

Sistemas Operativos - Apuntes para final

Gianfranco Zamboni

27 de abril de 2023

Índice

1. Introducción a la teoría de la información	2
1.1. La semilla del internet	2
1.2. Modelo TCP/IP	4
2. Nivel físico	5
2.1. Fundamentos de las señales	6
2.2. Teoría de la información	6
2.3. Los medios de transmisión reales	8
3. Nivel Enlace	11
4. Medios Compartidos	11
5. Nivel Red	11
6. Ruteo Externo e Interno	11
7. Nivel de transporte	11
8. Aplicaciones	12
9. Seguridad	12

1. Introducción a la teoría de la información

1.1. La semilla del internet

El telégrafo fue el antecesor del teléfono, un primer acercamiento a la comunicación de mensajes vía una codificación. Desde fines de siglo XIX hasta segunda mitad del siglo XX, aparecen las centrales de **conmutación de circuitos** (centrales telefónicas). A estas centrales llegaban señales (cables) correspondientes a todas las casas que participasen en el sistema de teléfonos. Las operadoras conectaban dos circuitos en sus tableros para cerrar el circuito y permitir la comunicación entre las dos partes involucradas. Sin embargo, este tipo de comunicación tenía una gran falla: Si una central de conmutación de circuitos dejaba de estar disponible por algún motivo de fuerza, todas las personas pertenecientes a esa zona se verían incomunicadas.

A fines de los 50 se empieza a desarrollar la **conmutación de paquetes** buscando resolver este tema, es decir se busca una **red más tolerante a fallas, más flexible** a la hora de conectar dos puntos distantes y que **escale más fácilmente** ante un incremento en el acceso a la comunicación.

La nueva red propuesta es una red descentralizada con múltiples caminos entre dos puntos que divide los paquetes en fragmentos que pueden llegar a destino a través de distintos caminos.

El modelo OSI

En 1983 aparece una publicación de ISO para establecer un estándar que especifique la estructura de una arquitectura de red, que uniformice la forma de construir las redes de comunicación: el modelo OSI-ISO (Open Systems Interconnection).

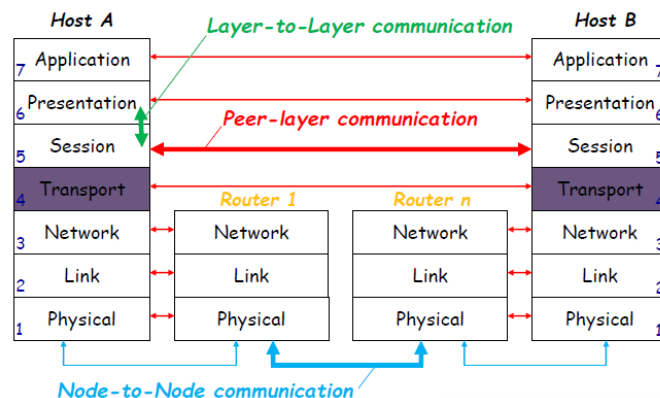


Figura 1: Modelo OSI

Este modelo está dividido en 7 capas, cada una de las cuales tiene una función definida que permitirán la comunicación coherente entre dos sistemas remotos.

1. La capa **física (Physical)** se encarga de enviar raw bits a través de los medios físicos disponibles en la red.

2. La capa de **enlace (Link)** se encarga de detectar errores en la transmisión y corregirlos, si es posible.
3. La capa de **red (Network)** se encarga de resolver problemas de congestión dentro de la red, que paquetes se aceptan y la ruta que deben tomar los paquetes que se envían por la misma.
4. La capa de **transporte (Transport)** se encarga de tomar la información provista por la capa de arriba, pasarla a la capa de red separada en pedazos más chicos (**chunks**) y se asegura que todas las partes lleguen a destino correctamente.

Esta es la primer capa **end-to-end**, es decir que entabla una conversación entre la máquina emisora (**Source**) y la destinataria (**Destination**). Las capas anteriores, usan protocolos de comunicación nodo a nodo, es decir, entre una máquina y su vecino inmediato y no entre el source y el destination que podrían estar separados entre sí por varios nodos.

5. La capa de **sesión (Session)** permite establecer sesiones entre dos máquinas distintas. Estas sesiones permiten sincronizar el pasaje de información entre ambas máquinas, deciden de quien es el turno para enviar información y evitar que ambas máquinas realicen operaciones críticas de manera simultanea.
6. La capa de **presentación (Presentation)** procesa la información recibida, la estructura y la codifica de la manera necesaria para que pueda ser usada por la máquina.
7. La capa de **aplicaciones (Application)** contiene los protocolos necesarios para que los usuarios puedan ver y leer la información.

Además de estas funcionalidades, cada capa ofrece una interfaz que le permite comunicarse con las capas vecinas para hacer el pasaje de los datos entre ellas y asumen que el host en el otro extremo de la comunicación tendrá una arquitectura similar y podrá interpretar los mensajes de cada una de ellas.

1.2. Modelo TCP/IP

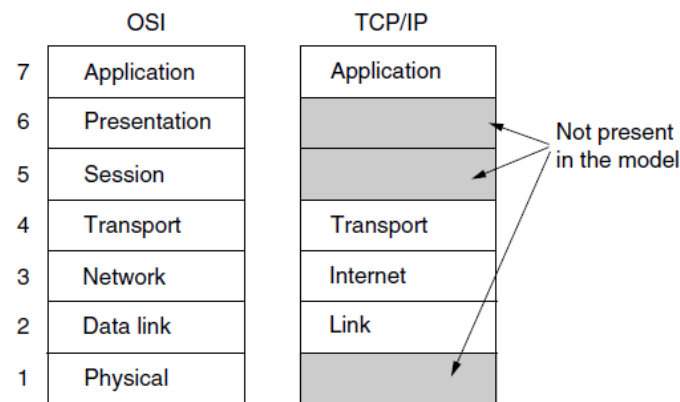


Figura 2: Modelo TCP/IP

Fue diseñado con el objetivo de mantener las conexiones intactas mientras ambos puntos finales de conexión estén funcionando, incluso si alguno de las máquinas o líneas entre ellos fuese dado de baja.

1. La capa de **enlace (Link)** describe que deben cumplir los enlaces (cable de ethernet, líneas telefónicas, etc) para poder usados como medios de trasporte para este tipo de conexión.
2. La capa de **intrared (internet)** permite al host inyectar paquetes en cualquier red y se encarga de hacerlos llegar a destino. Esto se hacer de manera independiente para cada paquete, es decir pueden no seguir el mismo camino e incluso podrían llegar en distinto orden. En el último caso corresponde a la capas superiores reordenarlos para que puedan ser procesados, si es necesario.
3. La capa de **transporte (Transport)** debe ser diseñada para permitir que dos entidades de la red puedan mantener una conversación. Aquí se definieron dos protocolos:
 - El **Transmission Control Protocol (TCP)** que permite enviar sin errores un stream de bytes desde una máquina a otra en la red.
 - El **User Datagram Protocol (UDP)** que permite evitar todo el flujo de conexión TCP y crear el suyo propio. En general, se usa en aplicaciones que requieren una respuesta más rápida que precisa.
4. La capa de **Aplicación (Application)** que se incluye la sesiones y funciones necesarias para codificar y procesar los paquetes enviados y recibidos. Entre los protocolos usados en esta capa se encuentran: TELNNet, FTP y SMTP.

2. Nivel físico

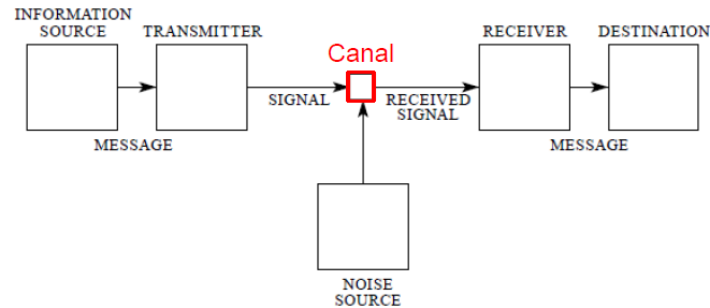


Figura 3: Esquema general de un sistema de comunicación

Cuando enviamos un mensaje a una máquina en una red, debemos pasar esos mensajes por un **transmisor** que convertirá el mensaje en una serie de **señales** que pueden ser enviadas a través del **canal** que nos comunica con la máquina de **destino**. La máquina de destino debe tener un **recibidor** que le permita captar las señales del canal y transformarlas nuevamente en el mensaje original.

Sin embargo, los canales de transmisión físicos no son perfectos y aportan **ruido** a las señales emitidas por nuestro transmisor pudiendo llegar a destino con errores.

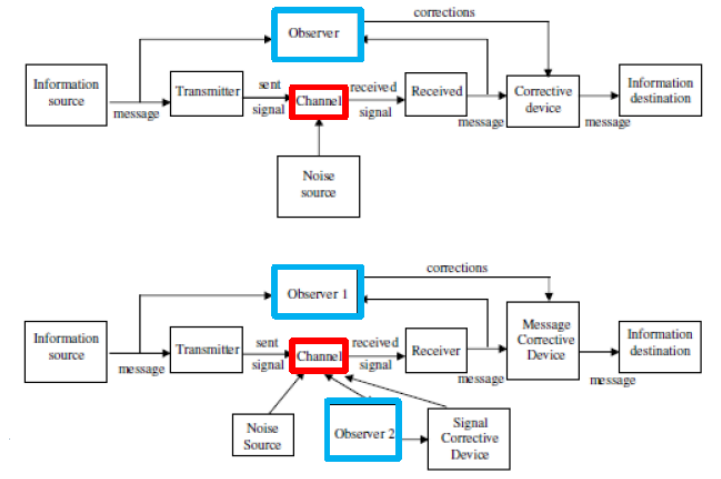


Figura 4: Esquema general de un sistema de comunicación

Una forma de tratar de resolver esto es agregar, al modelo, nuevas capas de actores que trabajan para reducir los efectos nocivos del ruido. Por ejemplo, podemos agregar al modelo observadores externos que sean capaces de ver lo que se transmite de un lado y se recibe del otro, deducir información a partir de las diferencias, y tener chance de enviarle correcciones al

Elemento Corrector. También es posible tener dos niveles de observadores: uno que se maneja a nivel mensaje y otro que se maneja a nivel característica del canal propio.

2.1. Fundamentos de las señales

Las señales que se envían por el canal físico para comunicar dos extremos de un canal son **ondas electromagnéticas** que se propagan a través del canal a una cierta velocidad determinada por el tipo de canal que estemos usando.

Esta onda se desplaza longitudinalmente sobre un plano **vibrando** a una **frecuencia** f determinada con un comportamiento periódico en el eje de su propagación.

En base a este período definimos su **longitud de onda** λ como:

$$\lambda = \frac{c}{f}$$

donde $c = 3 * 10^8 \frac{m}{s}$ es la velocidad de la luz.

Esta onda puede chocar con imperfecciones del material (del canal), producir reflexiones y refracciones que se traducen en **perdidas** (menos energía para la señal original en la que está codificado el mensaje)

Dado que las ondas electromagnéticas son continuas y que son modificadas a medida que se propagan por el canal, debemos encontrar una manera de mapear estas frecuencias a los 0 y 1 que componen nuestros mensajes.

Para esto, tanto el transmisor como el receptor definen un **ancho de banda** que es el rango de frecuencias que va a ocupar las señales que van a ser transmitidas por el canal y dentro de ese rango, cuales deben ser mapeadas a un 1 y cuales a un 0.

2.2. Teoría de la información

2.2.1. Información

Sea E un suceso que puede presentarse con probabilidad $P(E)$. Cuando E tiene lugar, decimos que hemos recibido

$$I(E) = \log \frac{1}{P(E)}$$

unidades de información.

Si introducimos el logaritmo de base 2, la unidad se denomina *bit*. Notemos, también, que si $P(E) = \frac{1}{2}$, $I(E) = 1$ bit. Es decir, un bit es la cantidad de información obtenida al especificar una de dos posibles alternativas igualmente probables.

2.2.2. Fuentes de Memoria Nula

Son fuentes de información que emiten una secuencia de símbolos pertenecientes a un alfabeto finito y fijo $S = \{s_1, \dots, s_n\}$ de manera estadísticamente independientes. Estas fuentes pueden describirse mediante el alfabeto S y las probabilidades con que los símbolos se presentan: $P(s_1), \dots, P(s_n)$.

La información que aporta cada símbolo de la fuente

$$I(s_i) = \log_2 \frac{1}{P(s_i)} \text{ bits}$$

2.2.3. Entropía

Dada una fuente de memoria nula S con alfabeto s_1, \dots, s_n , la entropía $H(S)$ es la cantidad media de información por símbolo de la fuente, es decir:

$$H(S) = \sum_{i=1}^n P(s_i) I(s_i) = \sum_{i=1}^n P(s_i) \log_2 \frac{1}{P(s_i)} = - \sum_{i=1}^n P(s_i) \log_2 P(s_i) \text{ bits}$$

Podemos interpretar la entropía como el **valor medio ponderado de la cantidad de información** del conjunto de mensajes posibles, como una medida de la **incertidumbre promedio (grado de incerteza)** acerca de una variable aleatoria o la **cantidad de información** obtenida al observar la aparición de cada nuevo símbolo.

Propiedades:

- La entropía es no negativa y se anula si y solo si la probabilidad de uno de sus símbolos es 1 y la del resto es 0.
- La entropía máxima (**mayor incertidumbre del mensaje**) se logra cuando todos los símbolos que puede ser emitidos por la fuente son equiprobables.
- Si hay n símbolos equiprobables $P(s) = \frac{1}{n}$ se cumple:

$$H(S) = - \sum_s P(s) \log_2 P(s) = -n \left(\frac{1}{n} \log_2 \frac{1}{n} \right) = -(\log_2 1 - \log_2 n) = \log_2 n$$

2.2.4. Codificación

Dado un alfabeto fuente Σ , un código es una correspondencia entre todas las secuencias posibles de símbolos de Σ a secuencias de símbolos de un alfabeto código X .

Cuando codificamos un alfabeto fuente, buscamos lograr una representación eficiente de la información mediante la eliminación de la redundancia.

Un código es $C : \Sigma \rightarrow X^*$:

- **Bloque:** Si asigna a cada símbolo una secuencia fija de símbolos de X .
- **No singular:** Si todas sus palabras son distintas. Es decir si C es una función inyectiva.
- **Separable o Univócamente Decodificable:** Si ninguna tira de símbolos del alfabeto código admite más de una única decodificación.
- **Instantáneo:** Si cumple la condición libre de prefijos: No existe ninguna palabra que sea prefijo de otra palabra de longitud mayor. Si tenemos este tipo de código entonces cuando el receptor recibe una palabra sabé como tiene que decodificarla, sin la necesidad de esperar más bits.

Si un código es instantáneo entonces es univócamente decodificable.

Inecuación de Kraft: Dado un alfabeto $S = \{s_1, \dots, s_n\}$ y un alfabeto de código $X = \{x_1, \dots, x_m\}$, es condición necesaria y suficiente, para exista un código instantáneo con palabras de longitud l_1, \dots, l_n , que se cumpla la siguiente inecuación:

$$\sum_{i=1}^n |X|^{-l_i} \leq 1$$

Codificación eficiente: Buscamos codificar un alfabeto S de tal forma que maximizar la relación entre la entropía $H(S)$ y la longitud media del código L : Sea L_i la longitud de la palabra que codifica al símbolo s_i de la fuente, p_i la probabilidad de aparición de s_i y r la cantidad de símbolos diferentes en el alfabeto del código entonces:

- $L = \sum p_i L_i$ es la longitud media del código.
- $\log r$ es la cantidad promedio máxima de información de un símbolo del código.
- $h = \frac{H(S)}{L \log r}$ es la eficiencia del código.

La máxima eficiencia se logra cuando $h = 1$. En general, esto sucede cuando se asigna las palabras de código más cortas a los símbolos de fuente más probables.

Codificador óptimo: Es un codificador que usa la menor cantidad posible de bits para codificar un mensaje, es decir: Un codificador se dice óptimo si no existe ningún código para la misma fuente con menor longitud media.

Sea $s_i \in S$, entonces la cantidad de bits necesarios para representarlo en un codificador óptimo es $\lceil \frac{1}{P(s)} \rceil$ y la entropía de

$$H(X) = \sum_{s \in S} P(s) \log_2 \left(\frac{1}{P(s)} \right)$$

Codificación de Huffman: Este método obtiene codificadores óptimos armando un árbol de prioridades en base a la frecuencia de aparición de cada símbolo en un mensaje.

2.3. Los medios de transmisión reales

Estos son los pasos necesarios a realizar durante el envío de datos a través de un canal:

1. Una fuente emite un mensaje hacia un codificador que lo transforma en una secuencia de bits a partir de una codificación idealmente eficiente.
2. El mensaje codificado en bits, pasa a un codificador de canal que se encarga de transformar esos bits a frecuencias para que puedan ser transmitidos. Muchas veces, se agrega redundancia para ayudar a combatir el ruido del canal.
3. Del otro lado del canal, sucede el proceso inverso: Un decodificador de canal lee las frecuencias y las transforma a una secuencia de bits para que, luego, el decodificador del mensaje, la transforme en el mensaje original enviado por la fuente.

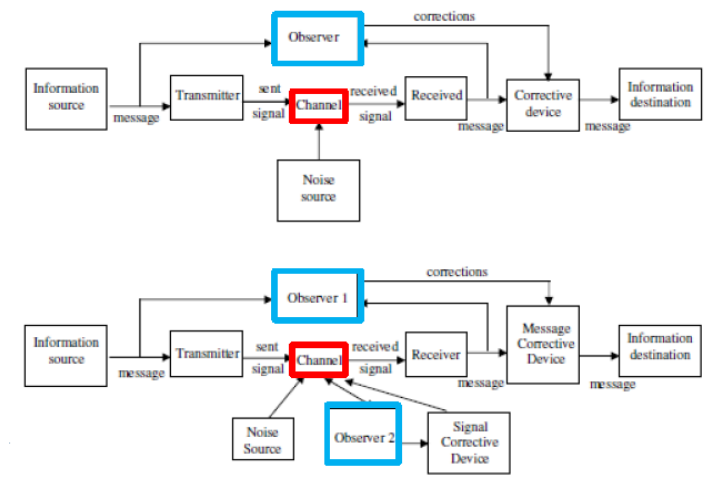


Figura 5: Esquema general de un sistema de comunicación

2.3.1. Problemas de los canales físicos

Como los canales de transmisión existentes no son perfectos, la señal recibida por el decodificador puede diferir de la transmitida debido a una gran variedad de tipos de errores:

- **Atenuación:** En medios análogos, las señales se degradan con la distancia recorrida lo que puede llevar a provocar errores en algunos bits recibidos. Por lo que la intensidad de la señal recibida debe ser suficiente para ser detectada y, además, debe ser suficientemente mayor al ruido del canal para que se reciba sin error. En general, las frecuencias más afectadas son las más altas por lo que se puede ecualizar estas frecuencias, es decir, amplificarlas.
- **Distorsión de retardo:** En medios guiados, la velocidad de propagación en el medio varía con la frecuencia por lo que los componentes del mensaje llegan en distintos instantes de tiempo, originando desplazamiento de fases entre las distintas frecuencias.
- **Ruido:** Los canales físicos poseen ruido natural. Es decir, transmiten señales adicionales debido a agentes externos:
 - **Ruido Térmico ó Ruido Blanco:** Se produce debido a la agitación térmica de electrones y aumenta linealmente con la temperatura absoluta del canal.
 - **Ruido por intermodulación:** Son señales que son la suma o la diferencia de sus frecuencias originales producidas por una falta de linealidad en el canal.
 - **Ruido por Diafonía:** Se produce cuando una señal de una línea interfiere en otra.
 - **Ruido impulsivo:** Son impulsos irregulares o picos que se pueden producir por interferencias externas (como pueden ser interferencia electromagnéticas, tormentas, etc). Este tipo de ruido es de corta duración, tienen gran amplitud y es disruptivo.

2.3.2. Capacidad de un canal

Estas perturbaciones afectan la velocidad de transmisión del canal por lo que debemos asegurarnos de no enviar más bits que la capacidad límite del mismo para no perder información durante la transmisión. Para esto vamos a definir los siguientes conceptos:

- Vamos a nombrar C a la capacidad del canal, es decir a la cantidad de bits por segundo que podemos transmitir a través del mismo.
- B es el ancho de banda por el cual vamos a transmitir nuestros datos, va estar medido en ciclos por segundo (Hertz) y va estar limitado por el transmiso y el medio.
- N es el nivel medio o potencia del ruido del canal
- BER es la tasa de errores de bits por segundo (Bit Error Rate).
- S es la potencia o amplitud de la señal.
- $SNR = S/N$ es la relación señal a ruido medido en decibeles.

$$C = B \log_2(1 + SNR)$$

En principio, si se aumentan el ancho de banda B y la potencia de señal S , aumenta la velocidad binaria. Sin embargo, un aumento de B aumenta el ruido y un aumento de S aumenta las no linealidades y el ruido de intermodulación.

Marco de Referencia

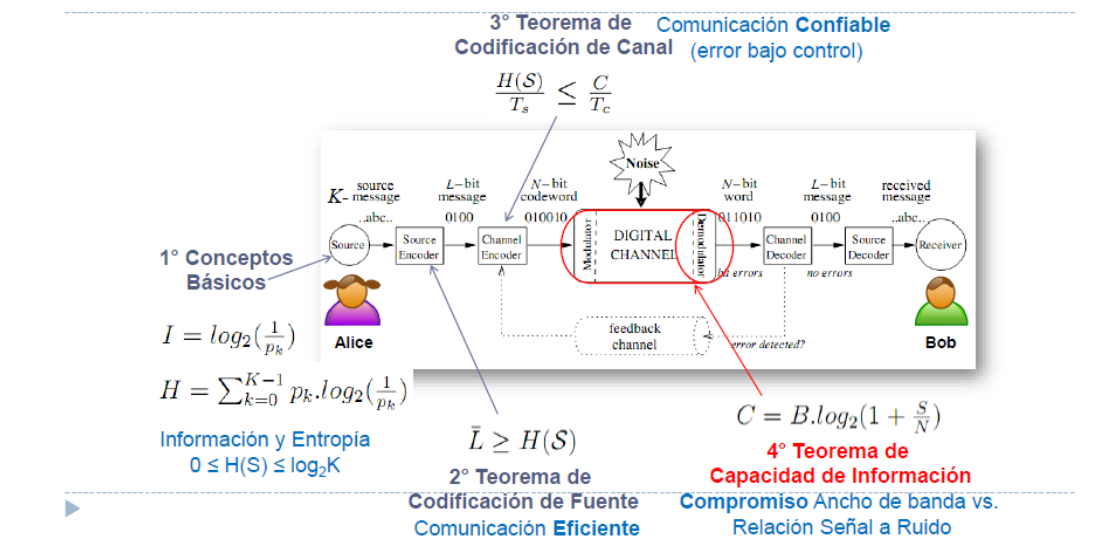


Figura 6: Esquema de un sistema de comunicación y sus conceptos asociados

Límite de eficiencia: La eficiencia de ancho de banda es la máxima cantidad de bits por segundo que podemos inyectar por cada Hz sin perder información. Mientras más eficiente sea el canal, se pueden transmitir más bits por segundo.

Límite de confiabilidad: Es la cantidad máxima de bits por segundos que podemos utilizar para transportar una señal de manera confiable a través de un canal ruidoso Ancho de banda. Medios de transmisión guiados y no guiados. Punto a Punto y Compartidos (Broadcast). Modems. Modulación: Fase, Frecuencia, Amplitud y QAM. Conversión Analógico/Digital. Capacidad de Canal. Latencia. Producto delay-ancho de banda.

3. Nivel Enlace

Punto a Punto. Servicios. Enrutamiento. Manejo del enlace. Control de flujo. Control de errores: detección y corrección. Distancia de Hamming. Protocolo Stop & Wait. Protocolos de ventana deslizante. Comparación de protocolos. Medidas de eficiencia.

4. Medios Compartidos

Compartidos. Redes locales (LANs). Generalidades. Medios de transmisión. Nivel físico. Nivel de enlace de datos. Subcapa LLC (Logical Link Control): IEEE802.2. Subcapa MAC (Medium Access Control): IEEE 802.3 (Ethernet). CSMA/CD. Redes Inalámbricas. Frecuencia dedicada versus expandida (Spread Spectrum). WLANs. CSMA/CA. Wi-Fi: IEEE 802.11b/g/n. Repetidores. Puentes. LAN Switches. Conceptos de VLAN y troncales de VLANs (IEEE 802.1Q). Spanning Tree Protocol.

5. Nivel Red

Conmutación y Forwarding. Subredes. Implementación: circuitos virtuales y datagramas. Control de flujo. Concepto de Ruteo. Protocolo IP. Direccionamiento. Broadcasting. Ejemplos de subnetting. Protocolo ARP. Forwarding. ICMP. Traducción de direcciones (NAT).

6. Ruteo Externo e Interno

Distance Vector y Link State. Los protocolos RIP y OSPF. Áreas. Inundación confiable.

7. Nivel de transporte

Servicios. Primitivas. Protocolos. Servidores de nombres. Manejo de conexión: establecimiento, uso y liberación. Manejo de conexión basados en tiempo. Direccionamiento. Control de flujo. Asignación de buffers. Recuperación de caídas. Multiplexado. Protocolos de nivel 4: Transport

Control Protocol (TCP). User Datagram Protocol (UDP) . Mecanismos de control de congestión. Cálculo del RTO. Control de Flujo. Control de errores. Determinación de la performance.

Congestión

Introducción al problema de congestión. Curvas de Trafico Enviado vs entregado. Resultado con buffer infinito. Causas de congestión. Control de flujo vs Control de congestión. Taxonomía de Yang y Redan. Soluciones de lazo cerrado y abierto. Concepto de sistemas realimentados. Métricas a sensar para las realimentación. Realimentación implícita y explícita. Determinación de la performance.

8. Aplicaciones

Aplicaciones. Correo Electrónico. Protocolos : SMTP, POP3 e IMAP, MIME. Servidores World Wide Web. HTTP. Servidor de Nombres: DNS. Jerarquía de dominios. Resolución de nombres.

9. Seguridad

Seguridad en Redes. Marco de Trabajo. Criptografía. Seguridad. Privacidad. Protocolos de Clave Pública y Privada. Algoritmos: DES, 3DES, AES, RSA, MD5 y SHA .Ventajas y desventajas de cada uno. Sus aplicaciones (Autorización, Firma, Confidencialidad e Integridad). Distribución de Claves Públicas. Firewalls. Tunneling. Conceptos de amenazas, ataques, intrusiones.