An Oracle White Paper
April 2010

Increasing Application Availability Using Oracle VM Server for SPARC:
An Oracle Database Example

ORACLE®

Introduction

Today, business challenges include creating and preserving value in a highly competitive environment, while keeping business applications up and running and reducing costs. One of the most challenging aspects of data center management is improving the availability of business applications during planned or unplanned outages.

This document provides technical insight about increasing application availability using Oracle VM Server for SPARC technologies (previously called Sun Logical Domains or LDoms).

Using an example and various scenarios, this paper describes how to take advantage of the domain mobility capability (also called "warm migration") in Oracle VM Server for SPARC version 1.3 to increase the availability of an Oracle Database 10g Release 2 single-instance database.

Note: Oracle 10gR2 database with Oracle VM Server for SPARC is certified for Oracle Solaris 10 5/08 (Update 5) and later. Oracle Solaris 10/09 (Update 8) was used for this paper to support the domain mobility feature of Oracle VM Server for SPARC version 1.3.

Fundamentals

This section provides background information.

High Availability Fundamentals

High availability is an approach to satisfy business availability needs. A distinction needs to be made between planned downtime and unplanned downtime.

Typically, planned downtime is a result of a preventive or corrective maintenance task that is an interruption to day-to-day system operation, and usually this downtime cannot be delayed due to constraints imposed by the system's software and hardware design.

A planned outage should be scheduled to have a minimum impact on the availability of a system, because despite the planned aspect of this type of outage, even a minor modification to subsystems can cause a disruption to availability and impact business needs.

Planned outages are the result of maintenance events, such as repair, backup, or upgrade operations. Repairs are intended to remove faulty components and restore a system to a functional state. Upgrades are implemented to replace the current hardware or software with newer (or enhanced) versions

Unplanned downtime is a result of hardware faults and environmental events. It is usually urgent to resolve such downtime, and addressing this type of downtime is out of the scope of this document.

Introduction to Oracle VM Server for SPARC

Oracle VM Server for SPARC is a virtualization and partitioning solution supported on Oracle Solaris CoolThreads technology-based servers powered by UltraSPARC T1, T2, and T2 Plus processors with Chip Multithreading Technology (CMT).

Oracle VM Server for SPARC allow the creation of multiple virtual systems on a single physical system. Each virtual system is called a *logical domain* (LDom) and runs its own copy of the operating system.

Oracle VM Server for SPARC enables fine-grained assignment of CPU and memory resources to a workload. Among many of the LDom features, we will demonstrate the migration of active domains using the domain mobility (warm migration) feature. The ability to do warm migration between two machines, with minimal disruption to applications, provides a powerful platform for production environments.

Introduction to Oracle Database

Oracle Database is the leading database application, and as such, it is frequently used on servers. With virtualization and consolidation, Oracle Database can now run on selected, certified virtual environments, such as Oracle VM Server for SPARC.

Domain Mobility Overview

Domain mobility provides the ability to migrate a guest LDom from one host machine to another. A distinction needs to be made between *cold migration* and *warm migration*.

With cold migration, the operating system needs to be in shutdown state before starting the migration, while with warm migration, the operating system and applications are running before and after the migration, and they are suspended during the migration process.

Warm migration has a set of requirements and restrictions on CPUs, memory, and I/O that must be satisfied on the source and target hosts.

The host where the migration is initiated is referred to as the *source machine* and the host where the domain is migrated to is referred to as the *target machine*.

Similarly, once a migration is started, the domain to be migrated is referred to as the *source domain*, and the domain created on the target machine is referred to as the *target domain* while the migration is in progress.

Benefits of Warm Migration

Here are some benefits of warm migration:

- Maintenance—In case of scheduled service downtime or some urgent environmental problem, the LDom mobility can result in less downtime. If there is a need to shut down some equipment, the administrator can preserve running applications by moving them to some other server.

- Hardware resource—If an administrator sees a better way to deploy an application over different servers, the administrator can use mobility to move a running domain to a machine with more physical memory, more CPU capacity, or a better I/O subsystem, thereby improving the general performance of the application.

- Application memory structures stay warm—Any application that uses memory intensively will keep all memory structures in-place. For example, after a warm migration, an Oracle Database will have its System Global Area (SGA) and other memory structures laid out the same way as before the migration.

- Long computations—Applications that do long computations, such as financial data analysis or scientific computations, do not need to restart from the beginning or have any intermediate data save/restore procedures.

Software Requirements

The following sections describe requirements.

Oracle VM Server for SPARC Requirements

Refer to the latest release notes for Oracle VM Server for SPARC. At the time of this writing, that was *Logical Domains 1.3 Release Notes* (http://sun.com/ldom).

Oracle Database Requirements

Running an Oracle database single-instance configuration on Oracle VM Server for SPARC requires the use of Oracle Database 10*g*R2 (Oracle 10.2.0.3). Later versions might also be supported. Refer to the Oracle documentation to check if a particular version is supported and for any additional requirements. Information can be found on the Oracle MetaLink Web site (registration required):

http://www.oracle.com/technology/support/metalink/index.html

Hardware Configuration

Any server supporting Oracle VM Server for SPARC can be used with the Oracle database. Some additional hardware, such as external storage arrays or network switches, might be needed depending on the selected configuration and deployment variant.

Figure 1 shows a typical hardware configuration and cabling compliant with domain mobility. This example uses the following hardware:

- Two Sun SPARC Enterprise T5240 servers from Oracle, each with 128 CPU threads and 64 GB of memory. These servers contain two UltraSPARC T2 Plus processors. Each UltraSPARC T2 Plus processor is made of 8 CPU cores, and each CPU core has 8 CPU threads, providing a total of 128 (2x8x8) CPU threads per system.

- One internal disk is used on each server as the system disk of the control domain. Oracle's Sun SPARC Enterprise T5240 server can have up to sixteen internal disks.

- One Sun StorageTek 6140 storage array from Oracle, with sixteen 280 GB Fiber Channel (FC) disk drives. The storage array is connected to both servers, and each server is connected to the same array controller (A or B) to which LUNs are mapped. The storage array provides one LUN used as the system disk of the guest domain and a second LUN to store the Oracle data.

- Two 4 GB FC-AL host bus adapters (HBAs), one for each server. These HBAs are used to connect the storage array to both servers.

- The on-board network interfaces present on each server are used for network communications. Oracle's Sun SPARC Enterprise T5240 server has four on-board 1-Gb Ethernet network interfaces.

Note: For a faster migration, you could use a 10-Gb Ethernet network interface that can be added into a PCI bus slot.
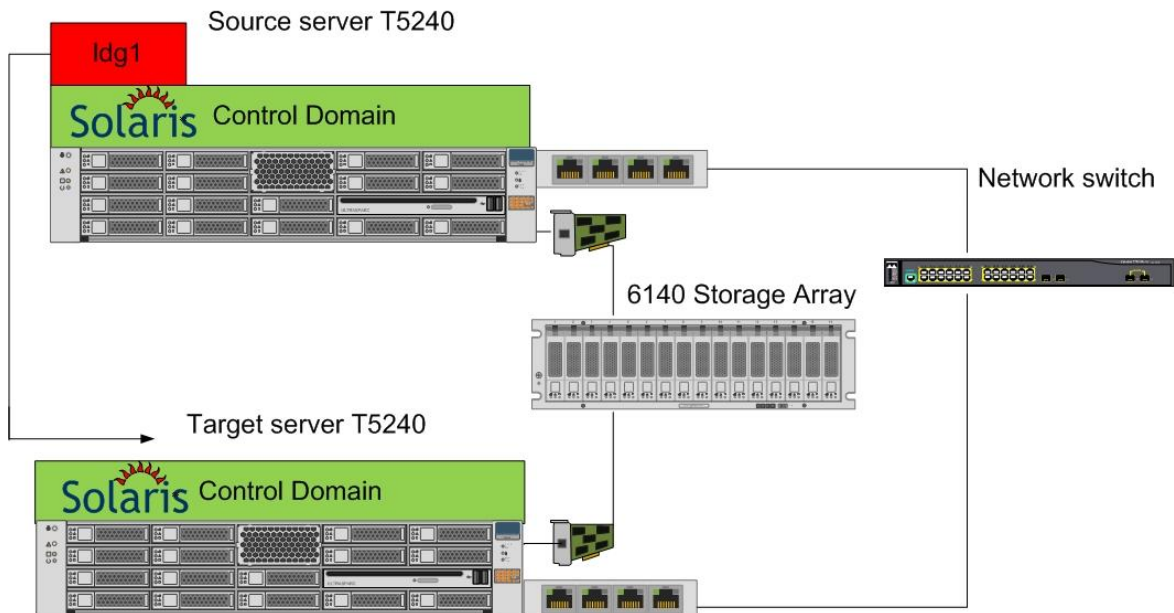

Figure 1. Hardware Configuration and Cabling

Running Oracle Database on Oracle VM Server for SPARC

Each domain running Oracle Database should have a minimum of 8 virtual CPUs (1 CPU core) and 4 GB of memory. See the Sun Blueprints paper *Running Oracle Real Application Clusters (RAC) on Sun Logical Domains* (registration required) for more information:

http://wikis.sun.com/display/BluePrints/Running+Oracle+Real+Application+Clusters+(RAC)+on+Sun+Logical+Domains

Each service domain providing virtual devices (virtual disk and virtual network) to an Oracle single-instance guest domain should have a minimum of 16 virtual CPUs (2 CPU cores) and 4 GB of memory.

Configuration Details

In the example presented in this document, the source server is configured with two domains: the control domain and one guest domain (named `ldg1`) hosting the Oracle database. The control domain is also the primary domain, and therefore, it is also an I/O domain. Furthermore, the control domain is used as a service domain providing virtual device services (virtual disk and virtual network) to the guest domain.

To act as a service domain, the control domain has the following virtual device services:

- One virtual disk service (`primary-vds0`) used to export physical disks as virtual disks to the guest domain.

- One virtual console concentrator service (`primary-vcc0`) used to provide access to the virtual console of the guest domain. This service will use the port range 5000–5100 for virtual consoles access.

- One virtual switch service (`primary-vsw0`) associated with the primary network interface (`nxge0`).

Software Installation and Configuration

The following tasks are part of the software installation and configuration:

- Oracle Solaris installation

- Oracle VM Server for SPARC installation

- Storage configuration

- Control domain configuration

- Guest domain configuration

- Oracle Solaris installation in the guest domain

- Oracle Database installation

- Target server preparation

Oracle Solaris Installation

Initially, Oracle VM Server for SPARC is not enabled on Oracle Solaris CoolThreads technology-based servers. The first step is to install the Oracle Solaris OS and the appropriate patches. Refer to the Oracle Solaris installation documentation for more information on how to install the OS on servers with SPARC processors. In our example, the operating system is installed on the first internal disk of the server.

New servers come pre-installed with the Solaris OS. For these systems, ensure that the appropriate release of Oracle Solaris is installed and check that the required patches are present.

After Oracle Solaris is installed, the system can be configured and enabled to use Oracle VM Server for SPARC.

Oracle VM Server for SPARC Installation

Refer to the *Logical Domains 1.3 Administration Guide* (http://docs.sun.com/app/docs/doc/821-0406) for a complete procedure on how to install Oracle VM Server for SPARC.

This example assumes that Oracle Solaris and the required patches are already installed on each server.

Storage Configuration

The storage array is connected to the control domain, which exports two LUNs (LUN 0 and LUN 1) as virtual disks to the guest domain.

The guest domain uses the first LUN (LUN 0) as the system disk for the operating system and the second LUN (LUN 1) as a disk to store Oracle data files.
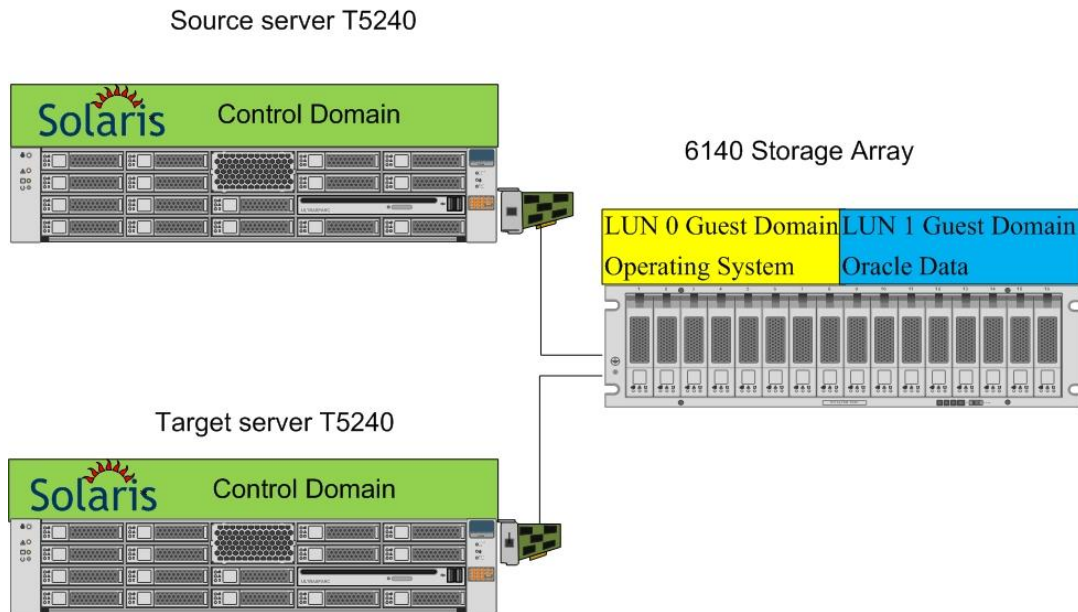
Figure 2 shows the storage layout.

Figure 2. Storage Configuration

Control Domain Configuration

After the system reboots, LDoms will be enabled and the system will now be configured with one domain: the control domain (primary domain). From the control domain, additional domains can then be created and configured.

Guest Domain Configuration

After the control domain has been configured, create the guest domain to be used as the Oracle database node. Only one guest domain is created on the physical server.

The guest domain `ldg1` is initially created with the following resources:

- 8 GB of memory

- 24 virtual CPUs

- One virtual network interface (`vnet0`) connected to the virtual switch `primary-vsw0`.

- One virtual disk, which appears as `c0d0` in the guest domain and is a LUN from the storage array. The domain `ldg1` uses LUN 0 of the storage array (`c2t6d0`).

- A second virtual disk, which appears as `c0d1` in the guest domain and is a LUN from the storage array. The domain `ldg1` uses LUN1 of the storage array (`c2t6d1`). This virtual disk will be used for storing Oracle database files on a UFS file system.

Refer to the *Logical Domains 1.3 Administration Guide* (http://docs.sun.com/app/docs/doc/821-0406) for a complete procedure on how to install guest domains.

If the guest domain (ldg1) was initially configured and installed with the first LUN (LUN 0), then the second LUN (LUN 1) can be added by executing commands such as the following from the control domain:

```
# ldm add-vdsdev /dev/dsk/c2t6d1s2 oradata@primary-vds0
```

```
# ldm add-vdisk oradata oradata@primary-vds0 ldg1
```

The new disk will be immediately added to the guest domain. Once the disk is visible, use the format(1m) command from the guest domain to partition the new disk and create the required partition for Oracle data files.

Note: LUNs can (and often do) appear with different names on different servers. LUN names in control domains are not required to be identical. But both control domains are required to virtualize the same LUNs (LUN 0 and LUN 1) with the same volume names (vol1 and oradata).

The following example shows how a domain, called ldg1 is configured with 8CPUs, 24GB of memory, and two LUNs from the Sun StorageTek 6140 storage array.

```
#ldm ls -l ldg1
NAME            STATE      FLAGS   CONS    VCPU  MEMORY   UTIL  UPTIME
ldg1            active     -n----  5000    24    8G       0.0%  18h 14m
SOFTSTATE
Solaris running
MAC
    00:14:4f:fb:96:89
HOSTID
    0x84fb9689
CONTROL
    failure-policy=ignore
DEPENDENCY
    master=
VCPU
    VID    PID    UTIL STRAND
    0      16     0.2%  100%
    1      17     0.0%  100%
    2      18     0.5%  100%
    3      19     0.0%  100%
    4      20     0.0%  100%
    5      21     0.0%  100%
    6      22     0.1%  100%
```

```
    7       23      0.0%   100%

    8       24      0.1%   100%

    9       25      0.0%   100%

    10      26      0.0%   100%

    11      27      0.0%   100%

    12      28      0.0%   100%

    13      29      0.0%   100%

    14      30      0.0%   100%

    15      31      0.0%   100%

    16      32      0.0%   100%

    17      33      0.0%   100%

    18      34      0.0%   100%

    19      35      0.0%   100%

    20      36      0.0%   100%

    21      37      0.0%   100%

    22      38      0.0%   100%

    23      39      0.1%   100%

MEMORY

    RA              PA              SIZE

    0x8000000       0x408000000     8G

VARIABLES

    boot-device=disk:a vdisk

    keyboard-layout=US-English

NETWORK

    NAME            SERVICE                     ID   DEVICE      MAC
MODE   PVID VID              MTU    LINKPROP
    vnet1           primary-vsw0@primary        0    network@0   00:14:4f:f9:c0:62
1                        1500

DISK

    NAME            VOLUME                  TOUT ID   DEVICE   SERVER
MPGROUP
    vdisk1          vol1@primary-vds0            0    disk@0   primary
    oradata         oradata@primary-vds0         1    disk@1   primary

VCONS

    NAME            SERVICE                 PORT
    ldg1            primary-vcc0@primary    5000
```

The following example shows what the configuration looks on the control domain. You can verify the services have been created by using the list-services subcommand.

```
# ldm ls-services primary
VCC
    NAME            LDOM            PORT-RANGE
    primary-vcc0    primary         5000-5100
VSW
    NAME            LDOM            MAC                 NET-DEV   ID    DEVICE
LINKPROP   DEFAULT-VLAN-ID PVID VID       MTU    MODE
    primary-vsw0    primary         00:14:4f:f9:32:b0 nxgeg0    0     switch@0
1               1        1500

VDS
    NAME            LDOM        VOLUME      OPTIONS        MPGROUP         DEVICE
    primary-vds0    primary     vol1                      /dev/dsk/c2t6d0s2
                                oradata                   /dev/dsk/c2t6d1s2
```

Oracle Solaris Installation in the Guest Domain

Ensure that the appropriate release of Oracle Solaris is installed in the guest domain, and check that the required patches are present.

After Oracle Solaris is installed in the guest domain, the system can be configured and enabled to use the Oracle database.

Oracle Database Installation

Installation of the Oracle database software is similar to a standard Oracle installation. Start with the installation of Oracle Database and then apply the latest patch set. Use the second LUN to create a UFS file system for storing Oracle data files.

Target Server Preparation

The target server needs to have access to both LUN 0 and LUN 1 from the same storage.

To add the first LUN (LUN 0) and second LUN (LUN 1) to the target machine, execute the following commands from the control domain of the target server:

```
# ldm add-vdsdev /dev/dsk/c2t6d0s2 vol1@primary-vds0
```

```
# ldm add-vdsdev /dev/dsk/c2t6d1s2 oradata@primary-vds0
```

Note: LUNs can (and often do) appear with different names on different servers. LUN names in control domains are not required to be identical. But both control domains are required to virtualize the same LUNs (LUN 0 and LUN 1) with the same volume names (vol1 and oradata).

Warm Migration Process

The main steps in the process are as follows:

- The target domain is created (but inactive).

- The source domain is suspended.

- The source domain memory and state is transferred to the target.

- The target domain is resumed.

Note that the domain is unavailable while the domain is suspended.

The Oracle VM Server for SPARC Manager on the source machine accepts the request to migrate a domain and establishes a secure network connection with the Oracle VM Server for SPARC Manager running on the target machine. Once this connection has been established, the migration occurs. The migration itself can be broken down into different phases.

Phase 1: After connecting with the Oracle VM Server for SPARC Manager running in the target host, information about the source machine and domain are transferred to the target host. This information is used to perform a series of checks to determine whether a migration is possible.
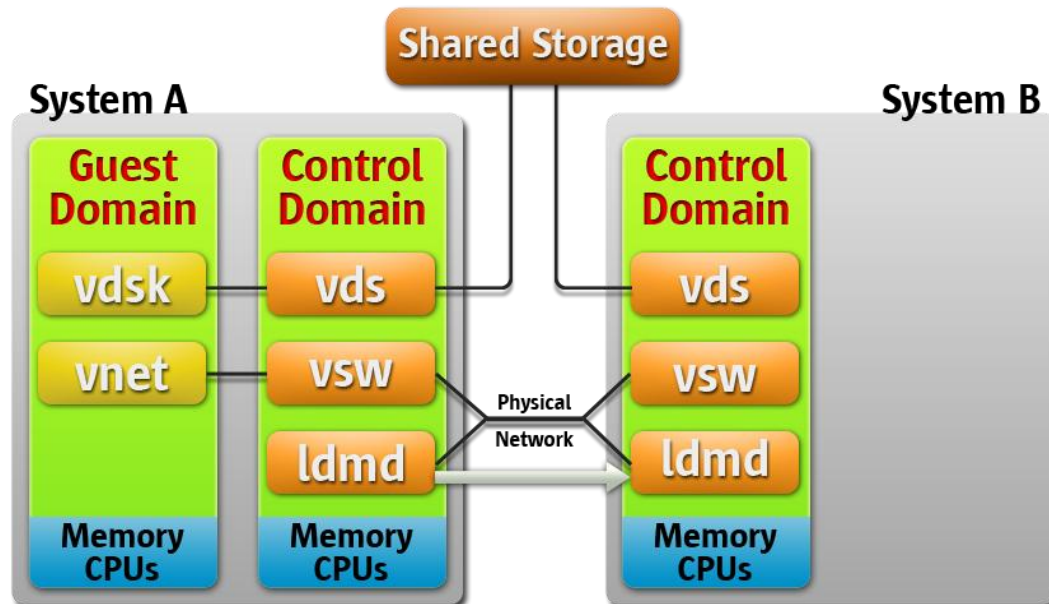

Figure 3. Phase One

Phase 2: When all checks in Phase 1 have passed, the source and target machines prepare for the Migration. This preparation includes shrinking the number of CPUs to one and suspending the domain. On the target machine, a domain is created to receive the source domain.

Phase 3: The next phase is to transfer all the runtime state information for the domain to the target. The information transferred is the CPU states and the domain memory.

Phase 4: Handoff occurs. After all state information is transferred, the handoff occurs when the target domain resumes execution and the source domain is destroyed. From this point on, the target domain is the sole version of the domain running.
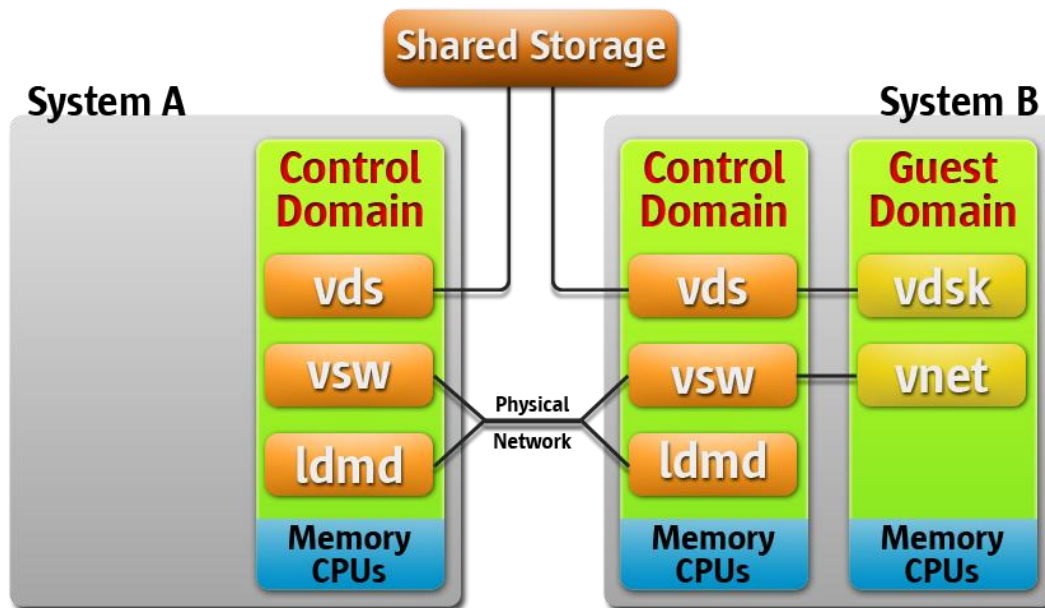

Figure 4. Phase Four

Migration Examples

A guest domain can be migrated from one system to another using the `ldm migrate-domain` command.

When you provide the `-n` option to the `migrate-domain` subcommand, migration checks are performed, but the source domain is not migrated. Any requirement that is not satisfied is reported as an error. This allows you to correct any configuration errors before attempting a real migration.

Note: Due to the dynamic nature of logical domains, it is possible for a dry run to succeed and a migration to fail and vice-versa.

The following example shows how to perform a dry-run migration so a domain called `ldg1` can be migrated (dry run) to a machine called `t5440`.

```
# ldm migrate-domain -n ldg1 t5440
Target Password: [password]
```

The following example shows how a domain called `ldg1` can be migrated to a machine called `t5440`:

```
# ldm migrate-domain ldg1 t5440
Target Password:
```

To perform this migration automatically, without being prompted for the target password, use the following command:

```
# ldm migrate-domain -p pfile ldg1 t5440
```

The `-p` option takes a file name as an argument. The specified file contains the super user password for the target. In this example, `pfile` contains the password for the target, `t5440`.

To monitor the migration progress, use the following command:

```
# ldm list -o status ldg1
```

Note: The entire process of warm migration took approximately six minutes to accomplish with the current settings. Time might vary depending on amount of memory assigned and network utilization.

You can reduce the migration by connecting the LDoms to a 10Gb network or using a private dedicated network for the migration.

Caution: The following is normal behavior and should be considered before starting process:

- The Oracle clients and application may experience timeouts and disconnects, depending on how long the migration process takes. Application should be ready to process such situation correctly.

- No transactions are performed while the domain is being migrated. Transactions will continue to run as if the domain was not disconnected if there are no timeouts.

- The operating system and all processes are suspended while migration is happening; as a result, the Oracle database will not detect CPU changes.

Conclusion

There are several approaches to ensuring you have highly available applications if there is a planned or unplanned outage.

Oracle VM Server for SPARC technology can provide useful features to increase application availability.

This paper presented an overview and benefits of the Oracle VM Server for SPARC domain mobility feature.

An example described the complete configuration process, including creation and configuration of domains, storage configuration, and software requirements to help administrators plan for their deployments of an application on Oracle VM Server for SPARC.

The Oracle Database can be installed on servers configured with Oracle VM Server for SPARC. This virtualization technology provides features that allow the mobility of Oracle database systems using Oracle VM Server for SPARC warm migration. This technology also provides better availability and flexibility for efficiently allocating hardware and deploying a production environment.

About the Authors

Orgad Kimchi joined Sun in September 2007. He is currently working in the Independent Software Vendors (ISV) Engineering organization helping software vendors adopt Sun technology and improve performance on Sun hardware and software. Orgad's blog can be found at http://blogs.sun.com/vreality.

Roman Ivanov joined Sun in January 2006. He is currently working in the Independent Software Vendors (ISV) Engineering organization helping software vendors adopt Sun technology and improve performance on Sun hardware. Roman's blog can be found at http://blogs.sun.com/pomah.

Acknowledgments

The authors would like to thank Alexandre Chartre and George Drapeau for their contributions to this document.

For More Information

Here are additional Sun resources.

Oracle Solaris resources:

- BigAdmin web applications and device lists:
  - Hardware Compatibility Lists for Solaris OS: http://www.sun.com/bigadmin/hcl/data/sol/
  - Hardware Certification Test Suite: http://www.sun.com/bigadmin/hcl/hcts/index.jsp
  - Solaris 10 Applications Library: http://www.sun.com/bigadmin/apps/
  - Sun Device Detection Tool: http://www.sun.com/bigadmin/hcl/hcts/device_detect.jsp

- Installation Check Tool: http://www.sun.com/bigadmin/hcl/hcts/install_check.jsp

- Device and Third-party Solaris Device Driver Reference Lists:
  http://www.sun.com/bigadmin/drivers

- BigAdmin Oracle Solaris technology resource centers:

  - Solaris Information Center: http://www.sun.com/bigadmin/hubs/documentation

  - Solaris Patching Center: http://www.sun.com/bigadmin/patches/solaris/index.jsp

  - Solaris 10 Upgrade Resources for System Administrators:
    http://www.sun.com/bigadmin/topics/upgrade/

  - Solaris Containers (Zones): http://www.sun.com/bigadmin/content/zones/

  - Logical Domains (LDoms): http://www.sun.com/bigadmin/hubs/ldoms/

  - DTrace: http://www.sun.com/bigadmin/content/dtrace/index.jsp

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

  - Predictive Self-Healing: http://www.sun.com/bigadmin/content/selfheal/

  - Solaris 8 Vintage Support: http://www.sun.com/bigadmin/topics/vintagepatch/

- BigAdmin Oracle Solaris resource collections (which include community submissions):

  - Solaris resource collection: http://www.sun.com/bigadmin/collections/solaris.jsp

  - Solaris 10 resource collection: http://www.sun.com/bigadmin/collections/solaris10.jsp

  - Solaris 9 resource collection: http://www.sun.com/bigadmin/collections/solaris9.jsp

  - Solaris 8 resource collection: http://www.sun.com/bigadmin/collections/solaris8.jsp

  - Solaris on x86 resource collection:
    http://www.sun.com/bigadmin/collections/solarisx86.jsp

- Discussions, such as the Solaris OS forums:
  http://forums.sun.com/category.jspa?categoryID=65

- BigAdmin Operating System wiki page:
  http://wikis.sun.com/display/BigAdmin/Operating+System

Virtualization resources:

- Information Exchange for Oracle Enterprise Manager Ops Center 2.5:
  http://wikis.sun.com/display/infoexchange/Home

- Information Exchange for Oracle Enterprise Manager Ops Center releases prior to version
  2.5: http://wikis.sun.com/display/xvminfoex/Home

- Virtualization web site: http://www.sun.com/solutions/virtualization/index.jsp

- Oracle Solaris Containers web site:
  http://www.sun.com/software/solaris/containers/index.jsp

- Information for system administrators on BigAdmin:
  http://www.sun.com/bigadmin/home/index.jsp

  - Hardware Compatibility List for Virtualized Platforms:
    http://www.sun.com/bigadmin/hcl/data/v12n/

  - Virtualization Resource Center: http://www.sun.com/bigadmin/topics/virtualization/

  - Virtualization Resource Collection (includes third-party resources):
    http://www.sun.com/bigadmin/collections/virtualization.jsp

  - Sun Docs articles (on virtualization and other topics):
    http://www.sun.com/bigadmin/sundocs/

- BigAdmin Virtualization wiki page: http://wikis.sun.com/display/BigAdmin/Virtualization

General links:

- Sun download site: http://www.sun.com/download/

- Oracle University web site: http://www.sun.com/training/

- Discussions, such as Sun forums (http://forums.sun.com/index.jspa) and the BigAdmin
  Discussions collection (http://www.sun.com/bigadmin/discussions/)

- Sun product documentation at http://docs.sun.com and the Sun Documentation Center
  (http://www.sun.com/documentation/)

- Sun wikis, such as the Sun BluePrints wiki
  (http://wikis.sun.com/display/BluePrints/Main)and the BigAdmin wiki
  (http://wikis.sun.com/display/BigAdmin/Home)

- Support:
  - Sun resources:

    - o   Register your gear: https://inventory.sun.com/inventory/

    - o   Sun Services: http://www.sun.com/service/index.jsp

    - o   SunSolve Online: http://sunsolve.sun.com

  - Community system administration experts:
    http://www.sun.com/bigadmin/content/communityexperts/

ORACLE

Increasing Application Availability Using Oracle
VM Server for SPARC: An Oracle Database
Example
April 2010
Authors: Orgad Kimchi and Roman Ivanov

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