Exadata Performance, Yes You Still Need to Tune

Kathy Gibbs
Senior Database Administrator, CONFIO Software
Who Am I?

- Over 18 years in IT and 12+ Years in Oracle & SQL Server
  - DBA and Developer
  - Worked for various industries (Telecom, Retail, Finance)
  - Oracle, SQL Server, Sybase, DB2 on VMware
- Sr DBA for Confio Software
  - KathyGibbs@confio.com
  - Makers of Ignite8 Response Time Analysis Tools
  - IgniteVM for Oracle/SQL/Sybase/DB2 on Vmware
  - Alarm VM for VM Admins
What is Exadata
How Exadata Solves Performance Problems
How is Performance Tuning different on an Exadata Machine
How Exadata Can Create Performance Problems
Questions
First, the different ‘Exa’ Products

- **Exalogic and the Elastic Cloud**
  - Think ‘WebLogic’ this is the application server solution
  - The Elastic Cloud is factory assembled and installed
  - Exabus is the defining architectural feature. It is basically the I/O subsystem

- **Exalytics In-Memory Machine**
  - Contains ‘optimized’ TimesTen db
  - Includes BI Enterprise Edition and Essbase
  - The rumor is this is Oracle competing with SAP’s HANA hardware

- **Database Appliance**
  - This is just a preconfigured 11gr2 database on OEL ‘Fully Redundant Integrated Database Appliance in a single box’
  - RAC and RAC One Licensing is included in the price.
What is Exadata?

- It is a preconfigured combination of hardware and software that provides a platform for running Oracle Db.

- Since Exadata includes a storage subsystem, new software has been developed to run at the storage layer.
  - This has allowed the developers to do some things that are just not possible on other platforms.
History of Exadata

- In 2008 Oracle introduced its 1st ‘Database Machine’.
  - This Exadata V1 was based on HP hardware
  - This version was really marketed to big Data Warehouse shops.
- In 2009 they came out with Exadata V2 with Sun hardware, added Smart Flash Cache and now OLTP recommended
- In 2010 Oracle completed the acquisition of Sun and they came out with X2-2 and X2-8
What is Exadata – V2

8 Compute Servers
- 8 x 2 sockets x 4 cores = 64 cores
- 576 GB DRAM

InfiniBand Network
- 40 Gb/sec each direction
- Fault Tolerant

I/O Capacity and Performance
- 15K RPM 600GB SAS or 2TB SATA 7.2K RPM disks
- 14 Storage Servers
  - 14x12=168 Disks
  - 100T SAS or
  - 336T SATA
- 5TB+ flash storage!
What is Exadata – X2-2

8 Compute Servers
- 8 x 2 sockets x 6 cores = 96 cores
- 768 GB DRAM

InfiniBand Network
- 40 Gb/sec each direction
- Fault Tolerant

14 Storage Servers
- 14x12=168 Disks
- 100T SAS or
- 336T SATA

5TB+ flash storage!

I/O Capacity and Performance
15K RPM 600GB SAS (HP model – high performance) or 2TB SAS 7.2K RPM disks (HC model – high capacity)
Note that 2TB SAS are the same old 2 TB drives with new SAS electronics.
What is Exadata – X2-8

2 Compute Servers
- 8 x 2 sockets x 8 cores = 128 cores
- 2 TB DRAM

InfiniBand Network
- 40 Gb/sec each direction
- Fault Tolerant

14 Storage Servers
- 14x12=168 Disks
- 100T SAS or
- 336T SATA

Oracle Linux or Solaris x86

5TB+ flash storage!

I/O Capacity and Performance
15K RPM 600GB SAS (HP model – high performance) or 2TB SAS 7.2K RPM disks (HC model – high capacity)
Note that 2TB SAS are the same old 2 TB drives with new SAS electronics.
<table>
<thead>
<tr>
<th>Features</th>
<th>V1</th>
<th>V2</th>
<th>X2-2</th>
<th>X2-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database servers</td>
<td>8 x DL360 G5</td>
<td>8 x Sun Fire x4170 1U</td>
<td>8 x Sun Fire X4170 1U</td>
<td>2 x Sun Fire x4800 5U</td>
</tr>
<tr>
<td>Database CPUs</td>
<td>16</td>
<td>16 Quad-Core Intel® Xeon® E5430 Processors (2.66 GHz)</td>
<td>16 x Quad-Core Intel Xeon E5540</td>
<td>16 x Eight-Core Intel® Xeon® X7560 Processors (2.26 GHz)</td>
</tr>
<tr>
<td>Database cores</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>96</td>
</tr>
<tr>
<td>Database Threads</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>192</td>
</tr>
<tr>
<td>Database RAM</td>
<td>576GB</td>
<td>576GB</td>
<td>576GB</td>
<td>768GB</td>
</tr>
<tr>
<td>Storage Server</td>
<td>14 x DL180 G5</td>
<td>14 x SunFire X4275</td>
<td>14 x SunFire X4270</td>
<td>14 x SunFire X4270 M2</td>
</tr>
<tr>
<td>Storage cell CPUs</td>
<td>28 x Intel Quad-core processors</td>
<td>28 x Xeon E5540 quad core 2.53GHz</td>
<td>28 x Quad-Core Intel Xeon E5640 2.53GHz</td>
<td>28 x Quad-Core Intel Xeon E5540</td>
</tr>
<tr>
<td>Storage cells CPU cores</td>
<td>112</td>
<td>112</td>
<td>112</td>
<td>168</td>
</tr>
<tr>
<td>Storage Cell Threads</td>
<td>224</td>
<td>224</td>
<td>224</td>
<td>336</td>
</tr>
<tr>
<td>Storage RAM</td>
<td>112GB</td>
<td>336GB</td>
<td>336GB</td>
<td>336GB</td>
</tr>
<tr>
<td>Smart Flash Cache</td>
<td>5.3TB</td>
<td>5.3TB</td>
<td>5.3TB</td>
<td>5.3TB</td>
</tr>
<tr>
<td>Database Servers networking</td>
<td>4 x 1GbE x 8 servers = 32 x 1GbE</td>
<td>4 x 1GbE x 8 servers = 32 x 1GbE</td>
<td>4 x 1GbE x 8 servers + 2 x 10GbE x 8 servers = 32 x 1GbE + 16 x 10GbE</td>
<td>8 x 1GbE x 2 servers + 8 x 10GbE x 2 servers = 16 x 1GbE + 16 x 10GbE</td>
</tr>
<tr>
<td>InfiniBand Switches</td>
<td>4 x 24 port QDR 40Gbit/s switches (Total ports 96)</td>
<td>3 x 36 port QDR 40Gbit/s switches (Total ports 108)</td>
<td>3 x 36 port QDR 40Gbit/s switches (Total ports 108)</td>
<td>3 x 36 port QDR 40Gbit/s switches (Total ports 108)</td>
</tr>
<tr>
<td>InfiniBand ports on database servers (total)</td>
<td>2 ports x 8 servers = 16 ports</td>
<td>2 ports x 8 servers = 16 ports</td>
<td>2 ports x 8 servers = 16 ports</td>
<td>8 ports x 2 servers = 16 ports</td>
</tr>
<tr>
<td>Ethernet Switch</td>
<td>2 x 16port Switch</td>
<td>1 x 48-port Cisco Catalyst 4948</td>
<td>1 x 48-port Cisco Catalyst 4948</td>
<td>1 x 48-port Cisco Catalyst 4948</td>
</tr>
<tr>
<td>Database Servers OS</td>
<td>Oracle Linux 5 Update 3</td>
<td>Oracle Linux 5 Update 3</td>
<td>Oracle Linux 5 Update 5</td>
<td>Oracle Linux 5 Update 5</td>
</tr>
<tr>
<td>Multiple Rack Capability (with requiring additional switches)</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>PDU</td>
<td>2 redundant 15 kVA PDUs (single phase or three phase, high voltage or low voltage)</td>
<td>2 redundant 15 kVA PDUs (single phase or three phase, high voltage or low voltage)</td>
<td>2 redundant 15 kVA PDUs (single phase or three phase, high voltage or low voltage)</td>
<td>2 redundant 24 kVA PDUs (three phase, high voltage or low voltage)</td>
</tr>
</tbody>
</table>

Thanks to Oracle and [http://blog.vishalgupta.com/](http://blog.vishalgupta.com/)
## What is Exadata – Key Capabilities

### Exadata Model Comparison - Full Rack

<table>
<thead>
<tr>
<th>Key Capabilities</th>
<th>High Performance SAS Disks</th>
<th>High Capacity SATA Disk</th>
<th>High Performance SAS Disks</th>
<th>High Capacity SATA Disk</th>
<th>High Performance SAS Disks</th>
<th>High Capacity SATA Disk</th>
<th>High Performance SAS Disk</th>
<th>High Capacity SAS Disk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard Disk Type</td>
<td>450GB 15k RPM SAS</td>
<td>1TB 10k RPM SATA</td>
<td>600GB 15k RPM SAS</td>
<td>2TB 7.2k RPM SATA</td>
<td>600GB 15k RPM SAS</td>
<td>2TB 7.2k RPM SATA</td>
<td>600GB 15k RPM SAS</td>
<td>2TB 7.2k RPM SATA</td>
</tr>
<tr>
<td>Uncompressed Flash data bandwidth</td>
<td>50 GB/sec</td>
<td>50 GB/sec</td>
<td>50 GB/sec</td>
<td>50 GB/sec</td>
<td>50 GB/sec</td>
<td>50 GB/sec</td>
<td>50 GB/sec</td>
<td>50 GB/sec</td>
</tr>
<tr>
<td>Disk IOPS</td>
<td>50,000 IOPS</td>
<td>25,000 IOPS</td>
<td>50,000 IOPS</td>
<td>25,000 IOPS</td>
<td>50,000 IOPS</td>
<td>25,000 IOPS</td>
<td>50,000 IOPS</td>
<td>25,000 IOPS</td>
</tr>
<tr>
<td>Flash IOPS</td>
<td>1,000,000 IOPS</td>
<td>1,000,000 IOPS</td>
<td>1,000,000 IOPS</td>
<td>1,000,000 IOPS</td>
<td>1,000,000 IOPS</td>
<td>1,000,000 IOPS</td>
<td>1,000,000 IOPS</td>
<td>1,000,000 IOPS</td>
</tr>
<tr>
<td>Raw disk data capacity</td>
<td>100 TB</td>
<td>336 TB</td>
<td>100 TB</td>
<td>336 TB</td>
<td>100 TB</td>
<td>336 TB</td>
<td>100 TB</td>
<td>336 TB</td>
</tr>
<tr>
<td>Uncompressed user data</td>
<td>28 TB</td>
<td>100 TB</td>
<td>28 TB</td>
<td>100 TB</td>
<td>28 TB</td>
<td>100 TB</td>
<td>28 TB</td>
<td>100 TB</td>
</tr>
<tr>
<td>Smart Flash Cache</td>
<td>5.3TB</td>
<td>5.3TB</td>
<td>5.3TB</td>
<td>5.3TB</td>
<td>5.3TB</td>
<td>5.3TB</td>
<td>5.3TB</td>
<td>5.3TB</td>
</tr>
<tr>
<td>Data Load Rate</td>
<td>5 TB/hour</td>
<td>5 TB/hour</td>
<td>5 TB/hour</td>
<td>5 TB/hour</td>
<td>5 TB/hour</td>
<td>5 TB/hour</td>
<td>5 TB/hour</td>
<td>5 TB/hour</td>
</tr>
</tbody>
</table>

Thanks to Oracle and [http://blog.vishalgupta.com/](http://blog.vishalgupta.com/)
## Exadata Model Comparison - Full Rack

<table>
<thead>
<tr>
<th></th>
<th>8 x DL360 G5</th>
<th>8 x Sun Fire x4170 1U</th>
<th>8 x Sun Fire X4170 1U</th>
<th>8 x Sun Fire X4170 M2 1U</th>
<th>2 x Sun Fire X4800 5U</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Database Server</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPU</td>
<td>2 x Intel Quad-core processors</td>
<td>2 x Quad-Core Intel Xeon E5540 2.53GHz</td>
<td>2 x Quad-Core Intel Xeon E5540 2.53GHz</td>
<td>2 x Six-Core Intel Xeon X5670 2.93GHz</td>
<td>8 x Eight-Core Intel® Xeon® X7560 Processors (2.26 GHz)</td>
</tr>
<tr>
<td>Memory</td>
<td></td>
<td>32 GB</td>
<td>72 GB</td>
<td>72 GB</td>
<td>96 GB</td>
</tr>
<tr>
<td>Disk Controller</td>
<td></td>
<td>HBA with 512MB Battery Backed Write Cache</td>
<td>HBA with 512MB Battery Backed Write Cache</td>
<td>HBA with 512MB Battery Backed Write Cache</td>
<td></td>
</tr>
<tr>
<td>Local Disks</td>
<td>4 x 146 GB SAS disks</td>
<td>4 x 146GB 10K RPM SAS Disks</td>
<td>4 x 146 GB 10K RPM SAS</td>
<td>4 x 300 GB 10K RPM SAS</td>
<td>8 x 300GB 10K RPM SAS Disks</td>
</tr>
<tr>
<td>Local Storage</td>
<td>292GB (RAID1)</td>
<td>292GB (RAID1)</td>
<td>292GB (RAID1)</td>
<td>600GB (RAID1)</td>
<td>1.2 TB (RAID1)</td>
</tr>
<tr>
<td>Infiniband Ports</td>
<td></td>
<td>2 x QDR (40Gb/s) Ports</td>
<td>2 x QDR (40Gb/s) Ports</td>
<td></td>
<td>4 x Dual-port 4X QDR PCIe 2.0 (40Gb/s)</td>
</tr>
<tr>
<td>Ethernet Ports</td>
<td></td>
<td>4 Embedded Gigabit Ethernet Ports</td>
<td>4x1Gb</td>
<td>4x1Gb + 2x10Gb (Intel 82599 Controller)</td>
<td>8x1GbE and 8x10GbE using SFP+ connectors (Intel 82599 Controller)</td>
</tr>
<tr>
<td>ILOM Ethernet Port</td>
<td></td>
<td>1 Ethernet port (iLO2 with Advanced Pack)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Power Supplies</td>
<td></td>
<td>2 x Redundant Hot-Swappable</td>
<td>2 x Redundant Hot-Swappable</td>
<td>4 x Redundant Hot-Swappable</td>
<td></td>
</tr>
</tbody>
</table>

Thanks to Oracle and http://blog.vishalgupta.com/
### Exadata Model Comparison - Full Rack

<table>
<thead>
<tr>
<th>Storage Server</th>
<th>14 x DL180 G5</th>
<th>14 x SunFire X4275 2U</th>
<th>14 x SunFire X4270 M2 2U</th>
<th>14 x SunFire X4275 2U</th>
<th>14 x SunFire X4270 M2 2U</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>2 x Intel Quad-core processors</td>
<td>2 x Quad-Core Intel Xeon E5540 2.53GHz</td>
<td>2 x Quad-Core Intel Xeon E5540 2.53GHz</td>
<td>2 x Six-Core Intel Xeon E5640 2.26GHz</td>
<td>2 x Quad-Core Intel Xeon E5540 2.53GHz</td>
</tr>
<tr>
<td>Memory</td>
<td>8GB</td>
<td>24 GB</td>
<td>24 GB</td>
<td>24 GB</td>
<td>24 GB</td>
</tr>
<tr>
<td>Flash Card</td>
<td>4 x 96 GB Sun Flash Accelerator F20 PCIe Cards</td>
<td>4 x 96 GB Sun Flash Accelerator F20 PCIe Cards</td>
<td>4 x 96 GB Sun Flash Accelerator F20 PCIe Cards</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smart Flash Cache</td>
<td>384 GB</td>
<td>384 GB</td>
<td>384 GB</td>
<td>384 GB</td>
<td>384 GB</td>
</tr>
<tr>
<td>Local Disks Count</td>
<td>12</td>
<td>12</td>
<td>12 (Non M2 model has 2TB SATA disks)</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Disk Controller</td>
<td>HBA with 512MB Battery Backed Write Cache</td>
<td>HBA with 512MB Battery Backed Write Cache</td>
<td>HBA with 512MB Battery Backed Write Cache</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infiniband Ports</td>
<td>Dual-port 4X QDR (40Gb/s)</td>
<td>Dual-port 4X QDR (40Gb/s)</td>
<td>Dual-port 4X QDR (40Gb/s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethernet Ports</td>
<td>1 Embedded Gigabit Ethernet Port</td>
<td>1 Embedded Gigabit Ethernet Port</td>
<td>1 Embedded Gigabit Ethernet Port</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ILOM Ethernet Port</td>
<td>1 Ethernet port (LO100c)</td>
<td>1 Ethernet port (LO100c)</td>
<td>1 Ethernet port (LO100c)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power Supplies</td>
<td>2 x Redundant Hot-Swappable</td>
<td>2 x Redundant Hot-Swappable</td>
<td>2 x Redundant Hot-Swappable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OS</td>
<td>OEL 5.5</td>
<td>OEL 5.5</td>
<td>OEL 5.5</td>
<td>OEL 5.5</td>
<td>OEL 5.5</td>
</tr>
</tbody>
</table>

**Thanks to Oracle and [http://blog.vishalgupta.com/](http://blog.vishalgupta.com/)**
# What is Exadata – Environmental Specifications

<table>
<thead>
<tr>
<th>Environmental Specifications</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
<td>42U, 78.66” - 1998 mm</td>
<td>42U, 78.66” - 1998 mm</td>
</tr>
<tr>
<td>Width</td>
<td>23.62” (600mm)</td>
<td>23.62” – 600 mm</td>
</tr>
<tr>
<td>Depth</td>
<td>47.24” – 1200 mm</td>
<td>47.24” – 1200 mm</td>
</tr>
<tr>
<td>Weight</td>
<td>2171 lbs (986.8 kg)</td>
<td>2,131 lbs. (966.6 kg)</td>
</tr>
<tr>
<td>Power - Maximum power usage</td>
<td>13.2 kW (13.6 kVA)</td>
<td>14.0 kW (14.3 kVA)</td>
</tr>
<tr>
<td>Power - Typical power usage</td>
<td>9.6 kW (9.9 kVA)</td>
<td>9.8 kW (10.0 kVA)</td>
</tr>
<tr>
<td>Cooling - At max usage</td>
<td>44,800 BTU/hr</td>
<td>47,800 BTU/hour (50,400 kJ/hour)</td>
</tr>
<tr>
<td>Cooling - At typical usage</td>
<td>32,800 BTU/hr</td>
<td>33,400 BTU/hour (35,300 kJ/hour)</td>
</tr>
<tr>
<td>Airflow - At max usage (front-to-back)</td>
<td>1680 CFM</td>
<td>2,200 CFM</td>
</tr>
<tr>
<td>Airflow - At typical usage (front-to-back)</td>
<td>950 CFM</td>
<td>1,560 CFM</td>
</tr>
<tr>
<td>Operating temperature</td>
<td>41° to 95° F (5° to 35° C) at sea level</td>
<td>5 °C to 32 °C (41 °F to 89.6 °F)</td>
</tr>
<tr>
<td>Operating humidity</td>
<td>10% to 90% relative humidity</td>
<td>10% to 90% relative humidity</td>
</tr>
</tbody>
</table>

Thanks to Oracle and [http://blog.vishalgupta.com/](http://blog.vishalgupta.com/)
Exadata Performance

- Cell Offloading
  - Smart Scan
  - Storage indexes
  - Hybrid Columnar Compression
- DBRM/IORM
- Smart flash cache
- Additional Views
Offloading is the generic term to describe what is now done at the storage tier that would have been done at the db tier.

- Smart Scans are the access mechanism used to offload the tasks
- DataFile Initialization – Through ASM speeds up initialization
- RMAN Offload - When Block Change Tracking is used with RMAN and Exadata the Cell does the incremental backups at a granularity of the individual block, rather than at the granularity of a group of blocks as is done without Exadata.
- HCC – will talk to in a later slide.
Why Exadata is so Efficient – Smart Scan

Select a.account_num, c.customer_name
From all_accounts a, all_customers c
Where a.acct_id = c.acct_id
And state = ‘PA’

10 TB of user data
Requires 10 TB of IO

1 TB with compression

100 GB with partition pruning

20 GB with Storage Indexes

5 GB with Smart Scans

Sub second
On Database Machine
Two Main Wait Events

- Cell Smart Table Scan
- Cell Smart Index Scan

The flow of data from Smart Scan can’t be buffered in SGA buffer pool. (Think PGA (heap))
Storage Index

- A storage index is an in-memory structure that holds some information about the data inside specified regions of physical storage.
- More importantly, it knows what is NOT located in that region.
- Think of this feature as a pre-filter.
Exadata Performance – Storage Index

- Storage indexes actually work in the negative, find where it doesn’t exist and therefore eliminate cells needed to be looked at.
- To use them, the queries must use smart scans which means not all queries will benefit. Typically they are utilized for queries using predicates, full table scans or fast fill scans of indexes.
This technology utilizes a combination of both row and columnar methods for storing data.

A logical construct called the compression unit is used to store a set of HCC rows. When data is loaded, column values for a set of rows are grouped together and compressed. After the column data for a set of rows has been compressed, it is stored in a compression unit.
To maximize loading, use below ‘DW’ options. However you can also ‘regular’ DML

- Insert statements with the APPEND hint
- Parallel DML
- Direct Path SQL*LDR
- CTAS

Types

- Query Low
- Query High
- Archive Low
DBRM/IORM

- **DBRM – Database Resource Monitor**
  - This is the resource manager you are familiar with.
  - Without all sessions are given equal priority.
  - The main reason is for use with consolidation.

- **IORM - I/O Resource Monitor**
  - Added with V2 and beyond.
  - Can prioritize I/O across dbs.
  - For first time can virtually guarantee I/O service levels within and among dbs.
DBRM/IORM

- **DBRM**
  - CPU Quantum – wait event. Resmgr: cpu quantum. Is the unit of CPU time that the DBRM uses for allocating CPU to consumer groups. Occurs when DBRM is actively throttling Cpu Consumpton.
  - Check DBRM Metrics V$RSRC_Consumer_group. Also v_$RSRCMGRMETRIC and V_$RSRCMGRMETRIC_HISTORY for monitoring effect of DBRM resource allocations have on sessions
  - Instance Caging – Provisions CPU at db instance level.
DBRM/IORM

- IORM
  - Interdatabase IORM – Manages priority among multiple dbs by db name
  - IORM Categories – This is a new attribute. Still by dbname. Ex oltp_category batch_category
  - Intradatabase IORM – On exadata when a DBRM plan is activated the db transmits a desc of this plan to all cells in the storage grid. So in a way this is a bit of the ‘default’
DBRM/IORM

- IORM
  - IORM manages at the storage cell
  - IORM distinguishes between small and large io request (<128k in size or > 128K)
  - For each cell disk cellsrv maintains an IORM queue for each consumer group and each background process. For each db accessing the cell.
    Cellcli> alter iormplan objective = low_latency
    Cellcli> list iormplan attributes object
    low_latency
Exadata Performance

DBRM/IORM

Or SQLPlus
This feature is what allowed V2 to be good for OLTP databases. 5TB on a full rack.
- ~ 366 G usable flash storage per storage cell

Can be used in two ways (or both)
- Configured as cache - recommended
- Carved out as solid state disks for ASM

Write through cache. Writes bypass the cache and go directly to disk. However can copy data into cache if likely to be used again
Exadata Performance – Smart Flash Cache

- Follow by querying v$sysstat (and related v$views)
  - Cell flash cache read hits

- CELL_FLASH_CACHE=KEEP
  - None, default, keep
  - Alter table oltp.busy_table (cell_flash_cache=keep);

- Cellcli  CELL_FLASH_CACHE=KEEP
  - Create flashcache all size=300g
  - List flashcache detail
  - List celldisk attributes name, diskType, size where name like ‘FD.*’
  - List metriccurrent where objecttype = ‘FLASHCACHE’
  - List flashcachecontent – shows objects in cache
Follow by querying `v$sysstat` (and related `v$views`)

- Can be overestimated if all in ‘KEEP’ state
- Cell flash cache read hits

```
SELECT `cell single + multiblock reads` c1, c2, c3, c5, c6,
     C6/decode(nvl(c2,0),0,1,c2) hit_ratio
FROM(
    SELECT sum(total_waits) c2,
            Avg(value) c6,
            Sum(time_waited /100) c3
    FROM `v_$system_event`, `v$sysstat` ss
    WHERE event in ('cell single block physical read', 'cell multiblock physical read')
        AND name like 'cell flash cache read hits'
        AND event not like '%Idle%')
ORDER BY c3
```
What to look for when tuning

- Sql statement response Time Monitoring
- DB Layer utilization and efficiency
- Storage Cell layer utilization and efficiency
- Advanced metrics and monitoring for Exadata
- Will be looking at
  - Sql statements v$views
  - Cellcli
  - OSWatcher
Exadata Performance

Top Waits | OEM1P1_DUB1DB01.CARDINALHEALTH.NET | November 6, 2011 to November 12, 2011

- cell single block physical read
- Backup: MML write backup piece
- CPU
- krf: async disk IO
- gc current request
- gc request
- library cache pin
- cell list of blocks physical read
- cell multiblock physical read
- log file sync
- Disk file operations I/O
- enq: TT - contention
- gc buffer busy release
- control file parallel write
- library cache lock
Actually none that only exist on exadata, built into db code. No time is allocated on other platforms for Exadata specific

Wont list all here.

Select event operation, count(*) from ( 
select sql_id, event, sql_plan_operation||''||sql_plan_options operation 
from DBA_HIST_ACTIVE_SESS_HISTORY 
where event like 'cell %' 
group by operation, event 
order by 1,2,3

Cell smart table scan event - see offloading

Cell smart index scan – fast full index scans that are offloaded
Exadata Performance
Exadata Causes Problems?

- This is not a reasoning not to get Exadata, just a note that Exadata is not the end all be all.
- Oracle states to drop all indexes, if you follow this you could have problems.
- If you have a very finely tuned query especially OLTP you may not see any benefit on this query.
- This is the exception but everyone I have talked to has that 1 query.
Conclusion

- At the root, Exadata houses an Oracle db.
- Always start with Cell Offloading
- Besides understanding roles, especially around patching, our biggest problems came from the upgrade to 11gr2
- Be aware, Exadata is not the be all end all, testing must be done and tuning must be done
- DBRM/IORM and Flash cache are your friends, don’t be afraid to use them even if you have had issues in the past
Confio Software

- Award Winning Performance Tools
- Ignite8 for Oracle, SQL Server, DB2, Sybase
- IgniteVM for Databases on VMware
  - Download at www.confio.com

Provides Answers for
- What changed recently that affected end users
- What layer (VM or DB) is causing the problem
- Who and How should we fix the problem

Download free trial at
www.confio.com