

Applying traditional DBA skills to Oracle Exadata

Marc Fielding

March 2013

About Me



- Senior Consultant with Pythian's Advanced Technology Group
- 12+ years Oracle production systems experience starting with Oracle 7
- 3+ years with Exadata
- Blogger and conference presenter
www.pythian.com/news/author/fielding
- Occasionally on twitter: @mfild

About Pythian

- **Recognized Leader:**
 - Global industry-leader in remote database administration services and consulting for Oracle, Oracle Applications, MySQL and Microsoft SQL Server
 - Work with over 250 multinational companies such as Forbes.com, Fox Sports, Nordion and Western Union to help manage their complex IT deployments
- **Expertise:**
 - Pythian's data experts are the elite in their field. We have the highest concentration of Oracle ACEs on staff—9 including 2 ACE Directors—and 2 Microsoft MVPs.
 - Pythian holds 7 Specializations under Oracle Platinum Partner program, including Oracle Exadata, Oracle GoldenGate & Oracle RAC
- **Global Reach & Scalability:**
 - Around the clock global remote support for DBA and consulting, systems administration, special projects or emergency response

Why this presentation?



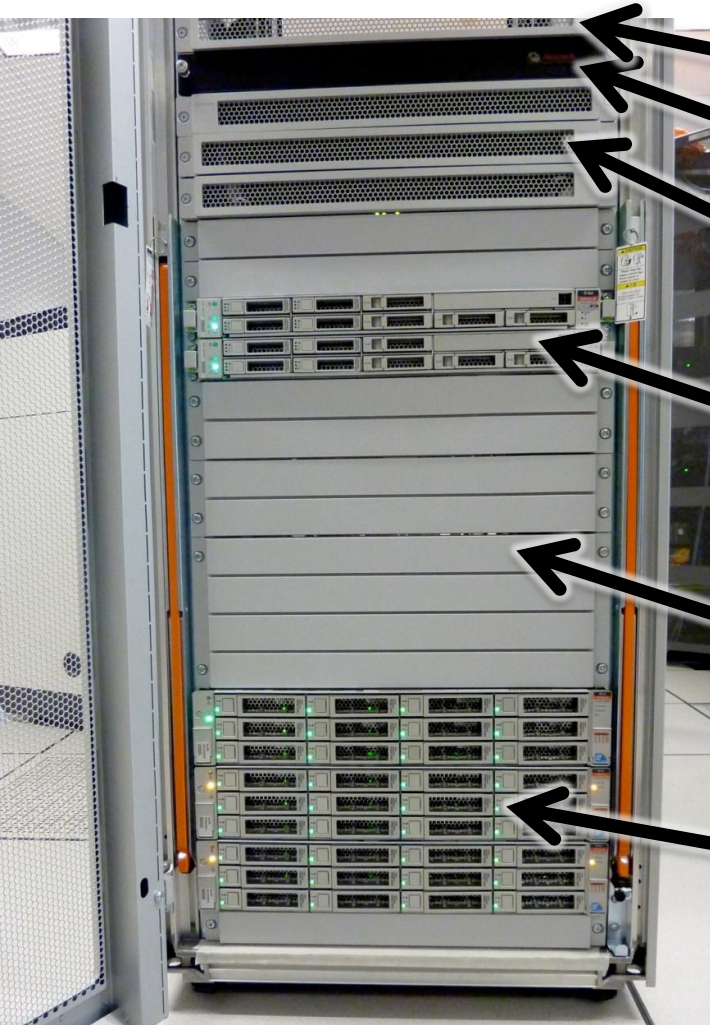
So you have Exadata. What next?

- Underpinning Exadata are well-known Oracle technologies
- SQL*Plus works just like on any other Oracle system
- RAC and Grid Infrastructure, 11gR2
- ASM is there
- Services, SCAN, etc no different than other Oracle RAC environments

So what *is* different?

- It's an engineered system: you run the same hardware as other Exadata customers
- Oracle development has tested on the exact same configuration
- It's a good thing: much fewer obscure OS issues, driver issues etc
- Single point of contact at Oracle support – hardware and software
- And yeah, there are are a few new concepts...

Major bits



Management switch

KVM

InfiniBand switches

Compute servers

Empty space

Storage servers

How you buy Exadata

	Full Rack	Half Rack	Quarter Rack	Eighth Rack
Database servers	8	4	2	2
Usable storage (High Capacity)	224 TB	252 TB	108 TB	54 TB

- X3-8: big servers – 80 cores each
- Multi-rack expansion
- Storage expansion
- High capacity/high performance

The bits you know already

- Database servers: Xeon-based servers running Oracle Linux 5
- 2, 4, 8- or more-node 11gR2 RAC cluster
- Clustered ASM instances
- Ordinary Ethernet connections to the outside

The secret sauce

- InfiniBand networking
 - Gives amazing latency: forget 1ms, try 1 μ s
 - Great capacity too: 40gbps
 - Used for RAC interconnect *and* storage access – replaces fiber channel
- Storage servers: also Linux servers
 - But they run Exadata-specific software
 - Take I/O requests from database servers and send to disk
 - Also have understanding of database concepts

Differences from SAN/NAS



Smart Scans (1)

- Somewhat analogous to parallel query
- Multiple parallel processes
- Communicate with shadow processes with SQL-type constructs – rows, columns etc
- Save communication capacity by only sending the data you need, not entire data blocks
- `Select col1 from big_table;`
- Sends only col1 data, for example
- Also more complex actions like predicate evaluation, decryption, bloom filter joins...

Smart Scans (2)

- Put processing closer to the storage
- Inherently parallel: processing happens on all storage servers that might have matching data
- Only apply to bulk, direct path operations – usually full table scans
- Excludes
 - Index range scans
 - Table access by index rowid
 - Blocks with uncommitted changes
 - LOBs

Storage Indexes

- Used in conjunction with smart scans
- Can avoid doing I/O entirely
- Split storage up into 1m “storage regions”
- Track high/low values
- Work best with pre-sorted data
- **Example:** `select * from tab where col1 < 20;`



Hybrid Columnar Compression

- New storage type with superb compression ratios: 10x or more
- Group blocks into compression units
- Store columns together instead of rows
- Typical databases have repeated values: the infamous “Y”/”N” column for example
- Then apply well-known compression like bzip2 or gzip on top

Hybrid Columnar Compression

- Much like smart scans, data must be loaded in bulk: CREATE TABLE AS SELECT, INSERT /*+APPEND*/, SQL*Loader direct path, etc
- Otherwise data will get OLTP compression instead
- CPU overhead to compress/decompress (though smart scans can offload decompression)
- Locks affect entire compression units rather than single rows

Recommendations: HCC

- Best for data that is inserted in large batches (data warehouses for example)
- Also useful with time-based partitions: re-compress old, no-longer-modified data
- There are 4 levels; ARCHIVE HIGH has significant CPU overhead so use for data that's rarely accessed
- Use DBMS_COMPRESSION to predict compression ratios, and see results afterwards

Flash cache

- Unlike the data warehouse-focused features we've covered, flash cache is designed for OLTP
- Flash memory has 5x better read latency than hard drives
- Transparent to applications
- Smart enough not to cache RMAN backup I/O, ASM secondary extents, and other non-repeatable reads
- DBAs can use resource manager and SQL to prioritize cache objects

Flash cache: recommendations

- With new Exadata X3, you may fit entire DB in flash. Enjoy!
- Entirely different from the similarly-named Database Smart Flash Cache in 11gR2
- You can put ASM disks on flash, but ASM redundancy reduces usable space
- You can use back-end commands to actually see what's being cached

I/O Resource Management

- You can use DBRM in Oracle 11g to manage CPU time, and manage runaway queries
- But what if your system is bottlenecked on I/O
- In Exadata, just turn on IORM and your CPU priorities apply to I/O too
- Can also allocate I/O capacity across databases

Management tools: what to expect

- Command-line tools all there: SQL*Plus etc
- For storage servers, a new tool: cellcli to manage storage servers
- cellcli can be combined with dcli to run a command on multiple machines.
- cellcli changes to one storage server aren't visible to others
- OEM 12c Cloud Control can now do most common tasks

Backups



Disaster recovery

- Storage-level replication and similar tools cannot be used with Exadata
- Data guard is the overwhelming favourite
- Active data guard works well with Exadata
- It's possible to have a non-Exadata standby database, but columnar compression is an issue – must ALTER TABLE MOVE before accessing data
- Most common approach: two datacenters with identically configured Exadata machines

Patching



Diagnostic tools

Oracle Exadata Assessment Report

System Health Score is 87 out of 100 [\(detail\)](#)

Cluster Summary

Cluster Name	2-cluster
OS Version	LINUX X86-64 OELRHHEL 5 2.6.32-400.1.1.el5uek
CRS Home - Version	/u01/app/11.2.0.3/grid - 11.2.0.3.0
DB Home - Version - Names	/u01/app/orclpdb/product/11.2.0.3/dbhome_1 - 11.2.0.3.0 - g
Exadata Version	11.2.3.2.0
Number of nodes	14
Database Servers	4
Storage Servers	7
IB Switches	3
exachk Version	2.2.0_20121109
Collection	exachk_SAMPLE_011113_134037.zip
Collection Date	11-Jan-2013 13:44:28

NOTE : exachk is only one part of the MAA Best Practices recommendation methodology. My Oracle Support "Oracle Exadata Best Practices (Doc [ID757552_1](#))" should be reviewed thoroughly as it is the driver for exachk and contains additional operational and diagnostic guidance that is not programmed within exachk.

Organizational support models



But overall

- As a DBA, you'll still be doing much the same work as you're used to on Exadata
- Performance tuning, backups, code deployments, disaster recovery all exist in Exadata land

Gaining Exadata skills

- Not many of us have datacenters in our basements to try out new technologies
- Get familiar with underlying technologies: RAC, and “real” ASM with redundancy, direct-attach disk
- But there are plenty of resources:
 - Exadata courses: Oracle university workshop, and others
 - Exadata books, particularly *Expert Oracle Exadata*
 - Conference presentations – even if you can’t attend, try and download the slides

Thanks and Q&A

Contact info



fielding@pythian.com



+1-877-PYTHIAN

To follow us



pythian.com/blog



@mfild
@pythian



linkedin.com/company/pythian