



An Oracle White Paper  
December 2011

# What Oracle Solaris Brings to Oracle Exadata Database Machine

Introduction .....	1
Exadata System Overview .....	2
Oracle Solaris – Over 25 Years of Working with Oracle Database .....	3
Unmatched Reliability .....	4
Performance and Scalability .....	5
Leading Security .....	6
Summary .....	7
Resources .....	7

## Introduction

For over 25 years, Oracle Solaris engineering has been working closely with Oracle Database engineers to ensure that Oracle Database achieves its full potential when running on Oracle Solaris. This included providing mission-critical reliability while scaling to the largest databases in the world. Oracle Solaris, one of two operating system options for Oracle Exadata, is a key building block that provides mission-critical reliability, performance, and security that are essential for database consolidation. This paper provides a high level overview of the Oracle Exadata Database Machine and then describes the value that Oracle Solaris brings to the Oracle Exadata offering.

## Oracle Exadata System Overview

The Oracle Exadata Database Machine is an easy to deploy solution for hosting the Oracle Database and delivers the highest levels of database performance available. The Oracle Exadata Database Machine is a “cloud in a box” composed of database servers, Oracle Exadata Storage Servers, and an InfiniBand fabric for internal networking. It delivers outstanding I/O and SQL processing performance for online transaction processing (OLTP), data warehousing (DW) and consolidation of mixed workloads. Extreme performance is delivered for all types of database applications by leveraging a massively parallel grid architecture using Real Application Clusters and Exadata storage. The Oracle Exadata Database Machine delivers breakthrough performance with linear I/O scalability.

The Exadata Storage Server is an integral component of the Oracle Exadata Database Machine. Extreme performance is enabled by several features of the product:

- Database aware storage services, such as the ability to offload database processing from the database server to storage, enable higher throughput and are transparent to SQL processing as well as database applications. Hence just the data requested by the application is returned rather than all the data in the queried tables.
- Exadata Smart Flash Cache dramatically accelerates Oracle Database processing by speeding I/O operations. The flash provides intelligent caching of database objects to avoid I/O operations on hard disk drives.
- Exadata Hybrid Columnar Compression, an advanced compression technology, typically provides 10x, and higher, levels of data compression. Exadata compression boosts the effective data transfer by an order of magnitude.
- Building on the superior security capabilities of the Oracle Database, the Exadata Storage Server provides the ability to query fully encrypted databases with near zero overhead. This capability enables users to take advantage of the world's most secure database machine without sacrificing performance.

The Database Machine is a pre-configured system ready to be turned on day one, taking significant integration work, cost, and time out of the database deployment process. Since it is a well-known configuration, Oracle Support is very familiar with how to service the system, resulting in a superior support experience with the system. The common infrastructure for deploying a database for any application, whether OLTP, DW, a mix of the two, or as a platform for consolidation of several databases, creates tremendous opportunities for efficiencies in the datacenter. It is truly a “cloud in box”.

While an Oracle Exadata Database Machine is an extremely powerful system, a building block approach utilizing Oracle’s network fabric allows Oracle Exadata Database Machines to scale to almost any size as illustrated in Figure 1. Multiple Full Rack and Half Rack systems can be connected using the InfiniBand fabric in the system to form a larger single system image configuration. The InfiniBand infrastructure (switches and port cabling) is designed to provide this growth option.



Figure 1 Eight connected Oracle Exadata Database Machine racks form a single system.

Exadata hardware is pre-assembled and delivered in standard rack configurations. Each Exadata configuration is a unit of elastic cloud capacity balanced for compute- or I/O-intensive workloads. Each contains a number of compute nodes, a scale out, high-performance disk storage subsystem, and a high-bandwidth interconnect fabric. The fabric comprises the switches needed to connect every individual component within the configuration as well as to externally connect additional Exadata or Exalogic Elastic Cloud racks. In addition, each configuration includes multiple 10 Gigabit Ethernet ports for integration with the client access network. Gigabit Ethernet ports are used for integration with the datacenter's management network. All Exadata configurations are fully redundant at every level and are designed with no single point of failure.

The Oracle Exadata Database Machine has also been designed to work with, or independently of, the Oracle Exalogic Elastic Cloud. The Exalogic Elastic Cloud provides the best platform to run Oracle's Fusion Middleware and Oracle's Fusion applications. The combination of Exadata and Exalogic is a complete hardware and software engineered solution that delivers high-performance for all enterprise applications, including Oracle EBusiness Suite, Siebel, and PeopleSoft applications.

## Oracle Solaris – Over 25 Years of Working with Oracle Database

Oracle Solaris is the strategic platform for today's demanding enterprise, delivering proven results on everything from mission-critical enterprise databases to high-performance Web farms. For customers facing challenging business and technical requirements—such as lowering costs, simplifying system administration, and maintaining high service levels—Oracle Solaris is the ideal cross-platform choice.

Oracle Solaris brings all the key advantages of a mission-critical operating system to Exadata while leveraging the many advantages of an engineered system. Oracle databases running on Oracle Solaris benefit from all the database and InfiniBand tuning, the interconnect fabric itself, the integration, and the extensive testing that go into Exadata.

Oracle Solaris has over 25 years of being a premier platform for Oracle Database, making the OS an especially robust environment for Oracle databases and associated applications. The many benefits of a close working relationship between the Oracle Database and the Oracle Solaris teams are reflected in the areas discussed below. These benefits include unmatched reliability for the database environment, excellent performance and scalability, and superior security features.

## Unmatched Reliability

Oracle Solaris brings proven datacenter reliability to Exadata with a long history of industry-leading reliability features and innovation. Providing a rock solid foundation for Exadata reliability is critical for a system that will be a single point of availability for many consolidated database instances.

Key advantages of Oracle Solaris include:

- **Continuous uptime**—Oracle Solaris [Predictive Self-Healing](#)<sup>1</sup> diagnoses, isolates, and helps to recover from hardware and application faults. The Oracle Solaris Service Manager Facility manages application services in a uniform way and attempts to recover failed services using dependency analysis to restart services in the proper order and ensure that all dependent services are addressed.
- **Safely diagnose problems on the fly**—[Oracle Solaris DTrace](#)<sup>2</sup> allows administrators to safely analyze, tune, and troubleshoot applications on production systems with little or no performance impact.
- **Safe and quick system updates**—Boot environments in Oracle Solaris 11 represent a significant evolution of the Live Upgrade experience used in Oracle Solaris 10 and earlier releases. An out of the box configuration of Oracle Solaris 11 allows for system updates to be applied in parallel on a live production system. Oracle Solaris 11 boot environments are designed specifically for the ZFS file system, utilizing its fast snapshot and clone features to save a copy of the boot environment any time a software update to the system is performed. A snapshot is taken before any package is installed or updated, and if the update includes modifications to the system that will require a reboot to take effect, then the snapshot is cloned to create a new boot environment. The package operations are then applied to the new boot environment. Once an administrator is satisfied that the update is ready to be used, the system can be rebooted into the new boot environment to activate it. In the event that an update turns out to be problematic for some reason, they can quickly activate a previous boot environment and reboot back into it to restore a prior state. ZFS file system snapshots and clones have low overhead and provide unprecedented flexibility for system management.

---

<sup>1</sup> <http://www.oracle.com/us/products/servers-storage/solaris/solaris-pred-self-healing-ds-075587.pdf>

<sup>2</sup> <http://www.oracle.com/us/products/servers-storage/solaris/solaris-dynamic-tracing-ds-067318.pdf>

- **Simplified, improved updates**—Oracle Solaris 11 uses the new Image Package System (IPS), which is a network-based package management that provides a framework for complete software lifecycle management such as installation, upgrade and removal of software packages. Combined with the ZFS file system and boot environments, IPS offers completely safe system upgrades. Administrators can install software from network-based package repositories with full automatic dependency checking for any additional libraries that might be required during a software package install. Administrators can also control access to different software package repositories, and mirror existing repositories locally for network-restricted deployment environments.

## Performance and Scalability

Oracle Solaris is designed to provide the highest levels of performance and scalability. The Oracle Solaris and Oracle Database teams have worked together for many years to help ensure that the Oracle Database gets maximum scale and performance with Oracle Solaris. Further work was done for Exadata, particularly around InfiniBand performance. Some of these built-in technologies that help deliver increased database performance include:

- **High-performance networking stack**—The networking stack in Oracle Solaris 11 has undergone significant re-architecture to unify, simplify and enhance the observability and interoperability of network interfaces and features. In particular, the InfiniBand networking stack has been enhanced along with significant performance improvements for Exadata in particular. As discussed in some detail in the [technical overview paper on the Oracle Exadata Database Machine and Exadata Storage Server](#)<sup>3</sup>, InfiniBand has a key role in delivering the extreme database performance on Exadata. Oracle Solaris provides the core software layers that support the InfiniBand capabilities and optimizations discussed in this paper. InfiniBand is not only high bandwidth, but also delivers low latency.
- **Non-Uniform Memory Access (NUMA) I/O**—Exadata is based on a NUMA architecture, where each CPU or set of CPUs is associated with its own physical memory and I/O devices. For best I/O performance, the processing associated with a device should be performed close to that device, and the memory used by that device for DMA and PIO should be allocated close to that device as well. Oracle Solaris 11 adds support for this architecture by placing operating system resources (kernel threads, interrupts, and memory) on physical resources according to criteria such as the physical topology of the machine, specific high-level affinity requirements of I/O frameworks, actual load on the machine, and currently defined resource control and power management policies.
- **Intimate Shared Memory Performance**—Significant integration work has been done in Oracle Solaris to improve the performance of the Oracle Database stack for Oracle Solaris systems with

---

<sup>3</sup> <http://www.oracle.com/technetwork/database/exadata/exadata-technical-whitepaper-134575.pdf>

large memory. Speed improvements to Intimate Shared Memory (ISM) creation, locking, and destruction have resulted in up to 8x start up performance improvement for the Oracle Database.

- **Multiple page-size support**—Memory intensive applications, which have a large working set, often perform sub-optimally unless they make use of larger memory management unit (MMU) pages. This is because they make inefficient use of the microprocessor's facility known as the 'Translation Lookaside Buffer' or 'TLB'. Exploiting larger page sizes for the microprocessor's memory management unit allows more efficient use of the TLB, ultimately resulting in improved application performance.
- **Memory placement optimization (MPO)**—MPO helps optimize latency and bandwidth performance by attempting to ensure that memory is as close as possible to the processors that access it, while still maintaining enough balance within the system to avoid the introduction of bottlenecking hotspots. This enables Oracle Solaris to optimize for locality with thread scheduling and memory allocation. MPO benefits the Oracle Database specifically by increasing locality of reference to shared memory (SGA) and private memory (PGA) access by Oracle processes
- **Oracle Solaris DTrace**—DTrace helps developers isolate performance bottlenecks when designing application code, and allows system administrators to safely analyze and resolve a broad range of issues up and down the Exadata stack in live production environments.

## Leading Security

Oracle Solaris includes leading security technologies and is designed to prevent security problems as well as protect against security vulnerabilities. Security is designed into the entire Exadata stack. Security benefits for Exadata that are provided by Oracle Solaris include:

- **Least privilege**— In today's systems, it is important to have different levels of access for different types of administration, and to record who performed what action. This is in accordance with the security principle of *least privilege* that demands that every program and user of the system operate using the smallest set of privileges necessary to complete the job. The solution in the Oracle Solaris is the use of roles for specific administrative tasks. In Oracle Solaris, roles are assigned execution profiles, a mechanism used to bundle the commands and authorizations needed to perform a specific function. This both provides a means to control what each individual user can do, and also enables visibility into to who is really performing which changes to the system.
- **Accountability**—The Oracle Solaris audit feature provides the ability to log system activity at a granular level. System activity refers to any auditable Oracle Solaris event, such as system calls on the server machine, packets sent over the network, or a sequence of bits written to disk.
- **Compliance**—By using the security features of Oracle Solaris, both tracking and enforcing compliance is simplified. Users have clearly defined roles and access rights. This improves overall security while increasing visibility into who is doing what on the system. Security policies are enforced, while at the same time tracking mechanisms are enabled to verify actual behavior.



- **Secure by default**—Oracle Solaris ships with minimal services running and only one networked service available (ssh).

## Summary

Oracle Solaris 11 on the Oracle Exadata Database Machine is a key component of a complete, integrated solution that offers the IT infrastructure needed to meet today's rising demands for consolidation in the datacenter. Oracle Solaris delivers key datacenter advantages in reliability, performance and scalability, and security for Exadata systems.

The flexibility and scalability of Exadata, combined with its superior performance, make it an excellent platform for database consolidation. This also means that Exadata has to be extremely reliable so that business operations are not disrupted by unplanned downtime. Oracle Solaris contributes to Exadata reliability, bringing years of enterprise mission-critical production experience to Exadata.

Many of the key features of Exadata are focused on achieving the extraordinary performance of the system for a wide range of database workloads. Oracle Solaris has been modified and tuned over 25 years of close collaboration with Oracle Database engineers to run the world's largest and most responsive databases. Now Oracle Solaris has been further enhanced to provide superior performance and scale on Exadata.

## Resources

Oracle Exadata Database Machine	<a href="http://www.oracle.com/exadata">http://www.oracle.com/exadata</a>
Oracle Solaris 11	<a href="http://www.oracle.com/solaris">www.oracle.com/solaris</a>
Oracle Solaris 11 Technical Resources	<a href="http://www.oracle.com/technetwork/server-storage/solaris11/overview/index.html">www.oracle.com/technetwork/server-storage/solaris11/overview/index.html</a>



What Oracle Solaris Brings to Oracle Exadata  
Database Machine  
December 2011

Oracle Corporation  
World Headquarters  
500 Oracle Parkway  
Redwood Shores, CA 94065  
U.S.A.

Worldwide Inquiries:  
Phone: +1.650.506.7000  
Fax: +1.650.506.7200

oracle.com



Oracle is committed to developing practices and products that help protect the environment

Copyright © 2011, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. UNIX is a registered trademark licensed through X/Open Company, Ltd. 1010

**Hardware and Software, Engineered to Work Together**