



Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Eléctrica

APUNTES EL55A
Sistemas de Telecomunicaciones

Prof. Néstor Becerra Yoma

1 INTRODUCCION

El curso “Sistemas de Telecomunicaciones”, EL 55 A, es el último en el cuál todos los alumnos de la carrera tienen la oportunidad de ver los conceptos y tecnologías de las telecomunicaciones modernas. Por ese motivo, y tal como dice su nombre, el objetivo principal de este curso, que es bastante ambicioso, es dar una visión amplia y lo más completa posible de las tecnologías que están cambiando las sociedades y economías de hoy en día. La variedad y la complejidad de los temas que se abordan requieren varios libros textos y, dada la imposibilidad de tener varios ejemplares de cada uno de ellos en la biblioteca, es que se pensó en la elaboración de apuntes que sirvieran de guía a los alumnos para los asuntos vistos en clase. Sin embargo, estos apuntes no pretenden reemplazar completamente las referencias que se citan, siendo que la profundización en ciertos temas, cuando sugerido en clase, es de responsabilidad de los alumnos.

La variedad y complejidad de las tecnologías que se estudian imponen algunas restricciones con relación al orden en que estos se presentan. Es de vital importancia, por razones didácticas, que haya una secuencia visiblemente coherente. Si esto no ocurre, se corre el riesgo que el curso se vea como un aglomerado de temas independientes. Así, el orden natural utilizado va desde la capa Física hasta la de Aplicación en los modelos TCP/IP y OSI. Esto se observa en los capítulos II a VIII. En el capítulo IX se ve la teoría de tráfico y de colas utilizada en el diseño de sistemas de telefonía fija y celular. En el capítulo X se ven los conceptos básicos de telefonía celular y, finalmente, en el capítulo XI se hace una introducción a redes ópticas WDM de última generación.

2 REDES DE COMUNICACIONES

Uno de los puntos importantes en el desarrollo de la humanidad es la necesidad de comunicación que nace entre personas, grupos o naciones, con el fin último de compartir y conocer ideas, es decir información. Los medios de comunicación han variado a lo largo del desarrollo de la historia del hombre, pero siempre han tendido a una mayor apertura, es decir, han logrado que más personas puedan comunicarse entre sí, con menor demora y sin importar la ubicación geográfica. De aquí es que surgen como una solución a esta necesidad intrínseca las redes de comunicaciones, las cuales nos permiten acceso a información de variada forma (voz, datos, imágenes, etc) cada vez más eficientes y a menor costo. Este capítulo intenta introducir algunos conceptos básicos para el entendimiento de redes de comunicaciones y de datos.

2.1 Arquitectura de redes

Cuando se habla de arquitectura esta hace relación a todos los componentes de diseño tanto de hardware como software que componen una red de comunicación. En esta sección se especificarán alguno de estos componentes.

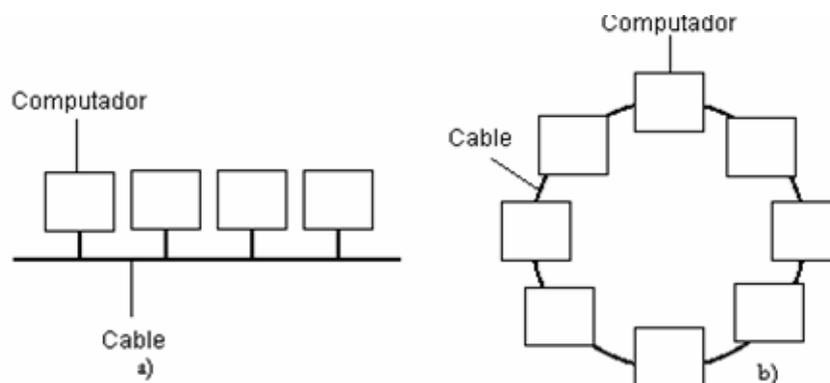


Figura 2.1 Redes de difusión a) Bus b) Anillo

2.1.1 Topología y Hardware

Una de las formas de clasificar las redes son a través de la tecnología de transmisión que utilizan. Estas pueden ser redes de *difusión* o redes *punto a punto*. Las primeras tienen un solo canal de comunicación compartido por todas las máquinas. Los mensajes que envía una máquina son recibidos por todas las demás pero un campo de dirección dentro del frame especifica a quién se dirige. En contraste, las redes punto a punto consisten de muchas conexiones entre pares individuales de máquinas. Para llegar a su destino puede ser que el paquete deba pasar antes por varias máquinas, lo cual hace muy importante el ruteo

que se de a este paquete. En las Figuras 2.1 y 2.2 se muestran posibles topologías para las redes de difusión y punto a punto.

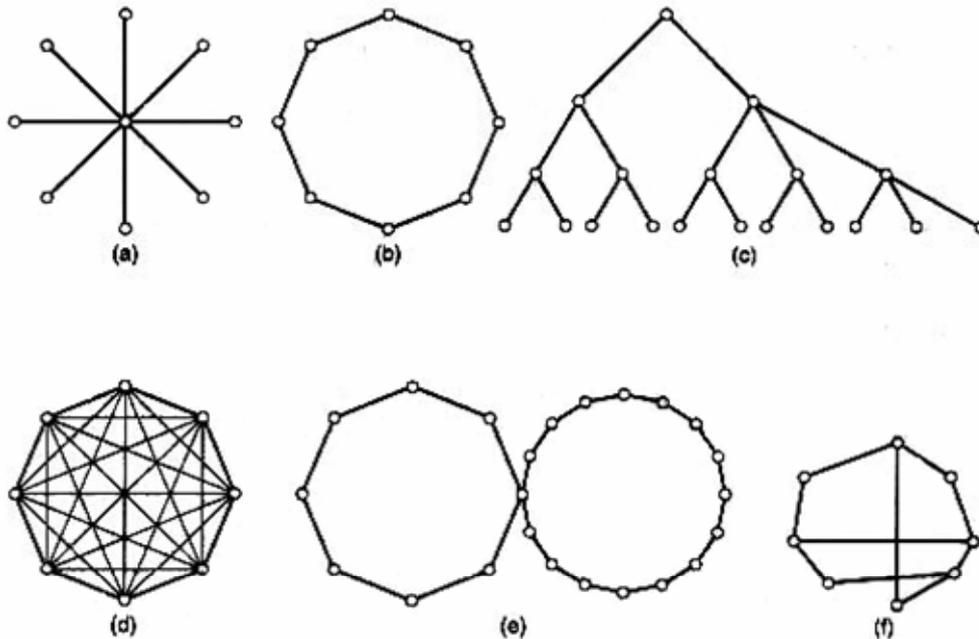


Figura 2.2: Posibles topologías para una subred punto a punto a) Estrella b) Anillo c)Arbol d)Completa e)Intersección de anillos f)Irregular

En función de las dimensiones físicas y del hardware, las redes se pueden clasificar en:

Redes de Área Local (LAN): son redes de propiedad privada dentro de un solo edificio o campus de hasta unos cuantos kilómetros de extensión. Se utilizan para interconectar computadoras dentro de espacios físicos relativamente reducidos (Ej: casa, oficina, campus, etc) con el objeto de compartir recursos. Pueden ser difusión o punto a punto, siendo las más conocidas y utilizadas las primeras, como por ejemplo redes de difusión Ethernet, IEEE 802.3 o IEEE 805.5 (Token Ring).

Redes de Área Metropolitana (MAN): es una versión más grande de una LAN y normalmente se basa en una tecnología similar. Podría abarcar un grupo de oficinas cercanas o una ciudad. Una MAN sólo tiene uno o dos cables y no contiene elementos de conmutación, lo cual simplifica el diseño.

Redes de Área Amplia (WAN): se extiende sobre un área geográfica extensa, a veces un país o continente. Contiene una serie de máquinas de usuario (hosts) los cuales se conectan a una subred de comunicación, la cual se encarga de rutear los paquetes entre hosts. La

subred se compone de líneas de transmisión y elementos de conmutación. Estos últimos conectan dos o más líneas y se encargan de asignar una línea de salida para cada paquete. Como termino genérico se les conoce como *enrutadores (routers)*. En la Figura 2.3 se muestra un ejemplo de subred.

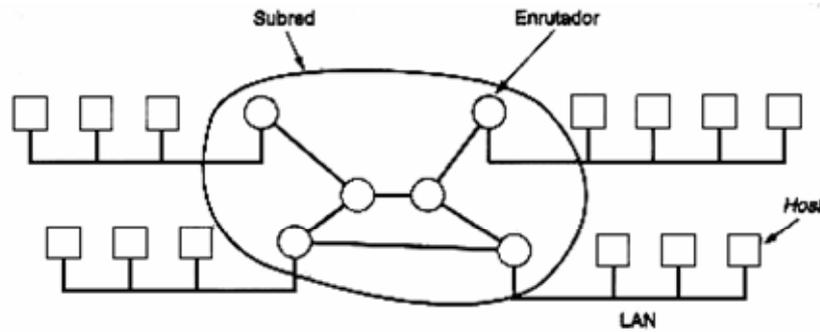


Figura 2.3: Subred y hosts

Pueden existir muchos tipos de redes alrededor del mundo con distintas tecnologías de software y hardware. Un conjunto de redes interconectadas se le llama *internet* (inter.-red). La Internet (con "I" mayúscula) es definida como la red formada por los hosts interconectados con TCP/IP.

2.1.2 Tipos de Conmutación

La comunicación entre máquinas en una red puede describirse desde dos enfoques: *conmutación de circuito* o también conocida como *orientada a conexión*, y *conmutación de paquetes* o *sin conexión*. La primera de ellas, como su nombre lo indica, necesita establecer un canal o circuito físico antes de establecer la comunicación. Una vez establecida, la información se transmite sin pérdida de datos y en orden. El mejor ejemplo de este tipo de conexión es el sistema público de telefonía.

Las redes de conmutación de paquetes funcionan de forma distinta. Estas dividen la información a ser enviada por la red en pequeñas unidades de hasta unos cientos de bytes, llamados paquetes, a los que se agrega identificaciones de destino y origen. El hardware envía estas unidades hacia su destino y cada uno de ellos viaja en forma independiente de su antecesor. El camino a seguir dependerá de las condiciones de la subred. Una vez en el destino el software se encargará de reensamblar los paquetes con la información que estos posean. Una analogía para este tipo de redes se observa en el servicio de correo.

2.1.3 Modelo de Capas

Como es posible deducir, el problema de comunicación entre equipos, los cuales pueden ser de fabricantes diferentes, y por ende, de tecnologías incompatibles puede llegar a ser muy complejo de resolver e implica un sin fin de consideraciones de diseño tanto de hardware como software. Una forma de abordar este desafío es dividir el problema en múltiples etapas, cada una independiente de la otra, y que en conjunto sean capaces de

entregar una solución a variados requerimientos. Por lo anterior, muchas redes se organizan como una serie de capas o niveles cada una construida sobre la inferior. El propósito de cada capa es ofrecer ciertos servicios a las capas superiores. Esto independiza las labores de cada capa pues a la capa superior no le interesa como la capa inferior se encarga de implementar su servicio. Luego, se llega a un nivel de abstracción tal que se puede mirar como si cada capa conversara con la capa adyacente en otra máquina gracias a un *protocolo* definido para cada nivel. Básicamente, un protocolo es un acuerdo entre las partes que se comunican sobre cómo va a proceder la comunicación. En la Figura 2.4 se muestra una red compuesta de cinco capas.

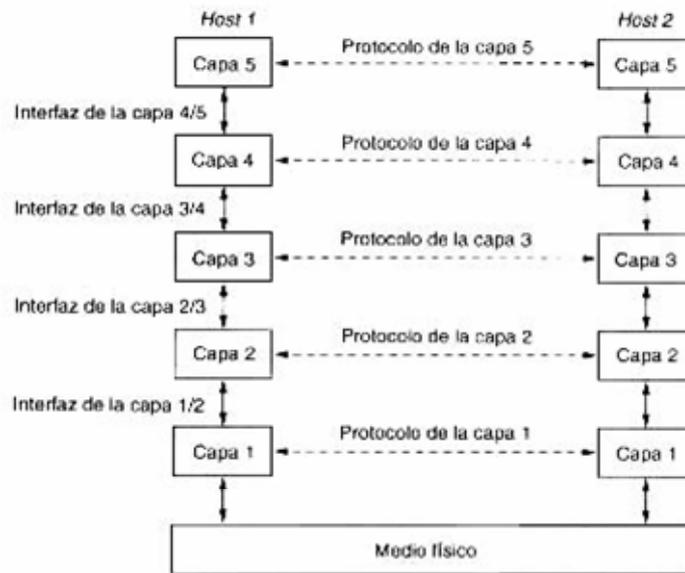


Figura 2.4: Capas, protocolos e interfaces.

De acuerdo a la Figura 2.4, los datos no se transfieren de la capa n de una máquina a la capa n de otra, sino que la capa transfiere los datos e información de control a la capa inferior, hasta llegar a la capa más baja. Luego es el medio físico en donde ocurre la comunicación real. Como se ve en la figura, entre las capas existen las interfaces las cuales definen operaciones y servicios primitivos que ofrece la capa inferior a la superior. La implementación de las funciones que definen las interfaces no se incluye en los protocolos, es decir, es de libre realización. Sólo deben cumplir con asegurar el o los servicios ofrecidos por cada capa.

Un conjunto de capas y protocolos recibe el nombre de *arquitectura de red*. La lista de protocolos empleados por cierto sistema, con un protocolo por capa, se llama *stack de protocolos*. Una analogía, tomada de “*Redes de Computadoras*” de Tanenbaum puede ayudar a explicar la idea de la comunicación multicapas. Imagine a dos filósofos (procesos pares de la capa 3), uno de los cuales es de India y habla indú e inglés, y el otro es de China y habla chino y francés. Ya que no tienen un idioma en común, cada uno contrata un traductor (procesos pares de la capa 2), cada uno de los cuales, a su vez, establece contacto con una secretaria (procesos pares de la capa 1). El filósofo indú desea comunicar a su par

su afecto por los conejos. Para hacerlo, pasa a su traductor un mensaje (en inglés), por conducto de la interfaz 2/3, que dice: "I like rabbits", como se ilustra en la Figura 2.5. Los traductores acuerdan el uso del holandés como idioma neutro y el mensaje se convierte en "Ik hou van konijnen". El idioma elegido es el protocolo de la capa 2 y la elección corresponde a los procesos pares de la capa 2. A continuación, el traductor entrega el mensaje a su secretaria para que lo transmita, por ejemplo, por fax (el protocolo de la capa 1). Cuando el mensaje llega, se traduce al francés ("J'aime les lapins") y se pasa a través de la interfaz 2/3 al filósofo chino. Observe que cada protocolo es independiente por completo de los otros mientras las interfaces no cambien. Los traductores pueden cambiar a voluntad del holandés, por decir, al finlandés, siempre que ambos lo acuerden y que nada cambie su interfaz ya sea con la capa 1 o con la 3. De manera similar, las secretarías pueden cambiar de fax a correo electrónico o a teléfono sin molestar (o incluso sin informar) a las otras capas. Cada proceso puede añadir algo de información dirigida únicamente a su par. Esta información no se pasa a la capa de arriba.

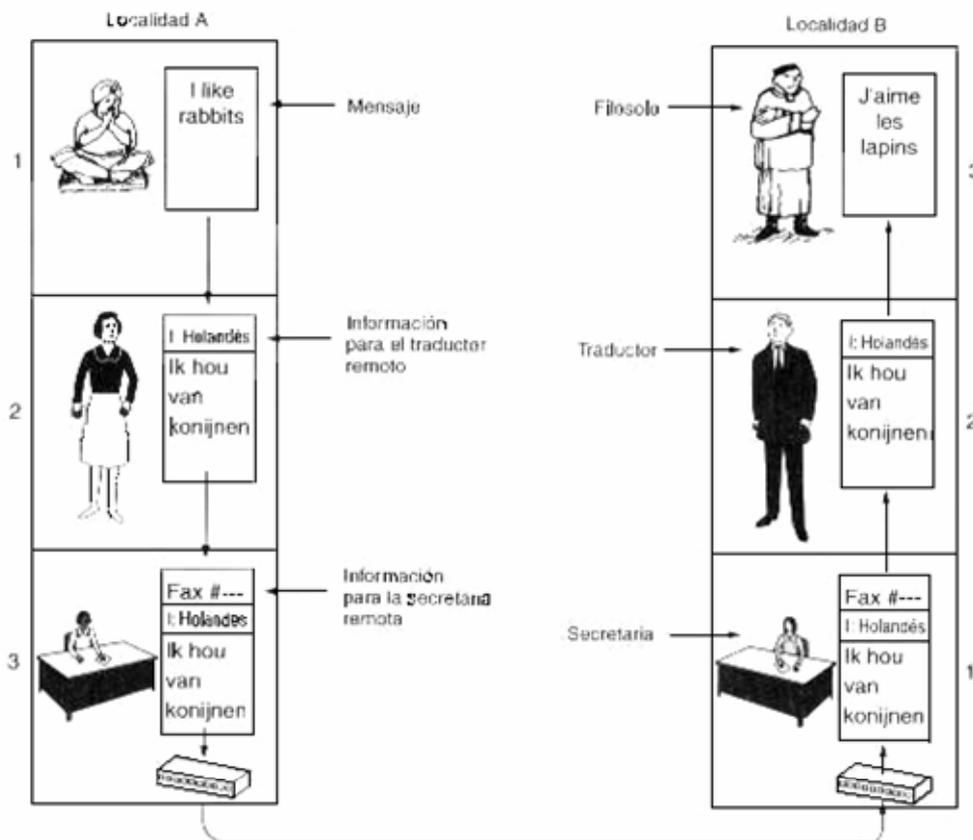


Figura 2.5: Arquitectura filósofo – traductor - secretaria

2.2 Modelo OSI

La Organización para la estandarización (ISO) en 1984 reconoció la necesidad de un modelo de red que ayudara a los productores a crear implementaciones de red inter-

operables y abiertas. De aquí nace el modelo de referencia **OSI** (*Open Systems Interconnection*) de la ISO, el cual sigue la idea de una arquitectura de red compuesta de 7 capas, las cuales debían tener funciones bien definidas y a la vez seguir estándares internacionales.

Aunque otros modelos de arquitectura han sido creados, la mayoría de los proveedores de redes asocian sus productos de red al modelo OSI cuando ellos quieren informar a los usuarios sobre sus productos. Luego, el modelo OSI es la mejor herramienta disponible para enseñanza de tecnologías de redes y de los sistemas de telecomunicaciones en general. Es importante mencionar que el modelo en sí no es una arquitectura de red pues no especifica los servicios y protocolos exactos que se han de usar en cada capa; sólo menciona lo que debe hacer cada una. Sin embargo, la ISO también ha elaborado estándares para todas las capas, aunque no sean parte del modelo de referencia mismo. La Figura 2.6 muestra la estructura de capas del modelo de referencia OSI.

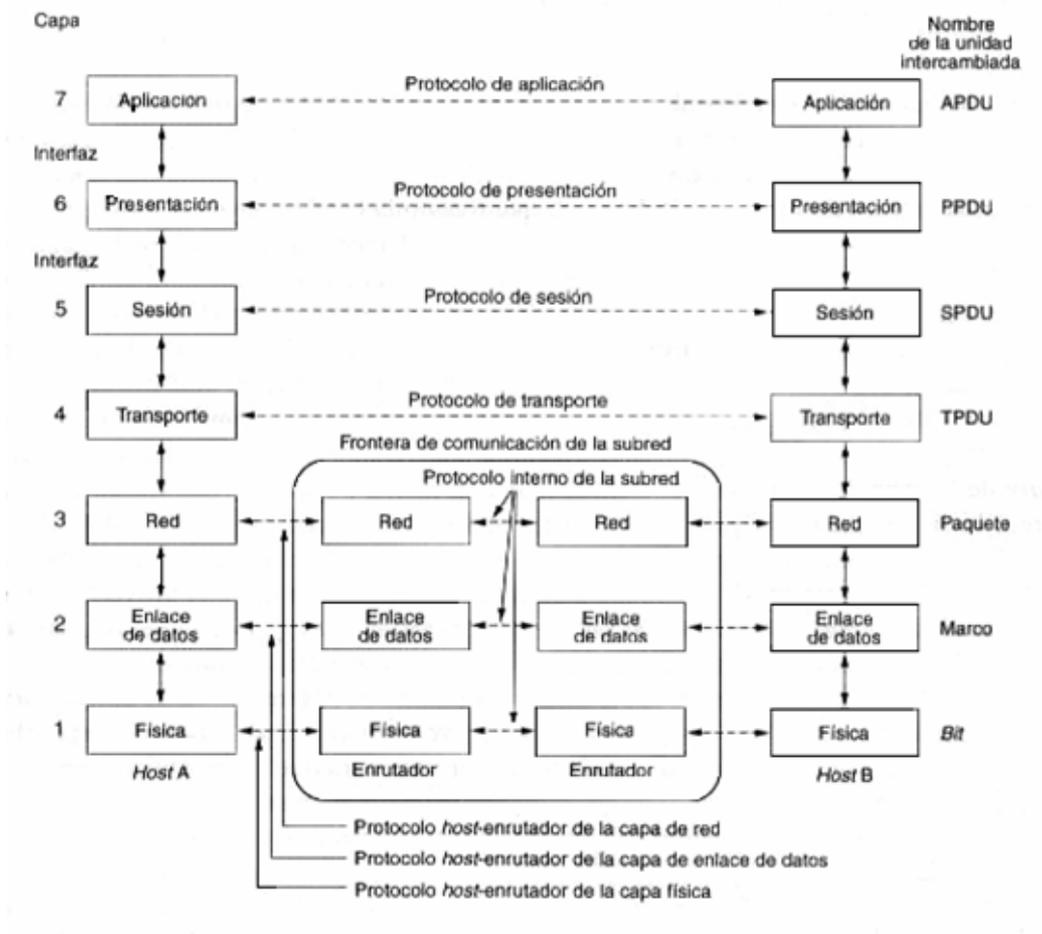


Figura 2.6: Modelo de referencia OSI

2.2.1 Capa Física

Toma en cuenta las interfaces mecánicas y eléctricas que harán posible la comunicación en un medio (el cual puede ser de difusión o punto a punto). Aquí la información es manejada a nivel de bits, debido a esto es necesario definir niveles de voltaje, tiempo de duración de cada bit, sentido de la comunicación, técnica de modulación, etc. Todas estas consideraciones toman en cuenta el medio físico por el cual se llevará a cabo la comunicación. (Ej: fibra óptica, UTP, Cable coaxial, etc)

2.2.2 Capa de Enlace de datos

Se encarga de agrupar los bits de la capa física en *frames o marcos* y a la vez lograr que los hosts sean capaces de identificar un principio y un fin de frame. Además, tratará de enviar estos frames de forma de evitar errores y efectuará control sobre el envío con tal que un emisor no sature el buffer de recepción del equipo destino. La sofisticación en esta capa dependerá del medio de transmisión. .

2.2.3 Capa de Red

Es la encargada del funcionamiento de la subred de comunicaciones, es decir la forma en que los distintos componentes de esta efectúan el enrutamiento de los paquetes y control de congestión en los nodos. La capa de subred también es la encargada de comunicar entre sí redes que no necesariamente utilizan el mismo protocolo en niveles inferiores.

2.2.4 Capa de Transporte

Esta capa se puede pensar como el límite entre las capas concernientes a la aplicación (sesión, presentación y aplicación inclusive) y las capas que guardan relación con el transporte de los datos. Específica asuntos como la confiabilidad en el transporte sobre una interred, es decir, provee mecanismos para establecer, mantener y terminar circuitos virtuales.

Es posible dar un enfoque de extremo a extremo en esta capa, en cuanto a protocolos se refiere, logrando una comunicación directa entre el emisor y usuario final, a diferencia de las capas inferiores en donde la comunicación se realiza entre máquinas o nodos de subred vecinos. Otro punto importante de mencionar es que en esta capa se realiza la multiplexación de mensajes provenientes de distintas aplicaciones por un mismo canal y viceversa.

2.2.5 Capa de Sesión

Esta capa establece, maneja, y termina sesiones entre aplicaciones. Una sesión consiste en un dialogo entre una o más entidades de presentación (servicio a la capa superior). Ejecuta un manejo de control de dialogo y sincronización para transferencia, por ejemplo, de archivos de larga duración.

2.2.6 Capa de Presentación

Esta capa se asegura que la información enviada por la capa de aplicación sea entendible, al mismo nivel de capas, en el otro extremo. Si es necesario se implementan traducciones a formatos comunes de representación. También verifica la estructura semántica y sintaxis de los datos a enviar.

2.2.7 Capa de Aplicación

Es la capa más cercana al usuario. Contiene varios protocolos los cuales pueden emular terminales virtuales o ser utilizados para transferencia de archivos u otras labores. Esta capa establece la disponibilidad de comunicación con el destino y verifica si existen los suficientes recursos para que la transferencia pueda llevarse a cabo.

La figura 2.7 muestra un ejemplo de cómo los datos son transmitidos usando el modelo OSI. El proceso transmisor tiene datos que enviar al proceso receptor, entonces entrega estos datos a la capa de aplicación, la cual agrega su encabezado AH y se lo entrega a la capa de presentación. Esta a su realiza el mismo trabajo anteponiendo su encabezado y entregando el resultado a la siguiente capa.

Este proceso se repite hasta llegar a la capa física en donde es enviado, y al llegar a la maquina receptora esta saca los encabezados uno por uno en las distintas capas hasta que finalmente llega al proceso receptor.

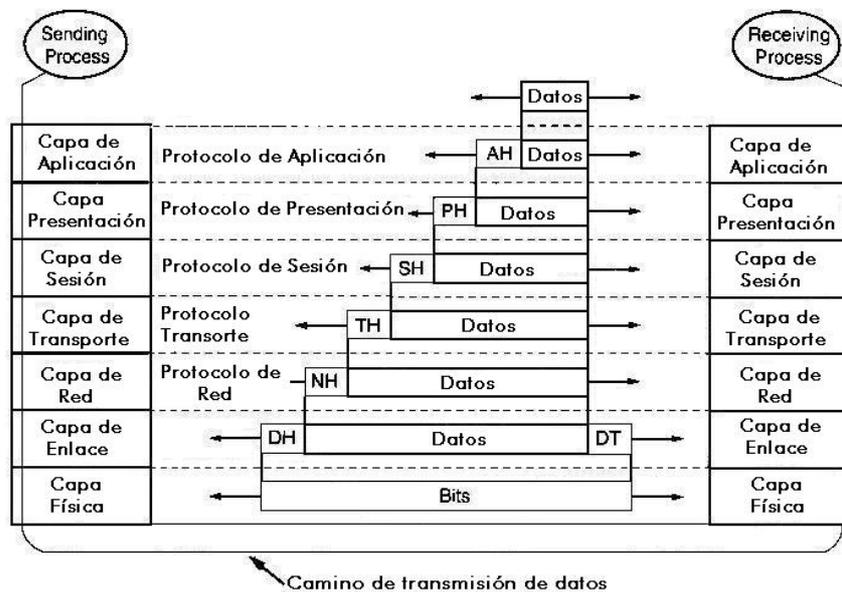


Figura 2.7: Ejemplo de uso del modelo OSI. Algunos de los encabezados pueden ser nulos

2.3 Modelo TCP/IP

Dada las características que podrían poseer las redes de conmutación de paquetes, la Agencia de Investigación de Defensa de los Estados Unidos (ARPA o DARPA) comenzó a interesarse por conformar una red de este tipo a nivel nacional. Esta red debía tener como características principales ser funcional para tecnologías disímiles de red y a la vez ser robusta a la pérdida de uno de sus nodos causada por ataques nucleares.

Con estos objetivos en mente, DARPA financió la investigación para el desarrollo de una serie de protocolos. El resultado de este trabajo fue la creación del grupo de protocolos de TCP/IP, nombre compuesto por las siglas de los dos estándares principales. (*Transmission Control Protocol / Internet Protocol*). En analogía con el modelo OSI, es posible dividir la arquitectura en capas, cada una con interfaces y servicios diferenciados. La Figura 2.8 muestra las distintas capas del modelo TCP/IP en contraste con el modelo OSI.

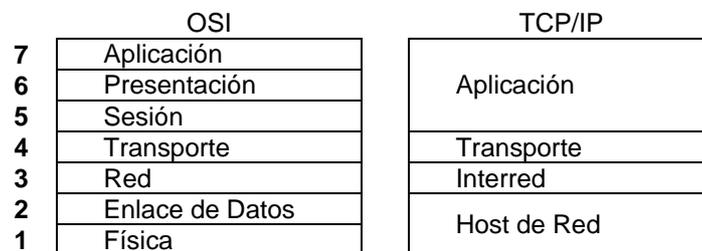


Figura 2.8: TCP/IP

2.3.1 Capa de Interred

Dada las características requeridas para la red se optó por orientar a esta capa hacia un servicio sin conexiones, que permitiera dar mayor flexibilidad a cada uno de los nodos componentes permitiéndoles ingresar paquetes a la red sin importar el destino de estos. Luego esta capa funciona con un filosofía del mejor esfuerzo (*best effort*), no confiable, dejando a las capas superiores la implementación de protocolos más sofisticados para recuperar paquetes perdidos. Luego el trabajo más importante en esta capa es el de llevar estos paquetes por los distintos nodos hacia sus destinos, lo que comúnmente se conoce como ruteo. En la capa de red se trabaja con unidades de largos no fijos, los *paquetes*, y el protocolo oficial para esta capa es el *Internet Protocol* o IP.

2.3.2 Capa de Transporte

Esta capa es similar a la capa del mismo nombre en el modelo OSI, en el sentido que es la que efectúa una comunicación directa entre entidades pares en el emisor y receptor. Para ello se basa en dos protocolos, cada uno de ellos orientado a distinto tipo de servicios. El primero de ellos es TCP (Protocolo de Control de Transmisión) el cual implementa un servicio orientado a conexión y confiable para las capas superiores. A la vez el protocolo TCP se encarga del control de flujo, lo cual evita que un emisor sature la sub-red de comunicación mediante control de congestión. Por otro lado, UDP (Protocolo de

Datagramas de Usuario), a diferencia del anterior, brinda un servicio no orientado a la conexión y no confiable. Es utilizado para aplicaciones que requieren una tasa constante de transmisión en vez de pérdidas iguales a cero. Básicamente UDP aprovecha el protocolo de capa de red para llevar sus mensajes (IP), agregando multiplexación de procesos, la cual permite diferenciar entre aplicaciones que se encuentren en una misma máquina y que estén siendo enviadas a través del mismo canal. Esto se realiza con la utilización de *puertos de protocolo*. Es importante mencionar que TCP también utiliza estos puntos abstractos de acceso llamados puertos, independizando cada uno de los programas o procesos de aplicación que corren en la misma máquina.

2.3.3 Capa de Aplicación

Aquí se encuentran todos los protocolos de alto nivel y que sirven para prestar determinados servicios al usuario. Entre ellos se pueden mencionar: TELNET (terminal virtual), FTP (transferencia de archivos), SMTP (correo electrónico), HTTP, etc. En la Figura 2.9 se observan los distintos protocolos mencionados y otros más que serán descritos en capítulos posteriores. Como se muestra en la figura TCP/IP no especifica protocolo para las capas inferiores a la de interred. Luego este variara y dependerá de la arquitectura de red física a utilizar.

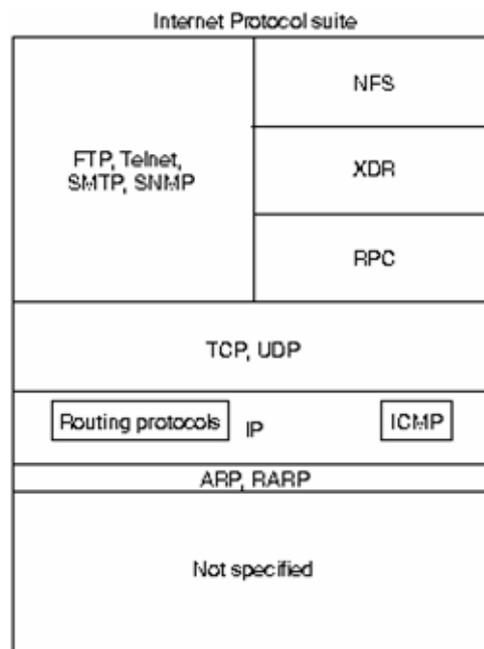


Figura 2.9: Protocolos TCP/IP

2.4 Tipos de Redes

Mientras avanza la técnica e investigación sobre nuevas formas para proveer servicios de comunicación, han surgido una serie de soluciones, las cuales se han ajustado a la problemática del momento llegando algunas de ellas a ser muy populares dadas sus características innovadoras o de menor costo. Algunas de estas tecnologías siguen siendo, en mayor o menor medida, utilizadas por las compañías prestadoras de estos servicios. En esta sección se describirán muy brevemente algunas de estas redes.

2.4.1 Redes X.25

Es un estándar que surge con el objetivo de interconectar terminales con las redes públicas ya existentes en Estados Unidos. Luego X.25 fue desarrollado principalmente por las compañías de teléfonos (portadoras o *carriers*), y fue diseñado para trabajar sin preocuparse del tipo de sistema del usuario. El estándar en cuanto a protocolos fue desarrollado por la ITU-T teniendo como logro resultar en un verdadero estándar mundial. Dado el ambiente de utilización (PDN: *Public Data Networks*) un sistema orientado a conexión, es decir, se necesita establecer una sesión entre el computador emisor y el receptor. Si el receptor acepta la comunicación esta se llevará a cabo en modo full-duplex.

La interacción entre dos puntos (punto a punto) es llevada a cabo entre el *data terminal equipment* (DTE) (en el lugar del usuario) y el *data circuit-terminating equipment* (DCE). A su vez este último se conecta al PSE (*Packet Switching Exchanges*), ambos equipos en las instalaciones del *carrier*, el cual irá conectado un DCE de salida que se conectará a su vez al DTE del receptor final. Esta configuración se muestra en la Figura 2.10.

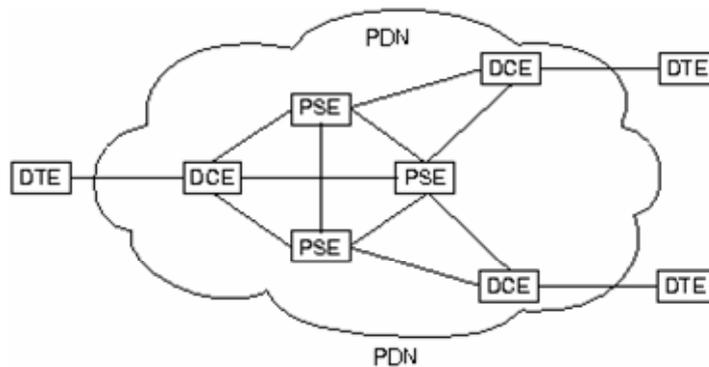


Figura 2.10: X.25

La capa física utiliza el protocolo *X.21 bis*, el cual a grandes rasgos es equivalente al conocido RS-232-C. X.21 bis identifica las características eléctricas y de circuito de intercambio para la interfaz DTE – DCE. Soporta conexión punto a punto, velocidades de hasta 64 Kbps, transmisión síncrona y full – duplex sobre un medio compuesto de cuatro cables. La distancia máxima entre un DTE y un DCE son 15 metros. La capa de enlace se

preocupa de control de transmisión para full – duplex, y manejo de errores en la línea telefónica entre el DTE del usuario y el DCE en la estación del carrier. La capa de red se encarga de los denominados *circuitos virtuales* definiendo identificadores y dependiendo si estos son de carácter permanente o conmutado. También se preocupa del direccionamiento y control de flujo.

2.4.2 Frame Relay

Frame Relay provee una comunicación de conmutación de paquetes, los cuales circulan entre la interfaz de usuario y los equipos de red. Aquí también se distinguen unos de otros como DTE para los equipo en el lado del usuario y DCE para los del lado de red. Este servicio fue pensado para ser más económico y eficiente que sus antecesores, aprovechando los cambio tecnológicos desde redes telefónicas analógicas poco confiables hacia enlaces digitales y menos propenso a errores. Esto permitió olvidarse de algoritmos de corrección de errores solo incluyendo un algoritmo de redundancia cíclico (CRC) que es capaz de detectar bits erróneos sin ejecutan corrección. Todo esto acompañado también con el desarrollo de las computadoras, lo cual permitió delegar a capas superiores trabajos que anteriormente eran ejecutados por las capas más bajas.

Una de las diferencias con X.25 (además de las arriba mencionadas) es que frame relay no incluye explícitamente control de flujo, delegándolo a capas superiores.

Frame Relay hace utilización de los canales virtuales conmutados (SVC), es decir DTE y DCE establecen, terminan y administran dinámicamente un SVC. Este también puede establecer circuitos virtuales permanentes los cuales son configurados por el administrador de la red.

2.4.3 ISDN

Red Digital de Servicios Integrados (RDSI) tiene por principio integrar todos los posibles servicios de voz, datos, texto, gráficos, música y video provistos para el usuario desde un simple terminal sobre la red telefónica pública conmutada.

Los servicios ofrecidos por ISDN son básicamente dos:

- ❑ *Basic Rate Interface (BRI)*: Servicio que ofrece dos canales B y uno D (2B+D). Un canal B trabaja a 64 Kbps y fue pensado para llevar datos. El canal D opera a 16 Kbps y se encarga de llevar información de control y señalización, aunque puede soportar datos bajo ciertas circunstancias.
- ❑ *Primary Rate Interface (PRI)*: servicio que provee se 30 canales B más uno D de 64 Kbps, haciendo un total de 2.048 Mbps (E1). Para EEUU y Japón el servicio consta de 23B + D (1.544 Mbps = T1).

2.4.4 ATM

El Modo de Transmisión Asíncrono (ATM) surge del esfuerzo de la ITU-T por desarrollar un estándar para B-ISDN (Red Digital de Servicios Integrados de Banda Ancha) para permitir transferencia a alta velocidad de voz, video y datos a través de la Red Pública de Datos (PDN).

La idea de ATM es transmitir la información en pequeños trozos de largo fijo llamados *células o celdas*, las cuales se componen de 48 bytes de campo para datos más 5 bytes para encabezado. A nivel de celda ATM, como su nombre lo indica, no necesita de sincronía (no están ligados a un reloj de referencia). Los nodos de una red ATM están compuestos por los llamados switches ATM.

Estas redes son orientadas a la conexión, es decir, es necesario establecer primero una conexión antes de enviar las celdas. Estas conexiones son ejecutadas a una velocidad de 155 Mbps, y trabajan con PVCs o SVCs (Permanent o Switches Virtual Channel) según sea el caso.

El modelo de referencia por capas es más complejo que los anteriormente vistos (TCP y OSI), pues como muestra la Figura 2.11, incluye un tercera dimensión para protocolos de gestión de capas y planos.

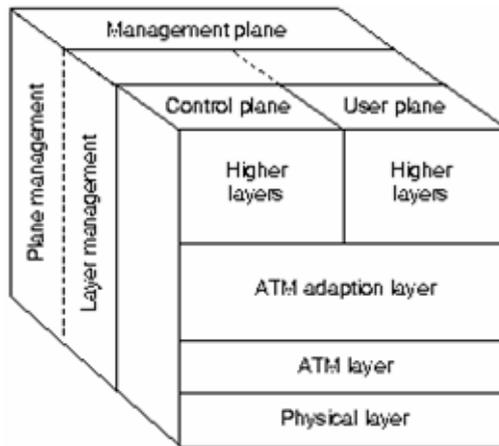


Figura 2.11: ATM

3 TRANSMISIÓN DIGITAL

En este capítulo se abordan los fundamentos básicos en los cuales se sustentan las comunicaciones digitales, a la vez que pretende mostrar las ventajas de la digitalización de la información.

Estas ventajas surgen debido al carácter discreto que adquieren las señales, la cual hace posible utilizar técnicas de procesamiento menos complejas (en los puntos extremos), a la vez que permite a las señales ser más inmunes al ruido (viajando por el mismo canal de comunicación).

Luego las ventajas de la transmisión digital son:

- Inmunidad al ruido
- Procesamiento digital y multicanalización
- Sencillas de medir y evaluar
- Facilidad para medir rendimiento y tasas de error

Y las desventajas:

- El procesamiento digital tiene implícito un retraso
- La conversión A/D y D/A introduce ruido de Cuantización
- Necesita sincronización
- Incompatibilidad con sistemas analógicos.

Como tema relacionado también se tratará un tipo de modulación muy utilizada para transmisión digital, la cual es conocida como PCM (*Modulación de Pulsos Codificados*) y el capítulo concluirá algunos tópicos sobre codificación de voz.

3.1 Ruido

El ruido eléctrico se define como cualquier energía eléctrica no deseada presente en el pasabanda de un circuito de comunicación.

El ruido puede dividirse en dos categorías generales: *Correlacionado* y *No – correlacionado*.

Correlación implica una relación entre la señal y el ruido, mientras que el ruido no correlacionado está presente en la ausencia de cualquier señal. Esto quiere decir que cuando está presente la señal no tiene efecto sobre la magnitud del ruido.

Ruido no correlacionado: Este ruido está presente sin importar si hay señal o no. Puede ser externo o interno.

El ruido externo se debe a ruido atmosférico o estática, ruido del espacio generado por el sol o galaxias. Las frecuencias son de 30 MHz para el ruido atmosférico y entre 8 MHz a 1,5 GHz para ruido del espacio.

El ruido interno es generado dentro de los dispositivos electrónicos y se clasifican en tres tipo: térmico, de disparo (shot – noise) y de tránsito.

Ruido térmico: es el movimiento aleatorio de los electrones por agitación térmica. Se define la densidad de potencia de ruido N_o .

$$N_o = kT [W / Hz]$$

k = Constante de Boltzmann ($1,38 \times 10^{-23}$ [J/°K])

T = Temperatura Absoluta °K

A temperatura ambiente la densidad de potencia de ruido disponible es:

$$N_o = 1,38 \times 10^{-23} \times 290 = 4 \times 10^{-21} [W/Hz]$$

En dBm:

$$N_o = 10 \log \left[\frac{kT}{10^{-3}} \right] = -174 [dBm]$$

La potencia total de ruido en un ancho de banda B es:

$$N = N_o B = kTB [W]$$

En dBm:

$$N_o = 10 \log \left[\frac{kTB}{10^{-3}} \right] [dBm]$$

Como el ruido térmico es igualmente distribuido en el espectro de frecuencias se le denomina ruido blanco en analogía a la luz blanca que contiene todas las frecuencias de la luz visible.

Ruido de disparo: Es producido por la llegada aleatoria de los portadores a los terminales de un dispositivo semiconductor. Esto se debe al efecto del camino libre medio en el movimiento de los portadores, por lo tanto, es proporcional a la corriente y ancho de banda del sistema de comunicaciones. La potencia de ruido de disparo es aditiva con el ruido térmico.

Ruido de tránsito: Generado en transistores BJT debido al retardo introducido por la base para el viaje de los portadores desde el emisor al colector. Es importante en alta frecuencia.

Ruido correlacionado: Es energía eléctrica no deseada que está presente como un resultado directo de una señal. Ej: distorsión armónica y de intermodulación. Ambas distorsiones son consecuencia de la amplificación no lineal de la señal. La diferencia entre ellas reside en que la primera aparece con sólo 1 señal y la segunda con dos señales.

Distorsión armónica: Es la generación de armónicas no deseadas por la amplificación no lineal de una señal sinusoidal. Ej: amplificador de señales grandes. La distorsión armónica total THD se define como:

$$THD = \frac{\sqrt{V_2^2 + V_3^2 + \dots + V_n^2}}{V_1} \times 100\%$$

Con V en valores efectivos.

Ruido de intermodulación: Son las frecuencias no deseadas del producto cruzado (sumas y diferencias) de dos señales cuando son amplificadas en un dispositivo no lineal. Ejemplo: estaciones radiobase de telefonía celular.

3.2 Teorema Muestreo y Máxima Capacidad del Canal

Para ponernos en el contexto en el cual son aplicables los teoremas a enunciar es necesario hacer una breve descripción de las componentes de un sistema análogo digital. La Figura 3.1 muestra los bloques fundamentales de un sistema de conversión A/D (Análogo a Digital).



Figura 3.1: Diagrama de bloques conversor A/D

En donde:

$x_a(t)$: señal analógica (t real)

$x(n)$: señal discreta (n discreto tq. $n=0,1,\dots$)

$x_q(n)$: señal digital

La primera etapa es el muestreo de la señal analógica. Este muestrea una tasa tal que el proceso pueda ser reversible. Aquí lo que se discretiza es la variable independiente (variable temporal). Esta tasa esta dada por el *Teorema del Muestreo*, el cual se enuncia a continuación.

Teorema del muestreo (Nyquist 1928, Shannon 1949): Si la frecuencia más alta contenida en una señal analógica $x_a(t)$ es $F_{\text{máx}}=B$ y la señal se muestrea a velocidad $F_2 > 2F_{\text{máx}} = 2B$ entonces $x_a(t)$ se puede recuperar totalmente a partir de sus muestras.

Si la tasa de muestreo es menor a $2B$ se produce “*aliasing*” o distorsión de la señal recuperada a esta tasa.

Luego, según el teorema anterior el proceso de muestrear un señal en el tiempo y obtener una representación discreta es reversible siempre y cuando la tasa de muestreo sea el doble de la máxima frecuencia contenida en la señal analógica.

Después de esta etapa se obtiene una señal discreta en tiempo pero que puede tener un rango infinitos de valores dentro de un intervalo (en el cual se define la señal) real. Para obtener una señal digital es necesario, entonces, cuantizar, con tal que el resultado este dentro de un conjunto finito y conocido de valores, y pueda ser posteriormente procesada.

3.3 Error de Cuantización y ruido

El proceso que transforma una señal de amplitud continua discreta en el tiempo en una señal digital al asociar a cada muestra un número de dígitos finito es conocido como *cuantización*. Como se puede deducir, este proceso introduce un error al representar los valores continuos de la señal a niveles discretos, el cual es conocido como *error o ruido de cuantización*.

Con el fin de obtener una expresión que nos sirva para representar el error de cuantización, se seguirá un alcance estadístico como el sugerido en el texto “*Digital Signal Processing: principles, algorithms and applications*” (de los autores John Proakis y Dimitris Manolakis, editorial Prentice Hall).

La cuantización es un proceso no lineal e irreversible, la cual mapea una amplitud dada $X(n) \equiv X(nT)$ en tiempo $t = nT$ dentro de una amplitud X_k , tomado desde un conjunto finito de valores. El procedimiento es ilustrado en la Figura 3.2, donde el rango de amplitud de la señal es dividido en L intervalos con $L+1$ *niveles de decisión* X_1, \dots, X_{L+1} .

$$I_k = \{x_k < x(n) < x_{k+1}\} \quad k = 1, 2, \dots, L \quad (3.0)$$

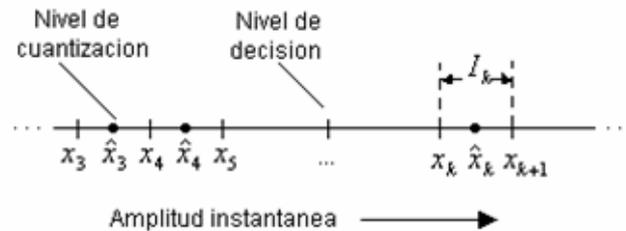


Figura 3.2: Proceso de cuantización

Las posibles salidas del cuantizador son denotadas por $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_L$.

El cuantizador frecuentemente utilizado es el *cuantizador lineal o uniforme*, definido por:

$$\begin{aligned} \hat{x}_{k+1} - \hat{x}_k &= \Delta & k = 1, 2, \dots, L-1 \\ x_{k+1} - x_k &= \Delta & \text{Para } x_{k+1}, x_k \text{ finito} \end{aligned}$$

donde Δ es el tamaño del paso del cuantizador.

Se puede ver fácilmente que error del cuantizador está siempre en el rango $-\Delta/2$ a $\Delta/2$:

Luego el error de cuantización $\frac{\Delta}{2} \leq e_q(n) \leq \frac{\Delta}{2}$ puede exceder la mitad del paso de cuantización. Si el rango dinámico de la señal, definido como $x_{\max} - x_{\min}$, es más grande que el rango del cuantizador, las muestras que exceden el rango del cuantizador son recortadas, resultando en un error más grande que $\Delta/2$.

El proceso de codificación en un conversor A/D asigna un número binario único a cada nivel de cuantización. Si tenemos L niveles, se necesitarán a lo menos L números binarios distintos. Con una palabra de largo b+1 bits se pueden representar 2^{b+1} números binarios distintos. Entonces podríamos tener $2^{b+1} \geq L$ o, equivalentemente $b+1 \geq \log_2 L$. Entonces el tamaño del paso o la resolución del conversor A/D esta dada por:

$$\Delta = \frac{R}{2^{b+1}} \quad (3.2)$$

Donde R es el rango del cuantizador.

3.3.1 Análisis del error de cuantización

La dependencia del error de cuantización en las características de la señal de entrada y la naturaleza no lineal del cuantizador hacen de análisis determinístico intratable, excepto en casos muy simples.

En el presente análisis se asume que el error de cuantización es de naturaleza aleatorio, y se modela como un ruido que es sumado a la señal original.

Se requieren los siguientes supuestos sobre las propiedades estadísticas de $e_q(n)$:

1. El error $e_q(n)$ está uniformemente distribuido sobre el rango $-\Delta/2 < e_q(n) < \Delta/2$.
2. La secuencia $\{ e_q(n) \}$ es una secuencia de ruido blanco estacionario. En otras palabras, el error $e_q(n)$ y el error $e_q(m)$ para $m \neq n$ están no correlacionados.
3. La secuencia de error $\{ e_q(n) \}$ está no correlacionada con la secuencia de señal $X(n)$.
4. La secuencia de la señal $X(n)$ es de media cero y estacionaria.

Bajo estas suposiciones (que no siempre se mantienen), el efecto del ruido aditivo sobre la señal puede ser cuantizado al evaluar la relación señal a ruido de cuantización (SQNR), la cual puede ser expresada sobre una escala logarítmica (en dB) como:

$$SQNR = 10 \log_{10} \frac{P_x}{P_n} \quad (3.3)$$

donde $P_x = \sigma_x^2 = E[X^2(n)]$ es la potencia de la señal y $P_n = \sigma_e^2 = E[e_q^2(n)]$ es la potencia del ruido de cuantización.

Si el ruido de cuantización está uniformemente distribuido en el rango $(-\Delta/2, \Delta/2)$ como muestra la Figura 3.3, el valor medio del error es cero y la varianza (la potencia del ruido de cuantización) es:

$$P_n = \sigma_e^2 = \int_{-\Delta/2}^{\Delta/2} e^2 p(e) de = \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} e^2 de = \frac{\Delta^2}{12} \quad (3.4)$$

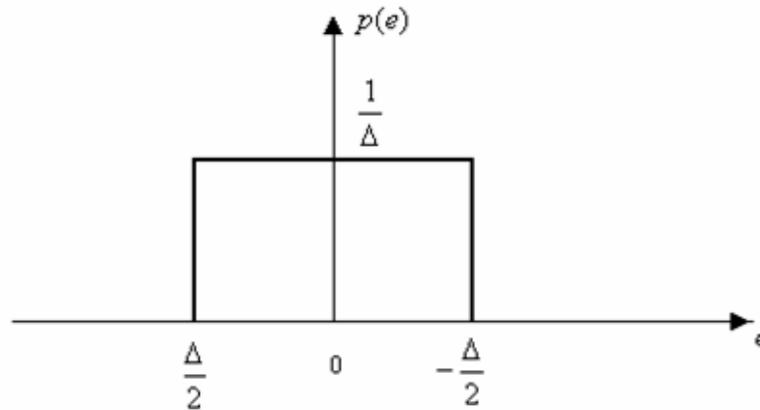


Figura 3.3: Función densidad de probabilidad del error de cuantización

Al combinar la ecuación (3.2) con (3.4) y sustituirla en (3.3), se tiene:

$$SQNR = 6.02b + 16.81 - 20 \log \frac{R}{\sigma_x} \quad (3.5)$$

La cual depende del rango R del conversor A/D y las estadísticas de la señal de entrada. Por ejemplo si asumimos $X(n)$ está distribuida en forma normal y el rango del cuantizador se extiende desde $-3\sigma_x$ a $3\sigma_x$ (i.e. $R = 6\sigma_x$) se tiene:

$$SQNR = 6.02b + 1.25 [dB] \quad (3.6)$$

La formula general para el SQNR es usada frecuentemente para especificar la precisión necesaria en un conversor A/D. Esta simplemente significa que cada bit adicional en el cuantizador aumenta la relación señal ruido en 6 dB.

3.4 PCM (Pulse Code Modulation)

La idea principal de la transmisión digital es lograr convertir la información en pulsos los cuales pueden ser enviados a través del canal de comunicación. Esta información ya puede estar el formato deseado, o también puede estar en forma analógica. Si es el caso este último, se a descrito en las secciones anteriores la base de la transformación de la información a formato digital.

Al momento de cuantizar, los niveles continuos de amplitud de la señal son encasillados en niveles discretos los cuales representan una aproximación al valor original.

Como último paso necesitamos transformar estos valores discretos en pulsos los cuales serán transmitidos a través del canal de comunicación.

Para ello se utilizan métodos de *modulación de pulsos*. Estos permiten transformar la información deseada en pulsos, siendo los más conocidos:

1. **PWM (*Pulse Width Modulation*)**: El ancho del pulso es proporcional a la amplitud de la señal analógica.
2. **PPM (*Pulse Position Modulation*)**: La posición de un pulso (de ancho constante) dentro de una ranura de tiempo determinada varía proporcionalmente a la amplitud de la señal.
3. **PAM (*Pulse Amplitud Modulation*)**: La amplitud de un pulso fijo (de ancho constante) varía de acuerdo a la amplitud de la señal.
4. **PCM (*Pulse Code Modulation*)**: A cada nivel de cuantización se le es asignado un número binario (código) de largo fijo. El número binario varía su valor de acuerdo a la amplitud de la señal.

La Figura 3.4 muestra un ejemplo para cada tipo de modulación.

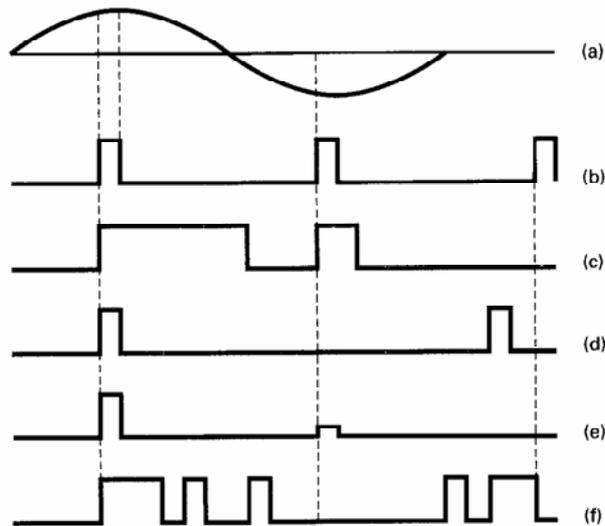


Figura 3.4: a) Señal analógica; b) Pulsos de muestreo; c)PWM; d)PPM; e)PAM; f)PCM

Como se puede notar, la formula obtenida para la relación señal ruido de cuantización (3.6), lleva implícita el hecho de la utilización de un código (de b bits) para cada nivel de cuantización. Luego la relación que se cumple entre los L niveles de cuantización y los b bits del código es:

$$2^b \geq L \tag{3.7}$$

$$\Rightarrow b = \lceil \log_2 L \rceil$$

Esto refleja que el método mayormente utilizado es PCM, y es el que se estudiará en forma breve.

En PCM, tanto la amplitud como la duración de cada pulso son fijos, los cuales son de carácter binario (1 's y 0's). La Figura 3.5 muestra un sistema simplificado de un solo canal PCM.

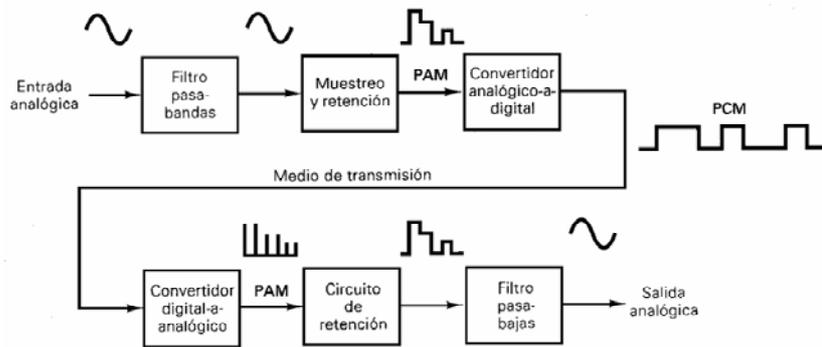


Figura 3.5: Sistema con canal PCM

Debido a la utilización de códigos para representar cada uno de los niveles es necesario el uso de codificadores y decodificadores en la transmisión y recepción respectivamente. Un sistema que realiza la codificación/decodificación es conocido como *codec*.

La eficiencia de un codificador PCM se representa como la razón entre el mínimo número de bits requeridos para lograr un número de niveles específicos y el número real de bits usados.

$$\text{eficiencia del codificador} = \frac{\text{mínimo número de bits} \times 100}{\text{número real de bits (incluyendo bit de signo)}}$$

También es posible distinguir entre códigos PCM *lineales* y *no lineales*. Los primeros son los que asignan en forma uniforme a los niveles de cuantización los códigos correspondientes, y son los que hemos visto hasta el momento. Los no lineales o no uniformes, son aquellos que asignan un mayor número de códigos a ciertos niveles de amplitud. Por ejemplo en codificación de voz es más probable que las señales sean de menor amplitud, por lo tanto se asignan mayor cantidad de códigos, provocando un mayor error de cuantización para señales de mayor amplitud (disminuyendo la relación SQNR).

Compansión

Se refiere al proceso de comprimir una señal (en el lado del transmisor) y luego expandirla (en el lado del receptor). Con PCM este proceso se logra por medio de procesos analógicos o digitales.

La característica de compansión vendrá dada por el tipo de señales a tratar. Por ejemplo, para señales de voz es necesario se necesita un rendimiento SQR relativamente constante, lo que significa que la distorsión debe ser proporcional a la amplitud de la señal

para cualquier nivel de señal de entrada[±]. Esto requiere una razón de compresión logarítmica. Existen dos métodos de compresión analógicos que se aproximan a una función logarítmica, y son conocidos como compansión de *ley-μ* y de *ley-A*.

- **Compansión de ley-μ:** Es usada en Estados Unidos y Japón.

$$V_{salida} = \frac{V_{máximo} \times \ln(1 + \mu V_{entrada} / V_{máximo})}{\ln(1 + \mu)} \quad (3.8)$$

en donde: $V_{máximo}$ = máxima amplitud de entrada analógica descomprimida.
 $V_{entrada}$ = amplitud de la señal de entrada en un instante particular del tiempo
 μ = parámetro usado para definir la cantidad de compresión.
 V_{salida} = amplitud de salida comprimida.

- **Compansión de ley-A:** Es utilizada en Europa.

$$V_{salida} = V_{máximo} \frac{AV_{entrada} / V_{máximo}}{1 + \ln A} \quad 0 \leq \frac{V_{entrada}}{V_{máximo}} \leq \frac{1}{A} \quad (3.9)$$

$$V_{salida} = V_{máximo} \frac{1 + \ln(AV_{entrada} / V_{máximo})}{1 + \ln A} \quad \frac{1}{A} \leq \frac{V_{entrada}}{V_{máximo}} \leq 1 \quad (3.10)$$

3.5 Máxima Capacidad de Canal

Una vez revisado los conceptos básicos para la transformación de la señal de información análoga en una señal digital que puede ser transmitida, es necesario describir cual es la capacidad del canal, en cuanto a tasa de bits se refiere.

Nyquist, en 1924, postuló que una señal que ha sido pasada por un filtro de ancho de banda B, y luego de ser cuantizada en L niveles discretos, la máxima tasa de datos a ser transmitidos por un canal sin ruido será:

$$máxima \text{ tasa de datos} = 2B \log_2 L \text{ bits / seg} \quad (3.11)$$

Pero esta relación considera canales sin ruido.

La forma de medir el ruido térmico presente en el canal es con la relación señal ruido S/N. Shannon (1948) postuló que para un canal con una relación señal ruido S/N, ancho de banda B, se tiene una capacidad de canal máxima dada por:

$$C = B \log_2 \left[1 + \frac{S}{N} \right] \quad (3.12)$$

[±] Fuente: “Sistemas de comunicaciones electrónicas” Wayne Tomasi.

La ecuación (3.12) sólo representa un límite teórico para la capacidad de canal, ya que los sistemas reales raramente lo alcanzan.

3.6 Codificación de Voz

Dado la importancia que ha tenido, sobre todo en comunicaciones móviles, se ha considerado introducir en este capítulo una breve descripción de los fundamentos básicos para la codificación de voz. Muchos de los temas a mencionar ya han sido tratados con anterioridad, así que sólo serán mencionados brevemente.

Gran parte de la información (sino su totalidad) aquí expuesta fue extraída del libro “*Wireless Communications*” (Theodore S. Rappaport), capítulo 7 “*Speech Coding*”.

3.6.1 Introducción

El fin de todo sistema de codificación de voz es transmitir voz con la más alta calidad posible usando la menor capacidad de canal. Estos dos objetivos son posibles manteniendo cierto compromiso entre nivel de complejidad de implementación y retraso en comunicación. Entre más complejo es un algoritmo, más retraso de procesamiento y costo de implementación requiere.

La jerarquía en codificadores de voz se muestra en la Figura 3.6.

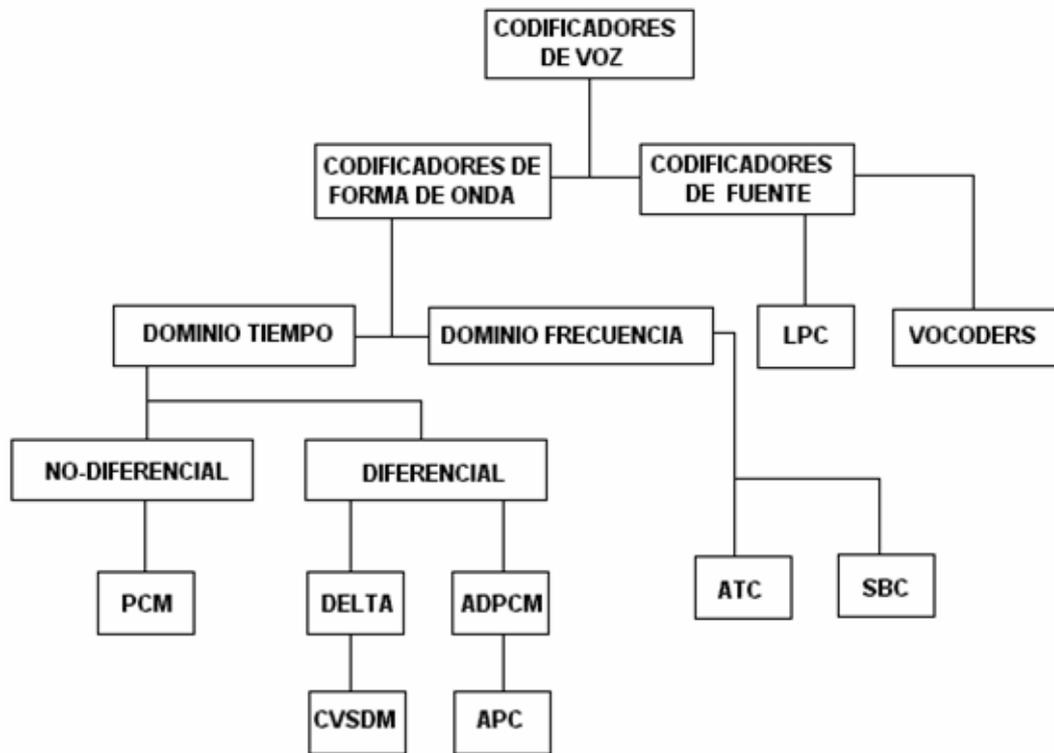


Figura 3.6: Jerarquía Codificadores.

Basados en la forma por la cual logran la compresión, los codificadores de voz son ampliamente clasificados en dos categorías: *Codificadores de forma de onda* y *Vocoders*. Los primeros esencialmente se esfuerzan por reproducir la forma de onda en el tiempo de la señal de voz tan cercanamente como sea posible. En principio fueron diseñados para ser independientes de la fuente y por lo tanto codifican igualmente bien una variedad de señales. Ellos tienen la ventaja de ser robustos para un amplio rango de características de voz y para ambiente ruidoso, con un mínimo de complejidad pero con moderada economía en tasa de bits. Algunos ejemplos de estos codificadores son: PCM, DPCM (PCM diferencial), ADPCM (PCM adaptivo diferencial), modulación delta (DM), modulación delta de gradiente continuamente variable (CVSDM), y código predictivo adaptivo (APC). Por otro lado, los Vocoders alcanzan tasa de bits menores y generalmente son más complejos. Ellos se basan en un conocimiento a priori de la señal a codificar, y por esta razón ellos son específicos para tipos de señales.

3.6.2 Características de la señal de voz

La forma de onda de la señal de voz tiene una serie de características y propiedades que pueden resultar útiles para el diseño de codificadores. Algunas de ellas son: distribución de probabilidad no uniforme de la amplitud de la voz, autocorrelación distinta de cero entre muestras, espectro de voz no plano. La propiedad más básica de la señal de voz que es explotada por todos los codificadores es que es de banda limitada, la cual hace posible el muestreo.

La siguiente propiedad más explotada es la *función de densidad de probabilidad no uniforme* de amplitudes que posee la voz. Esta es caracterizada por una probabilidad muy alta de amplitudes cerca del cero y una función decrecientemente monótona de amplitudes en los extremos. La distribución exacta, sin embargo, depende del ancho de banda de la entrada y condiciones de grabación. La función exponencial dada en la ecuación (3.13) provee una buena aproximación para señales de calidad telefónica.

$$p(x) = \frac{1}{\sqrt{2}\sigma_x} \exp(-\sqrt{2}|x|/\sigma_x) \quad (3.13)$$

La propiedad de las señales de voz la cual indica que existe mucha correlación entre muestras adyacentes de segmentos de voz permite realizar predicción utilizando las muestras previas con un pequeño error aleatorio.

La característica no plana de la densidad espectral hace posible obtener compresión significativa de código de voz en el dominio de la frecuencia.

3.6.3 Técnicas de cuantización

Existen básicamente dos técnicas: Cuantización uniforme (ya analizada) y no uniforme.

La cuantización no uniforme intenta distribuir los niveles de cuantización de forma más eficiente, tratándolos de acuerdo a la fdp de la señal de entrada. Para una señal de entrada se tiene una distorsión de la forma:

$$D = E[(x - f_Q(x))^2] = \int_{-\infty}^{\infty} [x - f_Q(x)]^2 p(x) dx \quad (3.14)$$

Donde:

- x(t): señal de voz
- f_Q: señal de voz cuantizada
- p(x): fdp

Luego se ve que la distorsión se reduce si se reduce el ruido de cuantización en donde fdp es grande.

Una implementación simple de un cuantizador no uniforme es el cuantizador logarítmico. Este usa pasos de cuantización fina para amplitudes bajas que ocurren frecuentemente en voz y usa pasos más gruesos para las menores frecuencias.

Se obtiene cuantización no uniforme al pasar primeramente la señal de voz análoga por un compansor (ley - A o ley - μ) y luego pasar la voz ya comprimida por un cuantizador estándar uniforme.

3.6.4 Vocoders

Los vocoders son una clase de codificadores de voz que analizan la señal de voz, transmitiendo parámetros derivados del análisis. En el receptor se sintetiza la voz usando estos parámetros. Todo Vocoder intenta modelar el proceso de generación de voz como un sistema dinámico tratando de cuantificar ciertas restricciones físicas del sistema. Son en

general más complejos que los codificadores de forma de onda y alcanzan más economía en la tasa de transmisión, pero son menos robustos y su ejecución tiende a depender del locutor. El más conocido es el LPC (*Linear Predictive Coder*).

La Figura 3.7 muestra el modelo de generación de voz tradicional que es la base de todo vocoder.

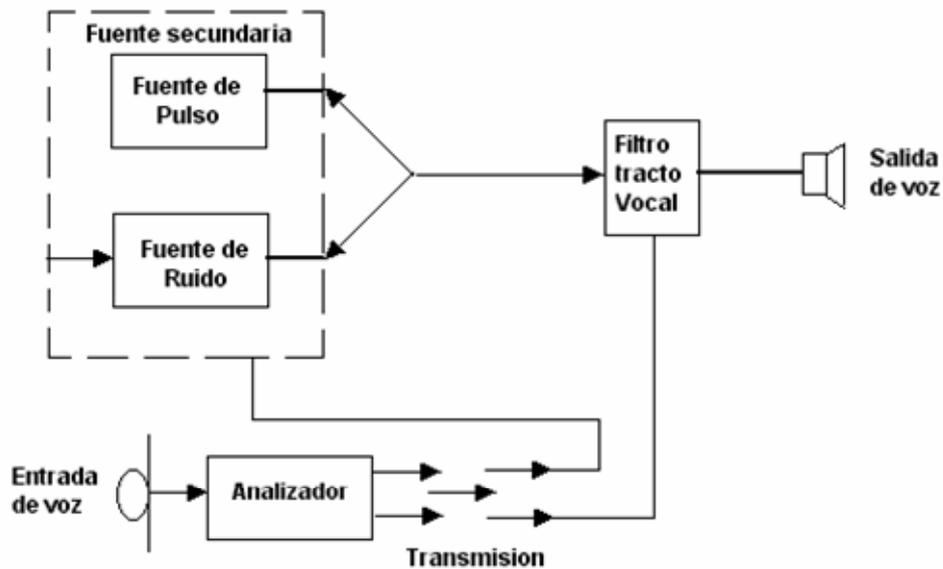


Figura 3.7: Modelo de generación de voz.