

Oracle Deployment on x86-64 Linux: Best Practices for Oracle On Demand

An Oracle White Paper

August 2005

Oracle Deployment on x86-64 Linux: Best Practices for Oracle On Demand

Introduction.....	3
Oracle On Demand Environment.....	3
Servers in Use Today.....	3
Limitations Encountered.....	3
64-bit Linux Systems.....	4
64-bit Systems Deployment.....	5
Changes to On Demand Infrastructure.....	5
Kernel Crash Dump.....	5
lkcd-netdump Server Configuration.....	5
lkcd Client Configuration.....	6
Network Interface Failover.....	7
Recommended Configuration.....	7
Manual Configuration.....	8
Automated OS Patching.....	9
YOU Server Configuration.....	9
Online Update Client Configuration.....	9
Performance Tuning and Sizing.....	9
Oracle OLTP Workload Setup.....	10
Order Entry Workload Setup.....	10
Results.....	11
Oracle OLTP Workload Results.....	11
Order Entry Workload Results.....	11
Conclusion.....	11
Appendix A: Key System Parameter Values.....	12
init.ora.....	12
sysctl.conf.....	12
limits.conf.....	12
Appendix B: Oracle Products Available on 64-bit SLES9.....	13

Oracle Deployment on x86-64 Linux: Best Practices for Oracle On Demand

INTRODUCTION

Today, Oracle On Demand uses x86-based servers running 32-bit Linux OS in its data center. These servers offer excellent value and performance. However, in order to address limitations in 32-bit architecture that may affect performance, scalability, and reliability in some cases, Oracle On Demand is now deploying 64-bit infrastructure to complement the existing 32-bit environment. This document describes some of the limitations faced by 32-bit servers and how AMD Opteron-based x86-64 servers running 64-bit Linux OS are being used to address them for the On Demand environment. This paper includes best practices from the 64-bit deployment within the Oracle On Demand data center.

ORACLE ON DEMAND ENVIRONMENT

To support customers' need to increase the value of their investment in Oracle products, Oracle is driving the convergence of software and services. IT services are the delivery mechanism to implement and manage enterprise applications. Leading customers view the delivery of the software as a lifecycle of continuous development, improvement and management that determines the value technology delivers to their organizations. These customers rely on Oracle to manage their infrastructure and software to keep their systems available, performing, secure and up to date, making their companies more productive and positioned to take advantage of the latest capabilities. Oracle On Demand delivers Oracle software as a service to Oracle customers around the world. Oracle software is managed by Oracle experts – at Oracle, at customer location, or at a third party. Day to day monitoring, patching, updates and problem management is handled by Oracle experts using Oracle's or the customers' infrastructure, freeing up capital and resources to allocate to more strategic projects.

Servers in Use Today

Oracle On Demand's data center consists mostly of x86-based servers running 32-bit Linux. This platform was chosen for its excellent value and performance. When dealing with larger workloads, multiple servers can be allocated for the middle tier and the database tier can be migrated to Oracle Real Application Cluster (RAC). Although this approach is effective, performance of individual servers at higher workloads face scalability issues due to their 32-bit nature.

Limitations Encountered

The x86 instruction set architecture supports 32-bit address space. Each process on x86-based Linux servers can access approximately 4GB of memory and the size of Oracle database's system global area (SGA) is limited to 1.7GB. Although

various techniques in both the operating system and the database are available to increase the SGA beyond this limit, the increase is limited and flat memory access is not possible at larger SGA sizes. In addition, scalability is affected due to lowmem restriction, which limits the number of concurrent database connections. There is a 2GB file size limit, which requires software workarounds. File size workarounds are already in place and are transparent to the end user, but they add an extra layer of complexity. hardware systems. These systems utilize multiple processors that share common memory and disk resources and hence are also known as ‘shared everything’ systems. Primary advantages of SMP systems include simplicity of application development and ease of administration. These systems, however, do not provide any inherent fault-tolerance—the failure of a single critical component such as a CPU could bring the entire system down. Further, they are currently somewhat limited in terms of scalability and growth due to limitations in available system bus bandwidth and operating system software scalability.

64-bit Linux Systems

Oracle On Demand is addressing these limitations by deploying AMD Opteron-based servers running Novell SUSE Linux Enterprise Server 9 (SLES9) for x86-64. These servers offer the following advantages over the machines in use today:

- Support for 40-bit (1TB) physical address space and 48-bit (256TB) virtual address space in the processor
- Native execution of both 64-bit x86-64 and 32-bit x86 applications
- Twice the number of general purpose registers and 128-bit XMM registers for 64-bit applications
- Integrated memory controller on each processor
- The use of HyperTransport links to connect processors and external devices

Support for larger address space allows SGA to grow significantly beyond 1.7GB without any of the workarounds required for 32-bit servers. Access to more memory allows a database to keep more data in memory and support larger number of concurrent users. This will enable each database server to support increased number of middle tier users without impacting performance or reliability.

The integrated memory controller provides each processor with fast low latency access to main memory. In multiprocessor systems, each processor has direct access to memory attached to its memory controller. When a processor needs access to memory that is attached to another processor’s memory controller, the data is transferred through HyperTransport connections that link the processors. In this configuration, system-wide memory bandwidth grows with the addition of new processors.

The native execution of 64-bit and 32-bit applications allows the server to run both types of binaries without performance degradation. 32-bit applications will each have access to full 4GB memory space, so memory intensive 32-bit applications will benefit from running inside the 64-bit system. Native 32-bit support allows resource intensive applications to be moved to 64-bit while other tools stay as is. This eases the migration to 64-bit environment since On Demand deployments consist of many tools and management agents, many of whom are not resource intensive.

64-BIT SYSTEMS DEPLOYMENT

These 64-bit Linux servers are first being used to run the database tier of Oracle E-Business Suite deployment. The database tier runs single instance databases and the middle tier runs on separate 32-bit Linux servers. The database tier and the middle tier run on separate servers.

The database tier currently runs Novell's SUSE Linux Enterprise Server 9 (SLES9) with Service Pack 2 for x86-64. It runs Oracle9i Database R2 9.2.0.6 for E-Business Suite 11.5.10 Cumulative Update 1 (CU1) on Linux x86-64. 32-bit tools for E-Business Suite Certified Configuration are also included. Future implementations will include Oracle Database 10g.

The middle tier runs Red Hat Enterprise Linux 3 Update 5 for x86 and Oracle E-Business Suite Certified Configuration 11.5.10 CU1 for Linux x86 along with associated tools.

CHANGES TO ON DEMAND INFRASTRUCTURE

The On Demand environment makes use of various tools to administer and maintain the servers and to diagnose and remedy any issues that come up. The introduction of 64-bit Linux servers in the On Demand environment created a need to adapt some of these tools for the new servers. They include kernel crash dump, network interface failover, and automated OS patching.

Kernel Crash Dump

32-bit Linux servers use netdump to send kernel crash dumps over the network. 64-bit Linux servers will use Linux Kernel Crash Dump (LKCD) that comes included in SLES9 for this task. Like netdump, LKCD can send crash dump data over the network. A separate LKCD server will be set up since LKCD clients do not work with netdump servers.

lkcd-netdump Server Configuration

SLES9 needs lkcdutils and lkcdutils-netdump-server packages to run an lkcd-netdump server. Run the following command to see if the packages are installed.

```
# rpm -q lkcdutils lkcdutils-netdump-server
```

The packages are not installed if this command does not return the names of both packages along with their version numbers. In this case, they must be installed using YaST.

First, make sure that YaST's install source module lists the latest service pack at the top of the priority list.

```
# /sbin/yast2 inst_source
```

If needed, move the service pack media to the top and select "Finish".

Next, install the packages.

```
# /sbin/yast -I lkcdutils lkcdutils-netdump-server
```

The necessary packages are now installed.

To configure the lkcd-netdump server, modify the values of DUMPPDIR and TARGET_PORT in /etc/sysconfig/dump as needed. DUMPPDIR defines the location where crash dumps will be saved, and TARGET_PORT defines the port where lkcd-netdump server will be listening.

On SLES9 SP2, lkcd-netdump can use secure connection to receive crash dump information over network. To enable this feature, set DUMP_FLAGS to "0x40000008" in /etc/sysconfig/dump.

When the lkcd-netdump server is configured, register the service for startup using the following command.

```
# /sbin/chkconfig netdump-server on
```

Start the service manually with the following command.

```
# /etc/init.d/netdump-server start
```

lkcd Client Configuration

The lkcdutils package is needed to send crash dump data to the lkcd-netdump server. Run the following command to see if lkcdutils package is installed.

```
# rpm -q lkcdutils
```

The package is not installed if the command does not return the lkcdutils package name along with its version number. In this case, the package must be installed using YaST.

First, make sure that YaST's install source module lists the latest service pack at the top of the priority list.

```
# /sbin/yast2 inst_source
```

If needed, move the service pack media to the top and select "Finish".

Next, install the lkcdutils package.

```
# /sbin/yast -I lkcdutils
```

The lkcdutils package is now installed.

To configure an lkcd client, first modify the values of TARGET_HOST and TARGET_PORT in /etc/sysconfig/dump. Set TARGET_HOST to the address of the lkcd-netdump server and set TARGET_PORT to the port number where the server will be listening.

Next, modify the values of SOURCE_PORT, DUMPDEV, DUMP_LEVEL, and DUMP_COMPRESS in /etc/sysconfig/dump as needed. SOURCE_PORT defines the port to be used by the machine sending the crash dump. DUMPDEV sets the name of the dump device. For network dumps, set the value to the name of the network interface used to send the crash dump. For example, if eth1 will be used, set DUMPDEV to "eth1". DUMP_LEVEL defines what will be included in the crash dump. See /etc/sysconfig/dump for details on the possible values. DUMP_COMPRESS defines which compression scheme the kernel should attempt to use on the crash dump. Possible values are 0 for no compression, 1 for RLE compression, and 2 for GZIP compression.

On SLES9 SP2, lkcd can use secure connection to send crash dump information over network. To enable this feature, set DUMP_FLAGS to "0x40000008" in /etc/sysconfig/dump and run the following command. When prompted, enter the lkcd-netdump server's root password to register the client.

```
# /sbin/lkcd login
```

When the lkcd client is configured, register the service for startup using the following command.

```
# /sbin/chkconfig lkcd-netdump on
```

Start the service manually with the following command.

```
# /etc/init.d/lkcd-netdump start
```

Network Interface Failover

Bonding driver will handle network interface failover for 64-bit Linux servers.

Bonding driver takes multiple network interface cards, known as slaves, and combine them into one virtual interface. For failover, the bonding driver will be configured in active-backup mode. In this mode, only one slave is active at any given time. A different slave becomes active when the active slave fails.

Recommended Configuration

Create /etc/sysconfig/ifcfg-bond* based on ifcfg-eth* files. The file name should match the name of the bonding driver interface. For example, when configuring bond0, the file should be called /etc/sysconfig/ifcfg-bond0. Then modify the new file.

The following is an example ifcfg-bond0 file. In this example, bond0 is running in active-backup mode and eth2 and eth3 are slaves for bond0.

```
BOOTPROTO=static
BROADCAST=192.168.0.255
IPADDR=192.168.0.1
NETMASK=255.255.255.0
NETWORK=192.168.0.0
STARTMODE=onboot
BONDING_MASTER='yes'
INTERFACE='bond0'
BONDING_MODULE_OPTS='miimon=100 mode=active-backup
primary=eth2'
BONDING_SLAVE0=eth2
BONDING_SLAVE1=eth3
```

When `ifcfg-bond*` is used to configure the bonding driver, `/etc/init.d/network` and `/sbin/ifup` will start the interface automatically.

Manual Configuration

Steps here describe the procedure for configuring `eth2` and `eth3` as slaves for `bond0` running in active-backup mode. First, add the following lines to `/etc/modprobe.conf.local`.

```
alias bond0 bonding
options bond0 miimon=100 mode=active-backup
primary=eth2
```

Next, shut down `eth2` and `eth3`.

```
# /sbin/ifconfig eth2 down
# /sbin/ifconfig eth3 down
```

When finished, load the bonding driver module and start up `bond0`.

```
# /sbin/modprobe bond0
# /sbin/ifconfig bond0 192.168.0.1 network
255.255.255.0 broadcast 192.168.0.255 up
```

Finally, start up the slave interfaces.

```
# /sbin/ifenslave bond0 eth2
# /sbin/ifenslave bond0 eth3
```

Automated OS Patching

An automated online update tool is available in SLES9 to keep the servers updated with latest OS patches. To use this tool, a separate machine is first configured to be the local YaST Online Update (YOU) server. The local YOU server can automatically synchronize with the main update server to make updates available to local machines. Local machines can then use the online update client to automatically pull the updates from the local YOU server.

YOU Server Configuration

Machines running SLES9 can be configured to be a local YOU server using the YaST tool. First, enter the network proxy settings if needed. Use YaST's proxy configuration tool to enter this information.

```
# /sbin/yast2 proxy
```

Next, start YaST's YOU server configuration tool.

```
# /sbin/yast2 you_server
```

If an HTTP server is not installed, YaST will ask you to install one. An HTTP server is needed to run the YOU server.

Two products, SUSE SLES 9 and SUSE CORE 9, are included by default. Edit the entries to add SUSE Portal authentication information. This information is needed to download software update for SLES9 from SUSE.

When finished, select "Synchronize Now" to synchronize the YOU server. Updates will be downloaded to `/var/lib/YaST2/you`. This may take some time. When completed, select "Start Server" to start the server. Select "Setup automatic synchronization" to synchronize the YOU server automatically.

Online Update Client Configuration

Start the online update client from YaST.

```
# /sbin/yast2 online_update
```

Select "User-Defined Location" for the installation source. Enter the name of the YOU server in the Location field.

```
http://<YOU server hostname>/YOU
```

Select "Configure Fully Automatic Update..." to run automatic updates. When done, select "Next" to select the patches to install. Finally, select "Accept" to install the patches.

PERFORMANCE TUNING AND SIZING

Oracle OLTP workload and Oracle E-Business Suite order entry workload were used to compare 32-bit and 64-bit Linux servers. Oracle OLTP workload was

used to measure the maximum number of database users that could be supported, and Oracle E-Business Suite order entry workload was used to measure the response times of application users as they performed predefined tasks. The two servers ran the backend database for both workloads.

The backend database was configured with buffer cache, java pool, and shared pool set to 1953MB, 150MB, and 1200MB respectively for both workloads. The same values were used on 32-bit and 64-bit systems. USE_INDIRECT_DATA_BUFFER was set to TRUE on the 32-bit system. Setting this parameter to TRUE enables the use of extended buffer cache mechanism, allowing 32-bit systems to access memory beyond 4GB. On the 64-bit system the parameter was set to FALSE. Values of key system parameters for the 64-bit system are available in the appendix.

Oracle OLTP Workload Setup

For the Oracle OLTP workload, database and database users ran on the same machine. The size of the swap space was kept intentionally low to avoid swap usage.

The following two machines were used for the workload.

	32-bit Server	64-bit Server
CPU	4 Xeon 1.5GHz	2 Opteron 248 2.2GHz
Physical Memory	16GB	16GB
Swap Space	2GB	2GB
OS	Red Hat Advanced Server 2.1 for x86	SUSE Linux Enterprise Server 9 SP1 for x86-64

Order Entry Workload Setup

For the order entry workload, the application users ran on three 32-bit middle tier machines. These machines processed the transactions by accessing the database, which ran on separate machines. A separate load generator drove the application users on the three middle tier machines. Database performance was measured by measuring the application users' response times.

The following machines were used for the workload.

	32-bit Middle Tier	32-bit Database Tier	64-bit Database Tier
CPU	2 Xeon 2.4GHz	4 Xeon 1.5GHz	2 Opteron 248 2.2GHz
Physical Memory	6GB	16GB	16GB
Swap Space	6GB	16GB	16GB
OS	Red Hat Advanced Server 2.1 for x86	Red Hat Advanced Server 2.1 for x86	SUSE Linux Enterprise Server 9 SP1 for x86-64

Results

	32-bit Server	64-bit Server
Oracle OLTP Workload	300 database users	2800 database users without swapping
Order Entry Workload	350 Applications users: <ul style="list-style-type: none">• 90th percentile: 42.1 seconds• Average: 13.60 seconds	350 Applications users: <ul style="list-style-type: none">• 90th percentile: 6.91 seconds• Average: 2.35 seconds Was able to support 450 users.

Oracle OLTP Workload Results

At 300 database users, the 32-bit Linux server began to run low on lowmem. The 64-bit Linux server reached 2800 database users without swapping out to disk.

Order Entry Workload Results

When using a 32-bit database backend, at 350 application users the average response time was 13.60 seconds. 90th percentile response time was 42.1 seconds.

When using a 64-bit database backend, at 350 application users the average response time was 2.35 seconds and the 90th percentile response time was 6.91 seconds. In this configuration, 450 application users could be supported with acceptable response times.

CONCLUSION

The deployment of 64-bit Linux servers in Oracle On Demand will help to extend the performance of database servers running large workloads. Migration from 32-bit Linux servers is smooth and key features such as kernel crash dump over network and network interface failover are available on both platforms through similar tools. The platform offers greater performance, scalability, and reliability for Oracle On Demand. Best practices from this environment can be applied to other environments as appropriate.

APPENDIX A: KEY SYSTEM PARAMETER VALUES

The values presented here are for 64-bit Linux systems with 2 CPU 16GB RAM configured for Oracle9iR2 database with RAC disabled and HUGEPAGES enabled.

init.ora

Parameter	Value for 64-bit SLES9
background_core_dump	partial
db_block_buffers	250000
db_block_size	8192
filesystemio_options	directio
java_pool_size	157286400
shared_pool_size	1258291200
use_indirect_data_buffers	False

sysctl.conf

Parameter	Value for 64-bit SLES9
fs.file-max	1677721
kernel.sem	1000 32000 100 142
kernel.shmall	3282294
kernel.shmmax	8858370048
kernel.shmmni	4096
vm.disable_cap_mlock	1
vm.nr_hugepages	2000
vm.overcommit_memory	0
vm.page-cluster	3

limits.conf

Parameter	Value for 64-bit SLES9
* hard nofile	327679
* soft nofile	327679

APPENDIX B: ORACLE PRODUCTS AVAILABLE ON 64-BIT SLES9

The following Oracle products are available for SLES9 on x86-64 as of July 2005. For latest certification information, please refer to Oracle MetaLink at <http://metalink.oracle.com/>.

Product	Type	Version
Database	Enterprise Edition	10g 64-bit 9iR2 64-bit
	Standard Edition	10g 64-bit 9iR2 64-bit
Real Application Clusters		10g 64-bit 9iR2 64-bit
Application Server	Oracle Application Server	10g (10.1.2.0.1) 10g (10.1.2.0.0)
	Oracle Containers for J2EE	10.1.2.0.0
	Oracle TopLink	10g (9.0.4.5)
	OracleAS Metadata Repository Configuration Assistant	10g (10.1.2.0.0)
Oracle Warehouse Builder		10g
Oracle Enterprise Manager 10g Grid Control		10.1.0.3



Oracle Deployment on x86-64 Linux: Best Practices for Oracle On Demand
August 2005

Author: Akira Ijuin

Contributing Authors: Takeshi Seo, Atul Behari, David Wong, Mike Gagas

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

Copyright © 2005, Oracle. 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, JD Edwards, PeopleSoft, and Retek are registered trademarks of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.