ORACLE

SOLARIS

An Oracle White Paper
September 2010

Create a More Efficient Storage Infrastructure with the Oracle® Solaris 10 Operating System

Executive Overview

Organizations in every industry rely on information stored in the datacenter to maintain a competitive stance and achieve corporate objectives. Whether using enterprise applications to manage people and processes, bring new products to market, or make discoveries faster, companies depend on their computing infrastructure to deliver data on demand. While many operating systems provide data management features, deriving the greatest value from information requires a platform designed for all of the pieces to work together.

Oracle is the only company that delivers an integrated and optimized technology stack—from applications and operating environments, to virtualization and management tools, and storage and networking solutions. At the heart of this platform is the Oracle® Solaris 10 operating system, the premier platform for enterprise computing, providing a robust foundation and built-in innovation for data management solutions. This technical white paper explains how upgrading to Oracle Solaris 10 can help IT organizations simplify administration and make better use of storage and data management infrastructure.

Organizations—from small businesses and government agencies, to global enterprises, research organizations, and universities—depend on information. As data volumes continue to rise, datacenter managers must cope with expanding storage infrastructure while providing around-the-clock access to data stored on reliable and secure media. In addition, stringent compliance regulations are forcing organizations to retain data for longer time frames. These issues are compounded by service-level demands that require greater operational efficiency, and economic pressures that necessitate cost reductions. The key to success is finding the right platform for storage infrastructure deployment. Now datacenters can take advantage of powerful features offered only by the Oracle Solaris 10 operating system to create intelligent, feature-rich storage servers that can ease infrastructure complexity and reduce operational costs.

The Need for Simplified Data Management and Administration

Organizations looking to consolidate systems or reduce management complexity have historically had to rely on third-party solutions that deliver file services. Yet exponential growth in the amount and type of data is changing organizational needs for file services—needs that these third-party solutions often fail to address. For example, companies are looking to optimize performance and capacity utilization based on the type, size, or read/write access requirements for files. Intelligent features, such as virtualization, automatic file migration, and policy-based recovery, can change the way cost-effective storage management is done.

Enterprises need more than basic file systems and data copy functions: a wider array of services is needed to meet increasingly complex needs. Because the operating system occupies a critical position in the computing system hierarchy—below applications and above hardware—it can integrate data services directly and provide an ideal platform for deploying network-attached storage or storage servers.

The Best Enterprise Operating System on the Planet

The Oracle Solaris operating system provides some of the most robust, reliable, and innovative enterprise operating systems in IT. With a single source base for SPARC® processor-based and x64 systems, Oracle Solaris combines key computing elements—operating system, networking, data and storage management, and user environment—into a stable, high-quality foundation that organizations can depend on to deliver infrastructure solutions. Incorporating innovative features such as virtualization technology, broad hardware applicability, rich security capabilities, debugging and analysis tools and more, Oracle Solaris is designed to deliver the power, flexibility, and availability necessary for small-scale to enterprise-wide computing.

Oracle Solaris 10 includes a host of new features—better internal communication, more complete networking, built-in file and data services with data integrity, simplified data management, and integrated storage pooling and volume management—to deliver an ideal platform for developing and deploying new storage solutions faster.

Making Data Accessible

Whether supporting enterprise users, developers, or educational institutions, storage infrastructure must be able to give users access to growing volumes of data. Oracle Solaris 10 is a major evolution in the Oracle Solaris family, incorporating several technologies aimed at simplifying data management and user access.

A New and Easier Way to Manage Data

With greater amounts of information needing to be stored and accessed, robust data management is key. Built into the operating system is Oracle Solaris ZFS, an enterprise-class, general-purpose file system designed to ease data and storage management complexity. Unlike file systems that require a separate volume manager, Oracle Solaris ZFS integrates traditional file system functionality with built-in volume management techniques and data services. Breaking free of the typical one-to-one mapping between file system and associated volumes, Oracle Solaris ZFS shares space dynamically between multiple file systems from a single storage pool to enable more efficient use of storage devices.

By automatically allocating space from pooled storage when needed, Oracle Solaris ZFS simplifies storage management and gives organizations the flexibility to optimize data placement. Many administrative tasks are automated to speed performance and eliminate common errors. For example, creating file systems is fast and easy. There is no need to configure, or reconfigure, underlying storage devices or volumes—these tasks are handled automatically when devices are added to a storage pool.

Windows Interoperability

Whether data is located on storage systems connected directly to a server, across a local or remote enterprise network, or on an external Web site, one or more protocols likely is in use to make data available to users. With data spread across a wide variety of devices located in storage tiers managed by many different computing platforms, a variety of storage protocols are needed to access and share data. To help this effort, Oracle Solaris 10 supports block- and file-level protocols, including Internet SCSI (iSCSI), Network File System (NFS), and SAMBA.

Network Access to Data

Oracle Solaris 10 delivers open, scalable, general-purpose file sharing with support for NFS. Enabling the transparent sharing of data and programs between heterogeneous systems, NFS allows access to files without regard to their physical location. In addition, a server can delegate the management of a file to a client and reduce the number of round-trip operations required. Operation compounding, which allows multiple operations to be combined into a single over-the-wire request, helps improve performance.

Oracle's implementation of NFS V4 is fully integrated with Kerberos V5, providing authentication, integrity, and privacy. As a result, servers can offer different security options for different file systems. In addition, file access, file locking, and mount protocols are integrated into a single, unified protocol to ease traversal through a firewall and to improve security.

File and Print Services

In today's heterogeneous environments, storage servers must be able to provide services to a wide variety of client systems. Oracle Solaris 10 includes SAMBA, an open-source software suite that provides seamless file and print services to Server Message Block (SMB) and CIFS clients. With SAMBA, the server appears in the Microsoft Windows network, giving clients easy access to file and print services running on Oracle platforms.

Virtualizing Systems and Storage

Innovative storage technologies and supporting services are bringing new economies of scale to enterprises. Opportunities exist to gain efficiencies by consolidating the datacenter storage infrastructure. Oracle Solaris 10 includes technologies that can help consolidate systems and virtualize storage resources to make more applications, storage capacity, and data available to users at less cost.

Oracle Solaris Containers

Oracle Solaris Containers are a set of technologies that help system administrators increase resource utilization by consolidating multiple applications onto a single system. Administrators can specify the percentage of physical system resources each application receives, and isolate each application in its own virtual environment with its own hostname, IP address(es), users, file system, and more. By providing isolation between software applications or services using flexible, software-defined boundaries, Oracle Solaris Containers create an execution environment within a single instance of the operating system.

Key features of Oracle Solaris Containers include

- Full resource containment and control for more predictable service levels
- Software fault isolation to minimize fault propagation and unplanned downtime
- Security isolation to prevent unauthorized access and unintentional intrusions

Storage Pools

Unlike most operating systems, which leave volume management to add-on tools, Oracle Solaris 10 integrates volume management functions. Breaking free of the typical one-to-one mapping between the file system and its associated volumes, Oracle Solaris ZFS introduces the storage pool model. Oracle Solaris ZFS decouples the file system from physical storage in the same way that virtual memory abstracts the address space from physical memory, allowing for more efficient use of storage devices. Space is shared dynamically between multiple file systems from a single storage pool, and is parceled out of the pool as file systems request it. Physical storage can be added to storage pools dynamically, without interrupting services, providing new levels of flexibility, availability, scalability, and performance. When capacity is no longer required by a file system in the pool, it is made available to other file systems.

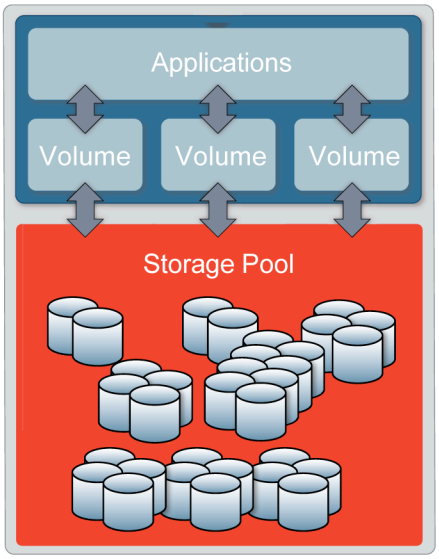

Figure 1. Virtual storage pools let multiple file systems share storage space.

Supporting a Wide Range of Storage Devices

Organizations of every kind are taking advantage of the wide range of storage devices now available to meet capacity and performance demands—from smaller, low-cost systems to high-performance, high-capacity devices, and everything in between. Incorporating so many device types into a storage infrastructure requires an intelligent platform with the facilities to simplify device integration. Oracle Solaris 10 supports a variety of protocols and interface technologies and provides key host bus adapter drivers with the software distribution.

- **Internet SCSI (iSCSI) target support.** Many storage deployments rely on the iSCSI protocol to move data across intranets and manage devices remotely. With iSCSI target support, the operating system can make SCSI storage devices available to clients over the network.
- **Fibre Channel.** Oracle Solaris 10 includes a number of Fibre Channel packages, including libraries based on the T11 FC-HBA specification, Emulex and QLogic device drivers, debugging aids, a FCIP IP/ARP over Fibre Channel device driver, a Fibre Channel transport layer, and much more.
- **Host bus adapter drivers.** A wide range of drivers are included in the operating system, including the Adaptec AdvanceRaid Controller SCSI HBA, Adaptec Ultra320, Advanced Host Controller Interface SATA, LSI MegaRAID SCSI HBA, LSI Hardware RAID HBA, LSI MegaSAS RAID Controller HBA, and Marvell 88SX SATA device drivers.
- **Multiple protocol support.** Several protocols are supported by the operating system, including the Serial Management Protocol for Serial Attached SCSI (SAS) that provides a way to communicate with SAS expanders through the Serial Management Protocol (SMP). A SMP target driver lets users issue SMP requests and receive SMP responses through an interface, giving privileged users the ability to configure and manage SAS domains.

Simplifying Storage Management

With so many devices—and types of devices—at work, finding ways to locate, manage, and make available storage solutions and the data they contain is paramount. Oracle Solaris 10 includes tools and technologies that make it easier to find, integrate, manage, and work with storage solutions.

Automatic Device Discovery and Configuration Management

Locating devices on the network and integrating them into the storage infrastructure can be a time-consuming task. Oracle Solaris 10 supports the Internet Storage Name Service (iSNS) protocol, an industry standard for automated discovery, management, and configuration of iSCSI and Fibre Channel devices on a TCP/IP network. The iSNS protocol helps administrators identify, connect to, and manage devices by supporting name registration and discovery, discovery domains and logon control, state change notification, and entity status inquiries. As a result, administrators can obtain information dynamically about accessible devices, manage devices with familiar network management tools, enhance security, and adapt to topology changes.

Utilities and Application Programming Interfaces

Oracle Solaris 10 includes several utilities and application programming interfaces that can ease the development and deployment of storage servers, including:

- Oracle Solaris Web-Based Enterprise Management (WBEM) services and an associated developer toolkit enable IT developers and managers to create and modify system information stored in the standard Common Information Model (CIM) format. Applications can be created based on the CIM schema and XML and HTTP communication standards.
- Several management utilities and Storage Networking Industry Association (SNIA) compliant management libraries ease device management. In addition, the operating system includes the iSCSI Management API, an object-based framework that facilitates initiator-based management of iSCSI storage systems with support for login and authentication configuration, statistics collection, and more.
- Support for SNIA's Multipath Management API (MP API) is built into Oracle Solaris 10, including a MP API common library, a plug-in for native OS multipathing solutions, and the `mpathadm` command line interface for multipathing discovery and management.

Delivering I/O Performance

As demand rises and services grow in complexity, datacenter infrastructure must provide massive capacity and fast access applications and data in order to keep pace with business priorities. Companies can add storage capacity easily by augmenting infrastructure with additional hard disk drives and arrays. Unfortunately, the devices that provide the highest capacity are expensive, and fail to provide the I/O performance needed to keep systems supplied with the data for processing.

Now Flash technology is emerging as a strong storage alternative that can help rebalance system and storage I/O performance. Robust data integrity, reliability, and availability, combined with breakthrough performance and power characteristics, make it possible to create a new class of storage platform that combines the strengths of hard disk drive technology with enterprise solid-state disks (SSDs) based on Flash technology. For example, enterprise SSDs can be placed in a new storage tier to assist hard disk drives by holding frequently accessed data to minimize the impact of disk latencies and improve application performance. By using enterprise SSDs to handle certain types of I/O, and hard disk drives to store massive data sets, a Hybrid Storage Pool delivers significant performance gains without sacrificing capacity.

Hybrid Storage Pool technology is designed to exceed the performance of Fibre Channel technologies without the additional management and administration complexity of a SAN. Several Oracle Solaris ZFS components are key to Hybrid Storage Pool operation and help accelerate performance.

- The Oracle Solaris ZFS Adaptive Replacement Cache (ARC) is the main file system memory cache and is stored in DRAM.
- The (Level Two Adaptive Replacement Cache (L2ARC) extends the ARC into read-optimized SSDs to provide a large read cache to accelerate reads. The Oracle Solaris ZFS Intent Log (ZIL) is transactional and uses write-based SSDs to provide a large cache to accelerate writes.
- The disk storage pool consists of conventional disk drives. Note that high-performance, more costly disk drives are no longer strictly required to achieve high performance levels given the interposition of SSDs in a Hybrid Storage Pool.

Sophisticated file system algorithms in Oracle Solaris ZFS use the ARC in memory and the L2ARC on SSD to determine pre-fetch or data placement during sustained read operations. SSDs accelerate write throughput for Oracle Solaris ZFS synchronous write I/O operations, helping to boost write performance.

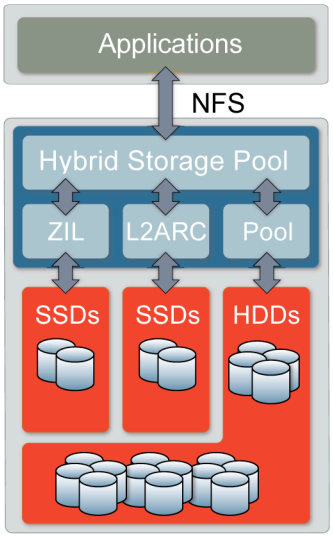

Figure 2. Hybrid Storage Pools optimize data placement to improve I/O performance

Other Oracle Solaris Performance Features

Oracle Solaris 10 includes a wide range of features aimed at improving infrastructure performance.

- **Intelligent algorithms.** Oracle Solaris ZFS simplifies the code paths from the application to the hardware, delivering sustained throughput. Block allocation algorithms accelerate write operations, and consolidate many small random writes into a single, more efficient sequential operation. An I/O scheduler bundles disk I/O to optimize arm movement and sector allocation to speed throughput. In addition, an intelligent prefetch performs read ahead for sequential data streaming, and can adapt its read behavior on the fly for more complex access patterns. Furthermore, data is striped automatically across all available storage devices to balance I/O and maximize throughput. Oracle Solaris ZFS immediately begins to allocate blocks from devices as soon as they are added to the storage pool, increasing effective bandwidth as each device is added to the system.
- **Better storage space utilization and increased data throughput.** Today, many companies take advantage of Redundant Arrays of Inexpensive Disks, or RAID systems, to achieve varied levels of performance and availability. With Oracle Solaris 10, the benefits traditionally associated with expensive hardware-based RAID implementations can be achieved with built-in software RAID features. The intelligent RAID-Z implementation automatically spans data across disks to deliver higher transfer rates, and uses parity, striping, and atomic operations to ensure the reconstruction of corrupted data. The result: higher transfer rates, greater on-line storage capacity, increased data availability and system reliability, easier management of large amounts of data, and reduced maintenance and downtime.
- **Built-in compression techniques.** Oracle Solaris 10 supports multiple compression algorithms, including lzjb, gzip, and gzip-N, to reduce the amount of disk space used to hold files and increase data throughput. Compression is performed on a per-block basis to help improve storage device utilization.

Keeping Systems and Data Secure

In today's hyper-connected economy, the ability to safeguard information and protect against security breaches is essential. Oracle Solaris 10 helps keep systems and data secure with built-in consistency checks, encryption capabilities, access controls, and labeling.

Data Integrity Measures

Oracle Solaris ZFS uses several techniques, such as copy-on-write and end-to-end checksumming, to keep on-disk data self-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 help reconstruct corrupted data.

Trusted Extensions

Organizations needing even higher levels of security can take advantage of Oracle Solaris 10's built-in Trusted Extensions, an optional layer of secure labeling technology that allows data security policies to be separated from data ownership in environments. By enabling Oracle Solaris 10 to support multilevel data access policies, Trusted Extensions can help organizations meet strict government regulatory compliance goals without modifying existing applications for underlying hardware platforms. It provides a platform for deploying high security desktops, database servers, firewalls, and communication gateways, as well as any application where access to sensitive information or networks must be strictly controlled. Key features include:

- Labeled security capabilities, enabling the implementation of strong Mandatory Access Control (MAC) security policies
- Labeled file system, providing the ability to store files on different parts of the disk based upon their security classification
- Labeled networking, enabling data to be exchanged with other multi-leveled (labeled) systems, and the creation of services that respond uniquely to a client based upon the client's classification level
- Labeled printing, providing the ability to assign a range of security classifications to a printer, thereby limiting what files can be sent to the printer based upon the file security label
- Labeled desktop, enabling the graphical user interface (GUI) to enforce and display data classifications

Secure Communications

Oracle Solaris 10 uses the Internet Protocol Security Architecture (IPSec), a suite of security protocols that secures communication channels and ensures that only authorized parties can communicate on

them. Shared-secret encryption is supported, with 128-bit Message Digest 5 (MD5) and Secure Hashing Algorithm 1 (SHA-1) algorithms available for datagram authentication and integrity, and 56-bit Data Encryption Standard (DES) and 168-bit Triple DES algorithms available for payload encryption. In addition, Oracle Solaris 10 includes APIs that enable application-level specification of IPSec policies—giving application developers the ability to dictate security policies independently of system administrators.

Understanding How Systems and Applications are Performing

In order to achieve the highest possible performance levels, organizations need to understand how well applications, systems, and storage are performing. Oracle Solaris 10 includes several tools for assessing operational performance and pinpointing opportunities for improvement.

FileBench

FileBench is a framework of file system workloads for measuring and comparing file system performance. This benchmarking framework simulates the effect of applications on file systems. Application workloads are configured as definitions in the FileBench "`f`" scripting language. Simulations run as several multithreaded processes. Combined with the use of interprocess communication primitives such as POSIX semaphores, and "`f`" language constructs to manipulate the I/O environment such as `O_SYNC`, these features enable the emulation of complex relational database applications and database benchmarks such as TPC1 and Web applications. Integrated statistics provide for microsecond-accurate latency and cycle counts per system call.

Oracle Solaris Dynamic Tracing Facility

The Oracle Solaris Dynamic Tracing (DTrace) facility is a dynamic tracing framework that provides top-to-bottom system observability for troubleshooting systemic problems in real time. Designed to quickly identify the root cause of system performance problems, DTrace combines over 100,000 trace points with a powerful scripting language and a simple, interactive command-line interface. It works by safely and dynamically instrumenting the running operating system kernel and applications with trace points (known as *probes*) that are completely passive until enabled. Probes can be enabled quickly for data collection, and then disabled again to minimize performance impacts on the system being examined. Developers and administrators can use this information to quickly identify performance bottlenecks, optimize resource utilization and performance, and quantify resource requirements.

Easing System Administration

Whether moving to Oracle Solaris 10 for the first time, or upgrading systems from an older version, IT organizations can take advantage of features built into the operating system to ease migration and data and system administration. For example, Oracle Solaris 10 includes snapshot capabilities—the ability to create a read-only copy of an Oracle Solaris ZFS file system or volume and restore it at a later time, if needed. Snapshots can be created almost instantly, and initially consume no additional disk space within the storage pool. Replicated streams of descendant file systems can be sent to named snapshots, preserving properties, snapshots, file systems, and clones. With snapshots, administrators can save the state of a file system at a particular point in time, and recreate it on another machine.

Another administrative headache is managing multiple file system types. Oracle Solaris ZFS now can be used as the root file system. With support for bootability, built-in redundancy capabilities, and on-disk consistency, Oracle Solaris ZFS brings greater reliability, availability, and data integrity to root and boot environments. Space for the root file system and swap and dump areas resides in a storage pool, enabling them to grow or shrink as needed. Administrators no longer need to pre-allocate space for these key areas on separate physical disk partitions.

Migrating from a UFS root file system is straightforward. Administrators upgrade systems to the latest version of Oracle Solaris and use the Oracle Solaris Live Upgrade feature to migrate an existing UFS root file system to an Oracle Solaris ZFS root file system. Oracle Solaris Live Upgrade also can be used to patch or upgrade an alternate ZFS boot environment. If an error occurs, administrators can boot back to the original environment.

For More Information

Oracle is a leading provider of innovative storage technologies—from Oracle Solaris 10 to intelligent storage appliances—that deliver alternatives to the high cost of managing file-based data. To learn more about Oracle Solaris 10 and Oracle storage solutions, visit the Web sites listed in Table 1 below.

TABLE 1. WEB SITES FOR MORE INFORMATION

WEB SITES	
Oracle Solaris Operating System	http://www.oracle.com/solaris
Oracle's Sun Unified Storage Systems	http://www.oracle.com/us/products/servers-storage/storage/unified-storage/
Oracle Solaris ZFS File System	http://www.oracle.com/us/products/servers-storage/storage/storage-software/031857.htm
ARTICLES, PAPERS, AND BLOGS	
ZFS Wiki	http://www.solarisinternals.com//wiki/index.php?title=Category:ZFS
Oracle Solaris ZFS Migration Strategies	http://www.solarisinternals.com/wiki/index.php/ZFS_Best_Practices_Guide - ZFS_Migration_Strategies
Oracle Solaris ZFS Boot	http://blogs.sun.com/video/entry/zfs_boot_in_s10u6
BOOKS	
Solaris Performance and Tools: DTrace and MDB Techniques for Solaris 10 and OpenSolaris, ISBN: 0131568191.	
Solaris 10 ZFS Essentials, ISBN: 0137000103.	
Solaris 10 System Administration Essentials, ISBN: 013700009.	

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

Create a Better Storage
Infrastructure with the Oracle Solaris
10 Operating System
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.