

DISEÑO DE UNA RED DISTRIBUIDA DE DATOS
PARA LA UNIVERSIDAD DEL VALLE DE GUATEMALA



UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ciencias y Humanidades



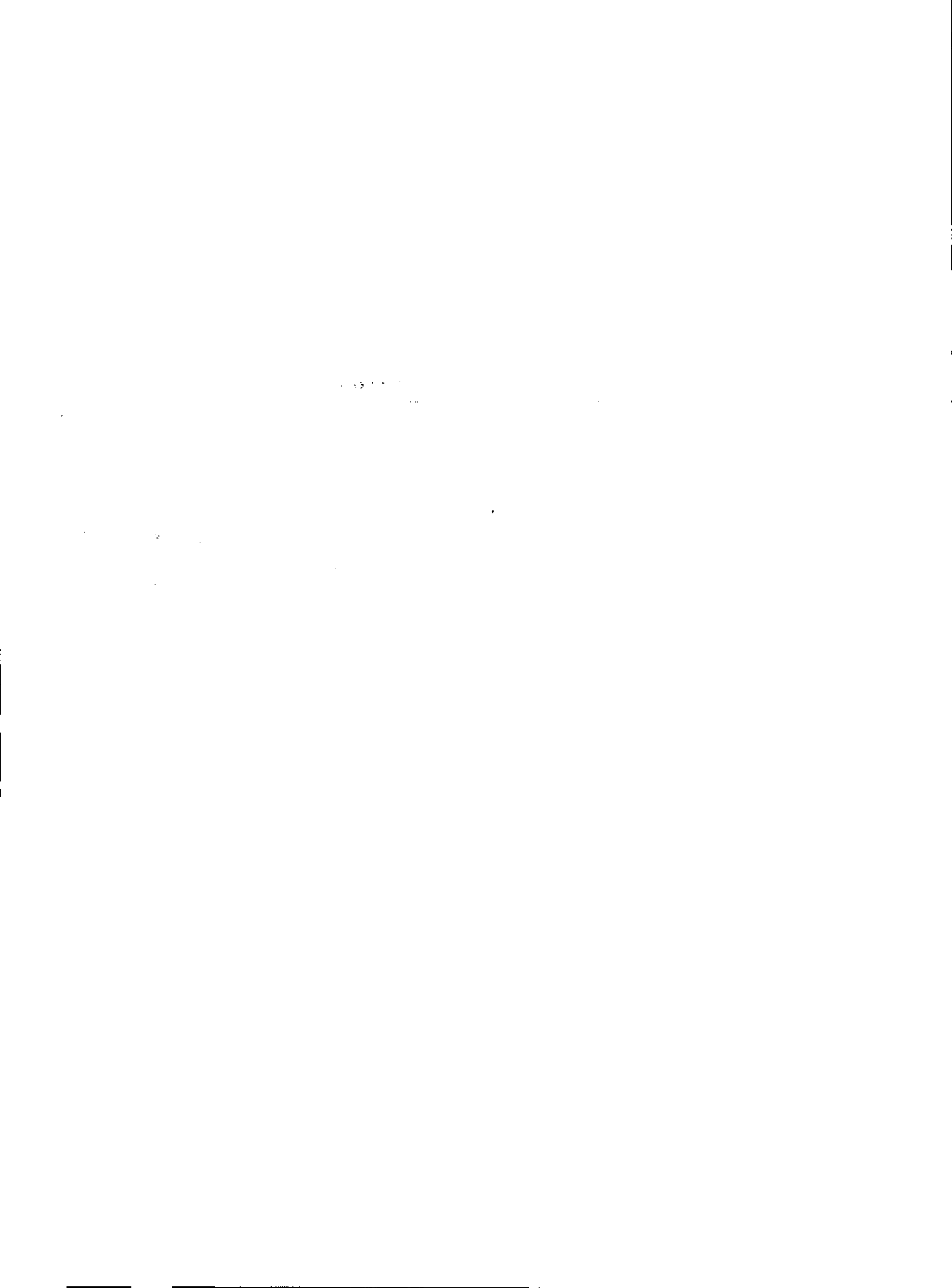
DISEÑO DE UNA RED DISTRIBUIDA DE DATOS
PARA LA UNIVERSIDAD DEL VALLE DE GUATEMALA

JUAN PABLO RODRIGUEZ VELIZ

Trabajo de investigación presentado para optar
al grado académico de Licenciatura en
Ingeniería Electrónica

Guatemala

1990



Vo. Bo. :

(f) *Ricardo Cordón*
- Ingeniero Ricardo Cordón
Asesor

Tribunal:

(f) *Ricardo Cordón*
Ingeniero Ricardo Cordón

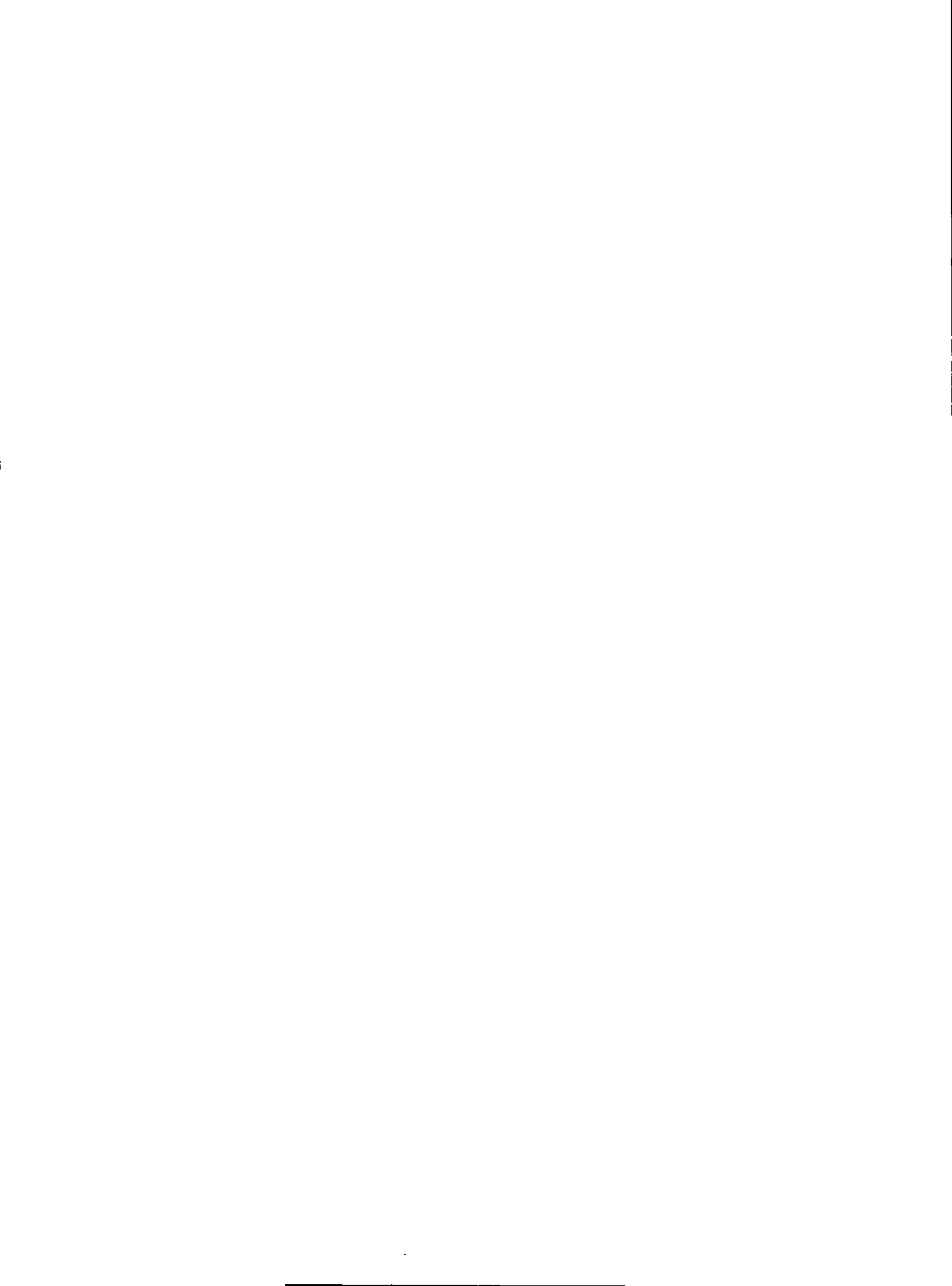
(f) *Roberto Tejada*
Ingeniero Roberto Tejada

(f) *Kari Campos de Fremme*
Ingeniera Kari Campos de Fremme

Fecha de aprobación: 22 de noviembre de 1990

A Dios

A mi esposa



PREFACIO

En este trabajo se realiza un estudio de los sistemas distribuidos: sus características, ventajas y desventajas, posibles variantes, etc.

Se describen, detalladamente, los lineamientos a seguir para la realización del diseño de una red distribuida de datos.

Finalmente se estudian las características y las necesidades computacionales de la Universidad del Valle de Guatemala y se diseña una red distribuida de datos.

CONTENIDO

	Páginas
PREFACIO	IX
I. INTRODUCCION	1
II. PROCESAMIENTO DISTRIBUIDO	3
A. Características de los sistemas distribuidos	4
1. Múltiples unidades de procesamiento	4
2. Imagen de un solo sistema	5
3. Interconexión electrónica	7
4. Libertad en el diseño	7
B. Objetivos del procesamiento distribuido	10
1. Organizacionales	10
2. Técnicos	12
C. Alternativas de procesamiento distribuido	14
III. MOTIVOS PARA LA DISTRIBUCION Y COMPARACION ENTRE UN SISTEMA CENTRALIZADO Y DESCENTRALIZADO	19
A. Razones para la distribución	19
1. Manejo centralizado de la red	19
2. Costos de comunicación	20
3. Confiabilidad	20
4. Mejorar interfaces para los usuarios	22

	Páginas
5. Seguridad y privacidad	24
6. Economías de dedicación	24
7. Crecimiento incremental y flexibilidad	26
8. Limitación de capacidad	27
9. Aumentar la facilidad de instalación	27
10. Corresponder a la organización de la empresa	28
B. Sistema centralizado y descentralizado	28
IV. TRANSMISION DE DATOS	33
A. Niveles de comunicación	40
1. Nivel intracomputadoras	40
2. Redes de área local	43
3. Teleproceso	51
B. Arquitectura de comunicaciones	62
V. REDES DE COMPUTADORAS	67
A. Objetivos	68
B. Clasificación de las redes de computadoras	72
C. Topología	75
D. Enrutamiento	81
E. Ejemplos de redes de computadoras	82

	Páginas	
VI.	DISTRIBUCION DE DATOS	85
	A. Criterios para la localización de los datos	87
	1. Centralizado	87
	2. Dividido	88
	3. Replicado	91
	4. Híbrido	92
	B. Sincronización	92
	1. Sincronización en un sistema centralizado	93
	2. Sincronización en un ambiente distribuido	97
VII.	DISTRIBUCION DE FUNCIONES	111
	A. Funciones necesarias para realizar una transacción	111
	B. Configuración de los nodos	117
VIII.	DISTRIBUCION DE CONTROL	125
IX.	LINEAMIENTOS PARA EL DISEÑO DE UNA RED DISTRIBUIDA DE DATOS	131
	A. Análisis de las necesidades del usuario	131
	1. Recolección de la información	133
	2. Análisis de la información	134

	Páginas
B. Diseño del sistema	136
1. Diseño lógico	138
2. Diseño físico	139
C. Análisis de la red	140
D. Determinación de los requerimientos del sistema	141
E. Comparación y elección entre las diferentes opciones de acuerdo a los criterios de selección	144
F. Cronograma de implementación	149
X. DISEÑO DE UNA RED DISTRIBUIDA DE DATOS PARA LA UNIVERSIDAD DEL VALLE DE GUATEMALA	151
A. Análisis de las necesidades del usuario	151
1. Recolección de la información	151
2. Análisis de la información	180
B. Diseño del sistema	183
C. Requerimientos del sistema	187
D. Elección entre las opciones	194
XI. CONCLUSIONES	197
XII. BIBLIOGRAFIA	201

	Páginas
ANEXO	
A. RED DIGITAL DE SERVICIOS INTEGRADOS (ISDN)	203
A. Servicios ISDN	210
B. Interfaces de usuario	215
C. Protocolos	221



LISTA DE TABLAS

Tabla	Páginas
10.1 Horas de uso de microcomputadoras, total y por aplicación. Mes de septiembre de 1989	157
10.2 Horas de uso de microcomputadoras, total y por aplicación. Mes de octubre de 1989	158
10.3 Resumen de horas de uso de las microcomputadoras. Meses de septiembre y octubre de 1989	159
10.4 Número de usuarios de microcomputadoras, total y por aplicación. Mes de septiembre de 1989	160
10.5 Número de usuarios de microcomputadoras, total y por aplicación. Mes de octubre de 1989	161
10.6 Resumen de número de usuarios de microcomputadoras. Meses de septiembre y octubre de 1989	162
10.7 Resumen de utilización de microcomputadoras, por aplicación. Meses de septiembre y octubre de 1989	163
10.8 Uso de computadora HP-3000. Mes de julio de 1989	166
10.9 Uso de computadora HP-3000. Mes de agosto de 1989	167
10.10 Uso de computadora HP-3000. Mes de septiembre de 1989	168
10.11 Resumen de los porcentajes de utilización de la HP-3000. Meses de julio, agosto y septiembre de 1989	169

Tabla	Páginas
10.12 Inscripción del Colegio Universitario y de las Facultades de Ciencias y Humanidades y de Ciencias Sociales, por carreras y grados, correspondientes al segundo ciclo de 1986.	170
10.13 Inscripción del Colegio Universitario y de las Facultades de Ciencias y Humanidades y de Ciencias Sociales, por carreras y grados, correspondientes al segundo ciclo de 1987.	171
10.14 Inscripción del Colegio Universitario y de las Facultades de Ciencias y Humanidades y de Ciencias Sociales, por carreras y grados, correspondientes al segundo ciclo de 1988.	172
10.15 Inscripción del Colegio Universitario y de las Facultades de Ciencias y Humanidades y de Ciencias Sociales, por carreras y grados, correspondientes al segundo ciclo de 1989.	173
10.16 Resumen de alumnos inscritos en el segundo ciclo. Años de 1986, 1987, 1988 y 1989	174
10.17 Resumen de porcentaje de alumnos inscritos en el segundo ciclo. Años de 1986, 1987, 1988 y 1989	175
10.18 Inscripción de la Facultad de Educación, por carreras, correspondiente al primer ciclo de 1986	176
10.19 Inscripción de la Facultad de Educación, por carreras, correspondiente al primer ciclo de 1987	177
10.20 Inscripción de la Facultad de Educación, por carreras, correspondiente al primer ciclo de 1988	178

Tabla	Páginas
10.21 Inscripción de la Facultad de Educación, por carreras, correspondiente al primer ciclo de 1989	179
10.22 Resumen de estudiantes inscritos en la Facultad de Educación. Años 1986, 1987, 1988 y 1989	180
10.23 Número de terminales requeridas	186



LISTA DE FIGURAS

Figura	Páginas
2.1 Imagen del sistema	6
2.2 Incremento de la distribución como función de la autonomía del nodo	9
2.3 Multiprocesador	17
4.1 Modos de transmisión	37
4.2 Concepto de pipa	42
4.3 Transmisión en serie y paralelo	53
4.4 El conjunto de caracteres ASCII	55
4.5 El alfabeto internacional No. 5	56
4.6 Interface digital RS-232/CCITT V.24	58
5.1 Red de comunicación entre computadoras	71
5.2 Configuraciones topológicas	78
5.3 Estructura de sistemas	79
8.1 Procesadores de funciones globales	128
8.2 Sistema con control distribuido	129
A.1 Modelo OSI	207
A.2 Grupos funcionales y puntos de referencia para ISDN	208
A.3 Dirección ISDN	214
A.4 Configuraciones de referencia	215
A.5 Estructura de trama	218

Figura	Páginas
A.6 Conector ISDN	220
A.7 Interacciones entre las entidades de la interface usuario-red	221
A.8 Diagrama funcional del adaptador A/S	225

I. INTRODUCCION

Actualmente, todos los procesos de manejo de información se están automatizando; en los ámbitos comerciales, industriales, financieros y educativos se tiende a realizar un manejo y transferencia electrónica de los datos.

Al iniciarse la automatización en el manejo de la información, únicamente se poseían sistemas centralizados; pero actualmente se tiene la posibilidad de distribuir los datos, las funciones o el control en diferentes nodos. Un sistema distribuido tiene como característica básica la división de datos o de funciones en unidades bien definidas que pueden estar localizadas en diferentes sitios y pueden enlazarse por medio de protocolos.

Al surgir en una organización la necesidad de automatizar el manejo de los datos, se puede escoger entre dos opciones: un sistema centralizado o un sistema distribuido. No existe un límite exacto que defina cuando se tiene un sistema centralizado y cuando se convierte éste en un sistema distribuido; el cambio entre uno y otro se realiza de una manera difusa entremezclándose ambos conceptos.

En la primera parte del presente trabajo se describe lo que es un sistema distribuido, indicándose sus ventajas y

desventajas comparado con un sistema centralizado. Adicionalmente se presentan los fundamentos para la transmisión de datos y las diferentes topologías de redes de computadoras. Posteriormente se describen los diversos criterios existentes para la distribución de los datos, las funciones y el control de la red.

En la siguiente parte se enumeran los pasos a seguir en el diseño de una red distribuida de datos, indicándose el tipo de información que se debe recabar y el análisis a realizar, además de los criterios de selección a utilizar.

Por último punto se presenta de manera detallada el diseño de una red distribuida de datos para la Universidad del Valle de Guatemala; se presenta la información recabada y el análisis realizado de la misma y se realiza el diseño en base a dicho análisis.

A pesar de las múltiples ventajas que presenta un sistema distribuido, debe realizarse un estudio de cada organización, ya que en algunos casos será recomendable utilizar un sistema distribuido y en otros casos, un sistema centralizado. La elección dependerá de las características que presente cada empresa o institución.

II. PROCESAMIENTO DISTRIBUIDO

Es difícil dar una definición precisa del procesamiento distribuido: el campo que abarca es grande y sus fronteras no están totalmente delimitadas.

En la literatura acerca del tema existen muchas definiciones tentativas para este término, algunas hasta con características contradictorias, otras lo entremezclan con el procesamiento paralelo o sistemas de multiprocesadores.

El concepto de procesamiento distribuido ha tenido una gran evolución: a finales de la década pasada, en algunos libros sobre el tema, se denominaba sistema distribuido a aquel sistema en que se poseía una minicomputadora y varios usuarios podían accederla en tiempo compartido.

Podemos decir que el procesamiento distribuido tiene como característica básica la división de datos o de funciones en unidades bien definidas que pueden estar localizadas en varios sitios lógicos y pueden enlazarse por medio de protocolos.

Un concepto muy importante es el de sitio lógico, porque no implica una separación física entre los componentes del sistema; se puede tener un sistema distribuido con elementos de hardware localizados en un mismo lugar. El sitio lógico consiste en un lugar conceptual en el sistema, conteniendo

una división bien definida de procedimientos y/o datos.

Debido a la amplitud de la definición anterior, es posible encontrar una gran cantidad de variantes de sistemas con características suficientes para denominarse distribuidos.

A. Características de los sistemas distribuidos

En general, los sistemas distribuidos tienen las siguientes características:

1. Múltiples unidades de procesamiento. Una característica fundamental de un sistema distribuido es la existencia de más de una unidad de procesamiento; aunque el hecho de existir varias unidades de procesamiento no implica necesariamente la existencia de un sistema distribuido.

Por ejemplo, una computadora que esté conectada a una impresora "inteligente", no la podemos considerar como sistema distribuido, aunque la impresora sea manejada por un microprocesador y el funcionamiento de la impresora esté definido por programas propios de la impresora y corriendo en su microprocesador.

Una forma de solventar esta dificultad consiste en pensar en "unidades de procesamiento significativas": una diferencia desproporcionada entre las unidades de procesamiento puede indicarnos que el sistema no es distribuido; la presencia de unidades de memoria secundaria y elementos del sistema operativo puede indicarnos la

presencia de un procesador "significativo".

2. Imagen de un solo sistema. Una característica del sistema distribuido es el grado en el cual se presenta a los usuarios como un único sistema.

Algunos opinan que el sistema debe presentarse de tal forma que los usuarios no se den cuenta de la existencia de múltiples unidades de procesamiento, que los usuarios tengan accesibilidad a diferentes datos y funciones sin importar el conocimiento del usuario acerca de la existencia de diferentes nodos.

Algunos diseñadores piensan que el sistema debe presentarse como una única entidad y, en caso extremo, un usuario podría acceder al sistema en cualquier nodo y llamar a datos y funciones sin importar donde se encuentren; un programa de aplicación utilizaría el mismo formato para pedir datos, sin importar donde estén los datos y donde esté el programa. De esta forma se puede lograr más fácilmente un balance de carga entre los nodos, ya que al usuario no le interesa el lugar físico donde están los programas y los datos.

Una completa transparencia se logra con mayor facilidad cuando los nodos son idénticos (tanto en software, hardware, sistema operativo, interfaces al usuario, manejador de bases de datos, etc.).

Si se trata de lograr un sistema distribuido con

completa transparencia a partir de nodos disímiles (hardware, software, sistema operativo, etc.), aparecen un sinnúmero de problemas y se debe poner en la balanza la dificultad técnica de la interconexión y la necesidad de dicha transparencia.

El problema con un sistema heterogéneo es la necesidad que las interfaces de usuario y aplicaciones sean las mismas para todos los nodos, lo que implica la creación de una nueva capa (ó layer) a colocar arriba de los sistemas de software existentes; las aplicaciones piden datos de la misma forma y estos datos pueden moverse libremente en el sistema.

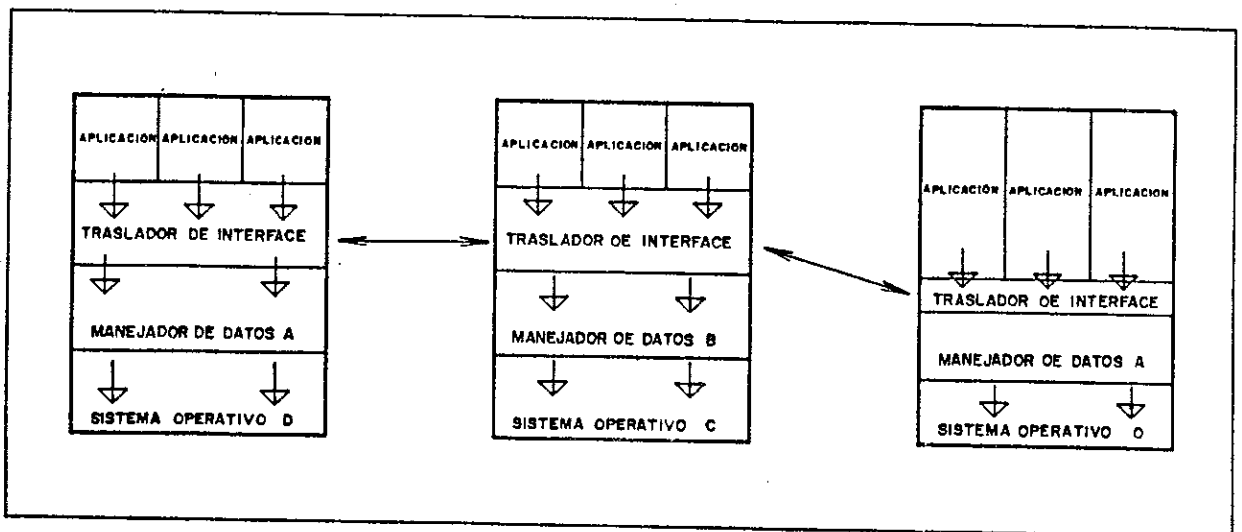


FIGURA 2.1
IMAGEN DEL SISTEMA

Otra posible solución consiste en que cada nodo permanezca con sus mismas interfaces de usuario e interfaces de los programas de aplicación, y el software del sistema

provea las traslaciones necesarias para cuando se necesita mover información de un nodo a otro. Una gran ventaja de este método resalta en el caso de interconexión de nodos disímiles, ya que los usuarios no necesitan cambiar sus patrones y convenciones.

El concepto de presentar una sola imagen del sistema depende del grado de interdependencia entre los nodos. La frecuencia de comunicación entre ellos indica en cierta medida la presentación de una sola imagen. Algunos diseñadores piensan que un sistema distribuido debe guiar a diseñar sistemas con un alto grado de dependencia e interacción.

3. Interconexión electrónica. Las unidades de procesamiento deben estar interconectadas electrónicamente; no es necesario que esta interconexión sea por teleproceso.

4. Libertad en el diseño. Se puede utilizar tanto el método arriba-abajo (top-down) o el método abajo-arriba (bottom-up).

Arriba-abajo: toma el diseño del sistema como una unidad; se puede asignar a cada nodo una serie de trabajos, y luego se diseña el flujo de datos y de control. El diseño de aplicaciones es similar al de los sistemas no distribuidos, pero los componentes del sistema están en diferentes nodos. Se logra una alta interdependencia y se

presenta al usuario como un solo sistema.

Abajo-arriba: es la técnica usualmente utilizada cuando se desea un sistema distribuido a partir de varios nodos existentes. Es el método recomendable cuando existen aplicaciones independientes manejadas por diferentes departamentos. Los motivos para diseñar un sistemas distribuido a partir de nodos independientes son: necesidad de mejorar el flujo de información, aumentar la cooperación interdepartamental, bajar los costos de compra de software, habilidad de mover aplicaciones entre departamentos, etc. El principal peligro en este tipo de diseño es la dificultad para realizar la interconexión.

En el diseño de sistemas distribuidos se tiene un mayor rango de alternativas en cuanto al lugar para colocar los datos, los programas y el control.

Una medida del grado de distribución de un sistema consiste en el grado en el cual los nodos individuales pueden decidir su propio destino. Una de las decisiones a tomarse en el diseño consiste en definir el grado en el cual existirá control local en el nodo, de manera que este nodo pueda controlar en cierta medida sus recursos. Para lograr lo anterior, es necesario tener de manera residente funciones del sistema operativo y que el nodo tenga otras capacidades además de poder recibir y enviar mensajes.

En una relación maestro/esclavo existe una

centralización del control, al ir dando más actividad a realizarse en el nodo, éste va asumiendo control sobre sí mismo; al poner más control en el nodo, la interacción entre nodos irá disminuyendo.

El sentido de distribución proviene del hecho de distribuir el control, los datos y las funciones del sistema.

Pero puede llegarse a un punto en el cual el nodo llega a ser tan independiente, perdiéndose el sentido de sistema distribuido debido a la poca interacción y dependencia.

De lo anterior se puede observar que los límites del sistema distribuido no están totalmente definidos.

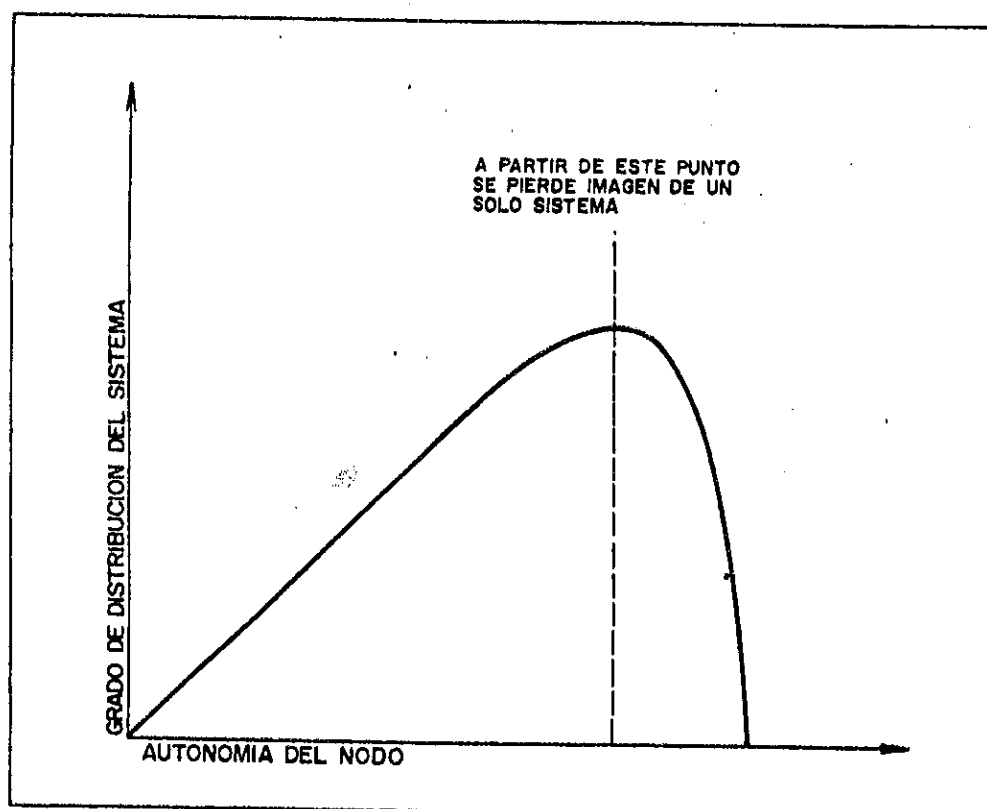


FIGURA 2.2
INCREMENTO DE LA DISTRIBUCION COMO FUNCION DE LA AUTONOMIA DEL NODO

B. Objetivos del procesamiento distribuido

Los objetivos perseguidos al utilizar el procesamiento distribuido se pueden dividir en organizacionales y técnicos.

1. Organizacionales. La distribución del hardware es sólo una de las decisiones a tomarse. Existen cuatro funciones del procesamiento de datos a distribuirse:

a. Análisis de sistemas:

Si el análisis es realizado por analistas localizados en el lugar donde se encuentran los usuarios, tendrán un mejor entendimiento de los problemas y una mejor comunicación con los usuarios. Este contacto más cercano puede resultar en mejores soluciones al procesamiento de datos. Lo anterior es especialmente cierto cuando los diferentes usuarios utilizan diferentes tipos de aplicaciones. La desventaja consiste en que sin una fuerza que obligue a la standarización, los diseños, métodos y procedimientos pueden diverger y crear problemas de compatibilidad y mantenimiento.

b- Programación:

Sin importar dónde se realice el análisis del sistema, se debe decidir dónde se efectuará la programación. Si se dan especificaciones completas y precisas del programa a efectuar, no existe gran diferencia en el lugar donde se efectúe el programa; pero lo anterior no se aplica cuando el análisis, diseño y programación son funciones combinadas en

las mismas personas. Se debe tomar en cuenta las economías de escala porque son una razón para centralizar el staff de programación, que incluye expertos de sistemas operativos, manejo de bases de datos, comunicación de datos, etc.

c- Operaciones:

Generalmente se piensa que una distribución del hardware implica necesariamente una distribución del staff de operaciones; pero esto no es siempre cierto. En cuanto al personal de ingreso de datos, es mejor realizar la captura de datos lo más cercano posible a su origen, sugiriendo que la entrada de datos debe realizarse en los sitios locales, cuando sea posible.

d- Planeación y coordinación:

Hasta cierto punto, estas funciones deben realizarse de una manera centralizada, porque de otra manera ocurriría poca standarización y compatibilidad. Esto podría llevar a sistemas en los cuales los nodos no podrían comunicarse entre sí. Algunas organizaciones poseen un centro de procesamiento de datos o un grupo de planeamiento y standarización encargado de seleccionar cierto tipo de hardware y software dependiendo del tipo de aplicación.

El objetivo organizacional más importante perseguido en un sistema distribuido es permitir al usuario tener un control más directo de lo que está haciendo. El director de un departamento o sección es responsable de lograr ciertos

objetivos y se le deben proporcionar los recursos necesarios incluyendo personal, dinero e información.

2. Técnicos. Los principales objetivos técnicos del procesamiento distribuido son:

a. Mejorar la seguridad:

La confianza en el sistema es crucial en el caso en el cual una de las aplicaciones es esencial para el funcionamiento de la organización (por ejemplo, la reservación de pasajes en aerolíneas).

En los sistemas centralizados se trata de dar una mayor seguridad a partir de una duplicación de ciertos componentes de hardware; pero se logra una mayor confianza al realizar dicha duplicación en un sistema distribuido que tiene sus nodos en diferentes lugares. Si un nodo cae, los únicos usuarios afectados serían los de ese nodo. La seguridad y disponibilidad del sistema se mejorará aún más si los usuarios pueden entrar al sistema por diferentes nodos. Si los usuarios del nodo que falló pueden entrar al sistema por otro nodo, la pérdida del nodo afectará únicamente el desempeño del sistema.

Si en el diseño, cada nodo está dedicado a un conjunto específico de funciones, al pedir un usuario el uso de una función se manda el pedido al nodo apropiado, sin importar de donde provenga dicho pedido. Al fallar un nodo las funciones de dicho nodo no estarán disponibles para nadie en

el sistema. Cuando todos los nodos pueden realizar todas las funciones, mientras algún nodo permanezca disponible, el sistema podrá seguir funcionando.

Otro aspecto a tomarse en cuenta es la disponibilidad de los datos: si la única copia de los datos residía en el nodo que falló, entonces el aumento de la seguridad en un sistema distribuido será una ilusión. Una posible solución consiste en la redundancia de datos, en poseer en el sistema varias copias de los datos.

En el diseño del sistema también debe tomarse en cuenta que la falla de un nodo no debe de afectar la comunicación entre los demás nodos: en un sistema totalmente mallado la falla de un nodo no afecta sustancialmente la comunicación de los demás nodos; pero en un sistema en estrella, la falla del nodo central puede tener efectos catastróficos.

b. Rápida respuesta y costos menores:

Si en el sistema se hace referencia frecuente a los datos locales, estos datos pueden almacenarse en el nodo donde son más utilizados. En este caso se procesarán los datos en forma local, mejorándose el tiempo de respuesta y disminuyendo los costos de comunicación.

c. Crecimiento del sistema:

El sistema distribuido provee una forma, en algunos casos la única, para poder incrementar el desempeño del sistema de procesamiento. Al aumentar los requerimientos de

procesamiento de la organización, el sistema debe aumentarse. En los sistemas centralizados esto puede implicar cambios en el software, incompatibilidad del hardware, etc. En un sistema distribuido bien diseñado, el crecimiento puede ser modular porque se pueden agregar nodos para aumentar la capacidad del sistema.

C. Alternativas de procesamiento Distribuido

Una organización puede seleccionar entre una gran variedad de sistemas distribuidos. A continuación se presentan ciertas alternativas desde el sistema centralizado hasta el distribuido, pasando por varias etapas intermedias. En ellas, puede observarse que no existe un límite exacto que defina cuando se tiene un sistema centralizado, y cuando llega a ser un sistema distribuido. El cambio de uno a otro se realiza de una manera difusa en la que ambos conceptos se entremezclan. En las siguientes alternativas no están especificadas todas las posibilidades, sino que únicamente se pretende mostrar este paso de centralizado a distribuido.

- En un sistema totalmente centralizado el poder computacional, el almacenamiento de datos y la entrada de los mismos están centralizados.

- Procesamiento y almacenamiento centralizado con entrada remota: en un inicio todo el hardware (procesamiento, memoria, I/O) estaba centralizado. El primer elemento de hardware a descentralizarse fueron los

dispositivos de entrada y salida (I/O).

- Procesamiento y almacenamiento centralizado con terminales en línea para procesamiento: todos los recursos de procesamiento y almacenamiento están centralizados, en cada lugar existen una o mas terminales conectadas a la computadora central. El número de terminales depende de la actividad realizada en cada sitio. En teoría, un usuario de una localidad remota puede utilizar cualquier función de la computadora central.

- Procesamiento y almacenamiento centralizado con entrada de datos distribuida: en el caso anterior se utilizan terminales "tontas", en este caso se le agrega a las terminales un mínimo de capacidad de procesamiento y almacenamiento. La mayoría de la entrada de datos y edición se realiza en los lugares remotos por personal local; se efectúa la transmisión de datos hacia la computadora central hasta realizar toda la edición de la información. En algunos casos la información se envía inmediatamente después de la edición, en otros se guarda temporalmente y se envía la información en tandas (batch).

- Entrada de datos y edición distribuida con partición local de la base de datos central: Se provee a la localidad remota de capacidad de almacenaje adicional; parte de la base de datos se guarda temporalmente en los sitios remotos para poder realizar una edición más completa: toda la

edición se realiza en el sitio local. Cuando el nodo remoto está totalmente satisfecho con los datos, estos se envían a la unidad central donde se realizará todo el procesamiento. Un punto importante en este caso es que parte de la base de datos se almacena localmente para fines de edición; todo el procesamiento y modificaciones de la base de datos se realiza en el computador central. Esta distribución puede mejorarse aumentando la capacidad de almacenaje y procesamiento de los nodos, y realizando el procesamiento de transacciones en ellos. Otra posibilidad consiste en permitir al nodo local procesar la copia que posee de la base de datos, aunque esto implicaría problemas de sincronización entre las copias.

- Procesamiento y almacenamiento distribuido en los sitios locales y enlace al sitio central: En este caso el nodo tiene capacidad de procesamiento local y significativa capacidad de almacenamiento secundario, y los datos y programas pueden almacenarse localmente de manera permanente. Esta opción es apropiada cuando el nodo remoto necesita realizar ciertas operaciones en sus datos. Se tiene poca comunicación entre el nodo remoto y el central y no está permitida la comunicación entre nodos remotos; se tiene una estructura en estrella.

- Procesamiento y almacenamiento distribuido con comunicación internodal: Es un caso igual al anterior, pero

permite a los nodos locales comunicarse entre si.

- Sistemas de manejo de bases de datos distribuidas: El usuario puede ingresar datos en cualquier parte del sistema sin tener que especificar su localización. En este caso puede existir o no la redundancia de datos. Es el caso más general. Para demostrar la dificultad en definir un sistema distribuido, tomemos el siguiente ejemplo:

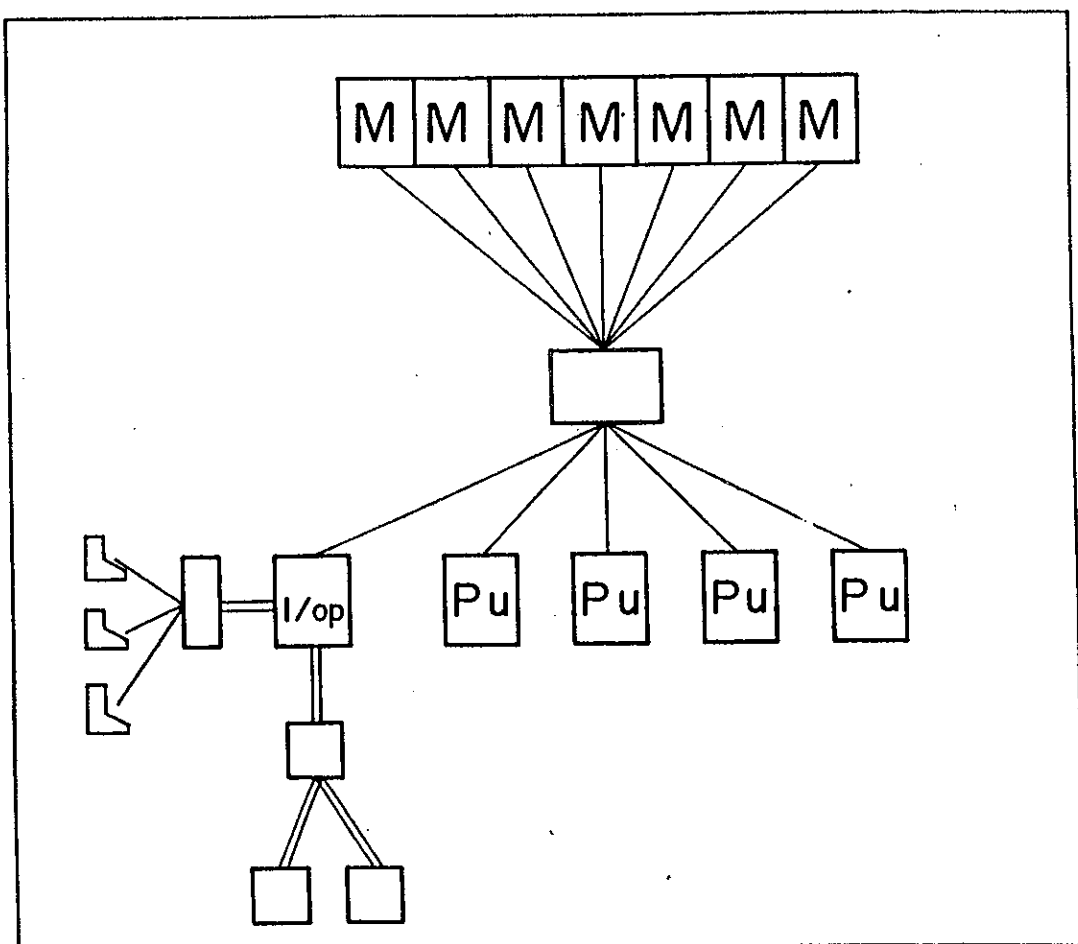


FIGURA 2.3
MULTIPROCESADOR

El sistema consiste en un multiprocesador con una unidad

de entrada/salida especializada que da servicio a todos los procesadores y una memoria común compartida por todos los procesadores.

Se puede afirmar que el sistema anterior no es distribuido por las siguientes razones:

- Se limita geográficamente porque las unidades de procesamiento deben estar físicamente cerca.

- Cada procesador no es una unidad completa; no es un nodo independiente y debido al nivel de compartición de recursos, no es un sistema confiable.

Pero también se puede decir que es un sistema distribuido por las siguientes razones:

- El grado de proximidad física no es relevante en la definición de un sistema distribuido, lo importante es la definición de sitios lógicos.

- En esta configuración puede crearse cierto grado de "tolerancia a fallas" y se puede mostrar a cada unidad de procesamiento como si tuviera su memoria propia y su dispositivo de entrada/salida propio.

III. MOTIVOS PARA LA DISTRIBUCION Y

COMPARACION ENTRE SISTEMA CENTRALIZADO Y DESCENTRALIZADO

En este capítulo se enumeran las principales razones esgrimidas en favor del sistema distribuido, se analiza cada una de ellas y se estudia su validez.

A. Razones para la distribución

Algunas de las razones que motivan la distribución son:

1. Manejo centralizado de la red. El manejo del sistema de manera centralizada es esencial cuando el sistema distribuido soporta funciones interdepartamentales y cuando se planea realizar un solo sistema a partir de múltiples nodos existentes (diseño bottom-up). El deseo de mantener un manejo centralizado, aun en los nodos remotos, se debe a las siguientes necesidades: mejor control, utilización del personal y evitar costos de incoherencia.

Un mejor control conlleva una generación de estándares para programación, equipos de operación e interconexión. Con un manejo centralizado de la red se pueden dar directrices para asegurar una ordenada adquisición y utilización de los sistemas computacionales. En cuanto a personal, si se tiene un único equipo para el procesamiento de datos, se puede atraer a buenos profesionales en diferentes áreas (transmisión de datos, manejo de sistemas, programadores,

etc.).

2. Costos de comunicación. Un motivo muy frecuente para la distribución es el deseo de reducir el costo de un conjunto de terminales no inteligentes (tontas), comunicándose desde una distancia grande al centro de procesamiento de datos.

A pesar de la posibilidad de disminuir los costos de comunicación al darles más inteligencia (memoria, capacidad de procesamiento) a las terminales remotas, al realizar la distribución de un sistema no necesariamente, disminuyen los costos de comunicación.

Si en un sistema las terminales no inteligentes se sustituyen por terminales inteligentes, el costo de comunicaciones puede bajar: el tráfico hacia el procesador central disminuirá, porque en vez de estar enviando transacciones sin procesar, se mandará un sumario de las transacciones realizadas. Pero si se tienen nodos independientes y estos se reúnen para formar un sistema distribuido (diseño bottom-up), los costos de comunicaciones aumentarán.

3. Confiabilidad. Un sistema distribuido puede proveer una mayor confiabilidad y facilidad de acceso porque en este tipo de sistema es más económica la replicación, y al ocurrir alguna falla en algún nodo, se puede diseñar el sistema para que los demás nodos puedan mantener funcionando

al sistema.

La práctica anterior es sumamente atractiva debido a que las unidades de procesador y memoria no son muy caras y las unidades adicionales dan un alto grado de seguridad. En un sistema centralizado, duplicar una unidad grande puede tener costos prohibitivos.

A pesar de lo anterior deben tomarse en cuenta los siguientes puntos:

Las unidades de memoria y procesamiento son las más confiables, y en general no es tan claro cómo el hecho de replicar estas unidades hará que el sistema sea más confiable. La réplica de unidades de procesamiento implica nuevos diseños para la transferencia, acceso y sincronización de los datos.

Para recuperación de fallas, es probablemente más barato realizar una réplica selectiva, seleccionando los puntos críticos. Si es necesario replicar el sistema entero para lograr la confiabilidad deseada, es posible que la replicación de un solo nodo (sistema centralizado) sea igualmente efectivo.

El funcionamiento del sistema es la probabilidad de funcionamiento de todos los nodos. La probabilidad de falla de todos los nodos es sumamente baja, pero si para lograr que todo el sistema funcione bien es necesario que todos los nodos estén funcionando, éste no es un sistema confiable. Se

debe diseñar el sistema de tal forma que la falla en uno de los nodos afecte en el menor grado posible al sistema (probablemente afectando únicamente a los usuarios de ese nodo).

4. Mejorar interfaces para los usuarios. La posibilidad de ofrecer mejores interfaces para los usuarios es citada como una ventaja de los sistemas distribuidos. Algunas de las características necesarias para una buena interface son:

- Proveer una respuesta del sistema en un tiempo apropiado para la aplicación.

- Tener una semántica consistente con el trabajo a realizar.

- Tener una sintaxis natural que parezca intuitiva al usuario.

- Poseer una sintaxis uniforme y consistente con el contexto del trabajo.

- Permitir diferentes niveles de ayuda dependiendo del conocimiento del operador.

Colocar una unidad de procesamiento en cierto lugar no provee necesariamente una buena interface para el usuario, pero la existencia de este procesador puede hacer más fácil la implementación de una buena interface.

En cuanto al tiempo de respuesta en programas de aplicación, éste puede mejorarse utilizando procesadores

dedicados en comparación con sistemas de tiempo compartido.

Si la carga de un nodo es predecible y homogénea, el sistema es más simple de analizar y se puede mejorar el tiempo de respuesta. Si las acciones del usuario se ejecutan en el nodo en el cual está trabajando, la respuesta será más rápida en comparación con el caso en que se necesitara realizar una comunicación con un nodo remoto.

En cuanto a la calidad del diálogo, un sistema distribuido lo hace más accesible, un mejor diálogo puede lograrse con una mejor interacción con el nodo procesador y una mayor presentación de datos y formatos.

Para los procesadores encargados de realizar las operaciones computacionales, el estar presentando formatos, mensajes y niveles de ayuda al usuario puede presentar una carga sumamente grande. Por lo que los diseñadores de interfaces tienden a usar formatos de mensajes concisos y breves, esta tendencia aumenta por el deseo de minimizar el flujo de datos de y hacia el sistema. Con este punto de vista, es posible mejorar el diálogo por medio de un procesador remoto sin incrementar la carga en los procesadores computacionales. Pueden fluir mensajes cortos y concisos entre el nodo y el procesador computacional, y es el nodo local el que soporta el diálogo con el usuario, expandiendo los mensajes y los niveles de ayuda.

Para poder dividir el diálogo con el usuario es

necesario que el nodo remoto tenga la capacidad para efectuar estas funciones y su carga no debe ser tan alta que aumente su tiempo de respuesta.

5. Seguridad y privacidad. El uso de diferentes nodos puede justificarse por el deseo de aumentar la privacidad de los datos y para incrementar la seguridad.

Muchos usuarios de sistemas grandes se muestran preocupados por el hecho de que no mantienen un control total sobre sus datos. Debido a los diferentes procedimientos utilizados para archivar los datos y a las prerrogativas del personal del centro de cómputo, es prácticamente imposible para el usuario controlar el acceso a sus datos.

Tomando en cuenta este deseo de privacidad, es sumamente atractiva la posibilidad de tener los datos almacenados con el acceso controlado por la unidad o el departamento que utiliza y crea dichos datos.

Los sistemas distribuidos son diferentes a los sistemas centralizados, ya que por naturaleza implican cierto acceso de un nodo a las bases de datos de otro. Se debe proveer un método de protección contra violaciones procedentes de otros nodos.

6. Economías de dedicación. La idea básica de este punto es que es mejor distribuir la actividad a través de una familia de procesadores, porque la economía de

escala no es un factor muy importante en el mercado de hardware; las unidades pequeñas de procesamiento pueden tener una relación precio/rendimiento igual o aun mejor que los procesadores grandes.

El costo del hardware está determinado cada vez menos por el precio de los procesadores y cada vez más por el precio de las unidades de almacenamiento, equipo periférico y equipo de comunicaciones; por tanto, a pesar de la pérdida de economías de escala en los procesadores, una colección de nodos pequeños con el poder equivalente a un sistema grande, costará más que dicho sistema grande.

Existen diferentes grados en la noción de dedicación: desde sistemas donde los nodos son altamente especializados hasta sistemas de un solo nodo donde todos los departamentos deben referirse a él.

Además de la posibilidad económica y técnica de tener una colección de pequeños sistemas, deben considerarse los costos de operación.

Dependiendo de las características del sistema, los costos de operación pueden variar: un número de unidades geográficamente distribuidas pueden necesitar programadores y operadores en cada sitio por lo tanto el costo total de operación puede subir; el costo total de operación del sistema es la suma del costo de operación en cada nodo y puede llegar a exceder al costo de operación de un sistema



grande de un solo nodo.

Una alternativa para reducir los costos de operación consiste en mejorar las interfaces de presentación del sistema. En vez de servicios remotos de operación existirán servicios programados de operación en cada nodo, automatizando las interfaces de operación, de tal forma que localmente se necesitará un número reducido de operarios sin habilidades y conocimientos especializados.

7. Crecimiento incremental y flexibilidad. Una de las ventajas principales del sistema distribuido sobre el centralizado es la posibilidad de crecer modularmente; se crece agregando módulos y no reemplazándolos. De esta forma los módulos son pequeños y un crecimiento paulatino no representa un fuerte desembolso para la organización.

Una configuración con un solo procesador presenta problemas para el crecimiento futuro: si se compra un procesador grande con capacidad de soportar crecimientos futuros, por un tiempo será demasiado grande y estará subutilizado. Una solución consiste en adquirir un uniprocador más pequeño con la capacidad de correr el sistema operativo deseado para las aplicaciones, y cuando quiera crecer, se puede interconectar otro procesador, pudiéndose dividir las aplicaciones y los datos; en este caso se tendría un sistema distribuido pero con nodos localizados en el mismo punto.

La mayor limitación en hardware consiste en los límites de los mecanismos de interconexión. Al incrementar el número de nodos aumenta la capacidad del sistema; pero no es segura la utilización de este camino en todas las topologías. Es necesario realizar un análisis de la carga total del sistema y de la carga internodal para poder repartir exitosamente la carga.

Otro aspecto a tomar en cuenta en este crecimiento modular es el aumento en costos al tener un nodo adicional. Idealmente el costo debería aumentar de manera lineal, pero no todas las topologías tienen este tipo de crecimiento.

8. Limitación de capacidad. Está íntimamente relacionada con el ítem anterior: es la idea de vencer las limitaciones en capacidad al colocar nuevas funciones en nodos adicionales. La idea clásica consiste en extender la vida del sistema centralizado grande al poner funciones en nodos periféricos. Se debe tener cuidado en el hecho de tener que soportar comunicación con nodos periféricos aumenta la carga del procesador central.

9. Aumentar la facilidad de instalación. Se escucha frecuentemente la idea que el sistema distribuido permite a las aplicaciones desarrollarse y hacerse operativas de una manera más rápida. Se piensa que es más fácil instalar una aplicación porque no es necesaria la negociación y la interacción con la organización de un

centro de cómputo grande para lograr recursos computacionales y personal de programación, debido a que el sistema es fácil de usar y las aplicaciones son usualmente pequeñas. Lo anterior no está totalmente comprobado y no se tiene certeza que la simplicidad de las máquinas pequeñas esté asociada a los sistemas distribuidos.

10. Corresponder a la organización de la empresa.

Debido a la estructura granular del sistema distribuido, existe la posibilidad de elaborar un mapa del sistema sobre la estructura organizacional.

Pero se deben tener las siguientes precauciones:

La organización de una compañía no es totalmente estable, puede sufrir cambios; algunas compañías tienen una estructura estable, pero otras están en constante cambio.

La estructura del sistema computacional debe ajustarse a la organización y no al revés.

Es menos pesado modificar una base de datos centralizada para representar nuevas unidades de la organización, que mover datos y hardware de un sitio a otro.

B. Sistema centralizado y descentralizado

En el capítulo II se dieron las características del sistema distribuido, se debe tener cuidado en diferenciar un sistema distribuido de uno descentralizado. La descentralización es un concepto organizacional y de administración. Los componentes de hardware, software,

estructuras de datos y aun las operaciones de control pueden estar dispersos sin crear necesariamente un ambiente descentralizado. Es únicamente cuando las habilidades del staff de planeamiento y desarrollo, el diseño del sistema y las decisiones comienzan a dispersarse, cuando ocurre la descentralización.

La distribución de habilidades a través de la empresa significa que los diferentes departamentos o locaciones tienen personal para desarrollar las funciones fundamentales del procesamiento de datos.

Algunos de los argumentos a favor de la centralización son:

- Control de la organización:

Los costos del procesamiento de datos son significativos y se hace necesario administrarlos explícitamente.

- Calidad del staff:

El diseño y desarrollo de sistemas de procesamiento de datos es una tarea compleja. Para lograr este objetivo se debe tener personal preparado y capaz; y para mantener este tipo de personal en la empresa se debe presentar una organización donde pueda hacerse carrera. Un sistema de administración y programación centralizado es el único camino para proveer desarrollo al profesional.

- Tamaño del staff:

Sin importar la existencia o no de economías de escala

en el hardware, existe economías de escala en el personal. Una organización centralizada permite la compartición de habilidades y conocimientos entre las aplicaciones y sistemas disminuyendo el costo de la actividad profesional por aplicación.

- "Puntos de vista de la empresa":

Es necesario que los intereses de la empresa sean representados en la planeación y desarrollo del centro de datos; las unidades individuales pueden diseñar y trabajar para su propio interés y sus propias metas, y éstas no necesariamente coincidirán con los intereses de la organización. Además, puede que no se den comunicaciones interdepartamentales eficientes ni compartición de datos, por lo que se necesita un sistema centralizado para el diseño y desarrollo.

- Stándares:

Para lograr los proyectos satisfactorios es necesario que se mantengan ciertos stándares en el hardware, software, lenguajes de programación, etc.

Entre los argumentos para la descentralización tenemos:

- Eliminación de interdependencia:

La interdependencia entre diferentes unidades y departamentos lleva inevitablemente a conflictos y debido a que se desea eliminar estos conflictos lo más posible, las organizaciones deben estructurarse para minimizar la

interdependencia, aun a costa de redundancia y duplicación de datos y funciones. Los costos de esta redundancia pueden ser menores que los costos de resolver conflictos entre departamentos.

- Evitar la percepción de pérdida de autonomía:

Las unidades departamentales se ven afectadas cuando sienten que sus áreas de decisión son reducidas. Los beneficios de una actitud agresiva e independiente superaría los costos. La sensación de pérdida de autonomía y propia determinación deben de mantenerse al mínimo.

- Responsabilidad:

Es necesario que los administradores de las unidades de la empresa sean responsables de su operación, por lo tanto todas las funciones y servicios necesarios para realizar su trabajo deben colocarse bajo su control. Si se necesita poder computacional para realizar una operación, entonces debe tener absoluto control sobre este poder computacional.

Se prefiere la descentralización cuando el trabajo del computador es esencial para la realización de las funciones de la unidad y cuando el manejo del poder computacional no requiere gran habilidad y conocimiento.

Es necesario recordar que algunas unidades o departamentos no pueden aceptar el control y la responsabilidad para manejar un computador, ya que adquirir las habilidades necesarias puede resultar sumamente costoso.

IV. TRANSMISION DE DATOS

La comunicación de datos es necesaria para la transferencia de información entre computadoras y terminales localizadas en diferentes lugares. Al diseñar el sistema de comunicaciones, la meta a alcanzar es el aumento de la velocidad del flujo de información, aumento del control y soporte de las operaciones debidas a la descentralización.

En un sistema de comunicaciones se necesita definir:

- El medio de comunicación.
- El equipo de comunicación.
- Los protocolos de comunicación.

En la mayoría de los países, las comunicaciones están reguladas; en los Estados Unidos de América el cuerpo regulador principal es la Comisión Federal de Comunicaciones (FCC, Federal Communication Commission). La regulación es necesaria para lograr una serie de productos compatibles entre sí, lo que incide en una baja de los costos al usuario.

En un sistema distribuido existen dos componentes de hardware básicos: los nodos y los caminos de comunicación. El nodo es el elemento procesador que contiene, tanto la capacidad de procesamiento, como la memoria, y dependiendo del tipo de nodo puede soportar memoria secundaria. Un nodo

puede ser tanto una mainframe, una minicomputadora o una microcomputadora. Para poder comunicarse con otros elementos de la red es necesario que el nodo posea software de comunicaciones. En los nodos grandes este software de comunicaciones puede coexistir con los demás programas de aplicación. En algunas redes pueden existir nodos dedicados específicamente a controlar las comunicaciones internodales y corren únicamente software de comunicaciones. A estos nodos se les denomina front end processors, cluster controllers, procesadores de comunicaciones o procesadores de control de red. Estos nodos necesitan algún tipo especial de hardware o únicamente utilizan software de comunicaciones.

El otro componente del sistema distribuido lo constituyen los caminos o enlaces necesarios para la comunicación entre las diferentes partes de la red. Entre las características que determinan este camino de comunicación están: el ancho de banda y el modo.

El ancho de banda es una característica física del medio; éste determina el número de bits que pueden transmitirse por segundo (bps). Una clasificación del ancho de banda del canal es la siguiente:

- Sub-vocal (menos de 200 bps).
- Vocal (600 a 4800 bps).
- Banda ancha (hasta 50 Kbps y aun más).

La clasificación anterior es sólo una posibilidad, y en diferentes libros se utilizan diferentes límites para dicha clasificación.

Si la velocidad de transmisión de la terminal es superior a la velocidad posible en el medio, pueden utilizarse buffers para "igualar" las velocidades. Cuando el medio tiene una capacidad mayor que la velocidad del dispositivo, para no desperdiciar esta capacidad, puede multiplexarse (FDM ó TDM) la información proveniente de varios dispositivos.

La segunda característica mencionada es el modo: un canal puede clasificarse en simplex, semi-dúplex, dúplex completo y écoplex.

- Transmisión en modo simplex: es cuando la información puede transmitirse únicamente en una dirección; sólo uno puede transmitir y el otro sólo recibe. La principal desventaja de este método es que no permite alguna interacción. Se utiliza principalmente en sistemas de emisión, cuando únicamente se desea enviar información y no se necesita respuesta, como por ejemplo en el caso en que la señal de un transductor se conecta a una terminal para su procesamiento, o en el caso de transmisiones de radio y televisión.

- Transmisión en modo semi-dúplex (half duplex): en este caso la transmisión ocurre en una sola dirección a la

vez, pero puede alternarse la dirección de la transmisión. En estos casos puede aprovecharse el ancho de banda completo del canal al transmitir en una sola dirección. Una desventaja es que puede crear retardos significativos debido a la conmutación en el canal, principalmente en el caso de transmisión de mensajes cortos.

- Transmisión en modo dúplex completo (full duplex): una línea dúplex completo, es aquella que puede transmitir información en ambos sentidos simultáneamente. Se puede transmitir y recibir al mismo tiempo. Este sistema es equivalente a dos canales en modo semi-dúplex, uno para transmisión y otro para recepción, por lo que no es necesario conmutar el sentido del trabajo del canal.

- Transmisión en modo ecoplex: consiste en una variante del modo dúplex completo en la que la información transmitida por la unidad remota es inmediatamente retransmitida en el otro sentido y es posible realizar un chequeo visual de la misma (se denomina ECO a esta retransmisión). Se logra dar ventajas al usuario, pero se da una mayor utilización del canal, disminuyendo el ancho de banda efectivo.

Existen otras características que sin ser intrínsecas al canal de comunicación determinan el desempeño del sistema. Estas características incluyen la velocidad de transferencia, el tiempo de establecimiento del enlace, el

retardo en la red y la confiabilidad.

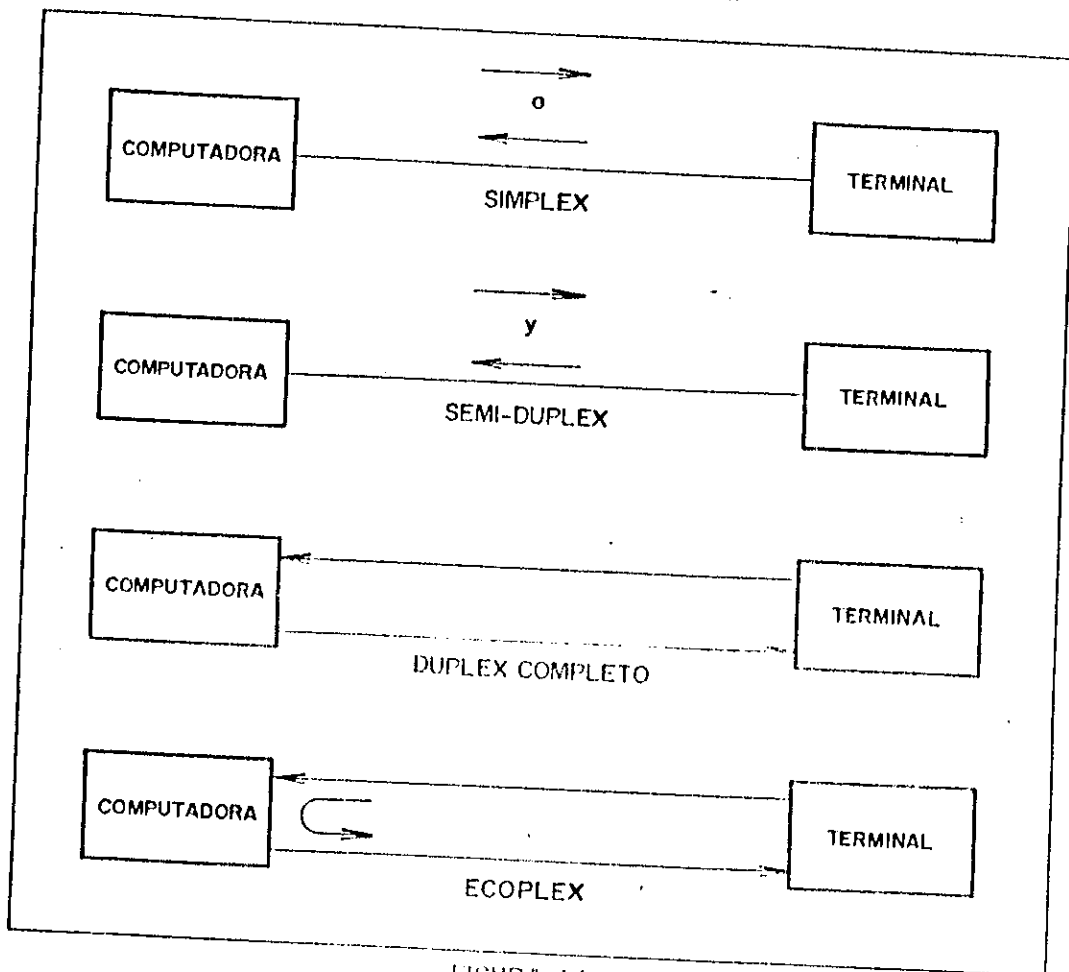


FIGURA 4.1
MODOS DE TRANSMISION

Para el usuario final, la velocidad de transmisión es un parámetro más importante que el ancho de banda. Se puede definir "tentativamente" el ancho de banda como el número de bits por segundo que se pueden transmitir a través de un canal. Pero la totalidad de bits que se transmiten no son de información. En muchos casos es necesario transmitir encabezados, señalización, señales de control, etc. (por ejemplo, los intervalos de tiempo t_1 y t_2 de la trama PCM de

2 Mb/s). El encabezado especifica la fuente y el destino del mensaje e identifica dicho mensaje. Además, el mensaje puede llevar información redundante (paridad, código cíclico, etc), para detección y/o corrección de errores. El ancho de banda del canal se utiliza para transmitir información importante para el usuario, para transmitir información redundante y para transmitir la información adicional. Por lo anterior puede afirmarse que la velocidad de transmisión (en bps) por el canal es mayor que la velocidad de transmisión de la información útil para el usuario.

El tiempo necesario para el establecimiento del camino es el tiempo que le toma al sistema de comunicaciones seleccionar y activar un camino desde el momento de la solicitud. Debe tomarse en cuenta la diferencia existente entre un camino físico y un camino lógico. Un camino lógico se establece desde el momento en que se inicia la sesión. La implementación física de dicho camino puede realizarse al mismo tiempo, y permanecer durante toda la sesión, pero en muchos casos este camino físico se crea dinámicamente para cada paquete de información. En una línea telefónica el camino físico se establece al crearse la sesión, y el tiempo de establecimiento es el tiempo necesario para marcar el número y para que los dispositivos de ambos lados del canal puedan realizar los procedimientos de handshaking necesarios. Este procedimiento puede llegar a tardar varios

segundos, por lo que puede tener un retardo prohibitivo para mensajes cortos. Con la conmutación de paquetes, existen enlaces fijos entre los nodos, y para establecer una comunicación se debe seleccionar únicamente el camino, porque los nodos terminales se conocen. Este proceso es más rápido que el anterior.

El retardo de la red es una medida del tiempo que tarda un mensaje a través de la red desde el lugar de origen hasta el destino. Este retardo puede dividirse en dos partes:

- El retardo debido a la propagación del mensaje de un nodo a otro, es el retardo de la señal a través del medio de transmisión y debido a que la señal se transmite a la velocidad de la luz, este retardo es generalmente despreciable y únicamente se toma en cuenta en la transmisión vía satélite.

- El otro tipo de retardo es el debido a la ejecución de cierto software de comunicación en los nodos de origen y destino y en los nodos intermedios. En la conmutación de paquetes es necesario empaquetar y desempaquetar la información y enrutarla a través de la red.

Otra característica del canal de comunicación es la confiabilidad; es decir, si puede asegurar la integridad del mensaje al enviarse de un sitio a otro. Se dice que un canal de comunicación posee buena disponibilidad, si acepta un mensaje cuando el usuario desea enviarlo. La confiabilidad

existe desde el momento en que el sistema aceptó el mensaje; si aceptó el mensaje, lo entregará en su destino aun en los casos de falla en alguna parte de la red. Se pueden utilizar técnicas de software para asegurar la integridad del mensaje, tanto códigos de detección como de corrección de errores.

Uno de los aspectos más importantes a tomar en cuenta en un sistema distribuido es el intercambio a ocurrir entre costos de procesamiento y costos de comunicaciones: el procesamiento se realizará principalmente en los nodos remotos, pero existirán costos de comunicaciones. Las diferentes tecnologías de comunicaciones pueden modificar drásticamente este intercambio. El diseño, costos y rendimiento varían dependiendo del sistema de comunicaciones a usar (Ej: par metálico y fibra óptica).

A. Niveles de comunicación

En cuanto a la tecnología de comunicación, podemos definir tres niveles básicos.

1. Nivel intracomputadoras. En este nivel existe movimiento de información entre partes de una computadora, la unidad de datos que consiste en uno o varios bytes más bits de control, se envía en paralelo. En este tipo de comunicación generalmente se utiliza cableado sumamente costoso y los dispositivos a comunicar no están separados por más de unas decenas de metros. Otra posible solución consiste en utilizar un bus serial de alta

velocidad que es compartido por todos los dispositivos.

Una opción para la interconexión de nodos localizados físicamente cerca consiste en la utilización de un "anillo" de alta velocidad. Este anillo, que interconecta a los diferentes nodos, es un dispositivo serial de alta velocidad con dos propiedades importantes: La primera es que los datos en el anillo pueden evitar un nodo inoperativo; dicho anillo se conoce como anillo T, sugiriendo que los datos fluyen por el nodo sin ninguna actividad necesaria del nodo para enviar la información "hacia adelante". La segunda propiedad es que el anillo es un dispositivo serial.

Los datos fluyen de nodo a nodo, si el procesador A envía un mensaje en el tiempo 0, este mensaje llegará al procesador B en el tiempo 1. En el tiempo 1 estos datos no estarán disponibles para el procesador C. La estructura del sistema es el de una pipa con un flujo de información único en una dirección. En este aspecto se diferencia el anillo del bus, ya que cuando se ponen datos en el bus estos están disponibles instantáneamente a todos los sistemas que tienen acceso al bus.

El procesador de control determina el enrutamiento adecuado al recibir el mensaje. Un nodo procesador recibe el mensaje del nodo "a su izquierda" y puede inspeccionar el código de enrutamiento establecido por el procesador de control y determina si el mensaje necesita procesamiento;

puede también determinar si el mensaje ha recibido procesamiento del procesador "a su izquierda". Cuando el mensaje pasa al nodo "a la derecha" puede estar de manera completa o incompleta. Si está de manera completa, entonces todos los procesadores que debían aceptar el mensaje lo guardarán; y el nodo de control recibirá esta aceptación. Es hasta este momento que se puede destruir la historia de la transacción.

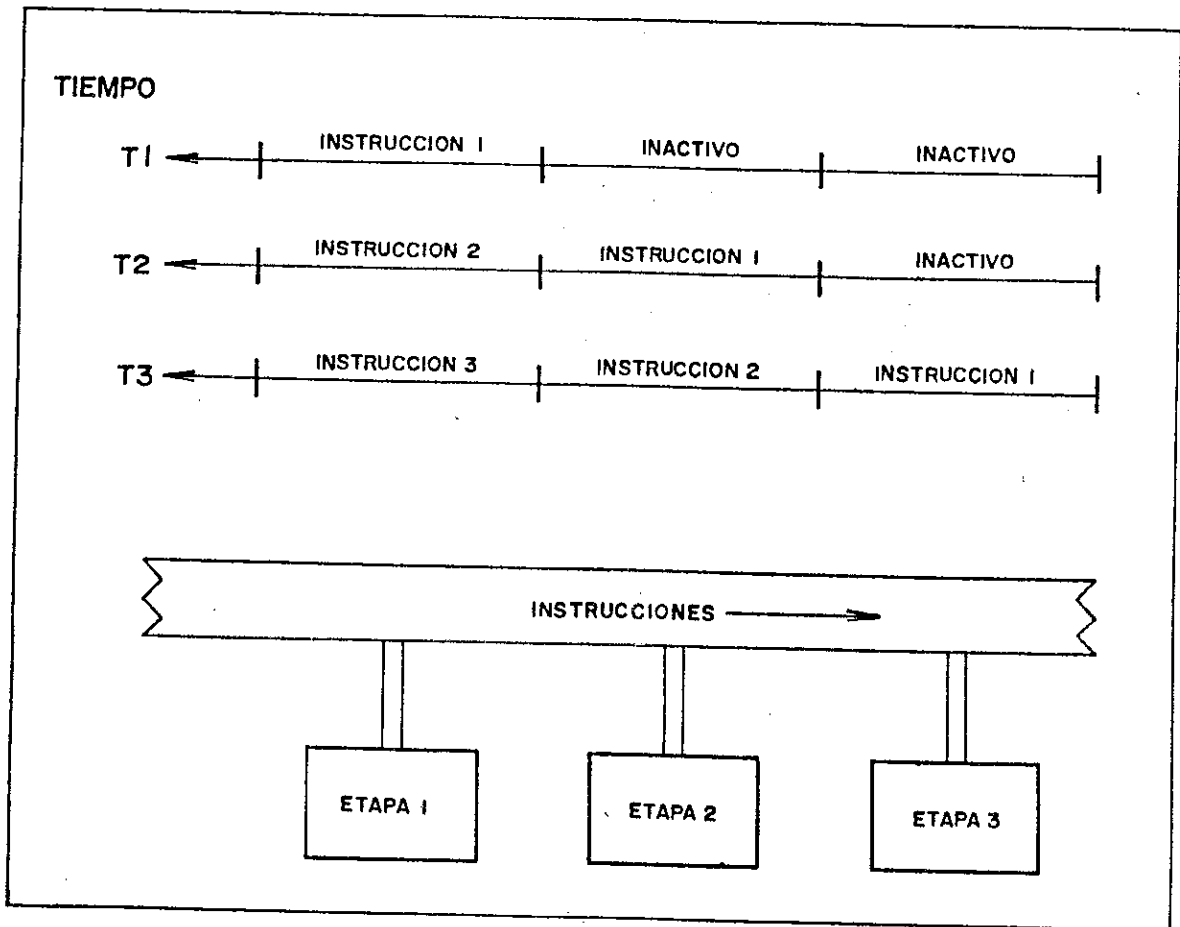


FIGURA 4.2
CONCEPTO DE PIPA

Cuando un nodo recibe una transacción, hace un intento para modificar sus archivos y/o bases de datos; pero por cierto tiempo retiene suficiente información para poder restablecer las condiciones iniciales por si más adelante se descubre que la transacción es inválida.

Entre las desventajas de la implementación anterior se tienen: para cada transacción se realiza un procesamiento multinodal, el desempeño del sistema es sumamente vulnerable, los problemas de coordinación pueden llegar a ser severos, la característica de pipa puede causar un retardo considerable y puede imponer una necesidad de almacenamiento adicional.

2. Redes de área local. Una red es un conjunto de entidades computacionales independientes (minicomputadoras, estaciones de trabajo, etc.) equipadas para comunicarse entre sí.

Una red de área local (LAN) limita la extensión física de la red, además de asegurar una alta velocidad de transmisión y una baja tasa de errores.

Una forma de clasificar las redes de área local es por el medio de transmisión utilizado, pudiendo utilizarse cables metálicos o fibra óptica; algunas redes pueden llegar a utilizar comunicación por radio, luz infrarroja o cables de alta tensión.

Cable paralelo: consiste en un par de cables, uno al

lado del otro y luego enrollados entre sí, puede ser blindado o no. Proporciona una velocidad de transmisión de unos pocos Mb/s (1-2) a una distancia de varios cientos de metros. Por su bajo costo es utilizado en redes con bajo volumen de tráfico.

Cable coaxial: consiste en dos conductores, uno totalmente rodeado del otro y separados por un aislante. En general es más costoso que el cable paralelo. Permite velocidades de transmisión de hasta varias decenas de Mb/s.

Fibra óptica: tiene una capacidad mayor que el cable coaxial, se volverá más utilizado mientras su costo continúe disminuyendo. Permite instalaciones en ambientes con interferencias electromagnéticas y cuando se necesita transmitir a altas velocidades en grandes distancias.

Una red de área local puede clasificarse como de banda base o de banda ancha. Se denomina de banda base cuando puede utilizarse el medio de transmisión en un momento dado. El medio de transmisión puede compartirse utilizando técnicas de tiempo compartido o por multiplexación por división de tiempo. En banda base se utilizan señales digitales y estas señales utilizan todo el espectro de frecuencias del medio. Al utilizar topología de bus, puede utilizarse cable coaxial de 50 ohms y una longitud máxima de 500 ms sin uso de repetidores y 2.5 Kms. con repetidores. Un LAN de banda base con cable paralelo se utiliza para redes

de bajo costo y menor desempeño, permitiendo longitudes de hasta 1 Km. y varias decenas de equipos conectados. Las redes de banda base tienen como ventaja su simplicidad y su bajo costo, y su posible desventaja es su limitada capacidad y corta distancia.

El LAN de banda ancha utiliza señales analógicas, posibilitando el uso de multiplexación por división de frecuencia (FDM), permitiendo la división del espectro de frecuencias en canales o secciones; diferentes canales pueden llevar información de datos, TV, radio, etc. Debido a la naturaleza analógica de las señales se permiten distancias de varios kms. El LAN de banda ancha permite cientos y aún miles de dispositivos. La ventaja de las redes de banda ancha es su tremenda capacidad; pero son más complejas y más difíciles de instalar y mantener que las redes de banda base.

Las redes de banda base pueden clasificarse por su topología y por su método de acceso.

a. Topología:

De acuerdo a su topología pueden clasificarse en:

- Bus: Todas las estaciones están conectadas en paralelo a un simple par de cables, cable coaxial o fibra óptica. El ejemplo más conocido de este tipo de red es la Ethernet. Para usuarios con un gran número de terminales y alto tráfico, el bus de banda ancha es la mejor opción; para

requerimientos menores, puede utilizarse el bus de banda base por ser un sistema más simple.

- Anillo: en este tipo de red una estación se conecta a la siguiente, con la primera estación conectada a la última. La transmisión es unidireccional, los datos se transmiten secuencialmente bit por bit; de un repetidor a otro; cada repetidor regenera y retransmite cada bit. En el anillo deberá ser posible insertar, recibir y remover datos. El repetidor pasa todos los datos que llegan a él y provee un punto de acceso para que las estaciones puedan recibir y enviar datos. Su beneficio principal es el uso de comunicaciones punto a punto, permitiendo mayores distancias que el bus de banda base. Entre sus problemas potenciales están: la falta de un repetidor deshabilita la red. Para instalar un nuevo repetidor se requiere la identificación de los repetidores adyacentes y al cerrarse el anillo se necesita un mecanismo para remover los paquetes circulantes.

- Estrella: Todas las estaciones se conectan a una estación central que provee conmutación y derechos de acceso a las demás estaciones.

- Malla: todas las estaciones se conectan entre sí en configuración punto a punto. Si no existe una interconexión total entre todas las estaciones, las estaciones intermedias deben servir como nodos de enrutamiento.

b. Método de acceso:

De acuerdo a su método de acceso, las redes de área local pueden clasificarse en:

- CSMA/CD (Carrier sense multiple access/ collision detected): es el método más utilizado en las topología de bus, también se le conoce como "escucha mientras habla" (LWT, listen while talk). La versión original en banda base fue desarrollada por Xerox como parte de la red local Ethernet. El protocolo básico consiste en que una estación con un mensaje a enviar debe monitorear el bus para saber si otra estación está enviando. Si otra estación está enviando, retardará el envío del mensaje hasta que se libere el bus. Si el bus se encontraba libre, enviará el mensaje inmediatamente. Si dos estaciones desean enviar un mensaje al mismo tiempo y están separadas en el bus, es posible que envíen su mensaje a aproximadamente el mismo tiempo y al sobreponerse las señales, éstas se interferirán y se producirá una colisión. El protocolo requiere que al ocurrir una colisión, cada estación cese su transmisión, espere un tiempo aleatorio y luego intente transmitir de nuevo. Este tiempo aleatorio de espera da una probabilidad razonable de que una estación empiece a transmitir con suficiente tiempo antes para que las otras estaciones determinen que el canal está ocupado y no transmitan.

Existen varios algoritmos para generar este tiempo aleatorio. Un método se basa en el conocimiento del mayor

número de estaciones de la red y escogiendo un número aleatorio con probabilidad de retransmisión después de colisión inversamente proporcional a ese número. Por ejemplo, en un sistema con 100 estaciones, luego de una colisión, una estación esperará para retransmitir un tiempo aleatorio entero entre 1 y 100 "intervalos de tiempo". Donde un "intervalo de tiempo" es el tiempo requerido para que la señal viaje a un extremo del bus y regrese. La principal desventaja de este método consiste en intervalos de espera demasiado largos.

Otro método es el binario exponencial (binary exponential backoff). En este método el número aleatorio de intervalos de tiempo que espera una estación para retransmitir no es función del número de estaciones sino del tráfico. Una estación tratará de retransmitir repetidas veces al ocurrir colisión, pero después de una colisión el valor promedio del tiempo de retardo se duplica.

- Paso de testigo en bus (Token bus): es una técnica en la cual las estaciones en el bus forman un anillo lógico, dando a cada estación una posición secuencial, con la última estación seguida por la primera. Cada estación conoce a su predecesor y a su sucesor. Un paquete de acceso conocido como token regula el derecho de acceso; cuando una estación recibe el token tiene derecho para utilizar, el medio de transmisión por un tiempo determinado; durante este tiempo

puede enviar paquetes de información, puede polear otras estaciones o puede pedir datos. Al terminar de enviar información la estación o cuando se terminó su tiempo, esta estación pasa el token a la siguiente. Debe notarse que el orden lógico de las estaciones es independiente de su orden físico. El sistema debe ser capaz de poner en su estado inicial al anillo al comenzar el sistema, adicionar una nueva estación, eliminar una estación, manejo de fallas, etc.

La principal desventaja de este método es su complejidad, otra desventaja es la necesidad de encabezados en los paquetes de información. Sus principales ventajas son las siguientes: se puede regular el tráfico dando diferentes prioridades a las diferentes estaciones utilizando diferentes tiempos de retención del token para cada estación; además no necesita un tamaño mínimo del paquete y no se necesita estar "oyendo mientras se habla"; el tiempo de espera para transmitir está determinado y no es un valor probabilístico (como en CSMA/CD).

- Paso de testigo en anillo (Token ring): está basado en un token que circula en el anillo; cuando todas las estaciones están libres, el token se denomina "token libre". Si una estación desea transmitir, debe esperar a que pase el token libre; entonces cambia el estado del token de libre a "ocupado", alterando su patrón de bits, la estación envía su mensaje inmediatamente después del token. La estación que

cambió el token debe convertir el token ocupado a token libre al retornar después de haber completado una vuelta alrededor del anillo. Cada estación del anillo debe decidir si el mensaje es transmitido para ella, si ese es el caso, marca el mensaje y lo retransmite indicando que se recibió adecuadamente. Debe observarse que cada estación necesita poder repetir adecuadamente los diferentes mensajes. Si una estación no funciona bien, el mensaje no retornará a la estación enviante o no llegará a la estación destinataria. Generalmente se designa a una estación como monitor del token; si el monitor detecta que no recibe el token en un tiempo programado, determinará que el token se ha perdido, reinicializará el anillo y producirá un nuevo token libre. Su principal ventaja es su posibilidad de regular el tráfico, ya sea permitiendo a diferentes estaciones transmitir diferentes cantidades de información cuando reciben el token o dando prioridades a las estaciones para reclamar el token circulante. La principal desventaja consiste en la necesidad de dar mantenimiento al token.

- Anillo ranurado (Slotted ring): en este método un número de intervalos de longitud fija giran continuamente en el anillo. Cada intervalo contiene un bit inicial para indicar si el intervalo está vacío o lleno; todos los intervalos son inicializados como vacíos. Cuando una estación desea enviar un mensaje espera hasta encontrar un

intervalo libre, le cambia la bandera e inserta el paquetes de datos en el intervalo. La estación no puede transmitir otro paquete hasta que el intervalo regresa; este intervalo puede contener bits de respuesta indicando si la estación destinataria aceptó el mensaje, lo rechazó, o estaba ocupada. Como cada estación conoce el número de intervalos existentes en el anillo, puede cambiar la bandera de lleno a vacío en el intervalo correspondiente.

Su principal desventaja es el desperdicio de ancho de banda. Un intervalo usualmente llevará más bits de encabezado que de datos, y una estación puede enviar solo un paquete por intervalo de tiempo alrededor del anillo. Su principal ventaja es su simplicidad.

Se han realizado estudios estadísticos acerca del desempeño de los tres protocolos estandarizados por el IEEE 802: CSMA/CD, token bus y token ring:

- Para parámetros dados, mientras menor sea la longitud del paquete, mayor será la diferencia entre la salida de los protocolos que utilizan token y CSMA/CD.

- Token ring es el menos sensitivo a la carga de trabajo.

- CSMA/CD ofrece el menor retardo bajo carga ligera, mientras que es la más sensitiva bajo carga pesada.

3. Teleproceso. El tercer nivel de comunicaciones es el teleproceso. En este sistema se utiliza

usualmente técnicas de menor velocidad, y se usan las redes telefónicas públicas. Se utiliza cuando los nodos están sumamente dispersos y las velocidades de transmisión son relativamente bajas: de unos 300 bps a 100 Kbps.

Los datos en las computadoras están codificados en una forma binaria y esta información puede ser transmitida y/o procesada en forma paralela o serial. En un sistema que maneja la información de manera serial, los códigos digitales son transmitidos bit por bit. En la transmisión de datos generalmente se utiliza la transmisión serial, ya que el equipo de transmisión se simplifica. Pero también es técnicamente posible transmitir datos en paralelo, en el canal de comunicación se pueden acomodar varias frecuencias o se pueden utilizar varios canales de comunicación simultáneamente. En los sistemas de transmisión en paralelo se envía un carácter a la vez y son los sistemas utilizados principalmente en los sistemas de alta velocidad.

Para que un sistema de comunicaciones funcione adecuadamente, es importante que tanto el transmisor como el receptor trabaje a una misma frecuencia y con una misma fase.

La transmisión de información puede realizarse de dos formas: sincrónica o asíncrona.

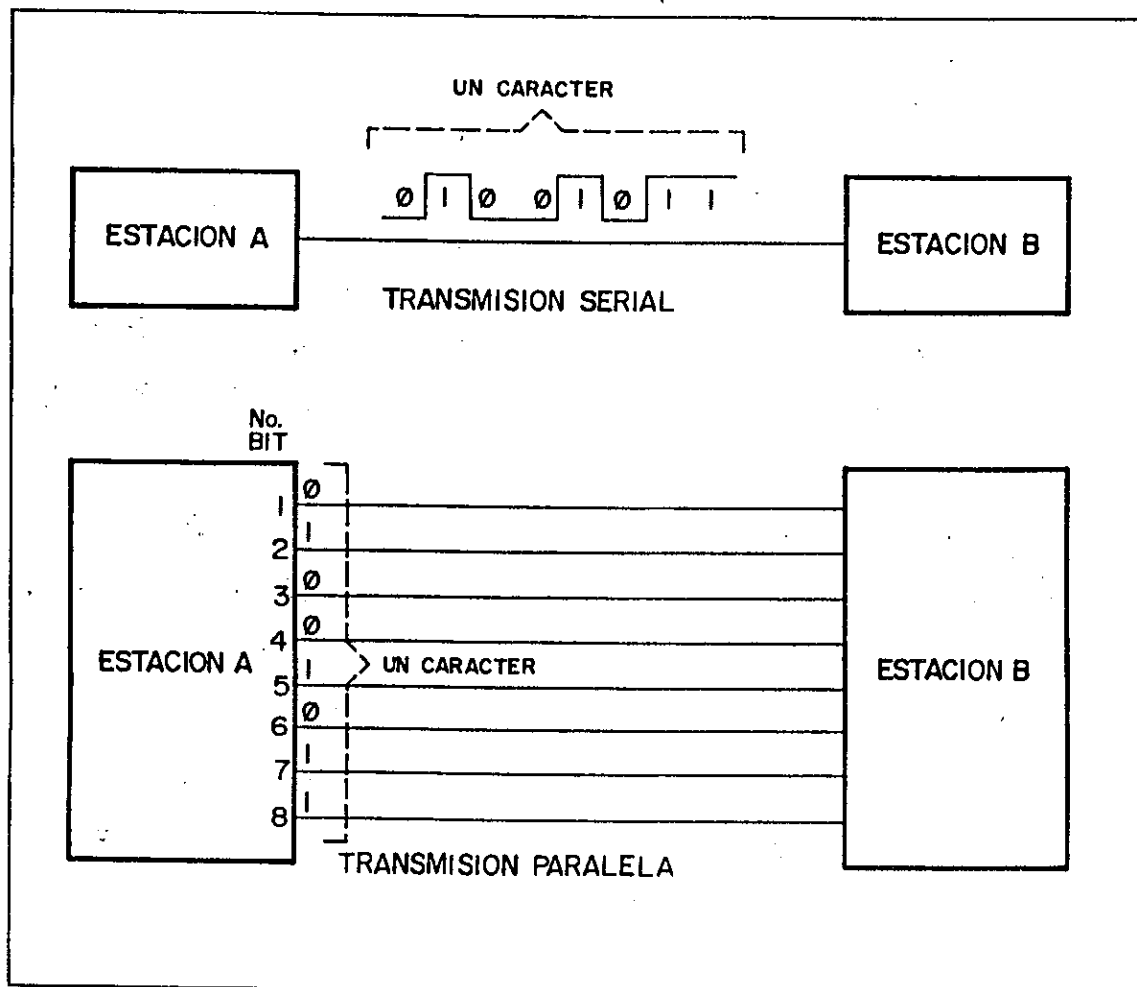


FIGURA 4.3
 TRANSMISION EN SERIE Y PARALELO

La comunicación asíncrona utiliza como elemento de sincronización información adicional que acompaña a los bits de información. Usualmente se utilizan los bits de arranque y los bits de parada. El bit de inicio indica el principio de la palabra o carácter y el bit de parada indica el fin de carácter. La transmisión asíncrona se utiliza principalmente en sistemas de teleproceso de baja y mediana velocidad y se adapta perfectamente a la forma de transmisión de información en sistemas con terminales interactivas, por

ejemplo, el ingreso de datos por un operador.

En la comunicación sincrónica, la información es transmitida de una manera continua y uniforme con base en un reloj de referencia, por lo que no es necesaria la utilización de los bits de inicio y de parada. En algunos casos, entre los bits de información se envía una señal de sincronismo. La transmisión sincrónica es mucho más eficiente, sofisticada y costosa que la asíncrona. En algunos casos, se utilizan memorias auxiliares (buffers) para convertir una transmisión asíncrona en sincrónica.

La velocidad de transmisión de datos se puede medir en baudios; la velocidad en baudios mide la cantidad de veces que una señal cambia de estado por segundo. Es una medida que señala la máxima velocidad con que las señales eléctricas pueden cambiar de estado dentro del canal de comunicación. En muchas ocasiones, se confunde la velocidad en bits por segundo y la velocidad en baudios; el único caso en que la equivalencia es verdadera, es cuando una variación de la señal eléctrica conlleva únicamente un bit de información. Cuando un estado del canal lleva la información de más de un bit de información (por ejemplo, en QPSK existen 4 estados posibles, y cada estado conlleva dos bits de información) la velocidad en bits por segundo y en baudios es diferente.

Existe una gran variedad de códigos para representar la

información, entre los que se encuentran:

- ASCII: (American Standard Code for Information Interchange), es el código standard actual para teleproceso y para transmisión de datos y tiene un equivalente en el CCITT (Comité Consultivo Internacional para Telegrafía y Telefonía) que es el código CCITT No. 5, y únicamente difieren por ciertos caracteres de control. En las figuras 4.4 y 4.5 se muestran dichos códigos.

- Código Baudot: código de 5 bits utilizado en teletipos lentos.

- Binary Coded Decimal: código de 6 bits utilizado en líneas de baja velocidad.

- EBCDIC (Extended Binary Coded Decimal Interchange Code): código IBM standard de 8 bits.

	+0	+1	+2	+3	+4	+5	+6	+7
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL
8	BS	HT	LF	VT	FF	CR	SO	SI
16	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB
24	CAN	EM	SUB	ESC	FS	GS	RS	US
32		!	"	#	\$	%	&	'
40	()	*	+	,	-	.	/
48	0	1	2	3	4	5	6	7
56	8	9	:	;	<	=	>	?
64	@	A	B	C	D	E	F	G
72	H	I	J	K	L	M	N	O
80	P	Q	R	S	T	U	V	W
88	X	Y	Z	[\]	^	_
96	`	a	b	c	d	e	f	g
104	h	i	j	k	l	m	n	o
112	p	q	r	s	t	u	v	w
120	x	y	z	{		}	~	DEL

FIGURA 4.4 EL CONJUNTO DE CARACTERES ASCII

				b ₇	0	0	0	0	1	1	1	1
				b ₆	0	0	1	1	0	0	1	1
				b ₅	0	1	0	1	0	1	0	1
					0	1	2	3	4	5	6	7
b ₄	b ₃	b ₂	b ₁									
0	0	0	0	0	NUL	TC ₇ (0x0)	SP	0	@	P	.	p
0	0	0	1	1	TC ₆ (0x1)	DC ₁	1	1	A	Q	a	q
0	0	1	0	2	TC ₅ (0x2)	DC ₂	2	2	B	R	b	r
0	0	1	1	3	TC ₄ (0x3)	DC ₃	3	3	C	S	c	s
0	1	0	0	4	TC ₃ (0x4)	DC ₄	4	4	D	T	d	t
0	1	0	1	5	TC ₂ (0x5)	TC ₈ (0x5)	%	5	E	U	e	u
0	1	1	0	6	TC ₁ (0x6)	TC ₉ (0x6)	&	6	F	V	f	v
0	1	1	1	7	BEL	TC ₀ (0x7)	'	7	G	W	g	w
1	0	0	0	8	FE ₈ (0x8)	CAN	(8	H	X	h	x
1	0	0	1	9	FE ₉ (0x9)	EM)	9	I	Y	i	y
1	0	1	0	10	FE ₁₀ (0xA)	SUB	:	10	J	Z	j	z
1	0	1	1	11	FE ₁₁ (0xB)	ESC	+	11	K	[k	(
1	1	0	0	12	FE ₁₂ (0xC)	IS ₁₂ (0xC)	,	12	L	\	l	
1	1	0	1	13	FE ₁₃ (0xD)	IS ₁₃ (0xD)	-	13	M]	m	}
1	1	1	0	14	SO	IS ₁₄ (0xE)	.	14	N	^	n	-
1	1	1	1	15	SI	IS ₁₅ (0xF)	/	15	O	_	o	DEL

FIGURA 4.5 EL ALFABETO INTERNACIONAL No. 5

Las líneas telefónicas fueron realizadas pensando en la transmisión de información de manera analógica entre 300 y 3,400 Hz (voz humana). Por lo que para enviar información de manera digital es necesario convertir la señal digital en analógica y luego la señal analógica a una señal digital para poder ser utilizada por un computador.

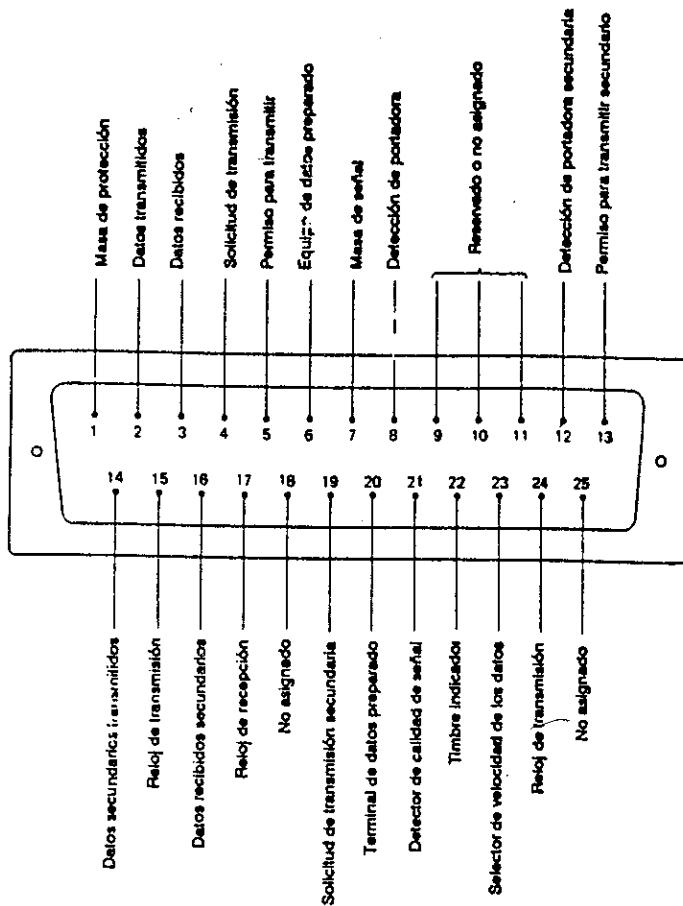
Al dispositivo encargado de realizar esta función de convertir la señal digital en analógica y viceversa se le denomina MODEM (contracción de las palabras MODulador y DEModulador). La señal que entra y sale del computador debe

ser señal digital, el modem se encarga de convertir esta señal digital en analógica para transmitirla a través del canal, y luego realiza la conversión de analógica a digital en el otro extremo.

El modem puede ser capaz de establecer un enlace de comunicación con otro modem localizado remotamente. Para esto se utilizan los "protocolos de comunicación", que permiten establecer adecuadamente dicho enlace. Otras funciones que puede realizar el modem son: mantener una supervisión del estado del enlace, realizar modificaciones a la información digital que se transmite, procesos de control de errores, etc. En los procesos de modulación, puede utilizar diferentes técnicas, llegando a las más avanzadas (QPSK, QAM).

Existen modems para transmisión sincrónica y asincrónica, y modems para línea dedicada y para línea conmutada. También existen modems half dúplex y full dúplex; modems con especificaciones CCITT o BELL, etc.

La interface generalmente utilizada entre el equipo terminal de datos (DTE) y el equipo de transmisión de datos (DCE) es la EIA-RS-232-C, que es la interface americana establecida por la EIA (Electronics Industries Association). Esta interface posee también un equivalente dentro del CCITT, siendo la V.24. Ver figura 4.6.



Línea	Círculo	Fuente	Descripción
1	AA	.	Masa de protección
2	BA	ETD	Datos transmitidos
3	BB	ETCD	Datos recibidos
4	CA	ETD	Solicitud de transmisión
5	CB	ETCD	Permiso para transmitir
6	CC	ETCD	Equipo de datos preparado
7	AB	.	Masa de señal
8	CF	ETCD	Detector de señal de línea recibida
9	.	.	Reservado para pruebas del equipo de datos
10	.	.	Reservado para pruebas del equipo de datos
11	.	.	No asignado
12	SCF	ETCD	Detector de señal de línea secundaria recibida
13	SCB	ETCD	Permiso para transmitir secundario
14	SBA	ETD	Datos secundarios transmitidos
15	DB	ETCD	Sincronismo del elemento de señal en transmisión
16	SBB	ETCD	Datos secundarios recibidos
17	DD	ETCD	Sincronismo del elemento de señal en recepción
18	.	.	No asignado
19	SCA	ETD	Solicitud de transmisión secundaria
20	CD	ETD	Terminal de datos preparado
21	CG	ETCD	Detector de calidad de señal
22	CE	ETD	Timbre indicador
23	CH	ETD	Selector de velocidad de la señal de datos
24	CI	ETCD	Selector de velocidad de la señal de datos
25	DA	ETD	Sincronismo del elemento de señal en transmisión
	.	.	No asignado

FIGURA 4.6 INTERFACE DIGITAL RS-232-C/CCITT V.24

En ciertas ocasiones se tienen varias computadoras y/o terminales localizadas en un punto lejano del nodo central, pero adyacentes entre sí. En este caso es muy costoso utilizar una línea o canal por dispositivo; se puede usar una compartición de la línea. Esta compartición de la línea puede realizarse con un sistema multipunto o un multiplexor.

En los sistemas multipunto, varias estaciones utilizan un mismo canal de comunicación. Este sistema puede resultar ventajoso cuando las estaciones pueden funcionar a través de un canal de comunicación, por ejemplo, varias terminales remotas localizadas en un mismo punto y se pueden interconectar todas las terminales por un solo canal. En este sistema se utiliza un controlador especial encargado de administrar los recursos de comunicaciones comunes. Además, el equipo terminal es más sofisticado, ya que es necesario que pueda participar y responder a un control común del canal de comunicación. La estación controladora de la red se denomina "estación amo" y las demás estaciones se denominan "estaciones esclavo".

En el polling, varias estaciones terminales comparten una línea, la computadora central pregunta a cada terminal o nodo remoto si tiene algo que enviar; si no se tiene nada que enviar se pregunta a la siguiente terminal, si tiene mensajes, la terminal envía los datos y luego la computadora pasa a la siguiente terminal.

La multiplexión puede verse como una compartición de un recurso por actividades similares independientes. En el campo de la transmisión de datos se utiliza la multiplexión para enviar a través de un canal de alta velocidad, simultáneamente, una serie de datos provenientes de varios dispositivos. Por ejemplo, en la transmisión en la primera jerarquía de PCM se multiplexan 30 canales de voz, estando la información de cada canal codificada a una velocidad de 64 Kbps, para transmitir un solo flujo de información a 2 Mbps. Se utilizan principalmente dos tipos de multiplexión:

- Multiplexión por división de frecuencia (FDM, frequency division multiplexing): en este caso un canal que tiene un ancho de banda amplio es compartido por varias transmisiones. Las transmisiones comparten la banda total al utilizar bandas más estrechas; por ejemplo, las transmisiones en AM.

- Multiplexión por división de tiempo (TDM, time division multiplexing): en el cual el tiempo se divide en fragmentos (time slots) y cada fragmento es utilizado para transmitir la información de una de las terminales o nodo remoto.

Entre las causas de error pueden mencionarse:

- Diafonía.
- Efectos del eco.
- Ruido

- Atenuación.
- Distorsiones no lineales.
- Interferencias externas.
- Cambios en los niveles de potencia.
- Falla en la línea.
- Distorsiones causadas por retardos.

Para solucionar los errores se pueden utilizar códigos detectores de errores, y hacer que el mensaje sea retransmitido, ó códigos detectores y correctores de errores.

En algunos casos el mensaje se retransmite únicamente cuando lo solicita el receptor, en otros casos, el mensaje se retransmite siempre. Para la detección de errores usualmente se utilizan chequeos de paridad o de redundancia. En el chequeo de paridad, el transmisor cuenta el número de unos del carácter, y dependiendo del tipo de paridad, se agrega un uno o un cero. En el chequeo de paridad par, se agrega o no un uno para que el número total de unos del carácter sea par. Con este método se puede detectar un error; pero si existen dos errores, éstos no se detectan.

En el chequeo de redundancia se utiliza un carácter completo, llamado carácter de chequeo de bloque, para chequear cada bloque de datos transmitido. Existen otros muchos métodos de codificación, como el código de Hamming, códigos cíclicos, códigos extendidos, etc.

Debe tomarse en cuenta que a pesar de las ventajas que presentan los códigos detectores y/o correctores de errores, al agregar bits o caracteres de chequeo (que no llevan información útil al usuario), se hace menos eficiente la comunicación.

En la comunicación de un grupo de computadoras localizadas físicamente cerca con otro grupo de computadoras cercanas entre sí, pero alejadas del primer grupo, se puede utilizar un procesador especial conocido como gateway. Este procesador provee una interface entre las terminales conectadas al sistema con el resto del mismo. Es utilizado en el mundo de las comunicaciones para representar el lugar donde dos redes, cada una con diferentes protocolos y convenciones, se unen. Existe un gran número de alternativas que pueden aliviar a los nodos procesadores de los problemas de comunicación; esta serie de alternativas se definen en relación a la cantidad de funciones que se ponen en el procesador gateway.

B. Arquitectura de comunicaciones

A continuación se estudia la clasificación de los sistemas de comunicación.

Un sistema de comunicación consiste en tres elementos principales: elementos de procesamiento o nodos, canales de comunicación y conmutadores o switches. Los canales de comunicación son los enlaces que sirven para enviar mensajes

entre los nodos. Las decisiones de enrutamiento se realizan en los conmutadores del sistema, estos conmutadores pueden ser elementos de hardware especializados o procesadores con funciones de conmutación. La taxonomía de un sistema de comunicación consiste en 4 decisiones básicas:

Decisión	Alternativa
1. Estrategia de transferencia	Directa, indirecta
2. Método de transferencia de control	Centralizada, descentralizada, ninguna
3. Estructura del camino	Dedicada, compartida
4. Arquitectura del sistema	Anillo, estrella, completamente mallada, bus

La primera decisión de diseño es acerca de la estrategia de transferencia. Cuando es directa, el mensaje se envía del nodo fuente al destino sin ningún procesamiento intermedio. Cuando es indirecta, se realiza un procesamiento indirecto de enrutamiento, realizado por los conmutadores intermedios.

El segundo nivel de decisión envuelve el método de transferencia de control usado en la transferencia indirecta. En la transferencia directa no interviene ningún control. Si existe transferencia indirecta al menos existe un nodo de conmutación intermedio que realiza funciones de enrutamiento y control. Esta función de control puede ser centralizada o descentralizada. Con el control centralizado

todo el enrutamiento se realiza en el conmutador central; con el control distribuido, existen varios nodos, posiblemente todos, que pueden realizar funciones de enrutamiento. Para que los nodos puedan realizar estas funciones de enrutamiento es necesario que los nodos de conmutación tengan cierta información; la información principal requerida es: definición de los nodos fuente y destino y la matriz de enrutamiento que indica la interconexión entre los nodos y las características de dicha interconexión.

La tercera decisión es acerca de la estructura del camino; éste puede ser dedicado o compartido. Un canal dedicado se usa para comunicar únicamente a dos nodos específicos. Un sistema en estrella o totalmente mallado puede utilizar estos canales. Un canal compartido es utilizado por varios nodos. Las estructuras de bus utilizan canales compartidos.

El cuarto nivel determina la arquitectura del sistema. Este nivel diferencia arquitecturas que son similares en los niveles anteriores. Por ejemplo, memoria compartida o bus global pueden implementarse con transferencia directa y canales compartidos.

Un sistema de comunicación puede evaluarse de acuerdo a los cinco criterios siguientes:

- Modularidad:

La modularidad se refiere a la facilidad que tiene un sistema para poder expandirse. Los dos extremos se tienen en la configuración de estrella y la totalmente mallada. Con la topología de estrella los nodos se comunican a través de un nodo central y existe una alta modularidad. Al agregar un nuevo nodo, únicamente se interconecta al nodo central; pero si se hace necesaria la adición de un nodo central para mejorar la confiabilidad, la interconexión puede tornarse sumamente compleja. Con un sistema totalmente mallado, la modularidad es sumamente baja, y el agregar un nodo a un sistema de n nodos existentes, se debe de proveer de n canales de comunicación adicionales, en este caso el número de canales de comunicación del sistema está en relación al cuadrado del número de nodos y los algoritmos de enrutamiento se vuelven sumamente complejos.

- Flexibilidad de interconexión:

La flexibilidad de interconexión determina las diferentes formas en que puede enrutarse un mensaje entre dos nodos. Al existir más caminos posibles, aumenta la flexibilidad. Una red en estrella es menos flexible que una red mallada. A pesar de la necesidad de algoritmos de enrutamiento más complejos, una mayor flexibilidad se reflejará en una mayor confiabilidad del sistema.

- Tolerancia a fallas:

La tolerancia a fallas trata de la capacidad del sistema

de seguir operando a pesar de la ocurrencia de una o más fallas. Pueden ocurrir fallas en los nodos o en los canales de comunicación. Si un nodo falla, el procesamiento en ese nodo debe detenerse o transferirse a otro nodo. Pero en cuanto a comunicaciones se refiere, esta tolerancia de fallas trata de la posibilidad de transferir un mensaje a pesar de la falla en un nodo o en un canal de comunicación. En una red mallada esto es posible debido a la multiplicidad de rutas de un nodo a otro. En una red en estrella la falla del nodo central resulta catastrófica.

- Cuellos de botella:

El cuarto criterio son los cuellos de botella; al aumentar la carga del sistema aparecen cuellos de botella en diferentes lugares. En una red en estrella es probable que ocurra en el conmutador central. Idealmente, el sistema debe ser modular para poder eliminar los cuellos de botella cuando se presentan.

- Complejidad lógica:

Se refiere al número y complejidad de las decisiones que deben realizarse para controlar la comunicación entre fuente y destino. En una red mallada se necesitan algoritmos de enrutamiento más complejos que en una red en anillo. Además, esta complejidad aumenta al introducir procedimientos de tolerancia de fallas.

V. REDES DE COMPUTADORAS

Una red de computadoras es una colección de computadoras y terminales conectadas a través de un sistema de comunicación. Una red de computadoras puede soportar una o más aplicaciones diferentes: puede crearse para ser utilizada por una organización, ó puede ser creada como una red comercial de tiempo compartido como TYMSHARE ó como una red de servicios (VAN, value added network) como TELENET.

En este tipo de red los usuarios pueden entrar a la red a través de una terminal y acceder bases de datos, hardware ó software de sitios remotos; el usuario puede acceder equipos y servicios que no se encuentran disponibles en la computadora local.

Los proveedores de servicios de la red pueden recuperar sus costos y aumentar ganancias al aumentar el número de los usuarios a través de la red; una de las ventajas de la red es la posibilidad de ofrecer servicios que serían económicamente imposibles de ofrecerlos en un solo sitio.

Una red está compuesta de nodos (ó centros) y enlaces o circuitos. Un circuito es un camino entre dos puntos de la red; los circuitos pueden ser físicos o virtuales. Los circuitos físicos son aquellos que son cableados (hard wired) y no son flexibles. Un circuito virtual es una

conexión lógica entre los enlaces de entrada y salida.

Un mensaje, puede enviarse a través de varios nodos intermedios, a este proceso se le denomina enrutamiento.

A. Objetivos de una red de computadoras

Los objetivos que se persiguen con una red de computadoras son los siguientes:

- Confiabilidad:

Un objetivo primordial es la confiabilidad; si la red no trabaja bien, los usuarios pierden confianza en la red.

En cuanto a comunicaciones, la confiabilidad de la red puede aumentarse utilizando códigos detectores y/o correctores de errores, con lo que se reducen los problemas de procesamiento. También puede utilizarse enrutamiento alternativo entre computadoras y terminales; si una ruta no está operando, puede utilizarse la ruta alterna.

En la comunicación telefónica se utiliza usualmente la conmutación de circuitos, en la que se establece un camino físico mientras dure la comunicación. En las redes de computadoras puede utilizarse la conmutación de mensajes: no existe un camino físico fijo, los mensajes tienen un código de destino, la computadora examina este código y determina la mejor ruta, dependiendo de los recursos de la red y del tráfico en la misma. La información que se transmite se descompone en mensajes (ó paquetes) y éstos se envían a través de la red. Los mensajes (ó paquetes) se chequean y se

reensamblan en el nodo destinatario.

- Facilidad de utilización:

El usuario necesita un interface simple hacia la red; además, los tipos de computadoras, software y lenguajes utilizados deben ser transparentes al usuario. Entre los procedimientos que necesita el usuario están: acceso a la red (identificación, password, etc.), obtención de software de sitios remotos, acceso a bases de datos, componer y procesar información en varios sitios, etc.

- Seguridad, protección e integridad:

Usualmente el usuario se siente más seguro si el software y los datos son locales. Para mejorar la seguridad se pueden utilizar diferentes niveles de protección. Se pueden utilizar claves (passwords) para cada usuario, file, récord e item de un récord. Lo anterior es sumamente costoso y agobiante y no siempre necesario, por lo que debe estudiarse cada caso. También se puede codificar la información que se envía para mantener su integridad.

- Costos:

El procesamiento local es más barato que el procesamiento remoto debido principalmente a los costos de comunicación. En las redes comerciales los costos de comunicación dependen principalmente de: la hora del día, la capacidad de almacenamiento, los costos de procesamiento y los servicios especializados utilizados.

- Tiempo de respuesta:

Puede tomarse el tiempo de respuesta como el tiempo desde que el trabajo penetra en el sistema hasta que el resultado aparece en el sitio del usuario. El tiempo de respuesta es afectado adversamente por tráfico pesado en la red, procesamiento en sitios remotos, etc.

Todo el manejo de datos puede ser realizado por un computador especializado: Un front end processor puede manejar la interface entre la computadora principal y la red, al igual que las terminales. Este procesador ensambla y ordena los mensajes, chequea la seguridad, etc. Un back-end processor puede manejar los trabajos de acceso y lectura de datos, asignados a él por el computador central.

Cuando un conjunto de computadoras se interconectan, el hardware y las facilidades de esa localidad se denominan "el anfitrión" (host) y los recursos que proveen las otras computadoras se deben obtener a través de la red de comunicaciones.

Las principales características que deben tener las redes de computadoras son las siguientes:

- La red debe ser capaz de aceptar sin retardo apreciable una transmisión de información.
- La red debe ser eficiente.
- Los datos que arriban a un anfitrión deben ser procesados rápidamente.

- Cuando una transmisión exige una respuesta, el retardo de dicha respuesta debe ser mínimo.

- El costo del establecimiento, mantenimiento y operación de la red debe ser aceptable.

En la figura 5.1 se muestra una red típica: los cuadrados son las computadoras y los círculos son los centros de conmutación que pueden consistir en computadoras altamente especializadas.

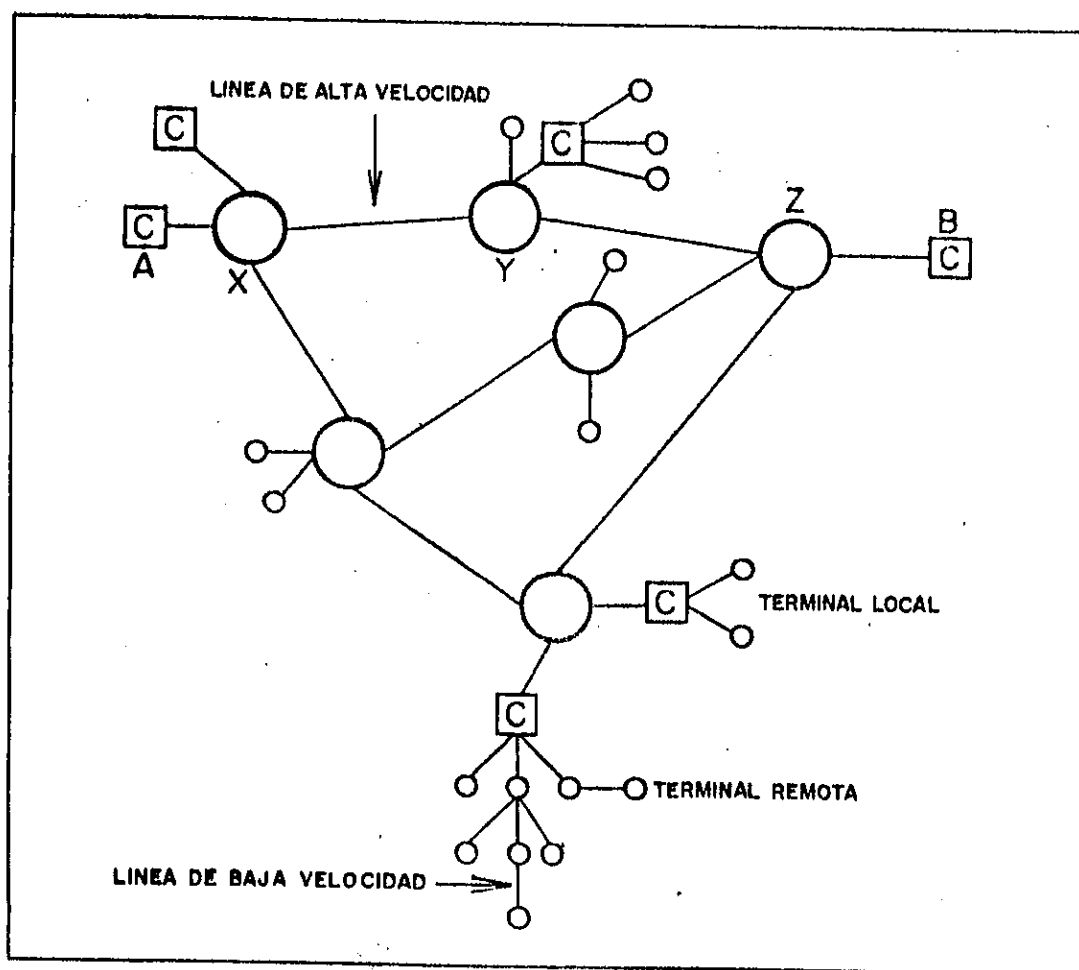


FIGURA 5.1
RED DE COMUNICACION ENTRE COMPUTADORAS

Algunas computadoras pueden tener conectadas varias terminales; las líneas que unen las terminales con la computadora local no necesariamente deben ser de alta velocidad y pueden utilizar un ancho de banda limitado. Para la interconexión con los centros de conmutación, se necesitan líneas de alta velocidad.

B. Clasificación de las redes de computadoras

Las redes de computadoras pueden dividirse en tres tipos: de conmutación de circuitos, de conmutación de mensajes y de conmutación de paquetes.

- Conmutación de circuitos:

Las computadoras se interconectan a través de un camino físico; se utiliza principalmente un par metálico para la comunicación y no se interrumpe mientras dura la transmisión. Este tipo de sistema es análogo a los sistemas de conmutación telefónica.

- Conmutación de mensajes:

Una característica de la comunicación entre computadoras es el hecho que la transmisión de datos ocurre en ráfagas relativamente cortas acompañadas por intervalos de tiempo largos sin transmisión. Típicamente se transmite información únicamente durante el 1 % del tiempo de ingreso de datos, por lo anterior resulta antieconómico tener un circuito dedicado para la comunicación.

En la conmutación de mensajes no se posee un circuito

totalmente dedicado a una comunicación. Tomemos el ejemplo de una computadora A que desea enviar información a la computadora B, se escogió la ruta que pasa por los conmutadores intermedios X, Y y Z. La computadora A envía un mensaje al centro de conmutación X, y el mensaje se almacena en el buffer de dicho conmutador. Este mensaje permanece almacenado en el buffer de X hasta que la línea de X a Y no se necesita para una transmisión anterior; cuando la línea está disponible se envía el mensaje de X a Y; repitiéndose el proceso hasta llegar el mensaje a su destino. Este sistema tiene la desventaja de presentar un retardo entre la transmisión y la recepción del mensaje. En la transmisión total de A a B, en un tiempo particular, generalmente sólo una de las líneas estará ocupada en transmitir el mensaje; las otras líneas pueden utilizarse para transmitir otros mensajes.

Algunas desventajas de la conmutación de mensajes son: Un mensaje puede ser muy largo o muy corto por lo que se hace necesaria la utilización de buffers sumamente grandes; aún estando la línea libre, un mensaje no puede reenviarse de un nodo a otro hasta que se ha recibido totalmente; con mensajes muy largos y varios centros de conmutación intermedios, los retardos pueden ser prohibitivos. También pueden existir problemas cuando se necesita la línea para enviar un mensaje muy corto, pero se está enviando un

mensaje muy largo y la transmisión no puede interrumpirse.

- Conmutación de paquetes:

En este caso el mensaje se subdivide en paquetes de longitud definida; el mensaje se transmite paquete por paquete. Al igual que la conmutación de mensajes, los paquetes se almacenan en un buffer hasta que pueden enviarse. En algunos sistemas de conmutación de paquetes, los paquetes que forman un mensaje pueden llegar al destino con diferentes retardos y utilizando diferentes rutas.

Cada paquete lleva además de los bits de información, una información de encabezado. Este encabezado debe indicar: el destino, para que cada centro de conmutación pueda enrutarlo adecuadamente; la fuente, para que sea posible el envío de mensajes de recepción; e identificación del usuario, para que pueda cobrarse el servicio. También puede contener bits de sincronización para indicar el inicio y el fin del bloque.

En la conmutación de mensajes, un mensaje, sin importar el tamaño, necesita sólo un grupo de bits como encabezado, en la conmutación de paquetes el mensaje se divide en paquetes, cada uno con su encabezado; por lo que en promedio la conmutación de paquetes utiliza más bits de encabezado por bit de información. Además, el hardware necesario para empaquetar y desempaquetar es sumamente complejo.

C. Topología

Al aumentar el número de elementos de la red, aumenta también las formas en que los elementos pueden interconectarse y enumerar la totalidad de topologías posibles es prácticamente imposible. Algunas de las topologías más importantes son:

- Totalmente conectada o mallada:

Existe una conexión directa entre todos los nodos de la red. Dicha topología es económicamente posible cuando existen pocos nodos en la red. Al aumentar el número de nodos, el número de enlaces necesarios aumentan aproximadamente en proporción al cuadrado del número de nodos. A pesar de su alto costo, provee un tiempo de respuesta y una confiabilidad mejor que las demás topologías. Esta topología es especialmente útil en las redes locales donde se necesita un rápido tiempo de respuesta.

- Arbol generalizado:

Esta es una estructura jerárquica multinivel que permite únicamente un único camino entre dos nodos. En la figura 5.2 se muestra esta topología con tres niveles.

- Bus:

Emplea una facilidad de transmisión compartida por los diferentes nodos; esta facilidad se denomina bus. Esta topología puede tener un control centralizado o distribuido

sobre el medio de transmisión; puede existir un nodo que maneje el flujo de información entre un par de nodos ó pueden ser todos los nodos idénticos. En este último caso se debe desarrollar una técnica para poder eliminar la contención en el uso del bus.

- Anillo:

Esta estructura puede ser de dos tipos: en una primera estructura, cada nodo actúa como un nodo que recibe y envía los mensajes, o sea, el nodo no transmite un nuevo mensaje hasta que ha procesado el mensaje que posee. En la segunda estructura, cada nodo recibe bits sólo en determinados intervalos de tiempo. Cada nodo tiene la capacidad de almacenar un cierto número de bits y reconocer alguna secuencia de control de bits que posibilita el intercambio de información entre nodos. El control de la comunicación puede ser centralizado o distribuido. Conceptualmente, el tráfico puede ser en ambas direcciones; pero en la práctica esto acarrea muchas dificultades por lo que la mayoría de diseños tienen un flujo de información unidireccional.

- Estrella simple (Single center, single star SCSS):

Es una topología jerárquica de dos niveles consistente en un solo nodo en el nivel superior y los demás nodos en el nivel inferior conectados al nodo de control. Esta topología es utilizada para conectar terminales a nodos de conmutación o nodos de procesamiento. Algunos ejemplos de esta topología

se encuentran en sistemas de tiempo compartido en universidades y sistemas de comunicación utilizando FABX.

- Multipunto (Single center, multidrop SCMD):

Es un sistema jerárquico multinivel consistente en un nodo en el nivel superior y varios enlaces, cada uno compartido por varios nodos. Esta topología es utilizada cuando se desea proveer un servicio a nodos localizados en un área extensa. Se permite realizar economías al compartir las facilidades de comunicación entre un gran número de usuarios.

- Multiestrella (Multicenter, multistar MCMS):

Es una topología jerárquica multinivel que consiste en varios nodos de conmutación interconectados sirviendo a un gran número de nodos terminales.

- Multicentros, multipunto (Multicenter, multidrop MCMD):

Es una topología jerárquica multinivel utilizada para proveer un servicio a usuarios localizados en una gran área. Los nodos de conmutación están interconectados entre sí, y cada nodo de conmutación da servicio a un número de usuarios que comparten las facilidades de comunicación de acuerdo a la topología SCMD. Al aumentar el número de nodos, disminuye el costo de las líneas compartidas porque los nodos están más cerca de los usuarios; pero el costo de la red de conmutación aumenta.

En la figura 5.2 se muestran las topologías anteriores.

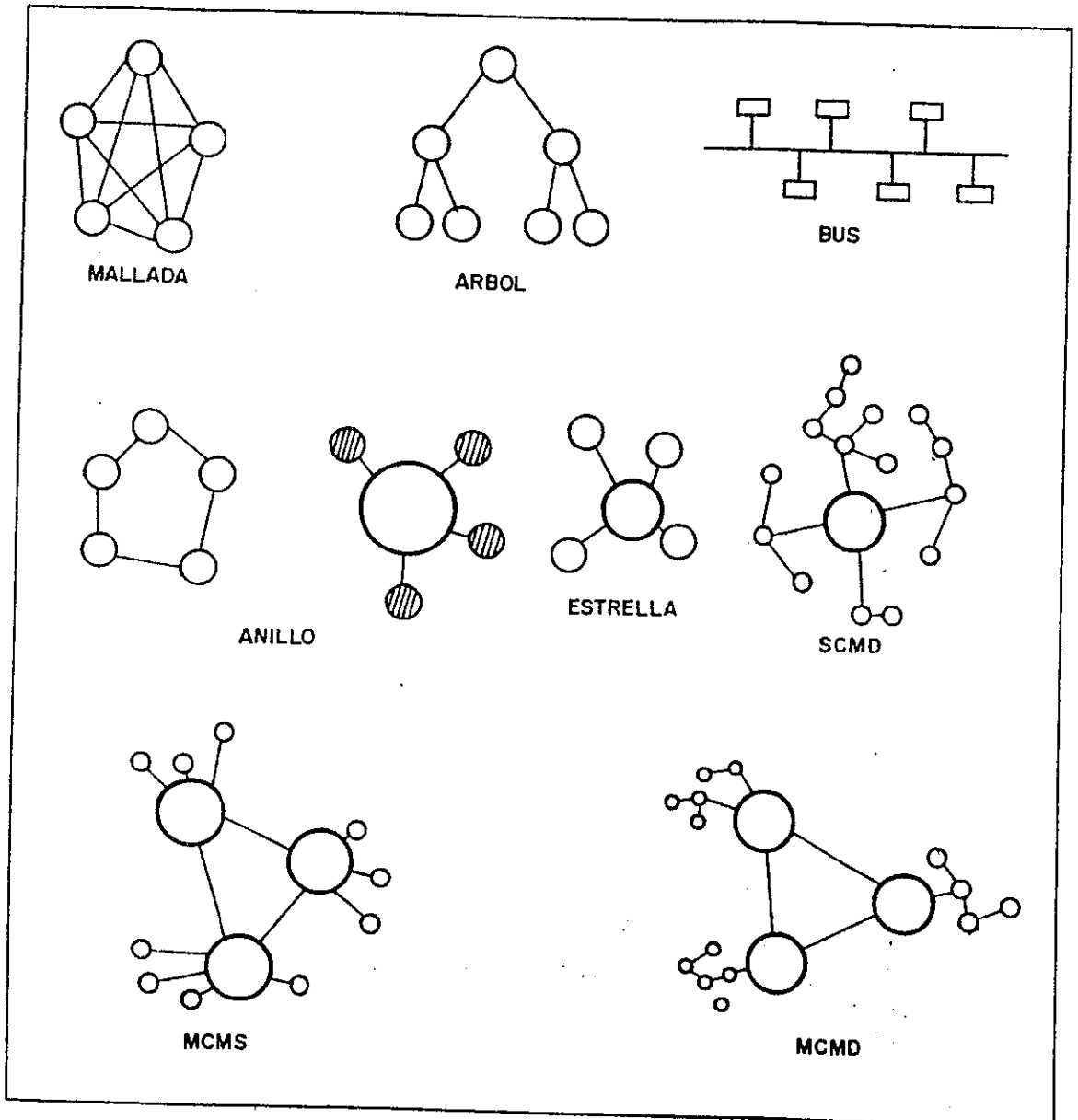


FIGURA 5.2
CONFIGURACIONES TOPOLOGICAS

Un sistema también puede definirse como jerárquico, de igualdad y compuesto.

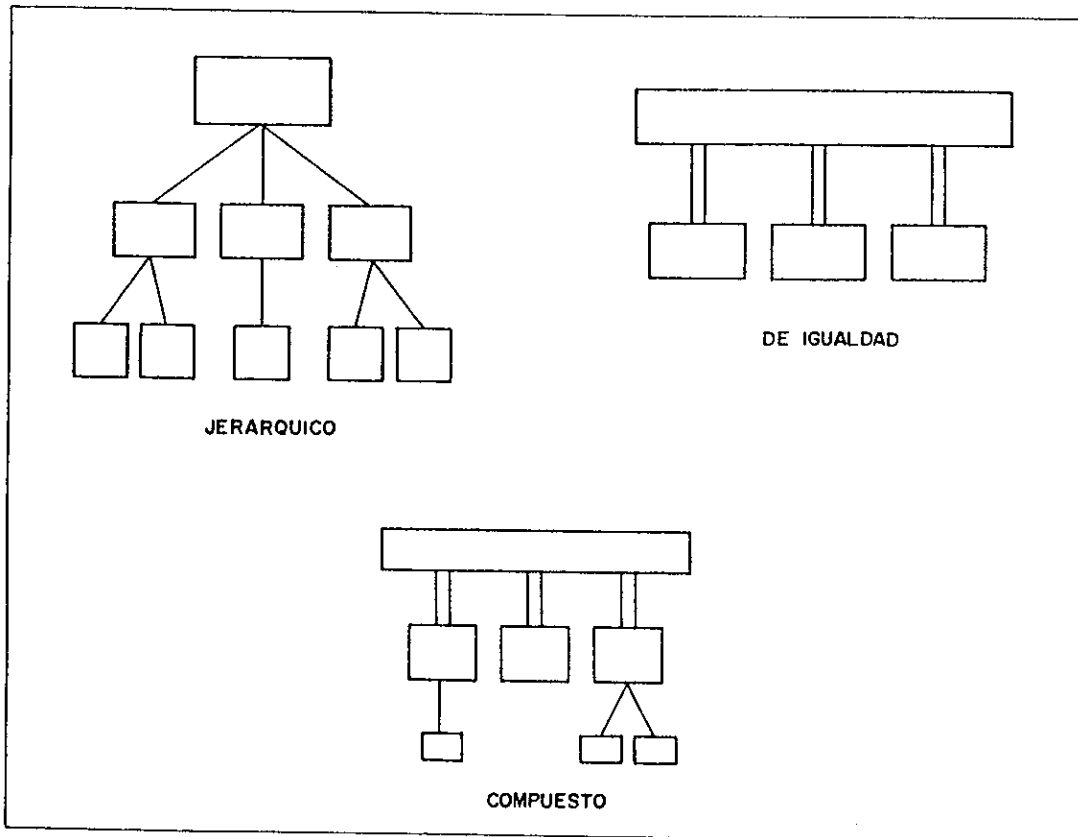


FIGURA 5.3
ESTRUCTURA DE SISTEMAS

Se debe tomar en cuenta que el flujo de datos y el flujo de control en un sistema pueden tener diferentes estructuras; por lo tanto, se necesita la forma del flujo de datos y la forma del flujo de control para definir el sistema. Aun más, diferentes aplicaciones pueden presentar diferentes flujos de datos; una aplicación puede utilizar división de datos de manera de igualdad entre nodos, mientras que otra aplicación puede utilizar división de datos de manera jerárquica.

- Sistema funcionalmente jerárquico:

Es cuando las funciones de manejo de la red están jerarquizadas pero la lógica de aplicación permanece esencialmente intacta en los procesadores operacionales.

- Sistema jerárquico por aplicación:

Nace de la forma jerárquica natural que puede tener una aplicación. De esta forma la aplicación puede derivarse de la estructura organizacional en la cual se pueden percibir claros niveles de función y responsabilidad a través de los niveles de dirección de la empresa.

- Sistema de igualdad:

Es aquel en el cual los nodos cooperan el uno con el otro como iguales. En su forma más pura, las aplicaciones y el control de la red se dispersan a través de la misma, sin ningún punto de control definido. Es posible definir sistemas en los que se tienen procesadores de aplicaciones en configuración de igualdad y procesadores de control en configuración jerárquica.

Una red de computadoras también puede definirse dependiendo de la función realizada por los nodos que la forman:

Un sistema es homogéneo si todos los nodos realizan la misma función. Esto no implica que los nodos sean duplicados físicos el uno del otro; puede variar la cantidad de memoria secundaria o la velocidad del procesador y hasta puede

contener elementos de diferentes fabricantes.

Un sistema es heterogéneo cuando cada nodo procesador realiza diferentes aplicaciones; en un sistema heterogéneo la interacción entre los nodos es mayor que en un sistema homogéneo.

D. Enrutamiento

Para realizar una conexión entre dos computadoras, pueden utilizarse diferentes procedimientos de enrutamiento:

- Flooding (torrente):

Con este método se transmite el mensaje de la fuente hacia todos los caminos posibles; al llegar a un centro de conmutación, el mensaje se retransmite a todos los caminos posibles excepto al camino de donde proviene. Con este método el mensaje llega a su destino con un retardo razonable, pero se congestiona la red de una manera innecesaria ya que se transmiten simultáneamente varias copias del mensaje.

- Enrutamiento aleatorio:

Un paquete almacenado en un nodo se envía a otro nodo escogido de una manera aleatoria. En este método únicamente se tiene transmitiendo un paquete en la red, es un método robusto ante fallas; pero ya que el enrutamiento es al azar, no se tiene una garantía que el paquete alcance su destino en un tiempo razonable.

- Enrutamiento fijo:

Cada nodo tiene una tabla de enrutamiento predeterminado, usualmente a partir de datos de tráfico anteriores. Este método es recomendable cuando es posible lograr una buena estimación del tráfico a priori, pero no en los casos en que el tráfico sufre variaciones severas.

- Enrutamiento centralizado adaptativo:

Cada nodo tiene su propia tabla de enrutamiento; pero es una tabla dinámica que se modifica como resultado de órdenes del "supervisor" que monitorea el tráfico en todos los nodos.

- Enrutamiento adaptativo aislado:

En este caso se tiene una tabla de enrutamiento en cada nodo y esta tabla también se modifica de manera dinámica. Existe un control en la red que envía periódicamente a todos los nodos datos del tráfico en la red, para que dichos nodos puedan realizar las modificaciones necesarias en sus tablas de enrutamiento.

E. Ejemplos de redes de computadoras

Algunos ejemplos de redes de computadoras son los siguientes:

- Tymnet:

Es una red de conmutación de paquetes que permite al usuario conectarse a las computadoras de Tymnet para procesamiento interactivo, o con otros usuarios utilizando únicamente la red de comunicación. La compañía Tymnet es

dueña de todos los nodos que se interconectan formando una red. La interconexión entre los nodos se hace a través de canales arrendados o a través de redes de transporte como AT&T; estas redes se denominan de valor agregado (VAN).

Tymnet es una red mallada. El tiempo promedio para transmitir un paquete de 1024 bits hasta su destino, es de 300 ms. La ruta a seguir es determinada por una computadora supervisora central. Cada vez que el usuario penetra a la red, se le puede dar un nuevo camino, pero mientras permanezca en la red, la ruta utilizada es la misma. El paquete se transmite de nodo a nodo a lo largo de toda la red. En cada nodo el paquete que arriba es almacenado y el nodo envía los paquetes en la modalidad "primero en llegar, primero en ser servido" (FIFO, first in, first out). El control central escoge el camino para proporcionar el menor costo y el menor tiempo de respuesta.

Debido al proceso de enrutamiento centralizado y al hecho que mientras el usuario permanezca en la red utilizará la misma ruta, los paquetes siempre arriban en orden a su destino. Si alguno de los paquetes no llega a su destino, se envía al usuario una indicación (NAK, negative acknowledgement) para que reenvíe el paquete.

- Arpanet:

Es una de las primeras redes de computadoras, fue creada por el Departamento de Defensa de los Estados Unidos. Es una

red que utiliza el enrutamiento adaptativo, con esta técnica cada nodo decide hacia que nodo desea enviar el mensaje para llegar a su destino final. Esta decisión se realiza a partir de una tabla basada en el menor tiempo estimado hacia el nodo de destino. Esta tabla se modifica periódicamente a partir de la información obtenida de todos los nodos. Ya que cada paquete puede llegar a su destino por diferente camino, éstos pueden llegar fuera de orden y deben de ser almacenados y reordenados antes de presentarse al nodo destinatario.

- Telenet:

La red de conmutación de paquetes de Telenet fue la primera red comercial en los Estados Unidos. Telenet comenzó en 1975, y para 1978 poseía más de 81 nodos de conmutación (Telenet Central Offices TCO). Los TCO se dividen en tres tipos: Los Clase 1 se interconectan a través de tres caminos; cada TCO Clase 2 o Clase 3 se conecta a un Clase 1. Cada mensaje desde una computadora o terminal se reorganiza en paquetes y se envía a través de la red. Los paquetes se envía de TCO a TCO de acuerdo a las tablas de enrutamiento almacenadas en cada TCO. Telenet soporta los protocolos de la recomendación CCITT X.25. cada paquete contiene 128 bits de datos y algunos bits de control.

Otras redes de datos que vale la pena mencionar son: Transpac, Datapac y la Red de datos noruega.

VI. DISTRIBUCION DE DATOS

En este capítulo se discute la distribución de datos: el lugar donde se almacenan los datos dentro de un sistema distribuido. Las alternativas de localización de los datos añaden complejidad al diseño de la base de datos.

La distribución de datos es el aspecto más complejo del procesamiento distribuido. Existen dos puntos de vista acerca de la importancia de la distribución de los datos. Algunos piensan que la distribución de datos es un aspecto importante del sistema y los usuarios deben conocer y controlar la localización física de los mismos. Otros piensan que la distribución de datos no es, por sí misma, importante, la meta importante es la eficiencia, el acceso a los datos sin importar donde estén localizados.

Existen tres alternativas para la localización de los datos: centralizados, replicados y divididos.

La replicación total implica la existencia de una copia completa de la base de datos en cada nodo de la red. Con la división pura, la base de datos se divide en segmentos que no se intersecan, cada uno de los cuales se localiza en diferentes nodos. La replicación implica varias copias de la base de datos, mientras que en los casos centralizado y dividido se permite únicamente una copia. Existen también

una variedad de híbridos donde se utilizan diferentes métodos para diferentes partes de la base de datos.

Hay cuatro criterios con los cuales debe evaluarse cada alternativa: costos de almacenaje, confiabilidad, costos de acceso y costos de modificación.

Existen intercambios básicos entre los costos de almacenaje y la confiabilidad, y entre los costos de acceso y los costos de modificación. Si sólo existe una copia de la base de datos, los costos de almacenaje son bajos pero la confiabilidad es relativamente baja. Si un nodo falla, entonces la parte de la base de datos almacenada en dicho nodo no es accesible para ningún usuario de la red. En un sistema distribuido, para lograr una buena confiabilidad es necesario que el usuario pueda acceder una copia de la base de datos a pesar de la falla de un nodo; también se puede mejorar la confianza al disminuir el tiempo durante el cual los datos no estaban disponibles después de una falla. La existencia de varias copias hace necesaria una capacidad de almacenamiento adicional.

El segundo intercambio importante es entre los costos de acceso y los costos de modificación. Los costos y el tiempo de acceso se minimizan cuando el uso del canal de comunicación se minimiza. Lo anterior se puede lograr cuando se poseen varias copias de los datos; y en el caso extremo, se tiene una copia de la base de datos en cada nodo; con lo

que la recuperación de los datos es más rápida y no se utilizan los canales de comunicación. En este caso, al tener múltiples copias de los datos, los costos de acceso disminuyen; pero los costos de modificación aumentan, porque si un usuario realiza una modificación, ésta debe de realizarse en todas las copias. La sincronización de las diferentes copias es una de las tareas más difíciles a realizar en las bases de datos distribuidas.

A. Criterios para la localización de los datos

Los diferentes criterios para la localización de los datos son:

1. Centralizado. Con el criterio centralizado, la base de datos completa se almacena en un solo nodo. Inicialmente esta era la única alternativa, pero ahora es únicamente una de las posibilidades para un sistema distribuido. Debido a que hay una copia de la base de datos, se minimizan los costos de almacenamiento. Pero existe un problema, la falla de hardware o de la base de datos en el nodo central produce la caída del sistema.

Los usuarios remotos necesitan utilizar siempre un canal de comunicación para acceder los datos, por lo que los costos de recuperación de la información aumentan. Además, al existir una única base de datos, los costos de modificación son bajos porque sólo se modifica una copia. Ya que existe una copia, no se utilizan técnicas de

sincronización; y estos mecanismos de sincronización llegan a ser sumamente complicados requiriendo comunicación extensiva, y en algunos casos resultan más costosos que los accesos remotos.

2. Dividido. Existen varios métodos para dividir los datos; estos métodos se definen por las decisiones que se toman acerca de:

- El nivel en que se define la división de los datos.
- El criterio utilizado para definir la división.
- La permanencia de una división de datos en un nodo.

Los datos pueden dividirse a nivel de archivo o a nivel de subarchivo. En un sistema con archivos bien definidos es posible dispersarlos a través de los nodos. También es posible realizar la división a nivel de récords de forma que diferentes partes de un archivo puedan estar en diferentes nodos. La división a nivel de subarchivos es preferible cuando algunos nodos tienden a referirse únicamente a ciertas partes de un archivo. Cuando se divide un archivo, el sistema tiende a considerarlo no como un solo archivo, sino como varios archivos menores.

Hay varios criterios sobre la forma de dividir los archivos. Un criterio toma en cuenta algún atributo de los datos; por ejemplo, su localización geográfica. Al realizar la división por localización, se tiende a formar un sistema homogéneo por aplicación, los nodos realizan prácticamente

las mismas funciones. Las divisiones también pueden estar definidas por aplicación; en un sistema de índole comercial, los datos pueden dividirse dependiendo de si son usados por el nodo de inventario, de contabilidad, de pagos, de mercadeo, etc. Al realizar la división por aplicación se tiende a formar sistemas heterogéneos por aplicación, cada nodo se especializa en algún campo.

Otro atributo que puede escogerse para definir una división es el tiempo. Se escoge el tiempo porque se ha observado que en ciertos sistemas, al ingresar los datos, reciben bastante atención, y mientras transcurre el tiempo comienzan a recibir menos atención. Los datos más recientes permanecen en los nodos de nivel más bajo o nodos locales por cierto tiempo, por ejemplo, una semana. Luego de este tiempo el interés sobre dichos datos disminuye y se trasladan al siguiente nivel, y después de un tiempo mayor se trasladan al último nivel. En el ejemplo anterior se puede observar una distribución jerárquica de los datos; pero a pesar de la distribución de los datos, el control puede o no ser jerárquico. Si son los nodos los que deciden cuando se envían los datos hacia niveles superiores, se tiene un sistema no jerárquico; si son los nodos superiores los que piden el envío de los datos, se tiene un sistema de control jerárquico.

Una posibilidad de dividir a nivel de subarchivo se

denomina extracción. Por ejemplo, en el capítulo II al enumerar las etapas desde un sistema centralizado a uno distribuido, en uno de los pasos, se enviaban ciertos datos a los nodos remotos para realizar en ellos la edición ya que no se necesitan todos los datos para realizar la misma. Los campos necesarios pueden extraerse del archivo central y colocarse en los nodos apropiados. Si los datos que se extrajeron no están presentes en el nodo central, se realizó una división; si se permite que los diferentes campos existan en ambos lugares, entonces se realizó una replicación.

Usualmente se utiliza la división cuando se hace referencia frecuente a los datos locales. Si hay una única copia de los datos, como en el caso centralizado, los costos de almacenamiento se minimizan; pero en ambos casos, debido a la existencia de una sola copia, existen problemas de confiabilidad y seguridad. Si se pierden los datos necesarios para una aplicación, aunque el hardware permanezca intacto, esta aplicación no puede utilizarse en el sistema.

A pesar de esta característica común con el sistema centralizado, el sistema dividido posee una ventaja en cuanto a confiabilidad. Si en el sistema centralizado falla el nodo que guarda la base de datos, todos los datos del sistema se pierden; pero si en un sistema dividido cae un

nodo, únicamente se pierden los datos de dicho nodo, y las aplicaciones que no hacen referencia a dichos datos pueden seguir corriendo. A pesar que el sistema dividido puede tener problemas de confiabilidad, es más confiable que el sistema centralizado.

Como generalmente se hace referencia a datos locales, los costos de acceso a los datos también disminuyen. Los costos de modificación de los datos disminuyen porque no se necesita sincronizar varias copias; la excepción ocurre cuando esta modificación envuelve datos de varios nodos, aumentando los costos de comunicación.

3. Replicado. Consiste en poner múltiples copias de los datos en varias partes del sistema.

En la replicación puede almacenarse una copia de la base de datos en cada nodo del sistema con costos de almacenamiento sumamente altos; el mayor beneficio de este método consiste en un gran aumento de la confiabilidad del sistema. Si un nodo falla, los otros nodos pueden continuar operando.

Los costos de acceso de datos son sumamente bajos y los accesos se realizan de manera local, sin utilizar el sistema de comunicaciones. Cuando se utiliza un sistema de sincronización centralizado, se debe enviar al nodo dominante los cambios realizados para que éste posea la última versión de los datos.

Otro caso es cuando se utiliza sincronización global y al realizar una modificación, ésta debe dispersarse a través de todo el sistema. Las modificaciones en un sistema con replicación de datos son sumamente costosas debido a la sincronización de las diferentes copias. La complejidad de los procesos de sincronización dependen directamente del número de copias; además, la cantidad de comunicaciones requeridas aumenta rápidamente al aumentar el número de copias.

El problema esencial en la replicación de datos es la distribución de cambios y la sincronización entre las diferentes copias.

4. Híbrido. Se pueden utilizar técnicas híbridas cuando diferentes partes de la base de datos pueden tratarse con diferentes métodos. Por ejemplo, un archivo puede ser utilizado por la mayoría de los nodos y se puede colocar una copia de dicho archivo en cada nodo (replicación); mientras otra parte de la base de datos puede dividirse debido a que se hace frecuentemente referencia local.

B. Sicronización

Uno de los objetivos principales del manejador de la base de datos es permitir que los usuarios compartan datos. Mientras éstos únicamente accesen datos, no existe problema. Pero si varios usuarios tratan de modificar los mismos

datos, ocurren problemas de sincronización. Si dos usuarios tratan de modificar los mismos datos, uno de los cambios se perderá.

En un sistema distribuido, se debe mantener una consistencia entre las diferentes copias que existen en el sistema.

Los principales problemas que debe solucionar la sincronización son: mantener los datos actualizados de acuerdo a las últimas modificaciones, mantener la consistencia interna de la base de datos y mantener la consistencia entre las diferentes copias de la base de datos.

1. Sincronización en un sistema centralizado. En un sistema centralizado el problema de sincronización se conoce como problema de concurrencia. El problema de sincronización puede ocurrir en dos casos, en ambos, uno de los usuarios trata de modificar la base de datos, y el otro usuario o está tratando de modificar los datos o de acceder los mismos. Si ambos usuarios desean modificar los datos, se necesita un control para no perder la validez y la exactitud de los mismos. Si un usuario trata de acceder los datos mientras otro usuario los está modificando, puede obtener datos no válidos.

El problema de concurrencia ocurre porque la modificación de los datos se realiza en dos etapas: primero

se modifican los datos, y luego, ya modificados, se reescriben en la base de datos. Estos dos pasos son necesarios cuando la modificación se realiza a partir de datos ya existentes.

En general, una modificación envuelve dos tipos de variables: variables base y variables objeto. La variable objeto es la que se modifica, y la variable base es un valor a partir del cual se determina el nuevo valor de la variable objeto. Dependiendo de la complejidad de la operación se pueden utilizar varias variables base y varias variables objeto. Un punto importante es que no es suficiente inmovilizar únicamente a las variables objeto, sino que se deben inmovilizar también las variables base. Durante la modificación es posible acceder las variables base sin poner en peligro la modificación.

El intervalo de tiempo durante el cual puede ocurrir el problema de concurrencia se inicia en el momento de la lectura de las variables base y termina en el momento de la escritura de las variables objeto modificadas.

Para evitar el problema de concurrencia el primer usuario de los datos debe indicar que su acceso a los datos es para realizar una modificación; de esta manera, se pueden inmovilizar los datos mientras se realiza la modificación. Con este método, los datos no pueden ser accedidos por otro usuario durante la modificación; las modificaciones se

realizan de una manera serial y no concurrente.

El manejador de la base de datos debe realizar la inmovilización de los datos, porque si no se realiza de una manera correcta, se puede destruir la base de datos, además, esta inmovilización de los datos debe realizarse de manera transparente al usuario.

Se debe tomar en cuenta que los datos necesarios para realizar una modificación pueden encontrarse en diferentes récords y en diferentes archivos, haciendo más complejo el proceso de inmovilización. Existe la posibilidad de que en una modificación se afecte a más de un récord.

Por ejemplo, el usuario 1 desea modificar los récords A y B, en ese orden, y el usuario 2 desea modificar los récords B y A, en ese orden. Si ambos usuarios comienzan su trabajo al mismo tiempo, el usuario 1 modificará el récord A y el usuario 2 modificará el récord B; pero luego el usuario 1 no podrá modificar el récord B porque está inmovilizado por el usuario 2, y el usuario 2 no podrá modificar el récord A porque estará congelado por el usuario 1. En este caso se llega a un punto muerto (deadlock) porque ninguno de los dos usuarios puede terminar la modificación.

Existen dos formas de solucionar el problema anterior:

1. Se permite la ocurrencia del punto muerto y luego se recupera del mismo, dando prioridades a los usuarios. Por ejemplo, en el caso anterior, al llegar al punto muerto, si

el usuario 1 tiene mayor prioridad que el usuario 2 se llevará al usuario 2 a su estado antes de comenzar la modificación y se permitirá al usuario 1 terminar su modificación. Después que el usuario 1 realice su modificación, se permitirá al usuario 2 realizar la suya.

2. Se previene la ocurrencia del punto muerto; se pueden utilizar varios métodos:

a. Se puede preespecificar todos los datos que se utilizarán y únicamente se realizarán las modificaciones cuando se pueden inmovilizar todos los datos; este método presenta problemas cuando no se sabe de antemano los datos a utilizar.

b. Se debe presecuenciar el orden de acceso a los datos, por ejemplo, primero se debe acceder el récord A y luego el récord B, y no al revés.

c. Es similar a la primera solución, pero no permite la ocurrencia del punto muerto; por ejemplo, permite al usuario 1 realizar sus modificaciones, al pedir permiso el usuario 2 para realizar sus modificaciones, el sistema analiza si ocurriría un punto muerto. Si este es el caso y el usuario 2 tiene menor prioridad, éste debe esperar que el usuario 1 termine sus modificaciones. Si el usuario 2 tiene mayor prioridad, se llevará al usuario 1 a su estado antes de iniciar las modificaciones y el usuario 2 realizará sus modificaciones.

2. Sincronización en un ambiente distribuido. En el caso en que en un sistema distribuido se utiliza únicamente una distribución de datos, los métodos discutidos anteriormente pueden resultar adecuados; pero cuando se permiten múltiples copias de la base de datos o se utilizan sistemas híbridos, el problema de sincronización se complica aún más.

En un sistema distribuido con división de datos el problema de punto muerto (deadlock) resulta más complejo comparado con un sistema centralizado. En el sistema centralizado el manejador de la base de datos (DBMS) se da cuenta de la ocurrencia del punto muerto; pero en un sistema distribuido los usuarios pueden encontrarse en diferentes nodos, y los datos en otros nodos, por lo que desde el punto de vista de un nodo es sumamente difícil reconocer la existencia de un punto muerto y éste únicamente se reconoce al observar al sistema como un todo. Por lo anterior, en un sistema distribuido se prefiere prevenir la ocurrencia del punto muerto.

Cuando existen múltiples copias se necesita la sincronización para proveer la exactitud y validez dentro de la misma copia y para mantener la consistencia dentro de las diferentes copias. Idealmente, se deberían de obtener los mismos resultados sin importar desde donde se ingrese la transacción. Pero este caso únicamente se da con la

"inmovilización global" (global locking) y en casos especiales de la "copia dominante".

- Técnicas de sincronización con múltiples copias:

a. Inmovilización global:

Es una extensión lógica del método utilizado para el sistema centralizado; todas las copias del sistema son inmovilizadas y se realiza la modificación en todas y cada una de las copias, luego se liberan las copias. Con este método se logra una gran exactitud y consistencia. La principal dificultad es su desempeño; se utiliza en gran medida el sistema de comunicaciones y el tiempo necesario para realizar la sincronización puede llegar a ser prohibitivo.

En el caso más simple se necesita enviar al menos 5 mensajes:

- El nodo donde se pide el cambio (nodo inicial) pide inmovilizar los datos de los demás nodos.
- Cada nodo responde inmovilizando sus datos.
- El nodo inicial pide la modificación de los datos.
- Se realiza la modificación en cada nodo del sistema y luego se envía el mensaje de aceptación al nodo inicial.
- El nodo inicial envía un mensaje a los demás nodos para liberar la inmovilización de datos.

Estos 5 mensajes son los mínimos necesarios, en general, se utilizarán más mensajes. Cuando es imposible que todos

los nodos inmovilicen su información, se debe regresar al paso inicial hasta lograr la inmovilización en todas y cada una de las copias.

Los mensajes a transmitirse a través de la red pueden enviarse de manera serial o de manera paralela. Si se utiliza un método serial, el mensaje se transmitirá de manera secuencial de un nodo a otro: de A a B, luego de B a C, después de C a D, etc. hasta llegar a todos los nodos. El último nodo envía al nodo inicial el mensaje de confirmación de inmovilización de todas las copias.

Con el enfoque paralelo, el mensaje se envía simultáneamente a todos los nodos. Si se envía el mensaje de manera física a todos los nodos, el mensaje se transmite una vez y alcanza a todos los nodos destinatarios. Si se envía el mensaje de manera lógica, el DBMS instruye al software de comunicaciones para que envíe el mensaje a todos los nodos; el software de comunicaciones puede enviar el mensaje de manera paralela o serial.

La confiabilidad es una de las características más importantes en un método de sincronización de datos. Si uno de los enlaces a un nodo falla, el mensaje debe poder enviarse por una ruta alterna; pero si todos los enlaces a ese nodo, ó el mismo nodo, fallan, la copia de los datos en dicho nodo no se podrá acceder o modificar.

Existen dos métodos para reconocer la falla en un nodo:

Primero, el nodo puede reconocer la existencia de una falla y avisar al sistema. Cuando el nodo ha sido reparado se reconecta al sistema y se realizan los procesos de recuperación necesarios. La otra forma de reconocer una falla es estudiar el tiempo de respuesta; si un nodo no responde dentro de un intervalo de tiempo preestablecido, se considera como un nodo fallido.

Cuando un nodo falla, el sistema debe realizar las acciones necesarias para seguir operando a pesar de la falla. Estas acciones consisten en: regresar el sistema al estado anterior al momento de la ocurrencia de la falla; poner una bandera (flag) en dicho nodo para no seguir enviándole información y probar periódicamente si el nodo ya fue reparado. Al estar el nodo reparado, éste debe recuperarse; debe de regresar al estado anterior a la falla y poner al día todas sus bases de datos.

b- Sincronización por copia dominante:

Es un método simple utilizado cuando se tienen múltiples copias de los datos. Puede utilizarse en bases de datos replicadas o en configuraciones híbridas. Este sistema provee mejoras con respecto a la inmovilización global; pero es necesario que se cumplan las siguientes condiciones: Primero, que la frecuencias de modificaciones sea relativamente baja; segunda, que la base de datos pueda ser dividida y la mayoría de modificaciones se realicen con

referencia local, es decir, la mayoría de modificaciones de una parte específica de la base de datos se realizan en un nodo específico; y por último, la mayoría de los nodos que requieren acceso de los datos para los cuales no son el nodo dominante, no requieren la última versión de los datos.

La principal característica del método es el hecho de que no todas las copias de los datos son iguales. Una copia de los datos se denomina copia dominante y el nodo donde está esa copia se denomina nodo dominante. Cualquier modificación de los datos debe enviarse a este nodo, el cual controla la modificación de las demás copias de los datos para las cuales es la copia dominante. Eventualmente todas las copias de los datos serán modificadas, pero la primera copia a ser modificada es la copia dominante. Los accesos a los datos pueden realizarse en cualquier nodo que tenga una copia de los datos. Debe recordarse que si se realizan una gran cantidad de modificaciones, el nodo dominante puede convertirse en un cuello de botella.

La copia dominante de diferentes partes de la base de datos puede localizarse en diferentes nodos. Luego de determinarse los datos a modificar, debe determinarse el lugar de la copia dominante. En la inmovilización global, todas las copias de los datos debían de inmovilizarse antes de realizar la modificación; en este método, únicamente se inmoviliza la copia dominante. Una modificación puede entrar

en cualquier nodo, pero la modificación se realiza en la copia dominante. Debido a lo anterior, las modificaciones, que pueden provenir de diferentes nodos, deben realizarse de una manera serial.

Cuando se ha modificado la copia dominante, ésta controla la modificación de las demás copias del sistema. También se logra una mejora del tiempo de respuesta porque al modificar la copia dominante, se da indicación al usuario de la realización de la modificación. En la práctica, se debe de modificar también la copia del nodo de donde se pidió la modificación, por si el usuario realiza un acceso de los datos después de la modificación y no encuentra los datos modificados.

Si un usuario desea utilizar la última versión de los datos, o copia dominante, y no la copia local, puede pedirlo de manera explícita. A cambio de los datos actualizados, se tendrá un mayor tiempo de respuesta.

En el caso de falla en el nodo dominante es necesario tomar ciertas precauciones. Si se tiene un solo nodo dominante, puede ocurrir un colapso del sistema. Si se tiene un nodo dominante principal para una parte de los datos y un nodo dominante de reserva; al fallar el nodo principal, el nodo de reserva pasa a ser el nuevo nodo dominante. No es obligatorio tener un único nodo de reserva, se pueden tener varios nodos de reserva, mientras más nodos de reserva se

tengan, mayor será la confiabilidad del sistema. Pero todos los nodos de reserva deben tener la última versión de los datos, por lo que aumentan los costos de comunicación y los tiempos de respuesta. Cuando todos los nodos con copias puedan ser nodos de reserva, este método es sumamente similar al método anterior.

Al realizar una modificación, el nodo de reserva recibe aviso del nodo principal, si no recibe aviso, supone que el nodo principal falló, y envía una notificación a todos los nodos indicando que pasa a ser el nuevo nodo dominante. Si este nodo falla, su nodo de reserva lo pasa a sustituir, y así sucesivamente; mientras exista una copia de los datos, se puede procesar cualquier petición de modificación.

Si falla un nodo diferente al nodo dominante, no ocurre un gran problema en el sistema. Cuando el nodo dominante recibe una petición de modificación, ésta se llevará a cabo sin importar la falla de cualquier otro nodo del sistema. La falla de un nodo con copia local es irrelevante; si el nodo falla después de la modificación no hay problema; si el nodo falla antes de la modificación, ésta se realizará dentro de los procedimientos normales de recuperación luego de la reparación del nodo.

c. Sincronización por consenso de la mayoría:

Con este método se utilizan marcas de tiempo en lugar de inmovilización; todos los nodos pueden votar para aceptar

o no la modificación, y mientras la mayoría lo acepte (siempre y cuando ninguno lo rechace) se realizará el cambio. Este método tiene la ventaja con respecto a la inmovilización global en cuanto al uso de los sistemas de comunicación y tiempos de respuesta. Además, es posible realizar cambios desde cualquier nodo y se da la misma prioridad a todas las copias.

El costo que debe pagarse por esta mejora en el desempeño del sistema es la necesidad de una capacidad de almacenamiento adicional para las marcas de tiempo. Estas marcas de tiempo pueden aplicarse a nivel de récords, archivos, etc.

Los pasos necesarios para realizar una transacción son los siguientes:

- Acceder las variables base.
- Computar el nuevo valor de las variables objeto.
- Realizar la petición de modificación.
- Ejecutar el protocolo de sincronización.
- Realizar la modificación.
- Notificar al usuario de los resultados.

El primer paso consiste en el acceso de las variables base necesarias para calcular la modificación. Si la base de datos está totalmente replicada se realizará un acceso local, si no fuera el caso, el acceso será compuesto o remoto. Al acceder la variable también se accesa su marca de

tiempo, que indica la última vez que las variables fueron modificadas.

El segundo paso consiste en los cálculos necesarios para obtener los nuevos valores de las variables objeto. En el tercer paso se realiza la petición de la modificación. En el cuarto paso se realiza el proceso de votación correspondiente. El quinto paso se realiza si la modificación es aceptada. En el sexto paso se informa al usuario del resultado de su solicitud de modificación.

Al recibir un nodo una petición de modificación, puede votar en tres formas distintas: puede aceptar, rechazar o pasar. Decide como votar de acuerdo al valor de las marcas de tiempo de las variables base. Si las marcas de tiempo de las modificaciones son anteriores a las marcas de tiempo de su copia local rechazará la transacción ya que estará basada en información obsoleta. Si algún nodo rechaza la modificación, ésta será rechazada por el sistema; y para poder realizarla se debe regresar al primer paso. Un nodo aceptará la petición únicamente cuando sus marcas de tiempo locales son anteriores a las marcas de tiempo de la petición de modificación y si ninguna de las variables base eran variables objeto de alguna transacción anterior que se haya solicitado pero que aún no se haya confirmado.

Cuando se ha tomado una decisión definitiva respecto a una modificación, esta es transmitida a todos los nodos. Al

aceptarse una modificación el nodo toma dos acciones: Primero, rechaza una modificación pendiente que tenga conflicto con la modificación aceptada y avisa a los demás nodos; y segundo, realiza los procesos de modificación adecuados. El nodo debe comparar la marca de tiempo de la modificación con la marca de tiempo de su copia de las variables objeto. Si su marca de tiempo es anterior, realiza la modificación; si la marca de tiempo de la copia local es posterior, no realiza la modificación por ser obsoleta.

Las marcas de tiempo se utilizan en vez de las inmovilizaciones de datos, y sirven para cumplir dos propósitos fundamentales: Primero, las marcas de tiempo de las variables base se utilizan para que cada nodo pueda aceptar o rechazar una modificación; y segundo, cuando se ha aceptado una modificación, las marcas de tiempo de las variables objeto determinan si la modificación puede realizarse localmente o debe ignorarse.

Si un nodo falla, el sistema puede seguir funcionando con los nodos restantes, ya que el algoritmo permite la realización de modificaciones mientras la mayoría de los nodos acepten. En este caso la comunicación se realiza de un nodo a otro y debido a que un nodo necesita saber la votación de los nodos anteriores antes de votar, sólo puede enviarse el mensaje a un solo nodo, aumentando el tiempo de respuesta.

Para conocer si un nodo falló puede utilizarse el siguiente método: Se envía el mensaje, si no se obtiene respuesta en un tiempo dado, se realiza la transmisión varias veces; si el nodo falló, se enviará la transmisión a un nodo alternativo y se proseguirá buscando el consenso de la mayoría; se realizará o no la modificación, y se pondrá una bandera indicando el nodo defectuoso.

La transmisión serial de la información puede crear retardos prohibitivos, por lo que el procedimiento puede modificarse enviando desde el nodo solicitante mensajes simultáneos a todos los nodos y éstos votarán sin tomar en cuenta la votación de los demás nodos. El nodo solicitante recibirá todas las respuestas y determinará si se obtuvo la aceptación de la mayoría o si algún nodo rechazó la modificación.

Si los datos están dispersos entre un conjunto de nodos, debe existir un mecanismo para determinar la localización de un determinado archivo; se hace necesaria la existencia de un directorio, de una lista de los archivos y sus localizaciones. Se debe tomar en cuenta la diferencia entre la localización del directorio y la localización de los archivos determinados por dicho directorio.

Uno de los métodos utilizados consiste en un control centralizado. Se posee un único directorio central localizado en el nodo de control de datos. Al pedir un

usuario acceso a un archivo, esta petición es enviada al nodo de control de datos; éste decide si el usuario puede acceder la información, y si este es caso, proporciona las tablas de localización. El nodo de control de datos puede rechazar la petición por varias razones: el usuario no tiene la prioridad necesaria, se está utilizando dicha información, etc. Otra posibilidad consiste en quitar a este nodo de control la posibilidad de negar un archivo a un usuario; en este caso, se convierte en un procesador de servicio y la autorización del usuario se realiza en el nodo local.

Otra posible solución consiste en proporcionar a cada nodo con un directorio completo del sistema. Se obtiene un ahorro en las comunicaciones porque se hace referencia a un directorio local y si el directorio local no está disponible puede utilizarse el directorio de otro nodo. La desventaja consiste en la necesidad de mayor almacenamiento en cada nodo para mantener un directorio local y los procesos de sincronización necesarios para mantener todos los directorios iguales al realizar un cambio en los archivos.

Otro método consiste en colocar en cada nodo de aplicación un directorio para todos los datos locales en dichos nodos. Cada nodo tiene la posibilidad de preguntar a los demás nodos si posee o no un determinado archivo. La forma de realizar dicha pregunta depende de la naturaleza de

la interconexión y de la topología de la red. Por ejemplo, en una red mallada se puede realizar la petición a todos los nodos simultáneamente, en una red en anillo se puede realizar la petición una vez y ésta se transmite de manera serial de un nodo a otro.

VII. DISTRIBUCION DE FUNCIONES

En este capítulo se discute la distribución de funciones en un sistema distribuido. Primeramente se presentan las diferentes etapas que se realizan al recibir una transacción de un usuario y luego se presentan varias formas en las cuales estas funciones pueden ser agrupadas.

A. Funciones necesarias para realizar una transacción

Las funciones necesarias para realizar una transacción en un sistema distribuido son:

1. Aceptar la petición del usuario.
2. Chequear la validez de la petición.
3. Chequear la autorización del usuario.
4. Traducir la petición.
5. Determinar la estrategia para el procedimiento a seguir por la red.
6. Codificar la petición.
7. Determinar el enrutamiento.
8. Transmitir la petición a través de la red de comunicaciones.
9. Decodificar la petición.
10. Determinar la estrategia local de procesamiento.
11. Localizar los datos en la estructura lógica.
12. Localizar los datos en la estructura física.

13. Realizar la entrada y salida física.

14. Convertir los datos a un formato adecuado para el usuario.

15. Traducir los datos.

16. Retornar la respuesta al nodo originario.

17. Consolidar, ordenar y editar la respuesta.

18. Presentar la respuesta al usuario.

1. Aceptar la petición del usuario:

En este punto el NDBMS (Network Database Management System) acepta la petición de usuario; ésta puede ser para datos locales o para datos localizados en otro nodo; puede desearse únicamente un acceso de datos o puede pedirse una modificación.

2. Chequear la validez de la petición:

Con esta función se edita la petición y se chequea para detectar posibles errores; se realiza un primer chequeo para examinar la sintaxis y el formato de la petición; en otros chequeos se estudia la existencia de los diferentes campos y récords. Para realizar el chequeo anterior se necesita tener un esquema de la arquitectura de la base de datos; si ésta no se posee localmente, este chequeo se realizará hasta después de enviar la petición a otro nodo.

3. Chequear la autorización del usuario:

El siguiente paso es chequear si el usuario tiene la autorización para acceder los datos y para realizar los

procesos deseados. El chequeo del usuario puede realizarse en el nodo local o en el nodo donde se localizan los datos y los procesos. Como una regla general, el chequeo y el control de acceso deben realizarse lo antes posible. En sistemas donde la seguridad es extremadamente importante, este chequeo puede realizarse varias veces: en el nodo local y en los nodos remotos.

4. Traducir la petición:

Esta función se utiliza únicamente en sistemas heterogéneos, para traducir una petición de un tipo de DBMS (Database Management System) a otro. Idealmente esta traducción debe resultar transparente al usuario.

5. Determinar la estrategia para el procedimiento a seguir por la red:

Probablemente es la función más compleja y su complejidad depende de la flexibilidad del NDBMS. Si únicamente se permiten referencias locales y remotas, entonces la función es relativamente simple, lo único que debe hacerse es determinar el nodo al cual debe enviarse la petición. Cuando se realizan peticiones compuestas, locales y remotas simultáneamente, el problema se complica y debe solucionarse en dos etapas: primero, se debe decidir como procesar y coordinar la petición, segundo, debe determinarse a cual nodo enviar la petición. Debe tomarse en cuenta que el nodo originario probablemente descompuso la petición en

varias sub-peticiones y se debió de buscar un destino a cada una de ellas. Se puede utilizar un directorio de datos de la red para determinar a cada uno de los nodos destinatarios. Cuando existe replicación de datos y/o funciones pueden tenerse varias posibilidades y la elección de los nodos destinatarios puede ser de acuerdo a una tabla donde se tengan los diferentes nodos ordenados por prioridades o de acuerdo al estado del sistema de comunicaciones (tráfico, tiempo de respuesta, etc.).

6. Codificar la petición:

Esta función puede o no usarse, dependiendo de la seguridad del sistema. Si se usa, el mensaje se debe codificar antes de enviarse al sistema de comunicaciones. El procedimiento utilizado es el siguiente: Se codifica el mensaje, se transmite y luego se decodifica en el nodo destinatario, ofreciendo seguridad únicamente en la etapa de comunicación.

7. Determinar el enrutamiento:

Determina la ruta a seguir por la petición al ser enviada. Este camino puede ser directo entre el nodo fuente y destino o puede pasar por nodos intermedios. Al existir nodos intermedios se tienen dos alternativas: se decide al inicio el camino a seguir, o los nodos intermedios deciden el enrutamiento de una manera estática o dinámica.

8. Transmitir la petición a través de la red de

comunicaciones:

El mensaje se envía a través del sistema de comunicaciones: pasa a través de las diferentes capas de comunicaciones en el nodo fuente, se transmite al destinatario y luego pasa a través de las capas de comunicaciones de dicho nodo.

9. Decodificar la petición:

Si el mensaje se codificó en el nodo fuente, es necesario realizar la decodificación en el nodo destino.

10. Determinar la estrategia local de procesamiento:

En este punto la petición se encuentra en el nodo que posee los datos; si se desea realizar una modificación, se realiza la sincronización de las diferentes copias. El DBMS local trabaja la petición como si fuera una petición local.

11. Localizar los datos en la estructura lógica:

Realiza el mapeo entre la forma que el usuario mira los datos y la estructura lógica de la parte local de la base de datos.

12. Localizar los datos en la estructura física:

Realiza el mapeo entre la estructura lógica y la estructura física de la base de datos local.

13. Realiza la entrada y salida física:

Los datos son leídos o escritos; en la lectura de datos, éstos se leen en un buffer; en una modificación, el procedimiento se complica y depende del método de

sincronización utilizado.

14. Convertir los datos a un formato adecuado para el usuario:

Esta función requiere los datos en su forma almacenada, la parte del esquema que los describe y la parte del esquema del usuario que indica el formato deseado.

15. Traducir los datos:

Se realiza en sistemas heterogéneos cuando se utiliza diferente hardware, software, o ambos. Esta función puede realizarse en el nodo fuente, destino o en un nodo intermedio.

16. Retornar la respuesta al nodo originario:

La respuesta, codificada o no, pasa por las capas de comunicación del nodo destinatario, se envía y se realiza la decodificación cuando es necesaria.

17. Consolidar, ordenar y editar la respuesta:

Prepara la respuesta a ser dada al usuario. En una petición compuesta, se obtienen varias respuesta que deben consolidarse para obtener una sola respuesta. Generalmente se realiza en el nodo local, porque disminuye los costos de comunicación.

18. Presentar la respuesta al usuario:

Se presenta la respuesta al usuario; si se realizó una modificación, se presenta el estado de la operación.

B. Configuración de los nodos

Existe una gran cantidad de posibilidades para poder distribuir las funciones; pero sólo unas pocas son realmente prácticas. Entre las posibles combinaciones se tienen:

1. Nodo mínimo de usuario (Minimal user node)
2. Nodo mínimo de datos (Minimal data node)
3. Nodo mínimo de directorio de datos de la red (Minimal network data directory node)
4. Nodo mínimo descriptor de la red (Minimal network description node)
5. Nodo de base de datos (Database computer node)
6. Nodo completo (Complete node)

Los nodos mínimos poseen únicamente software de comunicaciones y una de las siguientes características: interface de usuario, una parte de la base de datos y software del sistema de manejo de la base de datos (database management system DBMS), directorio de la red o descripción de la red. El nodo mínimo de usuario contiene la interface del usuario y un mínimo de software de comunicaciones, no posee funciones de manejo de datos. El nodo mínimo de datos sólo posee datos y funciones para accederlos, no posee interfaces de usuario. El nodo mínimo de directorio de datos de la red no posee interface de usuario ni datos, únicamente determina hacia que nodo deben enrutarse las peticiones; la función de enrutamiento de mensajes puede realizarse en un

nodo, creándose un nodo mínimo descriptor de la red.

A partir de estos nodos mínimos se puede crear cualquier tipo de sistema distribuido; el diseñador puede crear el tipo de nodo necesario al combinar los diferentes tipos de nodos mínimos. Por ejemplo, el nodo de base de datos tiene toda la base de datos y las funciones del DBMS, y al existir varios nodo de base de datos, estos tendrán una réplica de la base de datos. El nodo completo posee las características de todos los nodos mínimos.

1. Nodo mínimo de usuario

Provee únicamente la interface del usuario, no posee datos o funciones para manejar datos; todas las peticiones deben enviarse hacia otros nodos.

Este nodo realiza 4 funciones básicas: 2 con el usuario y 2 con el software de comunicaciones. Primero debe aceptar las peticiones del usuario y segundo debe poder presentar mensajes al usuario; no se realiza ningún tipo de conversión de datos o de formateo de la información.

Todos los nodos deben poseer cierto software de comunicación para enlazarse con el resto de la red; el nodo mínimo de usuario posee una interface para enviar o recibir mensajes del sistema; éstas son las dos funciones con el software de comunicaciones. Debido a que el nodo mínimo de usuario no posee un directorio o descripción de la red, no puede escoger al nodo destinatario ni enrutar mensajes; por

lo tanto, siempre envía los mensajes a un mismo nodo, actuando como una "terminal tonta".

Al incluir funciones adicionales en este nodo, puede mejorarse su desempeño: al incluir una descripción de la red puede realizar enrutamiento y si se posee localmente una parte de la base de datos, puede realizarse localmente cierta validación y chequeo de los datos ingresados.

2. Nodo mínimo de datos

Incluye parte de la base de datos, parte del esquema que lo describe y ciertas funciones del DBMS. Estas funciones deben ser suficientes para acceder los datos y el esquema de la base de datos; si se desea la capacidad de realizar modificaciones en los datos, deben agregarse otras funciones.

Todas las peticiones de acceso a datos deben referirse a los aspectos físicos del esquema porque es la única parte almacenada localmente; la petición, al llegar a este nodo, debe indicar la posición física de los datos.

En principio, existen dos tipos de nodos mínimos de datos, uno que permite únicamente acceso y otro que permite modificaciones; el que permite modificaciones es más complejo porque se necesita inmovilizar los datos y realizar procesos de sincronización. El nodo mínimo de datos debe realizar 4 funciones: la parte local de la DBMS debe aceptar las peticiones del software de comunicaciones; debe

responder a las peticiones que le llegan; debe utilizar el esquema local para localizar los datos y debe poder realizar cierta entrada/salida física.

3. Nodo mínimo de directorio de datos de la red

Es un tipo de nodo que no posee interface de usuario y tampoco posee parte de la base de datos. Debido a que posee una copia del directorio de la red, puede tomar una petición y determinar el nodo que puede satisfacerla. En su forma más simple únicamente puede responder a peticiones remotas. Este tipo de nodo puede identificar a los nodos; pero no puede realizar el enrutamiento porque no posee una descripción de la red; en este caso, la petición y la identificación del nodo que puede contestarla deben enviarse a otro nodo para realizar el enrutamiento. Una mejora obvia para este nodo consiste en integrar una descripción de la red en este nodo para poder realizar el enrutamiento; otra mejora consiste en incluir funciones adicionales para poder descomponer una petición compleja (una petición que necesita respuesta de varios nodos).

4. Nodo mínimo descriptor de la red

Provee únicamente las funciones de enrutamiento; toma el identificador del nodo, determina el enrutamiento y envía el mensaje.

5. Nodo de base de datos

Es un tipo especializado de nodo. En un ambiente

centralizado la computadora de la base de datos (DBC) tiende a formarse a partir de un anfitrión (host) de propósito general que ejecuta las aplicaciones, y otro computador que realiza las funciones de manejo de la base de datos. Este DBC puede implementarse utilizando hardware de propósito general o hardware especializado. En un sistema centralizado, el DBC controla la base de datos, sus definiciones (esquemas) y las funciones del DBMS.

El DBC puede poseer funciones desde un controlador de disco que busca y accesa los datos hasta un computador que acepta peticiones del host, realiza el procesamiento y devuelve una respuesta.

Al utilizar un DBC en un sistema distribuido aparecen dos decisiones importantes: el lugar de las funciones y el lugar de localización de los datos.

En un sistema distribuido se pueden localizar las funciones del DBMS local en el DBC: 1) Determinar la estrategia de procesamiento local. 2) Localizar los datos en la estructura lógica. 3) Localizar los datos en la estructura física. 4) Realizar la entrada/salida física. 5) Convertir los datos a un formato deseado por el usuario.

En el sistema distribuido se debe decidir donde colocar el NDBMS (sistema de manejo de la base de datos de la red, network database management system) y las funciones de comunicaciones, además de la conexión del DBC a la red. Si

se utiliza un DBC con hardware especial, se diseña para optimizar el paralelismo inherente de la base de datos, por lo que las funciones del NDBMS y de comunicaciones no deben colocarse en el DBC, al no hacer uso adecuado del paralelismo. En este caso se interconectaría como un DBC de propósito especial a un host de propósito general donde se pondrían las funciones del NDBMS y de comunicaciones; la interconexión a la red se realizaría a través del host.

Si se utiliza un DBC de arquitectura de propósito general, se pueden localizar las funciones del NDBMS y de comunicaciones en el mismo, y no se necesitaría un host para la interconexión a la red.

En cuanto a la localización de los datos, la solución más sencilla consiste en tener un solo DBC con toda la base de datos; pero se limita el tamaño de la base de datos y se tienen problemas de confiabilidad. Si se tienen varios DBC con la base de datos replicada, éstos se convierten en un nodo más del sistema y se tienen todos los problemas de sincronización que ocurren en los sistemas con bases de datos replicadas. El mayor beneficio del DBC es que permite una solución centralizada al problema de la base de datos.

6. Nodo completo

Posee todas las funciones enumeradas al inicio del capítulo; posee interface de usuario, una parte de la base de datos con definición de los datos que posee, además de



funciones del NDBMS y del DBMS local, directorio de los datos de la red y descripción de la red.

Un nodo completo posee todas las funciones del DBMS; pero sólo para la aplicación específica donde se encuentra. Por ejemplo, en un sistema donde sólo existen datos divididos, el Nodo completo no necesita poseer funciones de sincronización.

VIII. DISTRIBUCION DE CONTROL

El control consiste en tres elementos básicos: los datos necesarios para tomar las decisiones, los algoritmos necesarios para manipular los datos, y el acceso a los datos y algoritmos.

En los sistemas centralizados, los datos y los algoritmos se localizan en un solo punto y únicamente un nodo del sistema tiene acceso a ellos. En un sistema multinodal puede tenerse una relación maestro-esclavo entre el nodo que controla los datos y los demás nodos.

En los sistemas distribuidos se necesitan funciones de control adicionales para poder controlar a los nodos esclavos; entre estas funciones se encuentran: operadores remotos, determinación remota de problemas, reconfiguración remota, etc. En un sistema con control centralizado todas las funciones pueden colocarse en un nodo central y éste puede generar o demandar los datos de los demás nodos, y realizar las funciones necesarias para presentar al sistema como una única entidad.

El control puede distribuirse permitiendo a más de un nodo el acceso a los datos y a los algoritmos localizados en el punto central; este concepto es común en los sistemas de multiprocesadores. Por ejemplo, se puede tener un solo

sistema operativo, teniendo éste el control de los datos requeridos por el sistema y los algoritmos necesarios para manipular los datos y se permite a cualquier procesador del sistema el acceso a la estructura de control ya que las posiciones de memoria donde se localiza el sistema operativo se pueden estructurar haciéndose accesible a todos los procesadores. El sistema anterior puede centralizarse al permitir a un solo procesador realizar las funciones de control del sistema operativo: este procesador realizaría las funciones de control para sí mismo y para los demás procesadores del sistema; los nodos enviaría su petición de acciones de control al procesador de control designado. Si el procesador de control falla, otro procesador puede tomar las funciones de control. En este diseño se mantiene un control centralizado pero con posibilidades de recuperación al existir una falla.

Con el enfoque anterior se permite a cualquier procesador ejecutar funciones del sistema operativo cuando desea realizar funciones de control. Dependiendo de la estructura del programa del sistema operativo es posible que varios procesadores estén ejecutando simultáneamente varias porciones del programa de control. El sistema aún puede considerarse como centralizado porque hay un solo lugar en la memoria donde están almacenados los datos y los programas.

Se puede distribuir más el control al dividirlo en varias funciones discretas: manejo de memoria, manejo de programas y manejo de entrada/salida; esto se logra al estructurar los programas del sistema operativo en un conjunto de servicios especializados. Al estructurar el sistema operativo se puede permitir a todos los procesadores el acceso al mismo; o se pueden diseñar ciertos procesadores para controlar determinadas funciones de control al permitirles acceso a posiciones de memoria totalmente definidas.

La figura 8.1 presenta una serie de procesadores de aplicación combinados con un conjunto de procesadores de control; se tiene un manejador global del sistema, un procesador de manejo de mensajes y un procesador de manejo de datos. Cada procesador de control posee los datos y los programas necesarios para realizar su función a petición de los procesadores de aplicación; los datos en cada procesador de control son globales.

Por ejemplo, un procesador de aplicación puede solicitar la ejecución de un programa en otro procesador de aplicación; este procesador envía un mensaje por medio del procesador de manejo de mensajes al procesador de manejo global de la red; este mensaje pedirá el inicio del programa indicando las características del procesador donde debe ejecutarse. El manejador global conoce las características

de todos los procesadores de aplicación y los programas del sistema, y selecciona el procesador adecuado.

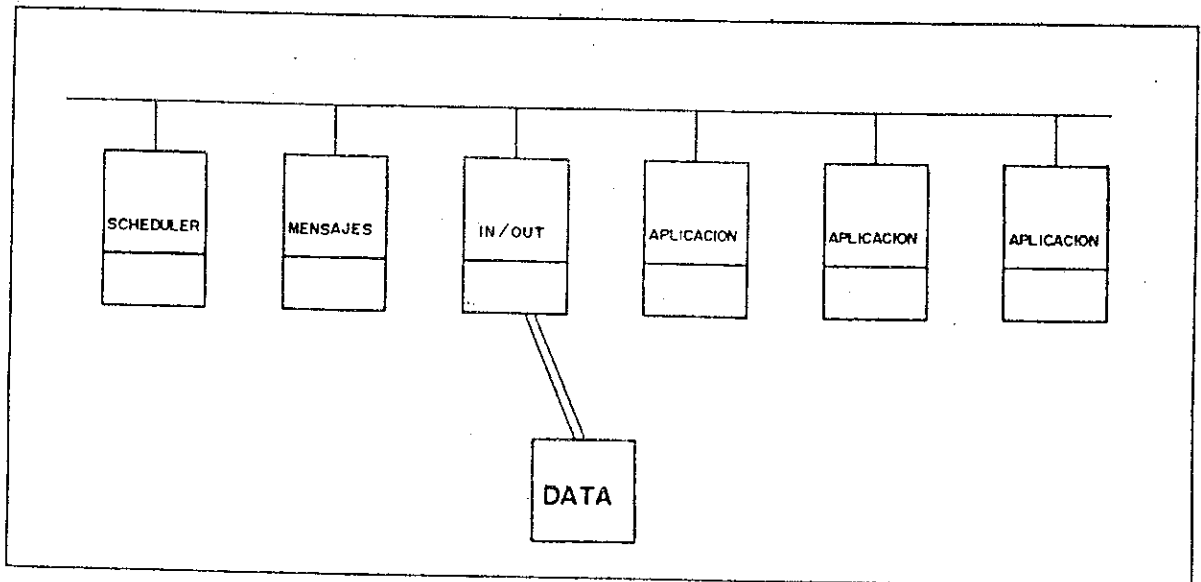


FIGURA 8.1

PROCESADORES DE FUNCIONES GLOBALES

Es necesario que cada procesador local posea ciertas funciones de control, por ejemplo, compartición de tiempo.

Al continuar distribuyendo el control se llega a un punto en que todos los datos y algoritmos de control se dispersan a través de todo el sistema; no existiendo "puntos globales", cada nodo posee un sistema operativo completo responsable del control local y de la interacción con los demás nodos.

En la fig 8.2 se muestra un sistema con control distribuido, se debe notar el procesador de interacción relacionado al procesador de aplicación. Este procesador de interacción es el encargado de toda la comunicación con los demás nodos, permitiendo al procesador de aplicación

realizar otras tareas.

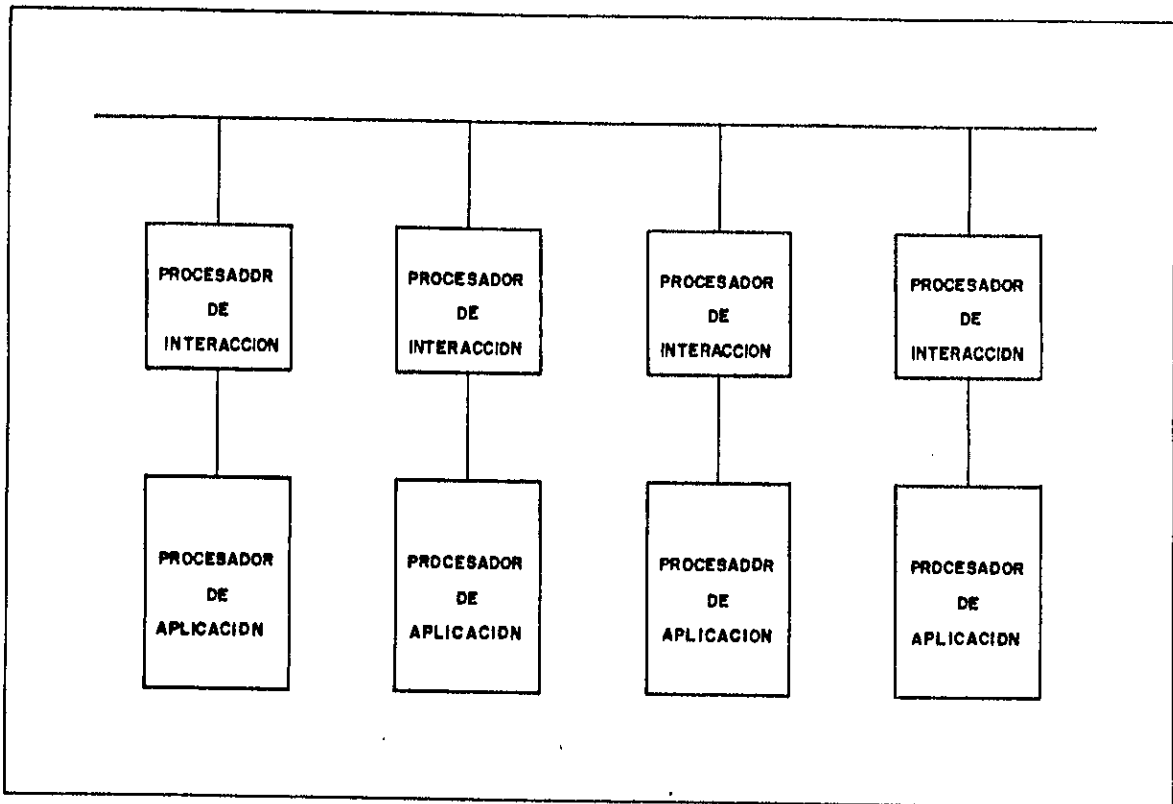


FIGURA 8.2
SISTEMA CON CONTROL DISTRIBUIDO

En este sistema se puede procesar una petición de la siguiente manera: un nodo desea ejecutar un programa que reside en otro nodo; el procesador de aplicación envía un mensaje a su procesador de interacción especificando el programa a ejecutar y las condiciones que debe cumplir el procesador a ejecutarlo. El procesador de interacción envía el mensaje a todos los nodos del sistema y los procesadores de interacción de dichos nodos, conociendo el estado y las condiciones de su procesador de aplicación, responden el mensaje del procesador de interacción fuente. Este

procesador decide cual es la mejor opción para ejecutar el programa. El uso del canal de comunicación y los mecanismos de interconexión pueden reducirse al suprimir las notificaciones negativas: un nodo que no desea realizar un trabajo no responde, y si un nodo no recibe dentro de un tiempo definido la autorización para ejecutar un programa, asume que no obtuvo el trabajo.

IX. LINEAMIENTOS PARA EL DISEÑO DE UNA RED DISTRIBUIDA DE DATOS.

En este capítulo se presentan ciertos lineamientos generales para el diseño de una red distribuida de datos; se indican los pasos a seguir, desde el análisis de las necesidades del usuario hasta el cronograma de implementación del sistema.

Debe tomarse en cuenta la generalidad del presente capítulo, ya que se trata de cubrir la mayor cantidad de posibilidades y dependiendo del tipo de sistema, puede no ser necesario estudiar en su totalidad cada uno de los siguientes puntos.

Los pasos a seguir en el diseño de una red distribuida de datos son los siguientes:

- A. Análisis de las necesidades del usuario.
- B. Diseño del sistema.
- C. Análisis de la red.
- D. Determinación de los requerimientos del sistema.
- E. Comparación y elección entre las diferentes opciones de acuerdo a los criterios de selección.
- F. Cronograma de implementación.

A. Análisis de las necesidades del usuario

Se deben de estudiar todas las funciones y los trabajos

que realiza el usuario en cada uno de los diferentes niveles:

- Operacionales: realizados día a día por el personal de bajo nivel (ingreso de datos).

- Tácticos: realizados por personal de nivel medio y concierne al planeamiento a corto término y a las funciones administrativas (análisis de datos).

- Estratégicos: tratan acerca de la administración y el planeamiento de la organización.

Se deben analizar dichas funciones y determinar cuales pueden distribuirse. La decisión de lo que se puede distribuir depende de cómo se realizan las funciones, la capacidad del usuario, la cantidad de información y aspectos técnicos y económicos.

Para la escogencia del sistema de procesamiento de datos a utilizar deben de tomarse en cuenta factores organizacionales, de aplicación y del sistema.

Entre los factores organizacionales se encuentran:

- La tendencia a la centralización o descentralización de la organización.

- La distribución de las habilidades computacionales dentro de la organización.

Los factores de aplicación incluyen:

- La relación entre los datos de una aplicación y la estructura de la empresa.

- La estructura lógica de la aplicación y como puede hacerse tender a un tipo de sistema distribuido.

- El futuro de la aplicación; si tiende a ser modificada o a permanecer estable.

Los factores del sistema incluyen:

- Relación precio/desempeño del sistema computacional.

- Relación precio/desempeño de los mecanismos de interconexión.

- Costos de programación.

- Costos de operación.

Para conocer las necesidades del usuario deben seguirse los siguientes pasos:

1. Recolección de la información.

2. Análisis de la información.

1. Recolección de la información. Se deben investigar las funciones que se realizan, que sistema poseen y la estructura de la organización. La recolección de información puede realizarse por medio de entrevistas o estudiando documentación escrita.

Entre las fuentes a investigar se pueden mencionar:

- Los diagramas de organización y la descripción de puestos.

- Procedimientos del usuario: cómo se ingresa la información, cómo se edita, cómo se preparan los reportes.

- Ejemplos de reportes: diarios, semanales, mensuales.

etc.

- Formas: órdenes, entrada de datos, etc.

Entre las preguntas a realizar se tienen:

¿Cuál es la situación actual?

¿Quién está envuelto en el problema?

¿Tiempo máximo para presentar una solución?

¿Qué equipo está disponible?

¿Cómo fluye la información dentro de la organización?

¿Qué información está disponible acerca del sistema?

Tipo, volumen y frecuencia de ingreso de los datos y de salida de reportes.

¿Quiénes son los usuarios de los recursos computacionales?

¿Dónde están localizados físicamente los usuarios?

¿Cuáles son los posibles nuevos usuarios?

¿Cuáles son las bases de datos existentes?

¿Cuáles son los costos de operación?

2. Análisis de la información. Se debe de analizar:

- Usuarios del sistema:

Número simultáneo de usuarios (promedio, pico, etc.)

Localización de los usuarios.

Máximo número de usuarios en una aplicación.

Nivel y experiencia de los usuarios.

- Funciones a realizar:

Flujo de información, entrada, salida (reportes,

archivos), etc.

Volumen y distribución geográfica de los datos.

Recursos necesarios para una aplicación (lenguaje usado, hardware, etc).

Tendencia a cambio en una aplicación.

Relación entre aplicaciones (compartición de archivos, operación simultánea, etc.).

Funciones a soportar por las terminales.

Cantidad y calidad de impresión.

- Intercambio de información:

Mensajes entre diferentes departamentos de la organización.

Volúmenes de intercambio por período de tiempo.

Determinación de horas pico.

- Operación del sistema:

Número de personas y cualidades para manejar el sistema.

Forma de recolectar la información.

Cómo se controlan y distribuyen los reportes.

Cómo se realizará la comunicación con otros sistemas.

- Costos:

Recursos económicos con que se cuenta.

Costo máximo permisible.

Costos directos e indirectos.

Costos fijos y costos variables.

- Desempeño del sistema:

Confiabilidad, tasa de errores (BER).

- Tiempo de acceso.
- Tiempo de respuesta.
- Volumen de transacciones.
- Disponibilidad.
- Lenguajes.
- Futuras necesidades de ampliación de la red.
- Ambiente del sistema:
- Localización.
- Aire acondicionado.
- Acceso físico.
- Piso falso.
- Protección contra fuego, terremoto, etc.
- Análisis de las restricciones:
- Técnicas (disponibilidad de equipo o tecnología necesaria).
- Jurídicas.
- Presupuestarias (limitación de presupuesto).
- Plazo de realización.
- Personal disponible.
- Equipo de procesamiento de datos o teleproceso ya existente.
- Adaptabilidad de la empresa.

B. Diseño del sistema

La distribución es el resultado de un diseño en el cual:

- Se dividen las aplicaciones: sistemas y aplicaciones

deben descubrirse y agruparse de manera que formen unidades separadas bien definidas.

- Se dividen los datos: debe determinarse los caminos de acceso a los datos, dónde es útil la división, dónde es útil la replicación y determinar la técnica de sincronización.

- Se identifican relaciones entre aplicaciones y datos: las divisiones de aplicaciones y datos deben definirse de manera conjunta ya que se interrelacionan.

- Se reconocen relaciones entre divisiones: se determina el grado de autonomía entre programas y la intensidad de su interacción e interdependencia.

- Se define una base de hardware potencial: a partir de la estructura lógica, se debe escoger la estructura de hardware necesaria.

Para el diseño de la arquitectura del sistema, deben tomarse en cuenta:

- Puntos de acceso al sistema.
- Control de procesamiento y almacenamiento de la información.

- Flujo de información.

- Manejo y mantenimiento del sistema.

- Distribución y terminales en cada nodo.

El diseño del sistema puede dividirse en dos etapas:

- Lógico o funcional: incluye archivos, bases de datos,

entrada/salida de reportes, etc.

- Físico: cuyo propósito consiste en trasladar el diseño lógico a la arquitectura de hardware y comunicaciones.

En el diseño lógico del sistema se indica el flujo general del sistema indicando sus entradas, reportes, procesamiento, etc.

1. Diseño lógico. El diseño lógico del sistema puede dividirse en:

- Salidas del sistema:

Se deben tomar en cuenta todas las diferentes salidas del sistema hacia el usuario, tanto en impresión como en pantallas. Se debe analizar las salidas periódicas como las no periódicas. En las salidas periódicas podemos tener reportes, los cuales pueden ser distribuidos a usuarios localizados en diferentes puntos del sistema, siendo necesario mantener consistencia entre los datos. Entre los puntos a estudiar en los reportes tenemos: nombre, a quienes va dirigido, frecuencia, datos que lo forman, etc. También deben analizarse las pantallas, indicando su disposición, arquitectura, nombre, datos que lo forman, etc.

- Entradas del sistema:

En un sistema distribuido debe tenerse cuidado en definir quien puede proporcionar datos y en que momento. Para los datos de entrada debe considerarse: tipo de dispositivo de entrada, fuente de los datos, formato de la

pantalla de entrada, datos de entrada, volumen y frecuencia de los datos de entrada, etc.

- Bases de datos/diseño de archivos:

Este es uno de los aspectos más importantes, además de determinar las características de los archivos (nombre, elementos que lo forman, frecuencia de modificación, volumen, etc) es necesario estudiar su localización y si se utiliza redundancia. La base de datos en un sistema distribuido puede ser dividida o replicada y aun diferentes partes de la base de datos pueden tratarse con diferentes criterios: una parte centralizada, otra parte dividida y otra parte replicada. La decisión depende de la localización de los usuarios de dichos datos y de la seguridad deseada. Además se necesita de un directorio de todos los datos del sistema.

- Interface con otros sistemas:

Se necesita identificar la siguiente información: a qué sistema se interconectará, elementos a intercambiar, frecuencia de intercambio, etc.

2. Diseño físico. Consiste en la unión del diseño lógico con el hardware; se deben dimensionar las memorias dependiendo de la base de datos y archivos a usar; el número y tipo de terminales dependiendo del número y tipo de usuarios; el tipo y número de impresores, lectoras de tarjetas, etc.

C. Análisis de la red

En un sistema distribuido, la comunicación puede realizarse entre nodos localizados físicamente cerca o entre nodos separados geográficamente.

Quando los nodos están separados, se debe tomar en cuenta la comunicación entre ellos, ya que se necesitará utilizar técnicas de teleproceso por medio de líneas discadas o arrendadas o utilizar la red de conmutación de paquetes.

En el análisis de la red, debemos de determinar la siguiente información:

- Datos a enviarse
- Realizar un listado de cada tipo de transacción, su tamaño y su destino.
- Cada transacción debe definirse indicando los elementos que lo forman, su longitud y el número máximo de datos a enviar en las horas pico.
- Al tener las características de cada transacción y su volumen, se debe realizar un diagrama de la red, identificando cada uno de los enlaces e indicando para cada enlace el número promedio y máximo de transacciones y caracteres por transacción.
- Posteriormente se deben de agregar todas las transacciones para estimar la carga total por enlace en toda la red. De la información anterior se puede obtener la

velocidad en bps para cada enlace de la red. Debido a la existencia de errores de transmisión debe agregarse cierto porcentaje para retransmisión; otro porcentaje para encabezados, control, sincronización y otros factores y además otro porcentaje para permitir un crecimiento del sistema.

La teoría de colas puede utilizarse para estimar la capacidad de las líneas de transmisión cuando la complejidad de la red lo amerite.

Se deben presentar varias alternativas, dependiendo de diferentes topologías de la red, diferentes vendedores y servicios (línea dedicada, línea conmutada, enlace a la red de conmutación de paquetes, etc.).

D. Determinación de los requerimientos del sistema

El siguiente paso a seguir consiste en el estudio de las necesidades de hardware, software y comunicaciones.

A partir de los requerimientos del usuario se tiene un diseño conceptual del sistema, y basados en éste se debe seleccionar el hardware y software necesario.

Para cada ítem de hardware se debe conocer donde estará localizado, que se supone que hará, cómo se interconecta a la red y su crecimiento potencial. Para el software se necesita saber que hace y que soporte necesita.

A continuación se presenta un listado general de los requerimientos del sistema:

- Descripción general de la red:
 - Configuración básica.
 - Posible expansión.
- Selección del hardware:
 - Tipo de máquina, cpu, velocidad.
 - Capacidad de memoria.
 - Discos:
 - Tipo, número de drives, capacidad, expansión.
 - Controlador de disco:
 - Tipo, número de drives que soporta, capacidad.
 - Cinta magnética:
 - Tipo, número por nodo, número de drives que soporta.
 - Impresoras:
 - Tipo, modelo, número por nodo, tipo de impresión (láser, impacto, etc.), soporte de funciones de usuario (formatos especiales, papel carbón, etc.)
 - Terminales CRT:
 - Tipo, modelo, número por nodo, funciones que soporta (edición, teclas con funciones especiales, etc.), tamaño del monitor, tipo de teclado, velocidad, protocolo de comunicaciones, etc.
 - Lectora de tarjetas:
 - Tipo, modelo, velocidad, número por nodo.
 - Otro tipo de equipo de hardware.

- Software del sistema:
 - Sistema operativo (versión, tipo, características) y compatibilidad entre sistemas operativos de cada nodo.
 - Software de la red.
 - Protocolos de telecomunicaciones.
 - Monitor de telecomunicaciones.
 - Método de acceso de telecomunicaciones.
 - Método de acceso a datos.
 - Compiladores.
 - Sistema de manejo de la base de datos (DBMS).
 - Software de manejo de la red.
 - Ayudas para desarrollo de software (generadores de reportes, generadores de datos de prueba, ayuda de debugging, macroassembler).
 - Lenguajes de alto nivel.
 - Software de telecomunicaciones.
- Comunicación de datos:
 - Front end y back end processors
Versión, número por nodo, capacidad, software del sistema residente.
 - Puertos de comunicación:
Tipo, número por nodo.
 - Adaptadores de comunicación.
 - Modems:
Tipo, velocidad, número por nodo.

- Multiplexores.
- Concentradores.
- Líneas de transmisión:
 - Tipo (discada, arrendada), simplex/dúplex, capacidad.
- Protocolos de comunicación.
- Software de aplicación:
 - Distribución de funciones por sitio.
 - Estructura general del software.
- Requerimientos de energía.

E. Comparación y elección entre las diferentes opciones de acuerdo a los criterios de selección

Para el estudio de las diferentes soluciones posibles es necesario contar con la siguiente información de cada propuesta:

- Lista completa de hardware, software del sistema y software de aplicación.
- Precios y términos de pago.
- Tiempo de entrega.
- Lista de usuarios actuales de sistemas similares, para contactarlos.
- Previsiones para crecimiento y expansión.
- Desempeño estimado.
- Correlación entre las necesidades del usuario y el sistema propuesto.

El proceso de evaluación consiste en los siguientes pasos:

- Realizar una evaluación inicial, eliminando vendedores y dejando finalistas.

- Realizar un análisis más profundo de las propuestas finalistas.

- Realizar el contrato con el ganador.

- Evaluación inicial:

Para eliminar las propuestas que claramente no cumplen con los requisitos. Algunas razones podrían ser: no posee las funciones pedidas, no cumple un número suficiente de requisitos del usuario, la situación financiera de la compañía vendedora es cuestionable.

- Análisis profundo:

Las propuestas finalistas deben cumplir los requisitos básicos del usuario y se debe realizar un análisis más completo de cada una de ellas, aun incluyendo visitas a fábrica si fuera necesario. Para la evaluación final puede utilizarse un método cuantitativo basado en los requerimientos del usuario y en el desempeño del sistema, dando a cada item un peso relativo; la propuesta con mayor punteo sería la elegida.

Algunos criterios de selección son:

- Facilidad de implementación.

- Homogeneidad de los componentes físicas.

- Capacidad del sistema.
- Ahorros en hardware y personal.
- Plazo de entrega.
- Plazo de desarrollo del software.
- Instrumentos de desarrollo de software.
- Soporte del vendedor.
- Disponibilidad de repuestos y apoyo técnico.
- Facilidad de uso.
- Facilidad de programación.
- Confiabilidad.
- Múltiples procesadores.
- Entrenamiento.
- Capacidad de ampliación.
- Expansión modular.
- Costos de implementación.
- Costos de operación y mantenimiento.
- Costos de transferencia.

Además de los puntos anteriores, debe realizarse una comparación entre hardware, software y comunicaciones de cada propuesta de acuerdo a los items del numeral anterior.

También puede elegirse a la mejor opción por medio de un proceso de eliminación. Algunas razones para eliminar una propuesta podrían ser: deficiencia en soporte en sitio, condición financiera del vendedor, inseguridad en soporte futuro, dificultades en instalaciones similares anteriores,

seguridad en la capacidad de la empresa vendedora, cursos de entrenamiento, disponibilidad de un sistema similar para desarrollo de software antes del arribo del sistema a instalar.

Uno de los criterios fundamentales para la elección son los costos, tanto los de equipo y software que se realizarán una única vez, como los de operación y mantenimiento que se realizan de una manera continua. En la selección de un sistema, al comparar las diferentes arquitecturas, es necesario saber no solamente los costos del hardware, sino también los costos del sistema en términos del personal necesario para implementar, operar y mantener el sistema.

La siguiente es una descripción general de los diferentes tipos de costos:

- Hardware de computación: es el costo de compra o renta de las unidades de procesamiento, memoria, impresoras, terminales, etc; además de los costos de energía, aire acondicionado, preparación de la instalación.

- Hardware de comunicaciones: incluye la interconexión de los nodos de procesamiento y los nodos de conmutación, si existen; incluye los buses, anillos, líneas de teleproceso, concentradores, multiplexores, modems, etc.

- Software de control del sistema: consiste en el sistema operativo, métodos de acceso de comunicación, sistema de manejo de datos, etc. Este software provee la

distribución de datos, localización de recursos, sincronización y técnicas de recuperación. Algunos de estos productos vienen con el hardware y otros se deben adquirir por separado. Además su compra o licencia de uso puede comprender diferentes aspectos; también deben tomarse en cuenta los costos de instalación y mantenimiento.

- Software de aplicación: existe una línea muy difusa entre el software de control y el software de aplicación. El usuario puede desarrollar su propio software de aplicación o puede adquirir paquetes específicos. Debe recordarse que un sistema centralizado permite un mejor control sobre el personal técnico.

- Costos de operación: están asociados con la operación del sistema de computación; incluye operadores de consola, personal para ingreso de datos, personal para realizar modificaciones necesarias para acomodar nuevas versiones de software, etc.

- Costos de administración: incluye operaciones diarias, mantenimiento de hardware, administración del sistema proporcionando directrices, etc.

- Costos de usuario: costos experimentados por el usuario al utilizar el sistema de procesamiento de datos; existen costos de educación, de acceso físico al sistema, de preparación de datos, etc.

Además de la inversión inicial, deben tomarse en cuenta

los costos de conversión tanto en el nuevo hardware y software, como en la evolución del personal del centro de cómputo y los costos de asistencia externa debido a las modificaciones a realizarse.

F. Cronograma de implementación

El cronograma de implementación debe de presentar los siguientes puntos:

- Contrato y negociación.
- Construcción física de los equipos.
- Desarrollo y prueba de datos para programas, subsistemas y el sistema.
- Desarrollo y prueba de programas individuales.
- Integración y prueba de programas como subsistemas.
- Desarrollo de manuales de procedimiento.
- Preparación de sitios (energía, alimentación ininterrumpida, mobiliario, aire acondicionado, piso falso, etc.)
- Tiempo de transporte y desaduanaje.
- Entrenamiento del usuario.
- Instalación y prueba del equipo de computación.
- Instalación y prueba del equipo de comunicaciones.
- Traslado de software.
- Pruebas de aceptación por el usuario.
- Realización de las últimas modificaciones.

X. DISEÑO DE UNA RED DISTRIBUIDA DE DATOS PARA
LA UNIVERSIDAD DEL VALLE DE GUATEMALA.

A. Análisis de las necesidades del usuario

1. Recolección de información.

- Recursos computacionales:

Como primer paso se realizó un estudio de los recursos computacionales existentes en la Universidad del Valle de Guatemala. Los recursos computacionales existentes a abril de 1990, son los siguientes:

a. HP-1000:

Hardware: memoria principal de 1 MB

disco de 50 MB

unidad de memoria de 10 MB

cinta magnética de 80 bpi

impresora de 200 líneas por minuto

lectora de tarjetas de 200 tarjetas por minuto usualmente tiene conectadas una o dos terminales.

Software: Fortran IV

Paquete estadístico

Sistema operativo: MPE 5 (sencillo)

b. HP-3000:

Hardware: memoria principal de 4 MB

disco de 132 MB

disco de 404 MB

cinta magnética de 1600 bpi

impresora de 400 líneas por minuto

36 terminales: 15 para estudiantes, 7 para Centro de Cómputo y 14 para administración (contabilidad, cuentas corrientes, control académico, programa DNL, programa DNF, etc.)

Software: SPSX (paquete estadístico)

Fortran IV

Basic

Cobol I y II

RPG II

Pascal

SPL

Edit 3000

Quad 3000

Sort Merge 3000

Image 3000

B-plus 3000

Sistemas de información:

DSG (Decision support graphics)

TDF (procesador de texto)

Directory 3000

Transac 3000

Inform 3000

Report 3000

Rapid 3000

HP-Word 3000

HP-Slate 3000

Librería científica.

Sistema Operativo: MPE 5 versión G.B3.03 llamado Delta 5

Bases de datos: principalmente una base de datos para contabilidad y control académico, con prioridades para lectura y escritura (FUNUDV.PUB.ADMON).

c. Microcomputadoras:

20 microcomputadoras Epson Equity I+:

10 microcomputadoras con 256 Kb de memoria, 2 disk drives de 5 1/4" y monitor monocromático.

10 microcomputadoras con 640 Kb de memoria, 2 disk drives de 5 1/4" y monitor a color.

A 10 microcomputadoras se les instaló disco duro de 30 MB, monitor a color, mouse y coprocesador matemático 8087-2

Sistema operativo: MS-DOS

Software: QUATTRO

WORDSTAR

WORD PERFECT

SISTAT (Paquete estadístico)

2 microcomputadoras PC XT con monitor monocromático y disco duro de 40 MB, una en caja y otra en registro

académico; estas microcomputadoras pueden trabajar independientemente o pueden conectarse a la HP-3000 emulando terminales por medio de la utilización del paquete de software Reflection. Se utiliza una configuración punto a punto.

5 microcomputadoras AT286 con disco duro de 40 MB y monitor monocromático, en: Lengua y Literatura, Oficina de Relaciones Públicas, Mal de Viñas, Administración del Mal de Viñas y Rectoría.

1 microcomputadora AT386 para uso de los alumnos.

1 microcomputadora Macintosh McIICX equipada con impresora láser para uso del departamento de publicaciones.

- Limitación de Hardware existente:

La computadora HP3000 tiene una capacidad máxima de 32 terminales, pero actualmente se tienen 36 terminales, por lo que no pueden conectarse todos los usuarios simultáneamente. Además, cuando está el sistema totalmente cargado, el tiempo de respuesta es muy lento.

- Modificaciones futuras:

Ampliación de memoria de 7 microcomputadoras Epson de 256 Kb a 640 Kb.

Cambio de la computadora HP-3000 por otra de mayores capacidades.

- Personal del Centro de Cómputo:

Como siguiente paso se determinó la cantidad y cualidad

del personal encargado del Centro de Cómputo:

Jefe: maestro de enseñanza media.

3 programadores: estudiantes universitarios de Ciencias de la Computación.

4 Operadores: 3 bachilleres en Ciencias de la Computación y 1 maestro.

- Modificación de programas:

Las modificaciones necesarias en los programas de aplicación son realizadas por los programadores del Centro de Cómputo.

Usualmente se realizan modificaciones a programas de aplicación de contabilidad y de registro académico; dichos programas utilizan la base de datos general (FUNUDV.PUB.ADMON).

Los lenguajes más utilizados para los programas de aplicación son: Pascal, Transac y Fortran.

- Ingreso de datos:

El ingreso de datos se realiza de la siguiente forma:

Administración: realiza el ingreso de datos para registro académico del Colegio Americano de Guatemala (CAG) y la Universidad del Valle (UDV), además de contabilidad.

Centro de Cómputo: ingreso de datos para cafetería, cuentas corrientes, tesis, DNL, FOES, inscripción de alumnos de UDV y CAG, exámenes de CAG, evaluación de profesores, además de apoyo para registro académico.

Para ingreso de datos a cafetería se utiliza la lectora de tarjetas de la HP-1000; se procesan aproximadamente 150,000 tarjetas al año.

Además se procesan anualmente aproximadamente 180,000 datos para el programa de POES, 10,000 exámenes del CAB y 5,000 evaluaciones de profesores.

- Reportes del Centro de Computo:

Se realizan reportes diarios de cafetería (nombre, cuenta, cantidad pedida, total, etc.).

Se realizan reportes mensuales de cafetería, incluyendo totales acumulados; nómina y cheques de la Fundación del Valle.

Se realizan reportes no periódicos de listado de alumnos (alfabéticos, por grado, etc.), actas de cursos, actas de calificación de estudiantes, etc.

Los reportes diarios y mensuales se distribuyen de manera escrita a contabilidad.

- Uso de microcomputadoras:

En las tablas 10.1 y 10.2 se muestra el uso de las microcomputadoras Epson Equity I, indicando el número de horas de uso por carrera y por aplicación, durante los meses de septiembre y octubre de 1989.

De las tablas anteriores puede observarse que los mayores usuarios son los estudiantes de las carreras de Ciencias de la Computación, Ingeniería Electrónica,

TABLA 10.1

HORAS DE USO DE MICROCOMPUTADORAS. TOTAL Y POR APLICACION
MES DE SEPTIEMBRE DE 1989

CARRERA	APLICACION									TOTAL
	1	2	3	4	5	6	7	8	9	
Cursos Libres	30.70									30.7
Baccalaureatus in Artibus	0.50					3.00				3.5
Baccalaureatus in Scientiis	5.50	4.45				1.00				10.95
Biología						1.60				1.6
Ciencias Agrícolas		3.60			0.25					3.85
Ciencias de la Computación	33.30	91.65	3.22	61.27	3.60	9.10				202.14
Física		0.28								0.28
Ingeniería Civil	9.18					4.10				13.28
Maestría Ingeniería Estructural										0
Ingeniería Electrónica	73.30	45.70		10.70	15.00	2.60	10.00			157.3
Letras										0
Matemática	4.00	3.00								7
Bioquímica	4.85	1.92								6.77
Fisicoquímica		5.50								5.5
Ing. y Ciencia de Alimentos	1.25	2.10			0.60					3.95
Ingeniería Química	65.05	14.25			19.70	1.83	16.00	2.00	5.00	123.83
Nutrición	17.75	0.45								18.2
Química	10.66	3.75						0.50		14.91
Química Farmacéutica	2.60	2.75								5.35
Antropología	1.03					2.35				3.38
Arqueología										0
Ciencias Sociales										0
Economía para el Desarrollo										0
Historia	0.85					0.50				1.35
Linguística										0
Psicología	3.70	6.80				8.85				19.35
Psicología Educativa										0
Educación	1.80					1.75				3.55
Maestría en Administración Educativa	23.60				4.90					28.5
MANEJ	21.85		11.25							33.1
Maestría en Currículo	48.23	2.25			2.67	1.50				54.65
Profesorado en Matemática y Física	103.75	3.50			5.85					113.1
Profesorado en Matemática y Computación	6.07	3.00	4.00			1.05				14.12
Profesorado en Problemas del Aprendizaje						1.50				1.5
TOTAL ESTUDIANTES	469.52	194.95	18.47	71.97	52.57	40.73	26.00	2.50	5.00	881.71
CONCHO/AID	1.70		5.75							7.45
DR. Luján	1.75		6.75							8.5
Facultad de Ciencias y Humanidades	2.00									2
Facultad de Educación	2.00									2
Departamento de Química	1.70									1.7
Departamento de Biología										0
GRAN TOTAL	478.67	194.95	30.97	71.97	52.57	40.73	26	2.5	5	905.36

TABLA 10.2

HORAS DE USO DE MICROCOMPUTADORAS, TOTAL Y POR APLICACION
MES DE OCTUBRE DE 1989

CARRERA	APLICACION									TOTAL
	1	2	3	4	5	6	7	8	9	
Cursos Libres	2.60				5.75					8.35
Baccalaureatus in Artibus	4.50									4.50
Baccalaureatus in Scientiis	18.45	8.00			1.60					28.05
Biología	16.25				2.00					18.25
Ciencias Agrícolas	18.85	3.65								22.50
Ciencias de la Computación	23.20	46.80		63.90	6.45					140.35
Física	25.00				20.85					45.85
Ingeniería Civil	3.40									3.40
Maestría Ingeniería Estructural		0.35								0.35
Ingeniería Electrónica	74.00	12.00		3.15	27.35	1.35				117.85
Letras										0.00
Matemática										0.00
Bioquímica	12.35				8.00					20.35
Fisicoquímica										0.00
Ing. y Ciencia de Alimentos	2.40									2.40
Ingeniería Química	63.20	2.25			20.00		2.35	2.00		89.80
Nutrición	10.50									10.50
Química										0.00
Química Farmacéutica	2.15				1.95					4.10
Antropología	2.65									2.65
Arqueología	13.30									13.30
Ciencias Sociales										0.00
Economía para el Desarrollo										0.00
Historia	2.00									2.00
Linguística										0.00
Psicología	9.30									9.30
Psicología Educativa										0.00
Educación	2.25									2.25
Maestría en Administración Educativa	27.20									27.20
NAMEI	14.80				3.25					18.05
Maestría en Currículo	28.00				2.60					30.60
Profesorado en Matemática y Física	6.00	1.00			38.30					45.30
Profesorado en Matemática y Computación		5.60			1.15					6.75
Profesorado en Problemas del Aprendizaje										0.00
TOTAL ESTUDIANTES	382.35	79.65	0.00	67.05	139.25	1.35	2.35	2.00	0.00	674.00
ONCHD/AID	4.75		3.50		1.80					10.05
DR. Lujan										0.00
Facultad de Ciencias y Humanidades	6.00		3.75		4.00					13.75
Facultad de Educación	0.70				4.60					5.30
Departamento de Química					4.25	2.00		1.00		7.25
Departamento de Biología	3.80									3.80
GRAN TOTAL	397.60	79.65	7.25	67.05	153.90	3.35	2.35	3.00	0.00	714.15

Ingeniería Química y Profesorado en Matemática y Física.

Tomando estas carreras como base, podemos agrupar a los usuarios de las microcomputadoras en los siguientes grupos:

Tabla 10.3

Resumen de horas de uso de las microcomputadoras
Meses de septiembre y octubre de 1989

Carrera	Porcentaje de las horas totales de uso.
Ciencias de la Computación	21.14 %
Electrónica, Física y Matemática	20.25 %
Departamento de Química	18.87 %
Facultad de Educación	23.37 %
Resto de la Facultad de CCHH	16.37 %

En las tablas 10.4 y 10.5 se presenta el número de usuarios de las microcomputadoras, indicando la carrera y la aplicación, para los meses de septiembre y octubre de 1989.

De las tablas anteriores debe tomarse en cuenta que un alumno puede hacer uso de las microcomputadoras para realizar diferentes aplicaciones, por lo que el número total alumnos no es igual a la suma del número total de alumnos por aplicación.

Realizando el mismo tipo de agrupación realizada en la

TABLA 10.4

NUMERO DE USUARIOS DE MICROCOMPUTADORAS, TOTAL Y POR APLICACION
MES DE SEPTIEMBRE DE 1989

CARRERA	APLICACION									TOTAL
	1	2	3	4	5	6	7	8	9	
Cursos Libres	2									2
Baccalaureatus in Artibus	1					1				1
Baccalaureatus in Scientiis	4	3				2				7
Biología						1				1
Ciencias Agrícolas	2				1					3
Ciencias de la Computación	9	17	1	9	2	4				27
Física		1								1
Ingeniería Civil	1					1				2
Maestría Ingeniería Estructural										0
Ingeniería Electrónica	22	15		4	7	2	1			34
Letras										0
Matemática	1	1								1
Bioquímica	2	2								3
Fisicoquímica		1								1
Ing. y Ciencia de Alimentos	1	1			1					3
Ingeniería Química	8	4			4	1	1	1	1	18
Nutrición	1	1								2
Química	3	1						1		4
Química Farmacéutica	2	1								2
Antropología	2									2
Arqueología										0
Ciencias Sociales										0
Economía para el Desarrollo										0
Historia	1					1				1
Linguística										0
Psicología	2	1				3				4
Psicología Educativa										0
Educación	1					1				2
Maestría en Administración Educativa	3				1	2				3
NAMEI	4		2							4
Maestría en Currículo	11	1			3	1				15
Profesorado en Matemática y Física	10	1			1					11
Profesorado en Matemática y Computación	3	1	1							4
Profesorado en Problemas del Aprendizaje						1				1
TOTAL ESTUDIANTES	96	52	4	13	20	21	2	2	1	159

TABLA 10.5

NUMERO DE USUARIOS DE MICROCOMPUTADORAS, TOTAL Y POR APLICACION
MES DE OCTUBRE DE 1989

CARRERA	APLICACION									TOTAL
	1	2	3	4	5	6	7	8	9	
Cursos Libres	1				2					2
Bacalaureatus in Artibus	2									2
Bacalaureatus in Scientiis	4	1			2					6
Biología	2				1					2
Ciencias Agrícolas	2	1								3
Ciencias de la Computación	12	7		9	3					22
Física	1				1					1
Ingeniería Civil	1									1
Maestría Ingeniería Estructural		1								1
Ingeniería Electrónica	27	8		1	11	3				43
Letras										0
Matemática										0
Bioquímica	3				1					3
Fisicoquímica										0
Ing. y Ciencia de Alimentos	2									2
Ingeniería Química	14	3			1		1	1		16
Nutrición	2									2
Química										0
Química Farmacéutica	2				1					2
Antropología	1									1
Arqueología	1									1
Ciencias Sociales										0
Economía para el Desarrollo										0
Historia	1									1
Linguística										0
Psicología	4									4
Psicología Educativa										0
Educación	2									2
Maestría en Administración Educativa	3									3
MAMEI	4									4
Maestría en Currículo	10				2					12
Profesorado en Matemática y Física	1	1			9					9
Profesorado en Matemática y Computación	1				1					2
Profesorado en Problemas del Aprendizaje										0
TOTAL ESTUDIANTES	103	22	0	10	35	3	1	1	0	147

tabla 10.3 se obtiene:

Tabla 10.6

Resumen de número de usuarios de microcomputadoras

Meses de septiembre y octubre de 1989

Carrera	Porcentaje del número total de alumnos.
Ciencias de la Computación	16.01 %
Electrónica, Matemática y Física	26.47 %
Departamento de Química	18.30 %
Facultad de Educación	23.53 %
Resto de la Facultad de CCHH	15.69 %

Analizando el número de horas de uso y el número de alumnos por aplicación se obtiene:

Tabla 10.7

Resumen de utilización de microcomputadoras, por aplicación
Meses de septiembre y octubre de 1989

Aplicación	Porcentaje del total de horas	Porcentaje del total de alumnos
1 Procesador de palabras	54.17 %	65.03 %
2 Lenguaje alto nivel	16.98 %	24.18 %
3 Manejo base de datos	2.36 %	1.31 %
4 Assembler	8.59 %	7.52 %
5 Hoja electrónica	12.76 %	17.97 %
6 Sistema operativo (DOS)	2.74 %	7.84 %
7 Dibujos y gráficas	1.75 %	0.98 %
8 Chemcad	0.34 %	0.98 %
9 Utilerías	0.31 %	0.33 %

De la tabla anterior se puede observar que la mayoría de alumnos utiliza las microcomputadoras únicamente como procesadores de palabras, aunque algunos alumnos las utilizan también para programar en lenguajes de alto nivel y para trabajar en hojas electrónicas.

- Uso de la computadora HP-3000:

En las tablas 10.8, 10.9 y 10.10 se muestra el uso de la computadora HP-3000 durante los meses de julio, agosto y

CUENTAS QUE UTILIZAN LA COMPUTADORA HP-3000

ADMON.CAG: Administración de CAG.

ADMON.CAJA: Administración de caja.

ADMON.PRIM: Administración de primaria del CAG.

ADMON.PUB: Administración del Centro de Cómputo (cuenta pública).

ADMON.SECU: Administración de secundaria del CAG.

ADMON.UDELV: Administración de la Universidad.

ADMON.UTILES: Administración del almacén del CAG.

AECAG.PUB: Asociación de Exalumnos del CAG.

CAG.PUB: Trabajos del Centro de Cómputo en la cuenta CAG.

CAG.REG: Estudiantes regulares del CAG.

CENTCOMP.OPERADOR: Operadores del Centro de Cómputo.

DNL.FUENTE: Programas fuente para estudio del desarrollo del niño longitudinal.

DNL.OPERADOR: operador de programas de desarrollo del niño longitudinal.

DNL.PUB: Trabajos del Centro de Cómputo en el programa DNL.

DNP.OPERADOR: Operadores en el programa del desarrollo del niño preescolar.

DNP.PUB: Trabajos del Centro de Cómputo en el programa DNP.

GRAD.MAEST: Estudiantes de maestría.

GRAD.PUB: Trabajos del Centro de Cómputo en la cuenta GRAD.

- INDIV.PUB: Trabajos realizados por el Centro de Cómputo.
- OEA.VIVIENDA: Programa de la OEA.
- POES.PUB: Programa de orientación de estudios superiores.
- REGU.PUB: Trabajos del Centro de Cómputo sobre la cuenta de alumnos regulares de UDELV.
- ROBELLE.PUB: Programas de U.S.A.
- SALVINIA.PUB: Programa de entomología médica.
- SYS.PUB: Cuenta del director del Centro de Cómputo.
- TELESUP.PRIV: Cuenta para mantenimiento del equipo del Centro de Cómputo.
- UDELV.PUB: Cuenta pública de la UDELV.

TABLA 10.8

USO DE LA COMPUTADORA HP-3000
MES DE JULIO DE 1989

USUARIO	TIEMPO CONEXION (MIN)	SEGUNDOS CPU
ADMON. CAG	1637	1474
ADMON. CAJA	17194	13231
ADMON. CONTA	10498	67936
ADMON. PRIM	217	707
ADMON. PUB	854	2335
ADMON. SECU	2408	10084
ADMON. UDELV	25647	37364
ADMON. UTILES	173	1390
TOTAL ADMON	58628	134521
AECAG. PUB	791	3587
CAG. PUB	3	5
CAG. REG	9213	1730
CENTCOMP. OPERADOR	295	141
DNL. FUENTE	127	179
DNL. OPERADOR	80	60
DNL. PUB	4558	24637
DNP. OPERADOR	704	860
DNP. PUB	12	15
OEA. VIVIENDA	2992	53581
POES. PUB	2810	3744
SALVINIA. PUB	311	1616
SYS. PUB	93	460
TOTAL	80617	225136

TABLA 10.9

USO DE LA COMPUTADORA HP-3000
MES DE AGOSTO DE 1989

USUARIO	TIEMPO CONEXION (MIN)	SEGUNDOS CPU
ADMON. CAG	2448	7509
ADMON. CAJA	15919	11224
ADMON. CONTA	10227	33827
ADMON. PRIM	1556	11089
ADMON. PUB	589	2205
ADMON. SECU	4399	16928
ADMON. UDELV	22176	56489
TOTAL ADMON	57314	139271
AECAG. PUB	284	872
CAG. PUB	30	133
CAG. REG	9179	3365
CENTCOMP. OPERADOR	3	10
DNL. FUENTE	3578	20684
DNL. OPERADOR	5	6
DNL. PUB	6266	17538
DNP. OPERADOR	634	840
DNP. PUB	2	3
GRAD. MAEST	46	927
INDIV. PUB	390	473
OEA. VIVIENDA	705	8663
POES. PUB	4263	16565
REGU. PUB	4	15
ROBELLE. PUB	46	92
SALVINIA. PUB	15	13
SYS. PUB	254	613
TOTAL	83018	210065

TABLA 10.10

USO DE COMPUTADORA HP-3000
MES DE SEPTIEMBRE DE 1989

USUARIO	TIEMPO CONEXION (MIN)	SEGUNDOS CPU
ADMON. CAG	4008	10085
ADMON. CAJA	15226	11096
ADMON. CONTA	9821	42505
ADMON. PRIM	535	3486
ADMON. PUB	743	2015
ADMON. SECU	501	654
ADMON. UDELV	18322	25028
ADMON. UTILES	85	493
TOTAL ADMON	49241	95364
AECAG. PUB	328	555
CAG. PUB	7	52
CAG. REG	17231	4939
DNL. FUENTE	5692	88006
DNL. OPERADOR	2	6
DNL. PUB	4191	36367
DNP. OPERADOR	961	787
GRAD. MAEST	245	954
GRAD. PUB	2	2
INDIV. PUB	2	6
OEA. VIVIENDA	546	3548
POES. PUB	8781	48978
ROBELLE. PUB	26	717
SYS. PUB	117	735
TELESUP. PRV	35	62
UDELV. PUB	10	3
TOTAL	87417	281081

septiembre de 1989, indicándose los minutos de conexión y los segundos de CPU.

En la tabla siguiente se presenta el porcentaje de los minutos totales de conexión:

Tabla 10.11

Resumen de los porcentajes de utilización de la HP 3000

Meses de julio, agosto y septiembre de 1989

Usuario	Julio	Agosto	Septiembre
ADMON	72.72 %	69.04 %	56.33 %
DNL	5.91 %	11.86 %	11.31 %
DNF	0.89 %	0.77 %	1.10 %
FOES	3.49 %	5.14 %	10.04 %
DEA	3.71 %	0.85 %	0.62 %
CAG	11.43 %	11.08 %	19.72 %
RESTO	1.85 %	1.26 %	0.88 %

La mayor utilización se dió en las cuentas ADMIN, DNL, FOES y CAG. Debe observarse que los alumnos de la Universidad del Valle prácticamente no utilizan la computadora HP-3000, utilizando principalmente las microcomputadoras.

- Número de alumnos inscritos en la Facultad de CCHH:

TABLA 10.12

INSCRIPCIÓN DEL COLEGIO UNIVERSITARIO Y DE LAS FACULTADES DE CIENCIAS Y HUMANIDADES Y DE CIENCIAS SOCIALES, POR CARRERAS Y GRADOS, CORRESPONDIENTES AL SEGUNDO CICLO DE 1986

CARRERA	GRADO					TESIS	MAESTRIA	TOTAL POR CARRERA	
	0	1	2	3	4				5
Cursos Libres	27							27	
Baccalaureatus in Artibus	5							5	
Baccalaureatus in Scientiis	35							35	
Biología	0	4	2	5	6	14	7	0	38
Ciencias Agrícolas	0	10	9	8	8	7	12	0	54
Ciencias de la Computación	0	22	17	9	9	11	11	0	79
Física	0	0	1	2	0	1	2	0	6
Ingeniería Civil	0	3	4	5	3	8	1	0	24
Maestría Ingeniería Estructural	0	0	0	0	0	0	0	10	10
Ingeniería Electrónica	0	57	41	23	1	0	0	0	122
Letras	0	0	0	6	4	4	4	0	18
Matemática	0	1	0	3	0	1	1	0	6
Bioquímica	0	4	3	2	3	1	2	0	15
Fisicoquímica	0	0	1	1	0	0	0	0	2
Inq. y Ciencia de Alimentos	0	20	13	1	3	3	0	0	40
Ingeniería Química	0	16	16	9	16	2	1	0	60
Nutrición	0	4	0	0	0	0	0	0	4
Química	0	1	2	0	1	0	3	0	7
Química Farmacéutica	0	7	4	2	0	0	0	0	13
Antropología	0	0	0	0	0	2	2	0	4
Arqueología	0	0	0	0	3	2	1	0	6
Ciencias Sociales	0	0	0	0	0	0	1	0	1
Economía para el Desarrollo	0	0	0	0	1	0	0	0	1
Historia	0	2	3	5	1	1	6	0	18
Linguística	0	0	0	0	0	0	1	0	1
Psicología	0	2	3	5	5	7	2	0	24
Psicología Educativa	0	0	2	1	0	1	0	0	4
TOTALES	67	153	121	87	64	65	57	10	624

TABLA 10.13

INSCRIPCIÓN DEL COLEGIO UNIVERSITARIO Y DE LAS FACULTADES DE CIENCIAS Y HUMANIDADES Y DE CIENCIAS SOCIALES, POR CARRERAS Y GRADOS, CORRESPONDIENTES AL SEGUNDO CICLO DE 1987

CARRERA	GRADO						TESIS	MAESTRIA	TOTAL POR CARRERA
	0	1	2	3	4	5			
Cursos Libres	24								24
Baccalaureatus in Artibus	7								7
Baccalaureatus in Scientiis	61								61
Biología	0	6	2	2	1	16	5	0	32
Ciencias Agrícolas	0	11	11	11	10	7	9	0	59
Ciencias de la Computación	0	14	16	13	10	12	8	0	73
Física	0	2	0	1	2	0	1	0	6
Ingeniería Civil	0	3	6	5	6	4	2	0	26
Maestría Ingeniería Estructural	0	0	0	0	0	0	0	14	14
Ingeniería Electrónica	0	43	43	42	18	0	0	0	146
Letras	0	0	0	9	5	5	2	0	21
Bioquímica	0	4	6	2	3	2	0	0	17
Matemática	0	0	0	1	2	1	0	0	4
Fisicoquímica	0	0	0	1	1	0	0	0	2
Ing. y Ciencia de Alimentos	0	17	16	7	1	3	1	0	45
Ingeniería Química	0	20	16	15	8	16	1	0	76
Nutrición	0	4	4	4	1	0	0	0	13
Química	0	0	4	2	0	0	4	0	10
Química Farmacéutica	0	4	6	4	2	0	0	0	16
Antropología	0	0	0	1	0	0	3	0	4
Arqueología	0	4	1	0	0	2	0	0	7
Ciencias Sociales	0								0
Economía para el Desarrollo	0								0
Historia	0	3	1	1	1	1	8	0	15
Linguística	0	0	0	0	0	0	0	0	0
Psicología	0	5	1	7	6	5	1	0	25
Psicología Educativa	0	0	0	2	0	0	1	0	3
TOTALES	92	140	133	130	77	74	46	14	706

TABLA 10.14

INSCRIPCIÓN DEL COLEGIO UNIVERSITARIO Y DE LAS FACULTADES DE CIENCIAS Y HUMANIDADES Y DE CIENCIAS SOCIALES, POR CARRERAS Y GRADOS, CORRESPONDIENTES AL SEGUNDO CICLO DE 1988

CARRERA	GRADO						TESIS	MAESTRIA	TOTAL POR CARRERA
	0	1	2	3	4	5			
Cursos Libres	31								31
Baccalaureatus in Artibus	3								3
Baccalaureatus in Scientiis	60								60
Biología	0	3	4	2	1	3	3	0	16
Ciencias Agrícolas	0	5	8	7	13	10	4	0	47
Ciencias de la Computación	0	18	12	13	11	12	10	0	76
Física	0	0	1	1	2	0	1	0	5
Ingeniería Civil	0	5	5	8	3	5	2	0	28
Maestría Ingeniería Estructural	0	0	0	0	0	0	0	4	4
Ingeniería Electrónica	0	39	33	23	28	12	0	0	135
Letras	0	0	0	5	5	5	4	0	19
Matemática	0	1	0	1	2	2	1	0	7
Bioquímica	0	9	6	5	3	4	0	0	27
Fisicoquímica	0	0	0	0	1	1	0	0	2
Ing. y Ciencia de Alimentos	0	15	15	14	4	2	1	0	51
Ingeniería Química	0	27	28	12	15	8	12	0	102
Nutrición	0	3	2	4	3	1	0	0	13
Química	0	0	1	2	2	2	0	0	7
Química Farmacéutica	0	3	2	4	4	2	0	0	15
Antropología	0	2	0	0	1	1	4	0	8
Arqueología	0	3	3	0	0	0	3	0	9
Ciencias Sociales	0								0
Economía para el Desarrollo	0								0
Historia	0	2	2	2	9	0	5	0	20
Linguística	0								0
Psicología	0	6	3	4	8	7	3	0	31
Psicología Educativa	0	1	0	2	2	0	0	0	5
TOTALES	94	142	125	109	117	77	53	4	721

TABLA 10.15

INSCRIPCIÓN DEL COLEGIO UNIVERSITARIO Y DE LAS FACULTADES DE CIENCIAS Y HUMANIDADES Y DE CIENCIAS SOCIALES, POR CARRERAS Y GRADOS, CORRESPONDIENTES AL SEGUNDO CICLO DE 1989

CARRERA	GRADO						MAESTRIA	TESIS	TOTAL POR CARRERA
	0	1	2	3	4	5			
Cursos Libres	18								18
Baccalaureatus in Artibus	4								4
Baccalaureatus in Scientiis	39								39
Biología	0	6	2	3	2	3	5	0	21
Ciencias Agrícolas	0	8	6	7	8	11	6	0	46
Ciencias de la Computación	0	13	10	4	8	9	13	0	57
Física	0	1	0	1	2	0	1	0	5
Ingeniería Civil	0	5	4	4	7	4	1	0	25
Maestría Ingeniería Estructural	0	0	0	0	0	0	0	6	6
Ingeniería Electrónica	0	36	19	21	20	26	4	0	126
Letras	1	0	0	8	4	7	5	0	25
Matemática	0	0	0	0	1	1	3	0	5
Bioquímica	0	10	6	4	3	4	2	0	29
Fisicoquímica	0	0	1	0	0	0	1	0	2
Inq. y Ciencia de Alimentos	0	18	14	12	11	5	0	0	60
Ingeniería Química	0	27	20	22	11	14	7	0	101
Nutrición	0	7	5	5	4	3	0	0	24
Química	0	1	0	1	1	3	0	0	6
Química Farmacéutica	0	7	2	2	4	5	0	0	20
Antropología	0	5	1	0	1	0	1	0	8
Arqueología	0	3	3	4	0	1	3	0	14
Ciencias Sociales	0	1	0	0	0	0	1	0	2
Economía para el Desarrollo	0								0
Historia	0	5	1	3	1	9	6	0	25
Linguística	0								0
Psicología	0	7	9	2	6	7	3	0	34
Psicología Educativa	0	1	1	0	2	1	0	0	5
TOTALES	62	161	104	103	96	113	62	6	707

En las tablas 10.12, 10.13, 10.14 y 10.15 se muestra el número de alumnos inscritos por carrera en la Facultad de Ciencias y Humanidades durante el segundo ciclo de los años 1986, 1987, 1988 y 1989.

En la tabla siguiente se muestra el número de alumnos inscritos en el segundo ciclo, agrupados de acuerdo al criterio de la tabla 10.3.

Tabla 10.16

Resumen de alumnos inscritos en el segundo ciclo
Años de 1986, 1987, 1988 y 1989

Carrera	1986	1987	1988	1989
Computación	79	73	76	57
Electrónica, Física y Matemática	134	156	147	136
Departamento de Química	141	179	217	242
Otros	270	298	281	272
Total	624	706	721	707

En la tabla siguiente se muestra el porcentaje del total de alumnos inscritos en el segundo ciclo, agrupados de acuerdo al criterio utilizado en el tabla 10.3.

Tabla 10.17

Resumen de porcentaje de alumnos inscritos en el segundo ciclo. Años 1986, 1987, 1988 y 1989

Carrera	1986	1987	1988	1989
Computación	12.66%	10.34%	10.54%	8.06%
Electrónica, Matemática y Física	21.47%	22.10%	20.39%	19.24%
Departamento de Química	22.60%	25.35%	30.10%	34.23%
Otros	43.27%	42.21%	38.97%	38.47%

- Número de alumnos inscritos en la Facultad de Educación:

En las tablas 10.18, 10.19, 10.20 y 10.21 se presenta el número de alumnos inscritos en la Facultad de Educación durante los años de 1986, 1987, 1988 y 1989.

En la tabla siguiente se muestra de manera resumida el número total de alumnos inscritos:

TABLA 10.18

INSCRIPCION DE LA FACULTAD DE EDUCACION, POR
CARRERAS, CORRESPONDIENTE AL PRIMER CICLO DE 1986.

CARRERA	TOTAL POR CARRERA
Cursos Libres	29
Licenciatura en Educación	28
Licenciatura en Educación para la salud	0
Maestrías:	
Medición, evaluación e investigación educativas	8
Administración educativa	28
Curriculo	0
Profesorado enseñanza media en:	
Biología y Ciencias Químicas	19
Ciencias Naturales	22
Educación de niños especiales	90
Niños con problemas auditivos	0
Historia y Ciencias Sociales	17
Historia y Estudios Sociales	42
Inglés	150
Lengua y Literatura	42
Matemática	41
Matemática y Ciencias Físicas	30
Matemática y Computación	0
Problemas del aprendizaje	116
TOTALES	662

TABLA 10.19

INSCRIPCION DE LA FACULTAD DE EDUCACION, POR
CARRERAS, CORRESPONDIENTE AL PRIMER CICLO DE 1987.

CARRERA	TOTAL POR CARRERA
Cursos Libres	33
Licenciatura en Educación	32
Licenciatura en Educación para la salud	0
Maestrías:	
Medición, evaluación e investigación educativas	8
Administración educativa	30
Curriculo	0
Profesorado enseñanza media en:	
Biología y Ciencias Químicas	17
Ciencias Naturales	11
Educación de niños especiales	98
Niños con problemas auditivos	0
Historia y Ciencias Sociales	19
Historia y Estudios Sociales	30
Inglés	130
Lengua y Literatura	42
Matemática	26
Matemática y Ciencias Físicas	44
Matemática y Computación	0
Problemas del aprendizaje	112
TOTALES	632

TABLA 10.20

INSCRIPCION DE LA FACULTAD DE EDUCACION, POR
CARRERAS, CORRESPONDIENTE AL PRIMER CICLO DE 1988.

CARRERA	TOTAL POR CARRERA
Cursos Libres	44
Licenciatura en Educación	29
Licenciatura en Educación para la salud	0
Maestrías:	
Medición, evaluación e investigación educativas	9
Administración educativa	25
Currículo	0
Profesorado enseñanza media en:	
Biología y Ciencias Químicas	24
Ciencias Naturales	5
Educación de niños especiales	111
Niños con problemas auditivos	0
Historia y Ciencias Sociales	19
Historia y Estudios Sociales	21
Inglés	125
Lengua y Literatura	42
Matemática	17
Matemática y Ciencias Físicas	39
Matemática y Computación	18
Problemas del aprendizaje	126
TOTALES	654

TABLA 10.21

INSCRIPCION DE LA FACULTAD DE EDUCACION, POR
CARRERAS, CORRESPONDIENTE AL PRIMER CICLO DE 1989.

CARRERA	TOTAL POR CARRERA
Cursos Libres	37
Licenciatura en Educación	37
Licenciatura en Educación para la salud	55
Maestrías:	
Medición, evaluación e investigación educativas	8
Administración educativa	20
Curriculo	0
Profesorado enseñanza media en:	
Biología y Ciencias Químicas	32
Ciencias Naturales	5
Educación de niños especiales	96
Niños con problemas auditivos	12
Historia y Ciencias Sociales	25
Historia y Estudios Sociales	9
Inglés	104
Lengua y Literatura	30
Matemática	11
Matemática y Ciencias Físicas	36
Matemática y Computación	32
Problemas del aprendizaje	114
TOTALES	663

Tabla 10.22

Resumen de estudiantes inscritos en la Facultad de Educación
Años de 1986, 1987, 1988 y 1989

Año	1986	1987	1988	1989
Número de alumnos	662	632	654	663

2. Análisis de la información.

- Usuarios:

Los estudiantes de la Universidad del Valle que usan los recursos computacionales, utilizan únicamente las microcomputadoras. Los estudiantes de Electrónica y Ciencias de la Computación son prácticamente los únicos que utilizan las microcomputadoras para programar en lenguaje de alto nivel y assembler, los demás estudiantes las utilizan principalmente como procesador de palabras y para trabajar en hojas electrónicas.

Los estudiantes de la Universidad del Valle pueden agruparse, por carrera, en los siguientes grupos:

Ciencias de la Computación.

Electrónica, Matemática y Física.

Departamento de Química.

Resto de carreras de la Facultad de CCHH.

Facultad de Educación.

El número de estudiantes inscrito en cada uno de los grupos anteriores ha permanecido relativamente estable en los últimos años.

La computadora HP-3000 es utilizada principalmente por el personal del Centro de Cómputo y por el personal administrativo de la Fundación del Valle. El personal del Centro de Cómputo realiza modificación de programas e ingreso de datos; el resto de usuarios realiza únicamente ingreso de información.

La programación es principalmente para modificar programas que utilizan como fuente la base de datos FUNUDV.FUE.ADMON.

Los estudiantes del colegio Americano utilizan la computadora HP-3000 para realizar programas de acuerdo a los cursos de computación que estudian.

- Funciones a realizar:

Los estudiantes de la Universidad del Valle utilizan principalmente las microcomputadoras, realizando en ellas sus propios programas. Cada estudiante posee sus propios datos y sus propios paquetes de software en diskettes independientes, sin necesitar datos de otros estudiantes.

Los estudiantes del Colegio Americano almacenan sus programas en memoria secundaria de la HP-3000.

La administración utiliza la base de datos FUNDV.FUE.ADMON; además se realiza entrada y salida de

información en las bases de datos para los programas de POES, DNL, DNP, etc. Las bases de datos anteriores son totalmente independientes.

- Intercambio de información:

No existe un intercambio electrónico de información entre los diferentes departamentos de la Universidad; únicamente se distribuyen reportes escritos.

- Operación del sistema:

El personal del Centro de Cómputo realiza la administración de los recursos computacionales. El ingreso de la información se realiza de manera manual a través de una terminal o por medio de la lectora de tarjetas. Se presentan reportes escritos, diarios y mensuales, a contabilidad.

Los estudiantes presentan de manera escrita sus trabajos: tareas, programas, reportes de laboratorio, investigaciones, etc.

No existe comunicación con otros sistemas.

- Desempeño:

No existe problema con sobrecarga de las microcomputadoras, se utiliza una microcomputadora por alumno; únicamente se sobrecarga el sistema cuando se debe entregar algún proyecto ya que el número de usuarios es mayor que el número de microcomputadoras disponibles.

La HP-3000 tiene 36 terminales cuando su capacidad

máxima es de 32. Cuando el sistema está totalmente cargado el tiempo de respuesta es muy lento.

- Restricciones:

Por muchos años se han realizado programas en la HF-3000, por lo que el personal del Centro de Cómputo desea que cualquier ampliación o modificación del sistema existente se realice con tecnología Hewlett Packard.

B. Diseño del sistema

El sistema actual presenta una división bien definida de funciones y datos: los estudiantes de la Universidad realizan sus propios programas de software, utilizan sus propios paquetes de software y sus propios datos; el personal del Centro de Cómputo utiliza la computadora HF-3000 en todos los trabajos administrativos, utilizando principalmente la base de datos FUNUDV.PUB.ADMON, además existen otras bases de datos para los programas de DNL, POES, DNF, etc. siendo cada base de datos independiente.

Cada aplicación utiliza únicamente sus propios datos; no existe una aplicación que utilice datos de diferentes bases de datos, lo que implica que sería recomendable realizar una división de datos y de aplicaciones. La principal razón para replicar sería por seguridad y para poder permitir acceso local desde varios nodos.

De acuerdo a la utilización de los recursos computacionales se pueden definir los siguientes nodos:

- Administración.
- Estudiantes de Computación.
- Estudiantes de Electrónica, Física y Matemática.
- Estudiantes del Departamento de Química.
- Resto de estudiantes de la Facultad de CCHH.
- Estudiantes de la Facultad de Educación.

La Administración continuaría realizando sus mismas funciones utilizando principalmente la base de datos FUNUDV.FUB.ADMON.

Los demás nodos utilizarían principalmente datos locales y desarrollarían sus propios programas.

Para darle seguridad al sistema cada nodo debe ser capaz de actuar independientemente, a pesar de problemas en el sistema de comunicaciones o de fallas en otros nodos.

Debe permitirse el acceso a las bases de datos de los otros nodos, tanto para lectura como para modificación, siempre y cuando se tengan las claves adecuadas.

La base de datos principal es la FUNUDV.FUB.ADMON por lo que por motivos de seguridad podría estar replicada en algún otro nodo, en memoria secundaria; esta copia sería únicamente de reserva y se modificaría al final del día.

Cada nodo deberá poseer un directorio local propio y posibilidad de acceso a los directorios de los demás nodos.

Se mantendrá un control centralizado de la red; el Centro de Cómputo se mantendrá como encargado del

funcionamiento de todos los recursos computacionales. El mantenimiento del sistema debe poder realizarse a partir del nodo principal (ADMÓN) y con las claves adecuadas acceder a los diferentes nodos.

Cada nodo necesita al menos una impresora para salida de reportes escritos y terminales de acuerdo al número estimado de posibles usuarios.

El ingreso de datos será principalmente por medio de terminales; el nodo de Administración necesitará además una lectora de tarjetas.

Actualmente no se posee una interconexión con otros sistemas; es necesario que la red tenga acceso a la red de conmutación de paquetes, todos los usuarios del sistema deberán de poder utilizar esta facilidad si poseen la clave adecuada. Se sugiere un acceso inicial de 9,600 bps, en línea conmutada (V.32) con posibilidad para tener un acceso en X.25.

Es recomendable se tenga para el futuro la posibilidad de acceso a la futura red ISDN en modo básico (2B+D).

- Número de terminales:

Únicamente se conoce que la Administración (HF-3000) utiliza 36 terminales y éstas son insuficientes; además los alumnos utilizan 20 microcomputadoras.

Para la Administración podemos esperar un sistema 40 % mayor, lo que implica tener para ese nodo 50 terminales.

Para los demás nodos se podría pensar en una utilización inicial pico del 25 % del total de usuarios, con capacidad para ampliarse al 40 %.

Tabla 10.23

Número de terminales requeridas

Nodo	Número Usuarios	Terminales Inicial	Terminales Ampliación
Computación	25	6	11
Electrónica	41	10	17
Química	28	7	12
CCHH	36	9	15
Educación	24	6	10
Total	154	38	65

- Dimensionamiento del sistema de comunicación entre los diferentes nodos:

Actualmente no se poseen datos para realizar un análisis cuantitativo debido a que no existe una comunicación electrónica entre los diferentes departamentos ni entre los estudiantes; pero puede esperarse dicha comunicación al tenerse el sistema funcionando.

Si tomamos que el 20% del total de posibles usuarios hacen uso del canal de comunicación a una velocidad de 9600 bps, obtenemos:

$$(65+50)*0.2 = 23 \text{ usuarios.}$$

$$23*9600\text{bps} = 220 \text{ Kbps.}$$

220 Kbps*1.5 = 330 Kbps (debido a encabezados, retransmisión, crecimiento futuro, etc).

De lo anterior puede observarse que se puede utilizar una LAN de baja velocidad (Ej. Ethernet, cable paralelo, baja velocidad).

C. Requerimientos del sistema

De acuerdo con el análisis realizado en el numeral anterior se pueden presentar, principalmente, dos diferentes soluciones:

- Opción 1:

Consiste en localizar una minicomputadora en cada nodo e interconectarlas a través de una LAN (Ethernet). Las características de cada nodo serían las siguientes:

- Administración:

1 minicomputadora.

50 terminales monocromáticas.

2 impresoras de alta velocidad.

1 impresora láser.

1 lectora de tarjetas.

- Estudiantes de Computación:

1 minicomputadora.

6 terminales monocromáticas (con posibilidad de ampliar a 11 terminales).

1 impresora de alta velocidad.

- Estudiantes de Electrónica, Física y Matemática:

1 minicomputadora.

10 terminales monocromáticas (con posibilidad de ampliar a 17 terminales).

1 impresora de alta velocidad.

- Estudiantes del Departamento de Química:

1 minicomputadora.

7 terminales monocromáticas (con posibilidad de ampliar a 12 terminales).

1 impresora de alta velocidad.

- Resto de estudiantes de la Facultad de CCHH:

1 minicomputadora.

9 terminales monocromáticas (con posibilidad de ampliar a 15 terminales).

1 impresora de alta velocidad.

- Estudiantes de la Facultad de Educación:

1 minicomputadora.

6 terminales monocromáticas (con posibilidad de ampliar a 10 terminales).

1 impresora de alta velocidad.

Para todos los nodos será necesario proporcionar

alimentación ininterrumpida (UPSes) de estado sólido con capacidad para mantener el sistema funcionando independientemente por lo menos 15 minutos. Además debe de proporcionarse toda la instalación eléctrica y de software del sistema. Debe de proporcionarse aire acondicionado y piso falso de acuerdo a las necesidades del equipo.

Las características del equipo anterior deben ser las siguientes:

- Minicomputadora de Administración:

CPU de 32 bits.

10 slots de expansión.

20 MB de memoria principal.

Disco SCSI de 1.2 GB.

Cinta de 400 MB.

Controlador Ethernet.

- Demás minicomputadoras:

CPU de 32 bits.

8 slots de expansión.

Disco SCSI de 300 MB.

Cinta de 300 MB.

Controlador Ethernet.

- Terminales monocromáticas:

Monitor de 13 pulgadas.

80 columnas por 25 filas como mínimo.

Verde, ámbar o blanca.

Teclado ASCII extendido con caracteres en español.

Full duplex hasta 19,200 bps.

Interface para computadora principal e impresora.

En lugar de terminales se pueden utilizar las microcomputadoras existentes y utilizando hardware y/o software emular una terminal para la minicomputadora (la opción anterior es posible para todos los nodos excepto el de Administración).

- Impresoras de alta velocidad:

Velocidad de al menos 600 lpm.

Memoria de 1 KB como mínimo.

Diferentes fonts y calidades de letras.

Interface paralela y serial.

- Impresora láser:

Velocidad mínima de 8 ppm.

Tamaño de papel mínimo de 8 1/2" X 11".

Interface paralela y serial.

Es necesario proporcionar acceso a la red de conmutación de paquetes, inicialmente con un modem que cumpla con las siguientes características:

Velocidad de hasta 9,600 bps.

Full duplex, asíncrono y sincrónico.

Para línea privada o conmutada.

Compatible Bell y CCITT.

Interface RS-232.

Comandos AT.

Posteriormente deberá ser posible el acceso X.25 a la red de conmutación de paquetes. También debe ser posible en un futuro proporcionar acceso a la red ISDN en acceso básico (2B+D).

Los requerimientos de Software son los siguientes:

- Sistema operativo:

Multitarea, multiusuario, capacidad de soportar lenguajes de cuarta generación (4GL), permitir el uso simultáneo de datos por varios procesos.

- Sistema de manejo de base datos distribuida.

- CASE/ Generador de Aplicaciones de 4GL.

- Procesador de palabras.

- Hoja electrónica.

- Programa de gráficas.

- Programas de estadística, matemática, química, circuitos eléctricos y electrónicos, etc.

- Interconexión a DOS y a UNIX.

Se debe de proveer la licencia de uso para todos los posibles usuarios.

Adicionalmente se debe proveer la posibilidad de interconexión de microcomputadoras (IBM compatibles) a la red.

- Opción 2:

Como se observa del análisis de las necesidades del

usuario, los estudiantes de la Universidad utilizan usualmente sus propios programas y datos; los estudiantes utilizan las microcomputadoras y la Administración utiliza la computadora HP-3000.

Por lo anterior puede pensarse en una minicomputadora para Administración con las mismas características de hardware y software presentadas en la Opción 1 y para cada uno de los demás nodos, proveer un servidor (server) y crear varias redes locales de microcomputadoras.

Se utilizarían las microcomputadoras existentes (aprox. 20) y se necesitarían 18 microcomputadoras adicionales, dando un total de 38 microcomputadoras, igual al número de terminales necesarias. Además se necesitarían 5 servidores (AT compatibles) para cada nodo de estudiantes.

- Características de microcomputadora XT:

Monitor monocromático.

Procesador 8086/8088.

Disk drive de 5 1/4", 360 Kb.

Disk drive de 3 1/2", 720 Kb.

Disco duro de 20 MB.

- Características de microcomputadora AT:

Monitor monocromático.

Procesador 80286/80386.

Disk drive de 5 1/4", 360 Kb.

Disk drive de 3 1/2", 720 Kb.

Impresora de alta velocidad (igual a Opción 1)

Disco duro de 200 MB.

Al igual que en la opción 1, debe proporcionarse el acceso inicial a la red de conmutación de paquetes a través de un modem de 9,600 bps y posibilidad de acceso futuro en X.25. También debe estar contemplado el acceso futuro a la red ISDN en acceso básico (2B+D).

Las microcomputadoras de cada nodo de estudiantes pueden formar una LAN para cada departamento utilizando Novell's Advance Netware como sistema operativo de la red.

La interconexión entre las LANs de microcomputadoras y la minicomputadora de Administración puede realizarse utilizando un Gateway (Ej: Micom TCP Gateway) que permite la comunicación bidireccional entre microcomputadoras en un ambiente Netware/DOS y servidores, estaciones de trabajo (workstations), minicomputadoras y computadoras grandes (mainframes) corriendo TCP/IP en una LAN Ethernet; permitiendo movimiento de archivos entre anfitrión-anfitrión (host-host), anfitrión-servidor, microcomputadora-servidor, microcomputadora-anfitrión, microcomputadora-anfitrión y todas las combinaciones posibles.

Para todos los nodos será necesario proporcionar UPSes de estado sólido con capacidad para mantener el sistema funcionando independientemente por lo menos 15 minutos. Además debe de proporcionarse toda la instalación eléctrica

y de software del sistema. Debe de proporcionarse aire acondicionado y piso falso de acuerdo a las necesidades del equipo.

D. Elección entre las opciones

Tanto la opción 1 como la opción 2 son técnicamente factibles y presentan una solución a las necesidades computacionales de la Universidad del Valle de Guatemala.

En ambas opciones se les da independencia a la Administración y a los diferentes departamentos de estudiantes y se permite un crecimiento modular del sistema.

Con la opción 1, se tiene una minicomputadora para cada nodo, estas minicomputadoras deben ser del mismo vendedor, con lo que se logra una gran homogeneidad y además se obtiene más fácilmente la imagen de un solo sistema, permitiéndose de una manera más sencilla el acceso remoto a bases de datos, el acceso remoto a archivos y periféricos, la transferencia de archivos, comunicación entre programas, etc.

Con la opción 2 no se posee un sistema homogéneo con lo que se dificulta presentar la imagen de un solo sistema, el acceso a las bases de datos de otros nodos no será tan transparente como en la opción 1. Al poseer diferente tipo de equipo computacional se complica un poco el mantenimiento del hardware y del software y el control de la red.

Fero a pesar de las ventajas de la opción 1 respecto a

la opción 2; con la opción 2 se obtiene un aprovechamiento mejor de los recursos computacionales existentes (microcomputadoras); además, su costo es bastante menor al utilizar microcomputadoras en los nodos de estudiantes. Adicionalmente, del análisis de la información del usuario, se tiene que el intercambio de información entre los diferentes nodos no será muy alto ya que usualmente se utilizan datos locales. Cuando se necesita la utilización de un programa de aplicación (para ing. civil, electrónica, física, química, etc.) que únicamente puede correr en la minicomputadora, el estudiante puede acceder la minicomputadora.

Con ambas opciones se puede obtener soporte del vendedor, entrenamiento, disponibilidad de repuestos y apoyo técnico.

Por lo anterior resulta más recomendable la implementación de la opción 2; ya que soluciona las necesidades computacionales de la Universidad del Valle a un costo menor.

XI. CONCLUSIONES

1. Un sistema distribuido presenta ciertas ventajas sobre un sistema centralizado; entre las más importantes se pueden mencionar:

- Mejora la seguridad debido a que se tienen varios nodos, al fallar un nodo, únicamente se ven afectados los usuarios de dicho nodo.

- Se obtiene una rápida respuesta al hacerse generalmente referencia a datos locales.

- Permite un crecimiento modular del sistema.

- Debido a su estructura granular, se puede mapear el sistema sobre la estructura de la organización.

- Proporciona mayor independencia a cada departamento o sección de la organización al dársele control sobre los recursos computacionales que necesita para realizar su función.

- Evita la duplicidad en programas de aplicación al permitir el acceso de otros departamentos al nodo de un departamento dado.

2. No se debe confundir la distribución con la descentralización. La descentralización es un concepto organizacional y de administración. Se puede distribuir el hardware y el software, manteniendo un control centralizado

de la red.

3. Al decidir acerca del método de distribución de los datos, deben tomarse en cuenta los intercambios básicos entre los costos de almacenamiento y la confiabilidad y entre los costos de acceso y los costos de modificación. En una base de datos centralizada o dividida se tiene una sola copia lo que implica un costo de almacenamiento y una confiabilidad baja, un costo de modificación bajo y un costo de acceso alto. Con una base de datos replicada se tiene un costo de almacenamiento y una confiabilidad alta, un costo de acceso bajo y un costo de modificación alto. Se debe elegir el tipo de distribución de datos de acuerdo a los requerimientos de la organización.

4. Al diseñar una red de datos debe preverse siempre el acceso a otras redes o computadoras, ya sea por medio de modem en una línea telefónica conmutada o privada o a través de la red de conmutación de paquetes.

5. Al realizar el diseño de una red de datos para una organización siempre debe de realizarse un análisis profundo, ya que la solución para una organización no es necesariamente cierta para otra. Si para una empresa puede ser recomendable una red distribuida, para otra puede serlo una red centralizada.

6. Para la Universidad del Valle de Guatemala es recomendable una división de datos y de funciones; se puede

tener una minicomputadora para la Administración y para cada nodo de estudiantes tener una LAN de microcomputadoras, estando todos los nodos enlazados a través de una LAN.

XII. BIBLIOGRAFIA

- Bray, O. Distributed databases management systems.
1982 New York, Lexington Brooks. 164 pp.
- Breslin, J. Distributed processing systems. New
1981 York, Amacon. 228 pp.
- Durando, C. Consideraciones sobre el uso actual y
1979 futuro del teleproceso en Guatemala. Gua-
temala. 80 pp.
- Fanjul, R. La transmisión de datos y las caracterís-
1986 ticas de una red de teleinformática privada
en Guatemala. Guatemala. 80 pp.
- Lientz, B. An introduction to distributed systems.
1981 New York, Addison-Wesley Publishing Company.
454 pp.
- Lorin, H. Aspects of distributed computer systems.
1980 New York, John Wiley & Sons. 276 pp.
- Sharma, R. Network systems. Pennsylvania, Van
1982 Reinhold Company. 321 Pp.
- Stallings, W. Data and computer communications. New
1986 York, Macmillan Publishing Company. 560 pp.
- Tanembaum, A. Computer Networks. New York, Prentice
1981 Hall. 517 Pp
- Taub, H. y Schilling, D. Principles of communications
1986 systems. Singapore, McGraw-Hill. 759 pp.

ANEXO A

RED DIGITAL DE SERVICIOS INTEGRADOS (ISDN)

En la última década se han desencadenado una gran serie de cambios en las redes de telecomunicaciones debido a nuevos requerimientos: las redes de telecomunicaciones pasan de un contexto nacional a formar parte de una red global; se busca un mejoramiento en la calidad de transmisión de información y se han multiplicado los modos de transmisión, incluyendo la transmisión telefónica, telex, facsímil, transmisión de datos, transmisión de imágenes, etc.

Debido a esta diversidad de servicios, no se puede esperar tener una red dedicada para cada uno de los servicios ya que los costos de operación y mantenimiento de cada una de las redes puede llegar a ser prohibitivo; cada red necesitaría una conexión física específica y terminales específicas. Además, cada red tendría procedimientos de acceso y direccionamiento diferentes, dificultando su utilización por los usuarios.

Por lo anterior se llega al concepto de la red digital de servicios integrados (RDSI ó ISDN, Integrated Services Digital Network).

La característica principal de ISDN es la posibilidad de soportar un amplio rango de aplicaciones vocales y no

vocales, a través de la misma red. Un elemento fundamental de la integración de servicios en ISDN es la capacidad de proveer un amplio rango de servicios a través de un conjunto limitado de tipos de conexión e interfaces usuario-red multipropósito.

ISDN se basa en tres elementos fundamentales:

- Transferencia digital de información: las señales análogas por naturaleza, como voz e imágenes, deben de digitalizarse para su transmisión.

- Señalización por canal común: este modo de señalización consiste en transmitir todas las señales de información de un grupo de circuitos a través de un solo canal. Esta separación en el transporte de la señalización y la información puede conducir a la construcción de una red específica de señalización. Para la señalización del usuario el CCITT ha definido una señalización por canal común conocido como protocolo D; un canal particular (canal D) transporta los mensajes entre el usuario y la red, completamente independiente a los canales de información.

- Interface usuario-red multipropósito: ISDN provee una interface multipropósito para todos los servicios; toda la información tanto señalización, voz, datos e imágenes se presentan a la interface en la forma de un múltiplex digital. Esta multiplexación provee servicios simultáneamente y por ejemplo, es posible tener una

comunicación telefónica mientras se recibe un fax y se accesa videotexto.

La construcción de la red ISDN debe ser gradual y en un principio se utilizará ISDN de banda estrecha (narrow band) con una velocidad de 2 Mb/s en la interface del usuario, y posteriormente se implementará ISDN de banda ancha (broadband) con mayores velocidades de transmisión.

ISDN de banda estrecha podrá soportar voz, datos, imágenes fijas e imágenes animadas con poca resolución.

Como un sólo equipo manejará varios tipos de servicios, se debe lograr una standardización. Esta standardización se hace necesaria para poder utilizar terminales y equipos de diferentes proveedores sin utilizar adaptadores y lograr un servicio sin importar la conexión a través de la red.

Los diseñadores de ISDN han hecho uso del principio de arquitectura abierta (OSI, Open System Interconnection) del ISO (International Organization for Standardization).

El ISO ha dividido las funciones relacionadas a los servicios de comunicación en subconjuntos o "capas funcionales". En la Recomendación X.200 del CCITT se presenta un modelo de 7 capas: 3 capas inferiores y 4 capas superiores. Las capas inferiores se refieren a las funciones necesarias para asegurar la transferencia de información entre dos terminales a través de la red, estas capas son:

- Capa 1 (Capa Física o Physical Layer): maneja los

aspectos físicos de la conexión de las terminales a las líneas de comunicación; maneja las interfaces mecánicas y eléctricas.

- Capa 2 (Capa de Enlace de Datos o Data Link Layer): corresponde a la transferencia de información en las líneas de comunicación y puede proporcionar protección contra errores de transmisión.

- Capa 3 (Capa de Red o Network Layer): asegura el establecimiento y fin de la comunicación y el enrutamiento a través de la red.

Las capas superiores se refieren a funciones específicas de las aplicaciones disponibles a los usuarios. Estas funciones pueden ser manejadas por las terminales, el equipo terminal o la red. Estas capas son:

- Capa 4 (Capa de Transporte ó Transport Layer): Provee monitoreo de la transferencia de información a través de la red (direccionamiento, sincronización, control de errores y control de flujo).

- Capa 5 (Capa de Sesión o Session Layer): define la organización de los intercambios y la estructura del diálogo entre aplicaciones.

- Capa 6 (Capa de Presentación o Presentation Layer): define la sintaxis del intercambio de información.

- Capa 7 (Capa de Aplicación ó Application Layer): contiene los mecanismos comunes que pueden implementarse

para varios servicios.

En la figura A.1 se muestra este modelo a nivel terminal, además se presentan tres conceptos esenciales:

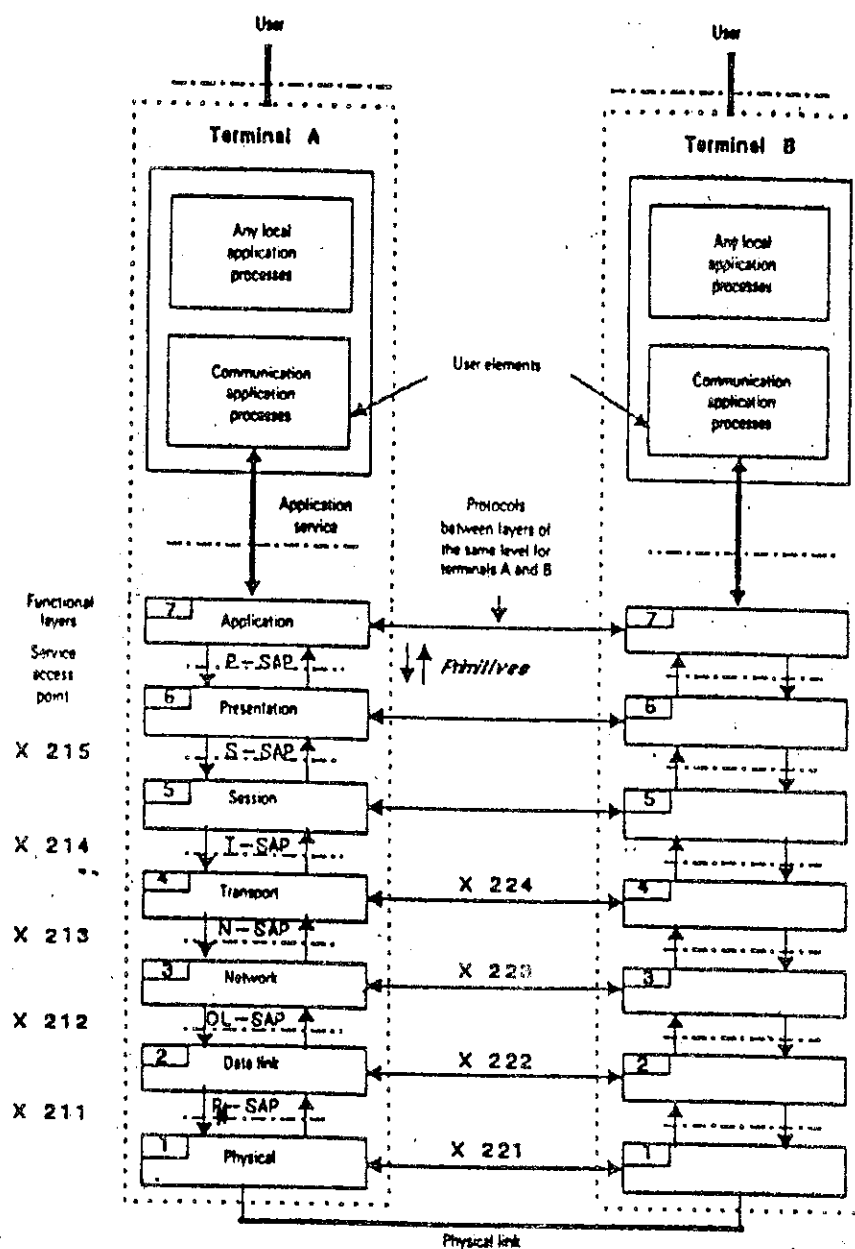


Figura A.1 Modelo OSI

- Puntos de acceso de servicio (SAP ó Service Access Points): definen la interface entre capas adyacentes en el mismo equipo.

- Primitivos: constituyen la base para el diálogo entre capas adyacentes del mismo equipo.

- Protocolos: son reglas que define el diálogo entre capas del mismo nivel, entre dos terminales.

La figura A.2 presenta los diferentes grupos funcionales del ISDN. El CCITT denomina "puntos de referencia" a las fronteras entre estos grupos:

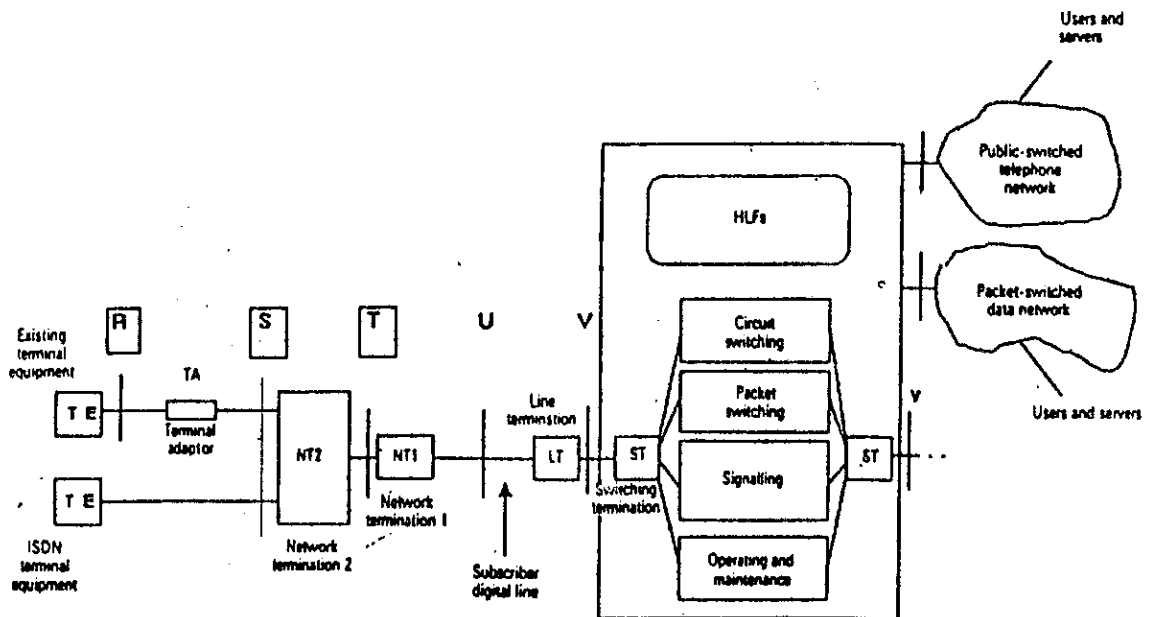


Figura A.2 Grupos funcionales y puntos de referencia para ISDN.

- Equipo terminal (TE, Terminal Equipment): accesa a la red en el punto S si se trata de terminales ISDN; si se trata de terminales existentes que satisfacen las recomendaciones de las series V y X del CCITT se necesita un adaptador R/S.

- Terminación de red 2 (NT2, Network Terminal 2): su punto de referencia es S hacia el lado terminal y T en el lado de la red. El punto T es en muchos casos el límite entre los dominios públicos y privados. NT2 controla el tráfico interno del usuario y controla el acceso a la red pública.

- Terminación de red 1 (NT1) en el lado del usuario y terminación de línea (LT) en el lado de la central: realizan la protección, monitoreo de la calidad de transmisión, etc. El interface entre ellos es el punto U.

- Señalización, conmutación y funciones de alto nivel (HLF, High Level Functions) de la red ISDN: el punto V es la interface entre la transmisión (LT) y la central (E).

Además, la red ISDN provee acceso a las redes existentes: red telefónica, red de conmutación de paquetes, red de telex, etc.

Según el desarrollo esperado hacia la red ISDN, se comenzará proporcionando al usuario un "acceso básico" a 144 Kb/s y para instalaciones de mayor capacidad se proporcionará el "acceso primario" a 1984 Kb/s ó 1536 Kb/s;

ambos proporcionando dos tipos de canales:

- Canal B a una velocidad de 64 Kb/s: este valor se deriva de la codificación de la voz en FCM, este canal puede utilizarse para voz, datos, fax, videotex, etc.

- Canal D a una velocidad de 16 Kb/s en acceso básico y a 64 Kb/s en acceso primario: utilizado para la señalización de la red del usuario; pero también planificada para transmisión de datos a baja capacidad o de naturaleza esporádica.

El acceso básico será del tipo 2B+D y el acceso primario de 30B+D (23B+D).

A. Servicios ISDN

Los servicios de comunicaciones proporcionados por una red ISDN son: servicios de transporte (bearer services), teleservicios y servicios suplementarios.

Un servicio de transporte es un servicio de transferencia de información limitado a las capas bajas del modelo OSI; se encarga de servicios de transferencia de información a diferentes velocidades y a diferentes niveles de calidad. El servicio de transporte se ofrece en la interface hacia la red (puntos S o T).

Un servicio de transporte se define por una serie de atributos. Estos atributos se dividen en tres categorías:

- Atributos de transferencia de información: modo de transmisión (circuito, paquete), velocidad de transferencia,

capacidad de transferencia, estructura, modo de establecimiento, configuración y simetría.

- Atributos de acceso: canal de acceso y velocidad, protocolo de señalización, protocolo de transferencia de datos.

- Atributos generales: servicios suplementarios, calidad de servicio, estructura de tarificación, etc.

El CCITT ha definido una estructura jerárquica para los servicios de transporte:

- Esenciales (E): debiendo introducirse lo antes posible.

- Adicionales (A): pueden ser provistos por ciertas redes ISDN.

- Servicios que requieren más estudio (F).

Los siguientes servicios se han clasificado como esenciales:

- Modo de circuito a 64 Kb/s con información estructurada a 8 KHz, se designará como circuito conmutado con canal B transparente (SCTB).

- Servicios en modo de circuito a 64 Kb/s estructurados a 8 KHz y para transmisión de voz y datos con canal B no transparente (SCNTB).

- Circuito para llamada virtual y para llamada permanente, en el canal B (VCB) o en el canal D (VCD).

Se observa que se puede acceder a la red telefónica con

SCNTB, a la red conmutada a una velocidad de 64 Kb/s por SCTB y a la red de conmutación de paquetes por VCB ó VCD.

En los teleservicios es necesario garantizar la compatibilidad de las terminales; deben tener los mismos protocolos para las 7 capas del modelo OSI. Un teleservicio se ofrece en la interface del usuario no más lejos de los puntos S/T.

ISDN puede proporcionar un teleservicio a través de 4 configuraciones:

- Conectando 2 terminales ISDN teniendo estas terminales las funciones de alto nivel ó HLF (capas superiores del OSI).

- Conectando una terminal ISDN a un HLF dentro de la red ISDN.

- Conectando 2 terminales ISDN que operan un teleservicio utilizando diferentes técnicas a través de un HLF localizado dentro de la red y que realiza la conversión de protocolos.

- Conectando una terminal a un HLF localizado en una red dedicada.

Los atributos que definen a un teleservicio se dividen en 3 grupos:

- Los atributos de las capas inferiores que definen la transferencia de información (como en los servicios de transporte).

- Los atributos que definen el tipo de información a transmitirse y los protocolos.

- Los atributos generales, como los servicios suplementarios que pueden agregarse.

Los teleservicios "más próximos" a implementarse son: telefonía, teletexto, fax, y mezcla de teletexto y videotexto.

Un servicio suplementario es un atributo adicional que complementa o modifica un servicio de telecomunicaciones. Los servicios suplementarios pueden clasificarse dependiendo del área de aplicación: afectando sólo una categoría de servicio de transporte, sólo un teleservicio o todos los servicios de telecomunicaciones.

Los servicios suplementarios pueden localizarse en una sola entidad o pueden distribuirse entre varias entidades. Entre los atributos utilizados para definir un servicio suplementario se encuentran: la entidad responsable de activar el servicio suplementario; los parámetros necesarios para implementarlo; el modo de activación; el área de aplicación; etc.

En los servicios suplementarios prioritarios a implementarse se encuentran: completación de llamadas a usuarios ocupados, llamadas con tarjeta de crédito, conferencia de n usuarios, transferencia de llamadas, transmisión de identificación con la llamada, etc.

Los mecanismos de direccionamiento de ISDN se cubren en las Recomendaciones I.330 e I.331 del CCITT. La dirección ISDN contiene los siguientes elementos: código de país, código de destino nacional, número de usuario ISDN, subdirección ISDN (ver figura A.3). La subdirección no es parte del número ISDN y se transmite transparente entre el usuario que llama y el llamado.

CCITT Recommendations: I.330 (E.164)-I.331

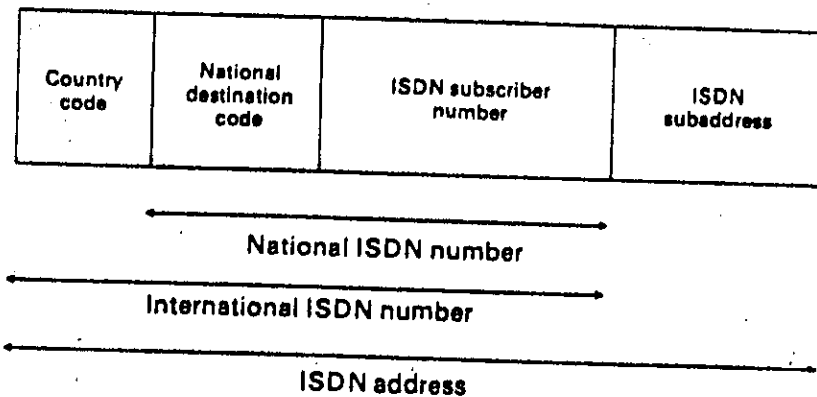


Figura A.3 Dirección ISDN

Al tener un usuario varias terminales que proporcionan diferentes servicios, el mecanismo de direccionamiento puede ser:

- Utilizar el número ISDN, por ejemplo, número de teléfono 2492, número de fax 2345, etc.
- Utilizar la dirección completa ISDN incluyendo el campo de subdireccionamiento, esta opción sólo puede emplearse entre centrales ISDN.

B. Interface de usuario

El objetivo principal de la red ISDN es permitir la implementación de un gran número de aplicaciones utilizando un número limitado de interfaces de usuario.

La configuración de referencia para el interface de usuario-red se muestra en la figura A.4:

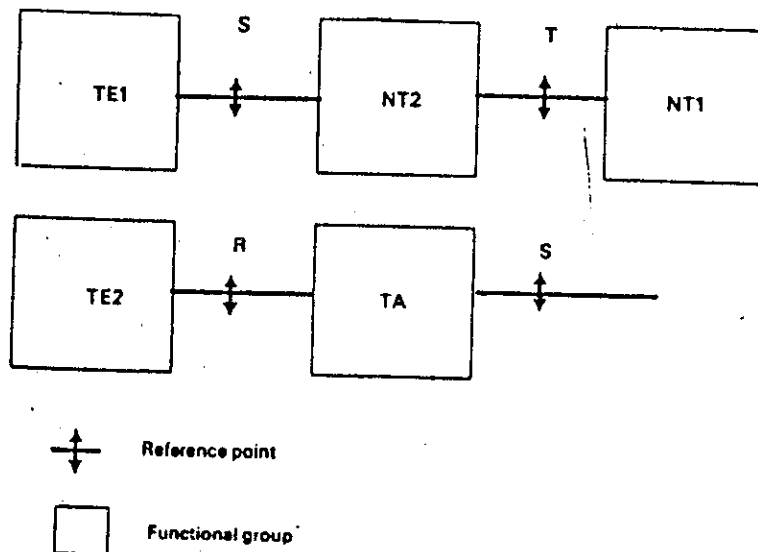


Figura A.4 Configuraciones de Referencia.

- Terminal de red 1 (NT1): maneja la capa 1 en el lado de la red y el sistema de transmisión de la línea del usuario.

- Terminal de red 2 (NT2): realiza las funciones de las capas 1, 2 y 3 de la interface usuario-red.

- Equipo Terminal: TE1 realiza las funciones de las capas 1, 2 y 3 del lado del usuario, además de las funciones de las aplicaciones que realiza; TE2 realiza las funciones de la aplicación además de las funciones de la interface no

ISDN.

- Adaptador de terminal (TA): adapta una terminal tipo 2 (TE2) a la red ISDN.

Los grupos funcionales anteriores no se localizan necesariamente en equipos físicos separados, y las interfaces que los separan no son siempre puntos físicos específicos. Varios grupos funcionales pueden realizarse en un único equipo.

Es sumamente importante la definición de la interface S, ya que es la conexión del terminal ISDN a la red. Los requisitos fundamentales que debe cumplir son:

- Posibilidad de colocar varias terminales en paralelo de una manera pasiva (sin electrónica en el conector).

- Distancia entre terminales acorde con las necesidades de las instalaciones de usuario existentes.

- Proporcionar energía a las terminales a través de la interface.

- Utilizar lo más posible el cableado existente.

Para la transmisión de datos en acceso básico se definen:

- Canales B1 y B2, bidireccionales a 64 Kb/s, definidos en modo de circuito, es decir, el canal se define para una única terminal mientras dure la comunicación.

- Canal D a 16 Kb/s bidireccional trabajando en modo de mensaje (basado en HDLC), compartido por todas las

terminales en una configuración multipunto.

Para el manejo de la interface se tiene:

- Función de resolución del problema de contención: se controla por el bit E, unidireccional de la red (NT) al equipo terminal (TE) a 16 Kb/s, transportando el eco del canal D a las terminales.

- Función de sincronización: las terminales se sincronizan en términos de bit, trama y multitrama. El reloj de bitio (192 Kb/s) se transmite en los datos. El reloj de trama (4 KHz) se transmite por medio de dos pares de bitios: F/L y Fa/N (a 4Kb/s cada bit). El reloj de multitrama se transmite en el bit M (4 Kb/s).

- Función de activación-desactivación: minimiza el consumo de energía al desconectar NT y TE al no existir transmisión. Es controlado por el bit A (4 Kb/s).

- Alimentación: permite transmitir energía a través de la interface; generalmente de NT a TE.

El bit S (4Kb/s) se define para funciones futuras.

Como se observa, la velocidad efectiva de transmisión ($2B+D = 144$ Kb/s) es inferior a la velocidad real de transmisión (192 Kb/s).

En la figura A.5 se muestra la estructura de trama para el acceso básico.

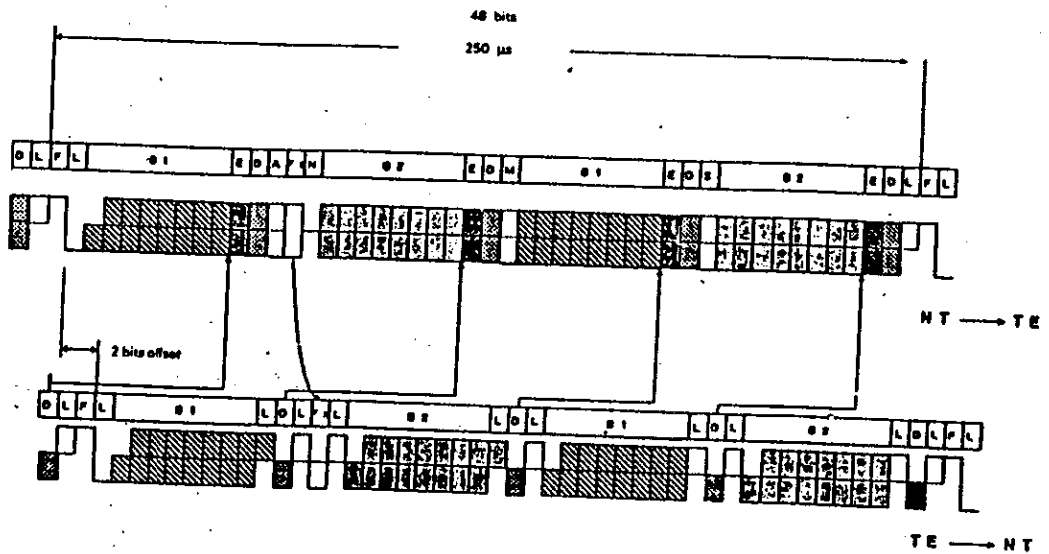


Figura A.5 Estructura de trama

En una configuración multipunto es necesario un mecanismo para asegurar que el acceso a un recurso se realiza de manera que las diferentes unidades conectadas puedan transmitir información en su turno y que dicha información no sea alterada por los intentos de acceso de las otras unidades.

Los accesos a los canales B se controlan por medio de la señalización del canal D. El problema resulta en el control del acceso del canal D. El mecanismo adoptado se llama CSMA-CR (Carrier Sense Multiple Access - Collision Resolution) y se basa en cuatro características:

- El canal D opera en la capa 2 de acuerdo al protocolo

HDLC.

- Las terminales pueden monitorear el canal D del TE al NT debido al canal de eco (bit E).

- El bus realiza una operación lógica AND de la información transmitida por las diferentes terminales.

- Una terminal inactiva transmite unos (representados por ausencia de señal).

El conflicto se soluciona de la siguiente manera:

Antes de transmitir una trama HDLC la terminal debe verificar que el canal D está libre. De acuerdo al protocolo HDLC el canal D está libre si existen ocho unos seguidos. Durante la transmisión de la trama HDLC la terminal debe comparar los datos transmitidos con el eco; si detecta una diferencia, detiene la transmisión.

Además, se han normalizado otros parámetros, por ejemplo, la forma del pulso en la terminal a 50 ohms o 400 ohms; parámetros eléctricos y físicos del medio de transmisión.

La Recomendación I.430 del CCITT no especifica el tipo de cable de conexión; pero el conector adoptado se especifica en el standard ISO DIS 8877 (Figura A.6).

Para el acceso primario se definen:

- Canales B: como en el acceso básico (64 Kb/s, bidireccional, en modo de circuito).

- Canales H0: bidireccionales a 384 Kb/s, en modo de

circuito.

- Canal H1: bidireccional a 1536 o 1920 Kb/s.
- Canal D: a 64 Kb/s en modo de paquete y soportando la señalización de los canales anteriores.

Los canales anteriores no pueden presentarse todos al mismo tiempo.

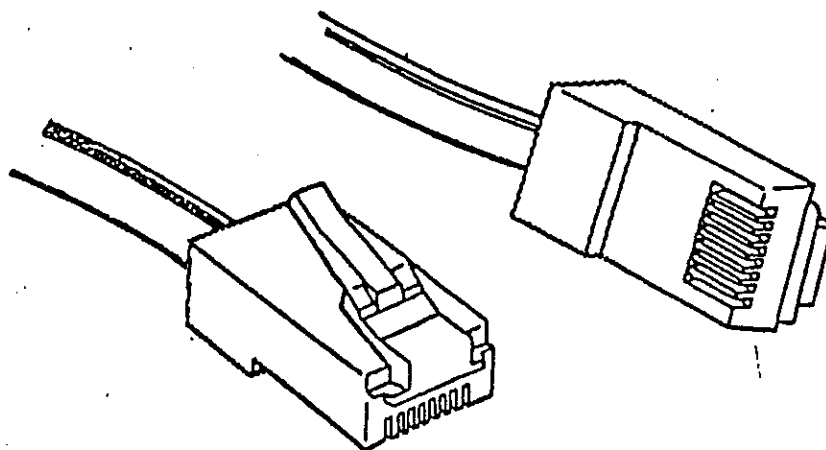


Figura A.6 Conector ISDN

Para controlar la interface se define:

- Sincronización de bit: se incluye en los datos (2048 Kb/s).
- Alineación de trama: provee la multiplexación/demultiplexación de los canales.
- Señal de 8 KHz: permite organizar los datos en octetos.
- Mantenimiento: bajo estudio del CCITT.

La trama es de 125 μ s, compuesto por 32 intervalos de

tiempo (TS) de 8 bits. La alineación de trama se compone de los bits 0011011, ocupando los bits 2 a 8 del TSO de cada trama par. El TSO se utiliza para alineación, el TS16 se utiliza para el canal D y los demás TS para los demás canales (D, HO, H1).

C. Protocolos

Debido a la digitalización de las líneas de abonado es posible utilizar CCS (señalización por canal común) hacia el usuario, la CCS se transmite por el canal D.

En la interface usuario-red se debe distinguir entre la señalización entre usuarios y la señalización relacionada con el mantenimiento y operación del equipo. Por lo anterior se obtiene un modelo tridimensional (figura A.7) con los siguientes planos:

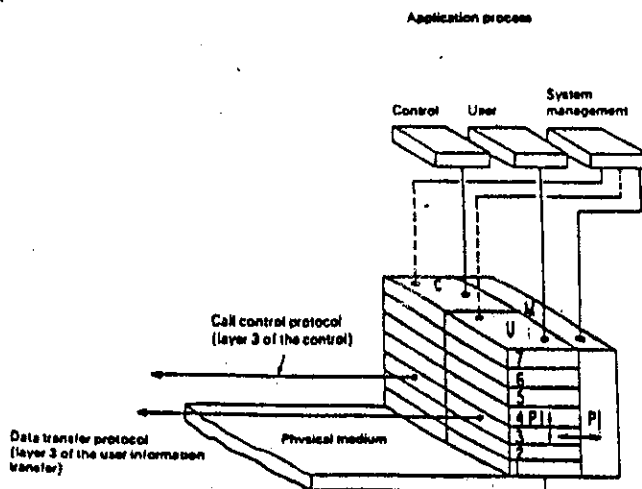


Figura A.7. Interacciones entre las entidades de la interface usuario-red
 C: Plano de control; U: plano de usuario; M: plano de manejo;
 Pi: primitivas entre planos adyacentes; Pj: primitivas de -
 servicio de manejo.

- Plano de control C: se organiza en 7 capas, concerniente a la señalización en el canal D y cubre los protocolos de control.

- Plano de usuario U: organizado en 7 capas, contiene los protocolos necesarios para el intercambio de información relacionada con las aplicaciones.

- Plano de manejo M: no se organiza en capas y concierne a funciones de operación locales en NT2 y las terminales.

Los planos C y U no pueden comunicarse entre sí y se intercomunican al plano de manejo por medio de primitivas de servicio.

La capa de enlace de datos implementa el protocolo de acceso en el canal D (LAP D, link access protocol) que permite el intercambio de tramas entre las entidades de la capa 2 situados a ambos lados de la interface S o T.

LAP D permite la demarcación de las tramas, multiplexación de varios enlaces de datos en el mismo canal, mantenimiento de las tramas en secuencia, detección de errores, etc.

El intercambio de información puede realizarse en dos modos:

- Modo de transferencia con reconocimiento: permite el intercambio de tramas de información numeradas (I) con corrección y detección de errores.

- Modo de no conexión: transfiere tramas no numeradas (UI) sin corrección de errores ni flujo de control.

La trama LAP D tiene los siguientes campos: bandera, dirección, control, información, chequeo y bandera. LAP D es una extensión de LAP B especificada en la Recomendación X.25; pero con una serie de simplificaciones para manejar errores.

El protocolo de control de llamada, denominado protocolo D, forma la capa de red en el plano de control, y realiza las siguientes funciones: Controla las primitivas de servicio en la interface de capas adyacentes; maneja los mensajes de la capa 3 y la comunicación con las funciones de control de llamada y manejo de recursos; controla los tiempos en el procesamiento de llamadas; maneja los recursos necesarios (canales, circuitos virtuales, etc.); monitorea y provee los servicios básicos solicitados por el usuario.

El protocolo D provee el intercambio de información usuario-usuario: en los mensajes de control de llamada, durante la presentación de la llamada y durante la comunicación.

Un usuario ISDN puede continuar utilizando su terminal y modem para acceder a la red pública de conmutación de paquetes (PSPDN) en modo analógico o con conexión totalmente digital. También puede utilizar el servicio de transporte de circuito virtual, utilizando el canal B o el canal D.

D. Terminales

Al inicio de la comercialización de los servicios ISDN se utilizarán las terminales existentes, por lo que deberá proveerse adaptadores de terminal (TA). Estos adaptadores deben cumplir los siguientes requisitos: Proveer al usuario con los servicios y aplicaciones existentes con un desempeño al menos igual a aquellos producidos por la red anterior, y proveer las ventajas de la red ISDN.

Existirán adaptadores para accesos clásicos: interface Z para telefonía analógica, interface X.21 para comunicación de datos en redes de conmutación de circuitos, interface X.25 para acceso en redes públicas de conmutación de paquetes, V.24 para comunicación de computadoras en redes privadas, etc.

Entre las desventajas del TA se cuentan: dificultad en optimizar el hardware e imposibilidad de poder explotar todo el potencial de la red ISDN. Por lo anterior, se desarrollarán terminales ISDN para los servicios existentes (telefonía, fax, telex, etc) y para nuevos servicios (ej. transmisión de imágenes).

El adaptador para una interface analógica clásica y una interface ISDN se denomina A/S (Figura A.8) y permite la conexión de varias terminales trabajando en línea con la red telefónica y terminales usando líneas de voz especiales a través de la interface S a la red ISDN:

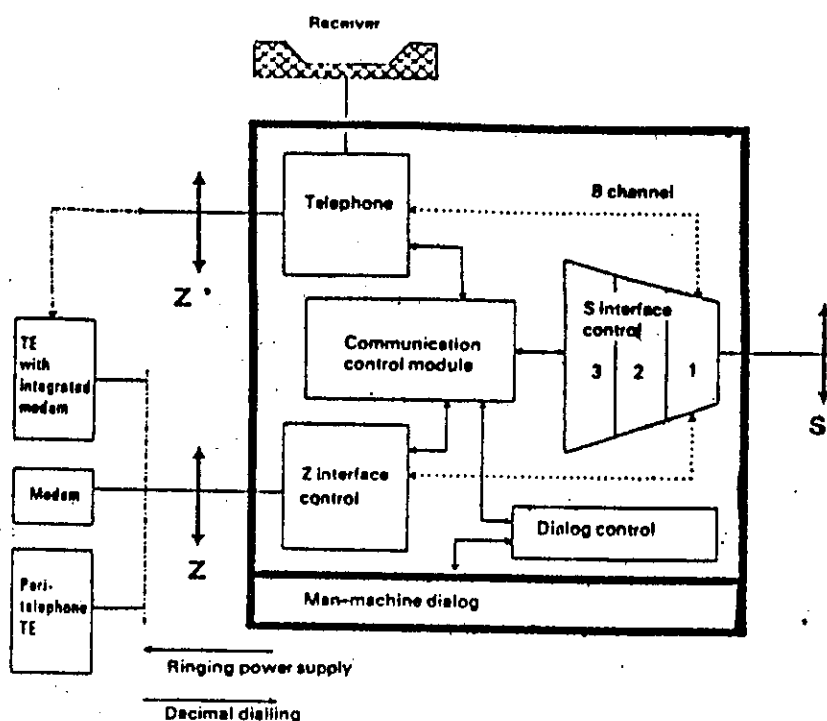


Figura A.8 Diagrama funcional del adaptador A/S.

El adaptador A/S tiene cinco módulos:

- Módulo de control de interface S: controla las capas 1 a 3 de la interface ISDN.
- Módulo de control de la interface Z: provee una interface analógica a 2 hilos; codifica y decodifica la señal de voz utilizando la ley A o u.
- Módulo de control de comunicaciones: monitorea el estado de la terminal y decide las acciones a tomar dependiendo de la demanda de los módulos periféricos.
- Módulo telefónico: contiene las funciones básicas de telefonía digital como generación de tonos, timbrado, etc.
- Módulo de diálogo: supervisa el intercambio con el

usuario durante el establecimiento y terminación de la comunicación.

La terminal telefónica ISDN tiene varias mejoras con respecto a la terminal telefónica existente: Como la interface S provee dos canales B y uno D, se puede usar un canal B para telefonía y otro para aplicaciones adicionales.

El uso del protocolo D permite instalaciones multiterminal; control de varias comunicaciones simultáneas, intercambio adicional de información, etc.

El servicio básico que proporciona una terminal telefónica ISDN consiste en: establecimiento de una comunicación bidireccional usando un canal B en modo SCNTB; codificación y decodificación de la voz de acuerdo a la Recomendación G.71x; alimentación de la terminal a través de la interface S.

Los servicios suplementarios que puede proveer se dividen en:

- Servicios suplementarios para ayudar a simplificar y mejorar el uso de la terminal durante la comunicación.
- Servicios suplementarios para mejorar la comunicación cuando existe un período de espera.
- Servicios suplementarios para situaciones cuando el usuario está ausente.
- Servicios suplementarios para transferencia de información relacionada a la comunicación.

- Servicios suplementarios adicionales.

Para la transmisión de documentos las terminales ISDN serán capaces de proporcionar servicios de teletexto y facsímil. El servicio de teletexto es el servicio de correo electrónico usado en la actualidad y provee una velocidad 20 veces mayor que el telex, un alfabeto completo y servicio automático de transferencia de información de memoria a memoria.

La red ISDN será capaz de transmitir imágenes, en un principio a 64 Kb/s, pero esencialmente limitado a imágenes fijas con códigos de reducción de velocidad.

La red ISDN podrá soportar terminales de acción remota, soportando aplicaciones como alarma remota, control remoto y telemetría.

Los servicios proporcionados por la red ISDN se aplican a diferentes medios: voz, datos o imágenes; por lo anterior pueden existir aplicaciones que combinen varios servicios. Estas aplicaciones se denominan multimedia.

En el diseño de terminales multimedia, existe el riesgo de tener tantas terminales diferentes como combinaciones de aplicaciones. Una solución consiste en construir terminales a partir de microcomputadoras de propósito general, utilizando el principio de tarjetas de expansión. En este caso, la microcomputadora se equipa con una tarjeta de comunicaciones ISDN con interface S. Esta tarjeta se encarga

de manejar los protocolos de la red (capas 1-3 del canal y permite la operación de los servicios de transporte servicios suplementarios. Para el caso de teleservicio (capas superiores), pueden utilizarse módulos de software combinaciones de software y tarjetas de expansión.

Actualmente existen centrales ISDN en varios países del mundo, con varios miles de usuarios.

En nuestro país se está estudiando la estrategia seguir para la transición hacia ISDN.