ORACLE®

ORACLE®

Designing and Tuning Large High Performance systems: Tips and Tricks

Graham Wood
Architect, Server Technologies

Designing and Tuning: Tips and Tricks

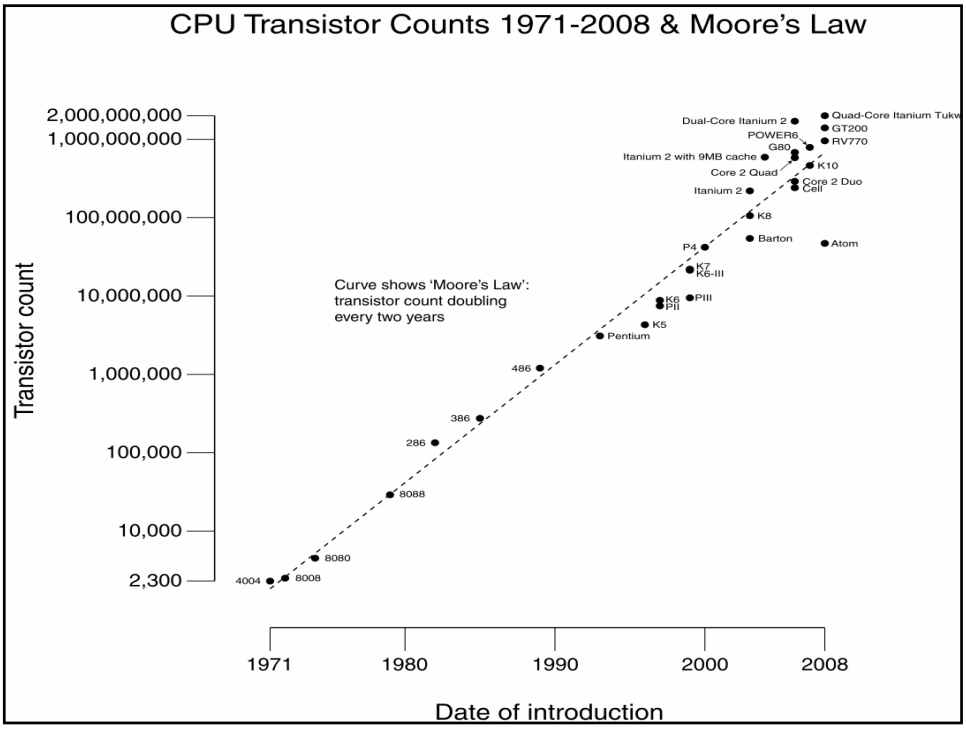
NOT

ORACLE

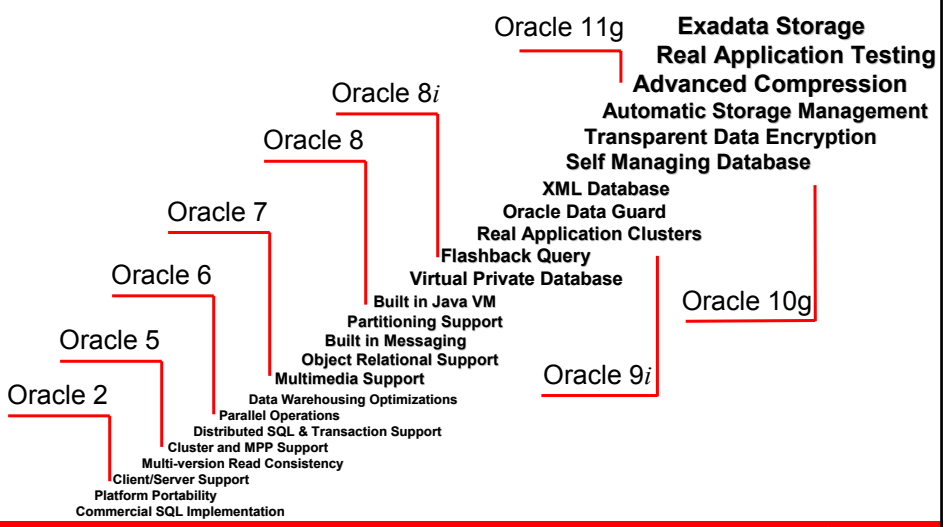
Designing for Oracle

- Changes over time
  - CPU
  - IO
  - Networks
- BUT Designing and Tuning, Large High Performance systems is NOT about tips and tricks, or fashion
- It is about good engineering practice

ORACLE



Continuous Innovation



Database Performance Basics
Schemas and SQL Statements

- Good schema/data design is the foundation to good database performance.
- Good schema/data design makes downstream activities both simpler and less error prone
  - Selection of Index and Partitioning Columns
  - Gathering and maintenance of schema statistics
  - Writing correct SQL statements
  - Optimization of SQL statements
  - Avoidance of serialization points and transactional bottlenecks

ORACLE

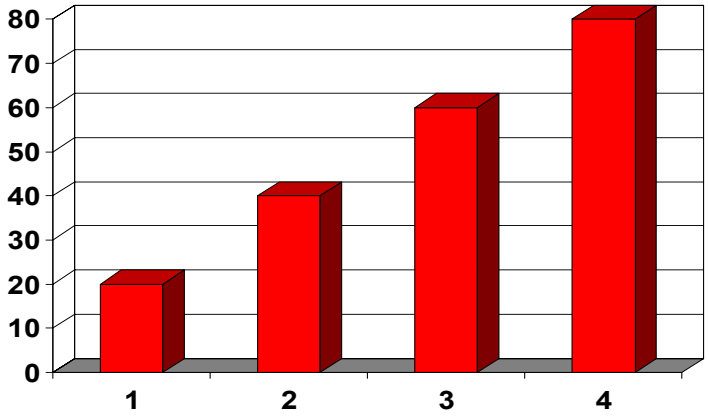
Scaling
What is it?

1. Linear
2. Unlimited
3. Maximum
4. Vertical
5. Horizontal
6. Unprecedented
7. Ultimate
8. Predictable
9. Diagonal

ORACLE

Scalability Quiz

Label the axes ?



ORACLE

Scaling

Some Possible Definitions

- When adding hardware resources to a system:
  - Number of transactions processed increases proportionately
  - Query times reduce proportionately
  - Query times remain constant when data volumes increase proportionately
- Hardware resources include
  - Nodes
  - CPU
  - Memory
  - Network
  - I/O

ORACLE

What Prevents Scaling

With respect to CPUs

- Scaling is severely impacted when the following scenarios arise
  - Hardware resources are operating beyond their practical maximum capacity for work
  - Software serialization takes place
    - Application — e.g. Row lock contention
    - Database — e.g. Latch contention
    - Operating System — e.g. Process allocation

ORACLE

Scaling the Oracle Database

In theory

- Requirements when scaling the Oracle Database
  - Effective hardware capacity planning
    - A **balanced** design
    - Hardware components sized to absorb peak workload
  - Effective application architecture
    - All software components tested and validated
    - Safe working limits identified

ORACLE

Scaling the Database
In reality

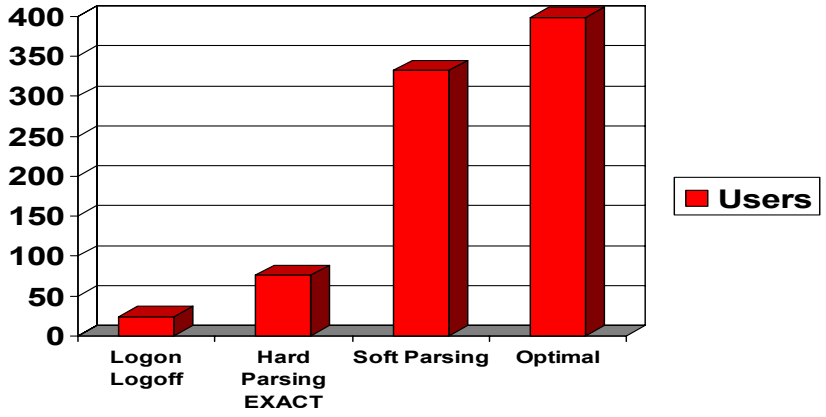
- Common Application Scaling Issues
- Single Instance Scaling Issues
- Multiple Instance Scaling Issues

Common Application Issues

- Incorrect use of Sessions
  - High connect/disconnect rates to the database
    - High load on O/S for process creation/teardown
    - High load mapping/unmapping SGA
    - CPU load on establishing database state
- Incorrect use of Cursors
  - Reparsing is best avoided
    - Hard parses use lots of CPU and will serialize
    - Soft parses scale better but they are not free

Database Performance Basics

Sessions & Cursors(This slide is over 10 Years old!)



ORACLE

Common Application Issues

- Poor SQL
  - Poor Access Paths
  - Excessive resource usage (CPU, I/O)
- Poor connection management
  - High numbers of connections can cause problems
  - Use connection pooling/concentration in the middle-tier
  - Use of Shared Servers is often just a band-aid

ORACLE

Common Configuration Issues

- Non-default initialization parameters
- I/O problems
  - Poor use of available resources resulting in hot spots
    - Bandwidth limitations
    - Poor response times
  - Use ASM to eliminate unpredictable hot spots
- Space management
  - Use Automatic Undo Management
  - Use Auto Segment Storage Management

ORACLE

Tuning Oracle

- Changes over time
  - Hardware
  - Software
  - Functionality
- BUT Designing and Tuning, Large High Performance systems is NOT about tips and tricks, or fashion
- It is about good engineering practice
- Correct diagnosis is key to successful tuning

ORACLE

Oracle Tuning Methods: A History

- Prehistoric (v5)
  - Debug code
- Dark Ages (v6)
  - Counters/Ratios
  - BSTAT/ESTAT
  - SQL*Trace
- Renaissance (v7/v8)
  - Introduction of Wait Event instrumentation
  - Move from counters to timers
  - STATSPACK
- Modernity (v10)
  - DB Time Tuning – Tuning using fundamental notion of time spent in database
  - Multiple scoping levels
  - Always on, non-intrusive
  - Built into infrastructure: instrumentation, ASH, AWR, ADDM, EM

Why Do We Care About Time?

- Human time is critical to the enterprise
- Systems performance affects business goals
  - Human time + technology resource time
- “Time is money”
- Performance improvement means doing things faster

Performance is always and only about time

Database Time and Average Active Sessions

ORACLE

Database Time (DB Time)

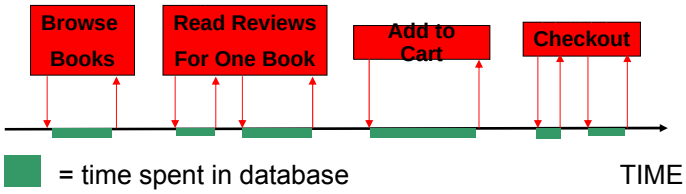
- Total time in database calls by foreground sessions
- Includes CPU time, IO time and non-idle wait time
- DB Time <> response time
- New lingua franca for Oracle performance analysis

Database time is total time spent by user processes either actively working or actively waiting in a database call.

ORACLE

A Single Session

Single session with Database Black Box server

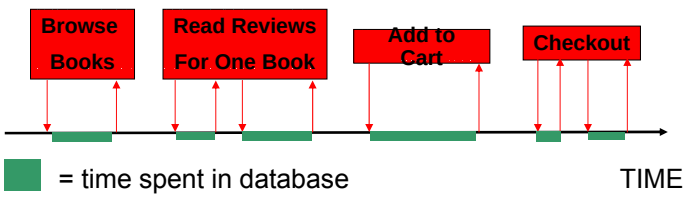


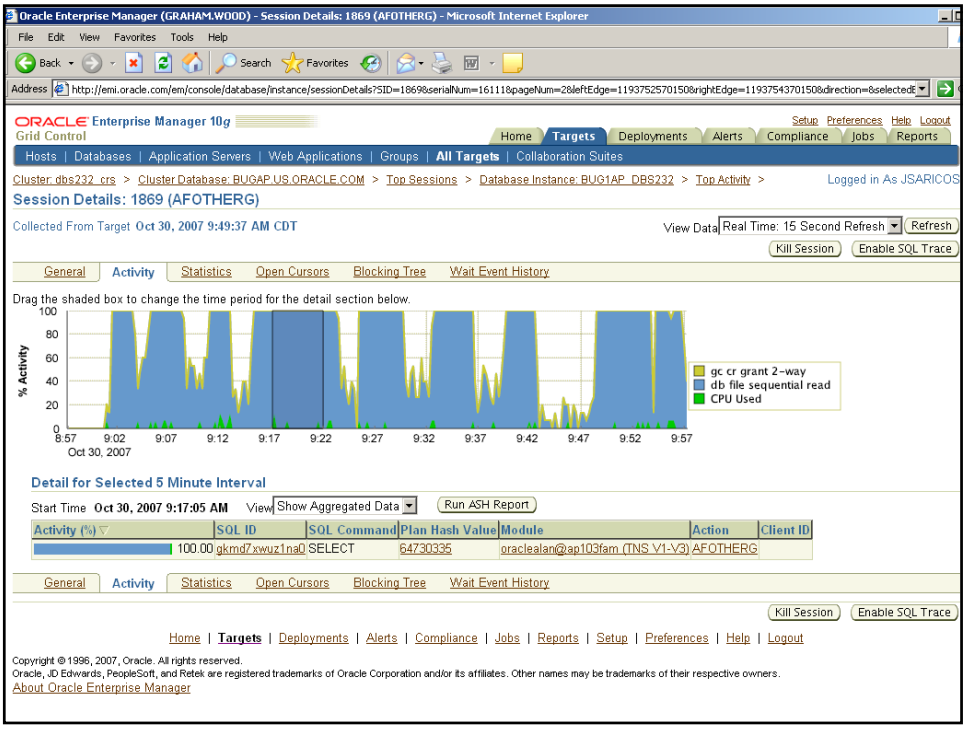
Fundamental Concepts

Database Time (DB Time) =
Total time session spent in all database calls

Active Session =
Session currently spending time in a database call

Average Activity of the Session (% Activity) =
The ratio of time active to total wall clock time



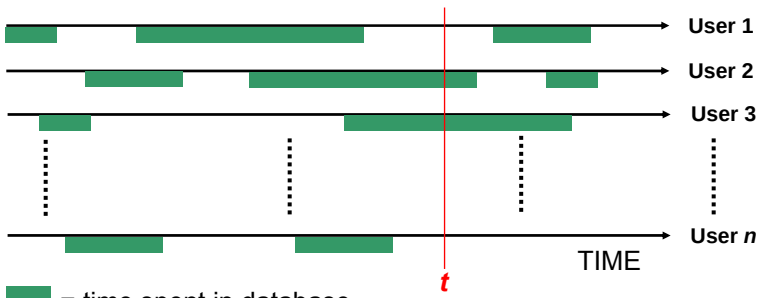


Multiple Sessions

DB Time = Sum of DB Time Over All Sessions

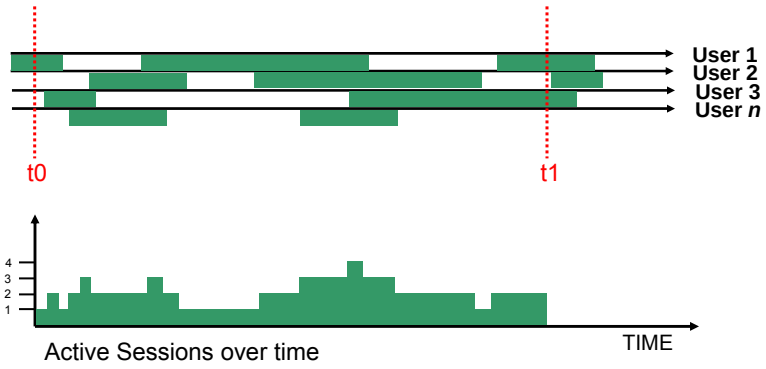
Avg. Active Sessions = Sum of Avg. Activity Over All Sessions

At time *t* we have 2 active sessions

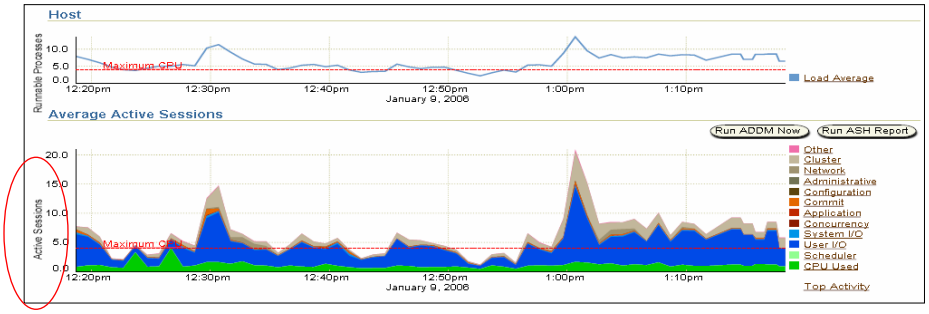


Visualizing DB Time

$$\text{Avg. Active Sessions} = \frac{\text{Total Database Time}}{\text{Wall Clock (Elapsed) Time}}$$



EM Performance Page



- Active Sessions by wait class over time
- Colored area = amount of DB time
- “Click on the big stuff”

Where to find DB Time?

- VSYS_TIME_MODEL, VSESS_TIME_MODEL
  - STAT_NAME = 'DB time'
- V$SYSMETRIC_HISTORY
  - "Database Time Per Second", "CPU Usage Per Sec"
  - 10g units = centi-secs/sec (100xAvg. Active Sessions)
  - 11g new metric "Average Active Sessions"
- V$SQL
  - ELAPSED_TIME and CPU_TIME
  - Wait class times:
    APPLICATION, CONCURRENCY, CLUSTER, USER_IO
- V$ACTIVE_SESSION_HISTORY

ORACLE

Active Session History

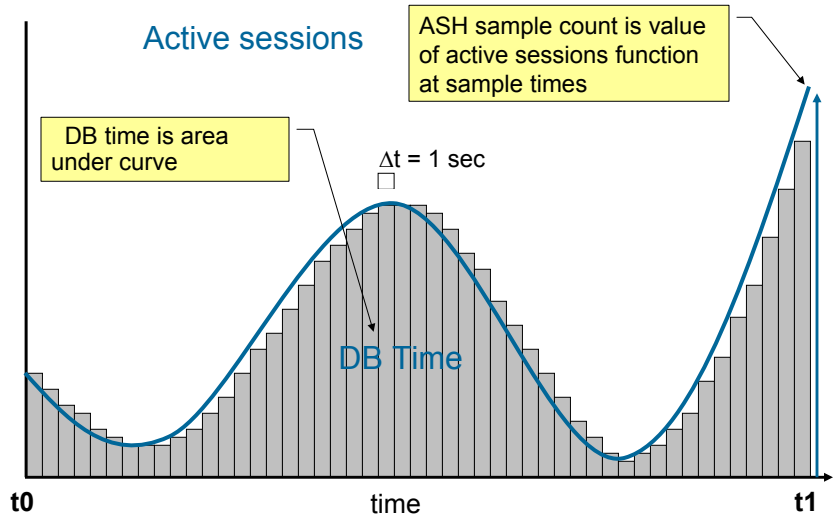
ORACLE

Active Session History (ASH)

- All ‘Active’ sessions captured every second
  - Foregrounds and backgrounds are sampled
  - Active foregrounds contribute to DB Time
- In-memory: V$ACTIVE_SESSION_HISTORY
  - Sampling interval = 1 second
- On-disk: DBA_HIST_ACTIVE_SESS_HISTORY
  - Sampling interval = 10 second
- ASH is a system-wide record of database activity

ORACLE

Active Sessions and DB Time



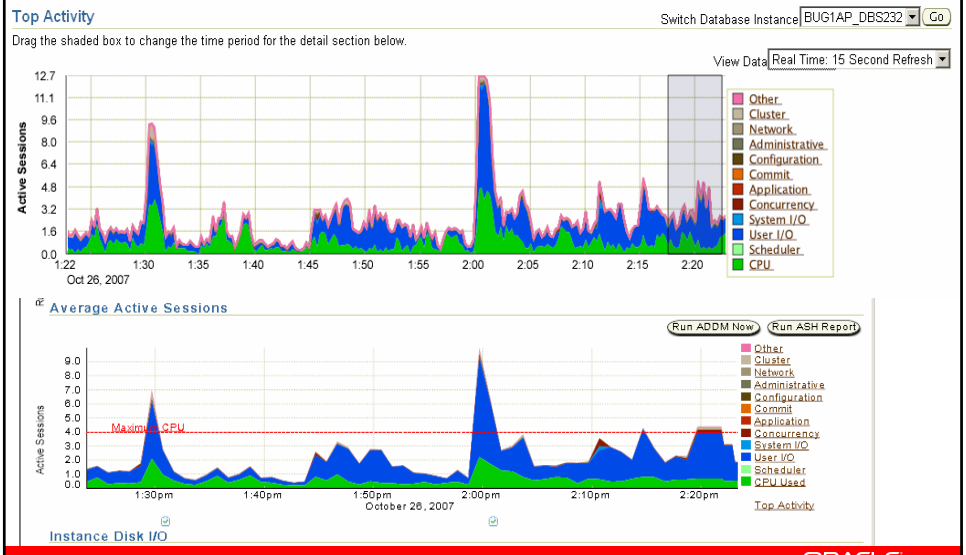
ORACLE

Estimating DB Time with ASH

- ASH sample counts = DB Time in seconds
  - Low sample counts are less reliable
- Enables DB Time analysis over many dimensions
  - Sqlid, session id, instance, service, module, action
  - 10gR2
    - Blocking_sid (10gR2)
    - XID
  - 11g
    - Row source
    - Execution ID
    - Operation type
      - Connect
      - Java/SQL/PLSQL
      - parse, bind, execute/fetch, close

ORACLE

DB Time: ASH vs Time Model



ORACLE

Where is DB Time used?

- ADDM
- EM Performance page and drill downs
- ASH report
- AWR and AWR compare periods reports
- SYSMETRICS and Server-generated Alerts

Techniques:

The DB Time Method

The DB Time Method: Short Course

or

just ask ADDM

ORACLE

The DB Time Method: Process

1. Identify performance issue
2. Scope the issue
3. Set goals
4. Data capture (NO OP)
5. Investigate DB time distribution
   - Identify the largest potential for improvement
6. Modify system to tune for largest gain
7. Evaluate against goals
   - Repeat from step 4 if goals not met

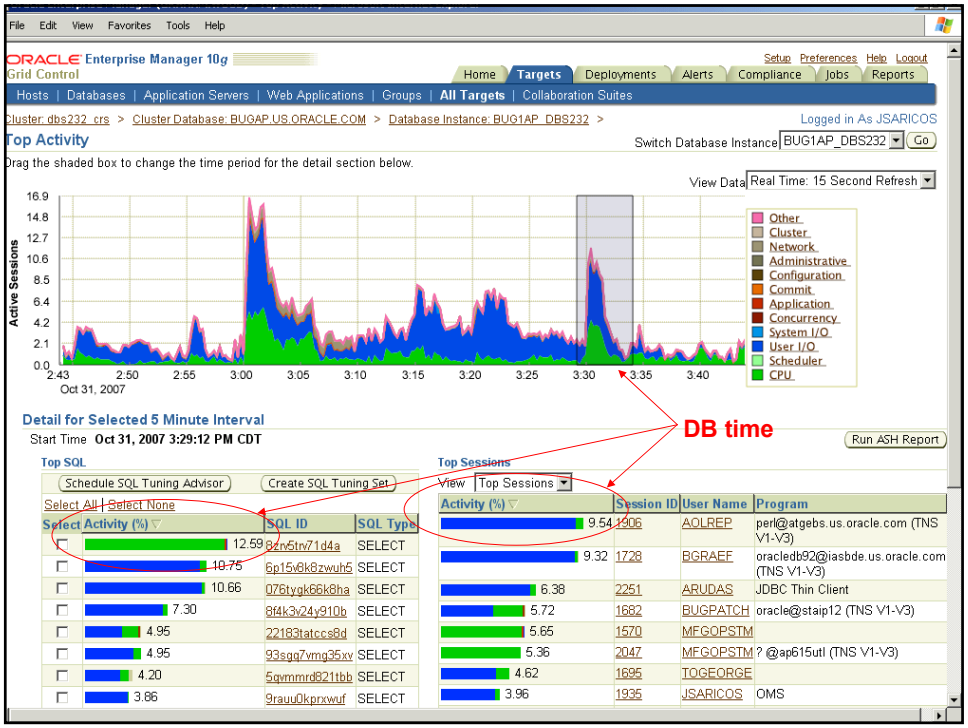
Performance tuning by removing excess DB time

ORACLE

Investigate DB Time Distribution

- Identify uneven distributions of DB time (skew)
  - => Largest potential improvement within scope
- System scope:
  - Resource limits – is problem outside the DB?
- Application scope:
  - Service, module, action
  - Resource contention (e.g. latches)
  - SQLID, rowsource
- Session scope:
  - Long running SQL
  - Resource contention (e.g. enqueues)

ORACLE



Identify Potential Solutions

- Session contention issues
  - Kill session
  - Fix application
- SQL issues
  - SQL Tuning Advisor => Indexes, SQL profile
  - Re-write SQL
- Design issues
  - Access Advisor => Indexes, physical layout
- System issues
  - Initialization parameters
  - Add resources

ORACLE

Modify System

- Start with the largest DB time issues first
  - Address root causes, not symptoms
- Match solution scope to problem scope
  - Don’t tweak optimizer parameters before tuning SQL
- Proceed iteratively one fix at a time
  - Concurrent fixes should be orthogonal
- Measure and validate results at each successive step
- Stop when goals are met

ORACLE

The DB Time Method: Advantages

- Tunes the one thing that affects users: Time
- Data capture scoping not necessary
  - 'Always on' data collection
  - No requirement to reproduce problem
- Works for concurrency problems such as locking
- Combines best of current methods
  - Less intrusive, more inclusive

ORACLE

Method Summary

- DB time is the fundamental performance metric
- The method allows DB time analysis at many scopes
  - Proper scoping of problems and solutions is critical to success
- DB time based diagnosis removes value judgments
  - Scientific method, not sorcerer's magic
- Performance improvement means doing the same work in less DB Time

ORACLE

Tools:

ADDM

Enterprise Manager

Reports

ORACLE

Tools for Applying DB Time Method

Two use-cases, one method:

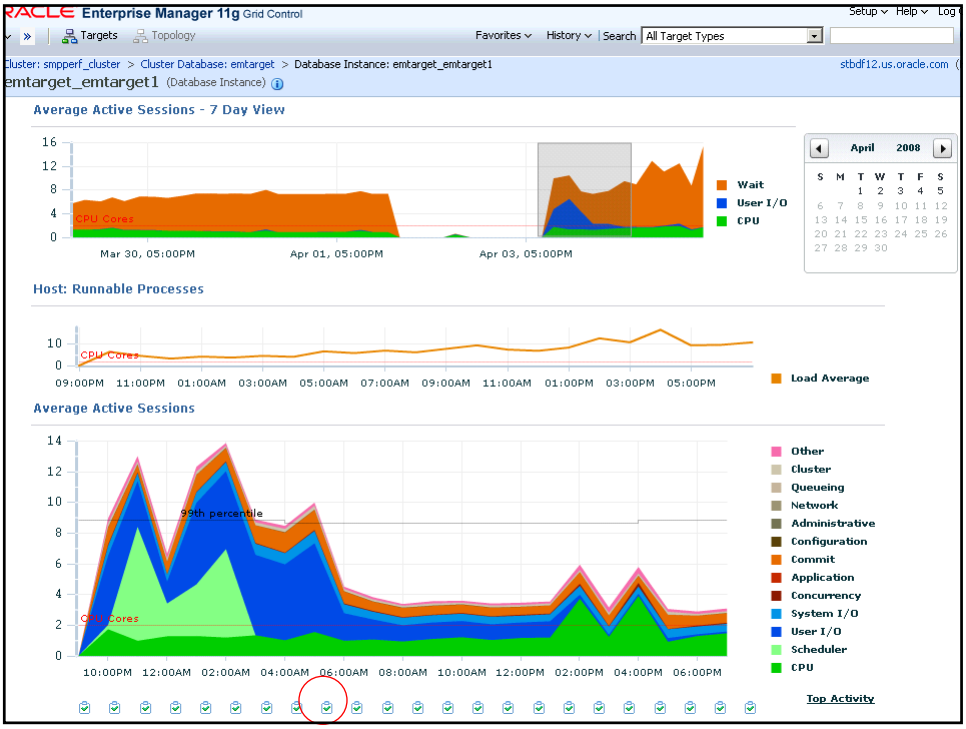
1. Tuning steady-state performance
   - Improve overall workload throughput or response time
   - Best practice: use ADDM
2. Diagnosing transient performance problems
   - Confirm and investigate reported performance issues
   - Best practice: use EM real-time screens

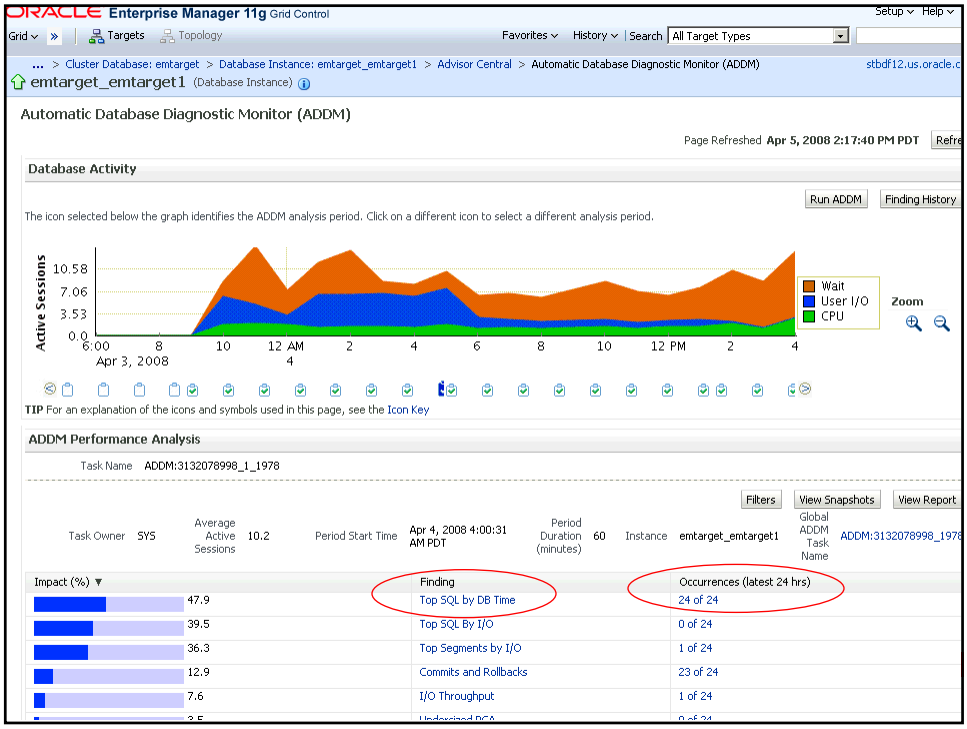
ORACLE

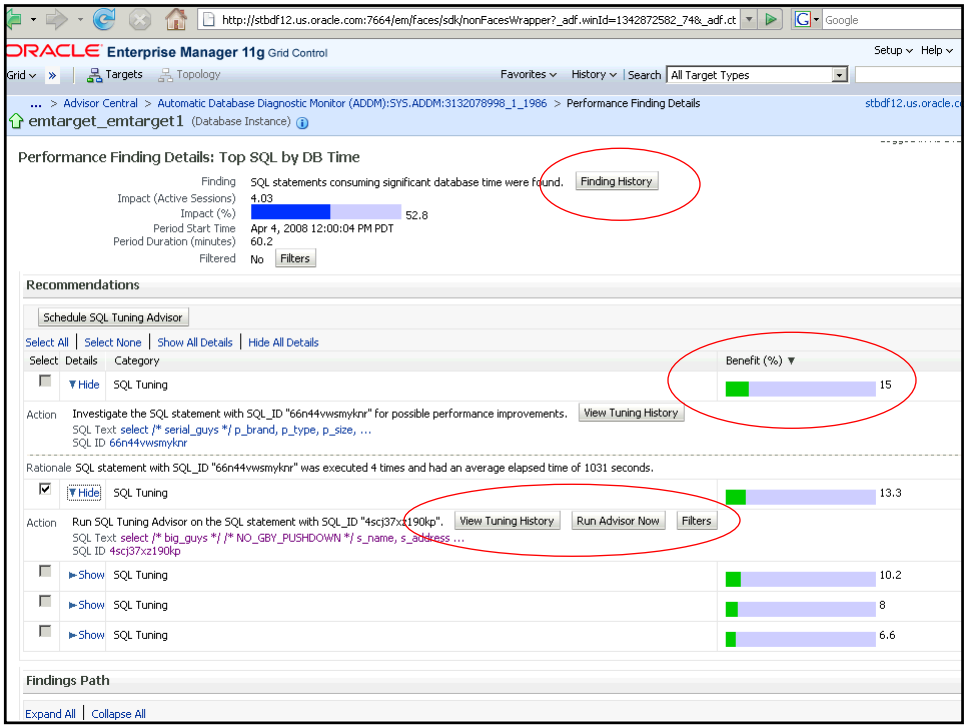
Best Practice: Use ADDM

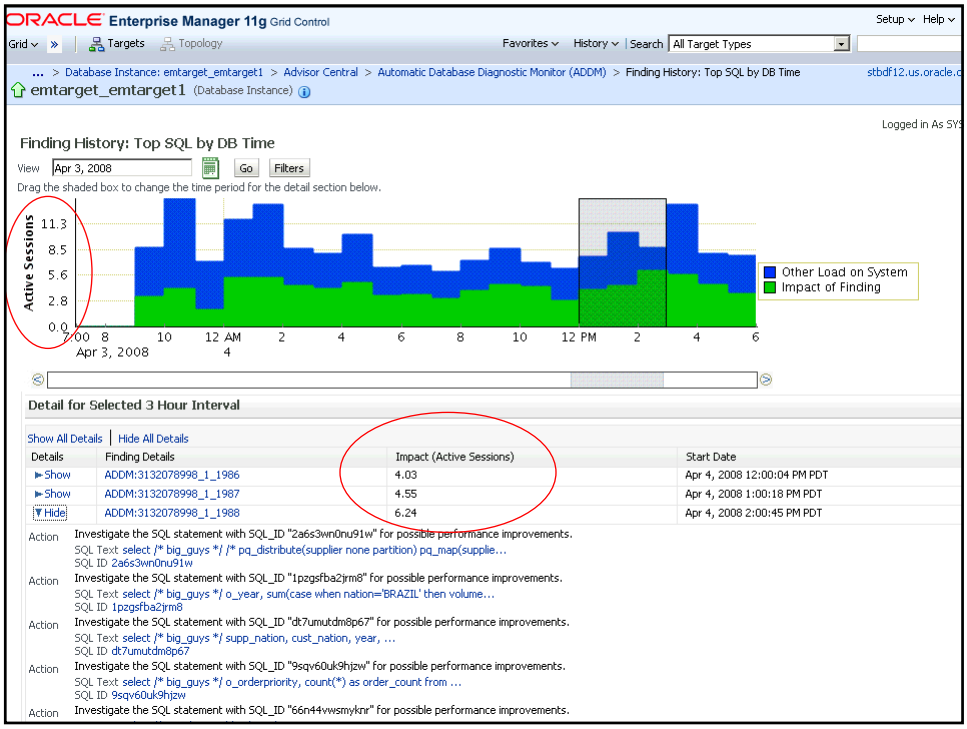
- Embedded expert system using the DB time method
  - Identifies root causes behind the symptoms
- Variably scoped:
  - Host to instance to SQL and even database block
  - Scoped to database for RAC (new in 11g)
- Findings prioritized by impact on DB time
  - Finding history allows flexible time scoping
  - Directives can filter findings
- Recommendations by benefit (reduction) to DB time

ORACLE



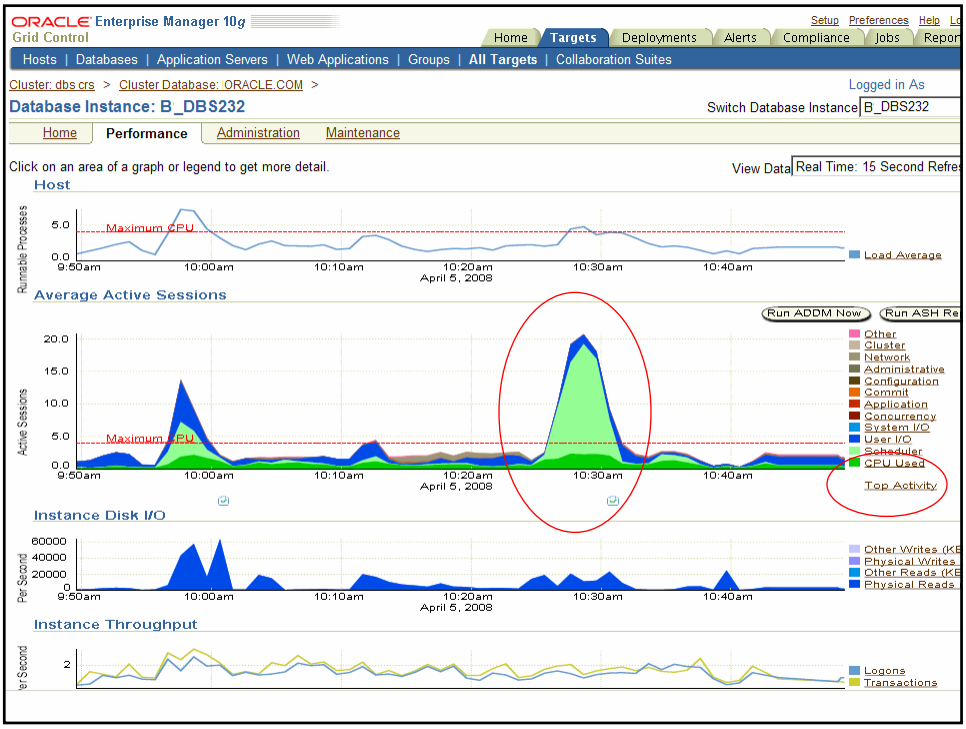

Enterprise Manager 11g Grid Control
Setup Help
Grid Targets Topology Favorites History Search All Target Types
... > Cluster Database: emtarget > Database Instance: emtarget_emtarget1 > Advisor Central > Automatic Database Diagnostic Monitor (ADDM)
stbdf12.us.oracle.c
emtarget_emtarget1 (Database Instance)
Automatic Database Diagnostic Monitor (ADDM)
Page Refreshed Apr 5, 2008 2:17:40 PM PDT
Refre
Database Activity
Run ADDM
Finding History
The icon selected below the graph identifies the ADDM analysis period. Click on a different icon to select a different analysis period.
Active Sessions
10.58
7.06
3.53
0.0
6:00
Apr 3, 2008
8
10
12 AM
2
4
6
8
10
12 PM
2
4
Wait
User I/O
CPU
Zoom
TIP For an explanation of the icons and symbols used in this page, see the Icon Key
ADDM Performance Analysis
Task Name ADDM:3132078998_1_1978
Filters
View Snapshots
View Report
Task Owner SYS
Average Active Sessions 10.2
Period Start Time Apr 4, 2008 4:00:31 AM PDT
Period Duration (minutes) 60
Instance emtarget_emtarget1
Global ADDM Task Name ADDM:3132078998_1978
Impact (%)
Finding
Occurrences (latest 24 hrs)
47.9 Top SQL by DB Time 24 of 24
39.5 Top SQL By I/O 0 of 24
36.3 Top Segments by I/O 1 of 24
12.9 Commits and Rollbacks 23 of 24
7.6 I/O Throughput 1 of 24

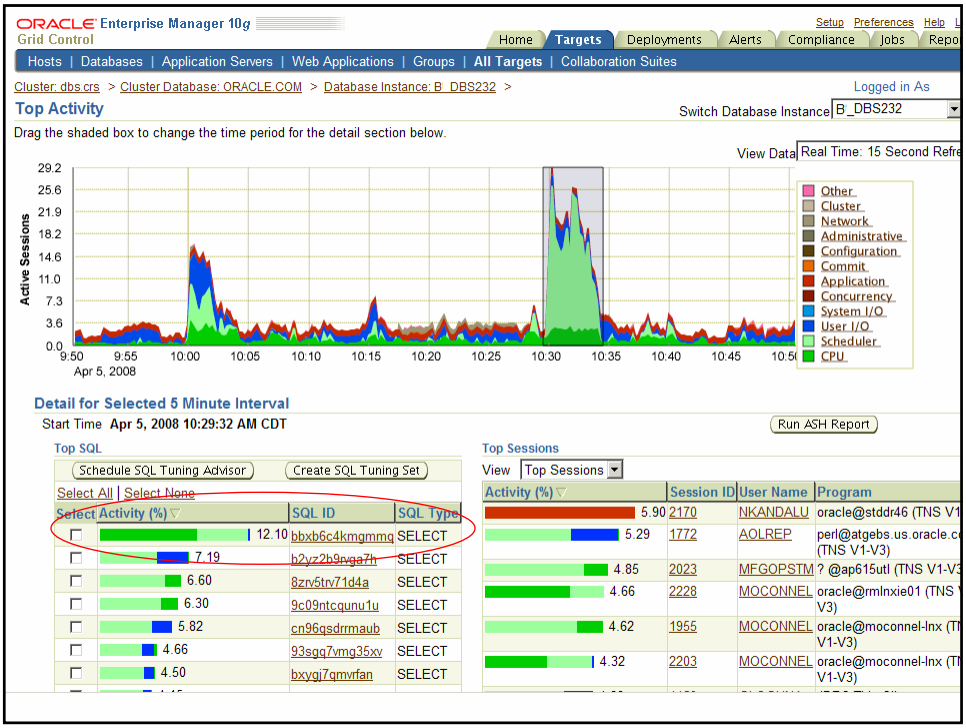

http://stbdf12.us.oracle.com:7664/em/faces/sdk/nonFacesWrapper?_adf.winId=1342872582_748&_adf.ct
Google
Enterprise Manager 11g Grid Control
Setup Help
Grid Targets Topology Favorites History Search All Target Types
... > Advisor Central > Automatic Database Diagnostic Monitor (ADDM):SYS.ADDM:3132078998_1_1986 > Performance Finding Details
stbdf12.us.oracle.c
emtarget_emtarget1 (Database Instance)
Performance Finding Details: Top SQL by DB Time
Finding SQL statements consuming significant database time were found. Finding History
Impact (Active Sessions) 4.03
Impact (%) 52.8
Period Start Time Apr 4, 2008 12:00:04 PM PDT
Period Duration (minutes) 60.2
Filtered No Filters
Recommendations
Schedule SQL Tuning Advisor
Select All | Select None | Show All Details | Hide All Details
Select Details Category Benefit (%)
Hide SQL Tuning 15
Action Investigate the SQL statement with SQL_ID "66n44vwsmyknr" for possible performance improvements. View Tuning History
SQL Text select /* serial_guys */ p_brand, p_type, p_size, ...
SQL ID 66n44vwsmyknr
Rationale SQL statement with SQL_ID "66n44vwsmyknr" was executed 4 times and had an average elapsed time of 1031 seconds.
Hide SQL Tuning 13.3
Action Run SQL Tuning Advisor on the SQL statement with SQL_ID "4scj37xz190kp". View Tuning History Run Advisor Now Filters
SQL Text select /* big_guys */ /* NO_GBY_PUSHDOWN */ s_name, s_address ...
SQL ID 4scj37xz190kp
Show SQL Tuning 10.2
Show SQL Tuning 8
Show SQL Tuning 6.6
Findings Path
Expand All | Collapse All

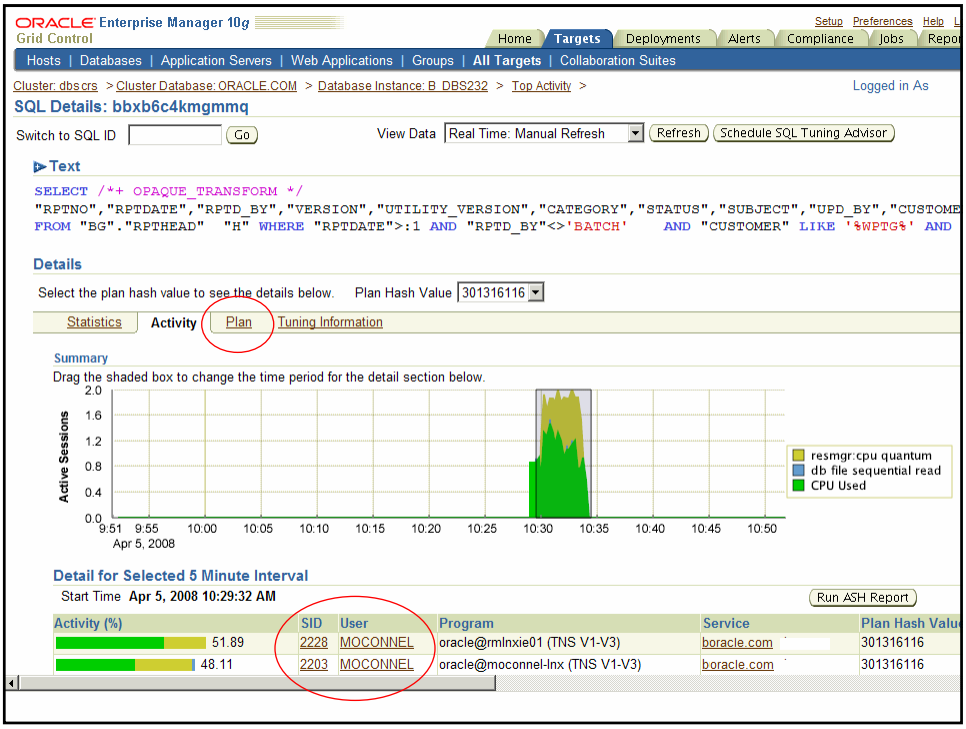


Best Practice: EM Real-time Interface

- Transient (sub-hour) or immediate time scope
  - Requires interactivity of UI
- 'Click on the big stuff'
  - Data visualizations display skew directly
- Takes some expertise to separate symptoms from root causes


ORACLE Enterprise Manager 10g
Grid Control
Setup Preferences Help
Home Targets Deployments Alerts Compliance Jobs Report
Hosts | Databases | Application Servers | Web Applications | Groups | All Targets | Collaboration Suites
Cluster: dbs crs > Cluster Database: ORACLE.COM >
Logged in As
Database Instance: B_DBS232
Switch Database Instance B_DBS232
Home Performance Administration Maintenance
Click on an area of a graph or legend to get more detail.
View Data Real Time: 15 Second Refres
Host
Runnable Processes
Maximum CPU
Load Average
9:50am 10:00am 10:10am 10:20am 10:30am 10:40am
April 5, 2008
Average Active Sessions
Run ADDM Now
Run ASH Re
Active Sessions
Other
Cluster
Network
Administrative
Configuration
Commit
Application
Concurrency
System I/O
User I/O
Scheduler
CPU Used
Top Activity
Instance Disk I/O
Per Second
Other Writes (KB
Physical Writes
Other Reads (KB
Physical Reads
Instance Throughput
Logons
Transactions


ORACLE Enterprise Manager 10g
Grid Control
Setup Preferences Help
Home Targets Deployments Alerts Compliance Jobs Repo
Hosts | Databases | Application Servers | Web Applications | Groups | All Targets | Collaboration Suites
Cluster: dbs crs > Cluster Database: ORACLE.COM > Database Instance: B_DBS232 >
Logged in As
Top Activity
Switch Database Instance B_DBS232
Drag the shaded box to change the time period for the detail section below.
View Data Real Time: 15 Second Refre
Active Sessions
Other
Cluster
Network
Administrative
Configuration
Commit
Application
Concurrency
System I/O
User I/O
Scheduler
CPU
9:50 9:55 10:00 10:05 10:10 10:15 10:20 10:25 10:30 10:35 10:40 10:45
Apr 5, 2008
Detail for Selected 5 Minute Interval
Start Time Apr 5, 2008 10:29:32 AM CDT
Run ASH Report
Top SQL
Schedule SQL Tuning Advisor
Create SQL Tuning Set
Select All | Select None
Select Activity (%) SQL ID SQL Type
12.10 bbxb6c4kmgmmq SELECT
7.19 b2yz2b9nvga7h SELECT
6.60 8zrv5trv71d4a SELECT
6.30 9c09ntcqunu1u SELECT
5.82 cn96qsdrrmaub SELECT
4.66 93sgq7vmg35xv SELECT
4.50 bxygj7qmvrfan SELECT
Top Sessions
View Top Sessions
Activity (%) Session ID User Name Program
5.90 2170 NKANDALU oracle@stddr46 (TNS V1
5.29 1772 AOLREP perl@atgebs.us.oracle.c (TNS V1-V3)
4.85 2023 MFGOPSTM ? @ap615utl (TNS V1-V3
4.66 2228 MOCONNEL oracle@rmlnxie01 (TNS V3)
4.62 1955 MOCONNEL oracle@moconnel-lnx (T V1-V3)
4.32 2203 MOCONNEL oracle@moconnel-lnx (T V1-V3)



Conclusions

- Design of Large, High Performance Systems is engineering, not a mystical art
- Good database design leads to
  - Simpler implementation
  - Scalability
- Oracle database instrumentation
  - Simpler performance diagnosis
  - Increases probability of success

Time for Your Questions