Universidad de la República
Escuela Universitaria de Bibliotecología y Ciencias Afines
"Ing. Federico E. Capurro"

Conceptos de Redes de Computadoras

Ing. Angel Caffa[1], MSc. , MBA

Escuela Universitaria de Bibliotecología y Ciencias Afines

Universidad de la República

Sede Central: Emilio Frugoni 1427, 11200 Montevideo, Uruguay / Telefaxes (5982) 4005810 / 4085576
Teléfonos (5982) 4010788 / 4011423
Anexo: J. Rodó 1839, Planta Alta, 11200 Montevideo, Uruguay / Teléfono: (5982) 4082925
Bedelía: En el Anexo / Telefax: (5982) 4020297

Correo eléctronico: eubca@adinet.com.uy / Sitio Web: http://www.eubca.edu.uy

[1] angel.fing.edu.uy

Agradecimientos

El autor agradece a los estudiantes del curso de Redes y Sistemas de Información que leyeron versiones preliminares del libro y aportaron nuevos puntos de vista.

Índice

Introducción

Este manual se ha escrito con el objetivo de recoger los conceptos expuestos en el curso de actualización “Conceptos de Redes para Archivólogos y Bibliotecólogos”. Sin embargo, también se espera sirva de apoyo al capítulo de redes del curso “Redes y Sistemas de Información” de la Licenciatura en Bibliotecología en EUBCA.

A juicio del autor, definir el área de influencia del Licenciado en Bibliotecología centrada en las bibliotecas es una visión limitada que restringe enormemente el futuro posicionamiento de estos profesionales en el mercado laboral. Parece más adecuado pensar en términos de *especialistas en manejo de información*. Aceptadas las consideraciones anteriores, las materias relacionadas con el área informática deberían formar una componente fundamental dentro de la currícula, y dentro de las mismas, la presentación de conceptos técnicos y la profundización en los mismos se presenta como inevitable.

El objetivo de estos cursos de redes es brindar una formación básica en el tema, presentando los conceptos fundamentales, pero sin descuidar algunos detalles técnicos para así dar un enfoque práctico y cercano a la realidad. Se busca aportar elementos que permitan al futuro profesional interactuar en un lenguaje común con especialistas en informática, y poder participar de decisiones que involucren diseño y/o evolución de redes de computadoras. Este manual prioriza de algún modo lo conceptual sobre lo técnico, pero no evita los aspectos técnicos.

La parte técnica del curso está basada en el libro “Redes de Computadoras” de Andrew Tanembaum, y el enfoque sigue el mismo estilo que el curso “Comunicación de Datos” de la carrera Ingeniería en Computación de la Facultad de Ingeniería de la Universidad de la República. Se ha intentado incorporar también la experiencia recogida por el autor en el dictado y coordinación de los cursos “Redes I”, “Redes II” y “Tópicos de Computación” de las carreras de “Ingeniería en Informática” del Universitario Autónomo del Sur entre los años 1996 y 2000.

La discusión de Internet desde el punto de vista del usuario está influenciada por el enfoque que el Cr. Roberto de Luca utilizó en la versión 2000 de su curso “Modelos de Negocios en eBusiness” en el Master en Administración de Empresas en la Facultad de Administración y Ciencias Sociales de Universidad ORT.

También se ha buscado capitalizar el intercambio con los estudiantes surgido del dictado del curso “Redes y Sistemas de Información” en EUBCA en los años 1999 y 2000, así como en el curso de actualización mencionado antes.

El capítulo 1 presenta una breve puesta al día en conceptos de informática. En el capítulo 2 se discuten conceptos generales relacionados con las redes de computadoras: conceptos, definiciones, clasificaciones, etc. El capítulo 3 introduce algunos modelos de redes y su utilidad. Los capítulos 4 a 10 analizan en profundidad el modelo OSI, estudiando cada una de las capas. El capítulo 11 trata la familia de protocolos TCP/IP, e introduce la red Internet desde el punto de vista técnico. En el capítulo 12 se discute el fenómeno Internet desde el punto de vista del usuario, con especial énfasis en los diferentes modelos de negocios que sustentan buena parte de los sitios de Internet.

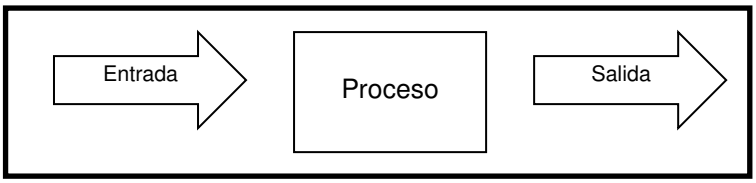
Los ejercicios buscan hacer este manual independiente (en la medida de lo posible) de las tecnologías actuales, entonces todos los datos relativos al estado actual del arte son dejados para investigar en forma guiada a través de los ejercicios. En todo caso no se debe perder de vista la rapidez con que evolucionan los productos y herramientas informáticas. Otros ejercicios buscan reafirmar conceptos o exponer problemáticas comunes.

Tenga claro el lector que este libro es un manual de un curso, y debería ser leído como tal. En particular, se recomienda investigar cada punto marcado como ejercicio.

Capítulo 1: Puesta al día en conceptos de informática

En la actualidad, el uso de las computadoras está ampliamente difundido, no sólo ya como herramienta programable de uso general, sino también como sistemas embebidos[2], resolviendo por ejemplo el tráfico de llamadas en la central telefónica de una empresa o controlando el flujo de combustible dentro de un auto.

En todos los casos, la función de cualquier computadora se puede abstraer en el siguiente esquema:



Los componentes tangibles (lo que se puede tocar) dentro de un sistema informático reciben el nombre de *hardware* y los no tangibles (los programas) se denominan *software*.

Hardware

A continuación se presentan los elementos más comunes del hardware de un sistema informático, tomando como guía el esquema anterior.

Los dispositivos de entrada más comunes son:

- Teclado: recibe datos correspondientes a teclas presionadas. Actualmente hay dos tipos, según la ficha de conexión: DIN y mini-DIN (PS/2).
- Ratón: recibe órdenes de movimiento horizontal o vertical (todo movimiento complejo se puede descomponer en una sucesión de movimientos horizontales y verticales muy pequeños), y si alguno de sus botones fue presionado. Actualmente hay dos tipos, según la ficha de conexión: mini-DIN o serial DB9.
- Scanner: permite digitalizar imágenes. Hay dos tipos, de acuerdo a la conexión a la computadora: paralelo o USB.
- Lector de código de barras: cada código de barras está asociado a un número, que en general corresponde a codificaciones que permiten identificar objetos rápidamente. Por ejemplo en un POS (point of sale= punto de venta).
- Lápiz óptico: permite marcar puntos al indicarlos con un lápiz sobre la pantalla. Con la popularización de los scanners ha quedado reducido a aplicaciones específicas.
- Tableta digitalizadora: permite marcar puntos al indicarlos con un lápiz sobre una tableta. Con la popularización de los scanners ha quedado reducido a aplicaciones específicas.
- Joystick: control para juegos.
- Cámaras fotográficas y de filmación. Normalmente se comunican con el PC vía USB.
- Micrófono: para recoger datos de audio.

Existen otros, relacionados con aplicaciones más particulares.

Los dispositivos de salida más comunes son:

[2] Los sistemas embebidos son sistemas incluidos dentro de otros más complejos.

- Monitor: Permite desplegar datos en pantalla. Actualmente son color en casi todos los casos, y sus características técnicas (resolución, etc.) están determinadas en gran medida por las de la tarjeta gráfica del mismo. El monitor recibe de la computadora datos (concretamente de la tarjeta de video), y a veces (puede conectarse directamente) corriente. Algunos –llamados multimedia- contienen parlantes y/o micrófono incorporados, teniendo conexiones especiales destinadas a tales efectos.
- Impresora: Permite desplegar datos en papel. Algunos de los tipos más comunes:
    - Láser: utilizan la misma técnica de impresión que las fotocopiadoras: se distribuye un material especial (tóner) que luego se fija por calor. Son la alternativa más rápida.
    - Inyección de tinta: la alternativa más económica, lográndose además una razonable calidad y velocidad.
    - Matriz de puntos: un conjunto de agujas se configuran de acuerdo al carácter a imprimir y éste se imprime por impacto sobre una cinta entintada. Robustas y rápidas al imprimir texto.
    - Margarita: También de impacto, pero en este caso existe una "margarita" de caracteres similar a la de una máquina de escribir.
    - Térmica: Imprimen quemando un papel especial, que usualmente es caro.
- Parlantes: Los parlantes son dispositivos de salida de los datos correspondientes a audio. Su interfaz con la computadora es usualmente una tarjeta de sonido, que en algunos casos puede venir ya integrada a la motherboard.
- Impresoras de código de barras: caso particular de impresora.
- Plotter: impresora diseñada con el objetivo de imprimir gráficos y planos.

En cuanto a procesamiento, se puede desagregar el esquema en las siguientes partes:

- Unidad Central de Proceso (UCP): Es quien regula todas las operaciones de la computadora. Tiene dos partes: Unidad Aritmético Lógica (UAL) que es donde se ejecutan todas las operaciones con los datos, y la Unidad de Control (UC), que es quien controla el flujo de ejecución. El procesador tiene unas celdas de memoria llamadas registros, que son de rápido acceso y en las que coloca los operandos y donde obtiene la salida en cada una de sus operaciones. Cada procesador (o familia) tiene su propio lenguaje ensamblador en el que se pueden especificar sus operaciones a bajo nivel de abstracción. Su velocidad se mide usualmente en Mhz. Una computadora puede tener más de un procesador.
- Memoria RAM (o memoria principal): Random Access Memory, esta memoria es la que contiene datos y programa en ejecución. De contenido volátil, es decir, necesita de corriente para conservarlo. A veces la RAM se complementa con la memoria caché (memoria de rápido acceso ubicada entre la RAM y la UCP), donde se almacenan resultados intermedios para mayor velocidad de acceso.
- Memoria ROM: Read Only Memory, contienen los programas de arranque de la computadora. No pierde su contenido cuando se apaga la computadora.
- Placa madre: es la placa donde están todos los componentes de la computadora.
- Bus: Son cables por donde se intercambia información entre dos componentes de la computadora. Existen diferentes protocolos a utilizarse aquí según el tipo de dispositivo: IDE, SCSI, SCSI2 y según el tipo de conexión: PCI, PCMCIA, ISA, EISA.
- Memoria secundaria: En esta memoria se guardan datos en forma persistente (es decir que se conservan más allá del apagado de la computadora). Hay muchas alternativas:
    - Discos: Compuesto de un conjunto de superficies magnetizables que giran y son leídas por un conjunto de cabezas, los discos constituyen

hoy una parte fundamental de la computadora. Permiten acceso aleatorio a datos, y en ellos residen programas (incluido el sistema operativo) y datos.

- Diskettes: Discos de menor capacidad para poder portar datos. La unidad lecto-escritora se denomina diskettera.
- CDs:Compact Disks. Aunque lo más difundido son aún las lectoras, es posible grabarlos una o varias veces, según la tecnología a utilizar. Basados en lectura óptica.
- DVDs: Digital Video Disks. Tecnología similar a los CDs pero que permiten guardar más datos.
- Cintas: De acceso secuencial[3], guardan gran volumen de datos aunque la velocidad de acceso no es buena. Por lo barato de los cartuchos, se utilizan para backups.
- ZIP y Jazz: utilizan la misma tecnología que los diskettes, pero con más capacidad de almacenamiento de información.

La unidad básica de almacenamiento de información es el bit. Un bit puede valer 0 o 1. Un byte u octeto son 8 bits. 1024 (= 2^{10}) bytes son un Kb (Kilo Byte), 1024 Kb son un Mb (Mega Byte), 1024 Mb. son un Gb. (Giga Byte) y 1024 Gb. son un Tb. (Tera Byte).

Ejercicios

1.1) Explicar la diferencia entre dispositivo de acceso aleatorio y dispositivo de acceso secuencial. Dar ejemplos.

1.2) Explicar la utilidad de las cintas a la hora de hacer respaldos siendo un dispositivo mucho más lento que los discos.

1.3) Para cada alternativa de almacenamiento primario y secundario indicar capacidad según la disponibilidad actual en el mercado.

1.4) Buscar información sobre la arquitectura Von Neumann, y su concepto central de "programa y datos en una misma memoria". Notar que este concepto es central en la arquitectura de toda computadora.

1.5) Hacer un relevamiento de precios de todos los dispositivos mencionados.

1.6) Recomendar una configuración doméstica, sobre un presupuesto de U$S 1000.- incluidos todos los impuestos. Compararlo contra el PC más barato del mercado.

1.7) Investigar acerca de la configuración de un PC más barata del mercado. Describirla y criticarla.

1.8) Recomendar una configuración para una máquina de uso más exigente, sobre un presupuesto de a) U$S 1500.- b) U$S 2000.- y c) U$S 3000.- incluidos impuestos.

1.9) Averiguar uso y funcionamiento de las arquitecturas redundantes y discos RAID. Investigar en qué medida se pueden incluir en presupuestos según las categorías vistas en 1.8.

1.10) Comentar brevemente la evolución en capacidad/prestaciones de alternativas correspondientes a la respuesta de 1.6 desde 1985 a la fecha.

1.11) Averiguar las alternativas actuales de dispositivos USB (teclados, ratones, etc.). Buscar argumentos a favor del uso de este protocolo de conexión de acuerdo a los vendedores.

[3] Notar que las demás tecnologías mencionadas son de accedo aleatorio.

1.12) Buscar información comparativa de discos SCSI frente a los discos IDE.

1.13) Discutir ventajas y desventajas de las placas madre “todo on board”.

1.14) Buscar información de UPSs, estabilizadores de tensión y protectores de línea.

1.15) Averiguar la utilidad de los racks dentro de un centro de cómputos.

1.16) Relevar las medidas de seguridad en un centro de cómputos, frente a contingencias como incendios o robos.

1.17) Relevar la oferta de discos duros en base a parámetros como: capacidad, velocidad de rotación, etc. Investigar qué son las tecnologías de almacenamiento redundante.

1.18) Investigar qué significa que una computadora pueda manejar fuentes redundantes de alimentación.

1.19) Averiguar la diferencia entre una fuente AT y una ATX. Investigar los voltajes desde/hacia la fuente, y determinar en función de los mismos los componentes del PC en cuya manipulación puede existir riesgo de electrocución.

1.20) Averiguar qué son los discos “hot swap”.

Para profundizar aspectos de arquitectura de computadoras se recomienda el libro [Tan92].

Software

A continuación se discuten conceptos generales relacionados con el software de los sistemas informáticos.

Se denomina software a todo lo no tangible en un sistema informático, es decir, los programas. Hay varias clasificaciones para los programas, y aquí se utilizarán algunas de modo de dar una idea general.

Los programas de uso general se dirigen a solucionar problemas comunes a muchos usuarios, mientras que los programas de uso específico son usualmente programados a medida de modo de solucionar algún problema concreto en una organización.

También es útil considerar programas de alto y bajo nivel (de abstracción). Los primeros son más cercanos al usuario y los otros son más cercanos a la máquina.

Un tipo de software muy importante son los sistemas operativos. Un sistema operativo es un conjunto de programas que administra el acceso de los usuarios a los recursos de la computadora.

Los sistemas operativos se clasifican en monotarea/multitarea según permitan la ejecución de varios programas a la vez o no y en monousuario/multiusuario según permitan el acceso de más de un usuario (reconociendo derechos de acceso diferenciados) o no.

Algunos ejemplos de sistemas operativos:

- Unix: Linux, Solaris, HPUX, AIX, etc. son implementaciones particulares de este sistema operativo multiusuario multitarea.
- Windows 95/98/Me: Multitarea/Monousuario.
- Windows NT/2000/XP: Multitarea/Multiusuario.
- MS-DOS: Monotarea/Monousuario.

La seguridad de sistemas es un área de importancia cada vez mayor, ya que las computadoras abarcan un conjunto cada vez más amplio de actividades. La seguridad se puede medir en dos dimensiones: dejar acceder a los recursos del sistema solamente a los usuarios registrados y controlar niveles de acceso y manejo de datos para estos usuarios registrados. La elección del conjunto de sistemas operativos a utilizar es la base del esquema de seguridad de los sistemas de una organización.

Para profundizar en el área de sistemas operativos, se puede recurrir a [Dei90].

Otros programas de uso general importantes son (todos corren sobre un sistema operativo):

- Paquetes de automatización de tareas de oficina (Office, etc).
- Servidores de aplicaciones de red (correo, páginas, etc.)
- Manejadores de bases de datos.
- Etc.

Ejercicios

1.11) Completar la lista de sistemas operativos, resolviendo la clasificación según monotarea/multitarea y monousuario/multiusuario.

1.12) Completar la lista de programas de uso general.

1.13) En la lista de 1.12, comentar sobre qué sistema operativo corren los programas.

1.14) Comentar la relación entre que un sistema operativo sea monousuario/multiusuario y la seguridad del mismo.

1.15) Comentar la relación entre que un sistema operativo sea monousuario/multiusuario, la seguridad y la posibilidad de trabajar en red[4].

[4] Por el momento considerar la "posibilidad de trabajar en red" como la posibilidad de intercambiar información entre dos o más computadoras. Este término se formaliza en el capítulo siguiente.

Capítulo 2: Redes de computadoras: generalidades

Definición de red

Una red de computadoras es un conjunto de computadoras autónomas interconectadas. La palabra interconectadas hace referencia a que existe algún mecanismo que les permite intercambiar datos. Una computadora es autónoma si tiene CPU y memoria propias.

Cada computadora en la red se denomina host.

El concepto de autonomía manejado implica excluir de la red a las terminales tontas, ya que las mismas sólo gestionan entrada/salida.

Las redes permiten compartir recursos.

Notar que no aparece la palabra servidor en la definición de red propuesta.

Ejercicio

2.1) Buscar otras definiciones de red, en especial en los manuales de productos de Microsoft y Novell.

El modelo centralizado vs el modelo de computación en red

Históricamente, la primer alternativa para compartir recursos fueron los mainframes. Este modelo centralizado del sistema implica la existencia de una computadora central a la que se accede desde muchas terminales tontas. Las terminales tontas sólo hacen entrada/salida, todo el proceso ocurre en el mainframe.

Las ventajas de los mainframes son:

- Permiten compartir recursos
- Existe un solo punto de acceso a todos los recursos del sistema
- Se pueden resolver los respaldos desde un solo puesto de trabajo.

Sin embargo, al haber una única CPU y memoria (que son los recursos clave) se observa una competencia de los procesos por estos recursos clave.

El modelo de trabajo en red (que sucedió cronológicamente al modelo centralizado) permite compartir recursos conservando las ventajas marcadas en el modelo centralizado, pero además:

- No hay competencia por CPU y memoria.
- La inversión se puede escalar, es decir el sistema se adapta gradualmente a nuevas exigencias de cómputo.
- Mayor flexibilidad en cuanto a la disposición espacial/geográfica de los equipos, al no requerirse un centro de cómputos para albergar una computadora de gran porte.
- Mejor relación costo/performance.
- Permiten compartir recursos.
- Se pueden implementar mecanismos de confiabilidad basados en la duplicación de la información.
- Una red es un mecanismo poderoso de comunicación entre sitios remotos.

Las causas de la llegada de las redes, pueden resumirse en:

- El bajo costo del hardware populariza el uso de la computadora.
- Las nuevas tecnologías de comunicación de datos permiten una eficaz interconexión.
- En su comienzo, las computadoras se utilizaban para el tratamiento de la información. Actualmente las computadoras se utilizan para tratamiento de la información y comunicación de datos.

Ejercicio

2.2) Dar ejemplos correspondientes a cada una de las ventajas y desventajas anteriores.

Protocolos de comunicación

Un protocolo es un conjunto de reglas formales de comportamiento. Un protocolo de comunicación es un conjunto de reglas formales que rigen la comunicación de datos.

En comunicación de datos, a las reglas que forman los protocolos también se las llama protocolos.

Al estudiar las redes, se encuentran protocolos de comunicación asociados a la resolución de actividades asociadas a la comunicación entre computadoras.

Los protocolos se estudian agrupados por niveles o capas, que en general quedan determinados por el nivel de abstracción de los mismos. Se dice que los protocolos más cercanos a la computadora (como podría ser el que especifica las dimensiones de un conector) son de más bajo nivel de abstracción que los protocolos más cercanos al usuario (como podía ser el que especifica la forma en que se envían los correos desde el servidor al cliente de mail).

Cuando un conjunto de protocolos de igual o distintos niveles se utilizan juntos se denominan familia de protocolos.

El paradigma cliente/servidor

Al considerar el conjunto de computadoras que integran una red, surge la necesidad de aprovechar al máximo el potencial de cómputo que constituye ese conjunto de computadoras independientes pero interconectadas.

El problema pasa por optimizar el aprovechamiento de esta enorme capacidad de cómputo, pero minimizando el tráfico de datos en la red.

La idea principal detrás del paradigma cliente-servidor es la búsqueda de paralelizar al máximo las aplicaciones, teniendo en cuenta la heterogeneidad de plataformas y el estado actual de la tecnología de la programación en paralelo.

La solución es partir las aplicaciones en dos, desde las etapas tempranas de diseño. Esas dos partes se llaman cliente y servidor. La aplicación servidor se especializa en la aplicación en sí misma, y da servicios al cliente, ocupándose de los aspectos generales. La aplicación cliente resuelve las necesidades particulares de cada punto de trabajo, como ser interfaz, aspectos dependientes de la arquitectura, etc., y pide servicios al servidor.

El correo electrónico y la publicación de páginas de WWW son ejemplos de aplicaciones que siguen el paradigma cliente-servidor.

Las computadoras donde se ejecutan los módulos servidor y cliente reciben el nombre de máquina servidor y máquina cliente respectivamente.

Ejercicios

2.3) Explicar detalladamente el funcionamiento del paradigma cliente/servidor en los dos ejemplos anteriores (correo y páginas WWW).

2.4) Identificar otras aplicaciones cliente/servidor y comentar su funcionamiento.

Clasificación de las redes

A continuación se discuten algunos criterios de clasificación de las redes. En la realidad se encuentran modelos híbridos entre las diferentes categorías.

- Según la tecnología de transmisión:
    - Redes punto a punto: basadas en conexiones punto a punto, es decir, todas las conexiones son entre pares de máquinas.
    - Redes de difusión: basadas en conexiones multipunto (o canales o broadcasts), que implican varias máquinas que comparten un canal y por lo tanto el mismo debe ser arbitrado.
- Según el tamaño:
    - LAN: red de área local (usualmente dentro de los límites de un edificio). En general utilizan tecnología multipunto en sus conexiones.
    - MAN: red de área metropolitana.
    - WAN: red de área amplia. En general utilizan tecnología punto a punto en sus conexiones.
- Según su topología:
    - Anillo
    - Estrella
    - Completa
    - Arbol/Jerárquica
    - Canal
    - Mixta
    - Irregular

La topología de la red describe la disposición física de sus equipos.

Ejercicio

2.5) Dar ejemplos para cada uno de los casos de las clasificaciones anteriores.

2.6) Elegir una red del mundo real. Clasificarla en términos de las categorías anteriores.

2.7) Completar dibujos de cada una de las topologías.

Grafo de una red

La Teoría de Grafos puede ser de gran ayuda a la hora de modelar redes.

Esto significa que dada una realidad (una red en este caso), se puede construir un modelo de la misma (utilizando Teoría de Grafos en este caso). Este modelo captura un conjunto de detalles que se ha decidido son relevantes, y deja de lado un conjunto de detalles que resultan irrelevantes para el estudio que se esté desarrollando. Así, el modelo es más simple de manipular. En el modelo se resuelve el problema que se está considerando, y finalmente se traduce dicha solución en términos de la realidad inicial. Una ventaja de esta metodología de trabajo es que en el modelo, muchos problemas se presentan de forma estándar, y ya han sido resueltos. Este es el caso de la Teoría de Grafos como herramienta para modelar.

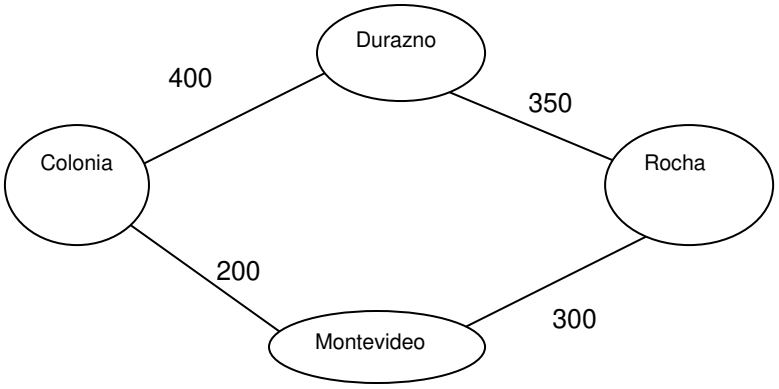
Un grafo es un par <V,A> donde V es un conjunto de vértices y A un conjunto de aristas.

Una arista es un par de vértices, por lo que A es un subconjunto de V X V (que contiene todos los pares posibles de vértices).

En los grafos orientados, importa el orden de los vértices que componen una arista, y se dice que las aristas tienen dirección. Cuando el orden no importa, el grafo es no orientado.

A veces se agrega una función f : A → R, siendo R el conjunto de los Reales, La función f se dice función de costos, y refleja el hecho de que "transitar" por una arista tiene un costo asociado.

Por ejemplo, si se quieren estudiar las rutas para un reparto entre las capitales de Montevideo, Colonia, Durazno y Rocha, un buen modelo en términos de grafos podría ser:



Algunos comentarios respecto al modelo anterior:

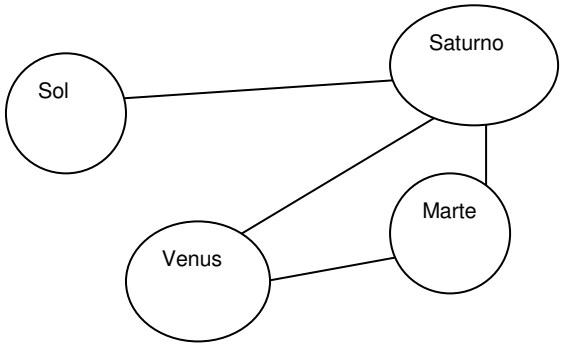
- La función de costos captura de alguna forma la dificultad de trasladarse de un punto a otro (distancia, estado de la ruta, probabilidad de atascos, etc.)
- De acuerdo al modelo, no se utiliza la ruta directa Colonia-Rocha.
- Las otras ciudades no son relevantes para el análisis, desde el momento que se decidió no incluirlas en el modelo.
- Se utiliza un grafo no orientado, lo que implica que no hay diferencias en la dirección de la ruta.

Para definir un grafo hay que:

- Dar el conjunto de vértices
- Dar una regla para la determinar las aristas
- Especificar la función de costos si la hubiera.

El grafo de una red es un grafo (en general) no orientado, y (en general) sin función de costos, cuyos vértices son los hosts de la red y existe arista entre dos vértices siempre que exista comunicación directa entre los dos hosts asociados a dichos vértices.

Así, el siguiente es un ejemplo de grafo de red:



Los vértices del grafo son etiquetados usualmente con los nombres de los hosts. A veces se colocan etiquetas en las aristas, por ejemplo cuando es relevante diferenciar los tipo de conexiones.

Los vértices entre los que existe arista se dicen adyacentes. Las computadoras entre las que existe conexión directa en la red se dicen adyacentes.

Un camino en el grafo es una secuencia de aristas. Los vértices extremos de algún camino en el grafo se dice están conectados, y consecuentemente, las máquinas asociadas a vértices conectados en el grafo de la red se dicen conectadas. Las máquinas conectadas son las que pueden intercambiar información. En el ejemplo, Venus y Marte son adyacentes, Venus y Sol están conectadas.

Capítulo 3: Modelos de redes

Las aplicaciones de red se pueden diseñar y también interpretar a la luz de algún modelo de redes.

Los modelos más populares son:

- Modelo de transparencia
- Modelo cliente/servidor
- Modelo de referencia OSI
- Modelo TCP/IP

Estos modelos, lejos de ser excluyentes, se pueden utilizar en forma complementaria.

El modelo de transparencia

La meta principal de este modelo es ocultar la presencia de la red. Las aplicaciones diseñadas según este modelo deben obtener acceso transparente a los recursos de la red.

Un ejemplo es el NFS de Sun (NFS=Network File System): las aplicaciones acceden a sistemas de archivos remotos de la misma forma que a los locales[5].

Se denominan aplicaciones distribuidas a aquellas que ocultan los detalles de la comunicación a través de la red. En contraste, una aplicación de red usa los servicios de la red sin ocultar su presencia.

El modelo cliente/servidor

Es el modelo más común para el software de red y es el mencionado antes bajo el título “paradigma cliente/servidor”.

Este modelo plantea la existencia de procesos (servidores) en ciertas máquinas, que proveen servicios a otros (clientes) en otras.

Un proceso cliente contacta un proceso servidor y solicita algún servicio.

Este modelo permite aprovechar de forma más eficiente la potencialidad que ofrece un conjunto de computadoras interconectadas en una red:

- Se optimiza el uso de la capacidad de procesamiento.
- Se minimiza el tráfico en la red, ya que paraleliza al máximo posible.

Al diseñar la aplicación, ésta se divide en dos partes: cliente y servidor. El servidor se especializa en la aplicación en sí misma. El cliente resuelve detalles de comunicación con el usuario.

Por ejemplo, un servidor de base de datos. El proceso de información que implica la resolución de las consultas se hace en el servidor (que es donde reside la base de datos). La comunicación con el usuario, atendiendo por ejemplo a detalles particulares de la interfaz, se realiza en el cliente. Puede haber incluso más de un tipo de cliente, debido a la existencia de usuarios o hardware diferentes.

[5] La palabra local hace referencia a la propia máquina, mientras que la palabra remoto indica otra máquina dentro de la red.

El modelo TCP/IP

Los protocolos TCP/IP son –como se discutirá más adelante- el núcleo de Internet.

La terminología del modelo OSI ayuda a describir TCP/IP, pero para un análisis más estricto, surge la necesidad de un nuevo modelo para describir estos protocolos.

Una descripción adecuada resulta de considerar 4 capas.

El modelo TCP/IP y sus capas se analizará más adelante.

El modelo de referencia OSI

Es un modelo desarrollado por ISO: International Standards Organization. OSI significa: Open Systems Interconection (Interconexión de Sistemas Abiertos).

El modelo propone una división del problema “interconectar computadoras a través de redes” en 7 capas o subproblemas.

Los cuestionamientos principales que se hacen a este modelo son:

- Muchas redes existentes no mapean bien su estructura en las capas del modelo.
- La división en 7 capas conduce a implementaciones ineficientes.

De todos modos constituye una herramienta de utilidad indudable a la hora de analizar el funcionamiento de una red y los distintos problemas que surgen. El modelo de referencia OSI provee un marco de referencia común para discutir las comunicaciones.

El modelo divide el estudio de una red en 7 capas. Es decir, que el problema original se subdivide en 7 subproblemas más simples (técnica “divide y reinarás”).

Cada capa brinda servicios a la inmediatamente superior.

Los objetivos detrás de esta división son:

- Favorecer la interoperabilidad entre productos de la distintas empresas.
- Minimizar el impacto de cambios.

Mientras la interfaz que define la comunicación entre capas adyacentes permanezca fija, es posible cambiar la implementación de las funciones dentro de una capa sin afectar al resto.

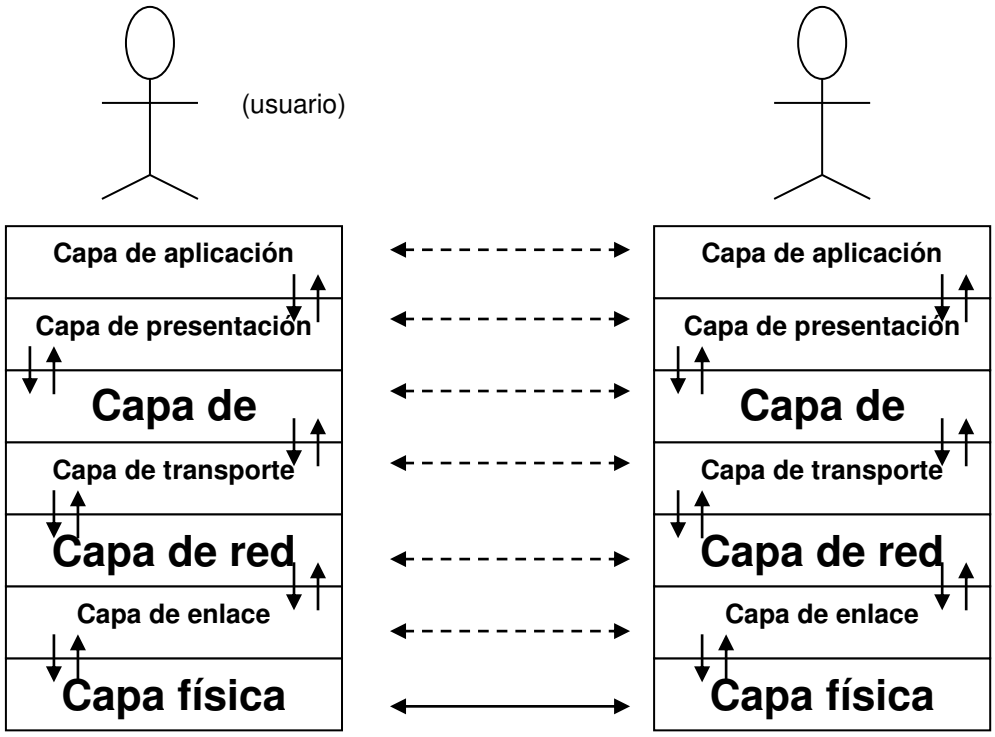
Al analizar la comunicación entre dos aplicaciones, se programa pensando a cada capa como virtualmente conectada con su par.

Entonces, cada capa está comunicada virtualmente con su par y, excepto la 1, cada capa está comunicada físicamente con la capa inmediatamente inferior. La única conexión física es a nivel de capa 1.

Las líneas punteadas en la figura muestran la comunicación virtual, y las líneas rellenas muestran la comunicación real.

En la figura también se presentan las 7 capas del modelo OSI.

ISO establece en el modelo OSI estándares para interfazs y para protocolos. La norma correspondiente es la norma ISO 7498.



Los datos “bajan” de una capa a otra hasta alcanzar el nivel inferior de la “pila”. Luego los datos se transmiten a través del medio de comunicación. Al llegar al otro extremo, los datos “suben” hasta llegar a la capa par con la que originó la transmisión.

Cuando un mensaje baja en la pila, cada capa agrega su propia información al mensaje.

A medida que un mensaje sube en la pila, cada capa quita (y utiliza) su información del mensaje y pasa el resto a la capa inmediatamente superior.

La capa 1 es la de menor nivel de abstracción y la capa 7 es la de mayor nivel de abstracción.

Las cuatro capas inferiores son relativas a la red y las tres capas superiores son relativas a la aplicación.

Los nodos (hosts) intermedios (por los que pasan mensajes que comunican máquinas no adyacentes), sólo involucran sus tres capas inferiores en esa comunicación.

Cada vez que la capa N detecta un error, lo notifica a la N+1.

Ejercicios

3.1) Considerar el caso de a) el envío de un email b) el acceso a una página Web. Esquematizar el pasaje de información entre las diferentes capas del modelo OSI.

3.2) Conseguir la norma ISO mencionada.

Las 7 capas son:

- Capa física: Define las características físicas de la red.
- Capa de enlace: Provee entrega de datos entre máquinas físicamente adyacentes.
- Capa de red: Provee entrega de datos a través de la red.
- Capa de transporte: Provee detección y corrección de errores end-to-end.
- Capa de sesión: Maneja sesiones entre aplicaciones.
- Capa de presentación: Estandariza la presentación de datos.
- Capa de aplicación: Consiste en los programas de aplicación que usan la red.

El modelo OSI y sus capas se analizan en detalle en los capítulos siguientes.

Capítulo 4: Capa física

La capa física controla la transmisión de datos de una forma adecuada de acuerdo al medio de transmisión. Los protocolos de capa 1 incluyen especificaciones eléctricas y mecánicas.

El servicio brindado es la transmisión de cierta cantidad de bits. Para fijar ideas, ayuda mucho suponer que la función de la capa 1 es transmitir/recibir un solo bit por vez. También se asegura la entrega de los bits en orden.

En esta capa se define qué es un bit (nivel de corriente, etc.), cuánto dura, sincronizaciones, cuál es el canal, cómo es el medio de transmisión, los conectores utilizados, etc.

Los puntos importantes a decidir con relación a la capa 1 son:

- Medio de transmisión y conectores
- Topología (física)
- Señal digital o analógica
- Sincronización
- Ancho de banda a utilizar
- Multiplexión

Definiciones

Medio de transmisión: es el medio donde se transmiten los bits “en bruto” de una computadora a la otra.

Canal: el camino que comunica dos puntos de la red se llama canal, y físicamente, consta de uno o más medios.

Velocidad de transmisión: Mide el número de bits por segundo (bps) que se pueden transmitir.

Baudio: cantidad de cambios de estado del medio por segundo.

Ancho de banda de un canal: Es el ancho del intervalo de frecuencias que el canal deja pasar sin atenuación. Se mide en Hz (Hertz). El ancho de banda de un canal es proporcional a la velocidad con que se puede transmitir en el mismo.

Relación señal/ruido: Ruido son aquellas señales espúreas y no predecibles que se superponen a la señal a transmitir. Cuanto mayor es la relación señal/ruido, mayor es la capacidad del canal. La unidad de medida son los decibeles (dB).

Los teoremas de Nyquist y Shanon establecen cotas para la cantidad máxima de información que se puede transmitir en un canal dado su ancho de banda, los niveles discretos de señal y la relación señal/ruido.

Atenuación, dispersión y repetidores: es la pérdida de potencia de la señal transmitida con la distancia. Las señales sufren atenuación que se ve incrementada con la distancia. Las demoras aumentan claramente con la distancia. Las señales sufren dispersión con la distancia. Para contrarrestar estos efectos existen dispositivos, por ejemplo los repetidores, que aplican la función identidad (transmiten la misma señal que reciben) pero la reconstruyen eliminando los efectos de atenuación y dispersión.

Las señales pueden ser analógicas o digitales. Una señal analógica tiene un rango denso[6] de niveles posibles. Una señal digital tiene un rango discreto de niveles posibles, que en general serán 1 o 0.

Tipos de conexión física

Hay dos tipos de conexiones:

- Punto a punto: Existe una conexión entre dos computadoras.
- Multipunto: Existe una conexión entre un grupo de computadoras. Lo que uno transmite lo reciben todos. Si transmiten dos a la vez hay problemas llamados colisiones, que implican pérdida de lo transmitido. También se denomina a este tipo de conexiones como broadcast, canal o bus.

Medios de transmisión

Los medios se comparan en función de un conjunto de características:

- Costo
- Velocidad de transmisión
- Ancho de banda
- Relación señal/ruido
- Distancia de operación
- Dificultad de instalación y uso

Los medios de transmisión se pueden clasificar en guiados o no guiados.

- Medios guiados:
  - Par trenzado (blindado o no blindado)
  - Coaxial
  - Fibra óptica
- Medios no guiados:
  - Ondas de radio
  - Microondas
  - Enlaces satelitales
  - Ondas infrarrojas

Cable par trenzado

Es un par de conductores, cada uno aislado, trenzados entre sí. Cada uno de los cables puede ser unifilar (conductor sólido) o multifilar (varios hilos trenzados entre sí).

El trenzado disminuye la inducción de los campos magnéticos espúreos. Por ejemplo, el fenómeno de diafonía que observamos en las comunicaciones telefónicas se debe generalmente a un pobre trenzado.

Su uso más común es: conexión de terminales asincrónicas, conexión telefónica, Ethernet sobre par trenzado, Token Ring.

Existen 5 tipos de categorías de cable par trenzado, de acuerdo a su posible uso:

[6] Un conjunto se dice denso cuando dados dos elementos siempre es posible encontrar un elemento que esté entre ambos en una relación de orden dada. Por ejemplo, un intervalo de Reales es un conjunto denso.

- Categoría 1: para aplicaciones de voz y no apto para datos.
- Categoría 2: para velocidades hasta 4 Mbps.
- Categoría 3: para velocidades hasta 10 Mbps.
- Categoría 4: para velocidades hasta 16 Mbps.
- Categoría 5: para velocidades hasta 100/150 Mbps.

Los cables de tipo par trenzado se clasifican en UTP (par trenzado no blindado) y STP (par tranzado blindado).

El caso más común de conector es el RJ45. La ficha contiene 8 contactos, normalmente asignados a 4 pares trenzados que van dentro de un mismo cable.

Cable coaxial

Es un conductor de cobre, recubierto por una capa aislante, afuera de ésta se encuentra un segundo conductor en forma de malla y recubriendo el conjunto la cubierta plástica exterior.

Tiene buena aislación frente a interferencias y un ancho de banda de centenares de Mhz. Se emplea frecuentemente en comunicaciones con terminales asincrónicas, redes Ethernet e instalaciones de TV por cable.

Un parámetro importante del cable coaxial es su impedancia característica. Los cables coaxiales usados en Ethernet tienen 50 ohmios de impedancia característica y los usados en TV tienen 75 ohmios.

Se utilizan “T” para las uniones y tapones de modo de evitar reflexiones de la señal y así mantener la impedancia constante en cada tramo.

Los conectores BNC son los más usados con cable coaxial.

Fibra óptica

La fibra óptica (en su versión más común) está formada por dos materiales transparentes y homogéneos. El de dentro se denomina núcleo (core) y el exterior cáscara o envoltura (cladding). El índice de refracción de la envoltura es menor que el del núcleo, y esto permite a la luz viajar “rebotando” en las paredes de la fibra.

Las fibras se clasifican en monomodo o multimodo de acuerdo a que permitan o no más de un modo de propagación. Las fibras donde el tamaño del núcleo sólo permite la propagación de un único modo son las llamadas monomodo.

Las ventajas de la fibra óptica son:

- Inmunidad a la interferencia y ruidos.
- Gran capacidad de transmisión.
- Baja atenuación.
- Seguridad de las transmisiones (es muy difícil de “pinchar”).

Las fibras requieren conectores especiales, los más usados en las LAN son los tipo ST, y también existen los SC y FC-PC.

Para empalmar dos fibras existen dos técnicas básicas: empalmes de fusión y empalmes mecánicos.

En cuanto a las fuentes de luz, existen dos tipos básicos: LEDs (LED= diodo emisor de luz) y Lásers. Estos últimos tienen menos dispersión cromática que los LEDs.

La siguiente tabla ensaya una comparación entre los medios de transmisión mencionados:

Medio	Costo	Facilidad de instalación	Frecuencia	Atenuación	Inmunidad a interferencias electromagnéticas
UTP	Muy Bajo	Muy simple	1 a 100 Mbps	Alta	Baja
STP	Medio	Simple a media	1 a 155 Mbps	Alta	Media a baja
Coaxial	Bajo a medio	Simple	1 Mbps a 1 Gbps	Media	Media
Fibra óptica	Medio a alto	Difícil	10 Mbps a 2 Gbps	Baja	Alta

Enlaces de radiofrecuencia

Se utilizan ondas de radio. Estas ondas son tienen las siguientes características:

- Fáciles de generar
- Viajan grandes distancias
- Son omnidireccionales
- No es necesario alinear con cuidado emisor y receptor
- Se ven afectadas por motores y otros equipos eléctricos
- Ofrecen un ancho de banda bajo

Los enlaces de radiofrecuencia tienen las siguientes características:

- Se dan en el rango de los 2 a 40 GHz
- Corresponden a longitudes de onda de entre 15 y 0.75 cm.
- A estas frecuencias, las ondas se pueden transmitir con poca dispersión, con el uso de antenas parabólicas.
- Se usa entre puntos separados por distancias no mayores a los 100 Kms.
- Debe haber visibilidad directa entre los puntos. Emisor y receptor deben estar alineados.
- El ancho de banda es del orden de las decenas a los cientos de MHz.
- La lluvia absorbe las ondas.

Enlaces satelitales

Los enlaces satelitales tienen las siguientes características:

- Desde el punto de vista de la banda de frecuencias, son enlaces de microondas.
- La particularidad es que al usar un satélite artificial como repetidor, permite superar las limitaciones de distancia de dichos enlaces.
- Los enlaces pueden ser punto a punto, como los que usa el servicio Datamundi de ANTEL, o pueden ser punto-multipunto, como los que usa el servicio Minisat de ANTEL.

Otros medios de transmisión

Telefonía Celular
Radiomodems
Enlaces UHF
Spread Spectrum
Infrarrojos

Canales telefónicos

La red telefónica se compone de un conjunto de centrales, conectadas entre sí. Cada central está comunicada con los teléfonos a través de pares trenzados. Cada par va de la central a un terminal telefónico individual (es decir, a un teléfono). Los pares se van agrupando en las borneras, cajas de distribución, etc. hasta llegar a la central. Para ello se utilizan cables multipar, que tienen muchos pares agrupados.

El conector RJ11 es el estándar utilizado para la conexión de los teléfonos. El RJ11 tiene cuatro conectores, pero son dos los que conectan al par telefónico. Los otros dos se utilizan por ejemplo para señalización interna en las centralitas.

Las líneas telefónicas son analógicas, y las centrales telefónicas son digitales. Entonces, hay que resolver las conversiones involucradas al respecto. Las distintas centrales telefónicas componen una red y se comunican en forma digital (por eso se dice “Uruguay 100 % digital”).

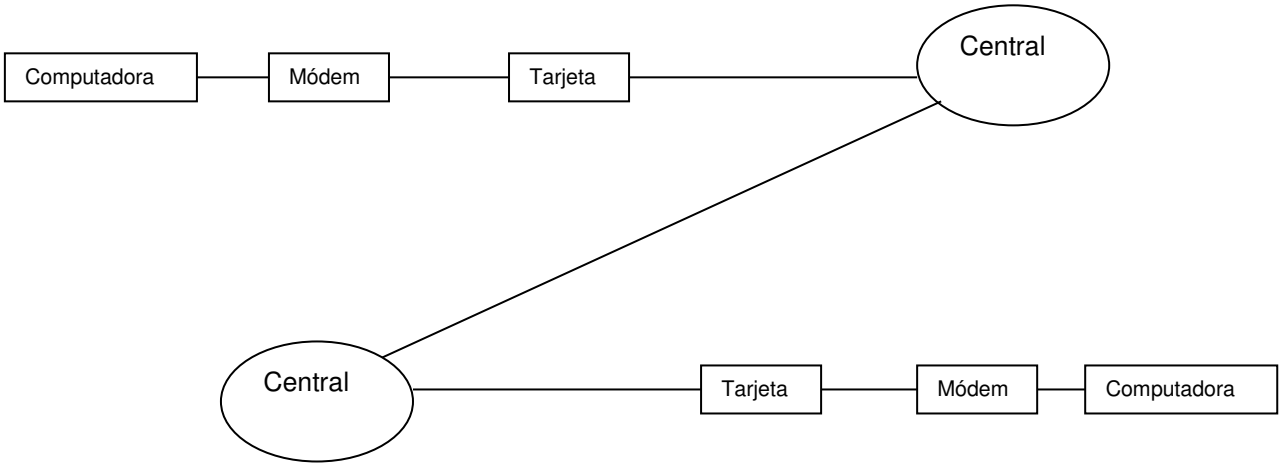
Entre centrales, las comunicaciones asociadas a las llamadas se “switchean”, esto es, se establece un camino dedicado a transmitir la conversación.

Al llegar las líneas analógicas a la central, pasan por una tarjeta de abonado, que realiza la conversión de señal analógica a digital y viceversa.

La computadora maneja niveles discretos de señal, 0 o 1, y por lo tanto la comunicación con ella debe ser digital. Como la línea es analógica, de este lado también se debe realizar conversión, y es el módem quien la realiza.

Un módem es un modulador-demodulador, lo que implica convertir de digital a analógico y viceversa. Un módem permite entonces transmitir datos por circuitos analógicos. El módem acepta como entrada una corriente de bits y produce una portadora modulada como salida o viceversa. El módem se coloca entre la computadora y la línea telefónica.

El siguiente esquema resume la situación descripta:



Las uniones entre cables, que en muchos casos tienen diferente sección de cobre impacta en la atenuación de las señales.

Para interconectar centrales de ciudades distintas se utilizan enlaces de microondas, fibra óptica o enlaces satelitales, dependiendo de las distancias y la accesibilidad. En el Uruguay el grado de digitalización de estos enlaces es –como se dijo antes- muy alto.

Para interconectar telefónicamente los países se utilizan básicamente enlaces satelitales y, cada vez más, los enlaces de fibra óptica, por ejemplo el cable Unisur.

El ancho de banda de un canal telefónico está determinado por el ancho de banda de los canales de transmisión de los enlaces de microondas y satelitales. Este ancho de banda está fijado entre 3.6 y 4 KHz.

La capacidad esté determinada, entonces, por el ancho de banda de los equipos de transmisión y la relación señal/ruido, en la que influye el trenzado (inducción del ruido) y la distancia y calidad del cobre utilizado (atenuación de la señal).

Transmisión serial o paralela

La transmisión serial implica la transmisión de un bit por vez. Un ejemplo es el protocolo RS232 o el RS232C.

La transmisión paralela implica la transmisión de más de un bit por vez. Un ejemplo es el protocolo Centronics.

La transmisión en paralelo es preferida para cortas distancias donde se precisan altas velocidades de transferencia. Por ejemplo el bus de una computadora, conexión de discos, comunicación con la impresora.

Los enlaces físicos preferidos actualmente para la implementación de redes de computadoras son todos seriales.

Sincronización

Para la correcta interpretación de los bits, el emisor y el receptor deben estar sincronizados. Es decir, se acuerda el comienzo y duración de los bits.

La sincronización depende de los relojes en las computadoras que intervienen en la comunicación. Una diferencia de velocidades entre ambos relojes de una millonésima del período, hará que cada 500.000 bits pueda ocurrir una lectura errónea.

Hay dos mecanismos diferentes de sincronización:

- Transmisión asincrónica:
    - Emisor y receptor tienen su propio reloj de la misma velocidad.
    - En la ausencia de transmisión el canal está en reposo.
    - Con el comienzo del primer bit transmitido el canal sale de reposo y eso es interpretado por el receptor para sincronizar su reloj.
    - Se transmiten por vez pocos bits, de forma que las diferencias entre ambos relojes no lleguen a ocasionar desviaciones durante la transmisión.
- Transmisión sincrónica:
    - Se transmite además la información de reloj.
    - Puede transmitirse en forma independiente (por ejemplo RS232 sincrónico).

- Puede enviarse información de reloj embebida en la señal (por ejemplo Manchester).
- Pueden incluirse preámbulos con sólo datos de reloj, para mejorar la sincronización del receptor.

Algunos tipos de codificación para los bits:

Sea M el nivel eléctrico de señal. Además, en todos los casos se considera u largo fijo t para cada bit.

Codificación binaria: el 0 se codifica como un pulso de valor 0 y largo t. El 1 se codifica como un pulso de valor M y largo t.

Codificación Manchester: El 0 se codifica como un pulso de valor 0 durante t/2 y M durante t. El 1 se codifica como un pulso de valor M durante t/2 y de valor 0 durante t/2.

Hay otras codificaciones como la codificación Manchester diferencial.

Multiplexión

La multiplexión se usa para compartir el canal entre varias transmisiones.

Las técnicas más usadas son:

- Multiplexión por división de frecuencias (FDM): Se asigna a cada transmisión un rango de frecuencias.
- Multiplexión por división de tiempo (TDM): Se asigna a cada transmisión una ranura de tiempo.
- Multiplexión estadística por división de tiempo (STDM): Se asigna a cada transmisión una ranura de tiempo variable que se calcula de acuerdo al uso que ella está haciendo del canal.

Ejemplos de protocolos de capa física

- RJ45: Especifica el conector para conectar computadoras a un hub vía UTP.
- RS232: establece las reglas para comunicación serial (un bit por vez) entre computadoras o entre una computadora y un módem. El protocolo abarca la definición de asignación de señales en los conectores, niveles de voltaje para 0 o 1, etc.
- Las especificaciones de capa 1 en las normas 802.3 (Ethernet) y 802.5.

Ejercicios

4.1) Elegir una red y analizar los protocolos de nivel físico que existen en ella.

4.2) Construir una lista de protocolos de nivel físico y el problema que resuelve cada uno.

Capítulo 5: Capa de enlace

La capa de enlace utiliza el servicio provisto por la capa física: transmitir bits entre dos puntos adyacentes de la red.

La función de la capa de enlace es transmitir tramas entre dos puntos adyacentes de la red.

Una trama es un conjunto ordenado de bits. Las tramas pueden ser de largo fijo o de largo variable. En este segundo caso las tramas se delimitan por configuraciones especiales de bits.

La capa de enlace se divide en dos subcapas:

- Subcapa de acceso al medio (MAC):
    - Encapsula el acceso al medio.
    - Existe siempre que el hayan conexiones basadas en canales.
- Subcapa de control del enlace lógico (LLC):
    - Intenta detectar y/o corregir errores en las tramas.
    - Se provee una interfaz orientada o no a conexión para la capa 3.
    - Se asegura que las tramas lleguen en la secuencia correcta.
    - También se implementa control de flujo.

Los problemas a resolver entonces en la capa 2 son:

- Arbitraje del canal si lo hubiera
- Proporcionar la interfaz del servicio
- Definir la manera que los bits se agrupan en tramas
- Proporcionar comunicación confiable y eficiente entre dos máquinas adyacentes
- Manejar errores en la transmisión
- Control de flujo
- Direccionamiento

Normalmente la capa 2 aparece implementada en los circuitos de las propias tarjetas de red (que también contienen la capa 1 en ese caso).

La norma ISO 8802-2 (IEEE 802.2) es un ejemplo de un conjunto de protocolos para subcapa LLC en capa de enlace.

La norma ISO 8802-3 (IEEE 802.3) es un ejemplo de un conjunto de protocolos para subcapa MAP en capa de enlace.

Construcción de las tramas

Con el fin de detectar y/o corregir errores, en esta capa se divide el flujo de bits en tramas (a veces llamadas marcos, frames).

Cada trama contiene información propiamente dicha e información redundante. La redundancia es necesaria para detectar y/o corregir errores.

Al enviar información, se genera la información redundante correspondiente (redundancia) y se colocan la información y la redundancia en una trama. Se envía la trama. El receptor extrae de ella la información y la redundancia. Recalcula la redundancia a partir de la información y de esa forma puede detectar la presencia de errores en el flujo de bits.

Hay varias formas de delimitar tramas de largo variable:

- Conteo de caracteres: se incluye en el cabezal de la trama un campo que indica la cantidad de caracteres de la misma. El receptor utiliza este campo para saber el final de la trama.
- Caracteres de inicio y fin, con relleno de caracteres: Se asignan caracteres especiales de principio y fin. Un problema es si los mismos deben aparecer dentro de la información.
- Caracteres de inicio y fin, con relleno de bits: Igual que el anterior pero reservando configuraciones especiales de bits.
- Violación de codificación de la capa física: Por medio de la notificación de error en capa 1 se detecta el comienzo y fin de trama.

Control de errores

El primer problema que aparece es el de asegurar que todas las tramas sean entregadas a la capa de red de la máquina destino en el orden apropiado.

Para solucionarlo, se proporciona al receptor retroalimentación sobre lo que está ocurriendo del otro lado de la línea. Para esto, cada vez que el receptor recibe una trama avisa si llegó bien o no.

Surge entonces el problema es detectar errores en las tramas recibidas. Esto se tratará más adelante en este capítulo.

Continuando con el protocolo basado en confirmaciones, si se perdiera una trama completa el protocolo se detendría, por lo que hay que usar temporizadores (timeouts) que permitan habilitar reenvíos.

También, si se pierde un mensaje o un acuse de recibo, se podría recibir un mensaje varias veces. La solución para esto es asignar un número de secuencia a las tramas.

Este tipo de controles aparece también en los protocolos de capa 4, aunque a más alto nivel (con paquetes y no con tramas).

Detección y corrección de errores

Como resultado de los procesos físicos que ocasionan el ruido, los errores tienden a presentarse en ráfagas más que en forma aislada.

Para detectar y/o corregir errores es necesario agregar redundancia a los datos a transmitir. Es claro que corregir errores es mucho más ambicioso que detectarlos, lo que implica la necesidad de mayor redundancia en caso de desear corregir.

Sea una trama de m bits de información y r bits de redundancia. El largo de la trama es $n = m + r$. A una secuencia de n bits se la denomina palabra del código.

Por ejemplo si m=3 y r=1, entonces r=4. Un ejemplo de palabra de ese código sería 0101.

Se denomina distancia de Hamming (o distancia) a la cantidad de diferencias (bit a bit) entre dos tramas.

Por ejemplo, las palabras 0011 y 0101 difieren en el segundo y tercer bit, lo que hace que d(0011,0101)=2.

La distancia de Hamming cumple las propiedades matemáticas de una función distancia.

La distancia de un código se define como la distancia mínima entre dos palabras, para todo par de palabras del mismo.

Para detectar d errores se necesita un código de distancia d + 1.

Para corregir d errores se necesita un código de distancia 2d+1.

Por ejemplo con un código de distancia 5 podemos detectar errores en cuatro bits y corregir errores en 2 bits.

Ejercicios

5.1) ¿ Cuántos bits de redundancia se necesitan para detectar 1 error en 4 bits de información?

5.2) ¿ Cuántos bits de redundancia se necesitan para detectar 2 errores en 4 bits de información ?

5.3) ¿ Cuántos bits de redundancia se necesitan para corregir 1 error en 4 bits de información ?

5.4) ¿ Cuántos bits de redundancia se necesitan para corregir 2 errores en 4 bits de información ?

Para detectar y corregir errores se debe –ya se dijo- agregar redundancia.

Si se transmiten por ejemplo m bits, se toma un subconjunto entre los 2^m posibles y se aceptan esos códigos como válidos, siendo el resto inválidos.

Si se recibe un código inválido, se reconoce un error en la transmisión. Para que un error no sea detectado, un código válido debe haberse cambiado por otro válido.

Hay varias formas de generar redundancias, entre las que se destacan la paridad par e impar, los códigos 2 de 5, los códigos de Hamming y los CRC (códigos de redundancia cíclica).

Ejemplo

A continuación se analiza en detalle la paridad par para 3 bits de información.

El bit de redundancia, llamado también bit de paridad, se calcula de modo tal que el total de unos en la trama sea par.

m=3 , r=1 , n=4

Por ejemplo, para enviar la información: 001 se envía la trama 0011 de modo de quedar con una cantidad par de unos.

El receptor recibe 0011 y separa información 001 y redundancia 1. Recalcula la redundancia a partir de la información y entonces concluye que es correcta la trama (o al menos no logra detectar errores en ella).

Si por el contrario, el receptor recibe 0111, separa la información 011 y la redundancia 1. Recalcula el bit de redundancia, le da 0 y rechaza la trama por ser incorrecta.

Sin embargo, si la trama 0011 sufre dos alteraciones en sus bits y llega 0000, la trama (incorrecta) es aceptada como válida.

El conjunto de posibles tramas de 4 bits es:

0000
0001
0010
0011
0100
0101
0110
0111
1000
1001
1010
1011
1100
1101
1110
1111

De ellas, sólo corresponden a tramas correctas la mitad:

0000 Correcta
0001
0010
0011 Correcta
0100
0101 Correcta
0110 Correcta
0111
1000
1001 Correcta
1010 Correcta
1011
1100 Correcta
1101
1110
1111 Correcta

Del universo de 16 palabras, 8 son correctas y 8 son incorrectas.

La distancia del código es 2, y por esa razón se detectan errores individuales y no es posible la corrección de errores.

Ejercicio

5.5) Repetir el estudio anterior para un código de paridad par con 4 bits de información.

Los métodos de corrección –en los que no se profundizará- asocian una trama errónea con la trama válida más cercana, en términos de la distancia Hamming.

Otro método de detección de errores son los códigos de redundancia cíclica (CRC). Estos códigos tienen una gran capacidad de detección de errores y su implementación por hardware es relativamente sencilla.

En algunos casos se proveen mecanismos de control de flujo en la capa de enlace. Se analizarán en la capa de transporte, donde se encuentran con más frecuencia.

Arbitraje del canal (subcapa MAC)

Las características de los protocolos de nivel de enlace dependen fundamentalmente del tipo de red con que se trabaja: punto a punto o difusión.

Cuando la red es de tipo difusión, se agrega el problema de arbitrar el canal.

En un canal lo que un integrante envía lo recibe el resto, y si hay emisiones simultáneas hay pérdida de información. El arbitraje del canal consiste en proveer un mecanismo de transmisión de tramas sobre el mismo sin perder información por emisiones simultáneas, etc.

La única función de la subcapa de acceso al medio es el arbitraje del canal. Por lo tanto esta capa sólo tiene sentido en las redes de difusión.

Los métodos de arbitraje del canal dependen fuertemente del medio físico elegido. Estos métodos se pueden clasificar en:

- Determinísticos o no determinísticos: de acuerdo a si se puede o no determinar con exactitud el tiempo máximo que debe esperar un nodo para poder efectuar una transmisión.
- Priorizado o no: de acuerdo a si todas las estaciones de la red tienen la misma prioridad de acceso al canal.

Se denomina colisión a la transmisión simultánea (en general parcialmente simultánea[7]) de dos o más tramas. En una colisión se pierden todas las tramas involucradas. Las colisiones desperdician ancho de banda, y se debe buscar evitarlas. El propósito de la subcapa MAC es proporcionar un servicio de emisión de tramas libre de colisiones. Sólo tiene sentido hablar de subcapa MAC en redes de difusión.

El canal a arbitrar se puede asignar en forma estática o dinámica:

- La asignación estática implica dividir a priori el derecho de uso del canal. Las variantes son división por tiempo y división por frecuencia. Esta técnica es poco usada ya que cualquier nodo con baja utilización desperdicia la parte del canal que le ha sido asignada.
- La asignación dinámica del canal es la más utilizada y permite un mejor aprovechamiento del ancho de banda.

Enfocando una segunda clasificación, hay tres formas básicas de controlar la transmisión:

- Centralizado: Un nodo determina los turnos, entonces no hay forma de que dos nodos puedan transmitir a la vez.

[7] O sea, se comienza a enviar una trama al canal y alguien ya estaba enviando una trama. Si ocurre una colisión se pierden las dos tramas.

- Distribuido: No hay un nodo central, pero se establece un sistema de turnos para transmitir, por ejemplo por medio del pasaje de un token.
- Contención: Se transmite y luego, si hubo colisión, se retransmite.

Los algoritmos más comunes en subcapa MAC asignan dinámicamente el canal, usando la técnica de contención. La solución más común en la actualidad para capa 2 es el protocolo Ethernet. La base de este protocolo es el algoritmo dinámico de contención ALOHA, creado en la Universidad de Hawaii en 1970.

ALOHA fue pensado para la interconexión de sistemas usando transmisiones de radiofrecuecia. La idea del algoritmo es la siguiente:

ALOHA Puro

Si un nodo tiene que transmitir lo hace
Monitorea su transmisión
Si la transmisión fue interferida:
    Entonces espera un tiempo aleatorio y vuelve a transmitir
    Si no OK.

Tiene una utilización máxima del 18 %.

Aloha ranurado

Una variante del ALOHA, introducida en 1972 consiste en incorporar el concepto de ranuras de tiempo.

Se trabaja con ranuras de tiempo, que son períodos de tiempo iguales a la duración de la transmisión de una trama.

Así, para transmitir una trama hay que esperar al comienzo de una ranura. Esto trae el problema adicional de la sincronización de los diferentes nodos. Esto se resuelve con el uso de un nodo especial que transmite una señal de sincronización al inicio de cada ranura.

La máxima eficiencia del canal es de un 37 %. El restante 37 % estará ocupado por ranuras vacías y queda un 26 % ocupado por información destruida por colisiones.

Protocolos CSMA (Carrier Sense Multiple Access)

Tomando como base el ALOHA, se agrega lo siguiente: si una estación tiene algo para transmitir primero escucha el canal, para ver si está teniendo lugar alguna transmisión.

Los protocolos CSMA admiten varias variantes:

- 1-Persistente: Al encontrar el canal ocupado espera hasta que este se desocupa. Si el canal está libre transmite de inmediato.
- No persistente: Al encontrar el canal ocupado espera un período de tiempo aleatorio antes de escuchar nuevamente. Si el canal está libre transmite de inmediato.
- P-persistentes: Se aplica a canales ranurados. Si el canal está libre, el nodo transmite con probabilidad p y con una probabilidad 1-p espera la siguiente ranura. Si el canal está ocupado espera la siguiente ranura.

Protocolos CSMA/CD (CSMA Collision Detection)

CSMA mejora respecto a ALOHA, y ahora la idea es agregar detección de colisiones de modo de mejorar CSMA.

Una vez que un nodo detecta una colisión, aborta la transmisión, espera un tiempo aleatorio e intenta nuevamente.

El hardware del nodo debe escuchar el cable mientras transmite. Si lo que lee es distinto de lo que él transmitió, sabe que está ocurriendo una colisión.

Otro enfoque es el de tratar de evitar las colisiones: CSMA/CA (Colission Avoidance).

La idea central en este caso es que las colisiones afectan adversamente el rendimiento del sistema. Entonces se busca la forma de intentar evitar las colisiones.

Direccionamiento

Las entidades de una red requieren distinguir e identificar a los distintos dispositivos de la red.

En la capa de enlace sólo interesa la dirección física del dispositivo.

Las direcciones físicas de los dispositivos son direcciones únicas en las tarjetas, típicamente asignadas por los vendedores (a su vez asignadas a ellas por las organizaciones de estándares).

La función de la dirección de capa de enlace es identificar una computadora dentro de un canal, y así resolver las entregas de datos dentro del mismo. Un ejemplo es la dirección Ethernet, compuesta de dos partes, una dirección de fabricante y un número asignado por el fabricante.

Ejercicio

5.6) Buscar un ejemplo de dirección Ethernet y separar la dirección del fabricante y el resto.

Dispositivos de capa 2

Tarjetas de red

Las tarjetas de red se requieren para conectar una computadora a la red. El caso más común son las tarjetas Ethernet, aunque también hay para otros protocolos. Los primeros modelos tenían conectores BNC y RJ45, para cableados coaxial y UTP respectivamente y las velocidades eran de 10 Mbps. Actualmente lo más común es encontrar tarjetas Ethernet de 100 Mbps con conector RJ45.

Hubs

Un hub[8] es un punto central de conexión en la red. Normalmente las diferentes computadoras de la red se conectan al hub a través de sus tarjetas de red, por cables UTP, con conectores RJ45 en sus extremos.

[8] También llamado repetidor multipuerto o concentrador.

El hub tiene un conjunto de “bocas” (receptáculos para conectores RJ45), en general 8 o 16. La función del hub es copiar lo que recibe en una boca cualquiera en el resto.

De esta forma, lo que en capa física es una topología en estrella (un conjunto de computadoras conectadas a un dispositivo central), en capa 2 se ve (por la funcionalidad del propio hub) como un canal. Por esta razón en general se utiliza el protocolo Ethernet para arbitrarlo.

El largo máximo de un PC al hub es usualmente 100 metros con cables UTP.

Algunos hubs además disponen de conectores BNC para conectar con segmentos de cable coaxial, que en ese caso usualmente no pueden sobrepasar un largo total de segmento de 185 metros.

Switches

Un switch es externamente igual a un hub. La diferencia es que permite conexiones en paralelo entre diferentes puntos de un mismo canal. Esto permite lograr mejor performance.

Tanto hubs como switches pueden colocarse en cascada, permitiendo armar lo que se denomina cableados estructurados.

Ejercicios

5.7) Relevar las velocidades actuales de los dispositivos anteriores y contrastarlos contra lo presentado en los párrafos anteriores.

5.8) Buscar y esquematizar un ejemplo de cableado estructurado.

5.9) Averiguar precios de hubs, switches, cable UTP y conectores RJ45.

5.10) Investigar variantes de hubs y switches, por ejemplo los llamados hubs inteligentes.

5.11) Investigar los protocolos de subcapa MAC Token Bus y Token Ring y los dispositivos que los mismos requieren.

La norma IEEE 802

Estos protocolos fueron diseñados en 1985:

- 802.1: Define las características generales de los protocolos y definen un conjunto de primitivas que implementan.
- 802.2: Define el subnivel LLC común a toda la serie.
- 802.3 Define protocolos a nivel físico y subnivel MAC del tipo CSMA/CD.
- 802.4 Define protocolos a nivel físico y subnivel MAC del tipo Token Bus.
- 802.5 Define protocolos a nivel físico y subnivel MAC del tipo Token Ring.

En 802.3 se incluyen un conjunto de protocolos de nivel físico y subnivel MAC (CSMA/CD 1-persistente). Todos ellos tienen su origen en Ethernet.

Existen 4 tipos de cableado:

- 10 base 5: Cable coaxial grueso. 10 Mbps. Largo máximo de cable 500 m. Máximo 100 nodos/s.
- 10 base 2: Coaxial fino. 10 Mbps. Largo máximo de cable 200 m. Máximo 30 nodos/s.
- 10 base T: Sobre UTP en topología de estrella. 10 Mbps. Distancia máxima entre el hub y cada estación: 100 m. Máximo 1024 nodos/s.
- 10 base F: Sobre fibra óptica. Largo máximo de cada tramo: 2000 m. 1024 nodos/s.

Capítulo 6: Capa de red

La capa de enlace resuelve el problema de transmitir (con chequeo de errores) tramas entre dos puntos adyacentes de la red. Dos puntos son adyacentes cuando los comunica un cable o cuando pertenecen al mismo canal, en general al formar parte de la misma LAN.

Surge ahora el problema de conectar dos puntos no adyacentes. En general este es el caso cuando se comunican dos puntos que pertenecen a LANs diferentes, pero que se encuentran interconectadas.

En capa 3 o capa de red se resuelve este problema de proveer intercambio de datos entre dos máquinas cualesquiera, adyacentes o no.

Entonces, analizando con más detenimiento, se pueden encontrar los siguientes puntos a considerar en esta capa:

- Conexión de redes entre sí
- Proporcionar un mecanismo de direccionamiento
- Determinar la ruta para un mensaje
- Realizar control de flujo
- Proporcionar primitivas de servicio a la capa de transporte
- Cobro por uso de servicios de red

Conmutación

En las redes de comunicación por difusión todos los nodos comparten el medio físico de comunicación. Un ejemplo de comunicación por difusión son las redes locales armadas en base a un hub.

En las redes de comunicación por conmutación los datos son transmitidos entre el nodo origen y destino mediante el uso de nodos intermedios. Estos nodos forman la ruta, que se puede asociar a un camino en el grafo de la red.

Hay tres formas básicas de conmutación, determinadas por la forma en que se crean los caminos entre transmisor y receptor:

- Conmutación de circuitos: Se crea un camino dedicado entre los nodos que participan en la comunicación.
- Conmutación de mensajes: No se establece un camino a priori.
- Conmutación de paquetes: Como en conmutación de mensajes, pero un mensaje se divide en unidades menores llamadas paquetes.

Conmutación de circuitos es la alternativa orientada a conexión y las otras dos son no orientadas a conexión.

Conmutación de circuitos

Se establece un enlace dedicado entre el nodo origen y el nodo destino, ocupando un canal en cada enlace físico entre nodos que pertenecen a la ruta.

Requiere de tres fases: establecimiento de conexión, transferencia y desconexión.

La capacidad del canal queda asignada en el momento de establecer la conexión (parámetros de calidad de servicio) y no depende de los datos a transmitir.

No existe retardo de transferencia excepto los de propagación. Puede existir retardo en el establecimiento de la conexión.

Un ejemplo es la red telefónica. Al momento de discar se gestiona una conexión entre los dos puntos. Ese camino es el que se utiliza a lo largo de toda la llamada, y este se conserva mientras dure la misma.

Conmutación de mensajes

En este caso, cada mensaje se considera una entidad en sí mismo. Cada mensaje contiene dirección origen y destino.

Los nodos intermedios en el camino reciben los mensajes, los almacenan, analizan la dirección destino y retransmiten el mensaje al siguiente nodo en la ruta (store and forward).

Si bien existe un retardo de transferencia debido al procedimiento anterior, se logra utilizar el canal de forma más eficiente.

Este tipo de conmutación no resulta adecuado para el uso interactivo (pues no garantiza tiempos de respuesta). Una ventaja es que se adapta mejor a fallas o cambios en la topología de la red.

Como no hay cota para el largo de los mensajes, se impone un sobredimensionamiento en la capacidad de los nodos intermedios.

La red de correo uucp (Usenet) funcionaba en base a conmutación de mensajes.

Conmutación de paquetes

Permite combinar las ventajas de los anteriores y disminuir sus desventajas. Es similar a conmutación de mensajes, pero cada mensaje se divide en paquetes de largo fijo (por ejemplo 128 bytes en X25 o 53 bytes en ATM).

Es decir, la información a enviar se divide en paquetes de largo fijo. Cada paquete tiene guardadas dirección origen y destino y va pasando de un nodo a otro en la red hasta alcanzar el nodo destino. Los paquetes no tienen un camino establecido a priori.

Hay dos variantes dentro del a conmutación de paquetes: por datagramas (asociado a servicios sin conexión) o por circuitos virtuales (asociado a servicios con conexión).

En los circuitos virtuales se establece una conexión *lógica* previo a la transmisión de los datos. Esto significa que se determinan la ruta por donde se transmitirán los paquetes. Esta ruta se mantiene durante toda la conexión.

En los datagramas cada paquete es tratado en forma independiente. Paquetes con igual nodo origen y destino podrían tomar rutas distintas. Esto puede hacer que paquetes pertenecientes a un mismo mensaje lleguen a destino en orden forma desordenada.

Al comparar circuitos virtuales y datagramas se deben tener en cuenta:

- Los circuitos virtuales requieren menos espacio en los dispositivos de conectividad.
- Los datagramas administran mejor el ancho de banda, pues no lo reservan como parte de una conexión.
- En circuitos virtuales se gasta tiempo en establecer la conexión, pero una vez establecida se garantizan parámetros de calidad de servicio.

- En datagramas se gasta tiempo en analizar la dirección de cada paquete.
- Los circuitos virtuales evitan el congestionamiento en la red.
- Con una red de datagramas es más difícil evitar el congestionamiento.

Discutidos los puntos anteriores, se pueden establecer con mayor detalle los requerimientos para los servicios de capa de red:

- Independencia de la tecnología utilizada en la red.
- Aislamiento del nivel superior (transporte) del número, tipo y topología de las redes involucradas.
- Presentación hacia las capas superiores un sistema de direcciones coherente y único.
- Dicho sistema de direcciones debe proporcionar una dirección para cada nodo que permita identificarlo en forma única.
- Proporcionar un mecanismo para establecer conexiones a nivel de red, y las mismas se deben poder identificar.
- Transferencia de datos.
- Administración de la calidad del servicio:
    - Tasa de error
    - Disponibilidad
    - Capacidad
    - Costo
    - Retardo
- Notificación de errores irrecuperables a la capa de transporte.
- Proporcionar la entrega en secuencia de paquetes.
- Control del flujo de la información.
- Permitir la transferencia de datos urgentes.
- Establecimiento y liberación de conexiones.

Algunos de los servicios anteriores son opcionales y dependen si se eligió a este nivel implementar servicios con o sin conexión.

Hardware de capa de red

Los routers son dispositivos que permiten la interconexión de redes. Trabajan a nivel de capa 3. Su función consiste en implementar el ruteo entre las distintas redes. Más adelante se incluye un estudio de caso que ejemplifica esto.

Usualmente la configuración de los routers se realiza “logueándose” a ellos vía un shell de Unix.

Los bridges unen dos redes, pero permiten pasar de un lado a otro sólo los paquetes que realmente lo necesitan. Esto disminuye el tráfico frente al caso en que cada envío de un paquete implica una copia a todos los integrantes del canal.

Algoritmos de ruteo

El algoritmo de ruteo es el software incluido en la entidad de red que determina los caminos por los que se enviará la información manejada en este nivel.

Las siguientes son propiedades deseables para los algoritmos de ruteo:

- Correcto: resolver el problema de enviar información e un punto a otro.
- Simple.

- Robusto: debe soportar situaciones contingentes, o de uso fuera de los rangos normales.
- Estable: debe mostrar un comportamiento predecible a lo largo del tiempo.
- Justo: no debe favorecer a determinados nodos a menos que esto se establezca explícitamente.
- Óptimo: debe consumir la menor cantidad de recursos posible.

Los algoritmos de ruteo se pueden clasificar según diferentes propiedades:

- Estáticos o dinámicos: Según si el algoritmo puede adaptarse (dinámico) o no (estático) a cambios en la red y/o su configuración.
- Locales (aislados) o globales: Según si las decisiones de ruteo se toman de acuerdo a datos obtenidos del nodo actual solamente (local) o de la red en su conjunto (global).
- Centralizados o distribuidos: Según si las decisiones de ruteo se toman en un solo nodo (centralizado) o si se toman en forma distribuida entre los distintos nodos de la red (distribuido).

Al razonar con los algoritmos de ruteo es útil hacerlo considerando el grafo de la red. Este concepto fue presentado en el capítulo 2.

A efectos del ruteo es usual manejar una variante donde los vértices representan los routers y las aristas enlaces entre los mismos. Se considera un grafo ponderado. El peso de las aristas puede asociarse a algunos de los siguientes atributos:

- Distancia física del enlace.
- Velocidad de transmisión.
- Tiempo medio en cola en el router origen.
- Tiempo medio en enviar un paquete estándar.
- Costo de envío ($).
- Tráfico.
- Combinaciones de los anteriores.

Tablas de ruteo

Se denominan así a las tablas que permiten a cada nodo decidir por donde mandar un paquete dado, a la vista de su origen, destino, y/o el punto por el que ha llegado al nodo.

Algoritmos estáticos, no adaptables

- Camino de costo mínimo: Este algoritmo es bien sencillo. Calcula para todo par de routers la trayectoria de costo mínimo. Por ejemplo se puede utilizar el Algoritmo de Dijkstra (1959) que permite determinar sobre un grafo el camino de costo mínimo entre todo par de nodos. Las estructuras de ruteo guardan, para cada nodo, para ir a cada nodo el nodo al que se debe mandar el paquete considerado.
- Multicamino: Se trata de elegir diferentes caminos para los distintos paquetes de modo de no saturar en único camino. Para cada nodo se guarda una tabla que da una serie de alternativas para mandar paquetes a un nodo determinado, todas con su correspondiente probabilidad.

Algoritmo centralizado

Un único nodo especial (llamado RCC: Routing Control Center), conoce y mantiene información global del estado de la red. Este nodo es el encargado de determinar todas las rutas.

Periódicamente, cada nodo envía al RCC sus vecinos activos, información de tráfico, largo de cola, etc. El RCC recalcula las tablas de ruteo y las distribuye a todos los routers de la red.

Desventajas:

- Vulnerabilidad (el ruteo depende de un solo nodo).
- El calculo total puede ser costoso en tiempo.
- Genera mucho tráfico desde/hacia el RCC.

Ventajas:

- El resto de los routers no tienen que calcular nada.
- Óptimo.

Algoritmos aislados

- Papa caliente: Cuando llega un mensaje se envía por la cola más pequeña, distinta a la que llegó. No asegura que el paquete llegue.
- Aprendizaje reverso: Cada paquete incluye entre otros los campos costo actual y router origen. Cada router por el que pasa el paquete actualiza el campo costo_actual, de acuerdo con la métrica elegida. El router analiza los paquetes que llegan a él, esto es, compara el costo en su tabla de ruteo con el campo costo actual del paquete y si es menor actualiza la tabla. De esta forma va descubriendo la mejor ruta hacia los demás nodos. Este algoritmo trabaja bien cuando la red mejora, pero si la misma empeora no se entera. Este problema se puede resolver borrando las tablas de ruteo cada determinado período de tiempo.
- Por inundación: Cada paquete es enviado por todas las salidas, exceptuando por la que arribó. Una solución para que los paquetes no circulen infinitamente por la red, es que todos los nodos quiten los paquetes que tengan un tiempo de vida mayor que uno dado, o que hayan atravesado un número determinado de nodos. Este algoritmo es muy robusto pues siempre que se envía un paquete, si existe un camino el paquete llega. Una aplicación son las rede militares. Hay variantes: inundación selectiva e inundación al azar.
- Ruteo distribuido: Cada nodo intercambia información de ruteo con sus nodos con sus nodos vecinos. De esta manera construye su tabla de ruteo. La tabla tiene una entrada por cada nodo y en ella se almacena la ruta preferida y la distancia estimada. Cada cierto tiempo los nodos vecinos intercambian sus tablas y de esa forma actualizan información.
- Ruteo jerárquico: Se divide la red en regiones. Existe camino para todo par de nodos de una misma región. Dentro de cada región existen nodos que mantienen información de rutas a las demás regiones. Puede haber más de una jerarquía dependiendo del tamaño de la red. La ventaja es que las tablas se reducen, y esto trae como consecuencia un ahorro de espacio y una disminución del tiempo de procesamiento. El costo de la trayectoria se incrementa.

Ejercicio

6.1) Ejemplificar los algoritmos de ruteo sobre una red de 5 nodos, proponiendo valores para las tablas de ruteo cuando corresponda.

Sistema de dominios

El sistema de dominios fue desarrollado para simplificar la administración de los nombres de los nodos en grandes redes.

Cuando la red contiene miles de nodos –como es el caso de Internet- no es eficiente que cada nodo contenga la tabla de conversiones de todos los nombres a las direcciones.

La idea es establecer una estructura jerárquica, en forma de árbol, con nodos que cumplen la función de "servidor de nombres".

Los servidores de nombres conocen cómo ubicar un nodo debajo de su jerarquía o bien otros servidores de nombres que sí lo pueden ubicar.

El nombre completo se forma especificando esta estructura jerárquica.

Ejercicio

6.2) Dar ejemplos de nombres de dominio.

6.3) Buscar información sobre las jerarquías de dominios en Internet.

La familia de protocolos X.25

Es un conjunto de protocolos especificados por el CCITT,y corresponden a los niveles 1, 2 y 3 del modelo OSI. En su momento fue muy utilizado en redes WAN.

A nivel físico se definen los protocolos X.21 (interfaz digital) y X.21bis (básicamente RS232C). A nivel de enlace se utilice LAP-B. A nivel de red se define el PLP (Packet Layer Protocol), que es un protocolo de conmutación de paquetes.

Se definen servicios con conexión a niveles 2 y 3.

El protocolo IP e Internet

El protocolo IP forma parte del conjunto de protocolos TCP/IP, definidos para Arpanet. Inicialmente utilizado en redes con sistema operativo Unix, su uso se ha extendido y hoy los protocolos TCP/IP son la base de Internet.

La red ARPANET nace como resultado del proyecto DARPA, del Departamento de Defensa de EEUU, orientado a interconectar equipos de características diferentes y lograr interoperabilidad entre ellos.

El protocolo IP es típico de capa 3 en el modelo OSI, más allá que en el contexto del proyecto DARPA, se utilizaba un modelo diferente, que se analiza en detalle en el capítulo próximo.

TCP e IP son dos protocolos individuales dentro de la familia más amplia que recibe el nombre de TCP/IP.

Una internet es una red que resulta de interconectar dos o más redes, en general compartiendo los protocolos de capa 3.

La red Internet es una internet de alcance mundial que surgió a partir de la mencionada red ARPANET.

El protocolo IP utiliza en capa 3 un servicio sin conexión, que se encarga de rutear datagramas a través de la red.

Uno de los problemas a resolver en capa 3 es proporcionar un mecanismo de direccionamiento, que en este caso se logra a través de las direcciones IP.

Las direcciones IP

Los nodos se identifican mediante direcciones de 32 bits[9], que se presentan agrupados de a 4 bytes[10], o sea 4 números entre 0 y 255.

La dirección está formada por tres partes:

- Dirección de red
- Dirección de subred
- Dirección del nodo o host

Hay tres categorías de 5 tipos de direcciones IP, asociados a las letras A a E. Las direcciones tipo A, B y C corresponden a categorías de redes, y las D y E corresponden a rangos de direcciones reservadas.

A continuación se presentan los criterios de clasificación y las características de cada uno de los tipos de direcciones:

- Redes tipo A:
  - Son pocas redes de gran tamaño.
  - El rango de direcciones va de: 1.0.0.0 a 127.255.255.255.
  - Se reconocen por el primer bit en 0 (correspondiente al rango 0..127).
  - El primer byte corresponde a la red.
  - Los otros tres bytes corresponden a subred y host.
  - Entonces hay capacidad para 127 redes y 255^3 hosts diferentes en cada red.
- Redes tipo B:
  - Una cantidad intermedia de redes de tamaño medio.
  - El rango de direcciones va de 128.0.0.0 a 191.255.255.255.
  - Se reconocen por los dos primeros bits en 10 (correspondiente al rango 128..191).
  - Los dos primeros bytes corresponden a la red.
  - Los otros dos bytes corresponden a subred y host.
  - Entonces hay capacidad para (191-128) x 255 redes y 255^2 hosts diferentes en cada red.
- Redes tipo C:
  - Muchas redes pequeñas.
  - El rango de direcciones va de 192.0.0.0 a 223.255.255.255.
  - Se reconocen por los tres primeros bits en 110 (correspondiente al rango 192..223).
  - Los tres primeros bytes corresponden a la dirección de red.
  - El byte restante corresponde a subred y host.
  - Entonces hay capacidad para (223-192) x 255^2 redes y 255 hosts diferentes en cada red.

[9] 1 bit puede tomar valores 0 o 1.

[10] 1 byte se compone de 8 bits. Pensando en términos de números binarios (base 2) se puede asociar a un byte un valor entre 0 y 255.

- Direcciones tipo D:
    - Reservado para direcciones “multicast”.
    - El rango de direcciones va de 224.0.0.0 a 239.255.255.255.
    - Se reconocen por los cuatro primeros bits en 1110 (correspondiente al rango 224..239).
- Direcciones tipo E
    - Reservado para uso futuro.
    - El rango de direcciones va de 240.0.0.0 a 247.255.255.255.
    - Se reconocen por los cinco primeros bits en 11110 (correspondiente al rango 240..247).
- El rango 248.0.0.0 a 255.255.255.255 no se utiliza.

Las direcciones de red en Internet las administra un organismo que regula la integración de Internet (NIC, Network Information Center).

Las organizaciones pueden (y usualmente lo hacen), definir a su interior un espacio de direcciones falso, y sólo asignar una IP de Internet a algunos hosts y/o routers seleccionados.

La subred y el host

La separación de la dirección de subred y el host se realiza utilizando una máscara.

La máscara se compone (igual que la dirección IP) de 32 bits, con la siguiente particularidad: contiene una secuencia de unos seguida de una secuencia de ceros.

Por ejemplo la máscara 255.255.0.0 corresponde a 11111111.11111111.00000000.00000000.

Al determinar el tipo de dirección IP, se puede separar la red de la subred.

Por ejemplo la dirección 164.73.220.3 es de tipo B, por lo tanto la dirección de red es 164.73 y la subred y el host son 220.3.

La máscara indica el corte entre subred y host, dado por los unos y los ceros en la secuencia.

Así, con una máscara 255.255.255.0, la subred será la 220 y el host será el 3.

La misma dirección anterior, pero con una máscara 255.255.127.0 indicará una división diferente (en la que hay que pasar a binario los valores para resolverla).

Direcciones reservadas

A continuación se presentan algunas direcciones reservadas:

- Todos los bits en 0: El propio host.
- Dirección de red en 0 . dirección del host: Un host en la misma red.
- Todos los bits en 1: Broadcast en la misma red.
- Dirección de red . dirección de red todo en 1: Broadcast en una red determinada.
- 127. cualquier cosa: Loopback.

Más sobre el protocolo IP

IP no establece por si mismo un algoritmo de ruteo.

Existen algunos algoritmos sugeridos para su uso dentro de Internet:

- RIP: Routing Information Protocol.
- RFC 1009.

IP está diseñado pensando en la interconexión de redes locales, conectadas entre sí conformando una red de amplia cobertura. Esta idea se refleja claramente en los esquemas para las direcciones.

Para la red local el ruteo es en general trivial.

En el caso en que un datagrama deba arribar a un nodo en otra red, alcanza con que llegue a la red destino y estamos en el caso anterior.

Cada computadora de la red mantiene una tabla con las redes que conoce y el nombre del router al que hay que dirigirse para llegar a dicha red.

También tiene un router por defecto para alcanzar las redes que no conoce.

El router tiene a su vez información que le permite determinar a qué otro router dirigir el datagrama, y así hasta llegar al nodo destino.

La tabla de rutas es dinámica. Se parte de una tabla estática predefinida y luego se van agregando entradas a medida que se va aprendiendo.

La forma de aprender es a través de mensajes intercambiados entre los routers mediante el protocolo ICMP (Internet Control Messages Protocol).

Este esquema encaja dentro de los algoritmos de ruteo distribuidos, ya que cada nodo conoce poco de la red global.

Ejercicios

6.4) Determinar red, subred y host para los siguientes casos:

- Dirección IP: 164.73.32.45 , Máscara: 255.255.255.0.
- Dirección IP: 164.73.32.4 , Máscara: 255.255.255.127.
- Dirección IP: 10.1.7.2 , Máscara: 255.255.255.0.
- Dirección IP: 195.1.1.1 , Máscara: 255.255.255.127.

6.5) Mostrar ejemplos de direcciones reales de Internet y las máscaras utilizadas.

6.6) Mostrar un ejemplo real de una red de cada tipo (A, B y C).

6.7) Investigar las alternativas futuras para sustituir el protocolo IP.

6.8) Buscar información de RIP y de la RFC 1009. Clasificar según las categorías vistas los métodos de ruteo allí propuestos.

6.9) Las RFC (Request For Comments) son la documentación más usual de los protocolos de Internet. Relevar la existencia de RFCs relacionadas con los protocolos vistos hasta el momento.

6.10) Averiguar la utilidad de las direcciones broadcast y loopback.

6.11) Conseguir información del protocolo IPX, y de otras alternativas para capa 3.

Estudio de caso: dos empresas comunicadas por una línea Data Express

Cada empresa tiene su LAN.

Las dos LANs son similares: se utiliza un switch al que se le escalonan hubs hasta cubrir todos los puestos minimizando el tendido. La velocidad es de 100 Mbps.

En la LAN de la empresa A se utiliza el espacio de direcciones falso 10.1.7.x, máscara 255.255.255.0. En la Lan de la empresa B se utiliza el espacio 10.1.8.x con la máscara 255.255.255.0.

Se contrata una línea Data Express de 64 Kbps entre los dos puntos. A la salida RJ11 de la línea, se conecta una DTU. Este dispositivo cumple un papel similar al del módem, pero para líneas digitales. La DTU se comunica con un conector DB25 al router. El router está conectado al switch. Esto es simétrico en las dos redes.

Los routers se ven a sí mismos como 10.1.7.249 y 10.1.8.249 respectivamente.

A su vez, una de las empresas tiene una salida a Internet por medio de un canal inalámbrico a un ISP[11].

Todas las computadoras de las dos redes tienen salida a Internet vía un proxy, que es la única máquina con una dirección real de Internet.

En esta máquina corre un firewall por motivos de seguridad.

Estudio de caso: salida doméstica a Internet

Un PC doméstico se conecta ocasionalmente a Internet para bajar y enviar correos y para navegar por WWW básicamente.

En estos casos, a menos que el uso sea importante, la solución menos costosa es la de una conexión discada con un ISP.

La computadora tiene un módem conectado a su bus, y el mismo está conectado a la línea telefónica. Usualmente además se conecta un teléfono de modo de poder usarlo mientras la computadora no está conectada vía módem. El conector del módem es un RJ11.

Se disca a un ISP, quien cuenta con una batería de módems asociadas a algunas líneas rotativas, y a su vez el ISP tiene conexión a Internet, por ejemplo a través de una ADSL.

Por lo tanto, mientras dura la conexión, a la computadora doméstica se le asigna una dirección IP y es parte de Internet, ya que accede vía el ISP.

Usualmente el ISP cobra por tiempo de conexión.

Ejercicios

6.12) Averiguar acerca del funcionamiento de un firewall.

[11] Internet Service Provider.

6.13) Esquematizar detalladamente los casos anteriores.

6.14) Averiguar características y costos de las siguientes soluciones de conectividad:

- ADSL
- Línea Data Express
- Línea Frame Relay
- Conexión Inalámbrica

6.15) Averiguar información y utilidad del protocolo DHCP, que permite la asignación de direcciones IP dinámicamente.

6.16) Averiguar qué es ISDN y los beneficios de esta tecnología para la conexión doméstica de equipos a Internet.

Capítulo 7: Capa de transporte

La capa de red deja resuelto (a través de los servicios que provee a la capa de transporte) el problema de comunicar dos máquinas cualesquiera dentro de la red, sean adyacentes o no en el grafo de la misma.

Al intercambiar información entre dos hosts de la red, la misma llega hasta la capa 3 de todos los nodos intermedios por los que dichos paquetes pasan. Sólo pasan a capa 4 del equipo que debe recibirlos.

Por esta razón, se dice que las capas 1 a 3 "residen en la red" y que las capas 4, 5, 6 y 7 "residen en el host".

En general la capa 4 es el nivel básico que se ejecuta en el host, y por eso varios autores indican que es el corazón de la jerarquía de protocolos. Reforzando esta idea, se encuentran redes (por ejemplo ARPANET/Internet) y modelos de redes que no contemplan lo que en el modelo OSI corresponde a los niveles 5 y 6.

La misión de la capa 4 o de transporte es proveer un servicio de transporte de datos entre el nodo origen y el nodo destino de forma confiable, independiente de la red sobre la que dicha información se transmitirá.

La capa de transporte provee servicios a la capa de sesión, y los mismos pueden ser o no orientados a conexión.

Los programas que implementan la capa de transporte se llaman entidades de transporte[12].

Las entidades de transporte intercambian TPDUs: Transport Protocolol Data Units.

Los servicios de la capa de transporte son accedidos desde niveles superiores a través de direcciones de transporte.

El servicio provisto por la capa 4 es similar al provisto por la capa 3. La diferencia fundamental es que el nivel de red, en el caso de las redes WAN, es manejado por las compañías de comunicaciones. Aun con los servicios de red orientados a conexión, la red está sujeta a errores, a la posibilidad de perder paquetes o necesitar reinicializaciones. El nivel de transporte debe aislar a los procesos usuarios de estos servicios de estos problemas de confiabilidad.

La presencia del nivel de transporte permite establecer una interfaz para la programación de las aplicaciones, independiente de las características de la red (LAN o WAN, servicio con o sin conexión, etc.).

Otro de los cometidos del nivel de transporte es lograr un nivel de QOS (Quality of Service) a brindar a los niveles superiores, independiente de la calidad del servicio de red. La calidad del servicio de red puede estar fuera del control del usuario.

Afinando lo anterior, la capa de transporte resuelve los siguientes problemas:

- Control de errores y secuenciamiento.
- Direccionamiento.
- Control de flujo.
- Establecimiento y mantenimiento de conexiones.

[12] Esto es válido para todo nivel.

Control de errores

La capa de enlace realiza control de errores, ya que uno de sus cometidos es entregar las tramas sin errores dentro de una LAN. Ese es justamente el punto: el control de error realizado a nivel de enlace se debe complementar con otros controles, pues se está pasando de considerar un canal (en capa 2) a considerar la red completa (capa 4).

Dicho de otra manera, cuando se intercambian tramas entre dos puntos de la red, a nivel de capa 2, se sabe que los mismos pertenecen al mismo canal, generalmente a la misma LAN. Cuando se intercambian paquetes entre dos puntos de la red, a nivel de capa 4, se sabe que los mismos pertenecen a la misma red, pero en general no al mismo canal ni la misma LAN. Entonces, se deben hacer más controles debido a que intervienen nuevas fuentes de error.

En la capa de transporte es necesario especificar la dirección destino. Si hay conexión, entonces se debe establecer y gestionar la misma. También pasa a ser importante la capacidad de almacenamiento de la red.

Ejercicios

7.1) Investigar la utilidad de los CRC y checksums, y en qué medida pueden aportar al control de errores en capa 4.

7.2) Investigar los protocolos de detección de errores utilizados en TCP/IP a nivel de transporte.

Direccionamiento

Cuando un proceso de usuario desea establecer una conexión con un proceso de usuario remoto, debe especificar con quién se quiere conectar.

Para esto se utilizan las direcciones de transporte, TSAPs: Transport Service Access Points. Los procesos de usuario se ligan a TSAPs para escuchar solicitudes.

La forma en que los procesos se ligan a las TSAPs depende del sistema operativo y de la implementación.

Hay varias formas de conectarse a los procesos de transporte. A continuación se analizan 3 de ellas, bajo un ejemplo de un servicio que proporciona día y hora al resto de la red.

La aplicación trabaja en la modalidad cliente/servidor. A efectos del análisis que sigue se supone el cliente corriendo en la máquina A y el servidor corriendo en la máquina B.

Conexión a proceso de transporte conociendo su dirección

Un proceso servidor que proporciona día y hora se une al TSAP 122 en la máquina B, para esperar una solicitud de conexión. La forma en que se liga el proceso al TSAP depende, como ya se indicó, del sistema operativo.

Un proceso en la máquina A necesita averiguar día y hora, por lo que emite un pedido de conexión. Su TSAP es el 6. Emite entonces el pedido de conexión especificando su TSAP 6 como origen y el TSAP 122 como destino.

La entidad de transporte que se encuentra en A, selecciona un NSAP[13] de su máquina, y establece una conexión de red, por ejemplo un circuito virtual X.25, entre ambas. Mediante el empleo de esta conexión puede comunicarse con la entidad de transporte ubicada en B.

La entidad de transporte de A pasa a la entidad de transporte de B el pedido de conexión entre el TSAP 6 de A y el TSAP 122 de B.

La entidad de transporte en B pasa al proceso corriendo en B una solicitud de conexión.

Entonces el proceso corriendo en B queda en condiciones de pasar día y hora al proceso corriendo en A.

Notar que la conexión de transporte va de TSAP a TSAP, y la conexión de red va de NSAP a NSAP.

Este procedimiento es correcto, pero hay un problema importante a resolver:

- ¿ Cómo el proceso de usuario en A llega a saber que el servidor de día y hora está unido al TSAP 122 ?

Una posibilidad es que el proceso haya estado así durante años y todos los procesos se hayan ido gradualmente enterando.

En este modelo, los servidores tienen TSAPs fijos. Sin embargo, en general los procesos de usuario existen por corto tiempo y no tienen una dirección que se pueda conocer anticipadamente.

Además, tener a cada proceso activo y escuchando una dirección puede llegar a ser un gran desperdicio.

Conexión a proceso de transporte utilizando un servidor de procesos

Una alternativa para solucionar lo anterior, introducida en ARPANET, es el servidor de procesos.

En lugar de que todos los servidores tengan que escuchar un TSAP bien conocido, cada máquina que desee ofrecer un servicio a usuarios remotos, tiene un servidor de procesos o registrador, a través del cual se hace la solicitud de todos los servicios.

Siempre que el servidor de procesos esté inactivo, escucha en un TSAP bien conocido. Los usuarios potenciales de cualquier servicio deben comenzar por solicitar conexión especificando el TSAP del servidor de procesos.

Una vez que la conexión quedó establecida, el proceso usuario transmite un mensaje al servidor de procesos indicándole el programa que desea correr (por ejemplo el que da día y hora).

El servidor de procesos selecciona un TSAP inactivo y crea un nuevo proceso, indicándole que escuche en el TSAP seleccionado. Luego termina la conexión. Los dos procesos han quedado comunicados y el continúa escuchando su bien conocido TSAP.

Este mecanismo funciona bien. Sin embargo, hay servicios que existen en forma independiente del servidor de procesos. Por ejemplo un servidor de archivos necesita correr sobre un hardware específico (máquina con discos adecuados), y no puede ser creado sobre la marcha.

[13] NSAP: Network Service Access Point. Es la forma de acceder la capa 4 a los servicios de capa 3.

Conexión a proceso de transporte utilizando un servidor de nombres

El problema anterior se soluciona usualmente utilizando un servidor de nombres, llamado también servidor de directorio.

Para determinar la dirección del TSAP correspondiente a un servicio dado (por ejemplo día y hora) un usuario establece una conexión con el servidor de nombres, que escucha un TSAP bien conocido. Así, el usuario transmite un mensaje identificando el nombre del servicio, y el servidor devuelve un nombre con la dirección del TSAP. Posteriormente, el usuario libera la conexión con el servidor de nombres y establece una nueva con el servicio deseado.

Cada vez que se crea un nuevo servicio se debe registrar en el servidor de nombres.

Ejercicio

7.3) Proponer un ejemplo de aplicación cliente servidor, y proponer esquemas de comunicación para las tres alternativas estudiadas antes.

Control de flujo

Al comunicar dos computadoras, las mismas pueden estar trabajando a velocidades diferentes. Así, podría darse el problema de que un transmisor rápido sature a un receptor lento al enviarle datos a mayor velocidad de la que este puede procesarlos.

Una posible solución a este problema de sincronización la dan los protocolos de parada y espera:

- El emisor envía una trama y luego espera un acuse de recibo (ACK[14])
- El receptor luego de pasar la trama a la capa de red envía un ACK.
- Si se agota el tiempo de espera por un ACK, el emisor reenvía.

Enseguida surgen varios problemas a resolver asociados a este mecanismo:

- Puede perderse una trama
- Puede perderse un ACK
- Puede llegar el ACK luego de reenviada la trama
- Puede llegar dos veces la misma trama

Ejercicio

7.4) Esquematizar y estudiar los casos problemáticos anteriores.

Protocolo de ventana deslizante

Este protocolo es básicamente una generalización del protocolo de parada y espera, en el que cada trama está numerada.

En todo momento el emisor mantiene un conjunto de números de secuencia, que corresponden a las tramas que tiene permitido enviar. Cuando se emite una trama, el

[14] De acknowledge

tamaño de la ventana del emisor se achica en 1. Cuando se recibe un ACK el tamaño de la ventana del receptor se agranda en 1.

El receptor también mantiene una ventana que se refiere a la cantidad de tramas que puede aceptar, y la maneja en forma simétrica a la del emisor.

Las ventanas de emisor y receptor no tienen por qué ser del mismo tamaño y además su tamaño puede ser variable.

El protocolo de parada y espera simple es un caso particular de este, donde la ventana del emisor tiene tamaño 1 y la del receptor tamaño 0.

Ejercicio

7.5) Proponer y analizar diferentes ejemplos de uso del mecanismo de ventana deslizante. Buscar los casos en que se toman acciones para controlar efectivamente el flujo (detener la emisión).

Establecimiento y mantenimiento de conexiones

La capa de transporte es un lugar donde frecuentemente se utilizan servicios orientados a conexión. La ventaja fundamental de este tipo de servicios es que a lo largo de la conexión se mantienen parámetros de QOS (calidad de servicio). La desventaja pasa por las dificultades que significan la gestión de la conexión y la desconexión.

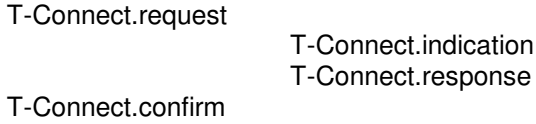
Las primitivas involucradas en la gestión de conexiones, de acuerdo al modelo OSI, son (pueden haber variantes en la realidad pero estas recogen la esencia del problema):

- T-Connect.request: pedido de conexión.
- T-Connect.indication: recepción de pedido de conexión.
- T-Connect.response: emisión de respuesta a pedido de conexión.
- T-Connect.confirm: recepción de respuesta a pedido de conexión.
- T-Disconnect.request: pedido de desconexión.
- T-Disconnect.indication: recepción de pedido de desconexión.
- T-Data.request: pedido de datos.
- T-Data.indication: recepción de datos.

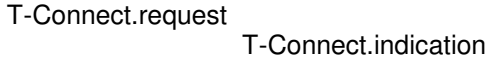
También hay primitivas para datos que pasen con prioridad sobre el resto (T-ExpeditedData).

Algunas secuencias interesantes (el tiempo avanza hacia abajo en los esquemas, y cada extremo es un participante de la conexión):

- Establecimiento de conexión:



- Conexión rechazada por el usuario al que llaman (la diferencia con el anterior está en los parámetros pasados en las primitivas).



T-Connect.response
T-Connect.confirm

- Conexión rechazada por la capa de transporte:

T-Connect.request
T-connect.indication

- Liberación normal de una conexión:

T-Disconnect.request
T-Disconnect.indication

- Liberación simultánea en los dos extremos:

T-Disconnect.request T-Disconnect.request

- Liberación iniciada en la capa de transporte:

T-Disconnect.indication T-Disconnect.indication

- Transferencia de datos normales:

T-Data.request
T-Data.indication

Los pedidos (request) en general esperan hasta un máximo de tiempo por una respuesta. La expiración de este tiempo de espera se denomina TimeOut. La ocurrencia de TimeOuts trae como consecuencia un cierto número de reintentos o un reporte de error.

Son varios los problemas que se podrían dar si se pierden requests o indications en la red. También si los mismos se retrasan, ocurren TimeOuts y luego hay peticiones repetidas, o fuera de sincronización.

Ejercicios

7.6) Analizar posibles casos problemáticos al perderse paquetes en la red.

7.7) Analizar posibles casos problemáticos al llegar paquetes retrasados, especialmente luego de la ocurrencia de TimeOuts.

Liberación de las conexiones

El problema de liberación de la conexión merece especial atención. Hay que notar que en este caso el problema es más sencillo que al establecer la conexión, pero potencialmente existe una cantidad importante de problemas.

El problema principal pasa porque al momento de solicitar la desconexión, es imposible un total convencimiento de las dos partes. Este problema es isomorfo al “problema de los ejércitos“:

> 3 unidades del ejército A se encuentra en un valle, y en dos extremos del mismo hay 2 unidades (en cada extremo) correspondientes al ejército B. Como $2 < 3$, si cualquiera de las unidades de B decidiera atacar, perdería. Pero si las 4 unidades de B lograran sincronizarse para el ataque, ganarían.

Pensando en mandar mensajeros de un extremo a otro, surge claramente que no existe forma de que las unidades puedan pactar una hora de ataque con 100% de seguridad. Si es posible establecer niveles crecientes de seguridad, pero no es alcanzable el 100 %, ya que siempre está latente la posibilidad que al último mensajero lo capturen.

El problema de liberación de la conexión es similar, y la solución más razonable es cortar la conexión unilateralmente luego de un número preestablecido de confirmaciones (que bien puede ser también una o ninguna).

Ejercicio

7.8) Diseñar un protocolo de establecimiento y liberación de conexión, adoptando para esto último alguna política razonable respecto a la cantidad de confirmaciones para la liberación de la conexión.

Multiplexión

Hay dos variantes:

- Multiplexión ascendente:
    - Esta técnica se utiliza para prevenir altos costos de comunicaciones por una estructura tarifaria que sobra por tiempo de conexión y no sólo por tráfico.
    - En este caso se inicia una única conexión de red para establecer varias conexiones de transporte entre dos nodos.
    - El inconveniente es que la performance puede ser mala si se asignan muchas conexiones. Por otra parte si se asignan pocas conexiones el costo es alto.
    - Utilizando esta técnica se puede estar desaprovechando por ejemplo la administración de circuitos virtuales que proporciona X.25.
- Multiplexión descendente:
    - Esta técnica se utiliza para lograr disponer de un ancho de banda adecuado, el que se puede ver afectado por el control de flujo con ventana, si los acks demoran en llegar.
    - La idea es utilizar más de una conexión de red por cada conexión de transporte.

Recuperación de caídas

Las caídas pueden darse en distintos niveles: en la red, en alguno de los niveles que están ejecutando en el host, o en el host en su totalidad.

En principio, si el que cae es el receptor, alcanzaría que cuando estuviera disponible nuevamente, el transmisor retransmitiera todos los mensajes de los cuales no haya habido acuse.

Sin embargo esto no es tan sencillo, ya que en el nivel de transporte ocurren dos eventos, el envío del acuse y el pasaje del mensaje al nivel superior.

Estos dos eventos deberían ser una sola operación indivisible, sin embargo no lo son.

Entonces, existe la posibilidad que el transmisor no retransmita mensajes de los cuales hay acuse, pero que estos no hayan sido pasados al nivel superior, perdiendo entonces información.

El protocolo TCP

Las características principales de TCP son:

- Es un protocolo orientado a conexión.
- Es confiable, debido a lo anterior.
- Byte-stream: flujo de bytes.

Básicamente TCP toma un mensaje del nivel superior (nivel 5 en OSI, nivel de aplicación en ARPANET[15]), lo separa en mensajes TCP y los envía hasta el otro nodo. En el otro nodo reconstruye el mensaje original y se lo pasa al nivel superior.

Está pensado para enfrentarse a retransmisiones, mensajes que llegan fuera de secuencia y duplicaciones retardadas.

Maneja el concepto de puerto (port) para identificar las conexiones (TSAP). Los puertos origen y destino se guardan en la información de control.

Utiliza un control de flujo por ventana de tamaño variable.

La confiabilidad se logra utilizando el mecanismo “Positive Acknowledgment with re-Transmission. La unidad de transmisión es el segmento. Cada segmento tiene un checksum para verificación. Si un mensaje llega OK se transmite un ack positivo en respuesta, Si en el emisor ocurre un timeout, retransmite.

El protocolo UDP

Es la contraparte sin conexión de TCP.

Básicamente puede considerarse como una interfaz de programación del protocolo IP, dado que solamente se encarga de “intentar” enviar un mensaje.

De todas maneras maneja la idea de puerto como forma de identificar los procesos que se están comunicando entre dos nodos (TSAP).

Permite acceso directo a un servicio de entrega de datagramas, logrando una transmisión con un mínimo overhead.

En la información de control de guardan los puertos origen y destino, y también el largo y el checksum.

Cuando el volumen de datos es pequeño, el overhead de negociar la conexión puede ser mayor que transmitir todos los datos de nuevo. En este caso puede ser UDP una alternativa válida. Las aplicaciones basadas en consulta/respuesta (query/response) también son candidatas a ser resultas sobre UDP.

Por último, aplicaciones que realizan sus propios controles de confiabilidad y no requieren de la capa de transporte, pueden evitar controles dobles si usan UDP, ganando en eficiencia.

[15] Modelo TCP/IP y Modelo ARPANET es lo mismo.

Capítulo 8: Capa de sesión

Es un nivel introducido con el modelo OSI, ya que los modelos previos de redes carecían del mismo. Es un nivel de relativamente pocas prestaciones.

Si bien pueden concebirse servicios sin conexión, las características de este nivel hacen que sea más apto para implementar servicios con conexión.

Los servicios brindados en el nivel de sesión son:

- Intercambio de datos.
- Administración del diálogo.
- Sincronización.
- Administración de la actividad, manejo de sesiones.
- Reporte de excepciones.

Estos servicios están implementados a través de primitivas, que establecen los servicios que se prestan a la capa superior.

Administración del diálogo

Dependiendo del sentido del intercambio de datos, la transmisión de datos puede ser:

- Simplex: se intercambian datos en un solo sentido.
- Half-Duplex: se intercambian datos en los dos sentidos, pero no simultáneamente. Esto se arbitra con el paso de un testigo o token[16].
- Full-Duplex: se intercambian datos en los dos sentidos simultáneamente.

En principio todas las comunicaciones en OSI son Full Duplex. Sin embargo, se admite la posibilidad de sesiones half-duplex. Para arbitrar el sentido de la comunicación se utiliza un mensaje llamado Data-Token.

Este testigo pasa de un extremo al otro de la conexión a través de primitivas especiales. También hay primitivas para solicitar el token al otro nodo.

Sincronización

Este servicio permite que en caso de ocurrir errores se pueda retornar a un estado anterior conocido.

La sincronización se implementa mediante puntos de resincronización.

Algunos ejemplos donde es útil:

- Pasaje de archivos grandes dentro de la red.
- Impresión remota.

Administración de la actividad

Esta facilidad permite definir actividades en el marco de una sesión.

El significado de las sesiones y actividades como unidades lógicas está asignado por el usuario del servicio.

[16] Al pasar el token se está cambiando el sentido de la comunicación.

Se brindan facilidades para comenzar, finalizar, suspender y retomar actividades.

Por ejemplo una transferencia de archivos suspendida para comenzar otra actividad y luego retomada, sin pérdida de información.

Reporte de excepciones

Una excepción es la ocurrencia de un evento no deseado o no esperado dentro de los sucesos normales en la comunicación.

Sirve para que el nivel de sesión comunique al nivel superior errores ocurridos en los niveles inferiores o en el propio nivel.

Si bien el manejo de excepciones se da a todo nivel, el nivel 5 es donde más usualmente pueden encontrarse primitivas específicas de manejo de los mismos.

Interfases de programación

La programación de aplicaciones distribuidas y aplicaciones de red requiere del uso de bibliotecas específicas para manejo de las comunicaciones.

A nivel de capa 4, es posible realizar programación utilizando las bibliotecas derivadas de las primitivas de manejo de sockets, por ejemplo. Sin embargo, las facilidades de programación sobre la red con un nivel razonable de abstracción se encuentran en capa 5.

Un ejemplo son las RPCs: Remote Procedure Calls.

El mecanismo de RPCs

El mecanismo de RPCs se asocia a capa de sesión. Esto se puede considerar así al tener en cuenta que las implementaciones de TCP/IP esta capa trabaja por encima de TCP, que corresponde al nivel de transporte.

RPC es una interfaz de programación para aplicaciones en reda un nivel superior a la que brindan los sockets.

Usualmente, RPC se implementa como una biblioteca de funciones, usualmente en lenguaje C, que permiten implementar aplicaciones cliente, aplicaciones servidor y el protocolo que las comunica, de forma más amigable, y con un mayor nivel de abstracción en cuanto a la programación, que usando directamente los sockets.

La semántica del RPC define las acciones a tomar al encontrar caídas en cliente y servidor.

La secuencia a seguir en el desarrollo de las aplicaciones distribuidas usando RPCs es:

- Diseñar la arquitectura de la aplicación.
- Diseñar el protocolo de comunicación entre cliente y servidor.
- Implementar servidor.
- Implementar cliente.
- Implementar otros componentes de la arquitectura si los hubiera.

A nivel de programación, las RPCs se comportan como llamadas a subprogramas, pero estos residen en máquinas remotas de la red. El intercambio de datos se implementa a

través del pasaje de parámetros, pero con la complicación adicional de que no se está trabajando sobre la misma memoria (son máquinas diferentes).

Ejercicios

8.1) Buscar ejemplos de programas RPC especialmente sus protocolos de alto nivel, que son compilados por la herramienta rpcgen.

8.2) Buscar ejemplos de aplicaciones que trabajando en capa 5, permiten restablecer conexiones a Internet desde el punto en que habían sido dejadas, siendo de gran utilidad para bajar programas grandes.

Capítulo 9: Capa de presentación

En la capa de presentación, se manejan conversiones sintácticas sobre los datos:

- La representación local de los datos (sintaxis abstracta) puede no coincidir con la representación utilizada en la transferencia (sintaxis concreta).
- La capa de presentación se encarga de las conversiones necesarias en el emisor y el receptor.

En este nivel se resuelven entonces tres problemas:

- Conversión de datos entre diferentes formatos.
- Compresión de los datos.
- Encriptación.

Conversión de datos entre diferentes formatos

Ya resuelta la conectividad entre dos hosts, aun persiste el problema de intercambiar información entre arquitecturas diferentes.

La representación interna

La computadora sólo puede almacenar en memoria ceros o unos. El problema de la representación interna consiste en aportar representaciones para los diferentes tipos de datos en términos de ceros y unos.

Hay muchos detalles al respecto, que pueden encontrarse por ejemplo en [Tan96]. A continuación se presenta un resumen del tema.

Para representar los naturales, se escriben en base dos, en binario. Con esto se logra una representación en términos de ceros y unos, y por lo tanto es posible guardarla en memoria.

Para representar los enteros, se emplea (hay otras técnicas) un bit para el signo y se expresa el valor absoluto (que es un natural) pasando como antes a binario.

Cuantos más bytes se destinen, mayor será el rango numérico abarcado. Por ejemplo, con 16 bytes, se pueden representar los naturales entre 0 y 32767 ($2^{15} - 1$).

Para los carácteres (letras, dígitos y símbolos), se emplea una asociación biunívoca con los naturales, dada por ejemplo por la tabla ASCII, y entonces el problema de representar un carácter se reduce al de representar su código ASCII, que por ser un natural ya está resuelto.

Para los strings (palabras) se puede (hay variantes) almacenar uno a uno sus carácteres y luego agregar un carácter especial de fin.

Los números reales, al constituir un conjunto denso, no son representables, pero se puede representar una discretización de los mismos utilizando la representación de punto flotante. En ella se fija una base *b*, y luego se asocia a cada número *r* una terna *(s,m,e)*, correspondiente a signo, mantisa y exponente, tal que se cumple que $r = s.b^e$. Luego, el problema de guardar s (1 bit), y m y e (Naturales) ya está resuelto.

Representación interna en diferentes arquitecturas

Hay varios criterios a definir a la hora de decidir la representación interna. Así, las diferentes arquitecturas deciden arbitrariamente algunos puntos, por ejemplo:

- Qué tabla usar para codificar caracteres (ASCII, EBCDIC, etc.)
- Little-endian/big-endian, que implica recostar a la derecha o a la izquierda los binarios .
- Diferentes formas de armar los strings.
- Cuántos bytes destinar para los enteros.
- Cuántos bytes destinar para mantisa y exponente.

A la hora del intercambio de datos entre computadoras con diferentes criterios para la representación interna, surgen problemas.

Soluciones

Hay dos alternativas para resolver estos problemas:

- El emisor convierte la representación local a una canónica y luego el receptor convierte la representación canónica a su representación.
- Convertir la representación local a la representación de la máquina destino directamente.

Al surgir una nueva forma de representación, en el primer caso hay que aportar dos algoritmos de conversión desde y hacia la representación canónica. En el segundo caso hay que aportar n algoritmos de conversión, desde y hacia cada una de las arquitecturas existentes.

Ejemplo: XDR

XDR (eXternal Data Representation) es un protocolo de capa 6 cuyo objetivo es resolver estos problemas de representación interna, aportando una solución basada en una representación canónica.

Está implementado como una biblioteca de C. Las rutinas, que funcionan como filtros, convierten de la representación local a la canónica y viceversa.

Ejercicios

9.1) Buscar ejemplos de representación interna de datos.

9.2) Investigar los formatos de datos canónicos utilizados por XDR.

9.3) Comparar los ejemplos anteriores con la conversión a la representación canónica propuesta por XDR.

Compresión de los datos

Algunas aplicaciones requieren grandes volúmenes de transferencia de datos, y entonces es importante el tema de la compresión. Por ejemplo, transmitir imágenes y video sobre una red implica un intercambio de un importante volumen de datos.

La efectividad de la compresión depende del algoritmo de compresión elegido y de las características particulares de los datos a comprimir.

La función de compresión debe tener inversa, que es la que permite al receptor descomprimir.

Ejercicios

9.4) Conseguir información del algoritmos de Huffmann, y comentar en qué medida el mismo puede ayudar a la compresión.

9.5) En algunos casos, la compresión se realiza manual y explícitamente, por ejemplo al enviar adjuntos comprimidos por email. Comentar este ejemplo.

9.6) Conseguir información de las herramientas de compresión de Unix y en qué medida pueden resolver el problema de la compresión a nivel de capa 6.

Encriptación

La encriptación pretende evitar la interpretación de los mensajes eventualmente interceptados. Este problema es de creciente importancia, ya que cada vez con mayor frecuencia se maneja en las redes información altamente confidencial, que constituye incluso ventajas competitivas para algunas organizaciones.

Hay muchas técnicas y algoritmos. En su visión más genérica, la idea es que antes de transmitir *x*, se encripta, transmitiéndose entonces *f(x).* El receptor aplica f^{-1} y $f^{-1}(f(x))=x$, obteniendo el dato original. Si alguien captura *f(x)* no puede –a menos que logre romper el cifrado- obtener *x*.

Un ejemplo muy simple

La forma más simple de cifrado es la sustitución monoalfabética. En este caso la función de cifrado asigna de forma biunívoca a cada letra otra letra. De ese modo es posible cifrar mensaje al cifrar letra a letra.

Este método es muy fácil de romper, ya que las tablas de frecuencias relativas de las letras son conocidas en varios lenguajes. Así, mirando la “forma” de la curva de frecuencias relativas, es fácil romper este tipo de cifrado.

Hay otras variantes descriptas en [Tan96].

Criptografía de clave privada

Se dice que es criptografía simétrica. Emisor y receptor se ponen de acuerdo en una clave secreta. El emisor encripta con la clave, el texto encriptado vieja por la red, y el receptor desencripta con la clave, obteniendo el mensaje original.

Un ejemplo es el algoritmo DES. El algoritmo es público, la fortaleza se basa en el cuidado de las passwords.

Se requiere de un acuerdo previo entre las partes.

Criptografía de clave pública

El principal objetivo de diseño de este método fue crear confianza en una red pública donde participan millones de personas.

En la criptografía de clave pública hay algunas diferencias respecto al sistema de clave privada:

- No se transfiere la clave privada.
- No se requiere de un acuerdo previo entre las partes.
- Proveen el servicio de no repudio.
- Para n usuarios, se requieren 2n claves.

Los sistemas de clave privada:

- Son mucho más rápidos.
- Se requiere un acuerdo previo entre las partes.
- La clave secreta se transfiere.
- Para n usuarios, se requieren n.(n-1)/2 claves para operar.

Hay varias formas de intercambiar texto cifrado haciendo uso de la encriptación de clave privada:

- El receptor la envía el emisor su clave pública. El emisor cifra con la clave pública del receptor. El receptor descifra con su propia clave privada.

- El emisor encripta con su clave privada y envía al receptor el mensaje encriptado y su clave pública. El receptor descifra el mensaje con su propia clave pública.

Certificados digitales y PKI

Un certificado digital permite vincular una identidad física (organización, persona, máquina, etc.) con una identidad digital.

Permiten que quien reciba un mensaje pueda certificar la autenticidad de la clave pública de quien lo envía. Un certificado digital se puede guardar en un PC, en una tarjeta inteligente, en un teléfono celular, etc.

Una arquitectura de PKI resuelve el problema de identificar a una persona o entidad, sin un conocimiento o registro previo. Brinda confidencialidad e integridad a las transacciones. Provee buenos servicios de autenticación y de no repudio. Por último, permite a las organizaciones establecer relaciones confiables entre dominios seguros.

Ejercicios

9.7) Esquematizar detalladamente el funcionamiento de la criptografía pública y de clave privada.

9.8) Investigar de qué manera se logra brindar el servicio de no repudio. Como guía, recordar que el único que puede no negar su identidad es el poseedor de una clave privada. Entonces, si se envía un mensaje codificado con la clave privada, cualquiera podrá abrirlo (basta con acceder a la clave pública), pro el emisor no podrá negar el envío.

9.9) Listar y describir los mecanismos de seguridad vistos según las distintas capas del modelo OSI. Por ejemplo los firewalls y su funcionamiento filtrando a diferentes niveles.

Capítulo 10: Capa de aplicación

A esta capa corresponden todas las aplicaciones con acceso al ambiente de red.

Algunos protocolos asociados a aplicaciones comunes son.

- Telnet: permite la emulación de terminales de acceso a un host remoto.
- SMTP[17], POP3, IMAP: Protocolos de intercambio de correo.
- FTP[18]: Protocolos para transferencia de archivos.
- HTTP[19]: Protocolos para publicación/acceso a páginas Web.
- Gopher: Une versión anticuada del anterior.
- NFS[20]: Sistema de archivos en red.

Ejercicio

10.1) Completar el relevamiento de aplicaciones anterior, agregando las aplicaciones estándar más conocidas.

[17] Simple Mail Transport Protocol.
[18] File Transfer Protocol.
[19] Hyper-Text Tag Protocol.
[20] Network File System.

Capítulo 11: La familia de protocolos TCP/IP

TCP/IP refiere a un conjunto de protocolos de comunicaciones. Originados en la red ARPANET, estos protocolos han sido progresivamente adoptados como forma nativa de comunicación, primero en Unix y luego también en Windows. Actualmente conforman la base de Internet, desde el punto de vista técnico.

Los protocolos principales de esta familia: IP y TCP ya han sido descriptos en sus correspondientes capas en el modelo OSI. Sin embargo, interesa en este capítulo repasar la relación entre TCP/IP e Internet, así como analizar el modelo de redes que esta familia de protocolos lleva implícito.

TCP/IP e Internet

Los siguientes son los hechos más relevantes que relacionan a los protocolos de TCP/IP con lo que es hoy la red Internet:

- En 1969 se crea la red experimental de conmutación de paquetes ARPANET, de la DARPA.
- En 1975 ARPANET pasa a ser una red operacional, y comienzan a crearse los protocolos básicos de TCP/IP.
- En 1983 los protocolos TCP/IP se adoptan como estándares militares y todos los hosts de ARPANET se convierten a los nuevos protocolos. Para facilitar dicha conversión, se implementa TCP/IP sobre BSD Unix[21].
- En 1990 deja de existir oficialmente ARPANET para dar lugar a Internet.
- Actualmente, el término Internet es aun más amplio, abarcando muchas redes interconectadas en todo el mundo.

Afinando un poco las definiciones presentadas antes:

- Una internet es una colección de redes físicamente separadas, interconectadas por un protocolo común, de forma de comportarse lógicamente como una sola red.
- La Internet es la colección mundial de redes interconectadas, que creció a partir de ARPANET y usa el protocolo IP (Internet Protocol) para unir dichas redes en una sola red lógica.

TCP/IP es entonces la familia de protocolos necesaria para acceder a Internet. La popularidad de TCP/IP radica en haber resuelto el problema de la comunicación de datos a nivel mundial.

Algunos aspectos importantes de los protocolos TCP/IP son:

- Protocolos estándar abiertos: Entonces, al ser ampliamente soportado, TCP/IP es ideal para comunicar computadoras cuyo hardware y/o software son diferentes.
- Independencia del hardware de red: Entonces, TCP/IP es adecuado para integrar diferentes clases de redes. TCP/IP corre sobre redes Ethernet, token ring, líneas discadas, redes X.25, etc.
- Provee un esquema de direccionamiento común.
- Estandarización de protocolos de alto nivel.
- Los protocolos de TCP/IP están orientados a comunicación de redes heterogéneas, y esa es la causa de su naturaleza abierta: están

[21] Esta era la versión de Unix de la Universidad de Berkeley.

disponibles para todos y son cambiados por consenso (no unilateralmente).
- La publicación de la documentación de TCP/IP se basa en tres vías:
  - MIL STDs (Estándares Militares).
  - IENs (Internet Engineering Notes).
  - RFCs (Request For Comments).

Ejercicio

11.1) Conseguir en Internet información de la historia de Internet. Hay sitios con abundante información al respecto.

11.2) Bajar de Internet alguno de los muchos glosarios que hay de términos relacionados con Internet.

Un modelo para describir TCP/IP

La terminología del modelo OSI ayuda a describir TCP/IP, pero para dar un análisis más estricto, hay que proponer un nuevo modelo para describir estos protocolos. Una descripción adecuada resulta de considerar 4 capas.

Como en el modelo OSI, los datos bajan y suben por la pila TCP/IP, y a medida que lo hacen se agrega y se quita (y analiza) información de utilidad para cada capa. Esta información de control, en TCP/IP corresponde a cabezales, ya que se ubica antes de los datos a enviar.

Arquitectura de TCP/IP

El modelo considera una arquitectura de 4 capas:

- Capa de acceso a la red:
  - Se compone de un conjunto de rutinas para acceder físicamente a la red. Se corresponde con las capas 1, 2 y parte de la 3 del modelo OSI.
- Capa de Internet:
  - Define el datagrama, controla el ruteo de datos y direccionamiento. Se corresponde con la capa 3 del modelo OSI.
- Capa de transporte host a host:
  - Provee servicio de transporte de datos entre dos puntos. Se corresponde con las capas 4 y 5 del modelo OSI.
- Capa de aplicación:
  - Aplicaciones y procesos que usan la red.
  - Se corresponden con las capas 6 y 7 del modelo OSI.

En la capa inferior, de acceso a la red, en lugar de definir protocolos específico, TCP/IP hace uso de estándares existentes.

La correspondencia entre capas de los modelos OSI y TCP/IP es aproximada.

Ejercicio

11.3) Esquematizar el intercambio de datos en el modelo TCP/IP de la misma forma que fue presentado en el modelo OSI, marcando el pasaje de datos entre diferentes capas.

Protocolos y Estructuras de datos en TCP/IP

Cada capa tiene sus propias estructuras de datos. Teóricamente, las estructuras correspondientes a capas distintas son totalmente independientes. Para ganar en eficiencia, las estructuras de cada capa son diseñadas de modo guardar compatibilidad con las de las capas adyacentes.

Los protocolos fundamentales son IP, TCP y UDP. Estos protocolos junto con el mecanismo de direccionamiento de IP fueron analizados en detalle en la descripción de las capas del modelo OSI.

Aplicaciones sobre TCP se refieren a sus datos como un flujo (stream). Aplicaciones sobre UDP llaman a sus datos mensajes.

TCP llama segmentos a sus datos. UDP llama paquetes a sus datos.

IP ve sus datos como bloques llamados datagramas.

TCP/IP trabaja sobre varios tipos de redes subyacentes. La mayoría de las redes se refieren a los datos a transmitir como paquetes o tramas (frames).

Capa de acceso a la red

En esta capa se define como usar la red para transmitir un datagrama IP. Esta capa debe saber los detalles de la red subyacente. Usualmente es ignorada por los usuarios.

Hay tantos protocolos de acceso como estándares para aspectos físicos de las redes. Al aparecer nuevo hardware de red, se necesita desarrollar también protocolos de acceso a la red.

Las funciones de esta capa son:

- Encapsular los datagramas IP en tramas emitidas por la red.
- Mapear las direcciones IP en direcciones físicas usadas por la red.

Ejemplos de protocolos:

- ARP (RFC 826).
- Transmisión de datagramas sobre Ethernet (RPC 894).

En el sistema operativo, los protocolos de esta capa aparecen como una combinación de device drivers y programas propios del sistema operativo.

Capa de Internet

El protocolo IP es el principal en esta capa, y es medular en todo TCP/IP (RFC 791). Toda comunicación de datos vía TCP/IP usa el protocolo IP.

El protocolo IP resuelve los siguientes problemas:

- Definición de la unidad básica de transmisión en Internet: el datagrama.
- Definición del esquema de direccionamiento para Internet.
- Transporte de datos entre capas de transporte y acceso a la red.
- Ruteo de datagramas a hosts remotos.
- Fragmentación y reensamblado de datagramas.

IP es un protocolo no orientado a conexión:

- No se intercambia información de control antes de comenzar una transmisión.
- IP confía en que de ser necesaria, la comunicación se establezca en capas superiores.
- También confía en las capas superiores para la implementación de métodos de detección y recuperación frente a errores.
- TCP/IP se diseñó para correr sobre la red ARPANET, que era una red de conmutación de paquetes.
- Un datagrama es el formato de paquete definido por el protocolo IP. Un datagrama se compone de datos de información y control.
- IP toma decisiones de ruteo para cada datagrama por separado.

Los routers pasan información entre redes distintas. Un host conectado a más de una red también se puede configurar para que rutee paquetes.

Cuando un router está conectado a varias redes físicas de diferentes características, puede ser necesario dividir el datagrama en piezas más chicas.

Se agregan algunos campos en la información de control para identificar y reensamblar los fragmentos de un datagrama.

El protocolo IP fue analizado en detalle en capítulos anteriores.

El protocolo ICMP (Internet Control Message Protocol) también trabaja en esta capa. Se envían mensajes que resuelven: control de flujo, detección de destinos no alcanzables, redirección de rutas, pings[22] a hosts remotos.

Capa de transporte host-a-host

Los dos protocolos más importantes en este nivel son:

- TCP: Transmission Control Protocol. Provee servicio de entrega de datos confiable, con detección y corrección de errores. Orientado a conexión.

- UDP: User datagram Protocol. Provee un servicio de entrega de datos basado en datagramas, con baja sobrecarga (overhead) y sin conexión.

Los dos protocolos trabajan sobre IP.

Estos protocolos fueron analizados en detalle en capítulos anteriores.

Capa de aplicación

Incluye todo proceso que utilice la capa de transporte para entregar datos.

Hay muchos protocolos de aplicación, coincidiendo con los listados en la capa de aplicación del modelo OSI.

[22] Un ping es una petición realizada a un host tal que el si el mismo contesta, se puede concluir que está alcanzable.

Capítulo 12: El fenómeno Internet desde el punto de vista del usuario

El enfoque de este manual es principalmente técnico, y por esa razón se arriba al concepto de Internet partiendo de un enfoque técnico. Se ha venido presentando un conjunto de conceptos que en forma "bottom-up" van analizando problemas relacionados con las redes, hasta llegar a una definición de Internet.

Sin ser el foco del manual, en este capítulo se pretende aportar algunos elementos de análisis del fenómeno de Internet desde el punto de vista del usuario. La idea es analizar el impacto de Internet en las organizaciones.

Varios datos comentados en este capítulo corresponden al curso del Cr. De Luca de e-Business y a la consultora Tea Deloitte & Touche.

El Web como una nueva herramienta de comunicación

Algunos eventos clave en la evolución de Internet, desde la década del 50-60 hasta nuestros días:

- Creación de una red (ARPANET) que no pudiera ser completamente destruida.
- Herramienta del mundo académico. Gente intercambiando ideas e investigaciones.
- WWW. Interfase Mosaic, nacimiento del browser y del World Wide Web.
- Creación de Motores de Búsqueda. Con ellos es posible indizar contenidos y guiar búsquedas dentro de la creciente complejidad del WWW.
- Sitios Web de organizaciones. Creación de contenido, lugares para visitar.
- Tiendas Virtuales. Posibilidad de comprar cosas vía WWW.
- Portales. Cosas para hacer, comunidades para integrar.

Esta evolución no ignora las posibilidades de la conectividad en general, ni el correo electrónico como nuevo medio de comunicación. Sin embargo muestra al Web como la herramienta de mayor potencial. La llegada del Web es una consecuencia innmediata de la evolución de las herramientas y tecnologías de Internet.

El Web como nueva tecnología ha tenido tiempos récord de adopción. La radio necesitó casi 40 años para alcanzar 50 millones de usuarios, los PCs 15 años, la TV 12 y el Web 4 años.

Estos datos nos muestran el enorme potencial de esta tecnología. Además, se espera que la misma tenga un crecimiento vertiginoso. Los datos que siguen son una muestra de esto[23]:

- El tráfico en Internet se duplica cada 100 días.
- Más de 300 millones de personas en línea a nivel mundial.
- El porcentaje de usuarios conectados crece sostenidamente.
- El comercio en Internet se duplica en menos de un año.
- La red se adoptó más rápido que cualquier tecnología previa.
- El comercio B2C[24] excederá los U$S 240 mil millones para el 2003.
- El comercio B2B[25] excederá los U$S 2960 mil millones para el 2003.

[23] Datos de 2001.

[24] B2C: Business to Consumer (comercio empresa a consumidor).

[25] B2B: Business to Business (comercio empresa a empresa).

- El eCommerce ya es mucho más que vender libros y CDs por la red.

Como se aprecia en las cifras anteriores, el comercio B2B tiene y tendrá un protagonismo mayor entre las transacciones hechas por la red.

Ejercicios

12.1) Relevar, ubicándolos históricamente, los hitos más importantes en la historia de Internet.

12.2) Averiguar datos de conectividad, acceso a Internet y uso de las herramientas más comunes sobre Internet en el mundo, en la región y en el Uruguay.

12.3) Investigar la cantidad de transacciones electrónicas B2C en el mundo, la región y Uruguay. Compararla con los niveles de conectividad. Comentar el problema observado en Uruguay, un país con una excelente infraestructura para conectividad, buena cultura en la población, alta tasa de hogares conectados y un índice llamativamente bajo de comercio electrónico.

12.4) Buscar explicaciones para el hecho observado de un mayor desarrollo de B2B frente a B2C. Investigar las trabas que existen al respecto y las posibles soluciones. Examinar luego el tema desde la perspectiva local.

Una propuesta para marco de análisis

En [WWWIndic] se expone un marco para el análisis del fenómeno de Internet. En el mismo, se separa el problema en 4 niveles, de diferente grado de abstracción.

- Nivel 1: Infraestructura de Internet
  - Comprende el hardware para implementar las soluciones de conectividad y acceso.
- Nivel 2: Aplicaciones de Internet
  - Comprende el software para implementar soluciones de conectividad, puesta en línea de sitios, servicios ofrecidos por los sitios, etc.
- Nivel 3: Intermediarios de Internet
  - Comprende los sitios que interactúan con los internautas pero sin vender productos[26] en línea.
- Nivel 4: Comercio en Internet
  - Comprende los sitios que realizan venta en línea de sus productos.

Cada uno de los actores cuya área de actividad se relaciona con Internet caen dentro de alguna de las categorías.

Luego se construyen y analizan indicadores en base a este modelo.

Ejercicios

12.5) Buscar los indicadores mencionados anteriormente.

[26] En este capítulo se utiliza el concepto amplio de producto, es decir, producto o servicio.

12.6) Comentar en qué medida es cierto que: "... El fenómeno de Internet ha sido como la fiebre del oro. En esta los que más ganancias obtuvieron fueron los proveedores de palas ... En esta explosión del fenómeno de Internet los proveedores de infraestructura han llevado la mejor parte ...

12.7) Conseguir información acerca del índice NASDAQ. ¿ Qué mide ?

12.8) Conseguir información de la caída del NASDAQ. ¿ Cuál fue el impacto si se compara la situación a mediados de 2000 y a mediados de 2001 ?. ¿ Cuál fue el impacto en el Uruguay ?

12.9) Comentar los datos del análisis publicado en 12.5) teniendo en cuenta la caída de las acciones de las empresas punto com según 12.8.

12.10) ¿ Cuál ha sido la perspectiva de las empresas Uruguayas respecto al florecimiento y caída del fenómeno de Internet ? Tome como ejemplo los proveedores de acceso gratuito, que surgieron para luego de unos meses desaparecer o reconvertirse.

Algunos sitios ejemplo

A continuación se presenta una lista de sitios que resultan interesantes para mostrar como Internet y en particular el Web pueden impactar en las organizaciones:

- [WWWDell] Es un sitio que muestra la posibilidad de ventas directas, autoconfiguración de productos (uno puede armar y dejar encargada una entre varias configuraciones de PC), y muestra una organización virtualmente integrada, organizada en torno a un conjunto de herramientas Web.

- [WWWCisco] Más del 75 % de las operaciones de Cisco se realizan on line, en particular el servicio de soporte al cliente y configuración del sistema. El ahorro en costos excede los U$S 550:000.000.- anuales.

- [WWWEspecta], [WWWCNN] Ejemplos de medios de comunicación en la red.

- [WWWBrit] La Enciclopedia Británica es un ejemplo de producto que se ha visto profundamente afectado por la llegada de las tecnologías de la Web. En el pasado era un gran producto y muy lucrativo. Hoy ya no es líder de mercado y el producto es gratis en Internet.

- [WWWSchwab] Comercio en Internet más de U$S 4.000:000.000 por semana. Comercio on line es el 67 % del total. El sitio web tiene más de 50 millones de visitas por día.

- [WWWBrad] Internet banking, internet gratis, mobile banking, shopping virtual.

Ejercicio

12.11) Buscar ejemplos análogos a los anteriores que involucren sitios uruguayos. ¿ Cuáles cree que son los factores determinantes del éxito de sitios como los anteriores en Uruguay ?

Reducción de costos

Las transacciones en Internet pueden ser drásticamente más baratas. Esto es válido para negocios de muy distinto tipo, y por supuesto depende del tipo de negocio. Por ejemplo, una transacción bancaria tiene un costo de cerca de U$S 1 realizada en ventanilla, U$S 0,5 realizada por teléfono, U$S 0,3 realizada por cajero automático y U$S 0,01 realizada por Internet.

Los eBooks han sido un ejemplo claro de cómo Internet puede cambiar los paradigmas de la publicación tradicional. El libro "Riding the Bullet" de Stephen King vendió 400.000 copias las primeras 24 horas, y el costo era sensiblemente inferior a la publicación tradicional, ya que se habían eliminado una cantidad de intermediarios entre el autor y el lector.

El impacto de Internet en la propuesta de valor se puede medir en varias dimensiones:

- Tiempo tiende a cero.
- Alcance tiende a infinito.
- Información tiende a la perfección.
- Precio tiende al verdadero costo.
- El costo tiende a cero.
- El acceso tiende a 365 x 24.

El adoptar o no las tecnologías de Internet en las operaciones de los negocios se ha tornado una decisión a analizar en muchas organizaciones. En general hay consenso en que las nuevas tecnologías cambiarán el escenario de los negocios, y llegar primero será bueno y llegar tarde será malo.

Ejercicio

12.12) Buscar más ejemplos en los que la adopción de tecnologías basadas en Internet contribuyan a la reducción de costos y al aumento de la competitividad en general.

Los cinco paradigmas básicos del eBusiness

Hay cinco paradigmas que plantean la base del eBusiness:

- Intermediación: Es posible controlar el proceso de ventas. Es posible actuar como proveedor de información, infomediario.
- Desintermediación: Capturar márgenes desplazando intermediarios. Llegar directo al consumidor.
- Personalización masiva: Productos y servicios hechos a medida. Segmentación uno a uno.
- Autoservicio del consumidor: Potenciamiento del consumidor. Reducción de errores de entrada de datos.
- Colaboración en la cadena de valor. Integración entre empresas, información compartida, Ejemplo: pronósticos de demanda.

El eBusiness es mucho más que una página Web o la venta de mercadería en un sitio Web. eBusiness transforma las relaciones y permite que los bienes y la información fluyan en múltiples direcciones.

Ejercicio

12.13) Proporcionar ejemplos para cada uno de los paradigmas mencionados.

La empresa extendida

Para analizar las posibilidades de la incorporación de herramientas de eBusiness en una organización, se propone la siguiente división:

- In Side, la empresa:
    - Fabricación
    - Distribución
    - Recursos Humanos
    - Administración
- Sell-Side:
    - Tienda Virtual
    - Autoservicio del consumidor
    - CRM y eCRM
    - Presentación electrónica de facturas
    - Pagos electrónicos
- Buy Side:
    - Supply Chain Management
    - E-Procurement
    - Intercambio electrónico de datos vía web

Ejercicio

12.14) Buscar, para cada uno de los puntos anteriores, ejemplos de soluciones basadas en Internet que correspondan a las categorías mencionadas.

Los modelos de negocios en Internet

Los modelos de negocios en Internet son muy dinámicos, constantemente surgen nuevas variedades y combinaciones. Además, al momento de escribir este libro no está claro que exista un modelo que pueda considerarse exitoso, y si hay varios modelos que se descartaron por inviables, aunque en otro contexto tal vez merecerían una segunda oportunidad.

Hay algunos términos que han surgido con la industria de Internet, y que se mencionan a continuación:

- Empresas Net-gen (las nacidas por negocios sobre Internet) vs empresas Brick and Mortar (las de la economía convencional).
- Infomediario: intermediario de información.
- Cibernegocios: negocios que tienen la red como protagonista.

Algunas clasificaciones

Según los participantes que convocan y que son convocados, los modelos se clasifican en:

- B2B, empresa a empresa
- B2C, empresa a consumidor
- C2C, consumidor a consumidor
- B2B2C, empresa a empresa a consumidor
- B2I, empresa a inversor.

Según la operación que realizan:

- Venta directa

- Intermediación

Según la proyección del negocio:

- Locales
- Regionales
- Globales

Según el grado de especialización:

- Horizontales
- Verticales

Según el objeto de la negociación:

- Reales
- Virtuales

Una clasificación más completa basada en modelos

Esta clasificación reconoce 10 variantes, que son las siguientes:

- Brokerage model. Modelos de intermediación: Los brokers o intermediarios son creadores de mercado: juntan compradores y vendedores y facilitan las operaciones. Usualmente los intermediarios cobran o bien una comisión o bien un porcentaje sobre las transacciones logradas.
    - Agregador de vendedores, comunidad negociante.
    - Shopping virtual
    - Agregador de compradores
    - Distribuidor
    - Metamediario
    - Remate OnLine
    - Remate Inverso.
    - Clasificados
    - Agentes de búsqueda
- Modelos de publicidad: se provee contenido mezclado usualmente con banners. Sólo funciona cuando el tráfico es alto o altamente especializado.
    - Portal generalizado
    - Portal personalizado
    - Portal Especializado
    - Modelos de servicio gratuito
    - Negocio de descuento.
- Modelo infomediario: Los datos de los consumidores son valiosos para campañas de Marketing, estudios de mercado, etc. Un infomediario puede obtener estos datos a cambio de algo gratis (email, información, etc).
    - Una variante es el sistema recomendador
- Merchant model: modelos de venta directa:
    - Comercio virtual
    - Application Services Provider
    - Ventas por Catálogo Virtual
    - Surf and Turf
    - Vendedor de bits

- Modelos basados puramente en información: servicios que se prestan a través de la red. Por ejemplo firewall y monitoreo de la red.
    - Modelo de venta de servicios virtuales.
- Manufactura directa al cliente.
- Generación de afiliados.
- Modelos de comunidad:
    - Contribuyente voluntario
    - Redes de conocimiento
- Modelos de suscripción.
- Modelo de Utilidad.

Ejercicio

12.15) Dar para cada categoría anterior y sus subcategorías, una descripción y aportar sitios ejemplo.

12.15) Buscar, para cada uno de los modelos anteriores, ejemplos de soluciones basadas en Internet que correspondan a las categorías mencionadas.

Bibliografía

[Amo00] Daniel Amor. La (R)Evolución del e-Business.
Prentice Hall. 2000.

[Blo92] John Bloomer. Power Programming With RPC.
O'Reilly & Associates. 1992.

[Coh00] Peter Cohan. El Negocio está en Internet.
Prentice Hall. 2000.

[Dei90] Harvey Deitel. Sistemas Operativos. Segunda Edición.
Addison Wesley. 1990.

[Gat99] Bill Gates. Los Negocios en la Era Digital.
Editorial Sudamericana.

[Hun94] Craig Hunt. TCP/IP Network Administration.
O'Reilly & Associates. 1994.

[Kal99] Ravi Kalakota, Marcia Robinson. E-Business. Roadmap for Success.
Addison Wesley. 1999.

[Ker84] Brian Kernighan, Rob Pike. The Unix Programming Environment.
Prentice Hall. 1984.

[Mar00] Chuck Martin. Las 7 Cibertendencias del Siglo XXI.
Mc. Graw Hill. 2000.

[Rag93] Stephen Rago. Unix System V Network Programming.
Addison Wesley. 1993.

[Rif00] Jeremy Rifkin. La Era del Acceso.
Paidós. 2000.

[Sey00] Patricia Seybold. Clientes.com.
Ediciones Granica S.A. 2000.

[Sha99] Carl Shapiro, Hal Varian. El Dominio de la Información.
Antoni Bosch. 1999.

[Sie00] Thomas Siebel. Cyber Rules.
Ediciones Gránica S.A. 2000.

[Sun96] Sun Microsystems. Manual Cursos SA-135 y SA-235.
Sun Education. 1996.

[Tan92] Andrew Tanenbaum. Organización de Computadoras: un enfoque estructurado. Tercera edición.
Prentice Hall. 1992

[Tan96] Andrew Tanenbaum. Redes de Computadoras. Tercera Edición.
Prentice Hall. 1996.

[WWWBrad] Sitio de Bradesco. www.bradesco.com.

[WWWBrit] Sitio de la Enciclopedia Británica. www.britannica.com.

[WWWCisco] Sitio de Cisco. www.cisco.com.

[WWWCNN] Sitio de CNN. www.cnn.com.

[WWWDell] Sitio de Dell. www.dell.com.

[WWWEspec] Sitio de Radio El Espectador. www.espectador.com.

[WWWIndic] Sitio Internet Indicators. www.internetindicators.com.

[WWWSchwab] Sitio de Charles Schwab. www.schwab.com