

一、概述

1. 概念、组成、功能和分类

计算机网络：一些相互连接的、以共享资源为目的的、自治的计算机的集合。



2. 标准化工作及相关组织

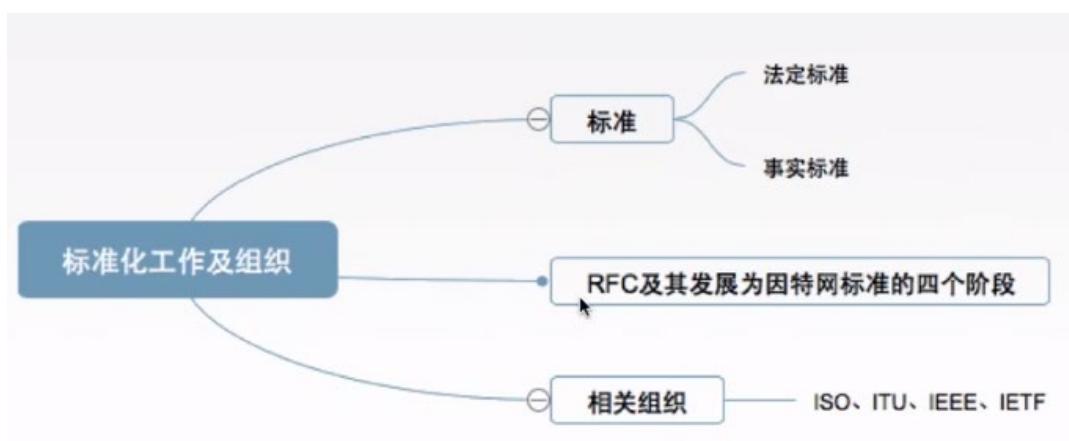
相关组织：

国际标准化组织 ISO OSI参考模型、HDLC协议

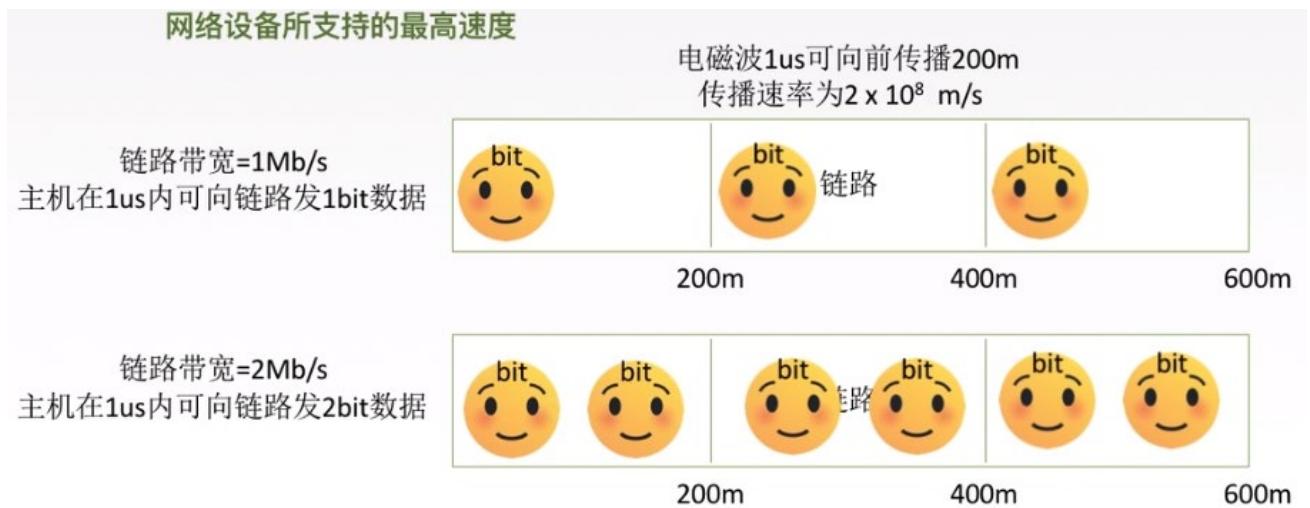
国际电信联盟 ITU 制定通信规则

国际电气电子工程师协会 IEEE 学术机构、IEEE802系列标准、5G

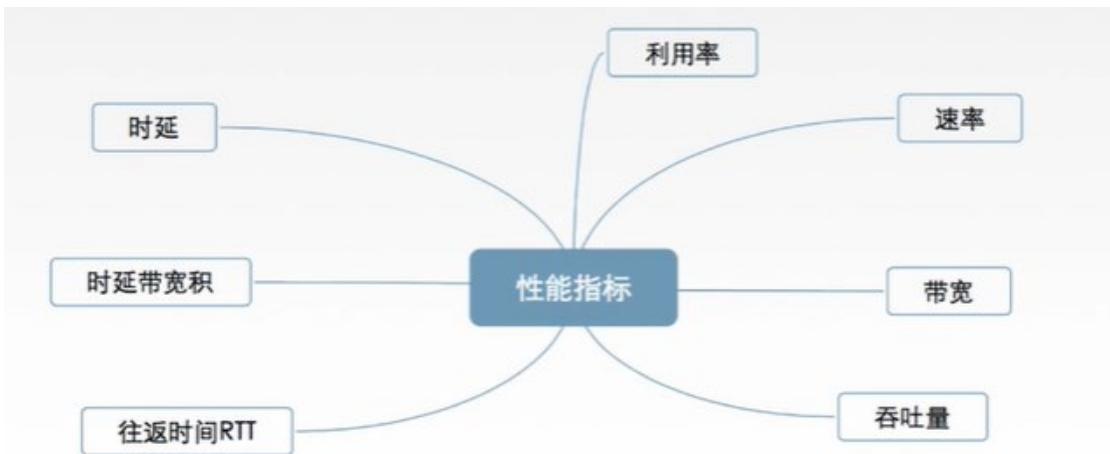
Internet工程任务组 IETF 负责因特网相关标准的制定 RFC XXXX



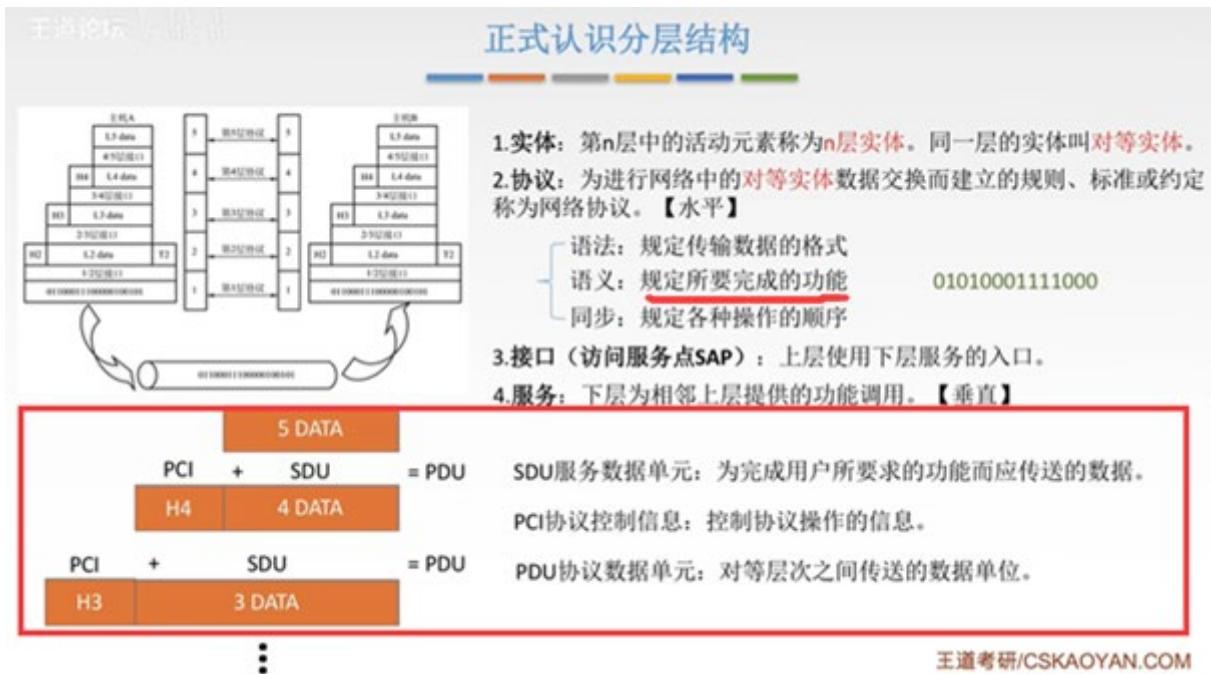
3. 性能指标——速率等



4. 性能指标——时延等



5. 分层结构、协议、接口、服务



概念总结

网络体系结构是从功能上描述计算机网络结构。

计算机网络体系结构简称网络体系结构是分层结构。

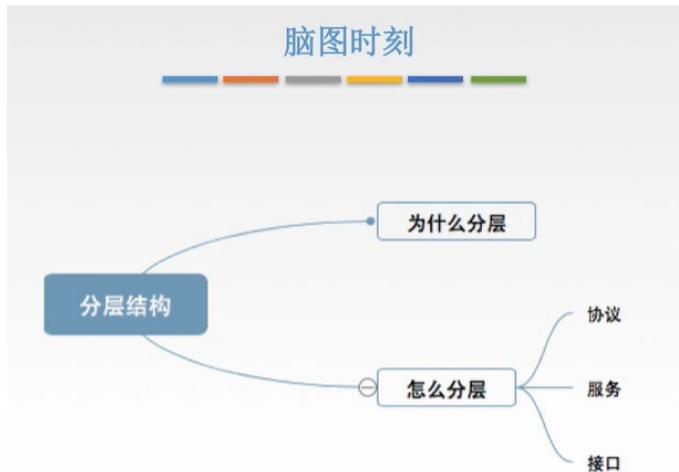
每层遵循某个/些网络协议以完成本层功能。

计算机网络体系结构是计算机网络的各层及其协议的集合。

第n层在向n+1层提供服务时，此服务不仅包含第n层本身的功能，还包含由下层服务提供的功能。

仅仅在相邻层间有接口，且所提供的服务的具体实现细节对上一层完全屏蔽。

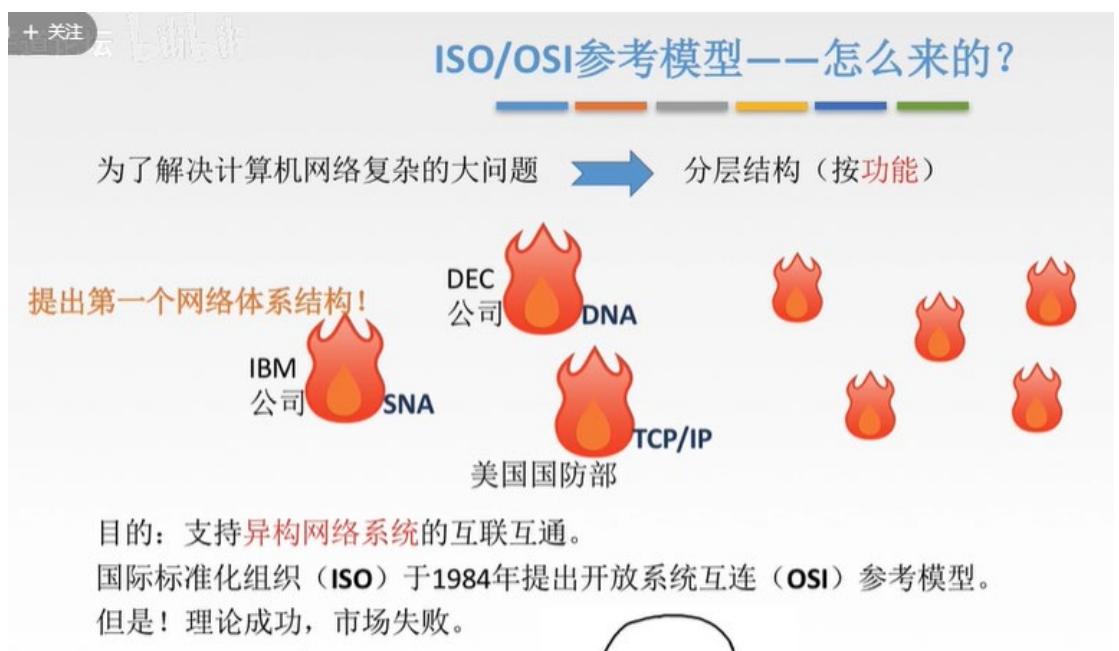
体系结构是抽象的，而实现是指能运行的一些软件和硬件。



6. OSI 参考模型(1)——解释通信过程

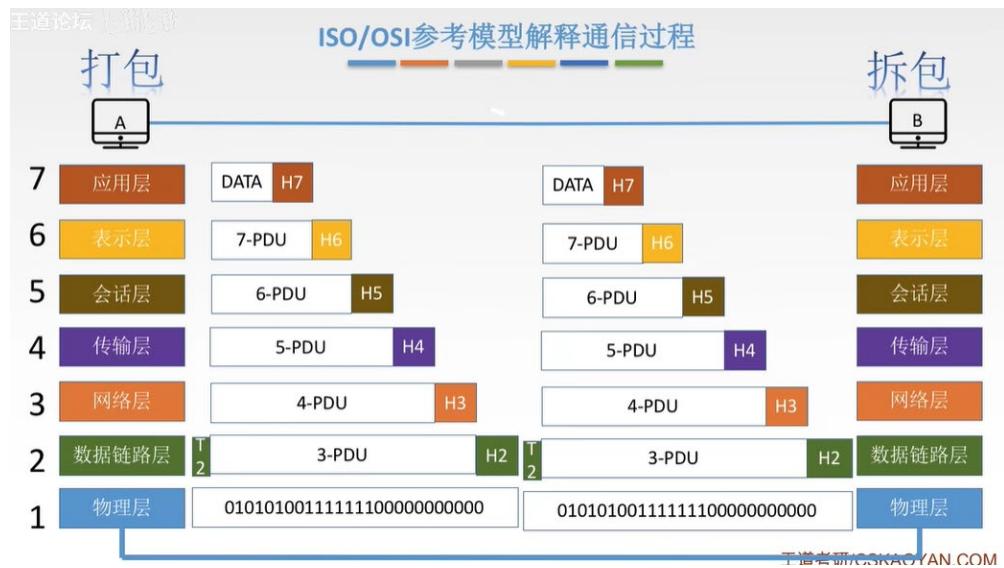


五层结构主要是为了学习思路更加清晰才产生的





对等实体(层次)之间的协议规定相同层次实现的功能相同。



7. OSI 参考模型(2)——每一层具体功能及使用到的协议

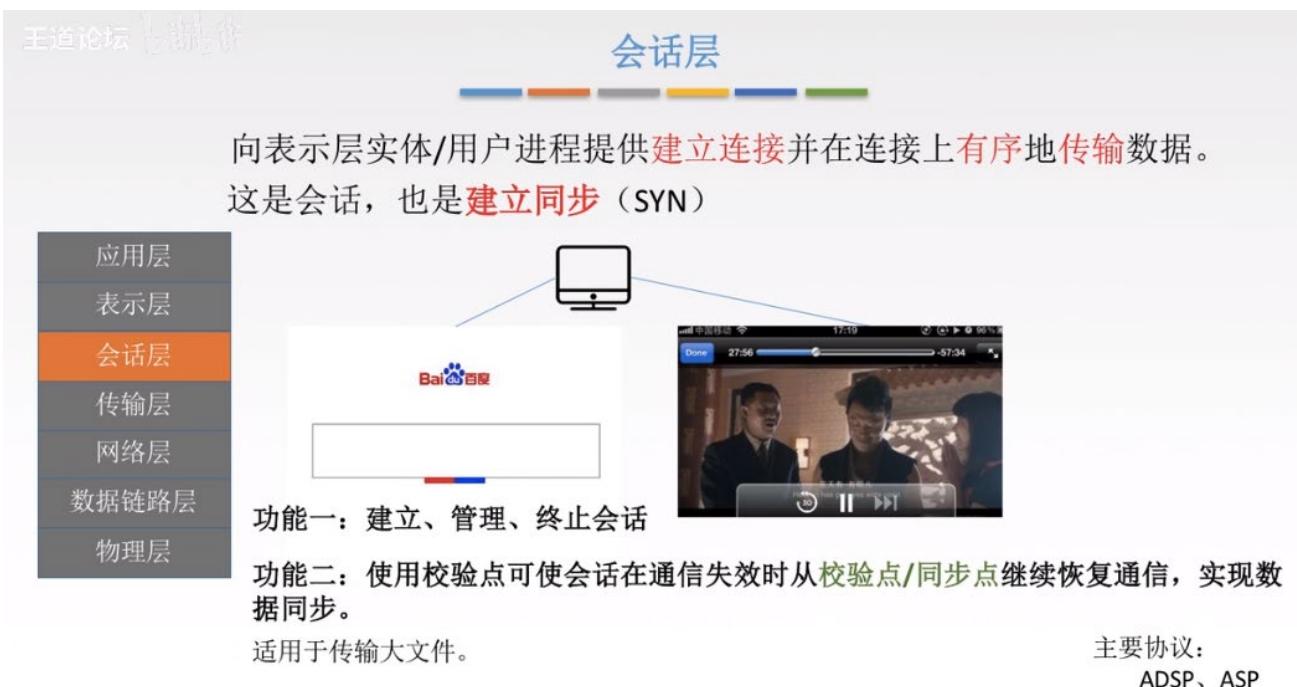
①应用层：连上网才能使用的程序。



②表示层：没有什么单独的协议，与表示层有关的协议：如下图。



③会话层



发送数据较长时，会在数据中间插入若干校验点/同步点。这样可以使会话在通信失效时从最近的校验点/同步点继续恢复通信，实现数据的同步。

④传输层：从下而上，第一个面向**端到端通信**的层次；也是资源子网和通信子网的**接口**。

主机内的进程，都会用一个编号标识，这里称为端口号，所以称为端到端通信。

区分：端到端(只管最终目的地)，点到点(需要管怎么一步一步到达最终目的地的)

可靠传输(基于确认机制的过程)：每一个报文段需要收到目的主机返回的确认收到信息才能继续发送下一个报文段；当一直没收到返回的确认收到信息时，发送端会再次发送该报文段，直到收到返回的确认收到信

息才能继续往下发送新的报文段。TCP?

不可靠传输: 不需要建立连接,也无需管是否会丢失,直接把数据报往上发即可。适用于发送小数据,如QQ聊天发送一句话。UDP?

差错控制: 报文段失序或丢失了,传输层负责纠正这些错误。

流量控制: 发送速度和接收速度是否匹配,实时调节发送速度

复用分用: 形成每个报文段时会包含应用层发送程序的端口号。如下图



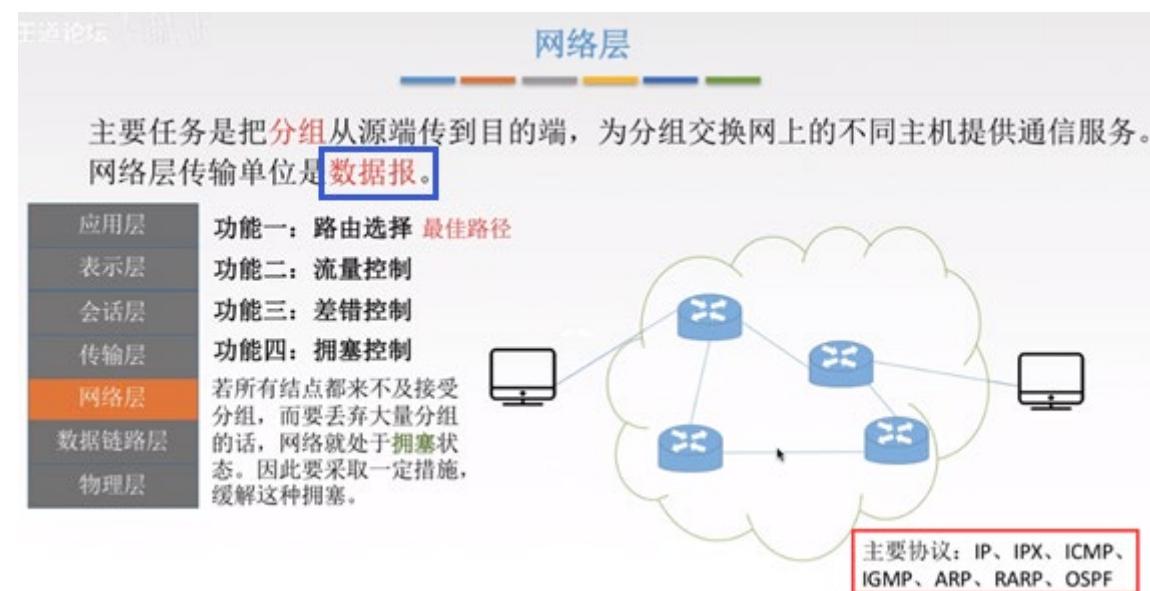
⑤**网络层:** IP 层、网际层(最重要的层次)。 数据报过长时,就拆分为分组。

路由选择: 根据网络情况,由路由算法选择最佳路径

流量控制: 发送端速度的控制

差错控制: 通信两节点之间约定特定规则,如奇偶校验码等。节点在收到分组时,进行校验。有错纠错,无法修正时扔掉分组,确保**传输层**提交的数据没有问题。

拥塞控制: 针对全局,如下图。



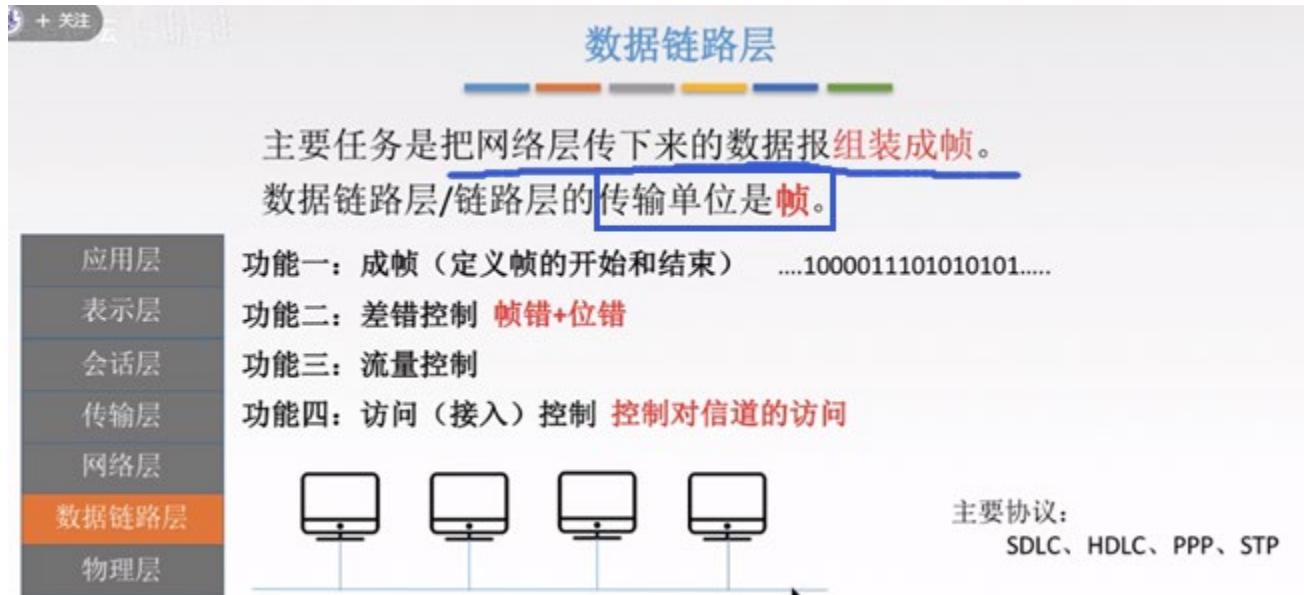
⑥**数据链路层:** 为网络层提供服务。

成帧: 接收端收到帧时，提取数据部分上交给网络层进行下一步的解封装

差错控制: 检错、纠错、丢弃

流量控制: 接收方缓存不够用时，再传输来的数据都会被丢弃，此时告诉发送方慢点发

访问（接入）控制: 如，广播式网络中多个主机共享一个公共信道，同一时间只能有一个主机在发送信息，其余都处于监听状态，即控制对共享信道的访问。



⑦**物理层**（“傻瓜层”）：将比特流转成电信号的形式在链路(物理媒体，任意一种物理传输介质：如铜轴电缆、双绞线、无线电波)上进行传输，不需要对数据有改动和切割。

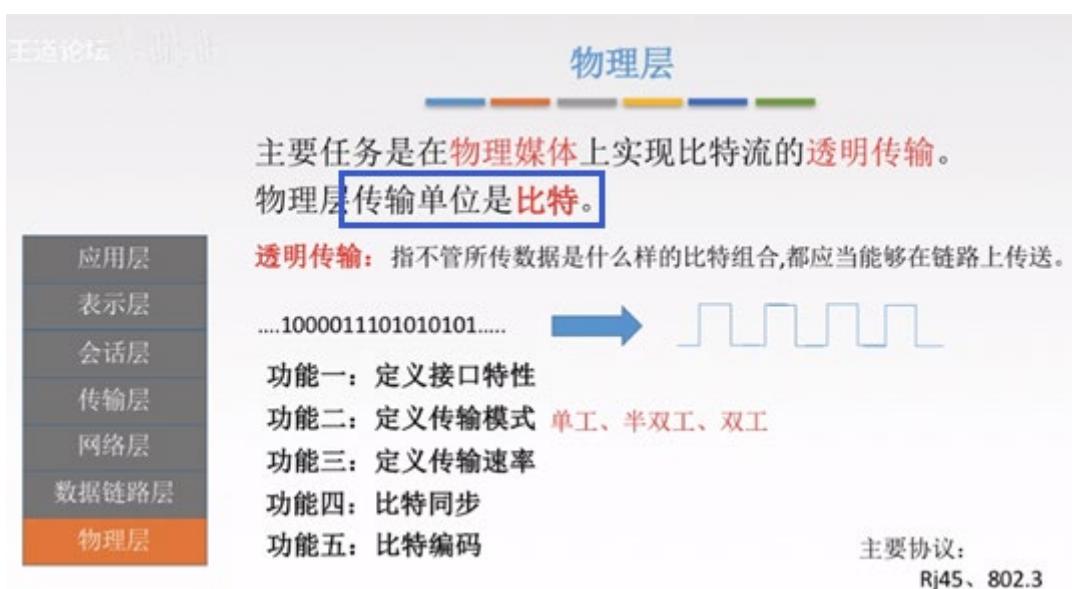
定义接口特性：如连接电缆的插头要有多少个引脚，引脚如何连接等。

定义传输模式

定义传输速率：如十兆网、百兆网，指发送端的发送速率(传输速率)。

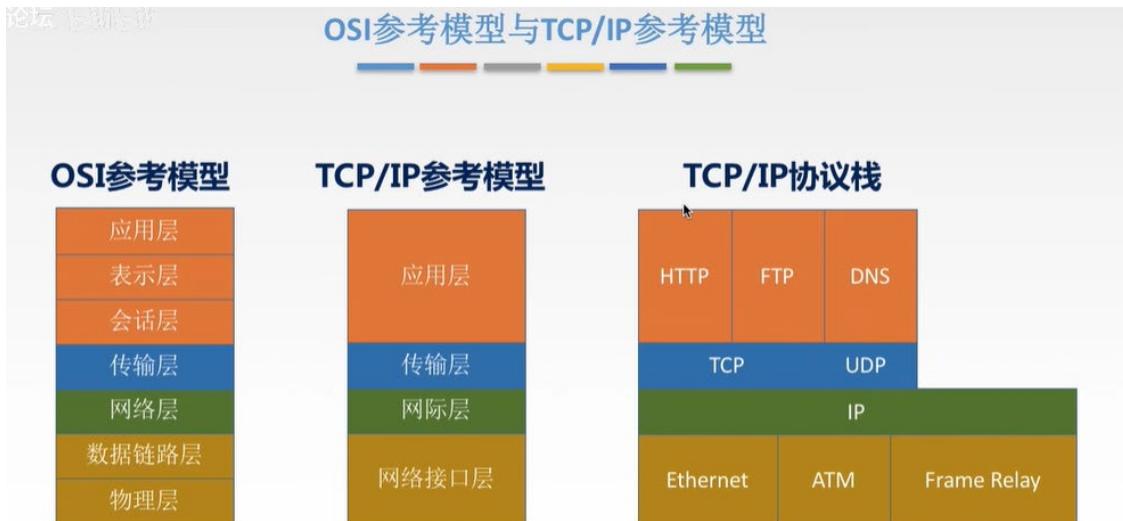
比特同步

比特编码：如多少电压表示 1，多少电压表示 0。



8. TCP、IP 参考模型和 5 层参考模型

①TCP/IP 参考模型



OSI 参考模型与 TCP/IP 参考模型相同点

- 都分层
- 基于独立的协议栈的概念
- 可以实现异构网络互联

OSI 参考模型与 TCP/IP 参考模型不同点

面向连接分为三个阶段，第一是建立连接，在此阶段，发出一个建立连接的请求。只有在连接成功建立之后，才能开始数据传输，这是第二阶段。接着，当数据传输完毕，必须释放连接。而面向无连接没有这么多阶段，它直接进行数据传输。



	ISO/OSI 参考模型	TCP/IP 模型
网络层	无连接+面向连接	无连接
传输层	面向连接	无连接+面向连接

1. OSI 定义 3 点：服务、协议、接口
2. OSI 先出现，参考模型先于协议发明，不偏向特定协议
3. TCP/IP 设计之初就考虑到异构网 互连 问题，将 IP 作为重要层次
4. ←

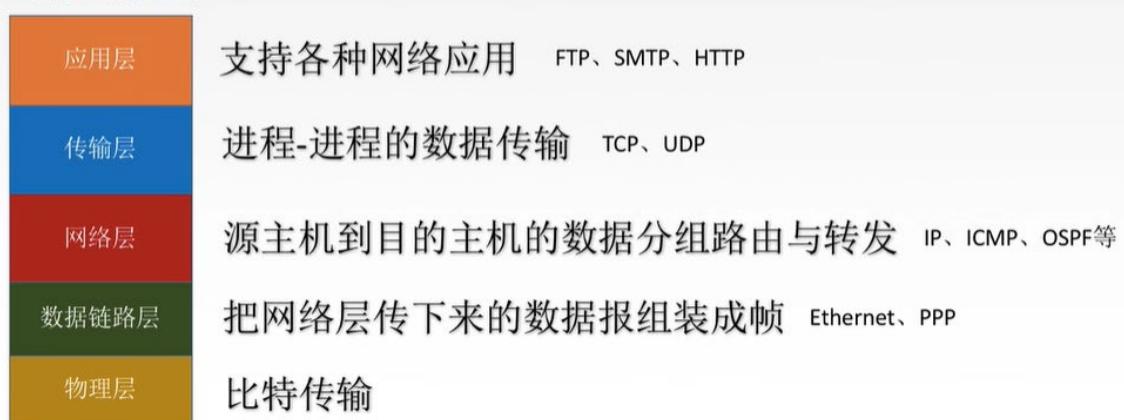
②五层参考模型

OSI 优点：每一层的功能都描述的很清楚

TCP/IP 优点：层次简单，层间重复的功能少

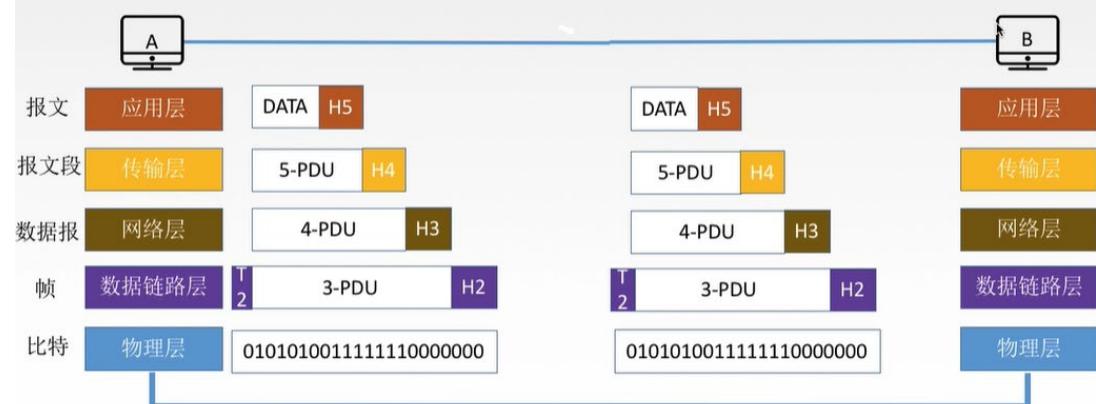
5层参考模型

综合了OSI和TCP/IP的优点



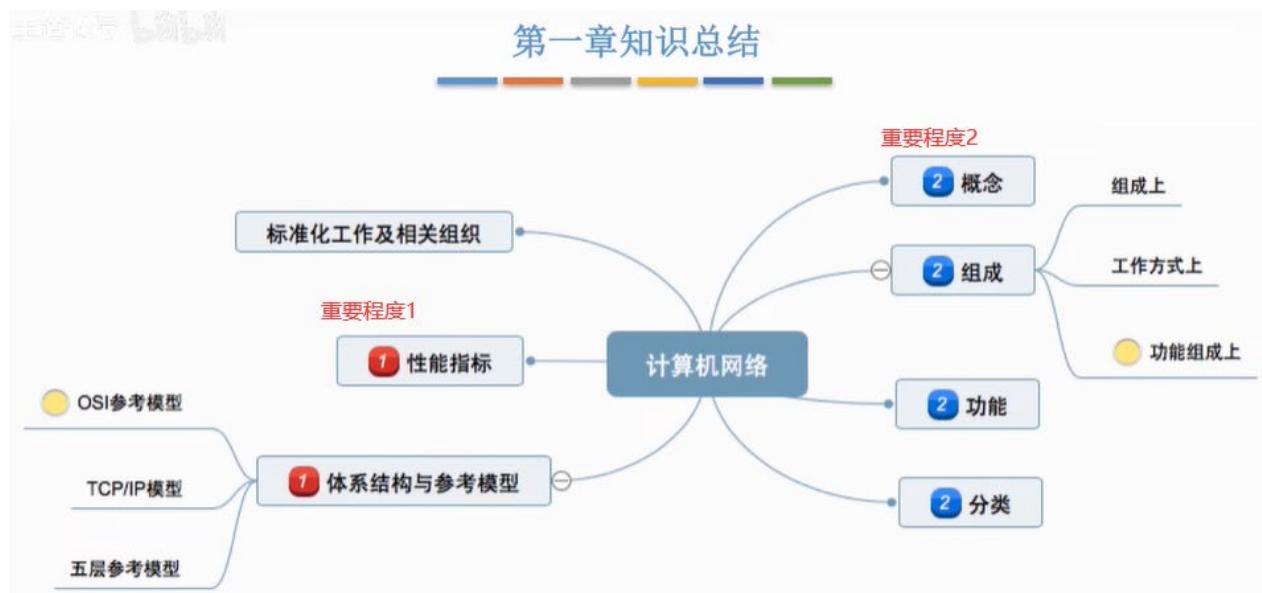
王道论坛

5层参考模型的数据封装与解封装



9. 第一章总结

第一章知识总结



二、物理层

- 第二章の刷透
- 1.通信基础 概念和公式
 - ★ 2.两个公式 \lim
 - 3.看图说话
 - 4.传输介质
 - 5.物理层设备

1. 物理层基本概念

物理层只关注本层的内容及与下面传输媒体的接口(传输媒体常称为参考模型第0层,要和物理层区分开)



2. 数据通信基础知识

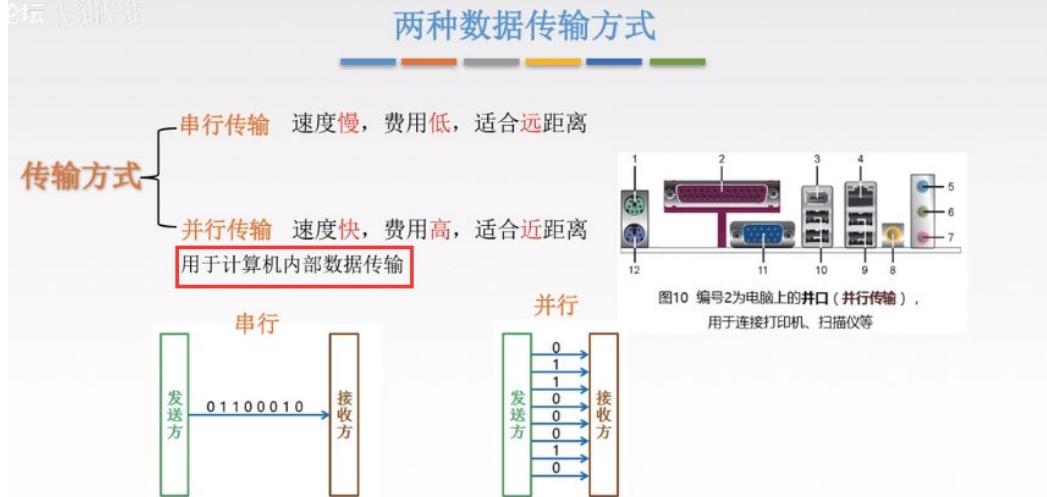
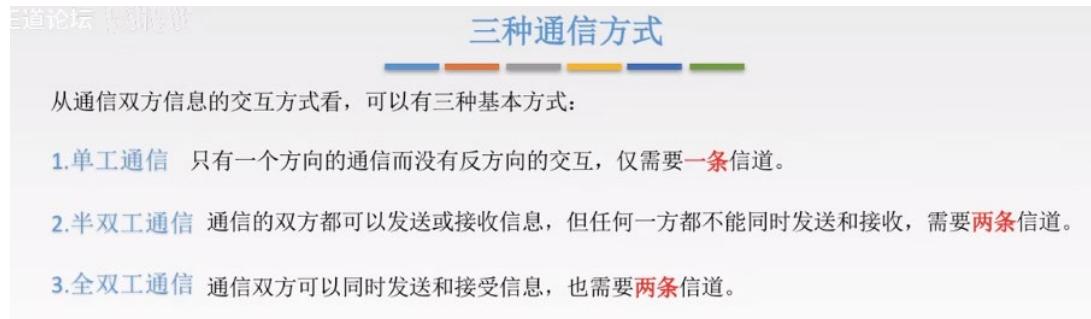
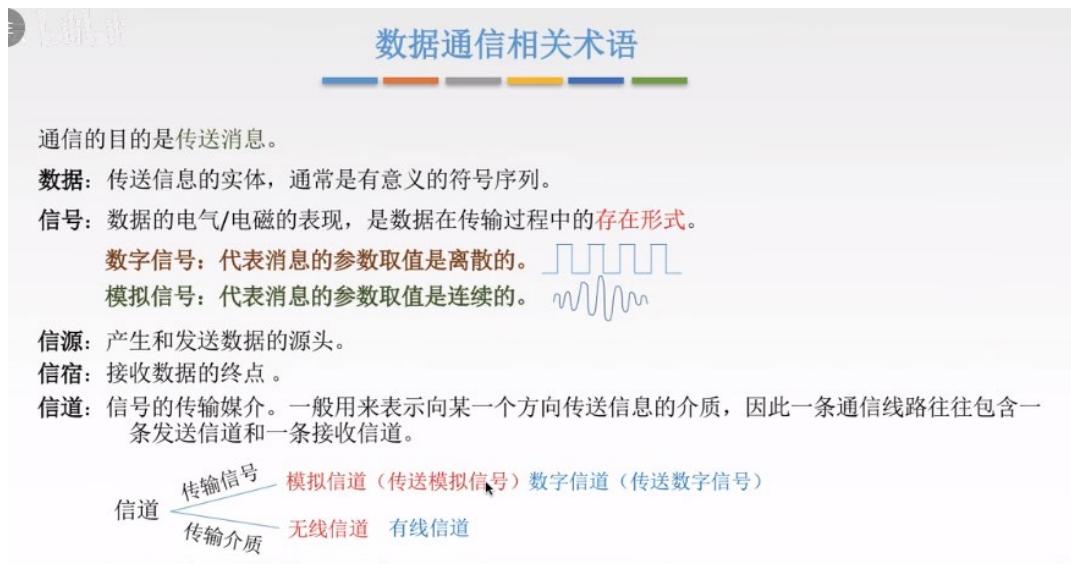
数据通信的范围很广,包括打电话、电脑、手机、传真机等



①通过电话线入网,电话线只能传输模拟信号,而计算机网卡发出的是数字信号,所以就需要调制解调器

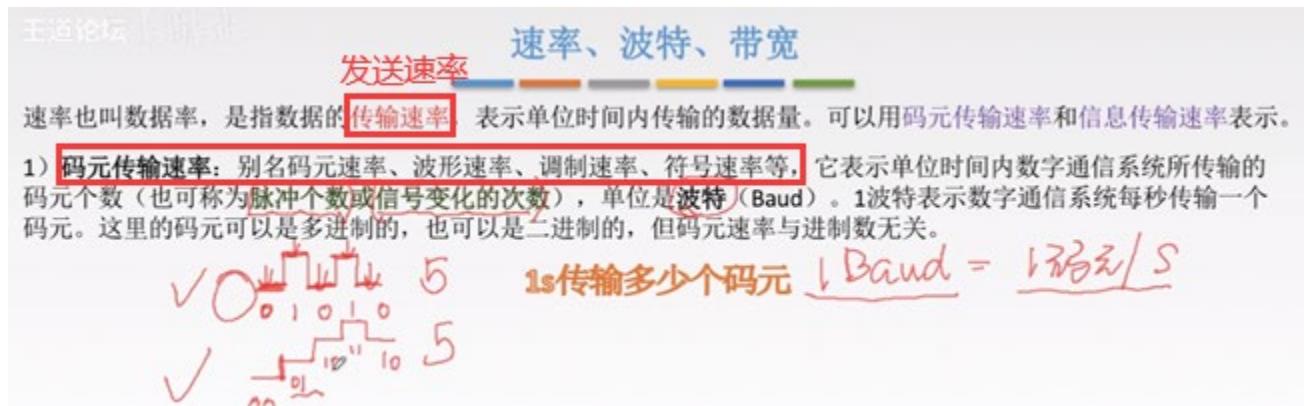
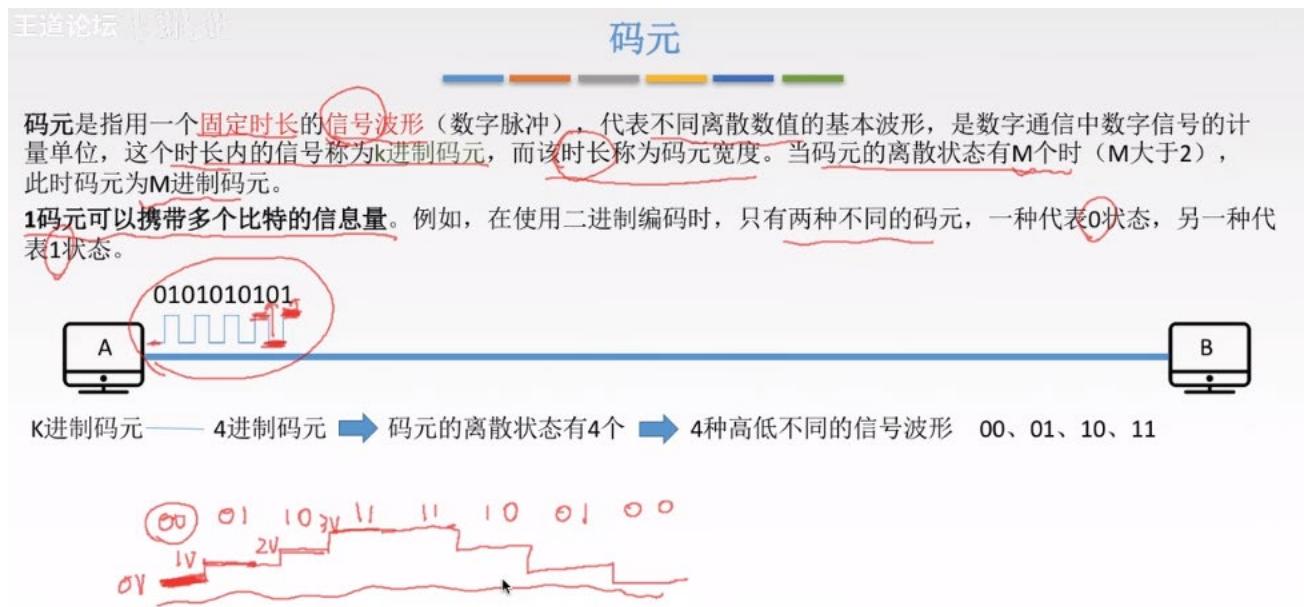
(俗称“猫”)转换信号

②宽带式入网，不需要“猫”，直接插上即可入网。



“并口上的孔就是信道”

3. 码元、波特、速率、带宽



上面二进制码元和四进制码元，一秒都是传输5个码元，所以码元速率一样。但它们所包含的比特数不一样。

2) 信息传输速率：别名信息速率、比特率等，表示单位时间内数字通信系统传输的二进制码元个数（即比特数），单位是比特/秒（b/s）。

1s传输多少个比特

关系：若一个码元携带 n bit的信息量，则 M Baud的码元传输速率所对应的信息传输速率为 $M \times n$ bit/s。

$$5 \times 2 = 10 \text{ b/s}$$

带宽：表示在单位时间内从网络中的某一点到另一点所能通过的“最高数据率”，常用来表示网络的通信线路所能传输数据的能力。单位是b/s。

计算练习题：

练习题

某一数字通信系统传输的是四进制码元,4s传输了8000个码元,求系统的码元传输速率是多少?信息传输速率是多少?若另一通信系统传输的是十六进制码元,6s传输了7200个码元,求他的码元传输速率是多少?信息传输速率是多少?并指出哪个系统传输速率快?

2000Baud, 4000b/s; 1200Baud, 4800b/s; 十六进制更快

四进制码元系统

码元传输速率就是 $8000/4=2000$ Baud, 信息传输速率就是 $2000 \cdot \log_2 4 = 4000$ b/s

十六进制码元系统

码元传输速率就是 $7200/6=1200$ Baud, 信息传输速率就是 $1200 \cdot \log_2 16 = 4800$ bit/s

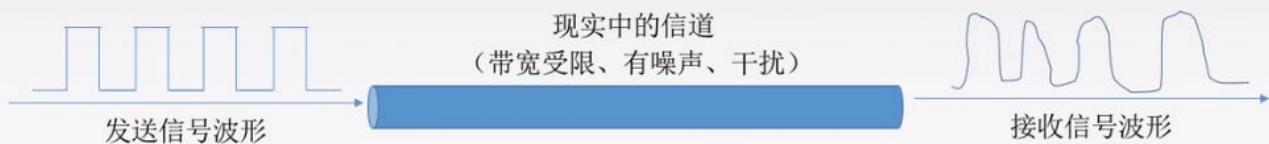
系统传输的是**比特流**,通常比较的是信息传输速率,所以传输十六进制码元的通信系统传输速率较快,如果用该系统去传输四进制码元会有更高的码元传输速率。

因为系统传输的是比特流,第二种系统的信息传输速率是不变的,所以换成四进制,单位时间内传输的比特数不变,那么每个码元又有两比特的信息量,所以码元传输速率就变了。

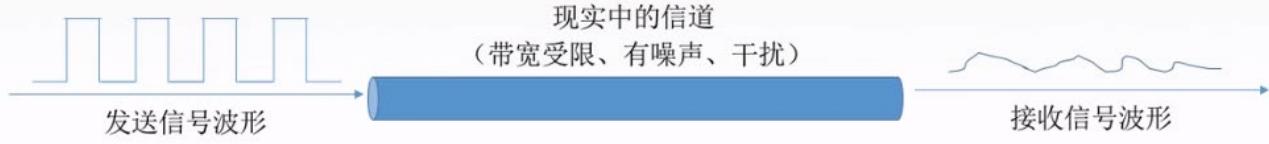
4. 奈氏准则和香农定理

失真

有失真但可识别



失真大无法识别

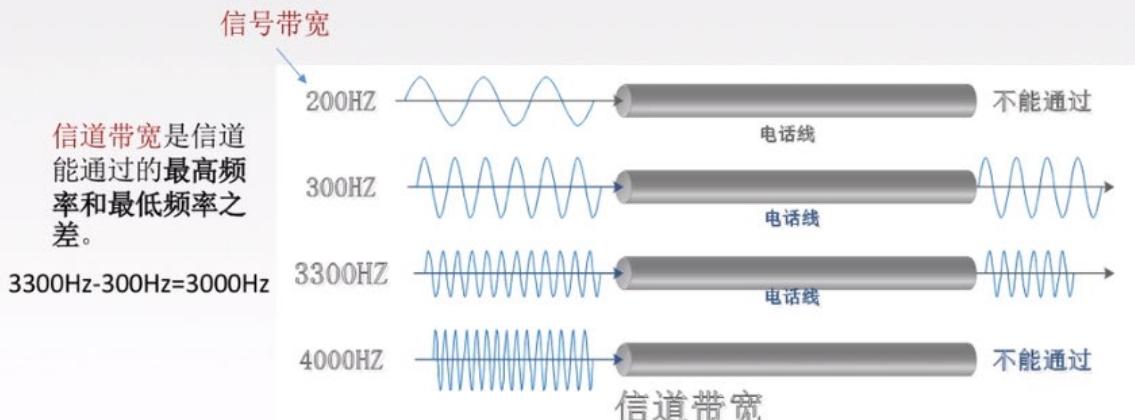


正相关

负相关

影响失真程度的因素: 1. 码元传输速率 2. 信号传输距离 3. 噪声干扰 4. 传输媒体质量

失真的一种现象——码间串扰



码间串扰：接收端收到的信号波形失去了码元之间清晰界限的现象。

信号频率过低，传播到终点可能都衰减没了；信号频率过高(码元传输速率过快)，码间串扰。

解决码间串扰——奈氏准则：

奈氏准则（奈奎斯特定理）

奈氏准则：在理想低通（无噪声，带宽受限）条件下，为了避免码间串扰，极限码元传输速率为 $2W \text{ Baud}$ ， W 是信道带宽，单位是 Hz 。

为了混淆大家，再求一步极限数据率吧~

只有在这两个公式
这带宽才用 Hz !

$$\text{理想低通信道下的极限数据传输率} = 2W \log_2 V \text{ (b/s)}$$

↓ ↓
几种码元/码元的离散电平数目
带宽(Hz)

1. 在任何信道中，码元传输的速率是有上限的。若传输速率超过此上限，就会出现严重的码间串扰问题，使接收端对码元的完全正确识别成为不可能。
2. 信道的频带越宽（即能通过的信号高频分量越多），就可以用更高的速率进行码元的有效传输。
3. 奈氏准则给出了码元传输速率的限制，但并没有对信息传输速率给出限制。
4. 由于码元的传输速率受奈氏准则的制约，所以要提高数据的传输速率，就必须设法使每个码元能携带更多个比特的信息量，这就需要采用多元制的调制方法。

练习题：

例. 在无噪声的情况下，若某通信链路的带宽为 3kHz ，采用4个相位，每个相位具有4种振幅的QAM调制技术，则该通信链路的最大数据传输率是多少？

信号有 $4 \times 4 = 16$ 种变化

最大数据传输率 = $2 \times 3\text{k} \times 4 = 24\text{kb/s}$

香农定理：

王道论坛

香农定理

噪声存在于所有的电子设备和通信信道中。由于噪声随机产生，它的瞬时值有时会很大，因此噪声会使接收端对码元的判决产生错误。但是噪声的影响是相对的，若信号较强，那么噪声影响相对较小。因此，**信噪比**就很重要。信噪比=信号的平均功率/噪声的平均功率，常记为S/N，并用分贝（dB）作为度量单位，即：

$$\text{信噪比 (dB)} = 10 \log_{10}(S/N) \quad \text{数值等价}$$

香农定理：在带宽受限且有噪声的信道中，为了不产生误差，**信息的数据传输速率有上限值。**

$$\text{信道的极限数据传输速率} = W \log_2(1+S/N) \quad (\text{b/s})$$

↓ ↓
带宽(Hz) 信噪比

S是信道所传信号的平均功率
N是信道内的高斯噪声功率

1. 信道的**带宽**或信道中的**信噪比**越大，则信息的极限传输速率就越高。
2. 对一定的传输带宽和一定的信噪比，信息传输速率的上限就确定了。
3. 只要信息的传输速率低于信道的极限传输速率，就一定能找到某种方法来实现**无差错的传输**。
4. 香农定理得出的为**极限信息传输速率**，实际信道能达到的传输速率要比它低不少。
5. 从香农定理可以看出，若信道带宽W或信噪比S/N没有上限（不可能），那么信道的极限信息传输速率也就没有上限。

练习题：

例. 电话系统的典型参数是信道带宽为3000Hz，信噪比为30dB，则该系统最大数据传输速率是多少？

$$30\text{dB}=10 \log_{10}(S/N)$$

$$\text{则 } S/N=1000$$

$$\text{信道的极限数据传输速率} = W \log_2(1+S/N) = 3000 \times \log_2(1+1000) \approx 30\text{kb/s}$$

区分与选择使用：

“Nice” 和 “香浓”

奈氏准则 内忧

带宽受限无噪声条件下，为了避免码间串扰，码元传输速率的上限 $2W \text{ Baud}$ 。

理想低通信道下的极限数据传输率 = $2W \log_2 V$

要想提高数据率，就要提高带宽/采用更好的编码技术。

香农定理 外患

带宽受限有噪声条件下的信息传输速率。

信道的极限数据传输速率 = $W \log_2(1+S/N)$

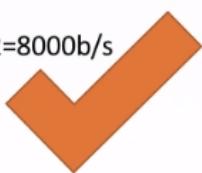
要想提高数据率，就要提高带宽/信噪比。

练习题：

题目：二进制信号在信噪比为 $127:1$ 的 4kHz 信道上传输，最大的数据速率可达到多少？

Nice: $2 \times 4000 \times \log_2 2 = 8000\text{b/s}$

香浓: $4000 \times \log_2(1+127) = 28000\text{b/s}$



5. 编码与调制(1)

王道论坛

基带信号与宽带信号

信道：信号的传输媒介。一般用来表示向某一个方向传送信息的介质，因此一条通信线路往往包含一条发送信道和一条接收信道。

信道 $\xrightarrow{\text{传输信号}}$ 模拟信道（传送模拟信号） 数字信道（传送数字信号）

传输介质 无线信道 有线信道

数字信号，所以计算机网络中基带信号指的就是数字信号

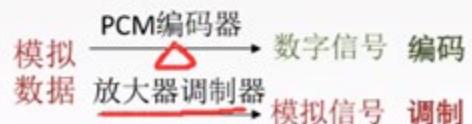
基带信号 将数字信号1和0直接用两种不同的电压表示，再送到数字信道上去传输（基带传输）。来自信源的信号，像计算机输出的代表各种文字或图像文件的数据信号都属于基带信号。基带信号就是发出的直接表达了要传输的信息的信号，比如我们说话的声波就是基带信号。

宽带信号 将基带信号进行调制后形成的频分复用模拟信号，再传送到模拟信道上去传输（宽带传输）。把基带信号经过载波调制后，把信号的频率范围搬移到较高的频段以便在信道中传输（即仅在一段频率范围内能够通过信道）。

在传输距离较近时，计算机网络采用基带传输方式（近距离衰减小，从而信号内容不易发生变化）

在传输距离较远时，计算机网络采用宽带传输方式（远距离衰减大，即使信号变化大也能最后过滤出来基带信号）

编码与调制

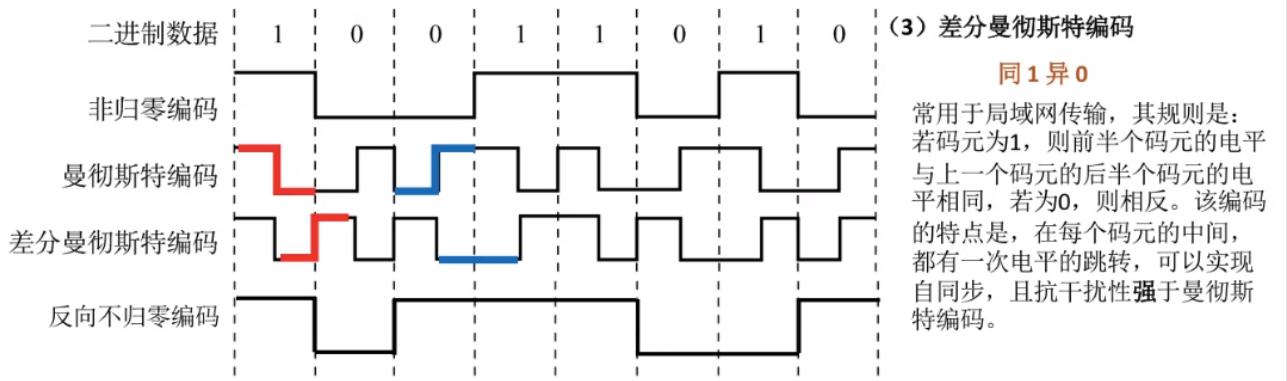


6. 编码与调制(2)——四种编码和调制的方法

①数字数据 编码为 数字信号

数字数据编码为数字信号

- (1) 非归零编码【NRZ】 (4) 归零编码【RZ】
 (2) 曼彻斯特编码 (5) 反向不归零编码【NRZI】
 (3) 差分曼彻斯特编码 (6) 4B/5B编码



(5) 反向不归零编码【NRZI】

信号电平翻转表示0，信号电平不变表示1。

(4) 归零编码【RZ】

信号电平在一个码元之内都要恢复到零的这种编码成编码方式。

(2) 曼彻斯特编码

将一个码元分成两个相等的间隔，前一个间隔为低电平后一个间隔为高电平表示码元1；码元0则正好相反。也可以采用相反的规定。该编码的特点是在每一个码元的中间出现电平跳变，位中间的跳变既作时钟信号（可用于同步），又作数据信号，但它所占的频带宽度是原始的基带宽度的两倍。每一个码元都被调成两个电平，所以**数据传输速率只有调制速率的1/2**。

(3) 差分曼彻斯特编码

同1异0

常用于局域网传输，其规则是：
 若码元为1，则前半个码元的电平与上一个码元的后半个码元的电平相同，若为0，则相反。该编码的特点是，在每个码元的中间，都有一次电平的跳转，可以实现自同步，且**抗干扰性强于曼彻斯特编码**。

(1) 非归零编码【NRZ】

高1低0

编码容易实现，但没有检错功能，且无法判断一个码元的开始和结束，以至于收发双方难以保持同步。

(6) 4B/5B编码

比特流中插入额外的比特以打破一连串的0或1，就是用5个比特来编码4个比特的数据，之后再传给接收方，因此称为4B/5B。编码效率为80%。

只采用16种对应16种不同的4位码，其他的16种作为控制码（帧的开始和结束，线路的状态信息等）或保留。

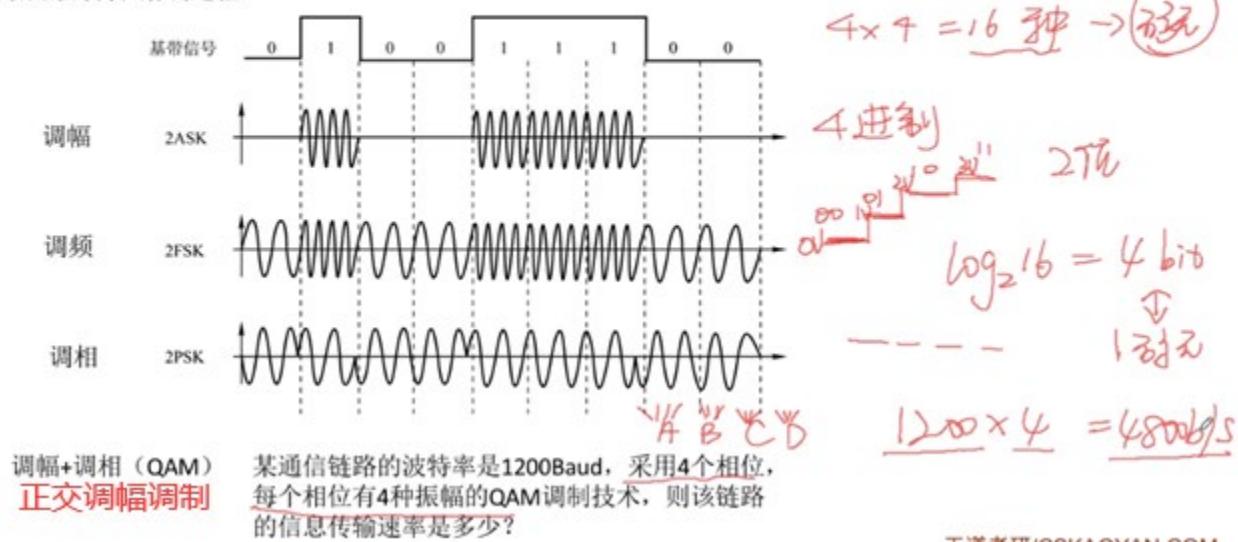
4比特数据符号	5比特编码
0000	11110
0001	01001
0010	10100
0011	10101
0100	01010
0101	01011
0110	01110
0111	01111
1000	10010
1001	10011
1010	10110
1011	10111
1100	11010
1101	11011

对于编码后不包含时钟信号的，还需要单独传输：时钟周期信号，即多久发送了一个比特，接收端应该多久接受一个比特

②数字数据 调制为 模拟信号

数字数据调制为模拟信号

数字数据调制技术在发送端将数字信号转换为模拟信号，而在接收端将模拟信号还原为数字信号，分别对应于调制解调器的调制和解调过程。



③模拟数据 编码为 数字信号

一种信号状态对应一种码元。

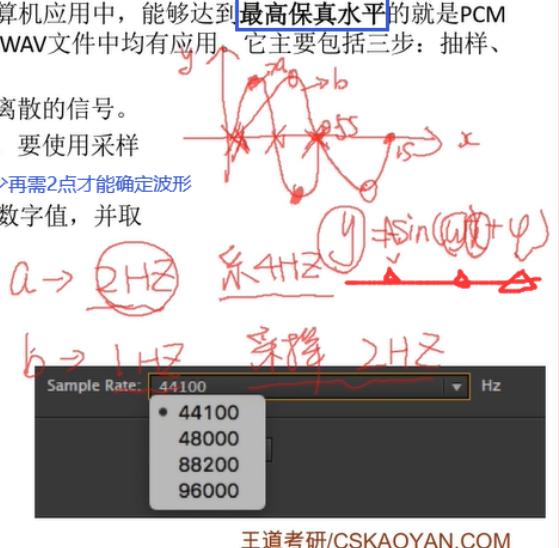
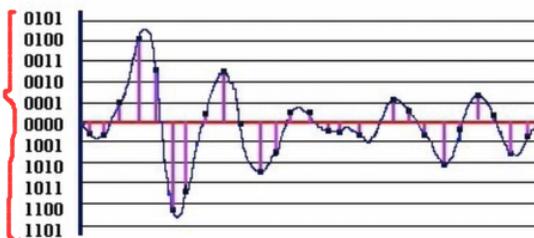
人能听到20Hz-20kHz频率的声波，所以音频设备采样频率必须在40kHz以上才能听到全部声音（保真）。高频失真就是采样频率不够。

模拟数据编码为数字信号

计算机内部处理的是二进制数据，处理的都是数字音频，所以需要将模拟音频通过采样、量化转换成有限个数字表示的离散序列（即实现音频数字化）。

最典型的例子就是对音频信号进行编码的脉码调制（PCM），在计算机应用中，能够达到最高保真水平的就是PCM编码，被广泛用于素材保存及音乐欣赏，CD、DVD以及我们常见的WAV文件中均有应用。它主要包括三步：抽样、量化、编码。

1. 抽样：对模拟信号周期性扫描，把时间上连续的信号变成时间上离散的信号。
为了使所得的离散信号能无失真地代表被抽样的模拟数据，要使用采样定理进行采样： $f_{采样频率} \geq 2f_{信号最高频率}$ 因为一个周期内至少再需2点才能确定波形
2. 量化：把抽样取得的电平幅值按照一定的分级标度转化为对应的数字值，并取整数，这就把连续的电平幅值转换为离散的数字量。
3. 编码：把量化的结果转换为与之对应的二进制编码。

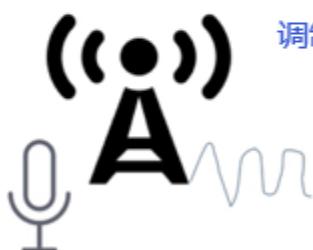


看量化后得到多少种状态的信号（码元），从而决定一个码元有多少比特位。

④ 模拟数据 调制为 模拟信号

模拟数据调制为模拟信号

为了实现传输的有效性，可能需要较高的频率。这种调制方式还可以使用频分复用技术，充分利用带宽资源。在电话机和本地交换机所传输的信号是采用模拟信号传输模拟数据的方式；模拟的声音数据是加载到模拟的载波信号中传输的。



调制为高频信号，抗衰减



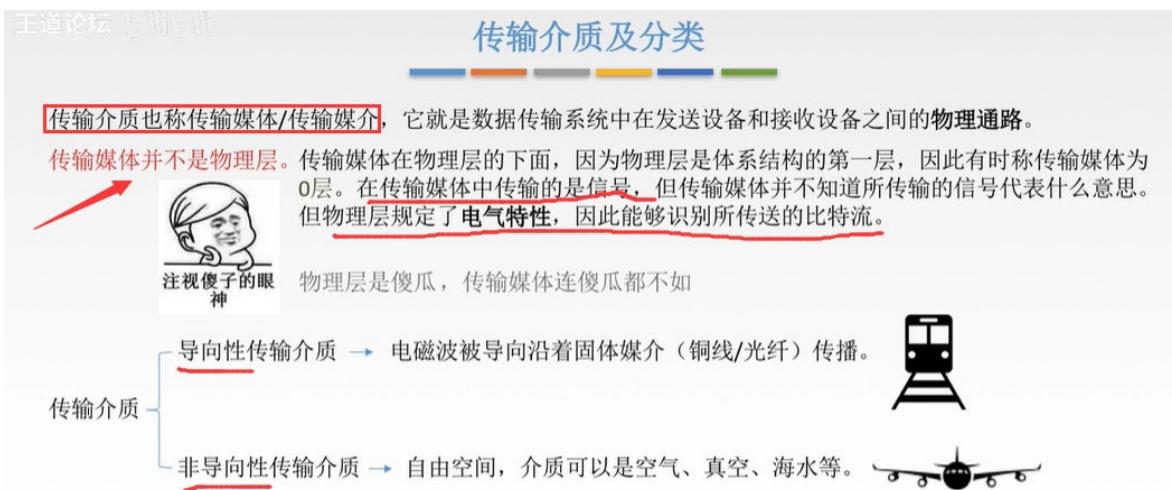
解调为基带信号传给信宿

此时是广播电台传输声音，均是模拟信号



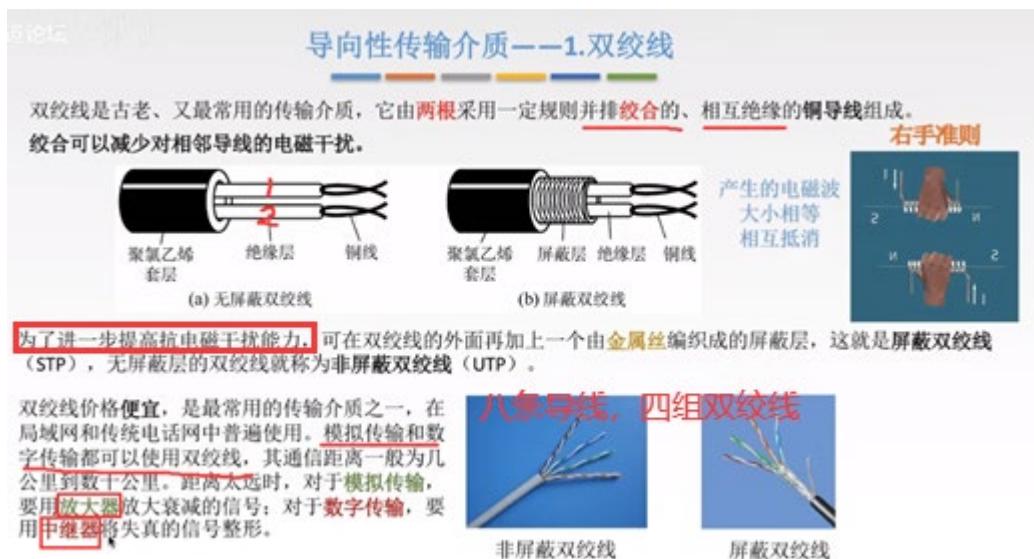
7. 物理层传输介质

① 概念及分类（传输媒体不是物理层）



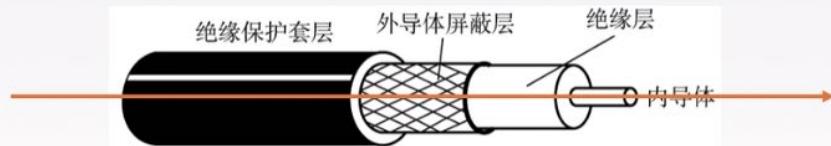
② 导向性传输介质

双绞线中两根导线电流方向相反，产生的电磁场相互抵消，避免对旁边导线造成影响



导向性传输介质——2 同轴电缆

同轴电缆由**导体铜质芯线**、**绝缘层**、网状编织屏蔽层和**塑料外层**构成。按特性阻抗数值的不同，通常将同轴电缆分为两类：50Ω同轴电缆和75Ω同轴电缆。其中，50Ω同轴电缆主要用于传送基带数字信号，又称为**基带同轴电缆**，它在局域网中得到广泛应用；75Ω同轴电缆主要用于传送宽带信号，又称为**宽带同轴电缆**，它主要用于有线电视系统。



同轴电缆Vs双绞线

由于外导体屏蔽层的作用，同轴电缆**抗干扰特性**比双绞线好，被广泛用于传输较高速率的数据，其**传输距离更远**，但**价格**较双绞线贵。



以上两种均是传输电脉冲(电磁波)，下面光纤传输光脉冲(光波)。

导向性传输介质——3. 光纤

光纤通信就是利用**光导纤维**(简称**光纤**)传递**光脉冲**来进行通信。有光脉冲表示1，无光脉冲表示0。而可见光的频率大约是 10^8 MHz，因此光纤通信系统的**带宽远远大于**目前其他各种传输媒体的带宽。

光纤在发送端有光源，可以采用**发光二极管或半导体激光器**，它们在电脉冲作用下能产生出光脉冲；在接收端用**光电二极管**做成光检测器，在检测到光脉冲时可还原出电脉冲。

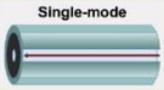
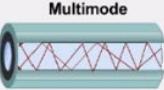
光纤主要由**纤芯(实心的！)**和**包层**构成，光波通过纤芯进行传导，包层较纤芯有较低的折射率。当光线从高折射率的介质射向低折射率的介质时，其折射角将大于入射角。因此，如果入射角足够大，就会出现**全反射**，即光线碰到包层时候就会折射回纤芯、这个过程不断重复，光也就沿着光纤传输下去。



光波的带宽非常大，则光纤的通信量非常大。

导向性传输介质——3.光纤



	定义	光源	特点	外观
单模光纤	一种在 横向模式 直接传输光信号的光纤	定向性很好 的激光二极管	衰耗 小 ，适合 远 距离传输	 Single-mode
多模光纤	有 多种 传输光信号模式的光纤	发光二极管	易 失真 ，适合 近 距离传输	 Multimode



一根光缆少则只有一根光纤，多则包括十至数百根光纤。

光纤的特点：

1. 传输损耗小，中继距离长，对远距离传输特别经济。
2. 抗雷电和电磁干扰性能好。
3. 无串音干扰，保密性好，也不易被窃听或截取数据。
4. 体积小，重量轻。

③非导向性传输介质

非导向性传输介质



无线电波：较强穿透能力，可传远距离，广泛用于通信领域（如手机通信）。

信号向所有方向传播

非导向性传输介质

微波

信号固定方向传播

微波通信频率较高、频段范围宽，因此数据率很高。

微波通信两个例子

地面微波接力通信



优点

- 1、通信容量大
- 2、距离远
- 3、覆盖广
- 4、广播通信和多址通信

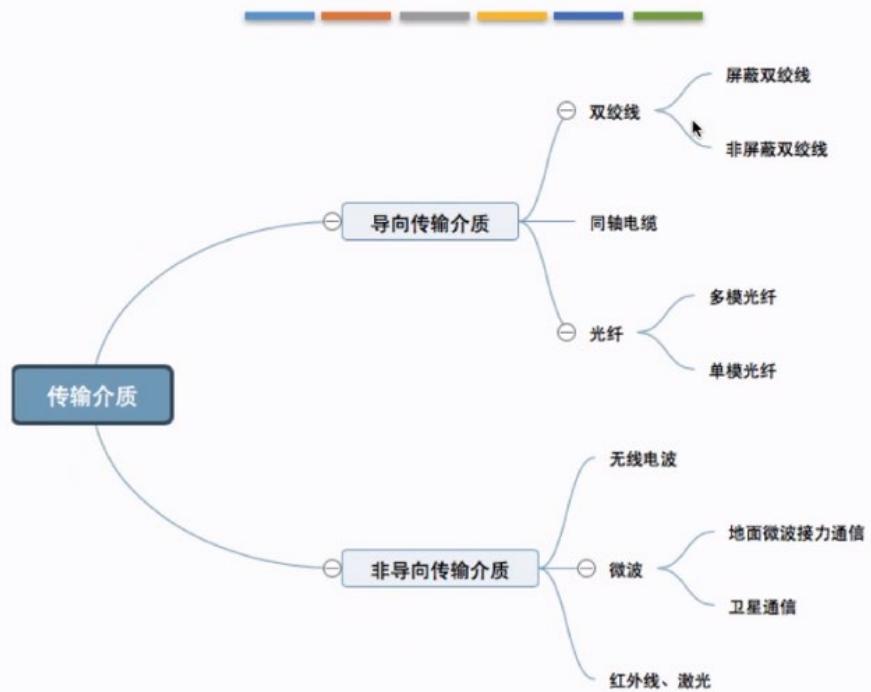
缺点

- 1、传播时延长（250-270ms）
- 2、受气候影响大（eg: 强风 太阳黑子爆发、日凌）
- 3、误码率较高
- 4、成本高

红外线、激光：把要传输的信号分别转换为各自的信号格式，即红外光信号和激光信号，再在空间中传播。

信号固定方向传播

脑图时刻



8. 物理层设备

① 中继器

先重新整形，再还原，以实现放大的效果。所以叫再生，而不叫放大。

如果中继器个数过多，会增加网络延迟。5-4-3 表示：只能最多不超过 5 个网段，并且其中最多只能有 4 个物理层网络设备，并且只有 3 个段可以挂接计算机(工作站)

不会存储转发!!!

中继器 针对数字信号



诞生原因: 由于存在损耗，在线路上传输的信号功率会逐渐衰减，衰减到一定程度时将造成信号失真，因此会导致接收错误。

中继器的功能: 对信号进行**再生和还原**，对衰减的信号进行放大，保持与原数据相同，以增加信号传输的距离，延长网络的长度。



中继器的两端: ① 两端的网络部分是网段，而不是子网，适用于完全相同的**两类**网络的互连，且两个网段速率要相同。

② 中继器只将任何电缆段上的数据发送到另一段电缆上，它仅作用于信号的电气部分，并不管数据中是否有错误数据或不适用于网段的数据。

③ 两端可连相同媒体，也可连**不同媒体**。

④ 中继器两端的网段**一定要是同一个协议**。（中继器**不会存储转发**，傻）

5-4-3规则: 网络标准中都对信号的延迟范围作了具体的规定，因而中继器只能在规定的范围内进行，否则会网络故障。



②集线器（多口中继器）

集线器（多口中继器）



再生，放大信号

集线器的功能: 对信号进行**再生放大转发**，对衰减的信号进行放大，接着转发到其他所有（除输入端口外）处于工作状态的端口上，以增加信号传输的距离，延长网络的长度。**不具备信号的定向传送能力，是一个共享式设备。**

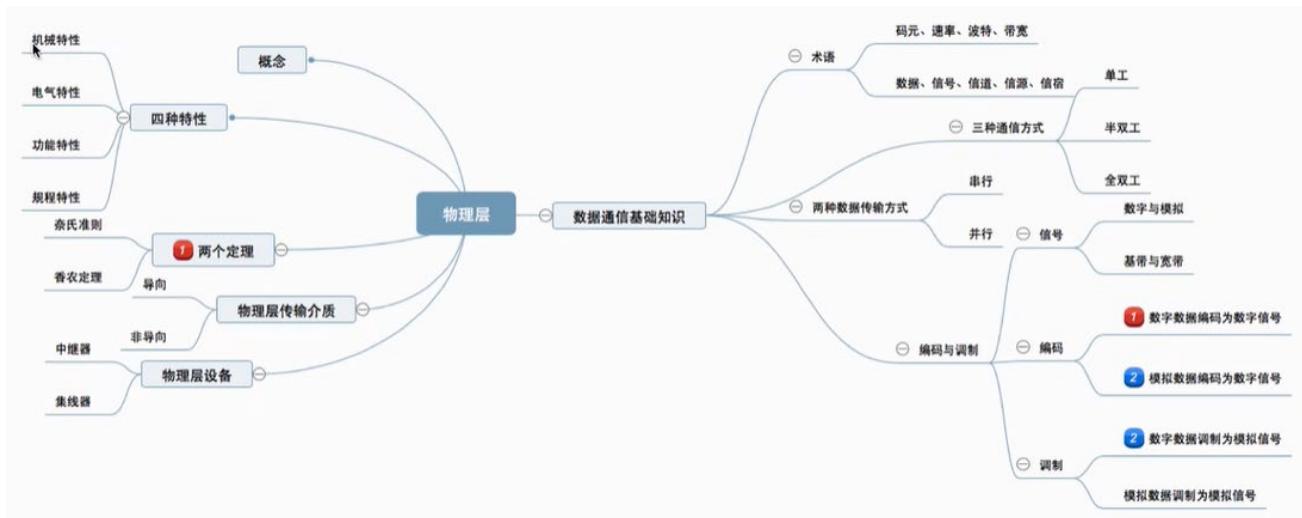
以广播的形式转发数据



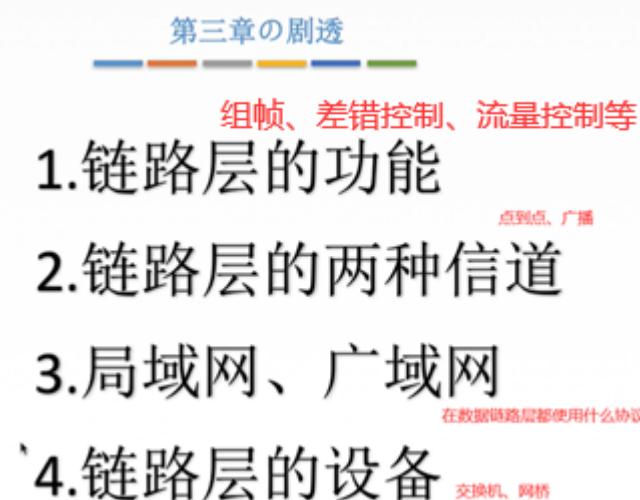
特点: 不能分割冲突域：若有两台计算机同时发送数据到集线器，会导致冲突。此时会停下，等一个随机时间重新发送，直到不冲突。所以集线器在一个时钟周期内只能传输一组信息（实现一组的通信），其他的不能保证。

要实现同时通信不发生碰撞，就需要**平分带宽**。

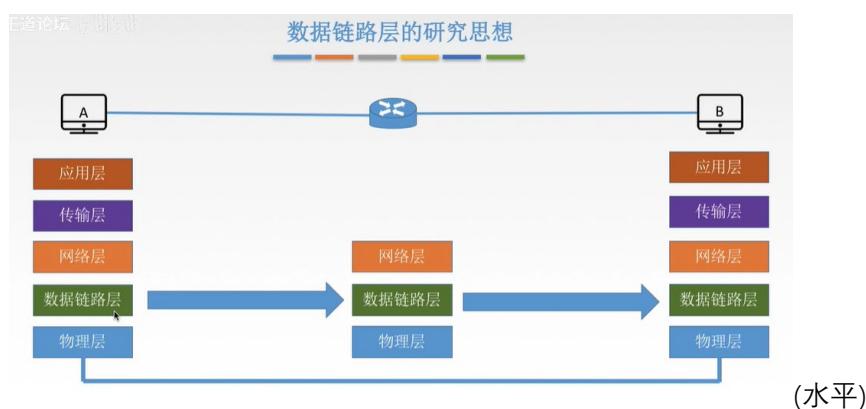
9. 第二章总结



三、数据链路层



3.1 数据链路层的功能概述



数据链路层基本概念

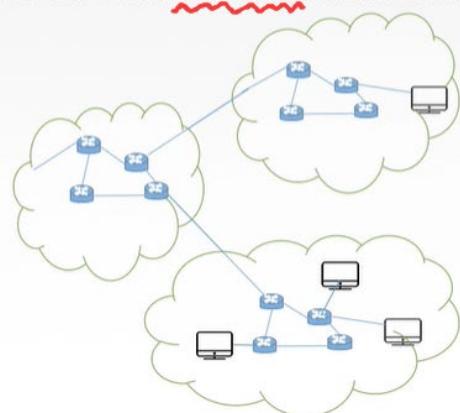
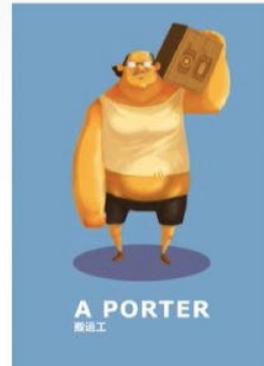
结点：主机、路由器

链路：网络中两个结点之间的物理通道，链路的传输介质主要有双绞线、光纤和微波。**分为有线链路、无线链路。**

数据链路：网络中两个结点之间的逻辑通道，把实现控制数据传输协议的硬件和软件加到链路上就构成数据链路。

帧：链路层的协议数据单元，封装网络层数据报。

数据链路层负责通过一条链路从一个结点向另一个物理链路直接相连的相邻结点传送数据报。



王道考研/CSKAOKYAN.COM

数据链路层功能概述

数据链路层在物理层提供服务的基础上向网络层提供服务，其最基本的服务是将源自网络层来的数据可靠地传输到相邻节点的目标机网络层。其主要作用是加强物理层传输原始比特流的功能，将物理层提供的可能出错的物理连接改造成为逻辑上无差错的数据链路，使之对网络层表现为一条无差错的链路。

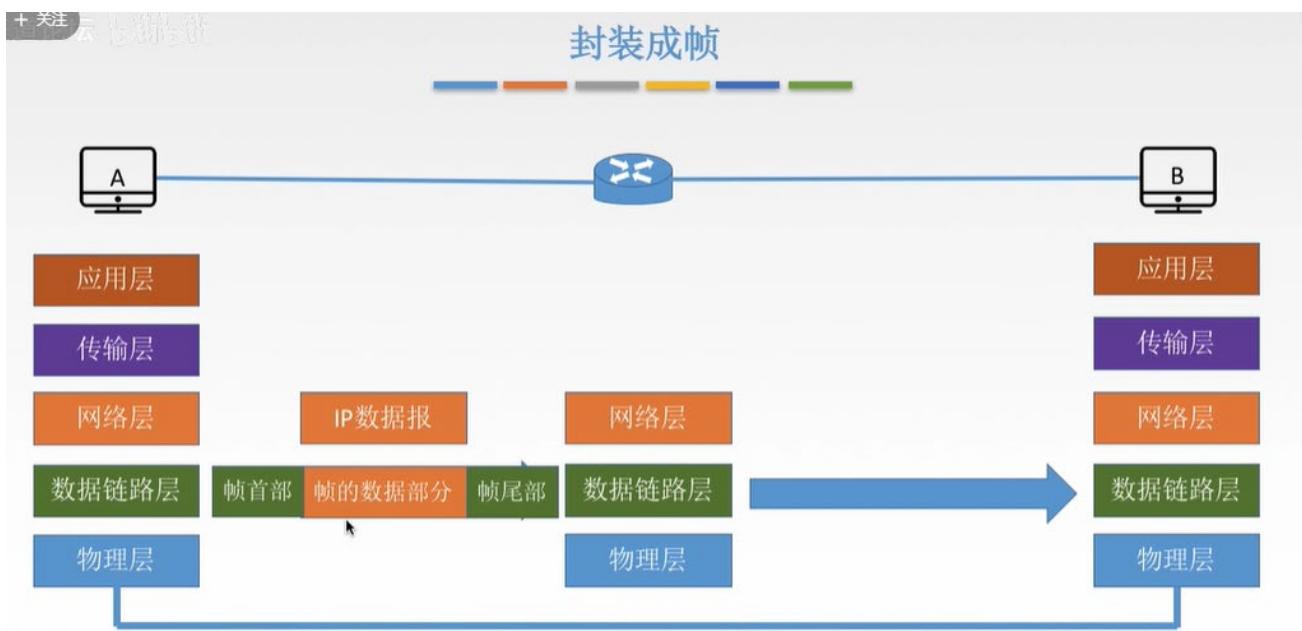


即，将网络层交给的数据报，进行加工和处理后递交给物理层传输。

有确认：每接收到一个帧，都要返回一个确认帧。发送端在哪个帧未收到确认帧，则重发该帧。

功能一：为网络层提供服务。无确认无连接服务，有确认无连接服务，有确认面向连接服务。有连接一定有确认！
功能二：链路管理，即连接的建立、维持、释放（用于面向连接的服务）。
功能三：组帧。
功能四：流量控制  限制发送方哦~
功能五：差错控制（帧错/位错）。

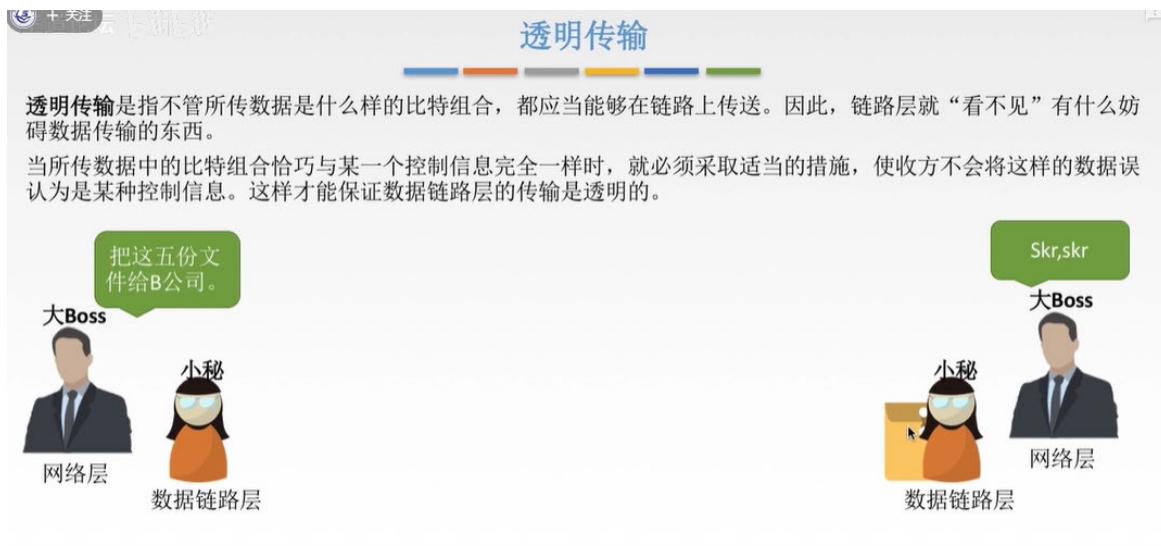
3.2 功能 1-封装成帧和透明传输



组帧：在发送端；帧同步：在接收端。



数据链路层不管是什么数据都会封装传输。



① 字符计数法



痛点：鸡蛋装在一个篮子里了。 前面错了后面都会错

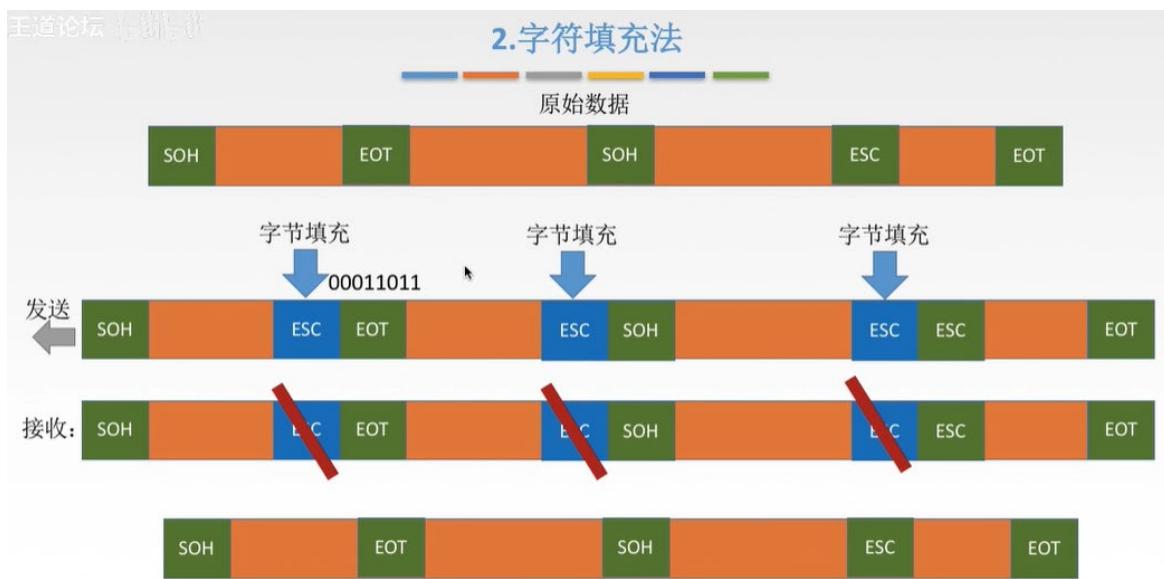
② 字符(节)填充法



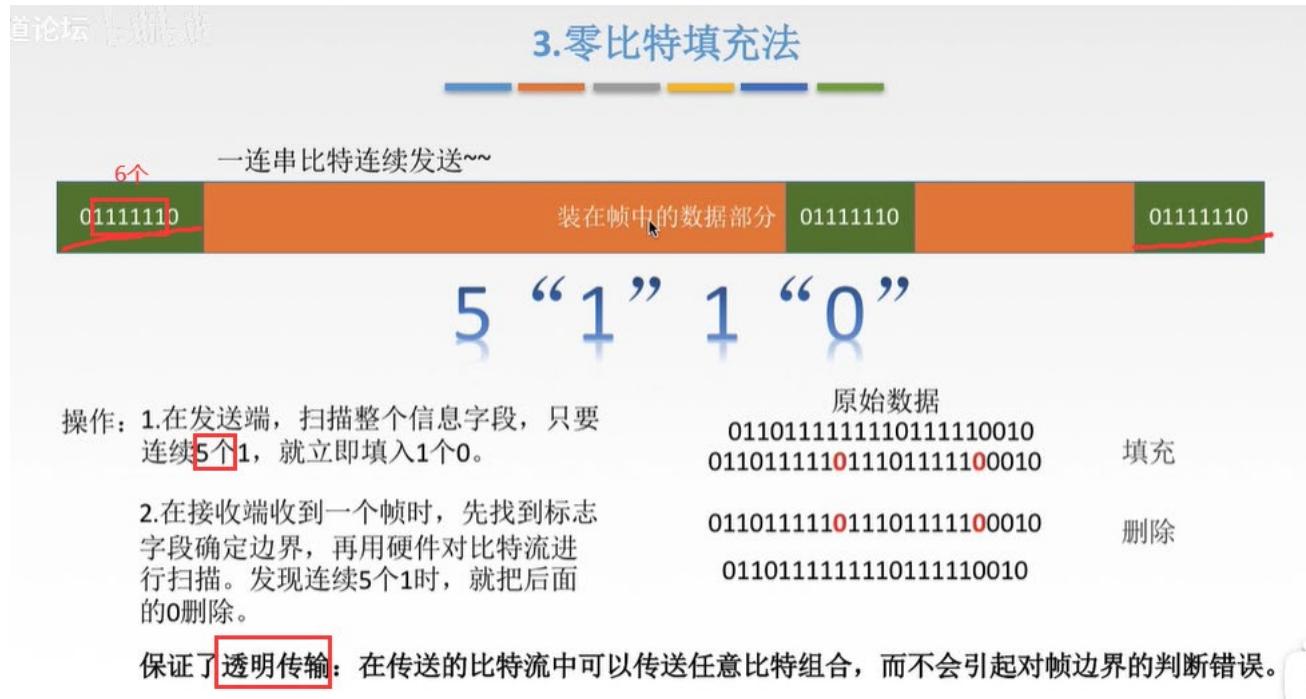
当传送的帧是由文本文件组成时（文本文件的字符都是从键盘上输入的，都是ASCII码）。不管从键盘上输入什么字符都可以放在帧里传过去，即透明传输。



当传送的帧是由非ASCII码的文本文件组成时（二进制代码的程序或图像等）。就要采用字符填充方法实现透明传输。

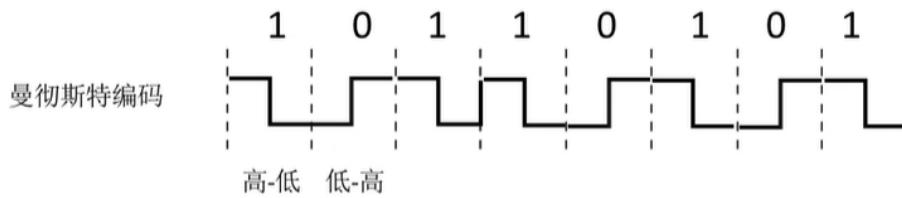


③ 零比特填充法



④ 违规编码法

4. 违规编码法



可以用“高-高”，“低-低”来定界帧的起始和终止。

由于字节计数法中Count字段的脆弱性（其值若有差错将导致灾难性后果）及字符填充实现上的复杂性和不兼容性，目前较普遍使用的帧同步法是比特填充和违规编码法。

3.3.1 功能 2-差错控制(检错编码)

① 差错来源

差错从何而来？

概括来说，传输中的差错都是由于噪声引起的。

全局性 1. 由于线路本身电气特性所产生的随机噪声(热噪声)，是信道固有的，随机存在的。

解决办法：提高信噪比来减少或避免干扰。(对传感器下手)

局部性 2. 外界特定的短暂原因所造成的冲击噪声，是产生差错的主要原因。

解决办法：通常利用编码技术来解决。

位错 【比特位出错，1变成0，0变成1。】

差错

丢失：收到[#1]-[#3] 以前(过去)OSI模型中使用的是帧编号，确认重传机制等实现差错控制
现在针对不同的网络会选择是否采用确认重传机制。

重复：收到[#1]-[#2]-[#2]-[#3]

失序：收到[#1]-[#3]-[#2]

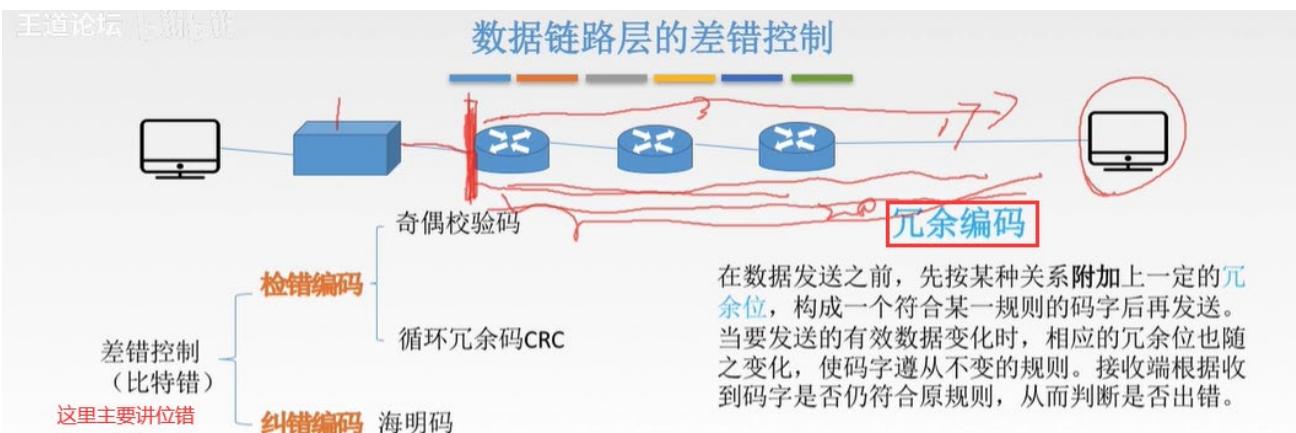
链路层为网络层提供服务：无确认无连接服务，有确认面向连接服务。

通信质量好
有线传输链路

→通信质量差的无线传输链路

② 链路层为什么要有差错控制？

在中间系统中(如路由器)的数据链路层如果发现有错误，就会丢弃该帧，以免继续传输耗费资源(尽早发现错误尽早解决)。



编码 VS 编码

数据链路层编码和物理层的数据编码与调制不同。物理层编码针对的是单个比特，解决传输过程中比特的同步等问题，如曼彻斯特编码。而数据链路层的编码针对的是一组比特，它通过冗余码的技术实现一组二进制比特串在传输过程是否出现了差错。

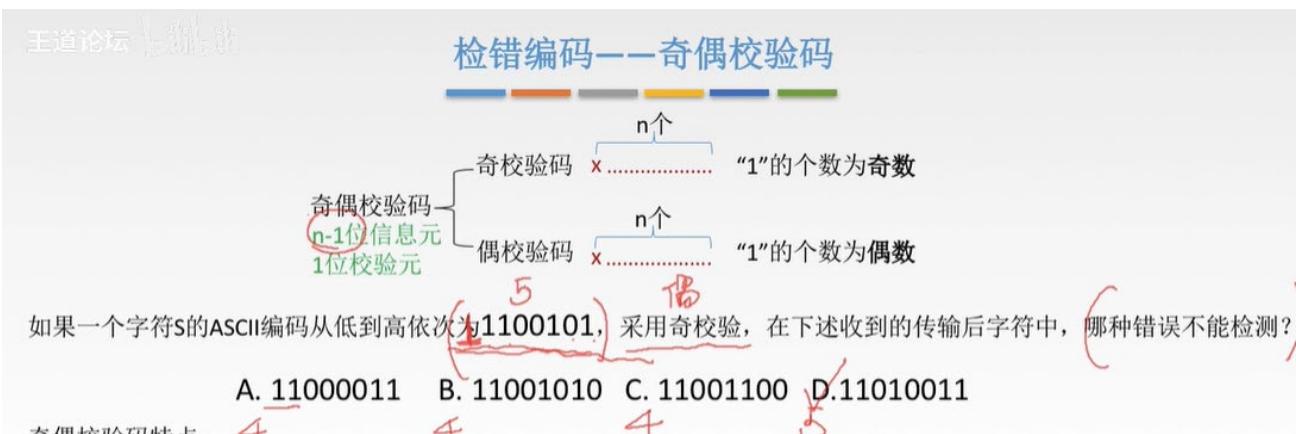


只有五本？拒收！



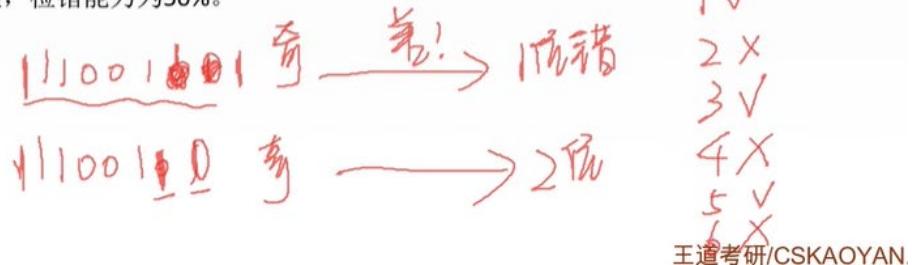
③奇偶校验码

信息元：要发送的有效数据；校验元：奇偶校验码中的冗余码



奇偶校验码特点：

只能检查出奇数个比特错误，检错能力为50%。



即发送时奇校验码有奇数个1，接收到的字符中有偶数个1了（接收时只知道是奇校验码，也并不知道具体奇数是多少），那么此时能检查出发生了错误。

⑤ CRC 循环冗余码

检错编码——CRC循环冗余码

发送端

$$\text{要传的数据} \quad \begin{matrix} \text{生成多项式} \\ 5 \quad \div \quad 2 \quad = \quad 2 \end{matrix} \quad \dots \dots \quad 1$$

最终发送数据: 5+1=6

接收端

$$\text{接收到的数据} \quad \begin{matrix} \text{生成多项式} \\ 6 \quad \div \quad 2 \quad = \quad 3 \quad \dots \dots \quad 0 \end{matrix}$$

余数为0, 判定无错, 就接受。

例: 要发送的数据是1101 0110 11, 采用CRC校验, 生成多项式是10011, 那么最终发送的数据应该是?

最终发送的数据:

要发送的数据+帧检验序列FCS

计算冗余码:

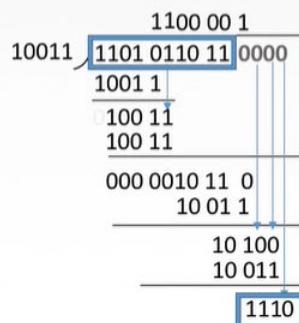
- (1) 加0 假设生成多项式G(x)的阶为r, 则加r个0。
- (2) 模2除法 数据加0后除以多项式, 余数为冗余码/FCS/CRC检验码的比特序列。

10011表示成多项式为

$$\begin{aligned} X^4 + X^3 + X^0 &= 1 \cdot X^4 + 0 \cdot X^3 + 0 \cdot X^2 + 1 \cdot X^1 + X^0 \\ &= X^4 + X^1 + 1 \end{aligned}$$

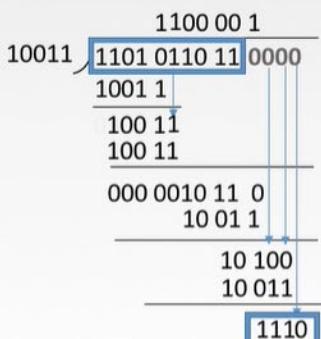
阶为4 1 0 0 1 1

TIPS: 多项式N位, 阶为N-1。



最终发送数据
1101011011110

王道考研/CSKAOYAN.COM



最终发送的数据: 1101011011110

接收端检错过程

把收到的每一个帧都除以同样的除数, 然后检查得到的余数R。

- 1.余数为0, 判定这个帧没有差错, 接受。
- 2.余数为不为0, 判定这个帧有差错 (无法确定到位), 丢弃。

FCS的生成以及接收端CRC检验都是由硬件实现, 处理很迅速, 因此不会延误数据的传输。



+ 关注

检错编码——CRC循环冗余码

在数据链路层仅仅使用循环冗余检验CRC差错检测技术, 只能做到对帧的无差错接收, 即“凡是接收端数据链路层接受的帧, 我们都能以非常接近于1的概率认为这些帧在传输过程中没有产生差错”。接收端丢弃的帧虽然曾收到了, 但是最终还是因为有差错被丢弃。“凡是接收端数据链路层接收的帧均无差错”。

“可靠传输”: 数据链路层发送端发送什么, 接收端就收到什么。无差错, 不丢失, 不重复, 有序到链路层使用CRC检验, 能够实现无比特差错的传输, 但这还不是可靠传输。

无比特差错传输, 不是可靠传输, 可能还会有帧丢失、重复和失序等或者说检测到差错帧后直接丢弃, 差错纠正由高层负者。

3.3.2 ※※※功能 2-差错控制(纠错编码)

纠错编码——海明码：可以发现双比特错(两位出错)，但只能纠正单比特错(一位出错)

纠错编码——海明码

海明码：发现双比特错，纠正单比特错。

工作原理：动一发而牵全身



有多个校验码或冗余码，不光可以校验自身，还可以校验其他几位比特。所以有些比特位(数据)可以同时被几位校验码作用。所以这种比特位(数据)发生错误时，可以通过多个校验码共同确定是哪一位发生了错误。

工作流程：



1. 确定校验码位数r

海明不等式

$$2^r \geq k+r+1$$

r为冗余信息位，k为信息位

要发送的数据：**D=101101**

数据的位数k=6，
满足不等式的最小r为4，
也就是D=101101的海明码应该有6+4=10位，
其中原数据6位，效验码4位。

2. 确定校验码和数据的位置

D=101101

假设这4位校验码分别为P₁、P₂、P₃、P₄；数据从左到右为D₁、D₂、……、D₆。
放在2的几次方的位置 按序把空填满

数据位	1	2	3	4	5	6	7	8	9	10
代码	P ₁	P ₂	D ₁	P ₃	D ₂	D ₃	D ₄	P ₄	D ₅	D ₆
实际值			1		0	1	1		0	1

3.求出校验码的值

D=101101

二进制	00010	00100	00110	01000	01010	01100	01110	10000	10010	10100
数据位	1	2	3	4	5	6	7	8	9	10
代码	P ₁	P ₂	D ₁	P ₃	D ₂	D ₃	D ₄	P ₄	D ₅	D ₆
实际值	0	0	1	0	0	1	1	1	0	1

令所有要校验的位异或=0。

$$P_1 \oplus D_1 \oplus D_2 \oplus D_4 \oplus D_5 = 0 \Rightarrow P_1 = 0$$

$$P_2 \oplus D_1 \oplus D_3 \oplus D_4 \oplus D_6 = 0 \Rightarrow P_2 = 0$$

$$P_3 \oplus D_2 \oplus D_3 \oplus D_4 = 0 \Rightarrow P_3 = 0$$

$$P_4 \oplus D_5 \oplus D_6 = 0 \Rightarrow P_4 = 1$$

故101101的海明码为00100111101。

所有包含与校验位二进制相同的1的位(包括校验位自身),都是该校验位要校验的位。

4.检错并纠错

D=101101

数据位	1	2	3	4	5	6	7	8	9	10
代码	P ₁	P ₂	D ₁	P ₃	D ₂	D ₃	D ₄	P ₄	D ₅	D ₆
实际值	0	0	1	0	1	1	1	1	0	1

故101101的海明码为00100111101。

假设第五位出错,因此接收到的数据位0010111101。

令所有要校验的位异或运算。

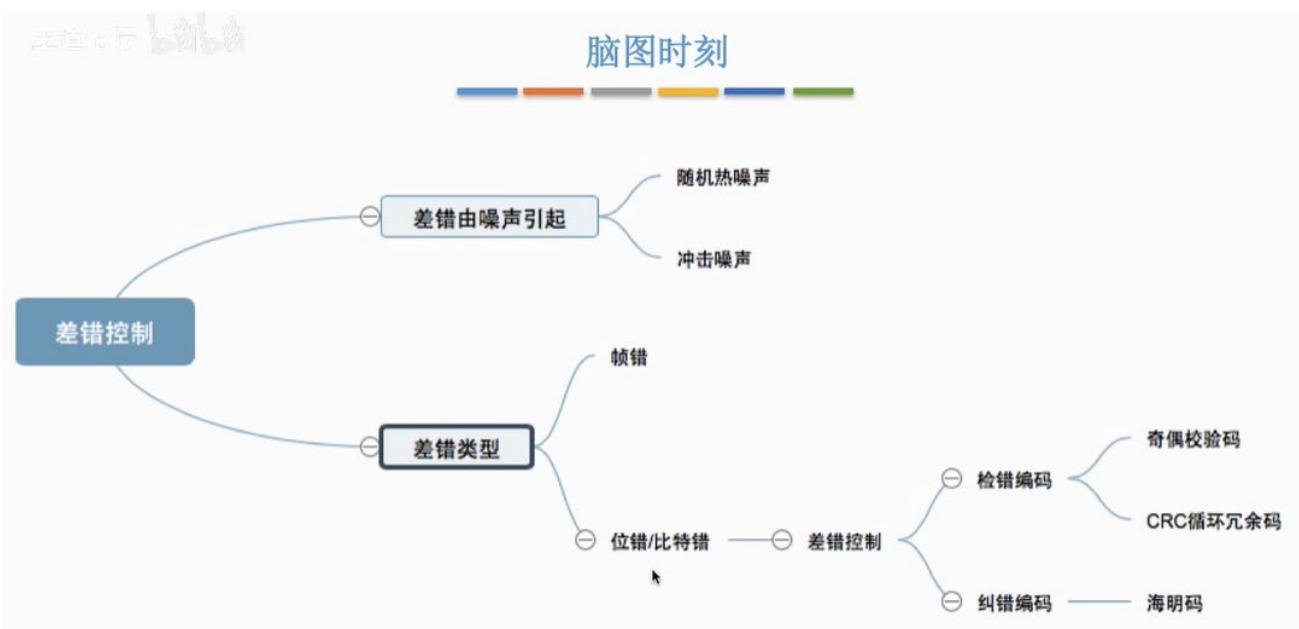
$$P_1 \oplus D_1 \oplus D_2 \oplus D_4 \oplus D_5 = 1 \quad P_4 \oplus D_5 \oplus D_6 = 0$$

$$P_2 \oplus D_1 \oplus D_3 \oplus D_4 \oplus D_6 = 0$$

$$P_3 \oplus D_2 \oplus D_3 \oplus D_4 = 1$$

$$\underline{0 \ 1 \ 0 \ 1} \rightarrow 5$$

二进制序列为0101,恰好对应十进制5,这样就找到了出错的位置,即出错位是第5位。



3.4.1 功能 3-流量控制与可靠传输机制

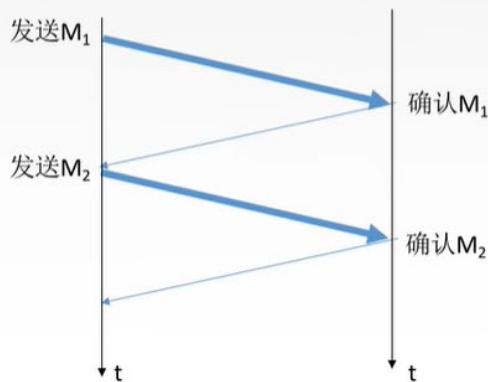
链路层和传输层都有流量控制。



流量控制的方法

停止-等待协议

每发送完一个帧就停止发送，等待对方的确认，在收到确认后再发送下一个帧。



滑动窗口协议

后退N帧协议 (GBN)

选择重传协议 (SR)



流量控制的方法

停止-等待协议

发送窗口大小=1，接收窗口大小=1；

后退N帧协议 (GBN) 发送窗口大小>1，接收窗口大小=1；

选择重传协议 (SR) 发送窗口大小>1，接收窗口大小>1；

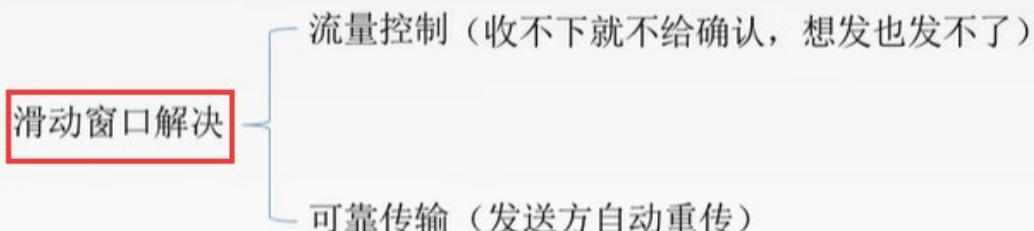
停止-等待协议是特殊的滑动窗口协议。

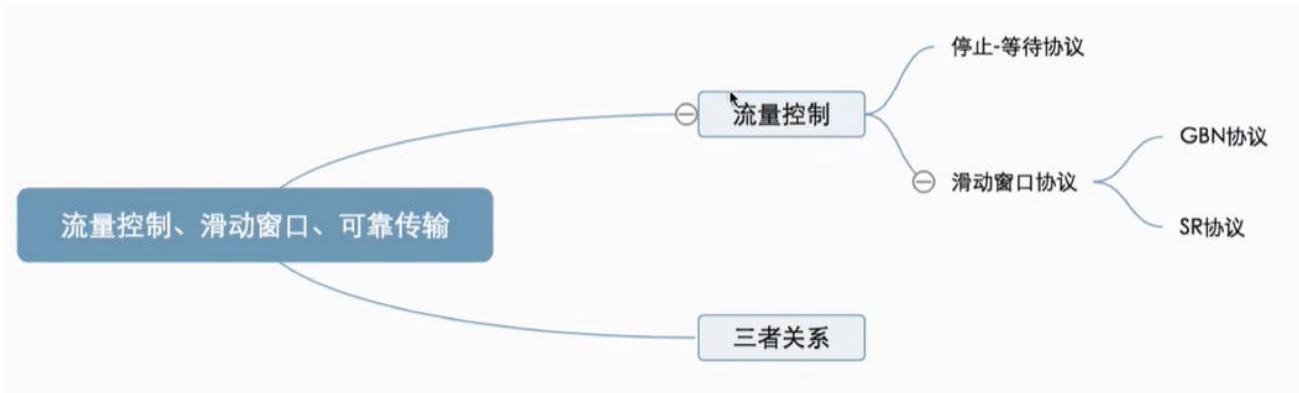
注意：数据链路层的窗口大小在传输过程中都是固定的。

可靠传输、滑动窗口、流量控制

可靠传输：发送端发啥，接收端收啥。

流量控制：控制发送速率，使接收方有足够的缓冲空间来接收每一个帧。





不用纠结以下三个协议具体是属于哪一层的(有些不同的教材划在不同的层)。以前链路质量不是很好时，链路层就需要担负起可靠传输的职责，所以就需要使用这三种协议；而现在链路质量变好了，发生错误的概率变小了，所以保证可靠传输的任务就可以交给传输层，链路层只需保证差错控制即可(这样也可提高数据传输速度)。

协议在不同层次，影响的仅是传输的数据对象。(帧或分组，本质都是要传送的数据)

3.4.2 功能 3 的协议 1-停止等待协议

王道论坛

停止-等待协议

1.为什么要有停止-等待协议？

除了比特出差错，底层信道还会出现丢包问题。
为了实现流量控制。

丢包：物理线路故障、设备故障、病毒攻击、路由信息错误等原因，会导致数据包的丢失。

2.研究停等协议的前提？

虽然现在常用全双工通信方式，但为了讨论问题方便，仅考虑一方发送数据（发送方），一方接收数据（接收方）。

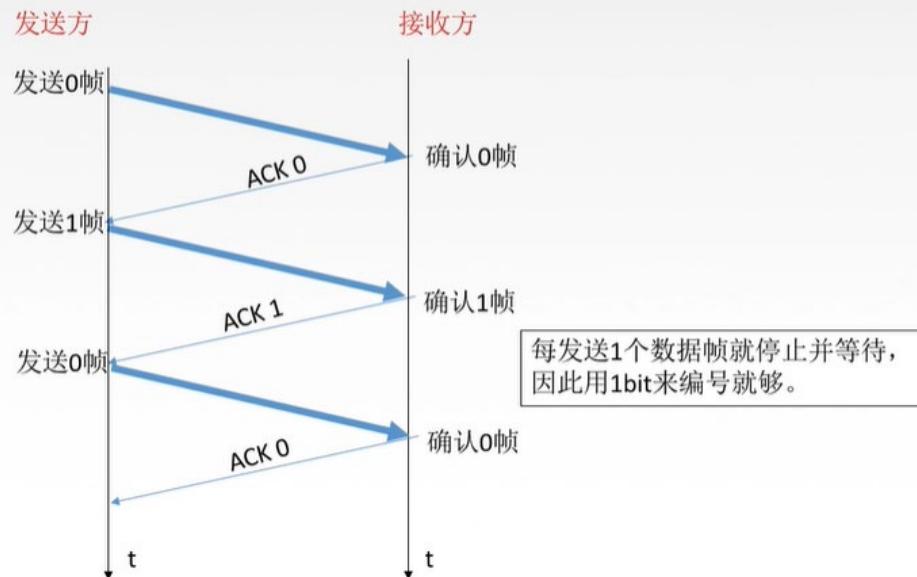
因为是在讨论可靠传输的原理，所以并不考虑数据是在哪一个层次上传送的。

“停止-等待”就是每发送完一个分组就停止发送，等待对方确认，在收到确认后再发送下一个分组。

3.停等协议有几种应用情况？

无差错情况&有差错情况
注意：编号是循环重复的。

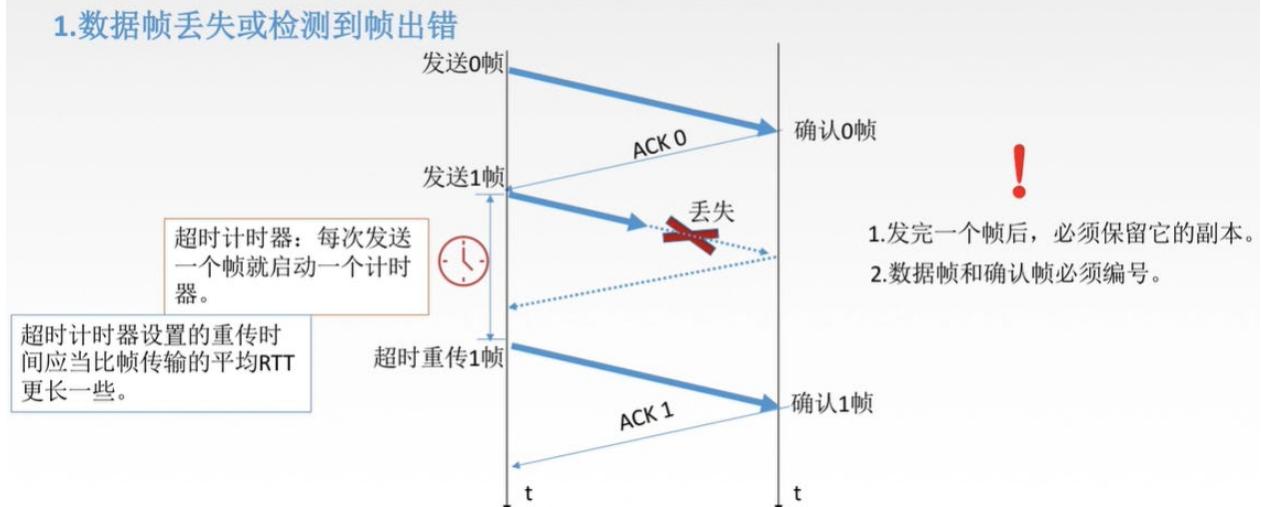
停等协议——无差错情况



超时、重传、确认机制：

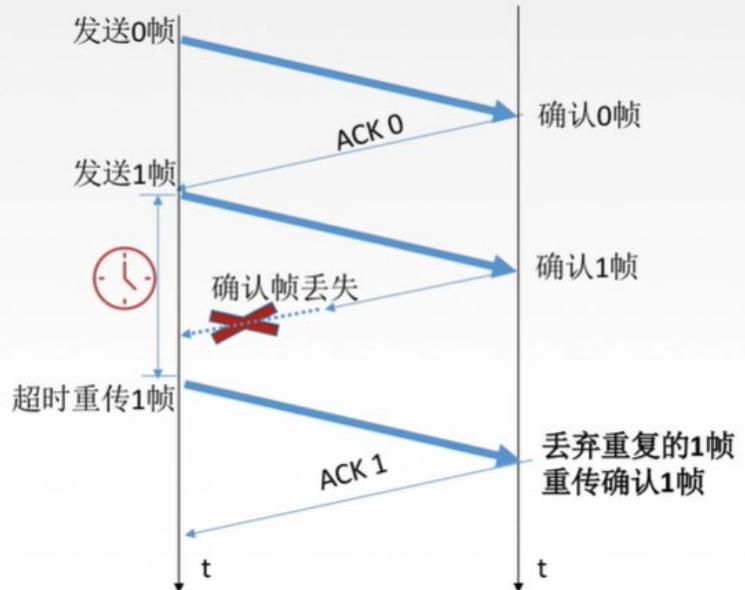
王道论坛

停等协议——有差错情况



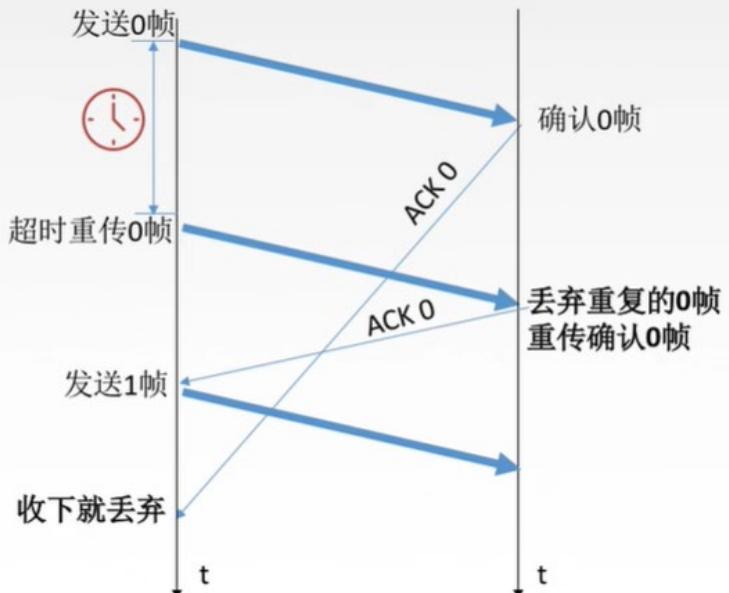
停等协议——有差错情况

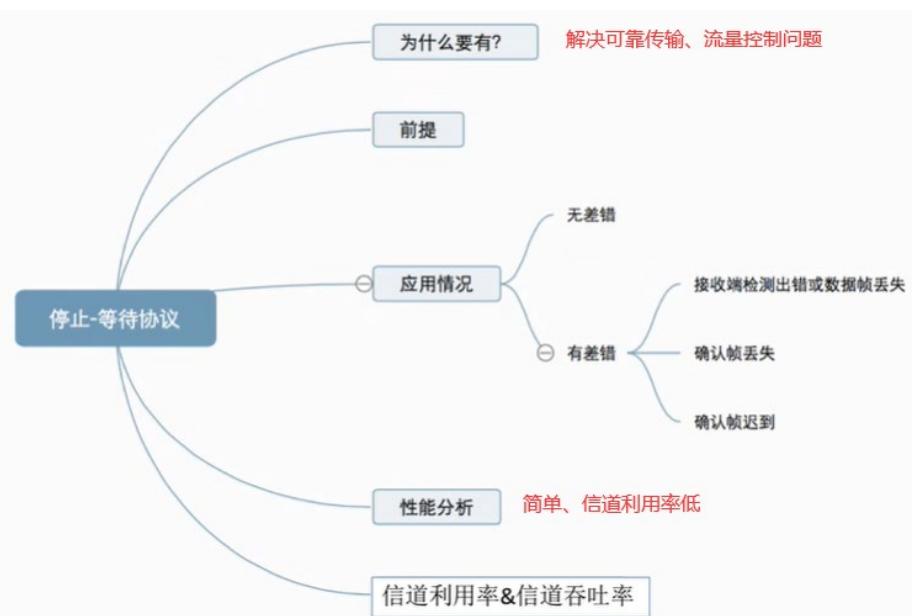
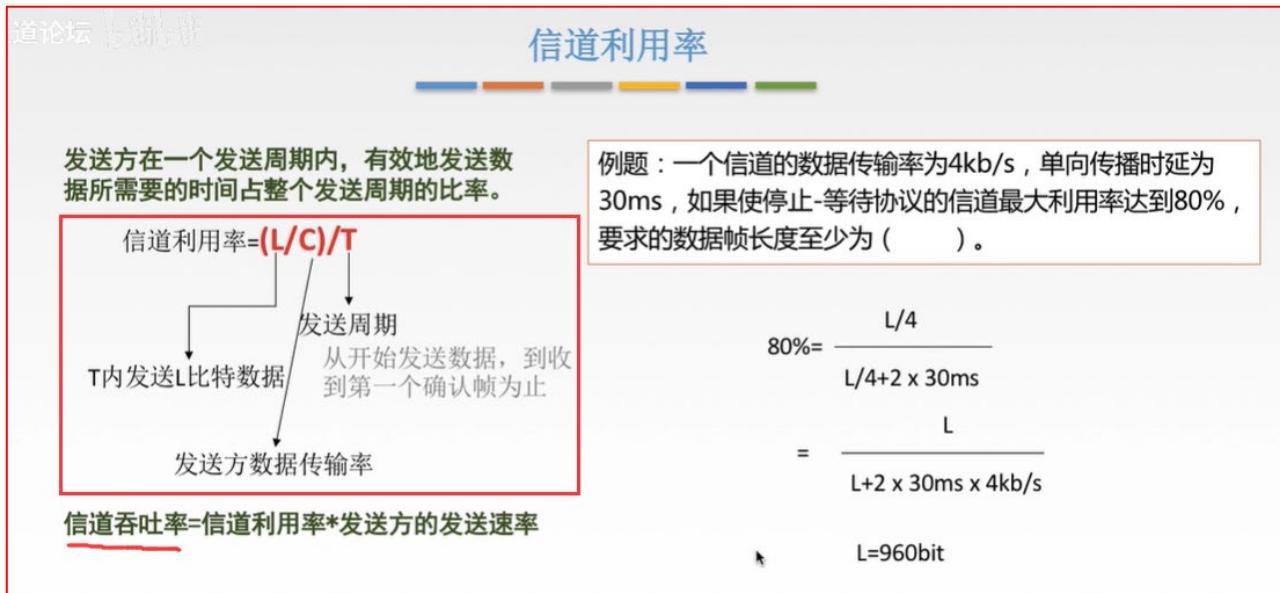
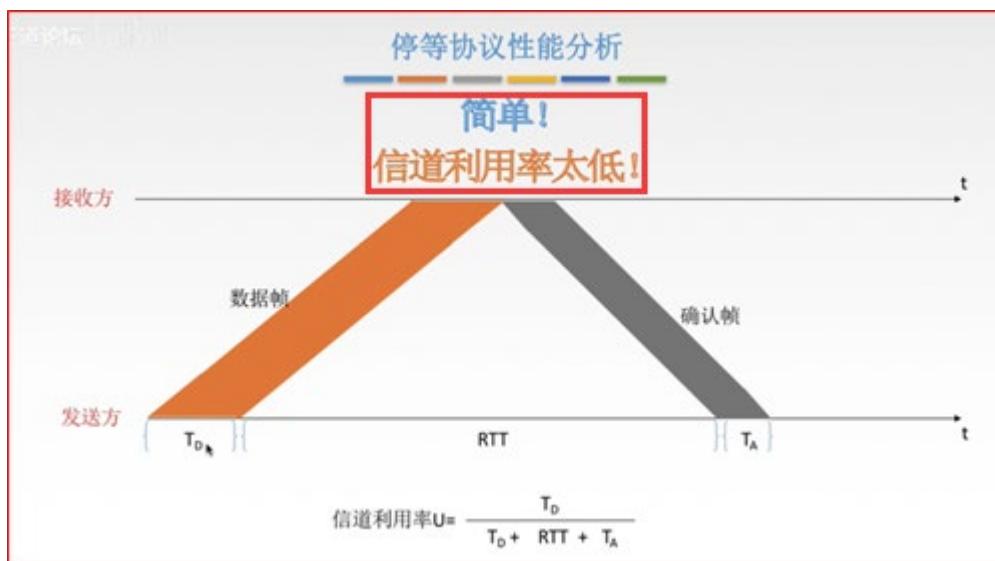
2.ACK丢失



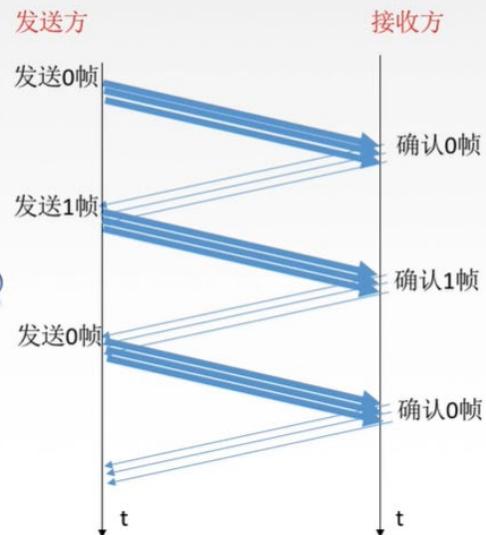
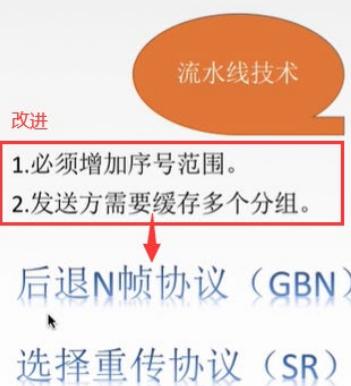
停等协议——有差错情况

3.ACK迟到





停等协议的弊端



太闲了。。

3.4.3 功能3的协议2-后退N帧协议(GBN)

后退N帧协议中的滑动窗口

发送窗口：发送方维持一组连续的允许发送的帧的序号。



GBN发送方必须响应的三件事



1.上层的调用

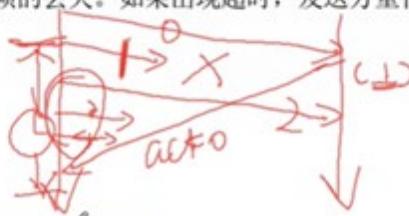
上层要发送数据时，发送方先检查发送窗口是否已满，如果未满，则产生一个帧并将其发送；如果窗口已满，发送方只需将数据返回给上层，暗示上层窗口已满。上层等一会再发送。（实际实现中，发送方可以缓存这些数据，窗口不满时再发送帧）。

2.收到了一个ACK

GBN协议中，对n号帧的确认采用累积确认的方式，标明接收方已经收到n号帧和它之前的全部帧。

3.超时事件

协议的名字为后退N帧/回退N帧，来源于出现丢失和时延过长时发送方的行为。就像在停等协议中一样，定时器将再次用于恢复数据帧或确认帧的丢失。如果出现超时，发送方重传所有已发送但未被确认的帧。



王道考研/CSKAOYAN.COM

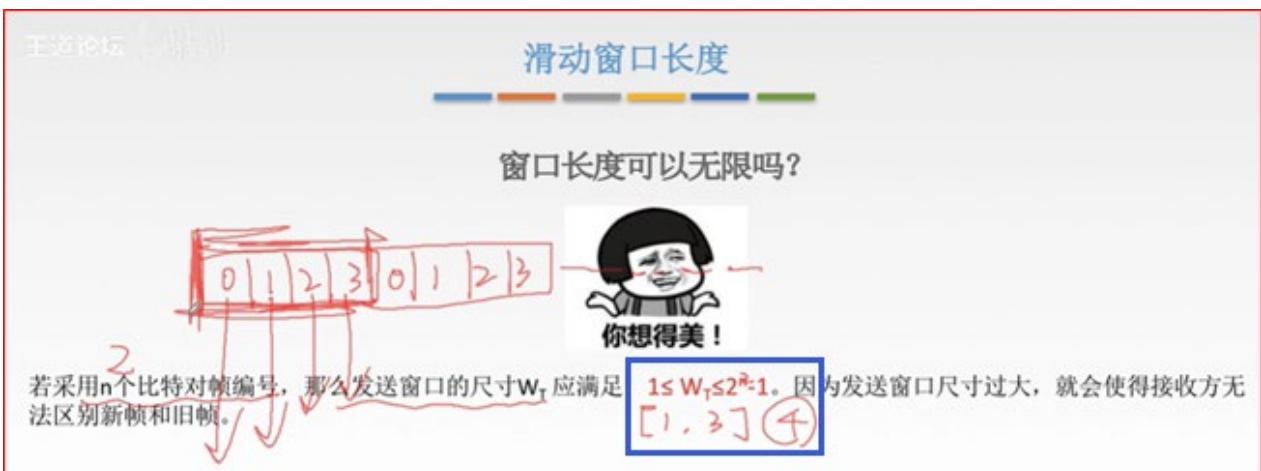
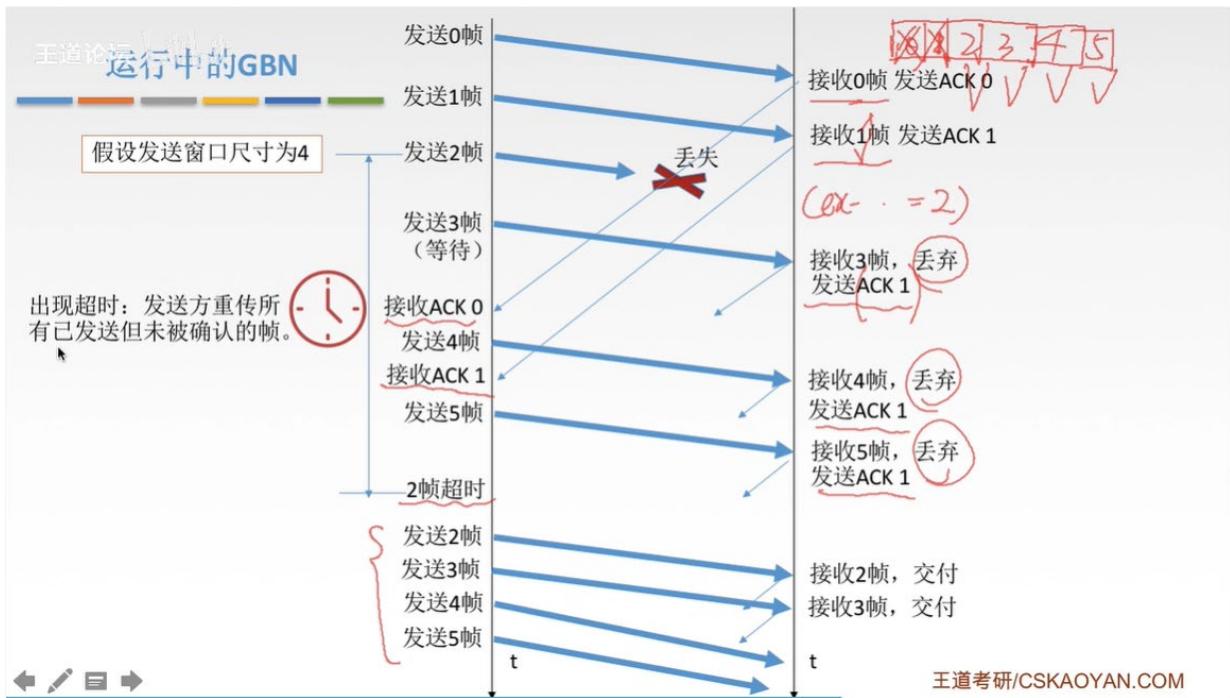
GBN接收方要做的事



如果正确收到n号帧，并且按序，那么接收方为n帧发送一个ACK，并将该帧中的数据部分交付给上层。



其余情况都丢弃帧，并为最近按序接收的帧重新发送ACK。接收方无需缓存任何失序帧，只需要维护一个信息：expectedseqnum（下一个按序接收的帧序号）。



解释：第一批0123接收方收到了，但是其返回的ACK全都丢失了。因为发送方没收到确认帧ACK。所以会把第一批重复发一遍，而接收方此时会误当成第二批0123。



习题1

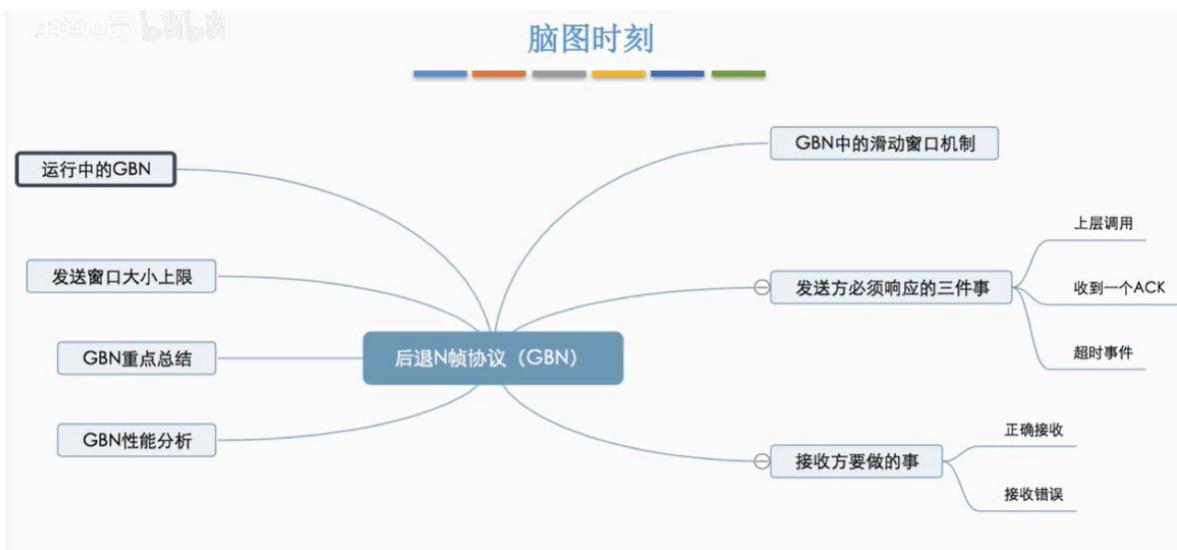
已发未确

数据链路层采用了后退N帧 (GBN) 协议，发送方已经发送了编号为0~7的帧。当计时器超时时，若发送方只收到0、2、3号帧的确认，则发送方需要重发的帧数是（ ）。

A. 2 B. 3 C. 4 D. 5

0~3 ✓

4~7 重新发送4、5、6、7帧



3.4.4 功能3的协议3-选择重传协议(SR, Selective Repeat)



解决办法：设置单个确认，同时加大接收窗口，设置接收缓存，缓存乱序到达的帧。



道论坛 | 讨论区

SR发送方必须响应的三件事

1.上层的调用

从上层收到数据后，SR发送方检查下一个可用于该帧的序号，如果序号位于发送窗口内，则发送数据帧；否则就像GBN一样，要么将数据缓存，要么返回给上层之后再传输。

2.收到了一个ACK

如果收到ACK，加入该帧序号在窗口内，则SR发送方将那个被确认的帧标记为已接收。如果该帧序号是窗口的下界（最左边第一个窗口对应的序号），则窗口向前移动到具有最小序号的未确认帧处。如果窗口移动了并且有序号在窗口内的未发送帧，则发送这些帧。



3.超时事件

每个帧都有自己的定时器，一个超时事件发生后只重传一个帧。

SR接收方要做的事



来者不拒 (窗口内的帧)

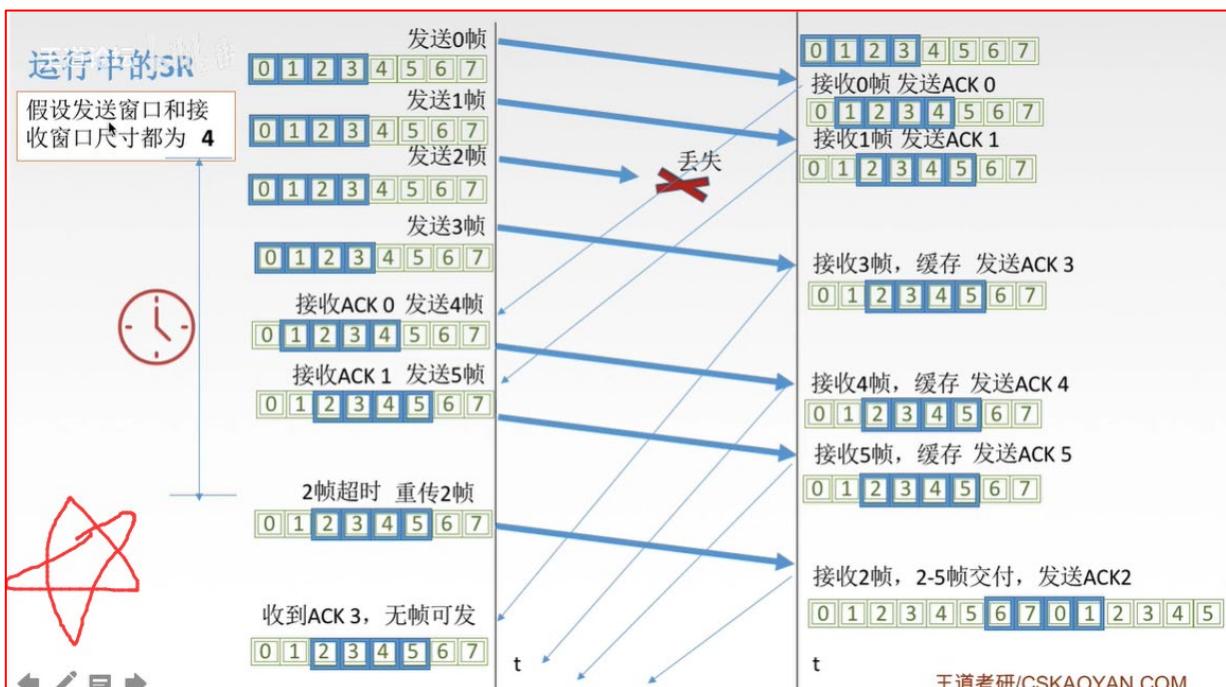
SR接收方将确认一个正确接收的帧而不管其是否按序。失序的帧将被缓存，并返回给发送方一个该帧的确认帧【收谁确认谁】，直到所有帧（即序号更小的帧）皆被收到为止，这时才可以将一批帧按序交付给上层，然后向前移动滑动窗口。



如果收到了窗口序号外（小于窗口下界）的帧，就返回一个ACK。其他情况，就忽略该帧。

运行中的SR

假设发送窗口和接收窗口尺寸都为 4

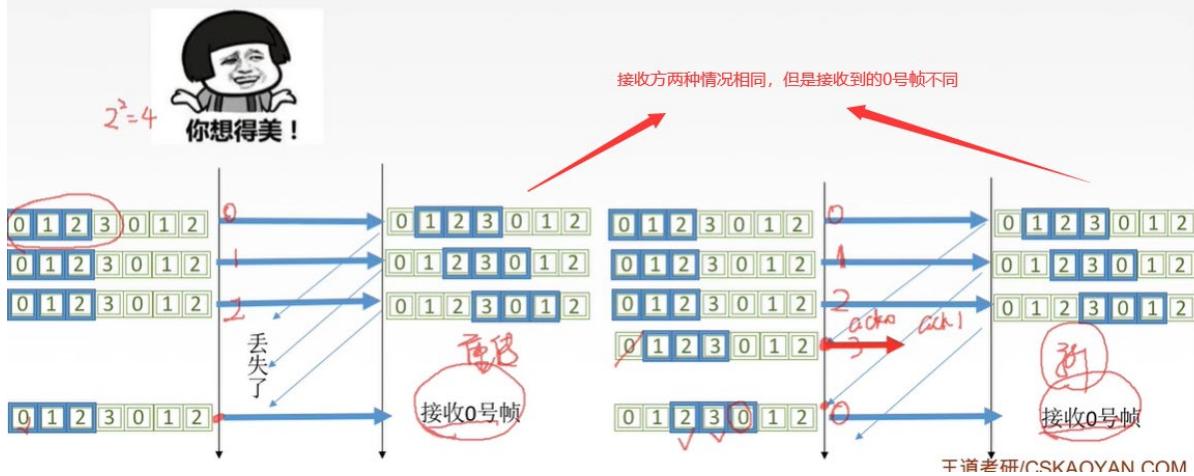


数据链路层滑动窗口示例

滑动窗口长度

窗口长度可以无限吗？

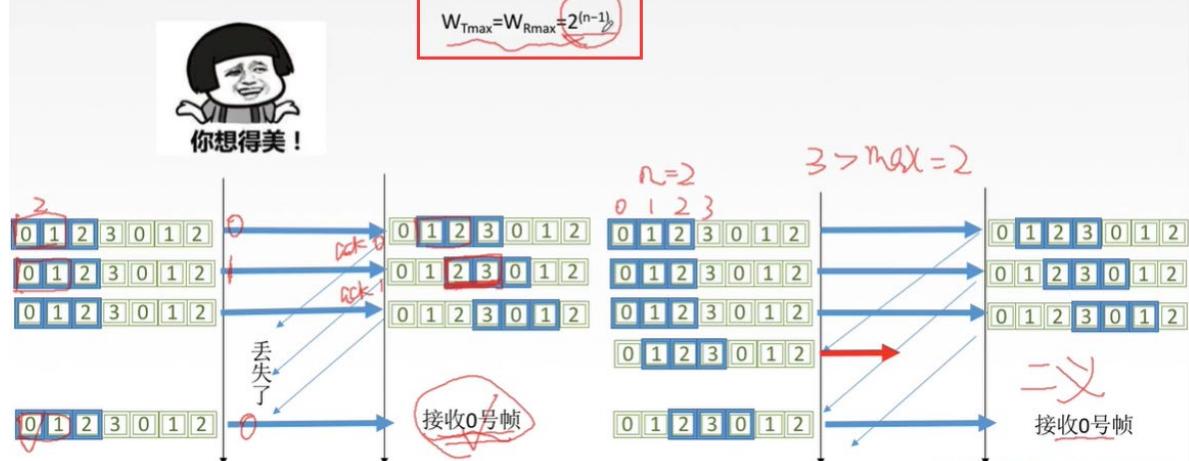
发送窗口最好等于接收窗口。（大了会溢出，小了没意义）



滑动窗口长度

窗口长度可以无限吗？

发送窗口最好等于接收窗口。（大了会溢出，小了没意义）



SR协议重点总结

1. 对数据帧逐一确认，收一个确认一个

2. 只重传出错帧

3. 接收方有缓存

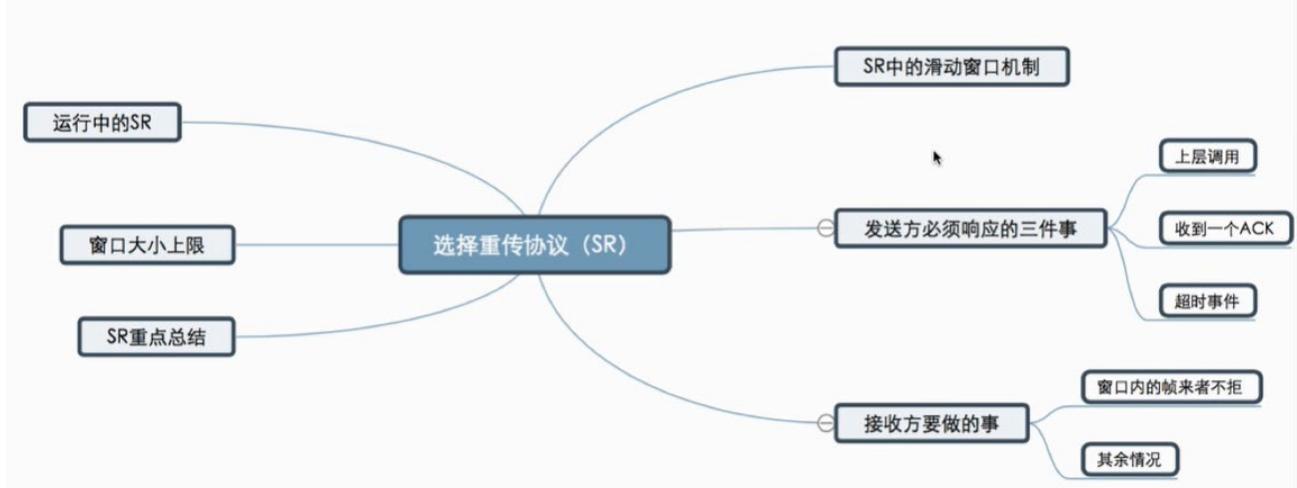
4. $W_{Tmax}=W_{Rmax}=2^{n-1}$

习题1

数据链路层采用了选择重传 (SR) 协议，发送方已经发送了编号为0~3的帧。现已收到1号帧的确认，而0、2号帧依次超时，则发送方需要重传的帧数是（ ）。

- A. 2 B. 3 C. 4 D. 5

02



3.5.1 信道划分介质访问控制

传输数据使用的两种链路

如两人打电话，第三人无法听到

点对点链路

两个相邻节点通过一个链路相连，没有第三者。

应用：PPP协议，常用于广域网

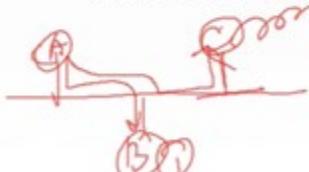


广播式链路

所有主机共享通信介质。

应用：早期的总线以太网、无线局域网，常用于局域网。

典型拓扑结构：总线型、星型（逻辑总线型）



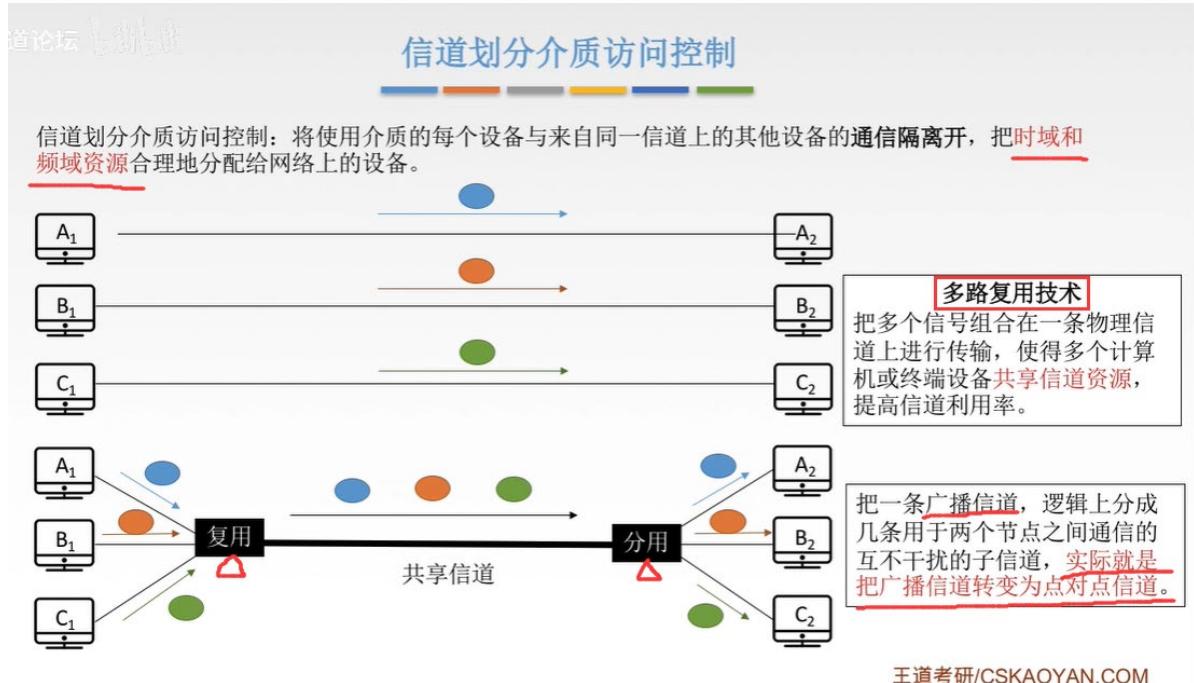
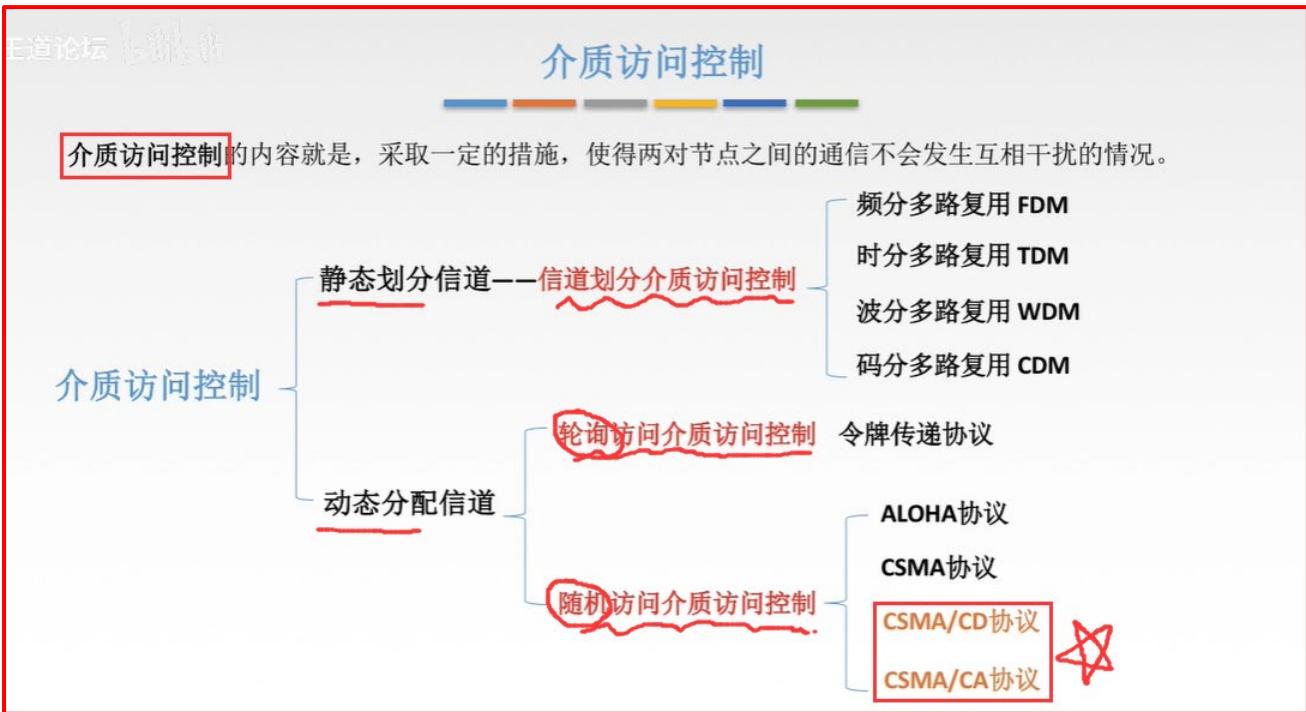
链路中间断了，后面的可能都不能用了



如村里广播、多台对讲机

广播式链路中，比如两台对讲机同时按下按钮讲话，此时就可能会造成冲突，使通话失败。而数据链路层就要解决该问题，采取一定的措施使节点之间的通信不会发生干扰。即对其共享的介质进行访问的控制。

主机/节点/站点之间

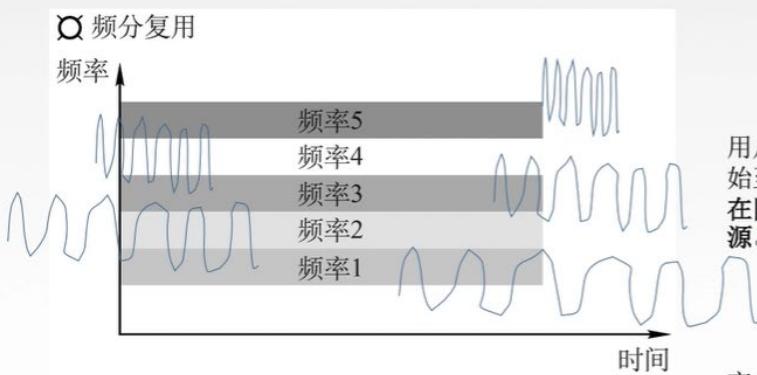


静态划分信道 信道划分介质访问控制

- 频分多路复用 FDM
- 时分多路复用 TDM
- 波分多路复用 WDM
- 码分多路复用 CDM

王道论坛

频分多路复用 FDM



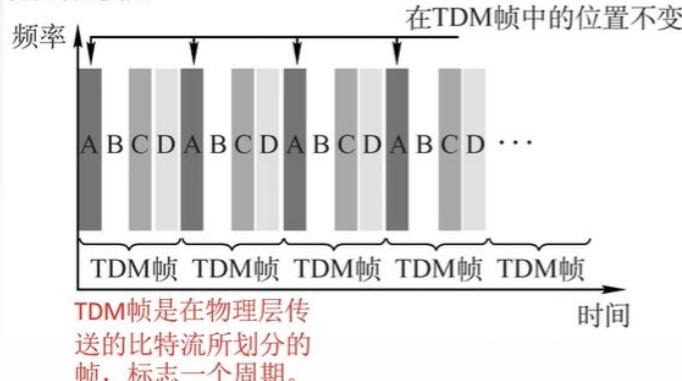
用户在分配到一定的频带后，在通信过程中自始至终都占用这个频带。频分复用的所有用户在同样的时间占用不同的带宽（频率带宽）资源。

指通信过程中的频率带宽, Hz!

充分利用传输介质带宽，系统效率较高；由于技术比较成熟，实现也比较容易。

时分多路复用 TDM

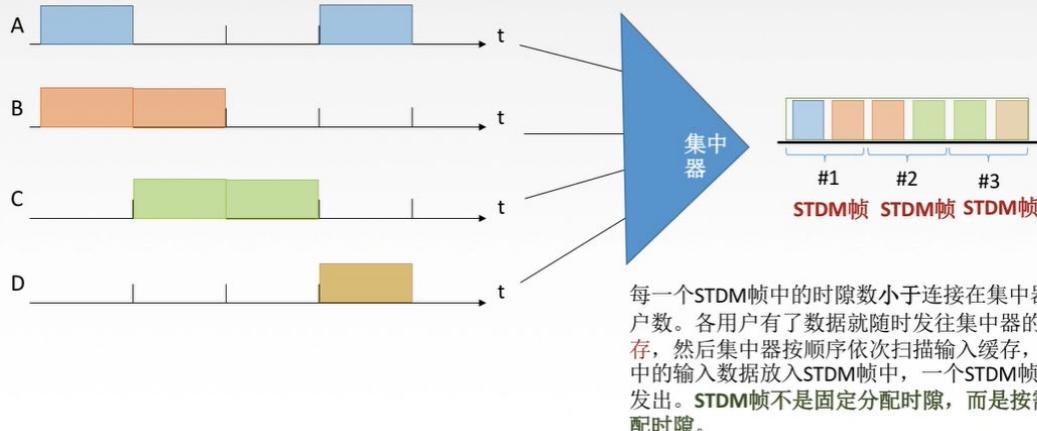
时分复用



将时间划分为一段段等长的时分复用帧（TDM帧）。每一个时分复用的用户在每一个TDM帧中占用固定序号的时隙，所有用户轮流占用信道。

频分复用——“并行”
时分复用——“并发”

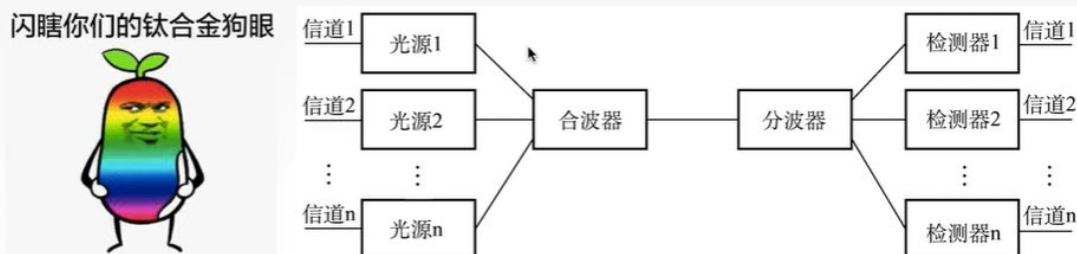
改进的时分复用——统计时分复用STDM



若线路传输速率为 8000b/s。则对于 TDM 来说(ABCD 四个用户)，每个用户分到最高 2000b/s；对于 STDM 来说(ABCD 四个用户)，每个用户最高 8000b/s。

波分多路复用WDM

波分多路复用就是光的频分多路复用，在一根光纤中传输多种不同波长（频率）的光信号，由于波长（频率）不同，所以各路光信号互不干扰，最后再用波长分解复用器将各路波长分解出来。



码分多路复用CDMA

码分多址（CDMA）是码分复用的一种方式。

1个比特分为多个码片/chip，每一个站点被指定一个唯一的m位的芯片序列。

发送1时站点发送芯片序列，发送0时发送芯片序列反码（通常把0写成-1）。

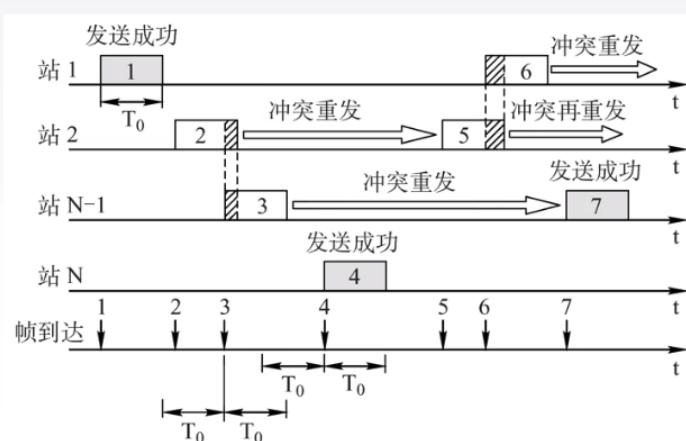
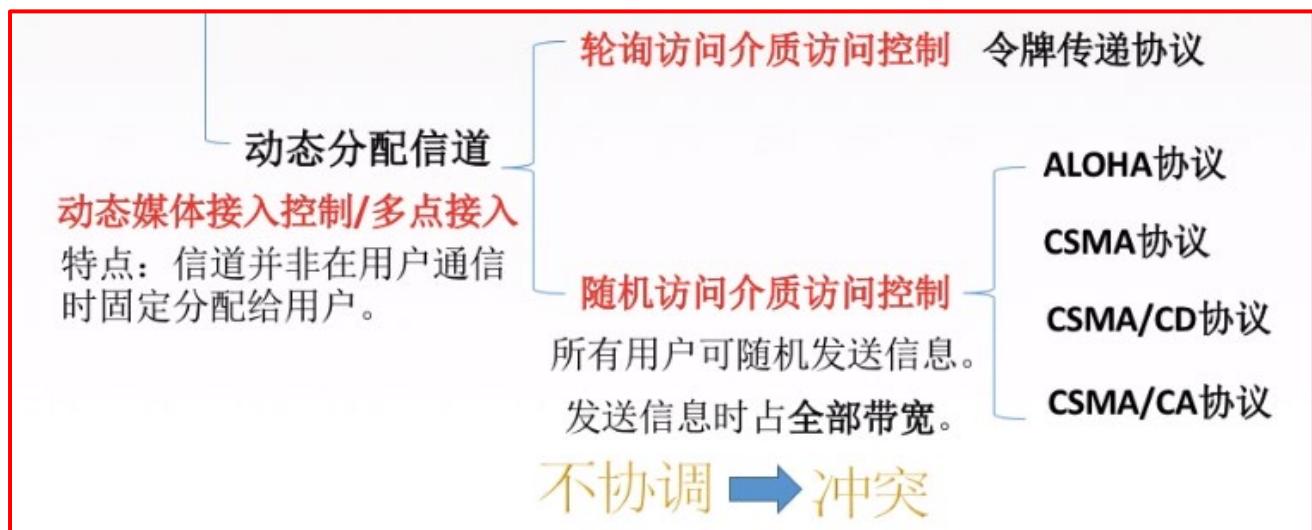
如何不打架：多个站点同时发送数据的时候，要求各个站点芯片序列相互正交。 $\sum_{i=1}^n a_i b_i = 0$

如何合并：各路数据在信道中被线性相加。

如何分离：合并的数据和源站规格化内积。

$$\begin{aligned}
 & \text{发送端} \\
 & \text{A: } (+1+1+1+1+1+1) \\
 & \text{B: } (-1-1-1-1-1-1) \\
 & \text{C: } (+1+1+1+1+1+1) \\
 & \text{D: } (-1-1-1-1-1-1) \\
 & \text{总信号: } (+1+1+1+1+1+1) + (-1-1-1-1-1-1) + (+1+1+1+1+1+1) + (-1-1-1-1-1-1) \\
 & \text{接收端} \\
 & \text{解调后的信号: } (+1+1+1+1+1+1) - (-1-1-1-1-1-1) + (+1+1+1+1+1+1) - (-1-1-1-1-1-1) \\
 & \text{结果: } 8 = 8
 \end{aligned}$$

3.5.2 ALOHA 协议



冲突如何检测？

如果发生冲突，接收方在就会检测出差错，然后不予确认，发送方在一定时间内收不到就判断发生冲突。

冲突如何解决？

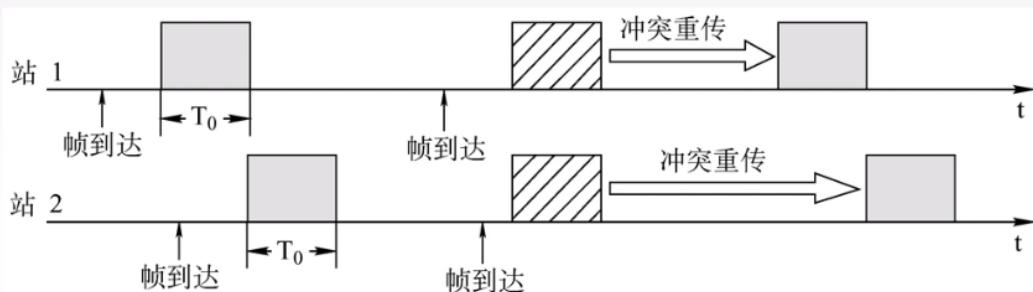
超时后等一随机时间再重传。

T_0 指的是：数据帧从刚开始发送到发送成功的时间，包括传输、传播时间。

时隙ALOHA协议

时隙ALOHA协议的思想：把时间分成若干个相同的时间片，所有用户在时间片开始时刻同步接入网络信道，若发生冲突，则必须等到下一个时间片开始时刻再发送。

控制想发就发的随意性

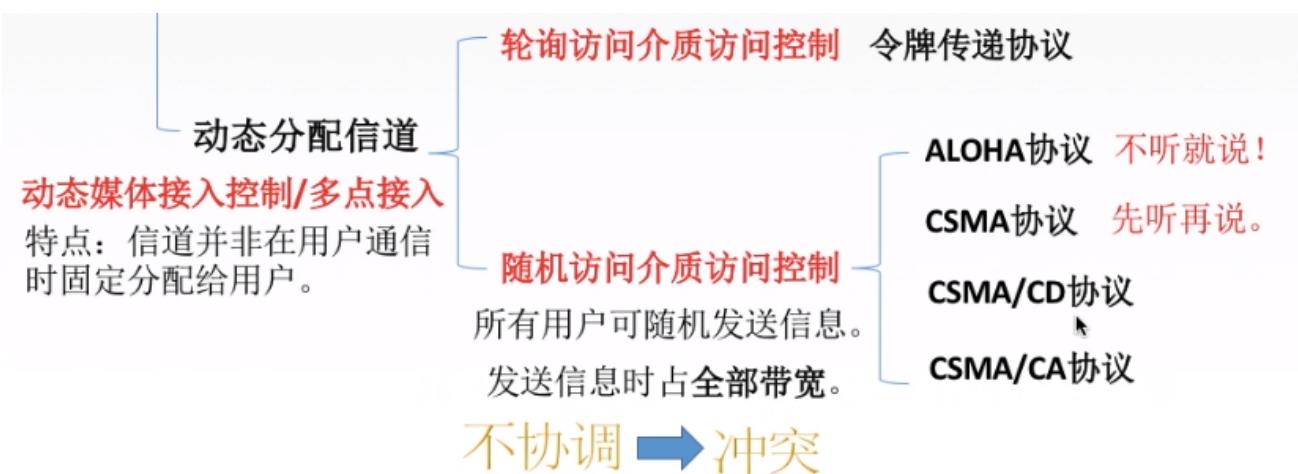


关于ALOHA要知道的事

这里的吞吐量：在一定时间内，成功发送的平均帧数

1. 纯ALOHA比时隙ALOHA吞吐量更低，效率更低。
2. 纯ALOHA想发就发，时隙ALOHA只有在时间片段开始时才能发。

3.5.3 CSMA (carrier sense multiple access) 协议



CSMA协议

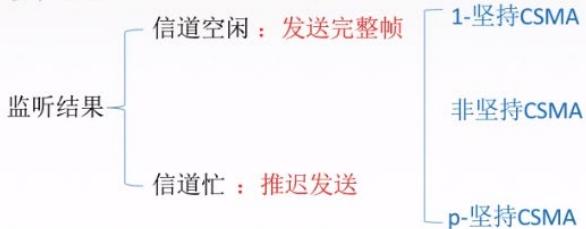
载波监听多路访问协议CSMA (carrier sense multiple access)

CS: 载波侦听/监听，每一个站在发送数据之前要检测一下总线上是否有其他计算机在发送数据。

当几个站同时在总线上发送数据时，总线上的信号**电压摆动值**将会增大（互相叠加）。当一个站检测到的信号电压摆动值超过一定门限值时，就认为总线上至少有两个站同时在发送数据，表明产生了碰撞，即发生了冲突。

MA: 多点接入，表示许多计算机以多点接入的方式连接在一根总线上。

协议思想：发送帧之前，监听信道。



1-坚持CSMA

坚持指的是对于监听信道**忙**之后的坚持。

1-坚持CSMA思想：如果一个主机要发送消息，那么它先监听信道。

空闲则直接传输，不必等待。

忙则一直监听，直到空闲再上传输。 因为传输时延，总线上可能并不是真正的空闲，所以存在冲突

如果有冲突（一段时间内未收到肯定回复），则等待一个随机长的时间再监听，重复上述过程。

优点：只要媒体空闲，站点就马上发送，避免了媒体利用率的损失。

缺点：假如有两个或两个以上的站点有数据要发送，冲突就不可避免。

PS：**忙**是有东西在信道(传输+传播)，**冲突**应该是两个都监听到是空闲(即同时可能有很多节点都在监听这个信道，都采用 1-坚持 CSMA)，然后两个同时传入信道就会发生冲突。

非坚持CSMA

非坚持指的是对于监听信道**忙**之后就不继续监听。

非坚持CSMA思想：如果一个主机要发送消息，那么它先监听信道。

空闲则直接传输，不必等待。

忙则等待一个随机的时间之后再进行监听。

优点：采用随机的重发延迟时间可以减少冲突发生的可能性。

缺点：可能存在大家都在延迟等待过程中，使得媒体仍可能处于空闲状态，媒体使用率降低。

p-坚持CSMA

p-坚持指的是对于监听信道空闲的处理。

p-坚持CSMA思想：如果一个主机要发送消息，那么它先监听信道。

空闲则以 p 概率直接传输，不必等待；概率 $1-p$ 等待到下一个时间槽再传输。

~~忙则等待一个随机的时间之后再进行监听~~

优点：既能像非坚持算法那样减少冲突，又能像1-坚持算法那样减少媒体空闲时间的这种方案。

~~忙则持续监听(推迟到下一个时隙再监听)~~

BUT !

~~发生冲突后还是要坚持把数据帧发送完，造成了浪费。~~

CSMA-CD

有没有什么办法可以减少资源浪费，一冲突就能发现呢？

此处冲突只能以是否接收到~~确认帧~~来判断，所以数据帧都需要发送完，造成浪费。所以就有了后面的CSMA-CD 和 CSMA-CA 协议。

三种CSMA对比总结

	1-坚持CSMA	非坚持CSMA	p-坚持CSMA
信道空闲	马上发	马上发	p 概率马上发 $1-p$ 概率等到下一个时隙再发 持续监听(推迟到下一个时隙再监听)
信道忙	继续坚持监听	放弃监听，等一个随机时间再监听	放弃监听，等一个随机时间再监听 继续坚持监听

3.5.4 ※CSMA-CD (carrier sense multiple access with collision detection) 协议 (对碰撞的检测)

只要用于：总线式以太网(有线网)。



CSMA/CD协议

载波监听多点接入/碰撞检测CSMA/CD (carrier sense multiple access with collision detection)

CS: 载波侦听/监听，每一个站在发送数据之前以及发送数据时都要检测一下总线上是否有其他计算机在发送数据。

MA: 多点接入，表示许多计算机以多点接入的方式连接在一根总线上。总线型网络

CD: 碰撞检测（冲突检测），“边发送边监听”，适配器边发送数据边检测信道上信号电压的变化情况，以便判断自己在发送数据时其他站是否也在发送数据。

先听后发为什么
还会冲突？

因为电磁波在总线上总是以有限的速率传播的。

主要用于

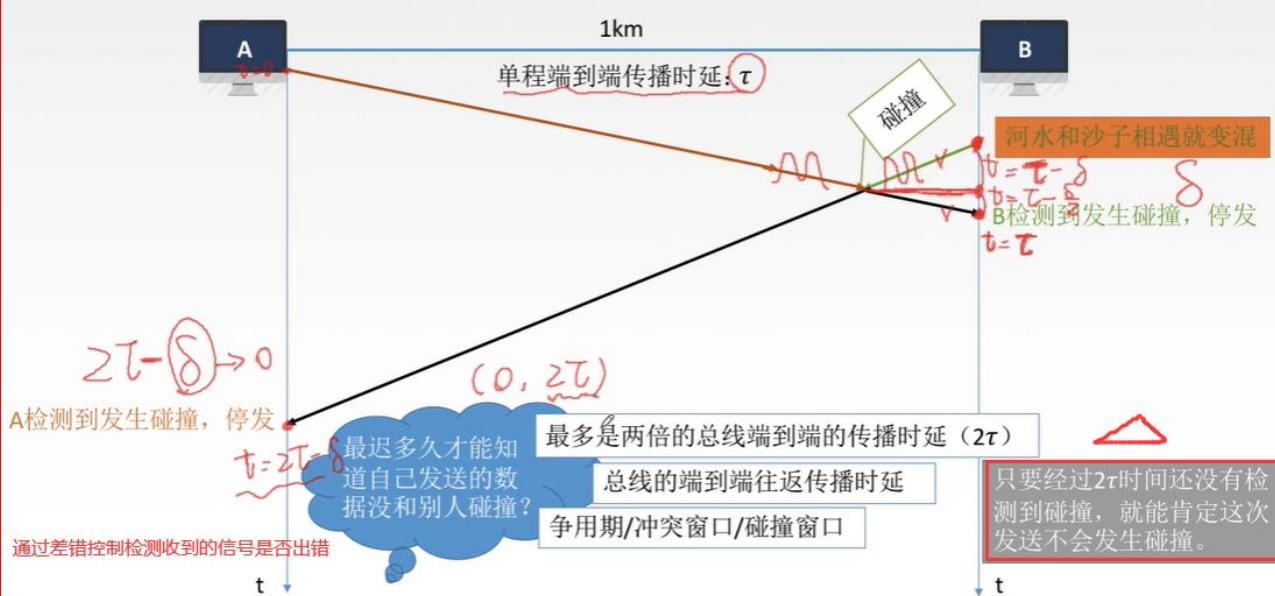
当两个站点同时监听到信道空闲时，均发送数据。此时双方都占时还不能检测到有信号进入自己的站点，因为电磁波在总线上传播速率有限。

传播时延对载波监听的影响

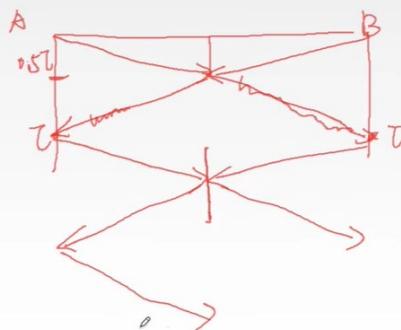


冲突！碰撞！

传播时延对载波监听的影响



如何确定碰撞后的重传时机?



如果碰撞后马上重传，则会再次发生碰撞。所以需要确定重传机制来避免传播时延造成的影响。

如何确定碰撞后的重传时机?

截断二进制指数规避算法

1. 确定基本退避（推迟）时间为争用期 2τ 。
 2. 定义参数 k ，它等于重传次数，但 k 不超过 10，即 $k = \min[\text{重传次数}, 10]$ 。当重传次数不超过 10 时， k 等于重传次数；当重传次数大于 10 时， k 就不再增大而一直等于 10。
 3. 从离散的整数集合 $[0, 1, \dots, 2^k - 1]$ 中随机取出一个数 r ，重传所需要退避的时间就是 r 倍的基本退避时间，即 $2r\tau$ 。
 4. 当重传达 16 次仍不能成功时，说明网络太拥挤，认为此帧永远无法正确发出，抛弃此帧并向高层报告出错。
- 第一次重传， $k=1$ ， r 从 $\{0, 1\}$ 选；
重传推迟时间为 0 或 2τ ，在这两个时间中随机选一个；
若再次碰撞，则在第二次重传时， $k=2$ ， r 从 $\{0, 1, 2, 3\}$ 选；
重传推迟时间为 0 或 2τ 或 4τ 或 6τ ，在这四个时间中随机选一个；
若再次碰撞，则第三次重传时， $k=3$ ， r 从 $\{0, 1, 2, 3, 4, 5, 6, 7\}$ 选.....
- 若连续多次发生冲突，就表明可能有较多的站参与争用信道。使用此算法可使重传需要推迟的平均时间随重传次数的增大而增大，因而减小发生碰撞的概率，有利于整个系统的稳定。

最小帧长问题

A 站发了一个很短的帧

但发生了碰撞

不过帧在发送完毕后才检测到发生碰撞

没法停止发送

因为发完了。。

还有这种操作！！！

所以为了是 CSMA-CD 协议有意义，定义了一个最小帧长，使得检测到碰撞时，帧发送还未结束。

最小帧长问题



我想到了！

最小帧长

即检测到冲突就停发了

帧的传输时延至少要两倍于信号在总线中的传播时延。

$$\frac{\text{帧长 (bit)}}{\text{数据传输速率}} \geq 2\tau$$

最小帧长=总线传播时延 \times 数据传输速率 $\times 2$

$$2\tau \times \text{数据传输速率}$$

以太网规定最短帧长为64B，凡是长度小于64B的都是由于冲突而异常终止的无效帧。

以太网规定最短帧长为 64B。

CSMA/CD协议

CS、MA、CD

传播时延对载波监听的影响

截断二进制指数规避算法

最小帧长= $2 \times$ 总线传播时延 \times 数据传输速率

3.5.5 CSMA-CA (carrier sense multiple access with collision avoidance) 协议（对碰撞的避免）

主要用于：无线局域网(无线网)。

CSMA/CA协议

载波监听多点接入/碰撞避免CSMA/CA (carrier sense multiple access with collision avoidance)

为什么要用CSMA/CA? → 无线局域网

无法做到360° 全面检测碰撞

当A和C都检测不到信号，认为信道空闲时，同时向终端B发送数据帧，就会导致冲突。

C相对于A就是一个隐蔽站

有礼貌的CSMA/CA



先听后说，听到空闲后，还要再等待一小会儿

CSMA/CA协议工作原理

发送数据前，先检测信道是否空闲。

空闲则发出RTS (request to send)，RTS包括发射端的地址、接收端的地址、下一份数据将持续发送的时间等信息；信道忙则等待。

接收端收到RTS后，将响应CTS (clear to send)。

避免其他未收到CTS的站点苦等

也可以让其他站点在这段时间内不发送数据，从而实现碰撞避免

发送端收到CTS后，开始发送数据帧（同时预约信道：发送方告知其他站点自己要传多久数据）。

接收端收到数据帧后，将用CRC来检验数据是否正确，正确则响应ACK帧。

发送方收到ACK就可以进行下一个数据帧的发送，若没有则一直重传至规定重发次数为止（采用二进制指数退避算法来确定随机的推迟时间）。

1. 预约信道

2. ACK帧

3. RTS/CTS帧（可选）

如果RTS (request to send) / CTS (clear to send) 冲突了，也采用二进制指数规避算法来重发

CSMA/CD与CSMA/CA

相同点：

CSMA/CD与CSMA/CA机制都从属于CSMA的思路，其核心是先听再说。换言之，两个在接入信道之前都需要进行监听。当发现信道空闲后，才能进行接入。

不同点：

1. 传输介质不同：CSMA/CD用于总线式以太网【有线】，而CSMA/CA用于无线局域网【无线】。

2. 载波检测方式不同：因传输介质不同，CSMA/CD与CSMA/CA的检测方式也不同。CSMA/CD通过电缆中电压的变化来检测，当数据发生碰撞时，电缆中的电压就会随着发生变化；而CSMA/CA采用能量检测(ED)、载波检测(CS)和能量载波混合检测三种检测信道空闲的方式。

3. CSMA/CD检测冲突，CSMA/CA避免冲突，二者出现冲突后都会进行有上限的重传。

3.5.6 轮询访问介质访问控制

结合信道划分介质访问控制和随机访问介质访问控制的优点。

王道论坛

介质访问控制

1 信道划分介质访问控制（MAC Multiple Access Control）协议：
 基于多路复用技术划分资源。
 [网络负载重：共享信道效率高，且公平
 [网络负载轻：共享信道效率低]]

2 随机访问MAC协议：冲突
 用户根据意愿随机发送信息，发送信息时可独占信道带宽。
 [网络负载重：产生冲突开销
 [网络负载轻：共享信道效率高，单个结点可利用信道全部带宽]]

3 轮询访问MAC协议/轮流协议/轮转访问MAC协议：
 既要不产生冲突，又要发送时占全部带宽。
 典型协议
 [轮询协议
 [令牌传递协议]]

只有我最棒！

令牌传递协议



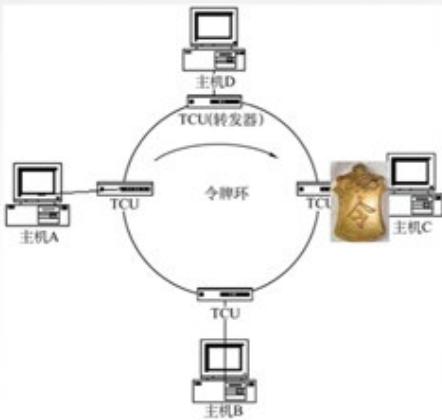
问题：

主节点故障(可设置备用主节点)

1. 轮询开销
2. 等待延迟
3. 单点故障

每次询问都会产生数据帧，当节点过多时，轮询开销较大

令牌传递协议



令牌：一个特殊格式的MAC控制帧，不含任何信息。

控制信道的使用，确保同一时刻只有一个结点独占信道。

令牌环网无碰撞

每个结点都可以在一定的时间内（令牌持有时间）获得发送数据的权利，并不是无限制地持有令牌。

问题：

- 1.令牌开销
- 2.等待延迟
- 3.单点故障

应用于令牌环网（物理星型拓扑，逻辑环形拓扑）。

采用令牌传送方式的网络常用于**负载较重、通信量较大的网络中。**

轮询访问、令牌环网

3.6.1 局域网基本概念和体系结构

局域网

局域网（Local Area Network）：简称LAN，是指在某一区域内由多台计算机互联成的计算机组，使用**广播信道**。

特点1：覆盖的地理范围较小，只在一个相对独立的局部范围内联，如一座或集中的建筑群内。

特点2：使用专门铺设的传输介质（双绞线、同轴电缆）进行联网，数据传输速率高（10Mb/s~10Gb/s）。

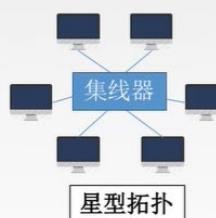
特点3：通信延迟时间短，误码率低，可靠性较高。

特点4：各站为平等关系，共享传输信道。

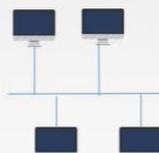
特点5：多采用分布式控制和广播式通信，能进行广播和组播。

决定局域网的主要要素为：**网络拓扑**，传输介质与**介质访问控制方法**。

局域网拓扑结构



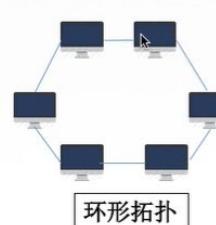
中心节点是控制中心，任意两个节点间的通信最多只需**两步**，传输速度快，并且网络构形简单、建网容易、便于控制和管理。但这种网络系统，**网络可靠性低**，**网络共享能力差**，**有单点故障问题**。
如集线器故障



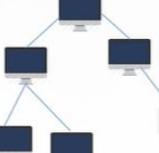
总线型拓扑

优势较大，成本较低。现局域网中较常用的以太网就是一种逻辑上的总线型拓扑结构

网络可靠性高、网络节点间响应速度快、共享资源能力强、设备投入量少、成本低、安装使用方便，当某个工作站节点出现故障时，对整个网络系统影响小。



系统中通信设备和线路比较节省。有**单点故障问题**；由于环路是封闭的，所以**不便于扩充**，**系统响应延时长**，且**信息传输效率相对较低**。



树型拓扑

易于拓展，易于隔离故障，**也容易有单点故障**。
如根节点

局域网传输介质

局域网
有线局域网 常用介质：双绞线、同轴电缆、光纤
无线局域网 常用介质：电磁波
空气中传播

局域网介质访问控制方法

1.CSMA/CD 常用于**总线型局域网**，也用于树型网络

2.令牌总线 常用于**总线型局域网**，也用于树型网络

它是把总线型或树型网络中的各个工作站按一定顺序如按接口地址大小排列形成一个逻辑环。只有令牌持有者才能控制总线，才有发送信息的权力。

3.令牌环 用于**环形局域网**，如令牌环网

局域网的分类 覆盖范围一般在几千米内

✓**以太网** 以太网是应用最为广泛的局域网，包括标准以太网（10Mbps）、快速以太网（100Mbps）、千兆以太网（1000 Mbps）和10G以太网，它们都符合IEEE802.3系列标准规范。逻辑拓扑总线型，物理拓扑是星型或拓展星型。使用CSMA/CD。
造价高，且不太实用

2.令牌环网 物理上采用了星形拓扑结构，逻辑上是环形拓扑结构。已是“明日黄花”。

光纤分布式数据接口 ——**较贵**

3.FDDI网（Fiber Distributed Data Interface） 物理上采用了双环拓扑结构，逻辑上是环形拓扑结构。

4.ATM网（Asynchronous Transfer Mode） 较新型的单元交换技术，使用53字节固定长度的单元进行交换。
WIFI是无线局域网的一种应用

✓**5.无线局域网（Wireless Local Area Network； WLAN）** 采用**IEEE 802.11**标准。

IEEE 802标准

IEEE 802系列标准是IEEE 802 LAN/MAN 标准委员会制定的局域网、城域网技术标准（1980年2月成立）。其中最广泛使用的有以太网、令牌环、无线局域网等。这一系列标准中的每一个子标准都由委员会中的一个专门工作组负责。



IEEE 802委员会

IEEE 802现有标准

IEEE 802.1：局域网体系结构、寻址、[网络互联](#)和网络

IEEE 802.1A：概述和系统结构

IEEE 802.1B：网络管理和网络互连

IEEE 802.2：逻辑链路控制子层（LLC）的定义。

IEEE 802.3：以太网介质访问控制协议（CSMA/CD）及物理层技术规范^[2]。

IEEE 802.4：令牌总线网（Token-Bus）的介质访问控制协议及物理层技术规范。

IEEE 802.5：令牌环网（Token-Ring）的介质访问控制协议及物理层技术规范。

IEEE 802.6：城域网介质访问控制协议DQDB（Distributed Queue Dual Bus 分布式队列双总线）及物理层技术规范。

IEEE 802.7：宽带技术咨询组，提供有关宽带联网的技术咨询。

IEEE 802.8：光纤技术咨询组，提供有关光纤联网的技术咨询。

IEEE 802.9：综合声音数据的局域网（IVD LAN）[介质访问控制协议](#)及物理层技术规范。

IEEE 802.10：网络安全技术咨询组，定义了[网络互操作](#)的认证和加密方法。

IEEE 802.11：无线局域网（WLAN）的介质访问控制协议及物理层技术规范。

IEEE 802.11，1997年，原始标准（2Mbit/s，播在2.4GHz）。

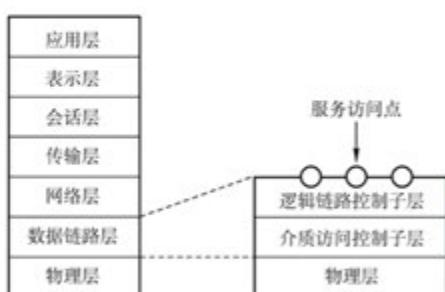
IEEE 802.11a，1999年，物理层补充（54Mbit/s，播在5GHz）。

IEEE 802.11b，1999年，物理层补充（11Mbit/s播在2.4GHz）。

王道论坛

MAC子层和LLC子层

IEEE 802标准所描述的局域网参考模型只对应OSI参考模型的[数据链路层与物理层](#)，它将数据链路层划分为逻辑链路层LLC子层和介质访问控制MAC子层。



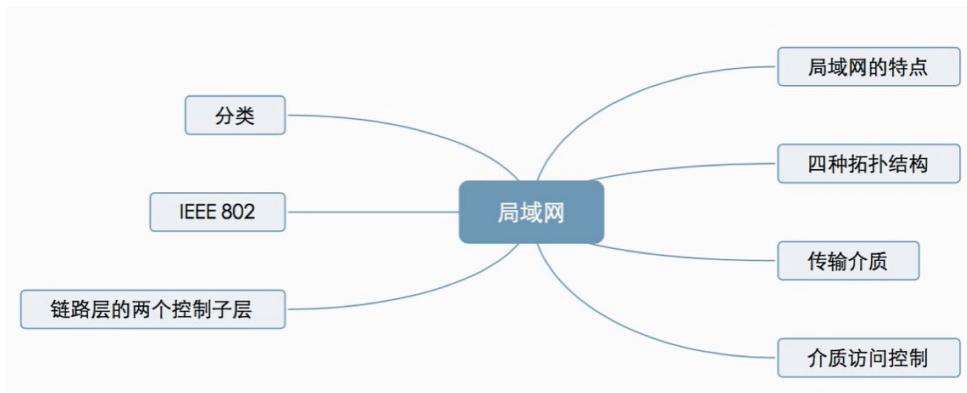
为网络层服务

LLC负责识别网络层协议，然后对它们进行封装。LLC报头告诉数据链路层一旦帧被接收到时，应当对数据包做何处理。为网络层提供服务：无确认无连接、面向连接、带确认无连接、高速传送。

将数据报封装成帧

MAC子层的主要功能包括数据帧的封装/卸装，帧的寻址和识别，帧的接收与发送，链路的管理，帧的差错控制等。MAC子层的存在屏蔽了不同物理链路种类的差异性。

与物理层相关



※※※3.6.2 以太网

以太网概述

以太网(Ethernet)指的是由Xerox公司创建并由Xerox、Intel和DEC公司联合开发的**基带总线局域网规范**，是当今现有局域网采用的最通用的通信协议标准。以太网络使用**CSMA/CD**（载波监听多路访问及冲突检测）技术。

以太网在局域网各种技术中占**统治性地位**：

1. 造价低廉（以太网网卡不到100块）；
2. 是应用最广泛的局域网技术；
3. 比令牌环网、ATM网便宜，简单；
4. 满足网络速率要求：**10Mb/s~10Gb/s.**

以太网两个标准

DIX Ethernet V2: 第一个局域网产品（以太网）规约。

IEEE 802.3: IEEE 802.3委员会802.3工作组制定的第一个IEEE的以太网标准。（帧格式有一丢丢改动）

802.3局域网 AKA 以太网

以太网提供无连接、不可靠的服务

无连接：发送方和接收方之间无“握手过程”。

不可靠：不对发送方的数据帧**编号**，接收方不向发送方进行**确认**，差错帧直接丢弃，差错纠正由高层负责。

传输层/运输层

只检错，不纠错

尽最大努力交付	以太网只实现无差错接收，不实现可靠传输。	
---------	-----------------------------	--

无差错接收：发送什么接收什么，但是发现有错时直接丢弃。主要由以太网实现，即以太网主要负者物理层和数据链路层。

可靠传输：只要发来的帧，全都要接收。帧丢失、重复和失序，都是可靠传输要解决的问题，主要由传输层/运输层实现。

以太网传输介质与拓扑结构的发展



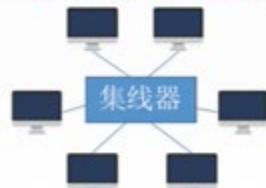
物理拓扑



更利于扩网和检测故障

使用集线器的以太网在逻辑上仍是一个总线网，各站共享逻辑上的总线，使用的还是CSMA/CD协议。

以太网拓扑：逻辑上总线型，物理上星型。



10BASE-T以太网

信源发出(数字信号)

10BASE-T是传送基带信号的双绞线以太网，T表示采用双绞线，现10BASE-T采用的是无屏蔽双绞线(UTP)，传输速率是10Mb/s。



物理上采用星型拓扑，逻辑上总线型，每段双绞线最长为100m。

采用曼彻斯特编码。

采用CSMA/CD介质访问控制。

适配器与MAC地址



网络接口板

网络接口卡NIC (network interface card)

NOW，不再使用单独网卡 现主板上已经嵌入好网卡

适配器上装有处理器和存储器（包括RAM和ROM）。

ROM上有计算机硬件地址MAC地址。

在局域网中，硬件地址又称为物理地址，或MAC地址。【实际上是标识符】

MAC地址：每个适配器有一个全球唯一的48位二进制地址，前24位代表厂家（由IEEE规定），后24位厂家自己指定。常用6个十六进制数表示，如02-60-8c-e4-b1-21。

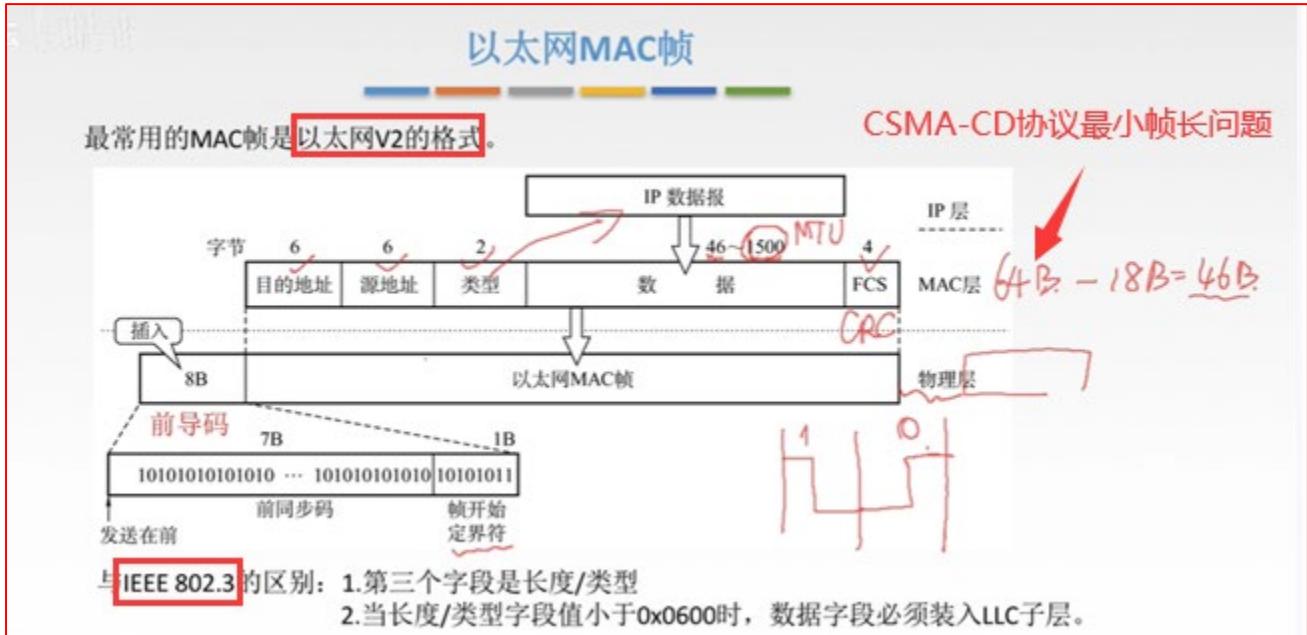


一个网卡一个mac地址

全球唯一

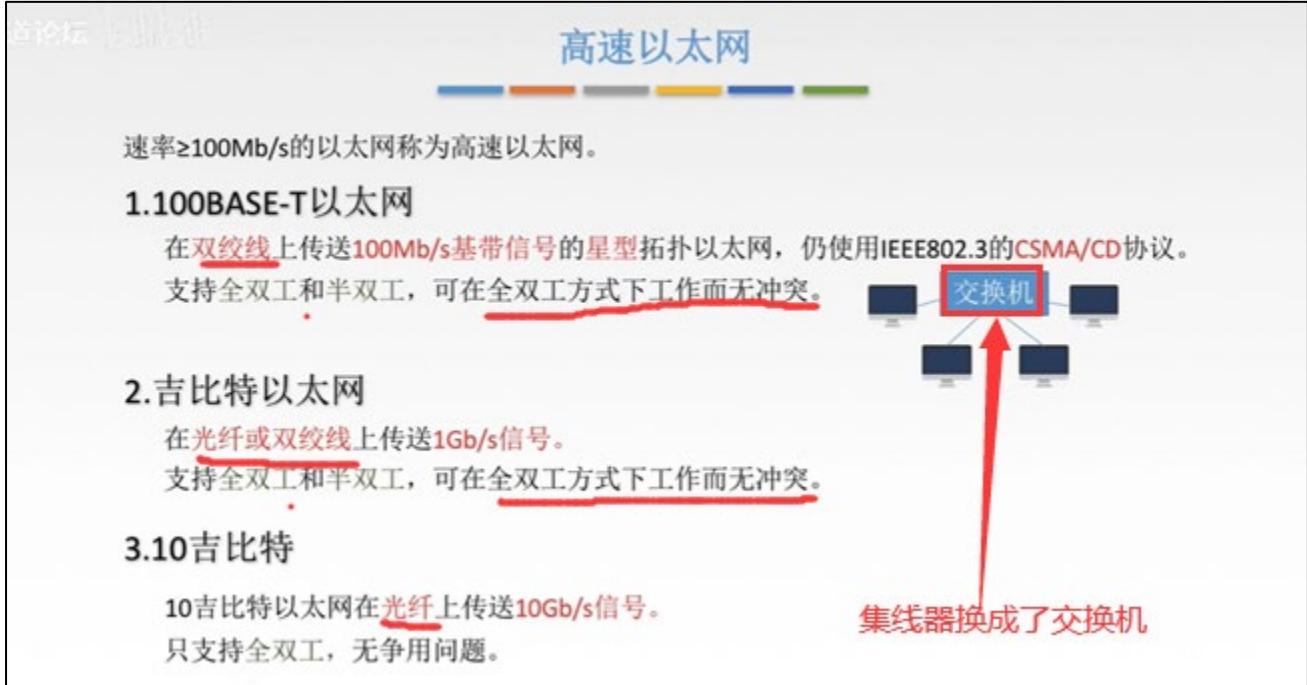
若计算机的网卡坏了，换了一个网卡，则 mac 地址也会相应改变。

一般鼠标等可以通过 mac 地址查询是否是该厂商的正品。



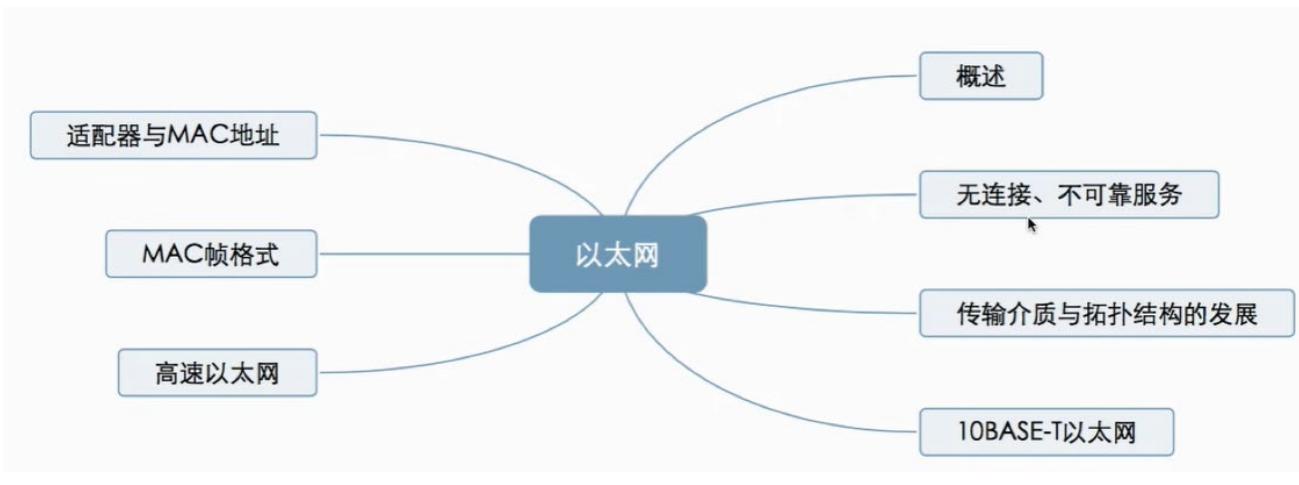
目的地址：1. 发给固定主机：mac 地址（单播地址）；2. 发给所有主机：广播地址（全 1 / 全 F）；3. 多播地址。
类型：指明上面的网络层使用的是什么协议，以便把收到的 mac 帧的数据上交给上一层的协议。
FCS：指的是 CRC 循环冗余检验的帧检验序列。

因为以太网使用的是曼彻斯特编码，即每个码元中间有个电压的变化，所以帧结束后就没有电压的变化了（帧与帧之间发送有空白时间），就知道该帧的结束位置。所以不需要再设置“后导码”。



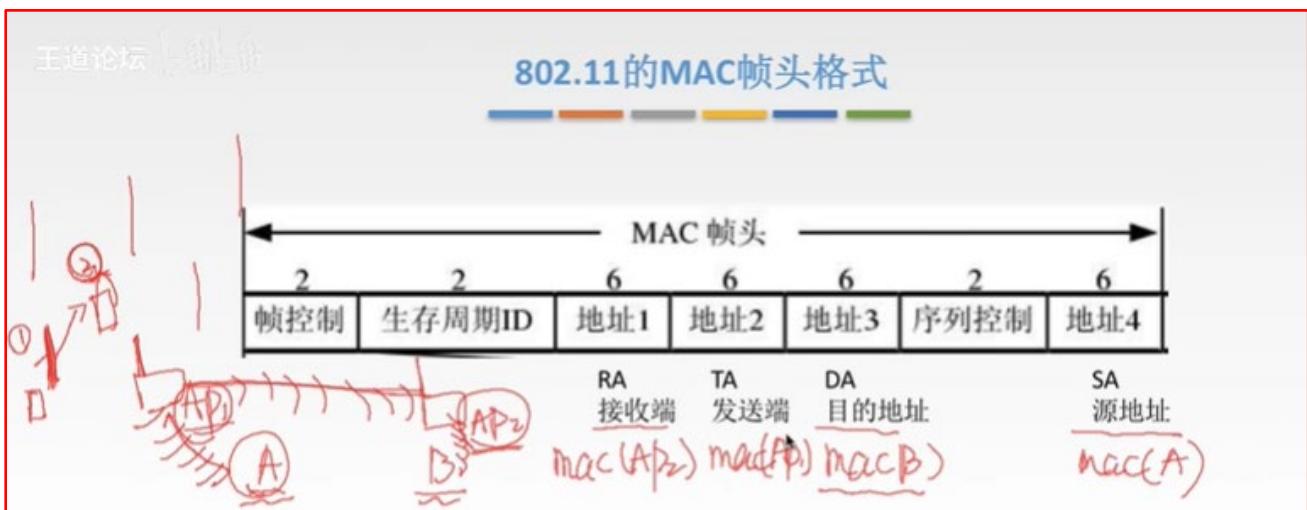
交换机可以隔离冲突域，所以每一个交换机端口就是一个冲突域，则一个主机占一个冲突域就不会产生冲突了。所以在全双工方式下不用使用 CSMA-CD 协议。

交换机工作在数据链路层，通过 MAC 地址转发数据。集线器工作在物理层，通过广播的形式转发数据。
目前，集线器已经被交换机取代，组网中很少使用集线器了。



3.6.3 无线局域网（WLAN）

不等于 WIFI，无线局域网的范围比 WIFI 大得多，一般是几千米。



AP 无线接入点或基站。基站的数据库不断更新，将当前距离自己最近的手机号存入到数据库中。（新到

一个城市收到的欢迎短信往往就是接入到该地方的第一个基站发送的)

802.11的MAC帧头格式

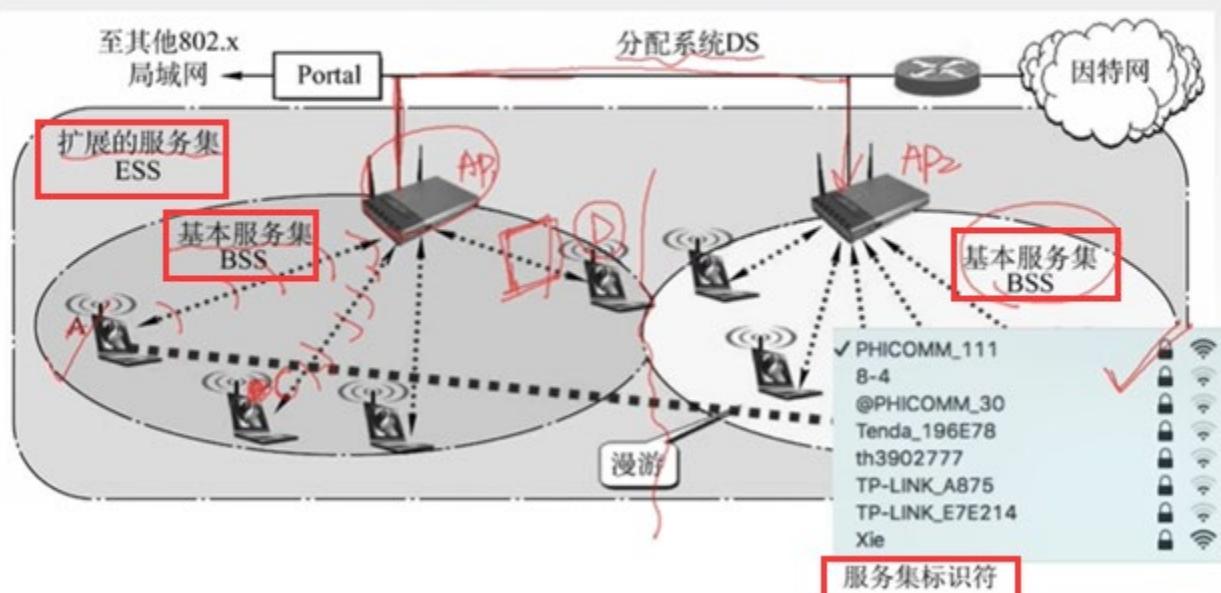
功能	To DS	From DS	Address1 (接收端)	Address2 (发送端)	Address3	Address4
IBSS	0	0	DA	SA	BSSID	未使用
To AP (基础结构型)	1	0	BSSID	SA	DA	未使用
From AP (基础结构型)	0	1	DA	BSSID	SA	未使用
WDS (无线分布式系统)	1	1	RA	TA	DA	SA

无线局域网的分类

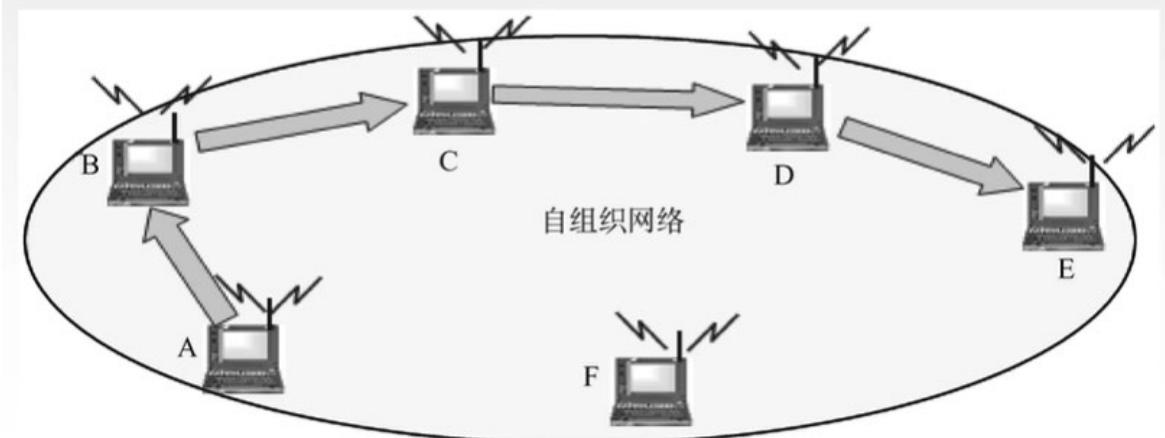
1. 有固定基础设施无线局域网

2. 无固定基础设施无线局域网的自组织网络

有固定基础设施无线局域网



无固定基础设施无线局域网的自组织网络



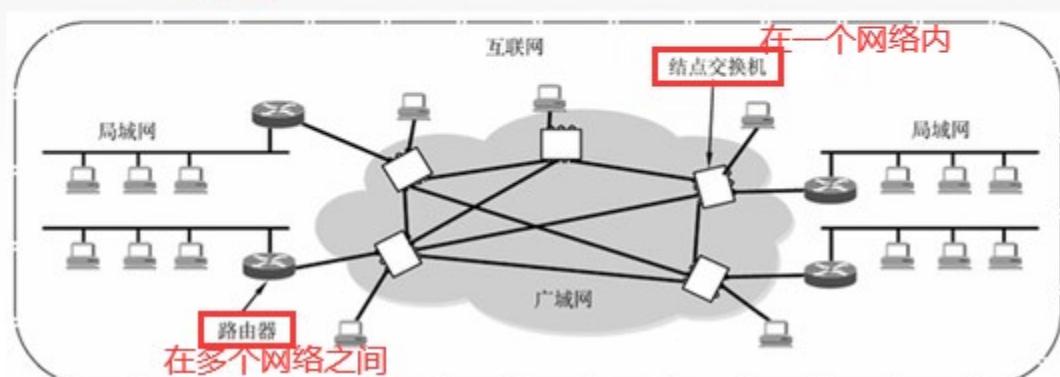
把所有主机安排在一个网段组网。

3.7 PPP 协议和 HDLC 协议——广域网的链路层协议

广域网

广域网（WAN, Wide Area Network），通常跨接很大的物理范围，所覆盖的范围从几十公里到几千公里，它能连接多个城市或国家，或横跨几个洲并能提供远距离通信，形成国际性的远程网络。

广域网的通信子网主要使用分组交换技术。广域网的通信子网可以利用公用分组交换网、卫星通信网和无线分组交换网，它将分布在不同地区的局域网或计算机系统互连起来，达到资源共享的目的。如因特网（Internet）是世界范围内最大的广域网。



广域网覆盖的网络结构层次：物理层（集线器）、链路层（交换机）和网络层（路由器）；普遍采用点对点（全双工或半双工通信模式）；强调资源共享；其数据传输速率更快，但距离也更远，所以传播延时更大。

局域网工作只覆盖物理层和链路层；普遍采用多点接入技术，逻辑上的总线型；强调数据传输。

PPP协议的特点

点对点协议PPP（Point-to-Point Protocol）是目前使用最广泛的数据链路层协议，用户使用拨号电话接入因特网时一般都使用PPP协议。

只支持全双工链路。

最复杂：TCP 协议当中。IP 协议层提供不可靠的数据报服务；所以链路层没有必要在 IP 协议之前实现可靠传输。

王道论坛

PPP协议应满足的要求

简单 对于链路层的帧，无需纠错，无需序号，无需流量控制。

封装成帧 帧定界符 一个字符/字节发送 比特发送

透明传输 与帧定界符一样比特组合的数据应该如何处理：异步线路用字节填充，同步线路用比特填充。

多种网络层协议 封装的IP数据报可以采用多种协议。~~不要求网络层使用什么协议~~

多种类型链路 串行/并行，同步/异步，电/光....

差错检测 错就丢弃。 不用实现可靠传输

检测连接状态 链路是否正常工作。

最大传送单元 数据部分最大长度MTU。 1500字节

网络层地址协商 知道通信双方的网络层地址。 IP地址 数据压缩协商

王道论坛

PPP协议无需满足的要求

纠错 只需检错

流量控制 上层负责

序号

不支持多点线路 只需满足点对点的连接过程

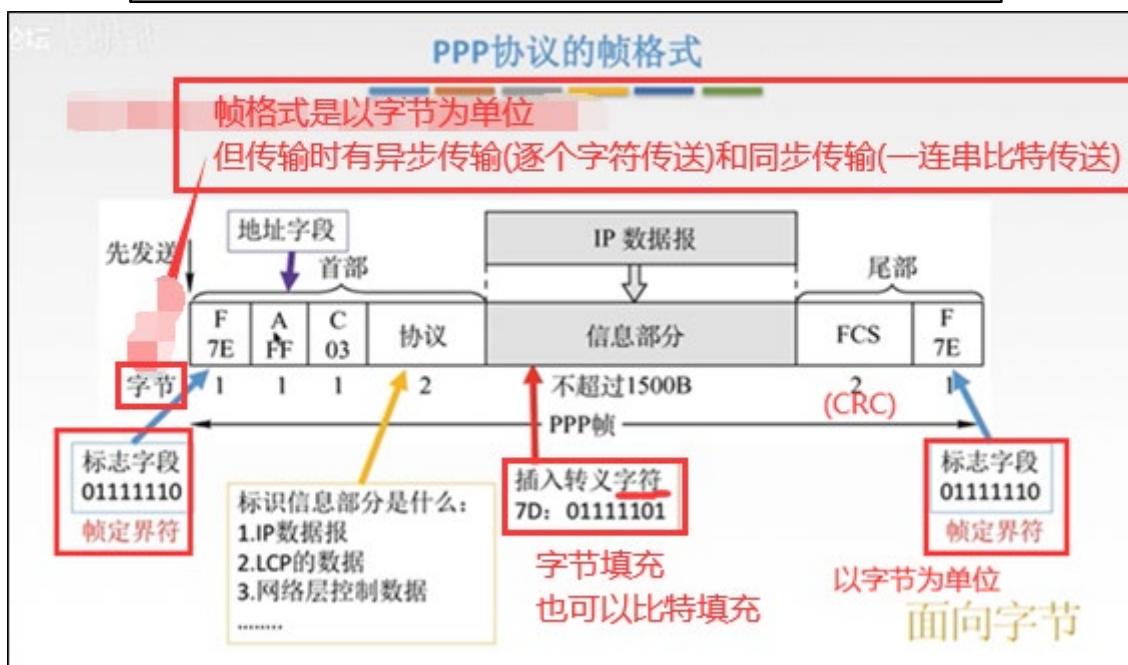
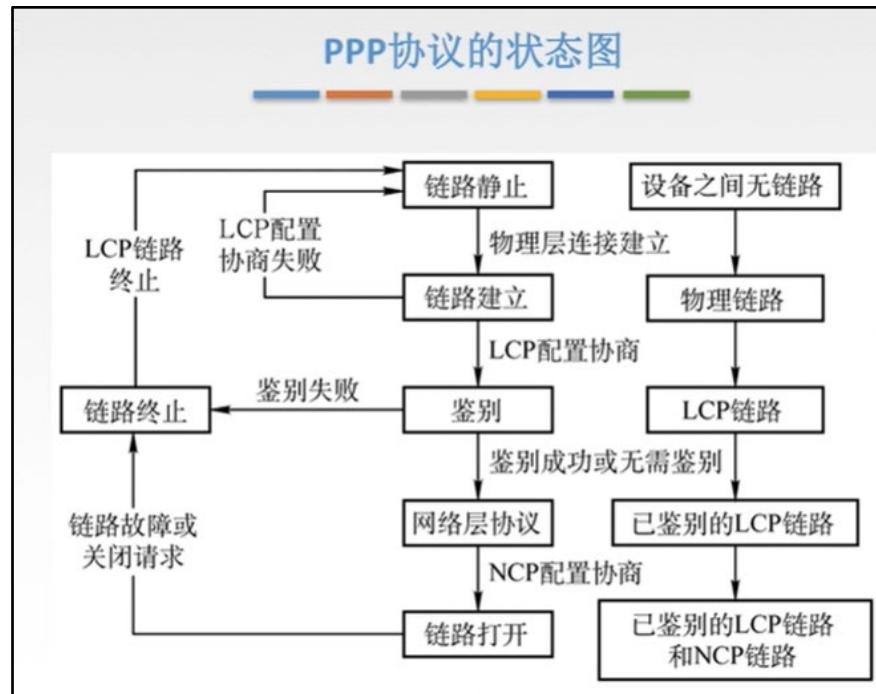
论坛

PPP协议的三个组成部分

1.一个将IP数据报封装到串行链路（同步串行/异步串行）的方法。

2 链路控制协议LCP 建立并维护数据链路连接。 身份验证

3 网络控制协议NCP PPP可支持多种网络层协议，每个不同的网络层协议都要一个相应的NCP来配置，为网络层协议建立和配置逻辑连接。



+ 关注

HDLC的站

主站、从站、复合站

1. 主站的主要功能是发送命令（包括数据信息）帧、接收响应帧，并负责对整个链路的控制系统的初启、流程的控制、差错检测或恢复等。
2. 从站的主要功能是接收由主站发来的命令帧，向主站发送响应帧，并且配合主站参与差错恢复等链路控制。
3. 复合站的主要功能是既能发送，又能接收命令帧和响应帧，并且负责整个链路的控制。

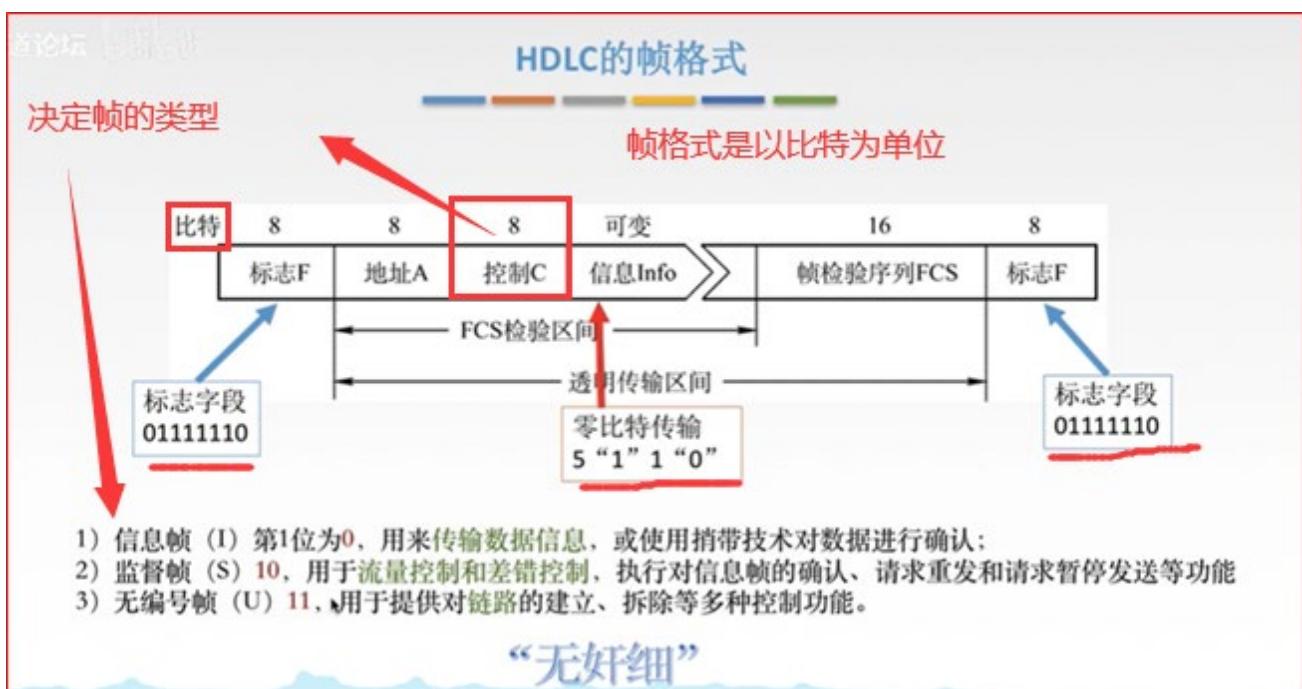
三种数据操作方式：

1. 正常响应方式
2. 异步平衡方式
3. 异步响应方式

正常响应方式：从站要发送消息的话，需要得到主站的同意；

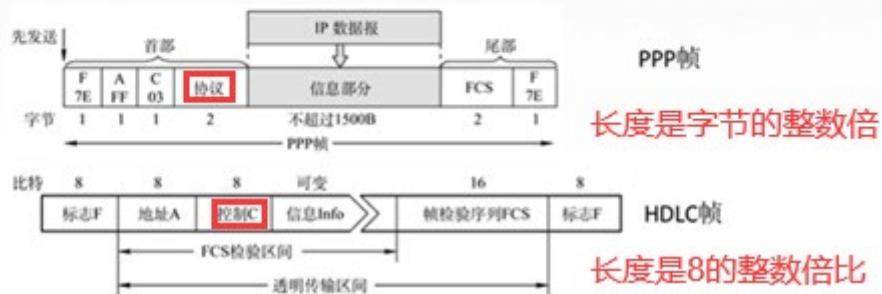
异步平衡方式：每一个复合站都可以对其他站进行数据传输，每个站平等；

异步响应方式：不经过主站同意，从站就进行数据的传输。

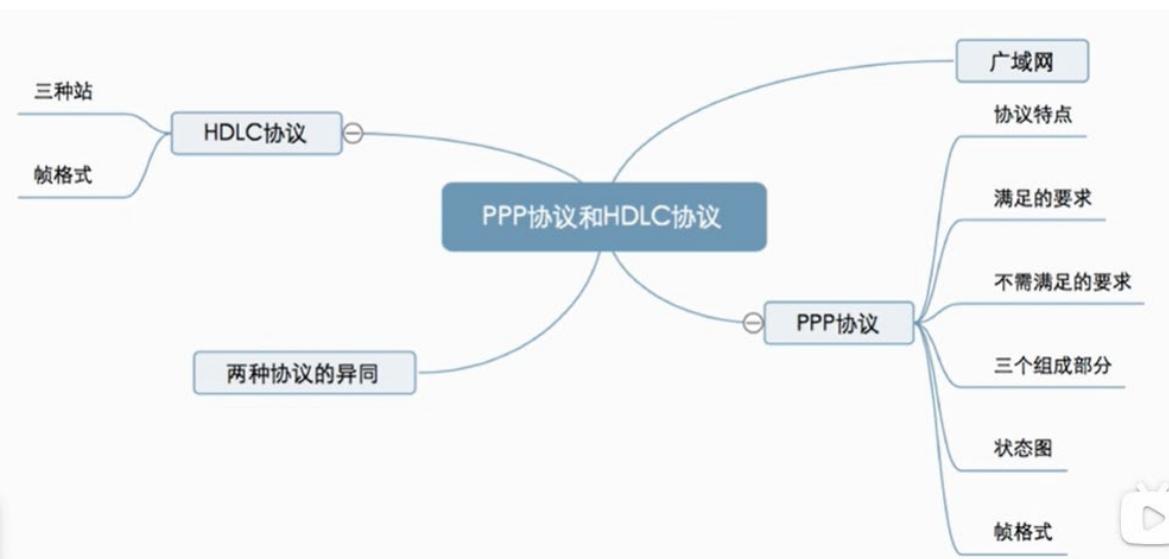


- 同**
- HDLC、PPP只支持全双工链路。
 - 都可以实现透明传输。
 - 都可以实现差错检测，但不纠正差错。

PPP协议	面向字节	2B协议字段	无序号和确认机制	不可靠
HDLC协议	面向比特	没有	有编号和确认机制	可靠

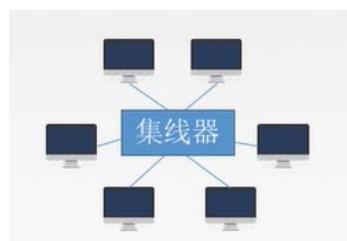


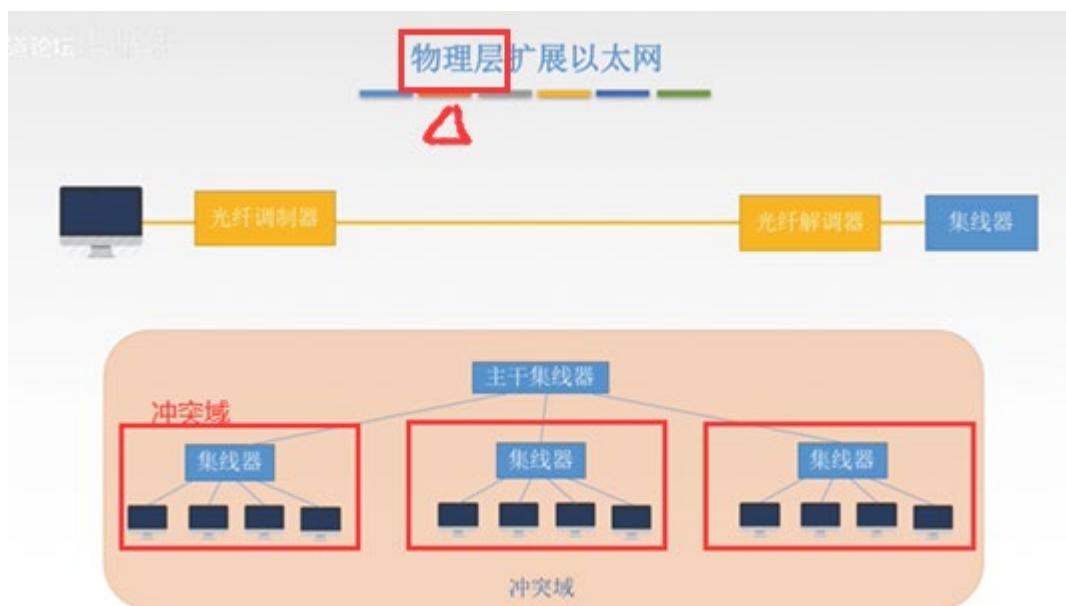
差错控制、流量控制等功能都有 TCP 协议来控制，网络和链路层都是尽最大努力交付的不可靠传输。因为现在对网络传输速率要求很高，在网络和链路层还要进行差错检测和纠正，延时就会增加。所以，使用 PPP 协议较多，很少使用 HDLC 协议。



3.8 链路层设备

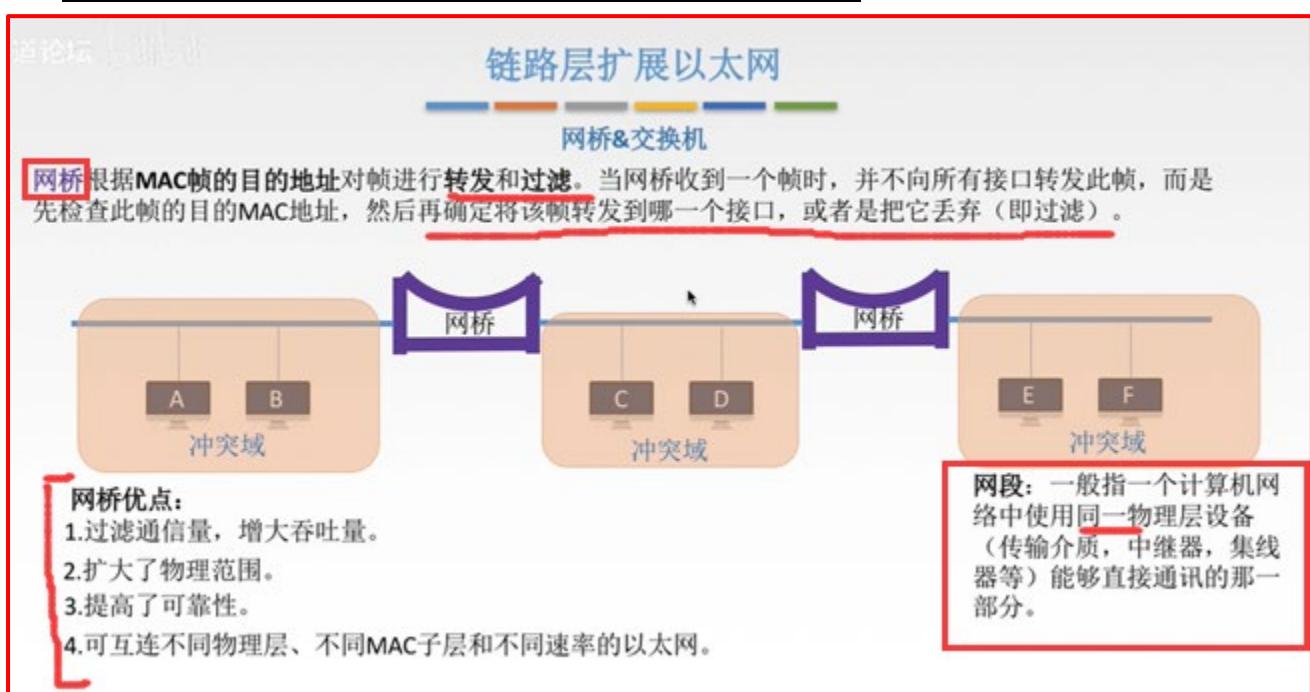
主机和集线器之间距离不能超过 100m，否则失真过大到无法恢复。





为了减少冲突，同时扩大以太网的范围。则在链路层上扩展以外网，即使用网桥或者交换机。

集线器：广播的形式转发数据。而网桥/交换机并非广播式转发。



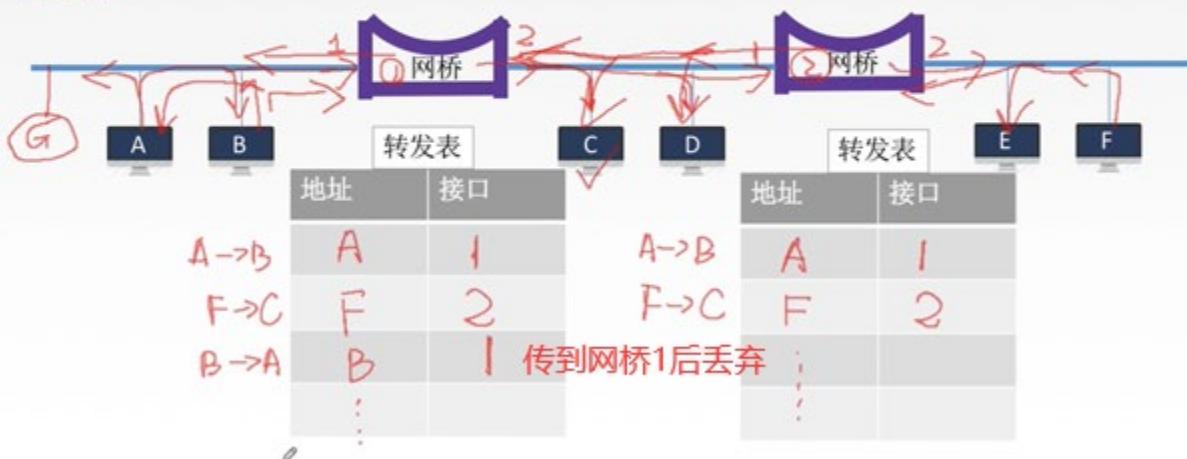
1. A、B 之间通信时，C、D 之间也可以同时通信。每个网段数据传输速率(带宽)为 10Mb/s，则 3 个网段合起来最大吞吐量则有 30Mb/s。

PS: **网段**定义：指一个计算机网络中使用同一物理层设备能够直接通讯的部分网络。同一个网段，通俗理解，就是不用通过路由器就可以相互通信的网络。不管是否具有公网 IP 地址，同一网段内的主机只通过集线器、交换机等设备二层交换，就能相互通信，不用通过路由器进行三层交换或者转发。(详见《局域网、网段和子网》文档)

网桥分类——透明网桥

透明网桥&源路由网桥

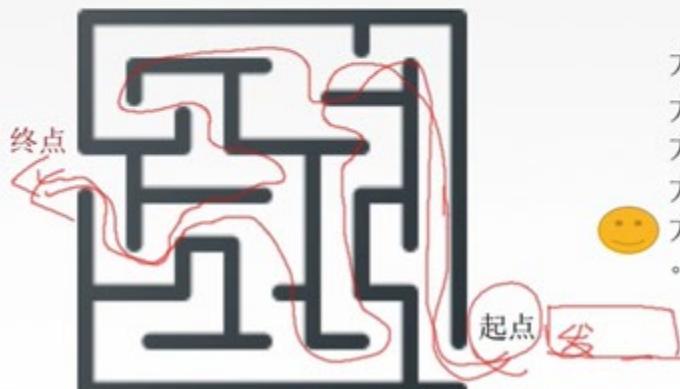
透明网桥：“透明”指以太网上的站点并不知道所发送的帧将经过哪几个网桥，是一种即插即用设备——**自学习**。



网桥分类——源路由网桥

源路由网桥：在发送帧时，把详细的最佳路由信息（路由最少/时间最短）放在帧的首部中。

方法：源站以广播方式向欲通信的目的站发送一个发现帧。



方案1:

方案2:

方案3:

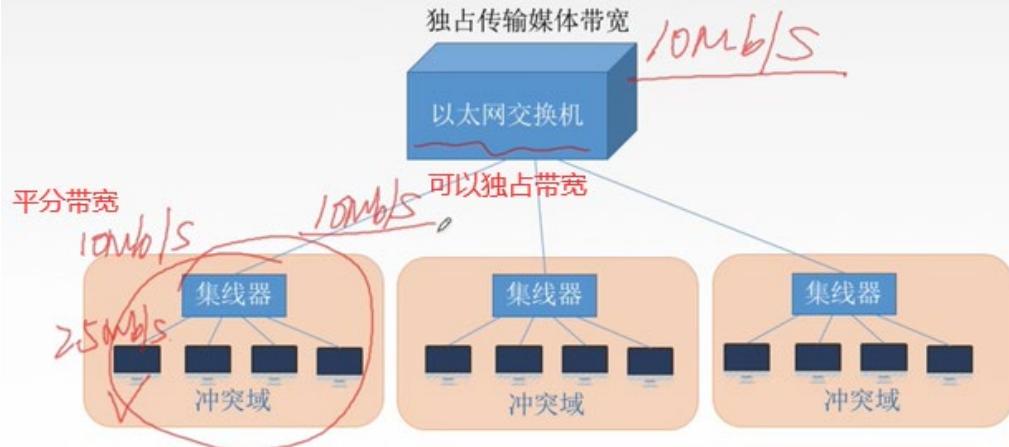
方案4:

方案5:

第一次搜索，所有搜寻成功的路径返回给起点，对比选择路由最少/时间最短的一条，作为该起点终点的帧首部

网桥端口少，所以出现了交换机(多端口的网桥)。

多接口网桥——以太网交换机



以太网交换机的两种交换方式

直通式交换机

查完目的地址 (6B) 就立刻转发。

延迟小，可靠性低，无法支持具有不同速率的端口的交换。



我觉得这样有失公正

存储转发式交换机

将帧放入高速缓存，并检查否正确，正确则转发，错误则丢弃。

延迟大，可靠性高，可以支持具有不同速率的端口的交换。

mac 目的地地址：48 位，6 个字节

王道论坛

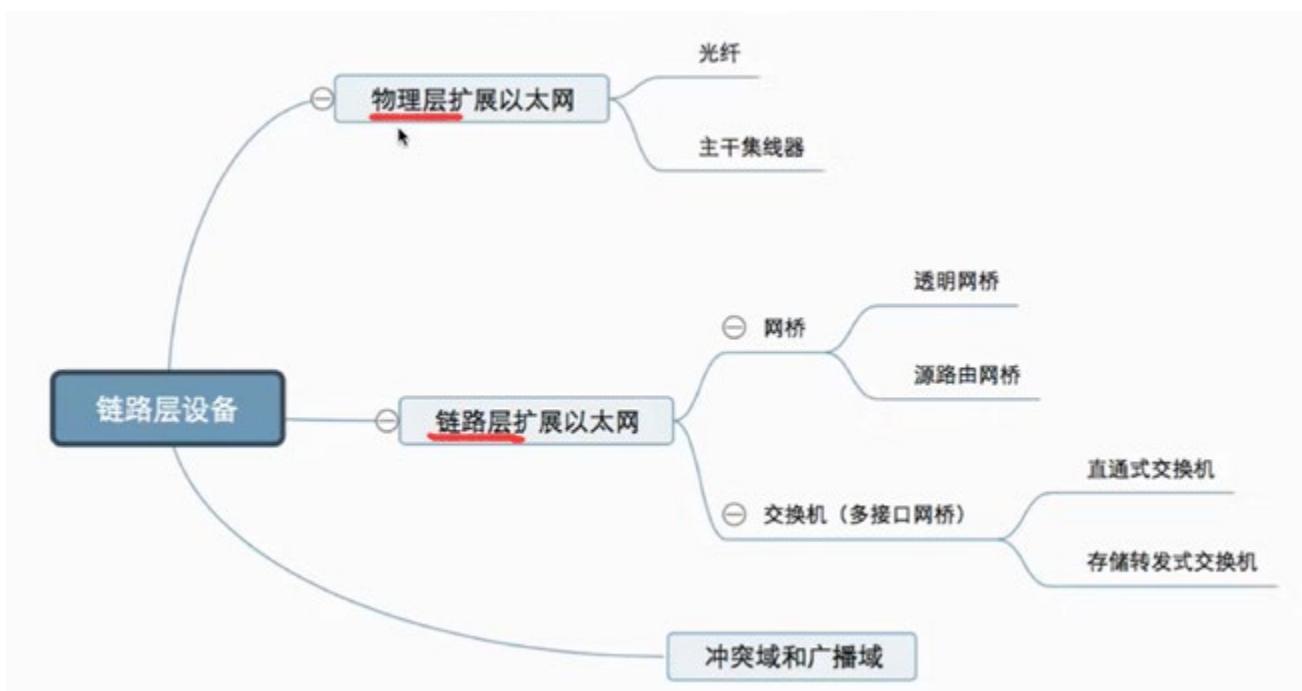
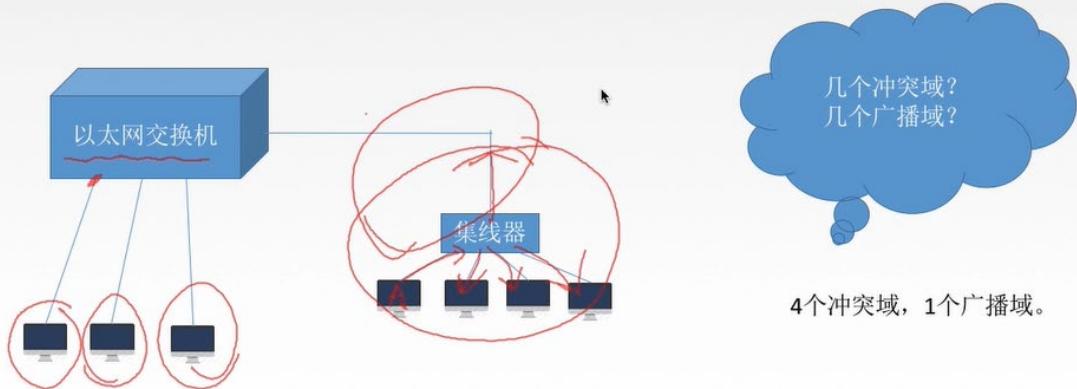
冲突域和广播域

冲突域：在同一个冲突域中的每一个节点都能收到所有被发送的帧。简单的说就是同一时间内只能有一台设备发送信息的范围。

广播域：网络中能接收任一设备发出的广播帧的所有设备的集合。简单的说如果站点发出一个广播信号，所有能接收到这个信号的设备范围称为一个广播域。

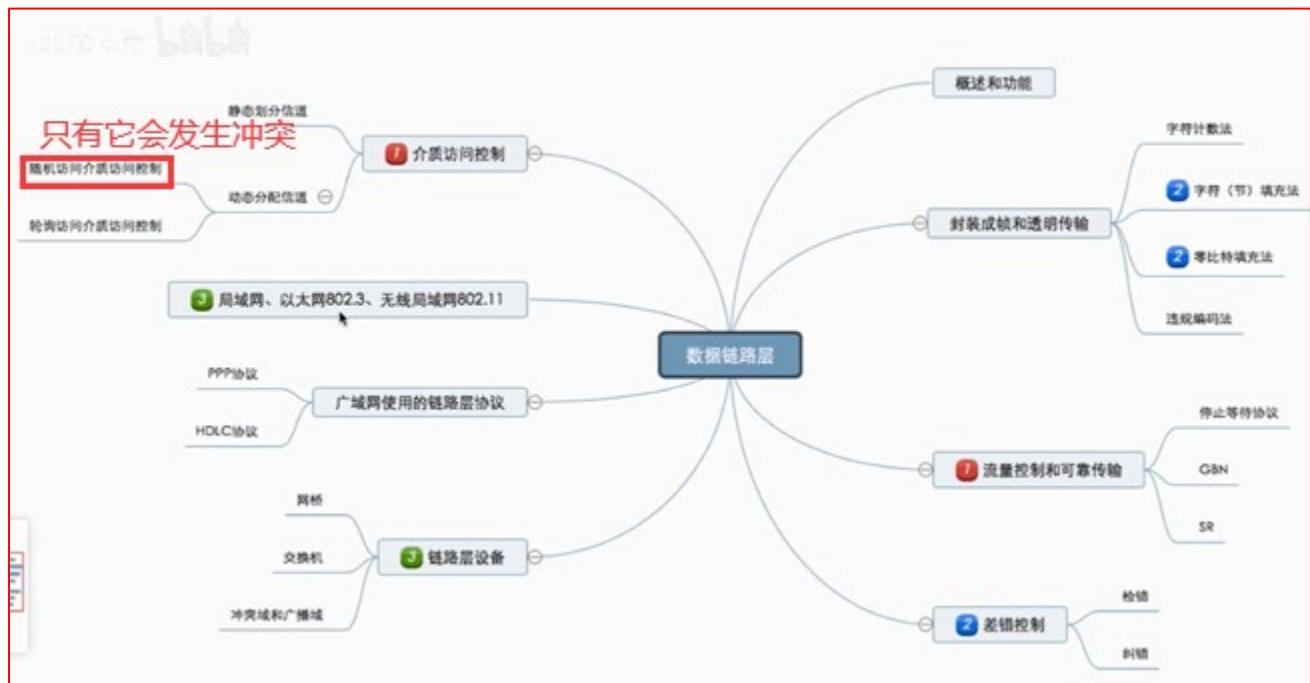
	能否隔离冲突域	能否隔离广播域
物理层设备【傻瓜】 (中继器、集线器)	✗	✗
链路层设备【路人】 (网桥、交换机)	✓	✗
网络层设备【大佬】 (路由器)	✓	✓

冲突域和广播域



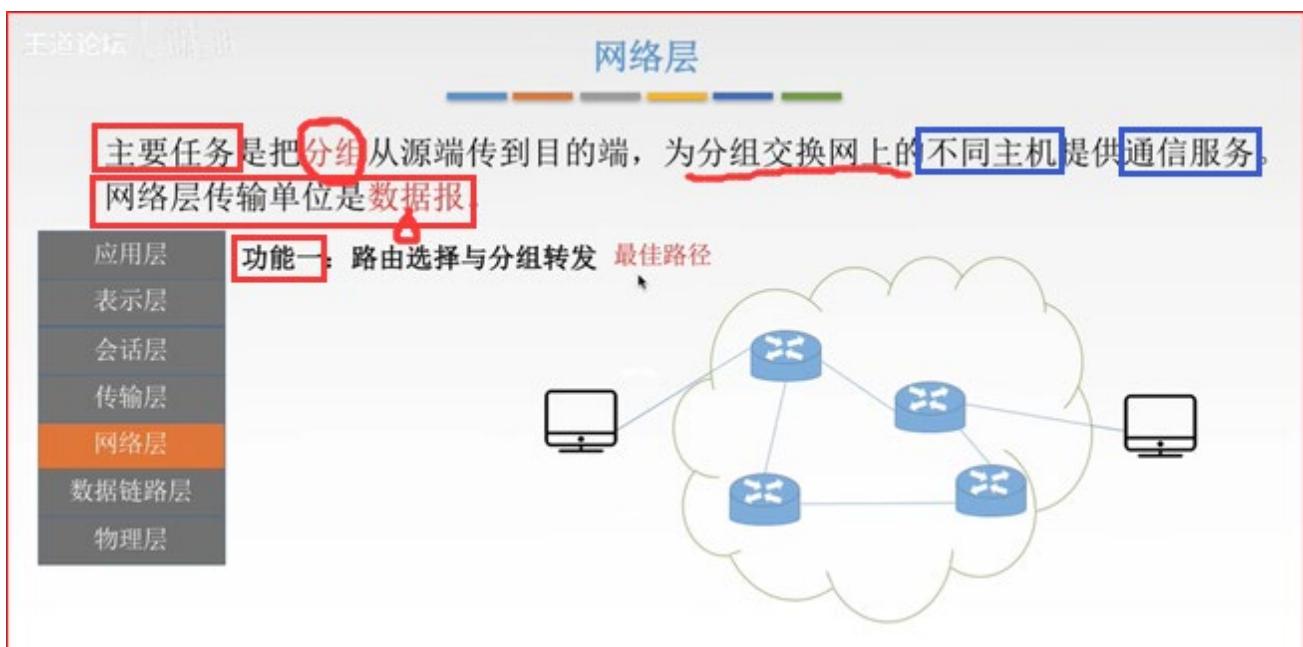
注意：交换机是在一个网络内，路由器是在多个网络之间。

3.9 第三章总结

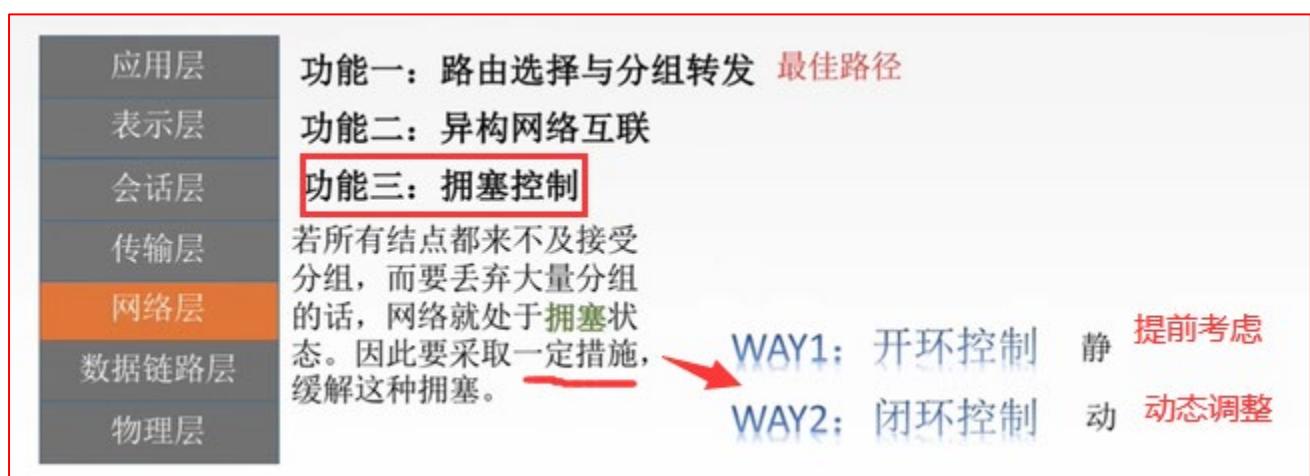


四、网络层

4.1.1 网络功能概述



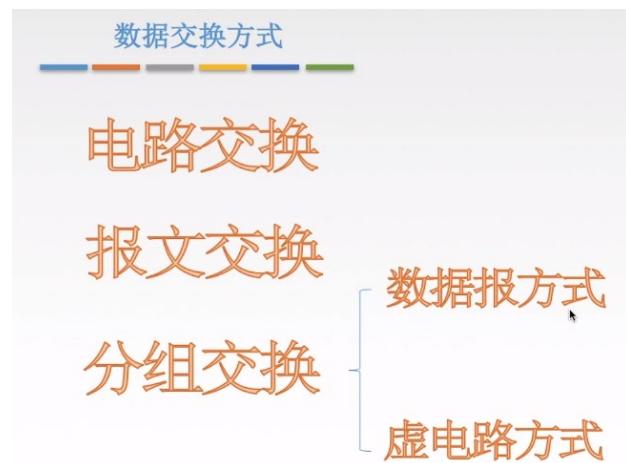
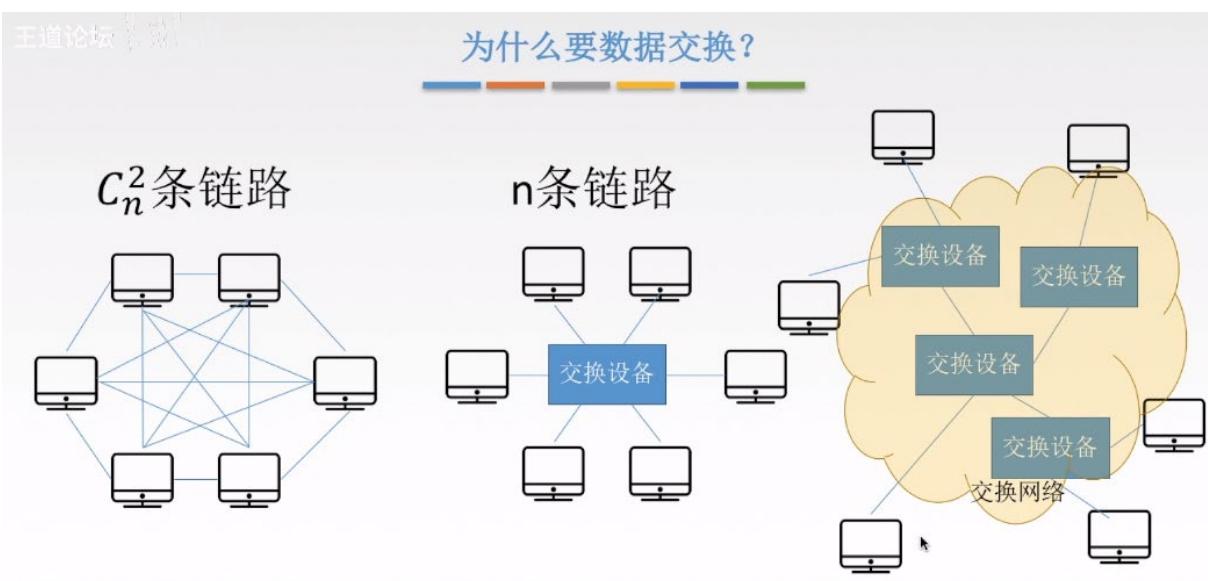
层传输的数据单元，都可以以笼统的用“分组”来表示。(书 P14、P32)

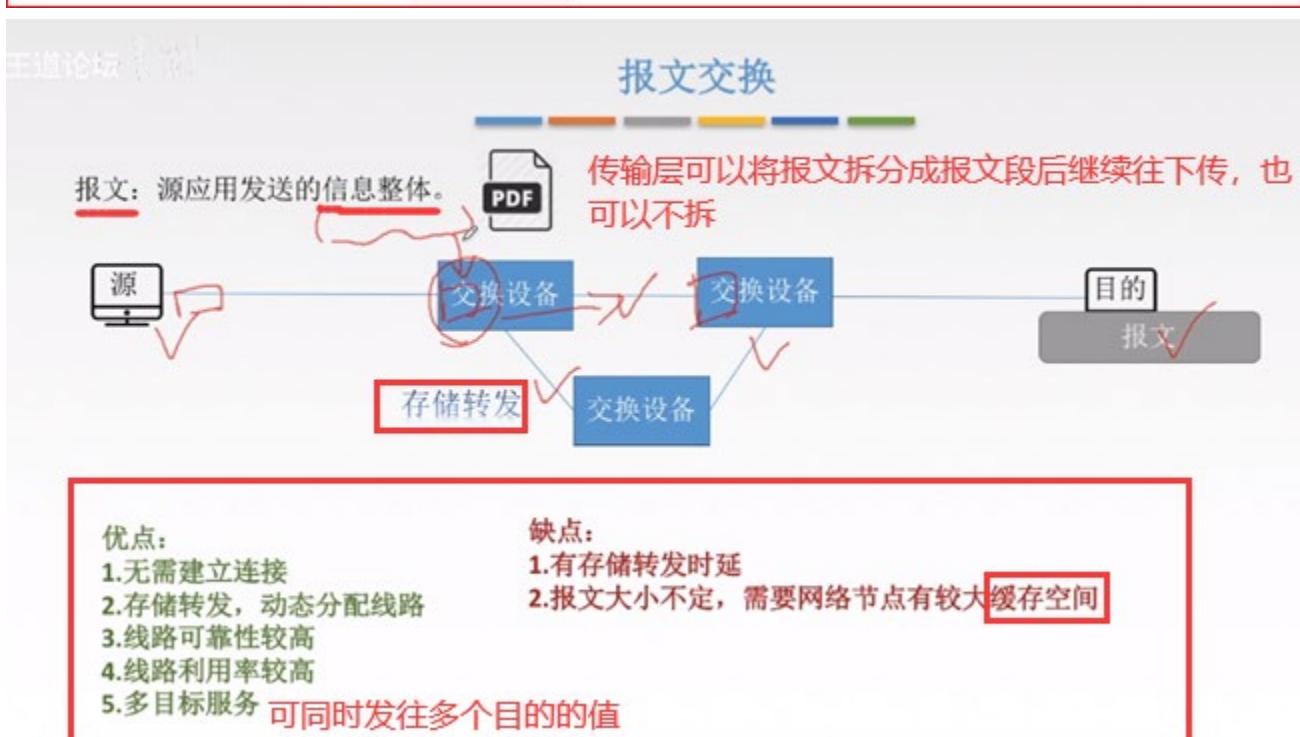
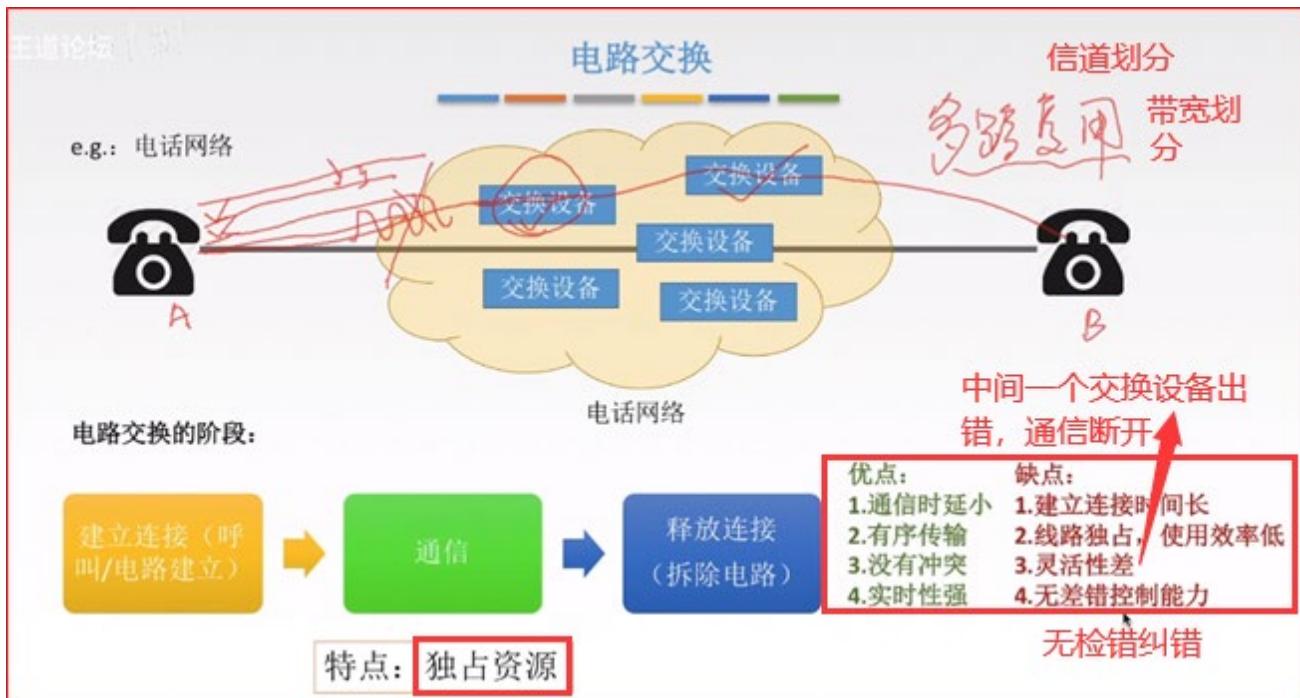


拥塞控制：全局性问题，涉及到网络中所有主机或路由器，以及导致网络传输能力下降的所有因素。

流量控制：发送方发送太快，导致接收方接收不过来。主要是解决发送方发送速率。

4.1.2 数据交换方式——电路交换、报文交换和分组交换





分组交换

分组: 把大的数据块分割成小的数据块。



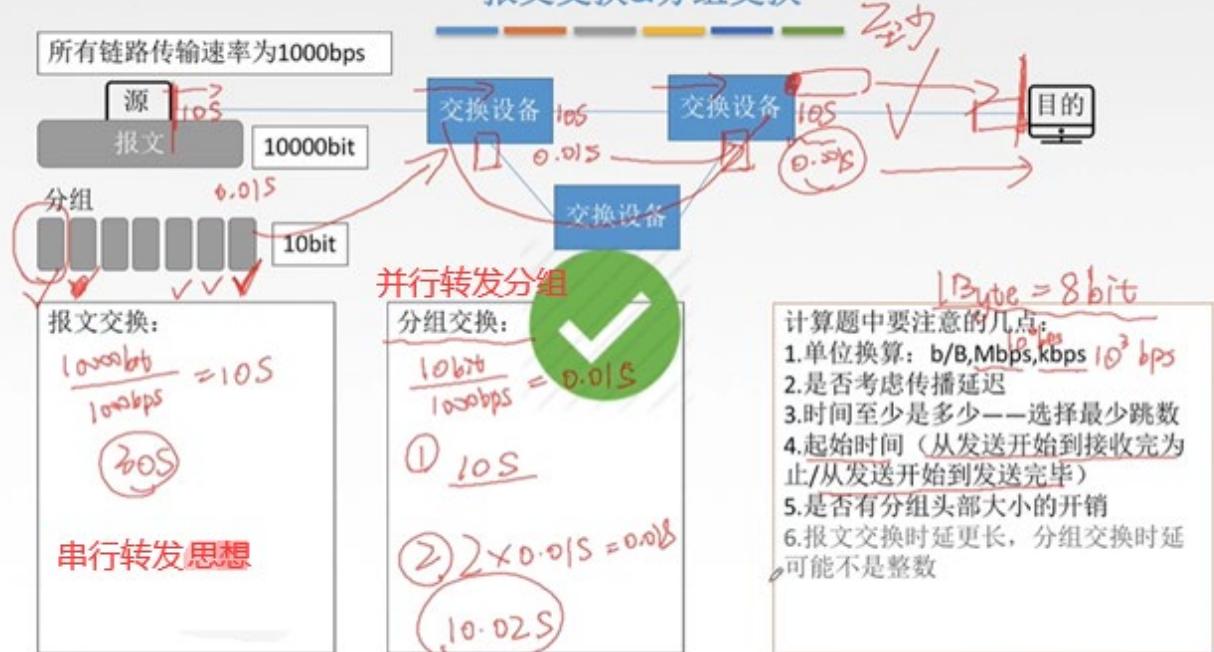
优点:

1. 无需建立连接
2. 存储转发，动态分配线路
3. 线路可靠性较高
4. 线路利用率较高
5. 相对于报文交换，存储管理更容易 需要缓存空间更小

缺点:

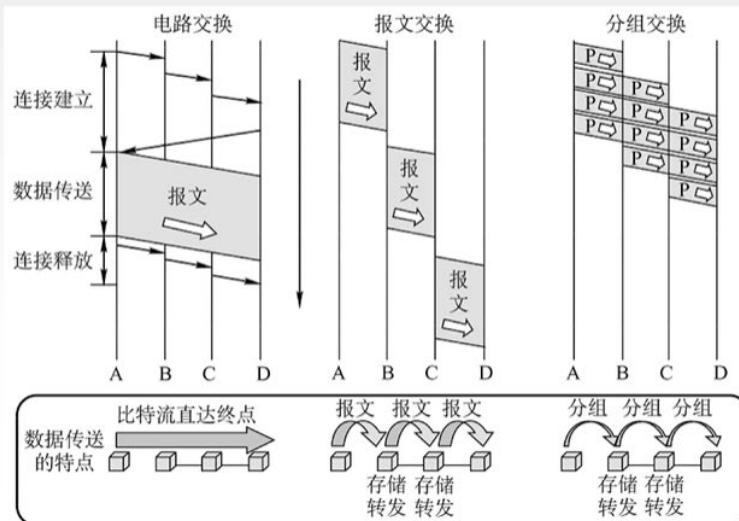
1. 有存储转发时延
 2. 需要传输额外的信息量 首部
 3. 乱序到目的主机时，要对分组排序重组
- 总时延相对报文交换还是会小一些

报文交换&分组交换



报文交换: 在每个交换设备时必须要等到该报文全部到达后才能继续转发出去。

三种数据交换方式比较总结

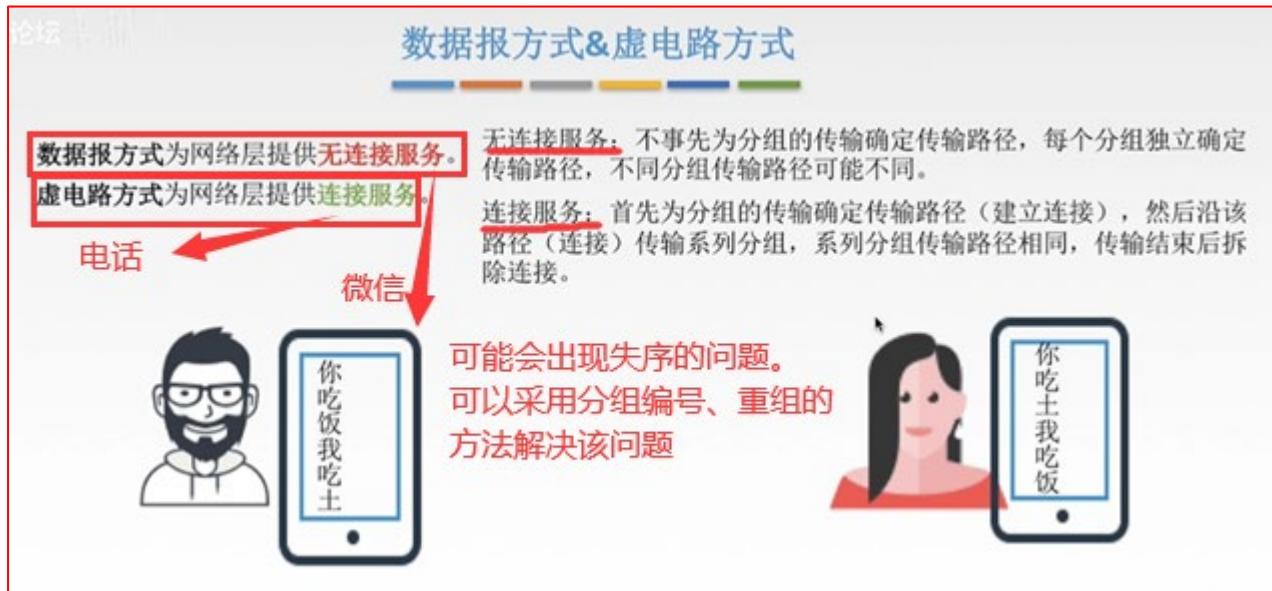


1. 报文交换和分组交换都采用存储转发。

2. 传送数据量大，且传送时间远大于呼叫时，选择 **电路交换**。电路交换传输时延最小。

3. 从信道利用率看，报文交换和分组交换优于 **电路交换**，其中分组交换时延更小。

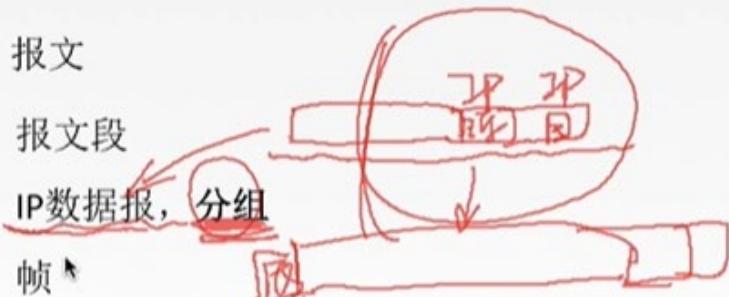
4.1.3 ***数据报与虚电路(接上节)



几种传输单元名词辨析



报文
报文段
IP数据报, 分组
帧
比特流



各层传输单元: 应用层发送**报文**(应用文件); 报文较大时, 传输层分割成**报文段**(较小时, 如只是一句话时就不分割); 网络层封装报文段称为**IP 数据报**, 如果数据报过大(超过 MTU), 对数据报进行切割成为**分组**; 继续往下到数据链路层对分组加头加尾, 封装成**帧**; 再往下到物理层则以**比特流**的方式转成信号的形式在链路上进行传输。

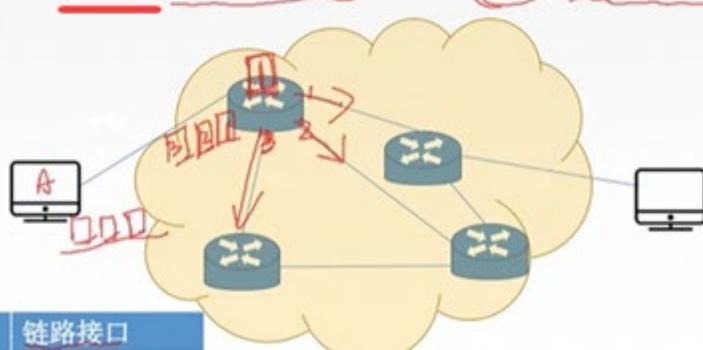
数据报 (因特网在用哦)

无连接

无连接服务: 不事先为分组的传输确定传输路径, 每个分组独立确定传输路径, 不同分组传输路径可能不同。

每个分组携带源和目的地址

路由器根据分组的目的地址转发分组: 基于路由协议/算法构建转发表; 检索转发表; 每个分组独立选路。



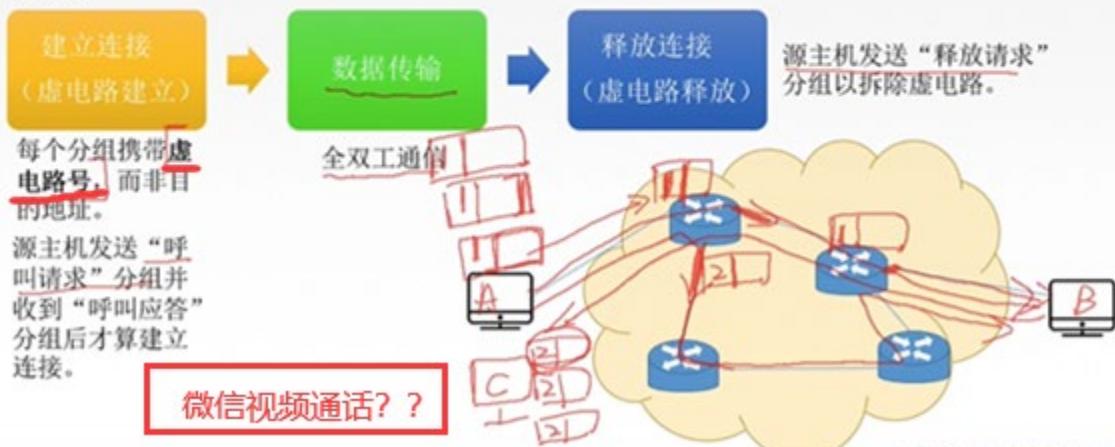
虚电路



虚电路将数据报方式和电路交换方式结合，以发挥两者优点。

虚电路：一条源主机到目的主机类似于电路的路径（逻辑连接），路径上所有结点都要维持这条虚电路的建立，都维持一张虚电路表，每项记录了一个打开的虚电路的信息。

通信过程：



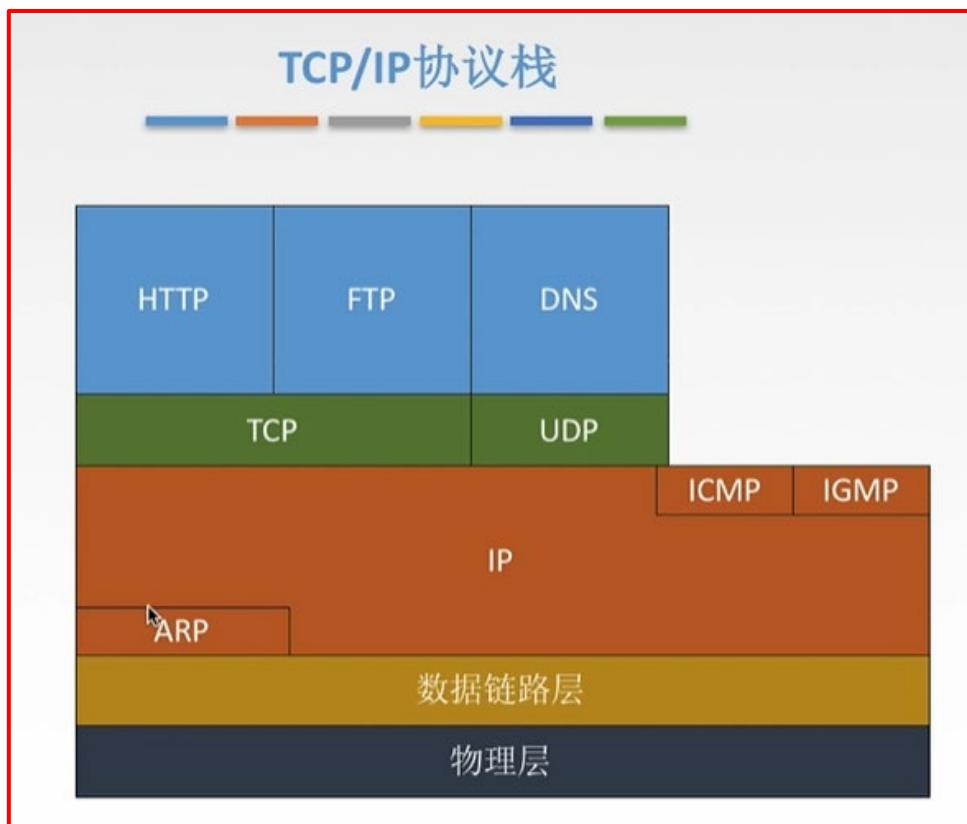
微信视频通话

数据报&虚电路



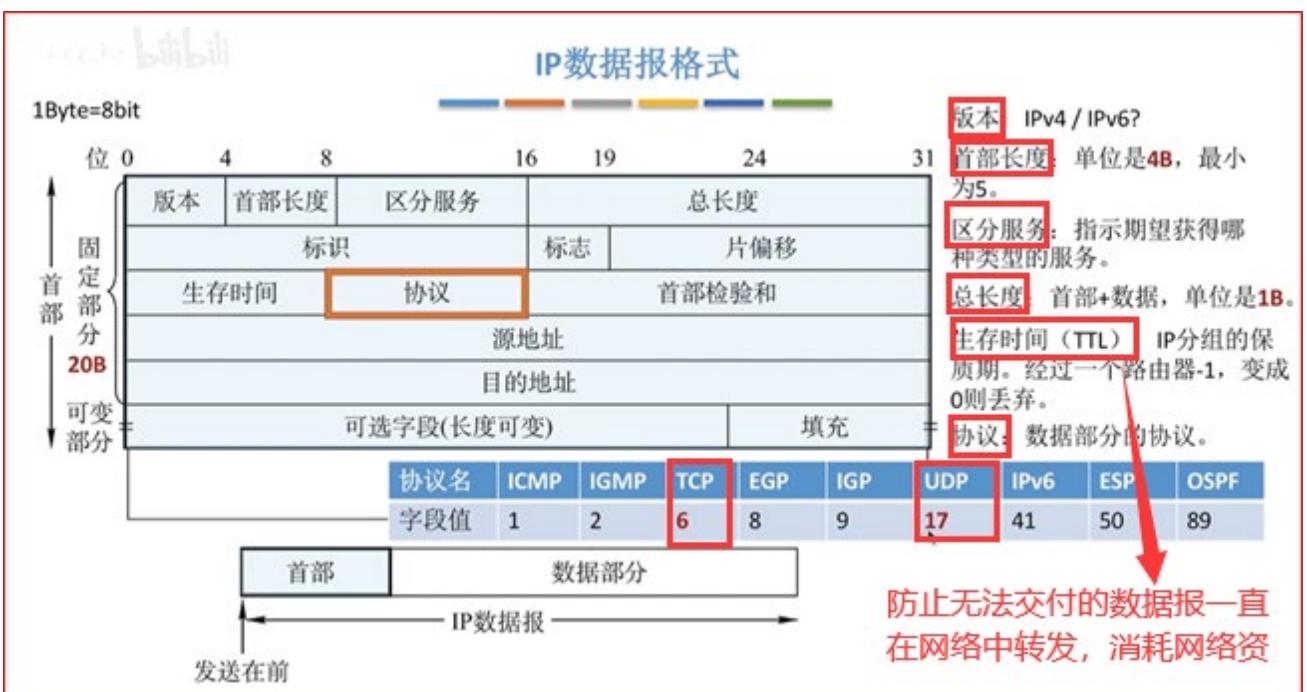
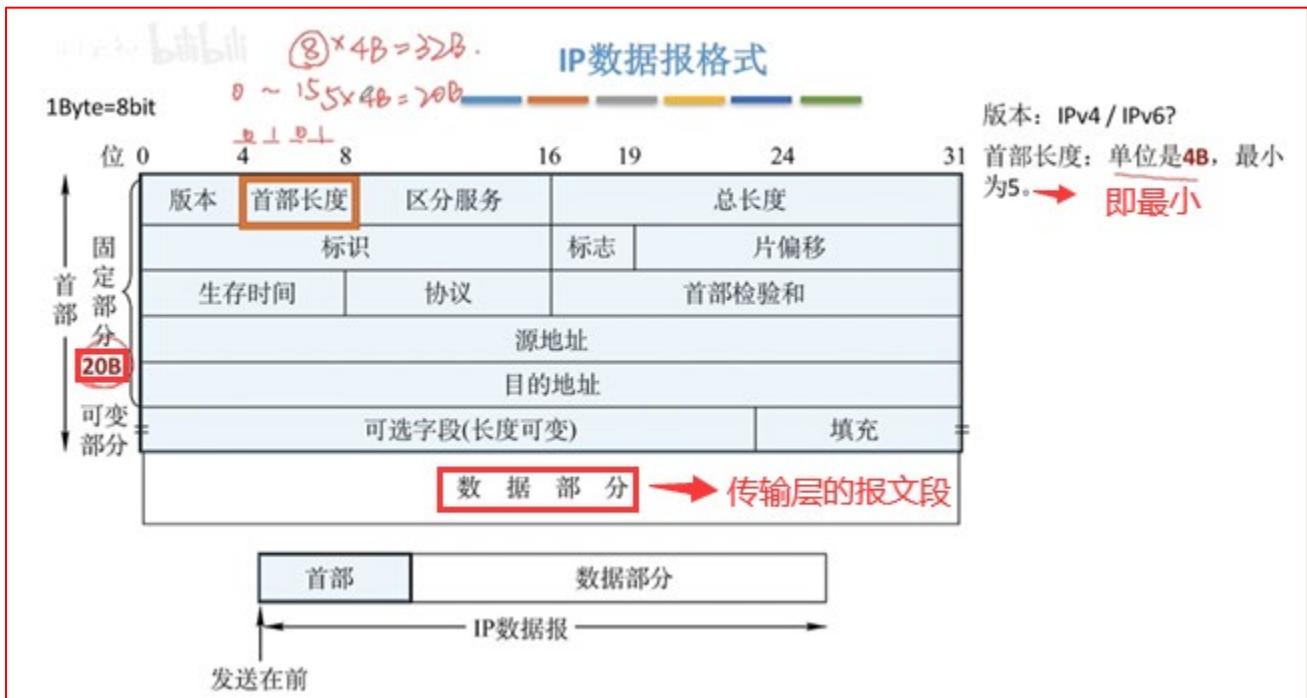
	数据报服务	虚电路服务
连接的建立	不要	必须有
目的地址	每个分组都有完整的目的地址	仅在建立连接阶段使用，之后每个分组使用长度较短的虚电路号
路由选择	每个分组独立地进行 路由选择和转发	属于同一条虚电路的分组按照同一路由转发
分组顺序	不保证分组的有序到达	保证分组的有序到达
可靠性	不保证可靠通信，可靠性由用户主机来保证	可靠性由网络保证
对网络故障的适应性	出故障的结点丢失分组，其他分组路径选择发生变化，可正常传输	所有经过故障结点的虚电路均不能正常工作
差错处理和流量控制	由用户主机进行流量控制，不保证数据报的可靠性	可由分组交换网负责，也可由用户主机负责

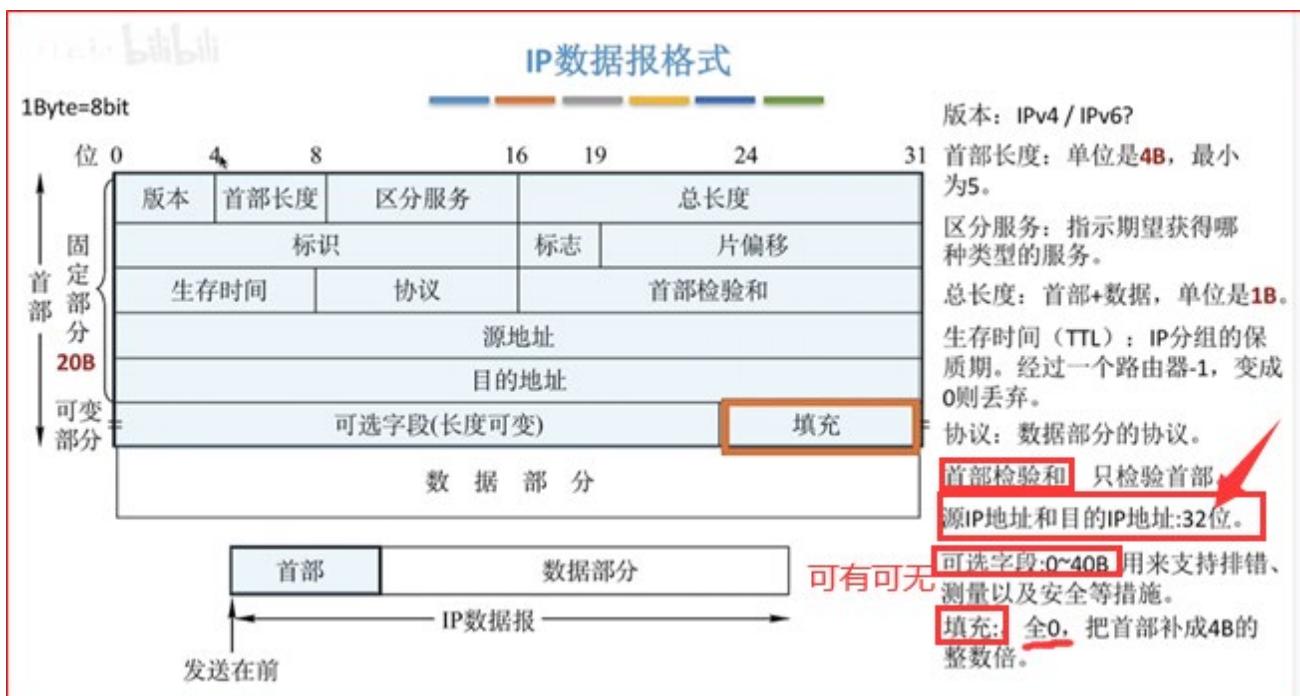
4.2.1 IP 数据报格式



每一个协议的功能和他们之间的关系。



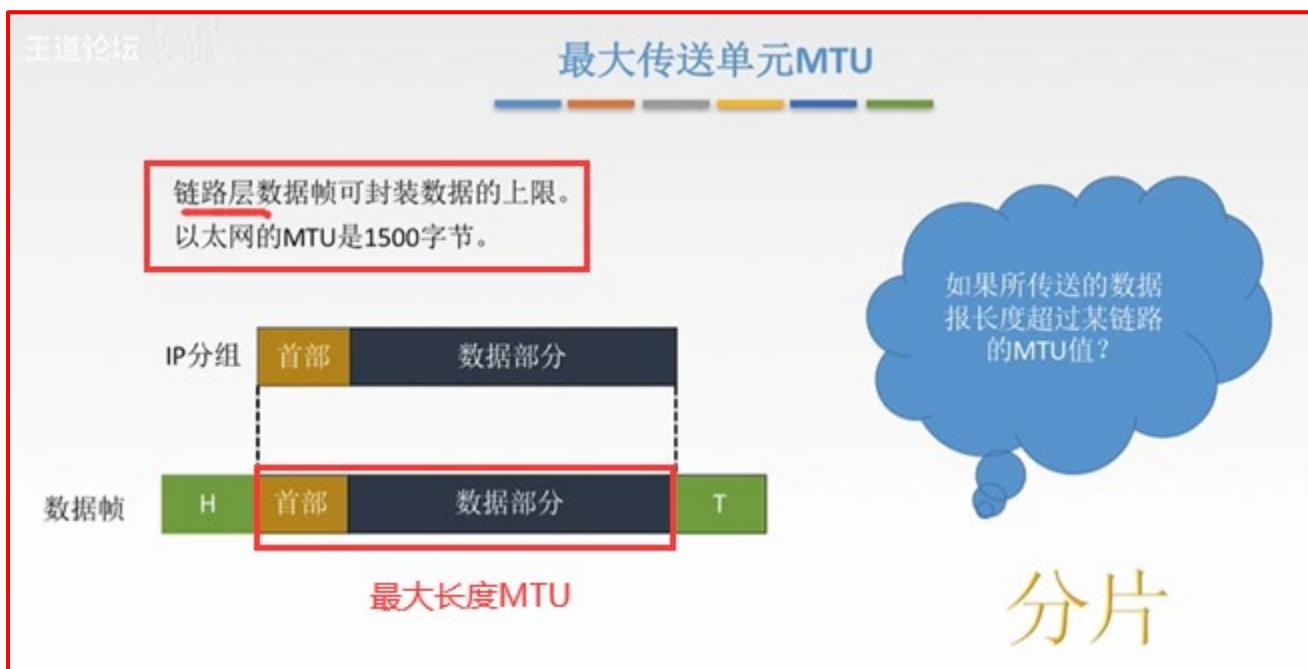


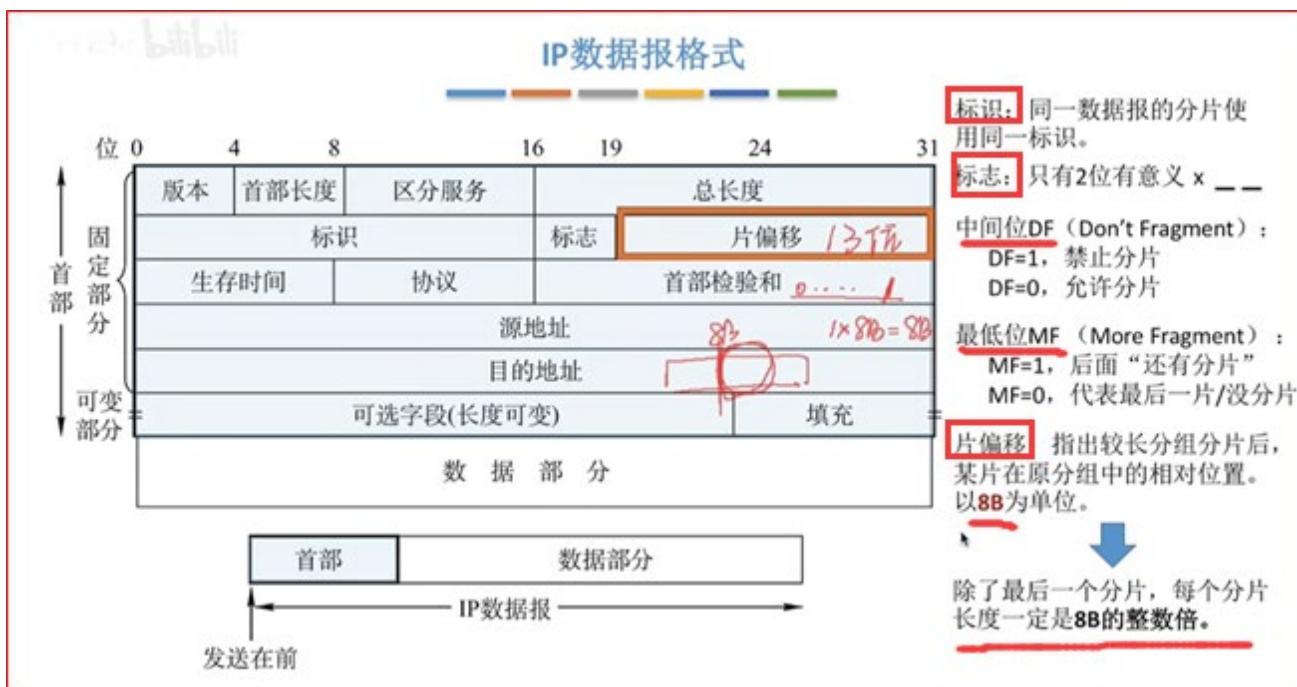


总长度有最大传送单元 MTU 限制。每经过一个路由器都要检验一下首部检验和。因为生存时间、片偏移等字段可能会发生变化，所以检验一下发生变化后数据报有没有出错，出错则丢弃。IP 地址一般指 IPv4 所对应的 IP 地址

4.2.2 IP 数据报分片

链路层可接受的最大传送单元，即 MTU



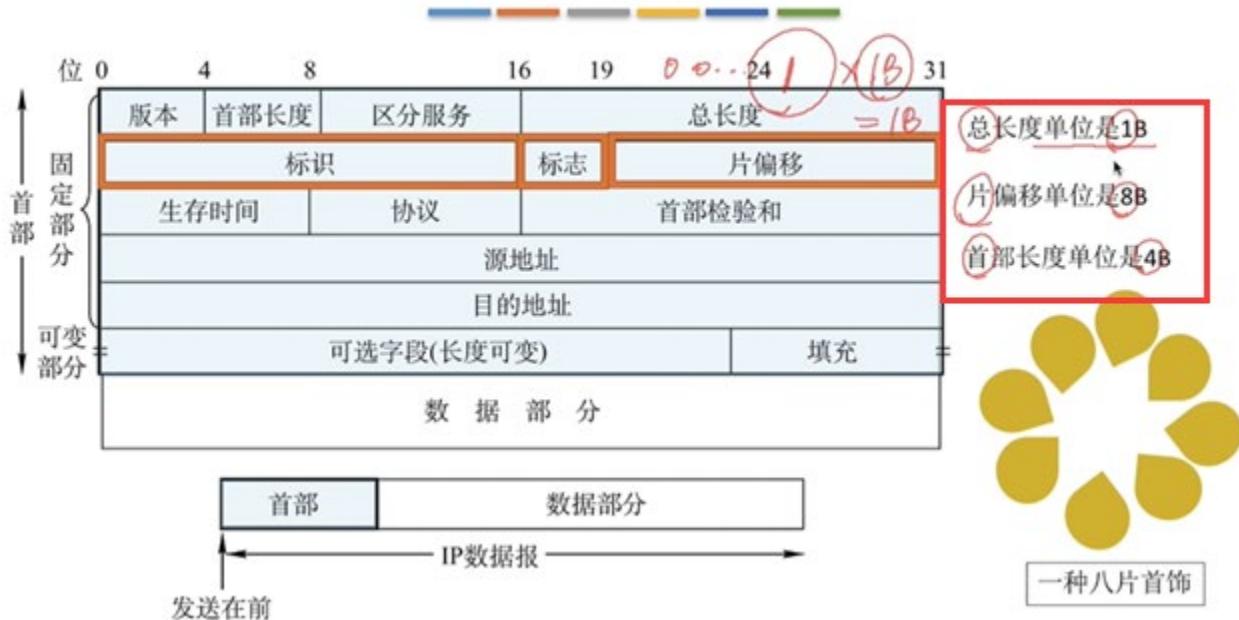


其中, MF 在 DF=0 时才有意义。片偏移值 $\times 8B$: 表示该片在原分组中从 偏移值 $\times 8B$ 位置开始。

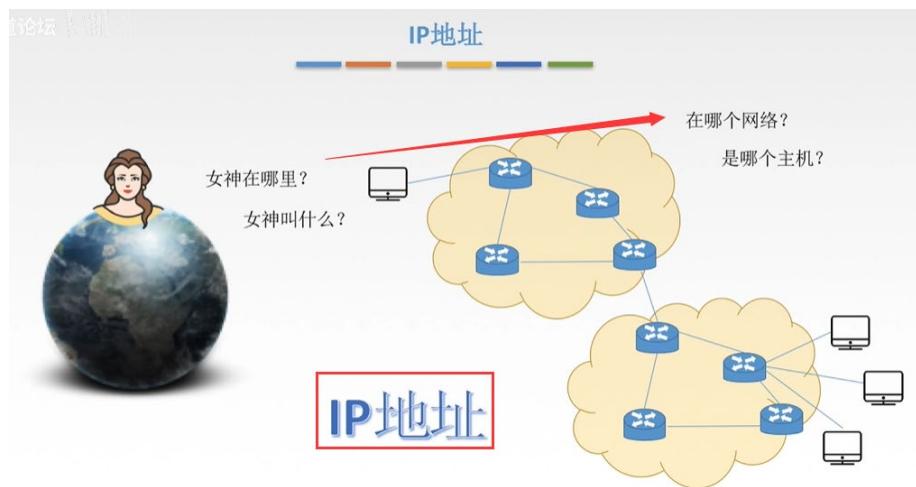
示例:



IP数据报格式



4.2.3 ※※※IPv4 地址



分类的IP地址



身份证号: 110000199601010442

男奇女偶

校验位



IP地址: 全世界唯一的32位/4字节标识符, 标识路由器/主机的接口

IP地址::={<网络号>,<主机号>}

11011111 00000001 00000001 00000001=223.1.1.1

223

*

1

1

点分十进制

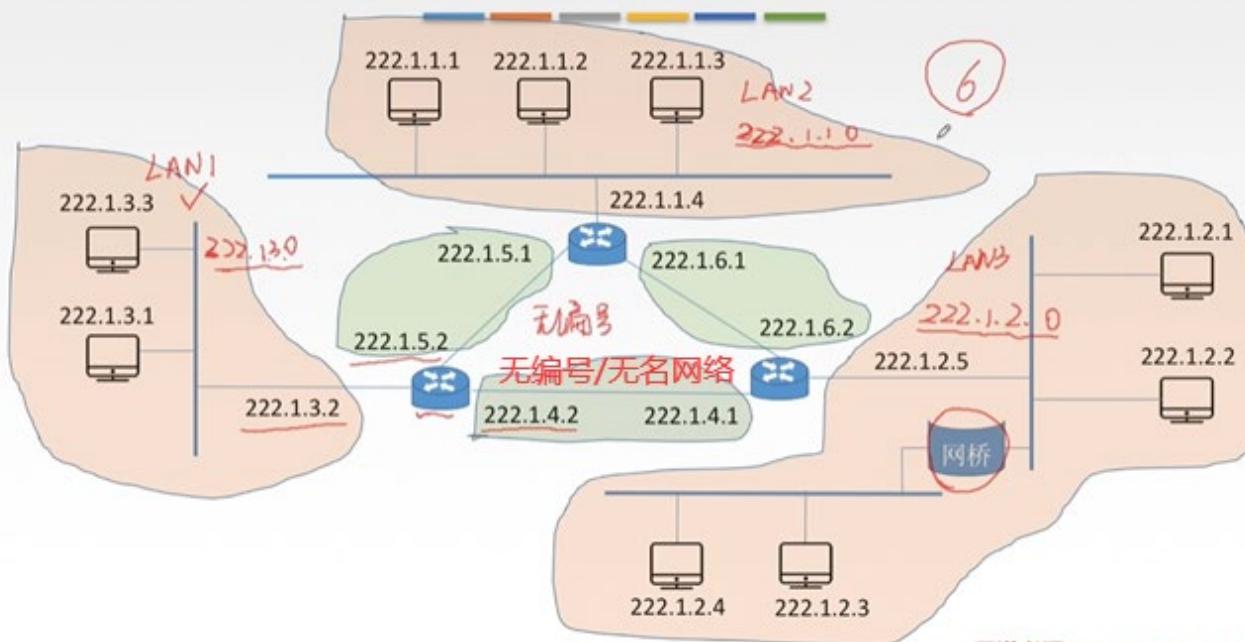
表示在哪个区域

表示是哪台主机

一个主机可以有很多接口, 如有线接口、无线接口。一个路由器也有
很多接口。

IP地址: 标识路由器/主机的接口。一个路由器/主机可以有很多个接口(即网口)。所以在设置IP地址时也要注意网线接的是哪一个网口。? ? ?

互联网中的IP地址



网桥/交换机不能分割广播域, 所以用网桥(链路层设备)相连接的网段还是属于一个局域网, 即只能有一个网络号。 路由器每个接口都会有一个不同的网络号的IP地址, 可以分割广播域, 即不同接口可以连接不同网络。

分类的IP地址

	0	1	2	3	8	16	24	32
A类(1~126)	0	1B	网络号			主机号		
B类(128~191)	1	0	2B	网络号		主机号		
C类(192~223)	1	1	0	3B	网络号		主机号	
D类(224~239)	1	1	1	0	多播地址	一对多通信		
E类(240~255)	1	1	1	1	保留为今后使用			

特殊IP地址

NetID 网络号	HostID主机 号	作为IP分组 源地址	作为IP分组目 的地地址	用途
全0	全0	可以	不可以	本网范围内表示主机，路由表中用于表示默认路由 (表示整个Internet网络)
全0	特定值	不可以	可以	表示本网内某个特定主机
全1	全1	不可以	可以	本网广播地址 (路由器不转发) 受限地址 路由器可以隔离广播域
特定值	全0	不可以	不可以	网络地址，表示一个网络 如上上图 即本网的网络号
特定值	全1	不可以	可以	直接广播地址，对特定网络上的所有主机进行广播
127 A类	任何数 (非全0/1)	可以	可以	用于本地软件环回测试，称为 环回地址

某个主机发送一个数据报，数据报上的目的地址是一个**环回地址**，则该数据报不会进入到网络当中，不会离开主机。主要用于软件(网络层功能)处理、测试。

特殊 IP 地址；私有 IP 地址

私有IP地址

和外部通信：使用NAT技术

这些IP地址在路由器上无效 仅适用于在内部网络中适用，如校园网、单位网。

地址类别	地址范围	网段个数
A类	10.0.0.0~10.255.255.255	1
B类	172.16.0.0~172.31.255.255	16
C类	192.168.0.0~192.168.255.255	256

即网络号个数

王道论坛

分类的IP地址

网络类别	最大可用网络数	第一个可用的网络号	最后一个可用的网络号	每个网络中的最大主机数
A	2^7-2 0,127	1	126	$2^{24}-2$ 全0, 全1
B	$2^{14}-1$ 128.0	128.1	191.255	$2^{16}-2$
C	$2^{21}-1$ 192.0.0	192.0.1	223.255.255	2^8-2

包括私有IP地址

128.0 和 192.0.0 是不指派的。开始确实设计的是 B 地址从 128.1 开始，因为两个 0 开始的地址都被 IANA 所保留了。但现在，128.0 开始的地址也可以指派给电脑网段。如果是为了考试，请按照教学大纲为准，即 B 类地址的网络数是 $2^{14}-1$ 。（即特殊 IP 地址都是不能指派的）

主机号，全 0：本网络；全 1：广播地址。

要熟悉 255, 191, 192, 127 等是十进制数对应的二进制长什么样。

4.2.4 网络地址转换 NAT

私有 IP 地址/本地 IP 地址： 使用于本地网络、专用网络(内网)等，如学校、单位、机房等。在大的互联网、广域网(外网)中路由器及其他主机无法识别出这些 IP 地址。那私有 IP 地址的主机如何同外部网络中的主机通信？？？→ 网络地址转化 NAT 技术

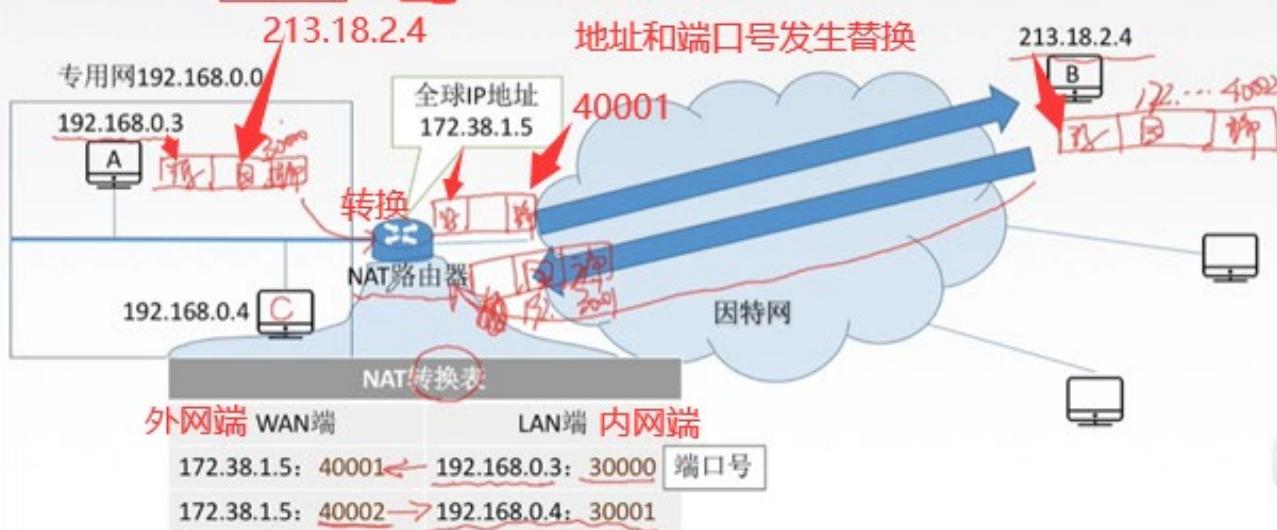
私有IP地址

地址类别	地址范围	网段个数
A类	10.0.0.0~10.255.255.255	1
B类	172.16.0.0~172.31.255.255	16
C类	192.168.0.0~192.168.255.255	256

路由器对目的地址是私有IP地址的数据报一律不进行转发。

网络地址转换NAT

网络地址转换NAT（Network Address Translation）：在专用网连接到因特网的路由器上安装NAT软件，安装了NAT软件的路由器叫NAT路由器，它至少有一个有效的外部全球IP地址。



点分十进制的 IP 地址：端口号

一个端口号唯一标识一个主机中的进程(传输层)

普通路由器仅工作在网络层，而 NAT 路由器转发数据时需要查看和转换传输层的端口号。

问题：能否将拥有私有 IP 地址的主机手动改为外网 IP 地址？？

<https://blog.csdn.net/rcfsyx/article/details/9617445>

https://blog.csdn.net/troublem_aker/article/details/68924553

4.2.5 子网划分和子网掩码

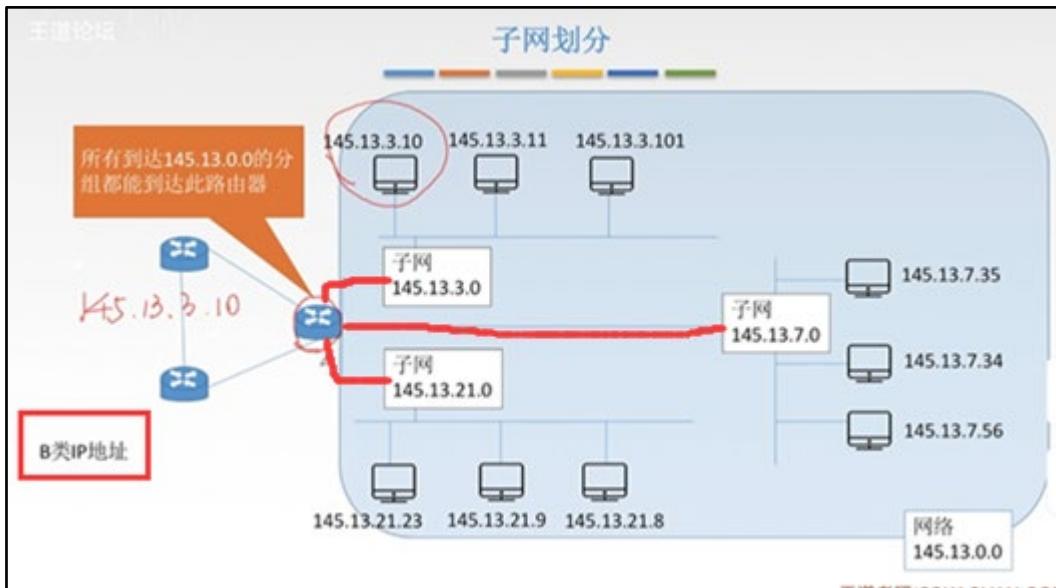


- 若一个公司申请到1个B类网络，则可分配的主机个数有约6万台。但实际是使用不完的，但公司又担心以后扩人，申请C类网络的话不够用。从而造成了IP地址的浪费和不够用。
- 若一个公司需要在一个新的地方马上开通网络，即要向当地ISP(因特网服务提供商)申请一系列新的IP地址。从而操作十分麻烦。怎么随时灵活增加自己单位的网络呢？→**子网划分**



子网号是某单位自己划分，可以一位都没有，但是必须保证至少有两位主机号（因为只剩一位的话，只有0和1，主机号全0全1是有特殊含义的）

下面以子网号占 8 位为例：



不同子网之间需要通过路由器或三层交换机设备来通信的。

此时，目的地址为 145.13.3.10 的数据报到达路由器，怎么找到相应的主机呢？ \rightarrow 子网掩码



子网掩码中：网络号对应全 1，子网号对应全 1，主机号对应全 0。

由此也可以由子网掩码看出子网号和主机号各占几位，如下题示例求子网的网络地址。

子网掩码习题

已知IP地址是141.14.72.24，子网掩码是255.255.192.0，求网络地址。如果子网掩码是255.255.224.0，求网络地址。

手写分析：

- IP地址：141.14.72.24
- 子网掩码：255.255.192.0
- 二进制表示：
 - IP地址：01000100 01110000 00000000 00010000
 - 子网掩码：11111111 11111111 11111111 00000000
- 结果：141.14.64.0

子网号占2位，主机号占14位

子网掩码表：

10000000	128
11000000	192
11100000	224
11110000	240
11111000	248
11111100	252
11111110	254
11111111	255

新版反馈 [新版](#) [反馈](#)

回到旧版 [回到](#) [旧版](#)

此题：同样的IP地址和不同的子网掩码可以相与得到相同的**子网网络地址**。但是达到的效果不同：子网号所占的位置不同，则可划分的子网个数不同及每个子网中可使用的最大主机数也不一样（详见后续CIDR）。

子网掩码习题

某主机的IP地址是180.80.77.55，子网掩码为255.255.252.0。若该主机向其所在子网发送广播分组，则目的地址可以是（ ）。

A. 180.80.76.0 B. 180.80.76.255 C. 180.80.77.255 D. 180.80.79.255

手写分析：

- IP地址：180.80.77.55
- 子网掩码：255.255.252.0
- 二进制表示：
 - IP地址：01001000 00001100 01000111 00110111
 - 子网掩码：11111111 11111111 11111100 00000000
- 结果：180.80.78.0
- 注释：子网2位 主机10位

子网网络地址：180.80.76.0

使用子网时分组的转发

路由器转发分组的算法：

1. 提取目的IP地址
2. 是否直接交付
3. 特定主机路由
4. 检测路由表中有无路径
5. 默认路由 0.0.0.0
6. 丢弃，报告转发分组出错

路由器转发分组的算法：

1. 提取目的IP地址
2. 是否直接交付
3. 特定主机路由
4. 检测路由表中有无路径
5. 默认路由 0.0.0.0
6. 丢弃，报告转发分组出错

路由表中：

1. 目的网络地址
2. 目的网络子网掩码
3. 下一跳地址

子网掩码1 145.13.3.10 子网掩码1 145.13.3.11 子网掩码1 145.13.3.101

子网掩码1 子网 145.13.3.0

子网掩码1 子网 145.13.21.0

子网 145.13.7.0

子网 145.13.7.35

子网 145.13.7.34

子网 145.13.7.56

网络 145.13.0.0

145.13.21.23 145.13.21.9 145.13.21.8

第二步：直接交付：目的地址是否在与该路由器直接相连的一个网络（子网）。**间接交付：**还需要转1个或多个路由器。**判断方法：**把目的地址与该路由器相连的所有子网的子网掩码（可以不同）相与，看是否和子网的

网络地址相同。相同则将分组转入该子网(直接交付)。

由子网的网络地址可以看出：上图中子网号至少占了 8 位！

第三步：在路由表中查找是否有该目的 IP 地址，有则按照该地址所要求的路径走。即特定主机路由：分组的目的地址和路由表中某一 IP 地址相同。

第四步：将目的 IP 地址和路由表中所有的子网掩码相与，得到的子网网络地址若与路由表中某个目的网络地址相同，则按照该地址所规定的下一个路径走。

第五步：走默认路由，将该分组发给另外的路由器，重复以上步骤，直至找到目的网络(子网)。但是转发也有限制(生存时间)，超过生存时间还未找到，则丢弃该分组。

4.2.6 ※※※无分类编址 CIDR ——构成超网

虽然划分子网技术能缓解 IP 地址资源枯竭的现状，但是仍不能解决问题。

前言：公司内不同部门可以根据不同人数，设置不通长度的子网掩码(最大主机数不同)，进一步减少 IP 地址的浪费。即**变长子网掩码**。

王道论坛

无分类编址CIDR

无分类域间路由选择CIDR:

1.消除了传统的A类，B类和C类地址以及划分子网的概念。
2.融合子网地址与子网掩码，方便子网划分。

CIDR把网络前缀都相同的连续的IP地址组成一个“CIDR地址块”。

128.14.35.7/20是某CIDR地址块中的一个地址

二进制: 10000000 00001110 00100011 00000111
最小地址: 10000000 00001110 00100000 00000000 128.14.32.0 **本网络**
最大地址: 10000000 00001110 00101111 11111111 128.14.47.255 **广播地址**
地址块: 128.14.32.0/20 “/20地址块”
地址掩码(子网掩码): 简写
11111111 11111111 11110000 00000000

CIDR记法: IP地址后加上“/”，然后写上网络前缀(可以任意长度)的位数。 e.g. 128.14.32.0/20

无分类编址CIDR

(²⁴ 192.199.170.82/27) ⁰¹¹⁰⁰¹⁰ $2^5 = 32$

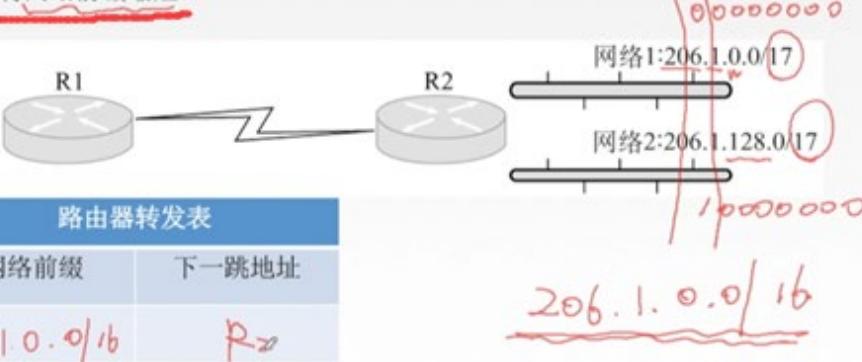
最小 192.199.170.64

最大 192.199.170.95

构成超网

将多个子网聚合成一个较大的子网，叫做构成超网，或路由聚合。

方法：将网络前缀缩短。



实际发送数据时，将分组的目的地址与转发表中网络前缀相匹配。因为构成超网时将网络范围扩大了，所以可能出现匹配时，转发表中有几行均匹配得上，此时就使用最长前缀匹配。

最长前缀匹配

使用CIDR时，查找路由表可能得到几个匹配结果，应选择具有最长网络前缀的路由。前缀越长，地址块越小，路由越具体。

路由器R0的路由表见下表：若进入路由器R0的分组的目的地址为132.19.237.5，请问该分组应该被转发到哪一个下一跳路由器（B）。

- A. R1 B. R2
C. R3 D. R4

目的网络	下一跳
132.0.0.0/8	R1
132.0.0.0/11	R2
132.19.232.0/22	R3
0.0.0.0/0	R4

00010011 11110011 11110011
132.0.0.0 132.0.0.0 132.19.236.0
132.0.0.0 132.0.0.0 132.19.236.0

因为构成超网时，是将范围扩大。所以，转发查找时应该找最具体的，才最接近目的地址。

在CIDR地址基础上，再使用子网划分

某网络的IP地址空间为192.168.5.0/24，采用定长子网划分，子网掩码为255.255.255.248，则该网络中的最大子网个数、每个子网内的最大可分配地址个数分别是（ ）。

- A. 32, 8 B. 32, 6 C. 8, 32 D. 8, 30

在最后8位采用子网划分

CIDR中，子网号可全0全1

11111000
5 3
25 2-2

CIDR 优势: 1. 变长子网掩可以更灵活地分配 IP, 更能充分利用 IP 资源, 有效节省。

2. 由于路由聚合, 可以提高网络性能。

没 cidr 就只能按 ABC 类的掩码划分, A 类掩码范围只能在/8 以上, b 类是/16, c 类是/24。也就是无论如何你一个 c 类地址通过汇总没办法出现掩码小于 24 位的情况, 拿 192.168.1.0-192.168.3.0, 这三个你要是没有 cidr 就没办法就行汇总, 因为汇总完是 192.168.0.0/22, 而 c 类默认掩码是/24, 通告路由的时候是通告三条路由, 用 cidr 可以汇总为一条路由 192.168.0.0/22 来进行通告。

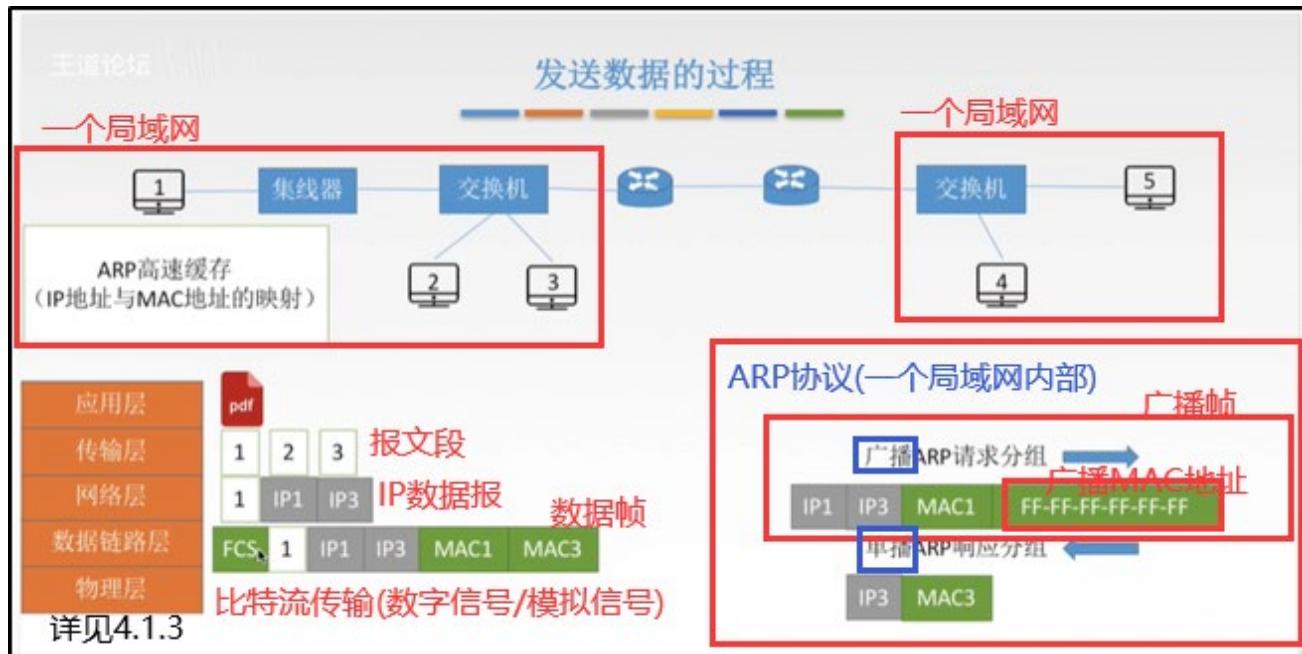
3. IPv6 一节也有相关说明。

4.2.7 ARP 协议(Address Resolution Protocol, 地址解析协议)——网络层协议

每个**主机和路由器**都有一个 ARP 高速缓存。主机的 ARP 高速缓存中存储**局域网内部所有主机或连在该局域网的路由器端口**的 IP 地址和 MAC 地址的映射关系。

网络层的 IP 数据报(分组)在封装成数据链路层的数据帧时, 就需要源 MAC 地址和目的 MAC 地址。当 ARP 高速缓存中也找不到目的 IP 与目的 MAC 地址的映射时, 就需要使用 **ARP 协议**去寻找 MAC 地址。

例一如下, 1 号主机想和 3 号主机通信(同一局域网内), 但是在封装 IP 数据报时, 在 1 号主机的 ARP 高速缓存中找不到 3 号主机 IP 和 MAC 地址的映射, 所以需要使用 ARP 协议:

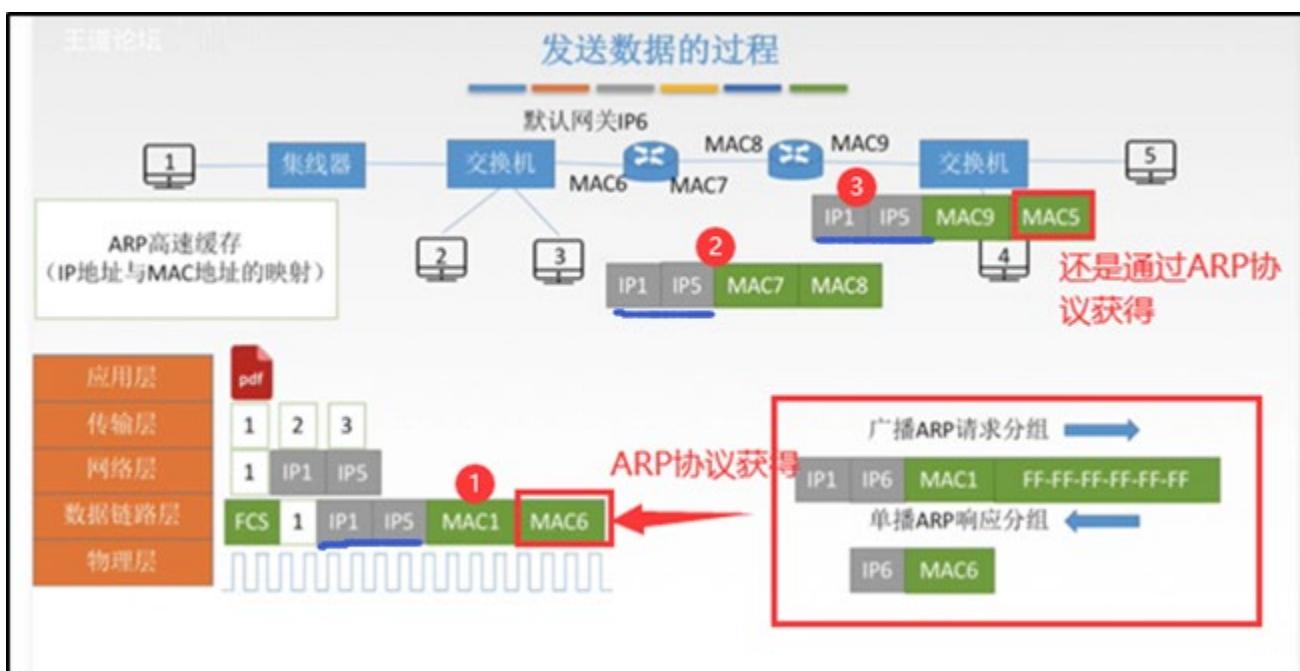
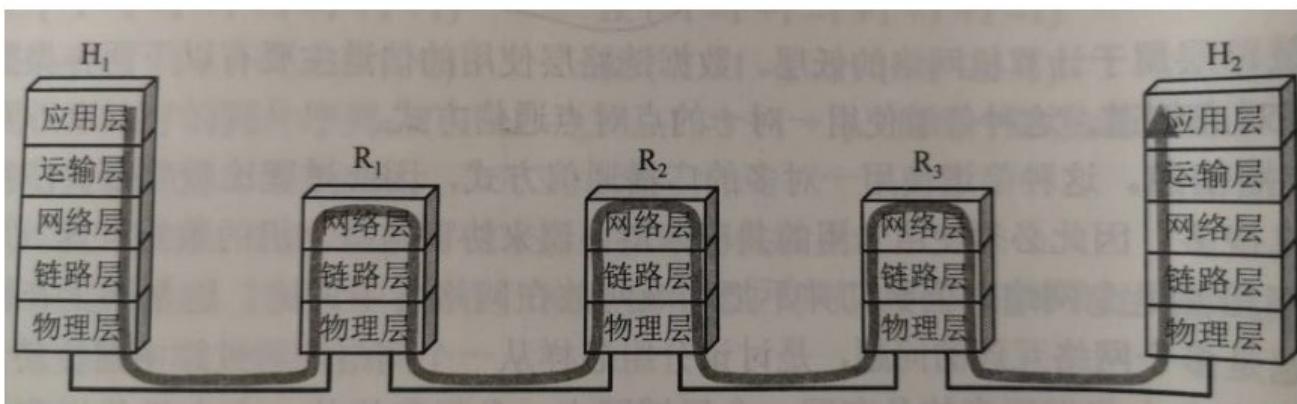


使用广播 MAC 地址, 构成**广播帧**, 此时交换机才会广播转发。

例二如下, 1 号主机想和 5 号主机通信(不同局域网), 在 1 号主机的 ARP 高速缓存中找不到 5 号主机 IP 和 MAC 地址的映射, 所以先用自己的子网掩码和目的 IP 地址相与, 发现子网号不同, 不在一个子网/网段内。所以使用 **ARP 协议**去寻找自己的默认网关的 MAC 地址, 以往外通信。注: 这里不考虑 NAT

MAC7->MAC8: 若是点对点的, 可以使用 PPP 协议(目的 MAC 地址全 1), 就不需要 ARP 协议; 若有多个路由器, 则还是使用 ARP 协议获取目的 MAC 地址。

路由器怎么通过 ARP 获取下一跳路由器的 MAC 地址? ? ?



网关：通常指路由器一个端口。

ARP 协议：IP 地址转 MAC 地址。

RARP 协议：MAC 地址转 IP 地址。

直论坛

ARP协议

由于在实际网络的链路上传送数据帧时，最终必须使用MAC地址。

ARP协议：完成主机或路由器IP地址到MAC地址的映射。解决下一跳走哪的问题

ARP协议使用过程：

检查ARP高速缓存，有对应表项则写入MAC帧，没有则用目的MAC地址为FF-FF-FF-FF-FF-FF的帧封装并**广播ARP请求分组**，同一局域网中所有主机都能收到该请求。目的主机收到请求后就会向源主机**单播一个ARP响应分组**，源主机收到后将此映射写入**ARP缓存**（10-20min更新一次）。

ARP协议4种典型情况：

1. 主机A发给**本网络**上的主机B：用ARP找到主机B的硬件地址；
2. 主机A发给**另一网络**上的主机B：用ARP找到本网络上一个路由器（网关）的硬件地址；
3. 路由器发给**本网络**的主机A：用ARP找到主机A的硬件地址；
4. 路由器发给**另一网络**的主机B：用ARP找到本网络上的一个路由器的硬件地址。

ARP协议自动进行

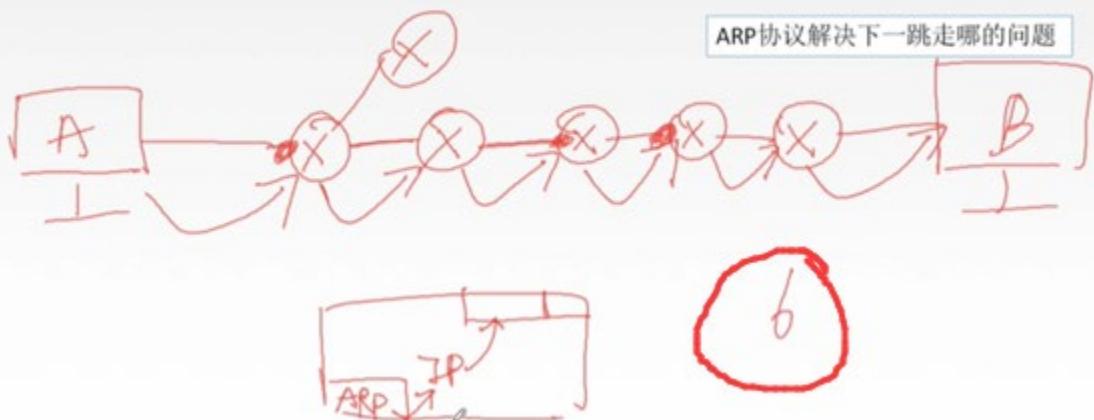
用户无感知

单播

ARP协议习题

考虑ARP高速缓存中初始均没有
映射关系

主机发送IP数据报给主机B，经过了5个路由器，请问此过程总共使用了几次ARP协议？



ARP协议介于网络层和数据链路层之间，主要划分为**网络层协议**，为IP协议提供服务。

4.2.8 DHCP 协议(动态主机配置协议)——是应用层的协议，但为网络层通信提供了基础



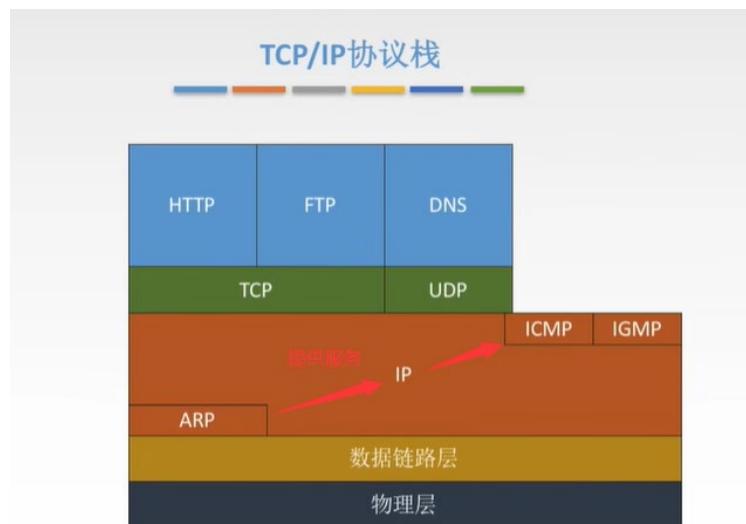
假如每个教学楼是一个网段，不同的人来使用教学楼中的网络时，DHCP 服务器会给他们自动分配和回收 IP 地址。



获得 IP 地址的整个过程都使用的是广播的方式。

DHCP 通常被用于局域网环境，主要作用是集中的管理、分配 IP 地址，使 client 动态的获得 IP 地址、Gateway 地址、DNS 服务器地址等信息，并能够提升地址的使用率。

4.2.9 ICMP 协议(Internet Control Message Protocol, 网际控制报文协议)——网络层协议



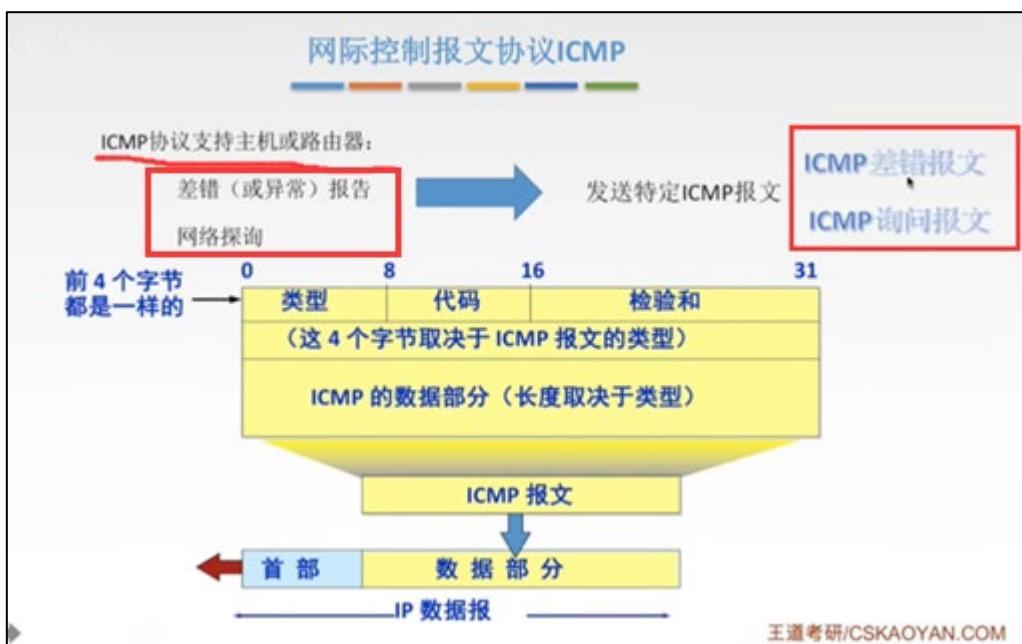
为了更有效地转发 IP 数据报和提高交付成功的机会.....

为了更有效地转发 IP 数据报和提高交付成功的机会.....

用于在 IP 主机、路由器之间传递控制消息。控制消息是指网络通不通、主机是否可达、路由是否可用等网络本身的消息。这些控制消息虽然并不传输用户数据，但是对于用户数据的传递起着重要的作用。

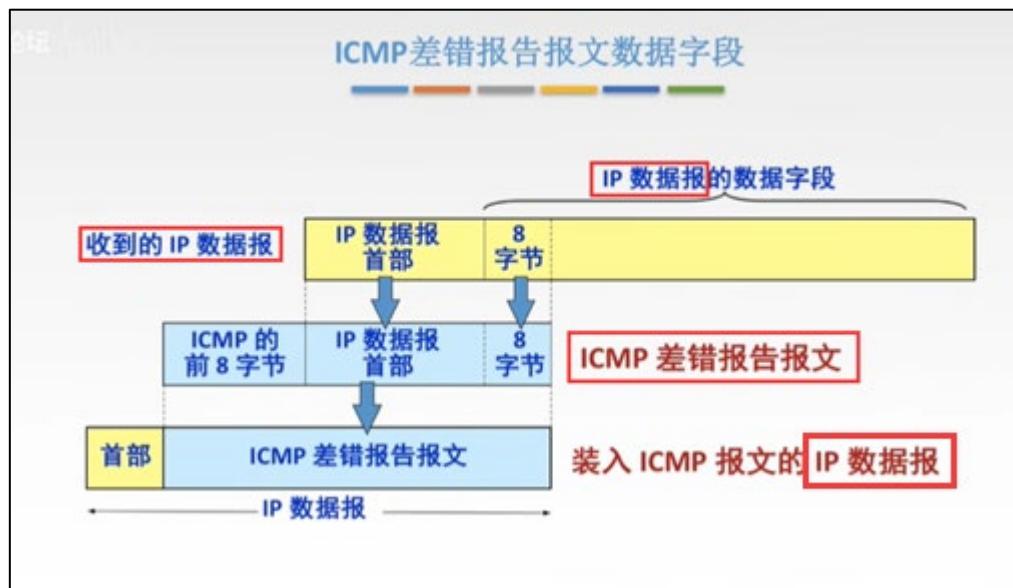
ICMP 协议是一种面向无连接的协议，用于传输出错报告控制信息。它是一个非常重要的协议，它对于网络安全具有极其重要的意义。它属于**网络层协议**，主要用于在主机与路由器之间传递控制信息，包括报告错误、交换受限控制和状态信息等。当遇到 IP 数据无法访问目标、IP 路由器无法按当前的传输速率转发数据包等情况时，会自动发送 ICMP 消息。

通信过程中，对于出错的 IP 数据报，网络层的处理就是直接丢弃，但是也要通过 ICMP 协议发送差错报告报文。



ICMP差错报告报文（5种）

- 1.终点不可达：当路由器或主机不能交付数据报时就向源点发送终点不可达报文。
无法交付
- ~~2.源点抑制~~当路由器或主机由于拥塞而丢弃数据报时，就向源点发送源点抑制报文，使源点知道应当把数据报的发送速率放慢。拥塞丢数据 目前基本不使用
- 3.时间超过：当路由器收到生存时间TTL=0的数据报时，除丢弃该数据报外，还要向源点发送时间超过报文。²当终点在预先规定的时间内不能收到一个数据报的全部数据报片时，就把已收到的数据报片都丢弃，并向源点发送时间超过报文。 TTL=0
- 4.参数问题：当路由器或目的主机收到的数据报的首部中有的字段的值不正确时，就丢弃该数据报，并向源点发送参数问题报文。 首部字段有问题
- 5.改变路由（重定向）：路由器把改变路由报文发送给主机，让主机知道下次应将数据报发送给另外的路由器（可通过更好的路由）。 值得更好的路由



不应发送ICMP差错报文的情况

- 1.对ICMP差错报告报文不再发送ICMP差错报告报文。
- 2.对第一个分片的数据报片的所有后续数据报片都不发送ICMP差错报告报文。
- 3.对具有组播地址的数据报都不发送ICMP差错报告报文。
一点到多点 广播：一点到所有点
- 4.对具有特殊地址（如127.0.0.0或0.0.0.0）的数据报不发送ICMP差错报告报文。

组播

ICMP询问报文

- 1.回送请求和回答报文 主机或路由器向特定目的主机发出的询问，收到此报文的主机必须给源主机或路由器发送ICMP回送回答报文。测试目的站是否可达以及了解其相关状态。

应用: ping

- 2.时间戳请求和回答报文 请某个主机或路由器回答当前的日期和时间。用来进行时钟同步和测量时间。

3.掩码地址请求和回答报文 子网掩码

这两种目前已经不再使用

4.路由器询问和通告报文

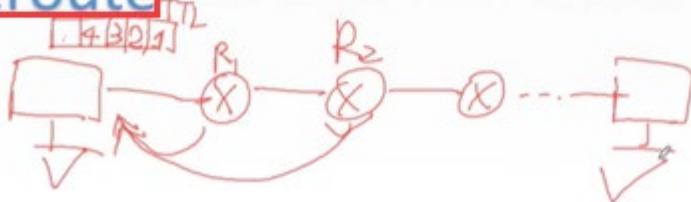
ICMP的应用

PING

测试两个主机之间的连通性，使用了ICMP回送请求和回答报文。

Traceroute

跟踪一个分组从源点到终点的路径，使用了ICMP时间超过差错报告报文。



源主机发送一连串数据报，但是这一连串数据报的 TTL(生存时间)不一样，利用每个数据报超过生存时间返回的 ICMP 差错报告报文，就可以测得原点到终点的路径。

4.3 IPv6

解决“IP地址耗尽”问题的举措有三种：

- 1、采用无类别编址CIDR，使IP地址的分配更加合理
- 2、采用网络[地址转换](#)(NAT)方法以节省全球IP地址
- 3、采用具有更大[地址空间](#)的新版本的IPv6

其中前两个方法只是延长了IPv4地址分配完毕的时间，只有第三种方法从根本上解决了IP地址的耗尽问题。

为了解决Internet地址资源枯竭的问题，技术人员先后提出了哪些解决方案（最好能有每种解决方案的详细说明

 我来答

 分享

 举报

2个回答

#热议# 二次感染新冠后会发生什么?



y1070933080

2011-01-18 · TA获得超过313个赞

关注

1.划分子网

可以把基于类的IP网络进一步分成更小的网络，每个子网由路由器界定并分配一个新的子网网络地址，子网地址是借用基于类的网络地址的主机部分创建的。划分子网后，通过使用掩码，把子网隐藏起来，使得从外部看网络没有变化，这就是子网掩码。

2.无类域间路由 (CIDR)

CIDR是一个在Internet上创建附加地址的方法，这些地址提供给服务提供商 (ISP)，再由ISP分配给客户。CIDR将路由集中起来，使一个IP地址代表主要骨干提供商服务的几千个IP地址，从而减轻Internet路由器的负担。

3.NAT (网络地址转换)

是一种将私有(保留)地址转化为合法IP地址的转换技术，它被广泛应用于各种类型Internet接入方式和各种类型的网络中。NAT不仅完美地解决了IP地址不足的问题，而且还能够有效地避免来自网络外部的攻击，隐藏并保护网络内部的计算机。

4.I Pv6

IPv6所拥有的地址容量是IPv4的约 8×10^{28} 倍。128位

什么是CIDR（无分类域间路由选择）？

无分类域间路由选择是在变长子网掩码的基础上提出的一种消除传统A、B、C类网络划分。

其特点主要有：

1、消除了传统A、B、C类地址及划分子网的概念，因而可以更有效地分配IPv4的地址空间。CIDR使用“网络前缀”的概念代替子网络的概念。因此，IP地址的无分类两级编址为：IP：：=<网络前缀>，<主机号>。

2、将网络前缀都相同的连续IP地址组成“CIDR地址块”。一个CIDR地址块可以表示很多地址，这种地址的聚合称为路由聚合，或称构成超网。路由聚合使得路由表中的一个项目可以表示多个原来传统分类地址的路由，有利于减少路由器之间的路由选择信息的交换，从而提高网络性能。

什么是NAT？

NAT即网络地址转换，是指通过将专用网络地址转换为公用地址，从而对外隐藏内部管理的IP地址。它使得整个专用网只需要一个全局IP地址就可以与因特网连通，由于专用网本地IP地址可以重用，因此NAT大大节省了IP地址的消耗。

使用NAT的时候需要在专用网连接到因特网的路由器上安装NAT软件，NAT路由至少有一个有效的外部全球地址。

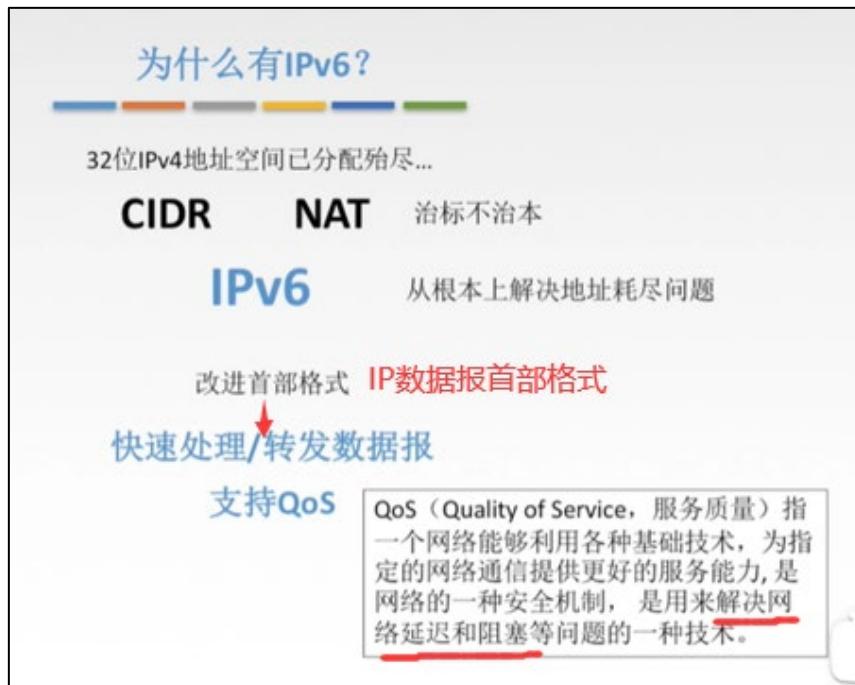
普通的路由器在转发IP数据报的时候，不改变其源IP地址和目的IP地址。而NAT路由器在转发IP数据报的时候，一定要更换其IP地址。普通路由器仅工作在网络层，而NAT路由器转发数据时需要查看和转换传输层的端口号。

什么是IPv6？

IPv6的主要特点：

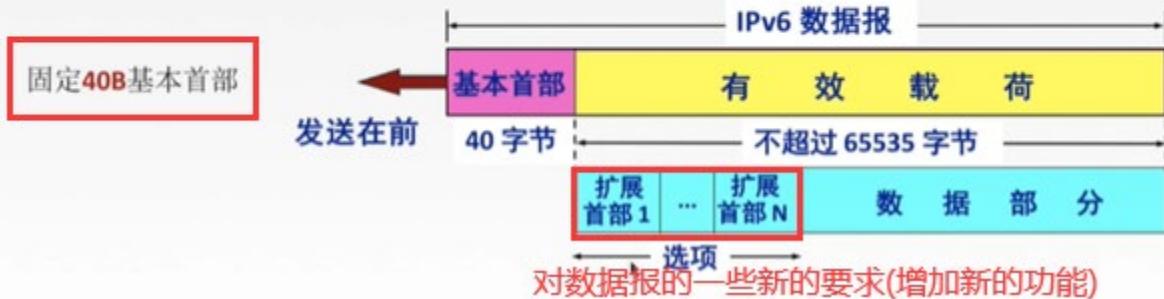
- 1、更大的地址空间。IPv4有32位，而IPv6有128位。
- 2、灵活的首部格式。IPv6将IPv4的校验和字段彻底移除。
- 3、允许协议继续扩充。
- 4、支持即插即用，即自动配置。
- 5、IPv6只有在包的源结点才能分片，是端到端的，传输路径中的路由器不能分片。IPv6只能在主机处分片，而IPv4可以在路由器和主机处分片。
- 6、增大了安全性。

<https://blog.csdn.net/xxxli/article/details/123649257>



改进首部格式：将首部中的可变部分挪到了有效载荷中。

IPv6数据报格式



IPv6和IPv4

1. IPv6将地址从32位 (4B) 扩大到 **128位 (16B)**，更大的地址空间。
2. IPv6将IPv4的校验和字段彻底移除，以减少每跳的处理时间。
3. IPv6将IPv4的可选字段移出首部，变成了**扩展首部**，成为灵活的首部格式，路由器通常不对扩展首部进行检查，大大提高了路由器的处理效率。
4. IPv6支持**即插即用**（即自动配置），不需要DHCP协议。
5. IPv6首部长度必须是**8B的整数倍**，IPv4首部是4B的整数倍。
6. IPv6只能在主机处分片，IPv4可以在路由器和主机处分片。
7. **ICMPv6**: 附加报文类型“分组过大”。如果传输的IPv6数据报过大，大于链路层最大传输单元MTU，路由器则将其丢弃，返回**ICMP差错报告报文**
8. IPv6支持资源的预分配，支持实时视像等要求，保证一定的带宽和时延的应用。
9. IPv6取消了协议字段，改成下一个首部字段。
10. IPv6取消了总长度字段，改用有效载荷长度字段。
11. IPv6取消了服务类型字段。

IPv6地址表示形式

128位

一般形式 冒号十六进制记法: 4BF5:AA12:0216:FEBC:BA5F:039A:BE9A:2170

压缩形式

4BF5:0000:0000:0000:BA 5F:039A:000A:2176
 4BF5:0:0:0:BA5F:39A:A:2176。

零压缩: 一连串连续的0可以被一对冒号取代。

FF05:0:0:0:0:0:B3
 FF05::B3

双冒号表示法在一个地址中仅可出现一次。

IPv6基本地址类型

单播 一对一通信 可做源地址+目的地址

多播 一对多通信 可做目的地址

任播 一对多中的一个通信 可做目的地址

IPv6 基本地址类型主要分为 3 种:

单播地址

多播地址

任播地址:一般是发送主机和任播组中, 离发送主机最近的一个进行通信。

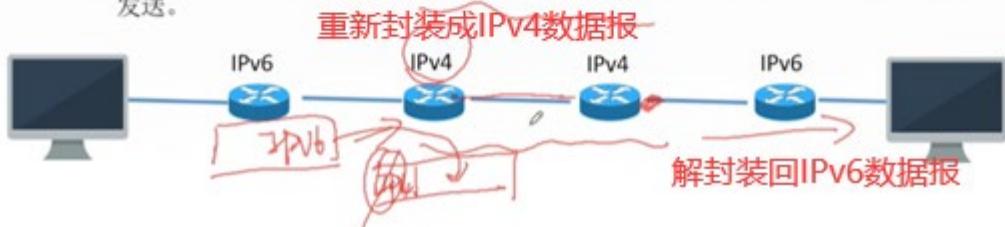
IPv6向IPv4过渡的策略

双栈协议

双协议栈技术就是指在一台设备上同时启用IPv4协议栈和IPv6协议栈。这样的话，这台设备既能和IPv4网络通信，又能和IPv6网络通信。如果这台设备是一个路由器，那么这台路由器的不同接口上，分别配置了IPv4地址和IPv6地址，并很可能分别连接了IPv4网络和IPv6网络。如果这台设备是一个计算机，那么它将同时拥有IPv4地址和IPv6地址，并具备同时处理这两个协议地址的功能。

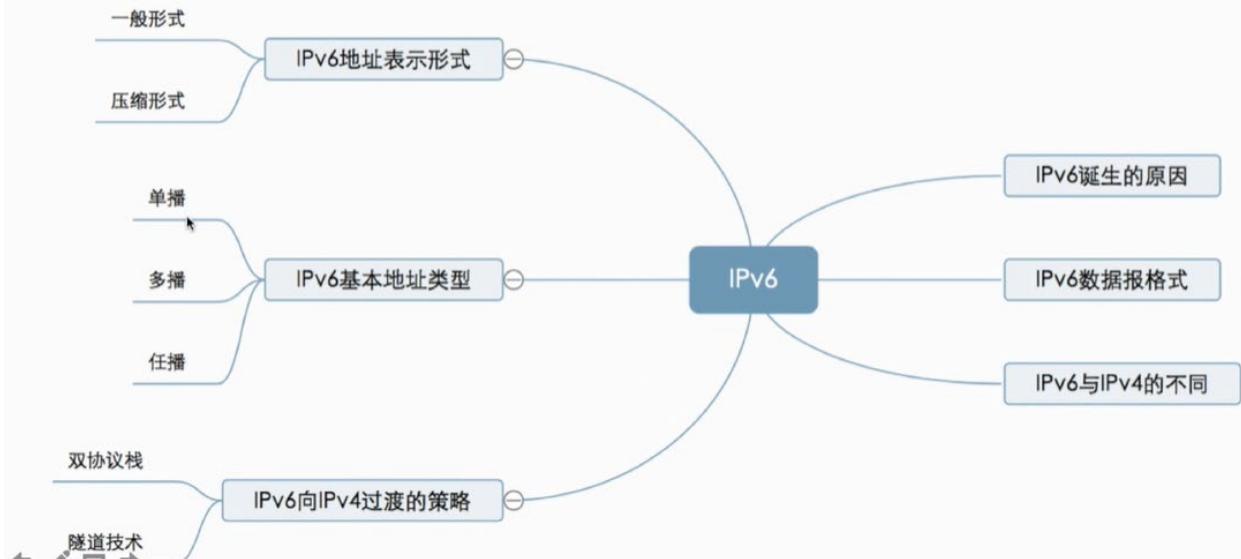
隧道技术

通过使用互联网基础设施在网络之间传递数据的方式。使用隧道传递的数据（或负载）可以是不同协议的数据帧或包。隧道协议将其它协议的数据帧或包重新封装然后通过隧道发送。



隧道技术

脑图时刻

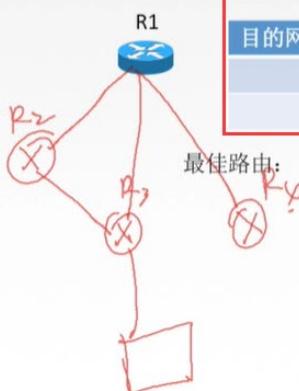


4.4 路由算法与路由协议概述

1. 路由算法

路由算法

标准的路由转发表



R1的路由表/转发表			
目的网络IP地址	子网掩码	下一跳IP地址	接口

最佳路由：“最佳”只能是相对于某一种特定要求下得出的较为合理的选择而已。

路由算法的分类

路由算法

静态路由算法（非自适应路由算法）管理员手工配置路由信息。

简便、可靠，在负荷稳定、拓扑变化不大的网络中运行效果很好，广泛用于高度安全性的军事网络和较小的商业网络。

路由更新慢，不适用大型网络。



动态路由算法（自适应路由算法）路由器间彼此交换信息，按照路由算法优化出路由表项。

路由更新快，适用大型网络，及时响应链路费用或网络拓扑变化。

算法复杂，增加网络负担。

对应协议：

动态路由算法

全局性 链路状态路由算法 OSPF

所有路由器掌握完整的网络拓扑和链路费用信息。

分散性 距离向量路由算法 RIP

路由器只掌握物理相连的邻居及链路费用。

分层次的路由选择协议

- (1) 因特网规模很大
- (2) 许多单位不想让外界知道自己的路由选择协议，但还想连入因特网



自治系统AS: 在单一的技术管理下的一组路由器，而这些路由器使用一种AS内部的路由选择协议和共同的度量以确定分组在该AS内的路由，同时还使用一种AS之间的路由协议以确定在AS之间的路由。

一个AS内的所有网络都属于一个行政单位来管辖，一个自治系统的所有路由器在本自治系统内都必须连通。



对外界透明：即外界看不到，不清楚。



4.5.1 OSPF 协议与链路状态算法——内部网关协议(全局性动态路由算法)——网络层协议

OSPF协议

开放最短路径优先OSPF协议：“开放”标明OSPF协议不是受某一家厂商控制，而是公开发表的；“最短路径优先”是因为使用了Dijkstra提出的最短路径算法SPF。

OSPF最主要的特征就是使用分布式的链路状态协议。

OSPF的特点：

- | | |
|-------|---|
| 和谁交换？ | 1. 使用 <u>洪泛法</u> 向自治系统内所有路由器发送信息，即路由器通过输出端口向所有相邻的路由器发送信息，而每一个相邻路由器又再次将此信息发往其所有的相邻路由器。 <u>广播</u> |
| | → 最终整个区域内所有路由器都得到了这个信息的一个副本。 |
| 交换什么？ | 2. 发送的信息就是与本路由器相邻的 <u>所有路由器的链路状态</u> （本路由器和哪些路由器相邻，以及该链路的度量/代价——费用、距离、时延、带宽等）。 |
| 多久交换？ | 3. 只有当 <u>链路状态发生变化时</u> ，路由器才向所有路由器洪泛发送此信息。 |

最后，所有路由器都能建立一个链路状态数据库，即全网拓扑图。

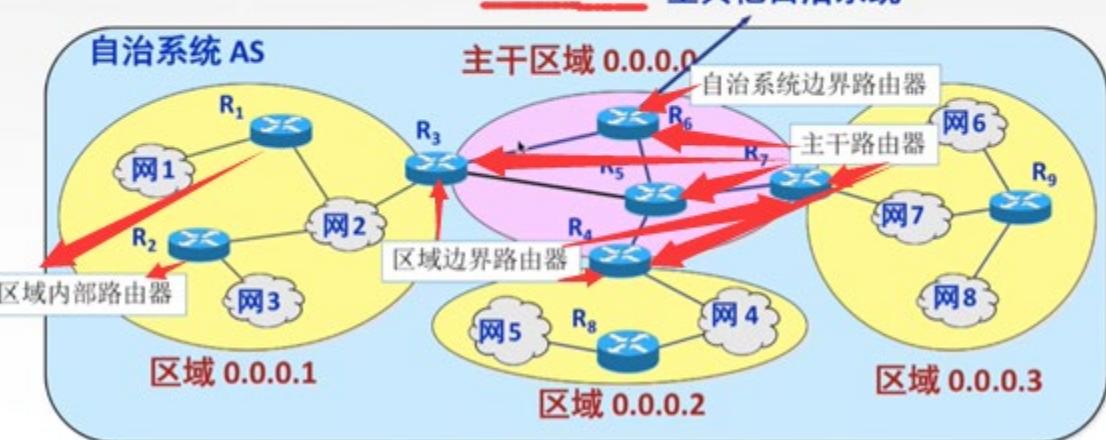
链路状态路由算法

1. 每个路由器发现它的邻居结点【**HELLO问候分组**】，并了解邻居节点的网络地址。
2. 设置到它的每个邻居的成本度量**metric**。
即全网的拓扑图
3. 构造【**DD数据库描述分组**】，向邻站给出自己的**链路状态数据库**中的所有链路状态项目的摘要信息。
4. 如果**DD分组**中的摘要自己都有，则邻站不做处理；如果没有有的或者是更新的，则发送【**LSR链路状态请求分组**】请求自己没有的和比自己更新的信息。
5. 收到邻站的**LSR分组**后，发送【**LSU链路状态更新分组**】进行更新。
6. 更新完毕后，邻站返回一个【**LSAck链路状态确认分组**】进行确认。
- 只要一个路由器的链路状态发生变化：
5. 泛洪发送【**LSU链路状态更新分组**】进行更新。
6. 更新完毕后，其他站返回一个【**LSAck链路状态确认分组**】进行确认。
7. 使用**Dijkstra**根据自己的**链路状态数据库**构造到其他节点间的最短路径。

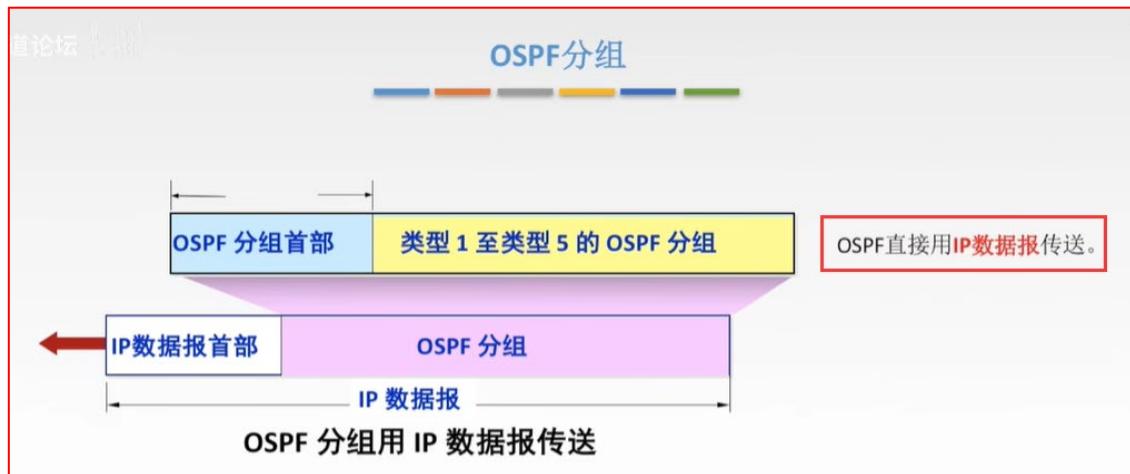
OSPF的区域

为了使 OSPF 能够用于规模很大的网络，OSPF 将一个**自治系统**再划分为若干个更小的范围，叫做**区域**。每一个区域都有一个**32 位的区域标识符**（用**点分十进制表示**）。

区域也不能太大，在一个区域内的路由器最好不超过**200 个**。
至其他自治系统



R3 R4 R7 既是主干路由器也是区域边界路由器



一般认为 OSPF 协议还是网络层的协议。



3.一般, RIP 收到相邻路由器传来的路由表后, 先要和自己的路由表进行**对照**, 才能确定一个到某网络的最短路径及下一跳路由器; 而 OSPF 会将收到的更新**直接**放到数据库中, 然后再使用 Dijkstra 算法**直接**计算最短路径, 收敛速度更快。

4.5.2 RIP 协议与距离向量算法——内部网关协议(分散性动态路由算法)——应用层协议



RIP协议和谁交换？多久交换一次？交换什么？



1. 仅和相邻路由器交换信息。

2. 路由器交换的信息是自己的路由表。



“我到Net1网络的（最短）距离是5跳，下一跳应该走R1路由器.....”

3. 每30秒交换一次路由信息，然后路由器根据新信息更新路由表。若超过180s没收到邻居路由器的通告，则判定邻居没了，并更新自己路由表。

路由器刚开始工作时，只知道直接连接的网络的距离（距离为1），接着每一个路由器也只和数目非常有限的相邻路由器交换并更新路由信息。

经过若干次更新后，所有路由器最终都会知道到达本自治系统任何一个网络的最短距离和下一跳路由器的地址，即“收敛”。



路由表具体是怎么进行更新的呢？→**距离向量算法**。

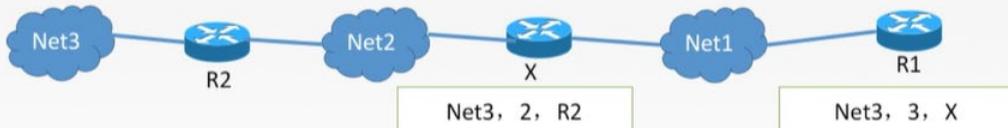
相邻路由器间交换的**RIP报文**包含路由表的全部信息。

距离向量算法



1. 修改相邻路由器发来的RIP报文中所有表项

对地址为X的相邻路由器发来的RIP报文，修改此报文中的所有项目：把“下一跳”字段中的地址改为X，并把所有的“距离”字段+1。



2. 对修改后的RIP报文中的每一个项目，进行以下步骤：

(1) R1路由表中若没有Net3，则把该项目填入R1路由表

(2) R1路由表中若有Net3，则查看下一跳路由器地址：

若下一跳是X，则用收到的项目替换源路由表中的项目；**使用最新的消息**

若下一跳不是X，原来距离比从X走的距离远则更新，否则不作处理。

3. 若180s还没收到相邻路由器X的更新路由表，则把X记为不可达的路由器，即把距离设置为16。

4. 返回

距离向量算法练习1

已知路由器R6的路由表，现收到相邻路由器R4发来的路由更新信息，试更新路由器R1的路由表：

R6的路由表		
目的网络	距离	下一跳路由器
Net2	3	R4
Net3	4	R5
...

解：

R4发来的路由更新信息-修改版		
目的网络	距离	下一跳路由器
Net1	4	R4
Net2	5	R4
Net3	2	R4

R4发来的路由更新信息		
目的网络	距离	下一跳路由器
Net1	3	R1
Net2	4	R2
Net3	1	直接交付

R6更新后的路由表		
目的网络	距离	下一跳路由器
Net1	4	R4
Net2	5	R4
Net3	2	R4
...

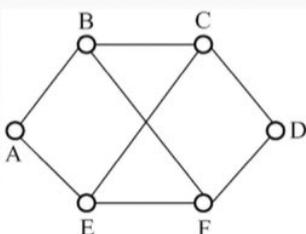


距离向量算法练习2

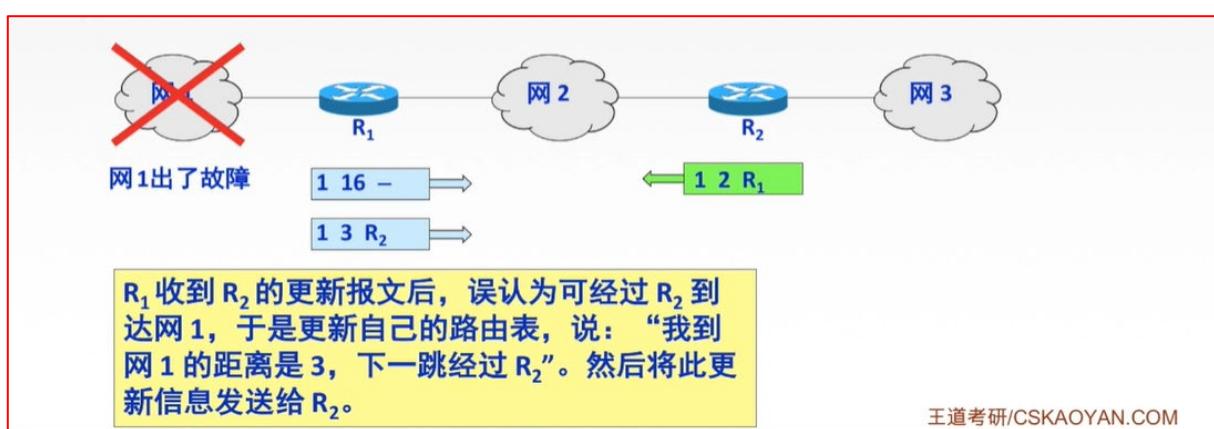
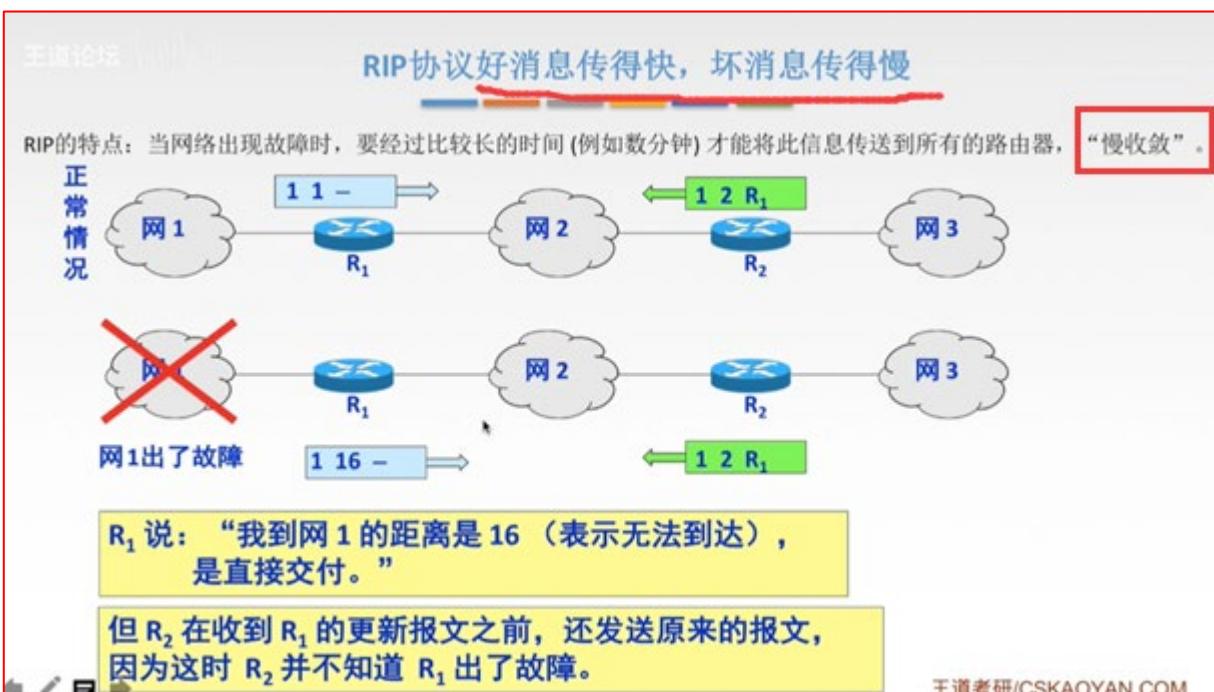
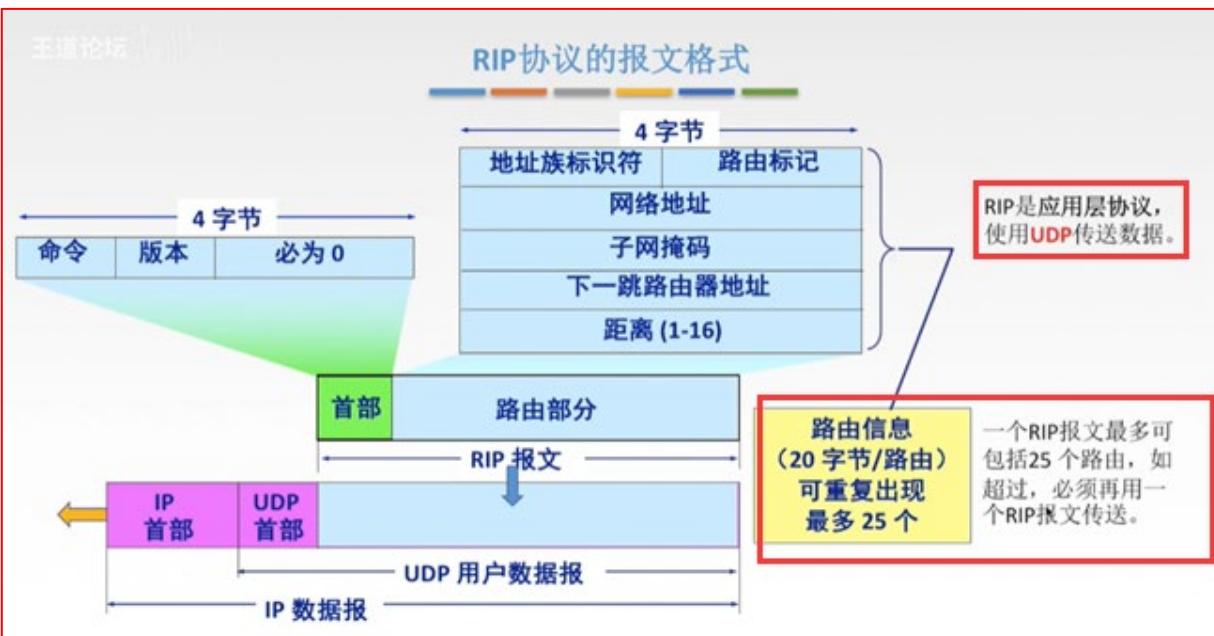
A B

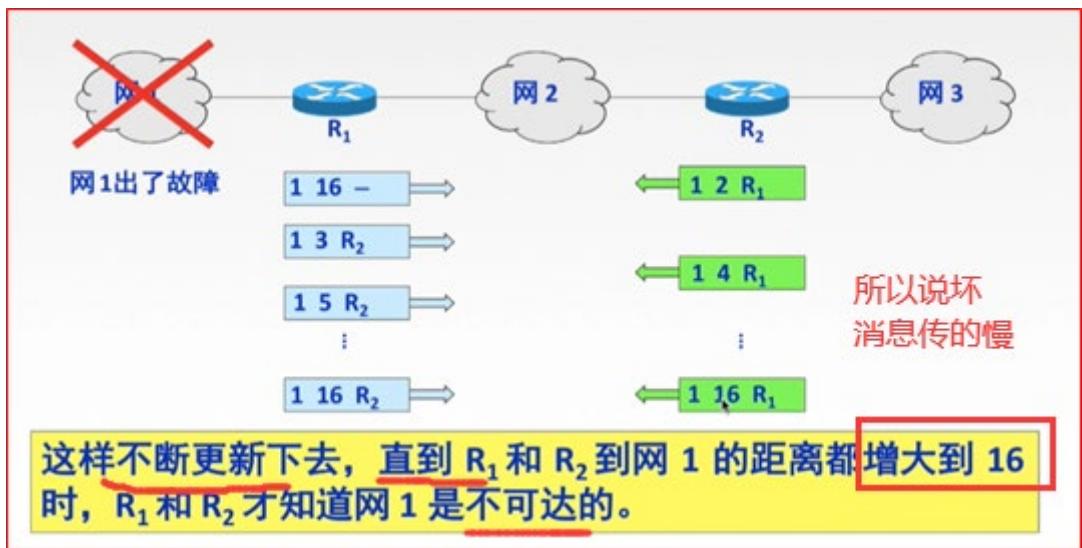
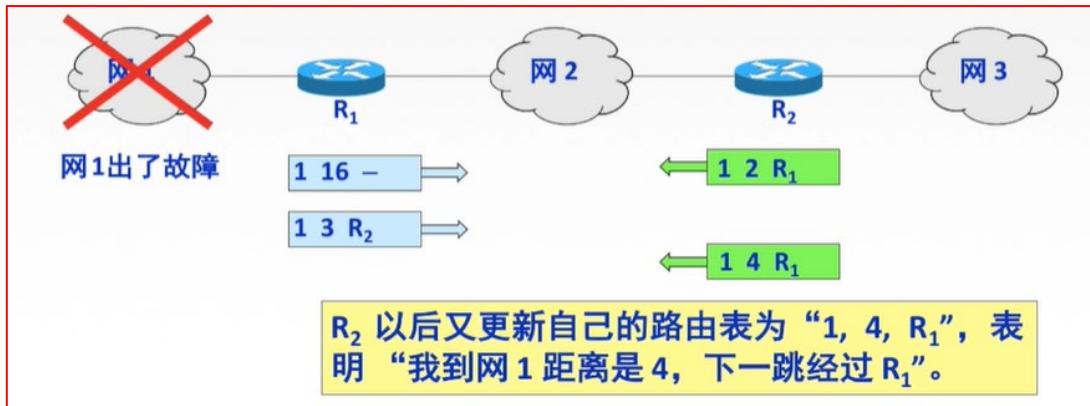
考虑如图所示的子网，该子网使用了距离-向量算法，下面的向量刚刚到达路由器C：来自B的向量为(5, 0, 8, 12, 6, 2)；来自D的向量为(16, 12, 6, 0, 9, 10)；来自E的向量为(7, 6, 3, 9, 0, 4)。经过测量，C到B、D和E的延迟分别为6, 3和5，那么C到达所有结点的最短路径是（ ）。

A. (5, 6, 0, 9, 6, 2) B. (11, 6, 0, 3, 5, 8)
 C. (5, 11, 0, 12, 8, 9) D. (11, 8, 0, 7, 4, 9)

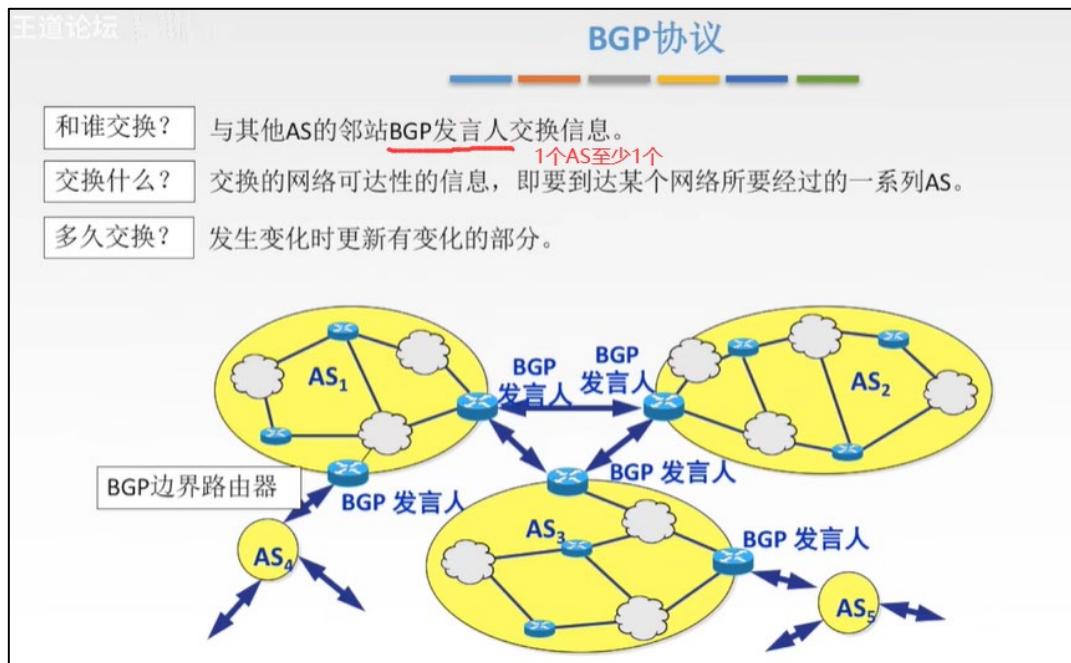


C到B: (11, 6, 14, 18, 12, 8)
 C到D: (19, 15, 9, 3, 12, 13)
 C到E: (12, 11, 8, 14, 5, 9)

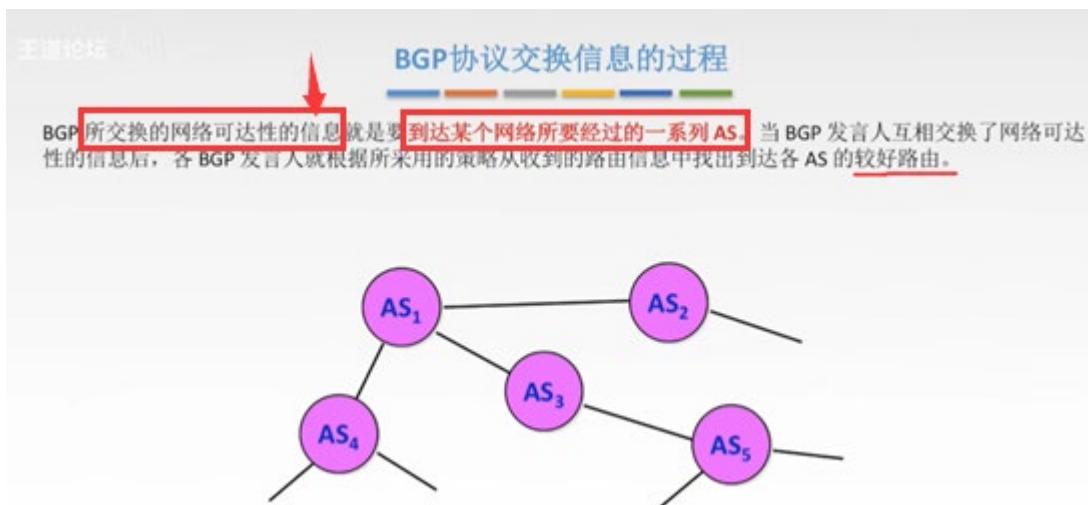




4.5.3 BGP 协议——外部网关协议——应用层协议



BGP 发言人既使用 BGP 协议，也使用 RIP 或 OSPF 协议。

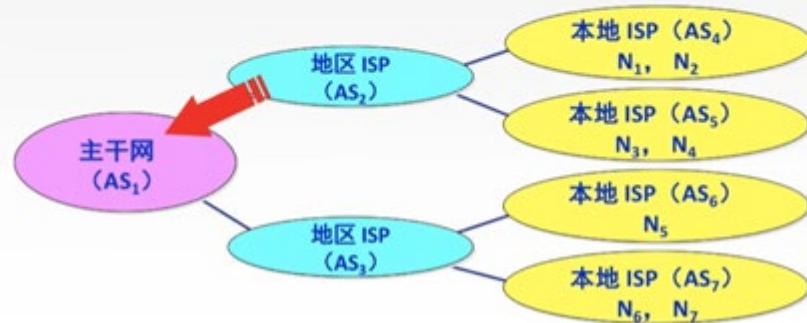


一般使用树形结构，避免绕圈子。

王道论坛 BGP 协议交换信息的过程

BGP发言人交换**路径向量**:

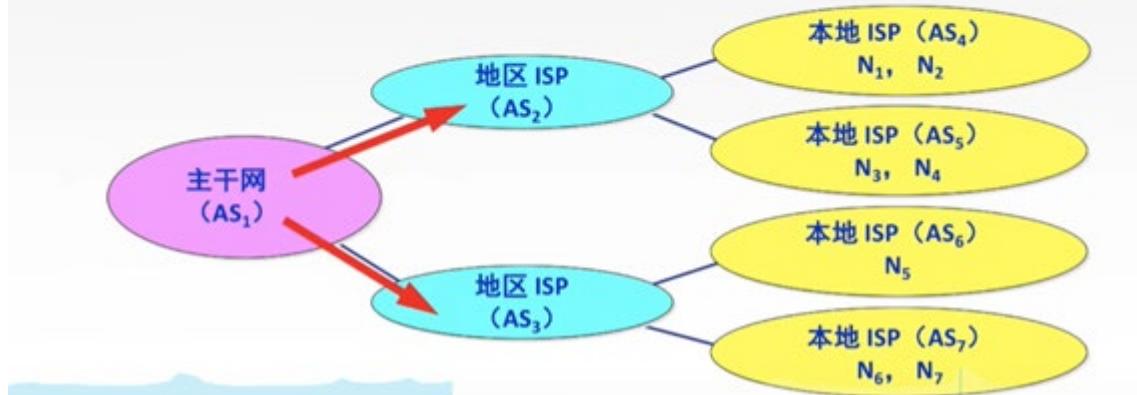
自治系统 AS₂ 的 BGP 发言人通知主干网 AS₁ 的 BGP 发言人: “要到达网络 N₁、N₂、N₃ 和 N₄ 可经过 AS₂。”



王道论坛 BGP 协议交换信息的过程

BGP发言人交换**路径向量**:

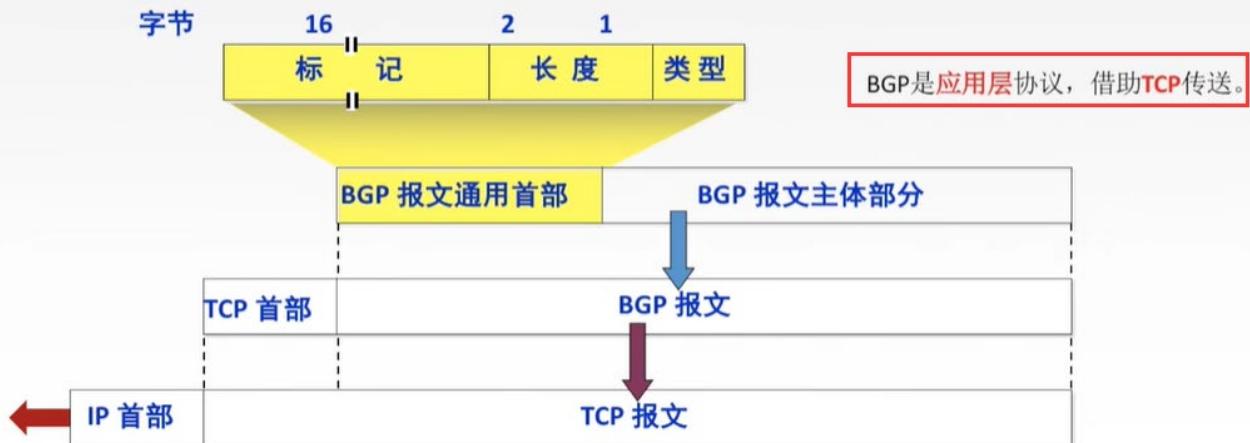
主干网还可发出通知: “要到达网络 N₅、N₆ 和 N₇ 可沿路径 (AS₁, AS₃)。”



BGP 协议交换的信息包含的是一个完整的路径

BGP协议报文格式

一个BGP发言人与其他自治系统中的BGP发言人要交换路由信息，就要先建立TCP连接，即通过TCP传送，然后在此连接上交换BGP报文以建立BGP会话(session)，利用BGP会话交换路由信息。



为什么使用TCP：使用TCP连接可以提供更加可靠的服务，可以简化路由选择协议。

BGP协议特点

BGP支持CIDR，因此BGP的路由表也就应当包括目的网络前缀、下一跳路由器，以及到达该目的网络所要经过的各个自治系统序列。

在BGP刚刚运行时，BGP的邻站是交换整个的BGP路由表。但以后只需要在发生变化时更新有变化的部分。这样做对节省网络带宽和减少路由器的处理开销都有好处。

BGP-4的四种报文

1. OPEN（打开）报文：用来与相邻的另一个BGP发言人建立关系，并认证发送方。

2. UPDATE（更新）报文：通告新路径或撤销原路径。BGP-4中最常使用的报文

3. KEEPALIVE（保活）报文：在无UPDATE时，周期性证实邻站的连通性；也作为OPEN的确认。

4. NOTIFICATION（通知）报文：报告先前报文的差错；也被用于关闭连接。

关闭TCP连接

三种路由协议比较

应用层协议

RIP是一种分布式的基于距离向量的内部网关路由选择协议，通过广播**UDP报文**来交换路由信息。

OSPF是一个内部网关协议，要交换的信息量较大，应使报文的长度尽量短，所以**不使用传输层协议**（如**UDP或TCP**），而是直接采用**IP**。**网络层协议**

BGP是一个外部网关协议，在不同的自治系统之间交换路由信息，由于**网络环境复杂，需要保证可靠传输**，所以采用**TCP**。**应用层协议**



三种路由协议比较

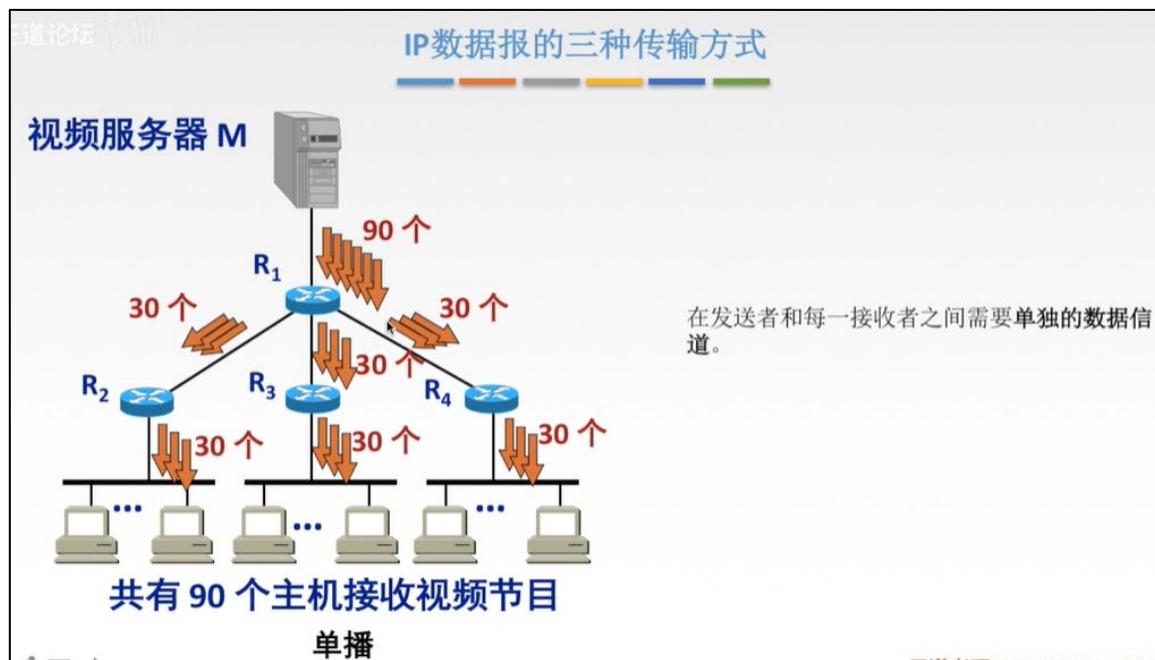
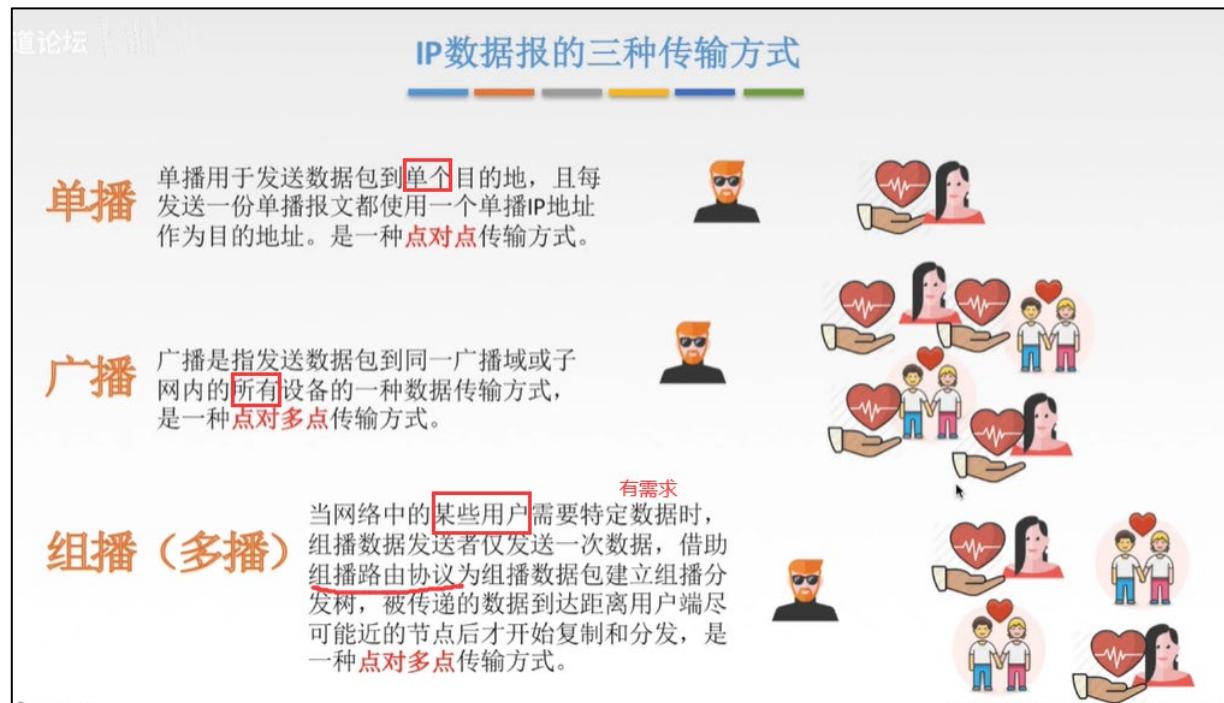
协议	RIP	OSPF	BGP
类型	内部	内部	外部
路由算法	跳数 距离-向量	链路状态	路径-向量
传递协议	UDP	IP	TCP
路径选择	跳数最少	代价最低	较好，非最佳
交换结点	和本结点相邻的路由器	网络中的所有路由器	和本结点相邻的路由器
交换内容	当前本路由器知道的全部信息，即自己的路由表	与本路由器相邻的所有路由器的链路状态	首次 整个路由表 非首次 有变化的部分

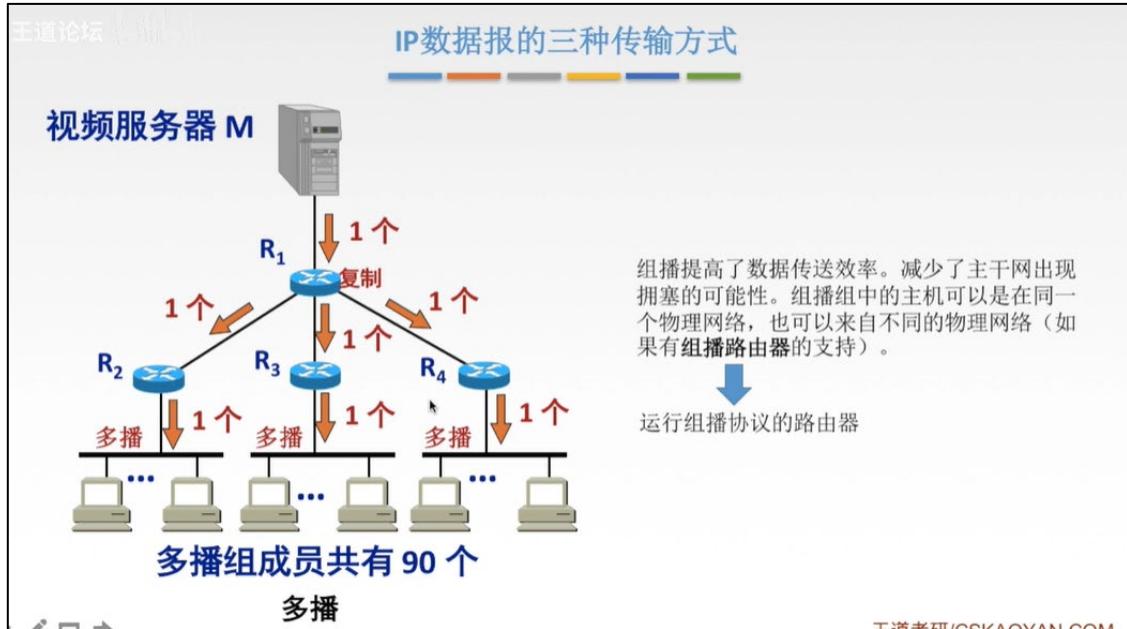
RIP：到某一网络的最短距离/跳数，下一跳地址

OSPF：交换内容就是与本路由器相邻的所有路由器的链路状态（本路由器和哪些路由器相邻，以及该链路的度量/代价——费用、距离、时延、带宽等）。

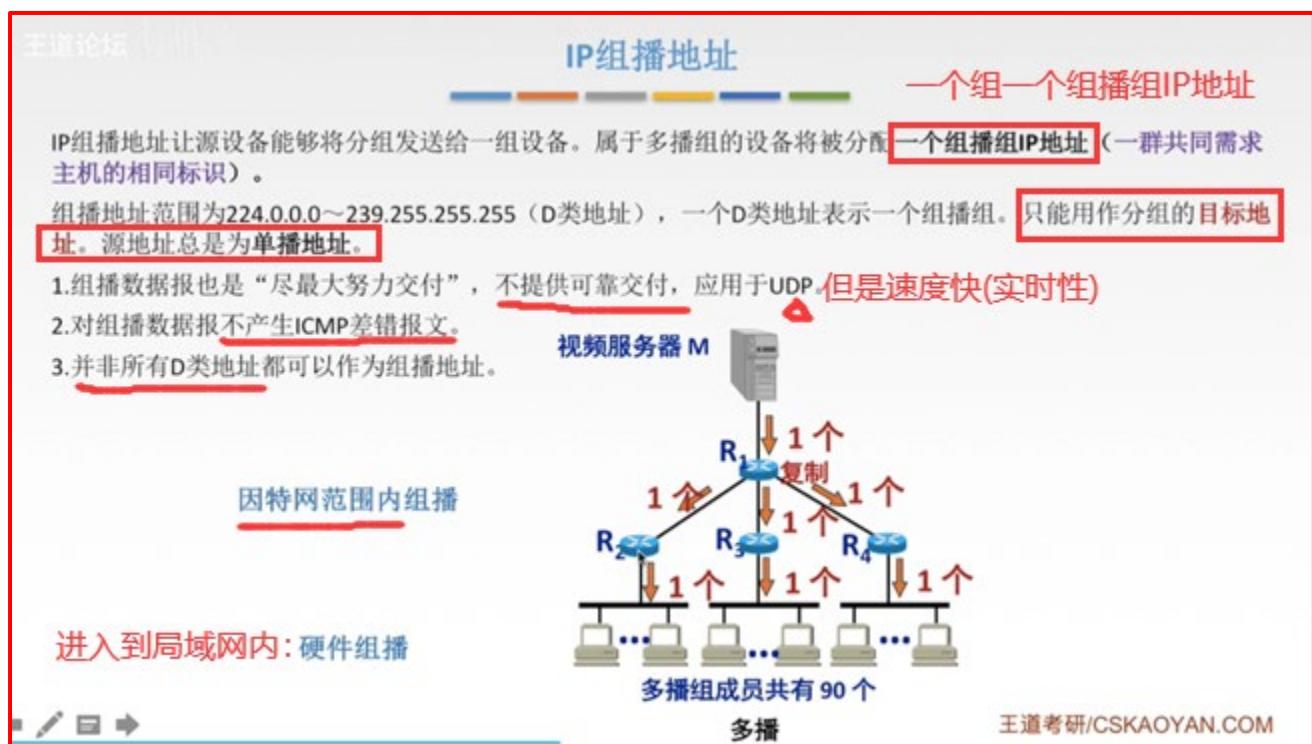
BGP：BGP的路由表也就应当包括目的网络前缀、下一跳路由器，以及到达该目的网络所要经过的各个自治系统序列。

4.6 IP 组播





发送时，怎么找到一个组播组中的主机？ \rightarrow 组播组 IP 地址。应用场景：看直播或线上会议等。

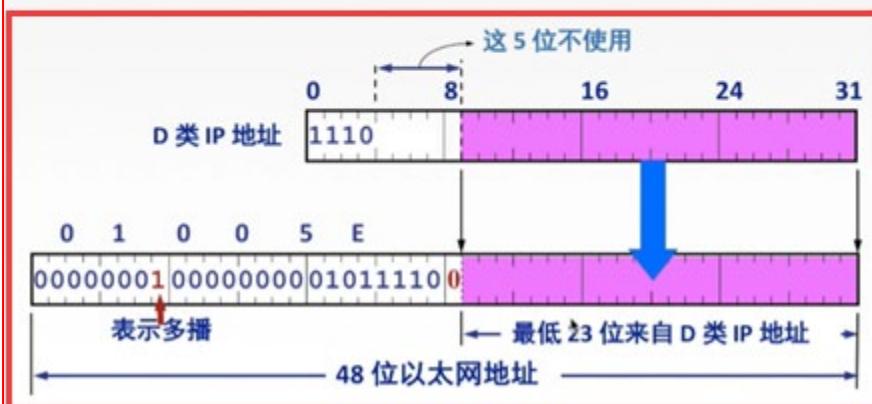


硬件组播

同单播地址一样，组播IP地址也需要相应的组播MAC地址在本地网络中实际传送帧。组播MAC地址以十六进制值01-00-5E打头，余下的6个十六进制位是根据IP组播组地址的最后23位转换得到的。

TCP/IP 协议使用的以太网多播地址的范围是：

从01-00-5E-00-00-00到01-00-5E-7F-FF-FF.



因为MAC地址相同时，组播组IP地址的中间5位还可能不同，所以可能不是一个组播组的

收到多播数据报的主机，还要在IP层利用软件进行过滤，把不是本主机要接收的数据报丢弃。

IGMP 协议与组播路由选择协议：

IGMP 协议(网络层协议): 在一个路由器内部使用，是为了让连在局域网上的组播路由器知道该局域网上是否还有主机参加或退出了某个组播组。即它的作用就是：对于一个组播路由器来说，收到一个组播数据报时，IGMP 告诉路由器是否将该数据报分发给所在局域网中的组播组成员。

组播路由选择协议: 以最小的代价将组播数据报传送给组播组中的全部成员。

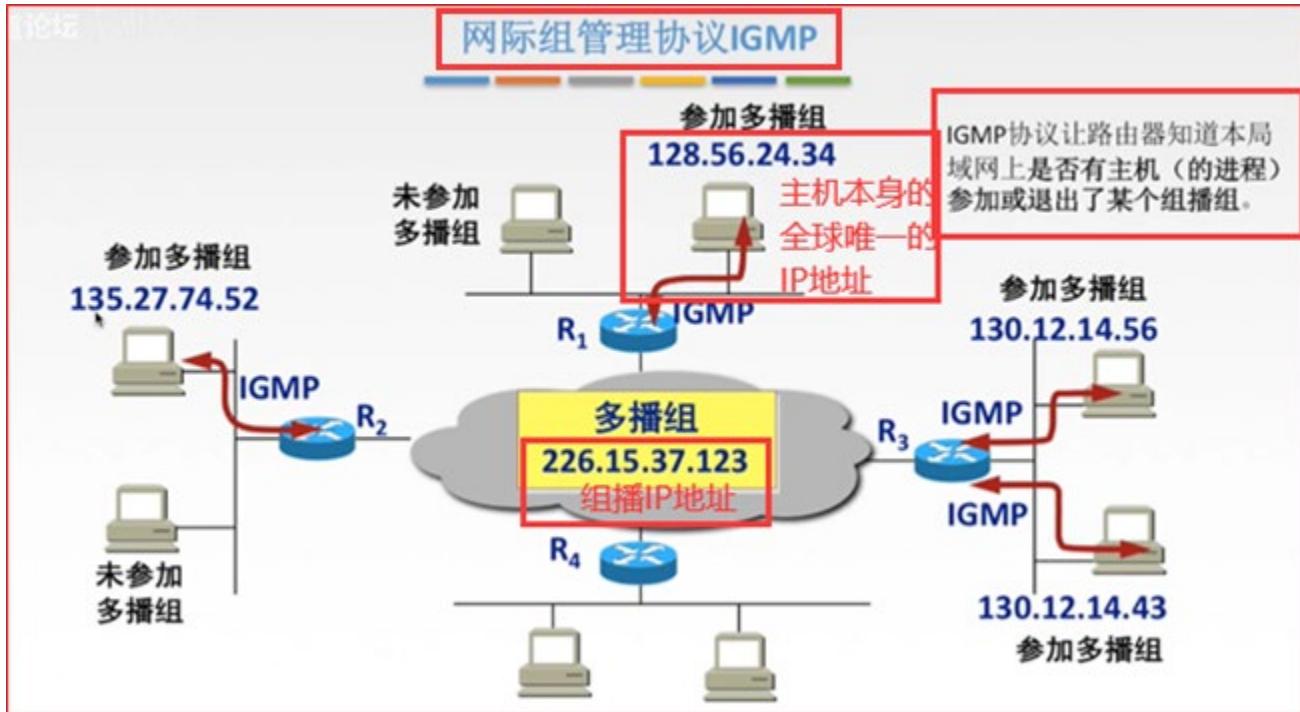
IGMP协议与组播路由选择协议

IGMP协议

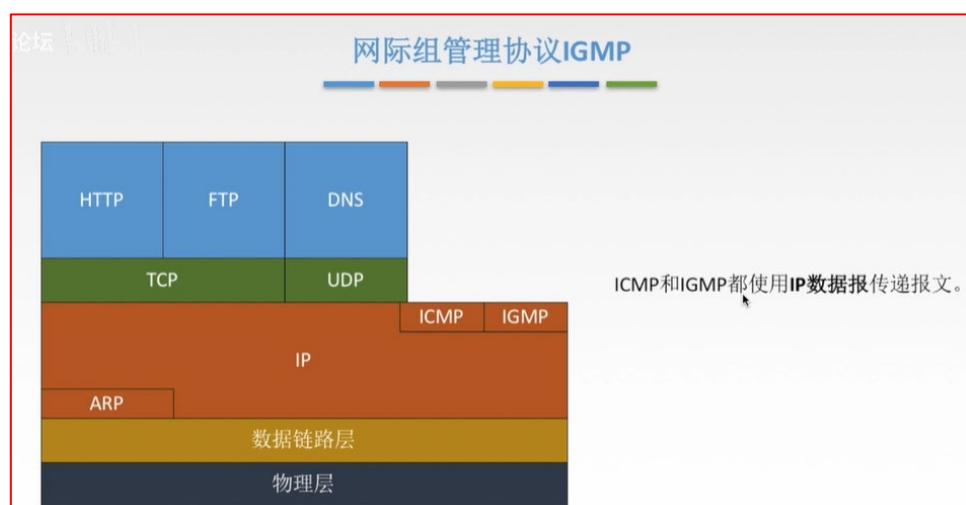


组播路由选择协议





可能一个主机有多个进程参加了组播组，那么每个进程都会有一个组播 IP 地址。IGMP 协议可以判断，如果路由器接收到了一个组播数据报，是否要给我局域网中的主机。但是 IGMP 不知道本局域网中组播组成员个数，也不知道组播组都在哪些网络上。他只是可以让路由器知道该局域网中是否组播组。<P181>



IGMP工作的两个阶段

ROUND 1:

某主机要加入组播组时，该主机向组播组的组播地址发送一个IGMP报文，声明自己要称为该组的成员。
本地组播路由器收到IGMP报文后，要利用组播路由选择协议把这组成员关系发给因特网上的其他组播路由器。

ROUND 2:

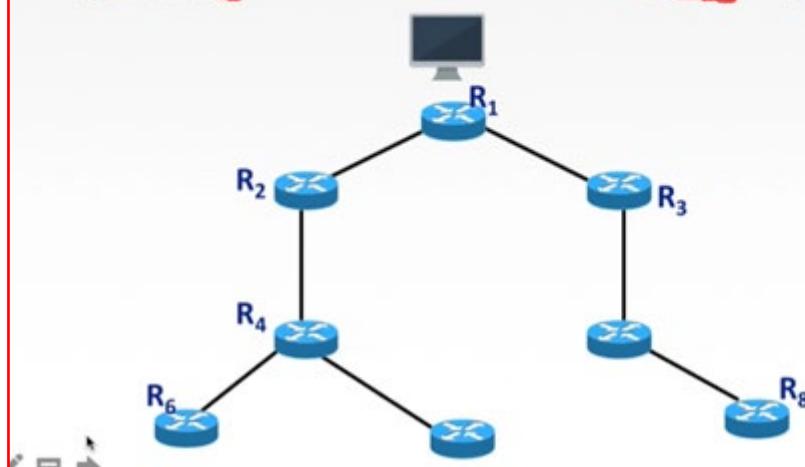
本地组播路由器周期性探询本地局域网上的主机，以便知道这些主机是否还是组播组的成员。
只要有一个主机对某个组响应，那么组播路由器就认为这个组是活跃的；如果经过几次探询后没有一个主机响应，组播路由器就认为本网络上的没有此组播组的主机，因此就不再把这组的成员关系发给其他的组播路由器。

组播路由器知道的成员关系只是所连接的局域网中有无组播组的成员。



组播路由选择协议

组播路由协议目的是找出以源主机为根节点的组播转发树
构造树可以避免在路由器之间兜圈子。
对不同的多播组对应于不同的多播转发树；同一个多播组，对不同的源点也会有不同的多播转发树。



首论坛

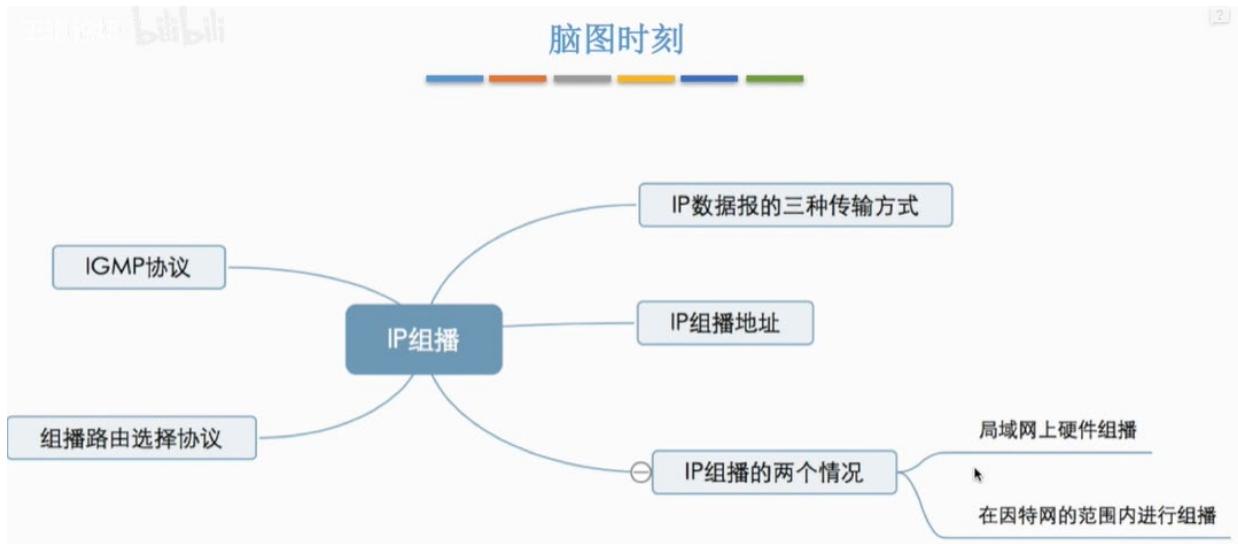
组播路由选择协议

组播路由选择协议常使用的三种算法：

- [基于链路状态的路由选择](#)
- [基于距离-向量的路由选择](#)
- [协议无关的组播（稀疏/密集）](#)

协议无关：可以建立在任何协议路由器之上，建立转发树时使用的是单播数据报和远程路由器通信，但是和具体是哪种单播路由选择协议无关。

稀疏/密集：组播组中的主机距离远近。



4.7 移动 IP

区分：动态 IP：使用 DHCP 协议在一个局域网内的一台主机，自动获取 IP 地址等信息。

移动 IP 的一个应用：在上海分公司连接网络后，获得和在北京总部相同的查看数据库等权限。这一过程，IP 地址是不变的。

移动IP相关术语

移动IP技术是移动结点(计算机/服务器等)以**固定的网络IP地址**, 实现跨越不同网段的**漫游功能**, 并保证了基于网络IP的网络权限在漫游过程中不发生任何改变。

移动结点 具有永久IP地址的移动设备。

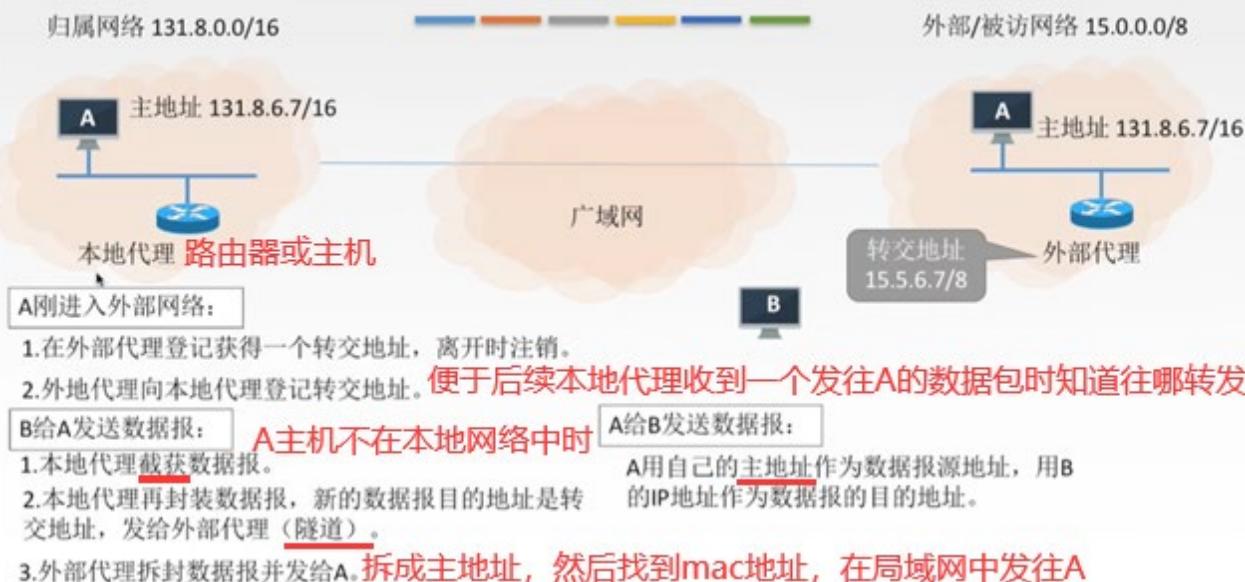
归属代理 (本地代理) 一个移动结点拥有的就“居所”称为归属网络, 在归属网络中代表移动节点执行移动管理功能的实体叫做**归属代理**

外部代理 (外地代理) 在**外部网络**中帮助移动节点完成移动管理功能的实体称为**外部代理**。

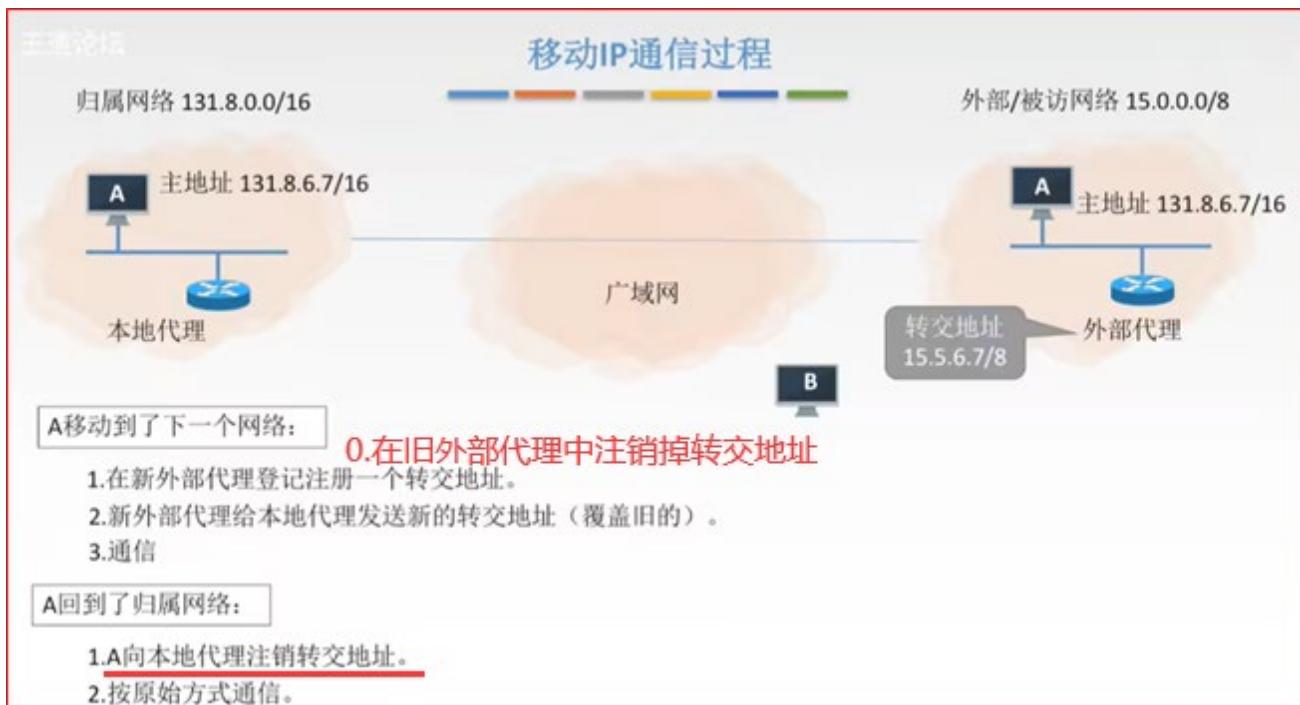
永久地址 (归属地址/主地址) 移动站点在**归属网络**中的原始地址。

转交地址 (辅地址) 移动站点在**外部网络**使用的临时地址。

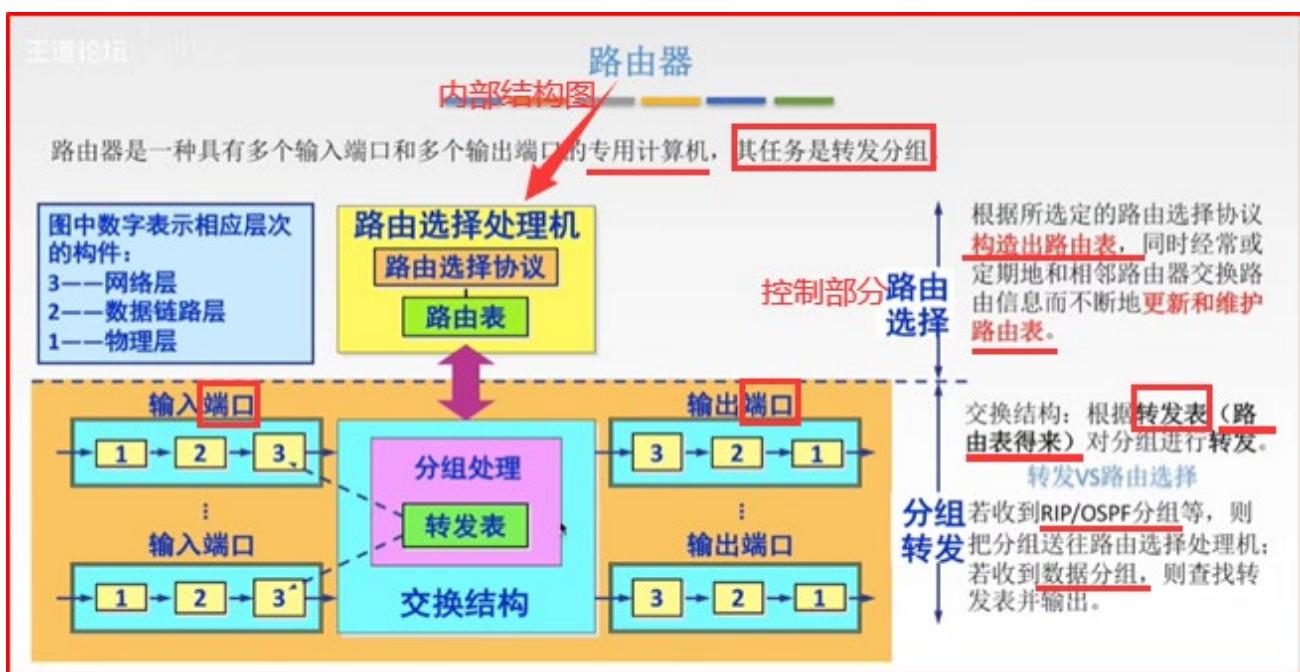
移动IP通信过程



隧道



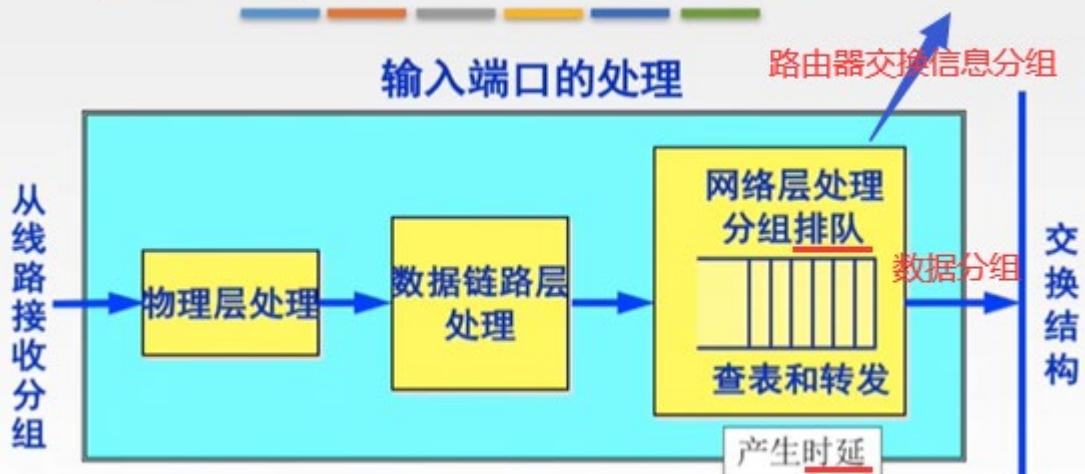
4.8 网络层设备



转发是在内部(输入/输出端口之间)的，路由选择是在外部(路由之间)的。

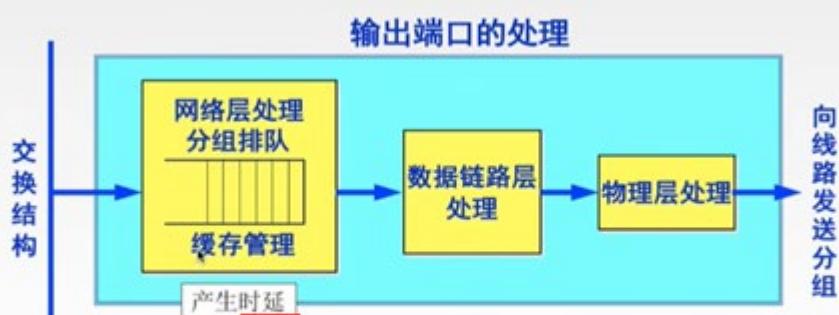
并不是所有分组从输入端口进来以后，都会从输出端口出去。路由器之间交换信息的分组(如RIP分组、OSPF分组等)，就是送往路由选择处理机。

输入端口对线路上收到的分组的处理



输入端口中的查找和转发功能在路由器的交换功能中是最重要的。

输出端口将交换结构传送来的分组发送到线路



若路由器处理分组的速率赶不上分组进入队列的速率，则队列的存储空间最终必定减少到零，这就使后面再进入队列的分组由于没有存储空间而只能被丢弃。

路由器中的输入或输出队列产生溢出是造成分组丢失的重要原因。

三层设备的区别

路由器 可以互联两个不同网络层协议的网段。

网桥 可以互联两个物理层和链路层不同的网段。

集线器 不能互联两个物理层不同的网段。(傻瓜式设备)

	能否隔离冲突域	能否隔离广播域
物理层设备【傻瓜】 (中继器、集线器)	×	×
链路层设备【路人】 (网桥、交换机)	√	×
网络层设备【大佬】 (路由器)	√	√

路由表与路由转发

路由表根据路由选择算法得出的，主要用途是路由选择，总用软件来实现。



转发表由路由表得来，可以用软件实现，也可以用特殊的硬件来实现。转发表必须包含完成转发功能所必需的信息，在转发表的每一行必须包含从要到达的目的网络到输出端口和某些MAC地址信息的映射。可以是下一跳的以太网地址

4.9 第四章总结





五、运输层/传输层

5.1 传输层概述

只有主机才会有的层次，中间的网络设备（如路由器等）最多只能到网络层。



到网络层后，通信其实还未完成，还需要到主机中某一个进程，甚至进程中的线程，才能实现通信的过程。



传输层的两个协议

传输层有两个好兄弟

大哥TCP和二弟UDP

大哥靠谱，二弟不靠谱

面向连接的传输控制协议TCP

传送数据之前必须建立连接，数据传送结束后要释放连接。不提供广播或多播服务。由于TCP要提供可靠的面向连接的传输服务，因此不可避免增加了许多开销：确认、流量控制、计时器及连接管理等。

可靠，面向连接，时延大，适用于大文件。如传输大文件

VS

无连接的用户数据报协议UDP

传送数据之前不需要建立连接，收到UDP报文后也不需要给出任何确认。

不可靠，无连接，时延小，适用于小文件。

如QQ消息，若失败了则有红色感叹号，可重新发送。

传输层的寻址与端口

复用：应用层所有的应用进程都可以通过传输层再传输到网络层。

分用：传输层从网络层收到数据后交付指明的应用进程。

服务访问点

区别于硬件端口 **逻辑端口/软件端口** **端口** 是传输层的**SAP**，标识主机中的应用进程。

端口号只有本地意义，在因特网中不同计算机的相同端口号是没有联系的。

端口号长度为**16bit**，能表示**65536**个不同的端口号。

熟知端口号：给**TCP/IP最重要的一些应用程序**，**让所有用户都知道**。

0~1023

登记端口号：为**没有熟知端口号的应用程序**使用的。

1024~49151

端口号
(按范围分)

服务端使用的端口号

服务器

用户/主机

客户端使用的端口号

主机的操作系统随机分配，进程结束端口号回收

49152~65535

王道考研/CSKAOYAN.

想要向数据发送给一个主机，需要知道它的IP地址。就可以在网络中根据**IP地址**寻找到它所在的网络(网络层)；进入到该主机所在的网络中，寻找主机靠的是**物理地址/MAC地址**(数据链路层)。相似的，为了进一步找到主机当中要接收这个数据的具体进程，在传输层中用**端口**来唯一标识主机中的具体应用进程。

端口号长度：16位

2014——65535，统称为**动态端口**，是因为它**一般不固定分配某种服务，而是动态分配**。动态分配是指当一个系统进程或应用 程序进程需要网络通信时，它向主机申请一个端口，主机从可用的端口号中分配一个供它使用。当这个进程关闭时，同时也就释放了所占用的端口号。

传输层的寻址与端口

应用程序	FTP	TELNET	SMTP	DNS	TFTP	HTTP	SNMP
熟知端口号	21	23	25	53	69	80	161

发现 **FTP** 谈恋爱 **TELNET** 删好友 **SMTP** 打电话 **DNS** 还要再见 **HTTP**

即最终通信的终点

在网络中采用发送方和接收方的套接字组合来识别**端点**，套接字唯一标识了网络中的一个主机和它上面的一个进程。

① 套接字Socket=（主机IP地址， 端口号）

5.2 UDP 协议

用户数据报协议UDP概述

UDP只在IP数据报服务之上增加了很少功能，即复用分用和差错检测功能。

UDP的主要特点：

- 1.UDP是无连接的，减少开销和发送数据之前的时延。
- 2.UDP使用最大努力交付，即不保证可靠交付。
- 3.UDP是面向报文的，适合一次性传输少量数据的网络应用。
也避免报文丢失，损失过大。
- 4.UDP要求延时不能过大，但允许可能丢失一些数据，如视频会议等。补救措施：向前纠错，重传等。
- 5.UDP首部开销小，8B，TCP20B。

8B
因为要求不高，实现功能不多



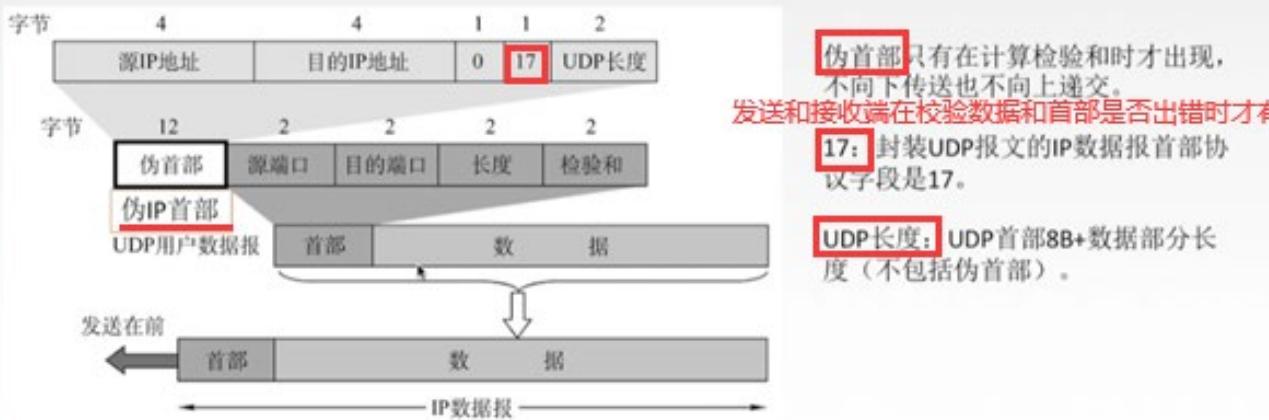
UDP首部格式



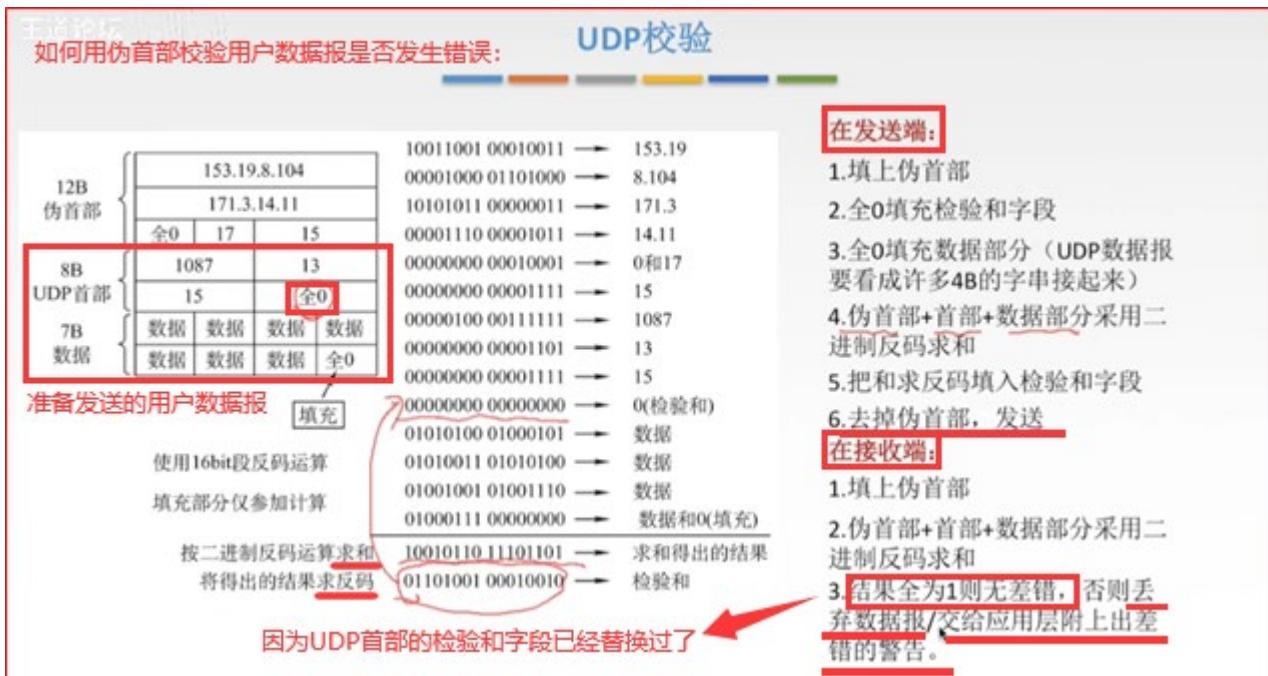
分用时, 找不到对应的目的端口号, 就丢弃报文, 并给发送方发送ICMP“端口不可达”差错报告报文。

UDP校验

首部协议字段, 说明数据部分使用什么协议, 这里UDP协议就是17



二进制反码运算求和: <https://www.cnblogs.com/jcchan/p/10400504.html>



传输层使用 UDP 协议时，就只能依靠应用层实现可靠传输。

5.3.1 TCP 协议特点和 TCP 报文段格式



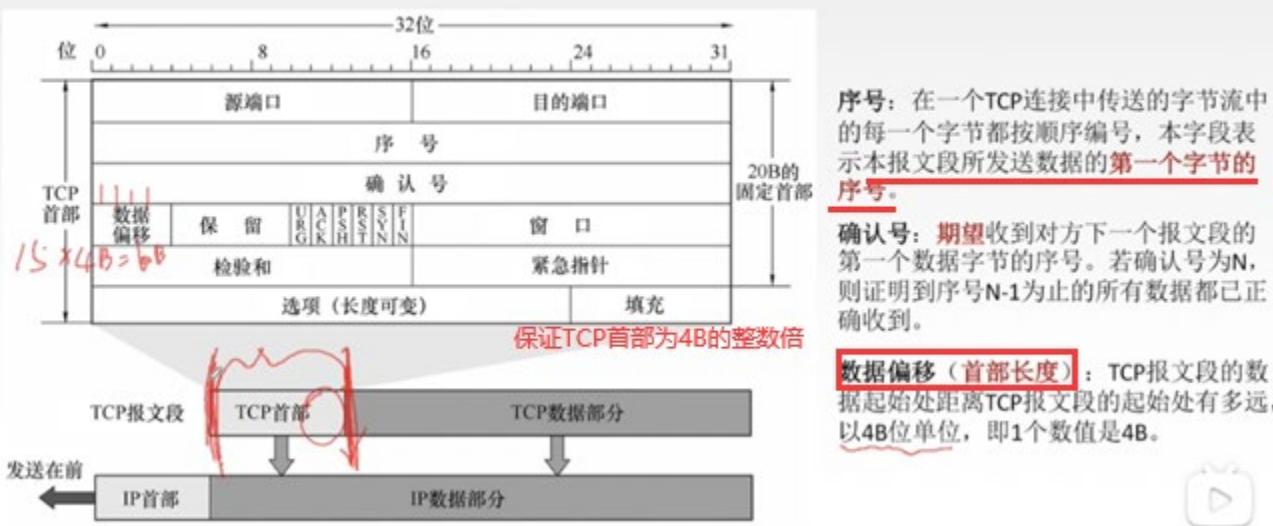
虚连接：像是两个进程间建立的点对点的直接连接。 完整的物理连接：向下一层层的封装，传输到接收端后再一层一层解封装。

TCP 面向字节流：TCP 把应用程序交下来的数据看成仅仅是一连串的无结构的字节流。

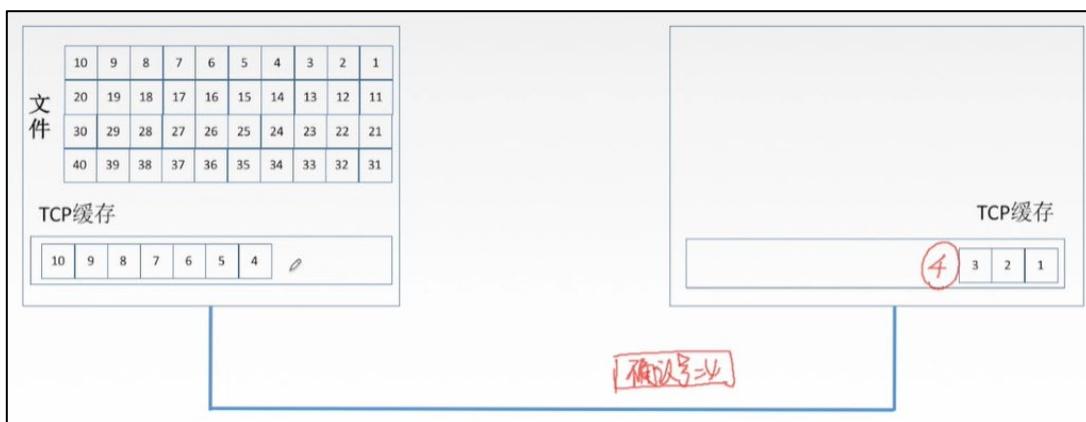
TCP协议的特点



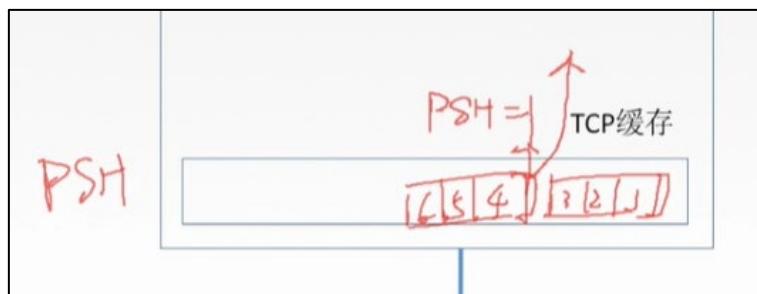
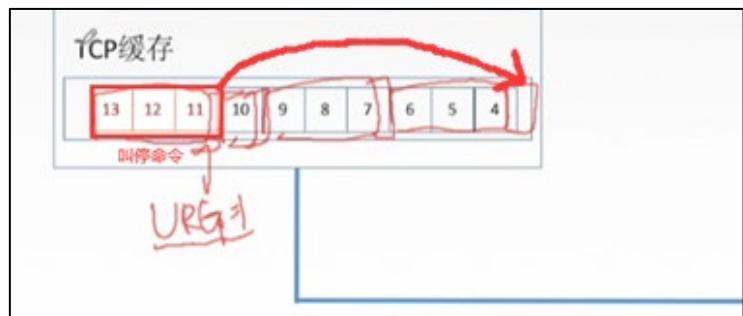
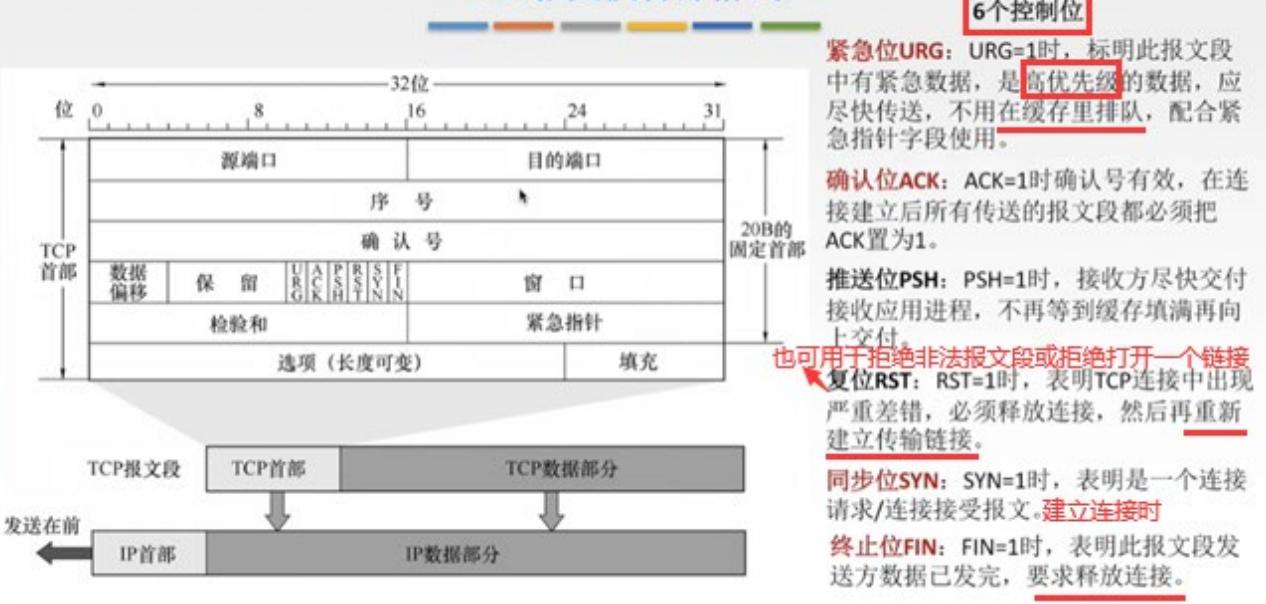
TCP报文段首部格式



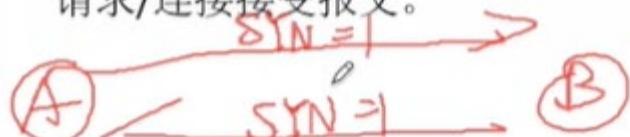
一个报文段从发送方发送, 收到接收方的确认报文段后才发送新的报文段。那么在确认报文段中, 就包含确认号, 即希望收到的下一个报文段的第一个数据字节的序号。

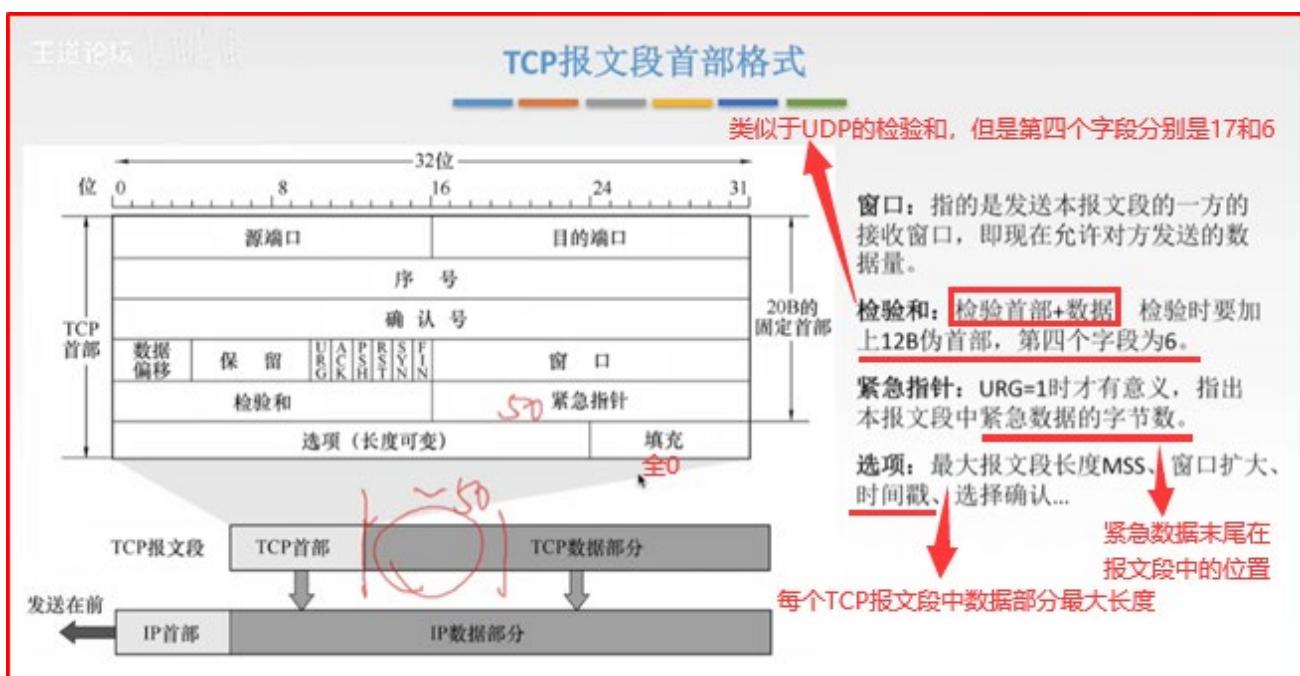
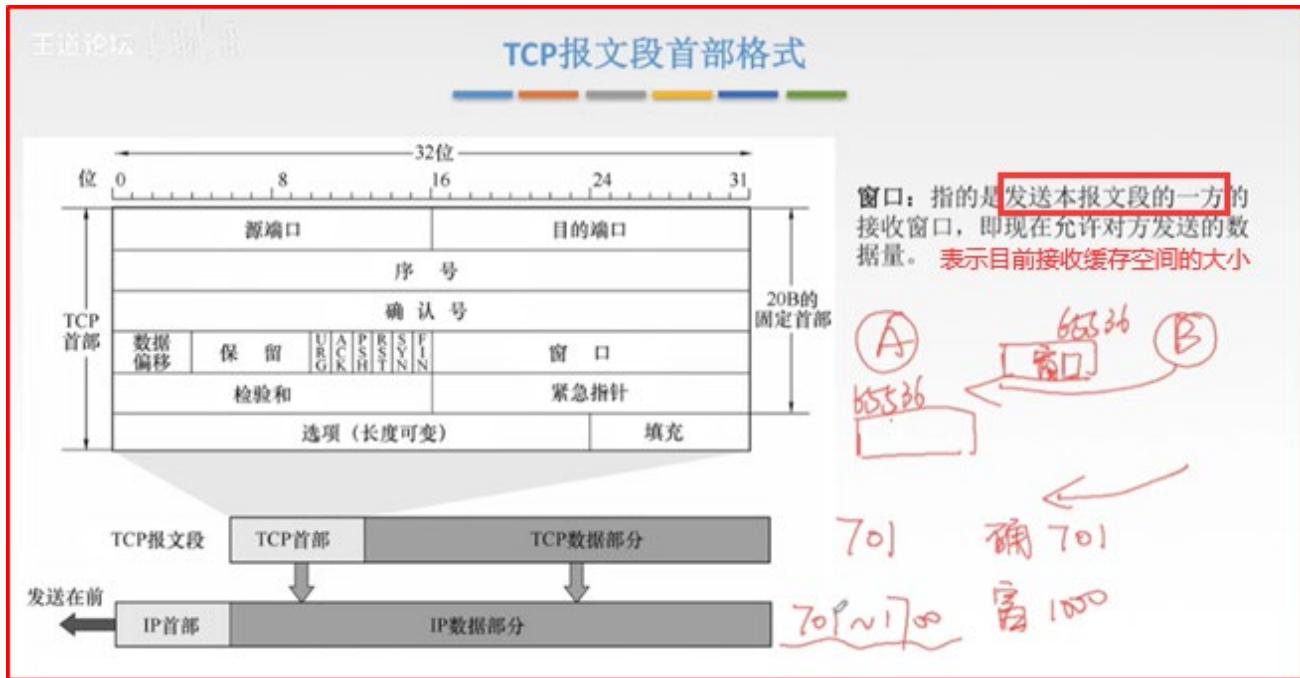


TCP报文段首部格式



同步位SYN: SYN=1时，表明是一个连接请求/连接接受报文。





5.3.2 TCP 连接管理

点对点、可靠、全双工

连接建立：三次握手
连接释放：四次挥手

TCP连接管理

TCP连接传输三个阶段：



TCP连接的建立采用**客户服务器方式**，主动发起连接建立的应用进程叫做客户，而被动等待连接建立的应用进程叫服务器。**但同时双方都可以发送和接收**

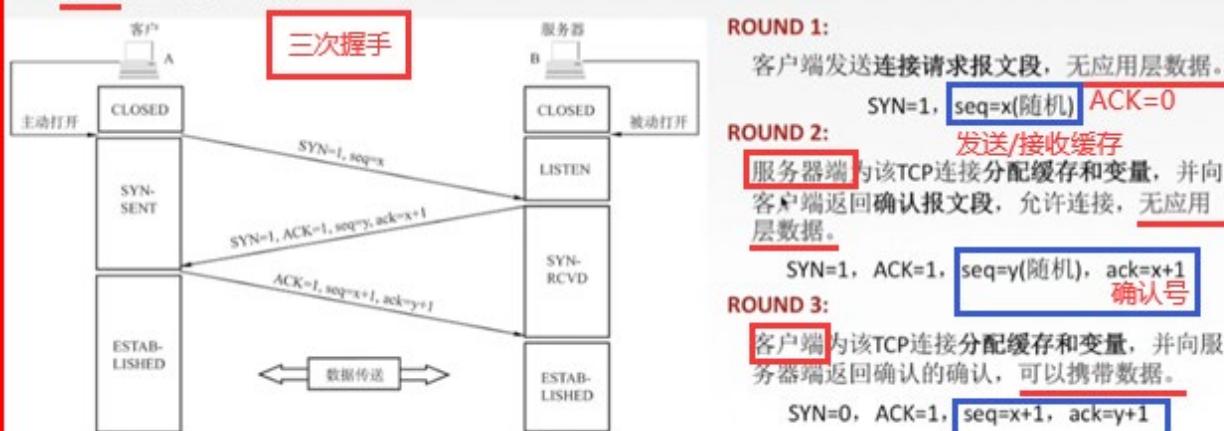


区分客户端和服务端：看是谁主动发起连接。

<https://www.itheima.com/news/20180720/165038.html>

TCP的连接建立

假设运行在一台主机（客户）上的一个进程想与另一台主机（服务器）上的一个进程建立一条连接，客户应用进程首先通知客户TCP，他想建立一个与服务器上某个进程之间的连接，客户中的TCP会用以下步骤与服务器中的TCP建立一条TCP连接：



SYN洪泛攻击

SYN洪泛攻击发生在OSI第四层，这种方式利用TCP协议的特性，就是三次握手。攻击者发送TCP SYN，SYN是TCP三次握手中的第一个数据包，而当服务器返回ACK后，该攻击者就不对其进行再确认，那这个TCP连接就处于挂起状态，也就是所谓的半连接状态，服务器收不到再确认的话，还会重复发送ACK给攻击者。这样更加会浪费服务器的资源。攻击者就对服务器发送非常大量的这种TCP连接，由于每一个都没法完成三次握手，所以在服务器上，这些TCP连接会因为挂起状态而消耗CPU和内存，最后服务器可能死机，就无法为正常用户提供服务了。

解决方法

SYN cookie

TCP的连接释放

四次握手



TCP的连接释放

参与一条TCP连接的两个进程中的任何一个都能终止该连接，连接结束后，主机中的“资源”（缓存和变量）将被释放。



最后要等待 2MSL 的时间：如果 B 未收到第四个确认报文段，会重传第三个连接释放报文段；当 A 在 2MSL 时间内再次收到连接释放报文段后，会重新发送第四个确认报文段，并重启 2MSL 计时器。

若没有这个等待时间(A 直接断开连接)，当第四个确认报文段丢失后，B 就会一直重发第三个连接释放报文段，从而无法正常进入关闭状态。

MSL：“最长报文段寿命”。任何报文在网络上存在的最长的最长时间，超过这个时间报文将被丢弃。

5.3.3 TCP 可靠传输

可靠：无差错、不丢失、不重复、按序到达

TCP可靠传输

传输层

使用TCP实现可靠传输

网络层

提供尽最大努力交付，不可靠传输

可靠

保证接收方进程从缓存区读出的字节流与发送方发出的字节流是完全一样的。

TCP实现可靠传输的机制

1.校验

2.序号

3.确认

4.重传

与UDP校验一样，
增加伪首部



发送方要收到接收方发来的确定收到信息(确认报文段或捎带确认报文段)，才能将相应的报文段从发送缓存中移除。



下一次收到 456 后，返回的确认报文段中首部确认号字段就为 9 了。

重传

? ——————

确认重传不分家，TCP的发送方在规定的时间内没有收到确认就要重传已发送的报文段。**超时重传**
重传时间

TCP采用自适应算法，动态改变重传时间RTTs（加权平均往返时间）。



重传

? ——————

确认重传不分家，TCP的发送方在规定的时间内没有收到确认就要重传已发送的报文段。**超时重传**
重传时间

TCP采用自适应算法，动态改变重传时间RTTs（加权平均往返时间）。

等太久！！！

冗余ACK（冗余确认）

改进

每当比期望序号大的失序报文段到达时，发送一个**冗余ACK**，指明下一个期待字节的序号。

发送方已发送1, 2, 3, 4, 5报文段

接收方收到1，返回给1的确认（确认号为2的第一个字节）

接收方收到3，仍返回给1的确认（确认号为2的第一个字节）

接收方收到4，仍返回给1的确认（确认号为2的第一个字节）

接收方收到5，仍返回给1的确认（确认号为2的第一个字节）

发送方收到**3个对于报文段1的冗余ACK** → 认为2报文段丢失，重传2号报文段 **快速重传**

每当收到失序报文段时，都会发送一个冗余 ACK。不然使用的就是累计确认。

5.3.4 TCP 流量控制

王道论坛

TCP流量控制

流量控制：让发送方慢点，要让接收方来得及接收。

TCP利用滑动窗口机制实现流量控制。

在通信过程中，接收方根据自己接收缓存的大小，动态地调整发送方的发送窗口大小，即接收窗口rwnd（接收方设置确认报文段的窗口字段来将rwnd通知给发送方），发送方的发送窗口取接收窗口rwnd和拥塞窗口cwnd的最小值。

发送方

接收方

接收方发送的报文段中窗口字段也可是0，此时发送方停止发送，接收方将接收缓存中的数据上交给应用进程，腾出空间。随后再返回一个报文段，使得发送窗口大小可以动态变化，得发送方继续发送。

王道论坛

TCP流量控制

TCP不使用停等协议

A向B发送数据，连接建立时，B告诉A：“我的rwnd=400（字节）”，设每一个报文段100B，报文段序号初始值为1。

主机A

主机B

发送窗口: 100B rwnd = 400

seq = 1, DATA → A发送了序号1至100，还能发送300字节

seq = 101, DATA → A发送了序号101至200，还能发送200字节

seq = 201, DATA → 丢失！

ACK = 1, ack = 201 rwnd = 300

seq = 301, DATA → 允许A发送序号201至500共300字节

seq = 401, DATA → A发送了序号301至400，还能再发送100字节新数据

seq = 201, DATA → A超时重发旧的数据，但不能发送新的数据

ACK = 1, ack = 501 rwnd = 100

seq = 501, DATA → 允许A发送序号501至600共100字节

ACK = 1, ack = 601 rwnd = 0

不允许A再发送（到序号600为止的数据都收到了）

窗口字段动态调整发送窗口大小 即进行流量控制

发送窗口滑动，具体可以复习链路层滑动窗口法

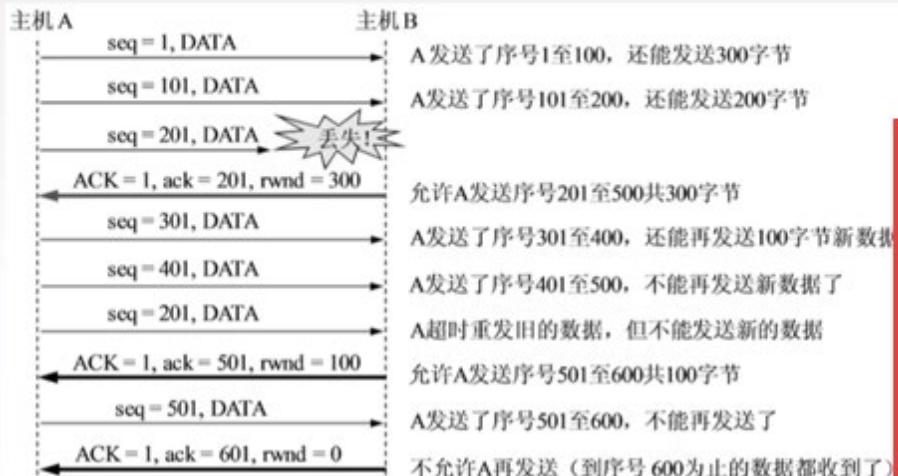
B 处理完(上交)接收缓存中的数据后，即直到 B 发出新的窗口值(如 rwnd=400)为止，A 才能继续发送数据。那万一新窗口值的报文段丢失了，AB 就会一直相互等待，怎么解决呢？

↓
持续计时器和零窗口探测报文段

滑动窗口的过程可以见：[数据链路层的 3.4.3\(后退 N 帧协议\)](#)、[3.4.4\(选择重传协议\)](#)

TCP流量控制

A向B发送数据，连接建立时，B告诉A：“我的rwnd=400（字节）”，设每一个报文段100B，报文段序号初始值为1。



TCP为每一个连接设有一个持续计时器，只要TCP连接的一方收到对方的零窗口通知，就启动持续计时器。

若持续计时器设置的时间到期，就发送一个零窗口探测报文段。接收方收到探测报文段时给出现在的窗口值。

若窗口仍然是0，那么发送方就重新设置持续计时器。

5.3.5 TCP 拥塞控制

网络状况不好，发生了阻塞。

资源：比如带宽、交换节点的缓存、交换节点中的处理机等。

TCP拥塞控制

出现拥塞的条件：

对资源需求的总和 > 可用资源

网络到处都走不通，“瘫痪了”

网络中有许多资源同时呈现供应不足

→ 网络性能变坏 → 网络吞吐量将随输入负荷增大而下降

拥塞控制：

防止过多的数据注入到网络中。全局性

拥塞控制 & 流量控制

网络拥堵使发送方发送的数据到不了接收方



发送方发送过快，使得收方来不及接收了



① 慢开始 ② 拥塞避免
 ③ 快重传 ④ 快恢复

→假定：

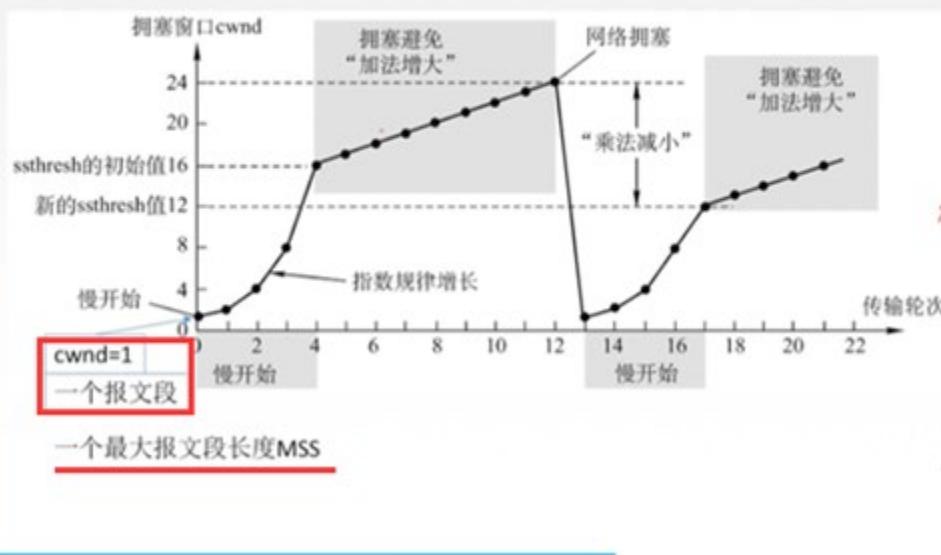
- { 1. 数据单方向传送，而另一个方向只传送确认
- 2. 接收方总是有足够的缓存空间，因而发送窗口大小取决于拥塞程度

发送窗口= $\min\{\text{接收窗口rwnd}, \text{拥塞窗口cwnd}\}$

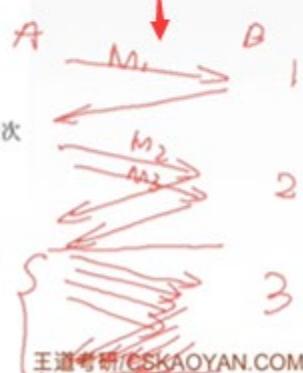
接收窗口 **接收方**根据接受缓存设置的值，并告知给发送方，反映接收方容量。

拥塞窗口 **发送方**根据自己估算的网络拥塞程度而设置的窗口值，反映网络当前容量。

慢开始和拥塞避免

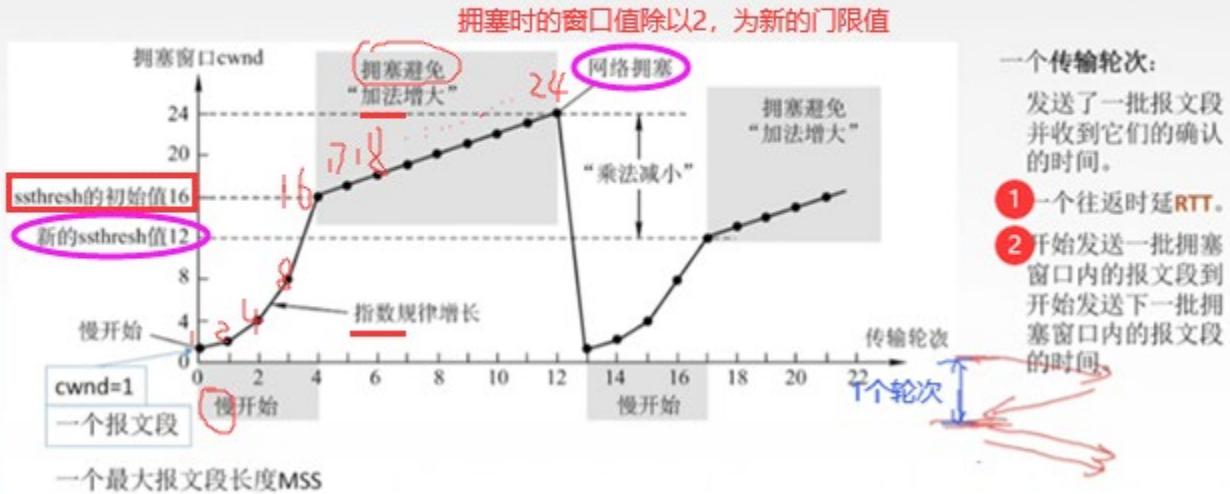


一个传输轮次：
发送了一批报文段并收到它们的确认的时间。



注意区分与流量控制中 rwnd 的单位。

慢开始和拥塞避免

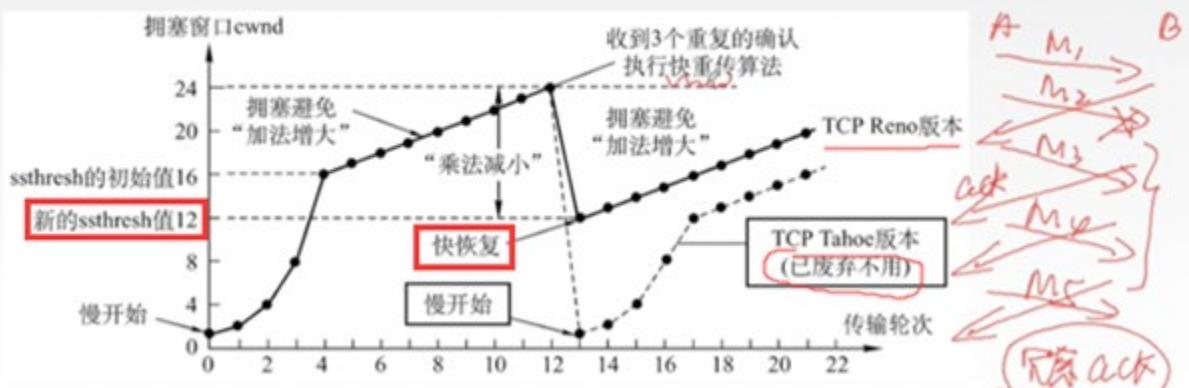


RTT: 报文段往返时延!

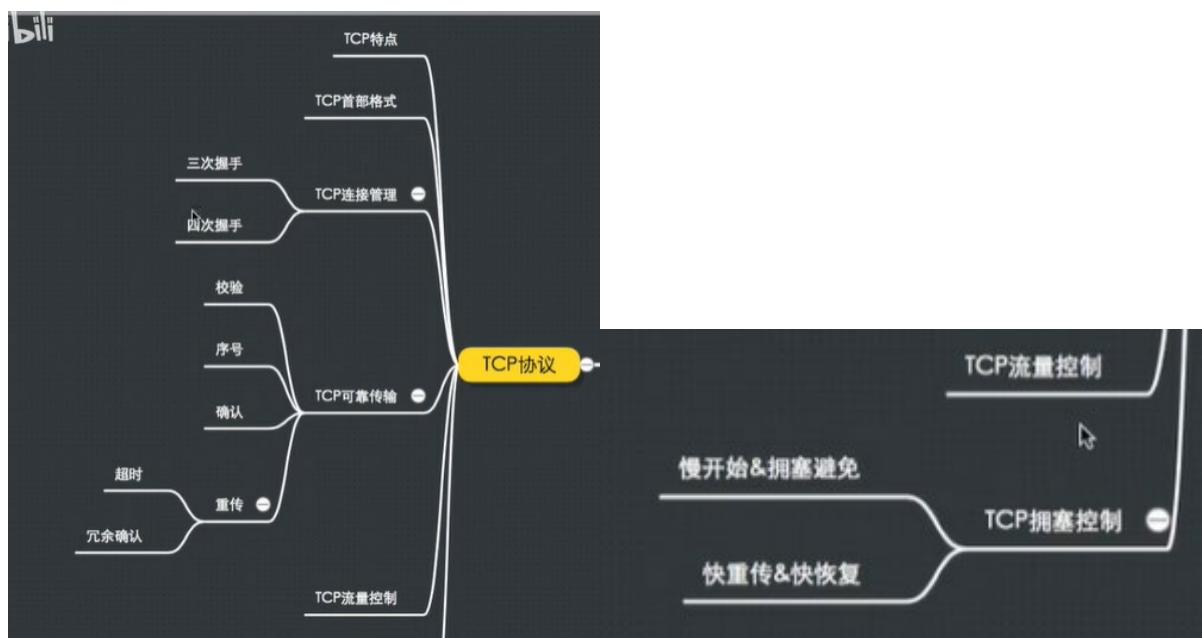
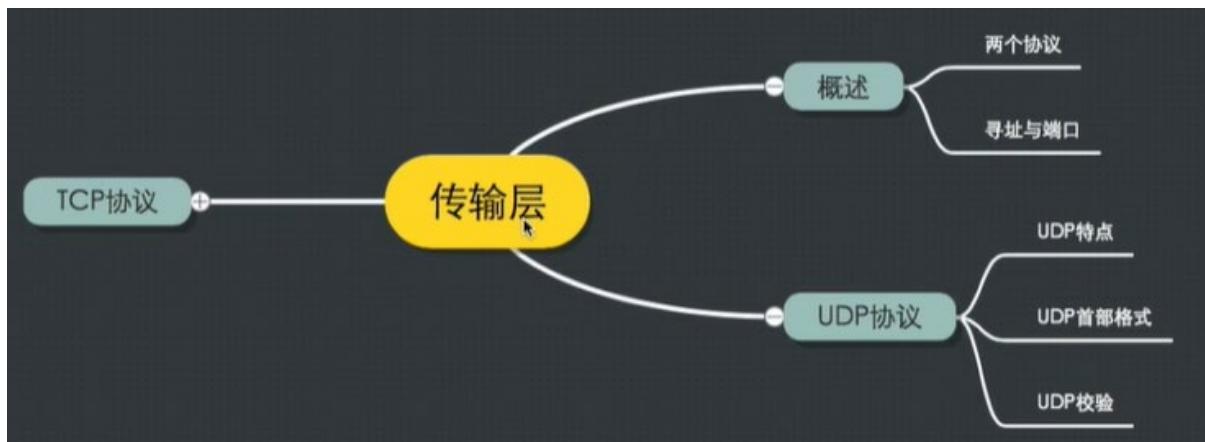
慢开始: 一开始出入少量报文段。先对网络进行探查。

慢开始时, 只要收到上一轮次发送的报文段的确认, 就立即将拥塞窗口 cwnd 翻倍, 然后再去发送报文段。

快重传和快恢复



5.4 传输层总结



六、应用层

6.1 网络应用模型

为什么会有应用层？ -> 因为不同网络应用的应用进程之间需要有不同的通信规则。

应用层概述

应用层
传输层
网络层
数据链路层
物理层

应用层对应用程序的通信提供服务。

应用层协议定义:

应用进程交换的报文类型, 请求还是响应?
各种报文类型的语法, 如报文中的各个字段及其详细描述。
字段的语义, 即包含在字段中的信息的含义。
进程何时、如何发送报文, 以及对报文进行响应的规则。

应用层的功能:

文件传输、访问和管理
电子邮件
虚拟终端
查询服务和远程作业登录

应用层的重要协议:

FTP
SMTP、POP3
HTTP
DNS

虚拟终端: 使个人计算机能更方便使用大型计算机的功能。

查询服务: 如访问网页。

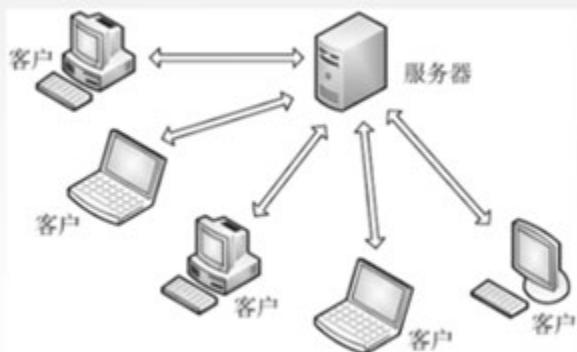
网络应用模型

应用层通常使用的两种网络应用模型:

客户/服务器模型 (Client/Server)

P2P模型 (Peer-to-peer) 对等模型

客户/服务器 (C/S) 模型



服务器: 提供计算服务的设备。

- 1. 永久提供服务
- 2. 永久性访问地址/域名

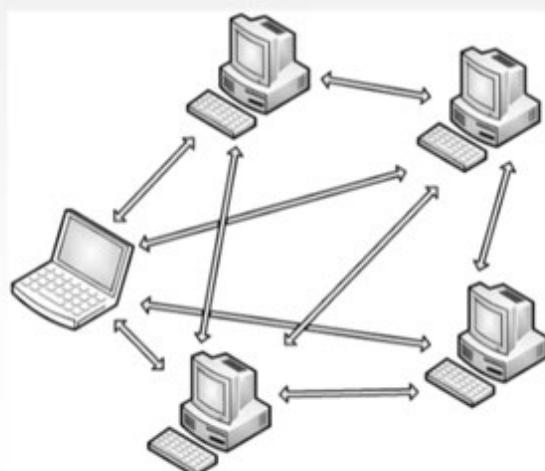
客户机: 请求计算服务的主机。

- 1. 与服务器通信，使用服务器提供的服务
- 2. 间歇性接入网络
- 3. 可能使用动态IP地址
- 4. 不与其他客户机直接通信

应用: Web, 文件传输FTP, 远程登录, 电子邮件

4. qq 诞生之前，一般是通过聊天室进行聊天(使用了服务器)，发送的数据都是需要通过服务器转发。

P2P模型



不存在永远在线的服务器

每个主机既可以提供服务，也可以请求服务

任意端系统/节点之间可以直接通讯

节点间歇性接入网络

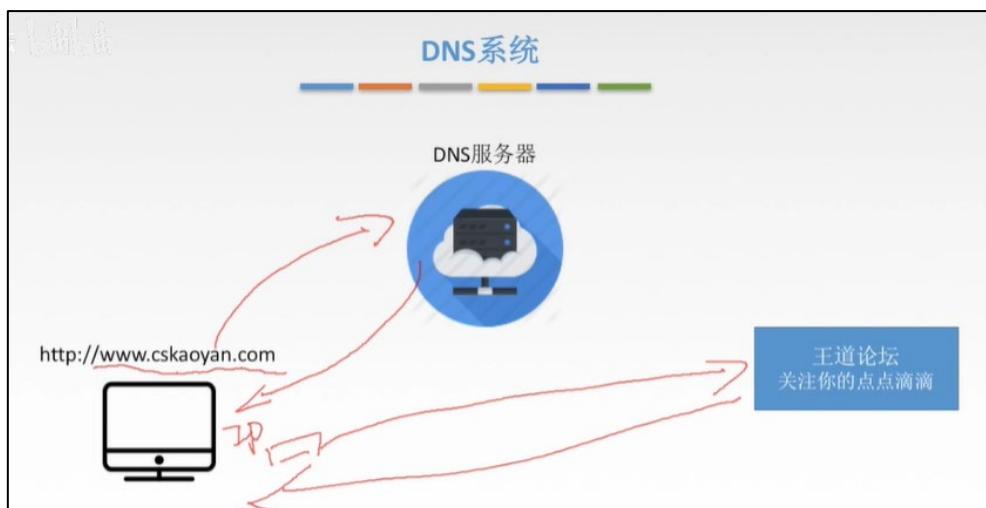
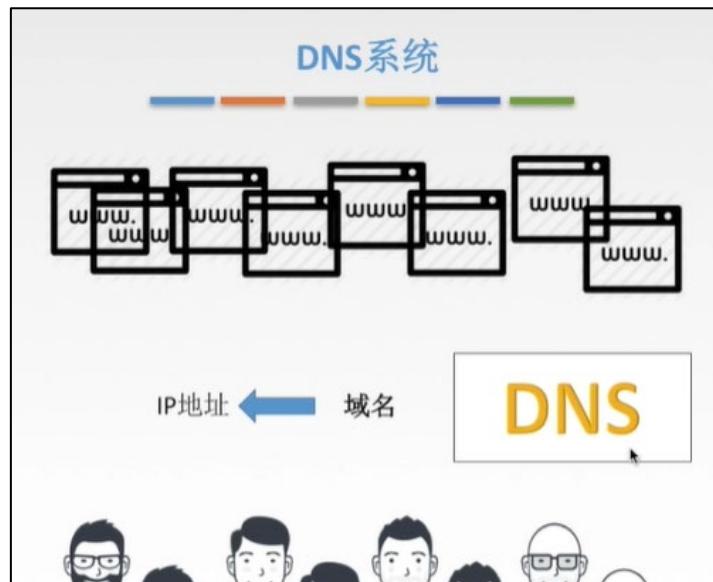
节点可能改变IP地址

可扩展性好 可以应对大量主机涌入的情况

网络健壮性强 一个节点出问题了，其他的不易受影响

6.2 DNS 系统/域名系统

访问网站: 实际就是主机与另一台主机远程通信和资源交换的过程。并且使通过 IP 地址(全球唯一标识符)找到对方主机。但是记忆 IP 地址太麻烦，所以就出现了域名。



点分隔的是**标号**, 每个标号理论不超过 63 个字符, 但为了记忆一般不超过 12 个字符(不区分大小写)。
标 低到高。





三级域名

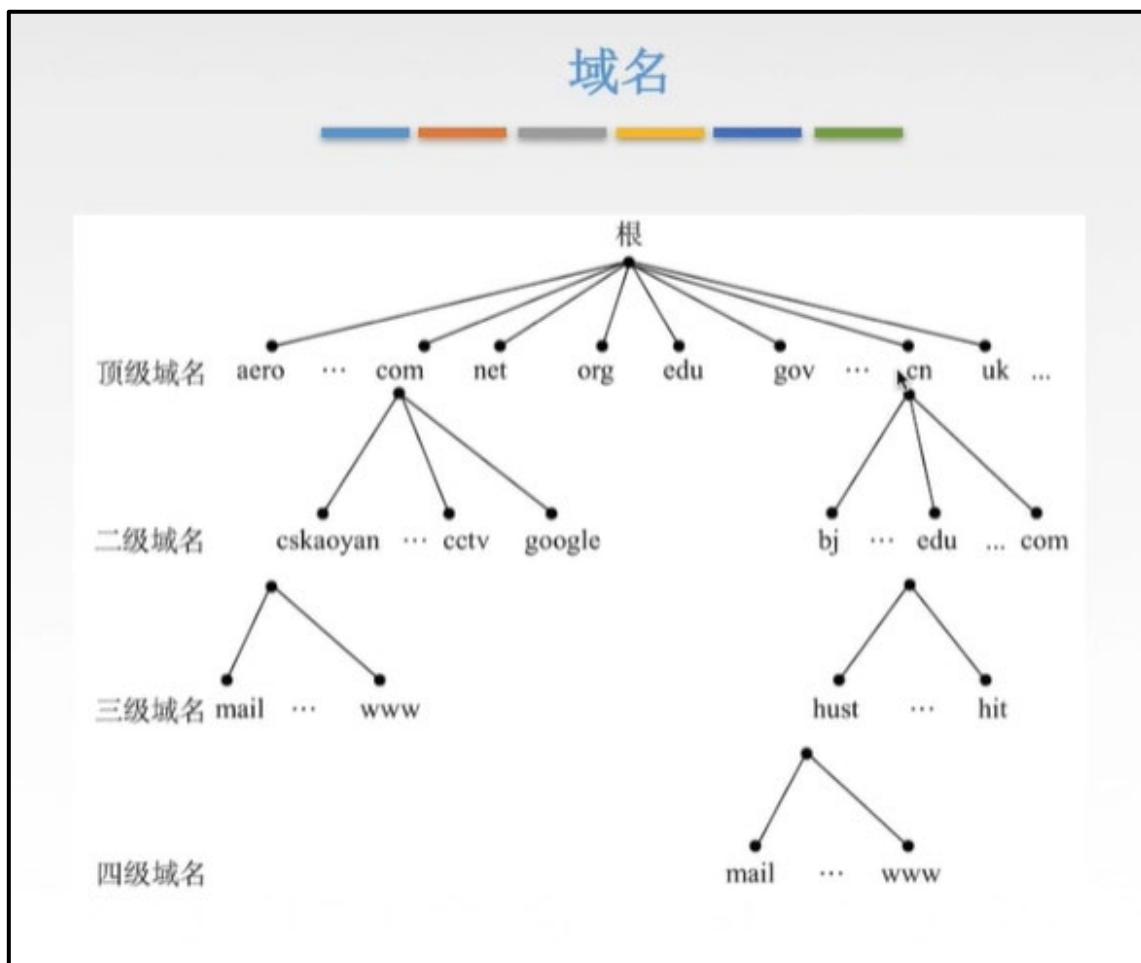
www.pku.edu.cn 北大主页

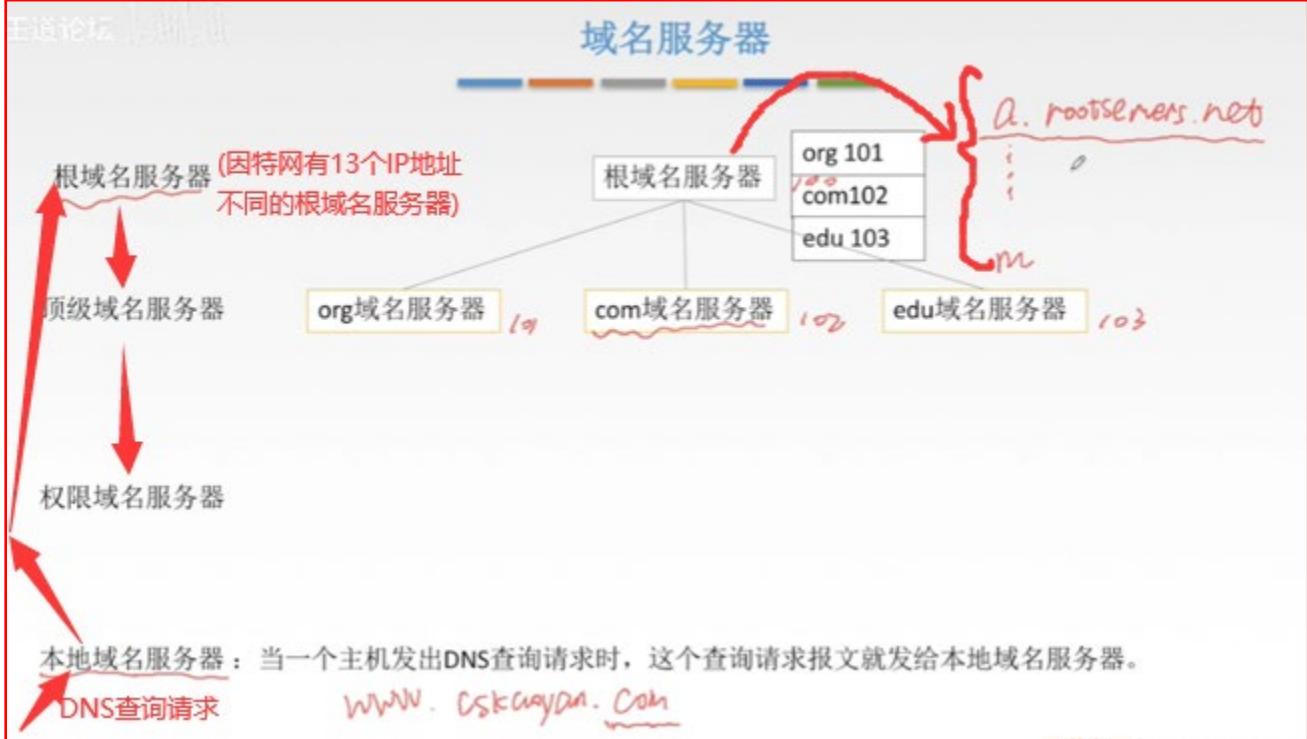
四级域名

mail 北大提供邮件服务的域名

ftp

北大提供文件传输服务的域名





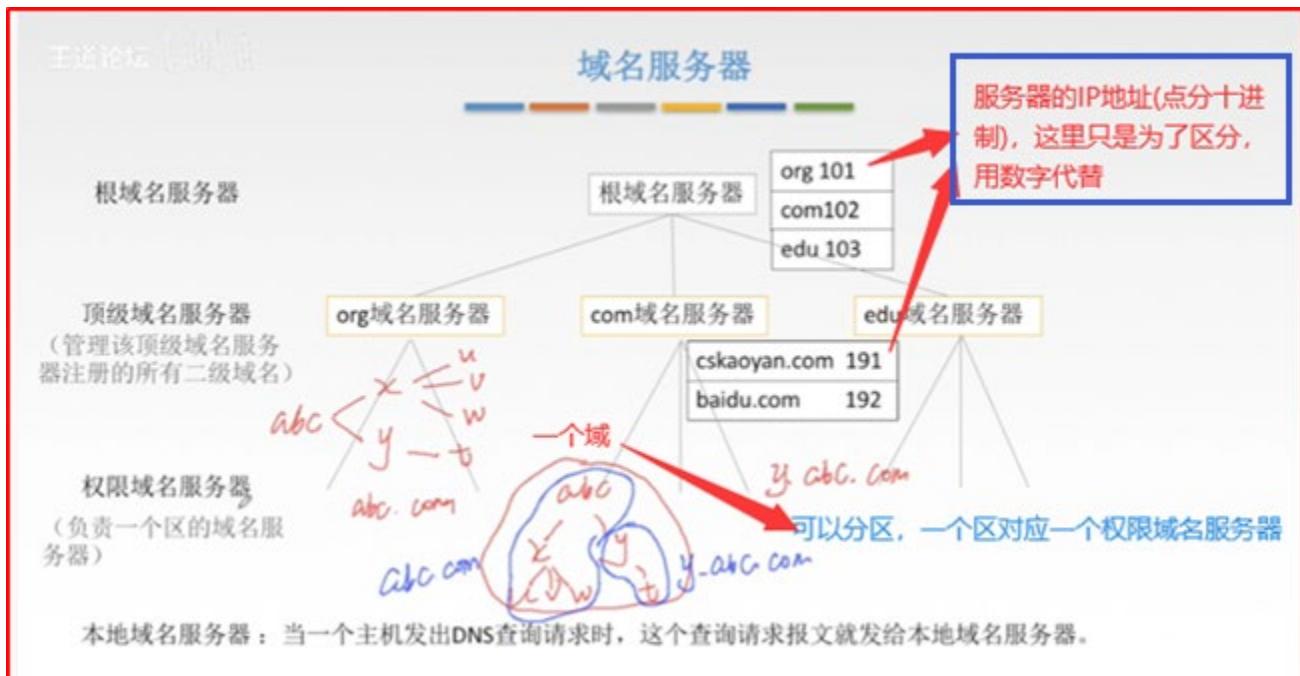
本地域名服务器(默认域名服务器)是离主机最近的 DNS 服务器(不超过几个路由器的距离)，一个学校一个院系都可以拥有。一台主机和要查询的另一台主机在一个本地 ISP 服务范围内时，本地域名服务器即可快速查询转换。

主机->(DNS 查询请求)->本地域名服务器->(缓存中查询不到)->向根域名服务器求助。

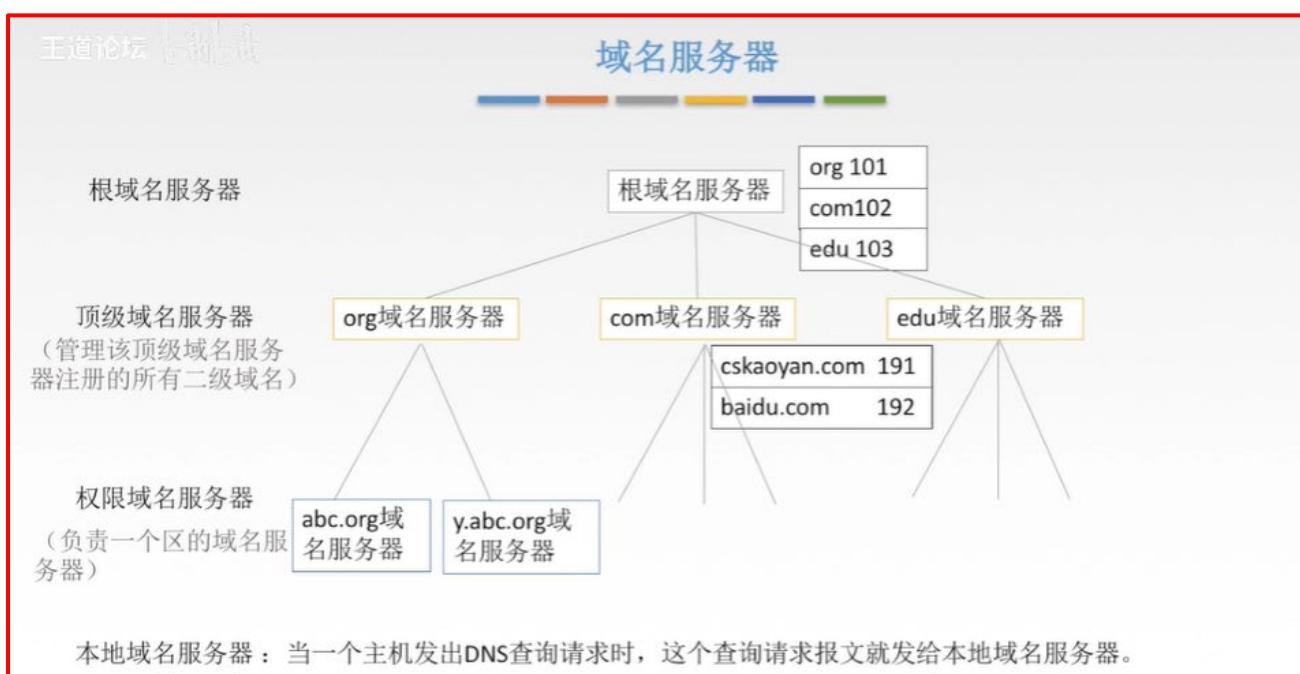
根域名服务器知道所有顶级域名服务器的 IP 地址和域名

根服务器主要用来管理互联网的主目录，最早是 IPv4，全球只有 13 台 (这 13 台 IPv4 根域名服务器名字分别为“A”至“M”)，1 个为主根服务器在美国，由美国互联网机构 Network Solutions 运作。其余 12 个均为辅根服务器，其中 9 个在美国，2 个在欧洲(位于英国和瑞典)，1 个在亚洲(位于日本)。但是中国有镜像根域名服务器。

在与现有 IPv4 根服务器体系架构充分兼容基础上，“雪人计划”于 2016 年在全球 16 个国家完成 25 台 IPv6 根服务器架设，事实上形成了 13 台原有根加 25 台 IPv6 根的新格局，为建立多边、民主、透明的国际互联网治理体系打下坚实基础。中国部署了其中的 4 台，由 1 台主根服务器和 3 台辅根服务器组成，打破了中国过去没有根服务器的困境



abc 公司可以整体为一个区，若公司扩张了也可分为多个区(即一个分支/一个域 abc 可以进一步分区，并且各分区是对等的关系)。类似于子网？



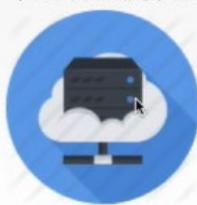
如上图中，二级域名和三级域名在域名服务器这里是对等的关系，分别对应两台权限域名服务器。

域名解析过程

递归查询

迭代查询

本地域名服务器



靠别人



靠自己



递归查询: 从高往低依次寻求帮助, 最终再按照相反的顺序依次返回发送到查询请求的主机。

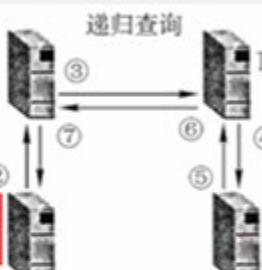
迭代查询: 若**本地域名服务器**中没有要查询的域名所对应的 IP 地址。则首先查询**根域名服务器**; 若还未解析完, 由**本地域名服务器**再去查询**顶级域名服务器**(**根域名服务器**告诉它去找哪个**顶级域名服务器**)……以此类推。

域名解析过程

递归查询

递归查询

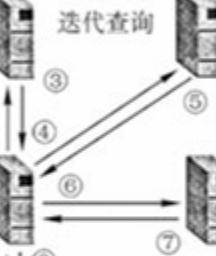
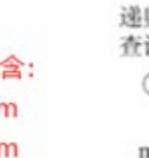
根域名服务器

本地域名服务器
dns.xyz.com顶级域名服务器
dns.com权限域名服务
dns.abc.com

可能就包含:
x.abc.com
y.abc.com

(a) 递归查询(比较少用)

根域名服务器

本地域名服务器
dns.xyz.com

顶级域名服务器
dns.com

权限域名服务
dns.abc.com



(b) 递归与迭代相结合的方式

递归和迭代都需要很多次的查询和 DNS 请求, 为了提高效率, 引入了**高速缓存(动态/定期更新)**。即在**本地域名服务器**中存储最近查过的域名以及从哪里获得域名映射信息的记录。下次访问时直接由**本地域名服务器**返回; 或者没有具体域名对应的 IP 地址, 但是有**顶级域名服务器**的 IP 地址, 也可以跳过**根域名服务器**, 直接到**顶级域名服务器**以获得下一步的信息。

很多主机也有类似的高速缓存, 每次开机时, 向**本地域名服务器**下载获取域名和地址对应的数据库。

6.3 文件传输协议 FTP

文件传送协议

主要的两种协议：

使用TCP

文件传送协议FTP (File Transfer Protocol)

适合UDP

简单文件传送协议TFTP (Trivial File Transfer Protocol) 面向小文件

很小，易于实现。适合于UDP环境，如文件需要同时被很多主机下载时。

代码块占用内存少，适用于小计算机或某些特殊设备

文件传送协议FTP (File Transfer Protocol)

提供不同种类主机系统（硬、软件体系等都可以不同）之间的文件传输能力。
可以屏蔽不同操作系统的差异性

简单文件传送协议TFTP (Trivial File Transfer Protocol)



拷贝 { 上传
 |
 | 下载

FTP服务器和用户端

FTP是基于客户/服务器（C/S）的协议。

用户通过一个客户机程序连接至在远程计算机上运行的服务器程序。

依照FTP协议提供服务，进行文件传送的计算机就是**FTP服务器**。

连接FTP服务器，遵循FTP协议与服务器传送文件的电脑就是**FTP客户端**。

FTP客户端软件



使用：在资源管理器，输入 **ftp+地址**，回车后就可以进入到 ftp 服务器下面的文件目录，即可对文件进行上传下载等操作。

FTP工作原理

登陆

ftp地址 用户名&密码

匿名登陆

互连网中有很大一部分FTP服务器被称为“匿名”（Anonymous）FTP服务器。这类服务器的目的是向公众提供文件拷贝服务，不要求用户事先在该服务器进行登记注册，也不用取得FTP服务器的授权。

Anonymous（匿名文件传输）能够使用户与远程主机建立连接并以匿名身份从远程主机上拷贝文件，而不必是该远程主机的注册用户。用户使用特殊的用户名“anonymous”登陆FTP服务，就可访问远程主机上公开的文件。

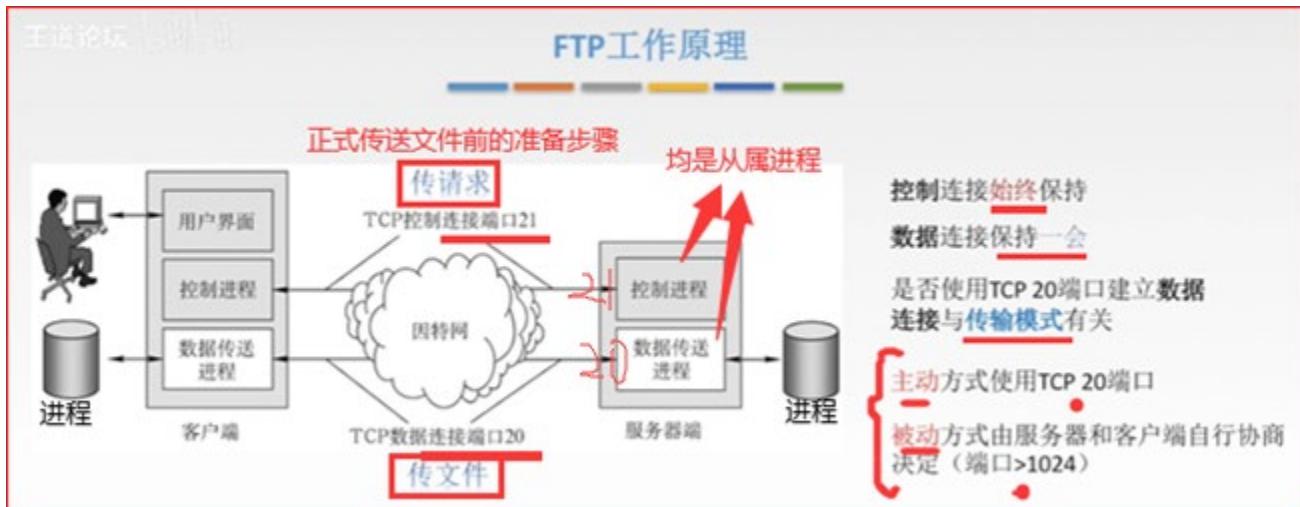
FTP使用TCP实现可靠传输。



主进程负责接受新的FTP连接请求。工作步骤：打开FTP服务器的熟知端口 21，可以使客户进程成功连接。

接下来等待客户进程的连接请求，连接成功后进行文件的传输。

接下来就可以启动从属进程，每个都可以处理单个的客户请求。



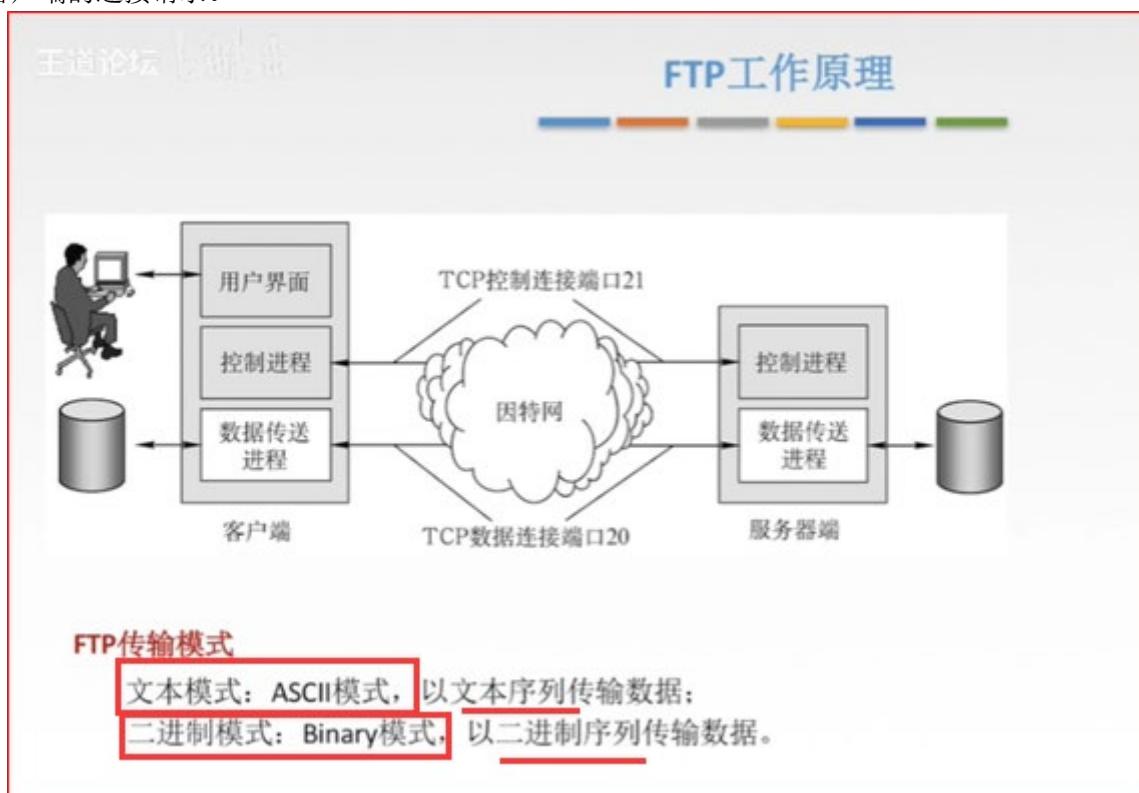
服务器的**控制进程**收到 FTP 客户端发来的**文件传输请求**后，创建**数据传送进程**并创建连接。

文件传输完毕后，关闭**数据传送连接**并结束进程；**控制连接**和**控制进程**全程保持打开，全程指的是客户端与服务器端建立会话并且会话还未结束。

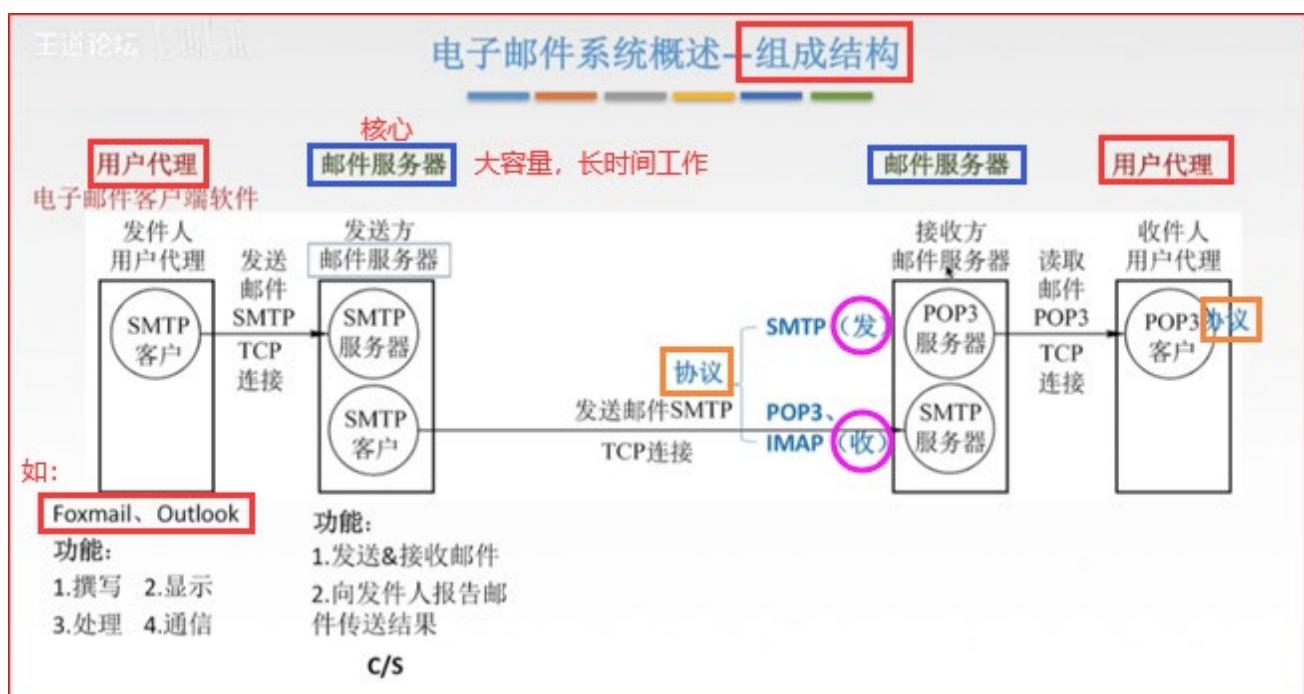
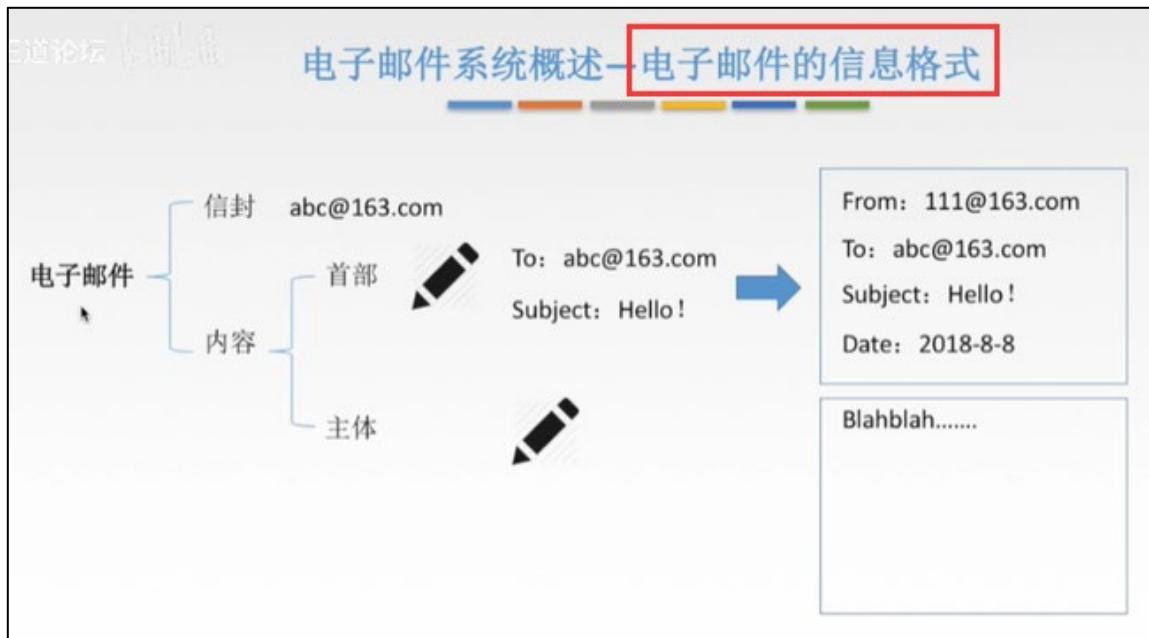
FTP 的控制连接和数据连接使用的是不同的端口号，所以会说 FTP 的**控制信息是带外传送的**。

主动模式：客户进程向服务器进程发送建立连接请求的时候，就要寻找连接服务器进程的熟知端口即 21，同时会告诉服务器自己用于建立他们之间的数据传送连接的进程的端口号。接着服务器端（主动）使用自己传送数据的熟知端口 20，与客户进程刚提供的端口号建立数据传输连接。

被动模式：控制连接的部分与主动模式一样。接下来客户端会给服务器端发送一个命令，**请求询问**服务器端给自己多少端口号来建立数据传送连接。服务器端收到该询问命令后，返回一个大于 1024 端口号。接着就等待客户端的连接请求。

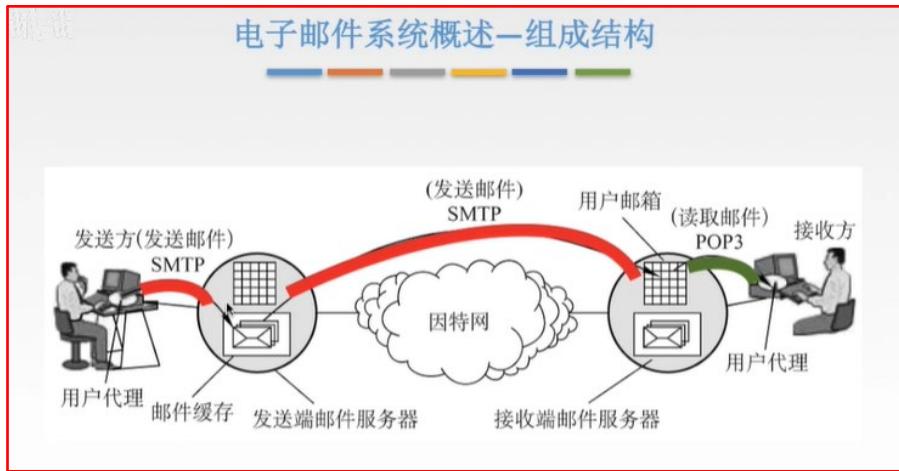


6.4 电子邮件



用户代理：用户和电子邮件系统的接口，通常是主机中的一个程序。

使用的均是 TCP 连接，可靠传输。



简单邮件传送协议SMTP

SMTP规定了在两个相互通信的**SMTP进程**之间应如何交换信息。

负责发送邮件的SMTP进程就是**SMTP客户**，负责接收邮件的进程就是**SMTP服务器**。

→ SMTP规定了14条命令（几个字母）和21种应答信息（三位数字代码+简单文字说明）。

① **TCP连接** ② **端口号25** ③ **C/S**

SMTP通信三个阶段：

连接建立 → 邮件传送 → 连接释放

简单邮件传送协议SMTP

1 连接建立



2 邮件发送

附上发送者的主机名

应答信息，准备接收

A : MAIL FROM: <wangdao@163.com>

B : 250 OK / B : 451 (452、500...) SMTP服务器是否已经准备好接收邮件

A : RCPT TO: <mooc@163.com> 可以有多个RCPT命令(群发)

B : 250 OK / B : 550 No such user here SMTP服务器确定是否有这个用户

A : DATA 要开始传输邮件的内容了

B : 354 start mail input; end with <CR><LF>.<CR><LF> SMTP服务器同意传输

A : Date.... 开始传输邮件内容 作为邮件内容结尾

B : 250 OK 接收结束

3 连接释放

邮件发完，SMTP客户发送QUIT命令，SMTP服务器返回“221”，表示同意释放TCP连接。

MIME

SMTP的缺点：

1. SMTP不能传送可执行文件或者其他二进制对象。
2. SMTP仅限于传送7位ASCII码，不能传送其他非英语国家的文字。
3. SMTP服务器会拒绝超过一定长度的邮件。

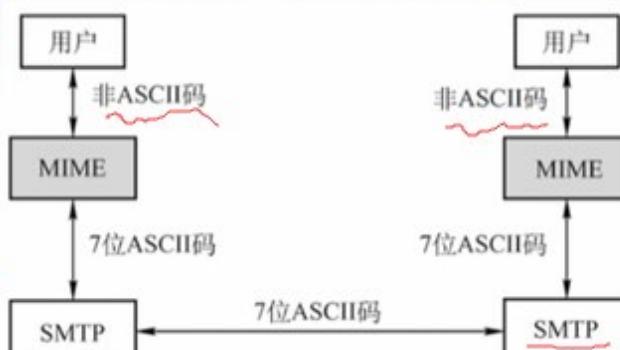
基于SMTP扩充

通用因特网邮件扩充MIME

使电子邮件系统可以支持声音、图像、视频、多种国家语言等等。

使得传输内容丰富多彩

现也多用于浏览器



MIME 协议最早用于**电子邮件**，现也多用于**浏览器**。浏览器服务器会将发送的多媒体数据类型告诉浏览器，使得浏览器知道收到的多媒体数据是什么 MIME 类型的 (MP3、MP4、JPG 等)，进而可以选择相应的插件去加载读取该数据。

论坛

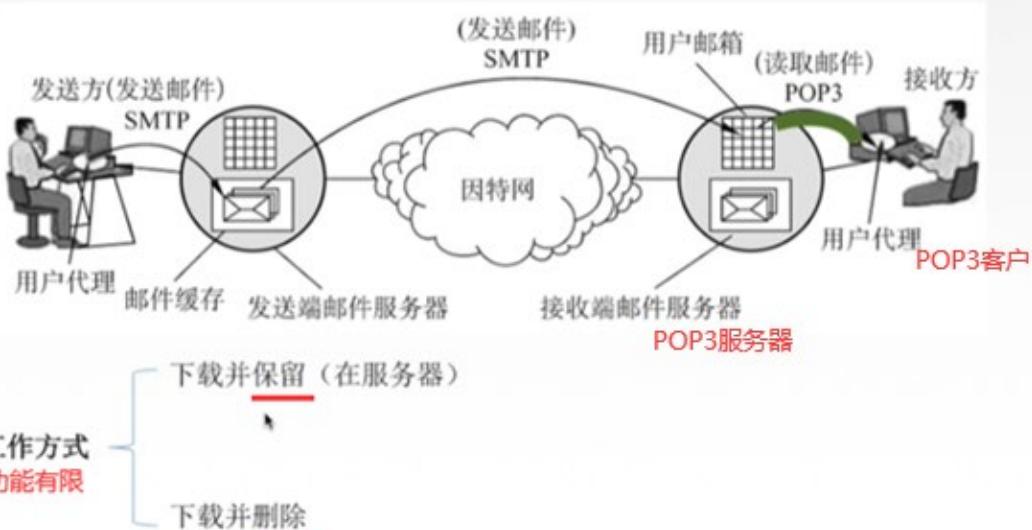
邮局协议POP3 第三版

TCP连接

端口号110

C/S

客户服务器方式



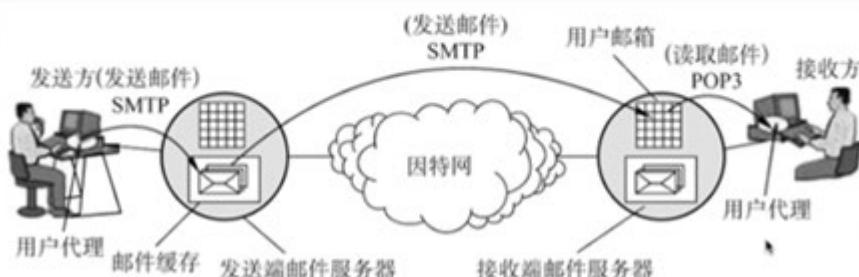
知道论坛

网际报文存取协议IMAP

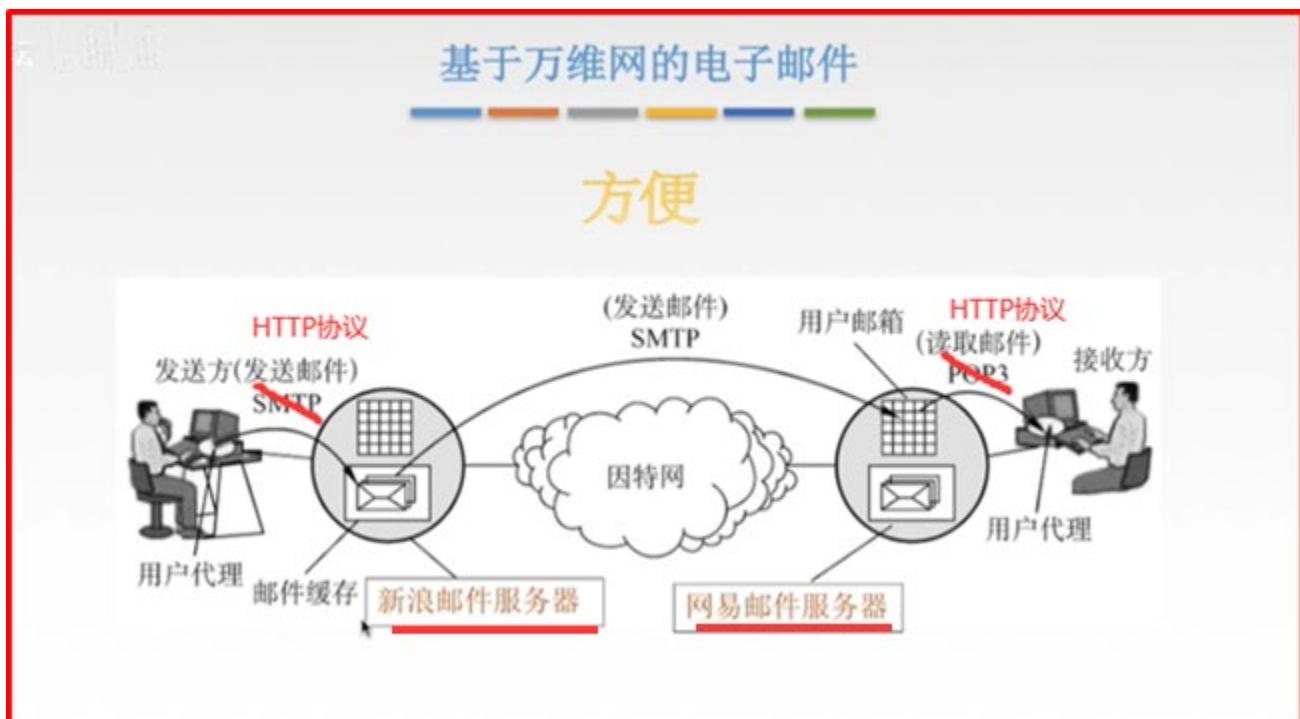
IMAP协议比POP协议复杂。当用户Pc上的IMAP客户程序打开IMAP服务器的邮箱时，用户可以看到邮箱的首部，若用户需要打开某个邮件，该邮件才上传到用户的计算机上。

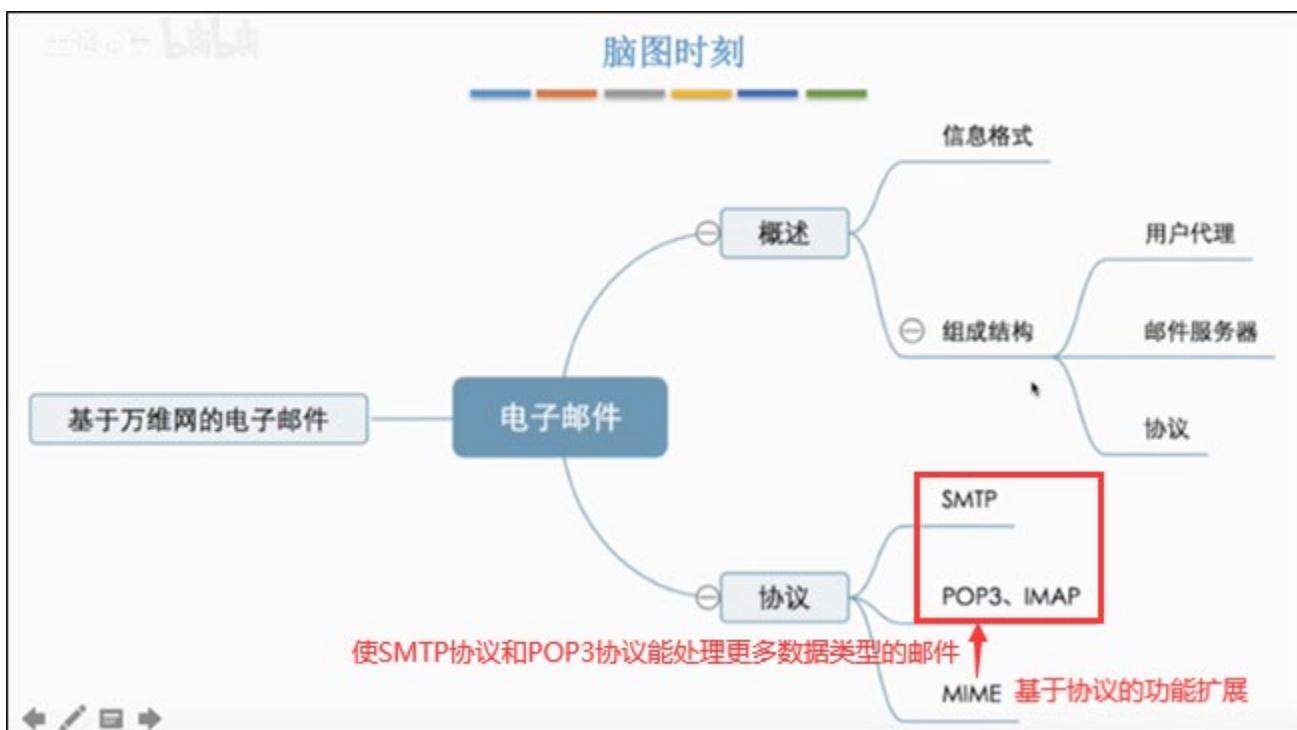
IMAP可以让用户在不同的地方使用不同的计算机随时上网阅读处理邮件，还允许只读取邮件中的某一个部分（先看正文，有WiFi的时候再下载附件）。

因为在下载到本地系统之前，可以在邮件服务器中直接读取部分内容



邮局协议 (POP3)	互联网消息访问协议 (IMAP)
POP是一种简单的协议，仅允许将邮件从服务器下载到本地计算机。	IMAP更为先进，它使用户可以查看邮件服务器上的所有文件夹。
POP服务器在端口110上侦听，而带SSL安全(POP3DS)服务器的POP在端口995上侦听	IMAP服务器侦听端口143，带有SSL安全(IMAPDS)服务器的IMAP侦听端口993。
在POP3中，一次只能从单个设备访问邮件。	可以跨多个设备访问消息
要阅读邮件，必须将其下载到本地系统上。	在下载之前，可以部分读取邮件内容。
用户无法在邮件服务器的邮箱中整理邮件。	用户可以直接在邮件服务器上组织电子邮件。
用户无法在邮件服务器上创建、删除或重命名电子邮件。	用户可以在邮件服务器上创建、删除或重命名电子邮件。
用户在下载到本地系统之前无法搜索邮件的内容。	用户可以在下载前搜索邮件内容中的特定字符串。
下载后，如果本地系统崩溃消息丢失，则该消息存在于本地系统中。	邮件服务器上会保留邮件的多个冗余副本，如果丢失本地服务器的邮件，仍可以检索邮件
可以使用本地电子邮件软件更改邮件。	Web界面或电子邮件软件所做的更改与服务器保持同步。
所有消息立即下载。	可以在下载前查看邮件头。





6.5 万维网和 HTTP 协议

王道论坛

万维网概述

万维网WWW（World Wide Web）是一个大规模的、联机式的信息储藏所/资料空间，是无数个网络站点和网页的集合。

统一资源定位符URL 唯一标识 资源（文字、视频、音频...）

URL一般形式：
 <协议>://<主机>:<端口>/<路径> <http://www.pku.edu.cn>

常用的 URL 构成部分：

http	域名	有时可省略
ftp	IP地址	

URL不区分大小写。

用户通过点击超链接（<http://www.baidu.com>）获取资源，这些资源通过超文本传输协议（HTTP）传送给使用者。

万维网以客户/服务器方式工作，用户使用的浏览器就是万维网客户程序，万维网文档所驻留的主机运行服务器程序。

万维网使用超文本标记语言HTML，使得万维网页面设计者可以很方便地从一个界面的链接转到另一个界面，并能够在自己的屏幕上显示出来。

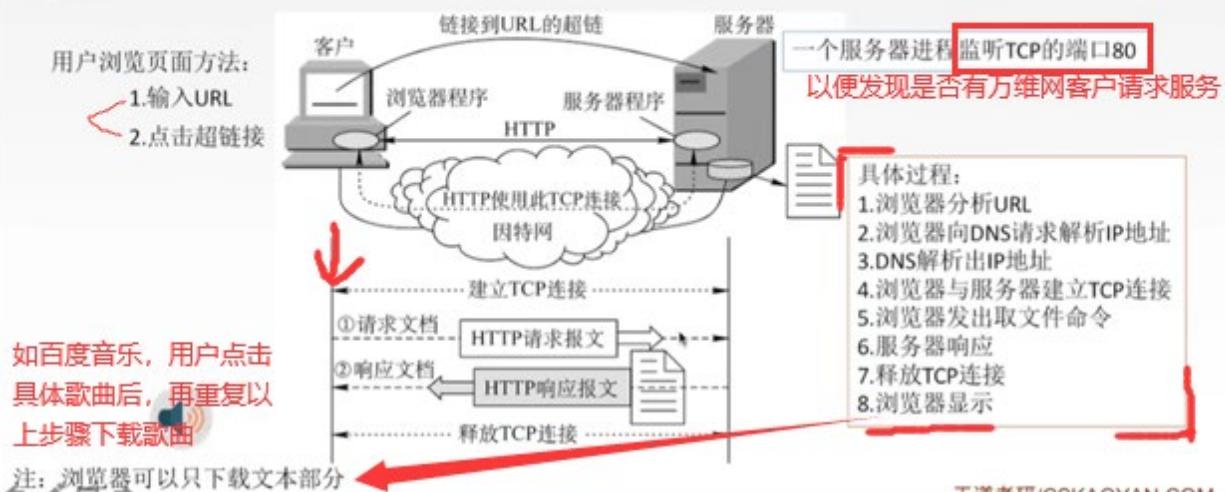
统一资源定位符 URL

超文本传输协议 HTTP

超文本标记语言 HTML

超文本传输协议HTTP

HTTP协议定义了浏览器（万维网客户进程）怎样向万维网服务器请求万维网文档，以及服务器怎样把文档传送给浏览器。



HTTP协议的特点

- ① HTTP协议是无状态的。



服务器每次响应速度是相同的

但是在实际工作中，一些万维网站点常常希望能够识别用户。



Cookie小饼干

Cookie是存储在用户主机中的**文本文件**，记录一段时间内某用户（使用识别码识别，如“123456”）的访问记录。

→ 提供个性化服务

如，淘宝服务器就可查看在主机中记录淘宝搜索记录的**Cookie**

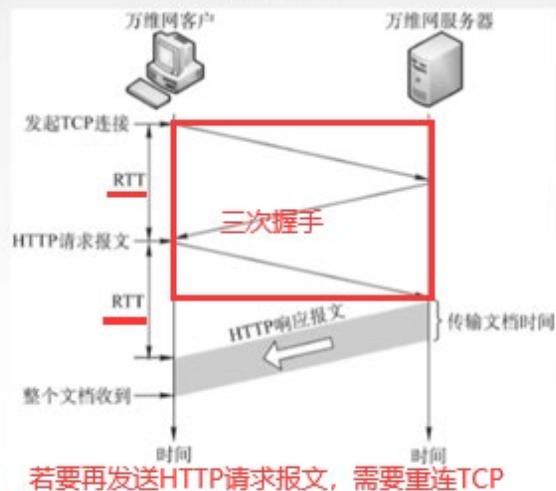
- ② HTTP采用TCP作为运输层协议，但**HTTP协议本身是无连接的**（通信双方在交换**HTTP报文**之前不需要先建立**HTTP连接**）。



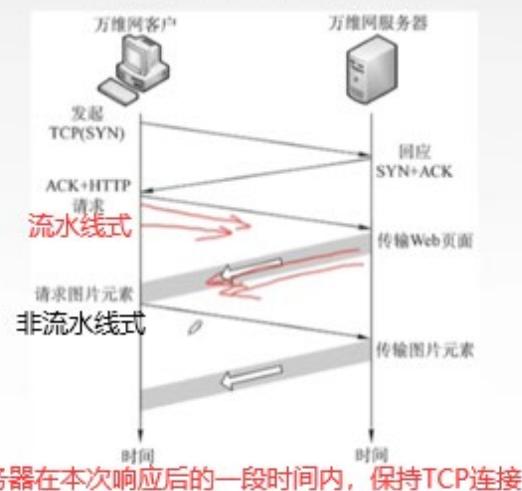
识别用户：比如淘宝，每次浏览选购后将商品放入购物车，那么就希望每次都是放入同一个购物车；或者进行用户需求分析，推销商品。 → **Cookie**

HTTP协议的连接方式

非持久连接



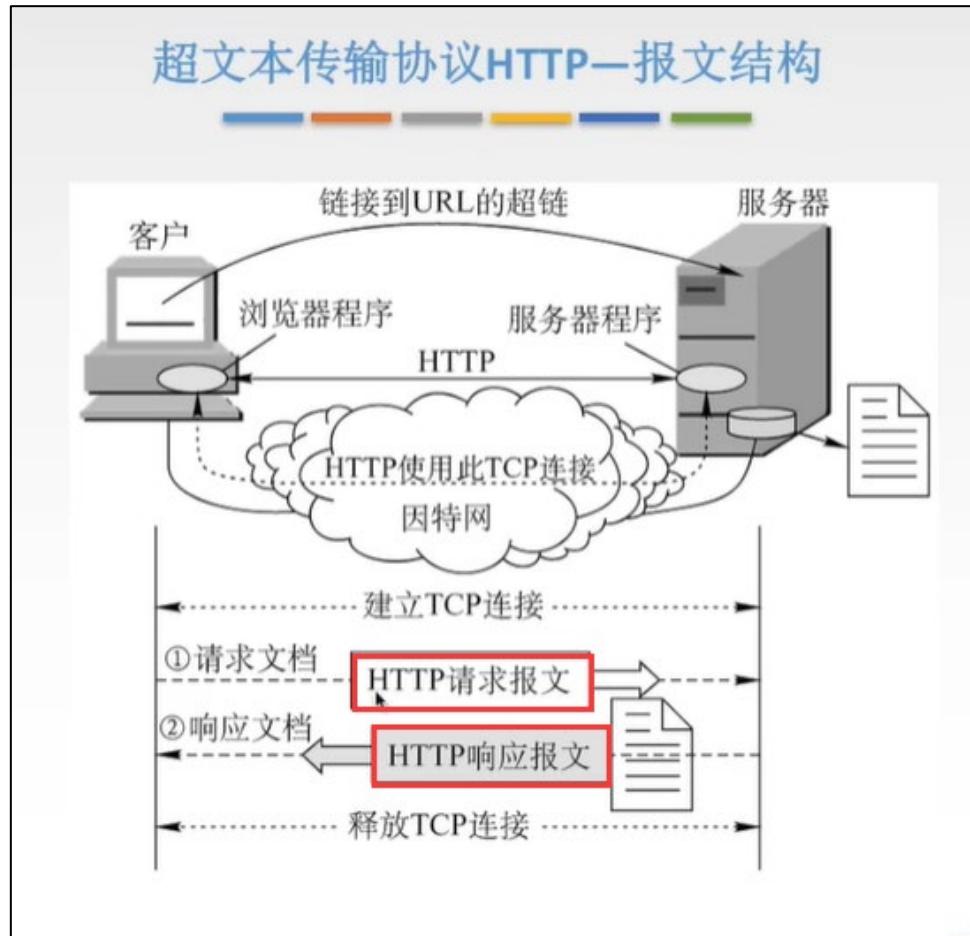
持久连接



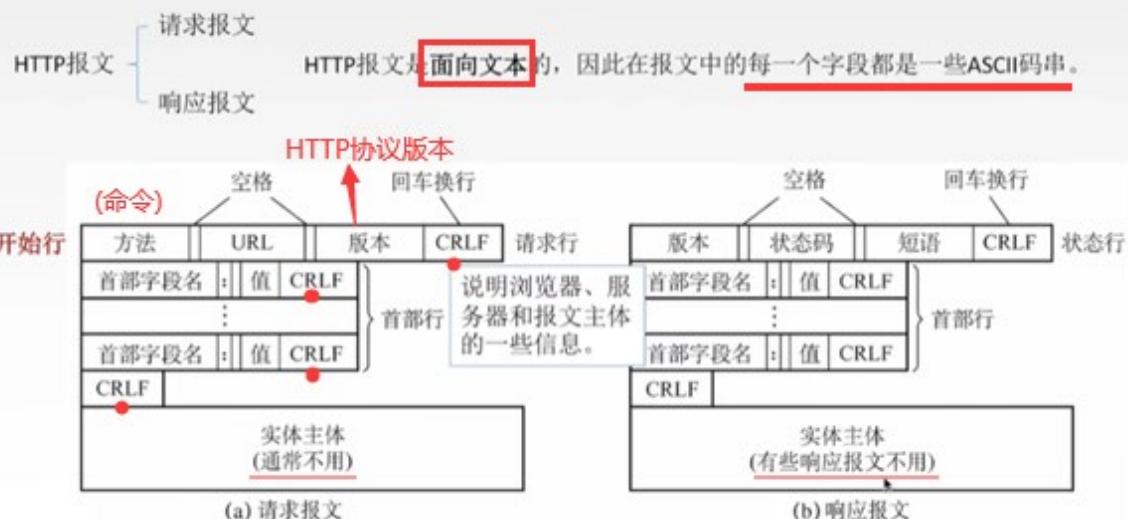
非流水线式：发了一个请求后，会收到一个响应；收到该响应后，才能发送下一个请求……

流水线式：减少 TCP 空闲时间，提高文档下载效率。

超文本传输协议HTTP—报文结构



超文本传输协议HTTP—报文结构

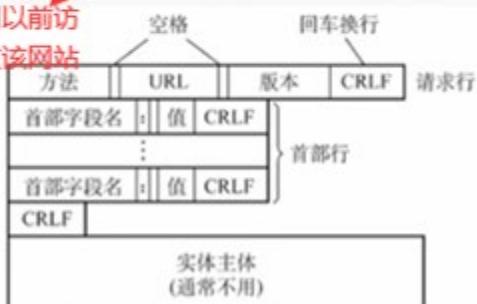


超文本传输协议HTTP—报文结构

某浏览器发出的请求报文

```
GET /index.html HTTP/1.1
Host: www.test.edu.cn 域名
Connection: Close 非持续连接
Cookie: 123456用户识别码
```

说明以前访
问过该网站



状态码：3个数字构成。5种类型，33种状态码

1xx表示通知信息的，如请求收到了或正在处理。

2xx表示成功，如接受或知道了。202 Accepted

3xx表示重定向，如要完成请求还必须采取进一步的行动。301 Moved Permanently

4xx表示客户的差错，如请求中有错误的语法或不能完成。404 Not Found

5xx表示服务器的差错，如服务器失效无法完成请求。

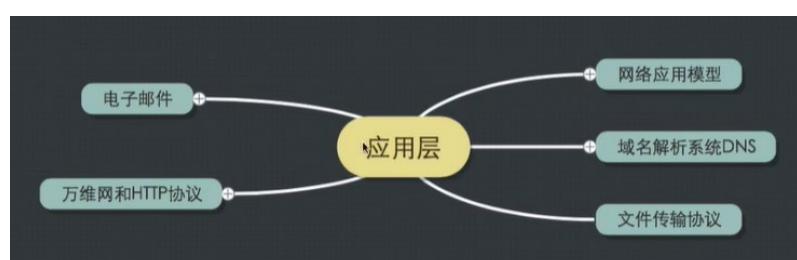


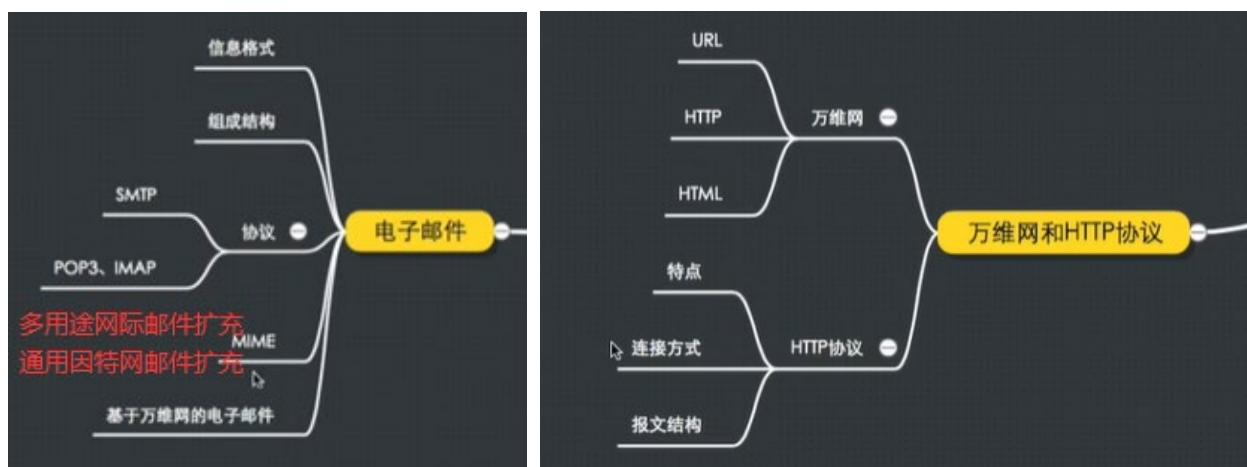
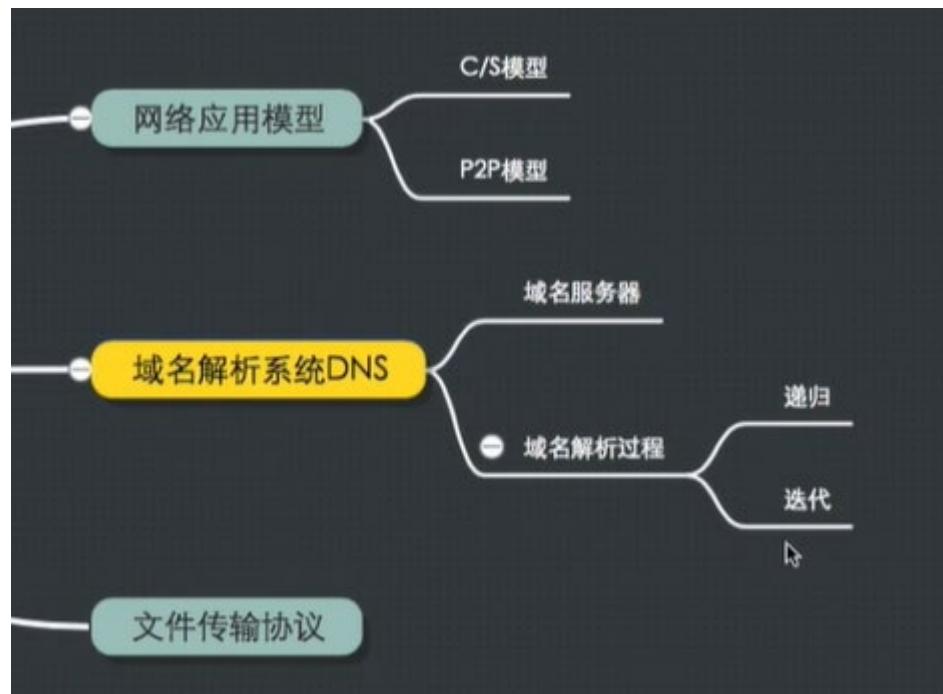
表示该网页转移到了新的地址。首部行中
会有relocation(首部字段名): URL(值)

301 Moved Permanently

如：平时访问网页
失败的404错误

6.6 第六章总结





使用TCP协议的应用层协议有：

- HTTP (超文本传输协议)
- FTP (文件传输协议)
- POP3 (邮局协议版本3)
- SMTP (简单邮件传输协议)
- Telnet (远程登录协议)
- SSH (安全壳协议)
- DNS (域名系统)

使用UDP协议的应用层协议有：

- DNS (域名系统)
- DHCP (动态主机配置协议)
- TFTP (简单文件传输协议)
- SNMP (简单网络管理协议)
- RIP (路由信息协议)
- NFS (网络文件系统)