CPS221 Lecture: The Physical and Data Link Layers

last revised 9/15/14

Objectives

1. To discuss various options for the lowest layers

Materials:

1. Projectable of Gordon's microwave tower
2. Projectable of Ethernet configuration using a coaxial cable
3. Projectable of Ethernet using a hub (KR 5.21)
4. Projectable of Ethernet frame format (KR 5.22)
5. Projectable of 802.11 subnet (KR 6.14)
6. Projectable of Hidden Terminal Problem (KR 6.11)
7. Projectable of IEEE 802.11 standards

I. Introduction

A. The lowest layers of the protocol stack deal with physical communication between two connected nodes.

1. In the OSI model, these are conceptualized as two distinct but tightly coupled layers: one being the lowest software layer, and the other the physical hardware layer. (They are tightly-coupled because the software is hardware-dependent, of course).

2. Often, though, the are thought of as a single layer, which may be called the Link Layer of the Network Access Layer.

3. This layer deals only with communication between two connected nodes, while higher layers (network layer on up) address communication between endpoint hosts.

4. This layer, together with the next layer up (the network layer) are found on all nodes, though only end-point hosts need the higher layers.

B. There are, in fact, many different ways of implementing this functionality, and most messages will actually pass through several different kinds of implementation in traveling between end-points.

   Example: if you request a web page from an off-campus server using a computer at Gordon, your request will first probably first travel via ethernet or WiFi to Gordon's internal backbone, and will then travel over optical fiber to our connection to our ISP, which is implemented by a microwave tower at 266 Grapevine.

   PROJECT: Gordon's microwave tower

   (How it travels the rest of the way will depend on what technologies our ISP and the server you are accessing use.)

C. The unit of information transfer at this level is commonly called a **frame**. The format of a frame (and, indeed, even its maximum size) varies depending on the technology in use.

D. In endpoint hosts, link-layer technologies are typically implemented by a network interface device - historically a separate card (NIC), but now often implemented on the motherboard.

   1. That is, most if not all of the functionality of this layer is implemented by hardware.

   2. It is not uncommon for a computer to have two or more separate network interfaces - e.g. the laptop I use for video projection has several interfaces as we shall see shortly.

3. Each link-layer interface is identified by a unique identifier, commonly called a Media Access Control (MAC) address.

   a) For most technologies, the MAC address is a 48 bit binary number, which is typically written a six pairs of hexadecimal digits.

      SHOW MAC addresses for interfaces on laptop (About This Mac | More Info | System Report | Networks | Location)

   b) Every interface device has a unique MAC address that is assigned to it during manufacture (typically by "burning" it into a chip)

      Manufacturers can purchase a range of 2^{24} MAC addresses from the IEEE for a nominal fee. The manufacturer then gives a unique address from this range to each device it manufactures.

   c) Note that the MAC address for the interface on a given computer doesn't change, even if the computer is moved to a different location. (In this respect, MAC addresses are more like individual Social Security numbers rather than postal addresses)

E. Link-layer technologies fall into two broad categories:

   1. Point to point connections, which directly connect a pair of nodes.

      Examples:

      a) The connections in the Internet core are all of this type.

      b) Some of the mechanisms by which an individual user might be connected to an ISP, e.g. Dial-up modem or DSL.

2. Broadcast or multiple access technologies.

   a) A broadcast technology services multiple nodes that share a common channel.

   b) Of course, only one pair of nodes is actually communicating at any given moment of time. However, all communication is "heard" by all the nodes - but is normally ignored by nodes other than the two that are communicating (except in the case of broadcast messages which are meant for all nodes).

   c) Examples

      (1) Ethernet

      (2) Various forms of wireless (e.g. WiFi, cell-phone 3G/4G

   d) Multiple-access protocols are similar to the protocols humans used when in a group - e.g. "don't interrupt when someone is speaking" or "don't monopolize the conversation".

   e) In this lecture, we will focus on link layer technologies that are commonly used for local-area networks. (Recall that, in the Internet was conceived as a mechanism for connecting local networks - an inter-net.) We will not talk about the technologies used to implement the Internet backbone. or other point-to-point technologies such as dial up or DSL

II. **Ethernet**

A. Ethernet has historically been the most widely-used technology for connecting computers in a local area network - and that is still the case, though wireless use is growing.

1. In its original form, Ethernet envisioned computers being connected by a common physical medium - a coaxial cable.

   PROJECT: Ethernet using coax

2. Today, however, it is more common to use a star configuration using hubs (or, as we shall see later) switches.

   PROJECT: Ethernet using a hub

3. In either case, a single medium is shared by all the nodes on a subnet.

B. An ethernet frame has the following format:

   PROJECT Ethernet frame format

1. The Preamble is a series of 7 bytes of 10101010 followed by a byte of 10101011. (This serves to synchronize the receiver with the sender clock and to mark frame boundaries) (8 bytes in all) It is not formally considered part of the frame.

2. The destination and source addresses are 6 byte MAC addresses

3. The next field is a 2 byte field specifying either the protocol at the next level to which the data should be passed, or the length of the data portion of the frame.

4. The data field is the actual data.

   a) The shortest permissible length is 46 bytes (which makes the overall frame not including the preamble 64 bytes). If the data to be sent is shorter than this, it is padded ("stuffed") to this length.

   b) The longest permissible length is 1500 bytes. If the data to be sent is longer than this, it is fragmented into two or more frames.

5. The CRC field is a cyclic-redundancy check field appended to error detection purposes. (If this check shows that the frame has been corrupted in transit, it is simply discarded. A higher level protocol may expect an acknowledgment for this frame and, receiving none, may choose to send it again.

C. Given that Ethernet is a multi-access technology, how does it deal with the possibility that two nodes might try to send data at the same time?

The answer is a strategy known as Carrier-Sense Multiple Access with Collision Detection (CSMA-CD)

1. When a node wishes to send data, it first "listens" to be sure that no other node is transmitting, and waits until it is clear to send. (Actually, it waits for the channel to be clear for 96 bit times before starting to send.)

   (This is analogous to the rule "don't start talking when someone else is talking")

2. Of course this doesn't guarantee there won't be a collision if two nodes are waiting for the medium to become clear before starting their own send.

Observe: if we are using 100 MBit Ethernet, then the time for a single bit is 10^{-8} seconds. If signals travel at the speed of light, then a signal will travel about 3 m in this time. If two nodes are 30 m apart, then one of the two will be able to send 10 bits before the other node receives the first.

3. This brings us to the other part of the strategy: collision detect. As a node is transmitting, it also reads the signal on the medium and compares it to what it's sending. If the transmitted and received signals differ, then the node knows that the data it has been sending has collided with what another node has been sending. At this point it "backs off" for a random period of time before trying to send again.

D. Hubs and switches

1. We saw earlier that the typical ethernet configuration today is a star, with all nodes connecting to a common hub.

   PROJECT AGAIN: Hub connection

   When a hub receives an incoming signal on any line, it sends a copy out over all the other lines. Thus, each node "hears" everything that any node transmits - though it ignores anything not addressed to it.

2. An alternative is to use a switch instead of a hub. A switch differs from a hub in that it stores a complete frame that comes in over one of its lines and then repeats it over another line only if the destination MAC address is the same as that of the device connected to that line - or the frame is a broadcast frame (destination `ff:ff:ff:ff:ff:ff`) - or the MAC address of the device connected to the line is not yet known.

   a) A switch is more complex than a hub, in that it

(1) Must keep track of the MAC address of each node connected to it (which it learns from the source address field of a frame sent over the link).

(2) Must be able to store a complete incoming frame to be forwarded over one or more links (possibly more than one at the same time)

(3) Selectively forwards frames only to relevant links.

b) A switch eliminates the possibility of two nodes colliding with each other, since complete frames are stored and then forwarded only when the outgoing link is clear.

c) A switch can support multiple transmissions between nodes at the same time - e.g. if a switch has 4 ports (A, B, C, D), then A can be allowed to communicate with B at the same time C communicates with D.

d) The fastest ethernet standards require the use of a switch instead of a hub.

E. Actually, Ethernet is a family of standards, covering various transmission media and covering a broad range of speeds.

1. Ethernet uses serial transmission - that is, individual bits are sent one right after another.

a) The original ethernet standard transmitted at 10 Mbits / second.

b) Fast ethernet transmits at 100 Mbits/second.

c) Gigabit ethernet transmits at 1 Gbit/second

d) Ten-Gigabit ethernet transmits at 10 Gbit/second

e) Even higher speed versions are emerging.

f) All speeds use the same frame format. All except 10 Gbit and above use CSMA/CD - the fastest standards requires dedicated connections through a switch to prevent collisions from occurring.

g) Higher speed interface devices are typically downward compatible - e.g. an 1 Gbit interface will shift down to either a 100Mbit or 10Mbit rate if connected to a slower interface.

2. Ethernet can be implemented over a variety of types of media

a) Thick and thin coaxial cable (rarely used today)

b) Twisted pair (CAT4, CAT5, CAT6)

c) Optical Fiber

III. **Wireless Link Layer Strategies**

A. There are a number of types of wireless link layer strategies, including Bluetooth and various cellphone strategies such as 3G and 4G.

B. However, the best known and most widely used is IEEE 802.11, also known as WiFi. This is the one we will discuss.

1. Because of its similarities with ethernet, it is sometimes called wireless ethernet, though it is really a distinct standard.

2. 802.11 makes use of access points which are connected via ethernet to a router.

PROJECT 802.11 subnet

C. To use a wireless network, a computer must connect to it.

1. Many wireless networks require a password in order to be able to connect.

2. Once connected to a wireless network, a computer can access it through any of its access points.

3. Many wireless networks also utilize some form of encryption to protect the content of transmissions - since otherwise anyone could "listen in".

D. Whereas ethernet uses CSMA/CD to control sharing of the medium, 802.11 uses CSMA/CA - Carrier Sense Multiple Access with Collision Avoidance.

1. One reason for this is that collision detection requires that a node be able to both transmit and listen at the same time. But with wireless, the signal being transmitted will be much stronger at the transmitting node than any signal received from another node, making collision detection difficult.

2. Another reason is the Hidden Terminal problem.

   PROJECT

   a) The problem arises because transmitted signals have a finite range - especially in the presence of obstacles. Thus, it is quite possible that two stations might be out of range with each other, while both still being in range of a station lying between them.

   b) Now suppose one of the outer stations is transmitting a frame to the station in the middle, and other outer station also wishes to transmit to this station. Since it cannot “hear” the transmission from the station that is sending, it will think the

channel is clear and begin sending. Now neither station is aware of the collision, though the recipient station certainly is!

3. Because collision avoidance is being used rather than collision detection, once a node begins transmitting it transmits its entire frame without stopping for a collision.

4. Two strategies are used to avoid collisions

   a) One is a node doesn't begin transmitting immediately when it senses that the channel is clear. Rather, it waits a random time and then begins transmitting. This way, if two nodes are both waiting for a clear channel, one will begin before the other and the other node will now sense that the channel is not clear.

   b) To deal with the hidden terminal problem, an optional feature allows a node to reserve a channel between two nodes for a period of time, using a handshaking protocol that causes other nodes in signal range of either node to avoid using the channel during the reserved time.

5. Of course, even with mechanisms like this, collisions do still occur. Therefore, 802.11 nodes acknowledge the successful receipt of a frame back to the sender. If a sender does not receive an acknowledgment for a frame it has sent, a collision may have occurred, so it sends the frame again.

E. As was the case with ethernet, 802.11 is not a single standard, but a growing collection of standards.

PROJECT IEEE 802.11 standards

Although a computer typically only implements one of these standards, an access point typically supports more than one so that different devices can connect.