SUN SPARC ENTERPRISE SERVERS

An Oracle White Paper
September 2010

Deploying Oracle Database on Oracle Sun SPARC Enterprise M-series Systems

Executive Summary

This document is intended for IT architects, system administrators, and developers who want to understand the details of how Oracle® Solaris and Oracle® Sun SPARC Enterprise M-series Servers combine to provide an ideal enterprise deployment environment for Oracle Database.

This paper provides technical information on the Oracle Sun SPARC Enterprise M-series systems and the mainframe-class RAS features and resource management capabilities that they offer. It also explores the virtualization capabilities provided by Oracle Solaris and the Oracle Sun Enterprise M-series Systems, as well as offers recommendations on how to combine virtualization technologies for enterprise deployments of Oracle Database.

Introduction

Today, organizations rely on information technology more than ever before. Computer systems, and the information that they manage, play a critical role in nearly every function in modern enterprise, from financial accounting, parts logistics, and order fulfillment to product design and marketing. In many cases, business success is dependent on the continuous availability of IT services.

While availability is a top priority, costs must also remain in budget and operational familiarity maintained. To deliver networked services as efficiently and economically as possible, organizations look to maximize use of every IT asset through consolidation and virtualization strategies. As a result, modern IT system requirements reach far beyond simple measures of compute capacity. Highly flexible servers are required with built-in virtualization capabilities and associated tools, technologies, and processes that work to optimize sever utilization. New computing infrastructures must also help protect current investments in technology and training.

Oracle's Sun SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 servers are highly reliable, easy to manage, vertically-scalable systems with many of the benefits of traditional mainframes — without the associated cost or complexity. In fact, these servers deliver a mainframe-class system architecture at open systems prices. With symmetric multiprocessing (SMP) scalability from one to 64 processors, memory subsystems as large as 4 TB, and high-throughput I/O architectures, Sun SPARC Enterprise M-series servers easily perform the heavy lifting required by consolidated workloads and data warehouse consolidation. Furthermore, these servers run the powerful Oracle Solaris 10 operating system and include leading virtualization technologies. By offering Dynamic Domains, eXtended System Boards, Dynamic Reconfiguration, and Oracle Solaris Containers technology, Sun SPARC Enterprise M-series servers bring sophisticated mainframe-class resource control to an open systems compute platform.

Oracle Sun SPARC Enterprise M-series Servers1

Overview of Oracle Sun SPARC Enterprise M-series Systems

The members of the Sun SPARC Enterprise server family share many of the same capabilities that foster power, reliability, and flexibility. Sun SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 servers feature a balanced, highly-scalable SMP design that utilizes the latest generation of SPARC64 processors connected to memory and I/O by a high-speed, low latency system interconnect that delivers exceptional throughput to applications. Also architected to reduce planned and unplanned downtime, these servers include reliability, availability, and serviceability capabilities to avoid outages and reduce recovery time. Design features, such as advanced CPU integration and data path integrity, memory extended ECC, and memory mirroring, end-to-end data protection, hot-swappable components, fault resilient power options, and hardware redundancy boost the reliability of these servers.

Sun SPARC Enterprise M4000, M5000, M8000, and M9000 servers also provide considerable configuration flexibility. As in other Sun high-end servers, administrators can use Dynamic Domains to physically divide a single Sun SPARC Enterprise M4000, M5000, M8000, or M9000 server into multiple electrically-isolated partitions, each running independent instances of Oracle Solaris. Hardware or software failures in one Dynamic Domain do not affect applications running in other domains, unless the failed resource is shared across those domains.

Dynamic Reconfiguration can then reallocate hardware resources among Dynamic Domains — without interrupting critical systems. Oracle Sun SPARC Enterprise M4000, M5000, M8000, and M9000 servers advance resource control one-step further with eXtended System Board (XSB) technology, supporting the allocation of sub-system board resources such as CPUs, memory, and I/O components to Dynamic Domains. The fine-grained resource control provided by eXtended System Board technology helps enterprises to further optimize resource utilization.

Adding even more value, the range of compute power offered by the Sun SPARC Enterprise server family provides the levels of vertical scalability required for many deployment classes, letting organizations match the right system to the task. Rack-mount Sun SPARC Enterprise M3000, M4000, and M5000 servers are economical, powerful, and reliable servers that are well-suited for entry-level and mid-range system requirements (Table 1). Sun SPARC Enterprise M8000 and M9000 servers deliver the processing power needed for high-end computing (Table 2).

[1] Some references to server names are abbreviated for readability. For example, if you see a reference to the SPARC Enterprise M9000 server or simply the M9000 server, note that the full product name is the Sun SPARC Enterprise M9000 server.

TABLE 1 CHARACTERISTICS OF SUN SPARC ENTERPRISE M3000, M4000, AND M5000 SERVERS			
	SUN SPARC ENTERPRISE M3000 SERVER	SUN SPARC ENTERPRISE M4000 SERVER	SUN SPARC ENTERPRISE M5000 SERVER
ENCLOSURE	2 rack units	6 rack units	10 rack units
SPARC64 VI PROCESSORS	• N/A	• 2.15 GHz • 5 MB L2 cache • Up to four dual-core chips	• 2.15 GHz • 5 MB L2 cache • Up to eight dual-core chips
SPARC64 VII PROCESSORS	• 2.75 GHz • 5 MB L2 cache • 1 dual or quad core	• 2.4 GHz with 5 MB L2 cache • 2.53 GHz with 5.5 MB L2 cache • Up to four quad-core chips	• 2.4 GHz with 5 MB L2 cache • 2.53 GHz with 5.5 MB L2 cache • Up to eight quad-core chips
MEMORY	• Up to 64 GB • Eight DIMM slots	• Up to 256GB • 32 DIMM slots	• Up to 512GB • 64 DIMM slots
INTERNAL I/O SLOTS	• Four PCIe	• Four PCIe • One PCI-X	• Eight PCIe • Two PCI-X
EXTERNAL I/O CHASSIS	One external x2 SAS port	Up to two units	Up to four units
INTERNAL STORAGE	• Serial Attached SCSI • Up to four drives	• Serial Attached SCSI • Up to two drives	• Serial Attached SCSI • Up to four drives
DYNAMIC DOMAINS	One	Up to two	Up to four

TABLE 2 CHARACTERISTICS OF SUN SPARC ENTERPRISE M8000, M9000-32, AND M9000-64 SERVERS			
	SUN SPARC ENTERPRISE M8000 SERVER	SUN SPARC ENTERPRISE M9000-32 SERVER	SUN SPARC ENTERPRISE M9000-64 SERVER
ENCLOSURE	One cabinet	One cabinet	Two cabinets
SPARC64 VI PROCESSORS	• 2.28 GHz with 5 MB L2 Cache • 2.4 GHz with 6 MB L2 cache • Up to 16 dual-core chips	• 2.28 GHz with 5 MB L2 cache • 2.4 GHz with 6 MB L2 cache • Up to 32 dual-core chips	• 2.28 GHz with 5 MB L2 cache • 2.4 GHz with 6 MB L2 cache • Up to 64 dual-core chips
SPARC64 VII PROCESSORS	• 2.52 GHz and 2.88 GHz • 6 MB L2 cache • Up to 16 quad-core chips	• 2.52 GHz and 2.88 GHz • 6 MB L2 cache • Up to 32 quad-core chips	• 2.52 GHz and 2.88 GHz • 6 MB L2 cache • Up to 64 quad-core chips
MEMORY	• Up to 1TB • 128 DIMM slots	• Up to 2TB • 256 DIMM slots	• Up to 4TB • 512 DIMM slots
INTERNAL I/O SLOTS	32 PCI Express	64 PCI Express	128 PCI Express
EXTERNAL I/O CHASSIS	Up to 8 units	Up to 16 units	Up to 16 units
INTERNAL STORAGE	• Serial Attached SCSI • Up to 16 drives	• Serial Attached SCSI • Up to 32 drives	• Serial Attached SCSI • Up to 64 drives
DYNAMIC DOMAINS	Up to 16	Up to 24	Up to 24

SPARC VI and VII Processors

The Sun SPARC64 VI and VII processors incorporate innovative multicore and multithreaded technology, while also providing extensive reliability features. In addition, the SPARC64 VI and SPARC64 VII processors are SPARC V9 level 2 compliant, helping to provide support for thousands of existing software applications.

The use of SPARC64 VI and SPARC64 VII processors within Sun SPARC Enterprise M-series servers offers organizations exceptional reliability, application choice, and outstanding processing power. For Oracle databases, the chip design provides significant performance gains from parallelization, while maintaining deep instruction pipelines and fast clock speeds. This combination provides excellent single-threaded performance for database objects that have not been optimized for parallel execution, as well as the ability to co-host other applications that may suffer from lack of parallelization.

Chip Multiprocessing (CMP) - SPARC VI and VII

Both the SPARC VI and VII chips use Chip Multiprocessing (CMP) technology. CMP is an architecture in which multiple physical cores are integrated on a single processor module. Each physical core runs a single execution thread of a multithreaded application independently from other cores at any given time. With this technology, dual-core processors often double the performance of single-core modules. The ability to process multiple instructions at each clock cycle provides the bulk of the performance advantage, but improvements also result from the short distances and fast bus speeds between chips as compared to traditional CPU to CPU communication.

Vertical Multithreading (VMT) – SPARC VI

The SPARC VI chip uses CMP and a two-way course-grained threading technique called Vertical Multithreading (VMT). In the VMT threading model, each core is able to simultaneously run two threads, which are viewed by the operating system as a virtual CPU. Multiple threads on the same core run in a time-sliced fashion, with only one executing at any given moment. A thread does not run if it is idle, or has encountered a cache miss and is waiting for main memory. A thread switch occurs on events such as an L2 cache miss, hardware timer, interrupt, or specific instruction to control threads. In this way, VMT improves system performance by maximizing processor utilization and effectively mitigating the impact of a cache miss. VMT is enabled automatically to improve performance when the number of threads in the system exceeds the number of cores.

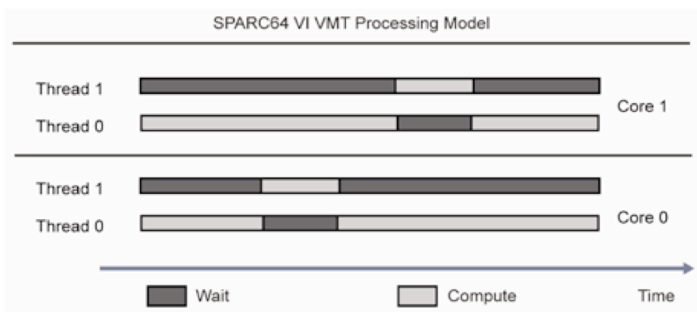

Figure 1: SPARC64 VI VMT Processing Model

Simultaneous Multithreading – SPARC VII

The SPARC VII chips use CMP and an improved threading technique called Simultaneous Multithreading (SMT). SMT technology supports the simultaneous execution of multiple threads in each core, reducing core wait times and increasing core utilization. From the software point of view, each thread is independent. This method of multithreading is facilitated by duplicating compute resources. A few resources remain shared between threads, such as instruction buffers, Commit Stack Entry (CSE), Branch History Storage (BRHIS), and caches.

With SMT, context switch time is eliminated and threads within a single core share the instruction pipeline smoothly. When both threads are ready to run, they alternate cycles for superscalar instruction issue, and share the functional units according to need.

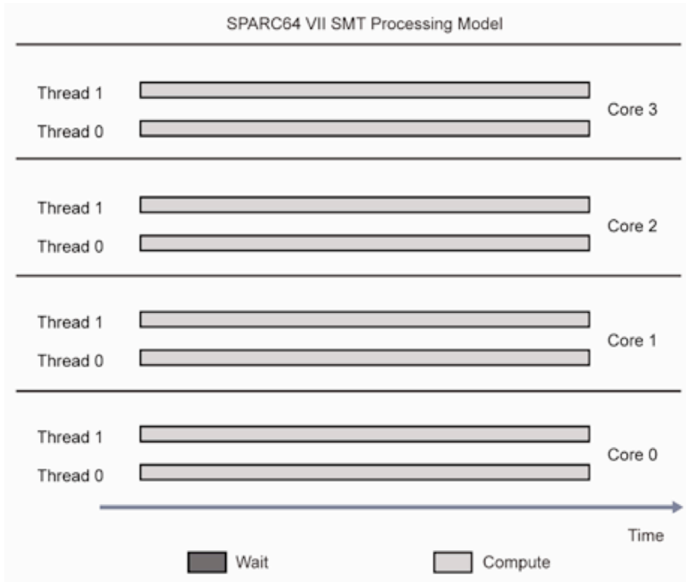

Figure 2: SPARC64 VII SMT Processing Model

Reliability, Availability, and Serviceability Features of the SPARC64 VI and VII Processors

The design of the SPARC64 VI and SPARC64 VII processor modules increase system reliability by delivering improved fault avoidance and error correction capabilities. In fact, significant footprint on the SPARC64 VI and SPARC64 VII processors is dedicated to error detection and data correction within the CPU. RAM units are ECC protected or duplicated, and most latches and execution units are parity protected. Rather than force the loss of operation of the entire processor, a single bad core can be isolated and taken offline.

Other reliability features of the SPARC64 VI and SPARC64 VII processors include support for error marking, instruction retry, and preventive maintenance. When memory read data has a multi-bit error, a special mark identifying the source of the error is written into the data and the ECC syndrome becomes a special value, providing valuable information for identifying the exact source of the fault. In addition, when a hardware error is detected, instructions that are currently in execution are cancelled, and retried automatically to prevent transient errors. Error data generated by the SPARC64 VI and SPARC64 VII processors is also sent to the service processor to support preventive maintenance.

SPARC64 VI and VII Socket Compatibility and Server Upgrades

SPARC64 VI and VII processors are socket compatible within the Oracle Sun Enterprise M4000, M5000, M8000, and M9000 servers. This provides an added level of investment protection and solution flexibility. SPARC64 VI and SPARC64 VII processors can co-exist within physical system boards, individual Dynamic Domains, and eXtended System Boards, without any impact on hosted applications, including Oracle Database.

I/O Subsystem

The use of PCI technology is key to the performance of the I/O subsystem within Sun SPARC Enterprise servers. A PCI Express bridge supplies the connection between the main system and components of the I/O unit, such as PCI-X slots, PCI Express slots, and internal drives. The PCI Express bus also supports the connection of external I/O devices by using internal PCI slots or connecting an External I/O Expansion Unit (IOU).

In order to facilitate hot-plug of PCI Express and PCI-X adapter cards, Sun SPARC Enterprise M4000, M5000, M8000, and M9000 servers utilize PCI cassettes. PCI cards which support PCI Hot Plug can be mounted by administrators into a PCI cassette and inserted into an internal PCI slot or External I/O Expansion Unit of a running server.

In addition to internal IOUs, Sun SPARC Enterprise M4000, M5000, M8000, and M9000 servers support the attachment of an optional External I/O Expansion Unit to provide additional I/O connectivity. The External I/O Expansion Unit is a four RU rack mountable device that accommodates up to two IOUs with six PCI Express or PCI-X slots. By using PCI cassettes, the external I/O chassis supports active replacement of hot-plug PCI cards.

TABLE 3: SUN SPARC ENTERPRISE M-SERIES I/O SUBSYSTEM EXPANSION CAPABILITIES

	Maximum Number of Internal PCIe	Maximum Number of External I/O Expansion Units	Maximum Number of PCI Slots
M3000	4 (plus 1 external SAS port)	0	4
M4000	4 (plus 1 PCI-X)	2	25
M5000	8 (plus 2 PCI-X)	4	50
M8000	32	8	112
M9000-32	64	16	224
M9000-64	128	16	288

For Oracle Database installations, this means that the I/O subsystem can provide massive bandwidth to attached storage. At 8 GB/s/channel, an Oracle SPARC Enterprise M9000-64 with 16 External I/O Expansion Units can provide a combined 2.304 TB/s transfer rate, using hot-plug hardware.

Advanced Reliability, Availability, and Serviceability (RAS)

Reducing downtime — both planned and unplanned — is critical for IT services. System designs must include mechanisms that foster fault resilience, quick repair, and even rapid expansion, without impacting the availability of key services. Specifically designed to support complex, network computing solutions and stringent high-availability requirements, Sun SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 servers include redundant and hot-swap system components, diagnostic and error recovery features throughout the design, and built-in remote management features. The advanced architecture of these reliable servers fosters high levels of application availability and rapid recovery from many types of hardware faults, simplifying system operation and lowering costs for enterprises.

Redundant Hot-Swap Components

The Sun SPARC Enterprise M-series servers feature redundant, hot-swap power supply and fan units, as well as the option to configure multiple CPUs, memory DIMMs, and I/O cards. Administrators can create redundant internal storage by combining hot-swap disk drives with disk mirroring software. High-end servers also support redundant, hot-swap service processors, and degradable Crossbar Units and Sun SPARC Enterprise M9000 servers include redundant Clock Control Units. If a fault occurs, these duplicated components can support continued operation. Depending upon the component and type of error, the system may continue to operate in a degraded mode or may reboot — with the failure automatically diagnosed and the relevant component automatically configured out of the system. In addition, hot-swap hardware within these servers speeds service and allows for simplified replacement or addition of components, without a need to stop the system.

Dynamic Domains and eXtended System Boards

In order to reduce costs and administrative burden, many enterprises look to server consolidation. However, organizations require tools that increase the security and effectiveness of hosting multiple applications on a single server. Dynamic Domains provide IT organizations with the ability to divide a single large system into multiple, fault-isolated servers each running independent instances of the Oracle Solaris operating system. With proper configuration, hardware or software faults in one domain remain isolated and unable to impact the operation of other domains. Each domain within a single server platform can even run a different version of Oracle Solaris, making this technology extremely useful for pre-production testing of new or modified applications.

Dynamic Domains provide a very effective tool for consolidation and facilitate the ideal separation of resources. In order to achieve this high level of isolation, previous generations of Sun servers designated entire system boards as the smallest unit assignable to a domain. However, some organizations do not require complete hardware isolation and can benefit from the ability to create a higher number of domains with compute power that more precisely matches current workloads. To meet these needs, the Sun SPARC Enterprise M4000, M5000, M8000, and M9000 servers introduce support for eXtended System Boards (XSB).

In the context of Oracle Database, and especially important for RAC deployments, Dynamic Domains can be used in conjunction with Oracle Solaris virtualization technologies to provide a full range of hardware and software-based virtualization that delivers deployment flexibility that is unmatched in the industry. Please refer to the *Virtualization Options for Sun SPARC Enterprise M-series Systems* section of this paper for further information.

Dynamic Reconfiguration

Dynamic Reconfiguration technology provides added value to Dynamic Domains by providing administrators with the ability to shift resources without taking the system offline. This technology helps administrators perform maintenance, live upgrades, and physical changes to system hardware resources, while the server continues to execute applications. Dynamic Reconfiguration even supports multiple simultaneous changes to hardware configurations without interrupting critical systems.

The ability to remove and add components such as CPUs, memory, and I/O subsystems from a running system helps reduce system downtime. Using Dynamic Reconfiguration simplifies maintenance and upgrades by eliminating the need for system reboots after hardware configuration changes.

Storage Options for Oracle Solaris SPARC

As mentioned in the *I/O Subsystem* section of this document, the Oracle Sun Enterprise M-series Systems provide massive bandwidth to back-end storage, via hot-pluggable hardware.

Oracle also offers a full range of storage solutions to connect to the Oracle Sun Enterprise M-series Systems, including flash storage arrays, the Oracle Sun Flash Accelerator PCIe card, disk arrays, and unified storage arrays.

Disk Storage

Oracle Sun Storage Fibre Channel Arrays provide mission-critical, tier-1 storage that scales in every dimension. Tailor performance profiles to site and application specific needs to achieve higher application performance while reducing power, space, and cooling requirements.

RAID 0, 1, 3, 5, 6 (p+q), and 10 support matches performance and availability to any mix of application requirements. As indicated in the subsequent section on file system recommendations, RAID configurations provided by Oracle Sun Storage Fibre Channel Arrays are fully supported and recommended for use in conjunction with Oracle ASM and provides the best storage solution, in terms of both performance and reliability.

Flash Storage

Oracle provides a complete Flash-based storage portfolio, from high-performance arrays through Flash-optimized system software and databases, all with Oracle support.

Using Oracle Flash-based storage can reduce I/O service times up to 15x, and accelerate applications up to 2x using Oracle's Flash storage and Flash-optimized database and systems software.

Oracle provides two Flash storage options: The Sun Storage 5100 Flash Array and the Sun Flash Accelerator F20 PCIe Card.

Oracle's Sun Storage F5100 Flash Array is a fast and power efficient flash array for accelerating databases and I/O intensive applications. It redefines database performance, cutting transaction times in half and potentially doubling application throughput. The Sun Storage F5100 Flash Array scales performance and capacity needs efficiently without impacting data availability to support growing business needs.

The Sun Flash Accelerator F20 PCIe Card improves response times, and reduces I/O latency. Based on Sun FlashFire technology, it delivers the I/O performance of over 300 disk drives to eliminate storage I/O bottlenecks and help your servers and applications run faster and more efficiently. In particular, the Sun Flash Accelerator F20 PCIs card can be used in Database Smart Flash Cache configurations, which are a new feature in Oracle Database 11gR2.

Unified Storage

Oracle Sun Unified Storage 7000 Series Systems combine an innovative storage architecture and file system with storage analytics and management technologies to deliver leadership performance and value.

Oracle Sun Unified Storage Systems simultaneously provide multiple storage interconnects (GigE, 10GigE, Fibre Channel and InfiniBand IB), thereby consolidating storage for file and block I/O driven applications onto a single high-capacity, high-performance storage system with reduced administration and the industry's most comprehensive and intuitive analytics environment.

Oracle Sun Unified Storage Systems run the ZFS file system. As mentioned in the following section on file system recommendations, Oracle ZFS is recommended for Oracle Database binaries and fully certified for Oracle Database data files.

Storage analytic and storage management tools are included with Oracle Sun Unified Storage Systems, which enables an integrated Oracle Database to Oracle Sun Storage 7000 view and provides monitoring of multiple storage systems from a single pane.

File System Recommendations

While the Oracle binaries can be installed on any supported file system, the recommended best practice for Oracle Database deployments on Oracle Solaris is to use Oracle Solaris ZFS for the Oracle Database binaries and Oracle ASM as an integrated volume manager for the Oracle Database data files.

However, if you choose to use Oracle Solaris ZFS for Oracle database deployment, which is certified for Oracle Database 10g and 11g , refer to the resources section for a list of collateral on best practices for using Oracle Solaris ZFS with Oracle Database.

Oracle Binaries – Oracle Solaris ZFS

Oracle Solaris Zettabyte File System (ZFS) technology offers a dramatic advancement in data management with a virtual storage pool design, integrated volume manager, and data services that provide an innovative approach to data integrity.

Oracle Solaris ZFS software enables more efficient and optimized use of storage devices, while dramatically increasing reliability and scalability. Physical storage can be dynamically added or removed from storage pools without interrupting services, providing new levels of flexibility, availability, and performance.

Oracle Solaris ZFS protects all data by 256-bit check sums, resulting in 99.99999999999999999-percent error detection and correction. Oracle Solaris ZFS constantly reads and checks data to help ensure it is correct, and if it detects an error in a storage pool with redundancy (protected with mirroring, Oracle Solaris ZFS RAIDZ, or Oracle Solaris ZFS RAIDZ2), Oracle Solaris ZFS automatically repairs the corrupt data. This contributes to continuous availability by helping to protect against costly and time-consuming data loss due to hardware or software failure, and by reducing the chance of administrator error when performing file system-related tasks.

Oracle Solaris ZFS software optimizes file system reliability by maintaining data redundancy on commodity hardware. It seamlessly and transparently supports new hybrid disk storage pools that include Flash technology for superior application performance.

Oracle Solaris ZFS can also be used to create snapshots and clones. The major difference is that snapshots are read-only, while clones are read-write copies of the file system.

In addition to typical disaster recovery and other business continuity purposes, Oracle Solaris ZFS snapshots are particularly useful as the Oracle Database environment is being built. In case of any configuration errors, the entire filesystem may be reverted back to its previous state.

Deployments to multiple target systems can be accelerated using Oracle Solaris ZFS cloning functionality. Instead of installing multiple systems from scratch, existing instances may be cloned in a fraction of the time.

Oracle Data Files – ASM

As mentioned above, the recommended practice for the Oracle Database data files is to use Oracle ASM as an integrated volume manager.

Once disk sets are configured, Oracle ASM will begin to stripe data across all available disk sets. The striping method is configurable to three levels:

- Normal – Oracle ASM will maintain two copies of all data
- High Availability – Oracle ASM will maintain three copies of all data.
- External – Oracle ASM will maintain only one copy of all data. In this scenario, it is expected for the external storage device to mirror data itself.

Please note that data redundancy can be defined at a much more granular level (down to the file level). Advanced configurations such as this, are beyond the scope of this paper. For further details, please see the *Storage Administrators' Guide*.

For storage systems that do not provide hardware RAID redundancy, it is recommended to present all storage devices to Oracle ASM individually, without the use of third-party volume managers Since Oracle ASM provides volume management in addition to a file system, including additional volume managers would be an unnecessary duplication of functionality.

For storage systems that provide hardware RAID, it is recommended to mirror the disks using hardware and set the redundancy level of the volume in ASM to "External".

- For maximum performance, the array should use RAID 10 and expose the disk set as one LUN.
- For more efficient disk usage (at a performance cost), RAID 5 is recommended. Again, each disk set would be exposed as one LUN.

- Also compliant with recommended practices, is the use of mirrored disk pairs. Using this method, each disk pair is mirrored and presented to Oracle ASM under its own LUN.

After creating disk groups, Oracle ASM will stipe data across all available disks, creating either one, two, or three copies of data, depending on availability settings.

Finally, it is recommended to separate data files and redo log files into two separate disk groups. Since redo log information has different properties (sequential reads and writes), significant performance gains may be possible.

Since Oracle Database 10g, Automatic Storage Management (ASM) provides the database administrator with a simple storage management interface, which is consistent across all server and storage platforms. The Automatic Storage Management feature in Oracle Database 11g Release 2 extends ASM functionality to manage ALL data: Oracle database files, Oracle Clusterware files and non-structured general-purpose data such as binaries, externals files and text files.

Improving Performance with Database Smart Flash Cache

For a database instance, the amount of data that is being actively accessed at any given time is referred to as the working set. For a large database instance, the working set may be small at times, for example, nights or weekends when the load is light and an individual user executes a query against an index and a small amount of data is returned. At other times, the working set of the same database instance may be large, for instance when queries result in full table scans. Optimal performance is obtained when the working set is resident in the buffer cache of the SGA and database performance will degrade as the working set outgrows the buffer cache such that data must be fetched from disk.

When the working set exceeds the buffer cache, Oracle Enterprise Manager tuning alternatives should be explored, for examples, the SQL Access Advisor may recommend indexes which will reduce the size of the working set. When tuning is not able to align the working set with size of the buffer cache, adding RAM and increasing the SGA_MAX_SIZE can result in substantial performance benefits. RAM, however, is relatively expensive. An affordable alternative to increasing the amount of RAM used by the SGA buffer cache is to use Database Smart Flash Cache to supplement the buffer cache.

Database Smart Flash Cache is a new SGA feature which expands DB buffer cache beyond main memory to flash-based storage. The graph in Figure 3 shows the throughput and response time benefits from adding Oracle Database Smart Flash Cache for a read-only workload. Transaction response times improve and throughout grows as the Database Smart Flash Cache size is increased.

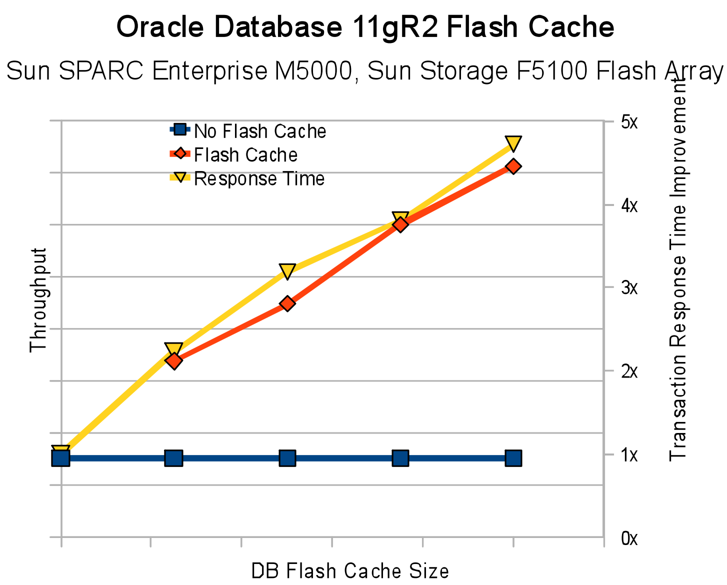

Figure 3: Transaction Response Time Improvements with Flash Cache

Consider adding Database Smart Flash Cache when the following are true:[2]

- Your database is running on the Solaris or Oracle Enterprise Linux operating systems. The flash cache is supported on these operating systems only.
- The Buffer Pool Advisory section of your Automatic Workload Repository (AWR) report or STATSPACK report indicates that doubling the size of the buffer cache would be beneficial.
- db file sequential read is a top wait event.

Oracle Database 11gR2 offers a way to put a flash device to optimal use, by configuring the device as an extra level of database block buffer cache. The configuration is very straightforward: it is sufficient to specify two initialization parameters, one with the name of the flash disk and another one with its size.

***.db_flash_cache_file='/dev/rdsk/c7t55d0s0'**

***.db_flash_cache_size=21474836480**

[2] "Configuring Database Smart Flash Cache," http://st-doc.us.oracle.com/11/112/server.112/e10595/memory005.htm.

The advantage of using flash technology in this way is that it has no impact on the existing setup and operations of a database. Sun Storage F5100 Flash Arrays and Sun Flash Accelerator F20 PCIe Cards are ideal devices for Flash Cache.

See *Oracle Database Administrator's Guide 11g Release 2 (11.2): Configuring Database Smart Flash Cache* for more information about Flash Cache tuning and sizing.

Virtualization Options for Oracle Sun M-series Systems

Figure 4 illustrates the spectrum of virtualization options available for Oracle Database deployments on Sun SPARC Enterprise M-series Systems.

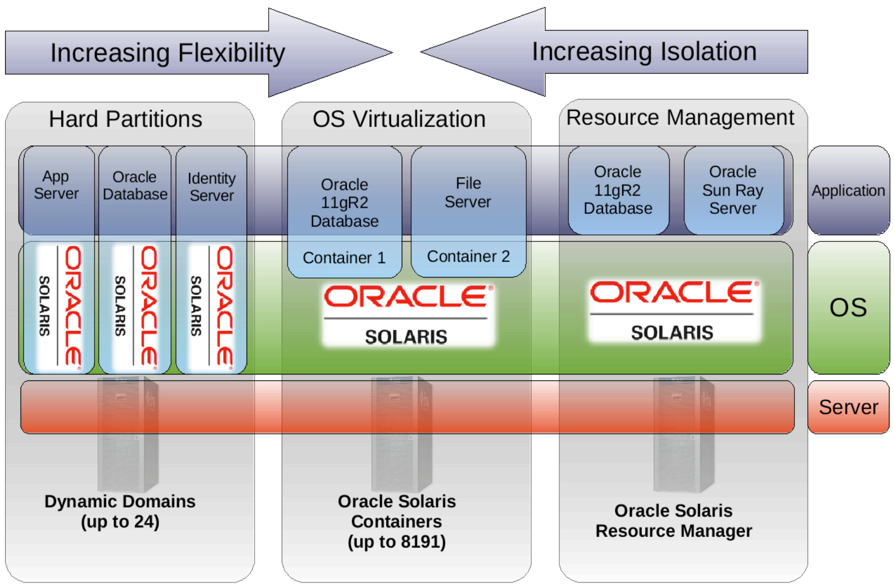

Figure 4: Virtualization Options for M-series Systems

Dynamic Domains offer superb hardware isolation by dividing a single system into multiple electrically isolated partitions, each running its own instance of the Oracle Solaris operating system. Hardware or software failures in one domain do not affect applications running in other domains.

On the other end of the spectrum, Oracle Solaris Resource Manager offers superb flexibility for sharing resources running within a single instance of the Oracle Solaris operating system by implementing administrative policies that govern the level of consumption of resources that each user or application is permitted. In some cases, enterprises can benefit by utilizing more than one virtualization technology on an individual server.

For more information on virtualization technologies, see *Virtualization Options for Oracle Database Deployments on Sun SPARC Enterprise M-series Systems.*

Deploying Oracle Database on the Oracle Solaris Platform

The Oracle Solaris Operating System runs across the entire range of Sun SPARC and x86 platforms from entry level servers to 64-processor servers like the Sun SPARC Enterprise M9000 server and 256 thread servers like the Sun SPARC Enterprise T5440.

The Oracle Solaris 10 Operating System introduces new features to enhance manageability, performance and availability. The key new features include Oracle Solaris Containers, Predictive Self-healing, DTrace for advanced observability, ZFS for next-generation volume management and file system support, and user and process rights management. An Oracle database deployment can take advantage of each of these features of Oracle Solaris 10 to enhance the manageability, scalability, availability and security of both single and multiple Oracle database instances – all across multiple platform and processor architectures.

Proven Performance and Scalability

The Oracle database has a proven track record of scaling well both vertically and horizontally on the Solaris 10 platform. Oracle Database 11gR2 with Oracle Real Application Clusters demonstrated excellent horizontal scalability across 12 Sun SPARC Enterprise T5440 servers on Solaris 10 running the industry-standard TPC-C workload. The Oracle database also scaled well on a single Oracle Sun SPARC Enterprise M9000 running the industry-standard TPC-H data warehousing benchmark. Oracle believes in empowering its customers to use scalability in both horizontal and vertical dimensions to best meet their performance and availability criteria. See the *References* Section below for the results.

Predictive Self-healing: Enhanced Availability

The Oracle Solaris Operating System has implemented predictive self-healing for CPU, memory, and I/O bus nexus components for a variety of hardware platforms incorporating SPARC, AMD Opteron and Intel Xeon processors, exploiting the specific hardware RAS features provided by the underlying system. Additionally, the Oracle Solaris Operating System provides a platform neutral technology, Memory Page Retirement (MPR), to ensure that both the Oracle Solaris Operating System and user applications continue to operate in the face of main memory faults.

Oracle Solaris MPR technology ensures that Oracle database deployments can continue uninterrupted even when the underlying system has memory errors. Consider the scenario of an Oracle database instance deployed on a system that is experiencing memory errors. The diagnosis engine of the Oracle Solaris fault manager, which is continuously examining both correctable errors (CEs) and uncorrectable memory errors (UEs), will see a series of correctable errors in a memory location as an indication of a potentially uncorrectable memory error. If the Oracle database has memory pages that contain CEs then Oracle Solaris MPR will retire those pages from memory without interrupting Oracle processes. If the Oracle database references memory pages that have uncorrectable memory errors, then Oracle Solaris MPR will retire clean pages containing UEs, again without interrupting Oracle processes. In the unlikely case of the Oracle database instance having dirty memory pages with UEs, the oracle processes will come down. However, if Oracle Database is configured with the Service Management Facility, as explained in the next section, it can restart automatically.

Adding the Oracle database instance and Oracle listeners as a service to the Oracle Solaris Service Management Facility (SMF) provides the following advantages:

1. Automatically restarts failed services in dependency order, whether they failed as the result of administrator error, a software bug, or were affected by an uncorrectable hardware error.
2. Provides more information about misconfigured or misbehaving services, including an explanation of why a service isn't running, as well as individual, persistent log files for each service.
3. Delegates the task of managing the Oracle services to Oracle administrators -SMF is integrated with Oracle Solaris Role Based Access Control (RBAC) which ensures that the services can be securely managed by non-root users, including the ability to configure, start, stop, or restart services.

User Rights Management: Enhanced Security

Default installations of the Oracle database can be made more secure by exploiting the user rights management feature of Oracle Solaris 10 security. In a typical Oracle deployment, all Oracle database administrators (DBAs) login as UNIX user oracle. Hence, it is not possible to track the DBA-related activities of an individual user; only the combined activities of all DBAs are tracked by the operating system and the database server. User rights management enables you to create an oracle role and assign it to users with DBA responsibilities. In this scenario, the users will login to the database server system with their regular UNIX logins and assume the oracle role when they need to do any Oracle DBA-related tasks. This approach ensures that multiple Oracle administrators do not share a single login. They login in as individual users and are accountable for their individual actions; yet they have the flexibility to perform all the functions of an Oracle administrator by assuming the oracle role. Complete accountability for individual users can be enforced by enabling auditing of the oracle role; which in turn will provide a detailed description all Oracle DBA-related activities for each individual UNIX user.

DTrace: Enhanced Observability

With the advent of multi-tier architectures today's applications have become very complex. While individual levels of the application tier may have excellent tools for observability and debugging, there are no tools to observe and optimize the entire application stack. This problem becomes even more complicated for observing applications in production which are likely sensitive to performance impacts. Also, it is not always easy to stop and start these applications to enable debug flags. Adding debug versions of applications into production may not be permitted. Even if permitted, bringing debug versions into production involves expensive and time consuming QA cycles. All of these issues complicate the problem of observation.

DTrace, a Dynamic Tracing framework, was developed to address this very problem. It can be used to observe any or all tiers of the application stack, it is truly dynamic and does not require application code changes or even an application restart. One can observe fully optimized applications using DTrace. The overhead of observation is low and there is no overhead when observation is turned off.

Instrumentation can be turned on and off dynamically thus only collecting information when it is needed. DTrace is safe and turns itself off when observation overhead affects system performance.

DTrace can be used to observe applications developed in, C, C++, Java, JavaScript, Ruby, PHP, Perl, Python among other programming and scripting languages. Other system layers, like I/O, networking, application and kernel locks, CPU counters etc, can also be observed using DTrace.

DTrace scripts are used to enable and program points of instrumentation. D-script format does not change based on the application tier being observed and a single script can be used to observe multiple tiers at the same time.

DTrace can be used to look at Oracle database processes in isolation or concurrently with any other processes running on the system and can be an invaluable tool for identifying performance bottlenecks and many other real world issues.

Enhanced System V IPC implementation

Prior to Oracle Solaris 10, installing Oracle Database on the Oracle Solaris Operating System required changes to the /etc/system file. Every reconfiguration required a reboot for the changes to take effect. However, the System V IPC implementation in Solaris 10 no longer needs changes to the /etc/system file. Instead, the new resource control facility is used, which allows changes to become effective immediately, without a system reboot. Furthermore, the default settings of the System V IPC parameters have been set to typical defaults enabling Oracle database instances to run out-of-the-box without requiring special parameters to be set.

Oracle database deployments on Oracle Solaris 10 work out of the box, with no additional system configuration, if the System Global Area (SGA) uses less than 25% of the system's total memory. If the deployment plans to use more than 25% of the systems memory, then the shared memory resource parameter can be dynamically set to the required value using the resource control facility.

Deploying Oracle Database on M-series Servers

As discussed in previous sections of this paper, Oracle provides a complete range of virtualization and resource management technologies. By using either Oracle Solaris container technology for software-based virtualization, Oracle Sun Enterprise M-series Dynamic Domains for hardware-based system virtualization and isolation, or a combination of both, it is possible to consolidate multiple Oracle Database single instances on a single physical Oracle Sun SPARC Enterprise Server while providing isolated, secure environments for each instance.

Dynamic Reconfiguration Using XSCF

By using Dynamic Reconfiguration technology, resources may be shifted between Dynamic Domains without incurring any system downtime. As mentioned, this technology helps administrators perform live upgrades and physical changes to system hardware resources for maintenance purposes. When Oracle database instances are hosted in Dynamic Domains, Dynamic Reconfiguration provides administrators with powerful resource allocation tools to enable more efficient server utilization.

The M-series servers provide system management capabilities through eXtended System Controller Facility (XSCF) firmware, which is pre-installed at the factory on the Service Processor boards.

The XSCF firmware consists of system management applications and two user interfaces to configure and control them:

- XSCF Web, a browser-based graphical user interface
- XSCF Shell, a terminal-based command-line interface

The XSCF firmware can be accessed by logging in to the XSCF command shell. This document includes instructions for using the XSCF interface as part of the initial system configuration. For more information about the XSCF firmware, see Chapter 2 of the *Sun SPARC Enterprise M3000/M4000/M5000/M8000/M9000 Servers Administration Guide*, and the *SPARC Enterprise M3000/M4000/M5000/M8000/M9000 Servers XSCF User's Guide*.

Dynamic Domains

As mentioned previously, a domain is an independent system resource that runs its own copy of the Oracle Solaris OS. Domains electronically divide a system's total resources into separate units that are not affected by each other's operations. Domains can be used for different types of processing; for example, one domain can be used to test new applications, while another domain can be used for production purposes. For Oracle databases, independent database engines or RAC nodes can be run on separate domains on the same physical M-series server, while maintaining complete hardware isolation.

The entry-level M3000 server supports only a single domain, one CPU, 8 dual inline memory modules (DIMMs), and I/O. Entry-level servers have a fixed system board configuration by default. Therefore, you do not need to reconfigure the system board on these systems.

Midrange and high-end servers support multiple domains and one to 16 physical system boards (PSBs). One PSB consists of 4 CPUs, 32 DIMMs, and I/O. The I/O varies by server, and can include PCIe slots, PCI-X slots, and built-in I/O.

To use a PSB in your midrange or high-end server, the hardware resources on the board must be logically divided and reconfigured as eXtended System Boards (XSBs). There are two modes of XSBs:

Uni-XSB

> In Uni-XSB mode, a PSB logically undivided and configured into one XSB. The single XSB contains all the resources on the board: 4 CPUs, 32 DIMMs, and I/O on a midrange and high-end server (1 CPU, 8 DIMMs, and I/O on an entry-level server).

Quad-XSB (midrange and high-end M-series servers only)

In Quad-XSB mode, a PSB is logically divided and configured into four XSBs. Each of the four XSBs contains one-quarter of the total board resources: 1 CPU, 8 DIMMs, and I/O[3].

A domain consists of one or more XSBs (single board in uni-XSB configuration or quarter of a board in quad-XSB configuration). A domain must have, at a minimum, 1 CPU, 8 DIMMs, and I/O.

In contrast to Solaris containers, mentioned previously in this paper, each domain runs its own copy of the Oracle Solaris OS.

For complete information on XSCF shell procedures, see Chapter 4 of the *Sun SPARC Enterprise M3000/M4000/M5000/M8000/M9000 Servers Administration Guide*.

Configuring Memory for Domains using XSCF

The Dynamic Reconfiguration (DR) functions classify system boards by memory usage into two types: Kernel memory board and User memory board.

A **kernel memory board** is a system board on which kernel memory (memory internally used by the Oracle Solaris OS and containing an OpenBoot PROM program) is loaded. Kernel memory cannot be removed from the system. But the location of kernel memory can be controlled, and kernel memory can be copied from one board to another.

To copy kernel memory from one board to another, use the Copy-rename operation. Copy-rename makes it possible for you to perform DR operations on kernel memory boards. For further information on this topic, see Chapter 2 of the *Sun SPARC Enterprise M4000/M5000/M8000/M9000 Servers Dynamic Reconfiguration (DR) User's Guide*.

A **user memory board** is a system board on which no kernel memory is loaded.

Before deleting user memory, the system attempts to swap out the physical pages to the swap area. Sufficient swap space must be available for this operation to succeed.

Some user pages are locked into memory and cannot be swapped out. These pages receive special treatment by DR.

Intimate Shared Memory (ISM) pages are special user pages which are shared by all processes. ISM pages are permanently locked and cannot be swapped out as memory pages. ISM is usually used by Data Base Management System (DBMS) software to achieve better performance.

Although locked pages cannot be swapped out, the system automatically moves them to the memory on another system board to avoid any problem concerning the pages. Note, however, that the deletion of user memory fails if there is not sufficient free memory size on the remaining system boards to hold the relocated pages.

---

[3] On a midrange server, only two XSBs have I/O per physical board

Although such moving of memory (called save processing) requires a certain length of time, system operations can continue during save processing because it is executed as a background task.

Deleting or moving a user memory board fails if either of the following statements is true:

- The swap area does not have sufficient free space to save data from the user memory to be deleted.
- There are too many locked or ISM pages to be covered by the memory on other system boards.

For further information on configuring memory for domains using XSF, see Chapter 2 of the *Sun SPARC Enterprise M4000/M5000/M8000/M9000 Servers Dynamic Reconfiguration (DR) User's Guide*.

Memory Configuration and DISM

Recent releases of Oracle Database can automatically assign more memory, up to a prescribed limit, to both the buffer cache and the shared pool if additional memory will improve performance. The Oracle Solaris Operating System supports such operations with the Dynamic Intimate Shared Memory (DISM).

Since shared memory is very heavily used in Oracle Database environments, it is important to optimize access to it and to minimize the amount of CPU consumed while referring to it. With this in mind, a specially-tuned variant of System V Shared Memory was introduced in Oracle Solaris many years ago, called *Intimate Shared Memory* (ISM). ISM has been widely used for database shared memory since. ISM is supported on both the the global system and in Oracle Solaris Containers (i.e. "zones"), via the project.max-shm-memory and zone.max-shm-memory parameters, respectively.

Oracle Solaris Dynamic Intimate Shared Memory (DISM) provides shared memory with the same essential characteristics as ISM except that it is dynamically resizable. That means that DISM, if configured correctly, offers the performance benefits of ISM while allowing the shared memory segment to be dynamically resized, both for the sake of performance and to allow dynamic reconfiguration (for example, adding or removing memory from a system or a domain).

DISM allows applications to respond to changes in memory availability by dynamically increasing or decreasing the size of optimized shared memory segments. Oracle Database has supported DISM since the Oracle Database 9i release, and was the first commercial application to do so.

The ability to dynamically change the amount of memory allocated to an Oracle Database instance has a number of benefits, including the ability to

- Allow Oracle Database to dynamically tune the database instance.
- Support Dynamic Reconfiguration operations, for example, adding or removing memory on Oracle's Sun SPARC Enterprise M-series servers using the eXtended System Control Facility (XSCF).
- Enable memory to be dynamically moved between different database instances according to changing processing requirements.

It is important to note that significant performance degradation can occur if DISM is not configured correctly, and for that reason Oracle recommends that DISM be turned off by default on SPARC-based servers. For users who need the functionality provided by DISM, it is possible to configure and monitor DISM to ensure its correct behavior.

DISM is turned on by default for Oracle Database 11.2.0.1 on Oracle Solaris, which makes it important for Database Administrators (DBAs) to understand its capabilities and behavior.

Regardless of whether Oracle Database uses ISM or DISM, it can always exchange the memory between dynamically sizable components such as the buffer cache, the shared pool, and the large pool after it starts an instance. Oracle Database can relinquish memory from one dynamic SGA component and allocate it to another component.

The MEMORY_TARGET parameter was introduced in Oracle Database 11g to allow automatic memory tuning. The amount of memory used is determined by the SGA_TARGET parameter. By default 40% of physical memory will be used. Further, by default 60% of the memory addressed by MEMORY_TARGET is allocated to the System Global Area (SGA) and the remaining 40% to the Program Global Area (PGA). The SGA/PGA mix can be changed dynamically.

Consequently the portion of SGA_TARGET which is expected to be locked (the SGA component) by default is therefore 60% of 40% of physical memory. That represents 24% of physical memory, just below the 25% that can be allocated without changing **project.max-shm-memory parameter** (or **zone.max-shm-memory** if running in an Oracle Solaris Container).

If memory is dynamically diverted from the PGA to the SGA, the 25% threshold is likely to be exceeded. It is very important to set the **project.max-shm-memory** parameter (or **zone.max-shm-memory**) high enough to ensure that all SGA memory can be locked under all circumstances.

For further information on configuring DISM for Oracle Database, see *Dynamic SGA Tuning of Oracle Database on Oracle Solaris with DISM*.

Configuring CPUs for Domains Using XSCF

Using DR to change a CPU configuration is easier than using it to change the configuration of memory or an I/O device.

An added CPU is automatically recognized by the Oracle Solaris OS and becomes available for use.

For a CPU to be deleted, the following conditions must be met:

- No running process is bound to the CPU to be deleted. If a running process is bound to the target CPU, you must unbind or stop the process.
- The CPU to be deleted does not belong to any processor set. If the target processor belongs to a processor set, you must delete the CPU from the processor set by using the **psrset**(1M) command.
- If the resource pools facility is in use by the domain, the CPU cannot be deleted unless the minimum processor set sizes can otherwise be maintained. Use the Oracle Solaris commands **pooladm**(1M) and **poolcfg**(1M) to check these parameters and, if necessary, adjust the sizes of the domain's resource pools.

Once Oracle Solaris has recognized the additional CPU and it has become available for use. Oracle Database automatically utilizes the new resources.

Configuring I/O for Domains Using XSCF

I/O devices can be added to and removed from Dynamic Domains without requiring system restarts.

The device driver processing executed by the Oracle Solaris OS is based on the premise that all device drivers dynamically recognize newly added devices. In the domain where DR is performed, all device drivers must support the addition of devices by DR. Upon the addition of an I/O device by DR, the I/O device is reconfigured automatically.

The path name of a device file under /dev is configured as the path name of the newly added I/O device to make the I/O device accessible.

If a device driver that does not support DR is used in the domain, all access to I/O devices controlled by the device driver must be stopped, and the device driver must be unloaded by using the **modunload**(1M) command.

For further information on configuring memory for domains using XSF, see Chapter 2 of the *Sun SPARC Enterprise M4000/M5000/M8000/M9000 Servers Dynamic Reconfiguration (DR) User's Guide.*

Oracle RAC

Oracle Real Application Clusters (RAC) is the premier database clustering technology that allows two or more computers ("nodes") in a cluster to concurrently access a single shared database. This database system spans multiple hardware systems, yet appears to the application as a single unified database.

This architecture extends availability and scalability benefits to all applications, specifically:

- Fault tolerance within the cluster, especially computer failures.
- Flexibility and cost effectiveness in capacity planning, so that a system can scale to any

desired capacity on demand and as business needs change.

A key advantage of RAC is the inherent fault tolerance provided by multiple nodes. Since the physical nodes run independently, the failure of one or more nodes does not affect other nodes. This architecture also allows a group of nodes to be transparently put online or taken offline, while the rest of the cluster continues to provide database service. Additionally, RAC provides built-in integration with Oracle Fusion Middleware and Oracle clients for failing over connections.

Oracle RAC also gives users the flexibility to add nodes to the cluster as the demands for capacity increase, reducing costs by avoiding the more expensive and disruptive upgrade path of replacing an existing system with a new one having more capacity. The Cache Fusion technology implemented in Oracle RAC and the support for InfiniBand networking enable capacity to be scaled near linearly without any changes to your application.

With its unique capabilities described above, Oracle RAC enables enterprise Grids. In addition, Oracle Real Application Clusters is completely transparent to the application accessing the Oracle RAC database, thereby allowing existing applications to be deployed on Oracle RAC without requiring any modifications.

Oracle Clusterware

Oracle Database 11g includes Oracle Clusterware, a complete, integrated clusterware management solution available on all Oracle Database 11g platforms. This clusterware functionality includes mechanisms for cluster messaging, locking, failure detection, and recovery. Oracle Clusterware 11g adds cluster time management to ensure that the clocks on all nodes in the cluster are synchronized.

Oracle RAC Nodes on Virtualized Systems

As discussed in previous sections of this paper, Oracle provides a complete range of virtualization technologies. Using Oracle Solaris' container technology for software-based virtualization, Oracle Sun Enterprise M-series Dynamic Domains for hardware-based virtualization and isolation, or a combination of both, it is possible to host several RAC nodes on each physical Oracle Sun SPARC Enterprise M-series System.

From the software standpoint, deployment of an Oracle RAC node to Dynamic Domains is functionally equivalent to deploying to a physical system.

Oracle ASM

Oracle supports dynamic online system reconfiguration for all components of your Oracle hardware stack. Oracle's Automatic Storage Management (ASM) has built-in capabilities that allow the online addition or removal of ASM disks. When disks are added or removed from an ASM Diskgroup – Oracle automatically rebalances the data across the new storage configuration while the storage, database, and application remain online.

Tuning Recommendations for Oracle DB on M-series

As detailed in this document, the Oracle Sun Enterprise M-series Systems provide mainframe-class system architecture. These systems are are fast, highly reliable, easy to manage, and vertically-scalable. While single threaded performance on the M-series Systems is very good, with up to 64 4-core SPARC VII processors and up to 4TB of memory on the systems, extensive use of parallelism and concurrency can provide significant performance benefits.

A fully configured M9000-64 presents 512 CPUs to the Oracle Solaris Operating System; a combination of concurrency and parallelism should be used to effectively utilize these resources. Online Transaction Processing (OLTP) applications typically result in many user processes, offering one way of consuming available CPUs. Decision Support Systems (DSS) and Data Warehousing applications can take advantage of the builtin parallel SQL capabilities of Oracle Database. And virtualization, described earlier in this paper, can also be used to configure a Sun SPARC Enterprise M-series system as multiple virtual servers.

The following processes are ideal candidates for parallelization:

Parallel data loading

Parallel Data Loading option of SQL*Loader allows multiple processes to load data into the database at the same table. To optimize:

- Create multiple input files
- Create a control file for each input file

Example with 3 parallel loads:

```
sqlload scott/tiger CONTROL=control1.ctl DIRECT=TRUE PARALLEL=TRUE

sqlload scott/tiger CONTROL=control2.ctl DIRECT=TRUE PARALLEL=TRUE

sqlload scott/tiger CONTROL=control3.ctl DIRECT=TRUE PARALLEL=TRUE
```

Index Creation

When creating indexes, use the parallel option for general parallel optimization and the "unrecoverable" option to eliminate redo. For further optimization, "compute statistics" can be done at the same time (which is the default in Oracle Database 11g.

Example code:

```
create index CUSTOMER_INDEX1 on CUSTOMER (CUST_ID, REGION)
parallel(degree 32) unrecoverable compute statistics;
```

RMAN Backups

When performing RMAN backups, use multiple channels and parallelism to maximize I/O throughput.

- Create multiple backup channels for RMAN and tell RMAN how to connect to the channels
- Configure the degree of parallelism

```
RMAN> CONFIGURE DEVICE TYPE DISK BACKUP TYPE TO BACKUPSET PARALLELISM 20;
```

- Issue the backup command

```
RMAN> BACKUP TABLESPACE <SPACENAME>;
```

Batch Workloads

Both single- and multi-threaded workloads work well on M-series systems. Of course, on M-series systems, multi-threaded workloads can take advantage of all processor, memory, and I/O resources provided to it.

In Oracle Applications, parallelism can be improved by increasing the number of Concurrent Managers and in adadmin, specifying the number of workers and specifying parallelism for maintenance functions.

Queries and DSS Workloads

Decision Support System (DSS) workloads can be optimized by providing "parallel" hints in the SQL queries.

Example #1: Table Scan

Pre-defining a parallel hint for a table:

```
alter table CUSTOMER parallel (degree 32);

select NAME, ADDRESS, BALANCE_OWING from CUSTOMER;
```

Using a hint to override any pre-defined table parallelism:

```
select /*+ PARALLEL(customer, 32) */

NAME, ADDRESS, BALANCE_OWING from CUSTOMER;
```

Example #2: Parallel DML

Create as select:

```
alter session enable parallel dml;

create table customer parallel (degree 32)
as select /*+ parallel (temp_table, 32) */ from temp_table;

alter session disable parallel dml;
```

Insert as select

```
alter session enable parallel dml;

insert /* parallel(customer, 32) */ into customer
select * from temp_table;

alter session disable parallel dml;
```

Conclusion

Oracle's Sun SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 servers are suitable for a large range of Oracle Database applications, especially when mainframe-class reliability, ease of management, and vertical-scalability are critical. These systems provide many of the benefits of traditional mainframes — without the associated cost or complexity. In fact, these servers deliver a mainframe-class system architecture at open systems prices. With symmetric multiprocessing (SMP) scalability from one to 64 four-core processors, memory subsystems as large as 4 TB, and high-throughput I/O architectures, Sun SPARC Enterprise M-series servers easily perform the heavy lifting required by consolidated workloads and data warehouse consolidation.

The M-series servers run the powerful Oracle Solaris operating system while including leading virtualization and resource management technologies. By offering Dynamic Domains, eXtended System Boards, Dynamic Reconfiguration, and Oracle Solaris Containers technology, Sun SPARC Enterprise M-series servers bring sophisticated flexibility and mainframe-class resource control to users of Oracle Database on an open systems compute platform.

References

RESOURCES

Database Installation Guide for Solaris Operating System	http://www.oracle.com/pls/db112/to_pdf?pathname=install.112/e10848.pdf
Deploying Oracle Databases on the Oracle Solaris Platform	http://www.oracle.com/technetwork/articles/systems-hardware-architecture/deploying-oracle-database-solaris-168405.pdf
Oracle® Grid Infrastructure Installation Guide 11g Release 2 (11.2) for Solaris Operating System	http://www.oracle.com/pls/db112/to_pdf?pathname=install.112/e10816.pdf
Oracle SQL Parallel Execution	http://www.oracle.com/technetwork/middleware/bi-foundation/twp-bidw-parallel-execution-11gr1-130766.pdf
Oracle's Sun SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 Server Architecture – Flexible, Mainframe-Class Compute Power	http://wikis.sun.com/display/SystemsComm/Oracles+Sun+SPARC+Enterprise+M3000,+M4000,+M5000,+M8000,+and+M9000+Server+Architecture+Flexible,+Mainframe-Class+Compute+Power
Parallel Execution Fundamentals in Oracle Database 11gR2	http://www.oracle.com/technetwork/middleware/bi-foundation/twp-parallel-execution-fundamentals-133639.pdf
Performance Considerations for Developers Utilizing Sun SPARC SPARC Enterprise® M-series Servers	http://wikis.sun.com/download/attachments/57517586/820-6882.pdf
Release Notes for Solaris Operating System	http://www.oracle.com/pls/db112/to_pdf?pathname=relnotes.112/e10847.pdf
Virtualization Options for Oracle Database deployments on T-Series systems	http://www.oracle.com /technetwork/articles/systems-hardware-architecture/virtualization-options-t-series-168434.pdf

ORACLE

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

Deploying Oracle Database on Oracle Sun SPARC Enterprise M-series Systems
September 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

SOFTWARE. HARDWARE. COMPLETE.