SYBASE®

Optimizing Performance of Adaptive Server® Enterprise 12.5 on Linux

enabling

WHITE PAPER

Table of Contents

Overview

Linux is an extremely powerful, multitasking, multiuser, secure operating system. It has all the features of an enterprise operating system, including virtual memory, shared libraries, proper memory management, and TCP/IP networking. It runs on many different processors, including Intel x86, Itanium, Sun Sparc, SGI, AMD Opteron, and PowerPC. There are many Linux distributions available on the market. Popular among them are Red Hat, SuSE, Redflag, and Caldera.

As Moore's law predicted, the number of transistors on a microprocessor has doubled repeatedly, contributing substantially to the rapid increase in the clock speed of microprocessors. Units of work that once required many processors can now be completed by fractions of a single processor. This advance combined with the enterprise readiness of Linux, a cheap commodity operating system, makes for a very cost-effective IT solution. The combination has forced IT managers to seriously consider migration from "big iron" boxes to commodity hardware running Linux.

This trend provides a new set of challenges that extend from migrating existing applications to a new operating system to ensuring that the applications run reliably. High performance of applications on Linux has become an absolute requirement. An application that does not perform well is judged to be unavailable.

Some of the most common trends in this migration phenomenon are:

- Migration from big iron boxes to commodity Intel hardware running Xeon/P3-P4 systems with double the clock speeds of previous CPUs
- CPUs with hyperthreading, where a single physical processor can provide parallelism in the hardware layer
- Sizing based on increased clock speeds—for example, one 1.5 GHz CPU is considered to be equivalent to two 750 MHz CPUs

This paper describes in detail how to propel Adaptive Server Enterprise applications running on Linux to the high performance levels that Adaptive Server customers have come to expect.

Adaptive Server Enterprise on Linux

Adaptive Server Enterprise 11.0.3.3 was the first commercially available enterprise class database on Linux. Adaptive Server versions 12.5, 12.5.03, and 12.5.1 deliver features that provide improved operational scalability, tight security, high availability, and exceptional levels of performance. Adaptive Server 12.5.1 introduces additional performance enhancements such as derived tables; high availability features such as mount/unmount, dynamic reconfiguration, dynamic data cache management, and parallel recovery; and manageability features such as Job Scheduler, Cache wizard, MDA views, and the Migration tool. This arsenal of features makes Adaptive Server a high-performance enterprise data management platform that is always available and easy to administer.

With industry-standard benchmarks such as TPC-C, it has been demonstrated time and again that Adaptive Server is a very high performance database server. Internal benchmarks have demonstrated the ability of Adaptive Server to provide over one million transactions per minute, and scale up to 64 processors on a benchmark based on an industry order processing workload.

Information on the availability of various versions of Adaptive Server on different versions of Linux can be obtained from **http://www.sybase.com/linux/ase**.

System Performance Tuning

System Performance Tuning is the answer to this fundamental question: How can I get my software and hardware to do more work without buying more hardware? Some performance problems do require you to buy a bigger or faster computer, but many can be solved simply by making better use of the resources you already have.

The three core components in system performance tuning are:

- Processors
- Memory
- I/O (disk and network)

Demystifying Processors

New processor technologies have made it a requirement to understand the characteristics of these technologies and their impact on applications. Intel Xeon and some revisions of Pentium have built-in hyperthreading to provide thread-level parallelism at the hardware layer. Each clock cycle is thereby able to deliver more work. Some of the other benefits of hyperthreading are:

- Support for multithreaded code and multitasking operations through better use of processor resources
- Ability to run multiple threads/tasks simultaneously to increase the number of transactions that can be executed
- Improved reaction and response times for end users
- Increases in the number of users a system can support

One of the most common problems users face is determining the number of processors required to run applications when moving from slower processors to faster processors. Some of the main points to consider are:

- One 1.5 GHz CPU does not necessarily yield the same performance as two 750 MHz CPUs. The variables that need to be considered include L1/L2/L3 cache sizes, instruction cache sizes, memory latency, and cycles per instruction.
- Hyperthreading provides the ability for one physical processor to run two ASE engines. It is very important to understand that this is not equivalent to running with two physical processors.

RH Advanced Server 2.1 includes optimizations to the process scheduler that have enhanced scalability of the operating system. RH Advanced Server 2.1 Linux and Adaptive Server provide a wealth of performance tools for monitoring and tuning processors and engines. This section lists some of the tools and tricks a DBA can employ to make better use of available resources.

Processor Tips

Task/Problem	Solution
Identify number of processors/MHz	cat /proc/cpuinfo
Identify CPU utilization	mpstat, top, vmstat
Identify whether CPU is hyperthreaded	If the the flag is returned from the cat /proc/cpuinfo operation, the CPU is hyperthreading enabled (the default)
Disable hyperthreading	Use the menu at BIOS setup
Set process-processor affinity	Optimize in RH AS 2.1

EngineTips

Task/Problem	Solution
Identify number of engines	Select * from master.sysengines
Determine engine usage	sp_sysmon "interval", kernel
Have one engine but see more than one dataserver process (RH 7.2 only)	Additional processes are running for POSIX AIO

In addition to these options, Adaptive Server provides a wealth of tools that let you manage engine use efficiently:

- Logical process management lets you assign high-priority tasks to specific engines. This is also an extremely effective option when managing mixed workloads such as OLTP and DSS.
- "Runnable process search count" determines the number of times the engine loops looking for runnable tasks before yielding to the operating system.
- "CPU grace time" and "time slice" can be tuned to prevent tasks from hogging CPU cycles.
- The default values usually require tuning only in specific scenarios such as those described in the tips above. sp_sysmon provides recommendations on tuning some of these parameters.

The Memory Subsystem

The memory subsystem plays a very important role in improving overall system performance. The memory subsystem for both Linux and Adaptive Server has to be carefully tuned to get the best system performance. Any I/O avoided by accessing data from memory provides significant performance gains. Linux and Adaptive Server have a wealth of tools that enable you to use the memory subsystem effectively. It is recommended that swap size equal the size of physical memory available on the system.

Linux Memory Tips

Task/Problem	Solution
View memory parameters	free -k
Increase the maximum shared memory	echo "<maxshm>" /proc/sys/kernel/shmmax Alternatively, update /etc/sysctl.conf to include Kernel.shmmax = <max_memory in bytes> and run /sbin/sysctl –p for persistence across machine boots
Allow ASE to use up to 2.7 GB	echo 268435456 >/proc/<pid>/mapped_base Note: This command needs to be run by someone with root privileges and <pid > is the process id of the shell from where dataserver starts. [This command need not be executed for ASE 12.5.1]

Adaptive Server Memory Tips

Task/Problem	Solution
Increase Adaptive Server shared memory	sp_configure "max memory", value
Increase the memory for data cache	sp_cacheconfig "default data cache", value
Increase the unit of I/O issued by Adaptive Server	sp_poolconfig "cache_cow", value, "4K"
Monitor usage of Adaptive Server structures	sp_monitorconfig "all"
Monitor usage of Adaptive Server data cache	sp_sysmon "interval", dcache

In addition to these configuration options, Adaptive Server offers a rich set of memory management features. Adaptive Server provides the ability to pin hot objects to memory. This feature helps Adaptive Server improve system performance and scalability significantly. Some of the common scenarios where named caches are used are:

- Using named caches and cache partitions when there is spinlock contention in the default data cache
- Binding tempdb to its own cache when there is significant usage of tempdb (order by, group by, create index, temporary objects)
- Binding the transaction log to improve log write performance
- Binding hot indexes and hot tables to pin these objects to memory

Adaptive Server also manages data caches so that frequently accessed pages are not replaced, thereby avoiding expensive disk I/O. In version 12.5.1, Adaptive Server provides a Cache wizard to help identify hot objects that can be bound to named caches. Asynchronous prefetch fetches pages expected to be used in the near future. This is extremely useful when doing table scans, non-clustered index scans, and crash recovery.

Intel's Itanium processor family (IPF) provides the ability to use large memory as it uses 64-bit addressing. Adaptive Server Enterprise 12.5.1 will be available on this platform to enable usage of very large memory.

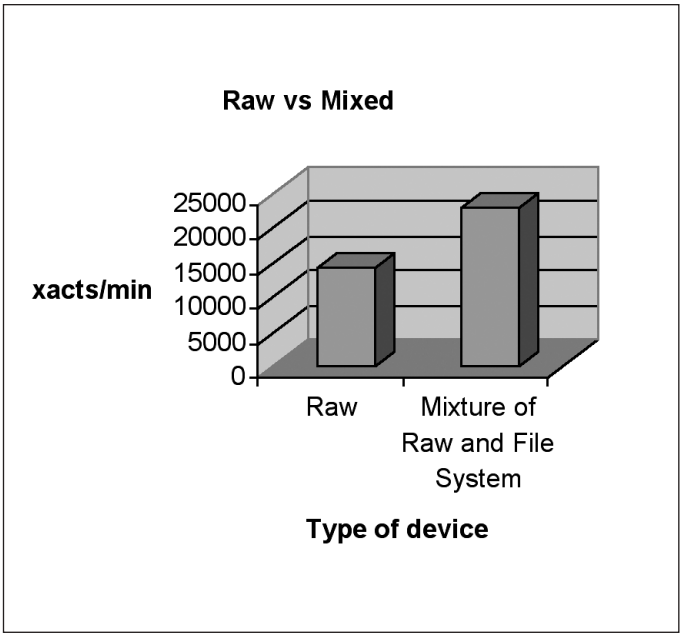
The Disk I/O Subsystem

Adaptive Server supports both raw devices and file systems. Asynchronous I/O (AIO) is supported on Adaptive Server Enterprise 12.5.0.3 by default. Adaptive Server also supports AIO on both raw devices and file systems on Linux. AIO helps improve overall system performance by making sure that CPU resources are not locked, waiting for I/O completion. The CPU cycles can be used servicing other tasks.

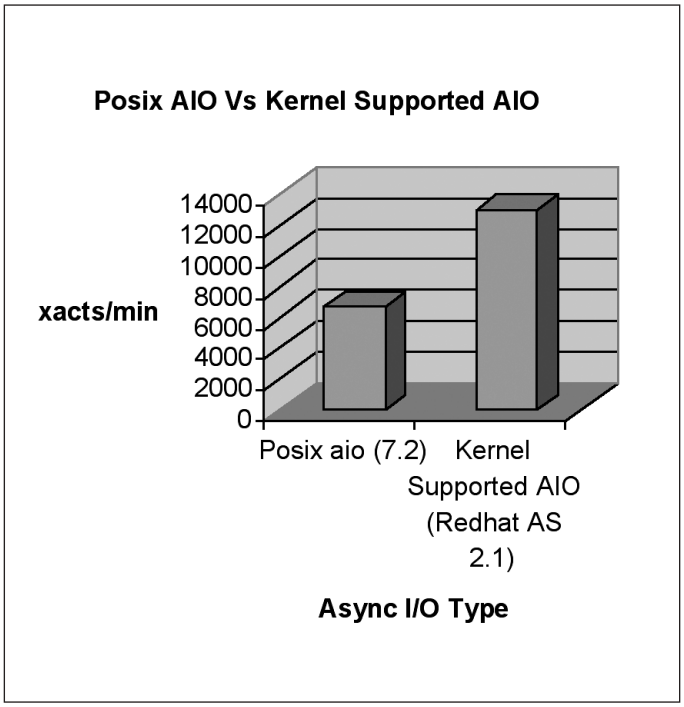
Adaptive Server Enterprise 12.5.0.3 also provides large file support that enables creation of devices greater than 2 GB. Adaptive Server Enterprise 12.5.0.3 can support up to 8 TB of data.

In general, some of the important points to remember while configuring the I/O subsystem are:

- A mixture of raw disks and file systems is the ideal configuration for overall system performance. Objects on which write access is infrequent and read access is frequent should be placed on file systems. Objects on which write access is frequent and read access is infrequent should be placed on raw devices.



- Entities such as transaction logs are best placed on raw devices, which provide very good service times.
- The use of file systems acts as a secondary buffer cache for Adaptive Server. This provides Adaptive Server with the ability to use much more than 2.7 GB. Pages not found in the Adaptive Server buffer cache can be accessed from the file system cache.
- Tempdb should ideally be placed on file systems with the "dsync" flag set to false using the sp_deviceattr stored procedure.
- The read response time of data devices is very important in ensuring high overall throughput. One of the best ways to achieve this is to ensure that the disk subsystem is not a bottleneck. Some of the optimizations to be considered are:
    - Ensure that there are a sufficient number of disks in an array. This will help improve I/O parallelism.
    - Use stripe size to help chunk data so that it is spread over multiple devices. Typical OLTP applications have stripe sizes of 64 KB for data and 8 KB for the transaction log.
    - If there are a number of disks being serviced by a controller, ensure that the controller is not a bottleneck.
    - Run iostat to ensure that service times are within reasonable bounds (good devices provide 12-15 milliseconds for data and 2-3 milliseconds for the log).
- Kernel-supported asynchronous I/O (Red Hat Advanced Server 2.1) is recommended. Adaptive Server has the ability to detect the AIO supported by the OS and adopt it without any configuration changes. On RH 7.2, Adaptive Server uses POSIX AIO, and on RH AS 2.1, it uses kernel-supported AIO.



Some of the commonly used commands are shown below.

Linux Disk I/O Tips

Task/Problem	Solution
Identify the file system types	/bin/mount
Create an ext3 file system	/sbin/mke2fs –b <Adaptive Server page size> -j /dev/sdg2
Monitor disk performance	iostat –d –x sdb [interval][count]
Control I/O throughput and response time	elvtune –r <val> -w <val> <dev>
Create a raw device	Run the command raw /dev/raw/raw1 /dev/sdb1 and in /etc/sysconfig/rawdevices add the line /dev/raw/raw1 /dev/sdb1
View existing raw devices	raw -qa

Adaptive Server Enterprise Disk I/O Tips

Task/Problem	Solution
Enable asynchronous I/O	Enabled by default
Monitor ASE device performance	sp_sysmon “interval”, diskio

Linux supports a variety of file systems such as EXT2, EXT3, and XFS. The recommended file system for Adaptive Server is EXT2 or EXT3 with journaling disabled.

The Network I/O Subsystem

Adaptive Server is capable of handling thousands of users and provides very good throughput and response times when dealing with a large number of users. For efficient network performance, use network cards that are capable of handling the network traffic. 100BASE-T and Gigabit Ethernet cards can handle large packets of data without significant overhead. This offloads CPU cycles from doing TCP processing. Use full duplex for data transmission to enable simultaneous read and write access to the network. Half duplex provides the ability only to read or write at a given time. These parameters are configured at the NIC level. The following command indicates if the NIC uses full or half duplex.

```
%  cat /proc/net/nicinfo/eth0.info | grep Duplex
```

To optimize tcp performance, use some of the following tuning parameters:

- Optimize for a large number of user connections

```
% echo 1024 65000 > /proc/sys/net/ipv4/ip_local_port_range
```

- Increase the amount of memory associated with socket buffers
- Reduce the amount of work to be done by the TCP stack

```
% echo 0 > /proc/sys/net/ipv4/tcp_sack
% echo 0 > /proc/sys/net/ipv4/tcp_timestamps
% echo 0 > /proc/sys/net/ipv4/tcp_window_scaling
```

The parameters that you can tune for efficient network processing with Adaptive Server include:

- Adaptive Server configuration parameters "max network packet size" and "default network packet size".
- CS_PACKETSIZE, which controls the TDS packet size for the client. Tune this property using the ct-library ct_conn_props() API when handling large packets of data such as text/image or when handling bulk data. This property can also be set for Java clients using the jConnect JDBC driver. You can use the property object's put() method for the packetsize property, or use the isql -A option.
- In version 12.5.1 and later, "send doneinproc tokens". Set this parameter to 0 to improve network performance by sending doneinproc tokens only for select statements.

Conclusion

Enterprise performance on Linux has come a long way. Adaptive Server has been working with a host of Linux vendors to achieve the very high levels of performance expected from Adaptive Server. Adaptive Server has added numerous features that improve performance on Linux. The following table provides some of the key features that contribute to the ability of Adaptive Server to scale well on Linux.

Adaptive Server Performance Features on Linux
• Raw device and file system support • Ability to exploit large memory using file system caching • Asynchronous I/O on raw devices and file systems • Support for kernel AIO and POSIX AIO • Large file support for VLDB applications • Dynamic management of configuration parameters • Dynamic management of data caches

Adaptive Server Enterprise 12.5.1 provides additional performance features,including parallel recovery, dynamic cache tuning and derived tables. These features, together with the operational scalability features introduced in 12.5.0.3, such as logging optimizations for tempdb, sampling for update statistics, multiple tempdb, enhanced housekeeper, faster create database, and faster checkpoint, ensure that Adaptive Server Enterprise on Linux is a platform on which customers can reliably run their businesses.

International Contacts

Argentina
+5411 4313 4488

Australia
+612 9936 8800

Austria
+43 1 504 8510

Belgium
+32 2 713 15 03

Brazil
+5511 3046 7388

Bulgaria
+359 2 986 1287

Canada
+905 273 8500

Central America
+506 204 7151

Chile
+56 2 330 6700

China
+8610 6856 8488

Colombia
+57 1 218 8266

Croatia
+385 42 33 1812

Czech Republic
+420 2 24 31 08 08

Denmark
+45 3927 7913

Ecuador
+59 322 508 593

El Salvador
+503 245 1128

Finland
+358 9 7250 200

France
+33 1 41 91 96 80

Germany
+49 69 9508 6182

Greece
+30 1 98 89 300

Guatemala
+502 366 4348

Honduras
+504 239 5483

Hong Kong
+852 2506 6000

Hungary
+36 1 248 2919

India
+91 22 655 0258

Indonesia
+62 21 526 7690

Israel
+972 3 548 3555

Italy
+39 02 696 820 64

Ivory Coast
+225 22 43 73 73

Japan
+81 3 5210 6000

Kazakstan
+7 3272 64 1566

Korea
+82 2 3451 5200

Malaysia
+603 2142 4218

Mexico
+5255 5093 8500

Netherlands
+31 20 346 9290

New Zealand
+64 4473 3661

Nigeria
+234 12 62 5120

Norway
+47 231 621 45

Panama
+507 263 4349

Peru
+511 221 4190

Philippines
+632 750 2550

Poland
+48 22 844 55 55

Portugal
+351 21 424 6710

Puerto Rico
+787 289 7895

Romania
+40 1 231 08 70

Russian Federation
+7 095 797 4774

Slovak Republic
+421 26 478 2281

Slovenia
+385 42 33 1812

South Africa
+27 11 804 3740

South Korea
+82 2 3451 5200

Spain
+34 91 749 7605

Sweden
+46 8 568 512 00

Switzerland
+41 1 800 9220

Taiwan
+886 2 2715 6000

Thailand
+662 618 8638

Turkey
+90 212 325 4114

Ukraine
+380 44 227 3230

United Arab Emirates
+971 2 627 5911

United Kingdom
+44 870 240 2255

Venezuela
+58 212 267 5670

Asian Solutions Center
+852 2506 8700

For other Europe, Middle East, or Africa inquiries:
+33 1 41 90 41 64 (Sybase Europe)

For other Asia Pacific inquiries:
+852 2506 8700 (Hong Kong)

For other Latin America inquiries:
+925 236 6820

Sybase, Inc.
Worldwide Headquarters
One Sybase Drive
Dublin, CA 94568-7902 USA
E-mail: partnersuccess@sybase.com
www.sybase.com