计算机网络知识

1. 网络概述

1.网络将计算机连接起来,互联网将不同的网络连接起来。

2.ISP

3.互联网的组成

1）边缘部分:所有连接在互联网上的主机

2) 核心部分:由大量的网络和连接这些网络的路由器组成

4.主机间的通信方式

1)客户机/服务器(C/S)

2)对等(p2p)

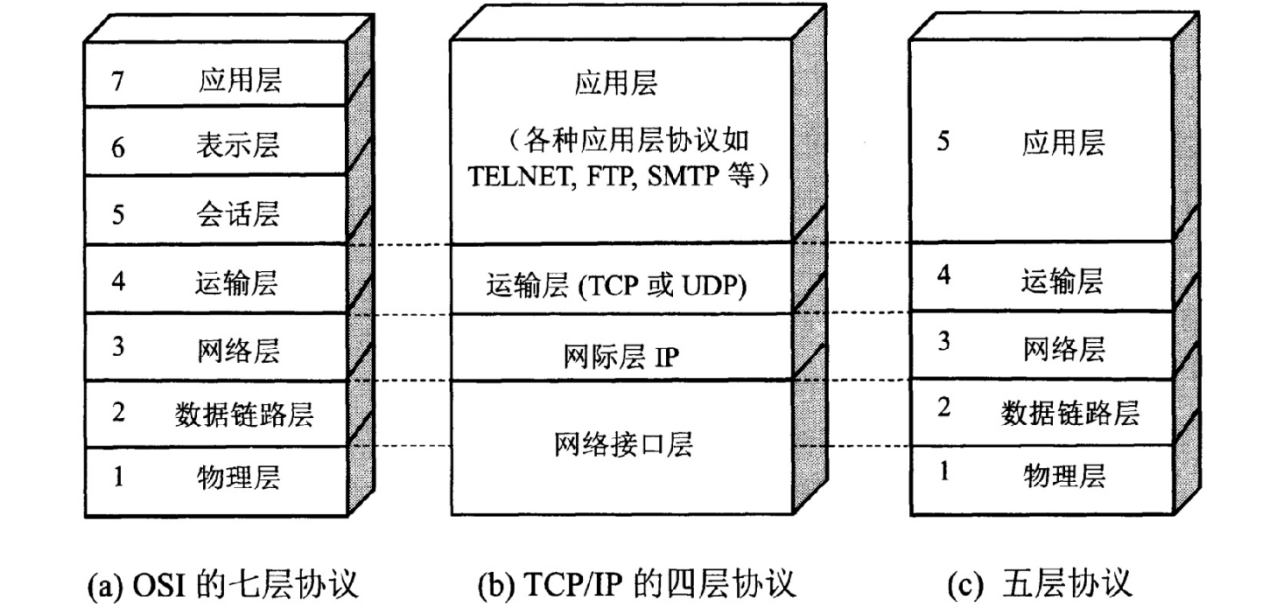
5.数据交换方式

1)电路交换

2)报文交换:存储转发(邮局传递信件的过程)

3)分组交换:也使用了存储转发,但是转发的不是整个报文而是分组，先对报文进行切分，在每个切分的数据之前加上首部由此来构成分组。

6.计算机网络体系结构

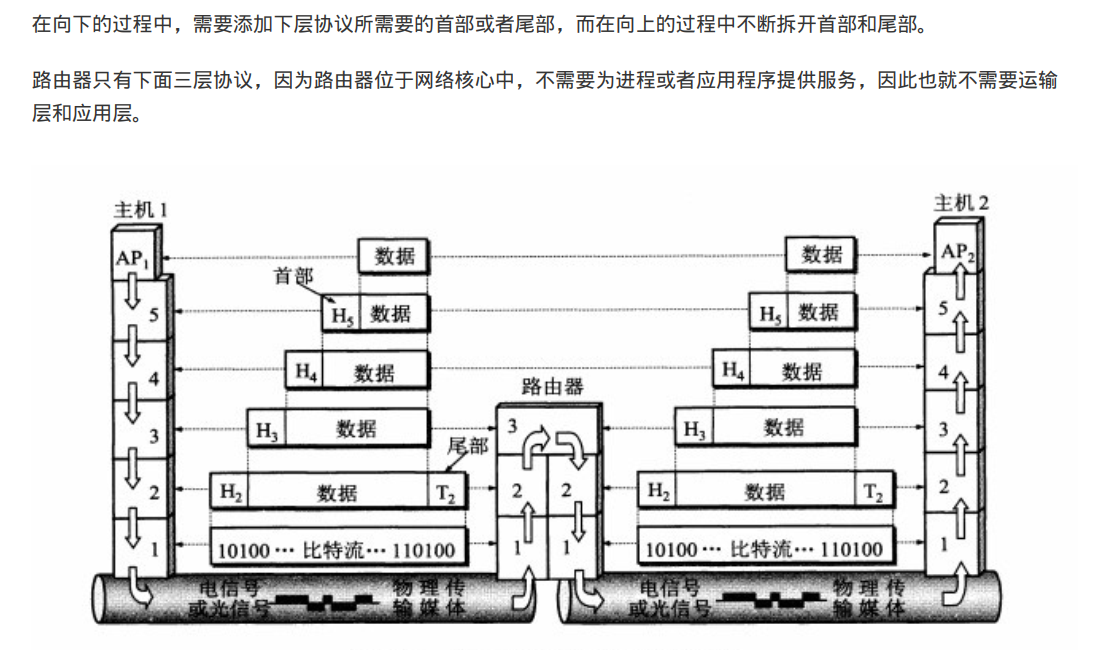


这个模型应该是要烂熟于胸的，同时应该知道为什么会有个7层，5层的划分。

各个层次的主要功能:

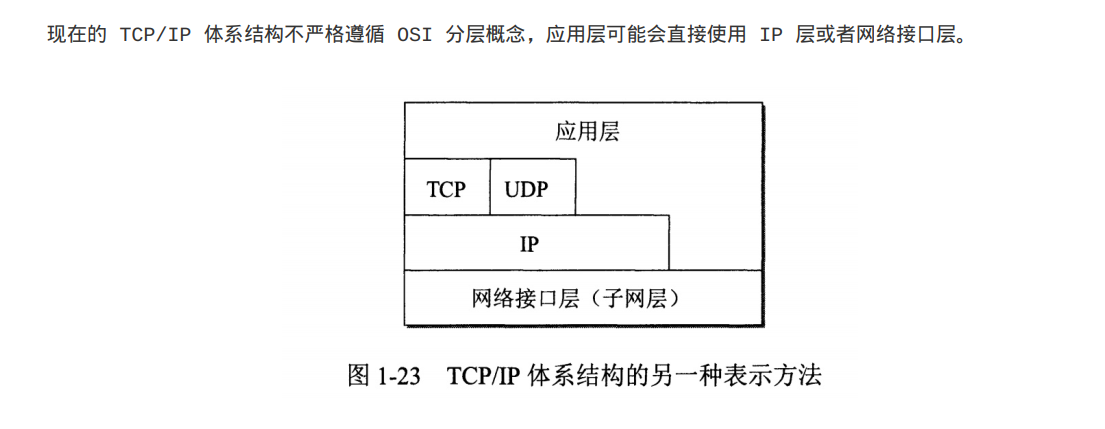
1. 应用层:为特定的应用程序提供数据传输服务,如HTTP,DNS。数据单位为报文。
2. 运输层:提供进程间通用数据传输服务。主要是两种协议:TCP,面向连接,可靠的。数据单位是报文段；UDP:无连接的，最大努力交付的,数据单位是用户数据报。
3. 网络层:为主机之间提供数据传输服务。
4. 数据链路层:主机之间可以有多条链路，链路层为同一链路的结点提供服务。
5. 物理层:在传输媒体上传输数据比特流，通过光信号或者电信号。

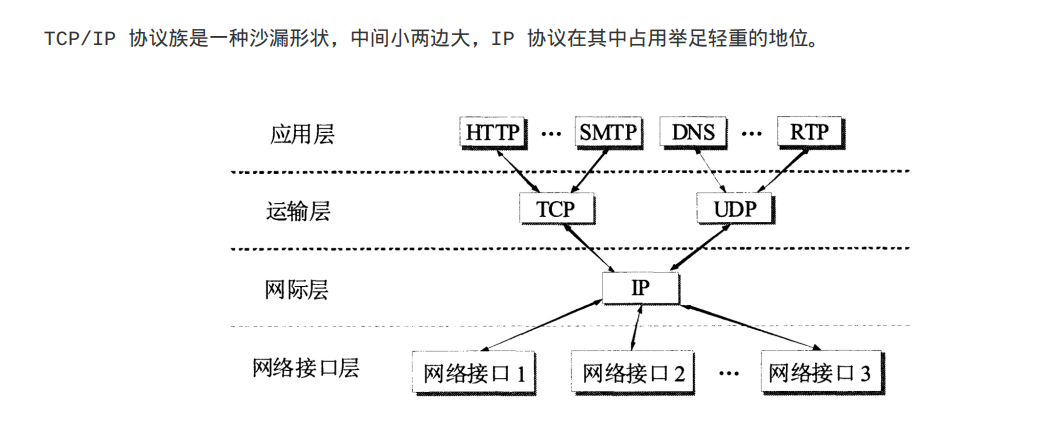
7.数据在各层的传递过程



8.TCP/IP体系结构

它只有4层,相当于5层协议中的物理层和数据链路层合并为网络接口层。





1. 各层详细知识
2. 物理层
3. 通信方式

单向通信:单工通信

双向交替通信:半双工通信

双向同时通信:全双工通信

2) 信道复用技术(了解)

频分复用,时分复用

统计时分复用

波分复用

码分复用

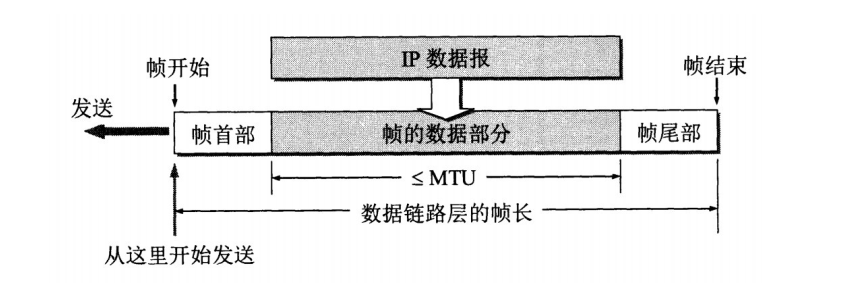
1. 数据链路层
2. 信道分类

点对点的通道:一对一通信方式

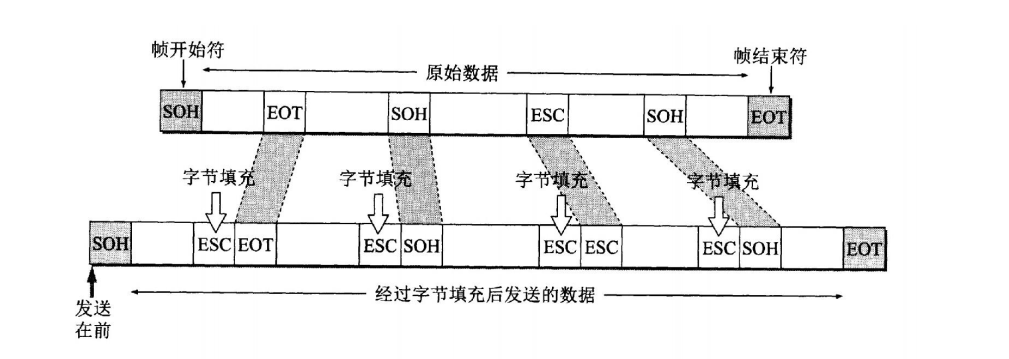
广播信道:一对多通信方式

1. 三个基本问题

封装成帧:将网络层传递下来的分组添加首部和尾部,用于标记帧的开始和结束。



透明传输:帧使用首部和尾部进行定界,如果帧的数据部分含有和首部尾部一致的内容时就会导致偏差，所以需要转义字符。

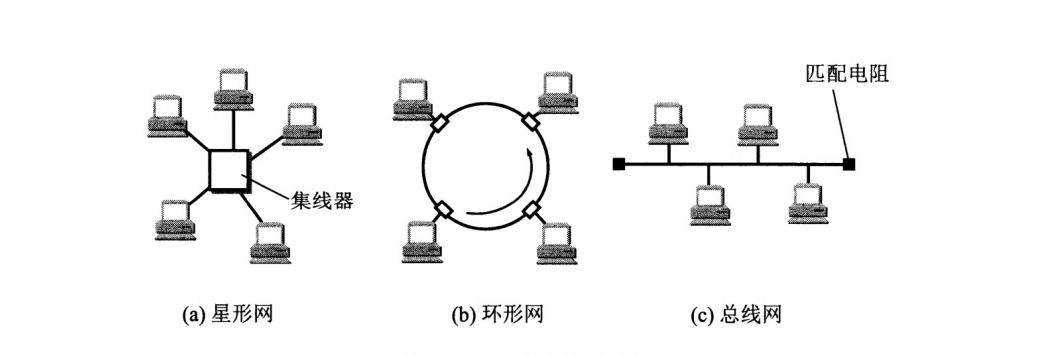


差错校验:目前使用CRC来检查比特差错。

1. 局域网和网络拓扑结构

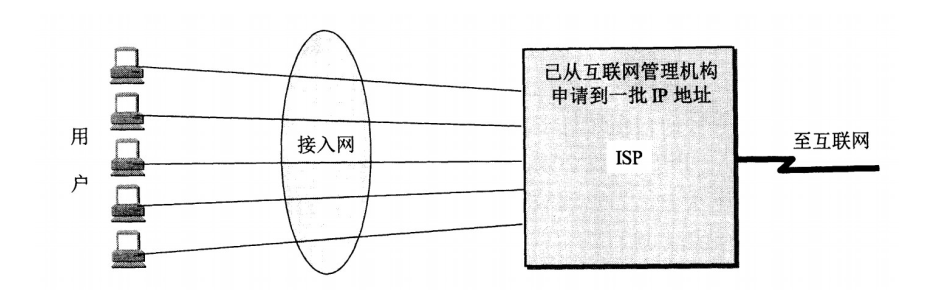
局域网是一种典型的广播信道,特点是为一个单位所拥有。地理范围和站点数目是有限个。

星型结构，环形结构，总线型

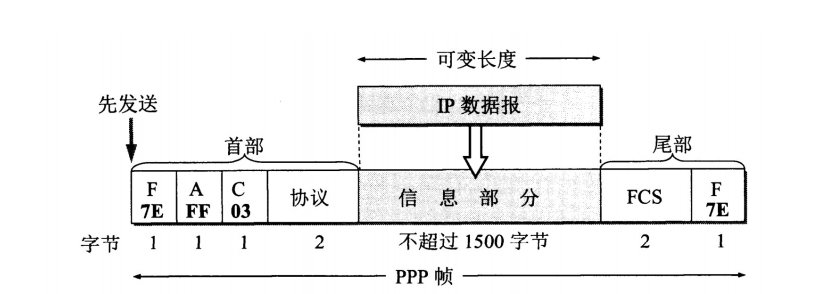


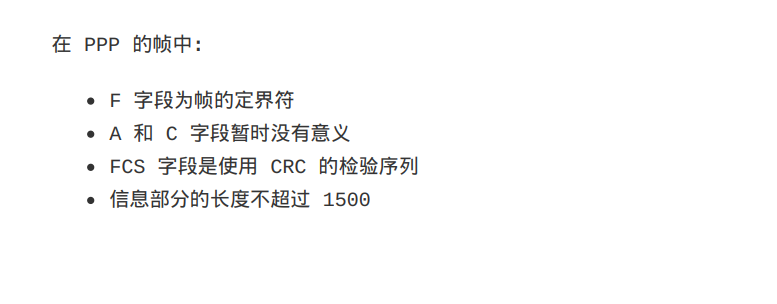
1. PPP协议

点对点的信道中,主要使用PPP协议,是将用户主机连接到ISP的协议。



PPP协议帧的格式





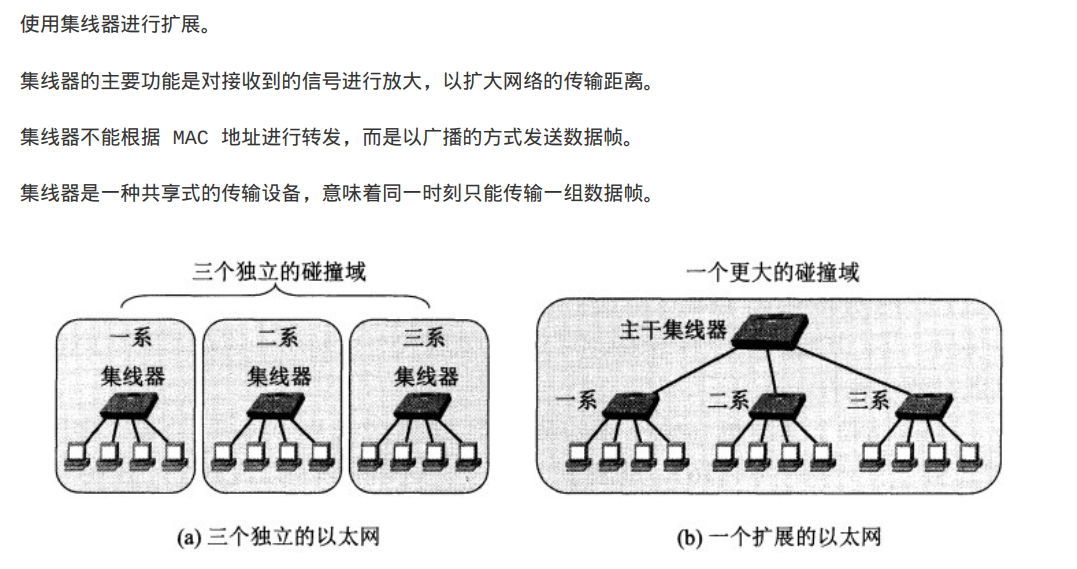
1. CSMA/CD协议

在广播信道上，同一时间只允许一台计算机发送数据。

CSMA/CD表示:载波监听，多点接入/碰撞检测

1. 局域网扩展方案

在物理层进行扩展

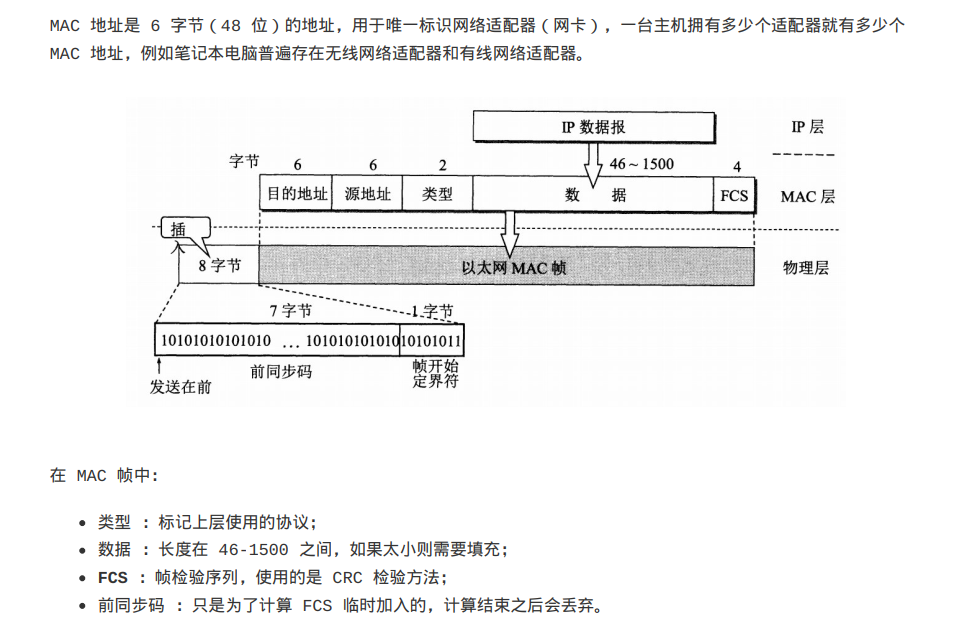


在链路层进行扩展

交换机:存储MAC地址到接口的映射

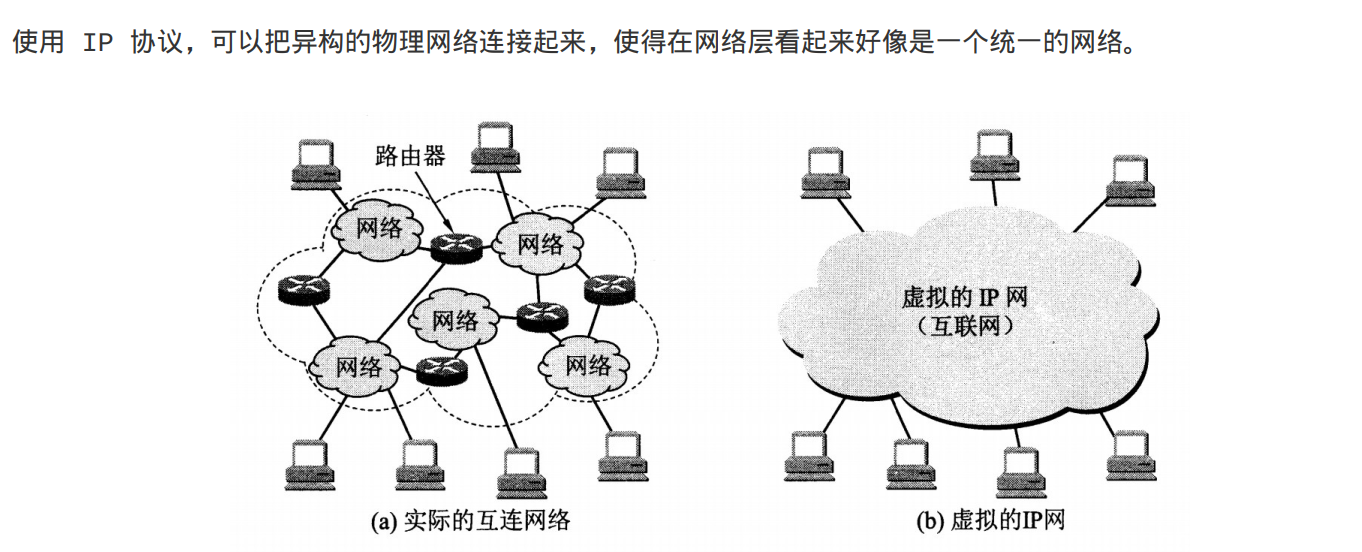
1. MAC层

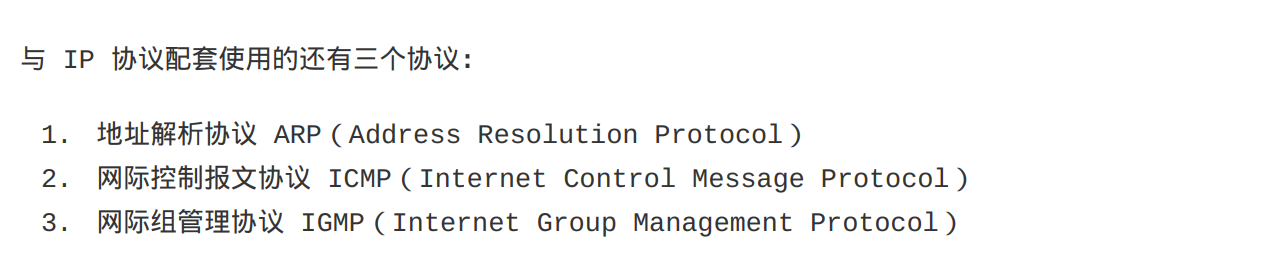
这里MAC层构造了8字节的地址插在以太网MAC帧前。

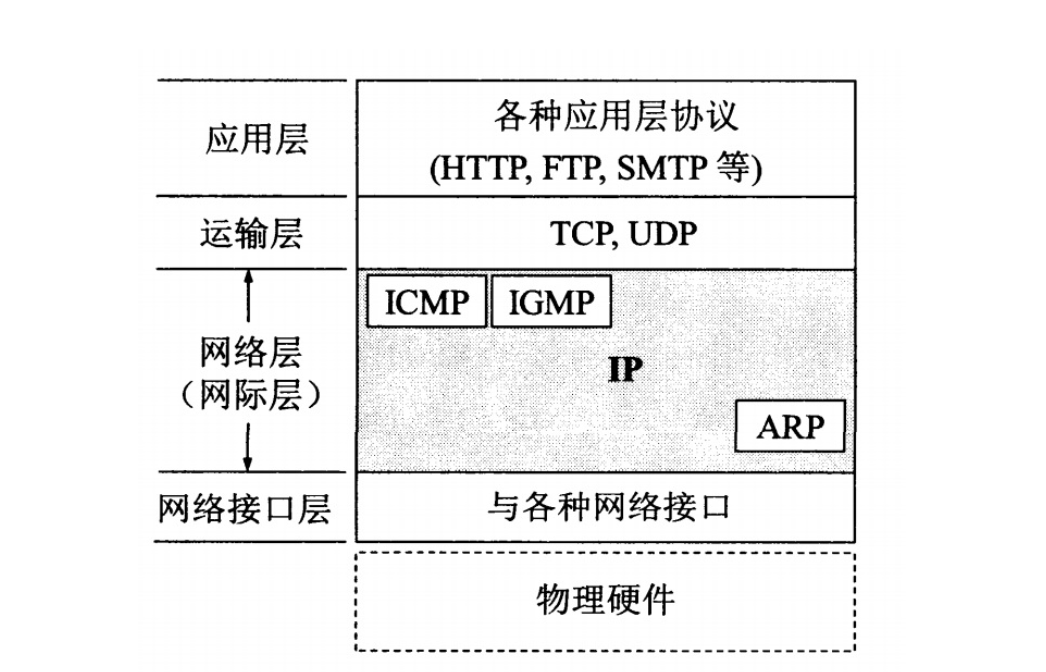


1. 网络层
2. 网际控制协议IP

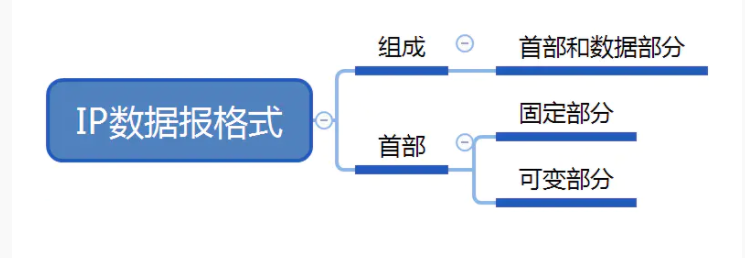
网络层是整个互联网的核心。网络层向上只提供灵活简单，无连接的，尽最大努力交互的数据报服务。

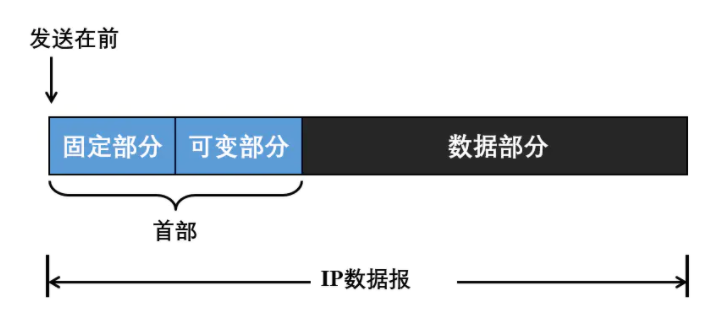




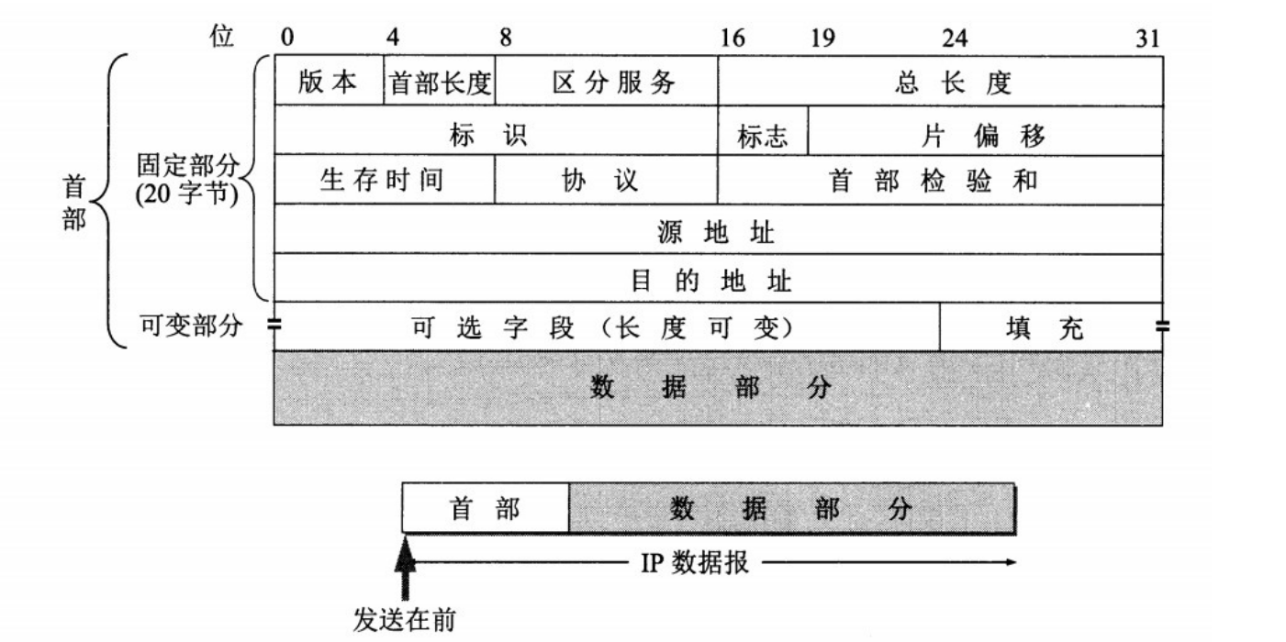


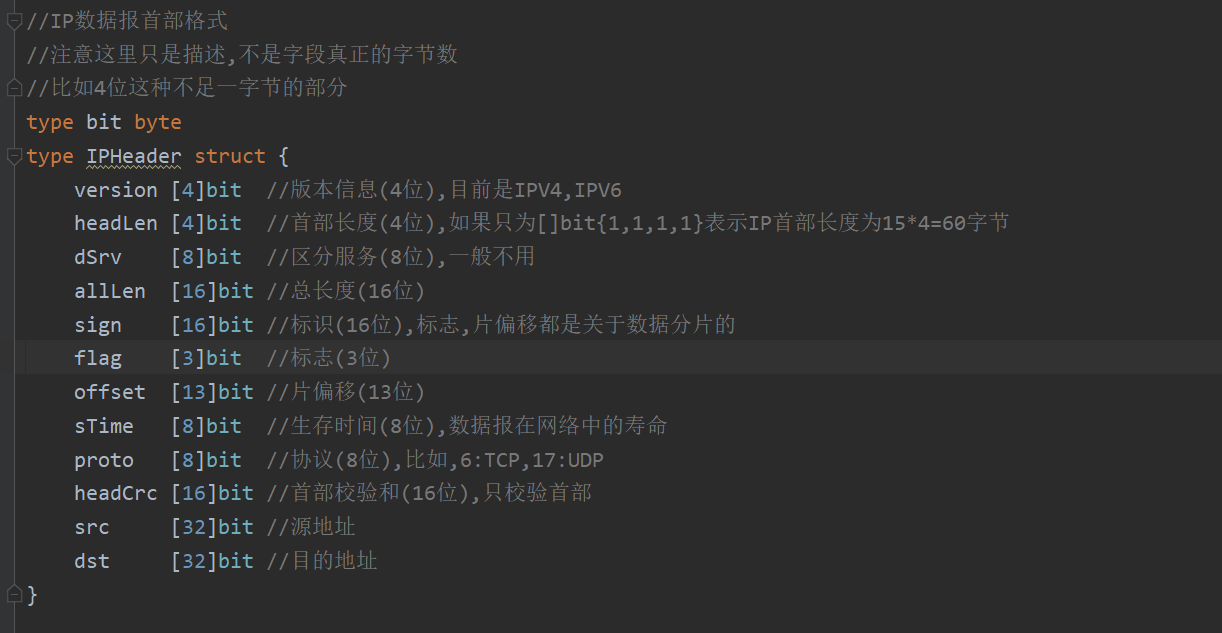
1. IP数据报格式(重点)





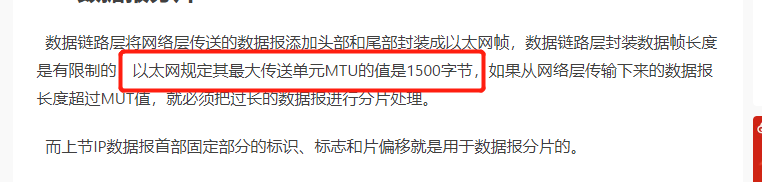
详细描述



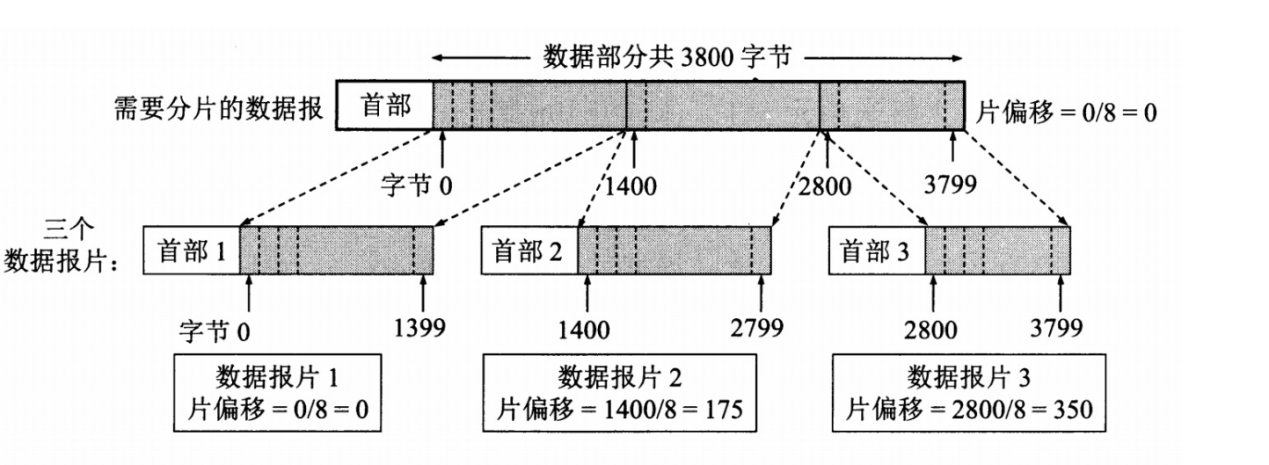


IP数据报分片

分片原因



分片过程



用到的字段



3）IP地址编址方式

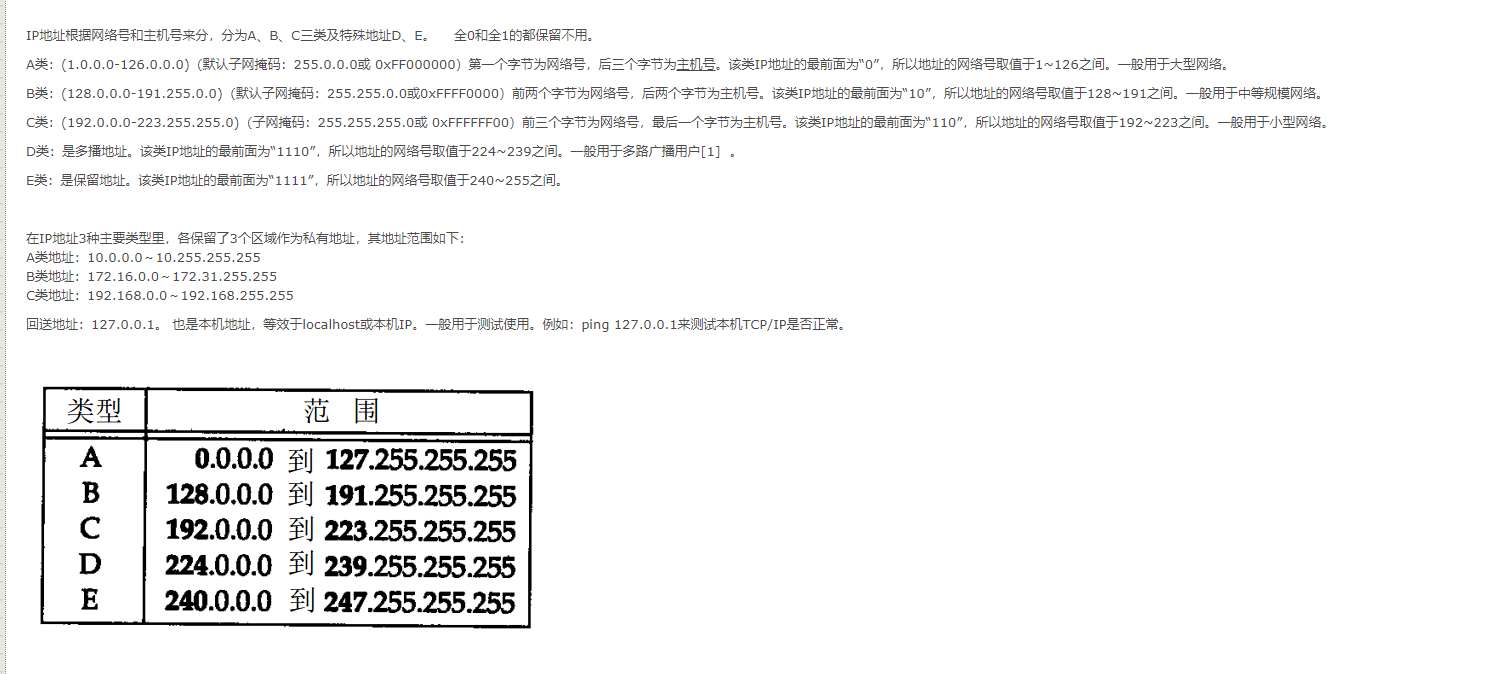
IP地址的编址方式经历了3个历史阶段

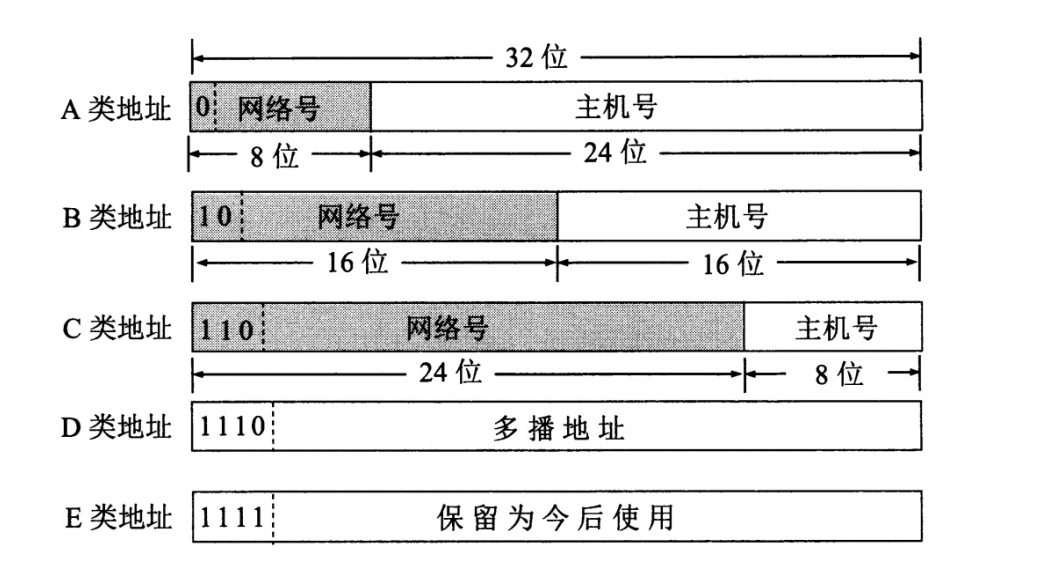
分类 ， 子网划分， 无分类

分类:

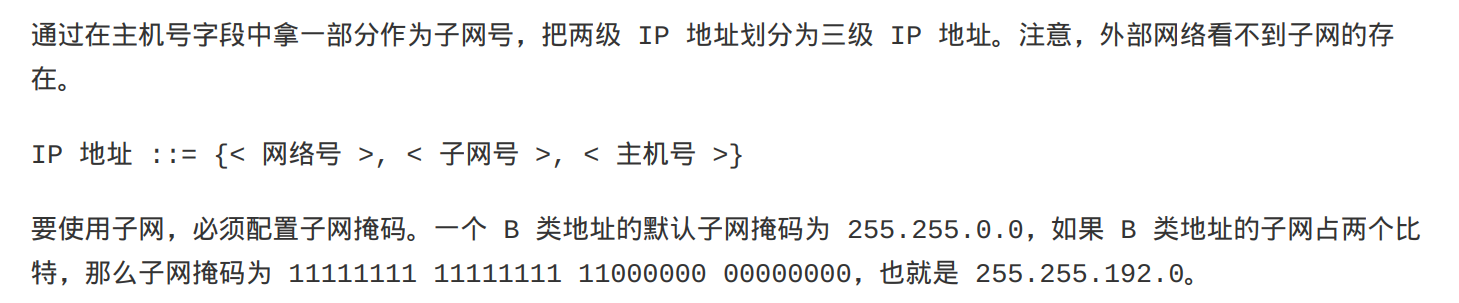
IP地址:网络号,主机号

总共32位,每8位一起看。

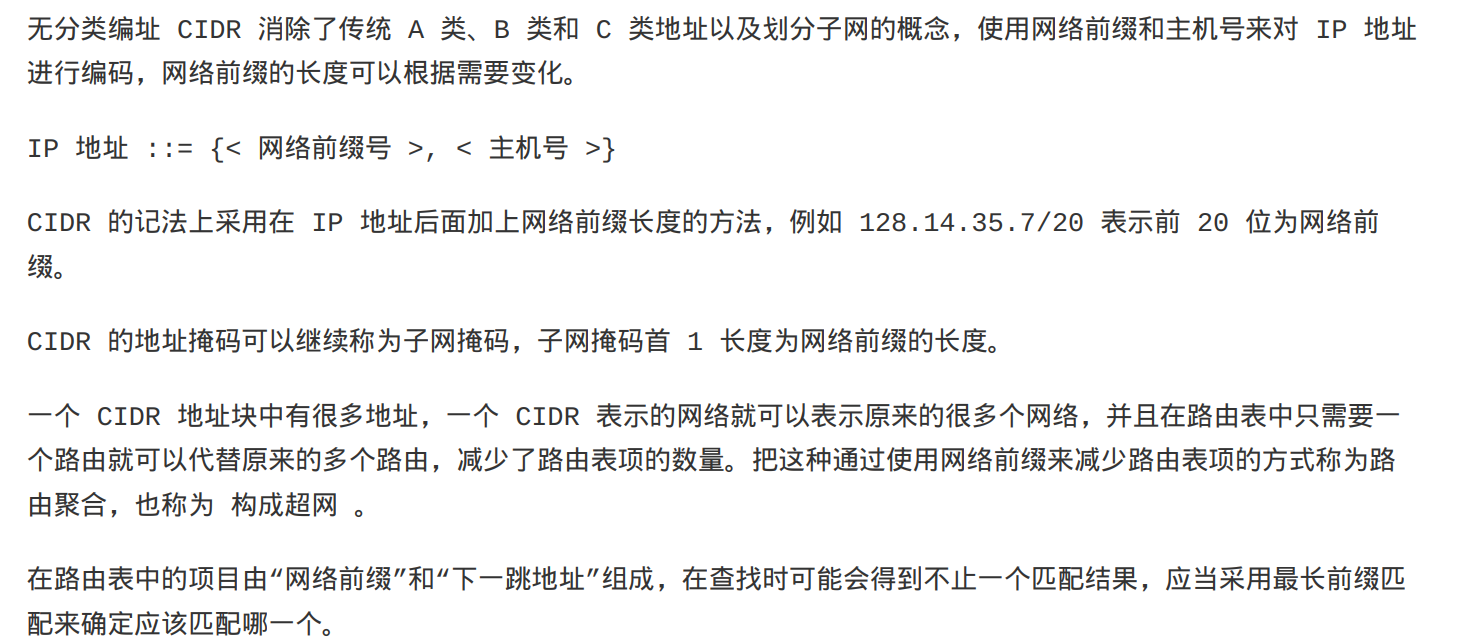




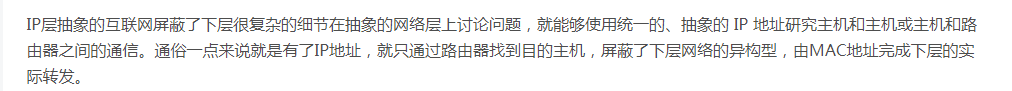
子网划分，要知道这个是什么意思



无分类

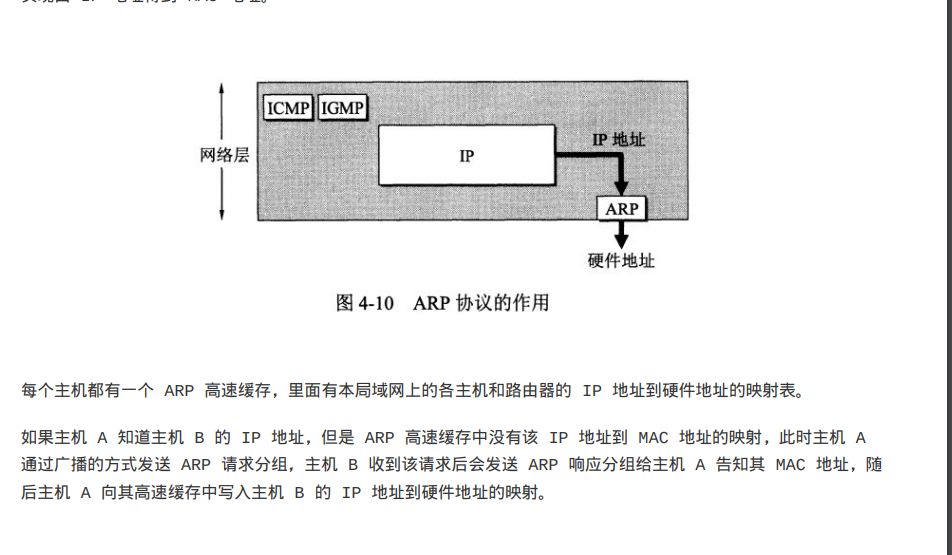


1. IP地址和Mac地址

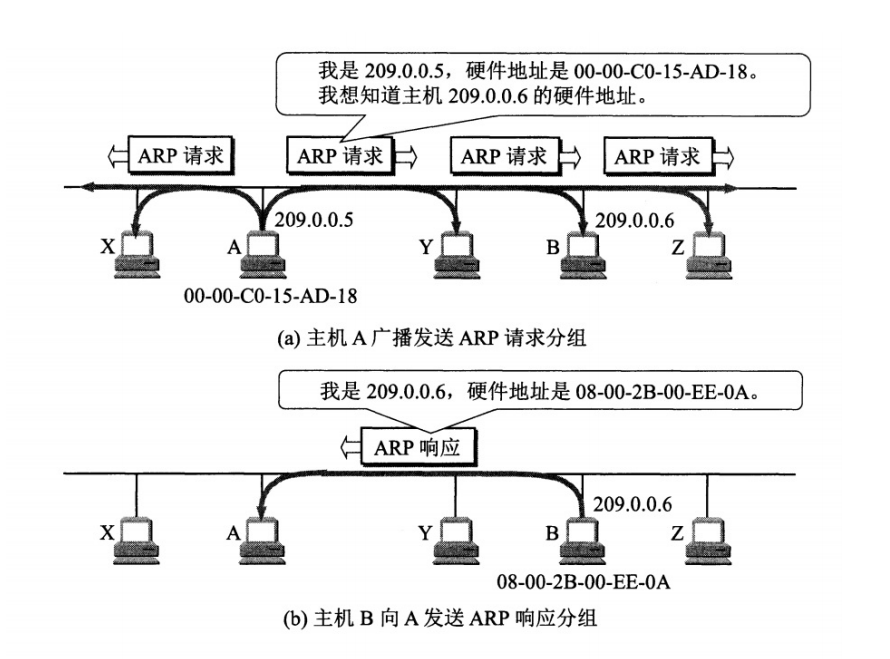


1. 地址解析协议ARP

实现由IP地址得到MAC地址。



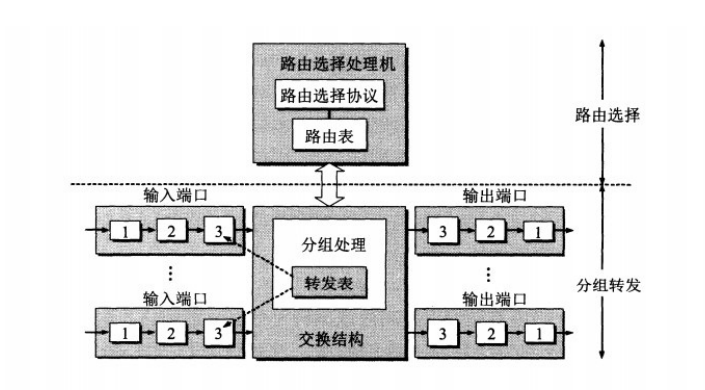
ARP的工作原理大致:广播的方式发送ARP请求分组,然后目的主机响应分组。



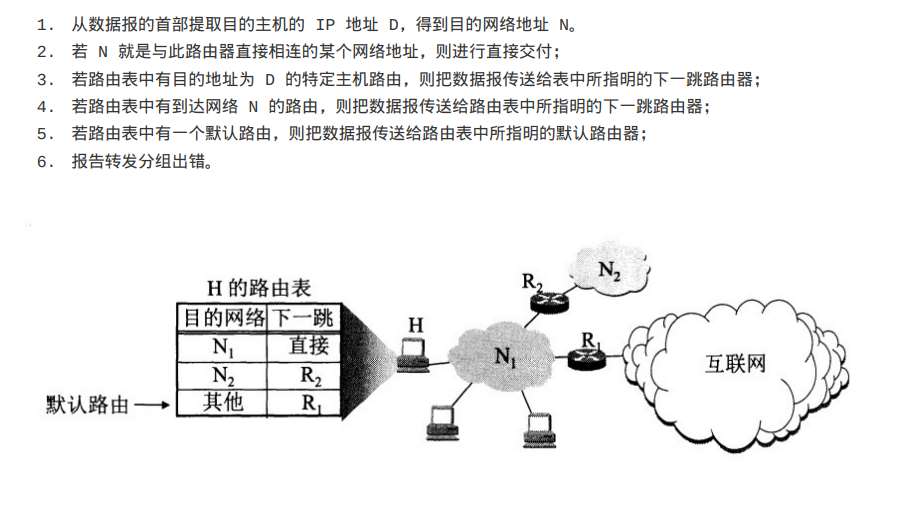
1. 路由器结构

路由器从功能上可以划分为:路由选择和分组转发。

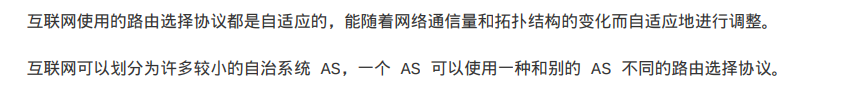
分组转发由三个部分组成:交换结构,一组输入端口,一组输出端口。



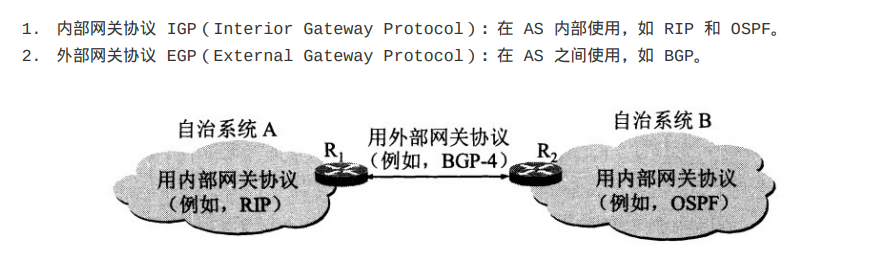
路由器的分组转发流程大致如下



1. 如果目的ip地址的网路地址N就是N1,则直接交付给网络N1
2. 如果是特定的主机路由，则给到下一跳的路由器
3. 默认路由
4. 向源主机报错
5. 路由选择协议



可以将路由选择协议划分为两大类：



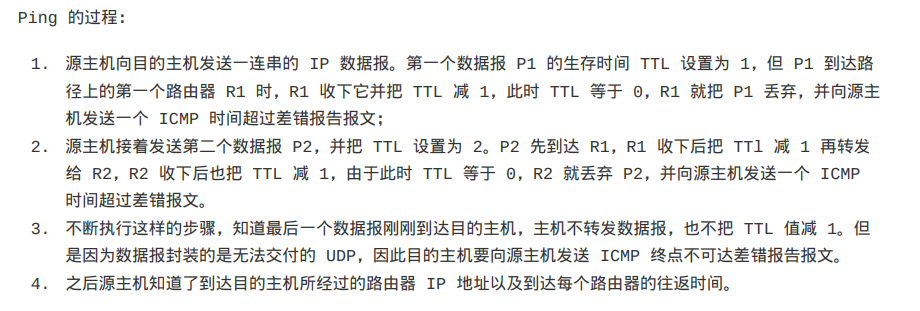
1. 网际控制报文协议ICMP

Icmp是为了更有效地转发IP数据报和提高交付成功的机会。它被封装在IP数据报中，但是不属于高层协议。

分组间探测PING

Ping是ICMP的一个重要应用,用来测试两台主机间的连通性。

Ping发送的IP数据报封装的是无法交付的UDP用户数据报。



用隧道技术实现虚拟专用网络VPN

1. 运输层

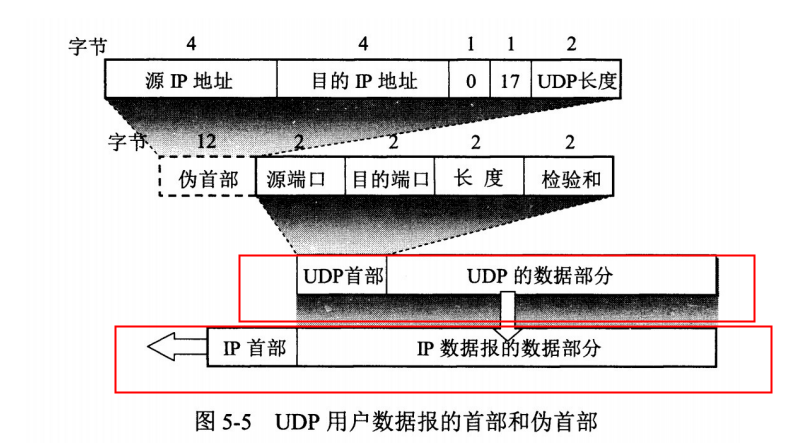
网络层只是把分组发送到目的主机，但真正通信的还是主机中的进程。运输层提供了进程间的逻辑通信，运输层向高层屏蔽了核心细节,使得应用程序看见的好像是在两个运输层实体间有一条端到端的逻辑通信信道。

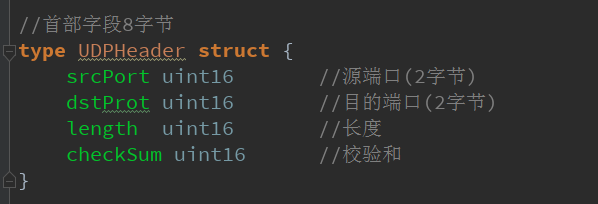
1. UDP和TCP的特点

用户数据报协议UDP是无连接的，尽最大能力交付的，面向报文(对于应用程序传递下来的报文不合并也不拆分,只添加UDP首部)

传输控制协议TCP是面向连接的，提供可靠的交互，有流量控制，拥塞控制，提供全双工的通信，面向字节流（把应用层传递下来的报文看成是字节流，把字节流组织成大小不等的数据块）

1. UDP首部的格式

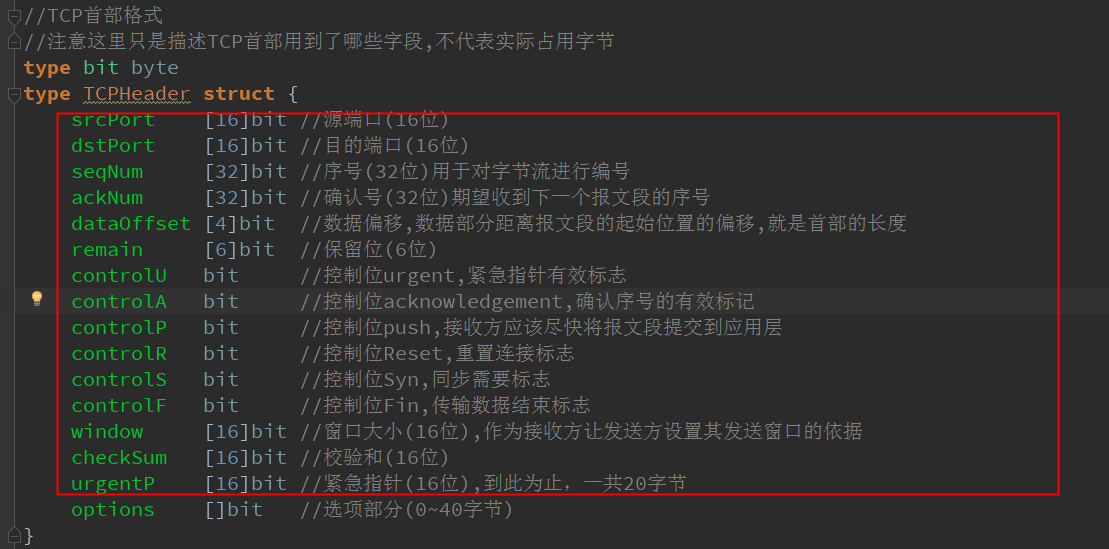


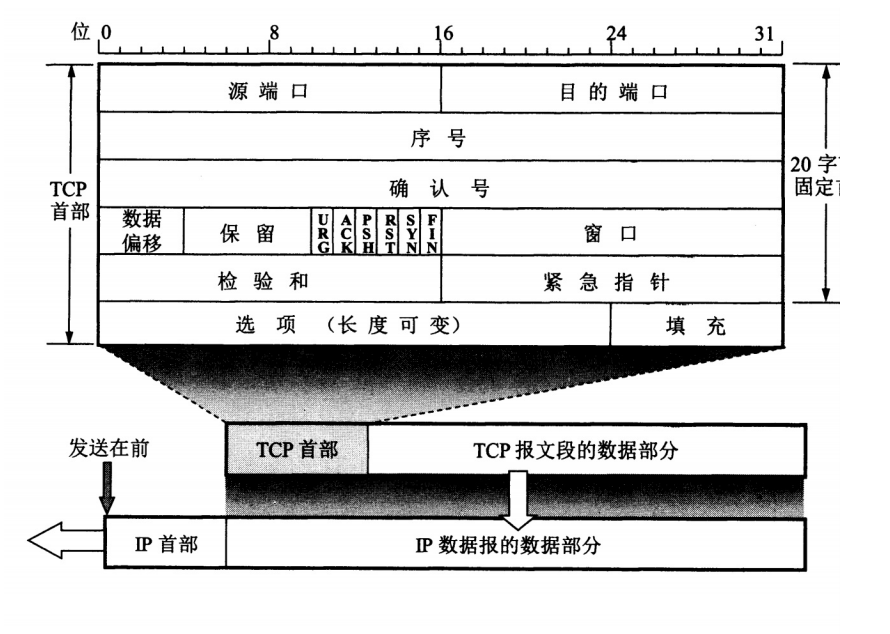


注意还有个12字节的伪首部,是为了计算校验和临时添加的。

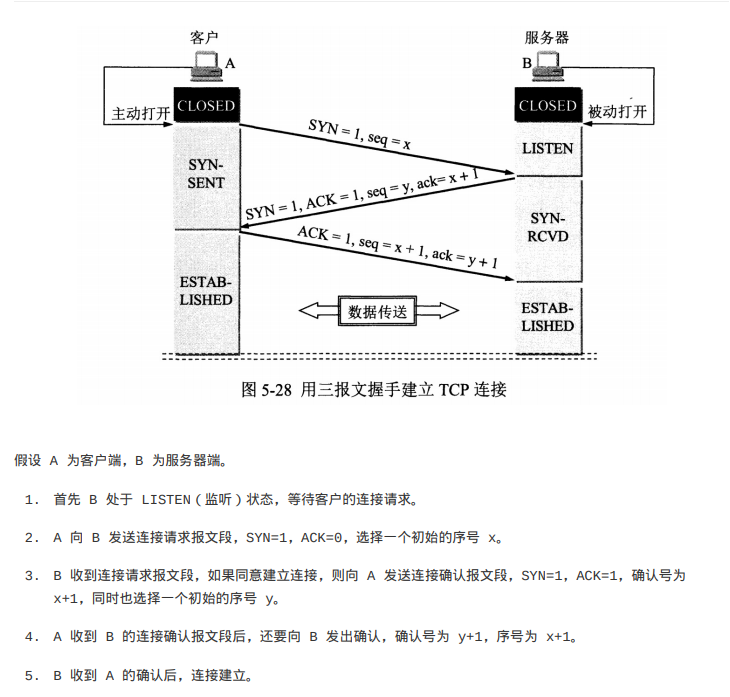
注意红线的的过程。

1. TCP首部格式





1. TCP三次握手



首先需要明确这些字段的使用:

控制位:

确认ACK:当ACK=1时表示确认号字段有效,TCP规定,在连接建立之后所有传送的报文段都需要把ACK置为1

同步SYN:在连接建立时用来同步序号

终止FIN:用来释放一个连接,当FIN=1，表示报文段的发送方数据已经发送完毕，要求释放连接。

传输数据的部分

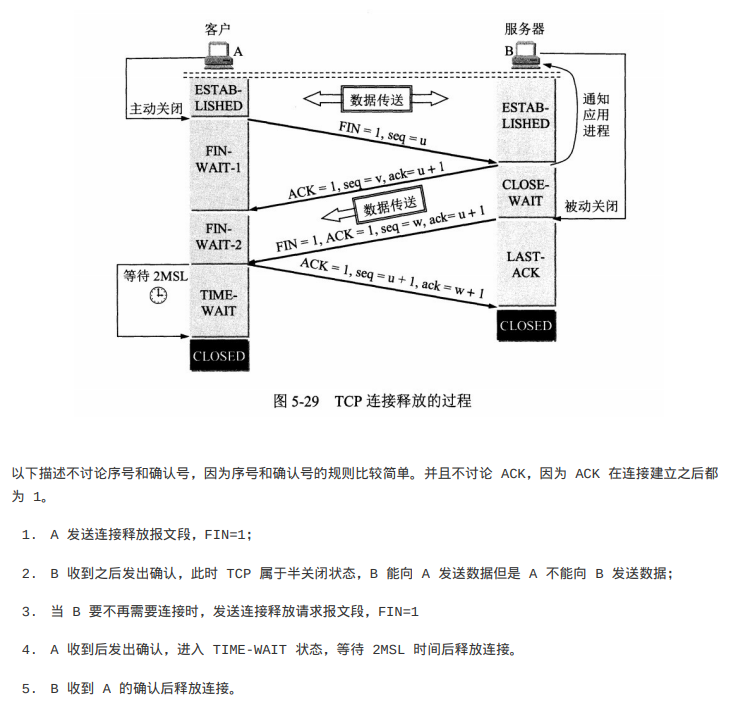
seq:序号,对字节流进行编号

ack:确认序号,希望收到的下一个字节流序号的编号

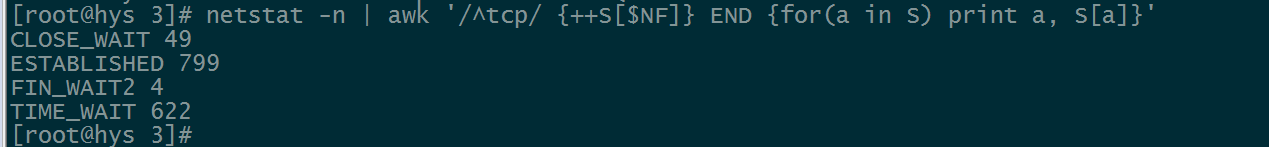
思考:为什么需要三次握手

考虑一下全双工的通信，发送消息的两个方向都要建立连接成功。

1. TCP的四次挥手



MSL:最大分节生命期,这是一个IP数据报在网络中的最长生存时间，RFC建议2min,berkeley的TCP实现传统的30s

思考:当出现大量的TIME\_WAIT是该如何解决

查看系统中time\_wait的数量

Ss -s

分析为什么会产生这种情况

TIME\_WAIT状态过多的危害

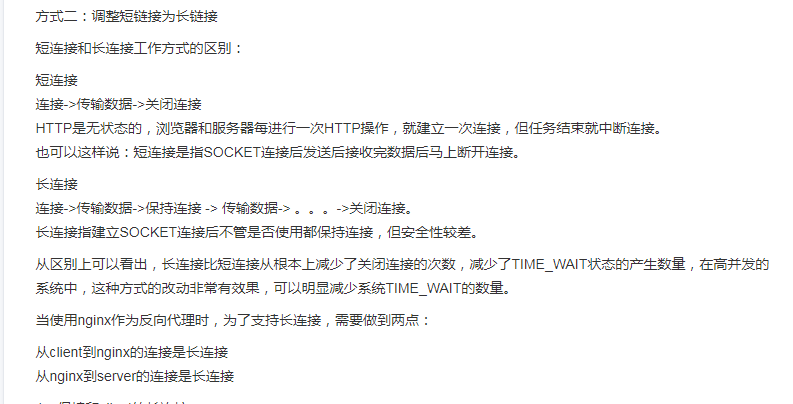
（1）在socket的TIME\_WAIT状态结束之前，该socket所占用的本地端口号将一直无法释放。这也是文章开头的提到问题的一个原因之一。

（2）在高并发（每秒几万qps）并且采用短连接方式进行交互的系统中运行一段时间后，系统中就会存在大量的time\_wait状态，如果time\_wait状态把系统所有可用端口 都占完了且尚未被系统回收时，就会出现无法向服务端创建新的socket连接的情况。此时系统几乎停转，任何链接都不能建立。

（3)大量的time\_wait状态也会系统一定的fd，内存和cpu资源，当然这个量一般比较小，并不是主要危害

如何优化：

1. 调整系统内核参数 
2. 调整短链接为长链接



1. TCP滑动窗口

发送方和接收方各有一个窗口，接收方通过TCP报文段中的窗口字段告诉发送方自己的窗口大小。

TCP是全双工通信,因此每一方的滑动窗口都包含了接收窗口+发送窗口。接收窗口负责处理自己接收到的数据，发送窗口负责处理自己要发送出去的数据。滑动窗口的本质就是维护着几个变量，通过这些变量将TCP处理的数据分为几类。

发送窗口:

|已发送并且确认的字节|已发送但还未收到确认的字节|允许发送但还未发送字节|

|----------------|--------------------|-----------------|

N nextQ n+size

N:发送窗口的起始字节，也就是字节序号N以前的都确认无误了

nextQ:下一次要发送报文的首部Seq字段,[n,nexeQ]都使用过了，但还未ack

n+size:当前发送窗口的最后一个可用字节

接收窗口:

|已经接收交付的字节|已完整有序收到,但还未ack给对端|不能被接收的

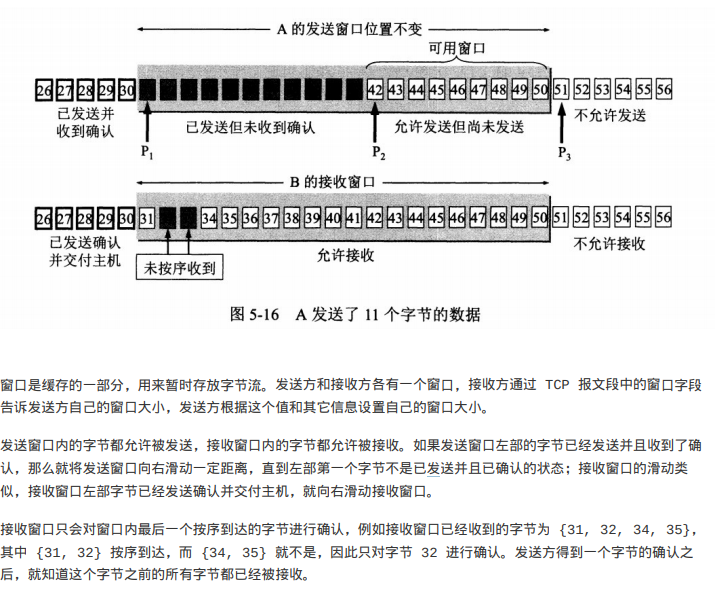
|--------------|------------------------|---------|

J1 J2 J3

J1:之前的字节都都已经接收并确认成功交付

J2:[J1,j2)表示这部分接收到了，但是还没有ack.

J3:接收窗口的最后一个可接收字节，超出J3字节是拒绝的



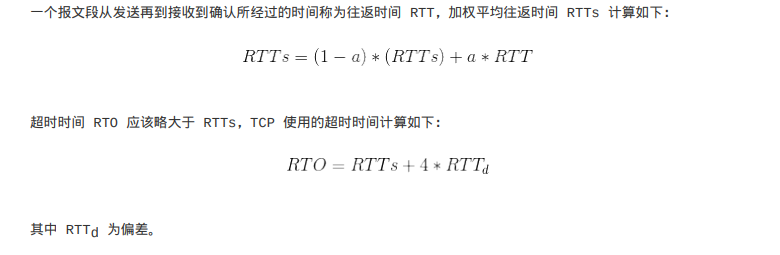
1. TCP流量控制

流量控制是为了控制发送方发送的速率，保证接收方来得及接收。

接收方发送的确认报文中窗口字段可以用来控制发送方窗口的大小，从而影响发送方的发送速率。如将窗口字段设为0,，则发送方不能发送数据。

1. TCP可靠传输

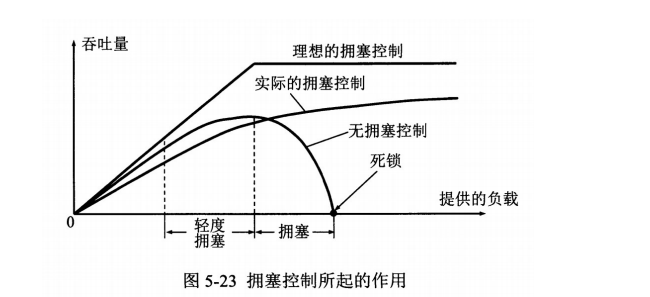
TCP使用超时重传策略来实现可靠传输,如果一个已经发送的报文段再超时时间内没有收到确认，那么就重传这个报文段。



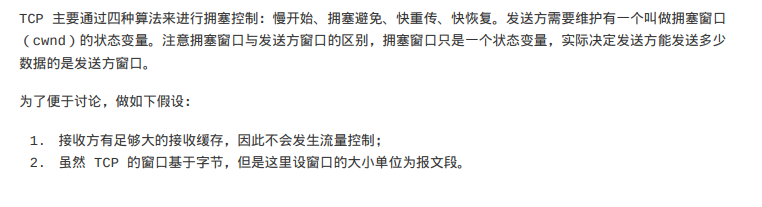
1. TCP拥塞控制

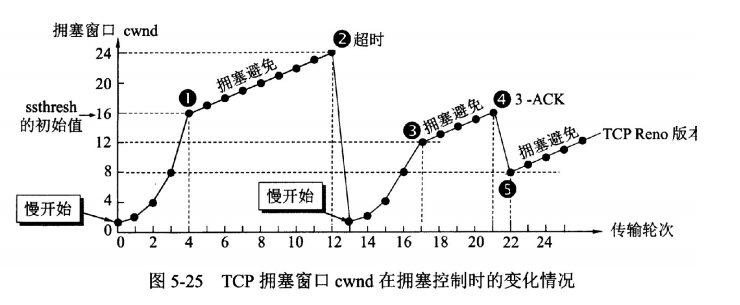
如果网络出现拥塞,分组将会丢失。此时发送方就会继续重传，这样会导致拥塞程度越来越高，不加管控会形成拥塞。因此在出现拥塞时，应当控制发送方的速率，这一点和流量控制很像。但是出发点不同。

区别:流量控制是为了让接收方来得及接收，拥塞控制是为了降低整个网络的拥塞程度。



拥塞窗口的含义：





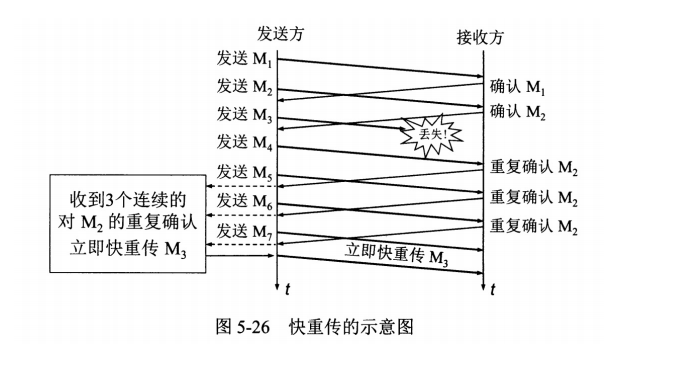
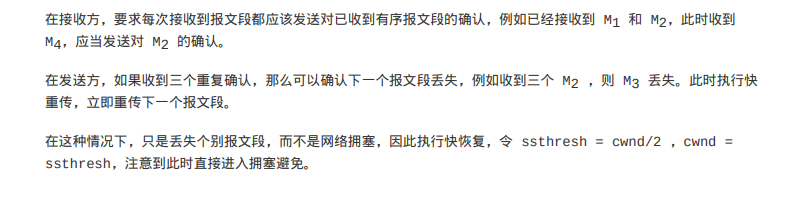
1. 慢开始与拥塞避免

最初,cwnd=1,发送方只能发送1个报文段,当收到确认后，cwnd就会加倍，之后发送的是就是2,4,8.

一旦开始每个轮次都会翻倍发送，这样cwnd增长就会很快。从而就会导致发送方发送速度过快,网络拥塞就会更高。设置一个慢开始的阈值ssthresh,cwnd>=ssthresh时就会进入拥塞避免阶段

进入拥塞避免阶段后，没轮次cwnd+1增长。如果出现超时，则让ssthresh=cwnd/2,重新开始从1~新的sshtresh

1. 快重传与快恢复

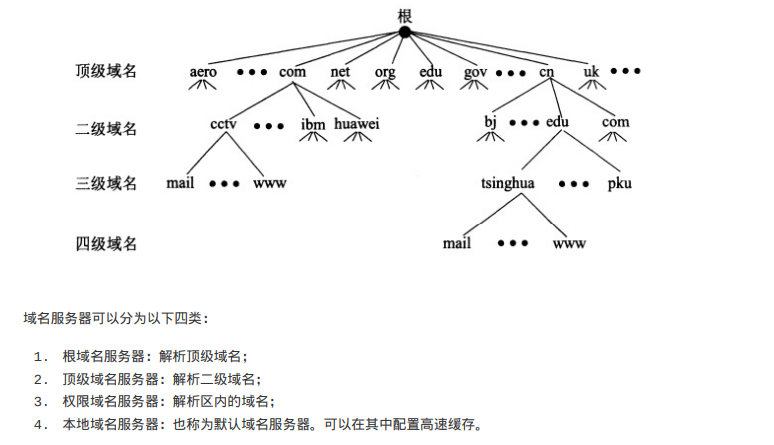


1. 应用层
2. 域名系统DNS

将主机名解析为IP地址,被设计成分布式系统。

层次结构



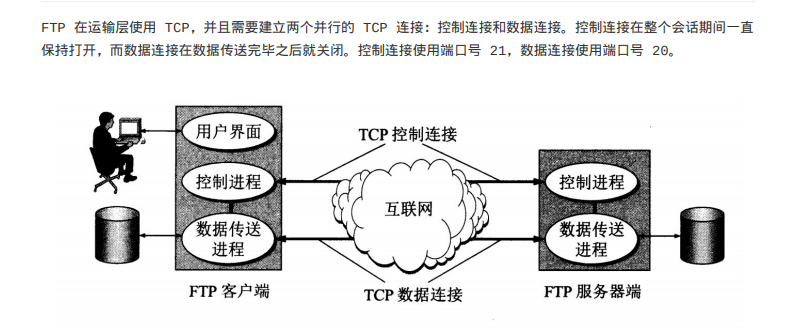


解析过程

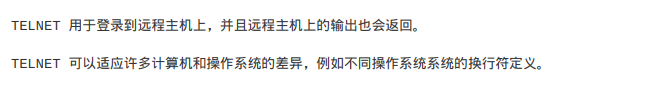
主机向本地域名服务器解析的过程采用递归。

本地域名服务器向其他域名服务器解析采用递归和迭代两种方案。

1. 文件传输协议FTP



1. 远程终端协议TELNET

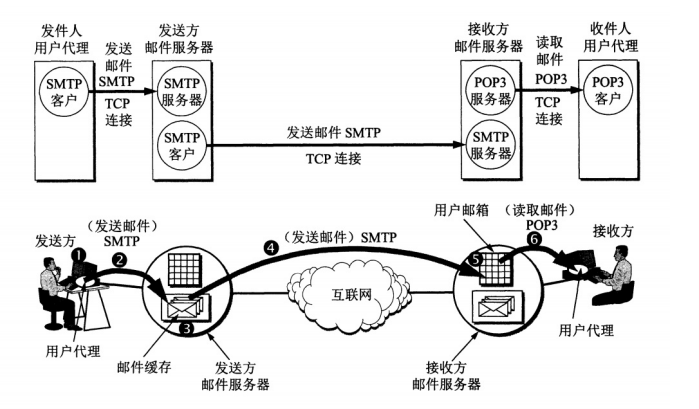


1. 电子邮件协议

电子邮件系统的组成:用户代理,邮件服务器,邮件发送/读取协议。

常用的发送协议:SMTP

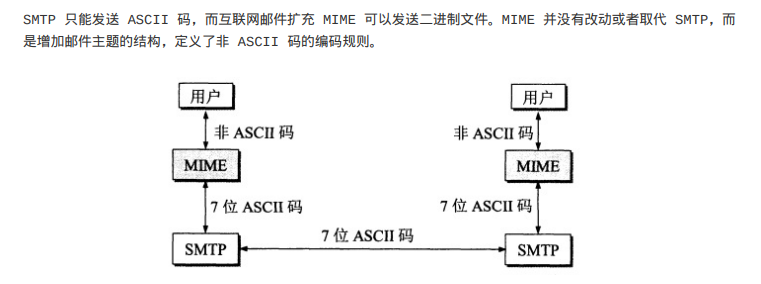
常用的读取协议:POP3,IMAP



邮件协议的特点:

POP3:用户从服务器上读取邮件，就将改邮件删除。

IMAP:不手动删除邮件，则服务器上邮件也不会删除。

SMTP:

1. 动态主机配置协议DHCP
2. 点对点传输P2P