ORACLE
SOLARIS

An Oracle White Paper
February 2011

Oracle Solaris: The Carrier-Grade Operating System

ORACLE

Executive Summary

Carrier grade refers to a system, or to a hardware or software component, that is engineered, tested and proven to deliver high availability, reliability, serviceability, security, failover and recovery, security, upgradeability, interoperability, scalability, soft real-time behavior, high performance and other capabilities.[1]

Telecommunications systems, both hardware and software components must provide carrier-grade service with all that the term carrier grade implies. The world's first commercial off-the-shelf (COTS) telecommunications systems certified to be Network Equipment Building System (NEBS) Carrier Grade by Telcordia, then Bellcore, was in 1997 by a company called Sun, which offered Netra systems that ran what is today called Oracle Solaris. Those systems were a direct response to requests from the telecommunications industry for Sun to provide carrier-grade systems to replace the proprietary systems that the telecommunications industry had developed previously. Those and subsequent systems operating with Oracle Solaris on other platforms have become the foundation of telecommunications networks throughout the world. With its modern advanced features, Oracle Solaris offers essential carrier-grade operating system (CGOS) capabilities, which focus on reliability, availability, and performance.

Oracle is committed to the communications market, supporting network equipment providers (NEPs), communication service providers, and carriers. Oracle Solaris 10 is a robust, mature, and capable carrier-grade OS, which has seamlessly evolved since its release in 2005. Now, with the release of Oracle Solaris 11 Express, new cost-saving, performance-enhancing features are available for review and evaluation, with full support from Oracle.

[1] SCOPE Alliance, *Carrier Grade Operating System Gap Analysis, Version 2.1*, http://scope-alliance.org/technical-documents/

This document is intended for IT architects, system administrators, and developers who want to understand how Oracle® Solaris and Sun servers can improve your application, operation, and infrastructure environment. This paper will provide technical guidance on how Oracle Solaris provides the capabilities for carrier-grade applications. The Overview section offers a brief description of some of the key functionality of Oracle Solaris, and subsequent sections provide more technical detail.

Powering Communication—The Oracle® Solaris Ecosystem

Thousands of telco customers, including carriers, service providers, and network equipment providers (NEPs), worldwide depend on Oracle Solaris to run their keep their customers up, running, and connected, with good reason:

- **Unmatched reliability** is the result of comprehensive testing across an integrated solution stack. Uptime is enhanced with innovation such as 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.
- **Superior performance** is proven with record-setting benchmarks, including TPC-H and TPC-C, PeopleSoft, Oracle BIEE, and many others, demonstrate maximum utilization. Oracle Solaris is optimized for throughput and scalability for the latest SPARC, Intel, and AMD processor technologies.
- **Built-in virtualization** enables efficient consolidation with Oracle Solaris Containers, Oracle VM Server for SPARC and other OS and network virtualization capabilities. These capabilities ensure flexibility, high-performance, without significant overhead.
- **Pervasive Oracle Solaris security infrastructure** that provides the compartmentalization and control required by governments and financial institutions, and helping to ensure peace-of-mind in multitenancy environments.
- **Committed support** for your infrastructure. Oracle offers sustaining support for Oracle Solaris releases for as long as you operate your systems, making it possible for you to keep your software infrastructure in place for as long as you consider it to make business sense.[2]

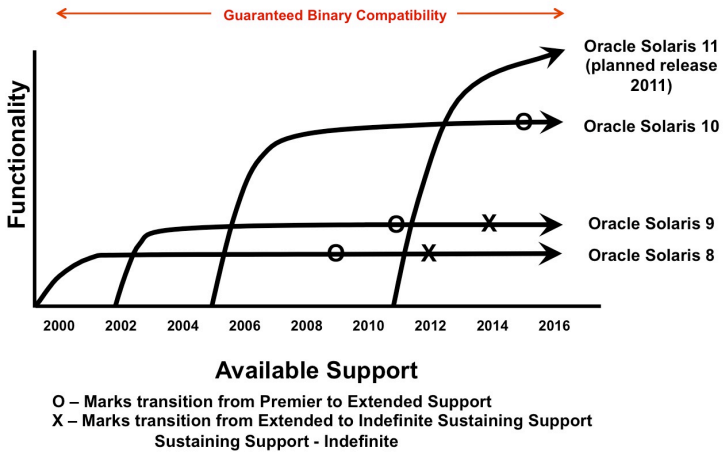

Figure 1: Oracle Solaris support timelines

These benefits help telco companies reduce capital and operational costs and enable improved margins and ROI. A leader in several segments, Oracle provides the technology, expertise, and products to support telecoms and help them achieve and maintain leadership.

- 10 of the top 10 network equipment providers rely on Oracle Solaris and Oracle's Sun hardware
- 5 of the top 5 handset manufacturers license Java™
- 98 of the top 100 telcos run Oracle Applications
- 97 of the top 100 telcos run Oracle Fusion Middleware
- 100 of the top 100 telcos run Oracle Database
- 10 of the 10 top telcos get better results with Oracle storage

Responding to increasing customer demand for applications across virtually any device, carriers, service providers, and NEPs are accelerating the deployment of next-generation services. The challenges are immense: competitive pressure, uncertain government regulation, and rapidly-changing technology. But new opportunities can create differentiated – and more profitable – services, and lower operating costs for legacy applications in both business and consumer environments.

Oracle has a long and successful history with the telco industry, helping companies move away from customized applications and systems to those based on a commercially available offerings. Oracle Solaris is a powerful foundation for the delivery of telco services, and provides many compelling capabilities to the telco industry, offering carrier-grade capabilities in performance, availability, and security, increasing service levels—which are being improved through integration and optimization across the entire stack.

[2] http://www.oracle.com/us/support/lifetime-support/

Integrated and Optimized Stack

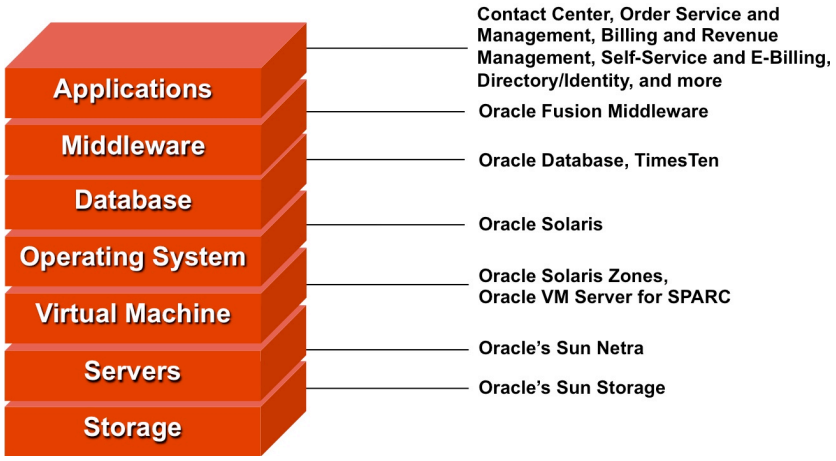

Figure 2: Application-to-disk integration and optimization drives value, performance, and service availability for telco environments.

Only Oracle provides a complete, integrated solutions portfolio spanning the entire communications systems landscape—from a carrier-grade operating system and servers, storage and IT infrastructure, to mission-critical business and operational support systems and service delivery platforms; from business intelligence applications and retail point-of-sale solutions to the Java™ platform running more than three billion mobile devices worldwide.

Oracle is continuing to improve platforms for scalability, reliability, and security for the telco industry. These improvements enhance and optimize the entire stack and leverage innovation throughout. A compelling benefit for the telco industry is the application-to-disk integration, optimizing performance, reliability, and security for a complete solution stack. Oracle is performing a wide range of test and optimizations not only within each tier of the solution stack, but across tiers and on targeted hardware platforms. This includes load and stress tests, fault injection, and interoperability and scalability testing. This end-to-end testing helps ensure maximum performance on a fully integrated system from NEBS-certified systems to certified releases of Oracle Solaris.

End-to-End 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 a firewall for network protection, Oracle Solaris sets a high standard for OS security by providing security capabilities at every layer, and taking advantage of these capabilities across all layers. For example, it is optimized to work with the built-in security and encryption capabilities of SPARC and x86 servers, speeding and hardening application login, networking, data storage, and more. 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 releases. 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.
- Oracle Solaris 10 offers role-based access control (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.
- Oracle Solaris 10 includes Oracle Solaris Trusted Extensions, which can be used to add an additional layer of protection via multi-label security, which enables data security policies to be separated from data ownership.

Oracle Solaris provides flexible and robust security to meet the dynamic requirements of the telco industry. This includes government-grade 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 an important part of why Oracle is a carrier-grade OS.

Unparalleled Performance and Scalability

Oracle Solaris has long held a unique position in the telco industry by delivering a single OS that can help CSPs and NEP 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 you to use both scalability dimensions to best meet your critical performance and availability criteria. Oracle Solaris supports telco customers who need to scale up 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.

Consider how integration and innovation in Oracle Solaris can help reduce costs and increase customer satisfaction. A traditional practice in telco environments is to provide one dedicated server per application, providing maximum scalability, security, and performance on demand. System utilization is typically very low except for relatively rare periods of peak demand. These can be consolidated onto a single, larger system using Oracle Solaris Zones, which offers each user a completely secure environment. Because Oracle Solaris Zones provides virtualized environments with extremely low overhead, hundreds of Zones can run on a medium-sized system, and in aggregate consuming as little as 50 percent of overall resources—leaving ample head room for peak demand.

Building on a proven track record, Oracle Solaris is ready to take advantage of the unique performance capabilities inherent in the latest generation SPARC and x86 processors. Oracle Solaris Dynamic Tracing (DTrace) helps developers isolate performance bottlenecks when designing application code, and allows system administrators to safely analyze and highlight a broad range of issues in production environments. As well, significant performance innovation comes from Oracle Solaris optimizations

for both individual processing cores and the overall multicore microarchitectures, increasing 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. There are performance improvements in critical areas such as the networking stack and disk subsystems, too, providing maximum application throughput.

Increased Reliability

In deployments with large numbers of users and service-level agreements, system reliability and availability are essential. A small memory glitch can affect hundreds or thousands of subscribers. In telecom, few things—if any—are more important than service availability. Nothing creates churn and customer dissatisfaction more quickly than unreliable services. Oracle Solaris delivers a foundation to cost-effectively deliver applications and services with carrier-grade availability.

Oracle Solaris features predictive self healing capabilities, significantly reducing downtime or unplanned outages. Self healing technology enables Sun systems and services to maximize availability when software and hardware faults occur. Combined with enhanced high availability solutions such as Oracle Real Application Clusters (RAC) and Oracle Solaris Cluster, maximum service availability is achieved while minimizing any required, underutilized resources.

Reliability is enhanced through the new image packaging system (IPS) available in Oracle Solaris 11 Express. This new system provides a more intelligent, streamlined approach to provisioning, patching, and upgrading local and remote systems, all while minimizing the amount of time a system is offline and out of service. These features, and many others, facilitate a simpler and more effective end-to-end experience for system administrators and service providers, thereby reducing costs.

Unmatched Flexibility

Oracle offers one of the industry's broadest product and technology portfolios, from application to disk. You can choose from single-to-multisocket SPARC and x86 systems ranging from 4-512 processing cores with terabytes of memory—Oracle Solaris powers them all. Oracle Solaris Legacy Containers enables application environments that were certified for Oracle Solaris 8 and Oracle Solaris 9 to run on Oracle Solaris 10, and Oracle Solaris 10 environments can run safely in Oracle Solaris 11 Express Zones, which means that these legacy services can take advantage of new hardware and capabilities. For example, DTrace can be used on such services when run in an Oracle Solaris Zone. Mature and proven virtualization environments are strictly secure and well behaved, enabling multitenancy through cost-effective consolidation.

SCOPE Alliance

SCOPE is a telco industry alliance committed to accelerating the deployment of carrier-grade base platforms for service provider applications. Founded in 2006 by NEPs, Oracle is a contributing member, along with several other technology companies. The SCOPE Alliance reviews industry standards and open specifications to determine their relevance and applicability to telecommunications systems. The objective is to identify subsets (called profiles) of existing standards and specifications

that are relevant to the Carrier-Grade Base Platforms (CGBP) of the telecommunications industry, as well as to identify missing requirements (called gaps) in those standards and specifications.

In 2007, the SCOPE Alliance formed the Carrier-Grade Operating System (CGOS) Working Group to review the Carrier-Grade Linux (CGL) version 4.0 specifications and to identify gaps in them. CGL version 4.0 comprises seven specifications: Availability, Clustering, Hardware, Performance, Security, Serviceability and Standards.

The SCOPE Alliance CGOS Working Group produced a profile for the Carrier-Grade Operating System, based on CGL version 4.0, and identified 16 gaps in those specifications, as they apply to telecommunications products and services. The gaps correspond to deficiencies in the CGL version 4.0 specifications, such as lack of APIs for SMP and multicore programming, OS tunable parameters, support for virtual IP routing and forwarding, and others. Many of these gaps are satisfied by Oracle Solaris. The SCOPE Alliance CGOS Gap Analysis document is publicly available on the SCOPE Alliance Web site.[3]

Oracle has made substantial contributions to the SCOPE Alliance CGOS profile requirements. For example, the fault-resistant file system and the unified cryptographic framework are based on Oracle's extensive experience with operating systems and the needs of the telecommunications industry.

Security

Modern telco infrastructures increasingly rely on IP backbones and internet connectivity. Unlike circuit-based architectures, modern networks are hyper-connected, with virtually unlimited points of access—leaving carriers and operators vulnerable to a wide variety of security threats—from both inadvertent acts that compromise security to intentional violations that expose sensitive information and corrupt data. The need for end-to-end security is necessary for virtually every aspect of your network in order to maintain the continued service operation and data integrity in the face of malicious attacks, intrusions, and unintended exposure.

Oracle Solaris provides a sophisticated network-wide security system that controls the way users access files 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, Oracle Solaris sets a high standard for operating system security by addressing security needs at every layer.

Security Hardening and Monitoring

Oracle Solaris 11 Express offers several features that make it easier to minimize and harden a system, including a fully "secure-by-default" network profile. With automatic secure by default, network

[3] http://scope-alliance.org/technical-documents/

services are reconfigured so that no network facilities are externally visible; they are either disabled, or set as internal services only. By default, the only externally accessible network service is secure shell (`ssh`), although others can be configured as necessary, through a well-defined and self-documenting Service Management Facility.

Process and User Rights Management

The weakest link in UNIX security has traditionally been the root (superuser) account, which is empowered with complete access to all system resources and privileges, making such accounts, and applications that run with root privileges, a prime target for hackers. 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.

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 role-based access control (RBAC) provide a compelling alternative model to the traditional superuser model. RBAC is a feature of Oracle Solaris and other Oracle products. 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.

Oracle Solaris 11 Express 2010.11 extends Oracle Solaris RBAC capabilities. The traditional UNIX root account is now accessible by default only as a role. Authorized users can assume the root role using the command line `sudo(1M)` utility rather than directly logging into a root user account. This feature enables authorized non-root users to complete tasks and scripts with superuser privileges, and makes it easier to document who performed what privileged operations.

Oracle Solaris also reduces opportunities for 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, Forced, and Basic Privileges

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.

An in-kernel `pfexec` implementation, new to Oracle Solaris 11 Express 2010.11, is used to execute administrative commands requiring a higher privilege level. A new process flag is used to specify that all subsequent program executions are subject to RBAC policy. This feature eliminates the need to modify shell scripts to invoke `pfexec` or profile shells. Additionally, Oracle Solaris 11 Express 2010.11 adds three new basic privileges (`file_read`, `file_write`, and `net_access`) beyond the five that exist in Oracle Solaris 10, which restrict read, write and outbound network access.

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. Secure by default networking configuration disables many unused network services, while configuring all other services for local system-only communications. From system boot, the administrator controls exactly which services run, who can manage those services, and what privileges those services run with.

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 Zones[4] 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

[4] *Zones* is an OS virtualization capability. In earlier versions of Oracle Solaris, it was also known as Containers. More detail is in the Virtualization section.

encryption algorithms, and a security policy database. Note that the Service Management Framework (SMF)[5] manages IPsec as a set of services, thereby increasing resiliency and reliability.

- **Labeled IPsec**. When labeled processes in a multilevel secure operating system, such as Oracle Solaris Trusted Extensions, communicate across system boundaries, their network traffic needs to be labeled and protected. Traditionally, this requirement is met by using physically separate network infrastructure to ensure that data belonging to different labeled domains stays in separate physical infrastructures. Labeled IPsec/IKE, new to Oracle Solaris 11 Express 2010.11, enables customers to reuse the same physical network infrastructure for labeled communications by transferring labeled data within separate labeled IPsec security associations, removing the need for redundant and expensive physical network infrastructure.

Cryptographic and Key Management Framework

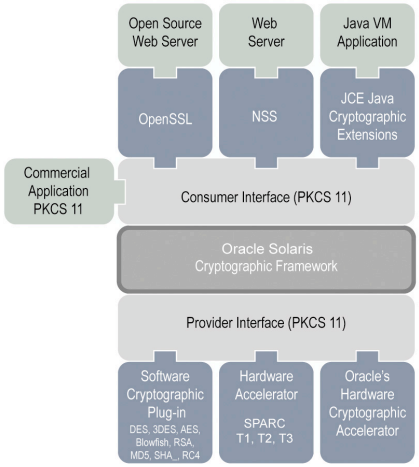

Figure 3: Oracle Solaris Cryptographic Framework is standardized and extensible—current and future cryptographic choices can easily plug in and take advantage of hardware and software capabilities.

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. To meet more

[5] Part of predictive self healing. More detail in the Reliability and Availability section.

stringent government standards, the Oracle Solaris Cryptographic Framework now supports the NSA Suite B algorithms.

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.

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.

Common Criteria and Trusted Extensions

Oracle Solaris has been tested 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. To enable greater flexibility and security, Trusted Extensions now enables per-label and per-user credentials allowing administrators to require a unique password for each label. This password is in addition to the session login password, and thus allowing administrators to set a per-zone encryption key for each label of every user's home directory. 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 broadens the proven Oracle Solaris 10 security model. It utilizes the user and process rights management feature in Oracle Solaris, Oracle Solaris Zones, file systems, and networking and doesn't require a new or separate kernel. Using trusted extensions doesn't require independent software vendors (ISVs) to requalify their applications to run them with sensitivity labels. Because it is an extension to Oracle Solaris 10 security policy, trusted extensions technology is flexible and quick to deploy. Developers and administrators can quickly add new applications, new users, and more, without extensive analysis of each application, without the need to write complex, error-prone security policies that require a system reboot.

Oracle Solaris 11 Express introduces labeled IPsec and labeled ZFS datasets. In Multi-Level Secure (MLS) environments, sensitivity labels are applied to operating system resources for the purposes of granting access in conjunction with least privilege. In Oracle Solaris Trusted Extensions, the existing CIPSO trusted networking protocol assumes that the underlying network is secure and that CIPSO packet headers cannot be manipulated or observed while packets are in transit. Oracle Solaris 11 Express introduces labeled IPsec, permitting sensitivity labels to be associated with IPsec-protected traffic, which facilitates labeled networking over untrusted networks. Oracle Solaris 11 Express can also apply sensitivity labels to ZFS datasets, further securing data assets in MLS deployments.

ZFS Dataset Encryption

Encrypted dataset support has been added to ZFS to protect against theft of physical storage, man-in-the-middle attacks on the SAN, and to provide dataset level secured deletion. Data is encrypted at the dataset level, allowing a mix of encrypted and unencrypted datasets in the same ZFS storage pool. A single dataset has a consistent policy that sets encryption only at the time of creation. All data and file system metadata is encrypted with a comprehensive encryption key management facility to cover different key management strategies.

Virtualization

Virtualization is important for carrier grade environments in that it balances the needs for high availability and security against the needs for flexibility, efficiency and resource sharing. Oracle Solaris virtualization offers extremely flexible resource management capabilities, including CPU, memory, network, and storage. For example, Oracle Solaris virtualization delivers more work from an existing IT infrastructure by increasing utilization. Virtualization also helps consolidate legacy applications from multiple obsolete hardware platforms onto a smaller number of up-to-date, more powerful, and more energy-efficient servers. Moving applications from a large set of underutilized servers to a smaller set of more powerful servers helps to reduce the number of servers to house, power, cool, and maintain.

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 Server for SPARC

Oracle's SPARC servers running Oracle Solaris are the only systems today that provide completely integrated application separation technologies at every level of the product stack, fully supported from one company—Oracle. Oracle VM Server for SPARC, previously called Sun Logical Domains, leverages the built-in SPARC hypervisor to subdivide and reconfigure 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 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. SPARC 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 SPARC T-Series 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.

Oracle SPARC servers running Oracle Solaris are the leading platform with the hard partitioning capability that provides the physical isolation needed to run independent operating systems.[6] Many customers have already used Oracle Solaris Zones 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 SPARC T-Series server with finer granularity for computing resources. For SPARC T-Series processors, the natural level of granularity is an execution thread, not a time-sliced microsecond of processing 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.

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

Oracle Solaris Zones

Supported on any server running Oracle Solaris, Oracle Solaris Zones (previously known as Oracle Solaris Containers) isolate software applications and services using flexible, software-defined boundaries. Oracle Solaris Zones provide virtualization and software partitioning, enabling the creation of many private execution environments from a single instance of Oracle Solaris. Oracle Solaris Zones provide low-overhead, low-latency environments required by carrier-grade environments.

Unlike virtual machines, Oracle Solaris Zones provide OS-level virtualization by giving the appearance of multiple OS instances rather than multiple physical machines. Isolation between Zones 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 Zones on a single system. Because of this negligible overhead, and unlike partitioning or virtual machines, Oracle Solaris Zones can be created in large numbers. For example:

- Service providers can provide isolated instances of Web servers or database instances.
- Individual developers can use safe, isolated test environments.

Zones create very low overhead compared to traditional virtual machines, maximizing the computing resources available to applications. Telcos and SPs can safely and more effectively consolidate

[6] This information is a technical description, and may or may not affect how applications are licensed.

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, dynamic domain, or logical domain where the Oracle Solaris Zones resides. Physical-to-virtual capabilities can be used to directly move an existing Oracle Solaris 10 system into an Oracle Solaris Zone; the hosted environment can be emulated as needed.

Hosting applications within individual Oracle Solaris Zones provides administrators the ability to exert fine-grained control over rights and resources within a consolidated server. The introduction of `zonestat` in Oracle Solaris 11 Express 2010.11 greatly facilitates the observation of system resources consumed by Oracle Solaris Zones. More specifically, administrators can observe memory and CPU utilization, utilization of resource control limits, total utilization and per-zone utilization break-down over specified time periods. In addition, with Oracle Solaris 11 Express you can delegate common zone administration tasks for specific zones to different administrators using RBAC. With delegated administration, for each zone, a user or set of users may be identified with the permissions to login, manage or clone that zone. These specific authorizations will be interpreted by the appropriate commands running in the global zone to allow access at the correct authorization level to the correct user.

Investment Protection

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 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. This means you can run existing applications that have been certified to run on a known environment on new, more powerful, energy-efficient, and productive systems, and transition these legacy environments to native Oracle Solaris Zones at your own pace.

For users already running their applications either in zones or on bare-metal Oracle Solaris 10 systems, virtual-to-virtual (v2v) and physical-to-virtual (p2v) tools are provided to help this transition over to an Oracle Solaris 11 Express Zone.

Virtualized Networking

Oracle Solaris 11 Express brings network virtualization and resource control into the operating system architecture, instead of adding on layered functionality with heavy overhead and undue complexity. The network stack for Oracle Solaris 11 has been substantially re-architected from that of Oracle Solaris 10 in an ambitious effort—the Crossbow project—to provide a fully virtualizable network environment that can share and manage network resources, decrease latency, and improve throughput as network load increases.

Oracle Solaris 11 Express enables the virtualization of the hardware network interface card (NIC) into virtual NICs (VNICs)—dividing a physical NIC port into multiple virtual interfaces to create kernel-enforced isolated and dedicated network stacks from physical interface to application. VNICs are built on top of dedicated resources like Rx/Tx rings, DMA channels, kernel queues and threads, CPUs, and available bandwidth. These hardware rings, such as on SPARC-based servers, can offload the CPU from network traffic. This allows multiple Oracle Solaris Zones to share the available bandwidth and host networking resources based on policies or resource partitioning without any negative performance impact due to virtualization. VNICs and virtual switches are some of the key components that enable virtual networks with higher level services, such as virtual routers, virtual firewalls, virtual load balancers, and more.

The power of VNICs has been tightly integrated with Oracle Solaris Zones. An Exclusive-IP Zone provides a complete, tunable, manageable and independent networking engine to each zone; such a zone can configure DHCP, IP Routing, IP multipathing, IPsec, and more. You can create an unlimited number of exclusive-IP zones, even on a system with a single physical NIC. In addition this capability is also extended to Oracle Solaris 10 Zones (the ability to run a Solaris 10 environment inside a zone on top of Oracle Solaris 11 Express).

There is also InfiniBand IPoIB support for Oracle Solaris Zones. This allows multiple IPoIB instances to be created on the same InfiniBand port, and enables each zone to talk to others by utilizing their own IPoIB instances.

The new network stack architecture introduces network resource management. This enables bandwidth limits for handling network traffic. For example, you can limit available bandwidth on a 10 Gb physical interface into smaller lanes, which can be assigned to different Oracle Solaris Zones. The Oracle Solaris will enforce the bandwidth and/or network CPU resources assigned, preventing one zone from using more network resources than expected.

Link protection is a new mechanism for preventing potentially malicious or misbehaving guest VMs from sending harmful packets to the network. This feature provides protection against these basic threats: IP, DHCP, MAC and L2 frame spoofing. Unlike a traditional firewall, link protection does not support inbound filtering or customizable filtering rules.

More information on Crossbow can be found in the Network section.

Reliability and Availability

Oracle Solaris is designed for carrier-grade reliability and availability, and there are many capabilities that contribute to this, including the following.

- Fault and service management overcome both hardware and software runtime issues, and restart applications and services—even those with complex, distributed architectures.
- A new image packaging system streamlines upgrades, patches, and new package additions while minimizing downtime and administrative overhead.
- Oracle Solaris ZFS delivers a robust and flexible file system with advanced features.

- High availability functionality ensures rapid failover, even over long distances.
- Oracle Solaris is tightly integrated with hardware reliability features offered on SPARC Enterprise and Sun x86 systems. Examples of this are memory mirroring, offlining problematic memory pages or processing cores. As well as other capabilities such as "lights-out management," fast reboot, and persistent shared memory.
- With Oracle's whole-stack testing, the comprehensive solution stack—from application to disk—is tested for increased reliability and availability.

Reliability also benefits from strong, pervasive security. Protecting systems and data from unauthorized or inadvertent use keeps them from going offline. Security capabilities are discussed in detail in Chapter 3.

New Image Packaging System

The Oracle Solaris Image Packaging System (IPS) is a new packaging administration model that enables you to install, update, and remove software packages for Oracle Solaris. IPS enables you to determine very specifically the packages that are installed on each system. You can specifically define what is needed—and don't need—on each system, and test that the software installs and operates as expected. You can also install other packages in addition to Oracle Solaris packages as part of a single distribution. The results can be verified, even for remote systems.

IPS also provides an easy method of sending new software packages to a repository through a series of package transactions to add package content, package metadata and dependent system services upon installation to a publisher. Administrators can easily create and manage new package repositories and associating publishers for local software delivery in a telco environment.

IPS provides a framework for complete software lifecycle management such as installation, upgrade and removal of software packages. Combined with the ZFS file system and boot environments, IPS offers completely safe system upgrades. Administrators can install software from network based package repositories with full automatic dependency checking for any additional libraries that might be required during a software package install. IPS also enables you to create your own software packages, create and manage package repositories, and mirror existing package repositories. Key features of the new IPS include:

- Lock down binaries or services so that they cannot be inadvertently removed through a subsequent patch or update without explicit authorization.
- Test that a package or series of packages can be installed without rebooting, or that if a reboot is required it happens successfully.
- Automatically obtain the minimum number of supporting files required to install an application or service, and update them as required.
- Support dependencies for applications or services, including minimum or maximum release levels, or within a specific range.

- Back out of an installation by reverting to the previous system image. The combination of IPS and ZFS allows Oracle Solaris 11 Express to easily, efficiently, and automatically create fallback alternate boot environments as the system is updated.
- Repositories record what images are placed on each system, and you can control which repositories each system can access.

IPS package repositories also provide the administrator an opportunity to greatly simplify software delivery with a completely centralized architecture for managing a selection of software, multiple versions of that software, and multiple different architectures. Administrators can control access to different software package repositories, and mirror existing repositories locally for network restricted deployment environments.

IPS includes a number of command-line utilities and graphical tools, Package Manager and Update Manager. With IPS, anyone producing a software package can make it available on internal or external Web sites as a single-click install. This is done via a new MIME association, `.p5i`. IPS makes it easy for developers to transparently support multiple architectures with a single package. For example, a single click can automatically install SPARC packages on a SPARC system, x86 packages on x86 server, and bring in custom components for specific configurations—all from a single installation package, and transparent to the end user. IPS also provides the ability to validate its consistency on a system and fix any software packages should any errors occur during that validation 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 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.

Boot Environments and Fast Reboot

Boot environments in Oracle Solaris 11 Express 2010.11 represent a significant evolution of the Live Upgrade experience used in Oracle Solaris 10 and earlier releases with an out of the box configuration to allow for system updates to be applied in parallel on a live production system.

Solaris 11 boot environments are designed specifically for the ZFS file system, utilizing its fast snapshot and clone features to save a copy of the boot environment any time a software update to the system is performed. A snapshot is taken before any package is installed or updated, and if the update includes modifications to the system that will require a reboot to effect, then the snapshot is cloned to create a new boot environment; the package operations are then applied to the new boot environment. Once an administrator is satisfied that the update is ready to be used, the system can be rebooted into the new boot environment to activate it. In the event that an update turns out to be problematic for some reason, they can quickly activate a previous boot environment and reboot back into it to restore a prior state. ZFS file system snapshots and clones have low overhead and provide unprecedented flexibility for system management.

The time to reboot on Oracle Solaris 11 Express 2010.11 is significantly faster by default, allowing tests to be skipped on SPARC systems, and firmware and boot loader to be bypassed on x86 systems. Administrators can change this behavior by modifying the `config/fastreboot_default` SMF property in the `svc:/system/boot-config:default` service.

Fault and Service Management

Predictive self healing is an innovative capability in Oracle Solaris that automatically diagnoses, isolates, and recovers from many hardware and application faults. This enables 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 Oracle 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. SMF acts 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.

Predictive self healing capabilities offer extensive reliability and availability capabilities on all Oracle 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 has 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.

Cluster and Failover

Keeping application data and services in a single system exposes telcos and SPs to potential failure from any component of the configuration. Carrier 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, and 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 Zones

- 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) service continuity for local datacenters, 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 Zone 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.

Oracle Solaris ZFS

A critical element to carrier-grade 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.

Oracle Solaris ZFS is the root file system on Oracle Solaris 11 Express 2010.11, offering a superior experience in terms of manageability, scalability and data integrity. ZFS presents a pooled storage model that completely eliminates the concept of volumes and the associated problems of partitions, provisioning, wasted bandwidth and stranded storage. Thousands of file systems can draw from a common storage pool, each one consuming only as much space as it actually needs. All operations are copy-on-write transactions, ensuring that the on-disk state is always valid. Additionally, blocks are checksummed to prevent silent data corruption, allowing data to self-heal itself in replicated (mirrored or RAID) configurations. If one copy is damaged, ZFS detects it and uses another copy to repair it. ZFS is also at the heart of Oracle Solaris 11 software installation and management with the IPS packaging system, greatly reducing planned and unplanned downtime with safe system upgrade capability.

Oracle Solaris ZFS uses checksums to protect all data, and reads and checks data to help ensure it is correct. If an error is detected 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 carrier-grade 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.

Oracle Solaris ZFS offers other important features:

- **Deduplication**, in which varying mechanisms are employed to reduce the amount of total data stored by eliminating and sharing common components. ZFS deduplication support has been added as of Oracle Solaris 11 Express. ZFS deduplication uses checksum based comparison of blocks with optional verification (for example with noncryptographically secure checksums). Deduplication is performed across the entire ZFS storage pool; administrators can select if individual datasets have deduplication enabled or not. This is useful in mixed-mode environments in which some datasets have highly duplicated data (such as virtualized images, home directories, or email folders) and others are unique (such as databases).
- **Encrypted ZFS datasets**, with on-disk encryption/decryption support and key management for ZFS datasets, Oracle Solaris 11 Express adds more exacting mechanisms for data protection. In the event of theft or in the case of untrusted paths to networked storage, encrypted ZFS datasets can help to safeguard data and prevent unauthorized access. The kernel implements raw encryption/decryption functions, which are applied to all data and file system metadata, including file owners, ACLs (Access Control Lists), file sizes, and more.
- **Snapshots**, 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 that need to perform backups, and other users that need 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.

- **ZFS Diff**, which lists the differences between ZFS snapshots, has been added as of Oracle Solaris 11 Express. Users with the appropriate privilege can now view what file and directory level changes have occurred between snapshots, such as files or directories added, removed, modified or renamed in a later snapshot.
- **Compression** can transparently compress file system data. Compression can be enabled or disabled dynamically, and compression policies can be applied on a per-file system basis. This can be used to save space on a file system that is used primarily to archive data.

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).

POSIX Compliance

ZFS incorporates a volume manager and other useful features. The ZFS POSIX Layer (ZPL) employs traditional standard POSIX file system semantics. The Data Management Unit (DMU) provides a general-purpose transactional object store. The Storage Pool Allocator (SPA) provides virtual block allocation and data transformations, such as compression. In essence, ZFS looks like a virtual memory system backed by disks instead of physical memory.

Network Support

To support IP convergence and carrier-grade network services, Oracle has improved network support in Oracle Solaris. The networking stack in Oracle Solaris 11 Express 2010.11 has been re-architected to unify, simplify, and enhance performance, observability, and interoperability of network interfaces and protocols.

Performance

Oracle Solaris delivers a highly scalable and enhanced networking stack that lowers overhead by reducing the number of instructions required to process packets. This efficiency also increases scalability, allowing more connections and enabling server network throughput to grow linearly with the number of CPUs or threads and network interface cards (NICs). The enhanced stack is tuned for 10 Gb/sec Ethernet, wireless, and offloading technologies and provides the foundation to add protocols without affecting the network stack's performance or scalability. This feature facilitates the integration of emerging protocols under development in standards bodies such as the Internet Engineering Task Force (IETF).

Oracle has concentrated on improving the performance of key server workloads that have a significant networking component. The sockets implementation has been improved and no longer uses STREAMS. The architecture monitors network traffic volume allowing it to shift from interrupt driven

to polling mode which is much more efficient when dealing with high network traffic volumes. The new sockets implementation is based on a procedural interface, which affords faster connection setup and teardown, as well as higher throughput and increased scalability. Historically Oracle Solaris has supported streams-based sockets, and with the move to the new architecture there are significant performance improvements along with a simplified developer interface for adding new socket types.

The InfiniBand stack in Oracle Solaris 11 Express has also undergone some significant improvements. Lending itself to the success of the Oracle RAC and Oracle Exadata products the InfiniBand stack has improved Sockets Direct Protocol (SDP) allowing the transparent redirection of TCP/IP usage to SDP and the efficiencies that brings. It also now implements the RDSv3 protocol providing better performance and observability for Oracle RAC databases.

Oracle Solaris 11 Express improves performance on non-uniform memory access (NUMA—essentially all modern, multi-CPU servers) systems by providing affinity between memory, CPU, resources and I/O. This addresses the issue of varying I/O performance depending on the location of kernel resources and the location of I/O devices.

Oracle Solaris 11 Express also introduces the idea of network flows. A flow can be defined for a subset of the network traffic, for example traffic for a specific service or IP destination. This becomes particularly powerful when used in combination with resource controls. A good example here is periodic backup operations. Before the advent of bandwidth limitations and flow control it would have been possible for the monthly backup operation to overwhelm the external network to the detriment of other, more important, applications. By defining a flow from say, the source server IP address to the backup server IP address, and applying resource management to that flow, undesired behavior can be prevented.

Oracle Solaris 11 Express 2010.11 includes an integrated L3/L4 load balancer. This capability can be complementary to existing higher layer load balancing solutions from different ISVs. The addition includes stateless DSR and NAT operation modes on a variety of load balancing algorithms, a command-line and configuration API to configure various features as well as view statistics and other configuration details.

These changes made to the networking stack through the Crossbow project can increase network throughput by scheduling and handling packets more efficiently.

Crossbow parallelizes the networking workload across multiple threads and cores using hardware or software classifier. It creates a unique path comprising of NIC hardware (Rx/Tx Rings, DMA channels, various registers), kernel threads and kernel queues, and has affinity to CPUs (cores and threads). There is no synchronization, shared locks, or shared counters between various kernel threads, which allows the stack to linearly scale not only with cores and threads, but also as more NICs are added to the system.

The best performance gains typically come with the latest generation intelligent NICs with packet classification and multiple receive and transmit ring buffers that the networking stack code can manage. In addition, Crossbow can move the networking stack away from interrupt model and instead introduces a dynamic polling model where packet processing is done in form of chains (instead of per-

packet interrupt) and without context switching and thread pinning associated with interrupt mode. As such, even single CPU performance increases dramatically.

Observability and Administration

Oracle Solaris 11 Express 2010.11 improves the area of IP observability, allowing the developer or administrator to use common packet sniffing tools such as `wireshark(1)` and `snoop(1M)` to view all IP traffic sent on real and virtual paths. In addition, Oracle Solaris 11 Express includes `dlstat(1M)`, a tool to provide runtime statistics for data links allowing administrators to get a better understanding of how well their networks are performing. Now all traffic can be observed at the IP layer, including to and from an Oracle Solaris Zone (per VNIC usage on a physical NIC). The `dladm` and `flowadm` commands are used to provide configuration information for data links and flows respectively.

There is new support for the Virtual Router Redundancy Protocol (VRRP), which provides high availability to our integrated L4/L3 load balancing and also our integrated router technologies. Improved datalink management now allows vanity naming, `dladm` support for legacy DLPI drivers, unified driver property configuration and simply makes the life of administrators that much easier.

IP Multipathing (IPMP) has been re-architected, providing improved administration and observability. Now it can work transparently with all IP based applications including core technologies such as DHCP. IPMP provides transparent redundancy for IP level communications between the application running on a system, and the first router in the communication path to the outside world. IPMP allows creating multiple paths to that first router so that port, NIC, cable or switch failures will not impact any connections. For high availability applications IPMP will assign one of the interfaces to the IP address and continually monitor the underlying interfaces to ensure a connection is maintained. If IPMP detects that the IP interface being used has failed it will use an alternate working IP interface. Applications do not need to be aware they are running on a system managed by IPMP.

For Oracle Solaris 11 Express 2010.11, IPMP has undergone a significant re-architecture to improve network administration and network observability. The network driver framework now has a kernel plugin mechanism that provides an architecture for implementing distinct MAC layers such as Ethernet, Wi-Fi, InfiniBand, and IP tunnels. In addition GLDv3 core driver APIs are now available for use by 3rd Party driver writers.

The introduction of a new network driver framework (Generic LAN Driver version 3—GLDv3) provides VLAN, link aggregation, and the ability to support MAC layers other than Ethernet, including IP tunnels, Wi-Fi, and InfiniBand. Flexible network administration is provided with `dladm(1M)`. Enhancements to `dladm` also include the ability to allow links to be renamed, including non-GLDv3 links, and set properties for NIC drivers using a common command.

Clearview

Clearview is a project to rationalize, unify, and enhance the way network interfaces are handled in Oracle Solaris at the programmatic and administrative levels. Clearview's capabilities have been implemented as part of Oracle Solaris 11 Express, enabling you to:

- Observe all IP layer network traffic, including loopback, IPMP group and IP tunnel traffic.
- Observe all IP layer network traffic flowing to, from, and within an Oracle Solaris Zone.
- Administrate all network interfaces using `dladm(1M)`.
- Use VLANs and form link aggregations on all Ethernet devices.
- Use IPMP with technologies such as DHCP and routing protocols.
- Select names for network interfaces, easing migration of network configuration information between systems or zones, and markedly improving Dynamic Reconfiguration.

Bridging and Tunneling

Bridging is a general Layer 2 technology that is used to connect together separate L2 subnetworks, allowing communication between attached nodes as if only a single subnetwork were in use. Basic Ethernet bridging support has been added to Oracle Solaris 11 Express 2010.11 using the Spanning Tree Protocol (STP, IEEE 802.1D-1998) and TRILL protocol.

IP tunneling functionality has been re-implemented in Oracle Solaris 11 Express 2010.11, delivering a generic LAN driver (`iptun`) that implements IP tunnel links atop which IP interfaces can be plumbed and managed through `dladm(1M)`. With this new architecture, tunnel links gain functionality common to other links, including link vanity naming, link-layer observability using transitional observability tools such as `wireshark(1)` and `snoop(1M)`, and the assignment of tunnel links to exclusive stack non-global zones.

Protocol Support

Oracle Solaris includes in-kernel support for key telco protocols, including IPv6, SCTP, and SIP, making it an ideal development and deployment platform for telephony applications.

IPv6

The IPv6 protocol is designed to meet the global demand for network connectivity. It leverages the design of IPv4—the current IP protocol—and extends it by providing a very large number of addresses that enable the vision of a vast global network of many different types of devices. With IPv6, the internet connects not only people and computers but also virtually any kind of electronic device that can take advantage of internet connectivity and flexibility.

Oracle Solaris supports current IPv6 specifications and APIs and provides full integration with the IPsec implementation, including the Internet Key Exchange (IKE). This enables encrypted and

authenticated network access between systems. Oracle Solaris provides the ideal foundation for a high-performance, secure, and robust IPv6-based Web services infrastructure.

To help ease the transition to IPv6, Oracle Solaris also provides a dual network stack with tunneling tools. Additionally, many tools and services support IPv6 such as IP Filter, Simple Network Management Protocol (SNMP), management information base (MIB), and DHCP client.

Streams Transmission Control Protocol

The Streams Transmission Control Protocol (STCP) is a message-based alternative to TCP that is used for IP-based SS7 signaling (IETF SIGTRAN) in telecommunications systems. STCP provides reliable transport over IP and is excellent for high-availability deployments such as IP telephony and SIGTRAN, which require more-reliable network connections. STCP allows more than one IP address, is resistant to SYN flooding attacks, and offers either unordered or ordered message delivery. It employs selective acknowledgements for more efficient retransmission of dropped packets, CRC-32C for integrity checking, and an efficient congestion control mechanism.

Session Initiation Protocol

The Session Initiation Protocol (SIP) is an application-layer protocol that establishes, modifies, and terminates interactive sessions, including VoIP sessions—the IP equivalent of SS7. It is typically used for streaming audio and video applications, incorporating SIP into Oracle Solaris offers the potential for next-generation applications and devices. The Oracle Solaris SIP server implementation is supported on a wide range of industry-standard platforms and interoperates with clients on a range of OSes, such as Microsoft Windows and Linux. This makes Oracle Solaris ideal for development projects that seek to exploit these new capabilities.

Improved Network Availability and Routing Protocols Support

Oracle Solaris includes Network Layer 3 redundancy, providing the ability to implement high-availability network solutions for more-resilient services and innovative new applications. Layer 3 multipathing (MP) enables end-to-end redundancy ensuring greater protection from network failures. This standards-based MP feature is implemented via a combination of virtual IP address selection and OSPF-MP. Virtual IP address selection enables system administrators to specify IP source addresses for packets on a per-network basis. OSPF-MP employs the protocol to route traffic around failed network interfaces. Oracle Solaris also includes OSPFv2 and BGP4 routing protocols, making it easier to administer complex routing policies.

Performance

Successful carrier-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.

Oracle Solaris has long held a unique position in the industry by delivering a single OS that can help telcos, SPs, and NEPs 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 telco providers to use both scalability dimensions to best meet their critical performance and availability criteria. Oracle Solaris supports customers who need to scale up back-end 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.

This scalability was proven out in a recent benchmark of Oracle Communications Order and Service Management. A SPARC T3-1 running Oracle Solaris 10 and Oracle Communications Order and Service Management 7.0 cluster on Oracle WebLogic Server 11*g*, and the SPARC T3-2 running Oracle Solaris 10 and Oracle Database 11*g*, handled a record number of orders, showing a 2.4x performance improvement over x86-based solutions.[7] Oracle Communications Order and Service Management enables communications service providers to address the enterprise-wide central order management challenges that are critical to their business success. It is a core component of Oracle's solution for Rapid Offer Design and Order Delivery, which enables some of the largest communications providers in the world to support ever-increasing order volumes, sometimes measured in millions of orders per day, with a predictable low cost of ownership. The benchmark simulates fulfillment of fixed-line phone activation orders that are executed as a sequence of tasks, and measures the number of completed tasks per second (tps).

Real-time Capabilities

Real-time capabilities are essential for a carrier-grade OS to deliver the responsiveness and predictability needed for telecommunications services. Solaris was designed with real-time capabilities from its inception. For example, data structures in the Oracle Solaris kernel are protected by mutexes, rather than by raising and lowering interrupt priority levels. And, the real-time scheduling class provides a fully preemptible kernel, with priority inheritance, and offers both FIFO and round-robin scheduling.

Several key features in the original Oracle Solaris design have been enhanced over the years and additional functionality has been added to make its time-critical processing capabilities more compelling. Original design features included:

[7] www.oracle.com/us/solutions/performance-scalability/t3-2-osm-92010-bmark-172819.html

- **Interrupts as threads.** Every entity that executes code in Oracle Solaris, whether as a part of a user process or running solely in the kernel, is represented by a thread.
- **Fully preemptible kernel. Oracle** Solaris is fully pre-emptible for processes in the real-time class. If a RT process becomes runnable, it will immediately be placed on a CPU if its priority is higher than the thread running on that CPU.
- **Real-time scheduling.** Real-time response latencies are provided by using the real-time scheduling class, which offers two options: round robin or FIFO scheduling. Entities in the real-time scheduling class are immediately scheduled to run if they become runnable and are of higher priority than what is currently running. Entities not in the real-time class will suffer delays when they become runnable even if they are of higher priority than what is running.
- **Priority Inheriting Synchronization primitives.** Oracle Solaris mutexes implement a basic priority inheritance protocol. When the high-level entity blocks, all of the entities blocking it are given the high level entity's priority. When they cease to block the thread, their priorities revert to their previous level.

Newer features include:

- **Processor sets.** Introduced in Oracle Solaris 10, processor sets provide the ability to select one or more CPUs and isolate them from the rest of the system in terms of what entities are scheduled to run on them. Processes and or threads within the process may be bound to individual CPUs within the set.
- **POSIX compliance.** Oracle Solaris has virtually full support for what was once known as POSIX 1003.1b (real time) and full support for POSIX 1003.1c (threads). These are now part of the IEEE 1003.1, 2004 Edition standard. Support of requisite real-time features is indicated by support of various option groups (such as thread priority inheritance). These options are also grouped into larger bundles, of which four exist relating to real-time and threads. These are real-time, Advanced Real-time, Real-time Threads And Advanced Real-Time Threads. Oracle Solaris satisfies all but one option group within the Real-time group. Solaris fully satisfies the Real-time Threads option.
- **Fixed priority scheduler.** Oracle Solaris offers a fixed priority scheduler to aid in processing. While this is not precisely a real-time feature, it can be used as an option for processes that would otherwise be placed in the real-time scheduling class only because of the need for fixed priorities. Unlike the normal time sharing scheduling class, the fixed priority class does not modify a process' priority as it runs. The priority changes only if the process asks for the change and has the appropriate privilege for the change.

Multithread Awareness

Oracle Solaris is optimized for systems with hierarchical processor architectures 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

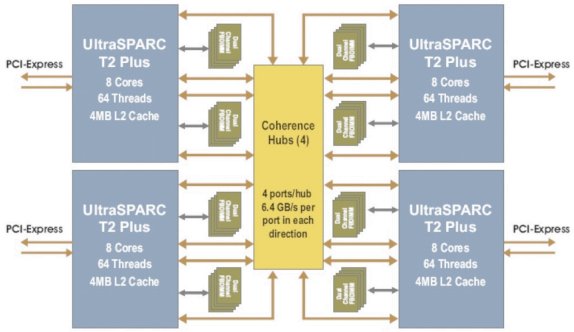

Figure 4: Oracle Solaris optimizes memory usage for maximum performance in multiprocessor systems using memory placement optimization—MPO.

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.

As multiprocessor systems become even more common, even low-end systems are designed with NUMA architectures. Oracle Solaris provides technology that can specifically help applications improve performance on NUMA architectures.

- **Memory Placement Optimization (MPO).** Oracle Solaris uses 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.
- **Hierarchical locality group 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.

Oracle Solaris Kernel Optimization

For over 20 years, Oracle Solaris kernel has 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—up to 256 MB—are used to reduce the cost of virtual to physical memory translation and increase overall system performance. 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.
- **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. Significant integration work was introduced in Oracle Solaris 11 to improve the performance of the Oracle Database stack for Oracle Solaris systems with large memory. Speed improvements to Intimate Shared Memory (ISM) and Dynamic Intimate Shared Memory (DISM) creation, locking and destruction have resulted in up to 8x start up performance improvement for the Oracle Database. Oracle Database uses DISM in its dynamic System Global Area (SGA) capability, which forms the part of the RAM shared by all processes belonging to a single Oracle Database instance.
- **Library optimization.** Oracle Solaris provides multiple implementations of common utility functions such as `memcpy(3C)`, each of which is optimized for different processors. 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 telco service throughput. Sun FlashFire products running on both 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. 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.

Integration

Oracle's product stack approach provides immediate benefits to telco companies by delivering high availability and minimizing the impact to users if failures occur. Total cost of ownership is lower

because all the software is designed to work together, making it less complicated to deploy and manage. It is also easier to reduce development and maintenance costs because telcos and SPs no longer have to support numerous skill sets required to manage multivendor solutions, and customers can take advantage of economies of scale with bundled pricing.

Oracle Solaris 11 Express offers telco customers an integrated stack, including tight integration with key components that can deliver streamlined solution environments. This includes Oracle Databases, Java environments, and complementary products such as directory, messaging, and application server.

Java Platform

Java is robust and mature computing environment that powers state-of-the-art applications. With its write once, run anywhere capability, Java powers applications that run on more than 850 million personal computers and three billion mobile devices worldwide. Telco companies can use Java to create applications and services that run on a wide variety of devices, and are easily ported to new opportunities. Users can play online games, chat with others around the world, calculate the interest on a mortgage, and view images in 3D. From laptops to datacenters, game consoles to scientific supercomputers, cell phones to the internet, Java applications can be found in virtually every type of telco platform.

Java is a key foundation of telco computing and is used in every major segment, providing a development environment that offers choice, performance, efficiency, and flexibility. Today, Java technology's versatility, efficiency, platform portability, and security make it the ideal environment for modern telco applications.

A core component of Java technology is the Java Virtual Machine. The Java HotSpot Virtual Machine is the primary Java implementation for desktops and servers produced by Oracle. It provides performance-enhancing technologies such as just-in-time (JIT) compilation and adaptive optimization. The name derives from its ability to continuously analyze the Java program's performance for code segments that are frequently or repeatedly executed. These hot spots are then targeted for optimization, leading to high-performance execution with a minimum of overhead for less performance-critical code. Java HotSpot VM is optimized for Oracle Solaris with built-in enhancements that work without requiring customization. Java HotSpot VM is available for Oracle Solaris on all platforms.

Java also provides specific telco capabilities, such as the Java API for Integrated Networks (JAIN) API specification for the SIP It is developed for the J2SE environment, and provides application developers with a standardized interface for SIP services that are functionally compatible with the RFC 3261 (SIP) specification.

Oracle JRockit is a family of Java runtime solutions available for SPARC systems for Oracle Solaris. It features industry-leading real-time infrastructure capabilities and unparalleled JVM diagnostics. Oracle JRockit Real Time offers deterministic response times for real-time Java applications on the order of milliseconds or even microseconds, suitable for the latency requirements of the most demanding telco applications. Oracle JRockit Mission Control is a set of powerful JVM diagnosis and management tools

that deliver advanced, unobtrusive, application monitoring suitable for use in development and production environments.

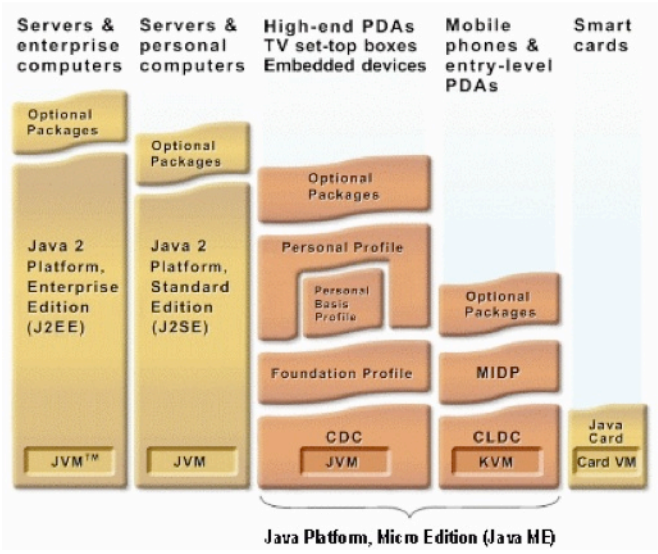

Figure 5: The Java platform offers development and delivery environments for virtually any application.

Oracle offers a choice of Java environments that are targeted at different application and service environments, from datacenters to mobile devices to media and more. Comprehensive information can be found at http://www.oracle.com/technetwork/java/

Oracle Solaris 11 Express helps Java developers create applications with superior performance, scalability, and security. It delivers a mature, secure, and stable OS with advanced development, debugging, and analysis tools and high-performance Java Virtual Machines for the most popular platforms.

Oracle Solaris Optimizations for Java

Oracle Solaris helps Java developers create telco applications with superior performance, scalability, and security. It delivers a mature, secure, and stable operating system with advanced development, debugging, and analysis tools and high-performance Java Virtual Machines for the most popular platforms. Some of the advanced capabilities of Oracle Solaris are as follows.

Oracle Solaris DTrace

An integral part of Oracle Solaris, DTrace helps speed development efforts by providing visibility into the processing of Java applications. DTrace is designed to safely instrument applications, even in production environments. It collects performance metrics by dynamically modifying the operating system kernel and user processes to record data at specific points of interest called probes. Java developers thereby gain visibility into application operation and identify trouble spots where the applications are spending too much time, helping them focus their development and debugging efforts.

DTrace can instrument any application without modifying or restarting it. It runs only the traces that are requested, analyzes the data, and delivers fast, accurate answers. Designed for system administrators and application developers, DTrace is easy to learn and easy to use, providing a C-like scripting language to save, share, and rerun tracing routines.

Instrumentation of the following JVM subsystems makes it possible for DTrace to collect performance data for applications written in the Java programming language.

- **VM Lifecycle Probes**, for VM initialization and shutdown
- **Thread Lifecycle Probes**, for thread start and stop events
- **Class-loading Probes**, for class loading and unloading activity
- **Garbage Collection Probes**, for system-wide garbage and memory pool collection
- **Method Compilation Probes**, for indicating which methods are being compiled and by which compiler
- **Monitor Probes**, for wait and notification events, plus contended monitor entry and exit events
- **Application Probes**, for fine-grained examination of thread execution, method entry/method returns, and object allocation

Enhanced Security

Modern corporate IT infrastructures are largely distributed networks linked to the internet. The connected nature of telco network operations centers, with virtually unlimited points of access, leaves them vulnerable to a wide variety of security threats—from unintentional acts that compromise security to targeted violations that expose sensitive information and corrupt data.

Oracle Solaris is widely regarded as offering some of the most advanced security features of any OS including those mentioned in the security section. The Java Cryptography Extension (JCE) provides is an integrated component of this framework, and provides encryption, key generation, key agreement, and Message Authentication Code (MAC) algorithms. When utilizing JCE, Java applications running on Oracle Solaris automatically take advantage of the crypto acceleration features available in the hardware. JCE supplements the Java platform, which already includes interfaces and implementations of message digests and digital signatures.

Developer Support

Oracle offers some of the top software development tools in the industry, including full-featured development environments designed to produce Java applications that execute as efficiently as possible. A set of powerful tools running on the HotSpot JVM or the Oracle JRockit JVM delivers advanced, unobtrusive monitoring and management suitable for use both in development and production environments.

Oracle JDeveloper 11*g*

Oracle JDeveloper 11*g* is an integrated development environment that combines features for Java, service-oriented architectures (SOA), Web 2.0, database, XML, and Web services into a single development tool. The various artifacts share the same project structure and development experience, simplifying the learning curve and development process of composite applications. As part of Oracle Fusion Middleware 11*g*, Oracle JDeveloper is hot pluggable with Oracle and non-Oracle environments, and supports most Java application servers and databases.

Oracle JDeveloper 11*g* is ranked the best IDE among major Java vendors in Forrester TechRankings. It provides a complete development environment for the latest versions of Java. Advanced code editors simplify the coding experience. State of the art debugger, code auditor, profilers, and JUnit integration ensure the delivery of high quality code. Oracle JDeveloper includes visual editors, declarative dialogs, and modelers that help create Java EE artifacts more quickly.

Oracle Solaris Studio

Developing high-performance applications requires a combination of compiler features, libraries of optimized functions, and tools for performance analysis. Oracle Solaris Studio is an integrated development platform for building C, C++ and Fortran applications and is also an invaluable tool to aid in Java application development. Oracle Solaris Studio can be used to profile and analyze Java applications, providing developers visibility into key aspects of application development, and operation.

In addition, Oracle Solaris Studio includes multilanguage debugging capabilities, including debugging native applications that call into Java APIs, and visual performance analysis tools. The primary tools included in Oracle Solaris Studio for assisting Java developers are:

- **Performance Analyzer.** Helps identify hotspots and debug Java applications by collecting and analyzing performance data including call stacks, microstate accounting information, thread synchronization delay data, hardware-counter overflow data, and memory allocation data, enabling developers to quickly identify bottlenecks and tune for optimal performance.
- **DLight.** Helps developers utilize the power of DTrace and better visualize DTrace results, providing new levels of insight into Java application operation reducing development times.

JRockit Mission Control

The JRockit Mission Control is a suite of tools designed to monitor, manage, profile, and eliminate memory leaks in Java application. In addition to monitoring and managing the JVM, JRockit Mission Control profiles the Java runtime environment with the lowest possible impact on the running system. The software also enables Java applications to run at full speed once the tool is disconnected from the JVM. This makes JRockit Mission Control suitable for use in both development and production environments. JRockit Mission Control contains three key tools:

- **JRockit Management Console.** A tool for monitoring and managing multiple JRockit instances, including dynamic control over CPU affinity, garbage collection strategy, memory pool sizes and more.

- **JRockit Runtime Analyzer (JRA).** An on-demand flight recorder that produces detailed recordings about the JVM and the application it is running. Recorded data includes profiling of methods and locks, as well as garbage collection statistics, optimization decisions, object statistics, and latency events.
- **JRockit Memory Leak Detector.** A tool for discovering, and finding the cause for, memory leaks. The trend analyzer can discover very slow leaks. It shows detailed heap statistics including referring types and instances to leaking objects, lists allocation sites, and provides quick drill down to the cause of the leak.

JRockit Mission Control is free for development use. It can be used reliably in production environments without leaving any trace in the system after it has been used, and with significantly less performance overhead than comparable tools when it is in use.

Strong Partner Ecosystem

Oracle has built a comprehensive partner community with more than 20,000+ partners, including 1000+ specialized partners. Specialized partners can be found worldwide and are those Oracle PartnerNetwork members who have met Oracle's business and competency criteria by demonstrating detailed knowledge of specific Oracle products, solutions, and services. Partners and consultants with deep technology and industry-specific application expertise can help implement carrier-grade solutions quickly and efficiently.

Sun Netra—Oracle's Carrier-Grade Servers

Designed for NEPs, carriers, and service providers, Oracle's Sun Netra Carrier-Grade server portfolio is a comprehensive line of ruggedized, NEBS-certified SPARC and x86 blade and rackmount servers supporting Oracle Solaris and other OSes. These servers, when properly configured, deliver five 9s availability for environments that require uninterrupted operations, and come in an optimized form factor for telco industry needs. The Netra portfolio offers extended product life cycles, which translates into significant cost savings for telco central offices and network datacenter deployments.

Oracle offers a broad line of Sun Netra servers, featuring both SPARC and x86 processors, in blade and rackmount configurations. Processor options include the newest SPARC T3 processors, UltraSPARC T2 and UltraSPARC T1 processors, Intel Xeon, and AMD Opteron. The following contains a brief overview of Sun Netra servers for the communications market; complete product line information can be found on oracle.com.[8]

[8] http://www.oracle.com/us/products/servers-storage/servers/netra-carrier-grade/

Sun Netra Rackmount Servers

Oracle offers a comprehensive line of Netra carrier-grade rackmount servers. These NEBS-certified systems enable telcos to address the widest range of computing requirements, and feature x86 and SPARC processors, including the new SPARC T3 processor. Sun Netra SPARC T3 servers are designed to support a multicore, multithreaded server on a chip (SoC) design that vastly increases reliability and performance as it minimizes cost, power demand, and components required—reducing operating costs. The Netra SPARC T3-1 is the world's first 16-core processor based system with unsurpassed throughput. This carrier-grade system is designed for network infrastructure applications such as media gateway controller, OSS/BSS, IP traffic management and BRM. It meets increasing demands from the telco networks and provides an ideal platform for virtualization and consolidation, offering 512 virtual systems in a single server.

Scaling with threads rather than frequency, the rackmount design packs performance by leveraging SPARC innovation such as 16 SPARC cores that concurrently run eight threads per core; a memory access crossbar; integrated PCI Express (PCIe) features and PCI-X; a separate floating point unit for each core; an L2 cache; and two to four independent, dual-channel memory controllers.

Ruggedized NEBS Level 3 certified, Netra SPARC T3 servers provide a high level of system reliability which helps ensure that the server continues to operate under extreme environmental conditions. Redundant hot-swappable AC or DC power supplies and hot-pluggable hard disk drives further enhance the system's uptime. In addition, the Netra SPARC T3 servers are offered with the Sun Integrated Lights Out Manager (ILOM), which enables simple remote monitoring and management from anywhere on the network. The Netra SPARC T3-1 system is designed for performance and feature PCIe 2.0 slots, 6 Gbps SAS2 drive interfaces, up to 512 GB of DDR3 RAM, and four onboard gigabit ports. Finally, the Netra SPARC T3 servers support Oracle Solaris, providing guaranteed binary compatibility with earlier Netra SPARC servers, preserving your investment.

Oracle also offers Netra rackmount systems equipped with x86 processors, including Intel Xeon and AMD Opteron processors. These systems support up to eight processing cores in 2U-high 20 inch-deep carrier-grade enclosures. As with the SPARC-based systems mentioned above, the Netra x86 rackmount servers are offered with the Sun Integrated Lights Out Manager (ILOM), which enables simple remote monitoring and management from anywhere on the network.

Sun Netra ATCA Blades Servers

Oracle's Sun Netra CT900 ATCA blade server is a high performance blade platform designed for the most demanding telecom networks. The Sun Netra CT900 Server is a NEBS Level 3-certified, ETSI compliant, rack-mountable, -48v powered ATCA blade system that supports PICMG 3.1 Options 1 and 9. As an integrated platform, it enables some of the most highly available applications for the telecommunications market. Broadly supported in the marketplace, the ATCA PICMG organization has more than 450 supporting companies delivering products and participating in the standard. Sun Netra CT900 ATCA blade server offers an AMC and ARTM for I/O expansion, and a CompactFlash slot to increase user flash support. Many configurations support 10 Gigabit XAUI Ethernet interfaces to the Extended Fabric to support a redundant dual star network topology. The Sun Netra CT900

ATCA blade server chassis can mix and match up to 12 of Oracle's SPARC or x86 processor-based blades, running Oracle Solaris and other operating systems, all in the same enclosure.

Sun Netra SPARC ATCA Blade Servers

Sun Netra SPARC ATCA blade servers are available with the new SPARC T3 as well as UltraSPARC T2, and UltraSPARC T1 processors. The Netra ATCA server blades can be plugged into any of the user node slots of the Sun Netra CT900 ATCA blade server, or any ATCA/PICMG 3.0- and 3.1-compliant chassis. Oracle's new Netra SPARC T3-1BA ATCA blade server is based on the SPARC T3 processor with 12 cores and 96 threads running at 1.4 GHz. The Netra SPARC T3-1BA ATCA blade server has a balanced I/O architecture through the use of new AdvancedRTM technology, which provides high-speed x8 PCIe, 10 Gigabit Ethernet, Gigabit Ethernet, and serial connections to Zone 3 in a standardized way. It can easily handle numerous workloads, including Control Plane, application server computing, Data Plane processing, using Sun Netra Data Plane Software, media processing using mediaLib with eight floating-point units, and security processing using the built-in cryptographic units in each core. It supports the basic PICMG 3.0 system management features, functioning with the same compatibility and availability of other ATCA-compliant blades.

Sun Netra x86 ATCA Blade Servers

The Sun Netra x86 ATCA blade servers combine cutting-edge performance with the ruggedness and reliability of the Sun Netra carrier-grade server family. The new Sun Netra CP3270 ATCA blade servers, based on the Intel Xeon processor LC5500 series with Hyper-Threading and Turbo Boost Technologies are NEBS Level 3 certified, ETSI compliant, and PICMG Compliant. The Sun Netra CP3270 ATCA blade server supports eight CPU cores and offers I/O compatibility, and power and thermal similarities with the previous generation blade, making upgrades a very simple procedure. The Sun Netra CP3270 ATCA blade server offers an AMC and ARTM for I/O expansion, a CompactFlash slot to increase user flash support, and eight DIMM sockets for large memory applications. AMC and ARTM I/O slots enjoy broad third party support for a variety of I/O cards.

Sun Netra x86 blade servers support Oracle Solaris and other operating systems. They are also available with AMD Opteron processors. All Netra ATCA server blades based on x86 or SPARC processors can be plugged into any of the user node slots of the same Sun Netra CT900 ATCA blade server.

Sun Netra Enterprise Blade Servers

The Sun Netra 6000 Modular System leverages the benefits of the Sun Blade 6000 modular system, which has been proven in the market over the past several years. This NEBS-certified blade system chassis fits in a compact form factor—only 10 RU (11 RU for DC)—while supporting up to 10 full-featured, top-performance blade server modules. The modular design enables easy upgrade and expansion, high flexibility with a breadth of compute, storage, and networking options. The externally accessible, hot-pluggable ExpressModules give each blade a unique I/O personality, while the network switch options dramatically help cut cabling, simplify inter-chassis storage and Ethernet connectivity,

and redundancy reduces downtime. The Sun Netra 6000 Modular System is designed for telecom customers who need a cost-efficient service delivery platform for value-added network, media services delivery, and Operations and Business Support Systems (OSS/BSS) deployments. Based on the Sun Netra 6000 chassis, it can be combined with the UltraSPARC T2+ or Intel Xeon processor 5600 series server modules to handle the most demanding workloads, including multithreaded Web applications and advanced IP-based telecom/Web services.

Billing and Revenue Management

Oracle Communications Billing and Revenue Management (BRM) on Sun Netra servers running Oracle Solaris provides a high-performance, scalable, and secure solution for enterprises using Oracle 11*g* and Oracle 11*g* Real Application Cluster (RAC). This highly scalable solution helps service providers accelerate the launch of new services; enable monetization of services for any customer type, network, service, business model or geography.

The Oracle Communications BRM application suite is a Tier 1, convergent billing framework that is architected for leading service providers. The BRM solution is a fully convergent prepaid and postpaid solution and employs an architecture that allows convergent modules to co-exist and offer integrated services.

Oracle Communications BRM has been designed and developed with parallelism in mind that enables several operations to utilize hardware strands concurrently, and is well suited to take advantage of SPARC processors, which provide up to 16 cores with each core supporting eight threads, and up to four sockets per system—512 total processing threads. This architecture allows to split the rating and loading in many threads resulting in high throughput—the rating pipeline can be configured to use the power of 512 hardware threads. The billing part can also be parallelized based on configurable parameter. Sun Netra servers handle large memory supported by Oracle Solaris very efficiently.

Oracle's BRM solution has been integrated into the TimesTen In-Memory Database, enabling real-time billing capabilities on both SPARC and x86 systems.

Additional Oracle Products and Services

A compelling benefit for Oracle Solaris users is the application-to-disk integration, optimizing performance, reliability, and security only Oracle can provide for your complete solution stack. This helps explain why Oracle Solaris is the dominant deployment platform for Oracle Database. Over 40 percent of Oracle Database customers are using Oracle Solaris, and over 11,000 applications are supported on Oracle Solaris. While there are many reasons for this broad market acceptance, a reliable platform is a big factor in the equation.

Only Oracle offers packaged software solutions that deliver end-to-end support for the key business processes, from service creation, offer management, and order orchestration, through provisioning and service delivery, to billing, revenue assurance, and reporting. And only Oracle offers the widest choice of enterprise and carrier-grade software applications, middleware, database technology, and decision-

support tools for the communications industry, with the combined strengths of PeopleSoft, Siebel, Portal Software, MetaSolv, TimesTen, and more.

With Oracle's unique portfolio of solutions for the communications industry, organizations can rapidly create, market, sell, fulfill, deliver, and bill for next-generation services and content, increase customer satisfaction and loyalty, reduce operating costs in the business and the network, and improve management reporting and control. Oracle offers many complementary products and services for telco service providers and NEPs. They are described as follows.

TimesTen

Oracle TimesTen In-Memory Database is a memory-optimized relational database that empowers applications with the instant responsiveness and very high throughput required by today's real-time enterprises and industries such as telecom, capital markets, and defense. Deployed in the application tier as an embedded or standalone database, Oracle TimesTen In-Memory Database operates on databases that fit entirely in physical memory using standard SQL interfaces. It can also be used as an in-memory database cache for the Oracle database to enhance the response time and throughput of user applications. Oracle TimesTen databases are persistent and recoverable. Durability is achieved through a combination of transaction logging and database checkpointing to disk.

Oracle Directory Services Plus

Directory Services provide the foundation for identity-enabled telco applications and the underlying identity management architecture. Oracle Directory Services Plus is the only integrated solution that provides a comprehensive set of directory solutions for high performance enterprises and carrier-grade environments.

Oracle Directory Services Plus is a single package comprised of:

- **Oracle Directory Server Enterprise Edition (formerly Sun Java Directory Server Enterprise Edition).** The market leading directory server that provides high performance LDAP services, ideally suited for heterogeneous environments.
- **Oracle Internet Directory.** LDAP compliant directory built upon the high performance Oracle Database and deeply integrated with Oracle applications and middleware. It is ideally suited for Oracle environments.
- **Oracle Virtual Directory.** Virtually aggregates identity information from multiple sources and presents a real-time unified view, eliminating the need to synchronize or move identity data.

Oracle Enterprise Manager 11*g* 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.

Oracle Carrier-Grade Framework

The Oracle Carrier-Grade Framework (CGF) is designed to accelerate the adoption of new networking infrastructures and applications by providing integrated, highly available and high performance software components into a framework that meets the stringent demands found in the telco industry. With Oracle Carrier-Grade Framework, communications service providers and telco equipment manufacturers can achieve critical business and time-to-market goals with lower risk.

Oracle Carrier-Grade Framework is the leading application-ready, future-proof foundation for next-generation networks and IT infrastructure. Oracle Carrier-Grade Framework is designed to accelerate the adoption of new networking infrastructures and applications by providing integrated, highly available and high-performance software components to a framework that meets the stringent demands of the telecommunications industry.

Oracle Real Application Clusters

Losing a mission-critical application, such as a billing system, can cost you tens of thousands of dollars every minute. With Oracle's unique server-clustering technology, Oracle Real Application Clusters, you can run a single database on dozens of parallel servers in a grid, making your applications unbreakable. With full server redundancy and failover times measured in subsecond, Oracle Carrier-Grade Framework delivers near 100% availability to meet carrier-grade requirements. The technology also allows you to scale capacity by adding or removing servers, enabling you to load balance for optimum server utilization.

Oracle Communications Service Delivery

The Oracle Communications Service Delivery product family is a portfolio of carrier-grade, open, standards-based service delivery platform (SDP) products, designed to help network operators and service providers by:

- Maximizing profitability from existing communication services
- Establishing new two-sided business models via service exposure and application stores
- More effectively monetizing unique assets in legacy fixed and mobile intelligent networks (IN), broadband and cable networks, as well as next-generation networks (NGN), such as Long-Term Evolution (LTE), WiMAX, and IP Multimedia System (IMS)

The principles of SOA are incorporated into the Oracle Service Delivery Platform, the industry's most comprehensive, open standards-based software environment for creating, delivering, and managing convergent communications and multimedia services in a fast, flexible, and efficient manner. Based on J2EE, and including prebuilt elements such as Voice over IP (VoIP), presence, and Virtual PBX, Oracle Service Delivery Platform enables you to create services quickly, get to market faster, and start profiting sooner.

The Oracle Communications Service Delivery product portfolio consists of the following:

- **Oracle Communications Converged Application Server** provides Voice over IP (VoIP) / Voice over LTE (VoLTE), multimedia conferencing, IP Centrex, IMS, and Next-Generation IN services.
- **Oracle Communications Services Gatekeeper** is an open, third-party service exposure platform designed to enable operators to better monetize their network assets by leveraging content and applications from third-party developers and partners. It is based on leading industry standards including Java EE, SIP/IMS, Parlay X, Web 2.0 (REST), SOA, and Web services.
- **Oracle Communications Marketing and Advertising** is a centralized marketing and advertising campaign management and execution platform that enables operators to offer partners and advertisers highly targeted and personalized marketing and ad campaigns.
- **Oracle Communications Service Broker**, which includes open standards-based, highly scalable and available carrier-grade service interaction and mediation capabilities that enable service interaction across diverse network types and orchestrate multiple services in real-time, enabling the creation of innovative blended services.
- **Oracle Communications Presence** is an integrated service-enabler platform providing presence, XML Document Management (XDM), and message waiting indication (MWI) capabilities to converged Web-telco applications.

The Oracle SDP approach is to enable operators to implement a unified SDP architecture, bringing together existing vertical, proprietary SDP initiatives into a horizontal, unified open, standards-based SDP framework, allowing operators to continue monetizing services in existing legacy IN platforms, and cross-leverage future investments into 4G (such as LTE, WiMAX), Application Stores (including network exposure, third party developers, multimedia content delivery), NGIN (including service mediation across IN-IP-Web, 2G/3G/4G-fixed-Web), and Online/Real-Time Charging, among many others. Oracle Communications Service Delivery products, with Oracle Communications Consulting and our large, global ecosystem of SI, ISV, NEP and solution partners, are delivering open, unified and future-proof SDP implementations worldwide, resulting in profitable monetization of networks through commercially successful two-sided business models and innovative converged services.

Oracle Communications Service Delivery products have been commercially deployed by customers worldwide, for a broad and comprehensive set of SDP solutions. SDP solutions enabled by Oracle Communications Service Delivery products encompass Network-Enabled Application Stores, Policy-Enforced Content Delivery Platforms, NGIN, IMS VAS (such as IMS Calling, IMS RBT, Voice-over-IMS), Hosted Enterprise Conferencing, IP Centrex and Hosted Unified Communications.

Summary

Oracle Solaris has continued to demonstrate great success as a carrier-grade OS for scalable performance, advanced reliability, and virtualization. The combination of Oracle Solaris on innovative Sun Netra servers, combined with complete, integrated solutions from Oracle, offers the IT infrastructure you need to meet telco industry challenges and requirements. Continuing efforts at application-to-disk integration with Oracle products and technology 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. The new IPS works with existing Oracle Solaris features to streamline patches and upgrades, and assure proper deployment of software images. No-cost virtualization provides fine-grain, end-to-end technology to efficiently consolidate your operations center applications. Pervasive security capabilities help lock down your systems while enabling your people and processes to do their job.

These benefits help telco companies reduce capital and operational costs and enable improved margins and ROI. A leader in several segments, Oracle provides the technology, expertise, and products to support telecoms and help them achieve and maintain leadership.

Resources

Oracle Solaris 10	www.oracle.com/technetwork/server-storage/solaris/overview
Oracle Solaris 11 Express	www.oracle.com/technetwork/server-storage/solaris11/overview/
Oracle Solaris Cluster	www.oracle.com/technetwork/server-storage/solaris-cluster/overview/
Oracle Solaris Studio	www.oracle.com/technetwork/server-storage/solarisstudio/overview/index.html
Oracle Solaris 10 Downloads	www.oracle.com/technetwork/server-storage/solaris/downloads/
Oracle Communications Solutions	www.oracle.com/us/industries/communications/
Oracle Communication Products	www.oracle.com/us/industries/communications/046718.html?ssSourceSiteId=opn
Oracle Solaris Security	www.oracle.com/us/products/servers-storage/solaris/security/
Oracle Solaris Common Criteria	www.oracle.com/technetwork/topics/security/oracle-cc-evalsolaris-083233.html
Oracle Solaris Virtualization	www.oracle.com/technetwork/server-storage/solaris/overview/virtualization-163570.html
Sun Netra Servers	www.oracle.com/us/products/servers-storage/servers/netra-carrier-grade/

Authors and Contributors

The following people contributed to this white paper: Chris Baker, Mark Butler, Joshua Grossman, Duncan Hardie, Joost Pronk Van Hoogeveen, Pete Salerno, Larry Wake, and Terri Wischmann.

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

Oracle Solaris: The Carrier-Grade OS
February 2011

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.