Troy's Tuning Tips
Take 2

Troy Ligon

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

Have you ever argued with your San guy about Read/Write ms?

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

Top 5 Timed Events

Event	Waits	Time(s)	Avg Wait (ms)	% Total Call Time	Wait Class
db file scattered read	421,019	10,980	26	31.8	User I/O
db file sequential read	873,640	10,729	12	31.1	User I/O
CPU time		9,397		27.2	
enq: TM - contention	1,059	1,554	1,467	4.5	Application
control file sequential read	1,989,848	1,121	1	3.3	System I/O

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

Have you ever argued with your San guy about Read/Write ms?

Did your Unix Admin side with the San guy?

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

Have you ever argued with your San guy about Read/Write ms?

Did your Unix Admin side with the San guy?

iostat 300 5

I/O Numbers
O/S vs. Database

It's all about averaging!

io-wait
io-svctime
io-util

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

```
==> cat io-wait
iostat -x 1 2|tail -`echo \`iostat 1 1|wc -l\`"+4"|bc`|
   egrep 'Device|emcpower|dm-'|egrep 'Device|\.'|
   awk '{printf "%8s  %8s  %8s  %20s \n", $10,$11,$12,$1}'|
   grep -v ' 0.00      0.00      0.00'|sort -g

==> cat io-util
iostat -x 1 2|tail -`echo \`iostat 1 1|wc -l\`"+4"|bc`|
   egrep 'Device|emcpower|dm-'|egrep 'Device|\.'|
   awk '{printf "%8s  %8s  %8s  %20s \n", $12,$10,$11,$1}'|
   grep -v ' 0.00      0.00      0.00'|sort -g

==> cat io-svctime
iostat -x 1 2|tail -`echo \`iostat 1 1|wc -l\`"+4"|bc`|
   egrep 'Device|emcpower|dm-'|egrep 'Device|\.'|
   awk '{printf "%8s  %8s  %8s  %20s \n", $11,$10,$12,$1}'|
   grep -v ' 0.00      0.00      0.00'|sort -g
```

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

./io-wait

await (…)	svctm	%util	Device:
9.24	7.52	24.80	emcpoweras
9.50	9.50	1.90	emcpowerbk
9.50	9.50	1.90	emcpowerbv
10.00	10.00	5.00	emcpoweraf
10.11	6.25	96.30	emcpowerbl
12.00	12.00	3.60	emcpowert
13.00	13.00	9.10	emcpowerar
14.63	14.53	27.60	emcpoweraw
41.33	34.33	10.30	dm-8
41.33	34.33	10.30	emcpowerbn

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

./io-svctime

svctm (…)	await	%util	Device:
5.56	9.83	20.00	emcpowert
5.64	11.68	99.80	dm-19
6.67	6.67	2.00	emcpowerbv
7.00	7.60	17.50	emcpowerbs
7.37	7.78	44.20	emcpoweral
7.61	8.01	60.10	emcpoweraw
8.08	8.08	9.70	emcpowerao
8.73	8.73	26.20	emcpowerbp
9.33	9.09	30.80	emcpowerar
10.00	10.00	2.00	emcpowerbk
13.71	18.79	79.50	emcpowerav

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

./io-util

%util (…)	await	svctm	Device:
38.20	6.45	3.24	emcpowerac
39.20	11.53	11.53	emcpowerar
59.20	17.85	17.94	emcpowerav
81.40	1.13	1.16	emcpowerx
89.00	5.48	3.10	emcpowerbw
89.10	4.05	3.18	emcpowerax
96.50	13.36	7.42	emcpoweras
98.40	14.10	9.55	emcpoweral
99.60	3.61	1.16	dm-13
99.80	3.88	1.54	emcpowerbl
99.90	14.13	0.88	dm-19
100.00	5.86	0.32	dm-12

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

Then they start talking about:

IOPS

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

"Disks are about 150 iops"

"I put 36 metas in your LUN"

"You can get 4800 iops from that LUN"

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

sqlplus "/ as sysdba" @iops2

SQL*Plus: Release 10.2.0.4.0 - Production on Thu Oct 25 15:43:12 2012

Copyright (c) 1982, 2007, Oracle. All Rights Reserved.

Connected to:
Oracle Database 10g Enterprise Edition Release 10.2.0.4.0 - 64bit Production
With the Partitioning, OLAP, Data Mining and Real Application Testing options

END_TIME	Small Read IOPS	Small Write IOPS	Total Small IOPS	Small Read I/O%	Small Write I/O%	Large Read IOPS	Large Write IOPS	Total Large IOPS	Large Read I/O%	Large Write I/O%	Total Read MBPS	Total Written MBPS	Total MBPS
09-OCT-12 02.00.59.649 AM	299.044	267.453	566.497	52.788	47.212	83.59	56.205	139.795	59.795	40.205	39.21	11.481	50.691
09-OCT-12 03.00.08.266 AM	398.756	353.389	752.145	53.016	46.984	147.815	113.7	261.515	56.523	43.477	80.419	32.971	113.39
09-OCT-12 04.00.13.986 AM	893.105	534.254	1427.359	62.57	37.43	90.377	175.876	266.253	33.944	66.056	54.714	49.95	104.664
09-OCT-12 05.00.18.612 AM	986.289	474.722	1461.011	67.507	32.493	104.192	155.936	260.128	40.054	59.946	62.606	37.696	100.3
09-OCT-12 06.00.25.680 AM	797.444	404.013	1201.458	66.373	33.627	119.982	98.817	218.798	54.837	45.163	64.645	20.194	84.83

SUNCOAST ORACLE USERS' GROUP

I/O Numbers
O/S vs. Database

sqlplus "/ as sysdba" @iops

physical read IOPS Range	No Of Occurences
0 - 499	260
500 - 999	552
1000 - 1499	289
1500 - 1999	160
2000 - 2499	100
2500 - 2999	80
3000 - 3499	43
3500 - 3999	11
4000 - 4499	17
5000 - 5499	2

physical read IOPS Range	No Of Occurences
7500 - 7999	1

physical write IOPS Range	No Of Occurences
0 - 499	1,332
500 - 999	175
1000 - 1499	8

redo write IOPS Range	No Of Occurences
0 - 499	1,515

SUNCOAST ORACLE USERS' GROUP

When AWR Snaps Are Not Enough

It's all about averaging!

twl_sysstat_create.sql
twl_sysstat_capture.sql
twl_sysstat_report.sql

SUNCOAST ORACLE USERS' GROUP

When AWR Snaps Are Not Enough

```
==> cat  twl_sysstat_create.sql
drop   table twl_sysstat ;
create table twl_sysstat (
   snap           date
  ,phy_read       number
  ,phy_read_byte   number
  ,phy_read_IO     number
  ,phy_write      number
  ,phy_write_byte  number
  ,phy_write_IO    number
  ,redo_write      number
)
tablespace SYSAUX;
exit
```

SUNCOAST ORACLE USERS' GROUP

When AWR Snaps Are Not Enough

==> cat twl_sysstat_capture.sql

```
-- CMDLINE PARAMETERS:
--   1: number of samples to capture
--   2: sample duration
set verify off feedback off
begin
  for cnt in 1..&1
  loop
    insert into twl_sysstat (
      select sysdate
        ,(select value from v$sysstat where stat_id=(select stat_id
          from v$statname where name='physical reads' ))                phy_read
        ,(select value from v$sysstat where stat_id=(select stat_id
          from v$statname where name='physical read total bytes' ))     phy_read_byte
        ,(select value from v$sysstat where stat_id=(select stat_id
          from v$statname where name='physical read total IO requests' )) phy_read_IO
        ,(select value from v$sysstat where stat_id=(select stat_id
          from v$statname where name='physical writes'))                phy_write
        ,(select value from v$sysstat where stat_id=(select stat_id
          from v$statname where name='physical write total bytes'))     phy_write_byte
        ,(select value from v$sysstat where stat_id=(select stat_id
          from v$statname where name='physical write total IO requests')) phy_write_IO
        ,(select value from v$sysstat where stat_id=(select stat_id
          from v$statname where name='redo writes'                  )) redo_write
      from dual
    );
    if (cnt<&1) then
      dbms_lock.sleep(&2);
    end if;
  end loop;
end;
/
exit
```

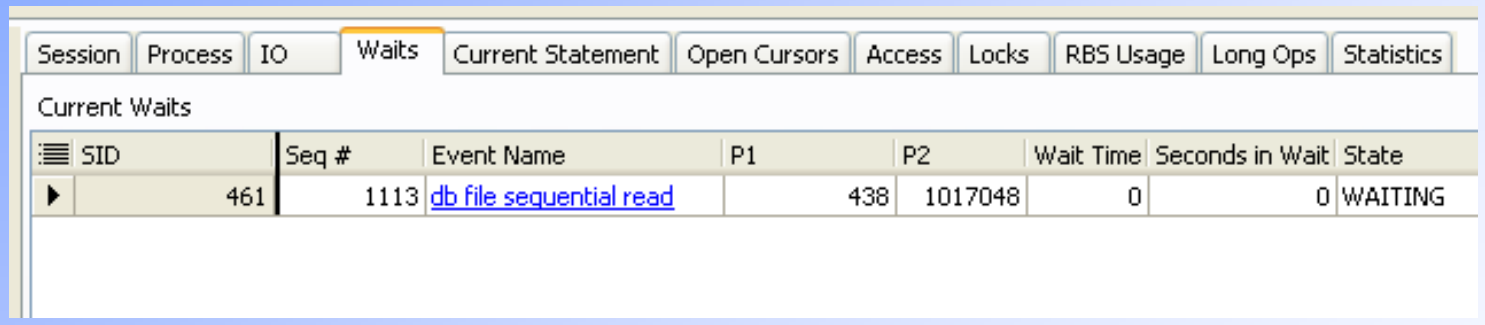
When AWR Snaps Are Not Enough

==> cat twl_sysstat_report.sql

```
set pagesize 50000 linesize 250 trimsp on
alter session set NLS_DATE_FORMAT='MON-DD HH24:MI:SS';
col REDO_IOPS format 999,999
col REDO_MBPS format   9,999.99
col WR_IOPS   format 999,999
col WR_PS     format 999,999
col WR_MBPS   format 999,999.99
col RD_IOPS   format 999,999
col RD_PS     format 999,999
col RD_MBPS   format 999,999.99
col SNAP      format a20     heading 'SAMPLE|START'
col SEC       format 999,999 heading 'SAMPLE|SECONDS'
select SNAP,SEC
  ,PHY_READ_IO/SEC  RD_IOPS,  PHY_READ/SEC  RD_PS,  PHY_READ_BYTE/1024/1024/SEC  RD_MBPS
  ,PHY_WRITE_IO/SEC WR_IOPS,  PHY_WRITE/SEC WR_PS,  PHY_WRITE_BYTE/1024/1024/SEC WR_MBPS
  ,REDO_WRITE/SEC REDO_IOPS,  REDO_WRITE*(select value from v$parameter where name='db_block_size')/1024/1024 REDO_MBPS
from (
select SNAP
  ,(SNAP-lag(SNAP) over(order by SNAP))*24*60*60  SEC
  ,PHY_READ-lag(PHY_READ)            over(order by SNAP) PHY_READ
  ,PHY_READ_BYTE-lag(PHY_READ_BYTE)   over(order by SNAP) PHY_READ_BYTE
  ,PHY_READ_IO-lag(PHY_READ_IO)       over(order by SNAP) PHY_READ_IO
  ,PHY_WRITE-lag(PHY_WRITE)          over(order by SNAP) PHY_WRITE
  ,PHY_WRITE_BYTE-lag(PHY_WRITE_BYTE) over(order by SNAP) PHY_WRITE_BYTE
  ,PHY_WRITE_IO-lag(PHY_WRITE_IO)     over(order by SNAP) PHY_WRITE_IO
  ,REDO_WRITE-lag(REDO_WRITE)        over(order by SNAP) REDO_WRITE
from twl_sysstat
)
order by 1;
```

What Am I Reading?

If you have Toad:



SUNCOAST ORACLE USERS' GROUP

What Am I Reading?

```
col event format a30
set linesize 250 trimsp on pagesize 50000

select sid,EVENT,P1,P2 from v$session_wait where WAIT_CLASS='User I/O';
```

SID	EVENT	P1	P2
385	db file scattered read	456	28617
388	db file sequential read	81	2050
404	db file scattered read	246	1053223
422	direct path read temp	523	399783
425	db file scattered read	29	973
430	db file sequential read	456	206339
433	db file scattered read	399	1534345
488	direct path read temp	538	458862

SUNCOAST ORACLE USERS' GROUP

What Am I Reading?

Db_file_sequential_read – Index read
Db_file_scattered_read – full table scan

P1 - File#
P2 - Block#

SUNCOAST ORACLE USERS' GROUP

What Am I Reading?

```
==> cat  fileblock.sql
set pagesize 50000 linesize 250 trimsp on
col object_name format a30
accept FILE_NUM  prompt " FILE_NUM? "
accept BLOCK_NUM prompt "BLOCK_NUM? "
select file_name from dba_data_files where file_id=&&FILE_NUM;
select
   a.object_type
  ,nvl(substr(a.object_name,1,30),'file=&&FILE_NUM  block=&&BLOCK_NUM') object_name
  ,a.subobject_name
from dba_objects a, x$bh b
where b.obj = a.data_object_id(+)
and b.file#(+) =&&FILE_NUM
and b.dbablk(+)=&&BLOCK_NUM
/
exit
```

SUNCOAST ORACLE USERS' GROUP

Troy's Tuning Tips
Take 2

Troy Ligon
tligon@soug.org

SUNCOAST ORACLE USERS' GROUP