Data Link Layer 物理上相邻两点传输

different link protocols over different links
Responsible for transferring datagram from one node to physically adjacent node over a link.

- Implement: Adaptor or chip

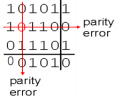
- Data Link layer services

➢ Framing, link access
Encapsulates datagram in frame
"MAC" addresses used in frame headers to identify source, destination

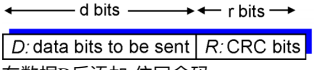
➢ Reliable delivery between adjacent nodes：Error detection & correction w3-link-p12
EDC= Error Detection and Correction bits 和数据拼接在一起进行传输

- Parity checking
single bit
two-dimensional bit

这样就能修改了



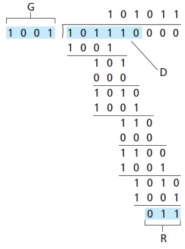
- Cyclic redundancy check 循环冗余校验 CRC



在数据D后添加r位冗余码
首先在原数据D后面添加r个0，相当于左移r位。此时数据长度变为原来的每组d个比特加r即（d+r）位。

1 0 1 1 1 0 0 0 0 D=101110 r=3

然后用该序列除以在计算之前规定的一个长度为（r+1）位的除数G (generator)，根据二进制的**模2 运算**(加减法相同，都是异或运算)，计算出余数R。



相同→0
不同→1

即D* 2^r XOR R

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

这个余数R就会作为冗余码拼接在原数据后面发送出去。
接收方用<D,R>除以G，若余数不等于0则有错误发生。

➢ Flow control
➢ half-duplex and full-duplex
半双工：数据可以在一个信号载体的两个方向上传输，但是不能同时传输。

全双工：通信的双方可以同时发送和接收信息

● Multiple access protocols

Distributed algorithm that determines how nodes share channel. 防止碰撞

➢ 碰撞：

Link types: point-to-point & broadcast (shared wire or medium)

Single shared broadcast channel,

Interference： two or more simultaneous transmissions by nodes

两个或者两个以上结点同时传输

Collision：node receives two or more signals at the same time

➢ Ideal multiple access protocol

broadcast channel： R bps

1. 每个点的传输速率由当下需要传输的点的数量决定(R.....R/M)
2. Fully decentralized: 没有专门用于协调传输的点，不需同步化。

➢ MAC protocols 下面分大块讲这三种 **重要**

- channel partitioning 通道分割

divide channel into smaller “pieces”& allocate piece to node

inefficient at low load

包括：TDMA，FDMA

- random access 随机获取

channel not divided, 需要做到 detect collisions & recover

when node has packet to send, transmit at full channel data rate R. Two or more transmitting nodes: collision

efficient at low load & collision overhead at high load

包括：ALOHA，slotted ALOHA，CSMA，CSMA/CD，CSMA/CA

- “taking turns” 轮流获取

● Channel partitioning MAC protocols

➢ TDMA: time division multiple access

机制： Access to channel in "rounds" , each station gets fixed length slot (length = packet transmission time) in each round

缺陷：Unused slots go idle

➢ FDMA: frequency division multiple access

机制： Channel spectrum divided into frequency bands, each station assigned fixed frequency band

缺陷：Unused transmission time in frequency bands go idle

● Random access protocols

➢ Slotted ALOHA

其基本思想是把时间分成若干个相同的时间片，所有用户在时间片开始时刻同步接入网络信道，若发生冲突，则必须等到下一个时间片开始时刻再发送。该方法避免用户发送数据的随意性，减少了数据冲突，提高了信道的利用率.

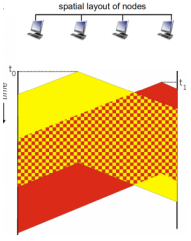
- 特征： frames same size；time divided into equal size slots；nodes start to transmit

only slot beginning；nodes are synchronized；all nodes can detect collision
- if collision: node retransmits frame in each subsequent slot with prob. p until success
- 优点：Single active node can continuously transmit at full rate of channel (利用率高)

Highly decentralized: only slots in nodes need to be in sync
- 缺点：Collisions, wasting slots
Idle slots
Nodes may be able to detect collision in less than time to transmit packet
Clock synchronization
- efficiency: at best: channel used for useful transmissions 37%of time

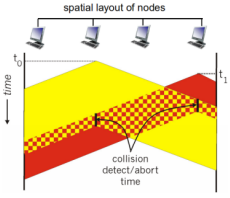
➢ Pure (unslotted) ALOHA
No synchronization, 当node需要发送frame时立即发送。
Collision probability increases, t0时开始发送的包会与在[t0-1,t0+1]开始发送的包产生冲突
- efficiency: 18%

※CSMA是常考的
➢ CSMA (carrier sense multiple access)
- 特征： listen before transmit 传输前检测信道
-> if channel sensed idle -> transmit entire frame 闲时传输
->if channel sensed busy-> defer transmission 忙时停止传输
- Collisions can still occur：传播延迟(propagation delay)，两个节点可能听不到彼此的传输
若发生冲突，就会浪费整个包的传输时间
距离和传播延迟在确定碰撞概率中起着重要作用



CSMA

CSMA/CD



➢ **CSMA/CD** (collision detection) used in Ethernet
- 特征 Collisions detected within short time
colliding transmissions aborted, reducing channel wastage
碰撞传输中止，减少信道损耗
Collision detection 在 wired LANs中较容易实现，在wireless LANs中较难实现
- Eg. Ethernet CSMA/CD algorithm w4-datalink-p22
- efficiency：better performance than ALOHA: and simple, cheap, decentralized

$$efficiency = \frac{1}{1 + 5t_{prop}/t_{trans}}$$

- Frame size **超级无敌螺旋爆炸重要**

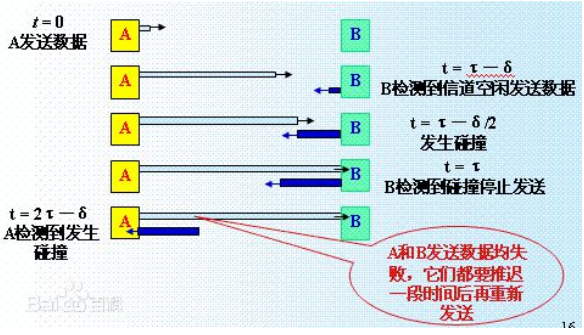
  Collision Window：经过碰撞窗口还没有检测到碰撞，就能够肯定这次发送不会发生碰撞。

  Minimum packet size must be greater than collision window

  min frame size = RTT * transmission rate

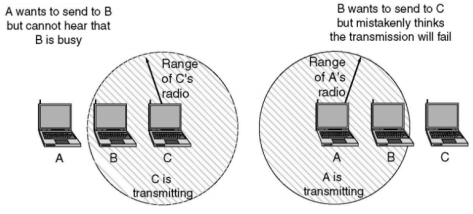
  RTT = 2*propagationTime = 2*(distance/speed)

  Collision window是2倍传输时间的原因：



➢ IEEE 802.11 MAC Protocol: **CSMA/CA** (Collision Avoidance) used in 802.11

- Hidden and Exposed Station problems



- 传输机制

  Sender：

  Sense channel idle -> wait for **DIFS** -> transmit entire frame 不检测冲突no CD

  Sense channel busy -> random backoff time

  Receiver:

  Frame received OK -> wait for **SIFS** -> return ACK

  即流程为：DIFS -> data -> SIFS -> ACK

  DCF = Distributed Coordination Function 分布式协调功能

  DIFS = DCF InterFrame Space DCF帧间空间

  SIFS = Short InterFrame Space 短帧间空间

- 冲突避免CA机制 **很重要**

  small reservation packets:“reserve” channel rather than random access of data frames

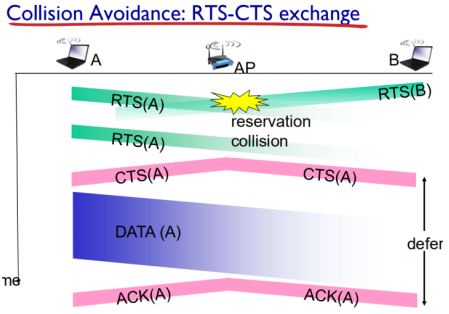
  允许发送方“保留”通道而不是随机访问数据帧:避免长数据帧的冲突

  要会以下的描述过程：

  sender first transmits small **request-to-send (RTS)** packets to base station (BS) using CSMA

  BS broadcasts **clear-to-send CTS** in response to **RTS**

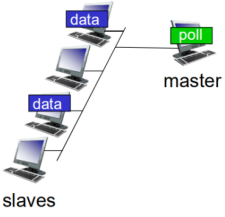
  CTS heard by all nodes, sender transmits data frame while other stations defer transmissions



● "Taking turns" MAC protocols

➢ Polling 轮询

master node "invites"slave nodes to transmit in turn



➢ Cable access network 有线接入网

DOCSIS: data over cable service interface spec 有线电缆数据服务接口规范

FDM over upstream, downstream frequency channels

TDM upstream: some slots assigned, some have contention

● MAC addresses and ARP 会描述

➢ MAC address

- Function: used 'locally' to get frame from one interface to another physically-connected interface (same network, in IPaddressing sense)
- 48 bits e.g.: 1A-2F-BB-76-09-AD 16进制，每个数字4bit
- Each adapter on LAN has unique LAN address 每个接口有一个adapter
- Uniqueness：每台物理设备的MAC地址是唯一的，不随物理地址改变，电子设备生商购买MAC地址以分配给商品。

➢ **ARP: address resolution protocol** 地址解析协议

- ARP table: < IP address; MAC address; TTL>

每一个IP点都有这样一个表 TTL：这个时间后失去该地址对

- 通过IP地址，获取下一跳MAC地址(A->B)：

A broadcasts ARP query packet, containing B's IP address

-> B receives ARP packet, replies to A with its (B's) MAC address, frame sent to A's MAC address (unicast)

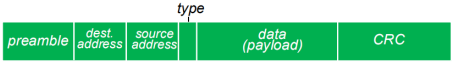
-> A caches (saves) IP-toMAC address pair in its ARP table until information becomes old

- ARP is "plug-and-play": 节点在不需要网络管理员干预的情况下创建它们的ARP表

➢ Addressing: routing to another LAN
W4-datalink-p41 注意四个地址的变化过程
IP src/des 地址一直不变，MAC地址逐跳而变

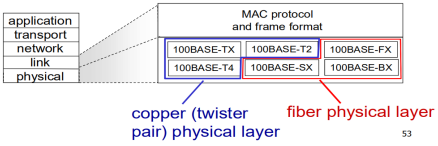
• Ethernet - “主导”有线局域网技术
➢ 特点：简单又便宜，单个芯片可以调配不同速度：10 Mbps–10 Gbps
Connectionless: no handshaking
Unreliable: no ACK/NACK 只能靠上层rdt找回丢失的包
➢ 物理网络结构
Bus(过时了) - Star(active switch in center, no collision)

➢ Ethernet frame structure
发送端adaptor将IP datagram 封装在 frame 里



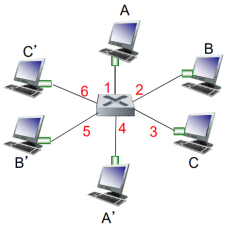
preamble序文：7 bytes，used to synchronize receiver, sender clock rates
addresses地址：6 byte source, destination MAC addresses
这个地址相同才向上层传，否则丢弃
type类型： higher layer protocol (mostly IP)
CRC: cyclic redundancy check at receiver

➢ 802.3 Ethernet standards 许多标准
相同MAC标准和frame格式
不同标准的传输速度不同，对应的物理介质不一定相同



• Ethernet switch **自学机制还是挺重要的**
➢ 特征：
- 只有两层-link-layer device
Store frames, examine incoming frame’s MAC address, selectively forward Ethernet frames, uses CSMA/CD to access segment
- Transparent: 没有adaptor，主机不知道switch的存在
- Plug-and-play, self-learning: Do not need to be configured

➢ 支持多路同时传输 multiple simultaneous transmissions
Host 直接与switch连接，star结构
No collisions; Full duplex 全双工 A->A’ B->B’



➢ Self-learning

Each switch has a switch table：(MAC addr, interface, TTL)
Switch learns which hosts can be reached through which interfaces
每接收到一个frame，switch能记住这个发送方接口能连接到这个主机。
A要传给A’ switch保留(A, 1, 60)
Table中找不到A’的接口-> flood
A’ 返回信息给A-> (A’, 4, 60)

➢ Switches vs. routers

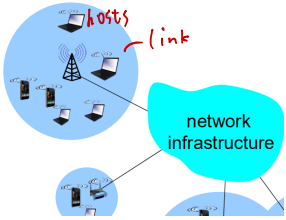
- Router：

  network-layer devices
  compute tables using routing algorithms, IP addresses
- Switch:
  link-layer devices
  learn forwarding table using flooding, learning, MAC addresses

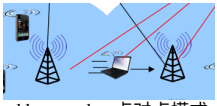
● Wireless network

➢ Elements

wireless hosts
wireless link: connect mobile(s) to base station
base station: typically connected to wired network



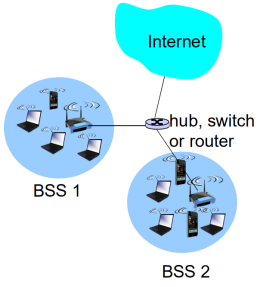
infrastructure mode: Base station connects mobiles into wired network



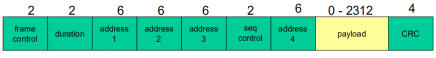
ad hoc mode : 点对点模式 no base stations eg Bluetooth

- IEEE 802.11 Wireless LAN
  802.11a 802.11b 802.11g 802.11n
  all use CSMA/CA for multiple access

- 802.11 LAN architecture
  base station = access point (AP)
  Basic Service Set (BSS) : wireless hosts, AP, ad hoc mode



- 802.11 frame: addressing 无线frame

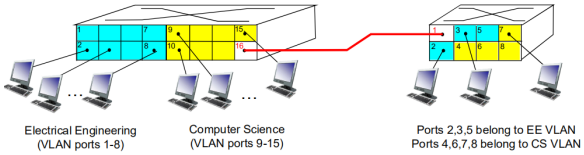


Address 1: 正在接收frame的host或AP的MAC地址
Address 2: 正在传输frame的host或AP的MAC地址
Address 3: 与该AP相连接的router的MAC 地址
Duration: reserved transmission time (RTS/CTS)
Frame control 中有一个type: frame type (RTS, CTS, ACK, data)
Seq control: frame seq #(for RDT)

- VLAN (Virtual Local Area Network)

➢ 特点：
支持VLAN功能的switch(es)可以配置为在单个物理LAN基础设施上定义多个虚拟局域网
port-based: 通过端口分配，一个switch可以当做多个虚拟switch来使用
traffic isolation：一个组内的端口只能互相传送信息
dynamic membership

➢ VLANS spanning multiple switches 多个switch组成的VLAN
trunk port: carries frames between VLANS defined over multiple physical switches
eg port16



● 802.1Q VLAN frame format
在地址后多了两块：2-byte Tag Protocol Identifier & Tag Control Information

● Multiprotocol label switching (MPLS) **重要**
high-speed IP forwarding using fixed length label (instead of IP address)
longest prefix matching -> fixed length identifier
在IP header 前多加了一个MPLS header

➢ MPLS capable routers
Forward based only on label value

➢ MPLS versus IP paths：
use destination and source addresses to route flows to same destination differently
fast reroute: precompute backup routes in case of link failure
- IP routing: path to destination determined by destination address alone
- MPLS routing: path to destination can be based on source and destination address

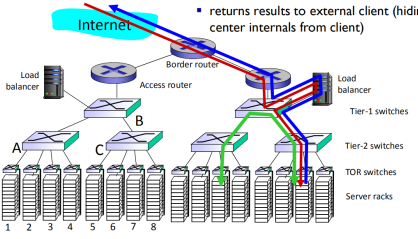
➢ MPLS forwarding tables w4-datalink-p92
同目的地。多路径

in label	out label	dest	out interface
	10	A	0
	12	D	0
	8	A	1

● Data center networks

load balancer: application-layer routing

- receives external client requests
- directs workload within data center
- returns results to external client (hiding data center internals from client)



➢ rich interconnection among switches, racks 通过增加冗余提高可靠性和吞吐量