

A Fujitsu White Paper
May 2011

Oracle Solaris and Fujitsu SPARC Enterprise Systems — Integrated and Optimized for Enterprise Computing

Executive Overview	2
Introduction—Datacenter Integration	2
Overview	2
The Oracle Solaris Ecosystem	2
SPARC Processors	3
Architected for Reliability	6
Oracle Solaris Predictive Self Healing	7
Highly Reliable Memory Subsystems	8
Oracle Solaris ZFS for Reliable Data	9
Reliable Networking	9
Scalable Performance	10
World Record Performance	12
Network Performance	14
Security	14
Integrated with Fujitsu SPARC Enterprise T-Series Servers	15
The Oracle Solaris Cryptographic Framework Library	16
Preventing Attacks	17
Least Privilege	17
Fujitsu Server Virtualization	18
Oracle VM Server for SPARC	18
Oracle Solaris Containers	19
Hardware Partitioning and Dynamic Reconfiguration	20
Conclusion	21
Resources	22

Executive Overview

This document is intended for IT architects, system administrators, and developers that want to understand the details of how Oracle® Solaris and SPARC® can improve your application solution environment. This paper will provide technical information on how Oracle Solaris and Fujitsu servers designed with the SPARC processor have been highly optimized for each other, improving throughput, security, and resiliency throughout the application solution stack, driving maximum ROI and minimum TCO. It includes brief technical descriptions of how specific Oracle Solaris features and capabilities are implemented in a system-wide approach to optimize the specific functionality of the SPARC processor family in the areas of scalable performance, advanced reliability, security, and cost-effective virtualization.

Introduction—Datacenter Integration

Oracle Solaris and Fujitsu have a long history of optimizing the platforms for scalability, reliability, and security. Working together, these improvements have enhanced and optimized the entire stack and leveraged innovation throughout. This paper offers a high-level discussion of the benefits of Oracle Solaris running on Fujitsu SPARC Enterprise M-Series and T-Series servers, and detailed information on specific optimizations and advantages for increased reliability, scalability, security, and virtualization. Resources that can provide more information are listed at the end of each section, and a comprehensive list is available at the end of the paper.

Overview

Thousands of customers worldwide depend on SPARC-based systems and Oracle Solaris to run their business, usually for one simple reason—these platforms simply don't quit. Maximum scalability is achieved when multicore servers and highly threaded operating systems host middleware and applications that are tuned to take advantage of these capabilities. Fujitsu SPARC Enterprise servers offer up to 512 hardware processing threads and four terabytes (4 TB) of memory. Oracle Solaris offers an industry-leading threading model, the result of nearly two decades of innovation. Oracle Solaris offers an exceptionally secure environment, including on-chip encryption capabilities, a robust cryptographic framework, Trusted Extensions, and virtualization capabilities. Finally, a comprehensive development platform enables organizations to create new applications that maximize solution performance while improving reliability.

The Oracle Solaris Ecosystem

Oracle Solaris is a proven, industry-leading operating system with features designed to handle enterprise, business-critical operations. In fact, Oracle Solaris 10 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 systems and compatibility with existing applications is guaranteed. This is why there are more than 50,000 businesses and institutions running more than 11,000 certified applications on Oracle Solaris today.

Powering Fujitsu SPARC Enterprise servers, Oracle Solaris continues to set world records for performance, scalability, and cost-effectiveness. Oracle is investing more in Solaris than Sun did prior to the acquisition, and will continue to develop innovative technologies and enhance Oracle Solaris.

Oracle Solaris includes many unique and innovative technologies that are uncommon to other operating system vendors—including: Oracle Solaris ZFS, Oracle Solaris DTrace, Predictive Self Healing, built-in virtualization, independent security verification, and binary compatibility. As enterprise system hardware often has a service life of 8-10 years or more, it is comforting to understand the commitment of Oracle Solaris to providing a long-lived platform for your software environment.

SPARC Processors

SPARC (Scalable Processor ARChitecture) is a RISC instruction set architecture developed by Sun Microsystems (now Oracle). The “Scalable” in SPARC comes from the fact that the SPARC specification allows implementations to scale from embedded processors up through large server processors, all sharing the same (non-privileged) core instruction set. A single version of Oracle Solaris runs across all Fujitsu SPARC Enterprise servers, including M-Series and T-Series servers. This means datacenters can run a single OS—Oracle Solaris—across all systems, from the smallest to the largest, greatly simplifying administration. Combined with Oracle Solaris, Fujitsu SPARC Enterprise servers provide record-setting performance, extreme scalability, mainframe-class reliability and availability, and strong security.

Table 1 provides an overview of the key features of the SPARC processor architectures.

TABLE 1: KEY FEATURES OF THE SPARC PROCESSOR ARCHITECTURE BY FAMILY

FEATURE	T-Series with UltraSPARC T2/T2 Plus*	M-Series with SPARC64 VII
Cores/Threads/sockets	8 cores/8 threads/4 sockets Up to 256 processing threads Chip Multithreading (CMT)	4 cores/2 threads/64 sockets Up to 512 processing threads Simultaneous Multithreading (SMT)
Maximum frequency	1.6 GHz	2.88 GHz
L2 cache	4MB on chip	6 MB on chip
On-chip support	PCI Express bridge, integrated dual 10GbE networking with XAUI, crypto acceleration, L2 cache	L2 cache
Maximum memory	512 GB	4 TB

Reliability features	Predictive Self Healing, hot-swap components, ECC everywhere, redundant components and networking.	End-to-end ECC protection; guaranteed data path integrity; automatic recovery with instruction retry; total SRAM and register protection; ECC and Extended ECC protection for memory, memory mirroring, and Predictive Self Healing; full hardware redundancy; fault-isolated Hardware Partitioning; dynamic reconfiguration; auto-diagnosis, and recovery; guaranteed data path integrity, total SRAM and register protection.
Security	Multiple on-chip cryptographic capabilities, plus additional protections	Available add-in crypto-accelerator cards
Virtualization Included at no extra charge—third-party products also available	Oracle VM Server for SPARC (previously called Logical Domains or LDOMs) and Oracle Solaris Containers	Hardware Partitioning and Oracle Solaris Containers
Target environments	Network-facing: consolidation and virtualization, Web, Media, security, OLTP, middleware/SOA, batch processing, datamart, application servers	Data-facing. Optimized for 24x7 mission-critical computing: DSS, ERP, CRM, BIDW, large databases, large-scale OLTP, and HPC/scientific/engineering applications, that require mission-critical RAS features.

* UltraSPARC T2 Plus is the multisocket version of the UltraSPARC T2 processor. Up to four UltraSPARC T2 Plus processors can be used in a single server.

As shown in Table 1, the SPARC processor family is designed and optimized for different types of application environments. The same Oracle Solaris provides commonality across both hardware platforms in myriad of applications and different datacenter tiers. The SPARC processor family spans a wide range of enterprise servers to create architectures that are suitable for best efficiency and security, such as with the T-Series, to massive scalability and availability, such as the M-Series. These two platforms create a potent mix of solutions such as CRM systems and Java/Web middleware infrastructure with the T-Series to ERP systems and backend OLTP/DW systems with the M-Series.

The SPARC processors provides a range systems—one to four sockets for T-Series, up to 64 sockets for M-Series—to run critical systems for the business from the edge of the network to deep in the datacenter. Server selection in an architecture is based purely on specific application scenarios and expectations, and servers can be mixed and matched.

A specific recommendation is out of the scope of this paper and we encourage you to understand the SPARC server application scenarios and case studies on www.fujitsu.com/sparcenterprise or by discussing with your Fujitsu representative. The exact sizing and capacity planning can be undertaken with the help of Fujitsu experts.

Fujitsu SPARC Enterprise T-Series Servers with Chip Multithreading (CMT)

The UltraSPARC T2 processor with CoolThreads technology implements the industry's first massively threaded "system on a chip." These processors power the Fujitsu SPARC Enterprise T-Series servers. With support for up to 8 cores/8 threads per core (64 threads per chip)—and up to four sockets—this processor provides breakthrough performance and energy efficiency. In addition, the UltraSPARC T2/T2 Plus processors are the first to integrate 10 Gb Ethernet, PCI Express I/O, and cryptographic acceleration directly onto the processor chip. Combined with Oracle Solaris, this approach provides leading levels of performance and scalability with extremely high levels of efficiency. CMT architecture is ultimately very flexible, and working with Oracle Solaris allows different modular combinations of processors, cores, and integrated components as stated above. The combination offers:

- Increasing computational capabilities to meet the growing demand from Web applications
- Supporting larger and more diverse workloads with greater floating point performance
- Powering faster networking to serve new network-intensive content
- Providing end-to-end datacenter encryption
- Increasing service levels and reducing downtime
- Improving datacenter capacities while reducing costs

Closely orchestrated with Oracle Solaris, these systems provide record-setting performance and excellent RAS characteristics, ideal for maximizing the uptime and ROI of mission-critical enterprise applications. Note that there are additional features that contribute to enhanced reliability, including advanced integration—significantly lower parts component count—and superior energy efficiency that contributes to a reduction of faults due to thermal conditions.

Fujitsu SPARC Enterprise servers running Oracle Solaris are built to achieve high levels of uptime and fast recovery from failures. Administrators can utilize Oracle Solaris commands to remove and replace disks, power supplies, and fan units while the system continues to operate. One PCI Express root complex per processor combined with the ability to configure multiple CPUs, memory FB-DIMMs, and I/O cards add to the resiliency of Fujitsu SPARC Enterprise T5140, T5240, and T5440 servers. Hot-swap and hot-plug chassis-mounted hard drives, fan units, and power supplies improve serviceability and availability.

Fujitsu SPARC Enterprise M-Series Servers with SPARC64 VI and SPARC64 VII

SPARC64™ processors power Fujitsu SPARC Enterprise M-Series servers. Running Oracle Solaris, these platforms offer mainframe-class features and sustainable levels of record-setting application performance. SPARC64 VI processors provide two cores and SPARC64 VII processors provide four cores. There are two strands (threads) per core for both processors. In combination with Oracle Solaris, SPARC64 VII provides simultaneous multithreading (SMT) scalability to support parallel execution of all eight threads across all available processors (from 1–64 processors). SPARC Enterprise servers feature memory subsystems as large as 4 TB, and high-throughput I/O architectures.

Fujitsu SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 servers running Oracle Solaris delivers a mainframe-class system architecture for high availability (HA) running Oracle Solaris 10. Furthermore, the range of compute power offered by these servers provides the levels of vertical scalability required for server consolidation and many other deployment classes. SPARC Enterprise M4000 and M5000 servers fulfill mid-range system requirements, while SPARC Enterprise M8000 and M9000 servers deliver the massive processing power needed for high-end computing.

Many design features of Fujitsu SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 servers work together with Oracle Solaris in contributing to a comprehensive and integrated architectural approach that is designed for high availability of key systems at lower total costs. Mainframe-class RAS features come standard in the Fujitsu SPARC Enterprise M-Series servers, including automatic recovery with instruction retry, up to 4 TB of system memory error-correcting code (ECC) protection with extended ECC support, guaranteed data-path integrity, total static random access memory (SRAM) and register protection, configurable memory mirroring, and many more.

What's more, most major system components are redundant and hot-swappable, for increased availability and serviceability. This includes processors, memory, disk drives, I/O cards, power supplies, and more. The Fujitsu SPARC Enterprise M8000 and M9000 servers add the ability to hot-swap CPUs, memory, and the service processors. These systems are able to recover from most hardware failures with no impact to users or system functionality. In case of almost the all of hardware failures, Fujitsu SPARC Enterprise M4000, M5000, M8000, and M9000 servers can continue operations with minimal impact. They recover quickly from many hardware failures, including serious faults such as the failure of a CPU or a critical ASIC. In fact, no single hardware component failure prohibits Fujitsu SPARC Enterprise M9000 servers from booting. Even if a CMU system board fails, system operation can continue while the failed board is replaced. This feature is available for M8000 and M9000.

Innovative CPU designs help Fujitsu SPARC Enterprise M-series offer better performance than other UNIX servers. At the same time, these servers offer full binary compatibility and complete investment protection for owners of previous generations of Fujitsu SPARC Enterprise systems.

Architected for Reliability

Oracle Solaris is designed for reliability. 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 ease serviceability.

There are many complementary features built into Oracle Solaris, SPARC64 VI/VII and UltraSPARC T2/T2 Plus processors, and M-Series and T-Series servers to promote mainframe-class reliability. On all Fujitsu SPARC Enterprise systems, Oracle Solaris Predictive Self Healing enhances reliability. On Fujitsu SPARC Enterprise M-Series servers, Hardware Partitioning (discussed in the Virtualization section) further improves uptime and availability.

Oracle Solaris Predictive Self Healing

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 10 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 Architecture (FMA) and Oracle Solaris Service Manager Facility (SMF) are the two main components of Predictive Self Healing.

The FMA, 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 FMA diagnosis engine produces a fault diagnosis once discernible patterns are observed from a stream of incoming errors. Following diagnosis, FMA provides fault information to agents that know how to respond to specific faults.

The FMA offers comprehensive reliability and availability capabilities on all Fujitsu SPARC Enterprise systems. For example:

- CPU “offlining” takes cores and threads (strands) deemed faulty offline. They are recorded and remain offline on reboot until the faulty processor has been replaced, at which point they are made available again.
- Memory page retirement retires pages of memory marked as faulty. They are recorded and remain offline on reboot until the faulty memory has been replaced, at which point it is made available again.

In addition, Fujitsu SPARC Enterprise M-Series servers running Oracle Solaris also provide FMA support on their service processors, or eXtended System Control Facility (XSCF). This allows the XSCF to report faults in the system even if there are no Hardware Partitions running. The alerts are in exactly the same format as the reports from FMA running in a partition.

The SMF facility creates a standardized control mechanism for application services by turning them into first-class objects that administrators can observe and manage in a uniform way. These services can then be automatically restarted if they are accidentally terminated by an administrator, if they are aborted as the result of a software programming error, or if they are interrupted by an underlying hardware problem. Specifically, SMF enables administrators to do the following tasks easily and efficiently with Fujitsu SPARC Enterprise servers running Oracle Solaris:

- 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 offers comprehensive reliability and availability capabilities on all Fujitsu SPARC Enterprise systems.

Solaris Memory Page Retirement

As a part of the Oracle Solaris Predictive Self Healing technology framework, the Oracle Solaris memory page retirement (MPR) capability works to isolate memory issues without system interruption. Fault Manager examines hardware on a continual basis, notifying the MPR subsystem of pages in need of retirement. MPR retires memory pages containing correctable errors and relocatable clean pages containing uncorrectable errors without interrupting user applications. In addition, MPR can also isolate relocatable dirty pages containing uncorrectable errors with limited impact on affected user processes and avoids forcing an outage of an entire system. By utilizing MPR on Fujitsu SPARC Enterprise servers, system interruption rates can be reduced by as much as 35-40 percent¹.

Highly Reliable Memory Subsystems

Oracle Solaris and Fujitsu SPARC Enterprise servers work together to ensure the reliability of system memory. Some Fujitsu SPARC Enterprise M-Series servers offer the following:

- **Memory patrol.** Memory patrol periodically scans memory for errors, proactively preventing the use of faulty areas of memory before they can cause system or application errors, improving system reliability.
- **Memory Extended ECC.** The memory Extended ECC function of these servers enables recovery from memory device failures so that processing can continue despite events such as burst read errors. Data is stored in noncontiguous locations and an extra error-correcting code (ECC), sometimes called a parity bit, is also stored to allow the detection of single-bit errors.
- **Memory mirroring.** Memory mirroring on the Fujitsu SPARC Enterprise M4000 to M9000 is an optional, high-availability feature appropriate for execution of applications with the most stringent availability requirements. Memory mirroring duplicates the data on write and compares the data on read to each side of the memory mirror. In the event that errors occur at the bus or dual inline memory module (DIMM) level, normal data processing continues through the other memory bus and alternate DIMM set.

¹ [Assessment of the Effect of Memory Page Retirement on System RAS Against Hardware Faults](#)

Oracle Solaris ZFS for Reliable Data

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

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

Oracle Solaris ZFS protects all data by 256-bit checksums, resulting in 99.9999999999999999-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.

Oracle Solaris 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.

Reliable Networking

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

In addition to traditional support for multiple network interfaces connected to different network subnets, Oracle Solaris running on Fujitsu SPARC Enterprise servers also provides support for redundant network interfaces that are connected to a single subnet. IP Multipathing provides both failover and IP link aggregation. A number of key features of redundant networking that work to improve the availability and performance of Fujitsu SPARC Enterprise servers 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 spread across multiple network adapters to achieve higher throughput. Load spreading occurs only when network traffic is flowing to multiple destinations using multiple connections.

Support for Virtualized Networking and I/O

Oracle Solaris contains technology to support and virtualize components and subsystems on the UltraSPARC T2/T2 Plus processor, including support for the on-chip PCI Express interface and cryptographic processors. As a part of a high-performance network architecture, multithread-hot device drivers are provided so that applications running within virtualization frameworks can effectively share I/O and network devices, maximizing utilization and increasing ROI.

Scalable Performance

Fujitsu SPARC Enterprise servers offer outstanding scalability and performance:

- Since their launch, SPARC64-based systems such as the Fujitsu SPARC Enterprise M9000 server have earned 24 world records and “product firsts.”²
- Since their launch, UltraSPARC-based systems such as the Fujitsu SPARC Enterprise T5440 server have earned over 176 world records and “product firsts.”

Oracle Solaris 10 is specifically designed to optimize the considerable resources of SPARC64 and UltraSPARC T2/T2 Plus processor based systems, and offers impressive scalability, as shown in Table 2.

TABLE 2: ORACLE SOLARIS SCALABILITY

SCALABILITY FEATURE	ORACLE SOLARIS 10
64-bit addressing	Since 1996
128-bit file system	Yes
Maximum CPU threads	512
Maximum RAM	4 TB
Maximum file system size	16 Exabytes
Maximum file size	16 Exabytes

Oracle Solaris has incorporated many features to improve scalability and performance on SPARC-based systems, such as the following.

² blogs.sun.com/BestPerf

Multithread Awareness

Oracle Solaris is optimized for the SPARC64 VI/VII and UltraSPARC T2/T2 Plus 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 chip), 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

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 with somewhat lower latency if the memory bank is physically close to the requesting processor. Fujitsu SPARC Enterprise 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.

Oracle Solaris uses Memory Placement Optimization (MPO) to improve the placement of 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 enough balance within the system. As a result, many database and technical computing applications are able to run considerably faster with MPO.

Oracle Solaris Internals Optimization

For over 20 years, Oracle Solaris internals have been improved to enhance scalability, enabling Fujitsu SPARC Enterprise 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 VI/VII and UltraSPARC T2/T2 Plus processors provide a range of pages sizes up to 256 MB that Oracle Solaris automatically uses in a variety of contexts without application change, including for user and kernel pages, and instruction and data pages.
- **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, the

UltraSPARC T2/T2 Plus and the SPARC64 VI/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.

- **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.

Successful enterprise-class servers efficiently process CPU, memory, and I/O workloads for middleware and databases. Building on a proven track record, Oracle Solaris unlocks the proven performance capabilities of the SPARC64 VI/VII and UltraSPARC T2/T2 Plus 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. Oracle Solaris running on SPARC-based systems are designed for optimized, end-to-end performance, reducing or eliminating bottlenecks in memory and I/O subsystems. This is highlighted in a number of world-record benchmarks, including the following.

World Record Performance

Oracle Database 11g, Oracle Solaris, and Fujitsu SPARC Enterprise servers continue to set world records in performance and affordability. This includes:

- **SPECjAppServer2004 JOPS@Standard.** Oracle Solaris, Oracle WebLogic 10.3.3 Application Server, and Oracle Database 11g Enterprise Edition power five Fujitsu SPARC Enterprise T5440 servers, six Sun Storage F5100 Flash Arrays, and one Fujitsu SPARC Enterprise M9000 server to a world record result of 28,648.74 SPECjAppServer2004 JOPS@Standard on the SPECjAppServer2004 benchmark.³
- **TPC-C.** SPARC Enterprise T5440 is the world's fastest OLTP system (7,646,486.7 tpmC), and achieved with the best \$/tpmC (\$2.81/tpmC) out of all top 10 performers.⁴ TPC-C demonstrates Oracle Solaris and SPARC combine to cost-effectively deliver heavily multithreaded and I/O capabilities on OLTP workloads.

³ blogs.sun.com/BestPerf (all records current as of publication date)

⁴ tpc.org/tpcc/results/tpcc_perf_results.asp

- **TPC-H@3000GB.** Fujitsu SPARC Enterprise M9000 server and Oracle Solaris delivered a single-system TPC-H 3000GB world record performance and price performance results.⁵ The Fujitsu SPARC Enterprise M9000 server, running Oracle Database 11g Release 2 proves the power of the Oracle solution.
- **Oracle Business Intelligence Enterprise Edition.** SPARC Enterprise T5440 systems, Oracle 11g Database, Oracle Solaris, Oracle Solaris Containers, and Oracle Solaris ZFS set world records in supporting 50,000, 28,000, and 10,000 concurrent users.
- **SPECweb2005.** Oracle has obtained a world record SPECweb2005 performance result of 100,209 SPECweb2005 with Oracle Solaris powering the Fujitsu SPARC Enterprise T5440, running Java System Web Server 7.0 Update 5, and Java Hotspot Server VM. This result demonstrates performance leadership of the Fujitsu SPARC Enterprise T5440 server and its scalability, by using Oracle Solaris Containers to consolidate multiple Web serving environments, and Sun Open Storage Flash technology to store large datasets for fast data retrieval.⁶
- **PeopleSoft Payroll (North America) 9.0 benchmark.** Oracle Solaris running on the Fujitsu SPARC Enterprise M4000 server with four 2.53GHz SPARC64 VII processors and the Sun Storage F5100 flash array on the PeopleSoft Payroll (NA) 9.0 benchmark with Oracle 11g Database.⁷ The Fujitsu SPARC Enterprise M4000 server combined with Oracle FlashFire technology demonstrated a speedup of 81 percent going from 1 to 8 streams on the PeopleSoft Payroll (NA) 9.0 benchmark using the Oracle 11g Database.
- **SAP SD 2-tier.** Oracle delivers world-record leadership, including:
 - SPARC Enterprise T5440 system and Oracle Solaris set 4-CPU world record—4,720 SD users—for SAP ERP 6.0 application Enhancement Package 4 (Unicode).⁸
 - Oracle Solaris and the Fujitsu SPARC Enterprise M9000 server with 2.88 GHz SPARC64 VII processors achieved 32,000 users.⁹
- **SPECjAppServer2004.** Single server world record using Oracle Solaris and Fujitsu SPARC Enterprise T5440 server.

⁵ blogs.sun.com/BestPerf/tags/tpc-c

⁶ blogs.sun.com/BestPerf/tags/tpc-c

⁹ blogs.sun.com/BestPerf/entry/oracle_peoplesoft_payroll_sun_sparc

⁸ blogs.sun.com/BestPerf/entry/sun_solaris_leadership_in_sap

⁹ blogs.sun.com/BestPerf/entry/sun_m9000_fastest_sap_2

TABLE 3: SPARC64 VI AND SPARC64 VII WORLD RECORDS

SYSTEM	WORLD RECORD PERFORMANCE ¹⁰
SPARC Enterprise M9000 (VII)	Linpack HPC: M9000 achieved 2.023 TFLOPS; Nearly 2X the best published result of IBM P6 595 @ 5GHz; beats HP Integrity Superdome result by 2.7X
SPARC Enterprise M9000 (VI)	Achieved Guinness World Record by powering the world's largest data warehouse - oracle.com/us/products/database/exadata/index.html
SPARC Enterprise M9000 (VII)	SPECjbb2005: M9000 achieved World Record single-instance JVM result of 1,757,035 SPECjbb2005 bops/JVM on the benchmark.

Network Performance

Oracle Solaris running on Fujitsu SPARC Enterprise servers introduces a new and highly scalable TCP/IP stack that significantly increases network throughput and capacity. This innovative stack speeds packet processing by reducing overhead when processing packets. The advanced design improves the performance of many networked applications by approximately 50 percent—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 NICs. Oracle Solaris 10 TCP/IP stack is tuned for 10 Gigabit Ethernet, and hardware offloading technologies.

Security

Oracle Solaris provides a sophisticated network-wide security system that controls the way users access files, protect system databases, and use system resources. 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. For example, it is optimized to work with the built-in security and encryption capabilities of the Fujitsu SPARC Enterprise T-Series servers, as outlined below. 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 associated transactions.

- Oracle Solaris 10 provides security features previously only found in Trusted Solaris OS. It delivers a secure environment right out of the box, and can be further hardened and minimized as needed, helping to reduce the risk that a system or application can be compromised.

¹⁰ For the latest information: oracle.com/us/solutions/performance-scalability/

- Oracle Solaris 10 offers RBAC, Process Rights Management, and least privilege. These technologies reduce security risk by granting users and applications only the minimum capabilities needed to perform tasks. System administrators can grant—or deny—a large number of discrete privileges to any process on the system to create effective security policies, minimize the likelihood of hostile actions, control access to data, and ensure compliance with regulatory requirements.
- As an optional layer of secure label technology in Oracle Solaris 10, Oracle Solaris Trusted Extensions allow data security policies to be separated from data ownership.

Integrated with Fujitsu SPARC Enterprise T-Series Servers

The Fujitsu SPARC Enterprise T-Series and M-Series service processors are also secure out of the box. All data services must explicitly be enabled, with only the serial port enabled by default. The service processors also implement a form of RBAC for account security. Users can be assigned specific roles with restricted access and capability as the business dictates. On the M-Series, this control can also be done on individual Hardware Partitions. This has the same effect as the Oracle Solaris 10 feature least privilege (discussed later).

The UltraSPARC T2 and T2 Plus processors were designed with a dedicated, integrated cryptographic accelerator unit for each of the eight cores. Integrated cryptographic acceleration means Oracle Solaris applications can run securely without the extra cost of a separate cryptographic processor, and without the high performance penalty previously associated with secure operation. Fujitsu's integrated cryptographic units support the ten most common ciphers and secure hashing functions, including NSA-approved algorithms. And, they outperform competing accelerators by more than 10x, with minimal performance impact.

The latest UltraSPARC T2 and T2 Plus processors extend algorithm support by introducing symmetric key-based encryption and decryption mechanisms, such as Data Encryption Standard (DES), Triple DES (3DES), Advanced Encryption Standards (AES-128, AES-192, and AES-256), RC4, as well as hashing operations such as Message Digest 5 (MD5) algorithm, SHA1, SHA256, and Elliptic Curve Cryptography (ECC) mechanisms, such as the ECCp-160 and ECCb-163 algorithms. An on-chip Random Number Generator supports random number generation operations intended for cryptographic applications.

RSA operation is an important component of the Secure Sockets Layer/Transport Layer Security (SSL/TLS) full handshake. Each core of the UltraSPARC T1, T2, and T2 Plus processors includes a Modular Arithmetic Unit (MAU) that supports RSA and Digital Signature Algorithm (DSA) operations. RSA operations utilize a compute-intensive algorithm that can be off-loaded to the MAU. Indeed, the MAU is capable of sustaining more than 30,000 RSA-1024 operations per second on systems with an UltraSPARC T2/T2 Plus processor. Moving RSA operations to the MAU speeds SSL/TLS full handshake performance and frees the CPU to handle other computations.

The Oracle Solaris Cryptographic Framework Library

The Oracle Solaris Cryptographic Framework (Figure 1) provides cryptographic services to applications and kernel modules in a manner seamless to the end user, and 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.

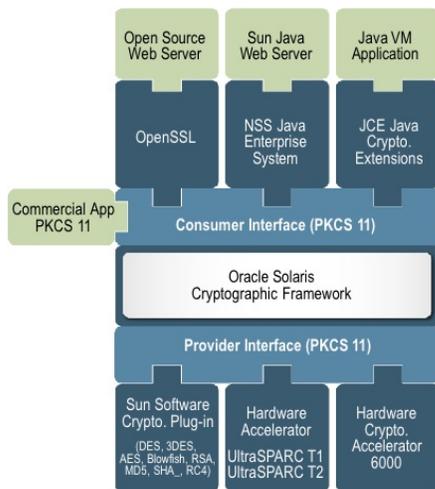


Figure 1: Oracle Solaris Cryptographic Framework is standardized and extensible—current and future cryptographic choices can easily plug in and take advantage of hardware and software capabilities.

Oracle Solaris Cryptographic Framework provides cryptographic services for kernel-level and user-level consumers, as well as several software encryption modules. Oracle Solaris Cryptographic Framework continues to include Kernel SSL proxy (KSSL), which off-loads SSL processing from user applications and enables them to transparently take advantage of hardware accelerators, such as those available in UltraSPARC T2/T2 Plus processors.

The Oracle Solaris Cryptographic Framework provides the PKCS#11 industry standard. It is accessible to Java applications on Oracle Solaris as the default Java Cryptographic Extension (JCE) provider. For OpenSSL applications a “pkcs11” OpenSSL ENGINE is available for them to offload cryptographic algorithms to the Oracle Solaris Cryptographic Framework. Applications using the Mozilla Network Security Services (NSS) API can be configured to use the crypto framework via PKCS#11. It provides cryptographic services to users and applications through commands, a user-level programming interface, a kernel programming interface, and user-level and kernel-level frameworks.

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.

For applications that utilize any of the above-mentioned cryptographic APIs, performance of cryptographic routines is automatically improved without recompilation. For applications that use a private cryptographic library, recompilation, or linking to one of these APIs will ensure that full hardware acceleration of cryptographic routines is achieved. 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 the UltraSPARC T2/T2 Plus processors.

Preventing Attacks

Oracle Solaris also takes advantage of the UltraSPARC's capability 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 UltraSPARC chipset 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 10 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 10 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.

Fujitsu Server Virtualization

Enterprise users need choice for server virtualization and consolidation. They need flexibility with respect to application, OS, and network virtualization methods. Fujitsu SPARC Enterprise servers offer comprehensive virtualization to address the needs of enterprise computing customers. They are the leading platform to have Hardware Partitioning capabilities, which provide the physical isolation needed to run independent operating systems.

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, the cost-savings benefit of virtualization make it a 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 the space, power, and cooling crunch. Organizations are increasingly using virtualization to increase business agility, which increase speed and flexibility in delivering IT services to support business goals.

Oracle Solaris 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 Fujitsu SPARC Enterprise servers—there are no additional costs to use them.

- Hardware Partitioning is provided on selected Fujitsu SPARC Enterprise servers, allowing physical hot-swap of components in the system without shutting down services. Hybrid virtualization is achieved by combining Hardware Partitioning and Solaris Containers.
- Oracle VM Server for SPARC offers a hybrid of partitioning and virtualization fully exploiting the unique advantages of CMT technology to provide a more optimized virtual machine environment and still providing all the advantages of the SPARC hardware platform and Oracle Solaris, including full binary compatibility.
- Oracle Solaris Containers provide security and resource isolation that allows multiple virtual Oracle Solaris environments to share the same OS instance, increasing security and utilization on all Fujitsu SPARC Enterprise servers by complementing the capabilities of Oracle VM Server for SPARC and Hardware Partitioning.

All Fujitsu's virtualization such as Hardware Partitioning, Oracle VM Server for SPARC, and Oracle Solaris Containers are multithreaded to maximize performance and utilization.

Oracle VM Server for SPARC

Oracle VM Server for SPARC, previously called 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 OS instances 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. Fujitsu 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 T1, T2, and T2 Plus processors and Oracle Solaris. This combination helps to increase flexibility, isolate workload processing, and improve the potential for maximum server utilization. To facilitate agile datacenters, Oracle VM Server for SPARC domains can be migrated between physical servers, and system resources such as CPUs, virtual I/O devices, memory, and cryptographic units can be dynamically reconfigured.

Fujitsu SPARC Enterprise servers running Oracle Solaris are the leading platform with the Hardware 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 customers the flexibility to deploy multiple operating systems simultaneously on a single Fujitsu 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 Solaris scheduler—to dispatch workloads to virtual CPUs, which are effectively physical CPU threads. The end result is a virtualization solution with “bare-metal” performance—lower overhead, and higher performance and scalability.

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

Oracle Solaris Containers

Supported on any Fujitsu SPARC Enterprise server running Oracle Solaris 10, 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.

Unlike virtual machines, Oracle Solaris Containers provide OS-level virtualization by giving 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—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 or Hardware Partition where the Oracle Solaris Container resides.

Hardware Partitioning and Dynamic Reconfiguration

A key feature of the Fujitsu SPARC Enterprise M-Series high-end servers with Oracle Solaris is the ability to partition the available hardware resources into smaller logical systems. Fujitsu SPARC Enterprise M-Series servers offer *hard partitioning* technology in the form of Hardware Partitioning. Instantiating a number of Hardware Partitions on a Fujitsu SPARC Enterprise M-Series server divides the system into multiple electrically isolated partitions. Each Hardware Partition 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 Hardware Partition do not impact applications running in other Hardware Partitions. SPARC Enterprise M-Series servers can provide up to 24 Hardware Partitions, each with configurable amounts of CPU, memory, disk, and I/O resources such as PCI Express and PCI-X slots, and networking.

Hardware Partitioning can be used with Solaris Containers to refine resource control and simplify the consolidation of several applications into one partition. As described previously, the Oracle Solaris Containers functionality in Oracle Solaris 10 enables multiple, software-isolated applications to run on a single server or Hardware Partition.

Dynamic Reconfiguration allows resources to be dynamically reallocated, or balanced, between Hardware Partitions. Utilizing this technology enables a physical or logical restructuring of the hardware components of Fujitsu SPARC Enterprise M-Series servers while the system is running and the applications remain available. This high degree of resource flexibility allows the partition or platform administrator to reconfigure the system easily in order to provision the resources to meet changing workload demands. Disaster recovery can also be used to remove and replace failed or upgraded hardware components while the system is online¹¹. CPU, memory, and I/O devices be added or deleted by Dynamic Reconfiguration.

¹¹ Fujitsu SPARC Enterprise M4000, M5000, M8000, and M9000 servers can perform Dynamic Reconfiguration to logically move system resources between domains. In addition, Fujitsu SPARC

The Reconfiguration Coordination Manager (RCM) is the framework that manages the dynamic removal of system components. By using RCM, administrators 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 Hardware Partition, enabling the SGA to be dynamically scaled back to allow the board to be removed without shutting down the database.

Conclusion

Oracle Solaris running on Fujitsu SPARC Enterprise servers has continued to demonstrate great success as a mission-critical, enterprise-class OS for scalable performance, advanced reliability, and virtualization, especially when deploying enterprise applications in the datacenter. The combination of Oracle Solaris on innovative Fujitsu SPARC Enterprise servers offers the IT infrastructure required for enterprises that need a reliable, available, and secure solution.

Oracle Solaris is leveraging more than 20 years of SMP expertise for proven performance in very large multicore processing environments. Developers and system administrators alike can use Oracle Solaris running on systems designed with the Fujitsu SPARC Enterprise systems for improved performance, reliability and throughput.

Whether serving enterprise applications, high-performance computing applications, or consolidating multiple lower-powered servers, IT systems must scale smoothly and intelligently, provide rock-solid security, and virtually nonstop reliability. Fujitsu SPARC Enterprise servers and Oracle Solaris are both widely recognized as the technologies of choice for enterprise and 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.

Enterprise M8000 and M9000 servers can perform hot-swap operations to physically add or remove boards from the chassis.

Resources

The following tables contain links to useful information related to this paper.

GET THE PRODUCTS

Oracle Solaris	oracle.com/solaris
Fujitsu SPARC Enterprise servers	www.fujitsu.com/sparcenterprise
Oracle Enterprise Manager Ops Center	oracle.com/us/products/enterprise-manager/opscenter/index.html

DEEP DIVE ON THE TECHNICAL

Oracle Technical Network	http://oracle.com/otn
--------------------------	---

AVAILABILITY

Oracle Solaris ZFS	http://www.sun.com/bigadmin/topics/zfs/
--------------------	---

SECURITY

Taking Advantage of Wire-speed Cryptography in Oracle WebLogic Server 10.3.x and Java Platform, Enterprise Edition 5 Application Environments	http://wikis.sun.com/download/attachments/188029349/821-0801.pdf?version=1&modificationDate=1259860946000
Using the Cryptographic Accelerators in the UltraSPARC T1 and T2 Processors	http://www.sun.com/offers/details/819-5782.html
The Least Privilege Model in the Solaris 10 OS	http://www.sun.com/bigadmin/features/articles/least_privilege.jsp
Oracle Solaris Trusted Extensions	http://www.sun.com/software/solaris/ds/trusted_extensions.jsp
Oracle Solaris Security FAQ	http://www.oracle.com/technetwork/server-storage/solaris/faq-sol-sec-159378.pdf

VIRTUALIZATION

Oracle Virtualization	http://www.oracle.com/virtualization
Oracle VM Server for SPARC—Enabling a Flexible, Efficient IT Infrastructure	http://www.oracle.com/us/products/servers-storage/servers/sparc-enterprise/vm-server-for-sparc-wp-075964.pdf
LDoms IO Best Practices - Storage Availability with Logical Domains	http://wikis.sun.com/display/BluePrints/LDoms+IO+Best+Practices++Storage+Availability+with+Logical+Domains
Solaris Containers (Zones)	http://www.sun.com/bigadmin/content/zones/index.jsp

DEVELOPER TOOLS	
Oracle Solaris Studio	http://developers.sun.com/sunstudio
DLight Tutorial for Sun Studio	http://developers.sun.com/sunstudio/documentation/tutorials/dlight
BigAdmin Portal—DTrace	www.sun.com/bigadmin/content/dtrace/index.jsp
Oracle Solaris DTrace	http://sun.com/bigadmin/content/dtrace

OTHER	
Using the Solaris 10 Applications Library to Find Apps That Run on the Solaris 10 OS	http://www.sun.com/bigadmin/features/techtips/solaris10appslib.jsp