Introduction

Circuit-switched Require dedicated end-to-end connections with dedicated channel at all times.

Packet-switched Data is split into packets, which are exchanged link-by-link and assembled in endpoints; connectionless.

Inter-CPU dist.	CPU in same	Type of network	Examples
1 m	Sqr. meter	Personal Area Network (PAN)	Bluetooth
10 m	Room	Local Area Network (LAN)	Over switch network
100 m	Building		
1 km	Campus		
10 km	City	Metrop. Area Network (PAN)	Cable TV
100 km	Country	Wide Area Network (PAN)	ISP network
1000 km	Continent		
10 000 km	Planet	Internet	

OSI model

		Layer	Data unit	Function	Protocols
Host layers	7	Application	Data	High-level APIs (resource sharing, remote file access)	DHCP, DNS, FTP, HTTP, RIP
	6	Presentation		Data translation (encoding, compression, encrypt/decrypt)	
	5	Session		Managing communication sessions	
	4	Transport	Segment	Reliable segment transmission (segmentation, ACK)	TCP,UDP
Media layers	3	Network	Packet	Multi-node network (addressing, routing, traffic control)	IP, ICMP
	2	Data link	Frame	Reliable frame transmission between two nodes	ARP,MAC,Ethernet, PPP
	1	Physical	Bit, symb.	Communication of raw bits over physical medium	

Physical layer

Each communication media has its own transference function $H(f)$ which impacts the original signal $S(f)$ according to $R(f) = H(f)S(f)$.

Nyquist theorem A sampler operating at frequency f_s can completely reconstruct a signal with bandwidth B when $f_s > 2B$

Nyquist bitrate The theoretical maximum capacity of a noiseless channel with M signal levels is $C = 2B \log_2 M$

Transm.	Description
Baseband	Signal has frequencies from 0 Hz to B; wires
Passband	Signal uses a (usually small) band of frequencies $[f_1, f_2]$ around the carrier wave frequency f_c; wireless/optical

Baseband transmission codes

NRZ-L	Non-return-to-Zero level: two levels for 0/1
NRZ-I	Non-return-to-Zero inverted: lvl change = 1
Manchester	Used in Ethernet; $+ \to -$ is 1, $- \to +$ is 0

Clock recovery Bursts of same symbol may confuse the receiver. Manchester naturally solves it; other codings can use 4B/5B maps (4 bits are coded into 5 bits for transmission).

Modulations

$s(t)$ – Signal function to be transmitted

Amplitude	$s(t) = A_i \cos(2\pi f_c t)$
Phase	$s(t) = A \cos(\theta_i + 2\pi f_c t)$
Quadrature	$s(t) = A_i \cos(\theta_i + 2\pi f_c t)$; M-QAM (quadrature amplitude modulation), uses M symbols

Shannon's Law The max data transmission rate C over a channel in the presence of noise is

$$C = B \log_2 \left(1 + \frac{P_r}{N_0 B}\right)$$

Guided transmission

Cable	B (GHz)	Atten. (dB/km)	Delay (µs/km)
UTP	$0.1 - 0.6$	$20 - 250$	5
Coaxial	≈ 1	≈ 150	4
Fiber optics	$\approx 30\,000$	< 1	5

Gain and attenuation

- $P/\text{dB} = P/\text{dBW} = 10 \log_{10}(P/\text{W})$
- $P/\text{dBm} = 10 \log_{10}(P/\text{mW})$
- $Gain/\text{dB} = -Attenuation/\text{dB}$

Wireless transmission

VLF (very low frequency), LF, MF follow Earth's curvature, HF bounce off the ionosphere.

Free space loss Two ideal isotropic antennas distanced d,

$$\frac{P_t}{P_r} = \frac{(4\pi d)^2}{\lambda^2} = \frac{(4\pi f d)^2}{c^2}$$

Data link layer

Byte stuffing An escape character is chosen, and flags/escape chars are encoded to avoid ambiguity.

Error detection

Parity check One parity bit is added every k bits; simple but does not detect even number of errors.

Bidimensional parity Any 4 errors in rectangular configuration are undetectable.

Cyclic Redundancy Check (CRC) given encoding parameters $r \in \mathbb{N}$ and G a number with $r+1$ bits, encodes message M by adding the remainer R of $M \times 10^r$ divided by G (where $M \times x^r \equiv A \times G + R \pmod 2$ if everything is interpreted as a polynomial) to the end of M (R always has r bits).

Automatic Repeat Request (ARQ)

Stop&Wait Each received frame must be ACK, sender only sends next frame if all previous frames were acknowledged.

$$a = \frac{T_{prop}}{T_f} \qquad P[A = k] = p_e^{k-1}(1 - p_e)$$

$$\mathrm{E}[A] = \sum_{k=1}^{\infty} k \times P[A = k] = \frac{1}{1 - p_e}$$

$$S = \frac{T_f}{E[A](T_f + 2T_{prop})} = \frac{1}{E[A](1 + 2a)} = \frac{1 - p_e}{1 + 2a}$$

Go Back N Allows transmission before previous frames were ACK. It sends ACK(NR) meaning it acknowledges all packets with $index < NR$. If an out-of-sequence frame is received, REJ is sent with the expected frame number on the first time; subsequent out-of-sequence frames are silently rejected.

$$W = M - 1 = 2^k - 1$$

$$S = \begin{cases} \dfrac{1 - p_e}{1 + 2ap_e} & : W \geq 1 + 2a \\ \dfrac{W(1 - p_e)}{(1 + 2a)(1 - p_e + Wp_e)} & : W < 1 + 2a \end{cases}$$

Selective Repeat Similar to Go Back N, but receiver accepts out-of-sequence frames and notifies the sender to send missing frames only.

$$W = M/2 = 2^{k-1}$$

$$S = \begin{cases} 1 - p_e & : W \geq 1 + 2a \\ \dfrac{W(1 - p_e)}{1 + 2a} & : W < 1 + 2a \end{cases}$$

Reliability in TCP/IP reference model

PLR – Packet loss ratio; K – number of links between sender and receiver; assume all links have same C and PLR.

Strategy	Description
Link-by-Link $S = 1 - PLR$	On error, the station closest to the sender notifies it. Repairs losses link by link. More complex, but better on very unreliable media.
End-to-End $S = (1 - PLR)^K$	On error, the receiver notifies the sender. Less complex, but not acceptable on very unreliable media.

TCP/IP assumes Data Link layer provides error-free packets with possible packet loss (but very low FER). End-to-End is used in most cases, implemented in Transport or Application; in lossy channels (e.g. wireless) link-by-link is implemented in Data Link.

Delay models

Multiplex strategies	Description	T_{frame}
Statistical	Transmitted on first-come first-served basis	L/C
Freq. division (FDM)	Link capacity C divided into m channels, each with bandwidth W/m and capacity C/m	Lm/C
Time division (TDM)	Link capacity C divided into m channels in the time axis, each with capacity C/m	Lm/C

Little's theorem $N = \lambda \cdot T$, $N_W = \lambda \cdot T_W$. The time a client waits on queue T_W depends only on the # of clients in queue N_W and client arrival rate λ, but not on the service rate (!).

Delay modelled as queue networks

Poisson arrivals can be described by a Poisson distribution with $\mathrm{E}[A] = 1/\lambda$, $\mathrm{Var}[A] = 1/\lambda^2$.

Kendall notation: A/S/s/K (A – arrival stat. process; S – service stat. proc.; s – number of servers; K – system buffer capacity

Queues $P(n)$ – Prob. of Markov chain being in state n. $N = N_W + N_S$. $T = T_W + T_S$. Element being transmitted is the first in queue; queue is popped when that element is fully transmitted, meaning max queue time in M/M/1/B is $T_S(B - 1)$.

M/M/1	$P(n) = \rho^n(1 - \rho)$	$N = \dfrac{\rho}{1 - \rho}$	$T = \dfrac{1}{\mu - \lambda}$	$T_W = \dfrac{\rho}{\mu(1 - \rho)}$	$N_W = N - \rho$
D/D/1		$N = \rho$	$T = 1/\mu$	$T_W = 0$	$N_W = 0$
M/M/1/B	$P(0) = \dfrac{1 - \rho}{1 - \rho^{B+1}}$, $P(n) = \rho^n \cdot P(0)$	$\rho = 1 \implies P(B) = \dfrac{1}{B + 1}$	$\rho \gg 1 \implies P(B) = \dfrac{\rho - 1}{\rho} = \dfrac{\lambda - \mu}{\lambda}$		
M/G/1		$N = N_W + \rho$	$T = T_W + 1/\mu$	$T_W = \dfrac{\lambda \mathrm{E}[X^2]}{2(1 - \rho)}$	$N_W = \lambda T_W$
M/D/1	$\mathrm{E}[X] = 1/\mu$, $\mathrm{E}[X^2] = 1/\mu^2$	$N = N_W + \rho$	$T = T_W + 1/\mu$	$T_W = \dfrac{\rho}{2\mu(1 - \rho)}$	$N_W = \dfrac{\rho^2}{2(1 - \rho)}$

M (Markovian) – Poisson process/exponential service time; **D (Degenerate)** – Fixed inter-arrival interval/service time; **G (General)** – General process, arrival/service times independent & identically distributed, with given parameters $\mathrm{E}[X] = 1/\lambda$, $\mathrm{E}[X^2]$

Kleinroch independence approximation

Each link is a queue. Several streams merging on a link restore independence of interarrival times and packet lengths. Thus each link is M/M/1 if entry points have Poisson arrivals, packet lengths are nearly exponentially distributed, network is dense and has moderate/heavy traffic.

Given each path p on the network has a known arrival rate x_p, the arrival rate for link $i \leftrightarrow j$ is $\lambda_{ij} = \sum_{p:(i,j)\in p} x_p$; μ_{ij} is a known property of the link; all other things are obtained with M/M/1 equations. The total number N of packets in the system is the sum of all N_{ij}, $\lambda = \sum x_p$ and T is obtained from Little's theorem.

Jackson networks

Now queues are not links, but nodes. Arrivals and services are Poisson, system is stable and has no cloggings ($\rho_j < 1$). For each node j in a network with K nodes, there are r_j packets per second entering the system through j, and packets from other nodes; on leaving j, packets have prob P_{ji} of going to i, and $1 - \sum_{i=1}^K P_{ji}$ to leave system. Implies $\lambda_j = r_j + \sum_{i=1}^K \lambda_i P_{ij}$.

Jackson's theorem Let $\vec{n} = (n_1, n_2, ..., n_k)$ be a possible state; then $P(\vec{n}) = \prod_{j=1}^K P_j(n_j)$. Thus, all queues are independent; grab the Markov chain probs for M/M/1 queues, and we get $P(\vec{n}) = \prod_{j=1}^K \rho_j^{n_j}(1-\rho_j)$.

MAC sublayer

Random Access Protocols

Aloha Sends packet to the networks, waits for a round-trip time, and if no ACK arrived then delay the packet sending by a random time to 'try later'. Pure: $S = Ge^{-2G}$, $S_{max} = 1/2e = 18.4\,\%$; Slotted: $S = Ge^{-G}$, $S_{max} = 1/e = 36.8\,\%$.

CSMA Carrier Sense Multiple Access, first listens if there is traffic, only transmits if the channel is sensed free.

CSMA Pers. If busy, waits until medium becomes free and then transmits (persistent).

CSMA NonP. If busy, waits a random time and repeats (non-persistent).

CSMA p-Pe. If free, transmits with prob p, or defers transmission to next slot with prob $1-p$.

CSMA/CD Collision detection; similar to CSMA Persistent but if a collision is detected **during** transmission, it is aborted and retransmission is delayed using a Binary Exponential Backoff algorithm (if there were i consecutive collisions, transmit in a random slot picked from slot set $[0, 2^i - 1]$). Does not use ACKs. Slots have size $T_{slot} = 2T_{prop}$. $\lim_{N\to\infty} S = 1/(1+3.44a)$. $T_f >$ RTT ensures all stations see collision.

CSMA/CA Waits some DIFS time to check if no-one is transmitting, and then transmits; if someone is transmitting, then the current station waits DIFS plus a random backoff and tries again.

Taking turns

The switch polls stations and decides which station gets to transmit first; or stations pass around tokens sequentially.

Switches

Data Link device, forwards Ethernet frames, is transparent to hosts (hosts are unaware of the switch presence, as if hosts were directly connected). It has a forwarding table and is self-learning: when it receives a frame, it matches the incoming frame MAC address to the switch interface the frame arrived to. This differs from ARP, as ARP matches IP to MAC.

Network layer

Provides two services: datagram network (connectionless service), and virtual circuit (connection-oriented, router maintains states, has forwarding tables).

Description	Special IP address
This host	`0.0.0.0`
Broadcast local network	`255.255.255.255`
Broadcast distant network	`<network>1111...`
Loopback	`127.<anything>`

Internet Protocol (IP) Can segment large packets, tho TCP already segments packets so IP only segments oversized UDP/ICMP datagrams.

Address Resolution Protocol (ARP) IP to MAC (can also get IP from MAC). When not on ARP table, device broadcasts ARP request; only device with that IP answers with corresponding MAC.

Dynamic Host Configuration Protocol (DHCP) Dynamically get IP address from a server, allowing IP address reuse. Client broadcasts `DISCOVER` to find all DHCP servers; all DHCP servers `OFFER` an IP each; client accepts one of the offers by `REQUEST`'ing the offered IP from one of the DHCP servers; server `ACK`s.

Network Address Translation (NAT) Separates private network from public network. Requests are rewritten to bear the NAT public address if going outside, or the dereferenced private address if going inside.

Internet Control Message Protocol (ICMP) Error/control messages. Used e.g. by `ping`. ICMP packets are encapsulated in IPv4 packets.

IPv6 128 bit addresses. **Link-local:** used for comms in same LAN/link, not forwarded by routers. **Global unicast:** global address. **Anycast:** packet is received by any (only 1) member of group. **Multicast:** packet received by all group members. **Neighbor Discovery protocol (ND):** replaces ARP IPv4, ICMP router discovery, ICMP redirect.

Transport

Pr.	Description	Used by
UDP	Unreliable, connectionless, direct interface to IP with minimal protocol overhead	DNS, SNMP, DHCP
TCP	Reliable (uses ARQ mechanism), connection-oriented, full-duplex, avoids receiver/network congestion	FTP, HTML

TCP uses a system similar to **Go Back N**, except when both ends support selective acknowledgement (SACK), in which case a variation of Selective Repeat is used. Connection is established using **three-way handshake** (SYN, SYN/ACK, ACK). Data is interpreted as a **byte stream**, and packets are numbered by the sequence number of the first byte of data the packet carries. TCP connection is **full-duplex**, and each direction has different sequence numbers.

Retransmissions

Adaptative ret. RTT $= a \cdot$ RTT $+ (1-a)$RTT$_{\text{sample}}$, Timeout $=$ 2RTT

Karn-Partridge alg. RTT not updated on retransmissions, only updated with unambiguous ack's (ack's for segments that were only sent once). If there is a sharp increase in RTT, a new method is used: if a timeout occurs and causes a retransmission, timeout is doubled. This avoids the possibility that TCP does not update RTT because there are only retransmissions, and as such would block into a very small RTT.

Flow control

The receiver regularly broadcasts its receiver window `RWND`. Congestion window `CWND` is only kept/updated by the sender, and is used to implement congestion control (maximize efficiency, fairly distribute bandwidth).

Additive Inc./Multip. Dec. For each RTT, increment `CWND`; on timeout, divide `CWND` by 2. Sawtooth behavior.

Slow Start Start with small `CWND`, and double it for each RTT; when segment is lost (detected by timeout), a threshold is defined at half `CWND`, then `CWND` is reset, increases exponentially until it reaches threshold, and then increases linearly (congestion avoidance phase)

Routing

Forwarding is to just use a forwarding table, and redirect a request to a certain node; routing is about computing the whole path of a request, and send it to the best node possible.

Shortest-path routing Destination-based, load-insensitive, uses minimum # of hops or sum of link weights.

Detecting topology changes Usually a periodic `hello` message is used in both directions.

Link-state routing

Station u broadcasts a list of its links (the LSA – link-state advertisement) to its neighbors. When a receiving station gets a more recent version of u's LSA it re-broadcasts to its neighbors; thus the network is flooded with u's LSA. When all stations know all links, each performs Dijkstra to create the routing table.

Distance vector

Station u keeps a table of least distance to all nodes v and the link that should be used, $D(u,v), \forall v$. u periodically informs its neighbors of its distance information $D(u,v), \forall v$; and each of its neighbors v update their own distance vectors $D(v,w)$ by checking if passing through u to reach w is better (i.e., if $D(v,u) + D(u,w) < D(v,w)$). Initially all distances are assumed infinite (except each station's neighbors), and the distance vector will eventually converge. Distributed version of Bellman-Ford, it is iterative and asynchronous (triggered by receiving distance vector).

Routing Information Protocol (RIP) Distance vectors sent every 30 s or when link cost changes. Does not scale well to large networks.

Border Gateway Protocol (BGP) Also known as Exterior Gateway Routing Protocol (eBGP), usually for high-level networks in hierarchy (e.g. inside an ISP, or between ISPs).

Ethernet

Requires the network to be modelled as a tree, otherwise frames could loop forever. Uses a minimum spanning tree algorithm to build said tree, rooted in smallest ID switch.

Flow network model

Networks are can be modelled as a flow network; packet networks are queue networks, as such this approach can only be used to estimate maximum data flow from one station to another.

Variables

C	[bit/s]	Max. theoretical channel capacity
B	[Hz]	Channel bandwidth
f_s	[symbol/s], [baud]	Sampler frequency; baudrate ($f_s = 2B$)
P_r	[W]	Signal power seen by receiver
N_0	[W/Hz]	Spectral density of white noise power
λ	[pac/s], [bit/s]	Client arrival rate
μ	[pac/s], [bit/s]	Service rate
ρ	[1]	Traffic intensity; $\rho = \lambda/\mu$
T_{prop}	[s]	Propagation time from send. to rec.
T_f	[s]	Time to transfer a frame
$P[A=k]$	[1]	Prob. of frame requiring k attempts
$E[A]$	[1]	Expected # of attempts
a	[1]	Ratio of T_{prop} to T_f
p_e	[1]	Frame error probability
S	[1]	Efficiency
W	[1]	Maximum window size
M	[1]	Modulo of sequence numbers
N	[1]	Avg client # in the system
N_W	[1]	# of clients in queue
N_S	[1]	# of clients being served
T	[s]	Avg delay experienced by client