INP Grenoble
DÉPARTEMENT
TÉLÉCOMMUNICATIONS

Advanced Computer Networks

Interconnection Layer 3: IP

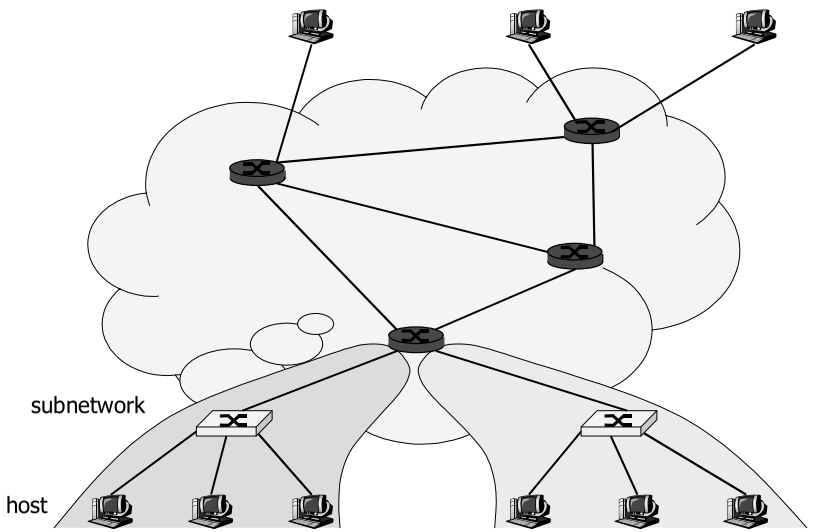
Prof. Andrzej Duda
duda@imag.fr

http://duda.imag.fr

Contents

- Principles of IP protocol
- Addressing
  - allocation
  - CIDR
- Host configuration
- Multicast and IGMP

Architecture - layer 3



IP principles

- Elements
  - **host** = end system; **router** = intermediate system; **subnetwork** = a collection of hosts that can communicate directly without routers
- Routers are between subnetworks only:
  - a subnetwork = a collection of systems with a common prefix
- Packet forwarding
  - **direct**: inside a subnetwork hosts communicate directly without routers, router delivers packets to hosts
  - **indirect**: between subnetworks one or several routers are used
- Host either sends a packet to the destination using its LAN, or it passes it to the router for forwarding

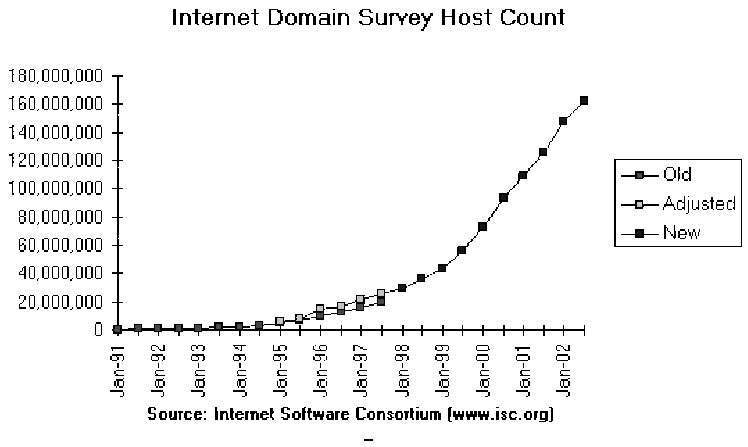
IP addresses

- Unique addresses in the world, decentralized allocation
- An IP address is 32 bits, noted in dotted decimal notation: **`192.78.32.2`**
- An IP address has a prefix and a host part:
  - **`prefix:host`**
- Two ways of specifying prefix
  - subnet mask identifies the prefix by bitwise & operation
  - CIDR: bit length of the prefix
- Prefix identifies a subnetwork
  - used for locating a subnetwork - routing

IP addresses

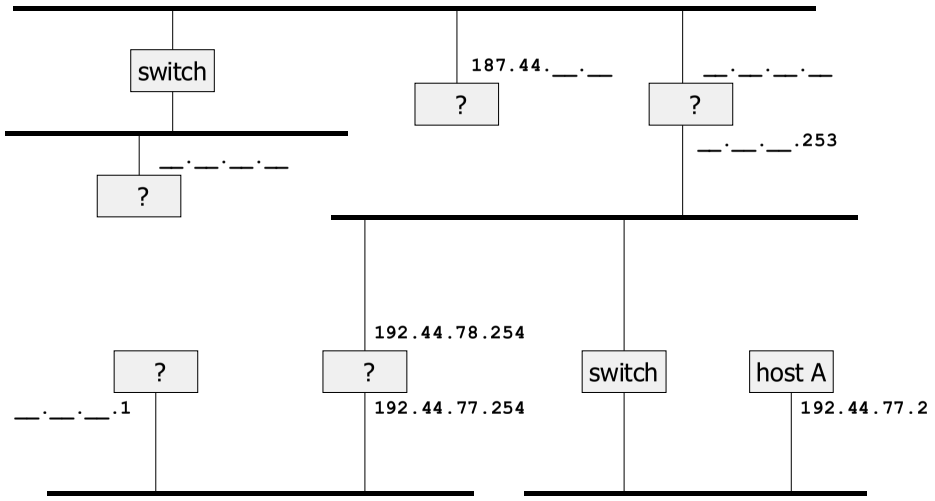
- Scalability
  - short prefix may aggregate many subnetworks (compare to flat MAC addresses)
- Mapping to MAC addresses
  - ARP maintains IP - MAC mapping
- Users use names instead of addresses
  - names mapped to IP addresses by DNS

Number of hosts



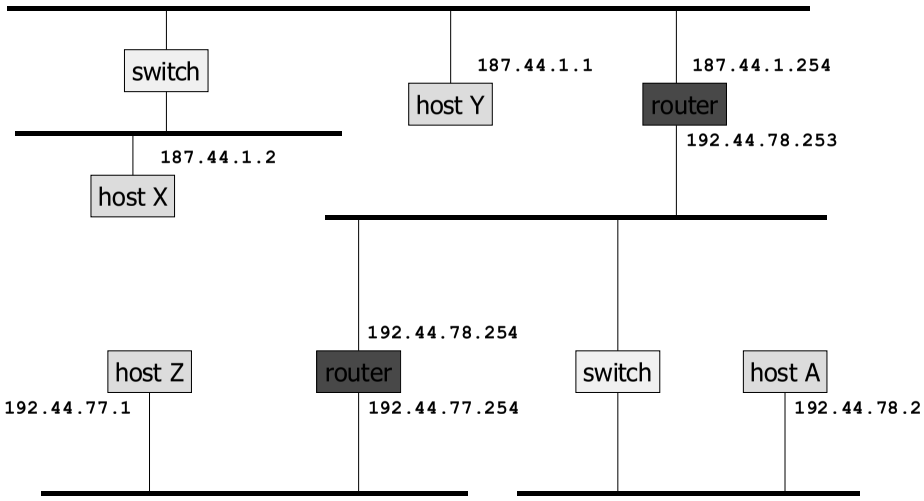
Used addresses in Internet

Example



- Can host A have this address?

Example



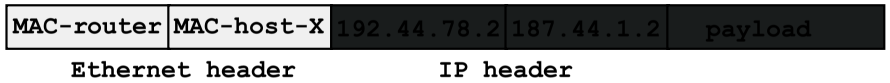
- Host A is on subnetwork `192.44.78`

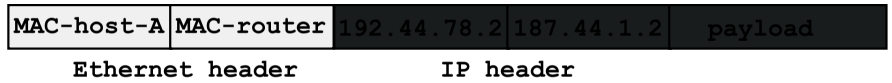
Packet delivery

Packet sent by `187.44.1.2 to 187.44.1.1`

X needs to know MAC address of Y (ARP)

Packet sent by `187.44.1.2 to 192.44.78.2`

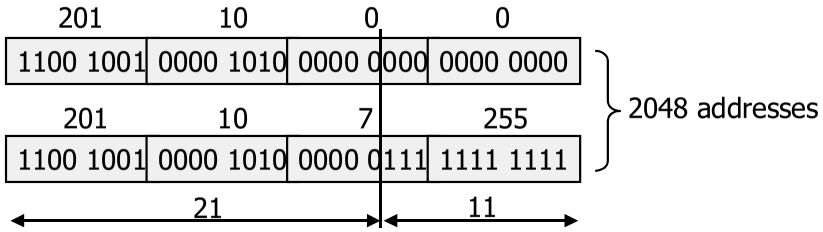




X needs to know MAC address of router (X knows the IP address of router - configuration)

Router needs to know MAC address of A

CIDR *Classless Interdomain Routing*

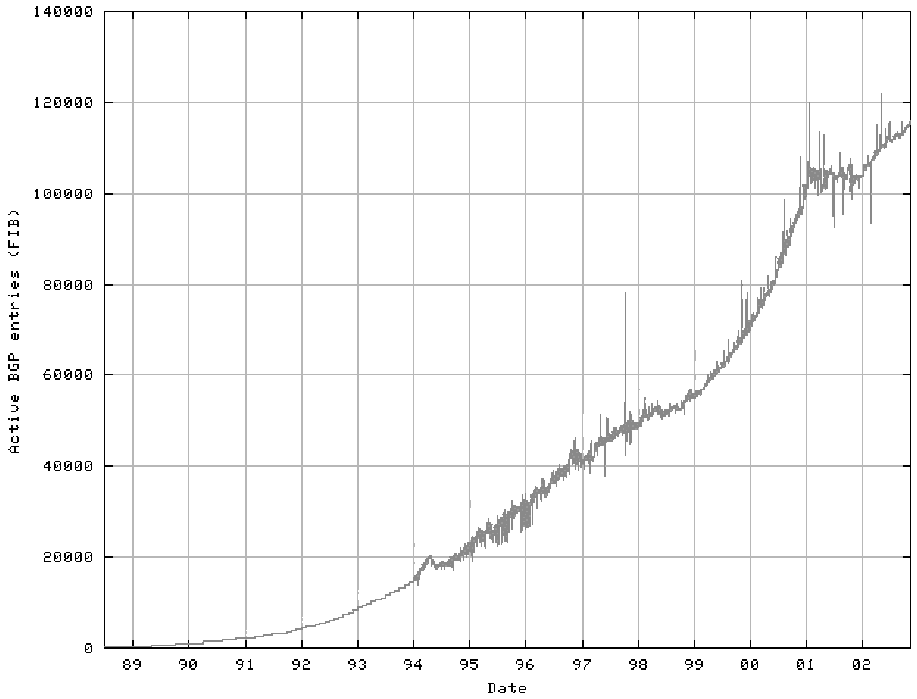


201.10.0.0/21 : 201.10.0.0 - 201.10.0.255
201.10.1.0 - 201.10.1.255
...
201.10.7.0 - 201.10.7.255

1 C class network : 256 addresses
256 □ 8 = 2048 addresses

Address allocation

- World Coverage
  - Europe and the Middle East (RIPE NCC)
  - Africa (ARIN & RIPE NCC)
  - North America (ARIN)
  - Latin America including the Caribbean (ARIN)
  - Asia-Pacific (APNIC)
- Current allocations of Class C
  - `193-195/8, 212-213/8, 217/8` for RIPE
  - `199-201/8, 204-209/8, 216/8` for ARIN
  - `202-203/8, 210-211/8, 218/8` for APNIC
- Simplifies routing
  - short prefix aggregates many subnetworks
  - routing decision is taken based on the short prefix



Address delegation

- Europe
  - `62/8, 80/8, 193-195/8, …`
  - ISP-1
    - `62.125/16`
    - `Site 1`
      - `62.125.44.128/25`
    - `Site 2`
      - `62.125.44.50/24`
  - ISP-2
    - `195.44/14`
    - `Site 1`
      - `195.46.216/21`
    - `Site 2`
      - `195.46.224/21`

Renumbering?

- Europe
  - `62/8, 80/8, 193-195/8, …`
  - ISP-1
    - `62.125/16`
    - `Site 1`
      - `62.125.44.128/25`
  - ISP-2
    - `195.44/14`
    - `Site 1`
      - `195.46.216/21`
    - `Site 2`
      - `195.46.224/21`
    - ***`Site 2' ?`***
      - ***`62.125.44.50/24`***

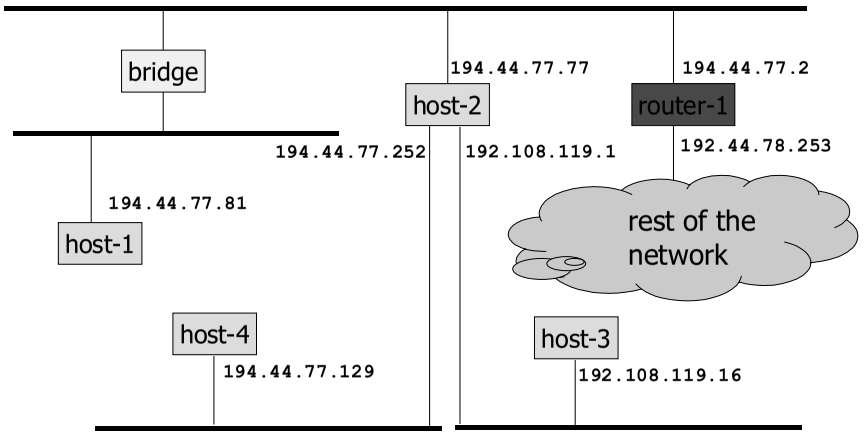
⇒ explicit route to Site 2'

Configuration of a Unix host

```
/usr/etc/ifconfig interface [ address_family ]
             [ address [ dest_address ] ] [ netmask mask ]
             [ broadcast address ] [ up ] [ down ] [ trailers ]
             [ -trailers ] [ arp ] [ -arp ] [ private ]
             [ -private ] [ metric n ] [ auto-revarp ]

host-1# ifconfig le0 host-1 netmask +
Setting netmask of le0 to 255.255.255.128
# + means netmask from /etc/netmasks
host-1# ifconfig -a
le0: flags=863<UP,BROADCAST,NOTRAILERS,RUNNING>
       inet 192.44.77.81 netmask ffffff80 broadcast 192.44.77.0
       ether 8:0:20:1c:74:84
lo0: flags=849<UP,LOOPBACK,RUNNING>
       inet 127.0.0.1 netmask ff000000
```


Example interconnection



Routing tables

```
host-1 (192.44.77.81) :
>netstat -n -r
Routing tables
Destination      Gateway       Flags    Refcnt Use   Interface
192.108.119.16   192.44.77.77  UGHD     1      1683  le0
127.0.0.1        127.0.0.1     UH       2      12971 lo0
default          192.44.77.2   UG       3      16977 le0
192.44.77.0      192.44.77.81  U        13     5780  le0

U - up
G - gateway (next router)
H - host route
D - route from ICMP Redirect
```


Routing tables

```
host-2 (192.44.77.77) :
>rsh host-2 netstat -n -r
Routing tables
Destination    Gateway        Flags  Refcnt  Use      Interface
127.0.0.1      127.0.0.1      UH     3       351344   lo0
default        192.44.77.2    UG     3       17388997 le0
192.44.77.128  192.44.77.252  U      26      504768   le2
192.44.77.0    192.44.77.77   U      24      10702069 le0
192.108.119.0  192.108.119.1  U      2       249777   le1
```

Modifying routing tables

```
/usr/etc/route [ -fn ] add|delete [ host|net ] destination
   [gateway [ metric ] ]
host-1# netstat -r
Routing tables
Destination     Gateway      Flags     Refcnt Use     Interface
localhost       localhost    UH        2      13569     lo0
192.44.77.0     host-1       U         18     13272     le0
host-1# ping 133.11.11.11
sendto: Network is unreachable
host-1# route add 0.0.0.0 router-1 1
add net 0.0.0.0 gateway router-1
```

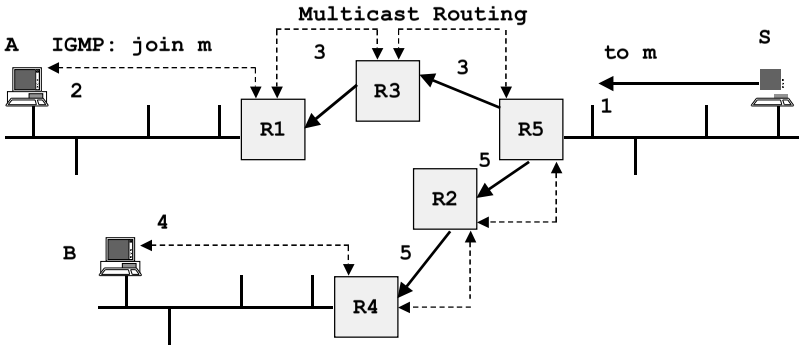

Modifying routing tables

```
host-1# netstat -r
Routing tables
Destination     Gateway      Flags     Refcnt Use     Interface
localhost       localhost    UH        2      13591     lo0
default         router-1     UG        0      0         le0
192.44.77.0     host-1       U         16     13566     le0
host-1# ping 133.11.11.11
133.11.11.11 is alive
```


IP Broadcasting, Multicasting

- Broadcast = send to all
  - sent to all hosts on one net/subnet; used by NetBIOS for discovery
- Multicast = send to a group
  - IP multicast address = class D = `224.0.0.0` to `239.255.255.255`
    `224.0.0.1` = all multicast capable systems on subnet
    `224.0.0.2` = all multicast capable routers on subnet
  - used for: routing, conferencing, radio distribution, …
- IP uses open group paradigm
  - multicast IP addresses are logical (= non topological)
  - for receiving data sent to multicast address *m*, a host must subscribe to *m*
  - for sending to multicast address *m*, a host simply put *m* in the dest addr field

IP Multicast Principles



- hosts subscribe via IGMP join messages sent to router
- routers build distribution tree via multicast routing
- sources do not know their destinations
- packet replication is done by routers

IP Multicast Forwarding Algorithm

Packet Forwarding (host, router)

```
Read address MA = destination IP@

/* assume it is multicast */
for every physical interface PI
      if MA is enabled on PI then
            send directly to PI
```

At lrcsuns: Physical Interface Tables

IP	subnetMask
128.178.156.24	255.255.255.0
224.2.166.207	
224.2.127.255	

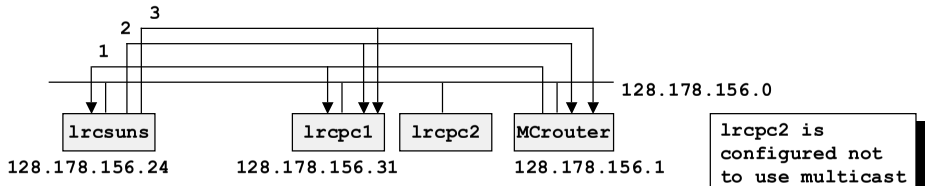
Send directly (Ethernet)

```
send directly(MA, MAC@):
  map last 23 bits of MA to last 23 bits
    of MAC address
  send MAC frame with      DA = 01-00-5E-xx-xx-xx,
                           SA = own i/f address
```

- Systems have to know which group they belong to
  - Hosts: application processes register to IP
  - Routers: learn if members present with IGMP
- Direct send to link layer:
  - algorithmic mapping of 23 last bits : ex : 224.2.166.207 -> 01-00-5E-02-A6-CF

IGMP: Internet Group Management Protocol

- Purpose: manage group membership inside one subnet
- routers: know if group is present on an interface
  - know whether to forward locally or not
- hosts: know if a multicast address is already in use locally



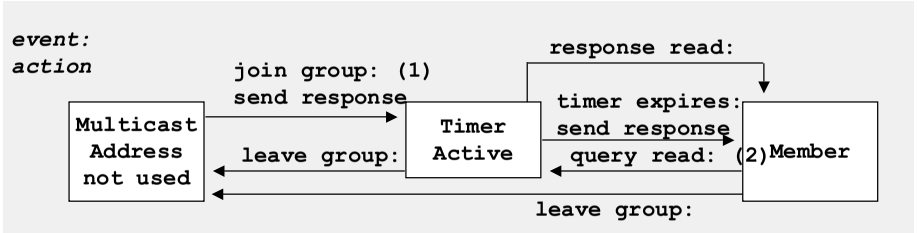
- 1: IGMP query, TTL =1, IGMP group @ = 0
  dest IP@ = 224.0.0.1; source IP@ = 128.178.156.1
- 2: IGMP report, TTL =1, IGMP group @ = 224.2.166.207
  dest IP@ = 224.2.166.207; source IP@ = 128.178.156.24
- 3: IGMP report, TTL =1, IGMP group @ = 224.2.127.255
  dest IP@ = 224.2.127.255; source IP@ = 128.178.156.24

IGMP Host Implementation

Host Implementation

- goal: avoid avalanche effects - one router originated query might cause a burst of reports
- solution = synchronization avoidance protocol
  - 1. hosts delay responses randomly
  - 2. hosts listen to responses, only first one answers

Host IGMP Finite State Machine



Conclusion

- CIDR - current designation of subnetworks
  - variable length prefix
- Addresses chosen in function of the subnetwork size
- Aggregation at the connection point to ISP
  - all subnetworks announced as one prefix
- Dynamic mapping of IP addresses to MAC
  - address resolution when needed