Network infrastructures

Introduction

The network infrastructures are divided into **2 parts** because technologies are different, they are the **access** and the **core**. The list of technologies change **frequently**. A network is built using different components, the terminals. They are:

- Computers
- End systems
- End devices
- Clients/servers

There is also **processes** that run on this network. The difference between client and server depends from the process running on it. Systems are interconnected by some **media**, like wifi etc.

In last years most of the network is becoming **software**. But what is a **network?**

First you have to set up it and then you can communicate. This is networking, by example using **TCP/IP**. Sometimes network elements are not physical but are called **virtual**. An end device could be a software.

We can have some different network natures, for example the well known **WAN,** wide area network, like IP, LTE, UMTS. That network covers a wide area. Then. Network could have a **shorter** area, even in terms of devices connected, and that is a **LAN,** local area network. We can have subcategories like Personal Area Network. In the middle we have **metropolitan area network**. The distinction is also in some details, like number of users.

Some network are **virtual.** There are concepts like virtual private networks. They are virtual not because they are a software but because the networking is done privately in a virtual network. It's a sort of **layering** of networks. This using protocols that differentiate users in the network. In a digital communication is important to have a **transport** element, something able to bring information, like a bit. We have a lot of wireless technologies, terrestrial, satellite, etc.

In a **wireless** communication the difference is the distance between devices. For instance, cellular networks. We divide an area in smaller areas, called **cells**. A cell is an area covered by a single network device, like access point. All the network elements in the area are connected. **Latency** is also important, the time to establish a connection.

Digital communication

In a digital communication system several building blocks are composed **together**. There is a first part dealing with **transmission,** then there is the part dealing with **communication** over the channel, then it goes to the channel and then there are **other** building blocks operating in the **reverse way**.

Typically there are three operations:

First, we have to identify if the media we want to communicate are **analog or digital**. Most of them are digital. Some analog media are transmitted in a **digital way**.

In any case, starting from analog or digital, we transfer our signal over a **channel**. Some building blocks are used to **convert** our digital information, bits, we also say "encode", in a way that is suitable to be transferred over the channel and this conversion depends on **2 things**. All the channel bandwidth that we have and the kind of media where we are transferring information.

In the **reverse** path, we obtain **signals** from the channel, we come back to have the bits and in case we go back to analog form.

What is **modulation**? It's the preparation of signals to be transferred over a given channel with some requirements. Typically, we want to transmit with **best performance**. Modulation tries to maximise performance of a channel. We use **digital modulation** because we have many advantages. The most important is that we can easily recover information after transmission over a medium. These channels are not ideal, so they **modify** the information, in several ways.

We can have an **attenuation**, signal starts with a size and end with another, than the channel modifies shape of our signal. Third, during the communication there could be **noise interferences**.

We start with a shape and arrive to the end with another. That's true with analog signals and digital. But with analog signals is very hard to **understand** the original form. On the other end, with a digital information we only have **two levels,** a 0 and a 1. After the transmission I'm able to recover the information, if it was a 0 or 1? **Yes,** better than the analog. Ho we do it?

We put a **threshold**. If it is above the threshold it's a 1, otherwise it's a 0. But this isn't enough. Another important aspect is **the timing**. If timing is wrong and the signal is not passed through an ideal channel we could have a **bit error**.

The less is the **bandwidth** of channel, the wider is the signal in the end. The original pulse is a perfect rectangular. During transmission we have **attenuation.** The effect of the signal becoming wider depends on the channel bandwidth.

If we transmit **very close** the pulses one to the other, do we have higher or lower bitrate? It's **higher**. But, depends, if the channel where I transmit is limited in bandwidth, the bit is not perfect anymore but his shape modifies. Does this happen? When the pulses are separated in time.

In other words, we can achieve an **higher bitrate** only if this effect is reduced and this happens when we have an high bandwidth. There is a theorem that relates channel's bandwidth with the bitrate, **Shannon capacity**. Your bit frequency can't be more of the double of channel's bandwidth. The bandwidth could be bandwidth of a signal or bandwidth of a channel.

Bandwidth of a **signal** indicates the frequencies that compose the signal, in the **channel** it refers to the range of frequencies I can transmit. If I put too much symbols in a bandwidth I cause a **self interference**, an interference cause by the signal itself.

Interference

In a digital communication, my goal is to transmit **as much bit** as possible. I put 2 values in time, 1 and 0, and this is my transmission. Instead of having 2 shapes, I could divide this range in **4 parts**. So I have 4 different values. How many bits can I encode?

So for example I have

11

10

01

00

In this way I have **2 bits per signal**. So, next step is having 3 bits. And so on, increasing number of levels I increase the bitrate and I solve the problem of **limits of bandwidth**.

The problem is that increasing the number of bits I increase **the threshold**. So even a small noise can alter the blue of transmission. But also **attenuation** could lead to a wrong transmission, also **interference**.

So, the **bit error rate** depends on 3 aspects:

- Real symbol interference
- Attenuation
- Noise

A channel bad or good depends on several aspects. Attenuation, distance, carrier frequency. If I transmit on a very high frequency, like 2.4 Ghz, determines **the distance** I can cover.

With a **very low** carrier frequency is related to very long transmission in terms of distance. Carrier frequency is related to the physical medium that can be crossed.

Conversion

If my signal is intrinsically analog, how do I convert it into a digital form? This is done by **sampling** and **quantisation**. Sampling means measuring a signal at a given frequency of time. With a **low frequency** I can't recognise the original signal. If I pick up the value at **high frequency** I instead can. Is there a suggested frequency to do this? In theory much is the frequency, better is.

There is a **theorem** saying there is a suggested frequency. This value is the double of f_b. So, for example to hear a speech which has the average frequency of 4 kHz for human voice, I have to sample at **8 kHz** minimum. More is the frequency, more are the **size** and **quality**.

This is not enough. Every time I sample, I have to convert that sample in **bits**. Analog to digital conversion. I do the opposite of previous operation. I check weather a sample is in the high or low part, putting a threshold. This operation is called **quantisation**. As we say for the sampling, quantisation has a role in the **quality** of the signal. Shortest are the distances, more are the bits. If we use **few levels**, we could miss some contents.

Today we close the recap started last week as for the **digital communications**. We were presenting topics related to the building blocks of a digital communication. What we have stressed is that we have 2 or 3 fundamental components in this **chain**: a first component is the part dealing with the starting from digital data and try to propose this data, to **encode** the data to transport them over a channel. The main problems are related to the **bandwidth** we have on the channel.

- If the bandwidth is **high** we can assure a high bitrate and a tight schedule of the symbol that transports the bit.
- If the bandwidth is **low** we have a low bitrate and the scheduling can't be so tight.

This is related to the fact we can improve the rate having a **multi level** communication, if the channel has not another kind of impediment indicated as noise. **Noise** is interference plus attenuation plus the presence of other communications in the same media. This noise does not allow to increase the data rate for a completely different reason than the one dealing with bandwidth. This two concepts are related to the **encode** in a channel.

Modulation

We skipped the last part, once we have formed our signal, we have to **transmit** the symbol on a media. So to transmit a symbol on a media the next step is named **modulation**. Once we derive that we can transmit our bits over a media next step is to determine how the modulation is done, but what is the modulation? Modulation is the fact that we present a signal in a form that is easy to recognise if the symbol is a one or is a zero. Indeed the modulation are quite complicated by themselves because there are **different ways** to modulate a signal to represent the fact that is a 1 or a 0.

The simplest way is a modulation that changes the **amplitude** of the signal to represent a 0 or a 1. If for instance the signal is transmitted on top of an electric cable we can do this way, it the signal is low is 0, it is high is 1. Even if this is quite intuitive this is **not so used** because it's very sensitive to attenuation for instance, so there is the risk of error.

The same concept can be used if we like to have a **multilevel** communication. Instead of transmitting one bit per symbol we transmit two bits per symbol, or 3, we have to differentiate the amplitude of the signal in the y axis. If it is very low it is 00, if it is medium low it is 01, if it is a quite high level it is 10, if it is very high is 11. This is a way to encode and modulate a signal to represent **two bits** per symbol.

However as we were discussing, the fact that the symbol is codified by using amplitudes sometimes is an impediment because if there are some attenuations this kind of signal can be interpreted wrong. This is an **amplitude modulation**.

Another kind of modulation is **frequency modulation**. We encode our bit by considering a shift frequency, a modification in frequency. We pick up a signal at a given frequency. Let's say this frequency, **carrier frequency**. How we do reflect the frequency? Looking at the time the signal crosses the x axis. Once we pick up this frequency, we have this nominal frequency, if we change a little bit this frequency, so we increase the frequency this is a 0, if we decrease it is a 1. So this time you see the amplitude is always **the same**, and this is an advantage, we encode our bit not anymore in amplitude of signal but in the **frequency**, starting from a well known frequency that is the carrier one, the nominal one. We have nominal frequency and some shifts. If the carrier is named f_0, if it is f_0 - x it is a 0. If it is f_0 + x it is a 1.

This is another way to modulate the signal. With the advantage of keeping the amplitude always the same. Can I set up a **multilevel** modulation in this scheme as before? Yes. For instance I set starting from f_0 if the frequency is f_0 - 2e it is a 00, if it is f_0 - e it is a 01 and so on.

Another approach is a **phase shift keying**, if we have again a signal at a given frequency this time the frequency doesn't change, what is changed is the phase. You see here the 2 symbols have completely **opposed phase**. So, thanks to these techniques we can compose a modulator. How this modulation scheme works depends on the system where the modulation scheme is applied. For instance, what is very important in a system is this fact, first as I said we should avoid interference, we can skip self interference caused by the bandwidth limitation.

Interferences and filters

Even if our system is not bandwidth limited we can have some **interferences.** One of these which is very crucial especially in outdoor environment is an interference named multi path. This means that the transmitted signal, the one you want to transmit has some **reflections** in the environment where the signal is transmitted, what happens is that this kind of reflection arrives with the original pulses, the ones where the symbol are. And there we have some replicas that arrive at the receiver **attenuated**, so this means with a different shape and shifted in time, delayed. This is the concept of multi path, with the original pulses. Then there are signals reflected by things in the environment. Multi path is a very **key disadvantage** in case of urban communication. In a urban environment there are a lot of buildings, cars that can reflect the signal. This is the first impediment. Unless you are able to estimate how the copies of the signal arrive at the receiver, and this is quite complex work, so for instance if you are able to estimate the delay of these replicas, this time what is a disadvantage can become an **advantage**.

There are these replicas of original signal, replicas depend on the multi path. If the replicas arrive in a **random way**, they are an impediment for my communication. If they arrive in a way that I can estimate, they can become an advantage. Why? Because you have multiple versions delayed of the same signal. You can **reinforce** what you are trying to figure out from the original signal by having the replicas. This becomes an advantage. This is the first point. The second aspect is this one, if we consider a bandwidth, in general a transmission on a given bandwidth is ideal if what passes through that bandwidth remains as **the original form**. Last time we mentioned the concept of **filter**. Filter means that we are operating in a given bandwidth, a range of frequencies. What is a **spectrum bandwidth**? By looking at the frequency domain, the bandwidth is the part of the spectrum where frequencies of interest remain unchanged. When we look at the channel we say the channel is a filter for some frequencies. Wireless channel, a cable, underwater channel. Anything behaves like a filter. Filtering some frequencies and keeping others. If I plot this for instance this is f_1, this is f_2, i say that this is a filter with a bandwidth b that is $f_2 - f_1$. This is the **bandwidth** of this channel. This means that all frequencies that are into this range are kept by this channel. The others are filtered out.

This is also named **band pass filter**. On the contrary there is also another kind of filter, a low pass filter, passing frequencies in the low range. Every channel is a filter. Keeps some frequencies and skips some others. An ideal channel doesn't behave like a filter, it has a bandwidth which is ideally **infinite**. Unfortunately all channels are finite, the bandwidth is limited but they can be ideal for that signal. The ideality is because they are flat. A channel that is a filter is **flat** when all the frequencies pass unaltered.

Example: In other words, when I speak I emit different frequencies. My voice is unaltered if all the frequencies that I emit pass through a channel that keeps the frequencies the same way they were emitted by me. In my voice there are 3 - 4 frequencies components. There is a little bit of f_1, f_2, f_3 and f_4. This is my voice. I speak through a channel, if the channel doesn't keep all the frequencies in the same way it is not an ideal channel. Try to figure out if my voice if one of the four frequencies changes, it is not the same voice that I emitted. So for instance let's consider a frequency attenuated of amplified, the speech is different. So the channel is ideal if all the frequencies have the same treatment. They are kept in the same identical way. All attenuated or all amplified. If one is amplified and other is attenuated the result is different.

Saying this, the channel is **ideal** if it has a bandwidth where the most of the frequencies I'm interested in pass, and keeps all the frequencies components in the same identical way. The wireless channels are not ideal meaning that there are some frequencies that are changed in a different way with respect to others. Sometimes this kind of behaviour is named **frequency selective**, some frequencies pass in a way and others in a different one. The problem to have channels with high bandwidth and flat are quite complicated. Looking at this picture (slide 27), you see on the left the behaviour in **time domain**. What happens? You have a change in amplitude. If the amplitude keeps changing it's hard to understand the message, for example if I quickly change from speaking with and without microphone it's hard to understand. If I always speak without microphone it's easier to understand, because after a while you find the best way to understand the signal.

Looking now at the **frequency domain.** If I restrict the bandwidth, the channel becomes more flat. But with less bandwidth I can transmit less bits. High bandwidth, channel not flat.

Channels include some self interference and unless you are able to do some sophisticated processes on the signal, these are impediments for you communication. So, this is the classification.

- Flat fading channel
- Frequency selective fading channel, so not all the frequencies are changed
- Fast fading channels, changing very fast
- Slow fading channels.

Quadrature amplitude modulation

Given this classification how to solve the specific **impediment**. As for the modulation we have that between the amplitude, frequency and phase modulation scheme the most used is the phase shift keying. But with some **modifications**, so again we can do as before a multilevel communication, as for the amplitude and the frequency, also for the phase we can change the phase of the signal with different levels, instead of changing the phase every a given number of degrees, I can change 45, 135 degrees etc. At the end the modulation scheme is the combination of phase and amplitude modulation. And these are named **quadrature amplitude**

modulation. What is a QAM? It is a signal where I can encode my bits by recognising the amplitude and the phase, a **combination** of the two. If I combine the concept of encoding with the amplitude and encoding with the phase I can put my symbols in a signal that changes both in amplitude and in phase, in this way I can encode my bits. So this kind of quadrature amplitude modulation can for instance be used by combining a signal amplitude and 4 different phases, how many symbols I can put on this? 2 bits. If we have 2 possible amplitudes and 4 different phases we encode 4 bits. We have 16 levels. So **QAM** are modulations where I base my modulation scheme on the amplitude of the signal as well the phase of the signal. The behaviour of our transmission depends on how we choice to transmit the signal.

We need to encode our bits on the signal and to transmit. In order to be robust as for the main impediments that are present in the system there are different techniques of modulation. The main are well known, like the **digital television** that uses the QAM. To synchronise your tv it appears QAM modulation.

Spread spectrum

Another well known technique valid in several environments is the so called **spread spectrum**. The idea of this technique is to encode the input data. You would like to shape the form of the signal to transmit your 1 and 0. To do this you have a spread spectrum, that means this signal originally is going at a very high frequency. Meaning that the shape of the signal is at a very high frequency, so during the period of a bit sub symbols are transmitted at a **very high frequency**. This part is named **chip code.** So the resulting signal is very spread, because this symbol uses very wide frequencies. The advantage is this one, we have allowed to transmit at a very high frequencies, to occupy a very high bandwidth if we transmit at a **very low power.** So it does not interfere with other transmissions. You distribute your power over a bandwidth very large but as an overall is low. You can perform your transmission and keep it still significant thanks to the fact that the chip code is known the receiver when receives the signal applies the **chip code** and you recognises your 1 and 0 that are completely different from 1 and 0 different of other transmission which are codified with another chip code. Is like applying a **key** to your signal only known by the receiver. This makes the signal with a shape that is significant only for the receiver. So, two **advantages**:

- I'm able to transmit in a very robust manner.
- I'm able to occupy a very large bandwidth because I transmit at a very low power.

Typically, this kind of techniques are used today in the new technologies that are rising for the internet of things. **Internet of things** are devices typically that need to be reached at very long distances with very low bitrate, it's important the fact that you would like to have a lot of concurrent transmissions contemporary. You can achieve using these chip codes and having concurrent transmissions of users. So you see in this example there is one transmitter who codifies his binary 1 with this chip and his binary 0 with this chip. When you receive a sequence of these chips and you apply the **decoding** by knowing this sequence you can recognise correctly the bit transmitted by one transmitter and the rest is seen as noise. So in this way you achieve a robust transmission at very low power. Decoding is made using at the receiver the same chip code used by the transmitter. Thanks to this you are able to recognise your transmission. This is named as **spread spectrum**.

The last idea is this. Instead of transmitting on the whole bandwidth my symbol, I divide the bandwidth in **sub bandwidths,** all with flat behaviour. This approach is name **orthogonal frequency division modulation** (OFDM). So, I divide my channel in sub channels. And this is the first point. For every sub channel I dedicate to it only a sub part of my transmission and I see that channel as flat, this is the advantage. Second, I put a symbol for my bit in that sub channel that is made with the form like this blue one, who has an important characteristic. The form of the symbol in time when I put this shape in frequency I see that the symbol crosses the 0s at multiple of the distances of the frequencies of the sub channels that I selected. The different colours here are **orthogonal**. This means that when a frequency is high all the others are zero.

Network infrastructures

Real introduction

This is propedeutic to understand some aspects of the technology that we'll meet. When we will describe the technologies we will look at different part of the network. We can recognise 2 or 3 parts of the network. There is a part called **access network** and there is a part named **core**. Sometimes in the access we have some subnetworks named **metropolitan area networks.** In the 3 elements of this networks we have 3 different requirements.

In the access we need to be very capillary towards the user. We need to find the way to reach all users. In the metropolitan area we need to touch the building elements of the area, depending on the area. And in the core we need to multiply at very high speed flows of data belonging to different users. In general the different parts of the network are build up using **different technologies**. We are aware that a big part of the access network in Italy is build up using **copper lines**, with a technology of the family of the digital subscriber lines. DSL means digital subscriber line, it means initial part of the network is digitally implemented, on the contrary today the evolution is to use in the access network **optical fibres** due to the characteristics. In the access we have these kind of technologies. Copper or optical.

On the contrary in the core network we have different technologies and different **topologies**. Typical topology used in the very very core part of the network is a so called **fully mesh** network. This means that all nodes belonging to that part are connected together, one to the others. In the access several times we have a star topology, in the edge several times we have a ring topology. Topologies in the network infrastructures are very important and are designed in different ways in order to assure different characteristics. Why in the access the topology typically is a star? Because it is natural to have a star. We have the end users that are connected to the network, so we think about them as **leafs** in a three. We can say that in the access part stars are the best topological structure to interconnect devices. In the edge or the metropolitan are used rings. Why? The deployment of a network, using stars or rings, which is the easier for a network operator?

Star mean that you have to reach one by one all network elements. Ring means that you have a single media, for instance the most of the networks set up in metropolitan area also private networks, like **Sapienza** network. Is a network where there are a lot of routers, located in different parts. All of these are mainly interconnected not in a star topology but in a ring one. Also because rings are easy to be **reconfigured**. This is something we will discover later. It's important in a metropolitan are or edge part is important to be easily reconfigured. If the star node fails, the entire network fails. This is a huge problem. In a ring, if one node fails, the ring can be reconfigured, it's much more robust to failures. This is an important requirement in this part of the network. It's much more important in the core, but in the core rings are not enough as for the reliability and reconfiguration of network. In that part is important to have fully mesh topology. They are reconfigurable but more they have capacity to reroute the traffic. As a principle, if you reconfigure a ring, where do you root the traffic? In a fully mesh you have more options. This is the reason why in a mesh is much **more robust**, in case it needs much more redundant capacity.

Also, in terms of **policies** adopted to manage the network, these are also different. For instance, while we mention some aspect related to the quality of service, like differentiating the kind of traffic sent in my network, the policies to mange this traffic. This can be done in the access, in the edge, quite hard to be done in the core. Why? Due to the **scalability** issues. In the core we have thousands, millions of flows, so how can we manage also a simple priority with that amount of data? This is hard to be done. On the contrary, can be done **toward the access**. The number of flows are less. Indeed, one approach is to provide the management of the intelligent functions like priority at the edge, the edge is the right place where the intelligence functions can be put. The edge may perform intelligent functions not reformed at the core due to scalability issues mentioned before.

After this part, you should know then into for instance a company you have a **subnetwork**, a local one with its own technology, protocols, then interconnected toward the public network, like fibre optic and then itself interconnected to ADSL, cellular network in the LTE technology, the optical networks and this part here it's a sort of three of fibre optics that is named **passive optical network**. In this picture as for the very initial part a first well known technology is the ADSL. The **ADSL** was born by using the topology provided for the network telephony. In initial structure the telephony was organised with having this kind of architecture, a sort of distribution network where I have a central element, the central office, other elements between the office and the end terminals. This part here is named **access local network**. The rest is interconnected in a regional area and then interconnected in the long distance area. This office is an office so is typically a building where all the interconnection of the telephone network both in ADSL and fibre optic structure are interconnected. For instance, this is an example of central office in a room that is somewhere, in centrals, inside this structure we have a lot of network elements where all the interconnections happen. End users are the users where the termination of the network are. The most important number reported in this table is this one, there were 105.000.000 of km of wires. Think about Italy, if you measure the amount of wire in 2007 dedicated to telephone there were 105 millions of km. This is the number that pushed the creation of the ADSL. Because who implemented the ADSL at that time, still true for several situations in Italy, they were considering the fact that there were already that **amount of cables** there.

Protocols and project

A project where some of you can participate is a sort of **survey of the protocols** that are designed to support the **internet of things**, I already know the names of these protocol and you can find material. The project could be a **comparison** of 2 or 3 of these protocols, by studying papers, only by comparing surveys, it's not about developing something, it's a study on these protocols in order to understand how they work.

The name of these protocols there are 3:

- **Lora**, low-range communication
- **Six fox**
- **Narrow Band LTE,** designed to support low-data rate.

We don't need high-data rate but instead we need **scalability** and **capacity** to support many devices at the same time. Remember the project is something done in a **group** of 2-3 students and after this you have to prepare a presentation of what you studied and the presentation will be in **December** in front of me and of the class.

Today the professor will present some **research activities**. We can develop this time also a **practical project** where we design some code, write some code, set up a simulation. As for the lesson, last Monday we arrived to these numbers, we will look at the new one, indeed this is a good idea the fact we can update my material thanks to your searches. This is something done. Yesterday there was another seminar, more oriented to **communication specific**.

There was a lot of discussion on the **broadband** access, using copper wires. What I'd like to stress is this point. At a given time period, true for all countries, since the **copper based** access was based on the **telephone network,** and this is the reason for having a big success, in the past mostly, there was the risk to cut a set of end users out of the broadband communication due to the fact that their operation were not able to operate in the access that was owned by specific operators. For example in Italy, true also in other countries, in the telephony terminology there are 2 kinds of operators. One is named incumbent - **ILEC** local exchange carrier. Carrier is another name for operator. The meaning of this term is that the ILEC was the **owner** of the local exchange network infrastructure, the one dealing with interface toward the users. Local exchange is also the name for device, central office, interfacing to users. In many cases there was one only big operator, like Telecom Italia, British Telecom and so on. Then the broadband access has increased and there was the need to **open the market** to new operators. To do this, there was the need to have more operators, concurrent carriers, to enter the market.

There were 2 solutions to enter the market, one was to provide **their own network infrastructure** in the access part. This was possible, but you know it's quite **complex** and expensive, this means that the new operator has to provide its own infrastructure. The other approach and the name of the laws were to allow the unbandling, this strong term here, means the fact that the operators **already owners** of the infrastructure **open it** to other operators. Allowing them to use their infrastructure, to **rent** their physical infrastructure and to provide their services on top of that infrastructure. This has been defined by law by defining this unbandling mechanism. Means again that the owner of the infrastructure opens to other operators. What does this mean in case? Why it is quite easy to understand how it works in the **wireless** communication? It's harder is to understand how the incumbent operator can open its own network to a concurrent one. Let's think how this can happen. In a fixed access.

The unbandling means they have to open, in case of **copper wires**, at the central office we have rooms, in the distribution network we have only rooms cabinets where devices are. This renting sometimes means only **renting the room** where the device must be put. Consider that this was not physical before this unbandling law. It was not possible to rent a part to the central office to someone else. Was not possible since the infrastructure covers **different elements** in a distribution network. A big space dedicated to host some devices. There are other elements in the network. Also this somewhere where the elements are put, is something to be **shared**, otherwise the network can't be used from other one. The idea was let's give the possibility to concurrent operators to rent rooms by the previous incumbent operator. By **law, o**therwise it would bee to complex for new operators. The most of the copper wired network was based in Italy on the preexisting infrastructure, not proposing a new one. The new infrastructures were proposed later only by other kind of technologies like the **optical ones**.

This of course has a **cost**. One meter of fiber costs about 0.005 €. One meter of **digging**, at the time this was presented, was about 90 €/m. This because you need the permission from the municipality, you have to switch the traffic to other roads and so on. To switch completely the actual technology with this new costs a lot. You can't afford it. No way. This is why the fiber is not so much used.

Optical access network

In a fiber optic infrastructure the network elements are mainly 2:

- Optical line terminal
- Optical network unit

Optical line terminal is at the network side. We discussed the access, that is between the end user and the edge of the network. The optical line terminal is at the edge of the network, the **optical network unit** are instead toward the end users.

For instance this is a 3 based topology that is typically used in the optical access network. Again, in the optical kind of networks we have **3 way of operating**. One way is very similar to the one used in the copper case is to have a **single fiber,** so a single link from an home or cab or a building, connected to the local exchange. A sort of star of fibers where the local exchange is the centre of this star and then the rest. This is the **simplest** but it's the one **more expensive**. That's because every user has to have its own fiber.

Another solution is this one, there is still the star but its closed to the end user. I have only one dig for reaching the active node and then from the active node toward the end user a star again. This is quite interesting solution even if the fact that we have an active node again means you have to provide an element and the **room** for this element, sometimes rooms are an **expensive** part of the network infrastructure. Because when you have a room, cabinet, you have to provide a person able to go there. The cost is also related to the fact you need someone able to **physically** manage that part. If you're able to provide an infrastructure without that kind of network elements is better in terms of **costs**.

Indeed, one solution very interesting in terms of cost for the optical network is the **passive optical network**, is a solution where there is a single piece of fiber, toward a point, and then a splitting of this finer toward the different end user. The advantage is that this point is a piece of fiber, something that has the **same size of a fiber**, some µm. Nothing to do with a cabinet, and could be provided below the road as a typical fiber. This solution is very interesting in terms of costs.

The architecture could be no fiber, a part a fiber arriving at the central office, from the rest we have copper, ADSL. Or a fiber arriving at network unit between the office and the end user, in a cabinet, **fiber to the cab**, or in a **carb**, similar to a cabinet, smaller, closer to the end user. And these are two possible solutions. The advantages of these solution are that by **reducing the length** of the copper cable we are able to increase the bitrate. Thanks to the fact we are able to reduce this we are able to say well this is a broadband access. Thanks to this. If the fiber arrives to the home or to the building these are solutions where the bitrate is the highest.

Thanks to the fact **we do not have at all the copper** that is a bottle neck of our infrastructure. If we decrease the length of the bottle neck is better. It has a **weight** on the whole infrastructure. The services uses today, how much are they symmetric or asymmetric? Maybe they have shifted more in the symmetric direction. Not at all. You know better than me.

Also the **wireless** infrastructure is **wired**. So strange. A big part of the wireless access technology is wired. Where is wired? From all the base stations that are

antennas we see on top of our buildings or on the highway masked as a tree, all base stations are connected to the rest of the network, you can imagine a big infrastructure that is made by using mainly fibers. If you imagine the base stations I don't know if this is true but I think yes. This says that in Italy this is the number of **mobile subscribers**. Now we have the seminar. We will see not the next week. On Monday you have the less practical lesson and on Friday you have the practical lesson.

Circuit switching

We continue with our lessons with discussing representing the old and new architecture of **telephone network**, this can be compared to the data network seen last Monday with Manuel. He presented an overview of the main principles of architecture. Then into details of the addressing and routing in case of IP.

What is missing is another part that has some important aspects also from the rest of the course that is the **circuit**. Even if it's presented in the initial part of this lesson, related to an **old technology**, the way data are transferred in a **circuit switching** network are still very important. Very briefly, this is well known indicating how a circuit switch operates. We have to recognise 3 main fases:

- Setup of the circuit
- Transmission of the data
- Teardown of the circuit

This is obviously **completely different** with respect what happens in packet switching, where there is **no setup** of a circuit, there is a transmission of the data made in a different way, indeed we recognise that the transmission of the data is a continue transmission, in case little bit delayed, this is the time running, our transmission is a bit delayed due to the crossing of the circuit **switching devices.** This delay is little, only something depending on how the crossing is performed, there is **no processing** on the data and circuit switch, only a delay due to physical transmission from input port to output port of a circuit switching device.

On the contrary, in case of packet switching you should remember that packets are **fully stored** in the switch, typically named router, processed and forwarded. The approach is named store and forward versus circuit transmission. **Store and forward** is done for 2 reasons. The first is the **processing**, for doing what? To understand where the packet is directed. Then there is **processing** to figure out the correct out port where the packet is directed, this is the **delay**. The rest is again a delay, due to what? To the queue, that happens in the note due to the fact that the packets are sent out to a given output port, packets are directed to a port, do we have the control of the traffic directed to a given port? Do we have the **control of the amount of traffic** directed from an input port to an output port in a router?

Typically no. Why? Unless we provide some traffic engineering, that is setup routes on the basis of the congestions, the packet switching does not admit the flows in accordance to a possible consideration of the congestion. In general you can wait in the queue related to an output port, a period of time that is **not predictable**. There is no control on the amount of packets admitted in the network, the **opposite** of what happens in the case of circuit switch. There is the control on the amount of data that are accepted in the circuits that are set up in the network. This store may take some unpredictable time, because there is a predictable time spent for processes, what you don't know is the **number of packets** directed toward the same direction or same output port. These are the differences you see in all books.

This should be known. In particular circuit switching will be used in all systems where there was a requirement on the **delay**. There are some applications, the first was the telephone application, requiring a control on the delay that the **data spend** into the network. Because if the data spend too much delay, in case of this basic application, the telephone, if you can't control the amount of time spent by data into the network the application itself **doesn't work** anymore. Indeed this was a basic element for building up the telephone system.

What is the telephone system we know, how in terms of **network architecture** this was structured, is also something well known, we can recognise the access part, as a topology made by a sort of star or tree, while in the core network there is a typically structure to have the connection robust in case of failures. The core network is also named **long-haul**, this part is the most peripheral one. It is a **hierarchical** structure, different levels of this architecture. For instance the stars at the beginning, also called local loops, are made by connecting directly as a star the telephone in this case to an element named **central office** and then the central office are interconnected to the rest. I think we already have seen this picture, in the previous slide, this was the architecture used in case of telephone, I mentioned this picture because we will see the **ADSL** that is not the telephone network but it's build on top of **this specific infrastructure**. We are seeing, the local exchange is the central element where the switching and managing of the services is implemented, in case of central office, we already have mentioned the difference of a carrier that is the operator managing a **local exchange** and also that some carriers are **incumbent** versus the others, competitive. The difference is that the first is the carrier **owning** the physical infrastructure, also the rooms where the devices are, while competitive is a local exchange carrier managing **only a part** of infrastructure, not owned by the concurrent carrier but taken by an incumbent.

This setup was very important because has influenced also the kind of infrastructure at the physical protocol level provided in the access network. The fact that these 2 kind of operator are present in the access has given a **design** to the network infrastructure itself. Obviously then after the local exchange part there is an **interexchange**, a network interconnecting the local exchange part so there are the carriers operating in a local area and carriers operating in an interexchange area. Typically it's a network owned by a carrier where also the interexchange with other operators, like a country and another country. A country has its own **main local exchange carriers**, thinking at Telecom, Fastweb and so on. Another country, like France, has its **own** local exchange carrier.

This is implemented in **internet exchange network**. the internet exchange network is the network where the carriers, long distance companies under their services. This is another important part of the network. In this kind of network there is another concept, named **point of presence**, also named **POP**. The concept of point of presence was already present in case of telephone network, the concept is the point where there is the **interexchange** as said before between 2 or more long distance carriers. In case of internet, we also use several times the points of present, this time the POP is where an internet provider, a sort of carrier in the internet terminology, provides the set of services to its clients. The point of presence in case of internet is also combined with **services provided** to the clients, for instance in case of internet service that should be provided is also the authentication.

The term was taken by the telephone, where indeed the point of presence was an element where the **inter connection** happens. This is a picture of a possible point of presence. What you recognise is mainly a combination of different routers, interconnected with different kind of connections, we can read this in this way. There are the interconnections toward the external part. This is how we can represent a point of presence in case of a **telephone network**. Notice that, in case of telephone network, we also differentiate the outer core from the inner. The **outer core** is a point of presence, close to the edge of the network. You see here the outer core has less interconnections, with respect to this, and also the outer core has **less capacity** in the links interconnecting the devices. You see here links have a name that is different with respect to the other case. Also only by reading this figure, you can recognise which of the two is **more complex**. The first, toward other point of presence like this.

Topology and cores

Presenting the **topology** of point of presence inner core and outer core in case of the Telecom Italia network. There are an overall amount of 32 point of presence. But only 2 of them, indeed 4, but 2, are the **inner core**. The inner core are here, one is in Milan and the other in Rome. There is an inner core here and an inner core here. We have only 2 network elements, a combination of routers, combined to operate as inner core point of presence. Indeed inner core are 4, in case of failure of one inner core the whole network doesn't work anymore, they are fully **replicated.** These are only two.

On the contrary, we have 28 **outer core** point of presence, the ones with a peripheral interconnection, is however dealing with a city, so it's peripheral but not so much, it's a point of presence in a city. This is the **logical architecture**. What does this logical architecture says? This is how logically the interactions of communication happens. Milan is for instance the inner exchange element towards the next of the world. Rome is the interconnection toward the rest of the world. This is the one showing the interconnection and in this case you recognise that the stars are mainly transformed, maybe they are physically located in different positions.

We said, the fact there are 2 in Rome and 2 in Milan, the right thing is to have them **physically separated**. They are two different parts of the city, quite far one with respect to the other, to be robust also for instance in terms of things that can happen. If we have time I would like to show, there is an exchange that I know we can visit. This is the **physical topology.** Let's go back to our slides.

This is where our inner and outer course are positioned. The position of this kind of elements is quite important and also the structure of the elements because on top of this infrastructure depends the rest of behaviour of the network. All devices are **interconnected.** The peripheral is a star and all devices are interconnected to the central office. Even if this architecture was quite old it has been kept as it is to support also the **ADSL**. Is important to mention that in case of telephone network this physical structure has some drawbacks. One important is **echo**. With telephone communications we have 2 directions. You can split then or in time or in frequency or you can have the 2 directions **contemporary**. In case of telephone network the 2 transmissions were performed contemporary. But the **problem** was that when you speak in a direction you can hear back the echo in **the other direction.** This is named echo.

Sometimes the echo was very present in the telephone network, maybe several years ago was something you can perceive as a **side effect.** This echo has to main way to present. One way is this one, one is named **cross talk,** I have a wire, this is the representation of the copper wire, twisted pair, twisted together, 2 wires interconnected in the same cable. First case, **far end cross talk**, a communication starting from a point arrives to the other point and contemporary to that communication also the **interference** of a contemporary communication of another wire, direct is one communication and this is another one. This communication interferes with this cable.

This is **far end interference**, also named far end cross talk, there is cross because is between different cables. Another cable interferes with my cable, another conversation, another talk, **interferes** with mine. This is one, the next is this one, that is there is a communication crossing a cable, interferes with another cable again, this is the near and the other is the far, you see the difference in the 2 cases? Let's write here, who is transmitting in this case, who is receiving? Which are the **differences**?

Both are interference from another cable. One is named far, the other is named near. What does this mean? The second is closer and the other is far. But then? Does this make difference in some way or not? If I have an interference from far, is it good, it doesn't matter? If I have an interference from near, is it good? It's better the interference comes **from far**, the interference from a near communication is worse. Why? It's weaker because it attenuates. There is a law that the attenuation is **proportional** to the distance. If we cross a long distance the signal will be attenuated more.

They have a different power, this power is **higher**. In that case this is better than this case. In these 2 cases the worst situation is this one, the next is something critical. We will see for instance in ADSL is much **more critical**. In telephone this was solved in some ways, in case of ADSL was much more critical and hard to be solved. Then, other concepts that were true for the telephone network and some systems. We already mentioned in the initial part of the course how an analog signal was converted to a digital one. We mentioned the **sampling** of a signal, the signal is transferred by sampling at a given frequency or a given time period, the symbols were transmitted by converting them in digits and these are the bits transmitted to the other.

There is a theorem saying if I have a signal covering a given bandwidth I have to **sample** the signal at a frequency that is the double of this bandwidth. In case of voice the bandwidth is 4 kHz. I have to sample at 8 kHz. This means that my time period is 125 µs. Every 125µs I take a sample. Why? Depends on the frequency that is 2 times the frequency band where the signal is. What is important to mention is that 125 µs has conditioned the whole structure of the network infrastructure for the telephone signal.

The whole telephone network, the **multiplexing scheme**, switching scheme and other aspects are all related to this **125µs,** it's like a magic number, given result to the whole infrastructure in terms of how the links are built, how the **interconnection** are built, how the device work and so on. Indeed what we do is that we take 1 sample every 125µs and then every sample is **converted** to 8 bits. The conversion is not normally done with a linear law. For instance this is a uniform quantisation of different levels, the quantisation is **not uniform**, there are much more levels in the higher part. It is the way to organise the distribution of the bits per sample in a quantisation structure.

Once the signal is quantised you have a different kind of structure, specified by **coding standards**. The standards that determine how the analog to digital conversion is done, how many samples you have to take and how the quantisation of the signal are performed. If it is a video for an high definition television, or other things like things. They were described in the past and based mainly on the same principle. Is to figure out the key elements that are sampling of the signal and these key elements are provided. These are details for some of you, what is important to mention is that the structure of 2 main aspects that are important in the telephone network that are the **multiplexing** and the **switching** are mainly based on this fact, samples are taken every 125 µs.

Multiplexing and switching

In a telephone network is important to provide **multiplexing** and **switching**. Multiplexing can be done in different ways. In the architecture provided for telephone network the multiplexing must be done with some **criteria**, indeed the criteria are the following. Starting from the so called digital hierarchy number 0, data flow at 64 kb/s, the flow resulting from the **sampling and quantisation** of a voice call, if you sample and provide 8 bit per sample, you achieve 64 kb/s. The basic building block for transferring the voice was 64 kb/s. Then 64 kb/s are multiplexed together in some structures that are called **digital signally hierarchies**.

These structures are, how do we combine 64 kb/s? Unfortunately there were different standards depending on where the standards were provided. In Europe the 64 kb/s were combined in a multiplexing structure composed by 32 channels. Indeed, 30 channels for the voice call, another channel for the alignment, we will see later. This one is the result of the **first level** of the multiplexing of these 32 channels. A line going at 2.048 mb/s. This is the result. In other countries this multiplexing was done with a different number of channels combined together, the basic channel is 64 kb/s but in one case Europe 30 channels multiplexed together, in other countries., there were 24 channels.

The result is that the first multiplexing hierarchy is influenced by the **number of channels** combined together. This is true for the first, second, third and so on. Next hierarchies are a combination of the previous one. The idea is this, I have single flows, the first multiplexing is done by picking the 64 kb/s channels and multiplex them together. To multiplex this I have to provide a structure to put these 64. In one case the structure is composed by having 32 flows and in the other case is composed by having 24 flows. This one is multiplexed in the next stage and so on.

Obviously what is implemented here influences what will be implemented **in the next stages**. The final data rates in the different levels and in different standards are **different**. This was something to be managed in this network. What is important is that the multiplexing, is typically done in the **time domain**, multiplexing time domain means you put together in the same frequency domain at **different time slots** the contribution, the flows of the different users, in this case you see different colours, meaning that one user is putting this slot, the same flow is put in this slot, this is true also for the yellow, and so on. They are put in **presented together** in terms of frequency, differentiated in terms of time slot.

Also to remind, not only a remind, another way to perform multiplexing is the **opposite**, we have contemporary different flows but they are differentiated in terms of frequency. These flows are emitted at the same time but what makes difference is that they are located at **different frequencies**. This is frequency division multiplexing. The current networks use **both** of them. Frequency division multiplexing is still used for all the wireless communications for instance, if you think at the modern wireless communication. Wifi has 11 channels, every channel is at a different frequency.

Not only his, also ADSL uses a sort of **frequency division multiplexing.** Channels are belonging to the same user but used for different purposes, in ADSL we have a channel dedicate to the voice, a channel dedicated to the upstream of ADSL and another channel dedicated to the bound stream of the ADSL. These are transmitted **contemporary** in terms of time. How can I differentiate where is the voice, the upstream? One way is that one is at a given frequency, the other is in another frequency, the third in another frequency band. This is current used frequency division multiplexing. This is the picture representing how this is **implemented**.

As for the time division multiplexing we divide our time in time slots and organise the time slots in frames. **Frame** is **repetition** of a group of time slots. So for instance the frame structure for the first level of multiplexing in Europe, the short name is E1, first multiplexing stage in Europe and is composed by having 32 time slots combined together in a frame lasting 125 μs. These 32 time slots are repeated with a period of 125 μs. The time of a frame is this. Why? Because is related to the **periodicity** of the voice samples that we need to take every 125 μs.

What I have every 125 μs is a group of 8 bit. I have 8 bit every 125 μs. 8 bits are occupied a slot, the number of slots in a frame are 32. These 32 are not all used for the data, indeed the first one is named **frame alignment** and used by the device to understand where the frame starts, I agree as device the frame lasts 125 μs but where this frame starts I don't know. I put an alignment word, a **well known sequence** of bits at the beginning of the frame, this allows a device to understand the frame starts. To do this the idea is the following, the device search for this word. It is not enough to find these bits once, because the idea is to find this multiple times. 4 times, it takes 400 μs, not enough. After 4 times, I can state well I found out for 4 times the same alignment, maybe I'm true. If you find 4 times you can state there is a frame starting.

This is for the **alignment**, another time slot out of the 32 is a time slot that was initially used for signally, what is **signally**? Are control informations transferred besides to control what the data are carrying. Indeed the first important signally component we can have in that kind of network is a sort of address. Is it clear? here we are transferring the voice, somewhere in some way we have to indicate to signal to the network **where this voice is directed**, from where to where. Where do I put this information? Not into the time slot of the data, but in another place. Originally, in this structure the idea was this one, let's put all the signal informations needed to support the voice in a single time slot of the frame. understand? I have 32 time slots, one is given for the alignment so not consider it anymore, the other 31, another 1, the number 16 is dedicated to signally, to say something relevant to the channels that are supported as for the data.

This was done by using these **time slots**. Only initially, it's an interesting way to see how signally works, only initially the idea was the following, how many channels we have to control by using signally? How many contemporary channels we have in this **time division multiplexing**? 30 possible different users. The number of channels I have to control is 30. I initially the idea was let's control these 30 channels by using one single signally slot. How can I share one time slot for controlling 30 different channels? Time slot composed by 8 bits. Can I share a time slot of 8 bits to control 30 users? How do you do?

Instead of having one bit divided in time, we have **4 bits** divided in time. 4 bits of a time slot control time slot 1. The other 4 bits control time slot 2. 4 bits control time slot 1, the other 4 bits control time slot 2. In this way we have controlled only 2 channels. We have **other 14 channels** to be controlled. Let's say in other way, if I have a channel and I'd like to have my own bits to control I can pick up 4 bits of this, another channel can pick up the other 4. This cannot be done every frame because otherwise we are only 2 using this signally bits, the idea is to do this **every 15 frames**. Indeed 16 if you put also the alignment. Every 16 frames there is this schema. First, frame in this group of 16 skip for the moment, second frame, frame 1, in this 16 frames, 4 bits for channel 1, channel 2, third, second frame in this 16, 4 bits for channel, 1, 2, 3 and 4. Frame 4, 4 bits for channel 5 and 6, and so on. Every 16 frames a channel let's say in this way has its own 4 bits. Every 16 frames. If I have to communicate something for my channel I have this possibility every 16 frames, how many seconds lasts for having again the possibility to send information? The rate is 64 kbit/s. This is the data rate for a channel. Which is the data rate of the **control channel**? Is 4 bits every 16 * 125 μs.

The resulting **data rate** is 2 kbit/s. Here we are. So I have a controlled channel so the channel transfers the data at a data rate of 64 kbit/s and I have a control channel for this one, this is the channel 1 and this the control for channel 1 transferring information at 2 kbit/s. I have the possibility to have 2 kbit/s to control this one. Where the data are. So, in other words, this is the **overhead** for controlling this. This is the payload, these are the overhead to control this.

So, to be precise the idea to have controlling channels at a data rate that is less than the controller is quite **usual**. The idea to exploit in a multi frame structure some time slots for putting **control information** is very common in several systems. They have some channels shared to put control informations. In the case of the **telephone systems** this was implemented in this way till a given time period. The same was done in the T1 frames. The T1 puts 24 channels in a time frame and then puts extra bits in a different way in a bit repeated with a **multi frame** period. In all cases, every channel can have 8 bits every 125 μs.

This idea was true till a given time period. It happened in the story of communication networks that signally started to become **more important than the data.** The applications, in case of the telephone as well, required more signally than data. In other words, an amount of signally, 2 kbit/s, to support 64 kbit/s was true where you have a telephone. The signal is very **few**, they started then to build up services on top of the telephone that were **more rich** in term of signally, the first example of rich signally added to the telephone network was the **green numbering**, green numbers are numbers for instance used in a different way. Then, numbers for voting, for devoting some money to an event. All these add nothing to do with the telephone itself, is not anymore, it is signally, I have to transfer **toward the network**, doesn't deal with the telephone call. Understand?

This has regarded the telephone network first and the idea is that this was **not anymore enough**. It may happen maybe the vice versa is useful, I need more data in the signally than here. Which is the story? At a given time period the signally network became a network **fully separated** by the telephone network, supporting all the signal messages that were **more important** than the data. I mention this because also we will see later story in the communications is represented every 20 years, there was a net separation of the part where the data are, named **data plain**, with respect to the part where the signally is named **control plain**. This was done let's say 20 years ago. It's not anymore visible to have signally and data integrated. Let's divide the 2. Let's have a data plain where the telephone calls run and the control plain to provide all services that we would like to provide for a telephone call.

Another service is the **green number**, you compose a green number, you don't pay the call, is paid by the receiving side. Or in some cases is paid more. All controlled by the signally, not anymore by the data. In the signally you decide wether this is a call you have to pay a lot. Or the fact that you receive a call **from a call centre** and you call a call centre, maybe the call centre is not in Italy, even if you compose an Italian number, you compose a green number, you don't know where the number is physically. All of this is managed by the signally. I mention this picture and express this point because this picture is something that happened to the telephone network 20 years ago. Let's say 2 years ago happened the same story to **the internet network**. Today we are talking, you will see in other courses, of a network named **software defined network**, where the data plain is splitter from the control pane. In internet, the control information where is? Did you analyse the routing in a IP network? The IP packet is read by the router to understand where it is directed. This was done till 2 years ago. Now, again, the idea is: let's provide a software control engine where the IP goes, they are controlled and give back commands to the router to do what it has to do. The same identical story proposed after 20 years. This is why we today have this networking. The networking is done in an other network with other data.

This is for the signally. Signally was proposed in a separate platform and this platform was named **signally system number 7**, the network infrastructure to provide signally mechanism for the telephone network. What is strange is that the signally system network model was provided by a network elements that are managed, that are a sort of that were organised in hierarchical way. One element is the mind of the network, some elements are the points where the signally arrive, this is a centralised architecture. This is the infrastructure for the signally system that is the platform to provide control information on a telephone network. The signally system number 7 knows the first **packet switched network**. Was a network build to support the telephone one, mainly provided as a packet switched one. The signally device information are packets, switched as we know in the IP world. This to support the telephone network, initially. Why this architecture had a lot of success?

First because the kind of services provided to the telephone increased. There were a lot of services provided besides the telephone calls, green numbers, etc. More important, this kind of network infrastructures were used to control the **mobile telephone networks**. The signally infrastructure must be much more rich with respect to the data to be transferred. Again, this is related to the story but still very actual.

Time slot interchange

Let's close the discussion of last lecture, the **construction of this structure.** Let's recap. We were dealing with the telephone networks, in that case we were discussing there are 2 main **important aspects:** first, how the data are generated. Once they are generated we have to define how data are multiplexed, typically the idea is to perform a **time division multiplexing,** and to form a frame where the different channels are multiplexed together.

This is in case of **E1** frame, the E stands for Europe, the frame formed in case of Europe, instead of **T1**, used in US and Japan. Differences are **not in the data rate** of a single sub channel, this is not possible because the data rate in that case is **dimensioned** in order to support the data rate needed to rebuilt the analog signal for instance in case of voice, while the differences are so not in the single channel but the differences are in the combination of the channels in a **frame**, what happens is that in case of E1 we have 32 identical channels in the same frame and the **frame lasts 125 μs**, in case of T1 we have again a frame lasting 125 μs but this time the number of channels into the frame are 24. So the consequence is that the final data rate of these 2 kinds of frames is different. You see in the first multiplexing scheme in case of E1 we have 2048 kb/s, in case of T1 we have 1544 kb/s, this is the difference.

Then the next step is this one, obviously in the **built of a network**, a telephone, another important aspect is the **switching**, the crossing of a network element, named node or switch, this kind of element has the task to **connect** an input port with an output port. You already know and maybe also have seen in the past lessons **how a router works**, maybe is less known how a switch works in case of a circuit switches architecture. Like in the IP network, the switch interconnects an input with an output basing on the information given to the switch by a control entity.

So there is a control panel where some informations given by **signaling** are processed to give input to the switch. It's **not the same to IP**. The concept is similar. We have to interconnect an input with an output. The difference in case of IP is the processing. The question is, do you have a similar scheme in case of IP? As for the switching? We can say yes, maybe. how? You remember how a router works? Is it similar to this or not? I would say **yes** if you think in a specific way. Depends on how you provide the explanation.

What happens in case of a **router**? Who decides which input port is interconnected to another output port? The router. On the basis of what? How a router decides to interconnect an input port to an output port? If this is decided by the router. By processing the **header of the IP packet**. The header is the **signaling**. Again, also in case of an IP packet we have the header, we have a router, when the packet enters the router the header that is the signaling is processed, this header is given in some way to a **control entity** that has to do what? The routing table. Thanks to the signaling I'm able to understand which is the right output port where to direct the packet. Indeed in the table you have **IP address** of the destinations and output port.

What I want to say is that this scheme is similar, this was already present **15 years before IP**, this is similar, the difference is that while this is done **only in one step**, the signaling arrives because it's together with the IP packet. The output port is derived, in case of the classical switch of telephone network the signaling is something **processed separately** from the data. Indeed, it's true that in case of a telephone network this signaling control is implemented in another network, indeed last week we said yer we know that after a while this signaling was implemented in a real separated network.

What's implemented here, we said there is **another network**, and indeed this is only to confuse ourselves a little bit more, a network packed switched one where packets transfer only signaling information in the control pane to force to derive how the data have to travel into the network, the data plane. In case of telephone there was a **control plane** run during the signaling phase and the decision to interconnect an input and output port are taken. During the signaling phase the decisions are taken, then implemented during the data transmission, instead in case of IP decisions and implementations are **taken contemporary**. The packet arrive, is processed, then is forwarded, instead here signaling arrives, is processed, decides and then when the data will arrive the decisions are implemented, are run.

How this decisions **are implemented**? The initial function was implemented having the **switching matrix** with a set of input ports and a set of output ports, the switch was decided on when an input has to be interconnected to an output. This can be seen as in this way, this was an hardware switch, where you have to interconnect an input and an output temporary during the transmission of the data. This was the so called **space division switching**. It happens that some data arrive, from the input port 2 and is interconnected to the output port 4. The decision on when and which cross point are configured is taken in an other phase that is the **signaling phase**, this decision is taken during the data transmission, then it's implemented.

It's **different** indeed with respect to the IP case. Do you have questions? Even if logically the kind of operations are the same, you have to decide on basis of signaling which input port has been interconnected with which output port. It's done in a completely different manner. This was in the space division switching. But in case of **telephone** thanks to the fact that the multiplexing, strictly related to multiplexing, the switching is also done in the **time domain**. What does this mean? This is a switch in the space domain, an input is interconnected to an output, on the contrary how a switch in the time domain is implemented? It's something more interesting now, how do you switch a communication in the time domain? How can you figure out to **switch in time**? What does this mean? Switching time, how do you think it can be? Think a little bit. How do you switch something in time? Let's consider 32 time slot, a switch in time is switch, so change the position in time of a content of a time slot in an other. The content of this goes here and so on. This is a **switch in time**.

Let's first think at this, a switch in time. This one is to store these bits for a number of time slots 1, 2, 3. So, I take this, I wait 3 time slots and I put this as an output, this is a switch in time. **Store and forward**, like in IP. But the store and forward is only done by having **already knowledge** of the position where I would like to put my content. Number 2, 3, 4, 5. Number 2 goes to 5. Number 4 goes to 6. This is **time slot switching**. The control information derived during the signaling phase is something that set up in the switch this table, 2 -> 5 and so on. Replaces, switches the time slot position in a frame. The only problem is that **I can't go back in time.** Since I can't do this, which is the solution? The only way is to skip a frame. Have a delay of 1 frame. The switch operates in this way, I have 1 frame of delay that is 125 μs. Every 125 μs the switch reads the content of a frame and **postpones** the content in the next frame in a different position. This is the **time slot switching**.

You see, this is the **example**, I plot there. I have in a time slot switch a configuration of time slots as **input** and a different configuration of time slots as **output**. This is the switch. Can you figure out how this works, when we have multiple inputs and multiple outputs? Think at an example with 2 inputs and 2 outputs for instance. Is identical. I read 2 inputs, in this example how many time slots are read in an input? 8. And these are configured in a different way in the output. Is this something that was **already known** by you? If you have 2 input and 2 outputs, input and output are done in this way. Every 125 μs you read these 2 inputs, 8 time slots, it's like **reading a matrix**. You read a matrix, 2 rows and 4 columns, and you change the columns of the matrix in the next time frame. The question is, why I have to **switch** one column from here to here? Why I have to put in another position and once we answered this if it is important that the position is different with respect to the initial one.

For sure, a movement from here to here is mandatory because maybe this input has to go **toward a different direction**, this is what a switch has to do. If this goes to Rome and this goes to Milan and I have to send my information to Milan I have to exit from this output. Is it clear? It's related to destination. I'm asking to the switch to direct toward the right direction. For sure it was easier to put 2 in 2, 3 in 3 and so on. Why do I have to **reconfigure**? The answer is this one. Did you catch the answer? Because maybe this 2 was already assigned to another communication crossing the switch and we can't control. During the signaling we can ask to the switch if is there a **time slot available** for that communication?

The overall switching in this way, this operation is named **time slot interchange**, what happens is that the control unit thanks to the signaling information exchanged previously defines **only the positions**, 1 goes into 3 and so on. A mapping, we can call this as a **mapping**, of time slots from the input port to the output port, this is defined during the **signaling phase.** The switch, during the data communication, only has to read and write, read the input and write the output, read and write of time slots, read in an order and write in a different order. Read one, write in 3, read 2 and write in 4.

Everything is **postponed** by 1 frame. I read time frame T and I implement the write in the time frame T + 1. Which is the delay of crossing this device? It's a sort of **store and forward**, with have a deterministic delay, constant delay that is 125 μs. Our **magic number**. This means that a switch operates with a constant delay of 125 μs. One single switch, it has to read all the input frames and write in the next so after 125 μs positions the time slots. This ends the combination of multiplexing and the switch. I already mentioned that last lesson that the idea to capture some signaling, to **process the signaling** in a control entity and then to act on the switch by having this kind of path is something present for sure in the telephone network, something not present in the classical ip network, for 30 years. The IP network was born 40 years ago more or less. The idea now is this one, I extract from the IP packets **some information**, in the header but not all, also in other. I give this information in a **control entity,** it decides something and configures the switch to operate.

Only to close the loop. We started 40 years ago, maybe more, 70 years ago when the telephone was born, with this kind of principle, then IP arrived and then we are similarly back to this kind of principle. This is everything we already said, I want to check but everything is that.

The SDN, **software defining network,** works in this way, extracts some information, gives information to control entity, it processes, derives some conclusions and says to the switch to operate as he decides. How can this be implemented? By acting on the routing table.

XDSL

The **digital subscriber line** is the last part of a line, also known as local group. Before the XDSL, the local group was **only analog**. This final part was implemented in this way, there was the telephone, then the analog line, then the well known **central office**. The analog to digital and digital to analog conversion was done at the central office, from the central office to the network the rest was digital. It happened that was an interest of the subscriber to have **also a digital line**. The digital subscriber line is a line transferring bits, not analog signals, starting from the subscriber. This gave raise to the digital subscriber lines.

As I said, initially the local loop was not digital, til some years ago. The first tentative to have this digital subscriber line, were done by having a technology named **ISDN**. That was the first digital subscriber line, then some evolution of this and then the **ADSL came up**. The ADSL was **the first real** digital subscriber line. All the emphasis in this technology was derived by the discussion we had previously that is the most a big part of the network, was done by using **wired local loops**, using copper wires. We have already seen numbers, if you remember there was that famous number of 100 millions of km of copper. So, 100 millions of km of copper already there. The idea is why not use that part of network to have a digital interconnection? There is still the so called **digital divide**, the idea to provide digital interconnection is this one. This is the number.

How to use these cables? But also taking into account there were some **constraints** deriving from the infrastructure, it was already there for the telephone. On average, for a network, in this case telecom Italia network, maybe also other operators have this infrastructure, the average distance of a central office to a subscriber, a terminal, **is about 1,2 km**. So this 1,2 km is the **average distance** of a central office with the subscriber. Indeed this distance has an impact on the kind of technology that has to support the digital network.

The first solution was proposed in 1980, was to provide an ISDN. The idea is to have a digital network **in the initial part** integrated the idea at that time was to have a network to support also the data service and so on. Integrated at that time was intended in that way. 160 kb/s was the **throughput maximum** achievable using the ISDN. Then this is why I mentioned this, was born the **asymmetric digital subscriber line**, the interest was not only in digital services, but to have an infrastructure able to support the **video on demand**. Remind that in the US they have the video on demand, I think, on cables from they typically cable TV, is something that was there so they had the cable television with respect to us that has always been by broadcasting, in the US they were aware of the analog at that time several years ago, for them it's easier the idea to transfer on that cable **also that data**.

Let's try to figure out how to use the copper wires cables to transport video on demand. Considering the infrastructure already there and considering the constraints of that infrastructure. Remind also that the initial use of that simple wire that is the twisted pair, the twisted pair was used in the analog form by using in the spectrum domain only up to 4 kHz of bandwidth. We already said this, transmit the voice in the analog form, it's sufficient to occupy this band.

The first tentatives were, let's transmit data in this bandwidth. I perform a **digital modulation**, my bandwidth is 4 kHz. After the design of this modulation scheme, also named as **voice band modem**, a modulation transmitted in the voice band, was maximum 56 kb/s. The first implementations of a transmission with modem operating in the voice band offered a data rate of **56 kb/s.** For some years we used that part, to transmit our emails, our ftp on that kind of band at 56 kb/s. It was very hard, it lasts some years, after a while obviously it was not enough. However the interest was to keep the infrastructure identical to the previous one.

How was the infrastructure? The subscribers, the cabinets, the wires and the central office. Theres's a central office and a star of twisted pairs. Twisted pairs are distributed in the **distribution network.** These are only cabinets where the wires are together. There are no switching, no operations like this. It's only an element where the interconnections are collected, together. In the voice band case the idea was let's put the modem operate in the voice band and thanks to these few little bands available, the capacity is **directly proportional to the bandwidth**, the capacity is if we have the capacity of 4 kHz I can have the maximum capacity given by the **Shannon formula.**

First disadvantage is this one, we have only 4 kHz per second. Second disadvantage, quite annoying, was this one, when you had to transmit the data you **can't transmit the voice contemporary,** because they were occupying the same bandwidth. It was a problem at that time. So, two problems, the **limited bandwidth** and the fact you can't use voice and data contemporary. In the central office there's the switch, takes the voice, on the other side takes the data, and then direct toward the right network, if it's voice sends the information toward the telephone network, if it's data sends through the IP network, this was as implemented in the analog voice band modem.

After a while the idea was to try to **increase the speed,** second try to figure out exactly a technology able to combine the 2 services contemporary. And this gave rise to the **XDSL family**. This X at the beginning has a different meaning depending on the kind of solution provided, indeed X in some other cases it's the speed, in some other cases it's related to other aspects. What was important to indicate is this one, since the **distance** of a central office with respect to the terminal was about 1.5 km we have to remind the **attenuation** and then the final capacity depends on the distance, so since in that formula of Shannon there is the dependency on the signal to noise relation, since the signal to noice relation depends on the power of the signal and the voice, since the power of the signal decreases as the distance increases, we can figure out a behaviour like this. The amount of bit per second I can exchange in that part of the network **decreases with respect to the distance** from the office. New standards are trying to offer more data rate.

First tentatives

So in the last year on the goals is reduce the lengths. As we put the fiber closer to the end user the faster is the **data rate**. The first tentative was named **HDSL,** created to support T1 services. T1 is that frame provided in US with 24 time slots, the idea was let's make available to the end user a frame where I can put my information **in a T1 form**. Anyway there was the problem of not sending voice and data contemporary. Then thanks to the idea of supporting video on demand they provided to design the **ADSL.** The idea was to support video on demand. Then, after a while, this technology was successfully used to support data and gave rise to the ADSL in the local group.

The first solutions were designed to support in a distance of about 2 km, with copper wires, the data rate maximum was 8 Mbps. Consider that is a nominal data rate, second is very high with respect to the 56 kbps of before. There is a **big gap**. Then the standards are trying to provide higher data rates.

As for current speeds, these should be **asymmetric**. As I mentioned this asymmetry was given only because the initial service for this kind of technology was a video on demand. A video typically you need much more bandwidth **in the downstream**, much less in the upstream, in fact ADSL works in this way, we have a **frequency division multiplexing.** We multiplex 2 services in different streams. We have a very short bandwidth up to 4 kHz for the voice, a bandwidth to 138 kHz approximatively for the upstream, and a bandwidth from 138 kHz toward about 1 MHz for the downstream, the direction from the central office toward the subscriber. Since the **frequency bands** are different now we know that if this is larger than this we can transfer more data. The amount of Mbps is more.

It's asymmetric because we have different spectrum of bandwidth. Given this, the result is also the data rates are **asymmetric**. Specifically, we have some **gaps** between the different spectrum. We have a gap about 20-25 between voice and data, upstream, and 250 between upstream and downstream. This is the frequency allocation. For the time let's think at the figure on top. This is the **allocation of the frequency** and the consequent occupation in terms of frequency. Thanks to this, first thing is that obviously they designed a technology able to go behind the 4 kHz, otherwise was not possible to have high data rate in those short bandwidth. The idea is to **occupy the wire** til 1 MHz.

If I occupy more bandwidth, can I do this? Are there limitations? Why they didn't invent this previously? There was **no need**. In the past that wire was used only for voice, and it was enough 4 kHz. These were not enough for the data. Let's try to use **more bandwidth** in terms of frequency. We can do this, but we have to take care of some aspects. First, considered also in other solutions, let's use again till 1 MHz but the whole bandwidth. First constraint, if I use one starting from 0 this means when I transmit in the 4 kHz I can't transmit the voice contemporary.

So, with a **frequency division multiplexing** you could have a contemporary transmission of voice and data. This continued to be there, doesn't matter anymore, the rest is put in another frequency band. Another problem was this one, in a cable, twisted pairs are grouped together, in general we have a cable named **binder group** where I have different wires of different users. With 100 users I have 100 wires, every user has its own wire. If they are physically collected in the same wire the kind of wire used here since it's an **unshielded twisted pair,** this means the signal over the wire can interference with the signal of another wire. They are grouped in the same bigger wire. If they are physically together and they transmit in the same bandwidth they can interfere.

In general it could be not a problem, I don't care the overlapping of the ADSL, unless you share a binder group with an user using and old telephone. This is **still a problem.** Another problem is the **cross talk**, the effect of a signal that passes from a wire to another one. This time if the wires are together, physically in the same cable this cross talk may have an effect, I'm not worried about a cross talk of me in via Eudossiana and a person living in another street. I'm worried with the communication of a neighbour, these 2 maybe share a **common binder group**. If this, cross talk could happen.

Near end cross talk, I receive my signal, affected by the transmission of another user. Let's map this configuration in our **ADSL network**. The ADSL is done in this way, I have a copper wire, another copper wire, a terminal, another terminal, in this way. Can this effect happen? How many combinations do you have? Try to do, I asked you at **the exam**.

This is a binder group, let's consider having 2 wires in the same group. This starts from the central office, somewhere these cables are separated, for instance in a cabinet, because they almost arrived at their home, in **different locations**. If this acts as transmitter and this as receiver the NEXT can't rise because they are not in the same binder group anymore. If this as transmitter and this as receiver the FEXT can rise because they are in the same binder group. If this acts as transmitter and this as receiver the FEXT in opposite direction can't rise. If this acts as transmitter and this as receive the NEXT can rise because they are in the same binder groups. We have 2 FEXT, 2 NEXT.

Why we are interested in this kind of problem? Because especially when you use binder groups at high frequency, for instance you occupy bandwidth **up to 1 MHz** or more, the higher is the frequency the higher is the amount of interference that can happen. If cables are neighbours and use high frequencies they can interfere each other. If the frequency is low, like in the voice calls, the effects were present but **less important**. Higher frequency, more amount of signal passing to one cable to another.

Initially, in the plain telephone network upstream and downstream were present in the same band, the idea is, can I increase the downstream occupying also the upstream band? Having **contemporary band**, downstream and upstream, in the same frequency? Can I do? This was already in the telephone network. I can do, if I do an operation named **eco-cancelation**.

An eco happens when a receiver receives the signal of another transmitter, let's consider this is the central office, this is the upstream, from here to here, and the central office is receiver. This is the downstream, the central office transmits toward this, and the downstream occupies **also the bandwidth** of the upstream. If the central office transmits in that way there is a **NEXT,** because the upstream and downstream contemporary, one is the receiver and the other is the transmitter, if this happens at the central office, we can do a **cancelation**. Why? We can cancel the eco derived in this NEXT. The central office, receiving from a transmitter, can cancel the signal that he transmitted in the other direction. He knows which is the signal he transmitted.

If the same entity is the central office, it can cancel the signal he transmitted in a direction in the receiver direction, this is the signal generated from the same entity. This configuration where the downstream and upstream coexist, even if they can generate NEXT interference, can be used if we operate with an eco cancelation. Thanks to this we can have a **higher data rate** in the downstream. The downstream can improve the data rate first because it is the downstream and because the downstream is transmitted from the central office to the other, and it can cancel the interference. To understand this it's important you have quite clear what happens in the picture.

Cross talks in ADSL

We were discussing about **effects of cross talks** in the ADSL or in general in the copper wires. Cross talks were present in the copper wires, already when the telephone network was used on top of that infrastructure and this effect was present in an easy way in case of telephone network. Remind that in case of telephone network the transmitter and receiver were **contemporary active** on a single wire. Because the transmission and reception is done in case of telephone on the **same frequency band**. But this effect were less important due the fact that the frequency used was a low frequency, this kind of effects are much more important **when the frequency is high**, there are measurements that show that as the frequency increases the effect of this cross talk that is this passage of signal from one wire to another wire is much more important.

So this is the reason why in case of ADSL a solution is to, there are different solutions, one is to **differentiate the different bandwidths.** This is done by some network elements dedicated to splitter out a frequency band with respect to another, these are **splitters**, devices able to filter out the telephone bandwidth from the ADSL bandwidth, filter out the cell upstream from the ADSL downstream. If this is not done and we were here last time we can still think at the transmission of a downstream **occupying also the upstream** bandwidth.

The first possibility is to have different transmission for downstream and upstream. They **do not interfere**. On the contrary if we would like to have a downstream occupying much more bandwidth and this is sometimes interesting, we can send at higher data rate, there is the **problem of the NEXT**, of a transmitter that is transmitting in upstream and interferes with the transmitter, but depends on the configuration because while this is less important when there is the transmission of an end user and reception of the central office this is more important if the central office transmitter and receiving as receiver, the advantage is that in case of central office we can do **eco cancelation**, that is cancel what is transmitted because the central office **knows the signal** that he is transmitting.

The central office being aware of the signal transmitting is able to cancel this signal from the reception, understand? Remind that this happens **in case of ADSL** by combining cables in the **same binder group**, this is true if the cables are in the same binder group, on the contrary if these are cables separated the effects are negligible, not possible that a signal from a wire is interfere with a signal on another wire, this is only present when we have the signal in the same binder group.

Modulation of the signal

Said this, let's go to the next, another important aspect of ADSL, this is very interesting because says how the signal is modulated. Why modulation is obviously done in every case, you remember modulation is how to **form a signal** that is transmitting a digital stream, ok, so we have a sequence of bits, we discussed in the initial lessons, these bits are transmitted by pulses, to be transferred at a given data rate, in a given spectrum band. In this case we have a modulation for **2 reasons,** we have to transmit bits and find out the right signal to transmit on copper wires, first reason, the second reason, less important is that we have to move our transmission **in the right band**.

We have to shift in frequency the signal in the right band, there is one band centred in the upstream frequency carrier and another band as for the downstream. A first example of modulation was named **carrier-less amplitude phase modulation**. During our brief recall of concepts we said there are different ways to modulate a signal, phase, amplitude or a mix of the two. The frequency where the signal is send was not transmitted together with the rest of the signal, but it is less important, only to give the name.

This was a way to transmit, the result was this one, there was a **sequence of bits** let's say, to be transmitted, for instance the upstream of an end user, so these are modulated and the modulation **gives raise to a signal** that is positioned in correspond of the upstream frequency band. Occupying the **whole band**. The same for the downstream obviously. Modulated to be transmitted in the downstream frequency band. However this kind of modulation was **not so effective**, the reason why this kind of modulation was no effective is that as an overall the amount of bandwidth available especially in the downstream is very large.

When you transmit in a **very large bandwidth** one problem is that typically the channel is not flat, not ideal. Has not an ideal behaviour. **Ideal behaviour** means we would like to have a behaviour common to all frequencies in this range, and this is commonly not true, the idea was to use another kind of modulation, **discrete multi-tone**, also known as multi-carrier modulation, in the family of the multi-carrier modulation, we already discussed.

A **multi-carrier modulation** is the following idea, we split our spectrum band in different sub carriers, given a spectrum band we divide this spectrum band in sub-portions, sub pieces, named **sub carriers**. Then we transmit, we modulate our signal by splitting the signal in different sub carriers. This number is multiple of 2, related to the fact we use a kind of process by using a digital processing named **fast Fourier transformer** and inverse, you know what is? It's a way to combine signal in different sub pieces. In this case is the inverse, done by using **multiple of 2**.

The idea is, let's consider our spectrum band divided in different pieces, you see here, on these bars are the different pieces. Given that the channel is not flat, if we measure the **signal to noise ratio** on the whole spectrum band the behaviour can be this one, is this a flat behaviour? No, we have a very deep part going down, this is for sure not flat. If we divide in different sub carriers, in 2 for instance, is it flat? Again no, 4? Maybe. What happens? If we divide in 256 sub carriers and we look at every bar, the question is do we see in the sub band of a sub carrier a flat behaviour of the channel? Probably yes. This is the fundamental reason for having this kind of modulation. We divide our whole spectrum in 256 sub carriers.

Every sub carrier is seen as a **sub band**, every sub band measures a signal to noise ratio that is **flat.** Does not change with the frequency, this is the meaning of flat. The SNR is the same as a function of the frequency, again this picture is very intuitive, here is not flat, in this case is flat. But what you see as difference in the 2 cases? In the second case is not flat, but? The **amplitude is different**. In the different sub bands I have different SNR. Then next step is this one, since the modulation scheme must be designed on the basis of the specific SNR, if you remember we mentioned this, the capacity of a channel is directly proportional to the SNR. A capacity gives raise to the data rate. The data rate is **directly proportional** to the SNR.

In this kind of figure, in other words, is like this case, we have a **different possible data rate for every sub band**. Understand? While maybe in this case what you do is well let's think at an average SNR, first case of modulation, I take this, the average, I compute the modulation to be sender. This case, I take every sub band and I see a different SNR in every sub band and then I **adapt my modulation scheme** to the specific sub band and a modulation scheme changes from a sub-band to another sub band. In other words, if I have 256 sub bands I have **contemporary 256 different modulation schemes**. This is represented here.

This is a diagram representing what happens to our bits sent in this case. I have a sequence of bits, converted **from serial to parallel**, in principle I would say I have 256 sub bands, I have 256 branches. Parallel. Output here. But then on each of this I can send more or less bits depending on the kind of modulation scheme QAM that I have to use on that sub band. QAM depends on the SNR that I have on that sub band. The data bit I send on different sub band is different. Maybe in a sub band I have a **very good quality**, a very high quality so I can send a lot of bits, in another sub band I have a very low quality so there I can send few bits.

I skip because I always remind your faces during this discussion. I put this diagram only to remind you this concept. The behaviour of a QAM as a function of the SNR. If I have a good SNR, a **high SNR**, I can spend a **high modulation scheme**. 128 QAM. If I have a very low SNR I have to use a modulation scheme. That is less bits per pulse. Everything in this plot. The different branches we have in discrete multi-tone modulation are adapted to the specific SNR of the system in that sub band. This is implemented in a very well known algorithm.

By going more in details the idea is this one, the 2 bands dedicated to upstream and downstream are divided in sub carriers, also named discrete multi-tone. These are spaced by 4.3 kHz, it's a sort of **frequency division multiplexing**, every sub band is at a different frequency. Some tones are not used at all, some tones in this figure, in particular the ones here are not used because they measured that in all the copper wires of telephone lines here you have this **rapid decreasing of the SNR**. There is a frequency part in the copper wire not useable. Cannot be used, because of a very bad condition. So there are some tones not used, the other are all used, in every tone you use the right modulation scheme. So this is written here, independent sub channels are manipulated, it may not be used at all, and this is a very computation of the achievable data rate, if we have 25 sub channels in the upstream and 249 sub channels in the downstream if we can achieve this is the maximum, the amount of bits per second per hearts that I can put, in every channel, so **every channel is of 4 kHz**, 249 channels for 4 kHz for the maximum spectrum efficiency, if we have the maximum this is the data rate we can have in the downstream and this is the data rate we can have in the upstream.

That is familiar to the data rate we use when we have an ADSL. Sometimes this is increased, for 2 reasons, instead of using only the part of the downstream we use **the whole**, downstream overlapping with the upstream. If we increase this and do eco cancelation we can have much more bandwidth so an **higher data rate**. And second, a theoretical dimensioning because depends on this, instead specific on the specific copper wire, this could change by user or every user has his own efficiency depending on his SNR, depends on the **kind of copper wire** they have, on the distance, on the number of users sharing the same binder group. The algorithm, there is an algorithm computing the amount of bits I have to send on different sub carriers, this is well known algorithm, named **water filling**, it is inspired to a principle.

Let's assume we have a pool, this pool indeed in our case is the **inverse of the SNR**. It's SNR^{-1}. When the NSR is low we have a **good channel condition**, when the NSR is low. When the NSR is high we have bad channel condition. This algorithm computes the amount of bits so, in other words you have an amount of bits that you have to spend, the power that you have to spend, that is the water you have to put in this pool. You put the water by putting the water in this kind of pool, the idea is if I put this water here, I have much more water here, then the water that I have here, it's obvious. If you look at the **blue colour** there is less water than here. The algorithm computes much more bits to be sent here, an higher QAM than the bits that are put here, depends on the SNR.

It's implemented by the model to **fill the pool available** on that specific copper wire. And every copper wire in principle has a different pool shape. It is a quite advanced mechanism, not available some years ago, that's why ADSL has improved only after a while, the 2 ideas that were important in the evolution of ADSL were first go to the band dedicated to the telephone, first case use a **normal modulation scheme**, not so efficient, the classical modulation scheme occupying the whole band, next and this is what has given success to this kind of technology, use an approach in the family of the **multi-carrier modulations**. We will see, we said this in some lessons ago, the multi carrier modulations, all the modulations that use this simple concept, divide your spectrum band in sub bands, are very effective in all cases you have a channel that has a behaviour that changes in frequency.

Then this kind of modulation **had success**, thanks to this. This is for sure the case of the wireless channels, this is why in the wireless channels we use some modulations that are inspired to this one, note we have also to mention that in terms of processing the processing of this behaviour is quite simple. As I said it's only a **serial to parallel**, the amount of bits sent here now we known are derived by the water filling algorithm, that forces to be different on the basis of what can be sent to different sub carriers and then every block in this chain is a classical QAM modulation, maybe you can't read here. Here we have a frequency that is different because it's the frequency of sub carrier 1, 2 , n.

So **every QAM is centred** in its specific sub carrier. This is a comparison of the behaviour of the 2 modulations, if we see here the different characteristics are mainly in the first carrier-less amplitude. This ends the modulation scheme, something quite specific of this kind of architecture. Let's go the architecture that is easier to be understood.

A part the modulation scheme, how we send the bits in the system, obviously is important to understand what are the network elements in the system. This architecture was designed, is the one already available in case of telephone network, we repeat this, in that case the architecture was we have the different telephones, **interconnected by coppers**, sometimes collected in a binder group, at this network elements that we name cabinet or crab, in this elements these are collected together. To **implement an ADSL** we need some new elements and the new elements are the following. On one side there is an element named **ADSL termination unit**, ATU, and this ADSL termination unit has 2 possibilities. One is named ATU-R, where R stands for remote side. ATU-R is what we call the **modem**.

ADSL Termination Unit

Our modem at home, the right name is **ATU-R.** R is remote side. Why we have a remote side, because the modems are 2, one is at the remote side, another identical device is at the central office side, named **ATU-C**, where the C stands for central.

We have for every copper wire 2 modems, 2 ATU, ADSL termination unit, one at home and one in the central office. This is quite particular, because given the fact that you known how many ADSL cable we have remind that **we have 2 devices on the 2 sides**.

The one that is most critical in some words is the one at the central office, because at the home obviously we for sure if we have a home, a office we have the space to host a modem, at office or at home. On the contrary if you look a the central office side the space to host the modem is **proportional to the number of subscribers** at that central office. The number of modems in your home is one, at the central office is 1000? 2000? 100? Depends on the number of clients connected to that central office. This is why to be honest the fact to **rent an infrastructure** to other operators, I have to give space to new operators to put their modems. Maybe here we have the envelop around, while in the central office is not done in this way, is something that has a space, you need some space.

The connection, the power and so on. Also the **power is an aspect**. You have to put the power for every device and give the keys for the operator, to open the cabinet, put the modem, in the right place, not disconnect the others and so on. This is something that needs **some regulations**. As needed, still need some regulations. These 2 are in the 2 sides, ATU-R, ATU-C, then there is another element, named **DSLAM**, DSL access multiplexer, a multiplexer that multiplexes the bits coming from different ADSL users. Before the DSLAM we had an element that is again specular to the 2 parts that is named **splitter**. A splitter is available at the home of the user, another splitter is available at the central office.

Again, **2 splitters for every copper wire**, each copper wire has 2 modems and 2 splitters. I always ask this during the oral and students put a lot of modem only at the user side. 2 splitters and 2 modems both sides. Splitters, they s**plit the frequency band**, they separate the telephone from the data part. The analog signals to the digital one. This is a sort of integrated services infrastructure.

Today to the voice sometimes it happens that it transferred by how do we transfer the voice typically today? **Over IP**. So this means the separation of the 2 services is less important, voice transferred over IP and IP used for the rest of the data. The differentiation is done at the application level. A service is the voice and another service is all the other applications. Here is done at lower levels, not application. We can have a break and then continue.

As I said let's briefly recall all the elements, what they do and what they are, in principle there were two kind of services, **voice and data**, the voice was designed previously in that architecture so the new architecture keeps the voice working **transparently** with respect to the ADSL, the voice was already there, the ADSL built the rest of the infrastructure and the two infrastructures are separated from the other. For this separation we need, since in the cable is done using **different frequencies**, in the network element this separation is done using the **splitters**, there is the splitter keeping the voice direct to the voice device, and the splitter keeping the data to be sent to the ATU, on the side at central office as well as for the remote element.

This is written what they do, the ADSL transmission unit at the remote side, the ADSL transmission unit at the central office side, at central office is also this **measurement of the quality of the wire** in order to understand how to implement the water falling. Remind also this, a modulation, must be **agreed on both sides**, on one side the transmitter has to decide how many bits I put here, which is the modulation scheme as transmitter and the same must be implemented by the receiver, so the receiver must know which is the **rate of bits** that are sent on a different quadrature amplitude modulation and where to read the symbols that are transmitted, If I transmit 4 symbols I have to pick up the point of these symbols, if I transmit 16 symbols I have to pick up the right points, the 2 must agree on the **modulation scheme**, there is a negotiation that is done at the beginning of the communication or at the setup of the ADSL, depends on the specific modem, some modems can do once at the setup, they **configure the line at the beginning**, other modems continue to measure and adapt the modulation scheme depending on the specific quality.

At the central office side there is the ATU-C that was implemented by having the modem directly connected to the line card of the digital subscriber line multiplexer, there is this multiplexer with different line cards that are the inputs multiplexed toward the data network. So this was already explained and then there are the different devices, there is this picture that represents a snapshot of a picture taken at a network unit at the central office side, here for instance what can be seen is the **network unit** where we have these are 4 customers, and the 4 splitters, so you have the splitters of the 4 customers, at the central office if I have 4 end users I have 4 splitters and 4 modems, the modem are all put one close to the other in the central office, there is this **battery of modems**, a new operator providing the access and put the line card connected at that point.

The **DSLAM,** has the primary function to multiplex and demultiplex the traffic to host the ATU-C interface, there are all the interfaces on the different ATU at the DSLAM and has the operations of managing the whole platform, it's an important element it has the need of obviously the power so needs to be powered and needs to **have space**, because as I said it has to host the different modems and power because it's an active element, sometimes especially when we would like to increase the bandwidth of ADSL we need o shortener this part of the network, we can **put the DSLAM in a cabinet,** the problem anyway is that in that cabinet you need the power we need to have an environment that is closed, can't be broken and so on. However it's a solution that is available.

I don't know if you notice there are also around Rome a lot of new cabinets. Do you see them around? If you have a look around there are **new grey cabinets**, that are mainly for the **fiber optics**. Totems, a sort of totem. Grey. These are for the next generation, for instance the VDSL, where the cabinet is used to put the DSLAM close to the end user, where the fiber terminates, and then the last part is made of copper. Try to take a picture if you find a cabinet and share in our google group.

We have 3 solutions to **implement the splitter**. Indeed the splitter is a **mix of a low pass and an high pass**, it's a low pass because filters only the low part of the spectrum, a low pass where as for the voice and instead is an high pass for the rest, you see from this picture the very short bandwidth for voice is low pass, very high part of the spectrum is the high pass dedicated to the data, sent to the internet device. Obviously in this part of the architecture is not mentioned the fact at the modem we can have **other kind of functions**, for instance routing functions, in this case the modem is the ending part as for the normal network. The modem is owned by the operator.

The network is terminated by the operator at the modem, the modem is controlled by the operator, who is able to connect the modem to give to the modem the right to for instance use the public internet and is defined by the operator, who gives the public addresses to the modem. That part is the **edge** between the private and the public, the operator. The modem at the remote side. **Splitter**, can be implemented in different ways, the most used even if in the standard there were defined 3 possibilities, the most used is the first one, named **splittered**.

At the entrance of the home network we have the first element, the splitter. Typically is put integrated separately with respect to the ATU, but is a single device put at the network entrance, where the copper wires arrive. In the plug where the copper wire is you put the first element to the splitter, then part is sent to the rest, so part is sent to the telephone and the other part is sent to the rest of the client.

Initially there were other 2 solutions, another is **distributed splitters**. Instead of splitting the spectrum of the single point of entrance, the different bandwidths are splitted, they travel together in the whole network into the house and they are splitted separately at the different plugs. You see I have a splitter here, a splitter here, a plug here where I have the modem, I can split at a different plugs into the house, or a network where the splitter does not exist at all, named **splitter-less.**

Splitter-less is a version of ADSL **without having the splitter**. Maybe less cost, a lower cost, the fact there is no splitter has an advantage, the cost, and a **disadvantage**, that is what? If I do not split? Interference of the different bandwidths. Sometimes interference is present, especially in the first versions of the ADSL there was an event due to the signaling of the telephone network. Signaling with telephone network, you should know is the one sent during the setup of the call and during the teardown. In the setup we have a signaling that is the **composition of telephone number**, this signaling is sent. While the telephone was designed to control that the frequency band occupied by the voice is **below 4 kHz** nobody cared of the components in frequency of the signaling.

At that time it was not important, if there is some use of frequencies due to to signaling in a frequency band above 4 kHz it was not a problem, but when the ADSL arrived, sometimes when we were composing your telephone number there was an interference, so splitters were important mainly for this fact, there was an **interference**, a pick of interference due to signaling in the band used by the ADSL. However splitter-less is if there is much more interference we expect in this kind of ADSL a **lower data rate**, there is a lower SNR. Our rule main rule in this course is **high SNR, high data rate**, high bandwidth, high data rate. Low SNR low data rate, low bandwidth low data rate. 3-2 concepts.

In terms of **protocols**, this is a representation of the protocol stack designed at the beginning in the ADSL architecture, you have here the client on that part and the central office, obviously between the client and central office there is an **interface**, which are the protocols on this interface? The protocols are, is quite obvious that on this side from central office toward the network, the network is here and obviously everything else is here, everything ends at IP level, how it works doesn't matter, there is IP, again the same at the **client side**, we have the IP network. What is in the middle? How the interface in the middle is structured in terms of protocol?

The solution is this one, we have 2, this is the **physical layer**, where the modulation, everything is designed, this is the layer 2, designed with a technology named ATM, that works in this interface. On top of this layer 2 there is another layer named **ATM adaptation layer**, that has the behaviour to adapt an IP packet to the ATM, a different kind of protocol. The main difference is that ATM uses packets, named **cells**, in this technology, of the same size. So the packets, the data units of ATM are **fixed size**. While the data units of internet or IP are of a fixed size or not? IP? variable. So IP has a packet with variable size, maximum but variable. ATM has fixed packets, so the role of this adaption is to adapt an IP packet to a fixed packet.

One important function implemented here. On top of this there is another protocol, **PPP**, point to point protocol. On top there is the classical IP protocol. So we have

IP, PPP, AAL5, ATM and ADSL on physical layer. This was implemented in this interface. At the user side this is quite old but there are different possibilities. At the beginning ADSL was designed to have also devices to use ATM, or to have devices communicating using internet, or devices using internet and communicating by using IP. This is at the user side, we can have local area network, internet based or a network into the house IP based like typically is done today, consider this internet based could be also for instance the case when we use a wireless internet, wifi. So it doesn't matter, it's important that on this side we can have other interfaces, the meaning is the interface I have from my pc to my modem is one technology, if I communicate with my modem is wifi, if I communicate with internet sometimes is not used anymore but in the beginning have you ever used an **internet cable** to connect your computer to the modem? So at the beginning we used also the internet cable to interconnect the pc to the mode, today typically we use wifi. Other technologies?

The evolution of ADSL is the very high data rate. We have listed some data rates, they depend to the distance. If I have a distance of the central office, where the digital subscriber line multiplexer is, where the ATU-C, is the part where **the copper** is. The distance is where the copper wire ends and where the fiber optic starts. This makes the **difference** in the overall data rate we have. If you remember we mentioned the fact that fiber to the home, fiber to the building, etc is a way to indicate where the fiber arrives, so if the fiber arrives to a network element named optical network unit, this optical network unit is the element that can be in a cabinet, in a totem we said before, in the building, or in the central office, this makes **different architecture.**

The VDSL has an optical network unit that is **closer** to the end user with respect to the classical network unit that is at the central office, the basic architecture was the fiber arrives to the central office, the VDSL is well try to shorten the path of copper that is between the end user and the network and this is implemented in this way for instance, if we have 1000 m we can increase the data rate, if we have 300 m the data rate increases much more, you see here. This is only because the **quality of the signal is better**, we known the length of the cable gives raise to higher quality and we can keep higher the data rate.

Nothing more than this, a typical configuration can be this one, we can have a cabinet, in some of these we have the VDSL termination unit, of the fiber in the optical network unit, then the last part is done having distribution cables implemented by using copper wires. At the end this is the a little change in the architecture, thanks to this change we are able to keep the **speed higher**. Obviously this can be also designed by using higher frequencies so in some solutions depends again on the kind of modulation that we use, we can use higher spectrum bands so you see here and we can use spectrum bands in different ways, there's the possibility to have upstream and downstream separated or we can have them combined and then be able to differentiate the 2 directions.

This one is a bit **overlapped with the basic ADSL** but uses higher frequencies. So this is a comparison of the different architecture, these are for instance the oldest standards defined in the xDSL family, with different names and speeds and different time periods. The initial data rate was high but however today these are numbers that look a little bit low for the speeds. Another break.

PPP protocol

I try to explain without having the slide. PPP is an important protocol because it's designed to **work in all interfaces** between 2 point to point elements. The advantage of this protocol is that it works in a way that is compatible to different lower layer protocols. The idea is that, in general when you have IP you don't know exactly what kind of subnetwork you have. You can have ATM or other technology, the idea of PPP is to work **like a classical protocol** to work below IP, the classical is internet, but with some functions less important than internet because designed to work in a point to point.

What defines this protocols is the format of data unit and the kind of interactions that can happen in the interface controlled by this protocol. The **format of the data unit** the idea is this one, you have a frame a PPP frame that is identical to an internet frame. With some differences, the differences are mainly first that is obvious, target as PPP instead of internet, second, since it's a frame exchanged **on an interface that has only point of ingress** and an output point, this internet frame does not need the internet address. In the IP packet you incapsulate the IP packet in an internet frame. You have to provide the MAC address of the internet frame, of the device where the internet frame has to go.

You are transmitting here, for sure there is another one, there is **no need of MAC address.** The PPP protocol is designed in every case you have only a point to point connection, if you interconnect a cable to exchange data from a pc to another pc in that case you don't need a MAC address, you need to use the PPP protocol, this is one characteristics, the rest is very similar in terms of **data units**, very similar to the internet frame, the addresses, and then the rest that is the name of the protocol and then a variable length payload, the control at the end and the flag ending the frame.

This is for **encapsulation**. The operations performed by this protocol are mainly sublayers in this protocol. A layer named **link control protocol** and a layer named **network control protocol.** Link control is the lower part and has the role of negotiating the parameters used during the communication, authenticate the end user and perform the control of the link, if you have to communicate using PPP first thing you have to do is to **establish the connection**, when you communicate with your modem with another modem, ATU-R with ATU-C you establish the connection, during this connection establishment you first negotiate the **parameters** of the link, then authenticate so there is an **authentication**, you know, in different ways.

And control the status of the link, if it fails during the communication this is something revealed at the link control layer of the PPP protocol. Instead, at the network control layer this part is the one **interfaced with upper protocols**, the upper protocols can be IP or other layer protocols, so in the past for instance there were other solutions for the layer, the PPP was designed to be able to interface also with this kind of protocols. What is the role of this network part with the upper part is also to negotiate the IP address used by the end user, during the setup of communication, if you notice, there is a **dynamic IP address** assigned to the end device, by the control unit, to the network by implementing some exchange by using PPP, as for the authentication this can be done by using a classical username and password, password authentication protocol or using a challenge handshake. In the first case you have to provide the password to access the ADSL, in the other case you have to provide an answer to a challenge that is sent by the remote unit, you send back the response and then you are accepted or rejected by the central office side.

The first case is named **password authentication protocol**, the rest is cheap. Ok, in the configuration we can configure the parameters used to navigate in the internet like the IP address but also other parameters like for instance the quality, kind of service to have or if you can configure this at the remote side. There are some quality parameters that you can configure.

New technologies graph

I pickup this slide because we didn’t discuss about this. I present this figure that is presents the **technologies** in the quite wide spread framework and this graph is presented in this figure. To read this indicates technologies, you see there are the technologies here, technologies that are rising, new technologies that are triggering.

At the beginning there are technologies that are rising, then there is the **pick** of these technologies, on the rise, left part, at the pick, then there is the sliding into the through of disillusionment, a period where some technologies were very promising, maybe there are **decreasing now**. Then something starts to have success, so the first part is related to the research, then there is the pick, a very big interest in that kind of technologies and something goes **toward the market** and toward the plateau of productivity. This is a quite old plot, every year all companies that work in this kind of area, let's say the **information technology area**, but not also, they put their result in this graph, every year I present the graph. This is for 2017. We can notice here that what is important is to see which are the rising technologies, you see there are a lot of them, some of them are very familiar with things discussed in this course, like **5G, edge computing**, smart robots, other are related to other kind of subjects.

We can notice at the pick for the time being, for 2017, we have connected home, deep learning, **autonomous vehicles**, in the left part but close to the pick area there are IoT platforms. You see commercial UAVs, drones, are now in this part. We will see what will happen in the next. This is interesting also because sometimes I look at this graph to propose in that year the projects for the students, only to let you have the projects. Obviously we don't do startup or things like this, we are **not in the rising part**, we will like to experiment and see what happens in the pick. We don't know what happens in the next. It's interesting for you next years to look at this also in the next years. The idea is to combine autonomous vehicles with the possibility to interconnect them with a 5G infrastructure, so the 3 technologies are 5G, autonomous vehicles and machine learning, we try to combine these 3 to have a PhD program for the next students in 2 years from now. I'm quite happy we are thinking at things in this kind of technologies that are here reported.

Ethernet and carrier ethernet

For our lesson today coming back to the contents of the course, last week we ended with the ADSL and its evolution if you remind what we have done, today we start to discuss another technology that again was very interesting some years ago, now is becoming again in a contest I will show you but it's important to know how this works. The name of this solution is **carrier ethernet**. Now I will present a little bit about ethernet, instead we have to understand what happens in the use of ethernet as a carrier, this is the name of the platform.

The idea is to use ethernet not only as typically has been designed as a layer to technology for a medium access control but for **other purposes**. The study started several years ago in this framework, were oriented in the possibility to build MAN and also WAN by using only ethernet. While the most of you know that ethernet is the standard used in the **local area networks**, the concept is how is it possible to use this kind of technology as a solution for a layer 2, a broadband and so on, replacing other kind of layer 2, we will see some of them in the next like this SONET/SDH, these are some layer 2 proposed to be honest for the **telephone network** that are used and replaced by other technologies in the core network, not local area. Is it possible to use this technology **also in the wide area network?** We will see how it is possible.

Also because the technology that was present for mainly the telephone network, is this SONET/SDH that was a technology **designed for the telephone.** We discussed the telephone was a solution with static multiplexing, designed for transporting bytes in the format of a frame and then multiple frame are multiplexed and so on, so it's very suitable design for telephone network but not so flexible for other kind of solutions. **Ethernet** is an alternative also because as we know it is **IP friendly**. While in the past there was this kind of technology, we have to discuss this, in the core network, this is a replacement in the core of what has been done at the edge as for the telephone, you build up frames, you have the slots for the data, for the signaling and so on. The idea is can we replace with a **technology more flexible** open to multiple services? While in the middle of the 2 there were some tentatives with other kind of technologies like ATM what has been proposed is trying to use ethernet as this kind of technology.

Because first I would like to mention this as first aspect, it's IP friendly, we know this fully matches the way of operating of IP, it's true for sure when we use ethernet in case of a local area network. Ethernet is the layer 2 for IP. So the idea is can we use this **also for the NEXT?** Second, it's very easy to be managed by a network administrator, how to configure, to manage and how to operate this kind of network is easy and also the idea to use this technology as for a carrier is this one, during the years the ethernet has evolved very fast in terms of speed. The very initial **standard was 10 Mbps**, then it evolved to 100 Mbps, not also, now we have a lot of ethernet to 1 Gbps, going to **direction of 10 Gbps.** So this kind of technology is designed to work at very high speeds. These kinds of high speeds are not so necessary only in a local area network but can be of interest in case of an edge network, more in case of the metropolitan area network, more in case of a wide area network. There is the interest.

Indeed as I said ethernet was born several years ago **by Xerox** and this kind of standard has evolved from the basic solution. This kind of solution initially was based on the 10BASE-T. 10 stands for 10 Mbps, base because works in a **baseband communication**, t stands for twisted pair, that was the kind of media used to support this technology. This technology was designed for local area network so this solution was simple to be implemented, the idea is to have variable sizes of packets in the terminology of ethernet are named frames, composed in a given way and on top of this IP packets can be transmitted. After a while it happened that there was a problem related to the specific kind of access adopted in this technologies. Is named carrier sensing multiple access collision detection.

CSMA/CD is a way to access a shared media by multiple users by sensing the carrier, hearing if there are other transmissions on that media. Multiple users would like to transmit on that media. CD is the fact that the device continues to listen to the medium during the communication to check if there are other transmission contemporary to its transmission. If there are other transmission the transmission is interrupted because there is a collision so the transmission is interrupted to avoid to continue to use the medium. You use the bandwidth without transmitting useful information. Very simple but the problem is that was used in the **local area network** that represents a collision domain, an area where multiple transmissions by different for instance terminals and systems, devices in the same area can rise. In the same collision domain can generate collisions.

Also because this kind of shared media is always a media where a transmission happens in a **broadcast manner**, every device transmitting in that media is received by all the other devices, it's a broadcast. So, to try to solve this problem the problem of collisions the standard proposed to use another kind of solution that was by adopting some devices named switch, also bridge, today they are typically switch, that **differentiate the collisions domains**. And so you see a single broadcast domain where multiple devices can collide thanks to the switch are separated, in this way a collision domain remains a place where collisions can happen but only by a subset of user.

This is done having these device, **switches**, which operate in a way that is this one, if you have a transmission on a broadcast domain the transmission goes to everyone. The two broadcast domains are separated by the switch, if a frame is received by the switch and the frame is directed to devices that are in this collision the switch has no necessity to transmit the frame in another collision domain. The transmission has already reached the device. On the contrary if the switch understands, we will see how, that the frame is directed to devices in another collision domain then the switch operates for forwarding that frame in another collision domain. To understand this the switch has to look at a table, **switching table**, not a routing table it's different, this table says look at the address that is named **MAC address** of the frame and see if this address is toward a port relevant to another collision domain or it is of a port that is already into the incoming collision domain.

There is a table to look at the addresses at the MAC level, the layer 2, so addresses at the layer 2 if they are into the table these are used to forward the frame in the next collision domain, these are separated. I will do this lesson, tomorrow in the bachelor degree. Said this, thanks to the use of the switch, thanks to the fact that the links that are interconnect a user to the switch are full duplex, this means that a user transmits in a direction and in the other contemporary **without interfering**. Thanks to these 2 points the problem related as the collision domain is overcome. And thanks to this let's say that the ethernet is **no longer limited in length**. So this fact that the length is also another important aspect. Till the advent of the switch ethernet was only used in a local area, due to the fact that the size of the area and the number of devices connected was **limiting the performance** of the system. If I have a very wide collision domain it is not feasible to use that technology, ethernet, as it was designed previously because the collision domain was a **bottle neck**. Solved this fact using the switch, gave the possibility to extend the length of a network built by using ethernet.

The switch does is this one, there are input and output ports, it happens that for instance from an input port a frame arrives, the frame arrives at the switch, it by looking at the switching table is able to understand which is the port where the frame must be sent and then **interconnects** the input port to the correct output port and performs CSMA in the next collision domain. Before the switch there were some elements similar to the switch conceptually, but they were not aware of the right port where to send the frame, secondly they, given this, the idea is that if a packet, a frame arrives on a input port this is sent toward the output port. This hub was not able to separate collision domains. An hub does not perform buffering and collisions can occur for packets, frames, on the whole set of ports. In case of switch collisions are controlled thanks to only **one port is selected as output**.

Thanks to the use of a switch it has been extended the possibility to use this technology in a **wider area**. There were some other problems, one was the problem of loops that can be in a network that has been built by using this technologies, and indeed in this kind of devices, bridges or switches, there was a protocol designed to avoid the loops, this protocol was able after some iterations on different devices to build up a sort of tree on top of the existing network, dynamic, so it's not a routing tree, even if it's named **spanning tree**, not a routing tree, is a tree to avoid loops during the transmission in the network. Thanks to this you can have a real network even if this network has not a routing mechanism, it's a simpler network because you have seen that routing is a different thing, the routing what does it do?

This operation is **not routing.** Why? Which is the different of this with respect to the routing, if you see **differences**? A router has a set of input ports and output ports? Yes. Has buffers? Yes. So? A **routing table**. Here is named switching table. Table is common to both, they are similar or different? This can be a question for the oral. Be prepare to answer these kind of questions at the oral. In the switching table the address is a MAC address, layer 2, in the routing table the address is an IP address, layer 3. This is still not a real difference.

The **real difference** is that the address is controlled to check where by reading the routing table the destination is. Where the destination is is reached by adopting routing algorithms used to fill the routing table. Thanks to this the router is able to solve, to interconnect an input port to an output port, that one is the best destination to this packet. There are **routing algorithms behind**. While in this case we only have to understand if the frame, our packet in this case, is directed toward another collision domain with respect to the one from where it is arriving. And where these other collision domain is. It doesn't matter how many hopes, the only fact is that is connected toward this output port or this output port.

On the contrary the **spanning tree** is not again used to find the best path toward a given destination but is a way to avoid that for instance this bridge sends a packet to this bridge and this bridge sends the packet to this bridge and so on. This is a loop, is to **avoid loops**, a different thing. This is also very important because if you don't use this then the performance of the network decreases because you see packets, frames traveling continuously into the network. In this technology it may happen that frames continue traveling through the network is **not controlled by the ethernet** as it is while in some ways this can be controlled at the IP level, you remember how? By using the **time to live**. In ethernet it's not included.

This is only a representation on how the ethernet is evolving, we will discuss, instead is important because some of you don't know, how the ethernet frame has been defined. So as we said ethernet frame has variable size, and is built by a combining **control fields and a payload**. The control fields are this, there is a preamble, well known pattern of bits, used by the devices to understand that a packet is traveling. The preamble is used to understand that a packet, a frame is starting bytes. Then there are source and destination address, composed by 6 bytes and represent the MAC address. Then there is a **message type**, specifying which kind of message is transferring this data and there is the data field that is variable and can include a maximum of 1500 bytes. It is a variable length of a size that has a maximum.

Even there is maximum length but there is **also a minimum length**. We will see in a while why this minimum length is also important. There is a max size, 1500 bytes, and a minimum frame that is 64 bytes. Finally in the end there are 4 bytes dedicated to the **check of the bits**, the control of the bits, the **CRC**. This is how the frame is built. Again, could be compared to other kind of variable packets that are used in other kinds of technologies. The important fact is that there is this minimum size. As for the protocol stack, that has been defined for an ethernet device is this one, there is the medium and there is a physical medium attached, the unit interconnected can be for instance the port that we have in our laptop, there is this **medium attachment** and then there is the interface of this port toward the rest. So the rest is a sublayer for the physical signaling used to have the synchronisation with the frames that are arriving, then there is the MAC layer.

The core is in this part, the intelligent part is this one. To use this part there is the rest to have the interconnection to the physical medium considered as for the signaling. This is a well known standard that **overcame other standards** presented at that time for the local area networks, when ethernet was born there were 2 possibles paths, one was using ethernet that is collision based very simple medium access control and the other one was a controlled based medium access control named token ring. **Ethernet was faster**. Mainly given the fact that it was easier to be implemented it received a very high attention and then became the standard for this kind of technology. After a while fast ethernet arrived and this was in the same bit rate of other technologies that were named FDDI and ATM. 155 kbps. Higher rates. And after a while there was fast ethernet and there was also an ATM at a higher data rate, 622 kbps and again thanks to an evolution of the standard also ethernet was also able to **overcome that technology**. So this is the fact that this confirms the fact that this ethernet was born with a given target market, LAN, then increased and was able to be compared to technologies like ATM that were born for other purposes, with the idea to be ready for a WAN, or a MAN. Let's see the first step. **Break.**

We were presenting ethernet as the original standard. After the high success of this standard for the LAN the standard has evolved also in other directions, the first direction is the fact that starting from the 10 Mbps the evolution was to provide a 100 Mbps. **Extending the original standards**, because the medium were evolving so the possibility to have physical layer that were again twisted pairs, copper wires but used to transmit at fast rate, or the possibility to user fiber optics, were able to support higher data rates. In order to include this part of the standard you see the **protocol stack changed a little bit**, the main changing are in the lower part because in the lower part we need to be able to adapt our transmission to the new media, the new media for instance fibers or twisted pair at higher capabilities, this is also important because you need to have a unique interface able to understand where it works, if you see several times the ports of ethernet in the switches, if you have ever seen a switch, they have a ports where they have 10/100, that port can work in both cases, for **10 and 100 Mbps.**

There is a part that is typically of the physical layer in order to adapt the transmission to that kind of speed. On the contrary the fact that the system improves the capacity is also determines some changing in cases of the times of the CSMA/CD. The fact that the length of the frame has as a **minimum** depends on the behaviour of the collision detection. CD is the way to continue to hear, to listen to the medium **during the transmission**. This listening must be done for a minimum time, why? Because the design of this protocol, CSMA, has to take into account the fact that when you transmit on a broadcast domain, there are physical finite time, there is a physical finite time to reach a device due to the **physical propagation** of the signal on the medium.

What is true is that a signal when is transmitted on a medium **doesn't propagate immediately**, not instant, but takes some time. This time is named propagation time. There is a propagation time, also named propagation delay, the time taken from the signal transmitted on that medium to propagate in the broadcast domain, on the medium. This transmission time is the one that in initial ethernet has conditioned the way the frame length was designed. Indeed, the time needed to transmit a frame must be **not less than the round trip delay**. When you transmit a frame in a broadcast domain and you suppose to continue to hear your transmission the duration of transmission itself must last a time greater than 2 times the propagation delay. That is also named **round trip time, RTT**. If I transmit a frame in my transmission must last at least 2 times the propagation delay. RTT.

Why? If I transmit something and my transmission lasts less than this, something that lasts less than 2 time the propagation delay my transmission I will not know if the transmission has a collision, had a collision or not. In other words, I have to speak to transmit at a minimum length of frame. So the length of frame L must be greater than or equal to the round trip time multiplied by the rate on the medium.

$$L \geq 2t_p * Rate.$$

Let's consider a possible length of collision domain. The worst RTT in that CD, let's fix the data rate, 10 Mbps, the product of these 2 was 64 bytes. This is the minimum length of the frame size. The minimum frame size. Said this, **till the data rate was 10 Mbps** everything works. The idea to change the interface toward 100 Mbps that could be simple, yes, I put another cable, this cable goes at faster speed, is only a good new, but this cannot be achieved with the previous standard unless you change the frame size, if you look at that formula if you increase the rate and keep the frame size this equality is not anymore true. You have to change something.

2 solutions: one solution is I can change, if I'd like to increase the data rate and keep this how it was I can do, how? **Reduce the RTT**. How can I do that? Reduce the length of the collision domain. Typically a project of the network like this is not a good project. But this was a solution. Or, **fast ethernet**, let's define a new length of the collision domain, of the maximum length of the network was reduced to about 200 m.

In other words, fast ethernet, continues to use the same standard of before, works at **higher speed**, from 10 to 100 Mbps, but the length of my collision domain is reduced, in the first standard was 2 km, now we are going to 200 m. So, the evolution of the standard was only in the direction of using the previous one, as I said you can use both and then negotiate and verify which is the data rate of the physical medium and adapt the rest to that physical medium. If you have a physical medium at 10 Mbps you can reach have a collision domain of length about 2 km, if you have a physical medium of 100 Mbps you have a length of the network of about 200 m. So this first step was quite easy with not so much modification in the standard. The **other modification were a little bit harder**. Why?

Let's continue the discussion. And try to go in the direction of increasing again the speed of our medium. Instead of having 100 Mbps **let's have 1 Gbps**. If I do again the same thinking of before I decrease the length of the network. From 2 km, 200 m, the next pass is to **cm maybe**. Could be a network of 2 cm, maybe is not useful at all. This time this **was not possible**. We would like to support 1000 Mbps, the technology in terms of transmission were there, I have to provide a new way of thinking. So there were a lot of media, you see fiber optic, twisted pair, able to transmit at that speed on a given maximum. In order to do this again, changing at the physical layer but more important the **changing are at the MAC layer**. What does this MAC?

First, again can be shared switched or hub based. If we use **switch** as said before the problem of the collision domain is less important. We have to put switches, there is a cost of the network, if I put switches I can split my collision domain, have shorter collision domains and this is important. If I would like to continue to use **hub**, where the collisions could be then the problem is different. I would like to introduce a concept, as I said the story was ethernet as the standard for the LAN, then the studies came in the direction of having ethernet for a technology for metropolitan and wide area network. In the meanwhile today there is a very big interest in the use of ethernet again at fast speed in a sort of local area. This sort of local area is named data centre.

A **data centre** that is a network, works in a local area, may occupy spaces in a range of hundred of meters, not so much, they need a very fast flexible, easy to be used and so on network. A **network infrastructure**. The best technology to do this today is ethernet. All the studies of this kind of solution even if they start with another purpose, ethernet in metropolitan and wide, today are useful in the context of data centres. **Layer 2 interconnected**, all the data centres are interconnected in layer 2. Why in your opinion is better to interconnect layer 2 and not at higher layers? This is only to say there is still very important to understand the use of this technology. If you want to use switch you can use **full duplex**, for hub you have to use the classical solution. With some changes, that are at the MAC layer. Which kind of changes? If you would like only to increase the data rate and keep the rest as it is the rest is I have to support a network of 20 m. Not of interest, unless in some cases.

The idea is this one, let's change in the standard our frame length, instead of having 64 bytes, let's have a frame minimum, named **slot time, of 512 bytes**. The change is this one, the slot time, our minimum frame length goes from 512 bits to 512 bytes. It has increased a lot. But as always I will like to have interfaces that are able to work in a network that can be 10, 100, 1000 Mbps, I don't want to change the whole, otherwise is another technology. I want the same technology and **being dynamic adapted** to the specific context. To do this the problem is the **backward compatibility** with the previous one. How this compatibility is solved in ethernet in a gigabit ethernet?

The problem is this one, if I emit frames of 512 bytes I transmit my frame and everything works. In case in the same system there are devices that are using the **previous standard**, 64 bytes, I have to find out a way to accommodate also the transmission of the previous. The system must continue to work also in that case, a solution could be this, if my frame lasts 512 bytes and I have a frame that has a length that is less than 512, is 64, I put my frame old style in the new one, I have the old style frame, 64, the new one, 512, I put the **old one in the new one** and I fill the frame with some **dummy characters** named *carrier extensions* that are used only to fill the frame till 512 bytes. I put my content in the frame and the rest is a repetition of these characters. Into the frame.

Thanks to this with this slot time done in this way you are able to **achieve a 1 Gbps** based solution with a range of 200 m. This is very nice, it works, even if then obviously when you receive you can understand if the frame is of the old type or new type. The new kind occupies all the slot. The old kind occupies only the initial part. However there are **some problems** here, the behaviour of the protocol is correct, it continues to work as before, but, if there is a but, what do you see? Do you see some *problems* in this kind of configuration? So, this works, how it works? But carrier extension can be very **bandwidth inefficient**. Is it clear what does carrier extension mean?

Carrier is related to the medium, is like I continue to be active on the medium carrier, until the safe frame length is reached. 512 bytes. This was the solution, but not very efficient in term of bandwidth. If the old style in this kind of solution generate frames of 64 bytes if you see, have you ever used **Wireshark**, if you look at Wireshark phases you see a lot of times frames of 1050 bytes, the length of ethernet. If you have shorter frames then in this case you can have a very big waste of bandwidth. Depends on the length of the frame.

There are **2 solutions** at the problem created trying to solve another problem. This idea was to solve the previous problem, due to this idea the problem of bandwidth inefficient is here. How to solve this? **First way, do pipeline.** Quite intuitive, let's do this, instead of filling a frame of the new kind with a single frame of the old one, what I can do? Put inside this frame **multiple frames**. Simplest thing you can think. I combine multiple frames into the same frame, this approach is named **packet packing**. Means I do a packing of different frames into single block. This for sure can be very efficient. Instead of putting RRRRRRRR, that are dummy things, I put there other parts, if I have. There is a simple delay because I have to wait multiple packets before transmitting, this may be an impairment, I have to wait to have a number of frames that is multiple or equal to 512 bytes to be ready to transmit, so it can **add some delay**.

More important is this one, the medium access control layer designed in case of CSMA/CD, needs to **acknowledge frames** by the receiver if they receive correctly of not. The design was, I acknowledge a frame. In this way if in a frame we put a single frame if I do an ack is ok. I ack a frame, one frame. If in a frame of new style are hidden frames of the old style when I do the ack what are ack? One, two, more? It's not clear anymore. This is a **standard violation**. I'm not doing what the standard says, I have to ack one frame. Yes but if into the frame are multiple frames it's not clear what I'm ack. And then this is something that doesn't work very well due to this standard violation. Because there is no this status quo, the standard says you ack one frame. If you send this status of ack after one frame of the new kind there are things that **do not work as the standard wants**. Obviously these are things we are trying to solve before the standard was designed. Today we will design a standard in a different way, since the idea is to keep the standard as it is and provide backward compatibility when you try to fix something in the code you have **overhead** and things like this. So next step, was this one.

Next step is let's provide another way named **frame bursting**. The reason of waiting 512 bytes expressed in time is as we said to have knowledge that the carrier is occupied only by 1 user and be sure that collisions did not happen. So I do as before so I do I put a short frame into a time slot. And then fill the time slot with dummy bytes, but this is done only once for a time slot. After this time slot if everything went well so no collision happened I'm sure collisions are not present. In other words I capture the medium after a RTT, if after a RTT I'm sure **nobody is transmitting** I am sure I am for that frame the owner of the medium, after this I can transmit a repetition of frames that don't need to check the carrier. This is named frame bursting. I do as before for one frame, the next sequence of frames are sent one after the other because they are sure that collisions are not present.

To do this in the right way first I have to define **how long is the burst**, otherwise there is the risk the device takes the fully control of the medium, the standard says yeah you can transmit a burst of a maximum length. Next, if you look at this picture there is a problem. When you transmit subsequent frames one after the other there is a time gap between the 2 frames. So we have to control this time gap. Why? If you … after the first 512 bytes I'm sure that **I am occupying the medium**. I continue to transmit frames, and to be sure that the others do not occupy the channels how can I do? This is a solution used also in other technologies, we will see. Carrier sensitive multiple access says, listen to the carrier for a time period and see if in this time period there are no transmissions. The standards also regulates the time period you have to listen to the channel. If the distance of 2 frames is too high I can give the possibility to **another user** to transmit its frames.

So I don't want to give this possibility to others, the only way is to have the time distance between different frames less than the period used by the standard to state that the medium is free. So there is a period that is an **inter packet gap** that says after this period the frame is free. If I regulate my transmission of the burst at a distance that is less back to back within this period, I'm sure that nobody enters in my transmission. It's not easy but in practise you want that after having done all these having put this overhead, carrier extension to be sure the channel is of your property for a burst period you don't want the burst is interrupted by other transmissions. Other transmissions can interrupt only if they can here a time gap, a silence for a time period that is greater than this. If you are able to transmit at a **distance less than this gap** I'm talking about this gap. If I assure this gap is less than the time that is typically dedicated by a device to state that the channel is free I can assure that nobody enters in the transmission of mine. This is the frame bursting. The MAC transmits the carrier extension. The next frame is sent till the maximum length of the burst, set in the standard itself. We can do another break and go to the conclusion of this.

First, only a reduction of the length of the network, fast ethernet, then to support carrier ethernet, an increasing of the slot time. So, to go fast in this part the only last thing is this one, is that the problem once the network speed increases sometimes the bottle neck becomes the network nodes. So we can have fast medium, fast media where to transmit the data, but the bottle neck specifically an idea to have a fast element so if we can solve in a collision domain for instance the problems related to the collision the problem is if we can have also fast devices like for instance the hubs or the switches. There is an idea to **build up an hub** by having a sort let's look at this figure, a sort of short network, very short because it's in the order of cm, by having a buffered distributor, it is a bus where into this bus are exchanged very fast informations, the bus is very fast, by adopting CSMA/CD, the problem is how we can exchange the information in a switch very fast at the same speed these information are exchanged in the network? The solution can be for an hub to apply the concept of CSMA/CD, inside the hub on top of the bus that is available there. Thanks to this there are input ports where the frames are received, these frames are buffered in FIFO queues, these are exchanged on the bus toward a given destination that is a given output port in this case and these are exchanged by using a way of accessing that is CSMA/CD.

This is only a detail on how to build a very fast device able to be compliant with the whole architecture of the network. Next step again is to increase the bandwidth and the increase of the bandwidth is 10 Gbps, this time if you look at this solution all solutions **having half duplex** and requiring CSMA/CD are **avoided**, otherwise we need to decrease the size of the network. For this solution everything is done having full duplex switched networks, without collision domains at all. If we have full duplex switched networks we can combine **have no collisions**, so a medium access control can transmit when it wants because it's alone on the switch, there are no links on the switch anymore. All the interconnections of switches, this has a cost because in the previous case we use hub that are less cost on the contrary if we put switch **switches cost more.**

An infrastructure like this technology is much more costly because we need all switched device, but it's used when you want to have a big network very fast like for instance can be the data centre. This is a remind of the evolution, ethernet, fast ethernet, gigabit ethernet. The next aspect to be considered it not related to the speed but to other aspects. Once the technology arrived at the solution that ethernet can be used also as a solution for supporting **very fast communications**, there are other aspects that must be keep into account. That are the following: always having in mind that could be a solution for metropolitan area network, wide area network, there are things **included in the other standards** and not present in ethernet.

One of these, first aspect very important, the **quality of service.** In general you need to have a traffic differentiation. Because the carrier ethernet is an ethernet supported by a carrier, carrier this time means an operator, wants to use ethernet as technology to provide its own network. In a core network an operator has to **differentiate the traffic**, give priority, give business and residential users, high priority users and so on. So this is something needed, so first to have a carrier grade technology, must be fast. Second, must be able to support quality of service differentiation. Ethernet, if you see at this frame structure, look at this, typically when you like to do a traffic differentiation you need what? Devices able to give priority to some packets with respect to others and what else?

If I would like to start a technology able to differentiate the traffic I need to have a switch, a router, what you want, different queues where packets are put, different scheduling queues and so on but also a simple thing that is needed have a **field used to differentiate the traffic**. You see a field here? No. Do you know if this field is present for instance in IP? Yes, named **type of service**. How many bytes? 4 bytes. So, in IP even if this was not enough also for IP they defined a field named type of service to be ready to differentiate the traffic, indeed in IP if this type of service is all 0 means **best effort**. If otherwise the type of service is used for different classes you give different values.

In ethernet because was born for **other kind of purposes** this field was not present. But the interest in having this solution for having a sort of quality of service was present also in ethernet. This is a fact, speed the first, second quality of service, we need to implement some mechanisms into the ethernet, adding some protocols, some fields, to be ready to support a quality of service or a service differentiation. You know the difference between quality of service and service differentiation? Quality of service with respect to service differentiation. **Service differentiation** means, I have 3 services, 3 kinds of clients, I would like to give priority one of this, priority 2 to the other and priority 3 to the third. **Quality of service** is you need to have that bandwidth, that delay, that packet loss, that quality of service parameter. That is much more harder to be realised and also complex in terms of technology.

So **service differentiation is easier**, quality of service means to figure out the bandwidth and so on. This is the difference. However for an operator sometimes they speak of quality of service and they do service differentiation. For sure they do service differentiation. An operator wants to do, I want to use the network for all my customers but there are business clients, **pay more**, if they pay more they should have a higher quality, higher priority, if they pay less they are residential, the service is with less quality parameters. Because they pay less. How to do this in carrier ethernet? Always because there was this competition with other standard solutions that were born by having this kind of parameters and way to manage them into the standard itself. This can be skipped. Another aspect but it’s too much not complex but less important for us, the **maintenance**. Sometimes an operator has the need to control the behaviour of the network and make some maintenance operations. For instance switch off some devices, route the traffic with one direction, all these are maintenance operations.

A **differentiation** of the service can be provided for instance by using the virtual lan concept, VLAN. A VLAN is the idea to support on the same physical network different logical network. What is the difference between **physical and logical network?** Physical network for instance interconnects 10 users, in a logical networks maybe only 2 or 3 out of the 10 are interconnected. And also the topology of the logical network may be different from the topology of the physical one. Logical groups that are different from physical ones and logical topologies different from the physical one. A VLAN can be built by having a **logical segmentation** of the broadcast domain. Here we have an example, let's consider this example here, a switched network here, all switches and this is a broadcast domain because potentially a transmission from one user may reach all the other users, even if a switch potentially can transmit toward another switch, can transmit toward the end user.

So this is one **broadcast domain** and what is important in ethernet is that some messages for instance broadcast messages are broadcast with all the MAC address equal to 1, FFFF that is all 1, this may reach all the devices. These are broadcast messages sent for some operations. Not only this. One solution to separate the users is for instance to **insert routers**. A router can be configure to send a message toward a direction and not toward another, so it's configured to set a **routing path**. So thanks to the routing path concepts we can have a message sent toward this subnet, this user, this subnet, this user. This is done using routers. The advantage that you are able to achieve what you want, the disadvantage is that you **replace a switch with a router**, a switch costs an amount, a router costs more, a switch is layer 2 technology, a router is layer 3. An hub is layer 1 technology.

If we can put the router the cost is higher. Instead the solution proposed by a virtual LAN is to replace some switches with a new protocol that is named 802.1Q that is for having this service differentiation. This protocol works by adding an **extra label**, an extra header to the classical internet frame that is an header where you can find a field that is user priority and virtual lan identifier. So as I say before, I can create on top of my single physical lan 2, 3, 4, 5 virtual lans and any virtual lan will have its own identifier. Networks, virtual networks on top of the same network, the price to do that is that I have to add because it was not present in the original ethernet frame, to add an **extra header**, and to have switches able to process and to put and to remove the extra headers. So this one have different name with legacy switch, a classical basic switch, ok, the one that you buy I don't know somewhere and the 802.1Q is **advanced switch**, that has the capability to process what is written in this part and to add and remove.

So this scheme is named tagging scheme and maybe if you heard about **multi protocol label switching**, have you heard about this? Is a similar concept but in case of IP. In case of IP also they invented a protocol that is above IP used to add at the IP header an additive label, that is mainly used to fasten the switching the routing. To have a faster routing when you cross a routing table in IP. Another technology. This is similar but instead is done at layer 2. Thanks to this we can provide a solution like this. Where for instance all in this picture all the devices marked in red are devices that belong to the **same virtual LAN**. This means that it is like they have their own lan, but it's a virtual lan. The cost is to have this overhead that is the tag. Then what happens, for instance you see here you have a switch interconnected to 2 virtual LANs and other switch interconnected to another virtual LAN, in different domains, so this means that there are **switching ports**, ports of the switch that are grouped into different broadcast domains, there is a switching port related to the virtual LAN 1, another switching port dedicated to virtual LAN 2.

Then there is the crossing of this part, the part that is between 2 VLAN switch and in the link frames of the 2 different virtual LANs are transferred. While in a VLAN in this port only frames belonging to this VLAN are sent, in the interconnection part also packets frames with that tag can be exchanged. If for instance this switch does not support this virtual LAN this switch is not authorised to send packets, frames with VLAN identifier of this VLAN toward an **interconnection**. This is the example. VLAN switch has the tag to remove the tag if they exchange information of the trunk link interconnecting the 2 virtual lans. You see here for instance we have 1, 2, 3, 4, 5, 6 trunk links that are on the links interconnecting 802.1Q switches.

So these links are links where packets **can be exchanged**, on the contrary this link has packets with this identifier do not cross and are not recognised by this. This switch, a legacy one, is not able to transfer the information. And as for this there have been lot of studies and this is the nesting of the tags so you see 802.1Q has one VLAN identifier, when you want to have an extended network you can also have another standard to add another identifier and another one to have an identifier that is proposed for the next. Every step increasing the network and crossing different kinds of elements adds an **overhead**, but this is mandatory in case you would like to provide a complex structure maybe using ethernet. Only to let you know that we can build up around the basic ethernet structure a much more complicated one by improving the devices, the switches, and by having these devices enriching the packet format by having different nested packets, like here.

If you have different tags every tag has a role, every role can be used to **differentiate the LANs** and then for instance to differentiate the quality and the priority every time you have the tag there is an overhead but on the other side there is the possibility to go into the direction of improving the standard. I don't mention this because after this one there was also a solution that was PBB that provided a backbone bridge, designed for having also traffic engineering that is to plan not only the priority, also the path to be crossed in order to provide traffic engineering the backbone of the network.

Which is the point? Here there are some references. The point is, there was interest in using this simple technology for kind of infrastructures that **are not the typical one** where this technology was designed for, initially the ethernet was designed for local area networks, after this then we wanted to have faster local area network, then interest for providing solutions to be ready to **be carrier grade**, all the facilities needed by an operator. Speed, obviously, must be very fast, also quality of service, maintenance, traffic engineering, all this for the time being can be achieved also by using ethernet. And this way of operating today is of interest maybe with some changes for instance for managing some challenging infrastructure like cloud computing and data centres. This is the summary of this lesson. Deep and fast travel toward the evolution of ethernet. See you on next Monday.

Recap on last lecture

Last week we presented a technology to be used in the access as in the core and this is called ethernet. So if you remember last time we have analysed the use of **carrier ethernet** that is the evolution of ethernet in the context of becoming a real infrastructure for data communications, is a technology, a protocol able to work in the local area network, we want to use it also to work in the core network. We have seen only to briefly recap what we have done last week that this infrastructure, carrier ethernet, is of interest in case with the possibility you remember to support high data rates, service differentiation and operation and maintenance. We didn’t discuss the support of operation and maintenance that is another stuff important for the **core network behaviour**, but also for instance in case this kind of technology are used in the data centre.

In case of data centre is important to have a very fast technology able to support **service differentiation** and operation and maintenance. There is an evolution of the ethernet standard that is indicated as a way to give priority to some delay constrained applications in case of transmission of ethernet. Ethernet works applying a CSMA/CD, while the traffic is not differentiated. There is the idea to have a sort of **2 queues before transmitting a packet**, one queue is having priority to have that kind of applications to be transmitted faster with respect to the other. The last wired, next lessons we go to the wireless, the next wired technology is the one related to the optical networking.

Optical networking

We have already seen in case of optical networking the interest is very high because we are able to support **very high distances**, so there is the possibility to go very far from the central office and at a very high bitrate. While the XDSL family has some limitations in case of too far distances, or given the fact that the spectrum band has a behaviour that doesn't allow very high bitrates, in case of fiber optics we are able to support much more services due to the rate and the distance from the central office. So, going back to the architecture that we already know, the **architecture** is composed by a central office, a set of cables crossing an area of about 1.5 - 2 up to 5 km, toward the end systems. This is the physical architecture of telephone network as well as of the ADSL. For instance in these days I checked on my operator if it's possible for me to go to the fiber and I discovered there is a tool on the web, you give your address and this tool indicates you the **position of the cab** where the fiber optic ends, from my house the fiber optic has the cab a distance of 1.8 km. So the operator told me I cannot have the fiber.

I also discovered that I found out a presentation of Telecom Italia of 2015 with a **forecasting** of the broadband coverage in Italy, they were planning to have only 1 city in 2017 fully covered with fiber optic till home, FTTH. This city I guess is **Milan**. The idea is that for the time being only 1 city the FTTH is something that in Italy was planned only for one city in 2015. You know all the drawbacks related to the XDSL solutions. What's **next**? This is for instance the kind of infrastructure so this was taken from Telecom Italia, saying what we have already studied, we have the central office, the **DSLAM**, the multiplexer of different communication, then the distribution network done by using copper wires and this distribution network is implemented with some constraints due to the distance that is limited since the copper cannot support a signal above this distance.

The idea is **try to use fibers**, fibers they if you know are characterised by an **attenuation** measured in db over km that has this kind of behaviour, the attenuation of a fiber is not always low, we would like to have a low attenuation, so be able to reach very far distances. So this is the behaviour of the attenuation, db for km, the amount of db we loss at every km, we can notice there are some windows, related as window, in the wavelength domain where this attenuation has this kind of behaviour, significant for **2 reasons,** first is that as always we would like to have a channel that attenuates in a **flat manner**. We already discussed, a target for the right and efficient communication is to have a flat channel. Second, is also important that this attenuation is low, is better low than high, if it is flat is significant. Low and flat is better. We have 2 windows, if the fiber can operate at this wavelength, around 1500 nm, in these ranges, the frequency domain for the optical signals, wavelength is the equivalent of frequency in this case, we can have **very good communications**. Since there are fibers able to work in these windows, we can have very efficient transmissions.

And this is as for the so called **single mode fibers**. We don't go into details on how a fiber works, what we can analyse is that the kind of communication that happens in these fibers guarantees a good channel behaviour, if we operate in this windows, we can have a good channel behaviour and the result if you remember the initial lessons we have done we would like to **avoid** as much as possible **the dispersion** of a pulse, the dispersion of a pulse implies an inter symbol interference, that implies interference that means a low SNR and a low data rate, everything related to this. If we can keep the symbol as close to the initial shape we can transmit at a very high bitrate. Transmitting on a given fiber, operating in that wavelength windows and after a while we can reach very far distances due to the fact we are able to transmit our pulses with a short, a **low loss**. We don't enter into this, I skip also this, the idea is that instead of having a situation where the fiber optic ends at the central office, the idea is to go to have fiber optic ending at the cab. We have seen this concept.

This means we have to work on fiber first and also the DSLAM. We need the **space** for the DSLAM and also the **digging** to support the fiber optic. In this picture there is an analysis on how there are different systems, how a fiber is presented, is put in a road for instance. This is fiber to the cabinet, I don't know if you have checked around any cabinet. I just looked around in these days and I noticed there are much more respect to the past. I put a picture in original slides, every time I mention this cabinet I see strange faces. These are the **cabinets** we are mentioning. If you look around is plenty of these. Inside there is the space for all the DSLAMs and the cables of course. We can end to the cabinet, we can end to the curb, we can end to the building. Or we can end to the home, in these pictures we have the different architectures to the **FTTx**, this x can be cabinet, curb, the building or can be the home. So, how the fiber is **interconnected** to the user? This is a qualitative representation, we see a single cable interconnecting all users.

How this is done practically? In **3 different ways**. The first way is exactly the same topology we have in case of ADSL, the following: I have the central office and for instance we have now speaking of architecture where the fiber arrives to the home or to the building, the very ending part of the system. In the first case all single users, all single buildings have fiber toward the central office, in this case this topology is identical to the topology we have in case of ADSL. Here you recognise exactly the **same topology of ADSL**. Second solution, I have the central office, a curb and the central office and the curb are interconnected by a **single fiber**, only one fiber and since fibers they have very high bitrate this single fiber can support multiple transmissions, high data rates, then there is this curb that operates like a switch, switching the communication toward different fibers, interconnected to different buildings. Last solution is this one, I have again a single fiber and this fiber is split toward the different houses buildings.

First case one fiber for each building, second case one fiber toward the curb and the curb in this case for different fibers, in this case one fiber toward a point that is named splitter and then 4 different fibers toward every end user. What is the **difference**? A part from the topology, in terms of cost for instance, which is better, which is worst. In the first case there is an high cost for the digging. In the third case we have a passive splitter. What is a **passive splitter?** Something represented in this picture. 4 fibers, fibers are made by using glass. The optical splitter is 4 pieces of fiber, 4, fused, you put some heat here and you fuse together the glass. So, if the cost of fiber is very very low, the cost of this is as well is very very low, is a composition of 4 fibers **fused together**. Maybe this part you see here is fused in the way that the optical signal, in a fiber we transmit by using a laser and optical signal, so the optical signal thanks to this solution doesn't go in this direction but instead is physically split, we can say replicated but it's not a replica, only passes toward 2 different directions, we are talking about a light. A light that enters in this tunnel and then exits from these 2 directions, this is an optical splitter, doesn't do **anything in an active way.** This is passive optical splitter, only is not a reflection, only transfers the signal toward 2 different directions, this is how is done.

For sure an optical splitter is something that in this kind of architecture **costs very low** as much as the cost of a fiber. More, this object is not a device, this piece of fiber is a piece of fiber, this piece of fiber has also an **occupation** in terms of space that is **very very low**, you remember a fiber is very very thin, this object is as much thin, of the same size, it is very simple. And so this in these 3 cases gives a lot of interest in the use of **passive optical network**, all 3 of these are optical networks, in this case for the access, so access networks implemented by using optical fibers, the first one is a network where there are no network elements a part from central office and end system, in this case there is an active network element, this one, in this case there is a passive optical element and the advantage that till the passive element we need only **one single fiber**. This is very interesting and introduced our passive optical network.

There is also another kind of architecture that was proposed, the network implemented by Fastweb to cover his clients by using a **single fiber in a ring**, so there is a single fiber made by using a ring and devices interconnected by another ring to the end systems, you could consider the different buildings are along this building and then there is another ring for instance covering a larger urban area, an architecture for having another kind of optical network, the difference is that in this case these elements are **active ones**, they are not switch but they do like a switch, they send traffic in this direction if they are end users. Ring topologies are used a lot in the access network, so a lot of access architectures, a part the ADSL and this passive are made by using rings. The advantage of using rings is which one? We can have the very last past of the network, the buildings, or again the cabinets close to the home but not so close as the situation where the system ends. So you see this is again Fastweb, 2000, maybe something has changed, however the idea was to put the ring, **crosses this network element**, put for instance in a building, this is the ring, crosses and passes through this element and continues, and then this is interconnected, in this case fiber to the home, in this case fiber to the building, the fiber ends in the building and the rest is interconnected with another technology, could be also ethernet.

We already discussed the **most interesting solutions** are fiber to the home, to the curb, to the building and then there were other kinds of solutions, fiber to the office, to the desk and so on. Much more used are FTTH, FTTB or the old solution named fiber to the e where this e stands for **exchange**. Was the solution the initial one were the solution, the fiber was **ending to the central office**, to the exchange. So FTTE, FTTC, FTTB, FTTH. So, said this let's see how a passive optical network operates.

Passive optical network

As i said there is, there were and there will be maybe still a lot of interest in using that kind of architecture. The idea is we use 2 a fiber and 2 different wavelengths in the 2 directions. In this kind of architectures the so called **full duplex**, we need to be able to communicate in the direction from the end users to the central office and vice versa. At the central office we have a network element name **optical line terminal**, OLT, and in case of the end user we have the **optical network terminals**, and a terminal. The line terminal, so the end systems and the terminal. In the downstream we have a wavelength operating in the window of 1.5 nm and in the other direction we have a wavelength operating in 1.3 nm. So there is one in this direction and one in the other one. Operating in 2 different wavelengths, this is identical to what we have seen in the ADSL, a **spectrum bandwidth** for the downstream and another one for the upstream. Identical, a wavelength for the downstream and another for the upstream. What else? In case of upstream we can have 155 Mbps, this number was derived to emulate another kind of speed for the **ATM** for instance, and in the downstream we can go to 622 Mbps.

You see very different speeds with respect to the ADSL or similar solutions. This is the basic architecture and kind of **communication** in the passive optical networks. Also in this kind of solution they are working they were working in the solution of improving also the distance so if with ADSL we can arrive till 1.5 - 2 km, in case of fiber optics we can go **till 1000 of km**. Indeed in case of passive optical networks we are around this distance, 10 km, much more of 1 km we have in case of ADSL. Second, how it works? If you remember we have seen several times that the classical way to operate a multiplexing of communications in case of for instance cables ethernet and so on is to implement a **time division multiple access**, but in the optical domain we have also the possibility and this is used in some technology today to operate a **wavelength division multiple access**, while in TDMA the multiple access is operated by dividing the time in time slots and having different transmissions by different users in different time slots, in case of WDMA we can have communications by different users where the communication happens at different wavelengths.

This solution implies that on the same fiber **different colours are transmitted**, every colour has a wavelength, means contemporary different are transmitted on the same fiber. Why we need to have a **multiple access?** Communication of multiple users. Because the first case have higher cost but the different users are physically separated. In these 2 cases the different users are not physically separated, here and here in this path they need to share a common medium. So a **common fiber**, there is one fiber that needs to be shared, divided in some way by different users, so as I said we can do as classical as in the classical way by dividing in time or we have also this option, we can divide the single shared fiber by **dividing in wavelengths**. Second, important aspect, that has a role, is this one. As always the downstream is directed from the OLT toward the home. The upstream is directed from the home to the OLT.

The sharing of the downstream is not a problem, everything is managed by a **single entity,** the OLT. The sharing of the upstream could be a problem because is managed by the single optical unit. The OLT manages the whole bandwidth in the downstream. Initially the idea was to use TDMA, if it looks easier there are some problems related to this. Now I will show the problems, after a while they studied the usage of WDMA, allows to exploit some advantages that are intrinsic in the use of optical fibers, why obviously this was not possible in other kinds of media, this is possible in case of fibers because it's only related to the **possibility of transmit and receive** different colours on the same fiber. Consider that the transmission and reception on the same fiber today is possible, we will see we have also we can send also hundreds of colours on the same fibers, while at the very beginning was possible to use only one wavelength. Let's do the break and then we continue.

The first version was designed to work by operating in the time division multiple access. First, what does **TDMA means?** There are packets transmitted by the different OLT toward the different end systems and these packets are different depending on the time slots where the packets are put. Since if you remember this is a passive optical network, differently from an active one this element does not do any processing on the crossing packets. It's not a packet passing but **light**. This element only splits the signal, sends the signal toward the different directions. In this way all network elements receive all network units terminals, receive the same signal. This means that into the time slot or some signaling information you have to know **where your information is**. So, all you know for instance the time slot position or you know the presence of time slots and you read at an upper layer, not the physical but the medium access control layer for instance, an address here that says this is information for c and c pick ups the information. Remind you this.

The passive is done in this way, you can also have a **chain of passive**, one passive combined with other two and so one, this gives raise to distribution of signal from an input port toward other output ports. I did not mention this fact, the splitter is very convenient because has a low cost and occupies no space. The only little drawback on this device if are you following? So, the splitter has the advantage to be low cost and occupying no space. Differently from an active element, that typically does a **degeneration of signal**, this is a switch, this does two things. First regenerates the signal and during the regeneration can cross very far distances because after you regenerate you cancel some way the whole attenuation that you have. Second, this does **switching**, meaning that if there is an information direct to this the information is only sent to this, if there is an information direct to this the information is only sent to this. On the contrary, this element does not regenerate the signal and doesn't perform a switching, is also **splitting all end systems** receive the same signal.

This kind of topology even if physically is different is a **broadcast topology**. Is a broadcast, everyone receives the signal of the other, like a bus. So, while in the downstream is like a bus, there is a transmitter, all end systems receive the same signal and then by receiving this signal they are able to pick up their specific information, for instance by knowing which is the time slot where the information is or by reading the information and understanding yes this is my packet because the other says the destination is myself. As for the upstream, it is similar. But there is a key important difference. This is it. Signals, transmitted by the different optical network units are multiplexed by a passive element, so what is the multiplexing done by passive elements, the inverse operation of a splitter? What does this mean **multiplexing** in this case? Which is the difference of a real multiplexer as we always know? Real multiplexer what does? The multiplexer regulates the access to the single output in order to have a single transmission in the output per time, on the contrary also this has multiple input and single output, the scheme is the same, but the signal **overlaps**. How can I distinguish the different communications? This is at the same wavelength. You cannot do. Is a light, a **light signal**, transmitted contemporary to other lights. It is stronger but you cannot distinguish the difference, as far as I know, you expert of fiber optics.

So, the only way to do this is to **control the access** and try to avoid that multiple users, multiple ONUs transmit contemporary, you can't do else then. If we have contemporary transmission is a sort of collision. The difference with respect to collision is that while in ethernet you can be **aware of the collision** happening on the same media in this case more critical this optical network unit are not aware of the collision because the collision happens here physically, in another place, far from there. Ok. This is much more critical because they transmit if they were able but is not what they do, to listen to the channel, carrier sensitive, if it's free I transmit, they cannot do this, operation, if they listen to this channel, nobody else is transmitting, but the collision happens here. So the only way to do this is to **schedule the transmission** of the different devices in different time slots. If I can regulate the different transmission in a way that a transmits in a time period different from B, different from C, in this case collisions do not happen.

Which are the **different fix** to be done in this context? Can I do this, feasible? How complex It is? Do I loose something? It merges, in this direction is a sort of merging of signals, indeed we have to avoid they upload the data at the same time. The only way is that originally upload their data in a time period that is different from all the others. Is it feasible? Yes or not? The clock is needed, not only a matter of clock, this is a problem, in this figure is not a casual fact that you have a gap here. If you look at this, this gap is not present. This is not only a matter of the draw, in this case all these, this burst of packets are transmitted by a single entity, so it can schedule, on the contrary this gap here means that in order to be sure that a transmission of one network element does not overlap with the others you should give some margins to the transmission in order to be sure this transmissions **does not overlap**. If I say to 2 students please send me information I want to be sure a transmission of a student they arrive with the right timing, considering also the latency, to the right timing to avoid their overlapping. Second, more important, there is the need of **control** coming from a centralised entity. Also gives an indication to say when they have to transmit.

This is **scheduling**. Is a centralised scheduling performed by giving indication, we need a control channel to indicate it's your turn, your turn and so on. On the contrary it doesn't work. This is a little bit complicated. Because again we are a little bit paying the price of a simpler architecture so this is simpler with respect for instance cheaper, this is cheaper with respect to that one but the price is that we have to **regulate the upstream** otherwise it doesn't work. In that case like in ADSL nobody has to indicate to someone it's your turn or not, every user has its own media, its own media. In this case also there is no need for a regulation because we have a switch that can store, this can schedule and let wait some transmission in order to avoid the contemporary transmission on the single shared media.

Another problem is this one, still related to **distance**. If the devices have different distances as you know the attenuation is proportional to the distance. So this means that if this transmits at a given power, and this transmits at a given power and the powers are identical, after their path it happens that the one that is closer arrives before the other. This could be a problem because you want to receive all packets, all signals, all the whole light path at the same power. It's a problem for the receiver to have this **up and down of the power** that is received, it doesn't work. So to do this again you need a control mechanism that operates in order to set the power of the different devices as a function of the range where these devices are, it's a sort of power control, it's similar to what is used in some communications, if you are closer to me and you are far I say to him please use less voice because you are closer to me and I say to him please use a **higher power** because you are far. Again this is some control information that must be set by the central entity. In this case even if the physical structure is simpler the kind of operation that I used to have the system working are a bit more much more complex.

We need some **cooperation of the OLT** with the ONU in order to set all these mechanisms working. This operation is also named **ranging**, so I have to find out the different positions of the ONUs in order to set the powers and not only this, consider this picture, imagine that after a while the OLT is able to set the power to all the 3 end users, what happens if a new user is added to the system? In both cases I have to balance the power with the rest of ONUs, if I have a **new user**, even if the adding is not something that happens at a frequency every hour or every minute, it happens every day, every week, depending on the infrastructure. This must be taken into account. However this kind of solution even if there are some complex has been used in several conditions, several infrastructures and at the end the administrator is in some standards, passive optical networks are standard used in several infrastructures for the network access.

There are **2 versions** of this passive optical network, one of these is ethernet passive optical network, frame is ethernet like, in other words these time slots are frames standard with the ethernet format, named **ethernet passive optical network.** Then there are gigabit optical network and this gigabit passive optical networks have different framings, some of them are again are based on the ATM, where the frame has a given format, or ethernet based, these have different kind of physical layer. The **topologies** that can be configured in case of ethernet are different topologies and so on. Once this is ethernet based every frame has the format of our well known ethernet frame, we have an header, payload and control part. In the header as always we have the destination and source address, in the initial part we have a set of bytes used to indicate the starting of the frame, transmission is scheduled by the OLT. This is the downstream, as before all devices receive all the frames, maybe this can be also significant aspect related to the operator.

Indeed a problem with this kind of infrastructure is exactly this, the fact that all users receive the same information, for sure is different from this infrastructure and from that one. An infrastructure as it has been defined thanks to the unbundling laws, must be ready to support also **different operators.** This and for sure the previous one can be infrastructures used by different operators. In this case you have a physical infrastructure that doesn't allow to have different operators on top of the same infrastructure. **Why**? For sure consider the ADSL. We said in this infrastructure I can have Fastweb, TIM, Infostrada, Vodafone, these are all different operators operating in the same infrastructure. If this is physically different for operator 1 and for operator 2 and operator 3 and so on this works. In this case this can still done because this is an active element. In the last case is not feasible.

Once the technology was able to support an infrastructure like this the problem due to the fact this kind of infrastructure was not ready to support concurrent operators on top of this. The problem, I remember when a consultant of the minister came to this course so he came here and said we have tested this kind of infrastructure in our labs and presented to the minister and the operators are happy, there is a problem, the infrastructure does **not respect the unbundling law**, is not feasible on the same infrastructure to differentiate different operators. And it's quite important the fact that all systems, all network units receive the signal of all the others, not only a matter of privacy, is also a fact that this does not respect how the law has been defined for the unbundling. In this infrastructure cannot be solved, in a while we will solve, I'll let you know. In the upstream again there are concurrent transmission scheduled in different time slots, in ethernet passive optical network the frames are transmitted following the ethernet standards. To coordinate on the mechanism to register, to do the ranging, to assign and so on there is a protocol used by the OLT in order to organise all the transmissions.

So you see there are some **control messages** used to register the different ONUs, the OLT sends this register message, the ONU sends register request toward the OLT and the OLT sends an ack, obviously these are messages defined in the standard. We skip these details on how it works. These are the mbps that can be achieved by the different standard, part of these mechanism the only fact that I mention is that at a given time period when this technology was born the only problem was related to the possibility to have it used by different operators contemporary. What I have proposed, when this problem was they were able to propose a solution, the solution was this: consider that at that time was feasible to support on the same fiber also a **wavelength division multiplexing**, to use different wavelengths. Consider we have n wavelengths in the downstream and n wavelengths in the upstreams. Instead of assigning time slots we propose a wavelength division sharing of the single fiber that is here, and we proposed a splitter able to differentiate to split wavelengths on the single output fiber.

You see this is different, first all the same fibers we multiplex in the wavelength domain different signals, second we have the so called **optical routers**, so elements able to provide a single out from a single flow, a single wavelength in a direction and another wavelength in another direction. In this case this is not passive, but can be also. It can be made by using passive elements let's see the principle. The principle is instead of having a **single lambda in the input**, and the same wavelength on all the outputs I have an input with different colours and an output that is like a **prism** so an output that separates the colours. This can be performed by active elements or also passive ones. For instance there are materials that can be put here so it's not a device, a material able to split the colours. This is a passive one. If it's actively done you need to reconstruct the signal, then you transmit in the active domain, you can do in **both cases**, the important thing is that you are able to transmit contemporary different colours, if you are able to split.

Can be an optical router made by using filters. Like a simple prism, we can see when we are doing the optical routers that are similar to this. Some solutions very visionary at that time to have optical routers where you remember what is a **router**? Something interconnecting an input to an output. For a time period. An optical router can be done for instance by having a wavelength entering here and exiting in a direction. So a visionary proposal was this, to have this cross point made by having a **bubble** of a material, rising and going down, of something able to reflect for that period the light. This was a router, proposed in the optical domain, made by using these going up and down bubbles, able to reflect the signal. For the time period you had to show the bubble in different places, a static routing based on time. Consider that today a big part of network is made **all optical**, you not pass in the electric domain. We will see. As for the passive optical network this is a solution, use different wavelengths and they proposed to the minister to provide a **different wavelength to the different operators.** Let's consider this, operator Telecom they have lambda 1 and lambda 2, operator Vodafone they have lambda 3, operator Fastweb they have lambda 4 and in this way the operators are differentiated. Even if this still not fully compliant with what the law says, the problem is still the following, even if this separates the different users there is still a problem for who is **responsible of for instance this fiber.** This is a single fiber, I have different flows but the fiber is only one, can I rent a lambda? A wavelength? Maybe this is still an open problem.

If the infrastructure is all owned by a single operator it's not a problem. It solves the fact we avoid the users receiving the information of the other. This is however critical for privacy and safety issues. This is how this kind of network can be proposed, we can skip and another aspect is sometimes since this kind of solutions are used on very big distances and sometimes they have also a big number of users it is fundamental also to have the possibility to keep the infrastructure working with some **redundancy** and this protection mechanism, redundant configuration can be achieved in different ways. You could have 2 optical line terminals redundant at the OLT or 2 redundant optical line terminals and 2 configurations that can be used. **1+1** means the 2 infrastructures are used and when one of the two doesn't work anymore the other comes in play, works as for the other. Instead of the other that is not working anymore. This is quite old but however was an indication of the growing of this technology, I'm not aware, for sure this is used was mentioned in the solution I was checking is that till my far cabinet there is a Gpoint in Rome.

This is only to say that today for a big part of the infrastructure till the cabinet from a central office you have **g passive optical network** and then you have the ADSL. If you look at the current infrastructure you can notice that in several cases this part that here is only indicated as fiber is a g point, a passive optical network, with the advantage of reaching **very far distances** from the central office to the cabinet and then you have to do only a very last mile less than a mile, last meters, last part of the network with a copper wire, thanks to this you can have higher data rates. This was only presentation of this lab, built in foundation some years ago to test these technologies, they were consultant of the minister of the communication to build up these new infrastructures and to show you a real testing platform, done by emulating an internet passive optical network, this is a configuration, these are splitters, an ADSL+ access, another ADSL+ VDSL with another DSLAM and then there was a ring, an optical ring interconnecting Rome and Pomezia with different distances of the network elements of between 50 and 300 km.

On top of this they were experimenting, this platform is still available, they were experimenting the **behaviour of the network**, performance that can be achieved, and tried to propose VLAN tagging, we have discussed the possibility to differentiate the quality of service by using virtual lans and adding an header to ethernet frames with the objective of having different service differentiations due to the fact that operators in case they have an infrastructure they want to differentiate their clients and provide different services to all of them. These were the results of some experiments they were doing at that time.

Wireless communications

Today we will talk about wireless communications. We present different wireless communication system and then we go into detail with LTE. For instance there are **two main ways** to support wireless communication. First is by providing a **broadcast distribution network**, using a big antenna like in tv case or in case of the use of satellite. We can have infrastructure with a satellite or antenna on for instance an high hill, like it happens in Rome from monte Mario, so on top of the high hill a unique transmitting entity or satellite, who broadcasts the signal toward the different end users. Or on the contrary in much more complex infrastructure we have different transmitting antennas named **base stations** and these are interconnected one with the other and then they have typically a part of the network named the call, interconnecting different base stations and then the base stations transmit their signal toward the end users.

These red links indicate the wireless part while the other are wired. The standards proposed in this context are many and can be framed in **different families**, we have many standards in the high degree context, starting from 802.11, we will see this a little big, and there are other standards in cellular context. Indeed 802.11 was born in the **framework of the ethernet**, while the other standards were born in the context of the cellular partnership and we have the 3GPP that is the context where GSM evolved toward EDGE and then LTE and we will also see 5G, something related to it. There are very complex parts defined for the standards and sometimes there is the tentative to merge these parts so for instance how it happened in case of WiMAX, this tentative was to **combine some standards** toward the direction of cellular network but maybe it was not so good in order to replace what on the contrary was going on the 3GPP framework.

Standards are a lot and indeed they initially were designed in order to support **different requirements**, for instance there are wireless standards able to work very well in the indoor environment, and standards designed to work in the outdoor environment. As well standards able to work with a reduced mobility, like bluetooth, and standards designed to support a high mobility, mobility is an issue, important issue, sometimes the mobility is at for instance walk speed or fixed or we need a very high mobility. The **most challenging** today is a mobility able to interconnect a train going at 300 km/h.

Another aspect important is the **bitrate**, the data rate. Again for some wireless technologies the bitrate is very reduced, initially in case of second generation cellular system they were from the beginning designed to work at a quite good mobility but the data rate was low. Going in this direction you see the data rate is growing, from **third generation** quite good data rates and again very high mobility, while in the WiMAX and other standards the data rate is quite from the beginning was designed in order to be quite high, 10 Mbps is very easy to be achieved from the beginning in 802.11, also improve to support much higher data rate so in the real broadband way in case of WiMAX you see, **LTE** is today the unique standard able to support high data rates on a side and high mobility on the other side. This is the framework where we move, the constraints are different and the technologies were presented in a different way. Let's recall some concepts we need to know in case of wireless communication. They are **different** with respect to the wired.

We already know a wireless link is much more sensible to the interference and the loss of signal and so on. Also an aspect that **was not present** in a wired network that is typical of wireless communication is to manage the mobility. All the effort in the wired infrastructure were mainly in the direction of supporting high data rates, everything was designed with this aim. In case of wireless communication it's intrinsic in the wireless you can design a technology able to support the mobility, even if mobility has other cons, indeed in mobility as for the wireless we will see there are different ways to provide an infrastructure and the wireless communication. The very initial proposal in the mobility were done in the framework of the **Mobile IP.** The mobile IP is similar to what happens in the cellular mobility, the first that was designed, before the mobile IP. Let's recap the elements of a network infrastructure.

First of all we have the **wireless host**. We have the wireless host, again the terminology of the IP we have the hosts, the end systems, wireless. Obviously this is also true, is not always true that a wireless host needs a mobility. Then we have the **base stations**, elements of a wireless infrastructure and they have sometimes different names, they change the name every time the technology changes, we call them base stations, that is the most common term, but you will see they added this strange name, node B, this name is also the one used in LTE, in wifi the name is typically **access point**, the meaning is similar to what is a base station even if they are a bit different. These are other key elements, indeed these are the nodes, the only nodes having a **wireless interface** toward the wireless hosts. So the base station, access point, how you call it, these are the only elements having a wireless interface toward the wireless host. The rest is a part of the network named backhole, for instance in case of wifi this is named distribution system, typically a wired infrastructure typically. In case of cellular network is called backend, a **wired network** and composed by different network elements.

These elements are interconnected to the real wired network that is an IP or telephone network, it depends on the specific context. We mention also this because another difference we have with respect to the wired, in case of wireless we can have a network **without an infrastructure**, a network where no infrastructure nodes exist, no base stations, on the contrary the network is setup and maintained only by the wireless hosts, the typical network, mentioned by the seminar, with this no infrastructure is named **ad hoc network**. It's a wireless network where the network nodes are only the end systems, these terminals they can behave like access points, we several times use today our smartphones as a router, so this becomes an access point for our interconnection like it can happen in a ad hoc network. The ad hoc network is also **much more complicated** a network where you can have multiple hopes, a routing like it happens in the IP routing, on wireless devices and mobile. For instance keeping a routing table, it takes an effort, it needs a protocol, you have to update the routing table and so on, imagine what happens in case of an ad hoc network, nodes can behave **as a router** and may change the position, you have to continuously update the routing tables or sometimes this routing is done **without using routing tables**.

Hundreds of papers there are dealing with this. Single hop is simple, you have a star topology, an access point operating as the element having the interconnection to the rest of the network or without an infrastructure, no base station but you have a **very little star network** for instance for interconnecting the network elements, in case of multiple hops as I said before you can use some end system as element as **relay nodes**. They have a role in the network only for relaying the data transmitted by other wireless networks, for sure a wireless link like a wired one is constrained by the length. In case of an ad hoc network you have a node, has a **wireless range**, the range where other devices can interact with this node and sometimes you have to use another node as a relay, like a router, wireless and this device must be in the radio range of the transmitting one and on the receiving one, in this case you have a 2 hopes path, you can extend this behaviour with multiples hopes till the destination.

Characteristics of wireless communications

Which are the characteristics that influence the wireless communications? First this is also true in the wired, as a function of the distance the **signal strength decreases,** like it happens in the wired. In case of wireless typically we name this phenomenon as **path loss**. The path loss measures and characterises and models how the signal decreases as a function of the distance. The path loss may be different depending on the specific wireless channel used for the specific technology. If you have questions please do them. This is the first point, decreasing of the signal as a function of the distance.

Another aspect quite new in this case with respect to the wired is that we have **much more interference** from other transmissions that happens in the same area, we have seen in case of wired the only kind of interference was the cross interference that we can have of different cables, on the contrary here the interference are **multiple**, we can have them by systems transmitting in the same frequency band and system of the same technology or systems of other technologies, think at 2.4 GHz bandwidth. Wifi operates in the same bandwidth where other devices can operate, without using wifi. If they share the bandwidth they share the interference. Microwave and other devices, every device able to transmit something at that frequency is an interferer for the wifi.

Another aspect, typical of a wireless communication is the **multi-path propagation**. In case of a wireless connection since the radio signal is not on a cable, it may be **reflected by the environment**, so it happens that if I transmit here my voice reflects on the walls of this room as on the table and on people. In case of wireless communication at a given frequency band this propagation depends on the adopted frequency and may have his interference that can happen on the signal itself. This transmission to arrive at the end of the room has a direct path and also several reflections given by the multi-path. At the end the communication arrives with **multiple versions** of the same signal. These multiple versions are different for 2 reasons, first they may be delayed, why? They are different in time, why? One path is direct and the other is reflection. If the path has a given length in the direct link maybe the reflected part has a little bit **higher length**. There is the direct path from the transmitter to the receiver and then if I'm in a room or an environment reflecting the signal in that environment, I can have multiple paths all of the same signal, but maybe this one is longer than this, so they arrive at the receiver **delayed**.

We will see the wireless technologies has **evolved**, the fact they arrive in multiple version, sometimes has been used as an **advantage**. Why? We can use it as a redundancy if we know the channel that is crossed and we can estimate the delay of the different versions. If we are able to characterise the multiple behaviour of the channel then all the multiple versions are used as **redundancy**, this improves the performance of the system instead of reducing it. The multi-path depends on the specific carrier frequency where the signal is transmitted, because reflecting objects can be different depending on the frequency, you know. For instance a frequency where a wall is a reflecting object maybe the same wall is not reflecting for another frequency. A tree with its leaves in some frequencies maybe is reflecting objects. Also people may be. Depends on the **specific frequency**, you know some frequencies some elements like the water, the rain, they may be reflecting objects.

These 3 are very specific to the wireless and determine the characteristics of the wireless channel. A wireless link is much more difficult to be managed, to be model, with respect to the wired. As we know we can send a given bitrate on top of a channel as a function of the well known **signal to noise ratio**. As always if we have a high signal to noise ration we can transmit much more symbols per bits, we can use QAM we can use a very high QAM if we are able to support to achieve a high signal to noise ratio, on the contrary if we have a low signal to noise ratio we go back to a low QAM and a low QAM means a low bitrate. The bitrate is also function of the SNR. Indeed is the right term, not anymore SNR but is **signal to noise plus interference ratio**. SINR. Indeed in a wireless communication we use this term that is the ratio of the signal over the noise plus interference, this I is the interference, something that decreases the behaviour of the system.

Another aspect is the **hidden terminal problem**. Well known in the context of the wifi, was presented there for the first time and is represented here. You can have 2 nodes as I said, what is important is that a receiving node is in the radio range of the transmitting one, if I transmit something my receiver is in my radio range. But it happens that for instance in this case B is the receiver, in the radio range of A and B is in the radio range of C, B can be a suitable receiver for A and for C. But A and C they are not in the same radio range. A does not see C and C doesn't see A. Both of them they can transmit to B. This is the hidden terminal problem, the signal transmitted by A **doesn't arrive to C** and vice versa. Both signals arrive quite high at B. This problem rises when A and C want to transmit something to B, they can assume that nobody else is transmitting to B, because A doesn't see C and vice versa. This is a **collision**. B receives 2 quite high signals, they are strong signals, seen by B, this is the hidden terminal problem. This problem has been solved in some cases or reduced by adopting protocols not needed in case of wired communication, that instead **are needed** in case of wireless due to this problem.

Several times in the wireless communication we have to design new protocols only because we are in a wireless context.

As for the multiple access in wireless communication there are several families of multiple access that are used. One of these, not the only, is the **code division multiple access**. Is a multiple access technique where users transmit their bits by using different codes. Every user has a codeword, used to encode the original data and transmit them. Obviously the decoding, who is receiving has to know the coding word and to decode the data by applying the coding word that is only of that receiver. This is an example that maybe you know, taken from the Kurose Ross book, in our website there is also the whole chapter of wireless communication, I put on the website. Here this is a **representation**, something similar to the spreading factor we are applying in a project. The idea is this one, I have to transmit my bit that is this one, so this bit can be a 1 or a 0, and I multiply the original bit for my codeword, I do this multiplication and transmit this signal. When this signal arrives to the receiver the receiver has the same identical code and by multiplying the received signal for the code and dividing by M to **reduce the power** of the different elements, I have a code of M elements, If I do this I can recover the original signal.

The power of this kind of transmission is that I can transmit at the **same frequency** at the same time signals belonging to different users. The different codewords are selected in a way they are **orthogonal** one with the other. If I have an orthogonal code word when I do this multiplication I keep only my original data and all the rest is put to 0. This is how the code division multiple access works. Consider that CDMA is very important because it was the **main multiple access** used in the CDMA2000, when we had the GSM in Europe they had the CDMA2000. This is not the only one, in case of GSM, LTE, we use another kind of multiple access technique based on **time division and frequency division**, also this one. However this is important, similarly the spreading factor is similar you know, but the reason why this is used in case of that kind of communication is that we spread the power of our signal in a large bandwidth at a lot power, this is indicated for the low power communication, they will present much better. Break now? Yes.

There are other standards, A and B, B was the one designed to operate at 2.4 and 5 GHz, that is in a spectrum band and the maximum nominal data rate was 11 Mbps, the kind of modulation scheme was the **direct sequence spectrum**. Again since we are operating in a license band the idea is to provide to spread the spectrum on the whole bandwidth in order to reduce the interference effect on the other systems, all hosts have a code done by using sub elements, these are used to modulate the signal. Then the **802.11a** instead operates in the 5.6 GHz bandwidth and goes at a bitrate up to 54 Mbps. The same was achieved later with 802.11g with an higher data rate and then there are still evolutions, for instance there is the 802.11n that is going to a very high data rate, 200 Mbps, and then there are also other standards, 802.11h and p, p is the one designed for the **vehicular network**, with some specific characteristics of this solution, in all cases the carrier sensitive multiple access collision avoidance approach is used and the architecture of the network is different if it operates with infrastructure mode or ad hoc mode. In the infrastructure mode the infrastructure is built **like cellular network.** Every system is managed by the access point, that are interconnected by for instance an hub, switch or router or by a little bit much more complicated infrastructure, distribution network.

In this basic set we have the wireless host and the access point, sometimes in a basic service set we can also use multiple access points wirelessly interconnected, so we have a multi hope into the same basic point. What happens in case of a communication of 802.11b and in the other cases? You have to **interconnect to an access point**, you have to perform an association procedure, that works in this way: the access point, broadcast, with some specific control frames, their ID. Every access point, broadcast with some repetition of frequency they send their ID and a device collecting all the wireless access points available in the radio range select the access point to associate with. You have to collect **all the frames** and once you see more than one frame you can select among the different frames, typically the selection is done manually but also it happens automatically by selecting the access point having the **strongest signal**. Since several access points operate in a secure manner, let's say, you have to authenticate with the access point, they are closed, not always open to the different users, when you are authenticated behind the access point the DHCP is the server assigning dynamically the IP addresses.

The association procedure works in this way, a device collects the frames sent by the different access points, you see in this picture by indicated by 1 the different frames arrive, then the terminal selects the access point that wants to be connected, sends an association request to the access point and the specific access point sends back an association response for the selected access point. This is named **passive scanning**. Instead in case of **active scanning** the process is started by the wireless host, the host sends the requests toward the possible access points that are in the area and they send back the response and again we have an association request and association response. This request is directed toward general access points that can be present in the area but also specific toward the access point **already known** by the device. This is the reason why you know we several times say well all the wifi networks that we register in our smartphones are sent out in case of active scanning, this means that every time you have visited a wifi this wifi is stored in the list of visited wifi.

Is sent out in the active scanning case, this means I can capture by using tcpdump for instance the list of all wifi you visited. I can be aware of the list of wifi that you have visited. There was a group of professors of the dipartimento informatica they did a very good work, they went in St Peter square and collected all the wifi. They were aware of the kind, they did a profiling of the users present in the pope assembly in that case to have information on the hotels, restaurants, all places where typical users participating in the pope assembly because you can do a data analysis. A profiling of the users. They are aware of the kind of users participating at that event. The list of wifi networks you have in your smartphone is a signature of your person.

The problem of hidden terminal was solved by providing a way based on this principle: when you transmit a data you have to perform this access, listen the carrier frequency for a time period, indicated as **DIS**, if you listen to the channel for a period longer for DIS period you can transmit your data. After this transmission if it's not possible to understand if a collision happened, you have to wait for another time period, shorter inter-frame spacing, after this you can receive an ack. If the ack arrives and you are aware the communication was successful you have to retransmit the information because maybe something went wrong. To **avoid the collision** was based on this concept, since it is not possible to have knowledge of the presence of another user transmitting on the radio range of your receiver, the CSMA/CD operates in this way. 2 terminals that would like to transmit toward a given receiver send **short packets** named request to send. Very short packets, so what they can do is that they can collide, a collision can happen but it happens only if they are transmitted at the same identical time instant and arrive at the same moment at the receiver. This condition happens very rarely that 2 identical signals request to send are sent contemporary. May happen but it's more rare.

If the transmission of a request to send succeeds the receiving entity, B for instance, sends back a clear to send. The limit of a clear to send is this one, on a side indicated to the transmitting entity that this entity can transmit. It is a sort of **ack to the transmitting entity**. On the other side the meaning of this clear to send is that this message makes aware another possible transmitter, in this case B, makes aware another transmitter that someone is going to transmit toward this receiver. In this way this hinibits the transmitter here to transmit the message. After the clear to send the channel is free for the transmission **only for the user** that had the request to send exchange, this user sending the request to send and receiving the clear to send is authorised to transmit its own data. An user receiving a clear to send is hinibited for transmission for a period that is written on the clear to send. On the clear to send there is the indication of the length of the communication that the entity would like to perform.

A wants to **transmit** toward the access point. In the request to send sends I would like to speak with you for 2 minutes. This is written in the request to send. The receiving part, the access point, sends out a clear to send and says ok you can transmit for 2 minutes, these 2 minutes are stored here and for 2 minutes this device will not try to access the channel. This has a double advantage, first it is a clear to send for who intends to transmit. Also the advantage of avoiding possible transmission of hidden node, but this gives the possibility to the hidden node to go to **sleep and to save energy**. So if there is someone that would use the channel for 2 minutes, this device can defer its tentative to transmit and also go to sleep and save energy. This is also an energy saver for the devices. As for the communication maybe you remember this, the 802.11 frame is very similar to the ethernet frame that we already have seen but a difference is that instead of having only 2 addresses this has 4 addresses, so 2 are the sending and receiving MAC addresses, the other 2 are used and in particular only one is used in the infrastructure mode, the second indicates the address of the AP. I don't know if you have used Wireshark, similar to tcpdump, I discovered that Wireshark has pre-stored a lot of traces, captures of wireless traces.

You can have different kind of mobility, initially 802.11 was not designed with the capability to support mobility as for instance is intended in case of cellular network, because of this association, made exchanging information between the **host and access point**. But also by in case authenticating. If there is this exchange of messages and the authentication it's not easy to move to an access point to another and avoid this message exchange that takes time. While in cellular network this is designed in a pro active way, when you connect to a base station there is a protocol assuring that if you move another base station can take care of you, in wifi initially this protocol was not designed, while in recent version of 802.11 there is possibility to **support mobility**, the access point should exchange some information so you have possibility to move toward access points without having to authenticate again. Initially there was no interest in having this mobility function. Moreover, when you move generally you would like to be interconnected with the network having the best signal to noise ratio, the mobility should measure the SNR and for instance if you are very close to an access point, if you move and your SNR decreases maybe you have to skip to another access point to **keep your quality** at a high level.

Cellular network infrastructure

When the quality of channel is decreasing the terminal tries to interconnect to another access point if this is available. There are several capabilities added to the initial standard in order to support mobility and power management that are important in this context. For instead **cellular networks** the architecture is a little bit more complicated. First aspect, as we discussed, the cellular infrastructure is made by combining different cells. Cells are areas served by a base station and different base stations are interconnected to other network nodes that are different tasks. These nodes are named **mobile switch centre,** even if the name changes depending from the technology.

The concept is this one, the base stations have the role of providing the wireless connectivity to the end users in the area, the mobile switch centre have the role to **allocate the resources** to the different users, and manage the mobility. Two main tasks. Then this mobile are interconnected to the rest of the network. So, connect cells to the rest of the network, manage the setup and permit mobility. In case of cellular networks the multiple access can be code division multiple access or a combination of frequency division and time division multiple access. In case of **LTE,** the recent technology, we have multiple frequencies, different channels and in every frequency we divide the channel in different time slots, so you have a sort of matrix where each matrix element is a combination of frequency time. So you have a resource block of an element of this matrix that is at a frequency in a given time. Everything started by the basic very first architecture of a cellular network that was the one designed for the 2G. The names were these one, base station systems, a combination of cells, every cell managed by a base station, then we have the base controller, a network controlling a set of base stations and then the mobile switching centre.

All this part in the left was the infrastructure added to the **very initial network** to support the mobility cellular network. When the 2G mobile network started the infrastructure was changed by adding all this stuff, so you see from here toward the left part all this was a new infrastructure, quite complex and also with high cost, added to support mobility in the 2G and **wireless communication.** This was designed only to support voice. After a while there was the interest to support data, this is how the 3G was provided. Provided on top of the 2G by adding new network nodes and these were named serving **GPRS supporting node**, so this was a modification of the original mobile wireless network designed only to support voice in order to support contemporary a packet switched network in the wireless domain. This was added to the existing infrastructure so a new node was added in the serving part and a new node was added for the packets in the gateway part. Obviously that gateway and that serving nodes were designed only to manage voice, when you have to manage packets there are different things that you have to manage, for instance when you have a packet switched network there are **many differences,** in case of packets you don't have the concept of connection, true in case of voice. In packets you have the addressing, not present in case of voice, and so on. This was meant to include new functionalities and support them. More, also the mobility, something not present in the initial packet network was added to this kind of network. Only to say, without going into details.

I want that you know the main aspects because it's obvious you must know this related to wireless network infrastructure. As for the mobility, let's recall main principles if you want we can do another break and then continue with this. How **mobility was designed** in the IP word, they invented some concepts that indeed are very similar to the concepts that are invented in case of cellular networks, they gave a different name to the devices so there are **3 concepts**, one is the concept of the **home network**, there is a network where I am registered and it's my home. In this home network in case of IP I have a so called **home permanent address**, an official address in my network, this address is named permanent address. In the home network there is an agent, then there is another network named **visited network** and in the visited network there is another agent named foreign agent and in this visited network I can have another IP address that is named care of address, so there are 2 agents home and visited, 2 networks, home and visited, 2 IP addresses, permanent and care of address. These 2, agent are different.

This is used in different mobility approaches and can give raise to different ways to **manage the mobility**. I would like to be able to find my mobile terminal, so for instance when I move from my home network to a visited network what is needed as first step is the registration of the mobile in the visited network, I belong to an home network and I move to the visited one, first thing is to register to the visited network. After this registration to the visited network a visited agent has to communicate to the home agent saying your user is now in my network. After this registration the foreign agent take cares of the visiting host and communicates to the home agent that the visiting host is in that network. Then how the routing works? In this way: if a user wants to contact a mobile device what the user should know is **only the address** of the device as is home address. So users starts the transmission trying to send the information to the permanent address. So if I want to communicate with my friend and I know my friend with his permanent address I try to communicate with the permanent address first, if not available then the connection is redirected to the visited one.

So the destination is the permanent, in this interaction instead the permanent is redirected to the visited one, this is named **indirect routing.** This is a source, this is destination. The final interaction goes in this way, there is this sort of **triangle** that can happen and the triangle on the one side efficient, on the other side inefficient. This is inefficient when a corresponding mobile device are very close to each other but they have to do every time this triangle for routing. I do an example very simple, I am in New York, I would like to send a message to my husband in New York and every time to my husband I do all these steps toward Rome, this is why is **inefficient.**

On the other hand, this is the disadvantage of this scheme, what is the **advantage**? If there is an advantage. I repeat, every time I send a message to another person in this area I have to do this, but which is the advantage of this schema? If I send always packets toward the same address, I don't have to know which is the temporary address of this device, I always communicate with the same address, I send the IP packet to the same address, this is **transparent** from the sender point of view where the entity toward who I am communicating is, I don't care where this entity is, **mobility is transparent to me**, on the contrary a different solution is the **direct routing**, I try to communicate in the home network, the home network says this device is not anymore in my network, consider for the time being this device is in New York, so use the address of New York instead of using the address one of Rome, use this address and communicate directly to this user that is close to you.

In this case first on the one side I **avoid the triangle routing**, I do direct routing, this is the advantage in this case, the disadvantage is that I'm aware that the entity that I'm contacting has an address close to me. In this case is not anymore in terms of privacy I'm aware where the user is, on the contrary in this case I cannot be aware of where the user is, from the sender point of view the user is always in the home, but maybe the user is somewhere else, but it's fully transparent, as for the transmitting entity. What happens when a **user moves?** If a user changes the foreign network, means a user moves from a foreign visited network to another foreign visited network everything continues as before, there is an agent in the foreign, a new agent in the new foreign, these 2 communicate and then again there is a communication toward the initial home network that is redirected to the initial foreign agent then is redirect to the second foreign agent and so on. It may happen that we can have **multiple hops** on the different visited foreign agents. The first visited one is names **anchor foreign agent,** the first for this communication and then a chain of subsequent foreign agents that assigns the care of address for the subsequent communications, this is how the mobility is managed in case of indirect routing.

Can be achieved also nesting different routing every time a user moves, this is designed for the **mobile IP,** or, this is very similar to what happens as for the cellular mobility, in this case we have an home address and an a visited address, but everyone communicates always toward the home address, you know for instance it doesn't happen that when you communicate with your phone you change the number you do when you have to communicate, you always communicate with that number, then that user **you don't know where it is**, you always communicate with the home and then the home in case redirects your communication toward another places and so on.

Obviously we mentioned that in case of direct routing there is disadvantage but is much more efficient, apart the case of **very high mobility**, what happens in that case? If this user changes again visited network, what happens in this case? If this one changes continuously the visited network. It is complicated, this means this one indicates again there is this mobility and then some packets sent here don't find anymore the user here so they have to redirect to the home and the home has to redirect to the new one, on the contrary in this case you see you need an **exchange of information** also between the different foreign agents because if they don't communicate the system doesn't work and maybe some packets are lost, due to the mobility, and this is something that has not to happen in the mobility case. This is to say that mobility can be managed also in case of IP networks even if these networks were not designed from the beginning to support the mobility so all aspects added later were no efficiently designed, other technologies are more efficient than this.

Wireless systems and LTE

Today we continue the presentation of the wireless systems. I want to go to the conclusion and then we start the LTE. Last week we were discussing how in mobile IP that is similar to what happens in other mobility structures the user is managed, specifically there is we said a **foreign agent** and an **home agent**, the home agent is the place where the original user is registered and when this user moves to a visited network this visited network has first of all to be registered so to be known by the visited network and after this registration the foreign agent contacts the home agent saying your user is in my network, so visited network says to the home I have a user with a ID of your home network.

This is the very initial part, after this now during this communication the foreign agent has knowledge of the visited network and this can be used as we were presenting last week by saying that for instance the client that wants to communicate with that user directs the communication **toward the home agent**, to redirect the routing toward the visited network. This redirection is done, I skip this, by a way that is named **tunnelling**, so tunnelling means the packets directed to a specific destination assumed to be in the home is encapsulated in another packet with another address that is the address of the destination of the packet in the visited network. Last week we said on the foreign agent to give a **temporary address**, IP address in this case, that is named **care of address**. Care of address is the address of that user in the visited network, used to prepare this tunnel, an IP packet put inside another IP packet and with a different address for routing the packet, otherwise if we don't use this second address the packet cannot be routed.

Originally the correspondent doesn't know the current IP address of the user and this user **never will know** the address of the corresponding. This user only knows that the correspondent has his own primary address, after this routing thanks to the interaction of the foreign agent and the home agent the foreign agent can receive the packets sent by the home agent, because they are **redirected by the home** using this tunnelling. Thanks to the tunnelling the packet arrives to the right visited router and the visited router by decapsulating this packet from the tunnel is able to recover the initial address and then to route to the final destination the packet. Thanks to this mechanism that is named **indirect routing** the correspondent doesn't know where the user is. Everything is transparent as for the corresponding user. This can be repeated several times, we can nest several tunnels, if this end user moves to another foreign visited network again you can do the same process.

This is defined in the **mobile IP protocol** and the interaction between the foreign and home agent happens thanks to some specific control packets used for announcing the movement of the specific terminal. You can see here there is this ICMP management control message so an **agent advertisement**, thanks to this the foreign agent advertises the corresponding address, then is stored as the care of address a client that represents the user and then this is sent to the home agent to say your home user is now in my network and my network has as a care of address this number. Is the care of address of the user as user in my visited network. Thanks to this there is a **registration reply** and this confirms. Obviously there is this registration reply to announce to the end user that his **process for registering** to be then able to locate this user in the foreign network is active. Indeed you see this interaction in every interaction there is a timestamp in order to indicate that this interaction stands for a given period. This is a detail of what we described last week as for the indirect routing in mobile IP.

There is also the possibility to manage the **direct routing**, it's not explicitly presented here into the detail, obviously in case of direct routing what we can imagine as for the management of the IP addresses, in case of direct routing, this case here, presented last week, which are in your opinion the differences with the indirect routing that we just discussed? We said that in this case this user can communicate directly with the correspondent. As for the routing. Routing happens by **addressing a packet** to the destination. In the indirect the routing is done by addressing the packet to this destination. This is fully transparent, as for the correspondent the destination is **always the same**, this is what happens also in the cellular network, when we call a cellular phone we call a number, we don't know where this person having this number is. So this is identical, in terms of routing in case of IP this correspondent sends a packet to the number the IP of the user as user of the home network and thanks to the interaction to this is fully transparent, this is sent toward the real position of the user.

In the other case which are the differences? In the direct routing. Can you see the **difference?** If there are. Which is the difference in the second case? This time no tunnelling exists because the end user is made aware of the current IP address of the visited network. This is the difference. So the routing then happens as always, this one always sends a packet with a destination that in this case is the old one, and then **tunnelled toward** this direction, in the direct routing the destination will be directly this, not the gateway, a destination. Destination this time will be in case of direct which number here? This is routed because the IP can **root where you want** so this is routed toward the right destination. Indeed we discussed last week that in this second case this is much more efficient, because it may happen that the visited network is very close to the sending network, you should remember that the IP addressing is related to specific geographical areas, so IP addressing is very close to the geographical area. This number is related to an area or another one.

So if you **replace the IP address** of the home network with the IP address of the visited network it may happen than these 2 are closer so for instance the routing path is shorter passing directly here with respect to having the triangle routing. But which is the **disadvantage**? This is the advantage, for sure. The disadvantage? I see two main aspects here. The correspondent has to be continuously **aware of the change of position**. Second, this is one. Another disadvantage maybe. Since we said the IP address is related to a location in this way you make aware this device of the position of the care of address. So in other words if you move in a place you have the address of that place and this address reveals where you are, it's a sort of **violation of privacy** let's say, while I just I'm aware my initial address is for instance in Italy but I don't want to let know that I'm here in another country for example.

Cellular network addressing

In case of **cellular network** everything is done in this similar to this: you always communicate by having a sort of this, even if in case of routing of IP is much more inefficient. This as for the mobile IP. As I said similarly it happens in the cellular networks, as indicated the last week, everything starting in this system and then has been moved. In cellular network there exists a **similar architecture** and concepts. We already presented, we are the **cells**, cells that cover a given area, in each cell we have a base station and different base stations are connected to the mobile switching centre, that is the element managing multiple base stations covering a given area. Typically an area covered by a mobile switching centre is named **location area**. Is a larger area where multiple cells are and also where multiple base stations are interconnected.

Mobile switching centres communicate in order to manage the mobility of users. What exists here is that there exist 2 kind of mobile switching centre. The home one where the **home location register**, HLR, exist. Its database where my official address, telephone number, is registered. If I buy a telephone number in Italy my home location register is in Italy, in Rome. Then there is a mobile switching centre where I can move and related to this there is a **visited location register** so is it identical to the previous case, an home and the visited. Home where I have been registered, a visited where temporary I exist. If I move in Milan I belong for a time period in the visited location register of Milan, in case for instance of cellular networks this happens continuously, update the visited location register of the place where we are.

Visited location register that is **managing the different cells** has the number of my phone because I'm there. This is a procedure named **location procedure**, used to localise where the user is. When someone wants to call me it happens as before, so the correspondent calls a number. This number is officially registered and then the network routes the signaling toward that number to the **home location register**. Then the home location register has knowledge of where the number should be sent, where temporary the user is. After this the network routes the call toward the visited location register and you can receive this. There is this transparent passage toward the home location register. And then moves your call toward the visited area. Then after this the call is directed to the specific base station that manages your communication. As before, the indirect routing that has been designed for the mobile IP is a **replica of the routing mechanism** already present in case of cellular networks to manage mobility.

I would like to move around areas and I would like to be found by a calling device in the area where I am. Then the mobile switching centres are also able to manage the mobility during the call, this is something related to the setup of the call, where I am. I have to be able to find where the end user is. This is the first step, the second step is in case I have a call I would like we re moving in the area managed by this mobile switching centre. I need to **change during my connection** I need to change my **serving base station**, so it may happen that I manage by a given base station I move in another cell, obviously this depends on the size of the cell, this process is named hand off or hand over. Hand over, is the procedure to manage a call or a data exchange, also during this, if I move from a cell where a base station is, so for instance I originally have been managed by this base station, I move and these 2 base stations has to **manage the hand off** of my device and the goal of the hand off is to route the call with the new base station **without interruptions.**

The interruptions are less evident in case of a data connection, much more evident in case of a call. The worst thing that can happen during a call is that this hand over if it fails what happens? The call is lost, **interrupted**. Sometimes it can be interrupted because the handover doesn't work well. This **handover** be done in the cellular network in different ways, initially a way to perform this handover was designed in order to the first important thing to be done is to keep the signal transmitted toward the end system high. If I'm able to understand that by moving from a given base station to another the **signal will become higher** this will be the base station managing my call. There are also other things that can be done, I can also manage to assign a user to another base station because maybe this has more capacity or less users than the first one.

In the very initial **GSM hand off procedure** in the framework of the same mobile switching centre was the following, you initially are interconnected to a base station and you can provide to the old base station a list of possible new base stations able to take care of your communication. The mobile switching centre has this list of possible base stations that can manage your communication and the mobile switching centre set ups a path toward the new base station in order to make this base station ready to **accommodate your incoming call**. If this as in this case is done by mobile switching centre has the advantage of preparing the new base station to take care of your communication, it's important hat you are able to manage the communication and don't interrupt that during this movement.

If you can't manage this you can't offer an handover service. The handover service should guarantee that **during the movement** the communication is kept alive. This can be done by you see how many **control messages**, this is a qualitative example. That are exchange in the network to keep the communication active during the movement. There is a initial exchange, step 1, step 2, 3, 4, the new base station sends the signal to the old base station saying I'm ready and the base station tells to the mobile please move to the new base station, ready to accept your communication, please move to this. Then the mobile is **activated to the new base station** and can transfer the data to the new one.

Remember that the **mobility** as I said also some days ago is in a cellular network the **most critical aspect**, we discussed last week the very initial solution for providing wireless connectivity, wifi is simple, even if it lacks all aspects related to mobility. Why? Because mobility, initially designed for cellular network, has **a lot of control information** and also the device should be able to manage a lot of control information, to repeat this you need to add to the wifi base station all the information able to manage this kind of control messages. If you then move to a mobile switching centre to another one, all this happens if you move into an area covered by the same mobile switching centre, because there is the central entity having control of the different base stations. On the contrary **if you move** and it may happen from a mobile switching centre to another then there is much more control information and again you have an anchor mobile switching centre, from which you started the communication, maybe a visited one, this is the first one where you have been located when the call arrived.

It exists a mobile switching centre when you were located when the call arrived and becomes the anchor, if it's not in your visited this can be the case, mobile switching centre and you can move from an area controlled by one mobile switching centre to another area controlled by another, so what happens in this case in your opinion? If you move from an anchor mobile switching centre to another mobile switching centre? In this case we said there is this mobile switching centre **managing the different base stations** under its control. Here it's not anymore the case, so how it can work? As before, it happens that there is a sort of **multi hope path** starting from the home mobile switching centre, toward the first. Anchor mobile switching centre, from this mobile switching centre then to the next and then to the next. There is this multi hope path for **managing your mobility.** During the call, if this happens without a call it doesn't matter.

The problem is when it happens during the connection. I notice in some places in the south of Italy the highway if you move very fast it's like probably that you change the mobile switching centre. If I move in a city and I move for instance by walking is quite complex to change the mobile switching centre, if you move in car, train, in a highway it may happen that you change the mobile switching centre. If this mechanism, doesn't work well **your call goes down**. For instance I experience this several times, during your movement the call fails. This is due to this fact, this mechanism here doesn't work very well, it's a little bit complex, you exchange a lot of information. So you see everything works in this way and you continue to be related to an anchor mobile switching centre having the control of the next visited mobile switching centre. I was saying the **two systems are very similar**, what is a home location register, the home agent, what is the visited mobile switching centre in the GSM, the foreign agent, what is a roaming number, that is something we don't see in case of GSM or the next cellular system. We don't see the temporary, we have another phone number, we don't know. This is done in this way, is named **roaming number** and is also the code used to apply roaming mechanisms. It's related to the mobile switching centre that assigns you the visited number that you don't see, a temporary ID for your call. Is the one where you apply the roaming mechanism, also used for billing purposes.

The billing is based on the **roaming number**, still you are in a roaming area where your number is recognised as home area, you are billed in the way. When you move to another roaming area you are billed in another way. It happened to me, I was in San Diego, I have done a flat pass billing mechanism to be billed during my period in the US and I was billed a lot in a given moment, very lot, for my data exchange. I had this flat account, so I don't know why. It happened I was interconnected to a mobile switching centre in **Mexico,** from San Diego. San Diego is very close to Mexico and my account was assigned to the Mexican area, even If I was physically in the US, so I paid a lot for this. This is related to the network, not able to recognise I was physically in the US.

To conclude this, the mobility is big part related to the management of the data exchanged in a wireless context. In case of data so not voice calls but for instance in case of IP this is harder to be managed because there are some mechanisms that don't work well when you manage mobility, you quite well know the behaviour of TCP, works by assuming several things related to your connection, connection oriented mechanism, to state if there is a congestion so you remember there is the congestion avoidance and so on, this works well in case of **fixed communications.** When you move you may have loss of packets, maybe you can have loss of packets in case of mobility because for a time period you are **not updating correctly the movement** of the device and you can loose packets during this exchange. TCP when some packets are lost assumes with some protocols that lost packets are elements indicating the congestion of a session. The network congestion, so what TCP does when there is a congestion? If TCP understands there is a congestion what does it do? **Reduces the data rate.** The reaction of TCP in a case where packets are lost but lost due to mobility the reaction of TCP is to reduce the data rate, that maybe is something is not needed, maybe you want to improve the data rate because you are moving, instead the reaction is to decrease the data rate. Only to say that unfortunately some protocols important protocols like TCP were designed having not in mind the idea to manage mobility.

While cellular network were designed from the very beginning taking in mind it was a wireless communication and where mobility can happen, this was true for the cellular networks from the second generation, third, fourth, and so on. Everything natively was designed having in mind wireless and mobility. IP and TCP were not designed natively taking in mind **wireless and mobility.** Wifi was not designed having in mind mobility. In the other two cases there are some drawbacks that were not encoded in the very initial standards. Every time you improve a standard starting from the basic solution sometimes it's easier to restart without adding pieces to avoid standard, sometimes it takes much more time to add what is needed to have the technology to manage mature that specific aspect. We can have a break.

Going back to the lesson, as for the cellular networks obviously it's a system much more complicated and one aspect important to manage the resource assignment and the mobility is the fact that the system is **structured in cells**, what is a cell? We said is an area covered by a base station, typically a cell is represented as an hexagonal form, in real they are not symmetric in the shape, it depends on the used antenna. In general a representation is this one, an **antenna** covers an area, typically at the beginning, in general you assume that you have a circular area around an antenna and this circular area is where the radio signal transmitted by a base station arrives. In the cellular network since we use a given spectrum frequency band the idea is to reuse a frequency band at a distance that depends on this pattern, and specifically this concept is named frequency reuse, a base station uses a given frequency band, used by that system. But the **frequency reuse** can be provided in different ways, for instance in this case in this picture there you have a frequency reuse with a given pattern, n = 4.

This means that the frequency used in cell 1, in the middle here, the frequency is used after this pattern so 1, 2, 3, 4, you see, around, and you have another one at the border, in this area if you look in the circle there are no others cells using frequency 1. On the contrary if you for instance have a **reuse pattern** you see the radio where the frequency is reused is larger. If you see here a frequency reuse and equal means that again the amount of cells into the same circular area is larger. So we have a much more distances of cells using the same frequency. Which is the advantage and disadvantage? The advantage is that if you reuse a frequency at a given distance the 2 cells using the same frequency **interfere less**. The frequency reuse is something designed to **reduce the interference**. Much more rich is the frequency reuse pattern the better is the interference. If you have for instance a single cell using the whole spectrum band you have a lot of interference, the advantage is that you can improve the capacity, in some ways scales as the number of cells we have in a pattern, in terms of capacity this is better than having 4 cells, one cell. So this is the advantage.

What is the **disadvantage**? In general I would like to have much more cells. Because every time I had the possibility to reuse a frequency at a larger distance is better. But? The trend is that we reduce the size of the cells, if here we can have 80 cells is better than having one. How can we manage to have 80 cells here? What does it means in terms of system? What is a cell? A cell is an area covered by a base station, so the difference in having a frequency reuse 4 or 7 or 19 is that we have in this case **much more base stations** than the other case than the other case. So the trend today is to decrease as possible the number of cells. Indeed in the story of cellular network they start with big cells, then go to micro cells, then to pico cells, ok? They are decreasing, this is due to this fact if I can decrease I can reuse the same frequency at a distance that allows the **absence of interference.**

But if I decrease I have to put an antenna, a base station in every cell. And the base station is a device that has a cost, so obviously the structure where we have a lot of cells means that I need a deep structure of base stations and sometimes the base stations are much more so the cost is higher than having less users. The **tradeoff** is, do I prefer having much more cells, more capacity and more users, spending in base stations or the contrary? This is the point. We will see in a while how this has been for instance proposed in the current cellular system that is **LTE**.

LTE system

What is LTE? You know is the 4th generation standard started some years ago and the idea of the standard was at that time to improve 3 aspects related to the connectivity everywhere at high bandwidth for every kind of service, so a **broad system**, not only designed for broadband communication, for every kind of user and every kind of application. There were some important goals and improvements compared with the other solutions. The improvements in case of LTE were mainly on the how to achieve **broadband communications** so ubiquitous broadband access by using some technological aspects already present in other solutions, this is what they have done, this is the main characteristics of LTE and we will see how these have been improved.

First, which is the **user bandwidth?** Of LTE can be of different sizes, there is a lower bandwidth amount, 1.4 MHz, the size of the bandwidth, there is one at 3 MHz, 5 MHz, 10 MHz, 15, 20, obviously as you already know every time I enlarge the bandwidth the capacity, the **bitrate improves.** So in this table we can assume that the bitrate at this bandwidth is less than the bitrate we have here and so on. Indeed this is achieved because in each bandwidth we can structure the channel in **sub-bandwidth pieces**, named in the LTE resource blocks. They are elements of the system that can be assigned to the users. So you see from this picture that in case of 1.4 MHz we have 6 resource blocks, in case of 20 MHz we have 100 of resource blocks, the amount of resources we can use in a bandwidth of 20 MHz is 100. In 1.4 MHz is 6. Now, which is the capacity we can use **in each resource block?**

This depends on the modulation scheme, so a modulation scheme is depends on how many bits I'm able to put on my symbol. So, if I can use a 16 QAM is better than using a QPSK. How this is managed in case of LTE? In a way that is very interesting and related to this channel quality indication. How it works? The base station in LTE this base station is named as it was named and the name is **eNB**. eNB is the name of a base station, how it works the **signaling**? The eNB sends in the cell a reference signal, a down link reference signal. This reference signal is sent to have to different mobile devices an indication of their quality. It's like I send you a signal and I say to each of you to send me back an indication of the quality of the signal. You send me back an indication, I see **medium quality signal**, someone else sends back. In the past cellular system the idea was to transmit the signal at the **same identical power** with the same identical modulation. The advantage was that it was easier, the disadvantage is that a modulation cannot be personalised, on the contrary in this way the base station can personalise the modulation scheme to the **specific channel quality indicator** sent by the different users.

If he has a high quality signal I can send to him use toward him a modulation scheme that is a high qualitative one because he has a high quality signal. If he has a low quality signal unfortunately I have to spend less performance because you are not in the condition to have a high quality reception, in the same communication I can send higher bitrate to him and less bitrate to you. The first time it has been proposed in a cellular system, in the past **the approach was different**, there was not this channel quality in explicit channel quality indication, and instead with this you can adapt the modulation scheme, adapted to the different end users, the end users in case of LTE again the term taken are named **user equipment**, they are the end device, the cellular phone, while the base station in this standard is named eNB. And then the architecture is taken so again there is a node in the core network managing the different base stations that are the different eNBs that cover the different user equipment.

This is the **basic architecture.** As for the modulation scheme these are different basing on the channel quality of users given to the different eNBs. Second aspect, how the multiple access is managed? Then now we see this, a third aspect is again as in other systems we have an **asymmetric capacity,** so we have much more capacity in the down link and less capacity in the up link, similar to what we have seen in other asymmetric systems, the only fact we have now in a wireless one. So these are the main characteristics, let's see what happens. As I said for the modulation is adapted to the specific channel quality, so if you have for instance a **low channel quality** you use a simple basic phase shift key modulation. If you have a high quality channel you can go up to a 64 QAM. So for instance here you can see that if you have a time during time a channel a signal that is varying in a time period you use modulation scheme, for instance here you have a very bad channel, so in a very bad channel you use or in a good channel you can use a 64 QAM. And this is a representation of a time varying. Adapting modulation is the first important aspect, in other words it's tailored the modulation it's tailored to the specific characteristics of the user.

And ok, this is repetition of what happens. So this is representation of for instance how modulation schemes are adapted depending on the SNR, represented in the standard by this channel quality indicator that is **4 bits** used in the control message to say low quality, if the 4 bits are all 0s the message is not reachable, you cannot send anything to me, I'm too low. Then 1, 2, 3 till 15, so when you have channel quality indicator 4 1s you are at **the higher signal** and then the modulation can be spend over this end system is 64 QAM. First aspect is possibility do adapt modulation to the end user, second aspect, this is also a novelty that has been included in the LTE, all these novelties are related to improve the data rate, this one is related to the so called trade off between **forward error correction** and automatic repeat request. What are these 2 things are? Forward error correction if you remember is something that allows you to put in your data **some extra bits** to correct the bits that arrive at the end systems, in wireless systems you typically need a lot of forward error correction if the quality is low and less if the channel quality is good.

But you have to plan at the beginning **how much error correction** you have to put. As for the automatic repeat request if an error happens. Again, planning to have always an automatic repeat request maybe not convenient if you have a good channel quality, into a system. Which is the idea in this case? The transmitter puts some forward error correction, if the packet arrives in a wrong way, not correct, the forward error correction was not enough the packet must be retransmitted, when the packet is retransmitted, this is the name of this solution, **hybrid automatic repeat request**, why hybrid? Instead of discarding this packet the receiver continues to keep the packet and adds the new one that hopefully arrives in a correct way or arrives with a forward error correction that can be combined to this in order to improve the capacity of recovering the errors, **error recovering** means a good amount of error correction, if you don't want to put this from the beginning because you don't know how is the behaviour of the channel this extra information is put in a way that may fail in the correction, this can be kept to improve to make it stronger when a copy of the same packets arrive. Increasing error correction done by using the **retransmission**, so every time you retransmit a packet this retransmission may allow to combine the information that previously received with the new one, thanks to this combination of bits received you can increase your capacity for correcting the data. Done in different ways.

Next step is the management of the bandwidth, this one. We have a channel bandwidth, expressed in MHz, this is divided into slots of pieces of bandwidth that are named resource blocks, that are of a different numbers depending on the channel bandwidth. How the transmission works? This happens in different ways depending on the downlink or the uplink. In the downlink we can use a so called **orthogonal frequency division multiple access**, that is I can think of my transmission in frequency and in time and allow a user to receive the information in a combination of time slots and frequency slots. So it's a sort of having a matrix and this matrix has cells that are a combination of a given time slot and a given frequency band. In the orthogonal frequency division multiple access the users have to be for instance assigned in some frequency slots and time slots, in an orthogonal frequency division multiplexing, users are transmitted in different times by using the whole time slots, in the second case you assign time slots and frequency, in the other case you assign always the same frequencies but divided in time domains.

In the downlink you can have **one of the two**, obviously this has the advantage of a higher granularity, because the granularity in this case is of an element that is a combination of time and frequency. In the other case only the time slot, here you can say to a user 1 or 2 or 3 time slots, instead here you can use frequency 2, 3, 4, 5, 6, 7 to the other user. This is much more flexible in the assignment. This is the advantage, this is as an advantage. The drawback is that in this case which is the drawback? We are comparing the case to a user we assign a map of time slots frequencies elements, in the first case you assign only the time slots. This one is much more flexible in assigning capacity, in terms of granularity. The other one is less flexible.

To this lo demanding user I can assign this. I can assign multiple resource elements, it's a little bit more complicated, the assignment happens in a time slot, a time slot then is divided in different resource elements, combined in time frame that is 10 ms of time slots, this is what the standard says. This go to low qualitative presentation, this one. So, the difference in these 2 cases is that this is much more flexible, but? Is more complex, only this. I need to signal to the user what they are going to use. While this is quite easy to be used in the **downlink**, what is the downlink? Is the direction from the base station toward the end user. So this is the down link and the other is the up link. The downlink is something controlled by the base station, for the base station to let's say implement this mapping, is feasible because everything is decided by the base station. The base stations says I'm transmitting some bits, some of these have to be put here, in this time slot, other here, then next time slot the same configuration, everything is managed by the **central entity base station**, in the up link instead this is much more critical, because how the up link is implemented, by having an end user in a position, another in another position and I have to say to all users in this time slot there is a user transmitting here, another here, another here, so it's **not centrally controlled,** indeed in the up link it's used a strange access method that is named orthogonal frequency division multiple access, again I have different frequencies, and the idea to have different frequencies is a valid approach in the LTE systems because these are managed as in case of ADSL as they have **orthogonal transmissions** on sub carriers, you remember the concept of sub carriers, if I have a big spectrum band I prefer to divide the band in sub carriers in order to have a sub carrier flat.

And this allows you to transmit with a given quality and so on, so I do the same but while in ADSL we have used this sub carriers for transmitting different signals of the same users here I can imagine to use the sub carrier as sub carriers managing signals coming from different users, this can be done in the downlink in an easy way, like it's represented here, I can imagine in the downlink I put a colour for the different users, int he uplink this is much more difficult so the uplink is a single carrier frequency division multiple access, that is the use of multi carrier, that has the advantage we know but **as it was a single carrier**, how? By transmitting the same user on all sub carriers. So you see the colour, red is a user. The user is a single user using all sub carriers. This is the advantage of using the multi carrier modulation like in ADSL given the fact is a modulation but done by avoiding the splitting of the users that is critical when you have to provide an up link. So now we stop again, we have **2 possibilities**, to use in the downlink orthogonal frequency division multiple access, managing the with orthogonal sub carriers with the whole flexibility we would like to have. Because I'm in the downlink and there is a single entity, the base station. In the up link we would like to do this but it's difficult to be implemented, in the uplink I use the advantages of the multi carrier modulations but in a single carrier mode. By transmitting all the elements on all sub carriers. Let's have another break.

Laboratorio applicazioni telematiche

Going back to this picture, the concept is the same of the previous representation. Let's put the frequency here, in general if I have a frequency band, if I want to use a **big frequency band,** the best way is to adopt a frequency division multiple access, in the orthogonal form. In quite huge number of sub carriers, so in a scheme that can be used in the down link well let's do, some sub carriers is used by a user, another set of sub carriers by another user, in the next time slot this used by another one, give the **whole flexibility** and this one. So I can put on the different sub carriers different colours, in the uplink given the fact is complex to implement on different sub carriers the contemporary use by different users I can use still the sub carrier mechanism but having a user transmitting on the whole sub carriers in a given time period, then another user on the whole sub carrier in another time period, then another user so you see this layer of colours that are the different users changing the use of the sub carriers. In this way I exploit the use of sub carriers so the advantage but I do an **access scheme** that is quite easy, so this is the scheme for the single carrier, I take the symbols of the user, I parallelise the symbols in n parallel branches, every branch for every sub carrier and I put the whole bits of the same users on the n sub carriers. This is the implementation of this **single carrier frequency division multiple access**.

Then this can be transmitted as in a classical OFDM scheme. I want to say this, that only this last aspect, here we have sub carriers on the y axis, the standard says sub carriers and time, sub carriers are 12 and 12 sub carriers represent a **resource block**, a resource block is this element in the LTE and a resource block occupies 180 kHz. The data rate that I can transmit on a resource block is something related to the bandwidth of that resource block and to the modulation scheme that I use on that blocks. The same resource block can transmit more or less data rate depending on the modulation scheme I can use and depends on the sub channel quality indication. It's quite complicate as a system to be honest LTE and 5th generation is a combination of all the advanced done in the wireless cellular technology, all combined in this standard, there were requirements particularly challenging. A standard like this is a combination of the best aspects of all technologies.

The last aspect is this one, while in the past it was used to transmit by using a single antenna, after a while the telecommunication researchers discovered that the so called **multiple input**, **multiple output,** also called MIMO, transmission is much more efficient. Consider that always all the aspects are related to the fact we would like to improve the data rate. So improving the data rate in a very challenging environment where the wireless communication in the indoor environment where we have the multi-path, the mobility, the interference and so on it's something that must be faced and all these aspects, the **adaptive modulation**, the management of single carrier versus multi carrier, all these aspects are related to the improvement we would like to include in the technology. So **MIMO** what is? Is this fact, if I transmit by using multiple antennas at the transmitter and receiver I can improve the capacity, there are 2 ways to see this multiple input multiple output, the first way is transmit diversity. If I put 2 antennas these will encounter a different channel.

They are transmitting at a distance. Even if this distance may be very close. But if they meet they pass through different channels I can use their transmitting diversity to make it robust the reception, how? You remember we discussed the fact there is an **attenuation**, if I can manage I receive a message from an antenna and a replica from another one, I can let's say exploit this channel diversity in order to announce the amount of information I can receive, it's a sort of **redundancy** we can have. This can be a redundancy in terms of replicas of the same original signals, if I receive two replicas, I can use the replicas to make it robust my receiving signal. In other words, if I receive two times the same signal a bit changed and I can understand which are the changing I can make the reception with a higher signal to noise ratio, in fact. I can improve the SNR because I receive multiple replicas.

Or, another aspect is I can do as **partial multiplexing**, with two antennas instead of transmitting the signal on a single antenna I transmit partially on antenna and partially on the other one, I do a real redundancy in space and again I exploit I transmit the two components and this simultaneous transmission can give me the advantage of receiving more data. So these 2 simple concepts have been used in the multiple input multiple output systems and these MIMO systems we don't go into details, is like I can model my channel like a behaviour represented as matrix, so we can have a channel that has a behaviour with a different behaviour at different time instance on different paths. If I have our paths for our transmitting entity I can have different behaviour on different paths. This can be used as I said before to improve my reception.

Thanks to MIMO I'm able to **improve my capacity**. In this picture we have the classical capacity of the SNR. If I use an m1 MIMO equal to 1, a single input single output this case I have this performance, if I put 2 antennas transmitting and 2 receiving I have this performance, if I put 4 antennas I improve again the performance. There is a performance improvement as the number of antennas. This is the advantage, the **disadvantage is the cost**. Putting these antennas on the same device maybe can be done MIMO, where these input and output are? In the transmitting and receiving device. The transmitting may be a base station. The receiving? The end device. This means that MIMO is multiple antennas on this. And obviously it has a cost, it drains the battery and so on, but it's the technology that now is mature also for these devices.

Not only this, thanks to this processing I can by managing the weights I give to the different components, branches multi path I can do another important thing that is named **beam-forming**, something like this. I can form the shape of the beam directed in a given cell. While in the past in the other technology a cell as I said was a base station with an omnidirectional antenna, covering in the all directions with the same behaviour, if I use a combination of the MIMO and the coefficient I give to different antennas I can form a shape, named beam, the shape of the signal I sent in my cell. And this is another advanced mechanism because by acting on the coefficient of these transmissions I can **form a shape** not only but also with the shape can be easily changed in time. So you remember we mentioned before what I would like is to have **small cells**, because my interest is in reusing the frequency. Thanks to this approach you can reuse a frequency for the different beams formed by the same base station, thanks to this a bas station may manage in the same cell different beams managing different frequencies, a sort of frequency reuse achieved using beams. This is an advantage, the second is that all these can be provided dynamically and on a software basis. Everything is done by managing the powers I give to the different antennas.

This is the LTE very briefly. After the LTE in the first definition after this there was an evolution of the standard named **LTE-A, advanced**. The advanced was in different aspects and specifically MIMO already present in LTE improved 8x8 antennas. LTE was using 1.4 till 20 MHz. LTE-A uses up to 100 MHz. The problem we will see on Monday is that when we use much more bandwidth the data rate increase but in the current cellular system there is no way to have a **unique bandwidth** of 100 MHz. Too big, not present in the current spectrum radio a bandwidth of 100 MHz. How this is solved? We will see on next Monday. Already present in LTE, is that speed of the user equipment is also important, we would like to support mobility, as you have seen has an impact on the amount of messages exchanged, and also on the channel behaviour that I have when I move. If I move very fast like for instance a fast car or a fast train this means I have to exchange very frequently the cell and I also change very frequently the channel that I need. To be able to model online the channel in order to be able to adapt my transmission this needs to be done in the right way. In the past this was quite complex, with LTE this was able to be done, this is another novelty of the standard. This is a story, different releases of this approach. Not all of them will be discussed. We can stop here today.

Project presentations

We don't give a fuck. Today we are going to close with the lessons and I would like to briefly summarise what we have presented last lesson on LTE, I give some indication on LTE-A, the evolution of LTE, as for LTE last lesson we have seen this is the standard for the 4th generation of wireless, based on the previous standards, UMTS, previous again like GSM, there are some key important evolutions, there are 4. They are first the fact it's feasible to **adapt the modulation,** on the basis of channel experience of the user, every user in LTE is able to send back to the base station an indication of the current channel quality, indicated in a specific field of a control path. Thanks to this indication the base station is able to adapt the modulation on the basis of the perceived quality, this is an exchange performed on top of this infrastructure and the result is that we have seen the system is able to tune the modulation scheme. In the field of the **control,** in LTE everything is designed in order to improve the quality, in this case the quality improvement is achieved by ah hybrid, means that the tradeoff between the amount of forward error correction that is to be provided in a packet and the amount of time the packet is retransmitted is adapted **dynamically** on the base of the amount of retransmission.

Since it's hard to figure out which is the quality of the channel the system considers both the approaches, forward error correction and retransmission, this is why is hybrid. Combining the two is able to improve the quality. If a packet arrives and is wrong a negative ack is sent, next part will be sent with less forward error correction but even if it was wrong there are some contents in the packet that can be used to improve the quality. There are **different ways to combine,** reaching an higher performance on bit error correction. Another aspect is that is possible to allocate with a quite flexibility the spectrum resources, organised in a complex way, by having some elements named **resource elements,** included in other elements named resource blocks, that are achieved by having a frequency allocation that is a frequency division multiple access allocation, then combining the frequency allocation with time allocation. A combination is a sort of matrix, and on top of this matrix the allocating entity that can be the **resource manager** is able to indicate to the different users the amount of resources. This can change dynamically, it happens that every frame period and the frame period is in the order of 10 ms, indicated here, so every 10 ms there are a number of time slots into the frame, like always, a time slot that can be figure out as a combination of time and frequencies, this resource blocks are dynamically allocated to the different users, on the basis of the requirements.

In order to give services to the users you can play in different ways. Last aspect, is the fact is possible to operate a **multiple input multiple output (mimo)** transmission, with different antennas as input and output. This is a way to capture all the characteristics of the complex channel that is met in this transmission, that can be modelled as a matrix, this is representing channel's behaviour on every path and by playing with this channel behaviour and arranging the parameters is possible to form the right way to transmit toward a given receiver. This is the scheme for instance you can provide multiple output on a smartphone, this is the novelty of this technology, thanks to this you can use mimo as a way to transmit in a **efficient way**. This can be measured, given a SnR while with the classical 1-to-1 transmission you have this performance, with mimo you have a sort of enhancement of the performance, it depends on the adopted scheme, the basic performance behaviour, this is the capacity, it's improved thanks to the usage of the mimo.

These 4 elements are all together used at the end in LTE in order to achieve a good a very high performance behaviour, you can notice from this picture for instance by comparing the performance of classical wireless old generation or wired system the performance are very interesting, you can measure up to 300 Mbps in a wireless system is a success of this kind, in the **downlink.** In the uplink up to 75 Mbps, a big success for a wireless communication. Than also other characteristics, the coverage, full performance up to **5 km**, also interesting, and also a challenging aspect of this technology was the possibility to support this kind of performance on a high mobile network. Also the mobility, introduces a lot of problems as for instance for the channel behaviour and all this has improved. The LTE-A is an evolution and the idea is to improve the capacity, the bandwidth, the possibility to optimise the network and so on. Which are the main aspects? One aspect, in order to improve the capacity the idea of LTE-A and the possibility introduced is the possibility to perform a **carrier aggregation,** you notice the availability of a high bandwidth in MHz is complex. Several times is not feasible to have a large bandwidth in the spectrum band.

Carrier aggregation

The idea of the standard it's let's try up to 100 MHz in the spectrum band. Quite difficult to have 100 MHz in the spectrum band, very large bandwidth, a high bitrate only when you have a high spectrum band. Carrier aggregation is a way to **aggregate** spectrum band portions that are not available on the same spectrum but by combining portions in different sub bands. So you have a spectrum and on top maybe you have some parts of the spectrum that are free from other communications, but they may be part of this **carriers**, so part of this carrier available here and another here. Carrier aggregation is like the possibility to have and consider n transmit in the spectrum band as this spectrum is a uniform. This is the novelty. In general when you have a transmission you occupy a spectrum band. The size of this spectrum band gives you an indication of the amount of the bitrate you can achieve. But if you are able to aggregate and implement a transmission scheme able to transmit on **different spectrum bands,** considering and operating as these spectrum bands are a unique band, this is the advantage.

The idea is this one, instead of transmitting in a single, you transmit on multiple spectrum band and you consider, this is complex, it's not on a single carrier frequency, is a multiple. The combination of this multiple are considered by the system and transmitting entities **like a unique spectrum**. This is a novelty. This can be done in LTE-A, if the spectrum bands are continuous in the same range or even if these are on different spectrum bands. So this is the first important novelty of the LTE-A named carrier aggregation, the benefits are multiple, you can play with the transmitting devices in order to operate at the best of different bands and considering this spectrum present in the specific considered region, area. Another important aspect is if you look more wide not all the spectrum bands are occupied in the same manner. It depends on the specific region, area, in that region you can have two spectrum bands and the third not available, all three spectrum bands available, in another only in this. This **flexibility novelty** and gives the possibility to go from 1.4 MHz up to 100 MHz spectrum band. At 100 MHz you have the possibility to transmit up to 1 kbps.

Another important aspect in this solution as mentioned before is **the use of mimo,** in this case you can operate with a mimo of 8 by 8 in the downlink, and the mimo has also the possibility to operate like this, we mentioned this fact, you have the possibility to give a shape to the beam, this is named beam forming. The beam where the signal is transmitted can be formed with the given shape. This is very interesting, because you can transmit your signal like the station here **toward a mobile user** and maybe also neighbour user doesn't receive this signal but it's own signal. This is different behaviour with respect when you have a omnidirectional antenna. A classical in the past, I have an antenna, this antenna is **omnidirectional,** so a user that is here is receiving the same signal received by this other user. This is important, this user when he transmits he interferes with this one. On the contrary, if you are able to a beam transmitting to this one, and another to this one, these 2 **do not interfere** anymore.

This is the advantage of this solution, has evolved in direction of **massive mimo,** an antenna very large of arrays, you have these arrays that can give a very precise shape to the different beams, thanks to this very precise shape you can contemporary transmit to multiple users. I would like to remind this, if you remember the use of **frequency reuse** concept, if I use a given frequency in a given cell I would like to avoid the use in a neighbour cell and reuse in a different cell to avoid interference. In massive mimo is like every arrow here is like a single cell. It is a way to look at the system at a system with very small cells, like a cell for every n users.

There is an area, that typically has an average number of user. But in a time period there is an increasing of the number, so a single antenna maybe by itself it's not able to serve all the users. But, communicating with another antenna close, this antenna can improve the capacity of the previous antenna in order to provide a server to all the users present in the antenna area of the previous one. This can be dynamically arranged to give **different possibilities** to the system. This can be performed with several technologies. There are lots of algorithms and solutions to perform this cooperation between antennas in order to provide on a dynamic way the capacity to the users on the basis of their specific requirements.

There is also in LTE another important novelty, that was not used and presented the first time in the cellular system, the possibility to perform the so called **device to device communication.** Something that is the possibility to perform a communication in an ad hoc manner. You don't rely on an infrastructure, there is a direct communication. The novelty here, while you know ad hoc has been proposed several years ago as the wifi. The ad hoc mode in wifi already existed. We discussed in that case you can also use another node as a **relay node**. And then go to another node. The difference here is that first it's allowed to have device to device communication in LTE. With the same technology used by for instance an end user to communicate with a base station, this between different devices without the involvement of the base station. This opened the system to a new services that can be provided here, because for instance there is today we are facing different solutions proposed in the field of the LTE as for the machine to machine communication and as for the internet of things, both machine to machine communication and IoT will be provided by different possibilities thanks to the fact there is the device to device possibility in the LTE.

This presentation is quite reach, I will skip some aspects, too detailed for us. It's not here but I would like to mention that for instance the device to device communication in LTE will be maybe used for the IoT, remember that in this case different standard this is a **license spectrum**, and the idea is for instance to use some band regards, some bands typically used like a band that is to separate with another frequency band, so let me see if there is a picture. In that case the idea is to use your system some bands that are typically not used for the communication on the systems in the LTE but are used as a guard band between another frequency band used by another system, you put the low rate communication there, it's a short one, typically 180 kHz. One of these, operates like a guard band, used for low energy, **low rate communications,** for having communications in that technology. This ends this presentation, obviously there is a lot to do.

Core network

Last part is relevant to the core network, what we missed is how the core network is operating. I have to run a little bit but I would like to come back to the initial solution, based on the idea that what a **core network** has to do. It has to multiplex flows. This is what a core network is to do. And the very initial solution to multiplex flows was presented in a very old standard named digital hierarchy where this digital hierarchy was intended as a methodology to multiplex flows derived by the e1 t1 frames. At the beginning was born by multiplexing flows deriving by the basic building blocks of the channels. On top of this full infrastructure for the home network has build up and this infrastructure is something that operates let me check if there is a picture of this, operates at the layer 2 let's say. It happens here is that we have a physical layer, this physical layer has a structure, on top of this there is a **quite complex infrastructure** of layering and this infrastructure you can think at this part as where different layer 3 or layer 2 technologies are multiplexed.

In other words there was the packet networks and other kind, on these were multiplexed in the core network infrastructure offering a concept of path, and a frame structure to support this path. So at the beginning this multiplexing was done very easily by having flows coming out from the circuit switching frame structure, we discussed with 32 slots in a frame, and to combine **flows of bits** on a line that goes at a higher speed. This multiplexing was achieved very easily by a multiplexer, a device preserving input from multiple flows and gives you back an output of a single flow at a higher speed. In a core network the main initial important element is the multiplexer, what is? Something that receives multiple flows as input and gives back a single output flow. At the very beginning these multiple flows were achieved by receiving bits coming from for instance the e1 frames. Several years ago, 40 years ago, the multiplexing was done in every node, the switching was done in the time domain, the multiplexing as a **single operation** already was a complicate operation. The different flows were arriving at the multiplexer in a sort of asynchronous mode, meaning that the clocks could determine the time period where the bits arrived, are not synchronous. How this can be solved?

If you have to multiplex 4 flows in a output line that goes 4 times at the speed the speed of this if the 3 flows arrive exactly with the same identical clock you can do. If this is not the case you can do it in this way, operate at the **frequency of the speed** a little bit higher that n times the speed of the flow and operate a so called pulse stuffing, I show you this. Imagine an incoming flow, in a multiplexer is named tributary flow. This is sent out at a speed that is higher than the incoming one, it has to be multiplexed. In picture there is the tributary flow and the tributary flow after this acceleration. So it may happen that, look here, there is a bit arriving, and the bit is sent out, another bit arrives and is sent out and so on. Now in this time period you should receive bits but the bit is not arriving. why? The clock is **a little bit different**. This one arrives at a speed that is maybe a little bit lower than expected. Less bits has arrived. So is a way, adopted in the multiplexer, where if you do not have a bit to put in the output line you put an extra bit to compensate the missing bit. Thanks to this it was feasible to multiplex flows arriving at a nominal frequency that is not exactly the speed of the output. If there are missing bits you can compensate these with a pulse here and there.

First question, **how do I recognise the added bits?** These are extra bits. When I have to demultiplex I have to omit bits that are inserted. How do I recognise the extra bits? Are these always in the same position? Consider that the input nominally and vary the arriving rate, a little bit faster or slower. On average it's exactly as expected, sometimes faster and sometimes lower. A multiplexer reads, receives an input and gives the output. The input should be received at the same speed, but sometimes the **speeds are different**. You can have less bits than expected, or more bits. Depends on the speed. To compensate this a solution is let me read so transmit the output always at a speed that is faster than the input. Because if I read the output at the same speed I may miss some bits or add some bits. If I decide to read at the speed a little bit faster what can happen is that I miss some bits. understand? Miss some bits but I can compensate this having **dummy bits.** I can write and read at a speed faster and if I miss the bits add an extra block. This is the extra bit.

The problem is this works, but the problem is that I don't know the position when and where the extra bits is added. This solution was a solution adopted to manage multiplexing in the network when the network was not sure that the arriving flows where on at a given speed. This kind of network was named that is not synchronous, a form of not synchronous network. There was a period where the core network was not synchronous and due to this the multiplexing, was done by adding extra bits. The only drawback is that there was another kind of information that must be transmitted with the information itself is **the position of extra bits**. Otherwise I can't get the real information. So, this moreover is also repeated at every stage of multiplexing, in a network you have a first multiplexing stage, then you go deeper in the network, another multiplexing stage, then you go deeper and have another. So imagine we have seen the access network, the edge, the core, a much more core, every stage **you have a multiplexing.**

What happens if you have to add at every stage extra bits? The amount of extra information, overhead, is great, and you see a problem in this? The problem is that if

I would like from this flow, the multiplexed one, I would like to pick up the single flow this blue one, so I have a **bit flow into the network.** I would like to pick up a flow from a single user. How do I do this in case this flow is achieved by operating like this? Consider this example it's clear, indicated only the composition of a multiplexed flow by looking only at a component of this multiplexed, that is one flow. You see already this one single flow is not anymore the original one but it's the original one **plus some pulsed stuffed bits.** What happens if my multiplex is of several tributary flows and more? Several tributary flows multiplexed at different stages. How do you imagine this output? In other words, let's consider 4 flows. There are the bits, if I would like to multiplex these flows the output is the same compressed at 4 times the speed. So I would like to represent I can represent this with the 4 bits all together in the same time period. So I have this one here, this one here, this one here. And so on. So this is the line going at 4 times a single line of this.

If I do this, what is another way to represent this, if I do this and by doing this operation some time periods I don't have the original tributary flow, because for instance it happens that this one instead of arriving at the same speed arrives a little bit anticipated or delayed. Faster or slower. If arrives not when it's intended to arrive I have to put in this time period a dummy bit. Somewhere in this composition there is a dummy element. Coming from one of this, the one who failed to arrive. Here and there you imagine you have this blue blocks in the multiplexed line. This multiplexed. And the multiplexed again multiplexed again and so on, at **different stages of multiplexing,** this is what really happens in the network, start from the periphery, then go in the first stage of the network and so on. If you imagine that picture, here, what you see? Something ordinated or concretely almost random?

This is a representation of what I have to do to switch a multiplexed flow in a switch from an input to an output. In this picture is represented a complex one but if we explain it's easy, is what is going to say, if I would like to switch a single flow, I have to go back so to the multiplex first stage, till the original one, then multiplex again, then multiplex again. Otherwise I cannot pickup the single flow as it was, if I don't demultiplex and remove the extra bits I cannot come back to the original flow. In other words, the fact I have to multiplex at different stages and cannot be sure the flows are synchronous, brings at the consequence I have to **demultiplex everything** at every router, every switch, in order to multiplex again and switch the content from one direction to another. This one was the solution adopted in the previous technology, this technology was named digital hierarchy, BDH, to multiplex and demultiplex a switch data flows into the core network.

To add and drop lower order streams, going at a lower speed, example E1. In this case if you have to demultiplex, to drop a stream E1 from an higher order, meaning that it was already multiplexed, and so on, you have to redo all the multiplexing till the very initial order. This is represented here. So this means that the system was very complex and very slow. This was the very initial solution for the multiplexing flows, after this to overcome the problem due to this technology has been proposed a second solution, named SDH, where this s stands for **synchronous digital hierarchy.** After a while the device into the network were supposed to be synchronous. At the beginning it was not possible to assume all the device were synchronous. Some network nodes had a different speed. So due to this was not possible to multiplex in a synchronous way. Once the device became synchronous it was feasible to multiplex in a synchronous way. In a technology is always important to be **backward compatible,** you have to put new device and be backward compatible.

This solution was provided with the intent to provide a compatibility with the past solution. And the idea to provide this compatibility was achieved with a structure that is quite complex but easy to be understood with this example, let's figure out that the transport solution to transport flows is operating at different levels and different levels can be represented in this way: let's assume our flows are composed by bits, bits in a digital structure named **container.** A container is named virtual container when besides the digital structure I put some overhead, named **path overhead.** A virtual container can be combined with other virtual containers, become a tributary unit. A tributary unit can then be put in a bigger container, that is named virtual container of higher order. On its turn the virtual container can have an overhead and this structure is carried in the network by a part that is named **administrative unit.** The fact is that, the single bits that are multiplexed in a network are structured in this way to be transmitted in a frame multiplexed flow and the advantage of this structure is this one, should I switch a all together a set of flows? yes.

How can I do that? I pick up this, I change the track header, and then I switch. The switching is I keep this like it is and I change the track header. or, should I switch a single tributary flow? Yes, I open this, then I take this, open this and take my container and I put the container in another structure, this structure I have the same solution to switch. This is a representation, thanks to this idea this was achieved in practise into a digital network. And thanks to this structure represented in a digital network it was able to have a very flexible solution to multiplex and switch that is the main operation done in a core network digital flows. Indeed this has been implemented for several years a in a solution named **SDH** and this was also translated when the network moved from the electronic version, the classical electric version, into the optical one. So this infrastructure was reflected in the same way with different methodology in the **optical network**. Today we have bits not anymore bits but wavelengths, multiplexed together in optical paths, and so on. So what was done in bits was then reflected in the optical network in the core and this is currently what we use in the core network. Below the layer 3. So you are quite aware of what happens but then what happens below the layer 3? How these packets are multiplexed? How do they travel in the network?

Last lecture

As for the SDH, last week we started talking about a solution that has been proposed to transport a digital information **in the core network**. This solution is still used and it's also the case of the optical transport so I would like to give you some brief indication on how it works, we are going to spend on hour on this and I ask you to understand what is important here. Thanks to this technology, we discussed on Monday there is the possibility to have a flexible structure where digital information are put. This structure is named **virtual container**, and virtual container is let me say a sort of matrix where bytes deriving from for instance a digital flow in case given by packets or other upper layer protocol, we can have upper layer protocols, there are conceived to transmit their data in this virtual containers. They are grouped in a structure more complicated and this one is put in a frame. Let me first indicate what is a frame.

A frame is this one, a frame so a digital structure that is repeated in the core network, by a big part of overhead and a payload. This overhead is structured logically in columns and rows. The basic frame is a frame composed by 9 columns and 9 rows of **overhead**, then there is the payload, 260 columns and one of these is again an overhead. How this is used in a flexible way? In general, I would like to see the content of my information fitting exactly the space of a matrix like this in a frame, so I would like to put my content in a frame. But my content could not arrive exactly at the frame start and not end exactly at the frame end. A virtual container is a fixed structure I cannot change but this fixed structure can move like a **floating matrix** into the SDN frame, by using pointers. A virtual container is inserted into the SDN frame by addressing the virtual container using pointers, digital information put in the header of the frame. I have a sort of coordinates, rows and columns, where the virtual containers starts.

Into the virtual container I can have other pointers that is an indication where the virtual container of lower order starts. In other words I have to figure out a fixed structure, and this cannot change in time. 9 rows of payload, this payload can be moved into this frame structure and where the payload is is indicated in the overhead thanks to the pointers. The information arrives at a data rate a little bit faster than the rate. How can I accommodate an information that arrives faster? In the past was done by adding pulses, in these structures this is done in these way. This is a container, borderline from a frame, and another frame. So where the container starts is indicated in the pointer, if indication arrives faster I have **more bytes than expected,** the idea is this one, I point at the position that is before the nominal starting, more information than expected, and this much more information is stored, **negative justification**, in a place that must not be used for the payload, that is internal. So while in general the header is dedicated to the control information, if I have the need to accommodate some information that is payload but has no space I put there.

If the information arrives **slower**, I don't have the information so, I delay a little bit the start of information and this delay means instead of having in the payload some bits I put some dummy bits. where? In correspondence of the pointer in an extra space that is provided into the payload close to the pointer. Thanks to this I can have a container into my frame and I accommodate temporary variations. so, very fast, this is the solution provided several years ago in order to have containers of different sizes tailored to different applications and thanks to these containers the problem that was present in this scheme, it was to switch a flow from an input port to an output port, in the past the only way was to demultiplex all the information come back to the original one, take the original one and put in the output and then multiplex again. Thanks to the containers switching can be done by following **the pointers**. If I have to switch a big flow thanks to the pointers I found in my digital structure the virtual container and I switch the virtual container, or I have to switch a low order flow? I follow the pointer and I switch the low order flow.

These network elements are the following, this is quite simple, named regenerator. A device able to regenerate an entering toward an output SDN. This only operates on the **digital signal**. Is the possibility to recover the original bits, they cross a path, due to the attenuation, the distance and so on I need somewhere in my path regeneration elements. And this is easy to be done. second, important, add broad multiplexer, an element where I have an entering SDN, my multiplexer device, an entering and an output SDN. I pick up flows from entering SDN and thanks to this method pointer by pointer I can pick up single data flows as well high rates data flows and I can also add flows, so I can add block multiplexer in a multiplexed line, I can somewhere add flows. By putting the pointer in a frame and by providing the container for that flow. The simplest add drop multiplexer is the multiplexer giving the fact this has a number of input ports and a single output port.

Finally, the much more complicate is the **switch**, also named digital cross connect. What is that? The device operating maybe below a router, at layer 2, this technology has to switch entering multiplexed flow in output multiplexed flow. It's a combination of different multiplexer. Typically in a core network we have 3 elements, multiplexers, in the very peripheral part, so we pick all the information coming from different users. Then we have digital cross connects, and then we have the regeneration sections. Each of these network sections is one from a multiplexer to another multiplexer is named **path**. This one to a multiplexer toward another multiplexer or a digital cross connect is named multiplex section and the sections between digital cross connect and regenerators are named regeneration sections. Every section add their own overhead. So there is an overhead to indicate what happens at the multiplexer section, at the regeneration section, what happens at the path section. For instance, an important aspect is to provide measurements over quality on a path for instance.

Quality in terms also of reliability of path, if it's active works and so on and everything is controlled **by the specific overhead**. This you can find a representation of the different devices, I want to only mention this, if you remember we discussed several times the initial part of the core network is made by using rings. How they are implemented in the network? In general entering this can be the classical portion of a network where let's consider the DSLAM in case of VDSL, a sort of multiplexer, we can have input lines coming from a peripheral area, demultiplexer, like described, and then interconnecting a very larger area by using an optical ring. So the input of output of a single fiber put in a ring section. This is the overall scheme for instance of a real transport network, taken in this case by telecom Italia. This is done like this.

This is how the infrastructure is implemented. I want to mention another aspect, what is important is to indicate that once you have formed an SDN you can multiplex multiple SDN of level one in a SDN of higher border by simply the content of a virtual container of that matrix showed before at the levels. If I have to multiplex in this case, an SDN 4, a frame going 4 times at the speed the initial one. I simply all the column of the basic SDN and I put first column on the first, second column on the second, and so on. I repeat continuously. Also the pick up of a flow at this level from an SDN 1 to an SDN 4, I repurpose and I pick up or compose an SDN 4 by in an SDN 16 a column each for columns. It's a way to structure digital information, every column is a composition of bytes with the meaning given by this structure. It's a sort of memory representing how the flows are transmitted in a network.

There are thanks to the structuring of this technology also the possibility to **provide protection** to the different paths, I can replicate a path, I can have 2 fibers in order for instance if there is a failure on a fiber provide a backup fiber in order to provide connectivity in a link, rings are made by using not a single fibers but 2 fibers. Even rings at this stage, so this ring is made by using one fiber, one maybe a backup one and how to use the backup is depends on the different protection schemes. On the contrary the method is fully replicated at that level, every element here is has a **protection path,** the nominal path and the protection one. We don't have time to give much more details, this ends this part. Instead what is important I would like to present you something related to the **optical network.** Today the core network is all made by using optical components, how these are used? The trend was is the following, I said you. At the beginning there was the digital hierarchy, used on top of a single fiber.

On other words, this structure was a single fiber using this technology, after a while this was not so flexible, not so easy to **accommodate synchronous flows,** was done by providing new philosophy, provided by providing the virtual container concepts, in a multiplexed flow of different orders I can recognise my information by following the structure, a sort of way to read in a flow bits the information of the different users at the different levels. And this was run in several network, SDH. On top of this the optical system became more nature to be provided. And at the end is not really the end, also for sure the multiplexing is **all done in the optical domain**. In the next, not so next, also the routing and switching will be done in the optical domain.

Indeed, in the optical network all elements we have seen before have an equivalent. A multiplexer operating in the optical domain. Interconnected by using optical cross connect. What is that? A sort of router done in the optical domain. This is the framework, indeed this is the physical structure where I support my logical structure that is by interconnecting routers, when I setup a routing path I setup at the routing layer, layer 3, this must be reflected. the mechanism is then mapped to a specific optical path, setup in this domain. Indeed someone says the network here is an analog one, it's made by all optical signals. The signals are implemented by using an important modulation scheme. **Wavelength division multiplexing**, what is? The multiplexing on a single fiber of different wavelengths. Meaning that, we can multiplex together in the same fiber multiple flows, in this case are multiple wavelengths. In a window where you can transmit lights, performance behaviour, in that window you can provide multiple wavelengths and these are multiplexed all together.

These devices, the transmitting and receiving one, are able to recognise different wavelengths, each of one supporting a multiplexer digital flow. Sometimes you can switch a digital form to a wavelength to another. If you have this picture here, a path at this level is named **light path**. The network is a combination of different light paths. In this light path there is a light path from Rome to Milan for instance, if a light path is enough, maybe it's not the case. This is switched and regenerated during its path. The switching may involve here the changing of the wavelength transmitting. This is the optical switching, **changing the wavelength in order to switch.** A path. This can be done in the optical domain or convert back the path to the electrical domain, perform a conversion, switch and then convert again. Every time you pass through this you regenerate the signal. This can be done by using a wavelength division multiplexing. In the same window you can put a high number of wavelengths.

So, what is important to say is this, all the concepts designed in the SDH are mapped in the optical domain. There is the concept of optical multiplexer, transmission section, optical channel and the optical channel is a container derived in the optical domain to accommodate the client information. The client information is in the electrical level. I form my digital structure and then my digital structure is fully mapped in the optical domain. Last is this, switching I mentioned is the **switching of light paths, of wavelengths.** Implemented by using optical devices and electrical ones. The switching may happen in this way, I receive in the optical domain, transform my domain in electrical, I switch the information, according to what is defined at the logical level, by routing the information where planning the routing, and then I convert in the optical domain. This is what works in our digital networks. There are continuously these converts. It was proposed to have **all this in the optical domain**.

Which is the reason for having all this in the optical domain? Which one could be the reason? Time. It's a matter of delay, these lines arrive at a very fast speed, this switching and conversion may take time, so the **bottle neck of a network** can become this one. This was true till some years ago, now the technology is also mature to have this one very very fast, but in the middle there was the proposal to provide a switching **fully operating** in the optical domain, I indicated you some weeks ago how an optical switch was supposed to work, the switching, cross connection of an input to an output was for instance planned by working this is a **fully optical switch**, the switching point was made by using very small mirrors, moved every time you have to modify the switching pattern, in order to reflect the light from an input to an output, this was the idea for an optical switch. Fully optical, so a matrix of mirrors in the order of mm, designed to be switched to interconnect an input to an output.

But however the control of these mirrors, how to move them, was done in the micro electronics domain, and maybe that part also is critical, so you have to move these mirrors, however the good opportunity was done by having only a reflection of the light, so **light always remains light**, is not converted back to the electrical. This was an idea, however as I said you today and for a while these solutions are the one adopted. On top of this, this is the structure of our layer 2 infrastructure, for providing what? A layer 3 plan that typically is done by using IP, on top of solutions that are derived these solutions are based on the protocol to **multiplex and demultiplex and route** traffic at the IP level, mapped on this digital optical platform. In this digital optical platform what you design and program is a path. A path that is light path, sometimes comes back to a path and then becomes again a light path, this is our infrastructure in the core network, exchange flows of light, on top of that flows there is a path designed at a higher level thanks to the different protocols like for instance also the SDN, I don't know if there are students to present something on SDN. Now maybe you will see how the SDN can be adopted also in this framework, for instance today there is the trend to use to provide the configuration of the **light paths at the SDH/optical** transport network context. I think I can stop here and we can start with the presentations.