计算机网络

计算机网络的概念

计算机网络：是一个分散的、具有独立功能的计算机系统，通过通信设备与线路连接起来，由功能完善的软件实现资源共享和信息传递的系统。

计算机网络是互连的、自治的计算机集合。

（互连：互联互通，通信链路 自治：无主从关系）

计算机网络的功能

- 数据通信（连通性）
- 资源共享（硬件、软件、数据）
- 分布式处理（多台计算机各自承担同一工作任务的不同部分）
- 提高可靠性（替代机）
- 负载均衡（各计算机之间更亲密）

计算机网络的组成

1. 组成部分：硬件、软件、协议
2. 工作方式：
    - 边缘部分，用户直接使用：C/S 方式、P2P方式
    - 核心部分：为边缘部分服务
3. 功能组成：
    - 通信子网：实现数据通信（各种传输介质、通信设备、相应的网络协议组成）
    - 资源子网：实现资源共享/数据处理（实现资源共享功能的设备和软件集合）

计算机网络的分类

1. 按分布范围分：广域网WAN（交换技术）、城域网MAN、局域网WAN（广播技术）、个人区域网PAN
2. 按使用者分：
    - 公用网：中国电信等
    - 专用网：中国银行、军队网等
3. 按交换技术分：电路交换、报文交换、分组交换
4. 按拓扑结构分：总线型、星型、环型、网状型（常用于广域网）
5. 按传输技术分
    - 广播式网络：共享公共通信信道
    - 点对点网络：使用分组存储转发和路由选择机制

标准化工作

标准的分类：

- 法定标准：由权威机构制定的正式的、合法的标准（OSI）
- 事实标准：某些公司的产品竞争中占据了主流，时间长了，这些产品中的协议和技术就成了标准（TCP/IP）

RFC（Request For Comments）—— 因特网标准的形式

RFC要上升为因特网正式标准的四个阶段：

- 因特网草案（Internet Draft）这个阶段还不是RFC文档
- 建议标准（Proposed Standard）从这个阶段开始成为RFC文档
- 草案标准（Draft Standard）
- 因特网标准（Internet Standard）

标准化工作的相关组织

- 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

性能指标

速率即数据率或称数据传输率或比特率

连接在计算机网络上的主机在数字信道上传送数据位数的速率

单位是`b/s, kb/s, Mb/s, Gb/s, Tb/s`

1Byte（字节）=8Bit （比特）

1KB=1024B=1024*8b

1MB=1024KB

1GB=1024MB

1TB=1024GB

速率

- “带宽”原本指某个信号具有的频带宽度，即最高频率与最低频率之差，单位是赫兹（Hz）
- 计算机网络中，带宽用来表示网络的通信线路传输数据的能力，通常是指单位之间内从网络中的某一点到另一点所能通过的“最高数据率”。单位是“比特每秒”，b/s，kb/s，Mb/s，Gb/s

吞吐量

表示在单位时间内通过某个网络（或信道、接口）的数据量。单位b/s，kb/s，Mb/s等

吞吐量受网络的带宽或网络的额定速率的限制

时延

指数据（报文/分组/比特流）从网络（或链路）的一端传送到另一端所需的时间。也叫延迟或迟延。单位是s

时延：

- 发送时延（传输时延）：从发送分组的第一个比特算起，到该分组的最后一个比特发送完毕所需的时间；$发送时延 = \frac{数据长度}{信道带宽（发送速率）}$
- 传播时延：取决于电磁波传播速度和链路长度 $传播时延 = \frac{信道长度}{电磁波在信道上的传播速率}$
- 排队时延：等待输出/入链路可用
- 处理时延：检错，找出口

时延带宽积

时延带宽积=传播时延*带宽；又称为**以比特为单位的链路长度**

往返时延RTT

从发送方发送数据开始，到发送方收到接收方的确认（接收方到数据后立即发送确认），总共经历的时延

RTT包括：往返传播时延=传播时延*2

末端处理时间

利用率

- 信道利用率：$信道利用率 = \frac{有数据通过时间}{（有+无）数据通过时间}$
- 网络利用率：信道利用率加权平均值

分层结构

分层的基本原则

- 各层之间相互独立，每层只实现一种相互独立的功能
- 每层之间界面自然清晰，易于理解，相互交流尽可能少
- 结构上可分隔开。每层都采用最合适的技术来实现
- 保持下层对上层的独立性，上层单向使用下层提供的服务
- 整个分层结构应该能促进标准化工作

分层结构

1. 实体：第n层中的活动元素称为**n层实体**。同一层的实体叫**对等实体**。
2. 协议：为进行网络中的**对等实体**数据交换而简历的规则、标准或约定称为网络协议。【水平】
    - 语法：规定数据传输的格式
    - 语义：规定所要完成的功能
    - 同步：规定各种操作的顺序
3. 接口（访问服务点SAP）：上层使用下层服务的入口
4. 服务：下层为相邻上层提供的功能调用。【垂直】

SDU服务数据单元：为完成用户所需求的功能而应传送的数据

PCI协议控制信息：控制协议操作的信息

PDU协议数据单元：对等层次之间传送的数据单位

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**以完成本层的功能
- **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

ISO/OSI 参考模型

目的：支持**异构网络系统**的互联互通

应用层：用户与网络的界面。所有能和用户交互产生网络流量的程序

典型应用层服务：

- 文件传输（FTP）
- 电子邮件（SMTP）
- 万维网（HTTP）

表示层：用于处理在两个通信系统中交换信息的表示方式（语法和语义）

- 功能一：数据格式的变换
- 功能二：数据加密解密
- 功能三：数据压缩和恢复

会话层：向表示层实体/用户进程提供**建立连接**并在连接上**有序**地**传输**数据。这是绘画，也是**建立同步（SYN）**

- 功能一：建立、管理、终止会话
- 功能二：使用校验点可使会话在通信失效时从**校验点/同步点**继续恢复通信，实现数据同步（适用于传输大文件）

传输层：负责主机中**两个进程**的通信，即**端到端**的通信。传输单位是报文段或用户数据报

- 功能一：可靠传输、不可靠传输
- 功能二：差错控制
- 功能三：流量控制
- 功能四：复用分用
    - 复用：多个应用层进程可同时使用下面运输层的服务
    - 分用：运输层把收到的信息分别交付给上面应用层中相应的进程

网络层：主要任务是把**分组**从源端传到目的端，为分组交换网上的不同主机提供通信服务。网络层传输单位是**数据报。**

- 功能一：路由选择（最佳路径）
- 功能二：流量控制
- 功能三：差错控制
- 功能四：拥塞控制（若所有节点都来不及接受分组，而要丢弃大量分组的话，网络就处于拥塞状态。因此要采取一定措施，缓解这种拥塞）

数据链路层：主要任务是把网络传下来的数据报**组装成帧。**数据链路层/链路层的传输单位是**帧**

- 功能一：成帧（定义帧的开始和结束）
- 功能二：差错控制（帧错+位错）
- 功能三：流量控制
- 功能四：访问（接入）控制（控制对信道的访问）

物理层：主要任务是在**物理媒体**上实现比特流的**透明传输。**物理层传输单位是**比特**

透明传输：指不管所传数据是什么样的比特组合，都应当能够在链路上传送

- 功能一：定义接口特性
- 功能二：定义传输模式（单工、半双工、双工）
- 功能三：定义传输速率
- 功能四：比特同步

- 功能五：比特编码

OSI参考模型与TCP/IP参考模型

1. 都分层
2. 基于独立的协议栈的概念
3. 可以实现异构网路互联

TCP/IP和ISO/OSI的区别

- OSI定义三点：服务、协议、接口
- OSI先出现，参考模型先于协议发明，不偏向特定协议
- TCP/IP设计之初就考虑到异构网络**互联**问题，将IP作为重要层次

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

物理层

物理层解决如何在连接各种计算机的传输媒体上**传输数据比特流**，而**不是指具体的传输媒体**

物理层主要任务：确定与传输媒体**接口**有关的一些特性 ——> 定义标准

- 机械特性：定义物理连接的特性，规定物理连接时采用的规格、接口形状、**引线数目、引脚数量**和排列情况
- 电气特性：规定传输二进制位时，线路上信号的**电压范围**、阻抗匹配、传输**速率**和**距离**限制等
- 功能特性：指明线路上出现的某**一电平表示何种意义**，接口部件的信号线的用途
- 规程特性（过程特性）：定义各条物理线路的工作**规程和时序**关系

数据通信

通信的目的是传送消息。

数据：传送信息的实体，通常是有意义的符号序列

信号：数据的电气/电磁的表现，是数据在传输过程中的**存在形式**

- 数字信号：代表信息的参数取值是离散的
- 模拟信号：代表信息的参数取值是连续的

信源：产生和发送数据的源头

信宿：接收数据的终点

信道：信号的传输媒介。一般用来表示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。

- 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

通信方式

从通信双方信息的交互方式看，可以有三种基本方式：

1. 单工通信：只有一个方向的通信而没有反方向的交互，仅需要**一条**信道
2. 半双工通信：通信双方都可以发送或接收信息，但任何一方都不能同时发送和接收，需要**两条**信道
3. 全双工通信：通信双方可以同时发送和接收信息，也需要**两条**信道

数据传输方式

- 串行传输：速度慢，费用低，适合远距离
- 并行传输：速度快，费用高，适合近距离

码元、速率、波特、带宽

码元是指用一个**固定时长**的**信号波形**（数字脉冲），代表不同离散数值的基本波形，是数字通道中数字信号的计量单位，这个时长内的信号称为**k进制码元**，而该时长称为码元宽度。当码元的离散状态有M个时（M大于2），此时码元为M进制码元。

1码元可以携带多个比特的信息量。例如，在使用二进制编码时，只有两种不同的码元，一种代表0状态，另一种代表1状态

速率也叫数据率，是指数据的**传输速率**，表示单位时间内传输的数据量。可以用**码元传输速率**和**信息传输速率**表示

1. 码元传输速率：别名码元速率、波形速率、调制速率、符号速率等。它表示单位时间内数字通信系统所传输的码元个数（也可称为**脉冲个数或信号变化次数**），单位是**波特（Baud）**。1波特表示数字通信系统每秒传输一个码元。这里的码元可以是多进制的，也可以是二进制的，但码元速率与进制无关。

   $1Baud = 1\text{码元}/s$

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)

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*n bit/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。

系统传输的是**比特流**，通常比较的是信息传输速率。

奈氏准则、香农定理

失真：有失真但可识别；失真大无法始别

影响失真程度的因素：

- 码元传输速率：越快，失真越大
- 信号传输距离：距离越远，衰减越久，失真越大

- 噪声干扰：越多，失真越大
- 传输媒体质量：越差，失真越大

失真的一种现象——**码间串扰**

信道带宽是信道能通过的**最高频率和最低频率之差**

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

奈氏准则（奈奎斯特定理）

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 Baud，W是信道带宽，单位是Hz。

理想低通信道下的极限数据传输率=$2Wlog_2V(b/s)$（W表示带宽(Hz)；V表示，几种码元/码元的离散电平数目）

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

香农定理

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：
$信噪比(dB)=10\log_{10}(S/N)$

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

$信道的极限传输速率=Wlog_2(1+S/N)（b/s）$ (W带宽(Hz)；S/N是信噪比)

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. 对一定的传输宽度和一定的信噪比，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，师姐信道能达到的传输速率要比它低不少

编码&调制

信道上传送的信号：

- 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）

数字数据：

- 数字发送器 ——> 数字信号 编码
- 调制器 ——> 模拟信号 调制

模拟数据：

- PCM编码器 ——> 数字信号 编码
- 放大器调制器 ——> 模拟信号 调制

数字数据编码为数字信号

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。**
2. **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方式
3. **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. **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，但它所占有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。**
5. **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**于曼彻斯特边编码。
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；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。

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过程和解调过程。

2ASK 调幅；2FSK 调频；2PSK 调相；QAM 调幅+调相

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**）。

在计算机应用中，能够达到**最高保真水平**的就是PCM编码。它主要包含三步：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：$f_{\text{采样频率}} \geq 2f_{\text{信号最高频率}}$
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：把量化的结果转化为与之对应的二进制编码

物理层传输介质

传输介质也称传输媒质/传输媒介，它就是数据传输系统中在发送设备和接收设备之间的**物理通路**。

传输媒体并不是物理层。传输媒体在物理层的下面，因为物理层是体系结构的第一层，因此有时称传输媒体为0层。在传输媒体中传输的信号，但传输媒体并不知道传输的信号代表什么意思。但物理层规定了**电气特性**，因此能够始别所传送的比特流。

传输介质：

1. 导向性传输介质：电磁波被导向沿着固体媒介（铜线/光纤）传播
    - 双绞线：古老、又最常用的传输介质，它由**两根**采用一定规则并排**绞和**的、相互绝缘的**铜导线**组成。（绞和可以减少对相邻导线的电磁干扰）
        - 为了进一步提高抗电磁干扰能力，可在双绞线的外面再加上一个由**金属丝**编织成的屏蔽层，这就是**屏蔽双绞线（STP）**，无屏蔽层的双绞线就称为**非屏蔽双绞线（UTP）**
        - 双绞线价格便宜，是最常用的传输介质之一，在局域网和传统电话网中普遍使用。模拟传输和数字传输都可以使用双绞线，其通信距离一般为几公里到数十公里。距离太远时，对于**模拟传输**，要使用**放大器**放大衰减信号；对于**数字传输**，要用**中继器**将失真的信号整形。
    - 同轴电缆：由**导体铜质芯线、绝缘层、网状编制屏蔽层和塑料外层**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。（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，但价格更贵）
    - 光纤：光纤通信系统的**宽带远远大于**目前其他各种传输媒体的宽带。光纤主要由**纤芯（实体的）和包层**构成。
        - 单模光纤：一种在横向模式直接传输光信号的光纤；光源是定向性很好的激光二极管；衰耗小，适合远距离传输
        - 多模光纤：有多种传输光信号模式的光纤；光源为普通的发光二极管；易失真，适合近距离传输

      特点：传输损耗小，中继距离长，对远距离传输特别经济；抗雷电和电磁干扰性能好；无串音干扰，保密性好，也不易被窃听或截取数据；体积小，重量轻
2. 非导向性传输介质：自由空间，介质可以是空气、真空、海水等
    - 无线电波：信号向所有方向传播；较强**穿透能力**，可以传远距离，广泛用于通信领域（如手机通信）
    - 微波：信号固定方向传播；微波通信率较高、频段范围宽，因此数据率很高
        - 地面微波接力通信
        - 卫星通信
    - 红外线、激光：信号固定方向传播：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

物理层设备

中继器

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：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。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于网段的数据。两端可连相同媒体，也可连不同媒体。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）

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。（不超过5个网段，最多有4个物理层网络设备，只有3个段可以连接计算机）

集线器（多口中继器）

集线器功能：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，以增强信号传输的距离，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。

集线器不能分割冲突域，连在集线器上的工作主机平分带宽。

数据链路层

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：网络中两个节点之间的**逻辑链路**，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

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。

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

功能五：差错控制（帧错/位错）

封装成帧就是在一段数据的前后部分添加首部和尾部，这样就构成了一个帧。接收端在收到物理层上交的比特流后，就能根据首部和尾部的标记，从收到的比特流中识别帧的开始和结束。

首部和尾部包含许多的控制信息，他们的一个重要作用：**帧定界**（确定帧的界限）

帧同步：接收方应当能从接收到的二进制比特流中区分出帧的起始和终止。

帧的数据部分 $\leq$ 最大传送单元MTU

组帧的四种方法：1. 字符计数法 2. 字符（节）填充法 3. 零比特填充法 4. 违规编码法

透明传输是指不管所传数据是什么样的比特组合，都应当能够在链路上传送。因此，链路层就"看不见"有什么妨碍数据传输的东西

当所传数据中的比特组合恰巧与某一个控制信息完全一样时，就必须采取适当的措施，使收方不会将这样的数据误认为是某种控制信息。这样才能保证数据链路层的传输是透明的。

1. 字符计数法：帧首部使用一个计数字段（第一个**字节**，八位）来标明帧内字符数
2. 字符填充法：添加转义字符ESC
3. 零比特填充法：操作：
    - 在发送端，扫描整个信息字段，只要连续5个1，就立即填入1个0
    - 在接收端收到一个帧时，先找到标志字段确定边界，再用硬件对比特流进行扫描。发现连续5个1时，就把后面的0删除

    保证了透明传输：在传送的比特流中可以传送任意比特流组合，而不会引起对帧边界的判断错误
4. 违规编码法：曼彻斯特编码；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（其值若有错误将导致灾难性后果）及字符填充实现上的复杂性和不兼容性，目前比较普遍使用的帧同步法时**比特填充**和**违规编码法**。

差错控制

概括来说，传输中的差错都是由于噪声引起的。

全局性：

- 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。（解决方法：提高信噪比来减少或避免干扰，对传感器下手）

局部性：

- 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（解决方法：通常利用编码技术来解决）

差错：

- 位错：比特位出错，1变成0，0变成1
- 帧错：丢失，重复，失序

检错编码

冗余编码：在数据发送之前，先按某种关系**附加**上一定的**冗余位**，构成一个符合某一规则的码字后再发送。当要发送的有效数据变化时，相应的冗余位也随之变化，使码字遵从不变的规则。接收端根据收到码字是否仍符合原规则，从而判断是否出错。

1. 奇偶校验码（n-1位信息元，1位校验元）
    - 奇校验码："1"的个数为奇数
    - 偶校验码："1"的个数为偶数

    奇偶校验码特点：只能检查出**奇数个比特**错误，检错能力50%
2. CRC循环冗余码：最终发送的数据 = 要发送的数据 + 帧检验序列FCS

计算冗余码：

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

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。

- 余数为0，判定这个帧没有差错，接受
- 余数不为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。

海明码

发现双比特错，**纠正单**比特错。

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
也就是D=101101的海明码应该有6+4=10位，
其中原数据6位，校验码4位
假设这4位校验码分别为P1,P2,P3,P4; 数据从左到右位D1,D2...D6
                    放在2的几次方的位置              按序把空填满

二进制    0001     0010     0011     0100     0101     0110     0111     1000     1001
 1010
数据位      1        2        3        4        5        6        7        8        9
10
代码       P1       P2       D1       P3       D2       D3       D4       P4       D5
D6
实际值      0        0        1        0        0        1        1        1        0
1

-> 令所有要校验的位异或=0
P1^D1^D2^D4^D5=0 -> P1=0
P2^D1^D3^D4^D6=0 -> p2=0
P3^D2^D3^D4=0 -> P3=0
P4^D5^D6 -> P4=1

故101101的海明码位0010011101

检错并纠错
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
P1^D1^D2^D4^D5=0 -> P1=1
P2^D1^D3^D4^D6=0 -> p2=0
P3^D2^D3^D4=0 -> P3=1
P4^D5^D6 -> P4=0
从p4到p1倒过来写，(0101)b=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
```

流量控制与可靠传输机制

数据链路层的流量控制

较高的发送速度和**较低的接收能力**的不匹配，会造成传输错误，因此流量控制也是数据链路层的一项重要工作。

数据链路层的流量控制是点对点的，而传输层的流量控制是端到端的。

数据链路层流量控制手段：接收方收不下就不回复确认。

传输层流量控制手段：接收端给发送端一个窗口公告

停止等待协议：没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

（发送窗口大小=1，接收窗口大小=1）

滑动窗口协议：

- 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

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. 停等协议有几种应用情况?

   无差错情况&有差错情况

   - 数据帧丢失或检测到帧出错（超时计时器：每次发送一个帧就启动一个计时器；超时计时器设置的重传时间应当比帧传输的平均RTT更长一些）（发完一个帧后，必须保留它的副本；数据帧和确认帧必须编号）
   - ACK丢失
   - ACK迟到

3. 停等协议性能分析：信道录用率太低$信道利用率U = \frac{T_D}{T_D+RTT+T_A}$

   信道利用率：发送方在一个发送周期内，有效地发送数据所需要的时间占整个发送周期的比率。

   $信道利用率 = (L/C)/T$（T：发送周期，从开始发送数据，到收到第一个确认帧为止；L：T内发送L比特数据；C：发送方数据传输率）

   信道吞吐率 = 信道利用率 * 发送方的发送速率

流水线技术：（停等的改进）

- 必须增加序号范围
- 发送方需要缓存多个分组

后退N帧协议（GBN）

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
- 接收窗口：接收方维持一组连续的允许接受帧的序号

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。上层等会儿再发送。（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**累计确认**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：出现丢失和时延过长帧时发送方的行为。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。

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，并为最近按序接受的帧重新发送ACK。接受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）。

滑动窗口长度

若采用n个比特对帧编号，那么发送窗口的尺寸W_T应满足：$1 \leq W_T \leq 2^n - 1$。因为发送窗口尺寸过大，就会使得接收方无法区别新帧和旧帧。接收窗口大小为1。

GBN协议的弊端：累计重传 -> 批量重传

解决方法：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

选择重传协议（SR）

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。
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**

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【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。

滑动窗口长度

发送窗口最好等于接收窗口。（大了会溢出，小了没意义）$W_{Tmax} = W_{Rmax} = 2^{(n-1)}$

信道划分介质访问控制

- 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）

介质访问控制的内容就是，采取一定的措施，使得两个节点之间的通信不会发生互相干扰的情况。

- 静态划分信道：信道划分介质访问控制：频分FDM；时分TDM；波分WDM；码分CDM多路复用
- 动态分配信道：轮询访问介质访问控制：令牌传递协议；随机访问介质访问控制：ALOHA协议；CSMA协议；CSMA/CD协议；CSMA/CA协议

信道划分介质访问控制：将使用介质的每个设备与来自同一信道上的其他设备的**通信隔离开**，把**时域和频域资源**合理地分配给网络上的设备

多路复用技术：把多个信号组合在一条物理信道上进行传输，使得多个计算机或终端设备**共享信道资源**，提高信道利用率。把一条广播信道，逻辑上分成几条用于两个节点之间通信的互不干扰的子信道，**实际就是把广播信道转变为点对点信道。**

1. 频分复用：用户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。**充分利用传输介质带宽，系统效率较高；由于技术比较成熟，实现比较容易。
2. 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。（TDM帧是在物理层传送的比特流所划分的帧，标志一个周期）
3. 统计时分复用：每一个STDM帧中的时隙数**小于**连接在集中器上的用户数。各用户有了数据就随时发往集中器的**输入缓存**，然后集中器按顺序依次扫描输入缓存，把缓存中的输入数据放入STDM帧中，一个STDM帧满了就发出。**STDM帧不是固定分配时隙，而是按需动态分配时隙。**
4. 波分多路复用：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。
5. 码分多路复用：**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/芯片（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。

   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

ALOHA协议

纯ALOHA协议思想：不监听信道，不按时间槽发送，随机重发

冲突如何解决：超时后等一随机时间再重传

时隙ALOHA协议思想：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。

纯ALOHA比时隙ALOHA吞吐量低，效率更低。

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。

监听结果

- 信道空闲：发送完整帧

- 信道忙：推迟发送

1-坚持CSMA：坚持指的是对于监听信道忙之后的坚持。

思想：如果一个主机要发送信息，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，直到空闲马上传输。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。

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。

思想：如果一个主机要发送信息，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：指的是对监听信道空闲的处理。

思想：如果一个主机要发送信息，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。

缺点：发生冲突之后还是要坚持把数据帧发送完，造成了浪费。

	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送
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

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。**半双工网络**

传播时延对载波监听的影响：

单程端到端传播时延：τ

最多是两倍的总线传播时延（2τ）。总线的端到端往返传播时延。争用期/冲突窗口/碰撞窗口。只要经过2τ时间还没有检测到碰撞，就肯定这次发送不会发生碰撞。

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[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, 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. 当重传达到**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。

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，因为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：帧的传输时延至少要两倍于信号在总线中的传播时延。

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

$$\text{最小帧长} = \text{总线传播时延} \times \text{数据传输速率} \times 2 = 2\tau \times \text{数据传输速率}$$

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。

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）

工作原理：发送数据之前，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）**。

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**二进制制数规避算法**来确定随机的推迟时间）。

CSMA/CD 与 CSMA/CA

相同点：核心都是先听再说

不同点：

- 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】
- 载波检测方式不同
- 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**

轮询访问介质访问控制

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

随机访问MAC协议：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

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

轮询协议：主结点轮流"邀请"从属结点发送数据。

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。

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。

局域网基本概念和体系结构

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。

决定局域网的主要因素为：网络拓扑，传输介质与介质访问控制方法。

局域网网络拓扑结构：

- 星型拓扑：中心节点是控制中心，任意两个节点间的通信最多只需要**两步**，传输速度快，并且网络构型简单、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。
- 总线型拓扑：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。
- 环形拓扑：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。
- 树形拓扑：易于拓展，易于隔离故障，也容易有**单点故障**。

局域网传输介质：

- 有线局域网：双绞线、同轴电缆、光纤
- 无线局域网：电磁波

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：

- 以太网：以太网是以用最广泛的局域网。

- 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”。
- FDDI网（传输介质是光纤）：物理上采用了双环状拓扑结构，逻辑上是环形拓扑结构。
- ATM网：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。
- 无线局域网（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）。

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。

以太网

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**技术。

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。

以太网只实现无差错接收，不实现可靠传输。

以太网拓扑：逻辑上总线型，物理上星型。

10BASE-T以太网

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：计算机与外界有局域网的连接是通过**通信适配器**的。

以太网MAC帧：最常用的MAC帧是以太网V2的格式。

高速以太网

- 100BASE-T以太网：在双绞线上传送100Mb/s基带信号的星型拓扑以太网，仍使用IEEE802.3的CSMA/CD协议。支持全双工和半双工，可在全双工方式下工作而无冲突。
- 吉比特以太网：在光纤或双绞线上传送1Gb/s信号。
- 10吉比特：在光纤上传送10Gb/s信号。

IEEE 802.11 无线局域网

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：自组织网络

PPP协议 & HDLC协议

广域网（WAN）：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，或跨越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、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。

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是用PPP协议。只支持全双工链路。

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
- 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。
- 多种类型链路：串行/并行，同步/异步，电/光....
- 差错检测：错就丢弃
- 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
- 网络层地址协商：知道通信双方的网络层地址
- 数据压缩协商

PPP协议无需满足的要求：

- 纠错
- 流量控制
- 序号
- 不支持多点线路

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S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。

HDLC 协议

高级数据链路控制，是一个在同步网上传输数据、**面向比特**的数据链路层协议。数据报文可透明传输，用于实现透明传输的“0比特插入法”易于硬件实现。**采用全双工通信**。所有帧采用**CRC检验**，对信息帧进行顺序编号，可防止漏收或重份，传输可靠性高。

主站、从站、复合站，三种数据操作方式：

- 正常响应方式
- 异步平衡方式
- 异步响应方式

HDLC帧的类型

- 信息帧（I）
- 监督帧（S）
- 无编号帧（U）

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

链路层设备

链路层扩展以太网：网桥&交换机

网桥根据MAC帧的目的地址对帧进行转发和过滤。当网桥收到一个帧时，并不向所有接口转发此帧，而不是先检查此帧的目的MAC地址，然后再确定将该帧转发到哪一个接口，或者是把它丢弃（即过滤）。

网段：一般指一个计算机网络中使用同一物理设备能够直接通讯的那一部分。

网桥有点：

- 过滤通信量，增大吞吐量
- 扩大了物理范围
- 提高了可靠性
- 可互连不同物理层、不同MAC子层和不同速率的以太网

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。

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）。

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。
- 存储转发式交换机：将帧放入高速缓存，并检查是否正确，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。**

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。简单的说如果站点出发一个广播信号，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（中继器、集线器）	×	×
链路层设备（网桥、交换机）	√	×
网络层设备（路由器）	√	√

网络层

网络层功能

主要任务是把**分组**从源端传到目的端，为分组交换网上的不同主机提供通信服务。网络层传输单位是**数据报**。

功能一：路由选择与分组转发（最佳路径）

功能二：异构网络互联

功能三：拥塞控制：若所有结点都来不及接受分组，而要丢弃大量分组的话，网络就处于**拥塞**状态。因此要采取一定措施，缓解这种拥塞。

数据交换方式

电路交换：eg.电话网络

电路交换的阶段：

- 建立连接（呼叫/电路建立）
- 通信
- 释放连接（拆除电路）

优点：通信时延小；有序传输；没有冲突；实时性强

缺点：建立连接时间长；线路独占，使用效率低；灵活性差；无差错控制能力

报文交换

报文：源应用发送的信息整体

优点：无需建立连接；存储转发，动态分配线路；线路可靠性较高；线路利用率较高；多目标服务

缺点：有存储转发时延；报文大小不定，需要网络节点有较大缓存空间

分组交换

分组：把大的数据块分割成小的数据块

优点：无需建立连接；存储转发，动态分配线路；线路可靠性较高；线路利用率较高；相对于报文交换，存储管理更容易

缺点：有存储转发时延；需要传输额外的信息量；乱序到目的主机时，要对分组排序重组

三种数据交换方式比较：

- 报文交换和存储交换都采用存储转发
- 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
- 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

数据报方式&虚电路方式

数据报方式为网络层提供**无连接服务**。虚电路方式为网络层提供**连接服务**。

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。

连接服务：首先为分组的传输确定传输路径（建立连接），然后沿该路径（连接）传输系列分组，系列分组传输路径相同，传输结束后拆除连接。

数据报

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，都要维持一张虚表，每一项记录了一个打开的虚电路信息。

通信过程：每个分组携带**虚电路号**，而非目的地址。

	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

IP数据报格式

- 版本：IPv4/IPv6
- 首部长度：单位是**4B**，最小为5
- 区分服务：指示期望获得哪些类型的服务
- 总长度：首部+数据，单位是**1B**
- 生存时间（TTL）：IP分组的保质期。经过一个路由器-1，变成0则丢弃。
- 协议：数据部分的协议（TCP：6；UDP：17）
- 首部检验和：只检验首部
- 源IP地址和目的地址：32位
- 可选字段：0~40B，用来支持排错、测量以及安全等措施。
- 填充：全0，把首部补成4B的整数倍。

IP数据报分片

最大传送单元MTU：链路层数据帧可封装数据的上限。以太网的MTU是1500字节。

IP数据报格式

- 标识：同一数据报的分片使用同一标识。
- 标志：只有2位有意义；
  - 中间位DF（Don't Fragment）：DF=1，禁止分片，DF=0，允许分片
  - 最低为MF（More Fragment）：MF=1，后面“还有分片”，MF=0，代表最后一片/没分片
- 片偏移：指出较长分组分片后，某片在原分组中的相对位置。以**8B**为单位（除了最后一个分片，每个分片长度一定是**8B的整数倍**）

总长度单位是1B，片偏移单位是8B，首部长度单位是4B

IPv4地址

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>}

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	2^7-1	1	126	$2^{24}-2$
B	$2^{14}-1$	128.1	191.255	$2^{16}-2$
C	$2^{21}-1$	192.0.1	223.255.255	2^8-2

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**。

WAN端	LAN端
172.38.1.5：40001	192.168.0.3：30000
172.38.1.5：40002	192.168.0.4：30001

子网划分和子网掩码

子网划分：某单位划分子网后，对外仍**表现为一个网络**，即本单位外的网络看不见本单位内子网的划分。子网号能否全0或全1要看情况，主机号不能全0全1

子网掩码与IP地址逐位相与，就得到子网网络地址。

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

路由表中：目的网络地址；目的网络子网掩码；下一跳地址

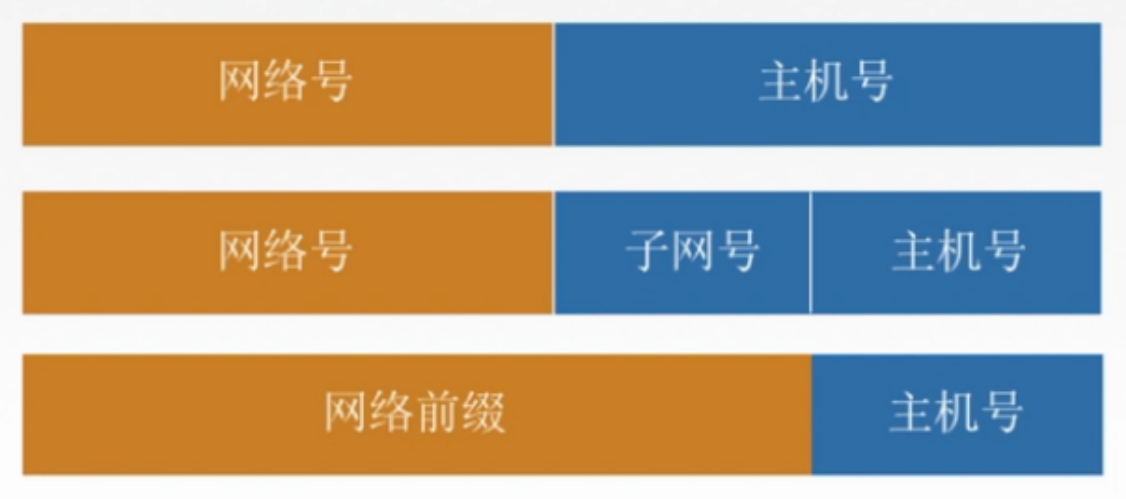
路由器转发分组的算法：

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

无分类编址CIDR（构成超网）

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。



   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

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 0010**0011 00000111（128.14.32.0）

   最小地址：**10000000 00001110 0010**0000 00000000（128.14.47.255）

   地址块：128.14.32.0/20（20位网络前缀+12位主机号=32位）

地址掩码（子网掩码）：11111111 11111111 11110000 00000000

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。

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。

ARP协议

ARP高速缓存：IP地址与MAC地址的映射

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. 路由器发送给**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

DHCP协议

动态主机配置协议DHCP是**应用层**协议，使用**客户/服务器**方式，客户端和服务端通过**广播**方式进行交互，基于**UDP**。

DHCP提供**即插即用**联网的机制，主机可以从服务器动态获取IP地址、子网掩码、默认网关、DNS服务器名称与IP地址，允许**地址重用**，支持**移动用户加入网络**，支持**在用地址续租**。

1. 主机广播**DHCP发现**报文（试图找到网络中的服务器，服务器获得一个IP地址）
2. DHCP服务器广播**DHCP提供**报文（服务器拟分配给主机一个IP地址及相关配置，先到先得）
3. 主机广播**DHCP请求**报文（主机向服务器请求提供IP地址）
4. DHCP服务器广播**DHCP确认**报文（正式将IP地址分配给主机）

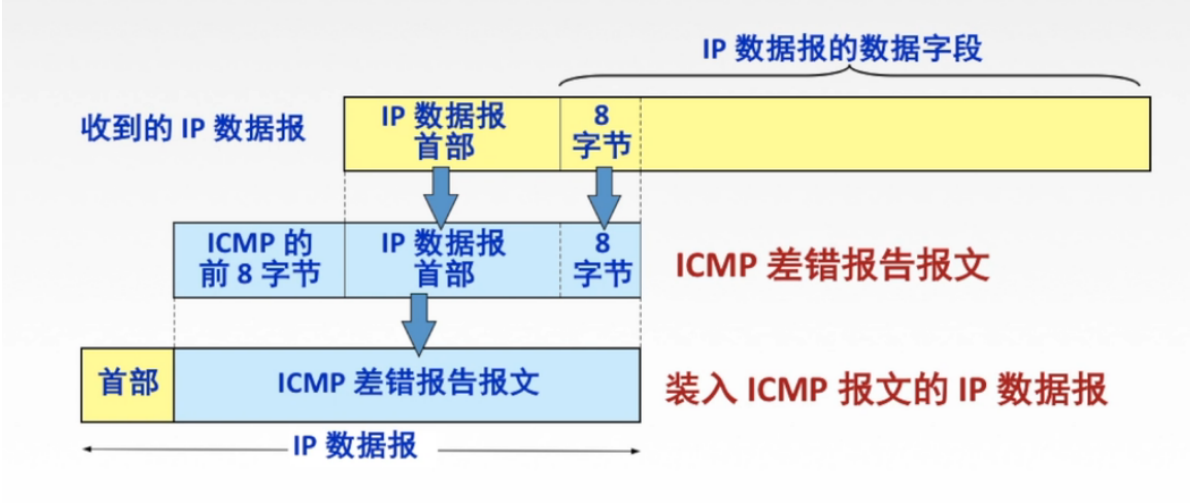
ICMP协议

ICMP协议支持主机或路由器：差错（或异常）报告，网络探寻 -> 发送特定ICMP报文

ICMP差错报告报文

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效率放慢。（拥塞丢数据）

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。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。
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。

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）
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
4. 路由器询问和通告报文

ICMP的应用

1. 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。**
2. 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。**

IPv6

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. ICMIPv6：附加报文类型"分组过大"。

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，改成写一个首部字段。
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

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:BA5F: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 FF05:0:0:0:0:0:0: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和IPv4过渡的策略

- 双栈协议：一台设备上**同时启用IPv4协议栈和IPv6协议栈。**如果是路由器，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。如果是计算机，它将同时拥有IPv4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：将其他协议的数据帧或包**重新封装**然后通过隧道发送。

路由算法及路由协议

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

- 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，使用大型网络，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）

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

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及确定分组在该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。

路由选择协议

- 内部网关协议IGP：一个AS内使用 RIP、OSPF
- 外部网关协议EGP：AS之间使用的 BGP

RIP协议及距离向量算法

RIP协议：一种分布式的基于**距离向量**的路由选择协议，是因特网的协议标准，最大优点是简单。

RIP协议要求网络中每一个路由器都维护**从它自己到其他每一个目的网络的唯一最佳距离记录**（即一组距离）。

距离：通常为“跳数”，即从源端口到目的端口所经过的路由器个数，经过一个路由器跳数+1。特别的，从一路由器到直接连接的网络距离为1，RIP允许一条路由最多只能包含15个路由器，因此距离为**16表示网络不可达。**（RIP协议只适用于小互联网）

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。**
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没有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**”

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：
   - R1路由表中若没有Net3，则把该项目填入R1路由表
   - R1路由表中若有Net3，则查看一下跳路由器地址：若下一跳是X，则用收到的项目替换源路由表中的项目；若下一跳不是X，原来距离比从X走的距离远则更新，否则不做处理。
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

RIP是**应用层协议**使用**UDP**传送数据。

一个RIP报文最多包括25个路由，如超过，必须再用一个RIP报文传送。

RIP的特点：当网络出现故障时，要经过比较长的时间（例如数分钟）才能将此信息传送到所有的路由器，“慢收敛”。

OSFP协议及链路状态算法

开放最短路径有限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：

1. 使用洪泛法向自治系统内**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。（广播）

   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。

2. 发送的信息就是与本路由器**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）。
3. 只有当**链路状态发生变化时**，路由器才向所有路由器洪泛发送此消息。

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。
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；如果有没有的或者是更新的，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直接用**IP数据报**传送。（网络层）

OSPF其他特点

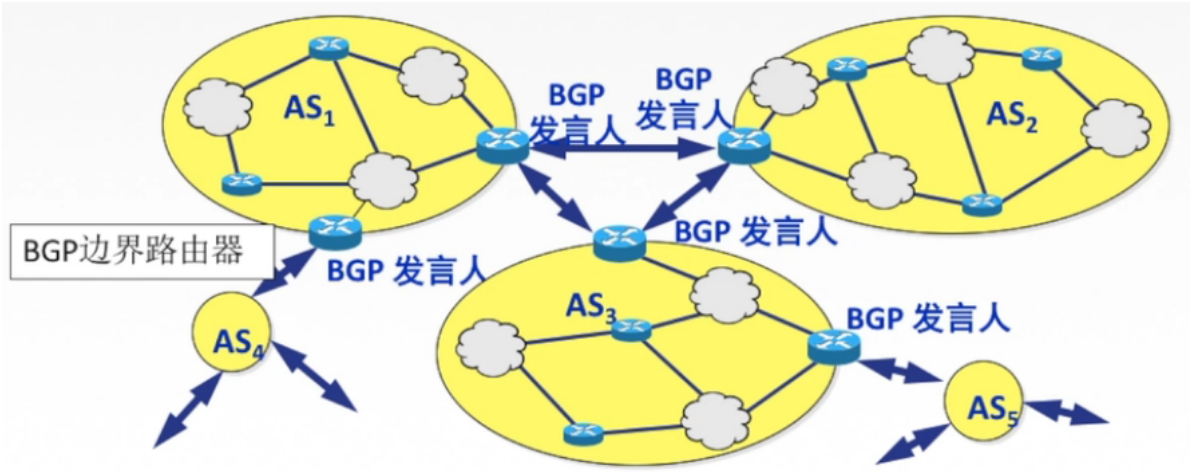
1. 每隔30min，要刷新一次数据库中的链路状态。
2. 由于一个路由器的链路状态只涉及到与相邻路由器的连通状态，因而与整个互联网的规模并无直接关系。因此当**互联网规模很大**时，OSPF协议要比距离向量协议RIP好得多。
3. OSPF不存在坏消息传的慢的问题，它的**收敛速度很快**。

BGP协议

与其他AS的邻站BGP发言人交换信息。

交换的网络可达性的信息，即要达到某个网络所要经过的一系列AS。

发生变化时更新有变化的部分。



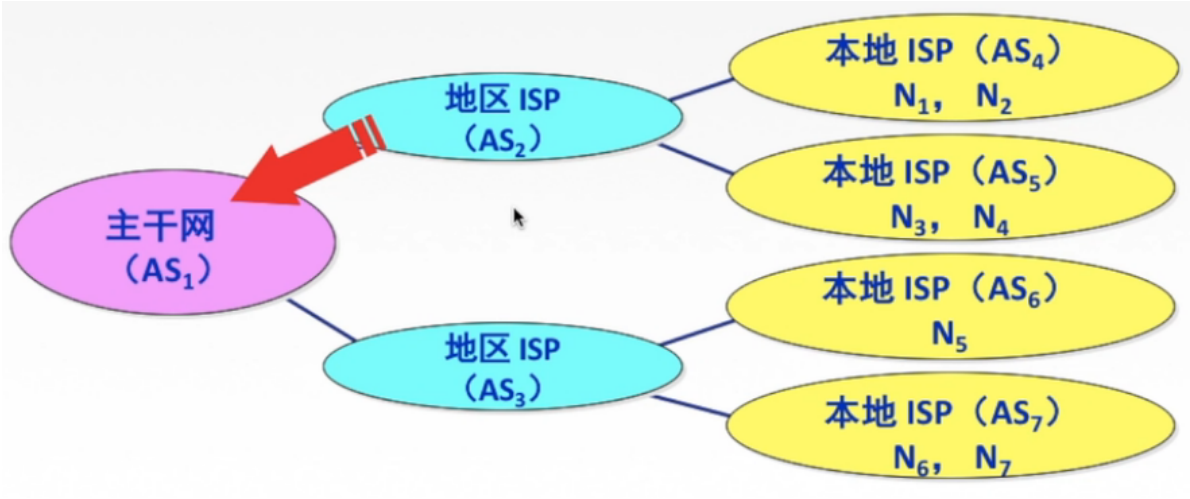
BGP协议交换信息的过程

BGP所交换的网络可达性信息就是要**到达某个网络所要经过的一系列AS。**当BGP发言人互相交换了网络可达性信息后，各BGP发言人就根据采用的策略从收到的路由信息中找出到达各AS的较好路由。

BGP发言人交换**路径向量**：

自治系统AS_2的BGP发言人通知主干网AS_1的BGP发言人：“要到达网络N1、N2、N3和N4可经过AS_2。”

主干网还可以发出通知：“要到达网络N5、N6和N7可沿路径（AS_1, AS_3）。”



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会话交换理由信息。

BGP是**应用层**协议，借助**TCP**传送。

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。
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。

三种路由协议比较

- 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。（应用层）
- OSPF是一个内部网关协议，要交换的信息量较大，应使用报文的长度尽量短，所以不使用传输层协议（如UDP或TCP），而直接采用IP。
- 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。

协议	RIP	OSPF	BGP
类型	内部	内部	外部
路由算法	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	网路中的所有路由器	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

IP组播

IP数据报的三种传输方式

1. 单播：用于发送数据包到单个目的地，且每发送一份单播报文都使用一个单播IP地址作为目的地址。是一种**点对点**传输方式。

   在发送者和每一个接收者之间需要**单独的数据信道**。
2. 广播：发送数据包到同一广播域或子网内的所有设备的一种数据传输方式，是一种**点对多点**传输方式。
3. 组播（多播）：当网络中的某些用户需要特定数据时，组播数据发送者仅发送一条数据，借助组播路由协议为组播数据包建立组播分发树，被传送的数据到达距离用户端尽可能近的节点后才开始复制和分发，是一种**点对多点**传输方式。

   提高数据传输速率，减少主干网出现拥塞的可能性。组播中的主机可以是来自同一个物理网络，也可以来自不同的物理网络（如果有**组播路由器**的支持）。

IP组播地址：让源设备能够将分组发送给一组设备。属于多播组的设备将被分配**一个组播组IP地址**（一群共同需求主机的相同标识）。

组播地址范围为224.0.0.0~239.255.255.255（D类地址），一个D类地址表示一个组播组。只能用作分组的**目标地址**。源地址总是为**单播地址**。

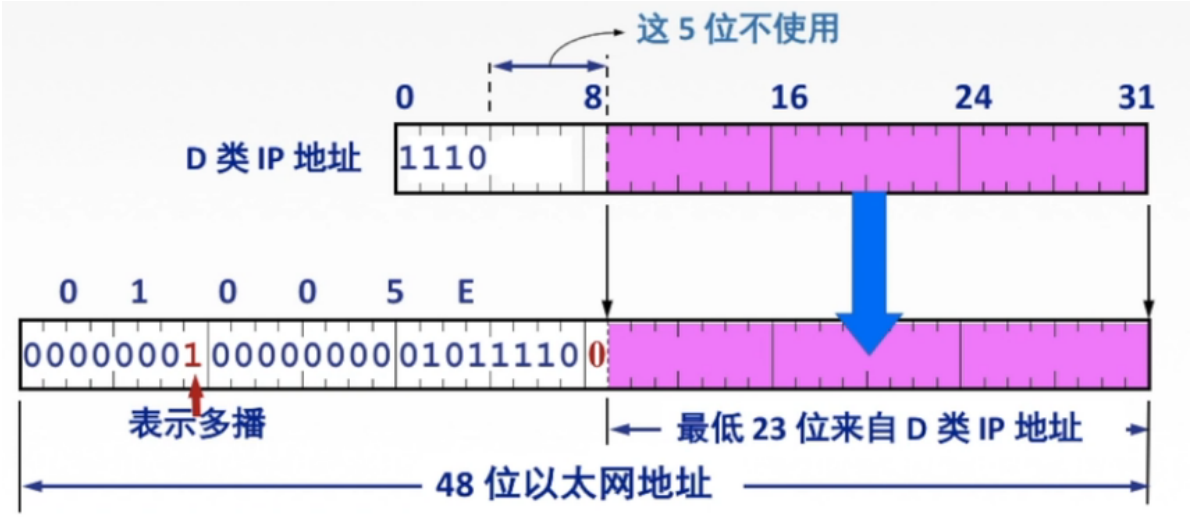
1. 组播数据报也是"尽最大努力交付"，不提供可靠交付，应用于UDP。
2. 对组播数据报不产生ICMP差错报文。
3. 并非所有D类地址都可以作为组播地址。

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**。

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。



网际组管理协议IGMP：让路由器知道本局域网上**是否有主机（的进程）参加或退出了某个组播组。**

ICMP和IGMP都使用**IP数据报**传递报文。

组播路由选择协议：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，对不同的源点也不会有不同的多播转发树。

组播路由协议常使用的三种算法：

1. 基于链路状态的路由选择
2. 基于距离-向量的路由选择
3. 协议无关的组播（稀疏/密集）

移动IP

移动IP技术是移动节点（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。

eg.A刚进入外部网络：

1. 在外部代理登记获得一个转交地址，离开时注销。
2. 外地代理向本地代理登记转交地址。

B给A发送数据报：

1. 本地代理截获数据报
2. 本地代理再封装数据报，新的数据报目的地址是转交地址，发送给外部代理（隧道）。
3. 外部代理拆封数据报并发给A。

A给B发送数据报：

A用自己的主地址作为数据报源地址，用B的IP地址作为数据报的目的地址。

A移动到了下一个网络：

1. 在新外部代理登记注册一个转交地址。
2. 新代理给本地代理发送新的转交地址（覆盖旧的）。
3. 通信

A回到了归属网络：

1. A向本地代理注销转交地址。
2. 按原始方式通信。

网络层设备

路由器：是一种具有多个输入端口和多个输出端口的专用计算机，其任务是转发分组。

根据所选定的路由选择协议**构造出路由表**，同时经常或定期和相邻路由器交换路由信息而不断地**更新和维护路由表。**

交换结构：根据**转发表（路由表得来）**对分组进行**转发。**若收到RIP/OSPF分组等，则把分组送往路由选择处理机；若收到数据分组，则查找转发表并输出。

路由器中的输入输出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。

	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目的网络到输出端口和某些MAC地址信息的映射。

传输层

传输层概述

只有主机才有的层次

传输层的功能：

1. 传输层提供进程和进程之间的逻辑通信。

2. 复用和分用
3. 传输层对收到的报文进行差错检测。
4. 传输层的两种协议。

传输层的两个协议：

- 面向连接的传输控制协议TCP：可靠，面向连接，时延大，适用于大文件。
- 无连接的用户数据报协议UDP：不可靠，无连接，时延小，适用于小文件。

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。

端口：是传输层的SAP，表示主机中的应用进程。（逻辑端口/软件端口）

端口号只有本地意义，在因特网中不同计算机的相同端口是没有联系的。

端口号程度为16bit，能表示65536个不同的端口号。

端口号（按范围分类）：

- **服务端**使用的端口号：
  - 熟知端口号（0~1023）：给TCP/IP最重要的一些应用程序，让所有用户都知道。
  - 登记端口号（1024~49151）：为没有熟知端口号的应用程序使用的。
- **客户端**使用的端口号（49152~65535）：仅在客户进程运行时才动态选择。

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

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。

套接字Socket=（主机IP地址，端口号）

UDP协议

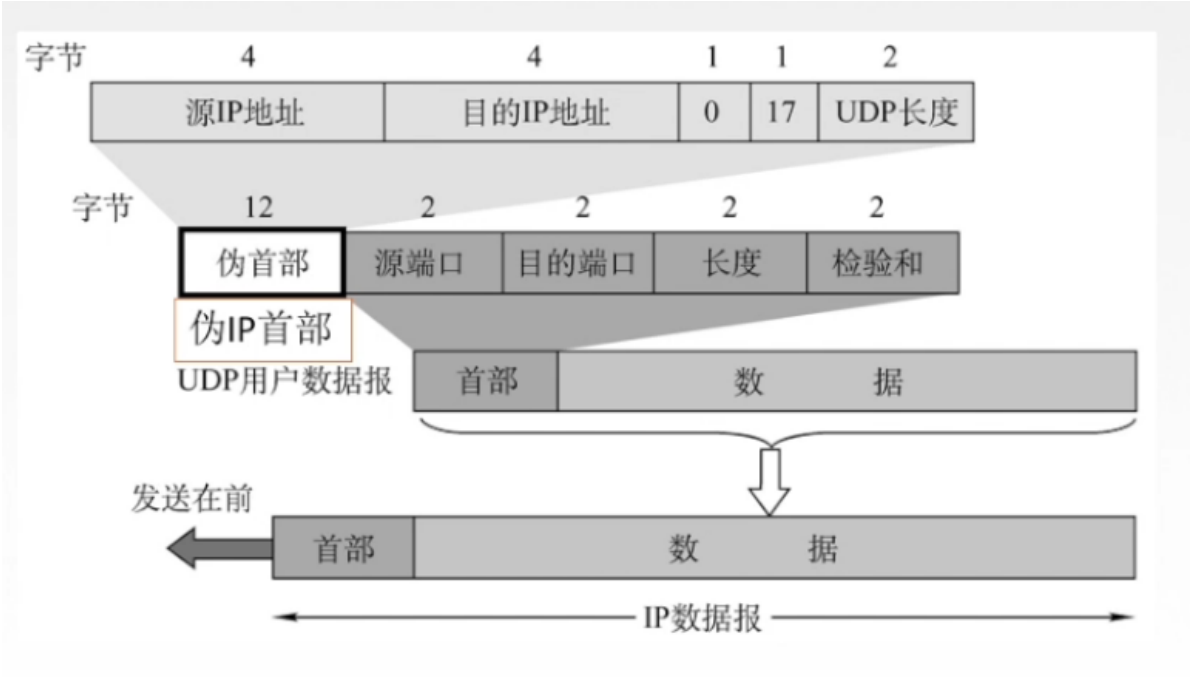
UDP只在IP数据报服务之上增加了很少功能，即复用分用和差错检测功能。

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。

   应用层给UDP多长的报文，UDP就照样发送，即一次发送一个完整报文。
4. UDP无拥塞控制，适合很多实时应用。
5. UDP首部开销小，8B，TCP20B。

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

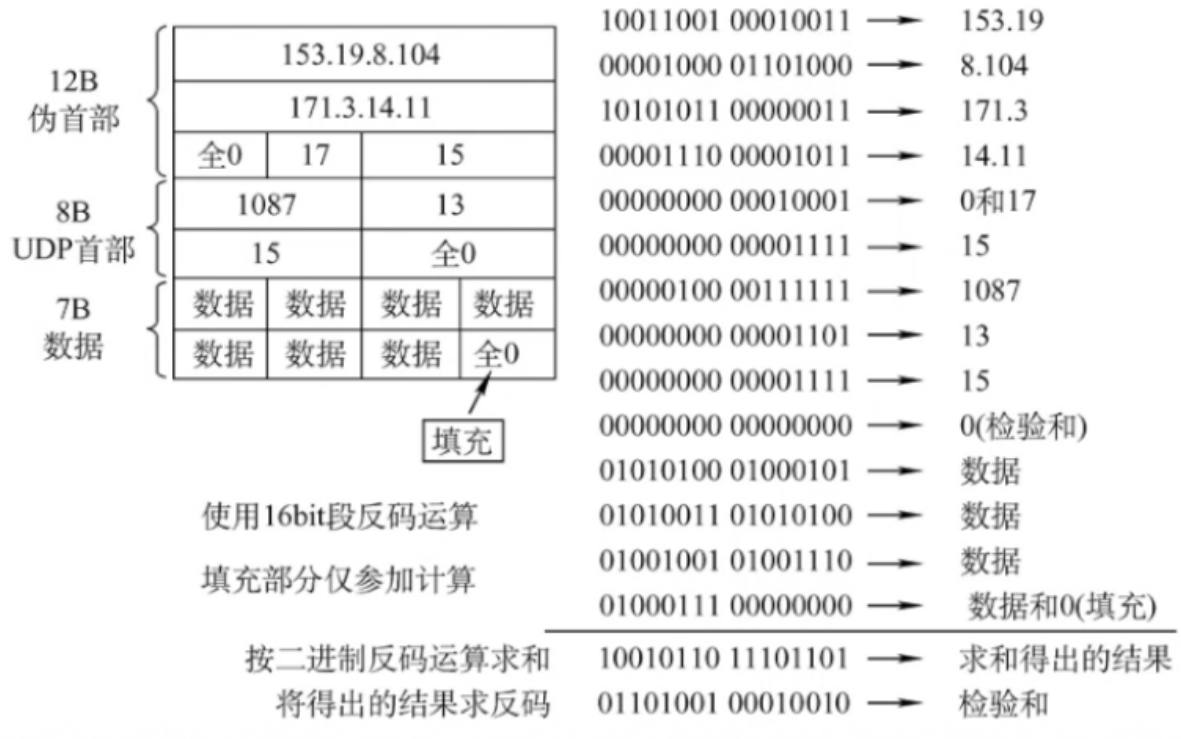


伪首部只有在计算检验和时才出现，不向下传送也不向上递交。

17：封装UDP报文的IP数据报首部协议字段是17。

UDP长度：UDP首部8B+数据部分长度（不包括伪首部）。

UDP校验



在发送端：

1. 填上伪首部
2. 全0填充检验和字段
3. 全0填充数据部分（UDP数据报要看成许多4B的字串接起来）
4. 伪首部+首部+数据部分采用二进制反码求和
5. 把和求反码填入检验和字段
6. 去掉伪首部，发送

在接收端：

1. 填上伪首部
2. 伪首部+首部+数据部分采用二进制反码求和
3. 结果全为1则无差错，否则丢弃数据报/交给应用层附上出差错的警告。

TCP协议特点和TCP报文段

TCP协议的特点

1. TCP是面向连接（虚连接）的传输层协议。
2. 每一条TCP连接只能有两个端点，每一条TCP连接只能是点对点。
3. TCP提供可靠交付的服务，无差错、不丢失、不重复、按序到达。（可靠有序，不丢不重）
4. TCP提供全双工通信。
    - 发送缓存：准备发送的数据&已发送但尚未收到确认的数据
    - 接收缓存：按序到达但尚未被接受应用程序读取的数据&不按序到达的数据
5. 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。**

   流：流入到进程或从进程流出的字节序列。

TCP报文段首部格式

序号：在一个TCP连接中传送的字节流中的每一个字节都按顺序编号，本字段表示本报文段所发送数据的**第一个字节的序号。**

确认号：**期望**收到对方下一个报文段的第一个数据字节的序号。若确认号为N，则证明到序号N-1为止的所有数据都已正确收到。

数据偏移（首部长度）：TCP报文段的数据起始处距离TCP报文段的起始处有多远，以4B为单位，即1个数值是4B。

6个控制位

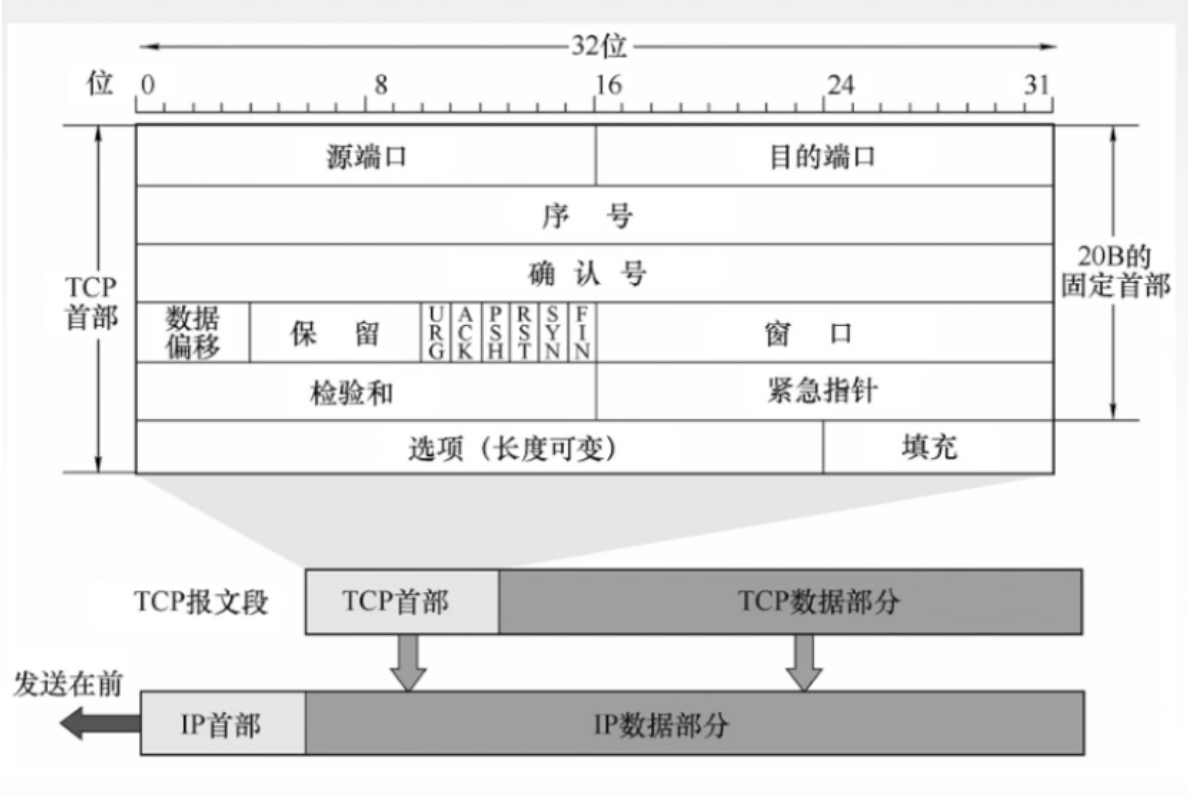
- **紧急位URG**：URG=1时，标明此报文段中有紧急数据，是高优先级的数据，应尽快传送，不用在缓存里排队，配合紧急指针字段使用。
- **确认位ACK**：ACK=1时确认号有效，在连接建立后所有传送的报文段都必须把ACK置为1。
- 推送位PSH：PSH=1时，接收方尽快交付接收应用进程，不再等到缓存填满再向上交付。
- 复位RST：RST=1时，表明TCP连接中出现严重差错，必须释放连接，然后再重新建立传输链接。
- **同步位SYN**：SYN=1时，表明是一个连接请求/连接接受报文。
- **终止位FIN**：FIN=1时，表明此报文段发送方数据已发完，要求释放连接。

窗口：指的是发送本报文段的一方的接收窗口，即现在允许对方发送的数据量。

检验和：检验首部+数据，检验时要加上12B伪首部，第四个字段为6。

紧急指针：URG=1时才有意义，指出本报文段中紧急数据的字节数。

选项：最大报文段长度MSS、窗口扩大、时间戳、选择确认...



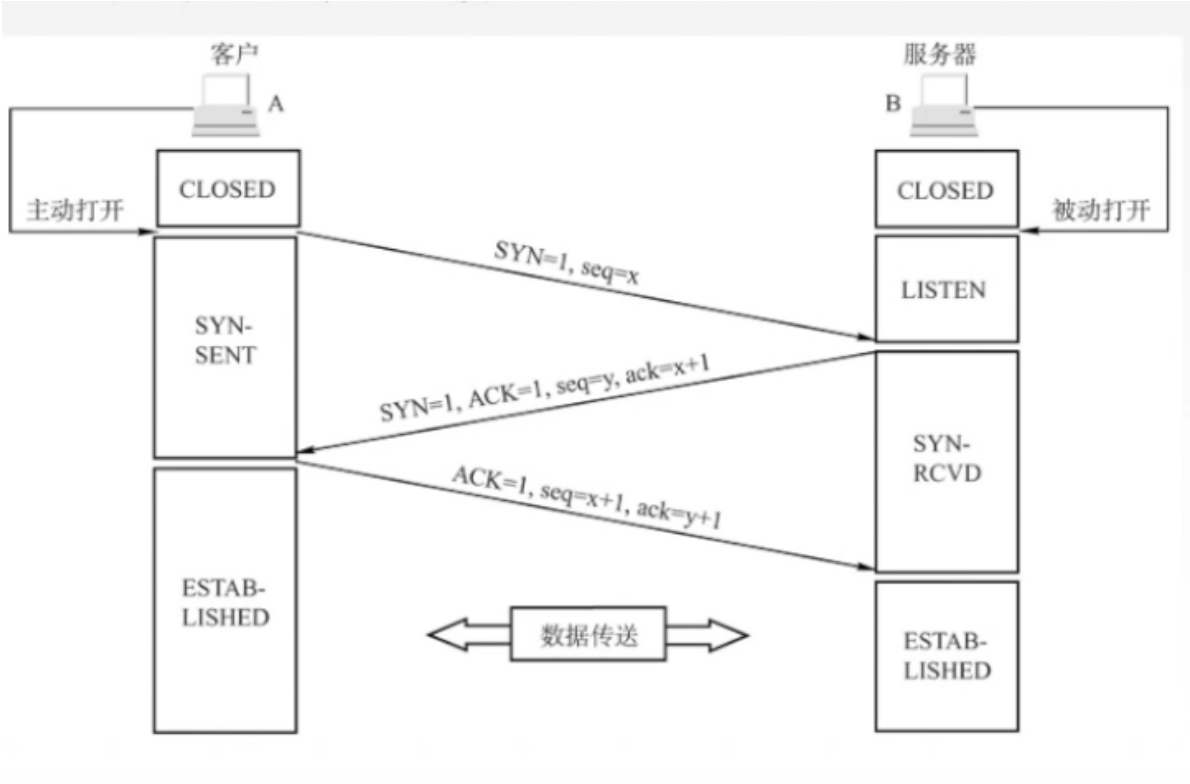
TCP连接管理

TCP连接传输三个阶段：连接建立，数据传送，连接释放

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。

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会用一下步骤与服务器中的TCP建立一条TCP连接：

1. 客户端发送**连接请求报文段**，无应用层数据。SYN=1，seq=x（随机）
2. 服务端为该TCP连接**分配缓存和变量**，并向客户端返回**确认报文段**，允许连接，无应用层数据。SYN=1，ACK=1，seq=y（随机），ack=x+1
3. 客户端为该TCO连接**分配缓存和变量**，并向服务器端返回确认的确认，可以携带数据。SYN=0，ACK=1，seq=x+1，ack=y+1



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，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。

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。

1. 客户端发送**连接释放报文段**，停止发送数据，主动关闭TCP连接。FIN=1，seq=u
2. 服务器端回送一个确认报文段，客户到服务器这个方向的连接就释放了——半关闭状态。ACK=1，seq=v，ack=u+1
3. 服务器端发完数据，就发出连接释放报文段，主动关闭TCP连接。FIN=1，ACK=1，seq=w，ack=u+1
4. 客户端回送一个确认报文段，再等到时间等待计时器设置的2MSL（最长报文段寿命）后，连接彻底关闭。ACK=1，seq=u+1，ack=w+1

TCP可靠传输

TCP可靠传输机制

1. 校验：与UDP校验一样，**增加伪首部**
2. 序号：一个字节占一个序号。**序号字段**指的是一个报文段第一个字节的序号。
3. 确认：TCP默认使用累计确认。
4. 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）

   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）。

冗余ACK（冗余确认）：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。

TCP流量控制

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通知发送给发送方），发送方的**发送窗口取接收窗口rwnd和拥塞窗口cwnd的最小值。**

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。

TCP拥塞控制

出现拥塞的条件：对资源需求的总和 > 可用资源，网络中有许多资源同时呈现供应不足 -> 网路性能变坏 -> 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）

拥塞控制四种算法：

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. 接收方总是有足够大的缓存空间，因而发送窗口大小取决于拥塞程度

应用层

网络应用模型

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

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

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
- 网络健壮性强

域名解析系统DNS

域名

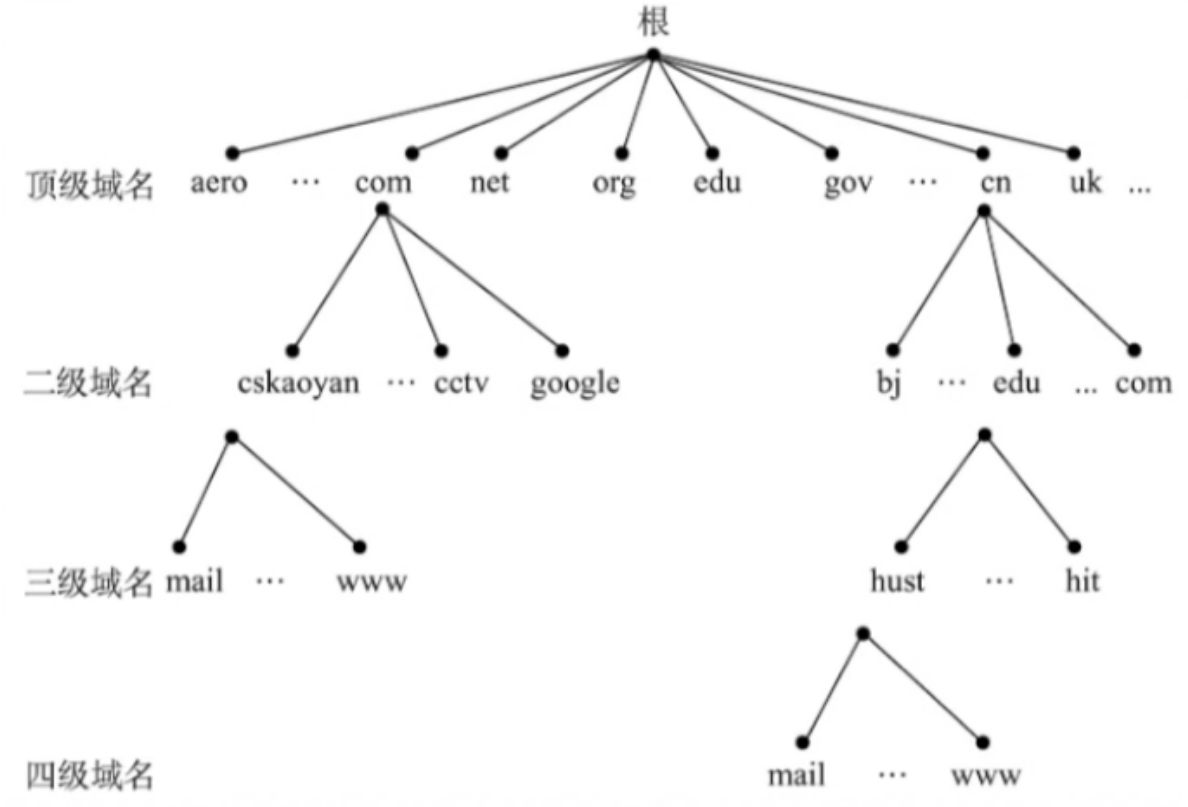
根：.

顶级域名：

- 国家顶级域名：cn,us,uk
- 通用顶级域名：com,net,org,gov,int,aero,museum,travel
- 基础结构域名/反向域名 arpa

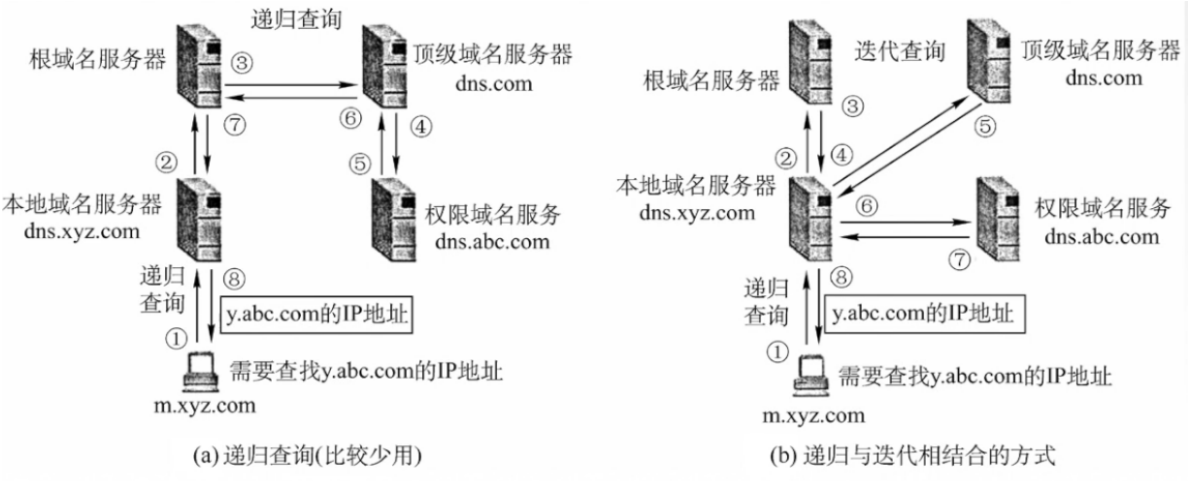
二级域名：

- 类别域名：ac,com,edu,gov,mil,net,org
- 行政区域名：用于我国各省、自治区、直辖市 bj,js



域名服务器

1. 根域名服务器
2. 顶级域名服务器：管理该顶级域名服务器注册的所有二级域名
3. 权限域名服务器：负责一个区的域名服务器
4. 本地域名服务器：当一个主机发出DNS查询请求时，这个查询请求报文就发给本地域名服务器。



文件传送协议FTP

文件传送协议FTP（File Transfer Protocol）：提供不同种类主机系统（硬、软件体系等都可以把不同）之间的文件传输能力。

FTP服务器和用户端

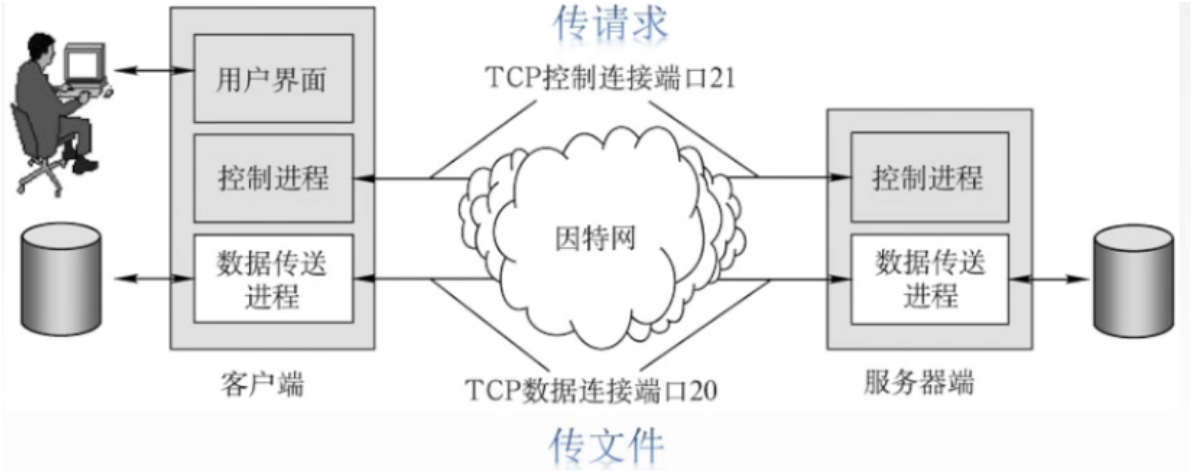
FTP是基于客户/服务器（C/S）的协议。用户通过一个客户机程序连接至在远程计算机上运行的服务器程序。依照FTP协议提供服务，进行文件传送的计算机就是**FTP服务器。**连接FTP服务器，遵循FTP协议与服务器传送文件的电脑就是**FTP客户端。**

FTP工作原理

登录：ftp地址，用户名&密码

匿名登录：互联网中有很大一部分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”登录FTP服务，就可以访问远程主机上公开的文件。

FTP使用TCP实现可靠传输。



控制连接**始终**保持。**数据**连接保持一会儿。

是否使用TCP 20端口建立**数据连接**与**传输模式**有关。

- 主动方式：使用TCP 20端口
- 被动方式：由服务器和客户端自行协商决定（端口>1024）

FTP传输模式

- 文本模式：ASCII模式，以文本序列传输数据；
- 二进制模：Binary模式，以二进制序列传输数据。

电子邮件

电子邮件的信息格式

组成结构

- 用户代理：电子邮件客户端软件。

  功能：1. 撰写2. 显示　3. 处理　4. 通信
- 邮件服务器：C/S

  功能：1. 发送&接收邮件　2. 向发件人报告邮件传送结果
- 协议：SMTP（发）POP3、IMAP（收）

简单邮件传送协议SMTP：TCP连接，端口号25，C/S

SMTP规定了两个互相通信的**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）。

SMTP通信三个阶段：

1. 连接建立：SMTP服务器若有能力接收邮件，回答“250 OK”；否则，回答“421 Service not available”
2. 邮件发送：

```
A:MAIL FROM: <wangdao@163.com>
B:250 OK/B:451(452、500...)      STMP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

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:DATA....      开始传输邮件内容

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。

SMTP的缺点：

1. SMTP不能传送可执行文件或其他二进制对象。
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。

通用因特网扩充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。

邮局协议POP3：TCP连接，端口号110，C/S

工作方式：

- 下载并保留（在服务器）
- 下载并删除

网际报文存储协议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箱，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部分（先看正文，有WIFI的时候再下载）。

基于万维网的电子邮件：方便

万维网和HTTP协议

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，是无数歌网络站点和网页的集合。资源（文字、视频、音频....）统一资源定位符**URL**唯一标识资源。

URL一般形式：<协议>://<主机>:<端口>/<路径>

用户通过点击超链接获取资源，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。

超文本传输协议HTTP：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。

具体过程：

1. 浏览器分析URL
2. 浏览器向DNS请求解析IP地址
3. DNS解析出IP地址
4. 浏览器与服务器建立TCP连接
5. 浏览器发出取文件命令
6. 服务器响应
7. 释放TCP连接
8. 浏览器显示

HTTP协议的特点

HTTP协议是**无状态**的。

Cookie小饼干

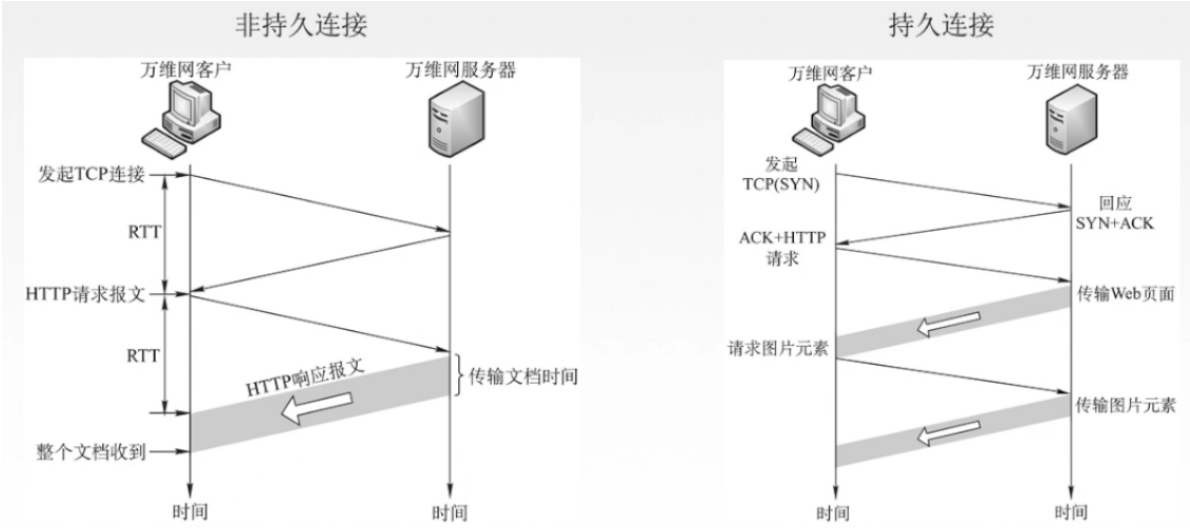
但是在实际工作中，一些万维网站点常常希望能够识别用户。

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。

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）。

HTTP的连接方式：

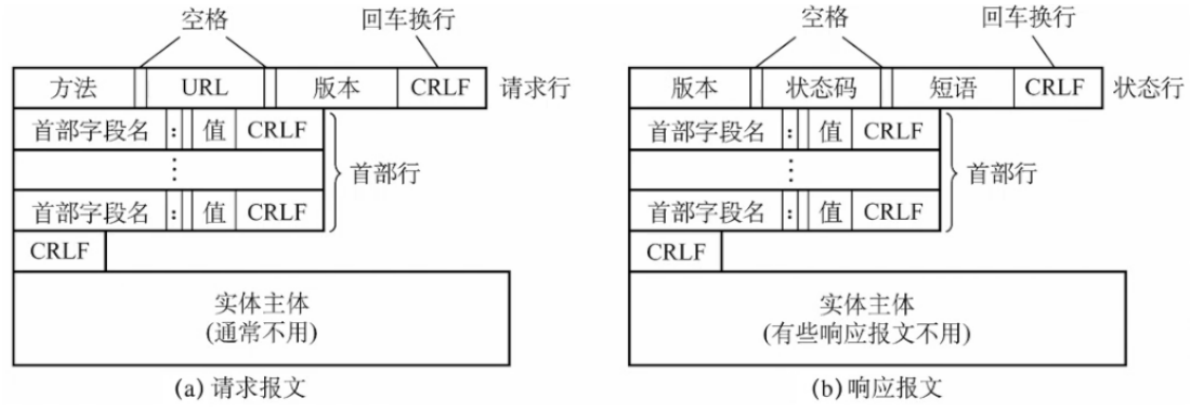
- 持久连接（Keep-alive）
    - 非流水线
    - 流水线
- 非持久连接（Close）



报文结构

- 请求报文
- 响应报文

HTTP报文是**面向文本**的，因此在报文中的每一个字段都是ASCII码串。



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。