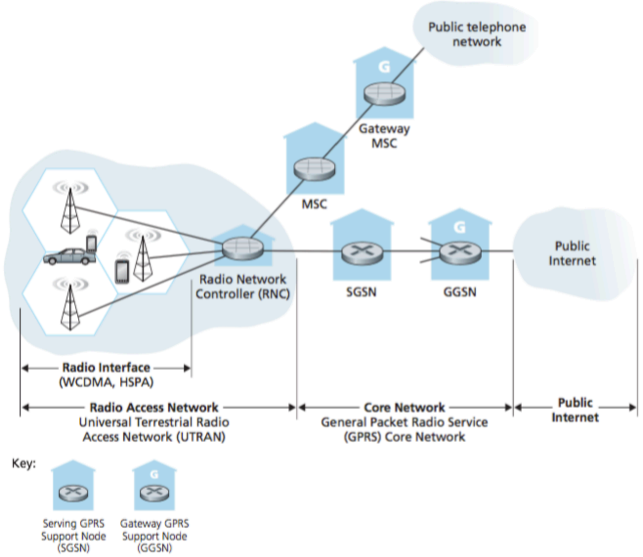
Security

Symmetric vs *public/private key* (more computation) encryption
Authentication: issues: spoofing, playback attack (send encrypted *nonce*), man in the middle.
Digital signature: since pub/priv key encr is expensive, hash the message and encrpypt the hashed value with your private key.
Attacker can record and play/re-order specific packets -> must put sequence numbers into MAC.
Attacker could replay all records -> must use once in a lifetime encrypted tags (nonce)
IPsec: provides integrity, authentication, confidentiality, replay attack prevention
Firewalls: prevents DoS attacks (SYN flooding), illegal modification of data behind it, only authorized access OK within network; special gateways for applications that need additional access to outside; limitations: IP spoofing, many gateways for many apps, software needs to communicate to gateways
Intrusion Detection Systems (IDS): prevent DoS attacks; do port scanning, network mapping.

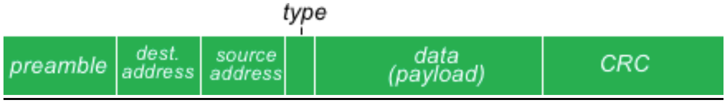
Wireless and mobile


Public telephone network
G
Gateway MSC
MSC
Radio Network Controller (RNC)
SGSN
G
GGSN
Public Internet
Radio Interface (WCDMA, HSPA)
Radio Access Network Universal Terrestrial Radio Access Network (UTRAN)
Core Network General Packet Radio Service (GPRS) Core Network
Public Internet
Key:
Serving GPRS Support Node (SGSN)
G
Gateway GPRS Support Node (GGSN)

802.11 – Each Access Point (AP) has unique SSID;
Beacon frames – sent from APs for new wireless hosts to detect.
ACKs used for reliability at frame level.
Reservations through request-to-send (RTS) and clear-to-send
4 Addresses: MAC of receiver; MAC of transmitter; MAC of router interface to which AP is attacked; ad hoc mode
CDMA/CA -- CA is better than CD; Transmissions of other nodes may not be detected due to Hidden Terminal Problem. After transmission, wait for ACK of frame from AP. If no ACK, then *binary exponential backoff.*
3G=voice+data (look at image); 4G=voice over IP
Cellular element (Mobile IP element)
Home system (home network) – stores host's IP address
Gateway Mobile Switching Center (home agent) – point of contact to obtain routable address of mobile user. Home Location Register: database in home system containing phone numbers.
Visited system (visited network) – network where host is now.
Visited MSC (Foreign agent) – analogous + Visitor Location Record
Roaming Number (Care-of-address) – current phone number in FA
Indirect routing: Correspondent routes through Home Agent. Used by Mobile IP (double encapsulation at home agent)
Direct routing: Route directly to foreign agent, use *foreign agent anchor* if foreign agent changes.
Handoff within common MSC controlled by MSC.
Handoff between MSCs route through chain of MSCs (Home, then Anchor, then all the other MSCs visited)

Link Layer

CRC error checking
Broadcast (shared wire or medium): old Ethernet, 802.11 wireless
MAC protocol taxonomy: channel partitioning (TDMA/FDMA), random access (ALOHA, CSMA), taking turns
Slotted ALOHA: transmit when you get frame, use CD to stop
CSMA/CD: listen before transmit; collision can still occur due to propagation delay; if during transmission collision is detected, abort; after abort enter *binary exponential backoff*: after m collisions, choose 0<K<2^m at random and wait that many frames.
Taking turns MAC protocols: issues=polling overhead, latency, single failure point at master. Bluetooth
Address Resolution Protocol (ARP)
ARP table: provides IP to MAC address translations
If you need MAC B's addr, broadcast ARP query; the target host responds with its MAC addr and you cache into your ARP table
Routing to another LAN: send frame with dest equal to first-hop router's MAC address; the router than creates new frame based on IP addr in datagram and forwards to actual destination.
Ethernet: Wired LAN; active switch in center. Ethernet frame:



Connectionless; unreliable; MAC protocol: CSMA/CD with backoff
Ethernet switch: link-layer device; hosts unaware of its presence; plug-and-play and self-learning; supports multiple simultaneous transmissions. *Self-learning* achieved by recording MAC addr to output link translations as frames come in, if switch doesn't know dest MAC, then sends frame to all links.
VLANs: single physical switch implements many virtual LANs and virtual switches. One VLAN can be distributed over many VLAN switches via *trunk port*.
Dynamic membership: ports can be assigned to any VLAN
Multiprotocol label switching (MPLS): provides section in frame that allows routers to forward without stripping down to a dataframe. MPLS capable routers—MPLS forwarding decisions can differ from those of IP.

Network Layer – Control Plane (routing)

Determines path taken by packets
Per-router control vs. *Software Defined Networking (SDN)*
Global (all routers have complete topology) – *Link State algorithms*(Dykstra); *Decentralized* (router knows only physically connected nodes) – *distance vector algorithms* (Bellman-Ford).
DV = Good news travels fast, bad news travels slow
Poisoned reverse advertise to neighbor Y an infinite distance to node X if optimal path to X travels through neighbor Y.
Link state speed of convergence is O(n^2); more robust than DV
Scalability – more nodes + administrative autonomy implies Autonomous Systems (AS).
Routing Information Protocol: DV based on hop count and poisoned reverse Intra-AS Routing. Limit hop count to prevent loops.

Open Shortest Path First (OSPF): Link-State algorithm; carried directly over IP (rather than UDP/TCP)

Hierarchical OSPF: two-level hierarchy: *local area* (run OSPF within itself) + *backbone* (runs OSPF independently); *area border routers* summarize distance info to their local areas; *boundary*

Border Gateway Protocol (BGP): inter-AS routing; **eBGP** for reachability info of neighboring ASs; **iBGP** for propagating reachability info to all AS-internal routers; AS_PATH and NEXT_HOP attributes used in BGP messages (over TCP); policy-based routing

Hot Potato Routing: choose gateway router with least intra-domain cost

Intra-AS (optimize for performance) vs *Inter-AS*(policy over performance) routing

Software Defined Networking (SDN): monolithic router contains all switching hardware and implements all protocols; easier network management; traffic engineering easy; other devices run generalized flow-based forwarding (e.g. OpenFlow) – switches, routers, NAT are all the same device (called *data plane switches*)

SDN controller runs network OS, maintains network state info, has network control apps above and below. Many distributed controllers for performance, scalability, fault-tolerance, robust

Internet Control Message Protocol (ICMP)

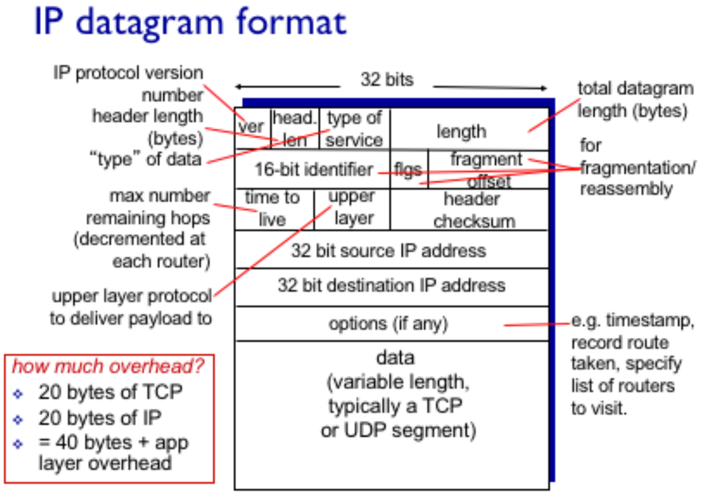
<u>Network Layer: Data Plane (forwarding)</u>

Longest prefix matching.

Switching fabric. Via memory (by CPU, limited by memory bus bandwidth; can process packets in parallel), bus (limited by bus bandwidth; can't process packets in parallel), crossbar (can send non-overlapping packets in parallel with bus speeds per packet)

Head-of-the-Line (HOL) blocking: datagram at head of input link waiting for another packet to finish transmitting to its output port stalls all others in its queue.

IP datagram format


IP datagram format
IP protocol version number
header length (bytes)
"type" of data
max number remaining hops (decremented at each router)
upper layer protocol to deliver payload to
32 bits
ver
head. len
type of service
length
16-bit identifier
flgs
fragment offset
time to live
upper layer
header checksum
32 bit source IP address
32 bit destination IP address
options (if any)
data (variable length, typically a TCP or UDP segment)
total datagram length (bytes)
for fragmentation/ reassembly
e.g. timestamp, record route taken, specify list of routers to visit.
how much overhead?
20 bytes of TCP
20 bytes of IP
= 40 bytes + app layer overhead

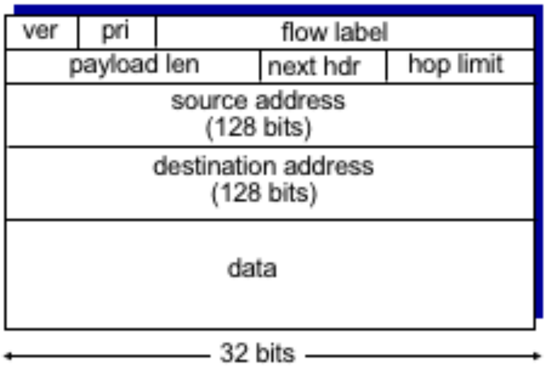
IP fragmentation: network links have *MTU (Max transfer size)*, so fragment datagram if too large and reassemble at destination.

Classless InterDomain Routing (CIDR): allows for subnet masks of arbitrary size.

Dynamic Host Configuration Protocol (DHCP): Helps new host get its IP address; host broadcasts DHCP discover message; DHCP server responds with DHCP offer; host requests IP address; DHCP server sends back ACK; Also sends IP addr of first-hop router for host, name and IP addr of DNS server, network mask of subnet.

Network Address Translation (NAT): advantages: only need one IP addr to support many hosts, can change addr of devices in local network without notifying outside network, can change ISP without changing host addr of local devices, devices inside local net protected from outside world; issues: routers should only process up to layer 3, P2P app issues, *NAT traversal* (i.e., contacting clients behind NATs is hard).

IPv6 datagrams



Removed checksums, options allowed but are not in header (pointed to by Next Header field), ICMPv6.

Tunneling: used to help gradual transition from IPv4 to IPv6; IPv6 datagrams encapsulated in IPv4 datagrams among IPv4 routers

OpenFlow data plane: idea of flow; generalized forwarding through pattern matching values in headers + action pairs

In OpenFlow switches, NATs, routers all the same general switch

<u>Transport Layer</u>

Mechanisms for reliable data transfer: error recovery, error detection, sequence numbers, timeouts, acknowledgments.

Go-Back-N: sender can have up to N unacked packets in pipeline; receiver only sends *cumulative ack*; sender has timer for oldest unacked packet only, when it expires retransmit all unacked pkts; no receiver buffer

Selective Repeat: sender can have up to N unacked packets; receiver sends individual ack for each packet; sender maintains timer for each unacked packet, when timer expires retransmit only associated packet. Requires receiver buffer.

TCP: full duplex, connection-oriented, flow controlled, pipelined

Sequence num: num of first byte in segment's data

ACK num: sequence num of next byte expected; cumulative

EstimatedRTT = 0.125*SampleRTT + 0.875*EstimatedRTT

DevRTT = 0.25*|SampleRTT-EstimatedRTT| + 0.75*DevRTT

TimeoutInterval = EstimatedRTT + 4*DevRTT

SampleRTT not computed for retransmitted segments due to ambigouty

The timeout value is doubled for each segment retransmission

TCP fast retransmit: if triple duplicate ACK, resend packet, enter fast recovery and set ssthresh=cwnd/2, cwnd=ssthresh+3

TCP fast recovery: increase cwnd by 1 for every duplicate ACK; when new ACK is received set cwnd=ssthresh

TCP Slow Start: increase cwnd by 1 for every new ACK until cwnd=ssthresh and switch to congestion avoidance.

Congestion avoidance: increase cwnd by 1 MSS per RTT until congestion loss; if loss from timeout: go to slow start, set ssthresh=cwnd/2, cwnd=1, retransmit lost packet (timeout implies heavy packet loss); if loss from duplicate ACKs go to fast retransmit (most probably only single packet lost)