ORACLE

SOLARIS

An Oracle White Paper
September 2010

Oracle Solaris: In a Class by Itself

ORACLE

Table of Multimedia Content

Executive Summary

Oracle® Solaris is the number one enterprise UNIX®, and the top deployment operating system for Oracle Database for several years running. The reasons for this are clear: Oracle Solaris delivers a reliable, secure, and high-performance platform for running Oracle Database, Oracle Applications, and other enterprise applications—a single operating system for the complete datacenter—that is in a class by itself. This paper offers a high-level discussion of the features that separate Oracle Solaris from other enterprise UNIX operating systems, with specific details on optimizations and advantages for increased reliability, scalability, security, and virtualization. Resources that can provide more information are available at the end of the paper.

There are several multimedia components included within this document, which provide topical additional information. These components can be played from a current version of Adobe Reader, without a web connection. You can find these components in the following locations:

John Fowler, Oracle Executive VP, on Oracle and Sun integration

Qualcomm on ZFS

Edward Screven, Oracle Chief Corporate Architect, on innovative virtualization

Charlie Boyle, Oracle Sr. Director of Product Marketing, on SPARC virtualization

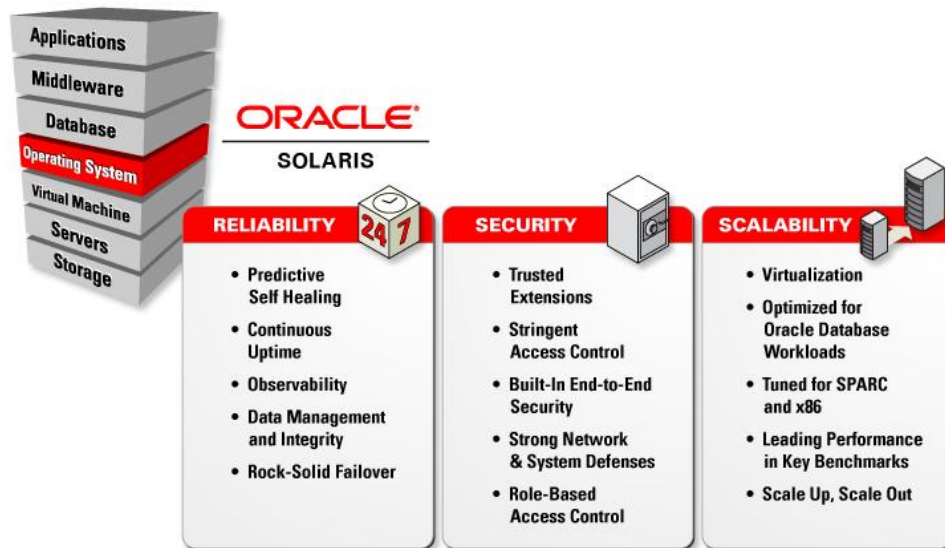# 1. Powering the Datacenter—The Oracle® Solaris Ecosystem

Figure 1: Oracle Solaris delivers reliability, security, and scalability, and application-to-disk integration.

Oracle's integrated portfolio of operating system, virtualization, cluster, and development technologies includes Oracle Solaris, Oracle VM, Oracle Solaris Cluster, and the Oracle Solaris Studio software development tools, which form the core of a large developer ecosystem. As shown in Figure 1, Oracle Solaris is the centerpiece on which Oracle delivers integrated hardware and software solutions that are reliable, scalable, and secure.

A compelling benefit for Oracle Solaris users is the application-to-disk integration, optimizing performance, reliability, and security for your complete solution stack. Oracle and Sun worked closely for years before the acquisition to improve the customer experience and increase return on investment. This helps explain why Oracle Solaris is the dominant deployment platform for Oracle Database. Over 40 percent of Oracle's customers are using Oracle Solaris, and over 11,000 Oracle Solaris applications—more than any other enterprise operating system. While there are many reasons for this broad market acceptance, a reliable platform is a big factor in the equation. Now, the next level of integration efforts is already underway, starting with improved testing of the integrated stack.

Oracle Solaris uses an extensive patching mechanism to help ensure the highest quality and reliability, performance, and security. New patches are released every week, totaling approximately 2,000 per year. Each patch undergoes an automated process on both SPARC® and x86 platforms. The testing is rigorous, with over 50,000 test cases run against every patch. The testing includes patch alignment, to prevent instances of one patch breaking another—regression testing. This testing has now been expanded to include:

- **Oracle Certification Environment (OCE).** OCE is a comprehensive suite of tests that validates stack reliability and conformance.

- **Oracle Automated Stress Testing (OAST)**. An automated TPC-C like test suite designed to build OLTP type workloads for systems using the Oracle Database. OAST creates tables, performs stress test runs, and outputs transaction-related performance data.

- **Oracle VM suite.** A comprehensive test environment that tests reliability, performance, and integration for virtualized solution stacks.

Figure 2: Click the image to watch John Fowler discuss Oracle and Sun integration. **See the entire Executive Webcast.**

Oracle is now stack testing with pre-release versions of Oracle Applications and Fusion Middleware with Oracle Solaris, to reduce the number of patches required and further enhance quality and reliability while tuning for maximum performance and scalability. These perform load and stress tests, inject faults, and test interoperability and scalability. This end-to-end testing helps ensure maximum performance on a fully integrated system. For example, internal testing shows that Oracle Solaris ZFS is delivering greater Oracle Database throughput and lowering response time on internal OLTP benchmarks, compared with previous versions. (Read more about Oracle ZFS in Chapter 2 – Reliability.)

In addition, Oracle Global IT is moving quickly to run their business on Oracle Sun products. This includes datacenter consolidation that will use Oracle Solaris, Sun SPARC Enterprise servers, and Sun storage. Coupled with Oracle Secure Backup, Oracle Solaris is also now an integral part of Oracle's IT backup environment.

Oracle Solaris is a proven, industry-leading operating system with features designed to handle enterprise, business-critical operations. In fact, Oracle Solaris provides key functionality for virtualization, optimal utilization, high availability, unparalleled security, and extreme performance for both vertically and horizontally scaled environments. Oracle Solaris 10 runs on a broad range of SPARC- and x86-based systems and compatibility with existing applications is guaranteed. This is why there are over 50,000 businesses and institutions running on Oracle Solaris today.

Oracle Solaris powers Oracle's Sun x86 and SPARC Enterprise servers to new world records for performance, scalability, and cost-effectiveness. Oracle has increased investment in Oracle Solaris, and will continue to develop innovative technologies and enhance Oracle Solaris for Oracle's optimized enterprise architecture. As shown in Table 1 below, the business value of Oracle Solaris place it in a class by itself.

TABLE 1: ORACLE SOLARIS BUSINESS BENEFITS

FEATURE	VALUE
Optimized for Oracle's Enterprise Architecture	• Oracle Solaris is the Number One deployment OS for Oracle Database • Unique Oracle Solaris innovation plus Oracle Database 11*g* Release 2 features deliver best database performance • Oracle Database, Oracle Application, Oracle Sun systems and storage integration and support • Whole-stack approach optimizes performance, minimizes issues, and delivers innovation faster than ever
Unmatched Reliability	• Oracle Solaris Predictive self healing framework and Oracle Solaris Cluster ensure resiliency and availability for the solution stack, not just server hardware • Live Upgrade maintains uptime during patches or upgrades • Oracle Solaris ZFS protects data from corruption and simplifies data management • RAC integration and Oracle Solaris Cluster delivers highest level of availability and data integrity
Mission-Critical Security	• Secure By Default, including cryptographic framework, IPsec, firewall, and more • Mandatory access controls using Trusted Extensions and Common Criteria • User and process management grants minimum privileges, and no more • Fingerprint database validates system and data file integrity
Optimized Performance and Scalability	• Highly-optimized hardware support provides 20 years of near linear speedup • Oracle Solaris DTrace feature unlocks performance bottlenecks • Integrated Oracle Solaris scalability and security and features drive Oracle WebLogic performance at a great price point • A single OS runs across SPARC and x86 servers—start with cost-effective x86 and scale up to hundreds of threads as needed with SPARC
Virtualization	• Oracle Solaris Containers delivers secure, low-overhead, virtualization, with over 8,000 VMs demonstrated on a single server • Low-cost virtualization on every system • Comprehensive solution—network, storage, cluster—plus innovative Oracle Solaris features such as DTrace, Predictive Self Healing, and more

Oracle Solaris includes many unique features based on innovative technologies that are not found in other operating system—including: Oracle Solaris ZFS, Oracle Solaris DTrace, Predictive Self Healing, built-in virtualization, independent security verification, binary compatibility, and the Oracle Solaris Cluster high availability and disaster recovery solutions. Oracle protects your IT investments by guaranteeing that existing Oracle Solaris 8 and 9 applications will run unmodified on Oracle Solaris 10. As enterprise system hardware often has a service life of 8-10 years or more, it is reassuring to understand Oracle's commitment to providing long-lived support for your software environment.

2. Reliability

Oracle Solaris offers reliability features not found in other enterprise operating systems. For example, Predictive Self Healing and Service Management Facility help ensure that your solution stack, from application to disk, is available. Other operating systems offer just server-centric reliability features.

Oracle Solaris is designed for reliability. With Oracle's *whole-stack* testing, the comprehensive solution stack—from application to disk—is tested for increased reliability and availability. Oracle Solaris is tightly integrated with hardware reliability features offered on Sun SPARC Enterprise and Sun x86 systems. Examples of this are memory mirroring, and offlining problematic memory pages or processing cores.

Built with a small, compact kernel, Oracle Solaris limits the potential for operating system faults and subsequent platform downtime. In addition, Oracle Solaris establishes a clear distinction between the kernel, shared libraries, and applications in order to limit the impact of application failures. Furthermore, the ability to install most patches and other incremental software updates for Oracle Solaris without taking the system offline helps organizations increase uptime and enables ease of serviceability.

Reliability also benefits from strong security. Protecting systems and data from unauthorized or inadvertent use keeps them from going offline. Security capabilities are discussed in detail in Chapter 4.

Predictive Self Healing

Predictive Self Healing in Oracle Solaris is an innovative capability that defines a philosophy and set of common technology components that support self-healing features within Oracle products. Oracle Solaris Predictive Self Healing software proactively monitors and manages system components to help organizations achieve maximum availability of IT services. Predictive Self Healing is an innovative capability in Oracle Solaris that automatically diagnoses, isolates, and recovers from many hardware and application faults. This enables business-critical applications and essential system services to continue uninterrupted in the event of software failures, major hardware component failures, and even misconfigured software. The Oracle Solaris Fault Manager and Oracle Solaris Service Manager Facility (SMF) are the two main components of Predictive Self Healing.

Oracle Solaris Fault Manager receives data (*telemetry*) relating to hardware and software errors on Sun systems, and automatically diagnoses the underlying problem. Oracle Solaris Fault Manager, a common system that works across platforms running Oracle Solaris, reduces complexity by automatically diagnosing faults in the system and initiating self-healing actions to help prevent service interruptions. This software helps increase availability by configuring problem components out of a system before a failure occurs. In the event of a failure, this feature initiates automatic recovery and application re-start (using SMF). The Oracle Solaris Fault Manager diagnosis engine produces a fault diagnosis once discernible patterns are observed from a stream of incoming errors. Following diagnosis, the Oracle Solaris Fault Manager provides fault information to agents that know how to respond to specific faults. For example, an agent designed to respond to a memory error might determine the memory addresses

affected by a specific chip failure and remove the affected locations from the available memory pool. Oracle Solaris Service Manager is integrated with SMF, acting as a framework that provides several functions, including:

- Definition of services for Oracle Solaris, which can be the state of a device, a running application, or a set of other services. Each service is referred to by a unique identifier.

- Creates a formal relationship between services, with explicit dependencies

- Automatic starting and restarting of services

- A repository for storing service state and configuration properties, eliminating the need for dozens of configuration files scattered throughout the system

When a low-level fault is found to impact a higher-level component of a running service, Oracle Solaris Fault Manager can direct SMF to take appropriate action. Failing services are automatically restarted whenever possible, reducing the need for human intervention. Should manual intervention be required, system administrators can quickly identify the root cause of the service's failure and significantly reduce the times-to-repair and recover from said failure.

Specifically, SMF enables administrators to do the following tasks easily and efficiently with Sun SPARC Enterprise systems:

- Observe and manage system-wide services

- Identify "misbehaved" or failed services

- Securely delegate administrative tasks to non-root users

- Automatically restart failed services in the appropriate order of dependency

- Persist the enable/disable of services across system upgrades and patches

- Preserve compatibility with legacy services

- Automatically configure snapshots for backup, restore, undo

- Provide consistent configuration handling

Predictive Self Healing capabilities offer extensive reliability and availability capabilities on all Sun systems. For example:

- Processor offlining capability is a self healing technology that dynamically removes a faulty processor chip, core, or thread from use. Services associated with the faulty processor are either moved to other processors or terminated and restarted if possible, to avoid an outage for the entire system. If needed, a reboot deconfigures the faulty processor. Internal reliability, availability, and serviceability (RAS) modeling analysis shows that with the processor offlining feature deployed, the system interruptions as well as system downtime caused by hardware faults can be reduced by over 30 percent.

- Memory page retirement (MPR) minimizes the effect of faulty memory. MPR technology allows memory pages suffering from correctable errors and relocatable clean pages suffering from

uncorrectable errors to be removed from use without interrupting user applications. It also allows relocatable dirty pages suffering from uncorrectable errors to be isolated with limited impact on affected user processes, avoiding an outage for the entire system. If the number of retired pages associated with a (partially) faulty DIMM have not exceeded a predefined threshold, the DIMM does not have to be replaced, avoiding a service action. Oracle field data indicates that MPR has helped reduce DIMM dispatch rates by 35-40 percent.

Oracle Solaris ZFS

Figure 3: Click the image to watch Qualcomm describe how Oracle Solaris ZFS improves their IT architecture. See the entire presentation.

A critical element to enterprise-class reliability is the data subsystems. Oracle Solaris ZFS provides unparalleled data integrity, capacity, performance, and manageability for storage. ZFS provides high-resiliency features, such as metadata logging to guarantee data integrity and speed recovery in the event of system failure. ZFS dramatically simplifies file system administration to help increase protection against administrative error. Oracle Solaris ZFS offers a dramatic advance in data management with an innovative approach to data integrity, greatly simplified administration, and a welcome integration of file system and volume management capabilities. For example, LUNs can be dynamically expanded— as more storage is added, ZFS will adapt automatically.

Oracle Solaris ZFS uses techniques such as copy on write and end-to-end *checksumming* to keep data consistent and eliminate silent data corruption. Data is written to a new block on the media before changing the pointers to the data and committing the write. Because the file system is always consistent, time-consuming recovery procedures like `fsck` are not required if the system is shut down in an unclean manner. In addition, data is read and checked constantly to help ensure correctness, and any errors detected in a mirrored pool are automatically repaired to protect against costly and time-consuming data loss and previously undetectable silent data corruption. Corrections are made possible by a RAID-Z implementation that uses parity, striping, and atomic operations to aid the reconstruction of corrupted data. There is support for triple-parity RAID, which means that three physical drives in a pool can fail without data loss.

Oracle Solaris ZFS protects all data by 256-bit checksums, resulting in 99.99999999999999-percent error detection and correction. Oracle Solaris ZFS constantly reads and checks data to help ensure it is correct, and if it detects an error in a storage pool with redundancy (protected with mirroring, Oracle Solaris ZFS RAIDZ, or Oracle Solaris ZFS RAIDZ2), Oracle Solaris ZFS automatically repairs the corrupt data. This contributes to relentless availability by helping to protect against costly and time-

consuming data loss due to hardware or software failure, and by reducing the chance of administrator error when performing file system-related tasks.

ZFS software also provides the data services needed to protect data far beyond what exists today in traditional storage systems. It optimizes file system reliability by maintaining data redundancy on commodity hardware through the delivery of basic mirroring, compression, and integrated volume management.

An interesting feature of Oracle Solaris ZFS is the snapshot capability—a read-only copy of an Oracle Solaris ZFS file system or volume. Snapshots can be created almost instantly and initially consume no additional disk space within the pool. Snapshots are a valuable tool for system administrators needing to perform backups and other users needing to save the state of a file system at a particular point in time and possibly recreate it later on the same or another machine. It is also possible to extract individual files from a snapshot.

Oracle Solaris ZFS seamlessly and transparently supports new hybrid disk storage pools that include Flash technology for accelerated application performance (discussed in the Scalable Performance section).

DTrace

Oracle Solaris DTrace is a comprehensive dynamic tracing framework for Oracle Solaris 10 designed for real time application debugging and performance troubleshooting. Oracle Solaris DTrace provides a powerful infrastructure to permit administrators, developers, and service personnel to concisely answer arbitrary questions about the behavior of the operating system and user programs. It is a powerful tool that can be used by both entry-level and experienced system administrators to troubleshoot both system and application performance problems in hours or minutes that might have previously taken days. Oracle Solaris DTrace is safe to use on development, test, and production systems.

DTrace probes are already installed through the Oracle Solaris environment. Administrators and application developers can instrument a live Oracle Solaris kernel and running applications without rebooting, recompiling, or even restarting the application of interest. DTrace instrumentation can be activated as needed and generates no application overhead when tracing is turned off.

With DTrace, system administrators, integrators, and developers can really see what the system is doing, as well as how the kernel and applications interact with each other. It enables users to formulate arbitrary questions and get concise answers, allowing them to find performance bottlenecks on development, pilot, and production systems.

DTrace provides accurate and concise information in real time. Questions are answered immediately, eliminating the need to collect large amounts of data for later analysis. DTrace can also highlight patterns and trends. This makes it easier and faster to identify bottlenecks, a task that can be difficult and time consuming with many other tools. By employing speculative tracing, DTrace can record, report, or discard trace data based on non-simultaneous events, enabling quick identification of transient problems and reducing the need for post-processing. In addition, DTrace can be used across

Oracle Solaris Containers to quickly correlate events and find bottlenecks on distributed applications such as Web environments. All of this leads to increased performance and service availability, while system downtime is reduced.

DTrace scripts are written in a language similar to C, simplifying the learning process for most system administrators and application developers. Within DTrace scripts, probes record data and set timestamps at locations of interest throughout application execution. Oracle Solaris includes over 30,000 built-in probes that are points of instrumentation within the kernel.

DTrace support has been added to the Java™ SE 6 HotSpot VM. The `hotspot` and `hotspot_jni` providers make available probes that can be used to monitor JVM internal state and activities as well as the Java application that is running. All of the probes are USDT probes and are accessed using the `process-id` of the JVM process.

Oracle Solaris Live Upgrade

Oracle Solaris Live Upgrade provides a method of upgrading an OS while the system continues to operate. The original system configuration remains fully functional and unaffected by the upgrade or installation of an archive. When ready, you can activate the new boot environment by rebooting the system. There are several benefits associated with: enables you to duplicate a boot environment without affecting the currently running system.

- **Minimized downtime.** Servers remain available during most of the upgrade or patch process, going offline only during the reboot.

- **Maintain numerous boot environments with different images.** For example, you can create one boot environment that contains current patches and create another boot environment that contains an update release. Or, you change the current boot environment's disk configuration to different file system types, sizes, and layouts on the new boot environment.

- **Quick rollbacks.** If a failure occurs, you can quickly revert to the original boot environment with a simple reboot.

Oracle Solaris Live Upgrade enables Oracle Solaris 10 and previous versions to continue to run while an administrator upgrades to the latest release of the OS, applies patches, or does routine maintenance on the inactive or duplicate boot environment. When satisfied with the process, you simply reboot the system to run the latest or updated operating environment.

You can also install a Flash archive on a boot environment. Oracle Solaris Flash archives can be clones of a master system, which can be reconfigured or customized for each system. To expedite the installation of a Flash archive, you can also create an update that only includes the differences between two system images.

The original system configuration remains fully functional and unaffected by the upgrade or installation of a Flash archive. The duplicate boot environment is then activated to become the active boot environment when the system is rebooted. If the new boot environment does work as expected, you

have a safety net. You can quickly fall back to the original boot environment with a simple reboot, thereby eliminating the service outages that are associated with the normal test and evaluation process.

Oracle Solaris Live Upgrade integrates with Oracle Solaris Volume Manager and other volume managers. Oracle Solaris Live Upgrade can be used to add or remove packages as well as patches.

Reliable Networking and I/O

Reliable, and high-performance connectivity is an essential aspect of an enterprise IT infrastructure. Oracle Solaris supports many innovative features that detect and repair network-related failures, even in virtualized environments. Integrated security technology contributes to data integrity and overall uptime.

Redundant Networking and Network IP Multipathing

Keeping applications up and running means keeping network services and I/O connections up and running in a variety of situations, while maximizing the use of all resources at all times. Oracle Solaris I/O multipathing provides the ability to set up multiple redundant paths to a storage system and networks, while simultaneously offering the benefits of load balancing and failover.

In addition to traditional support for multiple network interfaces connected to different network subnets, Oracle Solaris also provides support for redundant network interfaces that are connected to a single subnet. IP Multipathing (IPMP) provides both failover and IP link aggregation. A number of key features of redundant networking that work to improve the availability and performance of Sun systems are listed below.

- **Failure detection.** The ability to detect when a network adapter fails and automatically switch (failover) network access to an alternate network adapter.

- **Repair detection.** The ability to detect the repair of a previously failed network adapter and automatically switch back (fail back) the network access to this interface.

- **Outbound load spreading.** Outbound network packets are spread across multiple network adapters to achieve higher throughput. Load spreading occurs only when network traffic is flowing to multiple destinations using multiple connections.

Multiplexed I/O

Multipathing provides redundant paths and eliminates a single point of failure by failing over to an alternate path automatically when one of the paths fails. Oracle Solaris Multipath I/O (MPxIO) enables storage subsystems to be accessed through multiple host controller interfaces from a single OS instance. MPxIO helps protect against I/O outages due to I/O controller failures—should one I/O controller fail, MPxIO automatically switches to an alternate controller. MPxIO offers path management, failover support, load balancing, dynamic reconfiguration, and other capabilities.

Oracle Solaris Cluster

Keeping application data and services in a single system exposes businesses to potential failure from any component of the configuration. Enterprise IT services need to run in clustered physical servers that can efficiently and smoothly take over the services from failing nodes, with minimal interruption to the customer experience. Oracle Solaris Cluster extends the inherent availability capabilities in Oracle Solaris. Tightly coupled with Oracle Solaris, Oracle Solaris Cluster detects failures with minimal delay, provides much faster failure notification, application failover, and reconfiguration time. Significantly reducing services recovery time achieves much faster resumption of IT services. Oracle Solaris Cluster:

- Integrates tightly with the Predictive Self Healing framework and supports SMF-controlled applications in Oracle Solaris Containers

- Makes extensive use of Oracle storage management and volume management capabilities

- Supports Oracle Solaris ZFS as a failover file system and as a boot file system, allowing the use of ZFS storage as the single file system type used

- Leverages Oracle Solaris ZFS features such as pooled storage, built-in redundancy, and data integrity

- Leverages multipathing on public network interfaces for fast detection

- Provides a logical interface on private cluster network for traffic stripping and transparent handling interconnect failures and reconfigurations

- Integrates with Oracle Enterprise Manager Ops Center

- Offers secure administrative capabilities through Oracle Solaris role-based access control (RBAC), enhancing security

- Provides comprehensive failover capabilities for Oracle Real Application Clusters (RAC) and Data Guard installations

A Single High Availability and Disaster Recovery Solution for Multitier Oracle Applications and Databases

Oracle Solaris Cluster software enables high availability (HA) for local datacenters to business continuity and global disaster recovery solutions for evolving datacenter needs. The software leverages proven availability and virtualization features in Oracle Solaris and Sun servers, and supports an industry-leading portfolio of commercial and open source applications across the database and business logic tiers.

Oracle Databases and Applications, including Oracle E-Business Suite, Siebel CRM 8, and Oracle single instance and RAC Databases, delivers comprehensive support for virtualized applications running in Oracle Solaris Container Clusters and Oracle VM Server for SPARC. Oracle Solaris Cluster supports Oracle 9*i*, 10*g*, and 11*g* RAC (Real Application Clusters), running on top of Oracle Solaris Cluster, which links the server nodes supporting the RAC deployment. The integration of Oracle Clusterware with Oracle Solaris Cluster facilitates faster failure detection, effective and more reliable inter-node communication, and a broader choice of data storage options. Oracle Solaris Cluster

provides pretested, precertified software agents that support most leading, off-the-shelf applications, including WebLogic Server, Oracle E-Business Suite, Oracle Databases, and many others.

For complete information on Oracle Solaris Cluster, see the white paper *Oracle Solaris and Oracle Solaris Cluster: Extending Oracle Solaris for Business Continuity.*

3. Scalable Performance

Oracle Solaris and Sun servers offer maximum performance and scalability, without tradeoffs and compromise. Unlike IBM POWER7 and AIX 7.1 environments, Oracle's integrated framework does not force you to reboot and choose between performance or scalability. Oracle Solaris and Sun servers offer proven, mature scalability, and are available on systems that are in production with 256 cores/512 threads.

Oracle Solaris is designed for optimized, end-to-end performance, reducing or eliminating bottlenecks in memory and I/O subsystems.

Oracle Solaris has long held a unique position in the industry by delivering a single operating system that can help organizations maximize application performance by either scaling up (vertical scaling onto the largest SMP systems) and scaling out (horizontal scaling across multiple small servers). Oracle empowers its customers to use both scalability dimensions to best meet their critical performance and availability criteria. Oracle Solaris supports customers who need to scale up enterprise applications such as database management systems on servers with hundreds of processing cores. At the same time, for applications such as Web servers where scaling out is the preferred way to grow, Oracle Solaris sets world records using volume servers and new processor technologies.

Successful enterprise-class servers efficiently process CPU, memory, and I/O workloads for middleware and databases. Building on a proven track record, Oracle Solaris is ready to take advantage of the unique performance capabilities inherent in both SPARC and x86 processors. Significant performance innovation comes from optimizations of the individual cores and the overall multicore microarchitecture, which increase both single-threaded and multithread performance. As a result, the Oracle Solaris kernel and existing single- or multithreaded applications will run faster, with no code changes or recompilation necessary.

Additional performance improvements can be achieved by taking advantage of the latest Oracle Solaris Studio developer tools, which offers performance libraries and automatic parallelization capabilities. Taking advantage of these capabilities fully maximizes application performance—see more information in Chapter 6 – Developer Tools Optimizations.

Multithread Awareness

Oracle Solaris is optimized for SPARC, Intel, and AMD processor hierarchies so that the scheduler can effectively balance the load across all the available pipelines. Even though it exposes every physical processor strand as a logical processor (up to 64 per physical processor), Oracle Solaris understands the correlation between cores and the threads they support, and provides a fast and efficient thread

implementation. Independent software threads are first spread across processors, then across cores within a processor, then across pipelines within a core.
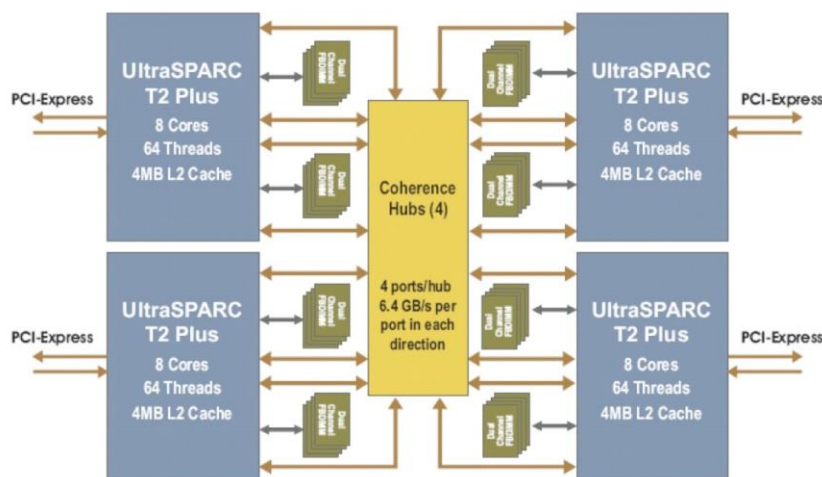
NUMA Optimization—MPO

Figure 4: Oracle Solaris optimizes memory usage for maximum performance in multiprocessor systems using memory placement optimization— MPO.

As systems grow larger, with more processor sockets and more memory, the ability of a processor to access memory becomes more challenging—all processors cannot directly access all memory at the same latency. Multiprocessor systems generally demonstrate some memory locality effects, which means that when a processor requests access to data in memory, that operation will occur somewhat faster if the memory bank is physically close to the requesting processor. As shown in Figure 4, you can see the processor in the lower left corner has direct access to memory in the lower left, while memory in the upper right is one hop away.

Sun SPARC Enterprise and x86 servers are designed with a NUMA architecture, enabling processors to directly access some memory at the lowest latency, while accessing the rest of the memory with more latency. Oracle Solaris provides technology that can specifically help applications improve performance on NUMA architectures.

- **Memory Placement Optimization (MPO).** Oracle Solaris uses Memory Placement Optimization (MPO) to selectively place data across the physical memory of a server, resulting in increased performance. Through MPO, Oracle Solaris works to help ensure that memory is as close as possible to the processors that access it, while still maintaining an appropriate balance within the system. As a result, many database and technical computing applications are able to run considerably faster with MPO.

- **Hierarchical lgroup support (HLS).** HLS improves the MPO feature in Oracle Solaris. 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 Internals Optimization

For over 20 years, Oracle Solaris internals have been improved to enhance scalability, enabling Sun servers to deliver maximum performance as they have grown to address terabytes of memory and hundreds of processing cores. These include:

- **Large Pages.** Large pages are used to reduce the cost of virtual to physical memory translation and increase overall system performance. The SPARC64 VII and UltraSPARC T-Series processors provide a range of page sizes up to 256 MB. Oracle Solaris automatically uses large pages for Oracle Database instruction pages and for the database System Global Area (SGA, the database buffer cache) on all SPARC and Sun x86 systems, and for the database Process Global Area (PGA, a dedicated memory cache) on Sun T-series servers. In addition, Oracle Database allows the end user to tune the selection of `pagesize` for the PGA on all systems through its use of the `memcntl(2)` system call.

- **Mutexes (mutual exclusion operations).** As system size grows, there is a growing likelihood of another thread holding a mutex when a second or third thread attempts to access it. To minimize the performance limitations of heavy mutex contention, Oracle Solaris applies a backoff algorithm that is tuned for the system size and processor characteristics before retrying contended locks. The larger the number of threads or strands, the greater the benefits of the improved mutex backoff algorithms.

- **Intimate Shared Memory (ISM).** The use of ISM allows the processes to share kernel data structures that store virtual to physical translations, reducing the cost of a TLB miss. In addition, UltraSPARC T-Series and SPARC64 VII processors implement a unique feature called the shared MMU context, which is used for ISM segments and allows threads to share translations in the hardware TLB cache, reducing the TLB miss rate. Large pages, ISM, and shared context combine seamlessly to optimize access to large memory and the Oracle Database SGA on Sun SPARC T-series and M-series servers.

- **Library Optimization.** Oracle Solaris provides multiple implementations of common utility functions such as `memcpy(3C)`, each of which is optimized for a different SPARC processor. The versions are kept in shared libraries that are updated as new processors are developed, and the linker dynamically selects the best version at application start time based on the processor that is present. No change to the application is required to get the fastest version for the latest processor.

- **64-bit Mode.** 64-bit capabilities offer extended precision, large dataset support, and a larger virtual address space.

Sun FlashFire Storage

Disk I/O performance is often a bottleneck to application throughput. Sun FlashFire products running on both Sun SPARC Enterprise servers and Sun x86 servers use high-performance Flash technology,

including on-board modules, solid-state drives (SSDs), and PCIe cards to significantly increase application throughput. Flash-based technology offers 10x faster data response times compared to traditional storage devices, and reduces cooling, power, and space requirements. Oracle is setting new benchmark records using Flash technology.[1] In addition, Oracle Solaris ZFS can seamlessly and transparently integrate Flash technology and conventional hard disk drives to create Hybrid Storage Pools, which delivers faster storage performance and lower overall costs.

Network Performance

Oracle Solaris introduces a new and highly scalable TCP/IP stack that significantly increases network throughput and capacity. This innovative stack speeds throughput by reducing overhead when processing packets. The advanced design improves the performance of many networked applications—without requiring you to modify a single line of application code. The resulting efficiency helps to drive down costs through increased scalability, allowing your systems to support more connections and enabling network throughput to grow linearly with the server's number of CPUs and network interface cards (NICs). The Oracle Solaris TCP/IP stack is tuned for 10 Gigabit Ethernet and hardware offloading technologies.

Oracle Solaris meets current and future networking challenges by significantly improving network performance without requiring changes to existing applications. Oracle Solaris speeds application performance via the Network Layer 7 Cache and enhanced TCP/IP and UDP/IP performance. The latest networking technologies, such as 10- Gigabit Ethernet (GbE) and hardware offloading, are all supported out of the box. In addition, Oracle Solaris supports current IPv6 specifications, high availability, streaming, and Voice over IP (VoIP) networking through extended routing and protocol support—meeting the needs of a growing customer base.

4. Security

Oracle Solaris provides flexible and robust security. This includes Common Criteria certification on real-word customer configurations, such a virtualized applications. No-cost, integrated cryptographic capabilities, and the flexibility to choose the latest security that supports your policies and procedures are unavailable from other system vendors.

Oracle Solaris provides a sophisticated network-wide security system that controls the way users access files, protect system databases, and use system resources—all while retaining a detailed audit trail of every system event. From integrated security services and applications, to enhanced encryption algorithms, to an enterprise firewall for network protection, Oracle Solaris sets a high standard for operating system security by addressing security needs at every layer. Extended security features are

---

[1] See the latest Oracle Solaris performance benchmarks on Sun systems.

also included, such as authentication, data integrity, data privacy, and single sign-on capabilities so that tampering, snooping, and eavesdropping do not compromise data or associated transactions. Oracle Solaris security comprehensive capabilities include the following.

Process and User Rights Management

Hackers often attempt to exploit root accounts because those accounts are empowered with complete access to UNIX systems. Oracle Solaris offers unique user and process rights management technology that reduces risks by granting users and applications only the minimum capabilities needed to perform their duties. Unlike other solutions, Oracle Solaris requires no application changes to take advantage of these security enhancements.

Process rights management enables processes to be restricted at the command, user, role, or system level. Oracle Solaris implements process rights management through privileges. Privileges decrease the security risk that is associated with one user or one process having full superuser capabilities on a system. Privileges and RBAC provide a compelling alternative model to the traditional superuser model. RBAC is a feature of Oracle Database, Oracle E-Business Suite, and Oracle Solaris. In the RBAC model in Oracle Solaris, users log in as themselves and assume roles that enable them to run restricted administration graphical tools and commands. RBAC is considered a best practice across all Oracle products.

Process and user rights management work to prevent attacks by disallowing application code to be executed from the application's stack. This type of attack, known as "stack smashing," could allow an otherwise unprivileged application to gain access to memory or processes that it should not have. Preventing this type of attack requires that Oracle Solaris and the processor (including SPARC, Intel, and AMD) work together—this protection is automatic for all 64-bit applications on the OS, and available for all older 32-bit applications with a simple system configuration setting.

Least Privilege

Most UNIX operating systems run a large number of their system processes with root privileges. These processes then have the capability to read and modify other processes, memory, I/O devices, and so on. While this gives these system processes the power needed to perform their tasks, it also provides them with unnecessary access to other protected parts of the system. Many software exploits count on this escalated privilege to gain superuser access to a machine via bugs like buffer overflows and data corruption. To combat this problem, Oracle Solaris includes a new least privilege model, which gives a specified process only a subset of the superuser powers and not full access to all privileges.

The least privilege model evolved from Sun's experiences with Trusted Solaris and the tighter security model used there. The Oracle Solaris least privilege model makes it convenient for normal users to do things like mount file systems, start daemon processes that bind to lower numbered ports, and change the ownership of files. At the same time, it also restricts access by programs that previously ran with full root privileges in order to perform a privileged task such as binding to ports lower than 1024, reading from and writing to user home directories, or accessing the Ethernet device. Since setuid root binaries and daemons that run with full root privileges are rarely necessary under the least privilege

model, an exploit in a program no longer means a full root compromise. Damage due to programming errors like buffer overflows can be contained to a non-root user, which has no access to critical abilities like reading or writing protected system files or halting the machine.

Network Security and Encrypted Communications

Oracle Solaris provides protection against inappropriate use of network resources through its secure by default networking configuration, which disables many unused network services to reduce exposure to attack. With secure-by-default networking, an administrator can enable or disable individual network services or change how they listen for network connections.

Oracle Solaris includes many network security and communication encryption features, including:

- **Firewall**. Oracle Solaris also ships with an integrated IP filter firewall software preinstalled, which reduces the number of network services that are exposed to attack. It also provides protection against maliciously crafted networking packets. IP filter firewall can also filter traffic flowing between Oracle Solaris Containers when configured in the global zone. In addition, TCP wrappers are integrated into Oracle Solaris, limiting access to service-based allowed domains or partner sites.

- **IPsec**. IPsec offers a robust set of capabilities for protecting network communication. This includes well-defined security protocols, a security association database, key management, authentication and encryption algorithms, and a security policy database. Note that the SMF manages IPsec as a set of services, thereby increasing reliability.
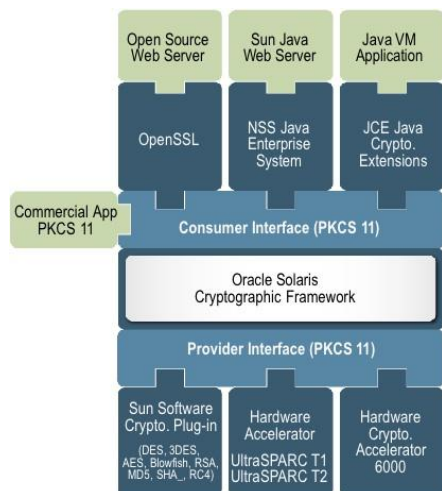
Cryptographic and Key Management Framework

Figure 5: Oracle Solaris Cryptographic Framework is standardized and extensible—current and future cryptographic choices can easily plug in and take advantage of hardware and software capabilities.

The Oracle Solaris Cryptographic Framework provides standardized and extensible cryptographic services to applications and kernel modules in a manner seamless to the end user. It brings direct cryptographic services, such as encryption and decryption for files, to the end user. The user-level

framework is responsible for providing cryptographic services to consumer applications and the end-user commands. The kernel-level framework provides cryptographic services to kernel modules and device drivers. Both frameworks give developers and users access to software-optimized cryptographic algorithms. These pluggable frameworks balance loads across hardware accelerators and software implementations, increasing encrypted network traffic throughput. They're available to applications written to use the PKCS #11, OpenSSL, and Java Cryptography Extension APIs.

The Cryptographic Framework takes advantage of hardware acceleration in both SPARC and x86 platforms, and features strong cryptographic routines—up to AES 256-bit key length. The Oracle Solaris Cryptographic Framework can provide performance and security benefits to both system administrators and developers. For example, applications and directory services can program to a standard interface (PKCS#11 providers) from Java or other development environments and take full advantage of a range of hardware cryptographic accelerators for SSL, token cards, or secure network transport between data repositories and business logic layers. Note that many system services in Oracle Solaris, such as IPSec/IKE and Kerberos authentication already take advantage of the Cryptographic Framework and will automatically use the hardware acceleration provided by UltraSPARC T-Series processors, as well as processors from Intel and AMD.

File Integrity

System administrators can detect possible attacks on their systems by monitoring for changes to file information. Oracle Solaris utilizes binaries that are digitally signed, so that administrators can easily track changes. All patches or enhancements are embedded with digital signatures, eliminating the false positives associated with upgrading or patching file integrity-checking software. Oracle Solaris also provides the basic audit reporting toolkit (BART) for checking the integrity of customer files.

The Oracle Solaris Fingerprint Database is a service that enables users to verify the integrity of files distributed with Oracle Solaris, including Solaris executable binaries back to Solaris 2.1. You can use this service to compare the hash of a suspected file to determine if it is as expected, and not running an unauthorized patch or hacked.

Common Criteria and Trusted Extensions

Oracle Solaris has been tested on all Sun servers against the stringent Common Criteria testing process and has achieved Evaluation Assurance Level 4+ certification against three protection profiles: Controlled Access, Role-Based Access Control, and Labeled Security. This allows customers to run their applications on one of the most highly certified operating systems in the world without the need for special programming or modifications to their applications.

Oracle Solaris with Trusted Extensions utilizes User and Process Rights Management, Oracle Solaris Containers (see Chapter 5 – Oracle Server Virtualization), file systems, and networking and does not require a new or separate kernel. Best of all, it does not require third-party developers to requalify their applications to run them with sensitivity labels. Because it's an extension to the Oracle Solaris security policy, Trusted Extensions technology is flexible and quick to deploy: You can add new applications,

new users, and more, very quickly, without extensive analysis of each application—without the need to write complex, error-prone security policies that require a system reboot.

5. Oracle Server Virtualization

Oracle has demonstrated support for over 8,000 Oracle Solaris Containers on a single, rack-mount Sun SPARC Enterprise T-Series server, highlighting superior performance, scalability, manageability, and minimal overhead. IBM has recently claimed they can deploy up to 1,000 VMs on a newly-announced system.

Figure 6: Click the image to watch Edward Screven discuss Oracle's comprehensive and innovative approach to virtualization. Watch the full podcast.

Virtualization provides the ability to deliver more work from an existing IT infrastructure by increasing utilization. As the power of today's servers continues to increase well beyond the needs of a single application stack, virtualization has become the must-have technology. Virtualization helps consolidate legacy applications from multiple obsolete hardware platforms onto a smaller number of up-to-date, more powerful, and more energy-efficient servers. It supports moving today's applications from a large set of underutilized servers to a smaller set of more powerful servers, helping to reduce the number of servers to house, power, cool, and maintain. Raising utilization levels helps to reduce inefficiency, helping with datacenter space, power, and cooling requirements. Organizations are increasingly using virtualization to increase business agility, which increase speed and flexibility in delivering IT services to support business goals.

Oracle provides strong capabilities for desktop, server, and storage virtualization. Enterprise users need choice when it comes to server virtualization and consolidation, and flexibility with respect to application, OS, and network virtualization methods. Oracle offers a full portfolio of virtualization solutions to address your needs. Virtualization capabilities are designed into Oracle Solaris to maximize productivity and flexibility.

Figure 7: Oracle offers the industry's most complete virtualization portfolio.

Oracle supports several complementary virtualization technologies, each of which provide different degrees of isolation, resource granularity, and flexibility. Oracle supports virtualization technologies that allow multiple OS (and application) instances to run on the same server, while each instance has the illusion of owning its own hardware resources. These capabilities are built into Oracle Solaris and Oracle's Sun servers—there are no additional costs to use them.

- Oracle VM is a family of products that work together to facilitate the creation and management of virtual environments. Consisting of Oracle VM Server and the integrated Oracle VM Manager browser-based management console, Oracle VM makes it easy to create and manage virtual server pools running on systems across the enterprise.

  - Oracle VM Server for x86 provides technology and an easy-to-use graphical interface for creating and managing virtual server pools running on x86 systems. Certified for a wide range of Oracle products—including Oracle Database, Fusion Middleware, Oracle applications, and Oracle Enterprise Linux, Oracle VM Server for x86 fully supports non-Oracle applications, as well as Oracle Solaris, Linux, and Windows guests.

  - Oracle VM Server for SPARC is purpose-built for Oracle servers with SPARC chip multithreading (CMT) technology. Oracle VM Server for SPARC (previously called Sun Logical Domains) provides a full virtual machine that runs an independent operating system instance and contains a wide range of virtualized devices. A hypervisor that largely resides in a chip on the server is tightly integrated with the hardware, enabling virtual machines to take advantage of underlying system advancements and reduce the overhead typically associated with software-based solutions. Oracle VM Server for SPARC supports virtualized CPU, memory, storage, I/O, console, and cryptographic devices, and redundant I/O paths—unlike solutions from other vendors—to make maximum use of platform resources.

  - Oracle VM Manager provides an easy-to-use, feature-rich graphical interface for creating and managing Oracle VM environments. Administrators can enable advanced functionality to load balance across resource pools and automatically reduce or eliminate outages associated with server downtime.

  - Oracle VM Templates facilitate rapid software deployment by providing preinstalled and preconfigured software images. You can shorten time to market, eliminate installation and configuration costs, and reduce ongoing maintenance and operational costs. Templates are available for immediate download for many key Oracle products, including Oracle Database, Oracle Enterprise Linux, Oracle Fusion Middleware, and others.

- Dynamic Domains provide hardware partitioning capabilities on selected SPARC Enterprise servers.

- Oracle Solaris Containers provide security and resource isolation that allows multiple virtual Oracle Solaris environments to share the same OS instance. Oracle Solaris Containers complements the capabilities of Oracle VM Server for SPARC and Dynamic Domains, and increases security and utilization on all SPARC Enterprise servers.

Oracle VM Server for SPARC and Oracle Solaris Containers are multithreaded to maximize performance and utilization.
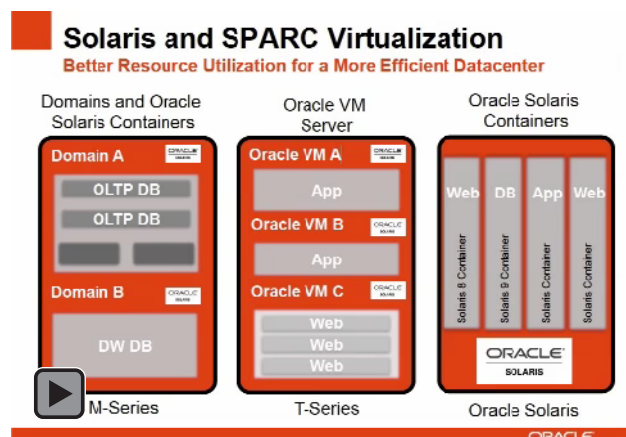
Oracle VM Server for SPARC

Figure 8: Click to watch Charlie Boyle talk about Oracle Solaris and SPARC virtualization. Click here to watch the entire webcast.

Oracle VM Server for SPARC, previously called Sun Logical Domains, leverages the built-in SPARC hypervisor to subdivide supported platforms' resources (CPUs, memory, network, and storage) by creating partitions called logical (or virtual) domains. Each logical domain can run an independent operating system. Oracle VM Server for SPARC provides the flexibility to deploy multiple Oracle Solaris operating systems simultaneously on a single platform. Oracle VM Server for SPARC also allows you to create up to 128 virtual servers on one system to take advantage of the massive thread scale offered by the CMT architecture. Sun SPARC Enterprise T-Series servers come with the right to use (RTU) for Oracle VM Server for SPARC, and the software is pre-installed.

Oracle VM Server for SPARC integrates both the industry-leading CMT capabilities of the UltraSPARC T-Series processors and Oracle Solaris. This combination helps to increase flexibility, isolate workload processing, and improve the potential for maximum server utilization.

SPARC is the leading platform with the hard partitioning capability that provides the physical isolation needed to run independent operating systems. Many customers have already used Oracle Solaris Containers for application isolation. Oracle VM Server for SPARC provides another important feature with OS isolation. This gives you the flexibility to deploy multiple operating systems simultaneously on a single Sun SPARC Enterprise T-Series server with finer granularity for computing resources. For SPARC CMT processors, the natural level of granularity is an execution thread, not a time-sliced microsecond of execution resources. Each CPU thread can be treated as an independent virtual processor. The scheduler is built into the CPU, without the extra overhead for scheduling in hypervisor. You just have one software scheduler—the Oracle Solaris scheduler—to dispatch workloads to virtual CPUs, which are effectively physical CPU threads. What you get is a virtualization solution with "bare-metal" performance—lower overhead, and higher performance and scalability.

Your organizations can couple Oracle Solaris Containers and Oracle VM Server for SPARC with the breakthrough space and energy savings afforded by Sun SPARC Enterprise T-Series servers to deliver a more agile, responsive, and low-cost environment.

About Oracle VM Server for SPARC

A logical domain is comprised of four key architectural components, which work together to manage and isolate the individual logical domains.

- **Hypervisor.** A small firmware layer that provides a set of hardware-specific support functions to operating systems, using extensions built into the sun4v CPU. It provides virtualization capabilities by abstracting the physical devices and presenting the virtual resource to the individual logical domains. The hypervisor also creates logical domain channels (LDCs), which provides the necessary communication between individual domains for services such as networks and shared devices. Data is transferred between domains using shared memory under the control of the LDC and the memory management unit (MMU) of the UltraSPARC T-Series processors. This enables secure and fast data and message copying between domains.

- **Physical and Virtual Devices.** Physical system hardware that is abstracted by the hypervisor. The devices are made available by the Service Domain to the individual guest logical domains, according to configurations specified in the Control Domain. Devices may be purely virtual, with no physical equivalent, as is the case with the virtual system consoles that enable access to each domain's Open Boot PROM (OBP), or may be tied to physical assets such as disks. Virtual devices in the Sun T-Series servers include:

  - CPU

  - Memory

  - Cryptographic cores

  - Network switches

  - NICs

  - Disk servers

  - Disk consoles

  - Virtual terminal server

Each domain can perform up to four roles. Depending on the system configuration, a logical domain may function in one or more of the following:

- **Control Domain.** As the configuration platform for managing server and domains, the Control Domain executes LDM. Currently, the Control Domain is also required to be a Service Domain. There must be one Control Domain on each server configured to run Oracle VM Server for SPARC.

- **Service Domain.** Provides virtual device services to other domains such as a virtual switch, a virtual console concentrator, and a virtual disk server. The Service Domain interfaces with the hypervisor to manage individual logical domain access to hardware resources such as CPU, memory, network, disk, console, and cryptographic units. For additional flexibility and control, administrators can

dynamically add and delete memory to domains, and processor cores and threads can be bound (affinity) to specific domains.

- **I/O Domain.** Controls direct access to I/O devices such as PCI-Express cards, storage units, and network devices as well as providing I/O access for other domains. The I/O Domain owns the physical I/O and provides physical I/O facilities to itself and virtual I/O facilities to other guest domains. The I/O Domain provides I/O load separation and redundancy within domains deployed on a platform.

- **Guest Domain.** An independent operating environment where applications are hosted. Multiple Guest Domains can exist, and each can be "powered up" and rebooted without affecting other domains. Guest domains use one or more Service Domains to obtain I/O.

With its straightforward and flexible architecture running on the Sun CoolThreads servers, Oracle VM Server for SPARC can be a solution for a variety of consolidation projects. A few scenarios where Oracle VM Server for SPARC virtualization technology can help create a more efficient and flexible datacenter are:

- Consolidation of up to 128 UNIX or Linux servers.

- Consolidation of a test, development, and QA environment where domain independence and flexible resource allocation are necessary.

- Consolidation of environments where there is a requirement for different operating systems that require different patch levels, in contrast with Oracle Solaris Containers, which run in the same Oracle Solaris instance and kernel patch level.

- Consolidation efforts where maximum isolation and security are needed to protect access to sensitive data.

Oracle VM Server for SPARC can provide both utilization and resiliency. By allowing several applications and operating system instances to share hardware resources, Oracle VM Server for SPARC promotes higher utilization rates and a better return on investment. By abstracting software from the underlying hardware, virtualization can offer higher application service availability by creating a redundant infrastructure that can provide failover domains if there is a planned or unplanned outage. Because new guest domains can be created quickly compared to the lengthy procedure of provisioning a new server, CPU resources can be moved from idle to busy domains—taking effect very quickly— enabling the flexible assignment of infrastructure to applications as needed. Oracle VM Server for SPARC helps organizations:

- Overcome datacenter power and cooling constraints

- Improve IT and business flexibility

- Reduce server sprawl and improve administration efficiency

- Improve availability and reliability in a consolidated environment

- Rapidly grow and shrink infrastructure

- Increase availability and application performance

- Simplify application management and increase application performance per dollar spent

Oracle Solaris Containers

Supported on any server running Oracle Solaris, Oracle Solaris Containers isolate software applications and services using flexible, software-defined boundaries. Oracle Solaris Containers provide virtualization and software partitioning, enabling the creation of many private execution environments from a single instance of Oracle Solaris. Oracle Solaris Containers supports over 8,000 instances on a single Sun SPARC Enterprise T-Series server, and can help increase on-demand scalability by dynamically resizing workloads on demand, as shown in Figure 9.

Figure 9: Oracle Solaris Containers have very little overhead, and can be dynamically resized based on changing workload requirements, providing capacity on demand.

Deploying applications in Oracle Solaris Containers can help improve security and reduce licensing costs. The sophisticated network-wide security system in the operating system controls the way users access files, protect system databases, and use resources. Security needs are addressed at every layer, from integrated security services and applications, to enhanced encryption algorithms, to an enterprise firewall for network protection. Extended security features are also available, including authentication, data integrity, data privacy, and single sign-on capabilities so that tampering, snooping, and eavesdropping do not compromise data or transactions. Systems also can be hardened and minimized as needed, helping to reduce the risk that a system or application can be compromised. Oracle Solaris can easily migrate physical Oracle Solaris 10 systems to an Oracle Solaris 10 Container, including `hostids`. And unlike server virtualization solutions from other vendors that charge per processor core, Oracle Solaris Containers are included as part of Oracle Solaris, at no additional cost.

Unlike virtual machines, Oracle Solaris Containers provide OS-level virtualization because it gives the appearance of multiple OS instances rather than multiple physical machines. Isolation between Containers is accomplished by restricting the scope of system calls, rather than the CPU-intensive task of emulating hardware architectures and instruction sets in software. This makes it possible to create hundreds, even thousands, of Oracle Solaris Containers on a single system. Because of this negligible

overhead, and unlike partitioning or virtual machines, Oracle Solaris Containers can be created in large numbers. For example:

- Individual developers can use safe, isolated test environments.

- Service providers can provide isolated instances of Web servers or database instances.

Hosting applications within individual Oracle Solaris Containers provides administrators the ability to exert fine-grained control over rights and resources within a consolidated server.   Containers create very low overhead compared to traditional virtual machines, maximizing the computing resources available to applications. Organizations can safely and more effectively consolidate applications onto a single server. Computing resources can be—CPUs, physical memory, network bandwidth, and more— can be dedicated to a single application one moment and then shared with others in an instant, all without moving applications or rebooting the system, dynamic domain, or logical domain where the Oracle Solaris Container resides.

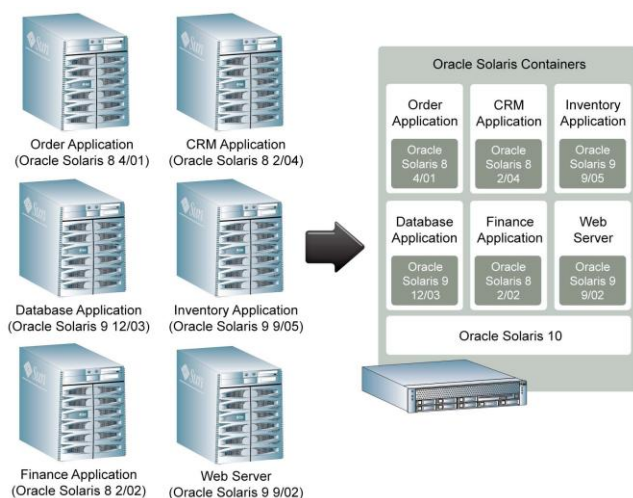**Support Legacy Applications with Modern Hardware**

Figure 10: Oracle VM Server lets companies consolidate legacy applications onto a single server to raise resource utilization rates and lower operating expenses

With Oracle Solaris 8 and Oracle Solaris 9 Containers, you can safely and easily move your existing applications and environments from a physical server to a software Container on the latest Sun SPARC Enterprise server running Oracle Solaris 10. This means you can migrate existing applications to new, more powerful, energy-efficient, and productive systems, and transition these legacy environments to native Oracle Solaris Containers at your own pace. This may save on licensing costs, and offers the benefits of Oracle Solaris, such as Oracle Solaris DTrace, Predictive Self Healing, and Oracle Solaris ZFS.

Dynamic Domains and Dynamic Reconfiguration

A key feature of the Sun SPARC Enterprise M-Series high-end servers with Oracle Solaris is the ability to partition the available hardware resources into smaller logical systems. Sun SPARC Enterprise M-Series servers offer *hard partitioning* technology in the form of Dynamic Domains. Instantiating a number of Dynamic Domains on a Sun SPARC Enterprise M-Series server divides the system into multiple electrically isolated partitions. Each Dynamic Domain executes a unique instance of Oracle Solaris. Since isolation is instantiated all the way to the hardware, configurations can be created in which software changes, reboots, and potential faults in one domain do not impact applications running in other domains. SPARC Enterprise M-Series servers can provide up to 24 Dynamic Domains, each with configurable amounts of CPU, memory, disk, and I/O resources such as PCI Express and PCI-X slots, and networking.

Fault Isolation and Error Management

Domains are protected against software or hardware failures in other domains. When a domain encounters a fatal error, a *domainstop* operation occurs that cleanly and quickly shuts down only the domain with the error. Domainstop operates by shutting down the paths in and out of the system address controller and the system data interface application-specific integrated circuits (ASICs). The shutdown is intended to prevent further corruption of data, and to facilitate debugging by not allowing the failure to be masked by continued operation.

Dynamic Reconfiguration and Automated Dynamic Reconfiguration (ADR) allow resources to be dynamically reallocated, or balanced, between domains. Utilizing this technology enables a physical or logical restructuring of the hardware components of Sun SPARC Enterprise M-Series servers while the system is running and the applications remain available. This high degree of resource flexibility allows the domain or platform administrator to reconfigure the system easily in order to provision the resources to meet changing workload demands. Disaster recovery (DR) can also be used to remove and replace failed or upgraded hardware components while the system is online[2]. CPU, memory, and I/O devices be added or deleted by Dynamic Reconfiguration.

The Reconfiguration Coordination Manager (RCM) is the framework that manages the dynamic removal of system components. By using RCM, you can register and release system resources in an orderly manner. Using RCM, it is also possible to write a script that allows Oracle Database to be alerted when new CPUs or memory are to be removed from the domain, so that the SGA can be dynamically scaled back to allow the board to be removed without shutting down the database.

---

[2] Sun SPARC Enterprise M4000, M5000, M8000, and M9000 servers can perform DR to logically move system resources between domains. In addition, Sun SPARC Enterprise M8000 and M9000 servers can perform hot-swap operations to physically add or remove boards from the chassis.

Working with Oracle Solaris Containers

Dynamic Domains can be used with Oracle Solaris Containers technology to refine resource control and simplify the consolidation of several applications into one domain. As described previously, the Oracle Solaris Containers functionality in Oracle Solaris enables multiple, software-isolated applications to run on a single server or domain.

Comprehensive Management with Oracle Enterprise Manager Ops Center

Managing physical and virtual server environments can be complex and time-consuming. Oracle Enterprise Manager Ops Center is the most comprehensive solution for managing physical and virtual Sun infrastructure, including Sun servers, Oracle Solaris, and Oracle VM Server. It also provides management for other operating systems. Designed to handle the entire deployment lifecycle, Oracle Enterprise Manager Ops Center is a single platform that helps to bring complexity under control by managing multiple server architectures, and multiple operating systems running on bare hardware or in virtualized environments.

The Oracle Enterprise Manager Ops Center Virtualization Management Pack streamlines operations, reduces downtime, and provides an end-to-end management solution for physical and virtual systems through a single web-based console. This solution automates the lifecycle management of physical and virtual systems and is the most effective systems management solution for Oracle's Sun infrastructure.

6. Developer Tools Optimizations

The Oracle Solaris ecosystem, including Oracle Solaris and Oracle Solaris Studio development tools, offers a compelling platform for developers to embrace the breakthrough capabilities of latest systems based on SPARC and x86 systems. In fact, all top Oracle software is built with Oracle Solaris Studio, including the Oracle Database, Oracle Solaris, the Java VM and key applications such as Peoplesoft, Siebel, JD Edwards, and Hyperion. With a focus on optimizing Oracle Solaris Studio for the latest Sun SPARC and x86 systems, Oracle Solaris Studio is the development platform of choice for Oracle enterprise customers and ISVs.

One of the most significant roles the Oracle Solaris Studio development tools play in the creation of robust applications is their ability to help streamline and automate the optimization process, including tuning the software for greater performance and reliability. Because these tools are designed with an intimate knowledge of both x86 and SPARC-based systems, developers can take maximum advantage of hardware features without specific hardware expertise. Internal testing using well-regarded industry benchmarks showed that Oracle Solaris Studio 12.2 outperforms open-source alternatives by up to 4.8x on floating-point benchmarks and 50 percent in integer benchmarks. Oracle Solaris Studio tools also help increase reliability with comprehensive debugging capabilities, such as highlighting memory leaks. The result delivers a robust choice for both established enterprise datacenters and reliable, leading edge applications.
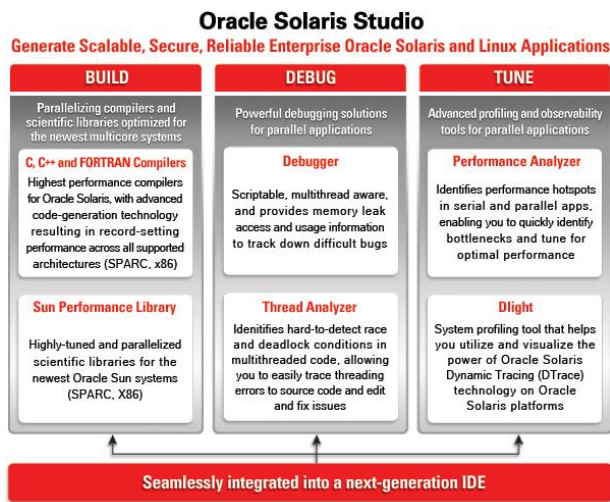
Figure 11: Oracle Solaris Studio is a platform for end-to-end application development

Oracle Solaris Studio

Oracle Solaris Studio improves both the development process and ultimate performance in multicore application development. It serves as a comprehensive build/debug/tune facility for SPARC and x86 systems, offering award-winning compilers (C, C++, Fortran) optimized for the latest multicore architectures, thread analysis tools, compiler auto-parallelization, OpenMP and MPI support, performance analysis tools, multithreaded debugging, and more. Oracle Solaris Studio delivers an advanced suite of tools designed to work together to provide an optimized environment for the development of single, multithreaded, and distributed applications. The debugging and analysis tools take advantage of compiler features to provide application context with high levels of accuracy, leading to more-robust software. Oracle Solaris Studio also comes with an integrated development environment (IDE) tailored for use with the included compilers, the debugger, and the analysis tools. This IDE increases developer productivity with a code-aware editor, workflow, and project functionality. In addition, the parallelizing C, C++, and Fortran compilers; enhanced math routines; and performance analysis tools enable users to maximize the performance of their applications on Sun SPARC server systems, and on other x86-based systems, generating higher ROI from deployment hardware systems.

As shown in Figure 11, Oracle Solaris Studio offers an optimized, comprehensive development environment for SPARC- and x86-based systems. This includes tools and environments to build, debug, and tune applications.

Oracle Solaris Studio IDE

Improving developer productivity, the next-generation IDE provides full edit, compile, and debug support including code completion, error highlighting, semantic highlighting, call graph, memory

window, `makefile` wizard/import, packaging of application as tar files, zip files, SVR4 packages, RPMs, or Debian packages, and much more.

C, C++ and Fortran Compilers

Oracle Solaris Studio delivers compilers that produce record-setting application performance— consistently exceeding that of open source alternatives. The C and C++ compilers provide a solid foundation for building robust, high-performance parallel code for the newest Sun SPARC and SPARC64 systems, as well as those based on Intel and AMD processors, from Oracle as other OEMs. In addition to supporting the latest language standards, Oracle Solaris Studio software includes GNU C/C++ compatibility features and is source- and object-level compatible with prior releases.

To take advantage of hardware concurrency in multicore systems, the compilers simplify the creation of parallel applications with autoparallelization features. These features enable the compiler to identify safe and productive parallelization opportunities in single-threaded code and automatically convert those segments into multithreaded code. In addition, the compilers support the OpenMP 3.0 specification that introduces task-based parallelism.

The compilers in Oracle Solaris Studio include an array of optimization options for increasing application performance, including:

- **Automatic Parallelization**. By selecting a compiler option, the compiler performs the dependence analysis to determine if a specific part of the program can be executed in parallel. If it can prove this is safe to do, it generates the underlying infrastructure, typically a series of calls to a multitasking library.

- **OpenMP**. OpenMP supports an extensive set of features to specify the parallelism, control the workload distribution, and synchronize the threads. Current OpenMP implementations are built on top of a native threading model.
  The Oracle Solaris Studio compilers support a combination of Automatic Parallelization and OpenMP. This minimizes the effort, since the compiler can first be used to parallelize the application. Those parts that are too complicated for the compiler to handle can then subsequently be parallelized with OpenMP.

- **MPI**. MPI provides for parallelization of applications running across many such systems, often referred to as a compute cluster. Outside of the basic functions to create the processes, as well as send and receive messages, MPI provides a very rich API through an extensive set of additional functions, including various ways to handle I/O.

- **Hybrid**. In a Hybrid parallel application, distributed and shared memory programming models are combined to parallelize an application at two levels. Typically, MPI is used for the distributed memory component, spreading the work over the nodes in a cluster. The process(es) running within one node are then further parallelized using a shared memory model, typically OpenMP. The Hybrid model is a very natural fit for a cluster consisting of multicore nodes. You can run an MPI application across all of the nodes and cores in the system, and use OpenMP for finer-grained parallelization where needed.

Another reason to consider the Hybrid model is that the memory within one node is used more economically by exploiting shared data through OpenMP, avoiding the need to replicate data. In such cases, this combined model is very suitable. It is often relatively easy to use OpenMP to implement the second level parallelism.

Oracle Solaris Performance Library

The Oracle Solaris performance library is a set of optimized, high-speed mathematical subroutines for solving linear algebra and other numerically intensive problems. It provides a performance boost to high-performance computing, financial, and other compute-intensive applications. The Oracle Solaris performance library contains enhanced and newly added standard routines such as BLAS, FFTPACK, LAPACK, ScaLAPACK, Sparse Solvers, and SuperLU.

Debugger

Ensuring application reliability, the Debugger helps track down difficult bugs in serial and parallel code. It also provides memory leak, access, and usage information. The Debugger is fully integrated into the Oracle Solaris Studio IDE, available via the command line, and available as a high-productivity standalone graphical debugger.

Thread Analyzer

Improving developer productivity and software robustness, the thread analyzer tool identifies hard-to-detect threading errors before they occur. It can detect potential race and deadlock conditions at runtime, map them to source lines in the application, and then enable the user to view the results by using command-line or graphical user interface (GUI) options.

Performance Analyzer

The performance analyzer tool identifies application performance bottlenecks, by specifying not only which functions, code segments, and source lines are having an impact on performance but also by providing the tools necessary to do tuning for optimal performance. From annotated compiler commentary listings in which the compiler indicates a range of information to optimization status and runtime thread performance, users can visualize performance hotspots with the GUI. The performance analyzer tool can be used to profile single-threaded as well as multithreaded applications.

Oracle Solaris provides the ability to observe performance characteristics of applications using performance counters. For example, counters can be used to determine the average cycles per instruction for a given workload, determine how cache/memory intensive an application is, or determine whether there are any serious memory alignment issues with the way that an application lays out its data. Oracle Solaris Studio uses DTrace probes and other techniques to monitor and analyze performance using these counters.

Oracle Solaris DLight

Oracle Solaris DLight is a plug-in for the Oracle Solaris Studio development environment that unites information you get from typical application profiling tools with system profiling tools such as DTrace. DLight is a visual profiling tool that unifies application and system profiling, using DTrace technology on Oracle Solaris platforms, and providing new levels of insight to dramatically reduce development timelines.

Sun Performance Library

Maximizing application performance, the Sun Performance Library provides a set of optimized, high-speed mathematical subroutines for solving linear algebra and other numerically intensive problems. The Sun Performance Library contains enhanced and newly added standard routines including, BLAS, LAPACK, FFTPACK, SuperLU, Sparse Solvers, and ScaLAPACK.

Customers worldwide depend on Oracle Solaris to run their business, with good reason:

- **Unmatched reliability** with Predictive Self Healing for hardware and software faults, data integrity with Oracle Solaris ZFS and DTrace for safe, live observability of the Oracle Solaris kernel and applications.

- **Optimized performance and scalability** for the latest SPARC, Intel, and AMD processor technologies along with efficient datacenter consolidation with Oracle Solaris Containers.

- **Mission critical Oracle Solaris security infrastructure** that provides the compartmentalization and control required by governments and financial institutions.

7. Conclusion

"Oracle Solaris 10 is reliable and predictable – and has what we need in an operating system. The Solaris 10 OS allows multiple applications to be consolidated onto a single system through virtualization, which makes it an ideal solution for standardization. Virtualization features ensure high availability, which is critical to Argonne's environment."

– David Salbego, Director of IT Infrastructure
Argonne National Labs

Thousands of enterprise customers worldwide depend on Oracle Solaris to run their business, with good reason:

- **Unmatched reliability** with Predictive Self Healing for hardware and software faults, data integrity with Oracle Solaris ZFS and DTrace for safe, live observability of the Oracle Solaris kernel and applications.

- **Optimized performance and scalability** for the latest SPARC, Intel, and AMD processor technologies along with efficient datacenter consolidation with Oracle Solaris Containers.

- **Mission critical Oracle Solaris security infrastructure** that provides the compartmentalization and control required by governments and financial institutions.

Oracle Solaris has continued to demonstrate great success as a mission-critical, enterprise-class OS for scalable performance, relentless reliability, and comprehensive security in the datacenter. Continuing efforts at application-to-disk integration with Oracle Applications will optimize these capabilities, and increase innovation for the complete solution stack.

Reliability features such as Predictive Self Healing can alert you before there is a problem, and restart services as needed. Record-setting benchmarks, including TPC-H and TPC-C, PeopleSoft, Oracle BIEE, and many others, demonstrate maximum utilization. No-cost virtualization provides fine-grain, end-to-end technology to efficiently consolidate your datacenter applications. Finally, Ops Center can help unlock the competitive advantage the *complete* Oracle system offers by bringing assets into production faster from a single comprehensive management interface.

Oracle Solaris is widely recognized as the enterprise operating system of choice for mission-critical applications. To learn more about each of the specific products, technologies, and capabilities discussed in this document, please refer to the next section, or contact your Oracle representative.

8. Resources

GET THE PRODUCTS

Oracle Solaris	www.oracle.com/technetwork/server-storage/solaris/downloads/index.html
Oracle Enterprise Manager Ops Center	www.oracle.com/us/products/enterprise-manager/opscenter/index.html
Oracle Solaris Cluster	www.oracle.com/technetwork/server-storage/solaris/downloads/index.html
Oracle Solaris Studio	www.oracle.com/technetwork/server-storage/solarisstudio/downloads/index.html
Oracle Database 11*g*	www.oracle.com/us/products/database/
Oracle Siebel CRM	www.oracle.com/us/products/applications/siebel/
Oracle PeopleSoft Enterprise Applications	www.oracle.com/us/products/applications/peoplesoft-enterprise/
Oracle Real Application Clusters (RAC)	www.oracle.com/technology/products/database/clustering/
Oracle Applications	www.oracle.com/us/products/applications/index.html

DEEP DIVE ON THE TECHNICAL

Oracle Technical Network	www.oracle.com/otn
Reliability	
Oracle Solaris Predictive Self Healing	www.oracle.com/us/products/servers-storage/solaris/solaris-pred-self-healing-ds-075587.pdf
Oracle Solaris and Oracle Solaris Cluster: Extending Oracle Solaris for Business Continuity	www.oracle.com/technetwork/server-storage/solaris-cluster/documentation/solaris-cluster-businesscontinuity-168285.pd
Oracle Sun Cluster Documentation Center	http://docs.sun.com/app/docs/coll/1124.5?l=en
Oracle Solaris Cluster – technical content	www.oracle.com/technetwork/server-storage/solaris-cluster/overview/index.html
Oracle Solaris ZFS	www.oracle.com/us/products/servers-storage/solaris/solaris-zfs-ds-067320.pdf
Performance	
Oracle Solaris Performance and Scalability	www.oracle.com/us/products/servers-storage/solaris/performance-scalability-066074.html
Oracle Database 11*g* Performance and Scalability	www.oracle.com/technetwork/database/features/performance/index.html
Flash Storage	www.oracle.com/us/products/servers-storage/storage/flash-storage/index.html
Security	
Oracle Solaris Security	www.oracle.com/us/products/servers-storage/solaris/security/index.html
Oracle Solaris Common Criteria	www.oracle.com/technetwork/topics/security/oracle-cc-evalsolaris-083233.html
Using Oracle Solaris 10 to Overcome Security Challenges	www.oracle.com/technetwork/articles/solarisadmin/solaris10securitychallenges-163842.pdf
Virtualization	
Oracle Virtualization	www.oracle.com/virtualization
Oracle VM Demo	www.oracle.com/pls/ebn/swf_viewer.load?p_shows_id=8910842
Oracle Server Virtualization Webcast – 11 parts	http://event.on24.com/event/23/07/86/rt/index.html?eventid=230786&sessionid=1&partnerref=email&key=ACD102D5564F47339E77B2101D7099D4&eventuserid=39734932
Oracle's Virtualization Blog	http://blogs.oracle.com/virtualization

Virtualization with Oracle Solaris 10	www.oracle.com/go/?&src=7011689&Act=4
Oracle VM Server for SPARC—Enabling a Flexible, Efficient IT Infrastructure	www.oracle.com/us/products/servers-storage/servers/sparc-enterprise/vm-server-for-sparc-wp-075964.pdf
Using Logical Domains And CoolThreads Technology: Improving Scalability and System Utilization	www.oracle.com/technetwork/articles/systems-hardware-architecture/vm-svr-for-sparc-163928.pdf
Consolidating Oracle's Siebel CRM 8 on Single Sun SPARC Enterprise Server	www.oracle.com/technetwork/articles/systems-hardware-architecture/siebel-crm-ha-163926.pdf
Oracle Solaris Containers (Zones)	www.oracle.com/us/products/servers-storage/solaris/solaris-containers-ds-075585.pdf
Developer Tools	
Oracle Solaris Studio	www.oracle.com/us/products/tools/050872.html
Oracle Solaris Dynamic Tracing WP	www.oracle.com/us/products/servers-storage/solaris/solaris-dynamic-tracing-ds-067318.pdf
Parallel Programming with Oracle Developer Tools	www.oracle.com/technetwork/systems/parallel-programming-oracle-develop-149971.pdf
DLight for Sun Studio	www.oracle.com/technetwork/server-storage/solaris/overview/observability-163553.html
OTN—DTrace	www.oracle.com/technetwork/server-storage/solaris/power-of-dtrace-136658.html

Authors and Contributors

The following people contributed to this white paper:

- **Performance and scalability**. Andy Bowers, Glenn Fawcett, Keng-Tai Ko, Jim Mauro, Allan Packer, Steve Sistare

- **Reliability.** Stephanie Choyer, Burt Clouse, Tarek Derbas, Eve Kleinknecht, Amour Kwok, Gia-Khanh Nguyen, Dong Tang.

- **Security.** Mark Thacker, Terri Wischmann

- **Virtualization.** Charlie Boyle, Honglin Su

- **Oracle Solaris Studio.** Ikroop Dhillon, Don Kretsch

- **Oracle Solaris.** Chris Baker, Art Beckman, Debbie Franklin, Mary Jane Greenfield, Jeff McMeekin, Mike Mulkey, Lynn Rohrer, Todd Tornga, Larry Wake, Markus Weber

- **Oracle**. Jeff Ramsey

ORACLE®

Oracle Solaris: In a Class by Itself
September 2010

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

SOFTWARE. HARDWARE. COMPLETE.