

**Informe  
No. 15**



**Diseño de Redes Novell**



**Club de Investigación Tecnológica**

# **Diseño de Redes Novell**

**Preparado por: David Baruch Goldberg  
Agosto 1993**

**Editado y Publicado por Rho-Sigma, S.A.  
a nombre del Club de Investigación Tecnológica.  
Todos los derechos reservados.  
Prohibida la reproducción total o parcial.  
San José, Costa Rica  
Agosto 1993**

# Resumen Ejecutivo

Las redes de computadoras han sido el esquema computacional que más ha sufrido de críticas y controversias.

Es la opinión del autor, que en su mayoría, las críticas que se han dado a las redes de computadoras provienen de los vendedores de equipos de mayor tamaño, centralizados, que ven en peligro sus utilidades, como resultado de un ataque a su mercado cautivo, de un esquema de muy bajo costo y alto rendimiento.

Debido a que típicamente, para implantar una red, se requieren de una serie de componentes de Hardware y Software provenientes de diferentes suplidores, las redes son por lo tanto vulnerables a este tipo de ataque de mercadeo.

Una red de computadoras bien implantada, ofrece no solo un alto rendimiento computacional a muy bajo costo, sino también la garantía, estabilidad, y disponibilidad, de los recursos computacionales.

La fiera competencia que se ha desatado en los últimos meses en el mercado mundial de las computadoras, ha dado como resultado que el principal componente de las redes de computadoras haya bajado de precio en

forma radical. El efecto final para el usuario es la posibilidad de implantar un ambiente computacional de un altísimo rendimiento, y a muy bajo costo.

En este informe se describen aspectos de diseño de redes de computadoras, y debido a la popularidad del sistema operativo Netware, y el medio de transmisión Ethernet, se enfocará principalmente en redes de microcomputadoras bajo esos dos estándares.

Este trabajo presenta los aspectos de mayor importancia en el diseño de redes de computadoras, que hoy en día son utilizadas para establecer sistemas informáticos confiables, eficientes, e integrados alrededor del mundo.

Muchas organizaciones están rediseñando sus aplicaciones que actualmente corren en mainframes, para disminuir su costo de operación por medio de la implantación de redes.

El material se divide en tres partes:

- El medio de transmisión de datos,
- Diseño de servidores, y
- Diseño de estaciones.

En cada área, se presentan tanto aspectos teóricos como prácticos, así como

limitaciones y recomendaciones específicas sobre aplicabilidad y eficacia.

Se deben pedir disculpas a la Real Academia Española, y a los Protectores de la Lengua. Algunos términos no son fácilmente traducidos del inglés (idioma en el que seguramente nació el concepto), y por lo tanto son usados con cierta libertad editorial para buscar una base común de entendimiento conceptual.

### **Agradecimientos**

Este trabajo, es el resultado de la investigación así como de la experiencia profesional y académica del autor. Se agradece muy especialmente Jay Trabad, quien en los últimos 6 años ha sido fuente de inspiradas ideas sobre aplicaciones de computadoras, y especialmente de las redes de microcomputadoras, en el foro de consultores de Compuserve. Se agradece

también la valiosa participación de los miembros del comité editorial de este informe: Dr. Ulises Agüero, Lic. Edgar Hernández y Dr. Roberto Sasso por la pertinencia de sus observaciones sobre el presente trabajo. Se agradece además a Irene Fajardo, quien transcribió innumerables grabaciones que resultaron en este texto.

### **Del Autor**

David Baruch, Presidente de Télesis S.A., casa de software basada en Costa Rica, es Ingeniero Civil por la Universidad de Costa Rica, Máster en Ciencias de la Computación y Máster en Administración de Negocios por la University of Pennsylvania. La experiencia del Ing. Baruch en el ambiente de redes discutido en este informe se remonta a 1982 cuando fué responsable de las primeras redes instaladas en Costa Rica. Desde 1989 es soportista certificado de Novell.

<b>Contenido</b>	<b>Página</b>
<b>I El Ambiente Computacional de las Redes.....</b>	<b>1</b>
<b>II El Diseño debe ser Orientado a las Aplicaciones.....</b>	<b>5</b>
<b>III El Efecto de las Aplicaciones sobre la Red.....</b>	<b>8</b>
<b>IV Ethernet: El Medio de Transmisión .....</b>	<b>13</b>
1 Ethernet Estándar (10Base5) .....	15
2 Ethernet "Delgado" (10Base2-ThinNet) .....	15
3 Ethernet 10BaseT .....	16
4 Diseño de segmentos de Red 10Base .....	17
5 Ethernet de Alta Velocidad .....	17
6 Fast Ethernet, 100Base-X .....	19
7 100BaseVG. ....	20
8 Switching Ethernet Hubs .....	21
9 Implantaciones con Ethernet .....	21
<b>V Diseño del Servidor de la Red.....</b>	<b>23</b>
1 Conceptos sobre Servidores de Red .....	24
2 Arquitectura del Servidor. Procesador 386, 486, Pentium, Alfa .....	31
3 Arquitectura del Servidor. Bus de ExpansiónISA, EISA, MCA, Local-Bus ..	32
4 Memoria .....	36
5 Almacenamiento. Discos DurosMFM, IDE, ESDI, SCSI .....	37
6 Almacenamiento. ARDIs (RAIDs) .....	42
7 Almacenamiento. CD-ROM Alta Capacidad a Bajo Costo .....	47
8 Almacenamiento. Co-Procesadores de Disco .....	48
9 La Conexión a la Red .....	49
<b>VI Diseño de las Estaciones .....</b>	<b>50</b>
1.La Relación de la Estación con la Red .....	51
2.La Estación Ideal .....	52
3. Laptops y Notebooks .....	53
4. WINDOWS 3.1 .....	54
<b>Bibliografía.....</b>	<b>58</b>

# I El Ambiente Computacional de las Redes

En la actualidad, existen muchos tipos esquemas y marcas de computadoras, que al fin y al cabo todas buscan el mismo objetivo: procesar información.

El esquema más conocido y tradicional, es el computador que cuenta con su CPU (Unidad Central de Proceso, que es el dispositivo de la computadora que hace el trabajo principal), su memoria, y su almacenamiento (típicamente en la forma de un disco duro). Este equipo es usado por varios usuarios por medio de terminales, que en la mayor parte de los casos se trata de un dispositivo sin capacidad de realizar proceso por sí mismo (conocido como terminal tonta). En algunos casos, se usan micro-computadoras para emular terminales de este sistema.

Este esquema se conoce como un esquema de proceso centralizado, y que los vendedores de hardware gustan llamar un sistema "Multiusuario".

No hay que dejarse engañar, ya que las redes de micro-computadoras son en realidad también sistemas multiusuario, en el sentido léxico de la palabra, es decir que múltiples usuarios pueden acceder los recursos disponibles en el ambiente computacional.

El ambiente centralizado que se menciona anteriormente, tiene un límite máximo de proceso, establecido por las cantidades y cualidades de los 3 elementos fundamentales:

- Capacidad de la CPU (Operaciones por segundo, MFLOPS, SPECMARCS, MIPS, etc)
- Memoria (Medida en MB).
- Almacenamiento (Medida en MB, o GB).

Pero fundamentalmente, cuando la cantidad de usuarios sobrepasa los límites de la combinación de la capacidad de la CPU y la memoria, este sistema centralizado o debe ser sustituido por uno de mayor capacidad, o se le hace participar en una red de computadoras, de tal forma que los requerimientos de los usuarios son distribuidos entre varios equipos.

Las redes de microcomputadoras, por su lado ofrecen a cada uno de los usuarios, su propia CPU, su propia memoria y en algunos casos hasta su propio almacenamiento en disco.

En términos muy simplistas, una red de micro-computadoras se compone de una serie de micro-computadoras conectadas



entre sí por un cable que permite la comunicación entre ellas y organizadas por medio de un software, que permite compartir los recursos disponibles entre ellas.

Para lograr este esquema, típicamente se utilizan computadoras de una marca, tarjetas para comunicaciones de red de otra marca, cables y conectores, sistemas operativos, utilitarios de comunicación y administración, y aplicaciones.

Bueno, es precisamente esta impresión de desorden las que aprovechan algunos

vendedores de equipos del estilo centralizado, para asustar a posibles compradores de equipos y convencerlos de no adquirir la tecnología de redes.

Sin embargo y para evitar el desorden, se ha establecido un estándar de la industria, conocido como el modelo OSI (Por Interconexión de Sistemas Abiertos), que se compone de 7 capas.

En orden reverso estas capas son:

CAPA	FUNCION
7: Aplicación	Servicios de comunicaciones.
6: Presentación	Interfase de la aplicación a la red.
5: Sesión	Envío y recepción de mensajes.
4: Transporte	Direccionamiento de la red.
3: Red	Conexiones inter-red
2: Enlace de datos	Interfase del equipo.
1: Física	Especificaciones del equipo.

A pesar de que este modelo pudiera salirse del ambiente y el interés de este documento, se incluye aquí ya que se trata de una norma ampliamente publicada y que actualmente es cumplida por casi la totalidad de los fabricantes de componentes de redes. La importancia radica en que al ser un estándar **ampliamente** seguido, se puede garantizar la aplicabilidad de los diferentes componentes, comprados a diferentes suplidores, al esfuerzo de integrar un sistema

computacional basado en redes de micros.

Simplificando, toda red Novell integra 4 tipos de componentes:

- **Servidor.** Es realmente el dispositivo donde se encuentra el almacenamiento de la red, que es compartido entre los usuarios. En el servidor corren los programas principales del sistema operativo de la red, y puede tener algunas funciones

especializadas como lo es el servidor de impresión. La velocidad del procesador (CPU) es crítica para determinar la capacidad computacional total de la red mientras que la cantidad de memoria en el servidor es función del tamaño del disco duro que esté instalado en el, y los requerimientos determinados por las funciones especiales, adicionales al servidor básico, que vayan a ser realizadas en el servidor de la red. Dentro de una misma red pueden existir más de un servidor; de hecho este esquema es muy utilizado para distribuir la carga de procesamiento en una red y optimizar el tiempo de respuesta de los usuarios. Los servidores de una red pueden ser desde computadores basados en el procesador 80286 hasta el 80486, y algunos especializados como el procesador ALFA de DEC; y con capacidades en disco que van hasta más de 10 GBytes. También pueden ser servidores de una red equipos con el sistema operativo UNIX, aunque esta es una condición especial de servidor.

- **Medio de Transmisión.** Para implantar la red de computadoras, es necesario que ellas se conecten entre sí es decir que puedan intercambiar datos de una a otra. El medio de transmisión, por lo tanto, tiene 2 componentes:

El primero es el medio de conexión misma que pueden ser un cable de 2 hilos, cable coaxial, cable óptico, o puede ser también radio frecuencia (en el caso de transmisión

de red sin hilos). El segundo componente es el dispositivo que permite conectar el cable, con cada una de las estaciones y el servidor de la red. Este dispositivo es la tarjeta de red.

- **Estaciones.** Las estaciones de una red de microcomputadores con el sistema operativo Netware, pueden ser PC, ATs, microcomputadoras tipo 386, 486, etc., así como algunas computadoras del tipo Apple Macintosh. Los requerimientos y capacidades de las estaciones estarán determinados por el tipo de programas que el usuario deberá utilizar en ellas. Dentro de cada una de las estaciones se conecta típicamente una tarjeta de red. Como requisito mínimo, las estaciones requieren de la PC, un monitor y teclado (En algunos casos, es posible eliminar el monitor y el teclado). En cuanto a dispositivos adicionales, las estaciones pueden tener o no unidades de diskette, así como discos duros, unidades de cintas para respaldo, CD-ROM, impresoras, y una multitud de otros dispositivos adicionales.

- **Servidores especializados.** Al ser la red de microcomputadoras un esquema de proceso distribuido, en algunos casos es posible identificar ciertos tipos de tareas que pueden ser centralizados en un solo lugar de la red para logra mayor eficiencia. Uno de estos casos corresponde a los procesos relacionados con el manejo de colas de impresión y las impresoras de la red. Para

esto se puede implantar un servidor de impresión, el que puede recibir en el mismo servidor de la red, o puede convertirse en un servidor externo, que se dedica exclusivamente a administrar las tareas de impresión, logrando con esto reducir las exigencias que sobre el servidor principal de la red hacen los diferentes usuarios. Otros dos ejemplos de servidores especializados son los servidores de comunicaciones y los servidores de base de datos.

#### **Pero adonde corren los programas?**

Un programa tiene 2 maneras de existir. Por un lado, mientras el programa no está siendo utilizado, reside en algún disco. Por ejemplo un procesador de texto, mientras no es llamado está grabado en un

disco para precisamente poder ser utilizado. Cuando el programa es llamado, es entonces cargado a la memoria de la estación y ejecutado por el CPU de la estación. Es decir, en la forma típica y tradicional, en el disco duro del servidor reside el programa antes de ser utilizado. Pero cuando una estación llama a un programa, es cargado desde el disco duro del servidor hasta la memoria de la estación y ejecutado por la unidad central de proceso de la estación. Es solamente en casos **muy** calificados, que un programa correrá en el servidor, como lo es en el caso que se menciona anteriormente del servidor de impresión, o cualquier otro servicio que el servidor ofrezca a los usuarios de la red.

## II El Diseño debe ser Orientado a las Aplicaciones

El diseño de la red debe partir no solo de un análisis detallado de los requerimientos del software que será utilizado en la red, sino que debe incluir también el entorno del usuario, en especial la intensidad de uso de cada una de las aplicaciones en términos de transacciones, transferencias de etc.

Los criterios de diseño que interesan especialmente al analizar el software al ser utilizado son:

- Intensidad de uso del disco (I/O)
- Intensidad del uso de impresión.

En términos simplistas, y desde el punto de vista de la estación, el servidor de archivos puede verse como un disco duro adicional (posiblemente con una enorme capacidad de almacenamiento y gran velocidad), y por lo tanto las aplicaciones usan ese disco duro ya sea para leer programas a ejecutar, leer bases de datos históricas, insertar registros nuevos en una base de datos producto de una transacción, o leer y modificar registros en la base de datos (como en el caso de actualizar saldos). Desde este punto de vista podemos establecer 2 entornos de usuario opuestos, en donde la intensidad del uso de disco duro varía en forma radical.

En el primer escenario podemos establecer una red de micros de índole gerencial en donde las aplicaciones incluyen procesadores de texto, hoja electrónica, correo electrónico, manejo de agendas, y envío y recibo de mensajes por fax. La pregunta de diseño es: Con qué frecuencia se accesa el disco de la red? Cuando se accesa el disco de la red, que cantidad de bytes son transmitidos del servidor hacia la estación o viceversa?

Tomando en cuenta las aplicaciones mencionadas anteriormente, podemos ver que aunque la red tenga una gran cantidad de usuarios, la mayor intensidad de uso del servidor se da a la hora de leer los programas que van ser ejecutados, ya que típicamente estos tienen un gran tamaño, y deben ser transferidos desde el disco duro en el servidor de la red hacia la memoria de la estación. Pero aún así no es típico en este tipo de red encontrar que los usuarios están saliendo y entrando, con gran frecuencia, entre las diferentes aplicaciones.

Durante el uso de las aplicaciones, podríamos encontrar que el procesador de texto tiene una actualización automática y periódica del archivo de texto que se está editando (para evitar que una falla de luz o

en el equipo resulte en una pérdida del trabajo realizado), o el eventual cambio del documento con el que está trabajando cada uno de los usuarios. Una situación similar se presenta con las hojas electrónicas en el sentido de que no son leídas y actualizadas muy frecuentemente. Lo mismo se aplica para el correo electrónico, etc.

La excepción a esta situación resulta ser una red con una enorme cantidad de usuarios, posiblemente más de 100 ó 150, en donde en efecto, las probabilidades de que varios usuarios estén simultáneamente tratando de leer un programa para ser ejecutado aumentan en forma importante, en cuyo caso deben utilizarse esquemas de diseño que se verán más adelante, para aumentar el rendimiento del servidor, y separar los medios de transmisión para que se reduzca el volumen de datos que es transmitido sobre el mismo medio.

Una segunda excepción son las aplicaciones que corren bajo WINDOWS, ya que WINDOWS en sí mismo es un sistema operativo que requiere de un uso intenso del disco. Más adelante se trata el tema de las redes con estaciones que usan el sistema operativo WINDOWS.

Para el segundo escenario, con una gran intensidad de transferencia de datos, podemos establecer una red utilizada por un puesto de bolsa muy activo, en donde por un lado se recibe información de las

operaciones que se estén realizando en la bolsa, en vivo, y que es transferida a los múltiples usuarios de la red para su análisis, graficación y toma de decisiones; la posterior realización de transacciones de compra y venta que deben ser registradas no solo en el servidor de la red, sino que también deben ser transmitidas a la bolsa para su ejecución; así como el uso intenso del historial de precios y volumen de diferentes instrumentos financieros para la toma de decisiones en línea al compararlos con el precio y dirección del mismo en un instante determinado.

Este entorno del usuario, y el tipo de aplicaciones, implica que no solamente es un ambiente con un alto volumen transaccional, sino que el uso en general del disco es bastante intenso.

En cuando a la intensidad del uso de impresión, hay dos componentes que deben ser tomados en cuenta al diseñar la red:

- **La intensidad de uso del disco**, y del medio de transmisión de la red, para la localización y lectura de los datos al ser impresos, y

- **La transmisión a través de la red** de los datos a ser impresos, que tienen un efecto sobre el medio de comunicación así como en el uso de colas de impresión y los dispositivos mismos de impresión.

En el ambiente de la red gerencial que mencionamos anteriormente, podemos pensar que si bien existe impresión de cartas, documentos, hojas de análisis, etc., la cantidad de hojas a imprimir típicamente no es muy grande (con las mismas excepciones mencionadas anteriormente).

Por otro lado, en el ambiente bursátil, se requiere de impresiones permanentes sobre las transacciones solicitadas, reportes históricos no sólo sobre el movimiento de

compra y venta de los clientes, sino sobre análisis históricos de los diferentes instrumentos.

Podemos ver, que ante una situación de cantidad de usuarios similar en ambos escenarios, los requerimientos sobre el ambiente computacional son muy diferentes, y por lo tanto los elementos específicos de diseño de red también resultarán diferentes, y por tanto su costo.

## El Efecto de las Aplicaciones sobre la Red

Para entender cómo diseñar en la red los componentes necesarios para obtener el resultado deseado, considerando las aplicaciones que se van a usar en la red, se explican a continuación los tipos de exigencia técnica que le imponen las diferentes aplicaciones:

Primero que nada, es importante entender lo que pasa al ejecutar un programa que se carga desde la red. Típicamente el programa mismo (el archivo EXE), reside en el disco duro del servidor. Cuando el usuario ingresa en su estación el comando para ejecutar el programa, el sistema operativo de la estación le solicita al servidor que envíe a través de la red la porción del programa que es necesario cargar. Si consideramos que algunos programas tienen tamaños de 500 Kbytes, y aún más podemos ver que cada vez que se solicita la ejecución de un programa, aunque sea por un período muy corto, se requiere transmitir a través de la red una gran cantidad de bytes. Algunos programas usan un esquema conocido como overlay, que permite ejecutar un programa de gran tamaño en un mínimo de memoria, en estos casos, al solicitar que se ejecute la aplicación, realmente se carga una porción de la misma, y posteriormente cuando el usuario solicita alguna función de la aplicación que

aún no ha sido cargada a memoria, el overlay entonces debe ser cargado también a la memoria de la estación. En este caso también se transmite a través de la red la totalidad de los bytes que componen dicho overlay. Ahora bien, aunque la cantidad de bytes que debe ser transmitida por la red al cargar una aplicación o uno de sus overlay, no se trata, en la mayoría de los casos de un requerimiento excesivo ya que la probabilidad de que una gran cantidad de usuarios estén **simultáneamente** iniciando la ejecución de aplicaciones, durante la operación normal de la red, es muy baja.

Cuando por la naturaleza de las aplicaciones o la cantidad de usuarios se da precisamente el caso que múltiples usuarios están cargando aplicaciones simultáneamente, debe considerarse la posibilidad de incluir en las estaciones alguna capacidad de almacenamiento en el disco duro para así evitar que el exceso de bytes transmitidos en la red vayan a generar una disminución del rendimiento total de la misma.

En el caso específico del sistema operativo WINDOWS 3.1, debe hacerse notar que este sistema operativo utiliza un esquema de intercambio de memoria (memory-swapping), que intercambia grandes

cantidades de bytes localizados en memoria a un área de trabajo en archivo, y viceversa. Como puede verse para realizar este intercambio de memoria, y únicamente cuando las áreas temporales de trabajo del WINDOWS se encuentran en el disco de la red, los requerimientos sobre la red serán grandes, ya que muy periódicamente existirá la transmisión a través de la red ya sea de la porción de memoria de las estaciones que está siendo trasladada al disco, o de la porción del disco que está siendo trasladada a la porción de memoria de las estaciones. Los esquemas para evitar que exista este requerimiento sobre la red incluyen la instalación de discos duros en las estaciones que van a utilizar WINDOWS, o por otro lado, la instalación de grandes cantidades de memoria en las estaciones con WINDOWS, y establecer un disco en RAM en cada una de las estaciones. Estos esquemas se hacen necesarios en ambientes de red con más de 20 usuarios **activos**, usando WINDOWS; con menos de 10 usuarios activos usando WINDOWS simultáneamente, el efecto sobre el rendimiento de la red no es extraordinario.

Los requerimientos de la estación en cuanto a RAM y de espacio temporal para "swapping" de Windows dependen de las aplicaciones que se vayan a utilizar, y de la cantidad de aplicaciones activas en diferentes ventanas simultáneamente durante la sesión Windows.

**Para el proceso transaccional** los requerimientos sobre la red, por estación,

son bastante menores que el requerimiento que implica la carga inicial de un programa, ya que típicamente la cantidad de datos involucrada en una transacción típica no supera los 4. Kbytes (por supuesto que siempre existe la excepción). La velocidad con que se procesan las transacciones es directamente proporcional a la cantidad de usuarios realizando transacciones, ya que de esta cantidad de usuarios depende la probabilidad de que 2 o más usuarios traten de realizar **simultáneamente** transacciones similares.

Cuando 2 o más usuarios tratan de acceder simultáneamente la red, típicamente se puede producir lo que se conoce como una colisión. Durante una colisión, ninguna de las transmisiones realizadas se hace efectiva, y las estaciones que intentaron transmitir entran en período de espera que es diferente para cada uno de ellos, ya que se genera aleatoriamente. El resultado es que todas las estaciones entran entonces en un período de espera, y el tiempo total de la ejecución de la transacción aumenta en forma directa. Puede esperarse que con el aumento en la cantidad de usuarios aumentará entonces la cantidad de colisiones en la red, y por lo tanto hay una porción del tiempo que se dedica a recuperarse de estas colisiones.

Un aspecto importante del manejo de colisiones es que con la mayoría de las tarjetas de red, cuando la tarjeta participa en



colisiones consecutivas, el tiempo de espera en la recuperación de la última colisión, será mayor que en la anterior colisión, por lo que al existir una situación de muchas colisiones en la red, el tiempo de respuesta se degrada no sólo por la interrupción que implica la colisión, sino que además por el tiempo adicional de recuperación generado por múltiples colisiones consecutivas.

A simple vista, pareciera que la cantidad de colisiones esperable en una red es función del medio de transmisión y de la cantidad de usuarios. Sin embargo, hay otros factores que pueden tener impacto. Si tenemos dos redes equivalentes en cuanto a capacidad de los equipos, medio de transmisión, y cantidad de transacciones generadas por las aplicaciones, A y B, pero el manejador de bases de datos usado en A es el doble de rápido que el usado en B para procesar transacciones, aunque el manejador de base de datos usado en B usa un esquema de compresión de datos, requiriendo transmitir únicamente el 25% de bytes por cada transacción, puede existir una cantidad de usuarios que en la aplicación A genere una gran cantidad de colisiones, y por lo tanto un mal tiempo de respuesta, mientras que la misma cantidad de usuarios en la aplicación B obtiene un tiempo de respuesta aceptable ya que tiene menos colisiones.

No hay duda, sin embargo, que las colisiones ocurren en el medio de transmisión, y por lo tanto, la cantidad de

colisiones depende de la forma en que las aplicaciones lo usan.

Resumiendo, la cantidad de colisiones en una red Ethernet depende de la cantidad de usuarios, pero también de la cantidad de bytes que compone cada transacción, la velocidad de transmisión de la red, las cualidades de las tarjetas de red en cuanto a capacidad de absorción temporal de paquetes para procesar, así como de las características de las aplicaciones en cuanto al tráfico generado en la red.

Los requerimientos de impresión que imponen los usuarios sobre la red, presentan exigencias de diferente tipo sobre los recursos disponibles en la red. Algunos tipos de impresión que se pueden dar son:

- **Impresión de un documento de procesamiento de texto.** En la mayoría de los casos esto no presenta grandes requerimientos de recursos sobre la red, ya que se trata de la impresión de documentos de algunas páginas. La impresión de un ambiente de procesamiento de texto puede volverse problema, sin embargo, cuando se están imprimiendo documentos de gran tamaño, se combinan diferentes tipos de letra en la impresión, y como parte de la impresión hay que enviar a la impresora la definición de esos tipos de letra (como en el caso de envío de fonts a una impresora láser) o cuando se están imprimiendo gráficos (por lo general los gráficos- en su representación

para ser impresos- tienen tamaños superiores al millón de bytes).

- **Impresión transaccional.** Cuando las transacciones que se procesan en la red incluyen la impresión de algún documento (como en el caso de comprobantes de pago, recibos, cheque, facturas, etc.), la impresión puede darse en una impresora local, en cuyo caso no tiene ningún impacto sobre la red, o en alguna impresora de la red, que implica que la cantidad de bytes que deben ser impresos son transmitidos desde la estación hasta el servidor de impresión, que lo coloca en un archivo conocido como cola de impresión, y cuando la estación termina de imprimir el documento, la cola de impresión se cierra y el servidor de impresión procede a enviar el archivo de la cola, a la impresora misma. Como puede verse, debido a que típicamente este tipo de impresión incluye una pequeña cantidad de bytes, no existe un gran impacto sobre el rendimiento de la red, excepto cuando hay una gran cantidad de usuarios, en cuyo caso es necesario diseñar cuidadosamente el servidor de impresión, así como impresoras locales y remotas.

- **Impresión de Estadísticas o Historial de la base de datos.** Presenta a la red 2 tipos de requerimiento. Por un lado, está el requerimiento de búsqueda de los datos a imprimir, mientras que por otro lado está la impresión misma de los registros.

En cuanto a la búsqueda de los datos, en la mayoría de los casos se trata de bases de datos indexadas, lo que implica que búsqueda de los datos se hará por medio de índices. Como desde el punto de vista de la estación, la base de datos que se encuentra en el servidor es como si se tratara de una base de datos que se encuentra en disco local, el proceso de búsqueda implica que se harán varias lecturas al "disco" para localizar cada uno de los registros. En realidad, cada acceso al disco representa un acceso al disco del servidor por medio de la red, lo que implica que la solicitud del dato debe ser transmitida por la red, y cada uno de los accesos de lectura del disco también deben ser transmitidos por la red.

Por otro lado, la impresión misma implica que cada registro que se imprime por un lado debe ser transmitido a través de la red, y además debe ser insertado en el archivo correspondiente de la cola de impresión, para que al terminar el proceso de impresión, el archivo de cola de impresión se imprime entonces en el dispositivo mismo.

Si tomamos entonces como ejemplo un reporte de 7.000 registros, y si el esquema de manejo de archivo o base de datos que se está utilizando implica que cada registro leído representa en realidad 3 accesos físicos al archivo, podemos decir a grandes rasgos, que se requieren por lo menos 21.000 transmisiones a través de la red (cada transmisión por lo menos del tamaño del

registro), más 7.000 transmisiones adicionales para la impresión misma, es decir que únicamente el impacto sobre el medio de transmisión corresponde a la transmisión de 28.000 registros (en lugar de los 7.000 que van a ser impresos). Por otro lado está también el impacto que sobre el uso del disco duro tiene la lectura de los 21.000 registros, así como la inserción de 7.000 registros en el archivo de impresión (en realidad se insertan más de 7.000 registros en el archivo de impresión, si tomamos en cuenta los encabezados, etc.).

Este tipo de requerimiento sobre la red, puede reducir sustancialmente la capacidad computacional total de la red, impidiendo que otros usuarios hagan un uso eficiente de los recursos disponibles. Cuando las aplicaciones lo requieren, es factible establecer un servidor especializado en la

red, que se dedica al manejo de la base de datos.

Como ejemplo, el manejador de datos conocido como Btrieve NLM (NLM: Netware Loadable Module), o Btrieve VAP en el sistema operativo Netware 2.2 (VAP: Value Added Process) que permite reducir en forma radical la cantidad de registros transmitidos por la red, aunque aumenta los requerimientos de capacidad en el servidor. Ambos son manejadores de bases de datos que corren en el servidor, y por lo tanto no requieren transmitir por la red todo lo relacionado con el mantenimiento y navegación de la base de datos y sus índices. En el ejemplo mencionado arriba, utilizando el Btrieve NLM, la cantidad de registros transmitidos por la red se reduce a un poco más de 14.000 registros.

## IV Ethernet: El Medio de Transmisión

La transmisión a alta velocidad entre las estaciones y el servidor, así como de estación a estación, es lo que hace del ambiente de redes de micros, un esquema computacional viable e integrado, que permite distribuir el proceso de tal manera, que el rendimiento total del Sistema de Red satisface las necesidades de la aplicación.

Seleccionamos para este documento el esquema de transmisión conocido como Ethernet dado su enorme popularidad a nivel mundial. Hay más estaciones de red conectadas por medio del esquema de Ethernet que por cualquier otro.

Una red Ethernet tiene uno o varios segmentos. Un segmento es una extensión de cable al que están conectados uno y más dispositivos de la red (entendiéndose dispositivos como estaciones o servidores).

El esquema que caracteriza y diferencia a Ethernet del resto de los esquemas de transmisión de red, es el hecho de que toda transmisión se hace desde la unidad que transmite, hacia la totalidad de los otros dispositivos conectados al mismo segmento de red ("broadcasting"). Es decir, que toda transmisión es "oída", por la totalidad de los dispositivos conectados al segmento Ethernet.

Para poder implantar este esquema, el estándar Ethernet utiliza un protocolo para organizar el uso que le dan al medio de transmisión el resto de los componentes de la red.

Este protocolo opera de la siguiente manera:

Antes de transmitir, la tarjeta de red "oye" el cable de la red, para ver si alguien está transmitiendo en ese momento. Si en efecto, algún otro componente de la red está transmitiendo en ese instante, entra entonces en un proceso de espera aleatorio, y reinicia el intento de transmisión. Si por otro lado, el cable de la red está "en silencio", la tarjeta de red se dispone entonces a transmitir, y envía el paquete de información al cable, transmitiéndolo de tal manera que la totalidad de los dispositivos conectados en el mismo segmento, pueden "escuchar el mensaje". La identificación del dispositivo receptor es parte del paquete de transmisión, por lo tanto todos los componentes de la red que están oyendo la transmisión buscan su identificación en la transmisión recibida, y si de hecho la transmisión le corresponde, procede entonces a "leer" la información de la red.

Este esquema deja abierta la posibilidad de que dos componentes de la red oigan el silencio simultáneamente, y de la misma manera simultáneamente intenten transmitir a

la red. Esto se conoce como una **colisión**. Al transmitir simultáneamente 2 o más componentes de la red en el segmento correspondiente, la información estará por lo tanto contaminada, para lo que Ethernet implanta entonces un esquema de detección de colisiones y recuperación. Esto quiere decir que las estaciones que reciben paquetes contaminados comienzan a transmitir una señal de bloqueo sobre la red, con lo cual las tarjetas de red que estaban transmitiendo se dan cuenta de la colisión, y entran entonces en un proceso de espera aleatorio, y reintentan transmitir el último paquete de información.

El esquema de Ethernet es un estándar, que ha sido establecido y es corregido y mantenido por un comité del Instituto de Ingenieros Eléctricos y Electrónicos de Estados Unidos, conocido como el IEEE 802.3. La popularidad del Ethernet se debe a las reglas muy claras establecidas por este comité, en un esquema de transmisión que es de propiedad pública.

La velocidad de transmisión establecida por IEEE 802.3, es de 10 Mbps.

La ventaja del Ethernet es que una vez establecida una transmisión, ésta se hace a la velocidad máxima permitida. La desventaja más importante del Ethernet, se da en que mientras más colisiones existen, hay un porcentaje del tiempo en que el medio de transmisión no puede ser utilizado, mientras

las tarjetas de red, usando el CSMA/CD, detectan y corrigen la situación de colisión.

Como se mencionaba anteriormente, la probabilidad de existencia de una colisión, y por lo tanto la cantidad de colisiones existentes en un plazo de tiempo determinado, depende tanto de la cantidad de dispositivos conectados a un segmento de red, así como de la cantidad de información transmitida y el tipo de paquetes transmitidos.

Ethernet permite transmitir paquetes de tamaño variable por lo que el peor de los casos para un medio de transmisión Ethernet, es cuando las aplicaciones requieren transmitir grandes bloques de datos, y existen muchas estaciones en el segmento, que desean transmitir simultáneamente.

El IEEE 802.3, está en el proceso de establecer varios nuevos estándares que permiten velocidades máximas de transmisión de hasta 100 Mbps, con lo que se reduce sustancialmente la probabilidad de colisión para la mayoría de las aplicaciones, y se aumenta por lo tanto el rendimiento global de la red.

El esquema Ethernet, permite actualmente 3 tipos diferentes de cableado:

- **Cableado 10Base5**. Este es el esquema bajo el cual originalmente se diseñó Ethernet y se caracteriza por permitir

distancias de hasta 500 metros en un sólo segmento de la red. Este esquema se conoce también como Ethernet estándar, y de los 3 esquemas utiliza el tipo de cable más caro.

- **Cableado 10Base2.** Este esquema utiliza un tipo de cable coaxial mucho más barato que el esquema 10Base5, pero la longitud máxima total de un segmento de cable es de 185 metros.

- **Cableado 10BaseT.** Este esquema se caracteriza por utilizar un cable similar al cable telefónico, y el largo máximo de un segmento es de 100 metros.

## 1 Ethernet Estándar (10Base5)

Las características de diseño del cable 10Base5 son:

- El largo máximo de un segmento es de 500 metros.
- Se pueden conectar un máximo de 2 repetidoras entre dispositivos conectados al mismo segmento del cable, con lo que se puede extender el tamaño máximo total de cable incluyendo las repetidoras a 4 kilómetros.
- Para conectar los dispositivos al cable de la red, se requiere de un transceptor, que se pega directamente a la red, y tiene un cable

que va al dispositivo que se va a conectar a la red. La distancia máxima entre el transceptor y el dispositivo de red es de 50 metros.

- La distancia mínima entre transceptores conectados a un cable 10Base5 es de 2.5 metros.

- La cantidad máxima de transceptores en un segmento de cable es 10.

## 2 Ethernet "Delgado" (10Base2-ThinNet)

Las características principales del esquema Ethernet 10Base2 son las siguientes:

- El cable utilizado es coaxial delgado, típicamente el estándar RG-58U.
- La distancia máxima de un segmento de cable coaxial es de 185 metros.
- Se pueden instalar hasta un máximo de 2 repetidoras en un mismo segmento Ethernet, permitiendo así un máximo de longitud del cable de 4 kilómetros.
- Los dispositivos de la red que se conectan al cable 10Base2, lo hacen a través de un conector tipo T, es decir un conector en que el cable es partido, y el conector provee tanto la continuidad entre las 2 porciones de cable, como la conexión hacia el dispositivo

de la red que se está conectando al cable.

- El esquema 10Base2 permite también conectar dispositivos al cable por medio de un transceptor. En este caso el transceptor es similar a un conector T, pero permite una distancia de hasta 50 metros entre el dispositivo de la red y el cable 10Base2.

- La distancia mínima entre conectores tipo T es de 0.5 metros.

- No deben existir más de 30 conexiones en un segmento de cable.

- Ambos extremos de un segmento de cable deben estar terminados por una resistencia de 50 ohm.

- Los conectores tipo T deben estar conectados directamente a la tarjeta de red, y no se debe utilizar cable adicional para separar el cable de la red del dispositivo que se desea conectar. Para estos casos debe utilizarse una conexión por medio de un transceptor tipo BNC.

### 3 Ethernet 10BaseT

Las características principales de diseño de los segmentos Ethernet con este tipo de cableado son las siguientes:

- La máxima distancia de un segmento 10BaseT es de 100 metros.

- Para la conexión se usa un cable tipo trenzado en pares (twisted pair), similar al cable telefónico, sin protección (unshielded), calibres 22 a 26 AWG. El calibre 24 AWG es apropiado para la mayoría de la aplicaciones.

- Los dispositivos se conectan a la red por medio de un concentrador central (hub), en una configuración de estrella, es decir cada cable va desde el dispositivo hasta el hub, utilizando el tipo de cable mencionado arriba.

- Los hubs pueden ser concatenados en forma jerárquica o de cascada, para formar redes de mayor tamaño, aunque un poco más complejas.

- Un dispositivo con una tarjeta tipo Ethernet estándar (10Base5, o 10Base2) puede ser conectado a un hub, o a la red, por medio de un transceptor (MAU).

- La mayoría de los hubs tienen también conexiones para 10Base5 y 10Base2.

- Se usan conectores tipo RJ45, similares a los conectores telefónicos modulares.

- De los 3 esquemas permitidos para el Ethernet de 10 Mbps, el cableado del esquema 10BaseT es el más barato, sin embargo es importante considerar el costo de los hubs, y MAUs, para hacer una comparación económica. Desde el punto de

vista de productividad, la diferencia entre los esquemas en cuanto a la velocidad efectiva de transmisión, es despreciable.

#### **4 Diseño de segmentos de Red 10Base**

Como se menciona anteriormente, una red Ethernet se puede componer de más de un segmento, entendiendo como segmento la porción del cable que conecta a un conjunto de dispositivos de red, los que en su totalidad, "oyen" una transmisión realizada en el segmento.

Al agregar un segmento adicional a la red, se utiliza un dispositivo conocido como puente (bridge), que oye todas las transmisiones en ambos segmentos, y decide cuáles transmisiones debe trasladar de un segmento al otro. De esta manera las comunicaciones que ocurren entre dispositivos en un mismo segmento, no son trasladadas al otro segmento de la red.

La forma más simple para implantar un puente en una red, es por medio de la instalación de varias tarjetas de red en el servidor mismo, de tal manera que el servidor también funciona como puente, disminuyendo así las probabilidades de colisión en la red.

En el diseño de la red, se debe considerar el caudal en cada segmento, de tal manera que

utilizando las técnicas de división de caudales por medio de la implantación de múltiples segmentos, se logre el objetivo de minimizar la cantidad de colisiones en cada segmento, y por lo tanto la totalidad de la red. Vea la sección de implantaciones con Ethernet más adelante, donde se presentan algunos ejemplos.

#### **5 Ethernet de Alta Velocidad**

Hasta ahora las topologías de red de mayor velocidad de transmisión son Ethernet de 10Mbps o el Token Ring de 16Mbps. Dependiendo de las características de la red, así como del comportamiento de las aplicaciones, ambos esquemas pueden ser equivalentes en el rendimiento total de la red (excepto en situaciones de alto tráfico persistente).

La amplia aceptación del esquema de redes de micros como ambiente computacional para el proceso de información ha cambiado los requerimientos de las aplicaciones sobre las redes, y en especial sobre el medio de transmisión.

Por un lado, en cuanto a tolerancia a fallas de la red, las redes Ethernet 10BaseX, y Token Ring, no ofrecen el rendimiento requerido para implantar el esquema de servidores espejo que es parte del diseño del Netware SFT-III, ya que sus velocidades no ofrecen la capacidad suficiente.



La orientación de las nuevas aplicaciones de apoyo del trabajo de oficina y de manipulación de información gráfica, por medio de los estándares de multimedia, video y sonido, hace también que los requerimientos de estas aplicaciones sobre la red sean de una proporción mucho mayor que las aplicaciones tradicionales de proceso de transacciones o de base de datos.

Otro aspecto que también impone grandes exigencias sobre el medio de comunicación de la red es el crecimiento de la cantidad de usuarios en cada red. El mismo Netware ha ido aumentando la cantidad de usuarios máximos en sus versiones recientes, permitiendo en la actualidad redes de miles de usuarios.

Esto ha motivado el desarrollo de diferentes esquemas para el medio de transmisión de más alta velocidad.

Sin embargo, si la aplicación requiere de un mayor rendimiento en el medio de transmisión, el esquema más popular que se ha utilizado hasta ahora ha sido la comunicación por medio de fibra óptica conocido como FDDI (Fiber Distributed Data Interface), que también puede utilizar cable de cobre, y que ofrece cerca de 100Mbps. Sin embargo el FDDI es un esquema relativamente caro, pudiendo costar los componentes de una red FDDI con fibra óptica, hasta 10 veces el costo de la misma

red (aunque con menor rendimiento) utilizando Ethernet.

Recientemente han aparecido varios esquemas, entre los que se encuentran 3 dentro de la familia Ethernet, que permiten aumentar de manera importante el ancho de banda o velocidad del medio de transmisión.

Existe alguna polémica sobre si uno de estos métodos pertenece al estándar Ethernet o no, debido a que cambian el esquema de transmisión simultánea utilizado por el Ethernet, y el correspondiente protocolo de detección y corrección de colisiones.

Antes de entrar en mayor detalle, debe quedar muy claro que el órgano reconocido para establecer estándares para medios de transmisión de redes de área local es el Instituto de Ingenieros Eléctricos y Electrónicos de Estados Unidos (IEEE). Dentro de este Instituto hay varios comités de los que participan fabricantes de los dispositivos, tanto para el estándar Token Ring, como para el estándar Ethernet. Particularmente en el caso de Ethernet, el comité 802.3 cuenta con cerca de 120 miembros procedentes de fabricantes y organizaciones de investigación y desarrollo. La falta de un estándar aceptado globalmente a nivel industrial, implica que el diseñador debe considerar que el uso de algunos de estos esquemas disponibles en la actualidad, implican que en el corto plazo el esquema instalado podría ser substituido por la

decisión final sobre el estándar.

Dentro del comité, hay dos grupos que han propuesto los estándares que se discuten en la actualidad, y para los que se pueden adquirir hoy en día los dispositivos necesarios para implantar una red bajo cualquiera de estos estándares.

El primer estándar, y el que aparentemente tiene mayor popularidad dentro el comité 802.3, es el llamado 100Base-X. Este estándar mantiene la filosofía de Ethernet en cuanto a la transmisión a todos los dispositivos conectados a un mismo segmento, así como el esquema de acceso múltiple al medio de transmisión por medio de detección de transmisión y detección de colisiones (carrier sense multiple access with collision detection CSMA/CD). El 100Base-X es entonces una verdadera extensión de Ethernet, con un aumento de la velocidad. La popularidad de este esquema se debe a la gran cantidad de investigación y desarrollo que se ha realizado para la fabricación de procesadores especializados del CSMA/CD de Ethernet.

Por otro lado, el segundo esquema conocido como 100Base-VG, si bien cambia la filosofía de acceso múltiple de detección de transmisiones y colisiones utilizados por Ethernet, utiliza un cableado similar al cableado telefónico. Esto ha despertado el interés de la más grande compañía telefónica de Estados Unidos, AT&T, debido a que los

recientes desarrollos de la tecnología han tendido a digitalizar todas las transmisiones de voz, y un estándar de alta velocidad para redes de área local basado en circuitos de voz podría implicar un enorme desarrollo de la telefonía.

## 6 Fast Ethernet, 100Base-X

El trabajo original del comité 802.3 del IEEE, estableció el protocolo para Ethernet independientemente de la velocidad de transmisión.

El 100Base-X, es similar al 10BaseT, en lo relacionado a la instalación física del cableado (ver arriba la especificaciones del esquema 10BaseT).

Si bien el esquema 100Base-X usa el mismo esquema "oír" antes de transmitir, y transmitir únicamente cuando no existe otra transmisión en el medio, se utilizan señales de control a frecuencias más altas (los controles sobre el medio de transmisión ocurren a una velocidad mayor), y los paquetes están espaciados por unidades de tiempo mucho menores que en el estándar 10BaseT.

Por lo tanto a pesar de utilizar el mismo tipo de instalación física, se debe poner especial atención a la instalación del cableado, ya que el esquema es altamente sensible a la interferencia generada entre una línea y otra.

Si bien este esquema mantiene los elementos negativos del Ethernet estándar, al realizar transmisiones a una velocidad mucho mayor, el resultado inmediato es que el tiempo requerido para transmitir la misma cantidad de datos en el esquema 100Base-X es mucho menor que en los 10Base, y por lo tanto ocurren 2 situaciones. En primer lugar, la probabilidad de que el medio de transmisión esté ocupado en el momento en que alguno de los dispositivo desea transmitir, se reduce en la misma proporción que la diferencia de velocidades entre los 10Mbps y los 100Mbps. Por otro lado, debido a que el tiempo requerido para completar la transmisión de un paquete de datos se reduce también de manera importante, se reduce de igual manera la probabilidad de que se dé una colisión en la red.

Sin embargo, los problemas típicos de la red no quedan eliminados, sino que han sido transferidos a un nuevo orden de magnitud. Podrían convertirse de nuevo en criterios de diseño, una vez que aumente el volumen de datos que debe ser transmitido a través de la red, ya sea porque aumenta en forma importante la cantidad de usuarios, o las aplicaciones requieren transmitir enormes cantidades de datos, como lo es el caso de transmisión de imágenes de video a través de la red que ocurrirían al implantar un esquema de teleconferencia por medio de una red de computadoras.

## 7 100BaseVG.

En el estándar Ethernet en todas las implantaciones que hemos visto hasta ahora, el medio de transmisión queda bajo el total dominio del dispositivo que está transmitiendo en un momento determinado.

El 100BaseVG, al igual que el 10BaseT y el 100Base-X, utiliza cable trenzado en pares (VG significa "Voice Grade", es decir calidad de voz, que es un cable de menor calidad y menor costo que el usado por 100BaseX) y la distribución física del medio de transmisión es similar en el sentido de que utiliza una especie de transceptores /concentradores llamados hubs.

El hub del 100BaseVG controla la transmisión, y por lo tanto evita la existencia de colisiones.

La estación o el dispositivo que desea transmitir, envía una señal de solicitud de transmisión al hub, el cual distribuye los permisos de transmisión por medio de un esquema de prioridades. Una vez que el hub envía el permiso de transmisión al dispositivo que desea transmitir, este último procede con su transmisión, la que es recibida por el hub y retransmitida a la totalidad de los dispositivos conectados al hub.

Debido a que el hub es quien distribuye los permisos de transmisión nunca se va dar una



colisión en este esquema, y por lo tanto no se desperdicia el tiempo, de hasta varios segundos, que típicamente se requiere en el Ethernet estándar para recuperarse.

## 8 Switching Ethernet Hubs

Como se menciona arriba, los atrasos producidos por las colisiones u los tiempos de espera para que se libere el medio de transmisión, son los principales problemas del Ethernet.

Un esquema intermedio entre los nuevos estándares de 100 Mbps y el ineficiente Ethernet estándar, permite conectar los dispositivos del segmento en un esquema de punto a punto.

En el centro de esta arquitectura, hay un "switching hub", que funciona como una central telefónica, estableciendo circuitos "privados" entre dos computadoras a la vez. El mismo hub, igual que la central telefónica, puede coordinar varias conexiones de pares, manteniendo múltiples circuitos privados.

En efecto, el esquema establece segmentos Ethernet de manera dinámica, haciéndolos y deshaciéndolos en función de los dispositivos que requieran conectarse.

Pero cada segmento tiene solamente dos dispositivos, y por lo tanto no se dan colisiones.

La mayoría de los Ethernet Switching Hubs funcionan para el estándar 10BaseT, por lo que el costo de implantación en una red que ya tiene 10BaseT puede ser menor que cambiar a un esquema 100BaseX o VG.

## 9 Implantaciones con Ethernet

Al diseñar la red Ethernet, se deben determinar dos dimensiones:

- Cuántos segmentos, y
- Cuántos dispositivos por segmento.

El objetivo principal es mantener la probabilidad de colisión en el segmento al mínimo, lo que depende de la cantidad de dispositivos, de la frecuencia de acceso de cada dispositivo, y del tamaño de cada transmisión.

Por dispositivo se entienden estaciones (usuarios de la red), servidores de archivos, servidores de impresión, de bases de datos, de comunicaciones, etc.

Por ejemplo, en una red en donde se combinan 8 usuarios con una aplicación de bases de datos y gran demanda por servicios de impresión, y 40 usuarios con una aplicación transaccional, un posible esquema de distribución del caudal implicaría establecer un servidor de archivos para ambas aplicaciones, un servidor de impresión para los 8 usuarios de bases de

datos, y un servidor de impresión para los 40 usuarios transaccionales.

El tráfico de la red se puede dividir en tres segmentos. Uno para los 8 usuarios de bases de datos, en el que un servidor de impresión sería un noveno dispositivo, y el servidor de la red un décimo. Los otros dos para los 40 usuarios transaccionales, 20 en cada uno. El servidor de impresión para la aplicación transaccional puede estar en el servidor de archivos, que tendría tres tarjetas de red, una para cada segmento. El servidor funciona como puente entre los segmentos. En este caso, el rendimiento de la red puede ser muy sensible a la capacidad de autonomía de las tarjetas de red en el proceso de empaque/desempaque de bloques de datos transmitidos o recibidos, así como del tamaño del buffer de transición en las tarjetas.

Cuando el diseño de la red requiere de varios

servidores, se puede implantar uno o varios segmentos dedicados exclusivamente al tráfico entre los servidores, ya sean éstos de archivos, de bases de datos, comunicaciones, de impresión, etc.

El conjunto de estos segmentos, la columna vertebral de la red (backbone), se diseña para minimizar las colisiones.

La forma más simple del backbone, es un segmento que conecta dos servidores de una red. Las colisiones se mantienen al mínimo, y el caudal del segmento se acerca al máximo posible.

Cuando se requiera, se puede utilizar una topología diferente en el backbone.

Por ejemplo, el backbone puede ser 100BaseX, y el resto de la red 10BaseT, o el backbone Token Ring de 16 MHz, y el resto de la red 10Base2.

## V Diseño del Servidor de la Red

El servidor de la red puede tener varias funciones que deben quedar muy claras durante el análisis de los requerimientos de las aplicaciones, para determinar los componentes que deben conformar el servidor.

Las características más importantes que deben ser establecidas por el diseñador de la red incluyen:

- Tipo de Procesador y su velocidad,
- Tipo de Bus de expansión,
- Tamaño de la Memoria,
- Tamaño del almacenamiento en línea (discos duros),
- Tamaño del almacenamiento fuera de línea (CD-ROM, unidad de cinta, etc.) y
- Conexión a la Red.

El refinamiento requerido para establecer el diseño del servidor, será determinado por el tipo de red que se desea implantar, y sobre todo por el ambiente del usuario y las aplicaciones.

Los diferentes componentes mencionados anteriormente tienen cada uno, una serie de

características técnicas sobre las que se pueden escoger múltiples opciones, y múltiples fabricantes. El objetivo al diseñar un servidor de la red no solo está en satisfacer los requerimientos de las aplicaciones, sino también en hacerlo al mínimo costo posible, y para ello es necesario lograr un balance de las capacidades de cada uno de los componentes, por ejemplo, es frecuente encontrar servidores con procesadores de alto rendimiento, y sin embargo no cuentan con memoria suficiente para lograr una eficiencia óptima, o cuentan con discos duros en donde la velocidad de transferencia entre el disco duro y la memoria del servidor es en términos relativos mucho más lenta que la velocidad de proceso del CPU. En estos casos, casi siempre con una inversión adicional marginal, se logra obtener un rendimiento muy superior del servidor implantado.

Para ciertos tipos de servidor algunos componentes toman más importancia que otros, y en realidad es importante entender los componentes de software que serán ejecutados en el servidor.

Por ejemplo, además del software del servidor de red que es componente del sistema operativo Netware, en el servidor de la red pueden ejecutarse también servidores

de impresión (el servidor de impresión se encuentra en el servidor de la red en la mayoría de los casos), manejador de base de datos (como por ejemplo Btrieve NLM), manejadores especializados de base de datos (como por ejemplo servidores tipo SQL), sistemas para respaldo, etc.

Además de estos elementos de diseño orientados al software que será utilizado en el servidor, pueden considerarse también los requerimientos impuestos por los manejadores (drivers) de los diferentes dispositivos que conforman el servidor, y que incluyen tarjetas de red, tarjetas controladoras de disco duro, tarjetas controladoras de unidad de cinta, topologías especiales, etc.

## 1 Conceptos sobre Servidores de Red

### Volúmenes y Discos Duros

El disco duro en un servidor de Netware se puede dividir en volúmenes, particiones y segmentos.

Un servidor de Netware 3.11 puede tener hasta 32 volúmenes Netware. Un volumen Netware es una entidad lógica que se compone por lo menos de un segmento y puede tener un máximo de 32.

Cada segmento puede abarcar un disco duro entero o únicamente un fracción del disco,

pero no puede extenderse a varios discos duros.

Debido a que el sistema operativo Netware 3.11 permite que un volumen tenga hasta 32 discos duros, es posible configurar un servidor con un volumen muy grande, en donde cada uno de los discos duros es un segmento del volumen.

Por otro lado, con un servidor en operación es posible aumentar el tamaño de un volumen de Netware 3.11, agregando un nuevo disco duro en el servidor, estableciendo una partición de Netware en el disco, y creando un nuevo segmento de volumen, el que entonces se convierte en parte del volumen existente. Además, en el caso de discos duros removibles, esto es posible realizarlo mientras el servidor está siendo utilizado por el resto de los usuarios de la red.

**Los discos duros del servidor están divididos en particiones.** Es posible que en el servidor existan múltiples discos duros, uno de los cuales se usará para cargar el MS-DOS (a pesar que durante la ejecución del Netware no es necesario contar con el DOS, éste sí es requerido para la carga inicial). Generalmente el disco duro del cual se carga el sistema operativo DOS, contiene por lo menos una partición para el DOS, y una o más particiones para Netware. El tamaño de la partición DOS debe ser entre 1 y 2 MB (únicamente en casos calificados, se requerirá más espacio



para la partición del DOS).

El sistema operativo Netware 3.11 implanta una serie de técnicas que le permite establecer procesos de transferencia de datos en múltiples canales. Es decir, cuando un volumen de Netware está compuesto por un solo disco duro, el servidor puede realizar únicamente una cantidad de I/Os por segundo; mientras que si el volumen está compuesto por 2, 3 ó 4 discos duros, el servidor puede transferir 2, 3 ó 4 veces más datos al volumen por segundo. También, cuando el volumen está compuesto de varios discos duros, el servidor se encarga de que todos los discos se llenen en una forma similar.

Cuando un archivo contenido en un volumen está en realidad esparcido entre varios de los discos duros que componen ese volumen, si uno de estos discos duros presenta un daño, el archivo resultará dañado indudablemente. Como el servidor tratará de llenar todos los discos duros de un volumen en forma simultánea y similar, y cuando los archivos contenidos en la red son base de datos (que típicamente tienen gran tamaño), existe una alta probabilidad de que una gran cantidad de los archivos que componen la base de datos estén distribuidos entre los diferentes discos duros que componen el volumen. Por lo tanto al dañarse uno de estos discos duros, las probabilidades de que la base de datos resulte dañada son bastante altas. Para la mayoría

de las aplicaciones, contar con un buen esquema de respaldos que proteja la base de datos resulta suficiente, y en estos casos será necesario substituir el disco duro dañado, y restaurar el último respaldo realizado. Sin embargo en algunas aplicaciones esto no será suficiente, y para estos casos es posible implantar un esquema de discos espejo, que mantiene una redundancia total sobre los discos de la red, y en caso de que uno de ellos falle, es substituido en forma transparente para los usuarios de la red, por su disco espejo.

Como ejemplo, digamos que es necesario implantar una red con una capacidad total de 1 GB. Si se usan 2 discos duros de 500 MB, se podrá obtener el doble del rendimiento en términos de entradas y salidas, que si se implanta este volumen con un solo disco duro, aumentando así el rendimiento del servidor, aunque también aumentando en alguna proporción la exposición al riesgo de falla del equipo. El diseñador de la red debe tomar la decisión de si este riesgo se traduce en un riesgo real de operación de la red, y reducirlo o eliminarlo implantando un esquema de discos espejo.

**Los volúmenes almacenan datos en unidades que se llaman bloques de asignación de disco (disk allocation blocks) que pueden ser configurados en uno de cinco tamaños: 4 KB, 8 KB, 16 KB, 32 KB, y 64 KB. El tamaño se define en el momento en que el volumen es agregado al**

servidor. Es factible definir tamaños diferentes para el bloque de asignación de disco para diferentes volúmenes en un servidor. Sin embargo, un volumen puede tener únicamente un tamaño de bloque. El bloque de asignación de disco representa el tamaño de archivo más pequeño que podrá ser utilizado en ese volumen.

Por ejemplo, tenemos un volumen que está configurado para tener un bloque de asignación de disco de 4 KB. Si se van a grabar en este volumen un archivo de 1.000 bytes, u otro de 3.000 bytes, ambos archivos requerirán, cada uno, un bloque de 4 KB de espacio en disco. De la misma manera, si se va a grabar un archivo de 5.000 bytes, o de 7.900, bytes ambos archivos requerirán 2 bloques de 4 KB.

Debido a que cada vez que el servidor requiere hacer una lectura del disco, lee múltiplos de estas unidades de bloques de asignación de disco, las aplicaciones que accesan archivos de datos o bases de datos de cierto tamaño, definitivamente van a tomar ventaja si el volumen donde se encuentran estos archivos es configurado con bloques de asignación de disco de mayor tamaño. Por otro lado, en otro volumen, se podrán definir bloques de asignación de disco de menor tamaño para archivos de aplicaciones de procesamiento de texto, hojas electrónicas, etc.

### **El sistema operativo Netware 3.11**

**ofrece varios esquemas de protección de datos.** La falla más seria en un sistema de red, es generalmente una falla en los discos duros, o en los dispositivos relacionados (tales como fuentes de poder en disco duros externos, tarjetas controladoras, cables de conexión etc.), ya que cada uno de los usuarios de la red depende del acceso continuo a los datos almacenados en los discos duros del servidor. Aunque un disco duro como dispositivo es un medio de almacenamiento confiable y durable, algunas áreas de la superficie magnética de los discos duros pueden, a través del tiempo, perder la capacidad de almacenar datos en forma confiable. Además existen probabilidades de que ocurran fallas más serias en los dispositivos de tal manera que la disponibilidad del disco como un todo, queda comprometida.

El sistema operativo Netware 3.11 permite diferentes esquemas de protección contra la pérdida de datos, dependiendo del nivel de riesgo aceptable para el entorno de las aplicaciones y sus usuarios.

Como esquemas básicos de protección, Netware 3.11 implanta, por default, varios esquemas de duplicidad que protegen contra el deterioro parcial de un disco duro.

Si se pierde confiabilidad en el área del disco en donde reside La Tabla de Directorio, o La Tabla de Asignación de Archivos (FAT: File Allocation Table), se pueden perder datos de

un archivo, o se puede asignar datos nuevos al archivo equivocado, o se pueden perder archivos completos. Por otro lado si se pierde la confiabilidad en el área del disco en donde se almacenan exclusivamente los datos, entonces es posible perder datos de algunos archivos.

Para proteger a los usuarios de la red contra la falla en estas áreas, Netware mantiene directorios de archivos duplicados, y redirecciona los datos desde los bloques malos, hacia bloques confiables en el disco duro. Además de esta capacidad, Netware tiene otros esquemas de protección de datos que son integrales:

- **Protección en contra de la corrupción de directorios y FAT.** Debido a que el directorio de un disco duro, y los FATs son muy importantes ya que contienen la información de direccionamiento que necesita el sistema operativo para determinar donde almacenar o leer datos, si alguno de los bloques del disco que contiene estas tablas es dañado, parte o la totalidad de los datos en el disco puede ser perdida. Netware reduce en forma importante la probabilidad de perder esta información, ya que mantiene copias duplicadas tanto de la tabla del directorio como del FAT en áreas separadas del disco duro. Si alguno de los bloques del disco en los que quedaron almacenadas las tablas originales es dañado, el sistema operativo automáticamente cambia a la copia de estas

tablas para obtener los datos que necesita. El sector malo se inserta entonces en la lista que compone "la tabla de bloques malos" del disco y los datos que contenía son almacenados en otra localización segura en el disco. Cada vez que el servidor carga el sistema operativo, se realiza un chequeo de consistencia en ambas copias de los directorios y FAT, para verificar que las 2 copias son idénticas.

La duplicación de directorios y FATs son una cualidad automática del sistema operativo Netware, y no se requieren pasos especiales durante la instalación del sistema operativo para implantar esta capacidad.

- **Protección en contra de efectos de superficie en el disco duro.** Durante la vida de cualquier disco duro, y como resultado de múltiples accesos para grabar y leer en las diferentes áreas del disco, es normal que se desarrollen defectos en áreas específicas del disco, que dan como resultado fallas al tratar de grabar o leer de ese lugar del disco. Los servidores Netware utilizan un área de su memoria que se denomina memoria caché, y que se utiliza para implantar una especie de extensión lógica del disco duro, y ya que el tiempo de acceso hacia la memoria tipo RAM es mucho menor que el acceso al disco duro, al grabar y leer de esta memoria se utiliza menos tiempo del procesador y por lo tanto el servidor tiene un mejor rendimiento. Sin embargo, cuando la memoria caché se llena,

es necesario trasladar los datos que residen en la memoria caché, hacia el lugar físico del disco duro que le corresponde.

Netware 3.11 implanta un mecanismo de lectura después de grabación, que implica que luego de grabar un dato que está en memoria, se vuelve a leer la posición del disco duro donde quedó grabado y se compara con el dato original que se estaba intentando grabar. Si los datos que se leyeron del disco no comparan con los datos que todavía están en memoria, el sistema operativo hace una serie de intentos adicionales, y si aún luego de estos intentos todavía no ha sido exitosa la grabación declara el área del disco como defectuosa.

Cuando esto ocurre, se hace un redireccionamiento del bloque de datos (que todavía está en memoria) a un área del disco que se conoce como área de redirección para arreglo en caliente (hot fix redirection area) en donde los datos pueden ser almacenados correctamente.

Cuando se define un volumen para Netware, se separan automáticamente una cantidad predeterminada de bloques del volumen, para implantar esta área de redireccionamiento para arreglos en caliente.

- **Discos Espejo.** La falla mecánica de un disco duro puede significar la pérdida total y permanente de los datos almacenados en el disco. El sistema operativo Netware

proporciona protección en contra de la falla total de discos duros, permitiendo al diseñador y al implantador duplicar los datos de la partición Netware en un disco duro, hacia particiones de Netware en 2 o más discos duros.

Cuando se establecen discos espejos, 2 o más discos duros, **en el mismo canal** son aparejados. Los bloques de datos que se escriben hacia el disco original (conocido como disco primario) también son escritos en el disco duplicado (el disco secundario). Los discos operan en tandem, ya que constantemente se almacenan y actualizan los mismos archivos en ambos discos. Si alguno de los discos llega a fallar, otro disco puede entonces continuar operaciones sin que se dé la pérdida de datos o la interrupción del servicio de la red.

Debido a que los discos espejos duplican discos en el mismo canal, no protegen en contra de fallas que pueden ocurrir a lo largo del canal entre el disco y el servidor, y esto incluye la tarjeta controladora de disco, los cables o fuentes de poder cuando se trata de discos duros externos. Un problema en el canal causaría entonces la falla de ambos discos.

- **Duplexamiento de discos.** Este esquema protege de manera similar a los discos espejo, en contra de la falla de un disco, pero protege también en contra de la falla de los elementos del canal, que incluyen la tarjeta controladora del disco, cables, y

fuentes de poder, mientras que los discos espejos protegen los datos únicamente en contra de la falla del disco duro.

Cuando se implanta un esquema de disco en espejo, los datos son escritos a ambos discos en forma secuencial, escribiendo primero en el disco primario, y posteriormente en el disco secundario. Con los discos duplexados, los mismos datos son escritos simultáneamente a todos los discos duplexados, evitando entonces la degradación del rendimiento del servidor.

Al implantar discos duplexados no sólo se evita la degradación del rendimiento del servidor, sino que también se logra un aumento en el rendimiento del servidor, ya que permite búsquedas "partidas". Esta característica del sistema operativo envía las solicitudes de lectura al disco de la pareja que pueda responder primero. También se pueden partir múltiples solicitudes de lectura entre los discos duplexados para su procesamiento simultáneo.

Si cualquiera de los componentes en un canal falla, el otro disco en el otro canal puede continuar operaciones sin que se dé pérdida de datos o interrupción en la operación.

En resumen, en cuanto al almacenamiento en discos duros, Netware 3.11 permite hasta 1.024 discos duros, con un máximo de 32 volúmenes, teniendo cada volumen un máximo de 32 segmentos de volumen. La

capacidad teórica máxima de un servidor es de 32 Tera Bytes, es decir más de 35 billones de bytes de almacenamiento en disco.

**En cuanto al uso de la memoria,** el sistema operativo Netware 3.11 requiere de una computadora que tenga una CPU 80386 u 80486. Como mínimo se requieren de 2 MB de RAM, dividida inicialmente entre el MS-DOS, el sistema operativo Netware 3.11, y la memoria caché (la memoria caché está dividida en lo que se conoce como buffers).

En versiones anteriores al Netware 3.11 los supervisores y administradores del sistema debían asignar durante la instalación la cantidad de memoria requerida para los recursos de la red tales como buffers de enrutamiento, cantidad de archivos y directorios etc. La cantidad de memoria asignada a estos recursos se mantenía fija mientras que el resto de la memoria se utilizaba para caché. En el sistema operativo Netware 3.11, aunque requiere aproximadamente 1.5 MB de memoria para el sistema operativo básico, el resto de la memoria del servidor se conforma en un grupo compartible de buffers de caché que son asignados en forma dinámica (es decir durante la operación del sistema operativo), según sean necesarios.

Cuando el Netware 3.11 inicia su operación asigna memoria de este grupo de buffers de caché para uno o varios de las siguientes recursos de la red:

- Caché de directorio,
- Procesos de Servicios de archivo,
- Turbo FAT,
- FAT,
- Buffers de direccionamiento,
- Tablas de optimización del directorio,
- Número máximo de archivos abiertos,
- Locks de archivos,
- Control de transacciones TTS, y
- Módulos Cargables Netware (NLM, Netware Loadable Modules), y
- Otros

La asignación dinámica de memoria permite que la cantidad de buffers de caché requerida para cada uno de estos recursos, crezca o disminuya según la demanda, por ejemplo la cantidad de buffers de caché requeridos para la tabla de directorio es determinada conforme el volumen aumenta en cantidad de archivos y directorios.

El esquema de asignación dinámica de memoria de Netware además, implanta un sistema de chequeos y balances. Debido a que la memoria asignada en forma dinámica

nunca retorna al grupo de memoria disponible a menos que el servidor sea apagado, los valores asignados no pueden crecer en forma automática sin restricción. Aunque el número de procesos del servidor y la memoria requerida para esos procesos crecen para satisfacer la demanda, **el servidor no inicia un nuevo proceso inmediatamente al existir una demanda.** El servidor espera entonces unos cuantos segundos para ver si algún proceso existente queda disponible, para darle servicio a la demanda. Si quedara algún proceso disponible, el servidor no inicia un nuevo proceso, que consume memoria adicional, sino que utiliza la memoria del proceso que queda disponible. Esta restricción en el consumo de memoria asegura que el servidor no asigna memoria permanentemente producto de actividad en la red súbita e infrecuente.

En resumen Netware 3.11 provee una asignación de memoria dinámica para los recursos de la red. Sin embargo dependiendo de la cantidad de recursos que se requieren para la red, y la cantidad de memoria que cada recurso consume, la asignación de la memoria total disponible en el servidor se hace dinámicamente, y por lo tanto el proceso es más eficiente en el uso de memoria que si fuera necesario establecer de ante mano la cantidad de memoria disponible para cada recurso.

Para la mayoría de los casos, la siguiente

regla de cálculo de memoria es satisfactoria: 1.5 MB de memoria básica, más 6 MB por cada 1,000 MB de capacidad de disco.

Además, si el servidor va a utilizar módulos adicionales (NLMs), deberán considerarse los requerimientos de memoria de éstos al diseñar el servidor. Por ejemplo Btrieve NLM, consume aproximadamente 2 MB de RAM como mínimo.

## **2 Arquitectura del Servidor. Procesador: 386, 486, Pentium, Alfa**

El sistema operativo Netware 3.11 requiere de un servidor cuyo CPU sea de la serie 80386, similar, o superior.

Esto quiere decir que en un computador tipo AT (CPU tipo 80286), no se puede implantar un servidor Netware 3.11, aunque sí se puede implantar Advanced Netware 2.2.

La velocidad del CPU (típicamente medida en MHz), es sin lugar a dudas un elemento crítico sobre el rendimiento total de la red, ya que al realizarse la mayoría de las operaciones de la red en la memoria del servidor, la velocidad del CPU determina la velocidad con que se realizan estas operaciones.

No se puede decir, sin embargo, al comparar un servidor de 16 MHz, con otro de 33

MHz, que la red con el segundo servidor tiene el doble del rendimiento que la red con el primer servidor, ya que existen otros elementos, aparte de la velocidad del CPU que determinan el rendimiento total de la red.

Si bien Netware 3.11 es "sensible al procesador 80486", en realidad no saca provecho de la totalidad de estas características que tiene este procesador. Sin embargo, en términos generales, los procesadores tipo 486 pueden obtener velocidades mayores que los procesadores tipo 386, y por lo tanto un procesador 486 en el servidor puede ser útil por razón de una velocidad mayor, pero también puede ser útil al proyectar hacia el futuro. Además un 486 a la misma velocidad que un 386, desde el punto de vista del sistema operativo de la red, tiene una mayor velocidad. Es decir que al comparar un 486 de 25 MHz, con un 386 también de 25 MHz, el 486 es preferible, ya que es más veloz al realizar el tipo de operaciones que requiere el sistema operativo de la red.

Algunos procesadores 486, ofrecen la capacidad de doblar velocidad (speed doubling). Estos procesadores identificados por las siglas DX2, procesan internamente (dentro del chip mismo), al doble de la velocidad de su equivalente de velocidad normal. Por ejemplo, los procesadores 486DX2-50 y 486DX-25 usan ambos un ciclo de 25 MHz para relacionar

los dispositivos externos con el CPU (memoria, etc.), pero el DX2-50 tiene una velocidad interna del CPU del doble que el DX-25. A pesar de esta aparente duplicación de capacidad, el usar procesadores DX2 incrementará el rendimiento del servidor en no más de un 30%. El incremento real depende de otros factores como lo son la intensidad del uso del disco del servidor, del medio de la red, etc.

El procesador Alfa de Digital Equipment (DEC), permite ser usado para servidor de red, y actualmente ofrece velocidades de 100 MHZ.

Dentro de la familia INTEL, quedan aún dos esquemas adicionales para mejorar el rendimiento de CPU. Uno de estos esquemas, el Overdrive, es un chip opcional que se agrega a la tarjeta madre, similar a un co-procesador matemático. El impacto final sobre el servidor de la red es similar al esquema DX-2.

El último procesador introducido por INTEL, el Pentium, que algunos han llamado 80586, por sus similitudes con el 80486, tiene rendimientos de cerca del doble que el 486 más rápido disponible actualmente, a la fecha de este documento.

### **3 Arquitectura del Servidor. Bus de Expansión: ISA, EISA, MCA, Local-Bus**

El servidor de la red se conforma de varios componentes, como se mencionó anteriormente. Algunos de estos componentes que incluyen el CPU, la memoria, y algunos casos, el circuito de control del video, residen en un componente que se conoce como tarjeta madre. Otros componentes tales como el controlador del disco duro, los controladores de interfase de red, y en ocasiones puertos de comunicaciones, se encuentran en circuitos apartes llamados popularmente "tarjetas".

Para lograr la integración entre estos diferentes componentes del servidor, es necesario que los circuitos que están en estas tarjetas individuales, pueden comunicarse tanto con el CPU, como con la memoria del servidor. Para lograr integrarlos, se utiliza lo que se conoce como el Bus de Expansión, que es en realidad una serie de conectores en la tarjeta madre, en donde se insertan las tarjetas mencionadas.

Cada conector en el bus de expansión se conoce como ranura del bus de expansión (slots). Es necesario que el servidor de la red cuente con una cantidad de ranuras de expansión que satisfaga la necesidades de diseño del servidor en términos de la cantidad de tarjetas que deberán ser insertadas en el servidor.

En cuando al bus mismo existen básicamente 3 estándares para este propósito, que son conocidos como ISA (que es el bus de



regla de cálculo de memoria es satisfactoria: 1.5 MB de memoria básica, más 6 MB por cada 1,000 MB de capacidad de disco.

Además, si el servidor va a utilizar módulos adicionales (NLMs), deberán considerarse los requerimientos de memoria de éstos al diseñar el servidor. Por ejemplo Btrieve NLM, consume aproximadamente 2 MB de RAM como mínimo.

## **2 Arquitectura del Servidor. Procesador: 386, 486, Pentium, Alfa**

El sistema operativo Netware 3.11 requiere de un servidor cuyo CPU sea de la serie 80386, similar, o superior.

Esto quiere decir que en un computador tipo AT (CPU tipo 80286), no se puede implantar un servidor Netware 3.11, aunque sí se puede implantar Advanced Netware 2.2.

La velocidad del CPU (típicamente medida en MHz), es sin lugar a dudas un elemento crítico sobre el rendimiento total de la red, ya que al realizarse la mayoría de las operaciones de la red en la memoria del servidor, la velocidad del CPU determina la velocidad con que se realizan estas operaciones.

No se puede decir, sin embargo, al comparar un servidor de 16 MHz, con otro de 33

MHz, que la red con el segundo servidor tiene el doble del rendimiento que la red con el primer servidor, ya que existen otros elementos, aparte de la velocidad del CPU que determinan el rendimiento total de la red.

Si bien Netware 3.11 es "sensible al procesador 80486", en realidad no saca provecho de la totalidad de estas características que tiene este procesador. Sin embargo, en términos generales, los procesadores tipo 486 pueden obtener velocidades mayores que los procesadores tipo 386, y por lo tanto un procesador 486 en el servidor puede ser útil por razón de una velocidad mayor, pero también puede ser útil al proyectar hacia el futuro. Además un 486 a la misma velocidad que un 386, desde el punto de vista del sistema operativo de la red, tiene una mayor velocidad. Es decir que al comparar un 486 de 25 MHz, con un 386 también de 25 MHz, el 486 es preferible, ya que es más veloz al realizar el tipo de operaciones que requiere el sistema operativo de la red.

Algunos procesadores 486, ofrecen la capacidad de doblar velocidad (speed doubling). Estos procesadores identificados por las siglas DX2, procesan internamente (dentro del chip mismo), al doble de la velocidad de su equivalente de velocidad normal. Por ejemplo, los procesadores 486DX2-50 y 486DX-25 usan ambos un ciclo de 25 MHz para relacionar

los dispositivos externos con el CPU (memoria, etc.), pero el DX2-50 tiene una velocidad interna del CPU del doble que el DX-25. A pesar de esta aparente duplicación de capacidad, el usar procesadores DX2 incrementará el rendimiento del servidor en no más de un 30%. El incremento real depende de otros factores como lo son la intensidad del uso del disco del servidor, del medio de la red, etc.

El procesador Alfa de Digital Equipment (DEC), permite ser usado para servidor de red, y actualmente ofrece velocidades de 100 MHZ.

Dentro de la familia INTEL, quedan aún dos esquemas adicionales para mejorar el rendimiento de CPU. Uno de estos esquemas, el Overdrive, es un chip opcional que se agrega a la tarjeta madre, similar a un co-procesador matemático. El impacto final sobre el servidor de la red es similar al esquema DX-2.

El último procesador introducido por INTEL, el Pentium, que algunos han llamado 80586, por sus similitudes con el 80486, tiene rendimientos de cerca del doble que el 486 más rápido disponible actualmente, a la fecha de este documento.

### **3 Arquitectura del Servidor. Bus de Expansión: ISA, EISA, MCA, Local-Bus**

El servidor de la red se conforma de varios componentes, como se mencionó anteriormente. Algunos de estos componentes que incluyen el CPU, la memoria, y algunos casos, el circuito de control del video, residen en un componente que se conoce como tarjeta madre. Otros componentes tales como el controlador del disco duro, los controladores de interfase de red, y en ocasiones puertos de comunicaciones, se encuentran en circuitos apartes llamados popularmente "tarjetas".

Para lograr la integración entre estos diferentes componentes del servidor, es necesario que los circuitos que están en estas tarjetas individuales, pueden comunicarse tanto con el CPU, como con la memoria del servidor. Para lograr integrarlos, se utiliza lo que se conoce como el Bus de Expansión, que es en realidad una serie de conectores en la tarjeta madre, en donde se insertan las tarjetas mencionadas.

Cada conector en el bus de expansión se conoce como ranura del bus de expansión (slots). Es necesario que el servidor de la red cuente con una cantidad de ranuras de expansión que satisfaga la necesidades de diseño del servidor en términos de la cantidad de tarjetas que deberán ser insertadas en el servidor.

En cuando al bus mismo existen básicamente 3 estándares para este propósito, que son conocidos como ISA (que es el bus de

expansión AT clásica, y que significa Industry Standard Architecture), MCA (el bus definido por IBM para la serie PS/2, y que significa Micro Channel Architecture), y EISA (un estándar que fue definido por un grupo de 9 de los principales competidores de IBM, y que significa Enhanced Industry Standard Architecture).

Al comparar las diferentes arquitecturas de bus es importante mantener en mente los tres criterios de diseño más importantes:

- Velocidad del bus,
- Direccionamiento,
- Tamaño de los datos.

El bus ISA, si bien fue muy adecuado para la AT clásica con una velocidad de 8 MHz, al introducirse procesadores de mayor velocidad, y al intentar instalar en bus tarjetas a una velocidad de conexión de 10 MHz o mayor, el bus comienza a fallar. Por otro lado, los datos de 16 bits, y el direccionamiento de 24 bits del bus ISA, no toman total ventaja de los procesadores 80386 y 80486, que prefieren datos de 32 bits, y direccionamiento de 32 bits.

El EISA y el MCA aumentan la velocidad del bus y la cantidad de bits, para así lograr incrementar la cantidad de datos que son transferidos a través del bus (medidos en mega bytes por segundo). Así, mientras el

bus ISA tiene una capacidad de 16 MB por segundo, el EISA tiene una capacidad de 33 MB por segundo, y MCA original tiene una capacidad de 20 MB por segundo (la especificación original del MCA ha sido aumentada recientemente con la introducción del sistema RISC System/6000 de IBM, implantando modos de transferencia de mayor velocidad que incluyen un modo del bus exclusivo de datos, y transferencia multiplexada de datos, que le dan a la nueva especificación MCA una capacidad teórica de 160 MB por segundo).

Al aumentar el direccionamiento a 32 bits, tanto el EISA como el MCA logra aumentar la cantidad de dispositivos que pueden ser conectados a un servidor. Si bien con la tecnología actual, no es lógico pensar en servidores que tengan más de 8 a 10 dispositivos conectados a través de su bus de expansión (es difícil conseguir servidores con más de 8 ranuras de expansión disponibles), la disponibilidad de una mayor cantidad de direcciones posibles en el bus facilitan la configuración interna de los dispositivos de la red, ya que los dispositivos provienen típicamente de diferentes fabricantes, lo que resulta en una mayor facilidad al instalar y configurar el servidor.

Las especificaciones EISA y MCA, también implantan un esquema conocido en inglés como **BUS MASTERING** (no existe en español una adecuada traducción para ese término), que es un concepto tomado de

arquitectura de computadoras grandes (mainframes): En este esquema el CPU y el bus de expansión son subsistemas totalmente separados, y cada uno es libre de las limitaciones del otro. En lugar de tener a un solo elemento en control del bus simultáneamente, el bus puede ser controlado por uno de varios dispositivos llamados "Bus Master". Además, existe un punto central de arbitrio para asignar el control del bus a los diferentes dispositivos que pueden tomarlo. Al independizar con este esquema la velocidad del CPU, de la velocidad del bus, se logra entonces una mayor integración entre los componentes del servidor.

El bus EISA, además, permite insertar la mayoría de las tarjetas diseñadas para el bus ISA, lo que permite incorporar las tarjetas tipo EISA para lo necesario (ya que son de mayor costo), y tarjetas tipo ISA para las funciones que no son críticas (como puertos en serie y paralelo, tarjetas de video monocromático, etc.)

**Local Bus.** Una ampliación adicional del tipo de EISA, es el estándar de facto conocido como Local Bus. Hasta ahora, hemos hablado del bus de expansión, que como se dijo antes permite la comunicación entre la CPU, y los dispositivos que se le agregan al computador. El CPU mismo, cuenta internamente con un bus interno, cuyo propósito es transferir datos entre los componentes directamente relacionados con el CPU. El bus interno de la CPU, es

típicamente mucho más rápido que el bus de expansión, ya que su velocidad es casi siempre la misma velocidad de la CPU. Sin embargo, los únicos dispositivos que hasta ahora tenían acceso a la CPU por medio del bus interno, son componentes que forman parte del diseño original de la tarjeta madre del computador, y típicamente instalados de fábrica. El esquema del Local Bus, permite que tarjetas controladoras de otros dispositivos, puedan transferir sus datos directamente al CPU, utilizando el bus interno, y por lo tanto se logra una capacidad de transferencia de datos mucho mayor que el bus de expansión.

Las implantaciones típicas del Local Bus se logran por medio de un conector adicional que "extiende" el conector típico del bus EISA. Aunque el uso más popular que en estos momentos se le está dando al Local Bus, es para tarjetas de video en aplicaciones que demandan una gran transferencia de datos entre la CPU y el monitor, han comenzado a salir ya al mercado tarjetas controladoras de disco duro, y se han anunciado tarjetas de red, que utilizan el esquema del Local Bus.

Debido a la gran velocidad con la que estos productos se desarrollan y penetran el mercado, es conveniente mencionarlos como un criterio de diseño, aunque con la importante consideración de que aún no existe un estándar claramente definido para este esquema, y por lo tanto es importante

buscar el soporte del proveedor sobre el esquema del Local Bus que se vaya a implantar.

Pero por qué es importante en el diseño del servidor la arquitectura del bus? Hay 2 componentes del servidor, que típicamente se implantan en tarjetas insertadas en los conectores del bus de expansión, y que son la tarjeta controladora del disco duro, y la tarjeta de interfase de red.

La tarjeta controladora del disco duro utiliza el bus para transferir los datos leídos del disco duro, hacia la memoria del servidor que posteriormente y cuando es necesario son enviadas a la red por medio de la tarjeta de interfase de red, y para llegar de la memoria del servidor a la tarjeta de interfase de red, debe pasar de nuevo por el bus.

De esta forma, en un bus de alta eficiencia y con un alto volumen de transferencia de datos, se obtiene una mejora en el rendimiento del servidor, y por lo tanto del rendimiento de la red.

En términos ideales, y si usamos exclusivamente el criterio de transferencia de datos como el criterio más importante en el diseño del bus, escogeríamos entonces el bus tipo MCA expandido. Sin embargo este estándar presenta todavía varios problemas, y tal vez el más importante es la falta de dispositivos que utilicen este esquema, ya que inclusive todavía no existen productos

comerciales que transfieran datos a la velocidad de 160 MB por segundo.

Por otro lado, está el costo de los dispositivos en estos estándares. Es decir el costo de una tarjeta controladora de disco duro, o de una tarjeta de interfase de red para bus ISA, es aproximadamente el 50% (y en algunas veces menos del 30%), que el costo de componentes equivalentes para los buses MCA o EISA.

Sea cual fuere el estándar de bus escogido, es importante establecer en el diseño del servidor, que las tarjetas que serán insertadas en los slots tengan el mayor tamaño posible, es decir que si bien el bus ISA permite 16 bits, y los buses EISA y MCA permiten hasta 32 bits, siempre permiten las opciones de insertar tarjetas de 8 bits en el caso de las 3 arquitecturas, y de 16 bits en el caso de las arquitecturas MCA y EISA.

Es decir, que aún si se escogiera una arquitectura de bus MCA de alta velocidad, si se le inserta una tarjeta controladora de red de 8 bits, y que permite una velocidad máxima de 8 MHz, entonces por lo menos en lo relacionado con la interfase de red, se obtiene un rendimiento muy parecido a un servidor con bus ISA, y una tarjeta de interfase de red de similares características. Es decir, si se escogen las arquitecturas EISA o MCA para el servidor, debe hacerse siempre que los dispositivos que se le vayan a agregar al servidor, sean conectados a través de tarjetas

controladoras que permitan bus de 32 bits y altas velocidades de transferencia.

#### 4 Memoria

Dos cualidades de memoria deben ser consideradas durante el diseño del servidor:

- Cantidad,
- Velocidad,

En cuanto a la cantidad, debemos recordar que el servidor Netware 3.11, asigna en forma dinámica la memoria, y por lo tanto no es necesario determinar de antemano cantidades fijas de memoria que serán utilizadas para cada uno de los recursos de la red, sino por el contrario la memoria será asignada durante la ejecución del sistema operativo, para adaptarse a los requerimientos de la red.

Como se mencionaba anteriormente, un servidor con 4 MB de RAM, puede soportar hasta 600 MB en disco duro, incluyendo el servidor de impresión corriendo en el servidor de la red, mientras que con 8 MB, se pueden implantar hasta aproximadamente 1.3 GB de almacenamiento en disco.

Si consideramos que los costos de memoria son muy bajos en la actualidad, y que los servidores Netware 3.11 permiten ejecutar en el servidor procesos especiales, entonces un criterio importante al decidir sobre el servidor

de la red, es la capacidad máxima de memoria permitida en la tarjeta madre del servidor. Esto es especialmente importante en servidores de alta velocidad, ya que en muchos casos, la velocidad de transferencia de datos entre el CPU y la memoria es diferente cuando la memoria se encuentra en la tarjeta madre, a cuando la memoria se encuentra en una tarjeta de expansión conectada a través del bus.

Otro aspecto importante sobre la memoria es la velocidad, ya que si la CPU es de alta velocidad, y la memoria no lo es, la CPU debe entrar en estados conocidos como tiempos o estados de espera, en los que el CPU no realiza ninguna operación, y está esperando a que la memoria esté disponible para ser accesada. Por lo general se deben buscar servidores con cero estados de espera en la memoria.

Con procesadores de más de 33 MHz, se requieren memorias con tiempos de acceso de menos de 30 ns, y debido a que típicamente la memoria RAM popular tiene un tiempo de ciclo entre 60 y 80 ns, esto hace que la memoria sea la mitad de rápida que el CPU. Esto implica que en una arquitectura estándar, el CPU debería esperar a la memoria durante su operación.

Alternativamente, en arquitecturas recientes se acostumbra a implantar un esquema conocido como memoria caché del procesador. Esta memoria se implanta utilizando SRAM (Static Random Access

Memory), que al no requerir refrescamiento de memoria puede correr a velocidades sumamente altas y hasta de 15 ns. Una CPU con una velocidad de 33 MHz puede entonces utilizar memoria caché entre 20 y 25 ns.

Típicamente se usa una pequeña cantidad de SRAMs, entre 8 K y 256 K, y toda transferencia entre el CPU y la memoria RAM, se hace a través de este caché de memoria, logrando entonces obtener cero estados de espera durante el proceso.

En el caso particular del procesador 80486, este cuenta con 80 K de caché interno, llamado el caché primario. Sin embargo, el mismo fabricante, Intel, ha aceptado que esta cantidad de memoria caché es menor que el óptimo, y ha sacado al mercado un "chip" adicional para implantar memoria caché especialmente diseñada para el procesador 486. Es decir que si el servidor de la red está basado en un procesador 486, es preferible utilizar una arquitectura con memoria caché adicional por encima de los 8 K del caché primario, ya que se obtiene un mejor rendimiento sobre el servidor.

## **5 Almacenamiento. Discos Duros: MFM, IDE, ESDI, SCSI**

Cuando hablamos del almacenamiento en disco duro del servidor, de nuevo son dos

los criterios que debe determinar el diseñador.

- Cantidad, y

- Rendimiento.

En cuanto a la cantidad de almacenamiento requerida en el servidor, ésta será determinada por las capacidades máximas de las bases de datos a ser utilizadas, los espacios reservados para el proceso de texto, hojas electrónicas, graficación, comunicaciones, etc., y para el sistema operativo y sus programas. Deben reservarse también espacios en disco para las colas de impresión.

El rendimiento del sistema de almacenamiento en disco depende de la velocidad con la que se pueden mover datos entre el sistema de almacenamiento y la CPU. Esta velocidad de transferencia depende a la vez de una serie de factores, de los cuales los tres más importantes son las características del disco en términos de tiempo de acceso, latencia, tiempo de cambio de pista a pista y otras similares; el estándar de interfase de disco utilizado, y el tipo de bus utilizado por la tarjeta controladora del disco.

Históricamente se ha requerido de una tarjeta controladora de disco duro, para realizar la interfase entre el disco duro y el computador. Esta tarjeta controladora

establecía los movimientos de las cabezas, puntos de grabación, puntos de control, etc., relacionados con el disco duro. Recientemente, sin embargo, la función controladora del disco duro, está siendo transferida hacia el mismo dispositivo de disco, de tal forma que cada controlador maneja únicamente un dispositivo, en oposición a la forma anterior para tener un controlador para varios dispositivos simultáneamente, que resulta en un esquema menos eficiente.

En los dispositivos de disco que cuentan con el controlador a nivel del dispositivo mismo, se utiliza en el servidor una tarjeta de interfase (ésta no es una tarjeta controladora, en términos estrictos), y se conoce como una interfase de sistemas. Es típico que los vendedores de equipo se refieran a ambos tipos de interfase como tarjeta controladora de disco duro, pero es importante establecer la diferencia a la hora de diseñar el servidor, ya que el rendimiento de las interfase de sistema es típicamente mayor que el rendimiento de las tarjetas controladoras del disco.

El tipo de tarjeta de interfase va a establecer dentro del sistema, un aspecto crítico para el rendimiento, que se conoce como el rango de transferencia, y que describe la cantidad de datos que pueden ser transferidos del disco duro hacia la memoria del servidor por segundo.

Los estándares de interfase de disco duro más utilizados en microcomputadoras son:

- **ST-506.** Esta es una interfase a nivel de dispositivo, es decir la tarjeta es el controlador del disco duro. Fue el primer tipo de interfase usado con computadoras personales, y provee un rango máximo de transferencia de datos de 625 KB por segundo, con el esquema MFM, o 984 KB por segundo, con el esquema RLL.

- **ESDI (Enhanced Small Device Interface).** Esta interfase también se implanta a nivel de dispositivo, por medio de una tarjeta controladora, y fue diseñado como un sucesor a la interfase ST-506, pero con un rango de transferencia de datos mayor, y que oscila de 1.25 MB hasta 2.5 MB por segundo.

- **IDE (Integrated Drive Electronics).** Esta es una interfase a nivel de sistema, que implanta las funciones de control y separación de datos en el disco mismo, y la tarjeta en el servidor es en realidad una tarjeta de transferencia de datos. Este estándar se conforma con el conocido como ANSI AT Attachment, que usa una variación de la definición del bus de expansión ISA, para conectar un disco duro al CPU, con un rango máximo de transferencia de datos de 4 MB por segundo.

- **SCSI (Small Computer System Interface).** Esta es también una interfase



a nivel de sistema, diseñada para aplicaciones de propósito general. Permite conectar hasta 7 dispositivos a la misma tarjeta de interfase, y usa una conexión en paralelo de 8 bits que permite un rango máximo de transferencia de 5 MB por segundo. Este tipo de interfase merece un tratamiento especial al considerar el diseño de servidores Netware 3.11.

El estándar ST-506, es obsoleto y no debe ser considerado en el diseño de un servidor Netware.

El estándar ESDI, sin embargo, y a pesar de ser un desarrollo del ST-506, puede ser utilizado aún en servidores, ya que con el diseño adecuado puede ofrecer un rendimiento satisfactorio para la aplicación. Sin embargo, cada día es más difícil conseguir en el mercado dispositivos tipo ESDI.

Una consideración del rendimiento de los discos duros del servidor, es el uso del concepto de memoria caché en la tarjeta de interfase al sistema de discos. Es decir, en algunos diseños de tarjeta de interfase, se incorpora memoria localizada en la interfase misma, para agilizar el proceso de transferencia de datos entre el disco duro y la interfase. Como resultado se puede obtener una velocidad de transferencia mayor a la del estándar en el que se esté usando la memoria caché.

Por ejemplo, en una tarjeta controladora ESDI con un MB de caché, es posible obtener velocidades de transferencia de datos entre la tarjeta controladora y el CPU, de más de 4 MB por segundo.

El estándar IDE, tiene el problema de que las especificaciones tempranas del estándar no estaban definidas claramente, y por lo tanto todavía nos encontramos con la situación de que los drivers para este tipo de discos disponibles de los fabricantes de tarjetas de interfase, presentan problemas de integración. El estándar IDE presenta también el problema, de que al aumentar los requerimientos de almacenamiento en el servidor, y al agregar un segundo disco IDE, se pueden presentar problemas de compatibilidad entre el disco anterior y el disco nuevo, ya que algunos discos de ciertos fabricantes no van a trabajar en conjunto con discos de otros fabricantes. Aunque el Instituto ANSI, ha estandarizado la conexión de los discos IDE, todavía existen en el mercado discos incompatibles. Esta es una situación que con el tiempo se resolverá por sí misma.

La interfase SCSI, es usada en una gran variedad de aplicaciones, desde discos duros y sistemas de respaldo en cintas, hasta impresoras, CD-ROMs, y dispositivos de entrada y salida especializados, tales como dispositivos de comunicaciones.

El concepto primario y su principal

(beneficio) es que el SCSI permite a varios periféricos diferentes compartir un bus común de entrada y salida.

La idea del bus SCSI, se inicia en los 60 cuando el equipo IBM 360 tenía un bus de entrada y salida que podía hablar a varios dispositivos periféricos diferentes simultáneamente.

A través de los años, este bus de IBM que se conoció en esa época como el "Canal OEM", fue modificado, mejorado y presentado al Instituto Americano de estándares ANSI, para ser adoptado como un bus estándar de entrada y salida para dispositivos periféricos. El comité ANSI rechazó la solicitud de estandarización debido a la naturaleza propietaria del canal OEM.

Al principio de los 80, Shuggart Associates, un fabricante de discos duros, desarrolló su propia interfase de entrada y salida que se llamó SASI, siglas para Shuggart Associates System Interface. Cuando se finalizó esta especificación y fue liberada para ser utilizada por varios fabricantes de disco duro, tuvo un éxito comercial. En 1982 el SASI fue presentado al ANSI como base para un estándar. El ANSI aceptó, formalizó, y extendió la especificación SASI debido a su uso extenso, y gran éxito comercial, y le cambió el nombre a SCSI (en buena medida para separar la especificación de cualquier vendedor en particular), pero no fue sino hasta el año 1986 que el ANSI

adoptó formalmente el estándar SCSI.

La gran ventaja del estándar SCSI, es que implanta un bus al cual se pueden conectar hasta 8 dispositivos. La tarjeta de interfase es uno de éstos dispositivos.

Un elemento importante en la eficiencia del estándar SCSI es la característica conocida como desconexión-reconexión. Debido a que los dispositivos conectados al bus SCSI son seleccionados en forma lógica y no física, y además en razón de que cada disco duro cuenta con su controlador íntegramente incorporado, la tarjeta de interfase hace una solicitud a un dispositivo, y en lugar de permanecer conectada al dispositivo mientras éste se alista y transmite el resultado de la solicitud de regreso a la interfase, ambos dispositivos (controlador SCSI y disco) se desconectan temporalmente, permitiéndole así a la interfase SCSI hacer solicitudes a otros dispositivos, o recibir el resultado de solicitudes anteriores. El dispositivo original será reconectado al bus, cuando esté listo para responder a la solicitud original de la tarjeta de interfase.

Cuando se diseña una configuración SCSI con varios discos duros, es importante entender que existe la probabilidad de que en algún momento los 2 dispositivos traten de acceder simultáneamente el bus. En estos casos, el conflicto se resuelve por el número de identificación del dispositivo (este número va del 0 al 7, siendo el 7 típicamente el

número de dispositivo asignado a la tarjeta de interfase). El menor número de identificación de dispositivo implica el mayor nivel de prioridad que tiene el dispositivo en el bus, de tal manera que si existe una tarjeta controladora con el número de dispositivo 7, un disco "0", y un disco "1", cuando ambos discos tratan de acceder simultáneamente el bus, el disco "0" obtendrá control del bus antes que el disco "1".

Esta teoría indica que si sobre el mismo bus SCSI, se fueran a implantar discos duros, unidades de cinta, e impresoras, no se le debe asignar un número a la unidad de cinta que sea menor a la de los discos duros, ni tampoco se le debe asignar un número a la impresora, que sea menor a la de los discos duros, o de la unidad de cinta.

En más de un estándar original SCSI, se ha introducido una extensión conocida como SCSI-2, establecido en 1990, y cuyo propósito es el aumentar la estandarización del esquema, y ampliar la cantidad de dispositivos que pueden ser instalados en el bus SCSI incluyendo escudriñadores (scanners), dispositivos de memoria óptica como los discos magneto ópticos, y dispositivos que cambian automáticamente el medio de almacenamiento como los cambiadores automáticos de discos de CD-ROM. Permite también la conexión de dispositivos de comunicaciones como interfases de red.

Por otro lado el SCSI-2, también amplía el bus de entrada/salida (este es el "lugar" por donde pasan los datos), permitiendo transferencias de 16 y 32 bits en paralelo, mientras que el SCSI-1 cuenta únicamente con la capacidad de 8 bits. Por otro lado los tiempos han sido optimizados, de tal forma que al combinar las características de mayor cantidad de bits en paralelo, con una mayor velocidad, es posible obtener transferencia de datos de hasta 40 MB, por segundo. Esta velocidad de transferencia es suficiente para mantener los buses de expansión MCA y EISA ocupados y llenos de datos sin necesidad de esperar la transferencia de datos desde los discos duros.

La discusión sobre el estándar SCSI para discos duros, ha sido más extensa con propósito, ya que en el ambiente Netware 3.11 se toma provecho de la capacidad de conexión y reconexión de los discos SCSI, y como se mencionaba anteriormente, cuando un volumen es partido en varios discos SCSI, el servidor puede hacer transferencias de datos entre el volumen Netware y la red a una velocidad mayor, directamente proporcional a la cantidad de discos que conforman el volumen.

Lo que se mencionó anteriormente es teoría. Para que funcione en la práctica, el diseñador y el implantador de la red deben garantizar que el equipo suplido cumple con el estándar SCSI deseado, ya sea SCSI-1 ó SCSI-2, pero también que los programas que

deben ser instalados en el servidor para conectar los discos SCSI (drivers), implantan las funciones deseadas de SCSI, porque de lo contrario no se obtienen las ventajas mencionadas. Además, es importante también la relación entre el disco duro y el tipo de bus, ya que nada se logra implantando un esquema SCSI de alta eficiencia, si en realidad el cuello de botella del servidor va a ser impuesto por una tarjeta de interfase de 8 bits en un bus ISA de baja velocidad.

De esta manera el esquema óptimo de disco duro de red, (exceptuando el esquema RAID, tratado más abajo) se trata de un esquema SCSI-2 implantando la facilidad de conexión, desconexión/reconexión a nivel del driver del servidor para el disco duro o para la interfase SCSI, con una interfase de 32 bits en un bus MCA o ESDI, y con un caché de memoria implantado a nivel de la tarjeta de interfase SCSI, diseñado según los requerimientos de transferencia de datos de la aplicación.

Este esquema se hace necesario únicamente para un ambiente de alto volumen transaccional y con bases de datos de extenso tamaño, además, su costo de implantación es relativamente alto en estos momentos, debido al costo del servidor ESDI o MCA, y de las tarjetas de interfase correspondiente, sobre todo al compararlos con los costos de componentes equivalentes en el estándar de bus de expansión ISA.

En términos de rendimiento, el nivel inmediato inferior mencionado arriba, es un ambiente similar en el sentido de implantar el sistema SCSI, pero esta vez con la tarjeta de interfase de 16 bits para el estándar ISA.

Los discos IDE, además de sus problemas de incompatibilidad actuales (problemas que eventualmente serán resueltos), no presentan en estos momentos la ventaja que sí representa para el SCSI el esquema de desconexión/reconexión, y que aumenta el rendimiento del servidor, cuando los volúmenes Netware están distribuidos en varios discos duros.

## **6 Almacenamiento. ARDIs (RAIDs)**

Los discos duros del servidor son los únicos componentes con partes móviles, y por lo tanto los más expuestos a fallas por razones mecánicas. Por otro lado, mientras mayor sea la capacidad requerida de almacenamiento en el servidor, y por lo tanto mayor el tamaño del disco duro, es típico encontrar varias dificultades durante la operación y mantenimiento de una red con discos duros de gran tamaño. Entre estas dificultades se encuentran los altos costos de reemplazar discos duros de gran tamaño, así como velocidades de acceso un tanto más lentas en los discos duros de gran capacidad.

Una nueva tecnología, que rápidamente se está popularizando en el diseño de servidores de redes de área local, es la tecnología correspondientes a los Arreglos Redundantes de Discos Independientes (ARDI), o en inglés Redundant Array of Independent Disks (RAID).

En teoría esta nueva tecnología resuelve tanto el problema de redundancia como el problema de rendimiento del Sistema de almacenamiento del servidor.

Es importante y crítico para el diseñador entender que no existe un estándar de la industria alrededor de los esquemas de almacenamiento de este tipo, si no que su implementación todavía es realizada por cada uno de los vendedores del dispositivo, y por lo tanto existen grandes diferencias en cuanto a las cualidades y características de los ARDIs, en especial en lo relacionado al rendimiento del dispositivo.

Un ARDI entonces es un arreglo redundante de discos independientes, lo que significa cualquier arquitectura para su Sistema de discos que combina 2 o más discos físicos estándar, y se presentan al servidor como un único dispositivo de almacenamiento, logrando entonces redundancia en los datos para implantar tolerancia a fallas. Es claro entonces que desde el punto de vista de disponibilidad de la red, y de su tolerancia a fallas, un ARDI es siempre más deseable que un único disco duro con la misma capacidad.

Existen varias características del Sistema ARDI que pueden ser utilizadas en el momento de configurar el subsistema de almacenamiento, y dependiendo de las características se determina el nivel del subsistema. Veamos:

**ARDI Nivel 0.** Este nivel no ofrece redundancia en los datos, y su principal función es ofrecer un alto nivel de rendimiento en Sistemas en donde la redundancia no es crítica.

El esquema se basa en distribuir los segmentos del disco entre la cantidad de discos que existan en el arreglo. Es decir si el servidor requiere grabar 4 segmentos de disco, en lugar de grabar los 4 segmentos en el mismo disco, como lo haría en el caso de un servidor con un único disco duro, graba el primer segmento en el primer disco, el segundo segmento en el segundo disco, el tercer segmento en el tercer disco, y el cuarto segmento en el cuarto disco. La siguiente vez que requiere grabar, comenzará en el quinto disco, y así hasta terminar la cantidad de discos en el arreglo, regresando entonces al primer disco para el siguiente segmento.

Si uno de los disco falla, el subsistema de almacenamiento en su totalidad fallará.

Podemos ver entonces que en función de la definición de ARDI que dimos inicialmente, el nivel 0 no es un verdadero ARDI, ya que

no ofrece redundancia. Sin embargo, el nivel 0 puede ser combinado con otros niveles, para ofrecer tanto redundancia, como rendimiento en el Sistema de Almacenamiento.

**ARDI Nivel 1.** Este nivel se parece mucho al esquema de discos de espejo que ofrece el Sistema Operativo Netware 3.X y 4.X. Cada disco en el arreglo, tiene un duplicado, de tal forma que cada byte que se graba en un disco, es grabado en su espejo simultáneamente. Este tipo de esquema presenta una ligera reducción en el rendimiento del Sistema de Almacenamiento cuando graba datos, pero puede presentar un enorme aumento en el rendimiento en la lectura de los datos, ya que de ambos discos se pueden leer porciones diferentes de datos simultáneamente.

Combinando el ARDI nivel 0 y el ARDI nivel 1, se pueden obtener un alto rendimiento, como redundancia en el Sistema de Almacenamiento, aunque el costo de la redundancia es idéntico al costo del sistema primario (excepto por el costo de la tarjeta controladora de discos).

**ARDI Nivel 2.** Este esquema cuenta con una cantidad de discos igual al tamaño del byte, más discos adicionales para ofrecer corrección de errores. Es decir por cada byte que se va a grabar al Sistema de Almacenamiento, el primer bit se graba en el disco 1, el segundo bit se graba en el disco

2, y así un disco para cada uno de los bits del byte. Además el esquema cuenta con discos adicionales que contienen bits de paridad, o algunos tipos más avanzados para detección y corrección de errores. Cuando uno de los disco falla, el ARDI utiliza ya sea el bit de paridad, o el código de corrección de errores, para regenerar el byte completo, y así presentar al servidor el dato correcto.

El ARDI Nivel 2 es un esquema que originalmente fue diseñado para supercomputadoras mainframes, y minicomputadoras, y en la actualidad presentan en el ambiente de redes, el mismo esquema robusto y eficiente a muy bajo costo. Sin embargo, a diferencia de los discos utilizados en supercomputadoras, mainframes y minis, los discos de PCs, típicamente ofrecen esquemas internos de detección y corrección de errores, por lo que el nivel 2 podría ser redundante, y hasta cierto punto innecesario.

**ARDI Nivel 3.** Este esquema cuenta con 2 o más discos duros para los datos, y únicamente un disco para paridad. De manera similar al nivel 2, la información es esparcida a través de todos los discos de datos, el esquema más popular para distribuir los datos es a nivel de byte, aunque también existen esquemas que lo hacen a nivel de bits, o algún otro tamaño lógico que pudiera depender del tipo de controlador del Sistema.

La redundancia se logra, de manera similar al

nivel 2, por medio del bit de paridad, así como asumiendo que los discos cuentan con su propio esquema interno de detección y corrección de errores. Esto quiere decir que al fallar uno de los discos, la información es reconstruida utilizando los datos de control contenidos en el disco de paridad.

Al distribuir por parejo los datos entre todos los discos de datos (valga la redundancia), cuando el servidor requiere grabar un bloque de datos, el controlador del ARDI emite una única instrucción de almacenamiento a todos los discos de datos, y les transfiere la información en paralelo. De manera similar, durante la lectura, el controlador del ARDI emite una instrucción de lectura a todos los discos de datos, los que transmiten la información al controlador en paralelo. Puede verse entonces que con este esquema se logra tanto la protección contra fallas, como muy alto rendimiento en la transferencia de datos para el subsistema de almacenamiento.

Sin embargo durante la transferencia de datos en paralelo, los discos del ARDI pueden ser utilizados únicamente en una transacción a la vez. Por lo tanto el nivel 3 es muy apto para redes en donde se almacenan grandes cantidades de datos, pero el nivel de proceso transaccional tampoco es muy alto, ya que existen otros esquemas de ARDIs que permiten un mejor rendimiento en el ambiente puramente transaccional.

**ARDI Nivel 4.** Como mencionábamos anteriormente, la principal desventaja del nivel 3 es su impedimento para realizar transacciones simultáneas de entrada y salida con el Sistema, por la forma en que se transfieren los datos, paralelamente entre el controlador y los discos. Con el ARDI nivel 4, también se distribuyen los datos en el momento de escribirlos a los discos, de tal forma que el primer bloque de datos (note que un bloque no es equivalente a un segmento de disco) se graba en el primer disco, el segundo bloque se graba en el segundo disco, y así hasta llegar a la totalidad de los discos de datos del arreglo, y reiniciando con el siguiente bloque de nuevo en el primer disco. Sin embargo, el nivel 4 cuenta siempre con un solo disco de paridad, que debe ser accesado para cada operación de escritura en cualquiera de los discos, lo que implica que en realidad el nivel 4 no permite transacciones de escritura al disco simultáneas. Esta es la principal desventaja del nivel 4.

Si un ARDI nivel 4 va a ser instalado en un servidor Novell, es importante que el driver para el subsistema de almacenamiento permita la lectura simultánea en forma de multitarea, de lo contrario el nivel 4 no presenta ninguna ventaja relativa para el diseño del servidor de la red.

Entonces, si el driver es adecuado, el ARDI nivel 4 puede procesar transacciones de lectura simultáneas para cada uno de los

discos del arreglo, por lo que en un arreglo con 6 discos de datos, el rendimiento de este arreglo durante la lectura, puede ser hasta casi 6 veces mayor que el rendimiento que se puede obtener con un solo disco.

**ARDI Nivel 5.** Como vimos, los niveles 1 al 4 de ARDI tienen la limitación de que únicamente se puede dar una transacción de escritura al Sistema de discos a la vez, debido a que estos esquemas utilizan únicamente un disco de paridad, y por lo tanto es un disco dedicado, es decir puede ser accesado únicamente por una transacción a la vez.

En el nivel 5, cada disco contiene tanto bloques de datos como bloques de paridad. De manera similar al nivel 4, los datos se dispersan en los discos con base en bloques de transferencia, pero cuando se graba un bloque de transferencia en un disco, su información de paridad se graba en otro disco del arreglo. Cuando uno de los discos falla, sus datos pueden ser reconstruidos con base en la información de paridad disponible en los discos restantes.

De esta manera, al eliminar un único disco de paridad, el ARDI nivel 5 permite múltiples transacciones de grabación al subsistema de almacenamiento, en una proporción de 2 a 1. Es decir como en cada operación de grabación se involucran 2 discos, el máximo de transacciones disponibles simultáneamente para grabar en el Sistema, es igual a la

cantidad de discos en el arreglo dividido entre 2.

Por lo tanto, si consideramos la comparación de un ARDI nivel 5 con 6 discos de datos, contra un único disco duro, podemos ver que el ARDI ofrece hasta 3 veces el rendimiento en grabación, y hasta 6 veces el rendimiento en lectura. Si comparamos un ARDI nivel 5 con 4 discos, podemos ver que puede tener un rendimiento 2 veces mayor al grabar y 4 veces mayor al leer.

Podemos ver entonces, que en el ARDI nivel 5, al crecer en cantidad de discos, también se crece en el rendimiento del subsistema de almacenamiento, aunque en realidad este crecimiento no es de manera lineal, como lo sugieren los ejemplos presentados anteriormente. Sin embargo, el diseño de un ARDI nivel 5, justifica utilizar discos de menor tamaño para obtener entonces una mayor cantidad de discos en el arreglo, lo que mejora la productividad del subsistema de almacenamiento.

El ARDI nivel 5 es el esquema más popular en el Sistema Operativo de Red Netware 3.X, y 4.X, ya que ambos son Sistemas Operativos de multitarea, y por lo tanto pueden sacar el mejor provecho de este esquema de discos duros.

Con ARDIs nivel 5 de cuatro discos, es posible lograr transferencias de hasta 16 MB/s.



Algunos drivers para este esquema presentan un degradación seria al sobrepasar un carga constante de más de 20 a 24 usuarios, por lo que se recomienda cautela y pruebas al seleccionar el hardware, especialmente ante la ausencia de un estándar oficial.

## **7 Almacenamiento. CD-ROM: Alta Capacidad a Bajo Costo**

Los discos conocidos como CD-ROM, son derivados de la tecnología de discos compactos (de ahí vienen las letras CD: Compact Disk), y típicamente se trata de dispositivos únicamente de lectura, pero con una gran capacidad de almacenamiento. Su uso es apto para aplicaciones bibliográficas, distribución de bases de datos históricas, estadísticas, publicaciones, almacenamiento de gráficos, etc. En la actualidad, existen 2 tecnologías básicas para implantar CD-ROMS en redes Netware, de tal forma que puedan ser compartidos por los usuarios de la red.

En uno de estos esquemas el dispositivo de CD-ROM está residente en el servidor, y es un dispositivo SCSI. Si bien la teoría de la interfase SCSI permite hasta 7 dispositivos simultáneos sobre el mismo bus, debido al desarrollo actual de los drivers tanto para los discos SCSI, como para los CD-ROM, es recomendable al implantar este esquema del CD-ROM, conectarlo a una interfase SCSI

separada y diferente de la interfase SCSI del almacenamiento principal en disco duro.

Esto implica requerimientos de direccionamiento para contar con por lo menos 2 tarjetas de interfase SCSI, así como requerimientos de memoria para que el NLM de manejo del dispositivo CD-ROM pueda correr en el servidor. La cantidad de memoria requerida para cada fabricante es diferente, aunque es típico encontrar requerimientos mínimos de 2 MB, y aproximadamente 1 MB por cada dispositivo de CD-ROM conectada.

Al residir el CD-ROM en el servidor principal de la red, este se convierte en un volumen más del servidor, y por lo tanto puede ser accesado por cualquier usuario de la red con derechos sobre el mismo.

**Precaución:** La aplicación para correr y buscar datos del CD-ROM debe ser específica para implantarse en este esquema. Sin embargo, una gran cantidad de aplicaciones de CD-ROM requieren del estándar de facto conocido como MS-CDEX.

Este estándar es típico del esquema para compartir CD-ROM en red, conocido como el esquema de redirector.

En este esquema, se establece un servidor de CD-ROM, que en realidad es una estación de la red en donde reside el dispositivo CD-ROM, el controlador SCSI, y el manejador

de dispositivo MS-CDEX.

En las estaciones de la red que van a compartir el CD-ROM, debe utilizarse también alguna extensión del MS-DOS, especializada para utilizar CD-ROM, como lo es el MS-CDEX. De esta forma, las solicitudes de acceso al CD-ROM realizadas por la aplicación que corre en la estación de trabajo, es redireccionada al servidor CD-ROM por medio del MS-CDEX, y utilizando el protocolo de la red.

El esquema de redirección es típicamente más lento que la implantación de CD-ROM en el servidor principal de la red.

Si bien la velocidad de transferencia típica de un dispositivo CD-ROM es bastante más lenta que un disco duro, se pueden obtener capacidades de transferencia de datos en los CD-ROM instalados en el servidor por medio de NLMs, que van desde 263 KB por segundo, con un usuario accediendo el disco hasta 144 KB por segundo, con 5 usuarios accediendo el disco, mientras que en los esquemas de redirección se pueden obtener capacidades menores de transferencia desde 182 KB por segundo con un usuario, hasta 42 KB con 5 usuarios simultáneos accediendo el CD-ROM.

Existen también esquemas que implantan una combinación de ambos métodos, es decir se establece un manejador del CD-ROM en el servidor principal de la red, por medio de un

NLM, se instala el CD-ROM en el servidor principal de la red, pero las estaciones requieren de la extensión MS-CDEX, para el DOS con el propósito de facilitar el acceso al CD-ROM.

## **8 Almacenamiento. Co-Procesadores de Disco**

Netware permite utilizar un esquema de manejo de los discos duros, por medio de una tarjeta de interfaz especializada, y que cuenta con su propio procesador, liberando al servidor de las tareas de manejo del disco.

Este tipo de tarjeta, conocido como tarjeta coprocesadora de disco (o DCB, Disk Coprocessor Board), permite usar esquemas de manipulación del acceso como la lectura dinámica adelantada (Dynamic Read Ahead), con los que se logran rendimientos mejorados.

La eficiencia de la porción de software de este esquema (el driver del DCB), es crítica para realmente aprovechar la capacidad ofrecida. En algunos casos, la misma tarjeta DCB con dos drivers diferentes puede tener rendimientos muy diferentes.

## **9 La Conexión a la Red**

Para conectar el servidor a la red, se usa una tarjeta de interfase, a la que popularmente se

le dice "tarjeta de red".

Un aspecto importante de diseño general de la red es el esquema de red o topología que se va a usar (Ethernet, Token-Ring, RX-Net, etc.), y el tipo de interfase de red depende de la topología. Por su popularidad, este documento trata el esquema Ethernet.

Esta sección se enfoca en la interfaz entre el servidor y la red desde un punto de vista del servidor. Más adelante se trata el tema del diseño de la red de comunicación.

Por lo general, si la red es de alto volumen, el diseñador deberá ser cuidadoso al especificar las tarjetas de red, ya que la tarjeta inadecuada puede dar al traste con la inversión en un servidor eficiente.

Aún si la red no es de alto volumen, la diferencia de costo de una buena tarjeta de red (comparado con una mediocre) es marginal al compararlo con el costo total de la red, y en la mayoría de los casos es despreciable, por lo que vale la pena analizar en detalle las características de las tarjetas de red.

Aunque las velocidades de procesamiento de los CPU (80486, Pentium) sigue aumentando (como el Pentium de 66 MHz), el tipo de bus de expansión del servidor sigue determinando la velocidad máxima con la que se pueden mover los datos entre el CPU y la tarjeta de red (y por consiguiente la

red misma).

Si el servidor tiene un bus de expansión ISA, la velocidad de datos en el bus es de 8 MHz. Existen tarjetas de red de 8 y 16 bits. Las de 16 bits son preferibles, y representan la capacidad máxima del ISA. En estos casos, es bueno especificar memoria en la tarjeta de red para buffers de comunicaciones. Así la tarjeta puede recibir datos del servidor aunque el bus esté ocupado, y puede recibir datos de la red cuando el servidor está ocupado.

Entre las tarjetas Ethernet disponibles para bus ISA, el rendimiento efectivo máximo en lo concerniente a transferencia de datos con el servidor es de cerca de 8 megabits por segundo.

Con EISA y MCA, la historia es totalmente diferente.

Para comenzar, el bus ofrece transferencia de datos, 32 bits cada vez, y a mayores velocidades que el ISA. Por otro lado, el esquema de bus mastering permite esquemas de transferencia entre la tarjeta de red y el CPU más eficientes.

Rendimientos de 33 megabits por segundo, en la transferencia entre el CPU y la tarjeta de red, son típicos. Al comparar estos rendimientos con los 8 megabits por segundo del ISA, podemos ver cómo se puede justificar un servidor con bus de expansión

EISA en lugar de ISA.

Resumiendo, las características deseables en la tarjeta de la red para el servidor, incluyen:

**Procesador en la Tarjeta.** Cuando la tarjeta de red contiene internamente un procesador especializado en descifrar los paquetes recibidos y cifrar los paquetes enviados para que estos tengan el formato correcto de la topología, el servidor no tiene que gastar tiempo en realizar este trabajo, y se lo deja a la tarjeta misma de red. Esto es especialmente importante cuando en el servidor corren además del Sistema Operativo, procesos de valor agregado (VAP), o módulos cargables de Netware (MLM), por ejemplo servidores de base de datos.

**- Memoria Caché.** La Memoria Caché en la tarjeta de red del servidor, almacena temporalmente los paquetes recibidos, mientras el servidor se desocupa del proceso que esta realizando en ese instante. De manera similar, almacena los paquetes transmitidos por el servidor hacia la red, mientras la red está ocupada. El Caché de memoria en la tarjeta es especialmente importante en las redes tipo Ethernet. Mientras más memoria tenga la tarjeta de red del servidor mejor.

**- Interfaz Tipo EISA, o Local Bus.** La interfaz EISA, de 32 bits es de gran ventaja por su capacidad de transferencia de datos, como se menciona

arriba, y es una interfaz especialmente deseable si se van a utilizar múltiples tarjetas de red dentro del servidor. La interfaz tipo Local Bus aunque es relativamente nueva, ofrece rendimientos de transferencia entre el servidor y la tarjeta aún mayores, y por lo tanto es más deseable que las tarjetas tipo EISA. Sin embargo debido a que la mayoría de las tarjetas madre de las computadoras que pueden servir como servidor de red, y que cuentan con el Local Bus, ofrecen apenas 1 ó 2 espacios disponibles para este tipo de interfaz, y por lo tanto no son muy adecuadas cuando se van a instalar más de 2 tarjetas de red en el servidor.

**- Control del Bus.** La capacidad de tomar control sobre el bus (conocida como Bus-Mastering), permite a la tarjeta de la red hacer transferencias de muy alta velocidad por medio del esquema DMA, lo cual ofrece un mayor rendimiento de transferencia en la memoria del servidor y la memoria de la tarjeta de la red.

## 8 Diseño de las Estaciones

El Diseño de las Estaciones debe ser función de las aplicaciones que vayan a correr en ellas. Dentro del esquema más efectivo de diseño y selección de equipo, el proceso debe iniciarse en un análisis de los requerimientos de las aplicaciones.

Ya que las estaciones de la red son microcomputadoras, nos enfrentamos, al diseñar la red, con una búsqueda de la forma óptima de integrar los recursos de las estaciones a la red.

En primera instancia, debido a que la estación requiere de 2 componentes de software para conectarse a la red, estos componentes de software van a consumir memoria RAM, de la estación, reduciendo así la memoria disponible para los programas que serán utilizados en ella.

Por otro lado también está la diversidad de Sistemas Operativos disponibles en la actualidad, y las oportunidades y problemas que presentan cada uno de ellos para integrarlos en forma eficiente a la red.

En cuanto al rendimiento de la estación y su relación con la arquitectura de la CPU, en términos generales, a mayor velocidad del CPU, se puede obtener un mayor rendimiento. Sin embargo, debe lograrse un balance entre el tipo de procesador (286,386,486, etc), su velocidad (8,10,12,16,20,25,33, ó 50 MHZ), el tipo de arquitectura de red y la tarjeta de interfaz a la red, y el tráfico esperado en la red.

Por ejemplo, en una red de alto tráfico y baja velocidad, la capacidad de transferencia de datos entre la estación y la red queda reducida, y por lo tanto el beneficio marginal de un procesador más eficiente se diluye, ya

que si bien la aplicación puede aprovechar esa mayor eficiencia en su proceso interno, en realidad no tiene un buen rendimiento de transferencia de datos. Esto es un esquema similar a diseñar una microcomputadora basada en un procesador 486 de 50 MHZ, con un disco duro de 50 ms de tiempo de acceso, ya que si bien el procesador es rápido, el disco es lento, y por lo tanto la inversión adicional requerida para un equipo más eficiente se diluye.

#### **La Relación de la Estación con la Red**

Como se menciona en el apartado "El diseño debe ser orientado a las aplicaciones", la relación de la estación con la red se mide en términos de la carga de transmisión que impone sobre el cable (u otro medio de comunicación), así como los requerimientos de recursos de los servidores que puedan existir en la red, tales como el servidor de archivos, servidor de impresión, servidor de bases de datos, etc.

También, en la estación pueden residir programas o recursos de equipo que son utilizados desde otras estaciones de la red (sin que por ello la estación se convierta en un servidor especializado), como en el caso de las impresoras "remotas", y la comunicación tipo "peer to peer" (típica de las aplicaciones de correo electrónico).

#### **La Estación Ideal**

Si bien es cierto que La Estación Ideal depende del Sistema Operativo y de las

aplicaciones que se vayan a utilizar en ella, a continuación se presenta la descripción de una estación de alto rendimiento que permite ser utilizada en múltiples ambientes.

- **El procesador ideal de estación en estos momentos es un procesador 486 DX2** de 66 MHz. - Esto significa que tiene un reloj externo de 33 MHz, pero sus operaciones internas las hace con un reloj de 66 MHz. Cuenta además con un co-procesador matemático, lo que acelera las hojas electrónicas, algunos manejadores de base de datos, y algunas implantaciones particulares de programas especialmente de orientación matemática. El co-procesador matemático también ayuda en ciertos ambientes de desarrollo de Sistemas.

- **8 MB son suficientes para la mayoría de las aplicaciones incluyendo windows.** Es importante que se cuente con suficiente memoria caché, por lo menos 128 KB.

- **Bus EISA con Local Bus.** El Local Bus ofrece el mejor rendimiento en aplicaciones que demandan mucho sobre los despliegues del monitor, mientras que el bus EISA se está convirtiendo rápidamente en el estándar industrial para interfaces de alta velocidad con sus 32 bits.

- **Tarjeta de Red de 16 bits.** En estos momentos, el precio de una tarjeta red de 32 bits es casi el doble que una tarjeta de 16

bits, mientras el rendimiento que le ofrece a la estación, no es en la misma proporción. Para la mayoría de las aplicaciones, las tarjetas de 16 bits con su manejador de software (driver) adecuado, y siempre que tenga un rendimiento también adecuado, será suficiente para la mayoría de las aplicaciones. Además es importante que la tarjeta de red cumpla con los estándares requeridos por la mayoría de los Sistemas Operativos, por ejemplo la tarjeta NE2000, de Novell, y las tarjetas de la compañía 3 Com, son permitidas por la mayoría de los Sistemas Operativos de estación, incluyendo Windows, Netware, otros Sistemas de red, y UNIVELL.

- **Sin Discos.** Si las estaciones no cuentan con unidad de diskette se protege al Sistema completo de 2 problemas potenciales muy serios. Por un lado, al no permitir la lectura de diskette, se limita de manera importante las probabilidades de que un virus ataque la red. Por otro lado, si en la red se mantiene información crítica para la operación de la organización o negocio que la utiliza como Sistema de información, no es posible extraer por medio de diskettes los datos críticos de la organización, limitando así la probabilidad del robo de datos. Además de esto, es importante considerar implantar los Sistemas de seguridad que ofrecen el Sistema Operativo de la red. Por otro lado y debido a la dificultad de respaldar múltiples discos duros en la red, es mejor diseñar una red con

múltiples estaciones sin discos duros, y considerar en la capacidad del Sistema de almacenamiento del servidor los requerimientos de los usuarios para sus archivos privados. Una posible excepción a esta regla puede ser la implantación del Sistema Operativo Windows, o utilizando algunas aplicaciones, en donde el disco duro local podría ayudar a disminuir la cantidad de paquetes enviados a través de la red para el manejo de la memoria virtual que implanta el windows. La mayor parte del tiempo, el supervisor debería ingresar a la red por estaciones sin discos, ya que debido a que el supervisor tiene derechos de acceso a la totalidad del Sistema de almacenamiento de la red, es posible que el mismo supervisor, por error, pueda introducir un virus a la red. Las estaciones con unidad de diskette deben existir en la red únicamente para propósitos de instalación de aplicaciones, o para el uso de la aplicaciones que así lo exijan.

### 3 Laptops y Notebooks

Para conectar una computadora compacta a la red, existen varias alternativas, dependiendo de la funcionalidad de la conexión, y las características de la computadora.

En primera instancia, la manera más simple es por medio de una interfaz de puerto en paralelo a la red. Este tipo de interfaz existe tanto para Ethernet 10base2 (coaxial) como

10baseT (cable telefónico), así como para Token Ring, y se compone de una pequeña caja que se conecta directamente al puerto paralelo de la computadora.

Algunos de estos dispositivos aprovechan las cualidades de la tecnología paralela extendida (EPT-Extended Parallel Technology), que permite obtener rendimientos de transferencia en la red de hasta el doble de bytes por segundo, que si no usaran esta tecnología. Para sacar provecho de este esquema, tanto la interfaz como la computadora deben contar con EPT.

Dos características adicionales deben ser consideradas al seleccionar este tipo de interfaz para la red:

- **Fuente de poder.** Algunas de estas interfaces obtienen su potencia por medio de un transformador que se conecta a un enchufe de corriente eléctrica normal, mientras que otro tipo de interfaces obtienen su potencia de la misma computadora por medio del conector para teclado externo. El primer esquema implica que se pierde un tanto la portabilidad ya que hay que trasladar el transformador, que le agrega peso. En el segundo caso, es más liviano, y más fácil de instalar al conectarse a la red

- **Disponibilidad del puerto paralelo.** Algunas de estas interfaces, permiten que el puerto paralelo de la computadora también pueda usarse para

conectar otros dispositivos, como la impresora, a la computadora.

El bus interno tipo PCMCIA es una alternativa adicional para conectar una Notebook que esté equipada con este tipo de bus. En este caso, la "tarjeta" de red es un dispositivo muy pequeño, de un tamaño similar a una tarjeta de crédito, que se puede instalar y remover fácilmente de la Notebook.

Por último, existen dispositivos llamados "docking station" en los que se introduce la computadora Notebook, y que contiene un bus de expansión, así como espacio adicional para diskettes, discos duros, CD-ROM, etc. Una vez introducida la Notebook en el docking station, el conjunto funciona como una computadora de escritorio, uniendo los recursos de la Notebook con los del docking station. En el docking station se puede instalar una tarjeta de red (como en cualquier computadora), conectando la Notebook a la red. La mayor desventaja de este esquema radica en que sólo se puede conectar la Notebook a la red desde el docking station.

#### 4 WINDOWS 3.1

Desde el punto de vista de la implantación de estaciones con el Sistema Operativo Windows 3.1 en una red cualquiera, hay dos elementos que tienen un gran impacto sobre la eficiencia de las estaciones, y sobre el

comportamiento de la red durante la operación de estas estaciones.

**Los programas relacionados con Windows, tanto de las aplicaciones como del Windows mismo, pueden estar localizados en el servidor de la red, o el disco duro de la estación. Si las aplicaciones están instaladas en el disco duro de la estación, no existe ningún impacto sobre la red, cada vez que el usuario llama al Sistema Operativo, o a cualquiera de las aplicaciones y sus opciones. Sin embargo, es importante entonces considerar que la estación debe contar con todos los dispositivos necesarios para poder lograr la instalación del software de manera correcta. Una forma popular de realizar estas instalaciones es por medio de una instalación inicial de los programas en el servidor, y la posterior copia de los mismos al disco duro de la estación por medio de la red. El diseñador o instalador deben estar conscientes de las implicaciones de legalidad de realizar este tipo de instalación, en función de la licencia del software.**

Si los programas están instalados en el servidor, el impacto sobre la red se da únicamente cada vez que el usuario inicia la sesión de windows, o solicita una opción de alguna aplicación que implica la lectura del programa desde el servidor de la red, lo que resulta en una transferencia de ese programa desde el servidor hacia la estación a través de la red. El impacto real que podría tener este



esquema sobre la red depende del tamaño de los programas (y esto depende a la vez de la aplicación misma), así como de la cantidad de usuarios que simultáneamente están cargando programas en las estaciones. En la mayoría de los casos, el impacto sobre la red de la carga de programas windows, es tan pequeña que se puede ignorar. El escenario de peor situación, es la red de aplicaciones de bases de datos con gran cantidad de datos siendo transferidos en la red (por ejemplo datos estadísticos, o multimedia), combinada además con una gran cantidad de usuarios windows. Como regla general, en este tipo de escenarios no deben existir más de 5 usuarios windows por segmento Ethernet 10BaseX.

Si la carga de aplicaciones windows en combinación con la cantidad de usuarios windows en la red se convierte en un problema, se debe entonces diseñar el esquema windows de tal forma de que las aplicaciones son cargadas desde un disco duro en la estación, para aquellos casos en donde el problema es crítico.

El otro tema importante, es el **relacionado con los archivos temporales de trabajo** utilizados por Windows 3.1. Este Sistema Operativo implanta un esquema similar al de memoria virtual, que le permite mantener activas múltiples aplicaciones simultáneamente. Cuando una aplicación tiene un requerimiento de memoria pero se encuentra en estado "adormecido" (es decir

no está realizando ningún proceso), pero está activa en una ventana, si existe alguna otra aplicación que requiere utilizar la memoria que está siendo ocupada por la aplicación adormecida, el windows toma la imagen de memoria de esa aplicación y la traslada a una área de trabajo en archivo.

Si la estación cuenta con suficiente memoria para satisfacer la totalidad de las necesidades de todas las aplicaciones que estarán activas en un momento determinado, nunca será necesario utilizar los archivos temporales de trabajo de windows. Si por el contrario, las estaciones no cuentan con suficiente memoria para la totalidad de las aplicaciones, el windows automáticamente estará realizando un proceso de entrada y salida (grabación y lectura) a estos archivos temporales.

Si la red cuenta con muchas estaciones Windows 3.1, sin suficiente memoria para la totalidad de las aplicaciones que corren en cada estación, entonces la localización de estos archivos temporales debe ser considerada durante el diseño de la red.

Si los archivos temporales se encuentran en discos duros localizados en cada estación, entonces no existe efecto sobre el medio de transmisión. Si por el contrario las estaciones son estaciones sin disco, y los archivos temporales de trabajo del windows están localizados en el servidor, entonces debe considerarse el tráfico en la red que implica la grabación y lectura a estos

archivos.

Como regla general, mientras mayor sea la capacidad de memoria de la estación, menor será el impacto sobre el tráfico de la red que tengan estos archivos temporales. Por otro lado los archivos temporales son accedidos en la mayoría de los casos, cuando el usuario cambia de una aplicación a otra y la estación no cuenta con suficiente memoria para satisfacer las necesidades de ambas aplicaciones (como se explicaba anteriormente). Por lo tanto, el impacto de los archivos temporales de trabajo de windows sobre la red, dependerá también del comportamiento de los usuarios en la utilización de sus aplicaciones.

Para diseñar la estación ideal para Windows 3.1, se debe entonces conocer los requerimientos de memoria de todas y cada una de las aplicaciones windows **que estarán activas simultáneamente** durante una sesión windows. Si las estaciones cuentan con suficiente memoria para satisfacer esta necesidad, el diseñador puede descontar el impacto de los archivos temporales de su proceso de diseño de la red.

### **Recomendaciones Generales para la instalación de Windows 3.1 en un servidor**

El Windows 3.1 ofrece varias opciones de instalación que son particulares para el caso mencionado arriba, donde el windows va ser utilizado por varios usuarios de una red, y va

a estar instalado en el servidor.

Hay 3 opciones que el instalador debe conocer para facilitar y acelerar su trabajo:

- **Instalación inicial (opción /a).** La instalación inicial de windows en el servidor es utilizada para copiar todos los archivos que componen el Sistema Operativo Windows, de los diskettes de distribución al servidor de la red. Esto se conoce como una instalación administrativa en la documentación de Microsoft, y cuando se realiza el proceso **NO INSTALA** una versión del Windows 3.1 que se pueda utilizar de inmediato, ya que requiere de procesos adicionales para especificar las características de las estaciones que lo van a utilizar.

- **Configuración de Red (opción /n).** La opción de configuración de redes utilizada desde alguna estación, cuando se desea preparar esa estación particular para utilizar windows, cargando el Sistema Operativo desde el servidor de la red. Este proceso instala únicamente los archivos del windows donde se especifican las características de la estación y las preferencias del usuario, y donde posteriormente también se indicarán las aplicaciones que serán utilizadas. Por ejemplo, instala los archivos WINDOWS.INI, Y SISTEM.INI. Estos archivos son instalados en un directorio privado para el usuario en la red, o un

directorio en el disco duro de la estación si el usuario así lo desea.

**Configuración en Grupo (opción /h).** La opción de configuración en grupo permite instalar windows para múltiples estaciones considerando los valores default que el instalador desea. Para esto, una vez instalado el windows con la opción de configuración administrativa (ver arriba), se puede modificar el archivo SETUP.SHH, que se encuentra en el directorio de windows

en el servidor. Por ejemplo, para configurar todas las impresoras de la red, el instalador puede crear un archivo llamado PRINTER.SHH, que contiene la configuración de las impresoras de la red e impresoras locales, y posteriormente en cada configuración de cada estación se pueden utilizar de manera combinada la opción H y la opción N, por ejemplo con el comando SETUP /h:printer.shh /n.

## Bibliografía

**Davis Dwight [1988]:** Benchmarking the Impacts of Distrinuted Database Architecture. *Netware Technical Journal* Octubre 1988.

**Hoarty Leo [1984]:** Philosophy of Local Area Networking, *Microsystems* volumen 5 número 8, agosto 1984

**Computer Networks: A Tutorial** IEEE Computer Society 1986

**Network Computing - The New Phase** en PDS Bullets volumen 3 número 7..

**Netware 3.11 Concepts** Novell 1991