An Oracle Technical White Paper
December 2013

How to Consolidate, Optimize, and Deploy Oracle Database Environments Including Multitenant Databases with Flexible Virtualized Services

Oracle Optimized Solution for Oracle Database

Introduction

Bringing together Oracle products and technologies at each tier of the solution stack, Oracle Optimized Solutions are designed to save money, reduce integration risk, improve user productivity, and shorten time-to-deployment. Each Oracle Optimized Solution acts as a blueprint to implement an effective infrastructure solution, applying validated configurations and proven best practices to address critical business challenges in the enterprise.

The Oracle Optimized Solution for Oracle Database creates an ideal infrastructure for consolidating Oracle Database services. It defines a flexible architecture for delivering fast, scalable, and continuous services for mission-critical database applications. The solution takes advantage of investments that Oracle has made in technology innovations and solution engineering to provide a robust infrastructure for consolidating database systems. This paper describes the underlying architecture, optimizations, and components, highlighting features in the solution stack that add value.

The Oracle Optimized Solution for Oracle Database helps to consolidate multiple Oracle Database releases securely within a single system, using less hardware, fewer licenses, and a smaller footprint. It is designed specifically to simplify datacenter cost and complexity and lower the total cost of ownership for database delivery. A companion paper, "How to Save Money, Reduce Risk, and Increase Agility When Deploying Oracle Database Systems," describes potential TCO savings and summarizes the business benefits.

Oracle Infrastructure for Oracle Database

In addition to the Oracle Optimized Solution for Oracle Database, Oracle has a full complement of database infrastructure offerings, including these Oracle engineered systems:

- **Oracle Database Appliance.** Oracle Database Appliance is an entry-level engineered system that integrates Oracle software, servers, storage, and networking to deliver highly available database services for a range of online transaction processing (OLTP) and data warehousing applications. It is a 4U rack-mountable appliance that combines Oracle servers and storage to create an off-the-shelf solution for small database requirements.
- **Oracle Exadata Database Machine.** Oracle Exadata Database Machine is an Oracle engineered system that factory-integrates Oracle hardware and software and supplies integral optimizations for Oracle Database workloads. Oracle Exadata is designed for one task and one task alone: to accelerate Oracle Database 11*g* and 12*c* services for data warehousing and OLTP applications. It is a fully preassembled and purpose-built system that serves this single targeted function.
- **Oracle SuperCluster.** Oracle SuperCluster engineered systems are Oracle's fastest and most scalable engineered systems, and are ideal for consolidating databases and applications, private cloud deployments, and Oracle software. They deliver extreme performance, the highest consolidation, availability, and efficiency, and eliminate the risks of build-it-yourself architectures. With the world's fastest processors and Oracle Exadata Storage Servers optimized for Oracle Database, Oracle

SuperCluster delivers unprecedented price/performance for mission-critical databases and applications.

Oracle designs these engineered systems to reduce the cost and complexity of IT infrastructures while increasing performance and customer value. These systems demonstrate how Oracle has made significant investments in engineering and optimizations at every layer of the stack, innovating to improve performance, simplify data center operations, reduce risk, and drive down costs.

While these factory-integrated engineered systems have tremendous value for many customers, they may not be the optimal solution for all sites and business needs. Engineered systems are purchased and shipped as specific configurations, so they aren't the best fit for organizations that may not require or be in a position to benefit from the Exadata Storage Servers included with those products. With the Oracle Optimized Solution, organizations can continue to invest in SAN storage or other installed technologies. In such cases, the Oracle Optimized Solution for Oracle Database is a better solution for customers to implement.

The Oracle Optimized Solution for Oracle Database

Like Oracle engineered systems, the Oracle Optimized Solution for Oracle Database follows an "Oracle-on-Oracle" approach that takes advantage of Oracle technologies and engineering expertise and supplies a single touch point for service and support. Unlike engineered systems, however, the Oracle Optimized Solution for Oracle Database is not a factory-built configuration. Instead, it is an architectural blueprint that combines Oracle technologies in proven and thoroughly tested configurations to help simplify and accelerate database deployments.

The Oracle Optimized Solution for Oracle Database offers tremendous flexibility in its ability to, adapt to existing environments, and adjust configurations as needed to meet business requirements. Customers that already have a significant investment in enterprise SAN storage, for example, can configure the solution to complement and leverage those resources. The Oracle Optimized Solution for Oracle Database can support Oracle Database 11*g*, and Oracle Database 12c making this solution ideal for sites that wish to modernize on current releases, and consolidate databases into a more cost-effective virtualized environment. Since deploying the Oracle Optimized Solution for Oracle Database with the latest Oracle software and hardware technologies provides the greatest benefit and highest level of optimization on infrastructure with the latest features that may not be available with previous generation components it is therefore recommended that customers deploy their Oracle Database Solutions with Oracle Database 11g and/or Oracle Database 12c as this paper discusses. This all said, should customers have requirements to support earlier versions of the Oracle Database in a virtualized environment Oracle can support customers with earlier architectures. Customers should reference the "Supported Virtualization and Partitioning Technologies for Oracle Database and RAC Product Releases" on Oracle's Technology Network Site for all of Oracle's supported configurations.

As shown in Figure 1, the solution leverages Oracle products and technologies at each tier of the Oracle technology stack—storage and compute systems, virtualization technologies, the Oracle Solaris operating system, and, of course, the Oracle Database software itself. Oracle Enterprise Manager manages the database environment through an intuitive administrative interface while Oracle

Enterprise Manager Ops Center supplies management of the Oracle server hardware, firmware, operating systems, virtual servers, and updates.

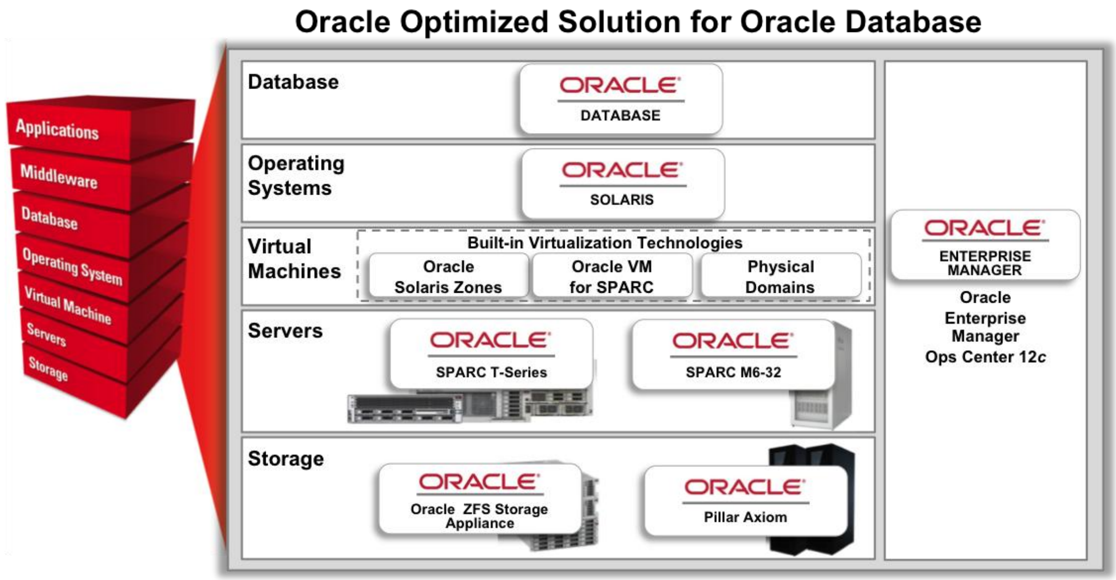

Figure 1. The Oracle Optimized Solution for Oracle Database leverages Oracle products and technologies to streamline mission-critical database deployments.

Solution Benefits

IT organizations often spend multiple weeks to plan, architect, troubleshoot, and deploy an infrastructure for Oracle Database and applications. Deployment teams must assemble and integrate a range of hardware and software components from different vendors (e.g., servers, storage, network, virtualization software, and operating systems). The process is not only time-consuming but also prone to errors, making it hard to meet schedules and achieve a profitable return on investment. Complex integration tasks can expose even the most carefully planned deployments to unanticipated risks that result in delays, downtime, costly upgrades, inefficient performance, or poor utilization.

The Oracle Optimized Solution for Oracle Database reduces the effort required to plan, architect, build, and test a database infrastructure, shortening time-to-deployment and reducing risks. It allows enterprises to leverage a flexible architecture and operating model that Oracle has developed and tested, rather than building an infrastructure from scratch. And because the components are all from Oracle, the solution includes the advantage of a single point of accountability, from purchasing to deployment, maintenance, and ongoing support.

Running Oracle Database on a complete Oracle stack brings the simplicity and reliability of having a single vendor to call. In multivendor solutions, production teams can spend hours debugging or tracing an issue before they can determine whether the source of the problem is the application, the virtual machine, the OS, or the physical server. When running an Oracle Optimized Solution, Oracle Support can be engaged right away, resolving problems more quickly and helping to avoid downtime.

A Smart and Flexible Virtualized Architecture

Figure 2 depicts a hypothetical deployment of the Oracle Optimized Solution for Oracle Database. To provide high availability for mission-critical workloads, Oracle RAC is used to support distributed database instances hosted on SPARC servers with access to shared storage. Defining Oracle RAC nodes across physical servers supplies redundancy that eliminates a single point of failure and helps to scale transaction performance. (On the SPARC M6-32 server, Oracle RAC nodes can reside on separate Physical Domains because of that server's availability design.) Please refer to Oracle's support matrix for Virtualization and Partitioning Technologies for Oracle Database and RAC for additional details.

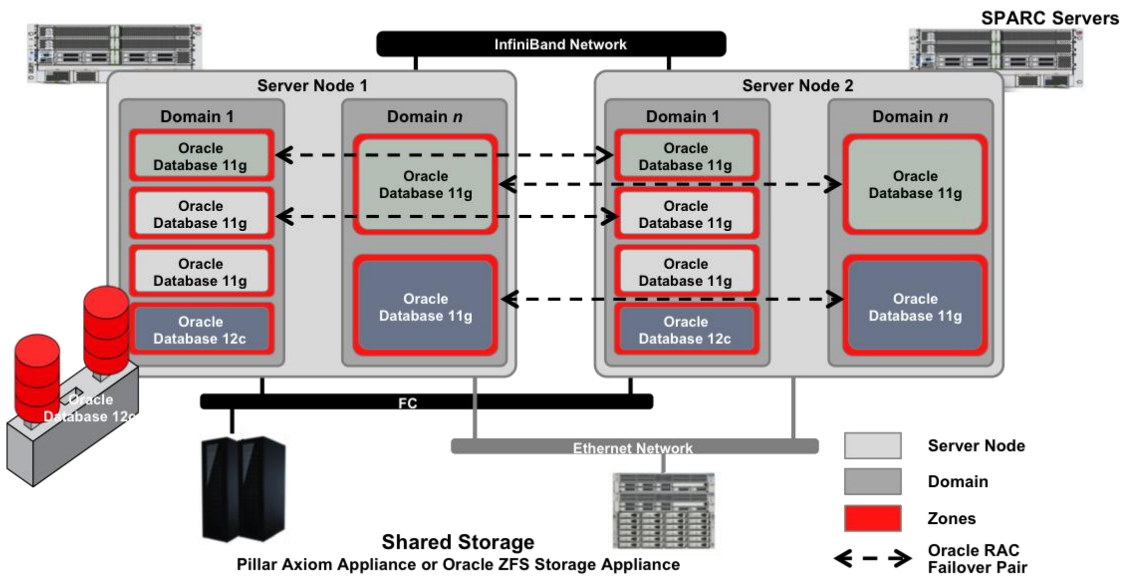

Figure 2. Hypothetical deployment of the Oracle Optimized Solution for Oracle Database.

The Oracle RAC interconnect is implemented on either Oracle InfiniBand or Ethernet. Shared storage resides on either Oracle's Pillar Axiom or Oracle ZFS Storage Appliances. Highly threaded SPARC T4, T5, or M6-32 servers connect via Fibre Channel to Oracle's Pillar Axiom storage system and via Ethernet to the Oracle ZFS Storage Appliance. In the process of validating Oracle RAC across multiple nodes, server platforms, and shared storage, Oracle Optimized Solution engineers documented configurations that optimize availability, performance, and ease of operations.

Figure 2 also shows how the solution architecture leverages virtualized environments for workload consolidation, isolation, and effective resource management. Because virtualization capabilities are especially important when combining many Oracle Databases on a single physical server, Oracle provides no-charge server virtualization technologies (Oracle VM for SPARC logical domains, Physical Domains in the SPARC M6-32 server, and Oracle Solaris Zones) These technologies are extremely lightweight (especially in comparison to other commercially available virtualization technologies), implying almost no overhead. Additionally, when one deploys Oracle Database 12c they receive the benefit of virtualization and consolidation of databases directly through the database. Oracle Database 12*c* Multitenant supports capability of unplugging and plugging of pluggable databases into consolidated container databases and the Oracle Optimized Solution for Oracle Database supports this level of operation along with cloning capabilities directly through native operations for Single Instance Database deployments.

Oracle's new SPARC T5 and M6-32 systems feature unprecedented thread scale and large memory footprints. The density of these servers enables efficient consolidation of existing systems, reducing footprint, conserving energy, and simplifying administration. Different system resources, such as virtual CPUs, memory, I/O and network devices, are assigned to virtual environments and reallocated as needed, even on live production systems in a majority of cases without any interruption, making sure that mission-critical applications get the resources they need. Administrators use the same management tool to monitor the health of physical servers—Oracle Enterprise Manager Ops Center—and to create virtual servers and adjust resources for each virtual machine (VM).

Oracle-Only Technologies and Optimizations

All components in the Oracle solution stack are engineered and tested to work together. In addition, Oracle has made considerable engineering investments to add value when multiple Oracle products are used together. These cross-product optimizations deliver advantages in areas like performance, security, system reliability, and ease of data center operations. While Oracle Database releases are available for a number of commercially available operating systems, virtualization technologies, and hardware platforms, there are distinct advantages of running Oracle Database services on an Oracle infrastructure. The Oracle Optimized Solution for Oracle Database allows sites to realize the benefits of Oracle's cross-product engineering efforts, including:

- **Fast cloning of production database systems.** DBAs frequently need to clone production databases to validate software changes, analyze query performance, or troubleshoot problems. Using Oracle Solaris Zones in conjunction with storage snapshots and clones (inherent in the Oracle ZFS Storage Appliance) and Oracle RMAN backups, Oracle engineers have developed procedures that automate the database cloning process, duplicating databases and the supporting Oracle Solaris Zones environments up to 50x faster than manual methods. (See page 30 for more detail.)
- **Secure Live Migration.** Migrating virtual environments provides a means of moving database workloads to other machines to support planned downtime or disaster recovery scenarios. Oracle VM Server for SPARC technology supports the migration of active domains while maintaining application services to users. In addition, in combination with on-chip cryptographic accelerators in SPARC T-Series and M6-32 servers, these migrations can be fully encrypted. Other virtualization solutions migrate VM data in the clear, which can leave sensitive data (such as passwords, account numbers, etc.) unprotected, compelling the use of dedicated networks or the addition of encryption devices. Instead, the Oracle Optimized Solution for Oracle Database enables secure, wire-speed encryption capabilities for live VM migrations, without any additional hardware or software investment required.

- **Built-in acceleration of Oracle Transparent Data Encryption.** The same on-core cryptography features in SPARC T-Series and M6-32 servers are applied to Oracle Databases workloads that use Oracle Transparent Data Encryption (TDE) to encrypt sensitive data. In recent Oracle TDE testing, the SPARC T4 proved to be 43% faster on secure database queries in comparison to Intel Xeon X5600 processors using AES/NI technology.[1] (For further information on TDE with Oracle Databases, please reference "High Performance Security For Oracle Database and Fusion Middleware Applications using SPARC T4.") Because on-chip encryption can occur at high speeds, organizations that must protect sensitive data can use encryption pervasively to reduce risk.
- **Kernel acceleration in Oracle Solaris for Oracle RAC.** The Oracle RAC distributed database uses the Lock Management System (LMS), a user-level distributed lock protocol that mediates requests for database blocks between processes on cluster nodes. Fulfilling a request requires traversing and copying data across the user/kernel boundary on the requesting and serving nodes, even for the significant number of requests for blocks with uncontended locks. Oracle Solaris includes a kernel accelerator to filter database block requests destined for LMS processes, directly granting requests for blocks with uncontended locks, eliminating user-kernel context switches, associated data copying, and LMS application-level processing for those requests.
- **I/O throughput at native speeds in VMs.** Oracle's SPARC T4 and T5 platforms deliver superior application workload performance with Single-Root I/O Virtualization (SR-IOV) and PCIe Direct I/O. SR-IOV enables the efficient sharing of PCIe network devices among I/O domains so application workloads can achieve near-native I/O performance, even in virtual environments. PCIe Direct I/O extends current PCIe support to enable the assignment of either individual PCIe cards or entire PCIe buses to an Oracle VM for SPARC domain. This helps to deliver native I/O throughput for applications needing maximum performance, and provides I/O configuration flexibility.
- **Memory optimizations that improve Oracle Database performance.** Oracle Solaris features memory management functionality and enhancements that are targeted at accelerating Oracle Database performance:
  - *NUMA optimization and memory placement.* Oracle Solaris 10 and Oracle Solaris 11 include memory placement intelligence that enhances performance for NUMA architectures like the SPARC T4, T5, and M6-32 servers (see page 22).
  - *Large memory pages.* Because of scalability enhancements to the virtual memory system in Oracle Solaris 11, Oracle Database can allocate 2GB page sizes to its System Global Area (SGA), a group of shared memory areas dedicated to Oracle Database instances. Large page sizes improve database performance by reducing costly Transaction Lookaside Buffer (TLB) misses in the CPU architecture that cause paging. In addition, enhancements to the memory predictor

[1] See https://blogs.oracle.com/roller-ui/bsc/spider.jsp?entry=403eff62-ff28-4007-b931-1abc606fb967 for test results.

in Oracle Solaris enable monitoring of memory page use and adjustments to the page size to match application and database needs, enhancing performance.

- *Support for ISM (Intimate Shared Memory), Dynamic ISM (DISM) and Optimized Shared Memory (OSM)*. Oracle Database uses shared memory to store frequently used data such as the database's buffer cache and shared pool. For this reason, the OS includes support for specially tuned variants of shared memory known as ISM and DISM. (Similar to ISM, DISM is a shared memory region that can grow dynamically; see page 22 for more information.) ISM and DISM shared memory segments take advantage of large pages locked in by the Oracle Solaris kernel that can't be paged out, making them always available and reducing context switches and overhead. With Oracle Solaris 11.1 and Oracle Database 12c customers receive the benefit of the latest software support for Optimized Shared Memory (OSM). This can be fully automated and OSM is primarily designed to allow Dynamic Sizing of SGA at no performance loss (if dynamic sizing is not done). It is optimized for database use cases and SGA allocated using OSM, will use the largest available and possible pagesize on the given system for SGA.
- *Multi-threaded shared memory operations*. A new multithreaded kernel process in Oracle Solaris 11 called "vmtasks" accelerates the creation, locking, and destruction of pages in shared memory. The vmtasks process parallelizes the creation and destruction of shared memory segments, dramatically speeding up database startups and shutdowns.

- **Hybrid Columnar Compression (HCC) technology.** When Oracle Database is used with either Pillar Axiom storage or the Oracle ZFS Storage Appliance, HCC enables substantial reductions in the overall storage footprint as well as performance enhancements since data is compressed. (See page 33 for more information.)
- **Quality of Service (QoS) and storage integration**. Oracle's Pillar Axiom storage system adjusts the QoS for SAN data storage according to defined priorities, making it easy to optimize I/O performance for business-critical databases. QoS policies can be modified at any time to reflect changes in business conditions and data value. The system's QoS manager interprets QoS policy settings and makes decisions on where to physically place data, based on the QoS service level requested as well as the actual devices in the system. Migrating database storage across different devices is a simple one-click process of changing the associated storage class. The system then migrates the data while continuously providing database access. (See page 25 for more information.)
- **Management efficiency**. Oracle Enterprise Manager Ops Center 12*c*, part of the Oracle Enterprise Manager portfolio, provides an intuitive interface for managing the solution hardware, firmware, virtual systems, operating system instances, administrative rights, patching, and updates. By adding other software in the Oracle Enterprise Manager family, administrators can also control Oracle Database instances and Oracle applications in the enterprise. An integrated and consistent set of management tools can save considerable labor costs in comparison to working with a variety of different multivendor tools.

A Choice of Customizable Configurations

The Oracle Optimized Solution for Oracle Database can be deployed using one of several recommended configurations, making it easy to match expected workload and growth projections.

Each configuration can be customized as needed to meet site objectives, including deployment requirements for applications, high availability, or disaster recovery. All configurations share the same underlying architecture and are based on enterprise-class Oracle components, creating a solution that is highly scalable, optimized for database performance, and engineered for high service levels for business-critical Oracle Databases.

TABLE 1. SAMPLE CONFIGURATIONS — ORACLE OPTIMIZED SOLUTION FOR ORACLE DATABASE

	SMALL	MEDIUM	LARGE	LARGE	X-LARGE
Server Nodes	2x SPARC T4-1, 1 CPU, 256GB each	2x SPARC T5-2, 2 CPUs, 512GB each	2x SPARC T5-4, 4 CPUs, 1TB each	2x SPARC T5-8, 8 CPUs, 2 TB each	1x SPARC M6-32, 32 CPUs, 8TB
Storage	Oracle ZFS Storage Appliance	Pillar Axiom 600 (1 FC SAN Slammer, 3 FC Bricks)	Pillar Axiom 600 (2 FC SAN Slammers, 6 FC Bricks)		Pillar Axiom 600 (4 FC SAN Slammers, 12 FC Bricks)
Software	Oracle Database 11*g*R2, Oracle RAC, and Oracle 12c (Single Instance), Oracle Solaris, Oracle VM Server for SPARC, Oracle Enterprise Manager 12*c*, Oracle Enterprise Management Ops Center 12*c*				
Advanced Services	• Fixed Scope Services: Install and Configure, Production Support Readiness, Patch Review and Deployment • Annual Services: Advanced Support Assistance, Business Critical Assistance, Solution Support Center, Advanced Monitoring and Resolution				

Key Components

The remainder of this paper examines architectural components at each tier of the solution stack that contribute to performance, scalability, reliability, and serviceability, including:

- Oracle's SPARC T4, T5, and M6-32 server platforms and virtualization technologies that help to consolidate database workloads. (An appendix covers the evolution of Oracle's multi-core, multithreaded processor designs and describes the Oracle SPARC T4, T5, and M6 processors that help to deliver performance and massive scale in these servers; see page 39.)
- The lightweight, no-charge virtualization technologies of Oracle VM for SPARC and Oracle Solaris Zones, as well as Physical Domains in SPARC M6-32 platforms.
- The Oracle Solaris operating system and how it contributes to the solution's scalability and reliability.
- NAS and SAN-based storage options for the solution, including Oracle ZFS Storage Appliances and Pillar Axiom storage systems.

The Oracle sales team will work with site architects to fine-tune a configuration to meet infrastructure requirements and business goals.

SPARC T4 and T5 Servers

The Oracle Optimized Solution for Oracle Database validates default configurations that incorporate the SPARC T4-1 (small), T5-2 (medium), and T5-4 or T5-8 (large) servers. (An extra large configuration is grounded on the massively scalable SPARC M6-32 server, described on page 13.) The solution architecture is flexible, however, and can be adapted using other SPARC T4 or T5 servers—consult your Oracle representative for recommendations on how to customize configurations to site and business requirements.

Tables 2 and 3 below give a quick comparison of SPARC T4 and T5 server models. Oracle's SPARC T4 servers are available in one-, two-, and four-socket implementations: the SPARC T4-1, T4-2, and T4-4 servers, respectively. SPARC T5 servers are available in blade, two-, four-, and eight-socket implementations: the SPARC T5-1B, T5-2, T5-4, and T5-8 servers. For additional details and resources, see http://www.oracle.com/sparc.

TABLE 2. SPARC T4 SERVER FEATURES

	SPARC T4-1	**SPARC T4-2**	**SPARC T4-4**
Size (Rack Units)	2U	3U	5U
Processor	SPARC T4 2.85 GHz		SPARC T4 3.0 GHz
Max. Processor Chips	1	2	4
Max. Cores/Threads	8/64	16/128	32/256
Max. Memory (with 32 GB DIMMs)	512 GB	1 TB	2 TB
PCIe Gen2 Slots	6	10	16
1 GbE/10 GbE Ports	4/2	4/4	4/8
Drive Bays (SAS)	8	6	8
Service Processor	Oracle ILOM		
Operating System	Oracle Solaris 11.1 or Oracle Solaris 10 1/13		
Virtualization Features	Oracle VM Server for SPARC (formerly known as Logical Domains), Oracle Solaris Zones		
Key RAS Features	Oracle ILOM, RAID 0/1, ECC correction Redundant, hot-plug fans and power supplies		

TABLE 3. SPARC T5 SERVER FEATURES

	SPARC T5-1B	SPARC T5-2	SPARC T5-4	SPARC T5-8
Size (Rack Units)	Blade server	3U	5U	8U
Processor	SPARC T5 3.6GHz			
Max. Processor Chips	1	2	4	8
Max. Cores/Threads	16/128	32/256	64/512	128/1024
Max. Memory (with 32 GB DIMMs)	512GB	1TB	2TB	4TB
Drive Bays	2	6	8	8
PCIe 3.0 Slots	2 modules	8	16	
10GbE ports	2x 10/100/1000 Ethernet connection	4x 10GbE ports	4x 10GbE ports	4x 10GbE ports
Service Processor	Oracle ILOM			
Operating System	Oracle Solaris 11.1 or Oracle Solaris 10 1/13			
Virtualization Features	Oracle VM Server for SPARC, Oracle Solaris Zones			
Key RAS Features	Oracle ILOM, RAID 0/1, ECC correction			
	In blade chassis	Redundant, hot-plug fans and power supplies		
	N/A		Hot-plug Disks and PCIe cards	

Overview of SPARC T4 Servers

SPARC T4-1 Server

The compact SPARC T4-1 server provides significant computational power in a space-efficient, low-power 2U rack mounted package. With a low price-to-performance ratio, low acquisition cost, and integrated 10 GbE, this server is well suited to delivering horizontally scaled transaction and Web services that require fast network performance. Because this server model reduces power consumption and has a small physical footprint, it addresses many operating cost challenges in modern datacenters. Coupled with Oracle Real Application Clusters (Oracle RAC), or Single Instance Oracle Database 12c SPARC T4-1 servers can provide a sound foundation for small deployments of the Oracle Optimized Solution for Oracle Database.

SPARC T4-2 Server

The expandable SPARC T4-2 server is optimized to deliver transaction and Web services, including Java Platform Enterprise Edition application services and enterprise application services such as enterprise resource planning (ERP), customer relationship management (CRM), and supply chain management (SCM). When coupled with Oracle RAC, or Single Instance Oracle Database 12c the SPARC T4-2 server is appropriate for deploying the Oracle Optimized Solution for Oracle Database in mission-critical environments. With its expansion capabilities and integrated virtualization technologies, this server is also an ideal platform for consolidated Tier 1 and Tier 2 workloads.

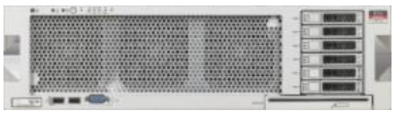

Figure 3. SPARC T4-2 Server.

SPARC T4-4 Server

With support for up to 256 threads, memory up to 2 TB, cryptographic acceleration, and integrated on-chip I/O technology, the SPARC T4-4 server is ideal for providing high throughput within significant power, cooling, and space constraints. With breakthrough levels of price/performance, this server is ideally suited for the Oracle Optimized Solution for Oracle Database, especially for consolidation and virtualization initiatives. When coupled with Oracle RAC or Single Instance Oracle Database 12c, medium-to-large Oracle Database workloads can be cost-effectively implemented on these servers.

Overview of SPARC T5 Servers

SPARC T5-1B Server

Oracle's SPARC T5-1B server provides the power of a single SPARC T5 processor in a blade form-factor. Hosted in Oracle's Sun Blade 6000 chassis, this dense and scalable SPARC blade server is ideal for customers that require high performance and an integrated platform with the deployment flexibility that blade servers provide. In such environments, the SPARC T5-1B server can accelerate both single- and multi-threaded enterprise applications, including Oracle Database, middleware, and Web services.

SPARC T5-2 Server Overview

The SPARC T5-2 is an ideal server for large departmental deployments, perfect for consolidating business processes, improving operating efficiency, and reducing data center overhead. For business-critical Oracle Database workloads and Oracle RAC implementations or Single Instance Oracle Database 12c, it is a superb choice to put the Oracle Optimized Solution for Oracle Database into practice. The server's density—with twice the cores, threads, memory capacity, and I/O bandwidth of the SPARC T4-2 server—helps to consolidate and virtualize enterprise databases with less hardware and in a smaller overall footprint.

SPARC T5-4 Server Overview

In a dense 5U form-factor, the SPARC T5-4 is a high-performing, four-socket server optimized for data-intensive, large Oracle Database workloads. It delivers impressive single-thread and multi-thread throughput performance, with a 1.2x improvement in single-thread performance, 2.5x throughput improvement, and 2x increase in I/O bandwidth compared to the SPARC T4-4 system. For large-scale environments that require extremely high service levels, the SPARC T5-4 is an optimal platform for mission-critical server and application consolidation, especially Oracle RAC deployments or Single Instance Oracle Database 12c that follow the architecture of the Oracle Optimized Solution for Oracle Database.

Figure 4. SPARC T5-4 Server.

SPARC T5-8 Server Overview

With eight SPARC T5 CPUs, 128 cores, and 4TB of system memory capacity, the SPARC T5-8 server provides extremely high compute density in an 8U enclosure. This system offers massive I/O bandwidth and high throughput, making it an ideal platform for virtualization and other I/O-intensive applications. In implementations of the Oracle Optimized Solution for Oracle Database, this server is designed for demanding business-critical workloads and ongoing in-chassis growth, especially in support of large consolidation initiatives or expanding database workloads. In conjunction with Oracle RAC or Single Instance Oracle Database 12c, large Oracle Database workloads can achieve the 24/7continuity needed to support strategic business applications. This server is also integrated in the Oracle SuperCluster for the Oracle Optimized Solution for Enterprise Database Cloud.

Figure 5. SPARC T5-8 Server.

SPARC M6-32 Servers

For the largest implementations of the Oracle Optimized Solution for Oracle Database, Oracle's SPARC M6-32 servers (Figure 6) offer massive symmetric multiprocessing (SMP) scalability with support for up to 32 SPARC M6 processors, 32 TB of memory, and a high-throughput I/O architecture. The SPARC M6-32 server is designed specifically for extremely large mission-critical workloads requiring 24/7availability, including large Business Intelligence/Data Warehousing (BIDW) and OLTP applications. It is also an ideal foundation for ERP, SCM, CRM deployments that require mission-critical reliability.

The SPARC M6-32 server is powered by the new 12-core, 8 threads-per-core SPARC M6 processor that features a 48 MB shared Level 3 cache. The system delivers high performance per processor core and extreme system throughput with up to 3,072 total processor threads in a 32-processor configuration. Combined with Oracle Solaris, the SPARC M6-32 server delivers leading performance and unmatched reliability, availability, and serviceability (RAS) with comprehensive enterprise-class virtualization capabilities built-in at no additional cost.

The SPARC M56-32 server combines the best of mission-critical data center server technology with the features from the high-throughput SPARC T-Series servers to deliver the most powerful and scalable enterprise-class server in the Oracle product family. Since the M6 processor leverages the same S3 core as the new SPARC T5 processors, it supports Oracle VM Server for SPARC technology along with Oracle Solaris Zones. (See the discussion of these no-charge server virtualization technologies on page 15.) The SPARC M6-32 server is managed using the common Oracle Integrated Lights Out Manager (Oracle ILOM) system management used in the SPARC T-Series server family and Oracle engineered systems. The ILOM management technology is directly visible through the Oracle Enterprise Manager Ops Center 12*c* software, providing a consistent and comprehensive view for managing all Oracle platforms.

The SPARC M6-32 provides the highest level of enterprise class RAS features, including redundant, hot-swap, and hot-plug capability. Other key high-availability features include use of Physical Domains for hardware fault isolation, end-to-end ECC memory protection, and the fault management architecture of Oracle Solaris. These capabilities enable server self-healing. No single system failure can prohibit system recovery, allowing this system to meet the most demanding service level agreements for non-stop mission-critical computing.

Figure 6. The SPARC M6-32 server is the optimal platform to deploy the Oracle Optimized Solution for Oracle Database in a large-scale, mission-critical environment.

Table 4 summarizes characteristics of the SPARC M6-32 servers. Additional information about this server can be found at http://www.oracle.com/sparc.

TABLE 4. SPARC M6-32 SERVER FEATURES

Processors	• SPARC M6 at 3.6 GHz • Maximum of 32 sockets • 384 cores/3072 threads • 48 MB shared L3 cache
Memory	• Up to 32 TB • 1024 DIMM slots for 16 GB and 32 GB DDR3 DIMMs
Modular architecture	• CMU: 16 maximum (2x SPARC M6 processors and 64 memory DIMMs) • IOU: 4 maximum (16 PCIe cards, 8 drive bays and 4 base IO/ 2x 10GbE slots)
Virtualization	Up to 4 maximum Physical Domains and 512 Logical Domains; also Oracle Solaris Zones
Internal Storage	Up to 32 drives (2.5” HDD or SSD)
Internal I/O Slots	64 LP PCIe 3.0 (x8 wired) slots (via a PCIe hot-plug carrier)
Network and I/O ports	32 10 GbE (100 Mbps/1Gbps/10Gbps) using configured Base I/O cards
Service Processor	Oracle ILOM
Operating System	• Oracle Solaris 11.1 • Oracle Solaris 10 1/13 (using Oracle Solaris guest)
Key RAS Features	• Hot-plug drives, PCIe cards, SP modules • Redundant and hot-swappable power supplies and fans

- Environmental monitoring, electronic prognostics
- Fault Manager architecture including Predictive Self Healing in Oracle Solaris

Oracle Integrated Lights Out Manager

The Oracle Integrated Lights Out Manager (Oracle ILOM) facilitates systems management across all Oracle servers—including the SPARC T4, T5, and M6 servers. ILOM accesses an integrated system service processor in each platform that also supports power management and power-capping capabilities to control power and reduce energy-related costs.

Oracle ILOM provides full remote keyboard, video, mouse, storage (rKVMS) support together with remote media functionality. In addition, Oracle ILOM is an integral part of Oracle Enterprise Manager Ops Center, which provides comprehensive management of all Oracle servers, operating systems, and Oracle Solaris virtualization technologies. Oracle Enterprise Manager Ops Center even permits control over logical domain and zone provisioning as well as resource allocation for these virtual environments. With its integrated lifecycle management and automated functionality, Oracle Enterprise Manager Ops Center helps to dramatically improve the efficiency of IT operations.

No-Charge, Lightweight Virtualization

The Oracle Optimized Solution for Oracle Database helps to increase service levels and business agility through its use of lightweight virtualization technologies. Recognizing the strategic importance of virtualization when consolidating servers, developing cloud-based delivery models, and shifting resources to critical applications, Oracle integrates virtualization technologies into its servers and the Oracle Solaris operating system—at no additional cost.

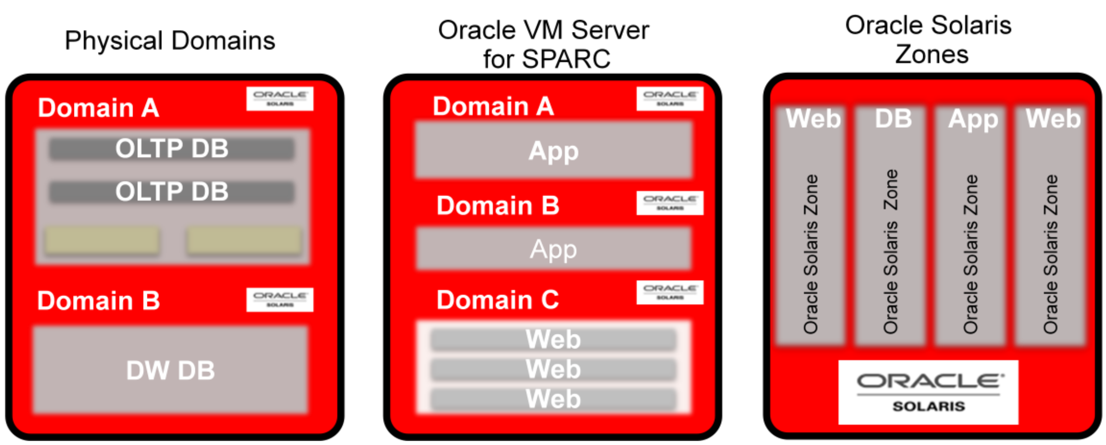

Figure 7. Virtualization technologies.

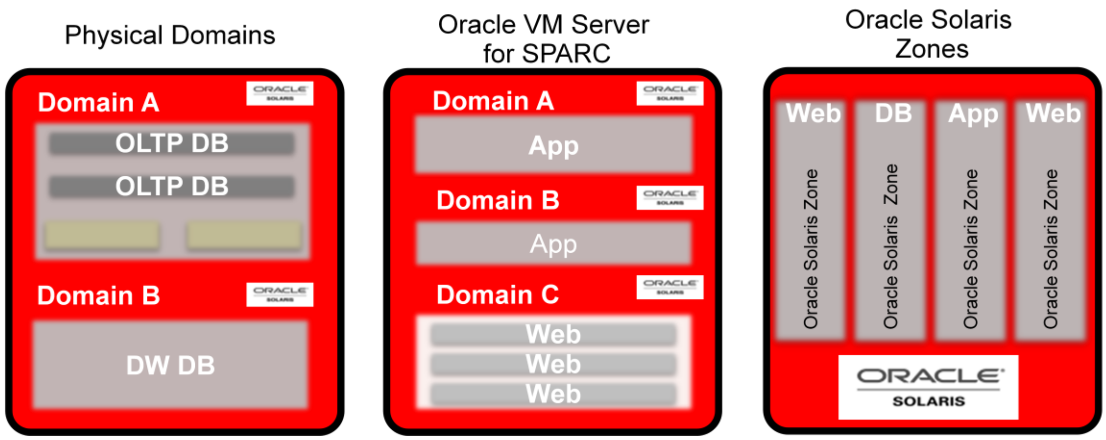

Figure 7 depicts the three built-in virtualization capabilities that the Oracle Optimized Solution for Oracle Database can put into practice:

- Physical Domains (PDoms) available in the SPARC M6-32 server
- Oracle VM Server for SPARC (Logical Domains, or LDoms), available in across all Oracle T-Series servers, including the SPARC T4 Series, T5 Series, and M6-32 servers
- Oracle Solaris Zones, available in Oracle Solaris 10 and 11 operating systems
- Additionally, with Oracle Database 12c Oracle offers very flexible consolidation and virtualization options directly through the Database Software in a native mode of operation through Pluggable Database options.

To simplify deployments and manage virtual environments, Oracle Enterprise Manager Ops Center provides an administrator-friendly interface to these different virtualization technologies. Each of these technologies is described in more detail in the following pages. It's important to remember that these virtualization technologies are often used together (Figure 8). For example, it's common to create Oracle Solaris Zones on top of Physical Domains (on SPARC M6-32 servers) or on Logical Domains defined using Oracle VM Server for SPARC.

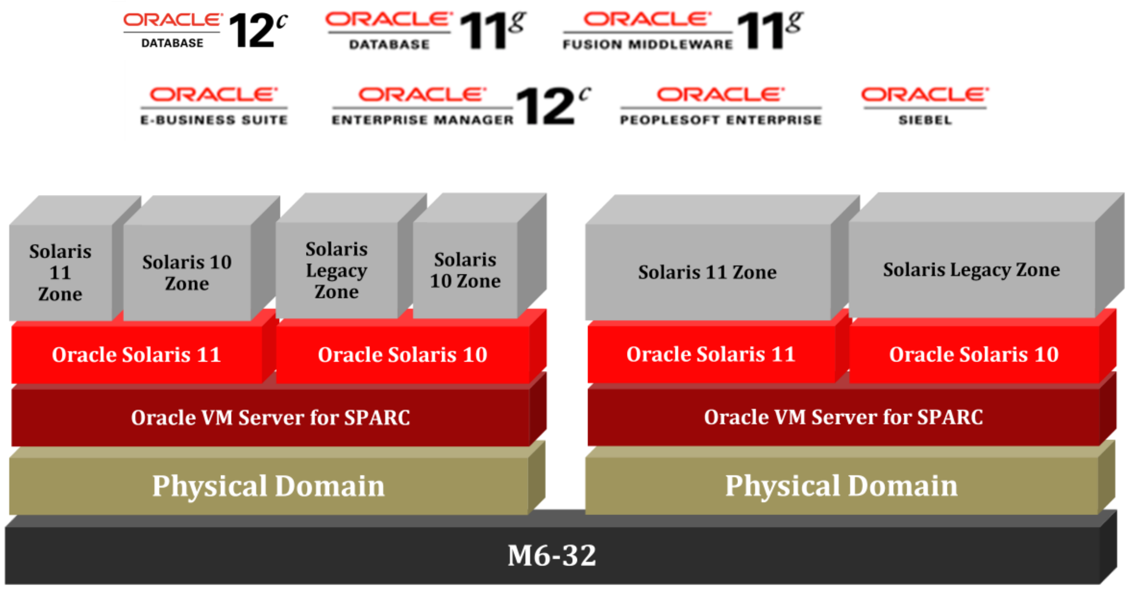

Figure 8. Oracle's virtualization technologies can be combined to enable workload isolation and flexible resource controls.

In conjunction with Oracle Database deployments, virtualization capabilities can play an extremely important role in helping to simplify database cloning (when creating application development and testing environments) or to update infrastructures without downtime.

Physical Domains

IT organizations can divide a single large SPARC M6-32 system into multiple fault-isolated domains called Physical Domains (or PDoms), each running independent instances of the Oracle Solaris operating system with access to designated I/O devices. Compute, memory, and I/O resources are assigned to different Physical Domains in the SPARC M6-32 server, and can be sized from 4 up to 32 SPARC M6 processors per PDom. When more than 8 CPUs are assigned to a PDom, the processors communicate across the system interconnect.

When system components are exclusively dedicated to individual PDoms, hardware or software faults in one domain remain isolated and unable to impact the operation of other domains. Each domain within a single SPARC M6-32 platform can run a different version of the Oracle Solaris operating system, making this technology extremely useful for pre-production testing of new or modified applications or for consolidation of multiple tiers of a database application.

Physical Domains enable organizations to tailor the compute capacity of SPARC M6-32 servers to meet specific enterprise needs. For instance, the SPARC M6-32 server can be configured as a single domain with up to 32 SPARC M6 processors to host an exceptionally compute-intensive application but more importantly to support database in-memory operations which yield extreme database performance capabilities virtually bypassing many latency IO operations of traditional non in-memory database systems . An organization with multiple databases that require isolation from one another might divide a single SPARC M6-32 server into multiple fully isolated domains, as shown in Figure 9.

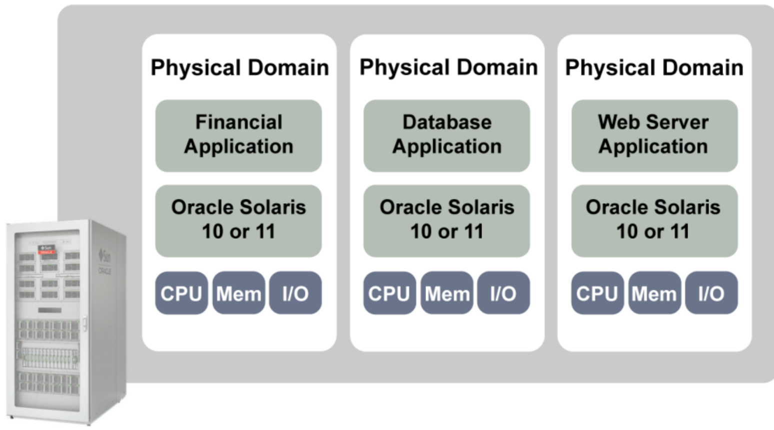

Figure 9. Organizations isolate different applications using independent Physical Domains.

Dynamic Reconfiguration in Oracle Database 11*g* Release 2 and Oracle Database 12c

Oracle Database 11*g* Release 2 and Oracle Database 12c includes several features that enable changes to be made to the instance configuration dynamically. For example, the dynamic SGA infrastructure can be used to alter an instance's memory usage. Dynamic SGA enables the size of the buffer cache, the shared pool, the large pool, and the process-private memory to be changed without shutting down the database instance. Oracle also provides transparent management of working memory for SQL execution by self-tuning the initialization runtime parameters that control allocation of private memory.

Another type of dynamic reconfiguration occurs when the Oracle Database polls the operating system to detect changes in the number of available CPUs and reallocates internal resources. In addition, some initialization parameters can be changed without shutting down the instance. The ALTER SYSTEM statement can be used to change the value of a parameter of the Oracle Database instance while the Oracle Database is actively running.

Oracle VM Server for SPARC

Oracle VM Server for SPARC is a lightweight virtualization technology supported on all Oracle CMT servers, including SPARC T4, T5, and M6 processor-based servers. While this functionality has existed in the SPARC T-Series servers for some time, it is now made possible in SPARC M6-32 systems because they incorporate multicore/multithreaded SPARC M6 processors that take advantage of the S3 core (see page 42 for more information on the SPARC M6 processor design).

Oracle VM Server for SPARC leverages a built-in hypervisor to subdivide system resources down to processor thread, cryptographic processor, memory, and PCI bus, and creating virtual partitions called logical domains (LDoms). Each logical domain runs in one or more dedicated CPU threads.

Oracle VM Server for SPARC delivers the following features and capabilities:

- **Live Migration** clones and migrates an active domain to another physical machine while maintaining application services. On-chip cryptographic accelerators enable secure wire-speed encryption—without any additional hardware—allowing domains (and any sensitive data they contain) to be migrated securely, even across public networks.
- **PCIe direct I/O** assigns either individual PCIe cards or entire PCIe buses to a guest domain. This delivers native I/O throughput.
- **Advanced RAS capabilities** for the logical domain include virtual disk multipathing and failover, as well as faster network failover with link-based IP multipathing (IPMP) support. The logical domain can also handle path failure between an I/O domain and storage. Moreover, the domain is fully integrated with the Oracle Solaris Fault Management Architecture (FMA), which enables predictive self-healing.
- **CPU whole core allocation and core affinity capabilities** enable organizations to optimize the assignment of virtual CPUs to deliver higher and more-predictable performance for all types of application workloads.
- **CPU Dynamic Resource Management (DRM)** enables a resource management policy and domain workload to trigger the automatic addition and removal of CPUs, which helps to better align business priorities.
- **Physical-to-Virtual (P2V) Conversion** transforms an existing SPARC server running the Solaris 8, Solaris 9, or Oracle Solaris 10 operating system into a compatible virtual environment running on Oracle Solaris 10. In this way, legacy applications can be consolidated onto virtualized environments and take advantage of new hardware platforms without the need for an immediate upgrade.
- **CPU power management** conserves power by disabling each core when all of its CPU threads are idle. Adjusting CPU clock speed, managing memory power, and establishing power use limits help to match energy consumption to utilization.
- **Advanced network configuration** supports greater flexibility, higher network performance, and scalability with these features: Jumbo frames, VLANs, virtual switches for link aggregations, and network interface unit (NIU) hybrid I/O.

Oracle Solaris Zones

Available in Oracle Solaris 10 and 11 releases, Oracle Solaris Zones are a lightweight virtualization technology that creates multiple private execution environments within a single Oracle Solaris instance. Oracle Database instances running within zones are completely isolated, preventing processes in one zone from affecting processes running in another. Virtual networks within zones help to isolate data, which can be a critical requirement for some applications.

Oracle Solaris Zones feature extremely fast initialization, and can be configured to instantly start, stop, or restart Oracle Database and Oracle RAC services. Oracle Solaris Zones provide flexibility in deployment, allowing easy configuration changes within defined physical or logical domains. Zones simplify the cloning and instantiation of new database environments, making it possible to prioritize database workloads within domains and allowing system resources to be reassigned to handle peak or seasonal workloads.

Together, these three virtualization technologies—Physical Domains, Oracle VM Server for SPARC, and Oracle Solaris Zones—are critical enablers to consolidate database workloads safely within a single platform. They provide a foundation for efficiently and securely consolidating databases on a single machine. An assortment of test, development, and production databases can be combined and run without being impacted by database services executing in other virtual environments. Organizations that must safeguard access and protect sensitive data, such as governments, financial institutions, and HR departments, can safely segregate Oracle Database instances in configured virtual environments.

Virtual environments permit fine-grained control of system compute, memory, networking and I/O resources. This allows administrators to create isolated environments for Oracle Database instances and Oracle RAC configurations.

When the Oracle Optimized Solution for Oracle Database is deployed, a site-specific layout is created, typically using physical or logical domains and Oracle Solaris Zones to provide the optimal design for flexibility and isolation. With Oracle Solaris Zones, it is possible to maintain a one-application-per-virtual-server deployment model while simultaneously sharing hardware resources.

Oracle Solaris for Scalability

Oracle Solaris is an industry-leading operating system designed to handle enterprise, business-critical operations. In addition to built-in virtualization with Oracle Solaris Zones, it provides key functionality, optimizing resource utilization, high availability, security, and performance for vertically and horizontally scaled environments. Oracle Solaris runs on a broad range of SPARC (and x86-based) systems. Oracle SPARC/Solaris binary compatibility is guaranteed, making it easy to migrate legacy software and in this way protect application investments.

Oracle Solaris includes the ZFS file system for superior data integrity, advanced security protection and management, and scalable performance due to efficient thread scheduling on multicore processors. Innovations like DTrace, Predictive Self-Healing, and the Service Management Facility (SMF) have made Oracle Solaris the operating system of choice for applications that demand business-critical performance, scalability, and availability.

Oracle Solaris 11 features several scalability enhancements, enhanced kernel data structures, memory management, and library optimizations to support large-scale database workloads. In addition to the OS virtualization capabilities of Oracle Solaris Zones described earlier, Oracle Solaris 10 and 11 incorporate features to improve performance on Oracle's multicore/multithreaded architectures.

Accelerated Cryptography

Accelerated cryptography is supported through the cryptographic framework in Oracle Solaris as well as native capabilities in the SPARC T4, T5, and M6 processors. These SPARC processors enable access to cryptographic cypher hardware implementations through user-level instructions. When Oracle Transparent Data Encryption (TDE) is used to encrypt sensitive data, Oracle Database operations directly access the cryptographic cyphers implemented in the instruction pipeline, dramatically speeding up the tasks of encryption and decryption.

Critical Thread Optimization

Oracle Solaris 10 and 11 have the ability to permit either a user or programmer to allow the Oracle Solaris Scheduler to recognize a "critical thread" by raising its priority to 60 or above through the either the Command Line Interface or system calls to a function. If this is done, the thread will run by itself on a single core, garnering all the resources of that core for itself. To prevent resource starvation to other threads, this feature is disabled when there are more runnable threads than available CPUs.

Critical threads have important performance and scalability implications for Oracle Database workloads. The Logwriter process, which logs transactions for recovery purposes, can often be a limitation for transaction throughput in Oracle Database 11*g*. Common practice today is to statically bind Logwriter to a processor core. Critical threads, however, can assign Logwriter much more dynamically, improving resource utilization. For Cache Fusion in Oracle RAC, the LMS lock manager can sometimes inhibit scalability, so using critical threads for the LMS lock manager can also enhance the ability for Oracle RAC to scale effectively.

Multicore/Multithreaded Awareness

Oracle Solaris 10 and 11 are aware of the SPARC processor hierarchy, so the scheduler can effectively balance the load across all available pipelines. Even though it exposes each of these processors as logical processors, Oracle Solaris understands the correlation between cores and the threads that they support, providing a fast and efficient thread implementation.

Fine-Grained Manageability

For the SPARC T4, T5, and M6 processors, Oracle Solaris 10 and 11 have the ability to enable or disable individual cores and threads (logical processors). In addition, standard Oracle Solaris features (such as processor sets) provide the ability to define a group of logical processors and schedule processes or threads on them. Oracle Solaris allows considerable flexibility in that processes and individual threads can be bound to either a processor or a processor set, if needed.

Support for Virtualized Networking and I/O

Oracle Solaris contains technology to support and virtualize components and subsystems on the SPARC T4, T5, and M6 processors, including support for 10 GbE ports and PCIe interfaces. As part of a high-performance network architecture, Oracle multicore/multithreaded-aware device drivers are provided so that applications running within virtualization frameworks can effectively share I/O and network devices. This also includes very high bandwidth and low latency support for InfiniBand technologies, which can be incorporated in the Oracle Optimized Solution for Oracle Database.

Oracle Solaris 11 added network virtualization enhancements including Virtual NICs (VNICs), virtual switching, network resource management, and an efficient and easy way to manage integrated Quality of Service (QoS) to enforce bandwidth limits on VNICs and traffic flows. VNICs can be assigned to Oracle Solaris Zones that share a physical NIC. Network virtualization in Oracle Solaris 11 allows a bandwidth limit to be set on a VNIC to ensure that there is a minimum bandwidth available. This capability is most useful so that one interface does not dominate network use and negatively impact other traffic.

Shared Memory

There is usually considerable commonality in the tasks undertaken by users connected to the same database instance and users running transaction-based workloads. They frequently access many of the same data blocks. For this reason, Oracle Database keeps frequently used data blocks in a cache in memory called the buffer cache, and it shares other frequently accessed information, such as table metadata and parsed (processed) SQL statements, in a second memory cache called the shared pool.

These memory caches are held in a single shared memory to allow multiple and concurrent access as well as interprocess communication. Since shared memory is heavily used in Oracle Database environments, it is important to optimize its access and to minimize the amount of CPU consumed when referring to it. With this in mind, a specially tuned variant of System V Shared Memory, called Intimate Shared Memory (ISM), was introduced in Oracle Solaris many years ago. ISM has been widely used for database memory that is shared by Oracle Database instances ever since.

First released in Solaris 8 Update 3 (1/01), Dynamic Intimate Shared Memory (DISM) provides shared memory with the same essential characteristics as ISM except that it is dynamically resizable. Thus DISM offers the same performance benefits of ISM while allowing the shared memory segment to be dynamically resized, both for the sake of performance and to support the dynamic reconfiguration of memory resources (for example, adding or removing memory from a logical domain). This dynamic resizing can be scheduled or unscheduled based upon particular operating system and associated hardware event. ISM and DISM can take advantage of large shared memory pages and other enhancements to memory management in Oracle Solaris 11.

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. 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. Users who need the functionality provided by DISM should consult the Oracle white paper "Dynamic SGA Tuning of Oracle Database on Oracle Solaris with DISM."

Finally, with Oracle Solaris 11.1 and Oracle Database 12c one can receive best practice advantages of providing for the simple use of automatic memory management through Oracle's Optimized Shared Memory integration of the best of breed Operating Systems, Solaris and Oracle's Database 12c to take care of memory allocations for consolidated pluggable databases automatically. The Oracle Optimized Solutions team has implemented the Oracle Optimized Solution for Oracle Database in this fashion with Single Instance Oracle Database 12c.

NUMA Optimizations

In Oracle's multi-processor systems, each SPARC T4, T5, and M6 processor manages memory and memory coherence—the processor implementations follow non-uniform memory access (NUMA) architectures. In NUMA architectures, a processor can access its own memory faster than accessing memory managed by another processor. Oracle Solaris provides memory placement technology that can specifically help to improve performance of NUMA memory management:

- Memory Placement Optimization (MPO)—Oracle Solaris 10 uses MPO to improve the placement of memory across the physical memory of a server, resulting in increased performance. Through

MPO, Oracle Solaris 10 works to help ensure that memory is as close as possible to the processors that access it, while still maintaining enough balance within the system. As a result, many database applications are able to run considerably faster.

- Hierarchical Lgroup Support (HLS)—In Oracle Solaris 11, HLS improves MPO memory placement. HLS helps Oracle Solaris optimize performance for systems with more-complex memory latency hierarchies. HLS lets Oracle Solaris distinguish between the degrees of memory remoteness, allocating resources with the lowest-possible latency for applications. If local resources are not available by default for a given application, HLS helps Oracle Solaris allocate the nearest remote resources.

Oracle Solaris ZFS

Oracle Solaris ZFS offers a dramatic advance in data management, automating and consolidating complicated storage administration concepts and providing unlimited scalability with the world's first 128-bit file system. Oracle Solaris ZFS is based on a transactional object model that removes most of the traditional constraints on I/O issue order, resulting in dramatic performance gains. Oracle Solaris ZFS also provides data integrity, protecting all data with 64-bit checksums that detect and correct silent data corruption.

A Secure and Robust Enterprise-Class Environment

Oracle Solaris supplies the scalability, security, and availability needed to support mission-critical database applications. For almost 20 years, Oracle Solaris has featured binary compatibility, so existing Oracle Solaris applications on SPARC can continue to run unchanged, protecting software investments. Certified multilevel security mechanisms protect Oracle Solaris environments from intrusion. A built-in Fault Management Architecture means that predictive self-healing elements can communicate directly with the hardware, helping to reduce both planned and unplanned downtime. Effective analysis and troubleshooting tools, such as Oracle Solaris DTrace, help organizations tune applications and optimize system resource use.

Predictive Self Healing and Redundant Interconnects

A standard part of the Oracle Solaris 10 and 11 operating systems, Oracle Solaris Predictive Self Healing software enhances the innate reliability of all SPARC servers. The implementation of Oracle Solaris Predictive Self Healing software for SPARC servers enables constant monitoring of all CPUs and memory. Depending on the nature of an error, persistent CPU soft errors can be resolved by automatically taking a thread, a core, or an entire CPU offline. Similarly, memory page retirement capability allows proactive off-lining of memory pages in response to repeated data access errors from a specific DIMM.

Oracle NAS and SAN-Based Storage Options

When consolidating Oracle Database services using the Oracle Optimized Solution for Oracle Database, customers have a choice in the underlying shared storage for the solution. Sites may choose to leverage existing investments in SAN-based storage, or they can take advantage of Oracle's highly integrated NAS or SAN-based storage solutions: the Oracle ZFS Storage Appliance or Oracle's Pillar Axiom storage systems.

Oracle ZFS Storage Appliance Overview

Oracle ZFS Storage Appliance family provides enterprise-class NAS capabilities and a complete suite of data services including snapshots, clones, replication, and industry-leading performance analytics, as well as native compression, deduplication, Oracle Database integration, and support for the Hybrid Columnar Compression (HCC) feature (HCC is further discussed on page33). These appliances feature a comprehensive and intuitive user interface and analytics environment, simplifying storage management and reducing operating costs and complexity. The Oracle ZFS Storage Appliance optimizes performance automatically over storage tiers, providing compelling economics and cost efficiencies for enterprise storage.

To deliver high performance, the Oracle ZFS Storage Appliance file system implements Oracle Solaris ZFS. The file system is designed to recognize I/O patterns automatically and places data on the best storage media using a Hybrid Storage Pool approach. For example, Oracle Solaris ZFS transparently executes synchronous writes to low-latency solid-state drives (SSD), acknowledging the write quickly to allow the application to continue processing. Then, the file system automatically flushes the data to hard disk drives as a background task. Another type of SSD media acts as a second-level cache to reduce read latency. Oracle Solaris ZFS transparently manages the process of copying less frequently accessed data (as it ages out of memory) into this cache to accelerate read requests. The Hybrid Storage Pool approach results in impressive read and write performance.

The appliance provides administrators with the tools needed to quickly identify and diagnose system performance issues and debug live storage and networking problems before they affect the application infrastructure. The award-winning DTrace Analytics software, a feature of the Oracle ZFS Storage Appliance, provides real-time analysis and monitoring functionality. DTrace Analytics uses built-in instrumentation to provide in-depth analysis of key storage subsystems. In addition, Oracle ZFS Storage Appliances include comprehensive self-healing capabilities of Oracle's fault management architecture, automatically and silently detecting and diagnosing underlying system problems and responding by taking faulty components offline.

To meet a variety of customer needs for capacity, price, and performance, the Oracle ZFS Storage Appliance is available in several configurations, including dual cluster configurations that offer maximum availability. Each configuration comes bundled with the same software, including data protocols, compression/deduplication capabilities, and the DTrace Analytics software for system troubleshooting and performance optimization. The Sun ZFS Storage 7320 Appliance is a midrange storage solution that provides a high-availability, entry-level cluster option with scalability up to 432 TB of raw capacity. It also supports Hybrid Storage Pools that can be configured with up to 4 TB of read-optimized cache and optional write-optimized cache for enhanced application performance.

Oracle's Pillar Axiom 600 Storage System Overview

Oracle's Pillar Axiom 600 storage system was designed from the ground up to deliver business-critical, deterministic Quality of Service (QoS) consistent with the business value of the application and its data. It allows organizations to buy the SAN-based capacity they need, scale over time, and conserve capital in the process—all while delivering consistent and reliable performance through an intelligent SAN architecture.

The Pillar Axiom 600 storage system delivers enterprise-class high availability for Oracle Database deployments. It is designed to eliminate single points of failure with redundant system components:

- The Pillar Axiom Pilot is a dual-redundancy policy controller that performs management.
- The Pillar Axiom Slammer is a dual-redundancy storage controller that virtualizes the storage pool, moving and managing data.
- The Pillar Axiom Brick is a storage enclosure that houses two RAID controllers and thirteen disk drives, including one hot-swappable spare. Bricks are available with SSD, FC, and SATA drives.

The modular architecture of the Pillar Axiom 600 storage system allows administrators to quickly replace failed components without disrupting or slowing system performance. Bricks are organized into two sets of six-disk RAID-5 or RAID-10 groups. A local hot-spare disk drive is available to replace a failed disk, allowing RAID rebuilds to begin immediately and data to be restored in hours instead of days.

Sophisticated software layered on top of the hardware helps to ensure high availability. Software processes on each Slammer control unit (CU) constantly communicate on status and performance. The Pillar Axiom 600 storage system uses a double-safe write system and secures the I/O in battery-backed, nonvolatile RAM (NVRAM). The Pillar Axiom 600 storage system's redundant CU architecture secures the I/O in both CUs before it is acknowledged as a complete transaction.

Key Differentiating Technologies

The Pillar Axiom 600 storage system provides distinct advantages:

- Modular scalability provides the ability to dynamically scale performance and capacity independently.
- Application prioritization and contention management enable multiple applications and databases to co-exist on the same storage system to ensure QoS. A set of policies is created to govern QoS for each file system or LUN. (Figure 10 shows an example of the user interface for managing QoS.)
- Storage domains prevent co-mingling of data, isolating customers, users, and applications to different storage enclosures.
- Application profiles provide application-oriented provisioning and management linked to application priority. These profiles determine how data is laid out on the drives to best support application performance priorities.
- Distributed RAID provides superior scalability and performance even during drive rebuilds by moving RAID into storage enclosures.

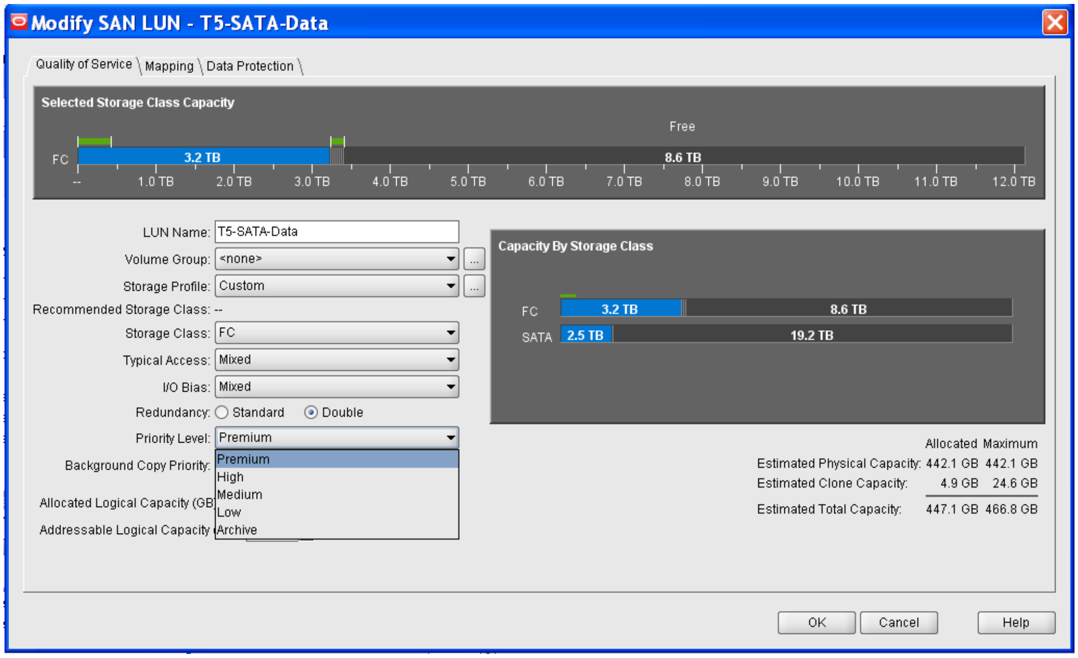

Figure 10. The Pillar Axiom user interface supports priority and storage class assignments to map the appropriate QoS to workload requirements.

Matching QoS to Requirements

Based on the overall performance and QoS required, Pillar Axiom 600 settings can be adjusted (Figure 10). The priority level can be set *archive*, *low*, *medium*, *high*, or *premium,* helping to ensure that critical workloads are guaranteed CPU allocations and favorable disk placement. If more resources are needed, the application QoS priority settings can be raised, either on a temporary or permanent basis. It's also easy to reassign and migrate a database from SATA to Fibre Channel or SSD devices (or from slower to faster RPM drives in the storage system) simply by changing the storage class assigned in the user interface. In this way, an administrator can easily assign less expensive SATA disks to certain workloads, optimizing QoS and allocating resources as needed to match strategic business goals.

Changing the storage class to migrate a data to different storage devices can be done very dynamically, with very little interruption of I/O service to the online running databases that the Pillar Axiom system is supporting. This capability helps to minimize downtime since it eliminates the need to shut down mission-critical databases to perform such a change. This functionality lends itself well to a storage upgrade, allowing data to be moved to new storage devices before removing the older storage devices on the Pillar Axiom system.

The screenshot in Figure 11 shows just how dynamic this process can be. When running an OLTP workload (generated using Swingbench, an Oracle Database workload generator) with an Oracle RAC Database hosted on two SPARC T5-2 servers, there is only a slight I/O interrupt while the database is migrated to different media. The migration is started by a simple point-and-click on the Pillar Axiom user interface. (For more information, see the video on Oracle's Pillar Axiom QoS and an overview of Oracle's Pillar Axiom 600 storage system.)

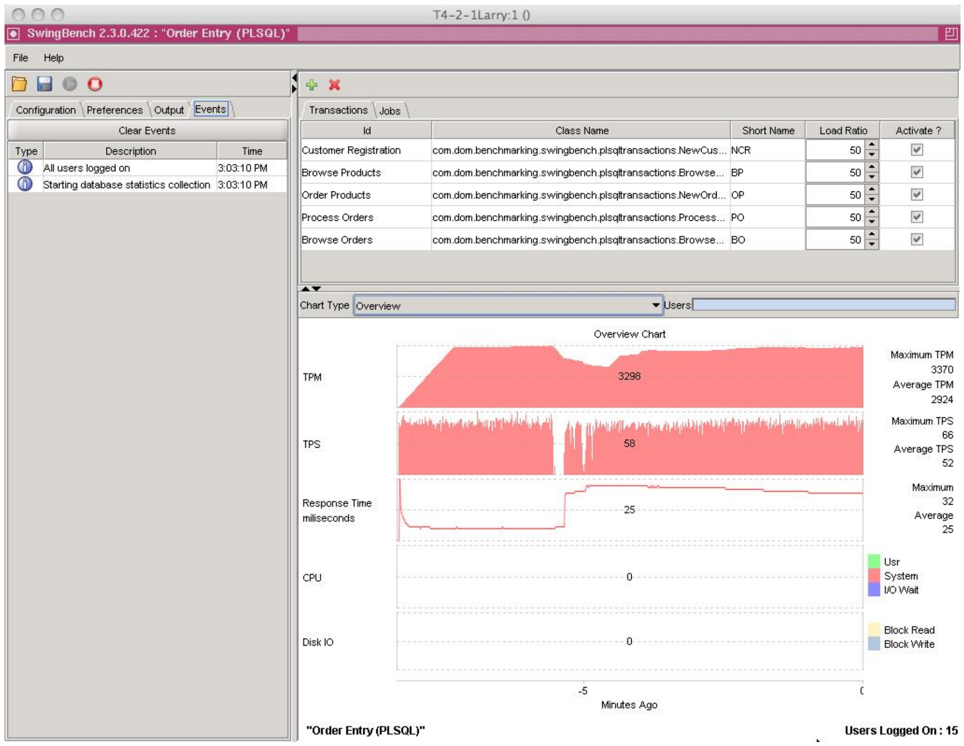

Figure 11. Only a brief I/O interrupt occurs when a database is migrated to a different storage class.

Application Profile for Oracle ASM

The storage architecture lends itself to consolidating applications and databases in a simplified but robust and reliable manner. The Pillar Axiom 600 system has specifically optimized the Oracle Application Profile to match I/O patterns generated when Oracle Automated Storage Management (ASM) is used. ASM's normal I/O pattern is to write and read 1MB RAID controller stripe settings. A special deep stripe that is exactly 1MB in size allows Oracle ASM I/Os to execute fewer internal writes to disk. The profile enables dramatically improved performance with lower disk overhead and higher overall sustained bandwidth. In addition, Pillar Axiom 600's large cache in write-back mode attempts to buffer multiple writes and write out complete RAID stripes whenever possible.

Multi-Path I/O for Throughput and Availability

Combined with multi-pathing capabilities in Oracle Solaris, software support for the Pillar Axiom 600 makes it possible to intelligently manage multiple data paths to storage and optimize I/O throughput for fast data access. Pillar Axiom software controls and manages all data paths to the storage LUNs, determining which data paths to use and routing I/O to optimized data paths when possible. The software groups multiple data paths for a Pillar Axiom LUN, presenting this group to the Oracle Solaris operating system as a single LUN or drive. In conjunction with the Pillar Axiom software, Oracle Solaris has built-in multi-pathing support, so it inherently directs I/O in a round-robin manner. Traffic is shared among available paths, ensuring that access to LUNs is not interrupted if some paths fail. In addition, parallel I/O paths yield significantly better throughput, especially for I/O-intensive operations like table scans. The implementation guide for the Oracle Optimized Solution for Oracle

Database contains additional recommendations and best practices for Pillar Axiom configurations in order to take full advantage of multi-pathing capabilities.

Oracle Multitenant

Oracle Multitenant—an Oracle Database 12*c*, Enterprise Edition option—introduces a new architecture that enables in-database virtualization for database consolidation, which delivers rapid provisioning and cloning as well as enhanced database patching and upgrades. It enables administrators to manage many databases as one and allows resource management between databases. This new architecture allows a multitenant container database (CDB) to hold many pluggable databases (PDBs), as shown in Figure 12. Please refer to the online whitepaper entitled "Oracle Multitenant on SPARC Servers and Oracle Solaris" for further details.

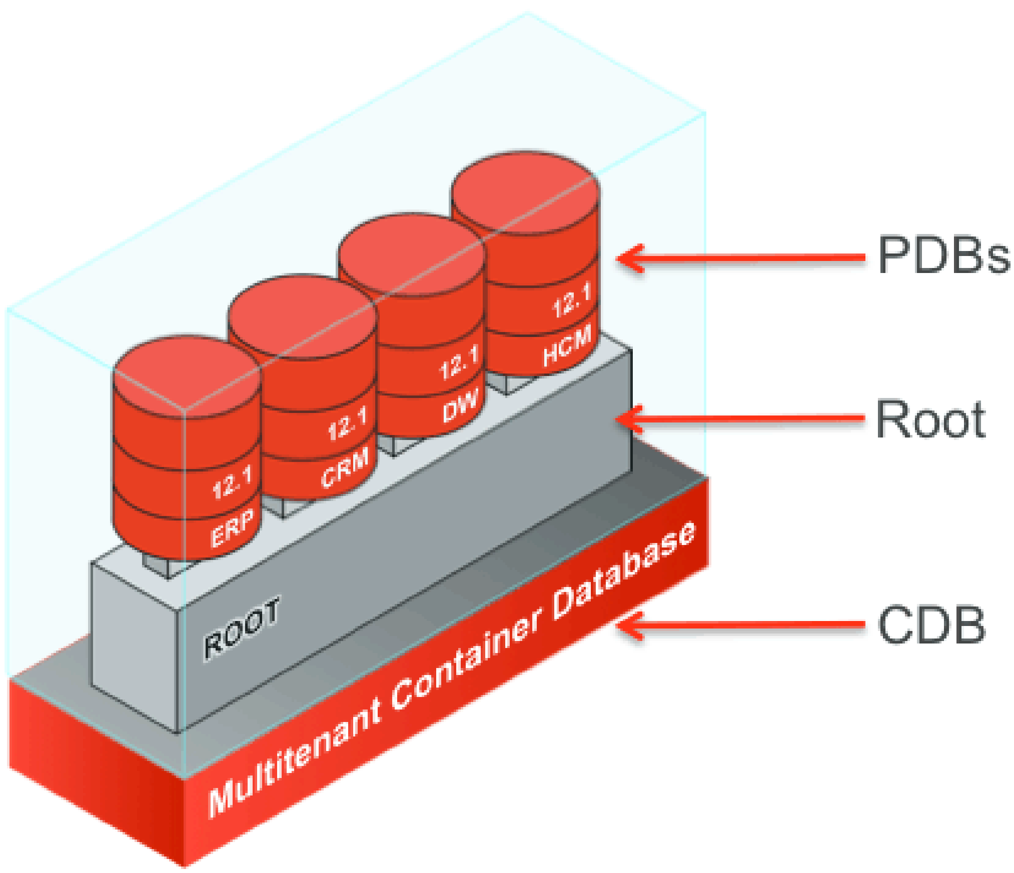

Figure 12 Oracle Multitenant

Beyond the In-Database Virtualization tests and results documented in the aforementioned whitepaper we also included the testing and deployment of this architecture on Oracle's Pillar Axiom 600 storage system on multiple media types. The Oracle Database 12c Database Configuration Assistant (DBCA) was first utilized to create container databases and pluggable databases for further consolidation on both SATA and FC media types within the Pillar Axiom storage system. Additionally, Swingbench 2.5.0.9 was used to load and access the Oracle Database 12c pluggable databases that were created via DBCA. An example of one of the oewizard and swingbench runs are shown in Figures 13 and 14.

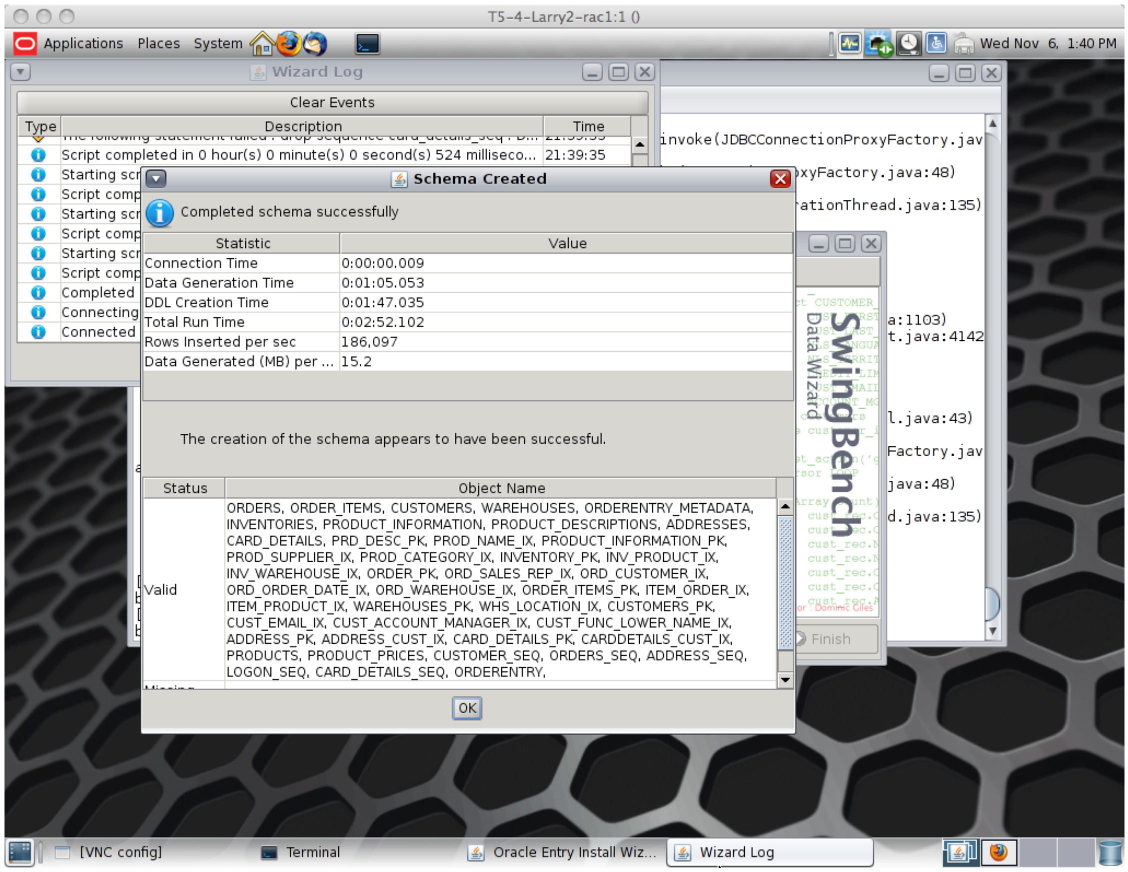

Figure 13 Oewizard Oracle Database 12c database schema creation

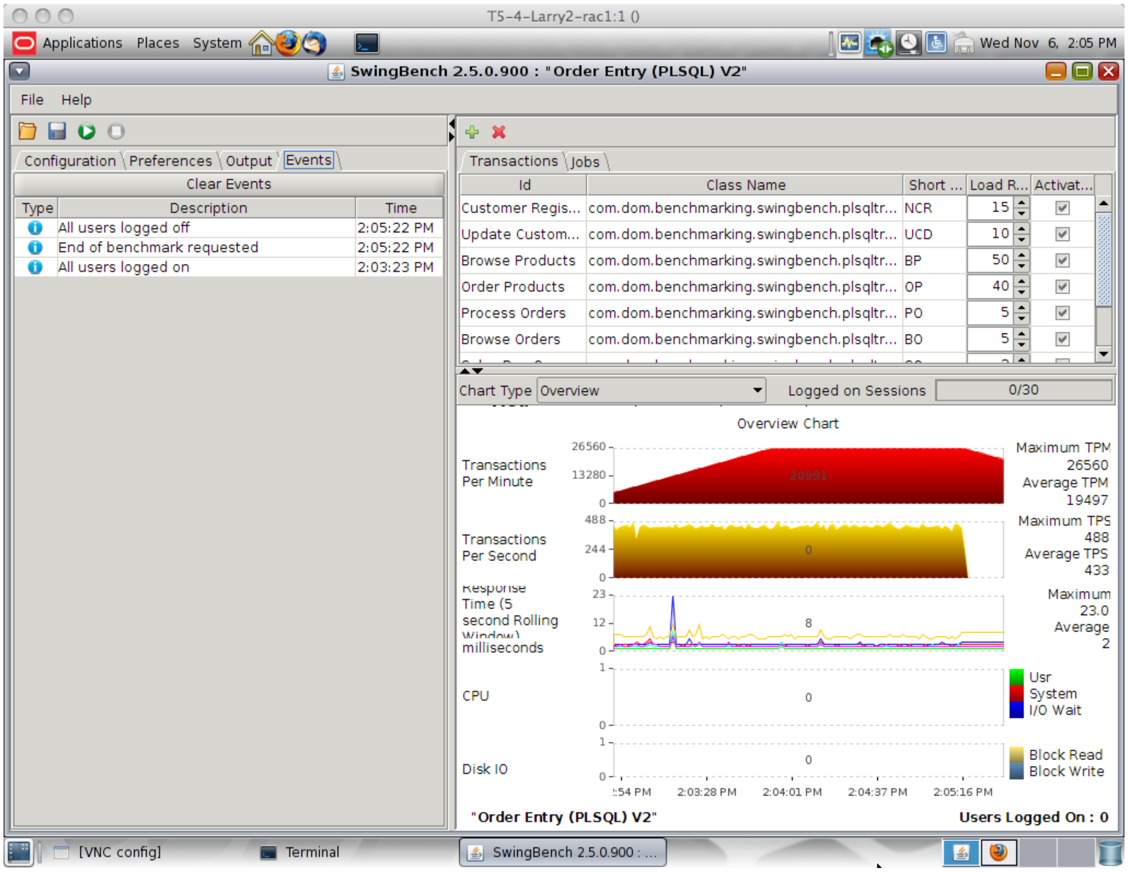

Figure 14 Swingbench Oracle Database 12c pluggable database OLTP access within Container Database

Many other things can be accomplished through this architecture. For example one could easily unplug a database that is currently being supported on SATA and easily plug the pluggable database into a FC media type within the Pillar Axiom or provide a much higher Quality of service by plugging the database into SSD media type supported on the Oracle Pillar Axiom 600. Newly introduced cloning capabilities can also be used for simple and quick deployment of development and test databases on the same local or remote container databases, as explained in the white paper "How To Accelerate Test and Development Through Rapid Cloning of Production Databases and Operating Environments."

Cloning Oracle Databases on Shared Storage

There are a many reasons that IT organizations must clone a production Oracle Database, including:

- Application development and testing. Customers often clone a production database to support application development and testing, evaluate performance impacts, and regression-test applications.

- Infrastructure updates without downtime. Mission-critical applications require continuous availability, so system updates or upgrades must occur without imposing downtime. Database clones enable upgrades while production systems continue to run.
- Troubleshooting. Replicating an existing software environment can help in identifying and resolving a problem.

One classic method of creating a copy of the Oracle Database is to restore the Oracle Database from a recent backup to a different Oracle Database server. This approach can often be extremely time consuming. Instead, using cloning capabilities of the Oracle ZFS Storage Appliance or the Pillar Axiom 600 storage system can help to streamline the database duplication process Additionally, with Oracle Database 12c one can easily clone a pluggable database natively with the Database Configuration Assistant (DBCA). The Oracle Optimized Solution for Oracle Database has taken advantage of this new feature and incorporates the use of this within the solution as shown in Figure 15. Additionally, Swingbench 2.5.0.9 was used to access the cloned Oracle Database 12c pluggable database created via DBCA as shown in Figure 16.

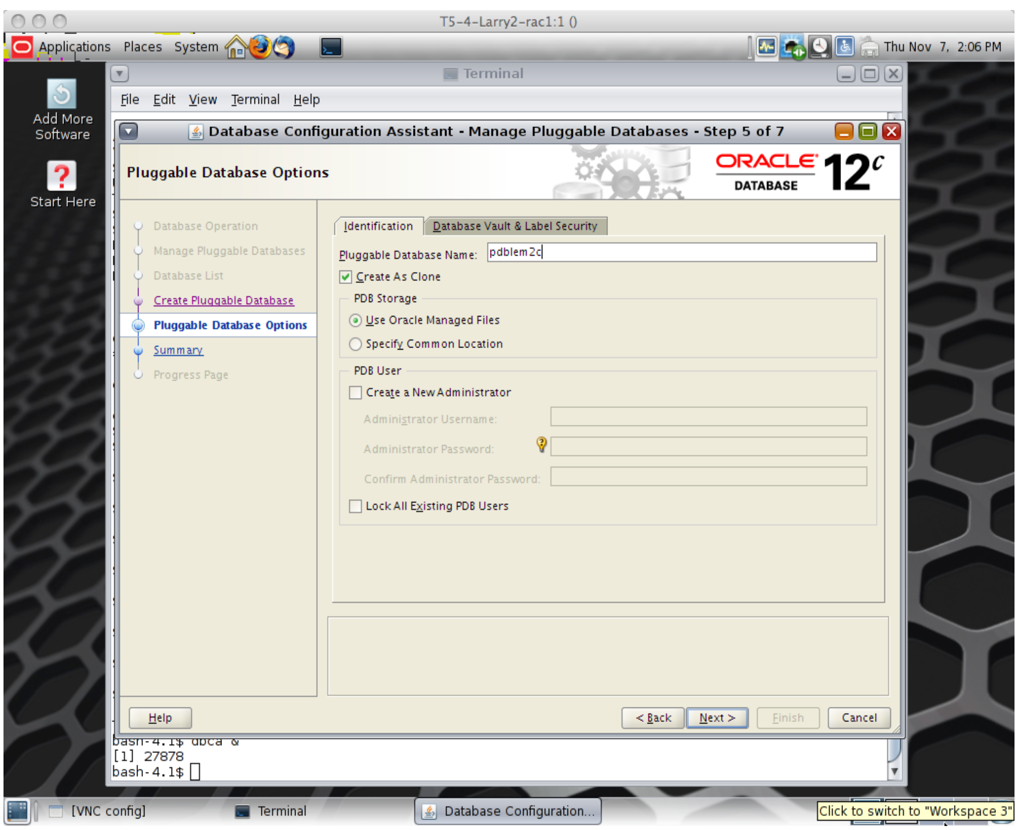

Figure 15. Natively cloning an Oracle Database 12 c pluggable database with Database Configuration Assistant (DBCA)

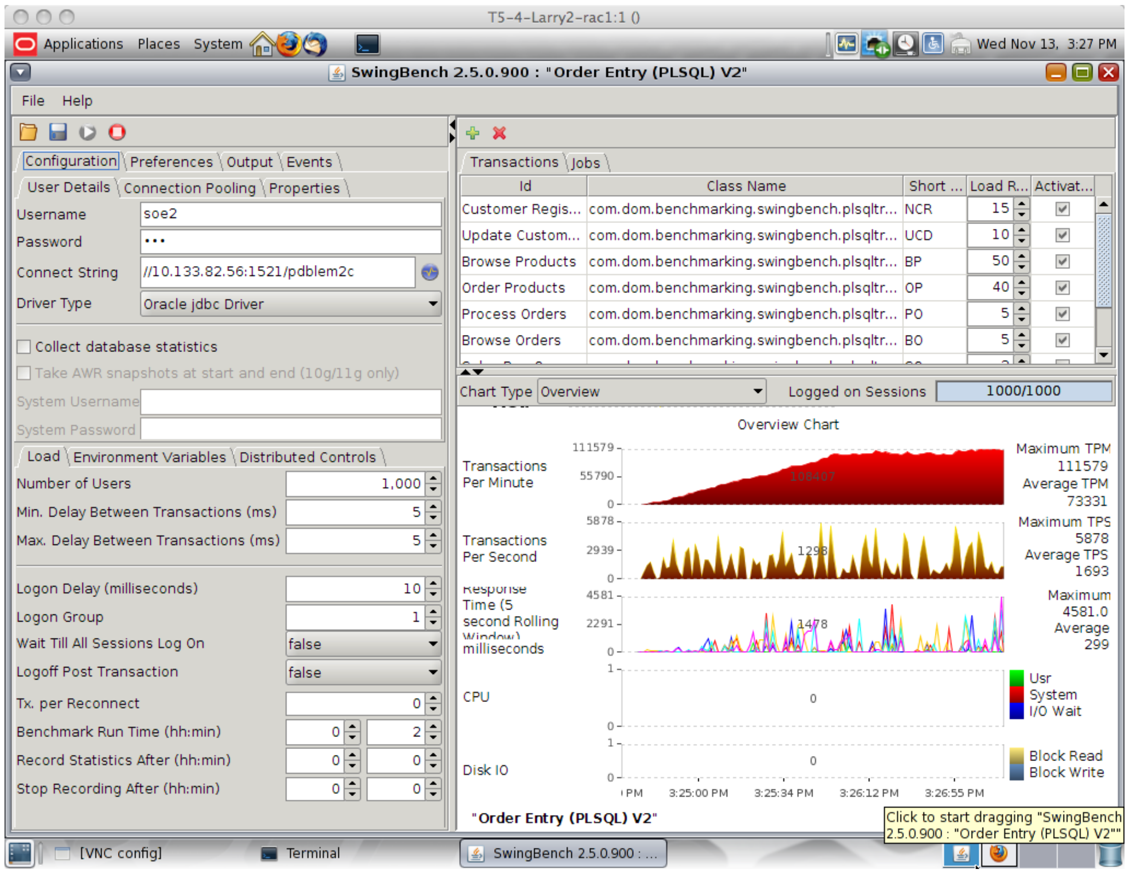

Figure 16 Swingbench 2.5.0.9 accessing the cloned Oracle Database 12c pluggable database created via DBCA in Figure 15

Hosting the database on a virtual environment created with Oracle Solaris Zones enables the creation of a golden image that can be replicated as needed. The entire duplication process can be executed using a script that creates, configures, and provisions independent Oracle Solaris Zones, and then duplicates the Oracle Database. Snapshot and cloning capabilities in the Oracle ZFS Storage Appliance make it easy to clone zones and the Oracle Database instance. The process is outlined in the paper "How To Accelerate Test and Development Through Rapid Cloning of Production Databases and Operating Environments." This paper describes a single script for database cloning that uses Oracle Recovery Manager (RMAN) backup technology, the CloneDB feature of Direct NFS (dNFS) Client introduced in Oracle Database 11*g*, Oracle Solaris Zones, and Oracle ZFS Storage Appliance snapshots and cloning capabilities. For deployments with the Oracle ZFS Storage Appliance, this method simplifies the process of replicating an Oracle Database and the entire Oracle Solaris Zone. Oracle engineers have observed as much as a 50x speedup in the database replication process through the use of this cloning method.

For deployments that use Pillar Axiom 600 Storage system, there are alternate approaches to cloning. The paper "Oracle Database Cloning Solution Using Oracle Recovery Manager, Oracle's Data Guard, and Oracle's Pillar Axiom 600 Storage System" describes two different methods to create a duplicate instance of a production Oracle Database using advanced features of the Pillar Axiom Storage system. The two methods are:

- Cloning a production Oracle Database managed by Oracle ASM and Oracle RMAN on a Pillar Axiom 600 storage system.
- Cloning a standby database instance when using Oracle's Active Data Guard on a Pillar Axiom 600 storage system.

Technology Highlights

Oracle Hybrid Columnar Compression for Oracle Database

Oracle Database-aware technologies have the advantage of compressing data before sending it to the storage system, which can have two positive outcomes: moving less data and higher performance. Unique to Oracle storage, Hybrid Columnar Compression (HCC) technology is a technique for organizing data within a database block. As the name implies, HCC uses a hybrid approach—a combination of both row and columnar methods—for storing data. The approach achieves the compression benefits of columnar storage but avoids the performance shortfalls often associated with a pure columnar format.

Enterprises with existing Oracle Databases and in-database archives for OLTP, data warehousing, or mixed workloads can achieve 10x to 50x reductions in data volumes by using Oracle HCC on the Pillar Axiom 600 or Oracle ZFS Storage Appliance products. Sites that implement HCC have seen 3x to 5x reductions in overall storage footprints. Because data moved between the Oracle Database server and storage is compressed, using HCC can result in significant performance advantages also. Figure 17 shows the results of an Oracle-conducted study in which customers achieved a compression ratio of 13x (on average) using HCC and Oracle storage.

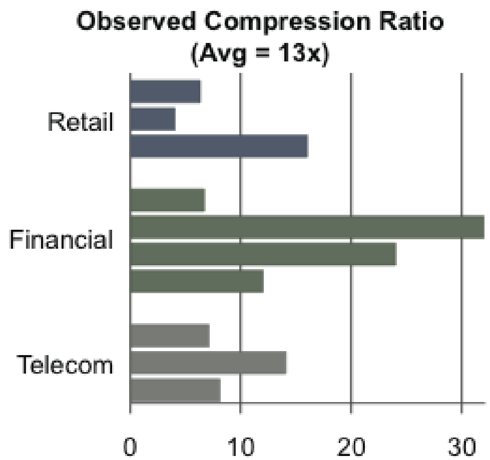

Figure 17. In real-world testing, HCC achieved compression ratios averaging approximately 13x.

When either the Pillar Axiom storage system or the Oracle ZFS Storage Appliance is attached to the database server, Hybrid Columnar Compression can be enabled in conjunction with the Oracle

Database. (It is also enabled for Oracle Databases on Oracle Exadata storage.) Fibre Channel connectivity is supported using Oracle ASM with the Pillar Axiom storage system. NFS connectivity, via the Oracle Database 11*g* Direct NFS (dNFS) Client, is supported with the Oracle ZFS Storage Appliance. For more information, see "How to Multiply the Capacity of Oracle Storage Systems With Hybrid Columnar Compression."

Oracle RAC for High Availability

Oracle Real Application Clusters (RAC) deploy database instances on multiple hosts that act upon a single database residing on shared storage. The database is highly available because access to the shared data is available as long as one Oracle RAC instance is online. When deployed in Oracle Solaris Zones, Oracle RAC adjusts to take advantage of new or changed resource levels, making it easier to optimize availability and scalability.

Each server in an Oracle RAC cluster is connected to a private network via a private interconnect. New Oracle RAC instances can be added on new hosts up to the configuration limits of the cluster and its interconnect. The interconnect serves as the communication path between nodes in the cluster to synchronize the use of shared resources by each instance.

Each database instance in an Oracle RAC database uses its own memory structures and background processes. When a data block located in the buffer cache of one instance is required by another instance, Oracle RAC uses Cache Fusion to transfer the data block directly between the instances using the private interconnect. This is much faster than having one database instance write the data blocks to disk and requiring the other database instance to reread the data blocks from disk. Cache Fusion technology enables the Oracle RAC database to access and modify data as if the data resided in a single buffer cache.

InfiniBand, Application Latency, and Fabric Convergence

Moving data between applications over a traditional network can be time consuming and can drain precious server resources. With traditional network technologies, data exchanges traverse operating systems on both source and destination servers, sometimes resulting in excessive application latencies because of system calls, buffer copies, and interrupts. The Oracle Optimized Solution for Oracle Database can take advantage of Oracle's InfiniBand networking (for the Oracle RAC private interconnect, for example) or 1Gb and 10Gb Ethernet, depending on site requirements.

InfiniBand Remote Direct Memory Access (RDMA) technology provides a direct channel from the source application to the destination application, bypassing the operating systems on both servers. An InfiniBand channel architecture eliminates the need for the operating system intervention in network and storage communication. This provides a very high-speed, low-latency interface for more efficient and robust movement of data between Oracle RAC nodes. The use of Oracle's InfiniBand networking also preserves server resources for other database processing.

Quad data rate (QDR) InfiniBand technology delivers 40 Gb/sec connectivity, providing application-to-application latency as low as 1 microsecond. It has become a dominant interconnect fabric for high-performance enterprise clusters. InfiniBand provides ultra-low latency and near-zero CPU utilization for remote data transfers, making it ideal for high-performance clustered applications.

In addition to providing unrivaled access to remote application data, InfiniBand's industry-leading bandwidth enables fabric convergence, allowing all network, storage, and inter-process communication traffic to be carried over a single fabric. Converged fabrics aggregate the functions of dedicated, sole-purposed networks and alleviate the associated expense of building and operating multiple networks and their associated network interfaces. In addition, InfiniBand, along with Oracle Solaris, supports IPoIB (Internet Protocol over InfiniBand) for application and Oracle RAC database access, providing higher bandwidth and lower latency of access than can be achieved through Ethernet networking today. The Oracle ZFS Storage Appliance can also support InfiniBand to take advantage of its lower latency and higher bandwidth I/O for database services.

Sun Datacenter InfiniBand Switch 36

The Sun Datacenter InfiniBand Switch 36 from Oracle offers low-latency, quad data rate (QDR), 40 Gb/sec fabric and cable aggregation for Oracle servers and storage. It supports a fully non-blocking architecture, and it acts as a self-contained fabric solution for InfiniBand clusters up to 36 nodes. The Sun Datacenter InfiniBand Switch 36 provides hardware support for adaptive routing, including InfiniBand 1.2 congestion control, which helps to eliminate fabric hotspots and to drive maximum throughput at the lowest-possible latencies. It is ideal for deployment with clustered databases and converged datacenter fabrics.

Advanced features support the creation of logically isolated subclusters, as well as traffic isolation and quality of service management. The embedded InfiniBand fabric management module is enabled to support active/hot-standby dual-manager configurations, ensuring a seamless migration of the fabric management service in the event of a management module failure. The Sun Datacenter InfiniBand Switch 36 is provisioned with redundant power and cooling for high availability in demanding datacenter environments. When deployed in pairs, the InfiniBand fabric provides a highly available and resilient network for business-critical applications.

Conclusion

Across all industries and market segments, database management systems are at the heart of organizational decisions and day-to-day operations. The Oracle Optimized Solution for Oracle Database defines a proven blueprint for efficiently consolidating and optimizing database systems using a proven and high-performance Oracle infrastructure. It combines and validates Oracle products and technologies to lower the total cost of ownership, reduce integration risk, boost employee productivity, and decrease time-to-production for database consolidation.

Oracle-only technologies and optimizations make this Oracle Optimized Solution a superior choice for an enterprise Oracle Database infrastructure. At each tier of the solution stack, Oracle has invested to provide deep integration of its technologies to accelerate Oracle Database performance, increase security and availability, and simplify IT operations.

References

For more information, see the resources listed below.

RESOURCES FOR FURTHER INFORMATION	
WEB RESOURCES	
Oracle Optimized Solution for Oracle Database	http://www.oracle.com/us/solutions/oos/database/overview/index.html
SPARC M6-32 Servers	http://www.oracle.com/us/products/servers-storage/servers/sparc/oracle-sparc/m6-32/overview/index.html
SPARC T4 and T5 Servers	http://www.oracle.com/us/products/servers-storage/servers/sparc-enterprise/t-series/index.html
Oracle Solaris	http://www.oracle.com/solaris
Oracle ZFS Storage Appliances	http://www.oracle.com/us/products/servers-storage/storage/nas/overview/index.html
Oracle Pillar Axiom Storage Systems	http://www.oracle.com/us/products/servers-storage/storage/san/overview/index.html
Oracle Database	http://www.oracle.com/us/products/database/overview/index.html
Oracle RAC	http://www.oracle.com/us/products/database/options/real-application-clusters/overview/index.html
Oracle Optimized Solution for Enterprise Database Cloud	http://www.oracle.com/us/solutions/oos/edb-cloud/overview/index.html
WHITE PAPERS	
How to Save Money, Reduce Risk, and Increase Agility When Deploying Oracle Database Systems	http://www.oracle.com/us/solutions/oracle-opt-sol-for-oracle-db-bw-172332.pdf
Oracle Optimized Solution for Oracle Database Implementation Guide	*Made available as part of the Oracle Optimized Solution sale*
How To Accelerate Test and Development Through Rapid Cloning of Production Databases and Operating Environments	http://www.oracle.com/technetwork/server-storage/hardware-solutions/o13-022-rapid-cloning-db-1919816.pdf
Delivering Quality of Service with Pillar Axiom 600	http://www.oracle.com/us/products/servers-storage/storage/san/improving-shared-storage-wp-488796.pdf
High Performance Security For Oracle Database and Fusion Middleware Applications using SPARC T4	http://www.oracle.com/technetwork/server-storage/sun-sparc-enterprise/documentation/o12-021-t4security-1577047.pdf
Oracle Advanced Security Transparent Data Encryption Best Practices	http://www.oracle.com/technetwork/database/security/twp-transparent-data-encryption-bes-130696.pdf
How to Multiply the Capacity of Oracle Storage Systems With Hybrid Columnar Compression	http://www.oracle.com/technetwork/articles/servers-storage-admin/howto-enable-hcc-pillar-axiom-1914466.html
Oracle Database Cloning Solution Using Oracle Recovery Manager, Oracle's Data Guard, and Oracle's Pillar Axiom 600 Storage System	http://www.oracle.com/technetwork/server-storage/san-storage/documentation/o12-057-cloning-rman-1721637.pdf

WHITE PAPERS (CONTINUED)	
Accelerating Deployment of Enterprise Database Cloud Using an Oracle Optimized Solution	http://www.oracle.com/technetwork/server-storage/engineered-systems/sparc-supercluster/db-cloud-using-sparc-supercluster-1875862.pdf?ssSourceSiteId=ocomen
Benefits of Deploying a Cloud to Deliver Database as a Service	http://www.oracle.com/us/dm/edbc-cloud-sparc-supercluster-1899065.pdf
Oracle Multitenant on SPARC Servers and Oracle Solaris	http://www.oracle.com/technetwork/articles/servers-storage-admin/multitenant-on-sparc-solaris-2016889.html
Managing Many Databases as One: Pluggable Databases (An IDC White Paper)	http://www.oracle.com/us/corporate/analystreports/corporate/managing-dbs-as-one-pdbs-2061401.pdf
Oracle on Oracle – Why the OS, Server, Network, Storage, and Virtualization Matters	http://www.oracle.com/technetwork/articles/servers-storage-admin/sol-why-os-matters-1961737.html

Appendix A
Oracle's Multicore, Multithreaded Processors

At the heart of the Oracle Optimized Solution for Oracle Database are SPARC processor-based platforms that deliver the performance and scalability needed to support consolidated database workloads. In 2013, Oracle announced innovative new processors—the SPARC T5 and M6 processors—that leverage the rich history and previous designs of SPARC CPUs, at the same time advancing thread scale and power management technologies that enable dense and efficient servers.

Evolution of SPARC CMT Processor Designs

For many years, processor performance was highly dependent on clock speeds. In last decade, processor speeds doubled about every two years, following the principle of Moore's law[2], while memory speeds typically doubled only every six years. As a result, memory latency began to dominate application and database performance, often erasing gains in processor clock rates. This relative gap between processor and memory speeds often left ultrafast processors idle, waiting for memory transactions to be completed. Idle processors continued to draw power and generate heat, increasing demands for datacenter power and cooling.

The approach of grouping two or more conventional processor cores on a single physical die—creating multicore processors or chip multiprocessors—yielded only slight improvements in performance since it replicated cores from existing (single-threaded) processor designs. However, the aggregate chip performance tended to increase since multiple programs (or multiple threads) could be accommodated in parallel.

Chip Multithreading

In 2005, Sun Microsystems introduced the UltraSPARC T1 processor, advancing processor designs in a critical manner—*scaling with threads rather than frequency*. The UltraSPARC T1 was available with four, six or eight CPU cores, each core able to handle four concurrently running threads. Unlike traditional single-threaded and multicore processors, hardware multithreaded processor cores allowed rapid switching between active threads as other threads stalled for memory accesses.

Figure 18 illustrates the difference between chip multiprocessing (CMP), fine-grained hardware multithreading (FG-MT), and multicore with multithreading in Chip Multithreading (CMT). The key to CMT is that each core is designed to switch between multiple threads on each clock cycle. As a result, the processor's execution pipeline remains active doing real useful work, even as memory operations for stalled threads continue in parallel. Since individual processor cores implement much-simpler pipelines that focus on scaling with threads rather than frequency, these processors are substantially cooler and require less power.

[2] Moore's Law states that the number of transistors that will fit in a square inch of integrated circuitry will double approximately every two years.

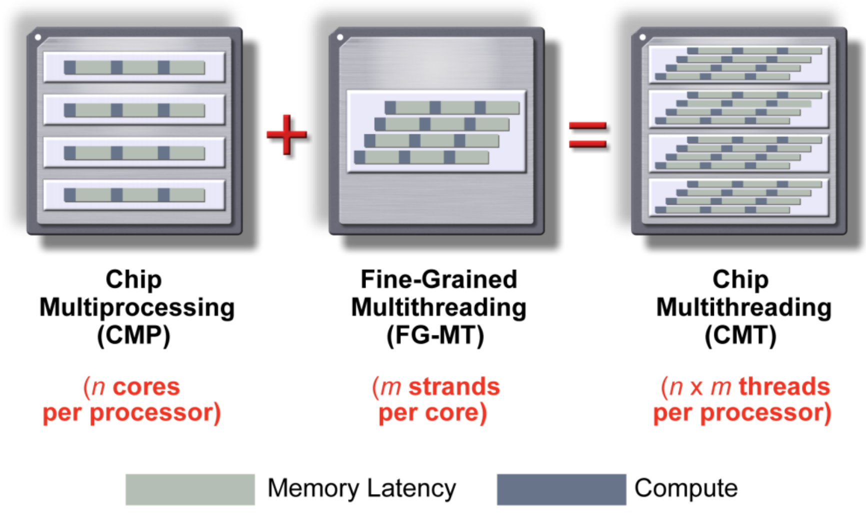

Figure 18. Oracle's multicore/multithreaded CMT approach combines CMP and hardware multithreading (FG-MT).

The successful CMT architecture of the UltraSPARC T1 influenced subsequent generations of SPARC processor designs, including those of Oracle's newest SPARC processors, the SPARC T4, T5, and M6 processors. Of course, each new silicon design takes advantage of breakthroughs in chip manufacturing, shrinking die sizes and enabling the expansion of features and the number of cores and concurrent threads.

Following the UltraSPARC T1 processor, the UltraSPARC T2 processor in 2007 introduced the integration of system level features such as high-speed 10 GbE networking directly on the chip. Even in today's SPARC processors, this System-on-a-Chip (SoC) functionality allows applications to perform more efficiently and improves reliability since it lowers the overall part count. Because encryption is increasingly important in providing data protection, the UltraSPARC T2 processor also added an embedded cryptographic unit in each core—another trend that continues in the today's SPARC T4, T5, and M6 processors.

Subsequent SPARC designs from Oracle—the SPARC T3 and now the SPARC T4, T5, and M6 CPUs—have expanded or enhanced these capabilities. They've added new features such as larger cache sizes, deeper pipelines, sophisticated branch prediction, and out-of-order execution. Over each generation, the goal has been to optimize throughput while improving per-thread performance for both single- and multithreaded applications. The next few pages highlight features in the current generation of SPARC T4, T5, and M6 processors and how they contribute to the value of the Oracle Optimized Solution for Oracle Database.

SPARC T4 and T5 Processors

Oracle's SPARC T4 and T5 processors deliver high throughput performance that has historically distinguished past generations of SPARC T-Series processors. Both processors take advantage of an innovative core design called the S3, first introduced in the SPARC T4.

The SPARC T4 processor includes eight S3 cores, with each core supporting up to eight threads for a total of 64 threads per processor. In addition, each S3 core provides two integer execution pipelines, so a single SPARC core is capable of executing two threads at a time. The SPARC T5 processor essentially doubles the number of cores and threads, supplying sixteen S3 cores with eight threads per core, for a total of 128 threads per processor.

The S3 core supports a 16-stage integer pipeline and one floating-point unit (FGU). The core architecture includes aspects that are conventionally associated with superscalar designs, such as out-of-order (OOO) instruction execution, sophisticated branch prediction, prefetching of instructions and data, deeper pipelines, three levels of cache, support for a much larger page size (2 GB), and multiple instruction dispatching. All these characteristics help to improve thread processing, networking, and throughput performance.

The S3 pipeline is engineered to automatically switch to single-thread mode when only a single thread is active, so all of the core's resources can then be dedicated to that thread's execution. With faster single-threaded processing, the SPARC T4 and T5 processors support fast application boot times and rapid batch processing while maintaining impressive throughput.

SPARC T4 and T5 processors feature a three-level cache architecture. Levels 1 and 2 are specific to each core and are not shared with other cores. Level 3 is shared across all cores of a given processor. Separate Level 1 data and instruction caches are each 16 KB per core. A single Level 2 cache, again per core, is 128 KB. A 4MB Level 3 cache is shared across all eight cores of the SPARC T4 processor, and a 8MB Level 3 cache is shared across all sixteen cores of the SPARC T5 processor.

Both the SPARC T4 and T5 processors are System-on-a Chip designs:

- **Integral PCI Express Support.** SPARC T4 processors provide dual on-chip PCIe Generation 2 interfaces. Each operates at 5 Gb/sec per x1 lane bi-directionally through a point-to-point dual-simplex chip interconnect, meaning that each x1 lane consists of two unidirectional bit-wide connections. SPARC T5 processors integrate dual x8-lane PCIe Generation 3 interfaces at 8GB/sec on-chip, increasing processor I/O bandwidth in comparison to the SPARC T4.
- **Integrated Networking.** SPARC T4 processors provide integrated on-chip networking. All network data is supplied directly from and to main memory. Placing networking close to memory reduces latency, provides higher memory bandwidth, and eliminates inherent inefficiencies of I/O protocol translation. Multiple DMA engines (16 transmit and 16 receive DMA channels) match DMAs to individual threads, providing binding flexibility between ports and threads. SPARC T5 processors use integrated PCIe Gen 3 support to support networking.
- **Stream Processing Unit.** The Stream Processing Unit (SPU) is implemented within the S3 core as part of each pipeline, and operates at the same clock speed as the core. The SPU is designed to achieve wire-speed encryption and decryption. SPARC T4, T5, and M5 processors support widely used cryptographic algorithms, accelerating them in hardware. When Oracle Transparent Data Encryption (TDE) is used to encrypt sensitive data in Oracle Database, it optimizes performance using this hardware-based cryptographic acceleration.
- **Power Management.** Beyond the inherent efficiencies of Oracle's SPARC T-Series multicore/multithreaded design, SPARC T4 and T5 processors incorporate unique power

management features at both processor core and memory levels. These features include reduced instruction rates, parking of idle threads and cores, and the ability to turn off clocks in both cores and memory to reduce power consumption. The SPARC T5 also adds Dynamic Voltage & Frequency Scaling (DVFS) capabilities. DVFS allows customers to set a power limit and then control power consumption within set policies. Processor power management features help to maintain power consumption relative to load.

SPARC T5 Doubles Cores and Threads, Scales to 8 Sockets

The SPARC T5 processor, announced in 2013, leverages the high throughput design of the SPARC T4 CPU. Both chip designs share the same S3 core architecture, but the SPARC T5 doubles the number of cores and available threads per processor. The on-die memory controllers are also similar, but there are now four controllers in the SPARC T5, doubling the available memory bandwidth. The SPARC T5 processor is also clocked at a higher frequency than the SPARC T4, running at 3.6 GHz.

To maintain cache coherency among multiple processors (especially to efficiently support the 8-way SPARC T5-8 server), the SPARC T5 processor moves from a snoopy-based coherence protocol to a glueless 1-hop directory-based protocol. (Snoopy protocols tend to not scale well since all caches share a single bus on which they “snoop” for the cache state. In contrast, with a directory-based protocol, a directory stored in SRAM tracks the state of all Level 3 caches. This approach helps to reduce unnecessary coherence chatter, lower memory latency, and dramatically increase bandwidth between processor nodes.) The result is that the SPARC T5 offers the ability to scale in a near linear fashion from one to eight sockets, making it an ideal CPU for supporting consolidation and virtualization of database architectures. The power management of the coherence link circuitry helps to conserve power use while supporting processor scalability.

SPARC M6 Processors

In 2013 Oracle also announced the new SPARC M6 processor as the cornerstone of its new high-end SPARC M6-32 servers. The SPARC M6 processor is designed to boost throughput, increase compute power, and accelerate memory accesses over previous generations of mainframe-class mission-critical SPARC Enterprise systems.

With an emphasis on increasing throughput by using a multicore/multithreaded design, SPARC M6 processors leverage the innovative S3 core technology spearheaded in the SPARC T4 and T5 chip designs. Clocked at 3.6GHz, SPARC M6 processors feature twelve S3 cores with 8 threads per core, for a total of 96 threads. They contain a large Level 3 cache of 48MB shared across all twelve cores. Compared to current SPARC M3 (SPARC64 VII+) designs, SPARC M6 represents a 100% increase in cores, twelve times as many threads, four times as much cache, and a 20% increase in frequency. As a result of these changes, today’s SPARC M6 chips are capable of significantly higher throughput and single-thread performance in comparison to the previous SPARC M3 generation.

Integrating the S3 core in the SPARC M6 brings desirable on-die features to this high-end processor, including:

- Dual x8-lane PCIe Generation 3 interfaces at 8GB/sec, enabling 16GB/sec each direction

- An integrated ISA-based encryption engine in each core, for wire-speed encryption and decryption. Again, Oracle Transparent Data Encryption (TDE) operations can tap this native hardware encryption capability when encrypting data for Oracle Databases. (For further information on TDE with Oracle Databases, please reference "High Performance Security For Oracle Database and Fusion Middleware Applications using SPARC T4".)

SPARC M6 processors use a directory-based cache coherence model similar to that of the SPARC T5, with up to 8 sockets supported in a glueless 1-hop fashion. To expand support to 32 sockets, the SPARC M6 architecture includes additional scalability links. Each link supports a bandwidth of 12 GB/sec per lane with 6 lanes per link. Because of the ramp in processor cores and threads per core in the SPARC M6 design, the processor enables a six-fold increase in thread-scale (from 512 to 3072 available threads) in the SPARC M6-32 server in comparison to the previous generation 64-socket SPARC Enterprise M9000 server.

Summary of SPARC T4, T5, and M6 Processor Features

In closing, Table 5 summarizes the features of the SPARC T4, T5, and M6 processors.

TABLE 5. COMPARISON OF SPARC T4, T5, AND M5 PROCESSORS

	SPARC T4 PROCESSORS	SPARC T5 PROCESSORS	SPARC M6 PROCESSORS
Processor	8 S3 cores, 8 threads @ 2.85GHz and 3.0GHz	16 S3 cores, 128 threads @3.6GHz	12 S3 cores, 48 threads @3.6GHz
S3 Core Features	1-8 Thread, Dynamically threaded pipeline, ISA-based Crypto-acceleration		
Shared L3 Cache	4MB	8MB	48MB
Integrated I/O	2x8 Lane PCI 2.0 @ 8GT/s	2x8 Lane PCI 3.0 @ 8GT/s	
Process technology	40 nanometer	28 nanometer	28 nanometer
System Scalability	Up to 4 sockets	7 Coherence Ports for scalability to 8 sockets, 1024 threads	Coherence Ports and additional Scalability Links for scalability to 32 sockets, 3072 threads
Power Management	Reduced instruction rates, idle thread and core parking, CPU downclock/overclock	Adds Dynamic Voltage Frequency Scaling (DVFS)	

ORACLE

How to Consolidate, Optimize, and Deploy
Oracle Database Envrionments with Flexible
Virtualized Services
March 2013, Version 5.0

Authors: Roger Bitar, Ritu Kamboj, Ken Kutzer, and Larry McIntosh

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

Hardware and Software, Engineered to Work Together