Everything is digital...

How to transmits on the net and live happily

massimo.marchi@unimi.it

Data coding

- Digital computers can handle only binary signals: sequences of 0 and 1 (bit = binary digit)
- In order to transform data by digital computers, it needs to **digitalize** data, i.e. transform real samples (images, sound, etc.) into sequences of bits, packed for technological and hostorical reasons into group of 8 bit, called bytes.
- The meaning of a sequence is given by the **format** used to code and interpreter the sequence, eg. ASCII, bitmap, mp3.

0100.0010
0011.1100
0100.0010

BITMAP 3x8

ASCII codes 66 60 66 ⟶ Characters "B<B"
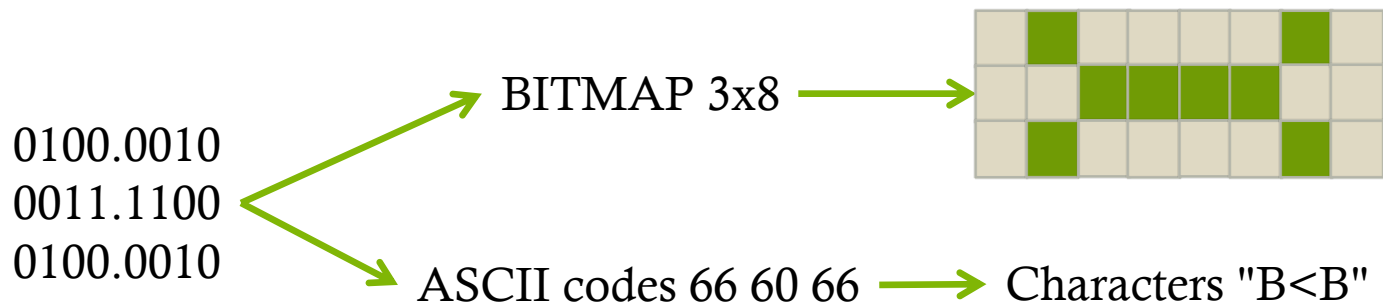
Digitalization

- Digitalization process loss informations due to the fact that you are taking only a finite sampled rappresentation of the original analog data.

- After digitalization, the information is expressed by a sequence of bits and code by some format that can contains some redoundance. By loss-less compression tecniques you can srink your data near to the teoretical minimal amount of bits needs to represent it (for eg. png format).

- Accepting some loose of information, by a loss compression tecniques, you can obtain a more srinked version respect to the teoretical limit (for eg. jpeg format for images) .

Coding Text

Bit, Byte, ASCII, UTF-8,...

- **Bit**: BInary digiT, a flag that can assume two values, 0 or 1.

- **Byte**: 8 bits, the "atom" of the information on PCs.

- **ASCII**: a way to code char with 7 bits, ex:
  010.0000 means <space>
  100.0000 means '@'
  100.0100 100.1111 100.0111 means "DOG"
    - try on google "ASCII TABLE"

- **UTF-8**: a way to code a lot of char with 1 to 4 bytes:
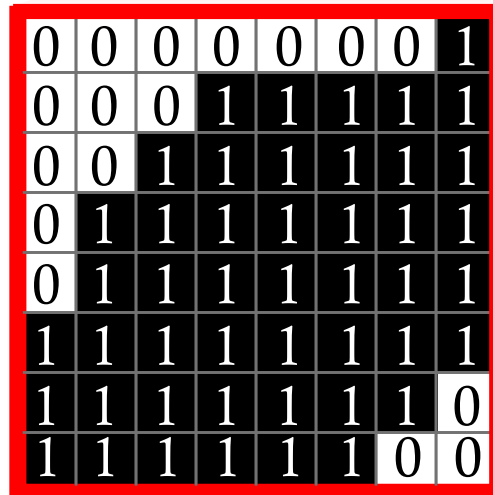   - try on google "UTF-8"

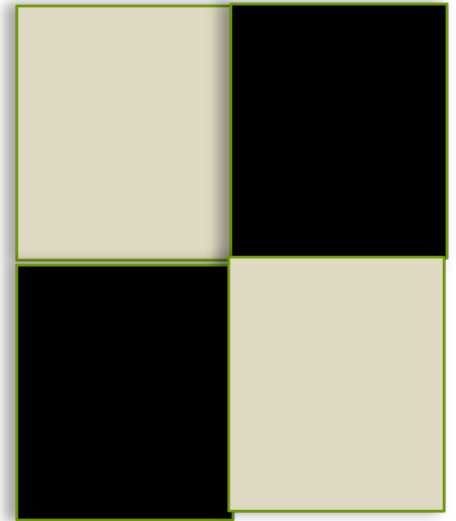# From pixels to images

bitmap

From pixels to images

bitmap

0	0	0	0	0	0	0	1
0	0	0	1	1	1	1	1
0	0	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	0
1	1	1	1	1	1	0	0

Coloring is Addictive

- All (visible) colors are combination of Red, Green, Blue
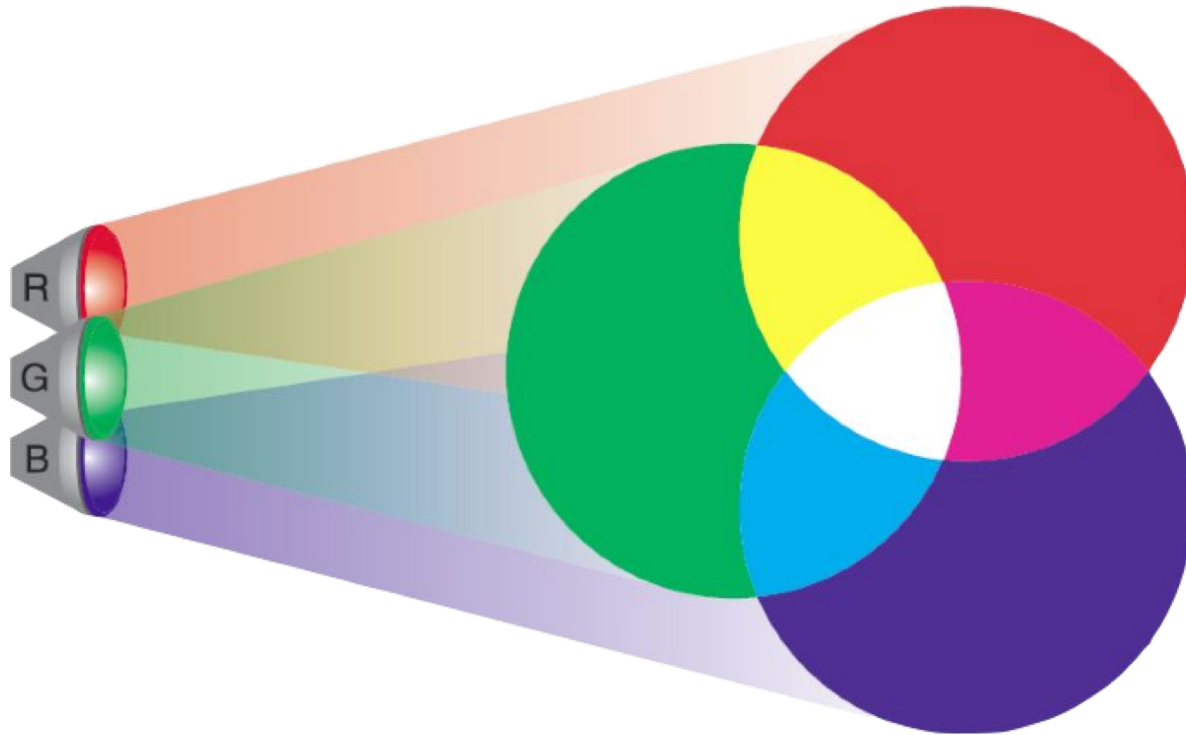
RGB Color Channels

RGB image

red dominant

blue dominant

Red Chanel

Green Chanel

Blue Chanel

Resolution and Color Depth

- Resolution: number of pixel composing the image:
  - 1024 column x 768 rows = 786.432 pixel

- Color Depth: number of bit used for pixel color:
  - Black And White: 1 bit/pixel
    - 1=black
    - 0=white
  - True color: 24bit/pixel (8 for red, 8 for green, 8 for blue)
    - 0000.0000 0000.0000 0000.0000 = black
    - 0000.0000 1111.1111 0000.0000 = green
    - 1111.1111 1111.1111 1111.1111 = white

Coding color with RGB

- **RGB**: color is a combination of **R**ed, **G**reen and **B**lue, (notation based on human perception and oriented to monitor devices)

- Each visible color can be obtained by addictively mix the basic color components Red, Green and Blue with a given ratio.

- The combination ratio can be expressed by 3-ple of numbers: <red_component, green_component, blue_component>
  - try: https://www.w3schools.com/colors/colors_rgb.asp

- The set of the possible component values defines our capacity to express color variations (color depth)

HTML Web Safe Colors

|											
#000000 0,0,0	#000033 0,0,51	#000066 0,0,102	#000099 0,0,153	#0000CC 0,0,204	#0000FF 0,0,255	#990000 153,0,0	#990033 153,0,51	#990066 153,0,102	#990099 153,0,153	#9900CC 153,0,204	#9900FF 153,0,255
#003300 0,51,0	#003333 0,51,51	#003366 0,51,102	#003399 0,51,153	#0033CC 0,51,204	#0033FF 0,51,255	#993300 153,51,0	#993333 153,51,51	#993366 153,51,102	#993399 153,51,153	#9933CC 153,51,204	#9933FF 153,51,255
#006600 0,102,0	#006633 0,102,51	#006666 0,102,102	#006699 0,102,153	#0066CC 0,102,204	#0066FF 0,102,255	#996600 153,102,0	#996633 153,102,51	#996666 153,102,102	#996699 153,102,153	#9966CC 153,102,204	#9966FF 153,102,255
#009900 0,153,0	#009933 0,153,51	#009966 0,153,102	#009999 0,153,153	#0099CC 0,153,204	#0099FF 0,153,255	#999900 153,153,0	#999933 153,153,51	#999966 153,153,102	#999999 153,153,153	#9999CC 153,153,204	#9999FF 153,153,255
#00CC00 0,204,0	#00CC33 0,204,51	#00CC66 0,204,102	#00CC99 0,204,153	#00CCCC 0,204,204	#00CCFF 0,204,255	#99CC00 153,204,0	#99CC33 153,204,51	#99CC66 153,204,102	#99CC99 153,204,153	#99CCCC 153,204,204	#99CCFF 153,204,255
#00FF00 0,255,0	#00FF33 0,255,51	#00FF66 0,255,102	#00FF99 0,255,153	#00FFCC 0,255,204	#00FFFF 0,255,255	#99FF00 153,255,0	#99FF33 153,255,51	#99FF66 153,255,102	#99FF99 153,255,153	#99FFCC 153,255,204	#99FFFF 153,255,255
#330000 51,0,0	#330033 51,0,51	#330066 51,0,102	#330099 51,0,153	#3300CC 51,0,204	#3300FF 51,0,255	#CC0000 204,0,0	#CC0033 204,0,51	#CC0066 204,0,102	#CC0099 204,0,153	#CC00CC 204,0,204	#CC00FF 204,0,255
#333300 51,51,0	#333333 51,51,51	#333366 51,51,102	#333399 51,51,153	#3333CC 51,51,204	#3333FF 51,51,255	#CC3300 204,51,0	#CC3333 204,51,51	#CC3366 204,51,102	#CC3399 204,51,153	#CC33CC 204,51,204	#CC33FF 204,51,255
#336600 51,102,0	#336633 51,102,51	#336666 51,102,102	#336699 51,102,153	#3366CC 51,102,204	#3366FF 51,102,255	#CC6600 204,102,0	#CC6633 204,102,51	#CC6666 204,102,102	#CC6699 204,102,153	#CC66CC 204,102,204	#CC66FF 204,102,255
#339900 51,153,0	#339933 51,153,51	#339966 51,153,102	#339999 51,153,153	#3399CC 51,153,204	#3399FF 51,153,255	#CC9900 204,153,0	#CC9933 204,153,51	#CC9966 204,153,102	#CC9999 204,153,153	#CC99CC 204,153,204	#CC99FF 204,153,255
#33CC00 51,204,0	#33CC33 51,204,51	#33CC66 51,204,102	#33CC99 51,204,153	#33CCCC 51,204,204	#33CCFF 51,204,255	#CCCC00 204,204,0	#CCCC33 204,204,51	#CCCC66 204,204,102	#CCCC99 204,204,153	#CCCCCC 204,204,204	#CCCCFF 204,204,255
#33FF00 51,255,0	#33FF33 51,255,51	#33FF66 51,255,102	#33FF99 51,255,153	#33FFCC 51,255,204	#33FFFF 51,255,255	#CCFF00 204,255,0	#CCFF33 204,255,51	#CCFF66 204,255,102	#CCFF99 204,255,153	#CCFFCC 204,255,204	#CCFFFF 204,255,255
#660000 102,0,0	#660033 102,0,51	#660066 102,0,102	#660099 102,0,153	#6600CC 102,0,204	#6600FF 102,0,255	#FF0000 255,0,0	#FF0033 255,0,51	#FF0066 255,0,102	#FF0099 255,0,153	#FF00CC 255,0,204	#FF00FF 255,0,255
#663300 102,51,0	#663333 102,51,51	#663366 102,51,102	#663399 102,51,153	#6633CC 102,51,204	#6633FF 102,51,255	#FF3300 255,51,0	#FF3333 255,51,51	#FF3366 255,51,102	#FF3399 255,51,153	#FF33CC 255,51,204	#FF33FF 255,51,255
#666600 102,102,0	#666633 102,102,51	#666666 102,102,102	#666699 102,102,153	#6666CC 102,102,204	#6666FF 102,102,255	#FF6600 255,102,0	#FF6633 255,102,51	#FF6666 255,102,102	#FF6699 255,102,153	#FF66CC 255,102,204	#FF66FF 255,102,255
#669900 102,153,0	#669933 102,153,51	#669966 102,153,102	#669999 102,153,153	#6699CC 102,153,204	#6699FF 102,153,255	#FF9900 255,153,0	#FF9933 255,153,51	#FF9966 255,153,102	#FF9999 255,153,153	#FF99CC 255,153,204	#FF99FF 255,153,255
#66CC00 102,204,0	#66CC33 102,204,51	#66CC66 102,204,,102	#66CC99 102,204,153	#66CCCC 102,204,204	#66CCFF 102,204,255	#FFCC00 255,204,0	#FFCC33 255,204,51	#FFCC66 255,204,102	#FFCC99 255,204,153	#FFCCCC 255,204,204	#FFCCFF 255,204,255
#66FF00 102,255,0	#66FF33 102,255,51	#66FF66 102,255,102	#66FF99 102,255,153	#66FFCC 102,255,204	#66FFFF 102,255,255	#FFFF00 255,255,0	#FFFF33 255,255,51	#FFFF66 255,255,102	#FFFF99 255,255,153	#FFFFCC 255,255,204	#FFFFFF 255,255,255
#000000 0,0,0	#333333 51,51,51	#666666 102,102,102	#999999 153,153,153	#CCCCCC 204,204,204	#FFFFFF 255,255,255	#FF0000 255,0,0	#00FF00 0,255,0	#0000FF 0,0,255	#FFFF00 255,255,0	#FF00FF 255,0,255	#00FFFF 0,255,255

- Try on Google: "RGB table"

Color Depth vs Size

- An example of B/W image with different gray-scale palette
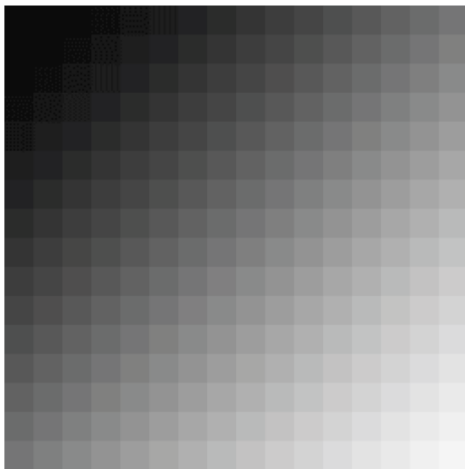
Resolution:512x512
Color depth: 8bit
Image size: 262144 bytes
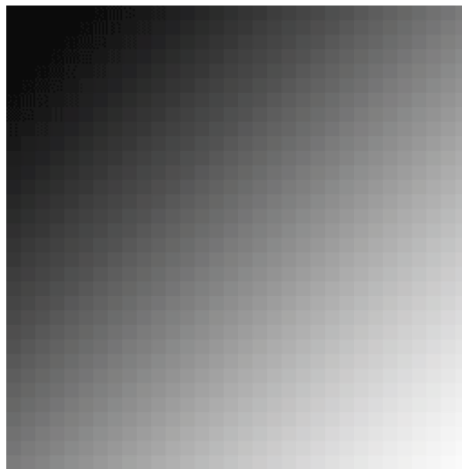
Resolution:512x512
Color depth: 10 bit
Image Size: 327680 bytes

Resolution:512x512
Color depth: 16bit
Image size: 524288 bytes

Other Color Spaces

- There are other many way to "code" colors; a short list:
    - **HSV** and HSL: color is a combination of **H**ue, **S**aturation and (**V**)brightness/**L** uminance (notation oriented to color perception, image manipulation)
    - **CMYK**: color is a combination of **C**yan, **M**agenta, **Y**ellow, blac**K,** oriented to print process by ink.
    - **RAL**: used in industries.

Reference: https://en.wikipedia.org/wiki/List_of_color_spaces_and_their_uses

Coding Sounds

Sampling and the Nyquist rule
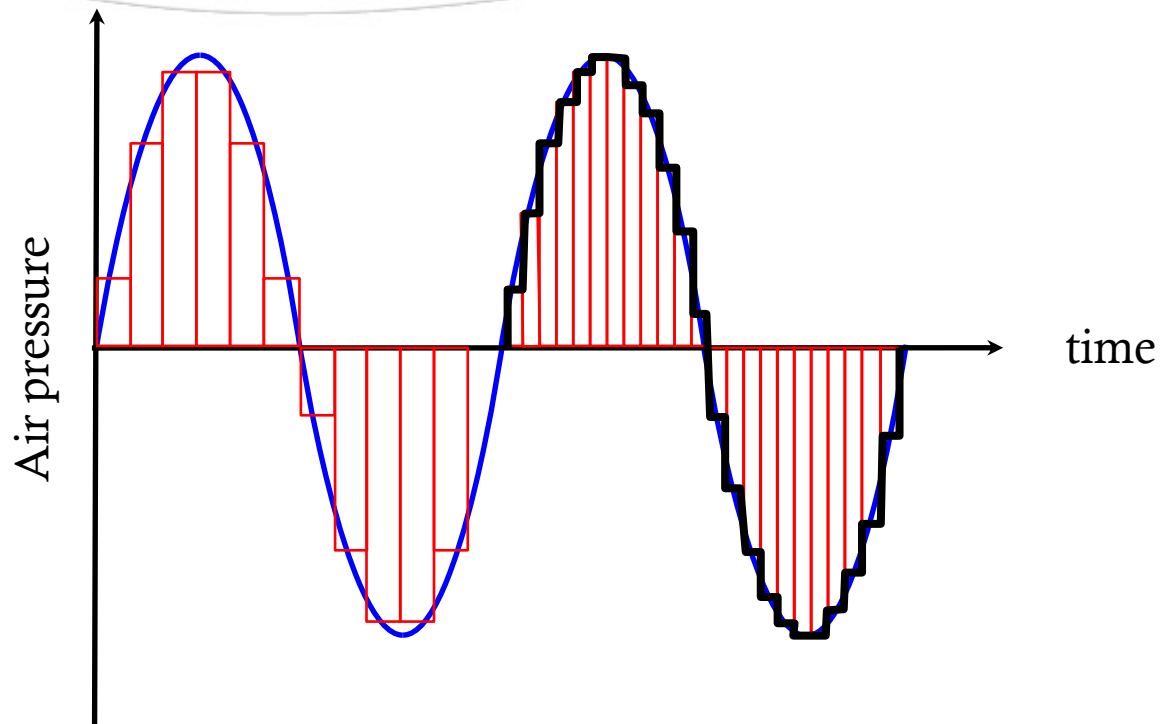
Nyquist rule

"In order to sample a looped signal, it is necessary to use a sampling rate of at least twice the maximum signal frequency"

Example:

- sampling at CD quality: 44khz

- Max audible frequency: 20khz

Coding Codes

Coding for humans

 ● Hi-Level language programs can be coded as text:

```
/* Hello World C-Language program */
#include<stdio.h>
main()
{
    printf("Hello World\n");
}
```

Coding for (not so) humans

- Low-Level language Language programs are also text:

```
/* Disassembled piece from hello.c */
/* in Assembly Language
....
4000561: sub %r12,%rbp
       : xor %ebx,%ebx
       : sar $0x3,%rbp
       : sub $0x8,%rbp
....
```

Coding for CPU

- CPU languages are made by sequence of bytes:
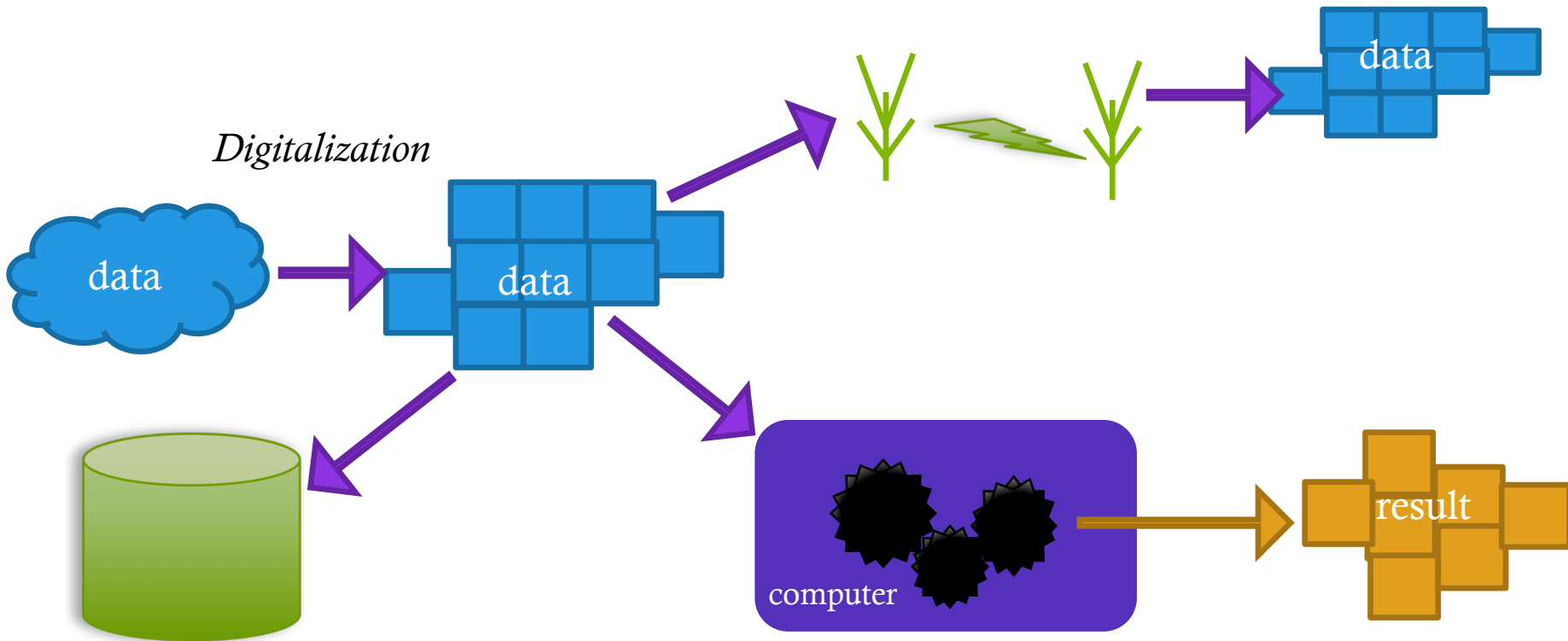
```
/* Byte Code piece from hello.c */
....
4000561: 4c 29 e5
       : 31 db
       : 48 c1 fd 03
       : 48 83 ec 08
....
```

Digitalization: the new world

- The digitization of the contents has allowed the automatic processing and transmission over long distances and the low-cost conservation of knowledge

The Net

*"Just think about it. Our whole world is sitting there on a computer."*

Angela Bennett

Packet Transmission on the Net

- Internet is self-organized meshed network.

- Internet transmits data by divide it in small packets of bytes.

- Each end-point have a worldwide unique address that can be used to reach it on the net (ipv4,ipv6 addressing)

- Each packet follows its own route through the net by jump from one node to another; the jump directions are locally decided by the traversing nodes in a best-effort logic.

- The target node have the duty to reconstruct the sequence (if need).

Internet

8.8.8.8

1.1.1.1

1.1.1.2
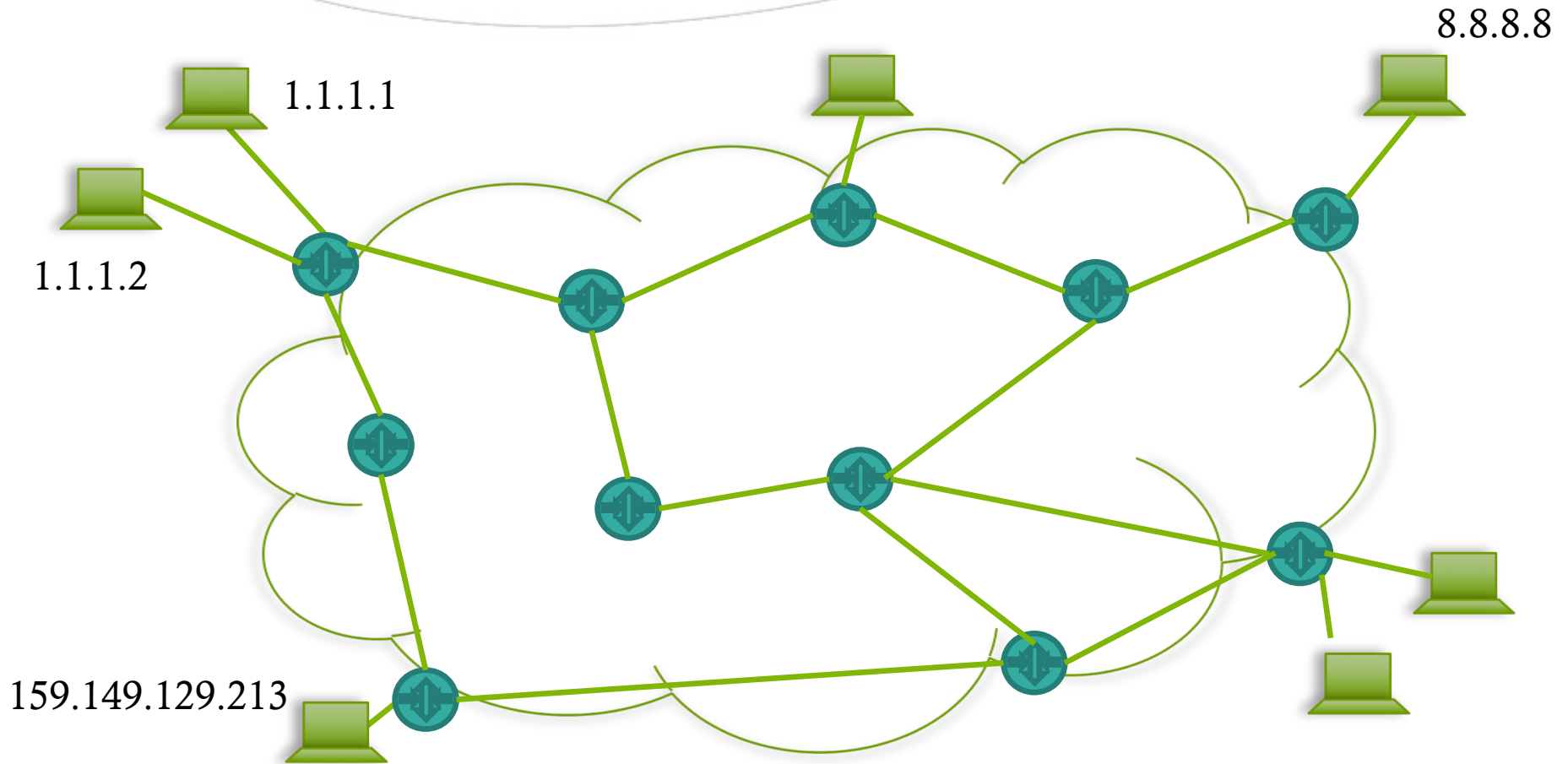
159.149.129.213

Adressing nodes: IPv4 and IPv6

- Each **public** node on the have its own unique address

- Each node on the net have a route for any possible destination; these routes are self-discovered and refreshed.

- IPv4 address are in the form of:

    <byte>.<byte>.<byte>.<byte> for eg. *159.149.129.213*

- Every possible IPv4 address has been already assigned to someone. Due to this problem, it has been proposed a new addressing schema, IPv6:

    - *2001:0db8:0000:0000:0000:ff00:0042:8329*
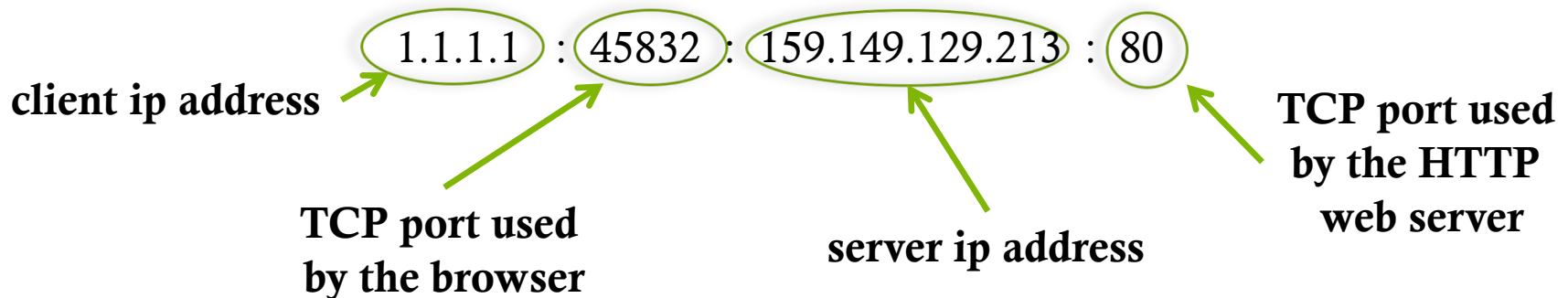
Transmission Control Protocol

🔹 Transmission Control Protocol (**TCP**): permits to establish stable conversations.
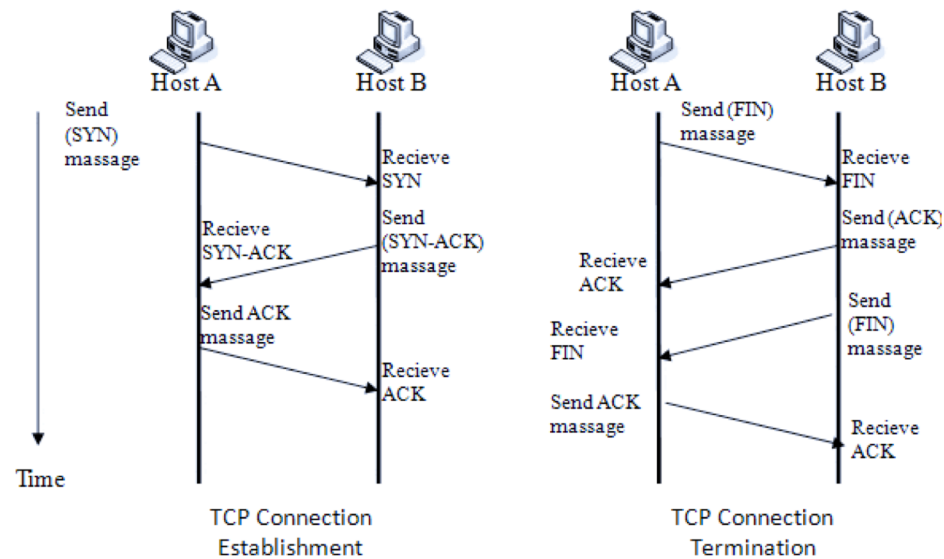
TCP session

- TCP use a number fro 0 to 65535 to identify services and client program inside node

- A TCP cornversation between two node is identified by the **session:**

<ip_source>:<tcp_port_src>:<ip_dest>:<tcp_port_dest>

1.1.1.1 : 45832 : 159.149.129.213 : 80

client ip address

TCP port used by the browser

server ip address

TCP port used by the HTTP web server

User Datagram Protocol

- User Datagram Protocol (UDP) is a protocol for conversation in which the delivery of every packet and the reconstruction of the correct sending sequence is not mandatory.

- It use the same TCP schema for identify a conversation:

  <ip_source>:<udp_port_src>:<ip_dest>:<udp_port_dest>

- Because there are less overhead informationi respect to TCP, it permits a more efficient transmission.

Domain Name System

- The Domain Name System permits (among other stuff) to resolve symbolic name into IPv4 address

- Any existing domain is resolved by a dedicated server; DNS servers are organised in a hierarchical schema.

- When a client needs to resolve a name, it query its own configured DNS server

ns.unimi.it 159.149.10.1

Doman: unimi.it

.....
grid003.ricerca.di.unimi.it → 159.149.129.213
marchi.ricerca.di.unimi.it →159.149.129.213

....

Hyper Text Transport Protocol

- Hyper Text Transport Protocol (HTTP) and its secure versione (HTTPS) permit to transfert information between browser and a web server.

- The server TCP port is used to select a particular web server process running on the server (they can be many, usually two: the HTTP and the HTTPS server processes)

- The symbolic name used in the address bar is used to select a specific site inside the selected web server (they can be many)

WWW communications

WWW Communications
happens at Application level

firefox

HTTP:...

TCP:43265

UDP:53

1.1.1.1

grid003.ricerca....

marchi.ricerca....

HTTP server

TCP:80

159.149.129.213

159.149.10.1

UDP:53

grid003.ricerca.di.unimi.it → 159.149.129.213
marchi.ricerca.di.unimi.it →159.149.129.213

L4 – Transport level

L3 – Internet level
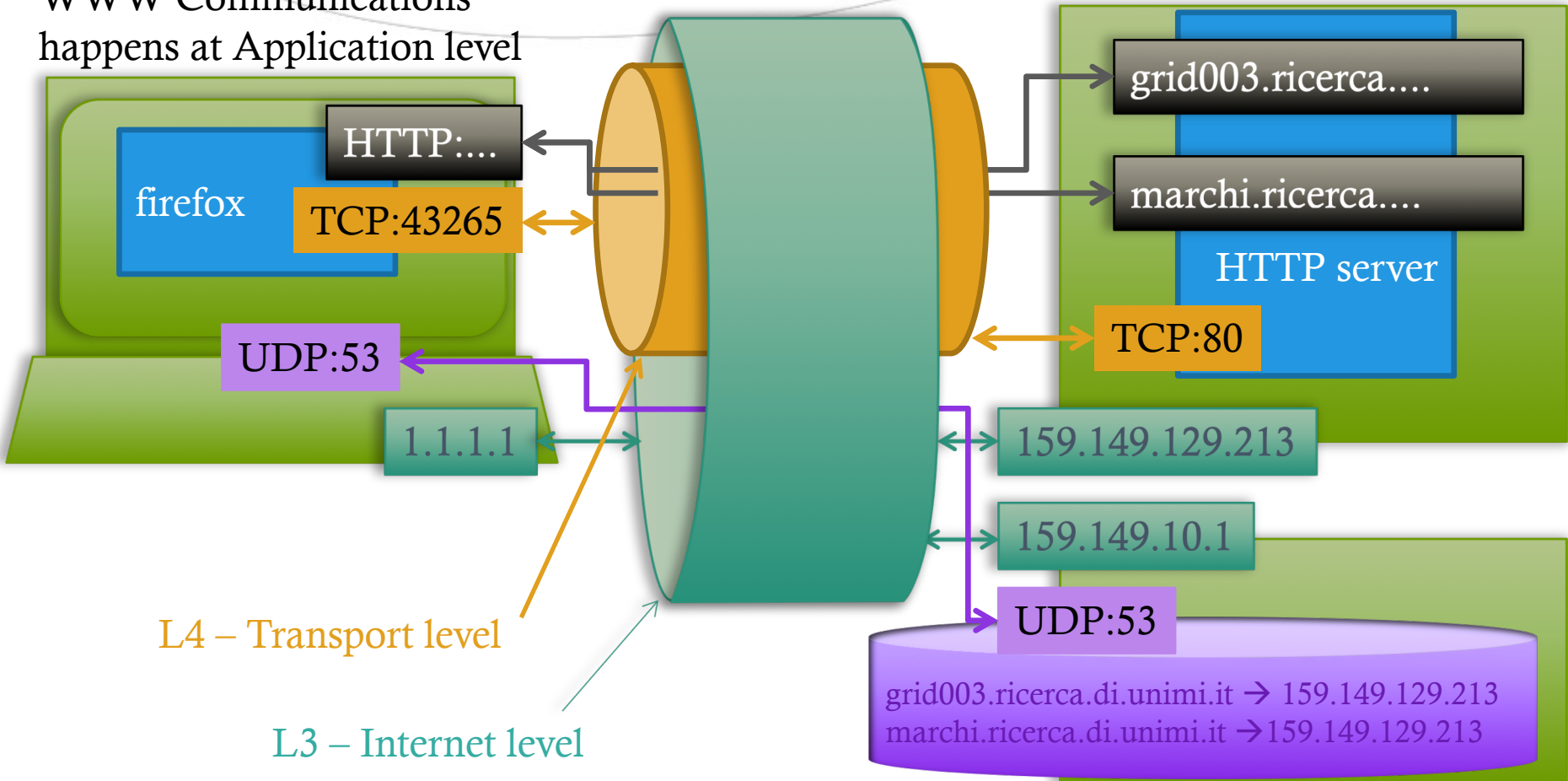
Network Address Translation
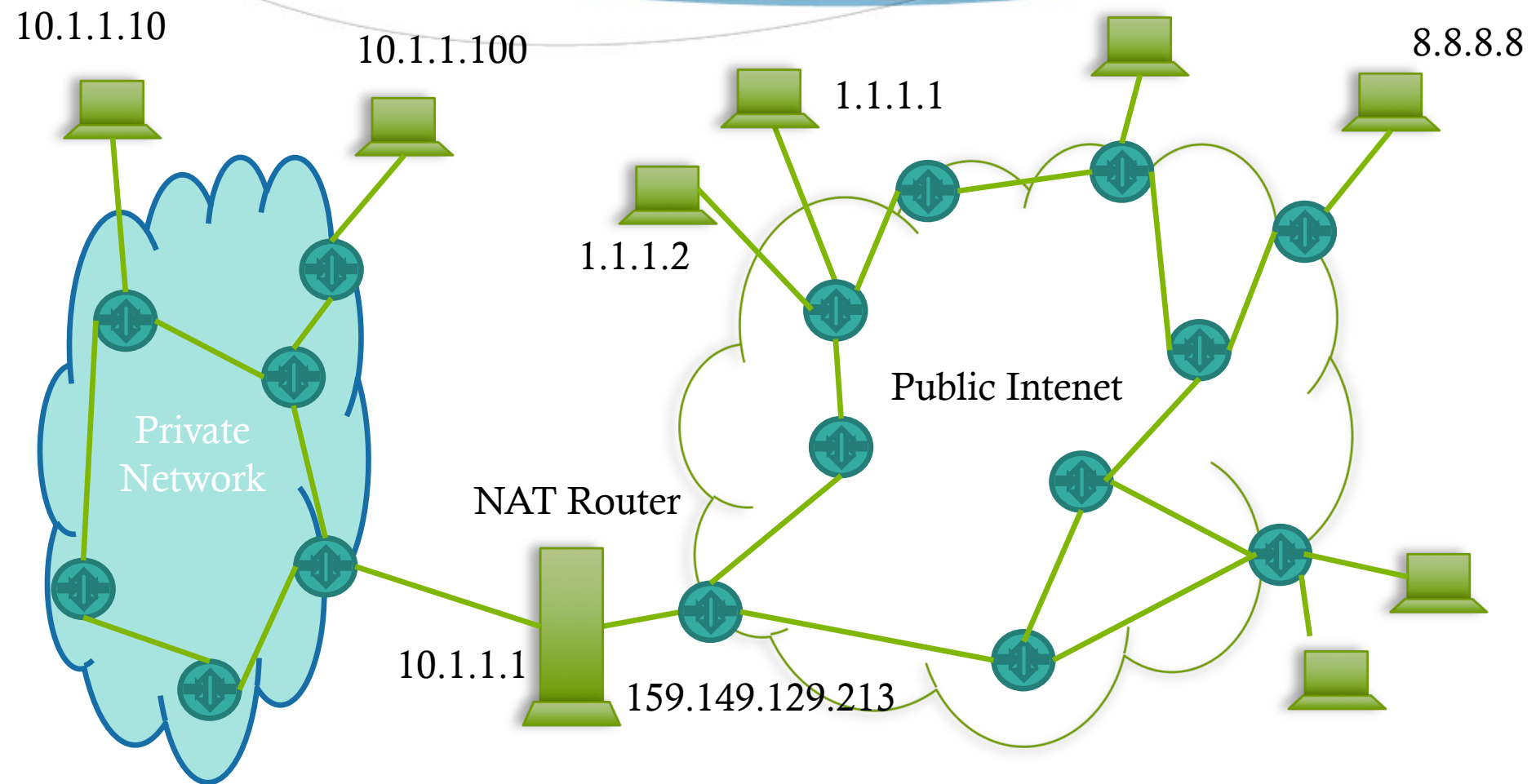
CDN

- "A **content delivery network** or **content distribution network** (CDN) is a geographically distributed network of proxy servers and their data centers. The goal is to provide high availability and high performance by distributing the service spatially relative to end-users."

https://en.wikipedia.org/wiki/Content_delivery_network