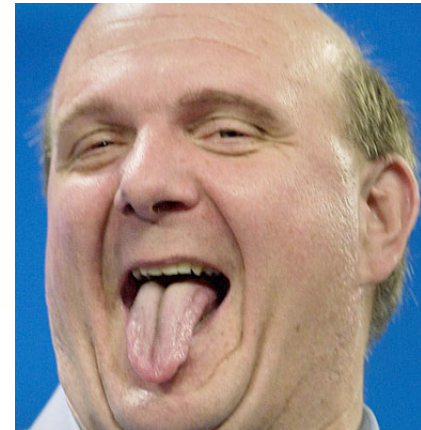


DATABASE DEATH MATCH

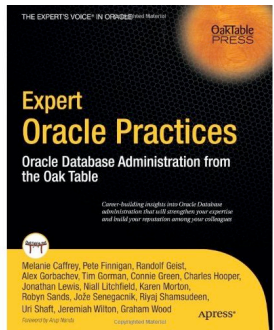
Oracle vs:

SQL Server



Jeremiah Wilton

jwilton@bluegecko.net



- 17 years working with Oracle
- Amazon.com's first DBA ~1997
- Technical Lead, Blue Gecko, Inc.
 - Remote administration, support and expertise
 - Oracle, E-Business Suite, MySQL, PostgreSQL, SQL Server
- Univ. of Washington Oracle instructor
- President, Western Washington Oracle Users Group
- Co-author: Expert Oracle Practices (Oak Table)
- Frequent speaker at Oracle conferences
- Current focus:
 - Oracle performance on the Cloud (Amazon EC2)
 - Time-based tuning for SQL Server



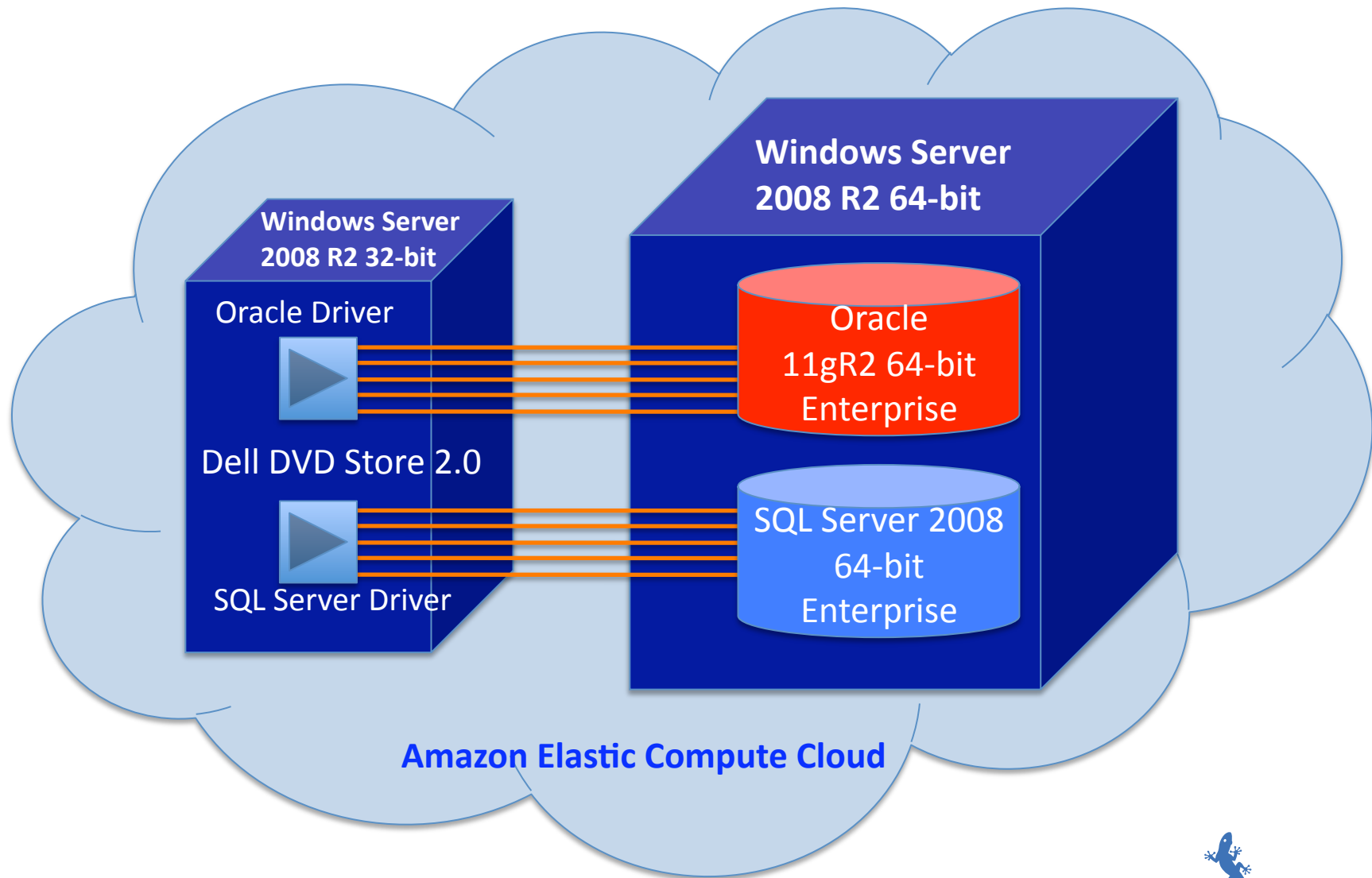
Rationale

- Fun
- Interesting
- Not boring
- Learn about SQL Server
- Break things
- Feel superior
- Engage in pretend science (à la Mythbusters)

Caveats

- Oracle and SQL Server are not the same
- I am not a Windows or SQL Server expert
- My tests may have a pro-Oracle bias
- I used a minimum configuration and engineering standard
 - No standby, replication, log shipping, etc.
 - Basic out-of-box functionality

Method



Tools I Use

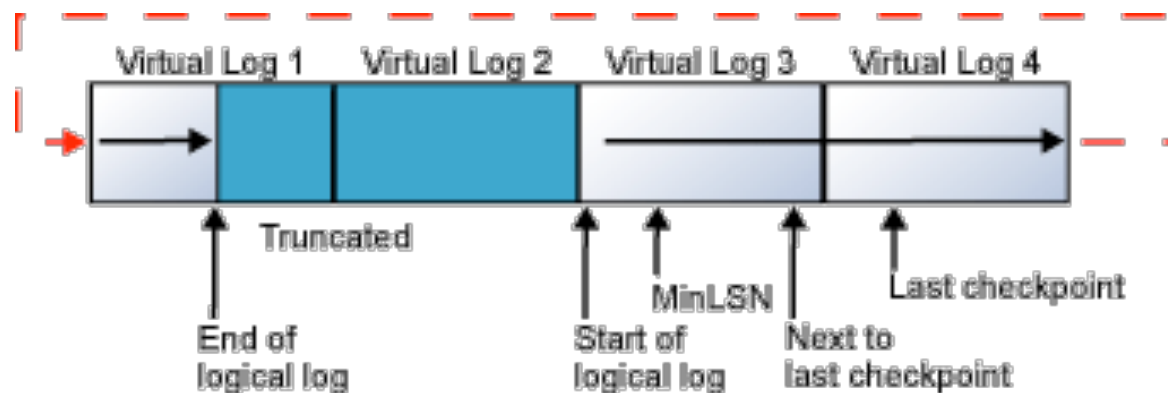
- GNU Core Utilities
 - tail
 - dd
 - sed
- Microsoft SysInternals Process Explorer
- sqlcmd
- sqlplus

The Four Ordeals

1. Loss of volume containing transaction log
2. Hang/crash of data file writer thread
3. Corrupt a single 8k block
4. Side-by-side competition for resources

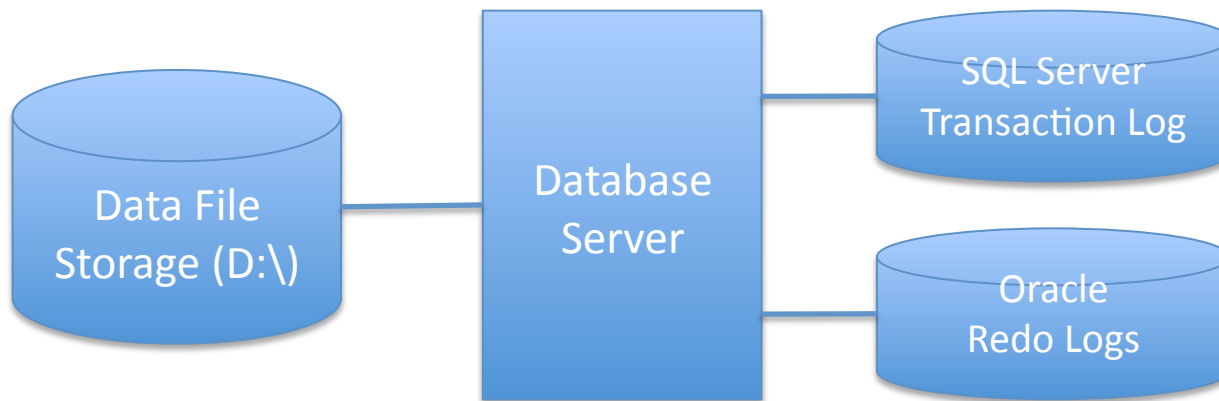
SQL Server Transaction Log

- Similar to Oracle's redo log (circular)
- Also provides undo function
- One per database (multiple databases per server)
- Usually one giant file, n 'virtual' logs within the one physical file
- Regular T-Log backups provide "archive" functionality
- If no backup before T-log is full, file grows (Full recovery model)
- **No native log redundancy mechanism**



SQL Server: Loss of log volume

- Each contender has a dedicated volume for logs



1. Apply load with DS2

```
ds2sqlserverdriver --target=dm --n_threads=5 --pct_newcustomers=0
```

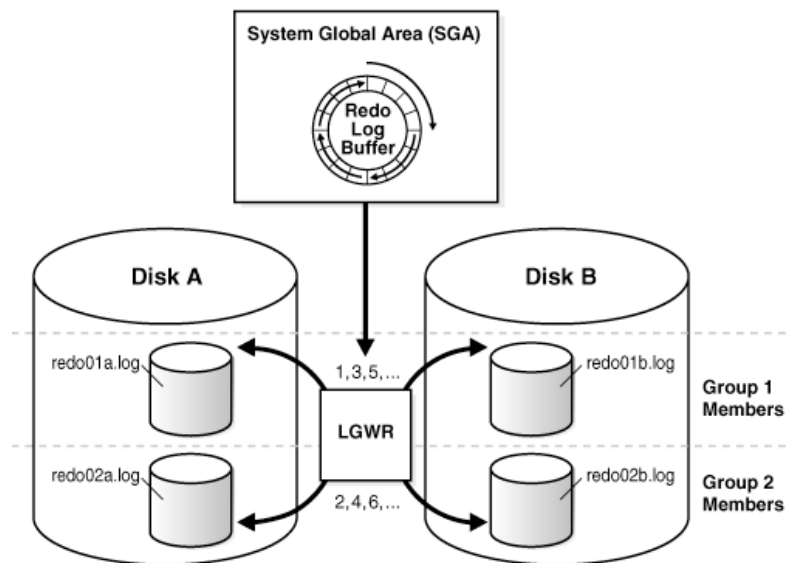
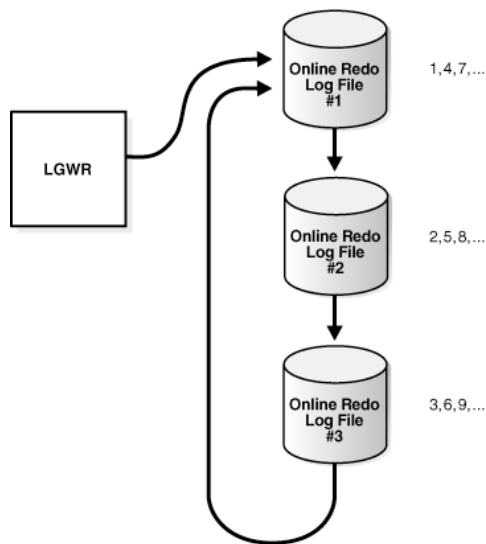
2. Tail the error log

```
tail -f "c:\Program Files\Microsoft SQL Server\MSSQL10.MSSQLSERVER\MSSQL\Log\ERRORLOG"  
| sed s/\x00//g
```

3. Detach log storage volume
4. Observe

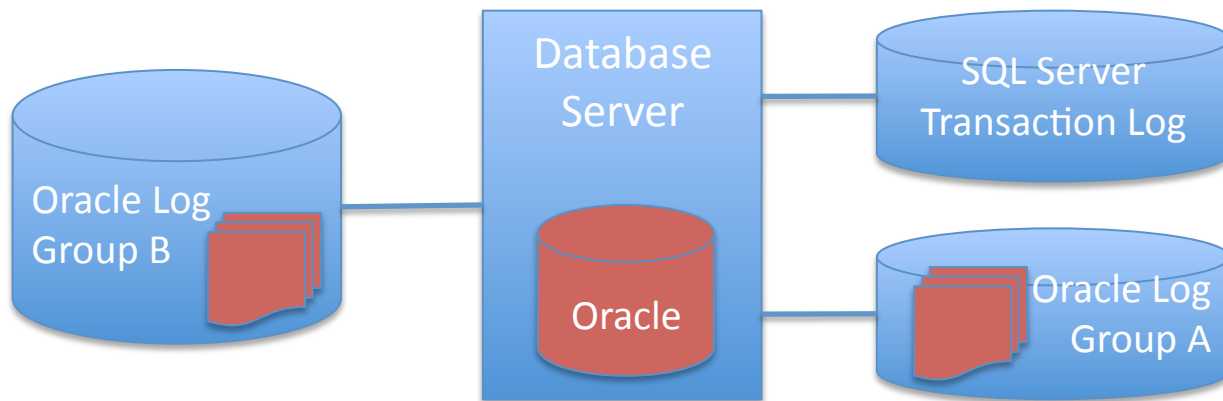
Oracle Redo Logs

- “Oracle recommends that you multiplex the online redo log.” -11gR2 Docs (standard practice)



Oracle: Loss of log volume

- Each contender has a dedicated volume for logs



1. Apply load with DS2

```
c:\ds2\ds2oracledriver --target=dm --n_threads=5 --pct_newcustomers=0
```

2. Tail the alert log

```
tail -f c:\app\Administrator\diag\rdbms\dm\dm\trace>alert_dm.log
```

3. Detach log group A storage volume

4. Observe

5. Drop missing logfile group members

SQL Server Lazy Writer

- Not the same as the Oracle Database Writer
- Writes some dirty buffers between checkpoints
- Worker threads also write dirty buffers
- Checkpoint also writes dirty buffers (at checkpoint)
- How to find it?

```
1> select session_id, command, os_thread_id
2> from sys.dm_exec_requests as r
3> join sys.dm_os_workers as w
4>   on r.task_address = w.task_address
5> join sys.dm_os_threads as t
6>   on t.thread_address = w.thread_address
7> where session_id <= 50
8> order by session_id
9> go
```

session_id	command	os_thread_id
1	RESOURCE MONITOR	2776
2	XE TIMER	2908
3	XE DISPATCHER	2416
4	LAZY WRITER	1316
5	LOG WRITER	2404
6	LOCK MONITOR	1592
7	SIGNAL HANDLER	1516
9	TRACE QUEUE TASK	1544
10	BRKR TASK	1540
11	TASK MANAGER	1524
13	CHECKPOINT	1520
14	BRKR EVENT HNDLR	1556
15	BRKR TASK	1552
16	BRKR TASK	1536

SysInternals Process Explorer

The screenshot displays the SysInternals Process Explorer interface. The main window shows a list of processes with columns for Name, PID, and CPU usage. The 'sqlservr.exe' process is selected, and its properties are shown in a separate window titled 'sqlservr.exe:6496 Properties'. The 'Threads' tab is active, showing a list of threads with columns for Thread ID, CPU, Cycles Delta, Service, and Start Address. The thread with ID 1372 is selected, and its detailed properties are shown below the list.

Process	PID	CPU
taskeng.exe	4084	
SLsvc.exe	984	
svchost.exe	244	
svchost.exe	428	
dwm.exe	3424	
svchost.exe	444	
rdpclip.exe	3616	
svchost.exe	1072	
spoolsv.exe	1240	
nmsvc.exe	1408	
cmd.exe	1892	
perf.exe	1908	
cmd.exe	996	
java.exe	1116	
emagent.exe	1500	
omtsreco.exe	1420	
TNSLSNR.EXE	1784	
oracle.exe	1816	0.77
svchost.exe	1844	
svchost.exe	1856	
RightScaleService.exe	2000	
sqlwriter.exe	1940	
svchost.exe	1956	
Ec2Config.exe	1452	
rhelsvc.exe	1396	
svchost.exe	3128	
msdtc.exe	3648	
sqlservr.exe	6496	
fdlauncher.exe	6616	
fdhost.exe	6844	
sass.exe	572	
lsn.exe		

T.	CPU	Cycles Delta	Service	Start Address
1372			MSSQLSERVER	MSVCR80.dll!lenthreadex+0x80
5032			MSSQLSERVER	MSVCR80.dll!lenthreadex+0x80
5068			MSSQLSERVER	MSVCR80.dll!lenthreadex+0x80
5092			MSSQLSERVER	MSVCR80.dll!lenthreadex+0x80
5152			MSSQLSERVER	MSVCR80.dll!lenthreadex+0x80
5236				sqlservr.exe+0xd78e00
5528			MSSQLSERVER	MSVCR80.dll!lenthreadex+0x80
5800			MSSQLSERVER	MSVCR80.dll!lenthreadex+0x80
5836			MSSQLSERVER	MSVCR80.dll!lenthreadex+0x80
6244			MSSQLSERVER	MSVCR80.dll!lenthreadex+0x80

Thread ID:	1372	Stack	Module
Start Time:	8:41:13 PM 8/19/2010		
State:	Wait:UserRequest	Base Priority:	8
Kernel Time:	0:00:00.015	Dynamic Priority:	8
User Time:	0:00:00.015	I/O Priority:	Normal
Context Switches:	4,175	Memory Priority:	5
Cycles:	657,620,005		

Permissions Kill Suspend

OK Cancel

CPU Usage: 4.62% Commit Charge: 40.19% Processes: 66 Physical Usage: 49.33%

SQL Server: Hang/Crash Lazy Writer

1. Apply load with DS2
2. Tail the error log
3. Suspend the Lazy Writer thread (procexp)
4. Observe
5. Resume the Lazy Writer thread (procexp)
6. Kill the Lazy Writer thread (procexp)
7. Observe

Oracle Database Writer(s)

- Responsible for all writes of dirty blocks
- How to find it in Windows?

```
SQL> select spid, b.name,  
2         description  
3   from v$process p,  
4         v$bgprocess b  
5  where p.addr = b.paddr  
6  and p.serial# = b.pserial#;
```

```
SPID NAME DESCRIPTION  
---- ----  
-----  
2116 PMON process cleanup  
2120 VKTM Virtual Keeper of TiMe process  
2124 GEN0 generic0  
2144 DIAG diagnosibility process  
2148 DBRM DataBase Resource Manager  
2160 PSP0 process spawner 0  
2164 DIA0 diagnosibility process 0  
2168 MMAN Memory Manager  
2172 DBW0 db writer process 0  
2176 LGWR Redo etc.  
2180 CKPT checkpoint  
2184 SMON System Monitor Process  
2188 RECO distributed recovery  
2192 MMON Manageability Monitor Process  
2200 MMNL Manageability Monitor Process 2  
2372 RVWR Recovery Writer  
3972 QMNC AQ Coordinator  
2488 ARC1 Archival Process 1  
2440 ARC0 Archival Process 0  
2500 ARC2 Archival Process 2  
2572 ARC3 Archival Process 3  
3960 CJQ0 Job Queue Coordinator  
3896 SMC0 Space Manager Process
```

Oracle: Hang/Crash Database Writer

1. Apply load with DS2
2. Tail the alert log
3. Suspend the DBW0 thread
4. Observe
5. Resume the DBW0 thread
6. Kill the DBW0 thread
7. Observe

Single-Block Corruption

- SQL Server calls them pages
 - Oracle calls them blocks
 - Both are 8k by default
1. Find the first block of the CUSTOMERS table
 2. SQL Server and Oracle “trade” the block
 3. Try to read the data
 4. Try to recover the block

Physical Address of a SQL Server Row

- %%PHYSLOC%% is the equivalent of ROWID
 - File:Block:Slot (sound familiar?)

```
1> select top 10
2>   sys.fn_physlocformatter(%%PHYSLOC%%)
3>   customerid, firstname
4> from ds2.dbo.customers
5> go

1> select physical_name
2> from ds2.sys.database_files
3> where file_id = 4
4> go

physical_name
-----
C:\sql\dbfiles\ds2_2.ndf
```

physloc	customerid	firstname
(4:896:0)	1	VKUUXF
(4:896:1)	2	HQNMZH
(4:896:2)	3	JTNRNB
(4:896:3)	4	XMFYXD
(4:896:4)	5	PGDTDU
(4:896:5)	6	FXDZBW
(4:896:6)	7	WVZTXZ
(4:896:7)	8	LIWLAI
(4:896:8)	9	NCGWRC
(4:896:9)	10	FUOHXX

Physical Address of an Oracle Row

- We use ROWID / dbms_rowid

```
SQL> select
2  dbms_rowid.rowid_relative_fno(rowid),
3  dbms_rowid.rowid_block_number(rowid),
4  customerid, firstname
5  from ds2.customers
6  where rownum <= 10;
```

FILE	BLOCK	CUSTOMERID	FIRSTNAME
6	131	1	VKUUXF
6	131	2	HQNMZH
6	131	3	JTNRNB
6	131	4	XMFYXD
6	131	5	PGDTDU
6	131	6	FXDZBW
6	131	7	WVZTXZ
6	131	8	LIWLAI
6	131	9	NCGWRC
6	131	10	FUOHXX

```
SQL> select name
2  from v$datafile
3  where file# = 6;
```

NAME

D:\APP\ADMINISTRATOR\ORADATA\DM\CUST.DBF

Swap the first block of CUSTOMERS

```
1> alter database ds2 set offline
```

```
2> go
```

```
C:\>dd bs=8k skip=896 count=1
```

```
    if=d:\sql\dbfiles\cust1.ndf of=c:\sb1
```

```
C:\>dd bs=8k skip=131 count=1
```

```
    if=d:\app\administrator\oradata\dm\cust.dbf of=c:\ob1
```

```
C:\>dd bs=8k conv=notrunc seek=896
```

```
    if=c:\ob1 of=d:\sql\dbfiles\cust1.ndf
```

```
C:\>dd bs=8k conv=notrunc seek=131
```

```
    if=c:\sb1 of=d:\app\administrator\oradata\dm\cust.dbf
```

```
1> alter database ds2 set online
```

```
2> go
```

SQL Server & Oracle: Read Corrupted Page

```
1> select top 10 firstname
2> from ds2.dbo.customers
3> go
```

```
Msg 824, Level 24, State 2, Server IP-0AF647E0, Line 1
SQL Server detected a logical consistency-based I/O error: incorrect
pageid (expected 4:896; actual 3:0). It occurred during a read of page
(4:896) in database ID 5 at offset 0x00000000700000 in file 'C:\sql
\dbfiles\cust1.ndf'.
```

Additional messages in the SQL Server error log or system event log may provide more detail.

```
SQL> alter system flush buffer_cache;
```

```
SQL> select firstname
  2  from ds2.customers
  3  where rownum <= 10;
```

*

ERROR at line 1:

```
ORA-01578: ORACLE data block corrupted (file # 6, block # 131)
```

```
ORA-01110: data file 6: 'D:\APP\ADMINISTRATOR\ORADATA\DM\CUST.DBF'
```

SQL Server: Recovery of Corrupt Block

```
1> backup log ds2
2> to disk='c:\bak\log.bak'
3> go
```

```
1> restore database ds2 page='4:896'
2> from disk='c:\bak\db.bak' with norecovery
3> go
```

```
1> restore log ds2
2> from disk='c:\bak\db.bak' with norecovery
3> go
```

```
1> restore log ds2
2> from disk='c:\bak\log.bak' with recovery
3> go
```

```
1> alter database ds2 set online
2> go
```

Oracle: Recovery of Corrupt Block

- Oracle has had *blockrecover* since 9i
- One line does it all!

```
RMAN> blockrecover datafile 6 block 131;
```

THE FINAL ORDEAL

- Side-by-side competition for resources
 1. Apply load to both Oracle and SQL Server
 2. Verify 100% CPU utilization
 3. Observe TPM numbers

Questions?