Estimating Cardinality: Use of Jonathan Lewis CBO methodology

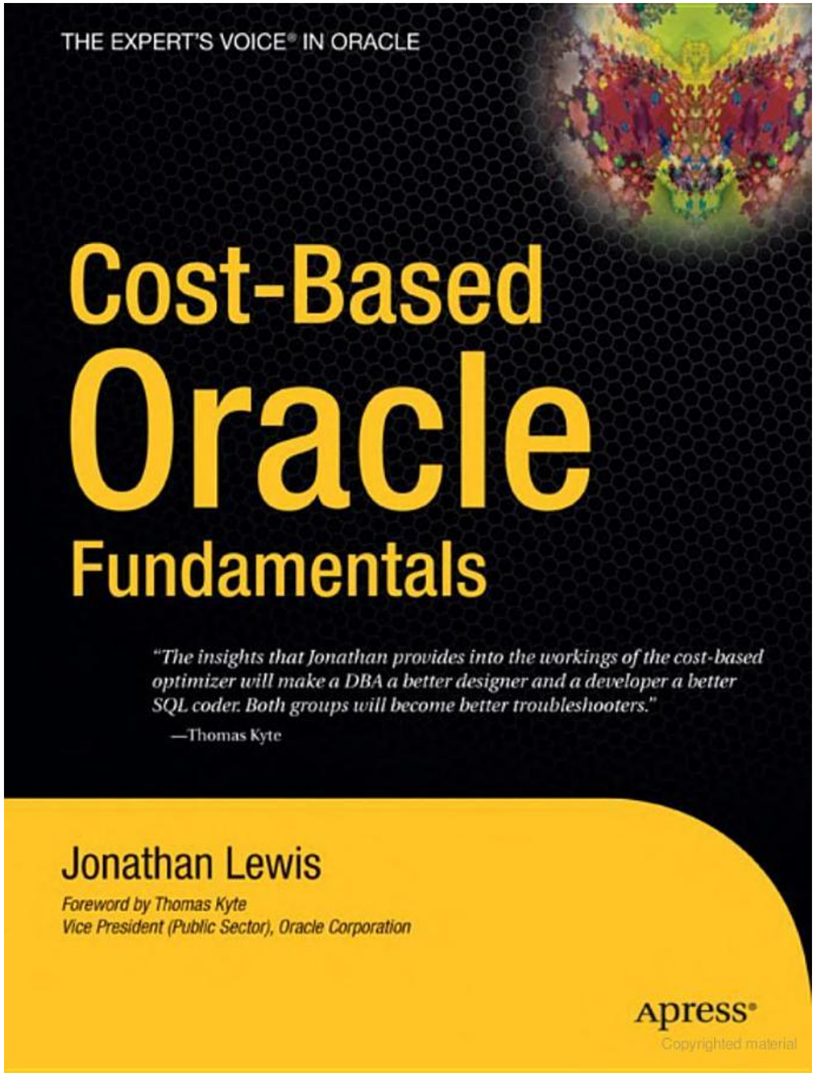
Dave Abercrombie
Principal Database Architect, Convio
NoCOUG Fall Conference 2010



© 2009 Convio, Inc.

Cost-Based Oracle Fundamentals

By Jonathan Lewis
ISBN13: 978-1-59059-636-4
ISBN10: 1-59059-636-6
536 pp.
Published Oct 2005
by Apress



© 2010 Convio, Inc.

How many red M&Ms?

Estimate total count

Count colors

Assume even distribution

Divide total by colors

(NoCOUG Physical Data Storage presentation on November 11, 2009)

© 2010 Convio, Inc.

Cardinality and Selectivity

- Selectivity = fraction of all rows matched by a predicate (or predicates)
- Cardinality = Selectivity * Rows

© 2010 Convio, Inc.

ClientX Group Members Query

```
-- 61zyzhy11juwf
select
   cons_id,
   member_id,
   active,
   first_name,
   last_name,
   primary_email,
   cat_id,
   admin_status
from
   clientX.constituent cons
   inner join clientX.group_members gm
      on gm.user_id = cons.cons_id
     and gm.group_id = :"SYS_B_0"
where active = :"SYS_B_1"
and cons.site_id = :"SYS_B_2"
and rownum <= :"SYS_B_3"
ORDER BY
   last_name_lower ASC,
   first_name_lower ASC;
```

© 2010 Convio, Inc.

Plan hash value: 990099801

```
---------------------------------------------------------------------
| Id  | Operation            | Name              | Rows  | Bytes |
---------------------------------------------------------------------
|   0 | SELECT STATEMENT     |                   |       |       |
|   1 |  SORT ORDER BY       |                   |  2001 |   170K|
|*  2 |   COUNT STOPKEY      |                   |       |       |
|*  3 |    HASH JOIN         |                   |   144K|    12M|
|*  4 |     INDEX RANGE SCAN | GROUP_USER_INDEX  |   144K|  1552K|
|*  5 |     TABLE ACCESS FULL| CONSTITUENT       |  4450K|   322M|
---------------------------------------------------------------------

Peeked Binds (identified by position):
--------------------------------------

   1 - :SYS_B_0 (NUMBER): 77569
   2 - :SYS_B_1 (NUMBER): 1
   3 - :SYS_B_2 (NUMBER): 1638
   4 - :SYS_B_3 (NUMBER): 2001

Predicate Information (identified by operation id):
---------------------------------------------------

   2 - filter(ROWNUM<=:SYS_B_3)
   3 - access("GM"."USER_ID"="CONS_ID")
   4 - access("GM"."GROUP_ID"=:SYS_B_0)
   5 - filter(("ACTIVE"=:SYS_B_1 AND "SITE_ID"=:SYS_B_2))
```

Plan narrative

- Scans GROUP_USER_INDEX for the GROUP_ID specified by bind variable.
- Reads index blocks, and constructs a hash table of USER_ID values that are in this group.
- Full scans CONSTITUENT, hashing each CONS_ID, feeding matching rows up to parent.
- Stops full scan after 2001 matches.
- Sorts the 2001 rows, passes them to Java

© 2010 Convio, Inc.

How many constituents in a group?

- Look at "Rows" column for Operation #4
  Oracle estimates 144,000 rows for access("GM"."GROUP_ID"=:SYS_B_0)

- Jonathan Lewis describes two methods (p. 43):
  1. cardinality = (rows in table) / (distinct values)
  2. cardinality = (rows in table) * (density)

- Which method gets used is not documented by Oracle, and can vary with version, etc.

- Method #2 is usually used with histograms, which do exist on this column

© 2010 Convio, Inc.

convio®

How many constituents in a group?

Value	Description
0.159%	DBA_TAB_COLUMNS.DENSITY of clientX.group_members.group_id
90,641,999	DBA_TABLES.NUM_ROWS (number of rows) in clientX.group_members
725	DBA_TAB_COLUMNS.NUM_DISTINCT (number of distinct values) in clientX.group_members.group_id
125,023	estimated cardinality of arbitrary group_id, **method 1, rows/distinct** (p. 43)
144,565	estimated cardinality of arbitrary group_id, **method 2, density * rows** (p. 43)

© 2010 Convio, Inc.

Method 2 was used

- Oracle used Method 2 (rows * density)
  144,565 equals the 144K in the execution plan

- In other words, the mean number of constituents in a group is about 144,000 (or 125,000, depending on estimation method).

- Two methods are usually equivalent.

- Watch out when DENSITY * NUM_DISTINCT is far from unity (here 0.159% * 725 = 1.15, OK)

© 2010 Convio, Inc.

Best way to get CONS data for 144K rows

- Oracle estimates 144K values of USER_ID.
- Need to find CONSTITUENT rows with matching CONS_ID, and that satisfy other predicates
- Two main choices:

  Full scan and do a hash join, or
  Nested loop into an index with table access by rowid

Full scan: multi-block

- Full scans are done with multi-block reads
- A multi-block read is more expensive than a single block read, but
- The cost per block is less when done via multi-block than when done one at a time.
- Based on system performance stats stored in sys.aux_stats$ (to which I had no access)

© 2010 Convio, Inc.

Multi-block read efficiency (Lewis, p. 20)

Metric	Value	Description (typical or actual)
IOSEEKTIN	10	Disk seek time in milliseconds (typical)
IOTFRSPEED	4096	Disk transfer time in bytes per millisecond (typical)
DB_BLOCK_SIZE	8192	Block size in bytes (actual)
DB_FILE_MULTIBLOCK_READ_COUNT	16	How many blocks are read in a multi-block read (actual)
SREADTIM	(tbd)	Single block read time in milliseconds (to be calculated)
MREADTIM	(tbd)	Multi-block read time (total) in milliseconds (to be calculated)

sreadtim = ioseektim + db_block_size/iotrfrspeed

mreadtim = ioseektim
+ db_file_multiblock_read_count *(db_block_size/iotrfrspeed)

© 2010 Convio, Inc.

Multi-block read efficiency (Lewis p. 20-1)

Value	Description
10	ioseektim (typical)
8192	db_block_size (actual)
4096	iotrfrspeed (typical)
16	db_file_multiblock_read_count (actual)
12	**sreadtim** = ioseektim + db_block_size/iotrfrspeed (Lewis, pp. 20)
42	**mreadtim** = ioseektim + (db_file_multiblock_read_count *(db_block_size/iotrfrspeed))
3.5	**mreadtim/sreadtim**

Full scan cost: clientX.constituent

Value	Description
608,252	number of blocks in clientX.CONSTITUENT
16	db_file_multiblock_read_count (actual)
38,016	multi-block-reads, number of IO requests to do full scan: (number of blocks / blocks per read)
3.5	mreadtim/sreadtim
133,055	equivalent number of single-block IO requests for full table scan access: multi-block-reads * (mreadtime/sreadtim)

A full scan costs 608,252 "gets", but costs only as much as 133,055 single-block gets

© 2010 Convio, Inc.

Nested loop cost

- Single-block reads
- Depends on:

  Availability of suitable CONS_ID index
  Number of rows per CONS_ID
  Cost of reading index (per CONS_ID)
  Cost of reading table blocks (per CONS_ID)
  Number of CONS_ID values to lookup

© 2010 Convio, Inc.

Index suitability

- No primary key **index** here!
- Only index leading with CONS_ID:
  ITOPS_CONS_ID_UNAME
- But this index is not unique - how many rows per CONS_ID?
- Oracle can estimate rows/CONS_ID by:
  Table does have a CONS_ID primary key
  Table stats (explained below)
  Index stats (explained below)

© 2010 Convio, Inc.

convio®

Index suitability – table stats

- From dba_tab_columns

  DENSITY of CONS_ID = 2.2151E-07

  NUM_DISTINCT of CONS_ID = 4,514,534

  NUM_ROWS of CONSTITUENT = 4,532,476

  computed cardinality method 1 (no histogram),
  num_rows/num_distinct, p 43 = **1.00397**

  computed cardinality method 2 (if histogram),
  num_rows * density, p 43 = **1.00399**

© 2010 Convio, Inc.

Index suitability – index stats

- From dba_indexes:

  DISTINCT_KEYS of ITOPS_CONS_ID_UNAME = 4,761,265

  NUM_ROWS of ITOPS_CONS_ID_UNAME = 4,761,265

  computed cardinality = NUM_ROWS/DISTINCT_KEYS = 1

- All three methods indicate that each CONS_ID has a single row, so a NESTED LOOP seems potentially feasible.

© 2010 Convio, Inc.

Index/table costs per CONS_ID

- BLEVEL clientX.ITOPS_CONS_ID_UNAME = 2

- 2 reads to get to leaf block
- 1 read to get leaf block
- 1 read to get table block (from index rowid)

- 4 reads per CONS_ID

© 2010 Convio, Inc.

Nested loop cost: clientX.constituent

Value	Description
144,565	Estimated CONS_ID values to lookup (from previous discussion)
4	Single-block reads per CONS_ID
578,258	Single-block reads to do all lookups

Previously shown: a full scan costs **608,252** "gets", but costs only as much as **133,055** single-block gets.

But a nested loop costs **578,258** single-block reads, so is much more expensive.

Note that this analysis depends heavily on the number of CONS_ID values to lookup.

© 2010 Convio, Inc.

CONVIO®

What about logical vs. physical?

- Optimizer ignores the difference between
  - blocks that are cached in the SGA, and
  - those that must be read from disk.

- Hard to account for this difference in a simple calculation, due to the variable nature of the cache. **Assumes all reads are physical.**

- However, this assumption is probably pretty good for large table scans, whose blocks are least likely to remain in the cache.

© 2010 Convio, Inc.

Just scratched the surface

- Study: range scans and clustering factor
- Study: ranges and high/low value
- Take home: **estimated row counts are key!**
- Cardinality Feedback

© 2010 Convio, Inc.

How many red M&Ms?

Estimate total count

Count colors

Assume even distribution

Divide total by colors

- Physical Data Storage
  November 11, 2009
- SQL Execution Plans, DBMS_XPLAN, and Cardinality Feedback
  May 20, 2010

© 2010 Convio, Inc.