The Network Layer

Redes de Computadores

2021/22

Pedro Brandão

1

References

- These slides are from "Computer Networking: A Top Down Approach 5th edition. Jim Kurose, Keith Ross Addison-Wesley, April 2009"
  - With adaptations/additions by Manuel Ricardo and Pedro Brandão

2

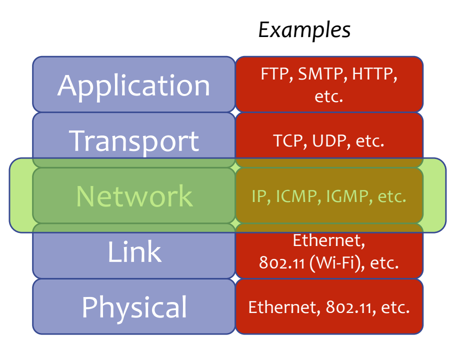
2

Driving questions…

- What are the main functions of the network layer?
- What are the differences between virtual circuit and datagram networks?
- How is forwarding handled in both type of networks?
- What are the main functions of a router?
- What are the formats of IP addresses?
- How to form subnets?
- What services are provided by ARP, ICMP, DHCP and NAT? How do these protocols work?
- What are differences the between IPv4 and IPv6?

Internet protocol stack

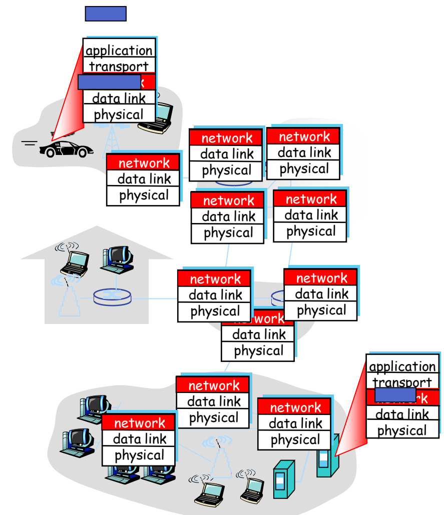
- Application: network processes
- Transport: data transfer between processes
- Network: packet routing between source and destination
- Link: data transfer between adjacent network elements
- Physical: bits on the "wire"



RCom 21/22 - The Network Layer

Network Layer Overview

- Network layer
  - transports packets (datagrams)
  - from sending host to receiving host
  - functions located in every host and router
- Sender
  - encapsulates transport data into packets
  - generates packets
- Receiver
  - receives packets
  - delivers data to transport layer
- Router
  - Receives packets from input line
  - examines network layer header
  - forwards packets through adequate output line(s)



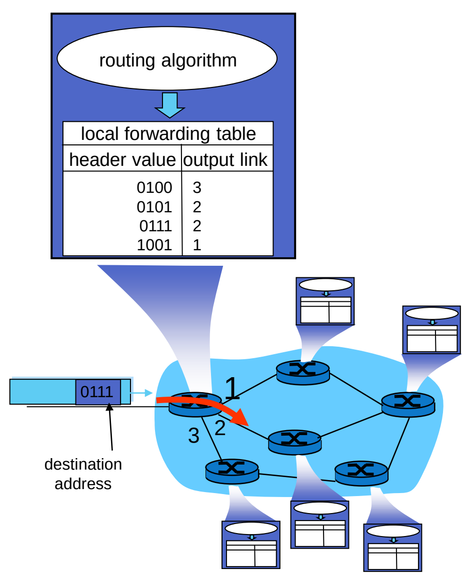
5

5

RCom 21/22 - The Network Layer

Network Layer – Main Functions

- Routing
  - determine route taken by packets, from source to destination
  - Algorithms using cost function (usually shortest path)
  - *Analogy: process of planning trip from source to destination*
- Forwarding
  - router forwards packet **from input port to output port**
  - *Analogy: process of getting through single interchange*



6

6

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

U.PORTO

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RCom 21/22 - The Network Layer

Virtual Circuits and Datagram Networks

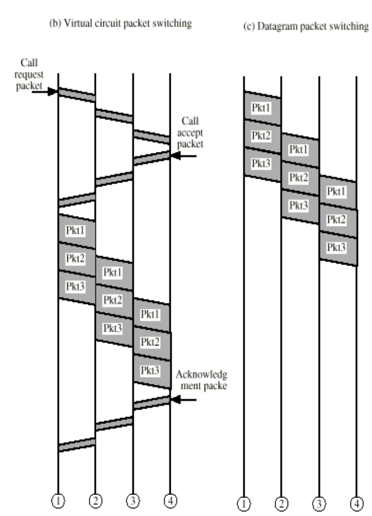
(Network layer)

7

7

Network Layer – Connection and Connectionless Service

- Services provided by network layer
  - Virtual Circuit network
    - ➔ connection-oriented service
  - Datagram network
    - ➔ connectionless service



RCom 21/22 - The Network Layer

U.PORTO

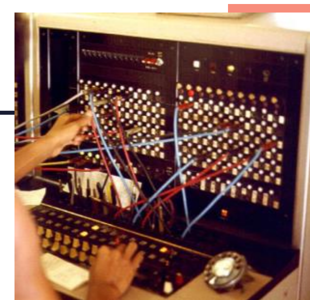
8

8

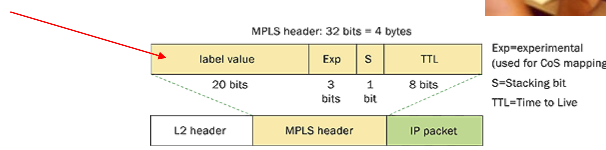
RCom 21/22 - The Network Layer

Virtual Circuit (VC)

Image from Joseph A. Carr, in [Wikipedia Telephone Exchange]



- Phases
  circuit establishment ➔ data transference ➔ circuit termination
- Packet carries identifier of Virtual Circuit



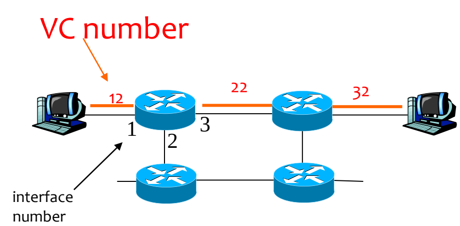
- Path defined from source to destination
  - sequence of VC identifiers, one for each link along path
- Router
  - maintains "state" for every supported circuit
  - may allocate resources (bandwidth, buffers) per Virtual Circuit

9

9

RCom 21/22 - The Network Layer

VC - Forwarding Table



Forwarding table in northwest router:

Incoming interface	Incoming VC #	Outgoing interface	Outgoing VC #
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
…	…	…	…

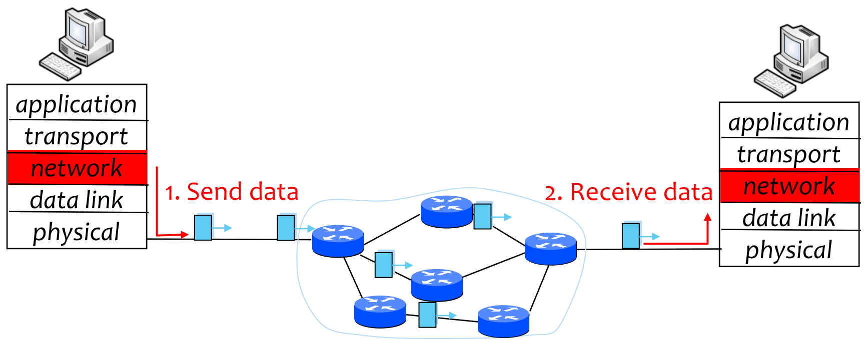
Routers maintain connection state information!

10

10

Datagram Networks

- No circuit establishment; no circuit concept
- Packets
  - forwarded using destination host address
  - packets between same source-destination pair may follow different paths



Forwarding Table

- IP Address
  - 32 bits

2^{32} possible entries in IPv4

Destination Address Range	Output Link interface
address X through address Z	0
address W through address Y	1
address A through address K	1
address P through address R	2
Otherwise	3

RCom 21/22 - The Network Layer

U.PORTO

- How to reduce the number of entries in the forwarding table?

13

13

RCom 21/22 - The Network Layer

U.PORTO

Longest Prefix Matching

- Which interface?
- DA: 11001000 00010111 00010110 10100001
  - ➔Itf: 0
- DA: 11001000 00010111 00011000 10101010
  - Can be itf 1 or 2
  - Longest prefix 1

Prefix	Output Link interface
11001000 00010111 00010	0
11001000 00010111 00011000	1
11001000 00010111 00011	2
Otherwise	3

14

14

RCom 21/22 - The Network Layer

Virtual-Circuit versus Datagram Networks

Issue	Datagram subnet	Virtual-circuit subnet
Circuit setup	Not needed	Required
Addressing	Each packet contains the full source and destination address	Each packet contains a short VC number
State information	Routers do not hold state information about connections	Each VC requires router table space per connection
Routing	Each packet is routed independently	Route chosen when VC is set up; all packets follow it
Effect of router failures	None, except for packets lost during the crash	All VCs that passed through the failed router are terminated
Quality of service	Difficult	Easy if enough resources can be allocated in advance for each VC
Congestion control	Difficult	Easy if enough resources can be allocated in advance for each VC

15

15

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

U.PORTO

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RCom 21/22 - The Network Layer

Internet Protocol

RFC 791/STD 5

IP RFCs

(Network layer)

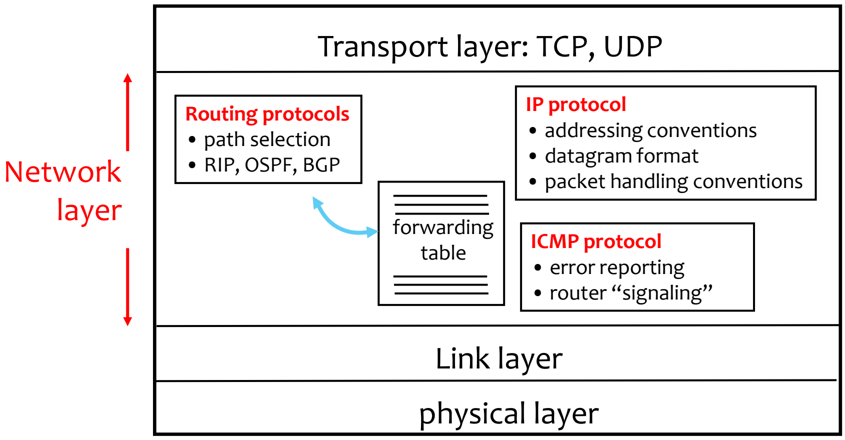
16

16

RCom 21/22 - The Network Layer

The Internet Network layer

- Host, router network layer functions

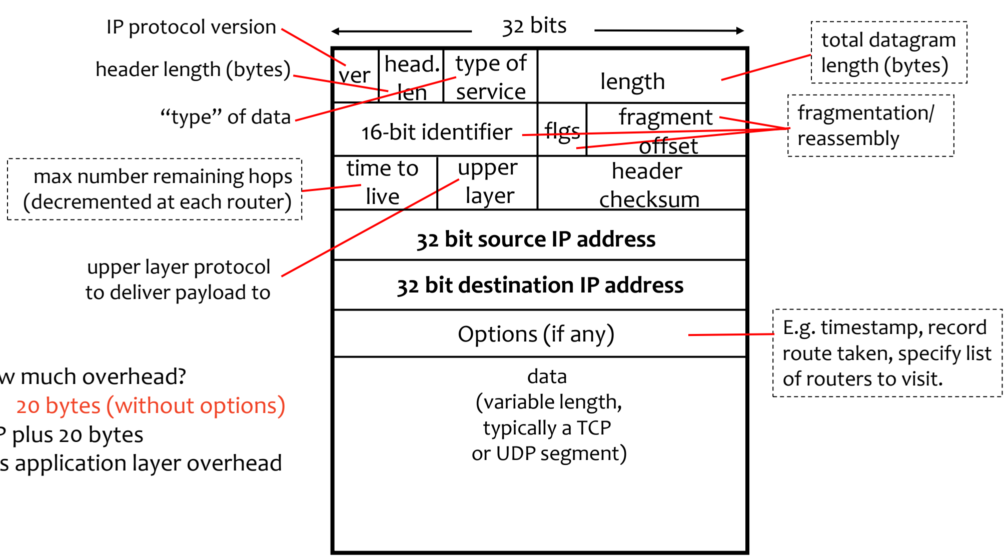


17

17

RCom 21/22 - The Network Layer

IP Datagram Format



- How much overhead?
  - 20 bytes (without options)
- TCP plus 20 bytes
- Plus application layer overhead

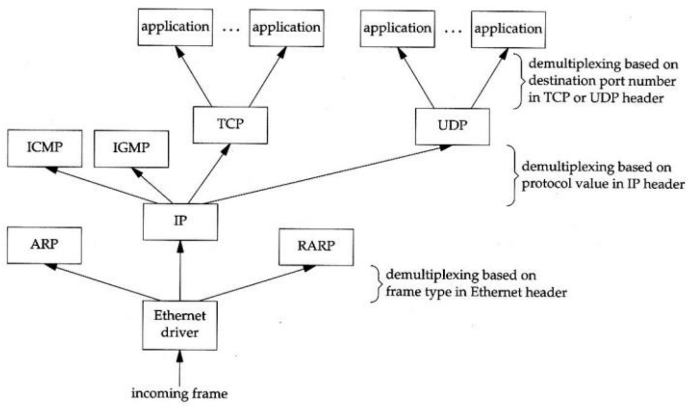
18

18

RCom 21/22 - The Network Layer

Demultiplexing for upper layers

- Ethernet header (type)
  - IP - 0x0800
  - ARP - 0x0806
  - RARP - 0x8035
  - IPX- 0x8037
  - IPv6 - 0x86DD
  - MPLS - 0x8847
- IP header (protocol)
  - ICMP - 1
  - IGMP - 2
  - TCP - 6
  - UDP – 17
- TCP/UDP header (port)
  - FTP – 21
  - Telnet - 23
  - HTTP - 80
  - SMTP - 25



19

19

RCom 21/22 - The Network Layer

Internet Checksum

- The Internet (not layer 2) uses a checksum
  - easily implementable in software
  - 1's complement sum of 16 bit words
  - Performance: d=2

```
u_short
cksum(u_short *buf, int count)
{
    register u_long sum = 0;

    while (count--)
    {
        sum += *buf++;
        if (sum & 0xFFFF0000)
        {
            /* carry occurred,
               so wrap around */
            sum &= 0xFFFF;
            sum++;
        }
    }
    return ~(sum & 0xFFFF);
}
```

- One's complement sum
  - Mod-2 addition with carry-out
  - Carry-out in the most-significant-bit is added to the least-significant bit
  - Get one's complement of "one's complement sum"

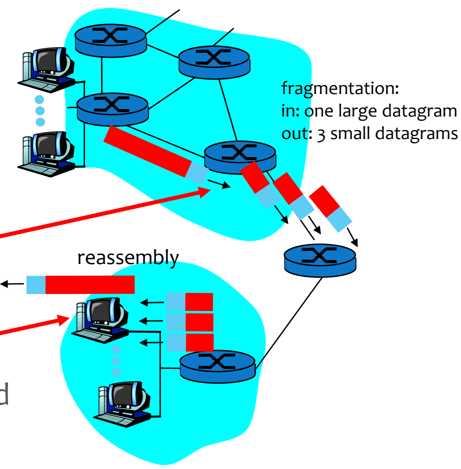
	1010011
	0110110
carry-out 1	0001001
Carry wrap-around	0000001
	0001010
One's complement =	**1110101**

20

20

IP Fragmentation and Reassembly

- Network links have MTU
  - MTU - max. transfer unit (size)
  - largest possible link-level frame
  - different link types, different MTUs
- Large IP datagram is fragmented
  - one datagram ➔ n datagrams
  - “reassembled” at final destination
  - IP header bits used to identify, order related fragments



21

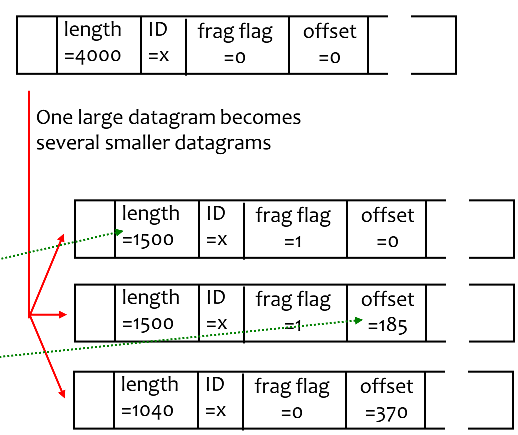
IP Fragmentation and Reassembly Example

Example

- 4000 byte datagram
- 3980 bytes data + 20 bytes IP header
- MTU = 1500 bytes

1480 bytes in data field

offset = 1480/8

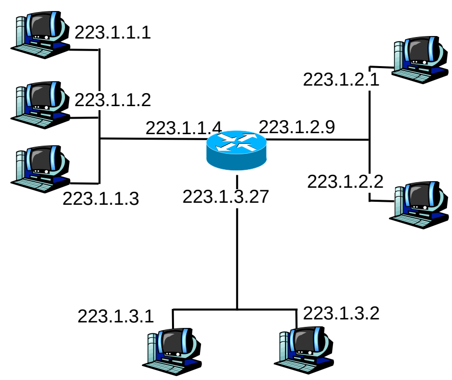


22

RCom 21/22 - The Network Layer

IP Addressing - Introduction

- IP address
  - 32-bit identifier for host/router interface
- Interface
  - connection between host/router and physical link
  - Routers have multiple interfaces
  - IP addresses associated with interface



223.1.1.1 = 11011111 00000001 00000001 00000001

$\underbrace{11011111}_{223}\ \underbrace{00000001}_{1}\ \underbrace{00000001}_{1}\ \underbrace{00000001}_{1}$

23

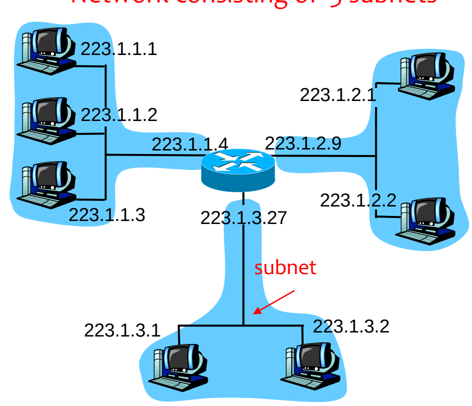
23

RCom 21/22 - The Network Layer

Subnets

- IP address
  - subnet part ➔ high order bits
  - host part ➔ low order bits
- Subnet ➔set of interfaces
  - with same subnet part of IP address
  - can reach each other without router intervention

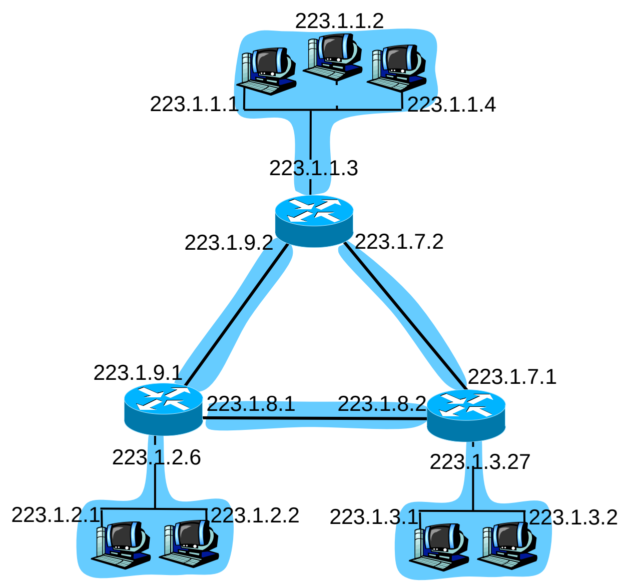
Network consisting of 3 subnets



24

24

6 Subnets



RCom 21/22 - The Network Layer

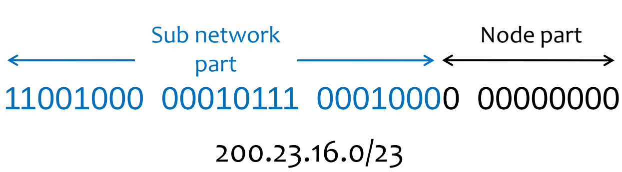
25

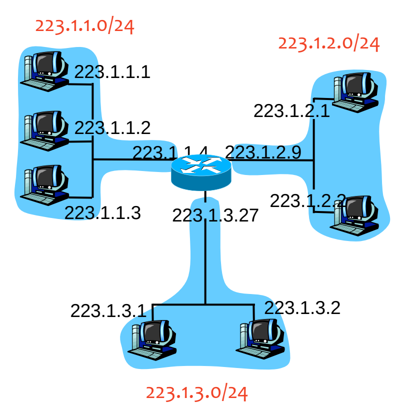
25

IP Addressing - CIDR

CIDR: Classless InterDomain Routing (RFC4632/BCP122*)

- subnet portion of address has arbitrary length
- address format ➔ a.b.c.d/x
  - where x is # bits in subnet portion of address





* BCP – Best Current Practice

RCom 21/22 - The Network Layer

26

26

RCom 21/22 - The Network Layer

U.PORTO

Special IP Addresses

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		This host
0 0 . . . 0 0	Host	A host on this network
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		Broadcast on the local network
Network	1 1 1 1 . . . 1 1 1 1	Broadcast on a distant network
127	(Anything)	Loopback

27

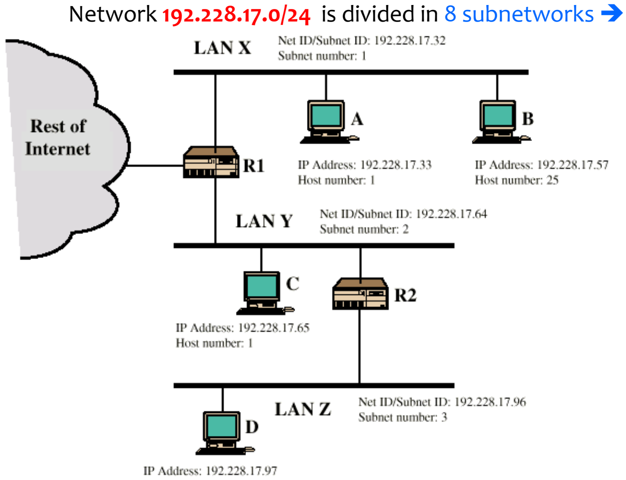
27

RCom 21/22 - The Network Layer

U.PORTO

Forming Sub-Networks

Network **192.228.17.0/24** is divided in 8 subnetworks ➔ masks of 27 bits



Subnetwork mask – 27 bits

subnetid – 3 bits (8 subnetworks)

11000000 11100100 00010001 01100000

192.228.17.96/27

hostid – 5 bits

30 hosts per subnet supported

all 0 – identifies subnet

all 1 – broadcast address

Example of subnetworks

192.228.17.0/27 (.00000000)

192.228.17.32/27 (.00100000)

192.228.17.64/27 (.01000000)

192.228.17.96/27 (.01100000)

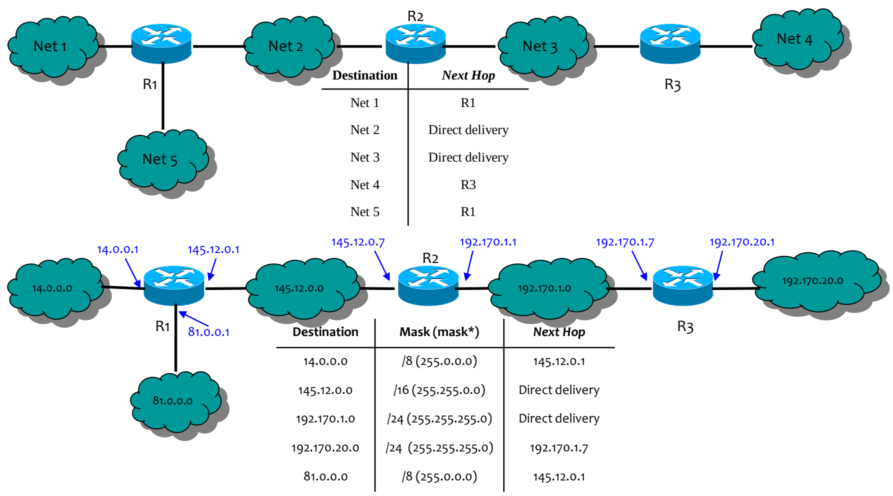
….

192.228.17.224/27 (.11100000)

28

28

Forwarding Table at R2



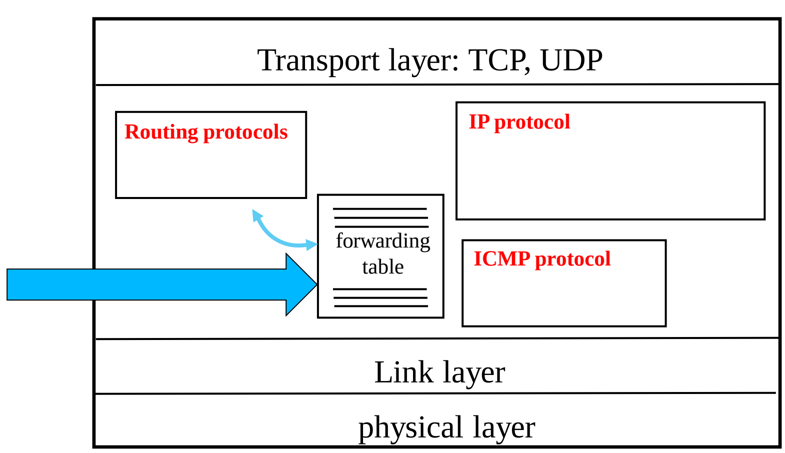
Destination	Next Hop
Net 1	R1
Net 2	Direct delivery
Net 3	Direct delivery
Net 4	R3
Net 5	R1

Destination	Mask (mask*)	Next Hop
14.0.0.0	/8 (255.0.0.0)	145.12.0.1
145.12.0.0	/16 (255.255.0.0)	Direct delivery
192.170.1.0	/24 (255.255.255.0)	Direct delivery
192.170.20.0	/24 (255.255.255.0)	192.170.1.7
81.0.0.0	/8 (255.0.0.0)	145.12.0.1

29

29

Forwarding Table use



30

30

RCom 21/22 - The Network Layer

- What is a loopback interface? What is its IP address?

P

31

31

RCom 21/22 - The Network Layer

IP Forwarding Function

- Forwarding table has entries in format
  <networkAddress/mask, port>
- Forwarding function
  - When a datagram arrives with destination address **A**, then
    - For each entry of the forwarding table
      `val= A & mask* // e.g., mask=8, mask*=255.0.0.0 = (bin)11111111.0.0.0`
      `if (val == networkAddress & mask*)`
      - add corresponding output port to the set of candidate ports
    - Select the port with the largest mask → most specific route

32

32

RCom 21/22 - The Network Layer

U.PORTO

IP Forwarding Function - Example

- frdTbl={<128.32.0.0/16,1>, <128.32.192.0/18,3>, <128.0.0.0/8,5>}
- Datagram with destination address A=128.32.195.1
- Set of candidate output ports
  - ➔ {1,3,5}.
    - Port 1: 128.32.0.1 - 128.32.255.254
    - Port 3: 128.32.192.1 - 128.32.255.254
    - Port 5: 128.0.0.1 - 128.255.255.254
- Selected port
  - 3 largest mask, 18 bits

One online IP Subnet Calculator

P

33

33

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

U.PORTO

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RCom 21/22 - The Network Layer

Address Resolution Protocol

RFC 826/STD 37

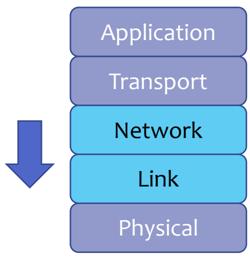
(Network layer)

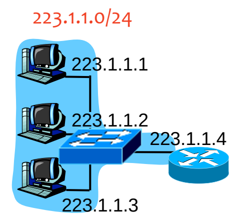
34

34

What is ARP needed for?

- On the same LAN you need to connect on the link layer
  - Known: IP address of destination (packet from Network Layer)
  - Unknown: MAC address of destination





ARP – Address Resolution Protocol

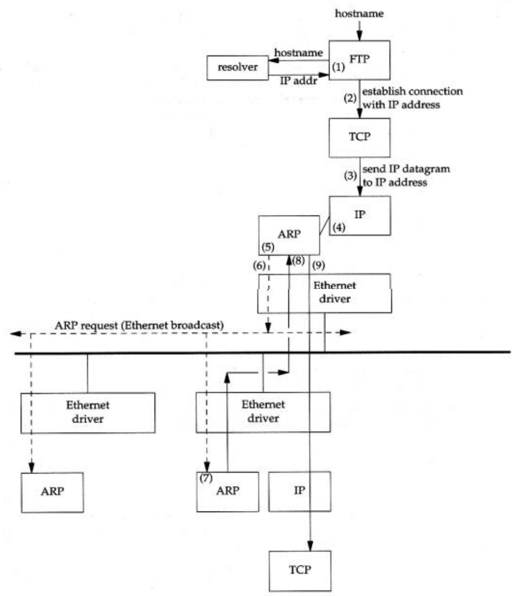
- A network interface has
  - one MAC address
  - one (or more) IP address(es)
- ARP: Address Resolution Protocol
  - Protocol used to obtain the MAC address associated to a given IP address
- RARP – Reverse ARP
  - Protocol used to obtain the IP address associated to a MAC address

ARP Example

frame hdr	ARP/RARP message	datalink frame

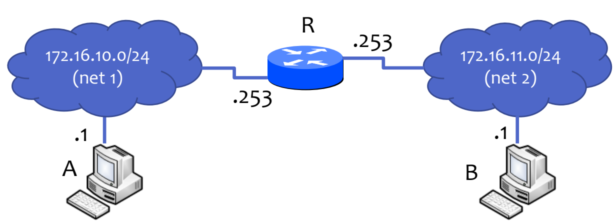
0	15 16	31
Hardware type:16		Protocol type:16
hlen:8	plen:8	ARP Operation:16
Sender MAC addr (bytes 0-3)		
sender MAC addr (bytes 4-5)		sender IP addr (bytes 0-1)
sender IP addr (bytes 2-3)		dest MAC addr (bytes 0-1)
dest MAC addr (bytes 2-5)		
dest IP addr (bytes 0-3)		

- **hardware type : Ethernet=1 ARCNET=7, localtalk=11**
- **protocol type : IP=0x800**
- **hlen : length of hardware address, Ethernet=6 bytes**
- **plen : length of protocol address, IP=4 bytes**
- **ARP operation : ARP request = 1, ARP reply = 2**
  RARP request = 3, RARP reply = 4



37

ARP/IP addresses



- Assume host A sends an IP packet to host B and that this packet is forwarded by router R. What are the MAC and IP addresses (source and destination) observed?

- What roles does ARP play in this scenario?

38

RCom 21/22 - The Network Layer

Obtaining IP Addresses

RFC 2131 Dynamic Host Configuration Protocol

(Network layer)

39

39

RCom 21/22 - The Network Layer

How to Obtain IP Addresses

- How does network get subnet part of IP addresss?
  - Gets allocated portion of its provider ISP's address space

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

40

40

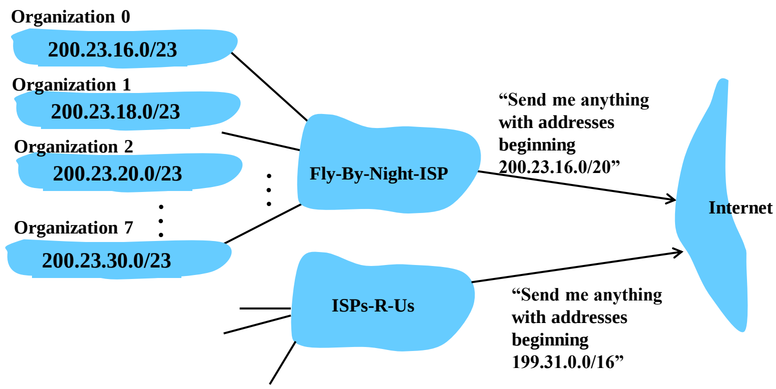
RCom 21/22 - The Network Layer

U.PORTO

Hierarchical Addressing - Route Aggregation

- Hierarchical addressing

  allows efficient advertisement of routing information



41

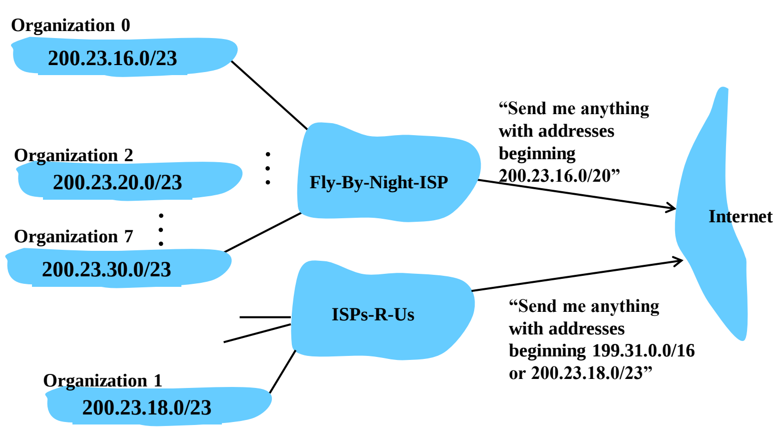
41

RCom 21/22 - The Network Layer

U.PORTO

Hierarchical Addressing – More specific routes

- ISPs-R-Us has a more specific route to Organization 1



42

42

IP Addressing

- How does an ISP get block of addresses?
  - From ICANN: Internet Corporation for Assigned Names and Numbers
  - ICANN
    - allocates addresses
    - manages Domain Name Service (DNS)
    - assigns domain names, resolves disputes

RCom 21/22 - The Network Layer

U.PORTO

43

43

IP Addresses

- How does a host obtain an IP address?
  - Hard-coded by system admin in a file
    - Windows: control-panel → network and sharing center → change adapter settings properties → tcp/ip → properties
    - UNIX/Linux/BSD: `/etc/sysconfig/network-scripts/`, `/etc/network/interfaces`, `ifconfig`, `ip addr`
  - DHCP: Dynamic Host Configuration Protocol
    - dynamically get address from a server
    - "plug-and-play"

RCom 21/22 - The Network Layer

U.PORTO

44

44

RCom 21/22 - The Network Layer

U.PORTO

DHCP - Dynamic Host Configuration Protocol

- DHCP allows
  - host to dynamically obtain its IP address from network server when it joins network
  - It supports address reuse

- DHCP overview
  - host broadcasts “**`DHCP discover`**” msg
  - DHCP server responds with “**`DHCP offer`**” msg
  - host requests IP address “**`DHCP request`**” msg
  - DHCP server sends address “**`DHCP ack`**” msg

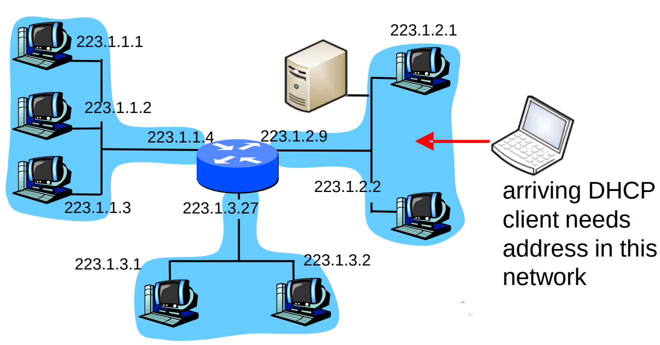
45

45

RCom 21/22 - The Network Layer

U.PORTO

DHCP - Client-server Scenario



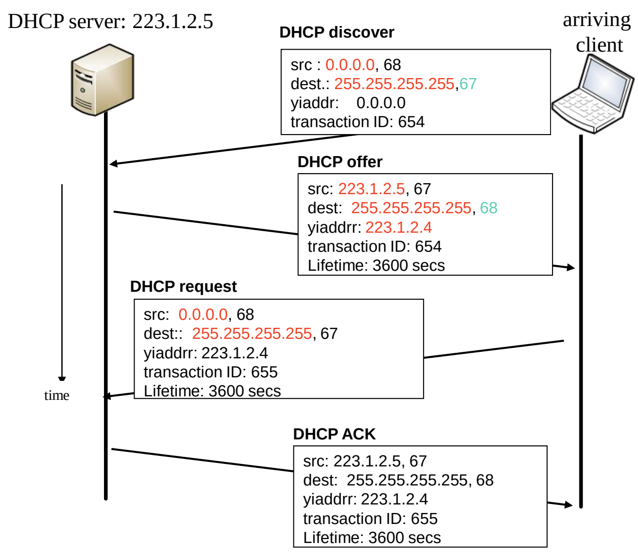
46

46

RCom 21/22 - The Network Layer

U.PORTO

DHCP Client-server



47

47

RCom 21/22 - The Network Layer

U.PORTO

- Is it sufficient for an arriving client to acquire an IP address? What other relevant information shall this client obtain in order to start working with full functionality?

P

48

48

FACULDADE DE CIÊNCIAS UNIVERSIDADE DO PORTO | U.PORTO | U.PORTO FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

RCom 21/22 - The Network Layer

Network Address Translation

RFC 3022 Traditional IP Network Address Translator (Traditional NAT)

(Network layer)

49

49

RCom 21/22 - The Network Layer

NAT: Motivation

- Shortage of IP addresses
  - Small/medium companies with ADSL connections, cable, want IPs for their machines (also domestic users).
- Uses private IP addresses
  - Not allowed in the Internet
    - 10.0.0.0/8
    - 172.16.0.0/12
    - 192.168.0.0/16

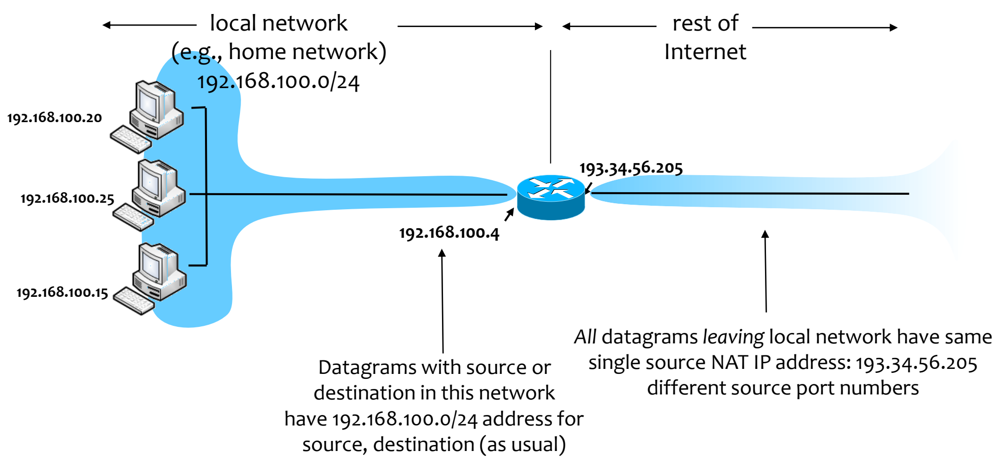
U.PORTO

50

50

RCom 21/22 - The Network Layer

NAT - Network Address Translation

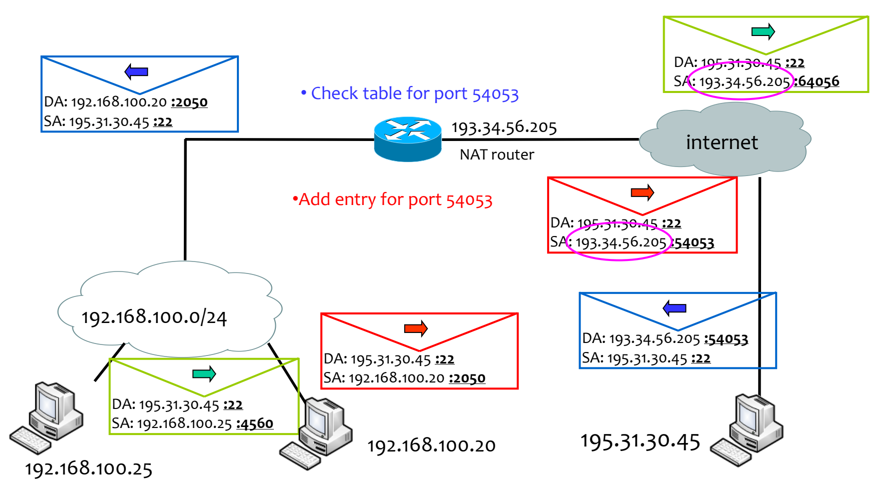


51

51

RCom 21/22 - The Network Layer

Private Address	Private Port	Outside address	Outside port	NAT port	Protocol
192.168.100.20	2050	195.31.30.45	22	54053	tcp
192.168.100.25	4560	195.31.30.45	22	64056	tcp



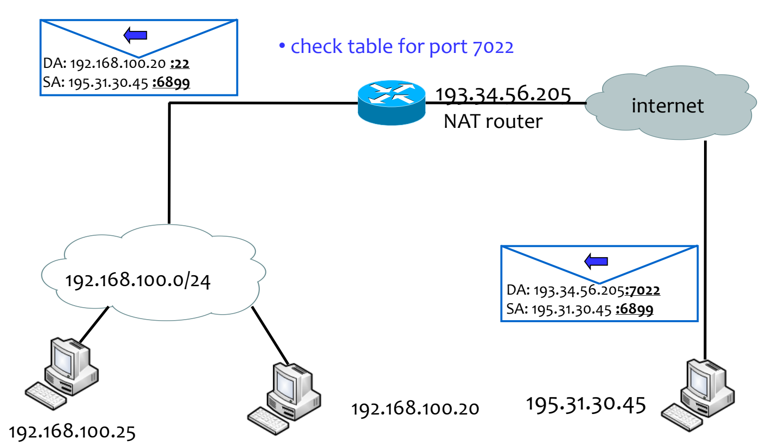
52

52

NAT traversal problem

- As inside address are private (not seen in the internet) how can an outside host connect to an internal one?
- Possible solution statically add entry to NAT table

Private address	Private port	Outside address	Outside port	NAT port	Protocol
192.168.100.20	22			7022	tcp
192.168.100.25	80			9080	tcp



FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

U.PORTO

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RCom 21/22 - The Network Layer

Internet Message Control Protocol

STD 5/RFC 792 Internet Control Message Protocol

(Network layer)

55

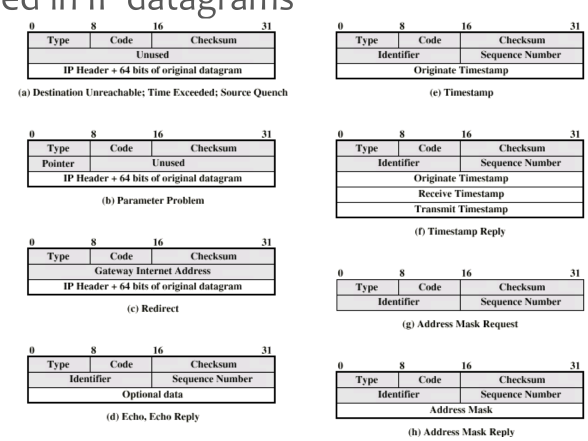
55

RCom 21/22 - The Network Layer

U.PORTO

ICMP - Internet Control Message Protocol

- Used by router or host
  - to send layer 3 error or control messages
  - to other hosts or routers
- Carried in IP datagrams



(a) Destination Unreachable; Time Exceeded; Source Quench
(b) Parameter Problem
(c) Redirect
(d) Echo, Echo Reply
(e) Timestamp
(f) Timestamp Reply
(g) Address Mask Request
(h) Address Mask Reply

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 - not used)
5	Redirect	
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

56

56

Ping – Echo Request, Echo Reply

```
machine:$ ping www.up.pt
PING www.up.pt.cdn.cloudflare.net (104.18.7.105) 56(84) bytes of data.
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=1 ttl=56 time=8.47 ms
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=2 ttl=56 time=7.06 ms
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=3 ttl=56 time=7.76 ms
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=4 ttl=56 time=7.08 ms
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=5 ttl=56 time=7.18 ms
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=6 ttl=56 time=6.99 ms
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=7 ttl=56 time=7.64 ms
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=8 ttl=56 time=6.95 ms
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=9 ttl=56 time=7.32 ms
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=10 ttl=56 time=6.75 ms
64 bytes from 104.18.7.105 (104.18.7.105): icmp_seq=11 ttl=56 time=6.86 ms
^C64 bytes from 104.18.7.105: icmp_seq=12 ttl=56 time=6.50 ms

--- www.up.pt.cdn.cloudflare.net ping statistics ---
12 packets transmitted, 12 received, 0% packet loss, time 11294ms
rtt min/avg/max/mdev = 6.504/7.214/8.471/0.506 ms
```


Traceroute and ICMP

- Source sends series of UDP segments to destination
  - first segment has TTL =1
  - second segment has TTL=2, …
  - unlikely port number
- When n^{th} datagram arrives to nth router
  - router discards datagram
  - sends to source: `ICMP TTL expired`
  - message includes: router name & IP address
- When ICMP message arrives, source calculates RTT
- Traceroute does this 3 times for each TTL
- Stop criterion: UDP segment eventually arrives at destination host
  - Destination returns ICMP `Dest port unreachable` packet
  - source stops

Traceroute and ICMP – example

```
machineFEUP$ traceroute tom.fe.up.pt
traceroute to tom.fe.up.pt (10.227.240.138), 30 hops max, 60 byte packets
 1  not.mshome.net (172.21.0.1)  0.337 ms  0.292 ms  0.270 ms
 2  172.29.0.1 (172.29.0.1)  12.832 ms  12.814 ms  12.636 ms
 3  * * *
 4  pinguim.fe.up.pt (10.227.240.138)  13.405 ms  12.966 ms  12.956 ms
```

Host did not respond with ICMP error

59

59

ICMP Redirect

- General routing principle of the TCP/IP architecture
  - routers have extensive knowledge of routes
  - hosts have minimal routing information ➔ learn routes also from ICMP redirects
- ICMP redirect message
  - Sent by router R1 to source host A
    - when R1 receives a packet from A with destination = B, and R1
      - finds that the next hop is R2 and
      - A is on-link with R2
  - R1 sends ICMP redirect to A saying next hop for destination B is R2
  - A updates its forwarding table with a host route

60

60

RCom 21/22 - The Network Layer

ICMP Redirect format

```
/                                                               /
|                IP datagram header  (prot = ICMP)              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type=5      |     code      |          checksum             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Router IP address that should be preferred           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        IP header plus 8 bytes of original datagram data       |
/                                                               /
```

61

61

RCom 21/22 - The Network Layer

ICMP Redirect Example – Routing table in host A

```
ha$ netstat -nr
Kernel IP routing table
  Destination        Gateway          Flags      Genmask         ...   Iface
 --------------- ----------------- ------- ----------------- ----- -------
  127.0.0.1       127.0.0.1         UH      255.255.255.255         lo0
  193.154.156.0   193.154.156.24    U       255.255.255.0           eth0
  0.0.0.0         193.154.156.1     UG      0.0.0.0                 eth0
```

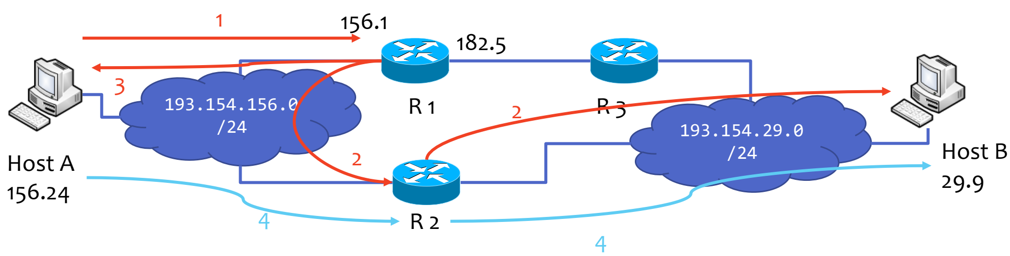
Other commands to check routing table in hosts:
- `ip route`
- `route`
  - (deprecated)

Flags:
- **U** - *route Up*
- **G** - *route to a Gateway (next hop router)*
- **H** - *route to a Host*

62

62

ICMP Redirect Example



- Host A to send packet to B
  - 193.154.156.24 → 193.154.29.9
- Mesg 3:
  - ICMP of type redir
  - Src: 193.154.156.1 → Dst: 193.154.156.24
- All messages (except 3) are
  - 193.154.156.24 → 193.154.29.9

RCom 21/22 - The Network Layer

U.PORTO

63

63

ICMP Redirect Example – Routing table in host A after mesg 3

```
ha$ netstat -nr
Kernel IP routing table
  Destination        Gateway          Flags      Genmask        ...    Iface
 --------------- ----------------- ------- ----------------- ----- -------
  127.0.0.1        127.0.0.1          UH      255.255.255.255          lo0
  193.154.29.9     193.154.156.100    UGH     255.255.255.255          eth0
  193.154.156.0    193.154.156.24     U       255.255.255.0            eth0
  0.0.0.0          193.154.156.1      UG      0.0.0.0                  eth0
```

Flags:
U - *route Up*
G - *route to a Gateway (next hop router)*
H - *route to a Host*

RCom 21/22 - The Network Layer

U.PORTO

64

64

IPv6

STD 86/RFC 8200 Internet Protocol, Version 6 (IPv6) Specification

(Network layer)

65

The Need of a New IP

- IPv4
  - Small addressing space (32 bits)
  - Non-continuous usage
  - Some solutions used to overcome these problems
    - private networks (NAT), classless networks (CDIR)
- IETF developed new IP version: IPv6
  - Same principles of IPv4
  - Many improvements
  - Header re-defined
  - First RFC 1883, December 1995

66

66

RCom 21/22 - The Network Layer

IPv6 – Improvements

- 128 bit addresses (16 octets, 8 shorts). No classes
- Better QoS support (native flow label)
- Native security functions (peer authentication, data encryption)
- Autoconfiguration (Plug-n-play)
- Routing
- Multicast

67

67

TAR 2014/15 -IPv6 intro - pbrandao

Addresses

- Represented in hexadecimal
  - Ex. `1528:8653:294c:0000:0000:90af:0900:7654`
- Zeros may be aggregated by ::
  - Ex.: `1528:8653:294c::90af:8900:7654`
  - Only one :: in an address
- Masks are the same as for IPv4 CIDR
  - Loopback address `::1 /128`
- Combining IPv6 and IPv4 addresses
  - `::ffff:5.6.7.8` (IPv4mapped address)
  - **`::5.6.7.8` (IPv4compatible address)**
  - `2002:5.6.7.8::1` (6to4 address)

128 bits → 665,570,793,348,866,943,898,599 addresses per m^2 on Earth

See RFC5952 for text representation

68

68

RCom 21/22 - The Network Layer

U.PORTO

Adresses – Link-Local, Global Unicast, Anycast, Multicast

- Link-Local
  - Used for communication between hosts in the same LAN/link
  - Address built from MAC address
  - Routers do not forward packets having Link-Local destination addresses
- Global Unicast
  - Global addresses
  - Address: network prefix + computer identifier
  - Structured prefixes
    - Network aggregation; less entries in the router forwarding tables
- Anycast
  - Group address; packet is received by any (only one) member of the group
- Multicast
  - Group address; packet received by all the members of the group

69

69

RCom 21/22 - The Network Layer

U.PORTO

Address Formats

```
|         n bits          |  m bits  |        128-n-m bits         |
+-------------------------+----------+-----------------------------+
|001 global rout prefix   | subnet ID |       interface ID         |
+-------------------------+----------+-----------------------------+
```

Global Unicast Address (2000::/3)

```
|   10     |
|  bits    |         54 bits          |          64 bits           |
+----------+--------------------------+----------------------------+
|1111111010|            0             |       interface ID         |
+----------+--------------------------+----------------------------+
```

Link-Local Unicast address (fe80::/10)

```
|                        n bits                 |   128-n bits    |
+-----------------------------------------------+-----------------+
|                    subnet prefix              | 00000000000000  |
+-----------------------------------------------+-----------------+
```

Anycast address

```
|   8     |  4 |  4 |                  112 bits                     |
+------- -+----+----+-----------------------------------------------+
|11111111|flgs|scop|                  group ID                      |
+--------+----+----+------------------------------------------------+
```

Multicast address
Scope – link, site, global, …
(ff::/8)

```
|   10     |
|  bits    |         54 bits          |          64 bits           |
+----------+--------------------------+----------------------------+
|1111111011|        subnet ID         |       interface ID         |
+----------+--------------------------+----------------------------+
```

OBSOLETE

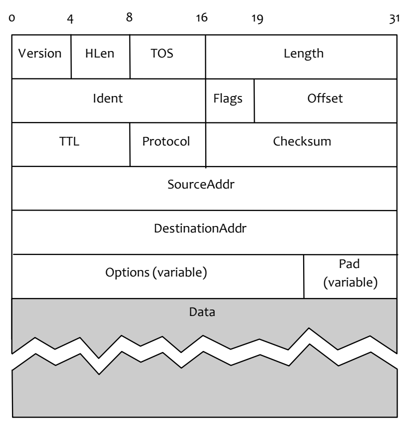
Site-Local Unicast address (fec0::/10) (not used)

OBSOLETE

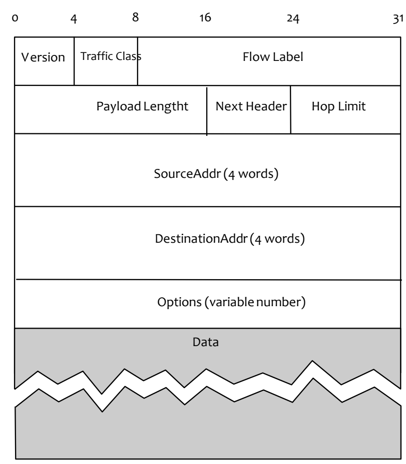
70

70

Headers IPv4 and IPv6



IPv4



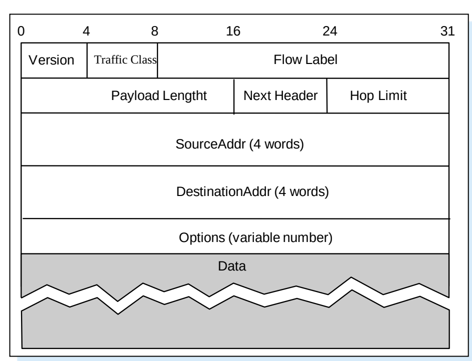
IPv6

71

71

IPv6 Header

- Flow label → identifies packet flow
  - QoS, resource reservation
  - Packets receive same service
- Payload length
  - Header not included
  - Limited to 2^{16} - 1 = 65 535 bytes, but there there's option for JumboDatagrams <= 2^{32} bytes
- Next header
  - Identifies next header/extension
- Hop limit = TTL (v4)
- Options → included as extension headers
- IPv4 checksum removed, as lower layers are responsible for verification.



72

72

TAR 2014/15 -IPv6 intro - pbrandao

Extension Headers

- Next Header – type of next header
  - Hop-by-Hop Options
    - the only header that is examined by intermediate nodes
  - Destination Options Header
    - Information for the destination node
  - Routing Header
    - List of nodes to be visited by the packet
  - Fragment Header
  - Authentication Header
  - Encrypted Security Payload Header
  - Transport layer headers

Example

IPv6 Header Next Header = Routing	Routing Header Next Header = Fragment	Fragment Header Next Header = TCP	Fragment of TCP header + data

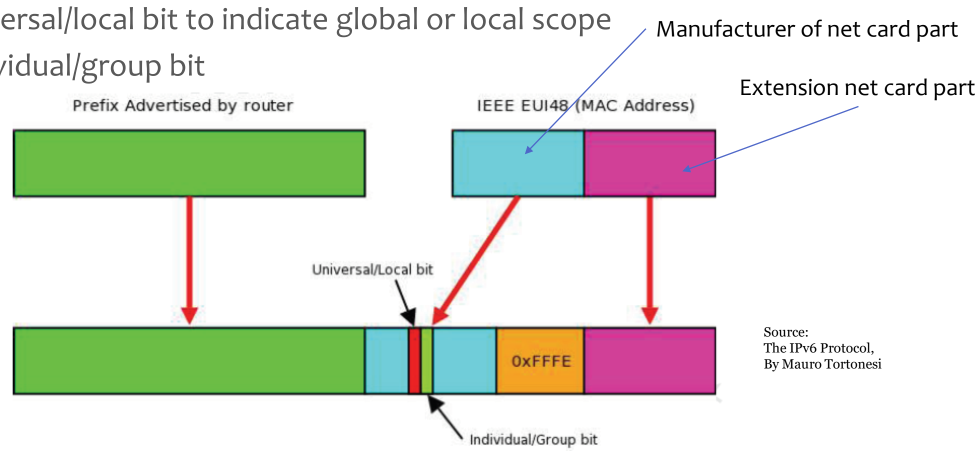
73

73

TAR 2014/15 -IPv6 intro - pbrandao

Auto Configuration Address

- IEEE-EUI64
  - universal/local bit to indicate global or local scope
  - individual/group bit



Source:
The IPv6 Protocol,
By Mauro Tortonesi

Lookup vendors by MAC address

74

74

RCom 21/22 - The Network Layer

U.PORTO

Protocol Neighbour Discovery (ND)

- RFC4861
- IPv6 node uses ND for
  - Find other nodes in the same link /LAN
  - Find a node MAC address
    - ND substitutes ARP
  - Find router(s) in its network
  - Mantaining information about neighbour nodes

- ND similar to the IPv4 functions
  - ARP IPv4
  - ICMP Router Discovery
  - ICMP Redirect

75

75

RCom 21/22 - The Network Layer

U.PORTO

ND Messages

- ICMP messages (over IP); using Link Local addresses
- Neighbour Solicitation
  - Sent by a host to obtain MAC address of a neighbour / to verify its presence
- Neighbour Advertisement
  - Answer to the request
- Router Advertisement
  - Information about the network prefix; periodic or under request
  - Sent by router to IP address Link Local multicast
- Router Solicitation: host solicits from router a Router Advertisement message
- Redirect: Used by a router to inform a host about the best route to a destination

76

76

RCom 21/22 - The Network Layer

U.PORTO

Summary

- Network layer overview
- Virtual Circuits
- Datagram Networks
- Forwarding
- ARP
- DHCP
- NAT
- ICMP
- Internet Protocol version 6

77

77

RCom 21/22 - The Network Layer

U.PORTO

Homework

1. Review slides
2. Read from Kurose & Ross
   - Chapter 4 – The Network Layer
     (this set of slides follows mainly Kurose & Ross)
3. Or, from Tanenbaum,
   - Chapter 5 – The Network Layer
4. Answer questions at Moodle

78

78

End of
Network Layer

79