7
2 ej.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE INGENIERIA

SISTEMAS DE ALTA DISPONIBILIDAD EN EQUIPOS DE COMPUTO

T E S I S

QUE PARA OBTENER EL TITULO DE:

INGENIERO MECANICO ELECTRICISTA
(AREA ELECTRICA - ELECTRONICA)

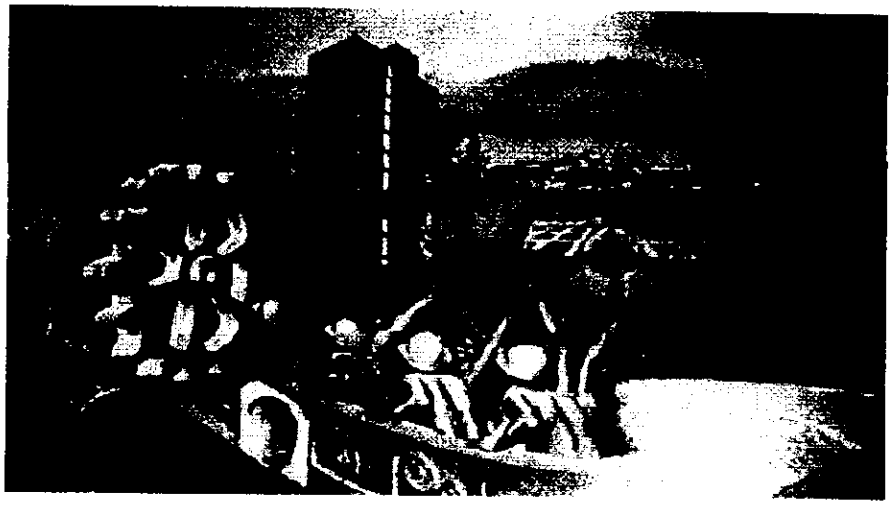
P R E S E N T A N:

AMARO BALLESTEROS RAUL
HERNANDEZ ESPINOSA HECTOR
LUNA SANCHEZ MARTIN ALBERTO
ROMAN VELAZQUEZ RENE
SOLTERO HORIO RICARDO

DIRECTOR DE TESIS:
ING. GLORIA MATA HERNANDEZ

Universidad Nacional Autónoma de México

Facultad de Ingeniería



Sistemas de Alta Disponibilidad en Equipos de Cómputo

Participantes:

Amaro Ballesteros Raúl

Hernández Espinosa Héctor

Luna Sánchez Martín Alberto

Román Velázquez René

Soltero Horio Ricardo

Mayo, 1998 **México D. F.**

AGRADECIMIENTOS

A mi madre, por el apoyo incondicional durante los tiempos difíciles.

A mi esposa, por su comprensión en esos tiempos.

Y a mis hijos, cuyo amor me impulsa día con día a ser un mejor ser humano.

ALBERTO

A Dios, por estar siempre conmigo.

A mis padres, con todo mi amor por su apoyo y confianza, sin los cuales no hubiera sido posible mi existencia. Gracias a mis maestros y amigos.

RENE

A mis padres y hermanos, por brindarme la experiencia de pertenecer a un familia.

A mi esposa e hijos, quienes son la fuerza principal que impulsa mis actos.

A mis compañeros de tesis, por su iniciativa y entusiasmo.

A todos quienes me apoyan en este y cada uno de los retos que presenta la vida y que se sabrán reconocer al leer estas lineas.

RICARDO

Agradezco a mi familia; que me ha formado y hecho quien soy, a todos los maestros; que gracias a sus enseñanzas me han permitido llegar hasta aquí y a la vida que me ha dado todo. Gracias por lo bueno y lo malo.

RAUL

Siempre he sido falto de palabras, y en este momento tengo que reducirlas aún más, esencialmente agradezco a Dios, a mis Padres, a mi amantísima esposa, a toda mi familia y a falta de palabras más lúcidas escojo las palabras de un amigo para dar las gracias:

¡Gracias te damos! ¡Gracias, Padre eterno!
Por todo lo que implica haber nacido,
¡gracias! Por vislumbrar, o no, el sentido,
¡gracias! Porque hay un Cielo y un Infierno,

¡gracias Por alternar, juicioso y tierno,
los días de sol, en que tu aroma es vivo,
con días de humor glacial y gesto esquivo,
para aprender a amarte, aún en invierno.

¡Gracias!

ÍNDICE

PRÓLOGO

En la última década el uso de los equipos de cómputo se ha vuelto vital para la operación de muchas compañías, como bancos, casas de bolsa, líneas aéreas, compañías de seguros, etc.. En estos lugares los sistemas deben estar trabajando ininterrumpidamente 24 horas al día los 365 días del año, ya que en todo momento deben estar listos para realizar transacciones en el mercado bursátil, para el caso de los bancos y casas de bolsa, o con el público en general en el caso de otras compañías, por tal motivo se han buscado algunos métodos para minimizar o evitar el tiempo de interrupción del servicio cuando se presenta algún evento que impide seguir operando normalmente. En caso de interrupción del servicio, la operación de las compañías se ve drásticamente afectada y repercute en la pérdida de dinero y la insatisfacción de sus clientes al no poder efectuar oportunamente las transacciones requeridas.

Algunas soluciones han sido comprar equipos tipo "mainframe" con procesadores redundantes, los cuales cumplen su objetivo pero son muy costosos y además el software y hardware que manejan es propio de la compañía que los fabrica, lo cual los hace incompatibles con los sistemas de otros vendedores.

De acuerdo a lo anterior, el objetivo de este trabajo es analizar y proponer diferentes alternativas para encontrar un balance costo-disponibilidad utilizando sistemas abiertos y la tecnología de vanguardia como lo es el concepto de arreglos

de discos, interfaces SCSI, interfaces SCSI-Fast-Wide, así como los diferentes estándares de redes locales: ethernet, token ring, FDDI.

Se analizan las principales causas que pueden provocar la interrupción del sistema y se plantean diferentes esquemas de interconexión de los dispositivos, utilizando fundamentalmente el concepto de hardware redundante, para lograr eliminar dichas causas. Las soluciones propuestas sólo se manejan a nivel de hardware y no recomendamos ningún producto de software en especial para controlar la actividad y el funcionamiento de los sistemas redundantes, debido a que al hablar de sistemas abiertos existe una gran variedad de productos de diferentes vendedores que podrían ser utilizados.

Un sistema que al experimentar la falla de un simple componente o recurso es capaz de mantenerse operando sin interrupción ó con una muy breve interrupción es conocido como un sistema de alta disponibilidad y en el presente trabajo se analizarán y propondrán diferentes sistemas, indicando ventajas y desventajas de cada uno de ellos.

Se espera que en un futuro próximo la tecnología de los dispositivos utilizados evolucione creando cada vez dispositivos más rápidos y sofisticados, pero los conceptos aplicados en este trabajo seguirán siendo vigentes ya que el objetivo principal, es el de eliminar las causas que provocan la interrupción del servicio en un sistema, y esto no cambiará.

Los capítulos que se incluyen en este trabajo son los siguientes:

- En el primer capítulo "Introducción", presentamos aspectos generales de la historia de los sistemas de cómputo, el planteamiento de la problemática a resolver, así como una introducción a los diferentes puntos de falla y niveles de

tolerancia en los sistemas abiertos, hasta llegar al concepto de sistemas de alta disponibilidad.

- En el segundo capítulo "Descripción de los elementos utilizados", se analizan los diferentes estándares utilizados en el presente trabajo, como son interface SCSI, discos y dispositivos de comunicaciones.

- En el tercer capítulo "Sistemas de alta disponibilidad propuestos" se analizan los diferentes niveles de disponibilidad para los sistemas abiertos y se proponen algunas configuraciones indicando sus limitaciones.

También incluimos al final las conclusiones a las que los miembros del equipo de trabajo llegamos, después de analizar el trabajo en perspectiva.

A lo largo del trabajo se utilizan términos en inglés, los cuales no tienen una traducción directa al español y al hacerlo pierden su significado original, por esta razón incluimos un glosario de términos a fin de dar una definición clara en caso de que no se conozcan dichos términos.

1
INTRODUCCIÓN

ANTECEDENTES HISTÓRICOS

En sus inicios la computación estuvo enfocada a cubrir necesidades específicas de negocios. Por ejemplo, sistemas de manejo financiero, control de inventarios, control de procesos y procesamiento de palabras, estos sistemas se apoyaron fundamentalmente en la computación tipo anfitrión, la cual consistía de un computador central controlando todas las unidades, terminales, y otros dispositivos como esclavos. Muchas de estas terminales tenían inteligencia limitada o no la poseían y por tanto dependían por completo de la disponibilidad del computador anfitrión para ejecutar la función esperada.

Los ambientes de computación del anfitrión condujeron a grupos aislados de tecnología, aplicaciones e información, que servían únicamente a aquellos usuarios que se conectaban al anfitrión. Con el paso del tiempo aumentaron las necesidades del usuario y muchas aplicaciones diferentes se concentraron en estos anfitriones. Los usuarios a menudo se molestaban debido al pobre o inconsistente desempeño y a la baja disponibilidad del servicio. Estos ambientes eran complejos de manejar debido a la variedad de aplicaciones y software de soporte de sistemas.

Los proveedores de estos sistemas competían al utilizar sus propios productos de hardware y software, los cuales se diseñaban sólo para una clase de sistema

computacional. Estos sistemas propietarios eran incompatibles por ser exclusivos del fabricante, lo cual restringía al usuario a la dirección tecnológica del proveedor o a una conversión costosa. Estas incompatibilidades condujeron a un aislamiento más profundo de las redes individuales del anfitrión, que hicieron de la interconexión de diferentes sistemas computacionales un desafío mayor, si no imposible.

La naturaleza aislada, independiente y consolidada de la computación del anfitrión se hizo muy restringida durante la década de 1980. En consecuencia, surgieron nuevos requisitos de negocios para la computación y nuevos modelos de arquitecturas computacionales para desafiar los bastiones de la computación del anfitrión.

La nueva realidad de los negocios creó atracción a un nuevo estilo de computación, basada en redes distribuidas y no anfitriones.

La supervivencia de las grandes empresas depende de vincular efectivamente las unidades de negocio y los socios comerciales. En otras palabras, es necesario concebir las operaciones de negocio no como entidades apartadas sino como funciones muy interdependientes e interconectadas.

Para tener éxito es esencial contar con la rapidez, la responsabilidad, la asociación, la productividad y otras tendencias claves, como lo es la alta disponibilidad de los sistemas.

COMPUTACIÓN EN RED

Un ambiente de computación en red suministra los medios para que los usuarios tengan acceso a una amplia gama de información, aplicaciones y recursos de computación sin tener que preocuparse por el lugar donde se encuentran o como están interconectados.

Fundamentalmente éste es un enfoque diferente de la computación tipo anfitrión. Múltiples puntos a lo largo de la red pueden entregar servicios de computación en contraste con la ubicación solitaria de un computador primario (anfitrión). Los diversos puntos de entrega de los servicios y aplicaciones del computador se consideran plataformas computacionales, las cuales se asignan a varios sitios para manejar diferentes grados de participación y suministrar el más adecuado tipo de capacidad computacional y de comunicaciones. Los tipos de plataformas incluyen estaciones de trabajo, servidores de trabajo en grupo en redes de área local, sistemas de información por departamentos en procesadores distribuidos, sistemas de información corporativa en procesadores empresariales y proveedores de servicio externo en redes públicas.

Los computadores anfitriones tradicionales involucraban largas aplicaciones monolíticas de software que corrían en ambientes altamente centralizados. La computación en red hace uso de un nuevo enfoque llamado procesamiento cooperativo, el cual involucra la expansión de componentes de aplicación a través de múltiples plataformas y la utilización de la red para vincular estos componentes.

Con la proliferación de la computación en red, algunos fabricantes se han dedicado a diseñar sus equipos basándose en estándares comunes, de tal forma que las aplicaciones que se encuentran funcionando en un sistema de una marca siguen operando si se trasladan a un sistema de otra marca. Para realizar esto, no es

necesario realizar ningún proceso complicado o laborioso, además de que se pueden comunicar entre sí a través de la red con sus servicios de comunicaciones nativos.

La utilización de estándares comunes por parte de los diferentes fabricantes de sistemas, da origen al concepto de sistemas abiertos.

SISTEMAS ABIERTOS

Los sistemas abiertos son ambientes de hardware y software basados en estándares que son independientes del proveedor. Es decir la computación ha madurado hasta el punto en que están adoptándose ampliamente los estándares, los cuales son regulados por organismos mundiales y por lo tanto no son monopolizados por ningún proveedor en especial, cada fabricante desarrolla su propia versión de productos, tomando como base los estándares. Este cambio profundo da paso a un nuevo mundo de posibilidades y desafíos para los clientes.

Las principales características de los sistemas abiertos son:

- Portabilidad. Las aplicaciones de software y la información pueden trasladarse con relativa facilidad a varios computadores de diversos tamaños o marcas. Esto posibilita que las tareas de las personas también sean portátiles, es decir sean aplicables a diferentes ambientes de computación.

- Interoperabilidad. Los computadores de diferentes tamaños y marcas pueden comunicarse entre sí, compartiendo recursos, información e incluso aplicaciones de software.

- Accesibilidad. Los sistemas abiertos reducen los costos en las áreas de hardware, software, administración de la información y costos humanos de administración del cambio, también suministran beneficios con valor agregado

como reducción del riesgo de dependencia del proveedor, flexibilidad arquitectónica, mejor integración del software, migración mas fácil hacia nuevas tecnologías innovadoras y una mejor selección de paquetes de software.

Los sistemas abiertos han llegado a ser tan indispensables para la operación de las grandes compañías, que el no contar con su servicio representa fuertes pérdidas.

Un sistema y sus componentes, los cuales afectan de manera drástica la operación de una compañía en el evento de una interrupción del servicio es conocido como ambiente de misión crítica.

AMBIENTE DE MISIÓN CRÍTICA

En la actualidad el uso de los equipos de cómputo se ha vuelto vital para la operación de las grandes compañías, ya que en estos equipos se procesan grandes volúmenes de información ya sea de los clientes o de los empleados y dan atención directa al público.

Este tipo de computadoras utilizan aplicaciones en donde la interrupción del servicio causa fuertes pérdidas en la operación de la compañía, estas computadoras y todo su ambiente son conocidos como Sistemas de Misión Crítica.

El impacto causado por la falla de alguno de los componentes del Sistema de Misión Crítica va desde la pérdida de muchos millones de pesos, por ejemplo en las operaciones de una casa de bolsa, hasta la insatisfacción de los clientes en espera de servicio en algún banco o línea aérea.

CAUSAS QUE PROVOCAN LA INTERRUPCIÓN DEL SISTEMA

La interrupción del servicio puede ser causada por dos tipos de eventos, aquellos que son planeados y los que se presentan de manera inesperada:

- Actividades Planeadas. Las actividades planeadas interrumpen el servicio de manera programada. Algunos ejemplos de estas actividades son: Mantenimiento Preventivo, Actualización al Hardware del equipo, Aplicación de notas de servicio, Actualización del sistema operativo, etc.. Aunque al realizar estas actividades el sistema sale de operación de manera planeada y ordenada, de cualquier forma se está quitando el servicio al usuario y en un ambiente de misión crítica provoca problemas en la operación de la compañía.

- Actividades no programadas o inesperadas. Las actividades no planeadas son las fallas en el Hardware, Software o alimentación eléctrica del sistema. En este tipo de eventos los problemas causados en la operación es mas fuerte ya que generalmente toma tiempo regresar el equipo a operación y desafortunadamente este tipo de eventos se presentan algunas ocasiones en horarios y fechas donde la carga de trabajo es mayor, horario de oficina, días de cierre de operaciones, etc..

Para mantener los sistemas de misión crítica operando ininterrumpidamente es necesario eliminar las causas que en un evento de falla interrumpirían el servicio, estos puntos los llamaremos "puntos de falla".

PUNTOS DE FALLA

Un punto de falla es cada uno de los elementos que pueden interrumpir el servicio a los usuarios de un sistema de cómputo. Los puntos de falla provocan la interrupción del servicio de manera inesperada, los principales son:

- Falla en la alimentación eléctrica del equipo.
- Falla en el CPU.
- Falla en la unidad de disco del sistema operativo o disco que contiene la aplicación crítica.
- Falla en la interface de alguno de los discos.
- Falla en alguna tarjeta de red.
- Errores humanos.

Para mantener un sistema de misión crítica operando ininterrumpidamente es necesario realizar un análisis de todos sus posibles puntos de falla y la magnitud del daño que causaría a la operación un evento de interrupción del sistema. Después del análisis anterior se proponen soluciones utilizando dispositivos redundantes para eliminar la posibilidad de que al presentarse una falla se de una interrupción en el sistema y este pueda seguir operando normalmente. Así las actividades correctivas se pueden planear para realizarse en una fecha posterior y de manera controlada.

Los sistemas que se conforman de dispositivos redundantes y componentes que disminuyen la probabilidad de interrupción del servicio son conocidos como "Sistemas de Alta Disponibilidad".

SISTEMAS DE ALTA DISPONIBILIDAD

Existen diferentes niveles de disponibilidad, desde conectar el equipo a una fuente de energía ininterrumpible, hasta utilizar tarjetas e incluso computadoras completas redundantes. El grado de disponibilidad lo da la importancia de la operación a proteger, pensando qué tan costoso es para la compañía el evento de la interrupción del servicio, por ejemplo, si para la compañía lo único importante es que la información no se pierda aunque se interrumpa el servicio, se podría utilizar sólo discos redundantes, así aunque el procesador falla tendrá su información protegida, pero el servicio se interrumpirá. Si por otro lado se pretende que también el sistema siga trabajando, tendrá que utilizar un procesador redundante y algún software de control.

Para las grandes compañías resulta más caro el perder dinero y clientes por la interrupción del servicio que invertir en equipo de alta disponibilidad.

2

DESCRIPCIÓN DE LOS ELEMENTOS UTILIZADOS

El objetivo principal de este capítulo es presentar el contexto técnico sobre el cual se centrará nuestra propuesta de solución para un equipo de cómputo de alta disponibilidad. Los componentes de un sistema de cómputo, en términos generales, se pueden clasificar en dispositivos de entrada/salida, unidades de almacenamiento de datos primarios/secundarios y unidad de procesamiento central. Estos elementos son las partes de cualquier sistema de cómputo, sin embargo la forma en cómo se hacen, los criterios de diseño y la tecnología existente son características que dan vida a las diferentes clases de equipo. En este trabajo de tesis sólo incluiremos las tecnologías que en la actualidad se han clasificado como "Sistemas Abiertos", y en forma especial aquellas tecnologías que se han situado como "El estándar", ya sea de facto o por convención entre la industria de cómputo.

Existe una demanda creciente por equipos de cómputo de alto rendimiento. Para lograr este tipo de equipos se requiere invertir grandes cantidades de dinero en

investigación y desarrollo. Es así que cada vez más compañías prefieren concentrar sus esfuerzos en una sola rama de la industria, para así lograr una ventaja competitiva. Esto crea compañías que solo producen un tipo de producto y para que pueda ser consumido por otras compañías necesita cumplir con ciertas características, a esto se le llama estándar. Por ejemplo, las compañías que producen CPU's requieren de periféricos de almacenamiento secundario que sean compatibles con sus CPU's.

Existe otra rama de la industria que tiene la tecnología y la capacidad para producir discos (almacenamiento secundario). Surge entonces la necesidad de establecer un estándar de comunicación entre la unidad de procesamiento central y las unidades de almacenamiento secundario.

En los "sistemas abiertos" lo que produce una compañía puede ser usado en un sistema de cómputo desarrollado por otra. Es decir, se puede encontrar un equipo donde la unidad de procesamiento central sea marca "IBM", los periféricos de almacenamiento secundarios sean "Seagate" y las comunicaciones estén implementadas sobre equipo "CISCO". Estas combinaciones en antaño eran impensables, pero la necesidad de equipos más poderosos, más flexibles y menos caros han hecho que este tipo de combinaciones sea cada vez menos raras.

Los principales estándares que presentaremos en este capítulo son:

- La interface SCSI. Éste es un estándar de comunicación entre la unidad de procesamiento central y los periféricos de almacenamiento secundario. Existen otros estándares en el mercado para este tipo de comunicación, pero por el momento éste es el más socorrido en los equipos de mediano rango.

- Los arreglos de disco Disk Arrays. Ésta es la manera genérica con la que se conoce en los sistemas abiertos la redundancia de discos, para asegurar la información en los dispositivos de almacenamiento secundario.

- Discos de medio magnético Standalone Disk. Propiamente no es un estándar, pero es el medio más común de almacenamiento secundario.

- Interfaces y dispositivos de comunicaciones. Todos los sistemas de cómputo necesitan comunicarse, ya que de otra manera no serviría de nada tener capacidad de cómputo. Dentro de los medios de comunicación veremos lo que es ethernet, la interface FDDI, puentes y ruteadores.

- La unidad de procesamiento central *CPU*. Veremos qué características tienen los CPU's de los sistemas abiertos.

LA INTERFACE SCSI

El desarrollo de controladores de dispositivos con tecnología VLSI, y su consecuente abaratamiento, favoreció el uso de pequeños sistemas de almacenamiento que incluyeran en su construcción toda la lógica necesaria para su funcionamiento de manera independiente del CPU. Esto permitió retirar del procesador central toda la lógica, antes necesaria, para el manejo de los dispositivos de almacenamiento haciendo a las computadoras mucho más baratas. A mediados de los 80's la industria de la computación se vio invadida por toda clase de dispositivos tales como cintas magnéticas, discos de media removible y discos duros de distintos tamaños. Tal variedad obligó a la organización ANSI, en 1982, a determinar un estándar para dichos dispositivos.

En la actualidad el estándar SCSI *Small Computer System Interface* está aprobado por ANSI para la transmisión física y eléctrica de datos entre el procesador central y los periféricos.

Especificación ANSI, SCSI-1

La interface SASI *Shugart Associates System Interface* fue adoptada por el comité ANSI X3T9 en 1981 como un estándar en interfaces. En 1982 un subcomité (X3T9.2) comenzó a trabajar en un nuevo estándar de I/O (Input/Output) llamado "Small Computer System Interface". No fue hasta 1986, después de muchas modificaciones que el comité X3T9 aceptó y presentó para ser aprobado por ANSI la revisión 17b de SCSI (se pronuncia Scozzy).

Especificación ANSI, SCSI-2

Debido a un cambio en la tecnología y a la necesidad de un mejor rendimiento, se generó un nuevo estándar SCSI-2, que mantiene las mismas características de SCSI-1 como un subconjunto de sus especificaciones, lo cual hace a ambas definiciones compatibles.

La especificación original de SCSI permitía un gran número de opciones de diseño. En un intento de reducir los riesgos de incompatibilidad, el estándar SCSI-2 elimina algunas de estas opciones.

Las principales características del estándar SCSI-2 son las siguientes:

- Notificación asíncrona de eventos. Esto permite la notificación de eventos aunque no existan comandos pendientes de ejecución (un evento es un cambio de estado ejemplo no disponible, ocupado, atención, etc...).

- Conjunto de comandos común CCS: Common Commnad Set. CCS es un subconjunto de los posibles comandos SCSI. Los distintos fabricantes utilizan CCS para incrementar la compatibilidad de sus productos con los de otros fabricantes.

- Cola de comandos CQ: Command Queuing. Es una opción que permite a múltiples comandos de un iniciador que se almacenen para una unidad lógica. Hay una cola de comandos de hasta 256 para cada unidad lógica de un solo iniciador. Esto mejora el rendimiento de la comunicación.

- Capacidad para nuevos dispositivos. Permite a la interface controlar nuevos dispositivos tales como CD-ROM's, Memoria óptica, Robots.

- SCSI Rápido Fast SCSI. Es una opción de rendimiento que permite doblar la velocidad síncrona de transferencia de datos.

- Conductor de 50 hilos de alta densidad. Esta opción permite mejorar la conexión que existía con el estándar anterior de 50 hilos de baja densidad.

Especificación ANSI, SCSI-3

Al final de 1995 se establece el estándar SCSI-3, aparece como una respuesta a la necesidad de incrementar la velocidad de comunicación entre el procesador central y los discos. Las principales mejoras de esta especificación sobre SCSI-2 son las siguientes:

- SCSI Ancho Wide SCSI. Es una opción para mejorar el rendimiento ya que permite cambiar los hilos ocupados para datos de 16 a 32.

- Cable y conector de 68 Pines. Permite soportar el cambio de 16 a 32 bits de datos.

- Velocidad de transferencia de 20 MB. Alcanzada al cambiar el lazo de voltaje por un lazo de corriente (wide differential SCSI Interface).

Requerimientos del cable

Los dispositivos SCSI son encadenados juntos usando un cable en común. Ambos extremos del canal deben estar terminados. Todas las señales son comunes entre todos los dispositivos en un canal SCSI. Los circuitos (drivers) permiten una longitud máxima de 6 metros en total en el canal terminados en forma simple *Single-Ended*. En los canales con circuitos diferenciales la longitud máxima del cable es de 25 metros.

Una impedancia ideal en el cable implica una impedancia característica de 132 ohms en canales con terminación simple y 122 ohms en canales diferenciales. Cables con una impedancia característica de mínimo 100 ohms son aceptables para cables en arreglo plano ribbon cable o par trenzado. Cables para arreglos blindados se aceptan que tengan una impedancia característica mayor de 90 ohms, para minimizar la interrupción de la señal por reflexión y cables de distintas impedancias no deben ser mezclados en el mismo canal.

Las configuraciones no blindadas son recomendadas para canales en gabinetes, las configuraciones blindadas son para canales externos o en ambientes con riesgos electromagnéticos y de descarga electrostática.

Interface SCSI con terminación simple

La interface SCSI de terminación simple Single-Ended (ya sea SCSI-1 o SCSI-2) permiten usar hasta 8 dispositivos SCSI, incluyendo el procesador central conectados a una interface o canal. Los dispositivos conectados al canal son descritos o denominados como iniciadores u objetivos Initiator, targets. El iniciador, normalmente el procesador central, controla la transmisión de datos comunicándose con un objetivo a la vez. El objetivo, normalmente un dispositivo en el canal, es capaz de requerir el canal para comunicarse con otro dispositivo convirtiéndose así en un iniciador. En la figura 2.1 se muestra el cableado básico del canal SCSI de teminación simple. Hay que hacer notar que el canal debe estar terminado por ambos extremos.

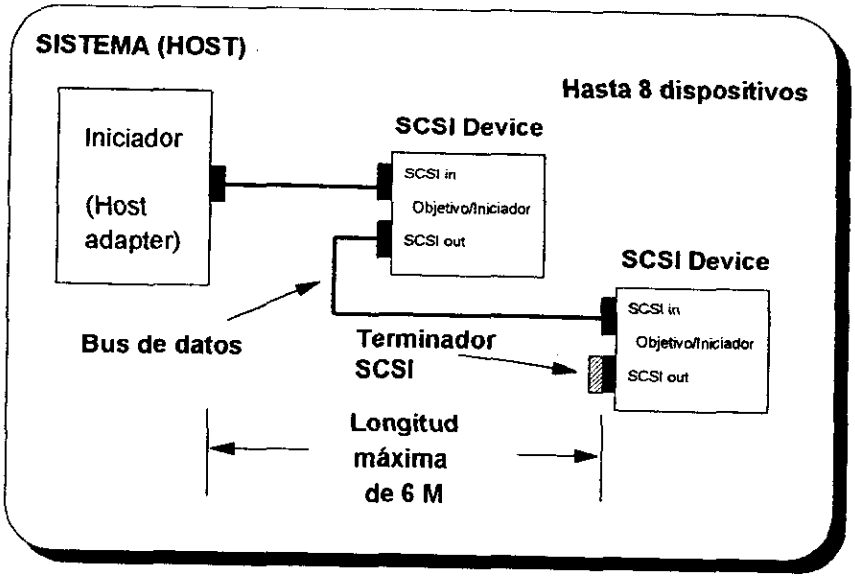


Figura 2.1 - Interface SCSI de terminación simple

La tabla 2.1 muestra la asignación de señales para un cable de terminación simple. Con respecto al cable el canal tiene 50 líneas, de las cuales la mitad son a tierra. Nueve de las líneas son usadas para datos (incluyendo paridad), el pin 26 para 5 [V] como fuente de poder para el terminador y las restantes son usadas

como señales de control. Las ocho líneas de datos, sin incluir la paridad, también funcionan como líneas de dirección de objetivo, así el iniciador puede contactar a un objetivo.

Los dispositivos de disco típicos pueden tener una transferencia máxima sobre el canal SCSI de terminación simple de 5 MB por segundo. Pero en términos reales la transferencia sostenida sobre un canal es mucho menor, en promedio 2.5 MB por segundo.

No. de pin	Nombre	No. de pin	Nombre
1	Ground	2	-DB(0)
3	Ground	4	-DB(1)
5	Ground	6	-DB(2)
7	Ground	8	-DB(3)
9	Ground	10	-DB(4)
11	Ground	12	-DB(5)
13	Ground	14	-DB(6)
15	Ground	16	-DB(7)
17	Ground	18	-DB(P)
19	Ground	20	Ground
21	Ground	22	Ground
23	Ground	24	Ground
25	Ground	26	TERMPWR
27	Ground	28	Ground
29	Ground	30	Ground
31	Ground	32	-ATN
33	Ground	34	Ground
35	Ground	36	-BSY
37	Ground	38	-ACK
39	Ground	40	-RST
41	Ground	42	-MSG
43	Ground	44	-SEL
45	Ground	46	-C/D
47	Ground	48	-REQ
49	Ground	50	-I/O

Tabla 2.1 - Cable de terminación simple

Interface SCSI rápida y ancha *Fast & Wide*

El estándar SCSI-2 define las especificaciones para la interface SCSI rápida fast y el estándar SCSI-3 define las especificaciones para la interface ancha wide. Esta interface utiliza circuitos diferenciales que son capaces de alcanzar mayores distancias (hasta de 25 metros) y mayores velocidades de transferencia (hasta 2 veces la velocidad que se logra en un SCSI de terminación simple). El término ancho (wide) implica que se utilizan más de ocho bits para la transmisión de datos y el término rápido (fast) implica que se duplicó la velocidad de reloj, comparada con la existente en el canal de terminación simple. Esta definición del canal SCSI puede direccionar hasta 16 dispositivos (incluyendo el procesador central) en comparación a los 8 del SCSI de terminación simple. En la figura 2.2 se muestra el cableado básico del canal SCSI rápido y ancho. Hay que hacer notar que, al igual que el canal de terminación simple, este canal debe estar terminado por ambos extremos.

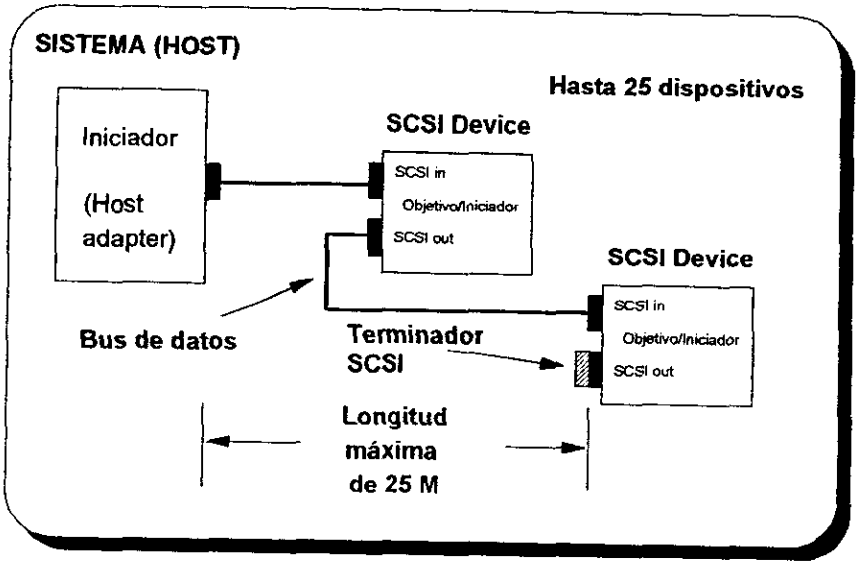


Figura 2.2 - Interface SCSI rápida y ancha

La interface SCSI rápida y ancha (se asume que se trata de líneas con manejadores diferenciales) usa un cable de 68 pines y alcanza una velocidad de transferencia pico de hasta 20 MB por segundo. Ya en la realidad la velocidad promedio de transferencia es de 10 MB por segundo.

En la tabla 2.2 se muestran los nombres de las líneas que se agregan en la interface SCSI rápida y ancha.

No. de pin	Nombre	No. de pin	Nombre	Pinout adicional para una interface " wide" diferencial
1	Ground	26	Ground	
2	+DB(0)	27	-DB(0)	
3	+DB(1)	28	-DB(1)	+DB(8)
4	+DB(2)	29	-DB(2)	+DB(9)
5	+DB(3)	30	-DB(3)	+DB(10)
6	+DB(4)	31	-DB(4)	+DB(11)
7	+DB(5)	32	-DB(5)	+DB(12)
8	+DB(6)	33	-DB(6)	+DB(13)
9	+DB(7)	34	-DB(7)	+DB(14)
10	+DB(P)	35	-DB(P)	+DB(15)
11	DIFFSENS	36	Ground	+DB(P1)
12	RESERVED	37	RESERVED	-DB(8)
14	TERMPWR	38	TERMPWR	-DB(9)
14	RESERVED	39	RESERVED	-DB(10)
15	+ATN	40	-ATN	-DB(11)
16	Ground	41	Ground	-DB(12)
17	+BSY	42	-BSY	-DB(13)
18	+ACK	43	-ACK	-DB(14)
19	+RST	44	-RST	-DB(15)
20	+MSG	45	-MSG	-DB(P2)
21	+SEL	46	-SEL	
22	+C/D	47	-C/D	
23	+REQ	48	-REQ	
24	+I/O	49	-I/O	
26	Ground	50	Ground	

Tabla 2.2 - Cable SCSI rápida y ancha "Fast & Wide"

ARREGLOS DE DISCOS *DISK ARRAY*

La parte más importante de los sistemas de cómputo son sus datos, su existencia, el manejo de los mismos y su disponibilidad es en gran medida lo que ha dado vida a los sistemas de cómputo. Éstos son los que definen a un sistema como de baja o alta disponibilidad. Son también de alguna manera los datos la parte más valiosa del equipo y sin embargo los dispositivos que los almacenan son la parte más vulnerable a fallas. En efecto, los discos de un sistema son los dispositivos periféricos que más fácilmente se dañan.

Los discos de todos los sistemas de cómputo son sensibles a vibraciones, campos magnéticos, calor, sonido, ruido eléctrico, envejecimiento, etc. A pesar de la evolución de la tecnología los discos de medio magnético siguen siendo la mejor opción entre todos los sistemas de almacenamiento. Hay otras soluciones además de los discos de medio magnético, pero o son muy caras o su capacidad de almacenamiento no es todavía lo suficientemente grande.

Dado que los discos magnéticos son todavía la mejor opción se ha pensado en otras formas de aumentar su confiabilidad, una de las más recurridas es el arreglo de discos.

Los arreglos de discos no son otra cosa más que conjuntos de discos que mediante algún tipo de redundancia aseguran la disponibilidad de los datos. Existen varias maneras de llevar a cabo los arreglos. El más básico es aquel que utiliza dos discos conectados en un canal de datos en paralelo, tal como se ilustra en la figura 2.3.

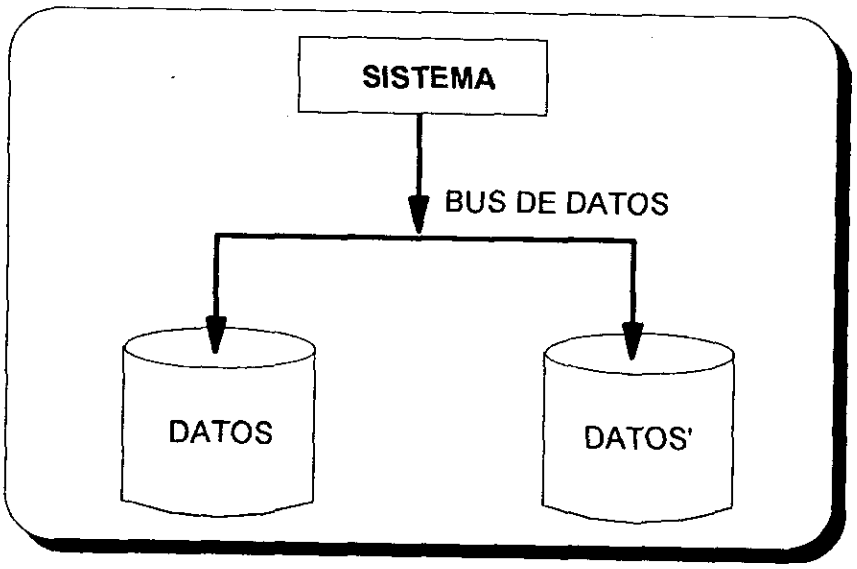


Figura 2.3 - Arreglo de discos básico

Este arreglo, con una realización, ya sea en hardware o en software, asegura la disponibilidad de datos manteniendo copias en discos separados. En la actualidad existen arreglos de discos mucho más sofisticados que utilizan múltiples redundancias y algoritmos realizados en hardware que permiten optimizar el uso del medio magnético. Hay incluso sistemas que permiten tener copia de los datos en dos puntos geográficos distintos, asegurando así la disponibilidad de la información aun en catástrofes. Los conceptos básicos para este tipo de arreglos se verán más adelante.

Los arreglos de discos más populares son los que utilizan múltiples controladores de interface al procesador central, múltiples controladores de disco y algoritmos de redundancia tipo RAID *Redundant Array of Independent Disks*.

Los arreglos de discos, contrario a los independientes, guardan la información en múltiples medios magnéticos. Los datos son distribuidos a través de los distintos

discos. Utilizando múltiples discos los arreglos consiguen varias ventajas significativas sobre los discos sencillos o independientes:

- Protección de datos. Una parte de los discos del arreglo puede ser usada para guardar información, con la cual se recuperan los datos en el momento de fallar uno de los discos. Esto mantiene el sistema funcionando y asegura la integridad de la información aun cuando se presente una falla en uno de los discos. Existen arreglos de disco que por desarrollos en hardware permiten mantener funcionando el sistema mientras se remplazan los discos dañados, lo cual permite la disponibilidad del sistema durante más tiempo.

- Flexibilidad de configuración. Los múltiples discos permiten una amplia variedad de formas de distribuir los datos a través de todo el arreglo. Esto permite maximizar el desempeño del sistema dependiendo de la carga de trabajo. El modo en que se distribuyan los datos determina la manera de comportarse del arreglo.

- Incremento en la capacidad de almacenamiento. Un arreglo de discos permite incrementar la cantidad de datos almacenados, al incrementar el número de discos conectados a una interface. Por ejemplo, en una interface SCSI se rompe la barrera de 8 o 16 dispositivos conectados.

Para lograr las ventajas anteriores, los discos independientes deben ser controlados adecuadamente. Esta labor la realiza el controlador de arreglo *array controller*. El controlador de arreglo balancea la carga de trabajo distribuyendo los datos a través de los distintos discos, y reconstruyendo la información en el evento de una falla de discos independientes. El controlador de discos hace la diferencia entre un arreglo de discos y un "montón" de discos.

Ya que el controlador del arreglo se encarga de manejar los discos del arreglo, el sistema es relevado de la necesidad de controlar a aquellos que trabajan

individualmente. Debido a la presencia del controlador del arreglo el sistema ve al arreglo como un solo disco de gran capacidad.

Algoritmos de redundancia de datos RAID

RAID son las siglas en inglés para *Redundant Array of Independent Disk* y es el nombre genérico para una serie de algoritmos que se han establecido como un estándar industrial para la protección de datos en arreglos de discos. Para poder hablar de ellos necesitamos definir algunos términos:

- Distribución de Datos *Data striping*: Es el proceso (o procesos) por el cual los datos son distribuidos a través de los distintos discos que conforman el arreglo. Este proceso es ejecutado por el controlador del arreglo y tiene el beneficio de balancear la carga de datos a través de los discos, mejorando el desempeño del sistema.

- Protección de Datos *Data protection*: Esto es exclusivo a los algoritmos RAID1, RAID3 y RAID5. Es una protección de datos que provee los medios para recuperar los datos en el evento de una falla. Esta información redundante es conocida como información de paridad, la cual es generada durante la escritura de los datos. Si un disco falla, la información de paridad y los datos restantes en los otros, son usados para reconstruir la información. El uso de información de paridad permite usar eficientemente los discos, además de proveer protección de datos. Usando sólo el 20% de la capacidad disponible en discos, la información queda protegida. En otro esquema de protección tal como la duplicación de datos, la capacidad utilizada para contener información redundante pude ser tanto como el 50%.

- Velocidad de Transferencia de Datos *Data transfer rate*: Es una característica fundamental de todos los subsistemas de disco. La velocidad de transferencia de datos refleja qué tan rápido pueden ser transferidos los datos del procesador central al disco. La habilidad de mantener una alta transmisión es particularmente importante en los sistemas que ejecutan largas transferencias de información.

- Concurrencia de entrada/salida *Concurrency I/O*: Es la habilidad del arreglo de discos de múltiples transacciones de entrada/salida simultáneamente. La concurrencia de un disco independiente impide a los restantes realizar otras transacciones. La alta concurrencia de entrada/salida es particularmente importante en sistemas multi-usuario.

Algoritmo RAID 0

El RAID0 permite el mejor desempeño y máxima capacidad sin protección de datos. La figura 2.4 muestra conceptualmente como son guardados en un arreglo con una realización de RAID0. Se muestra un arreglo de 5 discos aunque el concepto se aplica a uno de cualquier tamaño.

El controlador del arreglo escribe un bloque entero de datos a cada disco antes de continuar con el siguiente. El tamaño del bloque es definido por la profundidad de la distribución. Por simplicidad la figura muestra una distribución de un bloque, pero en realidad la distribución puede ser mucho más profunda.

Las principales característica de un arreglo en RAID0 son las siguientes:

- No protección de datos.
- Un muy alto índice de transferencia.
- Una muy alta capacidad para concurrencia de entrada/salida.

- Ya que no se utilizan discos para información redundante, se utiliza la máxima capacidad de disco para datos.

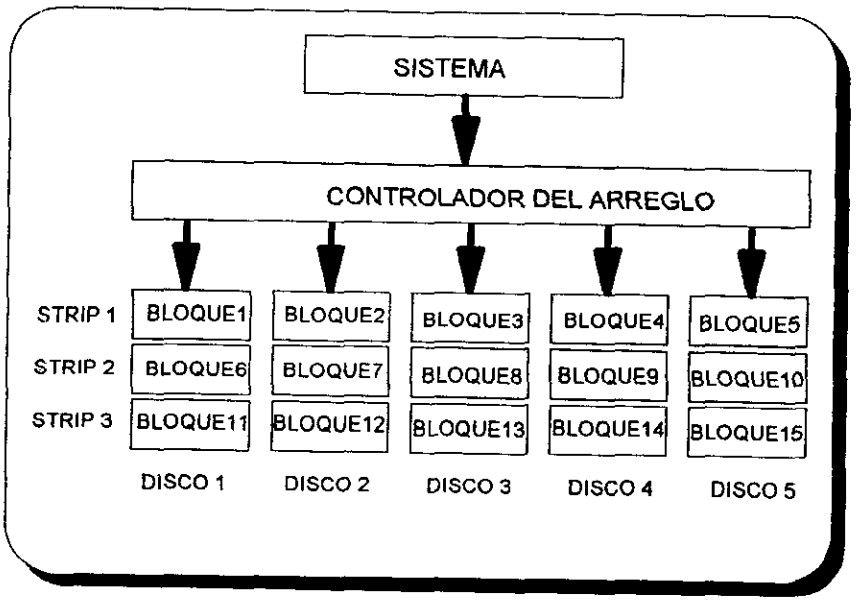


Figura 2.4 - Arreglo de discos RAID0

Algoritmo RAID 1

El RAID1 o par espejeado es un algoritmo en el cual se utiliza la copia en otro disco del arreglo para ofrecer la protección de los datos. En este tipo de implementación si un disco falla la información sigue disponible en otro disco del arreglo. En este tipo de arreglo la protección de discos es máxima, pero la capacidad de almacenamiento se ve reducida en un 50%.

En la figura 2.5 se muestra conceptualmente la implementación de un arreglo de discos con RAID1. Se muestra un arreglo de 4 discos aunque el concepto se aplica a arreglos de cualquier tamaño con número par.

Las principales características de un arreglo en RAID1 son las siguientes:

- Muy alta protección de datos.
- Un muy alto índice de transferencia.
- Una muy alta capacidad para concurrencia de entrada/salida.
- La capacidad en disco se ve reducida en un 50%.

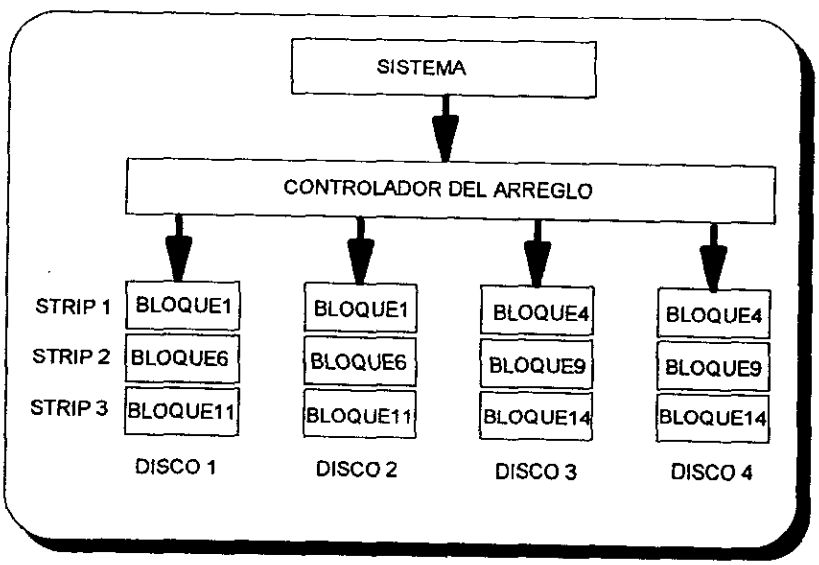


Figura 2.5 - Arreglo de discos RAID1

Algoritmo RAID3

El RAID3 ofrece protección de datos con alta transferencia. La figura 2.6 muestra conceptualmente un arreglo implementado con RAID3. Se muestra un arreglo con 5 discos, aunque el concepto se aplica a un arreglo de cualquier tamaño con un número impar de elementos.

El algoritmo RAID3 utiliza una técnica llamada distribución de Byte *byte striping* la cual distribuye la información en los discos byte por byte. Distribuyendo los datos

de esta manera implica que todos los discos del arreglo se ven involucrados en cada transacción de datos. Este tipo de implementación ofrece un muy alto índice de transferencia, ya que todos los discos trabajan en una transacción, pero ofrece un bajo nivel ante concurrencia de entrada/salida, debido a que todos los elementos están ocupados sirviendo sólo a una transacción a la vez.

La capacidad del arreglo se ve reducida en un disco, ya que es en éste donde se almacena la información de paridad.

Las principales característica de un arreglo en RAID3 son las siguientes:

- Alta protección de datos.
- Un alto índice de transferencia.
- Una baja capacidad para concurrencia de I/O.
- La capacidad en disco se ve reducida en un mecanismo.

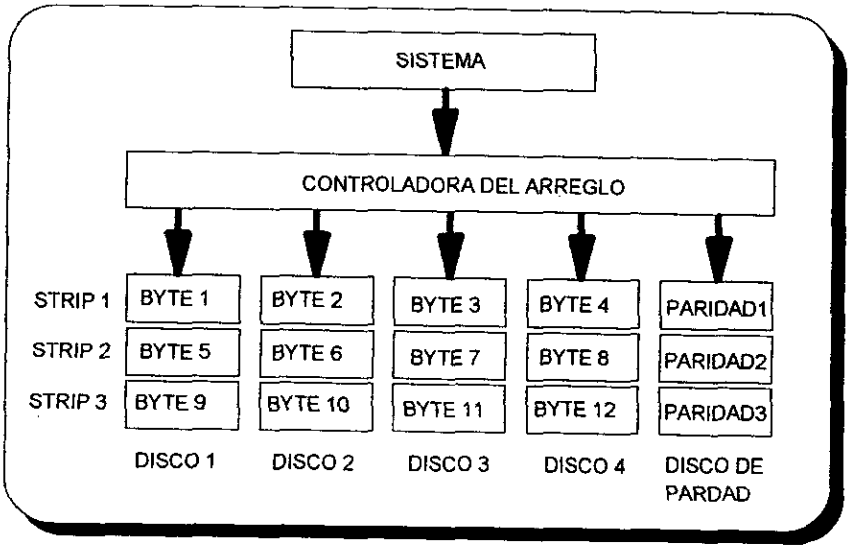


Figura 2.6 - Arreglo de discos RAID3

Algoritmo RAID5

El RAID5 ofrece una alta capacidad de protección de datos con una alta capacidad de concurrencia en entrada/salida. La figura 2.7 muestra conceptualmente un arreglo con implementación de RAID5. Se muestra un arreglo de 5 discos, aunque el concepto se aplica a arreglos de cualquier tamaño con un número impar de elementos.

El RAID5 guarda la información en forma de bloques a través de todos los discos del arreglo, así mismo la información de paridad está distribuida en todos los elementos. El controlador del arreglo escribe un bloque de información en un disco antes de continuar con el siguiente. El RAID5 utiliza distribución de bloques para distribuir datos a través de los discos del arreglo.

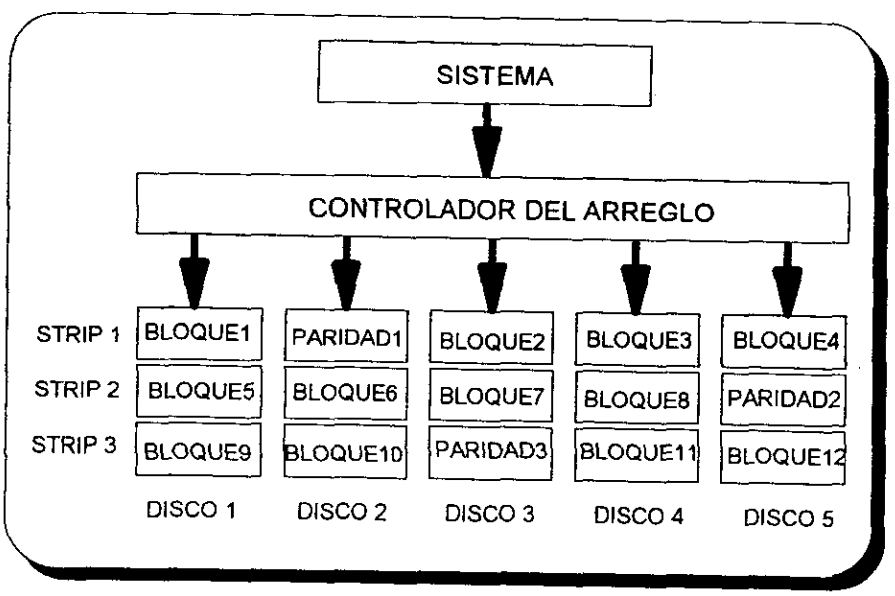


Figura 2.7 - Arreglo de discos RAID5

Las principales característica de un arreglo en RAID5 son las siguientes:

- Alta protección de datos.
- Un alto índice de transferencia durante lecturas y bajo durante escrituras.
- Una alta capacidad para concurrencia de entrada/salida.
- La capacidad en disco se ve reducida en 20 %.

DISCOS INDEPENDIENTES *STANDALONE DISC*

Los discos independientes, también conocidos como discos duros, son la manera más barata de almacenar información y tenerla disponible de manera casi inmediata. Existe una gran variedad de este tipo de discos. Difieren en cuanto el tipo de interface usada, al tamaño del medio magnético, la velocidad de acceso, temperatura de operación, capacidad de almacenamiento, "inteligencia", etc. Es tan grande, que intentar abarcar todos los tipos nos tomaría demasiado tiempo y no sería útil. En esta parte de la tesis limitaremos el alcance del trabajo a las características que tendrán los discos que usaremos en la solución sugerida. Definiremos lo que para nosotros es un disco de sistema abierto. Comenzando por las estructura física, la forma de uso del medio magnético, interface, configuración y recuperación del medio.

Estructura física del disco

Dentro de los discos de sistemas abiertos los más populares son los que almacenan la información en medio magnético de 3.7 pulgadas (95 mm) de diámetro, ya sea en 1 ó 4 de estos medios. A este tipo se le conoce en la industria como discos de 3.5 pulgadas. En todos los discos que utilizan este medio magnético el ensamble cabeza/disco está sellado en fábrica bajo una calidad de cuarto "blanco". El aire que circula en el interior del ensamble cabeza/disco se filtra para mantener libre de contaminación la superficie del medio magnético. En la figura 2.8 se muestra un ensamble típico de este tipo de discos. El resto de las características mecánicas varían dependiendo de la marca.

En estos discos el ensamble cabeza/disco nunca debe ser abierto ya que para poder darle servicio se requiere de un ambiente limpio de impurezas, de otra

manera se daña la superficie magnética. Usualmente estos discos son "monolíticos", no contienen partes remplazables.

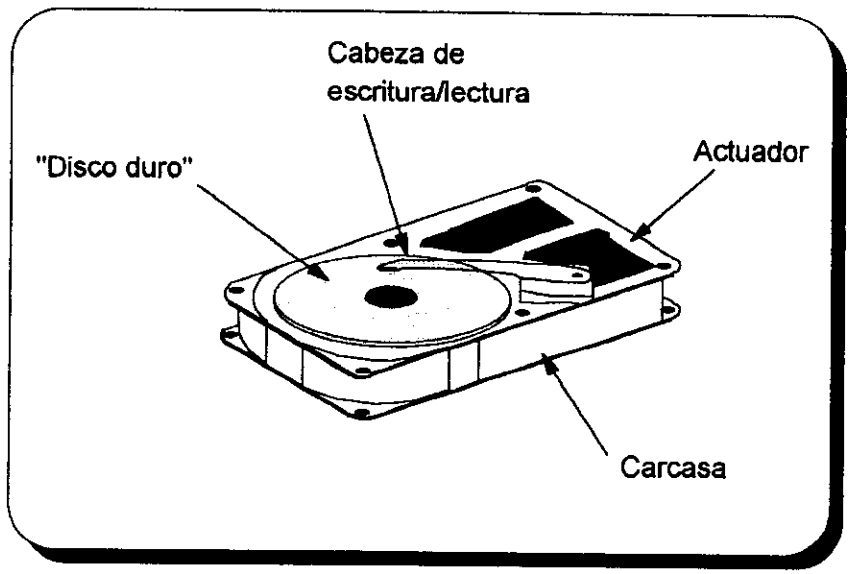


Figura 2.8 - Estructura típica de disco de 3.5"

Uso del medio magnético

La superficie de los discos está organizada en una serie de pistas circulares concéntricas *tracks*. Cada pista está subdivida en sectores. También se incluyen en la superficie pistas de repuesto *spare* y pistas de mantenimiento, éstas son áreas reservadas que se utilizan para guardar información de los diagnósticos así como código del controlador del disco. Existe una pista (usualmente la más interna) dedicada para aterrizar la cabeza de escritura/lectura con lo cual se evita que se dañe el área de datos cuando el disco se apaga. Algunos de estos discos incluyen en la superficie magnética información codificada de la posición de la cabeza de escritura/lectura con lo cual eliminan la necesidad de un transductor mecánico. En la figura 2.9 se muestra la distribución de las pistas sobre la superficie magnética del disco.

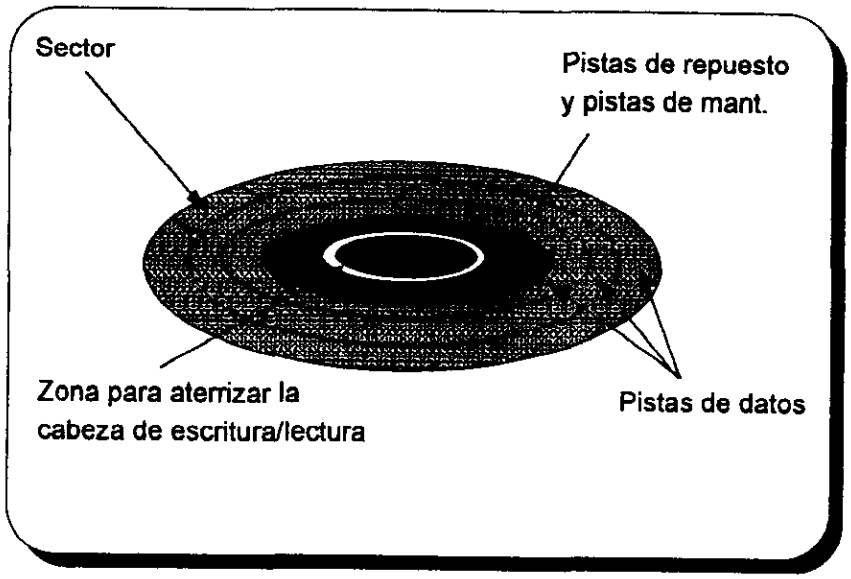


Figura 2.9 - Distribución de pistas en el medio magnético

Cada localidad de datos en el medio magnético está identificada por una dirección física y una dirección lógica. Las dos direcciones difieren porque la dirección física toma en cuenta la presencia de sectores que no contienen datos, tales como los sectores de remplazo y las pistas de mantenimiento, y las direcciones lógicas no. El procesador central en este tipo de discos se comunica usando direcciones lógicas. El controlador del discos convierte las direcciones lógicas a sus correspondientes coordenadas *cabeza, cilindro y sector*. La conversión toma en cuenta cualquier operación de remplazo *spare* ejecutada con anterioridad. De esta manera el procesador central no se tiene que ocupar del manejo de la información en el sistema de discos aumentando así el rendimiento del equipo.

Interface

Los discos que estaremos considerando para este trabajo de tesis son los que soportan interface SCSI, tal y como se describe en ANSI para SCSI-2 y SCSI-3.
La interface SCSI es sumamente flexible y permite que el procesador central se comunique con los periféricos y otros sistemas. Ésta también permite que los periféricos se comuniquen entre ellos. Existen cuatro distintas configuraciones de SCSI, en base a cuántos procesadores centrales y cuántos periféricos se encuentren en el canal. Usualmente los CPU's se comportan como iniciadores y los periféricos como objetivos. Las combinaciones posibles son las siguientes:

- Un sólo CPU iniciador/ Un sólo periférico objetivo
- Un sólo CPU iniciador / Múltiples periféricos objetivos
- Múltiples CPU´s iniciadores / Un sólo periférico objetivo
- Múltiples CPU´s iniciadores / Múltiples periféricos objetivos

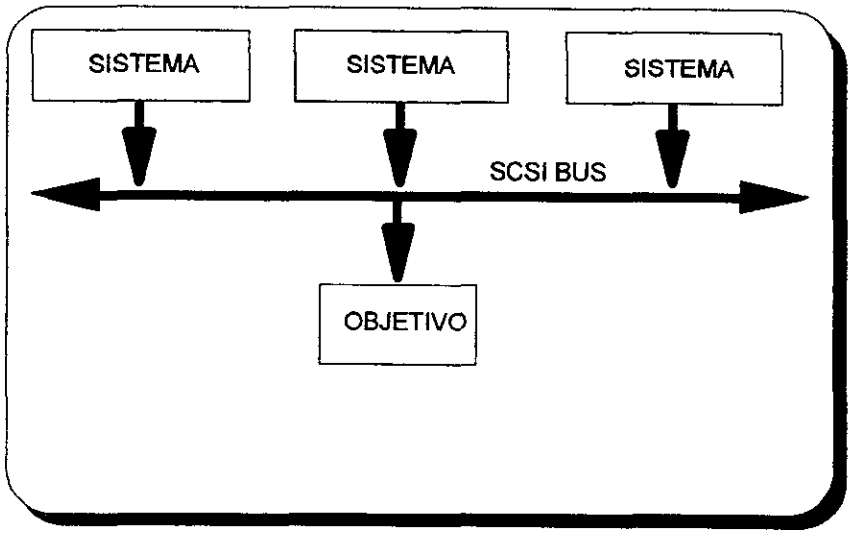


Figura 2.10 - Configuración con Múltiples iniciadores y un objetivo

Para nuestros objetivos las configuraciones que se utilizarán serán las de múltiples iniciadores, un sólo objetivo y múltiples iniciadores, múltiples objetivos. En las figuras 2.10 y 2.11se muestran dichas configuraciones.

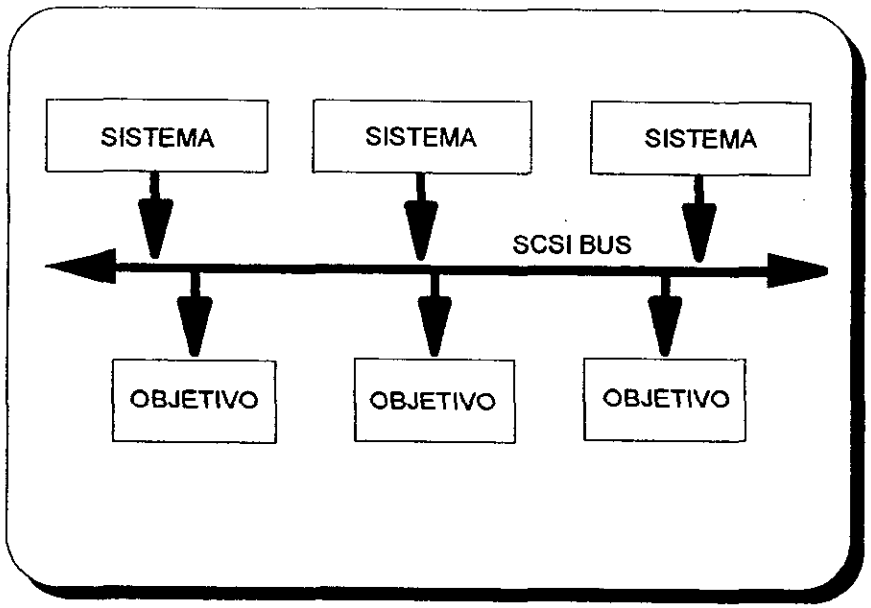


Figura 2.11 - Configuración con múltiples iniciadores y múltiples objetivos

INTERFACES Y DISPOSITIVOS DE COMUNICACIONES

Una parte muy importante de todo sistema de cómputo es la red de comunicaciones. Es este el medio por el cual la información es transmitida y compartida. De nada serviría tener un gran poder de cómputo y mucha información procesada si no la podemos tener en el momento y lugar adecuados. Imaginemos la situación de un banco, el cual invirtió una gran cantidad de dinero en un equipo de cómputo. Para que el banco funcione adecuadamente se necesita que la información del dinero de los cuenta habientes, que se encuentra centralizada en la casa matriz, este disponible en una y cada una de las sucursales. Y a su vez se requiere que una operación efectuada en cualquiera de sus sucursales se actualice lo más pronto posible en la casa matriz. Si la comunicación entre la casa matriz y las sucursales no fuera lo suficientemente confiable, podría provocar pérdidas económicas muy grandes al banco. Es por esto que tenemos que contemplar la seguridad en la red de comunicaciones.

En la actualidad el tamaño de la red de comunicaciones es algo que es muy difícil de definir con precisión. Con la existencia de Internet y de los sistemas abiertos el tamaño de la red depende del punto de vista del observador. Tomemos por ejemplo un estudiante de alguna universidad que desea imprimir un trabajo escrito en una impresora compartida en red ubicada en la misma sala de trabajo de donde se encuentra el estudiante, en este caso hablaríamos de una red local LAN *Local Area Network*. Pero si el mismo estudiante desde la misma computadora accesa los recursos de la biblioteca principal de su universidad, podríamos hablar de una red de área metropolitana MAN *Metropolitan Area Network*. Más aún, si ese mismo estudiante desde la misma computadora manda un mensaje electrónico a otro estudiante en otro país, estaríamos hablando de una red de área amplia WAN *Wide Area Network*. El mundo de las redes es muy amplio y no es el objetivo de este trabajo de tesis abarcar toda la variedad de dispositivos y medios de comunicación,

pero veremos las interfaces y dispositivos necesarios como para soportar la operación en una LAN o de una MAN en un ambiente de sistemas de alta disponibilidad. Comenzaremos por ver los conceptos utilizados en un protocolo de comunicaciones tal como lo es el ethernet. Después veremos los conceptos de una interface de comunicaciones de redes tipo MAN, como lo es la FDDI y terminaremos con los conceptos de dispositivos de comunicaciones básicos tales como los concentradores, puentes y ruteadores.

Ethernet

Ethernet es el estándar más utilizado hoy en día en comunicaciones y está definido por recomendaciones internacionales, específicamente IEEE 802.3. Ethernet fue desarrollado y patentado ante el dominio público por Xerox .

El primer sistema Ethernet, construido en el inicio de los setentas, fue un gran éxito, conectando más de 100 computadoras en una red a 2.94 Mbit/s. Posteriormente DEC, Intel, y Xerox colaboraron en un estándar para realizar la comunicación a 10 Mbit/s.

Actualmente existen dos especificaciones de Ethernet; el tipo original, llamado Ethernet, y la versión estandarizada por la IEEE, llamada 802.3. Estas dos versiones no pueden operar entre sí, es decir, si una computadora utilizando Ethernet manda un mensaje a una computadora utilizando tipo 802.3, el receptor no entenderá el mensaje. La diferencia entre las dos especificaciones está en la forma en que el contenido de los mensajes es interpretado.

Esta diferencia la podemos explicar de la siguiente manera:
En el mundo de la computación , el término campo aplica a un grupo de bytes entre un grupo más grande. Este grupo más grande podría ser un registro en una base de datos, un mensaje mandado a través de la red, o cualquier colección de datos

que tiene una estructura. Por ejemplo, en el mensaje ABCDEF, AB podría ser un campo, C otro y DEF otro. Cuando dos o más aplicaciones están usando los mismos datos, es importante que todos estén de acuerdo en el contenido de cada uno de los campos. De otra forma, la aplicación no podrá entender el significado de los datos. Por esta razón la computadora enviando mensajes en formato de ethernet no entenderá ni será entendida por una computadora que maneja sus mensajes en formato IEEE 802.3.

Se pueden mezclar los dos estándares en la misma red, sin embargo sólo se podrán comunicar entre sí las computadoras que están utilizando el mismo tipo de formato, el intercambio solo será entre esas dos computadoras y los otros equipos ignorarán esos mensajes.

Tecnología ethernet: cableado y topología

Ethernet puede usar dos tipos de cable coaxial (grueso y delgado) o cable tipo par trenzado. En el caso del ethernet grueso, la conexión al cable deberá ser hecha usando un transceptor, llamado MAU *Media Attachment Unit*. Este transceptor es sujetado al cable con un dispositivo conocido como vampiro. El transceptor contiene circuitos electrónicos que permiten que el cable de 9 pines que va del transceptor a la computadora se comunique con el cable grueso de ethernet. Sin los circuitos electrónicos del transceptor, las señales de los dos cables no son compatibles.

Con los cables de ethernet delgado, adaptadores tipo "T" son conectados directamente a la tarjeta de interface de ethernet para unir la computadora a la red. La tarjeta de interface de ethernet deberá estar lo más cercana posible a la red para evitar reflexiones.

Ambas instalaciones de ethernet, grueso y delgado, utilizan topología tipo canal, en ambos niveles, físico y lógico. En la modalidad de ethernet que utiliza cables tipo par trenzado las computadoras son conectadas a un concentrador. Esto significa que las instalaciones con par trenzado tienen una topología tipo estrella en el nivel físico. En el nivel lógico, sin embargo, el cableado aún tiene una topología tipo canal.

Los concentradores son generalmente interconectados con cable coaxial, de ahí que deben seguir las reglas aplicables al cableado con cable coaxial.

Cuando se alcanza el límite del número de computadoras que pueden ser conectadas a una red, ya sea porque la máxima capacidad de direccionamiento de ese segmento, ha sido alcanzado o es necesario reservar espacio para futuras conexiones, se necesita encontrar una forma de extender la red. Para hacer esto se utiliza un repetidor, un puente o un ruteador.

Para instalaciones de ethernet con cable coaxial, un repetidor es usado para hacer una simple red lógica más grande que los límites de una simple red física. Un repetidor es un dispositivo regenerador de señal. Éste reacondiciona señales y los pasa de un segmento de cable a otro. Cuando una señal viaja a través de un cable, varios efectos distorsionan la señal haciéndola difícil de ser detectada por la electrónica de la tarjeta de interface de red. Los repetidores toman una señal de un segmento, la filtran y la mandan a otro segmento.

En una red pueden ser instalados un número determinado de repetidores, no es recomendable instalar éstos en serie, ya que en un momento pueden introducir retardos y provocar problemas con la temporización de la señal, lo cual afecta notablemente el desempeño de la red.

La figura 2.12 muestra las principales características de una red ethernet utilizando cable coaxial delgado.

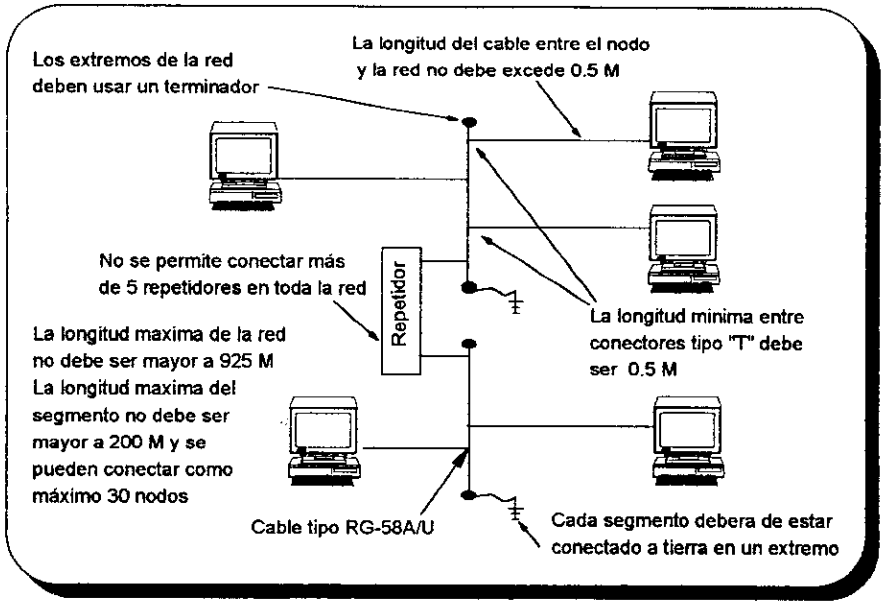


Figura 2.12 - Red Ethernet con cable delgado

Reglas para utilizar cable coaxial delgado:

- El máximo número de segmentos (redes físicas unidas por repetidores) es 5.
- La máxima longitud de un segmento es 180 m.
- La máxima longitud total (la longitud total de todos los segmentos) es 925 m.
- El máximo número de estaciones conectadas es 30 por segmento, ó 142 en total. (Cada repetidor cuenta como una estación para los dos segmentos que conecta)
- La mínima distancia entre conectores "T" es 0.5 m.
- En cada fin de segmento deberán usarse terminadores y uno de ellos deberá ser conectado a tierra.

Para cable grueso las reglas son ligeramente diferentes:

- El máximo número de segmentos (redes físicas unidas por repetidores) es cinco, pero sólo tres pueden tener computadoras conectadas. Las otras dos son simplemente para extender la longitud de la red.
- La máxima longitud de un segmento es de 500 m.
- La longitud total máxima (la longitud total de todos los segmentos) es de 2500 m.
- El máximo número de estaciones conectadas es 100 por segmento, o 492 en total (Cada repetidor cuenta como una estación para ambos lados de los segmentos que conecta).
- La mínima distancia entre transceptores es de 2.5 m.
- Un terminador deberá ser usado en cada fin de segmento, y uno de ellos deberá ser conectado a tierra.

Funcionamiento de ethernet

Ethernet utiliza el método de acceso a red conocido como CSMA/CD *Carrier Sense Multiplea Access with Collision Detection*. Con este método, alguna computadora que quiera acceder a la red deberá sensar el tráfico antes de transmitir. Para transmitir la estación verifica si hay señal en el cable, si no la hay, la transmisión puede empezar. La computadora deberá revisar inmediatamente para ver si ha habido una colisión debido a que otra computadora mandó datos al mismo tiempo. Si hay una colisión, la computadora para, espera una cantidad aleatoria de tiempo y transmite otra vez. Aunque esto suena complejo y de alto consumo de tiempo, todo es manejado por la tarjeta de interface de red.

TOKEN RING

La tecnología Token Ring fue originalmente propuesta a la IEEE como un estándar potencial en 1969. Aunque IBM públicamente demostró su interés en la tecnología en 1982, no fue sino hasta Octubre de 1985 que dicha compañía oficialmente anunció el Token Ring como parte de su familia de productos. La primera versión de Token Ring fue diseñada para trabajar a una velocidad de 4 Mbits/s y soportaba 260 computadoras.

Token Ring cumple con el estándar IEEE 802.5, por lo que ahora existen en el mercado muchos otros vendedores además de IBM. Algunos de ellos son 3Com Corporation, Madge Networks, y Ungerman Bass. IBM cuenta actualmente con una versión a 16 MBit/s, un estándar que ahora también ha sido soportado por varias terceras partes.

Tecnología de token ring: cablado y topología

Las redes Token Ring normalmente usan cable tipo par trenzado, blindado o sin blindaje, aunque también se puede usar cable coaxial (pero no siempre es recomendable, a menos que se use en distancias muy cortas). Las tarjetas adaptadoras de red se conectan a los cables con conectores DB9. El cable se conecta a la red con un conector *hermafrodita,* el cual a su vez se conecta a un concentrador llamado MSAU *Multistation Access Unit.*

Un MSAU generalmente tiene ocho puertos, aunque algunos vendedores a menudo ofrecen versiones de 24 puertos. Éste dispositivo contiene relevadores que conectan una computadora al anillo. Éste también tiene un puerto de entrada al anillo *Ring In RI* y un puerto de salida del anillo *Ring Out RO,* así varios MSAUS pueden ser conectados en serie (siempre RI a RO, nunca RI a RI o RO a RO).

También, cuando más de un MSAU se usa, todos los puertos RI y RO deberán ser conectados para completar el anillo.

Token Ring es lógicamente un anillo. Esto significa que hay una ruta (un cableado) que va de una computadora a la próxima, y sólo dos computadoras están en cada ruta. La distribución física de Token Ring es una estrella en cada MSAU, con los MSAUs conectados en anillo.

La figura 2.13 muestra un ejemplo de una red utilizando Token Ring.

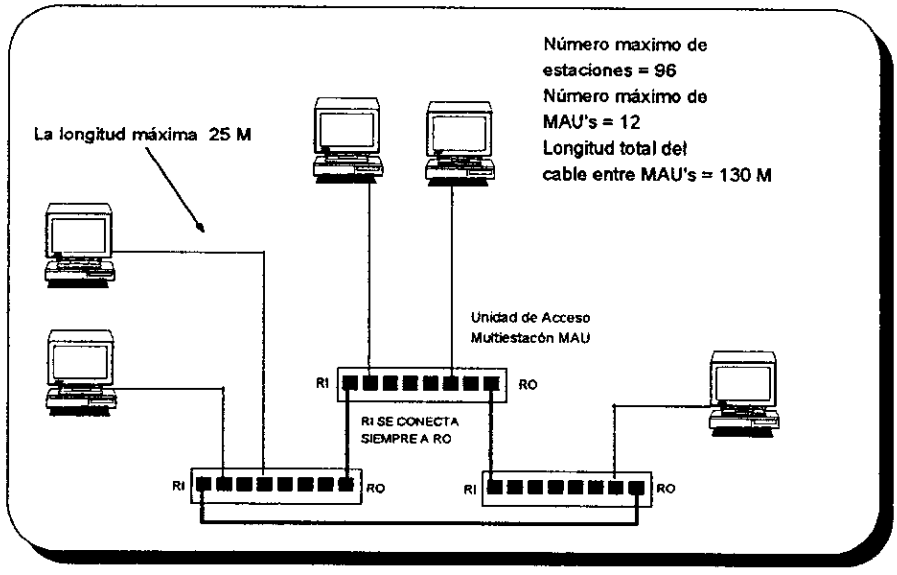


Figura 2.13 - Red Token Ring

Reglas para la instalación de una red utilizando Token Ring.

- El máximo número de computadoras es 260, si se utiliza cable tipo par trenzado con blindaje (72 para cable telefónico estándar).
- El máximo número de MSAUs es 33 (9 si es usado cable telefónico estándar).
- La máxima longitud de un cable entre un MSAU y una computadora es 45 m.

- El MSAU deberá ser cableado como un anillo usando las conexiones RI y RO. Un simple MSAU no necesita ninguna conexión RI o RO.
- Los puertos RI deberán conectarse al RO únicamente y viceversa.

Funcionamiento de token ring

En token ring pasa un mensaje de control de computadora a computadora alrededor del anillo, hasta que una de ellas lo toma. Esta computadora entonces reemplaza el mensaje por otro que contiene la información que desea enviar, este se pasa de un sistema a otro, hasta que llega a la computadora destino, la cual lo marca para indicar que lo ha recibido, el mensaje circula al rededor de la red hasta que regresa a la computadora que lo envió, ésta viendo que ha sido recibido, lo reemplaza con un nuevo mensaje de control y el proceso se vuelve a repetir.El mensaje de control es una señal conocida dentro del ambiente de las comunicaciones como *token*.

FDDI

En un esfuerzo por crear un estándar para alta velocidad, tolerante a falla en la red, la *American National Standars Institute* (ANSI) desarrolló el FDDI *Fiber Distributed Data Interface*. Este sistema está basado en un arreglo de dos pares de conexiones en dos anillos. La red funciona a 100 Mbit/s y puede cubrir distancias muy largas (125 millas). Los dos anillos están conectados, en forma tal que cuando se presenta alguna falla en uno de ellos se produce un enrutamiento hacia el otro asegurando continuo servicio.

La figura 2.14 muestra la conexión de los dos anillos, uno acarreando datos, llamado primario y el otro llamado el secundario es generalmente utilizado para

recuperación automática en el evento de falla en el anillo primario. Cuando una falla ocurre, las estaciones en uno u otro lado la detectan y automáticamente cubren la ruptura del anillo.

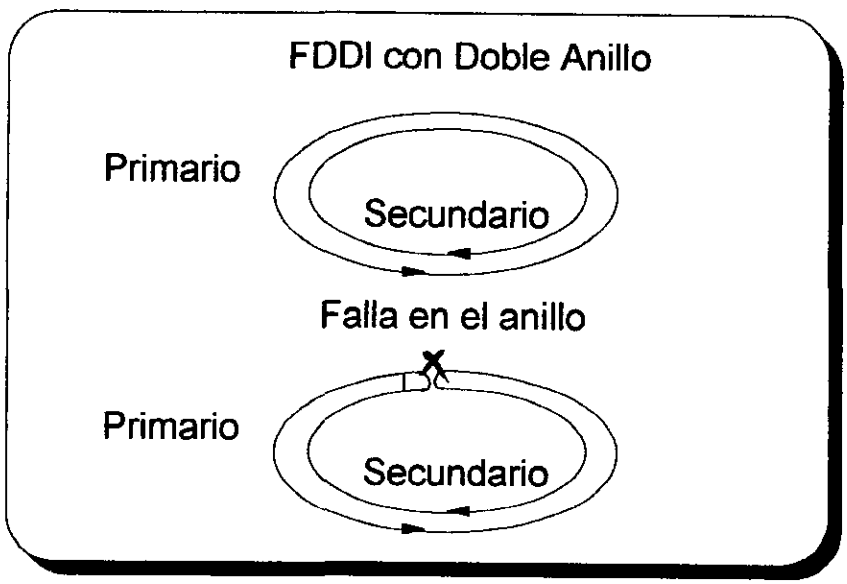


Figura 2.14 - Anillo Protegido

FDDI es muy similar en estructura y protocolo a token ring, pero éste tiene algunas diferencias que lo hacen más eficiente en el uso del ancho de banda de la red. El estándar para la configuración de la red permite conectar hasta 500 estaciones separadas en un anillo de 200 Km de longitud. El mecanismo de transmisión de FDDI es muy parecido al de token ring. Pero hay algunas diferencias significantes, la primera es que FDDI utiliza un código que le permite optimizar su ancho de banda. Otra diferencia mayor es que en FDDI un nuevo token puede ser iniciado inmediatamente después de la transmisión del último paquete. La estación no tiene que esperar a recibir su propia transmisión alrededor del anillo antes de enviar un nuevo token, haciendo más eficiente el uso del ancho de banda del anillo.

Otra característica relevante de FDDI es la incorporación de un esquema prioritario. También conexiones que son rotas después de que el anillo ha sido puesto en operación son auto-reparadas a través de una operación interna, la cual reinicializa el anillo y en caso de encontrar una ruptura construye una nueva ruta.

Dada la velocidad de las comunicaciones utilizando fibra, FDDI deberá ser considerado en aplicaciones que requieren la transmisión de grandes capacidades de datos o un rápido tiempo de respuesta. FDDI fue diseñado como la columna vertebral de un servicio que soporta y enlaza muchos sistemas.

Las principales ventajas de FDDI sobre token ring y ethernet son debido al uso de fibras ópticas en lugar de cables de cobre. Las principales desventajas para utilizar FDDI son su alto costo, complejidad.

La figura 2.15 ilustra la operación básica de la tecnología FDDI.

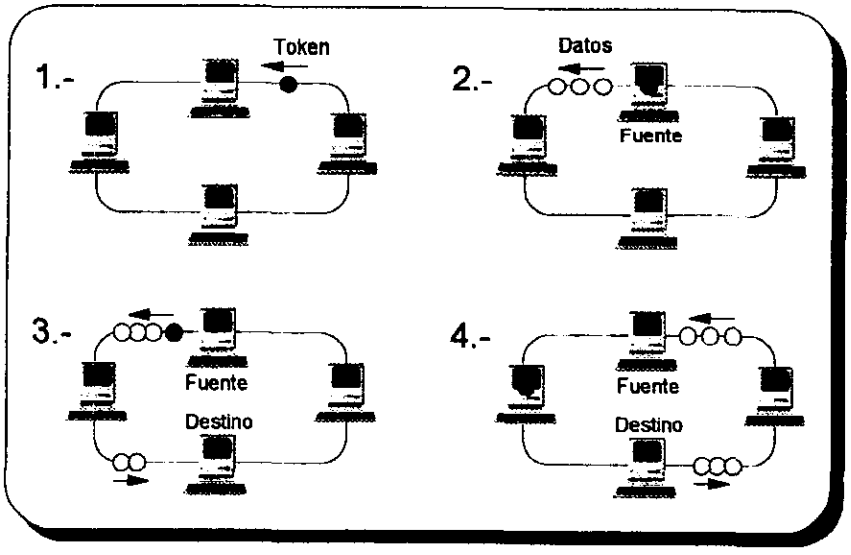


Figura 2.15 - Operación básica de FDDI

Concentradores, Puentes y Ruteadores

Los concentradores, los puentes y los ruteadores son los componentes básicos de toda red de comunicaciones digitales. Comenzaremos con los dispositivos básicos.

Los concentradores son repetidores multipuerto con capacidades de detección de colisiones y autosegmentación. Las señales que llegan al concentrador por cualquiera de sus puertos son automáticamente regeneradas y retransmitidas a los demás puertos del concentrador. El concentrador regenera los datos que llegan a el sin interpretar su contenido, por lo cual puede ser usado para redes tipo ethernet, IEEE 802.3 o cualquier otro protocolo. Los concentradores caen en la *capa física del modelo OSI, ofreciendo la conexión física para la red de* comunicaciones sin importar el medio. Existen concentradores para red delgada, red gruesa, par trenzado o fibra óptica. Los concentradores también son conocidos como cables

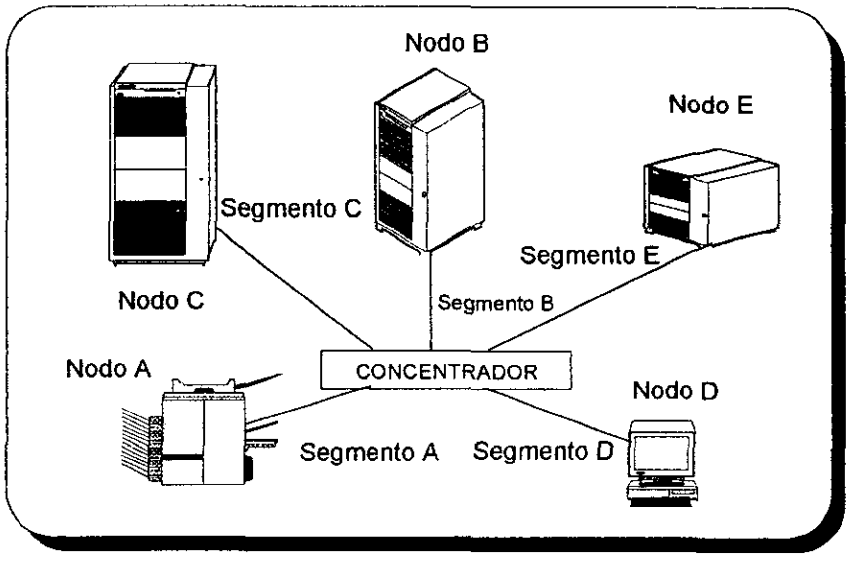


Figura 2.16 - Topología de red con concentradores

colapsados, ya que proporcionan el medio para extender la longitud de la red. En la figura 2.16 se muestra una topología típica construida con concentradores. Esta topología puede conceptualizarse como una estrella, pero electrónicamente es un canal. Un concentrador permite segmentar un puerto con fallas, esta característica permite que el resto de la red que no tiene problemas continúe trabajando. Otra característica de los concentradores es que son capaces de tener ligas redundantes, las cuales permiten asegurar un enlace.

Los concentradores más modernos ademas de ser retransmisores, también ofrecen capacidades de administración de red ofreciendo soporte a protocolos de manejo de red tales como el SNMP y estadísticas de paquetes.

Los puentes son equipos de comunicaciones que permiten la interconexión de varias redes, particionando el transito y asegurando los datos dentro de ellas. Estos equipos son capaces de proveer rutas redundantes al flujo de información entre redes. Los puentes operan dentro de la capa dos del modelo OSI. Ellos identifican a los nodos dentro de las redes usando lo que se conoce como direcciones de estación o MAC Address, *Medium Access Control Address*, las cuales son asingadas por el fabricante del equipo.

Cuando un puente recibe un paquete de datos, examina la dirección de estación origen y destino del paquete. Una tabla de direcciones es utilizada para determinar como procesar dicho paquete. La tabla de direcciones guarda la información necesaria para localizar los nodos y redes que se están comunicando. Cualquier cambio que ocurra en la red se ve reflejado con un cambio en la tabla. Si la tabla de direcciones muestran que la dirección destino y la dirección fuente pertenecen a una misma red, el puente descarta el paquete, pero si por el contrario muestran que son de distinta red el paquete es retransmitido a la red destino. Esta característica reduce el transito de paquetes innecesarios, haciendo la comunicación más confiable. En el caso de que el puente no sea capaz de determinar la red destino

de un paquete este retransmite el paquete a todas las redes (con excepción de la red origen). La mayoría de los puentes actuales construyen sus tablas de manera dinámica, es decir que "aprenden" de los paquetes que se transmiten entre las redes, para hacer esto, los puentes analizan la dirección fuente de cada paquete y lo relacionan con el puerto que lo recibe. En la figura 2.17 se muestra una topología típica donde se utiliza un puente para unir varias redes.

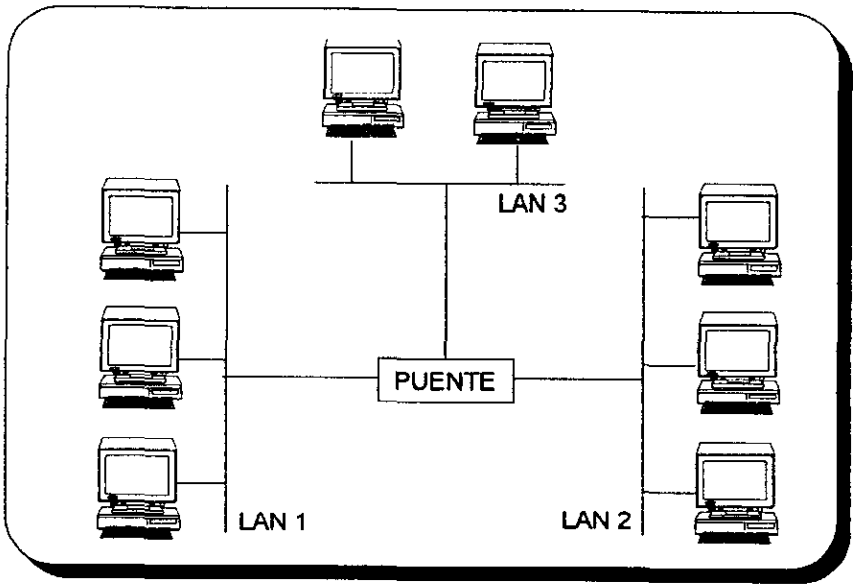


Figura 2.17 - Topología de red con puentes

Los puentes también tienen la capacidad de permitir rutas redundantes entre dos redes. Pensemos en un par de redes comunicadas entre si por dos puentes tal como se muestra en la figura 2.18. Si estos dos puentes solo trabajaran aprendiendo rutas y retransmitiendo mensajes y un nodo en alguna de las redes transmitiera un paquete cuya dirección destino es ignorada por cualquiera de los dos puentes, entonces ambos equipos retransmitirían el mensaje provocando un lazo *loop* infinito que después de un tiempo saturaría la red y la haría fallar. Cualquier red con lazos dentro de su topología tendría este problema. Para evitarlo nos tendríamos que asegurar que no se creara ningún lazo dentro de la topología,

pero en algunas ocasiones esto es deseable. Ligas paralelas entre redes con puentes pueden proveer redundancia para incrementar la tolerancia a fallas. El algoritmo *spanning tree* permite a una red contener lazos, asegurándose de que sólo exista una ruta activa entre cualquiera de dos nodos en dos redes distintas.

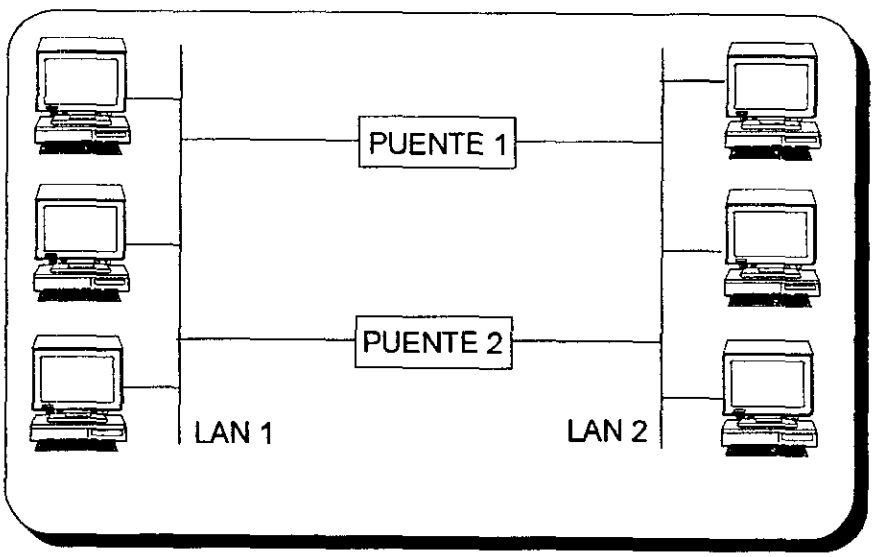


Figura 2.18 - Topología de red con rutas redundantes

Si la ruta activa entre las dos redes falla, el algoritmo detecta el problema y automáticamente reconfigura los equipos y activa una nueva ruta. El algoritmo trabaja de la siguiente manera:

- Un puente es seleccionado el la topología como la base. Todas las rutas activas son seleccionadas en base a este puente.
- Los puertos retransmisores de cada puente son seleccionados. Estos son los puertos por los cuales los paquetes son retransmitidos a las rutas activas. La topología de los puertos retransmisores se elige automáticamente, asegurándose que existe solo una ruta activa entre cualquiera de los nodos de la red.

- El resto de los puertos de los puentes son configurados como puertos de respaldo y no retransmitirán paquetes, a menos que se detecte una falla.
- Todas estas configuraciones se hacen de manera automática sin necesidad de intervención humana.

Los ruteadores son dispositivos de comunicaciones que trabajan dentro de la capa tres del modelo OSI, permitiendo la comunicación de diferentes protocolos de enlace de datos. Estos dispositivos son usados para interconectar distintas redes y tienen la capacidad de identificar nodos o segmentos dentro de cada red que interconectan. En redes de gran tamaño permiten determinar la mejor ruta entre dos distintas redes, minimizando el tráfico de comunicación entre redes. También pueden ser utilizados para establecer filtros entre redes que pueden ser usados para mantener estadísticas de utilización e incrementar la seguridad al aislar una red de otra.

En general los ruteadores permiten simplificar la administración de una red y pueden ser utilizados como herramientas para la detección de problemas. Algunos problemas tales como la excesiva transmisión de mensajes de identificación y errores de cableado no se propagan a través de los ruteadores permitiendo así incrementar la confiabilidad de la red. Otra manera en que estos dispositivos contribuyen a la confiabilidad de la red es que al permitir varias rutas entre dos redes la comunicación entre dos nodos no depende de una sola linea de comunicación. Para lograr esto los ruteadores utilizan algoritmos de ruteo tales como vector de distancia *distance-vector* y estado de concección *link-state*. Ademas estos dispositivos pueden trabajar con protocolos de administración de red tal como SNMP *Simple Network Manager Protocol*. En la figura 2.19 se muestra un mapa de red con varios ruteadores interconectando varias redes.

En esta topología seria necesario que al menos fallaran dos ruteadores para que la comunicación entre la red ETHER1 y ETHER2 se interrumpiera. Aunque en este

caso tendríamos que hacer un análisis de cual ruta es más eficiente, pero la comunicación es posible por más de una sola ruta con lo cual se garantiza el servicio.

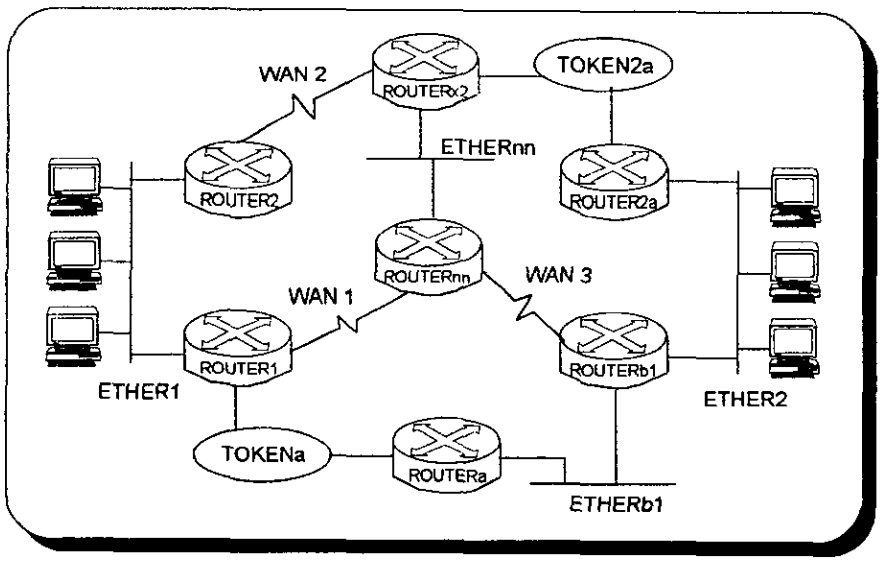


Figura 2.19 - Topología de red con ruteadores

UNIDADES CENTRALES DE PROCESAMIENTO EN SISTEMAS ABIERTOS

En la actualidad existe una gran variedad de equipos cómputo y el mercado de los sistemas abiertos es cada vez más grande. Las ventajas que ofrecen las tecnologías abiertas son mayores a las que ofrecen los sistemas propietarios, sin embargo también tienen desventajas. La diversidad de compañías involucradas en el desarrollo de los sistemas abiertos provocan que las tecnologías desarrolladas se obsoleticen rápidamente o jamas sean tomadas como un estándar. A pesar de los esfuerzos de la industria por coordinar sus desarrollos y estandarizar la tecnología emergente, lo que hoy se declara como un estándar mañana puede no cumplir con las expectativas y declararse como tecnología no compatible. Hasta el momento la tecnología que hemos contemplado en el capítulo dos de este trabajo ha comprobado su compatibilidad y parece que permanecerá como un estándar por varios años más. Pero en cuanto a los procesadores (CPU's) y sistemas operativos la situación cambia.

El surgimiento de los sistemas abiertos durante la década de los 80's y lo que va de los 90's ha sido el desarrollo más significativo en la evolución de la tecnología de la información y procesamiento de datos. Como resultado de esto, varios grupos de empresas y usuarios se unieron para monitorear y estructurar las actividades de los distintos proveedores de sistemas abiertos. Este grupo se conoce como The Open Group. Este grupo esta formado por X/Open Company y Open Software Foundation como principales fuentes de estandarización para sistemas y desarrollo de tecnología.

Las principales compañías detrás de The Open Group son Fujitsu, Hitachi, Novel, Digital, Hewlett-Packard, IBM, NCR, Simens-Nixdorf, Sun, Bull, SCO y Silicon Graphics. En la figura 2.20 se muestran los logotipos de estas empresas.

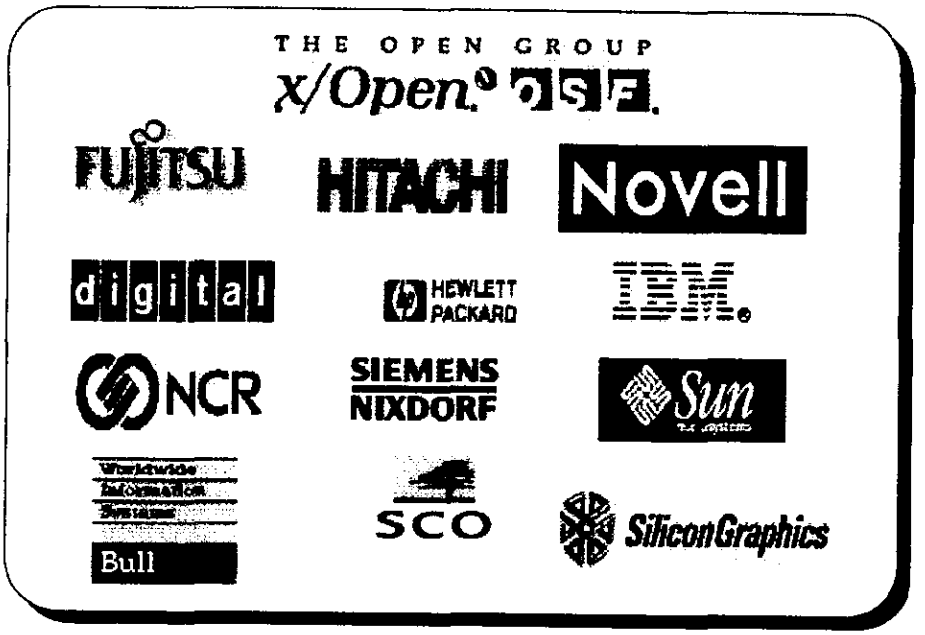


Figura 2.20 - "The Open Group"

Para poder asegurar que nuestra propuesta tenga permanencia y sean aplicables a los sistemas abiertos, lo que nosotros consideraremos como un CPU de sistemas abierto será cualquier equipo de procesamiento que cumpla con los estándares definidos por The Open Group. Estos equipos serán aquellos, que sin importar la arquitectura sean capaces de operar con el sistema operativo que The Open Group ha denominado UNIX95. Ejemplos de este sistema operativo son el HP-UX 10.XX de Hewlett-Packard y el Solaris de Sun microsystems. Como ejemplos de estos CPU's podemos nombrar los equipos serie Alfa de Digital, la serie 9000 de Hewlett-Packard o cualquier equipo de SiliconGraphics.

3
SISTEMAS DE ALTA DISPONIBILIDAD PROPUESTOS

El objetivo de este capítulo es presentar configuraciones de "hardware", que permitan mantener un nivel de disponibilidad en los sistemas de cómputo, el cual dependerá de las necesidades del usuario. Para poder entender las implicaciones del ambiente de alta disponibilidad primero tenemos que definir los niveles de disponibilidad.

El término disponibilidad describe a un sistema que provee un nivel específico de servicio requerido. En computación se entiende por disponibilidad el período de tiempo en el cual se puede hacer uso de los servicios, cualquier interrupción ya sea planeada o no, la llamaremos en adelante sistema fuera de línea. De acuerdo a lo anterior se definen cuatro niveles de disponibilidad en orden jerárquico, el nivel más bajo corresponde al sistema más susceptible a interrupciones y el nivel más alto a un sistema que teóricamente no tiene interrupciones. La figura 3.1 muestra los niveles de disponibilidad.

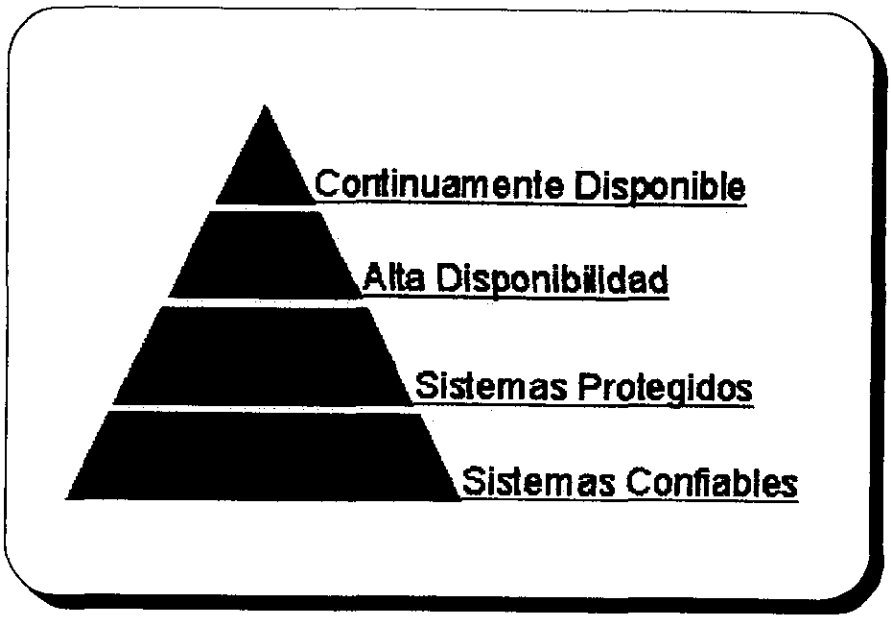


Figura 3.1 - Niveles de Disponibilidad

SISTEMAS CONFIABLES

Es el nivel mas bajo de disponibilidad y se considera que cualquier equipo de cómputo actual, cumple con este nivel. Podemos tomar de ejemplo una computadora personal en donde el tiempo entre fallas ("mean time between failure" MTBF) es en promedio de 45,000 horas, lo cual significa que se espera que el equipo trabaje continuamente durante 5 años antes de experimentar una falla. El dato se obtiene con base en estadísticas y pruebas realizadas en laboratorio. Todos los dispositivos de cómputo como tarjetas, discos, monitores, etc. tienen un MTBF, que al combinarse en un sistema, hacen que el sistema resultante tenga un MTBF menor que cualquiera de sus componentes. Es así que un equipo de cómputo con más componentes de "hardware" tendrá un MTBF menor. La tabla 3.1 muestra la confiabilidad típica de algunos dispositivos de cómputo.

Dispositivo	MTBF
Unidad de disco óptico	150,000
Unidad de cinta	175,200
Unidad de disco duro	400,000
Impresora láser	87,600
Interface FDDI	359,977
CPU	81,854

Tabla 3.1 - Ejemplos de MTBF

Considerando un sistema conformado por los dispositivos de la tabla anterior tendríamos un sistema con un MTBF de aproximadamente 10,000 horas, lo cual es poco más de un año. En este caso no se cuenta con ningún elemento redundante y la disponibilidad se confía en la calidad del equipo. Cualquier falla interrumpirá los servicios.

SISTEMAS PROTEGIDOS

Es un sistema cuya confiabilidad depende de la redundancia en componentes y el tiempo de restablecimiento va de 5 a 120 minutos. Un sistema protegido es aplicable en ambientes que toleran un tiempo de recuperación de hasta 2 horas con un mínimo impacto en la operación.

En algunos casos el ambiente de operación puede prescindir del sistema durante dos horas, tiempo suficiente para cambiar algún componente dañado pero no lo suficiente para recuperar los datos. El enfoque principal del presente nivel es la protección de la información utilizando redundancia. El uso de discos redundantes configurados en espejo o arreglos de discos nos permiten proteger elementos claves que permitirán cumplir con el nivel de disponibilidad requerido por los usuarios.

Este nivel de disponibilidad no evita la interrupción del servicio, pero asegura la información. Esto en última instancia permite restablecer la operación en un mínimo de tiempo, pero para lograr esto hay que considerar el tiempo de respuesta del personal de soporte técnico y la disponibilidad de partes. El tiempo para cambiar un componente en un sistema de cómputo se ha reducido de manera significativa en los equipos actuales, pero si no se cuenta con ese componente en el momento de la falla la interrupción del servicio podría incrementarse considerablemente.

En la figura 3.2 se muestra un sistema protegido en donde los datos residen en un arreglo de discos. Este proporciona la redundancia necesaria para asegurar la integridad de la información y la disponibilidad del sistema en caso de una falla en disco. En este caso si un componente distinto al disco falla el servicio se suspendería.

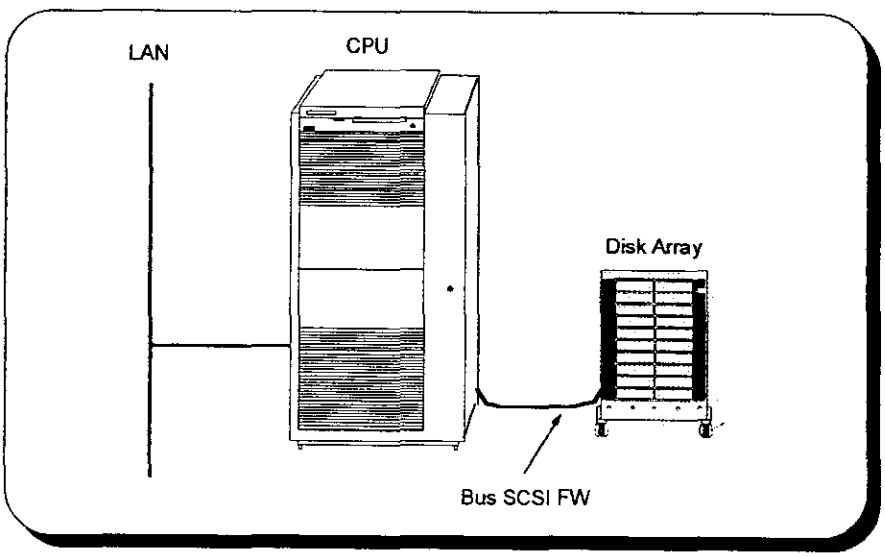


Figura 3.2 - Sistema protegido

SISTEMAS ALTAMENTE DISPONIBLES

Es un sistema que posee la habilidad de continuar procesando la aplicación aun cuando exista un evento de falla en uno de sus componentes. Se caracteriza por evitar o reducir la pérdida de los servicios, mediante una rápida recuperación de las aplicaciones.

Al presentarse una falla el sistema es capaz de recuperar los servicios en un lapso de tiempo menor a 5 minutos y en algunos casos puede ser menor a un minuto o incluso transparente. La recuperación del ambiente normal de operación algunas veces implicará una pequeña interrupción en la aplicación con un mínimo impacto. En un sistema altamente disponible la redundancia en los elementos principales es la clave para cubrir las necesidades de los ambientes críticos.

En éste caso el sistema reduce los tiempos fuera de línea causados por fallas y, dependiendo de la configuración utilizada se diminuye también los tiempos causados por actividades planeadas, tales como mantenimientos y mejoras al equipo. De un sistema de alta disponibilidad se espera que los servicios estén disponibles el 99.00 % del tiempo de utilidad. Para lograr este nivel de disponibilidad se necesita la redundancia en casi todos los componentes del sistema tales como el procesador, los discos, los buses de datos, los enlaces a la red y fuentes de alimentación. En la figura 3.3 se ilustra un sistema con todos estos componentes duplicados.

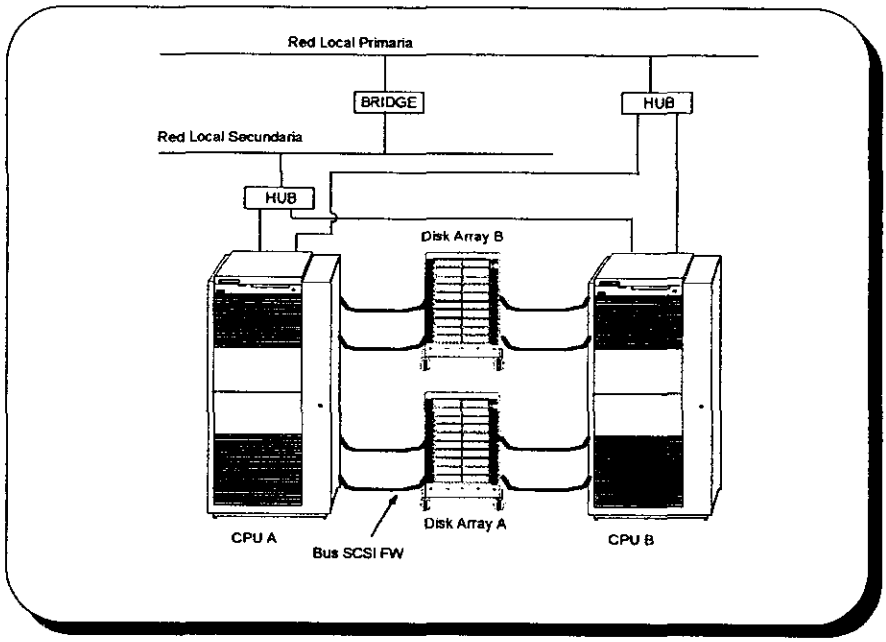


Figura 3.3 - Sistema altamente disponible

SISTEMAS CONTINUAMENTE DISPONIBLES

Es un sistema que presenta transparencia a la recuperación de un evento de falla, está disponible 24 horas al día los 7 días de la semana y el tiempo planeado fuera de servicio se elimina. Significa que no habrá pérdida de los servicios. La característica principal es que es tolerante a cualquier falla de hardware.

Existen ambientes de cómputo donde el 99.00 % de disponibilidad no es suficiente ya que significa que estará 3.6 días al año, fuera de servicio por fallas o por mantenimiento y se requiere del 100% de disponibilidad.

Un sistema continuamente disponible corresponde a arquitecturas especializadas de hardware que detectan la falla de cualquier componente y de forma instantánea hace el cambio a un elemento redundante. Un ejemplo de este nivel de

disponibilidad son los sistemas diseñados por "Tandem", en la figura 3.4 se muestra uno de estos equipos.

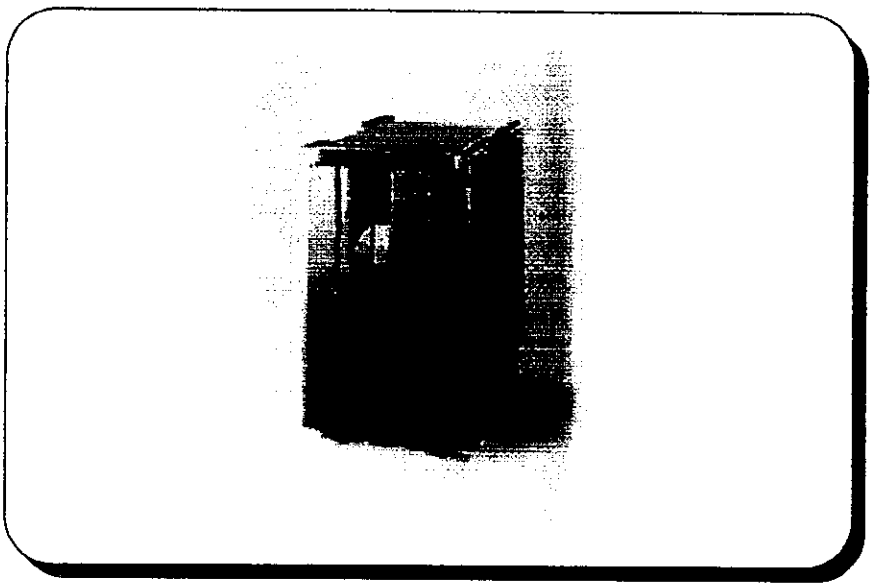


Figura 3.4 - Sistema continuamente disponible

La diferencia entre un sistema continuamente disponible y uno altamente disponible es que el primero esta diseñado para evitar cualquier falla, mientras que el segundo acepta una mínima interrupción de los servicios.

Aun cuando éstos sistemas nos proporcionan un 100% de disponibilidad no son muy utilizados, primeramente por su alto costo y por que existen muy pocas aplicaciones especializadas para estos equipos. Es por esto que para aquellos ambientes en donde se busca la mayor disponibilidad de los servicios pero no se cuenta con altos recursos económicos, un sistema de alta disponibilidad es el adecuado.

PUNTOS DE FALLA

Un punto de falla es cada uno de los elementos de "hardware" y "software" que por mal funcionamiento o daño pueden interrumpir el servicio a los usuarios de un sistema de cómputo. En términos generales, un componente que no se encuentra respaldado por un redundante, es considerado un punto de falla.

La tabla 3.3 muestra puntos de falla que pueden causar la interrupción del servicio, así como las consecuencias y una estrategia para eliminarlos.

Componente	**Consecuencias de la falla del componente.**	**Como es eliminado el punto de falla.**
CPU único.	El servicio se pierde hasta que el procesador sea reparado.	Proveer un CPU de respaldo.
Redes simples (LAN).	La conectividad del cliente se pierde.	Instalación de interfaces de red y subredes redundantes.
Interface de red única.	La conectividad con el cliente se pierde.	Instalación de tarjetas de red redundantes.
Disco de sistema operativo único.	El servicio se interrumpe hasta que el disco sea reemplazado.	Utilización de un disco de sistema operativo redundante.
Disco de datos único	Se pierden los datos.	Utilización de discos espejo, o uso de arreglos de discos en modo de protección de datos.
Fuente de energía	El servicio se pierde hasta que la alimentación sea restablecidas.	Uso de fuentes de poder redundantes y empleo de la tecnología de fuente de energía ininterrumpible.
Tarjeta controladora de discos.	El servicio se pierde hasta que la tarjeta sea reemplazada.	Utilización de tarjetas controladoras duales o redundantes con doble trayectoria de acceso a los discos.
Programas de Aplicación	El servicio se pierde hasta que la aplicación se restaure.	Proveer la capacidad de autorestablecimiento de la aplicación.
Errores humanos	El servicio se pierde hasta que los errores sean corregidos.	Automatización de la operación tanto como sea posible.

Tabla 3.3 - Eliminación de puntos de falla

Estas estrategias son muy generales, el procedimiento exacto dependerá de cada sistema y aplicación en especifico. En el caso del CPU, discos e interfaces es suficiente con conectar dispositivos redundantes y dependiendo de lo crítico de la operación se podrá utilizar "software" que haga el cambio automáticamente en el componente que presente falla.

En el caso de una falla en la red aun cuando el equipo este funcionando se considerara el servicio fuera de linea ya que quedará aislado de los usuarios. Para eliminar este punto de falla la estrategia es ligeramente distinta ya que no se coloca una red redundante sino se buscan trayectorias alternas que permitan la comunicación entre el sistema y los usuarios.

Análisis de los sistemas de cómputo en relación a su confiabilidad

Para poder definir un sistema de cómputo es importante conocer las necesidades operativas a satisfacer. Es necesario un análisis del número de usuarios que entran al sistema y de la cantidad de información que se debe de guardar en disco, además de la disponibilidad mínima del sistema requerida. Este análisis no está contemplado como parte de este trabajo, pero para efectos prácticos podemos consideraremos una configuración de sistema, y de ahí partir para analizar cada uno de los niveles de disponibilidad.

Consideremos un equipo Hewlett Packard con la siguiente configuración:
CPU:
Equipo HP 9000/K220 modelo A3453A con:
2 procesadores PA-RISC con 1 MB de memoria cache cada uno.
4 ranuras de I/O tipo HP-PB, con capacidad de 2 interfaces SCSI-FW y 2 SCSI-SE.
1 Interface LAN 802.3, conexión AUI o BNC.
1 Multipuerto con 8 conexiones RS-232C.
1 Módem interno.
1 CD-ROM de 650 MB.
1 Puerto para teclado PS/2.
256 MB de memoria RAM.
4 GB de disco interno modelo A3353A.

Discos externos:
Mini-Torre de discos Modelo C5264T con:
1 Fuente de poder única de 350W.
2 Discos de 4 GB modelo A3353A.

En la figura 3.5 se ilustra un equipo como el que se describió anteriormente. Este es un sistema mediano para una carga de datos que va de bajo a moderado. En la figura 3.6 se ilustran los discos de este sistema.

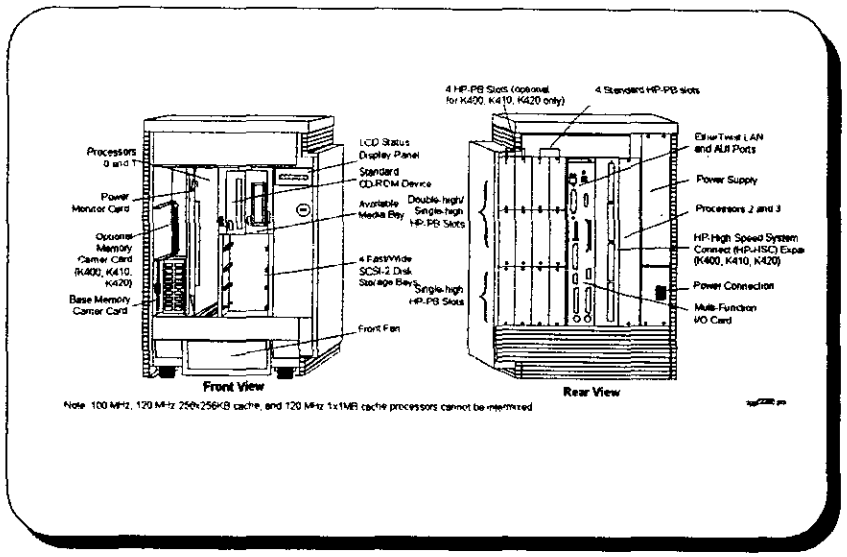


Figura 3.5 - Servidor clase K de Hewlett Packard

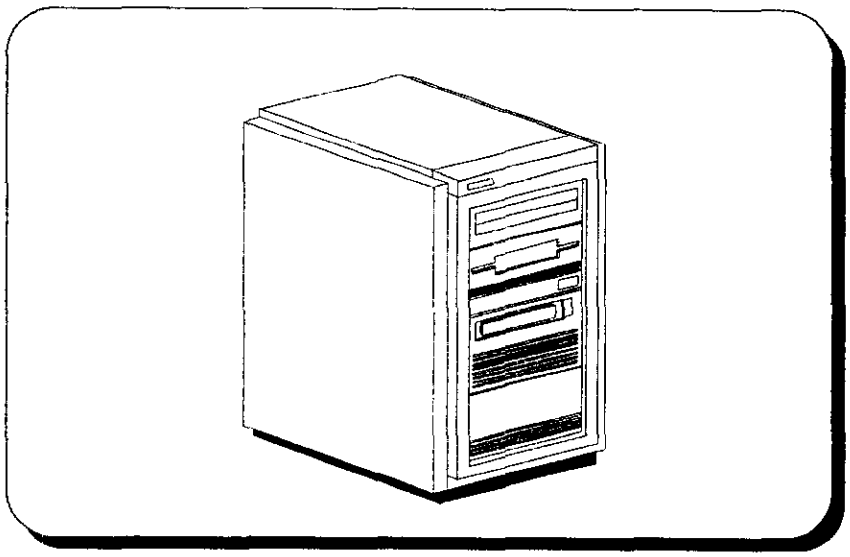


Figura 3.6 - Subsistema de disco HP 6000 SCSI

Sistema confiable

Dado que se esta planteando un sistema confiable, el escenario con el cual partirá es un sistema que corre una aplicación, con 260 usuarios y 6 GB de información en disco. El usuario requiere que el sistema este disponible el 90% del tiempo, trabajando 24 horas al día. Dada la importancia de la aplicación el tiempo fuera de línea sin interrupción soportado por el usuario sin sufrir perdidas es de 4 horas, esto quiere decir que MTTR debe ser menor a 4 horas. Además de esto la situación de falla no debe de repetirse más de una vez al año, es decir el AFR es igual a 1.

Ahora es necesario verificar que los parámetros del equipo tales como la disponibilidad y el MTBF cumplan con los requerimientos del escenario planteado. De acuerdo con esto el usuario necesita que el equipo este disponible 90% del tiempo con ciclo de trabajo de 24 horas diarias. Esto implica que el equipo requiere estar disponible un mínimo de 7884 hrs. en un año. Lo cual nos deja 876 hrs. para tiempo fuera de línea que puede ser consumido en labores de mantenimiento y fallas. Dado que un evento de falla no debe de repetirse más de una vez al año el MTBF mínimo requerido por el usuario es de 7884 hrs.

El fabricante proporciona los AFR de los siguientes componentes del equipo:
CPU K220:
AFR (global, sin incluir discos)%= 46.2%
Discos C5264T:
Fuente de poder AFR%=16.8%
Mecanismos C3353A AFR%=11.11%

El MTBF del CPU seria de:
MTBF = $\frac{8760}{0.462} = 18961.0316$ horas = 2.16 años

El MTBF de cada uno de los discos seria:

$$\text{MTBF} = \frac{8760}{0.111} = 78918.02 \text{ horas} = 9.00 \text{ años}$$

El MTBF de la fuente de poder de los discos seria:

$$\text{MTBF} = \frac{8760}{0.168} = 52142 \text{ horas} = 5.95 \text{ años}$$

El MTBF del sistema completo seria:

$$\text{MTBF} = \frac{1}{\frac{1}{18961.03} + \frac{3}{78918.92} + \frac{1}{52142.85}}$$

MTBF = 9096.57 horas = 1.03 años

Lo cual significa que el sistema podría trabajar poco más de un año antes de experimentar una falla. En principio esto indicaría que el sistema estaría disponible 100% del tiempo, pero siempre hay que considerar un determinado tiempo para las actividades que requieren que el sistema esté fuera de línea tales como los mantenimientos preventivos. Si observamos el MTBF del sistema en global, resulta que es mayor al MTBF requerido por el usuario, por lo cual este parámetro es satisfecho por el equipo.

El MTTR de cada uno de los componentes del sistema varía de uno a otro, pero el fabricante también especifica que en el peor de los casos el Hardware del sistema tiene un MTTR menor a 1.5 Hrs. Esto implica que el tiempo máximo de respuesta del personal de soporte debe ser menor a 2.5 hrs. con el fin de satisfacer la restricción del usuario, en el sentido de que el tiempo máximo fuera de línea sea menor a 4 horas.

De lo anterior podemos concluir que este equipo es suficiente para satisfacer las necesidades del usuario. Sin embargo este análisis es muy teórico y parte exclusivamente de la información proporcionada por el fabricante. No siempre se cumple en la vida real, ya que el equipo esta sometido a otros factores no considerados, como pueden ser condiciones ambientales, transportación, trato del

usuario, etcetera. Por lo cual para asegurar la disponibilidad del sistema es necesario utilizar algún nivel de redundancia.

En la figura 3.7 se documenta la configuración propuesta para este sistema confiable. El equipo cuenta únicamente con un bus SCSI-FW con el cual se comunican los discos internos y externos con el CPU. Cuenta solo con una interface de red con la cual se comunica con los usuarios. Hay que destacar que aunque el sistema es multiprocesador, este no es capaz de tolerar fallas en ninguno de los procesadores. En la configuración propuesta la fuente de poder del CPU (la cual alimenta a los procesadores tarjetas y discos internos) es independiente de la fuente de poder de los discos externos, pero la alimentación eléctrica es común a ambas. Este sistemas no podría continuar funcionando si fallara cualquiera de sus componentes.

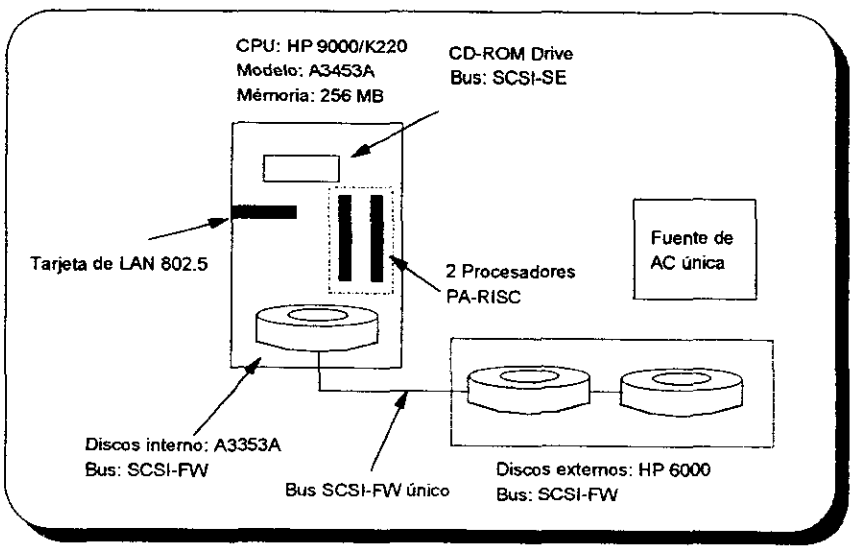


Figura 3.7 - Sistema confiable propuesto

Sistema protegido

En un sistema protegido el tiempo de recuperación que se puede obtener es de 2 horas gracias a la redundancia de sus componentes. Se considera principalmente que se tiene un sistema protegido cuando la información contenida en los discos no se verá afectada en caso de fallar alguno de los componentes. En los sistemas actuales si llega a dañarse cualquier otro elemento del sistema, y teniendo la refacción correspondiente, es posible realizar el cambio necesario en el tiempo estipulado, permitiendo que se reanude la operación del sistema inmediatamente.

Por el contrario, serían los discos el mayor obstáculo para lograr este objetivo debido a que el tiempo de reparación cuando se requiere cambiar algún mecanismo involucra no solo el cambio físico, además es necesario reconfigurarlo en el sistema y, sobretodo, hay que considerar el tiempo que tomará reinstalar la información contenida originalmente en el disco. Si se toma en cuenta que existen en el mercado discos desde 1 hasta 9 GB, y que los dispositivos de almacenamiento externo pueden bajar un respaldo a una velocidad de transferencia de entre 0.5 y 1.0 Mb/seg, entonces un disco de 1 GB podría recuperar su información en 17 ó 34 minutos, pero un disco de 9 GB tardaría entre 2.5 y 5 horas.

Es entonces que para el escenario que estamos presentando se encuentran como puntos de falla que afectan directamente el tiempo de recuperación a los discos externos y el disco interno.

Por lo general en el disco interno reside el sistema operativo, las aplicaciones del cliente y el manejador de base de datos. Si llegara a dañarse este disco el sistema estará fuera de servicio hasta que fuera reemplazado el dispositivo, reinstalar ademas el software, y recuperar de los respaldos la información adicional que contenía el disco original.

En el caso de los discos externos, en estos se almacenan las bases de datos. Al fallar alguno de éstos, el sistema seguirá en funcionamiento, pero la información de los usuarios no estará disponible. Después de ser reemplazado el disco, se recuperará del último respaldo la información correspondiente. Sin embargo, ésta no contendrá todas las actualizaciones o modificaciones hechas por los usuarios después de que se hizo ese último respaldo, retrasando aún más la correcta operación del sistema mientras se actualiza dicha información.

Para el escenario que se esta considerando se requiere entonces un cambio en las necesidades de operación del sistema. Éste debe estar disponible 95% del tiempo, trabajando 24 horas al día. El máximo tiempo que puede estar el sistema fuera de servicio sin provocar perdidas es de 2 horas, es decir, se requiere un MTTR menor a 2.0 horas. Ademas de que no deben ocurrir fallas en general más que una vez al año, se necesita un AFR igual a 1.

Para evitar los inconvenientes que ocasiona el tener el sistema fuera de servicio en tanto se corrige la falla y, sobretodo mientras se recupera la información, proponemos se sustituya la mini-torre con discos externos y el disco interno por un arreglo de discos A3231A, también de Hewllet Packard, que puede tener hasta 10 discos de 2.1, 4.2 ó 9 GB, utilizaremos discos de 2.1 GB.

Nuevo disco externo:
Disk Array Modelo A3231A con:
2 Fuentes de poder A3239A
1 Batería de respaldo A3332A
2 Procesadores de almacenamiento A3236A
8 Mecanismos A3240A

La figura 3.8 ilustra el arreglo de disco al que se hace referencia.

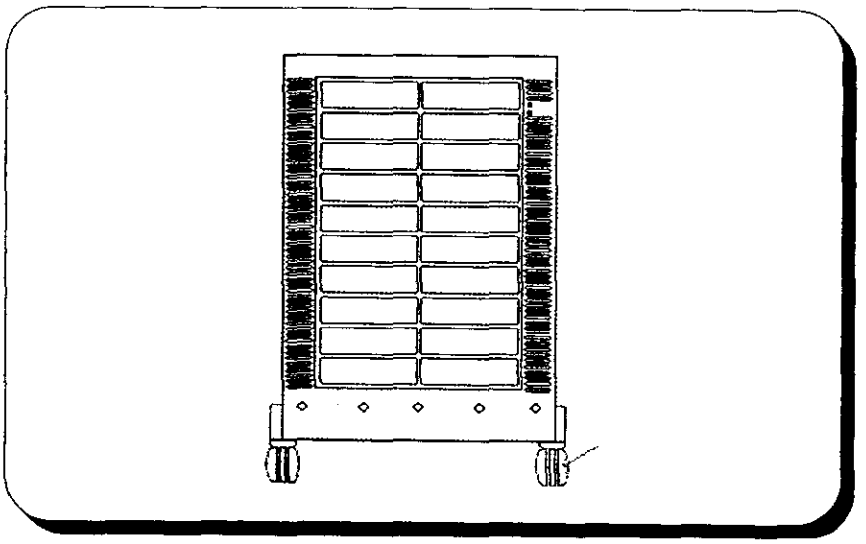


Figura 3.8 - Arreglo de disco

Considerando que el sistema debe estar disponible 95% del tiempo, 24 horas al día. En un año el equipo debe trabajar 8322 horas, dejando para actividades propias de mantenimiento y corrección de fallas sólo 438 horas. Si se requiere además que los componentes o dispositivos no fallen más de una vez al año, entonces el MTBF debe ser mayor o igual a 8322 hrs.

De la información que proporciona el fabricante de los componentes, tenemos los siguientes AFR y MTTR:

CPU K220:
MTTR = 1.5 hrs.
AFR %= 46.2%

Arreglo de discos A3231A:
MTTR = 0.5 hrs.

Arreglo de discos sin mecanismos AFR% = 29%

Grupo de discos A3240A AFR% = 3.0%

De los cuales podemos obtener entonces el tiempo promedio entre fallas de los componentes del sistema.

$\frac{8760}{0.462} = 18961.0316$El MTBF del CPU es:

MTBF = horas = 2.16 años

El MTBF de la electrónica del arreglo de discos sin considerar discos es:

MTBF Arreglo = $\frac{8760}{0.29}$ = 30206.90 horas = 3.29 años

El MTBF del grupo de discos del arreglo es:

MTBF Discos arreglo = $\frac{8760}{0.03}$ = 292000 horas = 33.33 años

Para el sistema completo tendremos el siguiente MTBF:

$$\text{MTBF sistema} = \frac{1}{\frac{1}{18961.04}+\frac{1}{30206.90}+\frac{1}{292000}}$$

MTBF sistema = 8830.65 Horas = 1.01 años

Podemos observar que el equipo propuesto cumple con los requisitos de no repetir un evento de falla en períodos menores a 1 año, tanto en forma global como para cada componente del sistema, al tener un MTBF mayor o igual al límite estipulado. Además, el MTTR de los dispositivos es siempre menor a 2.0 horas, por lo tanto se puede asegurar que es posible reemplazar cualquier componente del sistema cumpliendo con las expectativas del usuario.

Hablando del Arreglo de Discos, con éste podemos proteger la información gracias a que utiliza configuración RAID, en particular el RAID 5 que permite tener la información disponible aún cuando falle alguno de los discos que la integran, puesto que la información y la paridad de información se encuentran repartidas en todos los discos.

Por lo tanto, en caso de fallar alguno de ellos es posible recuperar la información en base a los restantes.

Al configurar los discos en el modelo propuesto en RAID 5 se pueden agrupar de 3 a 5 discos en Unidades Lógicas, utilizando para fines de protección de datos o paridad de información el equivalente al espacio de uno de los discos. Se configurarían entonces 3 discos de 2.1 GB para reemplazar al actual disco interno de la computadora HP 9000/K220, es decir 4.2 GB de espacio útil, equivalente al inicialmente propuesto. Los restantes 5 discos serán configurados para alojar la base de datos, con 8.4 GB de espacio disponible para información del usuario, aproximadamente igual a la configuración inicial. El sistema operativo reconocerá entonces únicamente 2 unidades de 4.2 y 8.4 GB totales.

Si llega a fallar cualquiera de los discos de alguno de las unidades el sistema podrá seguir trabajando sin interrupción, ya que la información se recuperaría de los elementos restantes. Posteriormente se podrá realizar el cambio del elemento defectuoso e iniciar el procedimiento de recuperación de información para el disco que haya sido reemplazado. Además, el reemplazo debe realizarse lo más pronto posible ya que si llegara a fallar otro disco de la misma unidad, ya no podría recuperarse ninguna información.

Con el modelo de arreglo de disco al que se hace referencia se puede obtener aún mayor redundancia de los componentes, por ejemplo, posee una tarjeta controladora de discos extra que permitirá conectar al sistema el arreglo de discos a través de una segunda interface, eliminando como puntos sencillos de falla la interface con el sistema. Sólo necesitará una tarjeta y cable scsi/fw extra en el sistema. También puede adicionar una fuente de poder extra para dar redundancia a este tipo de elementos, permitiendo trabajar aún cuando se llegue a dañar alguna de las fuentes de poder. Cuenta además con un ventilador adicional que proveerá de la suficiente capacidad de enfriamiento a los componentes internos aún si deja de trabajar alguno de ellos.

En lo que respecta a la energía eléctrica, además de que es sumamente recomendable el uso de una UPS para permitir que el sistema continúe trabajando por un tiempo razonable en casos de ausencia de energía eléctrica comercial, ya sea para permitir apagar el sistema normalmente o para continuar la operación por tiempo indefinido. Este modelo de arreglo de discos posee una batería de respaldo que permite escribir la información contenida en la memoria cache en los mecanismos mientras entra en operación la UPS.

La figura 3.9 ilustra las características de la nueva configuración, que permiten observar la redundancia de los componentes que permiten que este esquema se pueda considerar un sistema protegido. Cuenta con dos buses SCSI-FW para comunicarse con el arreglo de discos, que serán las interfaces A y B. En el arreglo se observan las dos controladoras, las fuentes de poder redundantes, los ventiladores y la batería de respaldo.

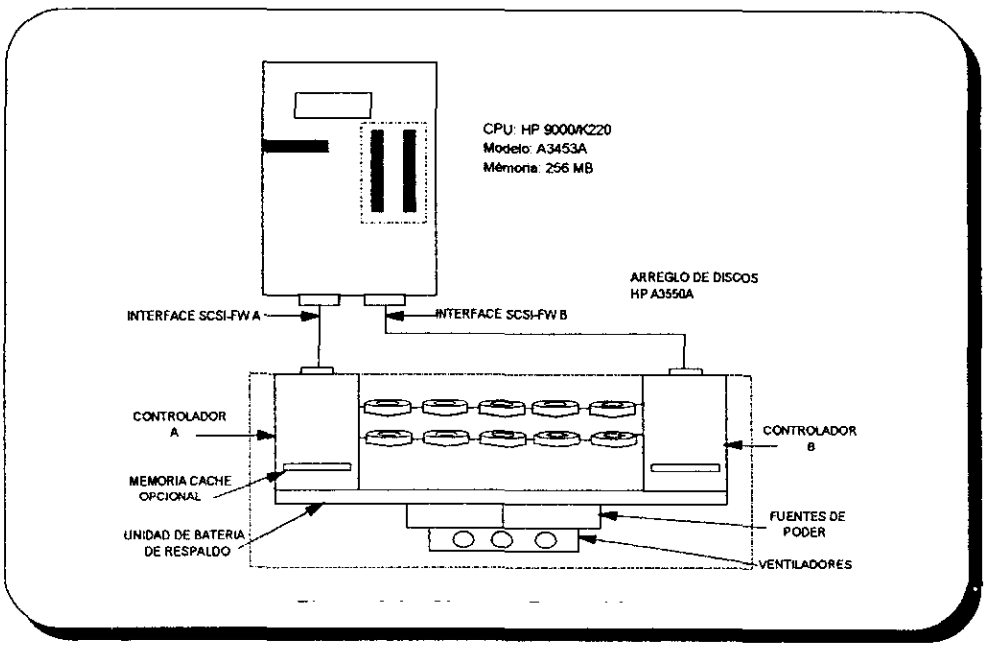


Figura 3.9 - Sistema protegido

Un arreglo de discos se puede considerar como una solución adecuada para un sistema protegido puesto que es posible tener un tiempo de recuperación menor a 2.0 horas, considerado suficiente para reemplazar cualquier elemento defectuoso del sistema. Permitiendo además continuar con la operación del sistema si llegara a fallar alguno de los discos, ya que la información contenida en el arreglo siempre estará disponible para los usuarios o procesos del sistema.

En esta configuración solamente se asegura la integridad de la información. Si se quiere proteger la operación de otro tipo de interrupciones es necesario utilizar sistemas de alta disponibilidad o continuamente disponibles.

Sistema de alta disponibilidad

En esta ocasión se partirá del siguiente escenario. El sistema tiene que estar disponible el 99% del tiempo, trabajando 24 horas al día. Dada la importancia de la aplicación el tiempo fuera de línea no debe de ser mayor a 30 minutos. La reparación de cualquier componente es de 1 hora, esto quiere decir que MTTR debe ser menor a 1 horas. Ademas de esto la situación de falla no debe de repetirse más de una vez al año, es decir el AFR es igual a 1.

Para la configuración se utilizara al proveedor Hewlett Packard nuevamente y en base al escenario se definirá lo siguiente. En primer término el análisis presentado en el ambiente de un Sistema Confiable, nos demuestra que un equipo puede presentar fallas que no permitirían alcanzar los objetivos planteados de alta disponibilidad, por lo tanto la configuración se verá reforzada para poder lograrlo. La base principal para alcanzar el objetivo de un 99% de disponibilidad es mediante la utilización de sistemas con elementos redundantes no solo en discos y tarjetas ahora también se utilizará un equipo redundante, un CPU que esté en espera de sustituir al primario o principal en caso de falla.

Para el ejemplo se utilizarán dos equipos de similares características tanto en memoria como en procesadores. La razón es para lograr el mismo rendimiento en las aplicaciones corriendo en uno u otro sistema. Nuevamente se utilizará la configuración del equipo planteado para el Sistema Confiable.

CPU:

Equipo HP 9000/K220 modelo A3453A con:

2 procesadores PA-RISC con 1 MB de memoria cache cada uno.

4 ranuras de I/O tipo HP-PB, con capacidad de 2 interfaces SCSI-FW y 2 SCSI-SE.

2 Interfaces LAN 802.3, conexión AUI o BNC.

1 Multipuerto con 8 conexiones RS-232C.

1 Módem interno.

1 CD-ROM de 650 MB.

1 Puerto para teclado PS/2.

256 MB de memoria RAM.

4 GB de disco interno modelo A3353A.

La figura 3.10 muestra el CPU del cual hablamos.

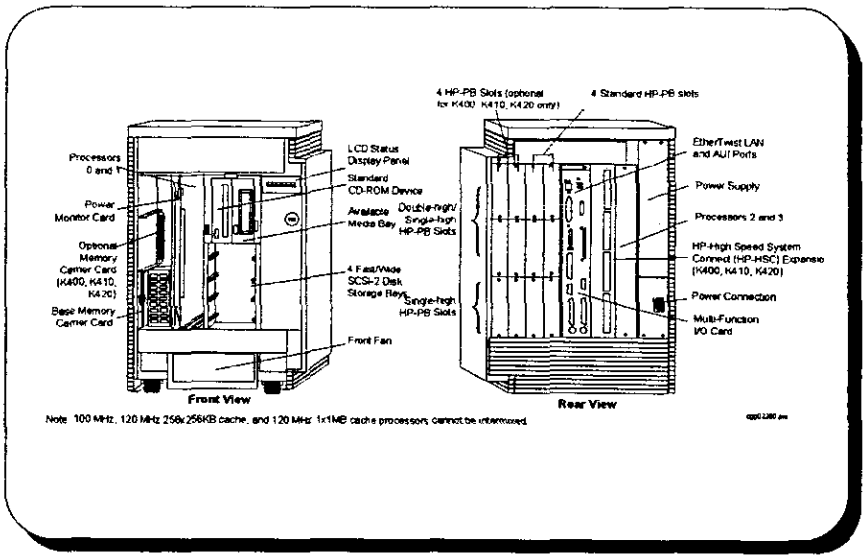


Figura 3.10 - Servidor clase K de Hewlett Packard

Discos externos:

Mini-Torre de discos Modelo C5264T con:

1 Fuente de poder única de 350W.

2 Discos de 4 GB modelo A3353A.

En la figura 3.11 se ilustran los discos utilizados en este sistema.

El análisis de la propuesta del caso de Sistemas Confiables indica que se tiene un equipo con un MTBF de 1.03 años. Para un Sistema de Alta Disponibilidad cuando

la falla se presente los resultados podrían ser desastrosos por las pérdidas que implicaría tener la aplicaciones fuera de linea. Es por ello que se utilizarán elementos redundantes y en vez de tener un subsistema de discos para los datos se tendrán dos. El primero manejando los datos y su copia en espejo quedará en el segundo. Ambos subsistemas serán controlados por diferentes interfaces SCSI FW para asegurar la continuidad de la información en caso de que una tarjeta falle. Se protege el disco de sistema operativo de una forma similar al instalar una cuarta tarjeta SCSI FW con un subsistema Mini-Torre de discos Modelo C5264T con un disco de 4 GB. Para recuperar la comunicación en caso de fallar una tarjeta de red se utiliza nuevamente un elemento redundante, entre ambas se tendrá un medio de comunicación que permitirá el viaje de información en cualquier de ellas.

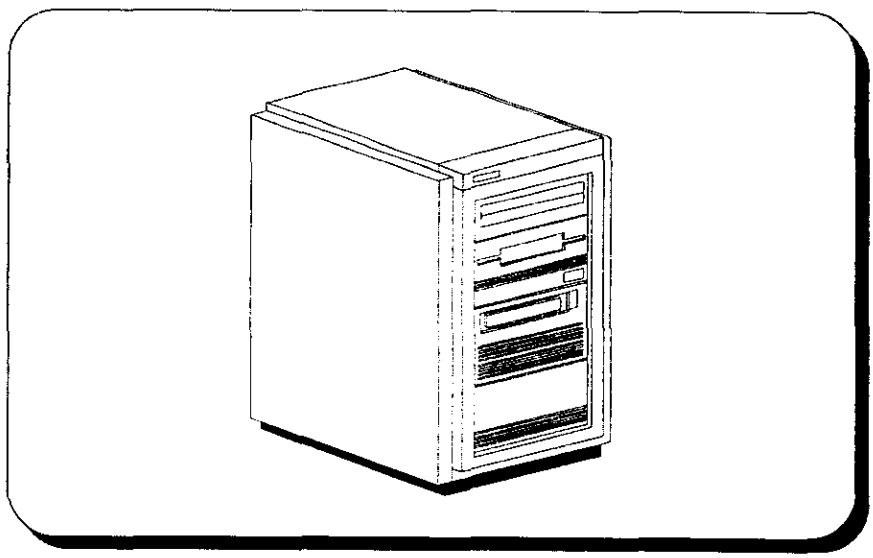


Figura 3.11 - Subsistema de discos HP 6000 SCSI

La configuración quedará como se muestra en la figura 3.12.

Es importante notar que en el presente caso el objetivo principal es la continuidad de las aplicaciones críticas. En un Sistema de Alta Disponibilidad no se busca evitar las fallas si no hacer la configuración tolerante a ellas. En la figura 3.13 se

muestra que se puede continuar con la operación aun cuando los discos, las interfaces y el CPU fallen, ya que se cuenta con los medios para sustituirlos y continuar de forma transparente la operación.

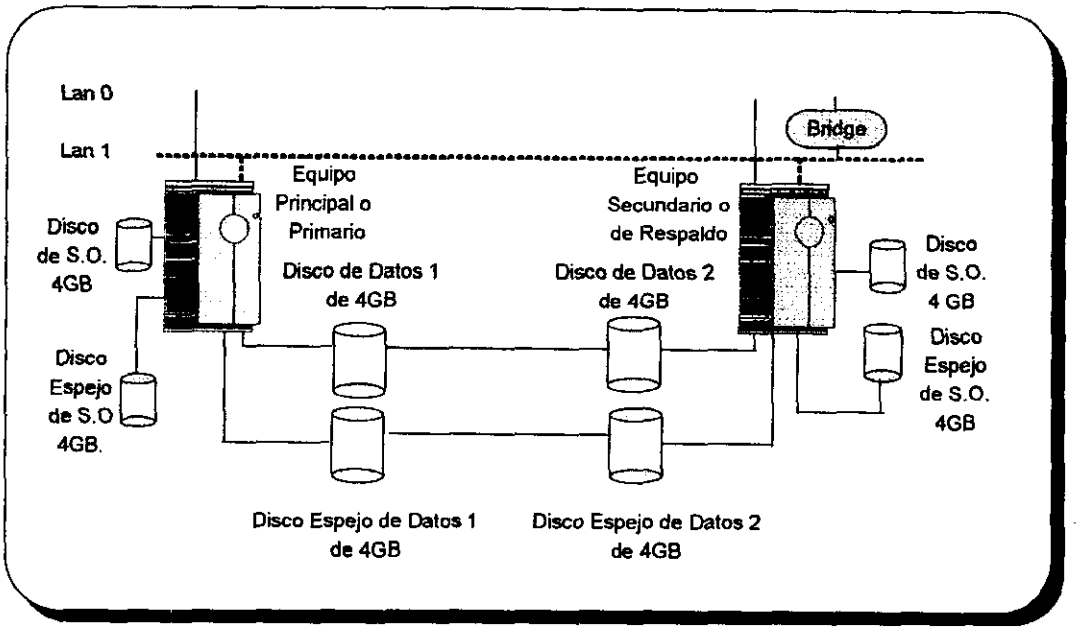


Figura 3.12 - Sistema de alta disponibilidad

El sistema completo trabajará de la forma siguiente, el equipo primario normalmente correrá las aplicaciones críticas y será el que controlará los discos y todos los recursos que sean necesarios para proporcionar los servicios. El sistema secundario puede trabajar con aplicaciones que no sean críticas que cesarán en el momento de contingencia en el equipo principal. Es decir fallas de procesadores, fallas en memoria que bloqueen el equipo, caídas de sistema o pérdida de la energía en el equipo principal ya sea por falla en la fuente de poder o de la linea de alimentación.

De esta forma los servicios continuarán aun después de fallar el equipo principal. Algo que es necesario notar es que las aplicaciones se interrumpirán por un momento y hasta que el sistema secundario tomé el total control levantará los servicios. El tiempo que transcurre desde el momento en que el equipo principal

deja de trabajar y el secundario proporcione los servicios en forma norma dependerá de varios factores. A continuación se podrá apreciar en la figura 3.13 paso a paso a través del tiempo como reaccionarían los equipos y la aplicación.

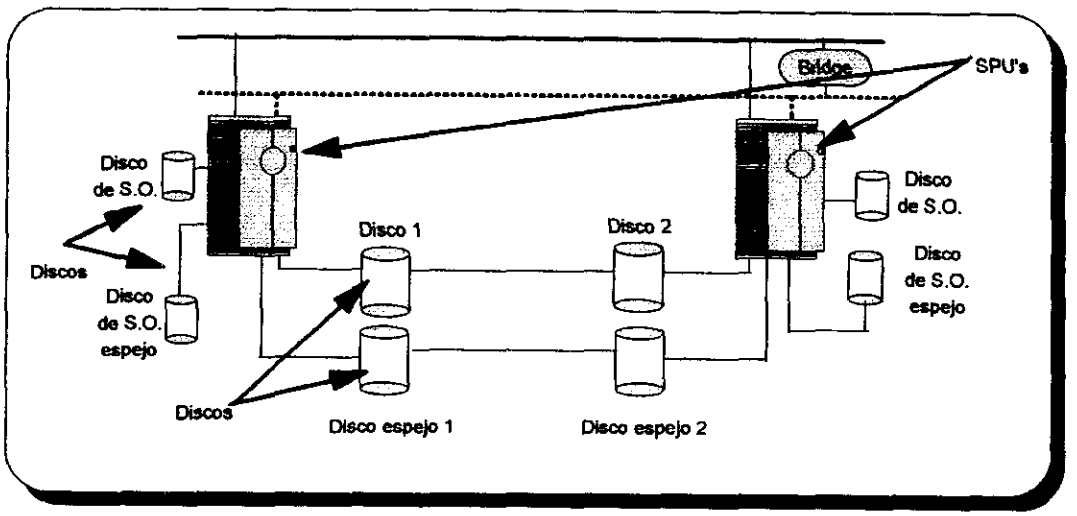


Figura 3.13 - Funcionamiento del equipo de alta disponibilidad

El equipo primario presenta un problema que ya no le permite enviar información por la red, el equipo secundario después de no recibir información asume que se ha producido una falla y comienza a auto-configurarse para tomar el papel del primario. Toma el control de los discos para empezar la recuperación de la aplicación en menos de un minuto. Como los discos fueron desactivados en linea será necesario recuperar la estructura de los archivos, en Unix se conoce como "verificación de sistema de archivos" y posteriormente se recuperará la base de datos, en el momento que se recupere se procede a levantar los servicios. Se puede apreciar que el tiempo en que el sistema de archivos se recupere, dependerá del estado de la información después de ser desactivados lo discos en linea y de la capacidad de recuperación de la base de datos. Por ello es importante manejar bases de datos confiables con una buena capacidad de recuperación. En la figura 3.14 se muestra la línea de tiempo para que se de la recuperación del sistema.

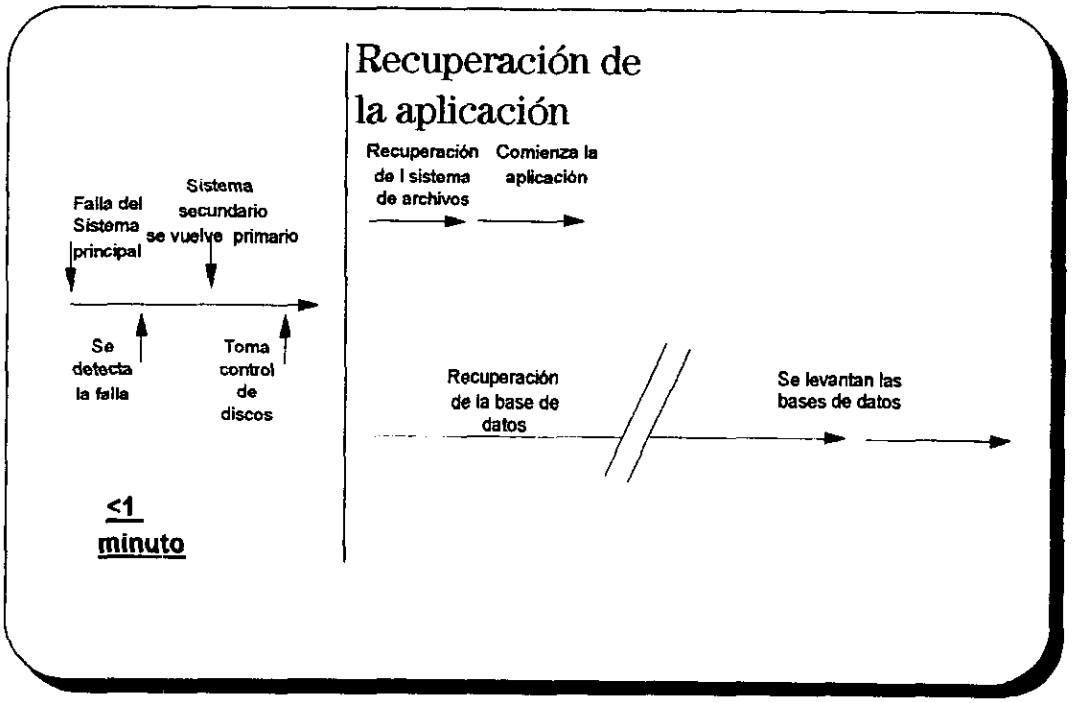


Figura 3.14 - Línea de tiempo en caso de falla

Para verificar que cumpla con el objetivo de 99% de disponibilidad tendremos que verlo en forma diferente, en primer término no debemos ver al sistema como un todo completo que puede presentar una falla y dejar las aplicaciones fuera, ahora lo veremos como que las aplicaciones críticas continuarán los 365 días del año por las 24 horas del día aun cuando falle un elemento. Así de esta forma no buscamos cuando va a fallar un elemento sino cuando se interrumpirán las aplicaciones.

Pueden fallar los discos y las tarjetas controladoras y aun así el sistema continuará trabajando por medio de los elementos redundantes, lo que significa que continuaremos con la operación. Lo único que puede detener los servicios es la pérdida del CPU principal y en este caso las aplicaciones serán controladas por el segundo sistema, dejándola por un tiempo fuera de línea. El equipo principal tiene un AFR (global sin incluir discos) de 46.2% por lo tanto se tendrá un MTBF (Horas de uso al año/AFR) de 2.16 años, la probabilidad de que fallen los dos equipos al

vez es mínima. Por lo tanto podemos continuar con los servicios, aun cuando quede fuera de linea uno de ellos. Lo importante es ver cuanto se tarda el sistema en recuperar las aplicaciones al sustituir al primario, como se comentó anteriormente puede tardar hasta 5 minutos si la recuperación es lenta. Es decir tendríamos las aplicaciones fuera de línea únicamente 5 minutos en una falla. Si se toma en cuenta que el 99 % de disponibilidad significa que se pueden tener las aplicaciones fuera de linea 87.60 horas o 3.65 días al año, entonces se tiene el tiempo suficiente para alcanzar los objetivos planteados al inicio.

Es importante recalcar que la idea en un sistema altamente disponible no es evitar la falla, sino continuar operando a pesar de ella.

CPU

AFR (global sin incluir discos) = 46.2%

MTBF = $\frac{8760}{0.462}$ = 18961.03 horas = 2.16 años

En cuanto a fallas de disco o tarjetas se cuanta con redundancia en cada una de ellas por lo que el equipo continuará proporcionando los servicios. Si es necesario apagar el equipo para una reparación, se podrán pasar los servicios al siguiente sistema y las aplicaciones continuarían aun cuando un sistema este fuera de linea por mantenimiento. Programando la reparación a horas adecuadas las aplicaciones podrían quedar fuera de linea únicamente cuando se realice el cambio de la aplicación al siguiente equipo. De esta forma se cumple con el objetivo planteado al inicio. En la figura 3.15 se muestra el equema del sistema confiable propuesto.

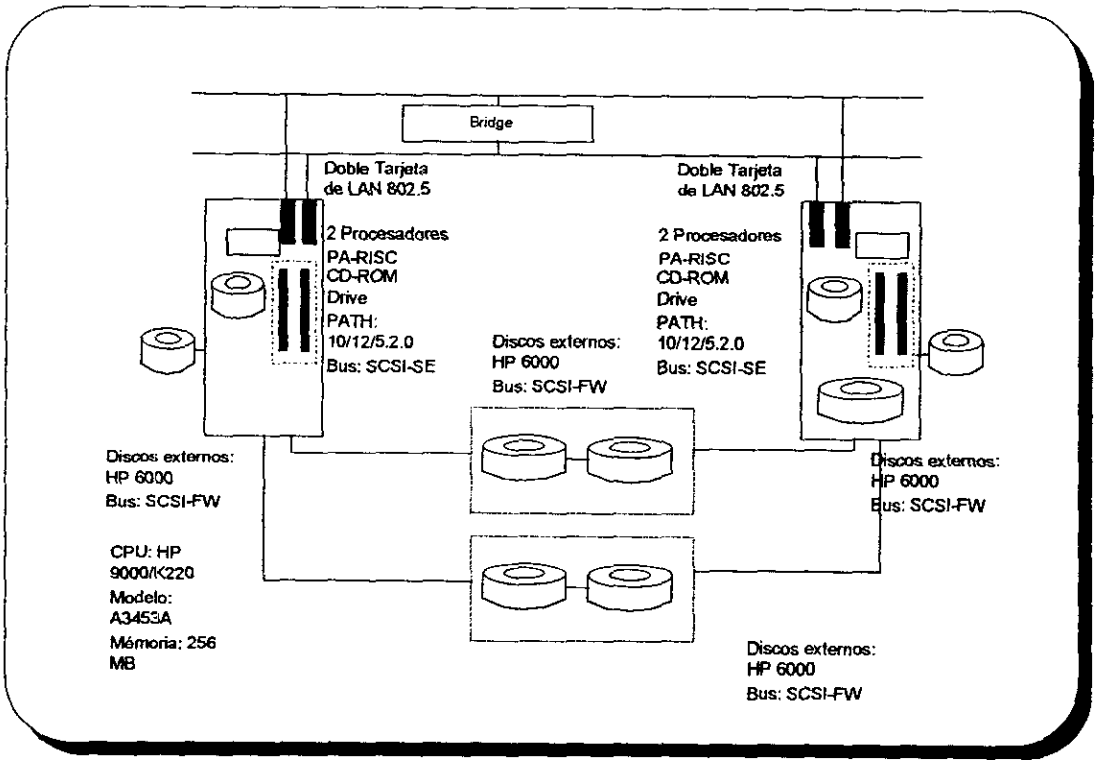


Figura 3.15 - Sistema confiable propuesto

Sistemas continuamente disponibles

El objetivo de los sistemas continuamente disponibles es mucho más ambicioso que el de los sistemas de alta disponibilidad ya que aquí no es permitida ninguna clase de interrupción, ni siquiera para realizar actividades planeadas.

El concepto de continuamente disponible es un estado ideal, sin embargo es aplicable aquellos sistemas en los cuales solo una muy pequeña cantidad de tiempo fuera de operación es aceptable.

El hecho de que un sistema posea alta disponibilidad no implica que sea continuamente disponible.

Hasta el momento dentro del ambiente de los sistemas abiertos no se cuenta con sistemas continuamente disponibles, ya que todos ellos; HP, IBM, SUN, NCR , etc., aún teniendo múltiples procesadores en paralelo se reinicializan en el momento de falla de un procesador, deshabilitandolo, y su regreso a operación depende en gran medida de la velocidad con la que pueden realizar las tareas de mantenimiento para dejar los archivos como estaban antes de la interrupción, lo cual en ocasiones lleva mucho tiempo. Aún utilizando los servicios proporcionados por los sistemas de alta disponibilidad es necesario considerar este tiempo de regreso.

Los sistemas continuamente disponibles en el remoto caso de que se llegase a presentar una falla proporcionan un tiempo de regreso a operación no mayor a dos minutos, debido a que su sistema operativo cuenta con regeneración en línea programas manejadores del sistema (drivers) que pueden ser cargados dinámicamente, menos rutinas que colapsan el sistema operativo, adición de discos en linea y dispositivos de *hardware* que se pueden cambiar mientras el equipo está

en operación. Esto se complementa con procesos que monitorean las actividades del sistema y replican completamente la información .

El costo de los sistemas continuamente disponibles es aproximadamente diez veces mayor al de dos sistemas abiertos redundantes que cuentan con los servicios de alta disponibilidad.

Las marcas de sistemas continuamente disponibles más conocidos en México son los Tandem e IBM.

Productos de Software Para Alta Disponibilidad.

Para controlar el hardware de los sistemas de alta disponibilidad es necesario utilizar software especializado, el cual se encarga de monitorear la actividad del sistema, determinar cuando se ha presentado una falla y realizar las acciones necesarias para mantener el sistema en operación.

Este software al igual que el hardware del sistema es producido por distintos proveedores, sin embargo usualmente se utiliza el software de la misma marca que el hardware.

Dentro del software existen varios niveles de seguridad, desde el software que solo replica la información entre discos, hasta el software que monitorea los procesos en diferentes CPU's detectando fallas e intercambiando servicios entre ellos para mantener la operación.
Algunos de estos productos son Parallel de SUN, el SwitchOver y Service Guard/UX de Hewlett Packard. Ya que no es objetivo de este trabajo de tesis profundizar en estos producto, a continuación se da una breve descripción del Service Guard/UX.

Service Guard/UX basa su funcionamiento en la premisa de que todos los servicios en una arquitectura cliente/servidor se dan vía red y por lo tanto no es necesario tomar toda la personalidad del equipo servidor para soportar la operación, únicamente la configuración que permita comunicarse con los clientes. Para llegar a esto se asigna una dirección IP y nombre que estará ligada a la aplicación crítica y no al equipo. Esta dirección y nombre se llama dirección flotante ya que viaja de un equipo a otro y los clientes que necesiten de ese servicio se conectarán por medio de la red a la dirección, y no a un equipo específico.

Este principio permite que los servicios sean intercambiables entre distintos equipos. Cuando el sistema primario falla, el equipo secundario toma las

aplicaciones y los recursos que necesita para proporcionar los servicios, entre ellos se encuentra la dirección flotante. De esta forma configura una tarjeta de red con la dirección en donde los clientes acostumbran tomar los servicios, no importando en que equipo estén conectados sino a que dirección de red. Así el tiempo en recuperar las aplicaciones se logra disminuir hasta a 5 minutos.

Otra ventaja es que no se limita a un equipo secundario, la configuración basada en software permite que trabajen hasta ocho sistemas conectados en red, con la capacidad de sustituir a cualquier equipo que falle. Lo que significa que cada sistema puede correr aplicaciones críticas y a la vez puede correr cualquier otra aplicación de otro equipo en caso de que presente fallas.

Cada aplicación se le asignan diferentes recursos, que pueden ser discos, aplicaciones y dirección de red flotante. A una aplicación crítica con todos sus recursos se le llama *paquete.* Para que éste pueda correr en otros sistemas, cada equipo debe de tener la posibilidad de accesar y manejar los recursos necesarios para cada paquete.

El software se encarga de hacer que cada sistema este monitoreando a otros, por medio de la red envía una señal preguntando si están bien y cada uno responde, de esta forma todos los equipos colaboran con la seguridad, en el momento en que alguno de ellos no responda, se da de baja y otro sistema en forma automática tomará el paquete y continuará proporcionando los servicios.

Cada sistema puede manejar los paquetes de otros, por lo tanto cuando se realice el mantenimiento de uno de ellos, en forma manual se pueden mover los paquetes a otro equipo y continuar con los servicios. Es importante que las aplicaciones usadas no estén ligadas a un equipo específico. Todos los recursos de un paquete deben ser accesados por los demás equipos y en un momento dado sustituir al original. En la figura 3.16 se muestra la configuración típica de un cluster de alta

disponibilidad usando el software Service Guard/UX. En la figura 3.17 se muestran las características de los paquetes por aplicación usados con Service Gurad/UX.

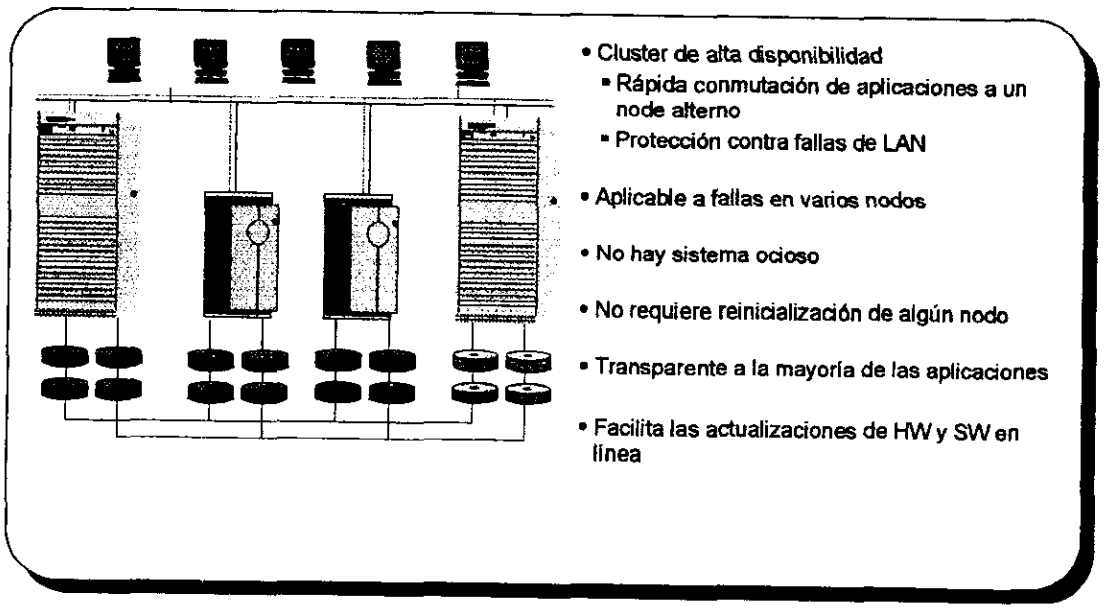


Figura 3.16 - Características de un cluster con Service Gurad/UX

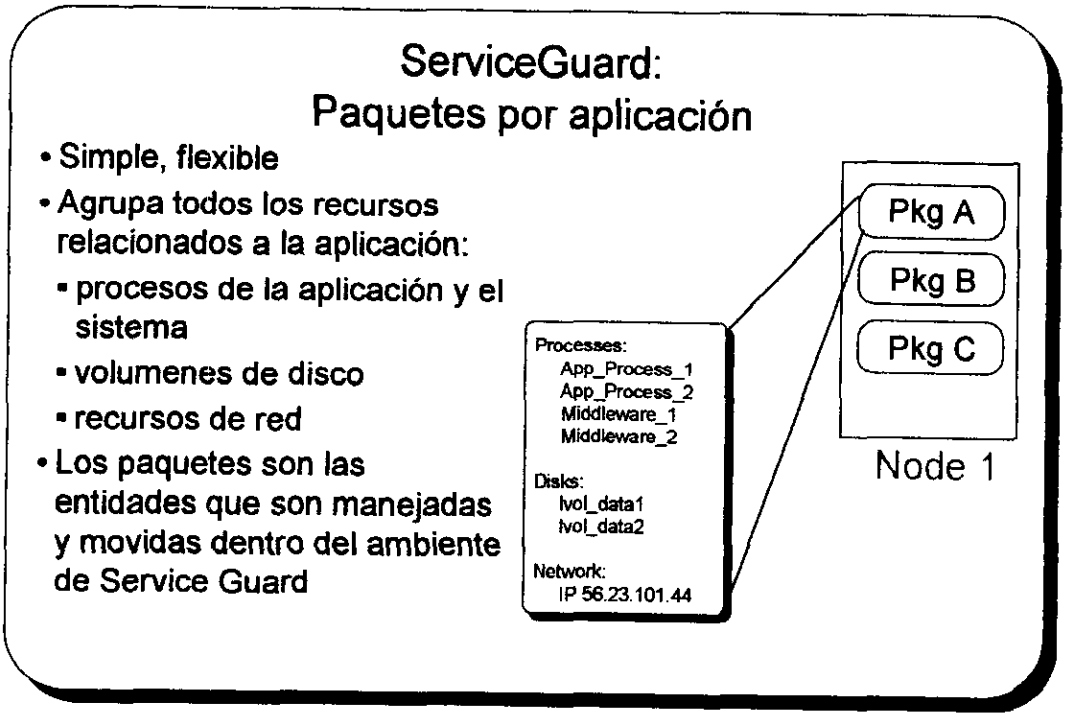


Figura 3.17 - Características de un paquete de software en Service Guard/UX

En resumen las ventajas principales son las siguientes: Menor tiempo de recuperación Hasta ocho sistemas que están trabajando conjuntamente para asegurar la disponibilidad de los servicios. Ningún sistema esta ocioso, cada uno proporciona un servicio. Por ser basado en software, los paquetes son pasados a otro sistema en forma automática evitando errores de administración.

Service Guard/UX protege para fallas de CPU, caídas de sistema, fallas en memoria, y bloqueo del equipo. Para cuidar la información es recomendable utilizar algún producto que permita la redundancia en la información. Un buen complemento es MirrorDisk/HP-UX del mismo distribuidor Hewlett Packard, que permite una copia o más en caso de ser necesario y maneja en forma transparente las fallas en los discos, permitiendo accesar la información por medio de las copias espejo.

CONCLUSIONES

Los equipos de cómputo se han vuelto indispensables para la operación de las empresas y con ello ha aumentado la dependencia de estos equipos. Las estrategias para lograr un servicio continuo son muy variadas, van desde desarrollar equipos tolerantes a fallas, hasta tener varios equipos de un mismo tipo en donde correr la aplicación crítica. En este trabajo se desarrollan las estrategias básicas que hoy por hoy se utilizan para lograr sistemas altamente disponibles en lo que se conoce como un ambiente de sistemas abiertos.

La aportación de este trabajo es importante ya que los sistemas abiertos son los que dominan el mercado de los equipos de cómputo de mediano rango y no existe hasta el momento literatura que concentre todas estas ideas. En este trabajo se exponen la tecnología y las estrategias de solución para problemas de disponibilidad en un sistema de cómputo.

El trabajo logra dar una idea clara de lo que se hace en los sistemas abiertos para obtener una alta disponibilidad. Creemos que el material desarrollado cubre de manera general las expectativas de los interesados en sistemas abiertos y pueden elegir entre las distintas alternativas que ofrece el mercado de cómputo.

Los ejemplos que se dan en el tercer capitulo son sistemas que permiten ver la aplicación de las ideas desarrolladas en el trabajo y a su vez permiten ver el alcance en sistemas de mayor tamaño. Las mismas ideas utilizadas en los ejemplos son aplicables a equipos de cualquier marca y cualquier tamaño, es por esto que este trabajo cumple con el objetivo de diseñar un sistema de alta disponibilidad para ambientes de cómputo de misión crítica.

Los sistemas de alta disponibilidad propuestos seguirán vigentes por un largo tiempo, ya que los fabricantes de sistemas abiertos cada vez enfocan más sus esfuerzos a la creación de hardware y software de estas características. No sería raro que en el futuro los sistemas se vendieran ya con las soluciones redundantes implantadas. Aun entonces este trabajo seguirá vigente ya que los conceptos fundamentales (como en toda tecnología) seguirán siendo aplicables.

APÉNDICE

A

Evaluación de la capacidad del sistema

La elección adecuada del sistema de cómputo depende de varios puntos, entre ellos está el número de usuarios y también el crecimiento a futuro. Una guía básica para seleccionar la capacidad del equipo se presenta a continuación:

Comúnmente las actividades que desarrollan los usuarios pertenecen a alguno de tres ambientes siguientes:

Tabla 1.1

Tipo de Ambiente	Carga de Trabajo
Desarrollo de software o aplicaciones comerciales con bases de datos relaciónales y un alto uso de los medios de almacenamiento.	Alto
Aplicaciones comerciales con alto uso de medios de almacenamiento	Medio
Aplicaciones comerciales sin bases de datos relaciónales y un promedio bajo o medio de manejo de medios de almacenamiento.	Bajo

Existen dos tipos de usuarios cuando se determina la capacidad de un sistema.

Los usuarios concurrentes, son aquellos que trabajan continuamente y no tienen interrupciones en el trabajo, siempre están trabajando en el sistema.

Los usuarios que no son concurrentes, tienen interrupciones y utilizan el sistema temporalmente.

Para determinar la capacidad del sistema, se debe tomar en cuenta el número de usuarios concurrentes y los no concurrentes. Junto con el tipo de ambiente en la que estará trabajando el equipo. La tabla 2.1 muestra la capacidad del sistema que se necesita de acuerdo al número de usuarios.

Tabla 1.2 Capacidad del sistema de acuerdo al número de usuarios.

Rendimiento en TPM´s	Carga Alta	Carga Media	Carga Baja
500	40	120	150
700	70	200	250
800	80	240	300
1,030	100	310	390
2,200	220	660	830
2,710	240	720	900
2,750	280	830	1,030
3,230	320	970	1,210
2,710 - 8,410	430	1,300	1,620
2,710 -8,410	460	1,380	1,730
4,090 - 12,680	650	1,960	2,450
4,800 - 18,150	1,070	3,200	4,000

Requerimientos de memoria RAM:

- Utilizando un sistema operativo unix como estándar, los requerimientos básico son como a continuación se indica:

UX incluyendo servicios de NFS y LAN 18 MB de RAM

En donde NFS es conocido como *Network File System*, espacio de disco magnético que estará compartido por servicios de red con otros sistemas. A la vez LAN son servicios que permiten la comunicación con los demás sistemas.

- La cantidad de memoria RAM por usuario:

- Con ambiente gráfico
  (2 MB X Número de usuarios)
- Utilización de aplicaciones sin bases de datos relaciónales
  (0.5 MB X Número de usuarios)
- Utilización de Informix, Ingres, Sybase
  (1.0 MB X Número de usuarios)
- Utilización de Oracle
  (1.5 MB X Número de usuarios)

- Requerimientos de aplicaciones

  Varía dependiendo de la aplicación y se necesita revisar cuales son los requermientos de cada uno. En promedio una aplicación común utiliza de 6 a 8 MB de memoria.

Se hace una suma de los resultados anteriores y se tiene la memoria que se requiere, a la vez se realiza una estimación del crecimiento del sistema dependiendo del número de usuarios o aplicaciones futuras.

Requerimientos del espacio en disco

- En promedio el sistema operativo unix, con servicios de redes Lan, NFS con ambiente gráfico utiliza 570 MB de espacio en disco.

- A esta cantidad se suma también lo que utilicen las aplicaciones. No existe una guía general, para esto se recomienda pedir los datos a cada proveedor.

- Para cada usuario se recomienda la siguiente regla

    15 MB X Número de usuario.

- Ademas es necesario sumar el espacio de la memoria que utilizará como swap.

    Mínimo = 2 X el total de memoria RAM.

Como ejemplo se puede pensar en un equipo de 200 usuarios que manejan bases de datos relaciónales con sistema operativo unix, no será necesario desarrollar aplicaciones, con un alto uso de los canales de entrada salida y a la vez 20 usuarios manejarán ambiente gráfico. La aplicación estará en Oracle que necesita 120 MB de espacio en disco. Se piensa que a futuro se tendrán 300 usuarios con 30 de ellos utilizando el ambiente gráfico. La información del cliente ocupa 4 GB.

Solución:

Al no tener desarrollo de aplicaciones, se puede pensar en un sistema de carga media, por lo tanto un sistema con 1,030 TPM´s puede ser utilizado para soportar 300 usuarios.

Para saber la cantidad de memoria RAM:

Utilización de unix con NFS y LAN		18 MB
Usuarios trabajando con Oracle	300 X 1.5	450 MB
Usuarios con ambiente gráfico	30 X 2	60 MB
Aplicación Oracle		8 MB
Total		536 MB

Para el espacio en disco

Sistema operativo unix		570 MB
Aplicación Oracle		120 MB
Espacio para usuarios	300 X 15	4500 MB
Espacio en swap	536 X 2	1072 MB
Datos del usuario		4000 MB
Total		10262 MB

Resultado:

Un sistema con 1,030 TPM´s, con 536 MB de RAM y 10,262 MB

GLOSARIO

A

AFR. Anualized Failure Rate. Índice anualizado de fallas.

ANFITRION. Véase Host.

ANSI. American National Standards Institute. Instituto Nacional Americano de Normas. Coordina el desarrollo de normas nacionales en los Estados Unidos de América, tanto en sectores privados como públicos. representa a las asociaciones de EUA para la ISO y la IEC.

ARRAY CONTROLLER. Controlador del Arreglo. Componente del arreglo de discos que transfiere o recupera la información de los discos que conforman el arreglo, de acuerdo al algoritmo RAID que se haya determinado.

AUI. Attachment Unit Interface. Interface de adicionamiento. Conector que se emplea en un adaptador de red para ser adicionado a un cable coaxial grueso ethernet.

B

BEACONING. Forma de señalización continua de condiciones de error en una LAN.

BIT. Dígito simple, 1 ó 0, en un número binario

BRIDGES. Puente. Dispositivo que conecta dos segmentos de red entre sí, que pueden ser similares o no. como Ethernet y Token Ring.

BNC. British Naval Connector. Conector utilizado para cable coaxial.

BUGS. Error de diseño en el hardware o software.

BUS. Ruta o conexión común entre dispositivos múltiples.

BYTE. Octeto. Por lo común ocho dígitos binarios utilizados para la transferencia de información en equipos de cómputo.

BYTE-BY-BYTE. Forma de repartir la información en los arreglos de disco en el cual los discos almacenan datos y otro paridad.

BYTE STRIPING. Forma de repartir la información en los arreglos de disco en el cual todos los discos almacenan tanto datos como paridad.

C

CABLE COAXIAL. Tipo de cable utilizado en comunicaciones que tiene un hilo conductor al centro cubierto de material aislante y un segundo conductor rodeando a éstos, además estará cubierto por otros materiales para darle mayor protección.

CD-ROM. Compact Disc - Read Only Memory. Disco compacto que se utiliza para almacenar texto, gráficos y sonido estéreo de alta fidelidad, utiliza un formato de pistas de datos diferente al CD de audio.

CLIENTE. En un ambiente cliente-servidor, estación de trabajo o computadora personal que utiliza los recursos del servidor.

CLUSTER. Controlador de grupos que administra varios dispositivos periféricos, como terminales.

CONCENTRADOR. Repetidores multipuerto con capacidad de detección y autosegmentación.

CONECTOR HERMAFRODITA. Conector utilizado en redes token ring que una vez conectado al dispositivo es posible conectar sobre de él otro más.

CUARTO BLANCO. Ambiente libre de impurezas.

CUNCURRENCY I/O. Concurrencia de Entrada/Salida. Característica del arreglo de discos de poder realizar múltiples transacciones de Transferencia de datos simultáneamente.

CPU. Central Processing Unit. Unidad Central de Procesamiento. La parte de la computadora que realiza el procesamiento.

CSMA/CD. Carrier Sense Multiple Access with Collision Detection. Método de acceso en las comunicaciones de banda base. Cuando un dispositivo trata de ganar acceso a la red verifica que el canal se encuentre libre, en caso contrario espera un lapso aleatorio para reintentarlo. En caso de que dos dispositivos accesen el

canal al mismo tiempo, ambos cancelan su intento y esperan de nuevo un lapso aleatorio para reintentarlo.

CYLINDER. Cilindro. Conjunto de todas las pistas que residen en la misma ubicación en la superficie de disco. En caso de discos múltiples, será la suma de las pistas de las superficies de los platos en dicha posición.

D

DATA PROTECCION. Protección de Datos. Es el proceso por el cual los datos son distribuidos a través de los distintos discos que conforman el arreglo de discos.

DATA STRIPING. Distribución de datos. Característica propia de los algoritmos RAID1, RAID3 y RAID5 en un arreglo de discos de discos. Con éste, es posible recuperar la información en caso de fallar alguno de los discos que conforman el arreglo.

DATA TRANSFER RATE. Velocidad de Transferencia de Datos. Característica de los dispositivos periféricos. Indica que tan rápido pueden ser transferidos los datos del procesador central al dispositivo periférico.

DISK ARRAYS. Arreglo de discos. Combinación de discos en una sola unidad para mayor capacidad, velocidad, u operación tolerante a fallas.

DISTANCE VECTOR. Vector de distancia. Algoritmo que determina la ruta más corta para la transmisión de información en un sistema interredes.

DRIVER. Controlador. Es una rutina de programa que enlaza un dispositivo periférico al sistema operativo. También designa al dispositivo que provee de señales o corriente eléctrica para activar una linea de transmisión o canal de información.

E

ETHERNET. Es una red de área local (LAN). Todos los mensajes se envían a todos los nodos en el segmento de red, conecta hasta 1024 nodos a 10 Mbits por segundo.

F

FDDI. Fiber Distributed Data Interface. Interface que utiliza cable de fibra óptica y transmite a 100 Mbits por segundo en distancias de hasta 2 kilómetros.

H

HARDWARE. La parte física de los equipos de cómputo.

HEAD. Cabeza. Dispositivo para leer y escribir datos sobre la superficie de un disco o cinta magnética.

HOST. Anfitrión. Computadora central en un entorno de sistema distribuido, o que controla una red.

HUB. Concentrador. Dispositivo de conexión central en una red con configuración estrella, de la cual parten los enlaces.

I

IEEE 802.3. Estándar para el protocolo de comunicaciones en una red ethernet.

IEEE 802.5. Estándar para el protocolo de comunicaciones en una red token ring.

IEC. International Electrotechnical Commission. Comisión Electrotécnica Internacional. Organización que establece estándares internacionales en electricidad y electrónica, con sede en Ginebra.

IEEE. Institute of Electrical and Electronic Engineers. Instituto de Ingenieros Electricistas y Electrónicos. Organización involucrada en el establecimiento de estándares en informática y comunicaciones.

INITIATOR. Iniciador. En la interface SCSI, nombre con el que se denomina al dispositivo que toma la iniciativa para la transferencia de información con otro dispositivo llamado objetivo o *target*.

INTERFACE. Conexión para la interacción entre el sistema de cómputo y sus dispositivos de entrada/salida, o con el usuario.

IP ADDRESS. Internet Protocol Addres. Dirección de Protocolo Internet. Dirección única para cada nodo de la red, que consiste de 4 cifras del 0 al 255 que designan además la sección de red a la cual pertenece. Por ejemplo: 192.1.1.2.

J

JUMPER. Puente de conexión. Es un pequeño conector de metal cubierto de plástico que presiona sobre dos pines, o agujas de metal, para cerrar un circuito

L

LAN. Local Area Network. Red de Area Local. Red de comunicaciones que sirve a usuarios dentro de un área geográficamente limitada.

LAN 802.3. Red de comunicaciones que utiliza el protocolo de comunicaciones IEEE802.3.

LINK STATE. Estado del Enlace. Condición que guarda el canal o enlace.

LOOP. Bucle, lazo. En redes de comunicaciones, repetición infinita de un mismo mensaje para el cual no se encuentra el destinatario.

LUN. Grupo de mecanismos o discos dentro de un arreglo de discos.

M

MAC ADDRESS. Medium Access Control Address. Dirección única de las interfaces de comunicación. Se compone de doce dígitos hexadecimales asignados por el fabricante del dispositivo para asegurar que no se repita a nivel mundial.

MAN. Metropolitan Area Network. Red de Area Metropolitana. Red de comunicaciones que cubre un área geográfica que puede ser una ciudad.

MAU. Media Attachment Unit. Convertidor de señal o transceptor utilizado para enlazar los dispositivos a las redes de coaxial grueso.

MODEM. Modulator-Demodulator. Modulador-Demodulador. Dispositivo utilizado para conectar un equipo periférico a la computadora, y que adapta las señales digitales al medio de transmisión, por ejemplo, la línea telefónica.

MSAU. Multistation Access Unit. Núcleo central en una red de área local de tipo anillo (TOKEN RING).

MTBF. Mean Time Between Failure. Tiempo medio entre fallas. Tiempo promedio que un componente trabaja sin fallar.

MTTR. Mean Time To Repair. Tiempo promedio para reparación. Tiempo promedio para reparar un componente dañado.

N

NFS. Network file System. Sistema de archivo en red. Sistema de archivos distribuidos que permite compartir datos a través de una red, independientemente de la máquina, sistema operativo, arquitectura de la red o protocolo. Hace parecer los archivos como si fueran locales en la máquina del usuario.

O

OSI. Open system Interconection. Interconexión de sistemas abiertos. Estándar para comunicaciones a nivel mundial que define una estructura para la realización de protocolos en siete estratos o capas. El control se transfiere de un estrato al siguiente: físico, enlace de datos, red, transporte, sesión, presentación, y aplicación.

P

PA-RISC. Precision Architecture - Reduced Instruction Set Code. Denominación que le da Hewlett Packard a su arquitectura de procesadores.

PAR TRENZADO. Par de alambres retorcidos uno alrededor del otro y de esta forma minimizar las interferencias provenientes de otros cables.

PC. Personal Computer. Computadora personal.

POWER DOWN. Apagar. Cortar el funcionamiento de un equipo en forma ordenada.

POWER UP. Encender. Iniciar el funcionamiento de un equipo en forma ordenada.

PROCESAMIENTO DISTRIBUIDO. Ambiente de cómputo donde por medio de la red los equipos de cómputo que la forman realizan el procesamiento de su información en forma local y la mantienen disponible para el resto de los sistemas.

R

RAM. Random Access Memory. Memoria de acceso aleatorio. Uno de los componentes principales de un equipo de cómputo, resulta ser el área principal de trabajo dado que la información resulta de más rápido acceso que estando en un disco o unidad de cinta.

RIBBON CABLE. Cinta delgada y plana que contiene varios hilos conductores.

RAID. Redundant Array of Independent Disks. Arreglo redundante de discos independientes. Alguno de los algoritmos de un arreglo de discos que proporcionan diversos grados de protección de la información que contienen los discos que lo conforman, por ejemplo RAID1.

ROUTERS. Ruteador. Dispositivo que en una red retransmite paquetes de datos entre diversos tipos de redes, por ejemplo, entre LAN y WAN.

RS-232C. Protocolo de comunicaciones utilizado en comunicaciones seriales, por ejemplo en terminales e impresoras.

S

SASI. Shugart Associates System Interface. Interface de sistemas asociados de Shugart. Interface de periféricos desarrollado por Shugart y NCR que evoluciono hasta convertirse en el estándar ANSI SCSI.

SCSI. Small Computer System Interface. Interface pequeña de sistemas de computadoras. Es una interface que permite la conexión de hasta 7 dispositivos periféricos, por ejemplo, discos o cintas al computador central.

SCSI-FW. Véase SCSI FAST WIDE

SCSI FAST WIDE. Interface SCSI que utiliza circuitos de corriente, menos susceptible al ruido, por lo que permite mayor velocidad de transmisión, y mayor número de bits de datos simultáneos.

SCSI-SE. Véase SCSI SINGLE ENDED.

SCSI SINGLE ENDED. Interface SCSI que utiliza señales eléctricas con referencia a una tierra común, más susceptible al ruido.

SCSI-1. Es la especificación SCSI original. Utiliza 8 bits de datos a 5 megabits por segundo(Mb/s), utiliza lazo de voltaje para detección de señales.

SCSI-2. Variante del original SCSI-1 y que define a la interface como *fast* o rápida, ya que permite una velocidad de transferencia de 10 Mb/s por segundo con 16 bits de datos, utiliza lazo de voltaje para detección de señales.

SCSI-3. Versión más reciente del standard SCSI, y la define como *Wide* o ancha, ya que provee 32 bits de información a 20 Mb/s, utiliza lazo de corriente para detección de señales.

SCSI/FW. SCSI FAST AND WIDE. Término común utilizado para designar a la interface rápida y ancha. Ver SCSI-3.

SECTOR. La unidad de almacenamiento más pequeña en un disco. Las pistas del disco o círculos concéntricos, se dividen en sectores.

SERVIDOR. En un ambiente cliente-servidor computadora que comparte sus recursos con los denominados clientes.

SINGLE-ENDED SCSI INTERFACE. Interface SCSI con terminación sencilla. Término utilizado para designar a la interface SCSI-2, por el uso de lazo de voltaje para detección de señales. Todos los circuitos eléctricos del canal de datos tienen una tierra común, lo que la hace más susceptible al ruido. Ver SCSI-2

SNMP. Simple Network Managment Protocol. Protocolo simple de administración de redes. Protocolo de monitoreo y control de redes.

SOFTWARE. Instrucciones ejecutables por una computadora.

SPANNING TREE. Nombre genérico con el que se denomina a los protocolos que utilizan los ruteadores para evitar *loops* en un sistema interredes.

SPARE. Reserva. Componente o sección de reserva para reemplazar a la que este dañada.

SPIN UP. Aceleración radial. Acción durante la cual el disco alcanza su velocidad de rotación característica.

STANDALONE DISK. Disco independiente. Disco duro utilizado para almacenar información en forma independiente.

STANDBY. En espera. Estado de operación del equipo o dispositivo en el que espera la señal o condición para realizar sus funciones.

T

TARGET. Objetivo. En la interface SCSI, nombre con el que se denomina al dispositivo que recibe la petición del iniciador o *inititator,* para la transferencia de información.

TCP-IP. Transmission Control Protocol / Internet Protocol. Protocolo de comunicaciones desarrollado bajo permiso del Departamento de Defensa de los EUA para intercomnunicar sistemas diferentes. Se ha convertido en un estándar de UNIX.

TOKEN RING. Red de anillo de señal. Método de acceso a LAN que utiliza la tecnología de paso de señal en un anillo físico. conecta hasta 255 nodos a una velocidad de 4 ó 16 mbits por segundo. Todas las estaciones se conectan a un dispositivo central llamado Multi Station Access Unit.

TRACKS. Pista. Canal de almacenamiento en disco o cinta, resultan ser círculos concéntricos en los discos duros.

TRANSCEPTOR. Dispositivo que realiza las funciones de transmisor y receptor de señales digitales en cierto medio de comunicación, sea cable coaxial o fibra óptica.

U

UPS. Uninterruptible Power Supply. Fuente de poder ininterrumpible. Dispositivo que provee de energía eléctrica regulada inclusive cuando la energía eléctrica de la linea comercial se interrumpe o baja a un nivel de voltaje inaceptable.

V

VLSI. Very Large Scale of Integration. Integración a escala muy grande. Entre 100 000 y 1 000 000 transistores en una sola oblea o chip.

W

WAN. Wide Area Network. Red de área ancha. Red de comunicaciones que abarca áreas geográficas muy amplias, por ejemplo un estado o país.

WIDE-DIFFERENTIAL SCSI INTERFACE. Interface SCSI diferencial-ancha. Término común utilizado para designar a la interface SCSI-3 debido al uso de lazo de corriente. Cada señal del canal de datos es detectado por un par individual de alambres, lo que proporciona mayor inmunidad al ruido. Ver SCSI-3

BIBLIOGRAFÍA

Libros

Ken, Sherman. Data Communications A User's Guide.
Tercera Edición. Prentice Hall 1990.

Weygant, Peter S. High Availability Workshop.
Primera Edición. Hewlett Packard - Prentice Hall 1992. PS3003.

Ceri, Stefano. Art and Science of computing.
Segunda Edición. Addison-Wesley Longman.

Pfister, Gregory. In Search of Clusters.
Primera Edición. Prentice Hall

Lewis, Harry. Elements of the theory of computation.
Primera Edición. Prentice Hall.

Storex, Neil. Safty Critical Computers Systems.
Addison-Wesley Longman

Gibbs, Mark. Absolute beginners Guide to Network.
Sams Publishing, 1996.

HP 9000 Enterprise Servers Configuration Guide.
Hewlett Packard Co., 1997, 4/97

Direcciones en Internet

High Availability Cluster Software
http://www.encore.com/realtime/products/sw/hacs.html

High Avalability Clusters Multiprocessing for AIX version 4.2.1
http://www.rs600.ibm.com/software/Apps/hacmp/hacmp.htm

High Performance Computin Cluster
http://scsweb.nyu.edu/scihpc.html

High Availavility and Cluster Solutions
http://www.digital.com/css/ha_sol_all.html

"SCSI for Sure"
http://bmdinfo.bbn.hp.com:4104/dw/it/d/cpcd/supnews/93news/
news85/8512.txt/control.htm

Harware Products Subarea
http://hprtdt85.grc.hp.com/document/mewe.html

SCSI-2 Specifications (Draft X3T9.2 Rev. 102)
http://pet.rose.hp.com/SCSI/SCSI2.html

IBM
http://www.ibm.com

Hewlett Packard Co.

http://www.hp.com

Tandem

http://www.tandem.com

Digital

http://www.digital.com