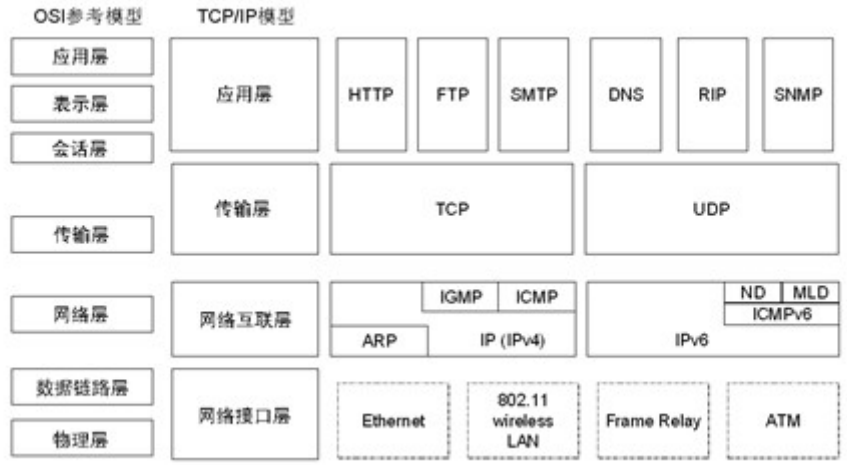
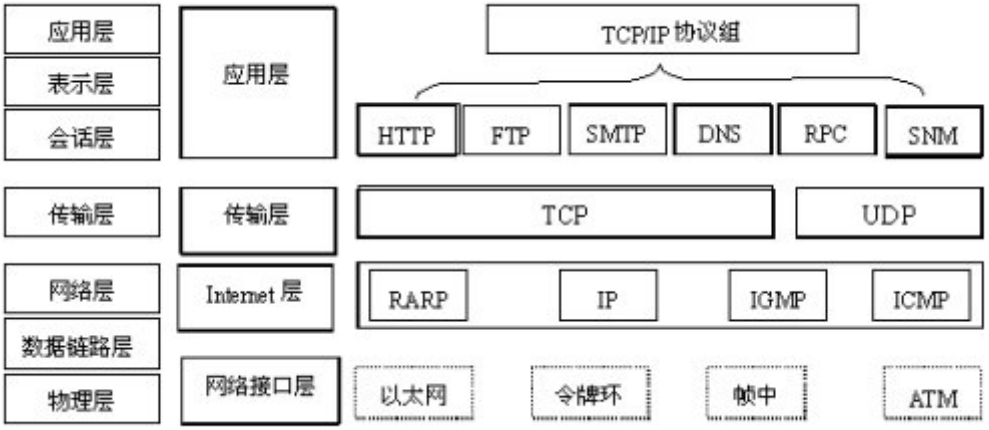


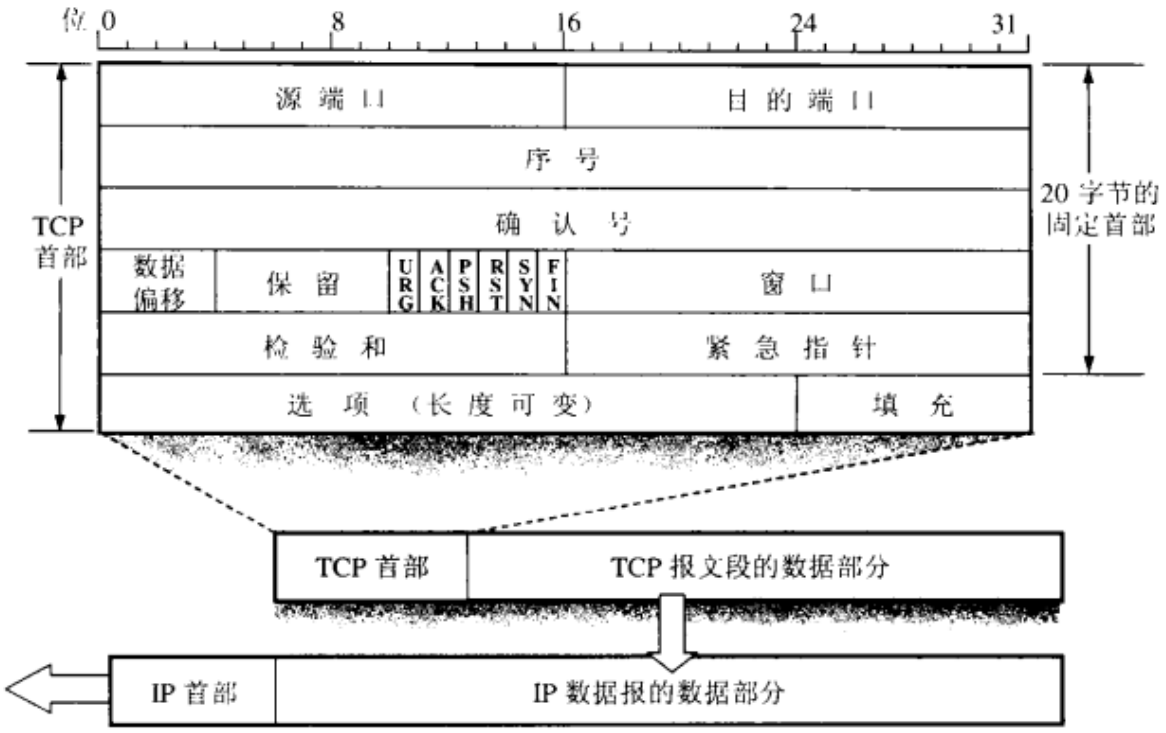


TCP：

TCP(Transfer Control Protocol)是面向连接的传输层协议，采用字节流传输数据。所谓面向连接，就是当计算机双方通信时必需经过先建立连接，然后传送数据，最后拆除连接三个过程。

- **TCP报文段格式**

TCP报文段包括协议首部和数据两部分，协议首部的固定部分有20个字节，首部的固定部分后面是选项部分。



报文段首部各个字段的含义。

1. 源端口号以及目的端口号，各占2个字节，端口是传输层和应用层的服务　　接口，用于寻找发送端和接收端的进程，一般来讲，通过端口号和IP地址，可以唯一确定一个TCP连接，在网络编程中，通常被称为一个socket接口。

2. 序号，占4字节，用来标识从TCP发送端向TCP接收端发送的数据字节流。

3. 确认序号，占4字节，包含发送确认的一端所期望收到的下一个序号，因此，确认序号应该是上次已经成功收到数据字节序号加1.

4. 数据偏移，占4位，用于指出TCP首部长度，若不存在选项，则这个值为20字节，数据偏移的最大值为60字节。

5. 保留字段占6位，暂时可忽略，值全为0

6. 标志位

URG（紧急）：为1时表明紧急指针字段有效

ACK（确认）：为1时表明确认号字段有效

PSH（推送）：为1时接收方应尽快将这个报文段交给应用层

RST（复位）：为1时表明TCP连接出现故障必须重建连接

SYN（同步）：在连接建立时用来同步序号

FIN （终止）： 为1时表明发送端数据发送完毕要求释放连接

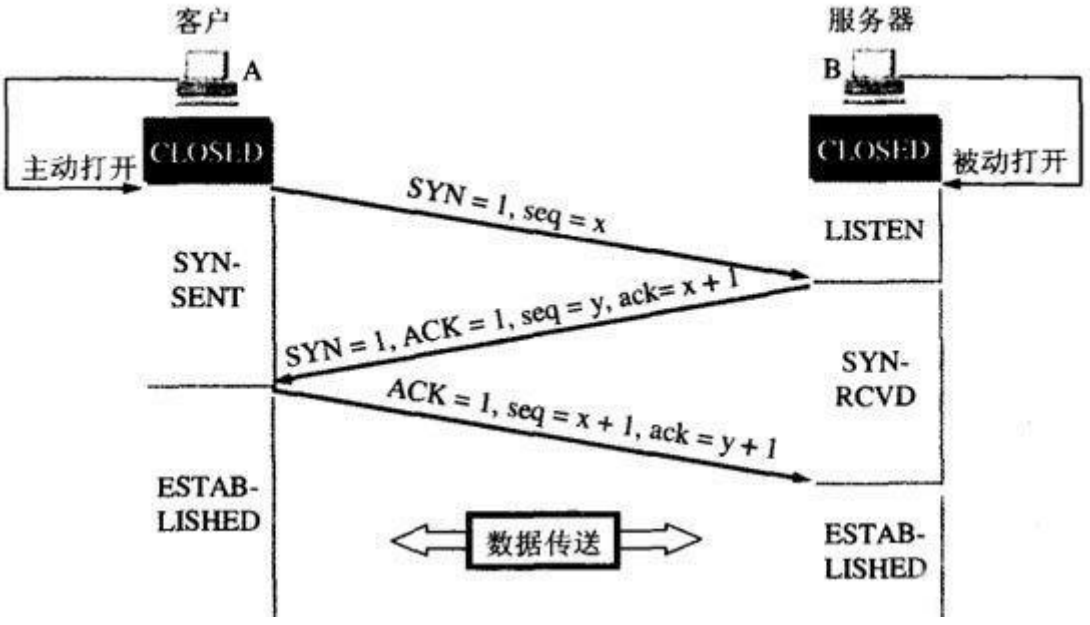
7. 接收窗口占2个字节，用于流量控制和拥塞控制，表示当前接收缓冲区的大小。在计算机网络中，通常是用接收方的接收能力的大小来控制发送方的数据发送量。TCP连接的一端根据缓冲区大小确定自己的接收窗口值，告诉对方，使对方可以确定发送数据的字节数。

8. 校验和占2个字节，范围包括首部和数据两部分。

9. 选项是可选的，默认情况是不选。

- **三次握手**

连接建立、数据传送和连接释放。

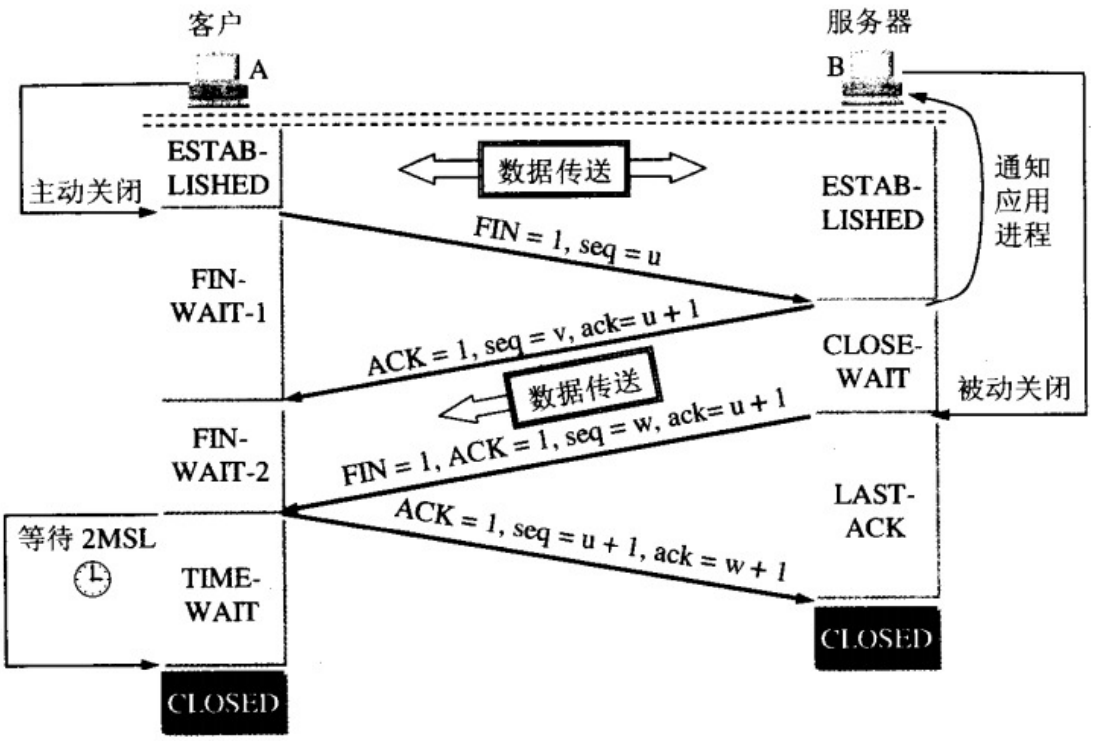


第一步，是请求端(客户端)发送一个包含SYN即同步(Synchronize)标志的TCP报文，SYN同步报文会指明客户端使用的端口以及TCP连接的初始序号。**并进入SYN_SENT状态，等待服务器确认。**

第二步，服务器收到客户端的SYN报文后，同意连接将返回一个SYN+ACK的确认报文，**并为该TCP连接分配TCP缓存和变量。**同时TCP序号被加一，ACK即确认。

第三步，客户端也返回一个确认报文ACK给服务器端，**并且也要给该连接分配缓存和变量。此包发送完毕，客户端和服务器进入ESTABLISHED（TCP连接成功）状态。**同样TCP序列号被加一，到此一个TCP连接完成。

- **四次挥手**



由于TCP连接是全双工的，因此每个方向都必须单独进行关闭。这原则是当一方完成它的数据发送任务后就能发送一个FIN来终止这个方向的连接。收到一个 FIN只意味着这一方向上没有数据流动，一个TCP连接

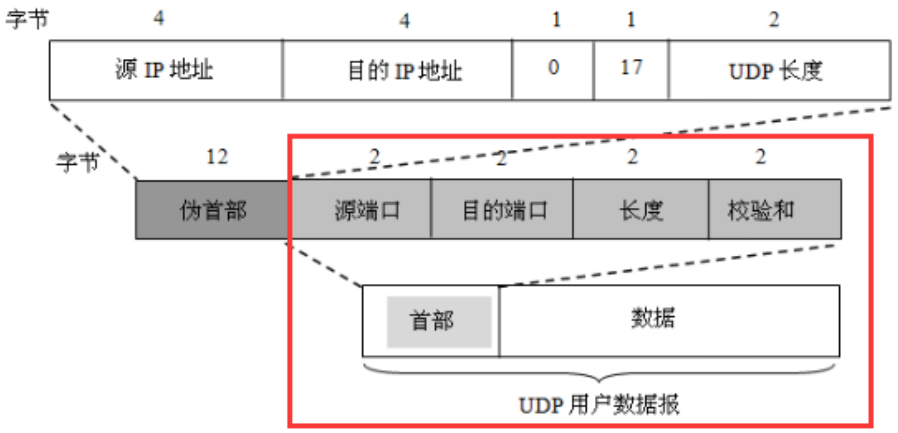
在收到一个FIN后仍能发送数据。首先进行关闭的一方将执行主动关闭，而另一方执行被动关闭。

1. TCP客户端发送一个FIN，用来关闭客户到服务器的数据传送。
2. 服务器收到这个FIN，它发回一个ACK，确认序号为收到的序号加1。和SYN一样，一个FIN将占用一个序号。
3. 服务器关闭客户端的连接，发送一个FIN给客户端。
4. 客户端发回ACK报文确认，并将确认序号设置为收到序号加1。

UDP：

是TCP/IP协议簇中无连接的运输层协议。

- **UDP协议格式格式**



由两部分组成：首部和数据。首部仅有8个字节，包括源端口和目的端口，长度（UDP用于数据报的长度）、校验和。

1. TCP（传输控制协议）是面向连接的，传输数据安全，稳定，效率相对较低。

2. UDP（用户数据报协议）是面向无连接的，传输数据不安全，效率较高。

UDP方式的同一个网络连接对象，可以发送到达不同服务器端IP或端口的数据包，这点是TCP方式无法做到的。

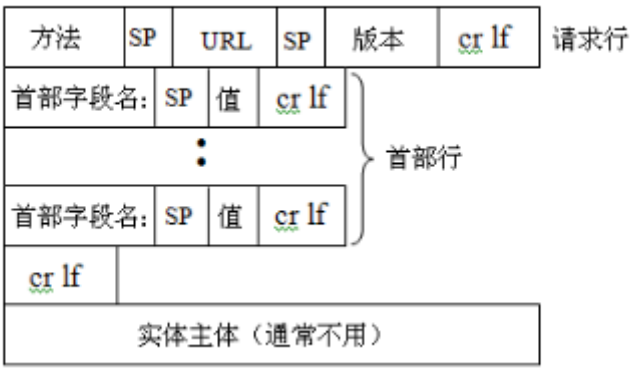
HTTP协议：

超文本传输协议。互联网上应用最广泛的网络协议，是应用层协议。基于TCP协议之上的请求/响应式协议，即客户端和服务端建立连接后，向服务端发送请求，服务器接到请求后，给予相应的响应信息。默认端口号80。

- **HTTP报文**

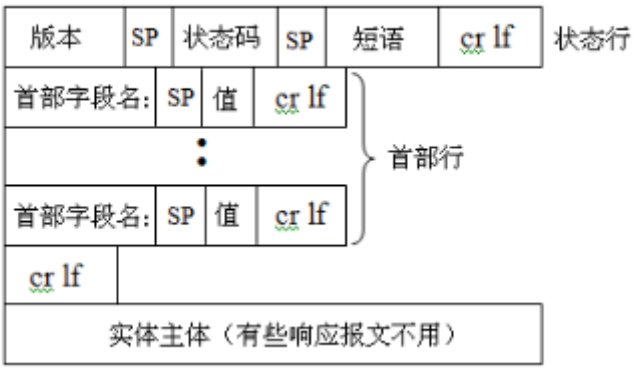
  HTTP协议是基于TCP协议之上的请求/响应式协议

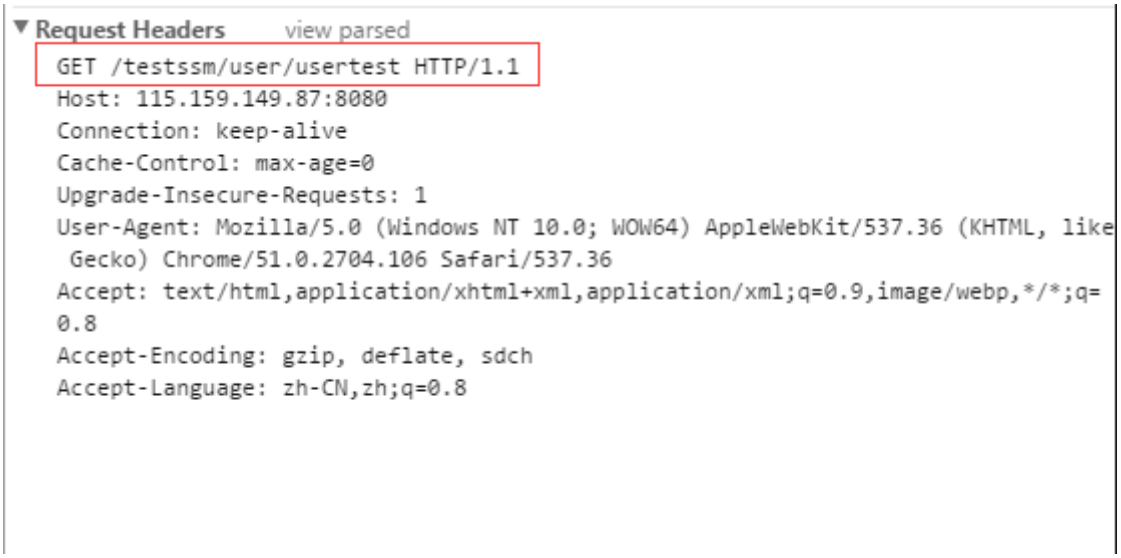
  请求报文格式：



HTTP请求报文由请求行、首部行和实体主体组成，由浏览器发送给服务器。上面这张图中　　SP表示空格，cr lf表示回车和换行。

响应报文格式：





由状态行、首部行和实体主体组成。

- **HTTP请求方法和响应状态码**

方法（操作）	含义
OPTION	请求一些选项的信息
GET	请求读取由 URL 所标志的信息
HEAD	请求读取由 URL 所标志的信息的首部
POST	给服务器添加信息
PUT	在指明的 URL 下存储一个文档
DELETE	删除 URL 指明的资源
TRACE	进行环回测试的请求报文
CONNECT	用于代理服务器

状态码	含义	例子
1xx	通知信息	请求收到了或正在处理
2xx	成功	接受或知道了
3xx	重定向	表示要完成的请求还要采取进一步的动作
4xx	客户差错	请求中有语法错误或不能完成
5xx	服务器差错	服务器失效、无法响应或完成请求

- **HTTPS和HTTP**

HTTPS（全称：Hyper Text Transfer Protocol over Secure Socket Layer），是以安全为目标的HTTP通道。即HTTP下加入SSL层，HTTPS的安全基础是SSL，加密的详细内容就需要SSL。它是一个URI scheme（抽象标识符体系），句法类同http:体系。用于安全的HTTP数据传输。https:URL表明它使用了HTTP，但HTTPS存在不同于HTTP的默认端口及一个加密/身份验证层（在HTTP与TCP之间）。

超文本传输协议HTTP协议被用于在Web浏览器和网站服务器之间传递信息。HTTP协议以明文方式发送内容，不提供任何方式的数据加密，如果攻击者截取了Web浏览器和网站服务器之间的传输报文，就可以直接读懂其中的信息，因此HTTP协议不适合传输一些敏感信息，比如信用卡号、密码等。

为了解决HTTP协议的这一缺陷，需要使用另一种协议：安全套接字层超文本传输协议HTTPS。为了数据传输的安全，HTTPS在HTTP的基础上加入了SSL协议，SSL依靠证书来验证服务器的身份，并为浏览器和服务器之间的通信加密。

HTTPS和HTTP的区别主要为以下四点：

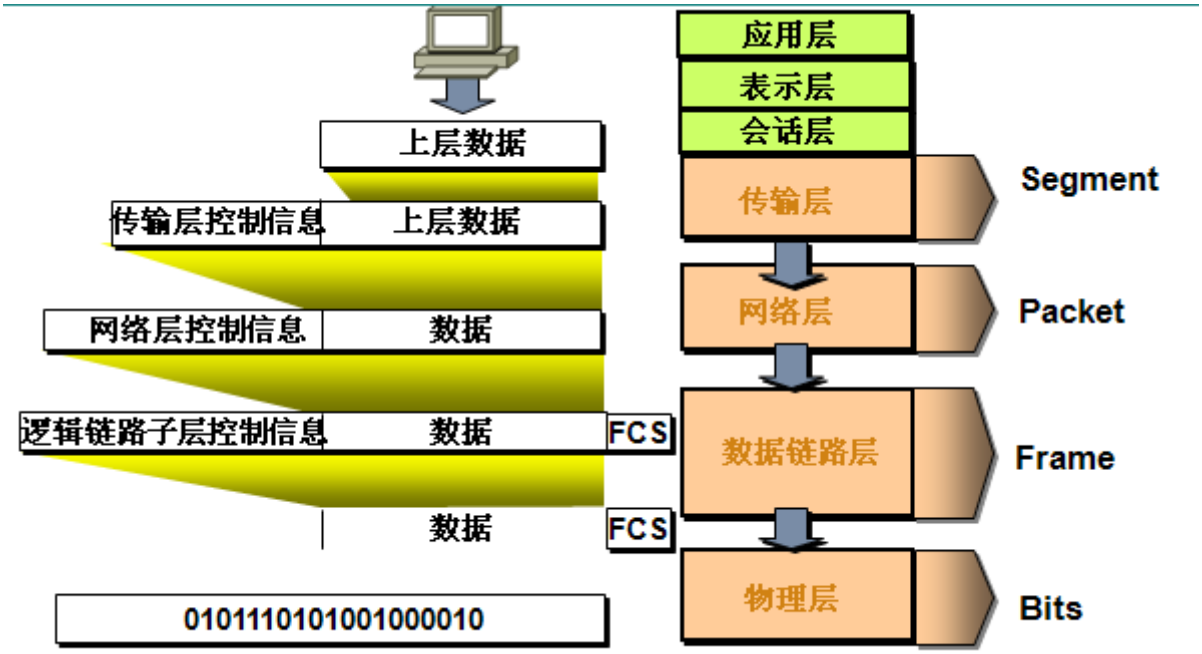
1、https协议需要到ca申请证书，一般免费证书很少，需要交费。

2、http是超文本传输协议，信息是明文传输，https 则是具有安全性的ssl加密传输协议。

3、http和https使用的是完全不同的连接方式，用的端口也不一样，前者是80，后者是443。

4、http的连接很简单，是无状态的；HTTPS协议是由SSL+HTTP协议构建的可进行加密传输、身份认证的网络协议，比http协议安全。

数据封装(Data Encapsulation)是指将协议数据单元(PDU)封装在一组协议头和协议尾中的过程。在OSI七层参考模型中，每层主要负责与其它机器上的对等层进行通信。该过程是在协议数据单元(PDU)中实现的，其中每层的PDU一般由本层的协议头、协议尾和数据封装构成。



- 数据发送处理过程

(1)应用层将数据交给传输层，传输层添加上TCP的控制信息(称为TCP头部)，这个数据单元称为段(Segment)，加入控制信息的过程称为封装。然后，将段交给网络层。

(2)网络层接收到段，再添加上IP头部，这个数据单元称为包(Packet)。然后，将包交给数据链路层。

(3) 数据链路层接收到包，再添加上MAC头部和尾部，这个数据单元称为帧(Frame)。然后，将帧交给物理层。

(4)物理层将接收到的数据转化为比特流，然后在网线中传送。

- 数据接收处理过程

(1)物理层接收到比特流，经过处理后将数据交给数据链路层。

(2)数据链路层将接收到的数据转化为数据帧，再除去MAC头部和尾部，这个除去控制信息的过程称为解封，然后将包交给网络层。

(3)网络层接收到包，再除去IP头部，然后将段交给传输层。

(4)传输层接收到段，再除去TCP头部，然后将数据交给应用层。

IP地址：

用来标识网络中的一个通信实体的地址。IPv4协议，该协议规定每个IP地址由4个0-255之间的数字组成，例如10.0.120.34。

1. 127.0.0.1 本机地址

2. 192.168.0.0--192.168.255.255为私有地址，属于非注册地址，专门为组织机构内部使用。

- **环回地址**

环回地址是主机用于向自身发送通信的一个特殊地址（也就是一个特殊的目的地址）。IPv4的环回地址为：127.0.0.0到127.255.255.255都是环回地址（只是有两个特殊的保留），此地址中的任何地址都不会出现在网络中。网络号为127的地址根本就不是一个网络地址（因为产生的IP数据报就不会到达外部网络接口中，是不离开主机的包）

可以这么说：同一台主机上的两项服务若使用环回地址而非分配的主机地址，就可以绕开TCP/IP协议栈的下层。（也就是说：不用再通过什么

链路层，物理层，以太网传出去了，而是可以直接在自己的网络层，运输层进行处理了）

- **127.0.0.1**

当操作系统初始化本机的TCP/IP协议栈时，设置协议栈本身的IP地址为127.0.0.1（保留地址），并注入路由表。当IP层接收到目的地址为127.0.0.1（准确的说是：网络号为127的IP）的数据包时，不调用网卡驱动进行二次封装，而是立即转发到本机IP层进行处理，由于不涉及底层操作。因此，ping 127.0.0.1一般作为测试本机TCP/IP协议栈正常与否的判断之一。

所以说：127.0.0.1是保留地址之一，只是被经常的使用，来检验本机TCP/IP协议栈而已。

如果我们可以ping通的话，就说明：本机的网卡和IP协议安装都没有问题。（跟我们当前主机有没有联网没有一点关系）

```
doubi@doubi-Inspiron-3421:~$ ping 127.0.0.1
PING 127.0.0.1 (127.0.0.1) 56(84) bytes of data.
64 bytes from 127.0.0.1: icmp_seq=1 ttl=64 time=0.044 ms
64 bytes from 127.0.0.1: icmp_seq=2 ttl=64 time=0.060 ms
64 bytes from 127.0.0.1: icmp_seq=3 ttl=64 time=0.054 ms
64 bytes from 127.0.0.1: icmp_seq=4 ttl=64 time=0.067 ms
64 bytes from 127.0.0.1: icmp_seq=5 ttl=64 time=0.061 ms
```


- **localhost**

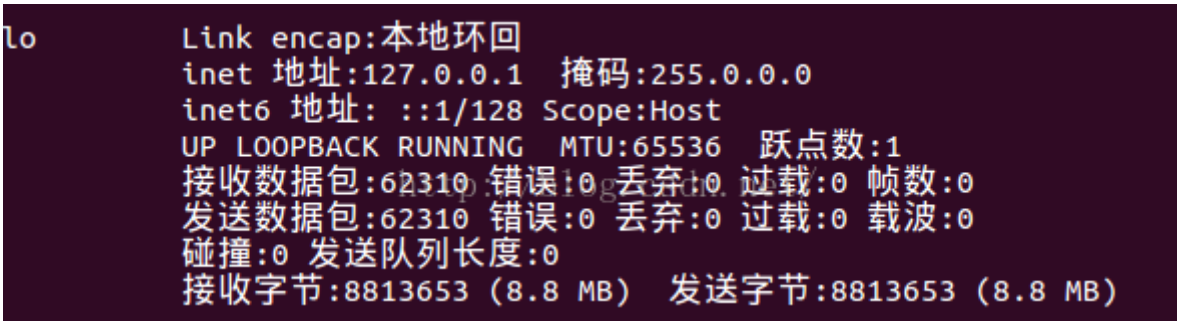
localhost首先是一个域名（如同：www.baidu.com），也是本机地址，它可以被配置为任意的IP地址（也就是说，可以通过hosts这个文件进行更改的），不过通常情况下都指向：（如下）

IPv4：表示　　127.0.0.1

IPv6：表示　　[::1]

整个127.*网段通常被用作loopback网络接口的默认地址，按照惯例通常设置为127.0.0.1。我们当前这个主机上的这个地址，别人不能访

问，即使访问，也是访问自己。因为每一台TCP/IP协议栈的设备基本上都有local/127.0.0.1



- **本机IP**

本机IP，我们可以理解为本机有三块网卡，一块网卡叫做loopback（虚拟网卡），一块叫做ethernet（有线网卡），一块叫做wlan（你的无线网卡），

联网，网卡传输，受防火墙和网卡限制

用于本机和外部访问

端口：

IP地址用来标识一台计算机，但是一台计算机上可能提供多种网络应用程序,用到端口区分这些不同的程序。**端口号的范围是0到65536，但是0到1024是为特权服务保留的端口号。**

域名：

Domain Name System，域名系统。一个IP地址可以对应多个域名，一个域名只能对应一个IP地址。在网络中传输的数据，全部是以IP地址作为地址标识，所以在实际传输数据以前需要将域名转换为IP地址，实现这种功能的服务器称之为DNS服务器，叫做域名解析。例如当用户在浏览器输入域名时，浏览器首先请求DNS服务器，将域名转换为IP地址，然后将转换后的IP地址反馈给浏览器，然后再进行实际的数据传输。当DNS不正常工作时，只能通过IP地址访问设备。所以IP地址的使用要比域名通用一些。

URL：

IP地址唯一标识了Internet上的计算机，而URL则标识了这些计算机上的资源。类 URL 代表一个统一资源定位符，在www上，每一信息资源都有统一且唯一的地址，该地址就叫URL （ Uniform Resource Locator)，它是www的统一资源定位符。URL由4部分组成：协议 、存放资源的主机域名、资源文件名和端口号。如果未指定该端口号，则使用协议默认的端口。例如http 协议的默认端口为 80。 在浏览器中访问网页时，地址栏显示的地址就是URL。

http://mail.163.com/index.html

1)http://:这个是协议，也就是HTTP超文本传输协议，也就是网页在网上传输的协议。

2）mail：这个是服务器名，代表着是一个邮箱服务器，所以是mail.

3)163.com:这个是域名，是用来定位网站的独一无二的名字。

4)mail.163.com：这个是网站名，由服务器名+域名组成。

5）/：这个是根目录，也就是说，通过网站名找到服务器，然后在服务器存放网页的根目录

6:）index.html：这个是根目录下的默认网页（当然，163的默认网页是不是这个我不知道，只是大部分的默认网页，都是index.html）

7）http://mail.163.com/index.html:这个叫做URL，统一资源定位符，全球性地址，用于定位网上的资源。