Computer Networks

Networks intro

- Definition of Computer Networks
  - Components:

Distributed systems (applications)
Networks (messages)
Communications (bits)

  - Example Networks:
    - car key with car; sensors network withtheir controllers (either one-way or two-ways)
- Design principle:
  - Dumb network & Smart users
    - networks don't store too much info but just pass the info
- Internet 'preferred' protocol stack
  - Application
    - delivery functionality
  - Transport
    - ensure end-to-end performance
  - Network
    - send packet over multiple links
  - Physical & Link
    - transmit frames
- Messages in Layers
  - Overview

Layer	What it transports (Protocol Data Unit)	How they connect
Application	Messages/Data	Proxy, gateway
Transport	Segments/Datagrams	
Network	Packets (!!)	Router
Link	Frames (cells, circuits)	Switch, Bridge
Physical	Bits	Hub (repeater)

  - Messages Encapsulation on Each Layers

- e.g. Ethernet frame's payload contains IP packet(s)

Information Transmission - Communication

- Problems:
  - Attenuation - loss of energy
  - Noise - gain of energy
  - Delay distortion - smearing
  - Frequency cut-offs - loss of information
  - Frequency-specific attenuation
- Approaches:
  - Circuit switching
    - Communication priented
    - Pros:

      hardware level guarantee fixed (reliable) quality during the communication
    - Cons:

      lots of waste of capacity (no sharing) & explicity resouce allocation $\Rightarrow$ expensive to scale

      networks need to store state (info of connection) $\Rightarrow$ multiple single-points of failure
  - Multiplexing
    - Spatial division multiplexing - (more wires)
    - Time division multiplexing - (take turns)
    - Frequency, Amplitude, Phase multiplexing
- Analog vs. Digital
  - Digital
    - easy to represent, store and regenerate
  - Analogue
    - represent the natural world
    - Sine wave appears every where $\Rightarrow$ Fourier transformation
  - $\Rightarrow$ Measureing & Creating Sine wave
    - encoding the feature of sine wave

      e.g. frequency, amplitude and phase
    - use Sine wave as carrier

      (especially in wireless communication, yet constant voltage is easier in wired transmission)
- Encoding of Bits into Signal
  - Modulation & Demodulation
    - Modulation: turning bits into signals

- Demodulation: turning signals into bits

- Single Bit Encoding in Modulation
  - Amplitude modulation (AM)
  - Frequency modulation (FM)
  - Phase modulation (PM)

    (detect phase shift: need sync ⇒ clock line can be represented by freq or in other forms)

- Symbol Encoding in Modulation
  - Symbol: bit pattern

    ⇒ 1 symbol / second > 1 bit / second

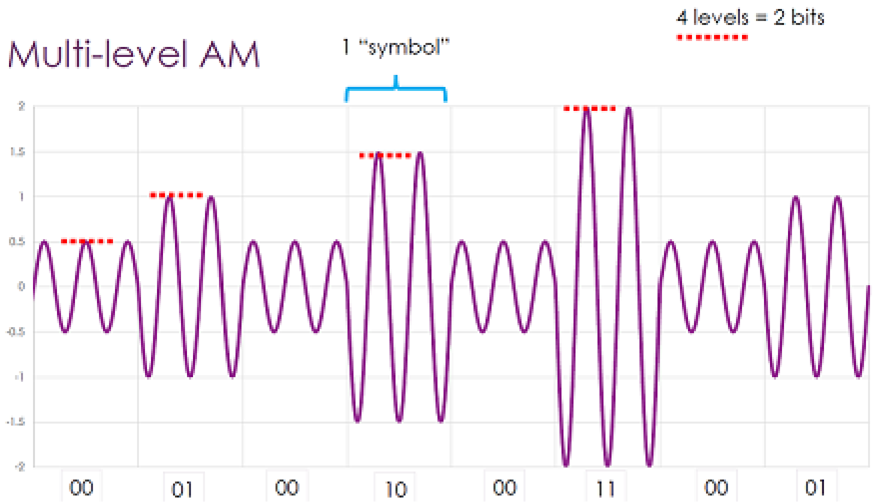
  - Multi-level modulation

    Multi-level AM = Amplitude Shift Keying (ASK)

    Multi-level FM = Frequency Shift Keying (FSK)

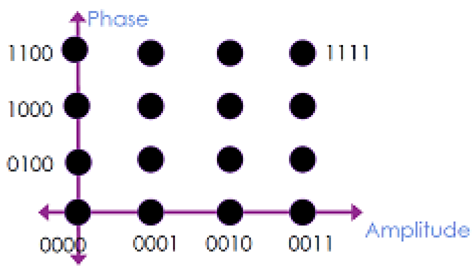
    Multi-level PM = Phase Shift Keying (PSK)

    e.g.:

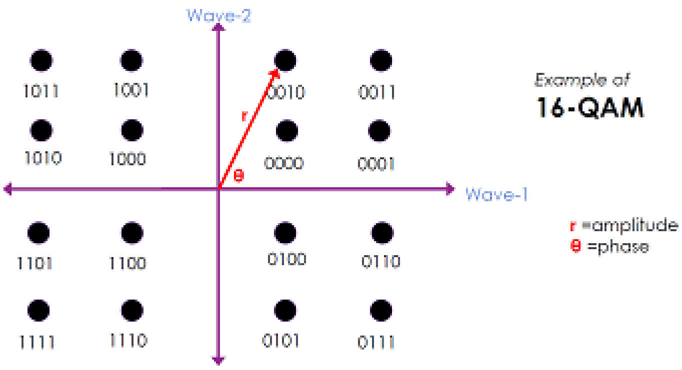


  - Phase + Amplitude Modulation:

    ⇒ Constellation diagram:



    ⇒ Quadrature Amplitude Modulation (QAM):



Other Modulation:

256-QAM CableTV system

4096-QAM Powerline data

65535-QAM ADSL

x-QAM depnding on the needs and techniques available

Note:

phase-amplitute-frequency modulation (on 3 axis) not commonly used

Because: frequency usually used to denote channel (using carriers)

⇒ frequency to avoid interference & harder to change

- Bands
  - Baseband: constant voltage
    1. Baseband signal: lowpass signal, using <u>constant voltage as carrier</u>

       ⇒ non-modulated signal

       ⇒ only non-zero near the origin of frequency spectrum

       e.g. ASK, OOK (On-off keying)
    2. Baseband channel: lowpass channel, typically an unfiltered wire
    3. Baseband transimission: <u>transferring bit steam in line coding on typically an unfiltered wire</u>
  - Passband: the range of frequencies that can pass through a filter
    1. Passband signal: use <u>single frequency as carrier</u>

       ⇒ a signal with energy only in a passband, up-converted to higher frequency

       ⇒ digital modulation employed

       ⇒ integrate low-frequency wave (info wave) into a higher-frequency carrier wave
    2. Passband channel: channel of range of frequency after bandpass filters employed
    3. Passband transmission: (carrier-modulated transmission)

       <u>using passband signal to transfer info, typically in wireless transmission</u>
  - Broadband:
    1. Broadband signal: <u>use multiple frequency carries across a range</u>

       ⇒ FSK

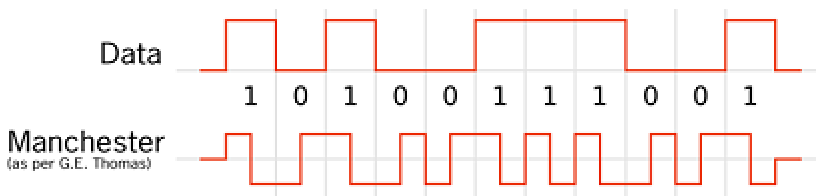
        - Bandwidth: a specific range of frequencies

          can be divided at your choice & capacity of the technology allowed
- Limitation in Transmission Quality:
    - Shannon "Capacity Limit"
        - lowest sampling frequency of twice as the imcoming signal to get a perfect reconstruction
    - Expressing Transmission Quality
        - Signal:Noise Rate (SNR) = Signal Energy : Noise Energy

          $\Rightarrow$ SNR in deciBel = $10 * \log_{10}(\text{Signal}/\text{Noise})$ dB
- Encoding of Bits Sequency into Bits Patterns (regardless of modulation)
    - Key Concepts:
        - Map bits into patterns to reduce repetition
        - Signal each pattern with a transition
    - Bits Pattern Example: 4b/5b Code:
        - Mapping Table:

Given	Send	Given	Send
0000	**11110**	0100	**01010**
0001	**01001**	1000	**10010**
0010	**10100**	1101	**11011**
0011	**10101**	1111	**11101**

        - Features:
            1. avoid runs of 0, but can have maximal 6 1's in a row...
            2. trade bandwidth for reliability $\Rightarrow$ enable self-checking
    - Transition Example: Manchester Code:
        - Example:



        - self-clocking:
            1. a sync pattern in the front to denote the start (sync)
            2. the receiver can then identify if it is misaligned by half a bit period (prevent phase shift)

Physical Layer

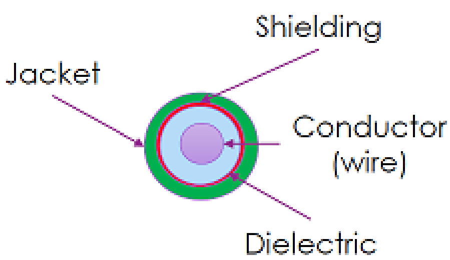
Copper

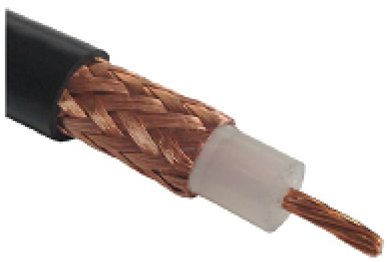
- Characteristics
    - Physical:
        - Soft & bendable around the corner
        - Light; Malleable; Easy to make thin wire

        - Easy to add insulation & preotection; Reasonably robust to oxidation
    - Social:
        - Cheap, compared to fibber (yet price is increasing)
    - Electrical
        - Shared medium (one voltage over the whole line)
        - Receive (RX) & Transmit (TX) on the wire:
            1. Half-duplex - each side takes turns to transmit & receive - Time Division Multiplexing
            2. Full-duplex - both ends can transmit & receive in parallel - Frequency Division Multiplexing
        - Reference of 'zero' ⇒ cables tend to have a pair of wires
        - Resistance:
            1. impedance, inductance (hate frequency change), etc...
            2. frequency related resistance: skin effect

               ⇒ in alternating current, higher frequency, higher resistance, more current close to skin

               ⇒ frequency attenuation
            3. Varies from cross-section: thinner wires, bigger resistance
        - Attenuation:
            1. loss of energy (in the form of heat, light and etc.)
            2. loss of frequency and etc...
- Noise in Signalling:
    - Random Wire Antenna: straight wires on the ground as an antenna
        - receiver for other signals
        - transmitter of its own signals
        - Electro-magnetic Interference （EMI, 电磁干扰） & Radio-frequency interference (RFI)
        - Coupling with adjacent wires ⇒ cross talk (expecially at near & far end - NEXT&FEXT)
    - Solving Antenna Problem
        - Pretection - "Coaxical" cables

          Pros: well sheilded - protection from noise & security (much less sending out), robust

          Cons: single RX/TX, expensive
        - Spatial Division Multiplexing ⇒ more wire in a cable

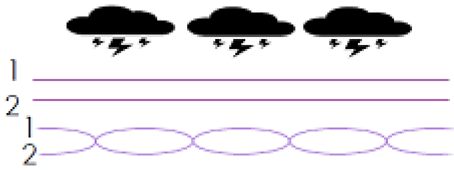
          Pros: full duplex, inverse multiplexing - multiple path to share (one-to-many & many-to-one)

          Cons:

          too many adjacent wires ⇒ cross talk

long straight unsheilded wires ⇒ antennas problems remains

        - Differential signalling
        - Twisting wires



Assumption: noise source has a direction

⇒ twisting to make sure noise added evenly

⇒ use the reference line to record the noise and then fileter it out

Example: UTP (unshielded twisted paris - 网线)

⇒ combin with shielding: STP/FTP (shielded twisted / foiled twisted pair)

    - Skew between Pairs:
        - Different lenghts between multiple pairs can result in un-aligned signal

          ⇒ affects inverse multiplexing

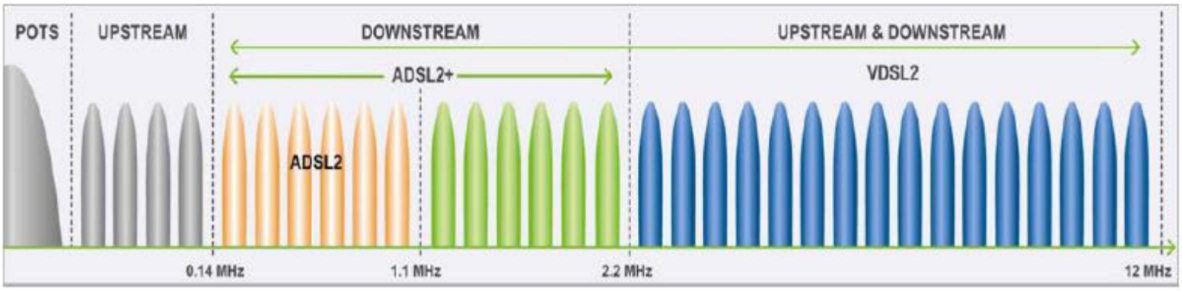
          ⇒ have to be in the same length within tolerance
    - Resistance (Inpedance) Mismatch
        - Results in signal bouncing back to the sender
- Transmission on Copper
    - Speed:
        - kHz to MHz, enhanced by different keying & multiplexing technology
    - Distance:
        - Low data rate (< 1Mb/s) for longer distances (km)
        - High data rate (~100Mbs) for short distances (~500's m) E.g. DSL+
    - Downside:
        - Propagation delay (speed of electricity in copper = ~3us / km)

          ⇒ collision of two sender signalling at the same time
- Costs
    - Deployment
    - Protect Damage
        - easy to have a shared backbone
        - last mile exposed in the real world - insects, weathers, stealing, etc...
    - Last Mile Trad-off (last mile also refered as local loops sometimes)
        - cost of exchanges, distance for the final cable, quality of signal throught the wire

          Note: up tp 4+ km from their exchanges
        - scalability
- Existing Last Mile Technology

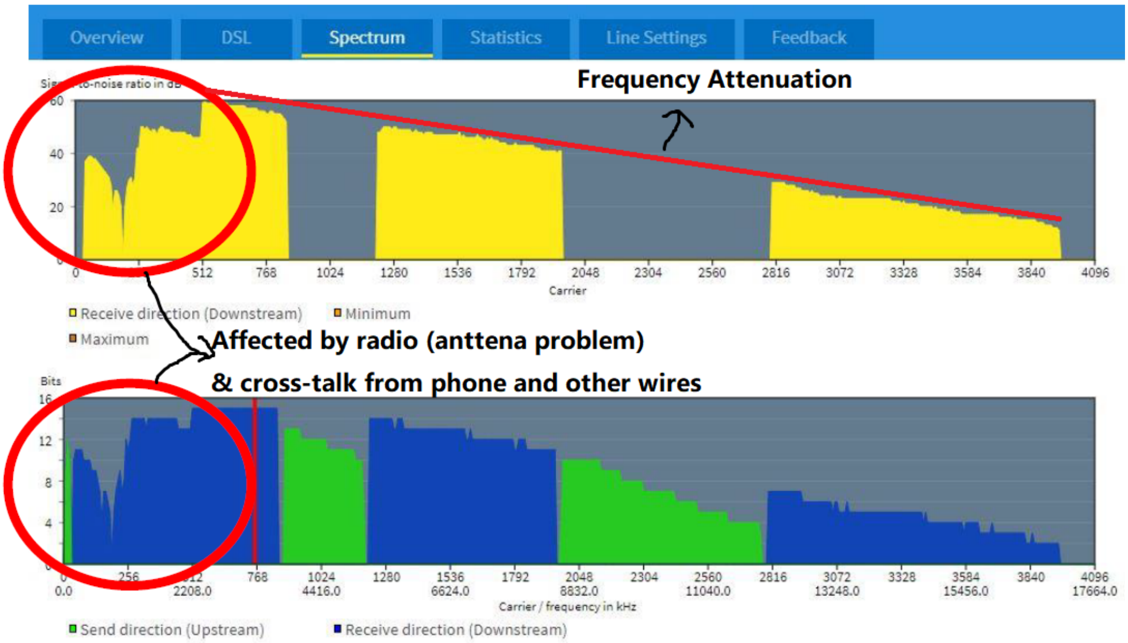
- DSL - based on existing telephone wires
  - evolving from ADSL to VDSL etc...

    later, DSL+: 16-bits (65535) QAM, FDM, ...
  - Assymetric: more on downstream performance



  - Pros: using the existing telephone line; co-exit with POTS (plain old telephone service)
  - Cons: limited performance; performance decreases over the distance

    $\Rightarrow$ may deploy more DSLAM to make average distance shorter
  - DSL Example:
    1. Computer -> modem (add info onto carrier)
    2. modem (s) -> DSLAM (aggregate signal from modems)
    3. DSLAM -> switch (decide which LAN it is in)
    4. switch -> router / switch (go to outer internet / transfer to another switch)
- NBN - National
  - Spectrum in Real World



  - Mixed-Technology
    1. NBN FTTx (Fiber to the x)
    2. Hybrid Fiber Coax and etc...

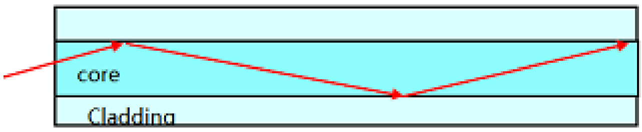
Fiber

- Characteristics

- Physical:
  - Light weight, very robust to oxidation and water
  - Easy to make thin cable
  - Fragile when twisting & bending, hard to connect (need to melt it)
  - Good at distance (several km is trivial)
- Social:
  - Expensive, compared to copper (yet price is decreasing)

    Note: fiber itself is okay, yet the end-point is expensive (⇒ usually use FTTx)
- Electrical
  - Robust to electircal interference
  - High throughput: much higher freqeuncy (light) signal - start at THz

* Noise in Signalling:
  - Oblique Light Leaks
    - use another layer of glasses to reflec the light (with in a 'critical angle'), wrapped with jacket



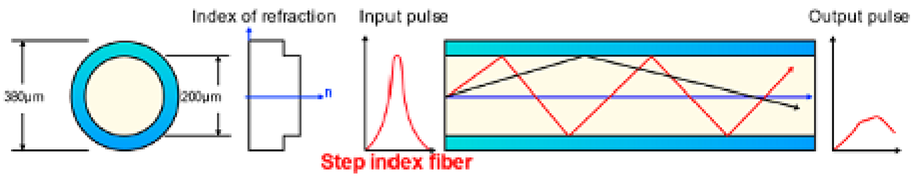
  - Modal Distortion (varying distance for light to travel because of reflection)
    - use graded-index (缓变折射率) fiber instead of step-index (阶跃折射率) fiber

      ⇒ different kind of glass at different layer so that...

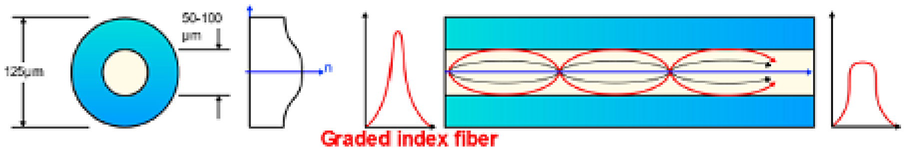
      ⇒ speed up the light bouncing in the fiber & slow down the light going straight

      ⇒ receiver can line up the light more easily

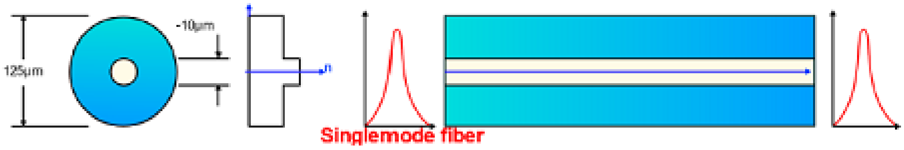
    - Multi-mode vs. Single-mode (each ray = a 'mode')
      1. multimode fiber (MMF) step-index: more bandwidth, significant modal distortion



      2. multimode fiber (MMF) graded-index: a few bandwidth, less modal distortion



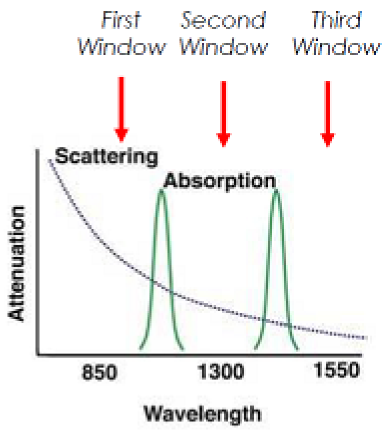
3. singlemode fiber (SMF): less bandwidth, good at travelling on long distance



Note: from 1. $\rightarrow$ 3. the performance increases, so does cost

does not mix up different cables $\Rightarrow$ performance suffers

- Fiber Connectors
  - Factors: dust; reflection at the end point
  - Solution: use curved faces at end point $\Rightarrow$ focusing the light on one point
- Attenuation
  - Scattering: structures + materials in the fiber
  - Absorption: materials in fiber
  - Can be frequency dependent:



- Chromatic dispersion (色散):
  - Factors: refraction index varies with wave length; hard to have a pure single wavelength lazer
  - Solution: Soliton pulses
- Polarization mode dispersion (偏振模色散)
  - Core shape helps

* Setting up Fiber
  - Multi-core Cable Design
    - Factors
      1. individual fibers are fragile $\Rightarrow$ cable bundles up to 1024
      2. costs the same to deploy one or a bundle of fibers
      3. people wants their own cable for security..., though one fiber can carry whole internet
* Transmission over Light
  - Electronic Data to Optical Signal

        - Keying in optical signal:

          E.g. OOK (on-off keying), QPDM (quatrature polarization division multiplexing)

        - Pulse an LED: cheap, yet broad wavelength range, no nice pulse ⇒ though used in MMF
        - chop a laser: can be small & at a high rate
            1. Cut the light using thin pins ⇒ noise at the edge of square wave (pysical effect)
            2. use the inveretd wave to cancel out the info wave

          ⇒ wavelength tunable on the fly ⇒ used in SMF

    - Speed
        - starts at THz, able to carry the whole internet traffic in on fiber (device can't catch up)
    - Distance
        - Normally...MMF: 1-2 km; SMF: 50-100 km
        - Regenerate/repeat: every 50-100 km

          using expensive optics & electronics for OEO interfaces (optical-electronical-optical)

        - Amplify: every 50-100 km

          using cheap electronics OR optics ⇒ amplifies both signal AND noise

    - Downside:
        - Not easily a shared medium ⇒ point-to-point

          ⇒ crosstalk still exists when sharing, in connectors and within fiber

        - Can use one fiber for RX & TX

          ⇒ need optical splitters at both ends; crosstalk effects

- Last Mile with Fiber
    - Costs
        - existing copper vs. deployment fiber
        - deployment copper vs. deployment fiber
        - maintain copper vs. maintain fiber
    - (G)PON Passive Optical Network
        - Technology used in the fiber part of the backbone, just before the last mile
        - Comparison with active network
            1. traffic from backbone splitted by splitters into cabinet depending on their destinations
            2. cabinet starts the last mile, sending only your info to you
        - Passive network
            1. traffic from backbone not splitted, sends all traffic on this fiber to all ends of this fiber (BFS)
            2. use TDM (time division multiplexing) at all fiber ends, to check if this piece of info is yours
            3. potentially RX&TX on the same fiber, using WDM (wavelength division multiplexing)
            4. security may suffer, yet business gains
    - General Approaches:
        - Push fiber as near as can afford / achieve

        - FTTx (Fiber to the x)

          ⇒ FTTP/B/C/N: Fiber to the Premises/ Building/ Curb/ Node

          ⇒ fiber node -> FDU (fiber distribution unit) -> copper cable into house

          Note: the position of FDU differes between 'x'

        - Combine with copper

          1. using DSL
          2. HFC (hybrid fiber coax): share coax copper

             Note: though coax affords 10 Gb/s, yet is shared

             ⇒ fiber node -> trunk coax -> trunk amplifier -> cable into house

             ⇒ peak speed might influenced

Wireless Communication

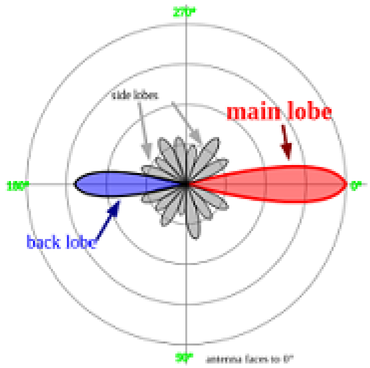
- Characteristics
    - Distance
        - can go a very, very long way (satellite transmission)
    - Electronic
        - sensitive to atmospheric conditions and EM interference
    - Unguided transmission
        - on a broadcast & shared medium (free space)
- Noise in Signalling
    - Absorption
        - Gases, dusts
        - Structure & terrains
    - Reflection, Refraction (折射) & Diffraction (衍射)
        - Temperature difference
        - Turbulance
        - Structure and terrains

          ⇒ causes multipath reception (multiple delayed refelcted waves interferes)

        - Even varies with time and different wavelengths
    - Noise
        - Extraneous signal in the free space
- Transmission in Wireless: Improvements
    - Transmiter & Receiver ⇒ Antennas
        - Omnidirectional (Broadcasting) antenna

          ⇒ broadcasting to all direction, yet poor coverage for directly under the antenna

        - Directional antenna ⇒ more focus



Note: generally, $O[n]$ in size, with $n = \text{wavelength}$

- Clearer Signals
  - More power - shout louder
  - Decrease bitrate - slow down
- Smarter to Deal with Environment
  - Frequency hopping ($\Leftrightarrow$ channel changing)
    1. detect traffic jam (lost / wrong messages)
    2. ask for re-allocation & try re-association (either actively or after the connection is lost)
    3. re-enter the session with the AP (access point), using the same credential info

       (Note: hopefully the APs reserve the same IP and session for a while)
  - Beam-shaping (directional antennas)
- Select the Right Wavelength (Frequency) & Power
  - Long wavelength (low frequency):
    1. Go around corner, through walls and waters $\Rightarrow$ long distance & through obstacles
    2. Low data rate as a trad-off
  - Short wavelength (high frequency)
    1. high data rate
    2. need line of sight (easily blocked)

  $\Rightarrow$ Consider requirement of point-to-point vs. area coverage; obstacles; effective distance
- Use the Right (Allowed) Spectrum
  - Spectrum allocation sets the rule of using the shared free space

    (some are reserved for special use, e.g. military use)
  - Channel allocation with each spectrum

    E.g. FM radio (85-108 MHz) in Canberra has 0.8 MHz channel spacing
- Covering Large Area with Wireless
  - Repeaters
  - Mixed with lined networks (link wireless to wired)
  - Coverage type
    1. fixed vs. mobile client $\Rightarrow$ directional vs. broadcast
    2. point-to-point vs. cell coverage

        - APs networks (mobile + cell coverage) ⇒ cell handovers
            1. negotiate with current APs to re-association (while the connection is still okay)
            2. APs aware the re-association - keep the same IP & session
            3. enter the same session with credential info
        - Spave wireless
            1. Forms: satellite to satellite; satellite to/from ground
            2. Handing over needed: satellite orbits
            3. Potentially high delay: long distance
        - Wireless between earth and space (e.g. Google balloon)
            1. Pros: stable-ish location with greater coverage
            2. Cons: power & maintaining
        - Longe range wireless: MIMO (multiple input multiple output, for 5G), MUSA-MIMO

Link Layer

- Focus & Role
    - Message - Frame
        - various length:
            1. length specified in the frame
            2. start & end of the content denoted
        - targeted messages:
            1. destination address
            2. source address

LANs

- Definition:
    - LAN: local area network
    - WAN: wide area network
    - PAN: personal area network

  ⇒ Start of any kind comminication
- Design Principles:
    - Simple
        - no guaranteed message delivery, correction or other specialized fedtures (real-time or etc.)

          ⇒ left to the software
        - focusing on transmitting one message from A to B
    - Efficient
        - multiplexing
- Multiplex in LAN
    - Fair Multiplexing vs. Statistical Multiplexing
        - not everyone talking at the same time
        - no one always spamming - don't need all the bandwidth or all the time

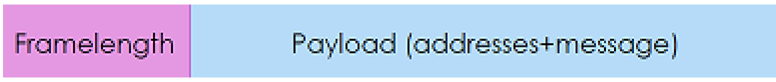
        - demand on capacity varies with time

    ⇒ Statistical multiplexing reduces capacity waste

    (more time / channels / wires for the current users)

  - Fair Access to Network
    - rules for trying to send

- Example Desings
  - Simple Frame: need to be in synchronization



  - Frame with Flag: need an escape symbol to distinguish (e.g. the "\" to denote "\n" in C)

- MAC (Media Access Control) & Sharing
  - Address Scheme
    - hardwired to the network **interface**
  - Access Scheme - Randomized Access on Shared Media
    - send and then detect
      1. send the frame
      2. detect collision on the wire
      3. wait for acknowledgement
      4. on collision or no acknowledgement ⇒ back-off for a random time & re-send

      Pros: simple, effective in low traffic networks

      Cons: actual performance depends on back-off scheme, not scalable
    - Carrier-Sense Multiple Access / Collection Detection (CSMA/CD)
      1. sense for carrier till no collision
      2. send frame
      3. detect potential collision - because transmit on wires takes time

         ⇒ upper limit time for any potential collision to occur (bounded by wire length)

         ⇒ can have a minimum frame size (need to wait for collision detection anyway)
      4. back-off for a random time in collision detected

      Pros: good for wired network

      Cons: not working in wireless
    - Carrier-Sense Multiple Access / Collection Avoidance (CSMA/CA)
      1. sense for carrier till no collision
      2. wait for a random time ⇒ reduce the possibility of sending frame at the same time
      3. send frame
      4. detect collision & re-try on detected

      Pros: better for wired network as wait before send

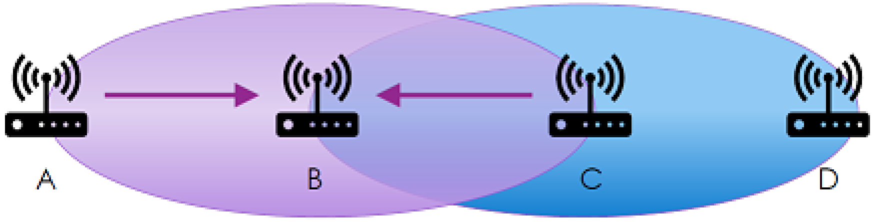
      Cons: not working in wireless either

- Back-off Scheme
  - Limitation: not too short & not too long
  - Ideal back-off time: depends on the number of devices in the LAN
  - Approximating the ideal time: Binary Exponential Back-off (BEB):
    1. counting the detected collision in a relatively recent history
    2. for the n^{th} collision, wait for a random number between $[0, 2^n - 1]$
- Access Scheme - Wireless
  - Problem of wireless environment
    1. cannot detect the whole network from a corner

       $\Rightarrow$ because of limited coverage of each cell

       $\Rightarrow$ different Tx can transmit to one Rx with out noticing interference

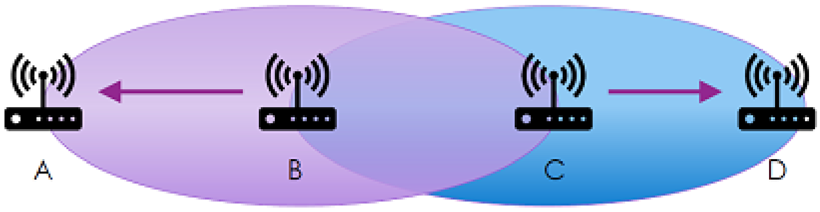
       $\Rightarrow$ hidden terminals: A, C are hidden from each other and can talk to B at the same time

    2. local Tx (e.g. its own Tx) are much louder than remote Tx

       $\Rightarrow$ detect fake collision, thus wasting bandwidth

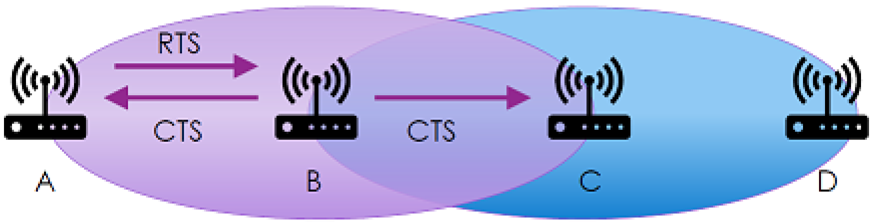
       $\Rightarrow$ exposed terminal: C detects collision because of B talking to A

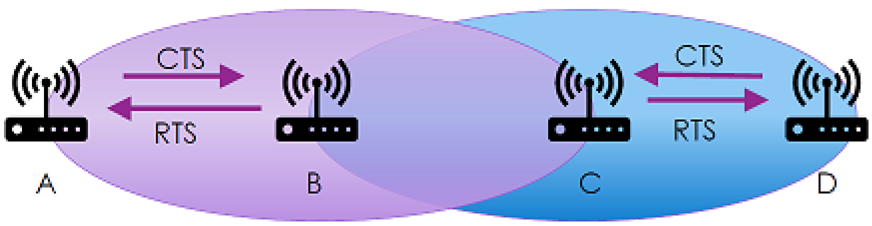
  - Multiple Access Collision Avoidance (MACA) - handsahe before yelling
    1. sender: request to send (RTS), providing the frame length N
    2. anyone hears RTS stay silent for receiver's CTS
    3. receiver: clear to send (CTS), providing frame length N
    4. sender transmits the frame & everyone hears CTS stay silent for N

    Pros: now, the receiver decides the collision instead of the sender itself

    $\Rightarrow$ fixing hidden terminals problem: C knows A is sending after CTS

    $\Rightarrow$ fixing exposed terminal problem: B, C not influenced by others' CTS



- Access Scheme - Contention-free access
  - Token rings
    1. generates tokens rings (special frame)
    2. pass the token along the rings, under the path-selecting scheme
    3. only talk when token at hands

    Pros: time multiplexing $\Rightarrow$ guaranteed no contention

    Cons: token may lost & hard to detect and re-generate & not adaptive to topologies change

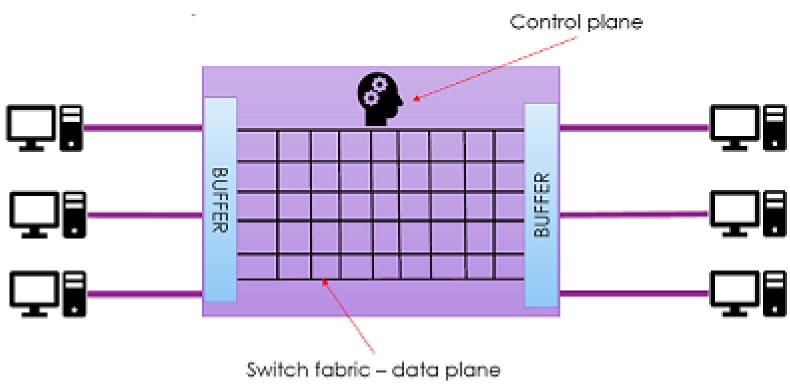
    $\Rightarrow$ fragile to error & not scalable

- Topologies
  - Bus topologies
    - needs repeater if too long
    - too much collision if many devices

    $\Rightarrow$ does NOT scale
  - Switch
    - a device sitting in the center to learn the source / destination addresses from traffic
    - makes every link point-to-point: source -> switch -> destination

    $\Rightarrow$ more scalable

    yet, people may employ policy on switch (slowing down the traffic)
- Different LANs design
  - General LAN (customer level)
    - bluetooth, 4G, Ethernet standards ...
  - Carrier-grade LAN (service level - guaranteed performance)
    - ATM (Asynchronous Transfer Mode), GPON (Gigabit-capable Passive Optical Networks), ...
  - Dat-center LAN (specific for high volumn, short distance)
    - FibberChannel, ...
- Switches



    - Learn the Address on the Air
        - Recording all source address of incoming message
        - New / Unknown address:
            1. broadcast to find the address & record the address $\Rightarrow$ may suffer if cyclic
            2. hope that address show up (send incoming message)
    - Cyclic Swtiches Hierarchy
        - Reasons:
            1. spaicial multiplexing - more wires
            2. redundancy
            3. short cuts
        - Broadcasting storm: with no global view, leads to recursive broadcasting
        - Spanning tree: disable some path $\Rightarrow$ reduce to tree architecture
            1. everyone think itself as root
            2. broadcasting & forward its current info to select a root on set-up (flooding)
            3. select the shortest path from root - using hop count
            4. turn off ports not on the tree

          Compared to flooding: maintain the reduced topologies instead of the whole map
- Virtual Lan
    - Reasons
        - Separation of traffic : logically separated network on the same infrastructure
            1. protect confidential info
            2. ensure devices in communication are compatible (computer cannot talk to phone)
            3. easy re-configure the LAN Structure on the
        - Prioritization of traffic
            1. drop frames accordingly when busy
    - Implementation
        - Tagging the port address into groups
        - Tagging the frames accordingly

LAN - Ethernet

- Advantages:
    - scalablity:
        - plug-and-play
    - backward compatbility
        - negotiate on connection
- Auto-Negotiation
    - Capability Negotiation
        - both ends communicate in "fast link pulse", containing requirement of:
            1. Speed

2. Duplex
3. Rx & Tx Detection

- Topologies Change - devices connect / disconnect
  - Heatbeat: device sends out a "normal link pulse" to remind the network of me

- Ethernet Frame

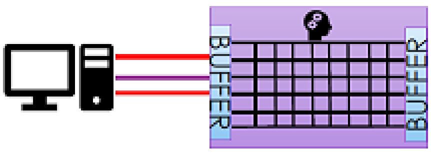
Preamble	Start of Frame	MAC dest	MAC src	802.1Q tag [opt]	Type / Length	Payload	Checksum
7 byte	1 byte	6 byte	6 byte	4 byte	2 byte	42-1500 byte	4 byte

  - Preamble:
    - 1-0 bits sequence
    - wake up the receiver & synchronize
  - MAC Addressing
    - originally plan to offer globally unique address
    - some address for sepcial use, e.g. "all ones" for broadcasting
    - have special bit for: multi-cast ⇒ send / receive messages from a group
  - Tag:
    - virtual lan tags
  - Type / Length:
    - different types of frame denoting the purpose

- Bigger Frames
  - Overhead of Ethernet Frame
    - ~30 bytes meta-data / 1500 bytes data ⇒ 3~5% bandwidth lost
    - more bandwidth ⇒ more frames ⇒ more read/write ⇒ traffic jam
  - Jumbo Frames
    - 9000 bytes payload

- Protocol:
  - Listening all frameson the wire until destination is my address
  - Can collect all frames transfering on the wire

- Link Aggregation / Trunking



  - Advantages:
    - performance
    - redundancy
  - Restriction:
    - need to use identical link for each port
    - frames order not changed
    - no partial frame (independent interfaces / devices / netwrok cards at other ends)

- Protocol:
  - on set-up, checking if using aggregation
- Model for Aggregation: selecting the path for frames
  - Round-robin: using each path in turns
  - Active back-up: use one path till broken
  - Random: randomly choose

LAN - Wireless LAN: WLAN

- Interference
  - Dealing noise
    - Adapt power: shout louder
    - Adapt rate: slow down - e.g. 1b/10b (encoding 1 bit into 10 bits)
  - Statistical Multiplexing and Frequency Hopping (⇔ channel changing)
    - choose the frequency to change to using statistical random method
  - Beam-Forming and Spaical Multiplexing
    - multiple input multiple output (MIMO) ⇒ multiple antennas for beam-forming
- Channels
  - 2.4GHz
    - most channels overlapped
  - 5GHz
    - larger spectrum space
    - channels does NOT overlap
    - can bind channels into a wider channel ⇒ higher bandwidth for each channel
    - built-in freqeucy hopping in the standard
- Frames in WLAN

X Mb/s *Y Mb/s*

Preamble	Start Frame	PLCP	Header	Payload	Checksum
7/16 byte	1 byte	6 byte	30 byte	**0-2312 byte**	4 byte

Frame Control	Duration/ Connection	Addr 1	Addr 2	Addr 3	Seq Ctrl	Addr 4
2 byte	2 byte	6 byte	6 byte	6 byte	2 byte	6 byte

  - Preamble
    - wake up the receiver
    - need to negotiate the frequency for EACH frame

      ⇒ to start the negotiation, Preamble is sent under a standard specific speed
  - Frame Control
    - denote the encoding / meaning of the rest of this frame

            1. control frame: control the communication with AP (acess point)

               e.g.request to send (RTS), clear to send (CTS), acknowledgement (ACK), ...

            2. management frame: manage relations with AP (acess point)

            3. data frame: sending the data

    - Reliability

        - detect error & drop frames

        - detect error & fix frames at receiver

        - detect error & sender sends again (WLAN default, as using acknowledgement)

            1. when resending, need to tag the frame as "resending", because acknowledge may lost
            2. may delay the performance because of delay

    - Association with AP

        - need to know:

            1. connection service (service set identifier - SSID)
            2. APs that accept this SSID

        - beacon / probe-request

            1. beacon: AP broadcast
            2. probe-request: client broadcast

        - authentication

          AP: connect to service database to check the ID-key (instead of storing info in local)

        - associate on to AP

            1. resource allocation
            2. re-associate
            3. dis-connect (free resource)

Network Layer

- Focus & Role

    - Message - Packets

        - Definition: fragments of message & smallest unit of data in the ne
        - Reasons: use spacial multiplexing more ⇒ more parallel

    - Traffic Control

        - Optimized routing ⇒ no order guaranteed

        - Prioritization

        - Compared with LAN:

            1. LANs focus on similicity, instead of optimization
            2. Spanning Tree can NOT guarantee optimal topology

    - Scaling Problems

        - Internet accross the world
        - Compatible to diffrent underlying LANs

    - Compatible for Different LANs structure

        - Routing as a layer upon LANs ⇒ network layer
    - Adaptive to Change
        - Coping the evolving network topology
    - Simplicity & Best-effort
        - Connection stage stored at ends
        - Minimal service level agreement ⇒ no guarantee but best effort

          (reliability provided only where it needed)
- Router
    - Forwarding
        - Happens within each router, based on its forwarding table
        - Distributed decision making
    - Routing
        - Happends on the globale level (in routers network)

          ⇒ optimizing routing causes each router optimize its forwarding table
        - Focus on packet ⇒ packets usually arrive in different order than that of when it's sent
    - Forwarding Table
        - packet forwarding table
            1. forward packet based on its destination address
            2. mor robust to router failure (find another path)
            3. learn / optimize forwarding table on the fly
        - circuit forwarding table
            1. forward packet based on the tag on packet
            2. storring states & policy in networks ⇒ virtual network

               ⇒ separate traffic, guaranteed performance (bandwidth, path, delay...), prioritized routing
            3. overhead of setup / tear down the circuit (resource allocation / cleanup
            4. more guaranteed performance
            5. more fragile ⇒ not able to recover from nodes failure automatically
- Netwrok
    - Routing on Packtes
        - statistical multiplexing - sharing links
        - decision made in destributed routers
        - no guarantee on arrived packets' order (or dependency)
    - Connectionless vs. Connetion-oriented
        - Connectionless ⇒ packet forwarding - network makes all decisions, in each distributed router
        - Connetion-oriented ⇒ circuit forwarding
- Hosts

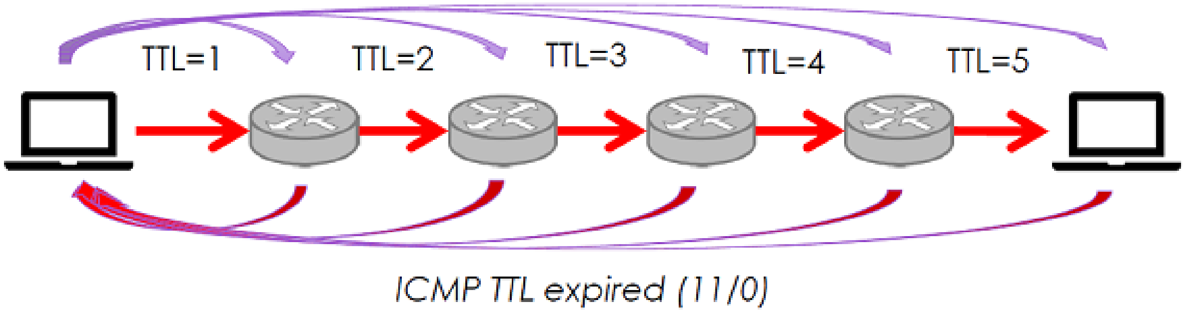
    - Sending Packets
        - send to local LANs service (switches)
        - switche decides the forwarding direction

          ⇒ router - outer internet

          ⇒ local hosts - in my LANs
    - Hosts Routing Table
        - longest matching prefix + broadcasting address
- Communicating with Link Layer
    - Reason
        - link layer deals with only MAC address

          ⇒ communication across layers (IP ⇔ MAC)
        - link layer sends only its frames

          ⇒ inter-changing of IP packet and link frames, especially address
    - The Address Resolution Protocol
        - source MAC: read from local hardware
        - destination MAC:
            1. sender broadcasts LAN frame to call for corresponding IP address
            2. receiver replies with its MAC address
        - optimizations
            1. caches the MAC address (with time-out)
            2. cache passing IP address (when others broadcast & reply)
            3. upon connection on LAN, broadcast my IP address (MAC address in frame's address feild)
- Allocation of Address
    - Consideration
        - globally-unique address
        - address aggregation
    - Authorities
        - allocate regional IP addresses blocks to regional internet registries
        - registries allocate IP addresses blocks to ISP
- Internet Control Message Protocol
    - Aims
        - special packet for router to inform the hosts (usually senders, including routers)
    - Traceroute



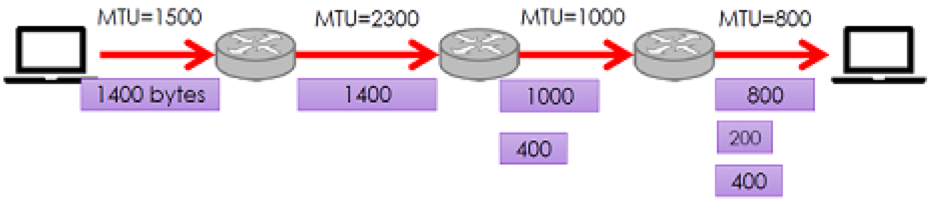
    - Sending message with increamenting TTL

      $\Rightarrow$ the i^{th} router sends back with corresponding exceptions (via control messge protocol)

  $\Rightarrow$ host can find out the path its packet is taking

    - TTL: time to live (in hops count)

- Fragmentation in IP
  - Slicing Packets
    - packet bigger than LAN's payload $\Rightarrow$ sliced packets



    - Realizing the need of slicing:

      packet size > MTU (maximum trasmission unit)

    - Flags to inform the next router:

      1. Identification field: key to identify a packet uniquely
      2. Fragment offset: the offset of this packet in the original big packet
      3. MF: more fragment flag $\Rightarrow$ more fragment of packets after me
      4. DF: don't fragment flag $\Rightarrow$ no more fragment after me

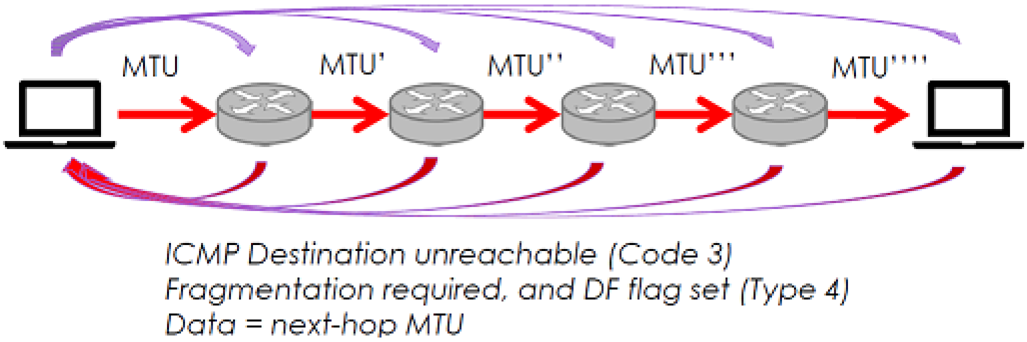
    - Trasmitting sliced packets

      1. copy IP Header, including identification each sliced packet belongs to the original packet
      2. adjust Length, Checksum & TTL (time to live) feilds of each sliced packet
      3. set fragment offset & MF/DF flags
      4. receiver re-assemble accordingly

    - Potential problem

      1. more work for routers
      2. more potential internal packet loss
      3. security issue (injection within packet)

  - Avoiding fragment in IP $\Rightarrow$ path MTU discovery
    - Using internet control message protocol - similar to traceroute

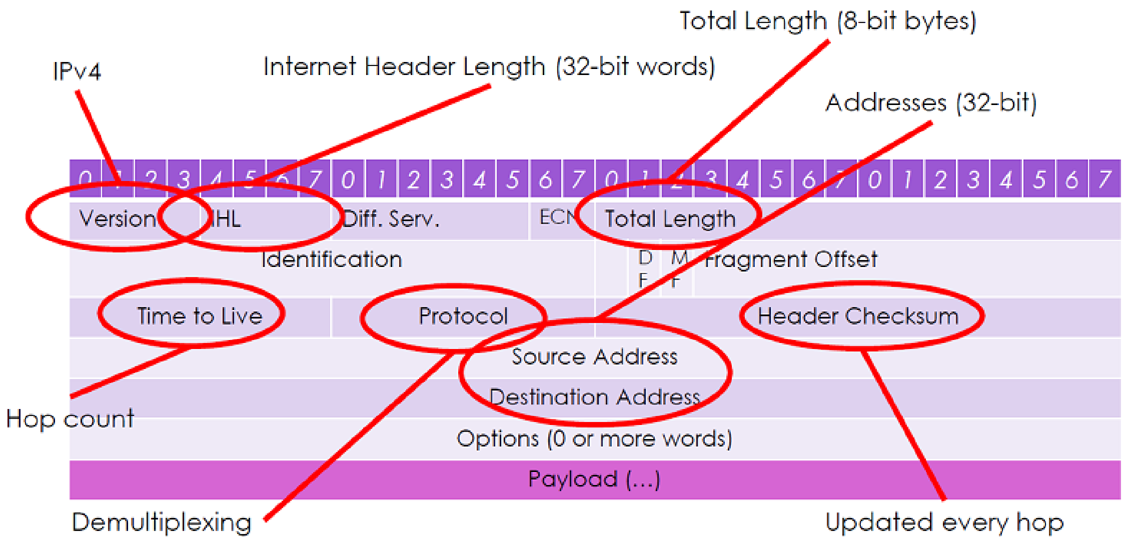


⇒ sender send at the lowest MTU

⇒ router focuses on sending packets

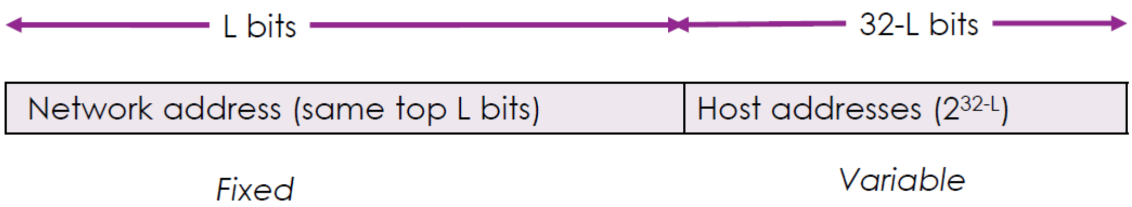
- IP multicasting
  - Definition
    - mutilcast to only a group of users, compared to broadcast
  - Challenge
    - sender may not be able to handle thousands of requests & data streams
  - Approach
    - usres subscribes to the sender, using special message / packet

      ⇒ all routers on the path know the subscription
- Internet Protocol - IP
  - IP - v4
    - Protocol overview (top-left -> bottom-right)



Note: checksum needs to be update at every router (hop counts changing)

- Addressing



1. total length - 32 bits

2. prefix denoting a netwrok, containing a range of the address

   ⇒ fewer entry in forwarding table

3. host addresses denoting the subnet under this network (denoted by prefix)

4. "/x" - x bits for host addresses; (32-x) bits for prefix

- Special addesses:

  1. private networks, multicasting, broadcasting, experimental, local interface, ...
  2. convention: sub-net wires, sub-net broadcast, local router, ...

- Classes

  1. denoted by first few bits of the address
  2. denoting different function of current packet (broadcastin, ...)

- Forwarding by prefix

  1. Prefix:

     network address

  2. Assumption:

     addresses aggregated into ranges

  3. Benefits:

     aggregation benefit of hierarchical addresssing ⇒ less entries, routing efficiency

     more flexible for directing specific trafic

  4. **Forward by longest matching prefix**

     default behavior for shorter (less-specific) prefixes

     specialized behavior for longer (more-specific) prefixes

     ⇒ allow to route sub-chunks (some specific) of address to other hops / routers

     choose the one that match the most ⇒ "best-effort service"

- Addresses exhaust

  1. problem ⇒ no more available addresses & large amount of wasted addresses

  2. current solutions

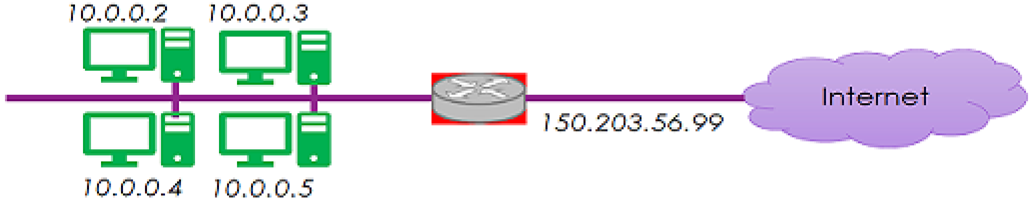
     re-allocating smaller chunk of addresses

     ⇒ addresses aggregation damaged

     ⇒ larger forwading table & more updates

     ⇒ routing efficiency ↓

     NAT (Network Address Translation) - use private address space behind a public IP address

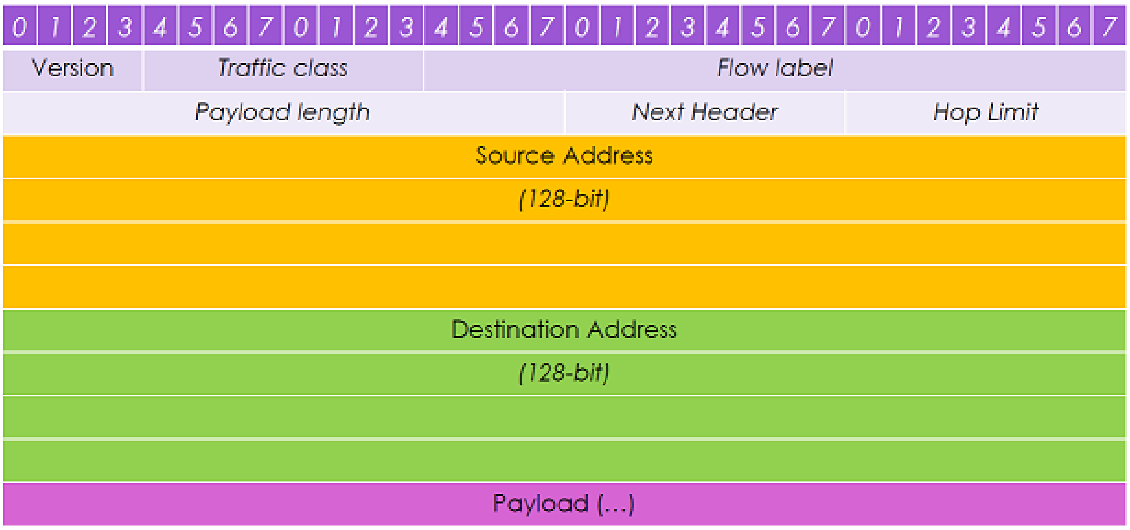


  3. future solution: IPv6

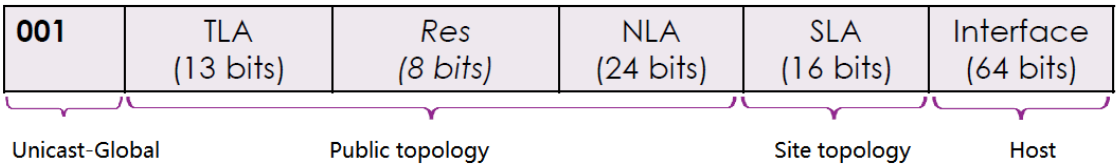
        - Potential & exisiting problems
            1. designed in a smaller & more trusting world

               ⇒ lack of security, mobaility and compatibility concern

            2. out of addresses & routing efficiency problem
    - IP - v6
        - Protocol overview



            1. larger address sapce: 128-bits
            2.
        - Addressing



            1. 3-bit header:
            2. TLA: top level aggregator - global ISP
            3. Res: reserved
            4. NLA: next level aggregator - site
            5. SLA: site level aggregator - subnet
            6. Interface: address in local subnet - host

          Advantage: explicit addresses aggregation

        - Transferring to IPv6
            1. dual stack: routers run both (2 separate pathways)
            2. translate: convert IPv4 ↔ IPv6
            3. tunneling: pack IPv6 packet inside IPv4 packet until a router recognize IPv6 (v4 everywhere)

Transport Layer

- Focus & Role
    - Message - Segment

- Definition & components:
  1. functionality & quality (including reliability) for applications
  2. host-to-host message
- Main Services
  - Reliability
  - Communication between hosts (on their ports)
- Port and IP addresses
  - IP address for host
  - Port for applications ⇒ port binding:
    1. port allocated on memory;
    2. client connects to an exposed port;
    3. server maintain the concurrency from inside
- Service Types
  - Reliability:
    1. reliable: packet loss repaired at transport layer
    2. unreliable: reliability offload to applications
  - Communication froms:
    1. messages: self-contained command and response
    2. byte-stream: generic flow of bytes, chunked into segments

  ⇒

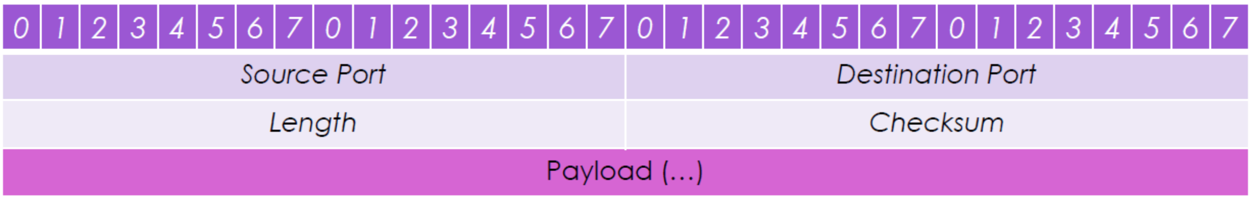
	Unreliable	Reliable
Messages	UDP (datagrams)	
Byte-stream		TCP (Streams)

- TCP vs. UDP
  - Comparison
    - TCP: many features, able to negotiate
    - UDP: enhanced packet

TCP	UDP
Connection-oriented *(significant state in transport layer)*	Connectionless *(minimal state in transport layer)*
Delivers BYTES: once, reliably, in order	Delivers MESSAGES: 0-n times, any order
Any number of bytes	Fixed message size
Flow control (sender/receiver negotiate)	Don't care
Congestion control (sender/network negotiate)	Don't care

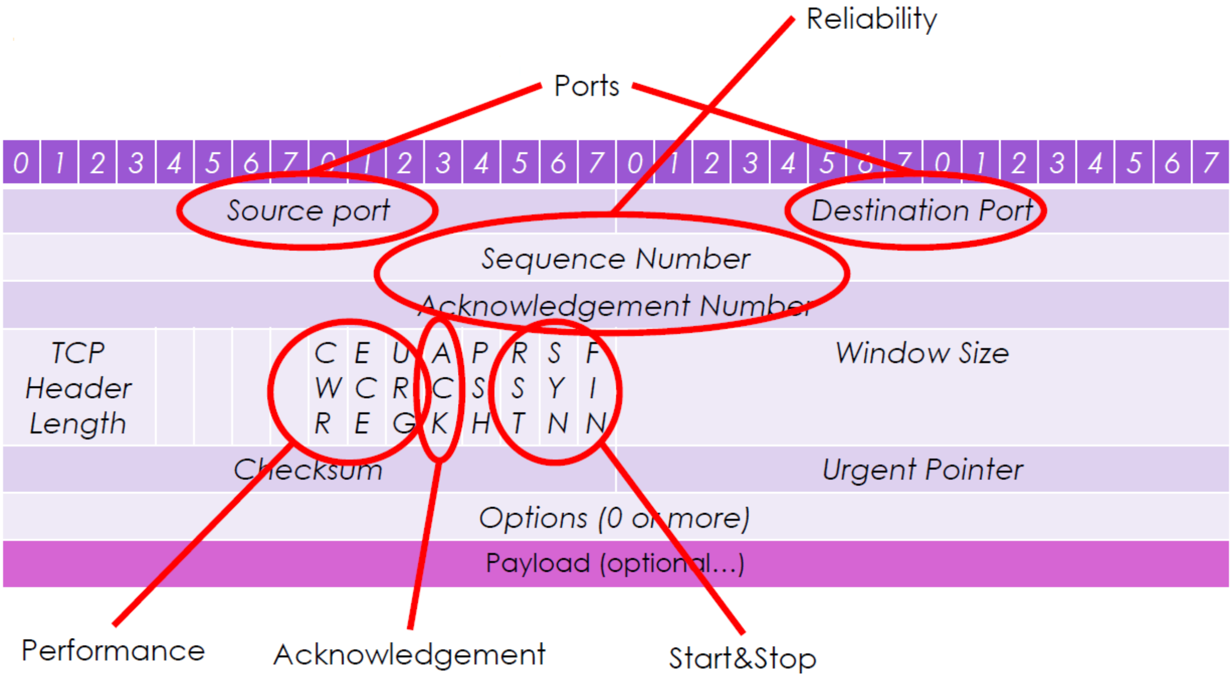
  - Situation for UDP - multicasting

        - connectionless
        - Replicate segments or packets are fine
        - Missing (some) segments or packets are fine
    - Sending Byte-stream - TCP
        - Chunks of byte-stream in segments - message boundaries not reserved
        - Read / write on buffer
- UDP Segment



    - Ports
        - associate segments with applications / sessions

          note: application can use multiple ports
- TCP Segment



    - Options
        - Maximum segment size
        - Window sacle - upon window is full will acknowledegement sent
        - Time stamp
        - Selective acknowledgement - advanced acknowledgement
    - Reliability
        - Important components: sequency number, acknowledegement
        - Sequence number: byte count in a stream (in $\bmod n$ space) $\Rightarrow$ can be used as relative time stamp

          Note: does NOT start from 0 $\Rightarrow$ security reason

- Acknowledgements: with sliding windows with size w & selective repeat

  $\Rightarrow$ more parallel

  1. <u>sender</u>

     allows w segments to be outstanding (no ACK provided) - sliding window

     a timer for every unACKed segment - re-send after time-out

  2. <u>receiver</u>

     buffers many segments $\Rightarrow$

     ACK received segments

     request missing segments - which is in gap of segments stream & the future

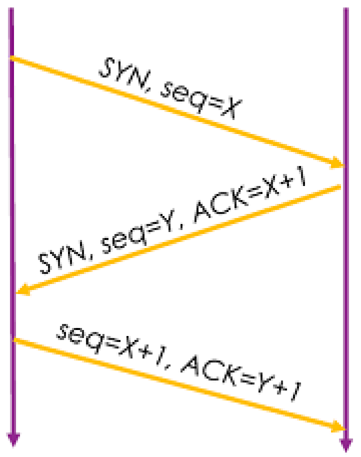
  3. Pros: no need to suspend on every segment $\Rightarrow$ more parallel

- Connection in TCP - 3 Way Handshakes

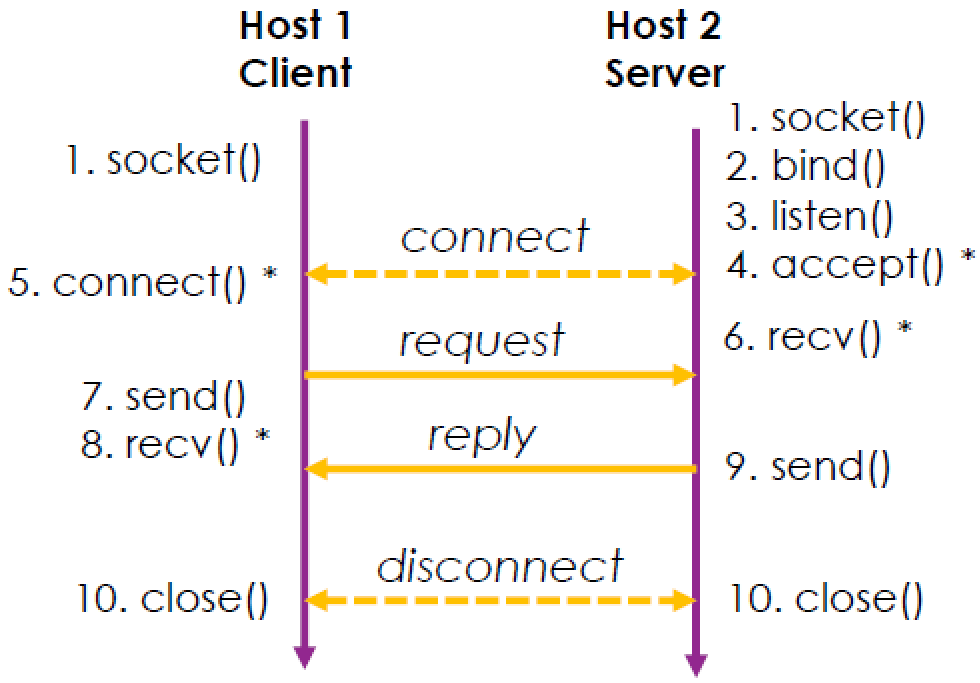
  - Reasons:

    1. TCP is full-duplex $\Rightarrow$ need to connect two independent paths for each direction
    2. need to start together $\Rightarrow$ negotiate synchronization & initial Seq (sequence number)

  - Procedure:



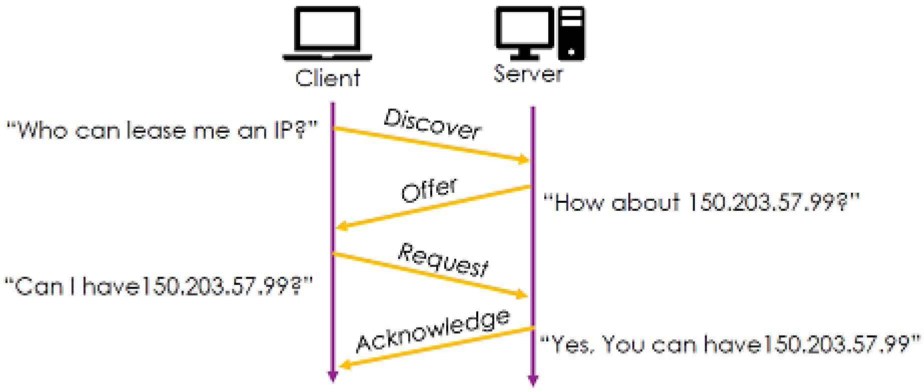
- Connection between Applications
  - Necessary components:
    - source & destination IP addresses
    - source & destination ports
    - protocol
  - Socket API:



        - Note: * = potential blocking calls
    - NAPT (Network Address and Port Translation)
        - NAT: hiding behind a public IP address
        - NAPT: hiding behind a public IP + translate outbound port into host's actual port

Application Layer

- Focus & Goal
    - Build Sessions
        - Sessions: a series of interactions
        - based on TCP reliable byte-stream or UDP unreliable messages, or combination / extension ...
    - Presentation of Content
        - Interpret content: interpret message/byte-stream inside TCP/UDP segments' payload
        - Handle Command: handle request & control from both ends
    - UDP-based Application
        - Short messages ⇒ light server touch ⇒ simple request-response transaction
    - TCP-based Application
        - Large content change ⇒ longer & complex sessions
- Dynamic Host Configuration Protocol (DHCP) - Getting IP Address
    - Goals:
        - allocate IP address
        - automatic configuration, instead of manual
    - Negotiation procedure



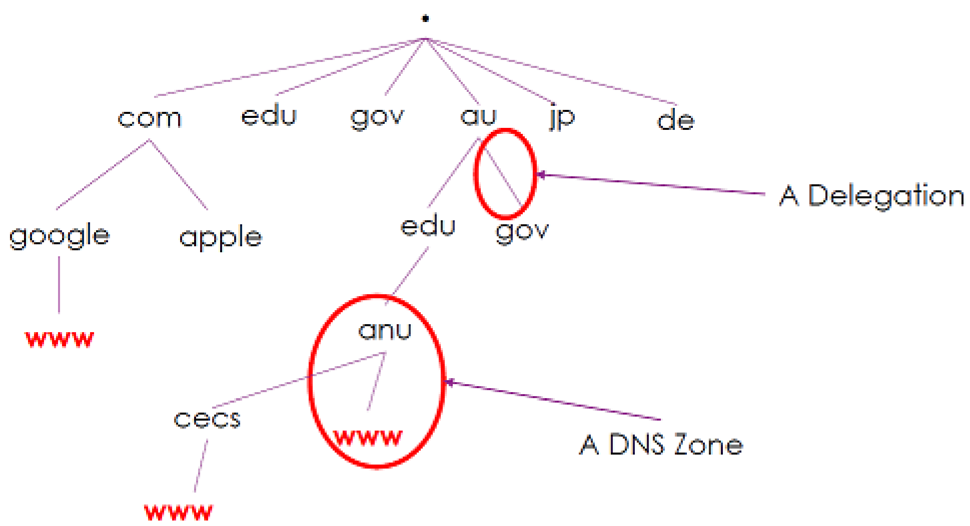
        - Need to broadcast to discover
        - need to request the IP address after being offered (backward compatibility)
        - can directly request the IP address in hand, when it is close to expired
    - DHCP relays (转接):
        - DHCP server in the middle;
        - Relays on router/switches...
        - Relays forward the request to the server (broadcst → unicast)
    - Multiple DHCP
        - Reasons: performance, robustness
    - DHCP service with orther services
        - on the default router (to the internet)
        - with DNS service
- Domain Name Service (DNS)
    - Goal & Reasons
        - Changing IP address
            1. IP address expired
            2. LANs re-configure
            3. physical movement to other places
        - Human readable name
    - Definition
        - Names - for humans
        - Addresses - for underlying protocols and applications
        - Resolution - finding the right servers (authority) to find the requested IP - names

      Note: 1 name can have multiple addresses; 1 addresses can have multiple names
    - Design
        - Distributed control
        - Hierarchical namespace - delegate to authorities (for them to be responsible for, legally)
        - Automated protocol & handling
    - Namespace
        - Root: '.' (usually dropped)

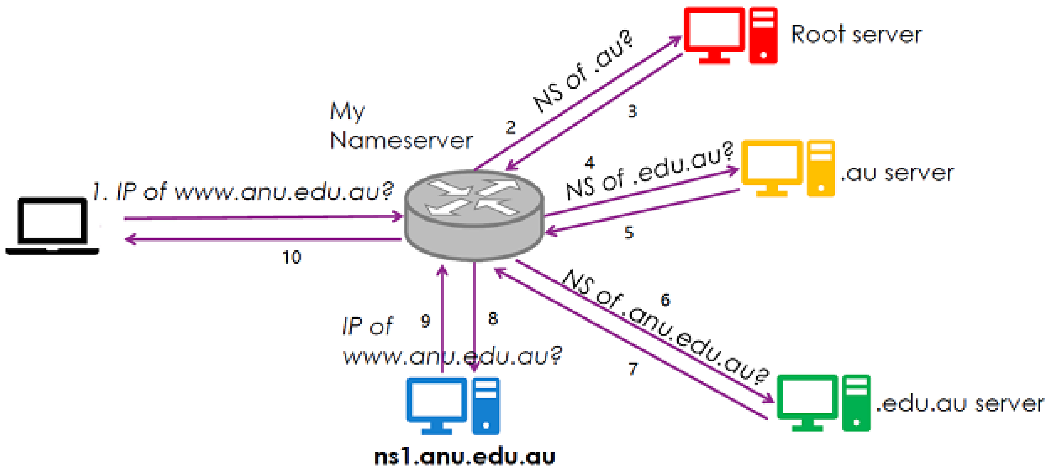
- Top level domain (TLD):
    1. classification - com, edu, org, net
    2. contry code - au, us, jp, cn, ...
- running down the hierarchical to the local host

- Domain
  - Delegated to authorities; authorities hold legal responsibility
  - Responsibility covers the subtree starting form the delegated point
- Zone
  - Shared pieces of DNS database - through technology
  - Recording the information about:
    1. meta data of the the zone
    2. further relations (delegations)
    3. resource records: addresses and corresponding DNS names, services & other meta data

       (includes time stamps - used for cache)
- Example



- Resolving the IP address (Resolution)
  - Iterative query to resolve IP address
    1. example:



2. Improvement - Caching

   ⇒ cache information with an expired time

   ⇒ dirrectly knows the authorities / IP address for the next query (within a time) - shortcut

- Nameserver, name & IP addresses replication:

  1. ask more than one server ⇒ spread the workload & risks
  2. more IP ⇒ more hosts, prevent nodes fail-over
  3. more names ⇒ less typo

  Note: enable prioritization, with other addresses

- Security issues

  can change the cache in router - man-in-the-middle

  1. make a query by yourself;
  2. draft a reply, including a valid source address & guess the ID (enumeration will do);
  3. (router usually check only destination addr, matching ID, is it answering query)

- More security: signature, public-private keys, ...

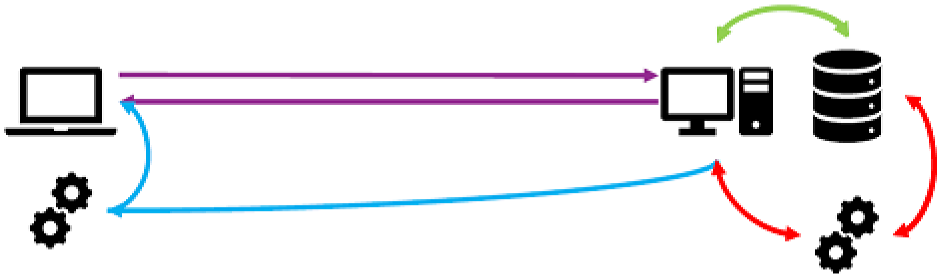
  ⇒ validate along the way ⇒ building a trust chain

- Policy: can change the query on the fly ⇒ enable policy, e.g. content check, DNS polluting
- Dynamic DNS - NAPT

  1. register a DNS name on DNS server for my host ⇒ server handle the address change
  2. host address change: local IP address (private address) change ⇒ router port change
  3. message sent to me will find the right port through DNS (DNS will query my router)

- HTTP Protocol - Most Common Used Web Application
  - Focus
    - Deliver associated content
    - Linking related content on the web, instead of fetching everything to local
    - Light weight (initially designed) - used with UDP as transport layer sometimes
  - URI vs. URL:
    - URI - uniform resource identifiers: identifier, identifying anything
    - URL - uniform resource locator: an example of URI ⇒ location on internet

- URLs - Schemes
  - Various schemes
    - server name in DNS, IP addresses, HTTP protocol, resource on the host, query to server
- Dynamic & Static Contents
  - Procedure



purple: communication & data between server and client

red & green: interpret & execute command on server side

blue: get input / command and execute program on the local - security concerns

1. parse URL & resolve DNS
2. connect to the host
3. make HTTP request
4. receiver contents
5. close TCP/UDP connection & render the page

- HTTP Request
  - Basic Command
    1. GET: get resource
    2. HEAD: get the headers about the resource (meta-data)

       enable backward compatibility

       enable re-direction - for performance reasons (e.g. redirect to closer server)
    3. POST: append my contribution to the host
  - Feedback
    1. 1xx - not used currently;
    2. 2xx - OK (successful); 3xx - redirection (no longer in this address)
    3. 4xx - hosts' problem, e.g. 403 bad request; 5xx - server has problem
- States in HTTP
  - Default
    1. Stateless - server should not hold state (too much - each session needs a key)
    2. Stand along query
  - Session Key in URL
    1. Key: encoded in URL
    2. Passing the Key: as part of query $\Rightarrow$ server interprets it from the query
  - Session Key in Cookies
    1. Key in cookies: set & offered by server; held by client

2. Passing the Key: include the cookies in the communication with server, if relevant
3. Type of Cookies

   session cookies - deleted when page closed

   persisten cookies - static cookies with expire time

- Efficiency in HTTP
  - Parallelism

    use idle bandwidth & potential distributed servers

  - Persistence

    use one TCP connect for all requests

    need only one connection set-up (3-way handshakes)

  - Pipeline

    send all requests and wait for all resource from server (full-duplex)

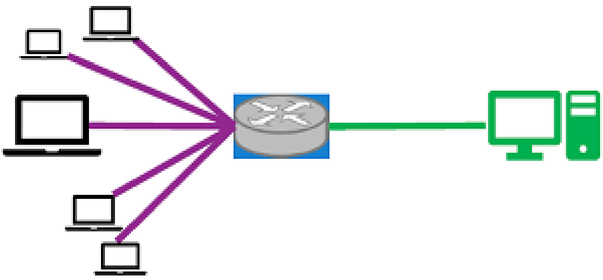
  - Caching
    1. Browser (local) level

       cache on demand (cache the most popular one)

    2. Proxy level - proxy cache ($\Leftrightarrow$ caching proxy)

       cache on routers on local LAN - closer to client

       share the cache in the LAN - enable security check & policies



  - Network server level - Content Distribution Network (CDN)
    1. distributed file system over internet - DNS redirects client to the best cache

       $\Rightarrow$ server offload; content closer to clients

    2. cache before the request $\Rightarrow$ as a way of sharing distributed servers