

Introduccion a TCP/IP

Carlos Vicente
Hervey Allen
Carlos Armas

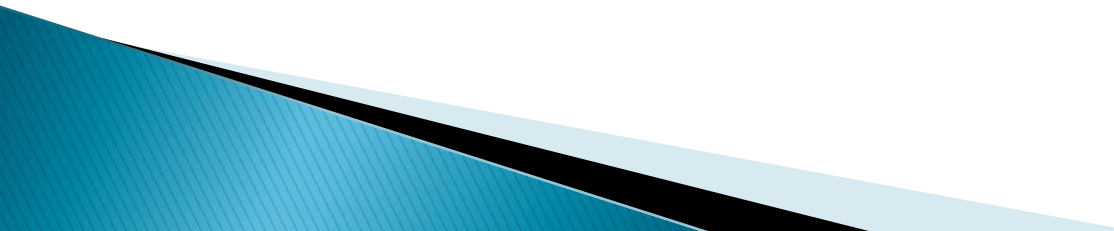


Este documento es producto de trabajo realizado por Network Startup Resource Center (NSRC at <http://www.nsrc.org>). Este documento puede ser libremente copiado o re-utilizado con la condicion de que toda re-utilizacion especifique a NSRC como su fuente original.

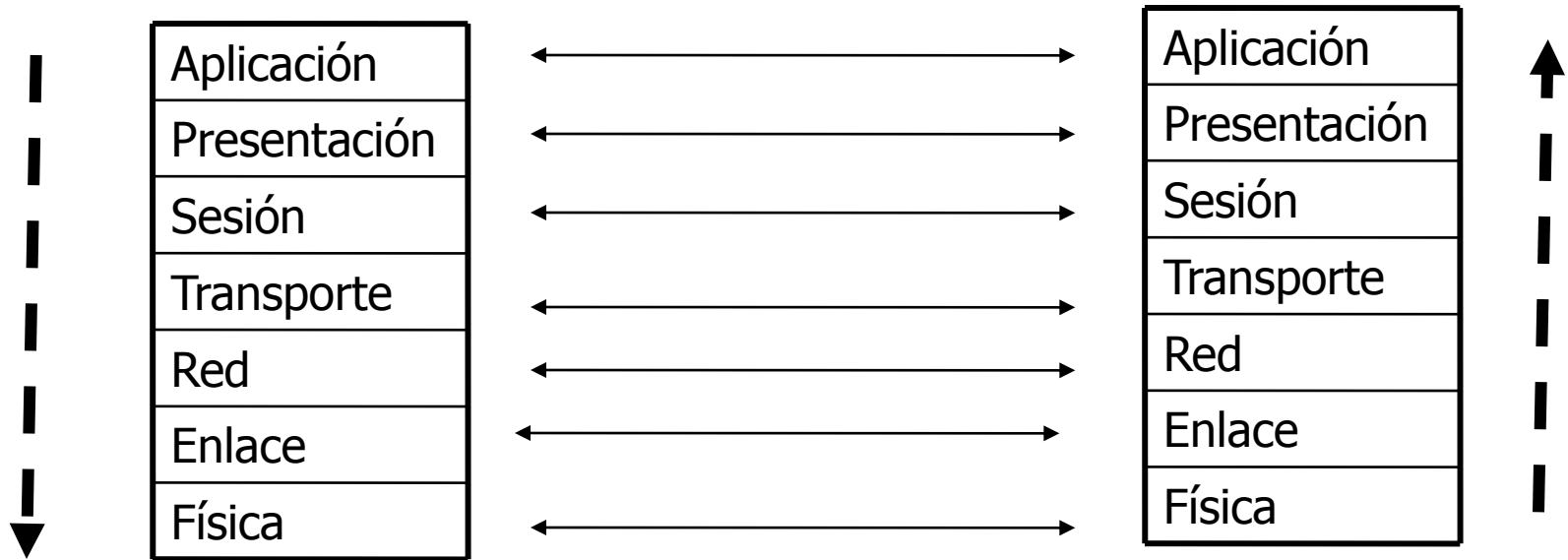
Conceptos previos

- ▶ Servicios orientados a conexión
 - Proveen garantía de que el paquete de dato se reciba, o re-envie
 - Se pueden reservar recursos
 - Necesitan interacción entre origen y destino
 - Implican un inicio y cierre de sesión
- ▶ Servicios no orientados a conexión
 - Sin garantías de recibo
 - No hay inicio o cierre de sesión, se aumenta eficiencia
- ▶ Tipos de redes en cuanto a control de canal
 - Conmutación de circuitos (red telefónica)
 - Conmutación de paquetes (Internet)

Tipos de envío

- ▶ Unicast
 - Uno a uno
 - ▶ Broadcast
 - Uno a todos
 - ▶ Multicast
 - Uno a varios
 - ▶ Anycast
 - Uno a alguno
- 

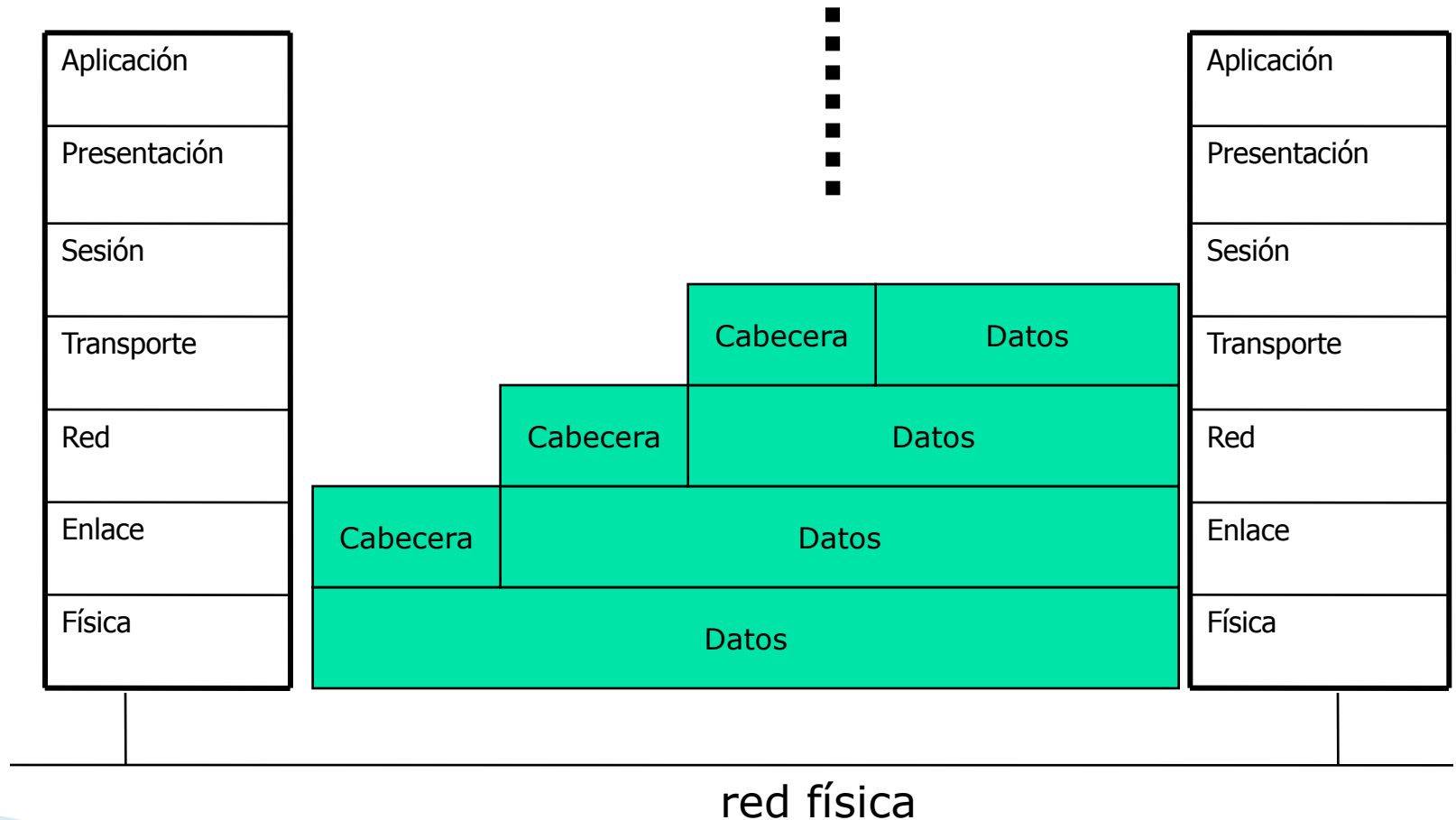
Modelo de Referencia OSI



- ▶ También conocido como “Modelo de Capas”
 - Cada capa provee servicios a la capa inmediata superior
 - Cada capa es cliente de la capa inmediata inferior
 - Conversación horizontal: cada capa dialoga con su homóloga remota
 - Flujo de datos vertical hacia/desde capa física
 - Un protocolo es la implementación de la lógica de una capa
 - Se pueden especificar uno o más protocolos por capa

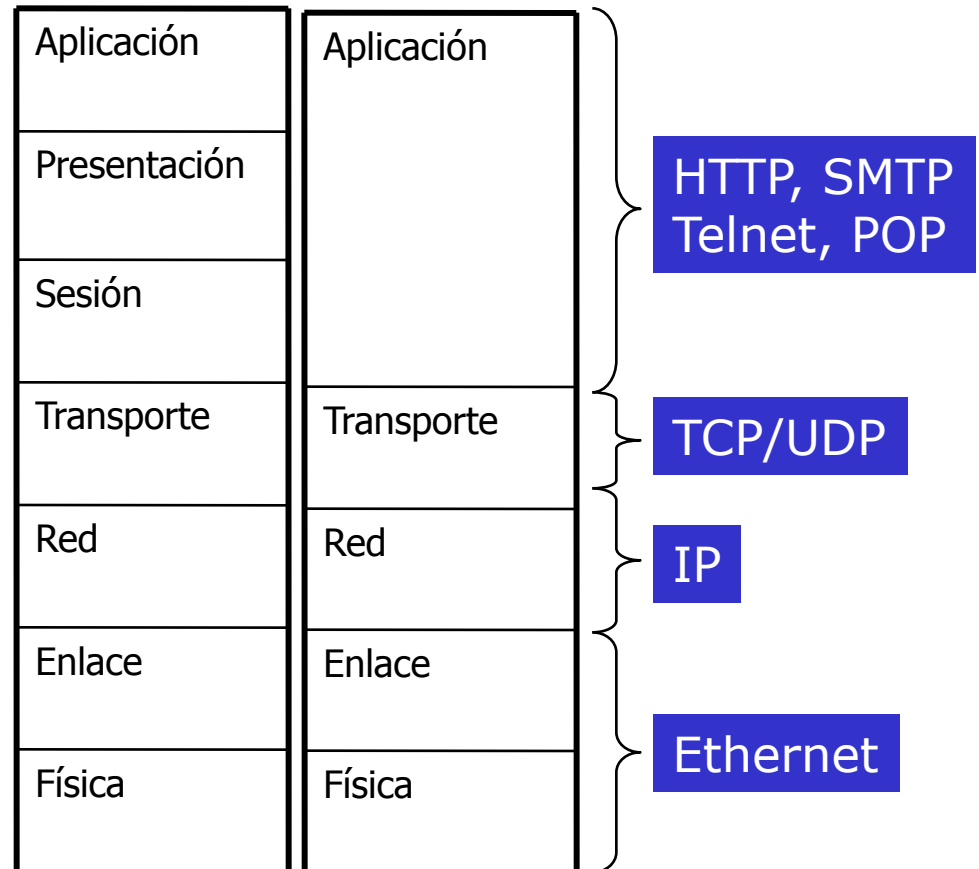
Modelo de Capas

► Encapsulación y cabeceras



OSI vs. TCP/IP

- ▶ ARPANET empezó una década antes que OSI
- ▶ Simple, no necesidad de capa de presentación o sesión
- ▶ Se convirtió en estándar 'de facto'



Capa 1: Física

- Implementada en hardware
- Codificación de canal
 - Representación de bits, voltajes, frecuencias, sincronización
 - Códigos Manchester, AMI, B8ZS...
- Define conectores físicos, distancias, cableado

Capa 2: Enlace

- ▶ Encapsula los los paquetes en tramas para pasarlos al medio físico
- ▶ Reconstruye las tramas originales a partir de secuencias de bits y pasa los datos a la capa de red
- ▶ Provee
 - Direcccionamiento (en el segmento de red local)
 - Detección de errores
 - Control de flujo

Capa 3: Red

- Oculta los detalles de la red física, direccionamiento global:
 - Una dirección IP es única en toda la red
 - Implica que hay que mapear las direcciones físicas con las IP
- Ofrece un servicio sin garantías (mejor esfuerzo)
 - No se ocupa de pérdida o duplicación de paquetes, confía esa función a las capas superiores
- Determina si el destino es local, o a través de un enrutador
- Provee funciones de control via ICMP
- Paquetes navegan de “salto en salto”, el trayecto completo puede constar de muchos

Capa 4: Transporte

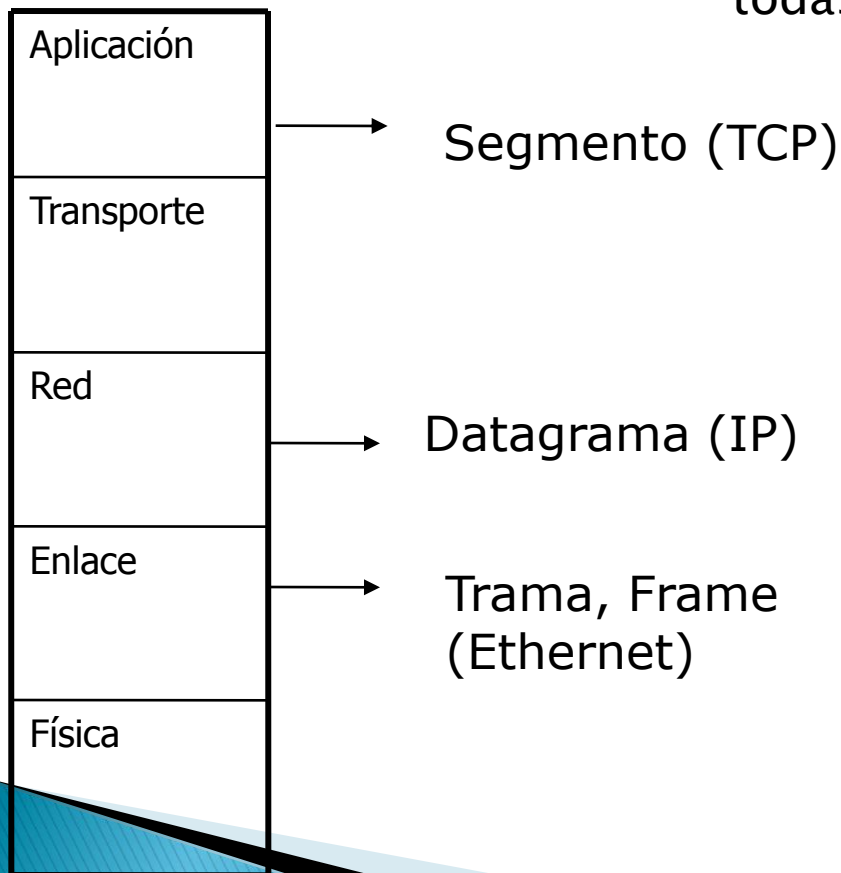
- ▶ Servicio con garantías (TCP)
 - Resuelve los problemas de:
 - Pérdida de paquetes
 - Duplicación
 - Desbordamiento (control de flujo)
- ▶ Servicio sin garantías (UDP)
 - Mucho más simple
 - A veces no hace falta fiabilidad
- ▶ Provee multiplexión de aplicaciones
 - Concepto de 'puertos'

Capa 5: Aplicación

- ▶ La más cercana al usuario
 - Define las funciones de clientes y servidores
- ▶ Utiliza los servicios de transporte
- ▶ Ej: HTTP (web), SMTP (mail), Telnet, FTP, DNS...

Terminología

- ▶ Nombres diferentes en cada capa
- ▶ No se sigue muy estrictamente. Suele hablarse indistintamente de 'paquete' en todas las capas.



Tipos de enlaces

- ▶ Difusión (broadcast)
 - Ej: Ethernet
- ▶ Punto a punto
 - Ej. PPP, SLIP, HDLC
- ▶ NBMA (Non-broadcast Multi-Access)
 - Ej: Frame Relay, ATM

Un vistazo a Ethernet

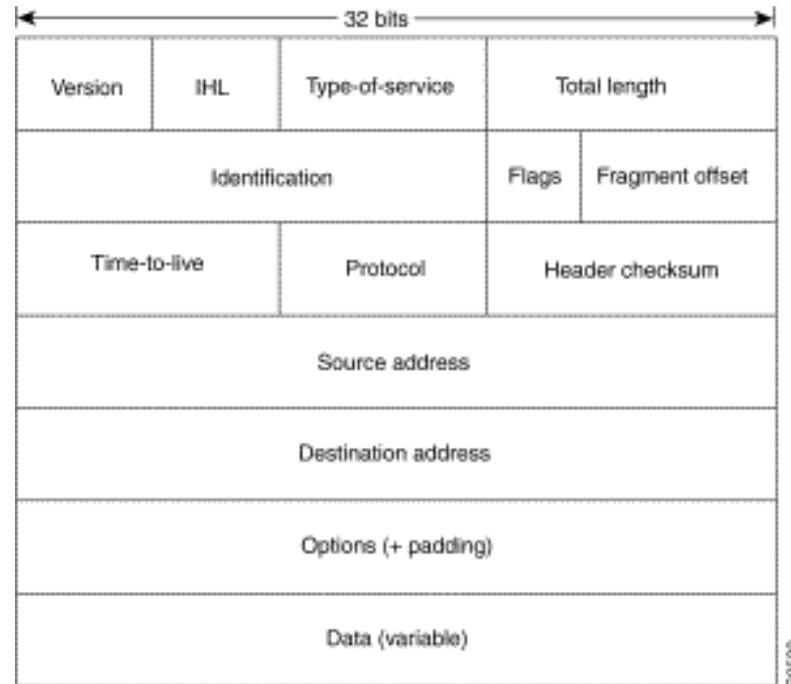
- ▶ Una red de difusión (broadcast)
 - Topologías
 - Bus (cable coaxial)
 - Estrella con repetidor
 - Estrella con conmutador
- ▶ ¿CSMA/CD?
- ▶ Razones para su éxito
 - Simplicidad
 - Costo
- ▶ De 10 Mbps a 10 Gbps

Un vistazo a Ethernet

Preámbulo (8 bytes)	Destino (6)	Fuente (6)	Longitud (2)	Tipo (2)	Datos (46-1500)	FCS (4)
------------------------	----------------	---------------	-----------------	-------------	--------------------	------------

- ▶ Direcciones MAC:
 - Únicas y grabadas en el hardware de la tarjeta
 - Por eso también se llaman “direcciones físicas”
 - 6 bytes x 8 bits/byte = 48 bits
 - Suelen escribirse en hexadecimal
 - FE:D2:89:C4:4F:2E
- ▶ Tipo: 0x800 especifica que la parte de datos contiene un datagrama IP

El datagrama IPv4



- ▶ El protocolo se refiere al que está siendo encapsulado (tcp, udp...)
- ▶ TTL se decrementa con cada salto
- ▶ Hay fragmentación al pasar de un MTU mayor a uno menor

La dirección IPv4

- ▶ Un número de 32 bits (4 bytes)
 - Se puede representar de varias formas:

- Decimal:

128	223	254	10
-----	-----	-----	----

- Binaria:

10000000	11011111	11111110	00001010
----------	----------	----------	----------

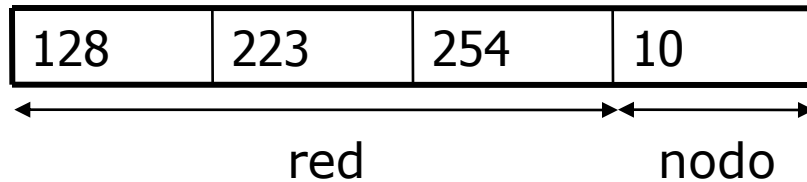
- Hexadecimal:

80	DF	FE	0A
----	----	----	----

La dirección IPv4

► Estructura

- Un sólo número, dos informaciones:
 - Dirección de la subred (prefijo)
 - Dirección del nodo dentro de esa red



► ¿Dónde está la división?

- Al principio era implícito (clases)
- Luego más flexible (máscaras) → CIDR

Estructura de la direccion IPv4

- ▶ Parte de Red (Prefijo)
 - Describe la subred
- ▶ Parte de nodo
 - Describe un nodo en la subred

205	.	154	.	8		1
11001101		10011010		00001000		00000001
Network					Mask	Dispositivo

- ▶ La frontera de la mascara puede ser en cualquier posicion (modelo sin clase o CIDR)
- ▶ Una notacion ya obsoleta (de clase) hacia coincidir la frontera en el bit 8 (clase A), bit 16 (clase B), y bit 24 (clase C).

Notación de Mascara

205	.	154	.	8		1
11001101		10011010		00001000		00000001
Network					Mask	Dispositivo

- ▶ Se puede especificar como la cantidad de bits a 1:
 - Dirección 205.154.8.1 con máscara 255.255.255.0
 - (o sea, el prefijo de red tiene 24 bits)
- ▶ O se agrega a la dirección IP con un símbolo “/”
 - 205.154.8.1/24
- ▶ Hoy día se utilizan indistintamente las dos notaciones

La mascara de subred

- ▶ La mascara de subred es util para definir el tamaño de la red
- ▶ Una mascara 255.255.255.0 o /24 implica
 - $32-24=8$ bits para direccion de nodos
 - $2^8 - 2 = 254$ posibles nodos
- ▶ Una mascara 255.255.255.224 o /27 implica
 - $32-27=5$ bits para direccion de nodos
 - $2^5 - 2 = 30$ posibles nodos

Direcciones especiales

- ▶ Todos los bits de nodo a 0: Representa la red
 - 128.223.254.0/24
- ▶ Todos los bits a 1: Broadcast local
 - 255.255.255.255
- ▶ Todos los bits de nodo a 1: Broadcast dirigido
 - 128.223.254.255
- ▶ Direcciones Loopback:
 - 127.0.0.0/8
 - Casi exclusivamente se usa 127.0.0.1

Más direcciones Especiales

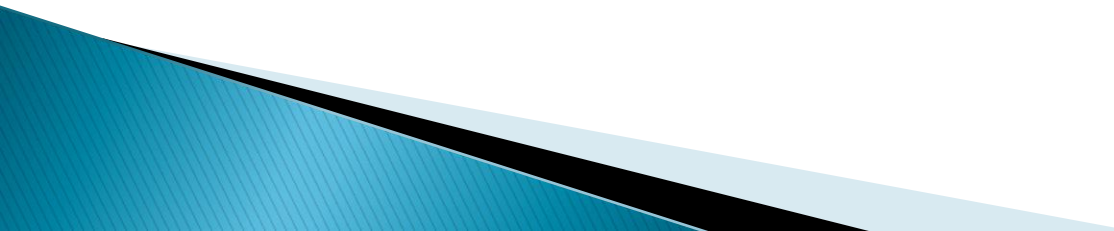
- ▶ Direcciones privadas (RFC 1918)
 - 10.0.0.0 – 10.255.255.255 (10/8)
 - 172.16.0.0 – 172.31.255.255 (172.16/12)
 - 192.168.0.0 – 192.168.255.255 (192.168/16)
- ▶ ¿Cuál es la necesidad?

Datagrama IP

0	4	8	16	19	31
Version	IHL	Type of Service	Total Length		
Identification			Flags	Fragment Offset	
Time To Live		Protocol	Header Checksum		
Source IP Address					
Destination IP Address					
Options					Padding

- Tipo de Servicio (TOS)
 - retardo, fiabilidad, velocidad (voz vs. datos)
- Manejo de fragmentacion
 - Identificacion, Banderas, Desplazamiento
- Tiempo de Vida (TTL)

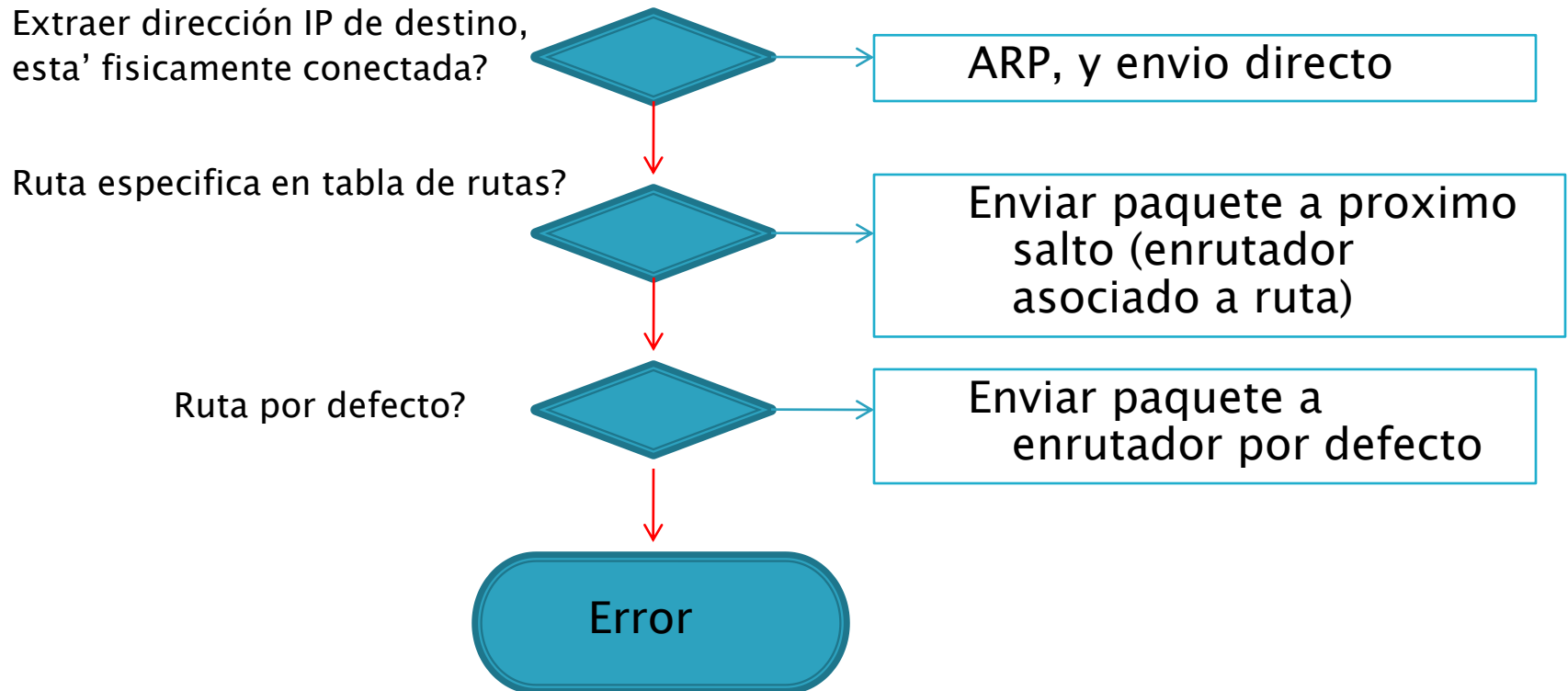
enrutadors

- ▶ Dispositivos con interfaces en varias redes físicas
 - ▶ Una dirección IP (y subred) por cada interfaz
 - ▶ Deciden el trayecto de los paquetes basados en tablas de rutas
- 

Envío

- ▶ En IP, distinguimos entre:
 - Envío directo:
 - La máquina envía a otra que está en su propia red física (Ej: mismo segmento Ethernet)
 - Envío indirecto:
 - El destino del paquete IP está fuera de la red física
 - Requiere la presencia de un enrutador

Algoritmo de Enrutamiento



Enrutadores

- ▶ Flujo de trabajo
 - Recibe un paquete en una interfaz
 - Determina si el paquete está dirigido a él
 - Decrementa el TTL
 - Compara la dirección destino con la tabla de rutas
 - Envía el paquete al proximo enrutador (o nodo de destino), o declara error

Enrutamiento

- ▶ Cada decisión es un salto en la dirección al destino
 - El enrutador A puede enviar paquetes a otro enrutador B solo si ambos tienen al menos una interfaz conectada a la misma subred física
- ▶ Cada enrutador tiene sus propias tablas de rutas
- ▶ Protocolos de enrutamiento: mantener tablas actualizadas

Tablas de Rutas

- ▶ Se compara la dirección IP destino del paquete con las entradas en la tabla
- ▶ Determinar el próximo salto
 - Se asume que está físicamente conectado
- ▶ Regla de “mas grande coincidencia”
 - (*longest match*)

IP	Máscara	Gateway
192.168.5.0	255.255.255.0	192.168.1.1
192.168.5.0	255.255.255.192	192.168.1.2
0.0.0.0	0.0.0.0	192.168.0.1

Tabla de Rutas

```
# netstat -nr
```

```
Kernel IP routing table
```

Destination	Gateway	Genmask	Flags	MSS	Window	irtt	Iface
128.223.60.0	0.0.0.0	255.255.254.0	U	0	0	0	eth0
127.0.0.0	0.0.0.0	255.0.0.0	U	0	0	0	lo
0.0.0.0	128.223.60.1	0.0.0.0	UG	0	0	0	eth0

```
Router# show ip route
```

Codes: I - IGRP derived, R - RIP derived, O - OSPF derived,
C - connected, S - static, E - EGP derived, B - BGP derived,
* - candidate default route, IA - OSPF inter area route,
i - IS-IS derived, ia - IS-IS, U - per-user static route,
o - on-demand routing, M - mobile, P - periodic downloaded static route,
D - EIGRP, EX - EIGRP external, E1 - OSPF external type 1 route,
E2 - OSPF external type 2 route, N1 - OSPF NSSA external type 1 route,
N2 - OSPF NSSA external type 2 route

Gateway of last resort is 10.119.254.240 to network 10.140.0.0

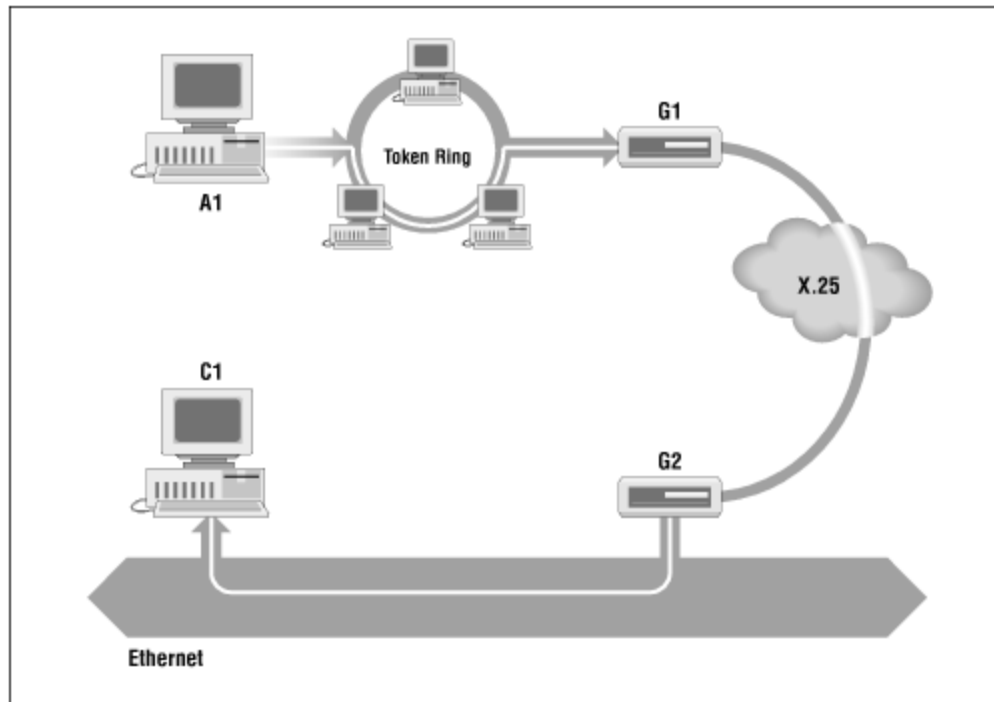
```
O E2 10.110.0.0 [160/5] via 10.119.254.6, 0:01:00, Ethernet2
E 10.67.10.0 [200/128] via 10.119.254.244, 0:02:22, Ethernet2
O E2 10.68.132.0 [160/5] via 10.119.254.6, 0:00:59, Ethernet2
O E2 10.130.0.0 [160/5] via 10.119.254.6, 0:00:59, Ethernet2
E 10.128.0.0 [200/128] via 10.119.254.244, 0:02:22, Ethernet2
E 10.129.0.0 [200/129] via 10.119.254.240, 0:02:22, Ethernet2
E 10.65.129.0 [200/128] via 10.119.254.244, 0:02:22, Ethernet2
E 10.10.0.0 [200/128] via 10.119.254.244, 0:02:22, Ethernet2
```

Encapsulación

1. Recibe trama de capa 2
2. Extrae datagrama IP, y analiza
3. Determina la interfaz de salida
4. Encapsula el datagrama en una trama de
5. tipo adecuado
 - Las redes de entrada y salida pueden ser completamente diferentes:
 - Ejemplos:
 - De Ethernet a PPP
 - De Frame Relay a Ethernet

Fragmentación

- Diferentes MTU en cada salto



ARP: Traducción de direcciones

- ▶ Mantiene tablas dinámicas

```
arp -a
```

```
Interface: 128.223.219.14 --- 0x2
```

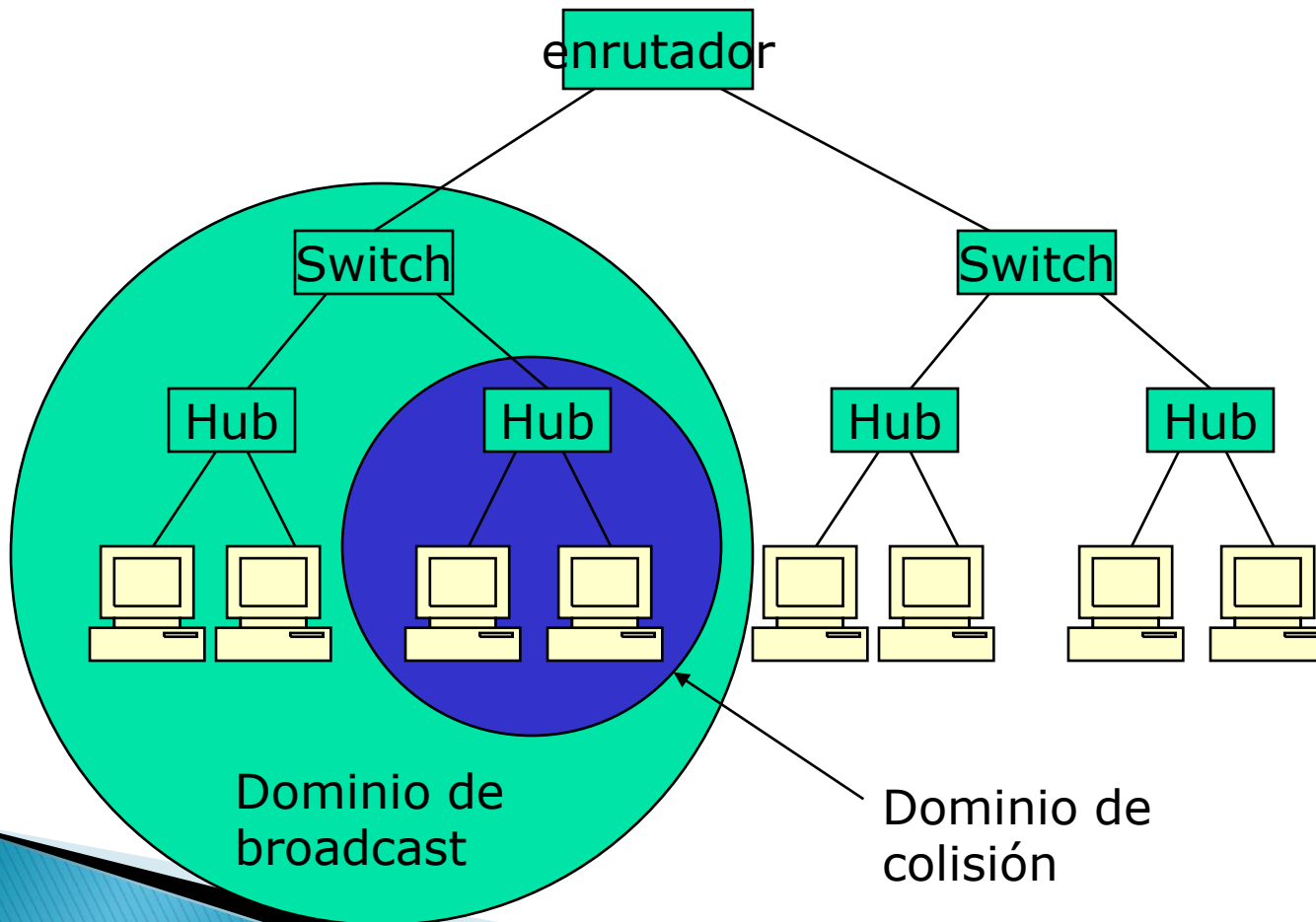
Internet Address	Physical Address	Type
128.223.216.1	00-04-75-71-e5-64	dynamic
128.223.216.24	00-04-23-62-14-4f	dynamic

- ▶ Las entradas tienen un tiempo de vida limitado (¿Por qué?)
- ▶ Mecanismo:
 - A quiere enviar a B, ¿existe la IP de B en la tabla de A?
 - Si no, A pregunta: ¿Quién tiene 192.168.0.1?
 - Como? Envía una trama a toda la red
 - Utiliza *FF:FF:FF:FF:FF:FF* (todos los bits a 1)
 - Todos reciben la trama. Sólo B responde

ARP

- ▶ Algunas mejoras de eficiencia:
 - A quiere saber la MAC de B
 - B recibe la trama. Toma las direcciones MAC e IP de A y las incluye en su tabla
 - Luego B responde a A
 - Como la petición es broadcast, en principio todos los demás pueden incluir a A en su tabla.
- ▶ Pregunta: El paquete ARP viaja dentro de una trama Ethernet o un paquete IP?

Dominios de tráfico



UDP

- ▶ User Datagram Protocol
 - Multiplexión de aplicaciones
 - Una dirección IP identifica una máquina
 - Los sistemas operativos son multitarea
 - Un puerto para cada servicio
- ▶ Servicio no orientado a conexión
 - No ofrece ninguna garantía
 - Sin acuses de recibo
 - Sin re-transmisión
 - Sin control de flujo

UDP

► Formato de UDP

Puerto Origen	Puerto Destino
Longitud	Checksum
Datos	
...	

TCP

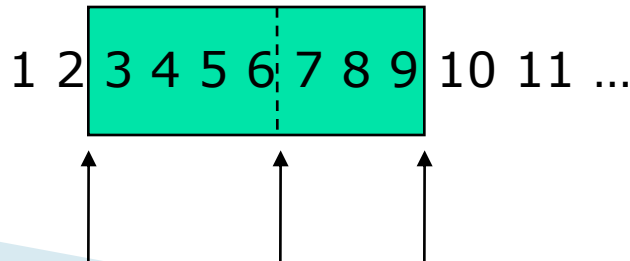
- ▶ Transmission Control Protocol
 - Orientado a conexión
 - Garantiza recibo de paquete
 - Previo acuerdo entre origen y destino
 - Control de flujo:
 - Tamaño de ventana se ajusta constantemente

TCP: Conceptos

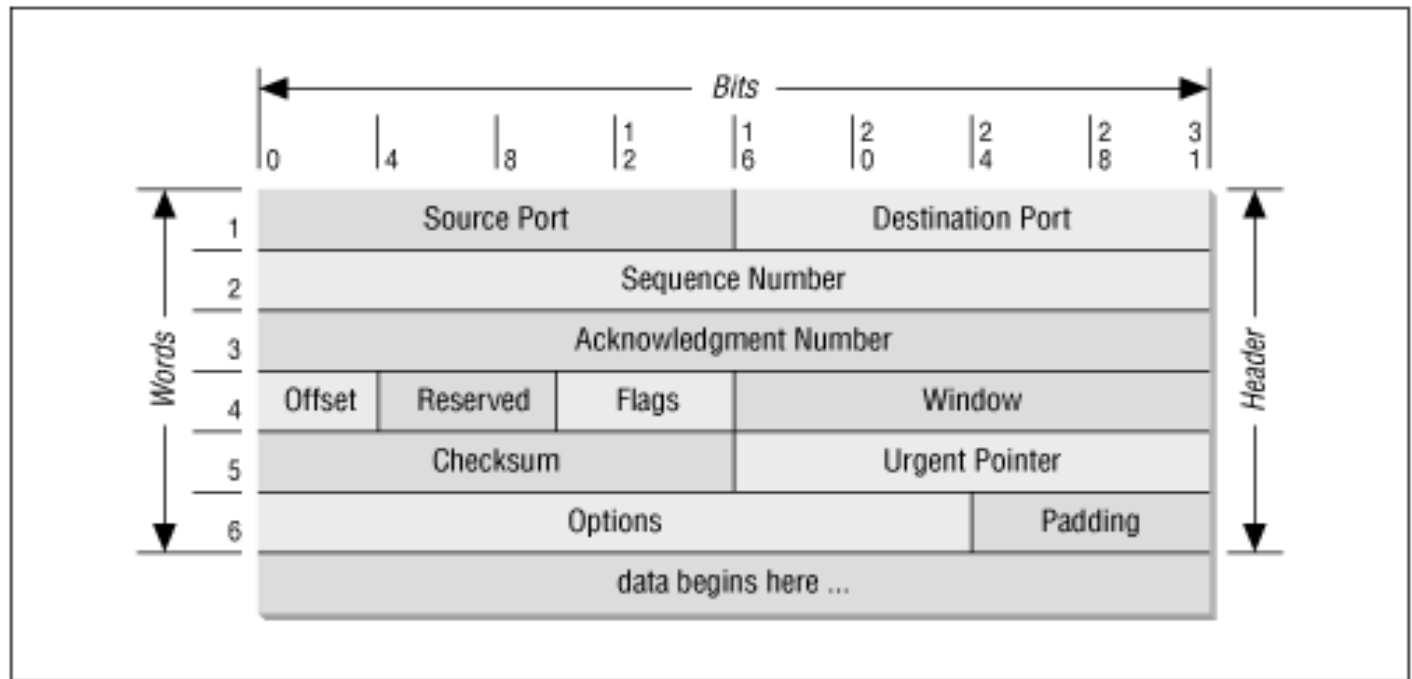
- ▶ Reconocimiento Positivo con Retransmision (PAR en Ingles)
 - Enviar segment, e iniciar conteo regresivo
 - Esperar por confirmación antes de enviar el siguiente segment
 - Re-enviar el mismo segmento si el conteo regresivo expira y no se ha recibido confirmacion
- ▶ ¿Segmentos duplicados? ¿Cómo?
 - Un retraso en la red produce retransmisión, mismo segmento llega dos veces

TCP: Ventana deslizante

- ▶ Esperar confirmación por cada paquete no es eficiente
 - Tamaño de ventana = 1
- ▶ Provee control de la congestión y control de flujo (¿cuál es la diferencia?)
 - El tamaño de la ventana se ajusta dinámicamente



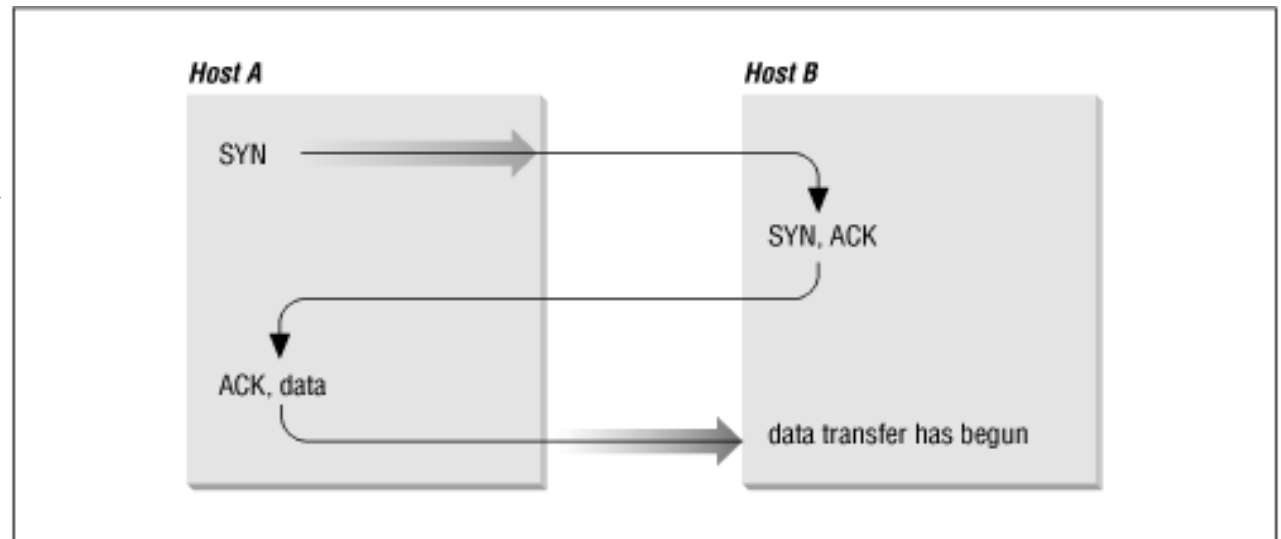
Formato de paquete TCP



TCP: Inicio de Sesión

Acuerdo en tres pasos:

- ▶ SYN
- ▶ SYN ACK
- ▶ ACK



¿TCP o UDP?

- ▶ Cuándo tiene sentido uno u otro?
 - FTP
 - DNS
 - SNMP
 - Voz sobre IP (H.323, SIP)
 - Multicast

ICMP

- ▶ Internet Control Message Protocol
 - Viaja sobre IP
 - pero no pertenece a la capa de transporte
 - Funcion: gestion de redes
 - Notificar errores
 - Control de Flujo
 - Redirección

ICMP

► Algunos tipos y códigos más usados

Tipo	Código	Descripción
0	0	Echo Reply
3	0	Destination Network unreachable
3	1	Destination Host Unreachable
3	2	Destination Protocol Unreachable
3	3	Destination Port Unreachable
8	0	Echo Request
11	0	TTL expired

ICMP: Aplicaciones

► Ping

```
# ping www.uoregon.edu
```

```
PING darkwing.uoregon.edu (128.223.142.13) from 128.223.60.27 : 56(84) bytes of data.  
64 bytes from darkwing.uoregon.edu (128.223.142.13): icmp_seq=1 ttl=254 time=0.229 ms  
64 bytes from darkwing.uoregon.edu (128.223.142.13): icmp_seq=2 ttl=254 time=0.254 ms  
64 bytes from darkwing.uoregon.edu (128.223.142.13): icmp_seq=3 ttl=254 time=0.226 ms  
64 bytes from darkwing.uoregon.edu (128.223.142.13): icmp_seq=4 ttl=254 time=0.232 ms  
64 bytes from darkwing.uoregon.edu (128.223.142.13): icmp_seq=5 ttl=254 time=0.222 ms
```

ICMP: Aplicaciones

► Traceroute

```
# traceroute www.google.com
traceroute: Warning: www.google.com has multiple addresses; using 66.102.9.99
traceroute to www.google.akadns.net (66.102.9.99), 30 hops max, 38 byte packets
 1  ge-4-6.uonet2-gw.uoregon.edu (128.223.60.3)  0.310 ms  0.236 ms  0.193 ms
 2  0.ge-0-0-0.uonet8-gw.uoregon.edu (128.223.2.8)  0.324 ms  0.331 ms  0.294 ms
 3  eugn-car1-gw.nero.net (207.98.66.11)  0.363 ms  0.296 ms  0.416 ms
 4  eugn-core2-gw.nero.net (207.98.64.169)  0.672 ms  1.029 ms  0.601 ms
 5  ptck-core2-gw.nero.net (207.98.64.2)  2.911 ms  2.994 ms  2.930 ms
 6  ptck-core1-gw.nero.net (207.98.64.137)  3.255 ms  2.874 ms  2.923 ms
 7  so-6-1.hsa2.Seattle1.Level3.net (63.211.200.245)  6.521 ms  6.153 ms  6.322 ms
 8  ge-6-1-1.mp2.Seattle1.Level3.net (209.247.9.85)  6.619 ms  6.565 ms  6.335 ms
 9  so-0-0-0.bbr2.NewYork1.Level3.net (64.159.0.238)  86.194 ms  86.239 ms  86.580 ms
10  so-2-0-0.mp2.London1.Level3.net (212.187.128.154)  147.899 ms  147.968 ms  149.461 ms
11  so-3-0-0.mp2.Amsterdam1.Level3.net (212.187.128.13)  155.019 ms  155.738 ms  155.406
    ms
12  ge-11-2.ipcolo2.Amsterdam1.Level3.net (213.244.165.116)  157.499 ms  155.627 ms
    155.857 ms
13  212.72.44.66 (212.72.44.66)  156.319 ms  156.168 ms  156.142 ms
```

Traceroute: Funcionamiento

```
# traceroute 128.223.142.13
traceroute to 128.223.142.13 (128.223.142.13), 30 hops max, 38 byte packets
 1  ge-4-6.uonet2-gw.uoregon.edu (128.223.60.3)  0.282 ms  0.206 ms  0.186 ms
 2  darkwing (128.223.142.13)  0.266 ms  0.197 ms  0.209 ms
```

(simultáneamente)

```
# tcpdump -lnv host 128.223.142.13 and icmp
tcpdump: listening on eth0
128.223.60.27.33962 > 128.223.142.13.33435: udp 10 [ttl 1] (id 12001, len 38)
128.223.60.3 > 128.223.60.27: icmp: time exceeded in-transit [tos 0xc0] (ttl 255, id 64235, len 56)
128.223.60.27.33962 > 128.223.142.13.33436: udp 10 [ttl 1] (id 12002, len 38)
128.223.60.3 > 128.223.60.27: icmp: time exceeded in-transit [tos 0xc0] (ttl 255, id 64236, len 56)
128.223.60.27.33962 > 128.223.142.13.33437: udp 10 [ttl 1] (id 12003, len 38)
128.223.60.3 > 128.223.60.27: icmp: time exceeded in-transit [tos 0xc0] (ttl 255, id 64237, len 56)
128.223.60.27.33962 > 128.223.142.13.33438: udp 10 (ttl 2, id 12004, len 38)
128.223.142.13 > 128.223.60.27: icmp: 128.223.142.13 udp port 33438 unreachable (DF) (ttl 254, id 14809, len 66)
128.223.60.27.33962 > 128.223.142.13.33439: udp 10 (ttl 2, id 12005, len 38)
128.223.142.13 > 128.223.60.27: icmp: 128.223.142.13 udp port 33439 unreachable (DF) (ttl 254, id 14810, len 66)
128.223.60.27.33962 > 128.223.142.13.33440: udp 10 (ttl 2, id 12006, len 38)
128.223.142.13 > 128.223.60.27: icmp: 128.223.142.13 udp port 33440 unreachable (DF) (ttl 254, id 14811, len 66)
```

Más información

- ▶ *TCP/IP Illustrated*. Richard Stevens. Addison–Wesley
- ▶ *Internetworking with TCP/IP*. Douglas Comer. Prentice–Hall
- ▶ *Cisco Internetworking Basics*
<http://www.cisco.com/en/US/docs/internetworking/technology/handbook/Intro-to-Internet.html>
- ▶ *TCP/IP Network Administration*. Craig Hunt
O'reilly & Associates.
- ▶ Requests for Comments (RFCs) www.ietf.org