CS339 计算机网络整理

1. 网络概述

1. 网络基本指标及其概念

- 什么是因特网：是一个世界范围的计算机网络，其中接入互联的计算设备叫做主机或者端系统。
- 互联网的重要特点：连通性和资源共享
- 计算机网络的基本性能指标

指标名称	单位	说明
速率	bit/s, kbit/s, Mbit/s, Gbit/s	数据的传输速率，其往往指的是额定速率或者标称速率，并不是实际运行速率。
带宽（无线网络）	赫，千赫等	指信号具有的频带宽度
带宽（有限网络）	bit/s	用来表示网络中某通道传送数据的能力，表示在单位时间内网络中某信道所能通过的最高数据量。
吞吐量	bit, byte	用来表示在单位时间内通过某个网络的数据量，其受网络的带宽或者速率的限制。
时延	ms	是指数据从网络的一端传送到另外一端所需的时间。 其等于发送+传播+处理+排队时延的总和。 ->发送时延：数据帧从结点进入传输媒体所需要的时间，等于从发送第一比特开始到最后一个比特发送完毕的时延 ->传播时延：电磁播在信道中传播一定距离花费的时间 ->处理时延：主机或者路由器在接到分组时需要进行一系列分析，提取，校验和查找的工作，这些操作需要花费一定时间。 ->排队时延：不一定存在，时分组在路由器输入输出的排队队列中等待所经历的时间。

  - 对于高速链路，提高的是数据的发送速率而非传播速率，即在高速链路中比特并不会传播得更快。
  - 提高链路带宽，减小了数据的发送时延。

2. 三种协议体系结构

- 网络协议：是为了进行网络中的数据交换而建立的规则、标准或约定。
- OSI七层协议体系和TCP/IP协议比较

OSI（七层）	TCP/IP（四层）	五层协议的体系结构
应用层		
表示层	应用层（DNS, HTTP, SMTP等）	应用层
会话层		
传输层	运输层（TCP或UDP）	传输层
网络层	网际层IP	网络层
链路层		数据链路层
物理层	网络接口层	物理层

- PDU：协议数据单元，对等层次之间传送的数据单位称为该层的协议数据单元PDU。
- SUD：服务数据单元，上一层的PDU到了下一层就是SDU。
- 五层模型中，只有数据链路层既添加了首部又添加了尾部。因为链路层要负责做差错检验。

2. 物理层

1. 物理层的基本功能和最小单位

- 物理层的功能：考虑怎样才能在连接各种计算机的传输媒体上传输数据流。其最小单位为比特。

2. 编码和调制的作用

- 数据（也分为数字数据和模拟数据）为了传输到目的地都必须转变为信号，把数据转变为模拟信号的过程称为**调制**，把数据转变为数字信号的过程称为**编码**。
- 数字数据->数字信号：区分0和1这两个数字，具体编码方式有非归零编码、曼彻斯特编码、差分曼彻斯特编码和4B/5B（2^n形式的编码）编码。
- 数字数据->模拟信号：其编码和解码对应了调制解调器的调制和解调过程。其基本调制方法可以分为：（1）ASK 幅移键控：即改变载波信号的振幅来表示1和0；（2）FSK 频移键控：改变载波信号的频率来表示1和0；（3）PSK 相移键控：改变载波信号的相位来表示1和0；（4）QAM正交振幅调制：结合ASK和PSK形成叠加信号。
- 模拟数据->数字信号
- 模拟数据->模拟信号：此调制方式可以使用频分复用，充分利用带宽资源。

3. 无线网络物理层实例和速率计算

- **ZIGBEE**: zigbee是一种低速短距离传输的无线网上协议，底层采用了IEEE802.15.4标准规范的媒体访问层和物理层。其特点有低功耗，低成本的无线网络技术。
- **WI-FI**：是无限局域网标准IEEE802.11的一种实现，其工作于UHF特高频段（分米波），物联网设备，蓝牙等也在使用这个频段，因此2.4GWI-FI干扰非常严重。WIFI的每个信道的有效带宽为20MHz,并且为了防止一组信道之间的干扰，设置了2MHz的隔离带，每两个信道的中心频率为5Hz，同时，2.4GHz频道下互不干扰的信道只有三组。

4. 时分复用、频分复用和码分复用

- 复用：是指通信技术中允许用户共享信道进行通信，从而降低成本提高利用率。
- **频分复用**：FDM 是指所有用户在相同的事件占用不同的频率带宽资源。
- **时分复用**：TDM 是指将时间划分成一段段等长的时分复用帧，每个用户在每个TDM帧中占用固定序号的时隙。每个用户在在不同时间段占用相同的频带宽度。这种复用可能会造成信道的利用率降低，因为用户在被分配到时可能并不需要使用这条信道。
- **码分复用**：CDMA 是指各个用户在相同时间的相同频段使用经过特殊挑选的不同码型。其具体的复用方法为：将一个比特时间划分为m个短的间隔，称为码片。每个站被指派一个唯一的m bit码片序列。若S站要发送信息的数据率为b bits/s，由于一个比特要转换成m个bits的码片，因此其所使用的频带宽度能够提高到原来的m倍。为了区分，每个站被分配的码片必须相互正交。

3. 数据链路层

- **为什么数据链路层需要进行帧定界**

  由于封装成帧后，下一步就是在物理层进行信号传输了，而信号传输是以比特流的形式进行的，因此因此需要在帧外封装定界符和结束符来标志一个帧的开始和结束。

- **每个首部中的“协议”指明的是什么?**

  每层的首部中“协议”指明的是，解包时，这个数据需要交付给什么协议（如MAC帧再交付给IPv4协议进一步解包）；从上向下封装时，则说明的是，是从哪个上层协议拿过来的。

#####

1. 数据链路层的基本功能和最小单位

- 数据链路层：其基本功能为将数据可靠地传递到相邻节点（无法做到远距离传输）。
- 最小单元：数据链路层地PDU为帧
- 数据链路：data link是物理线路加上通信协议控制来实现数据传输，通常这些协议控制和实现都是通过网卡实现的。
- 数据链路层不同协议都要解决的三个问题：

  （1）封装成帧：在一段数据的首尾添上帧首部和帧尾部（应该只有数据链路层才有帧尾）。这一步的主要作用是确定帧的界限。若数据的字符是ASCII码时，可以采用SOH和EOT定界符。用ESC可以用来防误判。

  （2）透明传输：用字节填充的方法，在数据中和定界符的ASCII码一致时进行填充防止错误识别到帧的开头和结尾。

  （3）差错控制：传输过程中可能导致某些比特位出错，需要有差错检验的措施。

2. 差错检验——循环冗余码（必考计算）

在数据链路层的帧中使用的CRC技术。

- 在**发送端**将待发送数据分成K比特的一组，在这组数据后再添加供差错检测的n位冗余码。循环冗余码（FCS）的计算方法如下：

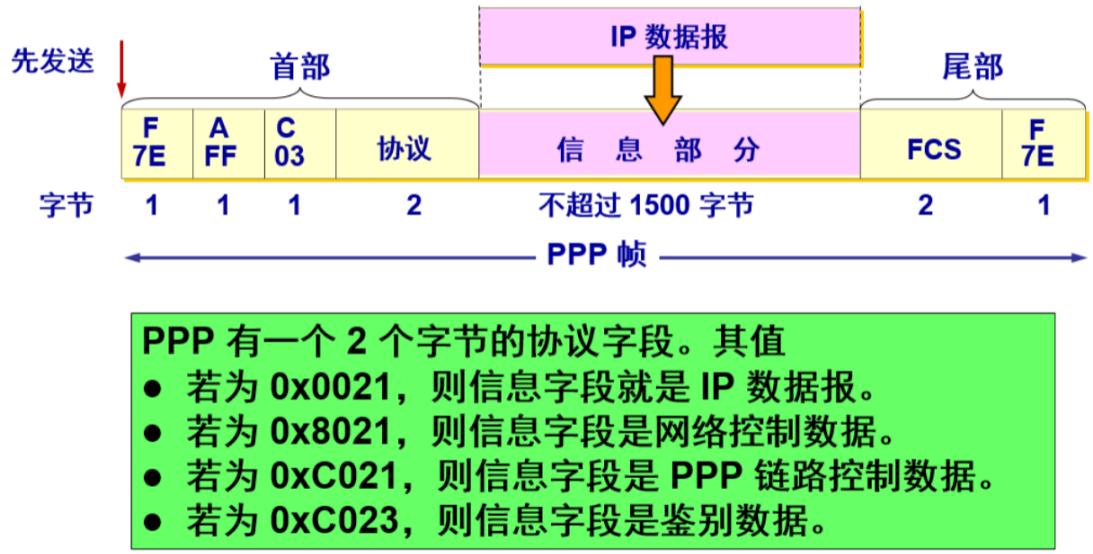
1. 分组，例如其中一组是M。
2. 在M的后面添加n个0。
3. 得到的（k+n）位数除以事先商量好的（n+1）位的除数P，得出的商是Q，余数是R，为n位。其中作除法时，并不是真正的除法，而是异或。

4. 将余数R作为循环冗余码拼接在数据M后发送。

- 当**接收端**收到帧时，对M+冗余码除以除数，若余数为0证明这个帧没有差错，而余数不等于0时这个帧被丢弃。

3. 点对点协议PPP

- **PPP**：是数据链路层使用的一种点对点通信方式，也是目前数据链路层使用得最广泛得协议。PPP协议是广域网上的数据链路层协议。
- **协议帧格式**



由于数据链路层是较低层，其上一层为网络层，网络层的PDU为packet，使用地址为IP地址。

- PPP协议包括了链路控制协议LCP，NCP网络控制协议和PPP的扩展协议。LCP包括了建立整个链路的几次握手，而NCP是用来协商PPP报文的网络层参数（IP地址）。

4. 使用广播信道的数据链路层

- **局域网和广域网的区别**：局域网是封闭的，通常是一个小型单位所有的有限站点。通常一个局域网就是一个广播域，也存在由多个路由器分割成许多个小广播域的局域网。只有局域网内可以广播，广域网上不能广播。而广域网通常是地区范围的，例如省市和国家。
- **拓扑结构**：集线器：星形；匹配电阻：总线网；干线耦合器：环形网。
- **以太网**：是一种计算机局域网技术，它规定了包括物理层的连线，电子信号，介质访问层协议的内容。以太网有两大类，一种是经典以太网，第二种是交换式以太网，交换式以太网采用了交换机。接入以太网的机器以MAC地址来互相鉴别。
- **CSMA/CD协议[适用于总线型结构]**：其英文含义是载波监听多点接入/碰撞检测。“多点接入”是指计算机接入在一根总线上。“载波监听”是指每个站在发送数据前需要监听总线上是否有其他计算机在发送数据，如果有则赞数不发送数据。“碰撞检测”是指计算机发送数据时会检测信道上的信号电压大小。其过程如下：

  （1）检测信道：在准备发送数据前检测信道，一直检测直到96比特时间内信道空闲，则发送这个帧。

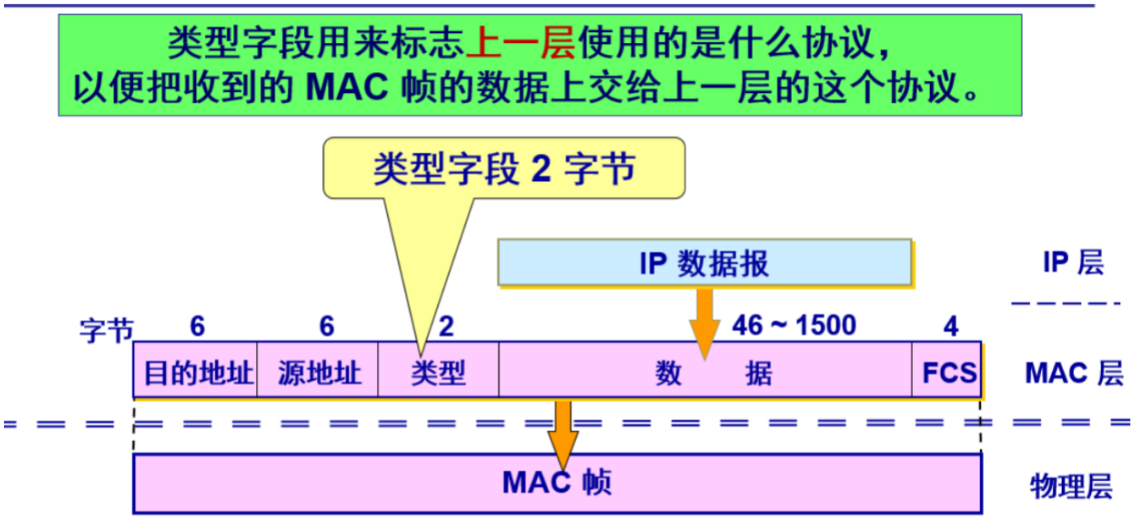
  （2）检查碰撞：在发送时也不停检测监听信道，若未碰撞则发送成功；若发生碰撞，则停止发送数据，并发送干扰信号，执行退避算法，等待n倍512比特时间后重传，若重传16次都不成功，则停止，并向上报错。

- **基于双绞线和集线器的星形拓扑**：速度更快，并且更方便，淘汰了总线型结构，但是在逻辑上以太网还是总线型的结构，因此CSMA/CD协议还是在起作用的。集线器其实在模拟总线，它像一个多接口转发器一样工作。
- **以太网信道被占用**：被占用时通常采用两倍于其单程端到端的传播时延作为争用期，隔一个争用期之后再发送，则以太网参数$\alpha = t/T$描述的就是发送效率，需要减小争用期，不能使得以太网连线太长，也不能让帧太短。

5. MAC地址

- **MAC地址**：又被称为物理地址，硬件地址，是局域网通信设备或端口的唯一标识符。其格式如下：

- **单/组/广播地址**：I/G（第一字节最低位）为0时表示单站地址，为1时表示组播地址，全为1时为广播地址。
- **以太网v2的MAC帧格式**：



在物理层传输时还需要再MAC帧首加上8字节用于比特同步，没有结束符。但PPP帧有结束符是因为，广播信道的局域网存在发送停止，不能连续发送，因此信道上无电频即为停止发送。

- **无效MAC帧**：帧的长度不是整数字节，FCS错误，数据字段长度错误，以太网会丢弃这个帧，并且不负责重传。

6. 集线器和交换机的区别

二者都是用来扩展以太网，即接入更多设备的。

	集线器	交换机
工作位置	物理层	数据链路层
工作方式	在物理上用端口将机器连接起来，每个数据帧都在总线上广播。因此一个集线器其实划分了一个冲突域。	交换机收到MAC帧时会根据其MAC地址向目的端口进行转发，并不是所有计算机都会收到帧。但交换机也能实现广播，即存在并行性。
优点	1. 使得原来不同碰撞域的以太网上的计算机能够跨碰撞域通信；2. 扩大了以太网的地理范围。	1. 转发速率快；2. 并行转发，使得以太网吞吐量提升；3. 能在繁忙时缓存帧；即插即用，其内部的帧交换表是自学习算法自动建立的。4. 与总线型以太网兼容性好；5. 能兼容不同速率的接口。
缺点	1. 碰撞域增大后总的吞吐量未提高；2. 并且集线器不能用来连接不同数据率的碰撞器。	

- 交换机的帧交换表自学习算法：最初是空的，当一个主机接入某个端口后，会发送数据帧，帧一定会从这个端口到交换机，此时交换机发现自己的表中没有这个MAC地址和端口的对应，就将其写入，再检查目的地址，目的地址也没有记录，就进行广播，等到目的主机回复后，交换机会将其写入交换表。其条目存在有效时间，因为端口连接可能变化。

7. 单播，多播，任播和广播

- **任波**：一对多传输数据，但接收方只有一台机器可以接收消息，然后返回一个单播信号给源主机。

8. VLAN

虚拟局域网从逻辑上划分了多个局域网，能够限制广播信息的数量，不会因传播过多而导致广播风暴。

4. 网络层

1. 网络层的基本功能和最小单位

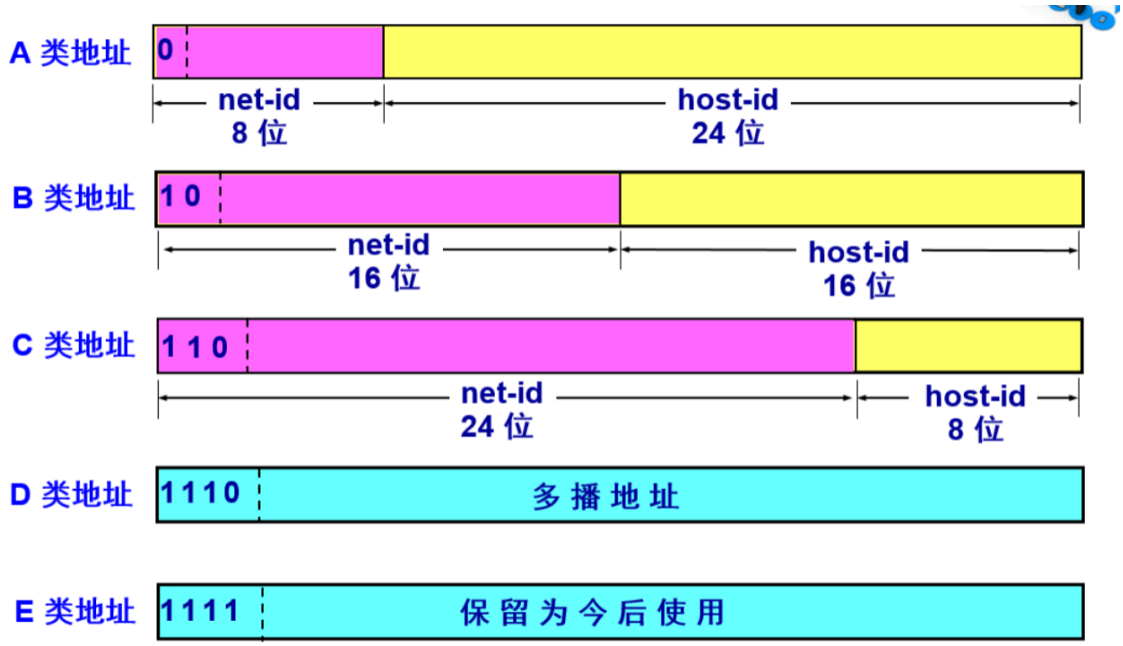
- **网络层的基本功能**：数据链路层只负责一跳的传输，仅限于一个以太网内，而网络层需要负责远端传输。网络层提供简单，无连接的，尽最大努力交付的数据报服务，而可靠传输则交给传输层。其不建立连接，每个数据报独立发送，而且不提供传输质量保障。
- **最小单位**：IP数据报

2. 网络层包含的协议

其包括了IP协议，地址解析ARP协议，网际控制报文协议ICMP和网际组管理协议IGMP。其中ICMP和IGMP更靠近其上层的传输层，而ARP协议更靠近数据链路层。

3. IP地址和作用

- **IP地址**：全世界范围内唯一的32位标识符。其可分为网络号和主机号两个部分。
- **IP地址的分类**：



其中A类地址网络有$2^7-2=126$个，其中127不用，全1的广播地址不用。每个A类地址可以容纳的主机数量为$2^{24}-2$，除了特殊的全0或者全1的地址。

IP 地址的指派范围

网络类别	最大可指派的网络数	第一个可指派的网络号	最后一个可指派的网络号	每个网络中最大主机数
A	126 (2^7-2)	1	126	16777214
B	16383 ($2^{14}-1$)	128.1	191.255	65534
C	2097151 ($2^{21}-1$)	192.0.1	223.255.255	254

一般不使用的特殊地址：

（1）全0地址：网络中的本主机；

（2）网络号为0+主机号为Host-id：在本网络上某台主机的Host-id；

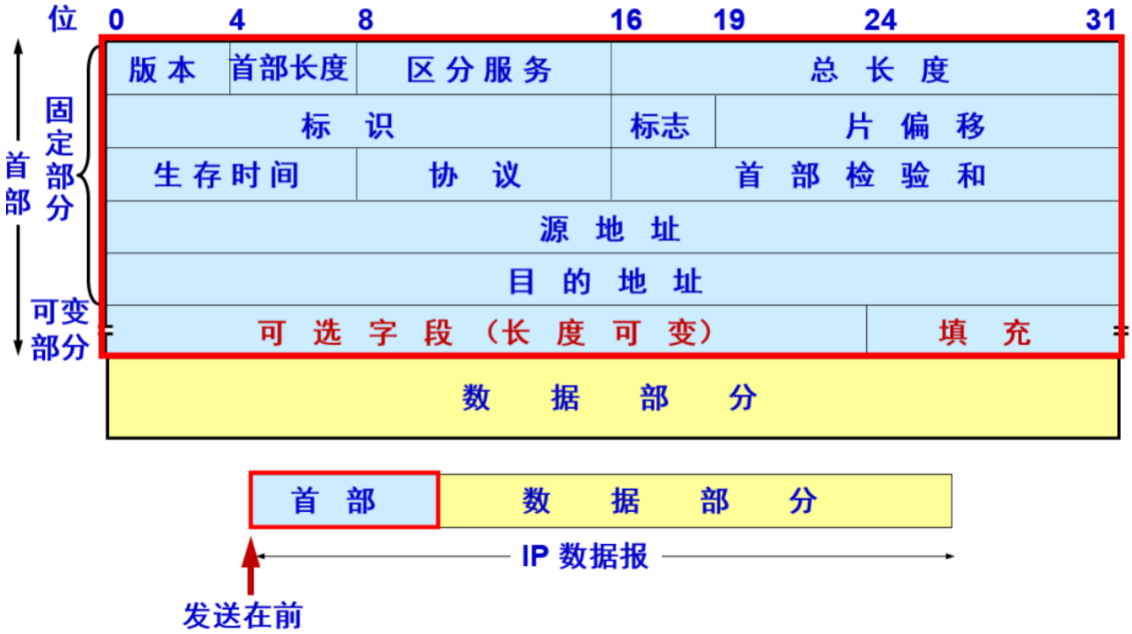
（3）全1地址：在本网络上进行广播；

（4）net-id+全1：在本网络上的所有主机进行广播；

（5）127+非全0或全1的任何数：测试用

- **IP地址的写法**：十进制点分法，使用4个十进制数提高可读性
- **网络中IP地址**：在同一个局域网中的主机的网络号是相同的，一个路由器具有多个IP地址，两个路由器连接的端口可以指定IP地址也可以不指定，若不指定则两个路由器为一个特殊的没有主机的局域网。
- **IP数据报的格式**：

  由首部和数据两部分组成，首部的固定部分有20字节，所有IP数据报都必须得有，后面还跟了长度可变的可选字段。



版本：指使用的IP协议版本，IPV4或者IPV6

生存时间：指该数据报最多经过的路由器个数

协议：指的是该条数据报上层的协议，如TCP，UDP等等。

4. ARP地址解析协议

- **ARP的作用**：通信时每当数据经过一个路由器都会被解析一次IP头，读取下一跳的MAC地址，由于数据链路层的协议只负责将数据传递至相邻节点，因此MAC地址在传播过程中会不断改变。而ARP协议的作用就是根据IP地址解析出其MAC地址的。
- **ARP高速缓存**：每个主机内部都有一个ARP高速缓存，里面有IP地址到MAC地址的映射表<IP address; MAC address; TTL>，其存储的是局域网内的主机和路由器的地址。
- **ARP协议内容**：每次主机想向另外一个主机发送数据报时，时首先查看自己的ARP缓存表，如果有可以直接将目的MAC地址写入MAC帧，再发送；如果没有，就广播一个ARP请求协议（如果需要的主机不在局域网内，路由器会发送一个响应，帮助转发至其他局域网），收到响应后主机会将收到的IP和MAC地址写入ARP缓存表内。

5. 路由器和路由表

- **路由器的作用**：连接不同的网络，并且选择数据传播路径。
- **路由表**：存储在路由器内部的，按照目的主机所在网络号来决定下一跳地址的转发表。注意**下一跳地址**是一为了将这个数据交付，需要找到的下一个网段地址。
- **直接交付和间接交付**：如果该路由器某个端口直接连接了某个网络，则可以将数据直接交付给某个端口，这是直接交付；而间接交付是，当该路由器没有直接相连目的网络时，会再次查找其路由表，看有没有到这个网络的网络，再根据这个网络来查找，看有没有对应能转发的端口，若还是没有，就一直迭代，直到找到一个能转发的端口，或者将其转发给一个缺省端口。

  [注意：一个路由表里存在两种类型的对应<网络地址：端口>和<网络地址：网络地址>]
- **划分子网的情况下的转发**：当各个网络存在子网掩码时路由器不能直接判断转发的IP地址，因此路由器会先将收到的IP地址（1）和其相连的各个网络的子网掩码分别逐位相与得到<网络号：子网号>，再检查这个网络号是否和相应的网络匹配，若匹配，则为直接交付；（2）若不匹配，则为间接交付，如果路由表中直接存有该全部位数的IP地址的路由，则按照路由表转发；（3）若不存在（2）中的路由，则对路由表中的每一行，将收到的数据报IP地址与子网掩码逐位相与，看是否存在匹配的网络，若有匹配网络，则迭代查找；（5）若不存在（4）中的匹配网络，则将其转发至默认路由（通常为网关）；（6）若不可以，则报告转发分组出错。

6. 划分子网

- **作用**：IP地址自身的分级太粗略了，因此局域网内部可以自己借用主机号来生成子网号。对外部显示的网络号还是跟原来一样，但是内部存在自己的子网号。这样做可以减少IP地址的浪费，从而使得网络的组织更灵活，并且更利用管理。
- **子网掩码**：用于解析时判断一个IP地址主机所在网络是否进行了子网划分。子网掩码的长度为32位，1覆盖的部分即为网络号部分。

7. ICMP网际控制协议

- **作用**：其用于在IP网络设别间传递各种差错和控制信息，并对于手机各种网络信息、诊断和排除各种网络故障等方面起到重要作用。
- **类型**：重定向（选择更优路径），差错检验（诊断网络联通性），错误报告。其报文的类型有两种，一种是差错报告报文（终点不可达，时间超过，参数问题，重定向），一种是询问报文（回送请求和回答报文，时间戳请求和回答报文）。
- **应用**：PING和Tracert都是ICMP协议的具体应用。其中PING使用的是回送请求和回答报文，它是应用层直接使用网络层ICMP协议，而没有通过运输层。Tracert是通过TTL和时间超过报告来实现对源主机和目的主机之间路由路径的追踪。

8. 路由协议

前面提到的路由表，只说明了其工作原理，并没有说明路由表是如何生成路由路径的。其路由算法必须是正确和简单的，并且能适应通信量和网络拓扑的变化，同时满足稳定和最佳性。但最佳路由并不是绝对意义的最佳，因为网络状态复杂变化，只能说尽量使其接近于理想算法。

- **协议种类**：

	内部网关协议IGP（域内路由选择）	**外部网关协议EGP （域间路由选择）**
定义	再自洽系统内部使用的路由协议，是使用最多的协议。	若目的主机与源主机处于不同的自洽系统内（指其内部网关协议选择不同），则消息传递到边界时需要更换协议，外部网关协议就是负责将路由信息传递到另外一个系统中的协议。
实例	OSPF, RIP	BGP-4

- **具体协议**：

	RIP	OSPF	BGP
范围	域内路由	域内路由	域间路由
算法	其协议要求网络内每个路由器都维护一个从它自己到其他每个目的网络的距离记录，其距离定义为经过的路由器的跳数，跳数越短的路径它认为其越好。当跳数多于16时，其认为这个目的地不可达。路由器之间会发送路由更新信息，根据收到的更新信息来更新自己的路由表。	其为RIP的改进版本。当链路状态发生变化时，一个路由器洪泛式向该网络中所有路由器发送信息，发送的信息是与自己相邻的路由的链路状态（指自己与哪些路由器相邻，以及该链路的度量）。 为了使得算法适用于更大的网络，OSPF会将自洽区域再划分为更小的网络，每个网络有32位标识符，并且分为了主干区域和其他区域。主干区域能连接自身区域内的其他区域，也能连接其他自洽区域。	其协议只争取寻求一条能够到达的目的网络的路径，而非最佳路径。 其需要每个自治系统都选出一个路由器作为BGP发言人，两个发言人路由器通过一个共享网络连接在一起，这两个路由器之间建立可靠的TCP连接。
适用场景	适用于小型网络	适用于大型网络	适用于两个自洽系统之间
缺点	好消息传播快，坏消息传播慢。即一个网络出现故障时，其消息传递的时间非常长。		
优点（特点）		1. 根据不同的服务类型可计算不同的路由 2. 若存在多条相同代价的路径，则实现多路径负载均衡，尽可能缩短排队时间 3. 支持可变长度子网划分 4. 30min刷新链路状态，并且一个路由器的链路状态只涉及自己到相邻路由器的联通状态，也没有坏消息传的慢的缺点。	1. 交换路由信息的节点相比域内路由少； 2. BGP刚运行时会发送整个BGP路由表，而之后只在变化时更新有变化的部分。

9. IPv6

- **相比IPv4的变化**：

  （1）地址位数从32位增长到128位，其固定首部长为40字节

  （2）定义了更多可选的扩展首部：放在数据部分的前面

  （3）扩展的地址层次结构

  （4）改进的选项：IPv6郧西数据报包含有选项的控制信息

（5）允许协议继续扩充；

（6）支持即插即用，自动配置，因而不需要DHCP协议来分配IP地址；

（7）支持资源的预分配

IPv6的首部取消了IPv4的扩展首部，而将其放入了有效载荷中，因此路由器收到数据包后不再处理扩展首部，因此提高了效率。

- **IPv6地址写法**：冒号16进制写法，每16位用16进制值表示（因此有8个16进制数），当某个分组内全0时，可以用一个0来表示，当有一连串0时，用“：：”来替代，但是一个地址只能用一次零压缩。IPv4地址可以通过在地址前补0的方式变为IPv6,用于过渡。
- **IPv4向IPv6过渡**：

  IPv6需要能够向后兼容IPv4，能转发IPv4的分组。

  （1）采用双栈协议：将IPv6的流标号删除，生成IPv4协议格式的报文，再转换回IPv6时流标号则置空。

  （2）隧道技术：在IPv6格式报文中提取必要信息，在外部包裹一层新的IPv4格式的报文，再进行传输，这样就不会损失原有信息。

10. IP多播和IGMP协议

- **好处**：可以节省网络资源，若是一个主机需要向某个网络内的多个主机发送相同的消息，则单播需要发送n次，而采用多播形式，则可以将利用路由器的硬件多播复制功能，将消息发出，源主机只需要发送一次。
- **多播地址**：D类地址，一个D类地址可以标志一个多播分组。并且多播地址只能用于目的地址，不能用于源地址。
- **多播数据报**：使用D类多播IP地址，以及使用IGMP网际组管理协议，其发送后不保证交付，并且PING多播地址也不会有回应，这是因为其不产生ICMP差错报文。
- **IGMP**：网际组管理协议IGMP，维护一个D类地址的成员名单，每个局域网中只有一个多播路由器直到这个名单，这个多播路由器还需要将数据报用最小代价传送给各个主机，因此还需要额外的多播路由选择协议（类比RIP,OSPF等）。IGMP协议并不知道IP多播组包含的成员和成员数量，它只负责让多播路由器直到是否有主机参加或退出了某个多播组。
- **IGMP协议内容**（1）某个主机加入新多播组时会向多播地址发送IGMP报文进行声明；（2）多播路由器接收到报文后会将多播组发送给互联网上的其他多播路由器；（3）多播路由器会周期性询问，确保多播组还活跃，只要本地网络上有一个主机相应，则这个多播路由器就认为这个多播组活跃，若几轮探寻后没有主机响应，则不再转发这个组成员的关系给其他多播路由器。
- **多播路由选择**：多播组的成员是动态变化的，随时都有主机再加入或者离开这个多播组，因此多播路由实际上是再寻找以源主机为根的多播转发树，寻找这棵树的方式有三种：

	洪泛与剪除	隧道技术	基于核心的发现技术
特征	1. 最初多播路由器广播多播数据报； 2. 采用反向路径广播，即一个路由器接收到多播数据报时检查其是否从源点经最短路径传播（即确认自身是否在多播转发树上），若是就向除了传入方向的所有其他方向转发刚才的多播数据报，否则就丢弃不转发。 3. 多播路由若发现其下游树枝没有多播组成员，即剪枝。	若两个路由器之间的网络不支持多播，则需要将整个多播数据报进行封装，封装为单播数据报在两个路由器之间进行传输。	对每个多播组指定一个核心路由器，并且给出多播组核心路由器的IP地址，由核心路由器创建出对应多播组的转发树。
适用场景	适用于规模较小的多播组，而且多播组成员所在局域网相邻	连适用于多播组在地理位置上很分散的情况。	适用于多播组的大小变化很大的情况。

11. VPN和NAT

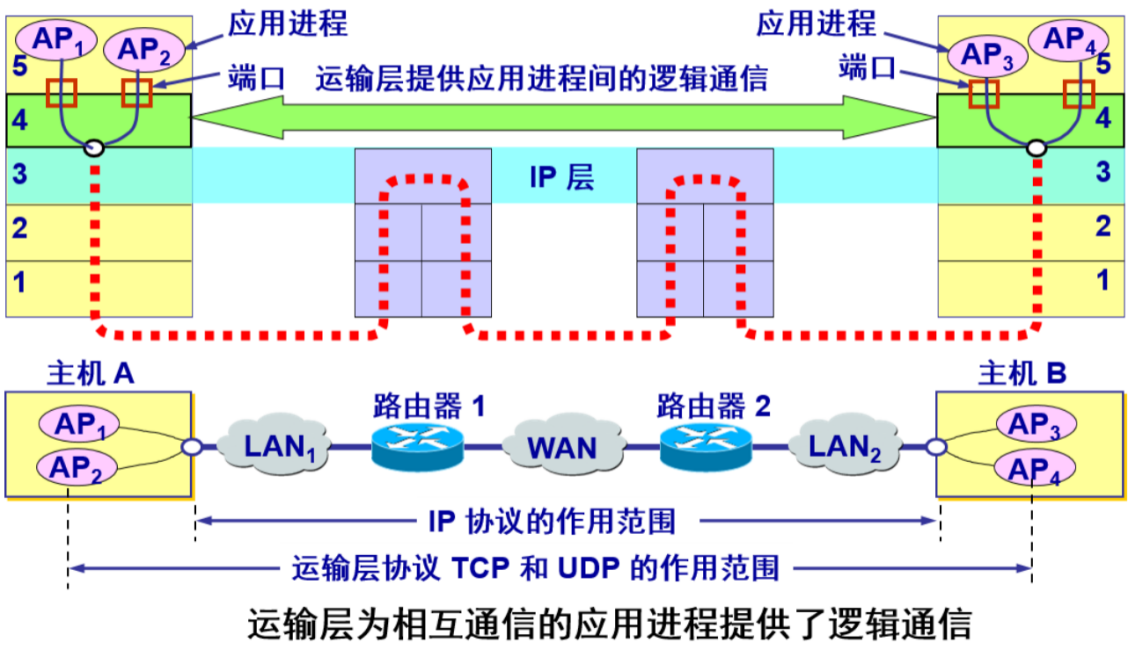
由于IP地址紧缺，有时一个机构不能申请到足够数量的IP地址，而一个机构内部也不需要把所有主机都接入到外部的互联网，因此部分只在机构内部使用的主机可以被分配一些专属机构内部的IP地址。

- **VPN**：即虚拟专用网，是利用互联网作为机构内主机通信载体的网络。专用网只用于机构内部通信，而不用于和外部通信。
- **NAT**：当专用网上使用专用地址的主机需要和互联网上的主机通信时需要采用网络地址转换NAT。NAT转换任务是由路由器上的NAT软件实现的，该路由器至少有一个有效的外部全球IP地址，VPN上的主机和外界通信时需要在NAT路由器上将其本地地址转换成全球IP地址才能和互联网连接。

5. 传输层

1. 传输层基本功能和最小单位

- **基本功能**：传输层提供了主机应用程序进程之间的端到端服务。只有位于网络边缘部分的主机的协议栈才有运输层，而网络核心部分的路由器在分组转发时只用到了物理层，数据链路层和网络层功能。（网络边缘：指的是主机和网络应用；网络核心：路由器和交换机）



运输层为相互通信的应用进程提供了逻辑通信

- **和网络层的主要不同**：网络层IP协议主要负责将连接两台主机，而一台主机中通常有多个应用进程，这些进程的运行通常需要和另外主机的多个进程建立多个连接，这就需要传输层的协议。
- **最小单位**：传输层PDU为数据段segment。TCP协议传送的数据单位协议是TCP报文段，UDP传送的数据单位协议是UDP报文。

2. 传输层端口，复用和分用

- **端口的作用**：对于主机上运行的进程来说，它是依靠进程标识符来区分不同进程的，但是不同主机的操作系统给进程指派的标识符很不一样，如果直接使用操作系统指派的标识符会导致不兼容的问题。为此需要在传输层使用同一的协议端口号，这个端口号只在本机范围内有意义，用于标志计算机应用层中的各个进程。两个主机中的进程若需要通信，必须直到对方主机的IP地址和端口号才能通信。
- **端口的形式**：端口号为16位，通常服务器使用的端口号为0~1023，为熟知端口号；1024~49151是登记端口号，为没有分配到熟知端口号的程序使用，使用这个端口号则必须登记，防止重复；客户端使用的是短暂端口号49151~65535。
- **复用和分用**：同一台主机上有多个进程在同时运行，因此理论上需要并行建立多个传输层连接，但是在传输层以下的网络连接是串行的，因此需要从并行变为串行，即复用。而经过下层传输后，又需要在目的主机侧变为并行，这就是分用。

3. TCP和UDP的区别

	TCP	UDP
中文	传输控制协议	用户数据报协议
连接	面向连接的协议	面向无连接的协议
可靠	可靠的连接	不可靠的连接
首部长度	20字节	8字节
拥塞控制	有	无
面向报文的协议		是
连接对象个数	一对一	一对一，一对多，多对一，多对多

4. UDP协议

- **UDP协议内容**：在IP数据报的基础上添加了复用和分用的功能，以及差错检验功能。
- **UDP的特点**：

  （1）无连接：发送数据前不需要建立连接，因此减少了开销和时延‘

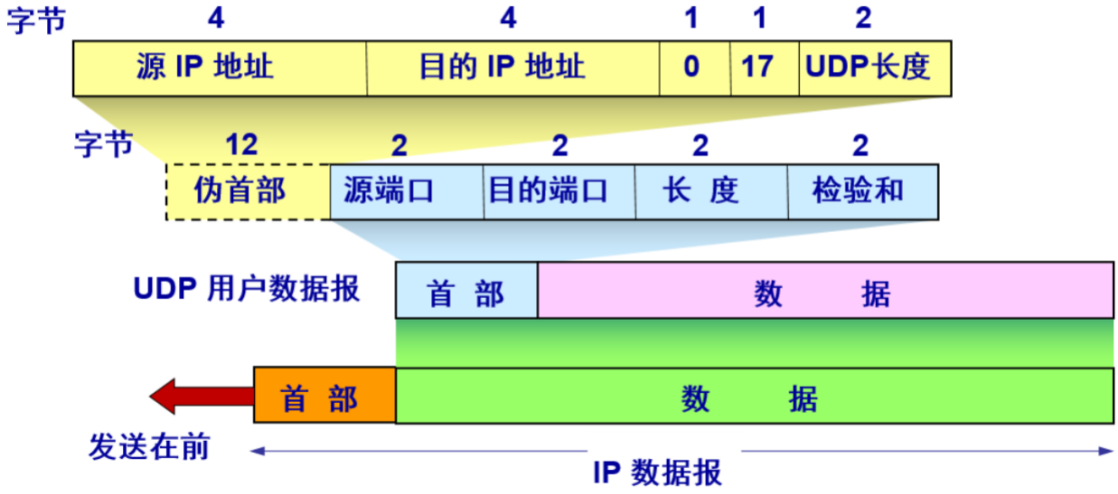
  （2）最大努力交付：和下层一样，不保证可靠交付，因此主机不维护连接状态表；

  （3）面向报文的：UDP对应用层交来的报文不进行拆分，而是保留报文边界，UDP一次交付一整个完整的报文。因此应用程序在将需要发送的数据传递给UDP协议之前需要考虑报文大小，若报文太大，则IP协议会将其分片，降低IP层效率；若报文太小，则通信过程不划算，也会降低IP层效率。

  （4）没有拥塞控制：网络拥塞不会使得源主机的发送速率降低，因此适合实时通信的要求。

  （5）支持一对一，一对多，多对一和多对多的交互通信。

  （6）首部只有8字节，少于TCP连接的20字节，因此首部处理开销小。
- **UDP报文**：



UDP用户数据报的首部和伪首部

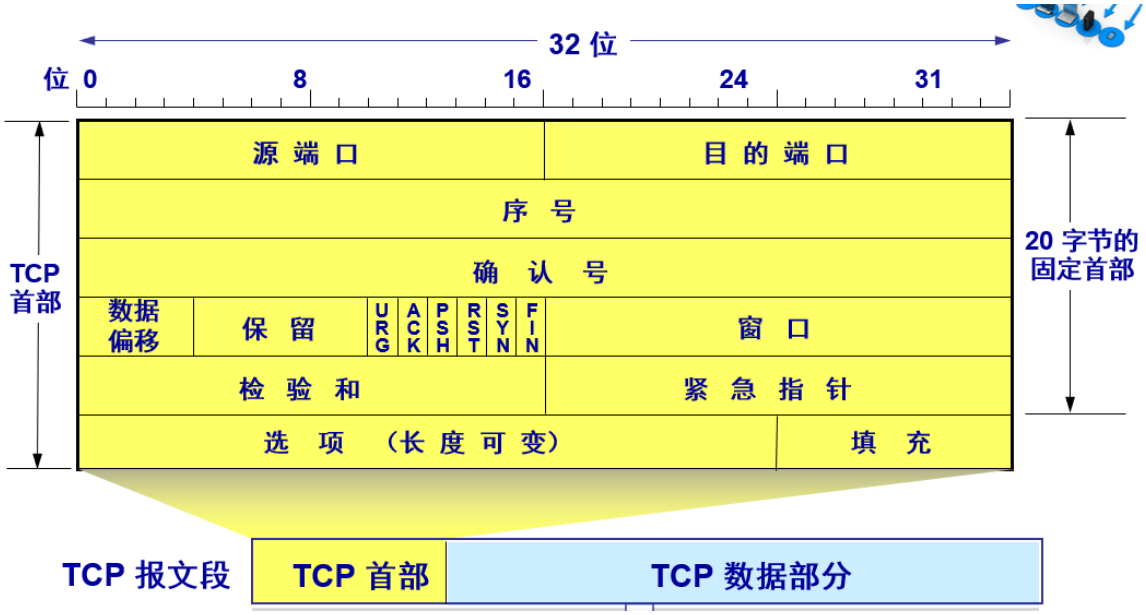
其首部只有8个字节，包括源端口，目的端口，长度和检验和。但是它还有个伪首部，伪首部有12个字节，伪首部提取了IP数据报中的源IP和目的IP信息以及协议字段，但不会实际发送出去，是用于计算校验用的（差错检验），应该是存在一些网络层无法检验的错误。

5. TCP协议

1. TCP的基础内容

- **TCP报文格式**

其报文是由首部和数据段拼接形成的，其首部有20个字节的固定部分，后面有4n个字节根据需要来增加。



- 序号seq：占4字节，指的是数据部分发送的第一个字节是整段数据的第几个字节。
- 确认号ack：是期望对方收下的**下一个**报文的段的数据的第一个字节的序号。
- 窗口：2字节，用于让对方设置发送窗口大小的依据。
- 检验和：检验范围为首部和数据，和UDP一样，也需要加上12字节伪首部。
- **TCP协议内容**：面向连接的协议，即在发送数据前确认发送端和接收端之间的连接。
- **TCP协议的特点**：

  （1）有连接：在发送数据前确认发送端和接收端之间的连接。

  （2）可靠交付；

  （3）一对一通信：每个TCP连接只能是端对端的；

  （4）提供全双工信道：即通信时收数据和发数据可以同时进行；

  （5）面向字节流：流是指流入或者流出进程的字节序列，TCP协议将应用程序交来的数据块当作无结构的字节流。并且TCP需要确认接收方收到的字节流和发送方发出的字节流完全一致。TCP将需要发送的数据写入发送缓存中，在首部添上TCP首部，并且每段数据都标上序号；接收方侧TCP报文段排队接收字节，并且也有TCP接收缓存。也就是说TCP会对原始报文进行长度变化，过长切割以及过短补齐。

  （6）报文长度决定：TCP发送的报文长度是自适应的，它会根据对方的窗口值以及当前网络的拥塞程度来决定一个报文该多长。

  （7）TCP连接的端点：不是传输层的协议端口，而是套接字（socket）或插口，但套接字本质上是<IP：端口号>拼接而成的。每个TCP连接唯一地被两个套接字确定。

2. TCP可靠传输的原理

（1）**停止等待协议**：发送方每发送完一个分组就停下来，等待对方确认收到后再开始发送下一个分组，因此其实是时间换可靠性的做法。

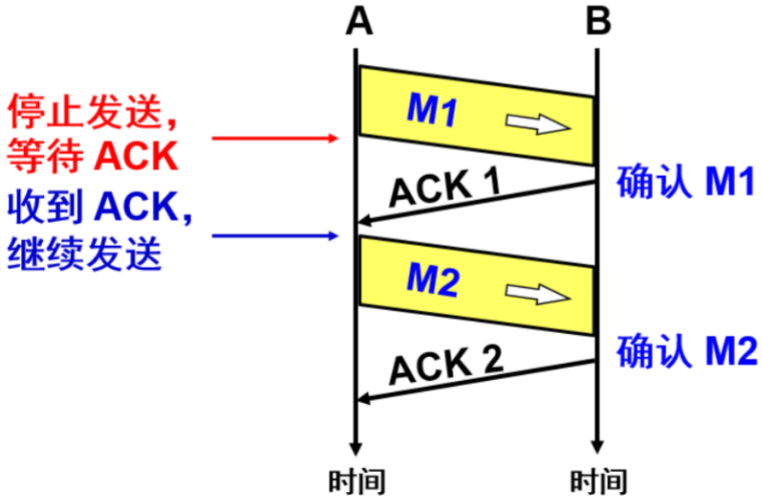
（2）**自动重传请求ARQ**：即A总是能收到全部发出的分组确认，当接收方发现错误或者没有收到数据时，不需要主动向发送方发出请求。

以上两个协议控制的过程如下：

- 当接收方B确认接收到A发来的数据，并且验证过数据完好时，B会向A发送一个ACK报文确认，A收到报文后再继续发送。

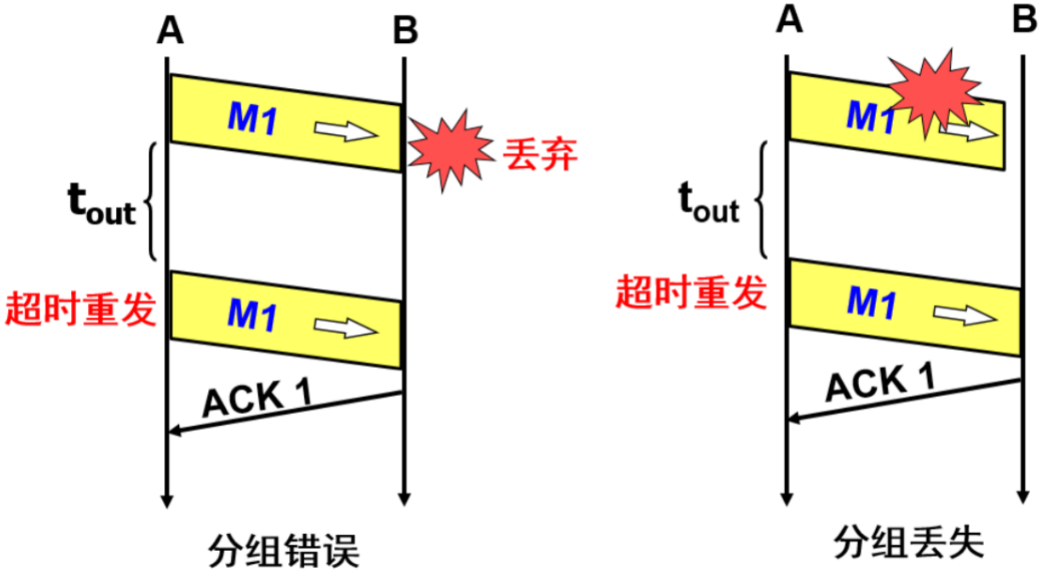
1. 无差错情况

A 发送分组 M1，发完就暂停发送，等待 B 的确认 (ACK)。B 收到了 M1 向 A 发送 ACK。A 在收到了对 M1 的确认后，就再发送下一个分组 M2。



- 当接收方B没有收到报文，或者收到后验证发现报文有错误，直接丢弃报文，它都不向A发回ACK，即什么都不作。那么A在发送时设置了计时器，只要在一定时间内没收到ACK报文，A就会重新向B发送上次没收到ACK的报文。

2. 出现差错



- 对于出现差错的情况，还有其他的可能性，例如是B的回复报文丢失，或者确认报文迟到了，A也不会收到确认报文。此时A依旧会重复发送报文给B。
- 对于以上情况的结合，A和B分别有可能收到重复的确认报文和数据，A收到重复的确认后丢弃即可，而B收到重复的数据会丢弃数据，然后再次发送一次确认报文给A。
- **TCP的信道利用率**：

  信道利用率的概念是：A发出发送数据所用时间比上发送时间，等待ACK时间和确认ACK时间 $U = \frac{T_D}{T_D+RTT+T_A}$。

而TCP使用停止等待协议，因此信道利用率很低。

- [改进策略]**停止等待协议->流水线传输**：

  为了提高信道利用率，可以采用流水线传输，即以此按照顺序发出不断发送数据，但是流水线长度需要控制，因为一直发送不停就没有办法处理出错的数据了，流水线长度一般是RTT时间长度。

  - **连续ARQ协议**：发送方会维护一个发送窗口，其意义是位于窗口长度内的数据分组可以连续发送出去，不等待ACK，而每收到一个分组的ACK就将窗口向前滑动一个分组的位置。
  - **累积确认**：接收方采取对顺序到达的最后一个而分组发送确认，表示在这个分组和它之前的所有分组都收到了。缺点是如果说收到分组出错，那么重传起来非常困难。比如发送方一次发了5个分组，而接收方没有收到第三个分组，因此它会发回前两个分组的ACK。
  - **GO-BACK-N**：对于上面提到的需要重传的情况，发送方需要重传3~5共3个分组。因此当通信线路质量不好时连续ARQ会带来负面影响。
- **TCP超时重传的时间选择**：RTT总的时间分布方差非常大，因此很难通过RTT来选择自身的超时重传阈值。若设置时间过短，则重传分组数量会过多，造成网络负荷增大；而设置时间过长，则会导致网络空闲时间过长。因此TCP采用了自适应的算法：

  它记录分组发出和收到分组的RTT时间，并且不断算加权平均的RTT_s：

  $RTT_s = (1-\alpha)RTT_{s-1} + \alpha RTT_{new}$

  当α接近0时，代表加权RTT更新的很慢，而α接近1时代表更新地快。一般取0.125.

  - **不采样策略**：但得到新的RTT样本就是很难得事情，因为当一个分组没收到确认，发送方又重传之后收到了确认，不能确定这个确认是对哪个时刻发出得分组的确认，因此只要某个分组重传，就不采样。
  - **Karn算法**：但不采样会导致，如果网络状况突然很差，那么大部分RTT都不采样，超时重传就会一直不更新。因此在原来的加权RTT算法中再进行修正：每发生一次重传，就要略微增大加权RTT时间。

    $RTT_s = \gamma(RTT_{s-1})$

    γ的值一般是2.
- [改进策略]**选择确认SACK**：针对之前的**GO-BACK-N**的问题，选择确认SACK可以让发送方只发送接收方想要的分组。
  - 使用选择确认，需要再TCP首部选项中加入允许SACK选项，并且由于首部最多40字节长，而指明一个边界就要用掉4字节（16+16，一个开头和结尾），因此最多指明4个字节块的边界信息。
  - 接收方接收到了一些不连续的分组块，如果确认这些分组块都在接收窗口内，则接收方会先接收这个分组。然后在SACK中指明自己**收到的**分组的边界。发送方就能通过这些边界来确认哪些包没有被成功接收。
  - SACK只针对于乱序传递的分组，而且对于片段特别分散的数据块就不合适。

3. TCP的流量控制

滑动窗口

需要流量控制的原因是需要让发送方的速度使得接收方来得及接收，以及不使得网络发生拥塞。流量控制是通过**滑动窗口**实现的。

- 接收方通过向发送方发送ACK报文来告知发送方自己目前允许接收的数据量以及自己收到的数据的边界$ack = n+1\ rwmd = window\ size$
- **互相死锁**：但若接收方给发送方发送了0大小的窗口，等它处理完缓存后，再告诉发送方自己的可以接收信分组的消息丢失了，则发送方等待这个消息，接收方等待发送方发送数据。TCP设置了**持续计时器**，只要发送方收到0窗口通知就启动计时器，隔一段时间发送一个0窗口探测报文进行询问。若对方窗口仍然是0，则重置计时器。

发送时机选择

对于发送方来说，何时发送数据也可以分不同时机：

- 缓存中数据量达到MSS字节时发送；
- 由发送方进程指明发送时发送；->这会导致很多问题
- 计时，定期发送

发送方糊涂窗口综合征：若TCP每次接到一字节的数据就发出，则效率很低，因此使用**Nagle算法**，若要发送的数据是逐字节地进入的，先发送第一个字节的数据给接收方；继续缓存后续的字节，等到收到ACK后，再发送缓存好的数据；若中途缓存满了，就直接发送。

接收方糊涂窗口综合征：若进程每次都只读取一个字节，那么接收方ACK告诉发送方$rwnd = 1$，发送方每次只发送一个字节，等接收完这个字节后，缓存又满了。因此解决办法是接收方等待一段时间，使得缓存的内容被消化一部分，有能力接收一个最长报文段后再发送ACK通知发送方。

4. TCP的拥塞控制

拥塞的概念

某段时间内，对网络中某种资源的需求超过了资源所能提供的可用部分，网络的性能会变化。同时**增加资源有时并不能解决问题**，反而可能还会加剧问题，有时是路由器缓存空间不足，重传分组会导致拥塞家具；增大路由器缓存，而链路处理速度不提升，排队时间会大大提高，也不能解决问题。

拥塞控制与流量控制的区别

	拥塞控制	**流量控制**
目的	拥塞控制的目的是防止过多数据注入网络中，从而使得路由器或者链路不至于过载。	流量控制的作用是使得发送方的发送的数据不会使得接收方来不及接收。这是个点对点通信量的控制
对象	全局控制，需要考虑主机，路由器，以及降低网络传输性能的所有因素。	局部控制，发送主机端和接收主机端
手段		1. 滑动窗口机制 2. Nagle算法控制发送时机

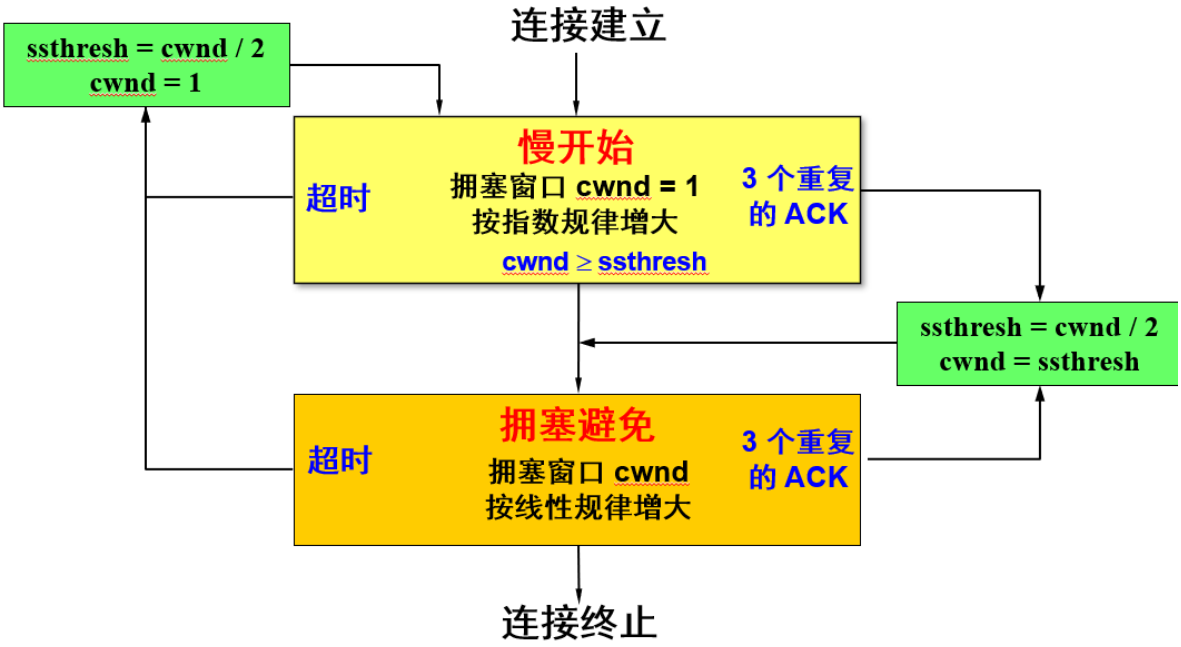
拥塞控制的手段和监测指标

- **开环控制**：在设计网络时就提前考虑好拥塞因素。
- **闭环控制**：基于反馈，实时监测网络的拥塞情况，并采取调整措施。
- **检测指标**：缺少缓存而被丢弃的包的数量；排队长度；超时重传百分数；平均时延等等

TCP拥塞控制的方法

- **基于窗口的方法**进行拥塞控制，发送方会维护一个拥塞窗口$cwnd$，这个窗口的大小取决于网络拥塞程度并且动态地变化。因此发送方真正发送窗口值既受流量控制又受拥塞控制：$Real\ Window = min(cwnd, rwnd)$.
- **判断拥塞的方法**：1. 重传定时器超时；2. 收到三个相同的ACK。

TCP拥塞控制算法——四部曲

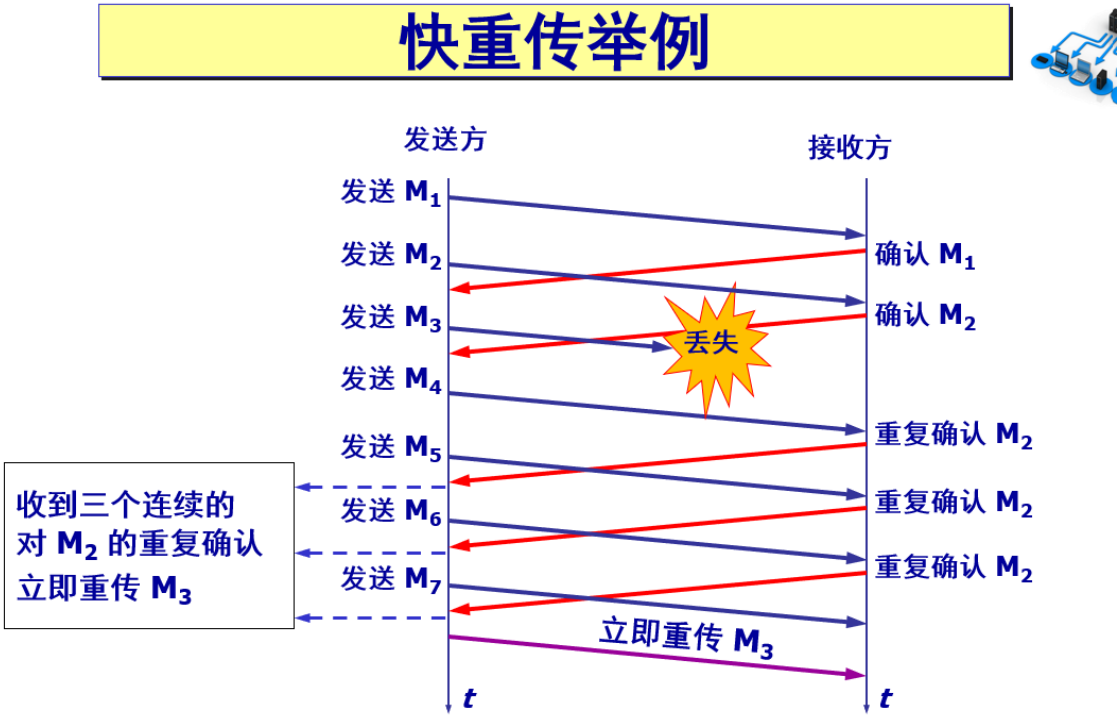


- **慢开始**：发送方设置初始拥塞窗口大小为不超过其最大报文段长度2~4个的数值进行发送，它每收到**全部发出的新报文的确认**就会加倍拥塞窗口大小，即每经过一个传输轮次就会加倍。（确认报文不包括重传的）这个指数增长存**门限值**，初始门限值为16个报文段，即ssthreash=16。等增长到门限值后开始进入拥塞避免阶段。当检测到拥塞后，重新开始慢开始，此时门限值调整为 $ssthreash = curr\ cwnd/2$。
- **拥塞避免**：在拥塞避免阶段，每经过一个RTT时间，就将拥塞窗口加1，进行线性增长。

注意：不论是慢开始阶段还是拥塞避免阶段，只要发送方确定传定时器超时，就从0开始进行慢开始阶段，将拥塞窗口置为1。

只要发送方确定受到了三个重复的ACK报文，就会开始执行快重传和快恢复算法。

- **快重传**：这个阶段是为了解决对方没有收到特定报文，因此需要立即进行重传，这样不会出现超时，发送方也不会误认为出现了网络拥塞（真的拥塞应该一个包都收不到才对）。下面的图片说明了为何会出现重复确认的ACK。



- **快恢复**：快重传算法运行后发送方认为网络应该没有拥塞，因此发送方将调整门限值为$ssthreash = curr\ cwnd/2$，并且直接将当前拥塞窗口调整到这个值（废话），此后立刻开始执行拥塞避免阶段。

5. TCP的连接的建立和释放

TCP的连接管理

需要有一项机制使得运输连接的三个阶段都能正常进行。协商内容要包括确定连接双方的存在，商议如最大窗口大小等参数，并且能对运输的实体资源，如缓存和项目进行分配。

TCP连接的建立

采用的是**三报文握手**，目的是防止失效的连接请求报文突然又送到的情况。

（1）客户端向服务端发送报文：SYN=1, seq=x请求建立连接

（2）服务端收到后向客户端发送确认：SYN=1，ACK=1，ack=x+1，seq=y

（3）客户端收到后再向服务端发送确认：ACK=1，ack=y+1，seq=x+1

ACK是确认位，置1表示肯定答复，SYN表示请求连接。因此以上过程其实是双向请求过程。

TCP连接的释放

采用的是**四报文握手**，在通信完成后，两方都可以向对方发送连接释放请求，只要某一方确定自己不会再向对方发送请求了，那么就可以向对方发送连接释放请求。

（1）A向B发送报文：FIN=1，seq=x请求释放连接

（2）B向A发送确认：ACK=1，ack=**x**，seq=y【注意：此处确认号就是x!】

....此时TCP连接是半关闭的状态，B还能向A再发送数据，但是A已经不会再向B发送数据了...

（3）B向A发送报文：FIN=1，ACK=1，ack=**x**，seq=w

（4）A向B发送确认：ACK=1,ack=w, seq=x+1

6. 应用层

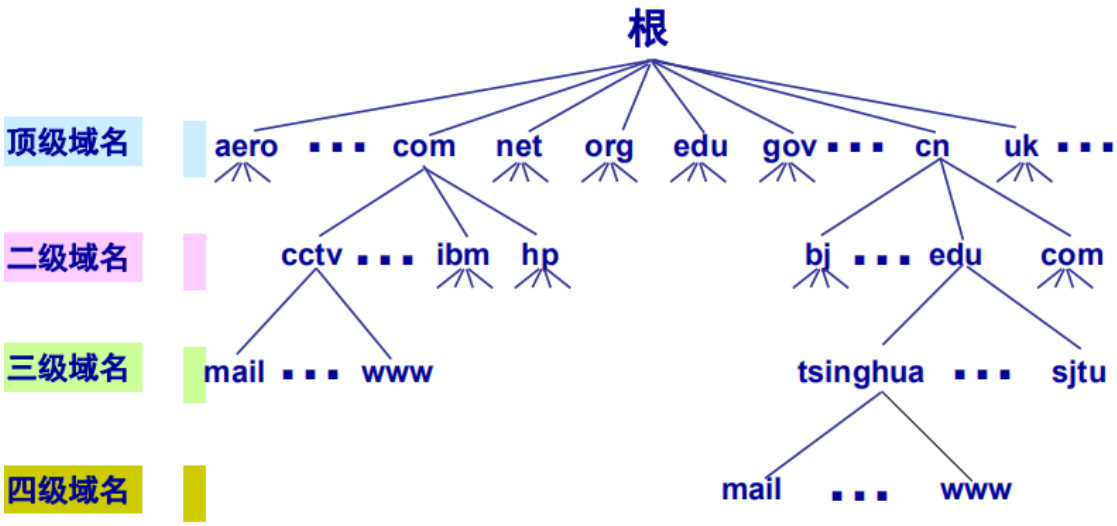
1. 应用层协议的作用和最小单位

应用层协议的作用是为了解决某一类应用问题，通过位于不同主机之间的多个应用进程的通信和协同工作来完成。其基本工作模式是客户-服务器方式。

应用层的PDU是消息/报文。

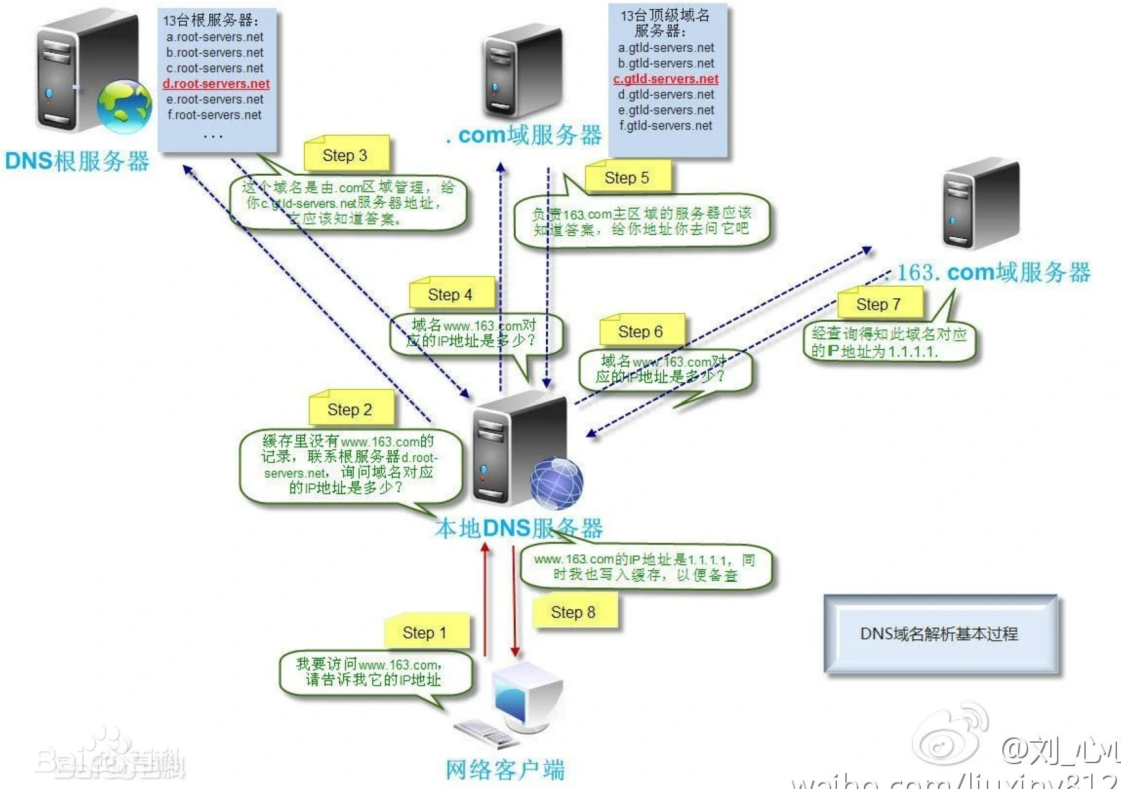
2. 域名系统DNS

- 互联网采用层次结构的命名树作为主机的名字，使用分布式的**域名系统DNS**，其作用其实是设置了一个让**域名服务器**能从域名到IP地址的解析对应。由于IP地址的唯一性，因此任何一个一一对应的域名都是唯一的。其存在意义其实是方便人的使用，这是一个逻辑概念。
- **域名的结构**：由标号序列组成，用点号隔开，级数大的域名在前，顶级域名放在最后。需要注意的是域名的点和IP地址的点没有对应关系。



- **域名服务器**：存在根域名服务器，顶级域名服务器，权限域名服务器和本地域名服务器
  - 根域名服务器：所有根域名服务器都知道所有的顶级域名服务器和IP地址
  - 顶级域名服务器：TLD服务器，负责管理在该顶级域名服务器注册的所有二级域名。当收到DNS查询请求时，它会给出相应的回答，也可能找到下一个域名服务器的IP地址。
  - 权限域名服务器：负责一个区的域名服务器。
  - 本地域名服务器：当主机发送DNS查询请求时，请求报文就发送给本地域名服务器。本地域名服务器也叫默认服务器。

以上的查询过程叫迭代查询，实际上，主机直接发送请求的对象是本地域名服务器，然后本地域名服务器分别给根服务器，顶级域名服务器，权限服务器发送查询。

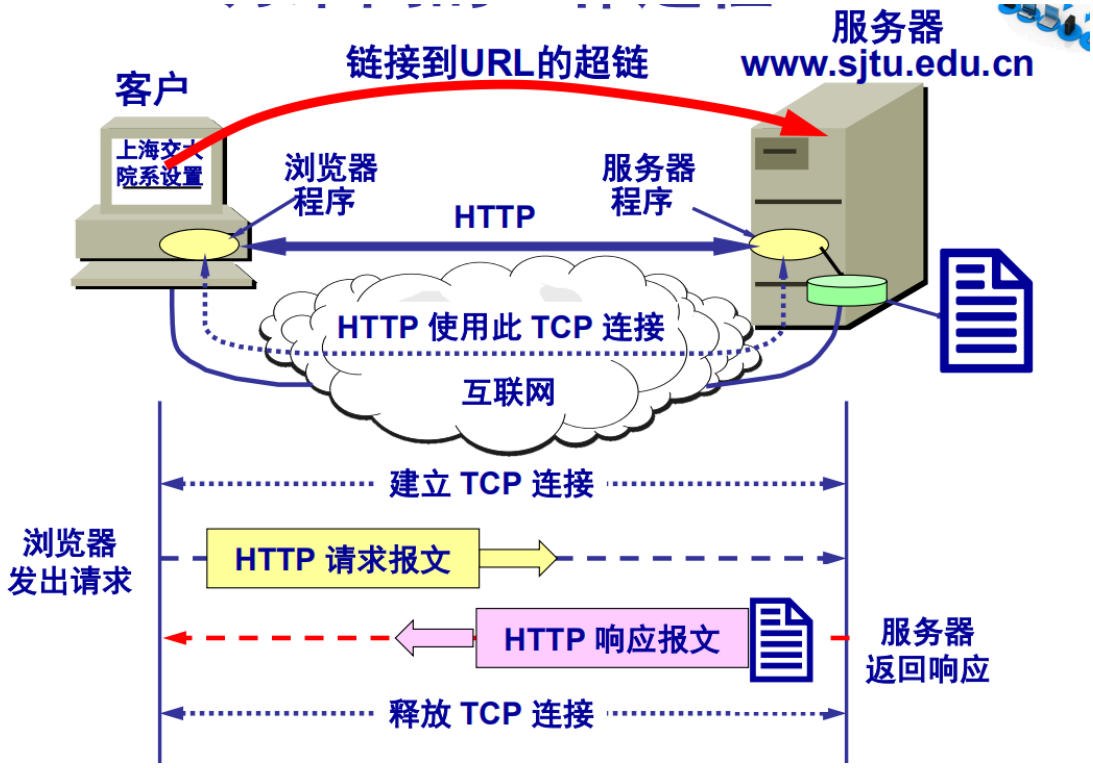


2. 文件传输协议FTP

- **FTP概述**：FTP是互联网上使用最广泛的文件传送协议，它提供交互式的访问，允许文件具有存取权限。它屏蔽了不同操作系统的细节，因此适合在异构网络中的任意计算机之间传送文件。
- **FTP特点**：使用两个TCP可靠连接，只提供基本功能，使用客户服务器方式，使用一个主进程负责接收新请求（控制进程），从属进程处理单个请求（数据传送进程）。
- **使用端口**：连接建立请求时使用21号端口，数据传送连接使用20号端口。

3. 万维网www

- **万维网概述**：不是特殊的计算即网络，是一种大规模，特殊的，联机式的信息储藏所。万维网用链接的方式从互联网上的一个站点访问另外一个站点。
- **基础**：万维网是超媒体系统，它是超文本系统的扩展。
  - 超文本：是由多个信息源链接而成的，利用一个链接可以使得用户从一个文档找到另外一个文档，这个文档本身可以存在于任何一个连接在互联网上的超文本系统中。
  - 超媒体：与超文本的不同之处在于其文档的内容不同，超文本只包含文本信息，而超媒体文档还包含其他表示方式的信息，如图形，图像，声音和动画。
- **客户-服务器工作模式**：浏览器是用户主机上的万维网客户程序，而万维网文档所驻留的计算机则运行服务器程序，因此这个计算机也称为万维网服务器。当客户程序向服务器发出请求时，服务器就向客户送回其所需要的万维网文档。而客户程序主窗口上显示的万维网文档被称为页面。
- **标注所有万维网文档——URL**：URL是统一资源定位符，用来标志万维网上的各种文档，使得每个文档在互联网范围内有唯一的标识符。其本质上是对互联网上得到的资源的位置和访问方法的简洁标识，其格式为<协议>：//<主机>：<端口>/<路径>。其中协议可以是http协议，而主机则是存放该资源的主机在互联网中的域名（IP），端口和路径有时可以省略。
- **实现超链接——HTTP**：万维网客户程序和服务器之间使用的协议是超文本传送协议HTTP，这是一个应用层协议，使用TCP进行可靠传输。其请求一个万维网文档的时间是一个发起TCP连接的RTT和一个HTTP请求报文的时间。



- **在各种计算机上显示超链接——HTML**：超文本标记语言使得万维网页面的设计者可以将一个超链从本页面地某处连接到互联网上任何一个万维网页面，并且能显示出来。HTML文档是嵌入了各种标签的万维网页面，是可以用任何文本编辑器创建的ASCII文件。

- **代理服务器/万维网高速缓存**：它代表浏览器发出HTTP请求，并会将最近的请求和响应暂存在本地磁盘中。当再次发出相同请求时，就不需要再按照URL去访问资源了。浏览器和代理服务器之间也要建立TCP连接，而代理服务器和万维网服务器之间也有TCP连接。
- **服务器存储用户信息——Cookie**：Cookie表示在HTTP服务器和客户之间传递的状态信息，服务器会为用户生成唯一的识别码，从而追踪用户在网站中的活动。
- **静态文档&动态文档&活动文档**：静态文档是不会改变的文档，动态文档是在浏览器访问万维网服务器是才由应用程序动态创建的文档，二者之间的差别只体现在生成文档的服务器，而对浏览器没有影响。而活动文档则是将生成文档的技术工作交给了浏览器，让其在主机本地运行程序进行交互。
- **用户查找所需信息——搜索引擎**

4. 电子邮件

- **电子邮件所用协议**：发送邮件的协议SMTP，读取邮件的协议POP3和IMAP。
- **电子邮件地址的格式**：<用户名>@<邮箱所在的主机的域名>，用户名在域名范围内是唯一的。
- **STMP协议**：连接直接建立在发送主机的STMP服务器和接收主机的STMP服务器之间的TCP连接，不使用中间的邮件服务器。连接建立后进行邮件传送，邮件传送完之后释放TCP连接。
- **POP3和IMAP协议**：POP邮局协议，接收邮件的用户主机中运行POP客户程序，在用户所连接的ISP邮件服务器中运行POP服务器程序。IMAP则让用户可以在自己的主机上操纵邮件服务器ISP，因此其为一个联机协议。因此其相比POP协议，可以让用户在不同的地方使用不同的计算机随时上网阅读和处理自己的邮件，它还允许收件人只读取邮件的一部分。但IMAP的缺点是如果用户没有将邮件复制到自己的主机上，用户需要经常访问服务器。
- **万维网上的电子邮件**：用户和邮件服务器之间建立HTTP协议，邮件服务器之间建立SMTP协议。

5. 动态主机配置协议DHCP

- **作用**：用于手动配置IP地址，其将软件做成通用的和便于一直的，编写者将协议软件参数化，给这些协议参数赋值的动作就叫做协议配置。
- **需要配置的参数**：（1）IP地址；（2）子网掩码；（3）缺省路由器的IP地址；（4）域名服务器的IP地址。这些参数存放在配置文件中。
- **效果**：DHCP提供了即插即用的连网机制，动态主机配置DHCP允许一台计算机加入新的网络和获取IP地址而不用手工参与。
- **过程**：DHCP报文是UDP用户数据报中的数据，主机以广播方式发送发现请求，DHCP中继代理单播给DHCP服务器。而DHCP服务器会给主机一个临时的IP地址，IP地址存在租期，这个租期可以是服务器决定的，也可以是用户自己制定的。

6. P2P应用

- **P2P概念**：使得音频和视频文件可以在普通用户之间进行传输，相当于互联网中的普通用户是服务器，这样可以规避集中式服务器的性能拼劲问题。

7. 5G

1. 5G的关键技术有哪些

1. 高频段传输；2. 新型多天线传输技术3D-MIMO；3. 同时同频全双工技术；4. D2D技术，允许终端之间复用小区资源直接进行通信；5. 密集和超密集组网技术；6. 新型网络架构