

---

# Link-layer

---

## Protocol

PPP      Point-to-point Protocol

组帧:

- 将网络层数据报封装到数据链路层帧中
- 可以同时承载任何网络层协议分组(不仅IP数据报)
- 可以向上层实现分用(多路分解)

比特透明传输:

- 数据域必须支持承载任何比特模式

差错检测:(无纠正)

连接活性(connection liveness)检测:

- 检测、并向网络层通知链路失效

网络层地址协商:端结点可以学习/配置彼此网络地址

MAC      Multiple Access Control Protocol 多路访问控制协议

**channel partition 信道MAC划分**

- TDMA      Time Division Multiple Access 时分多址
- FDMA      Frequency Division Multiple Access 频分多址
- CDMA      Code Division Multiple Access 码分多址
- WDMA      Wavelength Division Multiple Access 波分多址

**random access 随机MAC访问**

纯ALOHA 效率  $1/(2e)$

时隙ALOHA 效率  $1/e$

有数据立刻发, 冲突则失效/纯ALOHA协议中其他数据应该避免在 $[t_0-1, t_0+1]$ 区间发送  
**损人不利己的协议**

CSMA      Carrier Sense Multiple Access

载波监听多路访问

**不中断别人**

CSMA/CD      Carrier Sense Multiple Access with Collision Detection

载波监听多路访问/冲突检测 有线网络常用 无线网络由于信号衰减无法判断信号强度  
边发送数据边检测冲突  
边发边听不发不听(听不是听信道有没有数据 是听信道里有没有冲突)  
应用于以太网

**有礼貌的会谈, 有冲突就停止说话, 只能检测到冲突然后停止说话但是不能避免冲突**

CSMA/CA      Carrier Sense Multiple Access with Collision Avoidance

载波监听多路访问/冲突避免

应用于802.11无线局域网

## “taking turns” 轮转MAC协议

Polling 轮询      中心节点轮询开销/延迟/主节点失效则整个网络失效

Token passing 令牌传递 (令牌环)      令牌开销/延时/令牌失效/单点故障(令牌丢失)

令牌是一个特殊帧      数据传输是自生自灭，谁发的数据谁把数据从环上剔除

**实例蓝牙/FDDI(Fiber Distributed Data Interface 光纤分布式数据接口)/令牌环网**

ARP      Address Resolution Protocol 地址解析协议

“即插即用”的协议

LAN中的每一个IP结点(主机/路由器)维护一个ARP表

存储某些LAN结点的IP/MAC地址映射关系

<IP Address, MAC Address, TTL> 经过TTL时间后，映射关系会被遗弃，通常为20min

若源主机查询ARP表没有发现对应目的主机的IP地址/MAC地址映射关系，则使用ARP协议以试图补齐映射：

源主机广播ARP查询分组，其中包含目的主机的IP地址

(链路层广播MAC地址FF-FF-FF-FF-FF-FF)

目的主机接收ARP查询分组，IP地址匹配成功，向源主机应答自己的MAC地址(单播应答)

源主机在其ARP表中，缓存B的IP/MAC映射关系，直至超时

---

## Terminology

HDLC      High-level Data Link Control 高级数据链路控制

HFC      Hybrid Fiber-Coaxial 混合光纤同轴电缆

CRC      Cyclic Redundancy Check 循环冗余校验

NIC      Network Interface Card 网络适配器 网卡

MAC      Media Access Control

MAC地址/LAN地址/物理地址/以太网地址

48位12个16进制数/出厂时固化在网卡ROM中，有时也可以软件设置/标识局域网内一个帧从哪个接口发出，到达哪个物理相连的其他接口

Ethernet      以太网

connectionless

unreliable

差错数据帧直接丢弃，缺失的数据帧依靠高层可靠数据传输协议 e.g. TCP保证

采用二进制指数退避算法的CSMA/CD 在网卡中实现

---

# Network-layer

---

## Protocol

### Forwarding / Routing

IP            Internet Protocol  
              Connectionless

IPv4           Internet Protocol version 4   首部常为20 Byte但首部长度中选项域可变

IPv6           Internet Protocol version 6   首部长度40 Byte不支持路由器分片

IPv6 首部格式加速处理转发，改变首部利于QoS(Quality of Service)区分服务质量的要求

DHCP        *Dynamic Host Configuration Protocol* 动态主机配置协议 (帮助主机自动获取IP地址) **DHCP是一个应用层协议\***

              Client Port 68

              DHCP Server Port 67

**DHCP协议应用于网络层，在应用层实现，并且在传输层使用UDP协议**

              请求报文封装在UDP数据段中，在IP层将DHCP报文通过受限广播地址广播，以太网广播

ICMP        Internet Control Message Protocol 互联网控制报文协议  
              Connectionless

**ICMP是一个网络层协议\***

ICMP差错报告报文包括：差错IP数据报的首部及其封装数据的前8个字节，再加上ICMP报文的前8字节(ICMP报头)

ICMP差错报告报文封装在IP数据报内

IGMP        Internet Group Management Protocol 互联网多播组成员管理协议

**IGMP是一个网络层协议\***

              管理多播组中主机成员的加入和离开

              在IPv6中，该协议功能被集成到ICMPv6中

IGP        Interior Gateway Protocol 内部网关协议 (Autonomous System内部路由协议):

*RIP*            *Routing Information Protocol* 路由信息协议 适用于小规模AS

**RIP是一个应用层协议\***

*Distance-Vector*路由算法 动态规划 距离度量用跳步数  $max = 15 hops$

              每隔30s通过相应报文在邻居间进行RIP通告的交换

              每个advertisement包含最多25个AS内的目的子网列表

              转发表在传输层利用应用层进程route-d(daemon后台驻留程序)进行管理

              通告报文周期性地通过UDP使用Port 520发送，完成网络层功能

              OSPF        Open Shortest Path First 开放最短路径优先 (开放公用)

**OSPF是一个网络层协议\***

Link-State路由算法/路由器之间链路状态分组扩散(每个路由器构造一个链路状态数据库)

每个路由器构造完整的网络(AS)拓扑图/Dijkstra算法

OSPF通告中每个入口对应一个邻居/OSPF通告在整个AS范围泛洪

OSPF报文直接封装在IP数据报中传输

分层OSPF/区域边界路由器/边界路由器/主干路由器

IGRP Interior Gateway Routing Protocol 内部网关路由协议 (specific to Cisco)

EGP Exterior Gateway Protocol 外部网关协议(AS)/已被BGP取代

BGP Border Gateway Protocol 边界网关协议 (实际上使用的域间路由协议)

**BGP是一个应用层协议\* 在传输层使用TCP协议**

路径向量协议

将"Internet"粘合为一个整体的关键

给每个AS提供: eBGP (External Border Gateway Protocol)

从邻居AS获取子网可达信息

iBGP (Internal Border Gateway Protocol)

向所有AS内部路由器传播子网可达性信息

路由器peers利用半永久TCP连接(**Port 179**)交换BGP会话 (选路信息)

属性 AS-PATH (告知承诺对象(AS) advertisement中的prefix会经过的AS序号)

NEXT-HOP (advertisement中告知某个AS若要使用这个承诺的下一跳路由地址)

通告去往不同目的前缀(子网)的路径path vector

---

## Terminology

The Link-State (LS) Routing Algorithm 链路状态路由算法

Dijkstra单源最短路径算法/输出最短路径树转换为路由表

The Distance-Vector (DV) Routing Algorithm 距离向量路由算法

Bellman-Ford方程动态规划算法/坏消息的无穷计数问题/Poisoned Reverse技术

NAT Network Address Translation 网络地址转换

MTU Maximum Transmission Unit 最大传输单元 (涉及IP分组的分片)

CIDR Classless Inter-Domain Routing 无类域间路由

不再硬性规定IPv4地址的地址分类 (**路由聚合**)

最长前缀匹配优先原则 Longest Prefix Match (前缀最长 等价于 子网掩码最长)

可以避免路由聚合中出现的黑洞路由问题

VLSM Variable Length Subnet Mask 可变长子网掩码 (**子网划分**)

---

# Transport-layer

---

## Protocol

UDP      User Datagram Protocol

UDP报文首部 8 Byte

TCP      Transmission Control Protocol

TCP报文首部 20Byte

RDT      Reliable Data Transfer 可靠数据传输原理保证：分组要完整 / 正确 / 有序

TCP可靠数据传输的实现机制

连接管理 TCP三次握手

可靠数据传输原理 rdt信道/滑动窗口协议(GBN/Selective Resend)

流量控制 接收窗口大小大于发送窗口

拥塞控制 慢启动/拥塞避免/快速重传/快速恢复

ARQ      Automatic Repeat reQuest(rdt 2.0中引入的重传机制)

---

# Application-layer

---

## Protocol

HTTP      HyperText Transfer Protocol 超文本传输协议

TCP

Port 80

Stateless

FTP      File Transfer Protocol 文件传输协议

TCP

Control Connection Port 21

TCP

Data Connection Port 20

SMTP      Simple Mail Transfer Protocol 简单邮件传输协议

TCP

Port 25

Stateless

IMAP      Internet Message Access Protocol 邮件访问协议

TCP

Port 143

Pop3      Post Office Protocol—Version 3 邮局协议版本3

TCP

Port 110

Stateless

DNS      Domain Name System 域名系统  
UDP  
Port 53

P2P      Peer to Peer 对等网络

DHCP      Dynamic Host Configuration Protocol 动态主机配置协议  
Client Port 68  
DHCP Server Port 67

RIP      Routing Information Protocol 路由信息协议  
Port 520

BGP      Border Gateway Protocol 边界网关协议  
Port 179

---

## Standardization

PDU      Protocol Data Unit  
Physical-layer      bit  
Link-layer      frame  
Network-layer      packet  
Transport-layer      segment  
Application-layer      message

RFC      Request for Comment

Internet Draft 因特网草案  
Proposed Standard 建议标准  
Draft Standard 草案标准  
Internet Standard 因特网标准

IEEE 802 LAN /MAN Standards Committee 局域网/城域网标准委员会

ISOC      Internet Society 互联网协会 下设  
IAB      Internet Architecture Board 互联网结构理事会:

IETF      Internet Engineering Task Force 互联网工程任务组  
IRTF      Internet Research Task Force 互联网研究任务组

W3C      World Wide Web Consortium 万维网联盟

\*ISO: 产生的主要标准, 会交到下属机构

\*IETF: 负责定义并管理因特网技术的所有方面。包括用于数据传输的IP协议、让域名与IP地址匹配的域名系统（DNS）、用于发送邮件的简单邮件传输协议（SMTP）等。当前IETF正在推动两大标准是互联网协议IPv6和增加一个加密层的DNSSEC

\*W3C: Web技术领域内最具权威的中立的机构

ICANN     The Internet Corporation for Assigned Names and Numbers  
(DHCP服务器管理的IP地址来自上层ISP, ISP的IP地址最终来自ICANN)

NIC           Network Information Center  
InterNIC        ENIC        APNIC