hotsos
www.hotsos.com

GLOC
Great Lakes Oracle Conference

Journey to the center of the Cloud

A look at Oracle Data Blocks

Ric Van Dyke
Education Director
Hotsos Enterprises Ltd.

hotsos
www.hotsos.com

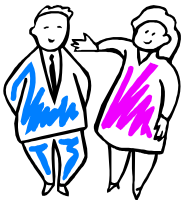
Overview

- The Data Block
- A Row of Data
- Updating Rows
- Block sizes
- Block storage parameters

hotsos
www.hotsos.com

Who is Ric?

- Oracle Ace
- Using Oracle since version 5
- Currently director of Education at Hotsos
- Prior experience:
  - Worked at Ford as a developer in Forms 2.3 and DBA
  - Worked at many companies as a consultant
  - Worked for Oracle for 10 years
    - Core DBA Senior Principal instructor
    - Education Manger, central region
    - ATS Technical Manger, north central region



hotsos
www.hotsos.com

Hotsos the Company

- **R**esponse-Time **P**rofiling **M**ethod **(RPM)**
  - Focus is on the business
  - Measures exact user experience
  - Trial and error tuning obsolete
- **Products**
  - HAWCS
  - Laredo

HAWCS

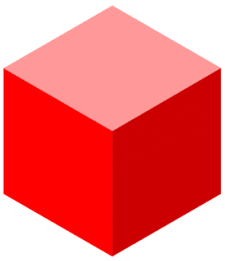
- **Services**
  - Performance assessment
  - DBA services
  - On-site / Remote consulting and education
- **Education**
  - Oracle performance curriculum
  - Hotsos Symposium

sixteenth annual mar 5 - 8 dallas, tx
hotsos symposium 2018
It's All About Oracle Performance.

2019 dates will be announced soon

hotsos
www.hotsos.com

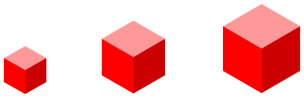
The Data Block



hotsos
www.hotsos.com

The most common block in your database

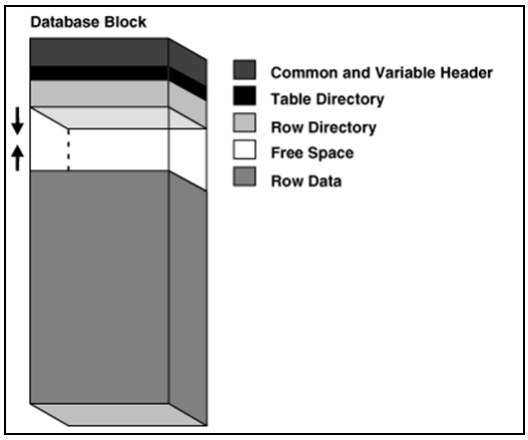
- Contains data, typically table data
- Default size is unchangeable
  - **`DB_BLOCK_SIZE`** initialization parameter
  - Default is typically 8K
  - Max: Windows 16K, Linux 32K
- Multiple block sizes
  - Tablespaces with different block sizes can exists in the same database
  - Intent is to support transportable tablespaces



hotsos
www.hotsos.com

Into the block

- The block overhead grows down.
- The row data grows up.
- Between the two is the free space in the block.



hotsos
www.hotsos.com

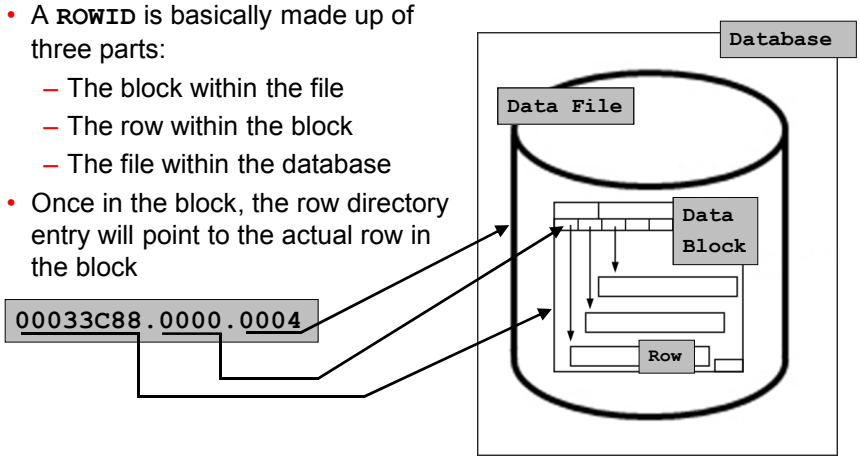
Data Block Overhead

- Over all size about 100 bytes
- Block header
  - General information about the block, including disk address and segment type
  - Also the transactional information
- Table directory
  - Information about the tables that have rows in this block
  - Normally only one table per block, a cluster can have multiple tables per block
- Row directory
  - These are the pointers to the rows in the block
  - The ROWIDs point to these entries

hotsos
www.hotsos.com

What is a ROWID?

- A ROWID is basically made up of three parts:
  - The block within the file
  - The row within the block
  - The file within the database
- Once in the block, the row directory entry will point to the actual row in the block



hotsos
www.hotsos.com

Using Oracles Block Dump

```
ALTER SYSTEM DUMP DATAFILE 6 BLOCK
MIN 12 BLOCK MAX 12;
```

- Use the alter system command to dump a block
- The above dumps block number 12 from file number 6
- Creates a dump file
- Use TRACEFILE_IDENTIFIER to help find the trace file

hotsos
www.hotsos.com

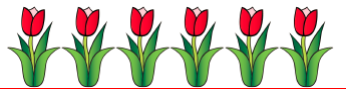
Example trace file

Display the file named below.

```
orcl_ora_14332_BLKTST.trc
```

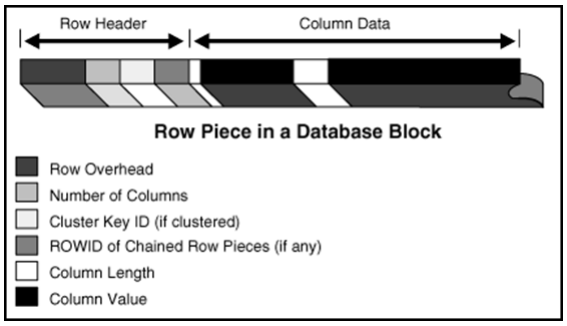

hotsos
www.hotsos.com

A Row Of Data



hotsos
www.hotsos.com

Into the row



- A row is variable length
- A row may be made have several pieces
  - Normally only one piece per block

hotsos
www.hotsos.com

Columns of data

- Each column has two parts:
  - The length
  - The data
- About Null column data
  - Between non-null data these columns just have a length
  - Trailing nulls are not stored

hotsos
www.hotsos.com

The Test Data

```
CREATE TABLESPACE BLOCKTEST DATAFILE
'C:\APP\ORACLE\ORADATA\ORCL\BLOCKTEST01.DBF' SIZE
10M REUSE;

CREATE TABLE BLOCKTEST1 (NAME VARCHAR2(12), ID
NUMBER, TEXT1 VARCHAR2(1000), TEXT2
varchar2(1000)) TABLESPACE BLOCKTEST;

INSERT INTO BLOCKTEST1 VALUES ('FIRST
ROW',101,NULL,'KNOW THYSELF');

INSERT INTO BLOCKTEST1 VALUES ('SECOND
ROW',102,NULL,NULL);

INSERT INTO BLOCKTEST1 VALUES ('THIRD
ROW',103,'ABCDEFGHIJKLMNOPQRSTUVWXYZ','NULL');
```


hotsos
www.hotsos.com

A set of rows in hex from a trace file

```
SQL> select * from blocktest1;
NAME                         ID TEXT1                          TEXT2
------------------------ ---- ---------------------------- ------------
FIRST ROW                 101                              KNOW THYSELF
SECOND ROW                102
THIRD ROW                 103 ABCDEFGHIJKLMNOPQRSTUVWXYZ NULL


-Address- ----------------Data----------------
001A1BF90 00000000 0014150A 012C0C08 48540904  [...........,...TH]
001A1BFA0 20445249 03574F52 1A0402C2 44434241  [IRD ROW.....ABCD]
001A1BFB0 48474645 4C4B4A49 504F4E4D 54535251  [EFGHIJKLMNOPQRST]
001A1BFC0 58575655 4E045A59 2C4C4C55 530A0201  [UVWXYZ.NULL,...S]
001A1BFD0 4E4F4345 4F522044 02C20357 04012C03  [ECOND ROW.....,.]
001A1BFE0 52494609 52205453 C203574F 0CFF0202  [.FIRST ROW.....]
001A1BFF0 574F4E4B 59485420 464C4553 B6B30605  [KNOW THYSELF....]
```

  orcl_ora_14332.trc

hotsos
www.hotsos.com

How the data is stored

- The data is stored in reverse order, sort of
- Rows go into the block from the bottom up
- This shows the hex translation of the bytes:

```
48474645 4C4B4A49 504F4E4D 54535251  [EFGHIJKLMNOPQRST]
 H G F E  L K J I  P O N M  T S R Q
```

hotsos
www.hotsos.com

Examining the First row

```
-Address- ----------------Data----------------
001A1BFD0 4E4F4345 4F522044 02C20357 04012C03  [ECOND ROW....,..]
           N O C E  O R   D    ,   W      ,
001A1BFE0 52494609 52205453 C203574F 0CFF0202  [.FIRST ROW......]
           R I F    R   T S  ,   W O
001A1BFF0 574F4E4B 59485420 464C4553 B6B30605 [KNOW THYSELF..A.]
           W O N K  Y H T    F L E S    A
```

Row over head:

2c – start of the row

01 – lock byte

04 – number of columns in this row

First column: 09 – length of column (9 bytes in BLUE)

Second column: 03 – length of column (3 bytes in GREEN)

Third column: FF – *NULL!* (in RED)

Forth column: 0C – length of column (12 bytes)

hotsos
www.hotsos.com

Length byte or bytes?

```
----------------Data---------------
52494609 52205453 C203574F 0CFF0202  [.FIRST ROW......]
574F4E4B 59485420 464C4553 8C410601  [KNOW THYSELF..A.]
```

In this version of the row, the 4rd column is "KNOW THYSELF" which is 12 characters and we have a single length byte 0C hex (12 decimal).

```
----------------Data---------------
46090401 54535249 574F5220 0202C203  [...FIRST ROW....]
03E8FEFF 7A78736C 4A4D464E 56445174  [....lsxzNFMJtQDV]
```

In this version of the row, the 4rd column is 1000 characters, notice there are now 3 bytes used for the length. The length is the 03E8 which is 1000 is decimal. The FE part tells Oracle the next two bytes are the length. This happens once the length is > 250.

NOTE: The second column is NULL in both, the FF makes it easy to find.

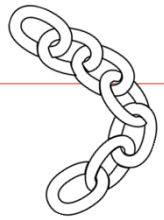
hotsos
www.hotsos.com

The "formatted" rows in the trace file

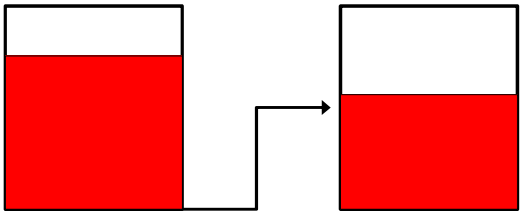
```
tab 0, row 0, @0x1f79
tl: 31 fb: --H-FL-- lb: 0x1  cc: 4
col  0: [ 9]  46 49 52 53 54 20 52 4f 57
col  1: [ 3]  c2 02 02
col  2: *NULL*
col  3: [12]  4b 4e 4f 57 20 54 48 59 53 45 4c 46
tab 0, row 1, @0x1f67
tl: 18 fb: --H-FL-- lb: 0x1  cc: 2
col  0: [10]  53 45 43 4f 4e 44 20 52 4f 57
col  1: [ 3]  c2 02 03
tab 0, row 2, @0x1f36
tl: 49 fb: --H-FL-- lb: 0x1  cc: 4
col  0: [ 9]  54 48 49 52 44 20 52 4f 57
col  1: [ 3]  c2 02 04
col  2: [26] 41 42 43 44 45 46 47 48 49 4a 4b 4c 4d 4e 4f 50 51
52 53 54 55 56 57 58 59 5a
col  3: [ 4]  4e 55 4c 4c
```

hotsos
www.hotsos.com

Chained Rows

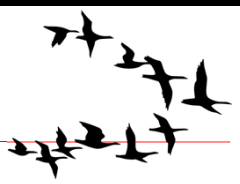


- A chained row doesn't fit in one block
- It has more then one piece
- Each piece will be in a block

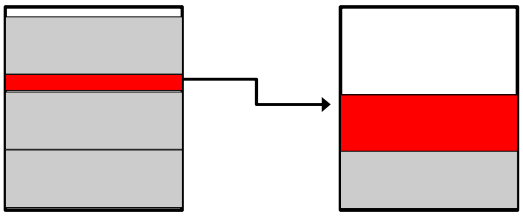


hotsos
www.hotsos.com

Migrated Rows



- Row fits in a block
- Has grown in size, can't stay in original block
- Entire row is moved, pointer left behind



hotsos
www.hotsos.com

Updating a row

hotsos
www.hotsos.com

Update the second row

```
SQL> UPDATE BLOCKTEST1 SET TEXT2='XXX' WHERE ID =
102;
```

New Row

```
016524180 2C800101 530A0402 4E4F4345 4F522044  [...,...SECOND RO]
016524190 02C20357 5803FF03 002C5858 48540904  [W......XXX,...TH]
0165241A0 20445249 03574F52 1A0402C2 44434241  [IRD ROW.....ABCD]
0165241B0 48474645 4C4B4A49 504F4E4D 54535251  [EFGHIJKLMNOPQRST]
0165241C0 58575655 4E045A59 2C4C4C55 530A0200  [UVWXYZ.NULL,...S]
0165241D0 4E4F4345 4F522044 02C20357 04002C03  [ECOND ROW....,..]
0165241E0 52494609 52205453 C203574F 0CFF0202  [.FIRST ROW......]
0165241F0 574F4E4B 59485420 464C4553 2E1E0601  [KNOW THYSELF....]
```

Original Row

hotsos
www.hotsos.com

Update the first row

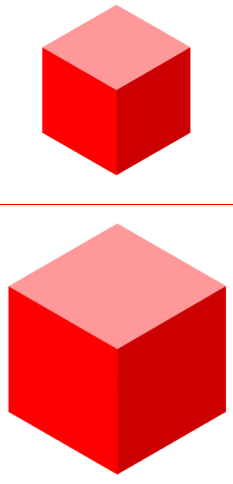
```
UPDATE BLOCKTEST1 SET TEXT2='THYSELF KNOW' WHERE
ID = 101;
```

```
015434180 2C800101 530A0402 4E4F4345 4F522044  [...,...SECOND RO]
015434190 02C20357 5803FF03 002C5858 48540904  [W......XXX,...TH]
0154341A0 20445249 03574F52 1A0402C2 44434241  [IRD ROW.....ABCD]
0154341B0 48474645 4C4B4A49 504F4E4D 54535251  [EFGHIJKLMNOPQRST]
0154341C0 58575655 4E045A59 2C4C4C55 530A0200  [UVWXYZ.NULL,...S]
0154341D0 4E4F4345 4F522044 02C20357 04012C03  [ECOND ROW....,..]
0154341E0 52494609 52205453 C203574F 0CFF0202  [.FIRST ROW......]
0154341F0 53594854 20464C45 574F4E4B 46130601  [THYSELF KNOW...F]
```

The changed column

hotsos
www.hotsos.com

Block Size



hotsos
www.hotsos.com

Bigger is better?

- General advice:
  - Smaller block size for transactional systems
  - Bigger for data ware house systems
- Smaller advantages:
  - Fewer rows per block
  - Ratio of ITL entries to rows better
  - Good for single row type work
- Bigger advantages
  - More rows per block
  - Shorter indexes
  - Good for full table scan type work

hotsos
www.hotsos.com

Compare Index scan with 4k, 8k and 16k block size

- Basic facts
  - Table (8K block size):
    - Number of Rows: 9,254,528
    - Number of Blocks: 131,840
    - Size in Meg: 1,030
  - Index:

	4K Block	8K Block	16K Block
Levels	4	3	2
leaf Blocks	197,987	95,873	47,177

- leaf blocks:
  - 4k to 8k : 102,114 less (reduction of 52%)
  - 8k to 16k : 48,696 less (reduction of 51%)
  - 4k to 16k : 150,810 less (reduction of 76%)

hotsos
www.hotsos.com

Index Scan - Run time Stats

	4K Block	8K Block	16K Block
Consistent Gets	14,017	13,904	13,846
Buffer Pinned Ct	12,065	12,065	12,065
Time*	22,172	22,297	22,165

- Consistent gets difference
  - 4k to 8k : 113 less (reduction of .8%)
  - 8k to 16k : 58 less (reduction of .4%)
  - 4k to 16k : 171 less (reduction of 1.2%)

***Time – a representative time from multiple runs in micro-seconds.**

hotsos
www.hotsos.com

Compare Full Table scan with 4k, 8k and 16k block size

- **4K block size:**
  - **Number of Rows: 9,254,784**
  - **Number of Blocks: 269,824**
- **8K block size:**
  - **Number of Rows: 9,254,784**
  - **Number of Blocks: 131,840 (51% reduction from 4K)**
- **16K block size:**
  - **Number of Rows: 9,254,784**
  - **Number of Blocks: 65,216 (50% reduction from 8K, 75% reduction from 4K)**

hotsos
www.hotsos.com

Full Table Scan - Run time Stats

	4K Block	8K Block	16K Block
Consistent Gets	534,811	261,354	121,494
Time*	11,590,478	15,411,618	14,280,009

- **`SELECT COUNT(*) FROM TABLE;`**
- Consistent gets difference
  - 4k to 8k : 273,457 less (reduction of 51%)
  - 8k to 16k : 139,860 less (reduction of 53%)
  - 4k to 16k : 422,317 less (reduction of 77%)

***Time – a representative time from multiple runs in micro-seconds.**

hotsos
www.hotsos.com

Block storage parameters

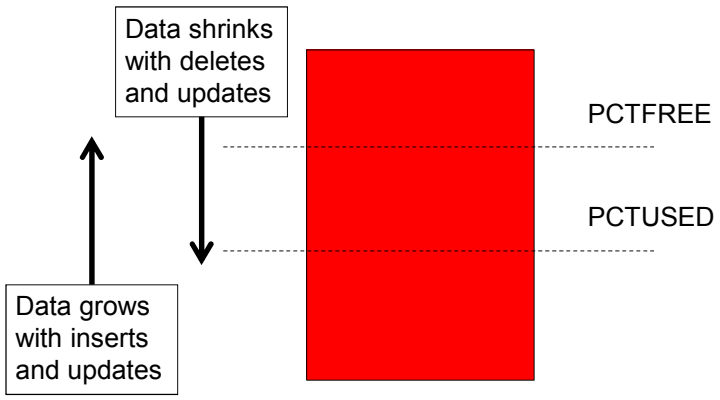
hotsos
www.hotsos.com

Space usage

- PCTFEE
  - Reserves space for future updates
  - When exceed the block goes off the free list
- PCTUSED
  - Used to put a block back on the free list
  - Not used with ASSM (Automatic Segment-Space Management)

hotsos
www.hotsos.com

PCTUSED and PCTFREE work together



hotsos
www.hotsos.com

Transactions

- **INITRANS**
  - Minimum number of concurrent transaction entries in a block (ITL slots)
  - Default is 1 for tables and 2 for indexes
- **MAXTRANS**
  - Obsolete as of 10
  - Is set to 255 for all blocks

hotsos
www.hotsos.com

Thanks!

- Questions?
- Comments?
- Confessions?

Ric Van Dyke
Hotsos Enterprises Ltd.
ric.van.dyke@hotsos.com