An Oracle White Paper
September 2010

Oracle Solaris and Oracle Solaris Cluster: Extending Oracle Solaris for Business Continuity

ORACLE

Executive Summary 1
Introduction 3
Traditional Solution Components 5
Oracle Solaris Cluster for Business Continuity 6
Oracle Solaris Cluster 6
Introduction—Basic Clustering 8
Data Availability 10
Applications Availability 11
Network Availability 12
Virtualization and Oracle Solaris Cluster 13
Disaster Recovery—Campus and Beyond 16
A Single HA and DR Solution for Multitier Oracle Applications and Databases 17
Oracle Solaris 19
Resource Management 19
Predictive Self Healing 20
Oracle Solaris Containers 21
Oracle VM Server for SPARC 22
Oracle Solaris Live Upgrade 23
ZFS 23
Oracle Solaris Networking Technology 24
Advanced Security Features 24
Other Oracle Availability Products 25
Conclusion 26
More Information 26

Executive Summary

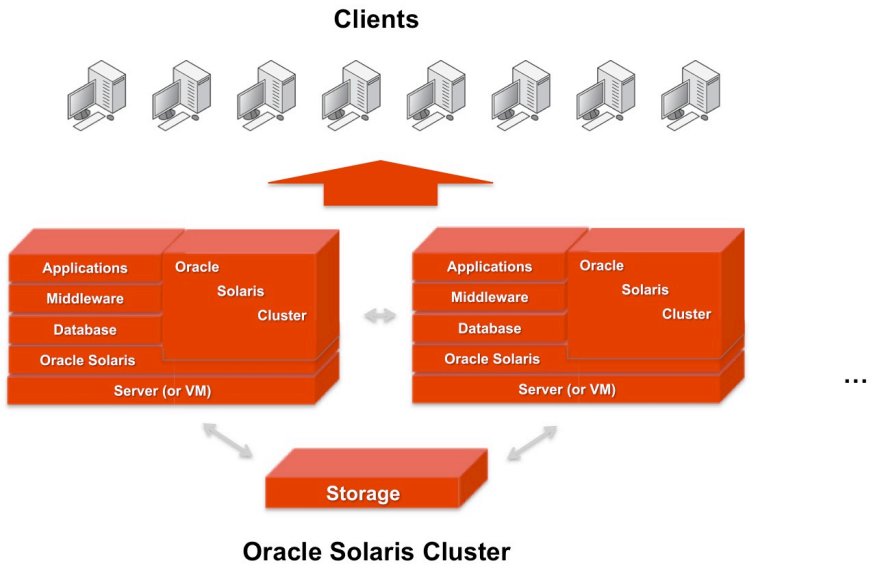


Figure 1: Oracle delivers a complete, integrated high-availability solution for Oracle Database, Oracle Applications, and other leading enterprise applications.

Oracle® Solaris Cluster extends the industry-leading availability features of Oracle Solaris, which includes Predictive Self Healing, reliable networking, resource management, and built-in virtualization functionality. Oracle Solaris Cluster offers the best high availability platform for Oracle Solaris, extending its reach from a single node to multisystem, multisite, and global disaster recovery solutions. Oracle Solaris Cluster offers comprehensive and robust capabilities for keeping your business IT, including those running Oracle Database and Applications, up and running in the face of nearly every conceivable situation. Oracle Solaris Cluster offers a full range of single and multisystem high availability (HA) and disaster recovery (DR) capabilities in traditional and virtualized environments. As a mature and robust solution, Oracle Solaris Cluster offers a high degree of flexibility in how it can be deployed and the technologies it supports.

Compared to point-product HA capabilities, Oracle Solaris Cluster coordinates the entire solution stack, orchestrating dependencies to ensure fast and reliable failover and recovery. Oracle Solaris Cluster works with a wide variety of servers, including Oracle's Sun SPARC® Enterprise servers, Sun x86 servers, Sun Storage and other storage solutions, and multiple network technologies. Organizations can mix systems and assemble clusters of up to 16 servers, or nodes. Applications run in an Oracle Solaris Cluster environment without modification. Cluster resources are managed in a unified administration environment.

Oracle Solaris Cluster works with Oracle Solaris virtualization products, offers fast failure detection and recovery, supports multiple file systems, and automates failover procedures locally, across campus or metro areas, and at even greater distances. To help ensure the highest reliability, Oracle Solaris Cluster is tested extensively with Oracle Databases and Applications, along with leading applications from other vendors. Oracle Solaris Cluster integrates with Oracle Real Application Clusters (RAC) and leverages Oracle Data Guard for disaster recovery. Oracle Solaris Cluster Geographic Edition can help minimize the impact to business IT services if there is a failure at one location.

Introduction

eBay's IT infrastructure supports $2,000 per second of goods being traded.[1]

– Dean Nelson, eBay senior director of Global Data Center Strategy and Operations

In a global economy, any downtime is undesirable. 24x7x365 is a requirement. IT downtime costs an organization in terms of lost sales and customers, harm to the company's reputation, and wages for people unable to perform their duties while the system or site is down. There is also the additional time is needed to play catch-up once the system is functional again.

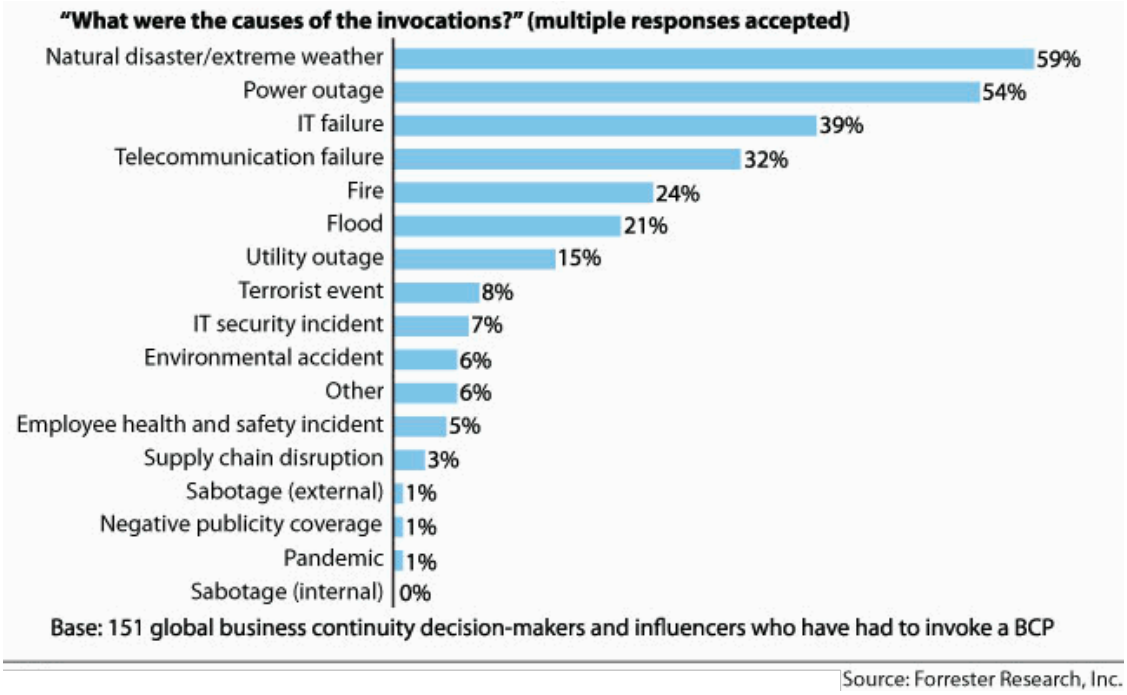


Figure 2: IT solution stacks can fail for many reasons.[2]

But the inescapable fact is that IT systems undergo both planned and unplanned downtime. Patches and upgrades, operator errors, system and software failures, security issues, and man-made and natural

[1] Melanie Kaplan, ZDNet, Aug. 18, 2010, zdnet.com/blog/btl/ebay-datacenter-chief-dean-nelson-we-are-living-moores-law/38127?tag=nl.e539

[2] Stephanie Balaouras, Forrester Research, Disaster Recovery Journal, Winter 2009, www.drj.com/index.php?option=com_content&task=view&id=2407&Itemid=419&ed=49

disasters all affect your ability to do business. Considering the complexity of highly networked, multitiered solution stacks, it can be difficult to keep application services up and running. Some critical applications are so intertwined in the business that if any component is interrupted, it can cause a series of disruptions that eventually lead to a complete breakdown of business processes. For example, a database operator could incorrectly enter a command causing an entire customer relationship management (CRM) system to come to a halt. While the problem is diagnosed and resolved, customer information must be recorded manually. The company must then face the challenge and cost of accurately and securely inputting the data while continuing normal operations. In addition, CRM systems are often dependent on other corporate databases such as order entry, marketing, inventory control, and shipping. A failure of one component in one tier can cascade through the entire business infrastructure.

Government and industry groups recognize the importance of IT to the global economy and citizens' welfare. Sarbanes-Oxley, Basel II, the Health Insurance Portability and Accountability Act (HIPAA), the Graham-Leach-Bliley Act, and other industry- and country-specific regulations increase the penalties for companies whose risk management practices are lacking or non-existent. Because of the integral role of IT in today's world, there is a demand for continuity planning. *Business continuity* is a comprehensive discipline that includes high availability of IT services—service or application continuance in the face of a failed component—and disaster recovery (DR), which can help business recover from large-scale issues or natural disasters.

TABLE 1: PERCENT DOWNTIME

PERCENT UPTIME	DOWNTIME/WEEK	DOWNTIME/YEAR	EXAMPLES
90%	16 hours 48 min	36 days	Offline weekly backup
99%	100 min	87.5 hours	Offline weekly software maintenance
99.9%	10 min	8 hours 45 min	One weekly restart, or one yearly offline backup
99.99%	1 min	0 hours 52 min	One offline backup per year, or several fast switchovers
99.999%	6 sec	0 hours 5 min	One yearly restart – maybe

Disaster recovery and business continuity plans require business users and IT managers to determine a recovery time objective (RTO) and recovery point objective (RPO) for each component of an application service, taking into account any interdependencies that might affect other applications. Costs and existing capabilities are taken into consideration when analyzing RTO and RPO.

- RTO is defined by how quickly applications must be available in order for the business to resume functioning in order to avoid unacceptable losses, including revenue losses.
- RPO is the amount of time that is deemed acceptable for which business applications and services can be unavailable.

For example, a business may determine that it is acceptable to lose one hour of data (RPO) and that a business IT service must be back online within two hours (RTO).

On average, businesses lose between $84,000 and $108,000 (US) for every hour of IT system downtime, according to estimates from studies and surveys performed by IT industry analyst firms...not including damaged reputation or customer loyalty. [3]

Traditional Solution Components

Once the RTO and RPO for each application are determined, the appropriate level of availability can be implemented to meet those objectives. Traditional business continuity and disaster recovery solutions still suffice for a number of application components.

Tape Backup

Tape backup is still widely used and provides high availability of infrequently changing data, but not application services. Tape backups capture a point-in-time instance of data; any data created between that time and the next backup can potentially be lost if a disaster occurs. The data must then be restored, which can take hours to days, depending on the quantity of data. As a result, offsite tape backup offers little ability to meet aggressive RTO and RPO goals. Tape is usually adequate for businesses or applications that can tolerate longer recovery times and recovery points, such as data archives.

Local Clustering

Businesses or applications that need to minimize recovery times and RPO need more proactive strategies, including highly available clusters. A cluster is a collection of tightly-coupled computing nodes that provide a single view of application services, including databases, Web, and file services. A cluster implemented using Oracle Solaris Cluster offers several benefits, including restart, failover, and scalability services; capacity for modular growth; flexibility for changing requirements, and a low entry price compared to fault-tolerant systems. In addition, it simplifies management of planned downtime of individual servers within a cluster, and creates a foundation for DR capabilities.

A local cluster involves two or more physical machines (nodes) that may share common disk storage and logical IP addresses. When an error is detected in one of the nodes, the cluster software will switch the clustered services over to the secondary machine. Oracle Solaris Cluster offers high availability for today's complex solution stacks, with failover protection from the application layer through to the storage layer.

[3] Alan Arnold, IT-Director.com, Apr. 6, 2010 www.it-director.com/business/costs/content.php?cid=12043

Campus and Metropolitan Clustering and Long Distance Failover

The next level of availability is to extend a cluster to another room or across a campus. When planning a disaster recovery infrastructure, it's important to note that local clustering can provide a solid level of continuous service availability in the event of application, operating system, or hardware failure in a single datacenter, but it does not provide site-level disaster tolerance against disasters. Campus clustering enables components, such as nodes and shared storage, to be located up to 400 kilometers apart. In the event of a localized disaster such as a flood, fire, or building power outage, the surviving nodes can support the service for a failed node. This solution offers some site-level tolerance, but the short distance limits survivability of the cluster and its services for larger disasters like earthquakes or power grid outages. Oracle Solaris Cluster Geographic Edition extends Oracle Solaris Cluster, offering the control and flexibility required when working with limited bandwidth over long distances.

Oracle Solaris Cluster for Business Continuity

"Oracle Solaris Cluster is an end-to-end solution—while usually, you have to find bits and pieces from different suppliers and build them into a cohesive, unified solution. Oracle Solaris Containers also fit well with Oracle Solaris Cluster technology by supporting the consolidation of workloads. Overall, business continuity is supported, and applications can continue to run, even if other workloads must be restarted, without affecting the ongoing business processes."

– Frank Oakley, enterprise architect, Liverpool Direct

While Oracle Solaris offers comprehensive reliability features, Oracle Solaris Cluster goes much farther. Highly available Oracle Solaris Cluster provides nearly continuous access to data and applications by keeping the cluster running through failures that would normally bring down an application or service. No single failure, hardware, software, or network, can cause a cluster to fail. A cost-effective high availability solution requires flexibility. Oracle Solaris Cluster enables customers to define and choose the right combination of high availability and hardware within budget.

Oracle Solaris Cluster

In today's global 24x7 economy keeping your enterprise applications up and running is more important—and can be more complex—than ever. As shown in Table 1, "always-on" availability means no time for planned maintenance, not to mention unplanned outages. Government regulations, corporate financial goals, and evolving requirements to address new opportunities mean your IT systems need to be constantly available. This can be a challenge with today's complex solution stacks and unique business requirements.

Oracle Solaris Cluster helps organizations deliver highly available application services. Oracle Solaris Cluster extends Oracle Solaris to provide enhanced availability of hosted applications. Oracle Solaris Cluster delivers instant support for several commercial and open-source applications through specific integration modules called *agents*, which monitor an application to detect whether it is operating correctly, and take action if a problem is detected. Applications do not need to be modified to benefit from the enhanced availability offered by Oracle Solaris Cluster agents. Applications can run either

directly on Oracle Solaris on a physical system, or in different virtual environments such as Oracle VM Server for SPARC, Dynamic Domains, or Oracle Solaris Containers.

Core Oracle Solaris services, such as devices, file systems, and networks work in a nearly seamless manner across a tightly-coupled cluster. While maintaining full Oracle Solaris compatibility for existing applications, Oracle Solaris Cluster transparently manages shared access to its global file service, global network service, and global devices. Enhanced availability for core services such as file and network services and virtualization allows existing applications written for Oracle Solaris to benefit from higher availability, more localized failure containment, and faster failover.

The Oracle Solaris Cluster framework can help reduce TCO by simplifying administration. Resources from multiple nodes can be shared and administered on a single console. Administrators can access system management tools and Oracle Solaris OS commands from any system in the cluster. Administrators can add or remove nodes while online to meet specific needs and respond to changing conditions.

Oracle Solaris Cluster supports both scalable and HA (failover) services, and can be extended to provide DR services. Oracle Solaris Cluster's capabilities can support multiple nodes in a cluster to handle growing numbers of simultaneous users—scalable services—but these capabilities are not discussed in this paper.

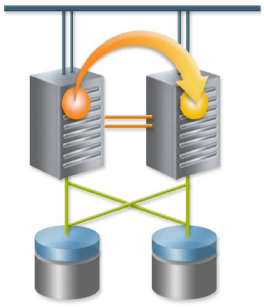
The following bullet points summarize the key capabilities and benefits of Oracle Solaris Cluster. More detail on each point is discussed in the pages following.

- **High availability framework.** Software that detects node failures quickly and activates resources on another node in the cluster. The framework includes a Cluster Membership Monitor, a distributed set of algorithms and agents which exchange messages over the cluster interconnect to enforce a consistent membership view, synchronize reconfiguration, handle cluster partitioning, and help maintain full connectivity among all cluster members.
- **Virtualization support.** Oracle Solaris Cluster provides comprehensive support for Oracle virtualization software: Oracle Solaris Containers, Oracle VM Server for SPARC, and Dynamic Domains (available on Sun SPARC Enterprise M-Series servers). This enables flexible HA for server consolidation efforts. Application can run unmodified in a virtualized environment.
- **Global devices, files, and networking.** All global devices, files, and network interfaces can be seen as local resources. This allows cluster nodes to access and utilize devices that are not physically attached, but instead are attached to another node within the cluster. These facilities create improved resource availability and simplified administration.
- **Flexible storage support.** Oracle Solaris Cluster can be used in combination with different storage technologies such as FC, SCSI, iSCSI, and NAS storage on Sun or non-Sun storage. There is broad file system and volume manager support, too.
- **High availability Oracle Solaris ZFS.** With the virtually unlimited scalability of ZFS, Oracle Solaris Cluster offers a file system solution with exceptional availability, data integrity, and flexibility for growth.

- **Component monitoring.** Extensive monitoring of application processes, disk path integrity, and network availability. For example, all disk paths can be monitored and configured to automatically reboot a node in case of a multiple path failure.
- **Oracle RAC 10*g* and 11*g* integration and administration.** Automated installation and wizard-led configuration enable faster set-up of Oracle RAC with Oracle Solaris Cluster. Specific RAC integration points enable improved coordination and simplified administration.
- **Failover and scalable agents.** Software programs that support Oracle or third-party applications to take full advantage of Oracle Solaris Cluster features.
- **Cluster-aware applications.** Third-party applications with direct knowledge of Oracle Solaris Cluster systems, such as Oracle Real Application Clusters (RAC) software.
- **Predictive Self Healing and failfast.** Oracle Solaris Cluster leverages Predictive Self Healing, which monitors application and system health. Failfast halts a component not healthy enough to continue correct operations.
- **Protection groups.** Oracle Solaris Cluster enables an administrator to define groups of resources that must be managed together, facilitating seamless startup, failover and shut down of related applications and data.
- **Highly available private interconnect.** Multiple types of interconnect technologies are supported by Oracle Solaris Cluster to establish a private communication channel between cluster nodes. Oracle Solaris Cluster can manage multiple interconnects to ensure high availability and improve performance of private internode communication.
- **Unified management view.** Oracle Solaris Cluster's Web-based graphical user interface (GUI) simplifies monitoring and management of all cluster components. A command line interface can be used on any node of the cluster. GUI-based wizards simplify configuration operations.
- **Certified application support.** Oracle Solaris Cluster, including local cluster and DR solutions, is stringently tested and certified together with Oracle Databases and Applications, Sun servers, Sun Storage, and connectivity solutions. In addition, Oracle Solaris Cluster supports more Solaris applications than any of its competitors.
- **Multisite disaster recovery.** Oracle Solaris Cluster Geographic Edition manages the availability of application services and data across unlimited geographic distances.
- **Extensible.** Oracle Solaris Cluster offers the ability to enhance the availability of other applications and technologies with published and documented APIs, generic service templates, and a graphical agent toolkit.

Introduction—Basic Clustering

Industry best practices dictate that business applications and services should be able to recover from any single point of failure. At its simplest, Oracle Solaris Cluster monitors the health of cluster components, including the stack of applications, middleware, operating system, servers, storage, and



network interconnects. Any failure executes a policy-based, application-specific recovery action. Recovery is enabled through redundant infrastructure and intelligent software algorithms.

Figure 3: Oracle Solaris Cluster server failover

From a physical perspective, an Oracle Solaris Cluster system consists of two or more servers that work together as a single entity to cooperatively provide applications, system resources, and data to users. Each server generally provides some level of redundancy. Storage is hosted by highly available redundant disk systems, which may be mirrored, supporting data access in the event of a service interruption on a single disk or storage subsystem. Redundant connections are provided to the disk systems so that data is not isolated in the event of a server, controller, or cable failure. A high-speed, redundant, private interconnect system provides access to resources across the set of servers. Redundant connections to the public network also provide each node with multiple paths for access to outside systems, helping ensure continued access in the event of a network connection or node failure.

Key to the design of Oracle Solaris Cluster systems is the fact that no single failure in the hardware, software, storage, or network can cause the cluster to fail. Oracle Solaris Cluster systems prevent loss of service through hardware redundancy, hardware and software failure detection, automatic recovery of services, and failover of applications. Oracle Solaris Cluster systems also provide a single management view for all of the services in the cluster. The entire cluster can be managed as a single entity, reducing the risk of errors.

Oracle Solaris Cluster technology depends on several technology components. Each plays an important role to ensure the correct inclusion of the right servers in the cluster.

- *Heartbeats* monitor the servers (or nodes) in a cluster through private interconnects. If one of the servers goes offline and ceases its heartbeat, the remaining devices in the cluster isolate that server and *fail-over* any application or data from the failing server to another server. This failover process is done quickly and is transparent to the users of the system.
- *Membership* establishes consistent inclusion of the hardware servers across the cluster and coordinates reconfiguration of other layers such as Global File Systems and Resource Group Management, discussed later.
- The *HA framework* guarantees that each node of the cluster maintains a consistent view of the cluster configuration and status. Internode message delivery and responses are handled in an atomic manner that accounts for delivery failures, node membership, and software revision level (to provide for rolling upgrades).

Oracle Solaris Cluster includes capabilities to detect, isolate, and contain failing cluster nodes. It accomplishes this using a robust, kernel-based membership monitor. Every node sends out low-level data link provider interface (DLPI) packets once per second (a heartbeat) to each of its peers on each of the private networks. These packets are sent in the kernel interrupt context, making them very resilient to the peaks in system load. A network, or path, between two nodes is only declared down if a heartbeat message does not complete the round trip between nodes, over that specific path, within the time-out period.

From a logical perspective, Oracle Solaris Cluster consists of a set of nodes, where each independently executes a copy of Oracle Solaris and communicates through the private interconnect. Utilizing high-speed messaging facilities to build a closely coupled, yet highly available, cluster eliminates the need for domains to share memory. Using fully distributed algorithms, the cluster continuously monitors the health of all members. Failing domains are actively prevented from participating in the cluster to prevent any chance of data corruption.

Data Availability

Oracle Solaris Cluster supports a wide range of file systems and services that enable data availability and help ensure data integrity.

- A *failover file service* provides high availability to non-global file system configurations and is suitable for single node applications. In general, when creating a failover service, you should configure one or more highly available local (failover) file systems to hold the data associated with the service. Oracle Solaris Cluster supports locally mounted UFS, VxFS, QFS, and ZFS for failover file services, combining availability with the type of performance provided by the local file system.
- A *global file service* uses shared storage devices—storage with physical connections to more than one domain—to make data highly available and accessible to application services running on any node of the clustered systems. Global file services improve ease of use through a simple, single point of management. Storing application configuration data on a global file service enables faster recovery of failed application instances. Global file services are offered above UFS and VxFS. This includes:
    - Cluster file systems, which are used to support cluster applications that need the capability to read and write the same file system from multiple machines concurrently. *Cluster file system* and *global file system* both refer to a file system that supports concurrent file system access from multiple nodes of the cluster. The specific implementation of a cluster file system in Oracle Solaris Cluster is the Proxy file system (PxFS). PxFS is a highly available, distributed, cache-coherent file system that is POSIX compliant. The most common use case for PxFS is to support general-purpose file system workloads. PxFS extends a single machine file system to work across the cluster. PxFS currently supports UFS and VxFS file systems as the underlying file system. No modifications were made to VxFS file systems to enable this support. To enhance performance, a limited number of additional interfaces were added to UFS. Another primary goal of a cluster file system is to simplify administration by keeping the administration of this cluster file system as close as possible to that of a single machine file system. Another primary function of this cluster file system is to provide a cluster-wide namespace for files.
    - *Shared QFS* file system, which supports Oracle RAC. Using shared QFS file systems within a Solaris Cluster implementation gives you another option for storing your Oracle RAC data and binaries. Unlike raw-disk volumes or Oracle's Automated Storage Management (ASM) software, a shared QFS file system gives your system and database administrators visibility of and access to the data held within it using the standard Oracle Solaris commands, such as `ls`,

`df`, `du`, `tar`, and `cpio`. Shared QFS offers high capacity, fast file system recovery, volume management, and support for paged and direct I/O. A shared QFS file system used for Oracle RAC data files can perform almost identically to an equivalent raw-disk or an Oracle ASM configuration.

Oracle Solaris Volume Manager, Oracle ASM, Veritas Volume Manager, and ZFS storage pools are supported alongside to address the need for data redundancy and protect data against failure through the mirroring capabilities. Oracle Solaris ZFS is supported as a failover file system and as a boot file system, enabling the use of ZFS storage as the single file system type. Note that the Oracle Solaris ZFS file system contains both file system and volume manager features. The *zpool* is a ZFS volume manager component. A single zpool can support multiple file systems.

Data Integrity

Because cluster nodes share data and resources, a cluster must never split into separate partitions that are active at the same time—multiple active partitions might cause data corruption. Oracle Solaris Cluster applies optional *fencing* and *quorum* to protect data integrity. When a failing node in a cluster is detected, Oracle Solaris Cluster uses disk fencing to prevent the node from accessing clustered data again. With Oracle Solaris Cluster the fencing protocol can be chosen per storage device.

In a more complex situation where all paths across the private interconnect fail and the cluster breaks into multiple partitions, Oracle Solaris Cluster uses quorum technology to resolve partitions or *split brain,* and to protect data integrity. Quorum also prevents *amnesia* through the ability to detect and reject the use of an outdated configuration that could lead to data corruption. You can tailor quorum functionality to fit into your storage and system topology, enabling solutions such as disk-based quorum and software quorum. Oracle Solaris Cluster provides a *quorum device protocol* that allows customers to use different types of disks, such as high-capacity 2 TB disk drives, SATA, and SSD as quorum devices. All quorum devices are continuously monitored to enhance availability.

Applications Availability

A *data service* is an application running on a cluster that is made highly available through agents that provide start, stop, and monitoring capabilities. A data service is one or more instances of applications running in the cluster simultaneously. If a node on which an application instance is running fails, the application is restarted on a surviving node or the load is redistributed among the surviving nodes, depending on the load-balancing policy. These applications run within the cluster under the same management control to provide ease of manageability.

Oracle Solaris Cluster provides a robust, extensible framework for all of this. One of the most important features of Solaris Cluster is its resource group manager, which maintains tight control of the starting, stopping, and monitoring of services, held in *resource groups*. By providing strong control, it helps ensure that a service is never started elsewhere before it is stopped in its former location. The Resource Group Manager provides a single procedure for instances of a distributed application to be installed and brought online or offline on multiple cluster nodes, saving time and reducing the complexity associated with distributed applications. The Resource Group Manager is tightly integrated

with Oracle Solaris Service Management Facility (SMF—discussed in the Oracle Solaris section), and leveraging SMF's recovery management capabilities.

Oracle Solaris Cluster supports an extensive list of data services from Oracle and independent software vendors. The data services supported directly by Oracle include:

TABLE 2: ORACLE SOLARIS CLUSTER SUPPORTED APPLICATIONS AND SERVICES*

CATEGORY	AVAILABLE APPLICATIONS OR SERVICES
Oracle	Oracle Application Server, Oracle Communication Messaging Exchange Server and Enterprise Mobility Server, Oracle Business Intelligence Suite Enterprise Edition Plus, Oracle E-Business Suite, Oracle Database, Oracle RAC, Oracle WebLogic Server, Oracle Grid Engine, Oracle TimesTen, Oracle Solaris Containers, Oracle VM Server for SPARC, PeopleSoft, PostgreSQL, Siebel CRM, MySQL Server, MySQL Cluster.
Web and infrastructure	Apache Proxy Server, Apache Tomcat, Apache Web Server, Dynamic Host Configuration Protocol (DHCP), Domain Name Service (DNS), NFS, Kerberos, Samba.
Enterprise applications	Agfa IMPAX, IBM Informix, IBM WebSphere Message Broker, IBM WebSphere MQ, SAP, SAP liveCache, SAP MaxDB, SAP Enqueue Server, Sybase ASE.
Financial services	SWIFTAlliance Access, SWIFTAlliance Gateway.

* Additional products are under consideration. For the most current information, visit the Oracle Solaris Cluster features page.[4]

Network Availability

Oracle Solaris Cluster leverages Solaris IP network multipathing (IPMP—discussed in the Oracle Solaris section later in this document) as public network interfaces for monitoring local failures, and for performing automatic failover from one failed network adaptor to another. IP network multipathing enables a server to have multiple network ports connected to the same subnet. First, IP network multipathing software provides resilience from network adapter failure by detecting the failure or repair of a network adapter. The software then simultaneously switches the network address to and from the alternative adapter. When more than one network adapter is functional, IP network multipathing increases data throughput by spreading outbound packets across multiple adapters.

For scalable data services, the requests will go through a round-robin load-balancing scheme for a balanced distribution of load to the various instances of the distributed application running within the cluster. Scalable data services can be made more secure through the use of IPsec services in combination with Oracle Solaris Cluster load balancing services.

[4] www.oracle.com/technetwork/server-storage/solaris-cluster/overview/features-cluster-166765.pdf

Cluster Private Network

When Oracle Solaris Cluster is installed, it automatically creates a virtual network interface called `clprivnet0`. This virtual interface can be supplied directly to any application running in the Oracle Solaris Cluster environment, including Oracle RAC. The virtual interface provides transparent aggregation of all the underlying cluster private interconnect interfaces. If a path between cluster nodes fails, that path is automatically removed from the trunk until such time as the path is found to be working again. If private interconnects are added or removed, this can be done online and `clprivnet0` automatically takes account of the change.

Finally, the combination of `clprivnet` and the Oracle Solaris Cluster heartbeats provides an end-to-end test of the health of the communication link for potential (Oracle) Cache Fusion traffic. In contrast, an IPMP solution only tests the ability of a network adapter to communicate with a remote ping-able target and not necessarily the host that is the recipient of the application messages.

The `clprivnet0` feature supports all NICs that are qualified with Oracle Solaris Cluster. This includes 100baseT, Gigabit Ethernet, 10 Gigabit Ethernet, and InfiniBand NICs as well as supporting jumbo frames. In addition, it does not require any special switch features (such as 802.3ad link aggregation), and also supports connecting each path to a separate switch. In cases where only two nodes are present, back-to-back Ethernet network connections are supported. The `clprivnet` interface can be configured for IPsec to provide secure communication on the cluster interconnect.

InfiniBand

InfiniBand drivers were added to the Oracle Solaris 10 1/06 release, and support is ramping up across the Sun server line for InfiniBand Host Channel Adapters (HCAs) support. InfiniBand is capable of running at speeds up to quad data rate (QDR), or 40 gigabits/sec. Because all current InfiniBand cards have two ports, only one InfiniBand HCA is needed for minimal Oracle Solaris Cluster configurations, although using more than one card is ideal. InfiniBand can be used on the private and public network, and for storage interconnect.

Virtualization and Oracle Solaris Cluster

"We are using Solaris Cluster to ensure that the system is highly available supporting the real time supply chain of results for the Games. The IDS applications are fully virtualized using Solaris Containers simplifying the backup process, enhancing security, and paving the way for consolidation and hardware cost savings in future games."

– Barry Caswell, Director of IT Operations and Security, Vancouver Organizing Committee for the 2010 Olympic and Paralympic Winter Games (VANOC)

Cluster nodes for Oracle Solaris Cluster can be physical servers and blades, and virtualized environments such as Oracle VM Server for SPARC, Oracle Solaris Containers, and Dynamic Domains (on some models of Sun SPARC Enterprise Server M-Series systems). Oracle Solaris Cluster provides expanded support for virtualization, including support for Oracle Solaris Containers

virtualization software. Oracle Solaris Cluster can leverage these virtualized environments to consolidate multiple applications within the same servers, optimizing resource use.

Oracle Solaris Containers

Oracle Solaris Containers partitioning technology is an Oracle Solaris virtualization feature. An Oracle Solaris Containers cluster is a virtual cluster that allows applications to run fully isolated across the clustered machines. Multiple Oracle and third-party applications and database versions can be consolidated into one physical cluster for highly reliable, consolidated service at a much lower cost, while still benefitting from Oracle Solaris Containers' advantages of security isolation, resource management, delegated administration, and fault isolation.

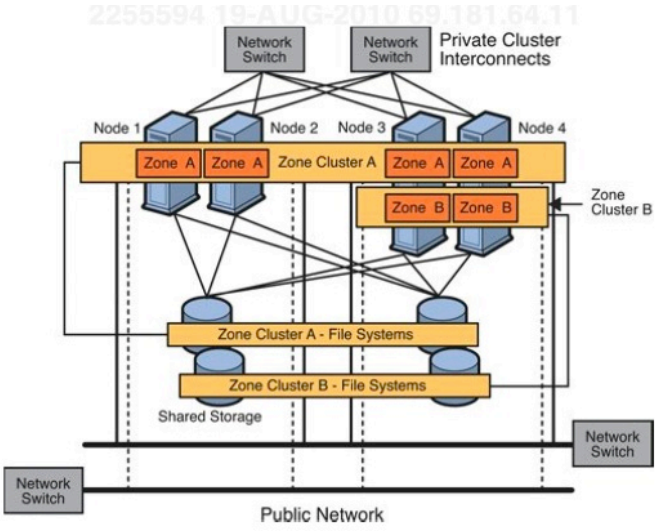


Figure 4: Zone cluster, or Containers cluster [5]

Oracle Solaris Cluster delivers advanced integration with Oracle Solaris Containers, by enabling the creation of multiple virtual clusters composed of Oracle Solaris Containers. This feature—also known as *zone clusters*—optimizes hardware compute resource utilization and can significantly reduce software expenses. Each application runs in fully isolated virtual clusters based on Oracle Solaris Containers virtual nodes that can be configured to cap CPU usage, reducing the number of application licenses required and the cost associated while providing highly available consolidated services. The applications are monitored within the Oracle Solaris Container. Application-level failures are detected and can be recovered based on the policies defined in the specific application agent. These

[5] Source: Tim Read, *Oracle Solaris Cluster Essentials*, 2010, Prentice Hall

virtualization capabilities allow scalable and failover applications and associated Solaris Cluster agents to run unmodified within Oracle Solaris Containers.

In addition to zone clusters, there are two other approaches for using Oracle Solaris Containers:

- The *black box* approach, where a Container is considered as a resource with all resource groups configured in the global zone. The HA Solaris Container agent can failover the complete container as a standalone unit. In this mode the cluster software can only start and stop the Container, and probe inside to ensure it is still running.
- The *white box* approach, where a Container is treated as a node. In this case, multiple resource groups can run in the same Container and fail over independently. Oracle Solaris Cluster can manage resources inside Oracle Solaris Containers with fine-grain control including monitoring of the application as if the container was a physical cluster node.

Oracle Solaris Containers cluster is the first and most comprehensive Oracle Solaris-based HA solution that leverages software licensing models based on CPU utilization. The cost of the applications and databases that co-exist in the same cluster of hardware can be effectively reduced by using Oracle Solaris Containers clusters.

Oracle VM Server for SPARC

Oracle VM Server for SPARC (formerly Sun Logical Domains or LDoms) is a facility provided by Sun SPARC Enterprise T-Series servers that allows one or more virtualized instances of Oracle Solaris 10 to run on the same underlying hardware.

Oracle Solaris Cluster integrates with Oracle VM server for SPARC in two ways. As a failover guest domain, it is managed as a standalone resource with the Oracle Solaris Cluster HA Agent that starts, stops, fail-over, and migrates as a whole (similar to black box, above). The Oracle VM Server for SPARC can also be managed as a node (similar to white box, above). Applications running within the guest domain, as well as their resources and dependencies, can be controlled and managed independently as if they were running on a classic Oracle Solaris Cluster hardware node.

Note that there is additional information on Oracle Solaris Containers and Oracle VM Server for SPARC in the section on Oracle Solaris.

Dynamic Domains

Many Sun SPARC Enterprise M-Series servers are available with Dynamic Domains, which offer electrically isolated domains that enable flexible and dynamic allocation of CPUs, memory, and I/O resources. (For additional flexibility, each Dynamic Domain can run Oracle Solaris Cluster.) Because domains are electrically isolated from one another, each domain is protected from any failures that might occur in other domains within the same chassis.

Each domain within a chassis runs a completely separate instance of the Oracle Solaris. Oracle Solaris Cluster can create distinct clusters between any two or more domains, whether the domains are in the

same chassis or in separate chassis. These clusters have their own Resource Group Manager to control the services contained in the resource groups.

Disaster Recovery—Campus and Beyond

Campus clustering enables components, such as nodes and shared storage, to be located tens of kilometers apart. In the event of a localized disaster such as a flood, fire, or building power outage, the surviving nodes can support the service for a failed node. This solution offers some site-level tolerance, but the short distance limits survivability of the cluster and its services for larger disasters like earthquakes or power grid outages.

Oracle Solaris Cluster automates failover procedures across a campus or metropolitan area to limit service outages due to local problems, minimizing human error and improving recovery time and overall services availability. Supported replication technologies include EMC SRDF, and Hitachi Universal Replicator and TrueCopy. Oracle Solaris Cluster also supports host-based mirroring, including SVM, VxVM, and ZFS. Hardware RAID is also supported.

Metro Clusters can use Wave Division Multiplexors (WDM) to extend the reach of a traditional campus cluster configuration and allows a cluster to span up to 400 km, enabling robust disaster recovery solutions.

DR and Geographic Edition

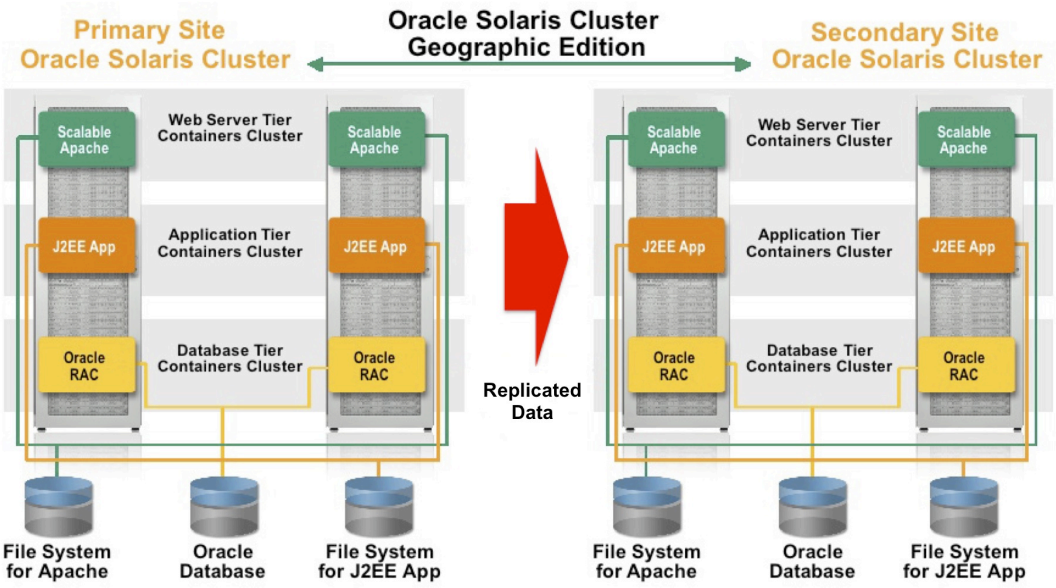


Figure 5: Oracle Solaris Cluster Geographic Edition provides comprehensive DR capabilities.

To provide true business continuity for applications that must continue to run despite the scope of a disaster, unlimited distance synchronization is required. Oracle Solaris Cluster Geographic Edition software provides this next step in the evolution of continuously available datacenters. Oracle Solaris Cluster Geographic Edition is a layered extension to Oracle Solaris Cluster. The Geographic Edition

software enables the controlled migration of production services from a primary cluster to a secondary cluster, either in the event of a disaster or as part of a planned procedure. Data is continuously replicated from the primary cluster to the secondary cluster, either synchronously or asynchronously (or a combination of both) depending on the RPOs of the application services supported by the clusters.

The Geographic Edition software supports a growing list of data replication technologies including: StorageTek Availability Suite (AVS) software, Hitachi TrueCopy and Universal Replicator, EMC Symmetrix Remote Data Facility (SRDF), Oracle Data Guard for Oracle Real Application Clusters (Oracle RAC), and MySQL database replication, all of which can be used to provide an integrated DR solution.

Oracle Solaris Cluster Geographic Edition provides an automated—not automatic—mechanism for migrating services from a cluster at the primary site to a cluster at the secondary, or disaster recovery, site. It facilitates the comprehensive process of coordinating the stopping and starting of the applications on each cluster, the direction of replication, and the sequence in which these should occur, which are defined by you. This enables you to make on-the-fly decisions about data recovery (for example, taking the time to reconstruct from other sources if required), use log files from unaffected application tiers, and so on.

Oracle Solaris Cluster Geographic Edition enables service continuity in case of planned outage. Services can be switched to secondary site when maintenance is required at the primary site. When that work is complete, services are switched back to the primary location, with minimal impact to clients. Oracle Solaris Cluster Geographic Edition also can be used to test DR readiness with minimal impact to production environments.

To further extend the flexibility of a disaster recovery solution, Oracle Solaris Cluster includes a script-based replication module that enables developers to integrate specific replication technologies with Oracle Solaris Cluster Geographic Edition for custom environments.

A Single HA and DR Solution for Multitier Oracle Applications and Databases

Oracle Solaris Cluster enables HA for local datacenters to business continuity and global disaster recovery solutions for evolving datacenter needs. The software leverages proven availability and virtualization features in Oracle Solaris 10 and Sun servers, and supports an industry-leading portfolio of enterprise applications across the database and business logic tiers.

- Oracle Solaris Cluster provides pretested, precertified software agents that support leading, off-the-shelf applications, including Oracle WebLogic Server, Oracle E-Business Suite, Siebel CRM and many others.
- In Oracle Database and Application environments, Oracle Solaris Cluster provides flexibility for the cluster infrastructure by supporting a wide range of networking and storage options such as InfiniBand, ASM, NAS, QFS, and hardware in thoroughly tested configurations.

Oracle Solaris Cluster supports Oracle 9*i*, 10*g*, and 11*g* R1 and R2 Database and Real Application Clusters (RAC), and is tightly integrated with Oracle Clusterware for improved coordination and

simplified administration. The integration of Oracle Clusterware with Oracle Solaris Cluster facilitates deployment of complex multitier solutions, including database and application components, and extends available configuration options. Oracle Solaris Cluster complements Oracle Database by:

- Supporting Oracle single instance and RAC Databases in virtualized environments such as Oracle Solaris Containers and Oracle VM Server for SPARC. Oracle Solaris Cluster enables the deployment of both single instance and RAC options in virtual clusters based on Oracle Solaris Containers, offering a unique solution for consolidated configurations that combines high availability and cost savings.
- Providing high availability to Oracle single-instance database deployments. Oracle Solaris Cluster offers comprehensive "whole-stack" HA capabilities for those environments where other Oracle HA products are not supported. For example, in environments where single-instance, pre Oracle 11g R2 is deployed, Oracle Solaris Cluster manages the failover to a secondary node, minimizing the potential for service outage.
- Supporting Oracle Data Guard with a complete, end-to-end Oracle RAC Database global disaster recovery solution. Oracle Solaris Cluster Geographic Edition enables automated DR in geographically-dispersed configurations. When disaster strikes, automated DR promotes higher confidence of success of the recovery procedures, because the procedures can be tested in advance, human errors happening in stress situation are avoided, and level of expertise for staff on duty is reduced.

Oracle Solaris Cluster complements Oracle Database for multitier deployments by offering:

- Unique coordination capabilities, resulting in faster failure discovery and recovery of overall service. For example, Oracle Solaris Cluster provides a mechanism to create dependencies between the database components (for example, Oracle Clusterware and the RAC Databases) and the underlying storage to ensure that both the start up and shutdown operations are correctly sequenced. The level of coordination is also enforced at other solution stack layers, including applications, middleware, and database.
- Load-balancing capabilities to optimize server load and ensuring that mission critical applications are restarted in priority to less important applications in case of failure of server.
- A global DR solution covering all application components supporting Data Guard for Oracle RAC and storage-based replications such EMC SRDF or Hitachi Universal Replicator for other needs.

Using Oracle Solaris Cluster in addition to Oracle Clusterware enables end-to-end availability of business service through inter-application and database coordination and management.

Oracle Solaris

Oracle Solaris delivers "designed in" reliability with features such as hardened device drivers[6] and Predictive Self Healing. Oracle Solaris Cluster is tightly coupled with Oracle Solaris, resulting in significantly faster failover detection, reconfiguration, service recovery, and end-to-end application failure. There are many technical advantages with this tight integration, such as low latency through kernel-level integration and hardened drivers that enhance overall robustness. Compared with other cluster and DR products, Oracle Solaris Cluster provides deeper and broader integration with the Oracle Solaris technologies. It works seamlessly with Oracle Solaris Containers, ZFS, Oracle VM Server for SPARC, and Dynamic Domains as found in Sun Enterprise M-Series servers.

Resource Management

By default, Oracle Solaris provides all workloads running on the system equal access to all system resources. This default behavior of Oracle Solaris can be modified by Solaris resource management capabilities. Resource management within Oracle Solaris provide:

- A method to classify a workload, so the system knows which processes belong to a given workload.
- The ability to measure the workload to assess how much of the system resources the workload is actually using.
- The ability to control the workloads so they do not interfere with one another and also get the required system resources to meet predefined service-level agreements.

Oracle Solaris resource management functionality enables you to constrain and schedule workloads, as well as ensure they get needed resources. There are two levels of granularity, projects and tasks, to identify a workload:

- *Projects* are a facility that allow the identification and separation of workloads. A workload can be composed of several applications and processes belonging to different groups and users.
- *Tasks* provide a second level of granularity in identifying a workload. A task collects a group of processes into a manageable entity that represents a workload component.

Oracle Solaris Cluster leverages these resource management capabilities to configure and monitor how much of a specific system resource is being used by an object type such as a node, disk, network interface, cluster resource group, or Solaris Zone. System resource usage can be part of your resource management policy. For example, you can control the CPU assigned to a resource group, and control the size of the processor set in which a resource group runs.

[6] Oracle Solaris device drivers are developed to help avoid system panics and hangs, and prevent the spread of corrupted data. You can learn more about Oracle Solaris defensive programming techniques at http://docs.sun.com/app/docs/doc/819-3196/gevsi?a=view.

Each construct has a configurable property to specify under which project the component should be started. So you can decide to use a specific project for a specific resource, or use one for all resources in the same resource group. If not specified, they are under the default project in Oracle Solaris.

Predictive Self Healing

Predictive Self Healing technology proactively monitors and manages system components to help organizations achieve maximum availability of IT services. Predictive Self Healing is an innovative capability in Oracle Solaris 10 that automatically diagnoses, isolates, and recovers from many hardware and application faults. This enables business-critical applications and essential system services to continue uninterrupted in the event of software failures, major hardware component failures, and even misconfigured software. The Oracle Solaris Fault Manager Architecture (FMA) and Oracle Solaris Service Manager Facility (SMF) are the two main components of Predictive Self Healing.

- **Fault Manager Architecture.** The FMA, a common system that works across platforms running Oracle Solaris, reduces complexity by automatically diagnosing faults in the system and initiating self healing actions to help prevent service interruptions. This software helps increase availability by configuring problem components out of a system before a failure occurs—and in the event of a failure, this feature initiates automatic recovery and application restart using SMF. The FMA diagnosis engine produces a fault diagnosis once discernible patterns are observed from a stream of incoming errors. Following diagnosis, FMA provides fault information to agents that know how to respond to specific faults.
- **Service Manager Facility.** The SMF facility creates a standardized control mechanism for application services by turning them into first-class objects that administrators can observe and manage in a uniform way. These services can then be automatically restarted if they are accidentally terminated by an administrator, if they are aborted as the result of a software programming error, or if they are interrupted by an underlying hardware problem. Administrators can easily move Oracle Solaris service management facility-enabled applications from a single-node Solaris 10 environment to a multi-node Oracle Solaris Cluster environment, increasing availability with little or no development effort.

A complementary technology is automatic system recovery, which facilitates immediate reboot of systems following a failure. These systems automatically configure around a failed component and notify the specified system administrator of the event. This approach prevents faulty or marginal hardware from causing repeated failures or keeping an entire system down.

Oracle Solaris Containers

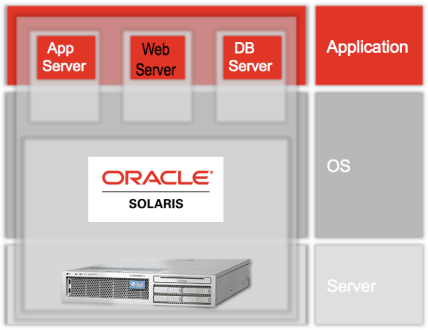


Figure 6: Oracle Solaris Containers provide secure, virtualized environments, with very little overhead.

Supported on any system running Oracle Solaris 10, Oracle Solaris Containers isolate software applications and services using flexible, software-defined boundaries. Oracle Solaris Containers provide virtualization and software partitioning, enabling the creation of many private execution environments from a single instance of Oracle Solaris.

Unlike virtual machines that require hypervisors, Oracle Solaris Containers provide OS-level virtualization by giving the appearance of multiple OS instances rather than multiple physical machines. Isolation between Containers is accomplished by restricting the scope of system calls, rather than the CPU-intensive task of emulating hardware architectures and instruction sets in software. This makes it possible to create hundreds, even thousands, of Oracle Solaris Containers on a single system. Because of this negligible overhead, and unlike partitioning or virtual machines, Oracle Solaris Containers can be created in large numbers. For example:

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

Hosting applications within individual Oracle Solaris Containers provides administrators the ability to exert fine-grained control over rights and resources within a consolidated server.

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

Oracle Solaris Containers cluster leverages and expands Oracle Solaris Container's technologies in application fault isolation, security isolation, and dedicated resource management. Each Oracle Solaris Containers cluster is a contained entity where a failure of an application is confined in that cluster and has no impact on other zone clusters in the same underlying physical hardware. Such isolation securely refrains read and write to the global cluster or other zone clusters to protect integrity of each zone

cluster. Oracle Solaris resource management can also be applied to the virtual cluster with dedicated administration policies.

Oracle Solaris 8 and Oracle Solaris 9 Containers

With Oracle Solaris 8 and Oracle Solaris 9 Containers, you can safely and easily move your existing applications and environments from a physical server running an older Solaris release to a software Container on the latest Sun SPARC Enterprise server running Oracle Solaris. This means you can run existing applications on 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 these existing environments the benefits of Oracle Solaris 10, such as Oracle Solaris DTrace, Predictive Self Healing, and Oracle Solaris ZFS.

Oracle VM Server for SPARC

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

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

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

Oracle Solaris Live Upgrade

Oracle Solaris Live Upgrade significantly reduces the downtime associated with an operating system upgrade or maintenance. Oracle Solaris Live Upgrade, combined with Oracle Solaris Cluster, enable you to migrate a boot environment to have different file system types, sizes, and layouts, without affecting the configuration of the installation software. You can also maintain on your system multiple installations of software packages, including Solaris operating system packages. It enables the Oracle Solaris 10 and previous versions to continue to run while an administrator upgrades to the latest release of the OS, applies patches, or does routine maintenance on the inactive or duplicate boot environment. When satisfied with the process, the administrator simply reboots the system to run the latest or updated operating environment.

Oracle Solaris Live Upgrade provides a method of upgrading that substantially reduces the usual service outage that is associated with an operating system upgrade. You can duplicate your current running boot environment; then, while the original boot environment continues to run you can upgrade the duplicate. Or, rather than upgrading, you can install a Web Start Flash archive on a boot environment. The original system configuration remains fully functional and unaffected by the upgrade or installation of a Web Start Flash archive. The duplicate boot environment is then activated to become the active boot environment when the system is rebooted. If the new boot environment does work as expected, you have a safety net. You can quickly fall back to the original boot environment with a simple reboot, thereby eliminating the service outages that are associated with the normal test and evaluation process.

ZFS

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

Oracle Solaris ZFS uses techniques such as copy-on-write and end-to-end *checksumming* to keep data consistent and eliminate silent data corruption. Data is written to a new block on the media before changing the pointers to the data and committing the write. Because the file system is always consistent, time-consuming recovery procedures like `fsck` are not required if the system is shut down in an unclean manner. In addition, data is read and checked constantly to help ensure correctness, and any errors detected in a mirrored pool are automatically repaired to protect against costly and time-consuming data loss and previously undetectable silent data corruption. Corrections are made possible by a RAID-Z implementation that uses parity, striping, and atomic operations to aid the reconstruction of corrupted data.

Oracle Solaris ZFS protects all data by 256-bit checksums, resulting in 99.99999999999999999-percent error detection and correction. 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, Solaris ZFS RAIDZ, or Solaris ZFS RAIDZ2), Oracle Solaris ZFS automatically repairs the corrupt data. This contributes to relentless availability by helping to protect against costly and time-consuming data loss due to hardware or software failure, and by reducing the chance of administrator error when performing file system-related tasks.

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

Oracle Solaris Networking Technology

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

Redundant Networking and Network IP Multipathing

In addition to traditional support for multiple network interfaces connected to different network subnets, Oracle Solaris also provides support for redundant network interfaces that are connected to a single subnet. IP Multipathing provides both failover and IP link aggregation.

There are a number of key features of redundant networking that work to improve the availability and performance of Sun systems are listed below.

- Failure detection, the ability to detect when a network adapter fails and automatically switch (failover) network access to an alternate network adapter.
- Repair detection, the ability to detect the repair of a previously failed network adapter and automatically switch back (fail back) the network access to this interface.
- Outbound load spreading, outbound network packets spread across multiple network adapters to achieve higher throughput. Load spreading occurs only when network traffic is flowing to multiple destinations using multiple connections.

Advanced Security Features

Oracle Solaris provides a sophisticated network-wide security system that controls the way users access files, protect system databases, and use system resources. From integrated security services and applications, to enhanced encryption algorithms, to an enterprise firewall for network protection, Oracle Solaris sets a high standard for operating system security by addressing security needs at every layer. 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 Cluster offers secure administrative capabilities through Oracle Solaris role-based access control. 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 10 offers RBAC, Process Rights Management, and least privilege. These technologies reduce security risk by granting users and applications only the minimum capabilities needed to perform tasks. System administrators can grant—or deny—a large number of discrete privileges to any process on the system to create effective security policies, minimize the likelihood of hostile actions, control access to data, and ensure compliance with regulatory requirements.

While traditional operating systems can contain inherent security weaknesses, Oracle Solaris 10 facilitates new approaches to protect the entire compute environment from the datacenter, through the network, and to the desktop.

Other Oracle Availability Products

Many Oracle Databases, Applications, and Fusion Middleware products, as well as MySQL, offer product-specific HA and scalability capabilities. Their HA functionality offers excellent failover of the product instance, with different capabilities for data integrity/resynchronization, failover time, virtualization, and so on.

Oracle Real Application Clusters (RAC) is an option to the award-winning Oracle Database Enterprise Edition. Oracle RAC is a cluster database with a shared cache architecture that overcomes the limitations of traditional shared-nothing and shared-disk approaches to provide highly scalable and available database solutions for all your business applications. Oracle RAC is a key component of Oracle enterprise grid architecture.

Oracle RAC utilizes *Oracle Clusterware* for the internode communication required in clustered database environments. Oracle Clusterware is the technology that transforms a server farm into a cluster. A cluster in general is a group of independent servers that cooperate as a single system. Oracle Clusterware is the intelligence in this system that ensures the required cooperation and is a key component of Oracle enterprise grid architecture.

Oracle Solaris Cluster integrates with Oracle Clusterware and Oracle RAC, ensuring that the entire solution framework is brought up, failed over, and shutdown in a coordinated fashion that recognizes the inherent dependencies that Oracle Clusterware has on the underlying storage components.

Oracle's *Maximum Availability Architecture (MAA)* is a framework and set of best practices for maximizing systems availability and meeting the most aggressive service level agreements (SLAs) for system availability, quality of service, and data protection.

Conclusion

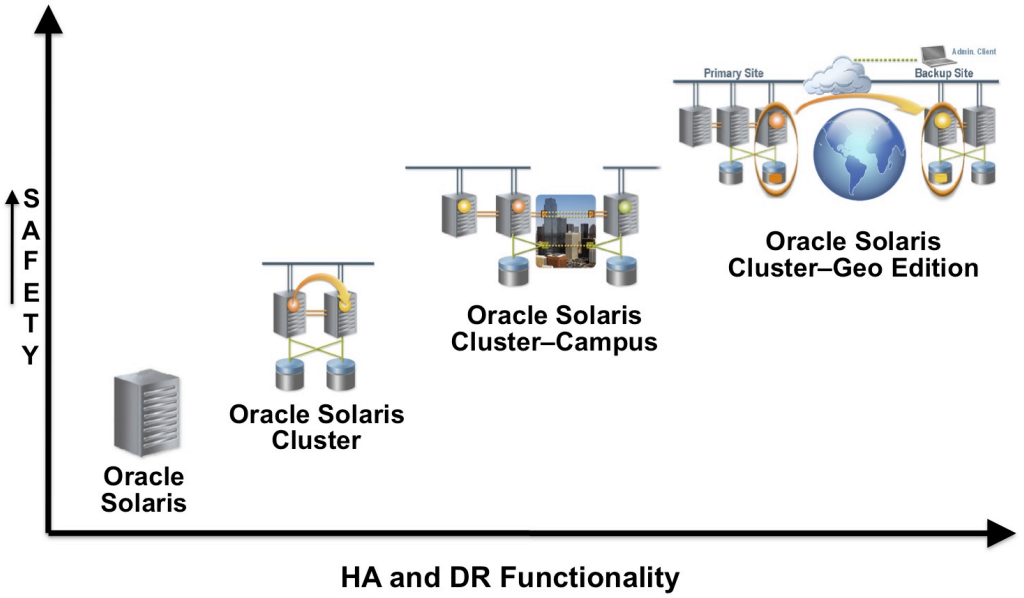


Figure 7: Oracle Solaris Cluster enhances the HA capabilities of Oracle Solaris, and provides solutions for enterprise HA and DR requirements.

Business risk can appear at your company's doorstep—or from within— at any time, and in many forms. Oracle Solaris Cluster delivers DR and best-in-class HA to business services in a 24x7 world. Only Oracle Solaris Cluster can be tailored to your specific needs, with the ability to work with a wide selection of IT technology. Oracle Solaris Cluster offers a full range of single- and multisystem HA and DR solutions for traditional and virtualized environments, with application to disk integration. Tightly coupled with the Oracle Solaris operating system, Oracle Solaris Cluster detects failures without delay; provides much faster failure notification, application failover, and reconfiguration time; and significantly reduces services recovery time.

Oracle Solaris Cluster is tested extensively with Oracle Database, Oracle Applications, and Fusion Middleware products, along with leading applications from other vendors. Oracle Solaris Cluster complements Oracle Real Application Clusters (RAC) and Oracle Data Guard, offering additional uptime assurance and flexibility for multitier deployments.

More Information

GET THE SOFTWARE	
Oracle Solaris Cluster	www.oracle.com/technetwork/server-storage/solaris-cluster/downloads/index.html
Oracle Solaris 10	www.oracle.com/technetwork/server-storage/solaris/downloads/index.html
HIGH AVAILABILITY	
Oracle Solaris Cluster Technical Information	www.oracle.com/technetwork/server-storage/solaris-cluster/overview/index.html

Oracle Solaris Cluster Documentation Center	http://docs.sun.com/app/docs/doc/821-1261/
Oracle Solaris Cluster Essentials, by Tim Read	www.amazon.com/Oracle-Solaris-Cluster-Essentials-Administration/dp/0132486229/
Patching and Upgrading Oracle Solaris Cluster	http://docs.sun.com/app/docs/doc/821-1256/z4000076997776?l=en&a=view
How-To Guide: How-To Install a 2 Node Cluster	www.oracle.com/technetwork/articles/servers-storage-admin/how-to-install-two-node-cluster-166763.pdf
How-To Guide: Configure Oracle VM Server for SPARC Guest Domains as Oracle Solaris Cluster Nodes	www.oracle.com/technetwork/articles/servers-storage-admin/configvmsrvrdomssolarisnodes-166764.pdf
Oracle Solaris Cluster Geographic Edition Webcast	http://medianetwork.oracle.com/media/show/13298?n=subCategory&nid=33
IDC: Addressing Virtualization and High Availability Needs with Solaris Cluster	www.oracle.com/us/products/servers-storage/solaris/idc-virt-ha-solaris-cluster-wp-070226.pdf
Oracle Real Application Clusters in Oracle VM Environments	www.oracle.com/technetwork/database/clustering/oracle-rac-in-oracle-vm-environment-131948.pdf
Oracle Maximum Availability Architecture—MAA	www.oracle.com/technetwork/database/features/availability/maa-090890.html
Solaris Cluster Concepts Guide (including Solaris Cluster topologies)	http://dlc.sun.com/pdf/821-0259/821-0259.pdf
Oracle Database High Availability	www.oracle.com/technetwork/database/features/availability/index.html
Oracle Fusion Middleware High Availability and Enterprise Deployment	http://download.oracle.com/docs/cd/E12839_01/relnotes.1111/e10133/ha.htm
Oracle Fusion Middleware Deployment Planning Guide for Oracle Directory Server Enterprise Edition, “Using Replication and Redundancy for High Availability”	http://docs.sun.com/app/docs/doc/821-1502/fpcqm?l=en&n=1&a=view
Oracle Fusion Middleware—High Availability	www.oracle.com/technetwork/middleware/index-098790.html
ORACLE SOLARIS	
Oracle Solaris 10 Technical Information	www.oracle.com/technetwork/server-storage/solaris/overview/index.html
Oracle Solaris: In a Class by Itself – White Paper	www.oracle.com/technetwork/articles/servers-storage-admin/solclassbyitsel-167850.pdf
Using Oracle Solaris 10 to Overcome Security Challenges	www.oracle.com/technetwork/articles/solarisadmin/solaris10securitychallenges-163842.pdf
VIRTUALIZATION	
Oracle Virtualization	www.oracle.com/virtualization
Virtualization with Oracle Solaris 10 White Paper	www.oracle.com/us/products/servers-storage/solaris/virtualization-solaris-wp-080927.pdf
Oracle VM Server for SPARC—Enabling a Flexible, Efficient IT Infrastructure	www.oracle.com/us/products/servers-storage/servers/sparc-enterprise/vm-server-for-sparc-wp-075964.pdf
ADDITIONAL RESOURCES	
Oracle Database 11*g*	www.oracle.com/us/products/database/
Oracle Siebel CRM	www.oracle.com/us/products/applications/siebel/
Oracle PeopleSoft Enterprise Applications	www.oracle.com/us/products/applications/peoplesoft-enterprise/
Oracle Real Application Clusters (RAC)	www.oracle.com/technology/products/database/clustering/

Oracle Applications	www.oracle.com/us/products/applications/index.html
Oracle Technical Network—Technical Information	http://www.oracle.com/technetwork/index.html

The following people contributed to this white paper:

Chris Baker, Stephanie Choyer, Burt Clouse, Duncan Hardie, Eve Kleinknecht, Amour Kwok, Steve McKinty

Oracle Solaris and Oracle Solaris Cluster:
Extending Oracle Solaris for Business Continuity
September 2010

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

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

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

SOFTWARE. HARDWARE. COMPLETE.