TCOM 370

NOTES 99-13

Ethernet LANs and Related Topics

1. IEEE 802.3, Ethernet, or CSMA-CD

This MAC protocol is commonly used for a LAN bus configuration where the medium consists of a coaxial cable or twisted-wire pair. It is based on the ALOHA protocol but achieves higher efficiency than ALOHA by implementing two extra features.

- Before a node starts to transmit a frame, it "listens" to the channel to determine if it is already carrying a transmission. If so, the node waits until the channel is clear before launching its frame on the medium. (This is *"Carrier Sense"* in CSMA-CD, referring to the sensing of a characteristic electrical activity on the bus when it is carrying a transmission from a node.)

- While a node is transmitting its frame, it ***monitors the medium*** in the initial part of its transmission to ***detect*** any overlapping transmission that may have been initiated by another. (This may happen because the transmission of the first node did not reach the other node before it started transmitting, due to propagation delay). If a "collision" or overlapping transmission from another node is detected by any node, it stops its transmission and backs off for a ***random time duration*** before attempting frame transmission again. This is *"collision detection"*, CD, in CSMA-CD

  (The "collision" is detectable because when more than one node has a signal on the bus, the electrical characteristic sensed on the bus is different from that sensed when only one signal is present).

When a collision is detected, in practice a node detecting the collision cuts off its transmission and transmits a short jam signal. (This allows the other node to detect a collision, which it could otherwise miss if the second node stayed on for a rather short time interval). The other node upon detecting the jam signal also aborts its transmission with a short jam signal. Thus in case of a "collision", the total length of the frame on the bus is limited in duration. Such a short frame and the jam signal when received by the other nodes also informs them that a collision occurred. The IEEE standard specifies a **minimum valid frame length of 512 bits** and a maximum length of **12144 bits** (excluding preamble). A **collision** must be **detected** in the time to transmit **450 bits. This limits the maximum span between nodes,** because it puts an upper limit on the maximum node-to-node propagation delay. The **jam** signal is a random bit pattern with a maximum of **48 bits**. Thus a collision is always detected and the frame fragment before frame-abort with the jam signal is under 500 bits in length.

Therefore in the event of collision, it can be detected within the fairly short span of 500 bits, so that if long data frames are being transmitted the result of occasional collisions will not amount to major loss of utilization of the bus. Of course under heavy load from many nodes, collisions will occur relatively frequently and utilization efficiency will be lower.

The **random back-off period** after a collision has been detected is determined by the following algorithm:

- If a packet has collided k successive times, where k<16, a node chooses a random number J equally likely to be any one from the set $\{0, 1, 2, \ldots, 2^q - 1\}$ where $q=\min\{10,k\}$. The node then waits for Jx512 bit times before attempting to transmit. Thus after one collision, the node chooses a delay of 0 or 512 bits, and after 2 successive collisions it chooses a delay of either 0, 512, 1024, or 1536 bit periods, etc. This is called the **binary exponential back-off algorithm**.

By doubling the range of the random delay after each collision for a frame, the probability becomes very high that eventually one node will pick a shorter delay and be able to transmit successfully. Note that if nodes decide (at random) to wait for different multiples of 512 bits, then since the collision window is less than 500 bits, the node choosing the smallest delay will transmit successfully. (After 16 successive collisions the attempt to transmit is aborted.)

2. Efficiency of CSMA-CD

To compute the efficiency of this protocol, we will assume an ideal slotted scheme to simplify the analysis. The **slot duration** is now defined to be **twice the maximum propagation delay, $2T_p$, between two nodes** on the bus. This is because after one node has initiated transmission on the channel, a second node located furthest away from the first may begin transmission on what it perceives to be a clear channel up to a time of one propagation delay after the first node's start of transmission. It takes a further one propagation delay interval before the first node can become aware of this collision and abort transmission. ***Note that a typical frame length is many slots wide.***

Once a node has "grabbed" a slot successfully, meaning that no collision is detected by the end of that slot, it will continue to transmit its whole frame of total duration T_{ix} successfully. Thus a total of $\mathbf{\frac{T_{ix}}{2T_p}}$ **or** $\mathbf{\frac{1}{2a}}$ **slots** will be used, where $\mathbf{a = \frac{T_p}{T_{ix}}}$ as in our previous notation.

- Suppose **M = 2** nodes compete for access on the medium.
- Suppose that for each node, there is a **probability p** that it wants to transmit a frame in a given slot, independently of other nodes.

Then the **probability a that exactly one node of the M will have a frame for transmission,** and hence be able to grab a slot and transmit a frame successfully, is

$$\alpha = \binom{M}{1} p^1 (1-p)^{M-1} = M\,p\,(1-p)^{M-1}.$$

This probability is a maximum when

$$\frac{d\alpha}{dp} = M(1-p)^{M-1} - Mp(M-1)(1-p)^{M-2} = 0$$

or when $\mathbf{p=\frac{1}{M}}$. For this p, the **maximum value** α_{max} of α is

$$\alpha_{max} = (1-\frac{1}{M})^{M-1} = (1-\frac{1}{M})^{M} \frac{M}{M-1}$$

For **large M** this is $\lim_{M\to\infty} \alpha_{max} = e^{-1} =$ **0.36.**

(We can take approximately α_{max}=0.4 for finite M. Check that for M=4 and up, this is a reasonable approximation)

When the medium is free again, the nodes will once again contend for a slot and there will be several collisions or empty slots before one node grabs a slot successfully. We now need the **probability there will be k such wasted slots before a successful grab** takes place. Now

P{k wasted slots before a "grab"}

= P{k times in a row **0 or more than 1** nodes want to transmit, followed by slot when only 1 wants to transmit}

$$= (1-\alpha)^k \alpha$$

The expected number of wasted slots is then

$$\sum_{k=1} k\,(1-\alpha)^k \alpha = \frac{1-\alpha}{\alpha}$$

which is a **minimum** of $\frac{1-e^{-1}}{e^{-1}} = 1.72$

So we have on the average **at least 1.72 wasted slots** before a successful transmission utilizing $\frac{1}{2a}$ lots, giving a **best-case theoretical efficiency** of

$$\mathbf{h}_{csma\text{-}cd} = \frac{(1/2a)}{1.72+(1/2a)}$$

i.e.

$$\boxed{\mathbf{h}_{csma\text{-}cd} = \frac{1}{1+3.44a}}$$

In practice the effective value of α is less than e^{-1}, and a more realistic expression for the efficiency is

$$\mathbf{h}_{\text{csma-cd}} = \frac{\mathbf{1}}{\mathbf{1+5a}}$$

For small "a", say a=0.1 (propagation delay smaller than frame duration) this gives an efficiency of 2/3, but for a=0.8 the efficiency is only 0.2. It is clear form the above that the efficiency of CSMA-CD or Ethernet depends critically on the ratio "a". This is because this protocol attempts to get higher efficiency than ALOHA by detecting collisions as soon as possible and aborting frame transfer upon collision, so that further time on the medium is not wasted. This will work well if the frame length is long compared to the time it takes to detect a collision. The time it takes to detect a collision depends on the maximum propagation delay, and therefore on the maximum span of the network. ***The maximum length limitation in Ethernet therefore also controls the magnitude of the ratio a.***

Connection with Slotted ALOHA:

Note that for a=1/2 we get approximately the slotted ALOHA efficiency. Actually, if we use the theoretical efficiency $\frac{1}{1+3.44a}$, we get $\mathbf{h}_{\text{csma-cd}}$ =0.3679 for a=1/2, exactly the slotted ALOHA efficiency. The reason should be clear. The slotted ALOHA scheme corresponds here to the case where the slot duration is the same as the frame duration. The number of slots $\frac{1}{2a}$ =1 means a=1/2).

3. Details of 802.3, 10 Mbps LAN Systems (CSMA-CD, "Ethernet")

In this section we note some practical details of baseband systems operating at the standard **10 Mbps** rate on a common bus. (Refer to figures at end).

10Base5

An ethernet system known as **"10Base5"** ("thicknet") uses a coaxial cable of diameter 10 mm as the common bus to which stations or nodes are attached. The attachment of a node (or DTE, "data circuit terminating equipment" such as a PC or a network printer) to the cable is via *a network interface controller* (NIC) which resides within the DTE . The NIC handles the digital or bit-level functions (such as checking destination and source addresses, handling FCS, assembling and receiving frames). The NIC is connected via an AUI *(auxiliary unit interface)* cable to the *analog transceiver* that is physically attached to the ethernet cable. The line coding scheme used is *Manchester coding*. Each cable segment is limited to 500 m length with a maximum of 100 nodes on it. Up to 4 repeaters may be used between 5 segments, so that the maximum distance between nodes is 2.5 Km.

10Base2

The **"10Base2"** system uses thinner cable (5 mm) and is less costly and flexible. Here the transceiver function is built into the NIC, and the ethernet cable is attached directly to the NIC (at the back of the PC, say) with a BNC connector. Thus the PC with the NIC in it attaches directly to the ethernet cable that runs by each station. The maximum segment length is now approximately 200 m (hence the "2" in 10Base2). There is a limitation of 30 nodes per segment. Repeaters may again be used, but the maximum distance between nodes is limited to be under 1 Km. (The thinner cable has higher attenutation).

10BaseT

In a **10BaseT** system, the *physical arrangement* is different but *the logical function* is the same as in the other two forms above. There is no common physical bus, instead an electronic "hub" unit with multiple connection ports on it serves as the "common electronic bus" to which individual nodes or stations connect their NICs via two twisted-pair wires (receive data and transmit data). A small hub may have 32 ports; larger hubs and hierarchies of hubs in a tree structure can serve many more nodes. The electronic hub basically acts as a repeater.

Repeaters and Bridges

To extend a span of coaxial cable on which a number of nodes are already connected, another span may be used with additional nodes attaching to it. Between cable spans a **repeater** is used; this is a ***physical layer*** device and simply acts as an amplifier in both directions, taking any transmission it sees incoming at any of its two ports and repeating it as an outgoing transmission on the other port. Beyond a maximum length, such extensions are not permissible. A **bridge** is a device that operates at a higher layer in the protocol stack. It can interconnect two LANs, but is seen by each as a simple "node". Any frame with an address belonging to a destination node outside the LAN from which it originates will be handled by the bridge and transmitted on the other LAN to which it is connected.

4. Token Ring Protocol

A ***token*** is a specific bit pattern that is passed from node to node in sequence on a unidirectional ring (it can also be used on a bus). When a node gets a token and it has a packet to transmit, it modifies the token into another pattern ("start of frame indication") to which it appends its packet. The start-of-frame and the token may differ only in the last bit, for example, so that a node can convert a token very efficiently with only a one-bit delay. After transmission the node releases the token to the next node. A node that is not transmitting repeats any packet it receives to the next node in sequence.

A node that has seized a token for its transmission releases the medium by inserting a new token on the ring when (a) it has completed transmission of its frame, and (b) after the frame header has returned to it around the ring. This is the ***Release after Reception (RAR)*** protocol. The RAR protocol is specified in IEEE 802.5. Here the transmitting node does not release the token until the complete frame has returned to it. This allows acknowledgments to be appended to the packet by the receiving node.

Release after Transmission or RAT protocol.
In the RAT protocol, a node releases the medium by issuing a token as soon as it has finished transmitting, without waiting to receive its own frame back. It has a higher efficiency if a>1. To calculate the efficiency of the RAT protocol with M nodes, note that in $M(T_{ix}+T_{it})+(T_{p1}+T_{p2}+ .. T_{pM}) = M(T_{ix}+T_{it})+T_p$ secs each node has transmitted one frame, where T_{it} is the token transmission time and T_{pj} is the propagation time between the j-th and the next node, and we are ignoring token modification time. Thus the efficiency is

$$\mathbf{h}_{\mathbf{RAT}} = \frac{MT_{ix}}{M(T_{ix}+T_{it})+T_p} = \frac{1}{1+\frac{a}{M}}$$

for relatively short token sequence. Token ring with the RAT protocol can be very efficient. Transmission media include twisted pair, coaxial cable and optical fiber.

5. Frame Structure for LLC/MAC

The DLC layer (layer 2) consists of a lower MAC layer and an upper LLC (logical link control) layer. The LLC layer is responsible for transmitting link-layer packets consisting of information bits for higher-layer protocols as well as the destination and source LLC users. These users are generally higher-layer protocols, and are addressed within the LLC layer as Service Access Points (SAPs). Thus a destination SAP may be a specific logical channel on a host computer over which an application communicates with the LLC. Each SAP is an address that is unique *within* a node. The LLC frame also contains a control field that specifies the type of frame. IEEE 802.2 specifies different services that may be supplied by the LLC. These are **connectionless (best-try, unacknowledged), connection-oriented,** and an intermediate service called **acknowledged-connectionless.**

Connectionless service has the least overhead associated with it. It does not guarantee that a packet will be delivered or that it will be delivered without errors.

Connection-oriented service requires a logical connection between a Source SAP and a Destination SAP to be set up and taken down. Once formed, the connection is used to pass control, information, and acknowledgments and provide flow control. The connection-oriented service provided by IEEE 802.2 LLC uses **HDLC-ABM** (high-level data link control, asynchronous balanced mode).

Acknowledged-connectionless service allows connectionless packet exchange with low overhead, but provides for acknowledgments to be returned.