ACE Director
Performance Tuning Boot Camp

Indiana Oracle Users Group

Health Alert

This presentation contains 416 slides and there is no detox facility

Daniel A. Morgan

- More than 40 years technology experience
  - First computer was an IBM mainframe in 1970: Punch cards and Fortran IV
- Oracle ACE Director
- Wrote the course and taught Oracle at the University of Washington for 10 years
- Guest lecturer on Oracle at Harvard University
- Senior Technologist at Performance Tuning Corp.

Official Beta Site ORACLE DATABASE 12c

ORACLE RACSIG

International zSeries Oracle SIG

Daniel A. Morgan

Morgan's Library

Search

www library

Morgan's 2010 - 2011 Calendar

May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr

EMEA Harmony Conference

Tallinn, Estonia

May 20-21, 2010

A joint conference of the Estonian, Finnish, Latvian and Russian user groups
EMEA Harmony will focus on Technology, Middleware and BI
Featured speakers include Tom Kyte, Mogen Norgaard, Tanel Poder, and Dan Morgan

Community
Events
Training
Evening Workshops

Resources
Library
How Can I?
Code Samples
Presentations
Links
Book Reviews
Downloads
User Groups

General
Contact
About
Services
Legal Notice & Terms of Use
Privacy Statement

Presentations Map

The Mad Dog ACE

Training Events
- EMEA Harmony - May 20 - 21, Tallinn, Estonia
- NoCOUG - August 2010,
- AIOUG Sep 3 - 4, Hyderabad, India
- OOW - Sep 19 - 23, San Francisco CA
- LAD Tour - October
- DOAG - Nov 16 - 18, Nurnberg, Germany
- UKOUG - Nov 29 - Dec 1, Birmingham UK

Oracle Events

EMEA Harmony - Tallinn Estonia - May 20- 21

Morgan

aboard USA-71

Library News
- Morgan's Notepad vi (Blog) UPDATED
- Join the Western Washington OUG
- Morgan's Oracle Podcast
- DBA Best Practice Guidelines
- Bryn Llewellyn's PL/SQL White Paper
- Bryn Llewellyn's Editioning White Paper
- Troubleshooting Performance

ACE News

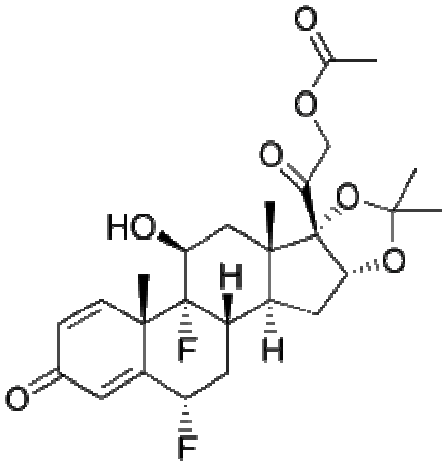
Would you like to become an Oracle ACE?

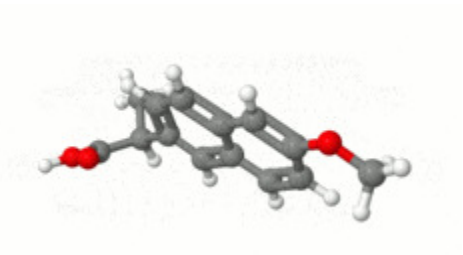
Learn more about becoming an ACE

- ACE Directory
- ACE Google Map
- ACE Nomination Form
- Stanley's Blog

Former Favorite Work Activity: Chemist

- Syntex Research
  - Analytical Organic Chemist
  - Member CNS Taskforce
- IND and NDA work on anabolic steroids and non-steroidal anti-inflamatories
  - Fluocinolone Acetonide (Synalar)
  - Fluocinolone Acetonide Acetate (Lidex)
  - Naproxen Sodium (Aleve, Naprosyn)





Favorite Work Activity: Travel and Lecture

With Chris Date at
Oracle Users Group Finland 2011

A Relational Model of Data for Large Shared Data Banks

E. F. Codd
IBM Research Laboratory, San Jose, California

Future users of large data banks must be protected from having to know how the data is organized in the machine (the internal representation). A prompting service which supplies such information is not a satisfactory solution. Activities of users at terminals and most application programs should remain unaffected when the internal representation of data is changed and even when some aspects of the external representation are changed. Changes in data representation will often be needed as a result of changes in query, update, and report traffic and natural growth in the types of stored information.

Existing noninferential, formatted data systems provide users with tree-structured files or slightly more general network

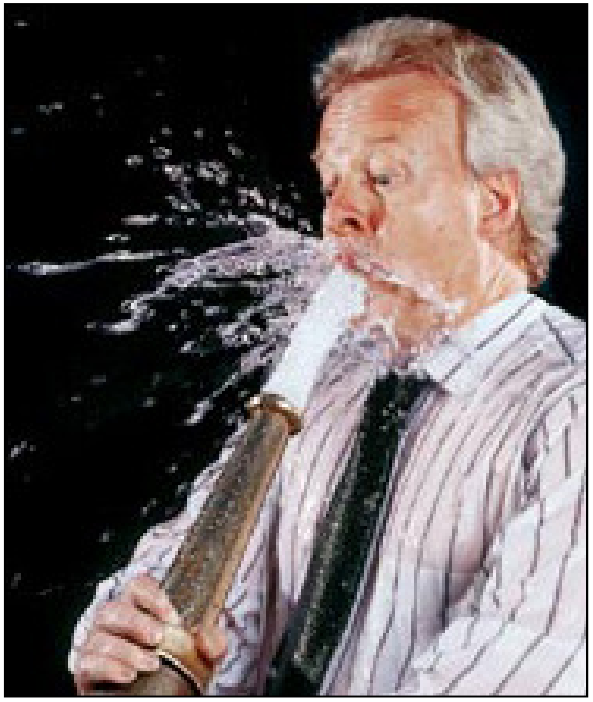
Favorite Non-Work Activity: Blue Water Sailing

Racing Larry's AC72 (LOL)

Syllabus

- Oracle "Best Practice" Guidelines
- Performance Improvement Methodologies & Tools
- Root Cause Analysis and RCA Tools
- Diagnosing and Tuning Database Performance
  - Network Layer
  - Storage Layer
  - Server Layer
  - Virtual Machine Layer
  - Operating System Layer
  - LVM Layer
  - Database Layer
  - Middleware Layer
  - Application Layer
- What's new in 12c ... you won't hear from Oracle

Take Notes ... Ask Questions



This class is not going to be hours of someone reading slides ...

Disclaimer

- This room is an unsafe harbour
- No one from Oracle has previewed this presentation
- No one from Oracle knows what I'm going to say
- No one from Oracle has supplied any of my materials

- and Oracle Legal has not reviewed this disclaimer

Performance Tuning Q&A (1:3)

- How many bytes of redo is generated by this update statement? ... this is not a trick question

```
CREATE TABLE t (
person_id   NUMBER,
last_name   VARCHAR2(20),
active_flag VARCHAR2(1));

CREATE BITMAP INDEX bix_t
ON t(active_flag);



UPDATE t
SET active_flag = 'Y'
WHERE active_flag = 'N'
AND rownum = 1;
```

- How does the answer to this question relate to performance?

Performance Tuning Q&A (2:3)

- When you perform the following update statement where is the undo stored and for how long?

```
CREATE TABLE sec_tab_kd (
rid  NUMBER(5),
bcol BLOB)
LOB (bcol)
STORE AS SECUREFILE bcol (
TABLESPACE securefiletbs
RETENTION MIN 3600
KEEP_DUPLICATES NOCOMPRESS DECRYPT CACHE READS)
TABLESPACE uwdata;


...
BEGIN
  UPDATE sec_tab_kd
  SET bcol = dst_file
  WHERE rid = i;
  COMMIT;
END;
/
```

- How does the answer to this question relate to performance?

Performance Tuning Q&A (3:3)

- Is this a database issue?

We have issues on instance XYZPCE1D.

Here is the error from the logs.

Caused by: oracle.ucp.UniversalConnectionPoolException: Cannot get Connection from Datasource: java.sql.SQLException: Listener refused the connection with the following error: **ORA-12516**, TNS:listener could not find available handler with matching protocol stack

```
at oracle.ucp.util.UCPErrorHandler.newUniversalConnectionPoolException(UCPErrorHandler.java:368)
at oracle.ucp.util.UCPErrorHandler.throwUniversalConnectionPoolException(UCPErrorHandler.java:49)
at oracle.ucp.util.UCPErrorHandler.throwUniversalConnectionPoolException(UCPErrorHandler.java:80)
at oracle.ucp.jdbc.DataSourceConnectionFactoryAdapter.createConnection(DataSourceConnectionFactoryAdapter.java:102)
at oracle.ucp.common.UniversalConnectionPoolImpl$UniversalConnectionPoolInternal.createOnePooledConnectionInternal(UniversalConnectionPoolImpl.java:1614)
at oracle.ucp.common.UniversalConnectionPoolImpl$UniversalConnectionPoolInternal.access$600(UniversalConnectionPoolImpl.java:1450)
at oracle.ucp.common.UniversalConnectionPoolImpl.createOnePooledConnection(UniversalConnectionPoolImpl.java:528)
at oracle.ucp.jdbc.oracle.OracleJDBCConnectionPool.createOnePooledConnection(OracleJDBCConnectionPool.java:129)
at oracle.ucp.common.UniversalConnectionPoolImpl.borrowConnectionWithoutCountingRequests(UniversalConnectionPoolImpl.java:332)
at oracle.ucp.common.UniversalConnectionPoolImpl.borrowConnectionAndValidate(UniversalConnectionPoolImpl.java:167)
at oracle.ucp.common.UniversalConnectionPoolImpl.borrowConnection(UniversalConnectionPoolImpl.java:142)
at oracle.ucp.jdbc.JDBCConnectionPool.borrowConnection(JDBCConnectionPool.java:157)
at oracle.ucp.jdbc.PoolDataSourceImpl.getConnection(PoolDataSourceImpl.java:916)
```

- How does the answer to this question relate to performance?

I do not believe that ...

- most of the time the root cause of a an Oracle database problem is the database itself
- most tuning involves rewriting SQL statements
  - We've tried it that way for more than 30 years and yet all of you are here because of performance-related issues
- most tuning tools are worth what they cost: Even when they are free
- tuning issues persist because of a lack of brilliant work by Jonathan Lewis, Maria Colgan, Tom Kyte, Cary Millsap, Tanel Poder, Greg Rahn, Christian Antognini, Richard Foote, etc., etc., etc. ....
- focusing on edge cases will ever address the underlying issues

I do believe that ...

- we are not failing for lack of effort
- we are not failing for lack of intelligence
- we are failing to cure performance issues just as medicine is failing to cure cancer
- we are failing because the ways that we are trying to attack performance tuning issues will not yield a cure

Oracle "Best Practice" Guidelines

"Best Practice"

- "Best Practice" does <u>not</u> mean
  - Buy an Exadata
  - Index scans are better than full table scans
  - Always use the latest/greatest feature Oracle licenses
- "Best Practice" does mean
  - Treat your database as a physician would treat a patient
  - EMT's have the ABC's
    - Airway ... is the airway open?
    - Breathing ... is the patient breathing?
    - Circulation ... is the heart beating?
  - DBA's have the SSS's
    - Security ... nothing matters if opening the database puts data at risk
    - Stability ... no one can use it if it isn't stable
    - Scalability ... it won't justify its existence if it can not scale to handle the load

Study Anatomy

- Physician
  - Cells & Tissues
  - Circulatory System
  - Digestive System
  - Endocrine System
  - Muscular System
  - Nervous System
  - Skeletal System
  - Urogenital System
- DBA
  - SGA
  - PGA
  - UGA
  - Buffer Cache
  - BigFiles
  - SecureFiles
  - Tablespace Groups
  - Refresh Groups

Study Physiology

- Physician
  - Aging
  - Blood Clotting
  - Hearing
  - Hormones
  - Inflammation
  - Lymphatic System
  - Nutrition
  - Pineal Gland
- DBA
  - CBO
  - LGWR
  - SMON
  - Autoextend
  - Dynamic Resizing
  - LRU
  - Multiversion Concurrency
  - Redo Log Switches

Learn the Tools of the Trade

- Physician
  - Scalpel
  - Forceps
  - Sphygmomanometer
  - Stethoscope
  - X-ray
  - Ultrasound
  - EKG
  - EEG
  - Cat Scan
  - MRI
  - Otoscope
  - Autoclave
  - Colonoscopy
  - Barium Enema
- DBA
  - RMAN
  - NETCA
  - DBCA
  - DBUA
  - Explain Plan
  - Autotrace
  - DBMS_MONITOR
  - DBMS_HPROF
  - OEM Cloud Control
  - TKPROF
  - AWR
  - SQL*Plus
  - SQL Tuning Advisor
  - ADDM

Take The Patient's History

- Physician
  - Patient name and age
  - When did it start?
  - Has this happened before?
  - Is there a family history?
  - What has changed recently?
  - Blood Pressure and Pulse
  - We need to do some tests
  - Take two aspirin and call me in the morning if the pain and fever persist

- DBA
  - Application name and version
  - When did it start?
  - Has this happened before?
  - Is there a reported bug?
  - What has changed recently?
  - Alert Log
  - We need an AWR Report
  - Reboot the system and call me in the morning if the report is still slow

"Best Practices" Guidelines

- What you should know about every database you manage
  - Network infrastructure: every switch, router, hub, firewall, NIC, HBA, IP address, manufacturer, model, bandwidth, latency, virtualized, routable, jumbo frames, VLAN
  - Operating system ... version, patches, kernel parameters, are HugePages enabled, is NUMA enabled, is OSWatcher installed
  - And for the database itself (and this is a very short list)

Multitenant Architecture	Archiving Enabled	Non-Default Init Parameters	Event Triggers In Use	Corrupt Database Blocks Present	Feature Usage Report
Stand-alone or RAC	Flashback Logging Enabled	Default Passwords In Use	Optimizer Stats Current	Log Switching Frequency	Skipped AWR Reports
Editioning Enabled	Supplemental Logging Enabled	Undo Retention	System Stats Collected	Objects Compiled in Debug Mode	Valid Registry
Replication Running	Force Logging Enabled	Backup Warnings and Failures	Auditing Enabled	Properly Sized UNDO Tablespace	Invalid Objects
Version and Patch Level	Control Files Multiplexed	Redo Logs Members Multiplexed	Disabled Triggers and Constraints	AMM or ASMM Enabled	Block Change Tracking Enabled

"Best Practices" Guidelines

- What you should monitor daily
  - Alert logs (Operating System, ASM, Clusterware, Database)
  - Audit trails (Operating System and Database)
  - Corrupt Blocks
  - Invalid Objects
  - Redo Log Switches
  - Backup Success/Warnings/Failures
- What you should alert on with DBMS_SERVER_ALERT
  - [review the full list of options]

"Best Practices" Guidelines: Corrupt Blocks

```
conn / as sysdba

col corruption_change# format 99999999999999999999

SELECT * FROM v$database_block_corruption ORDER BY 1,3;

     FILE#     BLOCK#     BLOCKS    CORRUPTION_CHANGE# CORRUPTIO
---------- ---------- ---------- --------------------- ---------
        22     162844          4         10594075667556 CORRUPT

-- if corruption is found use the following SQL to identify the corrupt segment(s)
SELECT de.owner, de.segment_name, de.segment_type
FROM dba_extents de, v$database_block_corruption vdbc
WHERE de.file_id = vdbc.file#
AND vdbc.block# BETWEEN de.block_id AND (de.block_id+(de.blocks-1));

-- then consider using DBMS_REPAIR to repair the corruption.
-- also consider RMAN> RECOVER CORRUPTION LIST;
```

"Best Practices" Guidelines: Log File Switches

```
SELECT TO_CHAR(first_time,'MMDD') MMDD,
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'00',1,0)),'99') "00",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'01',1,0)),'99') "01",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'02',1,0)),'99') "02",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'03',1,0)),'99') "03",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'04',1,0)),'99') "04",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'05',1,0)),'99') "05",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'06',1,0)),'99') "06",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'07',1,0)),'99') "07",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'08',1,0)),'99') "08",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'09',1,0)),'99') "09",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'10',1,0)),'99') "10",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'11',1,0)),'99') "11",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'12',1,0)),'99') "12",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'13',1,0)),'99') "13",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'14',1,0)),'99') "14",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'15',1,0)),'99') "15",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'16',1,0)),'99') "16",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'17',1,0)),'99') "17",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'18',1,0)),'99') "18",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'19',1,0)),'99') "19",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'20',1,0)),'99') "20",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'21',1,0)),'99') "21",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'22',1,0)),'99') "22",
TO_CHAR(SUM(DECODE(TO_CHAR(first_time,'HH24'),'23',1,0)),'99') "23"
FROM v$log_history
GROUP BY TO_CHAR(first_time,'MMDD')
ORDER BY 1;
```

"Best Practices" Guidelines: Log File Switches

MMDD	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
0609	16	11	9	8	8	10	12	8	8	10	8	10	14	10	11	15	15	8	12	8	7	6	9	7
0610	13	12	8	9	7	6	11	9	6	8	7	8	12	6	7	6	8	7	10	7	4	4	4	5
0611	12	8	5	9	9	7	11	7	6	7	8	5	12	9	10	8	9	12	12	10	6	6	9	8
0612	13	12	7	9	7	9	10	10	7	7	9	8	11	7	7	8	7	7	11	9	5	6	8	7
0613	12	11	7	8	8	7	13	7	9	7	8	7	13	10	9	8	8	8	11	8	7	5	7	6
0614	15	10	9	9	8	9	13	9	9	7	11	13	11	9	8	9	13	9	12	9	7	9	7	7
0615	15	10	10	8	10	9	12	8	9	8	9	7	13	6	8	7	7	7	15	10	7	7	7	5
0616	13	8	8	7	7	6	10	8	11	7	8	6	11	7	12	13	13	14	13	9	9	9	7	8
0617	15	13	10	9	8	9	16	8	8	10	9	10	16	11	10	10	8	11	13	8	9	9	7	9
0618	12	13	15	15	13	13	15	13	9	12	8	11	14	9	10	9	9	8	14	9	8	8	9	8
0619	16	11	10	11	9	9	13	12	10	9	12	12	17	8	9	9	11	11	14	9	9	11	10	12
0620	19	15	11	10	10	10	19	11	9	9	9	9	13	7	15	10	11	11	12	10	9	11	11	10
0621	13	16	11	9	10	13	16	8	14	9	11	12	17	10	10	11	8	11	14	8	11	14	8	11
0622	16	13	13	11	11	9	16	9	9	11	10	11	17	10	9	10	10	10	13	14	9	10	10	8
0623	19	13	12	13	13	11	16	12	11	11	11	11	16	9	10	13	2	14	14	8	9	8	8	8
0624	14	9	9	9	7	9	11	8	8	7	8	8	14	7	8	7	9	3	6	0	0	0	0	0

...

0630	7	4	23	19	9	10	5	6	7	17	19	17	15	17	15	43	40	32	17	15	14	20	13	15
0701	15	12	14	12	13	12	13	17	15	17	20	20	18	18	17	15	14	13	10	10	15	15	13	19
0702	21	22	20	18	14	14	12	13	11	11	14	14	14	10	9	10	9	10	11	9	11	9	10	12
0703	9	13	10	17	14	17	15	17	23	20	19	20	17	19	16	17	15	17	15	15	15	16	16	18
0704	22	19	19	18	16	15	13	13	14	11	13	10	12	14	10	12	14	11	9	11	12	13	12	9
0705	14	13	9	11	10	12	13	11	11	8	10	10	11	11	11	12	10	10	9	10	8	9	12	7
0706	14	15	11	12	9	15	13	12	12	9	12	14	12	12	12	12	13	11	8	9	12	13	2	0
0707	0	0	1	0	3	15	10	10	7	8	10	11	12	8	6	9	13	12	9	8	9	8	10	10
0708	16	9	8	15	10	11	9	8	8	14	9	10	10	8	8	14	15	10	9	9	8	9	10	10
0709	13	12	9	10	10	9	9	10	11	11	8	9	9	8	9	13	8	9	6	9	9	11	10	9
0710	12	10	9	10	9	12	9	8	8	11	7	10	11	9	9	13	10	9	8	9	11	12	10	10

"Best Practices" Guidelines: Backups

```
SELECT start_time, end_time, input_type, status
FROM v$rman_backup_job_details
ORDER BY 1;

START_TIME END_TIME STATUS
------------------- ------------------- -----------------------
03-FEB-2013 18:04:46 03-FEB-2013 19:44:57 COMPLETED
04-FEB-2013 00:05:56 FAILED
04-FEB-2013 02:00:54 04-FEB-2013 03:30:21 COMPLETED
04-FEB-2013 08:01:03 04-FEB-2013 09:27:16 COMPLETED
04-FEB-2013 14:09:03 04-FEB-2013 15:09:35 COMPLETED
04-FEB-2013 20:03:17 05-FEB-2013 00:46:34 COMPLETED WITH ERRORS
05-FEB-2013 02:01:54 05-FEB-2013 03:05:00 COMPLETED
05-FEB-2013 08:01:49 05-FEB-2013 09:16:59 COMPLETED
05-FEB-2013 08:36:15 05-FEB-2013 08:37:47 COMPLETED
05-FEB-2013 14:01:35 05-FEB-2013 15:31:33 COMPLETED
05-FEB-2013 20:01:23 05-FEB-2013 22:05:26 COMPLETED WITH ERRORS
06-FEB-2013 02:03:05 06-FEB-2013 03:37:16 COMPLETED
06-FEB-2013 08:02:56 06-FEB-2013 09:24:00 COMPLETED
06-FEB-2013 14:07:53 06-FEB-2013 15:12:10 COMPLETED
06-FEB-2013 20:01:50 06-FEB-2013 21:06:41 FAILED
07-FEB-2013 02:01:42 07-FEB-2013 03:16:17 COMPLETED
07-FEB-2013 08:02:35 07-FEB-2013 09:43:35 COMPLETED
07-FEB-2013 14:03:12 07-FEB-2013 15:14:44 COMPLETED
07-FEB-2013 20:03:56 FAILED
08-FEB-2013 02:04:27 08-FEB-2013 03:24:14 COMPLETED
08-FEB-2013 08:04:43 08-FEB-2013 09:44:53 COMPLETED
08-FEB-2013 08:37:11 08-FEB-2013 08:38:28 COMPLETED
08-FEB-2013 14:05:53 08-FEB-2013 15:52:53 COMPLETED
08-FEB-2013 20:04:20 08-FEB-2013 22:19:13 FAILED
09-FEB-2013 02:05:12 09-FEB-2013 04:15:11 COMPLETED
09-FEB-2013 08:07:40 09-FEB-2013 10:15:32 COMPLETED
09-FEB-2013 14:06:40 09-FEB-2013 15:17:38 COMPLETED
09-FEB-2013 20:05:35 09-FEB-2013 21:47:53 FAILED
10-FEB-2013 02:06:13 10-FEB-2013 03:46:37 COMPLETED
10-FEB-2013 08:06:56 10-FEB-2013 09:44:15 COMPLETED
10-FEB-2013 08:44:16 10-FEB-2013 08:45:33 COMPLETED
10-FEB-2013 14:06:14 10-FEB-2013 15:46:15 COMPLETED
10-FEB-2013 20:07:30 10-FEB-2013 21:54:22 FAILED
11-FEB-2013 00:53:58 11-FEB-2013 03:07:22 COMPLETE
```

"Best Practices" Guidelines: Tuning

- Tuning should happen before
  - Hardware is purchased
  - Code is written
  - Software licenses are purchased
  - Anything is installed or deployed
- If it isn't Secure you are wasting time and money
- If it isn't Stable you have already wasted money
- If it isn't Scalable you have already wasted money

- Point of view:
  If you are tuning a production system ... you've failed

And you must read the docs

Subject: Re: What's a quick simple way to test Oracle connectivity between client and database server?
From: Spin <Spin@invalid.com>
Reply-To: Spin <Spin@invalid.com>
Date: 11/6/2008 5:56 PM
Newsgroups: comp.databases.oracle.misc

I am flooded with 20 projects. I do not have time to read the documents. Besides, every person who answers my question I give acknowledgement too, and it becomes one more notch on the belt of someone looking for industry recognition. And then there is the intrinsic value of helping someone. Unless, of course, you're someone who determines another's worth by some measure of "laziness". You placed that label on me. You do not know me. I could be lazy. Or I could be swamped with 20 projects at one time and do not have time to RTM. Or I could be someone who volunteers at animal shelters full-time, or working on community projects, or helping the elderly or troubled juveniles. All of these are activities, which may in fact, indirectly benefit you, if not the common good. And as to top-posting, some people prefer it. You don't have to scroll down to see the answer. And if it was so bad, why did Microsoft make the Outlook Express newsreader default to top-posting?

Performance Improvement Methodologies and Tools

Tuning Methodologies and Tools

- Blogs
- Books
- Docs
- Magazines
- Methodologies
- Movies
- Posters
- Presentations
- Services
- Tools
- White Papers
- PTC's Method

Note: No product or service listed here is an endorsement

Blogs: Gavin Soorma

Oracle DBA – Tips and Techniques

Gavin Soorma

MY BLOGS | ABOUT | ORACLE GOLDENGATE | ORACLE VIRTUAL MACHINE | DOCUMENTS AND WHITEPAPERS | PRESENTATIONS | RECOVERY SCENARIOS | UNIX FOR THE DBA | ORACLE DATABASE 12C | NOTES

That's Me

Oracle Certified Professional
7.3, 8i, 9i,10g, 11g, 12c
11i Apps DBA OCP
10g RAC OCE
Certified GoldenGate Implementation Specialist
Oracle 11g Exadata Implementation Specialist
10g OCM
11g OCM

ORACLE ACE

ORACLE Certified Master
Oracle Database 11g Administrator

« Upgrading EM12c Release 1 (12.1.0.1) to 12c Release 2 (12.1.0.2) | Home | Configuring APEX 4.2 and the APEX Listener on Oracle 11g WebLogic Server »

ASH and AWR Performance Tuning Scripts

Listed below are some SQL queries which I find particularly useful for performance tuning. These are based on the Active Session History V$ View to get a current perspective of performance and the DBA_HIST_* AWR history tables for obtaining performance data pertaining to a period of time in the past.

I would like to add that these queries have been customised by me based on SQL scripts obtained from colleagues and peers. So if I am infringing any copyright material let me know and I shall remove the same. Also If anyone has any similar useful scripts to contribute for use by the community do send it to me and I shall include the same on this page

Top Recent Wait Events

```
col EVENT format a60

select * from (
select active_session_history.event,
sum(active_session_history.wait_time +
active_session_history.time_waited) ttl_wait_time
from v$active_session_history active_session_history
where active_session_history.event is not null
group by active_session_history.event
order by 2 desc)
where rownum < 6
/
```

Top Wait Events Since Instance Startup

Blogs: Oracle Optimizer Team

File Edit View History Bookmarks Tools Help

Oracle Optimizer

https://blogs.oracle.com/optimizer/

DuckDuckGo Google Email Humor News Oracle Science Headlines

BLOGS HOME PRODUCTS & SERVICES DOWNLOADS SUPPORT PARTNERS COMMUNITIES ABOUT Login

Oracle Optimizer

Insight into the workings of the Optimizer

ORACLE

Main | Next page »

Friday Sep 27, 2013

Oracle Open World 2013 - It's a wrap

By Maria Colgan on Sep 27, 2013

Thanks to all those that attended the Optimizer sessions and stopped by the Optimizer demopod during this years Oracle OpenWorld. It was an amazing conference this year with lots of great technical sessions and interesting discussion at the demo grounds.

Since so many folks have emailed me requesting a copy of the Optimizer session presentations I thought it would be a good idea to post them here. The presentation were unfortunately were larger than the 2MB file limit so I have broken them multiple files. Enjoy!

Session CON8622 :

- What to expect from the Optimizer in Oracle Database 12c part1
- What to expect from the Optimizer in Oracle Database 12c part2
- What to expect from the Optimizer in Oracle Database 12c part3
- What to expect from the Optimizer in Oracle Database 12c part4

About

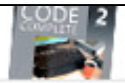
The Oracle Optimizer blog is written by members of the Optimizer development team. The goal of this blog is to provide an insight into the workings of the Optimizer and the statistics it relies on. The views expressed on this blog are our own and do not necessarily reflect the views of Oracle and its affiliates. The views and opinions expressed by visitors on this blog are theirs solely and may not reflect ours.

Search

Enter search term:

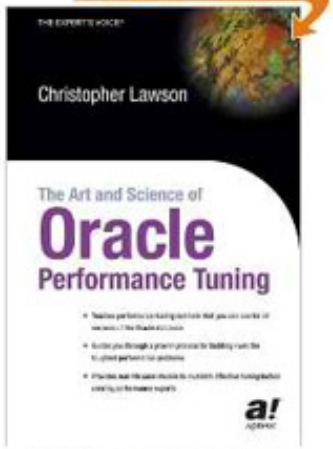
Books: Oracle Performance Tuning

Browse Programming Books C | Java | PHP | Python



Learn more

Click to LOOK INSIDE!



Click to open expanded view

The Art and Science of Oracle Performance Tuning [Paperback]

by Christopher Lawson

★★★★★ (13 customer reviews)

List Price: ~~$39.99~~

Price: $28.78 & **FREE Shipping** on orders over $35. Details

You Save: $11.21 (28%)

In Stock.

Ships from and sold by **Amazon.com.** Gift-wrap available.

Want it Monday, June 16? Order within **36 hrs 53 mins** and choose **One-Day Shipping** at checkout. Details

22 new from $25.16 23 used from $19.95

FREE TWO-DAY SHIPPING FOR COLLEGE STUDENTS

Learn more

amazonstudent

Books: Performance Tuning Tips & Techniques

Oracle Database 11g Release 2 Performance Tuning Tips & T... and over one million other books are available for **Amazon Kindle**.

Click to LOOK INSIDE!

Click to open expanded view

Oracle Database 11g Release 2 Performance Tuning Tips & Techniques (Oracle Press) [Paperback]

by Richard Niemiec

(23 customer reviews)

List Price: ~~$60.00~~

Price: $39.39 & **FREE Shipping**. Details

You Save: $20.61 (34%)

In Stock.

Ships from and sold by **Amazon.com**. Gift-wrap available.

Want it tomorrow, June 13? Order within **1 hr 55 mins** and choose **One-Day Shipping** at checkout. Details

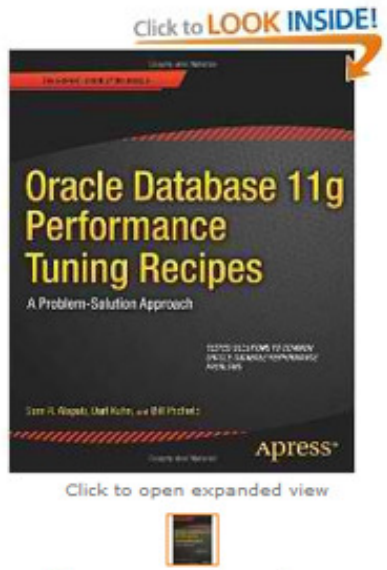
39 new from $30.00 22 used from $20.88

FREE TWO-DAY SHIPPING FOR COLLEGE STUDENTS

Learn more

amazonstudent

Books: Performance Tuning Recipes



Oracle Database 11g Performance Tuning Recipes: A Problem-Solution Approach (Expert's Voice in Oracle) [Paperback]

by Sam Alapati, Darl Kuhn, Bill Padfield

(9 customer reviews)

List Price: ~~$59.99~~

Price: **$55.19** & **FREE Shipping**. Details

You Save: $4.80 (8%)

In Stock.

Ships from and sold by **Amazon.com**. Gift-wrap available.

Want it Monday, June 16? Order within **36 hrs 55 mins** and choose **One-Day Shipping** at checkout. Details

26 new from $43.55 **17 used** from $32.66

FREE TWO-DAY SHIPPING FOR COLLEGE STUDENTS

Learn more

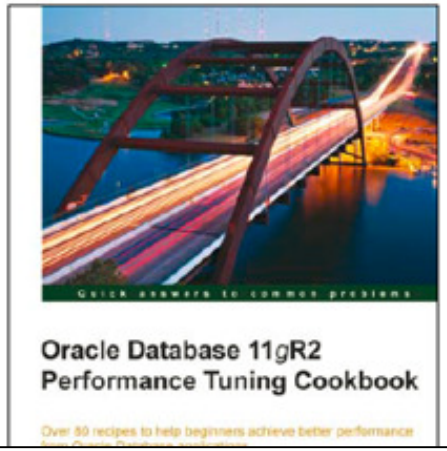
amazonstudent

Books: Performance Tuning Cookbook

[PACKT] PUBLISHING enterprise professional expertise distilled

Oracle Database 11gR2
Performance Tuning Cookbook

Ciro Fiorillo



Books: Method R - Cary Millsap

What is Method R?

Hits: 27785

Method R is a fast, effective, commonsense approach to optimizing a system. It was first documented in the book *Optimizing Oracle Performance*alt by Method R founding members Cary Millsap and Jeff Holt.

Method R prescribes four simple steps:

1. Identify the task that's the most important to you.
2. Measure its *response time* (R). In detail.
3. Optimize that response time in the most economically efficient way.
4. Repeat until your system is economically optimal.

Docs: Oracle

Oracle® Database 2 Day + Performance Tuning Guide
11*g* Release 1 (11.1)
B28275-02

Home Book List Contents Index Mas Ind

Previous Next PDF

2 Oracle Database Performance Method

Performance improvement is an iterative process. Removing the first bottleneck (a point where resource contention is at its highest) might not lead to performance improvement immediately because another bottleneck might be revealed that has an even greater performance impact on the system. For this reason, the Oracle performance method is iterative. Accurately diagnosing the performance problem is the first step toward ensuring that the changes you make to the system will result in improved performance.

Performance problems generally result from a lack of throughput, unacceptable user or job response time, or both. The problem might be localized to specific application modules, or it might span the entire system. Before looking at any database or operating system statistics, it is crucial to get feedback from the most important components of the system: the users of the system and the people ultimately paying for the application. Getting feedback from users makes determining the performance goal easier, and improved performance can be measured in terms of real business goals, rather than system statistics.

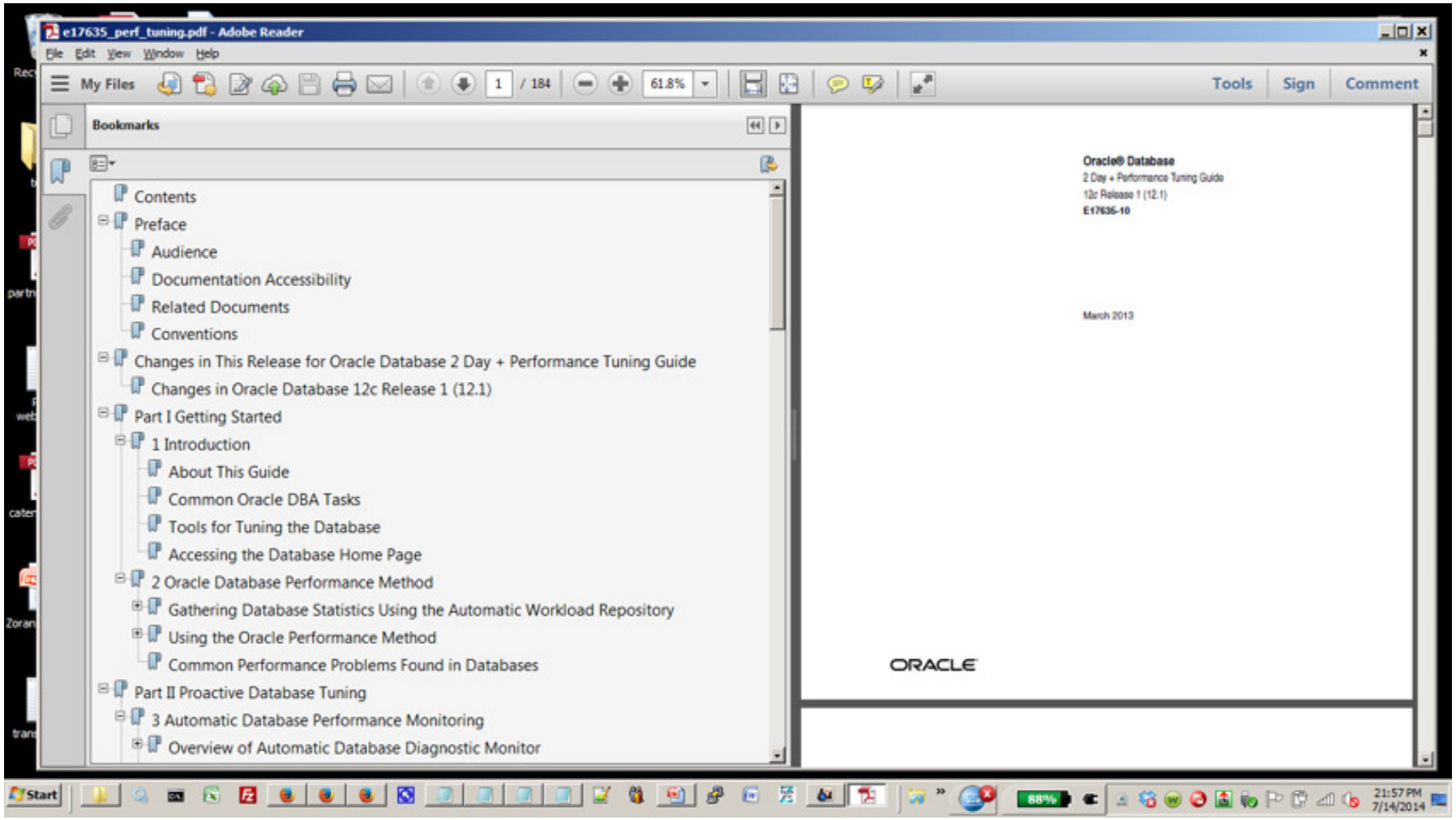
The Oracle performance method can be applied until performance goals are met or deemed impractical. Because this process is iterative, it is likely that some investigations will be made that have little impact on the performance of the system. It takes time and experience to accurately pinpoint critical bottlenecks in a timely manner. Automatic Database Diagnostic Monitor (ADDM) implements the Oracle performance method and analyzes statistics to provide automatic diagnosis of major performance problems. Using ADDM can significantly shorten the time required to improve the performance of a system, and it is the method used in this guide.

This chapter discusses the Oracle Database performance method and contains the following sections:

- Gathering Database Statistics Using the Automatic Workload Repository
- Using the Oracle Performance Method
- Common Performance Problems Found in Oracle Databases

Docs: Oracle

- E17635-10: 2 Day + Performance Tuning Guide



Magazines: Oracle Magazine

Oracle Technology Network > Oracle Magazine Issue Archive > 2013 > **January 2013**

- Oracle Magazine Online
- 2014
- 2013
- 2012
- 2011
- 2010
- 2009
- 2008

January 2013 | March 2013 | May 2013 | July 2013 | September 2013

TECHNOLOGY: Tuning

As Published In
ORACLE MAGAZINE
January/February 2013

Beginning Performance Tuning: Active Session History

By Arup Nanda

Identify the root cause of performance issues at a precise point in the past—even when the sessions have disconnected.

In my previous two articles on the topic of performance tuning (bit.ly/SEzR6t and bit.ly/U7U5IJ), you learned how to identify the cause of performance bottlenecks occurring in real time and view the summary of what happened in the past. Although these processes provide information to help resolve performance issues in many cases, there are times when you would like to know *precisely* what happened at a specific point in the past. In this article, you will learn how to identify the precise sequence of events in the past in a historical—not current—session.

To run the examples in this article, execute the setup.sql script. The script creates a user called ARUP and creates three tables: TEST1, TEST2, and TEST3. In the very rare event that a user called ARUP already exists, modify the script to create another user with a different name and replace all occurrences of ARUP with that name. The sample database created by setup.sql will take up less than 200 MB of space in your database.

After running the setup.sql script, open four SQL*Plus sessions, connected as ARUP. In three of these sessions, execute the test1.sql, test2.sql, and upd1.sql scripts, respectively. Here is an example of one session running upd1.sql on a UNIX-based system:

```
# sqlplus arup/arup
SQL> @upd1.sql
```

Methodologies: Guy Harrison

- Stages:
  1. Normalizing the application workload
  2. Reducing contention and bottlenecks
  3. Reducing physical IO
  4. Optimizing disk IO

Isn't it all about reducing I/O?

For most databases, the ultimate aim is to reduce disk IO. Disk IO remains the most expensive operation faced by the database.

When faced with an obviously IO-bound database, it's tempting to deal with the most obvious (symptom) – the IO subsystem – immediately. Unfortunately, this usually results in treating the symptom rather than the cause, is often expensive and usually ultimately futile. Methodically tuning the database layer by layer almost always leads to a healthier database, happier users and a more appreciated DBA.

Movies: YouTube

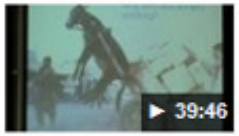
Google

youtube and oracle performance tuning

Web | Videos | News | Images | Shopping | More | Search tools

About 211,000 results (0.27 seconds)

Lecture 5 Oracle Database Performance Tuning ... - YouTube



www.youtube.com/watch?v=Ah1xjCl6Axg ▾ YouTube ▾
Nov 4, 2013 - Uploaded by Dell University
China Sun Yat-sen University and Dell co-established course IT System Management Introduction ...

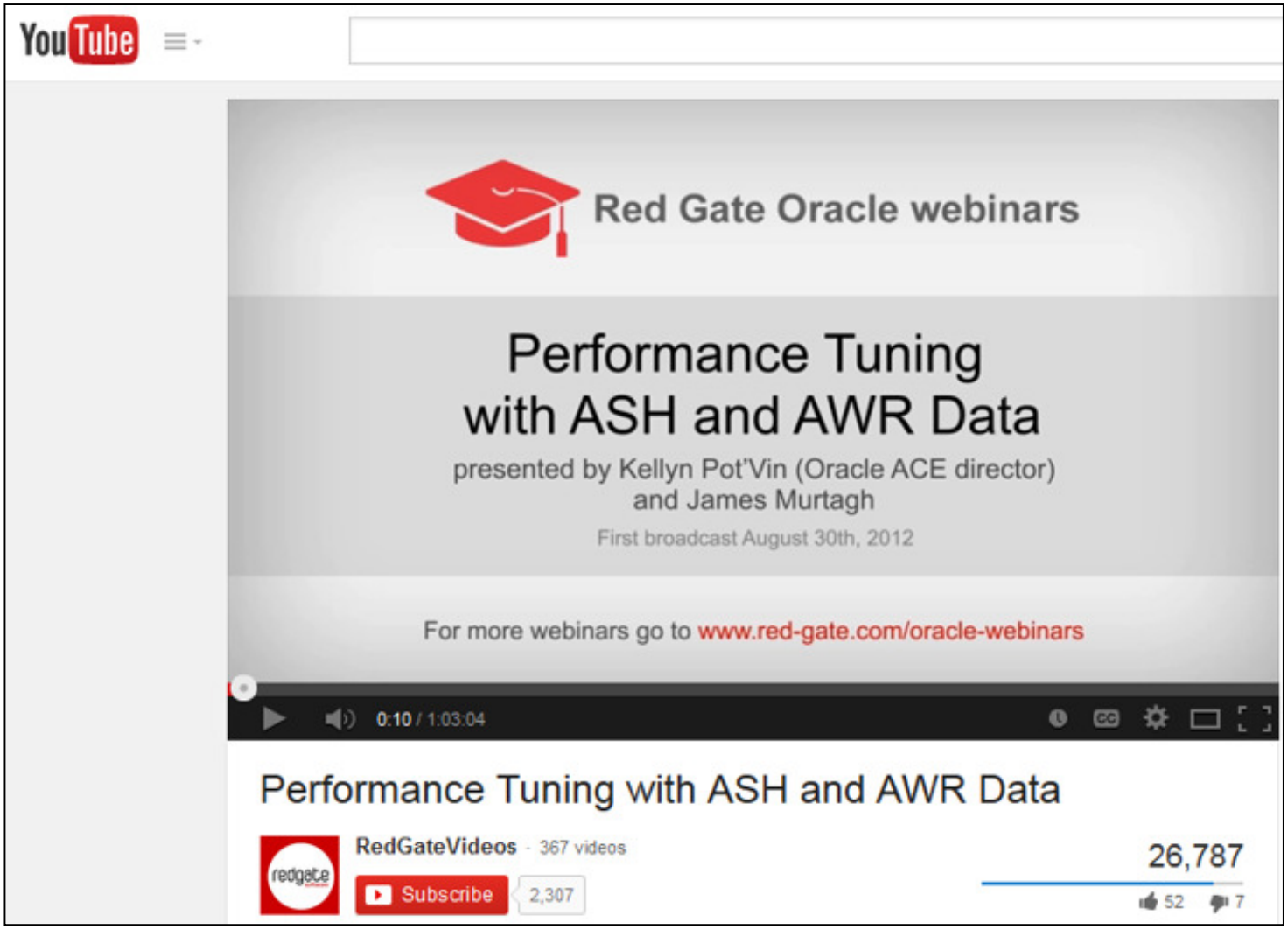
Understanding Oracle Performance Tuning - YouTube

www.youtube.com/watch?v=WkrA7xAciSs ▾ YouTube ▾
Feb 2, 2013 - Uploaded by Vijay Mahawar
Apologies for the poor video quality.*** This is recording of Google Hangout Event on "Understanding **Oracle** ...

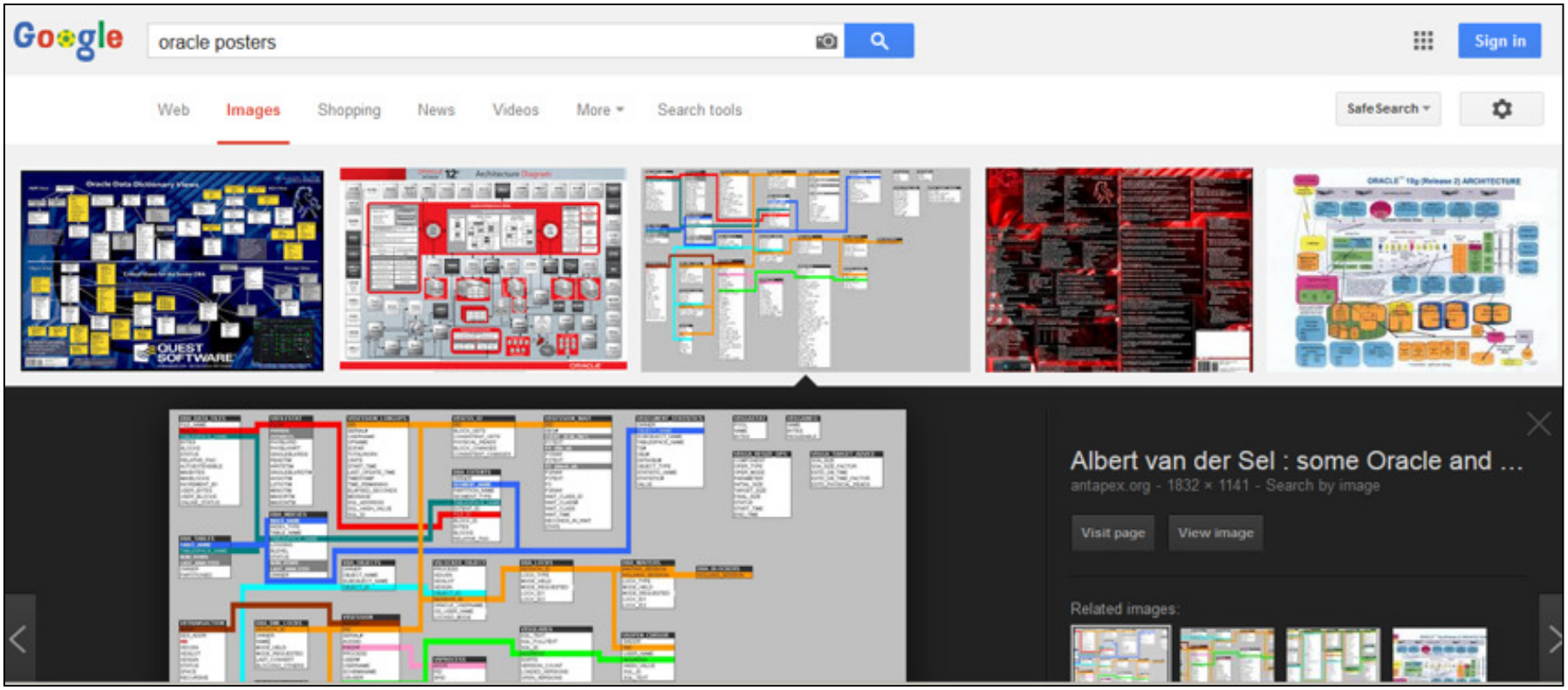
Oracle SQL Performance Tuning - YouTube

www.youtube.com/watch?v=GVZzl... ▾ YouTube ▾
Mar 2, 2009 - Uploaded by dwealth
A beginner's guide that shows you how to start on SQL **Performance tuning** with reference to **Oracle**. Talks ...

Movies: YouTube



Posters: Many Vendors



Presentations: Oracle

Oracle Tuning Methods

- Prehistory
  - Debug code
- Dark Ages
  - Counters/Ratios
  - BSTAT/ESTAT
  - SQL*Trace
- Renaissance
  - Increased instrumentation inc. Wait Events
  - Move from counters to timers
  - STATSPACK

Oracle Tuning Methods

- More Recent Methods – Time-based methods
  - YAPP
    - Instance tuning - instance statistics
      - Non intrusive
      - Always on
  - Method R
    - Session tuning – sql trace based
      - Tightly scoped
      - Must be highly selective
- Modern advances
  - DB Time Tuning
    - Instrumentation improvements
    - ASH, AWR,ADDM

ORACLE

Services: Datavail Performance Tuning

Datavail Video

Call Us Today at 877-634-9222

Performance Tuning Services

Need 24 X 7 Performance Tuning Help?

Datavail will fine tune your environment, code and storage to optimize your performance and put you on a schedule to maintain it. Our **"5S"** discovery process reveals the optimal plan for enhancing your performance. The Datavail team will review the following tactical areas for evidence of and solutions for best performance:

1. **SQL:** effective data utilization
2. **Space:** data storage and processing
3. **Sessions:** who is doing what, where and how
4. **Statistics:** leveraging data handling
5. **Scheduled Processes:** traffic control

The Datavail team work with you to explore, identify and resolve performance issues with the experience and expertise you need.

For best results, they will focus on the following technical elements:

- Database, Server and Instance
- SQL Statement Tuning
- Storage and I/O
- OLTP, DW and eCommerce
- Replication and Clustering
- Backup and Recovery

Request your FREE consultation

Denotes a required field.

First Name

Last Name

Company

Title

Phone

Email

Comments

Or call 877-634-9222

Interested in working for Datavail?

Services: Performance Tuning Corp.

June 14, 2014 (888) 750-3907 info@perftuning.com Oracle Performance Tuning and Microsoft SQL Server Performance Tuning Experts!

PERFORMANCE TUNING CORPORATION

SERVICES PARTNERS INSIGHTS EXPERTS CONNECT Let us empower you today!

- Tune
- Strategy & Planning
- Design & Architecture
- Build
- Manage
- Train

WE ARE THE INFORMATION PERFORMANCE EXPERTS

- DATABASE PERFORMANCE
- APPLICATION PERFORMANCE
- INFRASTRUCTURE PERFORMANCE

CLIENTS

g.com/services/

Tools: Confio's Solar Winds

File Edit View History Bookmarks Tools Help

Oracle Performance Monitoring T...

www.confio.com/performance/oracle/ignite/

DuckDuckGo Google Email Humor News Oracle Science Headlines

CONFIO solarwinds

Products Resources Support Partners Company News & Events Downloads

Ignite for Oracle

SolarWinds Database Performance Analyzer for Oracle (formerly Confio Ignite) uses response time analysis to find the Oracle problems causing application and end-user slowdowns. Improving Oracle performance has never been easier.

"IgniteVM eased my worries about moving our Oracle to VMware."

BUY NOW

Identify, analyze and resolve Oracle database performance problems. Free Trial View Demo

Tools: Embarcadero's DB Optimizer

embarcadero

Free Trials | Buy Now

Select country

Products | Company | Resources | News, Events, & Blogs | Solutions | Support | Partners

Search...

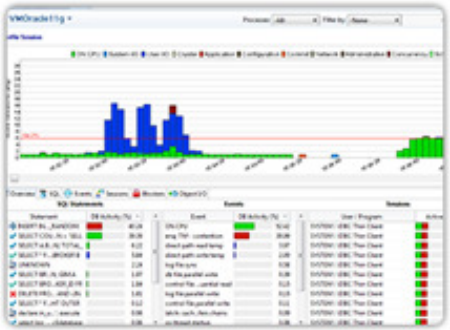
You Keep the SQL Running

DB Optimizer will help you

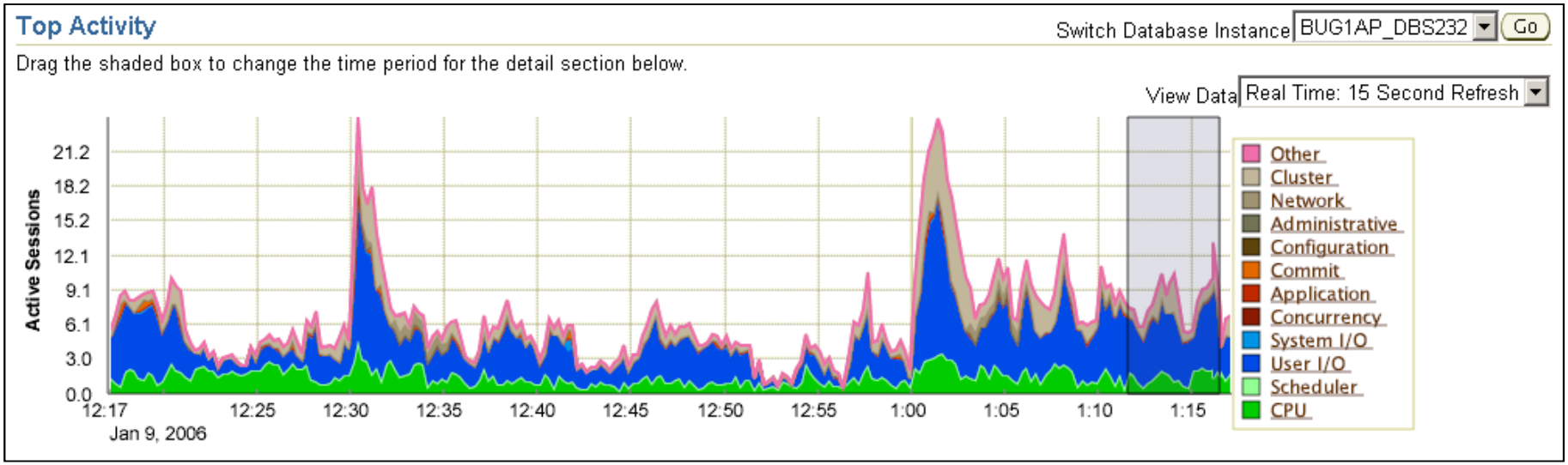
- Identify Performance Bottlenecks Immediately
- Tune SQL like a Pro!
- Visually Tackle Complex SQL Queries
- Simulate Production Environments
- Become Proactive in Data Governance Compliance

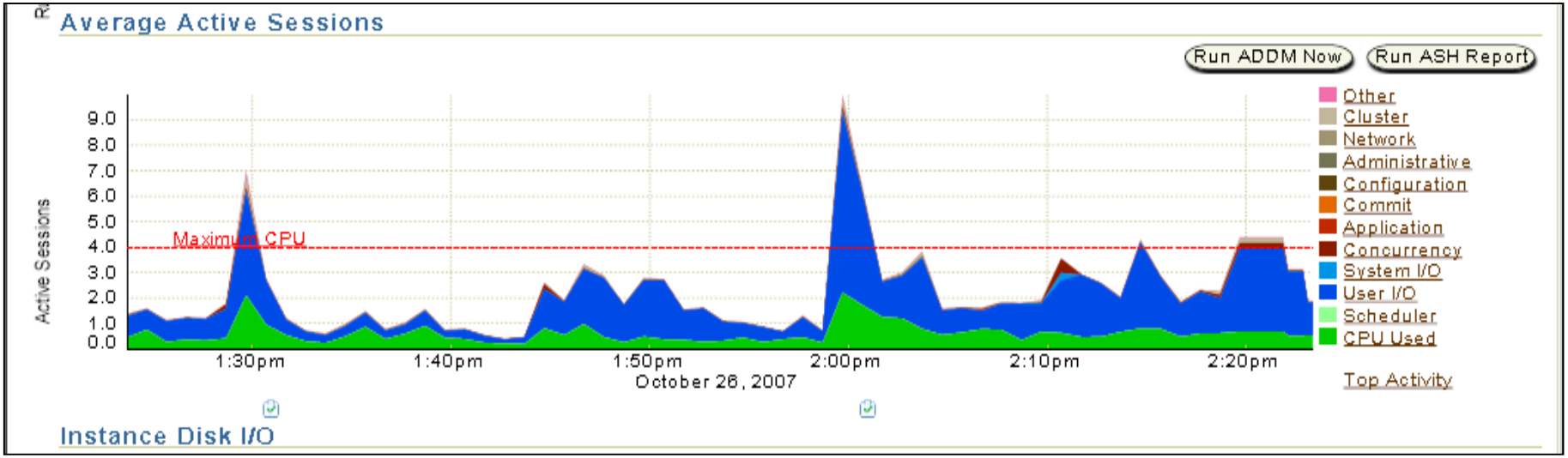
Identify Performance Bottlenecks Immediately

Database profiling provides a graphical visualization of wait-time analysis, making the SQL causing poor database performance easy to pinpoint. Continuous profiling monitors an entire data source within a configurable span of time. Explain plans are provided for a better understanding of how SQL will be executed and the performance costs. Utilize reporting to allow sharing of information captured during the profiling process.



Tools: Oracle's OEM Cloud Control





White Papers: Bryn Llewellyn of Oracle

Doing SQL from PL/SQL:
Best and Worst Practices

An Oracle White Paper
September 2008

ORACLE

White Papers: Quest Software (now Dell)

Overview

What is tuning anyway? A good place to start is to define a baseline so we have a starting ground from which to improve. Secondly we need to understand what our goal is. What exactly are we trying to achieve with performance tuning? The industry's many performance books, papers, best practices, scripts and seminars generally share a set of performance tuning goals:

- Identify and react to performance problems
- Proactively prevent performance issues
- Meet services levels

3

QUEST SOFTWARE

- Avoid hardware upgrades
- Ensure the application responds to customer needs

PTC's Tuning Methodology

- Caveat: There is, it seems, an infinite appetite for performance tuning but little for Change Management

- Step 1: Create a traceable written ticket
- Step 2: Define precisely
  - What is slow and what it touches
  - When did it first become slow
  - What is the frequency of the observation
  - What has changed
- Step 3: Prove it is the database (~50% of all tuning work ends here)
- Step 4: Compare what is happening with a baseline
- Step 5: Fix what is actually wrong (write an RCA doc on the event)
- Step 6: Close the ticket

Root Cause Analysis and RCA Tools

First Principles

- If the database is unavailable it is a database problem
- If the database is slow it is a database problem
- Oracle DBAs are expected to find database problems
- To fix the database so it never happens again
- To write an RCA (Root Cause Analysis) document
- Which will point to the database
- Even when the database isn't the root cause
- Because that is what DBAs are trained to do

- Did I mention ... the network is just fine? 

My Blog: January 10, 2011

If databases were run with the same degree of intelligence and attentiveness as a network router we would:

Log in as SYS, type

```
SELECT dummy FROM dual;
```

and if we did not get an exception, declare everything was fine.

[network admins]..., I might be inclined to recommend that we plug a couple of them into their own networks and see if they light up.

What Is Root Cause Analysis?

- Root Cause Analysis is finding, fixing, and reporting on, the event that precipitated a service related incident
- The incident or change may have resulted in one or more failures that affected database performance and/or availability
- Sometimes the root cause is within the database, for example a bug, it is at least equally probable the database was an innocent bystander

- The point is to avoid repeated fire fighting exercises

Root Cause Analysis Doc (1:7)

Root-Cause Analysis - PROVISIONAL

Node 5 Eviction 2010-12-06 18:32

Author:	Daniel Morgan
Date:	8 Dec 2010
Version:	v 1.0
Status:	Provisional
Reference:	RCA-PROV_201012061832_Node5
Sec. Class:	Commercial in Confidence

Root Cause Analysis Doc (2:7)

Contents

Root Cause Analysis Doc (3:7)

1 Introduction

1.1 Scope

This document is a root-cause analysis for the OPTIMA OPM01P database event that started at on Monday Dec 06 at 18:32:12 2010 on instance opm01p5. It is current a PROVISIONAL assessment and subject to change

1.2 Revision History

Revision Number	Date	Name	Revision
1.0a	8 Dec 2010	Daniel Morgan	Initial Version of Document

1.3 Review/Approval

Reviewer	Date	Feedback
S Bowke	8 Dec 2010	Approved for release as a PROVISIONAL root-cause assessment for discussion purposes

1.4 Outstanding Actions List

This section is to be used to capture the outstanding actions

Outstanding Actions List					
Action No	Priority	Owner	Description	Date Raised	Status
1	1	Daniel Morgan	Issue Full RCA	8 Dec 2010	In-Progress

Root Cause Analysis Doc (4:7)

2 RCA Timeline

At 6.32 Dec 6, Node opm01p5 alert log makes reference to the global resource directory, the global resource directory keeps details of how resources are used within a cluster database.

```
Mon Dec 06 18:32:12 2010
Reconfiguration started (old inc 16, new inc 18)
List of nodes:
4 5 6 7
Global Resource Directory frozen
Communication channels reestablished
Master broadcasted resource hash value bitmaps
Non-local Process blocks cleaned out
Mon Dec 06 18:32:21 2010
LMS 1: 33 GCS shadows cancelled, 1 closed, 0 Xw survived
Mon Dec 06 18:32:21 2010
Mon Dec 06 18:32:21 2010
LMS 0: 1 GCS shadows cancelled, 0 closed, 0 Xw survived
LMS 2: 2 GCS shadows cancelled, 0 closed, 0 Xw survived
Mon Dec 06 18:32:23 2010
Set master node info
Submitted all remote-enqueue requests
```

The LMON was unable to fix the issue and subsequently crashed at 6.37 (several related bugs apparent on metalink)

```
Mon Dec 06 18:37:12 2010
LMS2 (ospid: 13489) has not called a wait for sub 0 secs.
ERROR: GES receivers are not healthy.
ERROR: LMON (ospid: 13477) is terminating the instance.
ERROR: Please check LMON trace file for more detail.
LMON (ospid: 13477): terminating the instance due to error 481
```

This in turn caused the instance to restart

```
USER (ospid: 9864): terminating the instance
Instance terminated by USER, pid = 9864
Mon Dec 06 18:37:47 2010
Starting ORACLE instance (normal)
```

Root Cause Analysis Doc (5:7)

The instance then was unable to open as the global resource directory started to relocate the Global objects (this can take several minutes / hours).

DBAs, Oracle and System Administrators worked on the outage. DBA restarted Node 5 several times throughout the night however each time it came up queries to the GV$views were failing on the system.

At 7:30pm Dec 6 XXX Recommended a full Database bounce to restore the System

At 8:39 PM /var/adm/messages on usp9004a (opm01p5) had the following entry:

```
Dec  6 20:39:51 usp9004a ip: [ID 448101 kern.error] nxge5: <unknown
primitive> failed: DL_UNSUPPORTED
```

At 05:24am Dec 7 a secondary problem occurred on the system as /appl/oracle on usp9004b (opm01p6) was allowed to be filled up with logs and trace files being collected on the system.

Full Database shutdown was executed at 6:50am server time on Dec 7th.
OPTIM_REPORT and OPTIM_REPORT_PWRUSR services were up and available on node 5 at 6:51am (1 minute outage).
All Nodes restarted by 07:05am Dec 7th
All Loaders restarted at 07:07am Dec 7th
All Summaries restarted at 09:54am Dec 7th

Root Cause Analysis Doc (6:7)

3 Initial RCA Statement

With the current settings on the system Node 5 has the highest load in the system. When the resource manager was disabled there were periods when resource manager was off where the system was working at 95% CPU load. User queries were executing on the system with hints having a high degree of parallel requested. It is currently believed that high parallel query at the same time as a high sustained CPU load caused the initial problem leading to the node eviction.

The initial findings point to the fact that by turning off the RM and keeping the parallel max_server fairly high at 300 per instance resulted in Oracle hanging after a period of time as there was an increase in parallel sessions which in turn consumed all available cpu and resulted in timeouts for important tasks like inter node communication or slow response to other nodes.

When the DBAs tried to bring back node 5,the Oracle Database tried to rebalance that again, and caused all the process that access cluster information like GV$ views to hang on "reliable message". This is believed to be a bug previously experienced on the system for Node 2 on 07-Mar-2010 (Bug 9453004: INVESTIGATE DB INSTANCE 2 AND NODE B REBOOT). The inability to rebalance the Database meant the system could not come up cleanly as a 4 node RAC environment which is why a Database Bounce was requested.

Root Cause Analysis Doc (7:7)

4 Conclusion / Next Steps

Await response from Oracle RCA (bug 103776604) following their analysis of trace files/dumps taken during outage.

Recommend that Oracle Bug 9453004 raised March 7 be fixed by Oracle and applied when available.

User parallel query hints should be reduced and if at all possible eliminated. The loads on some nodes are high and may mean that we are under resourced and require extra hardware. Even with Resource Manager in place this may result in sub-optimal query times for users.

Re-assess if PSU need to be applied more regularly to minimize exposure to Oracle Bugs.

It is recommended that if two DBAs are working on the system at the same time, that they are in full communications with each other.

Explore the following potential related bugs:

Bug 6148054 - RAC hang waiting for "reliable message" [ID 6148054.8]

Bug 6773260: ORA-00600 [KOKEGPINLOB1] ON SELECT FROM GV$PX_SESSION, GV$SESSION AND GV$SQLAREA

We Will Examine Some Real-World Cases Today

- Case Study 1: The Puppet Master
- Case Study 2: Lost Blocks
- Case Study 3: 5010 ➡ 7010 Migration
- Case Study 4: Water water everywhere
- Case Study 5: Virtual Success
- Case Study 6: UCS (**U**nimpressive **C**ommon **S**ervers)
- Case Study 7: Port Exhaustion
- Case Study 8: It's RAC
- Case Study 9: Human Nature
- Case Study 10: Jobs

RCA Tools

- ASH Report
- ~~AWR Reports~~
- AWR Difference Reports
- Bash Shell
  - .bash_history
- Data Dictionary
  - DBA_OBJECTS
    - LAST_DDL_TIME
- DBMS_COMPARISON
- DBMS_METADATA.GET_DDL
- DBMS_METADATA_DIFF.COMPARE_ALTER
- OS Watcher

Case Study 1
The Puppet Master

Outage Fingerprint

Two physically separate two-node RAC clusters. They do not share servers. They do not share storage. They do not share network any component of the cache fusion interconnect ... and yet ... in 7 minutes 7 seconds ... they both go down.

Is it a database problem?

DC20PCE11

```
Thu Aug 08 16:52:30 2013 Archived Log entry 215974 added for thread 1 sequence 216019 ID 0x2d7ba8f dest 1:
Thu Aug 08 16:57:27 2013 Time drift detected. Please check VKTM trace file for more details.
Thu Aug 08 16:57:43 2013 ERROR: unrecoverable error ORA-29701 raised in ASM I/O path; terminating process 12257
```

DS20SCE11

```
Thu Aug 08 16:57:17 2013 Completed checkpoint up to RBA [0xae7f.2.10], SCN: 780145612
Thu Aug 08 17:04:34 2013 Time drift detected. Please check VKTM trace file for more details.
Thu Aug 08 17:04:46 2013 ERROR: unrecoverable error ORA-29701 raised in ASM I/O path; terminating process 2445
```

Production: 4 Seconds Earlier

ORAP1N1

```
2013-08-08 16:57:31.162: [    AGFW][1164335424] {0:12:9} Agfw Proxy Server received the message: RESOURCE_STATUS[Proxy]
ID 20481:147794
2013-08-08 16:57:31.162: [    AGFW][1164335424] {0:12:9} Received state change for ora.LISTENER_SCAN2.lsnr 1 1 [old
state = ONLINE, new state = OFFLINE]
```

ORAP1N2

```
2013-08-08 17:09:09.393: [UiServer][1175996736] {2:7473:48658} Done for ctx=0x2aaaac2532b0
2013-08-08 17:09:39.156: [GIPCHDEM][1115060544] gipchaDaemonProcessHAInvalidate: completed ha name invalidate for node
0x2aaaac25bb60 { host 'orap1n1', haName '9f34-b767-de19-a294', srcLuid 04a03a5c-f4851208, dstLuid e3aa430e-82601c00
numInf 2, contigSeq 62781, lastAck 56961, lastValidAck 62780, sendSeq [56961 : 56961], createTime 72155204, flags 0x28 }
```

P1N1 to P1N2 issue delta: 12 minutes 8 seconds

Staging: 4 Hours 11 Minutes Earlier

ORAS1N1

```
2013-08-08 13:04:45.315: [    AGFW][1159891264] {0:4:7} Agfw Proxy Server received the message: RESOURCE_STATUS[Proxy]
ID 20481:508508
2013-08-08 13:04:45.315: [    AGFW][1159891264] {0:4:7} Received state change for ora.asm oras1n1 1 [old state = ONLINE,
new state = UNKNOWN]
```

ORAS1N2

```
2013-08-08 13:12:07.199: [ CRSMAIN][96481872] Sync-up with OCR
2013-08-08 13:12:07.199: [ CRSMAIN][96481872] Connecting to the CSS Daemon
2013-08-08 13:12:07.202: [  CRSRTI][96481872] CSS is not ready. Received status 3
2013-08-08 13:12:07.202: [ CRSMAIN][96481872] Created alert : (:CRSD00109:) :  Could not init the CSS context, error: 3
2013-08-08 13:12:07.202: [    CRSD][96481872][PANIC] CRSD exiting: Could not init the CSS context, error: 3
```

S1N1 to S1N2 issue delta: 7 minutes 22 seconds

OS Log: Four Days Before Incident

```
Aug  4 04:09:16 orap1n1 Updating DNS configuration for: orap1n1.lux20.morgan.priv
Aug  4 04:09:16 orap1n1 Initial DNS Server: 10.2.198.34
Aug  4 04:09:16 orap1n1 Connecting to DNS server 10.2.198.34
Aug  4 04:09:16 orap1n1 Connected to DNS server 10.2.198.34
Aug  4 04:09:16 orap1n1 Updating both HOST and PTR record  for: orap1n1.lux2.morgan.priv
Aug  4 04:09:16 orap1n1 Deleting old reverse lookup records for orap1n1.lux2.morgan.priv on 10.2.198.34.
Aug  4 04:09:17 orap1n1 Adding GSS support to DNS server 10.2.198.34
Aug  4 04:09:17 orap1n1 Added GSS support to DNS server 10.2.198.34
Aug  4 04:09:17 orap1n1 Failed to delete reverse lookup record 11.78.2.10.in-addr.arpa. Reason Refused (5).
Aug  4 04:09:17 orap1n1 Deleting reverse lookup records for our current new IP Address(s) on  ad010.lux20.morgan.priv.
Aug  4 04:09:18 orap1n1 No reverse lookup records found for 11.0.168.192.in-addr.arpa on ad010.ams20.morgan.priv.
Aug  4 04:09:18 orap1n1  No reverse lookup records found for 21.34.254.169.in-addr.arpa on ad010.lux20.morgan.priv.
Aug  4 04:09:19 orap1n1 No reverse lookup records found for 12.0.168.192.in-addr.arpa on ad010.lux20.morgan.priv.
Aug  4 04:09:20 orap1n1 No reverse lookup records found for 181.139.254.169.in-addr.arpa on ad010.lux20.morgan.priv.
Aug  4 04:09:20 orap1n1 Failed to delete reverse lookup record 11.78.2.10.in-addr.arpa. Reason Refused (5).
Aug  4 04:09:21 orap1n1 Failed to delete reverse lookup record 10.78.2.10.in-addr.arpa. Reason Refused (5).
Aug  4 04:09:22 orap1n1 Failed to delete reverse lookup record 102.78.2.10.in-addr.arpa. Reason Refused (5).
Aug  4 04:09:22 orap1n1 Failed to delete reverse lookup record 100.78.2.10.in-addr.arpa. Reason Refused (5).
Aug  4 04:09:23 orap1n1 Failed to delete reverse lookup record 14.2.2.10.in-addr.arpa. Reason Refused (5).
Aug  4 04:09:23 orap1n1 Deleting host records for orap1n1.lux20.morgan.priv on ad010.lux20.morgan.priv.
Aug  4 04:09:23 orap1n1 Failed to delete host record for orap1n1.lux20.morgan.priv. Reason Refused (5).
```

7,824 lines of changes in /var/log/messages on one server
This happened 152 times on ORAP1N1, in DC20, in 6 days

OS Log: Two Days After Incident

```
Aug 10 12:03:23 orap1n1 Updating DNS configuration for: orap1n1.lux20.morgan.priv
Aug 10 12:03:23 orap1n1 Initial DNS Server: 10.2.198.33
Aug 10 12:03:23 orap1n1 Connecting to DNS server 10.2.198.33
Aug 10 12:03:23 orap1n1 Connected to DNS server 10.2.198.33
Aug 10 12:03:24 orap1n1 Updating both HOST and PTR record for: orap1n1.lux20.morgan.priv
Aug 10 12:03:24 orap1n1 Deleting old reverse lookup records for orap1n1.lux20.morgan.priv on 10.2.198.33.
Aug 10 12:03:24 orap1n1 Adding GSS support to DNS server 10.2.198.33
Aug 10 12:03:24 orap1n1 Added GSS support to DNS server 10.2.198.33
Aug 10 12:03:25 orap1n1 Failed to delete reverse lookup record 11.78.2.10.in-addr.arpa. Reason Refused (5).
Aug 10 12:03:25 orap1n1 Deleting reverse lookup records for our current new IP Address(s) on ad009.lux20.morgan.priv.
Aug 10 12:03:25 orap1n1 No reverse lookup records found for 11.0.168.192.in-addr.arpa on ad009.lux20.morgan.priv.
Aug 10 12:03:26 orap1n1 No reverse lookup records found for 21.34.254.169.in-addr.arpa on ad009.lux20.morgan.priv.
Aug 10 12:03:27 orap1n1 No reverse lookup records found for 12.0.168.192.in-addr.arpa on ad009.lux20.morgan.priv.
Aug 10 12:03:27 orap1n1 No reverse lookup records found for 181.139.254.169.in-addr.arpa on ad009.lux20.morgan.priv.
Aug 10 12:03:28 orap1n1 Failed to delete reverse lookup record 11.78.2.10.in-addr.arpa. Reason Refused (5).
Aug 10 12:03:28 orap1n1 Failed to delete reverse lookup record 10.78.2.10.in-addr.arpa. Reason Refused (5).
Aug 10 12:03:29 orap1n1 Failed to delete reverse lookup record 101.78.2.10.in-addr.arpa. Reason Refused (5).
Aug 10 12:03:30 orap1n1 Failed to delete reverse lookup record 14.2.2.10.in-addr.arpa. Reason Refused (5).
Aug 10 12:03:30 orap1n1 Deleting host records for orap1n1.lux20.morgan.priv on ad009.lux20.morgan.priv.
Aug 10 12:03:30 orap1n1 Failed to delete host record for orap1n1.lux20.morgan.priv. Reason Refused (5).
Aug 10 12:03:30 orap1n1 Updating host records for orap1n1.lux20.morgan.priv on ad009.lux20.morgan.priv.
Aug 10 12:03:31 orap1n1 Failed to update host records orap1n1.lux20.morgan.priv: Reason Refused (5).
```

Log File Research

To: <system and storage admins>
Cc: <management>
Subject: Need Hardware Information

Here is the signature of the ASM failure in DC20 over the last two days. Two different databases on different blades:

```
*** 2013-08-09 11:49:20.023
NOTE: ASMB terminating
ORA-15064: communication failure with ASM instance
ORA-03113: end-of-file on communication channel
Process ID:
Session ID: 82 Serial number: 9
error 15064 detected in background process
ORA-15064: communication failure with ASM instance
ORA-03113: end-of-file on communication channel
Process ID:
Session ID: 82 Serial number: 9
kjzduptcctx: Notifying DIAG for crash event
----- Abridged Call Stack Trace -----
ksedsts()+461<-kjzdssdmp()+267<-kjzduptcctx()+232<-kjzdicrshnfy()+53<-ksuitm()+1325<-
ksbrdp()+3344<-opirip()+623<-opidrv()+603<-sou2o()+103<-opimai_real()+266<-ssthrdmain()+252<-
main()+201<-__libc_start_main()+244<-_start()+36
----- End of Abridged Call Stack Trace -----

*** 2013-08-09 11:49:20.134
ASMB (ospid: 15341): terminating the instance due to error 15064
ksuitm: waiting up to [5] seconds before killing DIAG(15317)
```

Can you help me please with the following:

1. Are all database blades in the same or different chassis?
2. What is the storage solution? VSP, NetApp? What diagnostics can we pull?
3. What network infrastructure between the blades and the storage array? What diagnostics can we pull?

OS Log: Ruby on RAC?

```
Aug  8 13:04:22 orap1n1 ERROR:  While executing gem ... (Gem::RemoteFetcher::FetchError)
Aug  8 13:04:22 orap1n1 Errno::ETIMEDOUT: Connection timed out - connect(2) (http://rubygems.org/latest_specs.4.8.gz)
Aug  8 13:04:22 orap1n1 INFO:  `gem install -y` is now default and will be removed
Aug  8 13:04:22 orap1n1 INFO:  use --ignore-dependencies to install only the gems you list



Aug  8 15:42:41 orap1n1 ERROR:  While executing gem ... (Gem::RemoteFetcher::FetchError)
Aug  8 15:42:41 orap1n1 Errno::ETIMEDOUT: Connection timed out - connect(2) (http://rubygems.org/latest_specs.4.8.gz)
Aug  8 15:42:41 orap1n1 INFO:  `gem install -y` is now default and will be removed
Aug  8 15:42:41 orap1n1 INFO:  use --ignore-dependencies to install only the gems you list
```

This happened twice just before the outage
the first one 3 hours 53 seconds before the outage

The second time 1 hour 15 minutes before the outage

OS Log: NTP Time Synchronization

```
Aug  8 12:56:04 orap1n1 ntpd[1339]: ntpd exiting on signal 15
Aug  8 12:57:27 orap1n1 ntpdate[12406]: step time server 10.2.255.254 offset 82.262906 sec
Aug  8 12:57:27 orap1n1 ntpd[12408]: ntpd 4.2.2p1@1.1570-o Fri Jul 22 18:07:53 UTC 2011 (1)
Aug  8 12:57:27 orap1n1 ntpd[12409]: precision = 1.000 usec
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface wildcard, 0.0.0.0#123 Disabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface wildcard, ::#123 Disabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface bond2, fe80::217:a4ff:fe77:fc18#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface lo, ::1#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface bond0, fe80::217:a4ff:fe77:fc10#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface eth2, fe80::217:a4ff:fe77:fc14#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface eth3, fe80::217:a4ff:fe77:fc16#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface lo, 127.0.0.1#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface eth2, 192.168.0.11#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface eth2:1, 169.254.34.21#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface eth3, 192.168.0.12#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface eth3:1, 169.254.139.181#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface bond0, 10.2.78.11#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface bond0:1, 10.2.78.10#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface bond0:3, 10.2.78.102#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface bond0:4, 10.2.78.100#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface bond2, 10.2.2.14#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: kernel time sync status 0040
Aug  8 12:57:27 orap1n1 ntpd[12409]: frequency initialized 0.000 PPM from /var/lib/ntp/drift
```

Network Layer

Network Tuning Topics

- Network Infrastructure
  - Line Cards
  - HBA Cards
  - NIC Cards
- All Databases
  - Kernel Parameters
- RAC Databases
  - Jumbo Frames
  - Network Latency
  - VLANs
- Data Guard Primaries and Standbys
  - Net*Services Parameters

Network Infrastructure

- Network components
  - Switches
  - Routers
  - Hubs
  - Domain Name Servers (DNS)
  - Load Balancers
  - Firewalls
  - Line Cards
  - HBA Cards
  - NIC Cards

Network Considerations (1:2)

- Bandwidth
  - How much data, per unit of time, you can push across the connection
- Jumbo Frames
  - Ethernet frames with more than 1500 bytes of payload
  - Usually 9K packets
  - Oracle blocks are 8K: Do the math
- Keep-Alive
  - A network card (NIC or line) that rather than quickly and cleanly failing attempts to re-establish a connection preventing seamless failover
- Latency
  - The delay introduced between the point where the data originates and the receiving point

Network Considerations (2:2)

- TCP/IP Offloading
  - A technology used in NIC cards to offload processing of the entire TCP/IP stack to the network controller
  - Primarily seen in 10gEth networks
- Virtualization (VLAN)
  - Not all VLANs are equal
  - Your network admins have no idea how different types of VLAN affect Oracle RAC

Protocols

- FTP (File Transfer Protocol)
  - Used by Oracle for public interfaces
  - When transmitted over a network a database block is broken up into packets, sent in order, received in order, reassembled in order
- UDP (User Datagram Protocol)
  - Used by Oracle for the Cache Fusion RAC Interconnect
  - When transmitted over a network packets are broken up if necessary, transmitted in random order, and reassembled if found in the receive cache
  - If not found in the receive cache a failure may result

Stand-Alone Databases (1:2)

- Network Related Kernel Parameters

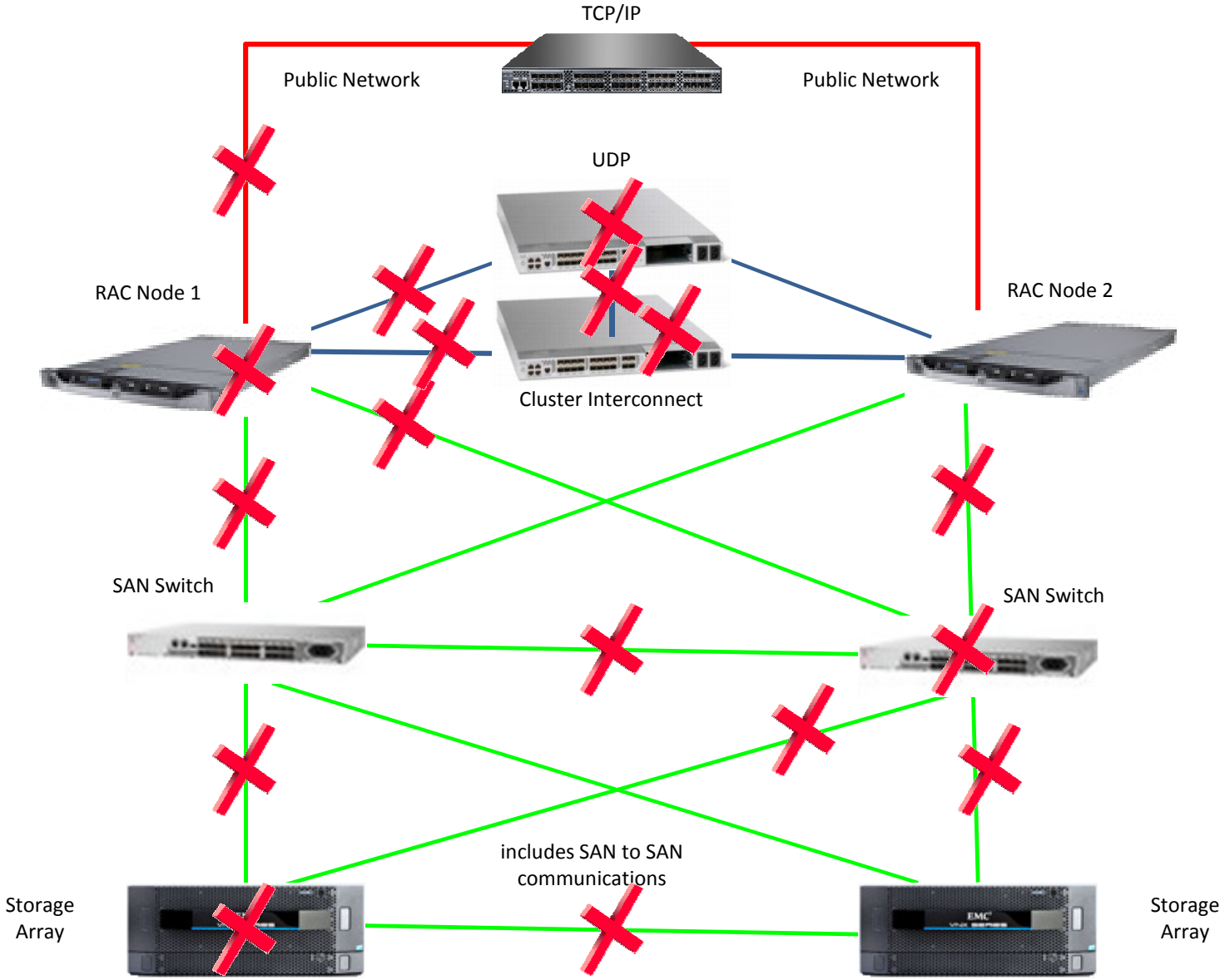
```
cat >> /etc/sysctl.conf <<EOF
fs.file-max = 6815744
fs.aio-max-nr = 3145728
kernel.shmall = 2097152
kernel.shmmax = 2147483648
kernel.shmmni = 4096
kernel.sem = 250 32000 100 128
net.core.rmem_default = 4194304
net.core.rmem_max = 4194304
net.core.wmem_default = 262144
net.core.wmem_max = 1048576
net.ipv4.ip_forward = 0
net.ipv4.conf.default.rp_filter = 1
tcp.ipv4.tcp_wmem = 262144 262144 262144
tcp.ipv4.tcp_rmem = 4194304 4194304 4194304
net.ipv4.ip_local_port_range = 9000 65000
EOF

-- activate changes
/sbin/sysctl -p
```

Stand-Alone Databases (2:2)

- Optional kernel parameters can improve TCP/IP performance

```
--enable TCP kernel auto-tuning
/proc/sys/net/ipv4/tcp_moderate_rcvbuf (1=on)

-- tune TCP max memory: tune to 2xBDP (Bandwidth x Delay Product)
-- For example, with 40 Mbits/sec bandwidth, 25 msec delay,
-- BDP = (40 x 1000 / 8 Kbytes/sec) x (0.025 sec) ~ 128 Kbytes
/proc/sys/net/ipv4/tcp_rmem and tcp_wmem 4096 87380 174760

-- tune the socket buffer sizes by setting to 2xBDP
/proc/sys/net/core/rmem_max and wmem_max

-- ensure that TCP Performance features are enabled
/proc/sys/net/ipv4/tcp_sack
/proc/sys/net/ipv4/tcp_window_scaling
/proc/sys/net/ipv4/tcp_timestamps
```

RAC (1:12)


TCP/IP
Public Network
Public Network
UDP
RAC Node 1
RAC Node 2
Cluster Interconnect
SAN Switch
SAN Switch
includes SAN to SAN communications
Storage Array
Storage Array

RAC (2:12)

- Jumbo Frames
  - MTU = 9000 except Linux (8972)
  - Any MTU greater than 9000 will lead to failure
- Routable vs non-routable IP addresses
- ~~V~~LANs
- AWR Report
  - Blocks sent and received
  - Receive time latency

```
# ping -M do -s 8972 10.0.0.4
(-M do = do not fragment)

# traceroute -U 10.0.0.4 -F 8972 -v
(U = UDP, s = source address, F = don't fragment, v = verbose output
```

RAC (3:12) Solaris

- Modify interconnect and set UDP High Water Marks to prevent buffer overflow

The User Data Protocol (UDP) parameter settings define the amount of send and receive buffer space for sending and receiving datagrams over an IP network. These settings affect cluster interconnect transmissions. If the buffers set by these parameters are too small, then incoming UDP datagrams can be dropped due to insufficient space, which requires send-side retransmission. This can result in poor cluster performance.
On Oracle Solaris, the UDP parameters are udp_recv_hiwat and udp_xmit_hiwat. The default values for these parameters are 57344 bytes. Oracle recommends that you set these parameters to at least 65536 bytes.

To check current settings for udp_recv_hiwat and udp_xmit_hiwat, enter the following commands:

```
# ndd /dev/udp udp_xmit_hiwat # ndd /dev/udp udp_recv_hiwat
```

To set the values of these parameters to 65536 bytes in current memory, enter the following commands:

```
# ndd -set /dev/udp udp_xmit_hiwat 65536
# ndd -set /dev/udp udp_recv_hiwat 65536
```

To set the UDP values for when the system restarts, the ndd commands have to be included in a system startup script. For example, The following script in /etc/rc2.d/S99ndd sets the parameters:

```
ndd -set /dev/udp udp_xmit_hiwat 65536
ndd -set /dev/udp udp_recv_hiwat 65536
```

RAC (4:12) AIX

- Modify interconnect UDP parameters

```
/usr/sbin/no -r -o ipqmaxlen=512
/usr/sbin/no -p -o udp_sendspace=65536
/usr/sbin/no -p -o udp_recvspace=655360
/usr/sbin/no -p -o tcp_sendspace=65536
/usr/sbin/no -p -o tcp_recvspace=65536
/usr/sbin/no -p -o rfc_1323=1
/usr/sbin/no -p -o sb_max=4194304
```

RAC (5:12) Jumbo Frames

```
bond0     Link encap:Ethernet  HWaddr 00:17:A4:77:08:E4
          inet addr:10.20.70.11  Bcast:10.20.70.255  Mask:255.255.255.0
          inet6 addr: fe80::217:a4ff:fe77:8e4/64 Scope:Link
          UP BROADCAST RUNNING MASTER MULTICAST  MTU:1500  Metric:1
          RX packets:13362219715 errors:0 dropped:0 overruns:0 frame:0
          TX packets:12741281318 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:3013285660943 (2.7 TiB)  TX bytes:6197747083306 (5.6 TiB)

bond0:1   Link encap:Ethernet  HWaddr 00:17:A4:77:08:E4
          inet addr:10.20.70.101  Bcast:10.20.70.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MASTER MULTICAST  MTU:1500  Metric:1

bond0:2   Link encap:Ethernet  HWaddr 00:17:A4:77:08:E4
          inet addr:10.20.70.10  Bcast:10.20.70.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MASTER MULTICAST  MTU:1500  Metric:1

bond1     Link encap:Ethernet  HWaddr 00:17:A4:77:08:E8
          inet addr:10.20.72.10  Bcast:10.20.72.255  Mask:255.255.255.0
          inet6 addr: fe80::217:a4ff:fe77:8e8/64 Scope:Link
          UP BROADCAST RUNNING MASTER MULTICAST  MTU:1500  Metric:1
          RX packets:27608445622 errors:0 dropped:0 overruns:0 frame:0
          TX packets:26908549838 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:22349990024684 (20.3 TiB)  TX bytes:21343299037950 (19.4 TiB)

bond1:1   Link encap:Ethernet  HWaddr 00:17:A4:77:08:E8
          inet addr:169.254.85.254  Bcast:169.254.255.255  Mask:255.255.0.0
          UP BROADCAST RUNNING MASTER MULTICAST  MTU:1500  Metric:1

bond2     Link encap:Ethernet  HWaddr 00:17:A4:77:08:EC
          inet addr:10.20.254.26  Bcast:10.20.254.255  Mask:255.255.255.0
          inet6 addr: fe80::217:a4ff:fe77:8ec/64 Scope:Link
          UP BROADCAST RUNNING MASTER MULTICAST  MTU:9000  Metric:1
          RX packets:259916305 errors:0 dropped:0 overruns:0 frame:0
          TX packets:226620202 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:1126764035639 (1.0 TiB)  TX bytes:643151981505 (598.9 GiB)
```

```
bond3     Link encap:Ethernet  HWaddr 00:17:A4:77:08:F0
          inet addr:10.24.50.20  Bcast:10.24.50.255  Mask:255.255.255.0
          inet6 addr: fe80::217:a4ff:fe77:8f0/64 Scope:Link
          UP BROADCAST RUNNING MASTER MULTICAST  MTU:1500  Metric:1
          RX packets:5322637832 errors:0 dropped:0 overruns:0 frame:0
          TX packets:20900487493 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:341114387221 (317.6 GiB)  TX bytes:31650497491000 (28.7 TiB)

eth0      Link encap:Ethernet  HWaddr 00:17:A4:77:08:E4
          UP BROADCAST RUNNING SLAVE MULTICAST  MTU:1500  Metric:1
          RX packets:13362072800 errors:0 dropped:0 overruns:0 frame:0
          TX packets:12741281318 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:3013265505816 (2.7 TiB)  TX bytes:6197747083306 (5.6 TiB)
          Interrupt:114 Memory:fb000000-fb7fffff

eth1      Link encap:Ethernet  HWaddr 00:17:A4:77:08:E4
          UP BROADCAST RUNNING SLAVE MULTICAST  MTU:1500  Metric:1
          RX packets:146915 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:20155127 (19.2 MiB)  TX bytes:0 (0.0 b)
          Interrupt:122 Memory:fa000000-fa7fffff

eth2      Link encap:Ethernet  HWaddr 00:17:A4:77:08:E8
          UP BROADCAST RUNNING SLAVE MULTICAST  MTU:1500  Metric:1
          RX packets:27608445572 errors:0 dropped:0 overruns:0 frame:0
          TX packets:26908549838 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:22349990008504 (20.3 TiB)  TX bytes:21343299037950 (19.4 TiB)
          Interrupt:122 Memory:f9000000-f97fffff

eth3      Link encap:Ethernet  HWaddr 00:17:A4:77:08:E8
          UP BROADCAST RUNNING SLAVE MULTICAST  MTU:1500  Metric:1
          RX packets:50 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:16180 (15.8 KiB)  TX bytes:0 (0.0 b)
          Interrupt:130 Memory:f8000000-f87fffff
```

Case Study 2
Lost Blocks

RAC (6:12) Cluster Interconnect

Ticket: 39398: FEDPRD1 oradb_server_generated_alerts_custom
Possible poor interconnect configuration. Will need to follow Metalink note "Troubleshooting gc block lost and Poor Network Performance in a RAC Environment [ID 563566.1]" if this problem persists.

According to the ticket, the performance metric “Global Cache Blocks Lost” reached 5. When I researched Global Cache Blocks Lost, it indicated an interconnect issue and led me to the Metalink Note I mentioned. I ran an ADDM report from around that time and there were many suggestions to improve performance.

I did notice that Global Cache Blocks Lost reached 5 on April 1, 2013 as well.

FINDING 6: 3.5% impact (29 seconds)

Inter-instance messaging was consuming significant database time on this instance.
SYMPTOM: Wait class "Cluster" was consuming significant database time. (3.5% impact [29 seconds])

Ticket 39721 oradb_server_generated_alerts_custom FEDPRD2

Global Cache Blocks Lost" is at 3. Connected and found event had passed.

RAC (7:12) VLANs

RAC: Frequently Asked Questions [ID 220970.1] Modified:Apr 29, 2013 Type:FAQ Status:PUBLISHED Priority:1

Are there any issues for the interconnect when sharing the same switch as the public network by using VLAN to separate the network?

Oracle RAC and Oracle Clusterware deployment best practices recommend that the interconnect be deployed on a stand-alone, physically separate, dedicated switch.

Many customers, however, have consolidated these stand-alone switches into larger managed switches. A consequence of this consolidation is a merging of IP networks on a single shared switch, segmented by VLANs. There are caveats associated with such deployments.

The Oracle RAC cache fusion protocol exercises the IP network more rigorously than non-RAC Oracle databases. The latency and bandwidth requirements as well as availability requirements of the Oracle RAC / Oracle Clusterware interconnect IP network are more in-line with high performance computing.

Deploying the Oracle RAC / Oracle Clusterware interconnect on a shared switch, segmented by a VLAN may expose the interconnect links to congestion and instability in the larger IP network topology.

If deploying the interconnect on a VLAN, there should be a 1:1 mapping of the VLAN to a non-routable subnet and the VLAN should not span multiple VLANs (tagged) or multiple switches.

Deployment concerns in this environment include Spanning Tree loops when the larger IP network topology changes, Asymetrical routing that may cause packet flooding, and lack of fine grained monitoring of the VLAN/port.

RAC (8:12) Jumbo Frames

Recommendation for the Real Application Cluster Interconnect and Jumbo Frames [ID 341788.1]

Purpose
This note covers the current recommendation for the Real Application Cluster Interconnect and Jumbo Frames

Scope
This article points out the issues surrounding Ethernet Jumbo Frame usage for the Oracle Real Application Cluster (RAC) Interconnect. In Oracle Real Application Clusters, the Cluster Interconnect is designed to run on a dedicated, or stand-alone network. The Interconnect is designed to carry the communication between the nodes in the Cluster needed to check for the Clusters condition and to synchronize the various memory caches used by the database.

Testing
Traceroute: Notice the 9000 packet goes through with no error, while the 9001 fails, this is a correct configuration that supports a message of up to 9000 bytes with no fragmentation:

```
[node01] $ traceroute -F node02-priv 9000
traceroute to node02-priv (10.10.10.2), 30 hops max, 9000 byte packets
1 node02-priv (10.10.10.2) 0.232 ms 0.176 ms 0.160 ms

[node01] $ traceroute -F node02-priv 9001
traceroute to node02-priv (10.10.10.2), 30 hops max, 9001 byte packets
traceroute: sendto: Message too long
1 traceroute: wrote node02-priv 9001 chars, ret=-1
```

RAC (9:12) Cluster Interconnect

Cluster Interconnect in Oracle 10g and 11gR1 RAC [ID 787420.1]

Interconnect High Availability

It is an Oracle Best Practice to make the private interconnect highly available. Depending on the operating system and vendor you can use OS network drivers such as bonding, teaming, IPMP, Etherchannel, APA, MultiPrivNIC. For that purpose the interconnect should be configured across two NICs as well as two switches for complete redundancy so that the cluster can survive a single point of failure.

The setup of the highly available interconnect is transparent to Oracle Clusterware and RAC. This means that an underlying failure is handled by the OS or networking drivers that manage the interfaces. Oracle software does not recognize the change underneath, because the failover is handled by the operating system transparently to Oracle.

RAC Assurance Support Team: RAC and Oracle Clusterware Starter Kit and Best Practices (Generic) [Doc ID 810394.1]

Ensure that all network cables are terminated in a grounded socket. A switch is required for the private network. Use dedicated redundant switches for private interconnect and VLAN considerations. RAC and Clusterware deployment best practices recommend that the interconnection be deployed on a stand-alone, physically separate, dedicated switch.

Jumbo Frames for the private interconnect is a recommended best practice for enhanced performance of cache fusion operations. Reference: Document 341788.1

Use non-routable network addresses for private interconnect; Class A: 10.0.0.0 to 10.255.255.255, Class B: 172.16.0.0 to 172.31.255.255, Class C: 192.168.0.0 to 192.168.255.255. Refer to RFC1918 and Document 338924.1 for additional information.

RAC (10:12) Node Evictions

Top 5 Issues That Cause Node Reboots or Evictions or Unexpected Recycle of CRS [ID 1367153.1]

Issue #2: The node rebooted because it was evicted due to missing network heartbeats.

This is due to missing network heartbeat or split brain condition. In two node environment, repeated reboots of node 2 normally means that node 2 is evicted due to split brain. The ocssd.log shows missing network heartbeat or a split brain message before the node is rebooted.

Cause: the network communication over private interconnect between nodes failed. The failure can be uni-directional or bi-directional.

Solution: Fix the network problem. Make sure all network components like switch and NIC cards are working. Make sure ssh works over the private interconnect. Note that the network often works again after the node is rebooted.

Note: If jumbo frames are used, please refer to the note 341788.1 (Recommendation for the Real Application Cluster Interconnect and Jumbo Frames). The network problem can occur and lead to node evictions or CRS start up problems if the switch is not set up properly to match the MTU (jumbo frame) set up of the NIC cards. Sometimes, the switch and NIC cards from different vendors do not provide the same support for jumbo frames.

RAC (11:12) VLANs

"Dedicated VLANs are what we are doing on the Nexus 7010s and what we did on the 5010s. The key to making it run reliably though ... is not to span line cards or allow the VLAN off the switches. Put all the tagged VLAN ports on a single line card.

So ... assuming you have dual interconnect ... I'd pick one line card on one switch ... and use X ports on that card. I'd pick a line card on another switch and use X ports on it for redundancy.

Don't allow those VLANs to be propagated off the switches (i.e. non-trunked VLANs as that would allow traffic to get to the distribution layer. If you need to span switches, interconnect the switches with a trunk that only carries the VLAN in question."

RAC (12:12)

Published by Oracle in the Technology Network forums

1. Daniel Morgan ... (yes: Oracle publishes my opinion on this subject)
   Using the same switch for public, storage, and cache fusion interconnect, no matter what the network admins claim, is a failure waiting to happen. Cache fusion interconnect traffic must be on its own private switch.

2. Hans Forbrich (Oracle University Instructor of the year 2010)
   The interconnect is not just used for heartbeat. It is the path on which the background processes perform the memory-to-memory data transfers AND lock management. Any application locking, wait state, or resource issues will be amplified by slow-downs on the interconnect.

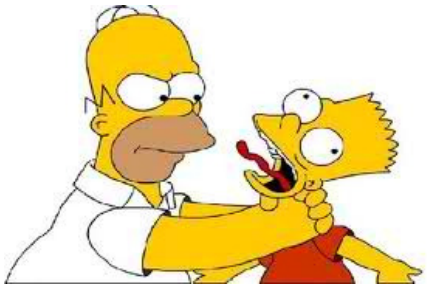
Data Guard

- Service Data Unit (SDU)
  - Unit of data prior to encapsulation
- Send Buffer and Receive Buffer Size
  - Tuned in Linux with the kernel parameters wmem and rmem
  - Tuned in Solaris with the kernel parameters wsize and rsize

```
DGLOGSHIPB =
  (DESCRIPTION =
    (SDU = 32767)
    (SEND_BUF_SIZE=9375000)
    (RECV_BUF_SIZE=9375000)
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = TCP)(HOST = 10.0.#.2)(PORT = 1526))
    )
   (CONNECT_DATA =
      (SERVICE_NAME = prodb)
    )
  )
```

OVN: Oracle Virtual Network

- One network solution to consider
  - Based on Oracle's purchase of Xsigo
  - Oracle Fabric Interconnect simplifies data center management by letting users dynamically connect any server to any network and storage
  - The result is up to 100 times faster server management, 70 percent fewer cables and cards, and 50 percent less capital cost
- It isn't Cisco so the network admins will be terrified but consider the possibility for your database infrastructure of something that recognizes our requirements for low latency, high bandwidth, and stability



Case Study 3
5010 <> 7010

What You Can't See Matters

- 6 Node RAC Cluster ... each node is a Sun M9000
- Storage is 3 clustered EMC VMax arrays = 1.25PB
- Networks are bonded and redundant 10gEth
- The initial interconnect switch is a single Cisco 5010
- The new interconnect will be redundant Cisco 7010s

WORKLOAD REPOSITORY report for

DB Name	DB Id	Instance	Inst num	Startup Time	Release	RAC
OPM01P	782247420	opm01p6	6	18-Aug-10 21:08	11.1.0.7.0	YES

Host Name	Platform	CPUs	Cores	Sockets	Memory (GB)
usp9004b	Solaris[tm] OE (64-bit)	128	64	16	503.16

	Snap Id	Snap Time	Sessions	Cursors/Session
Begin Snap:	7037	15-Sep-10 13:00:18	406	7.5
End Snap:	7038	15-Sep-10 14:01:28	318	8.5
Elapsed:		61.17 (mins)		
DB Time:		6,076.88 (mins)		

- What could possibly go wrong?

Storage Layer

Storage Topics

- Do you have enough spinning disk?
- Storage Virtualization
- Are you taking advantage of flash?
- SANE: Strip and Mirror Everything
- RAID
- Shared Services

Case Study 4
Water water everywhere but not a drop to drink

Processing Stops and the NOC says

[Ticket] Commented: (1246816) mount points filled 100% on dc1laxdb01 and dc1laxdb03

Hi,

Two mounts got filled 100%, please add space as early as possible.
/u108 on dc1laxdb01
/export/home on dc1laxdb03

There are only datafiles in both mount points,

Thanks
Murphy

Ticket 1246816

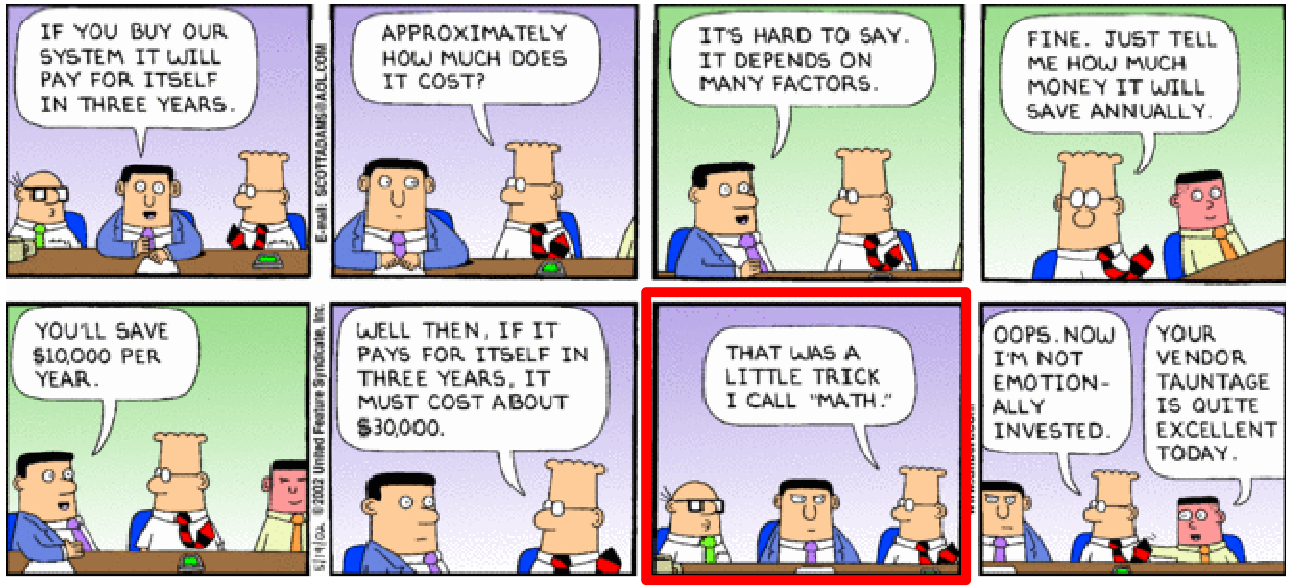
```
-bash-3.00$ df -h
Filesystem              size   used  avail capacity  Mounted on
/dev/md/dsk/d100         37G    11G    26G    29%    /
/devices                  0K     0K     0K     0%    /devices
ctfs                      0K     0K     0K     0%    /system/contract
proc                      0K     0K     0K     0%    /proc
mnttab                    0K     0K     0K     0%    /etc/mnttab
swap                     61G   2.1M    61G     1%    /etc/svc/volatile
objfs                     0K     0K     0K     0%    /system/object
sharefs                   0K     0K     0K     0%    /etc/dfs/sharetab
fd                        0K     0K     0K     0%    /dev/fd
/dev/md/dsk/d500         20G   4.6G    15G    24%    /var
swap                     62G   1.4G    61G     3%    /tmp
swap                     61G   142M    61G     1%    /var/run
/dev/dsk/c6t600601606AD11900E033B69AFA43DD11d0s2
                        115G    46G    68G    41%    /u01
/dev/md/dsk/d132         31G   2.2G    29G     8%    /var/crash
/dev/md/dsk/d60         9.8G   6.4G   3.3G    66%    /export/home
/dev/md/dsk/d402        422M   5.1M   374M     2%    /global/.devices/node@2
/dev/md/dsk/d404        481M   5.0M   428M     2%    /global/.devices/node@4
/dev/md/dsk/d401        415M    74M   299M    20%    /global/.devices/node@1
/dev/md/dsk/d403        481M   5.0M   428M     2%    /global/.devices/node@3
/dev/md/sf14/dsk/d112  4.2T   4.1T    34G   100%    /u112
/dev/md/sf14/dsk/d101  2.1T   2.0T    52G    98%    /u101
/dev/md/sf14/dsk/d109  2.1T   1.8T   239G    89%    /u109
/dev/md/sf14/dsk/d111  197G   3.5G   191G     2%    /u111
/dev/md/sf14/dsk/d100  2.1T   2.0T    31G    99%    /u100
/dev/md/sf14/dsk/d107  264G    73G   188G    28%    /u107
/dev/md/sf14/dsk/d102  1.0T  1005G    58G    95%    /u102
/dev/md/sf14/dsk/d106  264G    36G   225G    14%    /u106
/dev/md/sf14/dsk/d113  4.0T   3.6T   326G    92%    /u113
/dev/md/sf14/dsk/d110  3.0T   946G   2.0T    32%    /u110_arch
/dev/md/sf14/dsk/d104  2.0T   1.9T    37G    99%    /u104
/dev/md/sf14/dsk/d105  2.0T   2.0T   537M   100%    /u105
/dev/md/sf14/dsk/d108  2.0T   2.0T   2.0G   100%    /u108
/dev/md/sf14/dsk/d103  2.0T   1.9T    47G    98%    /u103
```

Storage Admin Tauntage: Let's Do Some Math

Allocated	Available
31	29
10	3
4200	34
2100	52
2100	239
197	191
2100	31
264	188
1000	58
264	225
4000	326
3000	2000
2000	37
2000	1
2000	2000
2000	47
27,266	**5,461**

The database is stopped because "they are out of space."

Yet 20% of the space allocated has never been used.



And That's Not Counting Free Space

```
SQL> select file_name, tablespace_name
  2  from dba_data_files
  3  where autoextensible = 'YES'
  4  order by 1;

FILE_NAME                                                TABLESPACE_NAME
-------------------------------------------------------- ----------------
/u113/oradata/SF14/datafile/o1_mf_lob_01_8jlsmo05_.dbf  LOB_01
/u113/oradata/SF14/datafile/o1_mf_lob_01_8jlst7ky_.dbf  LOB_01
/u113/oradata/SF14/datafile/o1_mf_lob_01_8jlsx6fr_.dbf  LOB_01
/u113/oradata/SF14/datafile/o1_mf_lob_01_8jlt035w_.dbf  LOB_01
/u113/oradata/SF14/datafile/o1_mf_lob_01_8jlt34sd_.dbf  LOB_01
/u113/oradata/SF14/datafile/o1_mf_lob_01_8rs5xndc_.dbf  LOB_01
/u113/oradata/SF14/datafile/o1_mf_lob_01_8vdx8bps_.dbf  LOB_01
/u113/oradata/SF14/datafile/o1_mf_lob_01_8vdx9r68_.dbf  LOB_01
/u113/oradata/SF14/datafile/o1_mf_lob_01_8vdxc5ks_.dbf  LOB_01
/u113/oradata/SF14/datafile/o1_mf_lob_01_8vdxf9v1_.dbf  LOB_01

10 rows selected.


SQL> select sum(bytes)/1024/1024/1024 FREE_SPACE
  2  from dba_free_space
  3  where tablespace_name = 'LOB_01';

FREE_SPACE
----------
6166.08484
```

Case Study 5
Virtual Success

Virtual Success

Consider the following scenario.

You have 12TB of disk available for redo logs in your FRA.

You have three databases each of which generates about 3TB of archived redo per day (3x3 = 9) leaving you 3TB of available disk. The storage admins want to virtualize the allocation.

Do you give each database 4TB or do you get creative with virtualization and give each of them 6TB?

```
3+3+3= 9
3+3+4=10
3+3+5=11
3+4+4=12
3+4+5=12
4+4+4=12
```

Even if there is a big data load on one of them it won't exceed 12TB: What could possibly go wrong?

Non-Spinning Persistent Storage

Comparing Oracle Database Appliance and Exadata

Sep 21, 2011 / By Marc Fielding

Tags: Group Blog Posts, Oracle, Oracle Exadata, Technical Blog

Oracle Database Appliance (ODA) is in many ways similar to an Exadata quarter rack: they both use two similar compute servers in an engineered system configuration, with shared storage and flash storage. But in other ways, especially networking and storage, they differ significantly. In particular:

- *Storage connectivity*: ODA uses SAS direct attachment, while Exadata uses an InfiniBand backbone connected to dedicated storage servers
- *Flash memory*: Exadata has significantly more flash memory than ODA does. While both can store ASM diskgroups in flash, Exadata also has flash cache capability. ODA's default configuration uses flash to store redo logs; this helps compensate for ODA's lack of
battery-backed disk write cache in the latency-sensitive redo write workload.

SANE & RAID

- SAME
  - Stripe And Mirror Everything
- RAID 0
  - Stripe (no parity or mirror)
- RAID 1
  - Mirror
- RAID 0+1
  - Mirror of stripes
- RAID 1+0
  - Stripe of mirrors
  - performance identical to 0+1
  - capacity identical to 0+1
  - higher fault tolerance than 0+1

Battle Against Any RAID

Why should I not use RAID 5?
Art S. Kagel

Sane SAN
James Morle Scale Abilities

Is RAID 5 Really a Bargain?
Cary Millsap Method R

Configuring Oracle Server for VLDB
Cary Millsap Method R

Finally! The RAID F Simulator is here !
Create by Jesper Haure Norrevang

Juan Loaiza on raid V
By Juan Loaiza

Optimal storage configuration made easy
By Juan Loaiza

Server Layer

Server Topics

- Hardware
- CPU: The term no longer has any meaning
  - sockets
  - cores
  - threads
- Memory
- Internal Buses
- Slots
- Blade Servers

Hardware

- Can we use our current servers?
  Or more importantly "should" we use our current servers
  - Pizza Boxes (1U, 2U, 3U)
  - Blades (H/P and Cisco UCS)
  - M10 (Fujitsu - Solaris)
  - P Series (IBM - AIX)
  - Z Series (IBM - zLinux)
  - T Series (Sun - Oracle)
  - M Series (Sun - Oracle)
  - ODA (Oracle - Linux)
  - Exadata (Oracle - Linux)

SMP vs NUMA

- SMP (Symmetric Multi-Processor)
  - Multiple CPUs
  - Shared Memory Bus
  - One Memory Controller
  - All CPUs see the same memory, thus "Symmetric"
- NUMA (Non-Uniform Memory Architecture)
  - Multiple CPUs
  - Multiple Memory Controllers (One or more CPUs / Memory Controller)
  - Memory Bus's shared between shared memory controllers
  - Different CPUs see different memory "Non-Uniform"
  - The Oracle database, by default, is not NUMA aware

Case Study 6
Cisco UCS

Guaranteeing Failure Before The Servers Arrive

- Network stability is critical to Oracle DBAs
- If you have network issues you can waste staggering amounts of time proving the issue isn't the database
- In 2013 I worked with Cisco UCS
  - ~10 databases stand-alone 11gR2
  - ~75 RAC Active-Active or Clusterware Active-Passive Failover
- The questions that needed to be addressed was
  - What is the value of failover to a cluster?
  - What is the value of functioning network diagnostics?

VLANs and the Cluster Interconnect

- It is essentially impossible do what is recommended in Oracle Support's "best practices" guidelines for RAC with blades: any blades from any vendor

RAC: Frequently Asked Questions (Doc ID 220970.1)

Cluster interconnect network separation can be satisfied either by using standalone, dedicated switches, which provide the highest degree of network isolation, or Virtual Local Area Networks defined on the Ethernet switch, which provide broadcast domain isolation between IP networks. VLANs are fully supported for Oracle Clusterware interconnect deployments. Partitioning the Ethernet switch with VLANs allows for:
- Sharing the same switch for private and public communication.
- Sharing the same switch for the private communication of more than one cluster.
- Sharing the same switch for private communication and shared storage access.

The following best practices should be followed:

The Cluster Interconnect VLAN must be on a non-routed IP subnet.

All Cluster Interconnect networks must be configured with non-routed IPs. The server-server communication should be single hop through the switch via the interconnect VLAN. There is no VLAN-VLAN communication.

Oracle recommends maintaining a 1:1 mapping of subnet to VLAN.

The most common VLAN deployments maintain a 1:1 mapping of subnet to VLAN. It is strongly recommended to avoid multi-subnet mapping to a single VLAN. Best practice recommends a single access VLAN port configured on the switch for the cluster interconnect VLAN. The server side network interface should have access to a single VLAN.

VLANs and the Cluster Interconnect

- It is extremely difficult to troubleshoot interconnect issues with UCS as the tools for separating public, storage, and fusion interconnect packets do not exist

Troubleshooting gc block lost and Poor Network Performance in a RAC Environment (Doc ID 563566.1)

6. Interconnect LAN non-dedicated
Description: Shared public IP traffic and/or shared NAS IP traffic, configured on the interconnect LAN will result in degraded application performance, network congestion and, in extreme cases, global cache block loss.

Action: The interconnect/clusterware traffic should be on a dedicated LAN defined by a non-routed subnet. Interconnect traffic should be isolated to the adjacent switch(es), e.g. interconnect traffic should not extend beyond the access layer switch(es) to which the links are attached. The interconnect traffic should not be shared with public or NAS traffic. If Virtual LANs (VLANS) are used, the interconnect should be on a single, dedicated VLAN mapped to a dedicated, non-routed subnet, which is isolated from public or NAS traffic.

My Experience

- Blade servers, of which Cisco UCS is one example, do not have sufficient independent buses and NICs to avoid the networking becoming a single point of failure
- It is good when the public interface has a "keep alive" enabled but this is a fatal flaw for the cluster interconnect
- When different types of packets: public, storage, and interconnect are mixed low-level diagnostics are difficult if not impossible
- When different types of packets, public, storage, and interconnect are mixed the latency of one is the latency of all
- Traffic from any one blade can impact all blades

Conclusion

- Blade servers may be a good solution for application and web servers
- They are acceptable for some stand-alone databases
- Do not buy blade servers if
  - High availability is the goal
  - RAC the technology for achieving it
  - Performance and stability are important
- Buy blade servers if
  - You are deploying small to medium size stand-alone databases
  - You are deploying RAC and getting paid by the hour
  - You are deploying RAC and your single biggest unfulfilled ambition in life is to work a service request with Oracle Support while being verbally abused by your management

Virtual Machine Layer

Virtual Machine Topics

- Virtual machines (VMs) require resources and configuration
  - CPU Resources
  - Interrupt Coalescing
  - Chipset Power Management
  - Memory Allocation
- Read the docs: They are all different

Virtual Machines

- Many choices ... which one is right for you?
  - Citrix
  - IBM LPar
  - IBM zVM
  - Microsoft HyperV Oracle VM (OVM)
  - Microsoft Virtual Server
  - Oracle Linux Containers
  - Redhat Enterprise Virtualization
  - Solaris Containers
  - VMWare vSphere

Only Oracle and IBM virtualization technology is supported for RAC ... and it matters

OVM

- For zLinux and AIX nothing is better than zVM and LPARs
- For non-IBM operating systems use OVM
  - One throat to choke
  - Templates
- Or, for Linux and Solaris, use operating system containers

- E-Business Suite 12.1.1
- E-Business Suite 12.X Sparse Middle Tier
- JD Edwards EnterpriseOne 9.1 Update 2 and JD Edwards EnterpriseOne Tools 9.1.4.4 for x86 servers, Oracle Exalogic, and Oracle Database Appliance
- PeopleSoft PeopleTools 8.53.01 for Oracle Exalogic
- PeopleSoft PeopleTools 8.52.10 for Oracle Exalogic
- PeopleSoft ELM 9.1 Bundle #8 with PeopleTools 8.52.06
- PeopleSoft FSCM 9.1 Feature Pack #2 (including up to Bundle #16 or Maintenance Pack 8) with PeopleTools 8.52.03
- PeopleSoft Portal Solutions 9.1 Feature Pack #1 (including up to Bundle #6) with PeopleTools 8.52.06
- PeopleSoft HCM 9.1 Feature Pack #2 (including MP5) with PeopleTools 8.52.06
- Oracle Business Intelligence Enterprise Edition 10.1.3.4
- Oracle Application Server 10g Release 3 WebCenter
- Oracle Fusion Middleware Service Oriented Architecture (SOA) 10.1.3.4

Database & Real Application Clusters (RAC)

- Oracle Database 12c & 11g Release 2 - Single Instance and Real Application Clusters (RAC)
- Oracle Real Application Clusters (RAC) 11g Release 1
- MySQL 5.5, Enterprise Edition

Operating System

- Oracle Linux 6
- Oracle Linux 5
- Oracle Linux 4
- Oracle Solaris 11
- Oracle Solaris 10

Virtual Machines

- Leave sufficient cpu resources for the bare-metal operating system or hypervisor (brand dependent) to perform I/O and manage network traffic
  - If you do not know, from your memory, how much cpu it takes to move one byte into and out of your computer's memory you shouldn't be using VMs
- Disable interrupt coalescing
- Disable chipset power management (brand dependent)
- Read the docs http://www.vmware.com/pdf/Perf_Best_Practices_vSphere5.0.pdf
- VMs on NUMA machines should be configured to recognize NUMA memory allocation
  - This example is from vSphere where 0 and 1 are the two processor sockets

```
numa.nodeAffinity=0,1
```

OVM Tuning for RAC

- Use non-sparse disks for RAC
  - Improves response times decreasing node evictions
- Minimize time drift by modifying sysctl.conf

```
xen.independent_wallclock=1
```

- Bond multiple network interfaces for production use by creating multiple xen bridges
- If the network is slow and there are heartbeat failures try increasing the misscount to 60 from default 30

```
[root@rac1 ~]# crsctl set css misscount 60
```

Operating System Layer

Operating System Topics

- Supported Operating Systems
- Kernel Parameters
- Swapiness
- DNS Redirection
- NUMA
- HugePages
- Tools

Supported Operating Systems

- *NIX
  - Linux
  - zLinux
  - Solaris
  - AIX
  - HP/UX
- Windows
  - Windows Server 2008
  - Windows Server 2012
  - Windows 7 Professional
  - Windows 8 Professional

At one time Oracle supported 92 different combinations of hardware platform and operating system

Operating System Topics

- Supported Operating Systems
- Kernel Parameters
- Swapiness
- DNS Redirection
- NUMA
- HugePages
- Tools

Kernel Parameters

- How many databases are on the server?

```
cat >> /etc/sysctl.conf <<EOF
fs.file-max = 6815744
kernel.shmall = 2097152
kernel.shmmax = 2147483648
kernel.shmmni = 4096
kernel.sem = 250 32000 100 128
net.core.rmem_default = 4194304
net.core.rmem_max = 4194304
net.core.wmem_default = 262144
net.core.wmem_max = 1048576
net.ipv4.ip_forward = 0
net.ipv4.conf.default.rp_filter = 1
tcp.ipv4.tcp_wmem = 262144 262144 262144
tcp.ipv4.tcp_rmem = 4194304 4194304 4194304
fs.aio-max-nr = 3145728
net.ipv4.ip_local_port_range = 9000 65000
EOF

-- activate changes
/sbin/sysctl -p
```

In Solaris **`rmem`** and **`wmem`** become **`rsize`** and **`wsize`**

Operating System Topics

- Supported Operating Systems
- Kernel Parameters
- Swapiness
- DNS Redirection
- NUMA
- HugePages
- Tools

Swapiness != Happiness

- Swapping (aka Paging)
- In a sense the operating system's version of the database's Temp tablespace

Value	Strategy
vm.swapiness=0	The kernel will swap only to avoid an out of memory condition
vm.swapiness=60	The default value
vm.swapiness=100	The kernel will swap aggressively

Operating System Topics

- Supported Operating Systems
- Kernel Parameters
- Swapiness
- DNS Redirection
  - Port Exhaustion is not the result of drinking too much port
- NUMA
- HugePages
- Tools

Case Study 7
Port Exhaustion

How it began

- Customer Reports stuck in the queue

Hi Ops

Report Jobs are getting stuck in Waiting in Queue. Also, having performance issues with Admin side

Thanks,
J

Step to Recreate

1. Log into Website
2. Navigate to Reports
3. Search for Account Data
4. Run the report for morgand
5. Notice that the report is stuck in Waiting in Queue

How it began

- The website declined to show this webpage (HTTP403)

As a partner we got communication that the previously assigned sandboxes will be brought down.

Instead -as a partner- we got a them demo environment assigned (Tenant ID: PARTNER0042 which we have integrated with a customer database instance (p42emo ace4morgan).

Everything was working fine (including integration).
Today I tried to access the instance via the partner and via the direct url (https://partner0001.demo.xxx.com/admin/nativelogin.jsp) but in both case an error is displayed on the screen (see attachment).

We need this be fixed as soon as possible!
(major customer demo session on Friday!)

How it began

Mozilla Firefox

File Edit View History Bookmarks Tools Help

[#CO-8985] Production is down - Succes...

Google Arch Oracle SAP World Clock Database Operations -...

CO

Production is down

Edit Assign Assign To Me Comment

Share Views

Attachments

Issue Links

has architecture desig... Master-Multiple p

Activity

All Comments Work Log History Activity Tran

added a comment -

Hi ,

We hereby acknowledge your issue. Your ticket will be assigned to the right assignee and

Proxy Error

The proxy server received an invalid response from an upstream server.
The proxy server could not handle the request GET /learning/admin/login.jsp

Reason: Error reading from remote server

Find: morgan Next Previous Highlight all Match case

Start "ntp" - sear... Microsoft E... SuccessFact... [#CO-8985... millenium_r... Microsoft ... rca_compil... Skype™ [1] ... Microsoft ... ashrdp006.... 98% 1:54 PM

How it began

Production is down

- Production-

Last checked

Uptime last 7 days 99.52%

Avg. resp. time last 7 days 944 ms

Check type: HTTP

Check resolution: 1 minutes

Uptime

Response

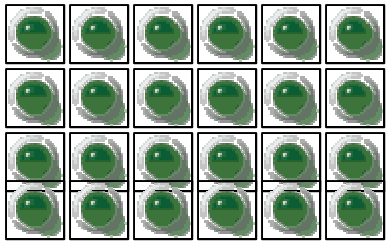
Response Time Per Country

How Does An Application Connect To RAC?

- Do you connect to the SCAN IP by name or number?
- If by name a DNS server resolves the name to an IP
- To avoid single points of failure you should have two DNS servers with a load balancer in front of them
- The SCAN IP points to a VIP
- A VIP points to a physical IP address
- Most servers cache DNS entries to improve speed

Connection Problem Triage

- Try to connect to the cluster?
  - From where?
  - With what tool?
  - To the SCAN, VIP, or physical IP?
- Ping the IP addresses
- Run Trace Route on the IP addresses
- Read the listener log
- Read the database alert log
- Call the network admins who will tell you:



everything looks good ...
the network is fine ...
the network is always fine
you DBAs are always having problems

RESOLV.CONF

NAME

resolv.conf- resolver configuration file

SYNOPSIS

`/etc/resolv.conf`

DESCRIPTION

The `resolver` is a set of routines that provide access to the Internet Domain Name System. See **resolver(3RESOLV)**. `resolv.conf` is a configuration file that contains the information that is read by the `resolver` routines the first time they are invoked by a process. The file is designed to be human readable and contains a list of keywords with values that provide various types of `resolver` information.

The `resolv.conf` file contains the following configuration directives:

`nameserver`

Specifies the Internet address in dot-notation format of a name server that the resolver is to query. Up to *MAXNS* name servers may be listed, one per keyword. See `<resolv.h>` . If there are multiple servers, the resolver library queries them in the order listed. If no name server entries are present, the resolver library queries the name server on the local machine. The resolver library follows the algorithm to try a name server until the query times out. It then tries the the name servers that follow, until each query times out. It repeats all the name servers until a maximum number of retries are made.

`domain`

Specifies the local domain name. Most queries for names within this domain can use short names relative to the local domain. If no domain entry is present, the domain is determined from **sysinfo(2)** or from gethostname(3C). (Everything after the first `.' is presumed to be the domain name.) If the host name does not contain a domain part, the root domain is assumed. You can use the `LOCALDOMAIN` environment variable to override the domain name.

RESOLV.CONF

`search`

The search list for host name lookup. The search list is normally determined from the local domain name. By default, it contains only the local domain name. You can change the default behavior by listing the desired domain search path following the search keyword, with spaces or tabs separating the names. Most `resolver` queries will be attempted using each component of the search path in turn until a match is found. This process may be slow and will generate a lot of network traffic if the servers for the listed domains are not local. Queries will time out if no server is available for one of the domains.

The search list is currently limited to six domains and a total of 256 characters.

`sortlist` ***addresslist***

Allows addresses returned by the libresolv-internal `gethostbyname()` to be sorted. A `sortlist` is specified by IP address netmask pairs. The netmask is optional and defaults to the natural netmask of the net. The IP address and optional network pairs are separated by slashes. Up to 10 pairs may be specified. For example:

```
sortlist 130.155.160.0/255.255.240.0 130.155.0.0
```

RESOLV.CONF

options

Allows certain internal resolver variables to be modified. The syntax is

```
options option ...
```

where option is one of the following:

debug

Sets RES_DEBUG in the _res.options field.

ndots: *n*

Sets a threshold floor for the number of dots which must appear in a name given to res_query() before an initial absolute (as-is) query is performed. See **resolver(3RESOLV)**. The default value for *n* is 1, which means that if there are any dots in a name, the name is tried first as an absolute name before any search list elements are appended to it.

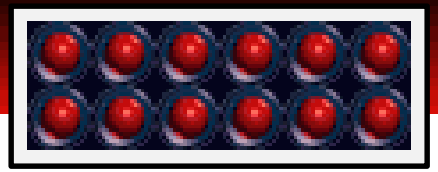
timeout: *n*
retrans: *n*

Sets the amount of time the resolver will wait for a response from a remote name server before retrying the query by means of a different name server. Measured in seconds, the default is RES_TIMEOUT . See <resolv.h> . The timeout and retrans values are the starting point for an exponential back off procedure where the timeout is doubled for every retransmit attempt.

attempts: *n*
retry: *n*

Sets the number of times the resolver will send a query to its name servers before giving up and returning an error to the calling application. The default is RES_DFLRETRY . See <resolv.h> .

Resolution: The DNS Admin



On August 7th, we experienced a 2 hour outage that impacted more than 150 customers. In researching this outage it was noticed that DNS caching had been disabled on the Oracle Database Servers. Also, in going through the logs on the F5 Local Traffic Manager (LTM), is was noticed that there were 39,696 port exhaustion errors on port 53 going to the three DNS servers, starting at approximately 4am and ending slightly after 3pm. There were also an additional 625,665 port exhaustion error messages that were dropped in the logs, bringing the total to 665,361 port exhaustion error messages.

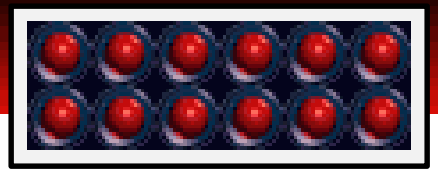
Further research discovered that there was a misconfiguration in the resolv.conf file on the servers in the data center. The resolv.conf file on these servers looked like this:

```
search morgan.priv
nameserver 10.24.244.200 (VIP pointing to servers listed below)
nameserver 10.24.244.21 (Bind server 01)
nameserver 10.24.244.25 (Bind server 02)
nameserver 10.24.244.29 (Bind server 03)
```

This results in the first DNS query going to the VIP for hostname and reverse IP resolution, and then to the three DNS servers. However, the 3 DNS servers which were supposed to be the alternative option to the VIP are also pointing to the same VIP. This basically **sets up an infinite loop** until the DNS queries time out.

The recommended resolution was to remove the VIP and have the servers query the DNS servers directly.

Resolution: The DNS Admin



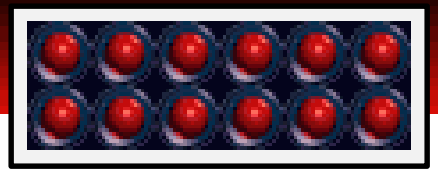
These graphs give an overview of what was happening throughout August 7th on the servers. I noticed that there is a sudden drop in connections right around 10:40am; and returning at around 10:45 am.

If you look at the files I've sent out previously, there is actually less evidence of port exhaustion between 10:22 and 10:42; with increasing levels of port exhaustion as connections and activity increases after about 12:07pm.

Additionally, I went back over the last few days and looked for port exhaustion for the DNS servers on port 53 and found the following:

```
Jul 29 -     16 port exhaustion errors
Jul 30 -      7 port exhaustion errors
Jul 31 -      8 port exhaustion errors
Aug 1  -      6 port exhaustion errors
Aug 2  - 38,711 port exhaustion errors
Aug 3  - 26,023 port exhaustion errors
Aug 4  - 22,614 port exhaustion errors
Aug 5  -     20 port exhaustion errors
Aug 6  - 11,282 port exhaustion errors
```

Resolution: The DNS Admin



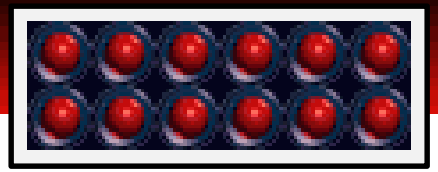
Additionally, I did some calculations on the additional port exhaustion log messages that were dropped – these were the throttling error that I mentioned previously.

On the 7th of August there were an additional 625,665 port exhaustion error messages that were dropped. On August 3rd, there were an additional 99,199 port exhaustion error messages that were dropped.

And on August 2nd, there were an additional 204,315 port exhaustion error messages that were dropped .

These numbers are in addition to the numbers of port exhaustion errors previously reported.

Resolution: The System Admin



Every unix box at the LAX data center has this resolv.conf file:

```
search morgan.priv
nameserver 10.24.244.200 (VIP pointing to both AD01 and AD02 windows servers)
nameserver 10.24.244.21  (Bind server 01)
nameserver 10.24.244.25  (Bind server 02)
nameserver 10.24.244.29  (Bind server 03)
```

The idea behind this design is to firstly query the VIP (for hostname resolution) and then, the 3 bind servers which are slave DNS servers of the AD DNS servers described above.

Now, I've found that the BIND servers (unix) which are supposed to be the alternative option to the VIP, have the same /etc/resolv.conf file and therefore are also pointing to the VIP on the first place. As you can imagine this basically ends up in an infinite loop until the load balancer get finally some relief or the DNS queries timeout.

Refer to the attachment "Morgan current arch" to see the workflow.

The fix should be easy and basically would consist of removing the VIP from the /etc/resolv.conf from the Bind servers and have them pointing to each AD server (bind01 -> AD01, bind02 -> AD02, etc).

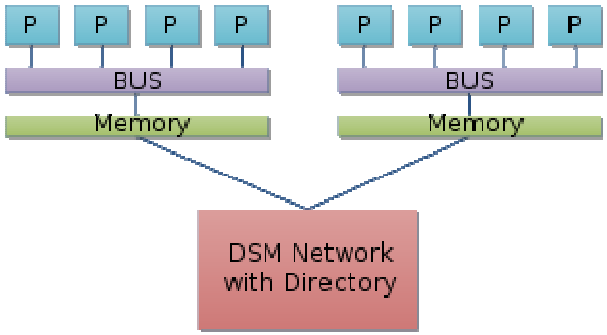
The ultimate solution would be to remove the VIP from all the /etc/resolv.conf files and query the BIND servers (Helen has been asking for this for months) and although we have done that in the DEN environment, apparently that hasn't been done on the LAX side yet.

Operating System Topics

- Supported Operating Systems
- Kernel Parameters
- Swapiness
- DNS Redirection
- NUMA
- HugePages
- Tools

NUMA (1:4)

- Non-Uniform Memory Access
  - A memory design used in multiprocessing, where the memory access time depends on the memory location relative to the processor
  - A processor can access its own local memory faster than non-local memory
  - The benefits of NUMA are limited to particular workloads, notably on servers where the data are often associated strongly with certain tasks or users



Graphic Source: Wikipedia

NUMA (2:4)

Not Configured

```
[oracle@alpha1 ~]$ numactl --hardware
available: 1 nodes (0)
node 0 size: 4095 MB
node 9 free: 648MB
node distances:
node   0
  0:  10

[oracle@alpha1 ~]$ numactl --show
policy: default
preferred node: current
physcpubind: 0
nodebind: 0
membind: 0
```

Configured

```
[dmorgan@lxorap1n5 ~]$ numactl --hardware
available: 2 nodes (0-1)
node 0 size: 48457 MB
node 0 free: 269 MB
node 1 size: 48480 MB
node 1 free: 47 MB
node distances:
node   0   1
  0:  10  20
  1:  20  10

[dmorgan@lxorap1n5 ~]$ numactl --show
policy: default
preferred node: current
physcpubind: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
cpubind: 0 1
nodebind: 0 1
membind: 0 1
```

NUMA (3:4)

- The Oracle Database does not, by default, understand NUMA
- Most System Admins don't either

```
SQL> SELECT a.ksppinm PNAME, c.ksppstvl PVAL, a.ksppdesc PDESC
  2  FROM x$ksppi a, x$ksppcv b, x$ksppsv c
  3  WHERE a.indx = b.indx
  4  AND a.indx = c.indx
  5  AND LOWER(a.ksppinm) LIKE '%numa%'
  6* ORDER BY 1;

PNAME                       PVAL           PDESC
--------------------------- -------------- ---------------------------------------------------------------------
_NUMA_instance_mapping      Not specified  Set of nodes that this instance should run on
_NUMA_pool_size             Not specified  aggregate size in bytes of NUMA pool
_db_block_numa              1              Number of NUMA nodes
_enable_NUMA_interleave     TRUE           Enable NUMA interleave mode
_enable_NUMA_optimization   FALSE          Enable NUMA specific optimizations
_enable_NUMA_support        FALSE          Enable NUMA support and optimizations
_numa_buffer_cache_stats    0              Configure NUMA buffer cache stats
_numa_shift_enabled         TRUE           Enable NUMA shift
_numa_shift_value           0              user defined value for numa nodes shift
_numa_trace_level           0              numa trace event
_px_numa_stealing_enabled   TRUE           enable/disable PQ granule stealing across NUMA nodes
_px_numa_support_enabled    FALSE          enable/disable PQ NUMA support
_rm_numa_sched_enable       FALSE          Is Resource Manager (RM) related NUMA scheduled policy enabled
_rm_numa_simulation_cpus    0              number of cpus for each pg for numa simulation in resource manager
_rm_numa_simulation_pgs     0              number of PGs for numa simulation in resource manager
```

NUMA (4:4)

- Either the S/A disables NUMA or you must enable database support

```
conn / as sysdba

ALTER SYSTEM SET "_enable_NUMA_support" = TRUE
COMMENT= 'NUMA Support Enabled 30-Jul-2014'
CONTAINER=ALL
SCOPE=SPFILE
SID='*';
```

Operating System Topics

- Supported Operating Systems
- Kernel Parameters
- Swapiness
- DNS Redirection
- NUMA
- HugePages
- Tools

HugePages (1:5)

- aka "Large Memory Pages" or "Large Pages"
- Each page table entry represents a "virtual to physical" translation of a process's memory
- PageTable overhead can be very high as standard O/S pages are 4-8K often resulting in thrashing
- Can be 2M or larger (up to 1GB when cpu has PDPE1GB flag)
- Can be used by the Oracle Database SGA
- One warning:
  The entire SGA <u>must</u> fit inside the HugePages. If it does not fit none of it will use the HugePage memory and you will essentially have walled your database off from using a large portion of the server's memory

HugePages (2:5)

- Set parameter: `/etc/sysctl.conf`

```
vm.nr_hugepages = <# of hugepages>
```

- Set memlock: `/etc/security/limits.conf`

```
* hard  memlock   5243000
* soft  memlock   5243000
```

- Restart Oracle
  - System reboot is recommended
  - Allows contiguous memory to be allocated for HugePages at system startup

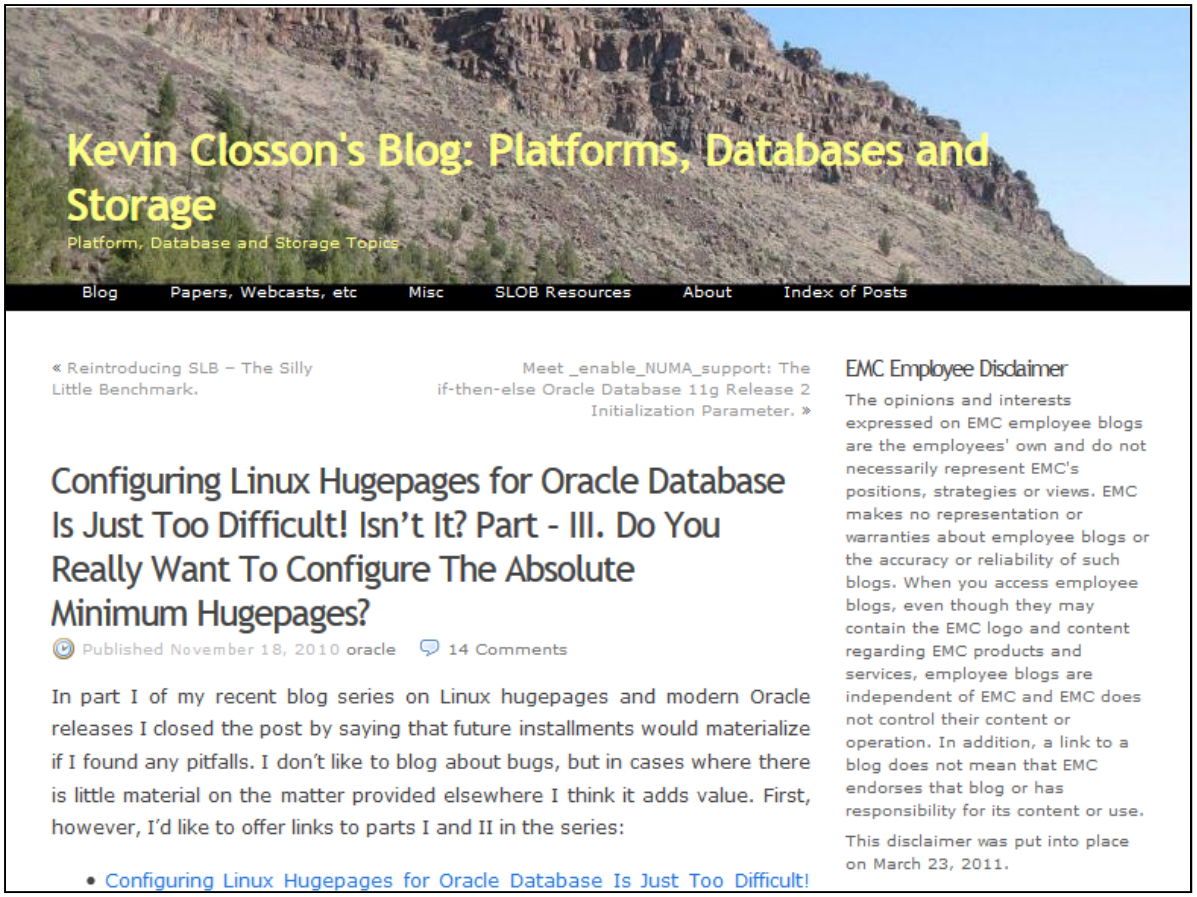
HugePages (3:5)

- Check `/proc/meminfo`
- Look for HugePages lines near the bottom of the listing

```
HugePages_Total:  6491
HugePages_Free:     10
Hugepagesize:     2048 kB
```

HugePages (4:5)

- The best technical reference on HugePages on the web is blogs by Kevin Closson



Kevin Closson's Blog: Platforms, Databases and Storage

Platform, Database and Storage Topics

Blog | Papers, Webcasts, etc | Misc | SLOB Resources | About | Index of Posts

« Reintroducing SLB – The Silly Little Benchmark.

Meet _enable_NUMA_support: The if-then-else Oracle Database 11g Release 2 Initialization Parameter. »

Configuring Linux Hugepages for Oracle Database Is Just Too Difficult! Isn't It? Part - III. Do You Really Want To Configure The Absolute Minimum Hugepages?

Published November 18, 2010 oracle 14 Comments

In part I of my recent blog series on Linux hugepages and modern Oracle releases I closed the post by saying that future installments would materialize if I found any pitfalls. I don't like to blog about bugs, but in cases where there is little material on the matter provided elsewhere I think it adds value. First, however, I'd like to offer links to parts I and II in the series:

- Configuring Linux Hugepages for Oracle Database Is Just Too Difficult!

EMC Employee Disclaimer

The opinions and interests expressed on EMC employee blogs are the employees' own and do not necessarily represent EMC's positions, strategies or views. EMC makes no representation or warranties about employee blogs or the accuracy or reliability of such blogs. When you access employee blogs, even though they may contain the EMC logo and content regarding EMC products and services, employee blogs are independent of EMC and EMC does not control their content or operation. In addition, a link to a blog does not mean that EMC endorses that blog or has responsibility for its content or use.

This disclaimer was put into place on March 23, 2011.

HugePages (5:5)

- For medium to large databases you should configure HugePages as it makes memory access more efficient
- For NUMA machines think NUMA and HugePages ... it is not a question of one or the other
- Configuration must be calculated as too few HugePages or too large an SGA will essentially fence your database off from much of the server's memory

Operating System Topics

- Supported Operating Systems
- Kernel Parameters
- Swapiness
- DNS Redirection
- NUMA
- HugePages
- Tools

Operating System Topics

- top
  - cpu utilization
- iostat
  - disk I/O
- memstat
  - memory usage
- sar
  - system activity reporter
  - **`sar -B`** (page swapping)
- vmstat
  - virtual memory
- O/S Logs
- OSWatcher

Operating System Logs

- Linux look at /var/log/messages
- Other operating systems may use other directories and file names

```
Aug  4 04:09:16 orap1n1 Updating DNS configuration for: orap1n1.lux20.morgan.priv
Aug  4 04:09:16 orap1n1 Initial DNS Server: 10.2.198.34
Aug  4 04:09:16 orap1n1 Connecting to DNS server 10.2.198.34Aug  4 04:09:16 orap1n1 Connected to DNS server 10.2.198.34
Aug  4 04:09:16 orap1n1 Updating both HOST and PTR record  for: orap1n1.lux2.morgan.priv
Aug  4 04:09:16 orap1n1 Deleting old reverse lookup records for orap1n1.lux2.morgan.priv on 10.2.198.34.
Aug  4 04:09:17 orap1n1 Adding GSS support to DNS server 10.2.198.34
Aug  4 04:09:17 orap1n1 Added GSS support to DNS server 10.2.198.34
Aug  4 04:09:17 orap1n1 Failed to delete reverse lookup record 11.78.2.10.in-addr.arpa. Reason Refused (5).
Aug  4 04:09:17 orap1n1 Deleting reverse lookup records for our current new IP Address(s) on  ad010.lux20.morgan.priv.

Aug  8 12:56:04 orap1n1 ntpd[1339]: ntpd exiting on signal 15
Aug  8 12:57:27 orap1n1 ntpdate[12406]: step time server 10.2.255.254 offset 82.262906 sec
Aug  8 12:57:27 orap1n1 ntpd[12408]: ntpd 4.2.2p1@1.1570-o Fri Jul 22 18:07:53 UTC 2011 (1)
Aug  8 12:57:27 orap1n1 ntpd[12409]: precision = 1.000 usec
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface wildcard, 0.0.0.0#123 Disabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface wildcard, ::#123 Disabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface bond2, fe80::217:a4ff:fe77:fc18#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface lo, ::1#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface bond0, fe80::217:a4ff:fe77:fc10#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface eth2, fe80::217:a4ff:fe77:fc14#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface eth3, fe80::217:a4ff:fe77:fc16#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: Listening on interface lo, 127.0.0.1#123 Enabled
Aug  8 12:57:27 orap1n1 ntpd[12409]: kernel time sync status 0040
Aug  8 12:57:27 orap1n1 ntpd[12409]: frequency initialized 0.000 PPM from /var/lib/ntp/drift
```

OSWatcher

- A tool for the analysis of the data produced such linux utilities as top, iostat, vmstat, and netstat
- Provides the analysed data as a graphic with OSWg tool
- Provides persistent storage for later analysis of operating system metrics
- Distributed, free, by Oracle Support

Helmut's Oracle Blog

HOME **RAC** RDBMS LINUX VIRTUALBOX JAVA OCP

OSWatcher – Installation and Usage

Contents

LVM Layer Configuration

ASM

- Intelligent Data Placement
  - When using spinning disk placing data on the outside of the disk yields faster access as the head can read more track area per unit of time
    - Works best for databases where different files are accessed at different rates

```
ALTER DISKGROUP data
ADD TEMPLATE datafile_hot ATTRIBUTE (HOT MIRRORHOT);

ALTER DISKGROUP archives
ADD TEMPLATE datafile_cold ATTRIBUTE (COLD MIRRORCOLD);

 ALTER DISKGROUP data
MODIFY FILE '+data/orcl/datafile/users.259.679156903'
ATTRIBUTE (HOT MIRRORHOT);
```

Database Layer Configuration

Database Configuration Topics

- Statistics and Stats Collection
- Memory Management
- Optimizer Parameters
- RAC
- Segments
- Tablespaces

Statistics

- System Stats
- Fixed Table Stats
- Dictionary Stats
- Table and Index Stats
- Resource Manager and Calibrate

System Stats (1:3)

- How does the optimizer know the capabilities of the hardware it is running on?
- DBMS_STATS.GATHER_SYSTEM_STATS gathers information for the CBO while the system is under load
- This information is used to determine the capabilities of the server's subsystems
- System Stats should be collected once on every system under full production load and then, again, every time a subsystem is modified

System Stats (2:3)

PNames	Descriptions
CPUSPEED	Workload CPU speed in millions of cycles/second
CPUSPEEDNW	Noworkload CPU speed in millions of cycles/second
IOSEEKTIM	Seek time + latency time + operating system overhead time in milliseconds
IOTFRSPEED	Rate of a single read request in bytes/millisecond
MAXTHR	Maximum throughput that the I/O subsystem can deliver in bytes/second
MBRC	Average multiblock read count sequentially in blocks
MREADTIM	Average time for a multi-block read request in milliseconds
SLAVETHR	Average parallel slave I/O throughput in bytes/second
SREADTIM	Average time for a single-block read request in milliseconds

System Stats (3:3)

```
SQL> SELECT pname, pval1
  2  FROM aux_stats$
  3  WHERE sname = 'SYSSTATS_MAIN';

PNAME                              PVAL1
------------------------------ ----------
CPUSPEED                              784
CPUSPEEDNW                     3264.04494
IOSEEKTIM                              10
IOTFRSPEED                           4096
MAXTHR
MBRC
MREADTIM
SLAVETHR
SREADTIM                             .495

exec dbms_stats.delete_system_stats;
alter system flush shared_pool;
exec dbms_stats.gather_system_stats();
-- wait 5 minutes then do work (create, select, rollup, cube, group by, order by)
exec dbms_stats.gather_system_stats('STOP');

SQL> SELECT pname, pval1
  2  FROM aux_stats$
  3  WHERE sname = 'SYSSTATS_MAIN';

PNAME                              PVAL1
------------------------------ ----------
CPUSPEED                             2634
CPUSPEEDNW                     3264.04494
IOSEEKTIM                              10
IOTFRSPEED                           4096
MAXTHR
MBRC                                   53
MREADTIM                            2.246
SLAVETHR
SREADTIM                           11.527
```

Fixed Table Stats

- Gathers statistics for all fixed objects (dynamic performance tables)

```
SQL> select type, count(*)
  2  from v$fixed_table
  3  group by type
  4  order by 1;

TYPE    COUNT(*)
----- ----------
TABLE       1144
VIEW        1261

SQL> select name
  2  from v$fixed_table
  3  where rownum < 11;

NAME
--------------------------
X$KQFTA
X$KQFVI
X$KQFVT
X$KQFDT
X$KQFCO
X$KQFOPT
X$KYWMPCTAB
X$KYWMWRCTAB
X$KYWMCLTAB
X$KYWMNF
```

```
dbms_stats.gather_fixed_objects_stats (
stattab       IN VARCHAR2 DEFAULT NULL,
statid        IN VARCHAR2 DEFAULT NULL,
statown       IN VARCHAR2 DEFAULT NULL,
no_invalidate IN BOOLEAN  DEFAULT to_no_invalidate_type(get_param('NO_INVALIDATE')));


SQL> exec dbms_stats.gather_fixed_objects_stats;
```

Dictionary Stats

- Gathers statistics for dictionary schemas 'SYS', 'SYSTEM' and schemas of RDBMS components
  - Two overloads

```
SQL> exec dbms_stats.gather_dictionary_stats(cascade=>TRUE);
```

Table, Partition, Column, and Index Stats

```
dbms_stats.gather_table_stats(
ownname          IN VARCHAR2,
tabname          IN VARCHAR2,
partname         IN VARCHAR2 DEFAULT NULL,
estimate_percent IN NUMBER   DEFAULT DEFAULT_ESTIMATE_PERCENT,
block_sample     IN BOOLEAN  DEFAULT FALSE,
method_opt       IN VARCHAR2 DEFAULT DEFAULT_METHOD_OPT, -- or 'FOR ALL COLUMNS SIZE 1',
degree           IN NUMBER   DEFAULT to_degree_type(get_param('DEGREE')),
granularity      IN VARCHAR2 DEFAULT DEFAULT_GRANULARITY,
cascade          IN BOOLEAN  DEFAULT DEFAULT_CASCADE,
stattab          IN VARCHAR2 DEFAULT NULL,
statid           IN VARCHAR2 DEFAULT NULL,
statown          IN VARCHAR2 DEFAULT NULL,
no_invalidate    IN BOOLEAN  DEFAULT to_no_invalidate_type(get_param('NO_INVALIDATE')),
stattype         IN VARCHAR2 DEFAULT 'DATA',
force            IN BOOLEAN  DEFAULT FALSE,
context          IN dbms_stats.ccontext DEFAULT NULL,
options          IN VARCHAR2 DEFAULT DEFAULT_OPTIONS);
```

```
exec dbms_stats.gather_table_stats(USER, 'SERVERS', 'P2010103106',
estimate_percent => dbms_stats.auto_sample_size,
block_sample=>TRUE,
method_opt=>'FOR ALL INDEXED COLUMNS',
degree=>2,
cascade=>TRUE
force=>TRUE);
```

Extended Stats

- Creates a virtual column for a user specified column group or an expression in a table
- Allows for the creation of stats that relate to a data distribution across multiple columns in a single table

```
SELECT dbms_stats.create_extended_stats(USER, 'SERV_INST', '(srvr_id, si_status)')
FROM dual;
```

Set Preferences

- Set Stats Collection preferences so DBMS_STATS runs the way you want it to run
  - SET_DATABASE_PREFS
  - SET_GLOBAL_PREFS
  - SET_SCHEMA_PREFS
  - SET_TABLE_PREFS

```
exec dbms_stats.set_table_prefs(USER, 'SERVERS', 'ESTIMATE_PERCENT','90');

exec dbms_stats.set_table_prefs(USER, 'SERVERS', 'DEGREE','8');
```

Set Stats

- Set stats when you know what will be there before stats will next be gathered
  - When you add a new partition to a range partitioned table do you already know the data range it will be holding? The optimizer doesn't

```
exec dbms_stats.set_table_stats(USER, 'EMP', numrows=>1000000, numblks=>10000, avgrlen=>74);

exec dbms_stats.set_index_stats(USER, 'ix_emp_deptno', numrows=>1000000, numlblks=>1000,
     numdist=>10000, clstfct=>1);

exec dbms_stats.set_column_stats(USER, 'emp', 'deptno', distcnt=>10000);

exec dbms_stats.set_table_stats(USER, 'dept', numrows=>100, numblks=>100);
```

Fix Stats

- Fix stats when you know there are errors: for example on a newly created partition
  - SET_COLUMN_STATS
  - SET_INDEX_STATS
  - SET_SYSTEM_STATS
  - SET_TABLE_STATS

```
DECLARE
 sr_arr   dbms_stats.statrec;
 new_low  DATE := TO_DATE('01-DEC-2013');
 new_high DATE := TO_DATE('31-DEC-2013');
 vals_arr dbms_stats.datearray;
BEGIN
  sr_arr.eavs := 0;
  sr_arr.chvals := NULL;
  sr_arr.epc := 2;

  vals_arr := dbms_stats.datearray(new_low, new_high);

  sr_arr.bkvals := dbms_stats.numarray(10000,1000000);
  dbms_stats.prepare_column_values(sr_arr, vals_arr);
  dbms_stats.set_column_stats(USER, 'PREPTEST', 'LAST_DDL_TIME', NULL, srec => sr_arr);
END;
/
```

New in 12c

- DELETE_PROCESSING_RATE
- GATHER_PROCESSING_RATE
- SET_PROCESSING_RATE
  - Set the preferences so that all collection jobs do it your way

```
dbms_stats.gather_processing_rate(
 gathering_mode IN VARCHAR2 DEFAULT 'START', -- valid values 'START' and 'END'
 duration       IN NUMBER   DEFAULT NULL);   -- default is 60 minutes

exec dbms_stats.gather_processing_rate('START', 1200);
```

ALL	CPU_JOIN	IO
CPU	CPU_NL_JOIN	IO_ACCESS
CPU_ACCESS	CPU_RANDOM_ACCESS	IO_BYTES_PER_SEC
CPU_AGG	CPU_SEQUENTIAL_ACCESS	IO_RANDOM_ACCESS
CPU_BYTES_PER_SEC	CPU_SM_JOIN	IO_SEQUENTIAL_ACCESS
CPU_FILTER	CPU_SORT	MEMCM
CPU_GBY	HASH, AGGR	MEMCPY
CPU_HASH_JOIN		

DBMS_RESOURCE_MANAGER

- Maintains plans, consumer groups, and plan directives that control and limit CPU and I/O resource allocation between PDBs and User Consumer Groups
- Provides semantics so that can group together changes to the plan schema
- Note in the Oracle docs the phrase "It can only be called when consolidation is enabled" which is jargon that when translated into English means "you have a CDB" and multitenant is enabled
- Introduced at version10gR1

Resource Manager Calibrate I/O

- If you do not run CALIBRATE_IO resource manager can not manage I/O resources

```
set serveroutput on

 DECLARE
  iops PLS_INTEGER;
  mbps PLS_INTEGER;
  alat PLS_INTEGER;
 BEGIN
   dbms_resource_manager.calibrate_io(1, 100, iops, mbps, alat);
   dbms_output.put_line('Maximum IOPS: ' || TO_CHAR(iops));
   dbms_output.put_line('Maximum MBPS: ' || TO_CHAR(mbps));
   dbms_output.put_line('Actual Latency: ' || TO_CHAR(alat));
 END;
 /

 SELECT * FROM gv$io_calibration_status;

 SELECT * FROM dba_rsrc_io_calibrate;
```

Memory Management

- Oracle's memory management architecture in 10g was ASMM (Automatic SGA Memory Management)
- In 11g Oracle tried AMM (Automatic Memory Management) assuming it would be an improvement: Oracle was really really wrong! Resize ops manufactures waits

```
SELECT trunc(start_time), status, oper_type, oper_mode, parameter
FROM v$sga_resize_ops
WHERE initial_size <> final_size;
```

- If AMM is deployed go back to ASMM to stop thrashing

```
SQL> ALTER SYSTEM SET memory_max_target=0 SCOPE=spfile;
SQL> ALTER SYSTEM SET memory_target=0 SCOPE=both;
SQL> ALTER SYSTEM SET sga_target=###G SCOPE=both;
```

Optimizer Parameters

- COMPATIBLE

```
SQL> show parameter compatible

NAME                                 TYPE        VALUE
------------------------------------ ----------- ----------------
compatible                           string      12.1.0.0.0
SQL> show parameter optimizer

NAME                                 TYPE
------------------------------------ ----------- ----------------
optimizer_adaptive_features          boolean     TRUE
optimizer_adaptive_reporting_only    boolean     FALSE
optimizer_capture_sql_plan_baselines boolean     FALSE
optimizer_dynamic_sampling           integer     2
optimizer_features_enable            string      12.1.0.1
optimizer_index_caching              integer     0
optimizer_index_cost_adj             integer     100
optimizer_mode                       string      FIRST_ROWS_10
optimizer_secure_view_merging        boolean     TRUE
optimizer_use_invisible_indexes      boolean     FALSE
optimizer_use_pending_statistics     boolean     FALSE
optimizer_use_sql_plan_baselines     boolean     TRUE
```

Did you upgrade the database but no one told the spfile?

Optimizer Parameters

- OPTIMIZER_MODE
  - The default of ALL_ROWS optimizes for maximum throughput for DSS and DW usage
  - FIRST_ROWS optimizes for best response time for OLTP

```
SQL> show parameter optimizer

NAME                                 TYPE        VALUE
------------------------------------ ----------- ----------------
optimizer_adaptive_features          boolean     TRUE
optimizer_adaptive_reporting_only    boolean     FALSE
optimizer_capture_sql_plan_baselines boolean     FALSE
optimizer_dynamic_sampling           integer     2
optimizer_features_enable            string      12.1.0.1
optimizer_index_caching              integer     0
optimizer_index_cost_adj             integer     100
optimizer_mode                       string      FIRST_ROWS_10
optimizer_secure_view_merging        boolean     TRUE
optimizer_use_invisible_indexes      boolean     FALSE
optimizer_use_pending_statistics     boolean     FALSE
optimizer_use_sql_plan_baselines     boolean     TRUE
```

Optimizer Parameters

- PARALLEL_ADAPTIVE_MULTI_USER
  - When set to true, enables an adaptive algorithm designed to improve performance in multiuser environments that use parallel execution. The algorithm automatically reduces the requested degree of parallelism based on the system load at query startup time. The effective degree of parallelism is based on the default degree of parallelism, or the degree from the table or hints, divided by a reduction factor
  - The algorithm assumes that the system has been tuned for optimal performance in a single-user environment
  - Tables and hints use the default degree of parallelism
  - Default value is TRUE

```
ALTER SYSTEM SET parallel_adaptive_multi_user=FALSE;
```

Optimizer Parameters

- PARALLEL_AUTOMATIC_TUNING
  - When set to true, Oracle determines the default values for parameters that control parallel execution. In addition to setting this parameter, you must specify the PARALLEL clause for the target tables in the system. Oracle then tunes all subsequent parallel operations automatically.
  - If you used parallel execution in a previous release and are now enabling PARALLEL_AUTOMATIC_TUNING, then you should reduce the amount of memory allocated from the shared pool to account for the decreased demand on that pool. This memory will now be allocated from the large pool, and will be computed automatically if LARGE_POOL_SIZE is not specified
  - Default: `FALSE`
  - Deprecated in 12c

```
ALTER SYSTEM SET parallel_automatic_tuning=FALSE;
```

Optimizer Parameters

- PARALLEL_DEGREE_LIMIT
  - Automatically decides whether a statement should execute in parallel and what degree of parallelism the statement should use
  - The optimizer automatically determines the degree of parallelism for a statement based on the resource requirements of the statement
  - The optimizer will limit the degree of parallelism used to ensure parallel server processes do not flood the system
  - Default value is CPU

```
ALTER SYSTEM SET parallel_degree_limit=CPU;
```

Optimizer Parameters

- PARALLEL_DEGREE_POLICY
  - Specifies whether automatic degree of parallelism, statement queuing, and in-memory parallel execution will be enabled
  - Default value is CPU which disables automatic degree of parallelism, statement queuing, and in-memory parallel execution and reverts the behavior of parallel execution to what it was prior to Oracle Database 11g Release 2 (11.2) which is the default
  - Note the above: In 12c Oracle has decided the 11gR2 algorithm was flawed

```
ALTER SYSTEM SET parallel_degree_policy=MANUAL;
```

Optimizer Parameters

- PARALLEL_EXECUTION_MESSAGE_SIZE
  - Specifies the size of messages used for parallel execution (formerly referred to as parallel query, PDML, Parallel Recovery, replication)
  - Default value is operating system dependent. On most platforms, the default value is as follows:
    - 16384 bytes if COMPATIBLE is set to 11.2.0 or higher
    - 4096 bytes if COMPATIBLE is less than 11.2.0 and PARALLEL_AUTOMATIC_TUNING is set to true
    - 2148 bytes if COMPATIBLE is less than 11.2.0 and PARALLEL_AUTOMATIC_TUNING is set to false
  - The default value is adequate for most applications. Larger values require a larger shared pool. Larger values result in better performance at the cost of higher memory use. For this reason, replication gets no benefit from increasing the size.

Optimizer Parameters

- PARALLEL_FORCE_LOCAL
  - If a query is parsed on node 1 of a RAC cluster and parallelized to node 2 ... how do the rows generated on node 2 get back to node 1 and to the client?
  - The cache fusion interconnect where the critical design decision is low latency not high bandwidth
  - Default value is FALSE meaning queries can be parallelized across multiple nodes
  - In my opinion this should be set to TRUE unless someone can prove, through testing, there are no negative consequences that result from Oracle's default

```
ALTER SYSTEM SET parallel_force_local=TRUE SCOPE=BOTH container=ALL;
```

Optimizer Parameters

- PARALLEL_INSTANCE_GROUP
  - A RAC parameter that can be specified in parallel mode only
  - Used in conjunction with services
  - Allows restriction of parallel query operations to a limited number of instances
  - By default parallel execution is enabled across all currently active instances
  - In my opinion this should be set to a non-existent group to disable this "feature" unless thorough testing demonstrates that it is safe

```
ALTER SYSTEM SET parallel_instance_group='ZZYZX';
```

Optimizer Parameters (new 12c)

- PARALLEL_IO_CAP_ENABLED
  - Specifies whether or not Oracle caps the default degree of parallelism to no greater than what the I/O system can support
  - Calculated based on the results of the resource manager's I/O calibration package
  - Database I/O must be calibrated for this to work

Optimizer Parameters

- PARALLEL_MAX_SERVERS
  - Specifies the maximum number of parallel execution processes and parallel recovery processes for an instance
  - Default: `0 to 3600`

Optimizer Parameters

- PARALLEL_MIN_PERCENT
  - The minimum percentage of the requested number of parallel execution processes required for parallel execution
  - Setting this parameter ensures that parallel operations will not execute unless adequate resources are available
  - The default value of 0 means that no minimum percentage of processes has been set
  - Default: 0

Optimizer Parameters

- PARALLEL_MIN_SERVERS
  - The number of parallel execution processes Oracle creates when the instance is started
  - These processes will be kept alive to service parallel statements
  - Default: `CPU_COUNT * PARALLEL_THREADS_PER_CPU * 2`

Optimizer Parameters

- PARALLEL_MIN_TIME_THRESHOLD
  - The minimum execution time a statement should have before the statement is considered for automatic degree of parallelism
  - By default, this is set to 10 seconds
  - Automatic degree of parallelism is only enabled if PARALLEL_DEGREE_POLICY is set to ADAPTIVE, AUTO, or LIMITED

Optimizer Parameters

- PARALLEL_SERVERS_TARGET
  - The number of parallel server processes allowed to run parallel statements before statement queuing will be used
  - When the parameter PARALLEL_DEGREE_POLICY is set to ADAPTIVE or AUTO, Oracle will queue SQL statements that require parallel execution, if the necessary parallel server processes are not available
  - Statement queuing will begin once the number of parallel server processes active on the system is equal to or greater than PARALLEL_SERVERS_TARGET
  - Range of values: 0 to PARALLEL_MAX_SERVERS

Optimizer Parameters

- PARALLEL_THREADS_PER_CPU
  - The default degree of parallelism for the instance and determines the parallel adaptive and load balancing algorithms
  - Describes the number of parallel execution processes or threads that a CPU can handle during parallel execution
  - The default is platform-dependent and is usually adequate
  - Decrease the value of this parameter if the system appears to be overloaded when a representative parallel query is executed
  - Increase the value if the system is I/O bound

Optimizer Parameters for Deprecated Parameters

- Check the ALERT LOG
  - When was the last time someone looked to see if the tuning parameters set are now obsolete following an upgrade

```
Sat May 13 09:17:34 2011
Starting up:
Oracle Database 11g Enterprise Edition Release 11.2.0.1.0 - Production
With the Partitioning, OLAP, Data Mining and Real Application Testing options.
Using parameter settings in server-side spfile C:\ORACLE\PRODUCT\11.2.0\DB_1\DATABASE\SPFILEORABASE.ORA
System parameters with non-default values:
  processes                = 150
  memory_target            = 1232M
  control_files            = "C:\ORACLE\ORADATA\ORABASE\CONTROL01.CTL"
  control_files            = "C:\ORACLE\FLASH_RECOVERY_AREA\ORABASE\CONTROL02.CTL"
  db_block_size            = 8192
  compatible               = "11.2.0.0.0"
  db_recovery_file_dest    = "C:\oracle\flash_recovery_area"
  db_recovery_file_dest_size= 3852M
  undo_tablespace          = "UNDOTBS1"
  remote_login_passwordfile= "EXCLUSIVE"
  db_domain                = ""
  dispatchers              = "(PROTOCOL=TCP) (SERVICE=orabaseXDB)"
  audit_file_dest          = "C:\ORACLE\ADMIN\ORABASE\ADUMP"
  audit_trail              = "DB"
  db_name                  = "orabase"
  open_cursors             = 300
  parallel_automatic_tuning= TRUE
  diagnostic_dest          = "C:\ORACLE"
Deprecated system parameters with specified values:
  parallel_automatic_tuning
End of deprecated system parameter listing
```

RAC: Dynamic Remastering (1:2)

- Blocks, by range, are mastered (tracked) by a single instance with a hash algorithm used to spread mastering for a segment among the instances
- Remastering happens when one instance is reading a segment aggressively and other instances are not reading that object at all?
- Remastering is expensive
- During remastering instance GRD is frozen during reconfiguration, and in a very busy instances, this can take many seconds leading to instance freeze

RAC: Dynamic Remastering (2:2)

- An increase in remastering indicates poor node affinity in the application and may be responsible for substantial performance issues

```
SQL> SELECT current_master, previous_master, COUNT(*)
   2  FROM v$gcspfmaster_info
   3  WHERE current_master <> previous_master
   4  GROUP BY current_master, previous_master
   5  ORDER BY 1,2;

CURRENT_MASTER PREVIOUS_MASTER   COUNT(*)
-------------- --------------- ----------
             4               5          7
             4               6         60
             4           32767       2355
             5               4         46
             5               6         24
             5           32767        171
             6               4        196
             6               5         15
             6           32767       1430
```

Segments (1:2)

- Block Size considerations
  - Ignore the nonsense on the internet ... Beta testers test 8K blocks ... accept the risk if you use a different size
- Extent Size considerations
  - When are new extents being added? There is no excuse for adding them during the business day
- INITRANS considerations
  - Initial number of transaction entries allocated within each block
  - Can be tuned to accommodate tables that experience contention due to high transaction volume
- PCTFREE considerations
  - How much unusable empty space is this creating?

Segments (2:2)

- Row chaining considerations
  - When is the last time you ran ANALYZE TABLE to detect chained rows? Chained rows affect index reads and full table scans
- Row migration considerations
  - Migrated rows affect indexed reads adding an extra I/O
- High water mark considerations
  - How much of the table storage is empty? A Full Table Scan will read it all
  - To determine whether space is being used inefficiently
    - DBMS_SPACE.SPACE_USAGE
    - DBMS_SPACE.UNUSED_SPACE
    - DBMS_SPACE.VERIFY_SHRINK_CANDIDATE

Tablespaces

- Undo Tablespace
  - Are you using DBMS_UNDO_ADV?
- Temp Tablespaces
  - What is the value in having only a single temp tablespace?
  - If you can answer that question then 1 may be the right answer for your database
  - If you can't answer the question ... reconsider your design
  - By definition, with Database 12c, you will not have only one temp tablespace
- Bigfile Tablespaces
  - The overhead created by managing numerous datafiles can be a serious impediment to high performance
  - No amount of SQL tuning will put all of the extents of a table into a single physical file or on a single logical stripe

Case Study 8
It's RAC

Ticket 108 (1 of 2)

- RCA Request for DC20 | Database | All databases are down due to memory issue and its 100% full. Here's the alert log.

```
system name: Linux
Node name: oras1n1.lux20.morgan.priv
Release: 2.6.18-274.el5
Version: #1 SMP Mon Jul 25 13:17:49 EDT 2011
Machine: x86_64
Redo thread mounted by this instance: 1
Oracle process number: 0
Unix process pid: 32402, image: oracle@oras1n1.lux20.morgan.priv (J000)


*** 2013-07-04 03:51:11.919
Unexpected error 27140 in job slave process
ORA-27140: attach to post/wait facility failed
ORA-27300: OS system dependent operation:invalid_egid failed with status: 1
ORA-27301: OS failure message: Operation not permitted
ORA-27302: failure occurred at: skgpwinit6
ORA-27303: additional information: startup egid = 1001 (oinstall), current egid = 1003 (asmadmin)

Errors in file /app/oracle/base/diag/rdbms/dc20sce11/DC20SCE11/trace/DC20SCE11_j000_32402.trc:
ORA-27140: attach to post/wait facility failed
ORA-27300: OS system dependent operation:invalid_egid failed with status: 1
ORA-27301: OS failure message: Operation not permitted
ORA-27302: failure occurred at: skgpwinit6
ORA-27303: additional information: startup egid = 1001 (oinstall), current egid = 1003 (asmadmin)

And current status of memory usage:
 oracle@oras1n1.lux20.morgan.priv[DC20SCE11]$ free -g
             total used free shared buffers cached
Mem: 141 140 1 0 0 66
-/+ buffers/cache: 73 67
Swap: 31 0 31
```

Ticket 108 (2 of 2)

Root Cause:
A review of Oracle Binary in oras1n1 revealed that Oracle Databases were started by user “oracle” and at that point in time the ORACLE_HOME/bin/oracle executable group was "oinstall". The ORACLE_HOME/bin/oracle executable group was accidentally changed to "asmadmin", due to a known Oracle bug. Due to this bug, cluster nodes originally started with Server Control must always be started with Server Control and, if started with SQL*Plus, can produce the result observed.

Corrective Action:
Need to change the group of executable "oracle" to "oinstall" for all the database homes, if they have been modified. The bug hit has been acknowledged by Oracle and, at least in theory, should be fixed in version 12cR1 and above. Further improvements will be tracked by CSI via a Corrective Measure (CM).

Case Study 9
Human Nature

Repeating Issue: User Configured Loads

RAC Server Node 1

MMDD	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
0804	0	0	0	0	0	0	0	0	0	0	0	0	5	32	18	65	91	13	12	20	84	9	14	9
0805	137	112	26	27	141	17	21	9	85	13	21	17	96	23	23	24	91	13	11	21	86	11	14	9
0806	151	111	21	24	96	41	50	14	84	22	20	22	91	18	17	18	92	24	10	11	83	9	14	20
0807	139	100	32	30	99	43	49	19	105	17	31	14	76	23	27	25	111	20	15	18	86	13	13	10
0808	145	99	29	30	109	52	48	11	102	25	47	24	101	23	20	23	117	31	30	16	91	12	11	9
0809	123	83	65	37	93	17	25	10	102	23	44	25	111	37	24	29	98	19	29	16	92	16	15	9
0810	169	120	52	32	125	58	38	9	109	17	26	14	104	13	17	15	93	13	16	11	61	10	10	9
0811	107	82	51	34	85	17	22	10	73	10	12	11	92	32	13	69	65	11	11	10	60	9	12	9
0812	149	121	26	15	70	16	24	11	95	34	15	18	34	67	21	21	87	11	13	9	77	9	14	9
0813	115	76	55	56	27	9	9	9	11	9	9	0	0	0	0	0	0	0	0	0	0	0	0	0

60 corresponds to one change per minute ... the ideal range is 4 to 12
Addressed by resizing redo logs from 400MB to 4GB
And rescheduling many of the jobs

Case Study 10
Jobs

Unobserved Job Failure: ASHPCE1D

```
1  select owner, job_name, job_type, trunc(start_date) SDATE, trunc(next_run_date) nxtrun, failure_count
  2  from dba_scheduler_jobs
  3* where failure_count <> 0;

OWNER           JOB_NAME                       STATE      SDATE                NXTRUN               FAILURE_COUNT
--------------- ------------------------------ ---------- -------------------- -------------------- -------------
SYS             SM$CLEAN_AUTO_SPLIT_MERGE      SCHEDULED  14-MAR-2011 00:00:00 14-AUG-2013 00:00:00            17
SYS             RSE$CLEAN_RECOVERABLE_SCRIPT   SCHEDULED  14-MAR-2011 00:00:00 14-AUG-2013 00:00:00            20
SYS             DRA_REEVALUATE_OPEN_FAILURES   SCHEDULED                                                        10
ORACLE_OCM      MGMT_CONFIG_JOB                SCHEDULED                                                         4
EXFSYS          RLM$SCHDNEGACTION              SCHEDULED  13-AUG-2013 00:00:00 13-AUG-2013 00:00:00             3
EXFSYS          RLM$EVTCLEANUP                 SCHEDULED  27-APR-2011 00:00:00 13-AUG-2013 00:00:00             2
RDBA5           LONG_RUN_SESS_JOB              SCHEDULED  12-AUG-2013 00:00:00 13-AUG-2013 00:00:00             1
EISAI_PROD_TMS  POPULATE_MORGAN_CATALOG        DISABLED   01-JUN-2009 00:00:00 08-AUG-2013 00:00:00          2559
```

Unobserved Job Failure: ASHSCE541

```
 1  select owner, job_name, job_type, state, trunc(start_date) SDATE, trunc(next_run_date) NXTRUN, failure_count
 2  from dba_scheduler_jobs
 3  where failure_count > 0
 4* order by 6;

OWNER           JOB_NAME                       STATE      SDATE                NXTRUN               FAILURE_COUNT
_______________ ______________________________ __________ ____________________ ____________________ _____________
SYS             PVX_STUDENT                    SCHEDULED  29-MAR-2013          09-AUG-2013                    122


Called out in Jira CO-1234 for the following exception:


  r-sd-ds:aukoras1n4 Logscan matched patterns in /app/oracle/base/diag/rdbms/auksce54/AUKSCE541/trace/alert_AUKSCE541.log RDBA WARN + Errors in file
/app/oracle/base/diag/rdbms/auksce54/AUKSCE541/trace/AUKSCE541_j000_12172.trc: + ORA-12012: error on auto execute of job "SYS"."PVX_STUDENT_REFRESH" W ORA-06550: line 1, column
797: +  PLS-00103: Encountered the symbol "PVX_STUDENT" when expecting one of the following: + + ) , * & = - + < / > at in is mod remainder not rem => + <> or != or ~= >= <= <> and or like like2 +
like4 likec as between from using || multiset member ------------------------- alert.pl v5.3.120207 mon_hub:auktusc01 (auktusc01) run_time:2013-Aug-08 08:01:43 client:R-SDS server:auktusc01
entity:aukoras1n4 entity_type:OPSYS processed by ftp_mail_proc.pl on delphi at 8-Aug-2013 07:03
```

Database Layer Object Selection

Object Selection

- Clusters
- Tables and Table Design
- Constraints
- Indexes
- Sequences
- Materialized Views

Clusters (1:4)

- A cluster is a section of a normal tablespace in which the developer or DBA modifies the default storage rules so that an individual block can contain more than a single object
- Cluster Types
  - Single table hash clusters
  - Multi-table hash clusters
  - Multi-table index clusters
  - Sorted Hash Clusters

Clusters (2:4)

- Oracle uses Cluster Tables to obtain maximum access performance in the data dictionary

```
SQL> SELECT DISTINCT cluster_name, tablespace_name
  2  FROM dba_clusters
  3  ORDER BY 1;

CLUSTER_NAME                    TABLESPACE_NAME
------------------------------- -------------------
C_COBJ#                         SYSTEM
C_FILE#_BLOCK#                  SYSTEM
C_MLOG#                         SYSTEM
C_OBJ#                          SYSTEM
C_OBJ#_INTCOL#                  SYSTEM
C_RG#                           SYSTEM
C_TOID_VERSION#                 SYSTEM
C_TS#                           SYSTEM
C_USER#                         SYSTEM
SMON_SCN_TO_TIME_AUX            SYSAUX
```

Clusters (3:4)

- Having related column, index, and table information co-located means that one single block read can bring back data about all of them

```
SQL> SELECT owner, table_name
  2  FROM dba_tables
  3  WHERE cluster_name = 'C_OBJ#'
  4  ORDER BY 2;

OWNER                      TABLE_NAME
-------------------------- ------------
SYS                        ASSEMBLY$
SYS                        ATTRCOL$
SYS                        CLU$
SYS                        COL$
SYS                        COLTYPE$
SYS                        ICOL$
SYS                        ICOLDEP$
SYS                        IND$
SYS                        LIBRARY$
SYS                        LOB$
SYS                        NTAB$
SYS                        OPQTYPE$
SYS                        REFCON$
SYS                        SUBCOLTYPE$
SYS                        TAB$
SYS                        TYPE_MISC$
SYS                        VIEWTRCOL$
```

Clusters (4:4)

- Sorted Hash Clusters are another type of table in which the data is not sorted but the hashes are and can eliminate the overhead of an ORDER BY clause

```
CREATE CLUSTER sorted_hc (
program_id  NUMBER(3),
line_id     NUMBER(10) SORT,
delivery_dt DATE       SORT)
TABLESPACE uwdata
HASHKEYS 9
SIZE 750
HASH IS program_id;


CREATE TABLE shc_airplane (
program_id  NUMBER(3),
line_id     NUMBER(10) SORT,
delivery_dt DATE       SORT,
customer_id VARCHAR2(3),
order_dt    DATE)
CLUSTER sorted_hc (program_id, line_id, delivery_dt);
```

Object Selection

- Clusters
- Tables and Table Design
  - Table types and Partitioning
  - Table Attributes
  - Too many columns
  - Column ordering
- Constraints
- Indexes
- Sequences
- Materialized Views

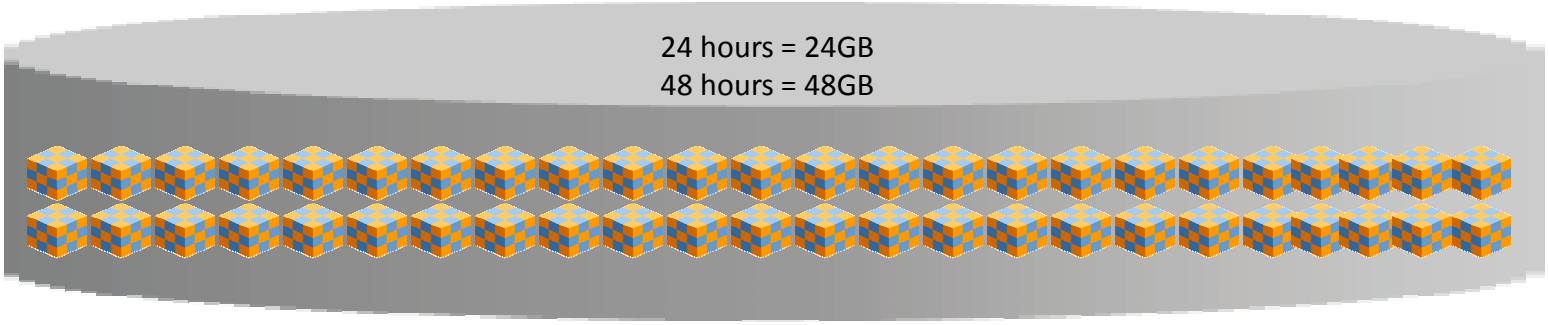
Table Types

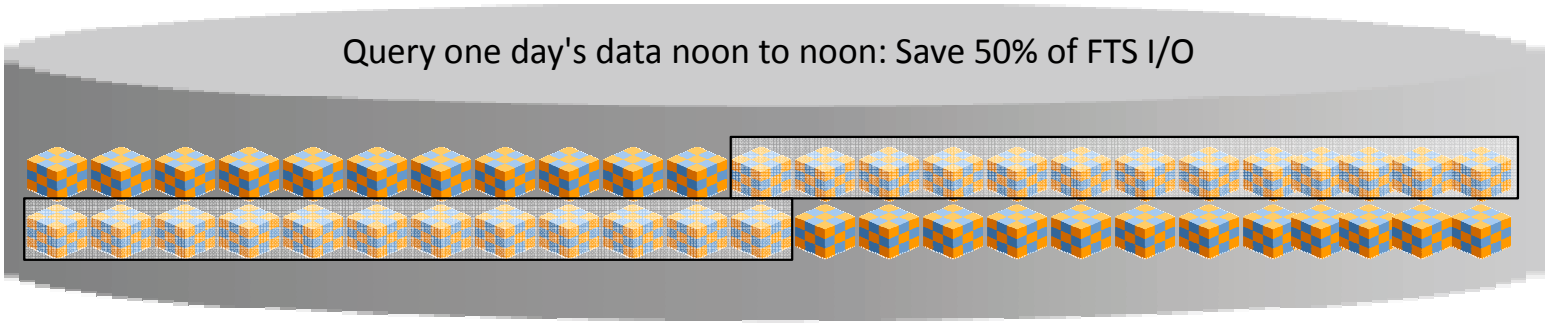
- Heap Tables
- External Tables
- Global Temporary Tables
  - ON COMMIT DELETE ROWS
  - ON COMMIT PRESERVE ROWS
- Index Organized Tables
- Partitioned Tables
- XML Tables

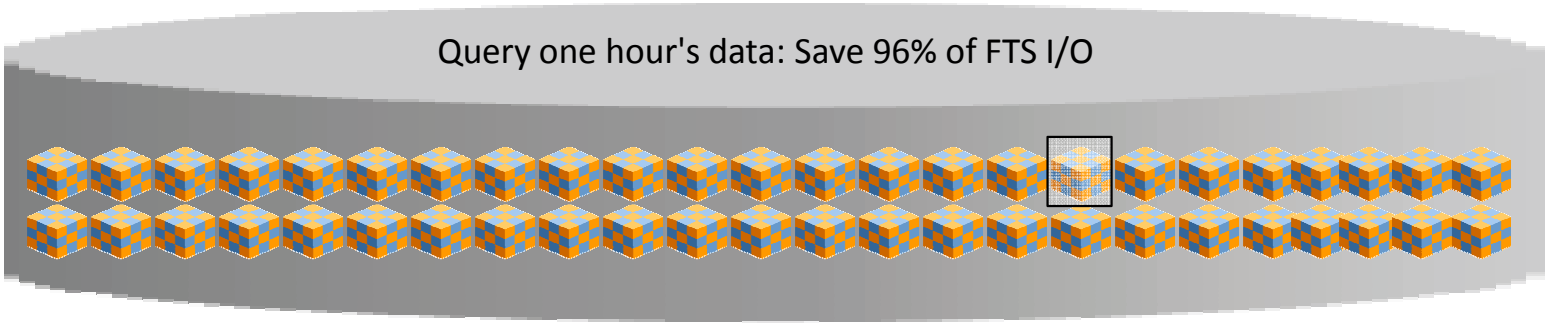
Partitioning and Subpartitioning Strategies

- The biggest mistake made: Not partitioning
- The second biggest mistake: Not partitioning enough
- Partitioning types
  - Hash
  - Interval
  - List
  - Range
  - Reference
  - System
  - Virtual Column
- Subpartitioning
  - List : Hash / List / Range
  - Range: Hash / List / Range

How Many Partitions Should I Create?







Partition XML By Virtual Column

```
DECLARE
 x XMLTYPE;
BEGIN
  x := XMLTYPE('<?xml version="1.0" encoding="utf-8"?>
       <Order orderId="1" orderRevision="1" orderTimeStamp="01-JAN-2012">
         <OrderHeader>
           <AlternateIds>
             <AlternateId altIdType="SiteId">2</AlternateId>
             <AlternateId altIdType="MerchantOrderNumber">Merch</AlternateId>
             <AlternateId altIdType="MarketplaceOrderNumber">Place</AlternateId>
             <AlternateId altIdType="CustomerReferenceId">Ref</AlternateId>
             <AlternateId altIdType="CartId">Cart</AlternateId>
             <AlternateId altIdType="SessionId">1</AlternateId>
           </AlternateIds>
         </OrderHeader>
       </Order>');
  INSERT INTO orders VALUES (x);
  COMMIT;
END;
/

CREATE TABLE orders OF XMLType
XMLTYPE STORE AS BINARY XML
VIRTUAL COLUMNS (SITE_ID AS (XMLCast(XMLQuery('/Order/@SiteId' PASSING OBJECT_VALUE RETURNING CONTENT) AS NUMBER)))
PARTITION BY RANGE (site_id) (
PARTITION p1 VALUES LESS THAN (10),
PARTITION p2 VALUES LESS THAN (20),
PARTITION pm VALUES LESS THAN (MAXVALUE));
```

Side Note: Avoid Hash Partitioning

- Impossible to drop legacy data by partition forcing delete statements on tables that are, by definition, large
- Impossible to set older data set to read only
- Impossible to compress older data sets to save room
- Impossible to implement life-cycle management and data tiering

Table Attributes

- Compressed
  - Minimize disk I/O on systems where cpu resources are available
  - Just using the default, no license fee, compression option will increase rows per block by 10% if byte-increasing updates are not anticipated
  - Do you have the Advanced Compression license?
  - Should you use Advanced Compression?
- READ ONLY
  - One fewer recursive SQL statements when SELECTing from a READ ONLY table

Too Many Columns (1:2)

- Oracle claims that a table can contain up to 1,000 columns: It is not true
- The maximum number of real table columns is 255
- Break the 255 barrier and optimizations such as compression no longer work
- A 1,000 column table is actually four tables joined together seamlessly behind the scenes just as a partitioned table appears to be a single segment but isn't
- Be suspicious of any table with more than 50 columns. At 100 columns it is time to reread the Codd-Date rules on normalization
- Think vertically not horizontally

Too Many Columns (2:2)

- Be very suspicious of any table with column names in the form "SPARE1", "SPARE2"
- The more columns a table has the more cpu is required when accessing columns to the right (as the table is displayed in a SELECT * query)

Column Order (1:3)

- Computers are not humans and tables are not paper forms
- CBO's column retrieval cost
  - Oracle stores columns in variable length format
  - Each row is parsed in order to retrieve one or more columns
  - Each subsequently parsed column introduces a cost of 20 cpu cycles regardless of whether it is of value or not

Column Order (2:3)

- These tables will be accessed by person_id or state: No one will ever put the address2 column into the WHERE clause as a filter

Common Design

```
CREATE TABLE customers (
person_id   NUMBER,
first_name  VARCHAR2(30) NOT NULL,
middle_init VARCHAR2(2),
last_name   VARCHAR2(30) NOT NULL,
address1    VARCHAR2(30),
address2    VARCHAR2(30),
city        VARCHAR2(30),
state       VARCHAR2(2));
```

Optimized Design

```
CREATE TABLE customers (
person_id   NUMBER,
last_name   VARCHAR2(30) NOT NULL,
state       VARCHAR2(2)  NOT NULL,
city        VARCHAR2(30) NOT NULL,
first_name  VARCHAR2(30) NOT NULL,
address1    VARCHAR2(30),
address2    VARCHAR2(30),
middle_init VARCHAR2(2));
```

Column Order (3:3)

- Proof column order matters

```
CREATE TABLE read_test AS
SELECT *
FROM apex_040200.wwv_flow_page_plugs
WHERE rownum = 1;

SQL> explain plan for
  2  select * from read_test;

PLAN_TABLE_OUTPUT
--------------------------------------------------------------------------------
| Id  | Operation         | Name      | Rows  | Bytes | Cost (%CPU)| Time     |
--------------------------------------------------------------------------------
|   0 | SELECT STATEMENT  |           |     1 |   214K|     2   (0)| 00:00:01 |
|   1 |  TABLE ACCESS FULL| READ_TEST |     1 |   214K|     2   (0)| 00:00:01 |
--------------------------------------------------------------------------------

-- fetch value from column 1
Final cost for query block SEL$1 (#0) - All Rows Plan:
  Best join order: 1
  Cost: 2.0002  Degree: 1  Card: 1.0000  Bytes: 13
  Resc: 2.0002  Resc_io: 2.0000  Resc_cpu: 7271
  Resp: 2.0002  Resp_io: 2.0000  Resc_cpu: 7271

-- fetch value from column 193
Final cost for query block SEL$1 (#0) - All Rows Plan:
  Best join order: 1
  Cost: 2.0003  Degree: 1  Card: 1.0000  Bytes: 2002
  Resc: 2.0003  Resc_io: 2.0000  Resc_cpu: 11111
  Resp: 2.0003  Resp_io: 2.0000  Resc_cpu: 11111
```

LOB Chunks

- Specifies the number of bytes to be allocated for LOB manipulation, always a multiple of the database block size
- Storing the block out-of-line will almost always be the optimal configuration
- As rule of thumb ... set CHUNK to 8K if the BLOB is 8K or smaller else set to 32K for best performance

LOB PCTVERSION

- Where do LOBs store undo when you update or delete them?
- Hint ... it isn't in the UNDO tablespace
- PCTVERSION specifies the maximum percentage of overall LOB storage space used for maintaining old versions of the LOB and when you read the LOB you read the undo too

SECUREFILES

- Introduced in Oracle 11g
- Designed specifically to enhance performance and scalability of LOB storage
- Poorly name they have nothing to do with security

```
CREATE TABLE secure_file_tab (
rid  NUMBER(5),
bcol BLOB)
LOB (bcol) STORE AS SECUREFILE bcol_lob (
TABLESPACE uwdata
DISABLE  STORAGE IN ROW
CHUNK 8192
RETENTION MIN 3600
KEEP_DUPLICATES
NOCOMPRESS
DECRYPT
CACHE READS);
```

SECUREFILE TUNING

- Set the db_securefile init parameter to PREFERED
- Determine whether enabling COMPRESSION and/or DEDUPLICATION will enhance your application's performance by testing with real data

```
ALTER SYSTEM SET db_securefile = 'PREFERRED'
COMMENT='Altered 01-APR-2014 for demo'
SCOPE=BOTH;
```

Object Selection

- Clusters
- Tables and Table Design
- Constraints
  - Unnecessary check constraints
- Indexes
- Sequences
- Materialized Views

Primary Keys Are Not Nullable

```
SQL> CREATE TABLE person (
  2  person_id NUMBER,
  3  last_name VARCHAR2(30));

Table created.

 SQL> ALTER TABLE person
  2  ADD CONSTRAINT pk_person
  3  PRIMARY KEY (person_id);

Table altered.

SQL> CREATE TABLE customer (
  2  customer_id NUMBER NOT NULL,
  3  cust_name   VARCHAR2(30));

Table created.

 SQL> ALTER TABLE customer
  2  ADD CONSTRAINT pk_customer
  3  PRIMARY KEY (customer_id);

Table altered.
```

```
SQL> desc person
 Name                      Null?    Type
 ------------------------- -------- --------------
 PERSON_ID                 NOT NULL NUMBER
 LAST_NAME                          VARCHAR2(30)

SQL> desc customer
 Name                      Null?    Type
 ------------------------- -------- --------------
 CUSTOMER_ID               NOT NULL NUMBER
 CUST_NAME                          VARCHAR2(30)

SQL> SELECT constraint_name, constraint_type
  2  FROM user_constraints
  3  WHERE table_name IN ('PERSON', 'CUSTOMER');

CONSTRAINT_NAME                C
------------------------------ -
SYS_C0011581                   C
PK_PERSON                      P
PK_CUSTOMER                    P
```

Will Oracle use CPU to validate the unnecessary check constraint?

Object Selection

- Clusters
- Tables and Table Design
- Constraints
- Indexes
  - Using the wrong type
  - Clustering Factor
- Sequences
- Materialized Views

Index Types

- B*Tree
- Bitmap
- Bitmap Join
- Compressed
- Descending
- Function Based
- Hash
- IndexType
- Invisible
- Reverse
- XML

Bitmap Indexes and Redo

- A bitmap index creates two bitmaps per index block
- In an 8K block there are two 4K bitmaps
- In a 32K block there are two 16K bitmaps
- The amount of redo generated by any change to a bitmap index is the size of the bitmap that was altered

Compressed Indexes (1:2)

- Index Compression
  - Indexes that compress well, especially those that have multiple repeating values in the leading columns can have their I/O vastly reduced by compression

```
SQL> CREATE TABLE postal_codes(
     city        VARCHAR2(30),
     state       VARCHAR2(2),
     postal_code VARCHAR2(5);

SQL> CREATE INDEX ix_postal_codes_norm
     ON postal_codes(state, city)
     TABLESPACE uwdata;

SQL> CREATE INDEX ix_postal_codes_comp
     ON postal_codes(state, city)
     TABLESPACE uwdata
     COMPRESS 1;
```

Compressed Indexes (2:2)

- Function Based Index Compression
  - Avoid building large indexes when most of the values indexed will not be efficiently accessed with the index

```
SQL> CREATE TABLE fbidemo AS
     SELECT object_name, object_type, temporary
     FROM dba_objects;

SQL> CREATE INDEX ix_fbidemo
     ON fbidemo (temporary);

SQL> CREATE INDEX fbi_fbidemo
     ON fbidemo (DECODE(temporary, 'Y', 'Y', NULL));
```

Descending Indexes

- Optimal when an index read is associated with an ORDER BY DESCENDING clause

```
SQL> CREATE TABLE orders(
     order_id    NUMBER(9,0),
     ship_date   DATE;

SQL> CREATE INDEX ix_orders_reg
     ON orders(cust_id, ship_date);

SQL> CREATE INDEX ix_orders_desc
     ON orders(cust_id, ship_date DESC);

SQL> EXPLAIN PLAN FOR
     SELECT order_id
     FROM orders
     WHERE ship_date IS NOT NULL
     ORDER BY ship_date DESC;

-------------------------------------------------
| Id  | Operation              | Name           |
-------------------------------------------------
|   0 | SELECT STATEMENT       |                |
|   1 |  SORT ORDER BY         |                |
|*  2 |   INDEX FAST FULL SCAN| IX_ORDERS_DESC |
-------------------------------------------------
```

Function Based Indexes

- Without a Function Based Index a full table scan would be required as well as a calculation of every row

```
conn scott/tiger@pdborcl

SQL> desc emp
Name                 Null?    Type
----------------- -------- ------------
EMPNO                NOT NULL NUMBER(4)
ENAME                         VARCHAR2(10)
JOB                           VARCHAR2(9)
MGR                           NUMBER(4)
HIREDATE                      DATE
SAL                           NUMBER(7,2)
COMM                          NUMBER(7,2)
DEPTNO                        NUMBER(2)

CREATE INDEX fbi_emp_sal_x_comm
ON emp(sal + comm);

SELECT ename
FROM emp
WHERE (sal + comm) < 300000;
```

Invisible Indexes

- Invisible indexes are real indexes and are maintained as like any other index but, by default, are not visible to the optimizer
- Excellent for use when inappropriate use of an index might slow down the system but a small number of sessions will benefit from it

```
CREATE INDEX ix_invis
ON invis(table_name)
INVISIBLE;

ALTER SESSION SET optimizer_use_invisible_indexes = TRUE;
```

Reverse Key Indexes

- Real Application Clusters can experience performance degradation where modifications to the index are concentrated in a small set of leaf blocks
- A Reverse Key Index addresses this problem by reversing the bytes of each index column (except the rowid), while keeping the column order
- As a result inserts are distributed across all leaf blocks

```
SQL> CREATE INDEX rix_index_demo_person_id
     ON index_demo(person_id)
     TABLESPACE uwdata
     REVERSE;

 SELECT index_name, index_type
 FROM user_indexes;
```

Indexes: Full Table Scans Are Bad


```
SQL> explain plan for
  2  SELECT doc_name
  3  FROM t
  4  WHERE person_id = 221;

SQL> select * from table(dbms_xplan.display);
PLAN_TABLE_OUTPUT
---------------------------------------------------------------------------
| Id  | Operation         | Name | Rows  | Bytes | Cost (%CPU)| Time     |
---------------------------------------------------------------------------
|   0 | SELECT STATEMENT  |      |   216 |  6264 |    64   (4)| 00:00:01 |
|*  1 |  TABLE ACCESS FULL| T    |   216 |  6264 |    64   (4)| 00:00:01 |
---------------------------------------------------------------------------
```

```
SQL> explain plan for
  2  SELECT /*+ INDEX(t ix_t_person_id) */ doc_name
  3  FROM t
  4  WHERE person_id = 221;

SQL> select * from table(dbms_xplan.display);
PLAN_TABLE_OUTPUT
-----------------------------------------------------------------------------------------------
| Id  | Operation                   | Name           | Rows  | Bytes | Cost (%CPU)| Time     |
-----------------------------------------------------------------------------------------------
|   0 | SELECT STATEMENT            |                |   216 |  6264 |   216   (0)| 00:00:03 |
|   1 |  TABLE ACCESS BY INDEX ROWID| T              |   216 |  6264 |   216   (0)| 00:00:03 |
|*  2 |   INDEX RANGE SCAN          | IX_T_PERSON_ID |   216 |       |     1   (0)| 00:00:01 |
-----------------------------------------------------------------------------------------------
```

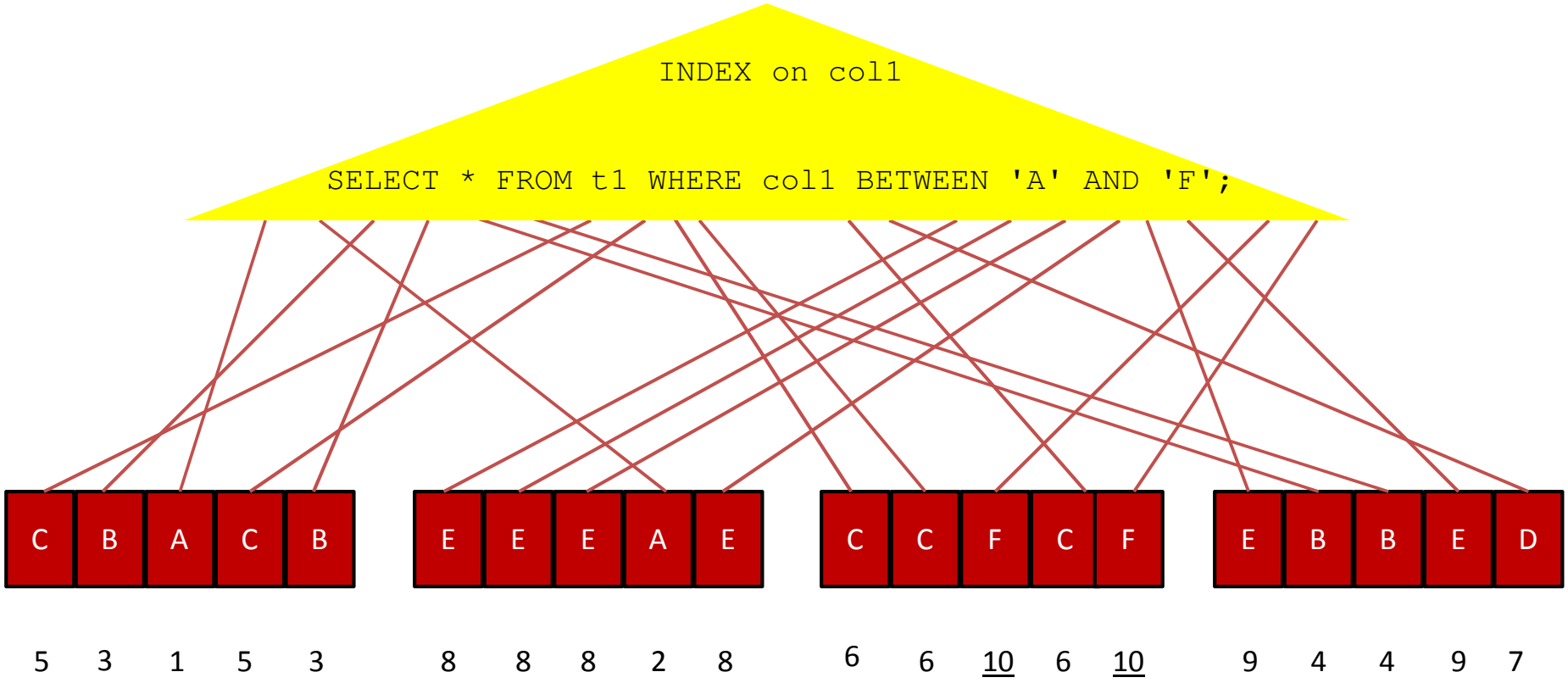
Clustering Factor (1:4)

- Nothing is more important to whether the optimizer chooses to use an index than the clustering factor: Nothing!
- Clustering Factor determines how well aligned, or ordered, the index entries are in relation to the rows in the table
- But the way clustering factor has been calculated up until now has been flawed
- The bug was fixed in 12cR1 and has been back-ported into 11.2.0.4

Clustering Factor (2:4)

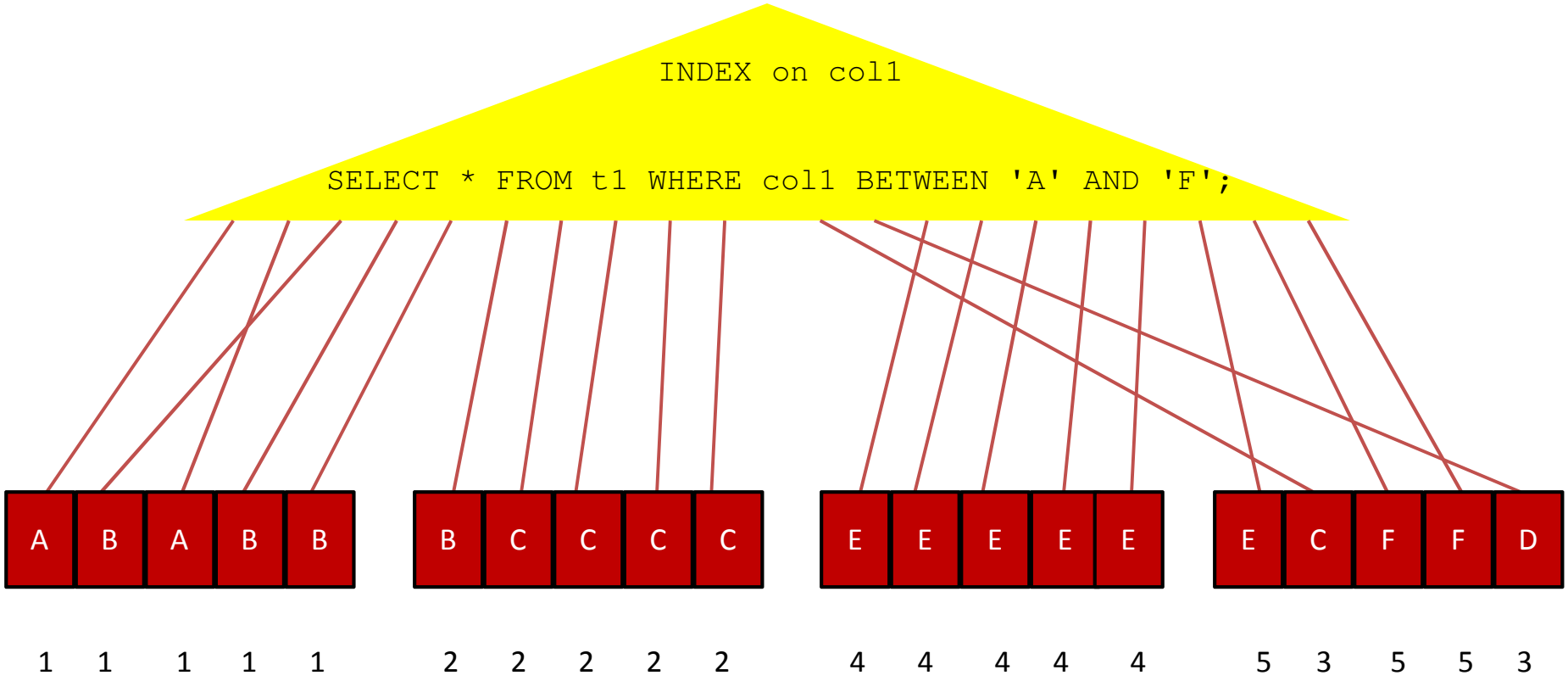
```
SQL> SELECT t.table_name, i.index_name, t.blocks TABLE_BLOCKS,
     i.clustering_factor
  2  FROM user_tables t, user_indexes I
  3  WHERE t.table_name = i.table_name
  4  ORDER BY i.clustering_factor, t.blocks;


TABLE_NAME        INDEX_NAME                    TABLE_BLOCKS CLUSTERING_FACTOR
----------------- ----------------------------- ------------ -----------------
SERVERS           PK_SERVERS                               5                 1
AIRPLANES         IX_PROGRAM_ID                         1055              1058
POSTAL_CODE       PK_ZIP_CODE                            126               251
SERV_INST         BIX_SERV_INST_WS_ID                     13               236
SERV_INST         PK_SERV_INST                            13               397
SERV_INST         BIX_SERV_INST_LOCATION_CODE             13               708
```

Clustering Factor (3:4)



A=2, B=4, C=5, D=1, E=6, F=2

Clustering Factor (4:4)



A=2, B=4, C=5, D=1, E=6, F=2

Invalid Indexes

- Will an invalid index affect performance? Yes!

```
SQL> SELECT owner, index_name, index_type
  2  FROM dba_indexes
  3  WHERE status NOT IN ('VALID', 'N/A')
  4  ORDER BY 1,3,2;

OWNER                      INDEX_NAME                     INDEX_TYPE
-------------------------- ------------------------------ ----------
GSM_DH                     ERC_PATH_HR_IDX                NORMAL
GSM_DH                     NOK_PATH_HR_IDX                NORMAL
GSM_DH                     NOR_PATH_HR_IDX                NORMAL
GSM_DH                     SIE_PATH_HR_IDX                NORMAL
KC4983                     LOCATION_SPATIAL_IDX           DOMAIN
KC4983                     SYS_C0021882                   NORMAL
UMTS_DH                    ERC_BBP_NODEB_KPI_IDX          NORMAL
UMTS_DH                    ERC_CELL_1030_IDX              NORMAL
UMTS_DH                    ERC_CELL_DH_IDX                NORMAL
UMTS_DH                    ERC_CELL_KPI_IDX_BACKUP        NORMAL
UMTS_DH                    ERC_SECTOR_1030_IDX            NORMAL
UMTS_DH                    ERC_SECTOR_DH_IDX              NORMAL
UMTS_DH                    ERC_SECTOR_KPI_IDX_BACKUP      NORMAL
```

Index Rebuilds

- The vast majority of index rebuilds are pure nonsense that wastes both cpu and I/O
- Most of what a rebuild does can be accomplished with far less overhead using COALESCE

```
-- 11g
ALTER INDEX ix_index_demo_gender_state COALESCE;

-- 12c
ALTER INDEX ix_index_demo_gender_state COALESCE CLEANUP;
```

Object Selection

- Clusters
- Tables and Table Design
- Constraints
- Indexes
- Sequences
  - Cache Size
  - Ordered
- Materialized Views

Sequences (1:6)

- Two clauses in CREATE SEQUENCE can have major impacts on performance
  - CACHE / NOCACHE
  - ORDER / NOORDER

```
CREATE SEQUENCE <sequence_name>
INCREMENT BY <integer>
START WITH <integer>
MAXVALUE <integer> / NOMAXVALUE
MINVALUE <integer> / NOMINVALUE
CYCLE / NOCYCLE
CACHE <#> / NOCACHE
ORDER / NOORDER;
```

Sequences (2:6)

- If you see something like this plan on investing some time in sequence tuning

```
SQL> SELECT sequence_owner, cache_size, COUNT(*)
  2  FROM dba_sequences
  3  WHERE sequence_owner IN ('ADMIN','PMCM','OPS','OSSBACKEND','OSS_DICTIONARY','SRE','WEBWIZ','ZZYZX')
  4  GROUP BY sequence_owner, cache_size
  5  ORDER BY 1,2;


SEQUENCE_OWNER                 CACHE_SIZE   COUNT(*)
------------------------------ ---------- ----------
ADMIN                                  20          1
PMCM                                 1000          2
OPS                                     0          1
OSSBACKEND                              0         11
OSS_DICTIONARY                          0         23
SRE                                     0         33
SRE                                    20          7
WEBWIZ                                 20         24
ZZYZX                                   0         21
ZZYZX                                  20         44
```

Sequences (3:6)

- If you see something like this plan on investing some time in sequence tuning

```
SQL> SELECT sequence_owner, order_flag, COUNT(*)
  2  FROM dba_sequences
  3  WHERE sequence_owner IN ('ADMIN','PMCM','OPS','OSSBACKEND','OSS_DICTIONARY','SRE','WEBWIZ','ZZYZX')
  4  GROUP BY sequence_owner, order_flag
  5  ORDER BY 1,2;


SEQUENCE_OWNER                 O   COUNT(*)
------------------------------ - ----------
ADMIN                          N          1
PMCM                           N          2
OPS                            N          1
OSSBACKEND                     N          5
OSSBACKEND                     Y          6
OSS_DICTIONARY                 N          7
OSS_DICTIONARY                 Y         16
SRE                            N         39
SRE                            Y          1
WEBWIZ                         Y         24
ZZYZX                          N         59
ZZYZX                          Y          6
```

Sequences (4:6)

- Specifies how many values of the sequence the database preallocates and keeps in memory for faster access
- In the event of a system failure all unused cached sequence values are lost
- There is substantial overhead, and locking each time the sequence pre-allocates a new set of values
- Oracle's default cache value of 20 is a bad joke and should never be used for anything other than demos of what is wrong with a value of 20

Sequences (5:6)

- On RAC clusters watch for the DFS Lock Handle wait as an indication of incorrectly sized caching: If found increase the caching size
- If a sequence is used to create a surrogate key and there are 1,000 inserts each second caching 20 values means that 500 times a second the cache will need to be refreshed
- If running on a RAC cluster 500 times a second the instance caches will need to be locked and synchronized
- This will negatively impact performance

Sequences (6:6)

- By default sequences are guaranteed to create unique numbers but when you force ordering you force serialization which will negatively impact performance
- Do not order sequences unless the application demands it

> "You would be amazed what setting a sequence cache via alter sequence to 100,000 or more can do during a large load -- amazed."
>
> ~ Tom Kyte

Object Selection

- Clusters
- Tables and Table Design
- Constraints
- Indexes
- Sequences
- Materialized Views
  - Fast Refresh on Commit
  - Fast Refresh on Demand
  - Force Refresh
  - Complete Refresh

Materialized Views (1:2)

- A materialized view is a database object that contains the results of a query
- They are local copies of data located remotely, or are used to create summary tables based on aggregations of a table's data
- Used to speed performance based on a specific query
- Materialized views which store data based on remote tables are also known as snapshots

Materialized Views (2:2)

- Materialized views share some characteristics of non-materialized views and indexes
- Materialized views are similar to indexes in the following ways
  - They contain actual data and consume storage space
  - They can be refreshed when the data in their master tables changes
  - They can improve performance of SQL execution when used for query rewrite operations
  - Their existence is transparent to SQL applications and users
- Unlike indexes, users can query materialized views directly using SELECT statements
- Depending on the types of refresh that are required, the views can also be updated with DML statements

Database Layer
SQL Tuning

SQL Tuning Topics

- How big is your SQL?
- Stop doing unnecessary work
  - Coalesce vs NVL
  - DML Optimizations
  - Fully Qualified Names
  - Hints
  - ORDER BY Clause
  - Reserved Words
  - Result Cache
  - Unnecessary Overhead
- Oracle Tuning Tools

I'm Not Afraid To Show Mine In Public

How Big Is Your SQL? (1:8)

0cjpvvqy854kh	SELECT 'ERICSSON_CSCORE.EOS' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from (select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then MSC end) ontime_count, count (*) num_of_rows from ERICSSON_CSCORE.EOS where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60)rawt, (select max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then MSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(s ysdate)+7/24 then MSC end) summarised_ontime_count, sum(entries) num_of_sumrows from ERICSSON_CSCORE.EOS_DY where datetime = trunc(sysdate)-1)sumt union all SELECT 'ERICSSON_CSCORE.MSCSTATS' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from (select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then MSC end) ontime_count, count (*) num_of_rows from ERICSSON_CSCORE.MSCSTATS where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60)rawt, (select max(datetime)sumday, count(dist inct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then MSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then MSC end) summarised_ontime_count, sum(entries) num_of_sumrows from ERICSSON_CSCORE.MSCSTATS_DY where datetime = trunc(sysdate)-1)sumt union all SELECT 'ERICSSON_CSCORE.MTRAFTYPE' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from (select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then MSC end) ontime_count , count (*) num_of_rows from ERICSSON_CSCORE.MTRAFTYPE where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60)rawt, (select max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then MSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then MSC end) summarised_ontime_count, sum(entries) num_of_sumrows from ERICSSON_CSCORE.MTRAFTYPE_DY where datetime = trunc(sysdate)-1)sumt union all SELECT 'ERICSSON_GERAN.CELLGPRS' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from (select count(distinct case when trunc(d atetime_ins, 'mi') > trunc(sysdate)+3/24 then BSC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then BSC end) ontime_count, count (*) num_of_rows from ERICSSON_GERAN.CELLGPRS where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60)rawt, (select max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then BSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then BSC end) summarised_ontime_count, sum(entries) num_of_sumrows from ERICSSON_GERAN.CELLGPRS_DY where datetime = trunc(sysdate)-1)sumt union all SELECT 'ERICSSON_GERAN.CELLSTATS' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_ count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from (select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then BSC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then BSC end) ontime_count, count (*) num_of_rows from ERICSSON_GERAN.CELLSTATS where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60)rawt, (select max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then BSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then BSC end) summarised_ontime_count, sum(entries) num_of_sumrows from ERICSSON_GERAN.CELLSTATS_DY where datetime = trunc(sysdate)-1)sumt union all SELECT 'ERICSSON_PSCORE.MM' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count,

How Big Is Your SQL? (2:8)

```
(ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count,
2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count curr ent_raw_elements,
summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff,
percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then SGSNID end) late_count, count(distinct case when trunc(datetime_ins, 'mi') <
trunc(sysdate)+3/24 then SGSNID end) ontime_count, count (*) num_of_rows from ERICSSON_PSCORE.MM where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday, count(distinct case
when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then SGSNID end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then SGSNID end) summarised_ontime_count, sum(entries)
num_of_sumrows from ERICSSON_PSCORE.MM_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'ERICSSON_PSCORE.SGSNSTAT' AS TABLENAME, sunday, CASE WHEN percent(summarised_ontime_count,
(ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0
else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+la te_count, 2)percentage_of_total_by5am,
percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count
current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2)
accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then SGSNID end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then SGSNID end) ontime_count,
count (*) num_of_rows from ERICSSON_PSCORE.SGSNSTAT where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday, count(distinct case when trunc(datetime_ins, 'mi') >
trunc(sysdate)+7/24 then SGSNID end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then SGSNID end) summarised_ontime_count, sum(entries) num_of_sumrows from
ERICSSON_PSCORE.SGSNSTAT_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'ERICSSON_UTRAN.RBS_CARRIER' AS TABLENAME, sunday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count),
2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count- late_count), 2) < 0 then 0 else
100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am,
percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count
current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2)
accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then NODEB end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then NODEB end) ontime_count,
count (*) num_of_rows from ERICSSON_UTRAN.RBS_CARRIER where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday, count(distinct case when trunc(datetime_ins, 'mi') >
trunc(sysdate)+7/24 then NODEB end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then NODEB end) summarised_ontime_count, sum(entries) num_of_sumrows from
ERICSSON_UTRAN.RBS_CARRIER_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'ERICSSON_UTRAN.RBS_EDCHRES' AS TABLENAME, sunday, CASE WHEN percent(summarised_ontime_count, (ontime_count-l
ate_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else
100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am,
percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count
current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2)
accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then NODEB end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then NODEB end) ontime_count,
count (*) num_of_rows from ERICSSON_UTRAN.RBS_EDCHRES where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday, count(distinct case when trunc(datetime_ins, 'mi') >
trunc(sysdate)+7/24 then NODEB end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then NODEB end) summarised_ontime_count, sum(entries) num_of_sumrows from ERICSSON_UT
RAN.RBS_EDCHRES_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'ERICSSON_UTRAN.RBS_HSDSCHRES' AS TABLENAME, sunday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100
THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else
100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am,
percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count
current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2)
accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then NODEB end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then NODEB end) ontime_count,
count (*) num_of_rows from ERICSSON_UTRAN.RBS_HSDSCHRES where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday, count(distinct case when trunc(datetime_ins, 'mi') >
trunc(sysdate)+7/24 then NODEB e nd) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then NODEB end) summarised_ontime_count, sum(entries) num_of_sumrows from
ERICSSON_UTRAN.RBS_HSDSCHRES_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'ERICSSON_UTRAN.RNC_IURLINK' AS TABLENAME, sunday, CASE WHEN percent(summarised_ontime_count, (ontime_count-
late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else
100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am,
```

How Big Is Your SQL? (3:8)

```
percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count
current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2)
accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then RNC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then RNC end) ontime_count, count (*)
num_of_rows from ERICSSON_UTRAN.RNC_IURLINK where dateti me between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24
then RNC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then RNC end) summarised_ontime_count, sum(entries) num_of_sumrows from ERICSSON_UTRAN.RNC_IURLINK_DY
where datetime = trunc(sysdate)-1 )sumt union all SELECT 'ERICSSON_UTRAN.RNC_RNCFUNC' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE
percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count,
(ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count,
ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count
SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') >
trunc(sysdate)+3/24 then RNC end) late_c ount, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then RNC end) ontime_count, count (*) num_of_rows from ERICSSON_UTRAN.RNC_RNCFUNC where datetime
between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then RNC end) summarised_late_count, count(distinct case when
trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then RNC end) summarised_ontime_count, sum(entries) num_of_sumrows from ERICSSON_UTRAN.RNC_RNCFUNC_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT
'ERICSSON_UTRAN.RNC_UCELL' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END
percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed,
percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count
current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-
late_count availabl e_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then RNC end) late_count, count(distinct case when
trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then RNC end) ontime_count, count (*) num_of_rows from ERICSSON_UTRAN.RNC_UCELL where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select
max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then RNC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then RNC end)
summarised_ontime_count, sum(entries) num_of_sumrows from ERICSSON_UTRAN.RNC_UCELL_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_CSCORE.BASE_OM' AS TABLENAME, sumday, CASE WHEN
percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count,
(ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count,
2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_
late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows,
num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then MSC end)
ontime_count, count (*) num_of_rows from NORTEL_CSCORE.BASE_OM where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') >
trunc(sysdate)+7/24 then MSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then MSC end) summarised_ontime_count, sum(entries) num_of_sumrows from
NORTEL_CSCORE.BASE_OM_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_CSCORE.C7SP' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN
100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else
100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_c
ount+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements,
summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select
count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then MSC end) ontime_count, count (*) num_of_rows from
NORTEL_CSCORE.C7SP where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then MSC end)
summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then MSC end) summarised_ontime_count, sum(entries) num_of_sumrows from NORTEL_CSCORE.C7SP_DY where datetime =
trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_CSCORE.CCS7' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count,
(ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end
percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2)
current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM,
late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end)
```

How Big Is Your SQL? (4:8)

```
late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then MSC end) ontime_count, count (*) num_of_rows from NORTEL_CSCORE.CCS7 where datetime between trunc(sysdate)-1 and
trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then MSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') <
trunc(sysdate)+7/24 then MSC end) summarised_ontime_count, sum(entries) num_of_sumrows from NORTEL_CSCORE.CCS7_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_CSCORE.LOCATIONAREACODE' AS
TABLENAME, sunday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cut off, case when
100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count,
ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements,
summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff,
percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') <
trunc(sysdate)+3/24 then MSC end) ontime_count, count (*) num_of_rows from NORTEL_CSCORE.LOCATIONAREACODE where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday,
count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then MSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then MSC end) summarised_ontime_count,
sum(entries) num_of_sumrows from NORTEL_CSCORE.LOCATIONAREACODE_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_CSCORE.MSCCP_CP' AS TABLENAME, sumd ay, CASE WHEN
percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count,
(ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count,
2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements,
summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff,
percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') <
trunc(sysdate)+3/24 then MSC end) ontime_count, count (*) num_of_rows from NORTEL_CSCORE.MSCCP_CP where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday, count(distinct
case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then MSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then MSC end) summarised_ontime_c ount, sum(entries)
num_of_sumrows from NORTEL_CSCORE.MSCCP_CP_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_CSCORE.MSCCP' AS TABLENAME, sunday, CASE WHEN percent(summarised_ontime_count, (ontime_count-
late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else
100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am,
percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count
current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2)
accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then MSC end) ontime_count, count (*)
num_of_rows from NORTEL_CSCORE.MSCCP where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(s ysdate)+7/24 then
MSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then MSC end) summarised_ontime_count, sum(entries) num_of_sumrows from NORTEL_CSCORE.MSCCP_DY where datetime =
trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_CSCORE.TRKGRP' AS TABLENAME, sunday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count,
(ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end
percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2)
current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM,
late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end)
late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then MSC end) ontime_count, count (*) num_of_rows from NORTEL_CSCORE.TRKGRP where dateti me between trunc(sysdate)-1 and
trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then MSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') <
trunc(sysdate)+7/24 then MSC end) summarised_ontime_count, sum(entries) num_of_sumrows from NORTEL_CSCORE.TRKGRP_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_CSCORE.VLR6' AS TABLENAME,
sunday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when
100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count,
ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements,
summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff,
percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then MSC end) late_count, count(d istinct case when trunc(datetime_ins, 'mi') <
trunc(sysdate)+3/24 then MSC end) ontime_count, count (*) num_of_rows from NORTEL_CSCORE.VLR6 where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sunday, count(distinct case
```

How Big Is Your SQL? (5:8)

```
when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then MSC end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then MSC end) summarised_ontime_count, sum(entries)
num_of_sumrows from NORTEL_CSCORE.VLR6_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_PSCORE.GSCGMM' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-
late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else
100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am,
percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count
current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumro ws, num_of_rows, 2)
accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then SGSN end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then SGSN end) ontime_count, count
(*) num_of_rows from NORTEL_PSCORE.GSCGMM where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24
then SGSN end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then SGSN end) summarised_ontime_count, sum(entries) num_of_sumrows from NORTEL_PSCORE.GSCGMM_DY where
datetime = trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_PSCORE.GSCSM_ACT' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE
percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count,
(ontime_count-late_count), 2) end percent_of_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count,
ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count
SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') >
trunc(sysdate)+3/24 then SGSN end) late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then SGSN end) ontime_count, count (*) num_of_rows from NORTEL_PSCORE.GSCSM_ACT where datetime
between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then SGSN end) summarised_late_count, count(distinct case when
trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then SGSN end) summarised_ontime_count, sum(entries) num_of_sumrows from NORTEL_PSCORE.GSCSM_ACT_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT
'NORTEL_PSCORE.GSCSM' AS TABLENAME, sumday, CASE WHEN percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END
percent_of_cutoff, case when 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_of_cutoff_not_summed,
percent(summarised_ontime_count, ontime_count+late_count, 2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count
current_raw_elements, summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-
late_count available_at_cutoff, percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then SGSN end) late_count, count(distinct case when
trunc(datetime_ins, 'mi') < trunc(sysdate)+3/24 then SGSN end) ontime_count, count (*) num_of_rows from NORTEL_PSCORE.GSCSM where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select
max(datetime)sumday, count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then SGSN end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then SGSN end)
summarised_ontime_count, sum(entries) num_of_sumrows from NORTEL_PSCORE.GSCSM_DY where datetime = trunc(sysdate)-1 )sumt union all SELECT 'NORTEL_PSCORE.GSDSTATS' AS TABLENAME, sumday, CASE WHEN
percent(summarised_ontime_count, (ontime_count-late_count), 2) > 100 THEN 100 ELSE percent(summarised_ontime_count, (ontime_count-late_count), 2) END percent_of_cutoff, case when 100-percent(summarised_ontime_count,
(ontime_count-late_count), 2) < 0 then 0 else 100-percent(summarised_ontime_count, (ontime_count-late_count), 2) end percent_o f_cutoff_not_summed, percent(summarised_ontime_count, ontime_count+late_count,
2)percentage_of_total_by5am, percent(summarised_ontime_count+summarised_late_count, ontime_count+late_count, 2) current_percentage, ontime_count+late_count current_raw_elements,
summarised_ontime_count+summarised_late_count current_summary_elements, summarised_ontime_count-summarised_late_count SUM_ELEMENTS_AT5AM, late_count late_raw_elements, ontime_count-late_count available_at_cutoff,
percent(num_of_sumrows, num_of_rows, 2) accuracy from ( select count(distinct case when trunc(datetime_ins, 'mi') > trunc(sysdate)+3/24 then SGSN end) late_count, count(distinct case when trunc(datetime_ins, 'mi') <
trunc(sysdate)+3/24 then SGSN end) ontime_count, count (*) num_of_rows from NORTEL_PSCORE.GSDSTATS where datetime between trunc(sysdate)-1 and trunc(sysdate)-1/24/60 )rawt, ( select max(datetime)sumday, count(distinct
case when trunc(datetime_ins, 'mi') > trunc(sysdate)+7/24 then SGSN end) summarised_late_count, count(distinct case when trunc(datetime_ins, 'mi') < trunc(sysdate)+7/24 then SGSN end) summarised_ontime_count, sum(entries)
num_of_sumrows from NORTEL_PSCORE.GSDSTATS_DY where datetime = trunc(sysdate)-1 )sumt
```

Customer's Original SQL (6:8)

```
SELECT DISTINCT E1_2.OBJECT_ID
   FROM PMCM.ELEMENT_DETAIL E1_1, PMCM.ELEMENT_DETAIL E1_2, PMCM.MARK_NETW_HIERARCHY H1,
        PMCM.ELEMENT_DETAIL E2_1, PMCM.ELEMENT_DETAIL E2_2, PMCM.MARK_NETW_HIERARCHY H2
   WHERE E1_1.OBJECT_ID = H1.PARENT_ID
     AND E1_2.OBJECT_ID = H1.OBJECT_ID
     AND E2_1.OBJECT_ID = H2.PARENT_ID
     AND E2_2.OBJECT_ID = H2.OBJECT_ID
     AND E1_1.CURRENT_IND = 'Y' AND E2_1.CURRENT_IND = 'Y'
     AND E2_1.CURRENT_IND = 'Y' AND E2_2.CURRENT_IND = 'Y'
     AND H1.CURRENT_IND = 'Y' AND H2.CURRENT_IND = 'Y'
     AND H1.HIERARCHY_TYPE = 'NETWORK' AND H2.HIERARCHY_TYPE = 'NETWORK'
     AND H1.PARENT_TYPE IN ('BSC','RNC')  AND H2.PARENT_TYPE IN ('BSC','RNC')
     AND E2_2.ELEMENT_TYPE = 'CELL' AND E1_2.ELEMENT_TYPE = 'CELL'
     AND H1.PARENT_TYPE IN ('BSC','RNC')
     AND E1_1.ELEMENT_NAME = E2_1.ELEMENT_NAME
     AND E1_1.ELEMENT_ID = E2_1.ELEMENT_ID
     AND E1_2.ELEMENT_NAME = E2_2.ELEMENT_NAME
     AND E1_2.ELEMENT_ID = E2_2.ELEMENT_ID
     AND E1_2.USEID LIKE '*%' AND E2_2.USEID NOT LIKE '*%';

| Id  | Operation                    | Name                          | Rows  | Bytes |TempSpc| Cost (%CPU)| Time     | Pstart| Pstop |
|   0 | SELECT STATEMENT             |                               |     1 |    78 |       |    74M (40)| 50:54:42 |       |       |
|   1 |  TEMP TABLE TRANSFORMATION   |                               |       |       |       |            |          |       |       |
|   2 |   LOAD AS SELECT             |                               |       |       |       |            |          |       |       |
|   3 |    PARTITION RANGE ALL       |                               |   22M|  1111M|       | 38153  (11)| 00:01:34 |     1 |    29 |
|*  4 |     TABLE ACCESS FULL        | ELEMENT_DETAIL                |   22M|  1111M|       | 38153  (11)| 00:01:34 |       |       |
|   5 |   LOAD AS SELECT             |                               |       |       |       |            |          |       |       |
|   6 |    PARTITION HASH ALL        |                               |  337K|  9231K|       |  3514  (15)| 00:00:09 |     1 |    16 |
|*  7 |     TABLE ACCESS FULL        | MARK_NETW_HIERARCHY           |  337K|  9231K|       |  3514  (15)| 00:00:09 |       |       |
|   8 |   SORT AGGREGATE             |                               |     1 |    78 |       |            |          |       |       |
|*  9 |    HASH JOIN                 |                               |  927G|   65T|  534M|    74M (40)| 50:53:00 |       |       |
|  10 |     VIEW                     |                               |   22M|   277M|       | 16808  (12)| 00:00:42 |       |       |
|  11 |      TABLE ACCESS FULL       | SYS_TEMP_0FDA7485F_6A66C42E   |   22M|  1111M|       | 16808  (12)| 00:00:42 |       |       |
|* 12 |     HASH JOIN                |                               |   21G|  1272G|  534M|  1616K (43)| 01:06:04 |       |       |
|  13 |      VIEW                    |                               |   22M|   277M|       | 16808  (12)| 00:00:42 |       |       |
|  14 |       TABLE ACCESS FULL      | SYS_TEMP_0FDA7485F_6A66C42E   |   22M|  1111M|       | 16808  (12)| 00:00:42 |       |       |
|* 15 |      HASH JOIN               |                               |  476M|    23G|  524M| 97327  (22)| 00:03:59 |       |       |
|* 16 |       HASH JOIN              |                               |   10M|   401M| 8704K| 34520  (10)| 00:01:25 |       |       |
|* 17 |        HASH JOIN             |                               |  234K|  5948K| 8256K|   783  (10)| 00:00:02 |       |       |
|  18 |         VIEW                 |                               |  337K|  4286K|       |   142  (14)| 00:00:01 |       |       |
|  19 |          TABLE ACCESS FULL   | SYS_TEMP_0FDA74860_6A66C42E   |  337K|  3956K|       |   142  (14)| 00:00:01 |       |       |
|  20 |         VIEW                 |                               |  337K|  4286K|       |   142  (14)| 00:00:01 |       |       |
|  21 |          TABLE ACCESS FULL   | SYS_TEMP_0FDA74860_6A66C42E   |  337K|  3956K|       |   142  (14)| 00:00:01 |       |       |
|  22 |        VIEW                  |                               |   22M|   277M|       | 16808  (12)| 00:00:42 |       |       |
|  23 |         TABLE ACCESS FULL    | SYS_TEMP_0FDA7485F_6A66C42E   |   22M|  1111M|       | 16808  (12)| 00:00:42 |       |       |
|  24 |       VIEW                   |                               |   22M|   277M|       | 16808  (12)| 00:00:42 |       |       |
|  25 |        TABLE ACCESS FULL     | SYS_TEMP_0FDA7485F_6A66C42E   |   22M|  1111M|       | 16808  (12)| 0        |       |       |
```

Customer's SQL Rewrite (7:8)

```
-----------------------------------------------------------------------------------------------------------------------------
| Id  | Operation                  | Name                        | Rows  | Bytes |TempSpc| Cost (%CPU)| Time     | Pstart| Pstop |
-----------------------------------------------------------------------------------------------------------------------------
|   0 | SELECT STATEMENT           |                             |     1 |    78 |       |    14T(100)|999:59:59 |       |       |
|   1 |  TEMP TABLE TRANSFORMATION |                             |       |       |       |            |          |       |       |
|   2 |   LOAD AS SELECT           |                             |       |       |       |            |          |       |       |
|   3 |    PARTITION RANGE ALL     |                             |   22M|  1111M|       | 38153  (11)| 00:01:34 |     1 |    29 |
|*  4 |     TABLE ACCESS FULL      | ELEMENT_DETAIL              |   22M|  1111M|       | 38153  (11)| 00:01:34 |       |       |
|   5 |   LOAD AS SELECT           |                             |       |       |       |            |          |       |       |
|   6 |    PARTITION HASH ALL      |                             |  337K|  9231K|       |  3514  (15)| 00:00:09 |     1 |    16 |
|*  7 |     TABLE ACCESS FULL      | MARK_NETW_HIERARCHY         |  337K|  9231K|       |  3514  (15)| 00:00:09 |       |       |
|   8 |   SORT AGGREGATE           |                             |     1 |    78 |       |            |          |       |       |
|   9 |    MERGE JOIN              |                             |  471P|   15E|       |    14T(100)|999:59:59 |       |       |
|  10 |     MERGE JOIN             |                             |   10P|  616P|       |   694G (81)|999:59:59 |       |       |
|  11 |      MERGE JOIN            |                             |  231T|   10P|       |   377G (64)|999:59:59 |       |       |
|  12 |       SORT JOIN            |                             |  334T|   11P|   28P|   377G (64)|999:59:59 |       |       |
|  13 |        MERGE JOIN CARTESIAN|                             |  334T|   11P|       |   140G (14)|999:59:59 |       |       |
|* 14 |         HASH JOIN          |                             |  989M|   23G|  534M| 96010  (38)| 00:03:56 |       |       |
|  15 |          VIEW              |                             |   22M|  277M|       | 16808  (12)| 00:00:42 |       |       |
|  16 |           TABLE ACCESS FULL| SYS_TEMP_0FDA7485B_6A66C42E |   22M|  1111M|       | 16808  (12)| 00:00:42 |       |       |
|  17 |          VIEW              |                             |   22M|  277M|       | 16808  (12)| 00:00:42 |       |       |
|  18 |           TABLE ACCESS FULL| SYS_TEMP_0FDA7485B_6A66C42E |   22M|  1111M|       | 16808  (12)| 00:00:42 |       |       |
|  19 |         BUFFER SORT        |                             |  337K|  4286K|       |   140G (14)|999:59:59 |       |       |
|  20 |          VIEW              |                             |  337K|  4286K|       |   142  (14)| 00:00:01 |       |       |
|  21 |           TABLE ACCESS FULL| SYS_TEMP_0FDA7485C_6A66C42E |  337K|  3956K|       |   142  (14)| 00:00:01 |       |       |
|* 22 |       SORT JOIN            |                             |  337K|  4286K|   12M|   844  (14)| 00:00:03 |       |       |
|  23 |        VIEW                |                             |  337K|  4286K|       |   142  (14)| 00:00:01 |       |       |
|  24 |         TABLE ACCESS FULL  | SYS_TEMP_0FDA7485C_6A66C42E |  337K|  3956K|       |   142  (14)| 00:00:01 |       |       |
|* 25 |      SORT JOIN             |                             |   22M|  277M|  855M| 65084  (16)| 00:02:40 |       |       |
|  26 |       VIEW                 |                             |   22M|  277M|       | 16808  (12)| 00:00:42 |       |       |
|  27 |        TABLE ACCESS FULL   | SYS_TEMP_0FDA7485B_6A66C42E |   22M|  1111M|       | 16808  (12)| 0        |       |       |
|* 28 |     SORT JOIN              |                             |   22M|  277M|  855M| 65084  (16)| 00:02:40 |       |       |
|  29 |      VIEW                  |                             |   22M|  277M|       | 16808  (12)| 00:00:42 |       |       |
|  30 |       TABLE ACCESS FULL    | SYS_TEMP_0FDA7485B_6A66C42E |   22M|  1111M|       | 16808  (12)| 0        |       |       |
-----------------------------------------------------------------------------------------------------------------------------
```

Final Rewrite (8:8)

```
WITH ed AS (SELECT object_id, element_id, element_name, element_type, useid
            FROM pmcm.element_detail
            WHERE element_type = 'CELL'
            AND current_ind = 'Y'),
     mnh AS (SELECT parent_id, object_id
             FROM pmcm.mark_netw_hierarchy
             WHERE current_ind = 'Y'
             AND hierarchy_type = 'NETWORK'
             AND parent_type IN ('BSC', 'RNC'))
SELECT COUNT(*)
FROM ed e1_1, ed e1_2, ed e2_1, ed e2_2, mnh h1, mnh h2
WHERE e1_1.object_id = h1.parent_id AND   e1_2.object_id = h1.object_id
AND   e2_1.object_id = h2.parent_id AND   e2_2.object_id = h2.object_id
AND   e1_1.element_name = e2_1.element_name
AND   e1_1.element_id = e2_1.element_id
AND   e1_2.element_name = e2_2.element_name
AND   e1_2.element_id = e2_2.element_id
AND   e1_2.useid LIKE '*%'
AND   e2_2.useid NOT LIKE '*%';

| Id  | Operation                                | Name                | Rows  | Bytes |TempSpc| Cost (%CPU)| Time     |
-------------------------------------------------------------------------------------------------------------------------
|   0 | SELECT STATEMENT                         |                     |     1 |   214 |       |   100K  (6)| 00:04:08 |
|   1 |  HASH UNIQUE                             |                     |     1 |   214 |       |   100K  (6)| 00:04:08 |
|*  2 |   HASH JOIN                              |                     |     1 |   214 |    12M|   100K  (6)| 00:04:08 |
|   3 |    PARTITION HASH ALL                    |                     |  337K|  9231K|       |  3514  (15)| 00:00:09 |
|*  4 |     TABLE ACCESS FULL                    | MARK_NETW_HIERARCHY |  337K|  9231K|       |  3514  (15)| 00:00:00 |
|*  5 |    HASH JOIN                             |                     |  207K|    36M|    22M| 95860   (6)| 00:03:56 |
|   6 |     PARTITION RANGE ALL                  |                     |  586K|    15M|       | 16233   (2)| 00:00:40 |
|   7 |      TABLE ACCESS BY LOCAL INDEX ROWID   | ELEMENT_DETAIL      |  586K|    15M|       | 16233      | ??:??:?? |
|*  8 |       INDEX SKIP SCAN                    | ED_ET_TECH_CI       |  586K|       |       | 12791   (1)| 00:00:3? |
|*  9 |     HASH JOIN                            |                     |  207K|    31M|    22M| 77982   (7)| 00:03:12 |
|  10 |      PARTITION RANGE ALL                 |                     |  586K|    15M|       | 16233   (2)| 00:00:40 |
|  11 |       TABLE ACCESS BY LOCAL INDEX ROWID  | ELEMENT_DETAIL      |  586K|    15M|       | 16233      | ??:??:?? |
|* 12 |        INDEX SKIP SCAN                   | ED_ET_TECH_CI       |  586K|       |       | 12791   (1)| 00:00:?? |
|* 13 |      HASH JOIN                           |                     |  179K|    22M|    12M| 60372   (8)| 00:02:29 |
|  14 |       PARTITION HASH ALL                 |                     |  337K|  9231K|       |  3514  (15)| 00:00:09 |
|* 15 |        TABLE ACCESS FULL                 | MARK_NETW_HIERARCHY |  337K|  9231K|       |  3514  (15)| 00:00:?? |
|* 16 |       HASH JOIN                          |                     |  184K|    17M|    10M| 55886   (8)| 00:02:18 |
|  17 |        PARTITION RANGE ALL               |                     |  184K|  9008K|       | 37137   (8)| 00:01:32 |
|* 18 |         TABLE ACCESS FULL                | ELEMENT_DETAIL      |  184K|  9008K|       | 37137   (8)| 00:01:32 |
|  19 |        PARTITION RANGE ALL               |                     |  576K|    28M|       | 17383   (8)| 00:00:43 |
|* 20 |         TABLE ACCESS BY LOCAL INDEX ROWID| ELEMENT_DETAIL      |  576K|    28M|       | 17383   (8)| ??:??:?? |
|* 21 |          INDEX SKIP SCAN                 | ED_ET_TECH_CI       |  583K|       |       | 13939   (9)| 00:00:35 |
-------------------------------------------------------------------------------------------------------------------------
```

Coalesce

- Why are developers using NVL rather than COALESCE?
- Hopefully they are ... but they do they know why?

```
SELECT NVL(comm, 0)
FROM scott.emp;

 Final cost for query block SEL$1 (#0) - First K Rows Plan:
  Best join order: 1
  Cost: 3.0008  Degree: 1  Card: 10.0000  Bytes: 20
  Resc: 3.0008  Resc_io: 3.0000  Resc_cpu: 31186
  Resp: 3.0008  Resp_io: 3.0000  Resc_cpu: 31186

SELECT COALESCE(comm, 0)
FROM scott.emp;

 Final cost for query block SEL$1 (#0) - All Rows Plan:
  Best join order: 1
  Cost: 3.0010  Degree: 1  Card: 14.0000  Bytes: 28
  Resc: 3.0010  Resc_io: 3.0000  Resc_cpu: 39387
  Resp: 3.0010  Resp_io: 3.0000  Resc_cpu: 39387
```

DML Optimizations

- INSERT ALL
- INSERT ALL WHEN
- INSERT FIRST WHEN
- RETURNING CLAUSE

```
INSERT
WHEN (deptno=10) THEN
  INTO emp_10 (empno,ename,job,mgr,sal,deptno)
  VALUES (empno,ename,job,mgr,sal,deptno)
WHEN (deptno=20) THEN
  INTO emp_20 (empno,ename,job,mgr,sal,deptno)
  VALUES (empno,ename,job,mgr,sal,deptno)
WHEN (deptno=30) THEN
  INTO emp_30 (empno,ename,job,mgr,sal,deptno)
  VALUES (empno,ename,job,mgr,sal,deptno)
ELSE
  INTO leftover (empno,ename,job,mgr,sal,deptno)
  VALUES (empno,ename,job,mgr,sal,deptno)
SELECT * FROM emp;
```

Be sure you know all the basic syntaxes

```
INSERT ALL
INTO ap_cust VALUES (customer_id, program_id, delivered_date)
INTO ap_orders VALUES (order_date, program_id)
SELECT program_id, delivered_date, customer_id, order_date
FROM airplanes;
```

Fully Qualified Object Names (1:2)

- When you do not use fully qualified names Oracle must do the work for you
- You write code once ... the database executes it many times

```
SELECT DISTINCT s.srvr_id
FROM servers s, serv_inst i
WHERE s.srvr_id = i.srvr_id;

SELECT DISTINCT s.srvr_id
FROM uwclass.servers s, uwclass.serv_inst i
WHERE s.srvr_id = i.srvr_id;
```

Hints

- The following was written by Jonathan Lewis: I've never heard better advice

Rules for Hinting

1. **Don't**

2. **If you must use hints, then assume you've used them incorrectly.**

3. **On every patch or upgrade to Oracle, assume every piece of hinted SQL is going to do the wrong thing. Because of (2) above; you've been lucky so far, but the patch/upgrade lets you discover your mistake.**

4. **Every time you apply some DDL to an object that appears in a piece of hinted SQL assume that the hinted SQL is going to do the wrong thing. Because of (2) above; you've been lucky so far, but the structural change lets you discover your mistake.**

Monitor Parallel Queries

```
col value format 9999999999

SELECT * FROM v_$pq_sysstat;

SELECT * FROM v_$px_process;

SELECT * FROM v_$px_sesstat;

SELECT * FROM v_$px_process_sysstat;
```

ORDER BY Clause

- Simply put ... don't write it if you don't need to
- Verify often database queries perform a sort and then the front-end does it again

```
SQL> SELECT * FROM TABLE(dbms_xplan.display);

PLAN_TABLE_OUTPUT
--------------------------------------------------------------------------------
Plan hash value: 995087943

-------------------------------------------------------------------------------------
| Id  | Operation          | Name    | Rows  | Bytes |TempSpc| Cost (%CPU)| Time     |
-------------------------------------------------------------------------------------
|   0 | SELECT STATEMENT   |         | 98128 |    12M|       |  3435   (1)| 00:00:42 |
|   1 |  SORT ORDER BY     |         | 98128 |    12M|    25M|  3435   (1)| 00:00:42 |
|   2 |   TABLE ACCESS FULL| SOURCE$ | 98128 |    12M|       |   577   (2)| 00:00:07 |
-------------------------------------------------------------------------------------
```

The full tables scan took 7 seconds
The swapping and ordering took 35 more

Reserved Words

- Do not make the database work hard by using/abusing reserved words

```
SELECT keyword
FROM v$reserved_words
WHERE keyword LIKE 'ID%';

KEYWORD
------------------------------
ID
IDENTIFIED
IDGENERATORS
IDLE_TIME
IDENTITY
IDENTIFIER

6 rows selected.
```

Result Cache

SQL ordered by Executions

- Total Executions: 29,717,627
- Captured SQL account for 77.4% of Total

Executions	Rows Processed	Rows per Exec	CPU per Exec (s)	Elap per Exec (s)	SQL Id	SQL Module	SQL Text
10,128,178	2,506,529	0.25	0.00	0.00	932srzg1krc33	ASN_07B_DIP(004110016)	SELECT NE_TIMEZONE FROM CMPM.E...
7,576,759	7,579,197	1.00	0.00	0.00	1h698sb62un99	asci_56_RANAPProtocolStats(01611000E)	SELECT DISTINCT NE_TIMEZONE FR...
3,914,621	3,848,268	0.98	0.00	0.00	5tbzddgguu8cc	asci_56_RANAPProtocolStats(01611000E)	SELECT SYS_VERSION FROM CMPM.T...
311,645	311,604	1.00	0.00	0.00	7gtztzv329wq0		select c.name, u.name from co...
301,428	301,325	1.00	0.00	0.00	36s446f9cnwhw		SELECT C.NAME FROM COL$ C WHER...
200,692	200,669	1.00	0.00	0.00	4vs91dcv7u1p6	OMS	insert into sys.aud$(sessioni...
65,044	65,035	1.00	0.00	0.00	fz9xwpt2cvt0k		SELECT par_type, param_clob, ...
64,949	3,945,482	60.75	0.00	0.00	f5ra7dru5fk5n	XML_P7R_RNC_RCS(003110008)	SELECT NAME, PATH, READ, WR...
64,801	64,807	1.00	0.00	0.00	fhzj09a7fnrnb	XML_V7I_IN_LP_DC(00811000V)	SELECT DBTIMEZONE, LENGTH(DBT...
64,632	64,542	1.00	0.00	0.00	15jnrrb6016nd	XML_V7I_IN_LP_DC(00811000V)	SELECT SESSIONTIMEZONE, LENGT...

```
SELECT /*+ RESULT_CACHE */ srvr_id
FROM (
  SELECT srvr_id, SUM(cnt) SUMCNT
  FROM (
    SELECT DISTINCT srvr_id, 1 AS CNT
    FROM servers
    UNION ALL
    SELECT DISTINCT srvr_id, 1
    FROM serv_inst)
  GROUP BY srvr_id)
WHERE sumcnt = 2;
```

Unnecessary Overhead

- V$ is the local instance
- GV$ view hit all instances
  - Creating delays and increased cache fusion interconnect traffic
- Do not query a gv$ unless you really need to collect information from more than one node

```
SELECT name FROM v$database;

SELECT name FROM gv$database;
```

SQL Tuning Topics (1:2)

- Stop doing unnecessary work
- Oracle Tuning Tools
  - Explain Plan & Autotrace
  - DBMS_ADDM
  - DBMS_ADVANCED_REWRITE
  - DBMS_ADVISOR
  - DBMS_AUTO_SQLTUNE
  - DBMS_MONITOR
  - DBMS_SERVICE
  - DBMS_SESSION
  - DBMS_SPD
  - DBMS_SPM
  - DBMS_SQLDIAG
  - DBMS_SQLDIAG_INTERNAL

SQL Tuning Topics (2:2)

- Oracle Tuning Tools
  - DBMS_SQLPA
  - DBMS_SQLTUNE
  - DBMS_WORKLOAD_CAPTURE
  - DBMS_WORKLOAD_REPLAY
  - DBMS_WORKLOAD_REPOSITORY
  - DBMS_XPLAN
  - SQLTXPLAIN
  - Trace and TKPROF
  - TRCSESS
  - TRACE ANALYZER
  - ASH Reports
  - AWR Reports
  - SQL Baselines

Explain Plan (1:2)

- Everyone thinks they know how to do this

```
EXPLAIN PLAN FOR
SELECT dummy FROM dual;

SELECT * FROM TABLE(dbms_xplan.display);
```

- But are you making use of the rest of DBMS_XPLAN's functionality?
  - DBMS_XPLAN.DIFF_PLAN_CURSOR
  - DBMS_XPLAN.DISPLAY_AWR
  - DBMS_XPLAN.DISPLAY_CURSOR
  - DBMS_XPLAN.DISPLAY_SQL_PATCH_PLAN (new 12c)
  - DBMS_XPLAN.DISPLAY_SQL_PLAN_BASELINE
  - DBMS_XPLAN.DISPLAY_SQL_PROFILE_PLAN (new 12c)

Explain Plan (2:2)

- Can you read an explain plan well enough to know when it is lying to you?

```
EXPLAIN PLAN FOR
SELECT srvr_id
FROM servers
INTERSECT
SELECT srvr_id
FROM serv_inst;

 ---------------------------------------------------------------------------
 | Id | Operation                | Name         | Rows | Bytes | Cost (%CPU)|
 ---------------------------------------------------------------------------
 |  0 | SELECT STATEMENT         |              |  141 | 4560  |     6  (84)|
 |  1 |  INTERSECTION            |              |      |       |            |
 |  2 |   SORT UNIQUE NOSORT     |              |  141 |  564  |     2  (50)|
 |  3 |    INDEX FULL SCAN       | PK_SERVERS   |  141 |  564  |     1   (0)|
 |  4 |   SORT UNIQUE            |              |  999 | 3996  |     4  (25)|
 |  5 |    INDEX FAST FULL SCAN| IX_SERV_INST |  999 | 3996  |     3   (0)|
 ---------------------------------------------------------------------------
```

This one is ... can you find Oracle's math error

Autotrace

- Explain Plan guesses .. Autotrace doesn't

```
SQL> set autotrace on
SQL> select dummy from dual;

D
-
X

Execution Plan
----------------------------------------------------------
Plan hash value: 272002086

---------------------------------------------------------------------------
| Id  | Operation         | Name | Rows  | Bytes | Cost (%CPU)| Time     |
---------------------------------------------------------------------------
|   0 | SELECT STATEMENT  |      |     1 |     2 |     2   (0)| 00:00:01 |
|   1 |  TABLE ACCESS FULL| DUAL |     1 |     2 |     2   (0)| 00:00:01 |
---------------------------------------------------------------------------

Statistics
----------------------------------------------------------
          1  recursive calls
          0  db block gets
          3  consistent gets
          0  physical reads
          0  redo size
        538  bytes sent via SQL*Net to client
        544  bytes received via SQL*Net from client
          2  SQL*Net roundtrips to/from client
          0  sorts (memory)
          0  sorts (disk)
          1  rows processed
```

DBMS_ADVANCED_REWRITE

- Introduced at version 10gR1
- Allows the in-optimizer rewrite of SQL statements
- Package includes
  - ALTER_REWRITE_EQUIVALENCE
  - DECLARE_REWRITE_EQUIVALENCE
  - DROP_REWRITE_EQUIVALENCE
  - VALIDATE_REWRITE_EQUIVALENCE

DBMS_ADDM

- Introduced at version 11gR1
- Automatic Database Diagnostic Monitor
- Package includes
  - ANALYZE_DB
  - ANALYZE_INST
  - ANALYZE_PARTIAL
  - COMPARE_CAPTURE_REPLAY_REPORT (new 12c)
  - COMPARE_DATABASES (new 12c)
  - COMPARE_INSTANCES (new 12c)
  - COMPARE_REPLAY_REPLAY_REPORT (new 12c)

DBMS_ADVISOR

- Introduced at version 10gR1
- Part of the server manageability suite of advisors, a set of expert systems that identifies and helps resolve performance problems relating to database server components
- Package includes
  - CREATE_TASK
  - EVALUATE_DIRECTIVE
  - GET_TASK_REPORT
  - IMPLEMENT_TASK
  - QUICK_TUNE
  - TUNE_MVIEW
  - UPDATE_DIRECTIVE

DBMS_AUTO_SQLTUNE

- Introduced at version 12gR2
- Interface to SQL Tuning Advisor (DBMS_SQLTUNE) when run within the Autotask framework
- Package includes
  - EXECUTE_AUTO_TUNING_TASK
  - REPORT_AUTO_TUNING_TASK
  - SET_AUTO_TUNING_TASK_PARAMETER

DBMS_MONITOR

- Introduced at version 10gR1
- Package includes
  - CLIENT_ID_STAT_ENABLE & DISABLE
  - CLIENT_ID_TRACE_ENABLE & DISABLE
  - DATABASE_TRACE_ENABLE & DISABLE
  - SERV_MOD_ACT_STAT_DISABLE
  - SERV_MOD_ACT_STAT_ENABLE
  - SERV_MOD_ACT_TRACE_DISABLE
  - SERV_MOD_ACT_TRACE_ENABLE
  - SESSION_TRACE_ENABLE & DISABLE

DBMS_SERVICE

- Introduced at version 10gR1
- Package includes
  - CREATE_SERVICE (new 12c)
  - MODIFY_SERVICE (new 12c)

If you are not familiar with service parameters such as GOAL_SERVICE_TIME, GOAL_THROUGHPUT, CLB_GOAL_SHORT and CLB_GOAL_LONG put this on your list of topics to research

DBMS_SESSION

- First available in 7.3.4
- Built in objects
  - SESSION_TRACE_DISABLE
  - SESSION_TRACE_ENABLE
  - SET_SQL_TRACE

DBMS_SPD

- Introduced at version 12cR1
- Package includes
  - ALTER_SQL_PLAN_DIRECTIVE
  - FLUSH_SQL_PLAN_DIRECTIVE
  - SET_PREFS

DBMS_SPM

- Introduced at version 11gR1
- Package includes
  - ACCEPT_SQL_PLAN_BASELINE (new 12c)
  - ALTER_SQL_PLAN_BASELINE
  - CANCEL_EVOLVE_TASK (new 12c)
  - CREATE_EVOLVE_TASK (new 12c)
  - DROP_EVOLVE_TASK (new 12c)
  - EVOLVE_SQL_PLAN_BASELINE
  - EXECUTE_EVOLVE_TASK (new 12c)
  - IMPLEMENT_EVOLVE_TASK (new 12c)
  - MIGRATE_STORED_OUTLINE
  - REPORT_AUTO_EVOLVE_TASK (new 12c)
  - REPORT_EVOLVE_TASK (new 12c)

DBMS_SQLDIAG

- Introduced at version 11gR1
- SQL Repair Advisor
- Package includes
  - ACCEPT_SQL_PATCH
  - CREATE_DIAGNOSIS_TASK
  - DUMP_TRACE
  - EXECUTE_DIAGNOSIS_TASK
  - REPORT_DIAGNOSIS_TASK

DBMS_SQLDIAG_INTERNAL

- Introduced at version 12cR1
- Package includes
  - I_CREATE_PATCH

This package is undocumented ... but it allows something otherwise not possible:
Patching a SQL statement in the optimizer by inserting a hint.
Don't leave home without it.

```
BEGIN
  sys.dbms_sqldiag_internal.i_create_patch(
      sql_text  => 'SELECT COUNT(*), MAX(empno) FROM emp WHERE deptno =: deptno',
      hint_text => 'BIND_AWARE',
      name      => 'test_patch');
END;
/
```

DBMS_SQLPA

- Introduced at version 11gR1
- SQL Performance Analyzer
- Package includes
  - CREATE_ANALYSIS_TASK (new 12c)
  - SET_ANALYSIS_DEFAULT_PARAMETER

DBMS_SQLTUNE

- Introduced at version 10gR1
- SQL Tuning Advisor
- Package includes
  - CREATE_TUNING_TASK
  - REPORT_AUTO_TUNING_TASK
  - REPORT_SQL_DETAIL (new 12c)
  - REPORT_SQL_MONITOR
  - REPORT_SQL_MONITOR_LIST
  - SCHEDULE_TUNING_TASK (new 12c)
  - SELECT_SQL_TRACE

DBMS_WORKLOAD_CAPTURE & REPLAY

- Introduced at version 10gR1
- Rhetorical questions
  - Have you ever tried to triage an issue that could not be replicated on demand?
  - Why is it that so few organizations have licensed workload capture and replay?

DBMS_WORKLOAD_REPOSITORY (1:2)

- Introduced at version 10gR1
- 14 Undocumented init parameters
- Header source: `$ORACLE_HOME/rdbms/admin/dbmsawr.sql`
- Package includes
  - ADD_COLORED_SQL
  - ASH_GLOBAL_REPORT_HTML (new 12c)
  - ASH_REPORT_HTML
  - AWR_DIFF_REPORT_HTML
  - AWR_GLOBAL_DIFF_REPORT
  - CREATE_BASELINE
  - CREATE_BASELINE_TEMPLATE
  - CREATE_SNAPSHOT
  - MODIFY_SNAPSHOT_SETTINGS

DBMS_WORKLOAD_REPOSITORY (2:2)

- By default AWR takes snapshots once an hour and retains them for 7 days: This is inadequate for serious diagnostic work
- Set to15-20 minutes and 30 days

```
set linesize 121
col retention format a20
col snap_interval format a20

SELECT retention, snap_interval, topnsql
FROM wrm$_wr_control
WHERE dbid = (SELECT dbid FROM v$database);

SELECT dbid
FROM v$database;

exec
dbms_workload_repository.modify_snapshot_settings((30*24*60),
20, 1000, 1701481905);

SELECT retention, snap_interval, topnsql
FROM wrm$_wr_control
WHERE dbid = (SELECT dbid FROM v$database);
```

DBMS_XPLAN

- Everyone knows how to do this

```
EXPLAIN PLAN FOR
SELECT dummy FROM dual;

SELECT * FROM TABLE(dbms_xplan.display);
```

- But are you making use of the rest of DBMS_XPLAN's functionality?
  - DBMS_XPLAN.DIFF_PLAN_CURSOR
  - DBMS_XPLAN.DISPLAY_AWR
  - DBMS_XPLAN.DISPLAY_CURSOR
  - DBMS_XPLAN.DISPLAY_SQL_PATCH_PLAN (new 12c)
  - DBMS_XPLAN.DISPLAY_SQL_PLAN_BASELINE
  - DBMS_XPLAN.DISPLAY_SQL_PROFILE_PLAN (new 12c)

SQLTXPLAIN

- SQLTXPLAIN, also known as SQLT for short, is a tool I developed by Oracle ACE Carlos Sierra
- A tool developed in PL/SQL that inputs one SQL statement and output a zip file that contains many pieces
- Some of the output is for diagnostics, and some for further use if a Test Case (TC) is created
- To get it to work, you need to download it first from My Oracle Support (MOS). You can find it under Doc ID 215187.1
- Reference: `http://carlos-sierra.net/2012/04/03/what-is-sqltxplain/`

Trace and TKPROF

- Oracle's built in trace facility can "trace' a large number of events and dump the output to a file

Event	Description
10013	Monitor Transaction Recovery
10032	Dump Sort Statistics
10045	Freelist Management
10046	Dump SQL Trace
10053	Dump Optimizer Statistics and Calculations
10079	Dump SQL*Net Statistics
10081	Dump Highwater Mark Changes
10084	Dump Hash Join Statistics
10128	Dump Partition Pruning and Partition-wise Join Calculations
10200	Dump Consistent Reads
10233	Trace Memory Usage Statistics
10390 & 10391	Dump Parallel Execution Statistics
10731	Trace Cursor Statements

Trace and TKPROF

- SQL Trace, in common usage, refers to only the 10046 event and its 6 levels

Event #	Level	Description
10046	0	No statistics generated
10046	1	Standard trace output including parsing, executes and fetches, etc.
10046	2	Same as Level 1
10046	4	Level 1 + Bind Variables
10046	8	Level 1 + Waits
10046	12	Level 1 + Bind Variables & Waits

- The second most frequently traced event is 10053 (see previous slide)

Trace and TKPROF

- The 10046 trace provides performance information on individual SQL statements including the following statistics
  - Parse, execute, and fetch counts
  - CPU and elapsed times
  - Physical reads and logical reads
  - Number of rows processed
  - Misses on the Library Cache
  - Username under which each parse occurred
  - Each commit and rollback
  - Wait event data for each SQL statement and a summary for each trace file

Trace and TKPROF

- This is an example of generating a Level 12 SQL Trace

```
-- set a trace identifier to make the file easy to find
SQL> ALTER SESSION SET tracefile_identifier = 'test_plan1';

-- start the event trace
SQL> ALTER SESSION SET EVENTS '10046 trace name context
     forever, level 12';

-- run the SQL statement to trace
SQL> SELECT srvr_id
     FROM servers
     WHERE srvr_id IN (
       SELECT srvr_id
       FROM serv_inst);

-- stop tracing
SQL> ALTER SESSION SET EVENTS '10046 trace name context OFF';
-- or
SQL> ALTER SESSION SET SQL_TRACE=FALSE;
```

The file generated by the above will contain the string "test_plan1" in its name

Trace and TKPROF

- TKPROF is a tool that can be used to turn the output generated by a 10046 trace into a human readable form
- By default the trace file will be located at `$ORACLE_BASE/diag/rdbms/<db_name>/<instance_name>/trace/`

The trace file's header

```
Trace file C:\APP\ORACLE\diag\rdbms\orabase2\orabase2\trace\orabase2_ora_13404_test_plan1.trc
Oracle Database 12c Enterprise Edition Release 12.1.0.1.0 - 64bit Production
With the Partitioning, OLAP, Advanced Analytics and Real Application Testing options
Windows NT Version V6.1 Service Pack 1
CPU                 : 4 - type 8664, 2 Physical Cores
Process Affinity    : 0x0x0000000000000000
Memory (Avail/Total): Ph:3102M/8075M, Ph+PgF:11097M/16148M
Instance name: orabase2
Redo thread mounted by this instance: 1
Oracle process number: 47
Windows thread id: 13404, image: ORACLE.EXE (SHAD)


*** 2014-04-18 21:59:16.130
*** SESSION ID:(366.38553) 2014-04-18 21:59:16.130
*** CLIENT ID:() 2014-04-18 21:59:16.130
*** SERVICE NAME:(pdbdev) 2014-04-18 21:59:16.130
*** MODULE NAME:(SQL*Plus) 2014-04-18 21:59:16.130
*** ACTION NAME:() 2014-04-18 21:59:16.130
*** CONTAINER ID:(4) 2014-04-18 21:59:16.130
```

Trace and TKPROF

The trace file's body (partial)

```
PARSING IN CURSOR #404677528 len=865 dep=3 uid=0 oct=3 lid=0 tim=364198033857 hv=715221786 ad='7ff89882f10' sqlid='616m6uhpa2usu'
select
i.obj#,i.ts#,i.file#,i.block#,i.intcols,i.type#,i.flags,i.property,i.pctfree$,i.initrans,i.maxtrans,i.blevel,i.leafcnt,i.distkey,i.lblkkey,
i.dblkkey,i.clufac,i.cols,i.analyzetime,i.samplesize,i.dataobj#,nvl(i.degree,1),nvl(i.instances,1),i.rowcnt,mod(i.pctthres$,256),i.indmetho
d#,i.trunccnt,nvl(c.unicols,0),nvl(c.deferrable#+c.valid#,0),nvl(i.spare1,i.intcols),i.spare4,i.spare2,i.spare6,decode(i.pctthres$,null,nul
l,mod(trunc(i.pctthres$/256),256)),nvl(i.evaledition#,1),nvl(i.unusablebefore#,0),nvl(i.unusablebeginning#,0),
ist.cachedblk,ist.cachehit,ist.logicalread from ind$ i, ind_stats$ ist, (select enabled, min(cols) unicols,min(to_number(bitand(defer,1)))
deferrable#,min(to_number(bitand(defer,4))) valid# from cdef$ where obj#=:1 and enabled > 1 group by enabled) c where i.obj#=c.enabled(+)
and i.obj# = ist.obj#(+) and i.bo#=:1 order by i.obj#
END OF STMT
PARSE #404677528:c=0,e=873,p=0,cr=0,cu=0,mis=1,r=0,dep=3,og=4,plh=0,tim=364198033857
BINDS #404677528:
 Bind#0
  oacdty=02 mxl=22(22) mxlc=00 mal=00 scl=00 pre=00
  oacflg=00 fl2=1000001 frm=00 csi=00 siz=24 off=0
  kxsbbbfp=152affe8  bln=22  avl=03  flg=05
  value=443
 Bind#1
  No oacdef for this bind.
EXEC #404677528:c=0,e=3809,p=0,cr=0,cu=0,mis=1,r=0,dep=3,og=4,plh=2556297882,tim=364198037773
FETCH #404677528:c=0,e=121,p=0,cr=3,cu=0,mis=0,r=0,dep=3,og=4,plh=2556297882,tim=364198037913
STAT #404677528 id=1 cnt=0 pid=0 pos=1 obj=0 op='SORT ORDER BY (cr=3 pr=0 pw=0 time=133 us cost=4 size=374 card=2)'
STAT #404677528 id=2 cnt=0 pid=1 pos=1 obj=0 op='HASH JOIN OUTER (cr=3 pr=0 pw=0 time=122 us cost=4 size=374 card=2)'
STAT #404677528 id=3 cnt=0 pid=2 pos=1 obj=0 op='NESTED LOOPS OUTER (cr=3 pr=0 pw=0 time=28 us cost=2 size=288 card=2)'
STAT #404677528 id=4 cnt=0 pid=3 pos=1 obj=19 op='TABLE ACCESS CLUSTER IND$ (cr=3 pr=0 pw=0 time=27 us cost=2 size=184 card=2)'
STAT #404677528 id=5 cnt=1 pid=4 pos=1 obj=3 op='INDEX UNIQUE SCAN I_OBJ# (cr=2 pr=0 pw=0 time=16 us cost=1 size=0 card=1)'
STAT #404677528 id=6 cnt=0 pid=3 pos=2 obj=75 op='TABLE ACCESS BY INDEX ROWID IND_STATS$ (cr=0 pr=0 pw=0 time=0 us cost=0 size=52 card=1)'
STAT #404677528 id=7 cnt=0 pid=6 pos=1 obj=76 op='INDEX UNIQUE SCAN I_IND_STATS$_OBJ# (cr=0 pr=0 pw=0 time=0 us cost=0 size=0 card=1)'
STAT #404677528 id=8 cnt=0 pid=2 pos=2 obj=0 op='VIEW  (cr=0 pr=0 pw=0 time=0 us cost=2 size=43 card=1)'
STAT #404677528 id=9 cnt=0 pid=8 pos=1 obj=0 op='SORT GROUP BY (cr=0 pr=0 pw=0 time=0 us cost=2 size=15 card=1)'
STAT #404677528 id=10 cnt=0 pid=9 pos=1 obj=31 op='TABLE ACCESS BY INDEX ROWID BATCHED CDEF$ (cr=0 pr=0 pw=0 time=0 us cost=2 size=15
card=1)'
STAT #404677528 id=11 cnt=0 pid=10 pos=1 obj=54 op='INDEX RANGE SCAN I_CDEF2 (cr=0 pr=0 pw=0 time=0 us cost=1 size=0 card=4)'
CLOSE #404677528:c=0,e=3,dep=3,type=3,tim=364198038701
```

Trace and TKPROF

- The syntax for TKPROF and the TKPROF commands used by the instructor on the test_plan1 trace file

```
TKPROF <trace_file_name> <output_file_name>
 [explain=<user/password> [table=<schema.table_name>]]
 [print=<integer>]    -- list only the first nth SQL statements
 [aggregate=<yes|no>
 [insert=<file_name>] -- list SQL statements and data inside INSERT statements
 [sys=<no>]           -- TKPROF does not list SQL statements run as user SYS
 [record=<file_name>] -- record non-recursive statements in the trace file
 [waits=<yes|no>]     -- record summary of for wait events in the trace file
 [sort=<sort_options_list>] -- zero or more of the listed sort options
```

```
C:\>cd stage

C:\stage>tkprof test_plan1.trc test_plan1.txt

TKPROF: Release 12.1.0.1.0 - Development on Fri Apr 18 22:14:16 2014

Copyright (c) 1982, 2013, Oracle and/or its affiliates.  All rights reserved.
```

Trace and TKPROF

- The trace file after processing with TKPROF

```
TKPROF: Release 12.1.0.1.0 - Development on Fri Apr 18 22:14:16 2014

Copyright (c) 1982, 2013, Oracle and/or its affiliates.  All rights reserved.

Trace file: test_plan1.trc
Sort options: default

********************************************************************************
count    = number of times OCI procedure was executed
cpu      = cpu time in seconds executing
elapsed  = elapsed time in seconds executing
disk     = number of physical reads of buffers from disk
query    = number of buffers gotten for consistent read
current  = number of buffers gotten in current mode (usually for update)
rows     = number of rows processed by the fetch or execute call
********************************************************************************

SQL ID: 9babjv8yq8ru3 Plan Hash: 0
```

Trace and TKPROF

- The trace file after processing with TKPROF: Partial listing

```
Misses in library cache during parse: 0
Optimizer mode: FIRST_ROWS
Parsing user id: 103

Elapsed times include waiting on following events:
  Event waited on                             Times   Max. Wait  Total Waited
  ----------------------------------------   Waited  ----------  ------------
  SQL*Net message to client                       3        0.00          0.00
  SQL*Net message from client                     3       43.78         58.09
********************************************************************************

SQL ID: 2kyc2y1tvfgj5 Plan Hash: 3598708576

select count(*) from user_editioning$ where user#=:1 and type#=:2

call     count       cpu    elapsed       disk      query    current        rows
------- ------  -------- ---------- ---------- ---------- ----------  ----------
Parse        4      0.00       0.00          0          0          0           0
Execute      4      0.00       0.00          0          0          0           0
Fetch        4      0.00       0.00          0         12          0           4
------- ------  -------- ---------- ---------- ---------- ----------  ----------
total       12      0.00       0.00          0         12          0           4
```

Trace and TKPROF

- The trace file after processing with TKPROF: Partial listing

```
Misses in library cache during parse: 19
Misses in library cache during execute: 21

    3  user  SQL statements in session.
   24  internal SQL statements in session.
   27  SQL statements in session.
********************************************************************************
Trace file: test_plan1.trc
Trace file compatibility: 11.1.0.7
Sort options: default

       1  session in tracefile.
       3  user  SQL statements in trace file.
      24  internal SQL statements in trace file.
      27  SQL statements in trace file.
      21  unique SQL statements in trace file.
    1832  lines in trace file.
      58  elapsed seconds in trace file.
```

Trace Analyzer

- Oracle provides an alternative to TKPROF: Trace Analyzer (see My Oracle Support Note 224270.1)
- Implemented with several scripts, called by TRCANLYZR.SQL
  - Creates custom repository tables
- Provides comprehensive reporting with improved information for:
  - Bind Variables
  - Wait Events
- Can be implemented for versions 9i and above

trcsess Utility (1:2)

- Useful for consolidating the tracing of a particular session for performance or debugging purposes
- The trcsess utility allows related trace files to be consolidated into one file based on:
  - Session ID
  - Client ID
  - Service name
  - Action name
  - Module name
- The resulting output can be formatted with TKPROF

trcsess Utility (2:2)

- Syntax

```
C:\app\oracle\product\12.1.0\dbhome_2\BIN>trcsess -h
oracle.ss.tools.trcsess.SessTrcException: SessTrc-00002: Session Trace Usage error: Wrong
parameters passed.
trcsess [output=<output file name >] [session=<session ID>] [clientid=<clientid>]
[service=<service name>] [action=<action name>] [module=<module name>] <trace file names>

output=<output file name> output destination default being standard output.
session=<session Id> session to be traced.
Session id is a combination of session Index & session serial number e.g. 8.13.
clientid=<clientid> clientid to be traced.
service=<service name> service to be traced.
action=<action name> action to be traced.
module=<module name> module to be traced.
<trace_file_names> Space separated list of trace files with wild card '*' supported.
```

- Example

```
trcsess session = 21.2371 main_12359.trc main_12995.trc
```

ASH Reports

- Current information can be found in V$ACTIVE_SESSION_HISTORY
- Historical information can be found in WRH$_ACTIVE_SESSION_HISTORY

AWR Reports

- I consider AWR Reports nearly useless
  - Sadly just about everyone can say AWR Report and hardly anyone realizes that what they are talking about is the least valuable part of what is available via the AWR
- Data is not actionable information
- AWR is not one thing ... it is multiple components
  - SYSAUX Tablespace
  - 149 WRH$ Tables (12c)
  - 100 WRI$ Tables (12c)
  - Built in packages
  - A V$ Dynamic Performance View
  - SQL Scripts

AWR Tables

```
SQL> SELECT name FROM dba_dependencies
  2  WHERE referenced_name LIKE 'WRI%'
  3  AND name NOT LIKE 'WRI%'
  4  UNION
  5  SELECT referenced_name from dba_depdendencies
  6  WHERE name LIKE 'WRI%'
  7  AND referenced_name NOT LIIKE 'WRI%';

NAME
------------------------------
...
DBA_ADDM_FDG_BREAKDOWN
DBA_ADDM_FINDINGS
DBA_ADDM_INSTANCES
DBA_ADDM_TASKS
DBA_ADVISOR_ACTIONS
DBA_ADVISOR_DEFINITIONS
DBA_ADVISOR_DEF_PARAMETERS
DBA_ADVISOR_DIR_DEFINITIONS
DBA_ADVISOR_DIR_INSTANCES
DBA_ADVISOR_DIR_TASK_INST
DBA_ADVISOR_EXECUTIONS
DBA_ADVISOR_EXECUTION_TYPES
DBA_ADVISOR_EXEC_PARAMETERS
DBA_ADVISOR_FDG_BREAKDOWN
DBA_ADVISOR_FINDINGS
DBA_ADVISOR_FINDING_NAMES
DBA_ADVISOR_JOURNAL
DBA_ADVISOR_LOG
DBA_ADVISOR_OBJECTS
DBA_ADVISOR_OBJECT_TYPES
...
```

AWR Built-in Packages

- DBMS_ASH_INTERNAL
- DBMS_AWR_REPORT_LAYOUT
- DBMS_MANAGEMENT_PACKS
- DBMS_SWRF_INTERNAL
- DBMS_WORKLOAD_REPOSITORY

AWR SQL Scripts (1:2)

- Located at `$ORACLE_HOME in the directory /rdbms/admin`

File Name	Description
awrddrpi.sql	Report on differences between values recorded in two pairs of snapshots. This script requests the user for the dbid and instance number of the instance to report on, for each snapshot pair, before producing the standard report.
awrddrpt.sql	Defaults the dbid and instance number to that of the current instance connected-to, then calls awrddrpi.sql to produce the Compare Periods report.
awrextr.sql	SQL/Plus script to help users extract data from the AWR.
awrgdrpi.sql	RAC Version of Compare Period Report.
awrgdrpt.sql	This script defaults the dbid to that of the current instance connected-to, defaults instance list to all available instances and then calls awrgdrpi.sql to produce the Workload Repository RAC Compare Periods report.
awrgrpt.sql	This script defaults the dbid to that of the current instance connected-to, then calls awrgrpti.sql to produce the Workload Repository RAC report.
awrgrti.sql	SQL*Plus command file to report on RAC-wide differences between values recorded in two snapshots. This script requests the user for the dbid before producing the standard Workload Repository report.

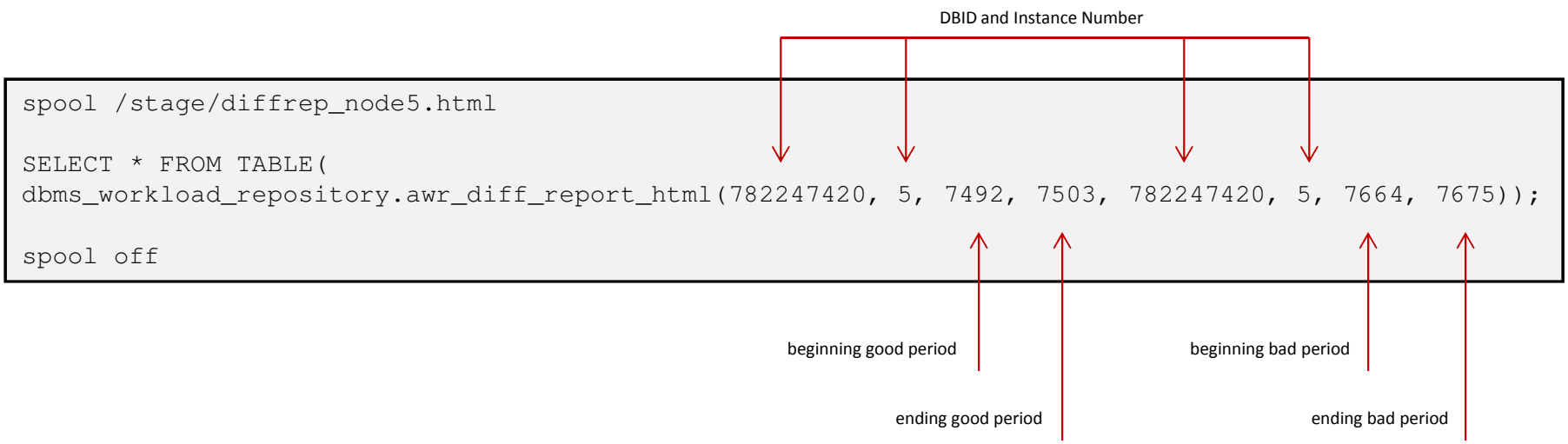
AWR SQL Scripts (2:2)

- Located at `$ORACLE_HOME in the directory /rdbms/admin`

File Name	Description
awrinfo.sql	Outputs general AWR information such as the size, data distribution, etc. in AWR and SYSAUX. The intended use of this script is for diagnosing abnormalities in AWR and not for diagnosing issues in the database instance.
awrrpt.sql	defaults the dbid and instance number to that of the current instance connected-to, then calls awrrpti.sql to produce the report.
awrrpti.sql	SQL*Plus command file to report on differences between values recorded in two snapshots. This script requests the user for the dbid and instance number of the instance to report on, before producing the standard report.
awrsqrpi.sql	SQL*Plus command file to report on differences between values recorded in two snapshots. This script requests the user for the dbid, instance number and the sql id, before producing a report for a particular sql statement in this instance.
awrsqrpt.sql	Defaults the dbid and instance number to that of the current instance connected-to then calls awrsqrpi.sql to produce a Workload report for a particular sql statement.
awrupd12.sql	This script updates AWR data to version 12c. It only modifies AWR data that has been imported using awrload.sql, or data from before changing the database DBID. In other words, it doesn't modify AWR data for the local, active DBID.

AWR Difference Report (1:2)

- Create AWR_DIFF reports and compare known good periods with known bad periods
  - The value is in the not in the absolute values but in the differences



AWR Difference Report (2:2)

- Is the problem redo and logical reads?

Key Instance Activity Stats

- Ordered by statistic name

Statistic	[redacted]											
db block changes	314,288,490	460,940,939	46.66	298.42	1,548.19	418.80	21, 832.10	31,929.66	46.25	177.45	254.18	43.24
execute count	149,150,104	114,455,440	-23.26	141.62	384.43	171.45	10,360.73	7,928.40	-23.48	84.21	63.12	-25.04
logons cumulative	9,118	10,197	11.83	0.01	0 .03	200.00	0.63	0.71	12.70	0.01	0.01	0.00
opened cursors cumulative	29,882,311	27,733,6 74	-7.19	28.37	93.15	228.34	2,075.78	1,921.13	-7.45	16.87	15.29	-9.37
parse count (total)	22,343,195	19,821,697	21.22	66.58	213.76	1,552 .07	1,373.06	-11.53	12.62	10.93	-13.39	
parse time elapsed	1,945,530	568,052	-70.80	1.85	1.91	3.24	135.15	39.35	-70.88	1.10	0.31	-71 .82
physical reads	127,937,187	134,482,086	5.12	121.48	451.69	271.82	8,887.18	9,315.66	4.82	72.23	74.16	2.67
physical writes	15,992,647	17,706,289	10.72	15.19	59.47	291.51	1, 110.93	1,226.53	10.41	9.03	9.76	8.08
redo size	68,097,611,028	80,741,189,252	18.57	64,660.41	271,190.58	319.41	4,730,411.02	5,592,991.34	18.23	38,448.68	44,523.99	15.80
session cursor cache hits	6,451,009	6,028,775	-6.55	6.13	20.25	230.34	448.12	417.62	-6.81	3.32	-8.79	
session logical reads	3,888,009,074	4,386,220,82 2	12.81	3,691.76	14,732.28	299.06	270,081.15	303,836.19	12.50	2,195.21	2,418.74	10.18
user calls	284,504,533	242,183,866	-14.88	270.14	813.44	201.12	19 ,763.15	16,776.22	-15.11	133.55	-16.86	
user commits	1,652,967	1,531,479	-7.35	1.57	5.14	227.39	114.82	106.09	-7.60	0.93	0.84	-9.68
user rollbacks	118,163	281,952	138.61	0.11	0.95	763.64	8.21	137.88	0.07	0.16	128.57	
workarea executions - onepass	422	493	16.82	0.00	0.00	0.00	0.03	0.03	0.00	0.00	0.00	0.00
workarea executions - optimal	19,070,549	14,1 36,361	-25.87	18.11	47.48	162.18	1,324.74	979.23	-26.08	10.77	7.80	-27.58

AWR Difference Report (2:2)

- The biggest change is user rollback

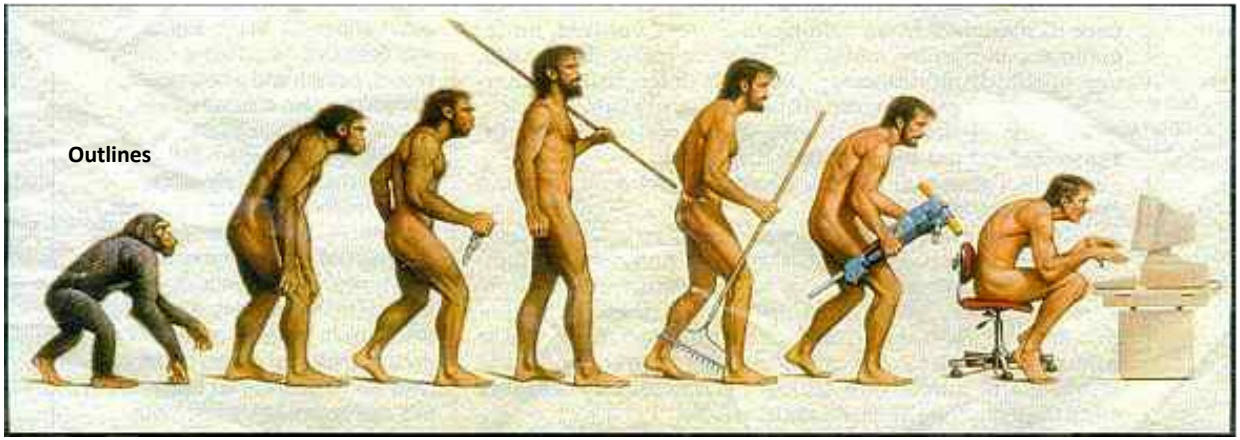
Key Instance Activity Stats

- Ordered by statistic name

	Value			per Second (DB time)			per Second (Elapsed Time)			per Trans		
Statistic	1st	2nd	%Diff	1st	2nd	%Diff	1st	2nd	%Diff	1st	2nd	%Diff
db block changes	314,288,490	460,940,939	46.66	298.42	1,548.19	418.80	21, 832.10	31,929.66	46.25	177.45	254.18	43.24
execute count	149,150,104	114,455,440	-23.26	141.62	384.43	171.45	10,360.73	7,928.40	-23.48	84.21	63.12	-25.04
logons cumulative	9,118	10,197	11.83	0.01	0 .03	200.00	0.63	0.71	12.70	0.01	0.01	0.00
opened cursors cumulative	29,882,311	27,733,6 74	-7.19	28.37	93.15	228.34	2,075.78	1,921.13	-7.45	16.87	15.29	-9.37
parse count (total)	22,343,195	19,821,697	21.22	66.58	213.76	,552 .07	1,373.06	-11.53	12.62	10.93	-13.39	
parse time elapsed	1,945,530	568,052	-70.80	1.85	1.91	3.24	135.15	39.35	-70.88	1.10	0.31	-71 .82
physical reads	127,937,187	134,482,086	5.12	121.48	451.69	271.82	8,887.18	9,315.66	4.82	72.23	74.16	2.67
physical writes	15,992,647	17,706,289	10.72	15.19	59.47	291.51	1, 110.93	1,226.53	10.41	9.03	9.76	8.08
redo size	68,097,611,028	80,741,189,252	18.57	64,660.41	271,190.58	319.41	4,730,411.02	5,592,991.34	18.23	38,448.68	44,523.99	15.80
session cursor cache hits	6,451,009	6,028,775	-6.55	6.13	20.25	230.34	448.12	417.62	-6.81	3.32	-8.79	
session logical reads	3,888,009,074	4,386,220,82 2	12.81	3,691.76	14,732.28	299.06	270,081.15	303,836.19	12.50	2,195.21	2,418.74	10.18
user calls	284,504,533	242,183,866	-14.88	270.14	813.44	201.12	19 ,763.15	16,776.22	-15.11	133.55	-16.86	
user commits	1,652,967	1,531,479	-7.35	1.57	5.14	227.39	114.82	106.09	-7.60	0.93	0.84	-9.68
user rollbacks	118,163	281,952	138.61	0.11	0.95	763.64	8.21	137.88	0.07	0.16	128.57	
workarea executions - onepass	422	493	16.82	0.00	0.00	0.00	0.03	0.03	0.00	0.00	0.00	0.00
workarea executions - optimal	19,070,549	14,1 36,361	-25.87	18.11	47.48	162.18	1,324.74	979.23	-26.08	10.77	7.80	-27.58

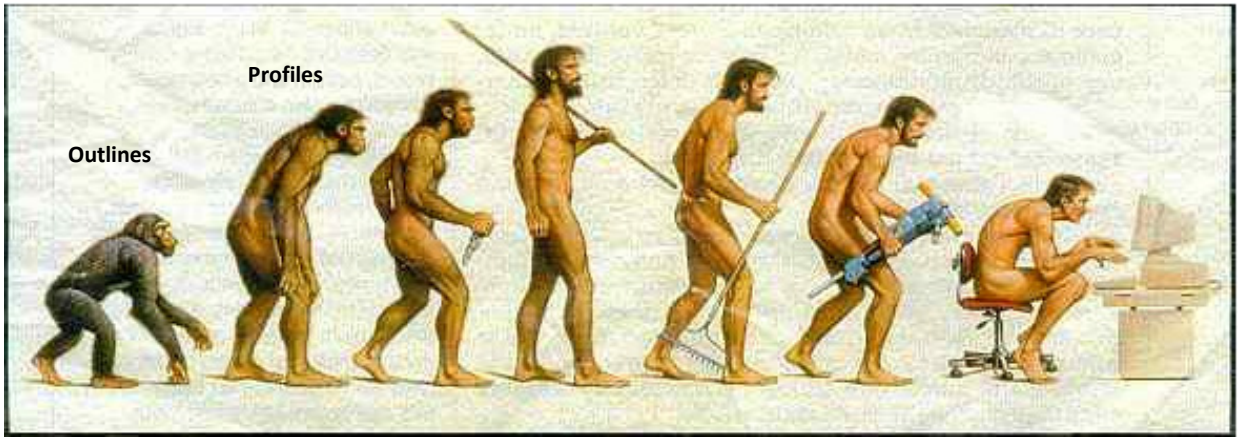
SQL Baselines (1:7)

- First came outlines
  - Built-in packages OUTLN_EDIT_PKG and OUTLN_PKG
  - An outline, more properly a Stored Outline, contains all of the hints required to force a specific SQL plan
  - Outlines are deprecated as of version 12.1 and while still functional will not be improved or maintained as new features are added
    - Outlines, for example, do not understand a multitenant Container database with CDB and PDBs



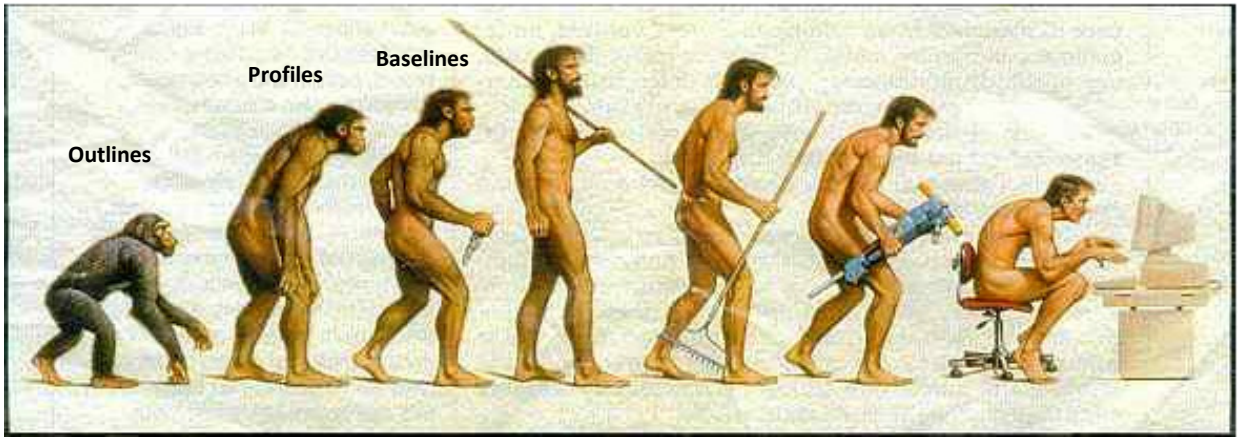
SQL Baselines (2:7)

- Then, in 10gR1, came profiles
  - Built-in package DBMS_SQLTUNE
    - Enhanced statistics for specific SQL statements



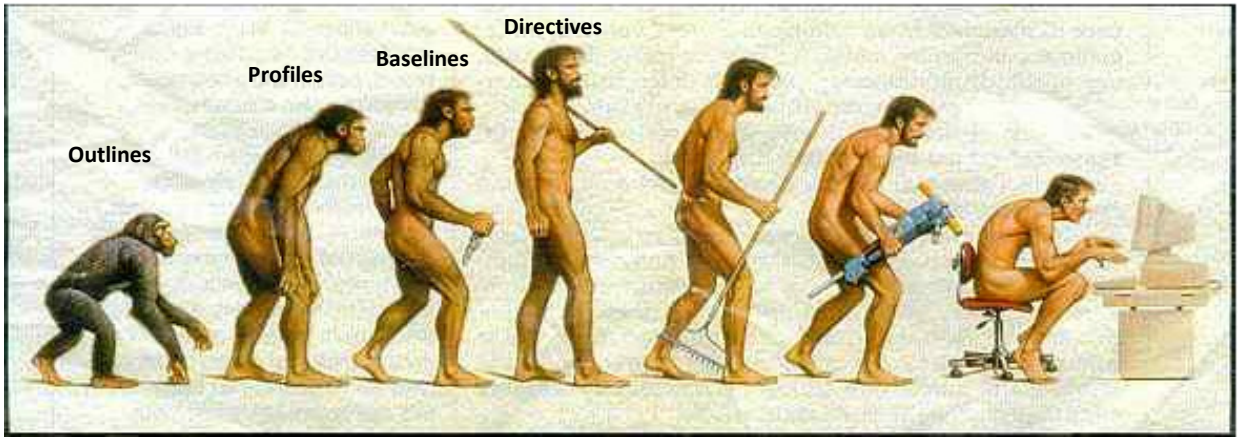
SQL Baselines (3:7)

- Next, in 11gR1, came Baselines and SQL Plan Management
  - Built-in package DBMS_SPM
    - Affect individual SQL statements
    - Adjusts CBO cardinality estimates
    - Motto: Only execute the best plan



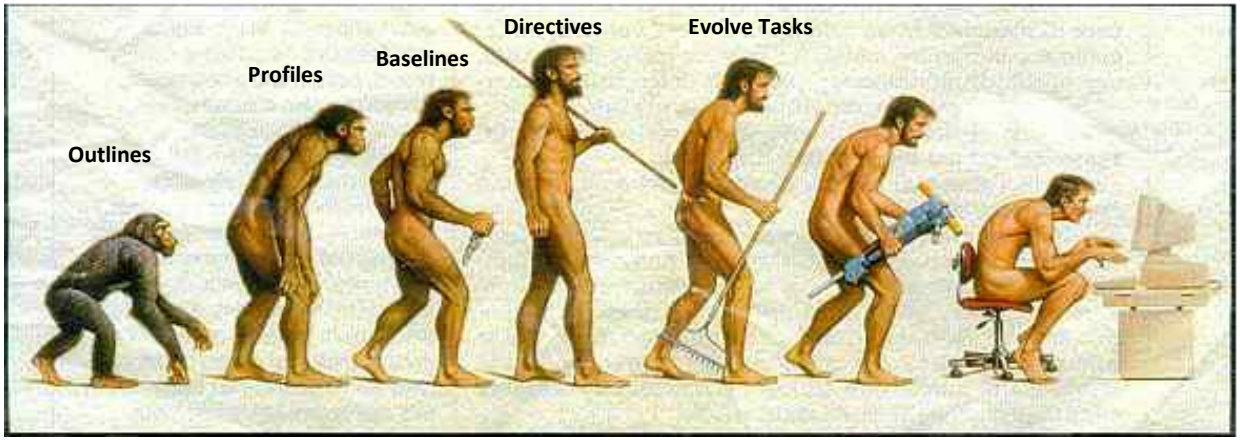
SQL Baselines (4:7)

- Now, in 12cR1, we have Directives
  - Built-in package DBMS_SPD
  - SPDs are objects generated automatically
  - Motto: Not satisfied until it arrives at the best plan



SQL Baselines (5:7)

- Migrating from there to here
  - DBMS_WORKLOAD_REPOSITORY
    - CREATE_BASELINE
  - DBMS_SPM
    - MIGRATE_STORED_OUTLINE
    - EVOLVE_SQL_PLAN_BASELINE
  - DBMS_SPD
    - ALTER_SQL_PLAN_DIRECTIVE



SQL Baselines (6:7)

- SPM Evolve Advisor
  - A SQL advisor that evolves plans that have recently been added to the SQL plan baseline
  - The advisor simplifies plan evolution by eliminating the requirement to do it manually
- SYS_AUTO_SPM_EVOLVE_TASK
  - Runs daily in a scheduled maintenance window
  - Ranks all unaccepted plans, and then performs test executions of as many plans as possible during the window
  - Selects the lowest-cost plan to compare against each unaccepted plan.
  - If a plan performs sufficiently better than the existing accepted plan the database automatically accepts it

SQL Baselines (7:7)

```
SQL> SELECT view_name
  2  FROM dba_views
  3  WHERE view_name
           LIKE '%DIRECTIV%'
  4  ORDER BY 1;


VIEW_NAME
----------------------------
CDB_ADDM_SYSTEM_DIRECTIVES
CDB_ADDM_TASK_DIRECTIVES
CDB_CDB_RSRC_PLAN_DIRECTIVES
CDB_RSRC_PLAN_DIRECTIVES
CDB_SQL_PLAN_DIRECTIVES
DBA_ADDM_SYSTEM_DIRECTIVES
DBA_ADDM_TASK_DIRECTIVES
DBA_CDB_RSRC_PLAN_DIRECTIVES
DBA_RSRC_PLAN_DIRECTIVES
DBA_SQL_PLAN_DIRECTIVES
USER_ADDM_TASK_DIRECTIVES
_BASE_OPT_DIRECTIVE
```

```
SELECT DISTINCT sql_id
FROM v$sql;

SELECT * FROM TABLE(dbms_xplan.display_cursor('gd90ygn1j4026'));

SELECT dbms_sqldiag_internal.i_generate_param_import('gd90ygn1j4026')
FROM dual;

DBMS_SQLDIAG_INTERNAL.I_GENERATE_PARAM_IMPORT('GD90YGN1J4026')
------------------------------------------------------------------
alter session set "optimizer_capture_sql_plan_baselines" = true;
alter session set "optimizer_index_caching" = 50;
alter session set "optimizer_index_cost_adj" = 50;
alter session set "optimizer_mode" = choose;

col directive_id format 99999999999999999999

SELECT directive_id, type, state, auto_drop, last_used
FROM dba_sql_plan_directives;

exec dbms_spd.alter_sql_plan_directive(16162032791226711612, 'AUTO_DROP',
YES');

SELECT owner, object_name, subobject_name, object_type
FROM dba_sql_plan_dir_objects
WHERE directive_id = 16162032791226711612;
```

Rhetorical Questions

- Do you expect your doctor to keep current?
- How current is your skill set?

Database Layer PL/SQL Tuning

PL/SQL Tuning Topics (1:2)

- Stop doing unnecessary work
  - Array Processing
  - Bind Variables
  - Compiler Mode Variations
  - Concatenations
  - Constants vs Variables
  - Data Types
  - Debug Mode Compilation
  - Deterministic Functions
  - Dynamic SQL
  - Implicit Casts
  - Invalid Objects
  - Native Compilation & SIMPLE_INTEGER

PL/SQL Tuning Topics (2:2)

- Stop doing unnecessary work (cont.)
  - Native Dynamic SQL
  - NOCOPY Compiler Hint
  - Pragma Inline
  - Ref Cursor Types
  - Result Cache
  - RETURNING Clause
  - Unnecessary SELECT COUNT(*)s
- PL/SQL Warnings
- DBMS_PROFILER
- DBMS_HPROF

Array Processing

- Can your fastest server perform 500K inserts per second? My laptop can
- Syntax elements you must learn and use
  - BULK COLLECT statement
  - FORALL statement
  - LIMITS clause
  - SAVE EXCEPTIONS clause

```
CREATE OR REPLACE PROCEDURE fast_way AUTHID CURRENT_USER IS
 TYPE myarray IS TABLE OF parent%ROWTYPE;
 l_data myarray;

 CURSOR r IS
 SELECT part_num, part_name
 FROM parent;

 BatchSize CONSTANT POSITIVE := 1000;
BEGIN
  OPEN r;
  LOOP
    FETCH r BULK COLLECT INTO l_data LIMIT BatchSize;

    FOR j IN 1 .. l_data.COUNT LOOP
      l_data(j).part_num := l_data(j).part_num * 10;
    END LOOP;

    FORALL i IN 1..l_data.COUNT
    INSERT INTO child VALUES l_data(i);

    EXIT WHEN l_data.COUNT < BatchSize;
  END LOOP;
  COMMIT;
  CLOSE r;
END fast_way;
/
```

Bind Variables (1:2)

- This is one way to find the places you should look

```
SQL> SELECT owner, COUNT(*)
  2  FROM dba_source
  3  WHERE owner NOT LIKE '%SYS%'
  4  AND UPPER(text) LIKE '%EXECUTE IMMEDIATE%'
  5  GROUP BY owner
  6  ORDER BY 1;

OWNER                            COUNT(*)
------------------------------ ----------
ORACLE_OCM                              8
OSSBACKEND                              9
PMCM                                   54
SRE                                    21
WEBWIZ                                 20
ZZYZX                                  39
```

Bind Variables (2:2)

- When using NDS bind variables do not concatenated literals or variables

```
DECLARE
 x NUMBER(10);
BEGIN
  FOR i IN 1 .. 50000
  LOOP
    EXECUTE IMMEDIATE 'SELECT ' || i || ' INTO :b1 FROM dual'
    INTO x;
  END LOOP;
END;
/
```

```
DECLARE
 x NUMBER(10);
BEGIN
  FOR i IN 1 .. 50000
  LOOP
    EXECUTE IMMEDIATE 'SELECT :b1 FROM dual'
    INTO x
    USING i;
  END LOOP;
END;
/
```

Use DBMS_ASSERT to thwart SQL Injection attacks

Non-Standard Optimizer Level Compilations

- The default optimizer mode for 11gR2 and 12cR1 is 2

```
SQL> SELECT owner, type, name, plsql_optimize_level
  2  FROM dba_plsql_object_settings
  3  WHERE plsql_optimize_level is not null
  4  AND plsql_optimize_level <> 2
  5  ORDER BY 1,2,3;

OWNER                TYPE         NAME                       PLSQL_OPTIMIZE_LEVEL
-------------------- ------------ -------------------------- --------------------
AIRCOM               TYPE         ALARM_VIEW_ALL_SC_TAB                         1
AIRCOM               TYPE         ALARM_VIEW_ALL_T                              1
GLOBAL               PROCEDURE    CUSTOM_GRANT                                  1
GLOBAL               PROCEDURE    REVOKE_GRANT                                  1
KG449W               PACKAGE      BULK_GATHER_V2                                1
KG449W               PACKAGE      BULK_GATHER_V2_STAGE                          1
KG449W               PACKAGE      CELLTRAFFICDATA                               1
KG449W               PACKAGE      LTE_BULK_ETL                                  1
KG449W               PROCEDURE    VENDORS                                       1
SYS                  TYPE         SYS_YOID0000009056$                           1
```

Concatenation (1:2)

```
FUNCTION delete_partitions_in_range(pschema    IN VARCHAR2,
                                    ptable     IN VARCHAR2,
                                    pretention IN INTEGER,
                                    parttype   IN INTEGER) RETURN PLS_INTEGER
IS
 ... variable declarations ...

BEGIN
  --add starting processing
  IF ptable <> 'COMMON_LOGS' THEN
    add_to_log_table(cloginformation,
                     827600,
                     upper(pschema) || '.' || upper(ptable) ||
                     ' - Processing (Delete partitions in range)');
  END IF;
  ... code here ...
  LOOP
    BEGIN
      ... code here ...
    EXCEPTION
      WHEN OTHERS THEN
        add_to_log_table(clogmajor,
                         827604,
                         upper(pschema) || '.' || upper(ptable) ||
                         ' - Unable to Delete Partition ' ||
                         partTable(i).Partition_Name || ':' ||
                         SQLERRM);
        RAISE;
    END;
  END LOOP;

  --Add log entry to summarise the partitions added
  IF ((vSummaryMinDate < vmindate) AND (ptable <> 'COMMON_LOGS')) THEN
    add_to_log_table(cloginformation,
                     827605,
                     upper(pschema) || '.' || upper(ptable) ||
                     ' - Deleted partitions from ' ||
                     to_char(vSummaryMinDate, 'YYYY-MM-DD HH24:MI:SS') ||
                     ' to ' || to_char(vmindate, 'YYYY-MM-DD HH24:MI:SS'));
  END IF;
  RETURN partTable.count;
END;
```

Concatenation (2:2)

```
FUNCTION delete_partitions_in_range(pschema    IN VARCHAR2,
                                    ptable     IN VARCHAR2,
                                    pretention IN INTEGER,
                                    parttype   IN INTEGER) RETURN PLS_INTEGER
IS
 ... variable declarations ...
 cObject VARCHAR2(61) := UPPER(pschema || '.' || ptable);
BEGIN
  --add starting processing
  IF ptable <> 'COMMON_LOGS' THEN
    add_to_log_table(cloginformation,
                     827600,
                     cObject ||
                     ' - Processing (Delete partitions in range)');
  END IF;
  ... code here ...
  LOOP
    BEGIN
      ... code here ...
    EXCEPTION
      WHEN OTHERS THEN
        add_to_log_table(clogmajor,
                         827604,
                         cObject ||
                         ' - Unable to Delete Partition ' ||
                         partTable(i).Partition_Name || ':' ||
                         SQLERRM);
        RAISE;
    END;
  END LOOP;

  --Add log entry to summarise the partitions added
  IF ((vSummaryMinDate < vmindate) AND (ptable <> 'COMMON_LOGS')) THEN
    add_to_log_table(cloginformation,
                     827605,
                     cObject ||
                     ' - Deleted partitions from ' ||
                     to_char(vSummaryMinDate, 'YYYY-MM-DD HH24:MI:SS') ||
                     ' to ' || to_char(vmindate, 'YYYY-MM-DD HH24:MI:SS'));
  END IF;
  RETURN partTable.count;
END;
```

Constants and Variables

- There are 10 things wrong in this code: Can you find them and fix them?

```
DECLARE
 counter NUMBER(3) := 2;
 pi      NUMBER(8,7) DEFAULT 3.1415926;
 test    NUMBER(10,8) NOT NULL := 10;
 today   VARCHAR2(30) := ' today is ' || SYSDATE;
BEGIN
  dbms_output.put_line('Counter: ' || counter);
  dbms_output.put_line('Pi:      ' || pi);
  dbms_output.put_line('Today:  ' || today);
END;
/
```

all three dbms_output statements force an implicit cast
all four assignments are constants not variables
test allocates excess memory
test serves no useful purpose
not null constraint requires revalidation with every assignment

Data Types

- The NUMBER data type requires 22 bytes be assigned for any value favoring portability and arbitrary scale and precision rather than performance
- To optimize for performance always use PLS_INTEGER or its subtypes for integer values
  - BINARY_INTEGER (PLS_INTEGER)
  - NATURAL
  - NATURALN
  - POSITIVE
  - POSITIVEN
  - SIGNTYPE

```
CREATE OR REPLACE FUNCTION myfunc RETURN signtype AUTHID DEFINER IS
BEGIN
  RETURN 1;
END;
/
```

Debug Mode Compilation

- Always compile in UAT and production using SQL*Plus
  - This is the result of using Toad, PLSQL Developer, etc.

```
SQL> SELECT owner, type, name
  2  FROM dba_plsql_object_settings
  3  WHERE plsql_debug = 'TRUE'
  4  ORDER BY 1,2,3;

OWNER                      TYPE         NAME
-------------------------- ------------ ------------------------------
E911_NATIONAL_REPORTING    TYPE         TYP_HANDSET_INTRADO
E911_NATIONAL_REPORTING    TYPE         TYP_HANDSET_TCS
E911_NATIONAL_REPORTING    TYPE BODY    TYP_HANDSET_INTRADO
E911_NATIONAL_REPORTING    TYPE BODY    TYP_HANDSET_TCS
JW430N                     PROCEDURE    CHECKERICMAPPING
JW430N                     PROCEDURE    CREATE_ERIC_MAPPING
KG449W                     PACKAGE      BULK_GATHER_V2
KG449W                     PACKAGE      BULK_GATHER_V2_STAGE
KG449W                     PACKAGE      CELLTRAFFICDATA
KG449W                     PACKAGE      LTE_BULK_ETL
KG449W                     PACKAGE      QUERY_CODE_PKG
KG449W                     PACKAGE BODY BULK_GATHER_V2
KG449W                     PACKAGE BODY BULK_GATHER_V2_STAGE
KG449W                     PACKAGE BODY CELLTRAFFICDATA
KG449W                     PACKAGE BODY LTE_BULK_ETL
KG449W                     PACKAGE BODY QUERY_CODE_PKG
KG449W                     PROCEDURE    CRC
KG449W                     PROCEDURE    GEN_CELLTRAFFICDATA
KG449W                     PROCEDURE    GEN_CONFIG_PARAM
KG449W                     PROCEDURE    NOR_DEL
KG449W                     PROCEDURE    VENDORS
M62085                     FUNCTION     ADDARR
```

Deterministic Functions

- Indicates that the function returns the same result value whenever it is called with the input value

```
CREATE OR REPLACE PACKAGE df_demo AUTHID DEFINER IS
 td DATE;

 FUNCTION get_date RETURN DATE;
 FUNCTION get_date_determ RETURN DATE DETERMINISTIC;
END df_demo;
/


CREATE OR REPLACE PACKAGE BODY df_demo IS
--=====================================================
 FUNCTION get_date RETURN DATE IS
 BEGIN
   RETURN df_demo.td;
 END get_date;
--=====================================================
 FUNCTION get_date_determ RETURN DATE DETERMINISTIC IS
 BEGIN
   RETURN df_demo.td;
 END get_date_determ;
--=====================================================
END df_demo;
/
```

Dynamic SQL

```
PROCEDURE gather_schema_stats(pschema IN VARCHAR2, pGlobalOption IN BOOLEAN := TRUE) IS
BEGIN
  IF pschema IS NOT NULL THEN
    DBMS_STATS.UNLOCK_SCHEMA_STATS (pschema);

    IF pGlobalOption = TRUE THEN
      EXECUTE IMMEDIATE 'BEGIN DBMS_STATS.GATHER_SCHEMA_STATS(OWNNAME => ''' ||
        pschema || ''',' ||
        '  ESTIMATE_PERCENT => 25, METHOD_OPT => ''FOR ALL INDEXED COLUMNS SIZE 254'',  DEGREE => 4,' ||
        '  CASCADE => TRUE, granularity => ''GLOBAL'', OPTIONS => ''GATHER STALE'', no_invalidate => TRUE); END;';
    ELSE
      EXECUTE IMMEDIATE 'BEGIN DBMS_STATS.GATHER_SCHEMA_STATS(OWNNAME => ''' ||
        pschema || ''',' ||
        '  ESTIMATE_PERCENT => 25,  METHOD_OPT => ''FOR ALL INDEXED COLUMNS SIZE 254'',  DEGREE => 4,' ||
        '  CASCADE => TRUE, OPTIONS => ''GATHER STALE'', no_invalidate => TRUE); END;';
    END IF;

    add_to_log_table(clogInformation, 827533, 'Schema stats gathered for schema ' || upper(pschema));
    -- Lock schema stats
    DBMS_STATS.LOCK_SCHEMA_STATS (pschema);
  END IF;
EXCEPTION
  WHEN OTHERS THEN
    --log error in COMMON_LOGS table
    add_to_log_table(clogmajor, 827516, 'Error in Gather_Schema_stats for schema ' ||
                     upper(pschema) || ' - message:' || SQLERRM);
END;
```

```
dbms_stats.gather_schema_stats(pschema, ESTIMATE_PERCENT => 25,
METHOD_OPT => 'FOR ALL INDEXED COLUMNS SIZE 254', DEGREE => 4,
CASCADE => TRUE, granularity => 'GLOBAL', OPTIONS => 'GATHER STALE', no_invalidate => TRUE);

 dbms_stats.gather_schema_stats(pschema, ESTIMATE_PERCENT => 25,
METHOD_OPT => 'FOR ALL INDEXED COLUMNS SIZE 254', DEGREE => gcDegree,
CASCADE => TRUE, granularity => 'GLOBAL', OPTIONS => 'GATHER STALE', no_invalidate => TRUE);
```

Implicit Casts

- Code that does not correctly define data types will either fail to run or run very inefficiently

The following example shows both the correct way and the incorrect way to work with dates. The correct way is to perform an explicit cast

```
orabase> create table t (
      2  datecol date);

Table created.

orabase> insert into t values ('01-JAN-2012');

1 row created.

orabase> insert into t values (TO_DATE('01-JAN-2012'));

1 row created.
```

In Oracle dates are dates ... not strings. Similarly numbers should either be explicitly cast with TO_NUMBER or only processed using numeric variables.

Invalid Objects

```
SQL> SELECT owner, object_type, object_name
  2  FROM dba_objects
  3  WHERE status = 'INVALID'
  4  ORDER BY 1,2,3;

OWNER                          OBJECT_TYPE         OBJECT_NAME
------------------------------ ------------------- -------------------
GLOBAL                         VIEW                CHANNEL_DIMVIEW
GLOBAL                         VIEW                CUSTOMER_DIMVIEW
GLOBAL                         VIEW                PRODUCT_DIMVIEW
GLOBAL                         VIEW                TIME_DIMVIEW
GLOBAL                         VIEW                UNITS_CUBE_CUBEVIEW
MLIB                           PROCEDURE           ALLCBLAST
MLIB                           PROCEDURE           ALLEBLAST
MLIB                           PROCEDURE           CBLAST
MLIB                           PROCEDURE           CLEAN_LIST
MLIB                           PROCEDURE           EBLAST
MLIB                           PROCEDURE           VANCBLAST
MLIB                           PROCEDURE           VANEBLAST
UWCLASS                        PACKAGE BODY        INDEX_UTIL
UWCLASS                        PACKAGE BODY        UTL_SPADV


SQL> alter procedure mlib.allcblast compile;
Warning: Procedure altered with compilation errors.

SQL> sho err
Errors for PROCEDURE MLIB.ALLCBLAST:

LINE/COL     ERROR
---------    -------------------------------------------------------
26/7         PL/SQL: Statement ignored
26/7         PLS-00201: identifier 'SYS.DBMS_LOCK' must be declared
```

Native Compilation (1:2)

- By default Oracle compiles PL/SQL into PCode
- Native compilation compiles to executable code
- Native compiled code can greatly improve recursive math
- Compiled code can not make disk I/O faster
- Compiled code can not be exported and imported between different operating systems and versions
- Use the SIMPLE_INTEGER data type when possible

```
ALTER SESSION SET plsql_compiler_flags = 'INTERPRETED';
or
ALTER SESSION SET plsql_compiler_flags = 'NATIVE';
```

Native Compilation (2:2)

```
CREATE OR REPLACE FUNCTION factorial_interpreted(p_n PLS_INTEGER)
RETURN NUMBER IS
BEGIN
  IF (p_n = 1) THEN
    RETURN 1;
  ELSE
    RETURN factorial_interpreted(p_n-1) * p_n;
  END IF;
END factorial_interpreted;
/
CREATE OR REPLACE FUNCTION factorial_native(p_n PLS_INTEGER)
RETURN NUMBER IS
BEGIN
  IF (p_n = 1) THEN
    RETURN 1;
  ELSE
    RETURN factorial_native(p_n-1) * p_n;
  END IF;
END factorial_native;
/
CREATE OR REPLACE FUNCTION factorial_simple(p_n SIMPLE_INTEGER)
RETURN NUMBER IS
BEGIN
  IF (p_n = 1) THEN
    RETURN 1;
  ELSE
    RETURN factorial_simple(p_n-1) * p_n;
  END IF;
END factorial_simple;
/

ALTER FUNCTION factorial_native
COMPILE PLSQL_CODE_TYPE=NATIVE REUSE SETTINGS;
ALTER FUNCTION factorial_simple
COMPILE PLSQL_CODE_TYPE=NATIVE REUSE SETTINGS;
```

NOCOPY Compiler Hint

- In parameters are passed by reference
- OUT parameters, by default, are pass by copying

```
CREATE OR REPLACE PROCEDURE nocopy_out (
  retval OUT NOCOPY VARCHAR2) AUTHID DEFINER AS
BEGIN
  retval := dbms_crypto.randombytes(32);
END nocopy_out;
/
```

Pragma INLINE

- Instructs the compiler to replaces a subprogram invocation with a copy of the invoked subprogram if in the same PL/SQL unit

```
DECLARE
 l_loops  NUMBER := 10000000;
 l_start  NUMBER;
 l_return NUMBER;

 FUNCTION add_numbers(p_1 IN NUMBER, p_2 IN NUMBER) RETURN NUMBER AS
 BEGIN
   RETURN p_1 + p_2;
 END add_numbers;
BEGIN
  l_start := dbms_utility.get_time;

  FOR i IN 1 .. l_loops LOOP
    PRAGMA INLINE(add_numbers, 'YES');
    l_return := add_numbers(1, i);
  END LOOP;

  dbms_output.put_line('Elapsed Time: ' || (dbms_utility.get_time - l_start) || ' hsecs');
END;
/
```

REF CURSORs

- There are two types of reference cursors
  - Strongly Typed
  - Weakly Typed (SYS_REFCURSOR)
- Strongly typed ref cursors are faster and require less latching

Result Cache (1:2)

- How many times do time zones change per hour?
- How many times does a software version change in an hour?
- Enough to justify 21,619,558 queries in 61 minutes?
- Enough to justify 72.75% of all queries running?

SQL ordered by Executions

- Total Executions: 29,717,627
- Captured SQL account for 77.4% of Total

Executions	Rows Processed	Rows per Exec	CPU per Exec (s)	Elap per Exec (s)	SQL Id	SQL Module	SQL Text
10,128,178	2,506,529	0.25	0.00	0.00	932srzg1krc33	ASN_07B_DIP(004110016)	SELECT NE_TIMEZONE FROM CMPM.E...
7,576,759	7,579,197	1.00	0.00	0.00	1h698sb62un99	asci_56_RANAPProtocolStats(01611000E)	SELECT DISTINCT NE_TIMEZONE FR...
3,914,621	3,848,268	0.98	0.00	0.00	5tbzddgguu8cc	asci_56_RANAPProtocolStats(01611000E)	SELECT SYS_VERSION FROM CMPM.T...
311,645	311,604	1.00	0.00	0.00	7gtztzv329wg0		select c.name, u.name from co...
301,428	301,325	1.00	0.00	0.00	36s446f9cnwhw		SELECT C.NAME FROM COL$ C WHER...
200,692	200,669	1.00	0.00	0.00	4vs91dcv7u1p6	OMS	insert into sys.aud$(sessioni...
65,044	65,035	1.00	0.00	0.00	fz9xwpt2cvt0k		SELECT par_type, param_clob, ...
64,949	3,945,482	60.75	0.00	0.00	f5ra7dru5fk5n	XML_P7R_RNC_RCS(003110008)	SELECT NAME, PATH, READ, WR...
64,801	64,807	1.00	0.00	0.00	fhzj09a7fnrnb	XML_V7I_IN_LP_DC(00811000V)	SELECT DBTIMEZONE, LENGTH(DBT...
64,632	64,542	1.00	0.00	0.00	15jnrrb6016nd	XML_V7I_IN_LP_DC(00811000V)	SELECT SESSIONTIMEZONE, LENGT...

Result Cache (2:2)

- Just as with the example earlier of using the Result Cache to improve SQL performance PL/SQL functions can also benefit from the Result Cache Hint

```
CREATE OR REPLACE FUNCTION rcache(p_srvr_id IN servers.srvr_id%TYPE) RETURN BOOLEAN
RESULT_CACHE RELIES_ON(servers) IS
  srvrow servers%ROWTYPE;
 BEGIN
   SELECT *
   INTO srvrow
   FROM servers
   WHERE srvr_id = p_srvr_id;
   RETURN TRUE;
 EXCEPTION
   WHEN OTHERS THEN
     RETURN FALSE;
 END rcache;
 /
```

Returning Clause

- It is inefficient to insert a row and then query to find information about the row that was created
- The returning clause allows you to capture information about the DML that is taking place
- The returning clauses exists for INSERT, UPDATE, and DELETE statements

```
DECLARE
 x emp.empno%TYPE;
 r rowid;
BEGIN
  INSERT INTO emp
  (empno, ename)
  VALUES
  (seq_emp.NEXTVAL, 'Morgan')
  RETURNING rowid, empno
  INTO r, x;

  dbms_output.put_line(r);
  dbms_output.put_line(x);
END;
/
```

Unnecessary Operations

```
PROCEDURE get_records(pCustID IN NUMBER) AUTHID DEFINER IS
 vCount PLS_INTEGER;
BEGIN
  SELECT COUNT(*)
  INTO vCount
  FROM customers
  WHERE cust_id = pCustID;

  IF vCount > 0 THEN
    process_customer_recs(pCustID);
  END IF;
END get_records;
/



PROCEDURE process_customer_recs(pCustID IN NUMBER) AUTHID DEFINER IS
 CURSOR cust_cur IS
 SELECT transno, trans_type, transdate
 FROM customers
 WHERE cust_id = pCustID;
BEGIN
  ... do stuff ...
END process_customer_recs;
/
```

PL/SQL Warnings (1:3)

- 5000 - 5999: Severe Warnings
- 6000 - 6999: Informational Warnings
- 7000 - 7999: Performance Warnings

```
col value format a80
SELECT value FROM v$parameter WHERE name = 'plsql_warnings';

ALTER SESSION SET PLSQL_WARNINGS='DISABLE:ALL';
ALTER SESSION SET PLSQL_WARNINGS='ENABLE:INFORMATIONAL';
ALTER SESSION SET PLSQL_WARNINGS='ENABLE:PERFORMANCE';
ALTER SESSION SET PLSQL_WARNINGS='ENABLE:SEVERE';
ALTER SESSION SET PLSQL_WARNINGS='ENABLE:ALL';

ALTER SESSION SET PLSQL_WARNINGS='ENABLE:SEVERE', 'DISABLE:INFORMATIONAL';

ALTER SESSION SET PLSQL_WARNINGS = 'DISABLE:5000', 'ENABLE:5001', 'ERROR:5002';

ALTER SESSION SET PLSQL_WARNINGS = 'ENABLE:(5004,6002,7206)', 'DISABLE:(5000,5003,7203)';
```

PL/SQL Warnings (2:3)

```
SQL> CREATE OR REPLACE PROCEDURE plw07203(sa IN OUT dbms_sql.varchar2a) IS
  2  BEGIN
  3    dbms_output.put_line('This is a test');
  4  END plw07203;
  5  /

SP2-0804: Procedure created with compilation warnings

SQL> sho err
Errors for PROCEDURE PLW07203:

LINE/COL ERROR
-------- -----------------------------------------------------------------------------------
1/1      PLW-05018: unit PLW07203 omitted optional AUTHID clause; default value DEFINER used
1/20     PLW-07203: parameter 'SA' may benefit from use of the NOCOPY compiler hint
```

PL/SQL Warnings (3:3)

```
SQL> CREATE OR REPLACE PROCEDURE plw07204 IS
  2   v VARCHAR2(10);
  3   i VARCHAR2(10);
  4  BEGIN
  5    SELECT srvr_id
  6    INTO v
  7    FROM servers
  8    WHERE rownum = 1;
  9
 10    SELECT COUNT(*)
 11    INTO I
 12    FROM servers
 13    WHERE srvr_id = v;
 14  END plw07204;
 15  /

SP2-0804: Procedure created with compilation warnings

SQL> sho err
Errors for PROCEDURE PLW07204:

LINE/COL ERROR
-------- -----------------------------------------------------------------------------------
1/1      PLW-05018: unit PLW07204 omitted optional AUTHID clause; default value DEFINER used
13/19    PLW-07204: conversion away from column type may result in sub-optimal query plan
```

DBMS_PROFILER

- Introduced at version 8.1.5
- Package includes
  - FLUSH_DATA
  - ROLLUP_RUN
  - ROLLUP_UNIT
  - START_PROFILER
  - STOP_PROFILER

DBMS_HPROF

- Introduced at version 11gR1
- Package contains
  - ANALYZE
  - START_PROFILING
  - STOP_PROFILING

PL/SQL Summary

- Most PL/SQL code I see in my consulting uses features from version 7.3.4
- Code that would compile in version 7.3.4 PL/SQL will not run efficiently
- Code that would compile in 8.0 or 8*i* is likely extremely inefficient
- High performance comes from learning about the optimizations and enhancements Oracle builds into their product version-after-version and using them
- If it is not necessary ... don't do it

Middleware Layer

Middleware Layer

- JDBC
- OCI
- ODBC
- Perl

JDBC

- Enlarge the array size
- Set SQL*Net CONNECT_TIMEOUT property
- Set SQL*Net READ_TIMEOUT property
  - Do not try to use it as a query timeout
- For query timeout use
  - `user statement cancel`
  - `statement.setQueryTimeout()`
- Set SDU
  - `jdbc:oracle:thin:@(DESCRIPTION...(SDU=32767)...)`

OCI

- Enlarge the array size
  - OCIDefineArrayOfStruct()
  - OCIDefineByPos()

OCIDefineArrayOfStruct()

Purpose

This call specifies additional attributes necessary for a static array define, used in an array of structures (multi-row, multi-column) fetch.

Syntax

```
sword OCIDefineArrayOfStruct ( OCIDefine *defnp,
                               OCIError  *errhp,
                               ub4        pvskip,
                               ub4        indskip,
                               ub4        rlskip,
                               ub4        rcskip );
```

ODBC

- Disable autocommit
- Enlarge the default packet size
  - The default is 512 bytes
  - When set to any integer from 1 to 10, the driver computes the packet size by multiplying the value by 512 bytes (for example, 6 sets the packet size to 6 * 512 = 3072 bytes)
- Enlarge the default array size
  - The default is 50 rows
  - And Oracle cursor loop by default retrieves 100 and a larger value may be better still

Perl (1:2)

- Autocommit
  - The default value of 1 forces an autocommit after every row
  - This is without question the worst possible default value
  - Always set this to 0 and issue explicit commits based on the largest possible logical transaction
- ORA_CHECK_SQL
  - It is best practice to take advantage of Oracle's "deferred parse" which decreases database calls, and which combines the "parse" and "execute" phases of SQL processing but is not enabled by default

Perl (2:2)

- RaiseError
  - The default value is 0 which is a poor choice for almost all applications
  - Reset this value to 1 and exception will be sent to your script terminating it should an Oracle exception be generated
- RowCacheSize
  - Creates a local pre-fetch cache and defines its size
  - Not having an enabled cache provides the worst possible performance

```
${AutoCommit} = 0;
${ora_check_sql} = 0;
${RaiseError} = 1;
${RowCacheSize} = 1000;
```

Application Layer

Application Layer

- Code written to be database agnostic will provide mediocre performance
- Code written to be database agnostic will require more expensive infrastructure
- Write code using database specific optimizations
- Utilize "real" connection pooling
  - Connections persist
  - Connections are reused
  - Connections that can not be reused are killed
  - Exception handling identifies the details about what failed

Database

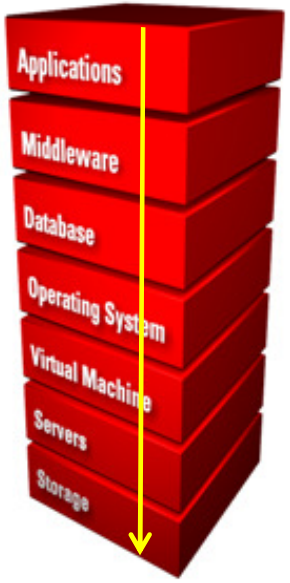
- Application developers like to cache tables on their application servers to improve performance
- But how do they know when the underlying table changes?
  - Often by repeatedly querying multiple tables from every application server every # minutes
  - A massive amount of DML that usually returns no changes
- DBMS_CHANGE_NOTIFICATION can be used to notify the application when a registered table has changed
  - Which can substantially reduce overhead on the DB server by reducing both cpu and i/o utilization

Middleware

- Coherence
  - Applications cache data in the data grid avoiding expensive database accesses
  - An in-memory scale-out data abstraction layer
  - Provides fast access to frequently used data
  - Cache automatically and dynamically partitions data
  - Ensures continuous data availability
  - Ensures transactional integrity
  - Full JCache support
- GoldenGate HotCache
  - Keeps the database and Coherence cached data in sync so "stale data" is never delivered to the application

Things I Recommend You Do Not Do

- Disable logging
- Follow much of the advice on the internet
- Use any undocumented hint
- Use any underscore parameter without getting Oracle to agree in an SR



Performance Tuning Conclusion (1:2)

- The majority of tuning errors happen during the purchase and configuration of infrastructure
- The majority of remaining tuning errors are made during application design and deployment
- Exadata, Neteeza, ODA, Teradata all reject the shared services concept: Consider why
- You can't make up for buying the wrong infrastructure by tuning a SQL statement
- Shared Services SANs are essentially a guarantee of performance issues
- No one is going to let you turn a heap table into a single table hash cluster a year after deployment

Performance Tuning Conclusion (2:2)

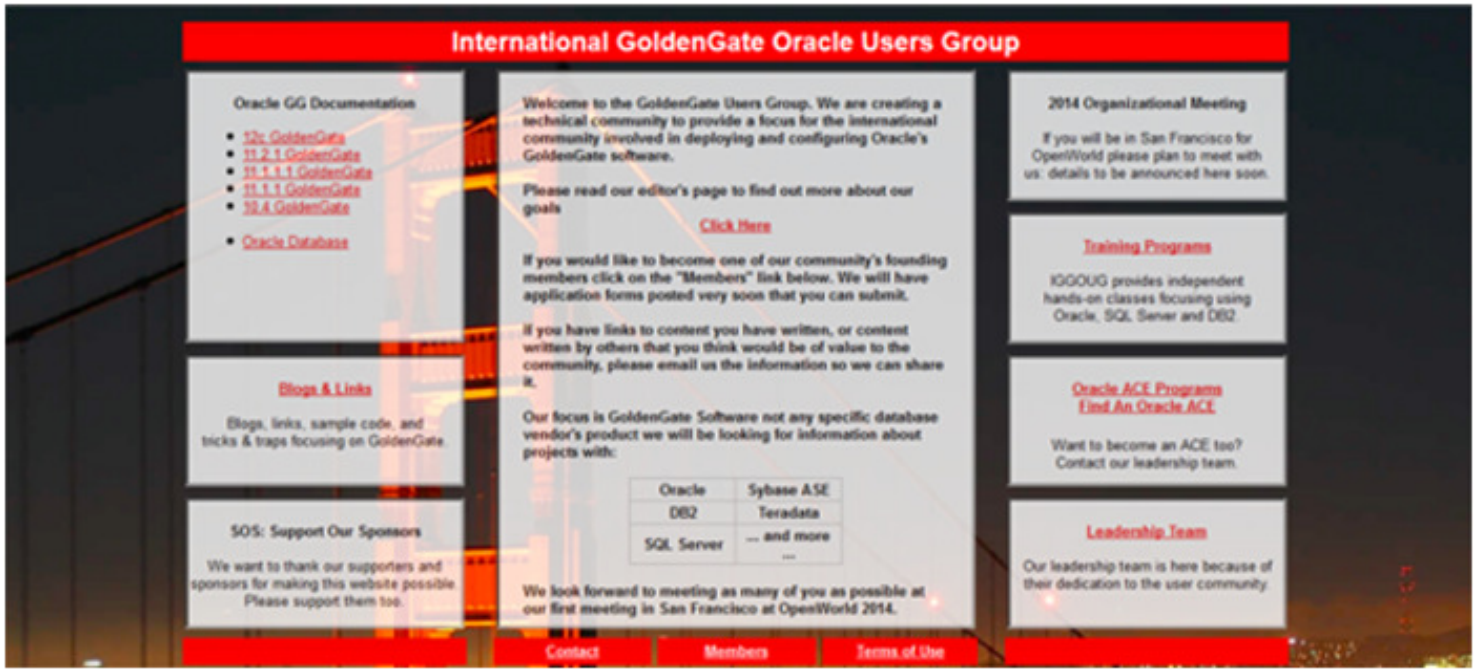
- Unless you like fire fighting give serious consideration to how little of this presentation had to do with turning an inner join into a scalar subquery or how many hints were recommended for you to use
- Your first job is to "OWN" the database and that means to educate yourself about the infrastructure it touches and to help your organization make better choices in the future which will
  - Save money
  - Improve your ability to meet SLAs
  - Allow you to focus on other aspects of your job

User Groups

International GoldenGate OUG

www.iggoug.org

- Do you or co-workers have in interest in
  - Oracle GoldenGate
  - Oracle Data Integrator EE
  - Oracle Data Profiling
  - Oracle Data Quality



Indiana OUG

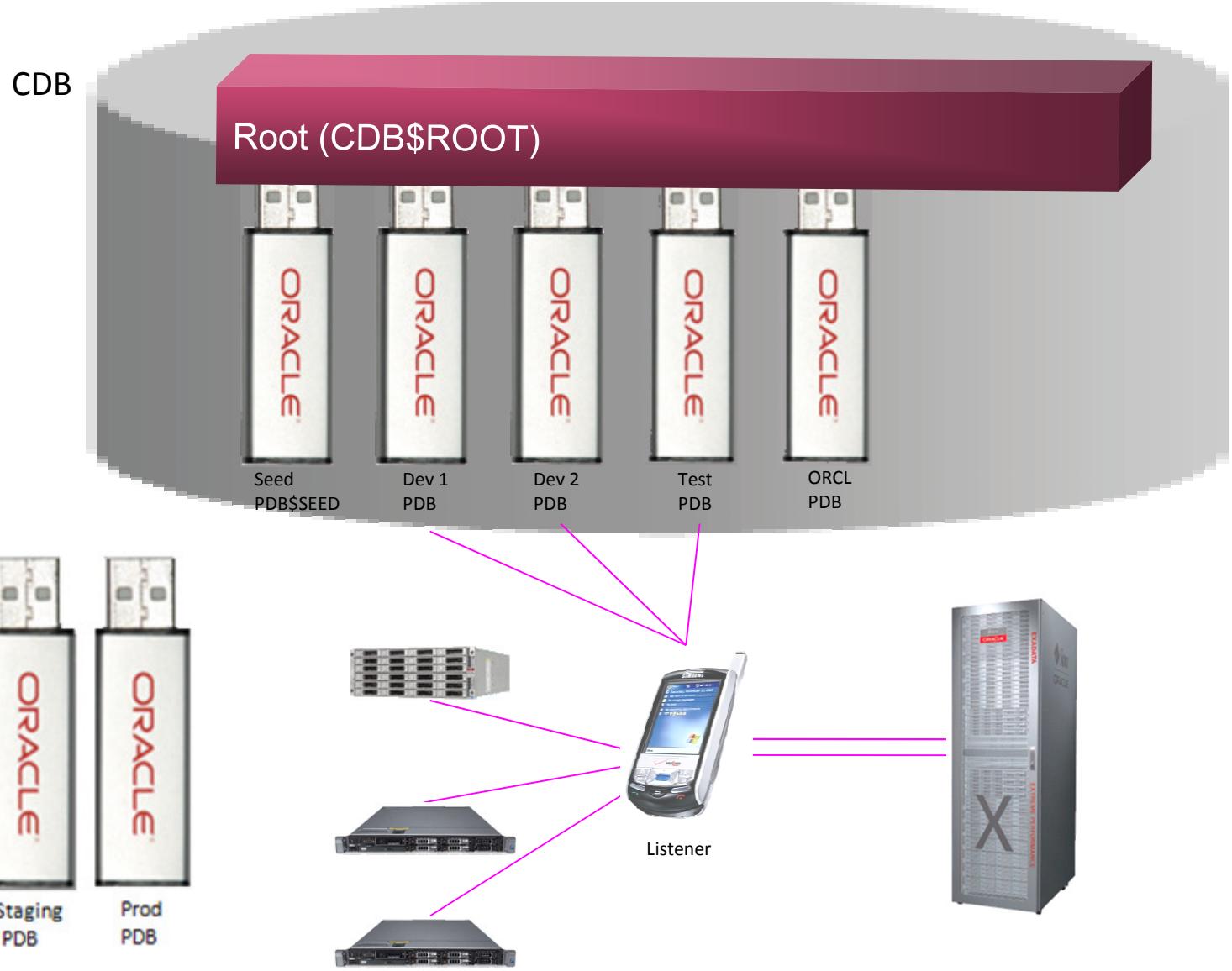
- Are you a member?
- Have you volunteered?
- Have you served on the board?

New in 12c
... and you won't hear about it from Oracle

New Infrastructure

- Oracle has made two equally important changes in the past
  - Version 4: Redo Logs
  - Version 6: Rollback Segments
- The word database should be deprecated
  - The statement "create a new database" has no meaning
  - The statement "start the database" has no meaning
  - The statement "shut down the database" has no meaning
  - The statement "backup the database" has no meaning
  - The statement "the database crashed" has no meaning
- The word "database" in an Oracle sense no longer has a definition
- We need to make sure that the words we use have meaning or mistakes are inevitable

The New Architecture



Consider Your Current Server Inventory

- Can we use our current servers? Or more importantly "should" we use our current servers
  - Pizza Boxes (1U, 2U, 3U)
  - Blades (including UCS)
  - P Series (IBM)
  - Z Series (IBM)
  - M10 (Fujitsu)
  - T Series (Sun-Oracle)
  - M Series (Sun-Oracle)
  - ODA
  - Exadata

Consolidation Effects

- What are the implications of consolidation?
  - Patching requirements
  - Processes
  - Transactions
  - Connections
  - Memory
  - CPU
  - I/O
- Do you really want to use only one TEMP tablespace? Consider what Oracle forces in 12c
- How do you allocate space for UNDO?
  - You went from many to one
- Networks

New Versions and Concepts

- 12.1.0.1
  - Multitenant
    - Perhaps you are ready for multi-tenant
    - I bet your infrastructure isn't
  - Integrated Lifecycle Management (ILM)
  - Credentials
  - Policies
- 12.1.0.2
  - Automatic Big Table Caching
  - Full Database Caching
  - In-Memory Aggregation and In-Memory Column Store
  - JSON Support
  - Rapid Home Provisioning
  - Real Application Security

Should I Deploy a Container DB?

- To create a container DB: Yes or Yes?
  - Separation of duties
  - cron jobs
- To multitenant consolidate: Yes or No?
  - Patching requirements
  - Upgrade requirements
  - Backup and restore requirements
  - Security requirements
  - Infrastructure

Security Model Changes

- New Users with Enhanced Privileges
- New Roles
- New System Privileges
  - And changes to existing ones such as SELECT ANY DICTIONARY
- New Unified Audit Policies
- New Credentials
- New Operating System Users and Groups
- Enhanced Encryption SHA-2 with DBMS_CRYPTO

Operating System Groups & Enhanced Priv. Users

- Legacy Operating System
  - root
  - oinstall
  - dba
  - oper
- Database
  - SYS
  - SYSTEM
  - DBA role

12c Users with Escalated Privileges

```
SQL> SELECT username FROM dba_users ORDER BY 1;

USERNAME
------------------------------
ANONYMOUS
APEX_040200
APEX_PUBLIC_USER
APPQOSSYS
AUDSYS
CTXSYS
DBSNMP
DIP
DVF
DVSYS
FLOWS_FILES
GSMADMIN_INTERNAL
GSMCATUSER
GSMUSER
LBACSYS
MDDATA
MDSYS
OJVMSYS
OLAPSYS
ORACLE_OCM
ORDDATA
ORDPLUGINS
ORDSYS
OUTLN
SI_INFORMTN_SCHEMA
SPATIAL_CSW_ADMIN_USR
SPATIAL_WFS_ADMIN_USR
SYS
SYSBACKUP
SYSDG
SYSKM
SYSTEM
WMSYS
XDB
XS$NULL
```

Roles Enhancing Separation of Duties

```
SQL> select * from dba_roles order by 1;

ROLE                           PASSWORD AUTHENTICAT COM O
------------------------------ -------- ----------- --- -
ADM_PARALLEL_EXECUTE_TASK      NO       NONE        YES Y
APEX_ADMINISTRATOR_ROLE        NO       NONE        YES Y
APEX_GRANTS_FOR_NEW_USERS_ROLE NO       NONE        YES Y
AQ_ADMINISTRATOR_ROLE          NO       NONE        YES Y
AUDIT_ADMIN                    NO       NONE        YES Y
CDB_DBA                        NO       NONE        YES Y
DBA                            NO       NONE        YES Y
DBFS_ROLE                      NO       NONE        YES Y
GATHER_SYSTEM_STATISTICS       NO       NONE        YES Y
GSMADMIN_ROLE                  NO       NONE        YES Y
GSMUSER_ROLE                   NO       NONE        YES Y
GSM_POOLADMIN_ROLE             NO       NONE        YES Y
HS_ADMIN_EXECUTE_ROLE          NO       NONE        YES Y
HS_ADMIN_ROLE                  NO       NONE        YES Y
HS_ADMIN_SELECT_ROLE           NO       NONE        YES Y
LBAC_DBA                       NO       NONE        YES Y
LOGSTDBY_ADMINISTRATOR         NO       NONE        YES Y
OLAP_DBA                       NO       NONE        YES Y
OLAP_XS_ADMIN                  NO       NONE        YES Y
OPTIMIZER_PROCESSING_RATE      NO       NONE        YES Y
ORDADMIN                       NO       NONE        YES Y
PDB_DBA                        NO       NONE        YES Y
PROVISIONER                    NO       NONE        YES Y
RECOVERY_CATALOG_OWNER         NO       NONE        YES Y
SCHEDULER_ADMIN                NO       NONE        YES Y
SPATIAL_CSW_ADMIN              NO       NONE        YES Y
SPATIAL_WFS_ADMIN              NO       NONE        YES Y
WM_ADMIN_ROLE                  NO       NONE        YES Y
XDBADMIN                       NO       NONE        YES Y
XS_CACHE_ADMIN                 NO       NONE        YES Y
XS_NAMESPACE_ADMIN             NO       NONE        YES Y
XS_SESSION_ADMIN               NO       NONE        YES Y
```

System Privileges

```
-- edited to show new privs

SQL> SELECT DISTINCT privilege
     FROM dba_sys_privs
     ORDER BY 1;

PRIVILEGE
----------------------------------------
ALTER ANY SQL TRANSLATION PROFILE
CREATE ANY CREDENTIAL
CREATE ANY SQL TRANSLATION PROFILE
CREATE CREDENTIAL
CREATE PLUGGABLE DATABASE
CREATE SQL TRANSLATION PROFILE
DROP ANY SQL TRANSLATION PROFILE
EXEMPT DDL REDACTION POLICY
EXEMPT DML REDACTION POLICY
REDEFINE ANY TABLE
SET CONTAINER
USE ANY SQL TRANSLATION PROFILE
```

Auditing

- Legacy Auditing

```
AUDIT CREATE PROCEDURE;
AUDIT TABLE;
AUDIT CREATE TABLE;
AUDIT CREATE TRIGGER;
AUDIT CREATE VIEW;
```

- Unified Audit Policies

```
CREATE AUDIT POLICY uw_full_clause
PRIVILEGES ALTER ANY TABLE
ACTIONS LOGOFF ROLES DBA
WHEN 'SYS_CONTEXT(''USERENV'', ''ISDBA'') = ''TRUE'''
EVALUATE PER STATEMENT
CONTAINER = ALL;
```

Credentials

- DBMS_CREDENTIAL
  - Interface for authenticating and impersonating EXTPROC callout functions, as well as external jobs, remote jobs and file watchers from the SCHEDULER

```
DECLARE
 cname   user_credentials.credential_name%TYPE := 'UWCRED';
 uname   user_credentials.username%TYPE := 'UWCLASS';
 pwd     sys.scheduler$_credential.password%TYPE := 'ZzYzX6*';
 dbrole  VARCHAR2(30) := NULL;
 windom  sys.scheduler$_credential.domain%TYPE := NULL;
 comment user_credentials.comments%TYPE := 'Test Cred';
 enable  BOOLEAN := FALSE;
BEGIN
  dbms_credential.create_credential(cname, uname, pwd, dbrole, windom, comment, enable);
END;
/
```

DDL Enhancements

- Basics
  - Enhancements to the CREATE USER syntax
    - Common Users and Roles
    - Editions
    - Containers
  - Column Size Expansion in the Data Dictionary
  - RMAN Single Table Recovery
  - ALTER DATABASE Move Datafile
  - ALTER DATABASE Control File Management
  - ALTER DATABASE FULL RECOVERY supports recovery from snapshot standbys
  - ALTER DATABASE Switchover and Failover Clauses
  - ADO (Automatic Data Optimization)
  - Configurable Threaded Execution

Object Enhancements (1:3)

- Tables
  - New ON_NULL clause for CREATE AND ALTER TABLE
  - New INVISIBLE clause
  - Temporal Validity Columns
  - USING TRUSTED CONSTRAINTS clause for Partitions
  - Identity Columns
  - 32K max size for VARCHAR2 and NVARCHAR2 columns
  - Evaluate Edition Clause
  - Table Unusable Editions Clause
  - ILM Policy Clause
  - Table Period Definition
  - Table Tiering

Object Enhancements (2:3)

- Indexes
  - ALTER INDEX COALESCE includes CLEANUP syntax
  - New CASCADE clause
  - New Asynchronous Global Index maintenance clause for partitioned tables
- Constraints
  - Online Drop Constraint
- Sequences
  - Identity Columns where sequences are created implicitly and bound to a table column
  - KEEP and NOKEEP affecting NEXTVAL during replays
  - Session and Global syntax for CREATE and ALTER sequences

Object Enhancements (3:3)

- Materialized View
  - USING TRUSTED CONSTRAINTS clause

SELECT Statement and XML Enhancements

- MATCH_RECOGNIZE function for Map Reduce
- Model Clause extended with new NEXT and PREV functions
- New INSERTCHILDXMLAFTER function
- New INSERTCHILDXMLBEFORE function

New PL/SQL (1:2)

- ACCESSIBLE_BY Clause
- FETCH_FIRST Clause
- $$PLSQL_UNIT_OWNER Inquiry Directive
- $$PLSQL_UNIT_TYPE Inquiry Directive
- Invoker Right's functions can be Result Cached
- Roles can be granted to PL/SQL packages and stand-alone programs
- More PL/SQL data types can be used for SQL
- Implicit statement results
- UPDATE triggers for LOB columns
- Logical LOB UPDATE triggers
- PDB_NAME added to ALERT_TYPE data type

New PL/SQL (2:2)

- A library can be defined as a Directory object and WITH CREDENTIALS
- BEQUEATH CURRENT_USER Views
- INHERIT PRIVILEGE and INHERIT ANY PRIVILEGE system privileges
- Objects, not Types, are editioned or non-editioned with a new ALTER syntax
- PL/SQL functions run faster in SQL
- New PDB_NAME attribute in the ALTER_TYPE data type

New / Modified Functions, Operator, Procs, & Types

- Many new feature usage procs
- Classifier function now part of map reduce
- New CLUSTER_DETAILS, CLUSTER_DISTANCE and FEATURE_DETAILS analytic functions
- New CON_DBID_TO_ID, CON_GUID_TO_ID, CON_NAME_TO_ID, and CON_UID_TO_ID functions supporting CDBs
- New ORA_DST_AFFECTED, ORA_DST_CONVERT, and ORA_DST_ERROR functions for use with DBMS_DST
- New ORA_INVOKING_USER and ORA_INVOKING_USERID functions
- New STANDARD_HASH function

Packages With New Constants, Exceptions, Types

- DBMS_ADR
- DBMS_ADVISOR
- DBMS_AQADM
- DBMS_COMPRESSION
- DBMS_CRYPTO
- DBMS_DATA_PUMP
- DBMS_DB_VERSION
- DBMS_METADATA
- DBMS_RLS
- DBMS_SPACE_ADMIN
- DBMS_SQLHASH
- UTL_RECOMP

Packages With New Overloads

- DBMS_COMPRESSION
- DBMS_DATAPUMP
- DBMS_LOGMNR_LOGREP_DICT
- DBMS_METADATA_INT
- DBMS_PICKLER
- DBMS_REPORT
- DBMS_SQL
- DBMS_SYS_SQL
- DBMS_SERVICE
- DBMS_WORKLOAD_REPOSITORY
- ODM_MODEL_UTIL
- UTL_XML

Packages With New Parameters (1:2)

- AS_REPLAY
- DBMS_APPLY_ADM
- DBMS_AUDIT_MGMT
- DBMS_CAPTURE_ADM
- DBMS_CLIENT_RESULT_CACHE
- DBMS_GOLDENGATE_AUTH
- DBMS_INDEX_UTL
- DBMS_RESOURCE_MANAGER
- DBMS_REDEFINITION
- DBMS_RLS
- DBMS_SCHEDULER
- DBMS_SERVER_ALERT

Packages With New Parameters (2:2)

- DBMS_SQLPA
- DBMS_SQLTUNE
- DBMS_STATS
- DBMS_XSTREAM_ADM

New Packages (1:2)

- DBMS_APP_CONT
- DBMS_AUTO_REPORT
- DBMS_CREDENTIAL
- DBMS_HEAT_MAP
- DBMS_ILM (integrated life-cycle management)
- DBMS_ILM_ADMIN (integrated life-cycle management)
- DBMS_PART (enhanced partition table management)
- DBMS_PDB
- DBMS_PRIVILEGE_CAPTURE
- DBMS_QOPATCH
- DBMS_REDACT (data masking)
- DBMS_ROLLING (application of rolling patches)

New Packages (2:2)

- DBMS_SPD
- DBMS_SQL_MONITOR
- DBMS_SQL_TRANSLATOR
- DBMS_SYNC_REFRESH
- DBMS_TSDP_MANAGE
- DBMS_TSDP_PROTECT
- UTL_CALL_STACK

Network Access Control List Management

- DBMS_NETWORK_ACL_ADMIN Deprecated Objects
  - ADD_PRIVILEGE
  - ASSIGN_ACL
  - ASSIGN_WALLET_ACL
  - CHECK_PRIVILEGE
  - CHECK_PRIVILEGE_ACLID
  - CREATE_ACL
  - DELETE_PRIVILEGE
  - DROP_ACL
  - UNASSIGN_ACL
  - UNASSIGN_WALLET_ACL

Network Access Control List Management

- DBMS_NETWORK_ACL_ADMIN New Objects
  - APPEND_HOST_ACE
  - APPEND_HOST_ACL
  - APPEND_WALLET_ACE
  - APPEND_WALLET_ACL
  - REMOVE_HOST_ACE
  - REMOVE_WALLET_ACE
  - SET_HOST_ACL
  - SET_WALLET_ACL

Advanced Index Compression (12.1.0.2)

- Prefix Compression
  Advanced Compression
- Works on many index types, including those that are not good candidates for the multi-column prefix including indexes with no, or few, duplicate values in the leading columns of the index
- Improves the compression ratios significantly while still providing efficient index access
- Some restrictions exist

```
CREATE INDEX ix_airplanes_custid_deldate
ON uwclass.airplanes(customer_id, delivered_date)
COMPRESS ADVANCED LOW;
```

Partial Index Clause (12.1.0.2)

- Local and global indexes can be created on a subset of the partitions of a table
- Partial indexes provide more flexibility in index creation for partitioned tables
- For example, index segments can be omitted for the most recent partitions to ensure maximum data ingest rates without impacting the overall data model and access for the partitioned object

```
[INDEXING <PARTIAL | FULL>]
```

Rapid Home Provisioning (12.1.0.2)

- Allows deploying of Oracle homes based on gold images stored in a catalog of pre-created homes
- Provisioning time for Oracle Database is significantly improved through centralized management while the updating of homes is simplified to linkage
- Oracle snapshot technology is used internally to further improve the sharing of homes across clusters and to reduce storage requirements

Zone Maps (12.1.0.2)

- For full table access, zone maps allow I/O pruning of data based on the physical location of the data on disk, acting like an anti-index
- Accessing only the relevant data optimizes the I/O necessary to satisfy a query, increasing the performance significantly and reducing resource consumption

What Else Is Deprecated

- Control Data Capture (CDC)
- Data Mining Java API (DBMS_DATA_MINING)
- DBMS_DBLINK built-in package
- Numerous startup parameters
- Resource Role no longer grants UNLIMITED TABLESPACE
- Streams (for multi-tenant)
- Stored Outlines (use Baselines)
- WMSYS CONCAT function

12c Conclusion

- Download and install it immediately
- Learn everything you can: Take nothing for granted
- The more experience you have as a DBA or Senior Developer the more likely you will blunder into mistakes because of assumptions that are not valid with the new architecture
- You are going to make mistakes
  - Do you want to make them at home
    - Or on the job?

Questions

ERROR at line 1:
ORA-00028: your session has been killed

Feel free to ask questions now or contact me at PTC
daniel.morgan@perftuning.com / +1 206-669-2949

Thank You