The Link Layer and LANs

SOFTENG 364: Computer Networks

Compiled: May 24, 2018

6.1. Introduction

1. What do we call the **protocol data unit** (packet) of the link layer?

> A **frame** (an **Ethernet** frame)

2. What services are provided by the link layer?

> **Framing** with header suitable for L2 addresses – perhaps trickier when dealing with bit stream arriving from a link
>
> **Medium/link access** i.e. coordination of transmission over shared links
>
> **EDC** beyond checksums
>
> **Reliable deliver** retransmission in the face of error ← not provided by Ethernet
>
> **Flow control** between pairs of nodes, driven by the receiver

3. Where is the link layer implemented?

> - In network interface card (aka. network interface controller, network adapter, LAN adapter, physical network interface)
> - Integrated **LAN-on-motherboard** configurations are now common

4. Do link layer reliable delivery services make those in the transport layer redundant? Explain. NB: Ethernet, the dominant L2 protocol, does not offer reliable delivery (only EDC)

> Although link layer protocols provide for reliable delivery over individual links, there is no guarantee that every individual link is reliable: Some links mightn't be under the control of the sending host.

5. If all links were to provide reliable delivery service, would the TCP reliable delivery service be redundant? Explain.

> - Packets are routed individually and may take different routes due to updates in forwarding tables or load balancing; hence, they may arrive out of order.

- The link layer cannot guarantee safe delivery in the face of:
  - Routing loops (a network layer problem)
  - Link or node failure
  - Errors arising between links i.e. while in a router's memory (the link layer is implemented in network interface cards)

6. Why not leave reliability to the transport layer completely?

See Link-level reliability is sufficient for end-end reliability, and is potentially much more efficient as we expect fewer link-level errors (and hence more likely to be corrected) and re-transmission is much less expensive over individual links.

> The **end-to-end principle**: *it is far easier to obtain reliability beyond a certain margin by mechanisms in the end hosts of a network rather than in the intermediary nodes, . . . especially when the latter are beyond the control of, and not accountable to, the former. . . . Positive end-to-end acknowledgements with infinite retries can obtain arbitrarily high reliability from any network with a higher than zero probability of successfully transmitting data from one end to another.* — wikipedia.org

7. Contrast **flow control** and **congestion control**.

Flow control concerned with a single pair of nodes; controlled by receiving side to prevent buffer overflow; "could happen at in layer"

Congestion control Network-wide concern; "happens above the link layer"

6.2. Error Detection and -Correction

8. Describe how a typical error detection scheme is employed to provide a reliable delivery service.

(a) Encapsulate the file in an L4 datagram, L3 packet, and a link-layer frame.

(b) Partition the binary representation into blocks of uniform size.

(c) Encode each block, generally adding redundant bits.

(d) Transmit blocks through noisy channel.

(e) If error is not detected, or if error can be corrected, discard redundant bits and recombine blocks into original frame.

(f) If errors are detected but not corrected, notify sender, which will re-transmit.

9. Compare/contrast the meanings of **source coding**, **channel coding**, and **cryptographic coding** and explain the relevance of each in communication networks, and the layer(s) in which it might be relevant.

source coding Compression, Application Layer: Elimination of redundancies (perfect reconstruction in lossless case)

channel coding Error control, transport layer, link layer: Introduction of redundancies (to increase transmission reliability)

cryptographic coding Encryption: Chapter 8

6.2.1. Parity checks

10. Describe how the **parity bit** is chosen for an **even** (or **odd**) parity check.

Append an additional bit to the message (of d bits) chosen so that the augmented $d+1$ bits have an even (**even parity**) or odd (**odd parity**) number of 1s.

11. How is the parity bit used to **detect** errors?

- Append the parity bit to the original message
- Transmit the augmented message
- An error is detected if the number of 1s is no longer ever (or odd).

12. Is the choice between these alternatives significant? Explain.

No: They perform equally well/poorly.

13. Explain how/why a (say) even **parity bit** scheme is only able to detect 50% of **burst errors**?

(Even) parity checks only detect errors with an odd number of errors. In the class of burst errors of any length, we expect errors even length to be about half-as-frequent as errors of odd length.

14. Describe how a **two-dimensional parity bit** scheme is able to both **detect** and **correct** any single bit in the original data.

- Row parity determines the row containing the error
- Column parity determines the column
- The erroneous bit has only one other possible value

15. How would a two-dimensional parity be used to identify and correct a single error in the **parity bits** in the final column?

- Row parity would determine the row index.
- That no errors occur in the parity bits beneath the original data indicate that the error is in the parity bits in the final column

16. How would two dimensional parity detect-but-not-correct any combination of two errors (i.e. in the original data or in the parity bits)?

- Note that two errors must occur in distinct elements (otherwise the errors would cancel).
- If two errors occur in the same row (resp. column), then the error would be detected in the bottom (resp. side) parity bits, but not in the side.
- Hence, we'd know both column (resp. row) numbers, but not the row (resp. column) numbers.

17. (Extra for Experts:) Show that the parity bit in the "bottom-right corner" of the augmented array acts a parity bit for both of the final row and final column.

$$\text{row parity}\ldots \quad \begin{bmatrix}\mathbf{1}^\top & 1\end{bmatrix}\begin{bmatrix}A \\ -\mathbf{1}^\top A\end{bmatrix} \equiv \mathbf{0}^\top \quad \text{and} \quad \begin{bmatrix}A & -A\mathbf{1}\end{bmatrix}\begin{bmatrix}\mathbf{1} \\ 1\end{bmatrix} \equiv \mathbf{0} \quad \ldots\text{col. parity}$$

$$\begin{bmatrix}\mathbf{1}^\top & 1\end{bmatrix}\begin{bmatrix}A & -A\mathbf{1} \\ -\mathbf{1}^\top A & x\end{bmatrix} \equiv \begin{bmatrix}\mathbf{0}^\top & 0\end{bmatrix}\Bigg\} \Rightarrow \quad x = \vec{1}^\top A \vec{1} \quad \Leftarrow \left\{\begin{bmatrix}\mathbf{0} \\ 0\end{bmatrix} \equiv \begin{bmatrix}A & -A\mathbf{1} \\ -\mathbf{1}^\top A & x\end{bmatrix}\begin{bmatrix}\mathbf{1} \\ 1\end{bmatrix}\right.$$

i.e. enforcing row-parity and column parity of the original data yield consistent formulae for "the parity bit on the parity bit rows- and columns".

6.2.1. Checksumming methods

18. Describe the **Internet Checksum**.

> *"The checksum field is the 16 bit one's complement of the one's complement sum of all 16 bit words in the header. For purposes of computing the checksum, the value of the checksum field is zero."* — RFC 791

19. In which internet protocols is is applied? To which fields?

TCP and UDP over header- and data fields

IP over header fields only

20. Checksums are not used in the link layer, where cyclic redundancy checks are used: Why not apply one "best" method in all layers?

- Transport and network layers are implemented in software: Less expensive checksums were preferred.
- Error detection in the link layer is implemented in specialized hardware.

"This is a simple to compute checksum and experimental evidence indicates it is adequate, but it is provisional and may be replaced by a CRC procedure, depending on further experience." — RFC 791

21. What kind(s) of error would a checksum not detect? Do you expect a checksum to be more reliable than a parity check?

- Any error that sums to zero; permutation of 16 bit blocks
- Yes: Random changes to numbers would typically change their sum.

6.2.3. Cyclic redundancy checks

22. In which protocols are CRCs actually used?

Ethernet; 802.11 WiFi; ATM

23. Define **modulo-2** addition and -subtraction. Are bits carried/borrowed if multi-bit patterns are added/subtracted?

$$0 \pm 0 = 0 \qquad 1 \pm 0 = 1 \qquad 0 \pm 1 = 1 \qquad 1 \pm 1 = 0$$

No – bits are neither carried nor borrowed.

24. Explain why modulo-2 arithmetic operates like an "old fashioned" clock face (with numerals at 12 o'clock and 6 o'clock).

Modulo-2 arithmetic "wraps around" just like a clock i.e. carries are discarded.

25. When calculating the EDC bits, the data bit-string is shifted by r bits: Explain why the shift is necessary, and why r is the degree of the generator polynomial.

- Required so that $D \cdot 2^r - R = [D, R]$ i.e. so EDC bits don't interact with data
- r is the degree of the generator: The remainder cannot have higher degree

26. Show that $\text{rem}(M_1 + M_2, G) \equiv \text{rem}(M_1, G) + \text{rem}(M_2, G)$.

$$M_1/G = Q_1 + R_1/G$$
$$M_2/G = Q_2 + R_2/G$$
$$(M_1 + M_2)/G = M_1/G + M_2/G = (Q_1 + R_1/G) + (Q_2 + R_2/G) = (Q_1 + Q_2) + (R_1 + R_2)/G$$
$$\text{rem}(M_1 + M_2, G) = R_2 + R_2$$

27. Suppose that the transmitted polynomial T is received as $T' = T + E$, where E is an error.
    (a) What is the maximum degree of E?
    (b) Show that $\text{rem}(T', G) = \text{rem}(E, G)$?

(c) Which errors would be undetected by the CRC scheme with generator G?

(a) $\deg(E) \leq \deg(T)$
(b) See workings in preceding question
(c) Those which G is a factor

28. Describe, in terms of bit patterns, the errors that a CRC cannot detect.

We can discuss these if there is time at the end of the course.

29. Comment on the performance of a well-designed r-bit CRC code in the face of burst errors.

A suitably-chosen r-bit generator can detect burst errors of up to $r + 1$ bits.

30. How are parity checks and CRC codes related? Explain why.

- Even parity corresponds to a CRC with generator $x + 1$.
- This generator bit-string `11` serves to eliminate `11` and to shift `10` to the right.

31. Can CRC codes perform error **correction**?

Yes, they can, but this capability is beyond the scope of our textbook.

6.3. Multiple Access Links and Protocols

32. Recall examples of shared physical media.

*Examples of shared physical media are **wireless** networks, **bus** networks, **ring** networks and point-to-point links operating in **half-duplex mode*** – wikipedia.org

33. Identify networks/protocols that support **broadcast links** (shared links).

wired Ethernet LAN (shared wire); 802.11 wireless LAN (shared RF channel); upstream HFC (Hybrid Fibre-Coaxial) networks

34. What is a **collision**?

When a node receives multiple signals simultaneously

A collision is the situation that occurs when two or more demands are made simultaneously on equipment that can handle only one at any given instant. – wikipedia.org

35. What is the purpose of a **multiple access protocol**?

To determine when (inc. how much) individual nodes can transmit over a broadcast channel (so as to minimize the impact of collisions).

36. Outline three or four desirable properties of an "ideal" multiple access protocol (with rate R bits/second, say).

Attains best-case throughput A single node is able transmit at R bps when the other nodes have nothing to transmit

Linear utilization When m nodes have data to transmit, the average transmission rate for each is R/m

Decentralized control: No single point of failure

Simple : Inexpensive implementation

37. Identify or describe three main categories of multiple access protocols.

Channel-partitioning TDMA, FDMA, CDMA

Random access slotted ALOHA, ALOHA

Taking-turns polling, token passing

38. Contrast the meanings of **efficiency** and **fairness** in the context of channel access.

If M nodes share a channel of rate R

Efficiency Aggregate transmission rate is close to R

Fairness Transmission for each of M active nodes uniformly divided

It is possible to be efficient-but-unfair and unfair-but-efficient!

6.3.1. Channel-Partitioning Protocols

39. Identify and describe three classes of **channel partitioning** protocols.

An analogy to the problem of multiple access is a room (channel) in which people wish to talk to each other simultaneously. To avoid confusion, people could take turns speaking (time division), speak at different pitches (frequency division), or speak in different languages (code division). CDMA is analogous to the last example where people speaking the same language can understand each other, but other languages are perceived as noise and rejected. Similarly, in radio CDMA, each group of users is given a shared code. Many codes occupy the same channel, but only users associated with a particular code can communicate. – wikipedia.org R bps channel shared by N nodes:

TDMA = "taking turns": Divide time into **frames**, frames into N **slots** (typically sufficient to transmit one packet); each slot is assigned to a node

FDMA = "different pitches": Divide channel into N **frequency bands**, each of bandwidth R/N;

CDMA Assigns a different **code** to each node

Channel partitions may be:

Statically pre-allocated to subscribers: **FAMA** (fixed assignment multiple access)

Dynamically allocated in response to requests: **DAMA** (demand assignment multiple access)

40. Explain the appeal and the drawback(s) of TDMA and of FDMA.

Assume channel has N users, channel has rate R:

Appeal eliminations collisions; perfectly fair

Drawback node transmission rate is limited to R/N i.e. channel is under-utilized unless all nodes have data to transmit – seems conservative

Kurose & Ross seem to have FAMA in mind.

6.3.2. Random Access protocols

41. Describe, in general terms, how random access protocols work.

- Each node transmits at the full rate R
- When collisions are detected, each node delays for a random time interval before retrying

42. Describe the **slotted ALOHA** protocol.

(a) Time slots sufficient to transmit one packet of size L bits (i.e. L/R seconds)
(b) Time slots are synchronized
(c) Node with packets to send (**active node**) prepares packet for transmission at beginning of next slot
(d) Collision detected during time slot
(e) In face of collision, retransmit with probability p at the beginning of next slot

43. Explain the appeal of slotted ALOHA relative to channel partitioning.

Continuous transmission at full-rate R when no other nodes are active.

44. In the context of ALOHA, explain the meaning of **successful slot**?

A slot in which exactly one node transmits i.e. the slot is utilized successfully

45. Each of the N nodes in a slotted ALOHA network transmits independently with probability p.

(a) Determine the probability that a specified node k transmits successfully in a given time slot.

(b) What, then, is the probability of successful transmission (by any node) in that time slot?

(c) Find the values of p that extremizes the expression in 45b.

(d) At this optimal p, determine the efficiency of a heavily-loaded system in the limit $N \to \infty$?

(a) $p(1-p)^{N-1}$

(b) $Np(1-p)^{N-1}$

(c)

$$\begin{aligned} E(p) &= Np(1-p)^{N-1} \\ E'(p) &= N(1-p)^{N-1} + Np(N-1)(1-p)^{N-2}(-1) \\ &= N(1-p)^{N-1} - Np(N-1)(1-p)^{N-2} \\ &= N(1-p)^{N-2}[(1-p) - (N-1)p] \\ &= N(1-p)^{N-2}[(1-p) - (Np-p)] \\ &= N(1-p)^{N-2}(1-Np) \end{aligned}$$

$$\begin{aligned} E'(p) = 0 \quad \Rightarrow \quad \text{either} \quad & N(1-p)^{N-2} = 0 \quad \Rightarrow \quad p = 1 \quad \ldots \text{“trivial” solution} \\ \text{or} \quad & (1-Np) \Rightarrow p = 1/N \quad \ldots \text{“desired solution”} \\ E''(p) &= (N-1)N(1-p)^{N-3}(Np-2) \\ \Rightarrow E''(1/N) &> 0 \quad \ldots \text{minimum} \end{aligned}$$

(d)

$$E(1/N) = N(1/N)(1-(1/N))^{N-1} = \left(1-\frac{1}{N}\right)^{N-1} = \left(1-\frac{1}{N}\right)^{N} \cdot \left(1-\frac{1}{N}\right)^{-1}$$

$$\text{Identities:} \quad \lim_{x\to\infty}(1-1/x)^x \equiv 1/e \quad \text{and} \quad \lim_{x\to\infty} 1-(1/x) \equiv 1-(1/\infty) = 1$$

$$\therefore \quad \lim_{N\to\infty} E(1/N) = (1/e)\cdot(1)^{-1} = 1/e$$

46. Repeat the calculations of 45 for “pure” ALOHA.

$$\begin{aligned} E(p) = \Pr(\text{success at node } i) = {} & \Pr(\text{node } i \text{ transmits}) \cdot \\ & \Pr(\text{no other node transmits in } [t_0-1, t_0]) \cdot \\ & \Pr(\text{no other node transmits in } [t_0-1, t_0]) \\ = {} & p \cdot (1-p)^{N-1} \cdot (1-p)^{N-1} \\ = {} & Np(1-p)^{2(N-1)} \\ E'(p) = {} & N(1-p)^{2(N-1)} + Np \cdot 2(N-1)(1-p)^{2(N-1)-1}(-1) \\ = {} & N(1-p)^{2(N-1)} - 2N(N-1)p(1-p)^{2(N-1)-1} \\ = {} & N(1-p)^{2(N-1)-1}[(1-p) - 2(N-1)p] \\ = {} & N(1-p)^{2N-3}[1-(2N-1)p] \end{aligned}$$

$$E'(p) = 0 \quad \Rightarrow \quad p = 1 \quad \textbf{or} \quad p = 1/(2N-1)$$

Substitute $p = 1/(2N-1)$ into $E(p)$ and take limit as $N \to \infty$.

47. Explain the trade-off that arises between **slotted ALOHA** and "pure" **ALOHA**.

- Slotted ALOHA has is more efficient – by a factor of 2 for $N \to \infty$
- Pure ALOHA is fully decentralized – no need for slot synchronization

48. Are collisions more- or less likely in pure- or unslotted ALOHA? Explain in one sentence or more.

Collisions are more likely in pure ALOHA: No synchronization; frame sent at a particular instant may collide frames sent both before and after that instant.

49. In the context of multiple access protocol, outline analogies between protocols of human conversation and:

    (a) **carrier sensing**

    (b) **collision detection**

CSMA "don't interrupt others" i.e. transmit only when channel is idle

Collision detection "polite conversationalist stops when others speak" – abort after detection to reduce "channel wastage"

50. With the aid of a diagram, explain how collisions are inevitable even if carrier sensing is employed.

- A free channel might be accessed by multiple nodes in the time interval required for signals to propagate between nodes
- This can be shown on a **space-time** diagram

51. How does **collision detection** improve the performance of a multiple access protocol.

Nodes can terminate transmission as soon as collision is detected:

- Saves transmission effort on the node
- Minimizes utilization of the channel

52. How soon after a collision is detected is a node in a CSMA/CD network able to re-transmit?

time to detect + time to abort + **backoff time**

53. Explain how a re-transmission delay is chosen by a **binary exponential backoff** algorithm, and why this process is sensible.

- After k collisions, select backoff time from $\{0, 2^0, 2^1, \ldots, 2^{k-1}\} \times M$ with equal probability, where M is some fixed value
- The interval is shorter when the number of collisions is small (reducing latency), and larger when the number of active nodes is large (reducing the probability of collision)

54. How is the backoff time defined in the **Ethernet** protocol.

- The constant M in 53 is 512 bit time.
- Backoff time is capped at $2^9 M$

55. Identify or describe the factor(s) influencing the efficiency of a DSMA/CD deployment?

(max) propagation time t_{prop} relative to (max) transmission time t_{trans}

56. Explain why carrier sensing cannot eliminate delay completely.

Channel propagation delay means that one node mayn't be aware that another has transmitted.

57. Why must a node wait for a random length of time (as opposed to a fixed length interval) after a collision is detected.

58. Identify or describe physical factors dictating the efficiency of a CSMA/CD network.

$$\text{efficiency} \approx \frac{1}{1 + 5d_{\text{prop}}/d_{\text{trans}}}$$

- d_{prop}: Propagation time between adapters
- d_{trans}: Frame transmission time

59. Describe the physical mechanism by which an adapter is able to detect collisions.

By sensing **signal energy** entering the adapter from the channel.

6.3.3. Taking-Turns Protocols: Polling and token-passing

60. Contrast potential sources of inefficiency of protocols based on "taking-turns" with those based on **random access**.

M nodes share a channel with rate R:

Random access A node may transmit at R/M only when $M = 1$; may be decentralized (e.g. pure ALOHA)

Taking turns Nodes may transmit at close to R/M for any M – but **polling delays** are relatively more significant for small M

61. Explain the workings of a protocol based on **polling**.

- **Master** node **polls** other nodes in **round-robin** fashion
- Specifies maximum number of frames that may be sent

62. Outline the positive and negative characteristic(s) of **polling** protocols.

Good No collisions, no empty slots

Bad **polling delay**

63. Explain the workings of a protocol based on **token-passing**.

Token passing is decentralized and highly efficient – [KR-493]

- Decentralized: No master node
- **token** = small, special-purpose frame
- Passed from node to node in sequence – to minimize transmission times
- On receipt of token, an active node transmits up to a maximum number of frames
- An inactive node forwards the token immediately

64. What challenge(s) does token passing raise for the protocol designer?

How to recover if the token is not forwarded – typically due to a hardware/software error in a node or link

65. Identify an existing protocol based on **polling** and on **token-passing**.

Polling IEEE 802.15 wireless personal area network (WPAN); Bluetooth

Token-passing IEEE 802.5 Token Ring local area network – *eclipsed by later versions of Ethernet*

6.3.4. DOCSIS: The Link-Layer Protocol for Cable Internet Access

66. The multiple access problem does not arise on the downstream channel of a cable access network: Why?

Only the **CMTS** transmits on the downstream channel

67. Explain how each of the following elements are incorporated into the DOCSIS protocol:
    (a) Channel partitioning
    (b) Random access
    (c) Centrally-allocated slots

- **downstream** = toward users (CMTS to modems); no multiple access problem
- **upstream** = modem to CMTS

Channel partitioning
- FDMA on both of downstream- and upstream channel; multiple modems per frequency channel i.e. number of modems typically exceeds number of frequency bands
- TDM-like partitions on upstream channel of time into intervals and further into mini-slots;

Centrally-allocated slots explicit allocation of **active** modems to **assigned mini-slots** (no collision) for data transmission...

Random access to mini-slots containing **mini-slot request frames** from active modems:
- no carrier sensing or collision detection on modems
- collision inferred by lack of response in next downstream control message
- binary exponential back-off

Summary of MAC protocols

Protocol	Low load	High load	Coordination	Issues	Examples
channel ptx	inefficient	efficient, fair	decentralized	empty slots	(DCMA wireless)
random access	efficient	inefficient	decentralized	collision, CS|CD, backoff	(slotted) ALOHA
polling	efficient	efficient, fair	centralized	polling delay, latency, single PoF (master)	Bluetooth
token-passing	efficient	efficient, fair	∼decentralized	token overhead, latency, single PoF (token)	Token Ring

6.4. Switched LANs

68. Recall the "addresses" employed at each layer of the Internet protocol stack.

	Layer	Address
4	Transport	port
3	Network	IP address, AS number
2	Link	MAC address

69. Complete the following sentence:

It is ....., not hosts and routers, that have link-layer addresses.

70. Provide at least two commonly used synonyms for "link-layer addresses".

hardware address, LAN address, MAC address, physical address

71. Explain how the MAC address space is managed to maintain uniqueness of addresses across devices.

Assigned by IEEE; first 24 (of 48) bits

72. Contrast **IPv4 addresses** and **MAC addresses**.

IPv4:

- 4 bytes = 32 bits
- Usually displayed in decimal `255.255.255.255`
- Hierarchical structure: network part (prefix) and host part (suffix) i.e. depends on the host IP subnet
- Assigned by Internet Assigned Numbers Authority (IANA)
- Analogous to postal address

MAC:

- 6 bytes = 48 bits
- Usually dispalyed in hexadecimal e.g. `FF:FF:FF:FF:FF:FF` or `FF-FF-FF-FF-FF-FF`
- Flat structure: Independent of the host network (prefix allocated to original vendor)
- Assigned by the IEEE
- Analogous to IRD number

73. How many MAC addresses are (typically) associated with a host? With a router?

host Just one on a single NIC

router One on the NIC on each network interface

74. Contrast the meaning of **adapter** and **interface**.

There is a one-to-one correspondence between the interfaces and adapters on a given computer. An interface is an IP-level abstraction, whereas an adapter is a datalink-level abstraction. — Documentation for Microsoft IP Helper

75. Why are IP addresses and MAC addresses both indispensable?

Both are essential to DHCP:

- A IP address is the "postal address" of the NIC – completely specifying its location in the context of the entire Internet.
- A MAC address is the "personal identity" of an NIC – our means of referring to the NIC in isolation, before it has been allocated an IP address.

76. How does addressing differ in the network layer and in the link layer?

network data is addressed to (interface on) destination host

link frame addressed to (interface on) the next hop

77. What happens when a frame's destination MAC address doesn't match that of a receiving interface? What if the frame has the MAC **broadcast address**?

- Unrecognized addresses are not even passed up to the network layer...
- ...but the broadcast address is.

78. What is the role of the **Address Resolution Protocol** (ARP).

To translate between IP addresses (network layer) and MAC addresses (link layer)

79. Compare and contrast the **Address Resolution Protocol** (ARP) and the **Domain Name System** (DNS).

DNS Translates host names to IP addresses; for entire Internet

ARP Translates IP addresses to MAC addresses; only for hosts on same subnet

80. Identify or describe the (most important) columns of an **ARP table**.

IP address; MAC address; TTL

81. What is the role of the **TTL** field in **ARP tables**? How does this compare with its role in **IP packets**? With its role in DNS?

- Preserve space in fixed-size ARP table
- Prevent IP packets from circulating forever in the presence of routing loops

82. Which Ethernet nodes contain an **ARP table**?

All nodes

83. What is an **ARP module**?

The operating system module responsible for maintaining the **ARP table**.

84. What fields are carried by an **ARP packet**?

MAC addresses of sender and receiver

IP addresses of sender and receiver

85. How would you distinguish a frame carrying an **ARP query packet** from one carrying an **ARP response packet**?

Destination MAC address of query's frame is **broadcast** MAC address. Can also look at ARP

86. Describe the workings of an ARP query.

- Query is sent in broadcast frame (destination MAC address `FF-FF-FF-FF-FF-FF`)
- Response is sent in standard frame
- Table is build automatically

87. What happens when a host receives an **ARP query packet**?

(a) Passed by adapter (because of broadcast address) up to ARP module
(b) Update entry in ARP table with IP and MAC address of source host
(c) Compare destination IP address with interface's IP address;
(d) if addresses match, then send response to host's MAC address

88. What is a typical (initial) value for **TTL field** in an ARP table entry?

20 minutes

> *The ARP module maintains a cache of mappings between hardware addresses and protocol addresses. The cache has a limited size so old and less frequently used entries are garbage-collected. Entries which are marked as permanent are never deleted by the garbage-collector.* – (Linux `arp` man page)

89. *ARP is plug-and-play*: Explain briefly.

No intial configuration is necessary: Entries in a

90. *ARP straddles the boundary between the link and network layers*: Discuss briefly.

- ARP packets are carred in link layer frames i.e. "they sit above the link layer, architecturally"
- ARP packets carry MAC addresses and IP addresses

91. How can a sending host tell whether a destination interface lives on the same subnet?

It knows the IP address of itself and of the destination.

92. How does an IP packet reach a destination off the subnet of the sender?

Note: The destination could be anywhere on the Internet.

- By comparing IP addresss of itself and the destination, knows that destination is on another subnet.
- IP packet with destination IP address is encapsulated in link layer frame with MAC address of the **first-hop router**.
- Routing adapter receives frame addressed to itself, so passes IP packet up to the router (network layer).
- Router consults it forwarding table; forwards IP packet to outgoing interface
- Output interface passes IP packet to its adapter, responsible for framing and sending into next subnetwork with appropriate MAC address (i.e. of destination or of next-hop router) – obtained by ARP.

6.4.2. Ethernet

93. Outline some of the factors that may have contributed to the success of **Ethernet** as the dominant wired LAN technology.

- Early leader: The first widely-adopted LAN technology
- Cheaper and simpler than early competitors
- Sufficient evolution in the face of competition
- Ethernet hardware has become cheap and accessible

94. Contrast early **hub** devices with modern **switch** devices. In what sense is a switch "smarter" than a hub?

Hubs are **physical layer** devices:

- Work with bits rather than **frames**
- Replicate bits arriving at one interface
- Boost and transmit on all other interfaces
- **Collisions** forces sending nodes to retransmit

95. Explain the role of the **preamble** field in an Ethernet frame.

*An Ethernet packet starts with a seven-octet **preamble** and one-octet **start frame delimiter** (SFD).*
*The preamble consists of a 56-bit (seven-byte) pattern of alternating 1 and 0 bits, allowing devices on the network to easily synchronize their receiver clocks, providing **bit-level synchronization**. It is followed by the SFD to provide **byte-level synchronization** and to mark a new incoming **frame**.* – wikipedia.org

96. *Ethernet technologies provide **connectionless** service to the network layer*: Explain briefly and compare this service to those of transport layer protocols.

97. What happens when an Ethernet frame fails the CRC check at the receiving interface?

The packet is discarded (without positive/negative acknowledgement to the sender).

98. Ethernet communication might be described as "unreliable" and "connectionless": Explain each of these terms.

Unreliable receiving NIC doesn't send acks or nacks to sending NIC

Connectionless no handshaking between sending and receiving NICs

If an Ethernet frame fails a CRC check, it is simply dropped – without negative acknowledgement

99. If Ethernet does not provide a reliable service, how are LAN users assured of reliable communication?

TCP and/or application layer services

100. Identify or describe the **MAC protocol** employed in Ethernet networks.

Unslotted CSMA/CD with binary backoff

101. Identify or describe each of the fields in the Ethernet frame structure.

Preamble 8 bytes used for synchronizing the clocks of the sending- and receiving interface

Destination MAC address (6 bytes)

Source MAC address (6 bytes)

EtherType 2-byte code indicating to which upper-layer protocol the payload should be demultiplexed – typically ARP (`0x0806`), IPv4 (`0x0800`) or IPv6 (`0x86DD`), MPLS unicast (`0x8847`), MPLS multicast (`0x8848`)

Payload 46-1,500 bytes

Frame check sequence 32-bit CRC value; note that it is neatly "appended" to the payload field `:)`

102. Which layer(s) do the Ethernet protocols cover?

physical layer + link layer

103. What do different members of the Ethernet family have in common and how do they differ?

	Same	Different	
MAC protocol	✓		
frame format	✓		
speed		✓	{2, 10, 100} Mbps, {1, 10, 40} Gbps
physical layer medium		✓	copper, fibre

104. Explain the identifier in, for example, the `100BASE-T` Ethernet standard.

`100`	speed	100 Mbps
`BASE`	configuration	"baseband Ethernet" – medium carries only Ethernet traffic
`T`	physical medium	twisted wire

105. Compare and contrast the configuration and features of (network layer) **routers** and (link layer) **switches**.

What follows is taken from the nice discussion in Kurose & Ross (7e), page 513–514:

Switches

- ✓ Selective forwarding
- ✓ Buffering
- ✓ Completely **plug-and-play**, i.e. self-learning; utilises incoming port number and sender's MAC address
- ✓ High filtering/forwarding rates – frames processed only up to L2 fields
- ✗ Restricted to spanning tree topologies
- ✗ ARP broadcast traffic doesn't scale well
- ✗ Susceptible to **broadcast storms**

Routers

- ✓ Selective forwarding
- ✓ Buffering
- ✗ Not plug-and-play: Routers and hosts that connect to them need **IP addresses** to be configured
- ✗ Larger per-packet processing time – datagrams processed up to L3 fields
- ✓ Not restricted to spanning trees (if routing loops are avoided) – optimal paths, redundancy
- ✓ Firewall protection against broadcast storms

106. Compare and contrast (level-2) **switches** and (level-3) **routers**.

Similarities

- Full duplex (interfaces can send and receive simultaneously)
- House forwarding tables

Switches:

- Examines MAC addresses
- Need process only up to level 2 (frames)
- Active topology of a switched network is restricted to a (spanning) tree, to prevent cycling
- Switching table is populated by self-learning process
- Network size is limited by ARP traffic and processing
- Susceptible to broadcast storms

Routers:

- Examines IP addresses
- Need to process up to level 3 (datagrams) – larger per-packet switching times
- Active topology not restrictred to spanning tree and can include redundant links (cycling only possible if routers are misconfigured), helpful for performance and reliability
- Forwarding table populated by network administrator or software controller
- Provide firewalls against level-2 broadcast storms

6.4.4. Virtual Local Area Networks (VLANs)

107. Explain the notion of traffic isolation.

The ability of a router to prevent interfaces on one subset of hosts from ever seeing traffic between between another subset of hosts. In comparison, LAN switches generate **broadcast traffic**.

108. Explain how **virtual LANs** (VLANs) provides for:

(a) Traffic isolation

(b) Efficient utilization of switching hardware

(c) Flexible user management e.g. physical transfers between organization units

Isolate traffic Just as for ordinary LAN's, broadcast traffic cannot flow between VLAN i.e. each VLAN forms a single broadcast domain, even if multiple VLANs share a single physical switch:

- Provides security from traffic sniffing
- Improves LAN performance – broadcast traffic doesn't traverse the entire organizational network

Hardware utilization A single large switch (split between VLANs) may be less expensive than multiple smaller switches (one for each organizational unit)

User management Physical cabling needn't be changed to connect a user to his/her old physical switch – can simply reassign a port on the nearest switch in his/her new location

109. How are packets transferred between VLANs?

Logically Done via (network layer) routing – just as with separate switches (physically separate LANS)

Physically In practice, vendors sell combined switches plus routers

110. Explain the notion of VLAN **trunking** and **trunk port**.

Trunking allows multiple VLAN's to direct traffic through a single switch port (the trunk port).

111. What would be required to connect users of a given VLAN that that are located in different physical locations in the absence of trunking? (i.e. What savings does trunking offer?)

- If we had one switch in each location
- and the uses of each switch were divided between M VLANs,
- then M ports on each switch would be need to be dedicated to connect their respective VLANs.
- In contrast, a single trunk port would be required.

112. How is trunking implemented?

- Via an extended frame format – **802.1Q** cf. the standard 802.1 frame format.
- Additional **VLAN tag** fields includes:

  TPID Value of `0x8100` in bits occupied by **EtherType** field of ordinary frame

  Tag control information (TCI) includes

  Priority point code (PCP)

  Drop eligible indicator (DEI) indicating frames able to be dropped in presence of congestion

  VLAN identifier (VID) ← $**$

113. Contrast **port-based** VLANs with **MAC-based** VLANs.

Port-based Switches' ports/interfaces are divided into groups

MAC-based MAC addresses of hosts in each group

VLANs may also be defined based on network-layer protocols and other criteria — [KR-519]

114. Explain the role of the (optional) **VLAN tag** (IEEE 802.1Q tag) in the Ethernet frame.

6.5. Link Virtualization: MPLS

115. Use an example to explain the meaning of **overlay network**.

*An overlay network is a telecommunications network that is built on top of another network and is supported by its infrastructure. An overlay network decouples network services from the underlying infrastructure by **encapsulating** one packet inside of another packet. After the encapsulated packet has been forwarded to the endpoint, it is **de-encapsulated**.* — searchsdn.techtarget.com

The Internet was originally built as an overlay upon the telephone network, while today (through the advent of VoIP), the telephone network is increasingly turning into an overlay network built on top of the Internet. — wikipedia.org

116. What is a **virtual circuit**? How does it relate to **circuit switching** and **packet switching**?

- A virtual circuit is a fixed path set up for all packets in a given connection.

  *A **virtual circuit** (VC) is a means of transporting data over a packet switched computer network in such a way that it appears as though there is a dedicated **physical layer** link between the source and destination end systems of this data.* – wikipedia.org

117. Contrast **virtual circuit** and **datagram** networking.

From Shay, *Understanding Data Communications and Networks* (3e), page 695:

	Virtual Circuit	Datagram
congestion control	Routers able to reserve space for anticipated packet arrival	Unexpected packet arrival makes control difficult
adaptability	changing conditions may make long-lived VCs suboptimal	each packet routed according to current network state
routing overhead	each virtual circuit is routed just once	separate routing decision for each packet
ordering	packets arrive in order	packets arrive out of order
robustness	packet loss if node/link failure breaks VC connection	packets may be routed around failures

118. Explain the original motivation for the development of the **Multiprotocol Label Switching** (MPLS) protocol.

To augment the (older, destination-based) IP infrastructure with one based on

feature	benefit
fixed-length labels	fast forwarding
and virtual circuits	low latency; congestion planning
but using IP addressing and routing	existing infrastructure

*The major goal of MPLS development was the increase of routing speed. . . . This goal is no longer relevant . . . because of the usage of newer switching methods, such as ASIC, TCAM and CAM-based switching. Now, therefore, the main application of MPLS is to implement limited **traffic engineering** and layer 3 / layer 2 "service provider type" VPNs over IPv4 networks.*

119. In what sense is MPLS (Multiprotocol Label Switching) "multiprotocol".

*MPLS can encapsulate **packets** of various network protocols, hence its name "multiprotocol".* – wikipedia.org
MPLS headers are defined for IP, ATM, etc.

120. With which protocol layer is MPLS associated? Discuss.

- MPLS encapsulates the network layer datagram
- Sits between link layer and network layer

121. Compare and contrast MPLS with software defined networking (SDN) technology.

SDN further extends the scope of MPLS.

122. Contrast the fields of the MPLS forwarding table to those of a traditional IP forwarding table.

- Additional fields: `in label` and `out label`
- Match label field of MPLS-enhanced header with in lael

123. Explain how **label switching** improves forwarding efficiency.

There is no need for an MPLS-capable router to perform longest-prefix matching

124. Explain how **label switching** increases routing flexibility and robustness.

Flexibility	Different routes for traffic (from different origins) to a single destination – policy/performance/security
Flexiblity	e.g. Traffic isolation in VPNs
Robustness	Pre-computed backup paths

125. How is MPLS routing actually implemented?

- Extension of link-state routing (cf. OSPF, IS-IS)
- Algorithms are currently vendor-specific

6.6. Data Centre Networking

126. What are the roles of a load balancer within a data centre?

- NAT-like translation of public/external IP address to internal address
- Balance workload across hosts/blades

127. Outline a similarity between data centre load balancing and **Network Address Translation** (NAT).

128. Explain why data centre load balancing might be described as **application-layer routing**.

- It makes forwarding decisions based on the destination **port number** (layer-4) in addition to the IP address (layer-3). See nginx.com glossary.

129. Discuss the relative strengths and limitations/challenges associated with hierarchical topologies and alternative highly-connected topologies.

- Benefit: scale
- Limitation: host-to-host

130. Briefly describe one recent trend in data centre networking technology.

Fully-connected topologies Every tier-1 switch to every tier-2 switch
Routing algorithms in highly connected topologies
Modular Data Centres (MDC)
Specialization/customization of hardware and software

6.7. Retrospective: A day in the life of a Web Page Request

This material draws from all other modules.

References