

link: null
title: 珠峰架构师成长计划
description: Open System Interconnection,适用于所有的网络
keywords: null
author: null
date: null
publisher: 珠峰架构师成长计划
stats: paragraph=134 sentences=195, words=805

1. 协议的概念和作用

- 为了让计算机能够通信, 计算机需要定义通信规则, 这些规则就是协议
- 规则是多种, 协议也有多种
- 协议就是数据封装格式+传输

1.1 服务+协议实现数据传输

- 摩尔电码 (<https://baike.baidu.com/item/%E6%91%A9%E5%B0%94%E6%96%AF%E7%94%B5%E7%A0%81/1527853?fr=aladdin>)

2. OSI七层模型

Open System Interconnection,适用于所有的网络

- 分工带来效能
 - [松江黄金大劫案 \(http://www.360doc.com/content/17/0616/08/3404269_663547705.shtml\)](http://www.360doc.com/content/17/0616/08/3404269_663547705.shtml)
- 将复杂的流程分解为几个功能相对单一的子进程
- 整个流程更加清晰, 复杂问题简单化
- 更容易发现问题并针对性的解决问题
 - 应用层(Application) 提供网络与用户应用软件之间的接口服务
 - 表示层(Presentation) 提供格式化的表示和转换数据服务, 如加密和压缩
 - 会话层(Session) 提供包括访问验证和会话管理在内的建立和维护应用之间通信的机制
 - 传输层(Transmission) 提供建立、维护和取消传输连接功能, 负责可靠地传输数据(PC)
 - 网络层(Network) 处理网络间路由, 确保数据及时传送(路由器)
 - 数据链路层(DataLink) 负责无错传输数据, 确认帧、发错重传等(交换机)
 - 物理层(Physics) 提供机械、电气、功能和过程特性(网卡、网线、双绞线、同轴电缆、中继器)

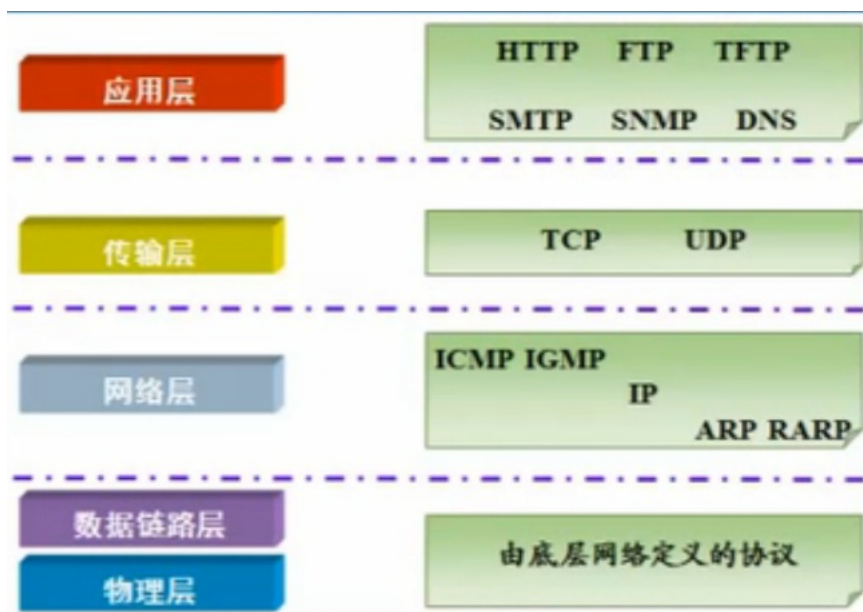
3. TCP/IP参考模型

- TCP/IP是传输控制协议/网络互联协议的简称
- 早期的TCP/IP模型是一个四层结构, 从下往上依次是网络接口层、网络层、传输层和应用层
- 后来在使用过程中, 借鉴OSI七层参考模型, 将网络接口层划分为物理层和数据链路层, 形成五层结构



常用协议

- TCP/IP协议被称为传输控制协议/互联网协议, 又称网络通讯协议
- 是由网络层的IP协议和传输层的TCP协议组成, 是一个很大的协议集合
- 物理层和数据链路层没有定义任何特定协议, 支持所有的标准和专用的协议
- 网络层定义了网络互联也就是IP协议
 - 网际协议IP(Internet Protocol) 负责主机和网络之间寻址和路由数据包
 - 地址解析协议ARP(Address Resolution Protocol) 获得同一物理网络中的硬件主机MAC地址
 - 反向地址转换协议(Rreverse Address Resolution Protocol) 允许局域网的物理机器从网关服务器的ARP表或者缓存上请求其IP地址
 - 网际控制消息协议ICMP(Internet Control Message Protocol) 发送消息, 并报告有关数据包的传送错误
 - 互联网组管理协议IGMP(Internet Group Management Protocol) IP主机向本地多路广播路由器报告主机组成员
- 传输层定义了TCP(传输控制协议)和UDP(用户数据报)协议
- 应用层定义了HTTP(超文本传输协议)、FTP(文件传输协议)、DNS(域名系统)等协议



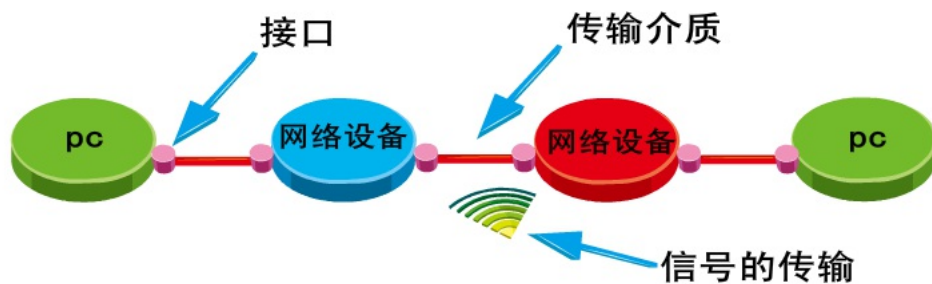
3.1 网络接口层 <#>

网络接口层是TCP/IP模型的最底层，负责接收从上一层传来的数据报并将数据报通过底层的物理网络发送出去，比较常见的就是设备的驱动程序，此层没有特定的协议。网络接口层又分为物理层和数据链路层。

3.1.1 物理层 <#>

计算机在传递数据的时候传递的都是0和1的数字，而物理层关心的是用什么信号来表示0和1，是否可以双向通信，最初的连接如何建立以及完成连接如何终止，物理层是为数据传输提供可靠的环境。

- 为数据端设备提供传送数据的通路
- 传输数据
 - 激活物理连接，在连接的设备之间连接起来形成通路
 - 传输数据，关心如何打包数据和传输速度
 - 关闭物理连接



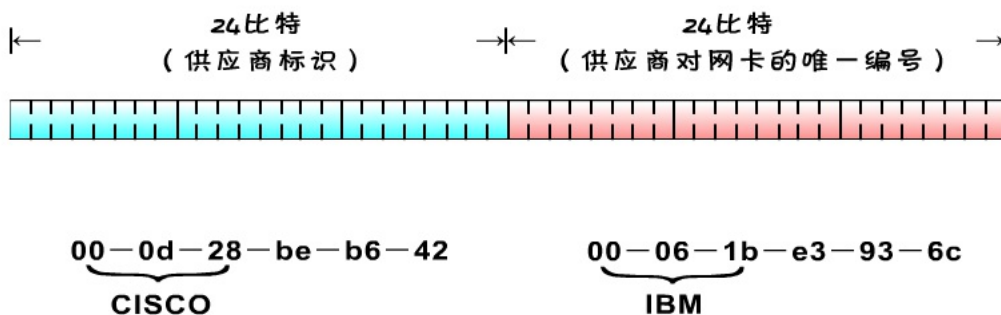
3.1.2 数据链路层 <#>

数据链路层位于物理层和互联网层之间，用来向网络层提供数据，就是把源计算机网络层传过来的信息传递给目标主机。

- 如何将数据组合成数据帧(Frame)，帧是数据链路层的传输单位
- 数据链路的建立、维护和拆除
- 帧包装、帧传输、帧同步
- 帧的差错恢复
- 流量控制

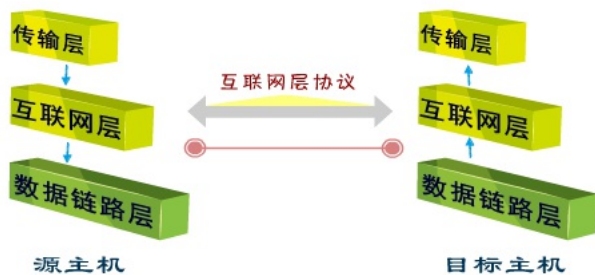
3.1.3 MAC地址 <#>

- 在通信过程中是用内置在网卡内的地址来标识计算机身份的
- 每个网卡都有一个全球唯一的地址来标识自己，不会重复
- MAC地址48位的二进制组成，通常分为6段，用16进制表示



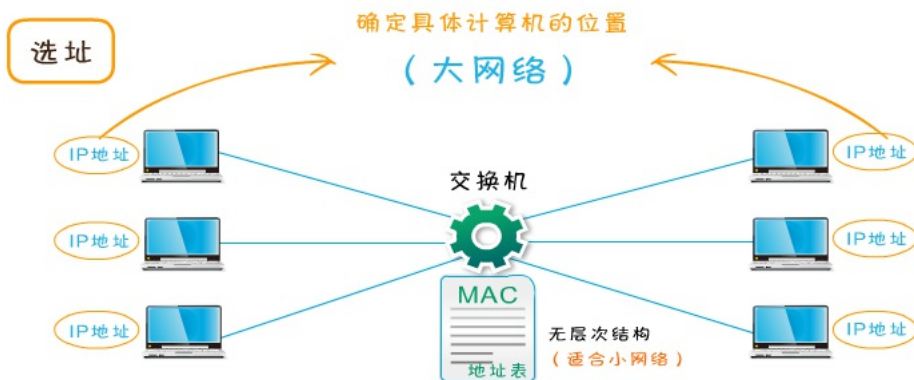
3.2 互联网层(网络层) <#>

- 位于传输层和网络接口层之间，用于把数据从源主机经过若干个中间节点传送到目标主机，并向传输层提供最基础的数据传输服务，它要提供路由和选址的工作



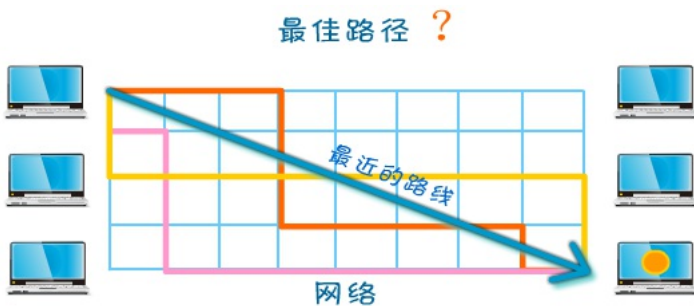
3.2.1 选址

交换机是靠MAC来寻址的，而因为MAC地址是无层次的，所以要靠IP地址来确认计算机的位置，这就是选址



3.2.2 路由

在能够选择的多条道路之间选择一条最短的路径就是路由的工作



3.2.3 IP

在网络中，每台计算机都有一个唯一的地址，方便别人找到它，这个地址称为IP地址。

3.2.3.1 IP头部

版本 (4)	首部长度 (4)	优先级与服务类型 (8)	总长度 (16)	
标识符 (16)			标志 (3)	段偏移量 (13)
TTL (8)		协议号 (8)	首部校验和 (16)	
源地址 (32)				
目标地址 (32)				
可选项				
数据				

20
字
节

20
字
节

- 版本
- 首部长度
- 优先级与服务类型
- 总长度 该字段用以指示整个IP数据包的长度，最长为65535字节，包括头和数据
- 标识符 唯一标识主机发送的每一份数据报
- 标志 分为3个字段，依次为保留位、不分片位和更多片位
 - 保留位：一般被置为0
 - 不分片位：表示该数据报是否被分片，如果被置为1，则不能对数据报进行分片，如果要对其进行分片处理，就应将其置为0
 - 更多片位：除了最后一个分片，其他每个组成数据报的片都要将该位置设置为1。
- 段偏移量 该分片相对于原始数据报开始处位置的偏移量
- TTL(Time to Live生存时间) 该字段用于表示IP数据包的生命周期，可以防止一个数据包在网络中无限循环地发下去。TTL的意思是一个数据包在被丢弃之前在网络中的最大周转时间。该数据包经过的每一个路由器都会检查该字段中的值，当TTL的值为0时此数据包会被丢弃。TTL对应于一个数据包通过路由器的数目，一个数据包每经过一个路由器，TTL将减去1。
- 协议号 用以指示IP数据包中封装的是哪个协议
- 首部校验和 校验和是16位的错误检测字段。目的主机和网络中的每个网关都要重新计算报头的校验和，一样表示没有改动过。
- 源IP地址 该字段用于表示数据包的源地址，指的是发送该数据包的设备的网络地址
- 目标IP地址 该字段用于表示数据包的目标的地址，指的是接收节点的网络地址

3.2.3.2 IP地址格式

- IP地址是一个网络编码，用来确定网络中的一个节点。
- IP地址是由32位二进制(32bit)组成

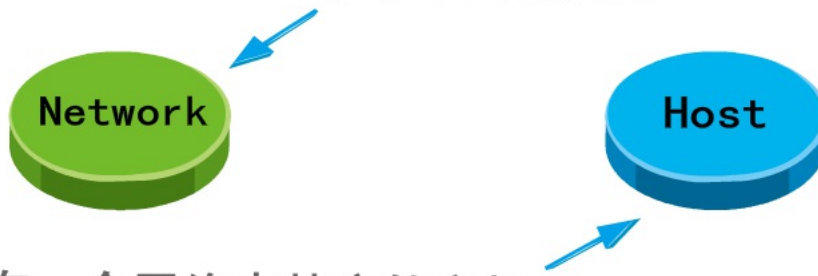
	32bits															
	Network								Host							
点分十进制	255				255				255				255			
最大值	1~8				9~16				17~24				25~32			
二进制	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	128	64	32	16	8	4	2	1	128	64	32	16	8	4	2	1
二进制例子	10101100				00010000				01111010				11001100			
十进制例子	172				16				122				204			

$128+64+8+4=204$

3.2.3.3 IP地址组成

- 网络部分(NETWORK)
- 主机部分(HOST)

标示不同的网络



标示在一个网络中特定的主机

3.2.3.4 IP地址表示

	32bits			
点分十进制	Network		Host	
最大值	255	255	255	255
二进制	1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1
二进制例子	10101100	00010000	01111010	11001100
十进制例子	172	16	122	204

$$128+64+8+4=204$$

3.2.3.5 IP地址的分类

- IP地址的网络部分是由Internet地址分配机构来统一分配的，这样可以保证IP的唯一性。
- ip地址中全为1的ip即255.255.255.255，它称为限制广播地址，如果将其作为数据包的目标地址可以理解为发送到所有网络的所有主机
- ip地址中全为0的ip即0.0.0.0，它表示启动时的ip地址，其含义就是尚未未分配时的ip地址
- 127是用来进行本机测试的，除了127.255.255.255外，其它的127开头的地址都代表本机

Bits:	1~8	9~16	17~24	25~32
A类:	0NNNNNNN	Host	Host	Host
	范围(1-126)	每个网络中可以有2的24次方-2台的主机.		
Bits:	1~8	9~16	17~24	25~32
B类:	10NNNNNN	Network	Host	Host
	范围(128-191)	最大的主机数量为2的16次方减2		
Bits:	1~8	9~16	17~24	25~32
C类:	110NNNNN	Network	Network	Host
	范围(192-223)	最大的主机数量为2的8次方减2		

3.2.3.6 公有地址和私有地址

私有IP

- A类私有IP: 10.0.0.0 ~ 10.255.255.255
- B类私有IP: 172.16.0.0 ~ 172.31.255.255
- C类私有IP: 192.168.0.0 ~ 192.168.255.255

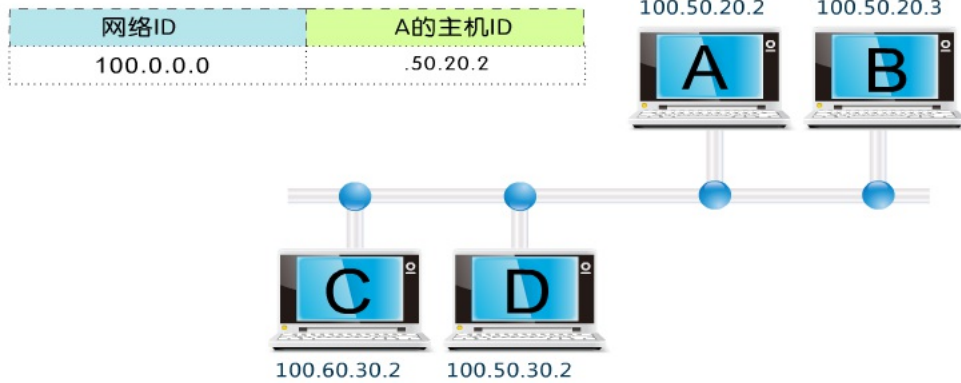
其他范围的IP均为公有IP地址

3.2.3.7 子网掩码

子网掩码(subnet mask)又叫子网遮罩，它是一种用来指明一个IP地址的哪些位标识的是主机所在的子网，以及哪些位标识的是主机位的掩码。子网掩码不能单独存在，它必须结合IP地址一起使用。子网掩码只有一个作用，就是将某个IP地址划分成网络地址和主机地址两部分。

- 子网掩码也是32个二进制位
- 对应IP的网络部分用1表示
- 对应IP地址的主机部分用0表示
- IP地址和子网掩码做逻辑与运算得到网络地址
 - 0和任何数相与都是0
 - 1和任何数相与都等于任何数本身
- A B C三类地址都有自己默认的子网掩码
 - A类 255.0.0.0
 - B类 255.255.0.0
 - C类 255.255.255.0

子网掩码: 255.0.0.0



这四台计算机只有在同一个网段，才能互相通信。

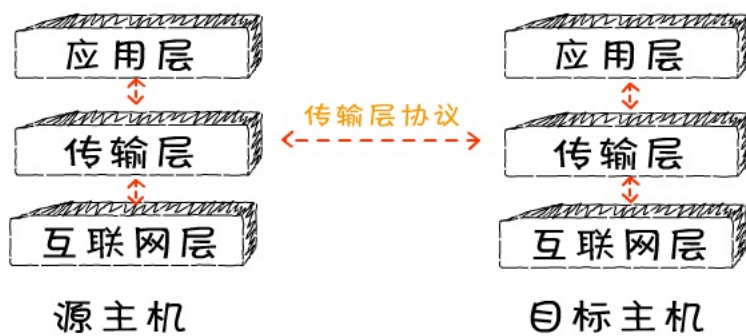
3.3 传输层 <#>

位于应用层和网络接口层之间

- 是面向连接的、可靠的的进程到进程通信的协议
- TCP提供全双工服务，即数据可在同一时间双向传播
- TCP将若干个字节构成一个分组，此分组称为报文段(Segment)

传输层概述

位于传输层和网络接口层之间



3.3.1 传输层的功能 <#>

提供了一种端到端的连接

□

3.3.2 对可靠性要求高的上层协议，实现可靠性的保证 <#>

- 如果数据丢失、损坏的情况下如何保证可靠性
- 网络层只管传递数据，成功与否并不关心

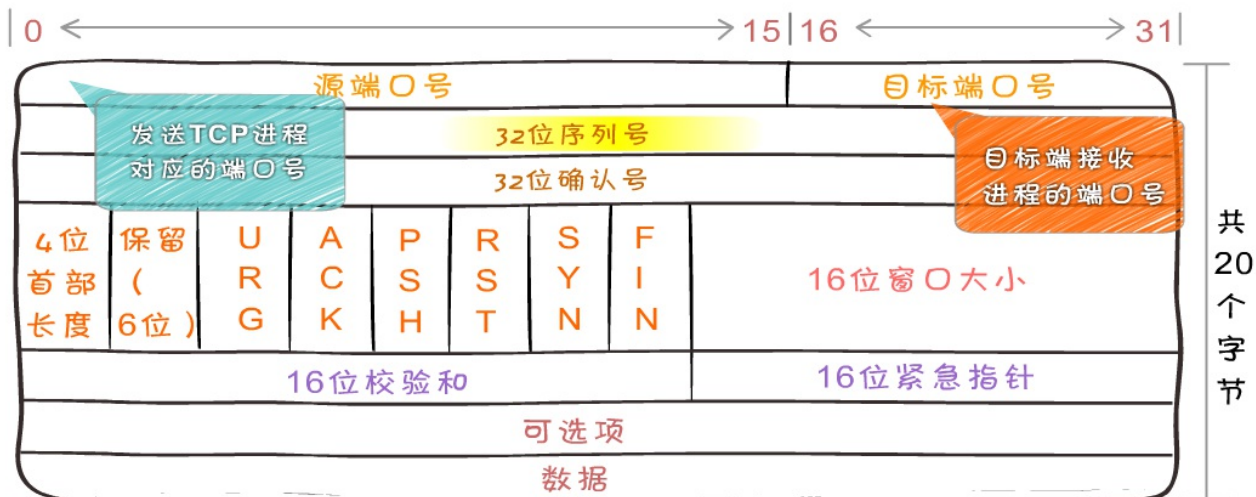
3.3.3 协议分类 <#>

- TCP(Transmission Control Protocol)
 - 传输控制协议
 - 可靠的、面向连接的协议
 - 传输效率低
- UDP(User Datagram Protocol)
 - 用户数据报协议
 - 不可靠的、无连接的服务
 - 传输效率高

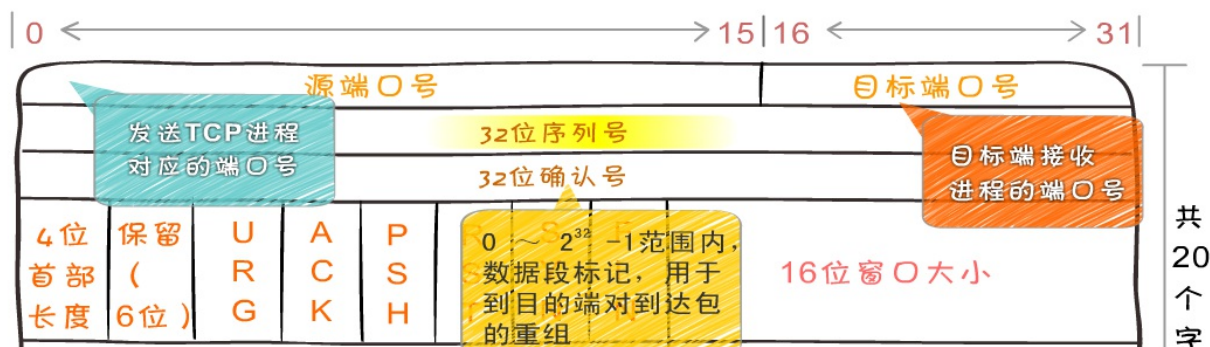
3.3.4 TCP功能 <#>

- 将数据进行分段打包传输
- 对每个数据包编号控制顺序
- 运输中丢失、重发和丢弃处理
- 流量控制避免拥塞

3.3.4.1 TCP数据包封装 <#>



- 源端口号和目标端口号，计算机通过端口号识别访问哪个服务,比如http服务或ftp服务，发送方端口号是进行随机端口，目标端口号决定了接收方哪个程序来接收



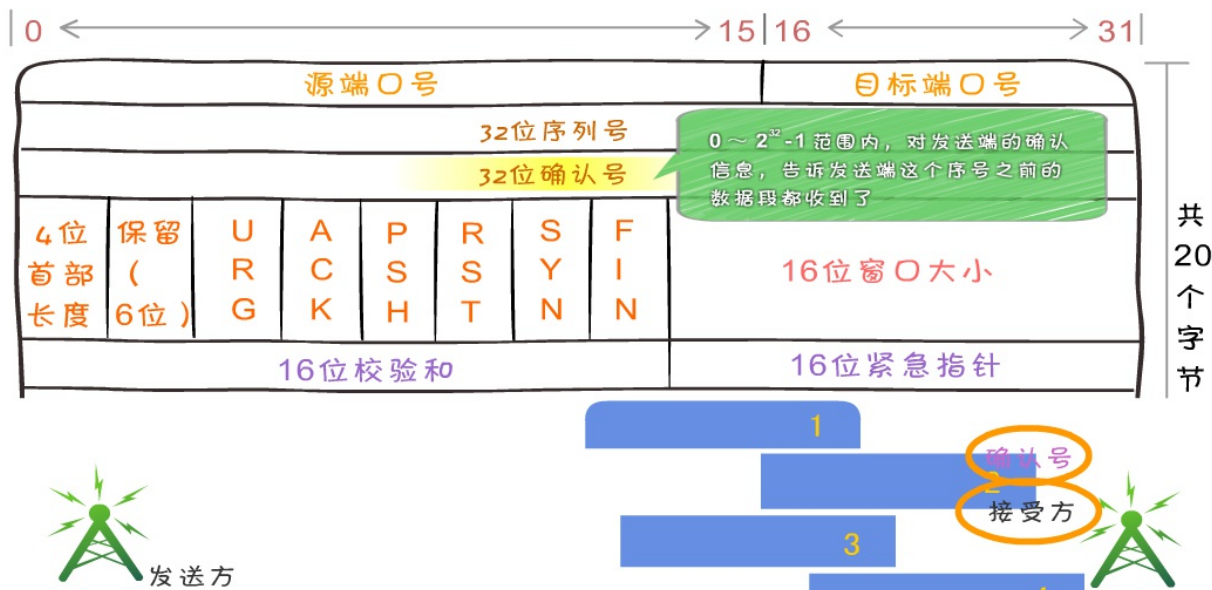
发送方



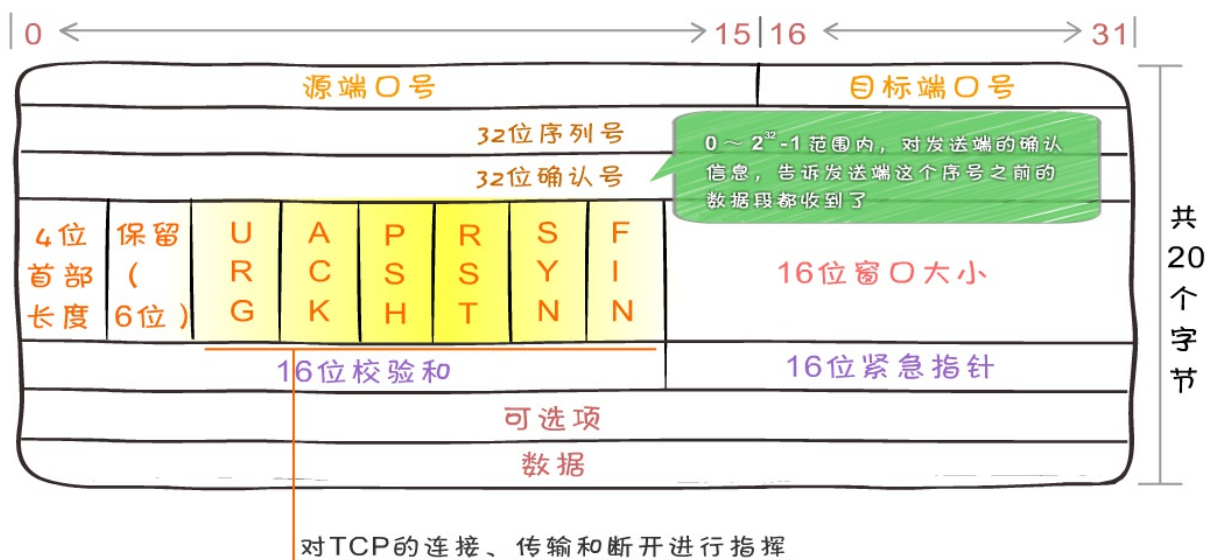
接收方



- 32位序列号 TCP用序列号对数据包进行标记，以便在到达目的地后重新重装，假设当前的序列号为 s ，发送数据长度为 l ，则下次发送数据时的序列号为 $s+l$ 。在建立连接时通常由计算机生成一个随机数作为序列号的初始值



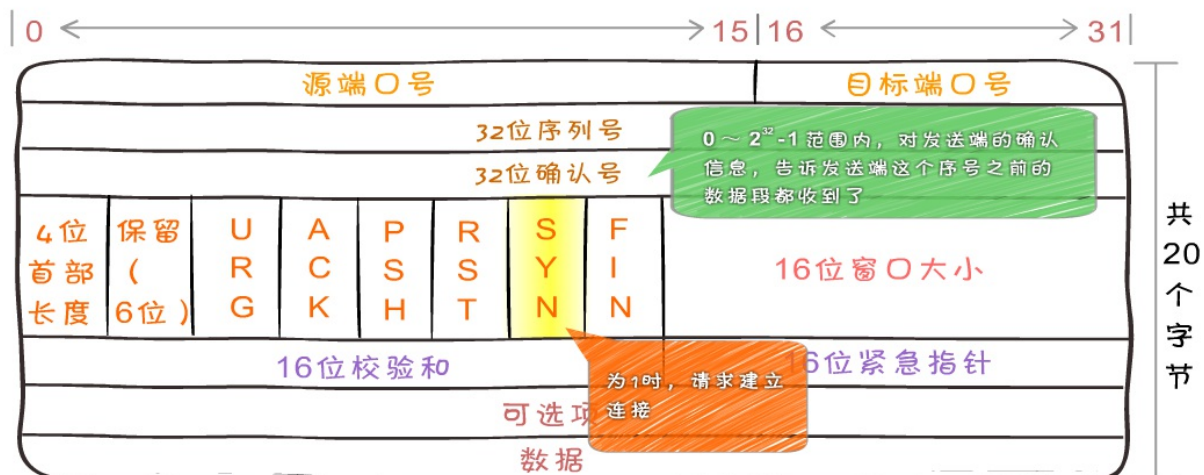
- 确认应答号 它等于下一次应该接收到的数据的序列号。假设发送端的序列号为 s , 发送数据的长度为 l , 那么接收端返回的确认应答号也是 $s+l$ 。发送端接收到这个确认应答后, 可以认为这个位置以前所有的数据都已被正常接收。
- 首部长度: TCP 首部的长度, 单位为 4 字节。如果没有可选字段, 那么这里的值就是 5。表示 TCP 首部的长度为 20 字节。



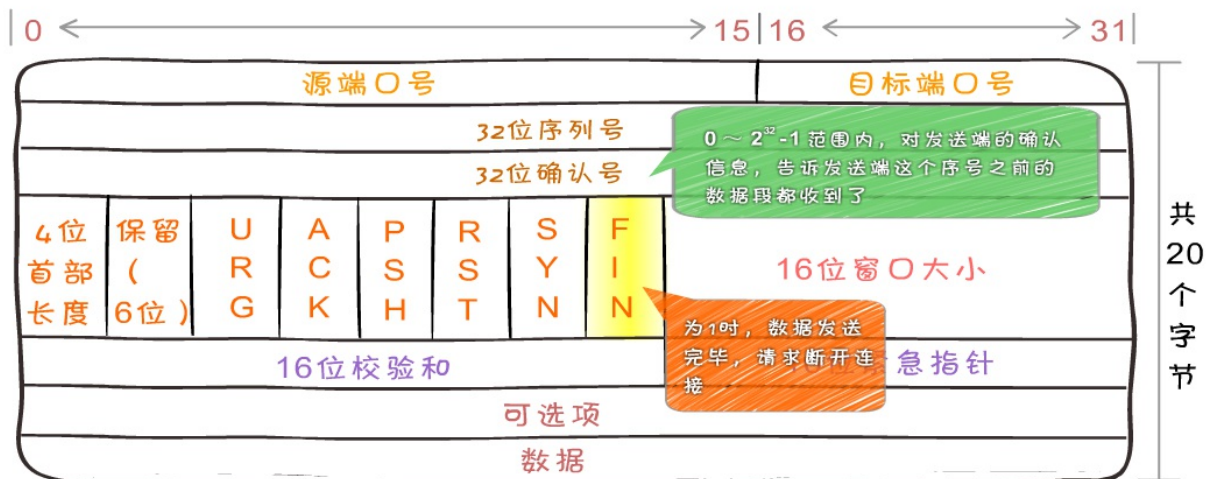
- 控制位 TCP的连接、传输和断开都受这六个控制位的指挥
 - PSH(push急迫位) 缓存区将满, 立刻传输速度
 - RST(reset重置位) 连接断了重新连接
 - URG(urgent紧急位) 紧急信号
- 紧急指针: 尽在 URG(urgent紧急) 控制位为 1 时有效。表示紧急数据的末尾在 TCP 数据部分中的位置。通常在暂时中断通信时使用 (比如输入 Ctrl + C)。



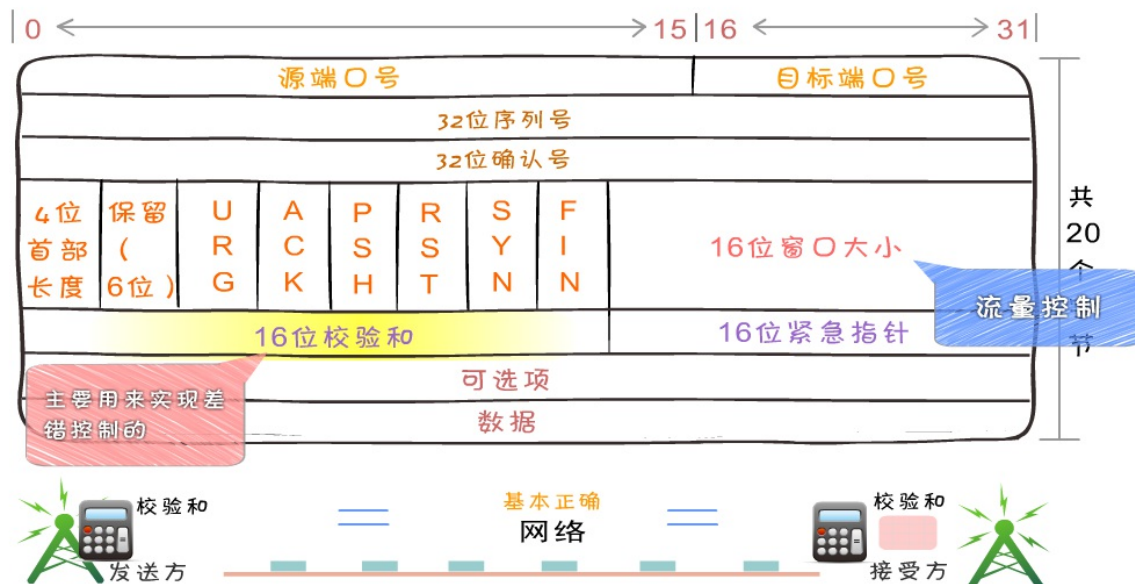
- ACK(acknowledgement 确认)为1表示确认号



- SYN(synchronous 建立联机) 同步序号位 TCP建立连接时要将这个值设为1



- FIN发送端成位，提出断开连接的一方把FIN置为1表示要断开连接
- 窗口值 说明本地可接收数据段的数目，这个值的大小是可变的。当网络通畅时将这个窗口值变大加快传输速度，当网络不稳定时减少这个值可以保证网络数据的可靠传输。它是来在TCP传输中进行流量控制的
- 窗口大小：用于表示从应答号开始能够接受多少个8位字节。如果窗口大小为0，可以发送窗口探测。



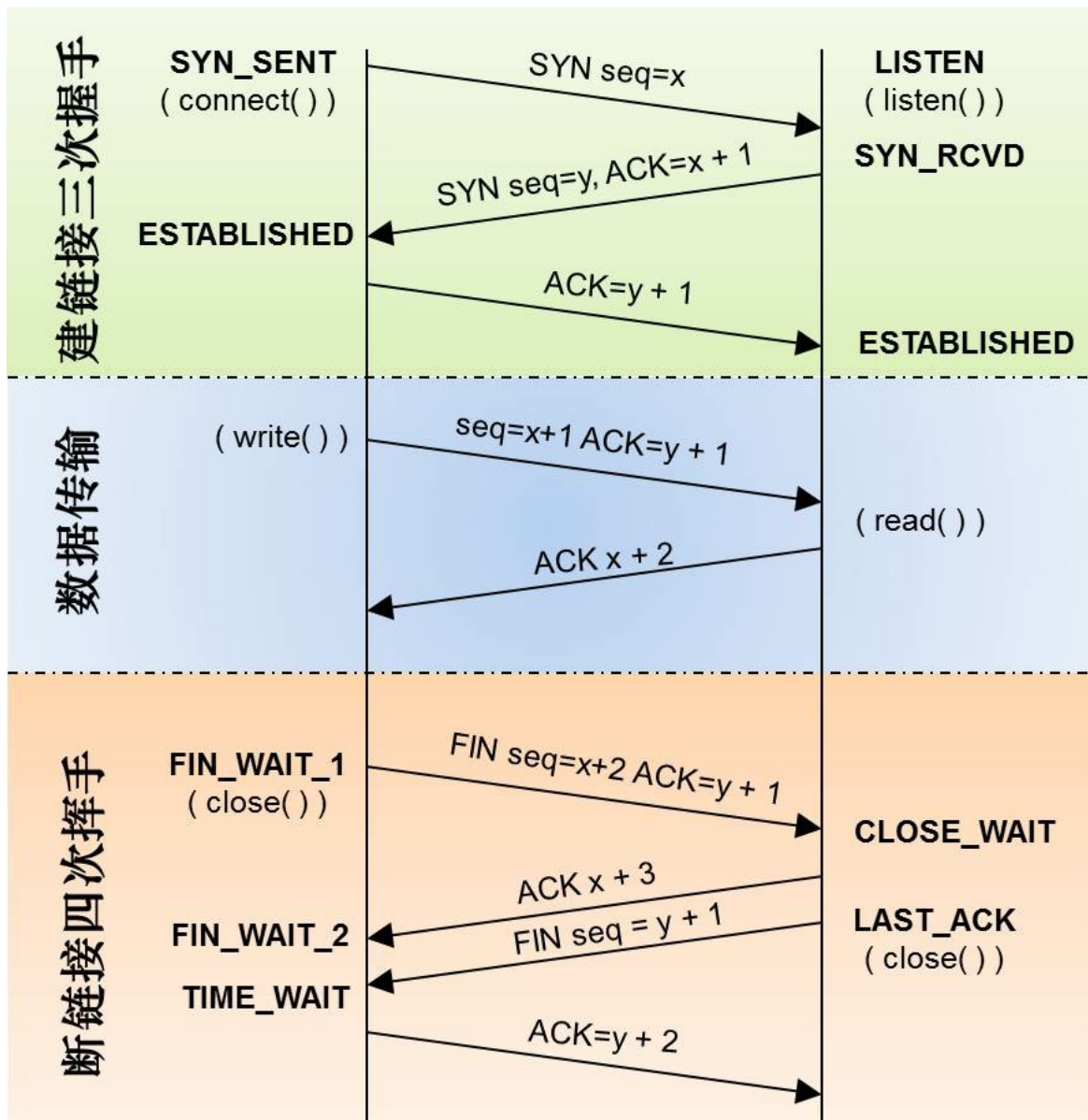
- 用来做差错控制，TCP校验和的计算包括TCP首部、数据和其它填充字节。在发送TCP数据段时，由发送端计算校验和，当到达目的地时又进行一次校验和计算。如果两次校验和一致说明数据是正确的，否则将认为数据被破坏，接收端将丢弃该数据

3.2.4.2 握手和断开

- TCP是面向连接的协议，它在源点和终点之间建立虚拟连接，而不是物理连接
- 在数据通信之前，发送端与接收端要先建立连接，等数据发送结束后，双方再断开连接
- TCP连接的每一方都是由一个IP地址和一个端口组成

Client

Server



3.2.4.2.1 三次握手

- 第一次握手主机A通过一个标识为SYN标识位的数据段发送给主机B请求连接，通过该数据段告诉主机B希望建立连接，需要B应答，并告诉主机B传输的起始序列号
- 第二次握手是主机B用一个确认应答ACK和同步序列号SYN标志位的数据段来响应主机A，一是发送ACK告诉主机A收到了数据段，二是通知主机A从哪个序列号做标记。
- 第三次握手是主机A确认收到了主机B的数据段并可以开始传输实际数据。

3.2.4.2.2 四次断开

- 主机A发送FIN控制位发出断开连接的请求
- 主机B进行响应，确认收到断开连接请求
- 主机B提出反方向的关闭要求
- 主机A确认收到的主机B的关闭连接请求

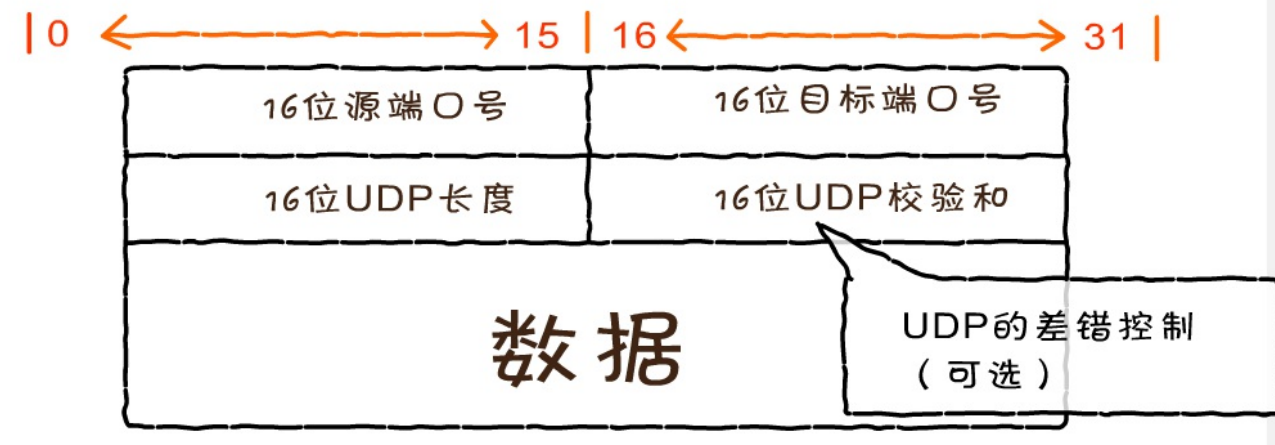
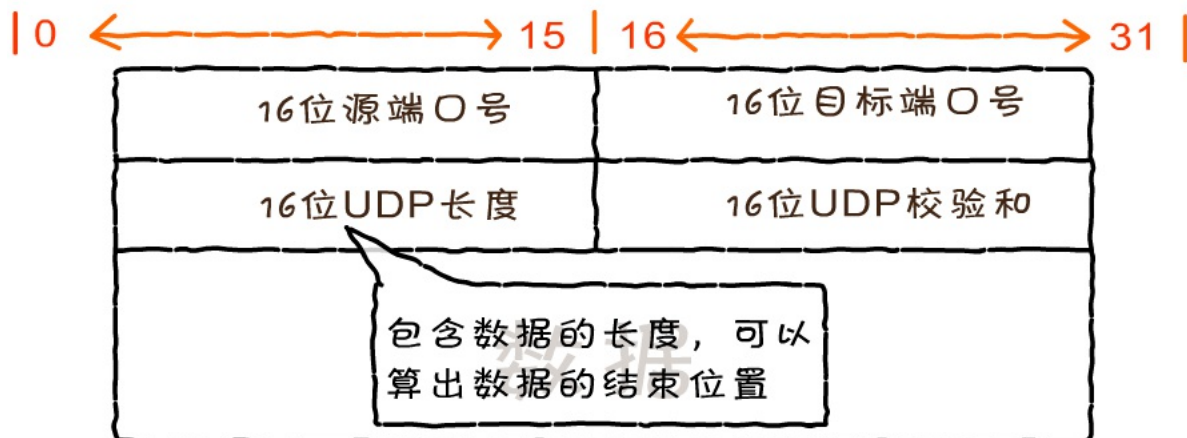
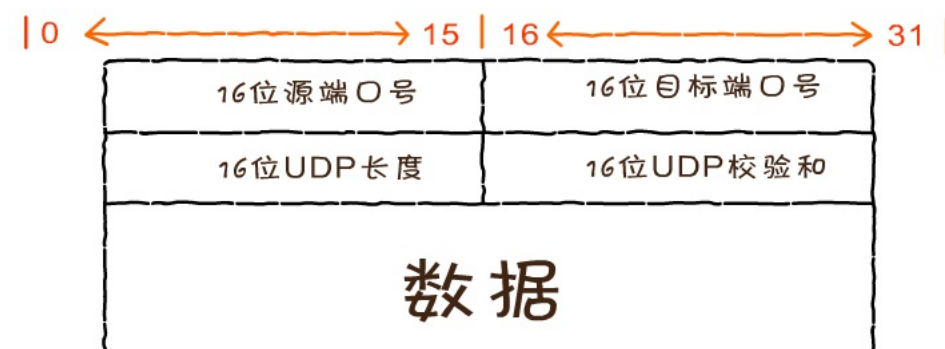
3.2.4.2.3 报文



3.4 UDP

- UDP是一个无连接、不保证可靠性的传输层协议，也就是说发送端不关心发送的数据是否到达目标主机、数据是否出错等，收到数据的主机也不会告诉发送方是否收到了数据，它的可靠性由上层协议来保障
- 首部结构简单，在数据传输时能实现最小的开销，如果进程想发送很短的报文而对可靠性要求不高可以使用

3.4.1 UDP的封装格式



3.4.2 UDP的应用 #

- QQ
- 视频软件
- TFTP 简单文件传输协议(短信)

3.4.3 DNS服务器 #

3.4.3.1 域名 #

- 域名空间结构
- 根域
- 顶级域
 - 组织域
 - 国家/地区域名
- 二级域名

名称类型	说明				示例	
根域	一般认为全球共有13台根逻辑域名服务器				单个句点 (.) 或句点用于末尾的名称	
顶级域	用来指示某个国家/地区或组织使用的名称的类型名称				.com	
第二层域	个人或组织在网上使用的注册名称				zol.com	
子域	已注册的二级域名派生的域名				www.zol.com	

DNS域名	cn/ru	com	net	edu	Mil	gov
组织类型	中国/俄罗斯	商业公司	网络公司	教育机构	军事政府机构	非军事政府机构

3.4.3.2 DNS服务器 #

DNS是Domain Name Service的缩写，DNS服务器进行域名和与之对应的IP地址转换的服务器

- IP地址不易记忆
- 早期使用Hosts文件解析域名
 - 主要名称重复
 - 主机维护困难
- DNS(Domain Name System 域名系统)
 - 分布式
 - 层次性

3.4.3.3 查找过程 #

- 客户端向本地域名服务器发出请求，我要访问www.163.com，请告诉我它的IP地址 (<http://www.163.com>，请告诉我它的IP地址)
- 本地DNS服务器向DNS根服务器发出请求，根DNS服务器会告诉本地服务器(.com)的服务器地址
- 本地DNS服务器会向(.com)发请求，会得到(163.com)的服务器地址
- 本地DNS服务器会向(163.com)发请求,会得到(www.163.com)的IP地址1.1.1.1 (<http://www.163.com>)的IP地址1.1.1.1)
- 本地DNS服务器向客户端回复域名(www.163.com)对应的IP地址是1.1.1.1 (<http://www.163.com>)对应的IP地址是1.1.1.1)

3.5 应用层 #

3.5.1 协议 #

3.5.2 应用层常见协议 #

- HTTP 超文件传输协议
- FTP 文件传输协议
- SMTP(发送邮件)和POP3(接收邮件)

3.6 案例 #

数据->传输层(包)->网络层(段Segment)->数据链路层(帧)

3.6.1 发送方是从高层到低层封装数据 #

- 在应用层要把各式各样的数据如字母、数字、汉字、图片等转换成二进制
- 在TCP传输层中，上层的数据被分割成小的数据段，并为每个分段后的数据封装TCP报文头部
- 在TCP头部有一个关键字段信息端口号，它用于标识上层的协议或应用程序，确保上层数据的正常通信
- 计算机可以多进程并发运行，例如在发邮件的同时也可以通过浏览器浏览网页，这两种应用通过端口号进行区分
- 在网络层，上层数据被封装上层的报文头部(IP头部)，上层的数据是包括TCP头部的。IP地址包括的最关键字段信息就是IP地址，用于标识网络的逻辑地址。
- 数据链路层，上层数据成一个MAC头部，内部有最关键的是MAC地址。MAC地址就是固化在硬件设备内部的全球唯一的物理地址。
- 在物理层，无论之前哪一层封装的报文头和还是上层数据都是由二进制组成的，物理将这些二进制数字比特流转换成电信号在网络中传输

3.6.2 接收方是从低层到高层解封装 #

- 数据封装完毕传输到接收方后，将数据要进行解封装
- 在物理层，先把电信号转成二进制数据，并将数据传送到数据链路层
- 在数据链路层，把MAC头部拆掉，并将剩余的数据传送到上一层
- 在网络层，数据的IP头部被拆掉，并将剩余的数据传送到上一层
- 在传输层，把TCP头部拆掉，将真实的数据传送到应用层

3.6.3 真实网络环境 #

- 发送方和接收方中间可能会有多个硬件中转设备
- 中间可能会增加交换机和路由器
- 数据在传输过程中不断地进行封装和解封装的过程，每层设备只能处理哪一层的数据
 - 交换机属于数据链路层
 - 路由器属于网络层

4. 附录

4.1 不同层中的称谓：

- 数据帧（Frame）：是一种信息单位，它的起始点和目的地都是 **数据链路层**。
- 数据包（Packet）：也是一种信息单位，它的起始和目的地是 **网络层**。
- 段（Segment）：通常是指起始点和目的地都是 **传输层** 的信息单元。
- 消息（message）：是指起始点和目的地都在网络层以上（经常在 **应用层**）的信息单元。

4.2 IP头服务类型

- IP首部中的服务类型（TOS）
- TOS包括共8位，包括3 bit的优先权字段（取值可以从000-111所有值），4 bit的TOS子字段和1 bit未用位但必须置0。
- 3bit的8个优先级的定义如下：
 - 111—Network Control(网络控制)一般保留给网络控制数据使用，如路由。
 - 110—Internetwork Control(网间控制)
 - 101—Critical(关键)语音数据使用。
 - 100—Flash Override(疾速)视频会议和视频流使用。
 - 011—Flash(闪速)语音控制数据使用。
 - 010—Immediate(快速)数据业务使用
 - 001—Priority(优先)数据业务使用
 - 000—Routine(普通)默认标记值。
- 4 bit的TOS分别代表：最小时延、最大吞吐量、最高可靠性和最小费用。4 bit中只能置其中1 bit。如果所有4 bit均为0，那么就意味着是一般服务。
- Telnet、Rlogin这两个交互应用要求最小的传输时延，FTP文件传输要求最大吞吐量，最高可靠性是指网络管理（SNMP）和路由选择协议。用户网络新闻要求最小费用

4.3 握手和断开

5. 参考

- [TCP/IP 之 大明王朝邮差 \(https://mp.weixin.qq.com/s?__biz=MzAxOTc0NzExNg==&mid=2665513094&idx=1&sn=a2accfd41107ac08d74ec3317995955e#rd\)](https://mp.weixin.qq.com/s?__biz=MzAxOTc0NzExNg==&mid=2665513094&idx=1&sn=a2accfd41107ac08d74ec3317995955e#rd)