



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

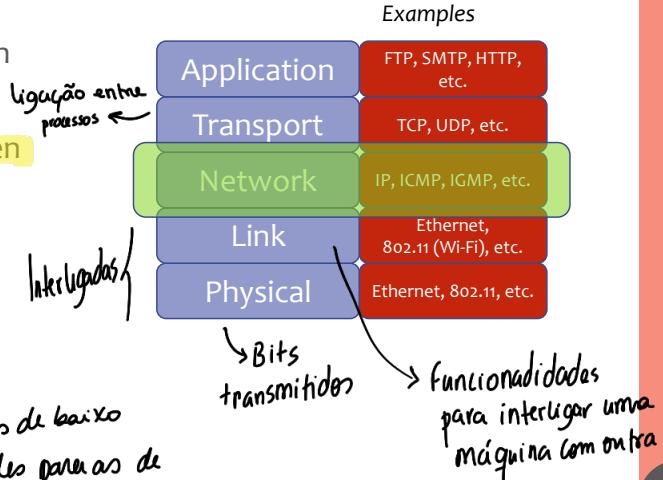
Driving questions...

- 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?

3

Internet protocol stack

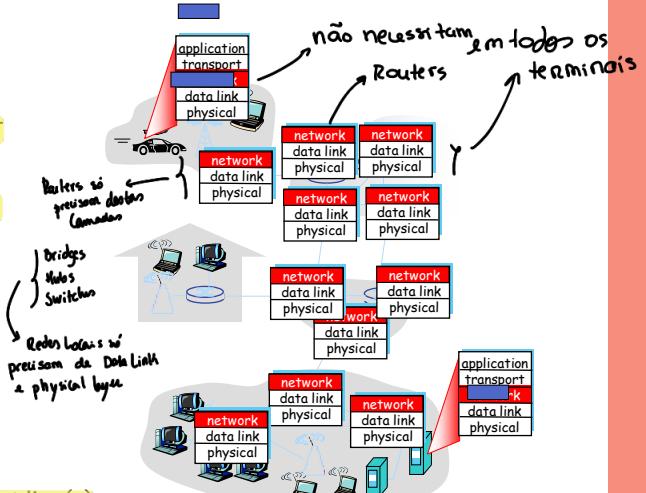
- Application: network processes
- Transport: data transfer between processes
- Network: packet routing between source and destination
- Link: data transfer between adjacent network elements
- Physical: bits on the “wire”



4

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

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

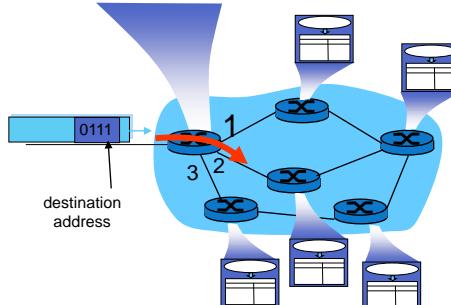
routing algorithm

local forwarding table	
header value	output link
0100	3
0101	2
0111	2
1001	1

tabela de encaminhamento para ver para que interface vai enviar

- Forwarding
 - router forwards packet from input port to output port
 - Analogy: process of getting through single interchange

↳ Numa viagem, os vários meios de transporte que temos de "apunhar" para ir da origem ao destino



6

Virtual Circuits and Datagram Networks

(Network layer)

estabelecimento de circuitos do inicio ate ao fim

Network Layer – Connection and Connectionless Service

① Cada um dos pontos intermédios guarda um estado associado à ligação para saber qual é o salto seguinte a tomar

- Services provided by network layer

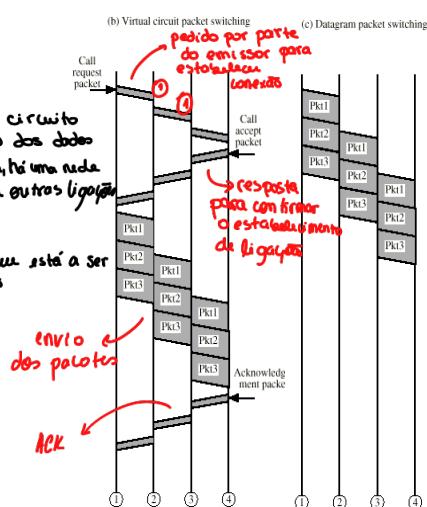
- Virtual Circuit network

→ connection-oriented service
 ↳ Estabelecimento de um circuito antes da transmissão dos dados.
 Virtual porque, na prática, há uma rede por baixo que suporta outras ligações.

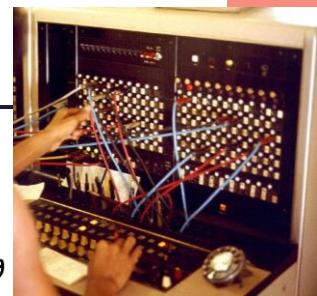
- Datagram network

→ connectionless service

↳ Não há estabelecimento de um circuito prévio. Cada um dos pontos intermédios trata de reencaminhar o pacote recebido.
 ↳ Independentemente do anterior



Virtual Circuit (VC)

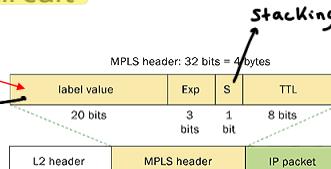
Image from Joseph A. Carr, in
Wikipedia Telephone Exchange

- Phases
 - circuit establishment → data transference → circuit termination

- Packet carries identifier of Virtual Circuit

Label permite ao router saber para que porta é que tem que enviar o pacote realizado. Assim, o router mantém um estado associado à tabela.

Identificação do que é o circuito virtual



Exp=experimental (used for CoS mapping)
S=Stacking bit
TTL=Time to Live

- 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

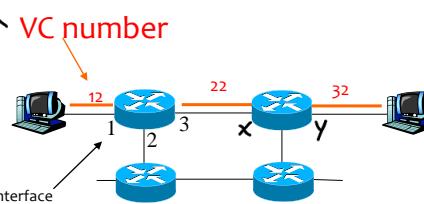
9

VC - Forwarding Table

Labels são estabelecidas quando se estabelece a conexão

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
...



label para o salto seguinte

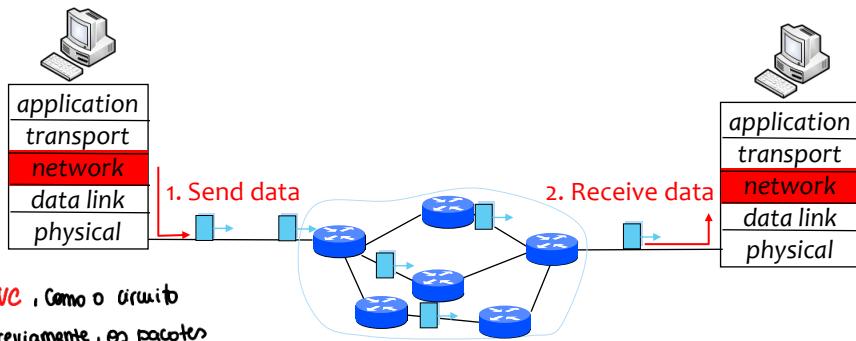
Routers maintain connection state information!

Tabela do router 1. No router 3: I I IVC OI OVC
 x 22 y 32

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



• Enquanto no VC, como o circuito foi definido previamente, os pacotes seguem todos o mesmo caminho, nos datagrams, pacotes entre a mesma source e destination podem seguir caminhos diferentes

→ Não há garantias da ordem com que os pacotes chegam

11

Forwarding Table

→ Especificações em relação ao endereço do destino

- 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

12

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

13

Longest Prefix Matching

- Which interface?
- DA: **11001000 00010111 00010** 110 10100001
 - → Itf: 0 *Output link interface $\Rightarrow \emptyset$*
- DA: **11001000 00010111 00011** 000 10101010
 - Can be itf 1 or 2
 - **Longest prefix 1** *Output link interface $\Rightarrow 1$*

Tabela de forwarding

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

*Recebendo um IP, Vão
reencaminhar por match
com os prefixes na tabela*

① e ② são praticamente iguais, mas 2 é + específico

O match é o + longo possível

11001000 00011111 00011000 10100001 $\Rightarrow 3$

*Não há match → router não sabe para onde
enviar, reencaminha para outro router*

14

Virtual-Circuit versus Datagram Networks

Identificar para onde vão

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

Identifica o circuito

Todos os pacotes pelo mesmo circuito

O espaço guardado lá é perdido a todos os VC's → RESTABELECER NOVO VC

O espaço alocado vai ser sempre utilizado → Só é usado quando necessário → por isso, pode alocar espaço parcialmente

Datagram tenta de utilizar o que sobra

Como o espaço é alocado previamente, por ex sabem de quanto espaço vai ser necessário, não vão ocorrer coisas em que o espaço é alocado e não é utilizado

Como é pré-estabelecido, podemos estabelecer apenas se houverem recursos para tratar do envio dos pacotes → Evita congestionamento

Se caminho 2 não tiver recursos suficientes, restabelece-se outro

15



Internet Protocol

[RFC 791/STD 5](#)

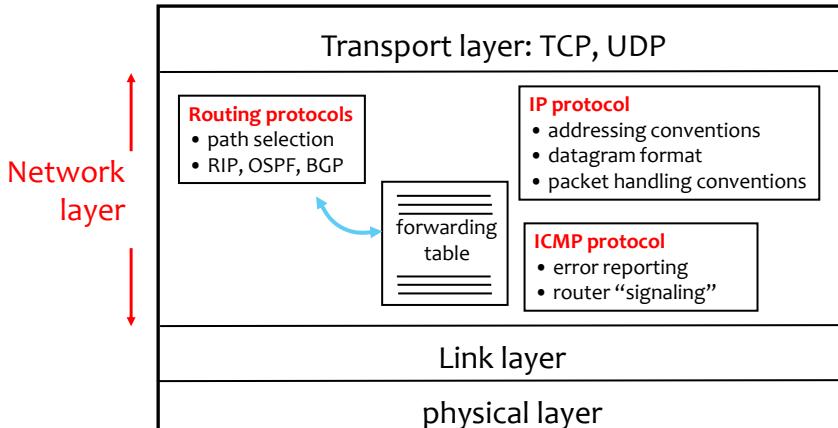
[IP RFCs](#)

(Network layer)

16

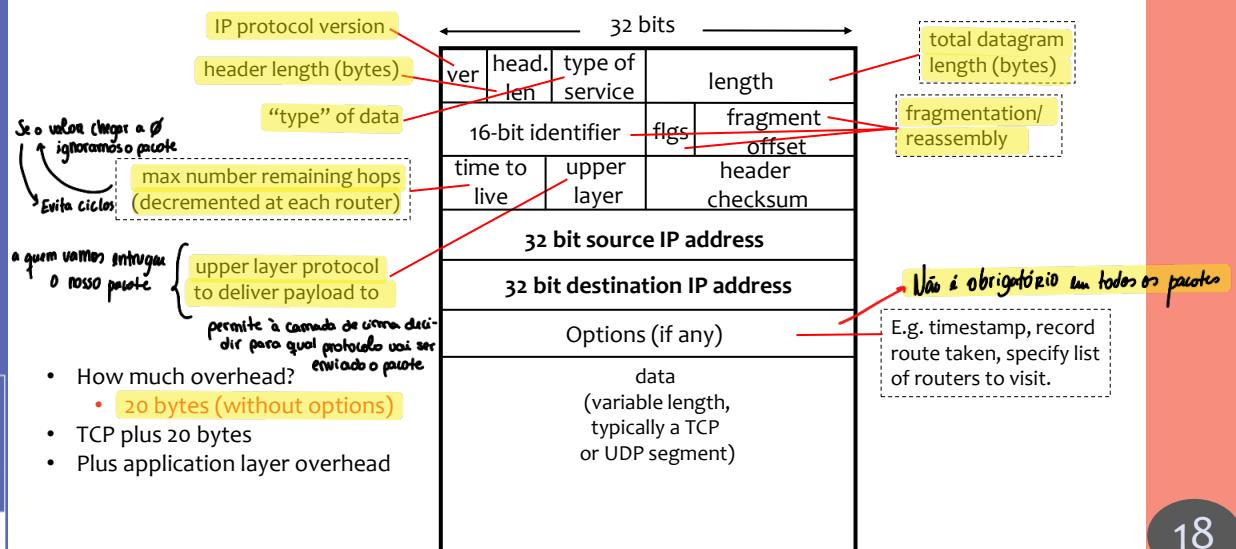
The Internet Network layer

- Host, router network layer functions



17

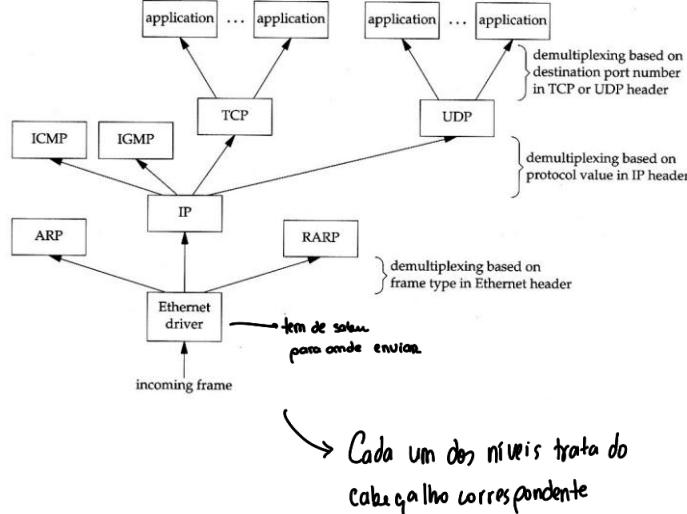
IP Datagram Format



18

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

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

checksum tem de bate certo

```
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);
}
```

trata do carry

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

20

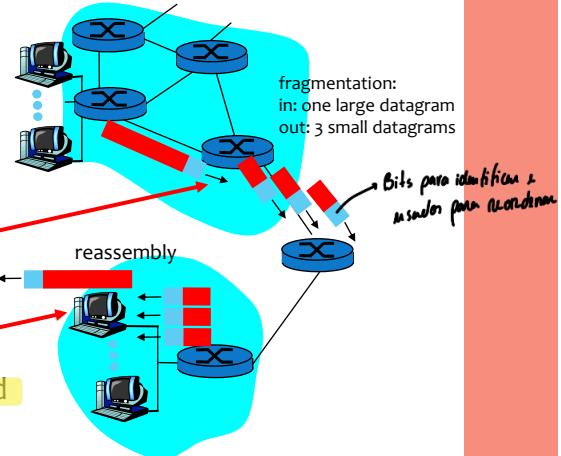
- One's complement sum
 - Mod-2 addition with carry-out
 - Carry-out in the most-significant-bit is added to the least-significant bit
 - Get one's complement of “one's complement sum”

→ Pacote a transmitir é maior do que aquilo que podemos transmitir

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

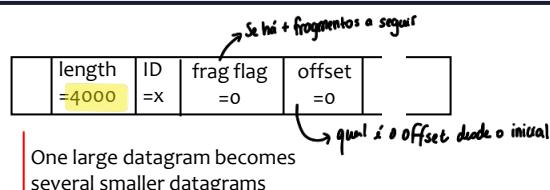
Example

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

1480 bytes in data field

$$\text{offset} = \frac{1480}{8}$$

Bytes enviados no 1'



length =1500	ID =x	frag flag =1	offset =0	
length =1500	ID =x	frag flag =1	offset =185	

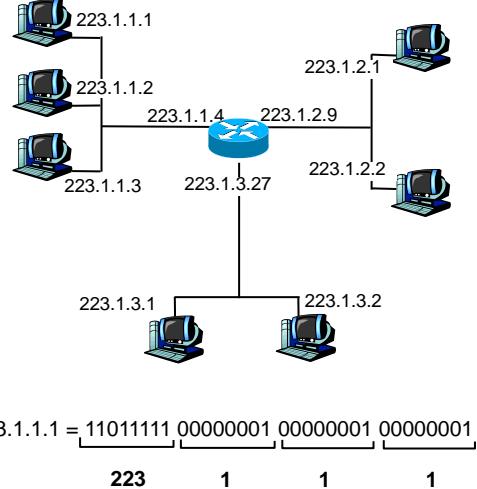
length =1040	ID =x	frag flag =0	offset =370	
-----------------	----------	-----------------	----------------	--

Permite reagregá-los no final

22

IP Addressing - Introduction

- IP address
 - 32-bit identifier for host/router interface
 - Se tiver + do que 1 interface tem + do que 1 IP
 - Interface wifi
 - Interface Ethernet
 - 2 IP's ≠
 - + do que 1 place de rede / + do que 1 IP
- Interface
 - connection between host/router and physical link
 - Routers have multiple interfaces
 - IP addresses associated with interface

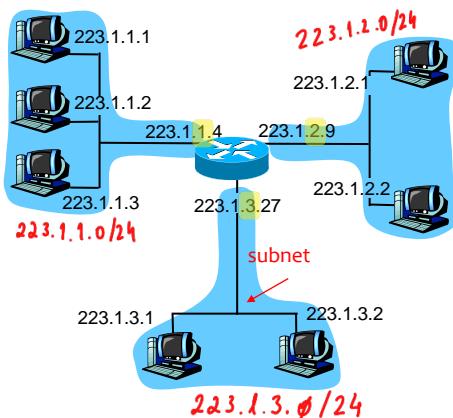


23

Subnets

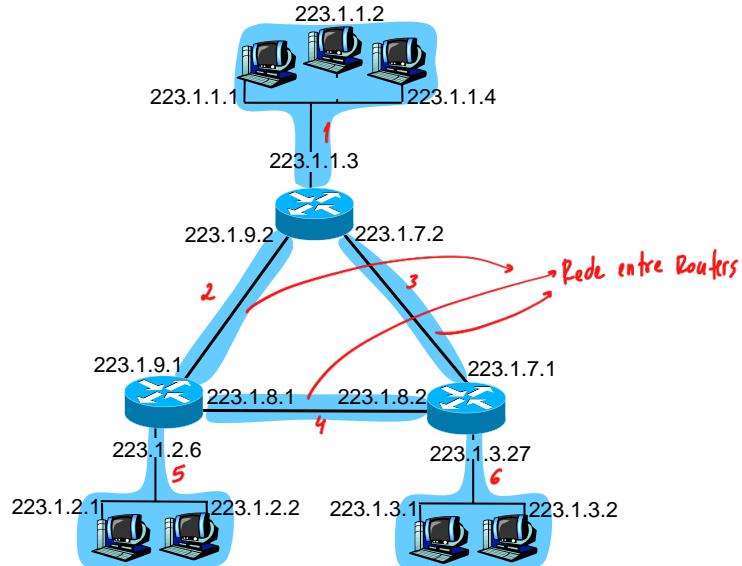
- IP address
 - subnet part → high order bits
 - host part → low order bits
- Subnet → set of interfaces
 - with same subnet part of IP address
 - can reach each other without router intervention

Network consisting of 3 subnets



24

6 Subnets



25

IP Addressing - CIDR

CIDR: Classless InterDomain Routing ([RFC4632/BCP122*](#))

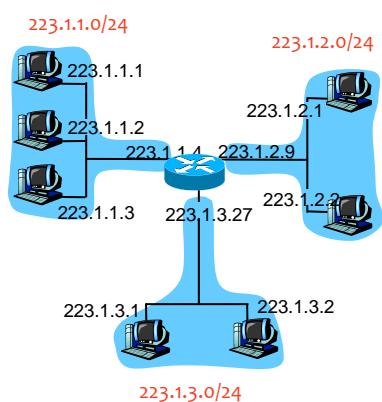
- subnet portion of address has arbitrary length
- address format → a.b.c.d/x
 - where x is # bits in subnet portion of address

Sub network part Node part

 11001000 00010111 00010000 00000000

200.23.16.0/23

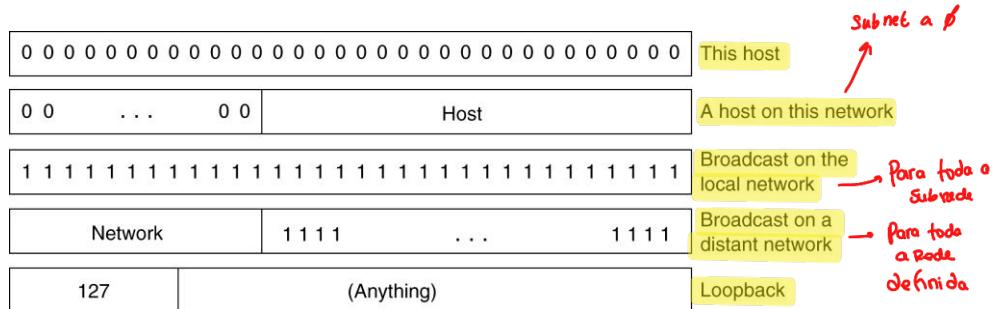
23 bits para identificar
a subrede
9 para identificar o host



* BCP – Best Current Practice

26

Special IP Addresses



27

27

Forming Sub-Networks

O resto da net só precisa de
saber que são Network 192.
necessários 24 bits

Router 1 vai tratar de repartir dos 8 subredes depois

LAN X
Net ID/Subnet ID: 192.228.17.32
Subnet number: 1

A
IP Address: 192.228.17.33
Host number: 1

B
IP Address: 192.228.17.57
Host number: 25

LAN Y
Net ID/Subnet ID: 192.228.17.64
Subnet number: 2

C
IP Address: 192.228.17.65
Host number: 1

R2
IP Address: 192.228.17.64
Host number: 2

LAN Z
Net ID/Subnet ID: 192.228.17.96
Subnet number: 3

D
IP Address: 192.228.17.97
Host number: 1

e vista externo, 24 bits para a sub-rede

máscaras internas de 27 bits

Subnetwork mask – 27 bits

subnetid – 3 bits (8 subredes)

hostid – 5 bits

30 hosts per subnet supported

all 0 – identifies subnet

all 1 – broadcast address

11000000 11100100 00010001 01100000

192.228.17.96/27

↳ + 3 bits para definir 8 subredes

8

$2^3 = 8$

254 possibilidades

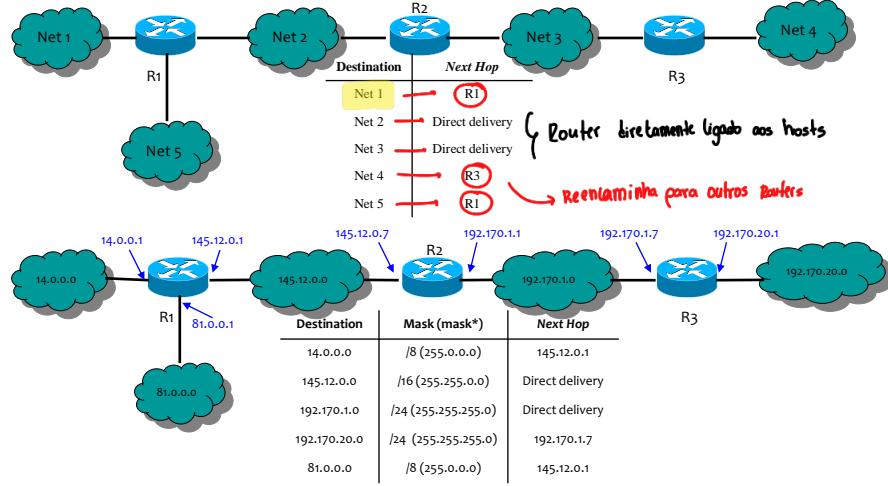
Não pode ser usado como endereço do host

Example of subnetworks

- 192.228.17.0/27 (.0000000) → Identifica a rede
192.228.17.32/27 (.0010000)
192.228.17.64/27 (.0100000)
192.228.17.96/27 (.0110000)
....
192.228.17.224/27 (.1110000)

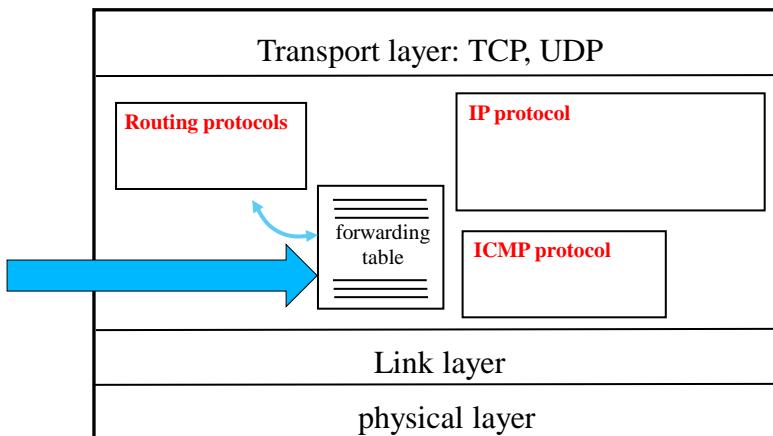
28

Forwarding Table at R2



29

Forwarding Table use



30

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

P

31

IP Forwarding Function

- Forwarding table has entries in format
`<networkAddress/mask, port>` → para onde o router vai enviar o pacote
- 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
```

Parte do endereço IP da rede
Parte dos hosts
 - Select the port with the largest mask → most specific route

{ Para cada uma das entradas, vai verificar se há match com o novo endereço

32

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
 - o → {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
 - o 3 largest mask, 18 bits

Match mais específico, mask mais longa
 - Broadcast porta 1 → 128.16.255.255
 - Broadcast porta 3 → 128.16.255.255
 - Broadcast porta 5 → 128.255.255.255
- 3 entradas de entrada
 mask 16 bits
 3 3 Port
 8 8 2 mask 18 bits
 mask 8 bits
 8 1 Match
 3 5 Match → Mask
 5 Match
- Tlatch
 3. 128.32.1100 0.0
 A. 128.32.1100 0011 .0
- One online [IP Subnet Calculator](#)



Address Resolution Protocol

RFC 826/STD 37

(Network layer)

→ Traduzir endereço de rede para endereço MAC, para a camada de ligação lógica que está por baixo

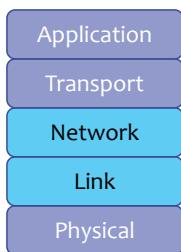
Serve para comunicar diretamente localmente (na mesma rede com ou sem necessidade usar o Router)

Serve para comunicar com um IP de outra rede, onde já é necessário saber qual o endereço MAC do router e este encaminhar para a outra rede (para outro router)

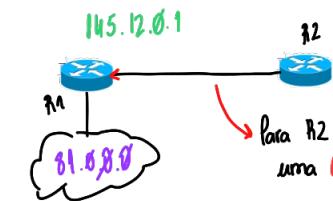
What is ARP needed for?

- 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

Ainda não sabemos qual é o endereço MAC para o qual teremos de enviar a trama no nível abaixo



Packete Vem de Linha



é necessário descobrir qual o endereço MAC associado a endereço IP como qual queremos comunicar

Forwarding Table :

Destination	Mask	Next Hop
81.0.0.0	/8 (255.0.0.0)	115.12.0.1



Switch que estabelece ligação física entre as máquinas de uma sub-rede

- Utiliza endereços da Camada de ligação para estabelecer estas ligações

35

Vê que tem que fazer o envio para um endereço que não é o destino final a antão:
Verifica quem tem este endereço MAC para estabelecer ligação com ele e depois enviar um pacote que tem este destino

ARP – Address Resolution Protocol

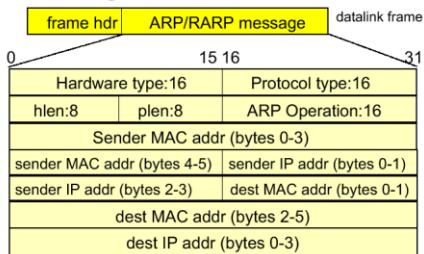
- A network interface has qualche interface de rede
 - one MAC address
 - one (or more) IP address(es)
- ARP: Address Resolution Protocol**
 - Protocol used to obtain the MAC address associated to a given IP address
- RARP – Reverse ARP**
 - Protocol used to obtain the IP address associated to a MAC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

permite obter o endereço MAC de um determinado IP

Inverso → permite obter um endereço IP através de um endereço MAC

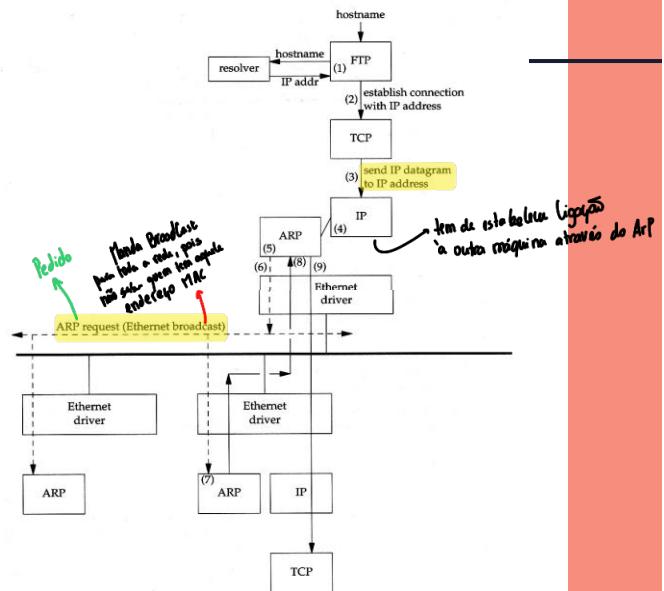
36

ARP Example



- hardware type : Ethernet=1 ARCNET=7, locltalk=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

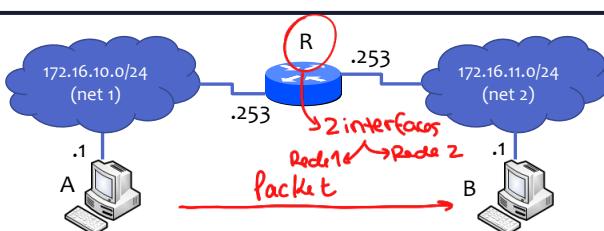
Para identificar o pedido



37

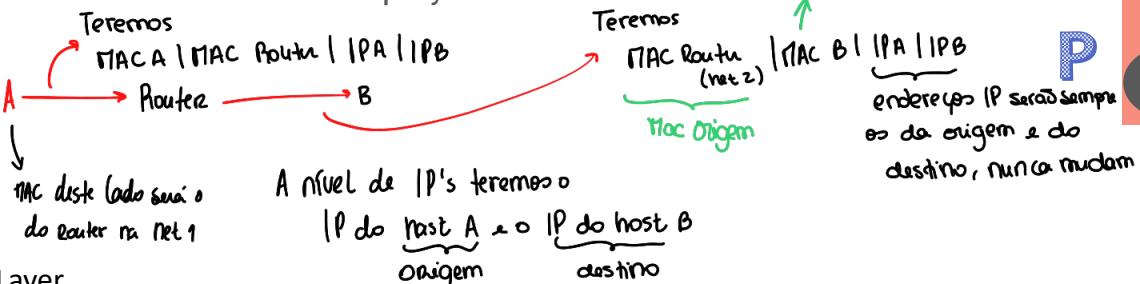
ARP/IP addresses

| Mac Origem | Mac Destino | IP origem | IP destino



- 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? → Considerando que o ARP já funcionou e uma máquina sabe os endereços Mac e etc da outra.

- What roles does ARP play in this scenario?



P

38

Obtaining IP Addresses

[RFC 2131 Dynamic Host Configuration Protocol](#)

(Network layer)

39

39

How to Obtain IP Addresses

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

— Não podem haver IPs repetidos se não o roteador não saberá para onde encaminhar o pacote

ISP's block	<u>11001000</u> <u>00010111</u> <u>00010000</u> <u>00000000</u>	200.23.16.0/20
Organization 0	<u>11001000</u> <u>00010111</u> <u>0001<u>000</u></u> <u>00000000</u>	200.23.16.0/23
Organization 1	<u>11001000</u> <u>00010111</u> <u>0001<u>001</u>0</u> <u>00000000</u>	200.23.18.0/23
Organization 2	<u>11001000</u> <u>00010111</u> <u>0001<u>010</u>0</u> <u>00000000</u>	200.23.20.0/23
...
Organization 7	<u>11001000</u> <u>00010111</u> <u>0001<u>111</u>0</u> <u>00000000</u>	200.23.30.0/23

Aumenta a Mask

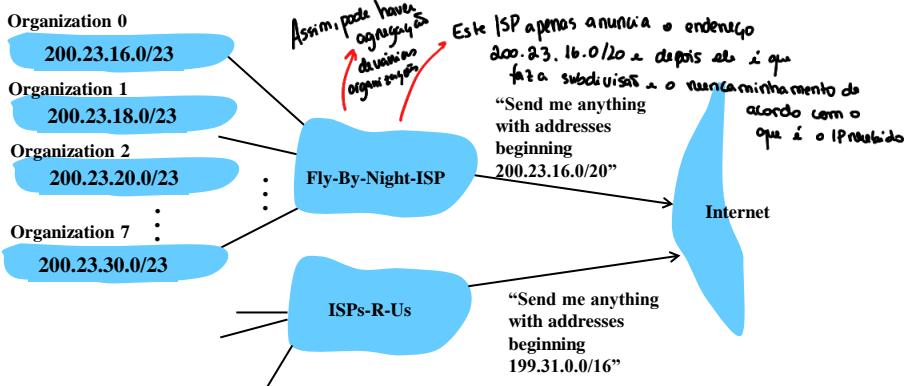
40

40

Hierarchical Addressing - Route Aggregation

- Hierarchical addressing

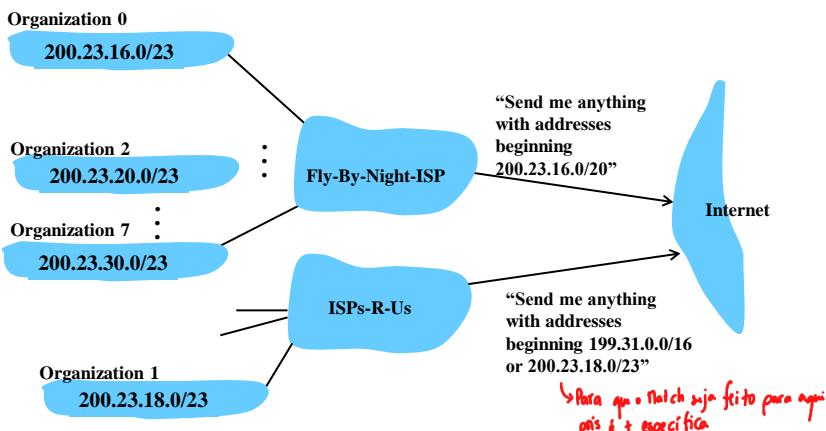
allows efficient advertisement of routing information



41

Hierarchical Addressing – More specific routes

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



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

entidade que faz alocações de endereços IP
a quem temos de fazer os pedidos, etc...

IP Addresses

- How does a host obtain an IP address?

- especificar qual o endereço e qual anexe em que a máquina está

- 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”

Descoberta através do uso de um servidor

alocação estática, se o host estiver sempre na mesma rede não há problema. se este mudar de rede deixa de funcionar → tendo em conta que não existem IPs repetidos

Se não houverem mudanças de rede, a máquina estiver sempre ligada à mesma rede, esta forma de obtenção de endereço é preferível, pois deixa de ser necessário ter um protocolo a conversar para obter estas informações

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 possui uma **pool de endereços disponíveis** dentro da gama que está a gerir da rede em que está

- **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

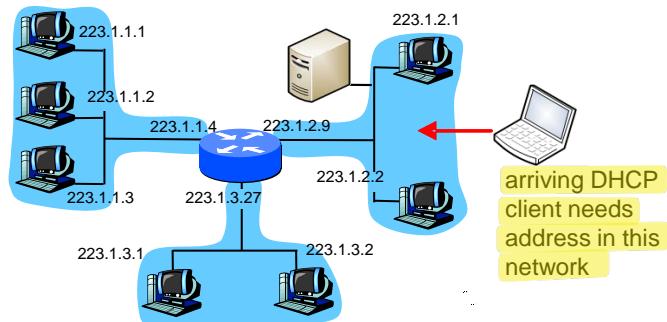
→ Do lado do servidor DHCP é possível colocar estaticamente a alocação. Sabendo endereço MAC da interface, coloca-o no servidor numa relação endereço MAC - endereço IP. aquele endereço MAC terá sempre aquele endereço IP associado (sempre que é pedido, é isso que é retribuído).

45

essa endereço deixa de estar disponível para outras máquinas

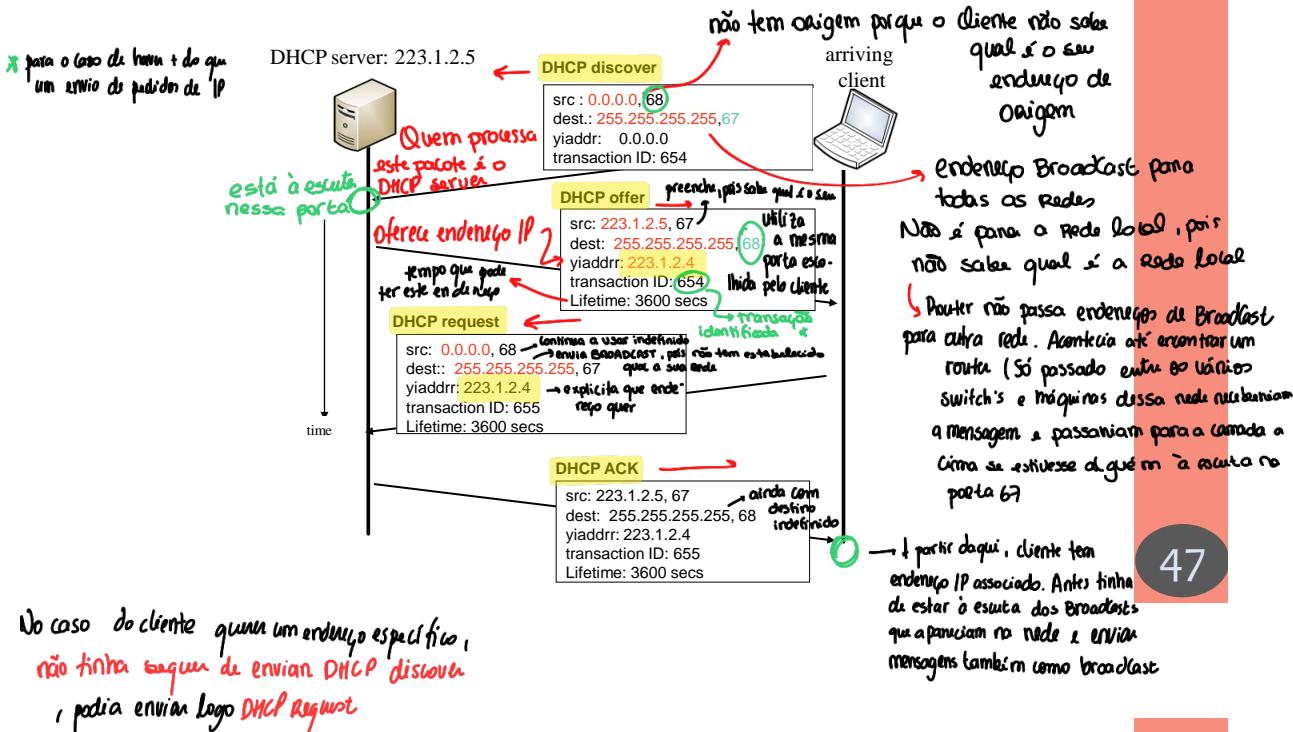
• Pode acontecer que o servidor DHCP só atribua um endereço IP a endereços MAC que foram estaticamente alocados

DHCP - Client-server Scenario



46

DHCP Client-server



47

No caso do cliente querer um endereço específico, não tinha saída de enviar DHCP discover, podia enviar logo DHCP request

- 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?

Vai → Tem de saber a máscara de rede, para saber qual a parte do endereço de rede e qual a parte do endereço do host → para perceber em que rede é que está localizado.

→ Tem de saber a **default gateway** → entrada genérica → se não conhece nenhuma outra entrada, envia para essa por omissão. Se não conhece destino específico que queria, envia para alguma máquina que saiba depois tratar desse endereço de destino. P

48

→ Servidor DNS → A quem perguntar para fazer a resolução de um nome para um endereço IP

48

Network Address Translation

[RFC 3022 Traditional IP Network Address Translator \(Traditional NAT\)](#)

(Network layer)

49

49

Network Address translation

NAT: Motivation

- Shortage of IP addresses

- Small/medium companies with ADSL connections, cable, want IPs for their machines (also domestic users).

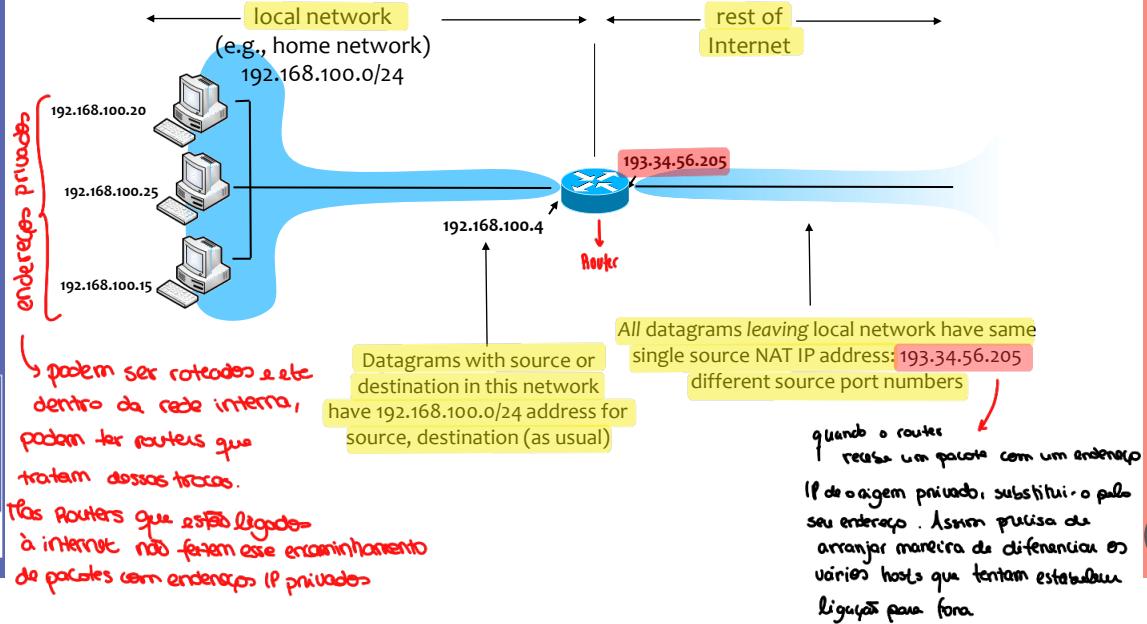
- Uses private IP addresses

- Not allowed in the Internet

- 10.0.0.0/8
- 172.16.0.0/12
- 192.168.0.0/16

endereços privados utilizados apenas em redes privadas e que não podem aparecer no core da internet / no routing da internet.
Se algum roteador os vir na internet não fará o forwarding e vai ignorar o pacote.

NAT - Network Address Translation



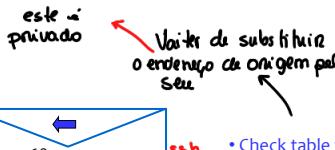
Desde que hajam portas livres no router, vai ser sempre possível estabelecer ligação com o exterior à rede interna.

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

Para depois saber para onde tem de enviar

Acrescenta nova informação e escolhe uma nova porta para este mapeamento

④ Router vai escolher uma porta para depois receber os dados

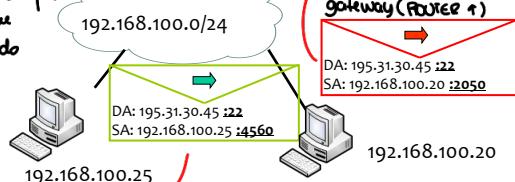


endereço IP rotável não privado que poderá ser utilizado

DA: 195.31.30.45:22
SA: 193.34.56.205:64056

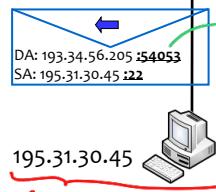
internet

Router vai, através da tabela de endereços, saber qual o endereço destino e substituir na mensagem recebida, encaminhando para a máquina que iniúlo o pedido



Com o seu endereço IP e com a porta que o Router atribuiu para a ligação

DA: 193.34.56.205:54053
SA: 195.31.30.45:22



A resposta é enviada para o remetente da outra mensagem

195.31.30.45
Comunicar com esta máquina

No ponto de vista desta máquina, sempre o mesmo endereço que está a efectuar os pedidos

Mensagem de outra máquina para a mesma (mesmo IP) e para o mesmo serviço (mesma porta)

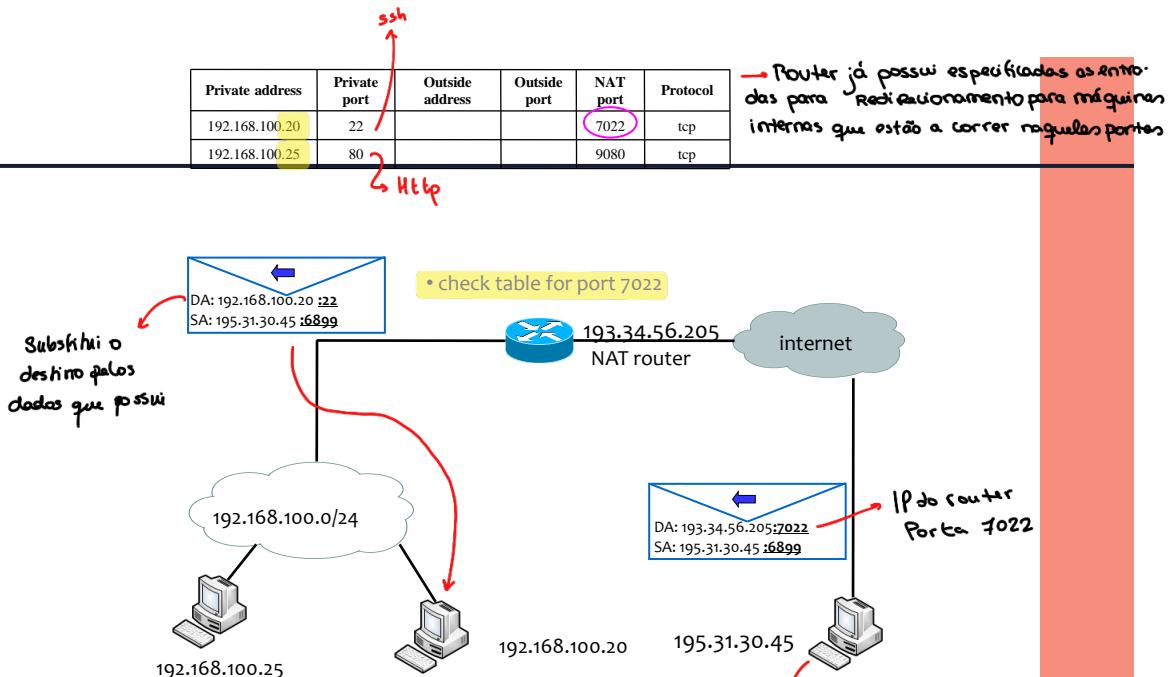
- A máquina pensa ter o seu IP
- utiliza-o na rede privada
- Destino: Outra máquina
- já possui as portas. Vai estar a escuta nessa porta por uma resposta
- Do outro lado está um programa a escuta na porta 22 para establecerem ligação

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

53



54

Desvantagens:

NAT não possui IPs para tudo
então tem de fazer modificações
dos dados enviados

Internet Message Control Protocol

[STD 5/RFC 792 Internet Control Message Protocol](#)

- Permite passar mensagens de controlo
- indicação de erros
 - indicação de redirecionamentos
 - indicação de informações para os hosts ou para os roteadores

(Network layer)

55

55

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

É acima do IP
+ utiliza o IP para comunicar

Quando não conseguimos chegar a um host porque expirou o tempo

Porta no destino não está aberta, não tem nenhum serviço a correr

(a) Destination Unreachable; Time Exceeded; Source Quench

Type	8	16	31
Identifier	Unused		
IP Header + 64 bits of original datagram			

Type	8	16	31
Identifier	Unused		
IP Header + 64 bits of original datagram			

(b) Parameter Problem

(e) Timestamp

Type	8	16	31
Identifier	Sequence Number		
Originate Timestamp			

(f) Timestamp Reply

Type	8	16	31
Identifier	Sequence Number		
Receive Timestamp			
Transmit Timestamp			

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	

(c) Redirect

Type	8	16	31
Identifier	Unused		
Gateway Internet Address			
IP Header + 64 bits of original datagram			

(g) Address Mask Request

Type	8	16	31
Identifier	Unused		
Address Mask			

(d) Echo, Echo Reply

Type	8	16	31
Identifier	Unused		
Sequence Number			
Optional data			

(h) Address Mask Reply

Type	8	16	31
Identifier	Unused		
Sequence Number			
Address Mask			

Necessário para saber qual a relação do que recebemos com aquilo que enviamos

56

56

Ping – Echo Request, Echo Reply

esta máquina envia echo requests e recebe echo replies do destinatário

```
machine:$ ping www.up.pt → destino
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
```

time to leave → número de saltos observados na rede

tempo para a mensagem ser enviada + recebida de um lado e do outro

term de estar na mensagem ICMP para saber do que está a resposta

Números sequência

--- 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

ping → resumo do que foram os pacotes enviados e recebidos

apenas 12 pacotes enviados (não se perdeu nenhuma)

12 pacotes recebidos

11294 ms → tempo total que demorou

57

Cada vez que um router faz um encaminhamento de um pacote, diminui -lhe o TTL
Se TTL chegar a 0, pacote é descartado*

Traceroute and ICMP

Qual o caminho entre 2 máquinas

Usa mensagens ICMP para ver quais os routers que estão entre uma máquina e outra

- 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
- pode ainda ser resolvido pelo DNS

When ICMP message arrives, source calculates RTT

Faz uma média.

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

utiliza a porta para saber que chegou. Se houver algo à esquerda não há resposta

* - há um aviso para trás a dizer que TTL expirou, e então o tempo de vida deste pacote expirou

Lá além disso, envia pacotes UDP e não ICMP, para poder distinguir o destino final. Pode como destino uma porta UDP, algo que não era expectável de receber pacotes UDP naquela porta.

Traceroute começa com um TTL baixo e vai aumentando até chegar ao destino (considerando que enquanto não chegar, vai enviando mensagens ICMP de erro)

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 * * * TTL = 3
 4 pinguim.fe.up.pt (10.227.240.138) 13.405 ms 12.966 ms 12.956 ms → respondeu a dizer
 máquinas destino — nome ≠ mesmo IP que tinha a porta fechada
 1 → endereço IP da origem Host did not respond with ICMP error → 3* → 3 pacotes enviados pelo traceroute sem resposta associada
 ↴ A máquina de destino pode não enviar ICMP port un reachable → traceroute não sabe que terminou (Chega ao limite e não é possível chegar a qualquer conclusão)
 ↴ Router pode estar configurado para não enviar informações para trás sobre o que é a sua rede

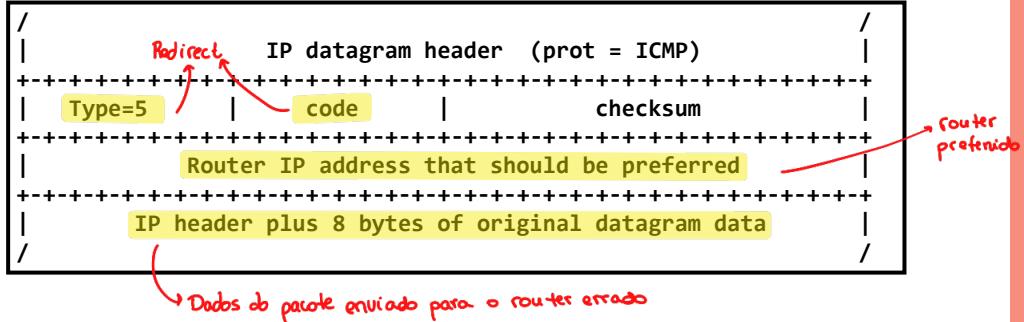
59

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

ICMP Redirect format



61

ICMP Redirect Example – Routing table in host A

→ Tabela de routing do host A → Para enviar para determinada Rede (Destination) e a máscara (GenMask) qual é a interface que deverá utilizar (Iface) e qual é a máquina para a qual tem de enviar o pacote (Gateway)

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

Rede por omisión
Router 1

Other commands to check routing table in hosts:

- ♦ ip route
- ♦ route
♦(deprecated)

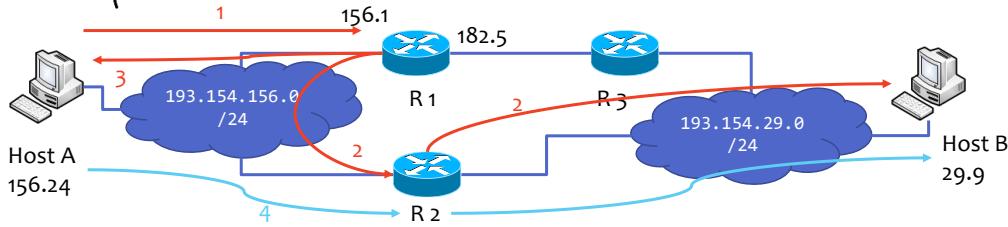
Flags:

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

62

ICMP Redirect Example

Host A quer se ligar ao Host B — tem como identificação da rota para chegar a B o caminho 1 — Caminho longo, ainda passa pelo Router R3
 R1 sabe que o melhor caminho para chegar a B é por R2 — Informa A do melhor caminho



- Host A to send packet to B
 - 193.154.156.24 → 193.154.29.9
- Mesg 3: **ICMP Redirect**
 - ICMP of type redir
 - Src: 193.154.156.1 → Dst: 193.154.156.24
 - Router 1**
 - Host A**

- All messages (except 3) are

- 193.154.156.24 → 193.154.29.9
 - Host A**
 - Host B**

- 1 → Host A envia pacote para R1 com intuito de enviar para host B
- 2 → Router 1 não quer que o pacote seja enviado por ali e para que o pacote não seja perdido reencaminha-o para o R2 que segue o caminho até ao host B
- 3 → Router 1 informa host A para Redirecionar a sua rota para o outro router. Host A atualiza a sua tabela de routing para enviar para o outro endereço de rede.

63

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

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

localhost

Route 2

Para um Host Único

loopback

Máscara de 32 bits, apenas serve para aquele endereço

Host A atualiza a sua tabela de endereços com a nova entrada para o host B que deverá ser por outra Gateway (Router 2)

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

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 → *Como há poucos endereços é mais difícil garantir continuidade naquilo que são as gamas de endereços em utilização*
 - Some solutions used to overcome these problems
 - private networks (NAT), classless networks (CIDR)
- IETF developed new IP version: IPv6
 - Same principles of IPv4
 - Many improvements
 - Header re-defined
 - First [RFC 1883, December 1995](#)

IPv6 – Improvements

- 128 bit addresses (16 octets, 8 shorts). No classes → IPv4 ten classes
- Better QoS support (native flow label)
- Native security functions (peer authentication, data encryption)
 - Dados pelo IPSEC → opcional no IPv4
obrigatório no IPv6
- Autoconfiguration (Plug-n-play)
- Routing
- Multicast → envio, não para todos (como o broadcast), mas para um grupo (várias máquinas) → endereço IP associado identifica várias máquinas ⇒ máquinas interessadas em pertencer a esse grupo só recebem os pacotes enviados para esse endereço

67

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 → 128 bits associados à parte da identificação da rede que é única para o host
 - 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)
- See [RFC5952](#) for text representation
- 

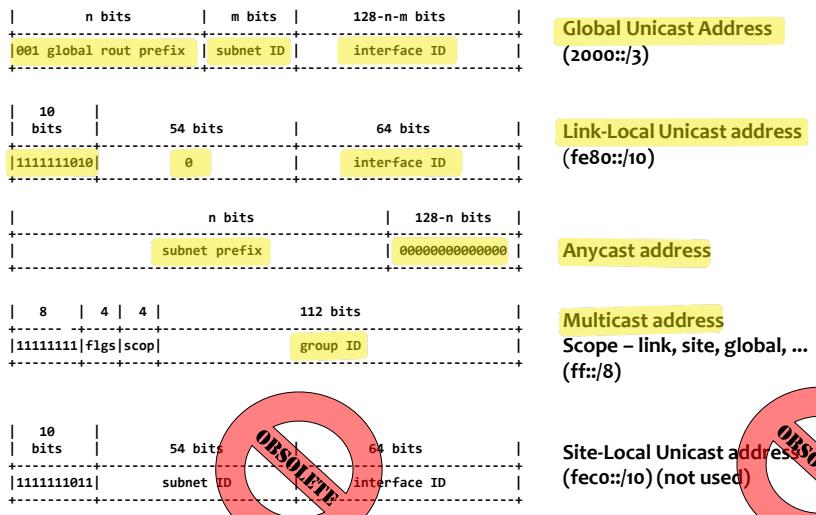
68

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

- **Link-Local** ↗ utilizado numa rede 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** ↗ utilizado na internet
 - Global addresses
 - Address: network prefix + computer identifier
 - Structured prefixes
 - Network aggregation; less entries in the router forwarding tables
- **Anycast** → enviar para um determinado endereço. Podem estar + do que 1 roteador à escuta desse endereço. O primeiro a receber, é onde o mensagem fica.
 - Group address; packet is received by any (only one) member of the group
- **Multicast** → tenta chegar a todo a gente à escuta
 - Group address; packet received by all the members of the group

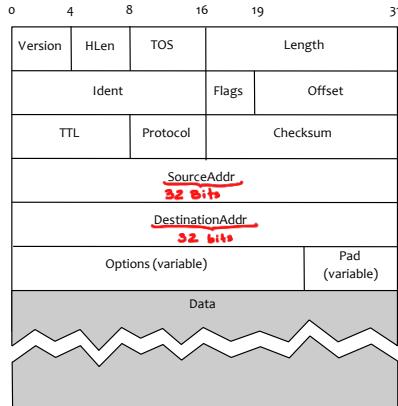
69

Address Formats

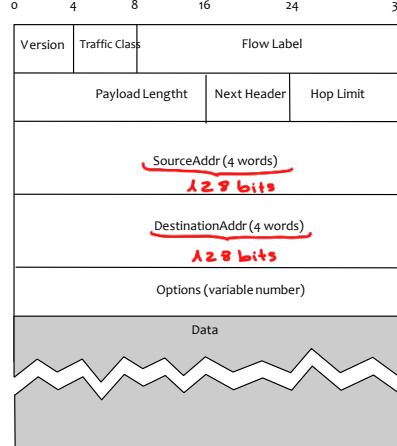


70

Headers IPv4 and IPv6



IPv4



IPv6

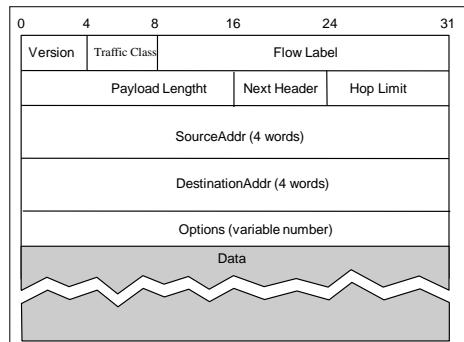
→ é maior
Menos headers que IPv4

71

IPv6 Header

Permite identificar um fluxo
→ Permite ao router ter a noção do fluxo

- **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's option for JumboDatagrams $\leq 2^{32}$ bytes
- **Next header** → o que vem a seguir
 - Identifies next header/extension
- **Hop limit** = TTL (v4)
- **Options** → included as extension headers
- **IPv4 checksum removed, as lower layers are responsible for verification.**



72

Extension Headers

- Next Header – type of next header**

- Hop-by-Hop Options** → quem estiver pelo caminho vai vendo se há alguma indicação para os nós intermédios
 - the only header that is examined by intermediate nodes
 - Destination Options Header** → apenas para o nó de destino
 - Information for the destination node
 - Routing Header** → permite listar os nós por onde o pacote deve ser enviado: em vez de confiar nos **routers** para definir a rota, podemos explicitar isso num routing header extra
 - List of nodes to be visited by the packet
 - Fragment Header**
 - Authentication Header** → Cifra dos dados que também pode ter a identificação → manter a integridade do que é o pacote enviado
 - Encrypted Security Payload Header**
 - Transport layer headers**
- único header verificado pelos routers intermédios
- apenas para o nó de destino
- permite listar os nós por onde o pacote deve ser enviado: em vez de confiar nos routers para definir a rota, podemos explicitar isso num routing header extra
- Quem recebe o pacote, vai ao routing header e vê a informação para o hop seguinte
- Example do encadeamento de opções extra
- so queremos integridade e a autenticação dos dados



73

73

Auto Configuration Address

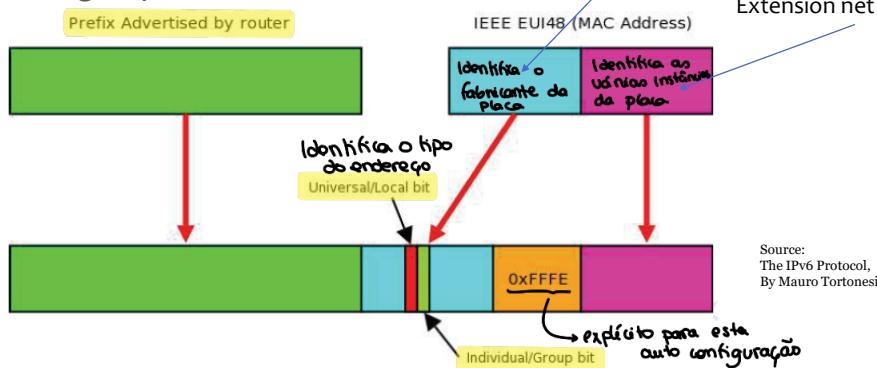
- IEEE-EUI64**

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

Endereço de 64 bits associado ao endereço MAC da placa de rede

Manufacturer of net card part

Extension net card part



Lookup vendors by MAC address

74

74

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
 - Maintaining information about neighbour nodes
- ND similar to the IPv4 functions
 - ARP IPv4
 - ICMP Router Discovery
 - ICMP Redirect

75

ND Messages

- ICMP messages (over IP); using Link Local addresses
- Neighbour Solicitation → "quem é que tem este endereço IP"
 - Sent by a host to obtain MAC address of a neighbour / to verify its presence → Semelhante ao ARP
- Neighbour Advertisement
 - Answer to the request → Resposta a)
- Router Advertisement → Router anuncia os prefixos (informações associadas à rede que está ligada)
 - 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

Summary

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

77

Homework

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

78



FACULDADE DE CIÉNCIAS
UNIVERSIDADE DO PORTO



FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

End of Network Layer