Network Notes

Oct.2019

XU Kangyan

1. **Client-Server Architecture**
   Server: Always-on host / Permanent IP Address
   Client: No directly communicate / may dynamic IP Address
2. **Peer to Peer (P2P)**
   No always-on server / arbitrary end systems directly communicate
3. **CDN:** Content Delivery Network (all over world only store replica)
4. **ISP:** Internet Service Provider
5. **Hub:** fastest, identify no address
   Switch: only identify MAC address
   Router: identify MAC and IP address
6. In **switched network**, the switch "reassemble" the frame and check the "header" of the frame to read the "destination" MAC address. Then, consult with "forward database" that maps MAC address to physical port and forward bits to the port.
7. **Internet** is worldwide, publicly accessible network of inter-connected computer networks that transmit data by *packet switching*, using the standard "Internet Protocol"(IP). It is a "network of networks".
8. IP(unreliable) / TCP(reliable) / UDP(unreliable)
9. The two most important network performance measures are **Delay**(sec) and **Throughput**(bps).
10. **End to End delay:** Tp, Tt, Tprocess(random), Queue Delay(random)
11. **Bandwidth:** The maximum amount of data that can travel through a channel (bps or Hz).
    Throughput: how much data actually does travel through the channel *successfully*.
12. **RTT:** Round Trip Time
13. loss and delay reason: packets queue in buffer / packets arrival rate exceeds output rate / packets queue wait for turn
14. **Rb:** *Data Rate*, the rate of which the bits are transmitted.
15. **Bandwidth delay product:** the maximum number of bits can be inserted in the link in a given interval of time (Tp or RTT).
16. **OSI(Open System Interconnection) Model: Application** / presentation / session / **Transport**(TCP segment or UDP datagram) / **Network**(IP packet) / **Datalink**(frame) / **Physical**
17. **IP address:** net id & host id (assigned by local administer) 32bits
    MAC Address: 48bits
    Port number: 16bits identify application

18. FOUR QUESTIONS

Q: How does a host/router *get the MAC address* of another host /router on the same LAN?

A: **ARP (Address Resolution Protocol)**

Q: How does a host *get the IP address* of another host across the Internet?

A: **DNS (Domain Name System)**

Q: How does a host *get its own IP address*?

A: **DHCP (Dynamic Host Configuration Protocol)**

Q: How do we distinguish between two or more applications running on the same host?

A: Port Numbers / Sockets

19. DHCP: C Discovery / S Offer / C Request / S Acknowledge

DHCP server: not necessary on each network / might more than one

Transaction ID

Both client & server side are well-known port number (S-67 C-68).

Return Your IP / Router IP / DNS Server IP / subnet mask ...

20. DNS

Local Name Server: The default name server that will receive the DNS query from the host.

Root Name Server

Authoritative Name Server: Where the host register its name/IP address.

Recursive / Iterative

21. A **socket** is an abstract representation of a communication endpoint.

A socket allows the application to "plug in" to the network and communicate with other applications.

A socket is uniquely identified by the IP address, port number and the underlying transport layer protocol.

Parent socket / child socket

Socket types: reliable(TCP) / unreliable(UDP)

22. Multiplexing

Statistical TDM (used in internet): Time slots allocated dynamically based on demand, time slots needn't of same duration, every time slots need to be processed by a header.

Synchronous TDM (used in telephone): many slots wasted.

23. HTTP: persistent(w or w/o pipelining) / non-persistent(parallel or serial)

24. single parity check

25. FCS: Frame Check Sequence

26. Bit stuffing: used to avoid confusion with data containing 01111110.

27. Protocols define format, order of messages and received among network entities, and actions taken on message transmission, receipt.
28. Error Control: Stop-and-wait ARQ / **Sliding Window ARQ** (**Go-Back-n / selective reject**)

ARQ: Automatic Repeat Request

Stop-and-wait: send one frame at a time

Sliding window: send several frames at a time

SWS: sender window size (max number of frames the sender can send before he has to stop)

RWS: receiver window size (max number of frames the receiver is willing to receive)

Go-Back-n: SWS <= $2^m - 1$ / RWS = 1

Selective Reject: SWS <= 2^{m-1} / RWS = SWS

29. Link Utilization: percentage of time the link is actually used for "useful" transmission of information (%).
30. Multiple Access Protocol:

Random Access Protocols: ALOHA CSMA/CD CSMA/CA(Wi-Fi) (collision allowed)

Controlled Access Protocols: Reservation Polling Token-passing

Channelization Protocols: FDMA(frequency) TDMA(time) CDMA(code)

31. ALOHA: random access send and hope you are the only one (uti 18%). Time collision may occur: 2Tf

Slotted ALOHA: Must transmit at the beginning of time slot (uti 36%). Time collision may occur: Tf

32. CSMA/CD: Carrier Sense Multiple Access with Collision Detection (*Ethernet*)

A node should not finish transmitting its frame (Tf) before a possible collision can be detected (2TP for detecting).

For the protocol to work: Tf >= 2Tp

An Ethernet Frame: cannot be too short (protocol will not work) / cannot be too long (unfair)

33. Learning Switch (How does a switch configure itself)

Switch is a *plug & play* device.

Modes of operation: *Filtering* (drop) / *Forwarding* / *Flooding* (as hub)

Learning based on source MAC address

L2 switch is a *transparent device* as far as the end hosts are concerned. Node not communicate with switch (unlike router).

L2 switch *doesn't isolate broadcast domains*.

34. Wireless LANs (Wi-Fi: Wireless Fidelity)

CSMA/CA: Carrier Sense Multiple Access with Collision Avoidance

AP: Access Point

35. BSS: Basic Service Set (using one AP, choosing one frequency)

36. CSMA/CD not work given *wireless Link Characteristics* & *Hidden Terminal Problem*.

37. In 802.11 there divide 11 frequencies, each 10 MHz wide, occupying ≈ 85MHz. There is overlapping channel, channel 1, 6, 11 not.

38. Channel Association

SSID: Service Set Identifier

MAC address

Passive Scanning: AP broadcast a **Beacon** at his frequency. Your NIC scan all channels, choosing stronger one.

Active Scanning: Your NIC broadcast a **Probe**. If no reply or weak, NIC change the channel.

Hot Spot: Someone willing to share this Wi-Fi.

39.MAC Protocol:

CSMA/CA: Collision Avoidance (Required) Based on random access.

RTS/CTS: *Request to send / Clear to send* (Option) To combat the Hidden Terminal Problem.

40. IFS: Inter-Frame Spacing

DIFS: *Distributed IFS* (50 μsec)

SIFS: *Short IFS* (10 μsec)

41. Wi-Fi cannot detect collision. Once transmitting, *NO STOP*. ACK to tell collision not occurred.

42. CSMA/CA:

Station ready to send starts sensing the medium.

If the medium is free for the duration of an IFS, the station can start sending.

If the medium is busy, the station has to wait for a free IFS, then the station must additionally wait a *random* back-off time (collision avoidance, multiple of slot-time, **Contention Window**).

If another station occupies the medium during the back-off time of the station, the back-off timer stops (count / freeze / resume counting).

43. RTS / CTS:

Allow sender to "reserve" channel rather than random access of data frames: avoid collisions of long data frames.

Sender first transmits small **RTS** frames (may still collide but short) with reservation parameter (amount of time the data frame needs the medium) after waiting for DIFS.

AP broadcasts **CTS** in response to RTS after SIFS (*CTS heard by all nodes*).

Sender transmits data, acknowledgement via ACK, other stations defer transmissions.

44. Exposed Terminal Problem

45. Ethernet Router not understand 802.11 frame. AP translates to 802.3 frame.

46. 802.11 frame addressing: Address 1 – MAC address of wireless host or AP to receive this frame. / Address 2 – MAC address of wireless host or AP transmitting this frame. / Address 3 – MAC address of router interface to which AP is attached. / Address 4 used only in ad hoc mode.

47. Drop: network congest / TTL expires / Header check

48. MTU: Maximum Transmission Unit, the **maximum size** of the **data field** (**payload**) in the **frame**. If Packet size > MTU, needs **Fragmentation**.

49. Fragmentation can be done by the source host and/or any router across the internet.

Reassemble of fragments can be done *ONLY* at destination host.

Fragments can be fragmented again.

50. Total length (16-bits): (header (20-60 bytes) + **payload)** (**in bytes**)

51. Packet Header Fields related to Fragmentation:

Identification (**16-bits**): All fragments of a packet has **the same ID number** which is the same as that of the original packet.

Flags (3-bits): D=1 Do not Fragment / M=1 More Fragment **M=0 Last**

Fragmentation Offset (**13-bits**): Relative position of the fragment to the whole **packet** measured in units of 8 Bytes.

(8 is because $2^{16}/2^{13} = 8$) can avoid lost.

52. The payload of each fragment (in bytes) must be **a multiple of 8** *except for the last fragment*.

53. Classful IP Addressing

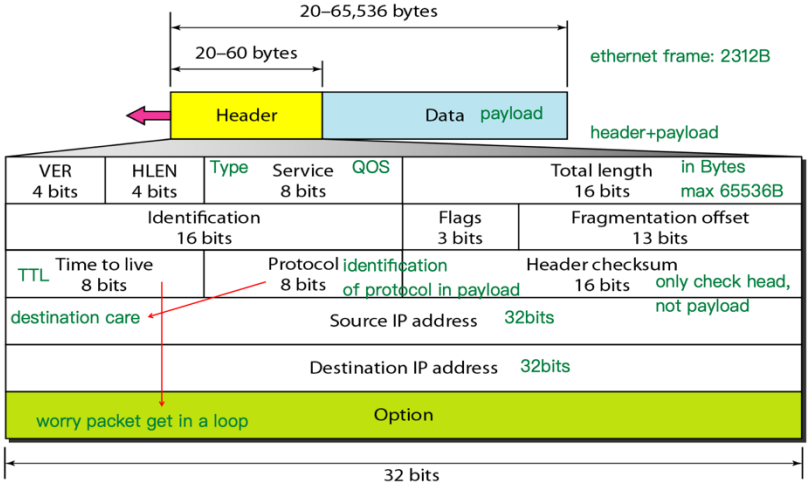
A: First 0, net id 7bits (1~126,0,127), host id 24bits.

B: First 10, net id 14bits, host id 16bits. 128~191.

C: First 110, net id 21bits, host id 8bits. 192~223.

D: 1110 224~239. E: 1111 240~255.

54. IP Packet Format



55. Private IP Addresses are non-routable

Range			Total
10.0.0.0	to	10.255.255.255	2^{24}
172.16.0.0	to	172.31.255.255	2^{20}
192.168.0.0	to	192.168.255.255	2^{16}

56. Special IP Addresses

Special Address	*Netid*	*Hostid*	*Source/Destination*
▪ *Network Address*	*Specific*	*All 0's*	*None*
▪ *Direct Broadcast Address*	*Specific*	*All 1's*	*Destination*
▪ *Limited Broadcast Address*	*All 1's*	*All 1's*	*Destination*
▪ *This host on this network*	*All 0's*	*All 0's*	*Source*
▪ *Specific host on this network*	*All 0's*	*Specific*	*Destination*
▪ *Loopback address*	*127*	*Any*	*Destination*

57. NAT: *Network Address Translation* (Implemented in the border router)

NAT is a protocol that maintains a translation table for mapping an internal private IP address to a globally unique IP address and vice versa.

58. NAT / PAT: *Port Address Translator*

59. Subnetting: the process of dividing a network into smaller size networks called subnets.

A subnet is a network.

Subnet is *transparent to the rest of the world*. Internet routing tables are not affected by subnetting.

Subnetting creates another level of hierarchy (net id, subnet id and host id). "stealing" some bits from the host id field.

60. Subnet Masking

A subnet mask is a 32-bit pattern, which has a "1" in every network and subnet (if any) locations and a "0" in every host location.

All devices including hosts & routers that are on the same subnet **MUST** have the same subnet mask. Devices on different subnet MAY have the same subnet mask.

61. Classless IP Addressing

Address Format: a.b.c.d/n, where n is # of bits in network and subnet portion of address. # of bits used for host is 32-n.

62. Subnet Router Routing Table

Logic AND with column 2 (Subnet Mask), check with column 1 (Subnet Address) and match, see column 3 (Router or others).

Longest prefix match: longest subnet mask.

63. CIDR: Classless Inter Domain Routing
64. Address Aggregation
65. IP Packet Delivery: **Routing** and **Forwarding**
Routing: the process of discovering and selecting the path to the destination according to some **metrics.**
Forwarding: the process of inserting the IP packet into a Layer2 frame and forwarding the frame to the next hop. (very fast)
66. Distance Vector Protocols (**RIP:** Routing Information Protocol)
Every router will tell its neighbor what it knows about all other routers. Based on **Bellman Ford.**
Link State Protocols (**OSPF:** Open Shortest Path First)
Every router will tell all other routers what it knows about its neighbors. Based on **Dijkstra.**
67. Count-to-∞ Problem: caused by the delay in A to tell B the failure of the link to access X.
68. Split Horizon: Router doesn't advertise to its neighbors.
Poison Reverse: advertise neighbor the distance is ∞.
69. LSP: Link State Packets
SPT: Spanning Tree
70. Transport Layer Protocols Function:
Provide for **Process-to-Process** communications (Port numbers are used).
Provide for end-to-end **Error Checking** (both TCP and UDP).
Provide for **Error Control** and **Flow** and **Congestion control** (only TCP).
71. TCP
TCP is a point-to-point, connection-oriented, reliable, end-to-end protocol. TCP is a **Byte-Oriented** Protocol.
72. TCP Segment Format
Sequence Number identifies the **number of the first byte** in the payload.
Acknowledgement Number is the **number of the next byte** expected to be received.
Receiver Window Size indicates the number of bytes the receiver is willing to accept.
A SYN segment doesn't carry data, but it consumes one sequence number.
A SYN/ACK segment doesn't carry data, but it consumes one sequence number.
An ACK segment can carry data.
73. Components of **Reliability in TCP:**
Sequence Numbers/Acknowledgements, Retransmissions, Timeout Mechanism.
TCP uses **single** retransmission timer.
Retransmissions are triggered by: **Timeout** events & **3 Duplicate ACKs.**
If retransmit a segment, ignore the RTT measurement. (might misleading)

74. Spare room in buffer = **Rcveiver Window Size (Wa)** = RcvBuffer – [Last Byte Rcvd –Last Byte Read]

Sender Window Size <= Receiver Window Size – [Last Byte sent – Last Byte Acked]

75. Congestion: too many sources sending too much data too fast for network to handle.

congestion window (Wc) determines the maximum number of bytes that can be transmitted without congesting the network.

Max number of bytes that can be sent = min(Wa,Wc)

76. TCP Congestion Control

Tahoe: **Slow Start, Congestion Avoidance, Congestion Control**

Slow Start: set Wc = 1 **MSS** (**Maximum segment Size in Byte**), if an ACK is received in one RTT, the TCP sender will double the size of his window. Wc = Wc +1 for **every ACK received.**

Congestion Avoidance: Wc = Wc +1 **for every RTT** (not every ACK, 1/Wc for every ACK), TCP sender sets **SS threshold**, when Wc reaches SS threshold, enters the CA phase.

Congestion Control: When a **Timeout** occurs or **3 dup ACKs** are received. TCP sender sets Wc = 1, sets a new SS threshold = Wc/2, and enters the SS again by retransmitting the segment that was timeout or 3 dup ACKs.

Reno: Doesn't take the Time-out & 3 dup Acks mechanism the same way.

If timeout: Reno behaves similar to Tahoe.

If receives 3 dup ACKs, TCP sender sets Wc = Wc/2 (ignore +3) (new SS threshold), and enters the **Fast Recovery Phase**.

Fast Recovery Phase: a dup ACK arrived Cw = Cw + 1; a new ACK arrived Cw = SS threshold, and enters CA.

77. TCPCC **AIMD** (**Additive Increase Multiplicative Decrease**)

No Slow Start, Saw Tooth Curve.

Additive Increase: Increase Wc by 1 MSS every RTT until loss detected.

Multiplicative Decrease: Cut Wc in half after loss detected.

78. Network Security

Confidentiality, Authentication, Message Integrity, Access and Availability.

Symmetric-key crypto: sender, receiver keys identical.

Trusted **key distribution center** (**KDC**).

Public-key crypto: encryption key public, decryption key private.

Trusted **certification authority** (**CA**).

Authentication using a **Nonce** (number R used only once in lifetime) to avoid playback attack.

Message Integrity: **Digital Signature** (not provide privacy). **Encrypt message digest.**

Firewalls: Packet-Filtering & **Application-Level Filtering.**