

UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA



DISEÑO DE UNA RED EN ALTA DISPONIBILIDAD CON SUN
CLUSTER

INFORME DE SUFICIENCIA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ELECTRÓNICO

PRESENTADO POR:

ROGGER DAVID VALDEZ BARRIENTOS

**PROMOCIÓN
2003 - II**

LIMA – PERÚ

2009

DISEÑO DE UNA RED EN ALTA DISPONIBILIDAD CON SUN CLUSTER

Dedicado a todas las personas que realmente
me apoyaron en todo momento. Por la
permanente búsqueda del conocimiento.

SUMARIO

El presente trabajo describe el diseño y los pasos necesarios para la instalación y configuración de la plataforma necesaria para permitir que una base de datos oracle accedida por 2 servidores y un arreglo de discos, este disponible y a prueba de fallos de alguno de los servidores. En la primera parte se introducen varios conceptos importantes para el entendimiento del diseño realizado de la instalación. Luego se describe la instalación del sistema operativo, configuración del arreglo de discos, instalación del programa Sun Cluster que permite la disponibilidad de la base de datos, la configuración de dispositivos de disco, el registro y creación de módulos del cluster para la base de datos, y pruebas finales del cluster.

INDICE

Introducción.....	1
-------------------	---

CAPITULO I

DESCRIPCIÓN DEL PROBLEMA Y MARCO TEORICO

1.1 Descripción del Problema.....	3
1.2 Objetivo del Trabajo.....	4
1.3 Evaluación de los Equipos Disponibles.....	4
1.4 Limitaciones del Trabajo.....	4
1.5 Elementos de un centro de datos.....	5
1.6 Medios de transporte de datos en redes de alta disponibilidad.....	5
1.6.1 SCSI.....	5
1.6.2 Fibra Óptica.....	6
1.7 RAID.....	7
1.7.1 RAID 0	7
1.7.2 RAID 1.....	7
1.7.3 RAID 5	8
1.7.4 Disco Spare	9
1.8 Protecciones en alta disponibilidad	9
1.8.1 Fuente Redundante.....	9
1.8.2 Tarjetas RAID	9
1.8.3 Gestor de Volúmenes o Volume Manager.....	10
1.8.4 Multipath o Multiruta.....	10
1.8.5 Protección en Storage	10
1.9 Cluster.....	10

CAPITULO II

CARACTERÍSTICAS DE LOS PROGRAMAS UTILIZADOS

2.1 Sun Solaris	12
2.1.1 Particionamiento de Discos	13
2.1.2 Archivo de inicio de Sistemas de Archivos o Filesystems	13
2.1.3 Archivo de resolución de nombres.....	13

2.2	Sun Cluster	13
2.3	Elementos de Sun Cluster	13
2.3.1	Nodos	14
2.3.2	Storage	14
2.3.3	Consola de Administración	14
2.3.4	Red de Administración	14
2.3.5	Red de Almacenamiento	14
2.3.6	Red de Interconexión	14
2.3.7	Red de Producción	15
2.3.8	Quorum	15
2.4	Solaris Volume Manager	16
2.4.1	Metadb	16
2.4.2	Metaset o Disk set	16
2.5	Dispositivos de Sun Cluster	16

CAPITULO III

INSTALACIÓN DE SISTEMA OPERATIVO Y SUN CLUSTER

3.1	Componentes utilizados	18
3.2	Instalación del Sistema operativo en nodos	18
3.3	Configuración de storage	24
3.4	Instalación del Software Sun Cluster	26
3.4.1	Permitir el acceso de un nodo a otro	26
3.4.2	Instalación de los paquetes de Cluster	27
3.4.3	Habilitación del Cluster	29
3.4.4	Instalación del Quorum Vote	32
3.4.5	Verificación de la Instalación del Cluster	34

CAPITULO IV

CONFIGURACIÓN DE RECURSOS DE CLUSTER

4.1	Estructura de la base de datos	36
4.1.1	Instancia de la base de datos	36
4.1.2	Archivos de datos o data files	36
4.1.3	Archivos Archive	36
4.1.4	Listener	37
4.2	Creación de los sistemas de archivos de la base de datos	37
4.2.1	Creación de los metaset	37

4.2.2	Creación de los metadevices.....	39
4.2.3	Creación de los sistemas de archivos.....	40
4.3	Instalación de la base de datos.....	41
4.4	Habilitación de los recursos.....	41
4.4.1	Servicios de datos.....	41
4.4.2	Tipos de recursos.....	41
4.4.3	Recursos.....	42
4.4.4	Grupo de recursos.....	42
4.4.5	Servicio de datos a prueba de fallos o failover.....	42
4.4.6	Supervisores de errores.....	43
4.5	Registro de Tipo de Recursos.....	43
4.6	Creación de Recursos.....	43
4.7	Verificación del Estado del Cluster.....	46
4.8	Pruebas de Sun Cluster.....	48
4.8.1	Primera Prueba: Intercambio de Servicios entre nodos.....	48
4.8.2	Segunda Prueba: Apagado del nodo con los servicios.....	50
4.8.3	Tercera prueba: Perdida de la red pública de un nodo.....	51
4.8.4	Cuarta Prueba: Eliminar procesos de la base de datos	51
4.8.5	Quinta Prueba: Desconectar los Cables de Interconexión.....	52
	Conclusiones y Recomendaciones.....	54
	ANEXOS	
A.	Cronograma de actividades.....	56
B.	Glosario de Términos.....	58
C.	Información de la Configuración.....	60
	BIBLIOGRAFIA.....	65

INTRODUCCIÓN

Actualmente, con el Internet podemos encontrar muchos servicios que operan durante todo el día todos los días del año. Podemos nosotros tener acceso al Internet y revisar diversas páginas web a cualquier hora del día, usar el Messenger, realizar consultas por web, descargar archivos etc. En algunas ocasiones sin embargo nos encontramos con que el servicio no esta disponible por un breve periodo de tiempo. Nos ha pasado varias veces que la dirección de Internet nos devuelve el mensaje como: "la página no esta disponible por el momento", o algunas veces en una ventanilla nos dicen que la operación que deseábamos realizar no se puede efectuar en ese momento pues "no hay sistema".

Para entender porque sucede estos problemas, podríamos clasificarlos en dos tipos, podría ser un problema de la red desde el punto donde se intenta acceder al servicio hasta el lugar donde están instalados las aplicaciones, o podría ser un problema de indisponibilidad de la aplicación, ya sea porque falla del equipo o por un error de ejecución de algún modulo de la aplicación. La mayoría de veces, se sabe que se debe al segundo tipo de problema, pues son más comunes que el primero.

La estructura tradicional de una aplicación es que este instalado en un servidor con un buen procesador y memoria RAM, doble fuente de poder y que cuente con soporte del proveedor en caso de fallas del servidor o de la aplicación. Dependiendo del tipo de negocio, muchos servicios sin embargo requieren disponibles en todo momento, resultando insuficiente este esquema, una avería del servidor puede suponer varias horas de parada hasta que se vuelva a ponerse operativo, de igual forma una falla de la aplicación.

Para solucionar esto, se crean los cluster que usan 2 o más servidores y dispositivos de almacenamiento de información con una interconexión especial entre servidores y los dispositivos de almacenamiento denominada red de alta disponibilidad, cuya función es no depender de un solo servidor o un solo componente para que el servicio o aplicación este operativo. De tal manera que si falla un equipo, este puede ser reparado mientras las aplicaciones siguen funcionando en otros equipos.

El presente informe abarca de un tipo de cluster denominado Sun Cluster, instalado sobre el sistema operativo Unix Solaris en dos servidores.

Se entiende que hay conocimiento por lo menos básico de Unix así como del Protocolo TCP/IP Para el estudio del presente informe.

Los materiales empleados durante la realización del trabajo, como programas, fuentes bibliográficas, información del fabricante y parches de los productos, se basaron en los últimos disponibles durante la etapa de implementación, que se desarrolló en el mes de Marzo del 2009.

Con el fin de no hacer extenso el informe se han recortado las salidas de los procesos de Instalación que no se han considerado relevantes para el desarrollo del informe.

CAPITULO I

DESCRIPCIÓN DEL PROBLEMA Y MARCO TEORICO

1.1 Descripción del Problema

El Presente Informe describe el trabajo del siguiente escenario descrito a continuación.

Las Boticas Inkafarma dentro de su plan de disponibilidad de servicios, posee 250 equipos con base de datos pseudos-sincronizadas a nivel nacional, trabajando cada base de datos de forma independiente durante el día y sincronizándose durante la noche, en donde se transfiere a la base central los registros de compras hechas durante el día y se actualiza los precios de los diferentes productos, Así como también se realiza en horas de la noche el respaldo de las base de datos a la central

Esto les permite a los locales donde se encuentran las boticas, tener cierta autonomía del local principal por un día o más en casos extremos, si hubiera necesidad, trabajando sin conexión remota y sin afectar el servicio de atención al cliente.

Dentro de su plan estratégico de ventas hace poco Inkafarma implemento el servicio de Fidelización de Clientes, consistente en acumulación de inkapuntos por monto de compra. Al llegar a cierta cantidad de inkapuntos se le otorga vales para compra de productos.

Este servicio sin embargo, requería una sincronización real, puesto que el usuario podía comprar en diferentes puntos durante un día, y debería poder tener acceso a los inkapuntos inmediatamente para hacer uso de los vales. Además se entendía que este servicio debe estar disponible permanentemente las 24 horas por el mismo motivo, es decir que debería ser un servicio en "alta disponibilidad".

Este servicio llamado Fidelización, estaba instalado en el local principal en un servidor Sun Fire V250 con 1 Procesador UltraSparc III de 800Mhz y 4GB de Memoria RAM, que comenzaban a resultar insuficiente para la demanda del servicio.

En estas circunstancias El jefe de Operaciones de Inkafarma, decide contactarse con la Empresa Sonda del Peru S.A. para contratar los servicios de puesta en alta disponibilidad de la base de datos de Fidelización, que maneja los inkapuntos por cliente.

1.2 Objetivo del Trabajo

El objetivo del presente trabajo es instalar y configurar los programas necesarios para obtener la base de datos de Fidelización en cluster conformado por 2 servidores designados por Inkafarma, para que en caso falle algún componente crítico del servidor designado como primario o algún módulo de la base de datos en ese servidor, automáticamente el servicio se inicie en el segundo servidor.

1.3 Evaluación de los Equipos Disponibles

Después de una reunión informativa con Inkafarma, se tenía el siguiente Hardware disponible:

- Servidor Sun Fire T2000 – 2 discos Internos – 1 Procesador T1 – 16 GB RAM - 1 Tarjeta SCSI
- Servidor Sun Sparc T5220 – 2 discos Internos – 1 Procesador T2 – 8 GB RAM
- Storage Sun 3310 – 5 Discos 73GB – 1 Controladora - 2 Interfaces SCSI
- La Base de Datos que utilizaba el Administrador es Oracle 10g R2
- Inkafarma no contaba con presupuesto para adquirir otro Servidor u otro Storage tampoco para adquirir licencia de algún software de Alta Disponibilidad

Las características del Storage, le permiten ser confiable y resistente a fallas de disco o adaptador SCSI. Es ahí donde pudieran almacenarse la base de datos, debiendo ser conectado a un servidor para que haga el procesamiento de la información. Sin embargo no se podía asegurar la confiabilidad de los servidores, puesto si bien contaban con fuente de poder y discos redundantes, podían quedar no operativos en caso falla de Procesador o de Mainboard.

Sin embargo, el servicio debería ser capaz de poder usar ambos servidores de tal manera que si falla un servidor, el servicio siga funcionando con el otro servidor. Es decir necesita un servicio de “cluster” entre los servidores y debería ser compatible con la base de datos Oracle. Dado que no se podía adquirir por costos el software de Cluster de Oracle, se eligió “Sun Cluster” como solución.

1.4 Limitaciones del Trabajo

La instalación empieza desde cero, es decir configura los equipos desde su instalación física, considerando que no tienen sistemas operativos activos y culmina con las pruebas referidas a la operación del cluster con la base de datos. Sin embargo, hubo un paso

dentro de todo el proceso que debía hacerlo Inkafarma, en particular su administrador de base de datos, el cual era la instalación de la base de datos de acuerdo a sus metodologías propias de la empresa. Luego de instalada la base de datos y recogida de esta la configuración necesaria se continuo con la instalación del cluster

1.5 Elementos de un Centro de Datos

Ahora veamos conceptos importantes

Un Centro de Datos es una sala usada para almacenar un gran equipamiento de equipo Electrónico. En un centro de datos o centro de cómputo podemos encontrar servidores, storage, librerías de backup entre otros equipos. Todos estos equipos deben estar preparados para operar continuamente y el centro de cómputo les debe proveer los requisitos eléctricos, y ambientales para ello.

Los servidores son los elementos que realizan el procesamiento de la información, para lo cual están equipados de procesadores y memoria RAM.

Los storage son equipos que están compuestos por varias unidades de discos duros y almacenan la información. Los storage por si solos no pueden procesar la información que contienen, por lo que necesitan ser conectados de alguna manera a uno o mas servidores.

Las librerías de Backup permiten el respaldo de la información a unidades de cinta. Generalmente las librerías son controladas por un software de administración de backup.

1.6 Medios de transporte de datos en redes de alta disponibilidad

Como se había mencionado en la sección anterior, los storage deben ser conectados de alguna manera a los servidores con un cableado de datos diseñado para este propósito.

Existen varios tipos de Cableado de datos, siendo los mas utilizados la conexión SCSI y la de Fibra óptica.

1.6.1 Tecnologia SCSI

El protocolo SCSI (Small Computer System Interface) Es un método de acceso de discos físicamente conectados a servidores.. Actualmente con la tecnología Ultra-320 se tiene una velocidad de transferencia de 320 Mbps con un máximo de 16 dispositivos en una cadena SCSI.

Las Configuraciones con cables SCSI se definen con canales donde una serie de dispositivos SCSI se conectan a un host a través de una cadena o canal.

Algunas características de Las Configuraciones por Canal (Chanell), es que tienen un rendimiento aceptable, utiliza distancia cortas, no mayores a 10 metros, la identificación

de dispositivos SCSI es manual y la conexión con el cable permite solo un servidor en una cadena de dispositivos.

1.6.2 Fibra Óptica

Las Conexiones por Fibra Óptica permiten velocidades de hasta 4Gbps, y últimamente se esta perfeccionando la velocidad de 8Gbps, sin embargo las interfaces de Fibra son mucho mas caras que las SCSI.

La Topología de Interconexión por Fibra Óptica puede ser conectando directamente los servidores a los Storage como es el caso de los cables SCSI o también puede ser a través de switches de Fibra, especialmente diseñados para este propósito. En esta última configuración los switches reciben el nombre de Switches SAN, y el conjunto de interconexiones a través de estos switches es denominado Red SAN

Las propiedades de los switches SAN, como zonas, redundancia entre switches o routing, permiten un entorno ideal para el transporte de datos, sin embargo su estudio escapa del objetivo del presente informe.

La siguiente figura 1.1 muestra el esquema de una Red SAN en Comparación con una Red LAN.

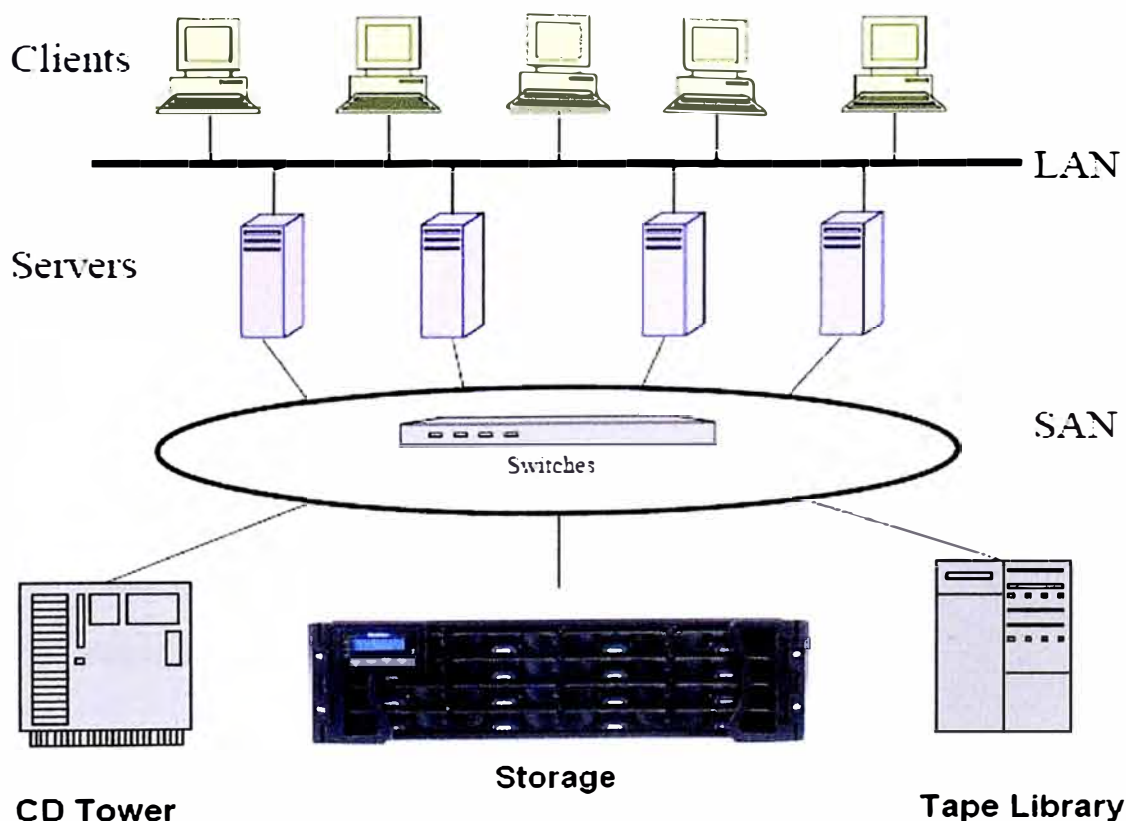


FIGURA 1.1 Arquitectura de una red SAN

1.7 RAID

Un problema bien conocido de pérdida de información, es la falla de un disco duro.

Los dispositivos diseñados para trabajar en alta disponibilidad tienen la capacidad de trabajar con discos virtuales en lugar de discos físicos, esto les permite no verse afectado en caso se produzca una falla de un disco físico.

Estos discos virtuales usan diferentes agrupaciones de Discos físicos denominadas RAID

A continuación los RAID más conocidos:

1.7.1 RAID 0

Los Discos se agrupan para formar un disco virtual con una capacidad que viene a ser la suma de los discos individuales

Existen 2 subtipos de RAID 0,

RAID 0 concatenado donde la información en el segundo disco empieza cuando se llena el primer disco, y así sucesivamente con el tercero hasta el ultimo disco.

RAID 0 striping La información es guardada en bloques usando todos los discos desde el comienzo. El primer bloque al primer disco, el segundo bloque al segundo discos, hasta el último disco, después se repite el ciclo, continuando en el primer disco. Tiene la ventaja de tener mayor velocidad de escritura pues usa varios discos a la vez. En la siguiente figura 1.2 La información a escribir es dividida en bloques A1, A2, A3, A4,A5 Y vemos la forma en que es colocada dentro del RAID

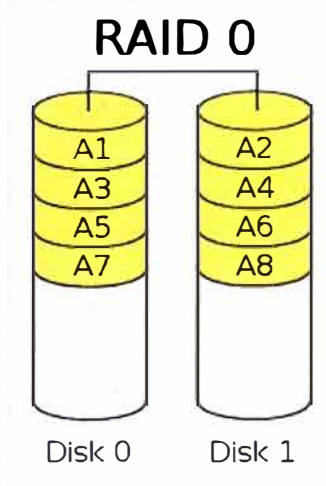


FIGURA 1.2 RAID 0

1.7.2 RAID 1

Utiliza 2 discos, la información es guardada en ambos discos a la vez, y se presenta como un disco virtual con una capacidad igual a la de cada disco que lo compone. Este RAID también tiene buena velocidad de lectura/escritura por usar 2 discos a la vez.

Además en caso falle un disco físico La información en el disco virtual no es afectado pues aun queda un disco físico. Este RAID también es conocido como mirror o espejo. En la figura 1.3 se observa la formación de un RAID 1.

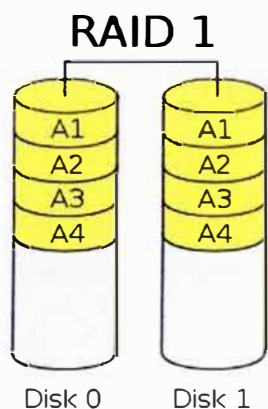


FIGURA 1.3 RAID 1

Hay combinaciones de ambos tipos de raid anteriores

El RAID 0 +1 Es un Raid 1 compuesto de 2 RAID 0

El Raid 1 +0 Es un Raid 0 de 2 o más RAID 1

Estos RAID son mostrados en la figura 1.4 Se muestra en la parte izquierda el RAID 0+1 y en la parte derecha el RAID 1+0

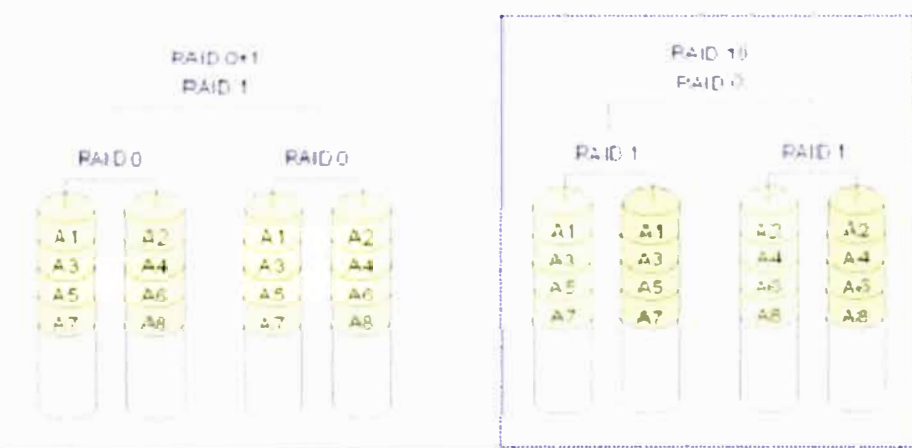


FIGURA 1.4 RAID 0+1 y 1+0

1.7.3 RAID 5

Esta compuesto de 3 o más discos en donde la información es grabada a través de los discos de tal manera que la capacidad equivalente a un disco es usado para paridad de la información y permite ser tolerante a la falla de un disco. La Capacidad del Discos virtual creado con RAID 5 es equivalente $(n-1)$ veces la argo capacidad de los discos físicos, en donde n es la cantidad de discos físicos. Este RAID quita rendimiento de

escritura, pues se pierde procesamiento al calcular la paridad. En la figura 1.5 Se aprecia la forma en que es colocada la información en un RAID 5. Los subíndices “p” indican la paridad de la respectiva fila.

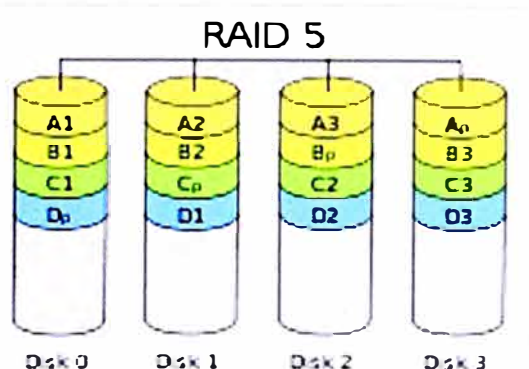


FIGURA 1.5 RAID 5

1.7.4 Disco Spare

En el Caso de RAID 1, 0+1, 1+0, 5 . Es posible configurar un disco que no pertenece al RAID como disco Hot-spare del RAID. La función del disco spare es reemplazar cualquier disco que falle del RAID, cuando esto ocurre, se procede a escribir en el disco spare la información que tenía el disco averiado, que es calculado a partir del contenido de los otros discos. Terminado El Proceso de Sincronización el disco spare se incorpora al RAID. Todo el proceso es automático, y durante el mismo, los datos dentro del RAID siguen estando disponibles en todo momento.

1.8 Protecciones de Alta Disponibilidad

1.8.1 Fuente Redundante

Al tener los servidores, storage y librerías 2 o más fuentes de poder redundantes, les permiten seguir operando a pesar de que falle una fuente de poder, o accidentalmente se desconecte un cable de poder.

1.8.2 Tarjetas RAID

En los Servidores hay 2 formas de crear los RAID de discos, el mas efectivo por rendimiento, es el RAID con una Tarjeta Interna RAID, de esta forma el sistema operativo solo ve los discos duros como RAID desde el inicio del arranque del sistema operativo. Este Método es también conocido como "RAID por Hardware". La tarjeta RAID es la encargada de gestionar el RAID en los discos duros.

1.8.3 Gestor de Volúmenes o Volume Manager

Un gestor de volúmenes permite la construcción de RAID con un programa que hace esta función también denominada "volume manager" del Sistema Operativo, dependiendo de este, el Volume Manager puede ser Veritas Volume Manager, Solaris Volume Manager, etc. Los RAID por Volume Manager son conocidos como "RAID por software". Cuando inicia el sistema operativo encapsula los discos físicos en discos virtuales, y trabaja con estos discos virtuales.

1.8.4 Multipath o Multiruta

Permite el acceso de un sistema operativo a un Storage a través de 2 caminos diferentes, es decir 2 cables diferentes, otorgando redundancia de cable y tarjeta HBA. Usado en conexiones de fibra óptica.

También puede utilizarse con los cables de red LAN para redundancia de interfaces de red.

1.8.5 Protección en Storage

Los storage, al ser agrupaciones de discos, se protegen usando controladoras capaces de manejar RAID, Discos spare, y además de tener fuente redundante. Hacen que estos equipos sean ideales para servicios en alta disponibilidad.

Los Servidores En cambio, tienen procesadores y memoria que en caso falle hacen quedar no operativo al equipo. Se ha mejorado esto, y en algunos servidores, se agrupan los procesadores y memorias en tarjetas, de tal manera que si falla un procesador o una memoria de una tarjeta, esta se deshabilita, sin embargo la falla de todas formas produce un reinicio del servidor. Además están otros componentes como la placa madre o alguna tarjeta que posiblemente provoquen un reinicio del servidor. Aunque estadísticamente este último tipo de fallas es bajo, hacen que los servidores no sean confiables al 100% y se ve la necesidad de disponer de más de uno, para asegurar el servicio.

1.9 Cluster

Un Cluster esta compuesto de 2 o más servidores que trabajan conjuntamente como un unico sistema y que están monitoreándose entre sí, con el fin de proporcionar aplicaciones a los usuarios de forma permanente, tal manera que si un servidor falla, el otro servidor asume el procesamiento de los servicios, lo cual le brinda como característica proporcionar alta disponibilidad.

Los cluster por el monitoreo que realizan pueden clasificarse en 2 tipos:

Cluster de Plataforma, que esta pendiente de alguna falla de hardware en los servidores.

Cluster de Aplicación, monitorea alguna falla de hardware y alguna falla de la aplicación.

Por el servicio de la aplicación que brindan, puede ser de 2 tipos:

Cluster Activo-Standby, constituido por 2 servidores, el servicio se ejecuta solo en un servidor y el otro no ejecuta ningún servicio de la aplicación hasta que el primer servidor falle.

El Cluster Activo-Activo, formado por 2 o más servidores, en este caso todos los miembros del cluster proporcionan el servicio y si una falla los servidores que quedan siguen proporcionando el servicio.

El Cluster Activo-Activo, tiene la ventaja de hacer un balanceo de carga, sin embargo esto es contraproducente cuando por ejemplo se tiene 2 servidores usando mas del 50 % de recursos, puesto que ante falla de uno de los equipos, el otro no será capaz de soportar toda la carga.

CAPITULO II

CARACTERISTICAS DE LOS PROGRAMAS UTILIZADOS

2.1 Sun Solaris

Solaris es un sistema operativo Unix de la compañía Sun Microsystems.

Sun Microsystems, Es una de la principales marcas de Fabricantes de Servidores, storage y librerías, junto con IBM, HP y Dell. Se caracteriza por disponer de muchos productos disponibles en su pagina web para descarga y uso, entre ellos se puede mencionar la base de datos MySQL, El servidor de aplicaciones Java GlassFish, el Sistema Operativo Solaris, entre otras cosas como OpenOffice y el Sun Cluster ya mencionado.

Sun Solaris es un sistema operativo, estable y de buen rendimiento y compatible con una gran variedad de productos. Además Solaris es el único sistema operativo que tiene versiones tanto un arquitecturas RISC, como en Entornos x86 con procesadores Intel o AMD.

La versión de Solaris sparc es un sistema operativo UNIX diseñado para trabajar en servidores Sun con procesadores sparc.

La versión de Solaris x86 puede operar en Servidores Sun x86 con procesadores Intel o AMD, y también en muchos servidores y Equipos de escritorio, y laptops. En la siguiente página se pueden encontrar los equipos certificados: www.sun.com/bigadmin/hcl

Existe una comunidad de Solaris así como en Linux que operan sobre Open Solaris, el cual es la versión de código abierto de Solaris.

Cabe mencionar que la descarga y uso de Solaris y Open Solaris es libre y no tiene costo alguno, desde la web de Sun, www.sun.com/download.

Actualmente estamos en la versión 10 de Solaris update 7 05/09 (Mayo 2009)

Solaris cuenta con un conjunto de aplicaciones incluidas en el sistema operativo como herramientas de office, navegador web, clientes de correo, manejador de volúmenes de disco, aplicaciones multimedia, etc. A continuación algunas características importantes de Solaris 10 que ayudaran a entender el trabajo realizado

2.1.1 Particionamiento de Discos

La nomenclatura que Solaris usa para los dispositivos de discos es de la forma `c##t##d##` donde los caracteres “#” representan la numeración del disco. La unidad de CD o DVD también tiene su representación `c##t##d##`.

Solaris divide todos los discos duros en 8 particiones del `s0` al `s7`, al juntarse con la nomenclatura de los discos, las particiones tienen la forma `c##t##d##s#`.

2.1.2 Archivo de inicio de Sistemas de archivos o file systems

El Archivo `/etc/vfstab` contiene todas las particiones de disco que serán montadas en el arranque del equipo, y en que directorio será montado.

Como ejemplo, puede indicarse que la partición `c1t0d0s5` se monte en el directorio `/u01`.

Indica que la partición 5 del disco `c1t0d0` será montado en el directorio `/u01`.

2.1.3 Archivo de Nombres

El archivo `/etc/hosts` que contiene 2 columnas, que representan las relaciones nombre vs direcciones IP que resolverá el equipo.

2.2 Sun Cluster

Sun Cluster al igual que Solaris es un software de Cluster de la Compañía Sun Microsystems, su descarga de Internet e instalación no están sujetos a costo alguno.

Sun Cluster está diseñado para operar en el Sistema Operativo Solaris operando a 64 bits. Sun Cluster supervisa los fallos de aplicaciones, rutas de disco y protocolos de Internet (IP) de ruta múltiple.

Sun Cluster posee agentes para diferentes tipos de Aplicaciones como NFS, Java, bases de datos como Sybase y Oracle, servicios web como Apache, servicios de monitoreo como Patrol, entre otros productos.

La última versión de Sun Cluster disponible a la fecha es la versión Sun Cluster 3.2 update 2 1/09

2.3 Elementos de Sun Cluster

Sun Cluster está compuesto de los siguientes equipos:

2.3.1 Nodos

Son los servidores que tiene el sistema operativo Solaris y son los que ejecutan el software de cluster así como los agentes del cluster para aplicaciones.

Se necesita un mínimo de 2 nodos para el funcionamiento del cluster.

Todos los nodos del clúster reciben información de cuándo un nodo se une al clúster o deja éste, conocen también los recursos que se están ejecutando tanto localmente como en los otros nodos del clúster.

No es un requisito que todos los nodos del cluster tengan igual capacidad de procesamiento, memoria o velocidad de entrada/Salida. Sin embargo deberían ser similares, puesto cada nodo debería ser capaz de soportar las aplicaciones.

2.3.2 Storage

Es un almacenamiento de discos externos que esta conectado a los nodos por cable SCSI o fibra óptica.

2.3.3 Consola de Administración

Estación de Trabajo que no necesariamente tiene que ser de uso exclusivo, desde el cual se administra cada uno de los nodos y el storage. Debería tener acceso a las interfaces de administración de cada uno de lo servidores y storage.

2.3.4 Red de Administración:

Es la red que conecta las interfaces de administración de los nodos y storage con la consola de administración, es altamente recomendable que sea una red o subred de rango diferente al de producción.

2.3.5 Red de Almacenamiento:

Es la Red que comunica los Servidores al Storage. Puede ser directamente a traves de cables SCSI o de Fibra Optica, o una Red SAN

2.3.6 Red de Interconexión:

Es la red que interconecta los nodos del cluster para las comunicaciones privadas de cluster y pueden ser conformados por switches o VLANs exclusivos o en el caso de 2 nodos, o cables conectados directamente entre los nodos. Se necesita 2 o mas

interfaces de red por cada nodo para redundancia. Las interconexiones redundantes permiten que continúe el funcionamiento en las interconexiones que queden, mientras los administradores de los sistemas aíslan los fallos y restablecen la comunicación. Sun Cluster detecta, repara y reinicia automáticamente la comunicación en una interconexión restablecida.

2.3.7 Red de Producción:

Es la Red que conecta los nodos a la red LAN, se recomienda usar 2 cables para redundancia por IPMP desde cada nodo.

En la siguiente figura 2.1 se puede observar los elementos de Sun Cluster para 2 nodos y un storage.

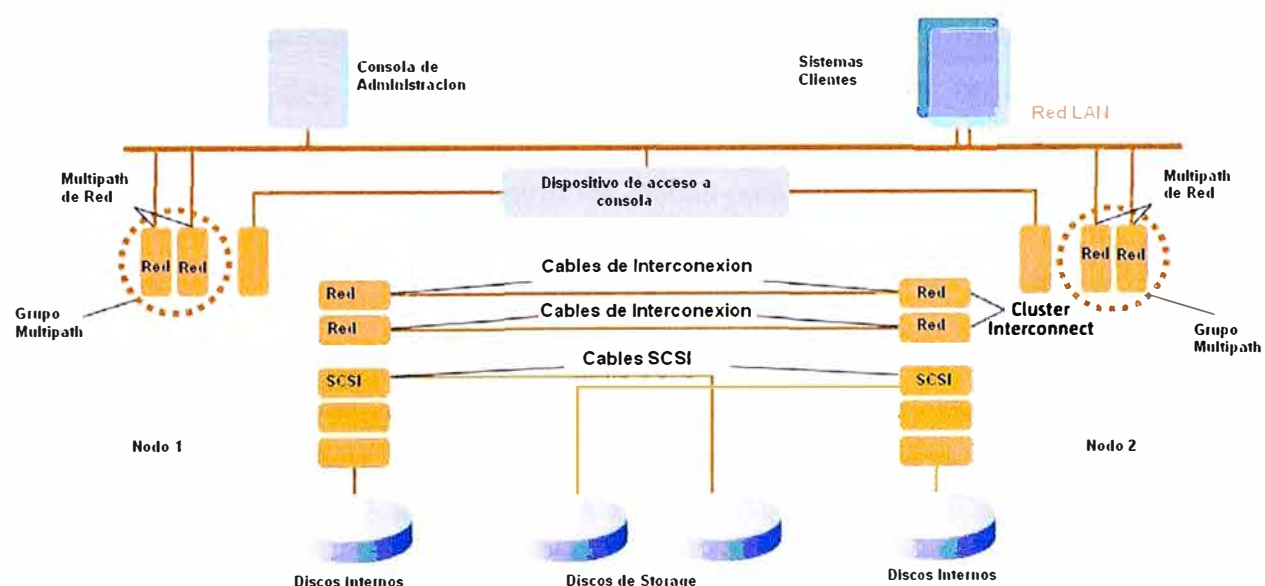


FIGURA 2.1 Arquitectura de Sun Cluster con 2 nodos

2.3.8 Quórum

En el caso de 2 nodos y un storage en cluster, se presentan 2 problemas que ponen en peligro la integridad de los datos.

El Primero es el data fencing o esquizofrenia, se presenta cuando se pierde la interconexión entre los nodos, el cluster se divide en 2 subcluster, cada uno de los cuales ignora la presencia del otro y puede crear conflictos en los recursos compartidos, como el acceso a discos y las direcciones de red.

El segundo problema es conocido como Amnesia, se presenta cuando hay un nodo apagado y se hacen cambios en la configuración en el nodo activo, luego si este se baja, y se enciende el que estaba apagado intentara cargar la configuración antigua pues no

esta al tanto de los cambios realizados, pudiendo manipular el cluster de manera incorrecta los recursos.

Para resolver los problemas anteriormente mencionados, Sun Cluster otorga votos a cada uno de los servidores y un voto extra a un dispositivo de quórum en el storage, que es asignado a uno de los nodos, de tal manera que el nodo que tenga 2 votos de los 3 disponibles, obtiene el quórum. En caso de pérdida de conexión entre nodos, el nodo que tenga el quórum, es el único que tiene acceso a los recursos compartidos, resolviéndose el problema de la esquizofrenia. De igual manera cuando se encienden los nodos se verifica que el nodo con la configuración mas reciente, es decir el ultimo nodo que tuvo el quórum, este disponible para evitar el problema de la Amnesia.

2.4 Solaris Volume Manager

Solaris Volume Manager es el gestor de volúmenes que viene incluido en la instalación de Solaris. SVM Tiene múltiples usos, aquí solo se explicara los necesarios para la Sun cluster, los cuales son los disk set o metaset y metadevices, los cuales a su vez son usado por Sun Cluster para administrar los discos compartidos.

2.4.1 Metadb

Para poder usar Solaris Volume Manager, en adelante abreviado SVM, primero se debe crear una pequeña base de datos interna que tendrá la configuración realizada por SVM. Para crearlo basta usar el comando "metadb -a" seguido de la partición en su nomenclatura `c###d###s#`. Esta partición se recomienda que sea de 50MB.

2.4.2 Metaset o Disk set

En el caso de discos compartidos, si los 2 nodos quisieran escribir simultáneamente crearían un conflicto y la información podría ser sobrescrita. Para evitar esto, SVM usa metaset en donde solo uno de los nodos es propietario de los recursos de disco que pertenecen a dicho metaset. Para crear el metaset, se utiliza el comando:

```
metaset -s "nombre del metaset" -a -h seguido de los nodos que pertenecerán al metaset.
```

Posteriormente se adicionan los discos compartidos al metaset.

Los metadevices son particiones logicas de los discos de un Metaset.

2.5 Dispositivos de Sun Cluster

Hay un sistema de archivos o file System llamado Global devices que es accesible en todos los nodos, cuya función es almacenar la configuración de Sun Cluster así como los

dispositivos que son accesibles por los nodos. Sun Cluster consigue que todos los dispositivos de un cluster sean accesibles por todos los nodos. Este tipo de acceso se denomina acceso global y se gestiona asignando ID a los dispositivos, denominados DID.

Los dispositivos pueden ser discos, CD-ROM o unidades de cinta. No obstante solo los discos de un storage compartido pueden ser considerados dispositivos de alta disponibilidad y son vistos desde cualquiera de los miembros del cluster con el mismo DID, aun cuando tengan diferente nomenclatura `c###d##` en los sistemas operativos de cada nodo.

De esta forma cuando un nodo falla al intentar acceder a un dispositivo global compartido, Sun cluster conmuta a otra ruta del dispositivo fácilmente puesto que el DID del dispositivo global es el mismo en cualquier nodo.

La configuración de los DID se guarda en el file system Globaldevices que debe tener un tamaño no menor de 500MB y debe estar disponible en todos los nodos miembros del cluster.

Se puede ver la relación entre la nomenclatura de dispositivos del sistema operativo y los dispositivos de cluster, es decir la relación `c###d##` y DID, con el comando `scdidadm -L`

CAPITULO III

INSTALACION DE SISTEMA OPERATIVO Y SUN CLUSTER

3.1 Componentes Utilizados

Se va a describir la instalación del sistema operativo, la configuración de discos del Storage y La instalación del software de cluster, mas no será descrita la configuración, que será detallada en el siguiente capítulo.

Recordando los elementos disponibles y la arquitectura necesaria para Sun Cluster como se muestra en la figura 2.1, Los elementos del Cluster fueron los siguientes:

Sun Sparc Enterprise T5220 - 1 Procesador Sparc T2 - 8 core, 8GB Memoria RAM expandible hasta 128GB, 2 Discos Duros internos de 73GB, 4 interfaces de red gigabit ethernet, 2 tarjetas de Fibra Canal, 1 Tarjeta con 2 puertos SCSI, 1 Tarjeta RAID para discos internos, 1 Tarjeta de administración ALOM. Soporta como sistemas operativos Solaris sparc, y Ubuntu.

Sun Fire T2000 - 1 Procesador Sparc T1 - 8 core, 16GB Memoria RAM expandible hasta 64GB, 2 Discos Duros internos de 73GB, 4 interfaces de red gigabit ethernet, 2 tarjetas de Fibra Canal, 1 Tarjeta con 2 puertos a SCSI, 1 Tarjeta RAID para discos internos, 1 Tarjeta de administración ILOM/ALOM que soporta RAID 0 y 1. Soporta como sistemas operativos Solaris sparc, Gentoo y Ubuntu.

Sun Storedge 3310 – 6 Discos de 73GB 15000 RPM, con capacidad de RAID y conexión a 2 host, en este caso los servidores anteriores con cables SCSI.

La red pública se habilitó conectando las 2 primeras interfaces de red de cada nodo a la red LAN de inkafarma.

La red de Interconexión se conformo con 2 cables UTP cruzados usando las 2 últimas interfaces de cada uno de los nodos.

La red de Administración se habilitó conectando las interfaces de administración de los servidores y el storage a la red LAN

3.2 Instalación de sistema Operativo en los Nodos

Aprovechando las tarjetas RAID internas de los servidores, se configura un RAID 1 entre los 2 discos internos. Así cuando empieza la instalación del Sistema operativo solo detecta un disco, que ya ha sido previamente encapsulado en un RAID1 por la tarjeta RAID.

De acuerdo a las mejores prácticas se instala la última versión estable de Solaris 10, el cual viene a ser Solaris sparc 10 10/08, Como los servidores no tienen interfaz grafica la instalación debe realizarse a través de la consola obtenida a través de la interfaz de administración.

Se había determinado los nombres que tendrían los equipos, el T5220 seria llamado inkaserver6 y el T2000 seria llamado inkaserver7.

Para mayor rapidez, se opta por hacer una instalación por la red. Usando un método conocido como Jumpstart. En una laptop con Solaris 10 se instala el Server colocando el DVD de Solaris 10 10/08 en la laptop y copiando la imagen de instalación a la ruta /export/install/sol10-1008

```
root@sopsun4 # /cdrom/cdrom0/Solaris_10/Tools/setup_install_server
/export/install/sol10-1008
Verifying target directory...
Calculating the required disk space for the Solaris_10 product
Calculating space required for the installation boot image
Copying the CD image to disk...
Copying Install Boot Image hierarchy...
Install Server setup complete
root@sopsun4 #
```

En el archivo /Jumpstart se almacenara los archivos de configuración que tendrán los nodos. Se edita los archivos profile dentro del directorio /Jumpstart

Inkaserver6 : /Jumpstart/profile-sc1

Inkaserver7 :/Jumpstart/profile-sc2

En estos archivos se define la instalación full con todos los paquetes incluidos, colocando el parámetro cluster en SUNWCXall y también definimos el particionamiento del disco. Para el Particionamiento del disco utilizan 6 sistemas de archivos o filesystems como sigue, las líneas que comienzan con filesys indican file systems la segunda columna es la partición, la tercera el tamaño en MB y la cuarta el directorio donde se montaran los file systems:

```
root@sopsun4 # more /Jumpstart/profile-sc1
install_type  initial_install
system_type   standalone
```

```

partitioning  explicit
cluster      SUNWCXall
filesystems  c1t0d0s0    11200  /
filesystems  c1t0d0s1    16100  swap
filesystems  c1t0d0s3    16100  /var
filesystems  c1t0d0s4    10000  /u01
filesystems  c1t0d0s5    10000  /u02
filesystems  c1t0d0s5    760    /globaldevices

```

De manera similar se edita el archivo / Jumpstart/profile-sc2. Luego se editan los archivos sysidcfg que tendrán las direcciones IP de los equipos. Se edita de la siguiente manera

/Jumpstart/nodo1/sysidcfg:

```
inkaserver6: 192.168.0.45
```

/Jumpstart/nodo2/sysidcfg:

```
inkaserver7 : 192.168.0.51
```

```

root@sopsun4 # more /Jumpstart/nodo1/sysidcfg
timezone=America/Lima
terminal=vt100
system_locale=en_US.ISO8859-1
timeserver=localhost
network_interface=primary {ip_address=192.168.0.45 protocol_ipv6=no
netmask=255.255.255.0 default_route=192.168.0.1}
root_password=""
name_service=NONE
security_policy=NONE
nfs4_domain=dynamic
root@sopsun4 #

```

De igual manera con el archivo /Jumpstart/nodo2/sysidcfg. Luego se procede a relacionar las direcciones MAC de los nodos con sus respectivos hostname.

```

root@sopsun4 # more /etc/ethers
0:3:ba:24:6a:b5 inkaserver6
0:3:ba:44:ff:56 inkaserver7

```

En el archivo /etc/hosts se relaciona las direcciones IP con los nombres de los nodos.

```

root@sopsun4 # more /etc/hosts | grep inkaserver
192.168.0.45 inkaserver6
192.168.0.51 inkaserver7

```

En el archivo /Jumpstart/rules se asocia el profile con cada nodo

```
root@sopsun4 # more /Jumpstart/rules
hostname    inkaserver6 - profile-sc1 -
hostname    inkaserver7 - profile-sc2 -
```

Finalmente se adiciona los nodos clientes a la configuración Jumpstart asociando las rutas de la imagen de instalación, archivo sysidcfg y tipo de plataforma.

```
root@sopsun4 #cd /export/install/sol10-1008/Tools
root@sopsun4 # ./add_install_client -c sopsun4:/Jumpstart -p sopsun4:/Jumpstart
nodosun1 sun4v
Adding "share -F nfs -o ro,anon=0 /export/install/sol10-1008" to /etc/dfs/dfstab
enabling network/rarp service
enabling network/rpc/bootparams service
updating /etc/bootparams
copying boot file to /tftpboot/inetboot.SUN4V.Solaris_10-2
root@sopsun4 #
root@sopsun4 # ./add_install_client -c sopsun4:/Jumpstart -p sopsun4:/Jumpstart
nodosun1 sun4v
enabling network/rarp service
enabling network/rpc/bootparams service
updating /etc/bootparams
copying boot file to /tftpboot/inetboot.SUN4V.Solaris_10-2
root@sopsun4 #
```

Luego en el cliente se indica iniciar por la red y comienza la instalación automáticamente del Sistema Operativo, a continuación la salida obtenida de la consola de administración en del inkaserver6.

```
ok>boot net - install
Resetting ...
```

```
Sun Fire T2000 (UltraSPARC-T1), No Keyboard
Copyright 2007 Sun Microsystems, Inc. All rights reserved.
OpenBoot 4.17.1, 8192 MB memory installed, Serial #52718261.
Ethernet address 0:3:ba:24:6a:b5, Host ID: 83246ab5.
```

```
Rebooting with command: boot net - install
Boot device: /pci@1f,0/network@c,1 File and args: - install
Requesting Internet Address for 0:3:ba:24:6a:b5
SunOS Release 5.10 Version Generic_120011-14 64-bit
Copyright 1983-2007 Sun Microsystems, Inc. All rights reserved.
Use is subject to license terms.
SUNW,e1000g0 : 100 Mbps full duplex link up
whoami: no domain name
Configuring devices.
Using RPC Bootparams for network configuration information.
```

(Salida Cortada)

```
Attempting to configure interface e100g0...
Configured interface e100g0
Setting up Java. Please wait...
Extracting windowing system. Please wait...
Beginning system identification...
Searching for configuration file(s)...
Using sysid configuration file sopsun4:/Jumpstart/nodo1/sysidcfg
Search complete.
Discovering additional network configuration...
Completing system identification...
Starting remote procedure call (RPC) services: done.
System identification complete.
Starting Solaris installation program...
Searching for JumpStart directory...
Using rules.ok from sopsun4:/Jumpstart.
Checking rules.ok file...
Using profile: profile-sc1
Executing JumpStart preinstall phase...
Searching for SolStart directory...
Checking rules.ok file...
Using begin script: install_begin
Using finish script: patch_finish
Executing SolStart preinstall phase...
Executing begin script "install_begin"...
Begin script install_begin execution completed.
```

Processing profile

- Selecting cluster (SUNWCXall)
- Selecting all disks
- Configuring boot device

Verifying disk configuration

- WARNING: Unused disk space (c1t0d0)

Verifying space allocation

- Total software size: 7305.75 Mbytes

Preparing system for Solaris install

Configuring disk (c1t0d0)

- Creating Solaris disk label (VTOC)

Creating and checking UFS file systems

- Creating / (c1t0d0s0)
- Creating /var (c1t0d0s3)
- Creating /u01 (c0t0d0s4)
- Creating /u02 (c0t0d0s5)
- Creating /globaldevices (c0t0d0s6)

Beginning Solaris software installation

Starting software installation

```

SUNWocfd.....done. 3805.41 Mbytes remaining.
SUNWlucfg.....done. 3805.33 Mbytes remaining.
SUNWcsu.....done. 3790.60 Mbytes remaining.
SUNWcsr.....done. 3786.41 Mbytes remaining.
SUNWcsl.....done. 3772.36 Mbytes remaining.
SUNWcnetr.....done. 3772.30 Mbytes remaining.
SUNWjdmk-base.....done. 3771.04 Mbytes remaining.
TSBWvplu.u.....done. 3771.01 Mbytes remaining.
TSBWvplr.u.....done. 3770.97 Mbytes remaining.
SUNWkvm.u.....done. 3768.69 Mbytes remaining.
SUNWcar.u.....done. 3768.09 Mbytes remaining.
FJSVhea.....done. 3767.46 Mbytes remaining.
SUNWcakr.u.....done. 3744.95 Mbytes remaining.
FJSVvplu.u.....done. 3744.92 Mbytes remaining.
FJSVvplr.u.....done. 3744.90 Mbytes remaining.
SUNWckr.....done. 3732.91 Mbytes remaining.

```

(Salida cortada)

```

SUNWter.....done. 1.00 Mbytes remaining.

```

Completed software installation

Solaris 10 software installation succeeded

Customizing system files

- Mount points table (/etc/vfstab)
- Unselected disk mount points (/var/sadm/system/data/vfstab.unselected)
- Network host addresses (/etc/hosts)
- Environment variables (/etc/default/init)

Cleaning devices

Customizing system devices

- Physical devices (/devices)
- Logical devices (/dev)

Installing boot information

- Installing boot blocks (c1t0d0s0)
- Installing boot blocks (/dev/rdisk/c1t0d0s0)

Installation log location

- /a/var/sadm/system/logs/install_log (before reboot)
- /var/sadm/system/logs/install_log (after reboot)

Installation complete

De esta forma concluye la instalación y el equipo ya es accesible por telnet o ssh. Una observación importante es que las 4 interfaces de red han sido nombradas como e1000g0, e1000g1, e1000g2, e1000g3.

El acceso al equipo se realiza por medio de una conexión telnet o ssh.

3.3 Configuración del Storage

El storage 3310 cuenta con una interfaz de administración desde el cual se puede configurar el storage con un menú propio del equipo.

Se sabe que el storage tiene 6 discos de 73GB 15000 RPM, con capacidad de hacer RAID y usar discos spare.

Los requerimientos del administrador de la base de datos son los siguientes:

- 1 Partición de 20GB Para Binarios de la Instancia de la Base de Datos
- 2 Particiones de 90GB para los data Files de la Base de Datos
- 1 Partición de mas de 50GB Para el Archive de la Base de Datos
- Además se pretende maximizar el uso de discos sin perder la protección contra falla de discos.

Con estos requerimientos se determina que la mejor opción de RAID es un RAID 5 con 5 discos y colocar el sexto disco como disco spare, se descarta el RAID 0+1 ya que se pierde efectivamente la mitad de los discos y no alcanza a cumplir los requerimientos de tamaño. Tampoco es conveniente un RAID 0 pues no tendría protección contra falla de discos. El disco spare permitirá reemplazar un disco averiado para no afectar el RAID 5. A Continuación en las figuras 3.1 se muestra la pantalla principal del configurador del storage.

Dentro del menú, primero se crea una agrupación de discos duros del tipo RAID5 con los 5 Discos del storage se tiene 270GB y se configura el sexto disco como Global Spare de este RAID 5.

En la figura 3.2 se muestra el RAID 5 con 5 discos en una agrupación denominada Logical Drive o disco lógico.

Luego, a partir del logical drive recién creado se divide en particiones del RAID 5 previamente creado, de la siguiente manera

- 1° 20GB
- 2° 90GB
- 3° 90GB
- 4° 72GB
- 5° 50MB

Finalmente las 5 particiones son presentadas a los canales 1 y 3, donde estan conectados los servidores inkaserver6 e inkaserver7 como se muestra en la figura 3.3. Cada partición presentada a un servidor es detectada como un dispositivo de disco. De tal manera que con la configuración realizada aparecen 5 nuevos discos en el sistema operativo de Solaris.

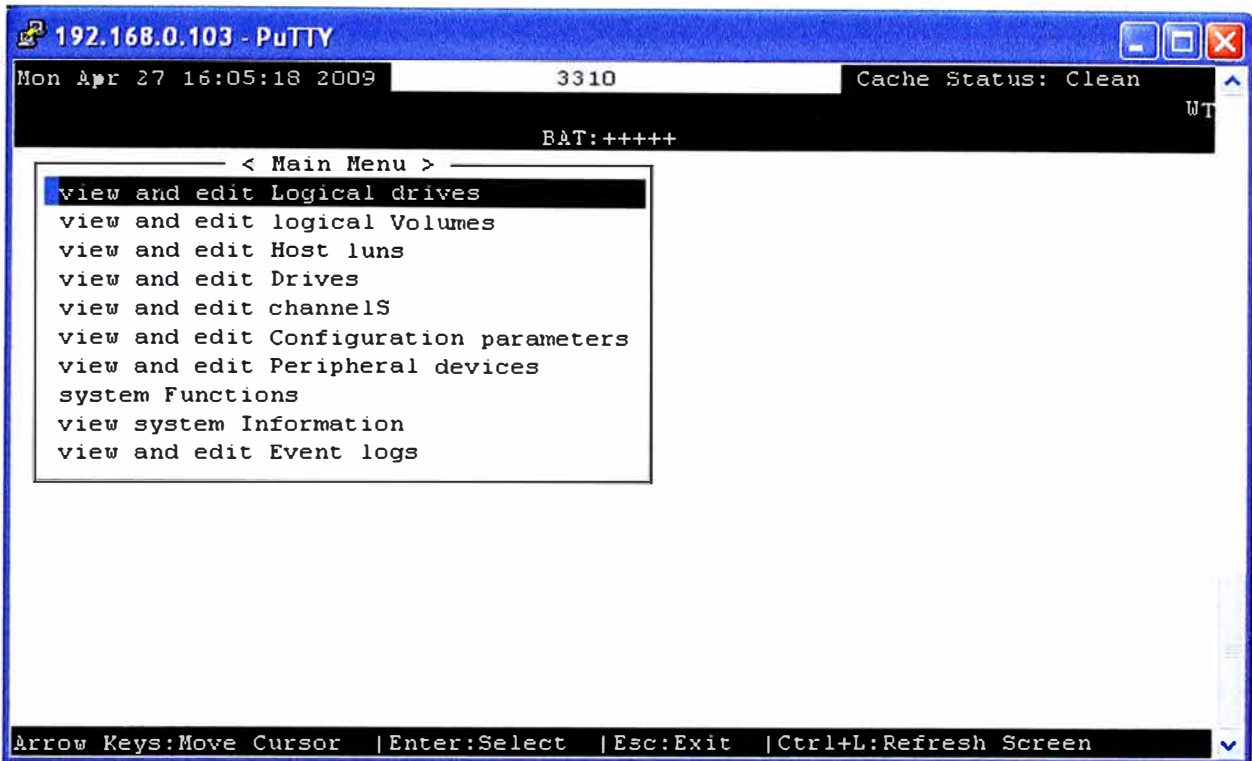


FIGURA 3.1 Pantalla Principal

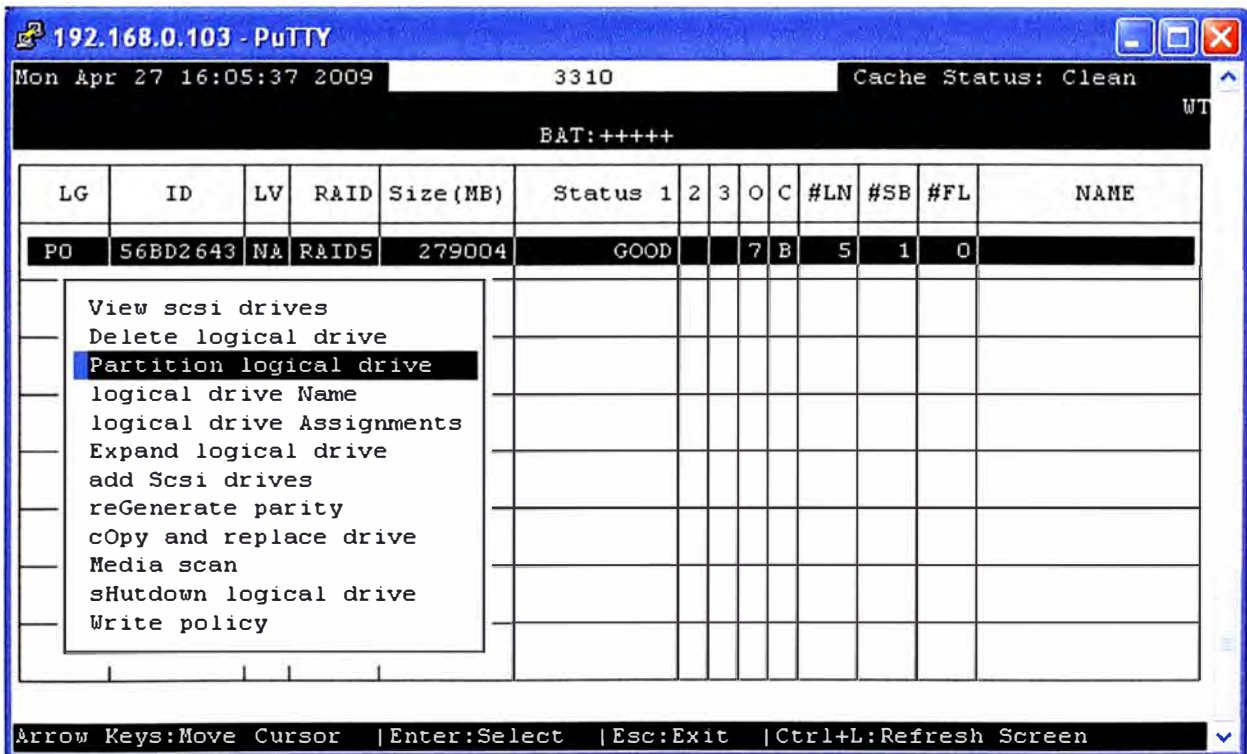


FIGURA 3.2 El Logical Drive 0 configurado como RAID 5

192.168.0.103 - PuTTY
 Mon Apr 27 16:06:13 2009 3310 Cache Status: Clean
 BAT:+++++

< Main Menu >
 view and edit Logical drives
 view and edit logical Volumes
 view and edit Host luns
 v
 v CHL 1 ID 0 (Primary Controll
 v CHL 3 ID 1 (Secondary Contro
 v CHL 3 ID 3 (Primary Controll
 s
 view system Information
 view and edit Event logs

LUN	LV/LD	DRV	Partition	Size (MB)	RAID
0	LD	0	0	20480	RAID5
1	LD	0	1	92160	RAID5
2	LD	0	2	92160	RAID5
3	LD	0	3	74150	RAID5
4	LD	0	4	54	RAID5
5					
6					
7					

Arrow Keys: Move Cursor | Enter: Select | Esc: Exit | Ctrl+L: Refresh Screen

FIGURA 3.3 Las 5 particiones del Logical Drive 0

3.4 Instalación de Software Sun Cluster

Antes de realizar la instalación de Sun Cluster se deben hacer algunos ajustes en el sistema operativo.

3.4.1 Permitir el acceso directo de un nodo a otro

Se habilita el acceso mediante ssh por root en ambos nodos, y luego se reinicia el servicio ssh.

```
inkaserver6#grep Root /etc/ssh/sshd_config
PermitRootLogin yes
inkaserver6#svcadm restart ssh
```

Se edita el archivo `/.rhost` agregando en cada uno de los nodos el nombre del otro.

```
inkaserver6#more /.rhost
inkaserver7
inkaserver6#
```

Con las configuraciones anteriores ya se pueden ejecutar configuraciones remotas en un nodo desde el otro. Esto es necesario para que la instalación del cluster en un nodo se efectúe también en el otro nodo.

La Instalación solo del Software de Cluster, sin administrar ningún recurso aun, se puede dividir en 2 etapas, la instalación de paquetes y la configuración del cluster.

3.4.2 Instalación de los paquetes de Cluster

La instalación de paquetes de cluster o Sun Cluster Framework Software, debe ser realizado en cada uno de los nodos.

Se descarga la última versión de Sun Cluster disponible dentro de www.sun.com/download. Sun Cluster 3.2 01/09 update 2 para Solaris Sparc. Se coloca el instalador en cada uno de los nodos. Para obtener una consola grafica, se utiliza un emulador grafico desde una laptop con Windows XP, en este caso se usa el Hummingbird Exceed, en donde se selecciona la ip 192.168.0.45 por ejemplo que corresponde a inkaserver6 y se procede a ejecutar el instalador, el cual usa una interfaz de Java.

Aparece una ventana con el mensaje de bienvenida, después el acuerdo de licencia y después los componentes a instalar, como se muestra en las figuras 3.4 y 3.5, se selecciona aparte del Sun Cluster core, el Sun Cluster HA for Oracle, también se selecciona otros componentes de Oracle que pudieran utilizarse en el futuro.

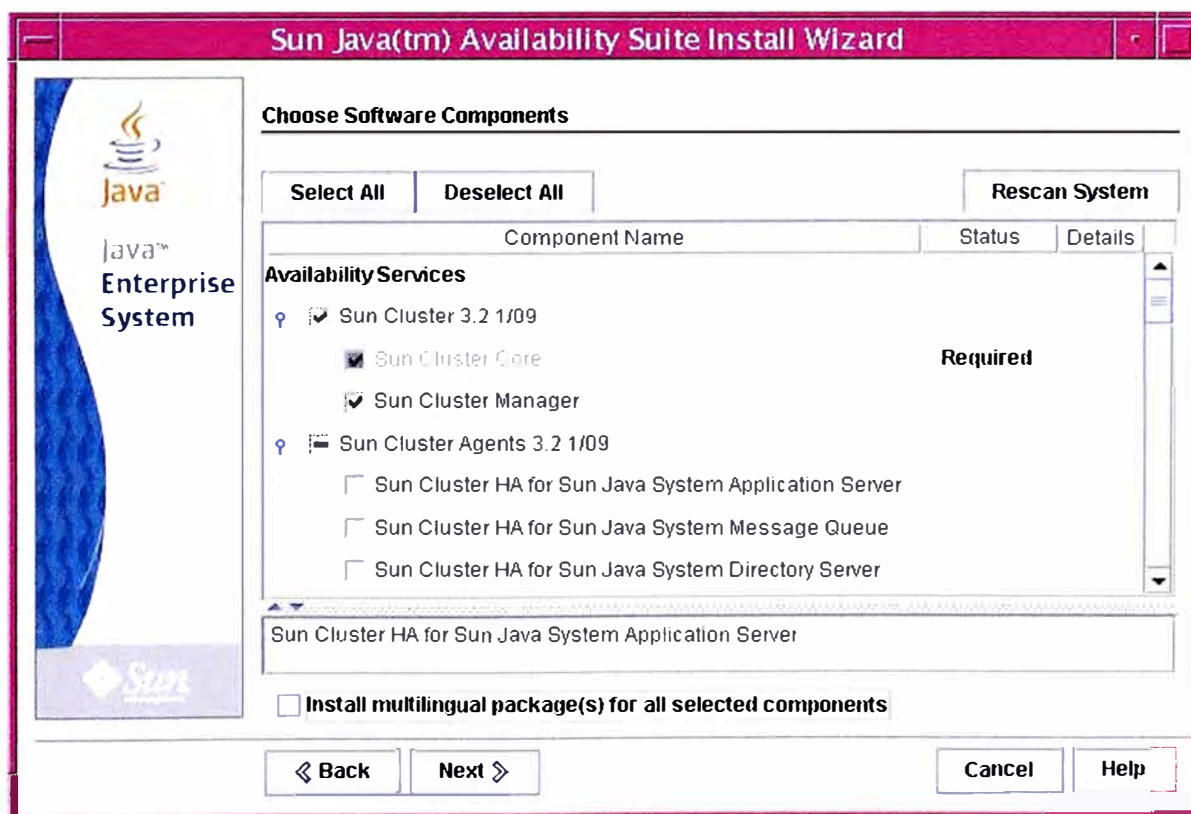


FIGURA 3.4 Componentes de Sun Cluster a instalas – primera parte

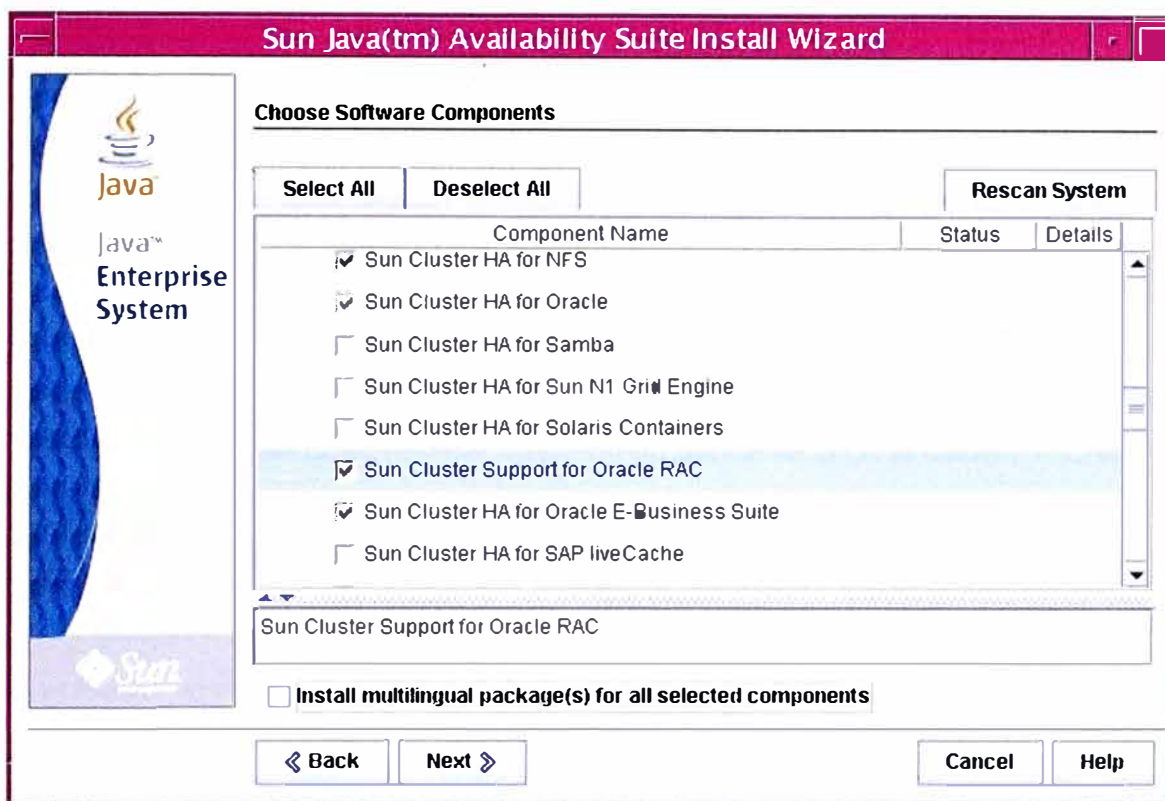


FIGURA 3.5 Componentes de Sun Cluster a Instalar – Segunda Parte

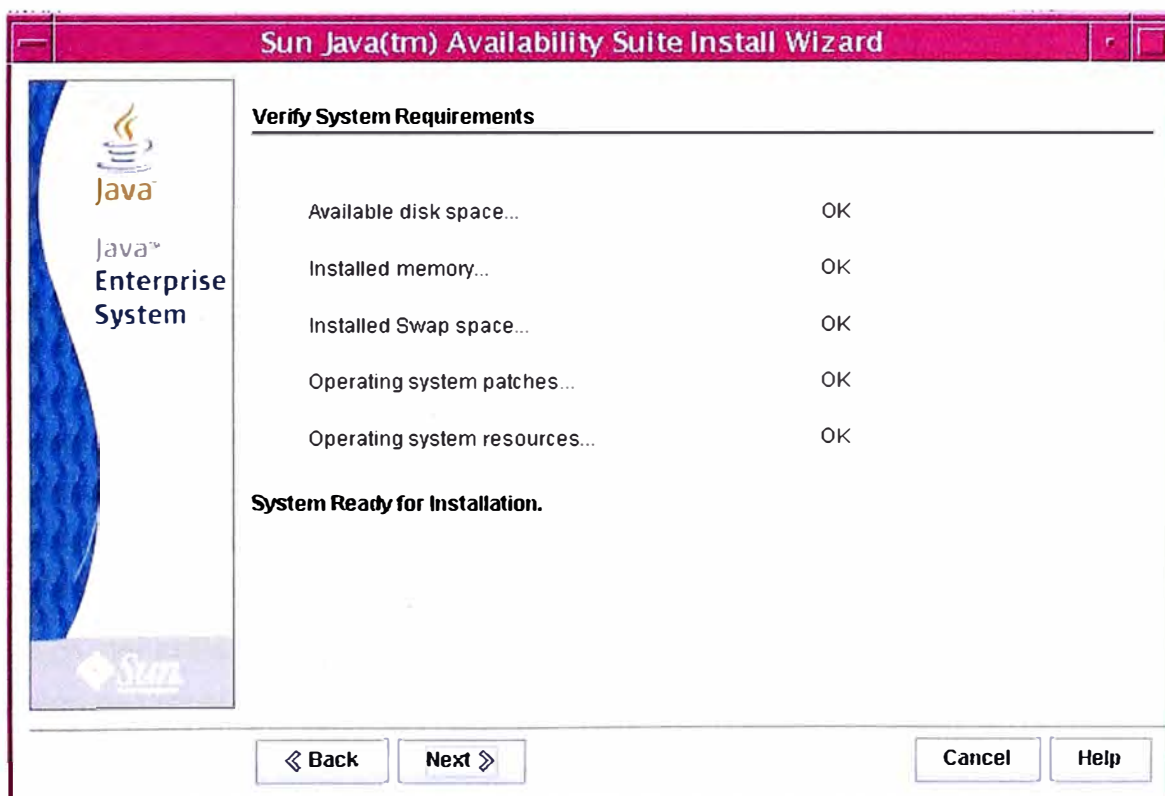


FIGURA 3.6 Validación de Requerimientos de Sistema operativo

Se realiza una validación de requisitos del sistema operativo como se muestra en la figura 3.6 y luego Comienza la instalación de los paquetes de Cluster.

3.4.3 Habilitación del Cluster

Después de Haber instalado los paquetes, comienza la habilitación del cluster ejecutando el comando `/usr/lib/cluster/bin/scinstall`.

Se ejecuta solo en un nodo, en este caso se ejecuta desde `inkaserver6`. Después de ejecutar `scinstall`, se nos presenta un menú en donde se elige la opción 1 "Create a new cluster", como se muestra en la figura 3.7.

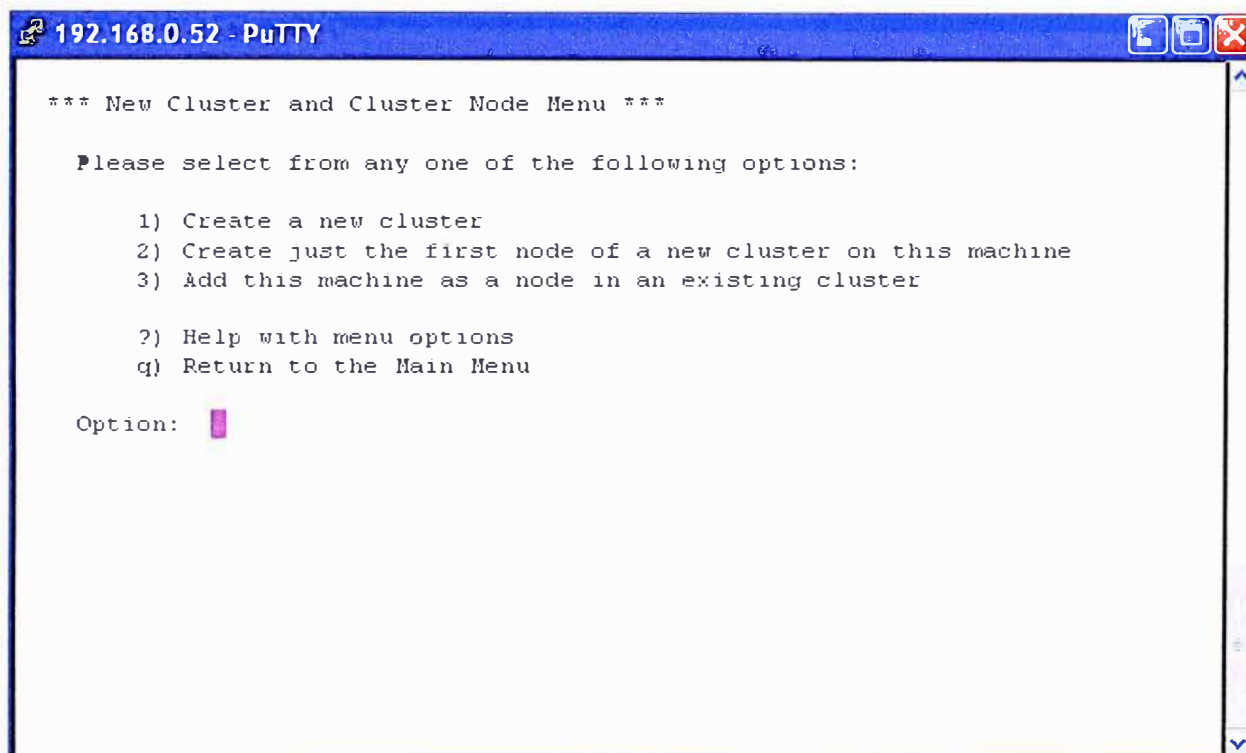


FIGURA 3.7 Pantalla principal de `scinstall`

Luego de advertirnos que los paquetes de cluster deben estar instalados en todos los nodos, se tiene que seleccionar el tipo de instalación que se desea realizar, típica o personalizada, en la configuración típica se selecciona automáticamente ips del rango 171.16.0.0 con mascara 255.255.255.240 usa 2 adaptadores de red para la interconexión de cluster el file system `/globaldevices` para almacenar la configuración de cluster. Como estamos preparados para esto, seleccionamos `typical`.

A Continuación se tiene que colocar los nombres de los otros miembros de cluster. Como estamos en `inkaserver6`, se coloca `inkaserver7` y a continuación la combinación de teclas `ctrl. + D`. Así se indica en la figura 3.8.

```

192.168.0.52 - PuTTY
>>> Cluster Nodes <<<

This Sun Cluster release supports a total of up to 16 nodes.

Please list the names of the other nodes planned for the initial
cluster configuration. List one node name per line. When finished,
type Control-D:

Node name (Control-D to finish): inkaserver7
Node name (Control-D to finish): ^D

This is the complete list of nodes:

inkaserver6
inkaserver7

Is it correct (yes/no) [yes]? yes

```

Figura 3.8 Selección de miembros del cluster

Luego se elige las interfaces de red que serán destinadas para la interconexión. Estas interfaces se habían definido que sean e1000g2 y e1000g3 en ambos nodos. Esto se aprecia en la figura 3.9.

```

192.168.0.52 - PuTTY
Select the first cluster transport adapter for "inkaserver6":

1) e1000g1
2) e1000g2
3) e1000g3
4) Other

Option: 2

Will this be a dedicated cluster transport adapter (yes/no) [yes]?

Searching for any unexpected network traffic on "e1000g2" ... done
Verification completed. No traffic was detected over a 10 second
sample period.

Select the second cluster transport adapter for "inkaserver6":

1) e1000g1
2) e1000g2
3) e1000g3
4) Other

Option: 3

```

Figura 3.9 Selección de adaptadores de red

Finalmente la instalación pregunta si debe interrumpirse la creación del cluster si hay errores durante la verificación, Seleccionamos yes y comienza la instalación.

Installation and Configuration

Log file - /var/cluster/logs/install/scinstall.log.16210

Testing for "/globaldevices" on "inkaserver6" ... done
Testing for "/globaldevices" on "inkaserver7" ... done

Starting discovery of the cluster transport configuration.

The following connections were discovered:

inkaserver6:e1000g2 switch1 inkaserver7:e1000g2
inkaserver6:e1000g3 switch2 inkaserver7:e1000g3

Completed discovery of the cluster transport configuration.

Started sccheck on "inkaserver6".
Started sccheck on "inkaserver7".

Configuring "inkaserver7" ... done

Rebooting "inkaserver7" ... Apr 26 10:10:06 inkaserver6 genunix: WARNING: e1000g2: fault detected external to device; service degraded
Feb 26 10:10:06 inkaserver6 genunix: WARNING: e1000g2: xcvr addr:0x01 - link down
Feb 26 10:10:06 inkaserver6 genunix: WARNING: e1000g3: fault detected external to device; service degraded
Feb 26 10:10:06 inkaserver6 genunix: WARNING: e1000g3: xcvr addr:0x01 - link down
Feb 26 10:10:08 inkaserver6 genunix: NOTICE: e1000g2: fault cleared external to device; service available
Feb 26 10:10:08 inkaserver6 genunix: NOTICE: e1000g2: xcvr addr:0x01 - link up 100 Mbps full duplex
Feb 26 10:10:08 inkaserver6 genunix: NOTICE: e1000g3: fault cleared external to device; service available
Feb 26 10:10:08 inkaserver6 genunix: NOTICE: e1000g3: xcvr addr:0x01 - link up 100 Mbps full duplex
Feb 26 10:11:34 inkaserver6 genunix: WARNING: e1000g2: fault detected external to device; service degraded
Feb 26 10:11:34 inkaserver6 genunix: WARNING: e1000g2: xcvr addr:0x01 - link down
Feb 26 10:11:34 inkaserver6 genunix: WARNING: e1000g3: fault detected external to device; service degraded
Feb 26 10:11:34 inkaserver6 genunix: WARNING: e1000g3: xcvr addr:0x01 - link down
Feb 26 10:11:36 inkaserver6 genunix: NOTICE: e1000g2: fault cleared external to device; service available
Feb 26 10:11:36 inkaserver6 genunix: NOTICE: e1000g2: xcvr addr:0x01 - link up 100 Mbps full duplex

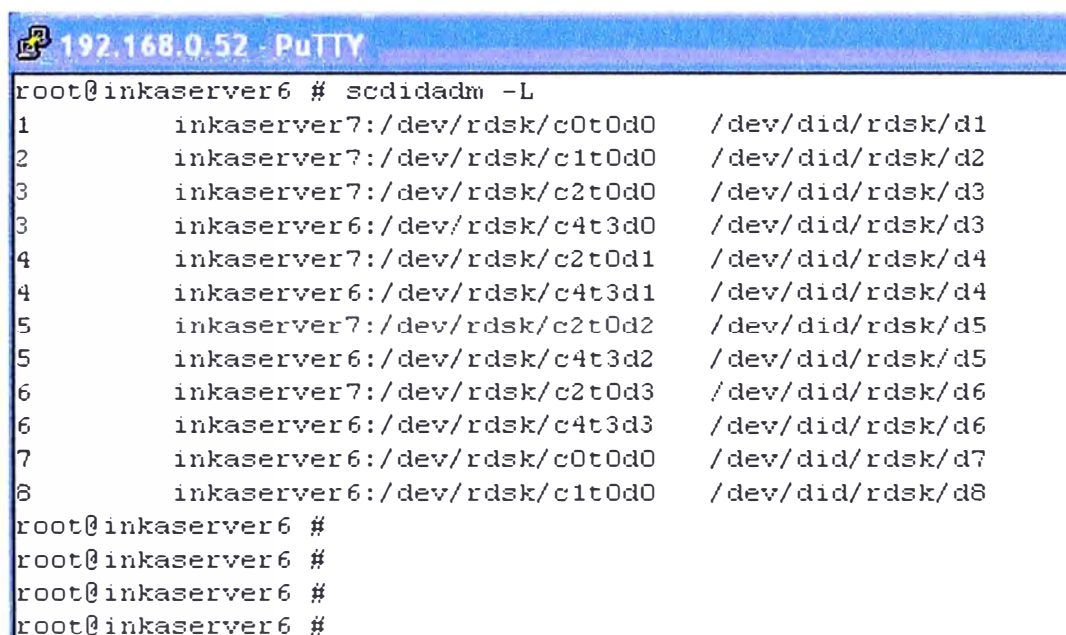
```
Feb 26 10:11:36 inkaserver6 genunix: NOTICE: e1000g3: fault cleared external to
device; service available
```

```
Feb 26 10:11:36 inkaserver6 genunix: NOTICE: e1000g3: xcvr addr:0x01 - link up
100 Mbps full duplex
```

```
scinstall -ik \
-C sondacluster \
-N inkaserver7 \
-A trtype=dlpi,name=e1000g2 -A trtype=dlpi,name=e1000g3 \
-m endpoint=:e1000g2,endpoint=switch1 \
-m endpoint=:e1000g3,endpoint=switch2
```

3.4.4 Instalación de Quórum vote

Se ejecuta el comando `scdidadm -L` para seleccionar el DID que sera destinado como quórum. En la figura 3.10 podemos apreciar la salida del comando realizado.



```
root@inkaserver6 # scdidadm -L
1      inkaserver7:/dev/rdisk/c0t0d0    /dev/did/rdisk/d1
2      inkaserver7:/dev/rdisk/c1t0d0    /dev/did/rdisk/d2
3      inkaserver7:/dev/rdisk/c2t0d0    /dev/did/rdisk/d3
3      inkaserver6:/dev/rdisk/c4t3d0    /dev/did/rdisk/d3
4      inkaserver7:/dev/rdisk/c2t0d1    /dev/did/rdisk/d4
4      inkaserver6:/dev/rdisk/c4t3d1    /dev/did/rdisk/d4
5      inkaserver7:/dev/rdisk/c2t0d2    /dev/did/rdisk/d5
5      inkaserver6:/dev/rdisk/c4t3d2    /dev/did/rdisk/d5
6      inkaserver7:/dev/rdisk/c2t0d3    /dev/did/rdisk/d6
6      inkaserver6:/dev/rdisk/c4t3d3    /dev/did/rdisk/d6
7      inkaserver6:/dev/rdisk/c0t0d0    /dev/did/rdisk/d7
8      inkaserver6:/dev/rdisk/c1t0d0    /dev/did/rdisk/d8
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
```

FIGURA 3.10 Relación `c##d##` con DID `d##`

Luego Ejecutamos el comando `scsetup` para adicionar el dispositivo de quórum, que es un disco compartido directamente conectado y que tiene el DID `d3`

```
root@inkaserver6 # scsetup
```

```
>>> Initial Cluster Setup <<<
```

This program has detected that the cluster "installmode" attribute is still enabled. As such, certain initial cluster setup steps will be performed at this time. This includes adding any necessary quorum

devices, then resetting both the quorum vote counts and the "installmode" property.

Please do not proceed if any additional nodes have yet to join the cluster.

Is it okay to continue (yes/no) [yes]? yes

Do you want to add any quorum devices (yes/no) [yes]? Yes

What is the type of device you want to use?

- 1) Directly attached shared disk
- 2) Network Attached Storage (NAS) from Network Appliance

q)

Option: 1

>>> Add a SCSI Quorum Disk <<<

A SCSI quorum device is considered to be any Sun Cluster supported attached storage which connected to two or more nodes of the cluster. Dual-ported SCSI-2 disks may be used as quorum devices in two-node clusters. However, clusters with more than two nodes require that SCSI-3 PGR disks be used for all disks with more than two node-to-disk paths.

You can use a disk containing user data or one that is a member of a device group as a quorum device.

For more information on supported quorum device topologies, see the Sun Cluster documentation.

Is it okay to continue (yes/no) [yes]?

Which global device do you want to use (d<N>)? d3

Is it okay to proceed with the update (yes/no) [yes]?

```
scconf -a -q globaldev=d3
```

Command completed successfully.

Press Enter to continue:

Do you want to add another quorum device (yes/no) [yes]? no

Once the "installmode" property has been reset, this program will skip "Initial Cluster Setup" each time it is run again in the future. However, quorum devices can always be added to the cluster using the regular menu options. Resetting this property fully activates quorum settings and is necessary for the normal and safe operation of the

cluster.

Is it okay to reset "installmode" (yes/no) [yes]?

```
scconf -c -q reset `
root@inkaserver6 #
```

El siguiente paso es verificar el estado del cluster

3.4.5 Verificar Estado de Cluster

Con el comando `scstat` se verifica la configuración del Cluster hasta el momento, se puede observar que los nodos están en estado online, Los adaptadores de interconexión están operativos, los 3 votos del quórum están presentes, no hay todavía ningún recurso configurado dentro del cluster y los adaptadores de red de la red pública están activas.

```
root@inkaserver6 # scstat
-----

-- Cluster Nodes --

      Node name      Status
      -----      -
Cluster node:  inkaserver7      Online
Cluster node:  inkaserver6      Online

-----

-- Cluster Transport Paths --

      Endpoint          Endpoint          Status
      -----          -
Transport path: inkaserver7:e1000g3  inkaserver6:e1000g3  Path online
Transport path: inkaserver7:e1000g2  inkaserver6:e1000g2  Path online

-----

-- Quorum Summary --

Quorum votes possible:  3
Quorum votes needed:    2
Quorum votes present:   3

-- Quorum Votes by Node --

      Node Name      Present Possible Status
      -----      -
```

```

Node votes:   inkaserver7      1    1    Online
Node votes:   inkaserver6      1    1    Online

```

-- Quorum Votes by Device --

```

          Device Name      Present Possible Status
          -----
Device votes: /dev/did/rdisk/d3s2 1    1    Online

```

-- Device Group Servers --

```

          Device Group      Primary      Secondary
          -----

```

-- Device Group Status --

```

          Device Group      Status
          -----

```

-- Multi-owner Device Groups --

```

          Device Group      Online Status
          -----

```

-- IPMP Groups --

```

          Node Name      Group  Status  Adapter  Status
          -----
IPMP Group: inkaserver7      ipmp  Online   e1000g0  Online
IPMP Group: inkaserver6      ipmp  Online   e1000g0  Online

```

Con esto ya Se tiene el Software de Sun Cluster ya instalado, aunque sin ningún recurso bajo control del cluster.

CAPITULO IV

CONFIGURACION DE RECURSOS DE CLUSTER

4.1 Estructura de la Base de Datos Oracle

Ahora se tiene el cluster activo pero aun no hay ninguna aplicación que este siendo monitoreada.

El Primer paso a realizar es crear los sistemas de archivos o File system donde estará instalada la base de datos. Para ello se revisara el funcionamiento de la base datos Oracle

4.1.1 Instancia de la Base de Datos

La base de datos Oracle esta compuesto de archivos ejecutables que son los encargados de ejecutar los procesos de oracle, en general el conjunto de procesos es denominado instancia de la base de datos. La instancia se encarga de reservar espacio de Memoria RAM en el sistema operativo, y se encarga de iniciar y detener el servicio de base de datos Oracle. La instancia debe ser controlado por un usuario del sistema operativo diferente al súper usuario, generalmente se crea un usuario oracle para este propósito.

Los archivos de Configuración de la Base de Datos se almacenan en una ruta definida en el perfil del usuario oracle, específicamente en la variable ORACLE_HOME. Así también los archivos de la Instancia de Oracle También se encuentran en esta ruta.

4.1.2 Archivos de Datos o Data file

Los archivos de datos o data file, pueden tener una ruta diferente de la indicada en ORACLE_HOME, es en estos archivos donde se almacena la información para el cual fue creada la base de datos. Se tiene un mejor control de espacios en disco cuando estos archivos se encuentran en un sistema de archivos diferente de la Instancia de la Base de datos.

4.1.3 Archivos Archive

Cuando se ejecutan transacciones en la base de datos se modifican los archivos de datos, y estos cambios son almacenados en archivos denominados archive de la base de datos. Estos archivos permiten recuperar los cambios realizados en caso haya algún

problema con los data file. De esta forma si la base de datos sufre una caída intempestiva, al iniciarse Oracle usa los Archive para reconstruir los data file.

4.1.4 Listener

El Listener es un proceso de Oracle que permite establecer las conexiones de los clientes a la base de datos, la función de Listener empieza con las peticiones de conexión de los usuarios y finaliza cuando se establece la sesión. El Listener es un proceso que no pertenece a la Instancia de la base de datos, Sin embargo su uso es imprescindible para las conexiones remotas.

4.2 Creación de los Sistemas de Archivos de la Base de Datos

Para obtener un mayor rendimiento se crearan 2 sistemas de archivos para los data file, además se utilizara otro para almacenar los archivos binarios de la instancia de la base de datos y un cuarto para los archive.

La división de la base de datos en varios sistemas de archivos permite mayor rendimiento de entrada y salida de datos al utilizar mas de un sistema de archivos a la vez.

La ubicación de los sistemas de archivos debe estar en el storage para asegurar que ambos nodos puedan tener acceso a cada uno de ellos.

4.2.1 Creación de Metaset

Sun Cluster usa metaset para control de los discos por lo que el primer paso es crear los metaset. Para una mejor administración, se crea 1 metaset para los archivos binarios de la base de datos y otro para los data file y archive.

Recordando la lista de dispositivos disponibles

```
root@inkaserver6 # sctidadm -L
1   inkaserver7:/dev/rdisk/c0t0d0   /dev/did/rdsk/d1
2   inkaserver7:/dev/rdisk/c1t0d0   /dev/did/rdsk/d2
3   inkaserver7:/dev/rdisk/c2t0d0   /dev/did/rdsk/d3
3   inkaserver6:/dev/rdisk/c4t3d0   /dev/did/rdsk/d3
4   inkaserver7:/dev/rdisk/c2t0d1   /dev/did/rdsk/d4
4   inkaserver6:/dev/rdisk/c4t3d1   /dev/did/rdsk/d4
5   inkaserver7:/dev/rdisk/c2t0d2   /dev/did/rdsk/d5
5   inkaserver6:/dev/rdisk/c4t3d2   /dev/did/rdsk/d5
6   inkaserver7:/dev/rdisk/c2t0d3   /dev/did/rdsk/d6
6   inkaserver6:/dev/rdisk/c4t3d3   /dev/did/rdsk/d6
7   inkaserver6:/dev/rdisk/c0t0d0   /dev/did/rdsk/d7
8   inkaserver6:/dev/rdisk/c1t0d0   /dev/did/rdsk/d8
root@inkaserver6 #
```

Los DID que pertenecen al storage son d3, d4, d5 y d6. Revisando en el sistema operativo se tiene los siguientes tamaños

d3 20GB
d4 90GB
d5 90Gb
d6 72GB

Se procede a crear el metaset orabin con el did d3

```
root@inkaserver6 # metaset -s orabin -a -h inkaserver6 inkaserver7
root@inkaserver6 # metaset -s orabin -a -m inkaserver6 inkaserver7
root@inkaserver6 #
root@inkaserver6 # metaset -s orabin -a /dev/did/rdisk/d3
root@inkaserver6 # metaset
```

Set name = orabin, Set number = 1

Host	Owner
inkaserver6	Yes
inkaserver7	

Mediator Host(s)	Aliases
inkaserver6	
inkaserver7	

Driv Dbase

d3 Yes
root@inkaserver6 #

De manera similar se crea el metaset oraprod con los did d4, d5 y d6

```
root@inkaserver6 # metaset -s oraprod -a -h inkaserver6 inkaserver7
root@inkaserver6 # metaset -s oraprod -a -m inkaserver6 inkaserver7
root@inkaserver6 #
root@inkaserver6 # metaset -s oraprod -a /dev/did/rdisk/d4 /dev/did/rdisk/d5
/dev/did/rdisk/d6
root@inkaserver6 # metaset
```

Set name = oraprod, Set number = 2

Host	Owner
inkaserver6	Yes
inkaserver7	

Mediator Host(s)	Aliases
inkaserver6	
inkaserver7	

Driv Dbase

d4 Yes

d5 Yes

d6 Yes

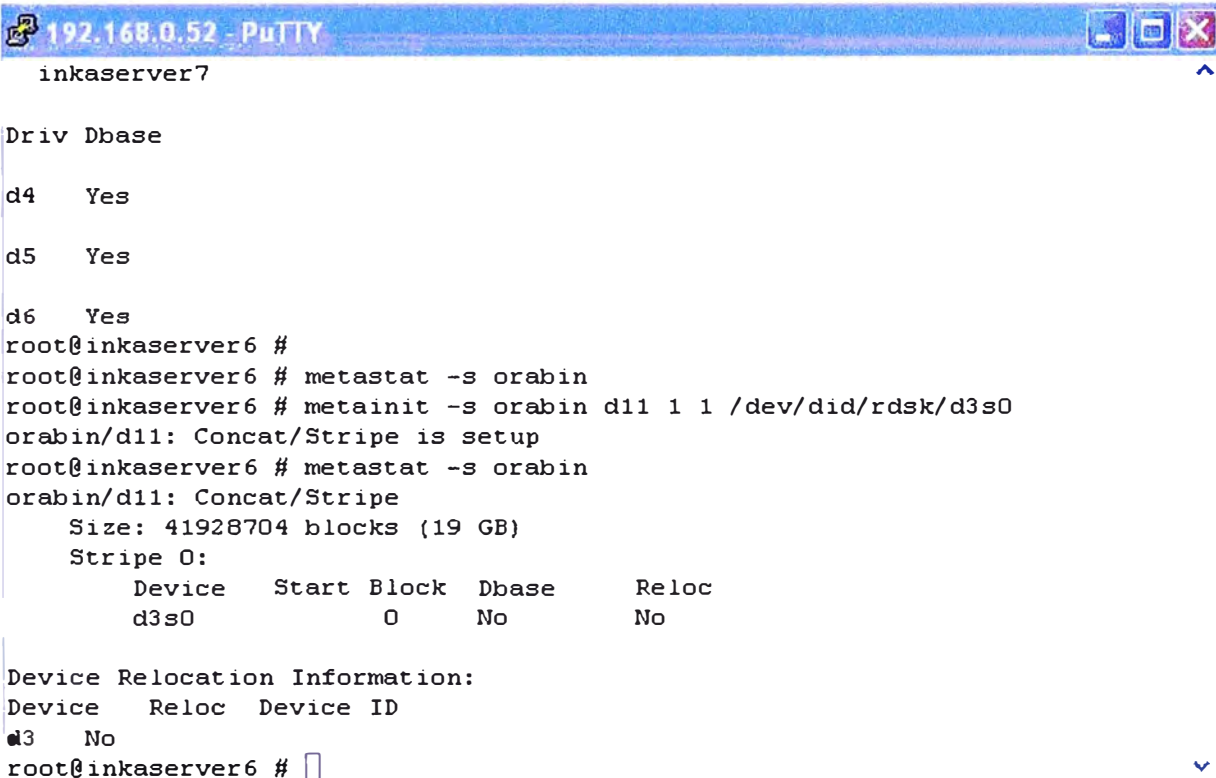
root@inkaserver6 #

4.2.2 Creación de los Metadevices

Ahora se crean los metadevices o particiones de los metaset. En este caso basta con crear una sola partición por disco de un metaset, ocupando toda la capacidad del disco respectivo, un metadevice tiene una nomenclatura d#. Sin embargo no debe confundirse con los DID que son denominados de la misma manera.

Para crear los metadevice se utiliza el comando metainit seguido del metaset, el número de dispositivo y el disco del metaset a utilizar. El comando completo se aprecia en la figura 4.1 para el metadevice d11 del metaset orabin, y en la figura 4.2 se observa la creación de los metaset d21, d22 y d23 del metaset oraprod.

El comando metastat seguido de la opción “-s” muestra la información del metaset respectivo, por ejemplo, en la figura 4.1 se muestra el tamaño de 19GB para el metadevice d11



```

192.168.0.52 - PuTTY
inkaserver7

Driv Dbase
d4 Yes
d5 Yes
d6 Yes
root@inkaserver6 #
root@inkaserver6 # metastat -s orabin
root@inkaserver6 # metainit -s orabin d11 1 1 /dev/did/rdisk/d3s0
orabin/d11: Concat/Stripe is setup
root@inkaserver6 # metastat -s orabin
orabin/d11: Concat/Stripe
      Size: 41928704 blocks (19 GB)
      Stripe 0:
            Device   Start Block  Dbase      Reloc
            d3s0           0         No         No

Device Relocation Information:
Device  Reloc  Device ID
d3      No
root@inkaserver6 #

```

FIGURA 4.1 Creación del metadevice orabin/d11

```

192.168.0.52 - PuTTY
root@inkaserver6 # metainit -s orabin d11 1 1 /dev/did/rdsk/d3s0
orabin/d11: Concat/Stripe is setup
root@inkaserver6 # metastat -s orabin
orabin/d11: Concat/Stripe
Size: 41928704 blocks (19 GB)
Stripe 0:
Device      Start Block  Dbase      Reloc
d3s0        0            No         No

Device Relocation Information:
Device  Reloc  Device ID
d3      No   -
root@inkaserver6 # metastat -s oraprod
root@inkaserver6 # metainit -s oraprod d21 1 1 /dev/did/rdsk/d4s0
oraprod/d21: Concat/Stripe is setup
root@inkaserver6 # metainit -s oraprod d22 1 1 /dev/did/rdsk/d5s0
oraprod/d22: Concat/Stripe is setup
root@inkaserver6 # metainit -s oraprod d23 1 1 /dev/did/rdsk/d6s0
oraprod/d23: Concat/Stripe is setup
root@inkaserver6 # metastat -s oraprod -p
oraprod/d23 1 1 /dev/did/rdsk/d6s0
oraprod/d22 1 1 /dev/did/rdsk/d5s0
oraprod/d21 1 1 /dev/did/rdsk/d4s0
root@inkaserver6 #

```

FIGURA 4.2 Creación de los metadevice d21, d22 y d23 de oraprod

4.2.3 Creación de los Sistemas de Archivos

Luego se procede a formatear los metadevices recientemente creados

```

root@ilinkaserver6#newfs /dev/md/orabin/rdsk/d11
root@ilinkaserver6#newfs /dev/md/oraprod/rdsk/d21
root@ilinkaserver6#newfs /dev/md/oraprod/rdsk/d22
root@ilinkaserver6#newfs /dev/md/oraprod/rdsk/d23

```

Se crea los puntos de montaje de los metadevice

```

root@inkaserver6#mkdir -p /global/oracle
root@inkaserver6#mkdir -p /global/oraprod/u01
root@inkaserver6#mkdir -p /global/oraprod/u02
root@inkaserver6#mkdir -p /global/oraprod/archivelog

```

Se modifican los permisos para que puedan ser accedidos por el usuario oracle

```

root@inkaserver6#chown oracle:dba -R /global/oracle
root@inkaserver6# chown oracle:dba -R /global/oraprod/u01
root@inkaserver6# chown oracle:dba -R /global/oraprod/u02
root@inkaserver6# chown oracle:dba -R /global/oraprod/archivelog

```

Finalmente se monta los Filesystem con el comando mount

```

root@inkaserver6#mount /dev/md/orabin/dsk/d11 /global/oracle
root@inkaserver6# mount /dev/md/oraprod/dsk/d11 /global/oraprod/u01
root@inkaserver6# mount /dev/md/oraprod/dsk/d21 /global/oraprod/u02
root@inkaserver6# mount /dev/md/oraprod/dsk/d23 /global/oraprod/archivelog

```

Con los sistemas de archivos ya se tiene todo lo necesario para iniciar la instalación de la base de datos Oracle

4.3 Instalación de la Base de Datos

La Instalación de la base de datos fue realizada por el administrador de la base de datos de Inkafarma. Se instala en directorios que pertenecen al storage, como /global/oracle (sección anterior 4.2.3). Terminada la Instalación se crea un usuario en la base de datos llamado sc_fm que Sun Cluster usara para la supervisión de recursos.

4.4 Habilitación de los Recursos

Antes de configurar los recursos se debe entender su significado, a continuación algunos conceptos importantes acerca de los recursos de cluster

4.4.1 Servicios de datos

Un servicio de datos es la combinación de software y archivos de configuración que permite a una aplicación ejecutarse en una configuración de Sun Cluster, ya que en esta circunstancia se ejecuta como recurso bajo el control del Gestor del grupo de recursos (RGM). Un servicio de datos permite configurar una aplicación como la base de datos de Oracle para ejecutarse en un clúster de varios nodos en lugar de un único servidor.

El software de un servicio de datos proporciona implementaciones de los métodos de gestión de Sun Cluster que efectúan las operaciones siguientes en la aplicación:

Iniciar la aplicación

Parar la aplicación

Supervisar los errores de la aplicación y recuperar después de estos fallos

Los archivos de configuración de un servicio de datos definen las propiedades del recurso que representa la aplicación en el gestor del grupo de recursos.

El RGM (gestor del grupo de recursos) controla la disposición de la recuperación de fallos y los servicios de datos escalables en el clúster y es responsable del inicio y de la parada de los servicios de datos en los nodos seleccionados del clúster, en respuesta a los cambios en la pertenencia al clúster; permite a las aplicaciones de servicios de datos utilizar la estructura del clúster.

4.4.2 Tipos de recursos

Un tipo de recurso es un conjunto de propiedades que describen una aplicación al clúster y que contiene información sobre cómo se debe iniciar, detener y supervisar la aplicación

en los nodos del clúster; incluye, también, las propiedades específicas de la aplicación que es necesario definir para usar ésta en el clúster. Los servicios de datos de Sun Cluster tienen varios tipos de recursos predefinidos. Sun Cluster HA para Oracle comprende los recursos del tipo SUNW.oracle_server y SUNW.oracle_listener. Para el caso de disco compartido se utiliza el tipo de recurso SUNW.HAStoragePlus y para el tipo de Recurso Logical Hostname que permite tener una ip virtual entre 2 nodos se utiliza SUNW.LogicalHostname

4.4.3 Recursos

Un recurso es una instancia de un tipo de recurso que está definido en todo el clúster y que permite que varias instancias de una aplicación se instalen en éste. Cuando se inicializa un recurso, RGM asigna valores a las propiedades específicas de la aplicación y el recurso hereda las propiedades en el tipo de recurso.

Los servicios de datos usan varios tipos de recursos. Las aplicaciones como la Base de Datos Oracle usan direcciones de red (nombres de sistema lógicos y direcciones compartidas) de las que dependen las aplicaciones. Los recursos de aplicación y red forman una unidad básica que gestiona RGM.

4.4.4 Grupos de recursos

Los recursos gestionados por RGM se colocan en grupos de recursos, de modo que se puedan gestionar como una unidad. Un grupo de recursos es un conjunto de recursos interdependientes o relacionados. Para la base de datos Oracle, es común usar un recurso derivado del tipo de recurso SUNW.LogicalHostname que se sitúe en el mismo grupo de recursos que el derivado de un tipo de recurso de base de datos de Oracle. El grupo de recursos migra como unidad si se inicia una recuperación de fallos o un cambio.

4.4.5 Servicios de datos a prueba de fallos o Failover

La recuperación de fallos es un proceso por el cual el clúster reubica automáticamente una aplicación de un nodo principal con fallos en un nodo designado secundario y redundante. Las aplicaciones a prueba de fallos pueden ejecutarse en un solo nodo del clúster

Si el supervisor de fallos detecta un error, intenta reiniciar la instancia en otro nodo del cluster (recuperación de fallos). Los servicios a prueba de fallos usan un grupo de recursos a prueba de fallos: un contenedor para los recursos de instancias de aplicaciones y recursos de la red (nombres de sistemas lógicos). Éstos son direcciones IP

que pueden configurarse como activas en un nodo y, posteriormente, configurarse automáticamente como inactivas en el nodo original y activarse en otro nodo.

Es posible que los clientes sufran una breve interrupción del servicio y es posible que necesiten volver a conectarse después de terminar la recuperación del fallo.

4.4.6 Supervisores de errores

El sistema Sun Cluster permite que todos los componentes de la “ruta” entre los usuarios y los datos queden totalmente disponibles, al supervisar las aplicaciones, el sistema de archivos y las interfaces de la red.

Sun Cluster detecta rápidamente un fallo en el nodo y designa un servidor equivalente para los recursos en el nodo que ha fallado. Sun Cluster se asegura de que los recursos que no resulten afectados por el nodo incorrecto y estén siempre disponibles durante la recuperación y que los recursos del nodo defectuoso queden disponibles tan pronto como se recuperen.

Cada servicio de datos de Sun Cluster proporciona un supervisor de fallos que explora periódicamente el servicio de datos para determinar su buen estado. Un supervisor de errores comprueba que el proceso de la aplicación o los procesos se estén ejecutando y que los clientes reciban servicio.

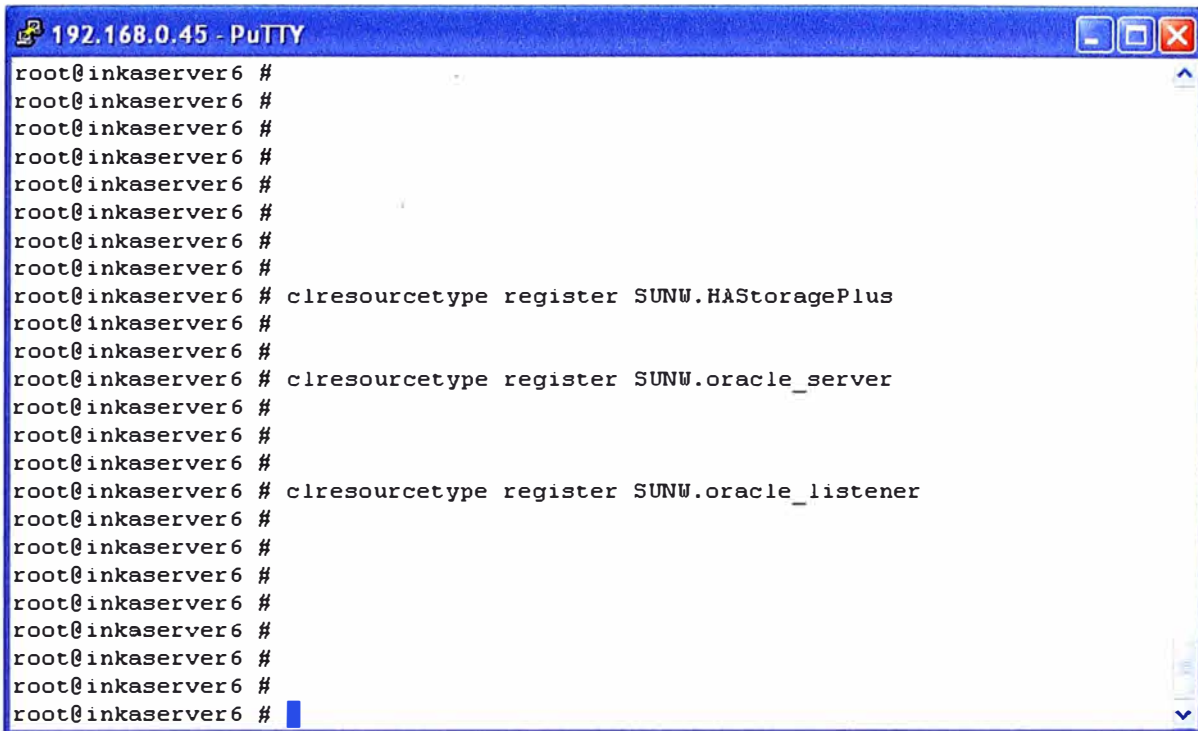
4.5 Registro de Tipo Recursos

Para el Registro de tipo de recursos usamos el comando `clresource`, por defecto el tipo de recursos `SUNW.LogicalHostname` ya se encuentra registrado en el RGM. Por lo que solo hace falta registrar los tipos de recursos de disco y de Oracle. El registro de los recursos se aprecia en la figura 4.3. Durante la instalación del Software de Sun Cluster se instaló los componentes de alta disponibilidad de Sun Cluster para oracle, por lo que solo hacia falta registrar estos componentes en la configuración CCR del Cluster. La mayoría de comandos de cluster empiezan con las letras “cl” esta característica se utiliza desde la versión 3.2 de Sun Cluster.

4.6 Creación de Recursos

A continuación se procede a crear el grupo de recursos `ora-rg` que contendrá todos los recursos necesarios para la base de datos, asociándolos a los nodos del cluster.

```
root@inkaserver6 # scrgadm -a -g ora-rg -h inkaserver6,inkaserver7
```

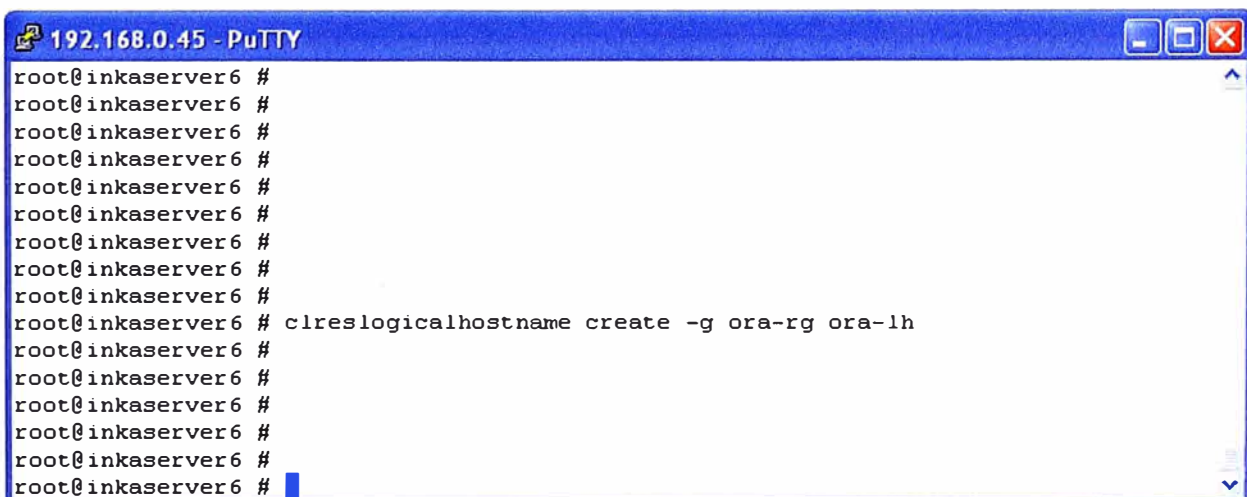


```
192.168.0.45 - PuTTY
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 # cresource type register SUNW.HAStoragePlus
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 # cresource type register SUNW.oracle_server
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 # cresource type register SUNW.oracle_listener
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
```

FIGURA 4.3 Registro de Tipo de Recursos

Luego se procede a crear los recursos del grupo de recursos. Primero se crea el recurso del tipo logical-hostname ora-lh que tendrá la ip de conexión de la base de datos, como se muestra en la figura 4.4. Este recurso se crea dentro del grupo de recursos ora-rg previamente creado.

En el archivo /etc/hosts se ingresa el nombre ora-lh asociandolo a la dirección ip. Al crearse el recurso Sun Cluster busca el archivo /etc/hosts para resolver el nombre ora-lh e iniciar la ip correspondiente.

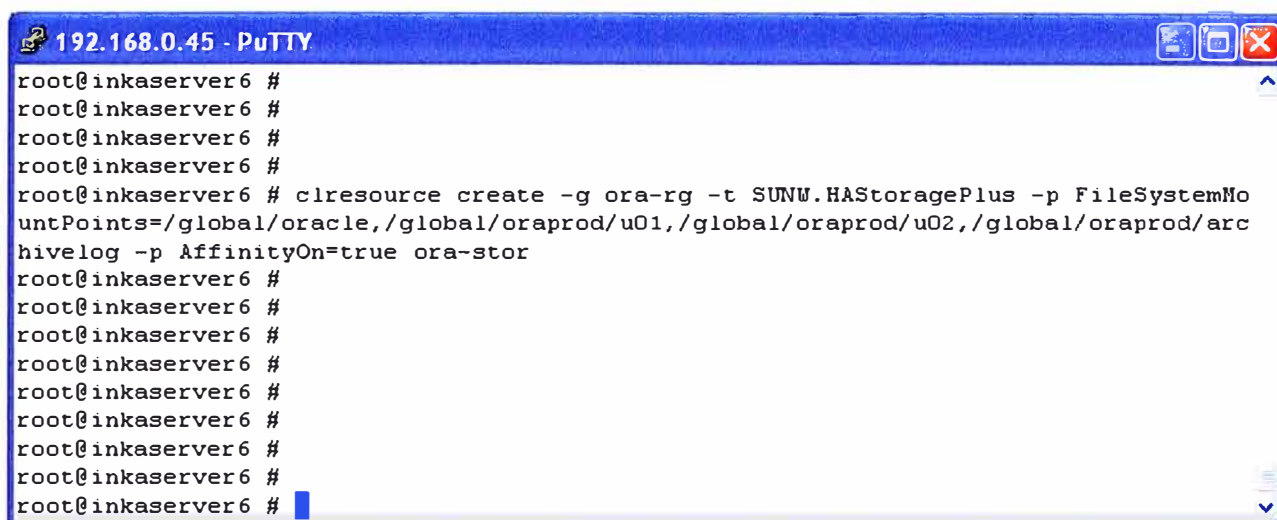


```
192.168.0.45 - PuTTY
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 # cresource logicalhostname create -g ora-rg ora-lh
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
```

FIGURA 4.4 creación del Recurso Logical Hostname

A continuación el recurso de disco ora-stor que tendrá como valores todos los puntos de montaje que deseamos estén en el control del cluster. Como se muestra en la figura 4.5.

Sun Cluster busca estos puntos de montaje en el archivo de inicialización de sistemas de archivos del sistema operativo (/etc/vfstab) de donde determina el metaset, o los metaset al cual pertenecen. Dado que los puntos de montaje pertenecen a los metaset orabin y oraprod, el control de estos metaset pasa a estar a manos de Sun Cluster.



```

192.168.0.45 - PuTTY
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 # clresource create -g ora-rg -t SUNW.HAStoragePlus -p FileSystemMountPoints=/global/oracle,/global/oraprod/u01,/global/oraprod/u02,/global/oraprod/arc
hivelog -p AffinityOn=true ora-stor
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #

```

FIGURA 4.5 Creación del Recurso de Disco

Luego se crea los recursos de la base de datos y el listener. El recurso de la base de datos es denominado ora-server-res y al momento de crearlo, también se especifica, la ruta donde está instalado los archivos binarios del oracle ORACLE_HOME, la ruta donde se depositarán las alertas referidas a la base de datos (Alert_log_file) el archivo de inicialización de variables de la base de datos (Parameter_file), el nombre de identificación de la base de datos (SID) y el usuario de la base de datos con el que Sun Cluster realizará la supervisión del servicio (Connect_string). Esto se puede apreciar en la figura 4.6, todas estas variables se configuraron durante la instalación de la base de datos.

Además se incorpora una dependencia con el recurso ora-stor, de tal manera que este servicio inicie solo si el recurso ora-stor está presente. Esto es deseado puesto que deben estar presentes los discos con la información para poder iniciar la base de datos. Con la creación de este recurso se inicia la instancia de la base de datos.

Este comando se ejecuta en un solo nodo por ser parte de la configuración del cluster, el CCR, que almacena toda la configuración del cluster, se encarga de replicarlo a los nodos restantes.

En la primera parte de la salida del comando `scstat` se observa que los nodos están activos e integrados al cluster, los cables de interconexión están operativos, Los votos del quórum están completos, tanto de los nodos como el dispositivo de quórum del storage.

```

root@inkaserver7 # scstat
-----
-- Cluster Nodes --
      Node name      Status
      -----      -
Cluster node:      inkaserver7      Online
Cluster node:      inkaserver6      Online
-----

-- Cluster Transport Paths --
      Endpoint      Endpoint      Status
      -----      -
Transport path:      inkaserver7:e1000g3      inkaserver6:e1000g3      Path online
Transport path:      inkaserver7:e1000g2      inkaserver6:e1000g2      Path online
-----

-- Quorum Summary --
Quorum votes possible:      3
Quorum votes needed:      2
Quorum votes present:      3

-- Quorum Votes by Node --
      Node Name      Present Possible Status
      -----      -
Node votes:      inkaserver7      1      1      Online
Node votes:      inkaserver6      1      1      Online

-- Quorum Votes by Device --
      Device Name      Present Possible Status
      -----      -
Device votes:      /dev/did/rdisk/d3s2      1      1      Online
-----

```

FIGURA 4.8 `scstat` primera parte

En la segunda parte del `scstat` se observa el estado de los metaset, los cuales están activos. Luego muestra información de los recursos. Empezando por el grupo de recursos `ora-rg` que se encuentra activo en el nodo `inkaserver6`, en el nodo `inkaserver7` nos indica que esta `offline`, es no esta activo en ese nodo pero es candidato apto para recibir los recursos del grupo de recursos.

Luego puede ver los 4 recursos creados, del grupo de recursos `ora-rg`, todos ellos a activos en el nodo `inkaserver6`

```

192.168.0.51 - PuTTY
Device Group      Status
-----
Device group status:  orabin           Online
Device group status:  oraprod          Online

-- Multi-owner Device Groups --

Device Group      Online Status
-----

-----

-- Resource Groups and Resources --

Group Name      Resources
-----
Resources: ora-rg      ora-lh ora-stor ora-listener-res ora-server-res

-- Resource Groups --

Group Name      Node Name      State      Suspended
-----
Group: ora-rg      inkaserver6    Online     No
Group: ora-rg      inkaserver7    Offline    No

-- Resources --

Resource Name    Node Name      State      Status Message
-----
Resource: ora-lh      inkaserver6    Online     Online - LogicalHostna
Resource: ora-lh      inkaserver7    Offline    Offline

Resource: ora-stor      inkaserver6    Online     Online
Resource: ora-stor      inkaserver7    Offline    Offline

Resource: ora-listener-res inkaserver6    Online     Online
Resource: ora-listener-res inkaserver7    Offline    Offline

Resource: ora-server-res inkaserver6    Online     Online
Resource: ora-server-res inkaserver7    Offline    Offline

```

FIGURA 4.9 scstat segunda parte

En la tercera parte nos muestra las interfaces de la red pública el cual multipath o multiruta por 2 adaptadores de red e1000g0 y e1000g1. Estando todos ellos activos, como muestra la figura 4.10.

4.8 Pruebas de Sun Cluster

A continuación Se realizaron las siguientes pruebas para verificar la alta disponibilidad del servicio y el correcto funcionamiento del cluster.

4.8.1 Primera Prueba: Intercambio del servicio entre nodos

Con los servicios que ejecuntandose en el nodo inkaserver6 se realiza el cambio forzado de los recursos a inkaserver7 para que sea este quien brinde los servicios

```

192.168.0.51 - PuTTY
-----
-- IPMP Groups --
      Node Name      Group  Status      Adapter  Status
      -----      -
IPMP Group: inkaserver7  sc_ipmp0 Online      e1000g1  Online
IPMP Group: inkaserver7  sc_ipmp0 Online      e1000g0  Online

IPMP Group: inkaserver6  sc_ipmp0 Online      e1000g1  Online
IPMP Group: inkaserver6  sc_ipmp0 Online      e1000g0  Online

-- IPMP Groups in Zones --

      Zone Name      Group  Status      Adapter  Status
      -----      -
-----
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #
root@inkaserver6 #

```

FIGURA 4.10 scstat tercera parte

Se ejecuta el siguiente comando para ello:

```
root@inkaserver6 #scswitch -z -g ora-rg -h inkaserver7
```

Resultado: Los servicios se detienen en inkaserver 6 y luego se inician en inkaserver7. Se comprueba que los servicios están presentes ahora en el nodo inkaserver7. A continuación parte del scstat.

```
-- Resources --
```

Resource Name	Node Name	State	Status Message
Resource: ora-lh	inkaserver6	Offline	Offline
Resource: ora-lh	inkaserver7	Online	Online - LogicalHostname
Resource: ora-stor	inkaserver6	Offline	Offline
Resource: ora-stor	inkaserver7	Online	Online
Resource: ora-listener-res	inkaserver6	Offline	Offline
Resource: ora-listener-res	inkaserver7	Online	Online
Resource: ora-server-res	inkaserver6	Offline	Offline
Resource: ora-server-res	inkaserver7	Online	Online

4.8.2 Segunda Prueba: Apagado del nodo con los servicios.

Se ejecuta el siguiente comando para apagar el equipo inkaserver7 cuando éste tiene los servicios corriendo en el:

```
root@inkaserver7 #init 5
```

Resultado: Los servicios que están presentes en el nodo inkaserver7 se migran a inkaserver6. El scstat nos advierte que el nodo inkaserver7 no es parte del cluster:

```
-- Quorum Summary --
```

```
Quorum votes possible: 3
Quorum votes needed: 2
Quorum votes present: 2
```

```
-- Quorum Votes by Node --
```

	Node Name	Present	Possible	Status
Node votes:	inkaserver7	0	1	Offline
Node votes:	inkaserver6	1	1	Online

```
-- Quorum Votes by Device --
```

	Device Name	Present	Possible	Status
Device votes:	/dev/did/rdisk/d3s2	1	1	Online

```
-- Resources --
```

	Resource Name	Node Name	State	Status Message
Resource:	ora-lh	inkaserver6	Online	Online - LogicalHostname
Resource:	ora-lh	inkaserver7	Offline	Offline
Resource:	ora-stor	inkaserver6	Online	Online
Resource:	ora-stor	inkaserver7	Offline	Offline

```

Resource: ora-listener-res inkaserver6      Online      Online
Resource: ora-listener-res inkaserver7      Offline     Offline

Resource: ora-server-res inkaserver6        Online      Online
Resource: ora-server-res inkaserver7        Offline     Offline

```

4.8.3 Tercera Prueba: Perdida de la Red pública de un nodo

Se desconecta físicamente los 2 cables de la red de datos del nodo inkaserver6, es decir de los adaptadores e1000g0 y e1000g1

Resultado: Los servicios están presentes en el nodo inkaserver7
 scstat indica que las interfaces de inkaserver 6 no están operativas

```

-- IPMP Groups --

      Node Name          Group  Status  Adapter  Status
      -----          -
IPMP Group: inkaserver7  sc_ipmp0 Online   e1000g1  Online
IPMP Group: inkaserver7  sc_ipmp0 Online   e1000g0  Online

IPMP Group: inkaserver6  sc_ipmp0 Offline  e1000g1  Offline
IPMP Group: inkaserver6  sc_ipmp0 Offline  e1000g0  Offline

```

4.8.4 Cuarta Prueba: Matar procesos de la Base de Datos

Se verifica los procesos de la base de procesos y se elimina uno de ellos denominado pmon. Para esto, con el comando ps verificamos su id (1571) y luego se ejecuta kill -9 para eliminarlo.

```

root@inkaserver7 #ps -ef | grep pmon
      oracle 1571  1  1 21:51:35 ?        0:00 ora_pmon_orcl
root@inkaserver7 #kill -9 1571

```

Resultado: El supervisor de errores detecta la caída del proceso y lo inicia nuevamente con un nuevo id (17403).

```

root@inkaserver7 #
root@inkaserver7 #May 26 21:58:32 inkaserver7
SC[SUNWscor.oracle_server.monitor]:ora-rg:ora-server-res: Fault monitor
detected error DBMS_ERROR: 99999 DEFAULT Action=RESTART : Monitor
detected death of Oracle background processes.
May 26 21:58:32 inkaserver7 SC[SUNWscor.oracle_server.monitor]:ora-rg:ora-
server-res: Restarting using scha_control RESOURCE_RESTART
root@inkaserver7 #
root@inkaserver7 # ps -ef | grep pmon
oracle 17403 1 1 21:58:55 ? 0:00 ora_pmon_orcl
root@inkaserver7 #

```

4.8.5 Quinta Prueba: Desconectar los cables de Interconexión

Con los servicios activos en inkaserver6, se retira los cables de red de interconexión entre nodos.

Resultado:

El nodo que no tenía los servicios, en este caso inkaserver7 deja de pertenecer al cluster, se reinicia automáticamente y queda encendido pero no operativo. A continuación parte del scstat en el nodo inkaserver6, se observa que no se permite al nodo inkaserver7 integrarse al cluster.

-- Cluster Nodes --

	Node name	Status
	-----	-----
Cluster node:	inkaserver6	Online
Cluster node:	inkaserver7	Offline

-- Cluster Transport Paths --

	Endpoint	Endpoint	Status
	-----	-----	-----
Transport path:	inkaserver6:ce2	inkaserver7:ce1	faulted
Transport path:	inkaserver6:ce0	inkaserver7:ce0	faulted

-- Quorum Summary --

Quorum votes possible:	3
Quorum votes needed:	2

Quorum votes present: 2

-- Quorum Votes by Node --

	Node Name	Present	Possible	Status
Node votes:	inkaserver6	1	1	Online
Node votes:	inkaserver7	0	1	Offline

Después de restablecer los cables de interconexión el nodo inkaserver7 automáticamente se une al cluster.

CONCLUSIONES Y RECOMENDACIONES

Después de realizado el trabajo se pueden obtener las siguientes conclusiones

- 1) Después de las pruebas realizadas se logro demostrar que el Cluster cumple con el objetivo para el cual fue instalado, el cual es asegurar la disponibilidad de acceso de los usuarios a la base de datos instalada en el conjunto de servidores y storage, pues en caso de falla del servidor principal inkaserver6 Sun Cluster migra los servicios al otro nodo inkaserver7, y viceversa cuando se esta ejecutando en el servidor inkaserver7 y los supervisores de recursos encuentran un problema en este nodo, se inicia los servicios en inkaserver6.
- 2) El tiempo que dura un cambio de servicios de un nodo a otro es menor a un minuto, y es básicamente el tiempo que demora a la instancia de la base de datos en detener los servicios en un nodo e iniciar en el otro.
- 3) Queda demostrado que Sun Cluster funciona bien en nodos que no tengan la misma cantidad de recursos, ya que los nodos utilizados tienen procesadores diferentes y diferente memoria RAM.
- 4) El Cluster al ser del tipo activo-standby permite hacer mantenimiento a los servidores, incluso pudiendo modificar los componentes físicos que tiene, agregar, modificar o retirar recursos al servidor, sin afectar el servicio que esta ejecutándose en el otro nodo. Una vez terminado el mantenimiento a un servidor bastaría hacer un intercambio de recursos al otro nodo.
- 5) Sun Cluster es un software que permite el uso de otras aplicaciones ya sea dentro del cluster o fuera de el, sin embargo solo las aplicaciones que estén dentro del cluster estaran en alta disponibilidad.
- 6) Oracle cuenta con su propio software de cluster denominado oracle RAC, sin embargo el costo de la licencia de Oracle RAC es muy elevado. Sun Cluster en cambio no tiene costo de licenciamiento y solo el soporte tiene costo, cuyo valor es mucho menor que RAC.

Así mismo se realizaron las siguientes recomendaciones:

1) Dadas las características de los Equipos, se puede determinar que el punto de menor rendimiento en la solución es el storage 3310, pues usa una controladora SCSI de 160Mbps. Por lo que se hace de conocimiento la recomendación de que si en el futuro existan problemas de rendimiento, se actualice el storage a otro de mayor velocidad de transferencia que puede ser SCSI de 320Mbps o un storage de Fibra. Los servidores ya disponen de interfaces de fibra, por lo que están listos para conectarse a un storage de Fibra. Las conexiones de los servidores a otro storage no requieren desconectar ni apagar el storage actual, tampoco requieren detener los servicios.

2) Para tener un respaldo completo del servicio, se debe tener un plan de recuperación de desastres que incluya un local de contingencia. De esta manera se protege contra desastres como incendios, cortes prologando de energía, asaltos, etc. que afecten todo el centro de cómputo, de tal manera que en esos casos se inician los servicios en el local de contingencia. Sun Microsystems cuenta con un producto diseñado para este propósito, el cual es Sun Cluster Geographic Edition. Que puede integrarse muy bien con Sun Cluster.

3) La base de datos Oracle es del tipo recuperación de fallos de tal manera mientras un nodo esta ofreciendo el servicio, el otro nodo esta a la espera sin procesamiento de aplicación. Si se instala otra base de datos en el cluster, esta base de datos que también tendría que ser del tipo recuperación de fallos, puede ejecutarse en el nodo que no este procesando la primera base de datos. Es decir Se puede tener 2 bases de datos, ejecutándose en nodos del cluster diferentes y en caso de falla de uno de ellos el otro asume toda la carga de las 2 bases de datos, de esta forma se optimiza el uso de los recursos, pues se obtiene un balanceo de carga de las bases de datos entre los nodos.

4) Sun Cluster es compatible con muchos productos, sin embargo en el diseño debe tenerse en cuenta la compatibilidad de productos tanto de software como hardware. Por ejemplo versión de sistemas operativos, aplicaciones, tipos de sistemas de archivos, versión de drivers y tipo de dispositivos, tarjetas, storage, servidores, volume manager, etc.

ANEXO A
CRONOGRAMA DE ACTIVIDADES

A.1 Cronograma de Actividades

Las actividades se desarrollaron En el presente año 2009, de acuerdo al cronograma establecido en la Tabla A1. No se incluye las reuniones de coordinación para la definición de parámetros que fue hecho durante el mes de marzo del 2008

TABLA A.1 Cronograma de Actividades

Item	Tarea	dias	Fecha
1	Instalación de Servidor T5220 T2000 y Storeedge 3310	1	23-Mar
2	Instalación de Tarjetas SCSI y Cableado	0.5	24-Mar
3	Instalación de Sistemas Operativos en ambos Servidores	1	14-Abr
4	Configuración de Storage 3310	0.5	15-Abr
5	Aplicación de Parches Recomendados	0.5	15-Abr
6	Instalación del Software de Cluster en Ambos Nodos	1	17-Abr
7	Instalación de File System en cluster para Base de Datos	0.5	20-Abr
8	Instalación de Software Oracle 10g en modo Standalone	1	21-Abr
9	Creación y Configuración de la Instancia BD para Fidelización	1	22-Abr
10	Creación de Recursos HA para Oracle y Puesta de Base de Datos en Alta Disponibilidad	1	23-Abr
11	Protocolo de Pruebas de HA de la Base de Datos	1	24-Abr
12	Revisión de Configuraciones y Administración básica de cluster	1	27-Abr

ANEXO B
GLOSARIO DE TÉRMINOS

B.1 Glosario de Términos

ALOM: Interfaz de Administración de Servidores SUN

ARP: Protocolo de Resolución de nombres.

BD: Base de datos

Controlador: En un arreglo de discos, es el elemento que distribuye el flujo de datos.

DB-9: Conector Serial de 9 hilos.

Exceed: Emulador gráfico de sistemas UNIX en Windows

GB: Conexión de red ethernet de 1 Gigabit por segundo

HA: High Availability – Alta Disponibilidad

Hostname: nombre de un equipo o servidor

ILOM: Interfaz de Administración web de Servidores SUN

IPMP: Conexión multiruta de una direcciones ip a través de 2 interfaces de red.

NFS: Protocolo de compartición de archivos por la red.

Pmon: Proceso de oracle.

Putty: Programa para conexión remota

Rack: Gabinete para equipos de computo.

Raw-device: Porción de disco sin formatear

Soft-partition: partición virtual de un disco duro

Storage: Modelo de storage o Arreglo de discos de Sun Microsystems

Storage: Arreglo de discos

SCSI: Protocolo de comunicación entre servidor y storage, aunque no es el único.

Sparc; Procesador con arquitectura RISC.

SSH: Protocolo de sesión remota segura.

Switch: Equipo de Comunicaciones para conexión de otros dispositivos.

TCP/IP: Protocolo de Internet.

Update: Nivel de actualización de un producto.

UPS: Sistema de Potencia ininterrumpida

UTP: Cable de Comunicaciones de 8 hilos.

VLAN: Grupo de puertos administrativamente aislados del resto de equipos.

ANEXO C
INFORMACIÓN DE LA CONFIGURACIÓN

C.1 Diagrama de Conexiones

En la figura C.1 se aprecia las conexiones que tienen los equipos

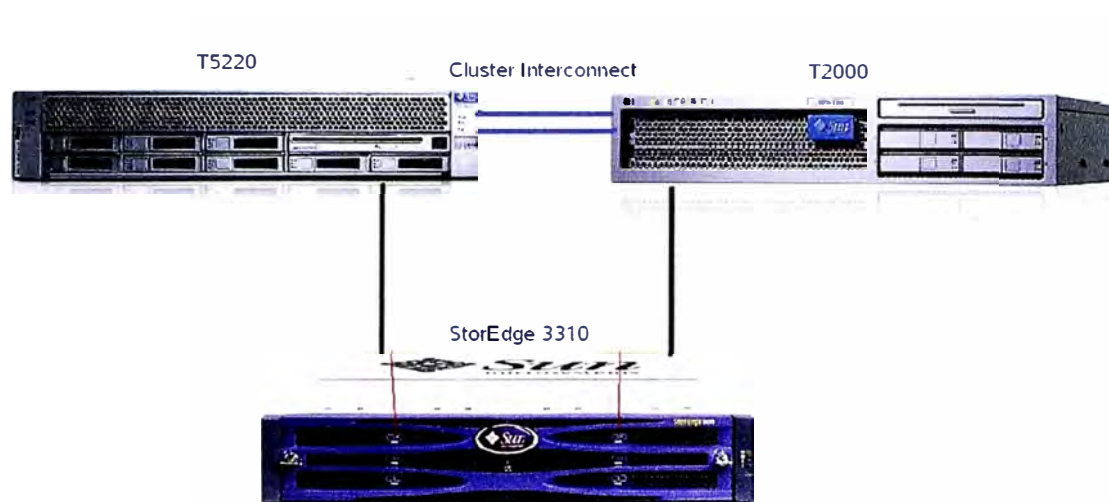


FIGURA C.1 Diagrama de Conexiones

Cabe mencionar que las conexiones del arreglo de discos hacia los servidores usan cables SCSI, mientras que las conexiones entre los servidores SUN es a través de cables UTP.

C.2 Interfaces de Administración

Las Interfaces de Administración cumplen un papel fundamental en los servidores, pues permiten obtener la salida, ver el inicio y el apagado del equipo, además de tener el monitoreo y control del hardware del equipo. Las interfaces de administración tienen 2 salidas, un serial DB-9 y un conector RJ-45 que permite conexión por red. En la Tabla C.1 se muestra las direcciones IP de los equipos.

C.3 Direcciones IP

Se configurarán las interfaces de a pares, para redundancia con IPMP.

En IPMP se utilizan 2 direcciones IP que tienen que tener conectividad con la puerta de enlace, como se muestra en la tabla C.2. Las direcciones Físicas se usan para test de las interfaces y la dirección de failover es usado para las aplicaciones.

TABLA C.1 Direcciones IP de Administración

Inkserver 6			
Inkafarma6	Tipo	Valor	Comentario
ILOM	IP	192.168.0.52	
Switch	Port	SW3 –port 2	
User		root	
Password		changeme	
Netmask		255.255.255.0	
Default Gateway		192.168.0.1	
ALOM user		admin	
Password		admin123	
Inkserver 7			
Inkserver7	Tipo	Valor	
ALOM	IP	192.168.0.54	
Switch	Port	SW3 – port 3	
User		admin	
Password		admin	
Netmask		255.255.255.0	
Default Gateway		192.168.0.1	
Storedge 3310			
Pasword		admin.\$1	
IP		192.168.0.103	
Netmask		255.255.255.0	
Default gateway		192.168.0.1	

TABLA C.2 Direcciones IP LAN

Inkserver6				
NIC	Tipo	IP/Mascara	Nombre	Switch/Port
GB-UTP 1	E1000g0	192.168.0.40/24	Test0	SW3-P8
GB-UTP 2	E1000g1	192.168.0.44/24	Test1	SW3-P9
GB-UTP 3	E1000g2	172.16.0.130/25	Interconect1	-
GB-UTP 4	E1000g3	172.16.1.2/25	Interconect2	-
virtual	Failover	192.168.0.45/24	<i>Inkserver6</i>	-
Inkserver7				
NIC	Tipo	IP/Mascara	Nombre	Switch/Port
GB-UTP 1	E1000g0	192.168.0.46/24	Test0	SW3-P12
GB-UTP 2	E1000g1	192.168.0.47/24	Test1	SW3-P13
GB-UTP 3	E1000g2	172.16.0.129/25	Interconect1	-
GB-UTP 4	E1000g3	172.16.1.1/25	Interconect2	-
virtual	Failover	192.168.0.45/24	<i>Inkserver6</i>	-

C.4 Configuración de Discos

Se considera Raid por hardware para 2 discos internos de 73GB. En la tabla C.3 Se muestran las particiones de los discos internos de cada servidor y del storage.

TABLA C.3 Particionamiento de Discos

Discos internos					
Part	Tag	Tamaño	Punto de montaje	Descripción	
s0	root	10GB	/	Root	
s1	swap	16 GB	swap	Swap	
s2	overlap	68 GB	-	Entire Disk	
s3	var	16 GB	/var	/var	
s4	-	10GB	/u01		
s5	-	10GB	/u02		
s6	-	750 MB	/globaldevices	Cluster Global Devices	
s7	-	100 MB	-	Metadb	
Storage					
Logic Drive	Raid	Part	Tam Part	Mapeo	Descripción
0	Raid 5 4x68GB	0	20GB	Inkaserver6 inkaserver7	Binarios
		1	90GB	Inkaserver6 inkaserver7	Data file
		2	90GB	Inkaserver6 inkaserver7	Data File
		3	72GB	Inkaserver6 inkaserver7	Archive
		4	50MB	Inkaserver6 inkaserver7	Quorum vote

C.5 Información de Productos Instalados

Las versiones y parámetros básicos de los productos se muestran en la tabla C.4.

Esta información fue definida y revisada en las reuniones de coordinación.

Se realizó previamente la compatibilidad entre los productos instalados.

TABLA C.4 Parámetros Generales

Solaris		
Solaris Release	Solaris 10 10/08 Update 7	
Cluster Release	Sun Cluster 3.2 01/09	
Os Type Installation	Entire Distribution Plus OEM	
Time Zone	America/Lima (Peru)	
Locale	ISO8859-1	
Volume Manager	Solaris Volume Manager	
Cluster		
Nombre del Cluster	Inkacluster	
Nodo 1	Inkaserver6	
Nodo 2	Inkaserver7	
Quorum Device	d3	
Subnet 1 – interconnect network	172.16.0.128/25	
Subnet 2 - interconnect network	172.16.1.0/25	
Oracle		
Usuario	oracle	uid = 101
Grupo	dba	gid = 101
Version	Oracle 10g R2	
Oracle Base	/global/oracle/oracle	
Oracle Home	/global/oracle/oracle/product/10.2.0/db_1	
SID	edbst008	
Listener name	LISTENER	
IP Conexion	ora-lh	192.168.0.59

BIBLIOGRAFIA

1. Compatibilidad de Sun Cluster 3.2
Sun Microsystems OSP – osp/engenio_interop.
2. Instalación de Solaris 10 con Jumpstart
Sun Microsystems Docs – doc/820-6047.
3. Administración de Solaris Volume Manager
Sun Microsystems Docs – doc/816-4520.
4. Instalación de base de datos Oracle para Solaris Sparc
Oracle Download – docs/cd/B19306_01
5. Documentación Completa de Sun Cluster 3.2
Sun Microsystems Docs coll/1647.3
6. Sun Cluster para el sistema operativo Solaris: Visión general
Sun Microsystems Docs doc/817-6910
7. Sun Cluster – Servicio de datos para Oracle
Sun Microsystems Docs doc/817-4639
8. Sun Educational Services SA-200.
“Advance System Administration for the Solaris 10 Operating System”
Sun Microsystems.
9. Sun Educational Services ES-388, “Sun Cluster Administration Guide”
Sun Microsystems.
10. Kristien Hens, “Oracle RAC on Sun Cluster 3.x”
Sun Microsystem Belgium – Sun Microsystems, 2005.
11. Kristien Hens, “Designing Enterprise Solutions with Sun Cluster 3.0”
Sun Microsystem.
12. Modelos De Configuraciones Oracle
Oracle Examples – Oracle/technology/obe