In Bed With Oracle
Lifting The Covers On Database Creation

Heavyweight Internet Group
Sean Hull
Why We Study Anatomy

- To understand creation
- To understand conception
- To make us better doctors
- To make us more responsible
- To make us more informed
Anatomy of Computer Systems

- Hardware level: build our own machine
- OS level: build kernel, compile software
- Perform Unix Systems Administration
- Database level: create a database, take it apart, put it back together
- Apprentice
- Generalized knowledge very helpful
Anatomy 101: Hardware

- Motherboard
- Cpu(s)
- Main memory
- Network, Video, Expansion cards
- Interfaces - USB, Serial, Firewire
- Disks or Disk Subsystem
Anatomy 101: Operating System

- Hardware bios, bootstrapping
- System kernel
  - Init process
  - System calls (fork, exec, read, write)
- Devices
- Device drivers
Anatomy 101: Operating System

- Virtual mem, paging, swapping
- Context switching
- Inter-process communication (signals)
- Filesystem
- Shell (command line interface)
Anatomy 101: Oracle Instance

Instance
- Unix processes
- Allocated memory
- Various Unix device files, sockets etc
- Can mount one database in lifetime
- Transient
Anatomy 101: Oracle Instance

*db_name=collab*

*SQL> startup nomount;*

[root@localhost root]# ipcs -m

<table>
<thead>
<tr>
<th>key</th>
<th>shmid</th>
<th>owner</th>
<th>perms</th>
<th>bytes</th>
<th>nattch</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x00000001</td>
<td>32768</td>
<td>root</td>
<td>600</td>
<td>655360</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>0x00000000</td>
<td>2162689</td>
<td>gdm</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>2195458</td>
<td>oracle</td>
<td>660</td>
<td>8388608</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>0x00000000</td>
<td>2228227</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>0x00000000</td>
<td>2260996</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>0x00000000</td>
<td>2293765</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>0x00000000</td>
<td>2326534</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>0x00000000</td>
<td>2359303</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>0x2abfaef0</td>
<td>2392072</td>
<td>oracle</td>
<td>660</td>
<td>10485760</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>
## Anatomy 101: Oracle Instance

```
[root@localhost root]# lsof -u oracle | wc -l
625
```

```
SQL> select name, value from v$parameter where isdefault = 'FALSE';
```

<table>
<thead>
<tr>
<th>NAME</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>shared_pool_size</td>
<td>100663296</td>
</tr>
<tr>
<td>background_dump_dest</td>
<td>/home/oracle/product/10.2.0/dbs/</td>
</tr>
<tr>
<td>db_name</td>
<td>collab</td>
</tr>
</tbody>
</table>

```
SQL> select * from v$sga;
```

<table>
<thead>
<tr>
<th>NAME</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Size</td>
<td>1218388</td>
</tr>
<tr>
<td>Variable Size</td>
<td>125831340</td>
</tr>
<tr>
<td>Database Buffers</td>
<td>50331648</td>
</tr>
<tr>
<td>Redo Buffers</td>
<td>7168000</td>
</tr>
</tbody>
</table>
Anatomy 101: Oracle Instance

```
SQL> select username, program from V$session;
USERNAME   PROGRAM
---------- ------------------------------------------------
oracle@localhost.localdomain (PMON)
oracle@localhost.localdomain (MMON)
oracle@localhost.localdomain (RECO)
oracle@localhost.localdomain (MMNL)
oracle@localhost.localdomain (CKPT)
oracle@localhost.localdomain (PSP0)
oracle@localhost.localdomain (LGWR)
oracle@localhost.localdomain (DBW0)
oracle@localhost.localdomain (MMAN)
oracle@localhost.localdomain (SMON)
SYS        sqlplus@localhost.localdomain (TNS V1-V3)
```

11 rows selected.
Anatomy 101: Oracle Database

Database
- Datafiles on disk
- Exists as a clone copy
- Exists as a backup copy
- Exists as a standby copy
- Persistent
Anatomy 101: Oracle Database

SQL> create database;

Wed Feb 21 00:59:43 2007
create database
Wed Feb 21 00:59:43 2007
WARNING: Default Temporary Tablespace not specified in CREATE DATABASE command
Default Temporary Tablespace will be necessary for a locally managed database in
future release
Wed Feb 21 00:59:43 2007
Database mounted in Exclusive Mode
Wed Feb 21 00:59:46 2007
Successful mount of redo thread 1, with mount id 3050081919
Assigning activation ID 3050081919 (0xb5cc8e7f)
Thread 1 opened at log sequence 1
Current log# 1 seq# 1 mem# 0: /home/oracle/product/10.2.0/dbs/log1collab.dbf
Successful open of redo thread 1
Wed Feb 21 00:59:46 2007
MTTR advisory is disabled because FAST_START_MTTR_TARGET is not set
Wed Feb 21 00:59:46 2007
SMON: enabling cache recovery
Wed Feb 21 00:59:46 2007
create tablespace SYSTEM datafile '?/dbs/dbs1@.dbf' SIZE 81920000 REUSE
default storage (initial 10K next 10K) EXTENT MANAGEMENT DICTIONARY online
Anatomy 101: Oracle Database

Wed Feb 21 00:59:48 2007
Completed: create tablespace SYSTEM datafile '*/dbs/dbs1@.dbf' SIZE 81920000 REUSE
default storage (initial 10K next 10K) EXTENT MANAGEMENT DICTIONARY online
Wed Feb 21 00:59:48 2007
create rollback segment SYSTEM tablespace SYSTEM
  storage (initial 50K next 50K)
Completed: create rollback segment SYSTEM tablespace SYSTEM
  storage (initial 50K next 50K)
Wed Feb 21 00:59:55 2007
create tablespace SYSAUX datafile '*/dbs/dbx1@.dbf' SIZE 40960000 REUSE
  EXTENT MANAGEMENT LOCAL SEGMENT SPACE MANAGEMENT AUTO online
Completed: create tablespace SYSAUX datafile '*/dbs/dbx1@.dbf' SIZE 40960000 REUSE
  EXTENT MANAGEMENT LOCAL SEGMENT SPACE MANAGEMENT AUTO online
Wed Feb 21 00:59:56 2007
ALTER DATABASE DEFAULT TABLESPACE SYSTEM
Completed: ALTER DATABASE DEFAULT TABLESPACE SYSTEM
Wed Feb 21 00:59:59 2007
SMON: enabling tx recovery
Wed Feb 21 01:00:00 2007
Threshold validation cannot be done before catproc is loaded.
replication_dependency_tracking turned off (no async multimaster replication found)
Starting background process QMNC
QMNC started with pid=13, OS id=1795
Wed Feb 21 01:00:00 2007
Completed: create database
Anatomy 101: Oracle Database

- Tablespaces created
- Sql.bsq has been run
- Default users created
- Default roles created
- $user$, $obj$, $tab$
Anatomy 101: Oracle Database

SQL> alter database open;
Database altered.

SQL> select count(*) from sys.obj$;
   COUNT(*)
----------
       885

SQL> select count(*) from sys.user$;
   COUNT(*)
----------
       13

SQL> select count(*) from sys.tab$;
   COUNT(*)
----------
       315
Anatomy 101: Oracle Database

SQL> @%/rdbms/admin/catalog.sql
SQL>%/rdbms/admin/catproc.sql

SQL> select count(*) from dba_objects;

<table>
<thead>
<tr>
<th>COUNT(*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3840</td>
</tr>
</tbody>
</table>

SQL> select count(*) from user_views;

<table>
<thead>
<tr>
<th>COUNT(*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1667</td>
</tr>
</tbody>
</table>

SQL> show user
USER is "SYS"
Anatomy 101: Oracle Database

SQL> alter database close;
Database altered.

SQL> alter database open;
alter database open
*
ERROR at line 1:
ORA-01531: a database already open by the instance

An instance may mount only one db in it’s lifetime
Anatomy 101: Oracle Recap

Instance is *transient*
- Exists only while procs + memory are allocated
- Can mount one database in it’s lifetime*

Database is *persistent*
- Lasts as long as datafiles exist, on disk, tape or other media

*In it’s lifetime, a database can of course be mounted by many instances, otherwise we’d be stuck after one shutdown. Furthermore, with RAC many instances can mount the same database at the same time.*
Biology Explained: OS Details

- Shared memory
- Datafiles + unix device files
- Processes
- CPU activity
- Disk I/O activity
- Network resources
Biology Explained: shared mem

[oracle@localhost oracle]$ ipcs -m

<table>
<thead>
<tr>
<th>key</th>
<th>shmid</th>
<th>owner</th>
<th>perms</th>
<th>bytes</th>
<th>nattch</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x00000000</td>
<td>98305</td>
<td>oracle</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>131074</td>
<td>oracle</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>163843</td>
<td>oracle</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>196612</td>
<td>oracle</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>229381</td>
<td>oracle</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>294919</td>
<td>oracle</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>327688</td>
<td>oracle</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>393222</td>
<td>oracle</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>753673</td>
<td>oracle</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>786442</td>
<td>oracle</td>
<td>600</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>819211</td>
<td>root</td>
<td>644</td>
<td>106496</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>1605644</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>12</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>163843</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>12</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>1671182</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>12</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>1703951</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>12</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>1736720</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>12</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>1769489</td>
<td>oracle</td>
<td>660</td>
<td>33554432</td>
<td>12</td>
<td>dest</td>
</tr>
<tr>
<td>0x2abfaef0</td>
<td>1802258</td>
<td>oracle</td>
<td>660</td>
<td>10485760</td>
<td>12</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>1933332</td>
<td>oracle</td>
<td>660</td>
<td>393216</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>1966101</td>
<td>oracle</td>
<td>660</td>
<td>12288</td>
<td>2</td>
<td>dest</td>
</tr>
<tr>
<td>0x00000000</td>
<td>1998870</td>
<td>oracle</td>
<td>660</td>
<td>12288</td>
<td>2</td>
<td>dest</td>
</tr>
</tbody>
</table>
```
[root@localhost root]# lsof -u oracle | grep ".dbf"
oracle  1775 oracle   16u   REG       8,17  6078464 15581194
/home/oracle/product/10.2.0/dbs/cntrlcollab.dbf
oracle  1775 oracle   17uW  REG       8,17  81928192 15581198
/home/oracle/product/10.2.0/dbs/dbs1collab.dbf
oracle  1775 oracle   18uW  REG       8,17  40968192 15581199
/home/oracle/product/10.2.0/dbs/dbx1collab.dbf
oracle  1777 oracle   16u   REG       8,17  6078464 15581194
/home/oracle/product/10.2.0/dbs/cntrlcollab.dbf
oracle  1777 oracle   17u   REG       8,17  52429312 15581195
/home/oracle/product/10.2.0/dbs/log1collab.dbf
oracle  1777 oracle   18u   REG       8,17  52429312 15581196
/home/oracle/product/10.2.0/dbs/log2collab.dbf
oracle  1779 oracle   16uW  REG       8,17  6078464 15581194
/home/oracle/product/10.2.0/dbs/cntrlcollab.dbf
oracle  1781 oracle   16u   REG       8,17  81928192 15581198
/home/oracle/product/10.2.0/dbs/dbs1collab.dbf
oracle  1781 oracle   17u   REG       8,17  40968192 15581199
/home/oracle/product/10.2.0/dbs/dbx1collab.dbf
```
Biology Explained: processes

[oracle@localhost oracle]$ ps auxw | grep collab
oracle  1769  0.0  0.6  275948  3260  ?  S    Feb21  1:44  ora_pmon_collab
oracle  1771  0.0  0.4  275308  2332  ?  S    Feb21  0:13  ora_psp0_collab
oracle  1773  0.0  0.5  275308  3004  ?  S    Feb21  0:14  ora_mman_collab
oracle  1775  0.0  3.2  291848  16048 ?  S    Feb21  0:20  ora_dbw0_collab
oracle  1777  0.0  1.4  290872  7068  ?  S    Feb21  0:18  ora_lgwr_collab
oracle  1779  0.0  0.7  275444  3708  ?  S    Feb21  1:15  ora_ckpt_collab
oracle  1781  0.0  4.2  275912  21216 ?  S    Feb21  0:09  ora_smon_collab
oracle  1783  0.0  0.8  275364  4268  ?  S    Feb21  0:00  ora_reco_collab
oracle  1785  0.0  1.8  275876  9444  ?  S    Feb21  0:24  ora_mmon_collab
oracle  1787  0.0  0.6  275316  3348  ?  S    Feb21  0:51  ora_mmnl_collab
oracle  1795  0.0  3.8  290188  19424 ?  S    Feb21  0:24  ora_qmnc_collab
oracle  1803  0.0  0.5  275344  2532  ?  S    Feb21  0:01  ora_q001_collab
collab07.txt
oracle  24446 0.0  1.1  11380  5772  pts/3 S    00:32 0:00  emacs -nw
collab07.txt
oracle  24721  0.0  0.1  3676  664  pts/4  S    01:10 0:00  grep collab
Biology Explained: CPU

```
[oracle@localhost oracle]$ vmstat 1 10
procs memory      swap          io system         cpu
r  b   swpd free   buff  cache   si so    bi    bo in    cs us sy id wa
2  0 118332  19956  61100  77628    0    0     1     8   24    10 26  0 32  0
1  0 118332  19956  61100  77628    0    0     0     0  160   355 52  1 47  0
2  0 118332  19956  61100  77628    0    0     0    16  104   326 53  2 45  0
1  0 118332  19956  61100  77628    0    0     0     0  159   355 55  0 44  0
1  0 118332  19956  61100  77628    0    0     0    44  107   326 55  0 45  0
2  0 118332  19956  61100  77628    0    0     0    16  160   355 55  0 44  0
1  0 118332  19956  61100  77628    0    0     0     0  103   318 55  0 45  0
3  0 118332  19956  61100  77628    0    0     0     0  159   357 55  1 44  0
1  0 118332  19956  61100  77628    0    0     0    16  106   325 55  0 45  0
1  0 118332  19956  61104  77628    0    0     0    32  164   360 55  0 45  0
```
Biology Explained: unix files

- Regular, directory, character device, and network device files

```
[root@localhost root]# lsof -u oracle | grep REG | wc -l
  456
[root@localhost root]# lsof -u oracle | grep DIR | wc -l
   34
[root@localhost root]# lsof -u oracle | grep CHR | wc -l
  125
[root@localhost root]# lsof -u oracle | grep IPv4 | wc -l
     8
```
Biology Explained: disk I/O

```
[root@localhost root]# iostat -d sda 1 5
Linux 2.4.21-32.Elsmp (localhost.localdomain)   02/27/2007

Device:            tps  Blk_read/s  Blk_wrtn/s  Blk_read  Blk_wrtn
sda 0.99         4.80        11.88    2507292    6202046

Device:            tps  Blk_read/s  Blk_wrtn/s  Blk_read  Blk_wrtn
sda 0.00         0.00         0.00          0        0

Device:            tps  Blk_read/s  Blk_wrtn/s  Blk_read  Blk_wrtn
sda 2.00         0.00        80.00          0        80

Device:            tps  Blk_read/s  Blk_wrtn/s  Blk_read  Blk_wrtn
sda 0.00         0.00         0.00          0        0

Device:            tps  Blk_read/s  Blk_wrtn/s  Blk_read  Blk_wrtn
sda 0.00         0.00         0.00          0        0
```
Biology Explained: net devices

[root@localhost root]# netstat -le | grep oracle

tcp 0 0 192.168.0.101:1521 *:* LISTEN
oracle 3750446
udp 0 0 localhost.localdomain:32792 *:*
oracle 9151
udp 0 0 localhost.localdomain:32797 *:*
oracle 3760895
unix 2 [ ACC ] STREAM LISTENING 3750442 /var/tmp/.oracle/sEXTPROC1
unix 2 [ ACC ] STREAM LISTENING 3750444 /var/tmp/.oracle/s#25067.1
unix 2 [ ACC ] STREAM LISTENING 3750447 /var/tmp/.oracle/s#25067.2
Biology Explained: Summary

- Behind scenes, OS doing lots of magic
- Managing resources
  - Allocation, deallocation, cleanup
- Context switching fairly
- Providing interface to filesystem
- Providing interface to devices
A More Sophisticated Example

Database name
System tablespace

connect / as sysdba;
startup nomount;
CREATE DATABASE "collab"
DATAFILE
'/u01/oradata/syscollab01.dbf' size 500M
EXTENT MANAGEMENT LOCAL
A More Sophisticated Example

- Undo tables pace
- Redolog files

UNDO TABLESPACE collabundo DATAFILE
  '/u01/oradata/undcollab01.dbf' size 250m
LOGFILE
  '/u02/oradata/rdocollab01.dbf' size 10M,
  '/u02/oradata/rdocollab02.dbf' size 10M
A More Sophisticated Example

- Character set
- Temp tablespace
- Archive/noarchive
- Miscellaneous options

CHARACTER SET "WE8ISO8859P1"
NATIONAL CHARACTER SET "UTF8"
DEFAULT TEMPORARY TABLESPACE mytemp TEMPFILE
'/u01/oradata/tmpcollab.dbf' SIZE 500M
NOARCHIVELOG
MAXDATAFILES = 1000,
MAXLOGFILES = 10;
A More Sophisticated Example

- Data dictionary views
- PL/SQL views

@?/rdbms/admin/catalog.sql
@?/rdbms/admin/catproc.sql
Anatomy by Analogy: Horse Racing

- Create database :: new racetrack
- Shutdown :: track is closed
- Startup nomount :: walk horse out of stable
- Startup mount :: saddle him, jockey preps in gates
Anatomy by Analogy: Horse Racing

*Startup (open) :: now we are racing*
Anatomy by Analogy: Horse Racing

- **Shutdown ::** announce that track will close when everyone leaves
- **Shutdown transactional ::** each horse leaves as it finishes current lap
- **Shutdown immediate ::** each horse begins leaves, does not complete lap
- **Shutdown abort ::** turn out the lights, worry about cleanup tomorrow
Anatomy by Analogy: Horse Racing

- Horses (instance) can run only one race (transient)
- Horses trained well (SQL Tuning) will win
- Racetrack (database) can handle one or more (RAC) horses
- Racetrack (database) outlasts the horses and races run on it (persistent)
Thanks

*Original inspiration from Ch2 of “Expert Oracle Database Architecture” by Tom Kyte*
Further Reading

- Operating Systems: Design and Implementation - Andrew Tanenbaum
- Essential System Administration - Aeleen Frisch
- Database System Concepts - Silberschatz, Korth, Sudarshan
Further Reading

- Oracle Database 10g - Kevin Loney
- Oracle db 10g: Insider Solutions - Kumar, Kanagaraj, Stroupe
- Expert Oracle Db Arch: 9i & 10g Prog Techniques & Solns - Thomas Kyte
Contact

Sean Hull
shull@iheavy.com

Heavyweight Internet Group

www.iheavy.com