73

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

Facultad de Ingeniería

SEGURIDAD E INTEGRIDAD DE DATOS EN UN CENTRO DE COMPUTO E IMPLEMENTACION DE UN SISTEMA DE ALTA DISPONIBILIDAD EN PLATAFORMA HP9000 T500

T E S I S

Que para obtener el título de
INGENIERO EN COMPUTACION
p r e s e n t a n

JORGE PEREZ RETANA
JOSE LUIS ESQUIVEL SERNA
MA. VANESSA GONZALEZ LOPEZ

ASESOR: ING. MA. JAQUELINA LOPEZ BARRIENTOS

México, D. F. 2000

Queremos agradecer:

A nuestra querida Universidad Nacional Autónoma de México por su cotidiana lucha por preservar los valores culturales del hombre y la desinteresada ayuda que nos ha proporcionado.

Al estímulo de superación de nuestros profesores de la Facultad de Ingeniería, nuestro reconocimiento por todo lo que representa su ejemplo, empeño y perseverancia por mantener latente la máxima Casa de Estudios.

A nuestra directora de tesis la Ing. Ma. Jaquelina López Barrientos por su dedicación, guía, estímulo y consejo para ver culminada esta meta.

Agradezco a Dios y a mis padres permitirme la vida y su confianza para poder realizar esta meta, la cuál les entrego en agradecimiento y eterno amor que les tengo.

A mi hija Nayeli, a mis sobrinas Areli y Verónica, a quienes tanto amo, les pido nunca pierdan la fe y la confianza en sí mismas para poder alcanzar las metas que se propongan.

A mi esposa y hermana doy gracias por el cariño y apoyo que en mi vida representan.

A mi abuelita Concepción Villicaña por su apoyo y cariño en todo momento, y más aún, en los momentos difíciles.

A todos aquellos amigos, familiares, compañeros y el bonito recuerdo de los muertos en vida, que fueron partícipes de este logro, gracias.

José Luis Esquivel Serna

Dedico mi tesis a:

Mi Esposa, quien ha sido mi único y verdadero amor de toda la vida y quien siempre ha estado en mi pensamiento. Gracias Nancy porque siempre has estado a mi lado de una u otra manera, y porque siempre que tenía problemas pude contar contigo. Ahora que eres mi esposa, quiero decirte que toda mi vida te amaré, cuidaré y respetaré. Te amo mi vida.

Gracias Nancy.

Mi Mamá, quien me dio la oportunidad de estudiar y me abrió todas las puertas de mi vida. Mamá, te dedico mi tesis, la cual es la culminación de muchos años de estudio. Mamá, sacrificaste muchas cosas por mi y quiero que sepas que toda mi vida estaré agradecido contigo, y además quiero decirte que gracias a ti soy un hombre de bien. Te quiero mucho Mamá.

Gracias Pera.

Mi hermana, quien siempre me ha dado todo su apoyo en todos los aspectos. Paty, quiero que sepas que te quiero mucho y que siempre tendrás mi amor y apoyo. Gracias Patolin por ser una excelente hermana conmigo. Te quiero mucho Paty.

Gracias Paty.

Mi Papá, quien se nos adelanto en el camino. Papá, yo se que estés donde estés siempre nos observaras y nos cuidaras. Quiero decirte que he cumplido y que soy una persona de bien, y que el día que nos volvamos a ver estarás orgulloso de mi. Te extraño mucho.

Gracias Papá.

Jorge Pérez Retana.

Agradezco a Dios, a mis padres por su apoyo en todo momento a lo largo de mi vida.

A mi abuelita Elena Urrea por todo su cariño y cuidados. Este logro se lo dedico a ustedes.

A mi hermano Efraín por enseñarme a ser crítica y por todo lo que he aprendido de él.

A mi novio Juvenal González por su comprensión y amor.

A mis amigos por todo los momementos que hemos tenido juntos.

A todos ustedes gracias.

Ma. Vanessa González López

OBJETIVO

Esta tesis tiene como finalidad el dar a conocer los fundamentos de un sistema de alta disponibilidad, así como también proporcionar una visión general de algunas tecnologías disponibles que permiten tener los centros de cómputo más seguros, confiables y protegidos del entorno que influye en su operación. Finalmente se espera obtener la puesta en marcha de un sistema que cumpla con las características de ser altamente disponible en plataforma HP 9000 T500 y que cumpla aún más allá con las expectativas esperadas de eficiencia y efectividad.

Objetivos específicos.

1) Presentar las bases teóricas de seguridad en sistemas y centros de cómputo, como punto de partida, para llegar a formular un entorno adecuado en la puesta en marcha del Sistema Altamente Disponible.

2) Hacer un análisis detallado de algunas tecnologías de monitoreo de redes, respaldo y recuperación de información, software y hardware para la implementar sistemas de alta disponibilidad, software y hardware de administración jerárquica de almacenamiento, etc., con la finalidad de crear una visión general de lo que puede ser utilizado para incrementar la seguridad y disponibilidad de las aplicaciones críticas que están inmersas en los centros de cómputo.

3) Mostrar las diversas formas en que se hace visible el elemento humano y su repercusión en la operación de los sistemas al presentarse como un intruso, usuario y/o operador del mismo.

4) Recalcar la importancia que tienen el desarrollar a tiempo planes de contingencia adecuados para recuperarnos de las fallas de los sistemas.

5) Detallar la facilidad que representa el tener un adecuado control en las diversas fases de la seguridad en los sistemas y en elementos de la misma tan básicos como lo es, por citar un ejemplo, el correcto resguardo y etiquetación de las cintas de respaldo y recuperación.

6) Presentar el impacto económico que conlleva el hacer uso de las técnicas en este trabajo citadas, en algunos casos las alternativas que podemos encontrar y los resultados que podremos esperar de ellas si son operadas y administradas de un modo ordenado, eficiente y eficaz.

7) Poner en práctica una metodología para la puesta en producción de nuestro Sistema de Alta Disponibilidad que consiste en:

1) Analizar
2) Diseñar
3) Implementar
4) Diseñar un Plan de Pruebas
5) Desarrollar Planes de Mantenimiento
6) Diseño de un Plan de Contingencia
7) Generación de Manuales

Con la finalidad de obtener como resultado Sistemas protegidos que trabajen eficiente y eficazmente.

Entendemos por eficiente el que cumpla con el objetivo de ser Altamente Disponible al asegurar operación continua de los sistemas en un 99.999%. Es decir, nuestra aplicación no estará fuera de funcionamiento por falla no más de 5 minutos al año.

Entendemos por eficaz el que nuestro sistema cumpla con el ser eficiente y además se asegura que está haciéndolo al menor costo posible.

INDICE

Página

Página

Página

Página

Página

CAPITULO 1
PROLOGO

El presente tema de TESIS es elegido por la novedad e importancia que representa hoy en día para las empresas que requieren de fundamentar su estructura y el logro de sus objetivos en sistemas abiertos. De la misma manera, se lleva a cabo una mezcla, análisis y recopilación de información de dos temas, seguridad y alta disponibilidad, que son inherentes debido a la interdependencia que guardan entre sí por su naturaleza.

El concepto de Alta Disponibilidad llega a nuestros días con la creciente variedad de aplicaciones de Misión Crítica corriendo en los sistemas de cómputo de las organizaciones, que ya no aceptan largas caídas" o "tiempos fuera". Por otra, que enfrentan el reto de hacer crecer el número y calidad de clientes / usuarios de las aplicaciones, de la Bases de Datos, de soluciones complejas del tipo de DataWarehouse, mediante trabajos puntuales de consultoría y capacitación, o bien mediante facilidades orientadas a las medianas empresas que pueden utilizar la opción de servicios de suscripción.

En el caso de Alta Disponibilidad (o High Avalaibility) se presenta un panorama general a manera de introducción al tema, para luego revisar con detalle la puesta en marcha de Alta Disponibilidad de un Sistema en clusters mediante la solución de Hewlett Packard llamada Service Guard. Dicha aplicación representa la solución de Hewlett Packard para atender Misión Crítica. Con respecto a seguridad en sistemas, se lleva a cabo una recopilación de las nuevas tecnologías, técnicas, puntos de cuidado, y demás elementos; que son esenciales para proteger nuestros sistemas abiertos de pérdidas de información y que ayudan a la indisponibilidad y disponibilidad de los mismos.

Hewlett Packard habla del entorno en que se presentan las alternativas de Alta Disponibilidad de los sistemas para soportar la toma de decisiones, el cual implica que si alguno de los componentes hardware, sistema operativo, red, aplicaciones se descuida, la solución para la empresa en su totalidad es proclive al colapso, por tanto, la construcción de un sistema efectivo y altamente disponible para soportar la toma de decisiones, requiere cimientos sólidos:

1. Tecnología confiable y adaptable para permitir el máximo desempeño, a pesar de los cambio.
2. Los parámetros de soportabilidad deben estar incluidos en el diseño de la tecnología
3. El sistema debe tener capacidad de autoadministrarse para evitar el impacto de errores humanos
4. La capacitación de nuestros operadores es fundamental en el logro de los objetivos.

La gente involucrada en desarrollar y mantener los sistemas de Misión Crítica en las organizaciones, tiende a identificar la Disponibilidad de los Sistemas de acuerdo a términos porcentuales que representan tiempos de operación continua, visto desde un lado de la moneda y pérdida de clientes, o bien, tiempos fuera o "caídas del sistema" en la otra cara. Sin embargo, este tipo de medición deja fuera de consideración una gran cantidad de factores que inciden en la disponibilidad de los sistemas (los cuales serán tratados en este trabajo en el Capítulo 10, ya que constituyen el interés principal del mismo).

Para tratar de entender claramente a qué se refiere el concepto de Alta Disponibilidad como propuesta y alternativa para el manejo de equipos/soluciones/sistemas en Alta Disponibilidad, más allá del análisis profundo de lo que significan estos conceptos como tendencias del mercado, resulta valioso contar con algunos elementos de definición y clasificación que faciliten la tarea, como los que se muestran a continuación:

Disponibilidad:

- Porcentaje de tiempo que un sistema de hardware está disponible en periodos de operación (medido en tiempo fuera de línea no planeado)
- En sistemas "unclustered" gracias a la arquitectura se puede obtener al menos 99.5% de disponibilidad

Alta Disponibilidad:

- Para soportar tareas de Misión Crítica y Sensitiva. Se fundamenta en el uso de componentes redundantes y sistemas de clusters.
- La disponibilidad esta entre 99.8 y 99.999% (es decir entre 18 horas y 5 minutos fuera de línea al año)

Recuperación de desastres:

- Proceso de restauración de la disponibilidad de los sistemas dentro de un lapso de "tiempo razonable", después de una condición que afecta al centro de cómputo en su totalidad.

Tolerancia al desastre:

- Arquitectura de clusters que permite la recuperación automática de la disponibilidad de los sistemas después de un desastre

Y como apoyo adicional, podemos revisar la siguiente tabla 1.1 que muestra en horas y minutos, al año, y que implica que un sistema pueda ser catalogado como de Misión Crítica y se le pueda incluir en el conjunto de Alta Disponibilidad:

Porcentaje de Disponibilidad	Tiempo al año en que el Sistema no esta Disponible
99.5%	44 horas
99.8%	18 horas
99.9%	9 horas
99.95%	4 horas
99.98%	2 horas
99.99%	1 hora
99.999%	5 minutos

Tabla 1.1 Disponibilidad de los Sistemas

¿Por qué es importante para algunas organización disminuir el "tiempo fuera" o de "caída" hasta llegar a una disponibilidad de los sistemas equivalente al 99.999%?. Visto a través de lo que sucede cuando la no disponibilidad equivale a una hora, tenemos que en ese lapso, la compañía para la cual trabaja el sistema TPS (Sistema de Teleproceso), transfiere y procesa, además del manejo monetario:

Transferencia y proceso de la misma en bytes:
2 Gbytes para los equipos de transferencia MT1 y MT2.

Proceso y tasación de la información transferida y procesada por MT1 y MT2 en MP1 y MP2:
2 Gbytes.

Información procesada y tasada en el transcurso del día por 4 horarios de recolección y tasación continua a lo largo del día por el Sistema TPS:
7Gbytes, que en días críticos puede llegar hasta los 10Gbytes.

Cantidad de información procesada traducida a cantidades monetarias en un día:
Millones de dólares al día.

Con lo anterior podemos observar que para la Compañía el esquema que se le presenta de cómputo continuo, con procesos de Misión Crítica, se ha traducido en una necesidad inmediata.

El nivel de disponibilidad lo debe establecer cada organización, ya que involucra tanto la definición de requerimientos de información, procesamiento y seguridad, como inversiones financieras que varían en función directa del grado de disponibilidad requerido: o mayor disponibilidad, mayor costo.

La evolución de lo que hoy llamamos Alta Disponibilidad se ha dado en los últimos años de acuerdo a lo siguiente figura 1.1

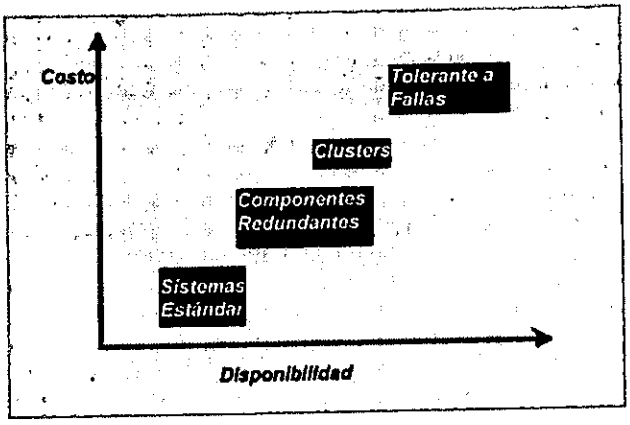


Figura 1.1 Costo vs Disponibilidad de los Sistemas

Para poder incrementar los niveles de disponibilidad, los proveedores de la industria de cómputo han tenido que trabajar en diferentes alternativas, que combinados, están siendo ofrecidas a los clientes como parte de las soluciones integrales, extremo a extremo, que hoy demandan. Entre los planteamientos sobre los que se ha trabajado para incrementar la confiabilidad de los sistemas están:

- La creación de hardware y software altamente disponible
- Minimizar el impacto de una falla
- Minimizar la necesidad de planear "tiempos fuera"
- Minimizar el tiempo de diagnóstico y reparación de fallas.

Hoy en día existen muchas empresas las cuales proponen soluciones integrales para sistemas de Misión Crítica entre las cuales destacan:

HEWLETT PACKARD
COMPAQ
SUN
IBM

Entre otras.

CAPITULO 2
VISION GENERAL DE SEGURIDAD E INTEGRIDAD DE LOS DATOS

La integridad y seguridad de los datos están estrechamente relacionadas por su propósito de proteger los datos de peligros potenciales. En el caso de la integridad de los datos el peligro es, a menudo, un simple error de cálculo, confusiones o errores cometidos por personas, o fallos de equipos que provocan la pérdida de datos, su corrupción o su incorrecta modificación. En relación con la seguridad, la gente puede tratar de infiltrarse de forma intencionada en los sistemas de otras compañías para robar o estropear información en su propio espacio.

2.1 Integridad de los datos

Se define integridad como un estado inalterado y la cualidad o el estado de estar completo o ser indivisible. El objetivo de la integridad de datos es mantener la información de los sistemas en un estado completo e inalterado.

A continuación se examinan algunas de las causas más comunes de pérdida en la integridad de los datos.

2.1.1 Tipos de problemas en la integridad de los datos

Las necesidades de la vida actual obligan a las empresas a dar un mejor servicio, y es por esto que se ven obligadas a utilizar variedades en las marcas de sus equipos, por lo que no existe una estandarización en éstos, ya que algunos equipos utilizan protocolos diferentes, lenguajes de compilación diferentes, etc. Esto origina un caos, es decir, las empresas se ven imposibilitadas de dar un soporte adecuado a los datos.

Las amenazas a la integridad de datos se dividen en:

1. Humanos
   - Inexperiencia
   - Estrés/Pánico
   - Falta de comunicación
   - Venganza
   - Accidentes
   - Avaricia
2. Errores en Hardware
   - Fallos de disco
   - Fallos de los controladores E/S
   - Fallos de energía

- Fallos de memoria
- Fallos en medios, dispositivos
- Mal funcionamiento de los chips y de la placa base

3. Errores de red
   - Fallos en los controladores y en las tarjetas de interfaz de red (NIC)
   - Problemas en componentes de red
   - Problemas de radiación

4. Problema de tipo lógico
   - Errores lógicos
   - Corrupción de Archivos
   - Errores de intercambio
   - Errores de almacenamiento
   - Errores del Sistema Operativo
   - Requisitos mal determinados

5. Contingencia
   - Incendios
   - Inundaciones
   - Tormentas
   - Accidentes Industriales
   - Sabotaje/Terrorismo

Daremos una descripción breve acerca de cada uno de ellos.

1. HUMANOS

Este es el mayor punto débil de los sistemas distribuidos, "la gente que está a cargo del sistema". Pero no sólo los usuarios finales cometen errores, también los administradores de la red o del sistema se equivocan. Ya que existen accidentes que no se pueden prever, otros que son sólo una distracción, hay una infinidad de argumentos que se pueden dar para decir que un sistema ha fallado por un error humano. Veremos los errores humanos más comunes a continuación:

Accidentes: Los accidentes suceden. Ya sea por que no se escuchó bien o por que no se puso la suficiente atención a las indicaciones. La única explicación real para este tipo de accidentes es que se puede haber estado pensando en una cosa mientras se trabajaba en otra.

Inexperiencia: Este tipo de problema está dado por la ansiedad en algunos casos de hacer las cosas que les asignan sin tener el debido conocimiento de éstas.

Estrés, pánico: El estrés de ser el administrador del sistema es agobiante y regularmente va de la mano con la inexperiencia.

Falta de comunicación: Uno de los problemas principales es éste, ya que es vital y en muchas empresas no se logra tener una buena comunicación con los empleados ya sea por jerarquías mal organizadas que no permiten la comunicación directa con el responsable del área que se encuentra en problemas o por encomendar recados que muchas veces no llegan a tiempo o son distorsionados. En la actualidad se cuenta con la facilidad del correo electrónico, pero desgraciadamente no siempre son revisados los mensajes a tiempo, o simplemente no son leídos. No debemos dar por hecho que al interesado le llegó el mensaje a tiempo.

Venganza: Esto es muy común, y se da en muchos casos por el enojo y la ira de empleados despedidos "injustamente", o por algún empleado que aún sigue trabajando en la empresa y simplemente desea vengarse de alguien por este medio.

Avaricia: Es una manera de alterar los datos que generalmente en este caso se refieren a los sueldos.

2. ERRORES DE HARDWARE

Cualquier clase de maquinaria de alto rendimiento sólo puede funcionar durante un cierto tiempo; esto incluye los componentes de las computadoras. Resumiremos algunos de los fallos eléctricos y mecánicos más comunes de las computadoras:

Fallos de disco: Uno de los fallos de procesamiento más comunes son los del disco. Un disco duro es una de las piezas más importantes de un equipo de cómputo y se espera que funcione como un reloj. Sin embargo, no se debe confiar en el tiempo medio entre fallos (MTBF), en cambio, debe acostumbrarse a reemplazar sus dispositivos antes de que sea demasiado tarde. Aparte un disco es relativamente barato, no así, los datos que están almacenados en él. Es por esto que los subsistemas RAID (Serie de Discos Redundantes: Redundante Array of Inexpensive Disks) que tienen mecanismos internos de redundancia para tratar fallos de discos están ganando popularidad.

Fallos de los controladores E/S: Esto ocurre cuando los datos escritos en el disco ya están en mal estado, y no existe ningún proceso que nos permita recuperar los datos originales, verídicos, etc.

Fallos de energía: Existen dos clases de pérdida de energía una es perder la energía de la fuente de alimentación que suministra corriente a la máquina, o bien falla la fuente de alimentación de la misma máquina. En ambos casos la probabilidad de perder datos de forma significativa es muy alta, debido al comportamiento impredecible del sistema cuando le falta la energía. Es buena idea instalar equipos de alimentación ininterrumpida y sistemas

con baterías de reserva en los servidores, que le ayuden a realizar una parada del sistema antes de que se pierda totalmente la energía.

Fallos de memoria: Los circuitos RAM fallan ocasionalmente, si ocurre un error de memoria en un área donde ha sido almacenado un dato, acabará con datos en mal estado y probablemente no se dará cuenta hasta que alguien más note el error en los datos. Los sistemas servidores que incorporan chequeos de paridad de la memoria podrían ayudarle a combatir este tipo de problemas, identificando los segmentos de código incorrectos en la memoria e impidiendo su ejecución, para que el sistema al tratar de ejecutar el segmento en mal estado no se detenga.

Fallos en medios, dispositivos: Los datos almacenados en medios extraíbles para realizar y recuperar copias de seguridad contienen copias de los datos. Cualquier problema con los dispositivos de almacenamiento o los medios que utilizan podrían tener como consecuencia la pérdida de datos si el servidor también estuviera dañado. Este tipo de problemas son muy comunes.

Mal funcionamiento de los chips y de la placa base: Las CPU pueden provocar errores, las placas base pueden fallar, y en general cualquier cosa debido a la manufactura como en el caso del procesador Pentium el cual saco a la luz los puntos débiles de la industria de las PC's en relación con la integridad de los datos.

3. ERRORES DE LA RED

En una red de computadoras las líneas que conectan las máquinas están expuestas a una variedad de riesgos, incluyendo interferencias y averías físicas. Cualquier anomalía en una red origina la pérdida o la corrupción de datos. Analizaremos cada problema que se puede presentar en una red:

Fallos en los controladores y en las tarjetas de interfaz de red (NIC): La tarjeta de interfaz de red y el dispositivo que la controla son virtualmente inseparables. La mayor parte del tiempo, los problemas de las NIC y de los dispositivos no dañan los datos; tan sólo impiden a los usuarios acceder a ella, y no sabremos con certeza qué archivos abiertos pudieron haber sido corrompidos cuando falla la tarjeta NIC en un servidor.

Problemas en componentes de red: La mayoría de las veces los administradores no prueban la fiabilidad y precisión de los componentes de la red bajo condiciones de carga de trabajo impuestas por los sistemas de almacenamiento y recuperación de copias de seguridad. Cualquier punto débil de estos componentes afectará probablemente al sistema de copias de seguridad.

Problemas de radiación: A partir de que lo que sucede en una computadora se fundamenta en el movimiento de los electrones, y puesto que la radiación tienen la capacidad de mover

electrones, se deduce que la radiación y las computadoras pueden combinarse para formar una relación peligrosa, o simplemente datos incorrectos. La mejor estrategia para evitarla es no juntarlos.

4. PROBLEMAS DE TIPO LOGICO

En las siguientes líneas se muestra una visión general de las formas en las que el software puede contribuir a la pérdida de la integridad de datos:

Errores lógicos: Los errores lógicos abarcan un amplio rango de defectos, relacionados generalmente con la lógica de la aplicación. Es difícil tratar de evitar éstos, ya que ningún fabricante puede probar todas las opciones posibles de utilización.

Corrupción de archivos: Los archivos pueden corromperse debido a problemas físicos o de red; problemas del control del sistema o de la lógica de aplicación. Si el archivo corrompido es utilizado por otros procesos para crear datos, los datos resultantes pueden ser incorrectos, ya que para el usuario final no suele ser evidente saber cuáles son los archivos que intervienen en todo el proceso.

Errores de intercambio: El intercambio de archivos entre aplicaciones sucede a menudo, un ejemplo es el de los procesadores de texto que al convertir los datos, ponen en riesgo la integridad de éstos.

Errores de almacenamiento: Este error se da cuando sobrecargamos una máquina, ya que es necesario que el sistema haga toda clase de trabajos extras para acomodarlos y no siempre aunque exista un plan de contingencia y la máquina se pare de manera correcta, probablemente haya archivos que no estén actualizados correctamente.

Errores del sistema operativo: Todos los sistemas operativos tienen su propio conjunto de errores y en algunos casos los datos pueden ser corrompidos. Un ejemplo de los lugares más frustrantes en los que se encuentran los errores es en el código de las interfaces de programación de aplicaciones (API). Una API es utilizada por software de terceros y desarrolladores de hardware para solicitar o suministrar servicios a los usuarios finales.

Requisitos mal definidos: Si los requisitos del software no describen correctamente el trabajo que el usuario necesita realizar, el sistema podría generar datos incorrectos.

5. DESASTRES

No existe nada como la destrucción completa de un edificio o del lugar de trabajo para desafiar la integridad de un sistema. A continuación se examina brevemente cada una de las causas posibles de desastre:

Incendios: El daño que puede ocasionar el fuego, combinado con el humo, el agua y el resto de los residuos resultantes después de un incendio, pueden hacer irrecuperables los datos.

Inundaciones y Tormentas: Muy comunes pero pueden ser igual de devastadoras que un incendio, ya que la pérdida de los datos es la misma. Una tormenta puede demoler un edificio o al menos destruir los servicios de energía y agua.

Accidentes industriales: Un accidente de esta índole, seguramente le haga imposible acceder a su equipo. Ej. Cables cortados por operadores, escape de gases peligrosos, gente de intendencia irresponsable, etc.

Sabotaje/Terrorismo: Desgraciadamente en esta época existen grupos que se dedican a corromper y estropear los datos por venganza o simplemente por gente que piensa que tiene el derecho a estropear o destruir edificios, datos, etc.

2.2 SEGURIDAD DE LOS DATOS

Se define seguridad como la cualidad o el estado de estar libre de daño así como las medidas
de protección tomadas contra el espionaje, el sabotaje, el crimen, el ataque o la fuga.

Las amenazas contra la seguridad de los sistemas es desafiante.

A continuación se examinan algunas de las causas más comunes de amenazas en contra de la seguridad de los datos.

2.2.1 Tipos de problemas de la seguridad de los datos

1. Físicas
    - Robo
    - Dumpster Diving
    - Espionaje
    - ID falsos

2. Basadas en los cables
    - Escuchas
    - Marcación de un número de teléfono
    - Imitación

3. Autentificación
   - ID falsos
   - Suposiciones hechas en algoritmos
   - Edición de contraseñas
   - Captura de contraseñas
   - Averiguación de contraseña

4. Programación
   - Virus
   - Caballos de Troya
   - Cargas y Actualizaciones
   - Códigos Bomba

5. Puertas de escape del sistema
   - Servicios no seguros
   - Configuración e iniciación
   - Inicialización
   - Piggybacking

A continuación se da una descripción breve acerca de cada uno de ellos.

1. FISICA

La seguridad física es un concepto sencillo: no deje conseguir a nadie lo que usted tiene, ni tampoco permita que le espíen.

Robo del equipo: Es muy común ya que es la forma más sencilla de obtener los datos de manera ilícita.

Dumpster Diving: No debemos revolver los disquetes con la basura ni información confidencial y que nos puede ser útil, ya que existen personas que se dedican a hurgar entre la basura para encontrar material costoso.

Espionaje: El espionaje industrial es muy real. Incluso los gobiernos lo hacen de vez en cuando; las organizaciones harán toda clase de acciones inmorales para ahorrar dinero y conocer los secretos de la competencia. El espionaje puede ser tan solo como el que un amigo escriba su password y nosotros lo memorizamos para después entrar y ver que es lo que está almacenado.

ID falsos: Se refiere a la gente que perpetra dichas actividades probablemente sea también bastante seria con sus planes y sepan lo que están buscando, por eso plantean una amenaza significativa a sus datos. Ej: Pasaporte, licencia de conducir, identificación, etc.

3. BASADAS EN LOS CABLES

La utilización de redes de computadoras crea amenazas adicionales de seguridad para sus datos.

Escuchas: La naturaleza del procesamiento distribuido se basa en la comunicación de diversas computadoras a través de un medio. Se deduce que se podría escuchar el tráfico de la sesión y recoger información . En una empresa se puede utilizar cifrado para evitar que sus mensajes sean fácilmente decodificados.

Marcación de un número de teléfono: Cualquiera con un módem y un número de teléfono al que llamar puede intentar acceder a una red a través de su facilidad de marcación remota.

Imitación: En este caso se le llama a la capacidad de una máquina de parecer otra en una red.

3. AUTENTIFICACION

Es el proceso mediante el cual la máquina determina si alguien está autorizado a solicitar o dar ciertos servicios del servidor.

Existen diferentes problemas que se dan con la autentificación:

Captura de contraseñas: Consiste en que alguien escriba y compile un código que tiene el mismo aspecto que su pantalla de presentación al sistema. Este se inserta en la secuencia de introducción al sistema al que se le pide introducirse en él realmente. Todos los usuarios finales ven dos pantallas de presentación, una después de la otra; la primera falla aparentemente, de la forma que se solicita al usuario final que se identifique de nuevo, la pantalla no falló, pero sus datos se escribieron en un archivo que puede ser recuperado posteriormente.

Averiguación de contraseñas: Se trata de adivinar la contraseña de una computadora, y para este tipo de trabajo los profesionales tienen muchas probabilidades de éxito.

Suposiciones hechas en algoritmos: El filtrado de contraseñas funcionan bajo una serie de requisitos que alguien ha codificado en alguna parte, y están basados en algún tipo de algoritmo.

Edición de contraseñas: De forma bastante sencilla, alguien dentro de la compañía establece una cuenta ficticia o cambia la contraseña de una cuenta inactiva. De esta forma, la máquina puede ser accedida por cualquiera que conozca el usuario y la contraseña de dicha cuenta.

4. PROGRAMACION

La mayor parte de las violaciones contra la seguridad provienen del código.

Virus: Un virus es un trozo de programa que se produce a sí mismo, accediendo a otros programas en la máquina y transfiriéndose a otras máquinas cuando el programa es transferido a ellas.

Códigos bomba: La mayor parte de los virus destructivos también funcionan como códigos bomba. La idea de los códigos bomba es que en una determinada hora y fecha, o basados en una secuencia de operaciones de la máquina, éste se disparará realizando su sucio trabajo.

Caballos de Troya: Se le denomina caballo de Troya a un rango de amenazas de códigos malévolos que incluye virus, bombas, gusanos, etc. Este se instala por sí mismo en una máquina y hace el trabajo del programador desconocido.

Actualizaciones y cargas: El instalar una actualización muchas veces resulta peligroso pues no se sabe qué contenga el software y si será del todo compatible con la computadora. Debemos tratar que las actualizaciones sean siempre autorizadas.

5. PUERTAS DE ESCAPE DEL SISTEMA

Conocidas también como puertas traseras, son introducidas en los sistemas operativos para permitir el acceso al sistema en caso de que un cliente pierda toda la información de sus accesos autorizados. Sólo la gente que las descubre conoce el proceso de las puertas traseras e incluso ellos no siempre lo saben.

Resumiremos las diversas amenazas contra la seguridad, planteadas por las puertas traseras:

Piggybacking: Significa llevar a cuestas a alguien, en este contexto se hace referencia a una situación en la que un usuario termina la comunicación con otro sistema; pero el puerto permanece activo en el otro sistema; entonces, otro usuario puede empezar la comunicación con este otro sistema en el mismo puerto sin pasar ningún control de seguridad.

Servicios no seguros: A veces, los servicios de un sistema operativo pueden evitar el sistema de seguridad de la máquina. Como ejemplo recordemos el gusano de Internet, el cual era capaz de pasar servicios en el sistema opertivo UNIX de Berkeley.

Configuración e iniciación: Cuando tenemos que parar a uno de los servidores por cuestiones de mantenimiento, se da el mismo caso que al iniciarlo, ya se han borrado

archivos, debido a los mecanismos de seguridad no se iniciaron correctamente, dejando agujeros de seguridad que son utilizados por otros.

2.3 SOLUCIONES GENERALES A LAS AMENAZAS CONTRA LA INTEGRIDAD Y LA SEGURIDAD DE LOS DATOS

A continuación se explican algunas de las técnicas que pueden ser utilizadas para mantener la integridad de los datos y la seguridad del sistema.

2.3.1 Herramientas para mejorar la integridad de los datos

La siguiente tabla 2.3.1 cataloga las técnicas para recuperar la integridad o la prevención de los datos.

TECNICA	CORRECTIVA/PREVENTIVA
Copias de seguridad	Correctiva
Técnicas en espejo	Preventiva
Archivado	Preventiva
Custodia	Correctiva
HSM	Preventiva
Chequeo de paridad	Preventiva
Plan de contingencias	Correctiva
Análisis predictivo de fallos	Preventiva
Alta disponibilidad	Preventiva
Alimentación ininterrumpida	Preventiva
Implementación de técnicas de seguridad	Preventiva

Tabla 2.3.1 Técnicas de integridad

Cada una de estas técnicas consiste en:

Copias de seguridad: La realización de copias de seguridad es el método más utilizado para restablecer un sistema. Si se pierden los datos, se recupera una copia anterior del sistema a partir de las copias de seguridad.

Técnicas de espejo (Niveles de raid): Son aquellas en las que se copian los datos en un dispositivo o máquina, a otra diferente, según se están escribiendo. Pueden ser realizadas de forma lógica para replicar segmentos del sistema de archivos de una máquina en otra parte de la red. También pueden realizarse estrictamente a un nivel físico mediante dispositivos de disco en espejo, subsistemas de E/S y máquinas enteras.

Archivado: Es el proceso de borrado de archivos del sistema de almacenamiento "online" (en línea) en red. Y su copia en elementos de almacenamiento a largo plazo, en cinta o medios ópticos. Se utiliza para aumentar la protección del sistema de archivos borrando datos del sistema de almacenamiento online y colocándolos en armarios dispuestos para ese fin.

Custodia: Nos referimos a la custodia, como el acto de salva guardar las cintas, discos, etc. en los que se realizaron las copias de seguridad.

HSM: Significa Administración Jerárquica de Almacenamiento (Hierarchical Storage Management), es un sistema automático de almacenamiento y recuperación de archivos entre el sistema de almacenamiento online y un sistema de almacenamiento de datos de uso frecuente.

Chequeo de Paridad: Es una característica de las máquinas servidoras de la gama alta. Suministra un mecanismo de guardia que asegura que fallos de memoria inesperados no tengan como resultado el fallo del servidor o pérdida de la integridad de los datos.

Plan de contingencias: Un plan de recuperación de información después de desastres es como una guía para reconstruir su sistema desde cero.

Análisis predictivo de fallos: Es muy difícil darse cuenta de que un dispositivo está fallando, los dispositivos de disco están siendo desarrollados para indicar que están empezando a fallar cuando esto ocurre.

Alimentación ininterrumpida: Se trata de suministrar las baterías de reserva en caso de pérdida de energía, también dan un voltaje consistente y sin fluctuaciones a la máquina, así, como evitar las variaciones de carga.

2.3.2 Herramientas para reducir las amenazas contra la seguridad

La siguiente tabla 2.3.2 cataloga las herramientas para conseguir que un sistema tenga un nivel de seguridad adecuado. La tabla explica del lado izquierdo la recomendación y del lado derecho si puede ser implementada por el sistema o si es una política personal que debe ser comunicada a los empleados.

RECOMENDACION	SISTEMA/POLITICA
Eliminación de las puertas traseras del sistema	Sistema
Chequeo de virus	Sistema
Seguridad física	Política
Política de máquinas desatendidas	Política
Política de eliminación de basura	Política
Política de contraseñas	Política
Cifrado	Sistema
Obligación de identificación	Sistema, también podrá ser práctica
Cortafuegos para el acceso a Internet	Sistema
Trampas para intrusos	Sistema

Tabla 2.3.2 Recomendaciones e implementación

A continuación se explica cada una de estas recomendaciones:

Eliminación de las puertas traseras del sistema: Como se explicó anteriormente es mejor cerrar una puerta trasera, a que alguien esté enterado de que existe, y tenga acceso al sistema.

Chequeo de virus: Se puede aplicar una estrategia que incorpore múltiples sistemas de protección y actualización periódicamente contra virus. Es importante poder detectarlos y aún mejor poder prevenirlos.

Seguridad física: Los equipos que se encuentran cerrados con llave en lugares a los que no puede acceder la mayor parte de la gente son más seguros, desde el punto de vista de las amenazas contra la seguridad.

Política de máquinas desatendidas: Se deberían apagar las máquinas en la noche y los fines de semana, así como acostumbrar a los empleados a poner contraseñas en los protectores de pantalla y en los teclados.

Política de eliminación de basura: Se debe triturar la basura en el caso de documentos importantes o confidenciales, en el formato electrónico debemos hacer lo mismo.

Política de contraseñas: Se deben cambiar las contraseñas de forma periódica, no deben ser reutilizadas ni basadas en cosas como apellidos o números de teléfono para evitar que sean descifradas por extraños.

Cifrado: El cifrado revuelve los datos de forma que no pueden ser utilizados, a no ser que sean primero descifrados. El punto débil de todos los esquemas de cifrado es la utilización de algoritmos que pueden ser finalmente decodificados por otra persona.

Obligación de identificación: La identificación que asegura la validez de la persona o el programa en el otro extremo de la sesión es extremadamente importante para muchas organizaciones.

Cortafuegos para el acceso a Internet: Si la empresa da acceso a Internet debe tener instalado cortafuegos, para evitar que los piratas informáticos tengan conocimiento de los detalles de su sistema.

Trampas para intrusos: La idea de tener trampas para intrusos es el de saber quien está tratando de entrar al sistema, así como también saber quiénes son, qué productos están utilizando y desde dónde están trabajando.

CAPITULO 3
SISTEMAS DE COPIAS DE SEGURIDAD DE RED

3.1 CONFIGURACIONES DE SISTEMAS DE COPIAS DE SEGURIDAD

Antes de analizar los sistemas de copias de seguridad de intranet, hay que identificar y familiarizarse con cuatro componentes básicos de la red.

a) **Destino**. Un destino es un sistema cuya información está siendo almacenada o recuperada de una copia de seguridad.
b) **Motor**. El motor es el sistema que realiza las tareas operativas de las copias de seguridad, como copiar los datos del destino a una cinta.
c) **Dispositivo**. El término dispositivo representa el dispositivo de almacenamiento que escribe los datos en los medios, generalmente cintas.
d) **Bus SCSI**. El bus SCSI son los cables y conectores físicos y eléctricos que unen los dispositivos a la red de computadoras. En una intranet, el bus SCSI conecta generalmente los dispositivos al sistema que realiza la función de motor.

A continuación se presentan las cuatro configuraciones principales de los sistemas de copias de seguridad:

Configuración autónoma.
El sistema de copias de seguridad más simple es el que combina los cuatro componentes de los sistemas de copias de seguridad en una sola máquina. La figura 3.1 muestra esta configuración.

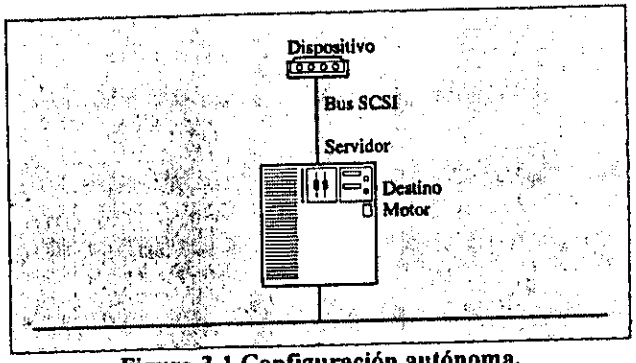


Figura 3.1 Configuración autónoma.

Configuración basada en una estación de trabajo.
Una variación de la configuración autónoma traslada el motor, el bus SCSI y el dispositivo a una estación de trabajo dedicada. La figura 3.2 muestra este tipo de configuración.

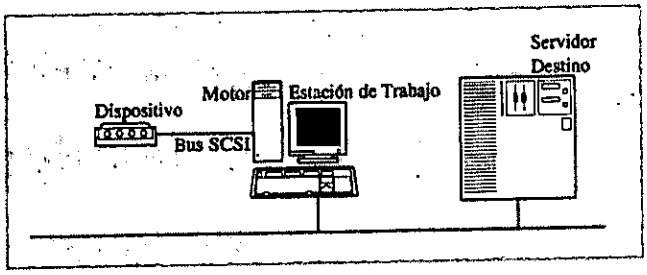


Figura 3.2 Configuración basada en una estación de trabajo.

Configuración servidor a servidor.
La figura 3.3 muestra un sistema para realizar copias de seguridad de servidor a servidor. Esta configuración es la unión de las configuraciones autónoma y la basada en una estación de trabajo, pues el servidor B realiza sus propias copias de seguridad en el dispositivo que tiene conectado (autónoma) y también realiza las copias de seguridad de los servidores A y C (basada en una estación de trabajo).

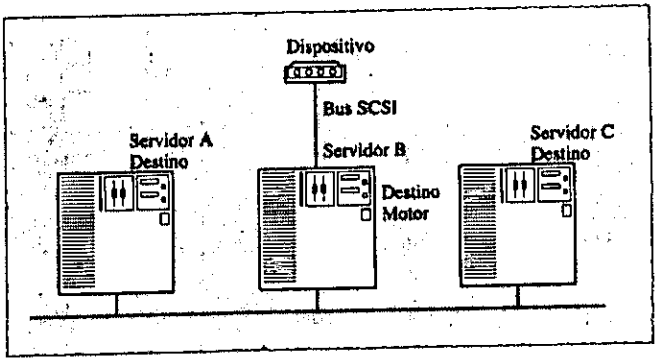


Figura 3.3 Configuración servidor a servidor.

Configuración servidor dedicado.
Debido a problemas que pueden ocurrir durante la realización de las copias de seguridad en los servidores de producción, algunas organizaciones sitúan el motor, el bus SCSI, y el dispositivo en sistemas servidores dedicados. Esta configuración está muy relacionado con la configuración basada en una estación de trabajo, excepto que la estación de trabajo es sustituida por un servidor por motivos de eficiencia y compatibilidad. La figura 3.4 muestra el diagrama de esta configuración.

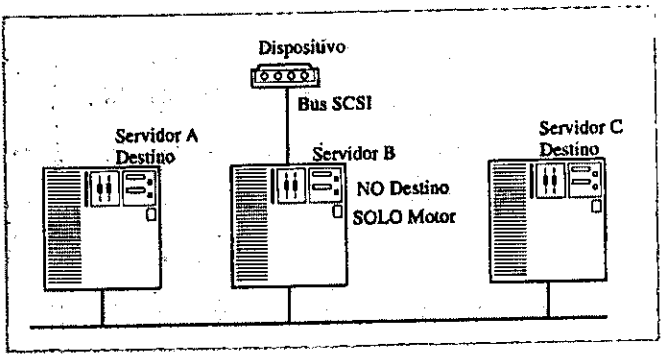


Figura 3.4 Configuración servidor dedicado.

3.2 LAS COPIAS DE SEGURIDAD COMO SISTEMA

Debido a la consideración de que las copias de seguridad son vistas como un sistema, entonces esto implica que se debe conocer cómo está compuesto dicho sistema. En la tabla 3.1 se muestran las diferentes partes de un sistema genérico de copias de seguridad, junto con una breve explicación del papel que desempeña cada una.

Componente	Descripción
Sistema host físico	Máquina donde se ejecuta la lógica principal de las copias de seguridad.
Sistema host lógico	Sistema operativo que está por encima.
Bus de E/S	Bus interno de la máquina, también pueden ser buses externos, como SCSI.
Dispositivos periféricos	Dispositivos de cinta, de disco, y ópticos.
Software del controlador del dispositivo	Código de bajo nivel asociado al dispositivo.
Medio de almacenamiento de las copias de seguridad	Cinta magnética, fuentes ópticas, etc.
Planificador de operaciones	Determina qué es lo que hay que hacer cada día que se realizan copias de seguridad.
Motor de operaciones	Código que realiza las copias de seguridad.
Sistema destino físico	Máquina de la que se extraen datos que se copian.
Sistema destino lógico	El sistema operativo y el entorno de dicha máquina.
Componentes de red	Ruteadores, puentes, conmutadores y cableado.
Protocolo(s) de red	Protocolos de transporte: IPX/SPX, TCP/IP, etc.
Metadatos del sistema	Base de conocimiento con información sobre los archivos contenidos en las copias de seguridad.
Consola del sistema	Interfaz con el administrador del sistema.
Administración de sistemas	SNMP u otras formas de administrar el sistema.

Tabla 3.1 Componentes de un sistema de copias de seguridad.

A continuación, se tratará con más detalle cada componente del sistema de copias de seguridad.

Sistema host físico.
Es el sistema donde reside el cerebro de las aplicaciones que realizan las copias de seguridad. Éste podría ser una PC con un procesador Pentium a 100 Mhz y un bus PCI, una estación de trabajo UNIX, un POWER PC, un Apple Macintosh, una máquina con un procesador Alpha de Digital, o cualquier otro hardware desde el que se puedan realizar copias de seguridad. Todas estas máquinas tienen CPU diferentes, con diferentes rendimientos y diferentes buses de E/S, los cuales se definen en base a las necesidades del sistema (por ejemplo, el volumen de la información que se maneja). El rendimiento de las operaciones de copias de seguridad puede estar incluso limitado por cuellos de botella inherentes a la propia máquina.

Sistema host lógico.
Es en el sistema operativo donde reside el núcleo del sistema que realiza las copias de seguridad. Debido a que el sistema operativo suministra la funcionalidad de E/S de acuerdo con su arquitectura interna, el sistema operativo puede convertirse en un factor que limite el rendimiento de las copias de seguridad.

Bus de E/S.
El bus de E/S tiene dos partes: la primera es el bus interno del sistema, utilizado por la máquina para transferir datos, y la segunda es el bus externo, utilizado a menudo para conectar los dispositivos de almacenamiento. Aunque hoy en día se utilizan diferentes tipos de buses, el más comunmente utilizado es el SCSI. Hay diversas implementaciones SCSI que es importante conocer, ya que son útiles para comprender mejor qué son y cómo podrían utilizarse. La tabla 3.2 muestra las distintas especificaciones SCSI.

Bus SCSI	Velocidad (MBps)	Número de dispositivos que pueden ser conectados
SCSI normal, transferencia de 8 bits	5	7
Fast SCSI (utilza protocolo mejorado)	10	7
Wide SCSI, transferencia de 16 bists	10	15
Fast-Wide SCSI	20	15
Ultra SCSI, 8 bits	20	15
Ultra SCSI, 16 bits	40	15

Tabla 3.2 Comparación de especificaciones SCSI.

Lo que es importante destacar de las velocidades SCSI es que la mayor parte de ellas exceden la velocidad del bus del sistema, esto es importante, pues se evitan cuellos de botella.

Se pueden conectar varios dispositivos SCSI a un único adaptador SCSI utilizando la técnica llamada "en cadena". La figura 3.5 muestra una cadena de adaptadores SCSI en un sistema servidor.

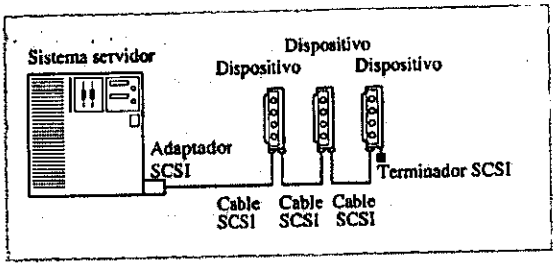


Figura 3.5 Dispositivos SCSI conectados en cadena.

Dispositivos periféricos.
Los dispositivos periféricos son las unidades de disco, unidades de cinta, los dispositivos ópticos, los sistemas RAID (en general, cualquier dispositivo que pueda leer y escribir datos). Cualquiera de estos dispositivos es más lento que el bus del sistema, y ninguno puede utilizar completamente el bus SCSI.
Las limitaciones de rendimiento de los dispositivos de almacenamiento son generalmente físicas. Cuanto más rápido se transfieren los datos más allá de la cabeza de grabación, mayor será el rendimiento. En los discos, el rendimiento está limitado por su velocidad de rotación; por este motivo, los dispositivos con altas RPM (Revoluciones Por Minuto) son más rápidos que los de bajas RPM. En los dispositivos de cinta, la limitación es la rapidez con la que la cinta pasa delante de las cabezas de grabación mientras se depositan en ellas los datos. Esto se conoce como el flujo de datos en la cinta.

Software del controlador del dispositivo.
El conjunto de todos los fabricantes de circuitos integrados SCSI utiliza controladores de dispositivo programables mediante los que se controla el hardware a través de algún tipo de interfaz de aplicación programable.
La figura 3.6. muestra la relación entre el software de copias de seguridad, los controladores de dispositivo, los controladores SCSI y los medios de grabación.

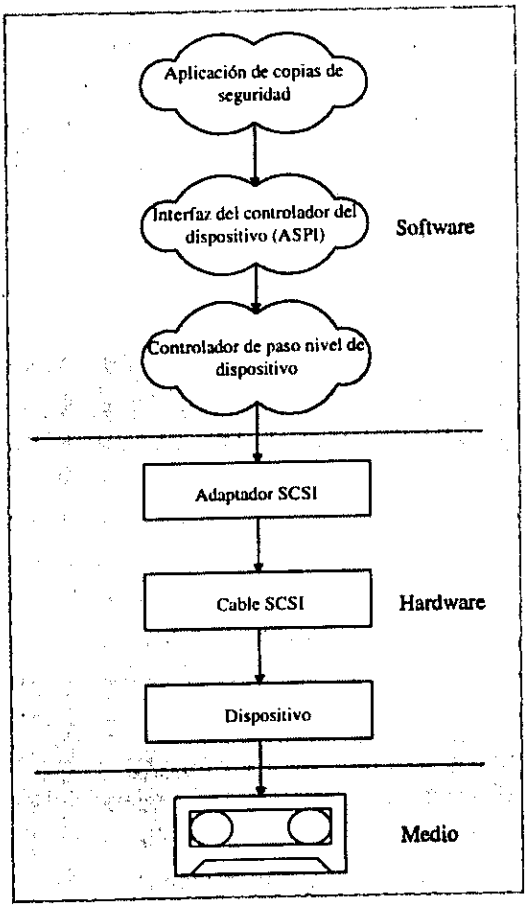


Figura 3.6 Subsistema de un dispositivo en sistemas de copias de seguridad.

Medios de almacenamiento de copias de seguridad.
Los medios y los dispositivos que escriben en ellos son virtualmente inseparables. Lógicamente, los dispositivos realizan más funciones y cuestan mucho más dinero que los medios. Por esta razón, se desarrollan productos de copias de seguridad que generalmente enfocan su atención en los dispositivos, sin prestar mucha atención a los medios.

Planificador de operaciones.
Esta es la parte del sistema de copias de seguridad que determina, en función de una serie de condiciones, cúales son los datos que tienen que ser copiados al medio de

almacenamiento. Estas condiciones incluyen generalmente los días de la semana, los ciclos de negocio tales como cierres de contabilidad mensuales, y otros datos cíclicos o basados en un calendario. Algunos sistemas que realizan copias de seguridad dan mucha flexibilidad en cuanto a la planificación y la automatización de las operaciones que deben ser ejecutadas.

Motor de copias de seguridad.
Es el software que la mayoría de la gente asocia con el motor de almacenamiento y recuperación de copias de seguridad. Es el programa responsable de gran parte del trabajo que se realiza con las copias de seguridad. Además es el núcleo de la mayor parte de las compañías que escriben software de copias de seguridad. Errores en esta parte del software pueden provocar operaciones poco eficientes o muy serias pesadillas durante las operaciones de recuperación.

Sistema físico destino.
Es la máquina donde residen los datos que se desean almacenar o recuperar de una copia de seguridad. La palabra físico hace referencia a la plataforma hardware real. Al igual que la plataforma hardware del host del sistema que realiza las copias de seguridad, la plataforma hardware del sistema destino puede tener un impacto significativo en el rendimiento global de las tareas que se ejecutan durante la realización de copias de seguridad.

Sistema lógico destino o agente.
Naturalmente, el sistema físico destino tiene un sistema operativo y unas aplicaciones que se ejecutan en él. Sin embargo, para el propósito de realizar copias de seguridad, el componente lógico principal del sistema destino es un agente que responde a las operaciones del ejecutor de operaciones (motor) descrito previamente. Este agente es el responsable de suministrar los archivos y otros datos del sistema a través de canales que son leídos por el motor que realiza las copias de seguridad. Debe ser capaz de manejar los detalles del sistema de archivos destino y otros datos que no residen dentro de dicho sistema de archivos.
Debido a los requisitos de Alta Disponibilidad de las bases de datos y otras aplicaciones de sistemas, puede que no sea posible cerrar dichos archivos mientras se hacen sus copias de seguridad. Esto introduce el requisito de tener agentes destino especiales que puedan manejar las dificultades específicas de dichos sistemas.

Componentes de red.
Los componentes de red son el conjunto de elementos que transportan el tráfico: ruteadores, puentes, concentradores, conmutadores, cableado y cualquier otra cosa que exista entre las computadoras de una red. Cuando se realizan copias de seguridad, es común descubrir los puntos débiles de los componentes de red. Un escenario común que suele producirse es cuando un dispositivo de red es rebasado con archivos y pierde paquetes. Cuando se pierden paquetes, pueden suceder muchos problemas, incluyendo la corrupción de archivos, destinos inexistentes e incluso el fallo del sistema de copias de seguridad. Por esta razón, es necesario conocer de forma realista la carga que se genera en la red durante la realización de las copias de seguridad antes de hacer grandes inversiones en los componentes de red.
Protocolos de red.

Esta es una de las mayores dificultades que se deben afrontar durante la realización de copias de seguridad de intranet. Lo mismo que la realización de copias de seguridad pone a prueba los componentes de red, también pone a prueba las torres de protocolos, debido al enorme volumen de tráfico generado. Puede que con el tráfico que se genera diariamente no se descubran ineficiencias en las torres de protocolos; pero la realización de copias de seguridad tendrá como resultado una disminución del rendimiento, e incluso podrá causar la desconexión o el fallo de las sesiones de comunicación, lo que provocaría comportamientos impredecibles en el sistema de copias de seguridad.

Metadatos del sistema.
Los metadatos del sistema son la base de datos que mantiene todos los registros detallados de qué archivos han sido llevados a copias de seguridad en qué dispositivo, cuándo han sido copiados, cuáles eran los atributos del sistema de archivos y cualquier otra cosa que el responsable del motor de copias de seguridad piensa que es importante.
Si se considera el enorme número de transacciones que ocurren durante la realización de copias de seguridad, es importante que el sistema de metadatos sea configurado de forma que permita su rápida actualización, y que disponga de alguna forma de controlar el tamaño de sus archivos. El diseño de una aplicación de seguridad que pase por alto esta parte puede dar lugar a graves problemas, ya que este subsistema es el núcleo de las operaciones realizadas por el sistema de copias de seguridad.

Consola del sistema.
El sistema de copias de seguridad tiene en algún lugar una consola desde la que se puede observar y operar sobre él. Generalmente, ésta se desarrolla como máquina/cliente en una plataforma con una interfaz gráfica de usuario, como en procesamiento cliente/servidor, mientras que el sistema que tiene conectados los dispositivos es el servidor.
La figura 3.7 muestra la consola de un sistema de copias de seguridad que se ejecuta en una estación de trabajo para controlar y supervisar el motor residente en el servidor B. (Obsérvese que la consola no necesita ser un destino de las copias de seguridad).

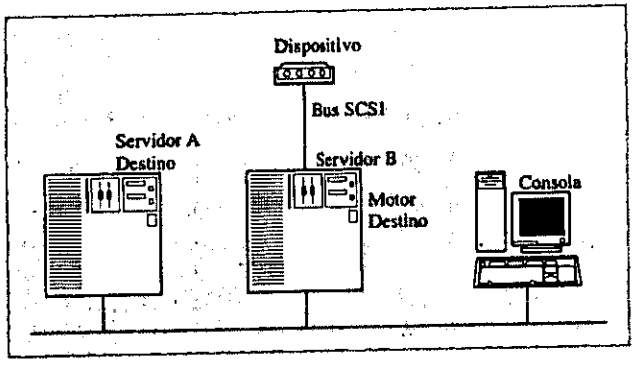


Figura 3.7 Consola del sistema de copias de seguridad.

Administración de sistemas.
A medida que crece el número de máquinas y la capacidad de almacenamiento de los sistemas de red, cada vez es más importante ser capaces de ver lo más rápidamente posible el estado de muchos sistemas de copias de seguridad de la red. Esto podría realizarse mediante un procedimiento propietario que examinará las instalaciones de sistemas de copias de seguridad, suministrando información detallada, o podría realizarse desde la propia consola de administración de red, probablemente con Protocolo de administración de red simplificada (Simple Network Management Protocol, SNMP), para indicar cualquier alarma o problema. Mientras que SNMP parece la elección lógica, no hay Base de información de administración (Management Information Base, MIB) estándar para sistemas de copias de seguridad. Sin una MIB estándar, los mensajes y avisos de diferentes sistemas de copias de seguridad serán inconsistentes.

3.3 TECNOLOGIAS DE MEDIOS DE ALMACENAMIENTO

A continuación se dan los conocimientos básicos sobre las tecnologías utilizadas en los medios de almacenamiento.

Tecnología QIC.
QIC soporta cartuchos de un cuarto de pulgada. Es el mismo ancho de cinta que ha sido utilizado durante años para grabaciones de audio. La tecnología es bien conocida y es estable. Este medio ha sido visto como una solución de bajas prestaciones para sistemas de copias de seguridad autónomos y, por lo tanto, no es aplicable a sistemas intranet porque su funcionalidad y velocidades son muy bajas. Recientemente, QIC ha incorporado dispositivos de 5GB de capacidad, y tiene un dispositivo de 13 GB.

Tecnología de 4 mm.
Introducidas inicialmente como cintas digitales de audio, o DAT, Hewlett-Packard y Sony dirigieron el desarrollo de una cinta DAT estándar que fuera utilizada para almacenamiento de datos. Otras compañías se unieron a los esfuerzos de estandarización de DAT y fabricaron equipos para este tipo de cintas. Dichas cintas se conocieron por el acrónimo almacenamiento digital de datos (Digital Data Storage, DDS).
Las cintas DDS originales tenían 60 metros de longitud y almacenaban 1.3 GB de datos sin comprimir. Después fueron introducidas cintas de 90 metros (DDS I), aumentando su capacidad a 2 GB. Añadiendo la compresión, estas cintas llegaban a una capacidad de 4 GB. La capacidad inicial de la cintas DDS II (120 metros) es de 4 GB, con una capacidad después de compresión de alrededor de 8GB. La tecnología DDS III ya está en el mercado y tiene una capacidad inicial de 8 GB, sin compresión.

Tecnología de 8 mm.
La capacidad de las cintas de 8 mm era originalmente de 2.2. GB, teniendo en la actualidad una capacidad inicial de 7 GB, y de 14 GB las cintas de longitud extendida (160 metros) y compresión.

Cinta lineal digital.
DLT (Digital Linear Tape) fue desarrollado Digital Equipment Corporation para ser utilizado en los sistemas VAX de gama media. La cinta tiene media pulgada de ancho y está contenida en un cartucho robusto que puede resistir grandes presiones y vibraciones
Los cartuchos de cintas DLT tienen una única bobina, la otra bobina se sitúa dentro del mecanismo de la unidad de cinta. Esto suministra un nivel de protección extra para la cinta porque ninguna de sus partes queda al descubierto cuando está fuera del dispositivo. La desventaja de esta tecnología es que las cintas DLT tardan generalmente algo más de tiempo en cargarse que otras tecnologías.
El rendimiento y la capacidad de las cintas DLT son excelentes. El dispositivo DLT 2000 escribe 10 GB sin comprimir, y 20 GB comprimidos en un único cartucho. La DLT 4000 tiene una capacidad inicial de 20 GB, y 40 GB con compresión.

3480/3490
Las cintas 3480/3490 son un medio utilizado en los dispositivos de alta velocidad del mundo de los sistemas basados en una computadora central. Históricamente, estas cintas han tenido velocidades de transferencia extremadamente altas y capacidad relativamente pequeña, del orden de los 250 a 500 MB; el cartucho de cinta mas reciente de esta familia, el 3490-5, ha tardado en ganar aceptación y tiene una capacidad inicial relativamente pequeña de 1 GB.

Medios ópticos.
En general, esta tecnología es mucho más robusta que las cintas por que el medio es estático, lo que significa que no se mueve a través de los mecanismos de transporte, ni se enrolla en carretes a gran velocidad. Además, el medio no genera su propia suciedad. En el lado negativo, los desafíos de escribir en medio ópticos son mucho mayores que los involucrados en las cintas.

Medios magneto-ópticos.
Los cartuchos magneto-ópticos (MO) tienen la mayor longevidad y resistencia al desgaste y a los tirones de todos los medios disponibles. A diferencia de las cintas no se desprenden partículas de materia en su superficie. Sin embargo la capacidad de los MO no es igual a la capacidad de las cintas, y por esta razón no es ampliamente utilizado como medio de almacenamiento de copias de seguridad. Los MO comenzaron como discos de 650 MB en medios de 5.25"; estos discos de 5.25" tienen 2 caras grabables, y en cada cara se pueden grabar 1.3 GB dando como capacidad final 2.6 GB.

CD grabables.
Muy seguramente, los medios de CD grabables (CD-R) se convertirán en los medios ópticos más utilizados en un futuro cercano debido a su costo y familiaridad.
La capacidad de los CD grabables es de 5.2 GB.

3.4 RENDIMIENTO EN LOS DISPOSITIVOS

Cuando se tienen que almacenar muchos datos en copias de seguridad, el rendimiento es importante. Esta sección mostrará algunas de las tecnologías que pueden aplicarse para conseguir aumentar el rendimiento de las copias de seguridad.

RAID

El concepto de cinta RAID es similar al disco RAID, ya que los datos son "repartidos" entre múltiples dispositivos de cinta, consiguiendo de este modo velocidades de transmisión extremadamente rápidas.

Los controladores RAID envían un bloque de datos a la unidad de cinta, que inmediatamente los almacena en un buffer de memoria; esta es una operación muy rápida para la mayoría de los dispositivos. Los controladores de cintas RAID se mueven de dispositivo en dispositivo, llenando los buffers de cada uno de ellos en forma secuencial. Las unidades vacían completamente sus buffers en el medio, después de recibir una transferencia, y se quedan a la espera de la siguiente transferencia de datos. La figura 3.8 muestra los datos que son repartidos entre cuatro dispositivos hacia sus cuatro cintas.

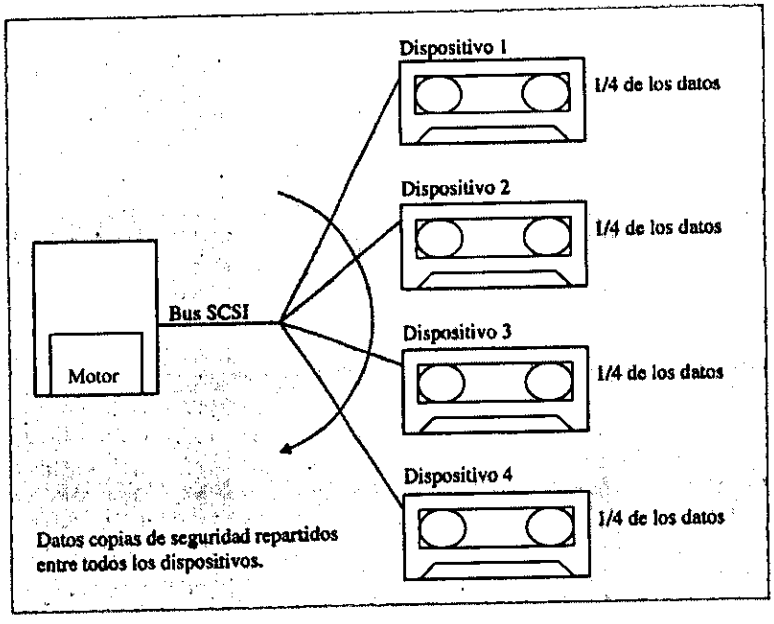


Figura 3.8 Muestra del funcionamiento RAID.

Uno de los peligros potenciales de esta solución podría ocurrir cuando la unidad vacía sus buffers y tiene que esperar que se llenen de nuevo. Esto provocaría que la unidad dejase de mover la cinta, lo que tendría como resultado una baja en el rendimiento, ya que la cinta estaría parando y arrancando continuamente. También hay dudas sobre el rendimiento de las operaciones de recuperación, debido a que es necesario situar de forma precisa en el

tiempo múltiples dispositivos de cinta. Dada la cantidad de variables involucradas en la grabación de datos, esto puede ser una tarea difícil.
Esta tecnología es prometedora en aquellas situaciones en las que sea necesaria una gran velocidad y capacidad, pues actualmente se está trabajando a velocidades de transferencia mayores a 100 MB por segundo con volúmenes de información del orden de terabytes.

DISPOSITIVOS A PLENO RENDIMIENTO.

Una unidad de cinta está a pleno rendimiento cuando la cinta se mueve a una velocidad directamente proporcional a la velocidad a la que se escribe o se lee de ella. De esto se desprende que las unidades de cinta necesitan estar a pleno funcionamiento para conseguir el máximo rendimiento. Para obtener el máximo rendimiento de un sistema de cintas RAID, implicara que todas las unidades del subsistema RAID deben estar a pleno funcionamiento.
Para mantener el dispositivo a pleno funcionamiento, el adaptador del host SCSI debe estar llenando continuamente con datos los buffers de los dispositivos. Para que los adaptadores SCSI de los host hagan esto, las aplicaciones de copias de seguridad les deben suministrar los datos suficientes. Desgraciadamente, la capacidad de transferencia de la mayoría de las intranet es incapaz de suministrar datos a las aplicaciones de copias de seguridad con la suficiente rapidez como para mantener llenos los buffers de los dispositivos, lo que significa que es difícil mantener los dispositivos a pleno rendimiento cuando el sistema está realizando copias de seguridad de otros sistemas de la intranet.

ENTRELAZADO DE CINTAS.

Una forma de resolver el problema descrito anteriormente es utilizando una técnica llamada entrelazado de cintas. El entrelazado de cintas combina los datos de varios sistemas destino en una única unidad y en una única cinta.
En esencia, el entrelazado trenza los datos en la cinta. La figura 3.9 muestra cómo se entrelazan los datos procedentes de tres sistemas destino. De forma individual, la transferencia de datos de cualquiera de los sistemas destino no sería capaz de mantener la unidad de cinta a pleno funcionamiento, pero en cambio sí se logra el pleno funcionamiento trenzando en forma conjunta a los tres sistemas destino.
Esta técnica permite que la unidad de cinta funcione a pleno rendimiento, incluso cuando se realizan copias de seguridad de destinos lentos a través de la red.

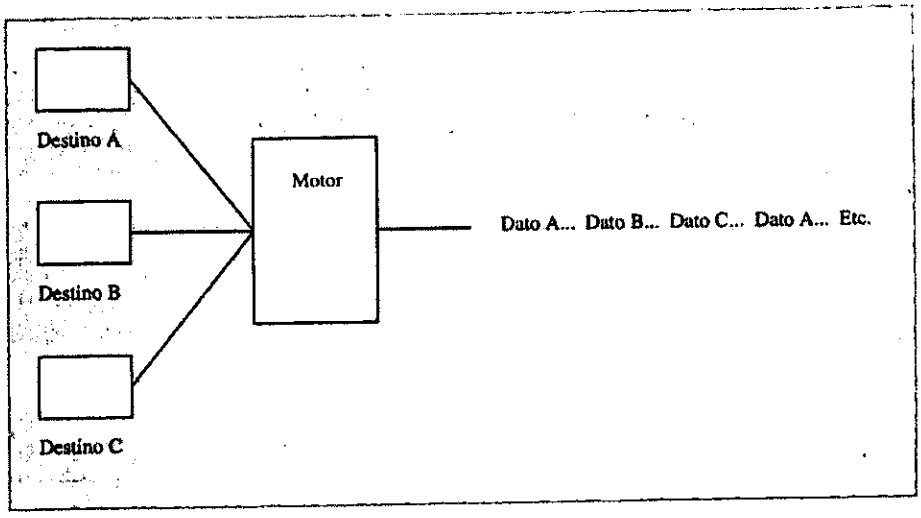


Figura 3.9 Entrelazado de cintas.

CONTROLADORES SCSI EN PARALELO.

Si realmente se quiere sacar el máximo rendimiento a los dispositivos SCSI, se debe limitar su número a menos de cuatro por cada adaptador SCSI (con la finalidad de evitar conflictos de prioridades y cuellos de botella). En grandes sistemas de copias de seguridad puede tener un impacto significativo.

RENDIMIENTO DE LA RED.

Se puede construir una intranet que ofrezca un alto ancho de banda de transmisión que permita que los dispositivos hagan fluir los datos a través de redes diferentes. Algunas organizaciones incluso instalan troncales especiales de alta velocidad, utilizados únicamente con fines de realizar copias de seguridad. Empleando una troncal dedicada, será muy difícil encontrarse con los sorprendentes problemas de degradación de otros sistemas o aplicaciones.
Algunas de las soluciones que se podrían tomar son: cualquiera de las redes con velocidades superiores a 100 MB, como FDDI, 100baseT, ATM ó 100VG AnyLAN; pero incluso pueden saturarse cuando hay suficiente tráfico.

3.5 DISPOSITIVOS AUTOMATICOS

Puesto que las copias de seguridad se suelen realizar a medianoche, puede ser deseable utilizar equipos automáticos de cambio de cintas que permitan completar correctamente la realización de las copias de seguridad sin la intervención humana. En este punto se deben

de considerar dispositivos como apiladores, jukebox, bibliotecas y autocargadores. Generalmente estos dispositivos se conectan al host a través de un único cable SCSI. La figura 3.10 muestra las diferencias entre estos dispositivos.

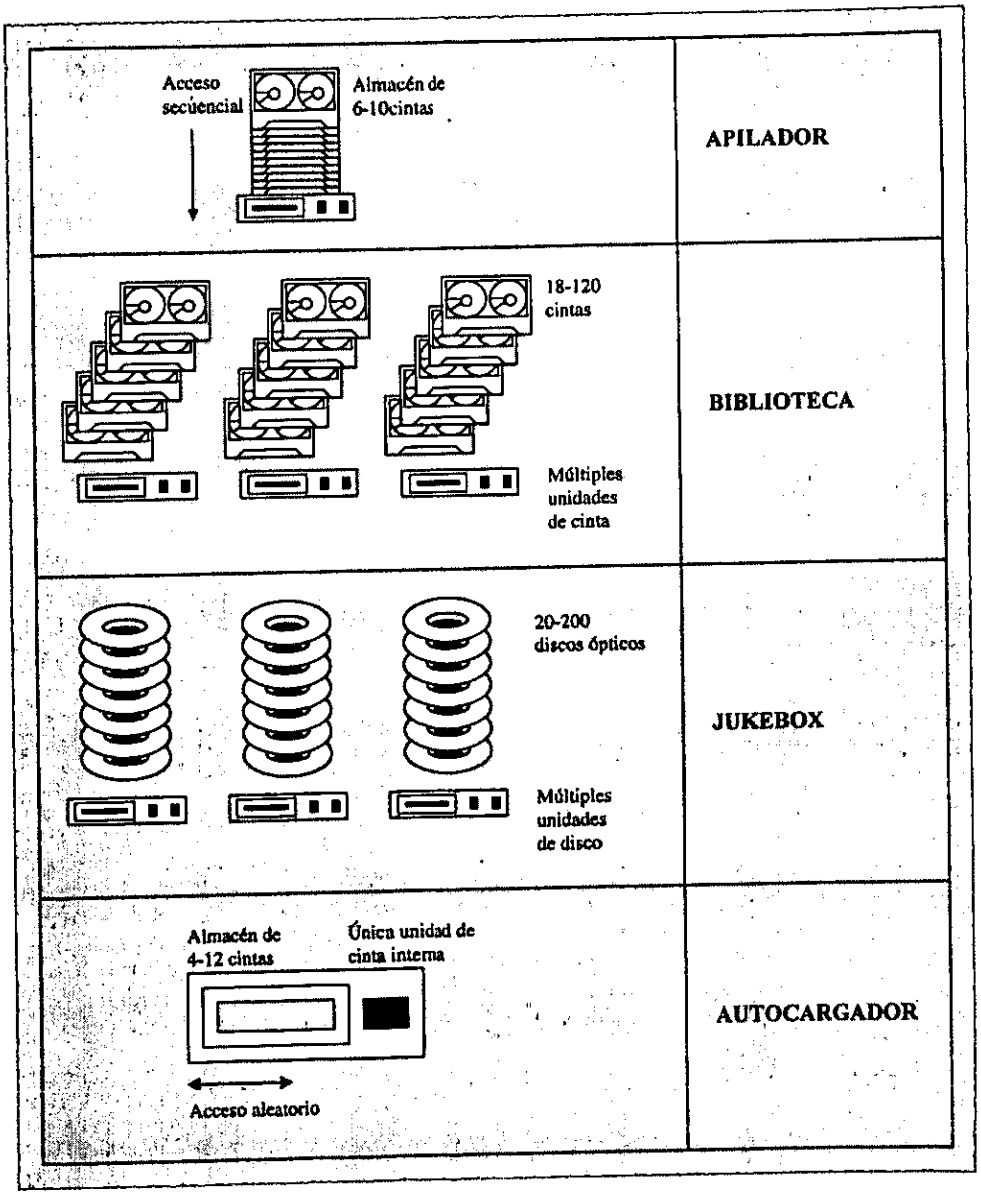


Figura 3.10 Comparación de diversos componentes automáticos.

APILADORES DE CINTAS.

Un apilador o cambiador es un dispositivo de acceso secuencial que utiliza las cintas en el orden en el que se han cargado. La principal ventaja de este sistema es que las operaciones que exceden la capacidad de un medio pueden continuar con el siguiente. Esto, obviamente, supone que las cintas del apilador pueden ser escritas de nuevo durante la siguiente copia de seguridad. Los apiladores no son muy compasivos si se ha insertado la cinta equivocada o si no están en el orden correcto; sin embargo, funcionan mejor cuando se utilizan para realizar copias de seguridad de un único sistema, donde se pueden producir pocas confusiones con las cintas.

Generalmente, los apiladores son también sistemas con una única unidad. Esto significa que las ventajas de rendimiento del paralelismo, tratadas un poco más adelante en este capítulo en el apartado "Técnicas de rendimiento software", no pueden ser utilizadas.

BIBLIOTECAS Y AUTOCARGADORES.

Las bibliotecas y los autocargadores, generalmente, son dispositivos más sofisticados que los apiladores. Un sistema de biblioteca es un dispositivo de acceso aleatorio que permite seleccionar y cargar cualquier medio en una unidad disponible. Es común que esta clase de máquinas tengan múltiples dispositivos que le permitan realizar operaciones en paralelo o de forma concurrente. La ventaja de esto es que un dispositivo podría estar escribiendo datos en un medio mientras que otro escribe, o posiblemente lee datos, de otro medio.

Sin embargo, la mayor ventaja de estas máquinas es su capacidad de acceso aleatorio. Esto permite que las cintas sean cargadas en cualquier orden, siendo el sistema quien las ajusta y utiliza de la forma apropiada. En sí mismo, esto reduce gran cantidad de errores humanos relacionados con la selección de medios.

Otra de las ventajas importantes de los autocargadores es su capacidad de configurarse para limpiar automáticamente las cabezas de las unidades de cinta. Se puede insertar una cinta limpiadora en el autocargador y, mediante software de control o por lógica firmware, se puede desencadenar la limpieza de forma automática.

JUKEBOX.

Un jukebox de almacenamiento de datos es normalmente un sistema autocargador. Los jukeboxes son muy parecidos a las bibliotecas en lo referente a su capacidad de acceso aleatorio. La ventaja principal de los jukeboxes frente a la integridad de los datos es el tiempo que tardan en recuperarlos: no solo el medio puede ser localizado y cargado muy rápidamente, sino que una vez en la unidad puede ser leído de forma mucho más rápida que en otros equipos de cinta comparables. Esto hace que los jukeboxes sean el componente de almacenamiento de datos de uso frecuente más destacado de los sistemas HSM, que serán tratados en el capítulo siguiente.

3.6 TIPOS DE COPIAS DE SEGURIDAD

La pregunta fundamental asociada a las copias de seguridad es ¿Cuántas copias de seguridad es necesario hacer y cuándo es necesario hacerlas para asegurar la recuperación de un sistema?. A continuación se muestran y se explican los tipos de copias de seguridad que se pueden realizar.

COPIAS DE SEGURIDAD COMPLETAS.

En este tipo de copias todos los datos del sistema son almacenados en el medio. De esta forma es segura la total recuperación de todo el sistema a partir de la copia de seguridad de cualquiera de los días que se realice. Sin embargo, las copias de seguridad completas tienen a menudo una gran cantidad de datos; por este motivo, sólo se realizan copias de seguridad completas los fines de semana.

COPIAS DE SEGURIDAD INCREMENTALES.

La solución siguiente es realizar copias de seguridad de solamente aquellos archivos que han cambiado desde la última vez . Las copias de seguridad incrementales son la forma más eficiente de realizar copias de seguridad.

Sin embargo, el tiempo que llevaría recuperar los datos de todas las cintas sería bastante grande. Otro de los problemas de las copias de seguridad incrementales es que la identificación de los archivos modificados es generalmente dependiente de los cambios producidos en los atributos del sistema de archivos, que no siempre es un método fiable. Es posible desarrollar algún tipo de base de datos o registro de un sistema de archivos que identifique los que han sido modificados. Para aumentar la velocidad de las copias de seguridad, así como para reducir el número de cintas que se necesitan, suelen combinarse las copias de seguridad totales con las incrementales.

COPIAS DE SEGURIDAD DIFERENCIALES.

Las copias de seguridad diferenciales son las que realizan copias de seguridad de todos los archivos que han cambiado desde la realización de la última copia de seguridad completa. Son similares a las copias de seguridad incrementales, salvo que los archivos que se almacenan son los que cambian desde el día en que se realizó la ultima copia de seguridad completa hasta el día en que se realice la siguiente; sin embargo, las copias de seguridad diarias que se realizan tardan gradualmente más en completarse, hasta que se realiza una nueva copias de seguridad completa.

La principal ventaja de las copias de seguridad diferenciales es que podría ser posible recuperar todo el sistema a partir de dos cintas: la que tiene la copia de seguridad completa y la que tiene la copia de seguridad diferencial .

COPIAS DE SEGURIDAD BAJO DEMANDA.

Las copias de seguridad bajo demanda son las que se realizan fuera de la planificación regular de las copias de seguridad. Hay muchas razones por las que se podrían usar este tipo de copias de seguridad no planificadas: por ejemplo, quizá sólo se requiere realizazar copias de seguridad de unos pocos archivos o directorios, o quizá se quiere realizar la copia de seguridad de un servidor antes de actualizarlo. Las copias de seguridad bajo demanda

también se pueden utilizar para potenciar las copias de seguridad planificadas, normalmente por motivos de redundancia o seguridad a largo plazo.

EXCLUSIONES.
Estas no son realmente copias de seguridad en sí mismas, son tan sólo datos que no se quiere almacenar en las copias de seguridad. Hay formas de asegurar que estos datos no se copien en el medio. Puede que el tamaño de los datos sea enorme, pero no importante, o quizá siempre causa problemas durante la realización de copias de seguridad y todavía no se ha resuelto el problema.

3.7 TIPOS DE OPERACIONES DE RECUPERACION.

Típicamente, las operaciones de recuperación se clasifican en dos grupos. El primero es el de las recuperaciones completas de sistemas, y el segundo es el de la recuperación de archivos individuales. Hay un tipo adicional de operaciones de recuperación que merece la pena mencionar, denominado recuperaciones re dirigidas.
En general, las operaciones de recuperación son mucho más problemáticas que las de almacenamiento. Mientras que la realización de copias de seguridad sólo copia información fuera del disco, las operaciones de recuperación tienen que crear realmente los archivos en el sistema destino, y hay muchas más cosas que pueden ir mal cuando se crean archivos. Estas incluyen sobrepasar el límite de almacenamiento, restricciones de permisos y errores relacionados con la re escritura de archivos.
Las copias de seguridad no necesitan saber mucho sobre el sistema antes de recuperar los datos, tan sólo copian lo que se supone que tienen que copiar. Por otra parte, para la recuperación de información es necesario conocer qué archivos deben recuperarse y cuáles no. Considere la versión anterior de una aplicación que ha sido borrada y reemplazada por una nueva aplicación que ocupa todo el espacio que ocupaba la anterior. Ahora suponga que el sistema falla y necesita ser recuperado de una cinta. Es muy importante que el sistema de copias de seguridad detecte que se ha borrado la versión anterior de la aplicación para no intentar recuperar tanto la nueva como la vieja, desbordando el servidor durante el proceso de recuperación y volviendo a estropear el sistema. No hay que suponer que el responsable del sistema de copias de seguridad, o al que se tenga en mente contratar, trate este problema adecuadamente.

RECUPERACIONES TOTALES.
Las recuperaciones totales de información son utilizadas después de sucesos catastróficos o durante la realización de actualizaciones, reorganizaciones o consolidaciones de sistemas. La idea es simple: se llevan los datos de un sistema a un medio, y se depositan de nuevo en el sitio, donde originalmente se encontraban. Dependiendo del tipo de operaciones de recuperación que se estuviera utilizando se pueden necesitar varias cintas.
Por regla general, si es posible, la primera cinta que debería utilizar para recuperar el sistema es la que contiene la última copia de seguridad, porque tendrá los archivos con los que actualmente se estaba trabajando, y los usuarios finales los necesitarán tan pronto

como se encuentre listo el sistema. Utilice otra cinta que contenga la mayor parte de archivos.
Una de las cosas con las que hay tener cuidado es la suposición de que existen todos los archivos en las cintas que contienen la última copia de seguridad junto con las que contienen las últimas copias diferenciales. Generalmente, suele haber archivos de los que no se hacen copias de seguridad porque un usuario no salió de la sesión por la noche dejándolos abiertos, o por alguna otra razón. Por eso, después de haberse realizado las operaciones de recuperación, se deben revisar los errores producidos recientemente, con el fin de buscar cualquier archivo que pudiera haberse colado.

RECUPERACION DE ARCHIVOS INDIVIDUALES.
Generalmente, necesitan la última versión de algún archivo escrito en el medio porque acaban de estropear o borrar la versión con la que estaban trabajando. Para la mayoría de los productos de copias de seguridad esto es una operación relativamente simple (se hojea el catálogo o la base de datos de copias de seguridad, se selecciona el archivo y se realiza la tarea de recuperación). La mayoría de los productos también permiten seleccionar archivos de un listado diario de medio.
A veces, se requiere ir un poco más atrás en el tiempo para recuperar una versión más antigua; de nuevo, hoy en día la mayoría de los productos proporcionan métodos para hacer esto.

RECUPERACIONES REDIRIGIDAS.
Una recuperación redirigida es aquella en la que el(los) archivos(s) que se están recuperando se devuelven a un sistema o ubicación diferente de la que fueron recogidos durante la realización de la copia de seguridad. Pueden ser recuperaciones completas o recuperaciones individuales de archivos.
Las recuperaciones redirigidas no suelen ser problemáticas si no presta atención a los detalles. Recuérdese que si se cambia el nombre de los servidores, o reorganiza completamente el almacenamiento del servidor y de los nombre de los volúmenes, necesitará utilizar una recuperación redirigida cada vez que recupere archivos de la cinta donde se almacenaron antes de la reorganización. No hay que olvidar el nombre y el camino del servidor primitivo; siempre se debería ser capaz de obtenerlo del medio diario, pero eso añade tiempo y suspense al proceso. También hay que recordar que el nuevo sistema utilizará probablemente diferente información de seguridad que el antiguo, y que esto puede causar que no se trasladen correctamente los permisos de los archivos después de la recuperación.

3.8 REUTILIZACION DE CINTAS.

Una de las primeras cosas que se debe hacer durante el establecimiento de una estrategia de copias de seguridad es la reutilización dc cintas. La reutilización de cintas es un esquema que, de acuerdo a una planificación predeterminada, selecciona el medio que debe ser utilizado en cada momento. Se debe hacer esto porque, en realidad, los datos están en cintas y, si se quieren recuperar alguna vez, se verá que un sistema de organización le puede ayudar inmensamente. Las cintas son relativamente baratas, pero eso no significa que tenga

sentido utilizar una cinta nueva cada día. No sólo hay un costo involucrado, sino que manejar una cantidad de cintas cada vez mayor hace extremadamente difícil la organización para realizar recuperaciones.
La función principal de la reutilización de cintas es conocer cuándo se puede grabar nueva información en una cinta encima de la existente o, dicho a la inversa, cuál es el periodo de tiempo en el que no se puede volver a grabar en una cinta.
La protección de los datos contenidos en las cintas es el motivo de la importancia de los métodos de reutilización de cintas.
Los métodos de reutilización de cintas ayudan a reducir la probabilidad de errores humanos, insertar una cinta incorrecta en un momento equivocado puede tener como resultado la pérdida de datos que puede que nunca sea capaz de recuperar de nuevo.
Otra ventaja de la reutilización de cintas es que permite desarrollar sistemas de carga automática. La combinación de un sistema de carga automática con una planificación de reutilización de cintas puede proporcionar un conjunto previsible de operaciones en las que se pueden confiar.

REUTILIZACION A/B.
En este modelo se tiene una cinta llamada "A" u otra llamada "B" que se intercambian diariamente. Esto significa que "A" es utilizada en los días pares y la cinta "B" en los impares. Este modelo no permite conservar los datos durante mucho tiempo, pero al menos se sabe cuál es la cinta que debe usar para recuperarlos. El modelo A/B lleva consigo escenarios de fuerza bruta en los que se realizan copias de seguridad totales de forma diaria.

REUTILIZACION SEMANAL.
Otra de las soluciones es cambiar las cintas una vez a la semana, dejando en la unidad la misma cinta durante toda la semana. Esto funciona bien si la cantidad de datos es lo suficientemente pequeña, o si está utilizando un autocargador que puede cambiar los medios si se llena la cinta anterior. Puesto que esta técnica permite utilizar una cinta toda la semana, el proceso de recuperación sólo necesitaría utilizar una única cinta. La idea es realizar primeramente una copia de seguridad completa y luego añadirle copias incrementales al final de dicha cinta.

REUTILIZACION DIARIA.
Otro modelo es el que obliga a cambiar las cintas cada día de la semana. Esto significa que se podrían tener siete cintas etiquetadas con los días de la semana. Este escenario funciona bien combinando con copias de seguridad completas y copias de seguridad diferenciales o incrementales.

REUTILIZACION MENSUAL.
Una extensión del método anterior es usar una reutilización mensual. Aunque esto no es muy común, a veces se emplea como forma de reducir el número de copias completas. Se suele llevar a cabo realizando una copia completa el primer día de mes, y copias incrementales en otras cintas durante el resto del mes, con cintas que se cambian diaria o semanalmente. También se puede realizar con copias difrenciales diarias, aunque éstas pueden ser bastante grandes al final de mes.

REUTILIZACION GFS (ABUELO, PADRE, HIJO).
El modelo de reutilización de cintas más comúnmente utilizado es el conocido con el nombre abuelo, padre, hijo (GFS). Es una combinación de los modelos diarios, semanales y mensuales vistos anteriormente, y la idea es la siguiente: se designan cuatro cintas que serán utilizadas de lunes a jueves, cuatro para ser utilizadas los fines de semana de cada mes (el primer fin de semana, el segundo fin de semana y así sucesivamente), y una cinta para cada fin de mes. Se realizan pequeños ajustes para acomodar los pocos días extras de cada mes, pero el patrón es el mismo.
La principal ventaja de GFS es que se ajusta perfectamente al calendario y los ciclos de negocio. Presenta un equilibrio entre los mecanismos de protección de copias de seguridad completas de los fines de semana con las copias diferenciales o incrementales el resto de días de la semana. Además, su funcionamiento es fácil de explicar al grupo de personas de operaciones, y funciona relativamente bien con sistemas autocargadores.

3.9 TECNICAS DE RENDIMIENTO DE SOFTWARE.

La respuesta al rendimiento de las copias de seguridad es el paralelismo. Si se necesita transferir muchos datos, la mejor forma de hacerlo es dar al sistema de copias de seguridad el mayor número de caminos posible que pueda utilizar. Por ejemplo, se tarda mucho más en almacenar diez servidores con un único sistema de copias de seguridad que si se utilizaran diez sistemas de copias de seguridad, uno para cada servidor. Sin embargo, el paralelismo no es la única forma de aumentar la velocidad de las cosas; también se puede sacar provecho de la fuerza bruta y el ingenio.

CONTROL REMOTO DE LOS DISPOSITIVOS.
El rendimiento de las copias de seguridad de intranet está restringido principalmente por la red. Si realmente se quiere conseguir el máximo rendimiento en las copias de seguridad de red, es necesario utilizar dispositivos conectados directamente al sistema. Una de las formas de hacer esto es poner el bus SCSI y el dispositivo en el destino, y controlarlos remotamente desde el motor en otra máquina. La figura 3.11 muestra este tipo de solución.

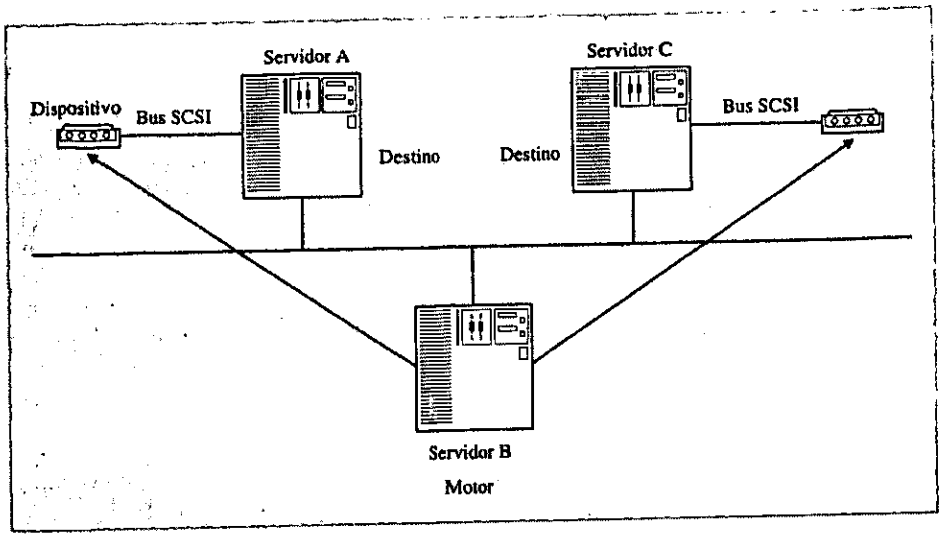


Figura 3.11 El motor del servidor B controla los dispositivos remotos de los servidores A y C.

DISPOSITIVOS SCSI EN PARALELO.

Otra solución efectiva que utiliza la fuerza bruta para aumentar la velocidad de las copias de seguridad es utilizar múltiples dispositivos simultáneamente. Esto se puede realizar porque el bus SCSI tiene mucho más ancho de banda que los dispositivos y está desocupado la mayor parte del tiempo. Sin embargo, no hay que esperar que el rendimiento del bus SCSI sea lineal. Porque se consigan velocidades de transferencia de 30 MBpm con un único dispositivo, no hay que esperar que cuatro dispositivos lleguen a 120 MBpm. Si los dispositivos son grandes y rápidos, se empezará a ver la degradación del bus a partir de tres dispositivos.

La figura 3.12 muestra muestra un sistema de copias de seguridad intranet que utiliza varios dispositivos en paralelo para realizar copias de seguridad simultáneas de varios destinos.

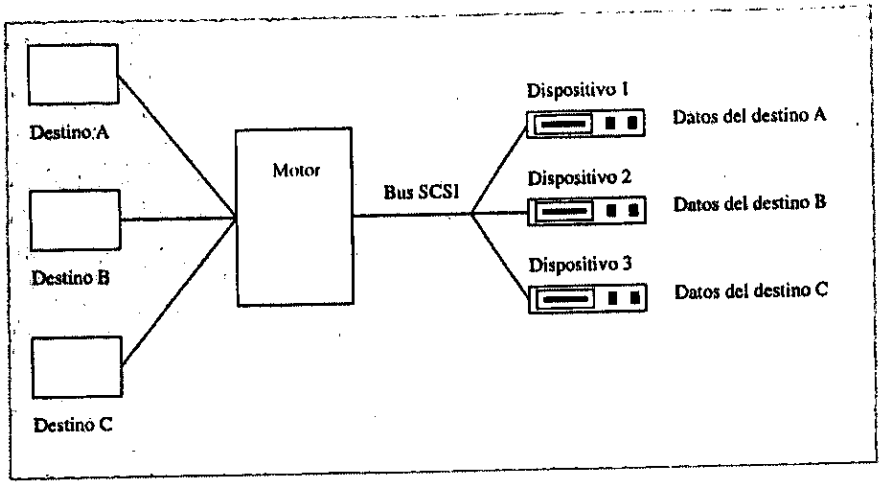


Figura 3.12 Dispositivos SCSI en paralelo.

ALMACENAMIENTO EN DISCO.

De forma similar a la multiplexación de sesiones visto anteriormente, el tráfico de copias de seguridad de múltiples destinos puede ser almacenado en un gran disco duro en el host del sistema de copias de seguridad. Aunque esto añade un dispositivo extra al proceso, las unidades de disco tienden a ser mucho más rápidas que las de cinta y consiguen mejores velocidades de transferencia en red. El principal beneficio de esta solución es que los datos de cada destino se escriben en grandes bloques contiguos del medio de copias de seguridad, hecho que podría apreciarse durante las operaciones de recuperación. También, en grandes sistemas de copias de seguridad donde la gestión de las cintas es un problema, es mejor cuanto menos esparcidos estén los datos entre múltiples cintas.

CAPITULO 4
ARCHIVADO Y ADMINISTRACION JERARQUICA DE ALMACENAMIENTO

4.1 Archivado.

Las copias de seguridad son una solución efectiva de fuerza bruta para realizar recuperaciones frente a desastres y fallos de sistema. Sin embargo, las copias de seguridad en sí mismas no suministran toda la funcionalidad que una corporación necesita para proteger y administrar sus datos.

La integridad de los datos puede ser comprometida por problemas en los sistemas interactivos, posiblemente por fallos o errores del sistema, pero también errores humanos o daños intencionados. Una forma de disminuir esta clase de problemas es borrar los datos de los sistemas interactivos y llevarlos a sistemas de almacenamiento offline. Si los datos no están disponibles online, se desprende que no están expuestos a lo que pudiera amenazar a este tipo de datos.

4.2 Archivado definido.

El término *archivo* ha sido utilizado en forma recíproca por algunas compañías de copias de seguridad de redes con el término *copia de seguridad*. Sin embargo, hay diferencias significativas entre las dos palabras que distinguen claramente sus funciones. Para nuestro propósito, definiremos la palabra *archivo* como una copia o un paquete de datos con el propósito de conservación histórica durante un prolongado periodo de tiempo. Y definiremos *copia de seguridad* como una copia de datos con el propósito de protección frente a desastre y recuperación circunstancial de archivos en un periodo de tiempo moderado.

Uno de los usos del archivado es intuitivo: se tienen datos que se sabe que son valiosos y se quiere mantenerlos a salvo durante un largo periodo de tiempo. Otro uso del archivado es ayudar al administrador de red a borrar archivos de los discos de los servidores, manteniendo un camino de acceso a ellos desde un almacenamiento offline. Con tantos datos en la red, podría no saberse si los archivos que se están borrando son importantes, por lo que es necesario asegurarse de que se pueden recuperar. Además de establecer lineamientos que eviten el doble trabajo (borrarlos y recuperarlos).

4.3 Diferencias entre copias de seguridad y archivado.

Los datos de las copias de seguridad se guardan sólo durante un periodo de tiempo relativamente corto: un día, una semana, un mes, o incluso unos pocos años, pero generalmente durante no más de un par de meses. El propósito principal de las copias de seguridad es recuperar los datos que por alguna razón están corrompidos o degradados; las operaciones de copias de seguridad se planifican para proporcionar este tipo de funcionalidad.

Los datos archivados, por otra parte, podrían ser almacenados indefinidamente. Esto tiene implicaciones serias para el tipo de medio que se utiliza y las operaciones necesarias para mantenerlo. Mientras que las copias de seguridad se realizan generalmente a diario, el archivado se realiza menos frecuentemente.
Archivando como administración de la capacidad.

La idea de archivar para preservar el espacio en disco no es nueva. La gente ha estado haciéndolo durante mucho tiempo, usando cualquier método disponible. En la actualidad las herramientas de administración de sistemas están cada vez más disponibles y tienen cada vez más potencia. Desgraciadamente, todavía no hay muchos productos que realizan completamente la función del archivado desde la selección de archivos hasta la gestión del medio. Son adaptaciones de productos de copias de seguridad y no tienen mecanismos sofisticados de selección de archivos, o son buenos seleccionado archivos, pero no suministran el soporte de los dispositivos y el medio que se necesita.

4.4 Métodos para seleccionar la forma de archivar.

Si se decide archivar datos, una de las primeras cosas en las que se habrá de pensar es en qué es lo que se va a archivar (los criterios de decisión más comunes son: tamaño, antigüedad, directorio y propiedad). Hay cuatro variables del sistema que se utilizan comúnmente para hacer estas selecciones. En la práctica se observará que una determinada combinación de estas variables funciona mejor en su caso, pero también puede descubrirse que una estrategia sencilla es más fácil de administrar.

La tabla 4.1 los lista, e indica si serían utilizados principalmente para realizar archivos históricos o administración de la capacidad.

Variable.	Uso.	Ejemplo.
Tamaño del archivo.	Administración de la capacidad.	Archivar todos los que ocupen mas de 100 MB.
Antigüedad del archivo (desde su última actualización).	Administración de la capacidad. Archivos históricos.	Archivar todos los que no han sido actualizados durante un año.
Directorio.	Administración de la capacidad. Archivos históricos.	Un directorio es utilizado como un repositorio de archivos; todo lo que está en él se archiva.
Propiedad.	Archivos históricos.	Archivar los creados por una persona, el miembro de un grupo o el miembro de un proyecto.

Tabla 4.1 Variables del sistema de archivos utilizados en el archivado.

4.5 Administración documental.

Casi todo el software de administración documental incluye un esquema de archivado, tanto para la administración de almacenamiento como para realizar archivos históricos.

Los sistemas de administración documental administran típicamente algún grupo de documentación definida en la red. Permite recuperaciones rápidas de documentos a través de diversos criterios de búsqueda que incluyen palabras clave, texto, títulos, propietarios y, en algunos casos, incluso algoritmos de lógica difusa con lenguajes de patrones. La idea es ayudar a las organizaciones que tienen fuertes requisitos de procesamiento de documentación a recuperarla de forma adecuada.

Estos sistemas limpian la estructura de directorios, seleccionando archivos con documentos que cumplen determinados criterios, generalmente, la antigüedad. Posteriormente, los sistemas de administración documental transportan estos archivos a un directorio asignado, que sirve de repositorio transitorio hasta que se escriben en el medio o son borrados (ver la figura 4.1)

Periódicamente, el administrador del sistema de administración documental debe comprobar el estado de este directorio y escribe estos archivos a un medio. Este proceso es principalmente manual, siendo el administrador quien realiza las operaciones de escritura en el medio con otros productos.

Esto incluye la verificación de que los archivos fueron escritos realmente en la cinta y dar a la cinta un nombre que permita al sistema de administración documental identificarla. Después de que los archivos han sido transferidos a la cinta, el administrador ejecuta una operación del sistema de administración documental que los borra del directorio de archivado, cambiando su localización por la que tienen en la cinta en donde fueron escritos.

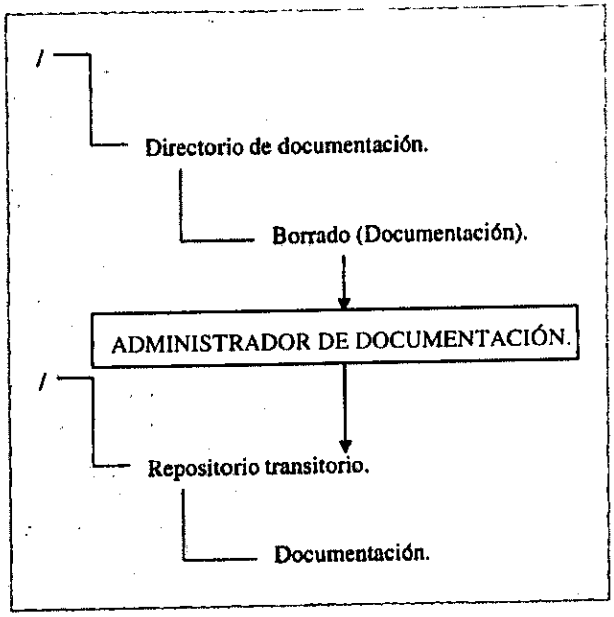


Figura 4.1 Los sistemas de administración documental borran un documento de los directorios originales y lo colocan en un directorio repositorio transitorio.

4.6 Archivado comprimido.

De forma abrumadora, este es el método más popular para archivar datos en un red, aunque es altamente susceptible a grandes fallos. El concepto básico es comprimir los datos con una utilidad de compresión de datos de forma que ocupen menos espacio en disco, luego se copian periódicamente en algún otro lugar, o se borran. A menudo, los archivos no se comprimen de uno a uno, donde un archivo comprimido corresponde a un único archivo original, sino que la utilidad de compresión los agrupa en archivos comprimidos de mayor tamaño.

Aunque es posible buscar estos archivos comprimidos por sus contenidos, dicho proceso deja mucho que desear como método de confianza para la administración de datos. Es realmente un escenario de fuerza bruta en el que el administrador de red debe seguir muy de cerca la pista de todas las variables, incluyendo los directorios, las cintas y los archivos

comprimidos donde se han transferido los datos. Por encima de eso, y a diferencia de un sistema de administración documental, los usuarios finales no tienen una herramienta sencilla que los ayude a encontrar los archivos que buscan.

4.7 Historia de HSM.

Contrario a la creencia popular, los sistemas de Administración Jerárquica de Almacenamiento (HSM, Hierarchical Storage Management) no es un concepto nuevo. Estos iniciaron hace un par de décadas en respuesta al alto costo que representa tener medios magnéticos de acceso aleatorio en línea.

Con la finalidad de maximizar la utilización de medios magnéticos, los desarrolladores del sistema operativo para mainframe se presentaron con la idea de migrar los contenidos de los archivos de los discos caros a cintas magnéticas menos caras. Esta simple idea tenía un gran obstáculo técnico para la limpieza; cuando un archivo era movido a cinta, el usuario se encontraba con que el archivo estaba ausente en el directorio jerárquico. Para eliminar la confusión, el file system tuvo que ser modificado. Las entradas en el directorio entonces presentaban que el archivo estaba residiendo en el disco, pero una bandera oculta para el usuario indicaba que el contenido del archivo había sido migrado a un medio en línea menos caro. Cuando una aplicación accedía un archivo migrado, el software debía bloquear la solicitud de la aplicación y recargar el contenido del archivo. Entonces, la aplicación se tardaba algunos segundos (o minutos) esperando a que los datos fueran recargados. Después de que el archivo había sido recargado el sistema debía permitir que la aplicación continuara.

Hace algunas décadas, los tamaños de los archivos eran relativamente pequeños. Un archivo de 100 kilobytes (100 KB) era considerado enorme. Para los estándares de hoy, un archivo de diez megabytes (10 MB) es considerado típico para la mayoría de las aplicaciones de escritorio de multimedia. Algunas imágenes de color de tres o cuatro componentes (por ejemplo, RGB, CMYK) pueden normalmente tomar miles de megabytes. Para aplicaciones médicas tales como tomografías, imágenes de rayos x, y datos escaneados CAT para un simple paciente pueden fácilmente tomar miles de megabytes.

Debido a que el espacio en el disco magnético era limitado, los archivos eran relativamente pequeños. Los usuarios eran dueños de algunas cuantas aplicaciones (la mayoría de los usuarios accedían a una o dos aplicaciones durante su día normal de trabajo). Los usuarios eran dueños normalmente de algunas docenas de archivos. Este hecho era simple para localizar directorios y archivos.

Hoy en día, debido al número y tamaño de las aplicaciones, los usuarios típicos son dueños de cientos de archivos. Un servidor de archivos para una operación que va de pequeño a mediano tamaño (servidor departamental) puede mantener miles de cientos de archivos.

Lo que no ha cambiado mucho del pasado es la utilización típica de los archivos.
En algunos casos, algunos archivos son accedidos una vez al día, en otros casos, los archivos son accedidos en un tiempo específico durante la semana, mes , o trimestre del año. También hay grupos de archivos que son raramente accedidos, sin embargo hay otros que no tienen un patrón de acceso predecible.

Con los sistemas de computadoras que trabajan en red, los patrones de acceso a la información son muy poco predecibles. La razón es que la información es ahora puesta disponible para un gran grupo de individuos con un rango vasto de intereses.
Por ejemplo, en una compañía química un ingeniero debe de tener acceso a los manuales técnicos, los abogados tendrán acceso a los documentos que hacen referencia a algún problema legal de la compañía, un diseñador gráfico deberá utilizar imágenes usadas en un folleto que la compañía empleó hace algunos años.

La forma en que los archivos eran almacenados y accedidos hace algunas décadas es completamente diferente a la forma que se necesita hoy en día.

4.8 ¿Qué es HSM?.

La Administración Jerárquica de Almacenamiento (HSM, Hierarchical Storage Management), es un sistema automático que suministra funcionalidad de archivado de forma transparente, tanto a los usuarios como a los administradores. El único componente clave que diferencía a HSM del archivado es que el sistema HSM no borra los archivos *per se*, sino que, en vez de eso, deja un pequeño archivo resguardo en el lugar del original. Este archivo resguardo se utiliza para desencadenar llamadas automáticas al original cuando un usuario final intenta acceder a dicho archivo. Otra diferencia es que los archivos HSM utilizan el término *migrar* en lugar del término *archivar*.

Brevemente, HSM funciona de la siguiente manera: HSM selecciona los archivos que se van a migrar, copiándolos en el medio HSM. Cuando el archivo se ha copiado correctamente, se crea un archivo resguardo con el mismo nombre que el original, pero que ocupa mucho menos espacio en disco. Posteriormente, cuando un usuario intenta acceder al archivo resguardo, interviene el sistema HSM y recupera el archivo original del medio HSM apropiado.

La parte jerárquica de HSM procede del hecho de que los datos se transportan desde un tipo de repositorio a otro según envejecen. HSM conlleva una estructura que hace un uso eficiente del medio y de los subsistemas de almacenamiento que optimizan el rendimiento y ahorran dinero.

La figura 4.2 muestra la relación entre las tres jerarquías más comunes en sistemas HSM: online, de uso frecuente, y offline. Los datos online están en los discos del sistema. Los datos de uso frecuentemente, han sido trasladados a repositorios que generalmente son

menos costosos que los disco y a los que puede accederse rápidamente. Y los datos offline han sido trasladados de nuevo a medios más baratos. No pudiendo recuperarse de forma automática si están almacenados en estanterías.

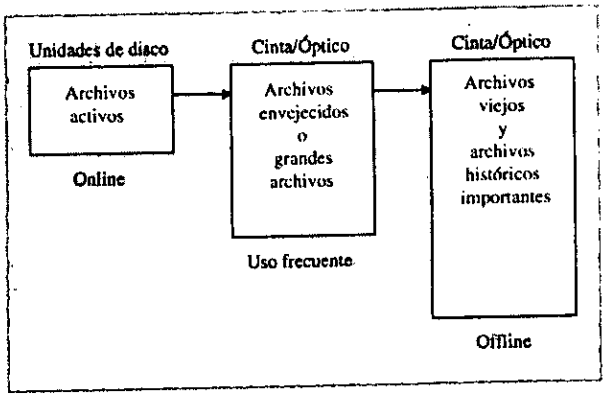


Figura 4.2 Jerarquía en los sistemas HSM.

4.9 Beneficios de HSM.

Los principales beneficios de HSM son: mejor uso del espacio disponible sobre el disco principal donde se utilizan los datos; accesos optimizados para los datos más importantes; una reducción en el costo total del almacenamiento y una administración simplificada del espacio en disco.

Con HSM el administrador del sistema no necesita estar constantemente monitoreando la utilización del disco y ejecutar continuos crecimientos de espacio cuando es reclamado.

Como una de las opciones, HSM automáticamente transfiere datos no frecuentemente utilizados del dispositivo de almacenamiento más rápido y caro, tal como el disco del servidor de archivos, hacia dispositivos de almacenamiento menos caros y menos rápidos, tales como jukeboxes y librerías de cintas. Haciendo esto, HSM permite una organización para crear almacenamientos virtuales los cuales proveen transparentemente almacenamiento de datos ilimitado para todos los usuarios corporativos. Al mismo tiempo, el disco principal es liberado para mantener solamente los datos más frecuentemente accedidos.

Por ejemplo, una estrategia HSM deberá utilizar un disco con cache ultra-rápido para proveer el más alto rendimiento para datos frecuentemente accedidos, como segundo almacenamiento, un jukebox puede ser utilizado para el acceso de datos con poca frecuencia de uso, y finalmente, el nivel más bajo de la jerarquía puede ser una librería de cinta para los datos que raramente se utilizan. Este tipo de solución optimiza el tiempo de

acceso para los datos más importantes y distribuye los demás datos a medios menos caros, mientras silenciosamente provee acceso directo a los datos a los usuarios.

Con un sistema HSM bien instrumentado, el administrador del sistema se enfocará más sobre el rendimiento de la red que en problemas de capacidad de almacenamiento.

HSM es altamente efectivo en costos, porque éste optimiza el uso del disco y utiliza medios menos caros en su lugar. A continuación se muestra una tabla comparativa de medios de almacenamiento, sobre el costo por megabyte y tiempo de acceso.

Medio	Costo por MB (dólares)	Tiempo de acceso
Estado sólido	$60 - $100	3 ms
RAID	$2 - $10	9-20 ms
Disco duro	$0.80 - $2	9-20 ms
Optico (plato sencillo)	$1 - $4	50-100 ms
Optico (jukebox)	$0.40 - $2	25 - 30 segundos
Cinta (sencilla)	$0.40 - $2	30 segundos - 3 minutos
Cinta (autocargador)	$.05 - $1	1-5 minutos

4.10 Arquitectura HSM.

Teóricamente, HSM ha estado presente sobre las redes antes de que fuera más sofisticado y conocido. Por ejemplo, cuando un administrador de red ha ejecutado un respaldo y un archivamiento, ha ejecutado un HSM de forma cruda. Típicamente, durante el respaldo, un administrador de red respalda datos del volumen de red a una cinta. En este proceso, el administrador también procura identificar los archivos inactivos sobre el volumen de red, y mover aquellos a un medio de almacenamiento menos caro de esta forma puede mantener un gran volumen de datos a un costo por megabyte más barato. Utilizando archivamiento es como utilizar una solución HSM que requiere de intervención manual para restaurar datos archivados.

El verdadero HSM, sin embargo, puede ejecutar ambas funciones (almacenamiento y restauración) de forma automática y transparente, esto quiere decir, que con HSM el usuario no se entera de que los archivos han sido migrados del disco del servidor de archivos a algún otro medio de almacenamiento, pues siempre los archivos quedarán visibles en el file system de la red, además los archivos siempre aparecerán cuando un usuario los busque desde la línea de comandos, o cuando una aplicación busque los archivos mediante los servicios del sistema operativo.

En HSM, los algoritmos utilizados para elegir que archivos se deberán migrar toman en cuenta diversos factores.

El primer criterio es usualmente la capacidad del disco: el administrador configura umbrales altos y bajos, también conocidos como marcas de agua (watermarks), y la máquina HSM mantiene la capacidad del disco en estos niveles. Cuando una marca de agua alta es cruzada, la máquina típicamente observará los archivos más viejos elegibles a migrar.

El software HSM puede identificar archivos como "viejos" basados en la frecuencia de acceso por parte del usuario final, así como también por la edad actual del archivo (fecha en que se creó). Además, el algoritmo puede ser configurado para reconocer y excluir cierto tipo de archivos, tales como ejecutables y DLLs, para asegurarse que ellos nunca serán migrados, pues provocaría tener problemas en el rendimiento del sistema operativo de la red.

Cuando un archivo es migrado fuera del medio de almacenamiento primario, HSM mantiene un archivo resguardo. Este archivo resguardo consiste de un apuntador hacia la nueva localización del archivo o una entrada índice en una BD que mantiene la actual localización del archivo. Cuando un archivo migrado es requerido, el sistema operativo le pedirá a HSM que de-migre el archivo desde el medio óptico o cinta hacia el medio de almacenamiento primario (disco duro), y de esta forma permitir al usuario final acceder el archivo requerido.

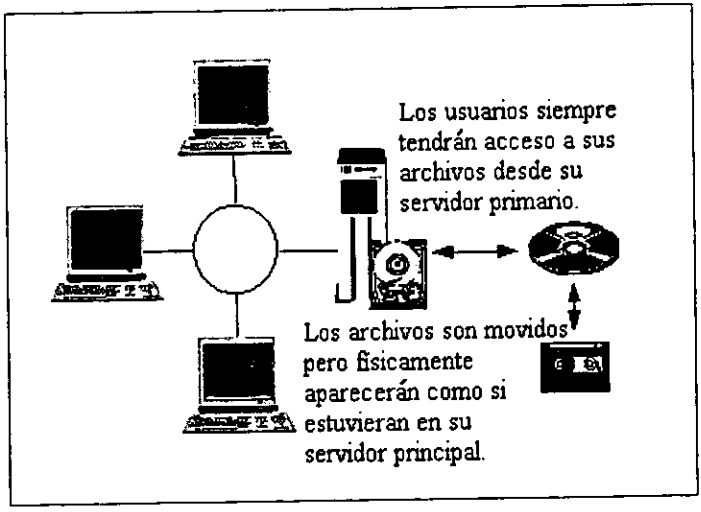


Figura 4.3 Vista de HSM en la red.

4.11 Componentes de los sistemas HSM.

La figura 4.2 muestra la relación entre los dispositivos y los medios en un sistema HSM, a continuación se muestran los componentes funcionales que hacen que todos los elementos funcionen de forma conjunta.
Un sistema HSM se construye sobre tres componentes principales:
a) Migración automática.
b) Acceso automático.
c) Archivo resguardo.

4.11.1 Migración automática.

La migración automática es la operación que realiza la función de archivado. En otras palabras, la migración copia los archivos a los medios, borrándolos posteriormente del servidor. En HSM, el componente de migración es también el responsable de la creación del archivo resguardo. La idea es que se puedan establecer algunos parámetros que determinen cuándo se pueden archivar y qué archivos se seleccionan. La figura 4.4 muestra este proceso.

La migración se dispara automáticamente cuando se superan determinados umbrales de la capacidad del sistema de almacenamiento o de acuerdo a un programa establecido. Un ejemplo es el siguiente: el sistema podría tener un proceso que vigilara si la capacidad del disco supera el 95%. Si esto sucediera, el sistema HSM se dispararía, empezando a copiar y a borrar archivos. Los umbrales superiores típicos están entre el 85% y el 95% de la capacidad del disco.

Una vez que la capacidad del disco supera un cierto punto, puede ser un problema continuar con la migración, ya que se empezarían a borrar archivos que fuera necesario tener online. Por eso, hay un determinado umbral en el que debe pararse la operación de migración. Este umbral es el umbral inferior. Los niveles típicos de los umbrales inferiores están en el rango del 60% al 70% de la capacidad del disco.

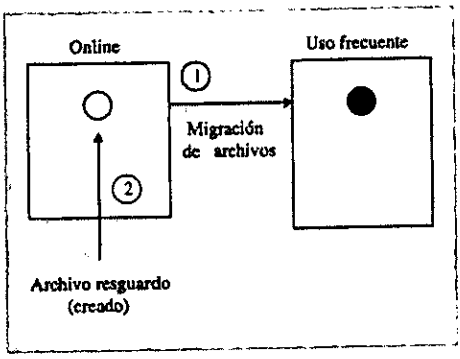


Figura 4.4 El sistema HSM migra un archivo a un medio de almacenamiento de uso frecuente y lo sustituye por un archivo resguardo.

4.11.2 Acceso automático.

El acceso automático es la función que recupera un archivo del sistema HSM cuando un usuario intenta acceder a un archivo resguardo. Cuanto más transparente y rápido se haga esto, mejor para los usuarios.

El funcionamiento es el siguiente: En el sistema hay instalado un mecanismo que identifica dicho archivo como archivo resguardado. Esto puede ser implementado mediante procesos que atrapan cada petición de apertura de archivo, o puede hacerse modificando el propio sistema de archivos, de forma que sea él mismo quien reconozca un archivo resguardo como tal. Una vez que se ha reconocido el archivo resguardo, se saca de él una pequeña información que pasa al sistema HSM. Posteriormente, el sistema HSM toma esta información e identifica el medio que debe usarse durante la operación de recuperación. Se carga al medio correcto, y el archivo se devuelve a disco, sustituyendo el archivo resguardo con el original (véase la figura 4.5).

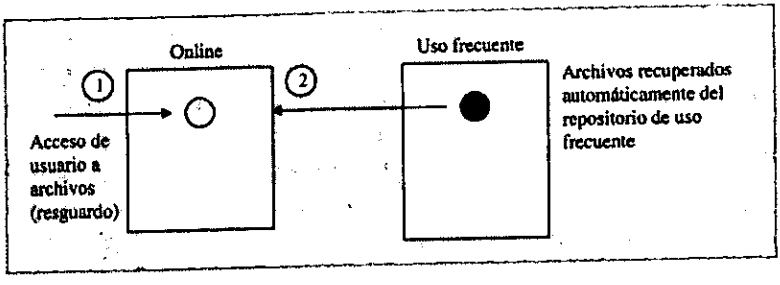


Figura 4.5 Cuando un usuario intenta acceder a un archivo resguardo, el sistema HSM recupera el archivo del sistema de almacenamiento de uso frecuente.

4.11.3 El Archivo resguardo.

Como se mencionó anteriormente, el archivo resguardo es el único componente HSM que se crea para sustituir el archivo original. El archivo resguardo tiene el mismo nombre que el original, pero es mucho más pequeño y contiene un pequeño conjunto de información relacionado con la localización del archivo original en el medio. La portabilidad de los archivos resguardo podría ser muy importante. Para que el archivo resguardo sea portable, debe ser autodescriptivo, y debe contener información de su localización en el momento de ser creado.

4.12 Operación.

La mayoría de los sistemas HSM están particionados en tres partes principales, las cuales se presentan en la siguiente figura:

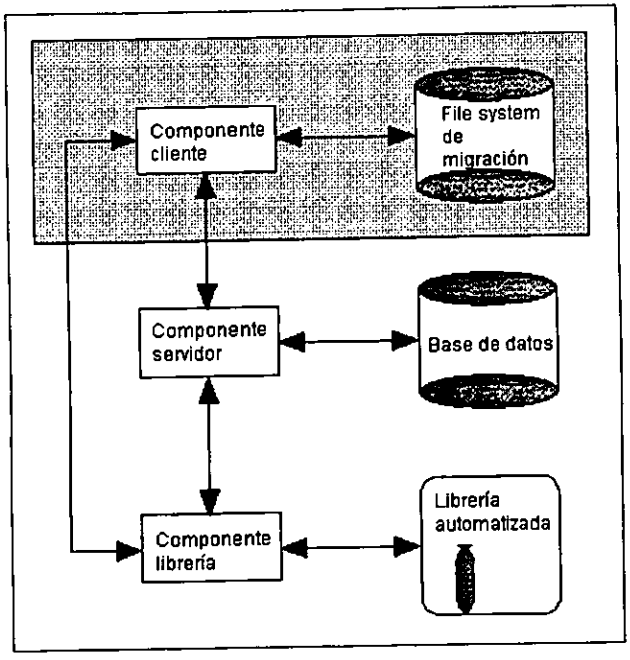


Figura 4.6 Partes de un sistema HSM.

La mayoría de los sistemas HSM ejecutan un muy bien definido conjunto de funciones (o tareas) que ellos operan. A continuación se describen las partes de un sistema HSM.

Componente cliente.

El componente cliente es la parte del software HSM que ejecuta las siguientes tareas:

a) Monitorea la capacidad del disco sobre un file system de migración.
b) Cuando la capacidad de disco alcanza la marca de agua alta, construye una lista de candidatos a la migración.
c) Copia archivos de la lista de candidatos para migración hacia volúmenes en librerías.
d) Reemplaza archivos migrados en el file system de migración con archivos resguardo.
e) Actualiza la entrada de directorio en el file system con la bandera que indica que el archivo ha sido migrado.

Con la finalidad de ejecutar esta lista de tareas, el software cliente HSM es típicamente implementado usando uno de los siguientes esquemas:

a) Modificar el file system y un conjunto de llamadas kernel del file system.
b) Usar un file system propietario y/o habitual.

Cualquiera de los dos esquemas anteriores tienden a ser dependientes del sistema operativo y en general no es fácil portar a diferentes plataformas (por ejemplo, Solaris 2.5.1 a Windows NT 4.0).

Componente servidor.

El componente servidor de un sistema HSM, típicamente ejecuta las siguientes operaciones:

a) Establece canales de comunicación para leer y escribir bitfiles (cliente asistidos).
b) Localiza bitfiles en volúmenes y librerías.
c) Aloja volúmenes y librerías para mantener bitfiles.

Componente librería.

Este componente de un sistema HSM dirige librerías específicas con la finalidad de ejecutar las siguientes tareas:.

a) Mover volúmenes de slosts a drives y viceversa.
b) Seleccionar drives a usar para lectura y escritura de bitfiles.
c) Comunicarse con los clientes para transferir bitfiles.

4.13 Configuraciones principales de HSM.

Esta sección describe las configuraciones típicas de sistemas HSM en varios tipos de negocios.
Muchas organizaciones y compañías que usan información en forma electrónica como parte de su operación regular día a día podrán beneficiarse del sistema HSM si es correctamente configurado y usado.

Los dos tipos principales de configuración de sistemas HSM son:

a) Configuración distribuida.
b) Configuración centralizada.

4.13.1 Configuración distribuida.

En el enfoque distribuido, cada sistema cliente es parte de la solución HSM. El vendedor HSM provee software para automáticamente migrar archivos del disco duro local a un disco magnético grande.
El disco magnético grande actúa como un caché. Esto es hecho para aligerar el potencial de un posible cuello de botella que puede ser causado por diferencias en las tasas de transferencia y accessos entre discos magnéticos y volúmenes magneto-ópticos o cintas en librerías.
A continuación se muestra la figura 4.7, en la cual se observa la configuración distribuida.

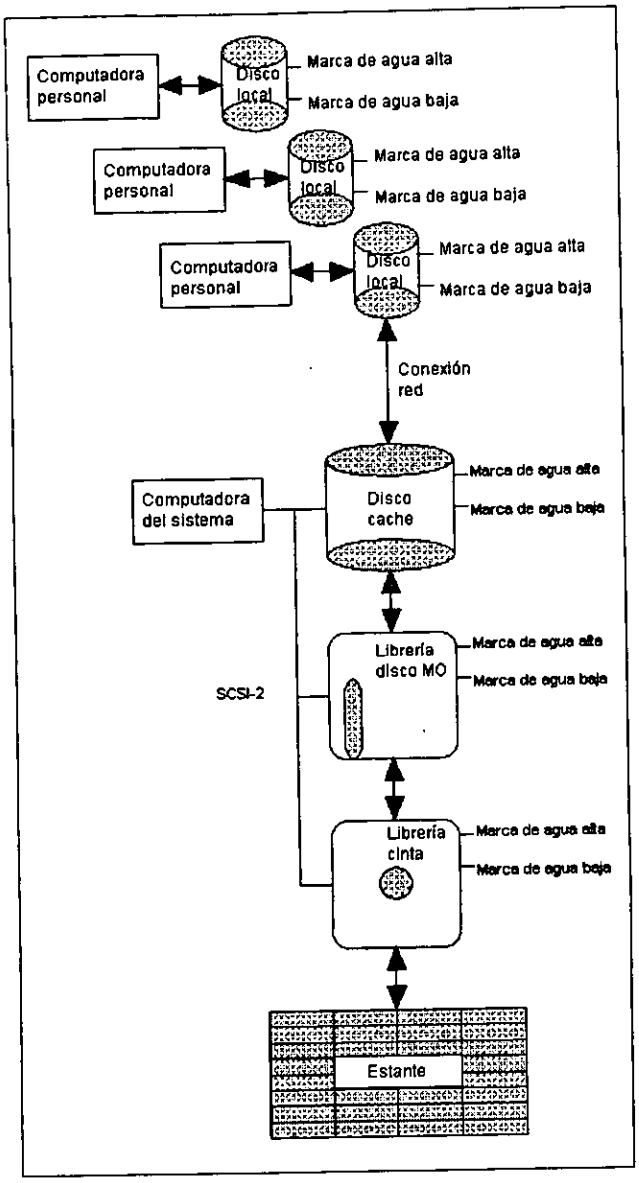


Figura 4.7 Configuración distribuida.

4.13.2 Configuración centralizada.

La principal diferencia entre los esquemas distribuido y centralizado se encuentra en la parte del cliente. En el enfoque centralizado el sistema HSM no es parte de las computadoras personales. El software cliente HSM reside sobre la computadora del sistema que está administrando el disco cache. Las computadoras personales que necesitan el acceso al file system de migración se montan sobre éste.
A continuación se muestra la figura 4.8, en la cual se observa la configuración centralizada.

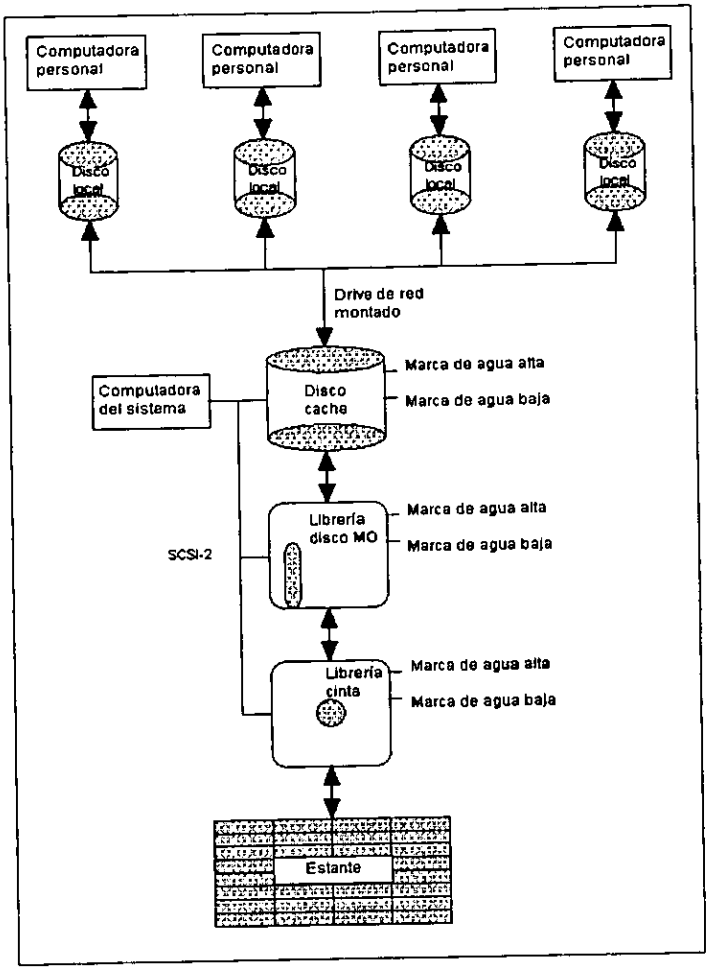


Figura 4.8 Configuración centralizada.

Con los precios de los discos magnéticos bajando y el incremento en el tamaño y número de archivos accesados por usuario típicos, el enfoque centralizado tiende a ser el más recurrido. El costo del despliegue y administración es más desahogado y provee una ventaja sobre un esquema distribuido. Las computadoras personales montan un drive de red si ellas necesitan el acceso al servidor. Si constantes manipulaciones son necesarias sobre archivos largos, una copia del archivo deseado es hecha al disco local. Cuando las manipulaciones hayan sido ejecutadas completamente, el archivo es retornado al repositorio.

CAPITULO 5
SEGURIDAD DEL SISTEMA OPERATIVO

Empezaremos este capítulo con la definición general de sistema operativo.

Un sistema operativo es el programa más importante que corre en una computadora, ya es quien administra los recursos de la computadora. Ver figura 5.1

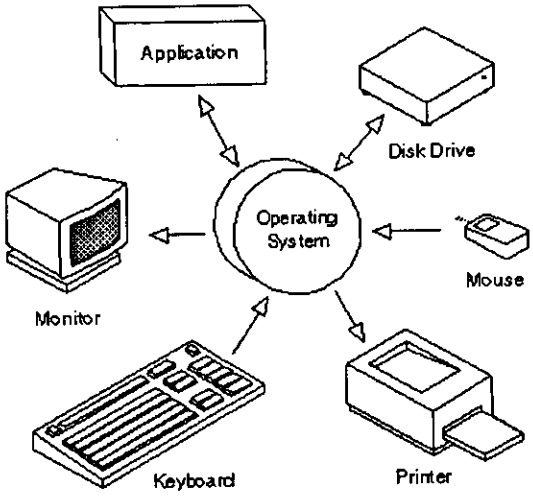


Figura 5.1 Sistema Operativo

A lo largo del capítulo se hablará en específico del sistema operativo UNIX y su seguridad parte vital de sistema operativo.

El sistema operativo UNIX se compone de cuatro partes principales: el Kernel, el administrador de archivos, el shell y las herramientas.

El kernel es el núcleo del sistema operativo. Controla el hardware de la computadora, y traduce las órdenes de UNIX en órdenes de hardware.

El file system es la manera en la cual UNIX almacena y administra información de cualquier tipo. Los archivos pueden contener documentos, gráficas, etc. UNIX trata a todos los archivos por igual y los almacena en el administrador de archivos, desde el cual puede ser recuperado por el usuario o por UNIX.

El shell es un programa que actúa como la interface entre el kernel y el usuario. El kernel está rodeado por el shell, y todos los comandos dirigidos al shell pasan a través del kernel, y éste los traduce en comandos para el hardware

Las herramientas son programas que pueden ser ejecutados por el shell para llevar a cabo varias tareas.

UNIX ha llegado a ser popular por muchas razones, una de ellas es el que realmente es un sistema operativo multitareas y multiusuarios que permite a múltiples usuarios tener varias tareas ejecutándose a la vez.

5.1 Seguridad en UNIX

5.1.1 Fuentes de daño

UNIX ha estado en constante uso por muchos años y su comportamiento es bien conocido. La naturaleza de un sistema operativo que permite multiusuarios y multitareas significa que en cualquier tiempo, varios sistemas de archivos estarán abiertos, y los datos a ser utilizados o ser escritos, pueden ser modificados por usuarios que no tengan permiso de acceso.

Además, UNIX mantiene un número de tablas de información acerca del administrador de archivos, ambos en memoria y en disco, los cuales son constantemente abiertos, o salvados en un disco.

Cuando los procedimientos del CPU son interrumpidos, los administradores de archivos y las tablas pueden perderse de la memoria, y los discos de archivos que correspondan a esta información se dejarían en un estado temporal. Los datos que comprenden estos archivos pueden estar escritos incorrectamente en el disco, y pueden no tener los identificadores y terminales propios de este archivo.

El daño al administrador de archivos puede ocurrir por un gran número de fuentes, algunas de las cuales están bajo control de los administradores del sistema.

La mayoría de las computadoras modernas están diseñadas para tolerar un rango de fluctuaciones en el voltaje y la corriente, incluyendo aquellos que son normales de encontrar en las líneas de abastecimiento de la ciudad. Sin embargo, aún en las líneas de mejor diseño, las fallas de corriente y voltaje son probables. A pesar de que existen fuentes de poder ininterrumpibles (*Uninterruptable Power Supplies)* UPS, las cuales mantienen encendidas las computadoras en caso de fallas en la corriente, estos dispositivos son caros y no siempre reaccionan lo suficientemente rápido para prevenir que el sistema se caiga. Debido a que existen caídas de voltaje y picos, cada sistema UNIX debería tener un sistema de filtrado de picos, ya que éstos son muy comunes.

La fuente principal de daño a un sistema se refiere al uso sin autorización de usuarios de cuentas. Muchos usuarios no protegen sus passwords tan bien como debieran. Muchos usuarios simplemente dejan sus terminales por largos periodos de tiempo sin salirse de la cuenta.

A pesar de que el sistema operativo UNIX provee, características de seguridad, aún requiere de que todos los usuarios sigan algunas prácticas de seguridad como:

- Uso apropiado de passwords
- Buen uso de permisos de directorios y de archivos
- Proteger los privilegios de la cuenta
- Estar alerta de otras sorpresas

El sistema UNIX puede encriptar archivos, previniendo con esto que cualquier usuario entienda los contenidos del archivo encriptado.

A continuación mencionaremos algunas de las prácticas que debemos hacer para proteger nuestros datos de los intrusos.

5.1.2 Passwords

Antes de que se puede entrar a UNIX debemos entrar en él. Resulta frustrante encontrar de pronto una pantalla con un sólo cursor, en el cual se le pide una contraseña (password) para poder acceder al sistema; el sistema operativo UNIX es una "jaula" acerca de los procedimientos de login. El archivo /etc/passwd contiene una lista de los nombres y de los passwords que los acompañan. Si un experto en violar passwords pretende entrar al sistema con un nombre que no está en el archivo /etc/passwd, el proceso de login aún solicita un password, previniendo que el intruso se percate si es que el nombre que tecleó es uno válido.

Adivinar los passwords no es difícil, ya que la mayoría de los usuarios escogen sus iniciales, sus cumpleaños, etc. Debemos escoger un buen password, el cual no sea difícil de recordar, no sean nombres y no se encuentre en ningún diccionario. El uso de nemónicos es bueno, ya que involucran patrones asociados de letras y números. Los passwords pueden ser tan largos como se quiera, y sólo los primeros 8 caracteres son tomados en cuenta.

Con el comando **passwd**, se puede cambiar el password, este comando solicita al usuario el password anterior antes que le permita hacer cualquier cambio . Esto nos previene de que alguien cambie nuestro password sin avisarnos.

Una vez que se tecleó el login y el password correctamente, estamos dentro del sistema. Cualquier archivo nuevo que nosotros creemos, será etiquetado con nuestro identificador y el acceso a los archivos estará determinado por el mismo identificador de usuario. No se debe dejar la terminal sin atención aunque sea por un minuto. Además se sugiere actualizar los passwords cada determinado periódo y en caso de que se tengan varios accesos se utilice un password diferente para uno.

5.1.3 Control de acceso, permisos y propiedad

El procedimiento de acceder al sistema correctamente es sólo el primer paso en UNIX. El procedimiento de control de acceso provee el segundo paso, determinando cuáles usuarios puede tener acceso a cuáles archivos y qué pueden o no hacer con éstos.

El sistema de acceso en UNIX está basado en un sistema Multics. Los objetos son divididos en tres categorías: el usuario (propietario del objeto), el grupo, y todos los demás. Los tipos de acceso son lectura, escritura y ejecutar (Multics incluye un tipo de acceso adicional, el derecho de añadir un archivo a un directorio). Juntos, las tres categorías y los tres tipos de acceso producen nueve permisos distintos, los cuales son desplegados con el comando ls.

Ejemplo:

	Usuario	**Grupo**	**Otros**
Lectura	r	r	r
Escritura	w	w	w
Ejecución	x	x	x
	rwx	rwx	rwx

Poner lo del sticky beat

El sistema UNIX es discreto ya que el propietario de un archivo puede hacer lo que quiera con éste. El propietario de un archivo puede dárselo a otro usuario, una vez que el archivo no le pertenece, el propietario no puede tener control sobre los permisos asociados con él.

La categoría de grupo puede ser utilizada como un sustituto para acceder al control de listas.

El sistema UNIX sigue un simple conjunto de reglas para determinar cuál categoría revisar para el acceso a archivos. Estas son:

- Si el usuario es también el propietario, revisar solamente los permisos del propietario.
- Si el usuario no es el propietario, pero es un miembro del grupo al que le pertenece el archivo, revisar solamente los permisos de acceso del grupo.
- Si el usuario no es el propietario ni pertenece al grupo propietario, revisar solamente los demás permisos.

Si un archivo permite escritura, pero no lectura, el archivo no puede ser editado. Si el archivo permite lectura, pero no escritura, el usuario deberá copiar el archivo y añadir un permiso de escritura, permitiendo la modificación de la copia. El solo permiso de ejecución permite la ejecución de programas compilados. Los programas del Shell requieren de ejecución y permisos de lectura, ya que los comandos del Shell no pueden ser leídos antes de ser ejecutados.

Los programadores tienen otra forma de llamarle a los permisos de un archivo, modo. El modo contiene información acerca del tipo de archivo y de los permisos para éste. El tipo de archivo es establecido cuando el archivo es creado, y no debe ser cambiado. El propietario del archivo puede cambiar los permisos utilizando el comando **chmod**. Este comando acepta un argumento numérico o simbólico para cambiar los permisos en un archivo. El argumento numérico consiste de tres dígitos, cada uno con un rango de uno a siete. El permiso de lectura está representado por el valor cuatro, el de escritura por un dos, y la ejecución por uno. Para combinar los permisos los valores son sumados.

En UNIX, los archivos son propiedad del usuario que los crea. Los permisos de los archivos son controlados por dos factores. Cuando un proceso crea un nuevo archivo, debe especificar los permisos por default deseados. Editores, procesadores de palabras, y hojas de cálculo, que generalmente solicitan que los nuevos archivos sean leíbles y escribibles por todos. El sistema UNIX maneja los permisos por default con el segundo factor que es **umask.**
El comando umask, elimina ciertos permisos de archivos nuevos, toma un número como un argumento, usando el mismo esquema que utiliza chmod., exceptuando los permisos de negación. Para negarle a otros el permiso de escritura, en archivos nuevos, usamos el argumento dos. Para negar los permisos de escritura grupales, y negar todos los demás permisos, usamos el argumento 27. La tabla 5.1 nos muestra los valores más comunes para esto.

Valor	Significado y tipo de seguridad
Umask 0	no restricciones (no seguro)
Umask 2	no escritura par los demás (mínima seguridad
Umask 22	no escritura para el grupo y otros (moderada)
Umask 27	no escritura grupal, ni otros permisos (fuerte)
Umask 77	no permisos para el grupo u otros (muy fuerte)

Tabla 5.1 Valores de umask y su significado

La mayoría de los sistemas UNIX tienen un valor por default de umask, el número dos, el cual no permite la escritura para otros. Para controlar el valor de umask, se pone el comando umask dentro del archivo de arranque para el programa del shell. Con un valor de restricción de umask, los archivos quedan automáticamente protegidos a menos que se

utilice el comando chmod para cambiar los permisos y hacerlos accesibles a otros usuarios o grupos.

Con respecto a los permisos de manejo de directorios se tiene lo siguiente:

- El permiso de lectura en un directorio permite el listado de los nombres de archivos en ese directorio.
- Escritura y ejecución en un directorio permite el cambio o eliminación de nombres de archivos en este directorio.
- Ejecución, permite el acceso a archivos que son referenciados por nombres de archivos en ese directorio.

5.1.4 Buena ruta

La primera "cosa" que no se debe hacer es terminar umask sin un valor de seguridad. Lo segundo es utilizar una ruta segura. La ruta define los directorios buscados por comandos dentro del shell. Estos directorios son buscados en orden, por lo que el primer directorio en la ruta, es el primero en ser buscado, y así consecutivamente. Tener una mala ruta permite a un intruso introducir una versión de un caballo de Troya[1] . Una buena ruta establece los directorios del sistema donde la mayoría de los comandos son encontrados de manera temprana en la ruta. Los directorios de sistema, como /bin y /usr/bin deben preceder de ambos directorios, locales y el directorio actual.

5.1.5 Protección de archivos de inicio

Los archivos de inicio contienen instrucciones cruciales para establecer un ambiente de seguridad. Sin embargo, si nuestros archivos son escribibles por un grupo o más gente, cualquiera puede alterar la información de inicio.

Un usurario no amigable podría alterar la ruta, o dejar el umask con un valor de cero por lo que los nuevos archivos no estarían protegidos. Un intruso podría introducir otras instrucciones que a los intrusos les gustaría que se ejecutaran y las estarías ejecutando tú, sin darte cuenta de esto.

Aún si nosotros hemos establecido un modo de protección para los archivos, de tal manera que sólo nosotros podamos leer éstos y los directorios, existen otros usuarios que aún pueden leer nuestros archivos.

---

[1] Es una instrucción dentro de un programa que permite que el programa se ejecute de manera normal, pero también ejecute funciones ilegales.

El usuario además de configurar los permisos de sus archivos de inicio, sus directorios y su umask, debe de respaldarlos, con la finalidad de que si se observa una anomalía en estos, se restaure el respaldo original y continúe trabajando sin problema alguno.

5.1.6 Encriptación de archivos

La encriptación[2] nos provee de una gran medida de privacía, y por lo tanto seguridad. Existen tres problemas con la encriptación en UNIX:

1. La encriptación está basada en un password, el cual debe ser bueno, de lo contrario los datos pueden ser fácilmente desencriptados.
2. Si existen versiones en un archivo de encriptaciones y desencriptaciones, el password puede ser descubierto y otros archivos encriptados.
3. Técnicas para desencriptación de archivos usando el esquema de encriptación de UNIX, sin utilizar el password son conocidas.

La encriptación se realiza mediante el comando **crypt,** este comando lee la información redirigida a él, encripta la información, y la escribe al archivó redirigido. Este comando o lee un password de la línea de comando, o bien solicita un password después de que es invocado.
El archivo de salida será del mismo tamaño que el de la entrada. Sin embargo, los archivos no parecerán de la misma longitud cuando son desplegados porque los archivos encriptados contienen caracteres no imprimibles.

Para resumir todo lo anteriormente mencionado, se debe recordar estos puntos básicos:

1. Utilizar un buen password; aquel que sea fácil de recordar, no un nombre ni cualquier palabra que aparezca en un diccionario; las frases sin sentido son mejores.
2. Nunca dejar una estación de trabajo o la computadora encendida, sin atención; debemos utilizar un buen programa de seguro (lock) o logout.
3. Proteger todos los archivos y directorios de la escritura de los grupos y otros.
4. Utilizar un umask restrictivo (27) que remueva todos los permisos para otros y la escritura para los grupos en todos los archivos nuevos y directorios.
5. Usar una ruta segura, con los directorios del sistema, como /bin y /usr/bin antes del directorio actual.
6. Revisar la última entrada (login), justo cuando entremos al sistema; si no vemos ningún acceso desplegado, editar los archivos de inicio.
7. Revisar los permisos de los archivos de inicio, y ver los archivos de inicio en directorios escribibles como (like/tmp).

[2] La encriptación es la traducción de los datos en un código secreto. La encriptación es el modo más efectivo para alcanzar seguridad en los datos.

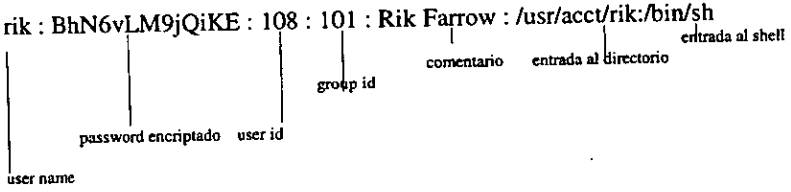
5.2 Seguridad para administradores de sistema

Un administrador de sistema es un usuario que tiene la cuenta de usuario root (el usuario root, también conocido como superusuario, tiene TODOS los permisos para trabajar con cualquier elemento del sistema operativo) y es responsable del mantenimiento de uno o más sistemas UNIX. También tiene la responsabilidad de dar los passwords a todos los usuarios, los permisos y las propiedades para cada archivo, pero en particular aquellos archivos propiedad de las cuentas del sistema.

5.2.1 Estructura del archivo de passwords

El archivo /etc/passwd tiene una estructura inflexible para cada línea. Cada línea, tiene siete campos separadas por dos puntos. Si a una cuenta le falta un campo será ignorada por el sistema.

Por ejemplo: cada cuenta que entre en el archivo **/etc/passwd** debe tener la siguiente forma:



Estos campos son:

a)User name (nombre del usuario): Este campo es utilizado como una parte de proceso de autentificación durante la entrada al sistema (login).

b)Password: El segundo campo contiene el password encriptado.

c)User id: El tercer campo contiene el número de user id. Este identifica únicamente cada usuarios y es utilizado para etiquetar los procesos iniciados por ese usuario en particular y los archivos que a éste le pertenecen.
El administrador del sistema: Asigna un número de user id cuando añade una nueva cuenta al archivo /etc/passwd. El administrado debe poner especial atención en no duplicar un número, se pueden hacer dos cosas para evitar que esto pase. Si añadimos una cuenta manualmente, utilizar el editor para buscar el user id para descubrir si aparece otra cuenta al entrar al sistema. Otra manera es la de utilizar el comando **sort** para producir una copia del archivo /etc/passwd buscando el número de user id, el cual es el tercer campo separado por dos puntos.

d)Group id (identificador de grupo): El cuarto campo en el archivo /etc/passwd contiene el identificador de grupo. Este puede ser compartido con otros usuarios. Etiqueta la entrada de procesos para la entrada del grupo a los archivos. El identificador de grupo

debe corresponder a la entrada en el archivo /etc/passwd el cual provee un nombre de grupo.

e)**Campo de comentario:** Es el quinto campo en una cuenta, la mayoría de las veces contiene información de identificación acerca de la cuenta, este campo si se desea puede estar vacío. Generalmente incluye el nombre completo del usuario, y posiblemente el número telefónico o en dónde localizarlo.
El campo de comentario: En particular el nombre completo del usuario, puede ser utilizado para suplir huellas en el password del usuario. Los intrusos usan variaciones basadas en el nombre completo, pero si los usuarios están acostumbrados a utilizar buenos passwords, el campo de comentario no supone ningún riesgo.

f)**Entrada al directorio Principal:** El sexto campo contiene la entrada de la cuenta al directorio principal. El programa de login establece el ambiente del directorio principal para este campo, e intenta cambiar el directorio a él.

g)**Entrada al shell:** El último campo contiene la entrada al shell, el programa de login ejecuta el programa encontrado en este campo. Si el campo está vacío, el shell que se ejecuta es el Bourne Shell.

En un sistema UNIX no se deben compartir cuentas ya que entonces se pierde cualquier registro de actividades del usuario dentro del sistema, así también se debe desactivar cualquier cuenta que no vaya a estar en uso durante un largo periodo de tiempo, para así evitar intrusos en nuestro sistema. Si no existen passwords, no existirá una línea de defensa inicial contra los intrusos.

5.2.2 Revisando el archivo de passwords

El archivo /etc/passwd es muy importante para el sistema de seguridad en UNIX, que debe ser revisado continuamente. Estas revisiones incluyen:

- **La propiedad y permisos del archivo /etc/paswd**
- **La corrección en los campos de cada entrada**
- **La existencia de un password para cada cuenta**

5.2.3 Propiedad del sistema de archivos y directorios de configuración de sistema

Existen unas simples reglas a seguir para hacer un sistema más seguro las cuales son responsabilidad del administrador del sistema.

- Los archivos de sistemas y directorios son propiedad del las cuentas del sistema y de los grupos.
- No escribibles para otros en directorios propiedad del sistema con la excepción de directorios temporales.
- No leíbles par otros.

- No escribibles para otros.
- Corregir los permisos y la propiedad de los archivos, para evitar que otro usuario maneje nuestros programas.
- Hacer rutinas de chequeo en permisos, propiedad de los archivos y hacer una suma de los archivos existentes.

Los usuarios son responsables de los archivos que les pertenecen. El administrador del sistema es responsable de todo lo demás, esto incluye configuración de archivos y de dispositivos, librerías y bases de datos.

5.3 Seguridad en la red y comunicación

Las comunicaciones y las redes extienden la utilidad de una computadora. Esta extensión también incrementa el acceso a cualquier computadora en la red.

Actualmente, no hay manera de tener una computadora segura en una red que se conecta a otra computadora no segura. En el caso de Internet, no hay manera de que las comunicaciones a través de la red, o computadora conectada directamente a Internet, puedan ser consideradas seguras. La seguridad en las PC's puede ser mejorada con passwords, pero al conectarla a la red no pueden ser completamente seguras.
El conectar un módem a una computadora la hace automáticamente insegura. El módem representa una extensión de la computadora, así como una terminal es la extensión de una computadora. El problema con el módem es que cualquiera que se pueda aprender el número telefónico conectado al módem, tiene acceso a la computadora.

Existen varias maneras para mejorar la seguridad en el módem. La mejor manera es no usarlo para todo. El primer paso es el de configurar bien el módem en el puerto correcto. Puede revisar la configuración del módem, cable, puerto y kernel con una simple prueba. Entre a la red utilizando el módem, como un usuario normal y cuelgue. Inmediatamente vuelva a llamar, y vea si no tiene un registro de entrada en el cursor, o el shell previo. Si su sistema no cierra la sesión en cuanto el usuario cuelga, deberá revisar varias cosas como:

- El cable de conexión del módem a la computadora debe llevar la señal del pin 8 al pin 8 de la computadora.
- El puerto utilizado debe ser capaz de detectar cambios en la portadora.

Existen características que se pueden añadir al módem para hacerlo más seguro, (aunque éstas elevan el costo), las cuales son:

Sistemas de marcar de nuevo (dialback): Estos sistemas tienen ambos modems de llegada y de salida. La ventaja de este tipo de sistemas es que nos previenen de que usuarios no autorizados marquen para entrar a la red. Solamente los usuarios en los archivos de callback pueden utilizar los modems.

Sistemas de respuesta de cambio: Estos sistemas requieren de hardware especial al final del sistema. Este se conecta entre el módem y el sistema, y previene el acceso hasta que un cambio ha sido respondido correctamente. Después de que el módem contesta, el dispositivo de cambio solicita un nombre. El cambio está basado en un código único construido en el dispositivo del usuario en tiempo real. El usuario introduce el cambio dentro de su dispositivo de respuesta, entonces teclea la llave presentada por el dispositivo; si la llave corresponde, le es permitido al usuario acceder al sistema y empezar la sesión de trabajo. La ventaja de estos sistemas es que la llamada es utilizada sólo una vez.

Dispositivos de encriptamiento: Este tipo de dispositivos añaden dos tipos de seguridad a los modems. Si una línea ha sido grabada, el dispositivo hace la recolección de información más difícil. También, iniciar una sesión o utilizar un módem que pasa a través de un dispositivo de encriptamiento no es posible, sin la llave de encriptamiento. La ventaja de estos dispositivos es su bajo costo.

5.4 Seguridad en la red

Los datos en la red son transmitidos en paquetes. Un paquete contiene los datos actuales a ser transmitidos, y un encabezado. El encabezado contiene la fuente y dirección de destino del paquete, e información acerca del tipo de paquete. Esto nos lleva a la primera característica en la red. Cualquier huésped (host: un sistema con una interface de red y software) puede leer cualquier paquete que pasa sobre la red. Un sistema que lee datos en paquetes no dirigidos a éste se conoce como, "escuchador promiscuo". A menos que la información en el paquete esté encriptada , ésta es potencialmente disponible para cualquier host. Esta es la razón por la que la mayoría de las redes no son oficialmente seguras, y no permiten que información clasificada resida en cualquier host conectado a la red. Solamente la red en la que cada host sea confiable sería considerada segura.

Otra característica tiene que ver con la identificación de la fuente de comunicación. TCP/IP utiliza una dirección de paquetes en Internet, un cuádruple de bytes (valores de 0 a 256), que se supone son únicos para cada host. En realidad, cualquiera con el privilegio de superusario puede alterar la dirección de Internete de su propio sistema. Algunos intentos para mejorar la autentificación han sido hechos, pero todavía no están disponibles universalmente.

5.4.1 Huéspedes confiables

Muchos sistemas UNIX soportan lo que es a menudo es llamado los comandos "r", **rlogin, rsh** y **rexec.** Los comandos r permiten que a usuarios remotos entrar al sistema o ejecutar comandos a través de la red. Los usuarios remotos están rodeados por las mismas contrariedades que los demás usuarios, además primero tienen que pasar por un proceso de autentificación dando su login y su password.

Mientras que el concepto del usuario remoto es interesante, por otro lado también es muy peligroso. El mecanismo de huésped confiable (trusted host) fue explotado por el gusano de Internet en la temporada pasada. Una vez que el gusano de Internet se estaba ejecutando en el sistema, el gusano revisaba a los huéspedes, y copiaba a sí mismo a otro sistema donde su identificador de usuario (user id) fuera confiable. No deben ser eliminados completamente los trusted host, solamente debemos tener cuidado en escoger en cuáles sistemas se puede confiar.

Este tipo de huéspedes están definidos en dos archivos diferentes en cada huésped. El archivo /etc/hosts.equiv provee una definición ancha de los huéspedes confiables. Y los archivos .rhosts permiten a los usuarios definir otros huéspedes confiables.

5.4.2 Seguridad en el administrador de archivos de la red

El administrador de archivos de la red permite la ascensión de los sistemas de los archivos a través de la red. Una vez que el archivo ha sido ascendido, la mayoría de las reglas de acceso familiares son aplicadas. Esto es, un usuario debe tener permiso para leer un archivos antes de que el archivo pueda ser leído, o ambos permisos de escritura y ejecución en un directorio antes de que un archivo pueda ser eliminado.
Una de las cosas que sí cambia, sin embargo, es que la ruta no es más larga que el usuario omnipotente. De hecho, la ruta se convierte en nada (el usuario con identificador -2)por default.

Cada sistema no tiene control sobre cuáles sistemas de archivos, o cuáles partes de los sistemas de archivos, están disponibles para ascender por sistemas remotos. El sistema local hace dos cosas para controlar la ascensión remota. Primero, el sistema local debe soportar NFS (Network File System) con una configuración apropiada del kernel. Lo segundo es que el sistema debe exportar el archivo del sistema.

5.4.3 Ambientes restringidos

Estos tipos de ambientes añaden grados de control acerca de lo que los usuarios pueden realizar con estos ambientes. Existen tres tipos diferentes de estos ambientes:

1. Reemplazar el login shell con un programa de propósito especial.
2. Reemplazar el login shell con un shell restringido.

Cada uno de estos tres métodos tienen sus propios beneficios . El uso incorrecto de éstos, pueden dar sentido de seguridad falsa, si es que no ha hecho nuestro sistema más seguro.

Los programas de propósito especial pueden ser los más seguros de las otras alternativas ya que como su nombre lo indica realizan una función en especial. El programa es ejecutado en el tiempo de login, y controla totalmente las actividades del usuario.

El shell restringido, /bin/rsh está ligado con el Bourne Shell, /bin/sh. Este shell engloba varias reglas, las cuales comprometen sus restricciones. Este shell trabaja como el Bourne Shell con las siguientes restricciones.

- El usuario no puede cambiar el directorio
- El usuario no puede cambiar las variables ambientales del shell o de la ruta
- El usuario no puede ejecutar un comando que incluya una diagonal, por ejemplo, /bin/chmod

La seguridad del sistema operativo es fundamental ya que en éste están construidas internamente las funciones de comunicación, de entrada -salida , etc. Finalmente el sistema operativo es la base de todo el sistema para su operación óptima.

CAPITULO 6
SEGURIDAD DE LA BASE DE DATOS

6.1 Seguridad de la base de datos.

Los datos de las bases de datos plantean un compromiso especial entre su necesaria disponibilidad y las amenazas de manipulación, corrupción y robo. El análisis de las amenazas y sus respuestas es el trabajo de cada día de la seguridad. Sin embargo, la propia seguridad procede del desarrollo de medidas sobre una base de continuidad. Demasiado a menudo, la seguridad, incluso la seguridad onerosa, falla por pequeños descuidos.

Los datos son esenciales para todos. Lo primero es reconocer que hay amenazas. Desgraciadamente, la mayoría de las amenazas son invisibles hasta que es demasiado tarde. Por eso es necesario reconocer que se está en peligro, incluso aunque no sea obvio.

6.2 Amenazas contra los datos en una base de datos.

A continuación se presentan las amenazas principales que afectan a los datos en una base de datos.

a) Manipulación/falsificación.
b) Corrupción.
c) Robo.

6.2.1 Manipulación/falsificación.

La manipulación es el cambio de los datos para falsearlos. Cualquier forma en la que se presente resulta un problema insidioso porque, a menudo, uno no suele darse cuenta de que tiene un problema hasta que se ha extendido.
A continuación se explican estas amenazas.

Ganancia personal.
Si es posible cambiar la dirección del cliente, los empleados pueden haber enviado artículos a sus casas y habérselos cobrado a los clientes

Ocultar la evidencia.
Si se pueden alterar las cantidades del inventario, se pueden robar los artículos sin que nadie se de cuenta hasta mucho tiempo después, muy a menudo demasiado tarde para seguirle la pista.

Bromas.
Se pueden cambiar los nombres de los clientes: "Vanessa González" por "Jorge Pérez". Es necesario hacer saber a los usuarios del sistema de cómputo que ese no es sitio para hacer bromas.

Ignorancia.
Permitir que la gente esté donde no le corresponde puede llevar a producir devastadores cambios sin mala intención. Cuando la gente se encuentra en situaciones raras puede estar alterando datos en un intento por salir. A menudo creen que no deberían haber estado allí, por eso permanecen en silencio debido a su propio sentido de culpabilidad. La seguridad, tanto a nivel de programas como al de datos, debe prevenir los cambios inadvertidos. El ejemplo perfecto son los usuarios que creen estar borrando archivos de sus propias estaciones de trabajo cuando, inadvertidamente, se han situado en un directorio en red, borrando los archivos equivocados.

6.2.2 Corrupción.

La pérdida de datos es otra de las amenazas que hay que afrontar. Tablas y bases de datos completas pueden ser borradas, movidas o alteradas, haciendo que su contenido deje de estar disponible. A continuación se muestran algunas fuentes de corrupción.

Sabotaje.
Tanto si se presenta de forma específica como si no, el sabotaje es fácil y difícil de manejar a la vez. Es fácil porque políticas sencillas lo prevendrían completamente, excepto a los saboteadores más decididos, y difícil porque si no se previene se tendrán grandes pérdidas y un campo de reacción más reducido.

Bromas.
Las bromas también provocan problemas de corrupción. Todos los programas acceden a los datos de alguna forma, y cambiando incluso los más pequeños detalles podrían dejar los datos ilegibles.

Virus.
Se deben limitar los accesos a fuentes de datos exteriores, discos o servicios interactivos, y forzar que todos los datos que se traigan pasen por un buen verificador antivirus.

6.2.3 Robo.

Difícil de detectar, incluso cuando se producen daños en el negocio, el robo de datos es un problema serio. El robo es realizado sobre los datos más sensibles, copiándolos a una fuente extraíble como un disquete o recogiendo informes impresos de la compañía.

Algunas compañías creen que perder los datos sería más desagradable que devastador. Nada más lejos de la realidad. Los datos de un negocio (los nombres de los clientes, los contactos, lo que se les vende y a qué precio) es una concentración de información que un competidor podría utilizar para acabar con la empresa.

6.3 Seguridad en la base de datos Oracle.

Una de las tareas principales de un Administrador de Base de Datos es administrar las cuentas de acceso, y los permisos de usuarios en una BD. A continuación se dan los puntos a revisar para tener mayor seguridad con lo que hacen los usuarios en una base de datos Oracle (se toma Oracle como referencia debido a que es una BD ampliamente conocida).
La figura siguiente (fig. 6.1) muestra lo primordial a revisar por un Administrador de Bases de Datos referente a controlar al acceso de usuarios.

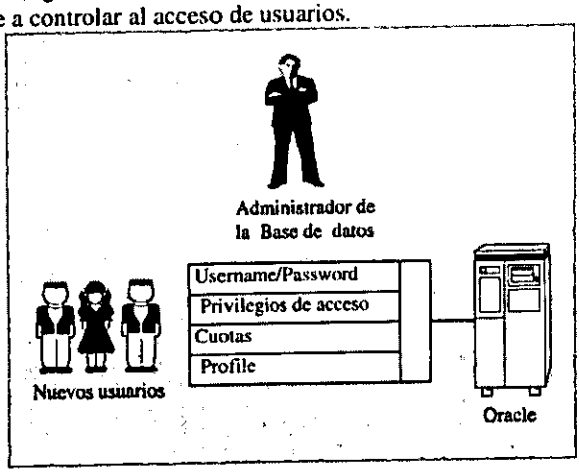


Figura 6.1 Controlando el acceso de los usuarios.

6.3.1 Control de acceso a los usuarios.

El Administrador de la Base de Datos controla el acceso a una BD Oracle a través de:
a) Creación de usuario.
b) Alteración de usuario.
c) Eliminación de usuario.
d) Monitoreo de usuario.

Creación de usuario.

Cuando se crea la cuenta de acceso de un usuario por vez primera se configuran sus siguientes elementos:

User	Identifica el nombre del usuario a ser creado.
BY password	Especifica el password para cuando se conecte a la BD.
EXTERNALLY	Verifica el acceso del usuario a través del sistema operativo.
DEFAUL TABLESPACE	Identifica la tablespace por default para los objetos del usuario.
TEMPORARY TABLESPACE	Identifica el tablespace temporal para segmentos temporales.
QUOTA	Permite al usuario a alojar espacio en el tablespace.
Integer	Especifica la cuota en KB o MB.
UNLIMITED	Permite al usuario alojar espacio dentro del tablespace sin límite.
PROFILE profile	Asigna el profile nombrado al usuario.

Un punto muy importante a configurar en este punto es el profile, pues con este elemento se puede controlar el uso de los recursos en el sistema por parte de un usuario.
Los recursos que se pueden controlar son:
a) Tiempo de CPU.
b) Operaciones de E/S.
c) Tiempo ocioso.
d) Tiempo de conexión.
e) Espacio en memoria
f) Sesiones concurrentes.

Los profile pueden ser creados, alterados o eliminados. Además permiten que los límites puedan ser habilitados o deshabilitados, estos límites pueden ser especificados individualmente, o algunos pueden ser especificados como un límite compuesto.

Otro punto esencial que debe revisar el Administrador de la Base de Datos es controlar los privilegios en la base de datos para cada uno de los usuarios.

El control del Administrador de la Base de Datos para manejar los privilegios incluye:

a) Proveer a un usuario los derechos para ejecutar un tipo de operación.
b) Habilitar y restringir el acceso y cambios a los datos.
c) Habilitar y restringir la habilidad de ejecutar funciones de sistema y el cambio a las estructuras de la base de datos.
d) Dar permisos a usuarios individuales y roles.
e) Dar permisos a todos los usuarios (PUBLIC).

Alteración de usuario.

Cuando una cuenta de acceso de usuario ha sido creada, pero por alguna razón se le dieron privilegios de más o de menos, es necesario modificar sus parámetros de seguridad. Para cambiar sus privilegios es necesario utilizar el comando ALTER USER. Este comando permite reconfigurar todos los elementos de la cuenta de acceso:

User	Identifica el nombre del usuario a ser creado.
BY password	Especifica el password para cuando se conecte a la BD.
EXTERNALLY	Verifica el acceso del usuario a través del sistema operativo.
DEFAUL TABLESPACE	Identifica la tablespace por default para los objetos del usuario.
TEMPORARY TABLESPACE	Identifica el tablespace temporal para segmentos temporales.
QUOTA	Permite al usuario a alojar espacio en el tablespace.
Integer	Especifica la cuota en KB o MB.
UNLIMITED	Permite al usuario alojar espacio dentro del tablespace sin límite.
DEFAULT ROLE	Establece los roles por default para el usuario.
PROFILE	Asigna el profile nombrado al usuario.

Eliminación de usuario.

El eliminado de usuario permite remover a un usuario de la base de datos. Muchos Administradores de Bases de Datos no prestan mucha atención a borrar a los usuarios que ya no tienen derecho a una cuenta, pero, éste puede causar grandes problemas, pues se les deja abierta la oportunidad a los usuarios para que amenacen con gran facilidad la disponibilidad e integridad de los datos.

Monitoreo de usuario

Para monitorear a los usuarios se revisa el diccionario de datos (data dictionary) de la base de datos, el cual almacena información de cada usuario.
El diccionario de datos incluye información sobre:

a) Todos los usuarios en la base de datos.
b) El tablespace por default para las tablas, clusters e índices de cada usuario.
c) El tablespace usado para segmentos temporales.
d) Cuotas de espacio.

Monitorear la base de datos por medio del diccionario de datos permite ver qué información ha sido grabada en la base de datos. El diccionario de datos para una base de datos Oracle es un conjunto de tablas y vistas que son usadas como una guía de referencia de solo lectura para la base de datos.
Las vistas conque un Administrador de Base de Datos debe estar familiarizado para monitorear a los usuarios son:

a) ALL_USERS
b) USER_USERS
c) DBA_TS_QUOTAS
d) USER_TS_QUOTAS

Los mecanismos utilizados para el monitoreo de usuarios son:

a) Indagación de tablas directamente desde SQL.
b) Revisión de bitácoras.
c) Utilización de utilerías que obtengan información del diccionario de datos.

CAPITULO 7
SEGURIDAD EN REDES

La seguridad en las computadoras va cobrando cada vez más importancia a medida que crecen las interconexiones entre ellos. En sus comienzos, la seguridad no tenía mucha relevancia ya que las computadoras estaban aisladas y había muy poca gente capaz de manejarlas, por lo que localizar a los responsables de irregularidades era tarea bastante sencilla.

Pero las redes de computadoras han seguido creciendo hasta dimensiones inimaginables hace algunos años, y a su vez los riesgos también se han multiplicado.

Es evidente que la falta de seguridad en empresas y en cualquier otro tipo de organización ocasiona pérdidas cuantiosas aunque muchas veces la principal amenaza para la seguridad informática de las empresas son los propios usuarios autorizados. No todas las entidades están de acuerdo con esta apreciación, sin embargo, hay algo en lo que prácticamente todas las empresas coinciden: los recursos humanos destinados al área de seguridad así como su presupuesto son escasos.
Mientras continúa aumentando el tamaño de las redes de computadoras y su integración con otras redes, los retos para mantener la seguridad de los datos aumentan significativamente.

Esto hace que las personas del mundo exterior tengan más posibilidades de obtener acceso a los servicios informáticos, incluso aunque no tengan ninguna relación con la empresa con cuyas computadoras intentan establecer contacto.

Este capítulo examina algunas técnicas y tecnologías disponibles para protegerse de los ataques contra la seguridad de una red.

7.1 Sistemas operativos de red

Los sistemas operativos de red (NOS) como NetWare de Novell, LAN Server de IBM Corporation y Windows NT Advanced Server de Microsoft Corporation ofrecen servicios de red a las LAN. Los sistemas operativos de red tienen características diseñadas específicamente para cumplir con los requisitos de las redes de computadoras y con el procesamiento de las aplicaciones en servidores en sistemas cliente/servidor. También incluyen herramientas de administración de la seguridad de los datos; gestión de red; resolución de problemas; y administración de usuarios, computadoras y periféricos de la red.

En la computadora que opera como servidor hay tres componentes principales:

1. Sistema de acceso
2. Sistema de gestión de archivos
3. Sistema de caché en disco

El primero, el sistema de acceso, ayuda a controlar quién utiliza qué datos y el acceso de múltiples archivos. La siguiente parte, el sistema de gestión de archivos, está diseñada para gestionar la lectura y la escritura de los datos a, y desde un disco duro, a la red. Finalmente, hay un sistema de caché en disco que gestiona el flujo de información en las cachés del disco duro. Además, también hay funciones que coordinan la seguridad de los recursos compartidos, así como las necesidades de impresión y comunicaciones.

7.2 Interconexión de redes

La interconexión es otro de los mayores progresos en redes de área local. Con la cantidad de LAN's que se están estableciendo en los diferentes departamentos, sucursales y ubicaciones de una empresa, surge la necesidad de enlazarlos de alguna forma. La interconexión es la respuesta, a través del uso de puentes y ruteadores. Esto permite que una LAN a miles de millas de distancia sea alcanzada tan fácilmente como la de la oficina de al lado.

7.3 Seguridad en redes

La seguridad en las redes está relacionada con la seguridad de los sistemas, programas y datos cuando éstos existen en una red establecida. Los riesgos de seguridad de las redes son mucho mayores que los de los sistemas aislados. Las siguientes figuras (7.1 a 7.3) muestran algunas de las diferentes topologías que existen en redes. No existe ninguna diferencia en cuanto a seguridad y la topología que se esté usando.

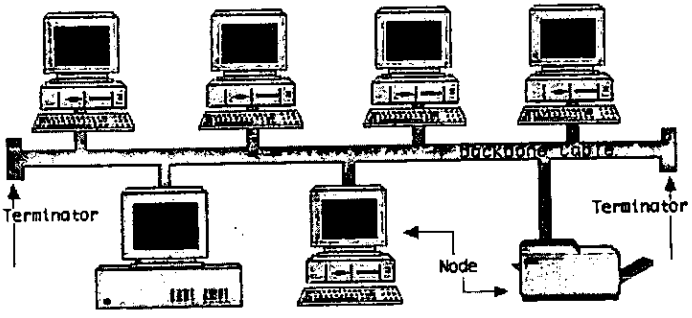


Figura 7.1 Topología BUS

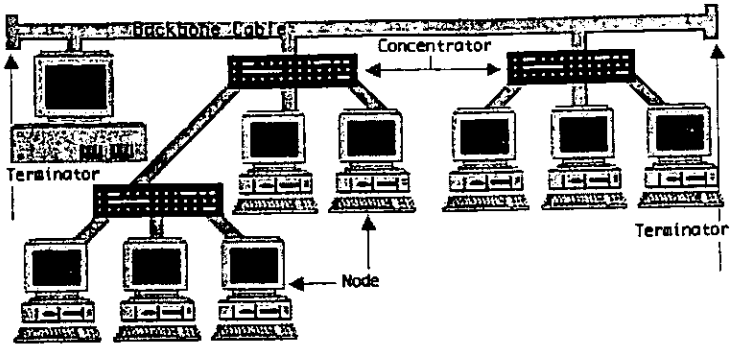


Figura 7.2 Topología de Arbol

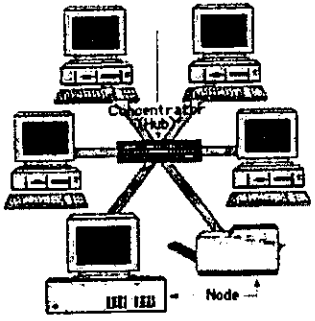


Figura 7.3 Topología Estrella

Para nuestro caso se utiliza la topología de Anillo Token Ring con FDDI

Los efectos de compartir información y recursos en las redes, desde el punto de vista de la seguridad, tiene como resultado la existencia de más usuarios potenciales (simpáticos o no) accediendo al sistema, o interceptando los datos de la red, accediendo de forma ilegal a los datos, a los programas y a los recursos de una localización remota.

El hecho de que la información tenga que viajar de un lugar a otro también incrementa la posibilidad de errores y corrupción.

7.4 Estrategias de protección

El concepto básico de estrategias de protección se puede dividir en varias tácticas:

Crear un entorno de red seguro: Hay métodos para la supervisión de los usuarios y de lo que pueden hacer en un sistema como:

1. El control de acceso
2. La identificación/autenticación
3. La visualización de ruteadores
4. La utilización de cortafuegos
5. Etc.

Cifrar los datos: Un problema de las redes es que un intruso puede interceptar un sistema y robar sus datos. Esto puede ser aliviado, no sólo utilizando diversos medios que impidan las interceptaciones, sino también cifrando la información de forma que, aunque incluso sea robada, no pueda ser leída.

Desarrollar planes de contingencia: Si existen planes de contingencia, medidas de copias de seguridad y otras formas de manejar desastres, es bastante más fácil recuperarse de una brecha de seguridad. Esto se verá con más detalle en el capítulo 10.

Planificar y administrar sistemas: Considerando la seguridad como un elemento importante. Aunque algunos de los aspectos mencionados anteriormente podrían incluirse dentro de éste, es importante planificar y administrar la red de forma apropiada para estar preparados frente a cualquier eventualidad.

Utilizar cortafuegos: Para prevenir las amenazas de las comunicaciones. Las puertas traseras de la seguridad relacionadas con la red Internet pueden permitir hacer daño a los intrusos. Dos de éstas incluyen el enmascaramiento y la denegación de servicios. Con el enmascaramiento, alguien finge ser otro diferente. Esto puede hacerse con una amplio rango de estrategias. Estas podrían ser tan simples como iniciar una sesión utilizando una contraseña robada o acceder al servicio informática a través de la tarjeta de identificación de otra persona. Hablaremos de los cortafuegos con más detalle en la página 87.

La vulnerabilidad de la red está sujeta a las siguientes brechas de seguridad:

1. **Equipos de comunicación**: La mayoría de los equipos de las redes de comunicación pueden ser una fuente de puntos débiles de seguridad, tanto si se habla de equipos de señalización y conmutación, dispositivos de interconexión de redes, servidores de archivos, como de cualquiera otro de los muchos componente de una LAN o una gran red. La avería física de estos equipos puede dañar o inutilizar una red.

2. **Medios de comunicación de red:** El par trenzado, los cables coaxiales o cualquier otro tipo de cables utilizados para conectar redes también son vulnerables al daño, el sabotaje y el asalto, debido a la interferencia.

3. **Conexiones de red:** Acceder a un sistema de forma remota es tan fácil como si se estuviera en el propio centro de cálculo. Las conexiones vía módem son particularmente vulnerables.

4. **Sistemas operativos de red:** Los sistemas operativos de red ayudan a mantener la seguridad a través de las facilidades de control de acceso, autenticación y control discrecional incorporadas en los propios sistemas.

5. **Ataque de virus:** El efecto de poderosos virus u otras formas de «fauna» informática pueden ocasionar muchos estragos en la red debido a la interdependencia de las partes de un sistema. Si, por ejemplo, un servidor es atacado por un virus, éste puede afectar a toda la red

6. **Seguridad física:**Para que las redes puedan funcionar y hacer su trabajo es necesario una gran cantidad de hardware, software y equipos de comunicación especializados. Los problemas con las redes, incluso los más pequeños, pueden ocasionar graves problemas, por lo que es muy importante proteger los accesos a los servicios de procesamiento en red y a los propios equipos. Hacer la red redundante.

7.5 Planificación y administración de sistemas

Aunque hay muchas técnicas disponibles para proteger la red, pueden ser poco efectivas si no hay un plan de seguridad detallado o una política clara de cómo administrar la red con la seguridad en mente. En resumen, si se quiere mantener la seguridad de una red, es importante planificar y administrar los sistemas de forma efectiva.

La complejidad de las redes requiere de una estrategia de seguridad que identifique los componentes de ésta que pueden ser susceptibles a brechas de seguridad.

Estos componentes son:

1. Servidores de red
2. Dispositivos de interconexión (repetidores, puentes, ruteadores, pasarelas)
3. Software de los sistemas operativos de red
4. Aplicaciones software ejecutándose en la red
5. Cableado y medios de comunicación
6. Información utilizada y transmitida por la red
7. Módems
8. Gente
9. Documentación

Se debe implantar una política completa de seguridad de servicios/nodos, escritos e implantados cuidadosamente en la red de computadoras.

Otra solución es crear un análisis de seguridad de la red determinando qué áreas necesitan una atención particular en función de la estructura actual o planificada de la red.

7.5.1 Control de acceso y autenticación

Una red puede hacerse más segura si se conoce quien accede al sistema, y si la persona es efectivamente quien dice ser. Hoy en día, el mantenimiento del control de acceso y de la autenticación es una parte importante de la administración de una LAN.
La asignación de niveles de seguridad suministra diferentes conjuntos de derechos y responsabilidades a los usuarios, dependiendo de la «posición» de cada uno o su «papel» dentro de la red. Son los siguientes:

Supervisor: Posee todos los derechos; puede crear supervisores equivalentes; crear todos los grupos de administradores; y tiene todos los derechos en el sistema y en todos los volúmenes. Esto podría ser semejante de alguna manera al «superusuario» de Unix, que tiene poderes absolutos en todo el sistema.
Equivalente al supervisor: Ofrece los derechos del supervisor, pero no es la cuenta del SUPERVISOR. En general, este usuario dispone de los derechos y las capacidades equivalentes de la cuenta del SUPERVISOR.
Administrador de grupos de trabajo:Este es un tipo de usuario que tiene los derechos de administrar un grupo de usuarios. El o ella pueden crear otras cuentas de administración de usuarios, crear y borrar usuarios, administrar las cuentas de usuario y usar la utilidad FCONSOLE de forma limitada.
Administrador de cuentas de usuario: Este tipo de cuenta permite a alguien trabajar como administrador sobre ciertos usuarios y/o grupos.
Operador de colas de impresora: Un operador de colas de impresora tiene derechos especializados relacionados con las colas de impresión.
Operador de servidor de impresoras: Un operador de servidor de impresoras tiene los derechos para administrar el servidor de impresoras.
Operador FCONSOLE: Este usuario tiene derechos para administrar la utilidad FCONSOLE. Esta se utiliza principalmente para llevar a cabo las opciones de SUPERVISOR.

7.5.2 Seguridad de los modems

Marcando un número de teléfono, alguien puede establecer una conexión directa con el módem de una computadora remota y «hablar» con el sistema como si estuvieran delante de él.

La accesibilidad de estos sistemas mediante números de teléfono normales, incluso a través de redes globales de conmutación de paquetes, ha posibilitado que cualquiera con un módem y una línea de teléfono tenga acceso a miles de computadoras.

Por esto, la seguridad en los módem es un aspecto importante cuando se permite la facilidad de acceso a través de ellos.
El objetivo principal de una seguridad efectiva con módems es impedir los accesos desautorizados a estos servicios, restringiéndolos sólo a los usuarios autorizados.

Seguridad de los módems en accesos telefónicos

Para aumentar de forma adicional la seguridad del sistema, se puede añadir una contraseña, que mantenga alejado a cualquiera que no tenga una contraseña de módem válida. Esta contraseña de módem es independiente y distinta de las contraseñas de inicio de sesión del sistema.

También hay módems conocidos como «módems con retrollamada». Estos módems no establecerán inmediatamente una conexión cuando reciban una llamada.
También hay módems que cifrarán la información que envían y reciben, de forma que ésta no pueda ser interceptada, o accedida de otra manera, en su forma original. Esto requiere la compra de módems de propósito especial.

Además, existen «módems silenciosos» especiales que no enviarán la señal característica de «conexión establecidas hasta que no se haya completado el inicio de la sesión. Esto ayudará a evitar a los que se dedican a buscar secuencias de números de teléfono de computadoras.

7.6 Seguridad de los medios de transmisión

Otro punto vulnerable e importante en términos de seguridad son las conexiones que unen las redes. Tanto en redes LAN de tamaño medio como en grandes (WAN) globales de interconexión de redes, la elección de los enlaces de comunicación, o medios, puede tener impacto en la seguridad de la red. Esto no es difícil de entender, ya que cualquier clase de medio de transmisión empleado, ya sea un cable físico, un par trenzado, un coaxial o un medio ilimitado como las microondas, y la utilización de fibra óptica (figura 7.4) puede tener ventajas y desventajas en términos de seguridad.

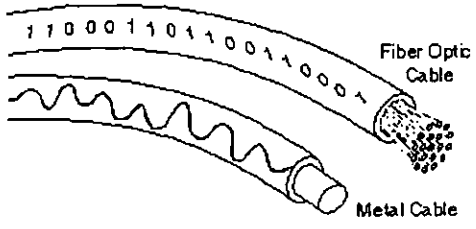


Figura 7.4 Fibra óptica

Todos estos enlaces varían en costo, ventajas y desventajas, y las diferencias en términos de susceptibilidad al daño, el sabotaje y las escuchas. Algunas clases de medios, como los pares trenzados, son baratos y fáciles de hacer funcionar, pero sensibles a las interferencias. Los cables coaxiales son algo mejores en términos de interferencias, y las fibras ópticas son incluso menos susceptibles a las interferencias electromagnéticas y otro tipo de interferencias. Por supuesto, como en cualquier tipo de cable físico, hay formas de cortar, o dañar de otra manera, dichos cables, o de escuchar o interceptar el flujo de datos que viaja a través de esos medios.

Por otra parte, los medios ilimitados, como las microondas, las ondas de radio y la transmisión de rayos infrarrojos, tienen problemas en sí mismos. Utilizados más frecuentemente en redes de área ancha y en operaciones nacionales/globales debido a su alto coste, estas tecnologías, que no usan cables físicos, no sólo están más sujetas a los problemas atmosféricos y climatológicos, sino que también son bastante sensibles a las interferencias externas, las escuchas y a la ocupación de frecuencias. Los medios ilimitados, aunque están libres de las limitaciones de los cables fisicos, pueden crear problemas de seguridad si la información que se transmite es altamente confidencial aún si es que todos los aspectos restantes de la red hayan sido asegurados.

7.7 Cortafuegos

Cortafuegos

Las redes, particularmente grandes, interconexiones de redes (muchas redes pequeñas conectadas entre sí), consisten en un gran sistema interconectado compuesto de hardware, software, recursos y datos.

Un cortafuegos o firewall es un sistema de defensa basado en el hecho de que todo el tráfico de entrada o salida a la red debe pasar obligatoriamente por un sistema de seguridad capaz de autorizar, denegar, y tomar nota de aquello que ocurre en la red.

Un firewall consiste en un conjunto de medidas HARDWARE y SOFTWARE destinadas a asegura una instalación de red.

Un Firewall actúa en los niveles 3 (red) a 7 (aplicación) de OSI. Sus funciones son básicamente las siguientes:

1. Llevar contabilidad de las transacciones realizadas en la red.

2. Filtrar accesos no autorizados a máquinas (mediante filtrado de paquetes, o bien observando el contenido de las unidades de protocolo de Transporte, Sesión, Presentación, y Aplicación).

3. Lanzar alertas en caso de ataques o comportamiento extraño de las comunicaciones.

Cualquier Firewall puede clasificarse dentro de uno de los tipos siguientes (o como una combinación de los mismos):

Filtros (Paket Filters)

Su cometido consiste en *filtrar* paquetes dejando pasar por el tamiz únicamente cierto tipo de tráfico.

Poxy (Circuit Gateways)

En este caso la pasarela actúa del mismo modo que un simple cable (vía software) conectando nuestra red interna con el exterior. En general se requiere que el usuario esté autorizado para acceder al exterior y que tenga una cuenta de salida en el proxy.

Pasarelas a nivel de Aplicación (Application Gateway)

Estas pasarelas se ocupan de comprobar que los protocolos a nivel de aplicación (ftp,http, etc.) se están utilizando de forma correcta sin tratar de explotar algunos problemas que pudiese tener el software de red.

El funcionamiento de un cortafuegos se basa en el principio de que no todos los segmentos del sistema deberían ser necesariamente una enorme «super-red». Más bien, las diferentes redes se mantienen separadas, pero comunicadas a través de ruteadores y pasarelas. La implantación de un cortafuegos pondría límites a los datos que pueden «viajar» a través de la «puerta» del cortafuegos y entrar en otra red. Igual que un guardia de seguridad en un complejo corporativo, verifica los ID de cualquier usuario que procede de una red adyacente.

Por otra parte, la barrera funciona como un guardia de seguridad, verificando «pasaportes» y otros documentos de la información transmitida. Pasará la información a través de ella si la fuente o el destino es la puerta, y la bloqueará en caso contrario. Puede realizarse un filtrado más selectivo, tal como paquetes o mensajes de cierto tipo.

7.8 Identificación y confidencialidad en las redes

A continuación se explican las técnicas criptográficas utilizadas para proporcionar identificación y confidencialidad. Se debe identificar a las personas o empresas con las que se tiene relación a través de la redes y asegurar que la información confidencial no es revelada erróneamente a personas no autorizadas. De esta inquietud nació el algoritomo internacional de cifrado de datos (International Data Encripyption Algorithm, IDEA).

IDEA fue desarrollado por Xuejia Lai y James L. Massey en el Instituto Tecnológico

Federal Suizo. Nació en 1990 y fue mejorado posteriormente en 1991. IDEA es un algoritmo de cifrado convencional de clave secreta que utiliza una clave secreta para cifrar un bloque de datos de 64 bits. Se emplea la misma clave para descifrar el bloque de texto cifrado de 64 bits y para recuperar el bloque de 64 bits de texto claro original. IDEA utiliza claves de 128 bits (16 bytes), valor sobre el que todos los analistas criptográficos coinciden en considerar debería ser suficiente por muchos años.
El algoritmo IDEA aplica una serie de fases que conllevan la generación de una subclave para cada fase a partir de la totalidad de la clave de cifrado y, al igual que DES, se emplea una función denominada idesmetiuzadora! para desordenar los bits durante cada fase.

Para que el remitente pueda desencriptar el mensaje enviado, debe disponer del algoritmo adecuado, que debe haber acordado previamente con el emisor de manera que sea secreto y pueda cambiar con cada mensaje. Para ello, se utilizan unos códigos llamados llaves, que pueden ser públicas o privadas.

La encriptación por llave privada o encriptación simétrica, es el mecanismo clásico. Se basa en la utilización de la misma llave por los dos extremos, es decir, usan una clave K que es conocida por el remitente de los mensajes y por el receptor, y con la que cifran y descifran respectivamente el mensaje. Para mantener la seguridad del cifrado, deben mantener esta clave en secreto. Las ventajas del uso de estas claves es la existencia algoritmos muy rápidos y eficientes para su cálculo. Si K es lo bastante larga (típicamente se usan valores de 56 a 128 bits), es imposible reventarlas usando la fuerza bruta. El principal inconveniente estriba en la necesidad de que todas las partes conozcan K, lo que lleva a problemas en la distribución de las claves. No parece muy coherente el envío dela llave utilizada para la seguridad través de una red no protegida.

El algoritmo más conocido de este tipo es el DES (Data Encryption Standard)desarrollado por IBM y adoptado por las oficinas gubernamentales estadounidenses para protección de datos desde 1977. Una mejora de este sistema de cifrado de clave simétrica es IDEA. También es conocido Kerberos, un sistema de autentificación diseñado en el MIT, con dos propósitos: proveer autentificación y distribuir claves. El sistema Kerberos actúa como autoridad de certificación que garantiza una relación correcta entre claves y usuarios o entidades.

La encriptación asimétrica surge por la debilidad del anterior sistema, se utilizan las llamadas llaves públicas. Existen dos llaves: una privada y una pública (distribuida por su propietario por toda la Red), el que envía utiliza la llave pública del destinatario para codificar el mensaje y el destinatario utiliza su llave privada para descifrarlo. Las dos claves están relacionadas matemáticamente, pero es casi imposible derivar la privada de la pública. El algoritmo que posibilitó la utilizaciónd dela llave pública fue el RSA (Rivest Shamir Adleman, sus creadores), y es aún hoy la base de la mayoría de sistemas de encriptación y autentificación que se utilizan. El principal inconveniente de este sistema es la existencia de una patente sobre este algoritmo, lo cual dificulta su uso fuera de los EE.UU. si no se ha obtenido la correspondiente licencia de exportación.

Este método elimina un problema pero introduce dos. En primer lugar, se necesitará encontrar la clave pública de una persona para encriptar el mensaje y, segundo, la encriptación y desencriptación por este método resultan muy lentas. El primer problema está siendo tratado con un nuevo servicio de directorios de Internet llamado Lightweight Directory Access Protocol (LDAP), es un soporte que sirve para hacer a bases de datos de llaves públicas que algunas compañías están creando. Para solucionar el segundo problema, la mayoría de los programas de correo utilizan la encriptación simétrica para encriptar el contenido del mensaje y después encriptan la clave utilizando una encriptación asimétrica; este enfoque intermedio, a veces llamado sobre digital, ofrece la seguridad de la encriptación asimétrica y la velocidad de la simétrica. Así, cuando le llegue el mensaje al destinatario, primero con su llave privada desencripta la llave simétrica, que le sirve para desencriptar el mensaje.
PGP-Pretty Good Privacy (magnífica confidencialidad)

PGP, escrito por Phil Zimmermann, es más que una magnífica confidencialidad, ¡es una confidencialidad excepcionalmente excelente!
PGP es una herramienta automática para la transmisión y recepción, tanto del correo electrónico firmado digitalmente, como del cifrado. El proceso de identificación es gestionado mediante tecnología RSA de clave pública/clave privada, pudiendo el usuario seleccionar libremente la longitud de la clave, desde 500 hasta 2000 bits. El cifrado por clave secreta o clave simétrica empleado para los mensajes secretos es IDEA, con sus claves de 128 bits, y el algoritmo de síntesis de mensaje empleado para realizar las firmas digitales es MD5.
Es posible que el lector haya recibido mensajes de correo electrónico en Internet de apariencia similar a ésta:

-COMIENZO MENSAJE FIRMADO PGP

Este es el texto de un mensaje firmado digitalmente.
Usted puede probar que procede de mi y que está totalmente intacto.

-COMIENZO FIRMA PGP

lkjLKJKjjjL:kjkjKSFKs:LKJFKLDFJHDLKJDFKJDLKIuyjhgUHSFDDF
LKJDFKJDFLKJDFL
:KJDFHJUEJJHhjsdfijjjerilkdflkwejrash-
drfljkerelkjare
kkdfjhkKHDFJHDHDjJJEIOREDLER:ELRE:RL:ERLE:lre;wleweHHdklds
-FIN FIRMA PGP-

Si se dispone de PGP y de la clave pública del emisor en alguno de los «llaveros» PGP, se puede verificar automáticamente la autenticidad y la integridad del mensaje simplemente pulsando una tecla.

7.9 PGP y el web de la confianza

Con Kerberos, o cualquier otro distribuidor de claves centralizado, se dispone de una autoridad central en la que se puede confiar para saber si otro es ficticio o auténtico, y también para proporcionar una clave de sesión para comunicarse. Sin embargo, existen sistemas que no son centralizados como Internet.

La solución PGP se denomina web de la confianza. La idea básica es que la gente tenga llaveros con las claves públicas de otros usuarios. Si usted conoce personalmente a Alicia, entonces puede certificar su clave pública y, por tanto, confirmar que la clave de Alicia es auténtica. Si Roberto recibe la clave que usted ha acreditado, y Roberto confía en usted, entonces él también puede certificar que la clave de Alicia es auténtica, incluso si él no conoce a Alicia, porque usted ha acreditado la clave de Alicia y Roberto confía en usted. Por lo tanto, si Roberto le conoce, él certificará su clave pública, garantizando el hecho de que es realmente su clave. Así pues, existe un web de la confianza en expansión donde alguien que conoce a Roberto puede confiar en la clave de Alicia porque Roberto se hace responsable de usted y usted responde por Alicia.

De esta manera, se tiene gente certificando las claves de otras personas, quienes a su vez acreditan las claves de otras personas, y así sucesivamente. Por supuesto, existe también el concepto de grado de confianza. Es mucho más lógico confiar en una clave que ha sido directamente acreditada por muchas personas conocidas, más que en una certificada por una persona cuya clave ha sido acreditada por uno de los amigos del hermano de la mujer de un primo.

A menudo, la gente difunde sus clave públicas PGP mediante otros medios, por ejemplo, publicándolas a través de Internet en los servidores web de sus empresas.
PGP es la técnica cuyo uso está más generalizado en Internet para verificar tanto la identidad del emisor como la integridad de un mensaje de correo electrónico o un correo electrónico cifrado. La identificación se logra utilizando RSA para cifrar una síntesis de mensaje MD5 de un correo electrónico empleando la clave privada del emisor. Los mensajes cifrados se envían cifrando el mensaje de correo electrónico mediante el algoritmo por clave secreta IDEA y, a continuación, agrupando el mensaje cifrado y la clave IDEA cifrados con la clave pública del receptor.

En este capítulo abordamos de manera general la seguridad tanto física y lógica de las redes.

CAPITULO 8
MONITOREO DE REDES

8.1 Monitoreo de Redes

En esta época la información acerca de nuestras redes es un componente integral en el éxito de las empresas para crear una ventaja competitiva.

La arquitectura cliente-servidor ha fortalecido un nuevo nivel de flexibilidad y robustez en los servicios de cómputo, proveyendo una cantidad sin precedentes en equipos de cómputo, redes, software y soluciónes. Esta libertad también ha creado un cambio substancial en el manejo de diversas organizaciones, es por esto que los equipos de cómputo y los sistemas se le salen de control al administrador.

Debido a esta proliferación, los centros manejadores de datos requieren altos niveles de disponibilidad, control y administración.

La familia HP 9000 (Servidores Empresariales) entregan una plataforma ideal para la centralización de datos proveyendo un excelente desempeño y una relación precio/desempeño muy buena. Con la combinación de la plataforma de monitoreo OpenView para la administración de soluciones, alcanzamos un alto nivel de funcionalidad y disponibilidad que cualquier centro de datos dentro de una empresa necesita. Aseguramos que los servidores, bases de datos, aplicaciones e infraestructura de redes estén funcionando correctamente. Esta plataforma también nos ayuda a la medición de los objetivos, así como también a monitorear el comportamiento y el desempeño de la red. En el mercado existen diferentes plataformas de monitoreo como se muestra en la Tabla 8.1:

Marca	Nombre
Hewlett Packard	HP OpenView
IBM	Tivoli, Patrol
Cabletron Systems	Spectrum Network Products
Computers Associates	CA Unicenter TNG x.x

Tabla 8.1 Dueños de productos

A lo largo de este trabajo hablaremos en particular de la Plataforma HP OpenView, así como también mencionaremos características principales de los otros productos.

8.2 HP OpenView

HP OpenView es un portafolio integrado de productos de software, de metodologías diseñadas, e implementación de servicios que proveen el acercamiento más compresible de los servicios de administración.

HP OpenView tiene como característica la administración centralizada, la automatización y un excelente control de los costos de administración, todos trabajando junto para lograr una fuerte contribución para mejorar la productividad y reducir el costro de propiedad de sus instalaciones.

Las soluciónes HP OpenView están basadas en tecnologías abiertas que pueden manejar ambientes globales distribuidos.

HP OpenView ofrece una familia de sistemas ejecutando monitoreos y herramientas de diagnóstico que toman ventaja del sistema operativo HP UX.

La plataforma NNM (Network Node Manager) es utilizada por todas las industrias para manejar redes, sistemas, aplicaciones y bases de datos. NNM, como una aplicación, está dirigida a los mercadosTCP/IP, Nivel 2 e IPX (sólo NT) .

8.2.1 Breve sumario de módulos de OpenView

La solución express de openview consiste en 4 módulos principales:

1. Administrador del sistema y aplicaciones
2. Administrador de redes
3. Administrador de almacenamiento
4. Integración

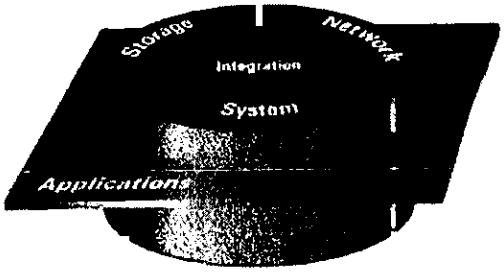


Figura 8.1.1 Solución express de openview

- **Access Manager (Administrador de Acceso):** Se utiliza para la autentificación de derechos de acceso para empleados, asociando a éstos a sus necesidades. Este módulo provee una configuración visual, monitoreando y generando reportes del empleado, acerca de como maneja los recursos de IT (Information Technology). Esto le facilita a los administradores de IT tomar decisiones.

- **Asset View (Vista de cualidades):** Es un software preciso y de fácil comprensión que puede liberar la infromación que necesitamos para manejar de manera efectiva las cualidades de la empresa. Utiliza varios programas integrados para colectar y presentar datos, así como proveer ligas para otros sistemas y mantenimiento de los datos. Emplea una modelo relacional de base de datos, permitiendo a los usuarios la opción de ver y combinar diferentes cualidades, para encontrar la solución adecuada para la empresa.

- **Customer Views (Vistas del Cliente):** HP OpenView Customer Views para Network Node Manager emerge del poder de éste con información del cliente para proveer un manejo inteligente de ambientes de red para ISP. Provee 5 vistas nuevas de recursos de la red para habilitar a los administradores de redes administrar la red de una manera consistente a las necesidades de la empresa. Estas 5 vistas nos proveen de multiples vistas jerárquicas de los recursos y las relaciones de la red. Customer Views organiza los recursos de la red por locación, agrupándola de manera geográfica y por tipo de red.

- **Desktop Administrator (Administrador de escritorio):** Ayuda a los administradores a tener el control sobre el software. Provee herramientas para ejecutar inventarios de software y hardware, distribución electrónica de software, mantenimiento de licencias y configuración remota de una variedad de sistemas opertivos de red (NOS) y sistemas operativos de PC's.

- **GlancePlus (Vistazo):** Es un poderoso sistema de monitoreo y herramienta de diagnóstico. Provee también información del desempeño de su sistema. Nos permite examinar las actividades del sistema, identificar y resolver los cuellos de botella y de esta manera cambiar sus sistema para tener un sistema más eficiente.

- **Internet Service Manager (Administrador de servicios de Internet):** Nos permite tener una vista integral de los servicios de Internet.

- **IT/Administrator (Adminstrador de IT):** Es una poderosa solución para tener una administración centralizada.

- **IT Service Manager (Adminstrador de sistemas de IT):** Automatiza los procesos de adminstración de IT en complejos ambientes para controlar la calidad y liberar los servicios de misión crítica.

- **ManageX**: Es el sistema más avanzado de administración de soluciones para asegurar el perfecto comportamiento de los sistemas de misión crítica. Diseñado únicamente desde el punto de vista del administrador, Manage X provee adminstración centralizada de los ambientes distribuidos NT y UNIX. Este tipo de adminstración nos baja el costo total de pertenencia (Total Cost of Ownership).

- **NetMetrix:** Provee detalles de la actural red y monitoreo y análisis de las aplicaciones y de los colectores de datos.

- **Omniback II:** La seguridad de los datos de misión crítica es vital es las operaciones de negocios. Omniback II nos provee una completo servicio protección de datos sin interrupción del sistema. Es una solución de almacenamiento cliente/servidor que incluye planeacion, configuración y soluciónes de backup y restauración de servicios, bakcups en líneas . Una característica única de la solución para HP 9000, Ominback II tiene una especial integración el file system de HP-UX, la base de datos de Oracle y los dispositivos de almacenamiento EMC, los cuales tienen cero impacto en los respaldos mientras las aplicaciones de la empresa siguen ejecutándose.

- **Service Simulator (Simulador de servicios):** Este servicio para redes es rápido, es una poderosa simulación y aplicación de diseño que de manera precisa predicen el impacto de la red y la aplicación cambia en un ambiente dinámico de red.

8.2.2 Beneficios obtenidos al usar OpenView

1. Saber como se ve la red.
2. Predecir problemas en la red.
3. Adminstración proactiva de la red.
4. Acceso remoto de herramientas de administración de la red.
5. Administración global de la empresa.

8.2.3 Administración de los servidores HP 9000

HP Openview ofrece una familia de sistemas para el monitoreo y herramientas de diagnóstico para tomar ventaja de la rica instrumentación HP-UX

8.2.4 Soluciones actuales para direccionar Alta disponibilidad soportando NNM (Network Node Manager)

MC/ServiceGuard es una solución de alta disponibilidad creada por HP. Esta solución está diseñada para monitorear los servidores HP 9000 series 500-800 para el sistema, los procesos y las fallas de la red (LAN).

Esta solución consiste y en un sistema primario y uno secundario en un cluster para permitir a las aplicaciones en caso de falla se migre de un sistema primario a un secundario. De esta manera en caso de falla en el hardware en el sistema primario la aplicación se podrá correr en un sistema secundario en un tiempo mínimo downtime del sistema.

En un ambiente de MC/ServiceGuard el NNM corre en el primer sistema. Para que el sistema continue trabajando se le asigna una dirección flotante IP a los procesos de NNM. NNM utiliza la dirección IP para monitorear la red e interactuar con las estaciones de recolección de datos del NNM. En caso de una falla en el sistema primario, los procesos del NNM y la conexión a la LAN son movidos al secundario sistema (Failover o Standby).

El sistema secundario adquiere la dirección flotante IP del NNM, activa los discos compartido y comienza los procesos de NNM. Los discos compartidos son sólo accesados por un sistema al mismo tiempo, sistema secundario o primario, a pesar de que los discos estén conectados a ambos sistemas.

Existen diferentes soluciones de alta disponibilidad en el mercado:

- Wolfpack es una solución para alta dsponibilidad de Microsoft para Windows NT.
- Sun también tiene una solución en Solaris que provee alta disponibilidad en platafomas Sun Solaris.

8.3 Productos IBM

En 1992, IBM embarcó su primera versión de TME 10 NetView. Digital e IBM en 1993 firmaron un acuerdo en el cual Digital aceptó la licencia de NetView. Este producto es llamado POLYCENTER.

Sin embargo, en enero 30 de 1996, IBM compro Tivoli Systems. Este sistema de manejo de ambiernte es producto de manejo de sistemas distribuidos. Tivoli tiene una tecnología basada en oricntación a objetos. Sin embargo, la integración de las dos diferentes tecnologías no es fácil. IBM tiene basada su estrategia en NetView como su plataforma base. IBM está tratando de incorporar IBM NetView con IBM Tivoli.

Este producto actúa como un manejador de SNMP para proveer descubrimientos distribuidos, status de las colas de impresión, recolección de datos.

8.4 Software de Redes Spectrum

En Junio de 1999 Cabletron Systems Inc. Anunció que se dividiría en dos empresas diferentes, para así lograr un mayor acercamiento a sus competidores en el ramo del software para monitoreo de redes. Es así como lanza al mercado SPECTRUM.

El producto básico de SPECTRUM consiste en varios componentes los principales son:

1. SpectroServer en la herramienta para monitoreo de servidores de SPECTRUM 5.0
2. Spectrograph, es la interfaz gráfica para SPECTRUM 5.0. Esta herramienta se puede ejecutar en una sistema diferente a SPECTRUM (SpectroServer).

Más de diez usuarios de spectograph se pueden conectar a un solo SpectroServer. La interface entre el GUI y el sevidor está soportada en enlaces WAN. Un GUI puede acceder a múltiples SpectroServers; sin embargo la consola ve vistas diferentes y ventanas diferentes para cada uno de los servidores. No hay combinación entre los servidores. No hay una vista común de la topología ni tampoco un buscador (browser) común. Es como si se tuvieran múltiples SPECTRUMS corriendo en una sola máquina, cada uno con su propia vista e interfaz, pero no existe interacción entre ellos.

Nos vemos en la necesidad de forzar a tener ambientes separados, y de esta manera tener ambientes más pequeños sobre servidores separados.

El SpectroServer y el Spectograph, consisten en un número diferente de aplicaciones, módulos, y herramientas. Como Autodiscovery, Spectro WATCH, y un generador de reportes de Spectrum.

Una de la fortalezas de SPECTRUM es un Tecnología de Modelado Inductiva (IMT). Este sistema de modelado es flexible y poderoso. Sin embargo, el resultado de esta flexibilidad se basa en una de sus más grandes debilidades, en un sistema difícil de configurar, ya que se necesita programarlo en lenguaje C++ y un conocimiento profundo de modelos.

Las características principales del Spectrum son:

- Editor de eventos, centraliza los eventos, alarmas, etcl.
- Event log
- Tolerancia a fallas
- Spectro Watch: Atributo de los logging
- Aislamiento de fallas
- Manejeador de Notificación de alarmas
- Agenda: Ejecución automática de eventos
- Visor de alarmas de la red.

Cabletron Systems también tiene un producto llamado **Bluevision** para el manejo de ambientes SNA. También está soportado en Sun Solaris, Microsoft Windows NT.

8.5 Productos de Computer Associates

Computer Associates tiene a la venta un producto para monitoreo de redes llamado TNG. El unicenter TNG es la base para el manejo de la red incluyendo: Descubrimiento, Mapas, Consola de eventos, Vista de objetos, Browser de DMI (Desktop Manager Interface).

CAPITULO 9
PLAN DE CONTINGENCIA

9.1 Plan de contingencia.

Un plan de contingencia es el proceso de determinar qué hacer si una catástrofe se abate sobre la empresa y es necesario recuperar la red y los sistemas.

La reanudación de la actividad normal ante un desastre puede ser una de las actividades más difíciles con las que un administrador de sistemas se ha de enfrentar. Tras un desastre, es probable que no haya posibilidades de regresar al lugar de trabajo habitual o que no se disponga de ninguna de las herramientas de administración de sistemas usuales para realizar el trabajo. Incluso, es posible tener que hacer el trabajo sin el equipo de administración o colaboradores. La preparación es la clave cuando todo está en contra.

No existe manera alguna de proteger completamente los datos contra todo tipo de amenazas. Por lo tanto, es prudente reflexionar sobre lo que sucedería en el caso de que un desastre ocurriera y nos encontráramos sin ningún tipo de acceso a la red del lugar de trabajo.

Con la cantidad de trabajo que la mayoría de los administradores tienen, el plan de contingencia tiende a dejarse para una ocasión posterior. Uno de los problemas asociados al plan de contingencia para redes es saber por dónde empezar.

Como todas las cosas que necesitan disciplina y práctica, restablecer un sistema de comunicaciones después de un desastre requiere práctica y análisis para tener aptitudes y poder realizarlo con un alto nivel de experiencia. Probablemente, llevó años diseñar y construir la actual red; de repente, será necesario reconstruir la red en unos días. Esto requerirá toda la pericia disponible para que sea un éxito. La presión de intentar recrear una red a partir de cero cuando otros, como el presidente y director general, están observando por encima del hombro, puede ser angustiosa; disponer de un proceso documentado en el cual basarse y al que se está habituado puede proporcionar la visión necesaria para salir airoso de esta situación.

La recuperación ante un desastre debería ser un ejercicio de equipo; sin embargo, resulta tremendamente difícil coordinar un equipo a no ser que todos sus miembros tengan como referencia un conjunto consistente de políticas y procedimientos. Cualquier procedimiento que forme parte del plan de contingencia deber ser comunicado de manera apremiante y clara a todos aquellos que necesiten conocerlo. Un plan de contingencia que cambia constantemente sin una comunicación efectiva es casi tan malo como no tener ningún plan.

El objetivo del proceso de generar un plan de contingencia es producir un documento denominado *plan de contingencia*. Proporciona la cohesión que permite al grupo de recuperación actuar como un equipo al adjudicar a cada miembro una lista concreta de responsabilidades y procedimientos a seguir.

La preparación ante un desastre comienza asegurándose de que se tienen los datos a recuperar. Un plan de contingencia no incluye necesariamente operaciones de copias de seguridad como parte de su contenido; sin embargo, la realización de copias de seguridad fiables debería ser un requisito previo del plan de contingencia; de no ser así, se está malgastando el tiempo pensando que es posible recuperar algo. Por lo tanto, al mismo tiempo que se considera el plan, éstas son algunas de las cosas que se pueden hacer si ocurre un desastre:

a) Realizar copias de seguridad todos los días y verificar su finalización.
b) Almacenar y renovar regularmente las cintas de copias de seguridad en una localización externa para asegurar la recuperación en el caso de un desastre en la instalación principal.
c) Familiarizarse con la recuperación de datos a través del sistema de copias de seguridad.

9.2 Metodología para el plan de contingencia.

Una vez controlados los aspectos de copias de seguridad y almacenamiento de datos, es el momento de reflexionar sobre lo que se necesitará cuando suceda un desastre. Existe una metodología general que puede emplearse para formalizar el proceso dentro de la organización. Aunque hay distintas variaciones, los conceptos básicos son los mismos, y a continuación se exponen.

1. Análisis de riesgos.
2. Valoración de riesgos.
3. Asignación de prioridades a las aplicaciones.
4. Establecimiento de los requerimientos de recuperación.
5. Elaboración de la documentación.
6. Verificación e implementación del plan.
7. Distribución y mantenimiento del plan.

9.2.1 Análisis de riesgos.

La primera fase del plan de contingencia, el análisis, nos sitúa en el lugar de un asesor de una compañía de seguros. En esta fase, la preocupación está relacionada con tres simples puntos:

a) ¿Qué está bajo riesgo?.
b) ¿Qué puede ir mal.?.
c) Probabilidad de que suceda.

¿Qué está bajo riesgo?.

En este primer punto, se necesita incorporar a todo el software (sistema operativo, base de datos, aplicaciones y utilerías) y todos los componentes de la red susceptibles a ser dañados, dando lugar a la pérdida de conexiones, computadoras o datos. Un diagrama de la arquitectura de todos los componentes del sistema de red facilitará la realización de un inventario de los elementos que pueden necesitar ser restituidos tras un desastre. No hay que olvidar que también el software necesita ser reemplazado , y que todos los productos software relevantes han de ser identificados.

Un inventario completo de la red muestra de manera clara la complejidad de ésta. Cualquiera que realice inventarios de componentes para redes, comprende los problemas en el seguimiento del hardware y software utilizado por los usuarios finales. Una omisión en el inventario fácilmente puede dar lugar a una recuperación fallida tras un desastre. El sistema de aplicación puede no encontrarse preparado para su uso si alguno de sus componentes no está disponible; en tal caso, es aconsejable estar constantemente a la expectativa de los nuevos elementos que pueden haberse olvidado.

Uno de los aspectos menos agradables a tener en cuenta, y que a menudo se pasa por alto, es que las personas esenciales se ven afectadas por el desastre y sea necesario recurrir a otras personas para realizar sus labores. Una formación diversificada en los sistemas dentro de la organización puede ayudar a reducir el impacto de la indisponibilidad de uno de los colaboradores.

¿Qué puede ir mal?.

La ley de Murphy nos ha proporcionado un conjunto de extraños e inesperados desastres.
Las clases más obvias de desastres son los desastres naturales que conllevan tormentas de todo tipo o los acontecimientos geológicos como terremotos o volcanes. Los propios incendios constituyen uno de los peores desastres posibles. El calor, el humo y el agua que rodea a los incendios son extremadamente perjudiciales para los sistemas de cómputo. Los dispositivos de almacenamiento se deterioran fácilmente debido a las altas temperaturas y

al humo. La eliminación de los residuos tóxicos tras el incendio en un centro de cómputo puede llevar meses.

Deben considerarse mecanismos alternos de acceso a la red en el caso de que, por alguna razón, sea imposible acceder al site, incluso aunque el site pueda estar de pie y operacional. Ejemplos de que sucesos que pueden impedir el acceso al interior del edificio son los accidentes químicos e industriales y las manifestaciones sociales.

Igual de perjudicial para la organización es la pérdida de equipos debido al robo, gente que destruye intencionalmente datos mediante su borrado o su modificación, los virus y los errores humanos.

Probabilidad de que suceda.

Si se tuviera una cantidad ilimitada de recursos y fuera posible protegerse contra todas las calamidades, este punto carecería de interés. Sin embargo, no se dispone de recursos infinitos; de hecho, los recursos son bastante escasos. Por lo tanto, se deben seleccionar los tipos de desastres contra los que se intentará protegerse. Obviamente, estos preciados recursos se querrán gastar en aquellos desastres que tengan la mayor probabilidad de afectar la organización.

9.2.2 Valoración de riesgos.

En este punto, la preocupación principal es comprender la cantidad de pérdida financiera que puede provocar la interrupción de los servicios de la red.

La valoración de riesgos es el proceso de determinar el costo para la organización de experimentar un desastre que afecte a la actividad empresarial. Por ejemplo, si la empresa realiza negocios a través de Internet, ¿cuál es el costo de tener el servidor web inhabilitado?.

Los costos de un desastre pueden clasificarse en las siguientes categorías:
a) Costos reales de reemplazar el equipo informático.
b) Costos de producción.
c) Costos por negocio perdido.
d) Costos de reputación.

El costo real de los equipos y el software es fácil de calcular, y depende de si se dispone de un buen inventario de todos los componentes de la red necesarios. Los costos de producción pueden determinarse midiendo la producción generada asociada a la red. La empresa tiene una correcta valoración de la cantidad de trabajo realizado diariamente y su

valor relativo. La pérdida de producción, debida a la interrupción de la red, puede ser calculada utilizando esta información.

Los costos por negocio perdido son los ingresos perdidos por las organizaciones de ventas y marketing cuando la red no está disponible. Si el sistema de solicitud de pedidos no funciona y la empresa sólo es capaz de procesar el 25% del volumen diario habitual de ventas, entonces se ha perdido el 75% de ese volumen de ventas.

Los costos de reputación son más difíciles de evaluar, sin embargo, sería deseable incluirlos en la valoración. Estos costos se producen cuando los clientes pierden la confianza en la empresa y se llevan su negocio a otro sitio. Los costos de reputación crecen cuando los retardos en el servicio a los clientes son más prolongados o frecuentes.

9.2.3 Asignación de prioridades a las aplicaciones.

Después de que acontezca un desastre y se inicie la recuperación de los sistemas, debe conocerse qué aplicaciones recuperar en primer lugar. No hay que perder el tiempo restaurando los datos y sistemas equivocados cuando la actividad comercial necesita primero sus aplicaciones esenciales.

Esto implica la necesidad de determinar por anticipado cuáles son las aplicaciones fundamentales del negocio. Desgraciadamente no todos los sistemas pueden ser el más importante; por lo tanto, es fundamental que la dirección ayude a determinar el orden en que los sistemas serán recuperados.

Es de esperar que esta información sea aceptada de buen agrado por todos los jefes de departamento. Independientemente de ello, el plan de contingencia debería incluir la lista de los sistemas y su prioridad.

Una vez conocido lo que se va a restaurar, debería disponerse de todo lo necesario para la disponibilidad de tales aplicaciones. Un sistema de aplicación en una red está compuesto por los sistemas servidores donde las aplicaciones almacenan sus datos, los sistemas de estaciones de trabajo que los procesan, la red que interconecta todo, y el software de las aplicaciones.

Las aplicaciones cliente/servidor o distribuidas añaden un nivel extra de complejidad al requerir que distintas partes de la aplicación residan en máquinas separadas

9.2.4 Establecimiento de los requerimientos de recuperación.

La clave de esta fase del proceso de elaboración del plan de migración es definir un periodo de tiempo aceptable y viable para lograr que la red esté de nuevo activa. Tal y como se ha planteado en la sección anterior, la preocupación básica debería ser disponer de las aplicaciones más importantes en primer lugar. El personal directivo de la organización deseará saber cuándo estarán sus aplicaciones funcionando para planificar las actividades de la compañía.

Es muy importante concederse una cantidad de tiempo adecuada y no realizar estimaciones poco realistas sobre las propias posibilidades. No es el deseo de nadie tener un grupo de personas alrededor esperando la finalización de las operaciones de recuperación; una distracción de este tipo probablemente perturbe las labores. El término para *este tiempo es tiempo de recuperación objetivo* o TRO (RTO, *Recovery Time Objective*). El TRO definido debe ser verificado para comprobar que es realista y factible, no sólo por uno mismo, sino por el resto de la organización, que puede ser requerido para realizar el trabajo.

Aplicaciones diferentes tendrán TRO diferentes. La figura 9.1 muestra seis aplicaciones ubicadas en tres servidores. El orden de recuperación de estas aplicaciones sería:

Servidor 2, Aplicación 4.
Servidor 2, Aplicación 3.
Servidor 1, Aplicación 1.
Servidor 3, Aplicación 6.
Servidor 1, Aplicación 2.
Servidor 3, Aplicación 5.

Es necesario asegurarse de que se dispone de tiempo para recuperar las cintas localizadas en la instalación de almacenamiento exterior para adquirir los sistemas necesarios.

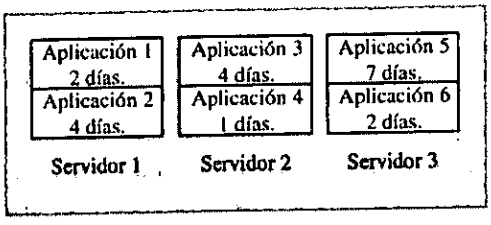


Figura 9.1. Sí es posible, realizar las operaciones de recuperación según la prioridad de las aplicaciones.

Es posible que sea necesario actualizar el sistema de copias de seguridad para satisfacer el TRO. Un sistema de cinta que recupera datos a 2 MB por segundo realizará la labor mucho más rápido que uno que lo ejecute a 500 KB por segundo. Hay que ser precavido y no suponer que se pueden hacer muchas cosas al mismo tiempo; uno se puede encontrar

cometiendo desafortunados errores que hacen más lenta la labor si no se presta atención al trabajo que se tiene entre manos.

9.2.5 Elaboración de la documentación.

Crear un documento que mucha gente pueda tener como referencia es el punto crítico del plan de contingencia. No hay que engañarse; implicará un esfuerzo significativo para algunas personas, pero ayudará a aprender cosas sobre el sistema y puede que algún día salve la empresa.

Uno de los problemas del plan de contingencia en un entorno de comunicaciones es que la tecnología de redes cambia tan rápidamente que resulta difícil permanecer al día. Esto incluye nuevos dispositivos, así como nuevos sistemas de aplicación que introducen su propio nivel de complejidad en este campo. Como ejemplo, considérese la recuperación de un gran sistema de base de datos relacional Unix. Este tipo de trabajo requiere un conocimiento mucho más complejo del que corresponde a la instalación de la base de datos y del que un administrador de red es probable que tenga; generalmente es necesario un administrador de base de datos, para el que también la labor será un desafío.

Dado el hecho de que la tecnología de red evoluciona tan rápidamente, debería planificarse la actualización del plan de contingencia periódicamente; por ejemplo, una vez al año. Aunque la redacción del plan inicial supondrá una gran cantidad de trabajo, una vez que se dispone del plan las actualizaciones son relativamente fáciles.

El plan de contingencia debe intentar definir las cinco áreas siguientes:

a) Listas de notificación, números de teléfono, mapas y direcciones.
b) Prioridades, responsabilidades, relaciones y procedimientos.
c) Información sobre adquisiciones y compras.
d) Diagramas de red.
e) Sistemas, configuraciones y copias de seguridad en cinta.

Hay que cerciorarse de que se sabe a quién notificar en primer lugar cuando ocurre un desastre. Si no se dispone de números de teléfono o direcciones actualizadas se puede pasar muy mal contactando con las personas afectadas.

Mapas mostrando las ubicaciones del centro de operaciones temporal y de la instalación externa pueden ahorrar mucho tiempo. También puede ser útil mostrar itinerarios alternativos de acceso para el caso de que las rutas principales no se encuentren disponibles.

Las personas deben disponer de instrucciones y responsabilidades precisas. La relación entre tareas debe hallarse documentada de manera que pueda identificarse cualquier cuello

de botella que pudiera surgir. Por último, deben incluirse, de manera detallada, las operaciones y tareas que muestren las labores de instalación y recuperación necesarias, debiendo ser fáciles de leer y seguir. También habrá que incluir aquí los números de teléfono de las organizaciones de asistencia que pudieran requerirse.

Como se ha mencionado anteriormente, debe saberse cómo expedir una solicitud de compra y obtener los equipos para el centro de operaciones temporal. Esto significa proporcionar a los vendedores la dirección y cualquier instrucción necesaria para el transporte. No hay que suponer que todos los vendedores del mundo van a enterarse de la difícil situación y venir a nuestro rescate. Es aconsejable disponer de copias de las facturas, recibos y demás para mostrarlos como prueba de compra.

También viene bien tener a mano una lista de los números de serie, modelo y características de los equipos de hardware; y para el software se debería de tener a mano la versión y licencias .

Los diagramas de red simplifican en gran medida la labor de reconstruir una red. Un diagrama detallado de la red, necesaria para las primeras aplicaciones, facilita y agiliza la reanudación de las actividades. La asignación de etiquetas a los cables y su almacenamiento en un lugar reservado, probablemente no llevará mucho tiempo y evitará muchas confusiones con posterioridad. La otra ventaja de un diagrama de conexiones es la posibilidad de emplear contratistas para realizar las instalaciones.

Es posible ahorrar horas o incluso días en el proceso de recuperación si existe la posibilidad de almacenar algunos sistemas de repuesto con la capacidad de administrar tareas diferentes. Planifíquese instalar una configuración genérica que, como mínimo, permita ejecutar las aplicaciones de mayor prioridad sin problemas.

Hay que asegurarse de la disponibilidad de un sistema de copias de seguridad de cinta en funcionamiento. Si es posible, debe mantenerse un sistema de reserva, incluyendo adaptadores SCSI, cables y software de unidades de dispositivo, en una ubicación exterior al lugar común de trabajo. No es inusual encontrarse con que los vendedores locales no disponen de existencias de los productos necesarios, obligando, por tanto, a esperar el envío de las piezas de recambio antes de poder empezar la recuperación de los datos. Si se sigue este consejo, no hay que olvidar actualizar este sistema cuando se actualicen los sistemas de copias de seguridad de producción; en caso contrario, uno se puede encontrar con formatos de cinta o bases de datos incompatibles u otros problemas que impedirán la restauración de la información.

9.2.6 Verificación e implementación del plan.

Una vez redactado el plan, hay que probarlo. Hay que estar seguro de que el plan va a funcionar. Para ello, se debe ser escéptico sobre el propio trabajo. De manera que pueda uno probarse a si mismo que funciona. Psicológicamente, esto no es fácil porque con toda

probabilidad se ha invertido una gran cantidad de tiempo y energía personal en este proceso, aunque lo mejor sería si es posible, situarse de manera imparcial ente la fiabilidad del plan. Por consiguientes, han de realizarse las pruebas para encontrar problemas, no para verificar que el plan funciona. Si existen errores en la información, tómese nota de ellos y corríjase el plan.

Compruébese el software para la realización de copias de seguridad para confirmar si pueden recuperarse las aplicaciones de mayor prioridad de la manera esperada. Esto debería hacerse con una red aislada para evitar problemas con el servidor de licencias. Una vez recuperada la información, verifíquese los usuarios pueden acceder a ella. Compruébese cada una de las operaciones del plan individualmente y examínese entonces si, como resultado, se tiene un sistema de red en funcionamiento.

9.2.7 Distribución y mantenimiento del plan.

Ya por último, cuando se disponga de un plan definitivo ya verificado, es necesario distribuirlo a las personas que necesitan tenerlo. Inténtese controlar las versiones del plan, de manera que no exista confusión con múltiples versiones. Así mismo, es necesario asegurar la disponibilidad de copias extras del plan para su depósito en la instalación exterior o en cualquier otro lugar además del lugar de trabajo. Cuando se actualice el plan, sustituya todas las copias y recoja las versiones previas.

El mantenimiento del plan es un proceso sencillo. Se comienza con una revisión del plan existente y se examina en su totalidad, realizando cambios a cualquier información que pueda haber variado. En ese instante, se debe volver a evaluar los sistemas de aplicación y determinar cuáles son los más importantes para la organización.

Este proceso llevará tiempo, pero posee algunos valiosos beneficios que se percibirán aunque nunca tengan que utilizarse. Más gente conocerá la red. Esto proporcionará a la organización una base técnica más amplia para mantener correctamente a la red.

CAPITULO 10
ALTA DISPONIBILIDAD

10.1 Alta Disponibilidad

Un sistema está altamente disponible si un simple componente o recurso falla y el sistema es interrumpido sólo por un breve periodo de tiempo.

Las interrupciones pueden ser vistas solo en ciertos niveles del sistema, los cuales pueden ser:

a) Usuarios,
b) Sistema Operativo,
c) Base de Datos,
d) Discos,
e) Nodo,
f) Etc.,

La aplicación debe estar disponible siempre, para el usuario final. No puede estar considerado un sistema altamente disponible el cual tiene como límite el detener sus aplicaciones para hacer un respaldo de Base de Datos o el mantenimiento preventivo del equipo; es decir, si el usuario final puede ver al sistema abajo en un tiempo X no es posible estar llamando al sistema altamente disponible. La tabla 10.1 muestra el porcentaje de disponibilidad de los sistemas en referencia a las horas año que se tiene fuera el sistema. Un sistema en cluster, con MC/Service Guard, puede llegar a tener una disponibilidad que va del 99.8% al 99.99%, lo anterior esta sujeto al diseño y a las aplicaciones que tiene que soportar.

Porcentaje de Disponibilidad	Tiempo al año en que el Sistema no esta Disponible
99.5%	44 horas
99.8%	18 horas
99.9%	9 horas
99.95%	4 horas
99.98%	2 horas
99.99%	1 hora
99.999%	5 minutos

Tabla 10.1 Disponibilidad de los Sistemas

10.2 Aplicaciones con Misión Crítica

Una aplicación crítica es aquella que es fundamental para el logro de los objetivos de la empresa, por tanto, en ella están fundamentados el éxito financiero de la misma. Una aplicación con misión crítica es aquella que al estar un tiempo sin operar contribuye, en gran medida, en pérdidas significativas en los valores de la empresa.

Entendemos como valores de la empresa todos aquellos elementos que contribuyen al éxito monetario y se conjugan en el logro de los objetivos. Un ejemplo práctico es el de un banco, el cual además de perder clientes por el descontento, deja de recaudar pagos de sus clientes al tener el sistema "DOWN".

10.3 Características de un Sistema en Alta Disponibilidad

Las características que pueden ser palpadas en todo sistema altamente disponible pueden ser resumidas de la siguiente manera:

a) Componentes de Hardware con características de alta Disponibilidad:

Dentro de los componentes de hardware necesarios para poder tener un sistema abierto altamente disponible existe una gran variedad los cuales deben ser elegidos de acuerdo a la aplicación crítica a manejar y el proveedor de Hardware y Software de nuestra elección. Cabe señalar que estos componentes varían en su precio de acuerdo a la efectividad que deseamos alcanzar y el proveedor de nuestra elección, sin embargo el gasto no es considerado tan grande si lo comparamos con el de la compra de un Mainframe los cuales tienen desde su nacimiento el nombre de Sistemas Altamente Disponibles ya que estos tienen redundancia inherentemente en sí mismos desde su arquitectura de fabricación. Como componentes de hardware con características de Alta Disponibilidad se pueden listar los siguientes:

Memorias con capacidad de autocorrección de errores, discos en espejo que soportan "Stripping" e información redundante, unidades de alimentación de energía redundantes, elementos de red redundantes, etc..

b) Aplicación de Alta Disponibilidad automática:

La aplicación, para soportar alta disponibilidad, debe tener la facilidad de ser configurada desde su implementación para poder detectar y autocorregir errores con la intervención de la parte humana en lo más mínimo posible. Lo anterior con la finalidad de aprovechar las velocidades en la ejecución de comandos ya previamente definidos mediante programas de arranque y autocorrección; evitando así, lentitud y errores humanos al intervenir en posibles contingencias de manera manual.

c) Diseño adecuado para explotar al máximo los atributos de los componentes de Hardware

El Cluster de Alta Disponibilidad debe estar bien estructurado y diseñado para soportar y aprovechar todos y cada uno de los componentes redundantes de hardware necesarios para la puesta en marcha de nuestro Sistema altamente disponible.

d) Fácil uso para la Operación y Administración de los equipos y las aplicaciones:

La manipulación de las aplicaciones y elementos que hacen al sistema altamente disponible no deben interferir ni estorbar al poner en práctica los procedimientos normales de operación, planes de recuperación o respaldo y mantenimientos preventivos y correctivos de los equipos. Si no por el contrario, debe proporcionar la facilidad de llevarlos a cabo sin tener que estar atendiendo a los usuarios molestos en la línea telefónica por que el sistema esta caído.

e) Diseño de las aplicaciones para hacer uso de ellas en Alta Disponibilidad

Las aplicaciones deberán ser desarrolladas para trabajar en conjunto con la aplicación de Alta Disponibilidad. Deben ser aplicaciones capaces de trabajar dentro de un Cluster altamente disponible. En este sentido, las aplicaciones deben ser de fácil arranque e instalación, de tal forma que en caso de contingencia la aplicación de Alta Disponibilidad pueda anular tiempo en el retraso de la actividad normal del sistema, haciendo las adecuaciones para que ésta opere de manera automática, y no tengan que hacerse manualmente por los operadores y administradores del Sistema.

10.4 Metas para la Implementación de un Sistema de Alta Disponibilidad

Deben ser considerados diversos factores antes de proponer el diseño de un sistema altamente disponible. Por tal motivo hay que hacer un análisis desde el punto de vista financiero pasando por la importancia del servicio que el sistema presta hasta llegar a determinar las características que el sistema debe cumplir y el hardware que es más adecuado. Por tanto para llegar a determinar el gasto y el nivel de disponibilidad del sistema se deben tomar muy en cuenta los siguientes puntos antes de llegar a una conclusión:

a) El costo actual del hardware para obtener un sistema de alta disponibilidad.

b) El retorno de la inversión contra el precio de tener el sistema caído por largos periodos de tiempo.

c) Los ingresos que se pierden por unidad de tiempo con un sistema poco confiable y de trabajo aleatorio.

d) La importancia en el servicio que el sistema presta a los consumidores finales, los cuales perdemos al estar mucho tiempo sin sistema.

e) Si los sistemas no funcionan las oficinas y los empleados no trabajan, lo cual implica pérdida monetaria en salarios no desquitados.

f) Multas a causa de prestar un servicio irregular.

g) Tiempos en la restauración de información corrupta o perdida.

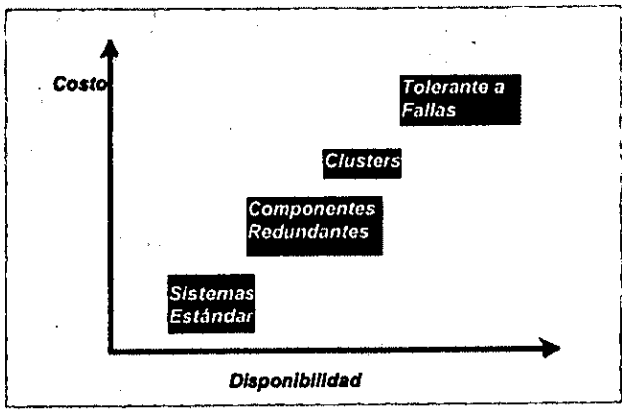
h) El trabajo acumulado que se debe atender después como consecuencia de una caída del sistema, cuando éste haya sido recuperado.

i) La importancia de satisfacer al cliente.

El costo de implementar una solución de alta disponibilidad en la organización inicia con el alto costo que implica el Hardware actualmente. La inversión regresa cuando los tiempos en la caída de los sistemas tiende acero. Actualmente las compañías requieren soluciones basadas en sistemas abiertos que requieren altos niveles de disponibilidad.

Las organizaciones con equipos mainframe están acostumbradas a tener alta disponibilidad en todas sus aplicaciones, pero también, existen empresas que desarrollan aplicaciones críticas en plataformas de sistemas abiertos, de ahí el nacimiento de soluciones para que estos sistemas mantengan las aplicaciones altamente disponibles.

En el mercado existen diversos fabricantes de hardware y software que proporcionan un ahorro monetario al ofrecer soluciones que pretenden hacer que un sistema abierto tenga casi la disponibilidad de un mainframe, con la diferencia de que su solución implica en costo un promedio de hasta diez veces menos el valor de la inversión comparado con el valor de una solución mainframe. La gráfica 10.1 muestra la relación Costo contra Disponibilidad en los sistemas.

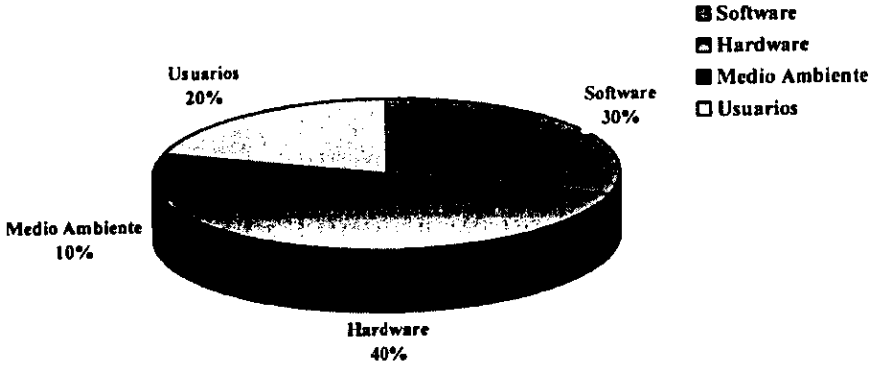


Gráfica 10.1 Costo vs Disponibilidad de los Sistemas

10.5 Causas de las Fallas

La causa de las fallas están relacionadas con los problemas descritos en el Capítulo 2, Sección 2.1.1 de esta Tesis y pueden ser agrupados de acuerdo a la gráfica 10.2.

El hardware es generalmente el principal responsable de que los sistemas fallen. Las fallas de software, generalmente, están íntimamente relacionados por "BUGS" en las aplicaciones o en el Sistema Operativo. También en este caso podemos considerar las ocasiones en que el descuido del administrador permite que los file system lleguen hasta un nivel de 100% de su capacidad de almacenamiento y lo mismo para los TABLESPACES de la Base de Datos.



Gráfica 10.2 Causas de las Fallas

Los usuarios, por descuido, por ser mal programadores o por ponerle pruebas al administrador, son causantes también de fallas. Lo anterior puede ser evitado implementando un adecuado nivel de seguridad, que considere una buena definición de cuotas, estructura de directorios, accesos, etc., y además se ajuste a las necesidades que se presentan. La revisión continua del administrador a sus usuarios en ocasiones resulta saludable.

El medio ambiente es un factor que en ocasiones no es muy considerado pero que sin embargo es un ente altamente destructible y que no debemos descuidar. Es necesario que también se analice, en el sentido geográfico, la ubicación de los sistemas evitando en lo más mínimo situar a los de Misión Crítica en lugares geográficamente propensos a regulares desastres naturales.

Existen fallas que no debemos olvidar, las cuales son escasas en México pero existen, los desastres por terrorismo. La falta de seguridad en los centros de cómputo se puede reflejar, además de terrorismo, en un incendio. La falta de clima controlado y el acceso de factores externos dañinos como lo es el polvo, lluvia, etc., también dañan los equipos. En México encontramos casos de saboteo en sistemas dado por aquel personal que algún día despedimos y continua teniendo accesos a nuestros sistemas.

En general, Hewlett Packard, base a la atención de reportes de soporte técnico, ha generado la gráfica 10.2 donde desglosa los causantes de las fallas y el porcentaje que representan del 100% de las mismas.

10.6 Clasificación de Fallas que Reducen la Disponibilidad de los Sistemas

Para poder evaluar de una mejor manera los problemas que reducen la efectividad de un sistema con misión crítica, podemos clasificarlos en tres grandes bloques:

a) CRISIS

Cuando los sistemas se encuentran en ejecución y existen condiciones que los limitan en su desempeño, se dice que el sistema está en un punto de crisis. Un ejemplo muy común es la falta de performance ya que dos aplicaciones están haciendo uso de un grupo de volúmenes al mismo tiempo e irracionalmente, el cual está constituido por cuatro discos físicos, además de que estos cuatro discos son manipulados por solo una tarjeta controladora. En este caso existe una crisis de performance y es ocasionado por el mal diseño y desconocimiento por parte del desarrollador. Otra causa de una crisis puede ser un cambio en los parámetros de inicialización de aplicaciones, Sistema Operativo o Bases de Datos. Los virus son también fundamentales para crear crisis en los Sistemas.

b) DISCONTINUIDAD

La discontinuidad se manifiesta cuando los sistemas quedan fuera por un periodo de tiempo, pero pueden ser restablecidos en un tiempo no muy significativo. La discontinuidad es producto, por lo general, de un reboot automático o manual después de que el equipo queda en un estado comúnmente llamado de "HALT", a causa de que uno de los componentes de Hardware están fallando. Otro factor causante de discontinuidades es la falta de energía en los equipos a causa del mal mantenimiento de los equipos que la proporcionan.

c) DESASTRE

Los sistemas en ocasiones pueden encontrarse fuera de servicio por un largo periodo de tiempo y las ganancias de la empresa se ven seriamente afectadas. Este tipo de discontinuidad, por un largo periodo de tiempo y con la posibilidad de no recuperar nunca el sistema, es producto de los desastres naturales ocasionados por huracanes, terremotos, etc., u otros por falta de seguridad como lo son los casos de destrucción por terrorismo o incendio. En ocasiones la falta de seguridad en los sistemas en el control de accesos a centos de cómputo y la inadecuada estructura organizacional de los usuarios pueden ocasionar desastres como un rm –r desde raíz en un sistema con plataforma UNIX o un proceso BATCH no controlado adecuadamente.

10.7 Hardware Sensible a Fallas y su Posible Solución con Componentes Altamente Disponibles

Un punto de falla es definido en ocasiones como la causa de que la aplicación no esté disponible para el usuario final. En ocasiones estas fallas no son detectadas de manera rápida ya que pueden no ser muy evidentes y requieren investigación y apoyo del proveedor.

Los productos de alta disponibilidad protegen de las fallas de hardware que no pueden ser determinadas en tiempo como son las de Disco y CPU. Las fallas que si pueden ser determinadas y prevenidas son las que ocasionan los "BUGS" de Sistema Operativo y de integridad de los datos con la adecuada instalación de parches recomendados por el proveedor y eficientes planes de respaldo por parte de los administradores.

10.8 Recuperación de Desastres

Disaster Recovery (Recuperación de Desastres) Es la habilidad que algunos sistemas tienen para recuperar la actividad cuando los desastres naturales se hacen presentes. Esta recuperación tiene la característica de que puede hacerse un breve periodo de tiempo si se diseña adecuadamente.

La única manera de poder recuperarnos de un Disaster Recovery es mediante la redundancia implementando un sistema espejo en un Centro de Cómputo ubicado en otro punto geográfico diferente al del nodo principal.

10.9 Opciones de Hardware para Alta Disponibilidad

Para alta disponibilidad existen diferentes componentes de hardware que pueden ayudar a implementar el cluster. En nuestro caso solo hablaremos de los disponibles por HP ya que al caso son los que nos interesan, sin olvidar que los demás proveedores tienen algunas opciones similares y con algunas variantes, en cuanto a tecnología principalmente se refiere.

Entre los componentes de Hardware, y herramientas para el análisis de componentes, se encuentran los siguientes proporcionados por HP:

a) Sistema inteligente de diagnósticos.
b) Tecnología de Discos soportada y sus características de aceptación de niveles de RAID, Stripping y Mirroring.
c) Redundancia recomendada en UPS (Uninterruptible Power Supply).
d) Redundancia de RED recomendada.

10.9.1 Sistema Inteligente de Diagnósticos

Esta herramienta tiene la capacidad de verificar, probar y corregir errores en los dispositivos de hardware para alta disponibilidad. Algunas de sus principales funciones se describen a continuación:

a)REVISION AUTOMATICA DE LAS PAGINAS DE MEMORIA:

Este producto tiene la capacidad de detectar y remover dinámicamente bloques dañados de memoria. En el momento de que los usuarios están haciendo uso de la memoria con sus procesos hace un diagnóstico, en background, de las páginas de memoria libres y de las usadas. Tiene la habilidad de proveer de reportes para consulta del administrador de hardware.

b)SOPORTE PREDICTIVO

Este producto tiene la capacidad de analizar los logs de errores y diagnóstico de las tarjetas controladoras de dispositivos como discos, componentes de entrada y salida de información, memoria cache, etc., entre otros. Proactivamente detecta potenciales daños en los elementos del hardware que pueden fallar en un futuro no muy lejano. Mediante este producto puede ser notificado vía modem que se esta teniendo un problema de hardware al personal de soporte técnico.

10.9.2 Niveles de Raid, Stripping y Mirroring.

Una técnica para proveer redundancia en el almacenamiento de discos es el uso de discos configurados en RAID para la protección de los datos. El significado de RAID es: "Redundant array of inexpensive disks". Se le denomina arreglo de discos a un grupo de discos que trabajan en conjunto de acuerdo al nivel de "RAID" o "mirroring" que se les configure y soporten. Algunos niveles proporcionan el aseguramiento de la información mediante el "mirroring", mientras que otros lo hacen mediante el uso de la paridad de los datos y algunos otros con la combinación del "mirroring" y el "RAID", para poder reconstruir los datos perdidos si un arreglo de discos falla.

"Stripping". Se le da este nombre a la acción de dividir la información a guardar en 2 ó más discos, con la finalidad de acelerar el proceso de escritura sin considerar la protección de la información. La figura 10.1, muestra un esquema de tres discos en stripping.

"Mirroring". Es la acción de duplicar la información en otros dispositivos de almacenamiento para protección de la misma en caso de que un, de al menos dos dispositivos, falle. La figura 10.2, muestra un esquema de tres discos en stripping y mirroring.

"Nivel de RAID". Se denomina nivel de RAID a la acción de tener redundancia de la información en un arreglo de discos sin necesariamente estar en configuración "mirroring".

10.9.2.1 Comunes Niveles de Raid.

Los niveles de RAID comúnmente más utilizados, son los mostrados en la tabla 10.2:

Niveles de RAID	Descripción.
0	El controlador escribe la información en múltiples discos en "stripes". Este nivel no provee de protección a la información.
1	El controlador escribe la información en más de un disco. Es decir, hace una copia fiel de un disco en otro disco. El performance es seriamente afectado aunque nuestra información esta bien resguardada.
3	Los datos están en "striped", y el controlador almacena la paridad en un disco por separado para poder reconstruir cualquiera de los discos del arreglo que pudieran perderse. Esta configuración se da con tres discos o más.
5	Los datos están en "striped", y el controlador almacena la paridad de la información en todos los discos que son miembros del arreglo para que cualquiera que se pierda de ellos pueda ser arreglado con la información de los demás. Esta configuración se da con tres discos o más.

Tabla 10.2 Niveles de RAID

10.9.2.2 Ventajas y Desventajas de los Niveles de Raid

Su poniendo que tenemos un arreglo de tres discos, la ganancia y costo que nos produce se resume a continuación:

RAID 0:

La información sólo esta en stripping. Es decir, la información se divide en los tres discos y no se tiene redundancia ni seguridad de poder recuperar la información, como lo muestra la figura 10.1. La única ganancia con este esquema es que la escritura y lectura de nuestra información se puede hacer de una manera más rápida y el desperdicio por salvaguardar la información es del 0%.

RAID 1:

Este caso implica tener por cada disco de almacenamiento uno para el espejo. Es decir, en el ejemplo práctico necesitamos tener seis discos para poder asignarle a cada uno de ellos un disco para el espejeo de su información. Este esquema no permite que la información sea dividida en más de un disco. La ventaja de este esquema es que en caso de la pérdida de uno de los discos, la información puede ser recuperada de su espejo. La desventaja es que se tiene un desperdicio de disco del 50%. El caso esta mostrado en la figura10.2.

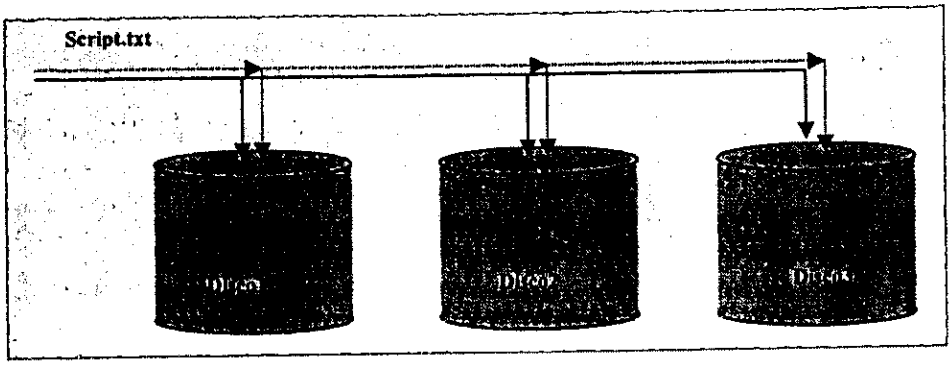


Figura 10.1 Stripping (RAID 0)

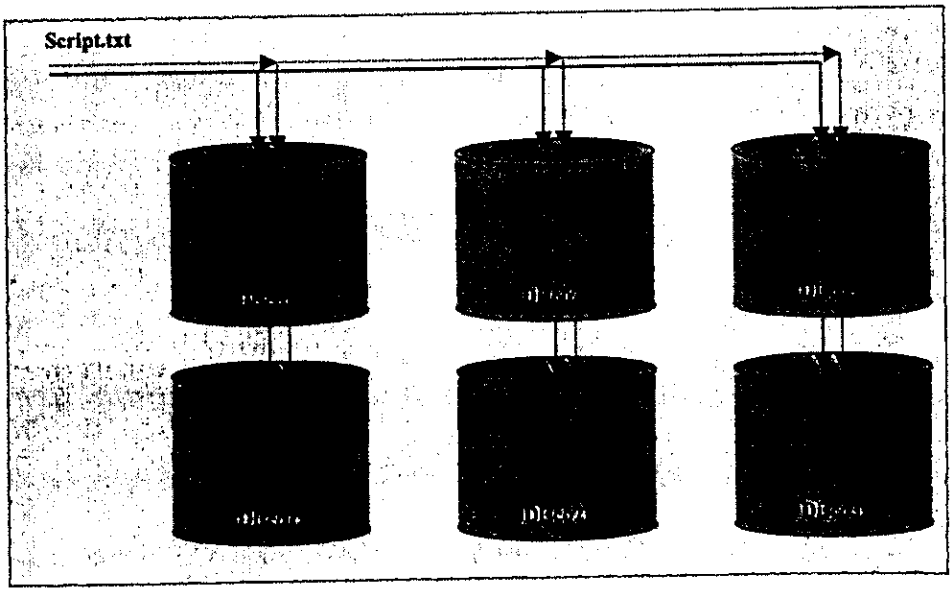


Figura 10.2 Mirroring (RAID 2)

RAID 3:

En este caso práctico, de los tres discos, la información se guardaría en "stripping" en dos de ellos y la paridad para recuperarla en el tercer disco. La ventaja de este esquema es la rapidez en la escritura por el "stripping", la recuperación por de uno de los discos, en este caso el de paridad. La desventaja es que si el disco de paridad se pierde además de otro que contiene la información ya no podemos recuperar el arreglo. El desperdicio en disco sería de un 33%. La paridad está dada por operaciones binarias hechas byte a byte de información escrita en el disco, a través de un algoritmo. Este esquema permite obtener un muy buen performance. La figura 10.3 es un caso muy extremo en donde se combina el RAID 3 con "mirror" y existe un desperdicio de espacio del 88%.

RAID 5:

El esquema permite el "stripping" en los tres discos, además, de almacenar la paridad de la información en los tres al mismo tiempo. La ventaja es que podemos recuperar la información de un disco con la paridad guardada en los otros dos. Las desventajas es que si más de un disco se pierde no podemos reconstruir la información. El desperdicio en disco de este esquema es del 33%. La paridad está dada por operaciones binarias hecha entre bloques de información escrita en el disco, a través de un algoritmo. Este esquema permite obtener un mejor performance, comparado con los otros niveles incluyendo RAID 3, ya que la paridad es obtenida a nivel bloque y no a nivel byte, además de que, es escrita más rápidamente en tres discos que en uno solo. La figura 10.4 muestra RAID 5 combinado con "mirror" en el cual se tiene un desperdicio del 88% del espacio.

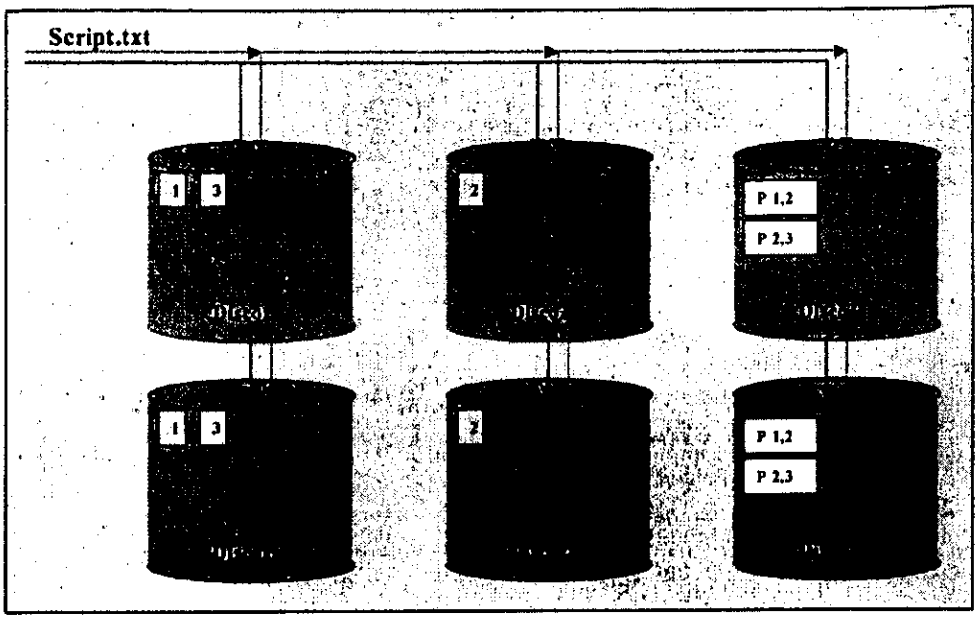


Figura 10.3 RAID 3 (BYTE)

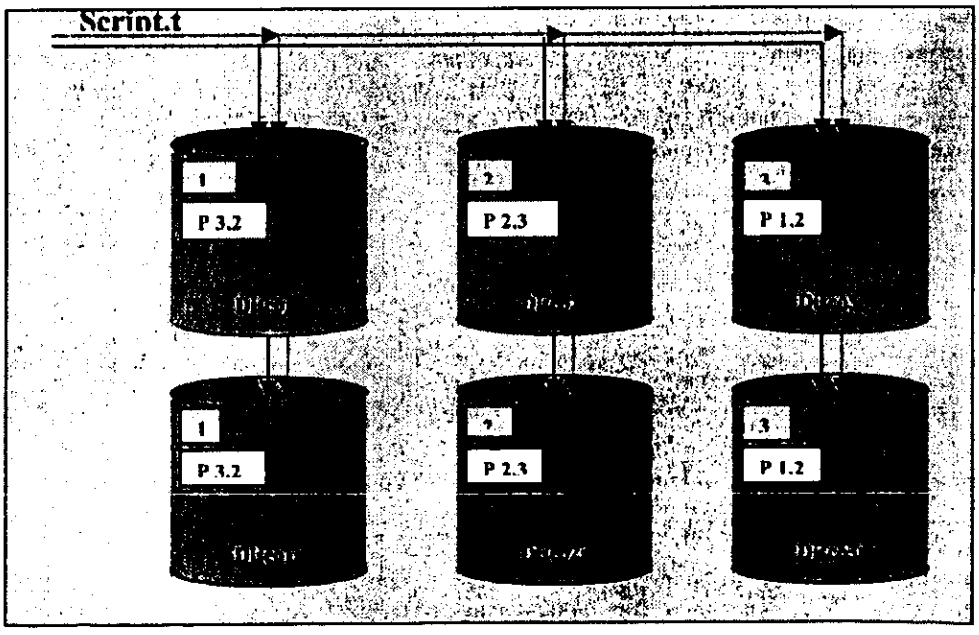


Figura 10.4 RAID 5 (BLOCK)

10.9.3 Redundancia recomendada en UPS

Si el centro de cómputo se encuentra diseñado para tener redundancia en sus UPS, para en el caso de que uno falle entre de manera automática otro, o en términos de un caos, los UPS fallen y entren de manera automática plantas de diesel, no es necesario que los equipos cuenten con UPS locales. Si el anterior no es el caso, se debe contar con al menos un UPS en el centro de cómputo y un local para cada equipo del cluster, como mínimo.

10.9.4 Redundancia de RED recomendada.

Es necesario considerar, para tener un cluster confiable, que existan al menos tres direcciones de RED configuradas, para poder tener un esquema de que una de ellas sea backup de la otra en caso de que una tarjeta de red falle, el cable se rompa, etc.

De la misma manera debe haber redundancia en los ruters para que si uno falla entre automáticamente el otro.

10.10 Arquitecturas de Alta Disponibilidad.

Para comenzar esta parte, es necesario definir un conjunto de elementos los cuales se les llama SPU (System Process Unit) en HP9000, y están sujetos a fallas de discontinuidad. El SPU está constituido de tres elementos principalmente y representan importantes puntos de falla:

a) Uno o más unidades de procesamiento. (CPUs)

b) Controladores de I/O

c) Tarjetas de Memoria

El software para implantar un CLUSTER de Alta Disponibilidad, debe permitir monitorear las aplicaciones activas en cada uno de los nodos y determina, de manera automática, en qué momento una de ellas está fuera de funcionamiento. El diseñador debe configurar la herramienta para que automáticamente intente levantar las aplicaciones que quedan no disponibles y el número de re intentos que tendrá que hacer, la aplicación de Alta Disponibilidad, antes de migrar los paquetes hacia el otro nodo del CLUSTER.

Las arquitecturas posibles entre los elementos de un CLUSTER pueden estar dadas de las siguientes maneras, aclarando que el grado de complejidad aumenta adicionando más nodos al CLUSTER:

a) Active/Standby

Uno de los elementos del cluster es configurado, de tal manera, que en caso de una falla en uno de los componentes de su SPU o aplicaciones configuradas como críticas, el otro nodo del Cluster comience a tomar el trabajo que éste deja pendiente. Cuando el nodo en "Standby" pasa a ser el activo el que era originalmente activo pasa a ser el nodo "Standby". En el caso de Active/Standby el nodo en Standby está en espera de que el Active falle para tomar su trabajo. El esquema Active/Standby se muestra en la figura 10.5.

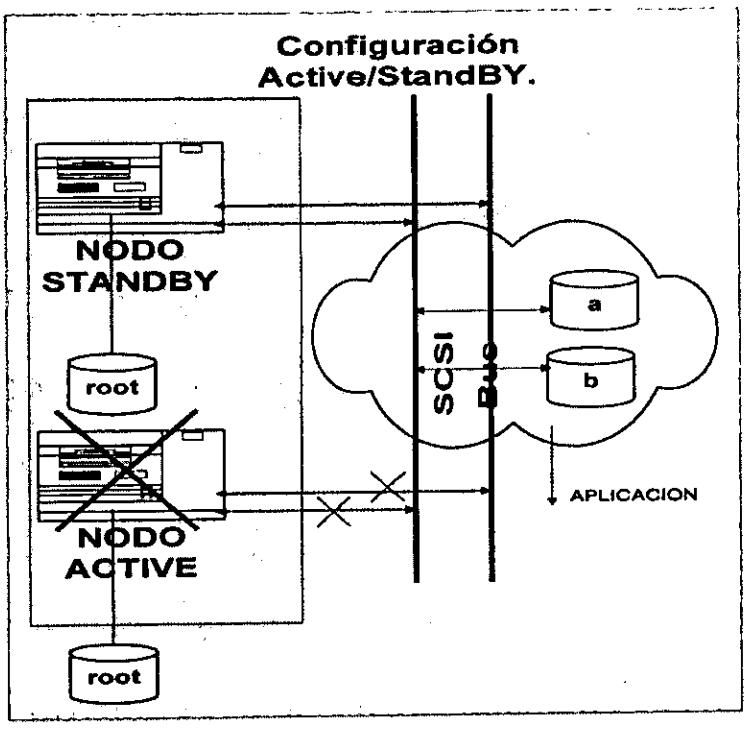


Figura 10.5 Configuración Active/Standby. El nodo activo falla, el nodo En Standby toma la aplicación y comienza a trabajar.

b) Active/Active

En este esquema los dos nodos del Cluster están en Activo. Es decir, los dos están corriendo una aplicación de manera independiente. En el momento que uno de ellos falla, el otro nodo en el Cluster toma la responsabilidad de seguir ejecutando la aplicación que estaba corriendo en el nodo de falla. En este caso el nodo que queda activo ejecuta su aplicación y la del nodo con falla al mismo tiempo, trabajando a un 50% de su rendimiento con respecto a su estado inicial. El esquema Active/Active se muestra en la figura 10.6.

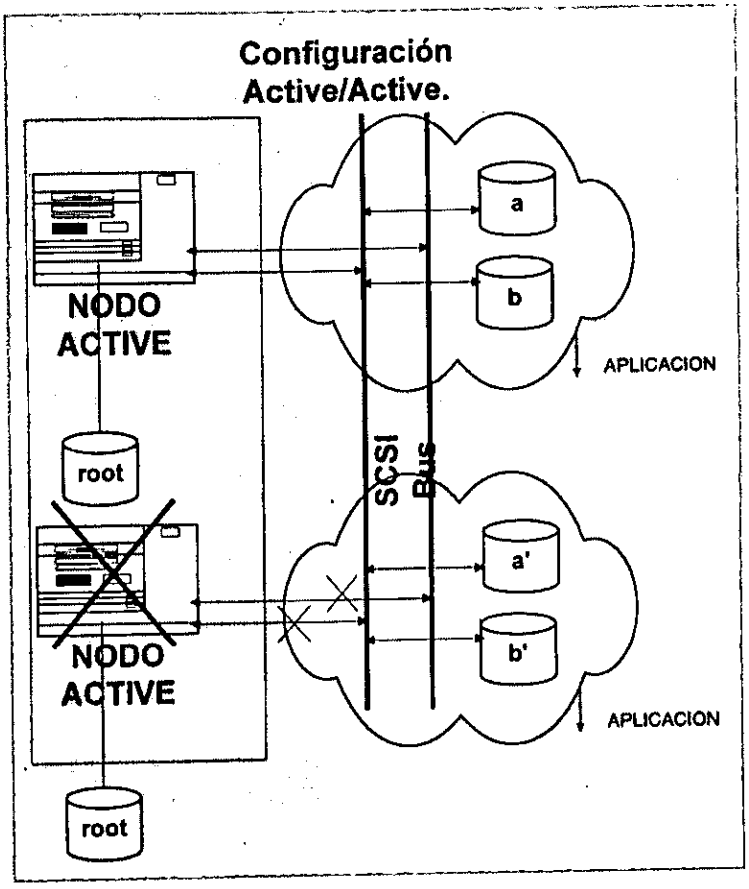


Figura 10.6 Configuración Active/Active.
Uno de los Nodos Falla y el otro toma su aplicación.

CAPITULO 11
IMPLEMENTACION DE UN SISTEMA DE ALTA DISPONIBILIDAD EN PLATAFORMA HP9000 T500.

11.1 PRESENTACION DEL PROYECTO

El presente Capítulo describe el diseño e implantación del Cluster de alta disponibilidad usando MC/ServiceGuard (Multi Computer/Service Guard) y como manejador de Bases de Datos Oracle.

11.1.1 Objetivo.

Implantar dos Cluster de alta disponibilidad con dos nodos y seis aplicaciones de misión crítica corriendo en cada nodo con las siguientes características:

a) La primera aplicación de misión crítica es la instancia de Oracle.
b) La segunda aplicación es TPS (Sistema de Teleproceso).
c) La tercera aplicación es SQL*NET.
d) La cuarta aplicación es la llamada DAR.
e) La quinta aplicación es ALARM.
f) La sexta aplicación es CONNECT DIRECT.

Todas estas aplicaciones se manejan como servicios que pertenecen a un paquete de MC/ServiceGuard. En cada nodo del Cluster existe un paquete que contiene estos seis servicios. Cada servicio tiene una cantidad específica de reintentos que indica cuántas veces el servicio correspondiente será levantado en caso de fallar.

11.1.2 Puesta en marcha de High Availability (Alta Disponibilidad)

Alta Disponibilidad "High Availability" es lo que puede garantizar la continuidad de las operaciones de los equipos de Transferencia y Procesamiento en caso de un simple problema de hardware. MC/ServiceGuard es un producto de software que habilita dos paquetes de aplicaciones dentro de un Cluster en uno solo de los nodos del Cluster en caso de que se tenga una falla de hardware, software o generada por algún agente externo al sistema.

11.2 ANALISIS

Antes de comenzar a plantear los CLUSTERS de alta disponibilidad en los equipos de producción MT1, MT2, MP1 y MP2, fue necesario hacer uso de los equipos de prueba incluidos en el proyecto TPS (Sistema de Teleproceso). Dichos equipos de prueba simulan ser un equipo de Transferencia y uno de Procesamiento. Como antecedente definimos el significado de MT y MP. Los equipos de prueba TESTMT y TESTMP están ilustrados en la figura 11.2.1.

MT "Machine Transfer" Equipos encargados de la transferencia y validación de archivos desde un Switch remoto, para enviarlos a proceso a los equipos MP.

MP "Machine Process" Equipos encargados de procesar y tarifar los registros recolectados y validados por los equipos MT.

El Sistema TPS (Teleprocess System) está ilustrado en la figura 11.2.1, en todo su conjunto.

En este Capítulo sólo presentaremos los cambios realizados para el Cluster MT (Cluster de Equipos de Transferencia), con la finalidad de evitar redundancia al duplicar la información incluyendo el Cluster MP ya que son semejantes.

11.2.1 Antecedentes.

Los Clusters, compuestos por los nodos MT1, MT2, MP1 y MP2 figura 11.2.2, fueron implantados siguiendo el diseño de la maqueta TCST que actualmente no se encuentra en producción sólo como equipo de prueba y fue previamente puesto en marcha por el proveedor y el equipo de Soporte y Administración del Sistema TPS. Del Cluster de prueba se tomaron los siguientes elementos y se personalizaron adecuadamente para configurar el Cluster que se describe en este Capítulo:

a) El archivo de configuración del Cluster.
b) El archivo de configuración de cada uno de los dos paquetes.
c) El script de control de cada uno de los dos paquetes.
d) Un archivo común a los dos paquetes que contiene la definición de variables particulares a cada paquete.
e) Los scripts que monitorean a cada uno de los seis servicios definidos en cada paquete.

En lo que respecta a Configuración de Volume Groups, Logical Volumes, Arreglos de Discos (RAID 5), Paths Primarios y PV Links de discos, Definición de ambientes (UIDs, GIDs, home directories, profiles y permisos), Valores de parámetros del kernel y Areas de swap, se hizo exactamente igual que en los nodos del Cluster puesto previamente en prueba ya que anteriormente habían sido definidos como los necesarios además de que los equipos de prueba son idénticos a los de producción.

A continuación vamos a indicar los cambios que se hicieron en la implantación del Cluster formado por los nodos MT1 y MT2, así como dar algunas recomendaciones necesarias en lo referente al hardware del Cluster.

11.2.2 Consideraciones para el Diseño del Cluster.

a) El Cluster está compuesto por servidores HP9000 T500 con 12 procesadores; 2 GB de memoria RAM; Sistema Operativo HP-UX 10.10 y MC/ServiceGuard 10.05, cada uno.

b) Existen dos paquetes (**MT1** y **MT2**). Uno de ellos tiene como nodo original a **mt1s** y como nodo adoptivo a **mt2s**. El otro paquete tiene como nodo original a **mp2s** y como nodo adoptivo a **mp1s**.

11.2.3 Información de Discos Locales.

Nodo mt1s:

Cuenta con tres discos internos de 4 GB cada uno en los cuales se encuentra el sistema operativo HP-UX 10.10, el software de Oracle y las áreas de swap.

Las direcciones de hardware para los discos internos son:

c2t6d0(10/0.6.0); c2t5d0 (10/0.5.0) y c2t4d0 (10/0.4.0)

Los volúmenes lógicos definidos, en base a los discos locales, son presentados en la tabla 11.2.1:

Volume Group	Logical Volume	Tamaño (MB)	Uso
vg00	lvol1	500	/
vg00	lvol2	1500	Area de swap
vg00	lvol3	1500	/usr
vg00	lvol4	500	/var
vg00	lvol5	2596	/tmp
vg00	lvol6	1496	Area de swap
vg01	lvol1	500	/home
vg01	lvol2	3000	/opt

Tabla 11.2.1 Definición de Volúmenes Lógicos en los discos locales del nodo mt1s.

Nodo mt2s:

Cuenta con tres discos internos de 4 GB cada uno en los cuales se encuentra el sistema operativo HP-UX 10.10, el software de Oracle y las áreas de swap.

Las direcciones de hardware para los discos internos son:

c2t6d0(10/0.6.0); c2t5d0 (10/0.5.0) y c2t4d0 (10/0.4.0)

Los volúmenes lógicos definidos tomando como referencia a los discos locales son presentados en la tabla 11.2.2:

Volume Group	Logical Volume	Tamaño (MB)	Uso
Vg00	Lvol1	500	/
Vg00	Lvol2	1500	Area de swap
Vg00	Lvol3	1500	/usr
Vg00	Lvol4	500	/var
Vg00	Lvol5	2596	/tmp
Vg00	Lvol6	1496	Area de swap
Vg01	Lvol1	500	/home
Vg01	Lvol2	3000	/opt

Tabla 11.2.2 Definición de Volúmenes Lógicos en los discos locales del nodo mt2s.

La tabla 11.2.3 integra los Volume Groups definidos para los discos internos, su Dirección de hardware, su archivo de dispositivo y el uso que se les está dando a los discos internos.

Volume Group	T500 MT1s Hardware Address	MT1s Device File	T500 MT2s Hardware Address	MT2s Device File	Uso
Vg00	10/0.5.0	c2t5d0	10/0.5.0	c2t5d0	Boot: Sistema Operativo, Software de Oracle y Swap
Vg00	10/0.6.0	c2t6d0	10/0.6.0	c2t6d0	
Vg01	10/0.4.0	c2t4d0	10/0.4.0	c2t4d0	

Tabla 11.2.3 Device Files de Discos Locales en el Cluster MT.

Cabe mencionar que tanto en el nodo **mt1s** como en el nodo **mt2s**, los discos de los Volume Groups vg00 y vg01 son internos y ninguno de sus Volúmenes Lógicos tiene espejo.

11.2.4 Información de Discos Compartidos.

Los discos compartidos del Cluster están formados por dos arreglos de discos NIKE MODEL 20/4. Cada arreglo cuenta con 15 discos de 4.2 GB cada uno. Así mismo, cada arreglo fue configurado para contener a tres LUN (Logical Unit Number) 0, 1 y 2. Cada LUN contiene 5 discos configurados en RAID 5.

Los dos primeros LUN de un arreglo NIKE forman al Volume Group vg1mt1, y el tercer LUN de ese mismo arreglo forma al Volume Group vg2mt1. Estos dos Volume Groups pertenecen al paquete que originalmente corre en el nodo **mt1s**. Cada uno de los dos Volume Groups anteriores contiene un file system.

Los dos primeros LUN del segundo arreglo NIKE forman al Volume Group vg1mt2, y el tercer LUN de ese mismo arreglo forma al Volume Group vg2mt2. Estos dos Volume Groups pertenecen al paquete que originalmente corre en el nodo **mt2s**. Cada uno de los dos Volume Groups anteriores contiene un file system.

En una de las cadenas SCSI de los discos compartidos se encuentra conectado también un disco que funciona como LOCK DISK del Cluster. Actualmente este disco no es usado por ninguna de las aplicaciones. Lo anterior está ilustrado en las tablas 11.2.4 y 11.2.5.

Ruta Primaria		Ruta PV Alterna Link		Uso
Hardware Address	**Device File**	**Hardware Address**	**Device File**	
8/12.4.0	c1t4d0	8/8.5.0	c0t5d0	Vg1mt1
8/12.4.1	c1t4d1	8/8.5.1	c0t5d1	Vg2mt1
8/12.4.2	c1t4d2	8/8.5.2	c0t5d2	Vg1mt1
8/12.2.0	c1t2d0	8/8.3.0	c0t3d0	Vg1mt2
8/12.2.1	c1t2d1	8/8.3.1	c0t3d1	Vg2mt2
8/12.2.2	c1t2d2	8/8.3.2	c0t3d2	Vg1mt2
10/4/4.15.0	c3t15d0			Vglock

Tabla 11.2.4 Volume Groups de Discos Compartidos en el Cluster vistos desde mt1s y mt2s.

Volume Group	Logical Volume	Tamaño (MB)	Ubicación	Uso
vg1mt1	data1	28000	c1t4d0, c1t4d2, c0t5d0, c0t5d2	/opt/MT1_mnt1
vg2mt1	data2	16000	c0t5d1, c1t4d1	/opt/MT1_mnt2
vg1mt2	data1	28000	c1t2d0, c1t2d2, c0t3d0, c0t3d2	/opt/MT2_mnt1
vg2mt2	data2	16000	c0t3d1, c1t2d1	/opt/MT2_mnt2

Tabla 11.2.5 Volume Groups de Discos Compartidos.

Cabe mencionar que esta distribución de espacios en disco y tamaños de los Volúmenes Lógicos, mostrados en las tablas 11.2.4 y 11.2.5 ya estaban hechos en el momento de la implantación del Cluster descrito en este Capítulo.

11.2.5 Información de Tarjetas de LAN.

Cada servidor del Cluster tiene dos tarjetas de LAN 802.3 y dos tarjetas FDDI SAS. Una de las tarjetas de red 802.3 (lan1) de cada servidor se usa para transmitir la señal de heartbeat del Cluster, y la otra tarjeta 802.3 (lan0) se usa como respaldo de la primera.

De manera similar, una de las tarjetas FFDI (lan3) se usa también para transmitir la señal de heartbeat, y la segunda tarjeta FDDI (lan2) se usa como respaldo de la primera.

En la tabla 11.2.6 se muestran las direcciones IP usadas en el Cluster:

Nodo	Interface	Dirección IP	Uso
mt1s	802.3-lan0	Desactivada	Backup de lan1
mt1s	802.3-lan1	13.49.152.101	Heartbeat
mt1s	FDDI SAS-lan3	Desactivada	Backup de lan4
mt1s	FDDI SAS-lan4	13.50.70.101	Heartbeat
mt2s	802.3-lan0	Desactivada	Backup de lan1
mt2s	802.3-lan1	13.49.152.110	Heartbeat
mt2s	FDDI SAS-lan3	Desactivada	Backup de lan4
mt2s	FDDI SAS-lan4	13.50.70.110	Heartbeat

Tabla 11.2.6 Información de tarjetas de LAN del Cluster.

Existen dos paquetes en el Cluster (**MT1** y **MT2**), cada uno de ellos tiene asignada una dirección IP en la subred 13.49.152 (802.3) y otra en la subred 13.50.70 (FDDI), tal como se indica en la siguiente tabla 11.2.7:

Paquete	Dirección IP (802.3) y hostname	Dirección IP (FDDI) y hostname
MT1	13.49.152.122-mxmt1e	13.50.70.122-mxmt1
MT2	13.49.152.129-mxmt2e	13.50.70.129-mxmt2

Tabla 11.2.7 Direcciones IP de los paquetes del Cluster.

11.2.6 Configuración de la Red .

El sistema consiste de 2 Equipos de Transferencia, 2 Equipos de Procesamiento, 1 Equipo de Almacenamiento y 2 Equipos de prueba, uno para Transferencia y otro para Procesamiento. Las X-Terminal también constituyen una parte fundamental para la configuración de la Red.

Los Equipos de Procesamiento, Equipos de Transferencia y de Almacenamiento están conectados al anillo FDDI primario y al anillo FDDI standby (de respaldo) por dos ligas FDDI. Los dos anillos FDDI están conectados por un Puente FDDI (FDDI Bridge). Esto garantiza que los sistemas puedan comunicarse si existe una falla en el Anillo Primario de FDDI. De modo similar, los Equipos de Transferencia y de Procesamiento están también conectados a la Ethernet LAN primaria y a la standby Ethernet LAN, y éstas a su vez están conectadas a través de un puente LAN. Los dos equipos de Test están conectados al FDDI Primario y a la Ethernet Primaria solamente. Las X-Terminal están conectadas a la Ethernet LAN primaria.

Los datos y los archivos transferidos entre los sistemas hacen uso del FDDI primario. La LAN primaria y FDDI LAN son usadas por la señal de Heartbeat intercambiada entre los equipos que estén dentro de la funcionalidad de Haig Availability y por las conexiones de las X-Terminal. La configuración de la red está mostrada en la figura 11.2.1:

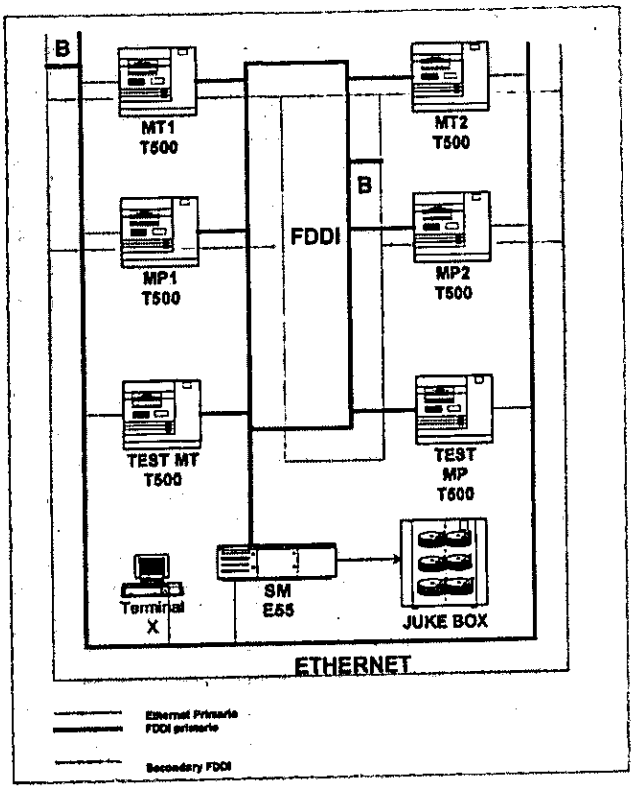
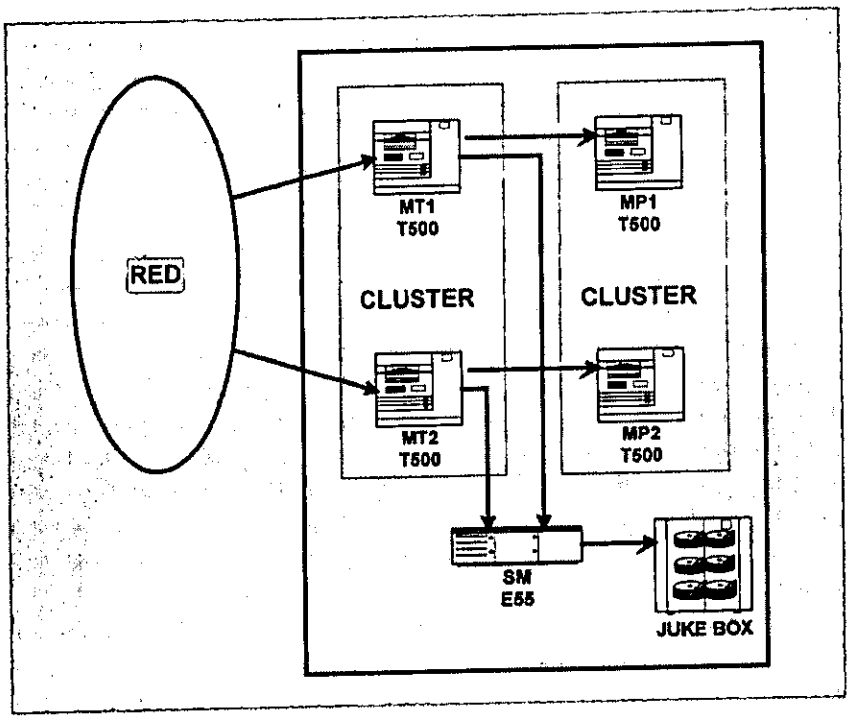


Fig. 11.2.1 TPS de México

La figura 11.2.2 muestra los elementos que componen el TPS (Sistema de Teleproceso), en su conjunto:



Fig, 11.2.2 TPS Operación Normal

La figura 11.2.3 es un ejemplo de la configuración del sistema dentro de operación normal. El conjunto de discos (a,b) y (a', b') son accesados por ambos sistemas. Pero cada uno de los conjuntos es reservado y usado sólo por uno de los sistemas.

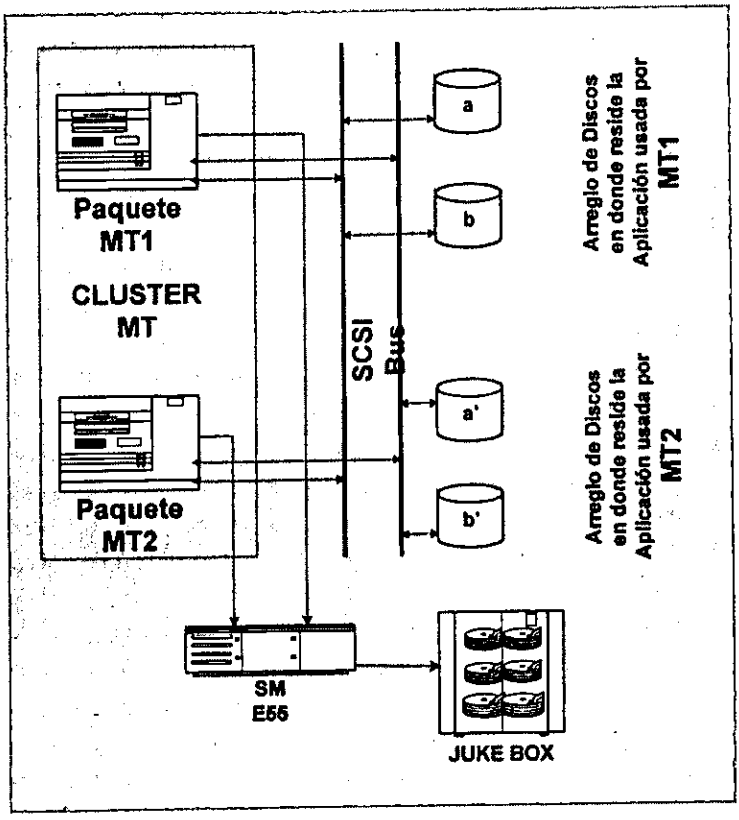


Fig. 11.2.3 Cluster en Operación Normal.

11.2.7 MC/ServiceGuard Software

El MC/ServiceGuard Software va a ser usado en el Sistema para proveer a éste de un Mecanismo de Alta Disponibilidad (High Availability) para los Equipos de Transferencia y Procesamiento, así como sus Aplicaciones de respaldo. Los equipos de Transferencia y Procesamiento deberán hacer uso de MC/ServiceGuard como protección contra fallas en los casos más críticos. El equipo de Almacenamiento no entrará en High Availability en caso de una falla ya que puede estar fuera de funcionamiento en un periodo máximo de 24 hrs. Los nodos en un Cluster estarán dando señales de su correcto funcionamiento mediante la señal de Heartbeat. La señal de Heartbeat será mandada sobre la Ethernet LAN y FDDI LAN a todos los nodos en el Cluster. Si en un determinado número de Heartbeats un nodo no detecta otro nodo, éste asume que tiene una falla. El Paquete de Aplicaciones, que está corriendo en el nodo que es detectado con fallas es transferido al nodo que lo detecta, convirtiéndose así en un Nodo Adoptivo.

En este caso, si el equipo de MT1 no está siendo visto por MT2 entonces MT2 puede tomar el Paquete de aplicaciones de MT1 y puede ser ejecutado por MT2.

El Software MC/ServiceGuard, monitorea ciertos aspectos del Sistema en un Cluster. En el caso de tener un problema, el sistema no queda fuera de servicio. El MC/ServiceGuard Software soporta dos tipos de Fallas:

a) Fallas Locales

El MC/ServiceGuard Software detecta fallas locales en uno de los equipos del Cluster.

b) Fallas Remotas

Las aplicaciones son ejecutadas en un Nodo adoptivo cuando el Nodo Primario falla.
NOTA: Existen ciertas aplicaciones que no monitorea MC/ServiceGuard (Aplicaciones no críticas). El monitoreo de estas fallas en las aplicaciones quedará en responsabilidad de Operación y esta situación va a ser descrita ampliamente dentro de la Sección de Responsabilidades del Administrador.

11.2.8 Migración en Caso de un Desastre en el Centro de Cómputo.

Esta situación está contemplada para la Fase 2 de este proyecto, próxima en el año 2000, incluyendo un espejo del proyecto en Monterrey para solventar situaciones de desastre en cualquiera de los dos Centros de Cómputo. Siempre que exista una falla (Fileover) que no pueda ser detectada por MC/ServiceGuard se le llama SiteFailover. Un SiteFailover puede suscitarse en caso de un catastrófico desastre natural, por citar un ejemplo. En este caso ambos equipos en el Cluster, incluyendo el MC/ServiceGuard Software fallarán por completo.

11.3 DISEÑO E IMPLEMENTACION

11.3.1 El Cluster definido para el Sistema y la Configuración del Paquete de Aplicaciones (Pakage).

La siguiente tabla 11.3.1 muestra los Clusters del Sistema y sus hostnames asignados, para la puesta en marcha del Cluster mostrado en la figura 11.2.2:

Nombre del Cluster	ID del Cluster	Nombre de los Nodos y (Hosts Estacionarios)
Machine Transfer (MT)	MT	mxmt1, mxmt2
Machine Process (MP)	MP	mxmp1, mxmp2

Tabla 11.3.1 TPS Clusters

11.3.2 Configuración del Cluster.

A continuación, se indicarán cuáles son los archivos de configuración y scripts del Cluster, así como las partes que fueron modificadas y adicionadas para éste Cluster en particular. Estos archivos y scripts fueron tomados del Cluster que actualmente se tiene en equipos de prueba.

1.- Existe un solo archivo de configuración del Cluster:

/etc/cmCluster/Cluster.conf.MT

En este archivo se especifican las características generales del Cluster, tales como: nombre del Cluster; nombres de los nodos; tarjetas de red de los nodos y direcciones IP asignadas a éstas. A continuación se indican las partes de este archivo que se modificaronon ya que estos archivos son creados en la instalación de MC Service Guard y los parámetros manejados deben ser definidos:

```
CLUSTER_NAME MT_12

FIRST_CLUSTER_LOCK_VG /dev/vglock

# SG node 1 definition
NODE_NAME               mt1s
  NETWORK_INTERFACE lan1
    HEARTBEAT_IP        13.49.152.102
```

```
 NETWORK_INTERFACE lan2
  HEARTBEAT_IP          13.50.70.102
 NETWORK_INTERFACE lan3
 NETWORK_INTERFACE lan0
 FIRST_CLUSTER_LOCK_PV /dev/dsk/c3t15d0

# SG node 2 definition
NODE_NAME              MT2s
 NETWORK_INTERFACE lan1
  HEARTBEAT_IP          13.49.152.109
 NETWORK_INTERFACE lan2
  HEARTBEAT_IP          13.50.70.109
 NETWORK_INTERFACE lan3
 NETWORK_INTERFACE lan0
 FIRST_CLUSTER_LOCK_PV /dev/dsk/c3t15d0


# Cluster Timing Parmeters (microseconds).
# Don't modify the following setting, it is the TPS HA design decision.
#

HEARTBEAT_INTERVAL 1000000
NODE_TIMEOUT       5000000


# Configuration/Reconfiguration Timing Parameters (microseconds).
# Don't modify the following setting, it is the TPS HA design decision.
#

AUTO_START_TIMEOUT 600000000
NETWORK_POLLING_INTERVAL     2000000

# List of Cluster aware Volume Groups. These volume groups will
# be used by Clustered applications via the vgchange -a e command.
#
# For example:
# VOLUME_GROUP               /dev/vgabin
# VOLUME_GROUP               /dev/vgdatabase.
# VOLUME_GROUP               /dev/vgdata.

# Add or remove the following VOLUME_GROUP entries to match the real
# number of Volume Groups aware to Cluster.
#
# VOLUME_GROUP               /dev/vg1mt1
# VOLUME_GROUP               /dev/vg2mt1
# VOLUME_GROUP               /dev/vg1mt2
# VOLUME_GROUP               /dev/vg2mt2
```

2.- Existen a continuación bajo el directorio **/etc/cmCluster** dos subdirectorios (**MT1** y **MT2**) en donde se encuentran tanto el archivo de configuración de cada uno de los paquetes, como el script de control también para cada uno de los paquetes:

/etc/cmCluster/MT1/MT1.conf
/etc/cmCluster/MT1/control.sh

/etc/cmCluster/MT2/MT2.conf
/etc/cmCluster/MT2/control.sh

A continuación se indican los parámetros del archivo **/etc/cmCluster/MT1/MT1.conf** que se personalizaron, con el fin de adecuarlo al paquete MT1 del Cluster necesario en TPS. En este archivo se determina la ruta de los archivos de control de arranque del Cluster, las aplicaciones definidas en el Cluster y el tiempo que esperará para tratar de reavivar la aplicación y si está habilitado el switcheo de la red y los paquetes que comprenden el Cluster.

```
PACKAGE_NAME  MT1

NODE_NAME       mt1s
NODE_NAME       mt2s

RUN_SCRIPT              /etc/cmCluster/MT1/control.sh
RUN_SCRIPT_TIMEOUT          NO_TIMEOUT
HALT_SCRIPT             /etc/cmCluster/MT1/control.sh
HALT_SCRIPT_TIMEOUT         1000000000

SERVICE_NAME            ORACLE_MT1
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT   50

SERVICE_NAME            TPS_MT1
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT   50

SERVICE_NAME            SQLNET_MT1
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT   50

SERVICE_NAME            DAR_MT1
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT   50
```

```
SERVICE_NAME                ALARM_MT1
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT   50

SERVICE_NAME                CD_MT1
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT   50

# SERVICE_NAME              OMNI_MT1
# SERVICE_FAIL_FAST_ENABLED    NO
# SERVICE_HALT_TIMEOUT   50

# SERVICE_NAME              OMNISTORE_MT1
# SERVICE_FAIL_FAST_ENABLED    NO
# SERVICE_HALT_TIMEOUT   50

SUBNET              13.49.152.0
SUBNET              13.50.70.0

PKG_SWITCHING_ENABLED           YES

NET_SWITCHING_ENABLED           YES

NODE_FAIL_FAST_ENABLED   NO
```

A continuación se indica la parte del archivo **/etc/cmCluster/MT1/control.sh** que se modificó el cual tendrá el objetivo de arrancar y detener las aplicaciones involucradas en el paquete MT1 del Cluster de manera automática:

```
#*********************************************************************
***
# *********     BEGIN of TPS Package Control script main function
*********
#*********************************************************************
***

#main()
#{
integer exit_value=0
integer TPS_flag=0

# Test to see if we are being called to run the package, or halt the package.
```

```
if [[ $1 = "start" ]]
then
   print
"\n\n##################################################"       >>
$LOG
     print "## SG node \"$(hostname)\": Starting package at $(date)" >> $LOG
   print   "##################################################\n"
>> $LOG

   # Initialize package control script variables
   Initialization WITH_CHECK

   # Module Volume Group
     Activate_Volume_Group

   # Module File System
     Check_And_Mount

   # Set TPS environment variables after file system mounted
   #
   if [ -x $TPS_ENV_PROFILE ]
   then
          . $TPS_ENV_PROFILE
   else
          let 0
          Test_Return 32
   fi

   # Send message to console monitor

   if (( ! $(Is_Original) ))
   then
          print "\tALARM: REMOTE FAILOVER: Package ${TPS_APPID}
is switched over to MC/SG node \"$(hostname)\" at $(date)." >>
$TPS_ALARM_CONSOLE
   fi
     print "\tINFORMATIVE: MC/SG node \"$(hostname)\": Starting package
${TPS_APPID} at $(date)" >> $TPS_ALARM_CONSOLE

   # Module Package IP Address
     Add_IP_Address

   # Module DB Start Up script
     Start_Up_DB

   # Module SQL*NET Start Up
   Start_Up_Sqlnet
```

```
    # Module C:D Start Up
    Start_Up_CD

    # Module OmniBack Start Up
    # Start_Up_OmniBack

    # Module TPS Start Up
    Start_Up_TPS

    # For MT, Startup DAR and Alarm
    if [ "$MT_MP" = "MT" ]
    then

            # Module DAR Start Up
            Start_Up_DAR

            # Module OmniStore Start Up
            # Start_Up_OmniStore
    fi

    # Module Alarm Start Up
    Start_Up_Alarm

    # Module Monitor Services Start Up
    Start_Up_Monitor_Services

    # Check/Restart missed cronjobs
    Restart_Cronjobs

elif [[ $1 = "stop" ]]
then
    print  "\n\n#####################################################"
>> $LOG
      print "## SG node \"$(hostname)\": Halting package at $(date)" >> $LOG
    print    "#####################################################\n"
>> $LOG

    # Set TPS environment variables
    #
    . $TPS_ENV_PROFILE

    # Initialize package control script variables
    Initialization

    # Send message to console monitor
```

```
    print "\tINFORMATIVE: MC/SG node \"$(hostname)\": Stopping
package ${TPS_APPID} at $(date)" >> $TPS_ALARM_CONSOLE

  # Moudle Stop Monitor Services
  Stop_Monitor_Services
  # Module Stop Alarm
  Stop_Alarm
  # For MT, Stop DAR and Alarm
  if [ "$MT_MP" = "MT" ]
  then
        # Module Stop DAR
        Stop_DAR

        # Module Stop OmniStore
        # Stop_OmniStore
  fi

  # Module Stop TPS
  Stop_TPS

  # Module Stop OmniBack
  # Stop_OmniBack

  # Module Stop C:D
  Stop_CD

  # Module Stop SQL*NET
  Stop_Sqlnet

    # Module DB Stop
  Stop_DB

  # Module Package IP Address
    Remove_IP_Address

  # Module File System
    Umount_FS

  # Module Volume Group
    Deactivate_Volume_Group
fi
# Check exit value
if ((exit_value == 1))
then
     exit 2  # no restart of TPS package on remote node
else
     exit 0  # the error is minor, ignore it.
```

```
fi
# ********* END of TPS Package Control script main function
```

A continuación se indican los parámetros del archivo **/etc/cmCluster/MT2/MT2.conf** que se personalizaron:

```
PACKAGE_NAME  MT2

NODE_NAME       mt2s
NODE_NAME       mt1s

RUN_SCRIPT              /etc/cmCluster/MT2/control.sh
RUN_SCRIPT_TIMEOUT      NO_TIMEOUT
HALT_SCRIPT             /etc/cmCluster/MT2/control.sh
HALT_SCRIPT_TIMEOUT     1000000000

SERVICE_NAME            ORACLE_MT2
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT  50

SERVICE_NAME            TPS_MT2
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT  50

SERVICE_NAME            SQLNET_MT2
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT  50

SERVICE_NAME            DAR_MT2
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT  50

SERVICE_NAME            ALARM_MT2
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT  50

SERVICE_NAME            CD_MT2
SERVICE_FAIL_FAST_ENABLED    NO
SERVICE_HALT_TIMEOUT  50

# SERVICE_NAME          OMNI_MT2
# SERVICE_FAIL_FAST_ENABLED    NO
# SERVICE_HALT_TIMEOUT  50

# SERVICE_NAME          OMNISTORE_MT2
```

```
# SERVICE_FAIL_FAST_ENABLED     NO
# SERVICE_HALT_TIMEOUT   50

SUBNET              13.49.152.0
SUBNET              13.50.70.0

PKG_SWITCHING_ENABLED           YES

NET_SWITCHING_ENABLED           YES

NODE_FAIL_FAST_ENABLED   NO
```

A continuación se indica la parte del archivo **/etc/cmCluster/MT2/control.sh** que se personalizó, ya que se pidió en última instancia deshabilitar OMNIBACK y OMNISTORE:

```
#*********************************************************************
***
# *********     BEGIN of TPS Package Control script main function
*********
#*********************************************************************
***
#main()
#{
integer exit_value=0
integer TPS_flag=0
# Test to see if we are being called to run the package, or halt the package.
if [[ $1 = "start" ]]
then
   print
"\n\n#####################################################"      >>
$LOG
     print "## SG node \"$(hostname)\": Starting package at $(date)" >> $LOG
   print  "#####################################################\n"
>> $LOG

   # Initialize package control script variables
   Initialization WITH_CHECK

   # Module Volume Group
     Activate_Volume_Group

   # Module File System
     Check_And_Mount
```

```
# Set TPS environment variables after file system mounted
#
if [ -x $TPS_ENV_PROFILE ]
then
        . $TPS_ENV_PROFILE
else
        let 0
        Test_Return 32
fi

# Send message to console monitor

if (( ! $(Is_Original) ))
then
        print "\tALARM: REMOTE FAILOVER: Package ${TPS_APPID}
is switched over to MC/SG node \"$(hostname)\" at $(date)." >>
$TPS_ALARM_CONSOLE
fi
  print "\tINFORMATIVE: MC/SG node \"$(hostname)\": Starting package
${TPS_APPID} at $(date)" >> $TPS_ALARM_CONSOLE

# Module Package IP Address
  Add_IP_Address

# Module DB Start Up script
  Start_Up_DB

# Module SQL*NET Start Up
Start_Up_Sqlnet

# Module C:D Start Up
Start_Up_CD

# Module OmniBack Start Up
# Start_Up_OmniBack

# Module TPS Start Up
Start_Up_TPS

# For MT, Startup DAR and Alarm
if [ "$MT_MP" = "MT" ]
then

        # Module DAR Start Up
        Start_Up_DAR
```

```
        # Module OmniStore Start Up
        # Start_Up_OmniStore
    fi

    # Module Alarm Start Up
    Start_Up_Alarm

    # Module Monitor Services Start Up
    Start_Up_Monitor_Services

    # Check/Restart missed cronjobs
    Restart_Cronjobs

elif [[ $1 = "stop" ]]
then
    print  "\n\n##################################################"
>> $LOG
      print "## SG node \"$(hostname)\": Halting package at $(date)" >> $LOG
    print    "##################################################\n"
>> $LOG

    # Set TPS environment variables
    #
    . $TPS_ENV_PROFILE

    # Initialize package control script variables
    Initialization

    # Send message to console monitor
      print  "\tINFORMATIVE:  MC/SG  node  \"$(hostname)\":  Stopping
package ${TPS_APPID} at $(date)" >> $TPS_ALARM_CONSOLE

    # Moudle Stop Monitor Services
    Stop_Monitor_Services

    # Module Stop Alarm
    Stop_Alarm

    # For MT, Stop DAR and Alarm
    if [ "$MT_MP" = "MT" ]
    then
        # Module Stop DAR
        Stop_DAR

        # Module Stop OmniStore
        # Stop_OmniStore
    fi
```

```
  # Module Stop TPS
  Stop_TPS

  # Module Stop OmniBack
  # Stop_OmniBack

  # Module Stop C:D
  Stop_CD

  # Module Stop SQL*NET
  Stop_Sqlnet

    # Module DB Stop
  Stop_DB

  # Module Package IP Address
    Remove_IP_Address

  # Module File System
    Umount_FS

  # Module Volume Group
    Deactivate_Volume_Group
fi

# Check exit value
if ((exit_value == 1))
then
      exit 2  # no restart of TPS package on remote node
else
      exit 0  # the error is minor, ignore it.
fi

#*********************************************************************
***
# *********      END of TPS Package Control script main function
*********
#*********************************************************************
***
```

3.- El archivo de configuración de cada uno de los paquetes invoca al script **/etc/cmCluster/cmrm_control.conf**. En este script se definen tanto los servicios como los Volume Groups, File System y Direcciones IP usados por cada uno de los dos paquetes.

A continuación se indican las partes del script **/etc/cmCluster/cmrm_control.conf** que se modifican para este Cluster:

```
SERVICE_LOG=/var/adm/cmCluster/services_${TPS_APPID}.log
SERVICE_NAME[0]=ORACLE_${TPS_APPID}
SERVICE_MTD[0]="/etc/cmCluster/monitor/ora.mon ${TPS_APPID} "
SERVICE_RESTARTS[0]="-r 0"
SERVICE_INSTANCES[0]=2

SERVICE_NAME[1]=TPS_${TPS_APPID}
SERVICE_CMD[1]="/etc/cmCluster/monitor/TPS.mon ${TPS_APPID} "
SERVICE_RESTARTS[1]="-r 2"
SERVICE_INSTANCES[1]=2

SERVICE_NAME[2]=ALARM_${TPS_APPID}
SERVICE_CMD[2]="/etc/cmCluster/monitor/alarm.mon ${TPS_APPID} "
SERVICE_RESTARTS[2]="-r 8"
SERVICE_INSTANCES[2]=2

SERVICE_NAME[3]=SQLNET_${TPS_APPID}
SERVICE_CMD[3]="/etc/cmCluster/monitor/sqlnet.mon ${TPS_APPID} "
SERVICE_RESTARTS[3]="-r 4"
SERVICE_INSTANCES[3]=2

#################################################################
###
# Setting Monitor Services for Dar & Alarm only for MT
#################################################################
###
case $TPS_APPID in
*MT?)
#################################################################
###

SERVICE_NAME[4]=DAR_${TPS_APPID}
SERVICE_CMD[4]="/etc/cmCluster/monitor/dar.mon ${TPS_APPID} "
SERVICE_RESTARTS[4]="-r 4"
SERVICE_INSTANCES[4]=2

SERVICE_NAME[5]=CD_${TPS_APPID}
SERVICE_CMD[5]="/etc/cmCluster/monitor/cd.mon ${TPS_APPID} "
SERVICE_RESTARTS[5]="-r 4"
```

```
SERVICE_INSTANCES[5]=1

# SERVICE_NAME[6]=OMNISTORE_${TPS_APPID}
#                 SERVICE_CMD[6]="/etc/cmCluster/monitor/omnistore.mon
${TPS_APPID} "
# SERVICE_RESTARTS[6]="-r 4"
# SERVICE_INSTANCES[6]=1

# SERVICE_NAME[7]=OMNI_${TPS_APPID}
# SERVICE_CMD[7]="/etc/cmCluster/monitor/omni.mon ${TPS_APPID} "
# SERVICE_RESTARTS[7]="-r 4"
# SERVICE_INSTANCES[7]=1

#################################################################
##
;;
esac
```

Se adicionó también a este script la siguiente sección que indica los recursos de los paquetes **MT1** y **MT2,** y la dirección ip estacionaria que tomarán los paquetes cuando se conmuten:

```
###############################################
MT1)
################ Begin of  MT1 ################

IP[0]=13.49.152.122
SUBNET[0]=13.49.152.0
IP[1]=13.50.70.122
SUBNET[1]=13.50.70.0

VG[0]=/dev/vg1mt1
VG[1]=/dev/vg2mt1

LV[0]=/dev/vg1mt1/data1
FS[0]=/opt/MT1_mnt1

LV[1]=/dev/vg2mt1/data2
FS[1]=/opt/MT1_mnt2

# LV[2]=/dev/vg2mt1/omnistore
# FS[2]=/opt/tmx/MT1/omnistore

###############################################
;;
################ End of MT1 ################
```

```
##############################################
MT2)
############### Begin of  MT2 ###############

IP[0]=13.49.152.129
SUBNET[0]=13.49.152.0
IP[1]=13.50.70.129
SUBNET[1]=13.50.70.0

VG[0]=/dev/vg1mt2
VG[1]=/dev/vg2mt2

LV[0]=/dev/vg1mt2/data1
FS[0]=/opt/MT2_mnt1

LV[1]=/dev/vg2mt2/data2
FS[1]=/opt/MT2_mnt2

# LV[2]=/dev/vg2mt2/omnistore
# FS[2]=/opt/tmx/MT2/omnistore

##############################################
;;
############### End of MT2 ###############
```

4.- Finalmente, bajo el subdirectorio **/etc/cmCluster/monitor**, se encuentran los scripts que monitorean los servicios definidos en cada uno de los paquetes. Ninguno de esos scripts se modificó ya que anteriormente los programó y entregó el proveedor de la aplicación tps.

Como puede observarse, de manera concreta, no se están arrancando ni monitoreando los servicios para OmniBack ni para OmniStore.

11.3.3 Cambios y Actualizaciones en el Cluster.

Es recomendable que cuando se requiera realizar algún cambio en la configuración del Cluster se haga un planteamiento de manera precisa, planeando los días de los cambios, contar con los recursos de apoyo necesarios y respaldar el sistema operativo, aplicaciones y archivos de configuración del Cluster antes del cambio para poder restablecer el servicio en caso de problemas. Así entonces, antes de realizar cambios en la configuración del Cluster se deben detener primero los paquetes y a continuación detener completamente el Cluster.

Una vez puesto en práctica esto, se pueden realizar los cambios y proceder a configurar el Cluster para incluir los cambios, según se indica a continuación.

Si se necesita incluir, remover o modificar algún recurso de un paquete como por ejemplo:

a) Volume Groups.
b) Volúmenes Lógicos.
c) File System.
d) Direcciones IP.
e) Subredes a monitorear.
f) Servicios a monitorear.

Se debe editar tanto el archivo de configuración del Cluster (**/etc/cmCluster/Cluster.conf.MT**), como el script de control del paquete (**/etc/cmCluster/MT1/control.sh** o **/etc/cmCluster/MT2/control.sh**) y el script **/etc/cmCluster/cmrm_control.conf** para indicar los cambios necesarios.

A continuación, debe revisar que tanto el archivo de configuración del Cluster como los de configuración de los paquetes estén correctos, usando el siguiente comando:

```
cmcheckconf -v -C /etc/cmCluster/Cluster.conf.MT \
-P /etc/cmCluster/MT1/MT1.conf\
-P /etc/cmCluster/MT2/MT2.conf\
```

Nota: en caso de existir algún error en el archivo de configuración del Cluster, lo indicará la salida del comando anterior.

Si el Cluster se encuentra corriendo se debe detener con el comando:

```
cmhaltcl
```

Después, se debe generar un binario del archivo de configuración del Cluster (**/etc/cmCluster/cmclconfig**) y distribuirlo a todos los nodos del Cluster con el siguiente comando:

```
cmapplyconf -v -C /etc/cmCluster/Cluster.conf.MT \
-P /etc/cmCluster/MT1/MT1.conf\
-P /etc/cmCluster/MT2/MT2.conf\
```

Finalmente, se debe arrancar el Cluster para reflejar los cambios anteriores:

```
cmruncl
```

Nota importante: Cuando se haya creado un Volume Group nuevo para un paquete, dicho paquete no arrancará, por lo que se deberá ejecutar el comando **vgchange -a n VG*XX*** y se tendrá que arrancar nuevamente el paquete con el comando **cmrunpkg paquete**.

Demonios Usados por MC/ServiceGuard.

/usr/lbin/cmcld	Demonio del Cluster de MC/ServiceGuard existente en cada nodo del Cluster
/usr/bin/cmclconfd	Demonio que genera y distribuye el binario del archivo de configuración del Cluster en todos los nodos. Solo existe mientras se ejecuta el comando **cmapplyconf**.
/usr/bin/cmlvmd	Demonio usado por el Cluster para comunicarse con LVM (Logical Volume Manager) y existe en todos los nodos del Cluster.

11.4 PLAN DE PRUEBAS

El documento que sirve de apoyo para asegurar que un nuevo proyecto, al ser puesto en producción, cumple con los requerimientos y especificaciones necesarias para poder llevar a cabo su trabajo en operación normal se le denomina Plan de Pruebas. El plan de pruebas no asegura la ausencia de defectos en el Software o Hardware pero debe intentar demostrar que no existen defectos que puedan impedir su puesta en producción.

Las pruebas a realizar son únicamente para asegurar el correcto funcionamiento del Cluster simulando los diferentes estados de crisis que éste puede llegar a tener en Operación normal. Los estados de crisis de los paquetes que conforman un Cluster deberán ser simulados a continuación con las siguientes pruebas. En todas y cada una de las pruebas deberá cumplirse el objetivo principal DISPONIBILIDAD CONTINUA DEL PROYECTO TPS.

11.4.1 Pruebas de Validación Realizadas.

El siguiente conjunto de pruebas fue hecho partiendo de un estado del Cluster en el que cada uno de los dos paquetes está corriendo en su nodo original. Esto es, el paquete **MT1** corriendo en el nodo **mt1s** y el paquete **MT2** corriendo en el nodo **mt2s**.

1. Se movió manualmente el paquete **MT1** del nodo **mt1s** al nodo **mt2s** usando la siguiente secuencia de comandos:

   cmhaltpkg MT1
   cmrunpkg -n mt2s MT1
   cmmodpkg -e MT1

2. Estando los paquetes **MT1** y **MT2** corriendo en el nodo **mt2s**, se movieron al nodo **mt1s**, usando la siguiente secuencia de comandos:

   cmhaltpkg MT1
   cmhaltpkg MT2
   cmrunpkg -n mt1s MT1
   cmmodpkg -e MT1
   cmrunpkg -n mt1s MT2
   cmmodpkg -e MT2

3. Estando los paquetes **MT1** y **MT2** corriendo en el nodo **mt1s**, se movió manualmente el paquete **MT2** del nodo **mt1s** al nodo **mt2s** usando la siguiente secuencia de comandos:

   cmhaltpkg MT2
   cmrunpkg -n mt2s MT2
   cmmodpkg -e MT2

4. Estando el paquete **MT1** corriendo en el nodo **mt1s**, se ejecutó el comando **shutdown -r 0** en dicho nodo, moviéndose automáticamente el paquete al nodo **mt2s**. Cuando el nodo **mt1s** terminó de realizar su proceso de boot, se integró manualmente al Cluster con el comando **cmrunnode**.

5. Estando el paquete **MT2** corriendo en el nodo **mt2s**, se ejecutó el comando **shutdown -r 0** en dicho nodo, moviéndose automáticamente el paquete al nodo **mt2s**. Cuando el nodo **mt2s** terminó de realizar su proceso de boot, se integró manualmente al Cluster con el comando **cmrunnode**.

6. Se hizo la misma prueba del punto 4 pero ahora apagando el switch de alimentación eléctrica del nodo **mt1s**, observándose el mismo resultado del punto 4.

7. Se hizo la misma prueba del punto 5 pero ahora apagando el switch de alimentación eléctrica del nodo **mt2s**, observándose el mismo resultado del punto 5.

8. Estando nuevamente los dos paquetes en su nodo original, se detuvo manualmente el servicio (**ORACLE_MT1**) de Oracle (matando al proceso **ora_pmon**) en el nodo **mt1s**, moviéndose automáticamente el paquete **MT1** al nodo **mt2s**.

9. Estando nuevamente los dos paquetes en su nodo original, se detuvo manualmente el servicio (**TPS_MT1**) de TPS (matando en dos ocasiones al proceso **manmbin**) en el nodo **mt1s**, moviéndose automáticamente el paquete **MT1** al nodo **mt2s**.

10. Estando nuevamente los dos paquetes en su nodo original, se detuvo manualmente el servicio (**SQLNET_MT1**) de SQL*NET (matando en cuatro ocasiones al proceso **/opt/oracle/bin/tnslsnr**) en el nodo **mt1s**, moviéndose automáticamente el paquete **MT1** al nodo **mt2s**.

11. Estando nuevamente los dos paquetes en su nodo original, se detuvo manualmente el servicio (**DAR_MT1**) de DAR (matando en cuatro ocasiones al proceso **darput**) en el nodo **mt1s**, moviéndose automáticamente el paquete **MT1** al nodo **mt2s**.

12. Estando nuevamente los dos paquetes en su nodo original, se detuvo manualmente el servicio (**ALARM_MT1**) de alarmas (matando en ocho ocasiones al proceso **alarm.sh**) en el nodo **mt1s**, moviéndose automáticamente el paquete **MT1** al nodo **mt2s**.

13. Estando nuevamente los dos paquetes en su nodo original, se detuvo manualmente el servicio (**CD_MT1**) de Connect Direct (matando en cuatro ocasiones al proceso **/opt/connect/ndm/bin/cdpmgr**) en el nodo **mt1s**, moviéndose automáticamente el paquete **MT1** al nodo **mt2s**.

14. Se hicieron las mismas pruebas de los puntos 8 al 13 pero ahora estando el paquete **MT1** corriendo en el nodo **mt2s**, moviéndose normalmente el paquete al nodo **mt1s** después de detener cada uno de los servicios del paquete.

15. Se hicieron las mismas pruebas de los puntos 8 al 13 con el paquete **MT2** corriendo en el nodo **mt2s**, moviéndose normalmente el paquete al nodo **mt1s** después de detener cada uno de los servicios del paquete (**ORACLE_MT2, BMP_MT2, SQLNET_MT2, DAR_MT2, ALARM_MT2** y **CD_MT2**).

16. Se hicieron las mismas pruebas de los puntos 8 al 13 con el paquete **MT2** corriendo en el nodo **mt1s**, moviéndose normalmente el paquete al nodo **mt2s** después de detener cada uno de los servicios del paquete (**ORACLE_MT2, BMP_MT2, SQLNET_MT2, DAR_MT2, ALARM_MT2** y **CD_MT2**).

Debe aclararse que el arranque automático del Cluster, al realizar el proceso de boot cada uno de los nodos, está habilitado. Esta acción se especifica en el archivo **/etc/rc.config.d/cmCluster** con la bandera:

AUTOSTART_CMCLD=1

11.4.2 Pruebas de Validación de Red.

Para estas pruebas cada paquete se encontraba corriendo inicialmente en su nodo original, y se hizo la siguiente matriz de pruebas, tabla 11.4.1, donde "conect" significa que la tarjeta se encuentra conectada y "desconect" significa desconectada, simulando así que se ha dañado el cable o la tarjeta.
La última columna indica el comportamiento y observaciones del Cluster y el estado de la tarjeta vista con el comando **lanscan**.

Lan4 10/16/12 FDDI Primaria HB	lan3 10/4/12 FDDI Backup	lan1 10/12/6 Ether Primaria HB	lan0 10/4/8.1 Ether Backup	Observaciones
conect.	conect.	conect.	conect	Situación normal, lan0, lan1, lan3 y lan4 UP.
conect.	conect.	conect.	desconect.	No sucede nada en el Cluster.
conect.	conect.	desconect.	conect.	El heartbeat continúa transmitiéndose por lan0. Falla lan1, lan1 conmuta a lan0.
conect.	conect.	desconect.	desconect.	El paquete corriendo en el nodo se mueve al otro nodo, SUBNET 13.49.152.0 DOWN.
conect.	desconect.	conect.	conect.	No sucede nada en el Cluster. Falla lan3 y se registra en /var/adm/syslog/syslog.log.
conect.	desconect.	conect.	desconect.	No sucede nada en el Cluster. Fallan lan3 y lan0 y se registra en /var/adm/syslog/syslog.log.
conect.	desconect.	desconect.	conect.	Fallan lan3 y lan1, se registra en /var/adm/syslog/syslog.log. lan1 conmuta a lan0.
conect.	desconect.	desconect.	desconect.	Fallan lan0 y lan1, se registra en /var/adm/syslog/syslog.log. El paquete corriendo en el nodo se mueve al otro nodo, SUBNET 13.49.152.0 DOWN.
desconect.	conect.	conect.	conect.	Falla lan4. lan4 conmuta a lan3.
desconect.	conect.	conect.	desconect.	Fallan lan4 y lan0. lan4 conmuta a lan3.

lan4 10/16/12 FDDI Primaria HB	lan3 10/4/12 FDDI Backup	lan1 10/12/6 Ether Primaria HB	lan0 10/4/8.1 Ether Backup	Observaciones
desconect.	conect.	desconect.	conect.	Fallan lan4 y lan1. lan4 conmuta a lan3 y lan1 conmuta a lan0.
desconect.	conect.	desconect.	desconect.	Falla lan4. lan4 conmuta a lan3. Fallan lan1 y lan0. El paquete corriendo en el nodo se mueve al otro nodo.
desconect.	desconect.	conect.	conect.	Fallan lan4 y lan3. El paquete corriendo en el nodo se mueve al otro nodo. SUBNET 13.50.70.0 DOWN.
desconect.	desconect.	conect.	desconect.	Fallan lan4 lan3 y lan0. El paquete corriendo en el nodo se mueve al otro nodo. SUBNET 13.50.70.0 DOWN.
desconect.	desconect.	desconect.	conect.	Fallan lan4 lan3 y lan1. lan 1 conmuta a lan0. El paquete corriendo en el nodo se mueve al otro nodo. SUBNET 13.50.70.0 DOWN.
desconect.	desconect.	desconect.	desconect.	Se pierde el heartbeat. Uno de los nodos (el que no gane el acceso al lock disk) realiza una secuencia de TOC (Transfer Of Control) y el paquete que estaba corriendo en él se mueve al otro nodo.

Tabla 11.4.1 Matriz de pruebas de red realizadas en el Cluster.

11.5 MANTENIMIENTO

Como en todo sistema el mantenimiento preventivo y correctivo es vital para el aseguramiento de calidad de servicio del mismo. El mantenimiento preventivo a los equipos debe darse cuatro veces al año como mínimo y consiste en hacer un cambio de filtros contenedores de polvo, limpieza interior de tarjetas, unidades de discos, etc., y revisión de bitácoras del sistema Operativo y de los componentes de Hardware. El mantenimiento correctivo solo se dará en el caso de que las aplicaciones fallen a causa de un "BUG" o de que un componente de hardware falle.

11.5.1 Responsabilidades del Administrador del Sistema.

El Administrador del Sistema deberá estar entrenado para estar a la expectativa de cualquier problema que deba ser atendido por MC/ServiceGuard por lo tanto se debe asumir que éste mismo debe estar completamente familiarizado con el funcionamiento de MC/ServiceGuard.

Los siguientes puntos son un resumen de las responsabilidades que el administrador debe de asumir. Las acciones que éste debe tomar van a ser discutidas con más detalle en las siguientes secciones:

1. Levantamiento y baja total del Sistema.

2. Monitoreo del Status del Sistema
   a) Monitoreo de la Consola de Mensages
   b) Monitoreo de los Archivos de Registro de Operación (Logs Fails)
   c) Monitoreo del Status de los Clusters.

3. Recuperación del Sistema
   a) Durante una Falla Local
   b) Durante una Falla Remota
   c) En un futuro, durante un Site Failover

4. Mantenimiento del Sistema

11.5.2 Reboot del Sistema

Los equipos de Transferencia (MT1, MT2), Procesamiento (MP1, MP2) y Almacenamiento (MA) pueden ser "booteados" después de cargar nuevo software requerido. La referencia está en éste subcapítulo de manteniniento, como debe ser cargado el software adicional en los equipos y como se debe "bootear" (reiniciar) el equipo.

11.5.3 Startup Clusters

Es recomendable que se lleve a cabo la secuencia de pasos para levantar los sistemas. Debe primero ejecutarse un normal startup en cada uno de los nodos del sistema y posteriormente levantar todos los paquetes que tengan habilitado el switcheo de MC/ServiceGuard.

11.5.4 Startup de los Clusters MP (Inicialización de los Clusters MP y sus paquetes)

Para poner en marcha los Clusters MP, se deben llevar a cabo los siguientes pasos después que el equipo ha sido puesto en operación:

1. Abrir sesión en alguno de los nodos del Cluster (MP1 ó MP2) como el usuario "root".
   # su - root
2. Ejecutar el comando "cmruncl"
   # cmruncl

Cuando éste comando es ejecutado el shell script "control.sh", localizado en el directorio /etc/cmcluster/MP?, es automáticamente ejecutado por MC/ServiceGuard.

El shell script "control.sh" realiza las siguientes operaciones:

a) Activa los grupos de volúmenes, volúmenes lógicos
b) Monta los file system
c) Asigna la dirección RIP a los paquetes
d) Levanta las aplicaciones de cada uno de los paquetes

La tabla 11.5.1 muestra los grupos de volúmenes y los volúmenes lógicos que deben ser activados, los respectivos file system que deben ser montados, las direcciones RIP que deben ser asignadas y los paquetes de aplicaciones que deben ser levantados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla 11.5.1 Inicialización de los Clusters MP

11.5.5 Startup de los Clusters MT (Inicialización de los Clusters MT y sus paquetes)

Para poner en marcha los Clusters MT, se deben llevar a cabo los siguientes pasos después que el equipo ha sido puesto en operación:

1. Conectarse a alguno de los nodos del Cluster (MT1 ó MT2) como el usuario "root".
   # su - root
2. Ejecutar el comando "cmruncl"
   # cmruncl

Cuando éste comando es ejecutado el shell script "control.sh", localizado en el directorio /etc/cmcluster/MT?, es automáticamente ejecutado por MC/ServiceGuard.

El shell script "control.sh" realiza las siguientes operaciones:

a) Activa los grupos de volúmenes, volúmenes lógicos
b) Monta los file system
c) Asigna la dirección RIP a los paquetes
d) Levanta las aplicaciones de cada uno de los paquetes

La tabla 11.5.2 muestra los grupos de volúmenes y los volúmenes lógicos que deben ser activados, los respectivos file system que deben ser montados, las direcciones RIP que deben ser asignadas y los paquetes de aplicaciones que deben ser levantados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla 11.5.2 Inicialización de los Clusters MT

11.5.6 Startup de el equipo de almacenamiento MA (Levantamiento del equipo MA y sus paquetes)

Este equipo no cuenta con el software de MC/ServiceGuard, ya que no está dentro de la alta disponibilidad requerida para el proyecto.

11.5.7 Como se deben verificar los procedimientos de Startup

1. Verificar el status del Cluster

Para verificar el status de el Cluster, el Administrador del equipo debe ejecutar el comando cmviewcl y realizando lo siguiente:

a. Conectarse a un nodo del Cluster como el usuario "root".
b. Ejecutar el comando

/usr/sbin/cmviewcl

La siguiente pantalla será desplegada en pantalla:

```
CLUSTER        STATUS

MT_12          up

  NODE           STATUS        STATE
  mt1s           up            running

    PACKAGE        STATUS        STATE         PKG_SWITCH    NODE
    MT1            up            running       enabled       mt1s

  NODE           STATUS        STATE
  mt2s           up            running

    PACKAGE        STATUS        STATE         PKG_SWITCH    NODE
    MT2            up            running       enabled       mt2s
```

Aquí el Administrador debe asegurarse de que todos los paquetes tienen el status de levantados y corriendo ("up" and "running").

"PKG_SWITCH" debe estar habilitado ("enabled") y el paquete debe estar ejecutando en su nodo original.

Para ver de manera más detallada la información, se puede ejecutar el siguiente comando:

/usr/sbin/cmviewcl –v

La siguiente salida será desplegada en el monitor del Operador:

```
~
CLUSTER       STATUS
~
MT_12         up
~

~
  NODE          STATUS       STATE
~
  mt1s          up           running

    Network_Parameters:
    INTERFACE     STATUS       PATH         NAME
    PRIMARY       up           0/20.1       lan0
    STANDBY       up           0/44.1       lan1
    PRIMARY       up           2/4          lan2
    STANDBY       up           2/12         lan3

    PACKAGE       STATUS       STATE        PKG_SWITCH    NODE
    Mt1           up           running      enabled       mt1s

      Script_Parameters:
      ITEM          STATUS       NAME          MAX_RESTARTS  RESTARTS
      Service       up           ORACLE_MT1            0          0
      Service       up           TPS_MT1               2          0
      Service       up           SQLNET_MT1            4          0
      Service       up           DAR_MT1               4          0
      Service       up           ALARM_MT1             8          1
      Service       up           CD_MT1                4          0
      Subnet        up           13.49.152.0
      Subnet        up           13.50.70.0

      Node_Switching_Parameters:
      NODE_TYPE     STATUS       SWITCHING     NAME
      Primary       up           enabled       mt1s        (current)
      Alternate     up           enabled       mt2s

  NODE          STATUS       STATE
  mt2s          up           running

    Network_Parameters:
    INTERFACE     STATUS       PATH         NAME
    PRIMARY       up           0/20.1       lan0
    STANDBY       up           0/44.1       lan1
    PRIMARY       up           2/4          lan2
    STANDBY       up           2/12         lan3

    PACKAGE       STATUS       STATE        PKG_SWITCH    NODE
    MT2           up           running      enabled       mt2s

      Script_Parameters:
```

```
ITEM          STATUS        NAME          MAX_RESTARTS   RESTARTS
Service       up            ORACLE_MT2           0           0
Service       up            TPS_MT2              2           0
Service       up            SQLNET_MT2           4           0
Service       up            DAR_MT2              4           0
Service       up            ALARM_MT2            8           1
Service       up            CD_MT2               4           0
Subnet        up            13.49.152.0
Subnet        up            13.50.70.0

Node_Switching_Parameters:
NODE_TYPE     STATUS        SWITCHING     NAME
Primary       up            enabled       mt2s        (current)
Alternate     up            enabled       mt1s
```

Aquí debemos asegurarnos de que todas las interfaces de red, servicios y subredes están habilitadas ("up").

El Nodo Primario "Primary" y el Alternativo "Alternate" deben estar levantados "up" y además habilitados "enabled" para "SWITCHING".

2. Monitoreo de las bitácoras de sistema (Log Files)

Para verificar la bitácora del equipo se debe hacer lo siguiente:

a) Conectarse a uno de los nodos de el Cluster como usuario "root"
b) Ejecutar el siguiente comando:

tail –30 /var/adm/syslog/syslog.log | more

Un desplegado similar al que a continuiación se presenta, se desplegará en pantalla:

```
Sep  3  14:30:23 hpsgnoh cmcld[1992] : Started package MT1 on
Node hpsgnoh

~
Sep  3  14:30:23 hpsgnoh cmcld[1992] : Started package MT1 on
Node hpsgnoh
```

Si los mensajes, como los anteriores, no son desplegados, esto significa que algún error ha ocurrido. Para su solución se puede consultar la sección de Fallas en el Sistema y Procedimiento de Recuperación de ésta Tesis.

Se pueden revisar los siguientes archivos de log para obtener información adicional, si el resultado del "tail" no es satisfactorio:

/var/adm/cmcluster/HA_MT?.log o
/var/adm/cmcluster/HA_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log representa la salida del resultado de la corrida del shell script "control.sh".

/var/adm/cmcluster/HA_MT1.log conserva los mensajes enviados durante el arranque del Paquete MT1.

/var/adm/cmcluster/MT?/control_MT?.sh.log o
/var/adm/cmcluster/MP?/control_MP?.sh.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene algunos otros mensajes relacionados con la corrida del script de control.

/var/adm/cmcluster/services_MT?.log o
/var/adm/cmcluster/services_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene la Salida de las Aplicaciones activadas y del monitoreo de los servicios de Oracle.

3. Verificar que los File System estén montados

Asegurarse de que los file system listados en las tablas 11.5.3 y 11.5.4, estén correctamente montados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla 11.5.3 Inicialización de los Clusters MT

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla 11.5.4 Inicialización de los Clusters MP

4. Revisar los Procesos que deben levantar cada una de las aplicaciones.

La tabla 11.5.5 enuncia los procesos en el Sistema que deben estar corriendo por aplicación, de las involucradas en los paquetes.

Software o Aplicación	Procesos	Equipo
Oracle	Ora_pmon_<ORACLE_SID> Ora_dbwr_<ORACLE_SID> Ora_arch_<ORACLE_SID> Ora_lgwr_<ORACLE_SID> Ora_ckpt_<ORACLE_SID> Ora_smon_<ORACLE_SID>	MT, MP
Aplicación de Transferencia Y procesamiento (TPS)	Manmbin Monsbin Mistsbin Manfsbin	MT, MP
CONNECT:Direct	Cdpmgr	MT, MP
SQLNET	Tnslsnr	MT, MP
Dar	Darput	MT
Omnistorege	Ded Qr	MT
Omniback	Crs Lockmgr	MT, MP
Alarm daemon	Alarm.sh	MT, MP

Tabla 11.5.5 Lista de Procesos

5. Si los paquetes no pueden ser levantados en su nodo original, MC/ServiceGuard puede tratar de levantarlos en el nodo adoptivo.

Si los pasos 1 a 4 no son realizados satisfactoriamente, se debe investigar la causa del problema y tratar de resolverlo. Cuantas veces sea necesario, se deben ejecutar los pasos de inicialización hasta que el Cluster quede trabajando correctamente.

11.5.8 Shutdown de los Clusters

Es recomendado que se lleve a cabo la secuencia de pasos para dar de baja los Clusters. Antes de ejecutarse un normal shutdown en cada uno de los nodos del sistema se deben de detener todos los paquetes que estén habilitados siguiendo los siguientes pasos:

11.5.9 Shutdown de los Clusters MP (Paro de los Clusters MP y sus paquetes)

Para detener los Clusters MP, se deben llevar a cabo los siguientes pasos antes de que el equipo sea detenido:

1. Conectarse a alguno de los nodos del Cluster (MP1 ó MP2) como el usuario "root".
   $ su - root
2. Ejecutar el comando "cmhaltcl"
   # cmhaltcl –f

Cuando éste comando es ejecutado el shell script "control.sh", localizado en el directorio /etc/cmcluster/MP?, es automáticamente ejecutado por MC/ServiceGuard.

El shell script "control.sh" realiza las siguientes operaciones:

a) Detiene todos los paquetes y las aplicaciones incluidas en cada paquete
b) Remueve la dirección RIP a los paquetes
c) Desmonta los file system
d) Desactiva los grupos de volúmenes
e) Detiene el Cluster

La tabla 11.5.6, file system MP, muestra los grupos de volúmenes y los volúmenes lógicos activados, los respectivos file system montados, las direcciones RIP asignadas y los paquetes de aplicaciones levantados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla 11.5.6 File System MP

11.5.10 Shutdown de los Clusters MT (Paro de los Clusters MT y sus paquetes)

Para detener los Clusters MT, se deben llevar a cabo los siguientes pasos antes de que el equipo sea detenido:

1. Conectarse a alguno de los nodos del Cluster (MP1 ó MP2) como el usuario "root".
   $ su - root
2. Ejecutar el comando "cmhaltcl"
   # cmhaltcl –f

Cuando éste comando es ejecutado el shell script "control.sh", localizado en el directorio /etc/cmcluster/MT?, es automáticamente ejecutado por MC/ServiceGuard.

El shell script "control.sh" realiza las siguientes operaciones:

a) Detiene todos los paquetes y las aplicaciones incluidas en cada paquete
b) Remueve la dirección RIP a los paquetes
c) Desmonta los file system
d) Desactiva los grupos de volúmenes
e) Detiene el Cluster

La tabla 11.5.7, file system MT, muestra los grupos de volúmenes y los volúmenes lógicos activados, los respectivos file system montados, las direcciones RIP asignadas y los paquetes de aplicaciones levantados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla 11.5.7 File System MT

11.5.11 Shutdown de el equipo de almacenamiento MA (Halt del equipo MA y sus paquetes)

Este equipo no cuenta con el software de MC/ServiceGuard, ya que no entra dentro de la alta disponibilidad requerida para el proyecto.

11.5.12 Como se deben verificar los procedimientos de Shutdown

1. Verificar el status del Cluster

Para verificar el status de el Cluster, el Administrador del equipo debe ejecutar el comando cmviewcl y realizando lo siguiente:

a) Conectarse a un nodo del Cluster como el usuario "root".
b) Ejecutar el comando

/usr/sbin/cmviewcl

El siguiente mensaje de error será desplegada en pantalla:

```
CLUSTER        STATUS
~
               down
~
```

2. Monitoreo de las bitácoras de sistema (Log Files)

Para verificar la bitácora del equipo se debe hacer lo siguiente:

a) Conectarse a uno de los nodos de el Cluster como usuario "root"
b) Ejecutar el siguiente comando:
tail –30 /var/adm/syslog/syslog.log | more

Un desplegado similar al que a continuiación se presenta, se desplegará en pantalla:

```
Sep  4  15:25:03 hpsgnoh cmcld[1992] : Halted package MT1 on
Node hpsgnoh
~
Sep  4  15:25:03 hpsgnoh cmcld[1992] : Halted package MT1 on
Node hpsgnoh
```

Se pueden revisar los siguientes archivos e log para obtener información adicional, si el resultado del "tail" no es satisfactorio:

/var/adm/cmcluster/HA_MT?.log o
/var/adm/cmcluster/HA_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log representa la salida de el resultado de la corrida de el shell script "control.sh".

/var/adm/cmcluster/HA_PT1.log conserva los mensajes enviados durante el arranque de el Paquete PT1.

/var/adm/cmcluster/MT?/control_MT?.sh.log o
/var/adm/cmcluster/MP?/control_MP?.sh.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene algunos otros mensajes relacionados con la corrida del script de control.

/var/adm/cmcluster/services_MT?.log o
/var/adm/cmcluster/services_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene la Salida de la Aplicación Iniciadas y de el monitoreo de los servicios de Oracle.

3. Verificar que los File System estén desmontados

Asegurarse de que los file system listados en las tablas 11.5.8 y 11.5.9, estén correctamente desmontados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla 11.5.8 Inicialización de los Clusters MT

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla 11.5.9 Inicialización de los Clusters MP

4. Revisar los Procesos que deben ser detenidos por cada una de las aplicaciones.

La tabla 11.5.10 lista los procesos en el Sistema que deben estar detenidos por aplicación, involucardas en los paquetes.

Software o Aplicación	Procesos	Equipo
Oracle	Ora_pmon_<ORACLE_SID> Ora_dbwr_<ORACLE_SID> Ora_arch_<ORACLE_SID> Ora_lgwr_<ORACLE_SID> Ora_ckpt_<ORACLE_SID> Ora_smon_<ORACLE_SID>	MT, MP
Aplicación de Transferencia Y procesamiento (TPS)	Manmbin Monsbin Mistsbin Manfsbin	MT, MP
CONNECT:Direct	Cdpmgr	MT, MP
SQLNET	Tnslsnr	MT, MP
Dar	Darput	MT
Omnistorege	Ded Qr	MT
Omniback	Crs Lockmgr	MT, MP
Alarm daemon	Alarm.sh	MT, MP

Tabla 11.5.10 Lista de Procesos

11.5.13 Recomendaciones.

1. NO REMOVER NINGÚN ARCHIVO NI DIRECTORIO QUE SE ENCUENTRE BAJO EL SUBDIRECTORIO /etc/cmcluster EN AMBOS NODOS DEL CLUSTER.

2. Cuando se tenga que dar de baja el sistema, no use el comando reboot, hágalo con **shutdown -r -y 0**.

3. Administre con la perspectiva de cluster y no con la perspectiva de sistemas aislados, de tal forma que cuando haga cambios en un nodo que impacten la operación del paquete y tenga la certeza de que funcionan, aplíquelos al otro nodo. De tal forma que si hace un cambio en el ambiente bajo el cual corre la aplicación o en el ambiente de un usuario, éste debe reflejarse en los otros nodos, ya que el paquete y los usuarios buscarán el mismo ambiente en cualquier nodo.

4. Cuando haga cambios a la configuración del cluster adicione Volume Groups compartidos por los nodos del cluster, expórtelos tan pronto como sea posible usando los procedimientos ya indicados anteriormente.

5. Cuando necesite importar Volume Groups en un nodo ejecute previamente el comando **vgcfgbackup** en ese nodo.

6. Cuando adicione Volume Groups a un paquete, después de aplicar la nueva configuración y arrancar el cluster, detenga el paquete y use el comando **vgchange -a n** en todos los Volume Groups del paquete y vuelva a arrancar el paquete.

7. No coloque entradas con File System que se encuentran en Volume Groups compartidos en el archivo **/etc/fstab**.

8. No ejecute el comando **kill** sobre los demonios **cmcld, y cmlvmd**, ya que son los usados por el cluster.

9. Periódicamente cheque el estado del cluster con los comandos **cmviewcl** y **cmviewcl -v**.

10. Periódicamente cheque el estado de los Volume Groups de los paquetes desde el nodo donde esté corriendo el paquete.

11. Al menos una vez al mes mover los paquetes en el cluster, con el objeto de asegurar que los posibles cambios hechos al cluster y los paquetes funcionen bien tanto en uno como en otro nodo y asegurar que en caso de contingencia de un nodo, el otro pueda proporcionar los servicios sin ningún problema.

En cuanto al hardware con el que cuenta el cluster, se recomienda lo siguiente:

- Contar con un esquema completo de redundancia en datos, teniendo espejo tanto del sistema operativo como del software de aplicaciones almacenado en los Volume Groups vg00 y vg01, con el cual no cuenta actualmente ningún nodo del cluster. Los discos que se usen como espejo se recomienda conectarlos a una tarjeta controladora diferente a la que están conectados los discos originales. Es recomendable contar con éste espejo, ya que en caso de dañarse alguno de los discos de los Volume Groups mencionados anteriormente el paquete que se encuentre corriendo en ese nodo se tendrá que mover al otro nodo de manera innecesaria.

- Conectar cables con terminadores en línea (**C2980A**) en cada tarjeta controladora de discos de cada servidor del cluster. Ya que actualmente, en caso de que alguna controladora se dañe y se tenga que cambiar, se tendrá que apagar el servidor al que pertenece, lo cual abre la cadena SCSI, dejándola sin terminador y causando problemas en el otro nodo del cluster, obligando a que el otro nodo tenga que apagarse también.

11.6 Fallas del Sistema y Procedimiento de Recuperación

Existen tres tipos diferentes de fallas en las cuáles se requiere la intervención directa del Administrador de los equipos:

a) Local Failover/Failback

b) Remote Failover/Failback

c) Site Disaster/Recovery

11.6.1 Local Failover (Tarjeta de red dañada)

Cuando MC/ServiceGuard detecta que las tarjetas de Red han fallado puede recuperarse de éste failover haciendo un failback con la tarjeta de respaldo que se encuentra en cada uno de los equipos del Cluster.

Todas las aplicaciones continúan corriendo en la misma máquina después de la falla en la tarjeta de Red.

MC/ServiceGuard automáticamente, cuando detecta una falla en una tarjeta, que puede ser la misma tarjeta o el cableado, reconfigura la tarjeta de respaldo (standby LAN card) asignandole la dirección de RIP.

El administrador puede hacer una revisión para ver cuáles son las condiciones del Failover.

Las figuras 11.6.1 y 11.6.2 muestran el status del Cluster antes y despúes de un local failover .

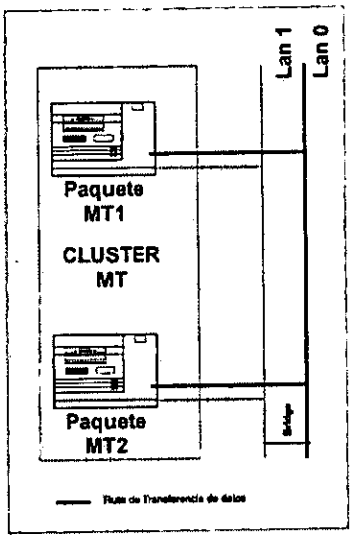


Fig. 11.6.1 Cluster Antes de un Failover

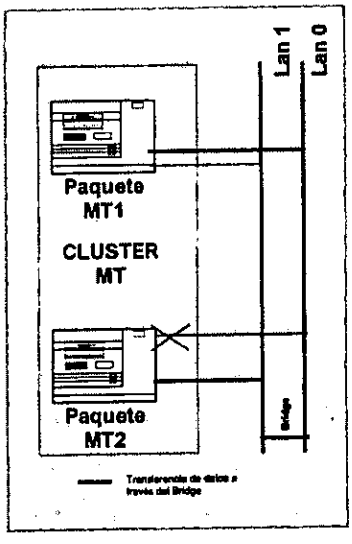


Fig. 11.6.2 Cluster Después de un Failover

11.6.1.1 Mensajes en Consola

El administrador puede revisar los mensajes en la consola. El siguiente mensaje es desplegado cuando la tarjeta de Red o el cableado falla:

```
Network NS_LS_DRIVER Disaster 1029, Pid [ICS]
  LAN card on interface unit 0 has network problem.
  Check cable for possible disconnection.
```

11.6.1.2 Revisión de los Mensajes en los Archivos de Log del Sistema

El archivo de log localizado en:

/var/adm/syslog/syslog.log

puede contener mensajes relevantes cuando la falla local o remota ocurren. Si tratamos de conectarnos de este equipo a sí mismo mediante un telnet o un login, como root, el mensaje que obtendremos será el siguiente:

Revisar el syslog.log haciendo lo siguiente:

a) Conectarse a uno de los Clusters como el usuario "root"
b) Ejecutar la siguiente operación:

tail –50 /var/adm/syslog/syslog.log | more

El mensage para la falla local en el syslog.log será:

```
Sep  4  11:26:03 hpsgnoh cmcld[4932] : lan0 failed
~
Sep  4  11:26:03 hpsgnoh cmcld[4932] : lan0 switched with lan1
Sep  4  11:26:05 hpsgnoh cmcld[4932] : Local switch has ocurred
```

El mensaje encontrado en caso de que la LAN en Standby fallara sería:

```
Sep  4  11:26:06 hpsgnoh cmcld[4932] : lan1 failed
```

11.6.2 Status del Cluster

El administrador debe estar monitoreando regularmente el status del Cluster.

Para revisar el status del Cluster, el administrador debe ejecutar el comando cmviewcl haciendo lo siguiente:

a) Conectarse a uno de los nodos del Cluster como "root"
b) Ejecutar la siguiente operación:

#/usr/sbin/cmviewcl –v

La siguiente salida va a ser desplegada en el monitor:

```
CLUSTER       STATUS
~
MT_12         up
~
~
  NODE          STATUS        STATE
~
  mt1s          up            running

    Network_Parameters:
    INTERFACE     STATUS        PATH          NAME
    PRIMARY       down            0/20.1        lan0
    STANDBY       up            0/44.1        lan1
    PRIMARY       up            2/4           lan2
    STANDBY       up            2/12          lan3

    PACKAGE       STATUS        STATE         PKG_SWITCH    NODE
    Mt1           up            running       enabled       mt1s

      Script_Parameters:
      ITEM          STATUS        NAME          MAX_RESTARTS  RESTARTS
      Service       up            ORACLE_MT1              0         0
      Service       up            TPS_MT1                 2         0
      Service       up            SQLNET_MT1              4         0
      Service       up            DAR_MT1                 4         0
      Service       up            ALARM_MT1               8         1
      Service       up            CD_MT1                  4         0
      Subnet        up            13.49.152.0
      Subnet        up            13.50.70.0

      Node_Switching_Parameters:
      NODE_TYPE     STATUS        SWITCHING     NAME
      Primary       up            enabled       mt1s          (current)
      Alternate     up            enabled       mt2s

  NODE          STATUS        STATE
  mt2s          up            running

    Network_Parameters:
```

```
INTERFACE     STATUS       PATH         NAME
PRIMARY       up           0/20.1       lan0
STANDBY       up           0/44.1       lan1
PRIMARY       up           2/4          lan2
STANDBY       up           2/12         lan3

PACKAGE       STATUS       STATE        PKG_SWITCH    NODE
MT2           up           running      enabled       mt2s

  Script_Parameters:
  ITEM          STATUS       NAME          MAX_RESTARTS   RESTARTS
  Service       up           ORACLE_MT2           0            0
  Service       up           TPS_MT2              2            0
  Service       up           SQLNET_MT2           4            0
  Service       up           DAR_MT2              4            0
  Service       up           ALARM_MT2            8            1
  Service       up           CD_MT2               4            0
  Subnet        up           13.49.152.0
  Subnet        up           13.50.70.0

  Node_Switching_Parameters:
  NODE_TYPE     STATUS       SWITCHING     NAME
  Primary       up           enabled       mt2s        (current)
  Alternate     up           enabled       mt1s
```

Si encontramos que la interface de Red primaria está abajo (status – "down"), eso significa que una falla está ocurriendo. El administrador debe revisar si esto es el resultado de un problema de conexión o de hardware.

11.6.3 Tiempo de Recuperación

El tiempo que se lleva en detectar y recuperar una falla local (Local Failover) es de aproximadamente 10 segundos. La aplicación no es reiniciada a causa de ésta falla. La aplicación puede continuar trabajando haciendo uso de la tarjeta de respaldo. La dirección RIP está en ese momento asociada a la LAN de respaldo.

11.6.4 LOCAL FAILBACK (Recuperación de la Comunicación en lan0)

Para recuperar el sistema de un Local failover, el administrador debe seguir los siguientes pasos:

1. Revisar el Log File

Revisar el syslog.log haciendo lo siguiente:

a) Conectarse a uno de los Clusters como el usuario "root"
b) Ejecutar la siguiente operación:

```
# tail -50 /var/adm/syslog/syslog.log | more
```

El mensaje para la falla local en el syslog.log será:

```
Sep  4  11:26:03 hpsgnoh cmcld[4932] : lan0 failed
"
Sep  4  11:26:03 hpsgnoh cmcld[4932] : lan0 switched with lan1
Sep  4  11:26:04 hpsgnoh cmcld[4932] : Local switch has ocurred
```

El mensaje encontrado en caso de que la LAN en Standby fallara sería:

```
Sep  4  11:26:05 hpsgnoh cmcld[4932] : lan1 failed
"
```

2. El siguiente comando nos puede ayudar a localizar fallas en la Red.

lanscan

Una salida similar a la siguiente es desplegada en pantalla con el uso de éste comando:

```
Hardware Station        Crd Hardware Net-Interface   NM MAC   HP DLPI Mjr
"
Path     Address        In# State    NameUnit State  ID Type  Support Num
"
0/20.1   0x080009B7612B 0   UP       lan0     UP     4  ETHER Yes     185
0/44.1   0x080009D01CBE 1   UP       lan1     DOWN   5  ETHER Yes     185
2/4      0x080009C44408 2   UP       lan2     DOWN   6  FDDI  Yes     191
2/12     0x080009C4D4EB 3   UP       lan3     DOWN   7  FDDI  Yes     191
```

La falla en la tarjeta da como resultado un "Hardware State" "DOWN".

3. Si la conexión del cable es el problema, podemos conectar otro cable y el failback es hecho por MC/ServiceGuard autómáticamente.

4. Si la tarjeta de LAN tiene problemas , hay que remplazarla y realizar los siguientes pasos:

a) Forzar un failover remoto del equipo.

b) Dar un shutdown al equipo y ponerlo en modo mantenimiento.

c) Remplazar la tarjeta de Red.

d) Levantar nuevamente el sistema.

e) Hacer un failback remoto.

11.6.5 Remote Failover (Un Nodo del Cluster Falla)

Un failover remoto sucede dentro de las siguientes dos condiciones:

a) . Cuando el equipo tiene problemas de Hardware o rebootea por completo.

b) .Cuando una aplicación falla

La figura 11.6.3 muestra el escenario en caso de que MT2 fallara.

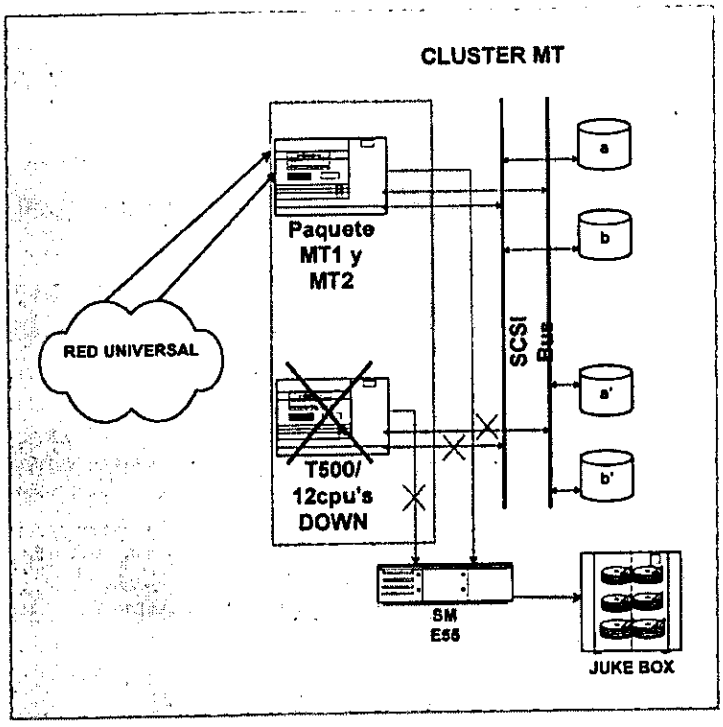


Fig. 11.6.3 MT2 Falla en el Cluster

La figura 11.6.3 es un ejemplo de la configuración del sistema estando en una situación de Failover Remoto. Los arreglos de discos (a,b) y (a',b') son accesados por ambos equipos dentro del Cluster. Pero ambos arreglos son bloqueados y usados por solo el Equipo adoptivo, que en este caso sería la MT1.

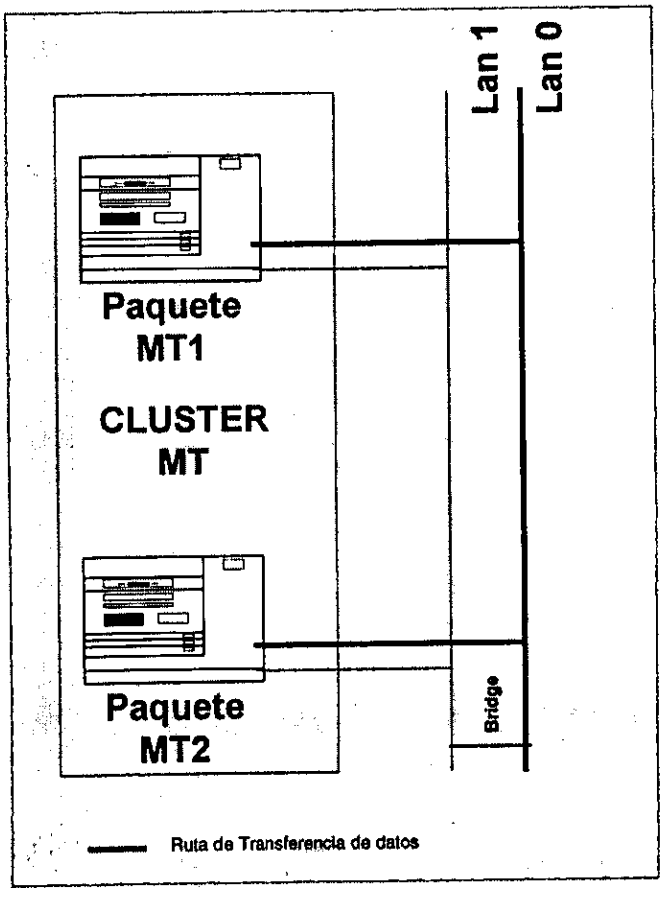


Fig. 11.6.4 Después de un Remote Failover

La figura 11.6.4 muestra esquemáticamente el regreso a la normalidad de los paquete MT1 y MT2 a sus nodos de orígen, después de un Fileover Remoto.

Una vez que MC/ServiceGuard detecta una falla de un nodo, éste ejecuta la operación "remote failover".

Existen dos operaciones principales durante el "remote failover":

a) Detenga los paquetes en el nodo que está fallando si están corriendo.

b) Levanta los paquetes en el nodo adoptivo.

Las siguientes operaciones son ejecutadas por el MC/ServiceGuard por el script de control. El script de control es el mismo script para detener o levantar los paquetes del Cluster. El Script de control está localizado en la siguiente ruta:

"/etc/cmcluster/MT? ó MT?/control.sh.

El script de control es capaz de ejecutar las siguientes operaciones:

a) Detener los paquetes en el nodo en falla si los paquetes están corriendo:

b) Detiene las aplicaciones haciendo un shutdown normal.

c) Desmonta los file system compartidos y desactiva los grupos devolúmenes.

d) Reconfigura las tarjetas de Red y mueve los paquetes hacia la dirección RIP.

e) Levantar los paquetes en el nodo adoptivo si los paquetes están corriendo:

f) Reconfigura las tarjetas de Red con las dirección de RIP.

g) Activa los grupos de volúmenes compartidos y monta los file system.

h) Revisa el status de las aplicaciones, y hace una recuperación de la Base de Datos si es necesario.

i) Reconfigura las aplicaciones en caso de ser necesario.

j) Levanta las aplicaciones del paquete, haciendo un startup normal.

11.6.6 Aplicaciones con falla

La tabla 11.6.1 muestra los tiempos en que Service Guard tarda en revisar el status de un servicio, el tiempo que espera para que un servicio sea restablecido y el número máximo de intentos en recuperar un servicio después de que falla.

Nombre del Servicio	Service Monitor/ Restart Script	Intervalo de Tiempo Para Revisar Proceso (segs)	Intervalo de Tiempo para Levantar el Servicio (segs)	Intentos de Recuperación del Servicio
ORACLE_MTI	/etc/cmcluster/monitor/ora.mon	5	0	0
TPS_MTI	/etc/cmcluster/monitor/tps.mon	5	450	2
OMNI_MTI	/etc/cmcluster/monitor/omni.mon	20	3600	4
OMNISTORE_MTI	/etc/cmcluster/monitor/omnistore.mon	20	3600	4
CD_MTI	/etc/cmcluster/monitor/cd.mon	20	1800	4
SQLNET_MTI	/etc/cmcluster/monitor/sqlnet.mon	20	1800	4
DAR_MTI	/etc/cmcluster/monitor/dar.mon	20	3600	4
ALARM_MTI	/etc/cmcluster/monitor/alarm.mon	5	900	8

Tabla 11.6.1 Configuración del Paquete (Servicios Críticos)

11.6.7 Configuración de los paquetes – Servicios o Aplicaciones Críticas

Todas las aplicaciones mostradas en la tabla anterior deben ser monitoreadas por el correspondiente script de monitoreo. El script revisa los procesos de las aplicaciones, la frecuencia para hacer la revisión está determinada por el intervalo de revisión de procesos (Process Checking Interval). Si uno de los procesos de alguna de las aplicaciones está abajo, MC/ServicGuard trata de restablecer la aplicación. Si la aplicación no puede ser restablecida satisfactoriamente, ésta espera un cierto tiempo (Restart Interval) mientras trata de hacer otro restart. Si el número de restart time sobrepasa su valor máximo configurado, MC/Service guard ejecuta un Remote failover.

El script de monitoreo es inicializado y continuamente revisado por el demonio de MC/ServiceGuard. Cada vez que el script de monitoreo detecta falla en las aplicaciones, éste hace, después de tratar de levantar la aplicación, un "exit" con código "1". Cada vez que MC/ServiceGuard detecta que el script manda una señal con código "1" éste incrementa el contador de restart times en 1.

Si uno de los procesos de la Aplicación Oracle, falla esta aplicación no puede ser levantada. Lo anterior sucede ya que si los procesos son matados de manera anormal, la memoria compartida, no puede ser liberada. Si se restablece Oracle, la memoria compartida no puede ser usada por completo, y si hay una falla en otro nodo, este nodo no va a estar disponible para tomar otro paquete de otro nodo debido a la falta de la memoria compartida. En este sentido, es necesario "rebootear" el equipo en caso de que Oracle falle.

Si el script de monitoreo detecta falla en los procesos de las aplicaciones configuradas como críticas, se pueden visualizar los siguientes mensajes en la consola del equipo:

```
ALARM:      Sep   1   13:02:08   -   SG node "hpsgnog": Omniback daemon
failure detected!

INFORMATIVE:      Sep  1      13:04:10   -   SG node "hpsgnog": Restarting
up Omniback.
```

En caso de que suceda un startup de las aplicaciones, se observaría lo siguiente:

```
INFORMATIVE:      Sep  1      13:10:23   -   SG node "hpsgnog": The Omniback
Server started.
```

Si el startup de las aplicaciones fallara:

```
ERROR:      Sep  1      13:10:23   -   SG node \"$(hostname)": Startup of
Omniback daemon failed!.
```

El administrador del sistema puede tratar de resolver el problema en caso de que una falla ocurra. Puede apoyarse en la revisión de los siguientes archivos de log para buscar la causa del problema:

/var/adm/cmcluster/HA_MT?.log ó HA_MP?.log

/var/adm/cmcluster/service_MT?.log ó service_MP?.log

11.6.8 Revisión de un Remote Failover

La revisión de un remote failover puede realizarse en el nodo adoptivo. Por ejemplo si el equipo MT1 fallara, la verificación puede hacerse en el equipo MT2. De manera similar, si el equipo MP1 falla, la revisión puede hacerse en la máquina MP2.

1. Revisión de los mensajes en consola

Los siguientes mensajes pueden ser desplegados en la consola adoptiva, cuando un remote failover ocurre:

```
ALARM: REMOTE FAILOVER: Package MT2 is switched over to MC/SG node
"hpsgnog" at Sep  1  14:50:45  SGP  1998.

INFORMATIVE:   MC/SG node "hpsgnog" : Starting package MT2 at Sep   1
14:50:45  SGP  1998.
```

2. Verificar el status del Cluster

Para verificar el status del Cluster, el Administrador del equipo debe ejecutar el comando cmviewcl , realizando lo siguiente:

a) Conectarse a un nodo del Cluster como el usuario "root".
b) Ejecutar el comando
/usr/sbin/cmviewcl

El siguiente mensaje de error será desplegada en pantalla:

```
~
CLUSTER        STATUS
~
MT_12        up
~

~
  NODE           STATUS        STATE
~
  mt1s         up            running
~

~
    PACKAGE       STATUS        STATE         PKG_SWITCH    NODE
    MT1           up            running       enabled       mt1s
    MT2           up            running       disabled      mt1s

~
~
  NODE           STATUS        STATE
~
  mt2s         down            halted
```

3. Monitoreo de las bitácoras de sistema (Log Files)

Para verificar la bitácora del equipo se debe hacer lo siguiente:

a) Conectarse a uno de los nodos del Cluster como usuario "root"
b) Ejecutar el siguiente comando:
tail –30 /var/adm/syslog/syslog.log | more

Un desplegado similar al que a continuación se presenta, se desplegará en pantalla, realizando los pasos anteriores en el nodo sobreviviente:

```
Sep 13 08:32:22 mt1s cmcld[17494]: (mccm1s) Halted package MCCM1 on node
mccm1s.
Sep 13 08:35:33 mt1s CM-CMD[20903]: cmhaltpkg MCCM2
Sep 13 08:35:33 mt1s cmcld[17494]: Executing
'/etc/cmcluster/MT2/control.sh  stop' for package MT2.
Sep 13 08:35:35 mt1s CM-MCCM2[20920]: cmhaltserv ORACLE_MCCM2
```

```
Sep 13 08:35:38 mt1s CM-MCCM2[20940]: cmhaltserv TPS_MCCM2
Sep 13 08:35:40 mt1s CM-MCCM2[20961]: cmhaltserv ALARM_MCCM2
Sep 13 08:35:43 mt1s CM-MCCM2[20976]: cmhaltserv SQLNET_MCCM2
Sep 13 08:35:46 mt1s CM-MCCM2[21012]: cmhaltserv CD_MCCM2
Sep 13 08:35:48 mt1s CM-MCCM2[21021]: cmhaltserv DAR_MCCM2
Sep 13 08:35:51 mt1s syslog: su : + tty?? root-TPSadm
Sep 13 08:35:54 mt1s last message repeated 2 times
Sep 13 08:35:59 mt1s syslog: su : + tty?? root-oracle
Sep 13 08:36:00 mt1s syslog: su : + tty?? root-oracle
Sep 13 08:36:11 mt1s CM-MCCM2[21124]: cmmodnet -r -i 13.49.152.130
13.49.152.0
Sep 13 08:36:11 mt1s CM-MCCM2[21129]: cmmodnet -r -i 13.50.70.130
13.50.70.0
Sep 13 08:36:19 mt1s LVM[21184]: vgchange -a n /dev/vg1cm2
Sep 13 08:36:19 mt1s LVM[21189]: vgchange -a n /dev/vg2cm2
Sep 13 08:36:19 mt1s cmcld[17494]: Halted package MCCM2 on node mt1s.
Sep 13 08:40:46 mt1s CM-CMD[21446]: cmhaltpkg MCCM2
Sep 13 08:41:05 mt1s syslog: su : + ttyp2 ohernan-oracle
Sep 13 08:41:44 mt1s LVM[21535]: /sbin/vgchange -a e /dev/vg1cm2
Sep 13 08:41:55 mt1s LVM[21569]: /sbin/vgchange -a e /dev/vg2cm2
Sep 13 08:41:59 mt1s LVM[21589]: /sbin/vgchange -a e /dev/vg2cm2
Sep 13 08:42:09 mt1s syslog: su : + ttyp2 ohernan-oracle
Sep 13 08:42:52 mt1s syslog: su : + ttyp1 ohernan-TPSadm
Sep 13 08:46:28 mt1s cmcld[17494]: (mccm1s) Started package MCCM1 on node
mccm1s.
```

Se pueden revisar los siguientes archivos e log para obtener información adicional, si el resultado del "tail" no es satisfactorio:

/var/adm/cmcluster/HA_MT?.log o
/var/adm/cmcluster/HA_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log representa la salida del resultado de la corrida del shell script "control.sh".

/var/adm/cmcluster/HA_MT1.log conserva los mensajes enviados durante el arranque del Paquete MT1.
También podemos revisar:

/var/adm/cmcluster/services_MT?.log o
/var/adm/cmcluster/services_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene la Salida de la Aplicación de Proceso y del monitoreo de los servicios de Oracle. En caso de que los procesos de TPS u Oracle fallasen, podemos buscar los siguientes mensajes en el log de services_MP.log ó services_MT.log :

/etc/cmcluster/MT?/control.sh.log ó
/etc/cmcluster/MP?/control.sh.log

regresan una salida en caso de que el paquete falle o trabaje correctamente al ser levantado en el nodo adoptivo.

Si el paquete falla en el nodo adoptivo al ser levantado, el administrador puede revisar el log del script de control y detectar el problema. El siguiente comando debe ser ejecutado para habilitar el paquete en el nodo adoptivo, después de revisar y arreglar la falla:

#cmmodpkg –e –n <nodo adoptivo> < paquete que no se puede levantar>

Si el paquete no falla otra vez al ser levantado en el nodo adoptivo, y regresado a su nodo de origen después de que la falla ha sido arreglada en el nodo de origen.

4. Verificar que los File system estén desmontados

Asegurarse de que los file system listados en las tabla 11.6.2 y 11.6.3, estén correctamente desmontados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla 11.6.2 Inicialización de los Clusters MT

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla 11.6.3 Inicialización de los Clusters MP

5. Revisar los Procesos que deben ser detenidos por cada una de las aplicaciones.

Las tabla 11.6.4, lista los procesos en el Sistema que deben estar detenidos por aplicación, involucardas en los paquetes.

Software o Aplicación	Procesos	Equipo
Oracle	Ora_pmon_<ORACLE_SID> Ora_dbwr_<ORACLE_SID> Ora_arch_<ORACLE_SID> Ora_lgwr_<ORACLE_SID> Ora_ckpt_<ORACLE_SID> Ora_smon_<ORACLE_SID>	MT, MP
Aplicación de Transferencia Y procesamiento (TPS)	Manmbin Monsbin Mistsbin Manfsbin	MT, MP
CONNECT:Direct	Cdpmgr	MT, MP
SQLNET	Tnslsnr	MT, MP
Dat	Darput	MT
Omnistorege	Ded Qr	MT
Omniback	Crs Lockmgr	MT, MP
Alarm daemon	Alarm.sh	MT, MP

Tabla 11.6.4 Lista de Procesos

6.- Si el paquete no puede ser levantado en su nodo original, MC/ServiceGuard podría tratar de levantarlo en el nodo adoptivo.

11.6.9 Comportamiento de las Aplicaciones Durante un Failover

Se hace la observación, de que las funcionalidades del Sistema continúan trabajando de manera adecuada sin verse afectado la capacidad y el performance de los equipos debido a que un equipo puede correr las aplicaciones y nivelar la carga de trabajo de la otra máquina en problemas, de manera correcta.

Esta sección desglosa el comportamiento de las aplicaciones y de los cambios de éstas por la manera en que está diseñada y configurada la Alta Disponibilidad (High Availibility) en este Capítulo:

11.6.9.1 Oracle Server

Dos instancias de Oracle pueden estar corriendo en el equipo adoptivo, sin ningún problema. Los archivos de configuración de las bases de datos (initSID.ora) pueden existir ambos en el nodo primario y en el secundario.

Las instancias de Oracle pueden ser levantadas en modo archive log. La tabla 11.6.5 describe las acciones a tomar por Service Guard después de un Fileover, y sus archivos de configuración.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
$ORACLE_HOME/dbs/initSID.ora	Local	No	
$ORACLE_HOME/dbs/sgadefSID.ora	Local	Ejecutar un Cleanup después del Restore	

Tabla 11.6.5 Archivos de configuración para la Base de Datos

11.6.9.2 SQL *NET

Pueden estar corriendo de manera simultanea dos SQL *NET LISTENER en el equipo adoptivo. La configuración de los LISTENERS permiten ser levantados en cualquiera de los equipos. El listener del paquete que falla puede ser levantado en el nodo adoptivo después de un failover. La tabla 11.6.6 muestra los archivos de configuración de SQL *NET.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
/etc/listener.ora	Local	No	
/etc/sqlnet.ora	Local	No	
/etc/tsnames.ora	Local	No	

Tabla 11.6.6 Archivos de configuración para SQL *NET

El nombre de los LISTENERS están definidos como "LISTENER_MT? y LISTENER_MP?. Todos los Listeners definidos para los paquetes MT ó MP tienen diferentes números de puertos asignados.

11.6.9.3 OMNIBACK

Solamente es posible tener una instancia de ésta aplicación corriendo en el equipo adoptivo. El horario y cintas de respaldo pueden ser organizados para realizar los respaldos de los dos nodos del Cluster al mismo tiempo. Durante la operación normal, el respaldo de los directorios del otro nodo está configurado en sí mismo. Después de un failover, los file system del nodo adoptado deben ser respaldados con OMNIBACK haciendo las inclusiones de éste en sus respaldos ya configurados. Los scripts para PREEXEC y POSTEXEC deben ser modificados y adecuados para soportar los respaldos del nodo adoptivo, sobre todo los de ORACLE para congelar la Base de Datos.

Puede ser configurado un "media pool" para un failover. Los respaldos y horarios en OMNIBACK pueden ser configurados para cualquier situación, esto siempre y cuando todavía se cuenten con los recursos. Ver tabla 11.6.7 los archivos de configuración de Omniback.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
/etc/opt/omni/config/datalist/*	Local	No	
/etc/opt/omni/schedule	Local	No	

Tabla 11.6.7 Archivos de configuración para OMNIBACK

Hay que tener cuidado al configurar demasiados "datalists" ya que el tamaño de los archivos en el directorio "/etc/opt/omni/config/db/*" pueden crecer demasiado. Es mejor hacer una liga a otro file system, o en su defecto, programar un CRON para purgar la base de datos de OMNIBACK regularmente.

11.6.9.4 OMNISTORAGE CLIENT

El VBFS del sistema en falla puede ser montado en el nodo adoptivo y puede ser direccionado al "volume set" del nodo adoptivo. Solamente puede estar corriendo una instancia de OMNISTORAGE CLIENT en el nodo adoptivo.

Los VBFS file system pueden tener diferentes FSID, y la configuración de ambos file system pueden existir en ambos equipos. Los puntos de montado pueden ser direccionados al correspondiente volume sets en ambos equipos.

La tabla 11.6.8 muestra los archivos asociados a OMNISTORE CLIENT.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
/var/opt/omnistorage/data/volsets/vsbind.FSID	Local	No	

Tabla 11.6.8 Archivos de configuración para la OMNISTORAGE

11.6.9.5 CONNECT DIRECT

Solo puede existir una instancia corriendo en el equipo adoptivo. Todas las aplicaciones pueden mandar peticiones a un demonio de C:D (Connect Direct).

11.6.9.6 TPS

Deben existir dos instancias de TPS corriendo en el equipo adoptivo. Toda la configuración de TPS se encuentra en los discos compartidos, y son transferidos hacia el equipo adoptivo después de un failover. El archivo mostrado en la tabla 11.6.9 evita que dos instancias de la aplicación TPS se ejecute al mismo tiempo en un equipo, por tanto después de un fileover es necesario limpiar esta bandera para poder ejecutar la instancia adoptada.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
$TPS_CONFIG/tps/TPS_master.pid	Shared	Clean up	No

Tabla 11.6.9 Archivo de Lock para TPS

11.6.9.7 DAR

Existen dos instancias corriendo en el nodo adoptivo. Toda la configuración se encuentra en los discos compartidos, y son montados en el adoptivo después de un failover. Los archivos temporales y de bitácora (log files) también están ubicados en los discos compartidos.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
$TPS_HOME/etc/setup/dar.cfg	Shared	No	

Tabla 11.6.10 Archivo de configuración de DAR

11.6.9.8 Alarm

Se deben levantar dos instancias en el equipo adoptivo, el cuál manda mensajes de su respectiva MT instancia a la consola.

11.6.9.9 CRON del Sistema

Solo existe un archivo de Cron (crontab) para cualquier sistema. No existe cambio en los trabajos del Cron después de un failover remoto. El equipo que toma la responsabilidad del que falla. Cada trabajo en el Cron de cada equipo está configurado para detectar cuando ha existido un SwitchOver y los file system del orto equipo han sido adoptados.

11.6.10 Tiempo de Recuperación

El tiempo tomado para que la aplicación sea levantada en el equipo adoptivo es de aproximadamente 10 minutos, tomando en cuenta que MC/ServiceGuard tiene que realizar el shutdown de las aplicaciones (tirar las aplicaciones) en el equipo que está fallando y levantarlas por completo en el adoptivo.

Si el nodo se cae o se queda trabado no es necesario que MC/serviceGuard haga un shutdown de las aplicaciones primero en el otro nodo, solamente tiene que levantarlas en el adoptivo. El tiempo requerido se reduce, en este caso, a la mitad, es decir, 5 minutos.

11.6.11 Failback Remoto (El nodo que había fallado es recuperado)

Para regresar a su estado normal un nodo que ha fallado y ha tenido que hacerse un SwuitchOver a uno adoptivo, es necesario que éste retome su paquete de aplicaciones. Para esto se requiere una intervención manual.

El administrador en turno, debe realizar los siguientes pasos para levantar el paquete en el nodo que ha sido recuperado ejecutando esta secuencia de comandos de MC/ServiceGuard:

1. Si el nodo que falló fue reseteado abruptamente o sufrió una interrupción en la alimentación de energía, el Administrador debe primero hacer un reboot en éste. Si el SwuitchOver fue debido a que Oracle o TPS fallan, es recomendable que también se de un reboot al sistema primero.

2. Conectarse al equipo a recuperar como el usuario "root".

3. Levantar MC/ServiceGuard en el nodo y reintegrar el Cluster, ejecutando "cmrunnode"

cmrunnode

Si el nodo aún está corriendo, se debe restaurar éste del siguiente modo:

cmchaltnode
cmrunnode

4. Detenga el paquete en el nodo adoptivo ejecutando "cmcahltpkg"
cmchaltpkg –n <nodo adoptivo> <nombre del paquete>

5. Levante el paquete en el nodo recientemente levantado. Existen dos opciones para hacerlo:

a) Habilite la recuperación del nodo a la hora de levantar MD/ServiceGuard automáticamente y posteriormente ejecute los siguientes comandos:

cmmodpkg –e <nombre del paquete>

#cmmodpkg –e –n <nombre del nodo recuperado> <nombre del paquete>

ó

b) Levante el paquete manualmente ejecutando los siguientes comandos:

cmrunpkg –n <nombre del nodo recuperado> <nombre del paquete>

cmmodpkg –e <nombre del paquete>

Las siguientes operaciones son ejecutadas por el script de control de MC/ServiceGuard. El script de control es el mismo que se utiliza para detener o arrancar los paquetes de un Cluster: "/etc/cmcluster/MT? ó MP?/control.sh"

El "control.sh" shell script realiza estas operaciones:

Deteniendo los paquetes en el equipo adoptivo:

a) . Detiene las aplicaciones de los paquetes, haciendo un shutdown suave.
b) . Desmonta los file system compartidos y desactiva los grupos de volúmenes.
c) . Reconfigura las tarjetas de RED elimina la dirección de RIP.
d) . Reconfigura las aplicaciones si es necesario.

Levantando los paquetes en el nodo recuperado:

a) . Reconfigura las tarjetas de RED con la dirección RIP.
b) . Activa los grupos de volúmenes compartidos y monta los file system
c) . Levanta las aplicaciones.

11.6.12 Verificación de un Failback Remoto

La verificación de un Failback remoto puede hacerse en el equipo recuperado. Si MT1 falló, la verificación puede hacerse en la MT1 ya recuperada. De igual forma, si MP1 falló la verificación puede hacerse ya recuperada MP1.
La verificación se lleva a cabo de la siguiente manera:

1. Revisión de los mensajes en consola

Los mensajes pueden ser desplegados en la consola adoptiva, cuando un remote failback ocurre.

2. Verificar el status del Cluster

Para verificar el status del Cluster, el Administrador del equipo debe ejecutar el comando cmviewcl y realizando lo siguiente:

a) Conectarse a un nodo del Cluster como el usuario "root".
b) Ejecutar el comando
/usr/sbin/cmviewcl

3. Monitoreo de las bitácoras de sistema (Log Files)

Para verificar la bitácora del equipo se debe hacer lo siguiente:

a) Conectarse a uno de los nodos del Cluster como usuario "root"
b) Ejecutar el siguiente comando:
tail –30 /var/adm/syslog/syslog.log | more

Se pueden revisar los siguientes archivos e log para obtener información adicional, si el resultado del "tail" no es satisfactorio:

/var/adm/cmcluster/HA_MT?.log
/var/adm/cmcluster/HA_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log representa la salida del resultado de la corrida del shell script "control.sh". /var/adm/cmcluster/HA_PT1.log conserva los mensajes enviados durante el arranque del Paquete.
También podemos revisar:

/var/adm/cmcluster/services_MT?.log o
/var/adm/cmcluster/services_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene la Salida de la Aplicación de Proceso y del monitoreo de los servicios de Oracle. En caso de que los procesos de TPS u Oracle fallen, podemos buscar los mensajes correspondientes en el log de services_MP.log ó services_MT.log :

/etc/cmcluster/MT?/control.sh.log ó
/etc/cmcluster/MP?/control.sh.log

regresan una salida en caso de que el paquete falle o trabaje correctamente al ser levantado en el nodo adoptivo.

Si el paquete falla en el nodo al ser levantado, el administrador puede revisar el log del script de control y detectar el problema. El siguiente comando debe ser ejecutado para habilitar el paquete en el nodo, después de revisar y arreglar la falla:

#cmmodpkg –e –n <nodo adoptivo> < paquete que no se puede levantar>

4. Verificar que los File system estén montados

Asegurarse de que los file system listados en las tablas 11.6.11 y 11.6.12, estén correctamente montados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla 11.6.11 Inicialización de los Clusters MT

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla 11.6.12 Inicialización de los Clusters MP

5. Revisar los Procesos que deben ser levantados por cada una de las aplicaciones.

La tabla 11.6.13 lista los procesos en el Sistema que deben estar levantados por aplicación, involucrada en los paquetes.

Software o Aplicación	Procesos	Equipo
Oracle	Ora_pmon_<ORACLE_SID> Ora_dbwr_<ORACLE_SID> Ora_arch_<ORACLE_SID> Ora_lgwr_<ORACLE_SID> Ora_ckpt_<ORACLE_SID> Ora_smon_<ORACLE_SID>	MT, MP
Aplicación de Transferencia Y procesamiento (TPS)	Manmbin Monsbin Mistsbin Manfsbin	MT, MP
CONNECT:Direct	Cdpmgr	MT, MP
SQLNET	Tnslsnr	MT, MP
Dar	Darput	MT
Omnistorege	Ded Qr	MT
Omniback	Crs Lockmgr	MT, MP
Alarm daemon	Alarm.sh	MT, MP

Tabla 11.6.13 Lista de Procesos

6. Si el paquete no puede ser levantado en su nodo original,
MC/ServiceGuard podría tratar de levantarlo en el nodo adoptivo.

11.6.13 Comportamiento de las Aplicaciones Durante un Failback

Esta sección desglosa el comportamiento de las aplicaciones y de los cambios de estas por la manera en que está diseñada y configurada la Alta Disponibilidad (High Availibility) en éste proyecto:

11.6.13.1 Oracle Server

Una instancia de Oracle debe estar corriendo en el equipo recuperado problema.

La instancia de Oracle debe ser levantada en modo archive log.

11.6.13.2 SQL *NET

Debe estar corriendo un solo SQL *NET LISTENER en el equipo. La configuración de los LISTENERS permite ser levantados en cualquiera de los equipos. El listener del paquete que falla puede ser levantado nuevamente en el nodo adoptivo después de que el equipo ha sido recuperado.

El nombre de los LISTENERS están definidos como "LISTENER_MT? y LISTENER_MP?. Todos los Listeners definidos para los paquetes MT ó MP tienen diferentes números de puertos asignados.

11.6.13.3 OMNIBACK

Solamente es posible tener una instancia de esta aplicación corrienado en el equipo recuperado. El horario y cintas de respaldo son recuperados al arrancar OMNIBACK. Después de un failover, los file system del nodo adoptado deben ser respaldados con OMNIBACK haciendo las inclusiones de éste en sus respaldos ya configurados, cuando el failback ocurre esto debe ser modificado en el nodo adoptivo. Los scripts para PREEXEC y POSTEXEC deben ser regresados a su normalidad, sobre todo los de ORACLE para congelar la Base de Datos.

11.6.13.4 OMNISTORAGE CLIENT

El VBFS del sistema puede ser montado en el nodo origen y arrancado OMNISTORAGE. Solamente puede estar corriendo una instancia de OMNISTORAGE CLIENT en el nodo origen.

11.6.13.5 CONNECT DIRECT

Solo puede existir una instancia corriendo en el equipo origen. Todas las aplicaciones pueden mandar peticiones a un demonio de C:D (Connect Direct).

11.6.13.6 TPS

Debe existir una instancia de TPS corriendo en el equipo adoptivo. Toda la configuración de TPS se encuentra en los discos compartidos, para ser transferida hacia el equipo adoptivo después de un failover.

11.6.13.7 DAR

Existe una instancias corriendo en el nodo origen. Toda la configuración se encuentra en los discos compartidos, que son montados en el origen después de un failback. Los archivos temporales y de bitácora (log files) también están ubicados en los discos compartidos.

11.6.13.8 Alarm

Se debe levantar una instancia en el equipo origen, la cuál manda mensajes de su respectiva MT o MP instancia a la consola.

11.6.13.9 CRON del Sistema

Solo existe un archivo de Cron (crontab) para cualquier sistema. No existe cambio en los trabajos del Cron después de un failback remoto. Cada trabajo en el Cron de cada equipo está configurado para detectar cuando ha existido un SwitchOver y los file system del otro equipo han sido adoptados.

11.7 MANUALES

11.7.1 Manuales Técnicos y de Operación.

El manual Técnico y de operación es un condensado de esta Tesis. Por tanto, para evitar redundancia entre la información presentada en este Capítulo, lo hemos incluido en el Apéndice A.

11.7.2 Manual de Usuario

Este proyecto no requiere de hacer una división entre el Manual de Usuario y el Técnico ya que es una aplicación implementada para el uso de los equipos HP 9000 T500 y el Administrador de la misma es el Administrador del Sistema Operativo y de las máquinas.

CAPITULO 12
CONCLUSIONES

12.1 Aprendizaje obtenido

La cantidad de datos y los usuarios se incrementan cada día. Lo anterior conlleva a un aumento en el potencial de problemas de integridad, seguridad de los datos y la necesidad de siempre disponer de ellos. Los tres anteriores, son puntos vitales en toda estructura que simiente en los sistemas de cómputo para lograr su permanencia en el mercado competitivo.

El nivel de disponibilidad de un sistema debe ser establecido en función de la definición de requerimientos de información, procesamiento, velocidad de respuesta, seguridad e inversiones financieras contables que varían en proporción directa con el grado de disponibilidad requerido, sin olvidar que a mayor disponibilidad se tiene mayor costo. Del mismo modo, para implementar un sistema de alta disponibilidad se debe considerar los costos, problemas y sanciones, representados en un sistema, por los tiempos de operación discontinua.

Para esto debe tenerse bien claro a que grado de disponibilidad sirve para evitar perder información valiosa, clientes por largos periodos de espera, pagos por sueldos no desquitados por los empleados al estar los sistemas fuera, multas por prestar mal servicio y duplicidad en el cobro, etc.

Debido al avance tecnológico se recomienda hacer un estudio de mercado para conocer cuáles son las nuevas tecnologías propuestas por las compañías de hardware y software, con el fin de seleccionar las que sean más adecuadas a nuestras necesidades presentes y futuras en nuestra empresa.

No basta con comprar un robusto sistema de altamente disponible, también es necesario llevar a cabo una gran planeación y organización de nuestra seguridad y accesos restringidos al mismo, ya que hoy en día es muy común encontrar profesionales que se dedican al daño y robo de nuestra información. Por tal motivo, debemos definir una jerarquía bien diseñada de usuarios para el acceso y capacitar a nuestros operadores para la continua vigilancia, soporte y control de los sistemas.

Es importante definir un excelente esquema de respaldos confiables, que no sean más que el resultado de una combinación de diferentes tecnologías y posibilidades establecidas por la estructura, usuarios y tiempos de operación.

Otro punto fundamental, es cumplir con los requerimientos establecidos de alimentación de energía y clima que los proveedores especifican para nuestros equipos, si no es así, el proveedor puede cancelar la garantía de los equipos, e inclusive no soportarnos, y perder nuestra valiosa inversión.

Finalmente, el resultado de esta Tesis es la obtención un sistema funcional, robusto eficiente y eficaz, que satisface las necesidades de seguridad y disponibilidad que se requiere para el logro de los objetivos de operatividad. Del mismo modo la documentación generada esta fundamentada en una metodología que permitirá medir, corregir y mantener desviaciones que el sistema tenga en sus etapas de implantación y producción.

GLOSARIO

Aplication Failback: Después que la falla en un equipo es reparada, las aplicaciones deben ser regresadas de el equipo adoptivo a el equipo primario.

Aplication Failover: Cuando en el equipo primario fallan las aplicaciónes, el MS/ServiceGuard puede automáticamente executar las aplicaciones en otro equipo dentro del Cluster. Esto es llamado Aplication Failover.

ATM: Asynchronous Transmission Mode.

Bus: Un canal o ruta común entre dispositivos del hardware, ya sea internamente entre componentes de la computadora o externamente entre estaciones de una red de comunicaciones.

CCITT: Comité Consultivo Internacional de telegrafía y Telefonía (siglas en francés). Organización internacional que desarrolla estándares de comunicaciones, como la recomendación X.25.

Cluster: Es un grupo de equipos que son conectados en una RED y comparten las aplicaciones mediante discos compartidos. Existe un coordinador central dentro de este grupo de equipos que se encarga de monitorear los demás equipos dentro del grupo o Cluster. En este sistema en particular, la configuración de Clusters está dada por dos equipos recolectores y dos equipos tarificadores los cuales forman en pares dos Clusters.

DAT: Digital A.... Tape. Es una unidad de cinta para realizar backup. Provee al administrador de la red de un backup simple y seguro.

DLT Digital Linear Tape.

Dirección IP Relocalizable (Relocate IP Address RIP): La RIP es una dirección IP asociada a un paquete. Cuando este paquete es movido a otro equipo dentro del Cluster la RIP es también movida a ese otro equipo.

Dirección IP estacionaria: Es la Dirección asociada a la tarjeta LAN en un equipo.

Equipo Adoptivo: El equipo dentro del Cluster que levante las aplicaciones, en caso de una falla de alguna de éstas en el equipo primario.

Equipo Primario: Es el equipo en el cual actualmente está corriendo el Paquete de aplicaciones.

FDDI: Conjunto de estándares desarrollado por ANSI (American National Standards Institute). Puede considerarse como un estándar para LAN de alta velocidad o para MAN empleando Fibra óptica.

Halt Script: Este script es el que permite inicializar el paquete de aplicaciones (pakage) definido dentro del Cluster como importantes en caso de una falla y prioritarias para ser levantadas en el equipo que este en buen estado.

High Availability (HA): Es el mecanismo utilizado para minimizar el tiempo en que las aplicaciones están caídas durante una falla. Un nodo se designa como el Equipo Adoptivo para correr las aplicaciones cuando el Equipo Primario falla.

IPX/SPX: Internetworking Packet Exchange. Intercambio de paquetes de interconexión de redes. Protocolo Novell de capa 3, similar a XNS e IP que se emplea en redes NetWare.

ISO: International Organization for Standarization (Organización Internacional para la Estandarización). Organización internacional responsable de una amplia gama de estándares, incluyendo aquellos relevantes para las redes. ISO es la responsable del modelo de referencia de redes más popular: el modelo de referencia OSI.

IEEE: Institute of Electrical and Electronic Engineers (Instituto de Ingenieros Eléctricos y Electrónicos). Organización profesional que define estándares de redes.

Internet: Conjunto de redes interconectadas y que en forma genérica funciona como una sola.

INTRANET: Es una red informática privada que utiliza normas y protocolos de Internet, para permitir a los miembros de una organización comunicarse y colaborar entre sí con mayor eficacia, aumentando la productividad.

JukeBox: Almacena numerosos CD-ROM's. Puede ser un brazo mecánico o un carrusel que lleve el disco a una estación óptica para lectura o escritura

LAN: Local Area Network. Red de Área Local. Segmento de red con estaciones de trabajo o nodos y más dispositivos de red enlazados. Conjunto de segmentos de red interconectados en una región no muy extensa.

Login: Nombre de identificación del usuario ante un sistema de cómputo. Es de carácter público.

MC/ServiceGuard: El MC/ServiceGuard es el software que permite soportar las aplicaciones durante una falla en las mismas en el Equipo Adoptivo.

MODEM: Modulator Demodulatro. Dispositivo que adapta una terminal o computadora a una línea telefónica. Permite la coneaxión de dos computadoras vía telefónica ya sea mediante línea privada o línea conmutada.

Multiusuario: Permite dos o más usuarios ejecutar programas al mismo tiempo. Algunos sistemas operativos permiten cientos o hasta miles de usuarios al mismo tiempo.

Multiproceso: Soporta la ejecución de un programa en más de un CPU

Multitareas: Permite más de un sólo programa ejecutándose.

OSI Referencel Model: Modelo de arquitectura de redes desarrollado popr ISO y CCITT. Consiste en siete capas, cada una de las cuales especifica funciones particulares de la red, tales como direccionamiento, control de flujo, control de errores, encapsulamiento, transferencia confiable de mensajes y muchas otras. La capa más alta (capa de aplicación) es la más cercana al usuario. La capa más baja (capa física) es la más cercana al medio físico. El modelo de referencia OSI es universalmente usado como método de enseñar y entender la funcionalidad de las redes.

Paquete: Un Paquete consiste en un conjunto de aplicaciones o programas de aplicación. Un Paquete puede ser migrado a un Equipo Adoptivo para su ejecución en el mismo siempre y cuando el Equipo Primario falla.

SCSI: Small Computer System Interface. Estándar para discos y equipo periférico de los sistemas de cómputo.

SNMP: Simple Network Management Protocol. Protocolo simple de manejo de redes. Ofrece medios para seguir y determinar la configuración de la red y los parámetros al tiempo de ejecución.

Start Script: Es un script que puede ser ejecutado cuando el paquete de aplicaciones es inicializado en un nodo.

TCP/IP: Transmission Control Protocol/Internet Protocol. Protocolo de Control de Transmisiones/Protocolo Internet. Los dos protocolos Internet más conocidos, que erróneamente suelen confundirse como uno solo. TCP corresponde a la capa 4 (capa de transporte) del modelo de refencia OSI y ofrece transmisión confiable de datos. IP corresponde a la capa 3 (capa de red) del modelo de referencia OSI, y ofrece servicios de datagramas sin conexión.

TPS: Sistema de Teleproceso

Tiempo real: Responde a la entrada instantaneamente.

Username: Se utiliza para acceder al sistema de la computadora

Workstation: Estación de trabajo. Un tipo de computadora utilizada para aplicaciones ingenieriles, publicidad, etc. Los sistemas operativos más comunes para WS son UNIX y Windows NT.

100baseT: Un estándar de redes que soporta transferencia de datos de hastea 100 Mbps. Está basado en el estándar de Ethernet, pero es 10 veces más que éste.

10Base2: Es una de las varias adaptaciones a Ethernet para LAN. Utiliza clable coaxial de 50 ohms , con distancias máximas de 185 metros. Opera a 10 Mbps y utiliza transmisión en base banda.

BIBLIOGRAFIA

- CLUSTERS FOR HIGH AVAILABILITY
  Weygant, Peter S.
  Hewlett-Packard, Professional Books. Prentice Hall,
  New Jersey, USA. 1996.

- HANDS-On with High Availability and MC/Service Guard. HEWLETT PACKARD
  Student Workbook.
  USA. 25 de Noviembre de 1996.

- UNIX Survival Guide
  Tim Parker
  Ed. Adisson Wesley Publishing Company Inc
  U.S.A. 1990

- UNIX Security System
  Rik Farrow
  Ed. Adisson Wesley
  U.S.A. 1994

- Intranet
  Tyson Green
  Ed. Mc. Graw Hill
  Madrid, España 1997

HEMEROGRAFIA

- Soluciones Avanzadas
  Año 6, Número 60.
  Agosto 1998.
  Pags. 11-20
  México, D.F.
  ISN 0188-8048

PAGINAS WEB CONSULTADAS PARA TEMA DE HERRAMIENTAS HSM:

- http://www.storage.ibm.com/software/adsm/adhsmov.ht
- http://webopedia.internet.com/TERM/H/HSM.html
- http://support.cai.com/hsmsupp.html

APENDICE A
MANUAL TECNICO Y DE OPERACION

Puesta en marcha de High Availability (Alta Disponibilidad)

High Availability es lo puede garantizar la continuidad de las operaciones de los equipos de Transferencia y Procesamiento en caso de un simple problema de hardware. MC/ServiceGuard es un producto de software que habilita dos nodos dentro de un cluster en uno solo de los nodos del cluster en caso de que se tenga una simple falla.

Terminologías

Para hacer más sencillo el entendimiento de éste capítulo, a continuación citamos algunas terminologías y conceptos difíciles de traducir al español que serán usadas en el mismo:

a) Equipo Adoptivo
El equipo dentro del Cluster que levante las aplicaciones, en caso de una falla de alguna de éstas en el equipo primario.

b) Aplication Failback
Después que la falla en un equipo es reparada, las aplicaciones deben ser regresadas del equipo adoptivo al equipo primario.

c) Aplication Failover
Cuando en el equipo primario fallan las aplicaciones, el MS/ServiceGuard puede automáticamente ejecutar las aplicaciones en otro equipo dentro del cluster. Esto es llamado Aplication Failover.

d) Cluster
Es un grupo de equipos que son conectados en una RED y comparten las aplicaciones mediante discos compartidos. Existe un coordinador central dentro de éste grupo de equipos que se encarga de monitorear los demás equipos dentro del grupo o cluster. En éste sistema en particular, la configuración de clusters está dada por dos equipos transmisores y dos equipos procesadores los cuales forman en pares dos clusters.

e) Halt Script
Este script es el que permite iniciar el paquete de aplicaciones (pakage) definido dentro del cluster como importantes en caso de una falla y prioritarias para ser levantadas en el equipo que esté en buen estado.

f) High Availability (HA)
Es el mecanismo utilizado para minimizar el tiempo en que las aplicaciones están caídas durante una falla. Un nodo se designa como el Equipo Adoptivo para correr las aplicaciones cuando el Equipo Primario falla.

g) MC/ServiceGuard.
El MC/ServiceGuard es el software que permite soportar las aplicaciones durante una falla en las mismas en el Equipo Adoptivo.

h) Paquete
Un Paquete consiste en un conjunto de aplicaciones o programas de aplicación. Un Paquete puede ser migrado a un Equipo Adoptivo para su ejecución en el mismo siempre y cuando el Equipo Primario falla.

i) Equipo Primario
Es el equipo en el cual actualmente está corriendo el Paquete de aplicaciones.

j) Dirección IP Relocalizable (Relocate IP Address RIP)
La RIP es una dirección IP asociada a un paquete. Cuando éste paquete es movido a otro equipo dentro del cluster la RIP es también movida a ese otro equipo.

k) Start Script
Es un script que puede ser ejecutado cuando el paquete de aplicaciones es inicializado en un nodo.

l) Dirección IP estacionaria.
Es la Dirección asociada a la tarjeta LAN en un equipo.

m) TPS
Sistema de Teleproceso

Configuración de la Red .

El sistema consiste de 2 Equipos de Transferencia, 2 Equipos de Procesamiento, 1 Equipo de Almacenamiento y 2 Equipos de prueba, uno para Transferencia y otro para Procesamiento. Las X-Terminal también constituyen una parte fundamental para la configuración de la Red.

Los Equipos de Procesamiento, Equipos de Transferencia y de Almacenamiento están conectados al anillo FDDI primario y al anillo FDDI standby (de respaldo) por dos ligas FDDI. Los dos anillos FDDI están conectados por un Puente FDDI (FDDI Bridge). Esto garantiza que los sistemas puedan comunicarse si existe una falla en el Anillo Primario de FDDI. De modo similar, los Equipos de Transferencia y de Procesamiento están también conectados a la Ethernet LAN primaria y a la standby Ethernet LAN, y éstas a su vez están conectadas a través de un puente LAN. Las dos equipos de Test están conectados al FDDI Primario y a la Ethernet Primaria solamente. Las X-Terminal están conectadas a la Ethernet LAN primaria.

Los datos transferidos y los archivos transferidos entre los sistemas hacen uso del FDDI primario. La LAN primaria y FDDI LAN son usadas por la señal de Heartbeat intercambiada entre los equipos que estén dentro de la funcionalidad de Haig Availability y por las conexiones de las X-Terminal. La configuración de la red está mostrada en la figura A.1

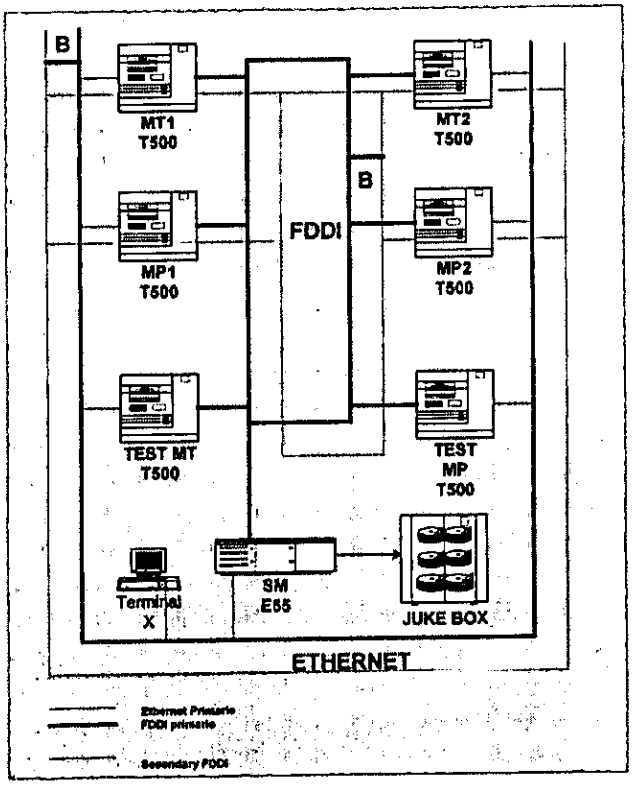


Figura A.1 TPS de México

La figura A.2 muestra los elementos que componen el TPS (Sistema de Teleproceso), en su conjunto:

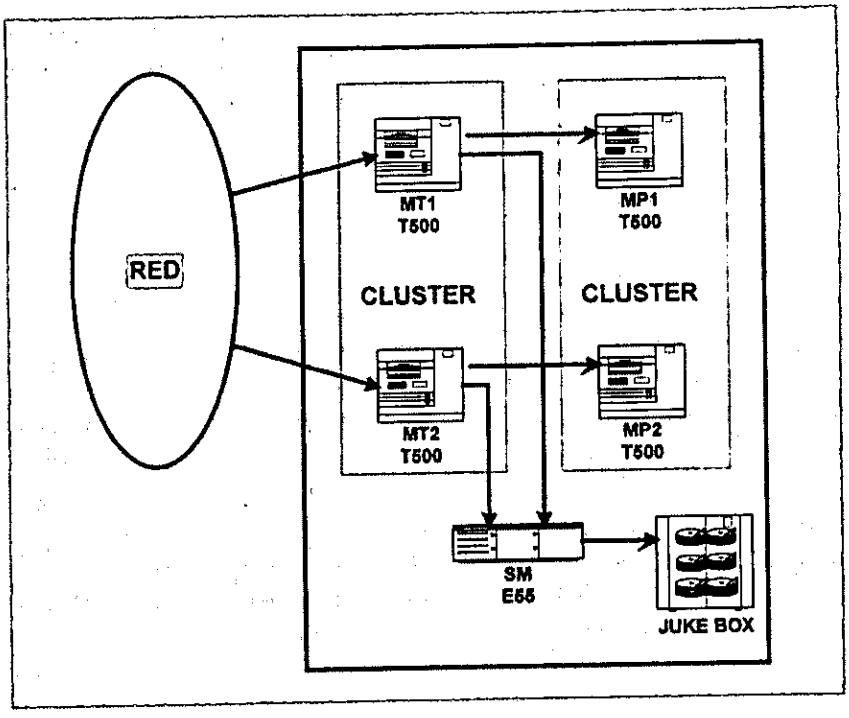


Figura A.2 TPS Operación Normal

Esta figura A.3 es un ejemplo de la configuración del sistema dentro de operación normal. El conjunto de discos (a,b) y (a', b') pueden ser accesados por ambos sistemas. Pero cada uno de los conjuntos es reservado y usado solo por uno de los sistemas en operación normal.

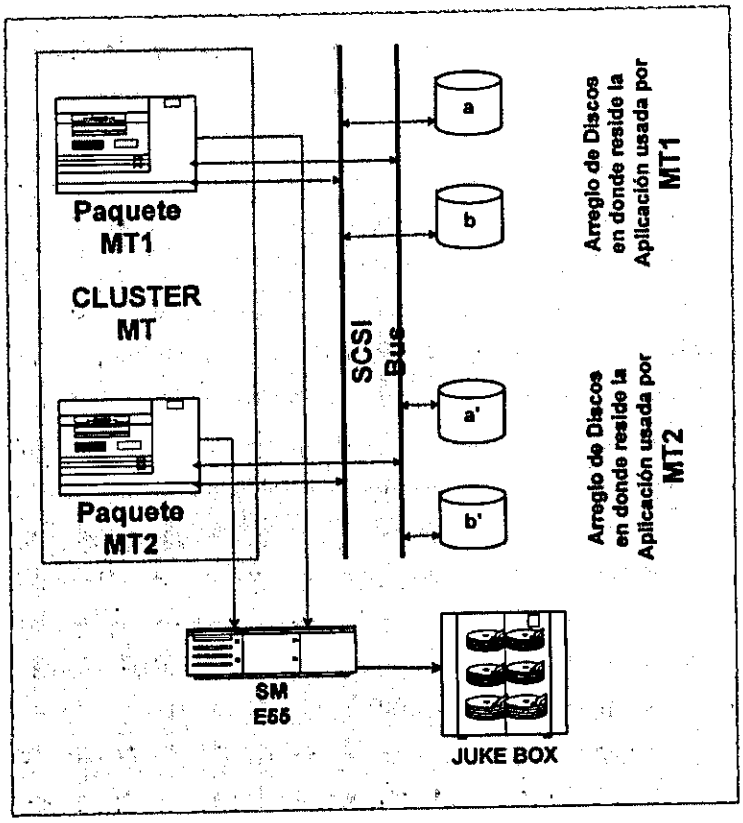


Figura A.3 Cluster en Operación Normal.

El Cluster definido para el Sistema y la Configuración del Paquete de Aplicaciones (Pakage).

La tabla A.1 muestra los Clusters del Sistema y sus configuraciones:

Nombre del Cluster	ID del Cluster	Nombre de los Nodos y (Hosts Estacionarios)
Machine Transfer (MT)	MT	mxmt1, mxmt2
Machine Process (MP)	MP	mxmp1, mxmp2

Tabla A.1 TPS Clusters

MC/ServiceGuard Software

El MC/ServiceGuard Software va a ser usado en el Sistema para proveer a éste de un Mecanismo de Alta Disponibilidad (High Availability) para los Equipos de Transferencia y Procesamiento, así como sus Aplicaciones de respaldo. Los equipos de Transferencia y Procesamiento deberán hacer uso de MC/ServiceGuard como protección contra fallas en los casos más críticos. El equipo de Almacenamiento no entrará en High Availability en caso de una falla. Los nodos en un Cluster estarán dando señales de su correcto funcionamiento mediante la señal de Heartbeat. La señal de Heartbeat será mandada sobre la Ethernet LAN y FDDI LAN a todos los nodos en el Cluster. Si en un determinado número de Heartbeats un nodo no detecta otro nodo, éste asume que tiene una falla. El Paquete de Aplicaciones, que está corriendo en el nodo que es detectado con fallas es transferido al nodo que lo detecta, convirtiéndose así en un Nodo Adoptivo.

En nuestro caso, si el equipo de MT1 no está siendo visto por MT2 entonces MT2 puede tomar el Paquete de aplicaciones de MT1 puede ser ejecutado por MT2.

El Software MC/ServiceGuard, monitorea ciertos aspectos del Sistema en un Cluster. En el caso de tener un problema, el sistema no queda fuera de servicio. El MC/ServiceGuard Software soporta dos tipos de Fallas:

a) Fallas Locales

El MC/ServiceGuard Software detecta fallas locales en uno de los equipos del Cluster.

b) Fallas Remotas

Las aplicaciones son ejecutadas en un Nodo adoptivo cuando el Nodo Primario falla.

NOTA: Existen ciertas aplicaciones que no monitorea MC/ServiceGuard (Aplicaciones no críticas). El monitoreo de las fallas en las aplicaciones quedará en responsabilidad de Operación y esta situación va a ser descrita ampliamente dentro de la Sección de Responsabilidades del Administrador.

Migración en Caso de un Desastre en el SITE.

Esta situación está contemplada para la Fase 2 de este proyecto, próxima en 2000, incluyendo un espejo del proyecto en Monterrey para solventar situaciones de desastre en cualquiera de los dos SITES. Siempre que exista una falla (Fileover) que no pueda ser detectada por MC/ServiceGuard se le llama SiteFailover. Un SiteFailover puede suscitarse en caso de un catastrófico desastre natural, por citar un ejemplo. En este caso, ambos equipos en el Cluster, incluyendo el MC/ServiceGuard Software, fallaran por completo.

Responsabilidades del Administrador del Sistema.

El Administrador del Sietema deberá estar entrenado para estar a la expectativa de cualquier problema que deba ser atendido por MC/ServiceGuard, por lo tanto, se debe asumir que debe estar completamente familiarizado con el funcionamiento de MC/ServiceGuard.

Los siguientes puntos es un resumen de las responsabilidades que el administrador debe de asumir. Las acciones que éste debe de tomar van a ser discutidas con más detalle en las siguientes secciones:

1. Levantamiento y baja total del Sistema.

2. Monitoreo del Status del Sistema
a) Monitoreo de la Consola de Mensages
b) Monitoreo de los Archivos de Registro de Operación (Logs Fails)
c) Monitoreo del Status de los Clusters.

3. Recuperación del Sistema
a) Durante una Falla Local
b) Durante una Falla Remota
c) En un futuro, durante un Site Failover

4. Mantenimiento del Sistema

Startup y Shutdown del Sistema

Reboot del Sistema

Los equipos de Transferencia (MT1, MT2), Procesamiento (MP1, MP2) y Almacenamiento (MA) pueden ser "booteadas" después de cargar nuevo software requerido. La referencia está en el tema, mantenimiento, a como debe ser cargado el software adicional en los equipos y como se debe "bootear" el equipo.

Startup Clusters

Es recomendado que se lleve a cabo la secuencia de pasos para levantar los sistemas. Debe primero ejecutarse un normal startup en cada uno de los nodos del sistema y posteriormente levantar todos los paquetes que tengan habilitado el switcheo de MC/ServiceGuard.

Startup de los Clusters MP (Inicialización de los Clusters MP y sus paquetes)

Para poner en marcha los Clusters MP, se deben llevar a cabo los siguientes pasos después que el equipo ha sido puesto en operación:

1. Conectarse a alguno de los nodos del Cluster (MP1 ó MP2) como el usuario "root".
2. Ejecutar el comando "cmruncl"
   # cmruncl

Cuando este comando es ejecutado el shell script "control.sh", localizado en el directorio /etc/cmcluster/MP?, es automáticamente ejecutado por MC/ServiceGuard.

El shell script "control.sh" realiza las siguientes operaciones:

a) Activa los grupos de volúmenes, volúmenes lógicos
b) Monta los file system
c) Asigna la dirección RIP a los paquetes
d) Levanta las aplicaciones de cada uno de los paquetes

La siguiente tabla A.2 muestra los grupos de volúmenes y los volúmenes lógicos que deben ser activados, los respectivos file system que deben ser montados, las direcciones RIP que deben ser asignadas y los paquetes de aplicaciones que deben ser levantados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla A.2 Inicialización de los Clusters MP

Startup de los Clusters MT
(Inicialización de los Clusters MT y sus paquetes)

Para poner en marcha los Clusters MT, se deben llevar a cabo los siguientes pasos después que el equipo ha sido puesto en operación:

1. Conectarse a alguno de los nodos del Cluster (MT1 ó MT2) como el usuario "root".
2. Ejecutar el comando "cmruncl"
   # cmruncl

Cuando este comando es ejecutado el shell script "control.sh", localizado en el directorio /etc/cmcluster/MT?, es automáticamente ejecutado por MC/ServiceGuard.

El shell script "control.sh" realiza las siguientes operaciones:

a) Activa los grupos de volúmenes, volúmenes lógicos
b) Monta los file system
c) Asigna la dirección RIP a los paquetes
d) Levanta las aplicaciones de cada uno de los paquetes

La siguiente tabla A.3 muestra los grupos de volúmenes y los volúmenes lógicos que deben ser activados, los respectivos file system que deben ser montados, las direcciones RIP que deben ser asignadas y los paquetes de aplicaciones que deben ser levantados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla A.3 Inicialización de los Clusters MT

Startup del equipo de almacenamiento MA (Levantamiento del equipo MA y sus paquetes)

Este equipo no cuenta con el software de MC/ServiceGuard, ya que no está dentro de la alta disponibilidad requerida para el proyecto.

Cómo se deben verificar los procedimientos de Startup

1. Verificar el status del Cluster

Para verificar el status del Cluster, el Administrador del equipo debe ejecutar el comando cmviewcl y realizando lo siguiente:

a) Conectarse a un nodo del Cluster como el usuario "root".
b) Ejecutar el comando
/usr/sbin/cmviewcl

La siguiente pantalla será desplegado en pantalla:

```
CLUSTER       STATUS

MT_12         up

  NODE          STATUS       STATE
  mt1s          up           running

    PACKAGE       STATUS       STATE        PKG_SWITCH    NODE
    MT1           up           running      enabled       mt1s

  NODE          STATUS       STATE
  Mt2s          up           running

    PACKAGE       STATUS       STATE        PKG_SWITCH    NODE
    MT2           up           running      enabled       mt2s
```

Aquí el Administrador debe asegurarse de que todos los paquetes tienen el status de levantados y corriendo ("up" and "running").

"PKG_SWITCH" debe estar habilitado ("enabled") y el paquete debe estar ejecutando en su nodo original.

Para ver de manera más detallada la información, se puede ejecutar el siguiente comando:

/usr/sbin/cmviewcl –v

La siguiente salida será desplegada en el monitor del Operador:

```
CLUSTER       STATUS

MT_12         up


  NODE          STATUS       STATE

  mt1s          up           running

    Network_Parameters:
    INTERFACE     STATUS       PATH         NAME
    PRIMARY       up           0/20.1       lan0
    STANDBY       up           0/44.1       lan1
    PRIMARY       up           2/4          lan2
    STANDBY       up           2/12         lan3

    PACKAGE       STATUS       STATE        PKG_SWITCH    NODE
    Mt1           up           running      enabled       mt1s

      Script_Parameters:
      ITEM          STATUS       NAME          MAX_RESTARTS  RESTARTS
      Service       up           ORACLE_MT1    0             0
      Service       up           TPS_MT1       2             0
      Service       up           SQLNET_MT1    4             0
      Service       up           DAR_MT1       4             0
      Service       up           ALARM_MT1     8             1
      Service       up           CD_MT1        4             0
      Subnet        up           13.49.152.0
      Subnet        up           13.50.70.0

      Node_Switching_Parameters:
      NODE_TYPE     STATUS       SWITCHING     NAME
      Primary       up           enabled       mt1s          (current)
      Alternate     up           enabled       mt2s

  NODE          STATUS       STATE
  mt2s          up           running

    Network_Parameters:
    INTERFACE     STATUS       PATH         NAME
    PRIMARY       up           0/20.1       lan0
    STANDBY       up           0/44.1       lan1
    PRIMARY       up           2/4          lan2
    STANDBY       up           2/12         lan3

    PACKAGE       STATUS       STATE        PKG_SWITCH    NODE
    MT2           up           running      enabled       mt2s

      Script_Parameters:
```

```
ITEM        STATUS      NAME          MAX_RESTARTS  RESTARTS
Service     up          ORACLE_MT2           0          0
Service     up          TPS_MT2              2          0
Service     up          SQLNET_MT2           4          0
Service     up          DAR_MT2              4          0
Service     up          ALARM_MT2            8          1
Service     up          CD_MT2               4          0
Subnet      up          13.49.152.0
Subnet      up          13.50.70.0

Node_Switching_Parameters:
NODE_TYPE   STATUS      SWITCHING     NAME
Primary     up          enabled       mt2s        (current)
Alternate   up          enabled       mt1s
```

Aquí debemos asegurarnos de que todas las interfaces de Red, servicios, Subredes están habilitadas ("up").

El Nodo Primario "Primary" y el Alternativo "Alternate" deben estar levantados "up" y además habilitados "enabled" para "SWITCHING".

2. Monitoreo de las bitácoras de sistema (Log Files)

Para verificar la bitácora del equipo se debe hacer lo siguiente:

a) Conectarse a uno de los nodos del Cluster como usuario "root"
b) Ejecutar el siguiente comando:
tail –30 /var/adm/syslog/syslog.log | more

Un desplegado similar al que a continuación se presenta, se desplegará en pantalla:

```
Sep  3  14:30:23 hpsgnoh cmcld[1992] : Started package MT1 on
Node hpsgnoh

~
Sep  3  14:30:23 hpsgnoh cmcld[1992] : Started package MT1 on
Node hpsgnoh
```

Si los mensajes, como los anteriores, no son desplegados, esto significa que algún error ha ocurrido. Para su solución se puede consultar la sección de Fallas en el Sistema y procedimiento de Recuperación de este manual.

Se pueden revisar los siguientes archivos e log para obtener información adicional, si el resultado del "tail" no es satisfactorio:

/var/adm/cmcluster/HA_MT?.log o
/var/adm/cmcluster/HA_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log representa la salida del resultado de la corrida del shell script "control.sh".

/var/adm/cmcluster/HA_MT1.log conserva los mensajes enviados durante el arranque del Paquete MT1.

/var/adm/cmcluster/MT?/control_MT?.sh.log o
/var/adm/cmcluster/MP?/control_MP?.sh.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene algunos otros mensajes relacionados con la corrida del script de control.

/var/adm/cmcluster/services_MT?.log o
/var/adm/cmcluster/services_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene la Salida de la Aplicación de Proceso y del monitoreo de los servicios de Oracle.

3. Verificar que los File system estén montados

Asegurarse de que los file system listados en las tablas A.4 y A.5, estén correctamente montados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla A.4 Inicialización de los Clusters MT

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla A.5 Inicialización de los Clusters MP

4. Revisar los Procesos que deben levantar cada una de las aplicaciones.

Las tabla A.6 lista los procesos en el Sistema que deben estar corriendo por aplicación, involucradas en los paquetes.

Software o Aplicación	Procesos	Equipo
Oracle	Ora_pmon_<ORACLE_SID> Ora_dbwr_<ORACLE_SID> Ora_arch_<ORACLE_SID> Ora_lgwr_<ORACLE_SID> Ora_ckpt_<ORACLE_SID> Ora_smon_<ORACLE_SID>	MT, MP
Aplicación de Transferencia Y procesamiento (TPS)	Manmbin Monsbin Mistsbin Manfsbin	MT, MP
CONNECT:Direct	Cdpmgr	MT, MP
SQLNET	Tnslsnr	MT, MP
Darput	Darput	MT
Omnistorege	Ded Qr	MT
Omniback	Crs Lockmgr	MT, MP
Alarm daemon	Alarm.sh	MT, MP

Tabla A.6 Lista de Procesos

5. Si los paquetes no pueden ser levantados en su nodo original, MC/ServiceGuard puede tratar de levantarlos en el nodo adoptivo.

Si los pasos 1 a 4 no son realizados satisfactoriamente, se debe investigar la causa del problema y tratar de resolverlo. Cuantas veces sea necesario, se deben ejecutar los pasos de inicialización hasta que el Cluster quede trabajando correctamente.

Shutdown de los Clusters

Es recomendado que se lleve a cabo la secuencia de pasos para dar de baja los Clusters. Antes de ejecutarse un normal shutdown en cada uno de los nodos del sistema se deben de detener todos los paquetes que estén habilitados ejecutando los siguientes pasos:

Shutdown de los Clusters MP
(Paro de los Clusters MP y sus paquetes)

Para detener los Clusters MP, se deben llevar a cabo los siguientes pasos antes de que el equipo sea detenido:

1. Conectarse a alguno de los nodos del Cluster (MP1 ó MP2) como el usuario "root".
2. Ejecutar el comando "cmhaltcl"
   # cmhaltcl –f

Cuando este comando es ejecutado el shell script "control.sh", localizado en el directorio /etc/cmcluster/MP?, es automáticamente ejecutado por MC/ServiceGuard.

El shell script "control.sh" realiza las siguientes operaciones:

a) Detiene todos los paquetes y las aplicaciones incluidas en cada paquete
b) Remueve la dirección RIP a los paquetes
c) Desmonta los file system
d) Desactiva los grupos de volúmenes
e) Detiene el Cluster

La tabla A.7 , file system MP, muestra los grupos de volúmenes y los volúmenes lógicos activados, los respectivos file system montados, las direcciones RIP asignadas y los paquetes de aplicaciones levantados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla A.7 File system MP

Shutdown de los Clusters MT
(Paro de los Clusters MT y sus paquetes)

Para detener los Clusters MT, se deben llevar a cabo los siguientes pasos antes de que el equipo sea detenido:

1. Conectarse a alguno de los nodos del Cluster (MP1 ó MP2) como el usuario "root".
2. Ejecutar el comando "cmhaltcl"

cmruncl –f

Cuando este comando es ejecutado el shell script "control.sh", localizado en el directorio /etc/cmcluster/MT?, es automáticamente ejecutado por MC/ServiceGuard.

El shell script "control.sh" realiza las siguientes operaciones:

a) Detiene todos los paquetes y las aplicaciones incluidas en cada paquete
b) Remueve la dirección RIP a los paquetes
c) Desmonta los file system
d) Desactiva los grupos de volúmenes
e) Detiene el Cluster

Lasa tabla A.8 , file system MT, muestra los grupos de volúmenes y los volúmenes lógicos activados, los respectivos file system montados, las direcciones RIP asignadas y los paquetes de aplicaciones levantados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla A.8 File system MT

Shutdown del equipo de almacenamiento MA
(Halt del equipo MA y sus paquetes)

Este equipo no cuenta con el software de MC/ServiceGuard, ya que no está dentro de la alta disponibilidad requerida para el proyecto.

Cómo se deben verificar los procedimientos de Shutdown

1. Verificar el status del Cluster

Para verificar el status del Cluster, el Administrador del equipo debe ejecutar el comando cmviewcl y realizando lo siguiente:

a) Conectarse a un nodo del Cluster como el usuario "root".
b) Ejecutar el comando
/usr/sbin/cmviewcl

El siguiente mensaje de error será desplegado en pantalla:

```
CLUSTER        STATUS
"
               down
"
```

2. Monitoreo de las bitácoras de sistema (Log Files)

Para verificar la bitácora del equipo se debe hacer lo siguiente:

a) Conectarse a uno de los nodos del Cluster como usuario "root"
b) Ejecutar el siguiente comando:
tail –30 /var/adm/syslog/syslog.log | more

Un desplegado similar al que a continuiación se presenta, se desplegará en pantalla:

```
Sep  4  15:25:03 hpsgnoh cmcld[1992] : Halted package MT1 on
Node hpsgnoh

"
Sep  4  15:25:03 hpsgnoh cmcld[1992] : Halted package MT1 on
Node hpsgnoh
```

Se pueden revisar los siguientes archivos e log para obtener información adicional, si el resultado del "tail" no es satisfactorio:

/var/adm/cmcluster/HA_MT?.log o
/var/adm/cmcluster/HA_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log representa la salida del resultado de la corrida del shell script "control.sh".

/var/adm/cmcluster/HA_PT1.log conserva los mensajes enviados durante el arranque del Paquete PT1.

/var/adm/cmcluster/MT?/control_MT?.sh.log o
/var/adm/cmcluster/MP?/control_MP?.sh.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene algunos otros mensajes relacionados con la corrida del script de control.

/var/adm/cmcluster/services_MT?.log o
/var/adm/cmcluster/services_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene la Salida de la Aplicación de Proceso y del monitoreo de los servicios de Oracle.

3. Verificar que los File system estén desmontados

Asegurarse de que los file system listados en las tablas A.9 y A.10, estén correctamente desmontados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	Mxmt2

Tabla A.9 Inicialización de los Clusters MT

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla A.10 Inicialización de los Clusters MP

4. Revisar los Procesos que deben ser detenidos por cada una de las aplicaciones.

La tabla A.11 lista los procesos en el Sistema que deben estar detenidos por aplicación, involucardas en los paquetes.

Software o Aplicación	Procesos	Equipo
Oracle	Ora_pmon_<ORACLE_SID> Ora_dbwr_<ORACLE_SID> Ora_arch_<ORACLE_SID> Ora_lgwr_<ORACLE_SID> Ora_ckpt_<ORACLE_SID> Ora_smon_<ORACLE_SID>	MT, MP
Aplicación de Transferencia Y procesamiento (TPS)	Manmbin Monsbin Mistsbin Manfsbin	MT, MP
CONNECT:Direct	Cdpmgr	MT, MP
SQLNET	Tnslsnr	MT, MP
Darput	Darput	MT
Omnistorege	Ded Qr	MT
Omniback	Crs Lockmgr	MT, MP
Alarm daemon	Alarm.sh	MT, MP

Tabla A.11 Lista de Procesos

3. Fallas del Sistema y Procedimiento de Recuperación

Existen tres tipos diferentes de fallas en las cuales se requiere la intervención directa del Administrador de los equipos:

a) Local Failover/Failback

b) Remote Failover/Failback

c) Site Disaster/Recovery

Local Failover (Tarjeta de red dañada)

Cuando MC/ServiceGuard detecta que las tarjetas de Red han fallado puede recuperarse de este failover haciendo un failback con la tarjeta de respaldo que se encuentra en cada uno de los equipos del Cluster.

Todas las aplicaciones continúan corriendo en la misma máquina después de la falla en la tarjeta de Red.

MC/ServiceGuard automáticamente, cuando detecta una falla en una tarjeta, que puede ser la misma tarjeta o el cableado, reconfigura la tarjeta de respaldo (standby LAN card) asignandole la dirección de RIP.

Las siguientes figuras A.4 y A.5 muestran el status del Cluster antes y despúes de un local failover.

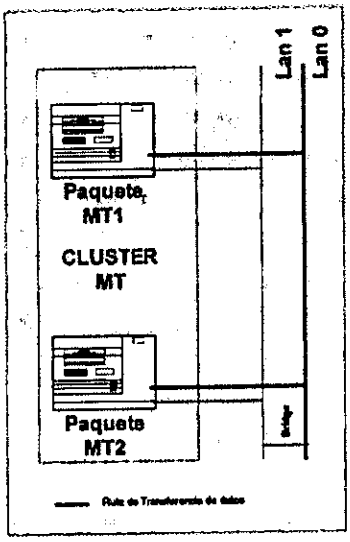


Figura A.4 Cluster Antes de un Failover

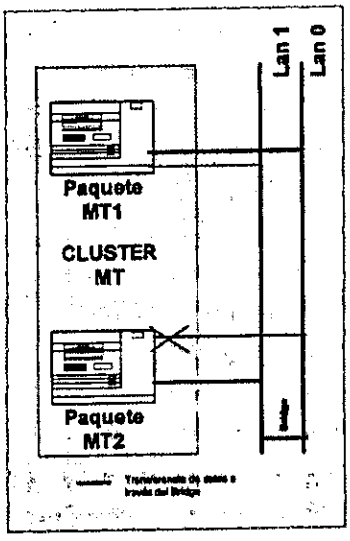


Figura A.5 Cluster Después de un Failover

El administrador puede hacer una revisión para ver cuales son las condiciones del Failover.

Mensajes en Consola

El administrador puede revisar los mensajes en la consola. El siguiente mensaje es desplegado cuando la tarjeta de Red o el cableado falla:

```
Network NS_LS_DRIVER Disaster 1029, Pid [ICS]
  LAN card on interface unit 0 has network problem.
  Check cable for possible disconnection.
```

Revisión de los Mensajes en los Archivos de Log del Sistema

El archivo de log localizado en:

/var/adm/syslog/syslog.log

puede contener mensajes relevantes cuando la falla local remota ocurren. Si tratamos de conectarnos de conectarnos de este equipo a si mismo mediante un telnet o un login, como root, el mensaje que obtendremos será el siguiente:

Revisar el syslog.log haciendo lo siguiente:

a) Conectarse a uno de los Clusters como el usuario "root"
b) Ejecutar la siguiente operación:

tail –50 /var/adm/syslog/syslog.log | more

El mensage para la falla local en el syslog.log será:

```
Sep  4  11:26:03 hpsgnoh cmcld[4932] : lan0 failed
~
Sep  4  11:26:03 hpsgnoh cmcld[4932] : lan0 switched with lan1
Sep  4  11:26:05 hpsgnoh cmcld[4932] : Local switch has ocurred
```

El mensaje encontrado en caso de que la LAN en Standby fallara sería:

```
Sep  4  11:26:06 hpsgnoh cmcld[4932] : lan1 failed
```

Status del Cluster

El administrador debe estar monitoreando regularmente el status del Cluster.

Para revisar el status del Cluster, el administrador debe ejecutar el comando cmviewcl haciendo lo siguiente:

a) Conectarse a uno de los nodos del Cluster como "root"
b) Ejecutar la siguiente operación:

#/usr/sbin/cmviewcl –v

La siguiente salida va a ser desplegada en el monitor:

```
CLUSTER        STATUS
~
MT_12          up
~
~
  NODE           STATUS         STATE
~
  mt1s           up             running

    Network_Parameters:
    INTERFACE     STATUS         PATH           NAME
    PRIMARY       down             0/20.1         lan0
    STANDBY       up             0/44.1         lan1
    PRIMARY       up             2/4            lan2
    STANDBY       up             2/12           lan3

    PACKAGE       STATUS         STATE          PKG_SWITCH    NODE
    Mt1           up             running        enabled       mt1s

      Script_Parameters:
      ITEM          STATUS         NAME          MAX_RESTARTS  RESTARTS
      Service       up             ORACLE_MT1              0         0
      Service       up             TPS_MT1                 2         0
      Service       up             SQLNET_MT1              4         0
      Service       up             DAR_MT1                 4         0
      Service       up             ALARM_MT1               8         1
      Service       up             CD_MT1                  4         0
      Subnet        up             13.49.152.0
      Subnet        up             13.50.70.0

      Node_Switching_Parameters:
      NODE_TYPE     STATUS         SWITCHING      NAME
      Primary       up             enabled        mt1s          (current)
      Alternate     up             enabled        mt2s

  NODE           STATUS         STATE
  mt2s           up             running

    Network_Parameters:
```

```
INTERFACE     STATUS       PATH         NAME
PRIMARY       up           0/20.1       lan0
STANDBY       up           0/44.1       lan1
PRIMARY       up           2/4          lan2
STANDBY       up           2/12         lan3

PACKAGE       STATUS       STATE        PKG_SWITCH    NODE
MT2           up           running      enabled       mt2s

  Script_Parameters:
  ITEM          STATUS       NAME              MAX_RESTARTS  RESTARTS
  Service       up           ORACLE_MT2        0             0
  Service       up           TPS_MT2           2             0
  Service       up           SQLNET_MT2        4             0
  Service       up           DAR_MT2           4             0
  Service       up           ALARM_MT2         8             1
  Service       up           CD_MT2            4             0
  Subnet        up           13.49.152.0
  Subnet        up           13.50.70.0

  Node_Switching_Parameters:
  NODE_TYPE     STATUS       SWITCHING    NAME
  Primary       up           enabled      mt2s          (current)
  Alternate     up           enabled      mt1s
```

Si encontramos que la interface de Red primaria está abajo (status – "down"), eso significa que una falla está ocurriendo. El administrador debe revisar si ésto es el resultado de un problema de conexión o de hardware.

Tiempo de Recuperación

El tiempo que se lleva en detectar y recuperar una falla local (Local Failover) es de aproximadamente 10 segundos. La aplicación no es reiniciada por a causa de esta falla. La aplicación puede continuar trabajando haciendo uso de la tarjeta de respaldo. La dirección RIP está en ese momento asociada a la LAN de respaldo.

LOCAL FAILBACK (Recuperación de la Comunicación en lan0)

Para recuperar el sistema de un Local failover, el administrador debe seguir los siguientes pasos:

1. Revisar el Log File

Revisar el syslog.log haciendo lo siguiente:

a) Conectarse a uno de los Clusters como el usuario "root"
b) Ejecutar la siguiente operación:

```
# tail –50 /var/adm/syslog/syslog.log | more
```

El mensaje para la falla local en el syslog.log será:

```
Sep  4  11:26:03 hpsgnoh cmcld[4932] : lan0 failed
~
Sep  4  11:26:03 hpsgnoh cmcld[4932] : lan0 switched with lan1
Sep  4  11:26:04 hpsgnoh cmcld[4932] : Local switch has ocurred
```

El mensaje encontrado en caso de que la LAN en Standby fallara sería:

```
Sep  4  11:26:05 hpsgnoh cmcld[4932] : lan1 failed
~
```

2. El siguiente comando nos puede ayudar a localizar fallas en la Red.

lanscan

Una salida similar a la siguiente es desplegada en pantalla con el uso de este comando:

```
Hardware Station         Crd Hardware Net-Interface    NM MAC   HP DLPI Mjr
~
Path     Address         In# State    NameUnit State   ID Type  Support Num
~
0/20.1   0x080009B7612B 0   UP       lan0     UP      4  ETHER Yes     185
0/44.1   0x080009D01CBE 1   UP       lan1     DOWN    5  ETHER Yes     185
2/4      0x080009C44408 2   UP       lan2     DOWN    6  FDDI  Yes     191
2/12     0x080009C4D4EB 3   UP       lan3     DOWN    7  FDDI  Yes     191
```

La falla en la tarjeta da como resultado un "Hardware State" "DOWN".

3. Si la conexión del cable es el problema, podemos conectar otro cable y el failback es hecho por MC/ServiceGuard automáticamente.

4. Si la tarjeta de LAN tiene problemas, hay que remplazarla y realizar los siguientes pasos:

a) Forzar un failover remoto del equipo.

b) Dar un shutdown al equipo y ponerlo en modo mantenimiento.

c) Remplazar la tarjeta de Red.

d) Levantar nuevamente el sistema.

e) Hacer un failback remoto.

Remote Failover (Un Nodo del Cluster Falla)

Un failover remoto sucede dentro de las siguientes dos condiciones:

a) Cuando el equipo tiene problemas de Harware o se rebootea por completo.

b) Cuando una aplicación falla

La figura A.6 muestra el escenario en caso de que MT2 en fallara.

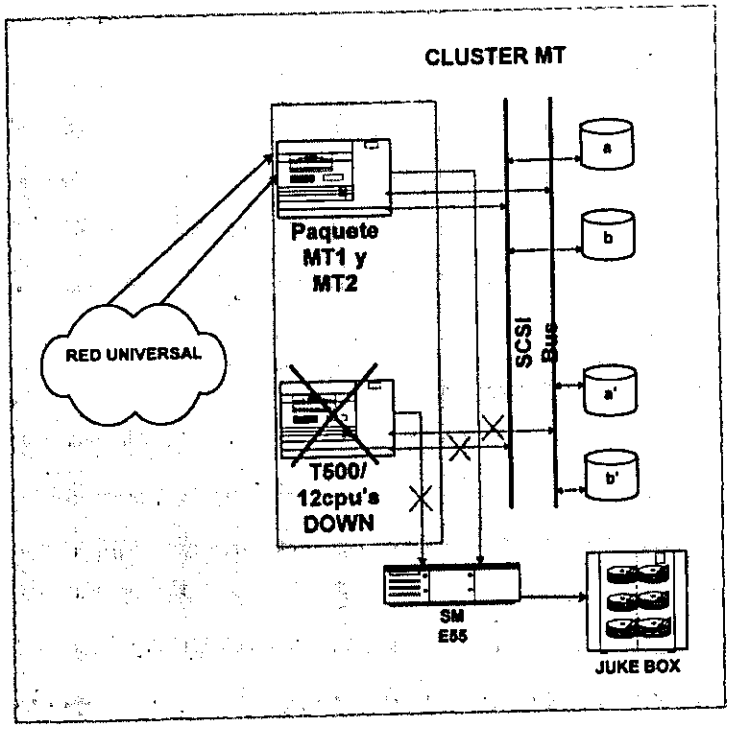


Figura A.6 MT2 Falla en el Cluster

La figura A.7 es un ejemplo de la configuración del sistema cuando está en una situación de Failover Remoto. El arreglo de discos (a,b) y (a',b') son accesados por ambos equipos dentro del Cluster. Pero ambos arreglos son bloqueados y usados por solo el Equipo adoptivo, que en este caso sería la MT1.

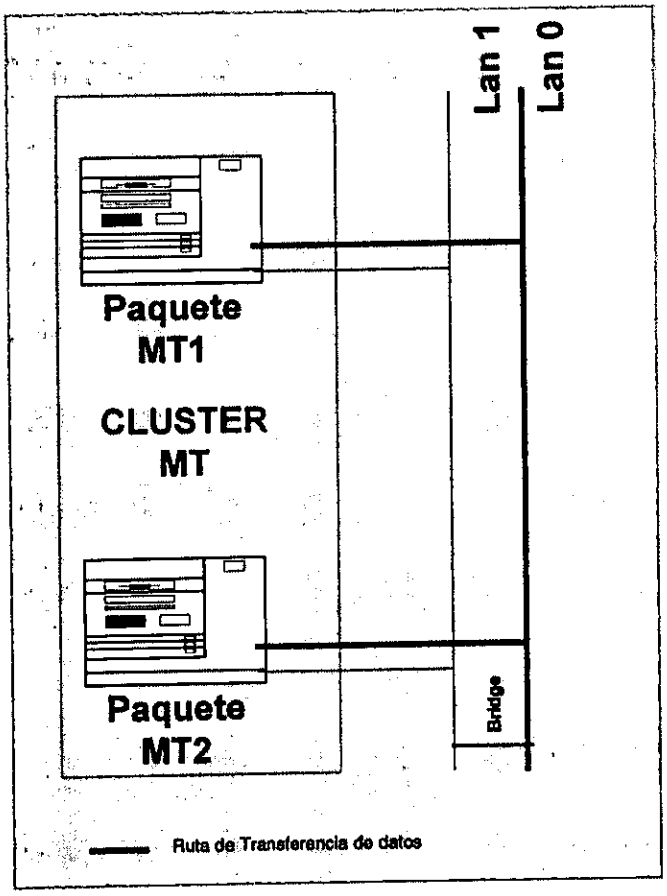


Figura A.7 Después de un Remote Failover

Una vez que MC/ServiceGuard detecta una falla de un nodo, éste ejecuta la operación "remote failover".

Existen dos operaciones principales durante el "remote failover":

a) Detenga los paquetes en el nodo que está fallando si están corriendo.

b) Levanta los paquetes en el nodo adoptivo.

Las siguientes operaciones son ejecutadas por el MC/ServiceGuard por el script de control. El script de control es el mismo script para detener o levantar los paquetes del Cluster. El Script de control está localizado en la siguiente ruta:

"/etc/cmcluster/MT? ó MT?/control.sh.

El script de control es capaz de ejecutar las siguientes operaciones:

a) Detener los paquetes en el nodo en falla si los paquetes están corriendo:

b) Detiene las aplicaciones haciendo un shutdown normal.

c) Desmonta los file system compartidos y desactiva los grupos de volúmenes.

d) Reconfigura las tarjetas de Red y mueve los paquetes hacia la dirección RIP.

e) Levantar los paquetes en el nodo adoptivo si los paquetes están corriendo:

f) Reconfigura las tarjetas de Red con las dirección de RIP.

g) Activa los grupos de volúmenes compartidos y monta los file system.

h) Revisa el status de las aplicaciones, y hace una recuperación de la Base de Datos si es necesario.

i) Reconfigura las aplicaciones en caso de ser necesario.

j) Levanta las aplicaciones del paquete, haciendo un startup normal.

Aplicaciones con falla

Nombre del Servicio	Service Monitor/ Restart Script	Intervalo de Tiempo Para Revisar Proceso (segs)	Intervalo de Tiempo para Levantar el Servicio (segs)	Número Máximo de Intentos de Recuperación del Servicio
ORACLE_MT1	/etc/cmcluster/monitor/ora.mon	5	0	0
TPS_MT1	/etc/cmcluster/monitor/tps.mon	5	450	2
OMNI_MT1	/etc/cmcluster/monitor/omni.mon	20	3600	4
OMNISTORE_MT1	/etc/cmcluster/monitor/omnistore.mon	20	3600	4
CD_MT1	/etc/cmcluster/monitor/cd.mon	20	1800	4
SQLNET_MT1	/etc/cmcluster/monitor/sqlnet.mon	20	1800	4
DAR_MT1	/etc/cmcluster/monitor/dar.mon	20	3600	4
ALARM_MT1	/etc/cmcluster/monitor/alarm.mon	5	900	8

Tabla A.12Configuración del Paquete (Servicios Críticos)

Configuración de los paquetes – Servicios o Aplicaciones Críticas

Todas las aplicaciones mostradas en la tabla A.12 deben ser monitoreadas por el correspondiente script de monitoreo. El script revisa los procesos de las aplicaciones, la frecuencia para hacer la revisión está determinada por el intervalo de revisión de procesos (Process Checking Interval). Si uno de los procesos de alguna de las aplicaciones está abajo, MC/ServicGuard trata de restablecer la aplicación. Si la aplicación no puede ser restablecida satisfactoriamente, ésta espera un cierto tiempo (Restart Interval) mientras trata de hacer otro restart. Si el número de restart time sobrepasa su valor máximo configurado, MC/Service guard ejecuta un Remote failover.

El script de monitoreo es inicializado y continuamente revisado por el demonio de MC/ServiceGuard. Cada vez que el script de monitoreo detecta falla en las aplicaciones, éste hace, después de tratar de levantar la aplicación, un "exit" con código "1". Cada vez que MC/ServiceGuard detecta que el script manda una señal con código "1" éste incrementa el contador de restart times en 1.

Si una de los procesos de la Aplicación Oracle, falla esta aplicación no puede ser levantada. Lo anterior sucede ya que si los procesos son matados de manera anormal, la memoria compartida, no puede ser liberada. Si se restablece Oracle, la memoria compertida no puede ser usada por completo, y si hay una falla en otro nodo, este nodo no va a estar disponible para tomar otro paquete de otro nodo debido a la falta de la memoria compartida. En este sentido, es necesario "rebootear" el equipo en caso de que Oracle falle.

Si el script de monitoreo detecta falla en los procesos de las aplicaciones configuradas como críticas, se puede visualizar los siguientes mensajes en la consola del equipo:

```
ALARM:      Sep   1    13:02:08    -    SG node "hpsgnog": Omniback daemon
failure detected!

INFORMATIVE:      Sep  1      13:04:10   -   SG node "hpsgnog": Restarting
up Omniback.
```

En caso de que suceda un startup de las aplicaciones, se observaría lo siguiente:

```
INFORMATIVE:      Sep  1      13:10:23   -   SG node "hpsgnog": The Omniback
Server started.
```

Si el startup de las aplicaciones fallara:

```
ERROR:      Sep  1      13:10:23    -    SG node \"$(hostname)": Startup of
Omniback daemon failed!.
```

El administrador del sistema puede tratar de resolver el problema en caso de que una falla ocurra. Puede apoyarse en la revisión de los siguientes archivos de log para buscar la causa del problema:

/var/adm/cmcluster/HA_MT?.log ó HA_MP?.log

/var/adm/cmcluster/service_MT?.log ó service_MP?.log

Revisión de un Remote Failover

La revisión de un remote failover puede realizarse en el nodo adoptivo. Por ejemplo si el equipo MT1 fallara, la verificación puede hacerse en el equipo MT2. De manera similar, si el equipo MP1 falla, la revisión puede hacerse en el la maquina MP2.

1. Revisión de los mensajes en consola

Los siguientes mensajes pueden ser desplegados en la consola adoptiva, cuando un remote failover ocurre:

```
ALARM: REMOTE FAILOVER: Package MT2 is switched over to MC/SG node
"hpsgnog" at Sep 1 14:50:45 SGP 1998.

INFORMATIVE:   MC/SG node "hpsgnog" : Starting package MT2 at Sep  1
14:50:45  SGP  1998.
```

2. Verificar el status del Cluster

Para verificar el status del Cluster, el Administrador del equipo debe ejecutar el comando cmviewcl y realizando lo siguiente:

a) Conectarse a un nodo del Cluster como el usuario "root".
b) Ejecutar el comando
c) # /usr/sbin/cmviewcl

El siguiente mensaje de error será desplegado en pantalla:

```
CLUSTER        STATUS

MT_12          up


  NODE           STATUS         STATE

  mt1s           up             running


    PACKAGE        STATUS         STATE          PKG_SWITCH    NODE
    MT1            up             running        enabled       mt1s
    MT2            up             running        disabled      mt1s


  NODE           STATUS         STATE

  mt2s           down           halted
```

3. Monitoreo de las bitácoras de sistema (Log Files)

Para verificar la bitácora del equipo se debe hacer lo siguiente:

a) Conectarse a uno de los nodos del Cluster como usuario "root"
b) Ejecutar el siguiente comando:
tail –30 /var/adm/syslog/syslog.log | more

Un desplegado similar al que a continuación se presenta, se desplegará en pantalla, realizando los pasos anteriores en el nodo sobreviviente:

```
Sep 13 08:32:22 mt1s cmcld[17494]: (mccm1s) Halted package MCCM1 on node
mccm1s.
Sep 13 08:35:33 mt1s CM-CMD[20903]: cmhaltpkg MCCM2
Sep 13 08:35:33 mt1s cmcld[17494]: Executing
'/etc/cmcluster/MT2/control.sh  stop' for package MT2.
Sep 13 08:35:35 mt1s CM-MCCM2[20920]: cmhaltserv ORACLE_MCCM2
Sep 13 08:35:38 mt1s CM-MCCM2[20940]: cmhaltserv TPS_MCCM2
Sep 13 08:35:40 mt1s CM-MCCM2[20961]: cmhaltserv ALARM_MCCM2
Sep 13 08:35:43 mt1s CM-MCCM2[20976]: cmhaltserv SQLNET_MCCM2
Sep 13 08:35:46 mt1s CM-MCCM2[21012]: cmhaltserv CD_MCCM2
Sep 13 08:35:48 mt1s CM-MCCM2[21021]: cmhaltserv DAR_MCCM2
Sep 13 08:35:51 mt1s syslog: su : + tty?? root-TPSadm
Sep 13 08:35:54 mt1s last message repeated 2 times
Sep 13 08:35:59 mt1s syslog: su : + tty?? root-oracle
Sep 13 08:36:00 mt1s syslog: su : + tty?? root-oracle
Sep 13 08:36:11 mt1s CM-MCCM2[21124]: cmmodnet -r -i 13.49.152.130
13.49.152.0
Sep 13 08:36:11 mt1s CM-MCCM2[21129]: cmmodnet -r -i 13.50.70.130
13.50.70.0
Sep 13 08:36:19 mt1s LVM[21184]: vgchange -a n /dev/vg1cm2
Sep 13 08:36:19 mt1s LVM[21189]: vgchange -a n /dev/vg2cm2
Sep 13 08:36:19 mt1s cmcld[17494]: Halted package MCCM2 on node mt1s.
Sep 13 08:40:46 mt1s CM-CMD[21446]: cmhaltpkg MCCM2
Sep 13 08:41:05 mt1s syslog: su : + ttyp2 ohernan-oracle
Sep 13 08:41:44 mt1s LVM[21535]: /sbin/vgchange -a e /dev/vg1cm2
Sep 13 08:41:55 mt1s LVM[21569]: /sbin/vgchange -a e /dev/vg2cm2
Sep 13 08:41:59 mt1s LVM[21589]: /sbin/vgchange -a e /dev/vg2cm2
Sep 13 08:42:09 mt1s syslog: su : + ttyp2 ohernan-oracle
Sep 13 08:42:52 mt1s syslog: su : + ttyp1 ohernan-TPSadm
Sep 13 08:46:28 mt1s cmcld[17494]: (mccm1s) Started package MCCM1 on node
mccm1s.
```

Se pueden revisar los siguientes archivos e log para obtener información adicional, si el resultado del "tail" no es satisfactorio:

/var/adm/cmcluster/HA_MT?.log o
/var/adm/cmcluster/HA_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log representa la salida del resultado de la corrida del shell script "control.sh".

/var/adm/cmcluster/HA_PT1.log conserva los mensajes enviados durante el arranque del Paquete PT1.
También podemos revisar:

/var/adm/cmcluster/services_MT?.log o
/var/adm/cmcluster/services_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene la Salida de la Aplicación de Proceso y del monitoreo de los servicios de Oracle. En caso de que los procesos de TPS u Oracle fallasen, podemos buscar los siguientes mensajes en el log de services_MP.log ó services_MT.log :

/etc/cmcluster/MT?/control.sh.log ó
/etc/cmcluster/MP?/control.sh.log

regresan una salida en caso de que el paquete falle o trabaje correctamente al ser levantado en el nodo adoptivo.

Si el paquete falla en el nodo adoptivo al ser levantado, el administrador puede revisar el log del script de control y detectar el problema. El siguiente comando debe ser ejecutado para habilitar el paquete en el nodo adoptivo, después de revisar y arreglar la falla:

#cmmodpkg –e –n <nodo adoptivo> < paquete que no se puede levantar>

Si el paquete no falla otra vez al ser levantado en el nodo adoptivo, y regresado a su nodo de origen después de que la falla ha sido arreglada en el nodo de origen.

4. Verificar que los File system estén desmontados

Asegurarse de que los file system listados en las tablas A.13 y A14, estén correctamente desmontados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla A.13 Inicialización de los Clusters MT

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla A.14 Inicialización de los Clusters MP

5. Revisar los Procesos que deben ser detenidos por cada una de las aplicaciones.

Las tabla A.15, lista los procesos en el Sistema que deben estar detenidos por aplicación, involucardas en los paquetes.

Software o Aplicación	Procesos	Equipo
Oracle	Ora_pmon_<ORACLE_SID> Ora_dbwr_<ORACLE_SID> Ora_arch_<ORACLE_SID> Ora_lgwr_<ORACLE_SID> Ora_ckpt_<ORACLE_SID> Ora_smon_<ORACLE_SID>	MT, MP
Aplicación de Transferencia Y procesamiento (TPS)	Manmbin Monsbin Mistsbin Manfsbin	MT, MP
CONNECT:Direct	Cdpmgr	MT, MP
SQLNET	Tnslsnr	MT, MP
Darput	Darput	MT
Omnistorege	Ded Qr	MT
Omniback	Crs Lockmgr	MT, MP
Alarm daemon	Alarm.sh	MT, MP

Tabla A.15 Lista de Procesos

Si el paquete no puede ser levantado en su nodo original, MC/ServiceGuard podría tratar de levantarlo en el nodo adoptivo.

Comportamiento de las Aplicaciones Durante un Failover

Se hace la observación, de que las funcionalidades del Sistema continúan trabajando de manera adecuada sin verse afectado la capacidad y el performance de los equipos debido a que un equipo puede correr las aplicaciones y nivelar la carga de trabajo de la otra máquina en problemas, de manera correcta.

Esta sección desglosa el comportamiento de las aplicaciones y de los cambios de éstas por la manera en que está diseñada y configurada la Alta Disponibilidad (High Availibility) en este proyecto:

Oracle Server

Dos instancias de Oracle pueden estar corriendo en el equipo adoptivo, sin ningún problema. Los archivos de configuración de las bases de datos (initSID.ora) pueden existir ambos en el nodo primario y en el secundario.

Las instancias de Oracle pueden ser levantadas en modo archive log.

En la tabla A.16 se muestran los archivos de configuración de la Base de Datos necesarios para levantar una instancia.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
$ORACLE_HOME/dbs/initSID.ora	Local	No	
$ORACLE_HOME/dbs/sgadefSID.ora	Local	Ejecutar un Cleanup Después Del Restore	

Tabla A.16 Archivos de configuración para la Base de Datos

SQL *NET

Pueden estar corriendo de manera simultanea dos SQL *NET LISTENER en el equipo adoptivo. La configuración de los LISTENERS permite ser levantados en cualquiera de los equipos. El listener del paquete que falla puede ser levantado en el nodo adoptivo después de un failover y los archivos de configuración que intervienen en el proceso son mostrados en la Tabla A.17.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
/etc/listener.ora	Local	No	
/etc/sqlnet.ora	Local	No	
/etc/tsnames.ora	Local	No	

Tabla A.17. Archivos de configuración para SQL *NET

El nombre de los LISTENERS están definidos como "LISTENER_MT? y LISTENER_MP?. Todos los Listeners definidos para los paquetes MT ó MP tienen diferentes números de puertos asignados.

OMNIBACK

Solamente es posible tener una instancia de esta aplicación corrienado en el equipo adoptivo. El horario y cintas de respaldo pueden ser organizados para realizar los respaldos de los dos nodos del Cluster al mismo tiempo. Durante la operación normal, el respaldo de los directorios del otro nodo está configurado en si mismo. Después de un failover, los file system del nodo adoptado deben ser respaldados con OMNIBACK haciendo las inclusiones

de este en sus respaldos ya configurados. Los scripts para PREEXEC y POSTEXEC deben ser modificados y adecuados para soportar los respaldos del nodo adoptivo, sobre todo los de ORACLE para congelar la Base de Datos.

Puede ser configurado un "media pool" para un failover. Los respaldos y horarios en OMNIBACK pueden ser configurados para cualquier situación, esto siempre y cuando todavía se cuenten con los recursos.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
/etc/opt/omni/config/datalist/*	Local	No	
/etc/opt/omni/schedule	Local	No	

Tabla A.17Archivos de configuración para OMNIBACK

Hay que tener cuidado al configurar demasiados "datalists" ya que el tamaño de los archivos en el directorio "/etc/opt/omni/config/db/*" pueden crecer demasiado. Es mejor hacer una liga a otro file system, o en su defecto, programar un CRON para purgar la base de datos de OMNIBACK regularmente. La tabla A.17 presenta los archivos que almacenan la información de los datalist (respaldos programados en Omniback) y los calendarios de los mismos para su arranque en automático.

OMNISTORAGE CLIENT

El VBFS del sistema en falla puede ser montado en el nodo adoptivo y puede ser direccionado al "volume set" del nodo adoptivo. Solamente puede estar corriendo una instancia de OMNISTORAGE CLIENT en el nodo adoptivo.

Los VBFS file system pueden tener diferentes FSID, y la configuración de ambos file system pueden existir en ambos equipos. Los puntos de montado pueden ser direccionados al correspondiente volume sets en ambos equipos.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
/var/opt/omnistorage/data/volsets/vsbind.FSID	Local	No	

Tabla A.18 Archivos de configuración para la OMNISTORAGE

CONNECT DIRECT

Solo puede existir una instancia corriendo en el equipo adoptivo. Todas las aplicaciones pueden mandar peticiones a un demonio de C:D (Connect Direct).

TPS

Deben existir dos instancias de TPS corriendo en el equipo adoptivo. Toda la configuración de TPS se encuentra en los discos compartidos, y son transferidos hacia el equipo adoptivo después de un failover. Si requerimos levantar mas de una instancia de TPS es necesario remover el archivo descrito en la tabla A.19. Dicho archivo es abierto para protección en caso de que el administrador levante dos veces una misma instancia de TPS.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
$TPS_CONFIG/tps/TPS_master.pid	Shared	Clean up	No

Tabla A.19Archivo de Lock para TPS

DAR

Existen dos instancias corriendo en el nodo adoptivo. Toda la configuración se encuentra en los discos compartidos, y sor montados en el adoptivo después de un failover. Los archivos temporales y de bitácora (log files) también están ubicados en los discos compartidos. El archivo de configuración, de la tabla A.20, es necesario tenerlo a la mano para poder arrancar una segunda instancia.

Ruta de archivo	En Disco Local Compartido	Reconfigurarlo Después de un Failover	Reiniciar Instancia para Activar Configuración
$TPS_HOME/etc/setup/dar.cfg	Shared	No	

Tabla A.20 Archivo de configuración DAR.

Alarm

Se deben levantar dos instancias en el equipo adoptivo, el cuál manda mensajes de su respectiva MT instancia a la consola.

CRON del Sistema

Solo existe un archivo de Cron (crontab) para cualquier sistema. No existe cambio en los trabajos del Cron después de un failover remoto. El equipo que toma la responsabilidad del que falla. Cada trabajo en el Cron de cada equipo está configurado para detectar cuando ha existido un SwitchOver y los file system del orto equipo han sido adoptados.

Tiempo de Recuperación

El tiempo tomado para que la aplicación sea levantada en el equipo adoptivo es de aproximadamente 10 minutos, tomando en cuenta que MC/ServiceGuard tiene que realizar el shutdown de las aplicaciones (tirar las aplicaciones) en el equipo que está fallando y levantarlas por completo en el adoptivo.

Si el nodo se cae o se queda trabado no es necesario que MC/serviceGuard haga un shutdown de las aplicaciones primero en el otro nodo, solamente tiene que levantarlas en el adoptivo. El tiempo requerido se reduce, en este caso, a la mitad, es decir, 5 minutos.

Failback Remoto (El nodo que había fallado es recuperado)

Para regresar a su estado normal un nodo que ha fallado y ha tenido que hacerse un SwuitchOver a uno adoptivo, es necesario que este retome su paquete de aplicaciones. Para esto se requiere una intervención manual.

El administrador en turno, debe realizar los siguientes pasos para levantar el paquete en el nodo que ha sido recuperado ejecutando ésta secuencia de comandos de MC/ServiceGuard:

1. Si el nodo que falló fue reseteado abruptamente o sufrió una interrupción en la alimentación de energía, el Administrador debe primero hacer un reboot en éste. Si el SwuitchOver fue debido a que Oracle o TPS fallan, es recomendable que también se de un reboot al sistema primero.

2. Conectarse al equipo a recuperar como el usuario "root".

3. Levantar MC/ServiceGuard en el nodo y reintegrar el Cluster, ejecutando "cmrunnode"

cmrunnode

Si el nodo aún está corriendo, se debe restaurar este del siguiente modo:

cmchaltnode
cmrunnode

4. Detenga el paquete en el nodo adoptivo ejecutando "cmcahltpkg"
cmchaltpkg –n <nodo adoptivo> <nombre del paquete>

5. Levante el paquete en el nodo reciente levantado. Existen dos opciones para hacerlo:

a) Habilite la recuperación del nodo a la hora de levantar MD/ServiceGuard automáticamente y posteriormente ejecute los siguientes comandos:

cmmodpkg –e <nombre del paquete>

#cmmodpkg –e –n <nombre del nodo recuperado> <nombre del paquete>

ó

b) Levante el paquete manualmente ejecutando los siguientes comandos:

cmrunpkg –n <nombre del nodo recuperado> <nombre del paquete>

cmmodpkg –e <nombre del paquete>

Las siguientes operaciones son ejecutadas por el script de control de MC/ServiceGuard . El script de control es el mismo que se utiliza para detener o arrancar los paquetes de un Cluster: "/etc/cmcluster/MT? ó MP?/control.sh"

El "control.sh" shell script realiza estas operaciones:

Deteniendo los paquetes en el equipo adoptivo:

a) . Detiene las aplicaciones de los paquetes, haciendo un shutdown suave.
b) . Desmonta los file system compartidos y desactiva los grupos de volúmenes.
c) . Reconfigura las tarjetas de RED elimina la dirección de RIP.
d) . Reconfigura las aplicaciones si es necesario.

Levantando los paquetes en el nodo recuperado:

a) . Reconfigura las tarjetas de RED con la dirección RIP.
b) . Activa los grupos de volúmenes compartidos y monta los file system
c) . Levanta las aplicaciones.

Verificación de un Failback Remoto

La verificación de un Failback remoto puede hacerse en el equipo recuperado. Si MT1 falló, la verificación puede hacerse en la MT1 ya recuperada. De igual forma, si MP1 falló la verificación puede hacerse ya recuperada MP1.
La verificación se lleva a cabo de la siguiente manera:

1. Revisión de los mensajes en consola

Los mensajes pueden ser desplegados en la consola adoptiva, cuando un remote failback ocurre.

2. Verificar el status del Cluster

Para verificar el status del Cluster, el Administrador del equipo debe ejecutar el comando cmviewcl y realizando lo siguiente:

a) Conectarse a un nodo del Cluster como el usuario "root".
b) Ejecutar el comando
/usr/sbin/cmviewcl

3. Monitoreo de las bitácoras de sistema (Log Files)

Para verificar la bitácora del equipo se debe hacer lo siguiente:

a) Conectarse a uno de los nodos del Cluster como usuario "root"
b) Ejecutar el siguiente comando:
tail –30 /var/adm/syslog/syslog.log | more

Se pueden revisar los siguientes archivos e log para obtener información adicional, si el resultado del "tail" no es satisfactorio:

/var/adm/cmcluster/HA_MT?.log o
/var/adm/cmcluster/HA_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log representa la salida del resultado de la corrida del shell script "control.sh".

/var/adm/cmcluster/HA_PT1.log conserva los mensajes enviados durante el arranque del Paquete.
También podemos revisar:

/var/adm/cmcluster/services_MT?.log o
/var/adm/cmcluster/services_MP?.log
dependiendo del Cluster en donde nos ubiquemos.

Este log contiene la Salida de la Aplicación de Proceso y del monitoreo de los servicios de Oracle. En caso de que los procesos de TPS u Oracle fallen, podemos buscar los mensajes correspondientes en el log de services_MP.log ó services_MT.log :

/etc/cmcluster/MT?/control.sh.log ó
/etc/cmcluster/MP?/control.sh.log

regresan una salida en caso de que el paquete falle o trabaje correctamente al ser levantado en el nodo adoptivo.

Si el paquete falla en el nodo al ser levantado, el administrador puede revisar el log del script de control y detectar el problema. El siguiente comando debe ser ejecutado para habilitar el paquete en el nodo, después de revisar y arreglar la falla:

#cmmodpkg –e –n <nodo adoptivo> < paquete que no se puede levantar>

4. Verificar que los File system estén montados

Asegurarse de que los file system listados en las tablas A.21 y A.22, estén correctamente montados.

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MT1	MT1	/dev/mt1vg1 /dev/mt1vg2 /dev/mt1vg2	/opt/MT1_mnt1 /opt/MT1_mnt2 /opt/tmx/MT1/omnistore	mxmt1
MT2	MT2	/dev/mt2vg1 /dev/mt2vg2	/opt/MT2_mnt1 /opt/MT2_mnt2 /opt/tmx/MT2/omnistore	mxmt2

Tabla A.21 Inicialización de los Clusters MT

Equipo	Paquete	Volume Groups	File System	RIP Address/ Hostname
MP1	MP1	/dev/mp1vg1 /dev/mp1vg2	/opt/MP1_mnt1 /opt/MP1_mnt2	mxmp1
MP2	MP2	/dev/mp2vg1 /dev/mp2vg2	/opt/MP2_mnt1 /opt/MP2_mnt2	mxmp2

Tabla A.22 Inicialización de los Clusters MP

5. Revisar los Procesos que deben ser levantados por cada una de las aplicaciones.

La tabla A.23 lista los procesos en el Sistema que deben estar levantados por aplicación, involucradas en los paquetes.

Software o Aplicación	Procesos	Equipo
Oracle	Ora_pmon_<ORACLE_SID> Ora_dbwr_<ORACLE_SID> Ora_arch_<ORACLE_SID> Ora_lgwr_<ORACLE_SID> Ora_ckpt_<ORACLE_SID> Ora_smon_<ORACLE_SID>	MT, MP
Aplicación de Transferencia Y procesamiento (TPS)	Manmbin Monsbin Mistsbin Manfsbin	MT, MP
CONNECT:Direct	Cdpmgr	MT, MP
SQLNET	Tnslsnr	MT, MP
Darput	Darput	MT
Omnistorege	Ded Qr	MT
Omniback	Crs Lockmgr	MT, MP
Alarm daemon	Alarm.sh	MT, MP

Tabla A.23 Lista de Procesos

6. Si el paquete no puede ser levantado en su nodo original, MC/ServiceGuard podría tratar de levantarlo en el nodo adoptivo.

Comportamiento de las Aplicaciones Durante un Failback

Esta sección desglosa el comportamiento de las aplicaciones y de los cambios de estas por la manera en que está diseñada y configurada la Alta Disponibilidad (High Availibility) en este proyecto:

Oracle Server

Una instancia de Oracle debe estar corriendo en el equipo recuperado problema.

La instancia de Oracle debe ser levantada en modo archive log.

SQL *NET

Debe estar corriendo un solo SQL *NET LISTENER en el equipo. La configuración de los LISTENERS permite ser levantados en cualquiera de los equipos. El listener del paquete que falla puede ser levantado nuevamente en el nodo adoptivo después de que el equipo ha sido recuperado.

El nombre de los LISTENERS están definidos como "LISTENER_MT? y LISTENER_MP?. Todos los Listeners definidos para los paquetes MT ó MP tienen diferentes números de puertos asignados.

OMNIBACK

Solamente es posible tener una instancia de ésta aplicación corrienado en el equipo recuperado. El horario y cintas de respaldo son recuperados al arrancar OMNIBACK. Después de un failover, los file system del nodo adoptado deben ser respaldados con OMNIBACK haciendo las inclusiones de este en sus respaldos ya configurados, cuando el failback ocurre esto debe ser modificado en el nodo adoptivo. Los scripts para PREEXEC y POSTEXEC deben ser regresados a su normalidad, sobre todo los de ORACLE para congelar la Base de Datos.

OMNISTORAGE CLIENT

El VBFS del sistema puede ser montado en el nodo origen y arrancado OMNISTORAGE. Solamente puede estar corriendo una instancia de OMNISTORAGE CLIENT en el nodo origen.

CONNECT DIRECT

Solo puede existir una instancia corriendo en el equipo origen. Todas las aplicaciones pueden mandar peticiones a un demonio de C:D (Connect Direct).

TPS

Debe existir una instancia de TPS corriendo en el equipo adoptivo. Toda la configuración de TPS se encuentra en los discos compartidos, para ser transferida hacia el equipo adoptivo después de un failover.

DAR

Existe una instancias corriendo en el nodo origen. Toda la configuración se encuentra en los discos compartidos, que son montados en el origen después de un failback. Los archivos temporales y de bitácora (log files) también están ubicados en los discos compartidos.

Alarm

Se debe levantar una instancia en el equipo origen, la cuál manda mensajes de su respectiva MT o MP instancia a la consola.

CRON del Sistema

Solo existe un archivo de Cron (crontab) para cualquier sistema. No existe cambio en los trabajos del Cron después de un failback remoto. Cada trabajo en el Cron de cada equipo está configurado para detectar cuando ha existido un SwitchOver y los file system del otro equipo han sido adoptados.