ORACLE
SPARC SERVERS

Oracle's SPARC T7 and SPARC M7 Server Reliability, Availability, and Serviceability

ORACLE WHITE PAPER | AUGUST 2016

ORACLE

Table of Contents

Introduction

Modern application users and businesses alike have a low tolerance for downtime, and enterprise IT infrastructure must meet application availability requirements that match business needs. For some applications, a certain amount of downtime might be acceptable. For others, any amount of downtime can cost enormous sums of money, damage corporate reputation, and cause lost opportunities. Organizations need assurances that the systems they deploy can meet stringent requirements that allow them to meet service level agreements (SLAs) and provide continuous services in line with stated recovery point objectives (RPOs) and recovery time objectives (RTOs).

In this context, reliable and serviceable systems are critically important. However, true application availability requires a completely integrated hardware and software stack along with system components that are designed to work together. To succeed, organizations need an *availability architecture* that combines robust and redundant systems and storage technologies with reliable operating systems, middleware, virtualization technology, and databases. Integrating all of these technologies to function seamlessly together can be time consuming and expensive, and the slightest mismatch can result in expensive and catastrophic failure and downtime.

With its SPARC M7 processor–based servers, Oracle is in a unique position to deliver the most advanced system architecture with a completely integrated technology stack. These servers provide a high level of reliability, availability, and serviceability (RAS) features that are the foundation for building the best and most secure mission-critical compute system. The integrated stack includes the Oracle Solaris operating system and Oracle's advanced middleware and database technologies, which combine virtual and physical system elements to create a platform that has the best possible RAS on which to run applications. In addition, Oracle Optimized Systems and Oracle Maximum Availability Architectures make use of the platform's unique RAS features to provide a tested, scalable, and highly available infrastructure that is based on proven best practices.

Deploying a RAS Architecture

It is common to focus solely on system reliability features when designing highly available IT infrastructures. In fact, delivering a high-quality application service relies on RAS functionality throughout the entire hardware and software stack. Although the ultimate focus is on application availability, reliability and serviceability are equally important concepts when considering the capabilities of a delivery platform or data center solution. The following are some things to consider:

- **Reliability**. There are numerous ways to increase reliability. For example, the number of components in a system can be reduced or minimized. A reduction in the number of components and, by extension, a reduction in the number of connections between components, induces a higher mean time between failures (MTBF). Implementing data integrity safeguards for all in-rest and in-transit data can dramatically improve reliability by allowing the system to automatically detect and correct errors before they have the chance to corrupt data and cause applications or systems to fail.
- **Availability**. Availability is typically defined as the ability to offer undisturbed application access during underlying component failures. This is typically achieved through redundancy of the platform, application availability features, virtualization, and clustering. Data availability includes mechanisms such as the DIMM sparing capability provided in SPARC M7 processor–based servers as well as memory lane sparing and processor interconnect lane sparing. Data availability is also increased by software such as I/O multipathing in Oracle Solaris or Oracle Solaris ZFS, which provides advanced RAID protection for disk devices. In many cases, availability is provided at the application level itself, but in the absence of this, platform availability can be further augmented by using clustering software that provides failover between virtual or physical systems.
- **Serviceability**. A high level of serviceability allows systems to be serviced or repaired with minimal downtime, improving availability. Serviceable components are either hot-swappable or hot-pluggable while the system remains operational. Hot-swappable components are the least disruptive because their replacement requires no assistance or preparation from the host operating system and server. Hot-pluggable components can also be swapped out while the system continues to operate, but they require that the operating system and platform be prepared before a component is removed or inserted.

Any application availability architecture is only as strong as its weakest link. Comprehensive design actually increases reliability by combining the complementary RAS features of a system's hardware and software layers into an availability architecture (Figure 1). In this way, each element builds on the strengths of the layer below to enhance the overall RAS capability within a complex system. An extremely robust system architecture can be delivered by utilizing the strengths and mitigating the vulnerabilities of the components at each layer.

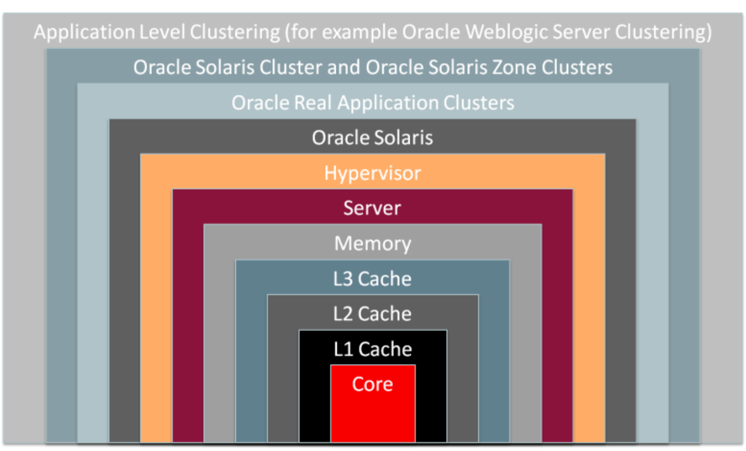

Figure 1. An availability architecture that combines the RAS capabilities of hardware and software elements at each layer of the system.

System software layers contribute enormously to RAS functionality. For example, the predictive self-healing capabilities that are implemented in the Oracle Solaris operating system and Oracle Integrated Lights Out Manager (Oracle ILOM) firmware as part of the Oracle Solaris Fault Management Architecture (FMA) are essential elements for ensuring high availability for hardware and application software. In addition, application, middleware, and clustering software, including database clustering, can all be employed to further increase the RAS levels. The following are few examples of technologies that are available with Oracle's SPARC servers:

» Oracle's layered virtualization technologies can help isolate system images—providing redundancy and hardware independence, and facilitating easy application mobility between systems.
» Oracle Solaris Cluster software provides system-level redundancy and failover between and physical or virtual systems.
» Oracle WebLogic Server provides a mission-critical cloud platform for applications requiring high performance, scalability, and reliability.
» Oracle Real Application Clusters (Oracle RAC) provides a clustered database with a shared cache architecture to enable highly scalable and available database solutions.

It is important to note that some features that provide RAS capabilities represent a trade-off with other features. For example, integrating multiple features into a single chip or surface-mounting all chips and ASICs on a single motherboard delivers higher reliability, but can cause reduced availability (due to mean time to repair) as well as reduced serviceability. Moreover, hot-swappable and hot-pluggable components can require additional connectors—improving availability but technically reducing reliability. As such, feature selection is always a careful design process, because there is a balance between system cost and the availability and reliability of data services.

SPARC M7 Processor–Based Servers: Availability by Design

SPARC servers have been designed from the outset to serve demanding enterprise environments and host mission-critical applications. Numerous RAS features are designed into these servers, ranging from advanced error detection and correction mechanisms to processor and memory protection and redundant components. These SPARC servers and Oracle Solaris include an architecture for deploying services that are capable of predictive self-healing. The Oracle Solaris FMA receives data related to hardware and software errors, automatically diagnoses the underlying problem, and responds by taking faulty components offline when possible.

Reliability, Availability, and Serviceability Features

SPARC servers provide a host of reliability features, extending from the processor, to the hypervisor, to the system as a whole. These features complement and strengthen each other, lending greater resiliency to the system overall. RAS features engineered into SPARC M7 processor–based servers include the following:

» Oracle's SPARC M7 processor
    » L1 cache tag, status and data: Parity protection and retry on error
    » L2 and L3 cache data: SEC/DED protection, inline correction and cache-line sparing
    » L2 and L3 cache status and directory: SEC/DED protection and inline correction
    » Architectural registers: SEC/DED protection, precise trap and hypervisor correction/retry
    » Message retry in hardware
    » Dynamic processor core deconfiguration
    » Independently managed power for each quadrant of the processor
    » Link-level RAS: Automatic frame retry, automatic link retrain, link-level multipathing, and single-lane failover
    » Dynamic voltage and frequency scaling

- Main memory
  - SDRAM error protection: Correction of single-device errors and triple-error detection across devices
  - Memory channel interconnect: Message retry in hardware, lane retire in hardware, and cyclic redundancy check (CRC)
  - DIMM sparing allows full memory DIMMs to be retired to prevent hard errors. If persistent recoverable errors occur, the FMA functionality can initiate DIMM sparing and stop using the DIMM that has questionable reliability. DIMM sparing is automatic and causes no interruption to application services. The system memory capacity is unchanged, error protection remains intact with no increased exposure to failures after DIMM sparing is performed.
  - If persistent recoverable errors are localized to a specific address, the system will use “page retire” functionality to prevent hard errors. This functionality remains available even after DIMM sparing has been applied to the same memory bank.
- Fault Management Architecture
  - Diagnosis engine on SP and Oracle Solaris
  - Auto reconfigure on failure
  - Soft error rate discrimination (SERD)
  - Bad page retirement
  - Operating system and SP watchdogs
- SPARC M7 hypervisor
  - Logical domains (LDoms) virtualization and failure containment
  - Dynamic assignment of CPU, memory, IO
  - Dynamic PCIe bus assignment
  - Processor support for error clearing, correction and collection
- System
  - Diagnosis to the level of a field-replaceable unit (FRU) on first fault
  - Redundant hot-serviceable service processors (SPs) and service processor proxies (SPPs) with automatic failover (for SPARC M7-8 and M7-16 servers)
  - Hot-pluggable internal 2.5-inch SAS hard-disk drives (HDDs) and SAS/NVMe solid-state drives (SSDs) (for SPARC T7-1, T7-2, and T7-4 servers)
  - Support for multiple physical domains (PDoms) (for SPARC M7-8 and M7-16 servers)
  - Independent clock sources and redundant clocking (for SPARC M7-8 and M7-16 servers). Each CMIOU board and switch board have dual clock sources and capable to recover from failure without an interruption to application services, and continue to operate with a single clock source.
- System I/O
  - PCIe end-to-end CRC
  - PCIe link retry
  - Up to five PCIe root complexes per I/O controller for redundant root domain and I/O multipathing
  - CPU-to-I/O Controller link lane sparing
  - Hot-pluggable low-profile PCIe cards (in SPARC T7-4, M7-8, and M7-16 servers)
- Power and cooling
  - Advanced power management
  - Redundant hot-swappable fans and fan modules
  - Redundant hot-swappable power supplies
  - Dual-grid power
  - Redundant CPU core power supply (in SPARC M7-8, and M7-16 servers)

Error Detection, Diagnosis, and Recovery

SPARC M7 processor–based servers feature important technologies that correct errors early and keep marginal components from causing unplanned downtime. Architectural advances that inherently increase reliability are augmented by error detection and recovery capabilities within the server hardware subsystems. Ultimately, the following features work together to raise application availability:

» End-to-end data protection detects and corrects errors throughout the system, ensuring complete data integrity.

» State-of-the-art fault isolation helps the server isolate errors within component boundaries and offline only the relevant item, chip, or component subsection. Isolating errors down to the smallest possible entity improves stability and provides the continued availability of maximum compute power. This feature applies to processors, memory buffer ASICs and DIMMs, switch ASICs, connectivity links, I/O controller ASICs, and SPs.

» Constant environmental monitoring provides a historical log of pertinent environmental and error conditions.

» The host watchdog feature periodically checks for the operation of software, including the operating system for each domain. This feature also uses the Oracle ILOM firmware to trigger error notification and recovery functions.

» The FMA capability of the system and dynamic CPU resource deconfiguration of the SPARC M7 processor enable powerful isolation and recovery. If necessary, the system can dynamically retire processor resources (for example, a core) without interrupting the applications that are running.

» The system performs periodic component status checks to detect signs of an impending fault. Recovery mechanisms are then triggered to prevent system and application failure.

» Error logging, multistage alerts, electronic FRU identification information, and system fault LED indicators contribute to rapid problem resolution.

Redundant Components

SPARC servers feature redundant, hot-swappable power supply and fan units, as well as the option to configure multiple storage devices, memory DIMMs, and I/O cards. All SPARC M7 processor–based servers except the SPARC T7-1 server support multiple SPARC M7 processors. Administrators can combine redundant hot-pluggable internal storage with RAID hardware and software to enhance data availability. SPARC M7-8 and M7-16 servers also support redundant, hot-swappable SPs and SPPs. If a fault occurs, these duplicated components can support continued operation. Depending upon the component and type of error, the system might continue to operate in a degraded mode or it might reboot—with the failure automatically diagnosed and the relevant component automatically configured out of the system. In addition, hot-serviceable hardware within these servers speeds service and allows for the simplified replacement or addition of components, without a need to interrupt or stop the system.

Replaceable components fall into the following categories:

» **Hot-serviceable** components can be removed and inserted while the server is running.

  » Hot-swappable components do not require any preparation prior to servicing.

  » Hot-pluggable components require preparation prior to servicing.

» **Cold-serviceable** components require that the server be shut down for service. In addition, some service procedures require that the power cables be disconnected between the power supplies and the power source.

Table 1 lists key hot-serviceable components in SPARC M7 processor–based servers.

TABLE 1. KEY CUSTOMER HOT-SERVICEABLE COMPONENTS IN SPARC M7 PROCESSOR–BASED SERVERS

Component	SPARC T7-1 Server	SPARC T7-2 Server	SPARC T7-4 Server	SPARC M7-8 Server	SPARC M7-16 Server
2.5-inch small form factor (SFF) drive	✓	✓	✓	NA	NA
Power supply	✓	✓	✓	✓	✓
Fan/fan module	✓	✓	✓	✓	✓
PCIe card	–	–	✓ [1]	✓	✓
CPU, memory, and I/O unit (CMIOU) board	NA	NA	NA	✓ [2]	✓ [2]
Switch unit (SWU)	NA	NA	NA	NA	✓ [3]
SP/SPP	–	–	–	✓	✓

✓ = Hot-serviceable, – = Gold-serviceable, NA = Not applicable

1. PCIe card in the card carriers. PCIe switch card in internal slots are cold-serviceable.
2. The PDom that the CMIOU board is in must be shut down. Other PDoms can be operational.
3. The SWU must be deconfigured, and the system must have been booted without the SWU. All PDoms can be operational during the SWU hot-service. Reconfiguration to include the SWU requires rebooting all the PDoms.

Multiple PCIe Root Complexes

SPARC M7 processor–based servers share the same basic I/O subsystem design. Each SPARC M7 processor is connected to one or more I/O controller ASICs via I/O links. The PCIe infrastructure is provided by the I/O controller ASICs, with each ASIC supplying five PCIe 3.0 root complexes. The PCIe 3.0 subsystem implementation varies by server model, but every SPARC M7 processor–based server offers multiple expansion slots with dedicated root complexes. This capability enables flexible I/O virtualization with separate root domains providing redundant, highly available network and storage connectivity.

DIMM Sparing

Each SPARC M7 processor supports up to 16 DDR4 memory DIMMs via eight buffer-on-board (BoB) ASICs. Up to 1 TB of memory is supported per processor with sixteen 64 GB DIMMs. The physical address space provided by the memory DIMMs and controlled by an individual SPARC M7 processor is interleaved to maximize performance. Half-populated memory configurations are eight-way interleaved, and fully populated configurations with16 DIMMs per processor are 16-way interleaved. The SPARC M7 processor also supports a 15-way interleaved configuration, and the switch from 16-way to 15-way interleaving can be done dynamically. This capability is the basis of a new feature called DIMM sparing (Figure 2), first introduced with the SPARC M7 processor–based servers.

D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13	D14	D15	D16	DIMM Number
94%	94%	94%	94%	94%	94%	94%	94%	94%	94%	94%	94%	94%	94%	94%	94%	% Used

DIMM #7 retired

D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13	D14	D15	D16	DIMM Number
100%	100%	100%	100%	100%	100%	0%	100%	100%	100%	100%	100%	100%	100%	100%	100%	% Used

Figure 2. DIMM sparing provides for automatic failover of a single DIMM, with no interruption of service.

DIMM sparing allows removing one of the 16 DIMMs, and remapping its content using the full capacity of the remaining 15 DIMMs. The process works by leaving 1/16 of the capacity of each DIMM unused in the initial configuration. A bad DIMM can then be retired, and its content can be remapped into the remaining 15 DIMMs.

DIMM sparing is done automatically when a DIMM is determined to be faulty, with no interruption of service. Through this process, the system memory capacity is unchanged and error protection remains intact after DIMM sparing is performed. The system simply continues to run with no loss of capacity and no increased exposure to failures. Consequently, there is no need to take the system down for service. The actual DIMM replacement process can wait until a subsequent failure in the same memory bank, which generates a serviceable fault.

DIMM sparing is enabled in SPARC M7 processor–based servers in fully populated memory configurations (16 DIMMs per processor). DIMM sparing is not supported in half-populated memory configurations. While not recommended, it is possible to disable DIMM sparing in fully populated memory configurations.

Redundant Service Processors

Platform management for all SPARC M7 processor–based servers is done via the service processor (SP) using Oracle ILOM. Oracle ILOM provides a command-line interface (CLI), a web-based graphical user interface (GUI), and Intelligent Platform Management Interface (IPMI) functionality to aid out-of-band monitoring and administration. Oracle Enterprise Manager Ops Center provides the management capability, communicating with Oracle ILOM to manage and monitor the SPARC M7 processor–based servers. All system telemetry and health diagnostics are recorded by Oracle ILOM and forwarded to Oracle Enterprise Manager Ops Center for further analysis and action. If a service event is deemed necessary, Oracle Enterprise Manager Ops Center works with Oracle Auto Service Request to notify the service that action needs to be taken.

In keeping with their focus on mission-critical enterprise workloads, SPARC M7-8 and M7-16 servers feature redundant hot-pluggable SPs with automatic failover. This configuration makes it possible for a SP to fail without impacting the ability to monitor and manage the system. The same SP hardware is also used as SP Proxies (SPPs) in the SPARC M7-16 server, which has a hierarchical SP infrastructure. This architecture provides redundant and isolated hardware management support regardless of how many physical domains are configured in the server.

For reliability and continuous operation, SPs and SPPs are always configured as a redundant pair with active/standby failover. The network between a pair of SPs or SPPs facilitates the exchange of system management information. In case of failover, the standby SPs or SPPs are always ready to become active. One or more Service Processor Modules (SPMs) are located on each of the SPs or SPPs. SPM configuration varies by server model, with each PDom connected to a redundant SP or SPP pair, depending on the server model:

» On a SPARC M7-8 server with a single PDOM, each of the two SPs has only one SPM to manage the single PDom in the system. One of the SPs functions as the active SP to manage the platform, and the other acts as a standby SP that assumes the active SP role in the event of a failure.

» On a SPARC M7-8 server with two PDoms, each of the redundant SPs has two SPMs, and the server uses two pairs of SPMs to manage the two PDoms in the system. The first SPM pair provides the active SP and standby SP functionality and manages CMIOU0 to CMIOU3. The second SPM pair manages CMIOU4 to CMIOU7. The two SPM pairs work together to provide a single management environment to all of the server components.

» On the SPARC M7-16 server, each SP on the switch chassis has only one SPM, and the two SPPs on each CMIOU chassis have two SPMs. This configuration is necessary because the system is capable of supporting up to four PDoms (Figure 3). One of the SPPs on each PDom is identified as a PDom-SPP, and it is responsible for managing tasks for the PDom and hosting the remote keyboard, video display, mouse, and storage devices (rKVMS) service for the PDom. PDom-SPPs are also accessible from the external network, as long as they have been configured with IP addresses.

Figure 3. The SPARC M7-16 server can be flexibly configured with one, two, three, or four PDoms (three PDoms shown), each with two SPPs for active/passive failover.

Oracle Solaris

SPARC M7 processor–based servers support both Oracle Solaris 10 and Oracle Solaris 11—with the appropriate level of patching. Oracle Solaris 11.3 is required in the control domain and for bare-metal operations. Oracle Solaris 10 9/10 or later is supported in guest domains. Oracle Solaris includes the following features that strengthen the reliability and availability of the SPARC M7 processor–based servers:

» **Fault Management Architecture**. The Fault Management Architecture (FMA) capabilities improve availability by automatically diagnosing faults in the system and initiating self-healing actions to help prevent service interruptions. This software helps increase reliability by configuring problem components out of a system before a failure occurs. In the event of a failure, the feature initiates automatic recovery and application restart using the Oracle Solaris Service Management Facility feature.

  The FMA diagnosis engine produces a fault diagnosis once discernible patterns are observed from a stream of incoming errors. Following diagnosis, the FMA provides fault information to agents that know how to respond to specific faults. While similar technology exists from other hardware and software vendors, each of their solutions is limited to either software- or hardware-based fault detection. Oracle's complete software stack provides a fault detection and management environment that allows full integration of hardware fault messages to be passed on to the operating system services, which can then adjust hardware resources as needed so that outages are greatly reduced.

  With Oracle ILOM running on the server's SP, the system supports initiation of FMA events to alert the administrator of platform faults. The SP has a diagnostic engine and can generate FMA events independently from Oracle Solaris. The SP sends the events to the control domain, so the Oracle Solaris guest can take action on processor or memory faults. Oracle Solaris will offline cores, threads, and memory pages that generate an inordinate number of correctable errors. Taking resources offline before they generate a fatal fault helps ensure increased application service uptime. The integration of FMA on the SP is an availability advantage of the SPARC M7 processor–based servers over comparable platforms from other vendors.

» **Service Management Facility**. With the Service Management Facility of Oracle Solaris, system administrators can use simple command-line utilities to easily identify, observe, and manage the services provided by the system as well as the system itself. The Service Management Facility describes the conditions under which failed services may be automatically restarted. These services can then be automatically restarted if an administrator accidentally terminates them, if they are aborted as the result of a software programming error, or if they are interrupted by an underlying hardware problem. Other operating systems on the market today use either a monolithic startup script or a series of smaller scripts that are executed sequentially. These systems cannot provide a dependency between scripts nor can they provide a restart of services once problems are corrected. The Predictive Self Healing capability of Oracle Solaris and the fault management combination of the Fault Management Architecture and the Service Management Facility can offline processor threads or cores in faults, retire suspect pages of memory, log errors or faults from I/O components, or deal with any other issue detected by the server.

» **Oracle Solaris ZFS**. The Oracle Solaris ZFS file system provides unparalleled data integrity, capacity, performance, and manageability for storage. It provides high-resiliency features, such as metadata logging, to guarantee data integrity and speed recovery in the event of system failure. What differentiates Oracle Solaris ZFS from other competitive file system offerings is strong data integrity and high resiliency.

  Oracle Solaris ZFS dramatically simplifies file system administration to help increase protection against administrative error by using techniques such as copy-on-write and end-to-end check summing to keep data consistent and eliminate silent data corruption. Because the file system is always consistent, time-consuming recovery procedures, such as using `fsck`, are not required if the system is shut down in an unclean manner. In addition, data is read and checked constantly to help ensure correctness, and any errors detected in a mirrored pool are automatically repaired to protect against costly and time-consuming data loss and previously undetectable silent data corruption.

  Corrections are made possible by a RAID-Z implementation that uses parity, striping, and atomic operations to aid the reconstruction of corrupted data. Oracle Solaris ZFS provides end-to-end data protection as well as facilities to self-correct data. It also offers a simplified management model that does not require a third-party volume manager. Oracle Solaris ZFS offers the scalability to grow to the largest data storage requirements. Oracle Solaris ZFS constantly reads and checks data to help ensure it is correct, and if it detects an error in a storage pool with redundancy (protected with mirroring, ZFS RAID-Z, or ZFS RAID-Z2), Oracle Solaris ZFS automatically repairs the corrupt data. It optimizes file system reliability by maintaining data redundancy through the delivery of basic mirroring, compression, and integrated volume management.

» **Oracle Solaris DTrace**. A feature of Oracle Solaris, DTrace is a dynamic tracing framework for troubleshooting systemic problems in real time on production systems. DTrace is designed to quickly identify the root cause of system performance problems. DTrace safely and dynamically instruments the running operating system kernel and applications without rebooting the kernel and recompiling—or even restarting—applications. This design greatly improves service uptime and lets developers and operators understand dynamic performance issues.

Oracle Solaris Multipathing

Oracle Solaris multipathing software enables organizations to define and control redundant physical paths to I/O devices such as storage devices and network interfaces. If the active path to a device becomes unavailable, the software can automatically switch (fail over) to an alternate path to maintain availability. To take advantage of multipathing capabilities, the server must first be configured with redundant hardware, such as redundant network interfaces or two host bus adapters connected to the same dual-ported storage array.

A number of different multipathing software options are available, including the following:

» **Oracle Solaris IP Network Multipathing (IPMP)**. Oracle Solaris IP Network Multipathing provides multipathing and load-balancing capabilities for IP network interfaces. For more information, please refer to the appropriate Oracle Solaris operating system documentation.

» **Oracle Solaris Datalink Multipathing (DLMP)**. Also referred to as *trunking*, Oracle Solaris Datalink Multipathing aggregations enable administrators to configure several interfaces on a system as a single, logical unit to increase throughput of network traffic. For more information, please refer to the appropriate Oracle Solaris operating system documentation.

» **Oracle Solaris I/O Multipathing (MPxIO)**. MPxIO enables administrators to configure the multipathing features for Fibre Channel devices to control all supported Fibre Channel host bus adapters. For more information, please refer to the appropriate Oracle Solaris operating system documentation.

» **Oracle VM Server for SPARC Virtual Disk Multipathing**. Virtual disk multipathing lets administrators configure a virtual disk on a guest domain to access its back-end storage by more than one path. For instructions on configuring and administering Oracle VM Server for SPARC Virtual Disk Multipathing, please refer to the appropriate *Oracle VM Server for SPARC Administration Guide* in the documentation library.

» **Oracle's Sun StorageTek Traffic Manager**. Sun StorageTek Traffic Manager is an architecture that is fully integrated within Oracle Solaris (beginning with the Oracle Solaris 8 release) and enables I/O devices to be accessed through multiple host controller interfaces from a single instance of the I/O device. For more information, please refer to the appropriate Oracle Solaris operating system documentation.

Oracle Solaris Cluster

While SPARC M7 processor–based servers deliver high levels of RAS, Oracle Solaris Cluster enables organizations to deliver highly available application services. To limit outages due to single points of failure, mission-critical services need to be run in clustered physical domains or physical servers that efficiently and smoothly take over the services from failing nodes, with minimal interruption to data services. While SPARC M7 processor–based servers feature redundancy at the hardware level, Oracle Solaris Cluster provides the best high availability solution for SPARC servers running Oracle Solaris and applications.

Oracle Solaris Cluster is focused on failover between zones and logical domains within a server as well as to external servers. Tightly coupled with Oracle Solaris, Oracle Solaris Cluster detects failures without delay (subsecond delay) and provides much faster failure notification, application failover, and reconfiguration time than competing solutions. Oracle Solaris Cluster offers high availability for today's complex solution stacks, with failover protection from the application layer through to the storage layer. This approach enables much faster resumption of IT services on SPARC M7 processor–based servers. Oracle Solaris Cluster does the following:

» Integrates tightly with the Oracle Solaris Predictive Self Healing framework and supports applications controlled by the Service Management Facility in Oracle Solaris Zones and logical domains

» Makes extensive use of Oracle's 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 as the single file system type used

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

» Integrates with Oracle Enterprise Manager Ops Center

Virtualization and Oracle VM Server for SPARC

Virtualization technology is increasingly popular and essential as organizations strive to consolidate disparate workloads onto fewer, more-powerful systems, while increasing utilization. SPARC M7 processor–based servers are designed specifically for virtualization, providing all three levels of partitioning and virtualization technologies: hardware partitioning, server virtualization, and OS virtualization.

» **Physical domains** (PDoms) are used to divide a single hardware system into multiple security and fault-isolated servers (on SPARC M7-8 and M7-16 servers). Each PDom runs a unique copy of the hypervisor (Oracle VM Server for SPARC).

» **Logical domains** (LDoms) are created using Oracle VM Server for SPARC and are used to virtualize a server or physical domain to host multiple virtual machines (VMs), each running its own instance of Oracle Solaris.

Oracle Solaris Zones enable OS virtualization so that a single instance of Oracle Solaris can securely isolate applications from each other and allocate appropriate system resources to each zone.

Hypervisor firmware is included in every system based on SPARC M7 processors, which enables logical domain partitions for software fault containment and resource allocation, including I/O virtualization. The hypervisor provides processor support for error clearing, correction, and collection.

Utilizing the hypervisor, Oracle VM Server for SPARC provides full virtual machines that run an independent operating system instance. Each logical domain is completely isolated. The maximum number of virtual machines created on a single platform relies upon the capabilities of the hypervisor, rather than on the number of physical hardware devices installed in the system.

Oracle VM Server for SPARC also has the ability to perform a live migration from one domain to another. As the term implies, the source domain and application no longer need to be halted or stopped. This functionality allows a logical domain on the server to be moved to a different physical domain on the same server, on a different SPARC M7 processor–based server, or to previous-generation servers based on Oracle's SPARC T2, T3, T4, T5, M5, or M6 processors.

Logical domains can also host Oracle Solaris Zones to capture the isolation, flexibility, and manageability features of both technologies, or to support legacy Oracle Solaris releases using Oracle Solaris branded zones (Figure 4). Deeply integrating Oracle VM Server for SPARC with the SPARC M7 processor, Oracle Solaris increases flexibility, isolates workload processing, and improves the potential for maximum server utilization.

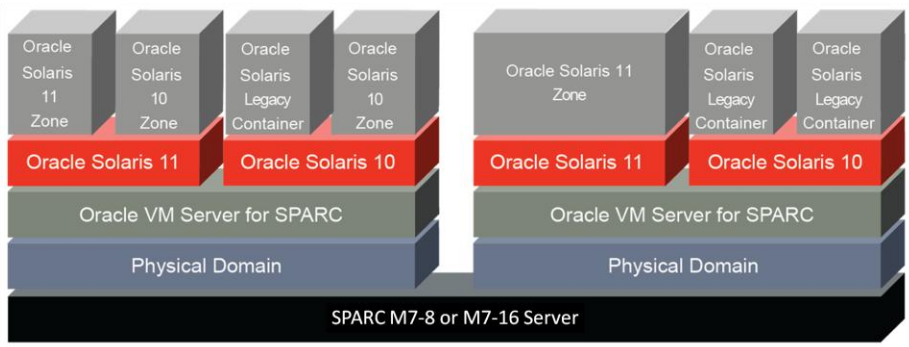

Figure 4. Oracle VM Server for SPARC runs on physical domains, and allows for further system virtualization and isolation using Oracle Solaris Zones and Oracle Solaris branded zones.

Oracle Products and Solutions for High-Availability Infrastructure

With access to the entire hardware and software application stack, Oracle is in a unique position to innovate with solutions that contribute directly to application availability. Beyond designing highly available systems and software, Oracle is committed to solutions and architectures that combine software into solutions that solve real problems for key applications and infrastructures.

Oracle Real Application Clusters (Oracle RAC)

Oracle Real Application Clusters (Oracle RAC) is a shared-cache clustered database architecture that overcomes the limitations of traditional shared-nothing and shared-disk architectures. Oracle RAC offers database performance, scalability, and reliability without requiring changes to existing Oracle Database applications. Oracle RAC has been successfully deployed by thousands of organizations, allowing the use of clustered database servers for a simplified, efficient, and successful delivery of database services in the cloud. Oracle RAC does the following:

» **Runs all database workloads**. Oracle RAC can be used for online transaction processing and data warehousing applications, or for mixed workloads. No changes to the application are required. Oracle RAC can also be deployed with complementary database technologies including Oracle Multitenant and Oracle Active Data Guard.

» **Provides a foundation for database services in the cloud**. Oracle RAC provides all the software components required to easily deploy Oracle Database on a pool of servers, exploiting the performance, scalability, and availability advantages that clustering provides. Oracle RAC utilizes the Oracle Grid Infrastructure feature of Oracle Database as the foundation for Oracle RAC database systems. Oracle Grid Infrastructure includes Oracle Clusterware and Oracle Automatic Storage Management (Oracle ASM), which enable efficient sharing of server and storage resources in a highly available and scalable database cloud environment.

- **Provides scalability on demand**. Oracle RAC enables the transparent deployment of Oracle Database across a pool of clustered servers. This ability enables organizations to easily redeploy their single-server Oracle Database silos onto a cluster of database servers—taking full advantage of the combined memory capacity and processing power the clustered database servers provide.
- **Provides the highest database availability**. Oracle RAC provides organizations with the highest database availability by removing a single database server as a single point of failure. In a clustered server environment, the database itself is shared across a pool of servers. If any server in the server pool fails, the database continues to run on surviving servers.
- **Provides cost-effective resource management**. Oracle RAC includes innovative technologies to manage workloads in the cluster while at the same time providing the best application throughput considering the configuration and high availability requirements of the application.

Oracle WebLogic Server Cluster

Oracle WebLogic Server 12*c* is the popular application server across conventional and cloud environments. With Oracle WebLogic Server, organizations can deliver next-generation applications on a mission-critical cloud platform, simplify operations with native cloud management, and accelerate time to market with a modern development platform and integrated tools.

An Oracle WebLogic Server cluster consists of multiple Oracle WebLogic Server instances running simultaneously and working together to provide increased scalability and reliability. A cluster appears to clients to be a single Oracle WebLogic Server instance. The server instances that constitute a cluster can run on the same system or be located on different systems. Organizations can increase a cluster's capacity by adding additional server instances to the cluster on an existing system, or they can add systems to the cluster to host the incremental server instances. Each server instance in a cluster must run the same version of Oracle WebLogic Server. An Oracle WebLogic Server cluster provides the following benefits:

- **Scalability**. The capacity of an application deployed on an Oracle WebLogic Server cluster can be increased dynamically to meet demand. Administrators can add server instances to a cluster without interruption of service—the application continues to run without affecting clients and end users.
- **High availability**. In an Oracle WebLogic Server cluster, application processing can continue if a server instance fails. Application components are clustered by deploying them on multiple server instances in the cluster. If a server instance on which a component is running fails, another server instance on which that component is deployed can continue application processing.

Oracle Optimized Solutions and Oracle Maximum Availability Architectures

Oracle's systems, middleware, and database technology come together in high-availability configurations in both Oracle Optimized Solutions and Oracle Maximum Availability Architectures. Featuring clustered SPARC M7 processor–based servers, Oracle Optimized Solutions are designed, tested, and fully documented architectures that are tuned for optimal performance and availability. The solution is based on uniquely matched components including Oracle's engineered systems, servers, and storage devices, as well as operating systems, virtualization, database, middleware, and enterprise applications. Oracle Optimized Solutions are available in a variety of areas, including the following:

- **Enterprise applications**. Oracle Optimized Solutions for enterprise applications provide optimized performance with integrated availability and security capabilities for complete application environments, including Oracle's JD Edwards EnterpriseOne applications, Oracle E-Business Suite, Oracle FLEXCUBE, Oracle's PeopleSoft and Siebel CRM applications, and SAP.
- **Middleware**. Oracle Optimized Solutions for middleware help organizations easily manage effective, secure, and high-performing application infrastructures with solutions that include Oracle WebLogic Server.
- **Data management**. Oracle Optimized Solution for Lifecycle Content Management, Oracle Optimized Solution for Oracle Database, and Oracle Optimized Solution for Enterprise Database Cloud provide data management

capabilities that use proven best practices to control costs and maximize performance for growing structured and unstructured storage needs.

» **Core infrastructure**. Oracle Optimized Solutions for core infrastructure include fully tested blueprints that provide the foundation for a more cost effective and agile data center. These solutions include Oracle Optimized Solution for Backup and Recovery, Oracle Optimized Solution for Disaster Recovery, Oracle Optimized Solution for Enterprise Cloud Infrastructure, and Oracle Optimized Solution for Tiered Storage Infrastructure.

Complementing Oracle Optimized Solutions, Oracle Maximum Availability Architectures leverage Oracle's best-practices blueprints, expert recommendations, and customer experiences on proven Oracle high availability technologies. The goal of each Oracle Maximum Availability Architecture is to provide optimal high availability for users at the lowest cost and complexity. As of this writing, Oracle Maximum Availability Architecture best-practices blueprints cover these topics and more:

» Oracle Database
» Oracle Fusion Middleware
» Oracle Applications Unlimited
» Oracle Fusion Applications
» Oracle Enterprise Manager
» Oracle VM
» Oracle Partners
» Oracle Consulting and Support

Conclusion

Oracle's full line of SPARC M7 processor–based servers was designed with a full understanding of the demanding enterprise requirements for high application availability. These servers coupled with the proven Oracle Solaris operating system, innovative SPARC M7 processor features, key redundant components, and features such as DIMM sparing means that these servers can often detect and recover from errors before the errors can interrupt key applications or services. Oracle's layered virtualization technology helps isolate different applications and system images, while key solutions such as Oracle WebLogic Server and Oracle RAC provide scalability and high availability with multiple virtual or physical systems. Providing extensively tested configurations, Oracle Optimized Solutions and Oracle Maximum Availability Architecture best practices combine complete Oracle hardware and software technology stacks into proven solutions that can help organizations realize their most demanding availability goals and SLAs.

For More Information

For more information, visit the resources listed in Table 2.

TABLE 2. RESOURCES

Web Resources	**URL**
Oracle's SPARC servers	oracle.com/us/products/servers-storage/servers/sparc/oracle-sparc and oracle.com/technetwork/server-storage/sun-sparc-enterprise
Oracle Solaris	oracle.com/solaris
Oracle Optimized Solutions	oracle.com/optimizedsolutions
Oracle Maximum Availability Architecture	oracle.com/technetwork/database/features/availability/maa-096107.html
Oracle Solaris Cluster	oracle.com/us/products/servers-storage/solaris/cluster/overview/index.html
Oracle Real Application Clusters (Oracle RAC)	oracle.com/technetwork/database/options/clustering/overview/index.html
Virtualization	oracle.com/technetwork/topics/virtualization
Oracle Enterprise Manager Ops Center	oracle.com/technetwork/oem/ops-center
Oracle Help Center, which provides product documentation	docs.oracle.com
White Papers and Technical Articles	**URL**
"Oracle's SPARC T7 and SPARC M7 Processor–Based Server Architecture"	oracle.com/technetwork/server-storage/sun-sparc-enterprise/documentation/sparc-t7-m7-server-architecture-2702877.pdf
"Consolidation Using Oracle's SPARC Virtualization Technologies"	oracle.com/technetwork/server-storage/sun-sparc-enterprise/technologies/consolidate-sparc-virtualization-2301718.pdf
"High Availability Best Practices for Database Consolidation"	oracle.com/technetwork/database/availability/maa-consolidation-2186395.pdf
"Setting Up, Configuring, and Using an Oracle WebLogic Server Cluster"	http://www.oracle.com/technetwork/articles/soa/vasiliev-wls-cluster-1867166.html
"Best Practices for Oracle Solaris Kernel Zones"	oracle.com/technetwork/articles/servers-storage-admin/solaris-kernel-zones-best-practices-2400370.html
"Oracle VM Server for SPARC Best Practices"	oracle.com/technetwork/server-storage/vm/ovmsparc-best-practices-2334546.pdf

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

CONNECT WITH US

blogs.oracle.com/oracle

facebook.com/oracle

twitter.com/oracle

oracle.com

Integrated Cloud Applications & Platform Services

Oracle's SPARC T7 and SPARC M7 Server Reliability, Availability, and Serviceability
August 2016

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