计算机网络

一、概述

- 计算机网络的功能：数据通信、资源共享、分布式处理、提高可靠性、负载均衡。
- 计算机网络的组成
  - 硬件、软件、协议
  - 通信子网、资源子网
- 计算机网络的分类
  - 广域网WAN、城域网MAN（1~10km）、局域网LAN（1km以内）、个人区域网PAN
  - 广播式网络、点对点网络
  - 面向连接服务、无连接服务
  - 电路交换（建立物理电路）、报文交换（存储转发）、分组交换（对报文分割，长度更小）。其中分组交换可以分为无连接的数据报方式和面向连接的虚电路方式。
- 一些概念：
  - IP电话/VoIP：使用IP协议实现的电话。
  - VPN/虚拟专用网络：建立虚拟的内部网络。
  - CS与P2P
  - NFC/近场通信
  - AP/接入点
- 协议：相同层之间通信规则的集合；

  接口：相邻两层之间交换信息的连接点；

  服务：下层为紧邻上层提供的功能调用。
- 时延：发送时延、传播时延、处理时延、排队时延。总时延 = 发送时延 + 传播时延 + 处理时延 + 排队时延。

  往返时延RTT，时延带宽积。

模型

- OSI参考模型：物理层、数据链路层、网络层、传输层、会话层、表示层、应用层。
- TCP/IP模型：链路层、互联网层、传输层、应用层。
- 5层模型：物理层、数据链路层、网络层、传输层、应用层。
- 只有传输层及以上各层的通信才能称为端到端；之下的层是点到点。

二、物理层

1.概述

- 物理层描述：机械特性、电气特性、功能特性、规程特性。
- 信道：信号从传输媒介。可分为单双工通信、半双工通信、全双工通信。
- 带宽：信号能通过的频率的范围，频带宽度。

- 码元：指一个固定时长的波形。波特率：码元传输速率。
- 传输介质：双绞线、同轴电缆、电力线、光纤（多模光纤、单模光纤）；无线电、微波、红外线、光线。
- 物理层的设备：中继器和集线器，作用是信号再生。

2.最大传输速率

- 奈奎斯特定理：极限数据传输率 = $2W\log_2 V$，其中W是理想低通信道的带宽，V是每个码元有多少种形式。
- 香农定理：极限数据传输率 = $W\log_2(1+S/N)$，其中S/N的信噪比。

3.基带传输

- 编码方式：不归零编码、不归零逆转、曼彻斯特编码、差分曼彻斯特编码、双级编码等。
- 信号频率越高衰减越大，因此需要尽量降低频率。
- 信号传输中需要考虑时钟恢复问题，即两端的时钟需要同步。解决办法：两端采用精准的时钟、采用曼彻斯特编码、采用不归零逆转编码、对编码进行4B/5B映射、对编码伪随机化。
- 信号传输中需要考虑功率平衡问题，即两端尽量只传输信号不传输功率，消除电流中的直流分量。解决办法：采用双极编码、采用8B/10B映射。

4.通带传输

- 频移键控/调频
- 幅移键控/调幅
- 相移键控/调相
- 正交条幅（调幅+调相）QAM

5.多路复用

- 频分复用FDM、正交频分复用QFDM：将带宽划分成多个频带给不用的信道使用。
- 时分复用TDM：划分等长的时间槽，以时间槽为单位分配给多个信道使用。
- 码分复用CDM：如码分多址CDMA为每个信道分配一个正交的码片序列（由多个+1、-1构成）作为地址，通过与广播中的码片做内积可以还原出信息。

6.常见网络

- 电话系统：本地回路、中继线、交换局。交换局目前已经被同步光网络SONET替代，对电话语音信号进行了数字化。

  通过电话线路可以拨号上网（56kbps），因电话中使用的信号频率较低，在两端增加了高频滤波器，限制了频率，故拨号上网速率很低；后新建设备去除了滤波器，产生了数字用户线DSL，如非对称数字用户线ADSL（<10Mbps）；新的网络业务是光纤到户。
- 移动电话系统：又称为蜂窝电话。目前已经发展到第4代，使用模拟语音的AMPS（1G）、支持数字语音的GSM（2G）、支持数字语音和数据的CDMA（3G）、4G LTE。
- 有线电视网络

三、数据链路层

1.概述

数据链路层的功能：为网络层提供服务、成帧、差错控制、流量控制。

数据链路层为网络层提供：

- 无确认的无连接服务，如以太网
- 有确认的无连接服务，如WiFi
- 有确认的面向连接服务

局域网相关协议：以太网802.3，无线局域网802.11，蓝牙，无线射频识别RFID；

广域网相关协议：PPP协议（SONET光纤链路），HDLC协议，ADSL协议。

工作在数据链路层的设备是网桥和交换机，可以连接使用不同物理层的多个网络，可以隔离冲突域，但是仍在一个广播域。交换机可以分为直通式交换机和存储转发式交换机。

2.成帧

帧定界的方式：

- 字节记数法：在帧头部添加一个字段来标识该帧有多少字节。如果丢失一个比特，其后的传输都会定界错误，基本不使用。
- 字节填充的标志字节法：用一些特殊的字节（flag）包裹有效载荷，如PPP协议。当需要在有效载荷中添加一个flag时，在其前面添加一个转义字节；当需要在有效载荷中添加一个转义字节时，在其前面添加一个转义字节。
- 比特填充的标志比特法：用比特模式01111110包裹有效载荷，在有效载荷中遇到连续5个1就添加一个0。
- 物理层编码违禁法：用物理层不应该出现的比特编码模式来区分帧边界。

帧中的帧头记录了帧的协议、序号、校验等信息。

3.差错控制

- 对于出现位错误，通常采用校验码来发现错误，通过自动重传请求ARQ重传帧。

  纠正单个错误所需的最少校验位个数r满足：$(m+r+1) \leq 2^r$，其中m位需要校验的位数。

  检错码的种类：奇偶校验码、校验和、循环冗余校验CRC；

  纠错码的种类：海明码、二进制卷积码、里德所罗门码、低密度奇偶校验码。

- 对于出现帧错误，如帧的丢失、重复或失序等，引入定时器和编号机制，保证帧的有序、正确交付。

4.流量控制与滑动窗口协议

停止-等待协议

发送方和接收方都只有一个窗口，发送方每发送一帧就停止等待，接收方每接收一帧就反馈一个应答信号。此时只需要1位帧序号，出现错误使用ARQ（快速重传）。

滑动窗口协议

发送方有一定大小的发送窗口，接收方有一定大小的接收窗口。发送方发送一帧会占用一个发送窗口，只有在收该帧的确认后才让出该窗口；当没有空闲的发送窗口，发送方停止发送。接收方接收到的帧放入接收窗口，当该帧传入上层后，才会对该帧确认并让出该窗口；当没有空闲的接收窗口，接收方停止接收帧，即丢弃窗口外的帧。

因网络通信都是双向的，帧的确认可以捎带在接收方给发送方发送的帧中，称为捎带确认。确认的方式有两种，形成了两种不同的协议。

- **回退N协议**：接收方只会按帧序号的顺序确认帧，如没有收到3号帧，即使收到4号、5号帧也不会对它们进行确认。此时发送方的计时器超时，会按顺序重传3号、4号、5号帧，即发送方回退的3号。这种方式下可以累计确认，即确认5号帧也就表明5号之前的也确认。

  接收窗口大小≥ 1即可，发送窗口大小$1 \leq W \leq 2^n - 1$，其中n位帧号的长度。

- **选择重传协议**：接收方会确认接收到的所有正确的帧，如没有收到3号帧，但是收到了4号、5号帧，接收方仍会确认4号、5号。此时发送方的3号帧的计时器超时，但是收到了4号、5号帧的确认，只会重传3号帧。

  接收窗口和发送窗口的最大值均为$2^{(n-1)}$，一般情况下大小相等。

5.介质访问控制子层MAC

介质访问控制子层用于广播信道中的信道分配。分为信道划分介质访问控制和随机访问介质访问控制。

数据链路层在IEEE标准中分为两个子层：链路控制子层LLC和介质访问子层MAC。

信道划分介质访问控制

将信道划分成多个，不需要考虑冲突。

- 频分多路复用FDM
- 时分多路复用TDM
- 波分多路复用WDM，实质是光的频分多路复用
- 码分多路复用CDM，码分多址CDMA

随机访问介质访问控制

多个端共享同一个信道，可能发生冲突。

- ALOHA协议：发送数据前不检测冲突直接发送，一段时间没有收到确认即认为发生冲突，等待随机的时间后重新发送数据。
- 分槽ALOHA协议：划分时间槽，发送数据不检测冲突，在时间槽开始时开始发送，一个时间槽内发送完一帧，同样的一段时间没有收到确认即认为发生冲突，等待随机数量的时间槽后重新发送数据。利用率比纯ALOHA协议要好，$1/e$。
- 载波侦听多路访问 CSMA协议
  - 1-坚持型：发送数据前侦听信道，若信道空闲则发送，若**信道忙则持续侦听**至信道空闲后发送。若发送后发生冲突，在等待随机时间后重复上述过程。
  - 非坚持型：发送数据前侦听信道，若信道空闲则发送，若**信道忙则放弃侦听**，等待随机时间后再次侦听。若发送后发生冲突，在等待随机时间后重复上述过程。
  - p坚持型：适用于分时间槽的信道。发送数据前侦听信道，若信道空闲，则以概率p发送数据，以概率（1-p）推迟到下一个时间槽，在下一个时间槽重复此判断，直至帧被发送出去。若信道忙，则等到下个时间槽重复此过程。

- 带冲突检测的CSMA CSMA/CD协议：与CSMA类似，发送前先侦听信道，发送后继续侦听，一旦检测到冲突就立刻停止发送，等待随机时间后重发。假设广播域内两个相距最远的站之间信号传播时延为τ，则发送方开始发送经过2τ后没有检测到冲突，此时广播域内的所有站点都已经知道信道被占用，就不会再发生冲突。开始发送至2τ称为争用期，一帧的发送时延要大于2τ。

  一旦检测到冲突，CSMA/CD采用二进制指数退避算法来决定重发的时延。基本退避时间取2τ，重传次数记为k，当重传次数大于10时，k始终记为10。当检测到冲突，从$0, 1, 2, 3, \cdots, 2^k - 1$中随机取一个数$r$，所需重传时间为$2r\tau$。当重传16次仍不成功，向上层报告错误。

  主要用于总线式以太网。

- 带冲突避免的CSMA CSMA/CA协议：无线通信中并非所有站点都能听见对方，存在隐蔽站，此时冲突检测就失效了。当发送方需要发送数据时，首先发送一个RTS短帧询问接收方，包含需要发送的数据的长度；接收方收到之后会发送一个CTS帧作为应答。其它的站不管是收到RTS帧还是CTS帧，都延缓发送请求。但是仍然可能发生冲突。

  检测对冲突后和带冲突检测的CSMA一样，使用二进制指数退避算法。

  主要用于无线局域网802.11abgn。

轮询访问介质访问控制

令牌传递协议中只有持有令牌的站才能使用信道，使用完之后将令牌传给下一个站。

6.以太网 IEEE 802.3

经典以太网和交换式以太网。现在只使用交换式以太网。

以太网采用无连接的工作方式，不对数据帧编号，也不要求确认，以广播的形式进行。

以太网的帧格式：

- 8个字节的前导码，用于界定帧起始位置，模式是010101...，用于同步时钟。最后两位均为1，表明开始一个帧。
- 6个字节的目标地址和6个字节的源地址，该地址有硬件确定，出厂时分配。
- 2个字节的类型/长度：值小于等于0x600的，为长度字段，是一个IEEE 802.3标准的帧；大于0x600的，为类型字段，是一个DIX以太网帧。
- 0~1500个字节的数据。
- 0~46个字节的填充，要求**帧最小为64字节。**
- 4个字节的校验和，采用32位的CRC。

速率：

- 交换式以太网：10Mbps
- 快速以太网：100Mbps，使用100Base-TX/FX（五类线/光纤）可达到
- 千兆以太网：1000Mbps，使用1000Base-
- 万兆以太网：10000Mbps，使用10GBase-

7.无线局域网 IEEE 802.11

物理层采用正交频分复用，数据链路层采用CSMA/CA协议。

分为有固定基础设施（AP）的无线局域网和无固定基础设施的自组织网络。

四、网络层

1.概述

网络层的功能：异构网络互连、路由选择、分组转发、拥塞控制（拥塞控制是指整个网络的拥挤程度，而不单指点对点之间）。提供无连接的数据报和面向连接的虚电路服务。

工作在网络层的设备是路由器，隔离广播域，主要功能是分组转发、路由选择。

2.路由算法与路由协议

路由算法可以分为静态路由算法和动态路由算法，其中静态路由算法有

- 泛洪算法：获取网络的全貌。
- 最短路径算法：在已知网络全貌的基础上，使用Dijkstra算法获取最短路径。

动态路由算法有

- 距离矢量路由算法
- 链路状态路由算法

距离矢量路由算法与RIP协议

每个路由器维护一张路由表，表中每个路由器对于一个表项，包含到该路由器的首选出口和距离估计值。经过一定时间间隔每个路由器会将自己的整个路由表分享给**直接相连**的路由器。

- 接收到其他路由器的路由表后，如果某个表项比自己的对应表项距离小，用此表项更新自己的路由表；如果不存在某个路由的表项，则将此表项加入路由表。
- 在工作过程中如果某个路由不可达，只能等待与其他路由交换路由表以获取新的路由。

距离矢量路由算法总会得到收敛，但是很慢。

- 网络中的好消息传的很快，当路由器A加入网络，网络中距离A有n跳的路由器，至多经过n次路由表交换就会感知到A的存在。
- 网络中的坏消息传的很慢，当路由器A停机，其相邻的路由器B无法确定A是停机还是AB链路断了，但是B可以从另外的路由器的路由表中获取一个到A稍远的路由，这条路由表项的距离总是慢慢增加的，直到超出阈值。故坏消息传得慢。

RIP协议使用跳数描述距离，一条路径最多包含15个路由器，任意两个路由器每30秒广播一次自己的RIP路由更新信息。RIP是应用层协议，使用UDP。

链路状态路由算法与开放最短路径优先协议OSPF

链路状态路由算法的过程如下：

1. 每个路由器会通过每个出口发送一个HELLO数据包，其相邻的路由器会返回一个ECHO数据包，包含路由器的名字。通过两个数据包的间隔可以为每个相邻的路由器设定一个距离估计值。
2. 路由器将刚刚获取的所有相邻路由和距离构造一个链路信息包，将这个包分享给其他所有路由，并接收所有其他路由器的链路信息包。这里的链路数据包包含序号（用于表明该包的新旧）和存活时间（控制泛洪规模）。
3. 根据接收到的链路信息包计算出到每个路由器的最短距离。使用Dijkstra算法。

OSPF的路由器每10秒交换一次HELLO数据包，每30分钟更新一次链路状态。OSPF是网络层协议，直接使用IP数据报传输数据。

分层路由与边界网关协议BGP

一个自治系统内部使用的路由协议称为内部网关协议IGP，可以使用RIP或OSPF。自治系统之间使用的协议是外部网关协议EGP或边界网关协议BGP。

BGP协议使用路径矢量算法，只要找到一条较好的路由即可，不追求最好的路由。BGP首次与相邻的BGP路由器交换整个BGP路由表，之后只交换变化的部分。

BGP协议是应用层协议，使用TCP。

其他

广播路由、组播路由、选播路由，用于非一对一的通信。

3.拥塞控制

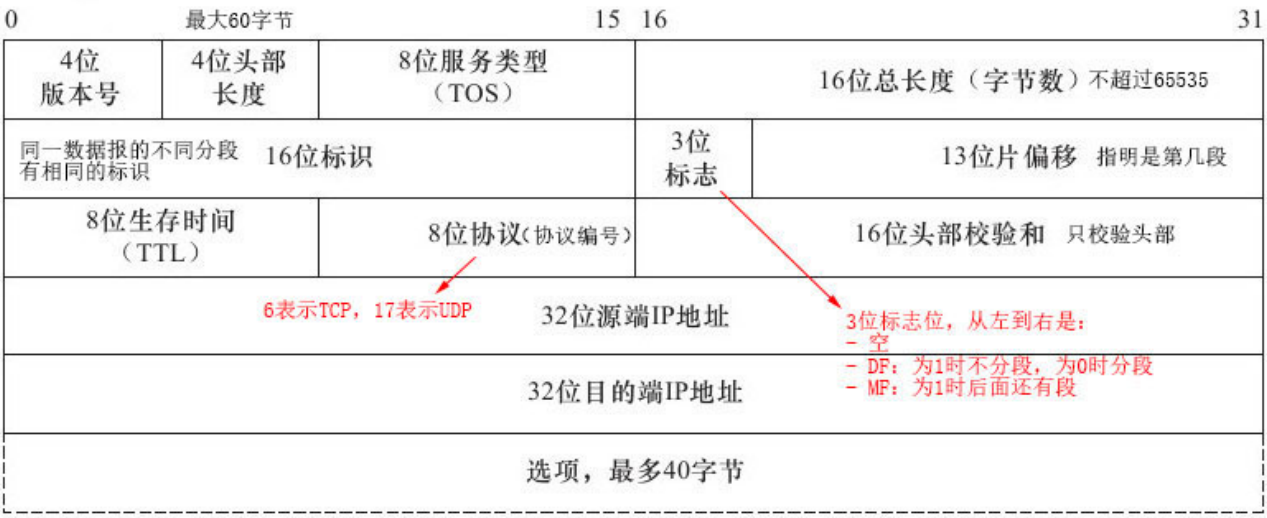
对于拥塞的解决方案：

- 提早增加带宽和升级设备。
- 使用流量感知路由可以将流量拆分到多个路径，即选择多个较短的路由而不是一个最短的路由。
- 使用准入控制和流量调节来降低网络的负载。准入控制只允许有权限的路由器发送数据包，流量调节是接收方发送一个流量抑制包来通知源主机降低发送速率。流量抑制包可以由任何其他包捎带，称为显示拥塞通知。抑制包可以设定为对其沿途的各个路由器都有抑制作用，称为逐跳后压。
- 当没有其他解决办法时，只能丢弃无法传递的数据包，称为负载脱落。

4.Internet

IPv4数据报

IPv4数据报包含一个头部和一个正文，头部包含如下内容



校验和字段首先置零，将整个头按16位字的补码相加求和再取补码，放入校验字段。

当一个数据报超过帧的长度，需要分段。

- 一个数据报的不同分段，有相同的数据报标识，表明是一个段。

- 标志位中的DF为0，表明该数据报已经分段；除最后一段外，其他段的MF为1，表明后面还有段。
- 数据报中的载荷，按字节编号，从0开始。片偏移量中记录的该段的起始字节号/8，比如以太网帧有效载荷为1500字节，IPv4数据报头为20字节，因此第一段最长1480字节（字节编号为0~1479），第二段的第一个字节为1480，片偏移量为1480/8=158。

IPv4地址

两级IP地址

IPv4地址由4组8位二进制组成，分为5类：

- A类：二进制开头为0，第一组是1~126；前8位是网络号，后24位是主机号。
- B类：二进制开头为10，第一组是128~191；前16位是网络号，后16位是主机号。
- C类：二进制开头为110，第一组是192~223；前24位是网络号，后8位是主机号。
- D类：二进制开头为1110，第一组是224~239，是多播地址。使用此类地址和UDP协议，实现IP组播。
- E类：二进制开头为1111，第一组是240~255，保留。

另外一些特殊的地址：

- 主机号全为0的，表示本网络自身。
- 主机号全为1的，表示本网络的广播地址。
- 127.0.0.0用于环路检测。
- 32位全0表示本网络上的本主机。
- 32位全1表示本网络的广播地址。

地址不够用，LAN用私有地址，公用一个公网地址，通过网络地址转换NAT将私有地址和公网地址的转换。有3类：

- A类：10.0.0.0~10.255.255.255
- B类：172.16.0.0~172.31.255.255
- C类：192.168.0.0~192.168.255.255

子网划分

两级IP地址限制了地址空间的利用率。通过子网掩码确定IP地址的前缀，包括网络号和子网号，网络号对应上述的地址划分，在此基础上划分子网。子网号不能全0或全1。

无类域间路由CIDR

指定网络号的位数，地址格式例如194.24.0.0/21中的21为网络号的位数，消除了ABC类地址划分。寻址时按照最长匹配前缀，寻址网络号最长的能够匹配的网络。主机号不能全0或全1。

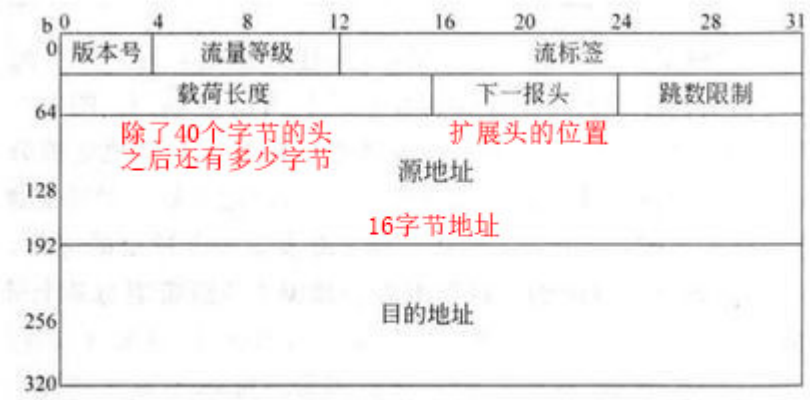
使用路由聚合将多个相邻地址块合并成一个大的地址块，减少路由表项。

移动IP

移动设备使用主IP和辅IP，辅IP是移动设备目前所处的网络给与的，实现使用固定IP在不同网段中漫游。

IPv6

#####IPv6头



IPv6 报头结构

IPv6地址

8组2字节的地址，共16字节，使用冒号间隔，一般用十六进制表示。前导零可以省略；如果一组全为零可以省略；IPv4的地址前加一对冒号。

ARP协议

地址解析协议，将IP地址转换为MAC地址，工作在网络层。每个主机都有一个ARP高速缓存，保存使用过的IP地址到MAC地址的映射，如果没有对应的项，则通过广播发现。

DHCP协议

动态主机配置协议，给主机动态地分配IP地址，工作在应用层，使用UDP协议。获取IP地址的过程如下

1. 客户机广播"DHCP发现"消息；
2. 服务器广播"DHCP提供"消息，包含想分配给此客户机的IP和相关配置信息；
3. 客户机若接收此IP，则广播"DHCP请求"消息，表示请求该IP地址。
4. 服务器广播"DHCP确认"消息，确认分配此IP地址。

动态分配的IP地址有租用期（如24小时），过期后重新请求。

ICMP协议

网际控制报文协议，是IP层协议，分为ICMP差错报告报文和ICMP询问报文。ICMP差错控制报文用于报告差错和异常情况，包含：终点不可达、源点抑制、时间超过、参数问题、改变路由（重定向）。ICMP包含回送请求和回答报文、时间戳请求和回答报文、掩码地址请求和回答报文、路由器询问和通告报文。PING命令使用的回送请求和回答报文，tracert命令使用的时间超过报文。

五、传输层

1.概述

传输层的功能：提供进程间的逻辑通信、复用与分用（使用端口地址区分进程）、差错控制。

传输层寻址

传输层使用端口号区分不同的进程，并实现对网络层连接的复用。端口号长度为16位，能够表示65536个不同的端口号。其中0~1023是熟知的端口号，已经指派给了一些常用应用程序；1024~49151为登记端口，49152~65535为短临时端口。

比较常用的端口号：

FTP	TELENT	SMTP	DNS	TFTP	HTTP	SNMP
21	23	25	53	69	80	161

进程之间在网络中的通信使用**套接字**，为（主机IP地址，端口号）。

2.UDP协议

UDP只提供信道的复用和分用、差错控制。UDP无需建立连接，尽最大努力交付。

UDP 数据报

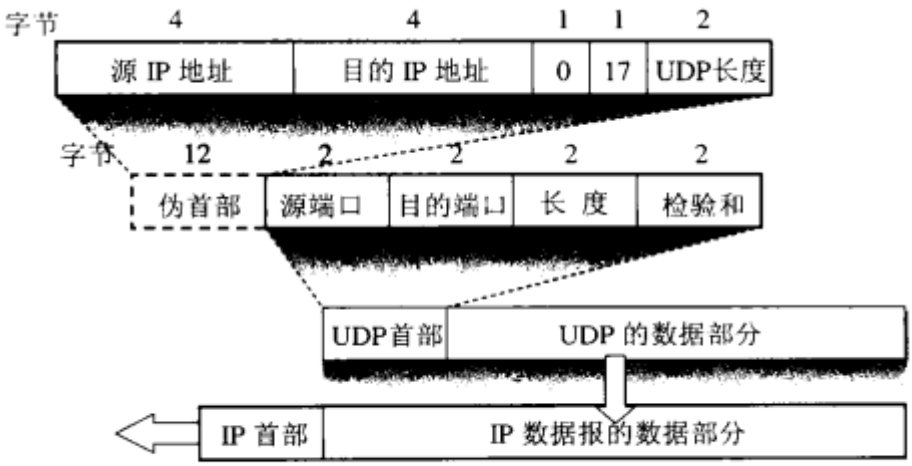


图 5-5　UDP 用户数据报的首部和伪首部

UDP首部包含2字节的源端口号、2字节的目的端口号、2字节的UDP长度、2字节的UDP校验和。

为计算校验和需要将UDP首部增加12字节的伪首部使其有16字节，伪首部包含4字节的源IP地址、4字节的目的IP地址、1字节全零、1字节的UDP协议号17、2字节的UDP长度。

校验和计算整个UDP数据报（伪首部、首部、数据），校验时先将校验和字段置零，将数据字段用零填充至偶数字节，按16位字的补码求和再求补码，放入校验和。如果不使用校验和，则该字段全零。

3.TCP协议

TCP协议是面向连接的协议，提供可靠的交付服务，保证数据无差错、不丢失、不重复、不失序。TCP是全双工的。

TCP报文段

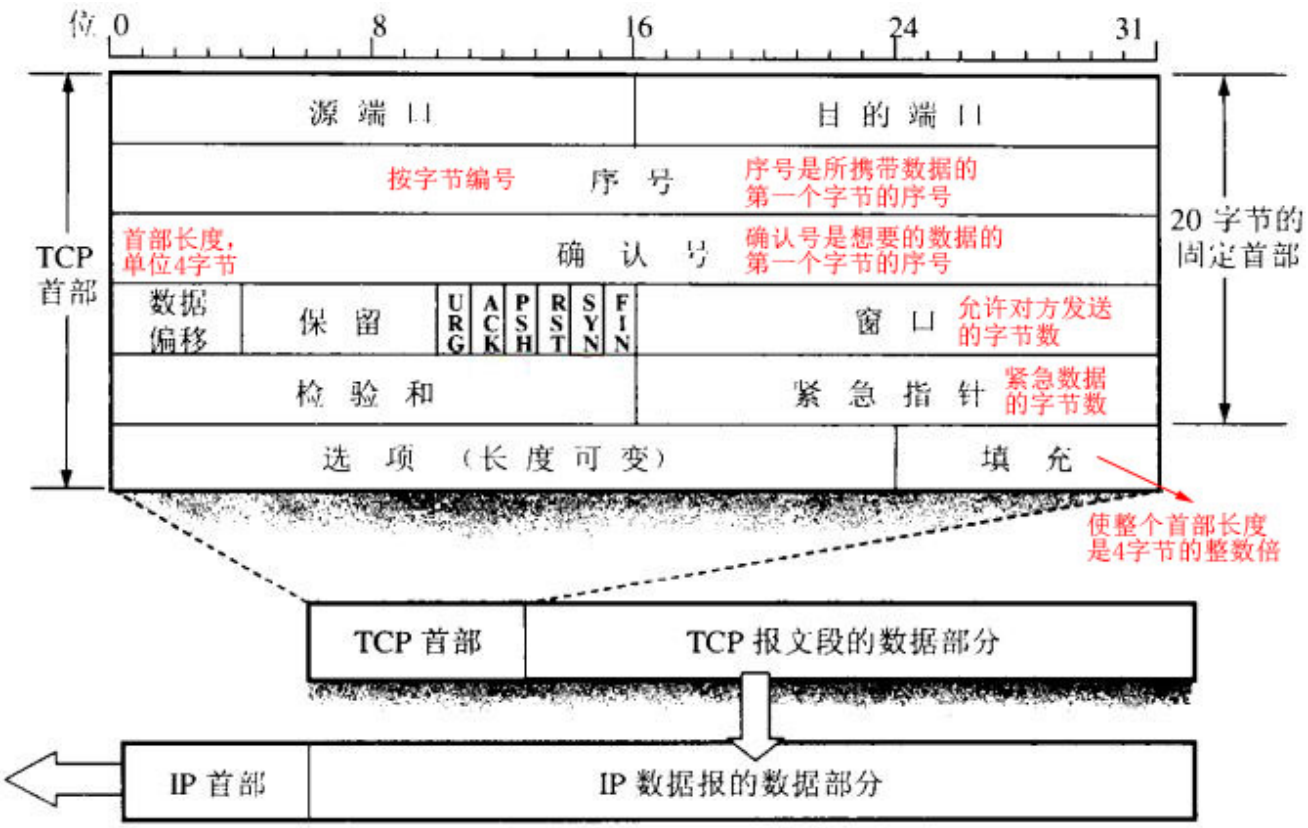


图 5-14　TCP 报文段的首部格式

- URG位：URG=于1时，紧急指针字段有效。
- ACK位：ACK=1时，确认号字段有效。
- PSH位：TCP接收到PSH=1的报文段，会尽快交付给进程。
- RST位：RST=1时，表明TCP连接严重错误，需要释放连接再重新建立。
- SYN位：SYN=1时，为连接请求报文或连接就收报文。

  当SYN=1、ACK=0时， 是一个连接请求报文，请求建立TCP连接；当SYN=1、ACK=1时，是一个连接接收报文。
- FIN位：终止位，当释放连接时FIN=1。

TCP 连接

TCP 连接的建立，称为三次握手。

- 客户机主动打开端口，发送连接请求报文，SYN=1，ACK =0，seq = x；
- 服务器收到请求后，如果同意连接，为此连接分配TCP 缓存和变量，发送连接确认报文，SYN=1，ACK=1，seq = y，ack =x+1；

- 客户机收到确认报文后，分配TCP缓存和变量，想服务器确认，SYN=0，ACK=1，seq=x+1，ack=y+1。

TCP连接的释放

- 客户机打算关闭连接，会发送一个连接释放报文，FIN=1，seq=u；服务器收到此报文，同意释放连接，发送确认，ACK=1，ack=u+1。此时只是客户机不再想服务器发送数据，并不影响服务器向客户机发送数据。
- 同样的，服务器打算关闭连接，也执行同样的步骤。

TCP可靠传输

- 报文段的序号保证数据有序。
- 确认机制保证报文段到达。
- 重传。当发生确认超时时，会触发重传；当收到3个对同一报文段的冗余ack时，说明后面的报文段丢失，触发快速重传。

TCP重传计时器

当发送方发送一个报文段，立即触发计时，若该报文段的确认时间R超过重传超时值RTO，则触发重传。RTO是一个动态的值，$\mathrm{RTO}=\mathrm{SRTT}+4\times\mathrm{RTTVAR}$，其中SRTT为平滑往返时间，RTTVAR往返时间变化。

当一个报文段的确认时间没有超过RTO，则需要对RTO进行更新：

$$\mathrm{SRTT}=\alpha\mathrm{SRTT}+(1-\alpha)\mathrm{R}$$

$$\mathrm{RTTVAR}=\beta\mathrm{RTTVAR}+(1-\beta)\mid\mathrm{SRTT}-\mathrm{R}\mid$$

$$\mathrm{RTO}=\mathrm{SRTT}+4\times\mathrm{RTTVAR}$$

其中，α通常取7/8，β通常取3/4。

Karn算法建议上述过程不更新重传段的时间，同时每次连续重传的超时间隔值加倍。

TCP流量控制

基于滑动窗口协议的流量控制机制，发送方根据接收方的接收窗口rwind和网络的拥塞窗口cwind中最小的，决定发送数据的多少。rwind由接收方的缓存决定，cwind由拥塞控制决定。TCP拥塞控制实际是维护cwind的大小。

TCP拥塞控制

拥塞控制是由发送方自己来实现控制的，自己维护cwind。而流量控制需要接收方反馈rwind。

- 慢启动算法：设置cwind=1，单位为最大报文长度；每收到一个报文段的确认，将cwind加1，实际表现为每经过RTT时间cwind翻倍；至达到慢启动阈值，切换至线性增加。
- 拥塞避免算法：即上述的线性增加，每经过一个RTT，cwind加1。
- 当网络出现拥塞时，即发生确认超时，将慢启动阈值设置为当前cwind的一半，cwind设置为1，重新开始慢启动算法和拥塞避免算法。
- 快重传机制：当发送方连续收到3个对同一报文段的冗余ACK，表明后面的报文段没有接收到，此时发送方不需要等到超时就对之后的报文段重传，称为快重传机制。
- 快恢复机制：当发送方连续收到3个对同一报文段的冗余ACK，慢启动阈值设定为当前cwind的一半，cwind设定为改变后的阈值，直接执行线性增加。

六、应用层

1.DNS

- 工作再UDP协议之上，端口号53。
- 域名与域名服务器。一共有13个根域名服务器，用于管辖顶级域名；下面是顶级域名服务器、授权域名服务器、本地域名服务器。
- 域名解析有两种方式：递归查询、迭代查询
    - 递归查询：当客户机向本地服务器请求域名查询时，该本地服务器会代替客户机完成域名查询工作，最后返还给客户机答案。
    - 迭代查询：本地服务器本地没有对域名的记录时，会向根服务器查询，但是根服务器只会返回部分答案，告诉本地服务器下一步去哪一台的服务器查询；本地服务器向根服务器告知的服务器查询，也可能不会得到完整的答案，继续到下一级查询。重复此过程至得到答案。

2.电子邮件

电子邮件系统有3个组成构件：用户代理、邮件服务器、邮件协议（如SMTP、POP3、IMAP）。

- SMTP使用推送的通信方式，即用户代理发送邮件时，SMTP协议主动向邮件服务器推送邮件。
- POP3使用拉取的通信方式，即用户读取邮件时，从邮件服务器拉取邮件。
- 电子邮件格式包含信封和正文两部分，正文首部包含`to:`、`from`、`subject`，正文主体是邮件内容。信封由SMTP自动填写。
- SMTP邮件格式简单，而MIME格式可以包含多媒体内容。

3.FTP

- FTP工作时使用两个并行的TCP连接，一个是控制连接，端口号21；一个是数据连接，端口号22。

4.万维网

- 超文本传输协议HTTP、统一资源定位符URL、Cookie、超文本标记语言HTML
- 服务器端是由一个前端模块和多个处理模块构成，前端模块接受所有请求，将该请求交给一个处理模块。
- HTTP协议本身是无连接的，但是使用TCP协议。