

## CAPA DE RED

La capa de red tiene como función transportar paquetes desde un host emisor a un host receptor. Cada equipo final implementa la capa de red y superiores; los routers solo hasta la capa de red.

REENVÍO: Acción que realiza un router al transferir un paquete desde una interfaz de entrada a una interfaz de salida.

ENRUTAMIENTO: Proceso que realiza la red en conjunto para determinar las rutas terminal a terminal que los paquetes siguen desde el origen al destino.

▷ Obs: todo router tiene una tabla de reenvíos.

ESTABLECIMIENTO DE LA CONEXIÓN: La tercera subfunción importante de la capa de red (en algunas arquitecturas de redes) es configurar la conexión.   
↳ como frame-delay, ATM, X.25, pero no Internet.

Antes del flujo de datagramas, las dos máquinas finales y los routers implicados en la conexión negocian entre sí para configurar la conexión.

MODELO DE SERVICIO DE RED: Define las características del transporte terminal a terminal de los paquetes entre los sistemas receptor y emisor. Algunos servicios son:

- Entrega garantizada
- Entrega garantizada con retardo limitado
- Entrega de paquetes en orden
- Ancho de banda mínimo garantizado
- Fluctuación máxima garantizada

Obs: Internet usa el modelo del mejor esfuerzo (best effort)

REDES DE CIRCUITOS VIRTUALES Y DE DATAGRAMAS: Manera de

enfocar la implementación de un nivel 3. Una red de datagramas proporciona un servicio sin conexión host a host. Las redes de circuitos virtuales proporcionan un servicio orientado a la conexión.

## REDES DE CIRCUITOS VIRTUALES

Un circuito virtual consta de:

- Una ruta: serie de enlaces y routers entre host origen y destino.
- Números VC: un n° por cada enlace a lo largo de la ruta.
- Entradas en la tabla de reenvío en cada router del camino.

La tabla de reenvíos de cada router incluye la traducción de VC. Cuando se configura un número de VC nuevo en un router, se añade una entrada a la tabla de reenvío. Cuando un VC termina, las entradas apropiadas se eliminan de cada tabla.

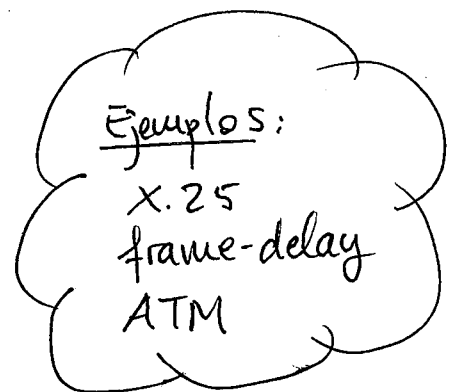
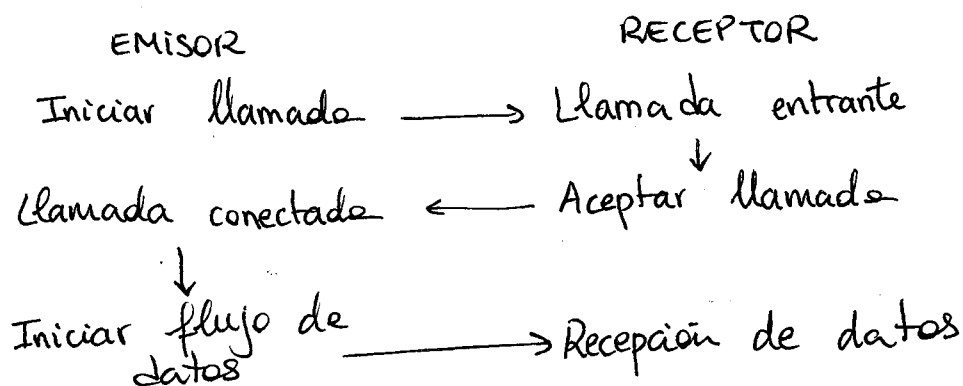
Un paquete no mantiene el mismo número de VC en cada enlace por dos razones:

1. Hace más pequeño el campo número de la cabecera del paquete.
2. La configuración VC se simplifica considerablemente.

En una red de CV, los routers mantienen información del estado de la conexión.

Tres fases:

- Configuración del VC (orientado a la conexión)
- Transferencia de datos
- Terminación del VC.



## REDES DE DATAGRAMAS

Las redes de datagramas no inicializan la conexión. Los routers no saben nada de las conexiones extremo-a-extremo.

Los paquetes se reenvían usando la dirección de la máquina destino, por lo que pueden tomar distintos caminos.

Específicamente, cada router tiene una tabla de reenvíos que asigna direcciones de destino a interfaces de enlace.

\* Cuando existen varias coincidencias, el router aplica la regla de coincidencia con el prefijo más largo.

## ¿REDES DE DATAGRAMAS O VC? ¿QoS?

### DATAGRAMAS (ej: IP)

- Datos intercambiados por ordenadores
- Servicio flexible
- Capaces de ejecutar control sobre el rendimiento
- Robustez y recuperación de fallos
- Más flexible ante cambios en los niveles inferiores al ser más simple.

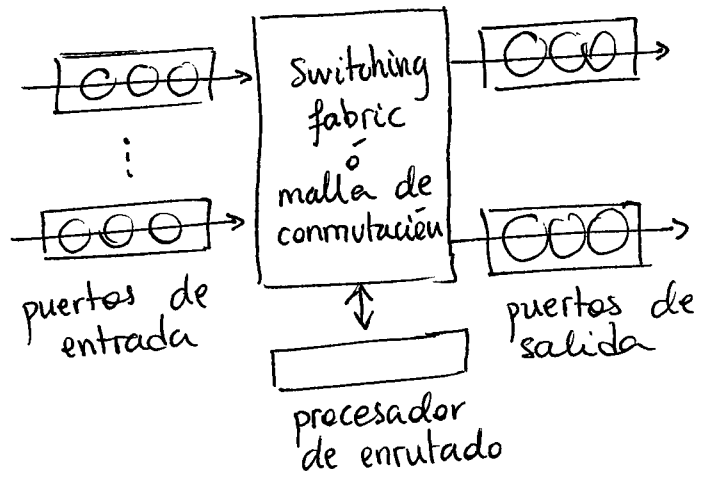
### VC (ej: ATM)

- Evolución telefónica móvil
- Servicio predecible
- Señalización inicial (peseado)
- Menos robusta ante fallos
- La complejidad se deja en la red.

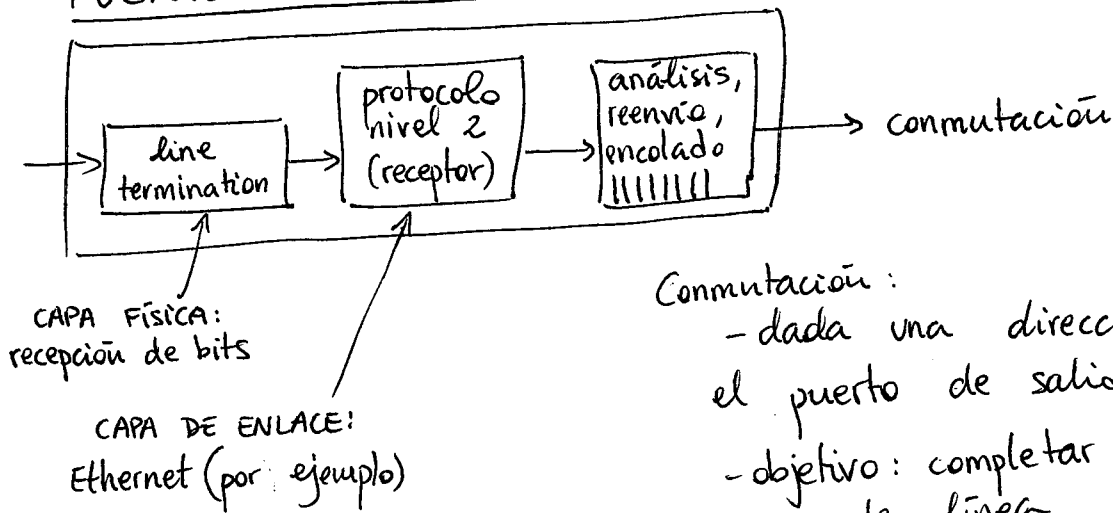
# INTERIOR DE UN ROUTER

Las funciones clave:

- Ejecutar algoritmos/protocolos de enrutado
- Reenvío de datagramas de la entrada a la salida.



## PUERTOS DE ENTRADA



Conmutación:

- dada una dirección destino buscar el puerto de salida adecuado
- objetivo: completar el proceso a tasa de línea
- encolado: si dos datagramas llegan más rápido que la tasa de reenvío estos se encolan en buffers

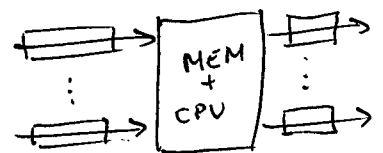
## MALLAS DE CONMUTACIÓN

La funcionalidad del enrutado de conmutación es transferir paquetes desde el buffer de entrada al buffer de salida.

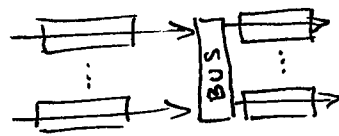
Velocidad de conmutación: tasa a la que los paquetes pueden ser transferidos desde los puertos de entrada a los de salida. Dadas  $N$  entradas la vel. de conmut. debería ser al menos  $N$  veces la tasa de línea por enlace.

Tipos de mallas:

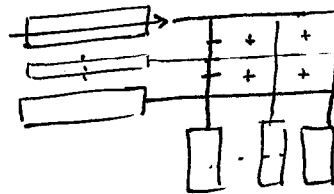
► Memoria + CPU: Se usa una memoria manejada por una CPU para el paso de datos de la entrada a la salida. Obsérvese que si el ancho de banda de la memoria es tal que pueden escribirse/leerse  $B$  bytes por segundo, entonces la tasa global de reenvío es menor que  $B/2$ .



► Conmutación bus común: la transferencia puertos entrada-puertos salida se realiza mediante un bus. Evita la memoria + CPU, pero el bus es común para todos los puertos. El ancho de banda de conmutación está limitado por la velocidad del bus.

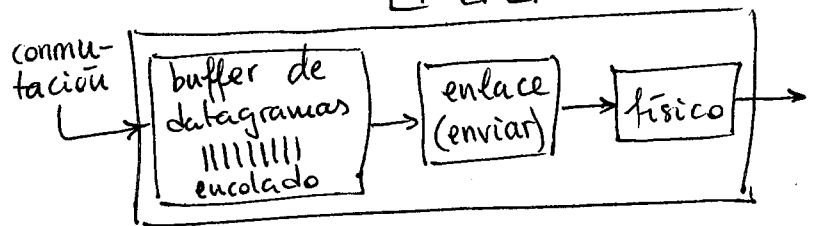


► Red de interconexión: consta de  $2n$  buses que conectan los  $n$  puertos de entrada a los  $n$  puertos de salida. Se evita la utilización de un único bus pero puede haber retenciones cuando dos puertos de entrada quieren enviar al mismo puerto de salida.



## PUERTOS DE SALIDA

Se requieren buffers cuando los datagramas llegan más rápido que la velocidad de transmisión.



Planificador de paquetes: debe determinar que paquete de los buffers de entrada transmitir en cada instante de tiempo.

Pérdidas en el caso de que el buffer del puerto de salida se desborde.

¿Cuánto buffer?

→ "rule of thumb":  $\frac{C}{RTT_{\text{típico}}}$   $C = \text{capacidad de enlace}$   
 $RTT_{\text{típico}} = 250 \text{ ms}$

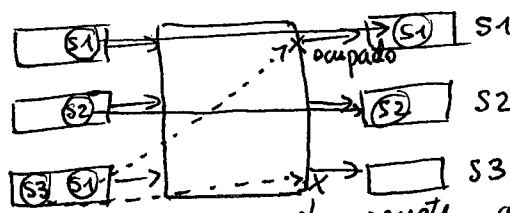
→  $N$  flujos:  $\frac{RTT \cdot C}{\sqrt{N}}$

→ RED: detección aleatoria temprana

## ENCOLADO EN LOS PUERTOS DE ENTRADA

Conmutación más lenta que el agregado de la tasa de llegadas. Puede haber retardo e incluso pérdidas (buffer overflow).

Bloqueo Head-of-the-line (HOL): datagramas bloqueados en la primera posición del buffer, bloquean a su vez a otros datagramas.



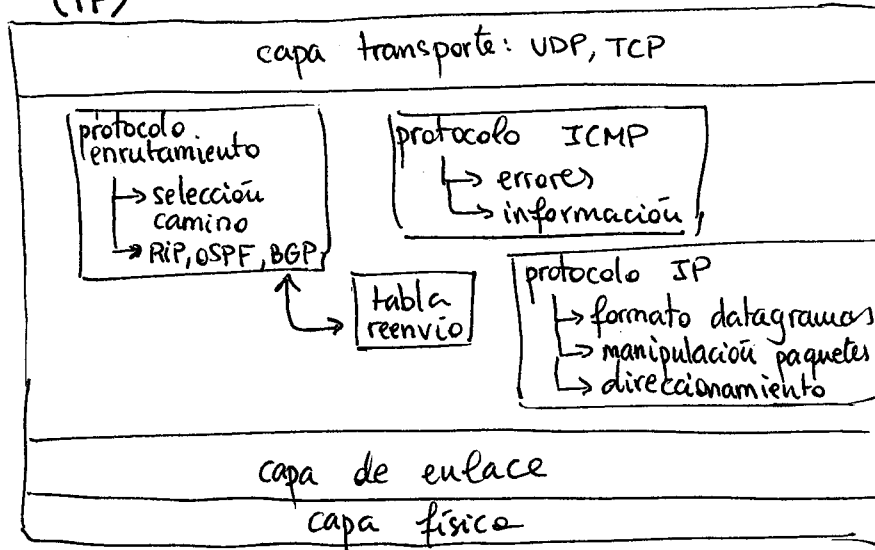
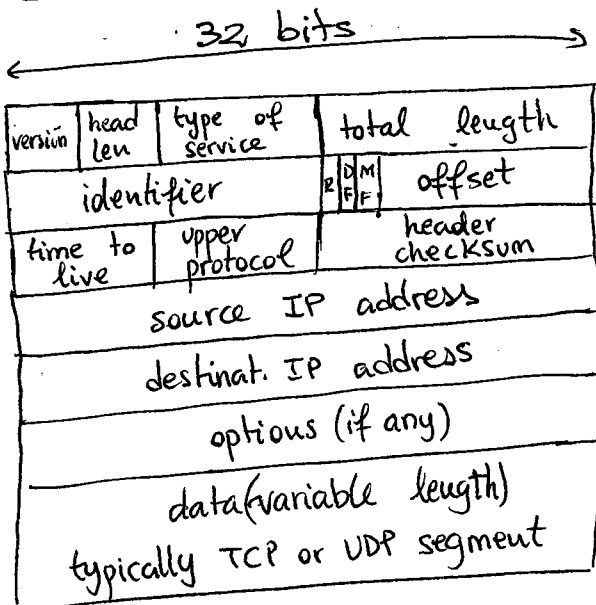
el paquete anterior le bloquea

# PROTOCOLO DE INTERNET (IP)

Funcionalidades:

- enrutamiento
- fragmentación
- ICMP

## CABECERA IPv4



versión (4 bits): versión del protocolo IP del datagr.

IHL (4 bits): dado que las opciones pueden tener long. variable. Este campo indica el tamaño de cabecera en nº de palabras 32b.

Tipo servicio: parámetro de calidad del servicio  
Longitud total datagrama: longitud en octetos de IP header + segment.

Identificador: identifica unívocamente un datagr. útil en caso de fragmentación.

Flags → bit de reservado  
 → bit de don't fragment  
 → bit de more fragments

Desplazamiento (13 bits): En paquetes fragmentados indica la posición. Múltiplo de 8 ( $2^3$ ).

Tiempo de vida: Indica el número máximo de enrutadores que un paquete puede atravesar. Típicamente 64 ó 128.

Protocolo superior: Indica el protocolo de transporte. Útil en el host final.

Checksum: suma de control de la cabecera

Dirección IP origen: 32 bits en formato de red.

Dirección IP destino: 32 bits en formato de red.

Opciones (longitud variable): Nos permite ampliar una cabecera IP.

Datos/Carga útil: Normalmente contiene una cabecera de la capa de transporte + datos a enviar.

\* Tamaño cabecera IPv4 sin opciones: 20 bytes

## FRAGMENTACIÓN IP y REESAMBLADO


La cantidad máxima de datos que una trama de la capa de enlace puede transportar se conoce como unidad máxima de transmisión MTU

Por esta razón, los datagramas IP grandes son divididos en fragmentos en algún router de la red. Un datagrama pasa a ser varios datagramas. El reesamblado solo se realiza en el host final.

### PROS Y CONTRAS

- Pros: independencia y flexibilidad de niveles inferiores
- Contras: trabajo significativo en los extremos  
trabajo extra en los routers que fragmentan  
debilidad ante ataques → sin final de datagrama  
→ numeración incorrecta

## DIRECCIONAMIENTO IP

Una dirección IP se divide en  parte de red (bits altos)  
parte del host/interface (bits bajos)

Una red es el conjunto de interfaces que tienen la misma parte de red

Una red local es el conjunto de elementos que pueden comunicarse entre sí sin la intervención de un router.

### ESTRATEGIA DE DEFINICIÓN DE RANGOS: CIDR

CIDR: enrutamiento de dominios sin clases

Notación CIDR: a.b.c.d/x donde x indica el nº de bits de la parte

Notación máscara en decimal: 255.255.255.0 ( $\equiv$  /24)

- Direcciones IP reservadas: no pueden asignarse a las interfaces direccionales que empiecen por 127 (dirección loopback).  
Tampoco se pueden asignar direcciones con parte de interfaz con todo 0's (identifican rangos) o todo 1's (dirección broadcast)
- Direcciones IP privadas: conjunto de direcciones que por convenio solo se pueden usar de manera interna en una red
- Direcciones válidas, pertenecientes al rango, asignables.

## CONFIGURACIÓN EQUIPO IP

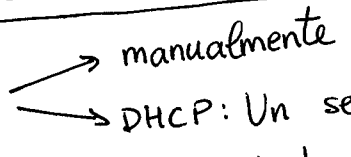
Las interfaces configuradas con IP requieren:

- dirección IP
- dirección del router de puerta de enlace (gateway o router del primer salto).
- máscara de red (para identificar si la IP es parte de la red local o no)

En los host típicamente también DNS.

ARP - Address resolution protocol - Routing to another LAN

## ¿CÓMO CONSEGUIR UNA DIRECCIÓN IP?

En HOST parte interfaz 

### DHCP: Dynamic Host Configuration Protocol

El objetivo es permitir a los host obtener dinámicamente su dirección IP cuando estos.

Funcionamiento:

→ DHCP discover

1. DESCUBRIMIENTO DEL SERVIDOR DHCP: Se envía un mensaje al servidor DHCP, que envía el cliente dentro de un paquete UDP-IP-ETH-Phy a la dirección de broadcast.

2. OFERTA DEL SERVIDOR DHCP: Cuando un servidor DHCP recibe un DHCP discover, responde con un mensaje de oferta DHCP que distribuye a todos los nodos de la red por la dir. broadcast. El mensaje de oferta lleva ID del mensaje DHCP discover, la IP propuesta, la máscara de red y el tiempo de arrendamiento de la IP (durante el que la dir. IP será válida).

3. SOLICITUD DHCP: El cliente seleccionará entre las ofertas del servidor y responderá con un mensaje de solicitud DHCP, devolviendo los parámetros de configuración.

4. ACK DHCP: El servidor contesta al mensaje de solicitud DHCP con un mensaje ACK DHCP de confirmación.



VNA RED : Las ISP reciben bloques de direcciones IP.  
parte de red

El direccionamiento jerárquico permite el anuncio de las rutas eficientemente.

## ADMINISTRACIÓN DE INTERNET

¿Quién facilita a las ISPs los bloques de direcciones?

↳ ICANN: Internet Corporation for Assigned Names and Numbers

- asigna direcciones
- gestiona los DNS's.
- asigna nombres de los dominios y disputas
- La gestión está parcialmente dividida en áreas geográficas.  
(RIPE Europa)
- Institución sin ánimo de lucro
- Solo cede su uso a socios

## NAT: NETWORK ADDRESS TRANSLATION

Motivación: Los usuarios de las redes locales usan una única dirección IP en lo que respecta al mundo "exterior":

- El ISP solo facilita una IP a cada red doméstica/oficina.
- Puede modificar direcciones internas sin notificación al resto del mundo
- Puede cambiar la parte de red sin cambios en las direcciones internas.
- Los elementos internos no son visibles en el exterior (↑ seguridad).

Implementación: Un router NAT debe:

- Datagramas salientes: reemplazar (IP origen, puerto origen) de cada datagrama saliente a (dir. NAT IP, nuevo puerto).
- Almacenar en la tabla NAT cada par (IP origen, puerto origen) x (dirección NAT IP, nuevo puerto).
- Datagramas entrantes: reemplazar (dir. NAT IP, nuevo puerto) en el campo destino de cada datagrama entrante con la correspondiente entrada (IP origen, puerto) almacenados en la tabla NAT.

Parte oscura:  
↓  
IPv6

↔ los routers nvl. 3 no deberían tocar elementos nvl. 4  
↘ los puertos deben determinar procesos  
↘ dificultad de accesibilidad externa  
↘ sensible ante errores

## NAT ESTÁTICA

Redirigir toda la conexión entrante a un puerto dado a una dirección concreta.

Útil para servidores e imposible con aplicaciones con puertos aleatorios

## NAT TRANSVERSAL SIMPLE

En el caso de que un extremo no esté detrás de un router NAT se puede usar un host intermedio.

El host intermedio debe ser accedido por el extremo con NAT (iniciar sesión de muchas aplicaciones).

El extremo sin NAT contacta con el servidor intermedio y el host con NAT es avisado para iniciar conexión con el host sin NAT que es públicamente accesible (inversión de la conexión).

## NAT TRANSVERSAL SIMÉTRICO (STUN)

En el caso de que ambos estén detrás de una NAT, ambos pueden añadir una entrada en la tabla NAT de forma "artificial" y darla a conocer al otro extremo.

Se necesita un servidor (IP pública) intermedio y ambos extremos abren una conexión, tras que uno de ellos haya hecho conocer al servidor que quiere conectarse con el otro.

El servidor comunica a ambos extremos de la entrada creada en la tabla del otro y se inicia la conexión.

## NAT TRANSVERSAL CON RETRANSMISIONES (TURN) ¿iSkype?

Clientes con NAT inician sesión en nodos de retransmisión.

Ambos extremos conectados y el retransmisor "puentea" los paquetes entre ambos.

## INTERNET GATEWAY DEVICE (IGD) UPnP

Permite gestionar y hacer pública la tabla de traducciones NAT.

# ICMP: INTERNET CONTROL MESSAGE PROTOCOL

Los hosts y los routers utilizan ICMP para intercambiarse información acerca de la capa de red. Su uso más típico es la generación de mensajes de error.

Ejemplos: eco, redirección, traceroute

Un mensaje de error ICMP nunca se genera como respuesta a:

- otro mensaje de error ICMP
- un datagrama con IP-dest broadcast
- un datagrama con dirección física broadcast
- un fragmento del datagrama  $\neq$  primero
- un datagrama cuya dir. IP no sea de un único host.

## FORMATO

tipo	código	checksum
depende del tipo y código		

\* En los msjs de error se copia la cabecera y los 8 primeros byte de datos del datagrama

## IPv6

El conjunto de direcciones de 32 bits va a ser próximamente completamente abocadas.

• Cabecera fija de 40 bytes

• Fragmentación no permitida.

## FORMATO

32 bits			
ver	priority	flow label	
payload len	upper hdr	hop limit	
source address (128 b)			
dest. address (128 b)			
data			

- versión (4 bits)
- Clase de tráfico (8 bits): prioridad del paquete
- Etiqueta flujo (20 bits): manejo de la QoS.
- Long. datos (16 bits): long. total. campo datos
- Cabecera siguiente (8 bits): protocolo capa superior
- Límite de saltos (8 bits): tiempo de vida

Obs: • Checksum eliminado para reducir el coste de proceso por salto.

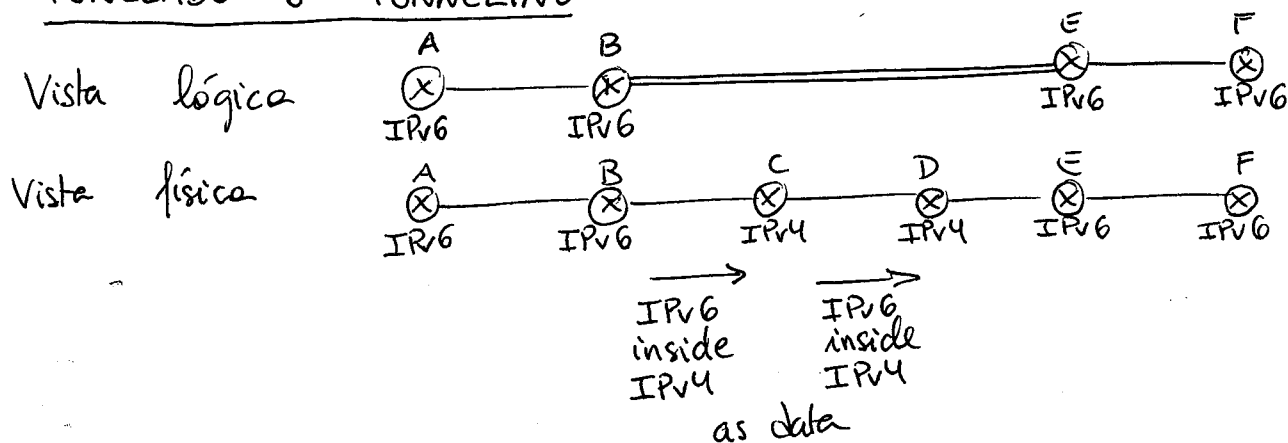
• Opciones permitidas pero fuera de la cabecera, indicado por el campo "next header".

• ICMPv6: nueva versión de ICMP.   
 → nuevos msjs ("packet too big")   
 → funciones de gestión de grupos

• Problema: no todas los routers pueden ser actualizados simultáneamente.

→ solución   
 → tunelado   
 → pila dual

## TUNELADO o TUNNELING



## PILA DUAL

Nodos IPv6 también disponen de una implementación IPv4 (conocidos como nodos IPv6/IPv4). Al comunicarse con un nodo IPv4, el nodo IPv6/IPv4 puede utilizar IPv4. Análogo para un nodo IPv6.

## ALGORITMOS DE ENRUTAMIENTO

### CLASIFICACIÓN DE ALGORITMOS

¿Información global o descentralizada?

GLOBAL → todos los routers conocen la topología completa y el coste de los enlaces  
Ejemplo: algoritmos de estado de enlace

DESCENTRALIZADO → Un router conoce sólo los routers vecinos con enlaces físicos y su coste  
→ proceso iterativo de cálculo que implique el intercambio de información con los vecinos  
Ejemplo: algoritmos de vector distancia

¿Dinámico o estático?

ESTÁTICO: La información en los routers cambia poco con el tiempo

DINÁMICO: Los routers cambian más rápido  
→ actualizaciones periódicas  
→ responden a los cambios en el coste de los enlaces.

## ESTADO DE ENLACES

El algoritmo por excelencia es el algoritmo de Dijkstra.

- La topología de la red y el coste de los enlaces son conocidos por todos los nodos. Se consigue vía difusión del estado de enlaces. Todos los nodos tienen la misma información.
- Calcula los caminos de mínimo coste de un nodo "origen" a todos los demás nodos.  $\Rightarrow$  Define la tabla de reenvíos del nodo
- Coste computacional:  $O(n^2)$       \* check algoritmo diapos/internet

## VECTOR DISTANCIAS

Un algoritmo de encadenamiento basado en vector distancias es iterativo (continúa mientras los nodos se intercambien información), asíncrono (no es necesario que los nodos intercambien información en momentos determinados ni con un orden fijo) y distribuido (cada nodo intercambia información solo con sus vecinos).

Estructura de datos: tabla de distancias  $\begin{cases} \rightarrow \text{cada nodo tabla propia} \\ \rightarrow \text{posible destino por fila} \\ \rightarrow \text{vecino inmediato por columna} \end{cases}$

Usa la ecuación de Bellman-Ford:

$$d_x(y) = \min_v \{c(x,v) + d_v(y)\}$$

Algoritmo: En cada nodo:  $\begin{cases} \rightarrow \text{esperar cambios en algún enlace a sus vecinos} \\ \downarrow \\ \text{recalcular la tabla de distancias} \\ \downarrow \\ \text{si ha cambiado algún camino de coste mínimo, notificar a los vecinos} \end{cases}$

## CAMBIOS EN LOS COSTES

- Las buenas noticias viajan rápido
- Las malas noticias lento (problema de cuenta al infinito)
- Puede exigir muchas iteraciones hasta que el algoritmo pare
- Inversa envenenada: Z dice a Y que su distancia a X es  $\infty$ , así Y nunca enrutará por Z para ir a X.

# COMPARACIÓN ENTRE ALGORITMOS

## COMPLEJIDAD DE LOS MENSAJES

EE: Con  $n$  nodos y  $E$  enlaces,  $O(nE)$  mensajes enviados en cada ciclo.

VD: intercambio de mensajes entre vecinos únicamente.

## VELOCIDAD DE CONVERGENCIA

EE: algoritmo  $O(n^2)$  que requiere de  $O(nE)$  mensajes.

VD: tiempo de convergencia variable

- └→ puede haber bucles de encadenamiento
- └→ problema cuenta al infinito

## ROBUSTEZ ¿Qué pasa si hay errores en el router?

EE: • Un nodo puede enviar costes de enlace erróneos

- Recuperables en la siguiente ejecución.
- Cada nodo calcula solo su propia tabla (errores de cálculo solo dañinos de forma local).

VD: • Un nodo puede enviar costes de rutas erróneos.

- Los errores se propagan por la red y difícil recuperarse.

- La tabla de un nodo sirve para el resto de tablas (errores de cálculo globalmente peligrosos).

## ENRUTAMIENTO JERÁRQUICO

El estudio hasta ahora ha sido una idealización. En escala real los hosts se cuentan por millones (imposible manejar todas esas direcciones y mensajes). En su lugar se opta por una autonomía administrativa: internet = red de redes (cada administrador controla su subred).

Los routers se agrupan formando SISTEMAS AUTÓNOMOS (AS). Los routers del mismo AS ejecutan el mismo algoritmo de enrutado. Para salir del AS existe la puerta de enlace o gateway: routers en los bordes de los AS que se conectan con otros.

La tabla de reenvíos está configurada tanto por los algoritmos intra-AS y inter-AS. Para configurar la tabla de reenvíos, el router debe determinar hacia que puerta de enlace debe ser dirigido para el destino x. Varias opciones:

- hot potato routing: puerta de enlace con menor coste
- Minimizar el nº de AS de la ruta
- Azar
- Precio/político

## ENRUTAMIENTO EN INTERNET

### ENRUTADO INTRA-AS (también conocido: INTERIOR GATEWAY PROTOCOLS (IGP))

Los protocolos de enrutado intra-AS (más comunes) son:

- RIP: Routing Information Protocol
- OSPF: Open Shortest Path First
- IGRP: Interior Gateway Routing Protocol (Cisco proprietary)

### RIP

Se basa en el algoritmo de vector distancias. Considera como métrica el nº de saltos (enlaces coste 1). Los VD's son intercambiados entre vecinos cada 30 sec (aprox.), y cada anuncio tiene como máximo 25 destinos (subredes).

Los fallos en los enlaces y recuperación de RIP son los siguientes:

- Si no se reciben anuncios después de 180s se considerará el vecino y su enlace como "muertos", y las rutas vía ese vecino son inválidas. Esto es informado al resto de vecinos, quienes informan de cambios en sus tablas.
- Emplea inversa envenenada (poison inverse), en este caso distancia infinita = 16 saltos.

### OSPF

Usa un algoritmo de estado de enlaces: se mandan paquetes en inundación a todo el AS, de este modo se recupera la topología y los costes.

Las rutas se computan usando Dijkstra.

Se envía información por cada cambio al menos 1 vez cada 40 minutos.

### ENRUTADO INTER-AS (Internet)

El protocolo aceptado por todos como algoritmo de enrutado inter-AS es el BGP (Border Gateway Protocol).

BGP → eBGP: obtener la alcanzabilidad de red de sus AS's vecinos  
→ iBGP: propagar esta información a todos los routers del AS

Permite a las redes darse a conocer en Internet.



## BGP SELECCIÓN DE RUTA

Si los routers tienen más de una opción para un AS destino, la selección BGP es en este orden:

1. Política establecida por el gestor: muy frecuente, falta de colaboración por la gran competitividad.
2. Menor número de AS: vector de distancias
3. Puerta de enlace con ruta más corta dentro del AS: patata caliente.
4. Criterios adicionales (p.ej: azar)

¿PORQUÉ DIFERENTE ENRUTADO INTRA- y INTER-AS?

► Política:

- Inter-AS: la administración puede controlar el tráfico que la atraviesa, en términos económicos y políticos.
- Intra-AS: una sola administración

Escala:

Los routers dentro de un AS pueden ser pocos, lo que motiva un algoritmo tipo EE. En cambio, en un AS grande, es más razonable un algoritmo VD.

Rendimiento:

- Inter-AS: primero términos políticos
- Intra-AS: esencialmente el objetivo final

# TIPOS DE RELACIONES

- PEERING: Ambas extremos se intercambian rutas de sus redes y clientes. Sin coste pero cumpliendo requisitos.
- CLIENTE-PROVEEDOR: El proveedor permite servir como tránsito para el tráfico entrante y saliente de sus clientes. El cliente paga por ese tráfico.  
¿Cómo se paga? Históricamente en función del agregado de volumen. En la actualidad por el percentil 95th del caudal. También usado el basado-en-pico-agregado: contribución a la máxima utilización del enlace.  
¿Dónde?
  - Puntos de intercambio
  - Enlace directo

