



Official Publication of the Northern California Oracle Users Group

# NoCOUG

J O U R N A L

Vol. 25, No. 2 • MAY 2011

\$15

# Gateway to Knowledge

## Interview

*Real-World Performance—An interview with Greg Rahn of Oracle.*

*See page 4.*

## Ask the Oracle ACEs!

*Why is my database slow?*

*See page 8.*

## Book Review

*Achieving Extreme Performance with Oracle Exadata by Rick Greenwald et al.*

*See page 13.*

*Much more inside . . .*

# Achieving Extreme Performance with Oracle Exadata

A Book Review by Brian Hitchcock

## Details

**Author:** Rick Greenwald, Robert Stackowiak, Maqsood Alam, Mans Bhuller

**ISBN:** 978-0-07-175259-6

**Pages:** 432

**Publication Date:** February 2011

**Price:** \$60

**Publisher:** McGraw-Hill



## Summary

**Overall review:** A very good introduction to Exadata.

**Target Audience:** Anyone with an Oracle database looking for performance improvement or database consolidation.

**Would you recommend this to others?** Yes.

**Who will get the most from this book?** Anyone who needs to understand what is new with Exadata.

**Is this book platform specific?** Yes—but that doesn't mean what it used to mean.

**Why did I obtain this book?** NoCOUG asked me to review this book.

## Overall Review

Exadata is both a new thing and a return to the past. Dedicated hardware preconfigured to run specific vendor software—sounds like old times, doesn't it? But there are new twists. The use of large amounts of flash memory and clever software makes Exadata a very new thing. While the familiar Oracle database runs on Exadata, the way the software and hardware work together presents many new concepts and terms that need to be understood. Much about Exadata will be familiar to Oracle DBAs, but there is much that is totally new. This book covers the new concepts very well. While there aren't any case studies or metrics from real-world implementations, the material presented is well worth your time as you prepare for the new world of Exadata.

Before we go any further, we need to define some terms. The overall product is the Oracle Exadata Database Machine. This product contains multiple database server nodes, multiple storage server cells, and networking gear. The product also contains Exadata Storage Server Software that is distinct and separate from the Oracle database software. Taking the full rack Exadata configuration as an example, it contains eight database server nodes and 14 storage server cells. The terms “database node” (server) and “storage server” (cell) will come up often.

## Chapter Reviews

- Introduction
- Chapter 1: Oracle and Tightly Integrated Hardware and Software Platforms
- Chapter 2: Oracle 11g Enterprise Edition Features
- Chapter 3: Exadata Software Features
- Chapter 4: Oracle Exadata Database Machine Platform Hardware Components
- Chapter 5: Managing the Exadata Database Machine
- Chapter 6: High Availability and Backup Strategies
- Chapter 7: Deploying Data Warehouses on the Oracle Exadata Database Machine
- Chapter 8: Exadata and OLTP
- Chapter 9: Consolidating Databases with the Oracle Exadata Database Machine
- Chapter 10: Migrating to the Exadata Database Machine

## Introduction

While the introduction is very short, it makes several interesting points. First, that this is the first comprehensive book on the Oracle Exadata Database Machine. Second, that Exadata is all about simplicity.

*Dedicated hardware preconfigured to run specific vendor software—sounds like old times, doesn't it? But there are new twists. The use of large amounts of flash memory and clever software makes Exadata a very new thing.*

And third, that Exadata “just works.” These are interesting statements, and we will see how the rest of the book explains each of them.

## Chapter 1: Oracle and Tightly Integrated Hardware and Software Platforms

We start with a brief history lesson. The evolution of Oracle database machines started in 2008. Perhaps that isn't a long enough period of time to really be called history, but there have been versions of this product from HP, Sun, and now Oracle and Sun. On the other hand, a timeline is shown for Oracle database development from 1977 that shows how this development has been coming for a long time. This also signals quite a change in focus for Oracle. Oracle has always been the database that runs on almost all combinations of server and storage platforms. With Exadata, Oracle is the database that runs on only one. This could be seen as “back to the future.” This turn of events follows the endless increase in the volume of data that businesses need to process. Business also wants to access more and more data in less time to support more sophisticated data analysis. This pushes traditional database server and storage platforms beyond their limits.

Exadata addresses the need for a large move forward in database server performance to satisfy current demands. In addition to a change to dedicated hardware, Exadata also signals a change in the way customers build their systems. Up to now, each customer needed to have access to various personnel to decide what server and storage platforms they should select. Exadata eliminates this. The server hardware, software, and storage are defined for you. We are told that this eliminates risk for the customer. It also simplifies support for Oracle by greatly reducing the number of combinations of hardware and software that customers can have. Some will see this as a lack of choice, a lack of control, and for those who are employed to make these system design decisions, they may see it as a head-count reduction. The integration between the Exadata hardware and software and the Oracle database software is described. The addition of Smart Flash Cache and various new indexes (too choose a few at random) illustrates the many new things that Exadata brings to the table. Further, we are told that applications don't have to be certified for Exadata. Therefore, any application should run on Exadata without modification. I will call this transparency; i.e., moving to Exadata is transparent to your application. Your application users won't notice any difference. We will watch as this concept is further refined throughout the rest of the book. The first refinement is that the database for the application must be 11gR2. The impact of Exadata on the traditional DBA role is discussed. We are assured that all existing DBAs will simply be assigned to more interesting work. Finally, we have some thoughts on where we are headed next. Most (if not all) applications have complex middle tiers that also require ever-increasing levels of performance. The Exalogic platform will do for WebLogic and Java-based applications what Exadata is doing for databases. If you don't know what WebLogic is, consider this a wake-up call. Oracle's focus is moving quickly to the middle tier, and the future there is WebLogic.

## Chapter 2: Oracle 11g Enterprise Edition Features

This chapter describes many of the features of the Oracle 11g Enterprise Edition database. This information is useful for readers who are approaching Oracle for the first time, but I think most experienced DBAs will not need 65 pages describing 11g features. For example, transactions and locks are explained. Having said this, there are pieces of useful information here and there, and this chapter does explain how each of the major features of 11g is enhanced by Exadata. We find that RAC is not a requirement for using Exadata. This means that your application doesn't have to be running on RAC before it can move to Exadata, but we are told that RAC and Exadata are

*The way the software and hardware work together presents many new concepts and terms that need to be understood. Much about Exadata will be familiar to Oracle DBAs, but there is much that is totally new.*

made for each other. You probably want to use RAC for your application to get the most from the Exadata hardware and software. Note how the transparency is not as complete as we first thought. Our application must be running on 11gR2 database, and it should be set up for RAC.

Next is ASM. Exadata uses ASM exclusively. An existing application that doesn't use ASM can move to Exadata, but you will need to learn how to manage ASM. Another interesting bit of trivia is the Fast Recovery Area. (Yes, it has changed its name. It was the Flash Recovery Area.) The future is full of subtle changes. Partitioning is explained next, and we are told that Exadata really performs best when applications use partitioning. Parallel execution is described in great detail, as is Data Guard.

Again, this information isn't bad, but for most readers, I think this chapter should be skimmed for the parts that relate specifically to Exadata.

## Chapter 3: Exadata Software Features

This chapter is the best part of the whole book. This is where we learn what is really new and different about Exadata. For all the interest in the Exadata hardware, it is the software that really offers some amazing performance improvements. Assuming your application can make use of all of them, the Exadata software really is a significant step forward. If you take away only one thing from this book (and/or this book review), it should be this: Exadata is hardware and software. And this software is separate from the Oracle database software. There is a lot of material here, and I have chosen to briefly describe each of the most significant software features. Again, this software is not the 11gR2 software. This software is called the Exadata Storage Server Software.

**Smart Scan:** During normal query processing, the database instance will examine all the blocks of a table that contain relevant data, and then eliminate data based on the query

predicate. Smart Scan processes the blocks at the storage level, before they are returned to the database instance, and eliminates those that aren't needed to satisfy the query. Smart Scan also eliminates columns that aren't needed. All of this will dramatically reduce the amount of data that is returned to the database instance. Smart Scan also reduces the amount of data returned for each table that will be used in a join using Join Filtering. Using a Bloom filter to determine which rows will be needed in each table that will be joined; rows that won't be joined are eliminated, further reducing the amount of data that the database instance will have to deal with.

*Exadata provides significant performance gains but will require significant organizational changes as well.*

Note that the effects of Smart Scan can be seen in the explain plan that now displays new operations, such as TABLE ACCESS STORAGE FULL.

**Encryption:** Smart Scan can process encrypted data directly without the processing overhead of decryption.

**Hybrid Columnar Compression:** As mentioned earlier, the amount of data to be processed is always increasing. Compression reduces the amount of storage required, as well as reducing the I/O bandwidth. Exadata can store table data by columns. Only the columns needed for the query will be processed. The table is stored as a set of Compression Units, and each unit stores the data for one or more columns for a set of rows.

While the compression is done by the database instance, decompression can be done at the storage level.

**Storage Indexes:** These indexes track the high and low values for columns in the rows stored in 1Mb storage regions. The Exadata Storage Server Software can use these indexes to eliminate storage regions that aren't needed for a query. This, in my opinion, is similar to partition pruning in the database.

**Smart Flash Cache:** Exadata has large amounts of flash memory storage. To make the best use of this flash storage, Exadata provides software that will decide what to keep in flash cache. Data that it determines will be used again is cached.

Each of these software features has been described very briefly. Detailed descriptions of each are provided in this chapter. Also note that for each feature, there are conditions under which the feature won't apply or won't be used by Exadata. There are queries where some or all of these features won't be used or won't improve performance.

## **Chapter 4: Oracle Exadata Database Machine Platform Hardware Components**

This chapter discusses the Exadata hardware in detail. The need for balance between the CPUs' ability to process data and the storage system's ability to supply that data is examined. This issue led to the fixed ratios seen between the database nodes and the storage server cells in the Exadata systems. It is interesting to read about how flash memory can wear out and how the Exadata hardware is designed to work around this by supplying 32Gb of flash storage for each 24Gb of flash cache that appears to be available. The evolution of the database server nodes and the storage server cells is shown for each of the iterations of the Exadata machine. And we are told that sizing the Database Machine should be done only by consulting Oracle. The CPU and storage capacity, as well as the physical dimensions of the various Exadata machines, are also presented.

## **Chapter 5: Managing the Exadata Database Machine**

Now we learn about the tasks you will need to do after your Exadata machine is first installed. This includes creating databases and migrating databases from other systems. An interesting note appears, pointing out that the focus here is not on the administration of the database servers. We are told that the database servers are managed similarly to a regular 11gR2 database utilizing RAC and ASM. This implies that if you want to move your database(s) to Exadata, they must be on RAC and ASM, in addition to being on 11gR2. Details of how Exadata formats the disks (LUNs, cell disks, grid disks) and presents them to ASM are covered, as well as detailed options for flash cache. We also see a new command line utility, CellCLI, which is used to configure and manage the Exadata Storage Server Software. You now have a new utility to learn. Further, we are told that while we will have root access to the storage servers, performing any modifications to the storage servers is prohibited.

The practical impacts of all of this are amusing to think about. Will your Exadata machine be monitored or inspected? What will happen to those who violate this rule? It turns out that CellCLI works only within each of the multiple storage cells (servers) in the Exadata machine. To manage multiple cells, we use the dcli utility. Many more things are covered, including the overall Exadata setup process and security that allows creation of realms of disks for separate databases and monitoring.

## **Chapter 6: High Availability and Backup Strategies**

We start out covering Data Guard. Exadata Hybrid Columnar Compression can cause problems if your standby server isn't also an Exadata machine. If the primary stores data using

*Exadata flash memory capacity is designed to hold about 18% of capacity of the storage servers. This ratio has been found to represent the "working set," which is the fraction of the data in a database that is used by an application at any given time.*

EHCC, this data will be delivered to the standby in the compressed form. If you were to fail over, you would need to uncompress this data on the standby. We are told that the best practice for EHCC and Data Guard is to have an Exadata machine for both the primary and the standby. The use of Golden Gate log-based replication is also discussed.

Next is patching. Note that Oracle supplies all patches for all aspects of Exadata software. Users are prohibited from applying OS patches directly, for example. Note that this means the patching process for Exadata will be separate from your patching process for your non-Exadata systems. Backup and recovery are covered. It turns out that you don't back up the OS for the storage servers, and, if needed, you will restore from the supplied CELLBOOT USB flash drive. Who keeps this flash drive? Lots of new procedures to implement. RMAN is the only way to make database backups on Exadata. Again, if you aren't using RMAN now, you need to be ready to support it for Exadata. Oddly, I found the description of RMAN image copies and backup sets to be the best I've run across.

### **Chapter 7: Deploying Data Warehouses on the Oracle Exadata Database Machine**

A review of data warehouse schema (stars, snowflakes) options and history precedes a discussion of how the features of Exadata help with data warehouse processing. We learn that there are actually more CPUs in the storage servers than in the database nodes, and that many math and analytic functions are processed within the storage servers. This is another ex-

ample of how Exadata (hardware and software) breaks up data processing tasks to improve performance.

### **Chapter 8: Exadata and OLTP**

Similar to the previous chapter, here we learn how Exadata features help with OLTP processing. Not surprisingly, Exadata Smart Flash Cache has the most impact on OLTP. I didn't

*While you don't have to use RAC on Exadata, the whole point of Exadata is having many database and storage nodes to share workload.*

know that the Exadata flash memory capacity is designed to hold about 18% of capacity of the storage servers. This ratio has been found to represent the "working set," which is the fraction of the data in a database that is used by an application at any given time. This is how Oracle decided how much flash memory to supply, versus physical disk storage. The flash memory in the Exadata server is also protected against power supply issues that could result in a loss of data by an integrated supercapacitor! Not just a really big capacitor, but a supercapacitor! Perhaps it will take your Delorean back in time?

### **Chapter 9: Consolidating Databases with the Oracle Exadata Database Machine**

Much of the marketing around Exadata emphasizes its sheer performance.

But it also has a lot of capacity. While many customers may have huge databases that can fill one or more Exadata racks, many customers may have many smaller databases that could be consolidated into a single Exadata machine. This chapter tells us that we can consolidate OLTP and data warehouse databases into a single Exadata machine but we have to be careful. Also, if your existing application is CPU bound, consolidating into Exadata needs to be done with care. Specifically, we are told that the CPU sizing process is critical when consolidating CPU-bound applications. Memory and I/O sizing are covered, as well as how to isolate databases from each other for security or performance reasons. While reading this book, I wasn't clear on whether RAC was a requirement of Exadata. In other words, I didn't know if my existing application had to be running on RAC to be ready to move to Exadata. In this chapter, this is discussed. The point is that while you don't have to use RAC on Exadata, the whole point of Exadata is having many database and storage nodes to share workload. If your application uses a single database node in the Exadata machine, you may not be realizing the full potential offered by the Exadata hardware and software.

### **Chapter 10: Migrating to the Exadata Database Machine**

This chapter covers the practical issues around actually migrating Oracle and non-Oracle databases to Exadata—things like how to gather metrics for the existing databases,

*(continued on page 18)*

**ROLTA**

**TISC**

**Enterprise IT Solutions**  
*The Value of Experience...*

- ✓ Oracle (Applications & Technology)
- ✓ IT Infrastructure
- ✓ Profitability & Cost Mgmt
- ✓ Rolta Software Solutions

**ORACLE** Platinum Partner

***If you are using snapshots for backups, Exadata doesn't support anything like this. You can't connect a third-party storage system to Exadata to support snapshots. You must transition to RMAN for backups.***

(continued from page 16)

documenting any data loads and extracts, how to take into account the various compression options offered by Exadata when sizing disk capacity, and so on. I found some interesting things here. We are told that the process of migrating to Exadata is no different from migrating to a regular 11gR2 database on Linux or Solaris configured with RAC and ASM. Depending on what your current configuration is, this migration could be very complex whether Exadata is involved or not. If you are using snapshots for backups, Exadata doesn't support anything like this. You can't connect a third-party storage system to Exadata to support snapshots. You must transition to RMAN for backups. When sizing the new environment, you should size for performance and then for capacity. Sizing for performance is harder, and many times only the sizing for capacity is done. It is easy to calculate the disk space of your existing system. However, even this is not as easy as it seems, since you can compress your data on Exadata and you may need a lot less disk space than you have now.

### **Conclusion**

This book is very good at presenting the Exadata Database Machine. Assuming you are familiar with recent versions of

the Oracle database, I would suggest you focus on chapters 3, 4, and 5. These contain most of the discussion of what is really new about Exadata and how it works. There are two topics that I think need more coverage for those considering a move to Exadata. First is what I have been calling transparency. Will your existing application run on Exadata without any changes, i.e., will the move to Exadata be transparent to your applications and your users? Let's assume this is true. However, to be able to move to Exadata, your applications have to run on

***To move to Exadata, your applications have to run on 11gR2 . . . to get the most from Exadata, your applications should make use of partitioning and RAC.***

11gR2. If you aren't on 11gR2 already, this migration could be a large project. Further, Exadata requires that you support ASM and RMAN. This could be another migration. Finally, to get the most from Exadata, your applications should make use of partitioning and RAC. Yet another migration? It may well be true that your existing databases could run Exadata without any modifications, once they were "modified" to run on 11gR2, ASM, RAC, etc.

Second is the support organization. Who will support ASM, Exadata storage cells, CellCLI, and dcli utilities? Will this be handled by your existing DBA staff or the SA staff? Who will support the Exadata patching process that spans OS, storage, and database? We normally patch development, then test, and then do production. What does this mean for Exadata? While we could create three databases in the Exadata machine, how about patches that affect the Exadata machine overall? Do we no longer need development and test environments to validate these high-level patches? Exadata provides significant performance gains but will require significant organizational changes as well. ▲

The statements and opinions expressed here are my own and do not necessarily represent those of Oracle Corporation.

*Brian Hitchcock worked for Sun Microsystems for 15 years supporting Oracle databases and Oracle Applications. Since Oracle acquired Sun he has been with Oracle supporting the On Demand refresh group and most recently the Federal On Demand DBA group. All of his book reviews, presentations and his contact information are available at <http://www.brianhitchcock.net>.*

Copyright © 2011, Brian Hitchcock

**QUILOGY<sup>®</sup>**  
SERVICES  
FROM ASPECT

**Oracle Professional Consulting and Training Services**

Certified training and professional consulting when you need it, where you need it.

**ORACLE** APPROVED EDUCATION CENTER **ORACLE** EDUCATION RESELLER

[www.quilogyservices.com](http://www.quilogyservices.com)  
[education@aspect.com](mailto:education@aspect.com)  
866.784.5649

