CS2105 Live Class Lec8&9: Link Layer

$_/|/|U_Ch@NgRu!$

May 5, 2021

1 Function of Link Layer

Link Layer is responsible of sending datagrams between **adjacent nodes**(Not end-to-end) over a **single link**

- IP data grams are encapsulated in link-layer frames for transmission
- Different link-layer protocols may be used on different links; each protocol may provide a different set of services

2 Service that link layer provides

Link layer may offer the following services

- Framing:
  encapsulate IP datagram into link-layer frame, adding header and trailer (different protocols have different header and trailer format; each header and trailer is valid for only one hop)
- Link access control:
  Coordinate which nodes can send frames at a certain point of time, when multiple nodes share a single link
- Reliable delivery: seldom used on low-bit error link such as fibre; often used on error-prone links such as wireless link
- Error detection
  - Errors are usually caused by signal attenuation or noise
  - Receiver detectes presence of errors. Receiver, when error is detected , may signal sender for retransmission or simply drop frame

Yu Changrui

  - The error or collision happens during the propagation, if I did not understand wrong, if collision happens, there will be a physical surge, which may be notified by the receiver or nearby hosts, and they may send a jamming signal to inform the error; This is not introduced in detail in the lecture, it's based on what I search, better search yourself

- Error correction:
  Receiver may identify and correct bit error without resorting to re-transmission

Link layer is implemented in adapters/network interface controller(NIC), which contains both link and physical layer(on chip)
adapter is semi-autonomous, which means it acts independently to some degree

3 Error detection and correction

Commonly used error detection schemes include

- Checksum(used in TCP/UDP/IP)
- Parity Checking
- CRC (commonly used in link layer)

Note that error detection scheme are not 100% reliable(see hash collision and birthday problem)
The fact is that the CRC is much more accurate and can detect more error comparing with parity check and checksum, but it is only implemented in link layer(both IP, UDP, TCP applies checksum instead), the reason si that the CRC requires much more calculation and is slow in software, but in link layer, there are hardware to support, so it's a trade-off

3.1 Parity Checking

3.1.1 Single Bit Parity

Suppose the data is of the form of d_n ...d_1, the parity bit p is chosen such that
$\sum_{i=1}^{n} d_i + p \equiv 0 mod 2$
The sender, after computing the parity, will send data and parity bit, in the form of d_n ... $d_1 p$
(ps: this technique is also applied in DDH error detection)
e.g. $\underbrace{0111000110101011}_{16\ bit\ data\ bit,\ and\ there\ are nine 1\ so\ the\ parity\ bit\ should\ be\ 1}$ $\underbrace{1}$ The problem is that if two bit data changes, the parity bit will remain the same, so can only detect 1 bit

Mu Changrui

3.2 Two-dimensional Bit Parity

In two dimensional bit parity, the data is arranged in a two-dimensional array, and the single bit parity is computed for each row and column. The augmented array with both data and parity bits will be sent
Two-dimensional bit parity can

- **Detect and correct** single bit errors in data
- detect two bits errors in data

e.g

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

if one bit is changed

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

3.3 Cyclic Redundancy Check(CRC)

Cyclic redundancy check is a power error-detection coding. It consists of

- D: data bits viewed as a binary number
- G: generator of length (r+1) bits, agreed by sender and receiver beforehand
- R: CRC checksum is in length of r bits

CRC checksum is computed as follows:

1. Append r bits of 0's after data D
2. Perform bitwise XOR operation on the augmented data $D\underbrace{0..0}_{r\ bits}$.
   This XOR operation is done without carry or borrow(no nouce), between augmented data

Mu Changrui

and G.

The operations starts from the MSBs of augmented data and cascading to LSBs

3.3.1 MSBs and LSBs

- LSBs: the least significant bit(LSB) is the bit position in a binary integer giving the units value, which determines whether the number is even or odd
- MSBs: The most significant bit(MSB) is the bit position in a binaey number having the greatest value,., The MSB is sometime refers to as the high-order bit or left-most bit. MSB is used to differentiate the positive and negative number in 1's or 2's complement

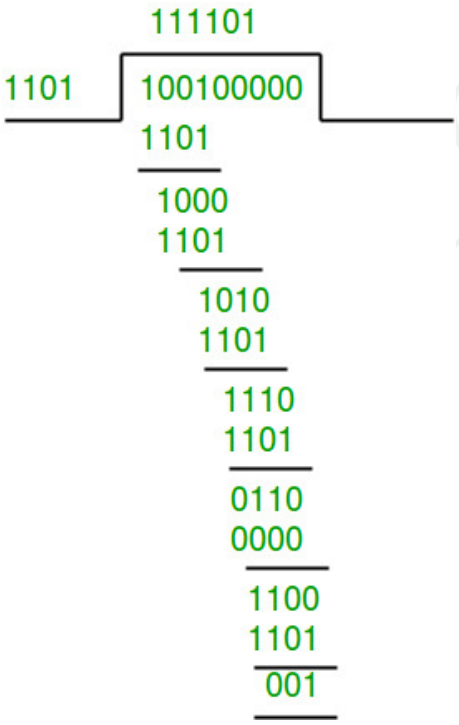
3. The remainder will be set as R, which replaces the trailing 0's in augmented data, to form the CRC checksum D+R, where + is the string contacenation if G∤ D+R, the error is detected

3.3.2 Example

e.g

data: 100100

generator: 1101



note that fail to pass the CRC check may due the corrupt of data or CRC or both, so if detecting

corruption, we cannot decide which part is the reason

4 Multiple Access Links and Protocols

There are two types of network links, namely

1. Point-to-point link, where a sender and receiver connected by a dedicated link, which do not need multiple access control(e.g the Serial Line Internet Protocol, AKA SLIP)
2. Broadcast link, where
   - Multiple nodes connected to a shared broadcast channe
   - When a node transmits a frame, the channel broadcasts the frame and each other nodes receive a copy

   In a broadcast channel, if two or more nodes transmit simultaneously, frame will collide at nodes and none will be correctly read
   Therefore, multiple access protocol is required, to serve as an distributed algorithm that determines how the nodes share the channel, using the channel itself.

4.1 Classification of multiple access links and protocol

Multiple access protocols can be categorised into three broad classes:

(a) Channel Partitioning: divide channel into smaller "pieces"; pieces are allocated to nodes for exclusive use
   i. TDMA
   ii. FDMA

(b) Taking turns: Nodes take turns to transmit
   i. polling
   ii. Token Passing

(c) Random access: channel are not divided so collision are still possible; The mechanism focuses on recovery from collision
   i. Slotted ALOHA
   ii. Pure ALOHA
   iii. CSMA

Mu Changrui

iv. CSMA/CD

v. CSMA/CA

5 Channel Partitioning Protocols

5.1 Time Division multiple Access(TDMA)

- The channels is divided into equal time slots
- Access to channel will be in ”rounds”
- each node gets fixed length slots(equivalently, equal transmission time) in each round
- Unused slots go idle(would not give to nodes in need)

5.2 Frequency Division Multiple Access(FDMA)

- Channel spectrum is divided into frequency bands
- Each node is assigned a fixed frequency band
- Used transmission frequency bands go idle

e.g

- node1: 4-8 kHz
- node2: 8-12 kHz
- node3: 12-16 kHz
- ...
- node6:24-28 kHz

If node2 was given 8-12 kHz, but it does not use it, it won't also give it to other node, it's a waste, it's what Chinese say: Zhan Zhe Mao Ken Bu La Shi

6 ”Taking Turns” Protocols

6.1 Polling

- there is a master node and rest of the nodes are the salves
- Master node invites slave nodes to transmit in turn

Mu Changrui

(No idea why they use such heavy-taste words, but at least easy to rememb...) Concerns of polling include:

1. polling overhead
2. single point of failure(at master node), if the master dead, the multiple access network dead

6.2 Token passing)

- there is a control token
- Control token is passing from one node to next sequentially Only the node with the control token can send

Concern of token passing include

1. token overhead(may generate multiple token)
2. single point of failure(the token may be lost)

7 Random access protocols

For random access protocols, when node has packet to send, there is no prior coordination among nodes. (who lucky, who serves)
The consequence is the presence of collision
The philosophy of RAP is lazy, which let it go until something wrong, then deal with the error
So the focus of the random access protocol is how to detect collisions and how to recover from collisions

7.1 Slotted ALOHA

- All frames require to have equal size
- Time is divided into slots of equal length
- Node start to transmit only at the beginning of a slot

The operation of each node includes:

- Collision detection: listens to the channel while transmissing
- Collision resolution: If collision happens, node retransmit a frame in each sub-sequent slot with probability p until success

Senders and receivers should synchronize the slots.The receivers also need to know when a slot starts and ends, in order to successfully decode the messages. Otherwise a receiver may take

Mu Changrui

some bits from a previous slot and some bits from the next slot and think that together this is one message. (One could introduce some special "markers" (e.g., bit patterns) that indicate the start of a slot. However, this then would indicate a synchronization point, meaning yes, we do need synchronization

7.2 Pure ALOHA

In pure ALOHA, there is no requirement of slot nor synchronisation: when there is a fresh frame, transmit immediately
The chance of collision(the main issue os efficiency) increases
A frame sent at t0 can collide with frame sent at [t0-d_pro,t0+d_pro]

7.3 Carrier Sense multiple Access(CSMA)

CSMA senses the channel before transmission and

- If channel is sensed idle, transmit frame
- If channel is sensed busy, defer transmission

Collision may still happen in CSMA. One scenario is when two nodes sense the channel idle at the same time and both start transmission
Even different nodes sense channel at different time, collision can still be possible due to propagation delay
One issue of CSMA is that node may not detect collision if the frame size is too small
Therefore, a minimum frame size is imposed on the frame. Payload will be padded 0 at the back if necessary to achieve minimum frame size
For example, Ethernet requires a minimum of frame size of 64 bytes

7.4 CSMA/CD)

CSMA(Collision Detection) is an extension from CSMA. It specifies the followiung rules:

- When collision is detected, transmission is aborted, to reduce channel wastage
- Retransmit after a **random amount of time**

7.4.1 How the random amount of time is selected

Assume it's the continuously n_th colission of the pkt, a numer K will be selected from 0,1,2,...,2^m-1, and wait for K*512 **bit** time
The concept 512 bit time refers to the time needed for the machine to transmit 512 bit data(so the real time is machine dependent, depend on the transmission delay in specific)

Mu Changrui

7.4.2 The Working Flow of CSMA/CD

1. NIC receives datagrams from network layer, captulating frams
2.
   - if NIC senses channel idle, start frame transmission
   - If NIC senses channel busy, waits until channel idle, then transmits
3. If NIC transmits the whole frame without detecting another transmission, NIC is done with frame
4. If NIC **detect other transmission**, abort and send a jam signal
5. After aborting, NIC enters binary back-off:
   (a) After m_th collision, NIC choses K randomly from 0,1,2,3,...,$2^m - 1$, and wait K*512 bit time then return to step 2
   (b) sometimes there is an upper bound of m, after which the frame is dropped.

7.4.3 Issue related to detect other transmission

The collision happens during the propagation, so how can the sender detect the collision?
Base on what I understand, when two signal collides, there will be a physical surge, which will be detected by the sender,receiver or any nearby node, and they may broadcast a jam signal to inform other nodes on the same channel.

7.4.4 The minimum frame problem regarding detecting collision

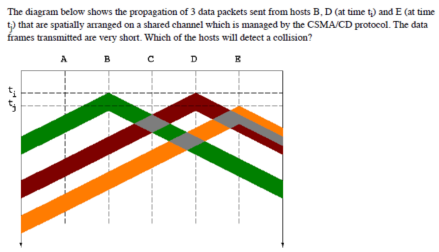
There is one question asking what is the minimum frame size to detect the collision
Based on my understanding above, if all nodes apply CSMA/CD, in fact they will try to sense whether the channel is idle or not before sending, that means if the frame is large, and have start sending for a while, other nodes will wait; The problem where the collision happens but not detected only happens at the beginning of sending(i.e. the time when leading bits are propagating, note that the physical layer is not store and forward, i.e. the frame is not sent as a whole; bits will be sent out once transmitted). So the worst case happens when the first bit of the frame is above to arrive destination but collision happens, so the surge/jam signal sent by receiver will one d_pro to go back to sender, when the surge/jam arrives sender

- if the sender has already finishes transmitting the frame: the collision cannot be detected
- if the sender is still transmitting the frame, the collision can be detected

Mu Changrui

Therefore, the let the frame size be L, if we wanna the collision neccessariky be detected, $\frac{L}{R} \geq$ RTT, note that the RTT here should be(the transmission delay for the first bit +2 propagation dalay, but the the transmission delay for the first bit is ignorable)

7.4.5 Exercise



C, D and E will detect the collision. Whenever two signal streams on the timeline of a host (the dashed line) mix together then a collision is, and must be, detected. Collisions should not just be detected by a sender. Let's assume B was sending a message to C, then we don't want C to receive garbled data.

So, all hosts need to constantly monitor the channel. If senders detect a collision they need to stop sending and try again later. If any other node detects a collision they need to ignore the data (it may have been addressed to them, but they must discard it).

Side note: this diagram shows a case that should never happen with Ethernet. Ethernet has a minimum frame length, i.e., the colored bands would be much deeper, which would ensure that all hosts on the network see all collisions. (That's why the question says that "The data frames transmitted are very short." because it is not supposed to be Ethernet.)

7.5 CDMA/CA

CSMA/CA(collision Avoidance) is an extension from CSMA. It specifies the receiver to return ACK if a frame is received OK and NAK is not
The sender initially transmits the intent to send the data, once an ACK received, the sender sends the data
CSMA/CA is used in IEEE802.11(WIFI) as collision detection by hosts is difficult in wireless context

Mu Changrui

7.5.1 Why Wifi uses CSMA/CA instead of CSMA/CD while CSMA/CD may be more efficient

IEEE802.11(WIFI) is half-duplex, which means it can either send or listen at one time(but not both)
Therefore it's not feasible for Wifi to listen to the channel to detect collision during sending

8 Local Area Network

LAN: computer that interconnects computers within a geographical area such as office or university campus
IMPORTANT: LAN may contains multiple subnets(e.g. the network in SOC is a LAN, which contains multiple subnet)

8.1 Link Layer Addressing & ARP

8.1.1 MAC address

Every adaptor has a MAC address aka physical or LAN address. MAC address is used to send and receive link layer frames.
MAC address is typically 48 bits long, the first three bytes of which identifies the vendor of an adaptor.
Adapter handles frames by checking if the destination MAC address of the frame matches its own MAC address.

- yes, adapter extracts the enclosed datagram and passes it to the protocol stack.
- no, adaptor discards the frame without interrupting the host.

8.1.2 Difference between IP and MAC

- IP
  - Dynamic Assigned
  - hierarchical
- MAC
  - fixed
  - identification of hardware(unique)

Mu Changrui

8.2 Address Resolution Protocol(ARP)

IP $\longrightarrow$ MAC
note that ARP was not designed to translate MAC back to IP, but computer can determine IP address from MAC address with DHCP
Each IP node has an ARP table, in which mappings of IP address and MAC address of other nodes in the same subnet is stored, in the following format:
IP address, MAC address, TTL
the TTL: time after which the address mapping will be forgotten(usually minutes in Windows)
Adaptor can "self learn" the ARP table
Three layer address in netword

- Host name: easy to remember for human
- IP address: end to end
- Mac address: node to node

"arp -a" order can be applied to search MAC address in linux

8.2.1 Workflow of ARP in Subset

Suppose node A sends frames to node B in the same subnet(A and B should be dirctly connected or through a switch or hub)

1. If node A has MAC address of node B in its ARP table, node A
   (a) creates a frame with B's MAC address and send it
   (b) Only node B will process this frame
   (c) All other nodes will receive but ignore this frame
2. If node A do not have MAC address of node B in the ARP table, node A
   (a) boradcast an ARP query pkt, the pkt contain
      - B's IP address
      - Destination MAC address set to FF-FF-FF-FF-FF-FF(broadcast MAC address)
      - A's Source MAC address
   (b) All nodes in the same subnet will receive this ARP query pkt, but only B will reply it, node B replies to node A with it's MAC address
   (c) node A caches B's IP-to-MAC in it's ARP table

Mu Changrui

(d) A returns to 1 to send pkt to B

8.2.2 Workflow of ARP for IP outside of subnet

Note that ARP only works between devices in the same IP subnet.
So for the IP address outside of subnet, ARP has a workflow to pass the work to router
Note that the ARP table will only store IP-MAC address within the same subnet
If node B is not in the same subset as node A, node A, when transmitting frames, should create a link layer frame with

- A's interfaced router's MAC address as destination MAC address
- B' s IP address as destination IP address.

When the packet reaches the router R, R will move datagram to outgoing link by constructing a new frame with B' s MAC address.
IMPORTANT, cannot use other MAC address of servers outside of the subnet, otherwise, every nodes in the subnet including the router will drop the pkt

9 Ethernet

Local Area Network(LAN) is a computer network that interconnects computers within a geographical area such as office building or university campus.
LAN technologies indlude:

- IBM Token Ring(expensive)
- Ethernet
- Wi-Fi
- Others

Ethernet is the dominant wired LAN technology. It comprises of a series of Ethernet standards of different speeds, different physical layer media, but MAC protocol and frame format remain unchanged.

9.1 physical media for Ethernet

- Twisted pair copper:
  - needs connecter(RJ45)
  - maxspeed around 10 Gbps

Mu Changrui

  - Max length: 100m
- Optical Fibre connecter
  - Needs reapeter, which enlarge signal
  - Max speed 10 to 40 Gbps
  - max length: $>$ 80 km

9.2 Types of Ethernet topology

9.2.1 Bus Topology

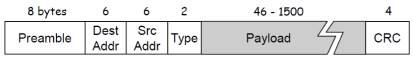
where all nodes can collide with each other

9.2.2 star topology

he switch in center eliminates collision

9.3 Ethernet Frame Structure

Ethernet frame are of the following format:



Preamble: **is not part of the Ethernet frame**

- 7 bytes with pattern 10101010 followed by 1 byte with pattern 10101011
- used to synchronize receiver and sender clock rates.

9.3.1 Structure of Ethernet frame

- Source MAC (6 bytes) and Distination MAC address (6 bytes)
- Type: indicates higher layer protocol(2 bytes)
- payload: (46-1500 bytes)
- CRC: used for error detection; corrupted frame will be dropped(4bytes)

10 Ethernet Data Delivery Service

- Connectionless: no handshaking required between sending and receiving nodes

Mu Changrui

- Unreliable: receiving nodes doesn' t send ACK or NAK to sending nodes. A consequence is that data in dropped frames will be recovered only if initial sender uses higher layer Reliable Data Trasfer Protocol.
- Ethernet' s multiple access protocol: CSMA/CD with binary backoff

11 Link Layer Switch

11.1 Function:

- Store and forward Ethernet frames
- Examine incoming frames' MAC addresses and selectively forward frame to one-or-more outgoing links

11.2 Properties

- transparent to hosts; no IP
- In Ethernet of star topology, hosts have dedicated connection to switch.)
- Switch buffers frames and is full duplex.
- CSMA/CD is used on each link, despite no collisions.

11.3 Switch Forwarding Table

Each switch has a switch table, which stores entry of the following format:
MAC address of host, interface to reach host, TTL

11.3.1 Self-learning of switch table

Switch learns which hosts can be reached through which interfaces. Specifically,

- When receiving a frame from A, the location and interface of A is noted down in switch table.
- If destination B is found in the table, forward the frame onto that link.(may refresh TTL)
- If destination B is unknown, broadcast the frame to all outgoing links, The destination B will send the ACK signal, from which the switch learns the location and interface of B and notes down them in the switch table.

Switches can be connected in hierarchy.

Mu Changrui

11.4 Difference between router and switch

Router	Switch
Check IP	Check MAC address
Computer routes to destination	Forward frame to category interface or broadcase
Mannual learn	self-learn

Both router and switch are store and forwarding

Mu Changrui