1Mbps = 0.125MB/s = 1000000bit/s = 125000Byte/s

- **Network Layer: Data Plane**

best-effort host-to-host transport

- Forwarding: Routers map incoming packets to output ports

forwarding table

Destination Address Range	Link Interface	
11001000 00010111 00010000 00000000 through 11001000 00010111 00010111 11111111	0	200.23.16.0 through 200.23.23.255
11001000 00010111 00011000 00000000 through 11001000 00010111 00011000 11111111	1	200.23.24.0 through 200.23.24.255
11001000 00010111 00011001 00000000 through 11001000 00010111 00011111 11111111	2	200.23.25.0 through 200.23.31.255
otherwise	3	

Router forwarding table:
○ Longest prefix matching: 能匹配的detailed *最长地址前缀*

-排队管理: FIFO; round robin; weighted fair queue//保证 min

-IP addr 32 位标识符 拆成 4 块 结构: subnet + host (part)

○ Subnet: 无须中间 router, 即可物理连接的 interface, 创建孤立网络

多少 hosts in /23? 2^(32-23)-2 = 510 (上下 255 边界不能取)

○ 128.119.40.128/25 切四个均等 subnet: /27; 2^(32-25)=128 每个 32bit; 128.119.40.(128,160,192,224)/27

○ 最多建立多少 subnet? 2^(32-25)=128; 每个 1 IP addr

○ 5+2=7<8(2^3) 128/8 = 16 subnets

-**DHCP: 动态 obtain ipaddr & network info**

○ 步骤:

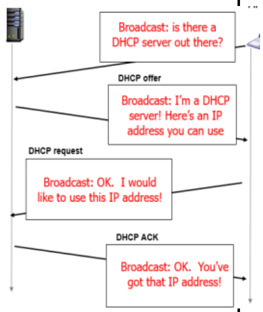
host broadcasts DHCP discover msg

DHCP server broadcasts DHCP offer msg

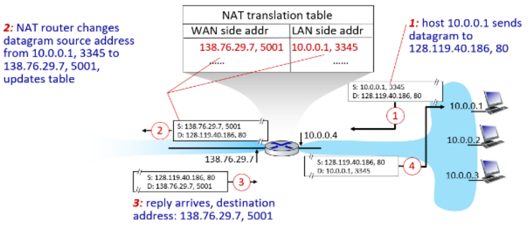
host requests IP address: DHCP request msg

DHCP server sends address: DHCP ack msg

○ ? broadcast: *ensure that all responding DHCP servers know that the client has chosen a server*



-**NAT 网络地址转换**:use **one** public IP addr for all devices in LAN



○ 当本地 host 从本地网络外 send/rcv packet, *all* datagrams *leaving* local network have *same* src NAT ip addr, but *different* src port numbers

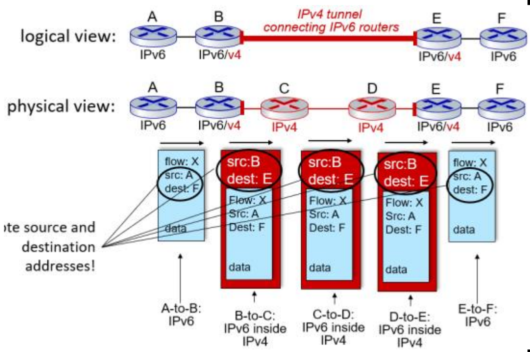
○ remember every pair in NAT table: 传出-将 src 的 ip 和 port 转换成 nat, 自此 remote side 响应用 nat 的信息; 传入-将传入 datagram 的 (nat IP 地址, 新端口) 替换为存储在 NAT 表中的相应

○ ? *violate strict separation of protocol layer*. router 应只处理到 layer3; addr shortage 应由 IPv6 解决 //violate end-to-end argument (port # manipulation by 网络层 device)

Queuing delay: 到达速度过快 Buffer 管理和 packet 调度

-ipv6: 128 位(ipv4 地址不够); 简化 header (处理更快 forwarding)

Tunneling: ipv6 数据报 carried in v4 datagram ? 并非所有路由器同时升级,ipv4 和 v6 路由器的网络混合

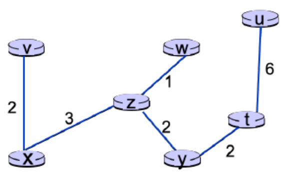


- **Network Layer: Control Plane**

-**路由算法//Convergence**: 链路成本变化, 更新在整个网络中泛滥, 所有节点 converge to new topology | delay: detect failure; flood link-state information; re-compute forwarding tables

○ *Link State*:每个节点知道整个网络拓扑和链路成本 Dijkstra

○ 仔细! 将表中最小的格子对应数据加到 N' 要会画最短路径图



Destination	Outgoing link
t	(x,z)
u	(x,z)
v	(x,v)
w	(x,z)
y	(x,z)
z	(x,z)

Outgoing link(root,倒数第二个节点)

Distance Vector: 将当前最短路径成本发给自己的邻居, 以 iterative distributed 进行更新 Bellman-Ford:

Let $D_x(y)$: cost of least-cost path from x to y.

Then:

$$D_x(y) = \min_v \{ c_{x,v} + D_v(y) \}$$

v's estimated least-cost-path cost to y

min taken over all neighbors v of x

direct cost of link from x to v

Distance table 已给的推断行(自身)中的 inf 要 update 替换掉, 其他 inf 保留 。把推断行作为 vector send 给其他包含 inf 的行

Node z	w	x	y	z
w	0	6	inf	1
y	inf	2	0	2
z	1	4	2	0

z will send vector {1,4,2,0} to w and y

对比: L-S: send info of my links to all nodes; D-V: send info of my known shortest paths to only my neighbors | message complexity: LS: all-to-all 通信 O(n2)条消息; DV 好: 仅在邻居之间交换 speed of convergence: LS 更快 robustness: LS router 为自己的链路通告 uncorrect cost, 只计算 own table; DV 通告不正确路径,表被利用

-**Internet Routing**: 异: policy: 域间 admin 想控制 how traffic routed, 谁通过; 域间单 admin, 无所谓; scale: Hierarchy 减小 table size&update traffic, 性能: 域内专注性能, 域间策略更重要|协同: Hierarchy lets routing scalable, 让 operator control their own networks: organized into Autonomous Systems

○ *intra-AS 域内*: 同 AS 内同协议 //RIF, EIGRP: DV based

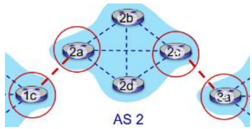
-**OSPF 开放最短路径优先**: link-state| broadcast other routers in as like normal(change or 30min). 链路成本由管理员定义: delay, bandwidth,, constant| 发消息用 IP |2-level hierarchy: local area, backbone-flood 只在一个层级, 每个节点有该层级 topology

优: security, load balancing, hierarchy for scalability

○ *inter-AS* 域间: gateway router 连接| like distance-vector|BGP 告知 path, not just distance (避免 loop)

差异: BGP advertise exact path to the dst

Ebgp:从相邻AS获取子网信息 | iBGP: 将可达性信息传播到所有 AS 内部 router



Ex. 小 X 连大 A, 后备大 B: X 不需要将 A 的 path advertise 给 B. (这样 B 会给 X, 再转给 A traffic) ISP 承载发往或来自其客户的流量. B 不是 X 的客户

Y-A| Z-B: 若大 ISP A 能去 Z-B 支路, 应当 advertise A this path. Y 是 customer, A carry Y 的所有 traffic to any prefix it know.

-**SDN** 软件定义网络:

○ *Generalized Forwarding*: forward based on dest. IP addr Match+action| 匹配任何层中到达数据包头中的 bit, 采取行动

Router M: longest prefix A: forward link | Switch M: dst mac addr A: flood or forward | firewall M: ip& port no. A: 允许/拒绝 | NAT M: ip& port no. A: rewrite ip addr

○ *Logically centralized control plane*: 每个路由器中的单独路由算法组件在控制平面中与 forwarding tables 交互

优: 1. easier network 管理: 避免 misconfiguration, traffic flow 更灵活 2. table-based forwarding: 允许对 router 编程- compute table centrally and distribute 3. 开放性, 创新

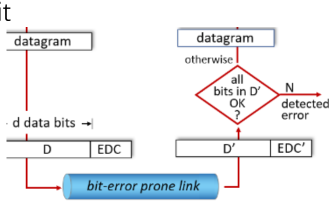
- **Link Layer**: link 将 datagram 传输到物理上相邻的节点

-**Error detection & correction**: EDCbit

- *Parity bits* 奇偶效验

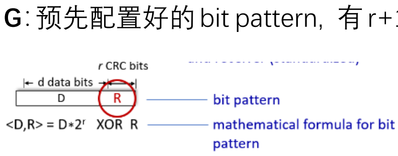
○ single bit parity: 设置 parity bit 位, detect **single bit** errors

○ 2d: detect&correct single bit error; detects (but can't correct) any combination of 2 errors



-*CRC 循环冗余校验*//更好: r bit detect burst errors up to r bits

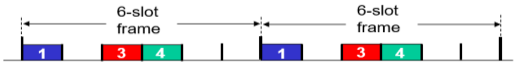
D: 要检验的 data bit; **R**: remainder; **r**: 所选择的 r 个 crc bit; **G**: 预先配置好的 bit pattern, 有 r+1 位 | 接收方知道 G, 将 <D,R>



$$R = remainder\left[\frac{D \cdot 2^r}{G}\right]$$

除以 G 如果 remainder 非零：检测到错误，然后 drop frames
-broadcast link: shared wire: old Ethernet\4G 5G\LAN\satellite
-**Multi access protocol:** ? 多个节点同时在 link 传输，导致 Collision：a node 同时收到多个不同信号…data garbled, and all frames lost

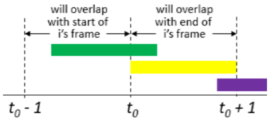
- ***Channel partitioning***:将频道分为更小片段，给节点独占
- *TDMA 时分多址*：每个节点在每一轮中获得固定长度的 slot



- *FDMA 频分多址*：每个节点分配固定频段//divide channel into frequency bands
- ***Randon access***：如何检测碰撞，恢复。
- ***slotted Aloha:*** **Np(1-p)$^{N-1}$ max=0.37**

efficiency = p(x transmit) * p(others do not transmit)//吞吐量:*R
好:单个节点在 channel 全速率下传输，decentralized(节点独立决定)，简单；坏: collision waste slots, free slot，节点可能在少于传输 packet 的时间内检测到冲突; clock synchronization

- *Aloha*: frame 到达时立即发送，碰撞后以 p 重传, no synchronization



- *CSMA/CD*: 有线容易无线难 $efficiency = \frac{1}{1+5t_{prop}/t_{trans}}$

carrier sense：传输前，节点 listen whether another node is transmitting (channel free?) 仍可能发生碰撞，由 propagation delay 导致。collision detection：是否 intersection of multi objects。检测到碰撞,传输中止，减少时间浪费。再次碰撞的几率指数降低

- ***Taking turns***:*Token ring(蓝牙)*: control token passed from a node to next sequentially 坏:令牌开销，单点故障(t)，latency
- *polling*: **coordinator node** 让其他节点依次传输 坏:同上 $\frac{NQ}{\frac{NQ}{R}+Nd_{poll}}$

//如果 only one node 有数据很亏
-***mac*** **addr**(48-bit, *hexadecimal*)：功能: local delivery: get frame from one interface to another physically-connected (same subnet, in IP-addressing) ?同时：*assure uniqueness*
异: Ip addr for global delivery | manufacturer 给 mac,完全 unique, 子网给 ip, locally unique | MAC portable, ip not |?
-***ARP*** 地址解析 resolution 协议: get MAC addr for given IP addr
ART table：ip, mac, tll
Step: **a.** A 想 send datagram to B, and B's mac not in A's ARP table. Then A **broadcasts ARP query**, 包括 B's IP addr as target (dst mac addr=FFFFFFFFFFFF//用于广播, all machines on LAN receive).
b. ip 不匹配的节点忽略，B 收到 query 后回复 A 自己的 mac.
另一个子网？ //src ip, src mac; dst ip, dst mac
a. ARP query: D D R2 FF **b.** ARP reply: R2 R2 D D **c.** datagram: D D A R2 **d.** ARP query: R1 R1 A FF **e.** ARP reply: A A R1 R1 **f.** datagram: D R1 A A
-***Ethernet:*** dominant wired LAN

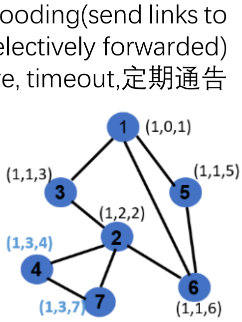
dest addr: MAC; *type*: 高层协议，多为 IP; *preamble*: 物理层-同步接收方、发送方 clock rate
unreliable, connectionless: no handshake & ACK; unslotted
CSMA/CD with binary backoff
-***Switch 交换机***: star topology,直连 host，在链路层| store, , buffer, forward 以太帧; transparent 主机不知道交换机的存在；无需配置
both are store-and-forward (exam header)：router 应用于网络层; switch 应用于链路层|both have forwarding tables: router uses routing algo based on IP addr; switch uses flooding based on mac
-*Self-learning*: learns which hosts can be reached through which interfaces

- **Switch table(可能多个相组)**: *MAC, interface, time* | 交换机记录 src MAC 传入链路；dst 未知时先 flooding(send links to all)；之后 sent to dst link (有 mapping, selectively forwarded)

-*Spanning tree:* 到达所有 vertice 的树，failure, timeout,定期通告
prevent looping within a network topology;
improve resilience when one connection fail

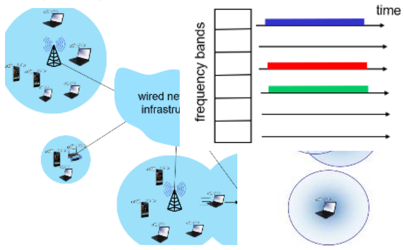
- 1. pick a root (dst, min mac addr); 2.计算 shortest path to root (pick a neighbor wiz lower id)

a.X 宣称自己为 root; b. update,谁 id 最小谁当 node；根据到达 root 的 link 更改 distance



- **Wireless**

-*无线网与有线网的不同*：基站是有线的；
1.信号强度传播时降低(path loss)；2.其他来源(设备的网络频率)的干扰：导致 hidden termina; 3.multipath propagation(自干扰);
-4.无线 higher **bit-error rate**(处理)：提高传输功率，但能耗变大&干扰其他设备；stronger (link-layer) error detection & recovery (损坏帧的链路层重传)；调整 transport protocol(tcp 替代/衍生)



- infrastructure mode & ad hoc mode

前者 Basic Service Set{ access point (AP= base station) 和 host}；后者 hosts only
主机只能选择一个 AP，但能检测到很多个 AP
-Broadcast medium(multiple access 问题): 导致 Hidden terminal problem 或 Signal attenuation（AC 都与 B 交互，但两者没交互）
-对协议的影响：host 连接到基站的 link 用的是 multiple access protocol；理论上：best effort service model(网络层)不影响；TCP 和 UDP(传输层)运行正常。实际上：packet loss due to bit-error (packet loss, delay of link-layer retransmission); tcp interprets loss as congestion, decrease congestion window; delay impairment for real-time traffic; typical less bandwidth available
-*为什么 CSMA/CD 不适用 WiFi?* //CD 适用传统以太网：仅在检测到冲突后才应用 random backoff interval；WLAN 难以感知冲突(can't sense all collisions in any case)，用**/CA**(*Avoid collisions*)
- Reservations(不咋用的机制)：用小 packet 预留通道 for frames
-同子网内 mobility: H1 remains in same IP subnet: IP address can remain same；AP 会更新
-蓝牙: ad hoc，距离近

- **Security（confidentiality integrity authentication）**

-攻击类型：Eavesdropping 窃听: intercept messages；insert messages into connection；impersonation 模拟: fake (spoof) source address in packet；hijacking 劫持：自己假冒为 receiver 或者 sender；service denial: 超载资源让别人也用不了
-具体例子（机密性 完整性 真实性）:

- 应用层：Server-side vulnerabilitie (buffer overflow, sql injection, XSS), spam, phishing, account theft
- 传输层：Abrupt Termination-任何知道 port 和 sequence number 的攻击者都可以破坏 tcp 连接。导致：客户端删除连接并将忽略所有未来的通信。处理：TCP Reset-Firewall, send reset so other side stops trying to send；Content blocking，ISP 阻止 P2P 文件上传
- Connection hijacking: taking over 已经建立的连接(原因: Eavesdropping & Lack of authentication)。导致：client 处理错误的数据，忽略 ACK 数据。处理：TLS（socket 提供 symmetric encryption & cryptographic hashing & public key cryptography）
- 网络层：网络层安全的好处是，传输层以下所有应用程序都可以使用，无需更改；routing architecture。
- IPsec: vpn(physically separate ip site) 用下面方法 private address and domain name，简化步骤搭建，并非真正私网。|允许远程主机连接到防火墙网络，用于远程连接到机构资源，只能使用私有 IP 地址。
  Tunnel mode: encrypt & authenticate datagram（重要）
- **BGP**：1. AS can claim to serve a prefix that don't have a route to 处理 Prefix filtering-让 AS "证明"他们有一条路，提供 whitelists 来过滤广告路由

2. AS forward packets along a route different from advertised
Prefix hijacking 因：BGP 不验证 AS 是否被授权，前缀所有权 Registries 陈旧不准确。导致：DV 导致黑洞 balckhole & Interception(拦截)
解决：S-BGP、BGPsec
DHCP: 攻击者可以监听、伪造响应、接管 DNS、网关（路由器）等核心信息，自己作为中间人插入；*链路层*：窃听/sniffing 嗅探是其他(tcp)攻击的 prereq；*ARP Spoofing 欺骗*：攻击者用自己的 MAC 地址响应；*MAC Flooding*: legitimate packets to be flooded instead of selectively forwarded 处理：手动设置 switches，例如手动配置允许的 MAC 地址，只允许第一个连接的 X