



An Oracle White Paper
May 2010

Increase Application Scalability and Improve System Utilization with Oracle VM Server for SPARC

Introduction	1
Using Oracle VM Server for SPARC and CoolThreads Technology ..	2
Oracle's Virtualization Solution	3
CoolThreads Technology.....	3
The Hypervisor	5
Logical Domains	5
Oracle Solaris Containers.....	6
Configuration Considerations	7
Domains Scenarios	8
Sharing Processor Resources	8
Simple Server Consolidation	8
Using Domains for Security and Isolation	10
Sharing I/O Devices Among Domains.....	11
Validating the Performance Gains	12
Summary.....	14
Acknowledgements	15
References.....	15

Introduction

Modern enterprises need low cost per unit of capacity and linear scalability. Advances in technology have led to a new generation of servers with very high performance. Processors such as the UltraSPARC® T2 processor, incorporating chip multithreading (CMT) technology, support up to 64 threads (parallel computations) and 64 fully buffered dual in-line memory modules (FB-DIMMs) on a single system. These new system architectures offer substantial improvements in price/performance, as well as large potential operational savings and reduced energy consumption.

Most applications are optimized for single-thread processors, and don't take full advantage of multicore, multithread parallelism — and the capacity gains that parallelism offers. Common datacenter practices, such as configuring a single network interface for the entire application stack, introduce bottlenecks that further limit applications' realized capacity and scalability. Virtualization is one approach to solving these problems.

Oracle® Solaris CoolThreads technology combined with Oracle VM Server for SPARC (previously called Sun Logical Domains), allows a single UltraSPARC processor to be subdivided into many independent virtual processors. Each of these virtual processors is allocated one or more independent threads.

By virtualizing a single large multicore processor into many smaller virtual processors, the system realizes much greater capacity — by several-fold in the field, and experimentally in the lab as much as 10 times greater capacity. This increase in capacity is realized at a significantly lower cost than an equivalent system built using many smaller physical servers.

This Oracle white paper describes the application of Oracle VM Server for SPARC to implement high-capacity enterprise systems.

Using Oracle VM Server for SPARC and CoolThreads Technology

Processors such as the UltraSPARC T2 processor with CMT technology support up to 64 threads on a single chip. However, some applications do not currently scale linearly on servers that use these multicore processors. When the workload increases, bottlenecks in the application itself or in the I/O and network configuration of the system might cause the application to be unable to take advantage of the full power of the server. Currently, some commonly used open source applications are designed to scale on 1, 2, or 4 core systems, but do not scale well to 8 or 16 core systems.

To overcome the scaling issue imposed by the software stack, it is often necessary to run these applications on multiple smaller systems, or run multiple instances of the software stack on a larger system. Web applications, for example, can be configured with paired application and database servers, or might combine many small application servers with a single database server, as illustrated in Figure 1.

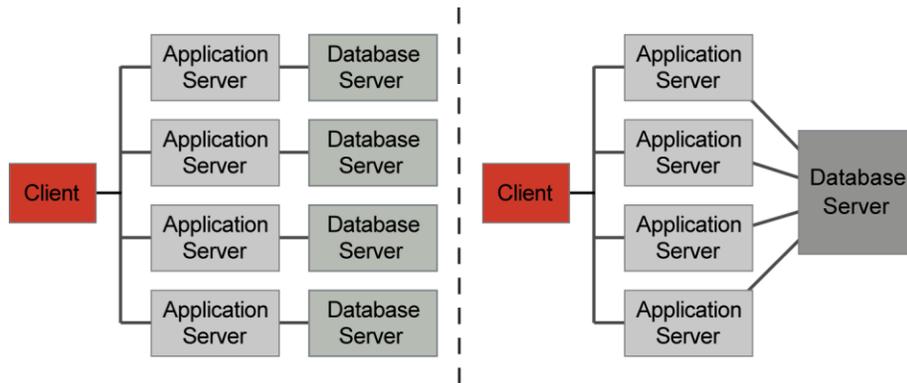


Figure 1. Two examples of multiple servers configured to support a Web application.

Running an application on multiple smaller systems can improve performance, but this configuration can present heating/cooling, power, and space challenges. In addition, managing multiple servers increases administration complexity and can increase management costs.

An alternative is to run multiple instances of the application server on a single larger system, as shown in Figure 2. Running multiple application instances on a single server can pose configuration and deployment challenges as well, such as memory and CPU resource allocation and load balancing. In addition, using a single operating system instance can mean the capacity of network-intensive applications, such as Web and application servers, is constrained by limited network resources.

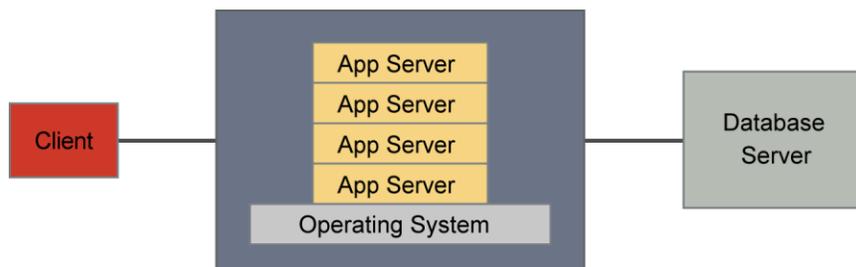


Figure 2. Multiple instances of one application running on a single server.

Using virtualization technology, such as Oracle VM Server for SPARC on a high-capacity CMT-based server is an effective solution to this scalability issue. Through this approach, multiple instances of the application can be deployed on multiple instances of the operating system — all on a single server. Using multiple operating system instances provides the advantage of multiple (virtual) network interfaces, avoiding complicated configurations. Because a single physical server is used, the configuration and management issues of deploying multiple servers are greatly reduced. Consolidating applications on a single energy-efficient server can also potentially save on total power and heating/cooling costs, reduce datacenter space requirements, and lower server management overhead.

Oracle's Virtualization Solution

Oracle has many virtualization solutions, with different features and goals. Oracle's servers with CoolThreads technology allow multiple Oracle Solaris operating system instances to be allocated across virtual processors or threads in the server's CPUs. These operating system instances are allowed to allocate memory, I/O, and CPU resources within the server, all under the control of the hypervisor. This technology allows applications that are not yet optimized for multicore processors to use a multicore processor efficiently.

With Oracle VM Server for SPARC, organizations deploying CMT technology can realize significant advantages in:

- Security
- Management
- Rack space
- Power consumption

CoolThreads Technology

Oracle Solaris CoolThreads technology is supported on servers that incorporate UltraSPARC T1, T2, and T2 Plus processors and run the Oracle Solaris operating system. The UltraSPARC T1 processor contains up to 8 processing cores with 4 threads (virtual processors) per core. The UltraSPARC T2 and T2 Plus processors contain 4, 6, or 8 processing cores with 4 or 8 threads per core, for a total of as much as 256 threads per processor. (See Table 1.)

UltraSPARC processors are available in a traditional form in Oracle's Sun SPARC Enterprise servers. In addition, the UltraSPARC T2 Plus processor is available in Oracle's Sun Blade T6340 and Sun Blade T6320 Server Modules. These blade servers can be configured using the Sun Blade 6000 and Sun Blade 6048 Chassis into very high density lights-out datacenter configurations of up to 48 Sun Blade servers and 96 UltraSPARC T2 Plus processors.

TABLE 1. REPRESENTATIVE SERVER CONFIGURATIONS

SYSTEM	PROCESSOR CONFIGURATION	MAXIMUM THREADS	MAXIMUM MEMORY
Sun SPARC Enterprise T5440 Server	1-4 UltraSPARC T2 Plus processor(s)	256	512 GB
Sun SPARC Enterprise T5420 Server	1-2 UltraSPARC T2 Plus processors(s)	128	256 GB
Sun SPARC Enterprise T5220 Server	1 UltraSPARC T2 Plus processor	64	128 GB
Sun SPARC Enterprise T5140 Server	1-2 UltraSPARC T2 Plus processor(s)	128	128 GB
Sun SPARC Enterprise T5120 Server	1 UltraSPARC T2 processor	64	128 GB
Sun Blade T6340 Server Module	1-2 UltraSPARC T2 Plus processor(s)	128	256 GB
Sun Blade T6320 Server Module	1 UltraSPARC T2 Plus processor	64	128 GB

In any CMT technology-based server, the number of domains that can be allocated is a function of the number of threads available, the available memory, and the specific requirements of the particular mix of applications supported.

At one end of the spectrum, all threads of all cores can be allocated to a single application. This configuration might be suited to specific, computationally-intensive programs that make efficient use of the intrinsic parallelism available in the UltraSPARC architecture. At the other end of the spectrum, it's theoretically possible to construct a domain for each individual thread, so that, for example, a single Sun SPARC Enterprise T5440 server could be configured with 256 independent domains, each with 2 GB of memory.

Practically, the best solution usually lies between these extremes. So, as in the example described below, a Sun SPARC Enterprise T5220 server with 64 threads might be configured to have 12 domains, with each of the domain operating system instances running a Web server such as Tomcat or an application server such as the GlassFish application server. This configuration can help fully utilize the server resources and provide significant advantages of management, rack space and power consumption.

Currently, the highest capacity CMT technology-based server available is the Sun SPARC Enterprise T5440 server. This server can be equipped with as many as four UltraSPARC T2 Plus processors and associated memory, for up to 256 threads and 512 GB of memory. This configuration is about four

times the capacity of the Sun SPARC Enterprise T5220 server used in the experiments from which the capacity estimates in this paper are derived. In addition, the Sun SPARC Enterprise T5440 server supports solid state devices (SSDs) based on Flash technology, that improve I/O throughput and application response time by as much as a factor of 70. At the same time, SSD storage consumes about 38% less power per gigabyte and 98% less power per I/O operation than traditional rotating media.

Features of the Sun SPARC Enterprise T5440 server include:

- Up to 256 simultaneous compute threads and 512 GB of memory
- World's first quad-socket, general-purpose server powered by the UltraSPARC T2 Plus system-on-a-chip processor
- Drive up utilization, drive down costs, and massively consolidate workloads with built-in, no-cost virtualization technology with domains and Oracle Solaris Containers
- Delivers outstanding performance and availability for mission-critical, back-office applications
- Reduces power and cooling costs, conserves datacenter space, improves service levels, increases server utilization, and maximizes return on IT investments
- Accelerate application response times and slash energy consumption with Flash SSDs

These features in combination make the Sun SPARC Enterprise T5440 server a powerful, cost-effective platform for datacenter consolidation, large-scale databases, back-office applications, and other demanding high-performance tasks.

The Hypervisor

Efficient virtualization is implemented through the hypervisor — a thin layer of firmware that abstracts physical hardware resources behind a virtual processor API. By using the hypervisor, no kernel or operating system changes are required to run virtual images. When the hypervisor is running on Oracle servers with CoolThreads technology, it enforces resource management at the hardware level and provides an interface between these resources and the operating systems software.

Logical Domains

Oracle VM Server for SPARC allows a system's hardware resources, such as memory and CPUs, to be subdivided among partitions called logical domains. Each domain has its own set of resources and runs its own independent instance of the operating system as shown in Figure 3.

A *control domain* (domain 0) runs Logical Domains Manager software, and must run Oracle Solaris. This software is used to create and manage additional domains called guest domains. In the example in Figure 3, each guest domain runs its own instance of any application software, as well as its own copy of the operating system, which in a guest domain can be any of several possibilities, including Oracle Solaris and Linux.

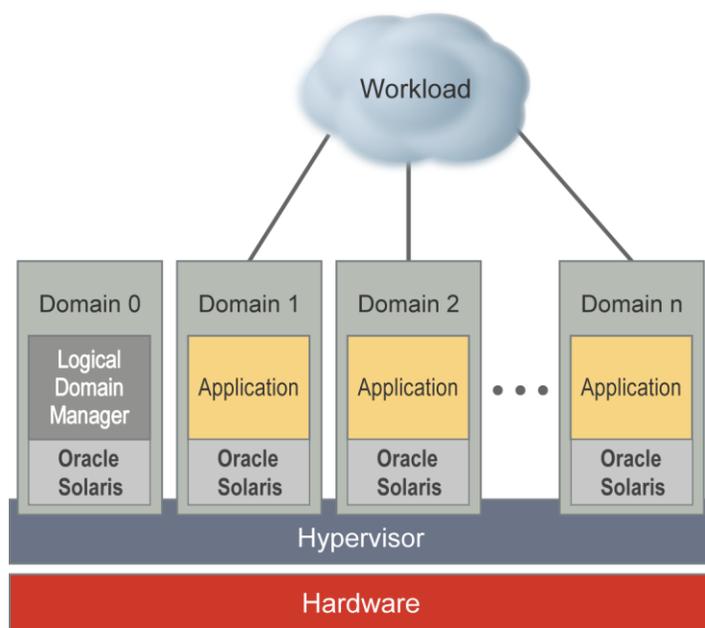


Figure 3. Oracle VM Server for SPARC running multiple domains.

The Logical Domains Manager software provides support for virtual devices, such as virtual disks and virtual networks. The actual physical hardware devices are abstracted by the hypervisor and presented to the domains on the system. With virtual disk devices, a physical device can be partitioned and shared by multiple domains. Communication between the domains and the physical devices is routed through the hypervisor. Virtual network switches and virtual network devices are used to abstract physical network interfaces and enable communications for the domains.

Domains provide a reliable, robust, and secure virtualization solution. Multiple guest domains provide strong workload isolation — a failure in one guest domain does not bring down other guest domains on the system. In addition, the hypervisor, which has access to all memory and I/O devices, is not accessible from the guest domains, or even from the control domain. This isolation increases security, and prevents unauthorized access to data and resources. As long as traditional OS security remains uncompromised on the server, no domain can access the content of another domain.

Note — For more detail about domains, see the *Beginners Guide to LDOMs: Understanding and Deploying Logical Domains for Logical Domains 1.0 Release*, available on the Web at <http://wikis.sun.com/display/BluePrints/Beginners+Guide+to+LDoms+1.0>

Oracle Solaris Containers

Domains are independent of other Oracle Solaris virtualization technologies. Oracle Solaris Containers can be used as well to virtualize a single kernel instance of Oracle Solaris, providing independently-managed entities sharing the resources of an Oracle Solaris instance. In total, Oracle's virtualization solution provides four levels of potential virtualization:

- UltraSPARC servers with CMT technology provide the capability of many parallel threads of computation, supported by the hardware.

- The hypervisor manages and allocates the system's hardware resources to software processes.
- Domains allow the hypervisor's managed resources to be allocated to virtual systems, with independent virtual CPUs and devices.
- Finally, within a single domain, Oracle Solaris Containers allow one more level of virtualization, with each container providing its own access control and security. Solaris Resource Manager can be used to further allocate resources within an individual virtual machine to various containers.

A CoolThreads technology-based server can provide multiple layers of virtualization, with each layer adding significant features to allow systems administrators to use the system to support applications effectively.

Configuration Considerations

Before attempting to configure the virtualized environment, it is important to estimate how many guest domains should be created. The total number of domains that can be created is platform dependent. At the lowest level of granularity, a domain requires at least one processor thread. Thus, a processor with 64 threads theoretically can be configured with a maximum of 64 domains. In practice, contention for other resources limits the useful number of domains that can be configured.

In real world applications, the number of guest domains that can be configured depends largely on the nature of the application. For example, in some cases it might be possible to share the same physical network interface between two or more guest domains, while in other cases, each guest domain would need a dedicated virtual network interface created on top of a dedicated physical network interface. Similarly, the amount of memory that an application server needs to perform at an acceptable level varies and might be a more limiting factor than the available number of CPU threads.

The requirements of a single application can be used to estimate a suitable number of guest domains. This information can easily be obtained by running a single instance of the application in a non-virtualized environment on a system similar to the target virtualization system.

During run time, collect the following information:

- Amount of memory required by the application (M)
- Utilization of the network interfaces used (N)
- Amount of CPU utilized (C)
- Disk space requirement (D)

Then, use this information to find the appropriate hardware that meets these performance requirements and budgetary constraints.

As an example, if disks (D) and network (N) resources are available in sufficient quantities, then consider CPU (C) and memory utilization (M) to determine the target number of guest domains. For example, if the application needs 10% of the available memory, then the maximum number of domains that can be configured is 10. However, if the percentage of CPU utilization is higher than the

percentage of memory utilization, then the CPU utilization is the limiting factor. For instance, if the memory utilization stays at 10% and 20% CPU is needed to run a single instance of the application with satisfactory results, then the maximum number of guest domains that can be configured is limited to five (each with 20% of the CPU resources).

Domains Scenarios

Making the best use of domains depends on the specific requirements of the application. To illustrate some of the options domains offer in an enterprise architecture, this section shows several possible requirements, and configuration scenarios to satisfy them.

Sharing Processor Resources

The simplest examples illustrating the advantages of domains are server consolidation projects. Many datacenters have existing server farms consisting of many servers running different applications and environments, often for different customers. Managing large numbers of servers is complicated, and collections of many small servers inevitably means that many redundant components must be purchased and maintained. A common — and effective — cost reduction strategy is to reduce the number of physical servers by consolidating applications into a smaller number of servers with greater capacity.

Simple Server Consolidation

The example in Figure 4 consolidates three servers into one. The existing environment consists of:

- One server running Oracle Solaris and an accounting program
- One server running Oracle Solaris and a proprietary database
- A third server running Linux, an open-source application, and a Web server

The consolidated server is configured with three domains: FIN for the financial application, DB for the database application, and WEB for the Linux system. The DB domain is allocated more memory and a greater share of CPU resources.

Existing SPARC processor-based, independent software vendor (ISV) applications move directly to the FIN and DB domains because the UltraSPARC T1 powered servers have full SPARC binary compatibility. As the WEB domain is an open-source environment, the open-source applications can be recompiled for a SPARC processor-based system.

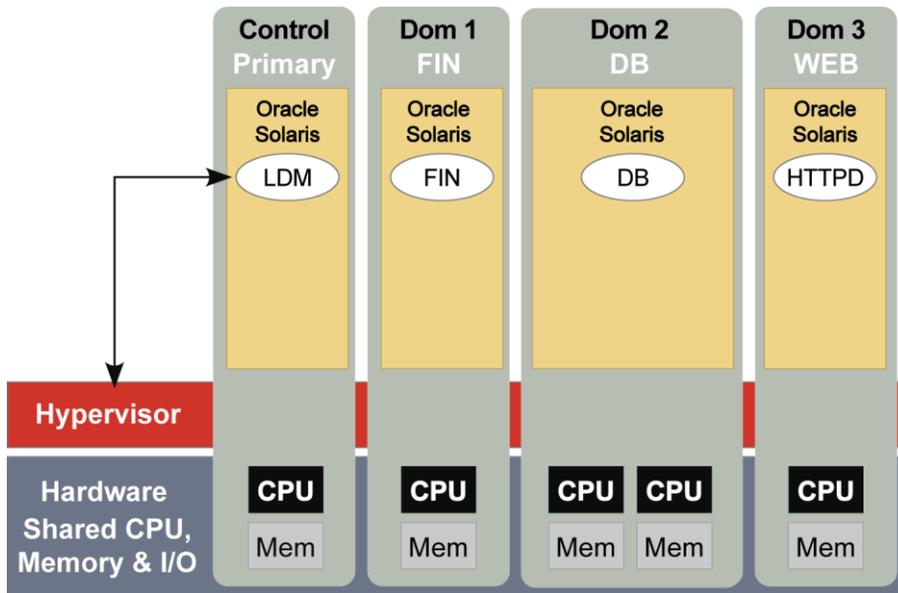


Figure 4. Many servers consolidated in a single system with domains.

Multiple Environments on a Single Server

In many enterprises, staging new applications requires several replicated servers: a development server, a quality assurance server, and a test server. Since both operating system changes and application changes need to be tested, this means the environment must support multiple versions of the operating system concurrently. Figure 5 shows such an environment, as it might be consolidated in a CoolThreads technology server using domains.

TEST, DEVELOPMENT, and QA environments run on the same machine. TEST runs Oracle Solaris N and DEVELOPMENT and QA run Oracle Solaris N-1. The Logical Domains Manager provides the capability to run different versions of the OS, as each virtualized system has its own, completely separate, operating system.

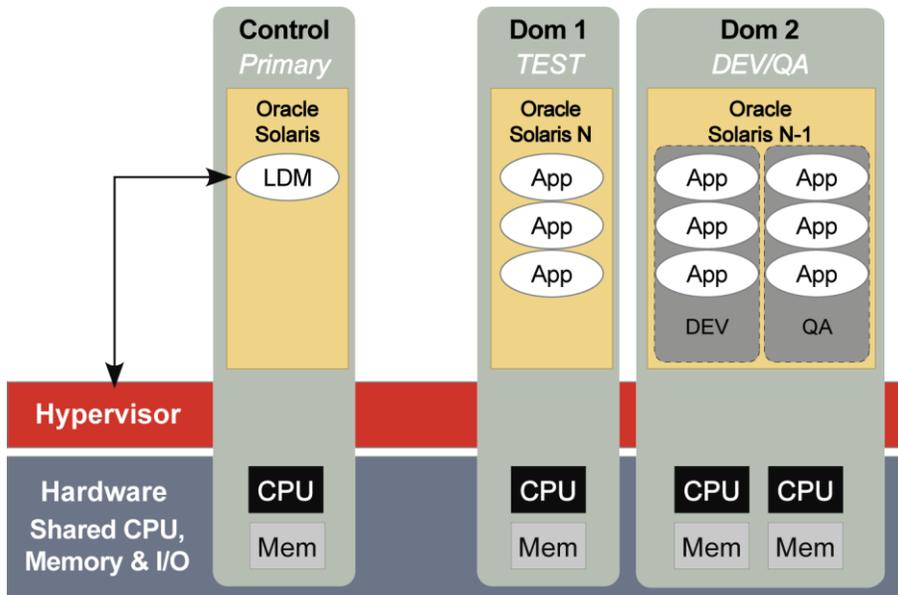


Figure 5. Multiple environments consolidated using domains.

Using Domains for Security and Isolation

Secure environments are a more and more important issue for enterprise datacenters. Security is an ongoing crisis, as attackers ranging from hobbyists to serious, malicious attackers with criminal intent continually challenge all but the simplest systems. At the same time, legal and regulatory changes mean that the risk to the enterprise is greater.

One of the most effective methods of securing a large system is to build firewalls between different functional areas so that attacks on one area can't compromise others. In this scenario, (see Figure 6) sensitive data contained in the directory and database domains would be isolated from the Web servers. The database server Web servers in this illustration are also in separate containers, which allow this configuration to take advantage of another layer of isolation within a single Oracle Solaris instance. In this example, three environments are required: one for a corporate directory server (DIR), another for a database with an application server (DB), and another for multiple Web servers (WEB). Isolation among the different logical domains is enforced by hardware features in the CoolThreads technology-based servers, further strengthened by features of the hypervisor and Logical Domain Manager, and finally enforced by the security features of Oracle Solaris in each domain.

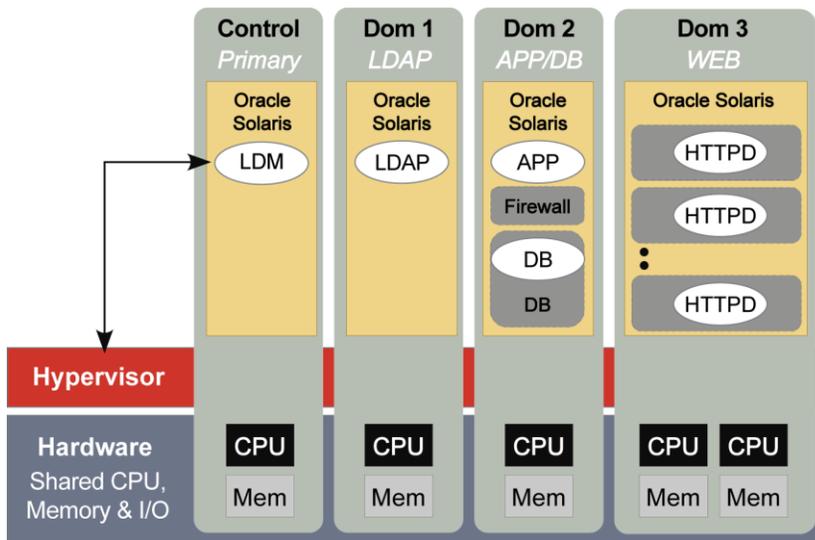


Figure 6. Domains help isolating functions for greater security.

Sharing I/O Devices Among Domains

Even virtual systems must eventually have access to real physical I/O devices. Oracle VM Server for SPARC allows single domains to have direct access to physical devices, but also allows the system administrators to construct virtual devices and virtual networks that appear to be real and independent devices to the domains. In reality, the domains transparently share the physical devices.

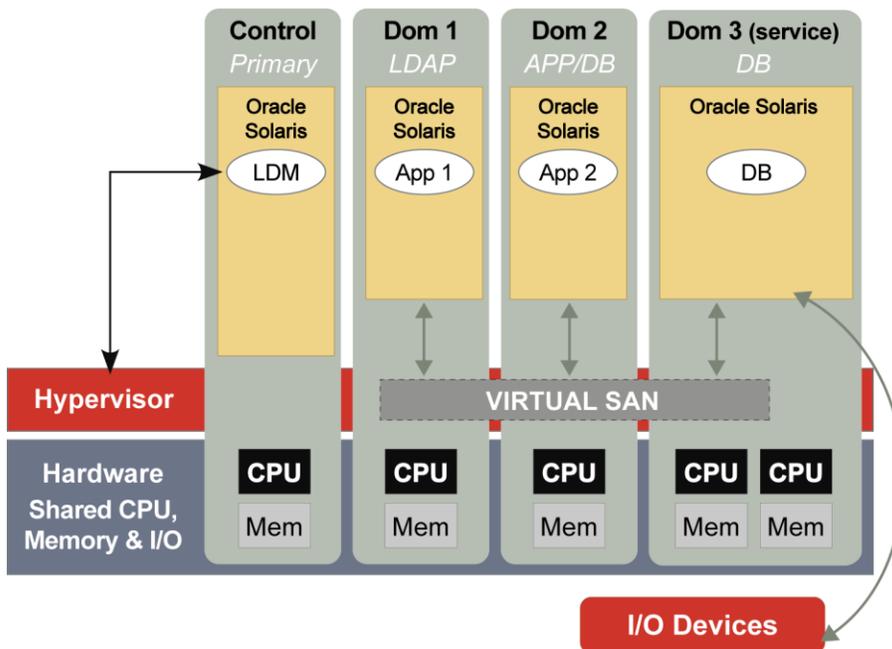


Figure 7. Shared virtual and physical I/O devices.

Virtual and Physical Devices

In this scenario, (Figure 7) the enterprise requires several environments. Some need to access devices directly, such as a database environment and a disk subsystem. Here the domains software offers some flexibility in how devices can be accessed, including virtual access and direct access. It is also possible to create a virtual storage-area network (SAN) and local-area network (LAN), allowing domains to share devices effectively.

Virtual devices are any hardware resources on the system that are abstracted by the hypervisor and presented to the domains on the system. They can take the form of physical devices partitioned by the hypervisor, such as CPU, memory, and I/O busses, and those devices that are provided from a service domain for use by other domains. Physical devices are translated to virtual devices by the hypervisor and provided by an I/O-service domain to other domains.

In this scenario, a domain runs a database (DB) and has direct access to the devices located under the peripheral component interconnect (PCI) controller. The initial domain then shares access to these devices through two other domains, APP1 and APP2, creating a virtual SAN.

Validating the Performance Gains

In order to validate the potential scalability advantages of CoolThreads technology and domains in a production environment, a test environment with synthetic workload was built and measured, using a Sun SPARC Enterprise T5220 server, using the JPetStore application, Tomcat application server, the MySQL database, and a synthetic workload generated by the Grinder Java™ load testing framework. These experiments were performed with two test configurations, using 6 and 12 guest domains respectively.

Note — For full details, please see *Using Logical Domains and Coolthreads Technology: Improving Scalability and System Utilization*, Sun BluePrints Online, April 2008. <http://wikis.sun.com/display/BluePrints/Using+Logical+Domains+and+CoolThreads+Technology>

Evaluating Capacity Gains from Domains

The throughput results for testing with six domains were evaluated. These results indicate a baseline rate of 952 transactions per second (TPS) was achieved with one guest domain. With six guest domains, the number of TPS increased by a factor of 5.84 to 5560 TPS.

Figure 8 illustrates these results graphically, showing linear throughput scalability as the number of domains was increased from one to six. The average throughput per domain decreases by only a small percentage as the number of domains increases, with the six domain configuration averaging 927 TPS per domain as compared to the 952 TPS for a single domain configuration.

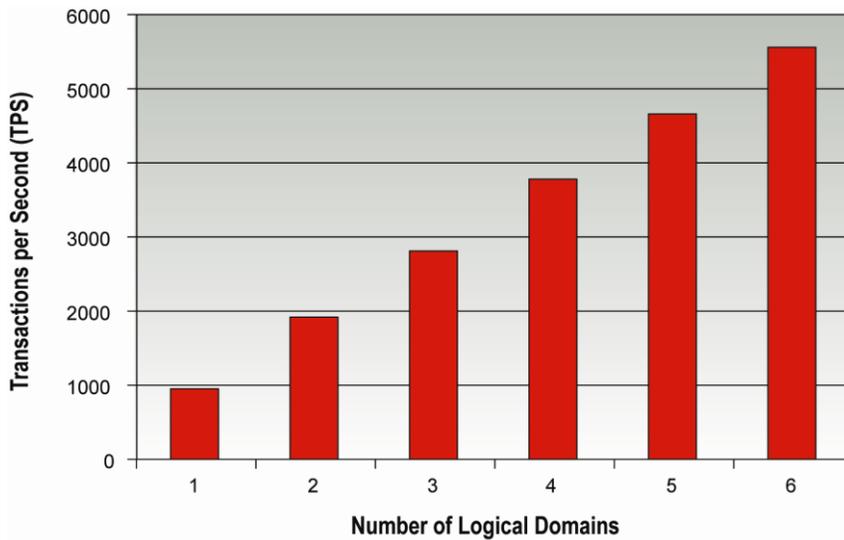


Figure 8. Experimental throughput measured for 1–6 logical domains.

Performance results for testing with 12 domains are included in Figure 9. CPU utilization is aggregated against the 48 strands that were assigned to the 12 guest domains.

These results show that in the laboratory, configurations with 6 logical domains exhibited scalable performance improvements and still did not fully utilize the system resources of the SPARC Enterprise T5220 server from Oracle. A configuration with 12 logical domains increased the overall throughput by over 50% compared to the six-domain configuration, while almost fully utilizing the available CPU resources assigned to the logical domains.

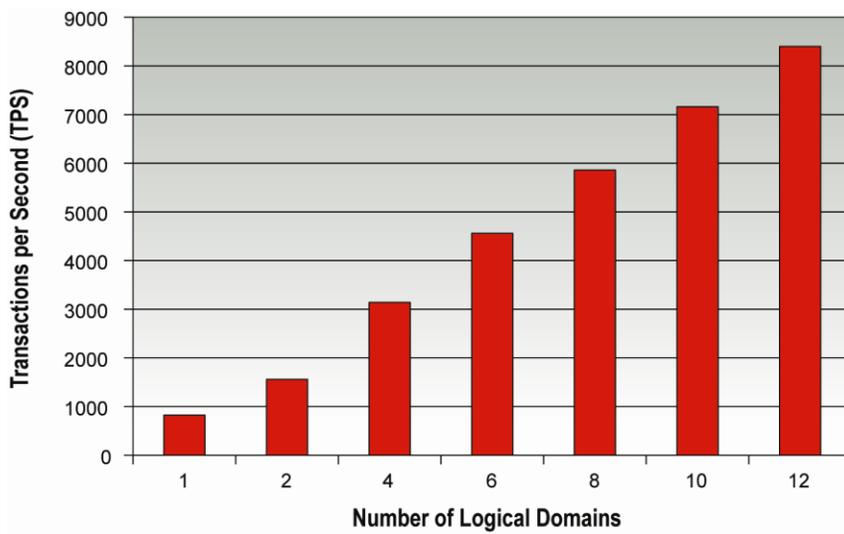


Figure 9. Experimental throughput measured for 1–12 logical domains.

Summary

While supporting new applications, and the need for greater capacity for existing applications, information technology management also needs to reduce equipment costs, operating costs, and environmental impact.

Oracle's CMT technology on UltraSPARC processor-based servers, along with domains, provide new and powerful tools that offer greater flexibility, easier management, and the opportunity to reduce costs.

These features build on one another:

- CoolThreads technology, to allow a single processor core to serve several threads of computation in parallel
- The hypervisor, to manage the allocation of processors and system resources among many cores with many parallel threads of computation
- Oracle VM Server for SPARC, to provide many independent virtual systems, each with their own I/O devices and operating system from the kernel up
- Oracle Solaris Containers, which allow several virtualized systems to share a single kernel within a logical domain, providing even greater flexibility and enhanced ease of management

Oracle's CoolThreads technology allows system managers to make efficient use of powerful multicore processors, even when applications have not yet been optimized to take advantage of new technology. This provides substantial increases in throughput: up to 10 times greater throughput for these tests.

Using CoolThreads technology and domains provides a new tool set for provisioning and administering enterprise datacenters by consolidating many different environments, and by allowing current software to efficiently leverage powerful UltraSPARC processors. Fewer servers to administer, to maintain, and to purchase, means reduced datacenter footprint, improved performance, reduced energy consumption, and better return on investment.

Acknowledgements

This white paper is derived primarily from two sources: Using Logical Domains and CoolThreads Technology: Improving Scalability and System Utilization, by Ning Sun and Lee Anne Simmons, and Beginner's Guide to LDomS: Understanding and Deploying Logical Domains, by Tony Shoumack. Their contributions are gratefully appreciated.

References

Logical Domains (LDoms) 1.0.1 Administration Guide. Sun Microsystems, Inc., September 2007, Revision A. Part No. 820-3268-10.

Shoumack, Tony. Beginner's Guide to LDomS: Understanding and Deploying Logical Domains, Sun BluePrints OnLine, July 2007. <http://www.sun.com/blueprints/0207/820-0832.pdf>

Sun, Ning and Lee Anne Simmons. Using Logical Domains and Coolthreads Technology: Improving Scalability and System Utilization, Sun BluePrints Online, April 2008.
<http://wikis.sun.com/display/BluePrints/Using+Logical+Domains+and+CoolThreads+Technology>



Increase Application Scalability and Improve
System Utilization with Oracle VM Server for
SPARC
May 2010

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 © 2010, 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. 0410

SOFTWARE. HARDWARE. COMPLETE.