AWR1page

Sanity checking time instrumentation in AWR reports

John Beresniewicz
ASHmaster

JB micro-bio

- 80's: DBA / Data modeler / Data Architect
- 90's: Startup co-founder / product designer / lead architect / PLSQL engineer
- 00's: 11-yr Oracle CMTS doing UX design and architecture for EM Grid Control and DB Diagnostic and Tuning Packs
- Recently released from a 2-yr sentence at Company X
- Available for hire or consult to collaborate on cool stuff

Acknowledgements

- Kevin Closson / Connor McDonald
  (for the motivating case studies)

- Graham Wood
  (for mentorship, advice and insight)

- Lothar Flatz / Wolfgang Breitling / Alberto Dell'Era /
  Toon Koppelaars
  (for test sample AWR reports)

Outline of talk

- Motivation: confusing AWR reports
- Model and instrumentation for time in Oracle
- When model and measures don’t match
- Concept and design of AWR1page
- Using AWR1page: case studies 1 & 2
- Final thoughts, future directions

Motivation: confounding AWR reports

- OakTable inquiries about AWR reports where "numbers don't add up"

- First case: double-counted DB time
  - *AWR Ambiguity* OakTable World 2015

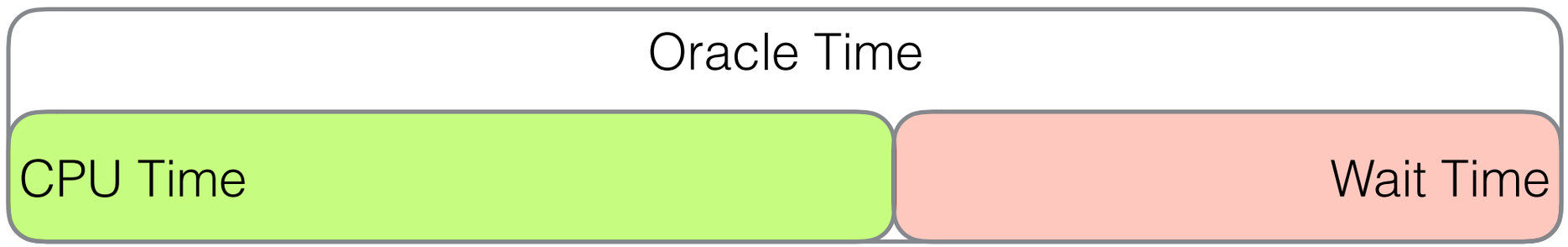
- Second case: AIX CPU reporting issues
  - *AWR1page* OakTable World 2016

- Solving such AWR report puzzles is mentally excruciating

AWR Ambiguity take-away slide

Symptom	*Possible issue*
DB CPU >> ASH CPU (and significant wait time)	CPU used within wait (this was the issue here)
ASH CPU >> DB CPU	System CPU-bound (ASH includes run-queue)
DB Time >> DB CPU + Wait (and not CPU-bound)	Un-instrumented wait (in call, not in wait, not on CPU)
DB Time >> ASH DB Time	1. Double-counted DB Time 2. ASH dropped samples

Model: Oracle Time = CPU + Wait



- Time spent executing Oracle code by either background or foreground processes

- Active processes are usually most interesting
  Active = (in DB call) && (on CPU or “active” wait)

- Multiple measure sources for each time component

Oracle time instrumentation

- Wait/Event Model: V$SYSTEM_EVENT
  FG/BG active/idle wait time

- Time Model: V$SYS_TIME_MODEL
  FG/BG CPU and Elapsed

- OS Timing: V$OSSTAT
  CPU sys/usr/IO/wait/load avg

- ASH: V$ACTIVE_SESSION_HISTORY
  FG/BG CPU and wait (estimated)

NOTE: Stop using V$SYSSTAT "CPU used by this session"

Time instrumentation essentials

- Run-queue time distorts everything!
- TM elapsed = (call end - call start) - idle wait time
- ASH ON CPU = active session not in wait (derived)
- TM CPU measured = actual instruction time used
- OSSTAT reliability platform-dependent
- Smaller call latencies increase distortion

Normalization: Avg Active Sessions

- Primary Oracle performance metric
- AAS = instrumentation time / elapsed time
- AAS / Core = core-normalized AAS

RMOUG 2007

Consistency checking AWR reports

- AWR reports are massive
  (most of it irrelevant for any specific situation)

- Time data scattered about, difficult to compare

- Units vary, breakdowns inconsistent, not normalized

- Labor intensive and error-prone

Cognitive load of consistency checking is huge

Sources of inconsistency

- CPU saturation: run queue time
  First and most important item to check

- CPU under wait event
  Time Model Wait << Wait ~ ASH Wait

- OS CPU accounting for multi-threaded cores
  IBM AIX issue: Time Model CPU << ASH CPU

- double-counted DB time
  10.1 bug with job queue process accounting

Sources of inconsistency (2)

- un-instrumented wait event
  ASH CPU >> Time Model CPU (OS CPU)

- active events classified idle
  Time Model, Wait and ASH under-reported

- dropped ASH samples*
  ASH << Time Model

- cpu-active idle events*
  FG Time << FG CPU Time

Check CPU first!

- OSSTAT: Average Active Sessions Busy / Core
- OSSTAT: Load Average / Core
- OSSTAT: OS_CPU_WAIT AAS (processes)
- Time Model: CPU / Core
- CPU Threads / Core = hyper-threading factor

Some consistency checks

? CPU saturation: Core utilization >> 1

? TM AAS ~ ASH AAS

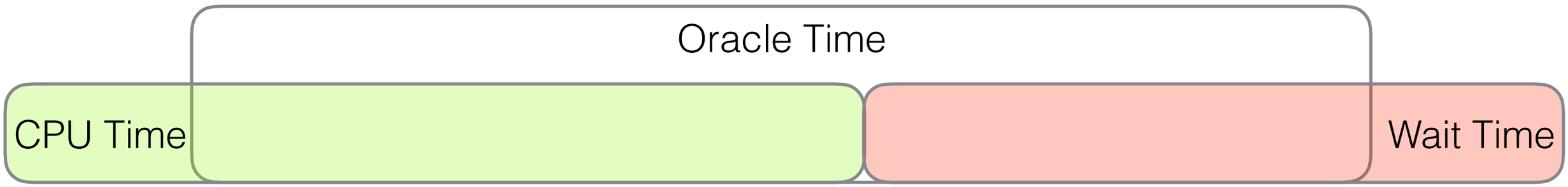
? TM CPU ~ ASH CPU ~ OS CPU

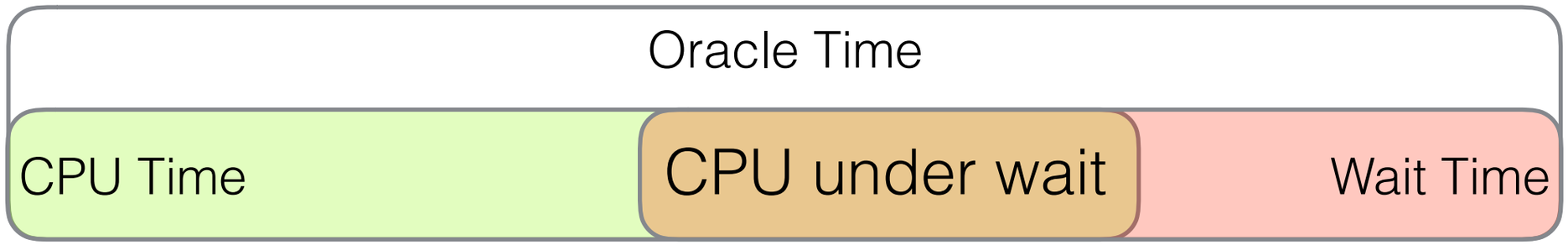
? TM AAS ~ (TM CPU + WT Wait)

? TM Wait ~ ASH Wait ~ WT Wait

KEY:
WT = Event/Wait Model
TM = Time Model
OS = OSSTAT
ASH = ASH

Oracle Time << CPU + Wait





- CPU under wait, time is being double-counted

  TM AAS < (TM CPU + WT Wait) &&
  ASH Wait ~ WT Wait

Oracle Time >> CPU + Wait

- Wait under-counted, un-instrumented wait?

  ASH CPU > TM CPU && TM CPU ~ OS CPU &&
  TM Wait > WT Wait

- CPU under-counted, AIX or course timers?

  ASH CPU > TM CPU &&
  ASH Wait ~ WT Wait && TM Wait > WT Wait

The AIX CPU issue

- AIX with SMT turned on reports CPU strangely
  - Semantics of CPU accounting altered
  - According to Graham: "*it's just broken*"
- OakTable blog references:
  Marcin Przepiorowski
  Jeremy Schneider
- Graham: Increase reported CPU values by 50% for SMT=4 for more realistic picture

Objectives for AWR1page

- Reduce cognitive load of checking time consistency in AWR reports: Total / CPU / Wait

- Facilitate comparative assessment of system sizes and busy-ness

- *In 1 page, using only AWR text report as input*

AWR report time measure sources

- $0 ~"^Operating System Statistics" {
- $0 ~ "^Time Model Statistics" {
- $0 ~ "^Foreground Wait Class" {
- $0 ~ "^Wait Classes by Total Wait Time" {
- $0 ~ "^Activity Over Time" {
- Also other important report singletons like elapsed time of AWR period, platform, core count, cpu count

Idea: measure x source matrix

data sources

measures

INFO	OSSTAT	TIME MODEL	ASH
HOST CPU SYS USR			
ORCL TOTAL CPU WAIT			
MISC			

Compare rows for equality

Columns add up heirarchically

Values from AWR report

INFO *	**OSTAT**	***TIME MODEL***	***ASH***
cores	**X**		
cpus	**X**		
threads/core	**X***		
memory	**X**		
ORCL .	**.**	**.**	**.**
total .	**.**	**X**	**X**
FG	**.**	**X**	**X**
BG		**X**	**X***
CPU load .	**X**	**.**	**.**
. busy	**X**	**.**	**.**
i/o wait	**X**	**.**	**.**
scheduler	**X**	**.**	**.**
ORCL .	**.**	**X**	**X**
FG	**.**	**X**	**X**
BG	**.**	**X**	**X***
WAIT			
MISC			

Design mockup

AWR1page	**cores: nn**	**cpus: nn**	**threads: nn**	**elapsed:nnnn.m**
measure	**OSSTAT**	**TIME MODEL**	**ASH**	**WAIT CLASS**
HOST LOAD	[nn:mm]			
CPU BUSY	[nn:mm]			
USER	nn			
SYS	nn			
IOWAIT	[nn:mm]			
OS_CPU_WAIT	[nn:mm]			
ORACLE				
AVG ACTIVE SESSIONS (BG & FG)		[nn:mm]	[nn:mm]	
CPU		[nn:mm]	[nn:mm]	
FG		nn		
BG		nn		
RES_MGR_CPU_WAIT	nn			
WAIT		[nn:mm]	[nn:mm]	[nn:mm]
FG		nn	nn	nn
BG		nn	dd*	dd*
Scheduler				nn
MISC: CPU per user call	n.mmmm	n.mmmm	n.mmmm	

Design notes

- AAS = Average Active [Sessions | Processes | Threads| Cores]
  - instrumentation time / elapsed time
- All numbers on the sparse matrix are in AAS
- Comparable measures from different sources on same line
- Indentation and vertical position represent decomposition
- Some cells show core-normalized AAS
  - [nn:mm] where nn=AAS, mm=AAS/cores

Version 2c: 634 lines of awk

core utilization

```
AWR1page

    file: awrKC.txt
platform: Linux x86 64-bit                                                                          cores: 36
     RAC: NO                                                                                         cpus: 72
 release: 12.1.0.2.0                                                                         threads/core: 2
 elapsed: 2.02 (min)                                                                       sessions(end): 56
   snaps: 1

========================================================================================================================
               measure                    |       OSSSTAT       |     TIME MODEL      |        ASH         |       WAIT CLASS
------------------------------------------------------------------------------------------------------------------------
HOST      LOAD                               [7.80:0.22]

               CPU BUSY                          [5.99:0.17]
                         USER                             2.27
                          SYS                             3.72

               IOWAIT                            [3.00:0.08]
               OS_CPU_WAIT                      N/A
........................................................................................................................
ORACLE
AVERAGE ACTIVE SESSIONS (BG+FG)                                       [4.98:0.14]         [4.58:0.13]

               CPU                                                        [2.02:0.06]         [0.50:0.01]
                    FG                                                             2.00
                    BG                                                             0.01
                    RSRC_MGR_CPU_WAIT                     N/A

               WAIT                                                       [2.96:0.08]         [4.08:0.11]         [4.32:0.12]
                    FG                                                             2.96                                 4.32
                    BG                                                             0.00                                *0.00
                    IOWAIT                                                                                          4.32
                          FG                                                                                            4.32
                    Scheduler                                                                                       0.00
........................................................................................................................
MISC
          CPU per call (ms)                             3781.98              10.50              312.50
             user calls/sec                1.58
```

case studies

Using AWR1page to diagnose timing inconsistencies

case study 1

Benchmarking Oracle I/O (a.k.a. being a “SLOB”)

```
AWR1page

    file: awrKC.txt
platform: Linux x86 64-bit                                                                                  cores: 36
     RAC: NO                                                                                                 cpus: 72
 release: 12.1.0.2.0                                                                                     threads/core: 2
 elapsed: 2.02 (min)                                                                                   sessions(end): 56
   snaps: 1


================================================================================================================================
               measure                    |      OSSTAT      |    TIME MODEL     |       ASH         |    WAIT CLASS
================================================================================================================================
HOST      LOAD                             [7.80:0.22]

               CPU BUSY                        [5.99:0.17]
                         USER                          2.27
                          SYS                          3.72

               IOWAIT                          [3.00:0.08]
               OS_CPU_WAIT                N/A
................................................................................................................................
ORACLE
AVERAGE ACTIVE SESSIONS (BG+FG)                                  [4.98:0.14]         [4.58:0.13]

               CPU                                                   [2.02:0.06]         [0.50:0.01]
                    FG                                                        2.00
                    BG                                                        0.01
                    RSRC_MGR_CPU_WAIT                 N/A

               WAIT                                                  [2.96:0.08]         [4.08:0.11]         [4.32:0.12]
                    FG                                                        2.96                                    4.32
                    BG                                                        0.00                                   *0.00
                    IOWAIT                                                                                         4.32
                          FG                                                                                           4.32
                    Scheduler                                                                                      0.00
................................................................................................................................
MISC
          CPU per call (ms)                        3781.98              10.50               312.50
             user calls/sec             1.58
```

Decent-sized Linux system running Oracle 12

```
AWR1page

    file: awrKC.txt                                                                                     cores: 36
platform: Linux x86 64-bit                                                                               cpus: 72
     RAC: NO                                                                                     threads/core: 2
 release: 12.1.0.2.0                                                                           sessions(end): 56
 elapsed: 2.02 (min)
   snaps: 1

========================================================================================================================
               measure                  |      OSSSTAT       |   TIME MODEL      |        ASH         |    WAIT CLASS
========================================================================================================================
HOST      LOAD                              [7.80:0.22]

               CPU BUSY                         [5.99:0.17]
                         USER                            2.27
                          SYS                            3.72

               IOWAIT                           [3.00:0.08]
               OS_CPU_WAIT                     N/A
........................................................................................................................
ORACLE
AVERAGE ACTIVE SESSIONS (BG+FG)                                    [4.98:0.14]         [4.58:0.13]

               CPU                                                     [2.02:0.06]         [0.50:0.01]
                    FG                                                          2.00
                    BG                                                          0.01
                    RSRC_MGR_CPU_WAIT                  N/A

               WAIT                                                    [2.96:0.08]         [4.08:0.11]         [4.32:0.12]
                    FG                                                          2.96                                  4.32
                    BG                                                          0.00                                 *0.00
                    IOWAIT                                                                                        4.32
                         FG                                                                                           4.32
                    Scheduler                                                                                     0.00
........................................................................................................................
MISC
          CPU per call (ms)                          3781.98           10.50             312.50
             user calls/sec            1.58
```

CPU ok

nontrivial IO wait

bedrock

IO wait = 1 aas CPU + 3 aas wait

DB Time < DB CPU + Wait => CPU under (IO) wait

case study 2

AskTom question about confusing AWR report

AWR1page

```
    file: QS02FNT_AWR_AskTom.txt
platform: AIX-Based Systems (64-bit)                                                    cores: 2
     RAC: NO                                                                             cpus: 8
 release: 12.1.0.2.0                                                                  threads/core: 4
 elapsed: 44.22 (min)                                                               sessions(end): 91
   snaps: 3


=================================================================================================================
               measure                    |      OSSTAT      |    TIME MODEL     |       ASH         |    WAIT CLASS
=================================================================================================================
HOST      LOAD                             [4.12:2.06]

               CPU BUSY                        [1.13:0.57]
                         USER                          0.66
                          SYS                          0.47

               IOWAIT                          [0.07:0.03]
               OS_CPU_WAIT                     [4.06:2.03]
.................................................................................................................
ORACLE
AVERAGE ACTIVE SESSIONS (BG+FG)                                  [0.62:0.31]         [0.72:0.36]

               CPU                                                   [0.13:0.07]         [0.66:0.33]
                    FG                                                        0.13
                    BG                                                        0.01
                    RSRC_MGR_CPU_WAIT                  0.00

               WAIT                                                  [0.48:0.24]         [0.06:0.03]         [0.05:0.03]
                    FG                                                        0.42                                    0.01
                    BG                                                        0.07                                   *0.05
                    IOWAIT                                                                                         0.03
                          FG                                                                                           0.00
                    Scheduler                                                                                      0.00
.................................................................................................................
MISC
          CPU per call (ms)                        115.47              0.01                60.02
             user calls/sec             9.80
```

AIX = CPU red fag

```
AWR1page

    file: QS02FNT_AWR_AskTom.txt
platform: AIX-Based Systems (64-bit)
     RAC: NO
 release: 12.1.0.2.0
 elapsed: 44.22 (min)
   snaps: 3
```

small system
SMT = 4

```
           cores: 2
            cpus: 8
  threads/core: 4
sessions(end): 91
```

```
================================================================================================================================
               measure                    |      OSSTAT      |    TIME MODEL     |       ASH         |    WAIT CLASS
================================================================================================================================
HOST      LOAD                              [4.12:2.06]

               CPU BUSY                                        [1.13:0.57]
                         USER                                          0.66
                          SYS                                          0.47

               IOWAIT                                          [0.07:0.03]
               OS_CPU_WAIT                                     [4.06:2.03]
--------------------------------------------------------------------------------------------------------------------------------
ORACLE
AVERAGE ACTIVE SESSIONS (BG+FG)                                [0.62:0.31]       [0.72:0.36]

               CPU                                0.56         [0.13:0.07]       [0.66:0.33]
                    FG                                                0.13
                    BG                                                0.01
                    RSRC_MGR_CPU_WAIT                   0.00

               WAIT                               0.06         [0.48:0.24]       [0.06:0.03]       [0.05:0.03]
                    FG                                                0.42                                0.01
                    BG                                                0.07                               *0.05
                    IOWAIT                                                                              0.03
                          FG                                                                                0.00
                    Scheduler                                                                           0.00
--------------------------------------------------------------------------------------------------------------------------------
MISC
          CPU per call (ms)                        115.47            0.01              60.02
             user calls/sec             9.80
```

?? → CPU bound? maybe

TM ~ ASH
ASH ~ Wait

DB Time > DB CPU + Wait => CPU under-report or run-queue

Final thoughts, future directions

- IT WORKS! Significant reduction of cognitive load to consistency check AWR report
- Re-write into Python, awk too fragile
- Accept html version of AWR report
- AWS Lambda service: AWR IN -> 1page OUT
- SQL version: check AWR directly
  (should be in DB regression tests)

contact info

- LinkedIn: https://www.linkedin.com/in/john-beresniewicz-986b0

- Gmail: john.beresniewicz@gmail.com

thank you

questions?