# Oracle9*i* Release 2 on Linux x86-64 (AMD64)

*An Oracle White Paper*
*May 2004*

ORACLE

# Oracle9*i* Release 2 on Linux x86-64 (AMD64)

## EXECUTIVE OVERVIEW

Over the past several years, x86-based servers running the Linux operating system have gained in popularity. This is because over time, Linux has gained key enterprise features, and x86-based servers offer an excellent price to performance ratio. However, the x86 instruction set architecture suffers from a 4GB address space limitation. Ability to access large blocks of memory is critical for database applications, and although work has been done to get around this limitation, the x86 instruction set architecture is still limited in its ability to access large amounts of memory.

AMD has introduced a 64-bit extension to the x86 instruction set architecture called AMD64. AMD's first processor in the AMD64 family, the AMD Opteron, supports 256TB virtual address space, eliminating the large memory access problem. It is able to run both 64-bit AMD64 applications and 32-bit x86 applications from within a 64-bit operating system, and 32-bit applications will run with no degradation in performance. This paper describes key features of the AMD64 architecture, and looks at the performance of Oracle9*i* Release 2 on Linux/AMD64.

## AMD64 ARCHITECTURE

AMD64 is AMD's 64-bit extension to the x86 instruction set architecture. It enables 64-bit computing for natively compiled applications while delivering full backward compatibility with existing x86 applications. On April 2003, AMD released the AMD Opteron, its first 64-bit processor in the AMD64 processor family. Some of the key features of the AMD Opteron are:

- Support for 40-bit (1TB) physical address space and 48-bit (256TB) virtual address space.
- Native execution of 64-bit AMD64 applications.
- Native execution of 32-bit x86 applications.
- 8 additional general-purpose registers in 64-bit mode, for a total of 16.
- Support for SSE and SSE2 instructions.
- 8 additional 128-bit XMM registers in 64-bit mode for a total of 16.

- Integrated memory controller in the processor.
- Use of HyperTransport links to connect processors and external devices.

### Access to More Memory

AMD Opteron can address 1TB physical address space and 256TB virtual address space. In SuSE's AMD64 Linux kernel, each 64-bit process can access up to 0.5TB virtual memory, and the system can access up to 128TB virtual memory. 32-bit x86 applications running in the 64-bit kernel will have access to full 4GB address space. Additional memory is beneficial for database applications, as this allows the database to keep more data in memory and support large number of concurrent users.

### Support for 64-bit and 32-bit Applications

AMD64 allows 64-bit AMD64 applications and 32-bit x86 applications to run within a 64-bit operating system. AMD Opteron is able to run 64-bit and 32-bit applications at full speed without incurring performance degradation. This allows users to migrate to 64-bit applications as they become available while they continue to run other 32-bit applications on the same machine. For example, a database can be migrated to 64-bit while applications that use the database continue to run in 32-bit. The database will be able to take advantage of flat 64-bit memory access while 32-bit applications run with no performance degradation. If needed, these 32-bit applications can take advantage of the full 4GB address space.

While most 32-bit x86 Linux applications will run unmodified on 64-bit AMD64 Linux, some applications rely on certain 32-bit Linux features to run. For example, some applications expect the command "uname –m" to return "i686". These applications will fail because "uname -m" returns "x86_64" in AMD64. In another case, an application assumed that it ran within a 3GB address space typical in x86 Linux, and failed when it had access to the full 4GB.

A tool has been included in AMD64 Linux distributions to allow these applications to run properly. In United Linux 1.0 and SuSE Linux Enterprise Server 8, the tool is called linux32, and Red Hat Enterprise Linux has a similar tool called setarch. Both tools change the output of "uname -m" to "i686". For example, "linux32 uname -m" and "setarch i686 uname -m" will both return "i686" instead of "x86_64". The two tools also have an option to run 32-bit applications in 3GB address space. "linux32 --3gb bash" and "setarch -3 i686 bash" will start a bash shell session, and any 32-bit application started from this shell will run in 3GB address space.

### Integrated Memory Controller

AMD Opteron has an integrated dual channel 128-bit wide memory controller for accessing main memory. Using PC2700 DDR DIMMs, the integrated memory controller is capable of delivering 5.3GB/s memory bandwidth per processor.

This bandwidth along with the integration of the memory controller allows very fast low latency access to memory.

As more processors are added, each additional processor is able to access its own set of memory through the memory controller. When a processor needs access to data on another processor's memory controller, the data is passed through HyperTransport links that connect the processors.

### HyperTransport Links

Each processor supports up to three HyperTransport links. These links are used to connect processors and external devices. Each link can provide up to 6.4 GB/s bandwidth, and can be used to connect up to eight Opteron processors on a single system. A processor will at most be connected to 3 other processors, and communication with other processors will be done through one of these processors. The use of HyperTransport links along with the integrated memory controller allows memory bandwidth to scale with the number of processors.

## ORACLE9*i* RELEASE 2 ON AMD64 LINUX

The performance, scalability, and reliability characteristics of Oracle9*i* Release 2 were tested on Linux for AMD64. Oracle9*i* Release 2 for Linux/AMD64 was stressed using the Oracle OLTP workload to see how it would perform under various conditions. These conditions were simulated by adjusting the workload's user count, and by adjusting the System Global Area (SGA) of the database instance.

The tests were performed on 2-way and 4-way SMP systems. The test results cannot be used to compare the performance of the two systems, as the systems use different peripherals and storage configurations. The results are meant to compare the relative performance of each system under different workloads.

### Setup

The first system used for this test is a 4-way system with 32GB RAM, and the second is a 2-way system with 16GB RAM. Both machines used 1.4GHz AMD Opteron processors. The two systems were each connected to an external storage device through a fibre channel controller, and used SCSI disks for local storage. The systems ran SuSE Linux Enterprise Server 8 for AMD64 with Service Pack 2. The distribution uses a 2.4.19 SMP kernel.

The Oracle OLTP workload is designed to run various types of OLTP workloads on the database and measure the average number of transactions per minute (TPM). For this test, the workload users ran cached workload against the database. This workload keeps most of the required data in the SGA. The database and workload users ran on the same machine.

### System Configuration

The following table provides the system configuration used:

| System | 4-way | 2-way |
|---|---|---|
| CPU | Opteron 1,400MHz x4 | Opteron 1,400MHz x2 |
| Physical Memory | 32GB (8GB per CPU) | 16GB (8GB per CPU) |
| Swap Space | 8GB | 6GB |
| Local Storage | 36GB x3 | 36GB |
| External Storage | 145GB RAID5 | 251GB RAID0 |
| Fiber Channel Controller | QLogic QLA2200 | QLogic QLA2312 |
| OS | SLES8 for AMD64 with SP2 | SLES8 for AMD64 with SP2 |
| Kernel | k_smp-2.4.19-249 | k_smp-2.4.19-249 |

### Database Configuration

The database was configured with asynchronous I/O and direct I/O turned off, and `DB_BLOCK_SIZE` was set to 8192. `DB_CACHE_SIZE` was used to adjust the size of SGA. `SHARED_POOL_SIZE` and `SESSIONS` parameters were kept small, and were increased as needed to run the workloads. Change in `SHARED_POOL_SIZE` and `SESSIONS` account for the difference in SGA size on workloads that have the same `DB_CACHE_SIZE`.

Certain kernel parameters were adjusted to work with the database instance. SEMMSL, SEMMNS, SEMOPM, and SEMMNI were set to 250, 256000, 100, and 1024 on both systems. SHMMAX was set to the size of the physical memory.

### Results

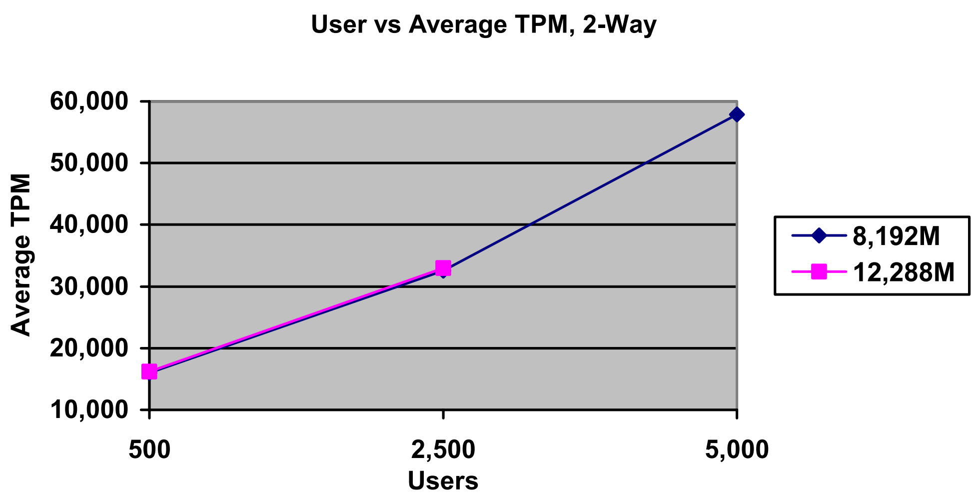
Average transactions per minute from successful runs of the Oracle OLTP workload were averaged to determine the result for a given configuration. The results are displayed in the following table:

| CPU | RAM | Swap | Users | db_cache_size | SGA | Average TPM |
|---|---|---|---|---|---|---|
| 2 | 16,384M | 6,144M | 500 | 8,192M | 8,554M | 15,984.16 |
| 2 | 16,384M | 6,144M | 2,500 | 8,192M | 8,698M | 32,592.50 |
| 2 | 16,384M | 6,144M | 5,000 | 8,192M | 8,794M | 57,900.33 |
| 2 | 16,384M | 6,144M | 500 | 12,288M | 12,666M | 16,223.49 |
| 2 | 16,384M | 6,144M | 2,500 | 12,288M | 12,826M | 32,976.65 |
| 2 | 16,384M | 6,144M | 5,000 | 12,288M | 12,923M | Out of Memory |
| 4 | 32,768M | 8,192M | 500 | 16,384M | 16,788M | 6,255.53 |
| 4 | 32,768M | 8,192M | 2,500 | 16,384M | 16,939M | 8,527.95 |
| 4 | 32,768M | 8,192M | 5,000 | 16,384M | 17,114M | 13,556.51 |
| 4 | 32,768M | 8,192M | 7,500 | 16,384M | 17,322M | 17,352.40 |
| 4 | 32,768M | 8,192M | 10,000 | 16,384M | 17,387M | 22,613.53 |
| 4 | 32,768M | 8,192M | 500 | 24,576M | 25,018M | 7,168.40 |
| 4 | 32,768M | 8,192M | 2,500 | 24,576M | 25,162M | 9,026.93 |
| 4 | 32,768M | 8,192M | 5,000 | 24,576M | 25,355M | 14,022.22 |
| 4 | 32,768M | 8,192M | 7,500 | 24,576M | 25,563M | 16,820.63 |
| 4 | 32,768M | 8,192M | 10,000 | 24,576M | 25,627M | 21,023.60 |

#### 2-Way System Results

The Oracle OLTP workload was run on the 2-way system with 500, 2,500, and 5,000 users. The database instance's db_cache_size was set to 50% and 75% physical RAM (8,192MB and 12,288MB respectively) for each run.

The two db_cache_size settings yielded very similar results for 500 and 2,500 user runs. In both cases the larger db_cache_size resulted in approximately 1% improvement in result. With 5,000 users, the system ran out of memory when db_cache_size was set to the larger value, and the workload was not able to finish. The database instance does not appear to benefit from the larger db_cache_size.

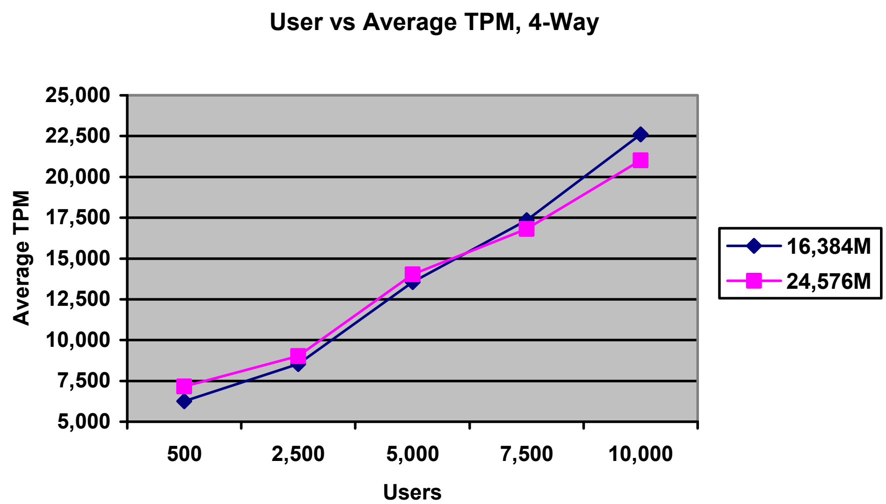



#### 4-Way system Results

The Oracle OLTP workload was run on the 4-way system with 500, 2,500, 5,000, 7,500, and 10,000 users. The database instance's db_cache_size was set to 50% and 75% physical RAM (16,384MB and 24,576MB respectively) for each run.

At 500 users, larger db_cache_size resulted in approximately 15% higher result. This improvement shrank as the number of users was increased. At 2,500 users the improvement was 6%, and at 5,000 users it was 3%. At 7,500 users the smaller db_cache_size performed better, yielding a 3% higher score. Smaller db_cache_size also performed better with 10,000 users. In this case the difference grew to almost 8%. The gradual improvement in the performance of the database instance with smaller SGA is most likely caused by the larger memory needed to support the increasing number of users.

With 10,000 users the system began to run out of memory with both SGA sizes.




## Recommendations

Several performance and scalability characteristics of Oracle9*i* Release 2 on AMD64 can be deduced from the results of the Oracle OLTP workload. On the 2-way system, the system began to run out of memory with 5,000 users and 12,923MB allocated for SGA. On the 4-way system, memory began to run out at 10,000 users, although several workload runs did complete with 17,387MB and 25,627MB SGA.

While the results are specific to the hardware configuration, database setup, and the type of application and workload used, it is hoped that the results will serve as a base reference point for configuring and tuning Oracle on AMD64 Linux. For example, in the 4-way system, at 7,500 users a database instance with an SGA of approximately 17GB performed better than one with an SGA of roughly 25GB. The larger SGA is most likely hurting performance because the combination of larger SGA and high user count is causing the system to swap. If the larger SGA is required, as with Oracle on other platforms, the number of users should be brought down, or the users should be moved to a separate client machine.

## CONCLUSION

AMD64 and the Opteron processor provide a new choice for users of x86-based Linux servers. Users can take their existing 32-bit applications and migrate to 64-bit in stages as needed. Once an application is migrated to 64-bit, it is able to take full advantage of the architecture, including access to larger amount of memory. Oracle9*i* Release 2 for Linux/AMD64 was able to go beyond the 4GB limit with no change in configuration. With the availability of the production release of Oracle9*i* Release 2 Linux/AMD64, users will be able to support 32-bit applications and enjoy the benefits of 64-bit computing on the same hardware.

ORACLE

**Oracle9*i* Release 2 on Linux x86-64 (AMD64)**
**May 2004**
**Author: Akira Ijuin**

**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**
**www.oracle.com**