stats_method` or `stats_sample_size` parameter
- Default method and sample (has lowest priority)

The following table describes the default method:

<table>
<thead>
<tr>
<th>Number of rows in table</th>
<th>Analysis Method</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row s &lt; 10,000</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>10,000 &lt;= Row s &lt; 100,000</td>
<td>E</td>
<td>P30</td>
</tr>
<tr>
<td>100,000 &lt;= Row s &lt; 1,000,000</td>
<td>E</td>
<td>P10</td>
</tr>
<tr>
<td>1,000,000 &lt;= Row s &lt; 10,000,000</td>
<td>E</td>
<td>P3</td>
</tr>
<tr>
<td>10,000,000 &lt;= Row s &lt; 100,000,000</td>
<td>E</td>
<td>P1</td>
</tr>
<tr>
<td>100,000,000 &lt;= Row s &lt; 1,000,000,000</td>
<td>E</td>
<td>P.3</td>
</tr>
<tr>
<td>1,000,000,000 &lt;= Row s &lt; 10,000,000,000</td>
<td>E</td>
<td>P.1</td>
</tr>
</tbody>
</table>
Analysis method C means compute the statistics exactly. Analysis method E means estimate the statistics using the sample size specified.

For example, "E P10" means that BRCONNECT takes an estimated sample using 10% of rows.

The sample size is doubled if single table partitions are chosen for update statistics.

4. BRCONNECT uses the number of new rows for each table in the working set, as derived in the previous step, to see if either of the following is true:
   - For tables without the MONITORING attribute:
     - $<\text{new\_number\_rows}> > = <\text{old\_number\_rows}> * (100 + <\text{threshold}>)/100$
     - $<\text{new\_number\_rows}> <= <\text{old\_number\_rows}> \times 100 / (100 + <\text{threshold}>)$
   - For tables with the MONITORING attribute (see SAP Note 744483):
     - $<\text{old\_number\_rows}> + <\text{number\_of\_inserted\_rows}> >= <\text{old\_number\_rows}> * (100 + <\text{threshold}>)/100$
     - $<\text{old\_number\_rows}> + <\text{number\_of\_updated\_rows}> >= <\text{old\_number\_rows}> * (100 + <\text{threshold}>)/100$
     - $<\text{old\_number\_rows}> + <\text{number\_of\_inserted\_rows}> >= (<\text{old\_number\_rows}> + <\text{number\_of\_inserted\_rows}> - <\text{number\_of\_deleted\_rows}>) * (100 + <\text{threshold}>)/100$

The standard threshold is 50, but the value defined in the -c|change option of -f stats or the stats_change_threshold parameter is used if specified.

The threshold is automatically halved for partitioned tables, if they are being checked as complete tables (not for single partitions).

5. BRCONNECT immediately updates statistics after checking for the following tables:
   - Tables where either of the conditions in the previous step is true
   - Tables from the DBSTATC table with the value ACTIVE field U

6. BRCONNECT writes the results of update statistics to the DBSTATTORA table and also, for tables with the DBSTATC history flag or usage type A, to the DBSTATTHORA table.

7. For tables with update statistics using methods EI, EX, CI, or EX, AI, or AX, BRCONNECT validates the structure of all associated indexes and writes the results to the DBSTATTORA table and also, for tables with the DBSTATC history flag or usage type A, to the DBSTATAIHORA table.
RECOMMENDATIONS FOR PRODUCTION OPERATION

In production operation, we strongly recommend you to update statistics at least **weekly** using the DBA Planning Calendar (that is, transaction DBACOCKPIT or DB13). For Oracle 10g we recommend you to do this even daily. We provide several action patterns that include a weekly or daily run of update statistics.

If you use a tool such as **cron** (UNIX) or **at** (Windows), we recommend the following standard command:

`brconnect -u / -c -f stats -t all`

We recommend you to normally use the above method to update statistics.

We also recommend you to use this call immediately after a SAP System upgrade or after a large amount of data has been changed in the database.

In the above call, BRCONNECT checks statistics for all SAP tables and indexes using the standard method. If the change threshold for a particular table is exceeded, statistics are updated for that table.

BRCONNECT also deletes damaging statistics. For more information, see "Delete damaging statistics" [page 19] below.

The following are further examples of update statistics commands:

- Update (that is, create) statistics only for tables and indexes with missing statistics
  `brconnect -u / -c -f stats -t missing`

- Check and update statistics for all tables specified in the `DBSTATC` table
  `brconnect -u / -c -f stats -t dbstatc_tab`
ORACLE TABLE MONITORING

Checking whether statistics need updating can take a long time for large databases. This is because statistics have to be updated for unique indexes, which can take a long time. You can practically eliminate the time required to check statistics by using Oracle table monitoring.

With this feature, the Oracle database system constantly updates information on the numbers of rows inserted, deleted, and updated for database tables, entering the results in the `DBA_TAB_MODIFICATIONS` table. It only takes BRCONNECT a very short time to read this results table and determine whether update statistics is required for a given database table.

For more information, see SAP Note 408527.

In Oracle 10g or higher, table monitoring is automatically enabled for all tables.
PARALLEL UPDATE STATISTICS

BRCONNECT supports parallel processing with:

- Parallel threads, especially for large numbers of tables
- The DBMS_STATS package, especially for large tables

Parallel threads, if activated, significantly improves run times. You can implement this as follows:

- **Command call** `brconnect -p <number of threads>`
- **Parameter** `stats_parallel_degree` in the `init<SID>.sap` file

Here are some examples:

```
brconnect -u / -c -f stats -t all -p 4
stats_parallel_degree = 4
```

Each of these sets the number of threads – that is, the degree of parallelism – to 4. The default is 1, which means that update statistics runs in non-parallel – that is, serial – mode. In either case, a single thread processes each table.

For more information on parallel processing with BRCONNECT, see SAP Note 403713.

You can also use the DBMS_STATS package to update statistics in parallel for individual tables. The parallel processing option in this package considerably reduces run times for very large tables (that is, several GB and more). Each table can be processed by several threads.

For more information on parallel processing with DBMS_STATS, see SAP Note 408532.

You can use both the above types of parallel processing in the same BRCONNECT run.
THE DBMS_STATS PACKAGE

Up to Oracle 9i, the ANALYZE command is used as standard to update statistics on the tables and indexes of an SAP Oracle database. But BRCONNECT can also do this by calling the supplied DBMS_STATS package. This can reduce the BRCONNECT runtime for updating statistics. DBMS_STATS is the standard to update statistics as of Oracle 10g.

The DBMS_STATS package especially improves the runtime if you use the internal parallel setting for collecting statistics for very large tables (that is, several GB and more). However, this feature does not have so much runtime benefit for smaller and medium-sized tables.

To activate update statistics with the DBMS_STATS package, specify parameter stats_dbms_stats in the init<SID>.sap profile. You only need to do this for Oracle 9i.

For more information on the stats_dbms_stats parameter, see “BRCONNECT Parameter stats_dbms_stats” [Page 14] below.

For more information, see SAP Note 408532.
**BRCONNECT PARAMETER STATS_DBMS_STATS**

You can control the use of the DBMS_STATS package for updating statistics with BRCONNECT through the `stats_dbms_stats` parameter in init<SID>.sap:

```
stats_dbms_stats = all:R|B[<buckets>|A|S|R]:0|<degree>|A|D|
all_part:R|B[<buckets>|A|S|R]:0|<degree>|A|D|
info_cubes:R|B[<buckets>|A|S|R]:0|<degree>|A|D|
[<owner>]<table>:R|B[<buckets>|A|S|R]:0|<degree>|A|D|
[<owner>]<[prefix]*[<suffix>]:R|B[<buckets>|A|S|R]:0|
<degree>|A|D|(<object_list>)|NO
```

Default value: `ALL:R:0`

Possible values:

- `R|B[<buckets>|A|S|R]` lets you specify the sampling type and number of buckets for histograms:
  - `R`: row sampling
  - `B`: block sampling
  - `<buckets>`: the number of buckets for histograms
  - `A`: auto
  - `S`: skew only
  - `R`: repeat

- `<degree>|0|1|A|D` lets you specify the degree of parallelism for DBMS_STATS:
  - `<degree>`: degree of parallelism
  - `0`: table default degree
  - `1`: no parallelism
  - `A`: auto-degree option
  - `D`: default-degree option

- `all_part`: only processes SAP partitioned tables
- `info_cubes`: only processes InfoCube tables

The default entry `ALL:R:0` activates the DBMS_STATS package for all selected tables and defines the standard sampling type and degree of parallelism.

The sampling type and degree of parallelism that you enter for the specified tables or table groups overrides the default sampling type and degree of parallelism.

The `NO: ANALYZE` statement is used to update statistics (not recommended in Oracle 10g or higher). This is the default method in Oracle 9i.

For more information, see SAP Note 424239.
UPDATE STATISTICS FOR INFOCUBES WITH BRCONNECT

InfoCube tables, which are used in SAP Business Intelligence (BI) and Advanced Planner and Optimizer (APO), require special handling during update statistics. Normally it is best to create statistics for such tables with so-called histograms. For more information, see SAP Note 129252.

Previously, statistics for the InfoCube tables could only be updated using the ABAP report SAP_ANALYZE_ALL_INFOCUBES.

With BRCONNECT, statistics for the InfoCube tables can be updated together with other tables in one run. In this case, the statistics for the InfoCube tables are always created with histograms.

For more information, see

- SAP Note 428212
UPDATE STATISTICS FOR INDIVIDUAL PARTITIONS

Previously, BRCONNECT only monitored the global statistics of partitioned tables. This meant that significant changes within a partition only resulted in the statistics being updated if the number of changes exceeded a threshold value, \texttt{stats\_change\_threshold}, in relation to the total number of records in the table. Therefore, changes within a partition were often ignored.

BRCONNECT can now check and update the statistics at partition level, if necessary. However, to use this feature, you must activate the monitoring attribute for all partitioned tables (standard in Oracle 10g). For more information, see SAP Note \texttt{408527} and \texttt{524341}. 
VALIDATING THE STRUCTURE OF ORACLE DICTIONARY TABLES

You can use BRCONNECT (that is, `brconnect -u / -c -f stats -v`) to validate the structures of tables belonging to an SAP database user (that is, `SAPR3/SAPSR3, SAP<SID>`).

In addition, Oracle dictionary tables are also automatically included in the validation if all SAP tables were selected, as in the following example:

`brconnect -u / -c -f stats -t all -v cascade`

Oracle dictionary tables can also be validated without validating SAP tables:

`brconnect -u / -c -f stats -t oradict_tab -v cascade`

For more information, see SAP Note 526021.
CREATING SPACE STATISTICS WITH BRCONNECT

You often require information about how the space is used in the database by tables and indexes. For example, this information is necessary for planning an upgrade, or when reorganizing or converting individual tables.

You can create space statistics for all database objects as follows:

```
brconnect -u / -c -f stats -t all -m +I -s P1
-f allsel,collect,method,precision,space -p 4
```

The options used have the following meaning:

- `-t all`: collects statistics for all database objects. You can specify subsets if required.
- `-m +I`: collects index space statistics by validating its structure. Without this option, only space statistics for tables are collected.
- `-s P1`: collects statistics with a minimum accuracy of analysis of 1% of the table records
- `-f allsel,collect,method,precision,space`: forces the functionality described above
- `-p 4`: accelerates update statistics by using parallel processing with four threads

The space statistics for tables are stored in the DBSTATTORA table and for indexes in the DBSTATIORA table.

You can retrieve the space statistics from these tables for further processing using SQL.

You can find summary information on space statistics at the end of the BRCONNECT log.

For more information, see SAP Note 554031.
BRCONNECT OPTION -FORCE WITH UPDATE STATISTICS

This section gives examples of how you can use the BRCONNECT -force options to override the internal rules for update statistics, described above.

⚠️
Only use these options in exceptional circumstances.

- Check and update statistics for all tables in tablespace PSAPBTAD, including pool and cluster tables (up to Oracle 9i) or for tables that have the ACTIVE flag in the DBSTATC table set to N
  brconnect -u / -c -f stats -t psapbtabd -f allsel

- Update statistics without check for all tables relevant to the application monitor
  brconnect -u / -c -f stats -t dbstatc_mon -f collect

- Update statistics for a pool table specified in the DBSTATC table – for example, to determine space usage – using method "estimate 10% rows"
  brconnect -u / -c -f stats -t atab -m EI -s P10 -f allsel,collect,method,sample,space
  BRCONNECT stores the space usage data in tables DBSTATTORA (for tables) and DBSTATIORA (for indexes). Finally, the statistics are immediately deleted (up to Oracle 9i).
  If you want to keep the statistics – for example, for test purposes – you must include the keep option in the command, as follows:
  brconnect -u / -c -f stats -t rfblg -m EI -s P10 -f allsel,collect,method,sample,space,keep

- Check and update statistics in two separate phases:
  a) First phase: Check statistics
     brconnect -u / -c -f stats -t all -f nocoll
  b) Second phase: Update statistics on tables identified in the first phase
     brconnect -u / -c -f stats -t all -f nocheck

- Delete damaging statistics (up to Oracle 9i)
  Pool and cluster tables and tables that have the ACTIVE flag set to N in the DBSTATC control table should not normally have statistics, since such statistics can negatively affect database performance. In the standard update statistics run, using brconnect -f stats -t all (see above), BRCONNECT checks whether such damaging statistics exist and deletes them if so.
  You can delete such damaging statistics immediately:
  brconnect -u / -c -f stats -t harmful -d
  To delete statistics for other tables as well (only for test purposes), you can use the option -f allsel:
  brconnect -u / -c -f stats -t sdbah,sdbad -d -f allsel
  See SAP Note 892296 for further values that you can set in the -f|--force option.
ADDITIONAL INFORMATION

SAP Library
You can find more information on Oracle database administration and the contents of this document in the SAP Library as follows:

1. Call up the SAP Help Portal at help.sap.com/nw73 → Application Help → SAP Library: English.
2. Choose SAP NetWeaver Library: Function-Oriented View → Database Administration → Database Administration for Oracle.
3. Choose one of the following:
   o SAP Database Guide: Oracle → Approach to Oracle DBA → Update Statistics
   o CCMS: Oracle → Update Statistics for the Cost-Based Optimizer in CCMS (Oracle)

You can also find these plus selected extracts from the SAP Library at: www.sdn.sap.com/irj/sdn/ora → SAP on Oracle Knowledge Center → SAP Documentation in Help Portal

SAP Notes
You can find SAP Notes at: service.sap.com/notes