Medium Access Control (MAC) protocol

The Internet consists of a large number of interconnected computers. Most of those computers appear on local area networks (LAN). A LAN interconnects a small, say, 20, computers owned by the same organization. A LAN is generally spread over a small geographic area, such as a university campus, a corporate office, and so on. If an organization operates hundreds of computers, then those are organized into tens of LANs. The mode of interconnection of the computers can be wire or wireless. To a large extent, the Ethernet technology provides a wired medium to form a LAN. If the mode of interconnection is wireless, say, provided by the IEEE 802.11 technology, then we say that we have a wireless LAN (WLAN). (At this point we do not get into the different kinds of IEEE 802.11 protocols, such as 802.11a/b/g.)

Computers within one LAN or WLAN can exchange data packets without the involvement of any other communication entity. However, for computers on two different LANs to be able to communicate, there is a need to interconnect the two LANs.
using a communication entity called a router. Thus, each and every LAN or WLAN is connected to a router. An organization can own several LANs and WANs and hence several routers as shown in Figure 1.

The routers owned by an organization are interconnected so that the different LANs can work together. Finally, a small number, say one or two, of routers owned by an organization are connected with external routers for Internet access. In this manner, all the computers of the world can be, or, in fact, have been, interconnected. This situation has been depicted in smaller scale in Figure 1.

Medium Access Control (MAC) protocol

Computers and routers are interconnected in a wired or wireless manner. In an Ethernet LAN, computers and the router are interconnected by a shared medium—the Ethernet. Thus, all the computers and the router need to compete to be able to transmit a data packet between two computers or between a computer and the router. (On the other hand, two routers connected by a dedicated medium, say, a fiber link, do not compete to transmit data, because the medium is dedicated to the pair of routers.)

A shared medium is also called a broadcast channel, multi-access channel, or random access channel.

Problem with a shared channel: Assuming that computers are independent entities, how to determine who gets to use the medium (channel) next when they compete to access the channel? (If there is a dedicated channel between two computers, the computers can transmit data at their will.)

The goal of designing a MAC protocol is to resolve the above issue. Before discussing the design issues of a MAC protocol, it is important to understand FIVE key assumptions as described below.

1. Station model: The computers are independent. Once a computer generates a packet for transmission, subsequent packets are queued up for transmission.
2. Single Channel model: A single channel is available for communication. All stations can transmit on it and all can receive from it. Thus, from the hardware point of view all stations are equivalent.
3. Collision assumption: If two packets are transmitted simultaneously, the resulting signal is garbled (if the two signals can reach the same point.) This event is called a collision. In a wired LAN we assume that computers can detect collisions. In a WLAN, it is not possible for a transmitter to detect collision at a receiver.
4. Time model:
   a. Continuous: Packet transmission can begin at any instant.
   b. Slotted: Packets are transmitted at well-known instants identified by slots.
5. Carrier sense:
a. Carrier sense: Each computer monitors the channel for carrier to know if the channel is in use.
b. No carrier sense: Computers do not sense the channel before trying to use it.

We will study a variety of MAC protocols in the class including (pure) Aloha, slotted Aloha, CSMA/CD, and CSMA/CA.

**Aloha protocol**

In the 1970s, the Aloha protocols were developed at the University of Hawaii to interconnect computer terminals with the main computers. Each terminal was equipped with a radio interface as were the main computers. Thus, all the terminals belonging to a main computer were connected to the computer via a shared medium—the air interface.

There are types of Aloha protocols: pure (unslotted) Aloha and slotted Aloha. In the pure Aloha protocol, terminals can transmit at any time instant, whereas in slotted Aloha, terminals can transmit at the beginning of well defined time slots.

**Pure Aloha protocol**
- A terminal can transmit a packet at any instant.
- If the terminal does not receive an ACK before a timeout, it runs a backoff algorithm. At the end of the backoff period, retransmit the packet.
- If the terminal receives an ACK before a timeout, it is ready to transmit the next packet.

*Note*
- If the total traffic is low, the protocol works well.
- If the total traffic is high, the protocol performance is poor. In fact, it is 18%. (Throughput: Total packet arrival rate * Probability of success of a packet.)
- Packet transmission continues even when there is collision. \( \Rightarrow \) Loss of bandwidth.
- The protocol is very simple. It is not used for forming computer networks, but is used by cell phones in the GSM system to communicate with base stations of their presence when you switch on a cell phone.

**Slotted Aloha**

This protocol is similar to the pure Aloha protocol except for the fact that time is assumed to be slotted. Terminals transmit packets at the beginning of slots, and not whenever they want to. This improved the throughput of the protocol to 37%.
CSMA/CD Protocol

To improve the performance of the Aloha protocols, it is important to abort all transmissions as soon as a collision is detected. This is because once two packets start colliding, it is useless to continue with their transmissions as both the packets will be received by their respective receivers with bit errors in them. Such continued transmissions after a collision will lead to a waste of bandwidth. Thus, there is a need to detect collision while transmitting a packet.

It is important for a sender to detect collision at a receiver, because it is collision at the receiver that results in the loss of a packet. However, it is not possible for a sender to detect collision at a receiver, because senders and receivers are located far apart and there is no other protocol more basic than a MAC protocol so that a receiver can inform the sender of a collision. Thus, a sender makes an assumption and uses some other means to detect a collision at a receiver: (i) We assume that signal from any computer can reach all other computers. (ii) Collision is detected a little while later than its occurrence.

Figure 2: Non-Persistent CSMA/CD
In case of the Ethernet technology, where computers are wired to form a LAN, a sender detects collision by switching on its receiver while transmitting—if what it receives is different from what it is transmitting, a collision is said to have occurred. It is assumed that the difference between what a computer transmits and what it receives is caused due to another computer transmitting a packet at the same time.

A MAC protocol with carrier-sensing and collision detection is given in Figure 2. The protocol waits for a while before accessing the medium if the medium is found to be busy, and this is done multiple times for the same packet. Hence, the protocol is known as Carrier Sense Multiple Access with Collision Detection (CSMA/CD). As the name of the protocol suggests, the CSMA/CD protocol senses the medium to find out whether or not it is free before transmitting a packet. If the medium is found to be busy, the computer waits for a small random period before sensing the medium once again – this is known as “non-persistent” carrier sensing.

If carrier sensing indicates that the medium is free, the computer initiates packet transmission and starts detecting collision at the same time. A computer achieves this by switching on its receiver while transmitting a packet. Collision detection is done by the physical (PHY) layer. If no collision is detected while a packet is being transmitted, the packet transmission is said to be successful.

If there is a collision, the computer aborts the transmission process so as not to waste bandwidth, and transmits a jamming signal so that all other computers are aware of the collision. Next, the computer runs a backoff procedure to hold itself back for a while from carrier sensing. The duration of backoff is random, and it falls in the range $T_p[0, 2^i-1]$, where $T_p$ is propagation time and $i$ is a variable denoting how many times the protocol has gone around the backoff loop before. The reader may note that the range of the random number increases exponentially as the medium is found busy in successive iterations of the loop.

At this point, it is interesting to compute the time a computer may take to detect collision after starting a transmission. Assume that there are two computers $A$ and $B$ separated by the maximum possible distance $d_{max}$. Let $A$ sense the medium to be free at time instant $t_0$, and start transmitting immediately. It takes a while for $A$’s signal to reach $B$. Assume that $B$ detects the medium to be free right before $A$’s signal reaches $B$, and starts transmitting its packet immediately. $B$ will detect a collision almost immediately. However, for $A$ to detect collision, $B$’s signal must reach $A$. Let $t_{prop}$ be the time taken by signal to travel between the distance $d_{max}$ between $A$ and $B$. Thus, for $A$ to detect a collision in the above case, it will take $2*t_{prop}$ time units. This is because, it will take $t_{prop}$ time unit for $A$’s signal to reach $B$ during which $B$ sees the medium to be free and another $t_{prop}$ time unit for $B$’s signal to travel to $A$. Thus, $A$ can continue transmitting for $2*t_{prop}$ time units before it can detect a collision.
For a certain value of $d_{max}$, for example, 100 meters, and given that signal travels at $2/3$ the speed of light in free-space (the speed of light in free-space is $3 \times 10^8$ meters/second) on the physical medium, one can easily calculate the value of $2t_{prop}$.

**Physical interconnection of computers using Ethernet technology**

In the 1970s, wired LANs were constructed using co-axial cables as shown in Figure 3. Taps at different points on a cable were used to access the medium. To be able to connect a tap, it was necessary to drive a hole in the cable or cut it at that point to form a T-junction. One computer is connected to the cable at one T-junction, and several T-junctions had to be formed to connect more computers. Cable installation and adding more T-junctions were the central problems in managing a co-axial cable based LAN. Also, locating cable fault was a major concern.

**Figure 3: Co-axial cable interconnection in Ethernet technology**

Thus, a new technology was invented to eliminate the need for co-axial cables in wired LANs, as shown in Figure 4. Here, the physical medium consists of an electrical box and a set of twisted pairs of cables connected to the electrical box. The task of the box is to distribute the signal coming in from one pair to all the pairs. It is much easier to install twisted pairs. Thus, a LAN constructed this way is cheaper to construct and maintain.

**Figure 4: Twisted pair interconnection in Ethernet technology**