Capítulo 4: Capa Red - II

Este material está basado en:

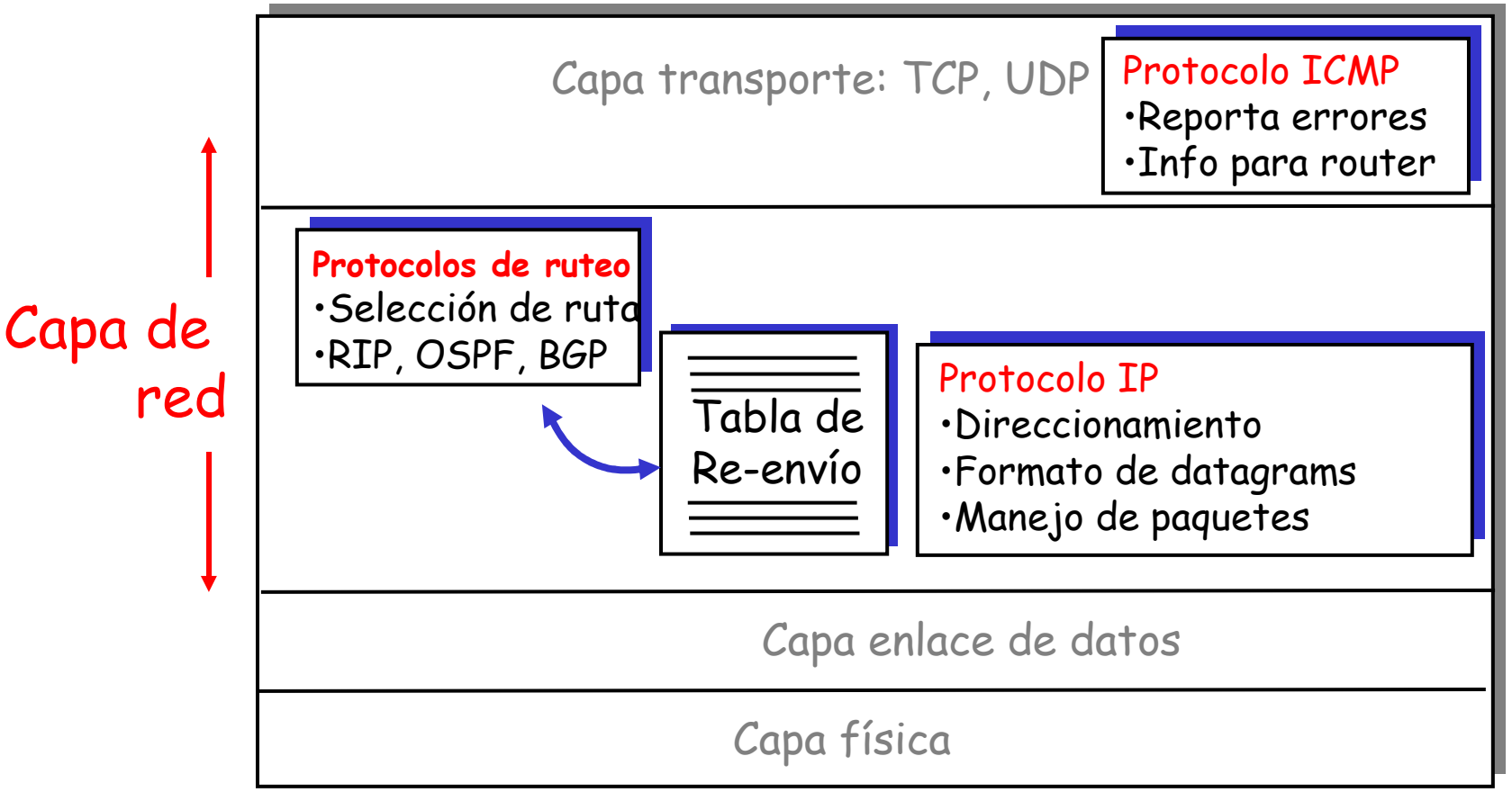
- material de apoyo al texto *Computer Networking: A Top Down Approach Featuring the Internet 3rd* edition. Jim Kurose, Keith Ross Addison-Wesley, 2004.
- material de wikipedia: www.wikipedia.org

Capítulo 4: Capa de Red

- 4. 1 Introducción
- 4.2 Circuitos virtuales y redes de datagramas
- 4.3 ¿Qué hay dentro de un router?
- 4.4 IP: Internet Protocol
  - Formato de Datagrama
  - Direccionamiento IPv4
  - ICMP
  - IPv6
- 4.5 Algoritmo de ruteo
  - Estado de enlace
  - Vector de Distancias
  - Ruteo Jerárquico
- 4.6 Ruteo en la Internet
  - RIP
  - OSPF
  - BGP
- 4.7 Ruteo Broadcast y multicast

Capa de red en Internet

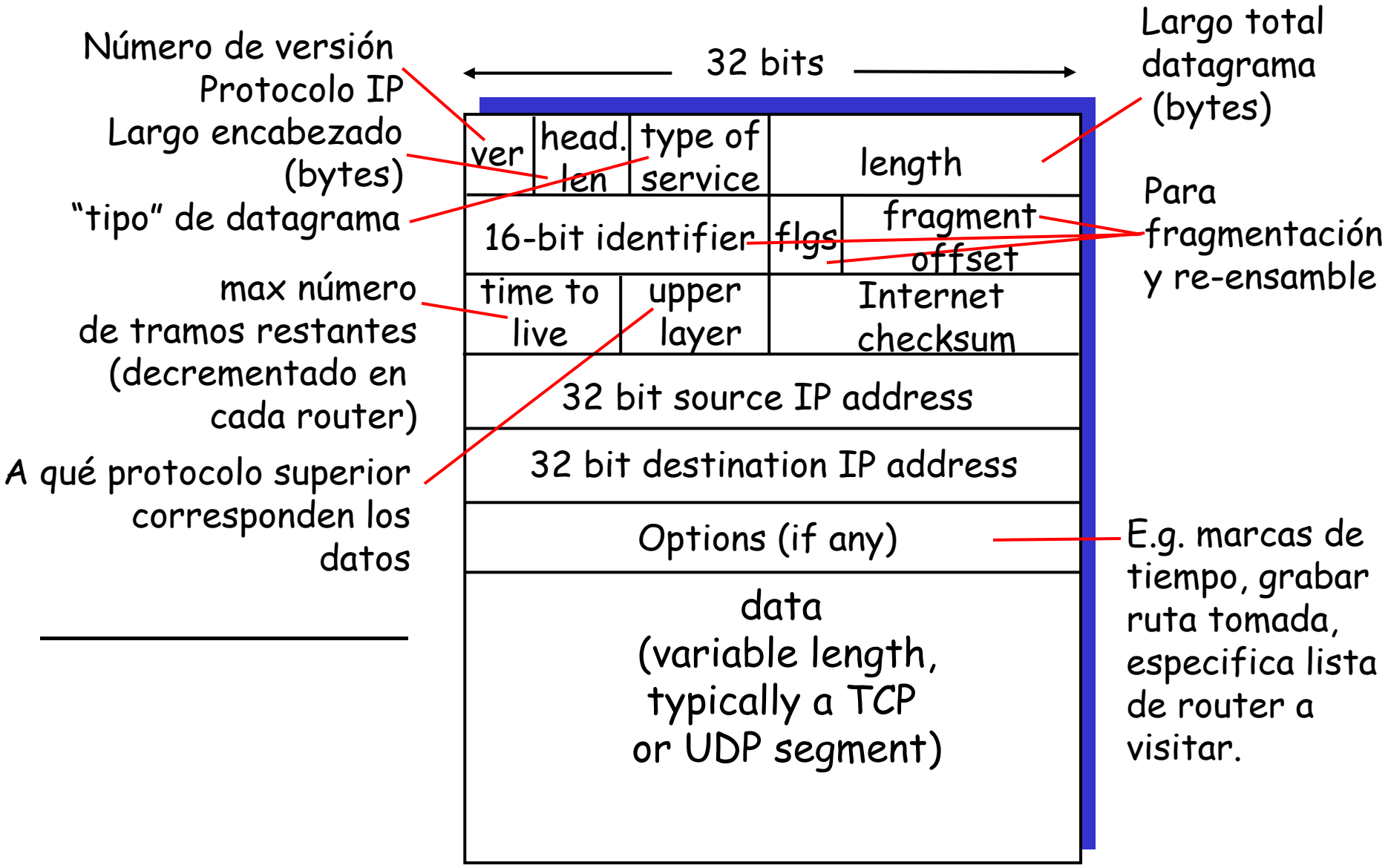
Funciones de la capa de red del host y router :



Capítulo 4: Capa de Red

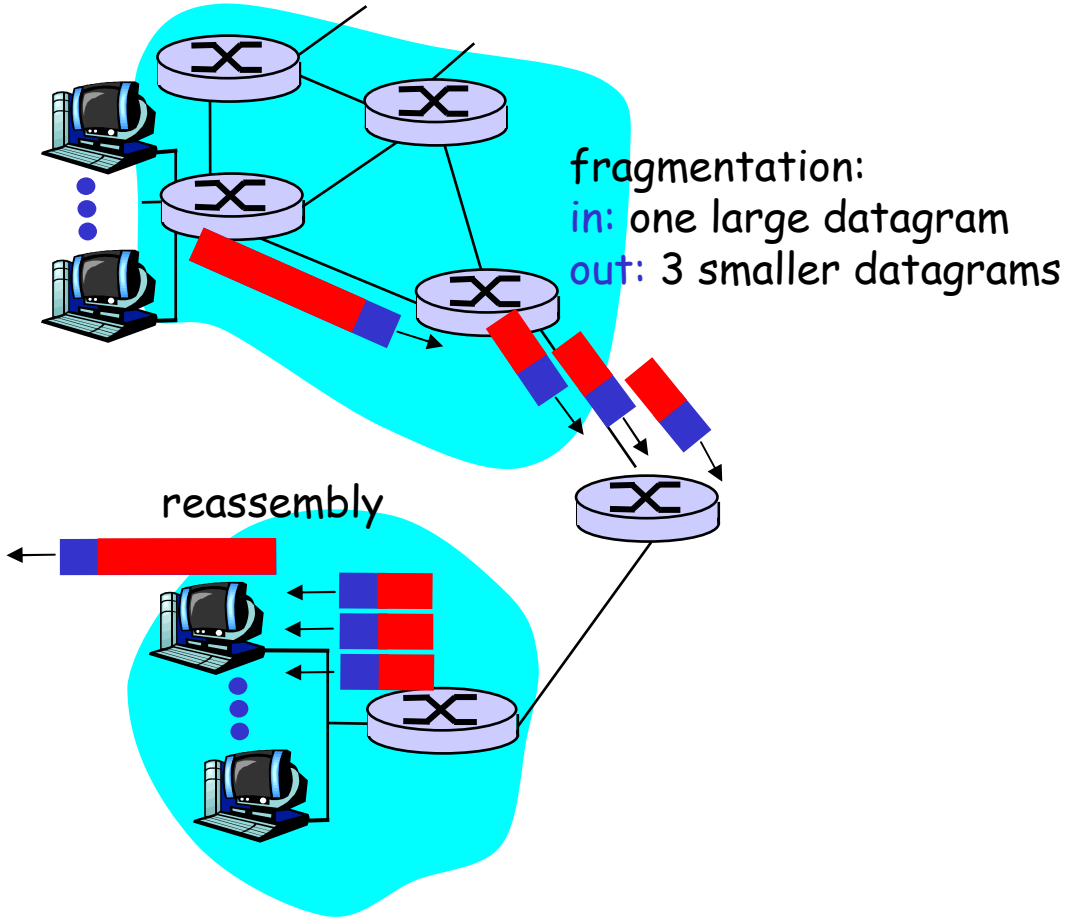
- 4. 1 Introducción
- 4.2 Circuitos virtuales y redes de datagramas
- 4.3 ¿Qué hay dentro de un router?
- 4.4 IP: Internet Protocol
  - Formato de Datagrama
  - Direccionamiento IPv4
  - ICMP
  - IPv6
- 4.5 Algoritmo de ruteo
  - Estado de enlace
  - Vector de Distancias
  - Ruteo Jerárquico
- 4.6 Ruteo en la Internet
  - RIP
  - OSPF
  - BGP
- 4.7 Ruteo Broadcast y multicast

Formato del datagrama IP v.4



Fragmentación y re-ensamble IP

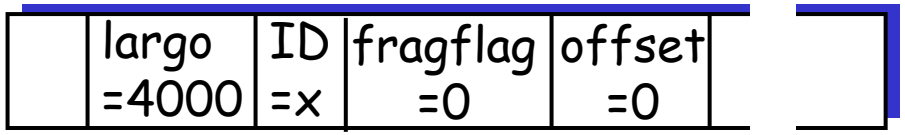
- Enlaces de red tienen MTU (max.transfer size) – mayor tamaño de un frame en la capa enlaze.
  - Diferentes tipos de enlace tienen diferentes MTUs
- Por esto es que un datagrama IP grande es dividido ("fragmented") en la red
  - Un datagrama se convierte en varios datagramas
  - Se "rearma" en el destino final
  - Bits del encabezado IP se usan para identificar y ordenar fragmentos relacionados



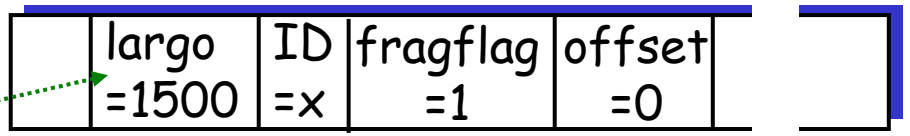
Fragmentación y re-ensamble IP

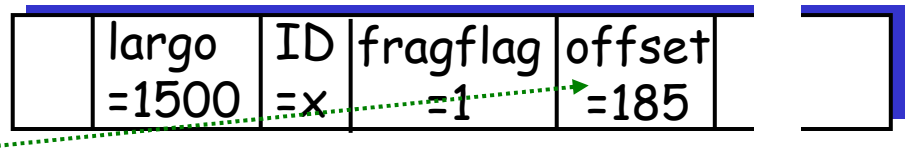
Ejemplo

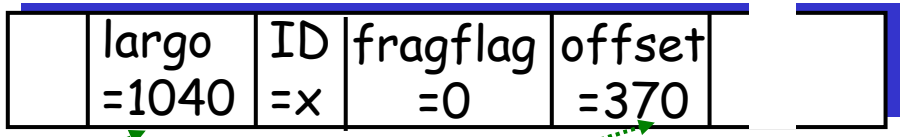
- 4000 byte datagram (20 bytes header IP + 3980 en campo datos datagrama)
- MTU = 1500 bytes



Un datagrama grande es transformado en varios datagramas más pequeños







1480 bytes en campo de datos de datagrama

offset en bloques de 8 bytes
1480/8 = 185

largo ultimo = 3980-1480-1480 = 1020

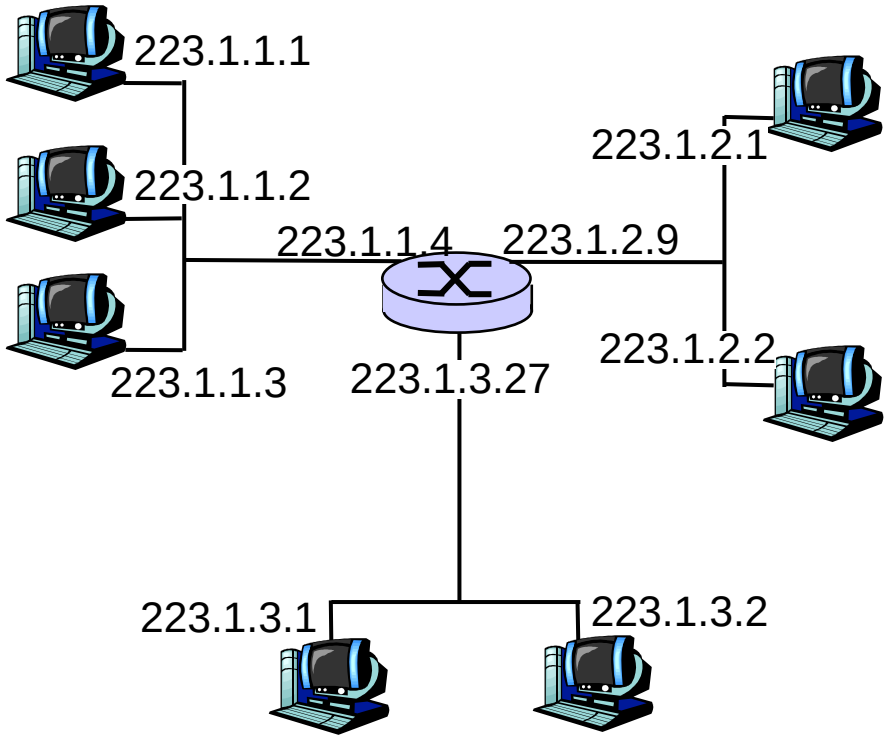
byte insertado en posición 370*8=2960

Capítulo 4: Capa de Red

- 4. 1 Introducción
- 4.2 Circuitos virtuales y redes de datagramas
- 4.3 ¿Qué hay dentro de un router?
- 4.4 IP: Internet Protocol
  - Formato de Datagrama
  - Direccionamiento IPv4
  - ICMP
  - IPv6
- 4.5 Algoritmo de ruteo
  - Estado de enlace
  - Vector de Distancias
  - Ruteo Jerárquico
- 4.6 Ruteo en la Internet
  - RIP
  - OSPF
  - BGP
- 4.7 Ruteo Broadcast y multicast

Direccionamiento IP: introducción

- Dirección IP: identificador de 32-bit del host, *interfaz* del router
- Interfaz: conexión entre host y router, enlace físico
  - Router típicamente tiene múltiples interfaces (bocas)
  - Host puede tener múltiples interfaces
  - Direcciones IP están asociadas a cada interfaz

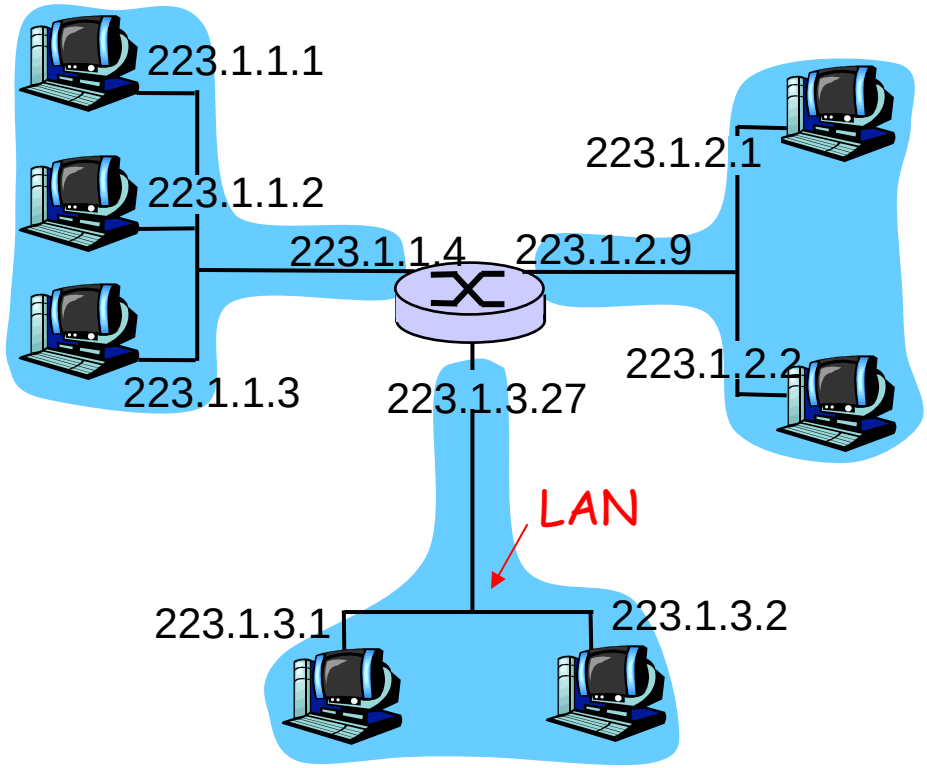


223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

Subnets

- IP address:
  - Parte de subnet (bits alto orden)
  - Parte del host (bits bajo orden)
- *Que es un subnet ?*
  - Red local que utiliza la misma parte de la dirección ip
  - Se podrían interconectar sin tener un router (e.g. con un switch o hub)

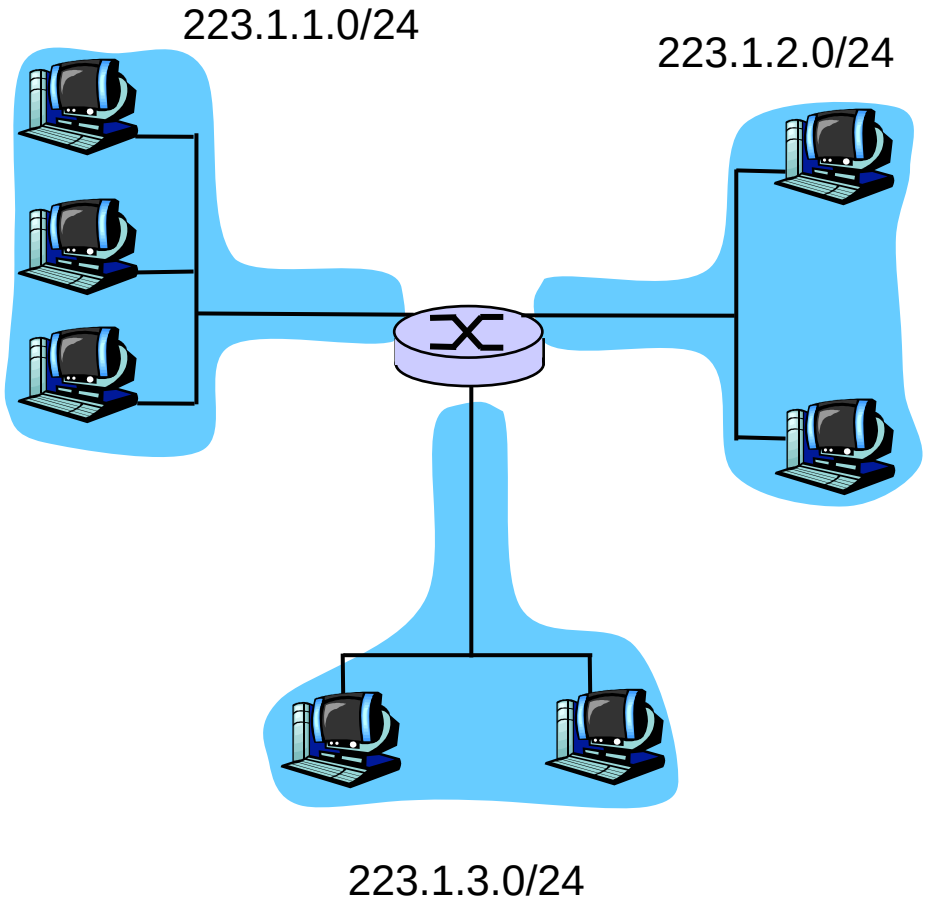


network consisting of 3 subnets

Subnets

Receta

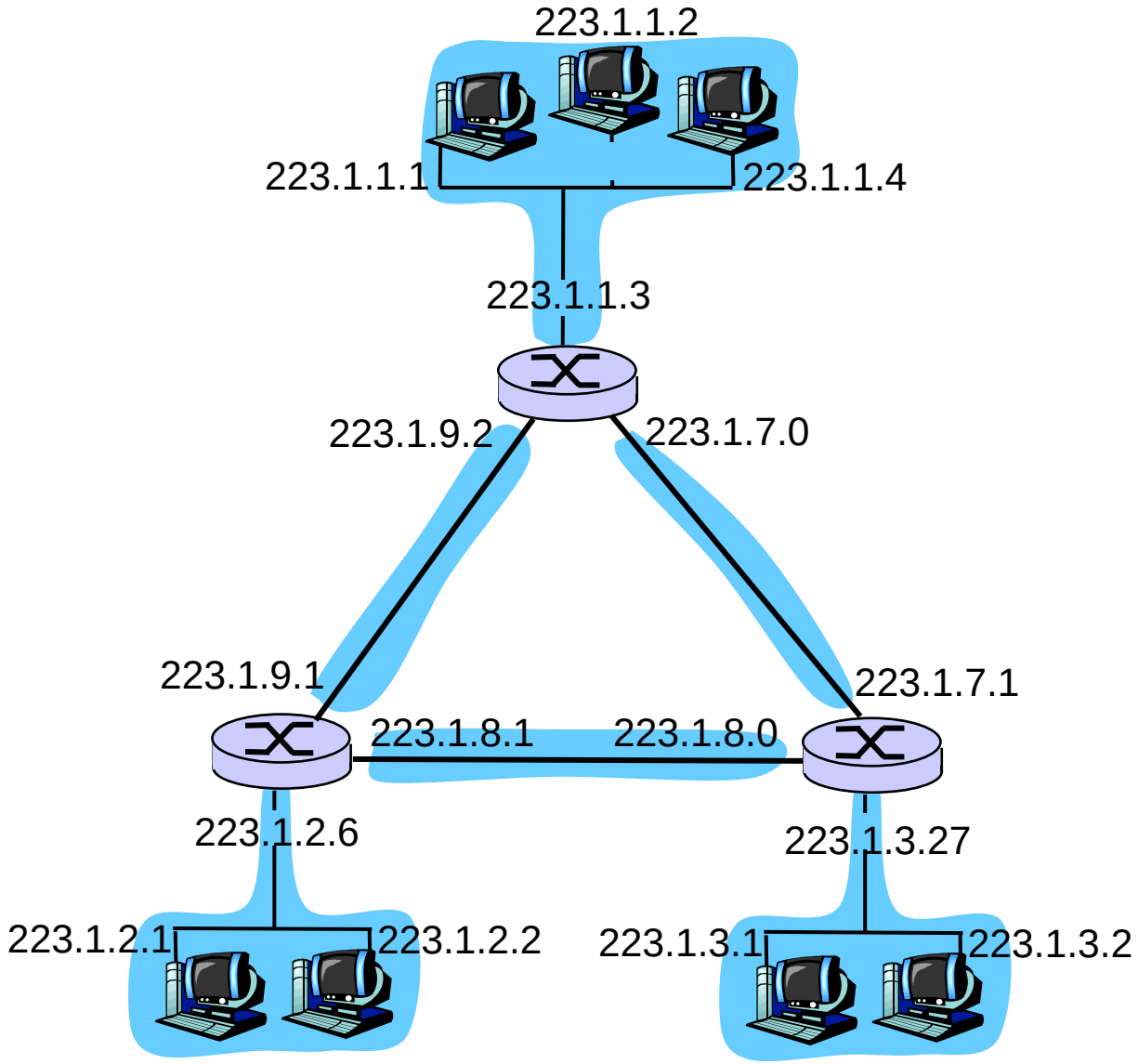
- Para determinar los subnets, desconectar los interfaces del router para crear redes tipo islas independientes.
- Cada red independiente se llama un subnet.



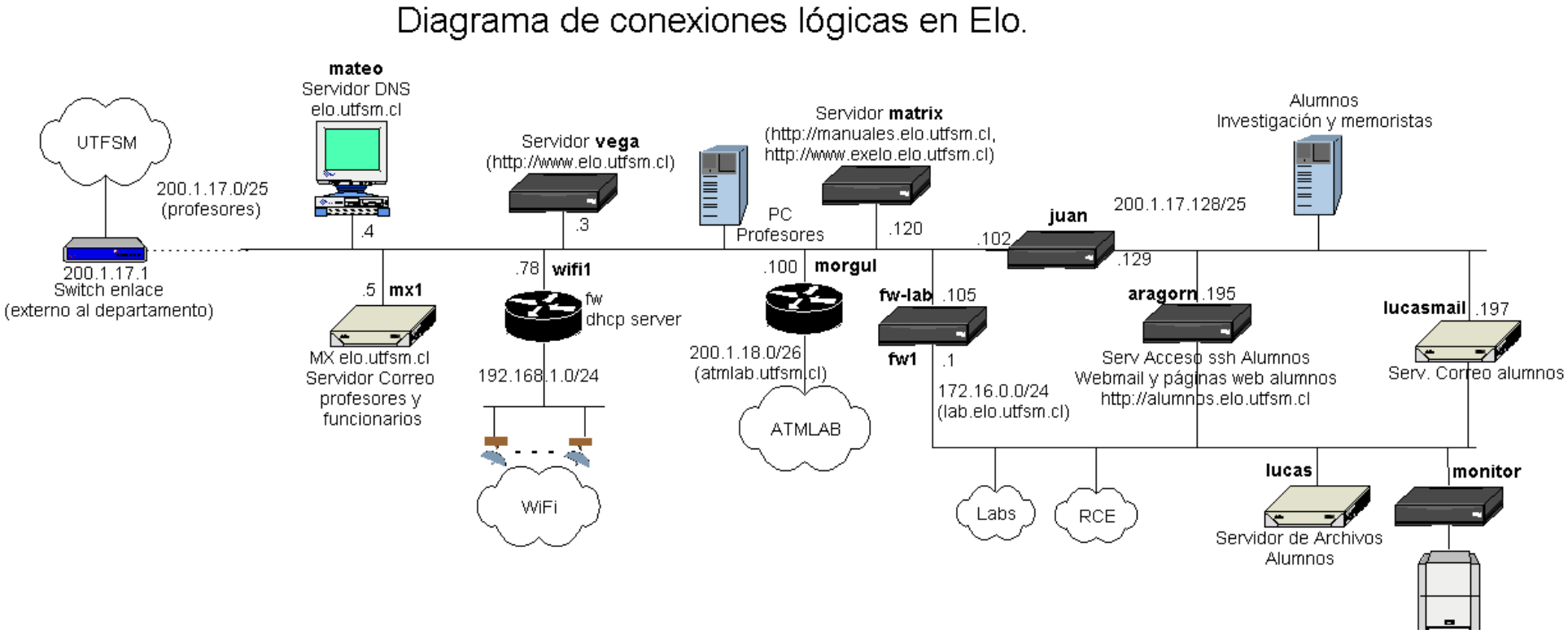
Subnet mask: /24

Subnets

Cuantas hay?

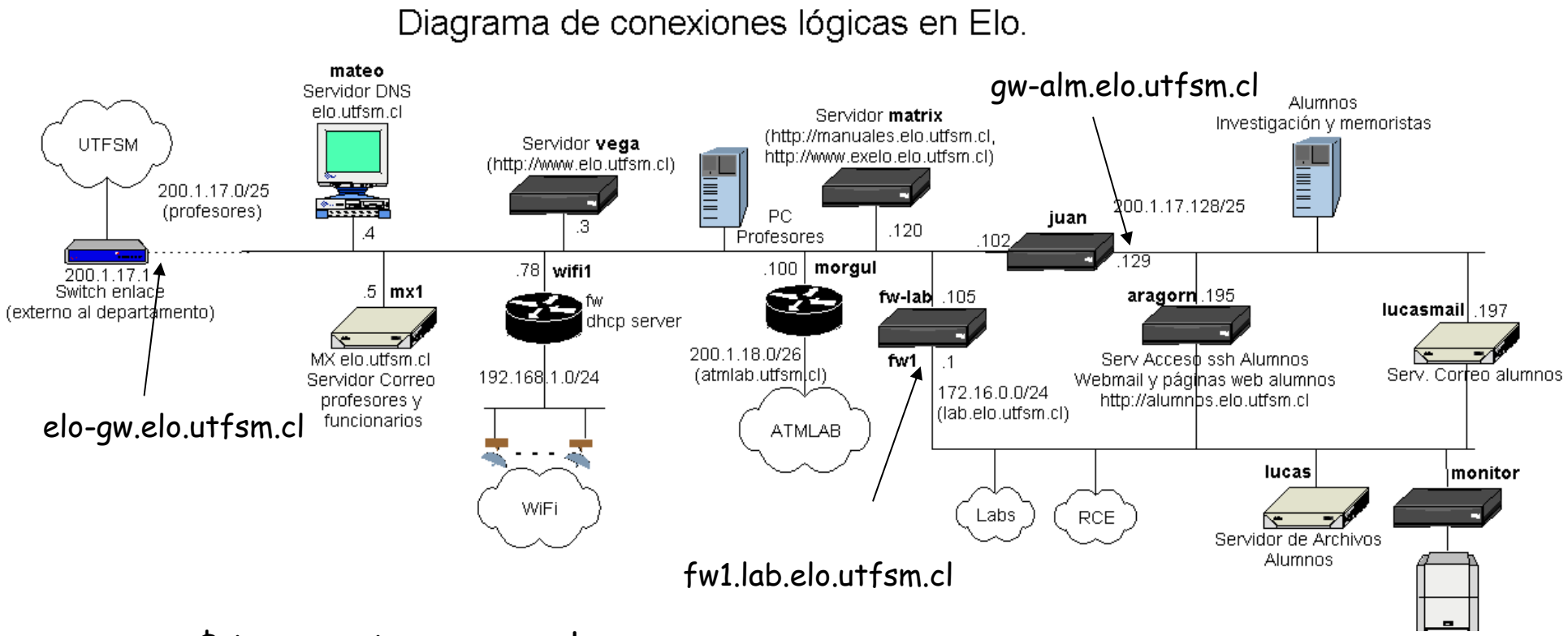


RED ELO (http://www.elo.utfsm.cl/~rce/images/stories/rce/diagrama_red_elo_todo.png)



Nota: Los IPs: 192.168.1.0 y 10.0.0.0 son para redes privadas. Para que estas maquinas sean vistas de afuera se tiene que usar NAT.

Ejemplo: traceroute desde red ELO

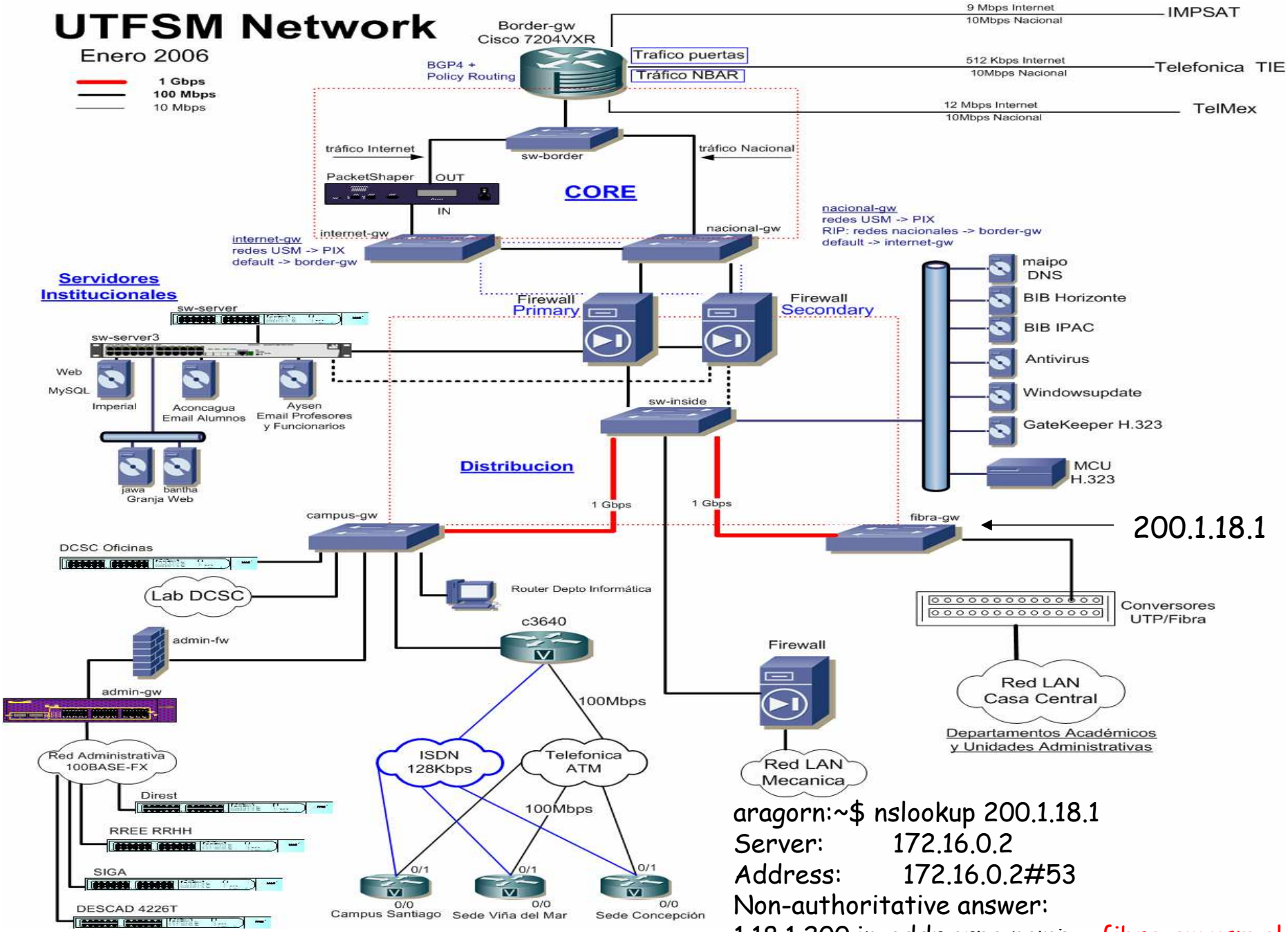


aragorn:~$ traceroute www.google.com
traceroute: Warning: www.google.com has multiple addresses; using 216.239.37.104
traceroute to www.l.google.com (216.239.37.104), 30 hops max, 38 byte packets
1 gw-alm (200.1.17.129) 0.151 ms 0.128 ms 0.130 ms
2 elo-gw (200.1.17.1) 0.668 ms 2.125 ms 0.590 ms
3 * * * (Mensajes ICMP usado por Traceroute bloqueado por firewalls en red USM)
Nota: Esto se sabe porque se puede hacer un traceroute a informática.

Ejemplo: traceroute desde red ELO (cont)

- Como saber adonde se bloquea nslookup?

```
aragorn:~$ traceroute www.inf.utfsm.cl
traceroute to spender.inf.utfsm.cl (200.1.19.14), 30 hops max, 38 byte packets
 1  gw-alm (200.1.17.129)  0.140 ms  0.120 ms  0.120 ms
 2  elo-gw (200.1.17.1)  0.642 ms  0.593 ms  0.584 ms
 3  inti.inf.utfsm.cl (200.1.21.155)  0.526 ms  0.475 ms  0.353 ms
 4  spender.inf.utfsm.cl (200.1.19.14)  0.422 ms  0.401 ms  0.400 ms
```




```
aragorn:~$ nslookup 200.1.18.1
Server:         172.16.0.2
Address:        172.16.0.2#53
Non-authoritative answer:
1.18.1.200.in-addr.arpa name = fibra-gw.usm.cl.
```

Ejemplo traceroute usando UTFSM traceroute

(http://www.dcsc.utfsm.cl/cgi/nph-traceroute.sh)

traceroute: Warning: www.google.com has multiple addresses; using 64.233.187.99
traceroute to www.l.google.com (64.233.187.99), 30 hops max, 40 byte packets
1 nacional-gw.usm.cl (200.1.20.195) 1.046 ms 0.742 ms 0.693 ms
2 border-gw.usm.cl (200.1.21.165) 1.566 ms 1.509 ms 1.733 ms
3 impsat-internet.usm.cl (192.168.171.5) 7.110 ms 5.294 ms 8.631 ms
4 192.168.171.2 (192.168.171.2) 17.520 ms 9.270 ms 9.623 ms
5 192.168.171.2 (192.168.171.2) 7.194 ms 6.954 ms 16.963 ms
6 64.76.144.65 (64.76.144.65) 14.856 ms 14.380 ms 9.180 ms
7 Impsat-Chile-NY-OUT.us.impsat.net (200.31.0.177) 134.514 ms 168.204 ms 155.740 ms
8 65.115.0.53 (65.115.0.53) 157.419 ms 128.662 ms 125.006 ms
9 205.171.230.18 (205.171.230.18) 124.855 ms 124.970 ms 127.340 ms
10 205.171.209.114 (205.171.209.114) 134.878 ms 141.798 ms 125.300 ms
11 205.171.251.22 (205.171.251.22) 144.985 ms 146.945 ms 144.116 ms
12 72.165.86.2 (72.165.86.2) 136.801 ms 136.018 ms 134.264 ms
13 216.239.49.248 (216.239.49.248) 130.080 ms 216.239.49.36 (216.239.49.36) 129.965 ms 216.239.47.120 (216.239.47.120) 152.721 ms
14 72.14.238.97 (72.14.238.97) 194.282 ms 168.627 ms 150.937 ms
15 72.14.236.175 (72.14.236.175) 158.440 ms 72.14.232.97 (72.14.232.97) 160.260 ms 66.249.95.149 (66.249.95.149) 158.018 ms
16 216.239.49.226 (216.239.49.226) 177.017 ms 66.249.94.244 (66.249.94.244) 179.668 ms 216.239.49.226 (216.239.49.226) 164.739 ms
17 72.14.236.175 (72.14.236.175) 161.769 ms 72.14.236.201 (72.14.236.201) 153.927 ms 64.233.187.99 (64.233.187.99) 162.508

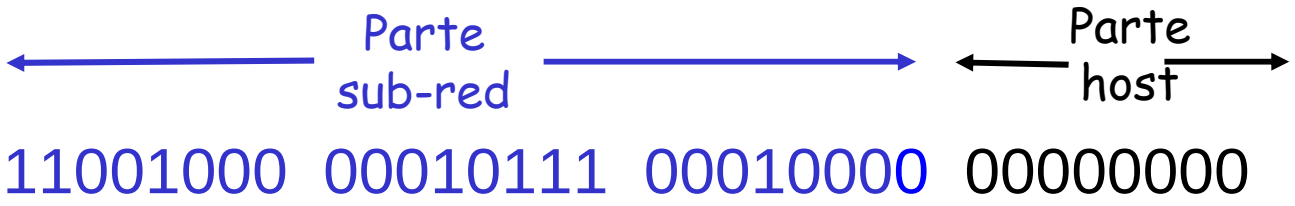
Ejemplo: Servidores DNS usados en ejemplo previo

```
aragorn:~$ nslookup
> 200.1.17.102
Server:         172.16.0.2
Address:        172.16.0.2#53

Non-authoritative answer:
102.17.1.200.in-addr.arpa       name = juan.elo.utfsm.cl.

Authoritative answers can be found from:
17.1.200.in-addr.arpa   nameserver = inti.inf.utfsm.cl.
17.1.200.in-addr.arpa   nameserver = mateo.elo.utfsm.cl.
17.1.200.in-addr.arpa   nameserver = ns.usm.cl.
17.1.200.in-addr.arpa   nameserver = ns2.usm.cl.
ns.usm.cl       internet address = 200.1.21.80
ns2.usm.cl      internet address = 200.1.21.150
inti.inf.utfsm.cl       internet address = 200.1.21.155
inti.inf.utfsm.cl       internet address = 200.1.19.1
mateo.elo.utfsm.cl      internet address = 200.1.17.4
> server ns.usm.cl
Default server: ns.usm.cl
Address: 200.1.21.80#53
> 200.1.17.102
Server:         ns.usm.cl
Address:        200.1.21.80#53

102.17.1.200.in-addr.arpa       name = juan.elo.utfsm.cl.
> exit
```

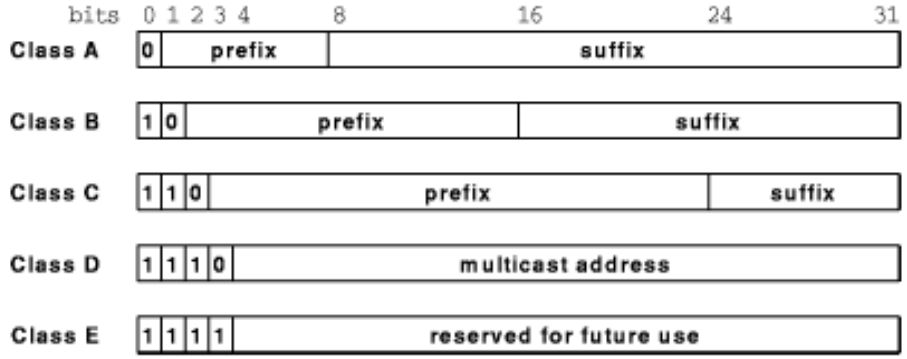
Direccionamiento IP: CLASES

Clases

- Porción de dirección de la red (subnet) se hace de tamaño fijo
- Ejemplo: Clase C



Classful addresing: Esquema original (con clases A, B, C, D, E)

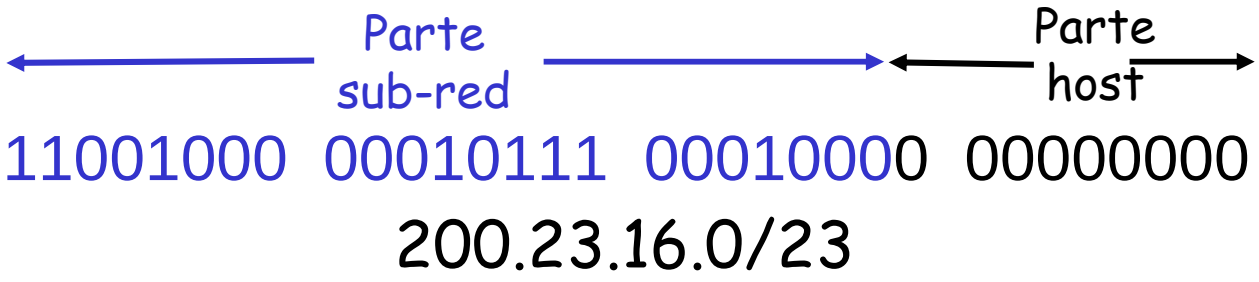


Clase A = subnet /8
Clase B = subnet /16
Clase C = subnet /24

Direccionamiento IP: CIDR

CIDR: Classless InterDomain Routing

- Porción de dirección de la red (subnet) se hace de tamaño arbitrario
- Formato de dirección: a.b.c.d/x, donde x es el # de bits de la dirección de subnet



Direcciones IP: ¿Cómo obtener una?

Q: ¿Cómo es que un *host* obtiene su dirección IP?

- Configurada por el administrador en un archivo
  - Windows: control-panel->network->configuration->tcp/ip->properties
  - UNIX: /etc/rc.config
- DHCP: Dynamic Host Configuration Protocol: el host obtiene la dirección dinámicamente desde un servidor
  - "plug-and-play" (más adelante)

Direcciones IP: ¿Cómo obtener una?

Q: ¿Cómo la red obtiene la dirección de subred parte de la dirección IP?

A: Obteniendo una porción del espacio de direcciones del proveedor ISP.

ISP's block	11001000 00010111 00010000	00000000	200.23.16.0/20
Organization 0	11001000 00010111 00010000	00000000	200.23.16.0/23
Organization 1	11001000 00010111 00010010	00000000	200.23.18.0/23
Organization 2	11001000 00010111 00010100	00000000	200.23.20.0/23
...	.....	....	....
Organization 7	11001000 00010111 00011110	00000000	200.23.30.0/23

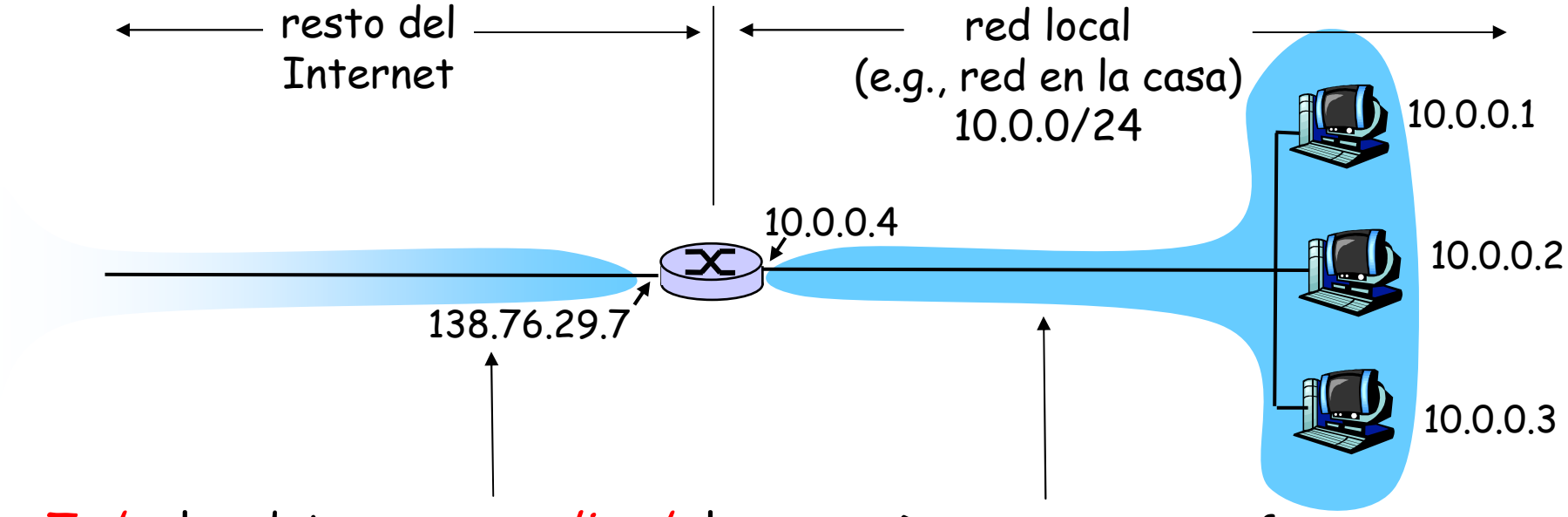
Direccionamiento IP: la última palabra...

Q: ¿Cómo un ISP obtiene un bloque de direcciones?

A: ICANN: Internet Corporation for Assigned Names and Numbers

- Asigna direcciones
- Administra DNS
- Asigna nombre de dominio, resuelve disputas

NAT: Network Address Translation



Todos los datagramas *saliendo* la red local tienen la misma dirección NAT IP: 138.76.29.7, pero diferentes números de puerto

Datagramas con fuente o destino en esta red tienen direcciones 10.0.0/24 fuente, destino (También se puede usar: 192.168.1/24)

NAT: Network Address Translation

- Motivación: la idea es usar sólo una dirección IP para ser vistos desde el mundo exterior:
  - No necesitamos asignación de un rango del ISP: sólo una dirección externa es usada por todos los dispositivos internos (computadores)
  - Podemos cambiar la dirección de dispositivos en red local sin notificar al mundo exterior
  - Podemos cambiar ISP sin cambiar direcciones de dispositivos en red local
  - Dispositivos dentro de la red no son explícitamente direccionables o visibles desde afuera (una ventaja de seguridad).

NAT: Network Address Translation

Implementación: ruteador NAT debe:

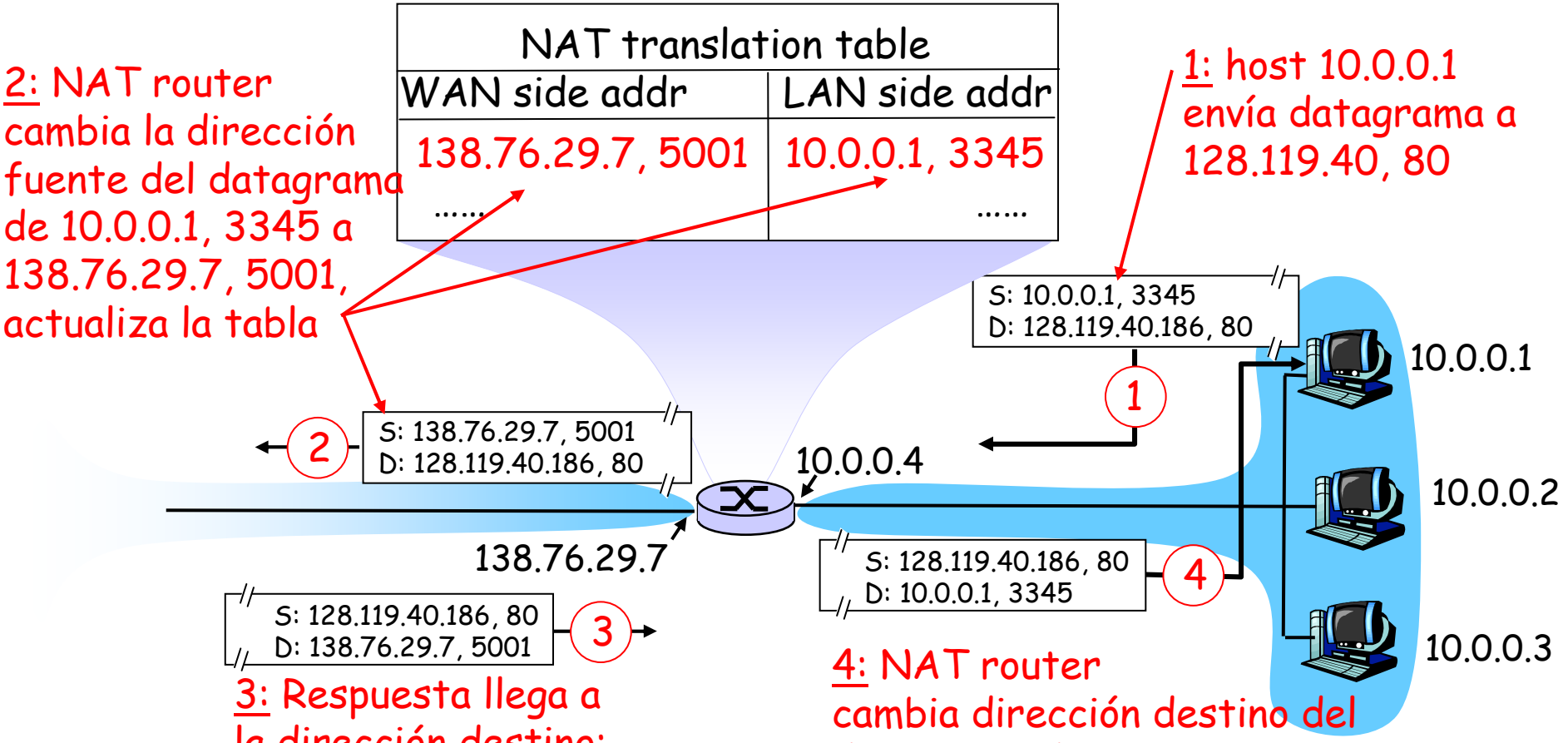
- *Datagramas salientes: remplazar* (IP fuente, # puerto) de cada datagrama saliente por (IP NAT, nuevo # puerto)

  . . . Clientes y servidores remotos responderán usando (IP NAT, nuevo # puerto) como dirección destino.

- *Recordar (en tabla de traducción NAT)* cada par de traducción (IP fuente, # puerto) a (IP NAT, nuevo # puerto)

- *Datagramas entrantes: remplazar* (IP NAT, nuevo # puerto) en campo destino de cada datagrama entrante por correspondiente (IP fuente, # puerto) almacenado en tabla NAT

NAT: Network Address Translation



NAT: Network Address Translation

- Campo de número de puerto es de 16 bits:
  - ~65,000 conexiones simultáneas con una única dirección dentro de la LAN!
- NAT es controversial:
  - Routers deberían procesar sólo hasta capa 3
  - Viola argumento extremo-a-extremo
    - Posiblemente los NAT deben ser tomados en cuenta por los diseñadores de aplicaciones, eg, aplicaciones P2P
  - En lugar de usar NAT, la carencia de direcciones debería ser resuelta por IPv6

Capítulo 4: Capa de Red

- 4. 1 Introducción
- 4.2 Circuitos virtuales y redes de datagramas
- 4.3 ¿Qué hay dentro de un router?
- 4.4 IP: Internet Protocol
  - Formato de Datagrama
  - Direccionamiento IPv4
  - ICMP
  - IPv6
- 4.5 Algoritmo de ruteo
  - Estado de enlace
  - Vector de Distancias
  - Ruteo Jerárquico
- 4.6 Ruteo en la Internet
  - RIP
  - OSPF
  - BGP
- 4.7 Ruteo Broadcast y multicast

ICMP: Internet Control Message Protocol

- Usado por hosts & routers para comunicar información a nivel de la red
  - Reporte de errores: host inalcanzable, o red, o puerto, o protocolo
  - Echo request/reply (usado por ping)
  - Usado por traceroute (TTL expired, dest port unreachable)
- Opera en capa transporte:
  - ICMP son llevados por datagramas IP
- Mensajes ICMP: tipo y código de error, más primeros 8 bytes del datagrama que causó el error

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - seldom used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

Traceroute e ICMP

- La fuente envía una serie de segmentos UDP al destino
  - Primero usa TTL =1
  - Luego usa TTL=2, etc.
  - Número de puerto no probablemente usado
- Cuando el n-ésimo datagrama llega a n-ésimo router:
  - Router descarta el datagrama, y
  - Envía a la fuente un mensaje ICMP "TTL expirado" (tipo 11, código 0)
  - Mensaje incluye nombre del router y dirección IP
- Cuando mensaje ICMP llega, la fuente calcula el RTT
- Traceroute hace esto 3 veces

Criterio de parada

- Segmento UDP eventualmente llega al host destino
- Host destino retorna paquete ICMP "puerto inalcanzable" (tipo 3, código 3)
- Cuando la fuente recibe este ICMP, para.

Capítulo 4: Capa de Red

- 4. 1 Introducción
- 4.2 Circuitos virtuales y redes de datagramas
- 4.3 ¿Qué hay dentro de un router?
- 4.4 IP: Internet Protocol
  - Formato de Datagrama
  - Direccionamiento IPv4
  - ICMP
  - IPv6
- 4.5 Algoritmo de ruteo
  - Estado de enlace
  - Vector de Distancias
  - Ruteo Jerárquico
- 4.6 Ruteo en la Internet
  - RIP
  - OSPF
  - BGP
- 4.7 Ruteo Broadcast y multicast

IPv6

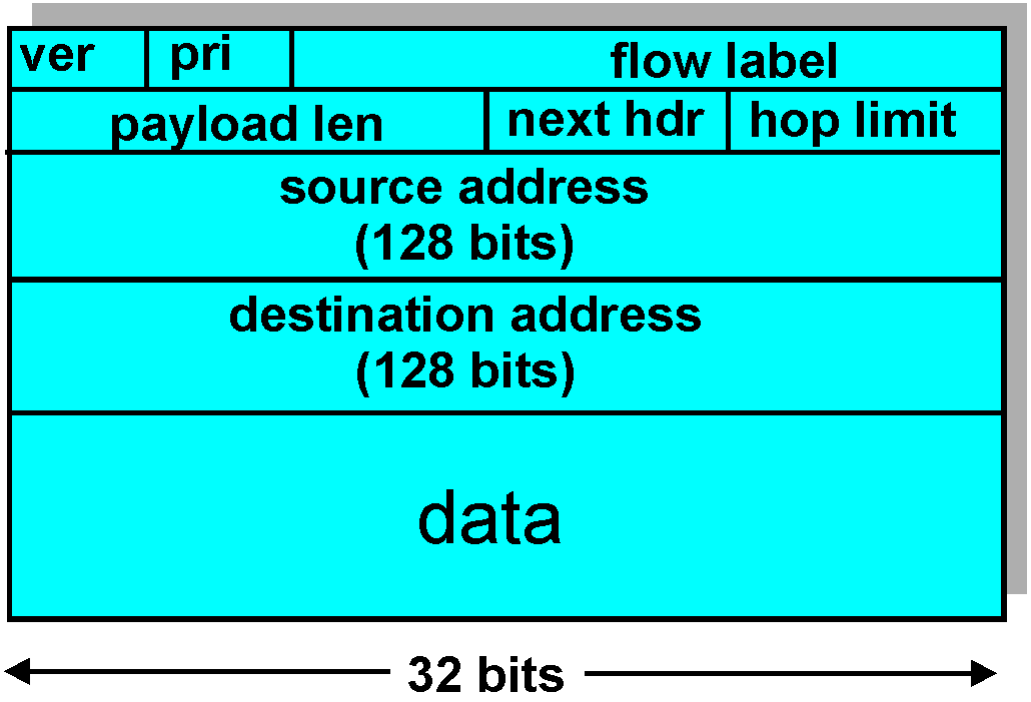
- Motivación Inicial: espacio de direcciones de 32-bit pronto serán completamente asignadas.
- Motivación adicional:
  - Formato de encabezado ayuda a acelerar el procesamiento y re-envío
  - Encabezado cambia para facilitar QoS

Formato de datagrama IPv6:

  - Encabezado de largo fijo de 40 bytes
  - Fragmentación no es permitida

Encabezado IPv6

- *Prioridad:* identifica prioridad entre datagramas en flujo
- *Flow Label:* identifica datagramas del mismo "flujo." (concepto de "flujo" no está bien definido).
- *Next header:* identifica protocolo de capa superior de los datos



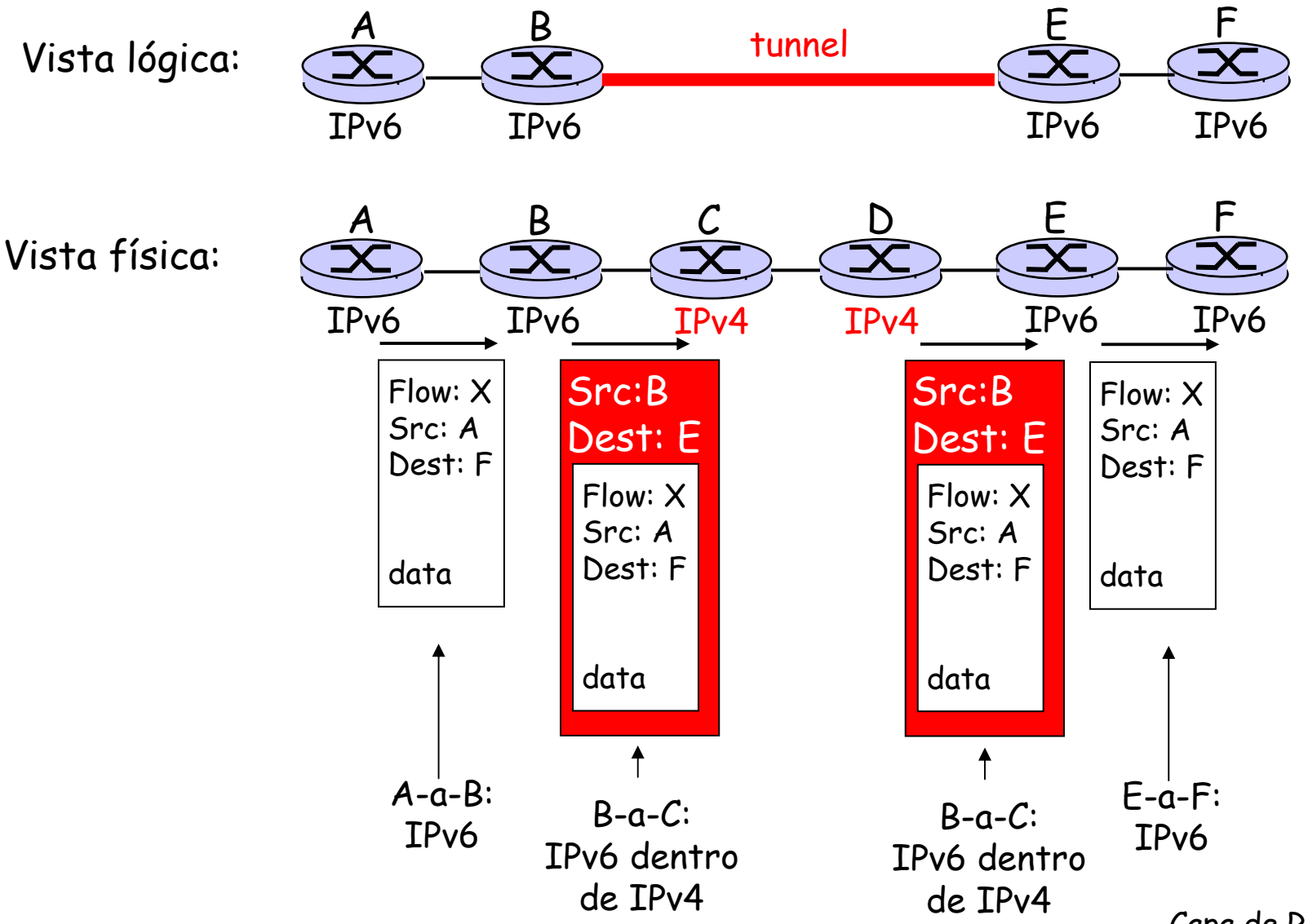
Otros cambios de IPv4 a v6

- *Checksum*: eliminada enteramente para reducir tiempo de procesamiento en cada router al ser redundante en capa transporte y enlaze (Ethernet)
- *Options:* permitidas, pero fuera del encabezado, indicado por campo "Next Header"
- *ICMPv6:* nueva versión de ICMP
  - Tipos de mensajes adicionales, e.g. "Paquete muy grande" (usado en el descubrimiento de MTU: unidad máxima de transmisión)
  - Funciones para administrar grupos multicast

Transición de IPv4 a IPv6

- No todos los routers pueden ser actualizados (upgraded) simultáneamente
  - No es posible definir un día para cambio "día de bajada de bandera"
  - ¿Cómo operará la red con routers IPv4 e IPv6 mezclados?
- *Tunneling:* IPv6 es llevado como carga en datagramas IPv4 entre routers IPv4

Tunneling


Vista lógica:
A
B
tunnel
E
F
IPv6
IPv6
IPv6
IPv6
Vista física:
A
B
C
D
E
F
IPv6
IPv6
IPv4
IPv4
IPv6
IPv6
Flow: X
Src: A
Dest: F
data
Src:B
Dest: E
Flow: X
Src: A
Dest: F
data
Src:B
Dest: E
Flow: X
Src: A
Dest: F
data
Flow: X
Src: A
Dest: F
data
A-a-B:
IPv6
B-a-C:
IPv6 dentro
de IPv4
B-a-C:
IPv6 dentro
de IPv4
E-a-F:
IPv6

Capítulo 4: Capa de Red

- 4. 1 Introducción
- 4.2 Circuitos virtuales y redes de datagramas
- 4.3 ¿Qué hay dentro de un router?
- 4.4 IP: Internet Protocol
  - Formato de Datagrama
  - Direccionamiento IPv4
  - ICMP
  - IPv6
- 4.5 Algoritmos de ruteo
  - Estado de enlace
  - Vector de Distancias
  - Ruteo Jerárquico
- 4.6 Ruteo en la Internet
  - RIP
  - OSPF
  - BGP
- 4.7 Ruteo Broadcast y multicast