信息工程系

教

案

课程名称： 网络基础

教　　师： 黄苗苗

总 学 时 ： 72

理论学时： 36

实训学时： 36

上课班级： 计算机 251

授课学期： 25-26(2)

课程说明

计算机《网络基础》是计算机科学与技术、软件工程、网络工程、信息管理与信息系统等专业的必修课程。在当前的国民经济中，计算机网络通信技术应用越来越广泛，地位越来越重要。因此，掌握计算机网络通信技术是每一个算机科学与技术、软件工程、网络工程等专业学生必修的专业课程。

课程要求侧重掌握计算机网络体系结构、体系结构中各层次意义及其相互间关系以及网络互连等知识。《网络基础》课程为将来从事计算机网络通信领域的开发和研究、网络的使用和维护提供必要的基础知识，打下良好的基础，而且还是实践技能训练中的一个重要的教学环节。因此要求上述专业学生都必须掌握本课程的内容。

课程要求：

1. 通过学习，学生要全面系统地掌握计算机网络的发展历史、计算机网络体系结构、物理层、数据链路层、局域网、网络互连、运输层、高层协议、网络新技术和计算机网络安全等内容。

2. 通过学习，强化学生用分层次的体系结构来分析资源子网和通信子网的能力。通过网上练习和实验，验证和掌握计算机网络的安装、配置、调试、开发与应用，提高利用计算机解决实际网络通信问题的实践技能。

课程性质：

计算机基础专业课程，本课程是对计算机网络的体系结构做了全面的介绍，关系以后相关课程和相关技术开发，是网络技术编程的基础、是计算机网络应用开发的前景知识准备，需要学生对此投入较大的学习量，才能掌握好相关概念，并将其应用到实际操作中。

课程思政体现：

我们的《网络基础》课程改革的方向，融入了思政改革的时代要求，以培养、提升岗位职能能力为目的，具体课程改革涵盖：

1.培养有国家意识、人文情怀的人才，激发学生的爱国热情和民族自信。

2.着重培养具有科学精神、具有计算机网络方向专业能力和专业素养的专业人才。

3.注重学生的个性化培养，符合立德树人的中心思想，同时也尊重学生的个性和兴趣发展。

4.辅助完善学生职业素养的形成。

课程教学手段说明

依据课程性质及周课时安排：

理论课（2 节/周）：

课程主要是从总体概述之后，分层介绍各个网络层结构。

结合课程的各层结构，在每个章节的授课中，分几个主要方式进行：

- **提问式引导**：直接就本章要进行的问题，留一定时间给学习者进行章节概述简介阅读和短暂思考。结合日常接触的网络情境，对问题直接进行提问。然后纵观全章，给出整章概貌的描述。
- **承上启下式进入**：对于上一个章节，进行总结，然后在上一章节的基础上，对新的可能出现的情境进行引导式引入，然后就新的问题的本质、现象出现的方式以及其如何工作引入思考，然后进入章节介绍。

引入章节之后，对于每个章节，理顺一个相同的网络层次共同点，进行对比介绍，帮助学习者尽快进入新知识的学习。主要的各章特点可归纳总结如下：

- 网络层次负责的总体功能概述
- 网络层次包含的主要工作部件、协议有哪些
- 主要协议的介绍
- 网络工作数据单元
- 封装的格式及各个意义
- 实际应用问题及情境
- 课后相关习题巩固

实验课（1 节/周）：

实验课是为了巩固课堂实际理论而开设的，针对本课程特性，特设置了对实际接触网络的认识、对实际通信物理介质认识、网络工作原理认识、网络应用程序的应用认识。

其中，网络工作原理是本课程的重点，主要采取了 Cisio 出品的 Packet Tracer 工具导入作为实验的主要模拟环境。锻炼学生对小学习环境的适应能力，然后以此，实际结合网络知识以及日常工作可能布置的网络拓扑结构进行网络环境模拟和网络实际情况模拟分析。比较切合实际 的网络运行环境，同时培养学生对新工作环境的适应学习过程。故此，实验课的难度系数属于中等难度系数，需要学生克服依赖心理执行方可。

第一章 概 述

❒ 课时安排：

4 学时

❒ 教学目的：

1. 认识因特网
2. 熟练掌握因特网的组成
3. 熟练掌握因特网的体系结构
4. 认识因特网发展历史、因特网类别、因特网性能

❒ 教学重、难点：

1. 因特网的体系结构是本章的重点和难点。
   a) 体系结构的分层
   b) 体系结构各层的功能
   c) 体系结构如何协作运行
2. 因特网的性能指标是本章的重点：由此，学习者可以对网络的衡量标准达到一个质的认识，对于以后网络相关知识和相关实际应用中的情景进行实际问题的解决。

❒ 教学内容：

1．1 建立计算机网络的目的

1．目的

- 资源共享
- 高可靠性
- 节约经费
- 通信手段

2．计算机网络与分布式系统的区别

分布式系统是建立在计算机网络之上的软件系统，它具有高度的整体性和透明性。因此计算机网络和分布式系统的区别在于软件（尤其是操作系统）而不是硬件。

1．2 计算机网络的发展过程（四代）

1．2．1 通信与计算机的结合——产生计算机网络（电路交换）

- 通信网络为计算机之间的数据传递和交换提供了必要的手段。

- 数字计算机技术的发展渗透到通信技术中，又提高了通信网络的各种性能。
- 电路交换：建立连接　　数据通信　　释放连接
- 电路交换的分类：**空分交换**是交换比特流所经过的**端口号**；**时分交换**是交换比特所在的**时隙**；**波分交换**是交换荷载比特的光的**波长**。

1．2．2　分组交换网的出现（包交换）

- 传统的电路交换技术不适合计算机数据的传输。
- 分组交换网的试验成功：存储转发原理——即断续（或动态）分配传输带宽。
- 分组交换的主要特点：高效、灵活、迅速、可靠。
- 分组交换网：以通信子网为中心，主机和终端都处在网络的外围。
- 电路交换、报文交换和分组交换的主要区别：参见课本 P5 图 1-4。

1．2.3　计算机网络体系结构的形成：OSI/RM（ISO）、TCP/IP（Internet）、SNA（IBM）、DNA（Digital）等。（分层网络体系结构的形成）

- OSI/RM：开放系统互联基本参考模型。
- TCP/IP：INTERNET 的体系结构。
- SNA：系统网络体系结构，1974 年出现，世界上第一个网络体系结构。
- DNA：分布式系统体系结构。

1．2．4　B-ISDN：综合化：即各种业务综合　　高速化：即宽带化

采用高速分组交换、高速电路交换、异步传输模式 ATM 和光交换的高速综合业务数字网就称为 B-ISDN，ATM 交换是电路交换和分组交换的结合。

补充：ISDN 知识。

- ISDN：综合业务数字网。是由综合数字电话网（IDN）演变发展而来的一种网络，它提供端到端的数字连接以支持广泛的业务，包括语音的非语音的业务。它为用户进网提供了一组少量的标准多用途网络接口。
- ISDN 的特性：端到端的数字连接、综合的业务、标准的入网接口。
- ISDN 的用户—网络接口：

（1）基本速率接口 BRI（Basic Rate ISDN）：2B+D，B 信道为载荷信道，速率为 64kbps，D 信道为信令信道，速率为 16kbps。BRI 一般用于较低速率的系统中。

（2）一次群（基群）速率接口 PRI（Primary Rate ISDN）：一般用于大容量系统。

23B+D：美国和日本采用，可适应北美的 T1 系统（1.544Mbps）。

30B+D：欧洲采用，可适应 E1 系统（2.048Mbps）。

- B-ISDN：采用高速分组交换、高速电路交换、异步传输模式 ATM 和光交换四种传输模式。

高速分组交换：利用分组交换的基本技术，简化了 X.25 协议，采用面向连接的服务，在链路上无流量控制、无差错控制，集中了分组交换和同步时分交换的优点。

高速电路交换：采用多速时分交换方式（TDSM）允许信道按时间分配，其带宽可为基本速率的整数倍，但信道的管理和控制十分复杂。

光交换：采用光交换机将光技术引入传输回路和控制回路，实现数字信号的高速传输和交换。

异步传输模式 ATM：集电路交换的实时性和分组交换的灵活性于一体，能适应各种类型的业务。

- B-ISDN 与 N-ISDN 的区别：

（1）N-ISDN 使用的是电路交换。只是在传送信令的 D 通路使用分组交换；B-ISDN 则使用一种快速分组交换，称为异步传递方式 ATM。

（2）N-ISDN 是以目前正在使用的电话网为基础，其用户环路采用双绞线（铜线）；但在 B-ISDN 中，其用户环路和干线都采用光缆（但短距离也可以使用双绞线）。

（3）N-ISDN 各通路的比特率是预先设置的。如 B 通路比特率为 64kb/s；但 B-ISDN 使用虚通路的概念，其比特率只受用户到网络接口的物理比特率的限制。

（4）N-ISDN 无法传送高速图像。但 B-ISDN 可以传送。

1．3 协议与体系结构

1．3．1 计算机网络体系结构的形成

- 开放系统互联基本参考模型：OSI/RM，理论上的国际标准。
- TCP/IP：事实上的国际标准。

1．3．2 网络的体系结构

计算机网络的各层及其协议的集合，称为计算机网络的体系结构。计算机网络的体系结构就是这个计算机网络及其部件所应完成的功能的精确定义。

体系结构是抽象的，而实现是具体的，是真正在运行的计算机硬件和软件。

1．网络协议：**以教师课堂上与学生交流为例来说明。**

为进行网络中的数据交换而建立的规则、标准或约定。包括语法、语义和同步。

- 语法：数据与控制信息的结构或格式。
- 语义：要发出何种控制信息，完成何种动作以及做出何种应答。
- 同步：事件实现顺序的详细说明，即时序问题。

2．分层的优缺点：

优点：各层之间是独立的；灵活性好；结构上可分割开；易于实现和维护；能

促进标准化工作。

缺点：分层的层次数难以确定；有些功能会在不同的层次中重复出现，而产生了额外开销。

1．3．3 计算机网络的原理体系结构

1．说明物理层、数据链路层、网络层、运输层和应用层的功能。

物理层：在物理媒体上透明地传输比特流。透明表示某一个实际存在的事物看起来却好像不存在一样（以玻璃为例）。

数据链路层：在不太可靠的物理链路（两个相邻结点间）上，实现无差错的帧传输。

网络层：负责为互连网上的不同主机之间提供分组交换。主要任务就是路由选择和分组转发。

运输层：负责主机中两个进程之间的通信，数据传输单位是报文段。

应用层：直接为用户的应用进程提供服务。

2．以示意图的形式说明 SDU（服务数据单元）、PDU（协议数据单元）、IDU（接口数据单元）之间的关系。（参见 P17 图 1-12）

（N）PDU=（N）SDU+（N）PCI

IDU=PDU+ICI （接口控制信息）

3．以示意图的形式说明数据通信过程。（课本 P15 图 1-11，P25 图 1-17）

以乘飞机旅行为例来说明。

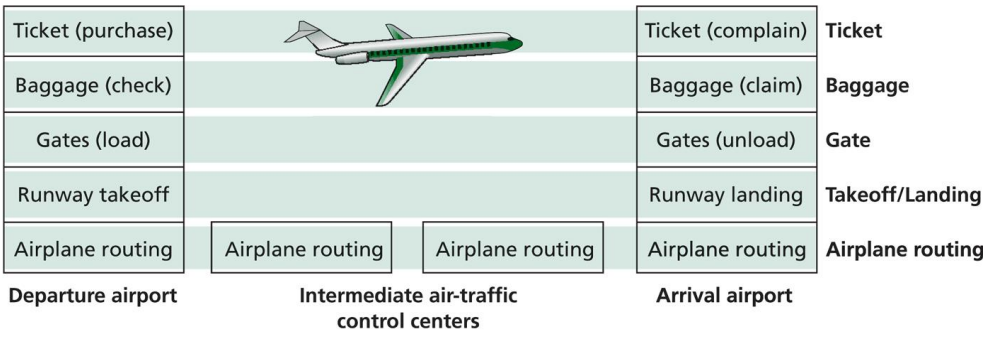


Figure 1.16 ♦ Horizontal layering of airline functionality

4．说明实体、协议、服务和服务访问点的概念。

- 协议保证服务得以向上一层提供，本层的服务用户只能看见服务而无法看见下面的协议，下面的协议对上面的服务用户是透明的。
- 协议的水平的，即控制对等实体之间的通信。而服务是垂直的，是由下层向上层通过层间接口提供的一组原语（操作）。下层能向上层提供两种不同形式的服务，即面向连接的服务和无连接的服务。（参见课本 P19 图 1-13）

- 服务：能够被高一层看得见的功能才能称之为“服务”。
- 服务访问点 SAP：在同一系统中相邻层的实体进行交互（即交换信息之处），一个服务访问点只能被一个实体所使用。
- 服务原语

原语	**含　义**
请求	一个实体希望得到完成某些操作的服务
指示	通知一个实体，有某个事件发生
响应	一个实体希望响应一个事件
证实	返回对先前请求的响应

- 服务类型：面向连接服务和无连接服务。

连接类型	服务类型	应用实例
面向连接的服务	可靠的消息流	页码序列
	可靠的字节流	远程登陆
	不可靠的连接	数字化的声音
无连接服务	不可靠的数据报	电子方式的函件
	有确认的数据报	挂号邮件
	问答	数据查询

1．3．4　网络体系结构详述

1. OSI 的体系结构：物理层，数据链路层，网络层，传输层，会话层，表示层，应用层。

（课本 P22 图 1-16）分 7 层，分层原则如下：

（1）根据不同层次的抽象分层。

（2）每层应当实现一个定义明确的功能。

（3）每层功能的选择应该有助于制定网络协议的国际标准。

（4）各层边界的选择应尽量减少跨过接口的通信量。

（5）层数应足够多，以避免不同的功能混杂在同一层中，但也不能太多，否则体系结构会过于庞大。

2. TCP/IP 参考模型：主机至网络层，互连网层，传输层，应用层。

（课本 P28 图 1-19）

3. TCP/IP 与 OSI 的对比：（课本 P27 图 1-18）

（1）两者之比较

- TCP/IP 一开始就考虑到多种异构网的互连问题，并将网际协议 IP 作为

TCP/IP 的重要组成部分。但 ISO 和 CCITT 最初只考虑到使用一种标准的公用数据网将各种不同的系统互连在一起。

- TCP/IP 一开始就对面向连接服务和无连接服务并重，而 OSI 在开始时只强调面向连接服务。
- TCP/IP 有较好的网络管理功能。而 OSI 到后来才开始考虑这个问题。
- TCP/IP 对一些基本概念没有很清楚的区分，而且其模型的通用性较差。

（2）OSI 模型和协议的缺点

- 糟糕的提出时机
- 糟糕的技术
- 糟糕的现实
- 糟糕的策略

（3）TCP/IP 参考模型的缺点

- 该模型没有明显地区分服务、接口和协议的概念。
- TCP/IP 模型完全不是通用的，并且不适合描述除 TCP/IP 模型之外的任何协议栈。
- 主机至网络层在分层协议中根本不是通常意义下的层。
- TCP/IP 模型不区分（甚至不提及）物理层和数据链路层。
- 虽然 IP 和 TCP 协议被仔细地设计，并很好的实现了。但是其他很多协议却很特别，没有被很好的实现就免费发送，造成现在很难被替换。

1．4 计算机网络的分类

1．按照传输技术分类：

- 广播式网络：仅有一条通信信道，由网络上的所有机器共享。如总线型、环型。
- 点到点网络：由一对对机器之间的多条连接构成。为了能从源到达目的地，这种网络上的分组可能必须通过一台或多台中间机器。

2．按网络的交换功能分类：

- 电路交换
- 报文交换：
- 分组交换：
- 混合交换：在一个数据网中同时采用电路交换和分组交换。

3．按拓朴结构分类：

- 集中式网络：
- 分散式网络：

- 分布式网络：

4．按作用范围分类：

- 广域网 WAN
- 局域网 LAN
- 城域网 MAN

5．按使用范围分类：

- 公用网：
- 专用网：

1．5 关于计算机网络的若干术语

1．计算机网络：由若干个主机，一个通信子网和一系列的协议组成。

2．计算机网络的标准制定机构：

- 国际标准化组织（ISO）：
- 国际电报电话咨询委员会（CCITT）：现改名为国际电信联盟电信标准化局（ITU-T）
- 美国国家标准局（NBS）：
- 美国国家标准学会（ANSI）：包括电子工业协会（EIA）、电气和电子工程师学会（IEEE）。
- 欧洲计算机制造商协会（ECMA）：

3．计算机网络与分布式计算机系统

相同点：物理结构相同，都是建立在网络结构之上的系统。

不同点：高层软件不同，即网络操作系统与分布式操作系统的区别。

1．6 计算机网络在我国的发展

- 中国公用计算机互联网 CHINANET
- 中国金桥信息网 CHINAGBN
- 中国教育和科研计算机网 CERNET
- 中国科学技术网 CSTNET
- 中国联通数据网
- 网通公用互联网 CNCnet
- 利用军队资源组建的数据网

1．7 计算机网络的主要技术指标

1．带宽：某个信号具有的频带宽度。相近概念：**最高数据率**、**吞吐量**

2．时延：

（1）发送时延（传输时延）：结点在发送数据时使数据块从结点进入到传输媒体

所需要的时间。

发送时延＝ 数据块长度/信道带宽

（2）传播时延：电磁波在信道中需要传播一定距离而花费的时间。

传播时延＝ 信道长度/电磁波在信道上的传播速率

（3）处理时延:数据在交换结点为存储转发而进行一些必要的处理所花费的时间。

例：以高速公路收费站为例来说明传输时延和传播时延。

3．时延带宽积

时延带宽积＝ 传播时延 × 带宽

从公式可知：时延带宽积又可以称为**以比特为单位的链路长度**。

4．往返时延：表示从发送端发送数据开始，到发送端收到接收端的确认，总共经历的时延。

5．其他相关概念：

- 信息：信息是对客观物质的反映，可以是对物质的形态，大小，结构，性能等全部或部分特性的描述，也可以是物质与外部的联系。信息有各种存在的形态，如文字，声音，图像等等。
- 数据：信息也可以用数字的形式表示，数字化的信息称为数据。数据是装载信息的实体，信息则是数据的内在含义或解释。为了确切的表示信息，数据有时是一些连续值，另一些则取离散值，如声音的强度，灯光的亮度等都可以连续变化，而成绩，名次等的取值都是离散的。连续值的数据叫做模拟数据，离散值的数据叫做数字数据。
- 信号：数据通信中的“信号”是指数据的电编码或磁编码。它分为模拟信号和数字信号两种。模拟信号是连续变化的电磁波，数字信号则是一串电压脉冲序列。两种信号在一定技术措施下可以相互转换。信号可以在各种传输媒介上传送，如双绞线，电话线，同轴电缆，光缆，甚至可以通过卫星以微波的方式传送。
- 噪声：信号在传输过程中受到的干扰称为噪声，干扰可能来自外部，也可能由信号传输过程本身产生。噪声过大将影响被传送的信号的真实性或正确性。
- 信道：信道是传送信号的一条通路，由传输介质及相应的附属设备组成。同一个传输介质上可以同时存在多条信号通路，即一条传输线路上可以有多个信道。例如一条光缆可以供上千对人同时通话，有上千个电话信道。
- 信号带宽：信号通常都是以电磁波的形式传送的，电磁波都有一定的频谱范围，该频谱范围称作该信号的带宽。理论上任意持续期有限信号（如脉冲信号，方波等）的频谱总是无限宽的，但在实际应用中，频谱宽度被看作是信号能量比较集中的那样一个频谱范围。

- 信道带宽：指信道上能够传送的信号的最大频率范围，如普通电话信道的带宽是300HZ～3400HZ。当信号带宽大于信道带宽时，信号就不能在该信道上传送，或者传送出的信号将失真。
- 信道容量：是信道允许的最大数据传输率，是信道性能的极限。如果要求的数据率大于信道容量，这样的传输在该信道上根本无法实现。
- 吞吐量：吞吐量是单位时间内整个网络能够处理的信息总量。在信道总线网络中，吞吐量＝信道容量×传输效率。
- 传输效率：指原始数据量占整个传送的数据的比率，数值上等于数据包中数据的长度与整个包长度的比值。显然传输效率越高越好。

1．8 应用层的客户－服务器方式

客户是服务请求方，服务器是服务提供方。

第二章 物理层

❒ 课时安排：

3 学时

❒ 教学目的：

1. 认识物理层功能
2. 衡量物理层的几个性能标准分别是什么
3. 掌握几类常见的组网线缆类型

❒ 教学重、难点：

1. 物理层的衡量标准，如何决定组网的线缆类型
2. 对于性能指标，几个制约因素当如何对待

❒ 教学内容：

2．1 物理层的基本概念

1、物理层考虑的是怎样才能在连接各种计算机的传输媒体上传输数据的比特流，而不是指连接计算机的具体的物理设备或具体的传输媒体。

2、物理层的作用：

- 采用 OSI 术语，其作用为确定与传输媒体接口的一些特性：

  （1）机械特性：指明接口所用接线器的形状和尺寸、引线数目和排列、固定和锁定装置。

  （2）电气特性：指明在接口电缆的那条线上出现的电压的范围。

  （3）功能特性：指明某条线上出现的某一电平的电压表示何种意义。

  （4）规程特性：指明对于不同功能的各种可能事件的出现顺序。

- 采用 OSI 术语，其作用为给其服务用户在一条物理的传输媒体上传送和接收比特流的能力。

2．2 信道的极限容量

2．2．1 有关信道的几个基本概念

1、信道：用来表示向某一个方向传送信息的媒体

2、通信的基本方式：

- 单向通信
- 双向交替通信

- 双向同时通信

3、通信信号；模拟信号、数字信号

- 基带信号：将数字信号 1 或 0 直接用两种不同的电压来表示，送到线路上传输。
- 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。

2．2．2 信道上最高码元传输速率

奈氏准则：每赫带宽的理想低通信道的最高码元传输速率是每秒 2 个码元。

- **理想低通信道的最高码元传输速率**=2W Baud

其中：W 为理想低通信道的带宽，单位为赫兹；25.

Baud 是波特，码元传输速率、信号传输速率或调制速率的单位，1 波特为每秒传送 1 个码元，一个数字脉冲即为一个码元（注意强调波特与比特的区别）。

码元传输速率表示单位时间内通过信道传输的码元个数。若信号码元（脉冲）的宽度为 T 秒，则码元传输速率为 1/T 波特。

- **理想带通信道的最高码元传输速率**=W Baud
- **无噪声信道容量**=2W $\log_2$N bps （其中 N 表示携带数据的码元可能取的离散值的个数）
- **若在有噪声情况下，且误码率为 P，则信道容量＝**2W[1−P $\log_2$(1/P) − (1−P) $\log_2$(1/(1−P))]bps

例如：P＝0.1，信道容量＝2W×0.53，即有效信息速率降低了将近一半。

2．2．3 信道的极限信息传输速率仙农公式：

信道的极限信息传输速率＝W$\log_2$(1+S/N) bps

W：信道的带宽

S：信道内所传信号的平均功率

N：信道内部的高斯噪声功率

仙农公式表明：只要信息传输速率低于信道的极限信息传输速率，就一定可以找到某种办法来实现无差错的传输。

例：信噪比为 30dB，带宽为 3kHz 的信道的最大数据传输速率为：

10 $\log_{10}$ S/N=30 dB，则 S/N=$10^{30/10}$

3k* $\log_2$(1+$10^{30/10}$)=3k* $\log_2$(1+1000)≈30k bps

2．3 传输媒体

1．导向传输媒体（有线传输）

- 双绞线（ＳＴＰ：屏蔽双绞； ＵＴＰ：无屏蔽的双绞线）

采用绞起来的结构是为了减少对相邻的导线的电磁干扰

UTP 双绞线的等级

类别	描述	最大传输速率
第一类	铜线没有缠绕，只能传送声音，不适合在 LAN 中传送数据	/
第二类	无缠绕，但可用于低速数据传输，不适合 LAN	4Mbps
第三类	铜线缠绕绞距为每分米缠绕 1 次，是架设 10BaseT 网络最基本的线材需求等级	10Mbps
第四类	缠绕较密	16Mbps
第五类	铜线缠绕最紧密，架设高速以太网络一定要使用此等级的线材	100Mbps

国标 T568A 和 T568B 定义的管脚排列

管脚排列	T568A 定义的色线位置	T568B 定义的色线位置
1	绿白 W-G	橙白 W-O
2	绿 G	橙 O
3	橙白 W-O	绿白 W-G
4	蓝 BL	蓝 BL
5	蓝白 W-BL	蓝白 W-BL
6	橙 O	绿 G
7	棕白 W-BR	棕白 W-BR
8	棕 BR	棕 BR

注：RJ-45 水晶头的引脚顺序为：将水晶头水平放置，有塑料弹片的一面向下，插入网卡的一头向右，则靠近自己的那只引脚即为“8”。

- 同轴电缆

（1）50 同轴电缆；基带同轴电缆（数字通信）

同轴电缆特性

	直径	最大传输距离	阻抗
RG-58（细缆）	0.26 厘米	185 米	50 欧姆
RG-11（粗缆）	1.27 厘米	500 米	50 欧姆

A、曼彻斯特编码（参见课本 P28 图 2-3）

将每一个码元分成两个相等的间隔，码元“1”在前一间隔为高电平而后一间隔为低电平，码元 0 正好相反。

B、差分曼彻斯特编码

若码元为 1，则其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；若码元为 0，则其前半个码元的电平与上一个码元的后半个码元的电平相反。

（2）75 同轴电缆：模拟通信、宽带同轴电缆

- 光缆

（1）光缆的优点：传输频带宽，通信容量大；适宜于远离传输；抗雷电和电磁干扰性好；保密性好；体积小，重量轻。

（2）光缆的分类：

多模光纤：多条不同入射角度（大于临界角度）的光线在一条光纤中传输，此光纤称为多模光纤。

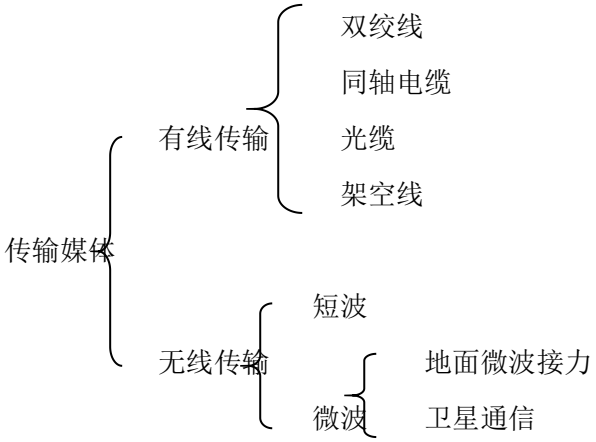
单模光纤：若光纤的直径减小到只有一个光的波长，则它可使光线直线传播，这种光纤称为单模光纤。

- 架空明线

2．非导向传输媒体（无线传输）

- 无线传输

包括短波通信和微波通信（包括地面微波接力通信卫星通信）。



3．数据传输的总时延：

- 传播时延：信号在信道中传播所需的时间，取决于信号在信道上的传播速率以及所传播的距离。
- 发送时延：发送数据所需要的时间，取决于数据块的长度和数据在信道上的发送速率。
- 重发时延：因为数据出错，需要重传，直到传送正确为止。

2．4 模拟传输与数字传输

2．4．1 模拟传输系统

1．调制解调器（数字数据采用模拟信号传输时）

- 调制解调器的作用（modem）

调制器：将基带数字信号的波形变换适合模拟信道传输的波形。

解调器：将经过调制器变换过的模拟信号恢复或原来的数字信号。

2．调制方法：

- 调幅ＡＭ：载波的振幅随基带数字信号而变化。
- 调频ＦＭ：载波的频率随基带数字信号而变化。
- 调相ＰＭ：载波的起始相位随基带数字信号而变化。

2．4．2 数字传输系统

1． 编码解码器（模拟数据采用数字信号传输时）

编码器：将模拟数据转换数字信号。

解码器：将数字信号还原成模拟数据。

2． 脉码调制ＰＣＭ（数字化过程：采样、量化和编码）。

注：数字信号的特点：

（1）抗噪声（干扰）能力强。

（2）可以控制差错，提高了传输质量。

（3）便于用计算机进行处理。

（4）易于加密，保密性强。

（5）可以传输语音，数据，影像。通用，灵活。

第三章 数据链路层

❒ 课时安排：

5 学时

❒ 教学目的：

1. 数据链路和帧
2. 点对点协议 PPP
3. CSMA/CD 协议
4. 以太网
5. 高速以太网

❒ 教学重、难点：

1. 点对点协议 PPP，是链路层的重要协议，也是基本的链路层协议，在此之上构建了最出名的以太网局域网类型。
2. 以太网协议只是局域网类型中最典型的类型，是局域网类型的其中一种类型，替代不了局域网。这两者必须区分清楚。
3. 以太网由于其使用范围和应用范围的广泛性，故此，事实的应用中，人们通常将以太网代称局域网。并将其标准广泛化了。
4. 局域网常见的几种组网模式：星型、总线型、环型。
5. CSMA/CD 协议：载波冲突检测协议，是现在几乎所有的网络都需要用到的检测协议，其中包含了一个数据帧如何在网络中正确发送，发送时延的决定因素、网络快慢的决定因素都在其中，必须深度掌握。是本章的重点。

❒ 教学内容：

3．1 数据链路层

3．1．1 数据链路层使用的信道主要有以下两种类型：

- **点对点信道。**
  - 这种信道使用一对一的点对点通信方式。
- **广播信道。**
  - 这种信道使用一对多的广播通信方式，因此过程比较复杂。
  - 广播信道上连接的主机很多，因此必须使用专用的共享信道协议来协调这些主机的数据发
- 数据链路层的简单模型

3．1．2 使用点对点信道的数据链路层

- **数据链路和帧**
  - **链路(link)**
    - ◊是一条无源的点到点的物理线路段，中间没有任何其他的交换结点。
    - ◊一条链路只是一条通路的一个组成部分。
  - **数据链路(data link)**
    - ◊除了物理线路外，还必须有通信协议来控制这些数据的传输。
    - ◊若把实现这些协议的硬件和软件加到链路上，就构成了数据链路。
    - ◊现在最常用的方法是使用适配器（即网卡）来实现这些协议的硬件和软件。
    - ◊一般的适配器都包括了数据链路层和物理层这两层的功能。
- **数据链路层像个数字管道**
  - 常常在两个对等的数据链路层之间画出一个数字管道，而在这条数字管道上传输的数据单位是帧。
  - 早期的数据通信协议曾叫作通信规程(procedure)。
  - 因此在数据链路层，规程和协议是同义语。

3．1．3 三个基本问题

- **(1) 封装成帧**
- **(2) 透明传输**
- **(3) 差错控制**
- **封装成帧**
  - **封装成帧(framing)**
    - ◊就是在一段数据的前后分别添加首部和尾部，然后就构成了一个帧。
  - **首部和尾部的一个重要作用就是进行帧定界。**
    - ◊确定帧的界限。
- **透明传输**
  - 发送端的数据链路层在数据中出现控制字符“SOH”或“EOT”的前面插入一个转义字符“ESC”(其十六进制编码是 1B)。
  - 字节填充(byte stuffing)或字符填充(character stuffing)
  - 接收端的数据链路层在将数据送往网络层之前删除插入的转义字符。

- 如果转义字符也出现数据当中，那么应在转义字符前面插入一个转义字符。
- 当接收端收到连续的两个转义字符时，就删除其中前面的一个。

- **差错检测**
  - 在传输过程中可能会产生比特差错：
  - 1 可能会变成 0
  - 而 0 也可能变成 1。
  - 在一段时间内，传输错误的比特占所传输比特总数的比率称为误码率 BER (Bit Error Rate)。
  - 误码率与信噪比有很大的关系。
  - 为了保证数据传输的可靠性，在计算机网络传输数据时，必须采用各种差错检测措施。
  - **循环冗余检验的原理**
    - ◊在数据链路层传送的帧中，广泛使用了循环冗余检验 CRC 的检错技术。
    - ◊在发送端，先把数据划分为组。
    - ◊假定每组 k 个比特。
    - ◊假设待传送的一组数据 M = 101001（现在 k = 6）。我们在 M 的后面再添加供差错检测用的 n 位冗余码一起发送。
  - **帧检验序列 FCS**
    - ◊在数据后面添加上的冗余码称为帧检验序列 FCS (Frame Check Sequence)。
    - ◊循环冗余检验 CRC 和帧检验序列 FCS 并不等同。
    - ◊CRC 是一种常用的检错方法，而 FCS 是添加在数据后面的冗余码。
    - ◊FCS 可以用 CRC 这种方法得出，但 CRC 并非用来获得 FCS 的唯一方法。
    - ◊仅用循环冗余检验 CRC 差错检测技术只能做到无差错接受 (accept)。
    - ◊“无差错接受”是指：“凡是接受的帧（即不包括丢弃的帧），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”。
    - ◊也就是说：“凡是接收端数据链路层接受的帧都没有传输差错”

（有差错的帧就丢弃而不接受）。

◊要做到“可靠传输”（即发送什么就收到什么）就必须再加上确认和重传机制。

3．2 点对点协议 PPP

3．2．1 PPP 协议的特点

- **现在全世界使用得最多的数据链路层协议是点对点协议 PPP (Point-to-Point Protocol)。**
  - 用户使用拨号电话线接入因特网时，一般都是使用 PPP 协议。
- **协议应满足的需求**
  - 简单——这是首要的要求
  - 封装成帧
  - 透明性
  - 多种网络层协议
  - 多种类型链路
  - 差错检测
  - 检测连接状态
  - 最大传送单元
  - 网络层地址协商
  - 数据压缩协商
- **PPP 协议不需要的功能**
  - 纠错
  - 流量控制
  - 序号
  - 多点线路
  - 半双工或单工链路
- **PPP 协议的组成**
  - 1992 年制订了 PPP 协议。经过 1993 年和 1994 年的修订，现在的 PPP 协议已成为因特网的正式标准[RFC 1661]。
  - PPP 协议有三个组成部分
    - ◊一个将 IP 数据报封装到串行链路的方法。
    - ◊链路控制协议 LCP (Link Control Protocol)。
    - ◊网络控制协议 NCP (Network Control Protocol)。

3．2．2 PPP 协议的帧格式

- PPP 的首部和尾部分别为 4 个字段和 2 个字段
- 标志字段 F = 0x7E
- 十六进制的 7E 的二进制表示是 01111110
- 地址字段 A 只置为 0xFF。
- 地址字段实际上并不起作用。
- 控制字段 C 通常置为 0x03。
- PPP 是面向字节的，所有的 PPP 帧的长度都是整数字节。
- PPP 有一个 2 个字节的协议字段。
- 当协议字段为 0x0021 时，PPP 帧的信息字段就是 IP 数据报。
- 若为 0xC021, 则信息字段是 PPP 链路控制数据。
- 若为 0x8021，则表示这是网络控制数据。

- **透明传输问题**
  - 当 PPP 用在同步传输链路时，协议规定采用硬件来完成比特填充（和 HDLC 的做法一样）。
  - 当 PPP 用在异步传输时，就使用一种特殊的字符填充法。
  - **字符填充**
    - ◊将信息字段中出现的每一个 0x7E 字节转变成为 2 字节序列 (0x7D, 0x5E)。
    - ◊若信息字段中出现一个 0x7D 的字节, 则将其转变成为 2 字节序列(0x7D, 0x5D)。
    - ◊若信息字段中出现 ASCII 码的控制字符（即数值小于 0x20 的字符），则在该字符前面要加入一个 0x7D 字节，同时将该字符的编码加以改变。
  - **零比特填充**
    - ◊**P**PP 协议用在 SONET/SDH 链路时，是使用同步传输（一连串的比特连续传送）。这时 PPP 协议采用零比特填充方法来实现透明传输。
    - ◊在发送端，只要发现有 5 个连续 1，则立即填入一个 0。接收端对帧中的比特流进行扫描。每当发现 5 个连续 1 时，就把这 5 个连续 1 后的一个 0 删除，
- **不提供使用序号和确认的可靠传输**
  - PPP 协议之所以不使用序号和确认机制是出于以下的考虑：
  - 在数据链路层出现差错的概率不大时，使用比较简单的 PPP 协

议较为合理。

- 在因特网环境下，PPP 的信息字段放入的数据是 IP 数据报。数据链路层的可靠传输并不能够保证网络层的传输也是可靠的。
- 帧检验序列 FCS 字段可保证无差错接受。

3．2．3 PPP 协议的工作状态

- 当用户拨号接入 ISP 时，路由器的调制解调器对拨号做出确认，并建立一条物理连接。
- PC 机向路由器发送一系列的 LCP 分组（封装成多个 PPP 帧）。
- 这些分组及其响应选择一些 PPP 参数，和进行网络层配置，NCP 给新接入的 PC 机分配一个临时的 IP 地址，使 PC 机成为因特网上的一个主机。
- 通信完毕时，NCP 释放网络层连接，收回原来分配出去的 IP 地址。接着，LCP 释放数据链路层连接。最后释放的是物理层的连接。

3．3 使用广播信道的数据链路层

3．3．1 局域网的数据链路层

- **局域网最主要的特点是：**
  - 网络为一个单位所拥有，且地理范围和站点数目均有限。
- **局域网具有如下的一些主要优点：**
  - 具有广播功能
    - ◊从一个站点可很方便地访问全网。局域网上的主机可共享连接在局域网上的各种硬件和软件资源。
  - 便于系统的扩展和逐渐地演变，各设备的位置可灵活调整和改变。
    - ◊提高了系统的可靠性、可用性和残存性。
- **共享信道——如何共享通信媒体资源 的问题**
  - **静态划分信道**
    - ◊频分复用
    - ◊时分复用
    - ◊波分复用
    - ◊码分复用
  - **动态媒体接入控制（多点接入）**
    - ◊随机接入(重点{)
    - ◊受控接入

如多点线路探询(polling)，或轮询。

- **以太网的两个标准**
  - DIX Ethernet V2 是世界上第一个局域网产品（以太网）的规约。
  - IEEE 的 802.3 标准。
    - ◊DIX Ethernet V2 标准与 IEEE 的 802.3 标准只有很小的差别，因此可以将 802.3 局域网简称为“以太网”。
    - ◊严格说来，“以太网”应当是指符合 DIX Ethernet V2 标准的局域网
- **数据链路层的两个子层**
  - 为了使数据链路层能更好地适应多种局域网标准，802 委员会就将局域网的数据链路层拆成两个子层：
    - ◊**逻辑链路控制 LLC (Logical Link Control)子层**
    - ◊**媒体接入控制 MAC (Medium Access Control)子层。**

    与接入到传输媒体有关的内容都放在 MAC 子层，而 LLC 子层则与传输媒体无关，不管采用何种协议的局域网对 LLC 子层来说都是透明的

    由于 TCP/IP 体系经常使用的局域网是 DIX Ethernet V2 而不是 802.3 标准中的几种局域网，因此现在 802 委员会制定的逻辑链路控制子层 LLC（即 802.2 标准）的作用已经不大了。

    很多厂商生产的适配器上就仅装有 MAC 协议而没有 LLC 协议。
- **适配器的作用**
  - 网络接口板又称为通信适配器(adapter)或网络接口卡 NIC (Network Interface Card)，或“网卡”。
  - 适配器的重要功能：
    - ◊进行串行/并行转换。
    - ◊对数据进行缓存。
    - ◊在计算机的操作系统安装设备驱动程序。
    - ◊实现以太网协议。

3．3．2 CSMA/CD 协议

最初的以太网是将许多计算机都连接到一根总线上。当初认为这样的连接方法既简单又可靠，因为总线上没有有源器件。

- **以太网的广播方式发送**

- 工作过程：
  - ◊总线上的每一个工作的计算机都能检测到 B 发送的数据信号。
  - ◊由于只有计算机 D 的地址与数据帧首部写入的地址一致，因此只有 D 才接收这个数据帧。
  - ◊其他所有的计算机（A, C 和 E）都检测到不是发送给它们的数据帧，因此就丢弃这个数据帧而不能够收下来。
  - ◊具有广播特性的总线上实现了一对一的通信。

- **为了通信的简便以太网采取了两种重要的措施**
  - 采用较为灵活的无连接的工作方式
    - ◊即不必先建立连接就可以直接发送数据。
  - 以太网对发送的数据帧不进行编号，也不要求对方发回确认。
    - ◊这样做的理由是局域网信道的质量很好，因信道质量产生差错的概率是很小的。

- **以太网提供的服务是不可靠的交付，即尽最大努力的交付。**
  - 当目的站收到有差错的数据帧时就丢弃此帧，其他什么也不做。
  - 差错的纠正由高层来决定。
  - 如果高层发现丢失了一些数据而进行重传，但以太网并不知道这是一个重传的帧，而是当作一个新的数据帧来发送。

- **载波监听多点接入/碰撞检测 CSMA/CD**
  - **CSMA/CD ：**
  - **Carrier Sense Multiple Access with Collision Detection**
    - ◊“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。
    - ◊“载波监听”是指每一个站在发送数据之前先要检测一下总线上是否有其他计算机在发送数据，如果有，则暂时不要发送数据，以免发生碰撞。
    - ◊总线上并没有什么“载波”。因此， “载波监听”就是用电子技术检测总线上有没有其他计算机发送的数据信号。
    - ◊“碰撞检测”就是计算机边发送数据边检测信道上的信号电压大小。
    - ◊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的门限值时，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。

◊所谓“碰撞”就是发生了冲突。因此“碰撞检测”也称为“冲突检测”。

- **检测到碰撞后**

◊在发生碰撞时，总线上传输的信号产生了严重的失真，无法从中恢复出有用的信息来。

◊每一个正在发送数据的站，一旦发现总线上出现了碰撞，就要立即停止发送，免得继续浪费网络资源，然后等待一段随机时间后再次发送。

- **电磁波在总线上的有限传播速率的影响**

当某个站监听到总线是空闲时，也可能总线并非真正是空闲的。

◊A 向 B 发出的信息，要经过一定的时间后才能传送到 B。

◊B 若在 A 发送的信息到达 B 之前发送自己的帧(因为这时 B 的载波监听检测不到 A 所发送的信息)，则必然要在某个时间和 A 发送的帧发生碰撞。

◊碰撞的结果是两个帧都变得无用。

- 重要特性

◊使用 CSMA/CD 协议的以太网不能进行全双工通信而只能进行双向交替通信（半双工通信）。

◊每个站在发送数据之后的一小段时间内，存在着遭遇碰撞的可能性。

◊这种发送的不确定性使整个以太网的平均通信量远小于以太网的最高数据率。

- 争用期

◊最先发送数据帧的站，在发送数据帧后至多经过时间 2τ（两倍的端到端往返时延）就可知道发送的数据帧是否遭受了碰撞。

◊以太网的端到端往返时延 2τ 称为争用期，或碰撞窗口。

◊经过争用期这段时间还没有检测到碰撞，才能肯定这次发送不会发生碰撞。

- **二进制指数类型退避算法 (truncated binary exponential type)**

◊发生碰撞的站在停止发送数据后，要推迟（退避）一个随机时间才能再发送数据。

  - 基本退避时间取为争用期 2τ。

- 从整数集合[0,1,…, (2k –1)]中随机地取出一个数，记为 r。重传所需的时延就是 r 倍的基本退避时间。
- 参数 k 按下面的公式计算：

$$k = \text{Min}[\text{重传次数}, 10]$$

- 当 $k \le 10$ 时，参数 k 等于重传次数。
- 当重传达 16 次仍不能成功时即丢弃该帧，并向高层报告。

◊以太网取 **51.2 μs** 为争用期的长度。

◊对于 10 Mb/s 以太网，在争用期内可发送 512 bit，即 64 字节。

◊以太网在发送数据时，若前 64 字节没有发生冲突，则后续的数据就不会发生冲突。

♦ **最短有效帧长**

◊如果发生冲突，就一定是在发送的前 64 字节之内。

◊由于一检测到冲突就立即中止发送，这时已经发送出去的数据一定小于 64 字节。

◊以太网规定了最短有效帧长为 64 字节，凡长度小于 64 字节的帧都是由于冲突而异常中止的无效帧。

♦ **强化碰撞**

◊当发送数据的站一旦发现发生了碰撞时：

- 立即停止发送数据；
- 再继续发送若干比特的人为干扰信号(jamming signal)，以便让所有用户都知道现在已经发生了碰撞。

3．4 使用广播信道的以太网

3．4．1 使用集线器的星形拓扑

- 传统以太网最初是使用粗同轴电缆，后来演进到使用比较便宜的细同轴电缆，最后发展为使用更便宜和更灵活的双绞线。
- 这种以太网采用星形拓扑，在星形的中心则增加了一种可靠性非常高的设备，叫做集线器(hub)
- 星形网 10BASE-T
  - ♦ 不用电缆而使用无屏蔽双绞线。

    ◊每个站需要用两对双绞线:

    ◊分别用于发送和接收。
  - ♦ 集线器使用了大规模集成电路芯片，因此这样的硬件设备的可靠

性已大大提高了。

- 以太网在局域网中的统治地位
  - 10BASE-T 的通信距离稍短，每个站到集线器的距离不超过 100 m。
  - 这种 10 Mb/s 速率的无屏蔽双绞线星形网的出现，既降低了成本，又提高了可靠性。
  - 10BASE-T 双绞线以太网的出现，是局域网发展史上的一个非常重要的里程碑，它为以太网在局域网中的统治地位奠定了牢固的基础。
- **集线器的一些特点**
  - 集线器是使用电子器件来模拟实际电缆线的工作，因此整个系统仍然像一个传统的以太网那样运行。
  - 使用集线器的以太网在逻辑上仍是一个总线网，各工作站使用的还是 CSMA/CD 协议，并共享逻辑上的总线。
  - 集线器很像一个多接口的转发器，工作在物理层。

3．4．2 以太网的信道利用率

- 以太网的信道被占用的情况：
- 争用期长度为 2τ，即端到端传播时延的两倍。检测到碰撞后不发送干扰信号。
- 帧长为 L (bit)，数据发送速率为 C (b/s)，因而帧的发送时间为 $L/C = T0$ (s)。
- 一个帧从开始发送，经可能发生的碰撞后，将再重传数次，到发送成功且信道转为空闲(即再经过时间 τ 使得信道上无信号在传播)时为止，是发送一帧所需的平均时间。
- 当数据率一定时，以太网的连线的长度受到限制，否则 τ 的数值会太大。
- 以太网的帧长不能太短，否则 $T0$ 的值会太小，使 a 值太大。

3．4．3 以太网的 MAC 层

- MAC 层的硬件地址
  - 在局域网中，硬件地址又称为物理地址，或 MAC 地址。
  - 802 标准所说的“地址”严格地讲应当是每一个站的“名字”或标识符。
    - ◊但鉴于大家都早已习惯了将这种 48 位的“名字”称为“地址”

◊本书也采用这种习惯用法，尽管这种说法并不太严格。

- 48 位的 MAC 地址

◊IEEE 的注册管理机构 RA 负责向厂家分配地址字段的前三个字节(即高位 24 位)。

◊地址字段中的后三个字节(即低位 24 位)由厂家自行指派，称为扩展标识符，必须保证生产出的适配器没有重复地址。

◊一个地址块可以生成 224 个不同的地址。

◊这种 48 位地址称为 MAC-48，它的通用名称是 EUI-48。

◊“MAC 地址”实际上就是适配器地址或适配器标识符 EUI-48。

- 适配器检查 MAC 地址

◊适配器从网络上每收到一个 MAC 帧就首先用硬件检查 MAC 帧中的 MAC 地址.

如果是发往本站的帧则收下，然后再进行其他的处理。

否则就将此帧丢弃，不再进行其他的处理。

◊“发往本站的帧”包括以下三种帧：

- 单播(unicast)帧（一对一）
- 广播(broadcast)帧（一对全体）
- 多播(multicast)帧（一对多）

- MAC 帧的格式

◊常用的以太网 MAC 帧格式有两种标准 ：

- DIX Ethernet V2 标准
- IEEE 的 802.3 标准

◊最常用的 MAC 帧是以太网 V2 的格式。

◊无效的 MAC 帧

- 数据字段的长度与长度字段的值不一致；
- 帧的长度不是整数个字节；
- 用收到的帧检验序列 FCS 查出有差错；
- 数据字段的长度不在 46～1500 字节之间。
- 有效的 MAC 帧长度为 64～1518 字节之间。
  - 对于检查出的无效 MAC 帧就简单地丢弃。
  - 以太网不负责重传丢弃的帧。

◊帧间最小间隔

- 帧间最小间隔为 9.6 μs，相当于 96 bit 的发送时间。

- 一个站在检测到总线开始空闲后，还要等待 9.6 μs 才能再次发送数据。
- 这样做是为了使刚刚收到数据帧的站的接收缓存来得及清理，做好接收下一帧的准备。

3．5 扩展的局域网

3．5．1 在物理层扩展局域网

3．5．2 在数据链路层扩展局域网

第四章 网络层

❒ 课时安排：

12 学时

❒ 教学目的：

1. 认识什么是 IP 网络
2. IP 的作用
3. 路由器及其功能
4. 四类 IP 划分
5. 掌握子网划分
6. 掌握 CIDR 划分
7. 熟悉 ARP\IP\ICMP 协议，运用协议功能解决实际问题

❒ 教学重、难点：

4 IP 及路由器是网络层需要掌握的核心内容，也是 TCP/IP 协议的重要核心协议之一。

1. IP 的划分方法关系到对子网和主机的理解，关系网络中主机搜索。
2. 子网的划分必须依据网络的规模进行，如何合理地利用地理位置和其他因素，制约了子网划分方法的选择
3. CIDR 划分法是目前最切合实际的选择，掌握它，必须理解好 IP 的划分法和路由器的选择

5 路由协议

4. 三个重要的协议工作在网络层
5. ARP 决定了如何在一个局域网中寻找对应的主机
6. IP 决定了在互联网中如何查找一个网络
7. ICMP 决定了如何在实际工作中解决网络难题

三大协议，相辅相成，对于实际网络的分析和设计起着决定性作用。必须重点掌握。

❒ 教学内容：

4.1 网络层提供的两种服务

4.2.5 电信网的成功经验--让网络负责可靠交付

- 面向连接的通信方式
- 建立虚电路(Virtual Circuit)，以保证双方通信所需的一切网络资源。
- 如果再使用可靠传输的网络协议，就可使所发送的**分组无差错按序**到达终点。
- 虚电路是逻辑连接
  - 虚电网络层向上只提供简单灵活的、无连接的、尽最大努力交付的

数据报服务。

- 网络在发送分组时不需要先建立连接。
- 每一个分组（即 IP 数据报）独立发送，与其前后的分组无关（不进行编号）。
- 网络层不提供服务质量的承诺。
- 即所传送的分组可能出错、丢失、重复和失序（不按序到达终点）
- 当然也不保证分组传送的时限。
- 路表示这只是一条逻辑上的连接
  - 分组都沿着这条逻辑连接按照存储转发方式传送，
  - 而并不是真正建立了一条物理连接。
- 请注意，
  - 电路交换的电话通信是先建立了一条真正的连接。
  - 因此分组交换的虚连接和电路交换的连接只是类似，但并不完全一样。

4.2.6 因特网采用的设计思路

- 网络层向上只提供简单灵活的、无连接的、尽最大努力交付的数据报服务。
  - 网络在发送分组时不需要先建立连接。
  - 每一个分组（即 IP 数据报）独立发送，与其前后的分组无关（不进行编号）。
  - 网络层不提供服务质量的承诺。
  - 即所传送的分组可能出错、丢失、重复和失序（不按序到达终点）
  - 当然也不保证分组传送的时限。
- 尽最大努力交付的好处
  - 由于传输网络不提供端到端的可靠传输服务，这就使网络中的路由器可以做得比较简单，而且价格低廉（与电信网的交换机相比较）。
  - 如果主机（即端系统）中的进程之间的通信需要是可靠的，那么就由网络的主机中的运输层负责（包括差错处理、流量控制等）。
  - 采用这种设计思路的好处是：
  - 网络的造价大大降低，运行方式灵活，能够适应多种应用。
  - 因特网能够发展到今日的规模，充分证明了当初采用这种设计思路的正确性。

4.2 网际协议 IP

4.2.5 网际协议 IP 是 TCP/IP 体系中两个最主要的协议之一。与 IP 协议配套使用的还有三个协议：

- 地址解析协议 **ARP**
  - (Address Resolution Protocol)
- 网际控制报文协议 **ICMP**
  - (Internet Control Message Protocol)
- 网际组管理协议 **IGMP**
  - (Internet Group Management Protocol

4.2.6 虚拟互连网络

- 互连在一起的网络要进行通信，会遇到许多问题需要解决，如：
  - 不同的寻址方案
  - 不同的最大分组长度
  - 不同的网络接入机制
  - 不同的超时控制
  - 不同的差错恢复方法
  - 不同的状态报告方法
  - 不同的路由选择技术
  - 不同的用户接入控制
  - 不同的服务（面向连接服务和无连接服务）
  - 不同的管理与控制方式
- 中间设备又称为中间系统或中继(relay)系统。
  - 物理层中继系统：
  - 转发器(repeater)。
  - 数据链路层中继系统：
  - 网桥或桥接器(bridge)。
  - 网络层中继系统：
  - 路由器(router)。
  - 网桥和路由器的混合物：
  - 桥路器(brouter)。
  - 网络层以上的中继系统：
  - 网关(gateway)。
- 当中继系统是转发器或网桥时，一般并不称之为网络互连，因为这仅仅是把一个网络扩大了，而这仍然是一个网络。
  - 互联网都是指用路由器进行互连的网络。
  - 网关由于比较复杂，目前使用得较少。
  - 由于历史的原因，许多有关 TCP/IP 的文献将网络层使用的路由器称为网关。
- 所谓虚拟互连网络也就是**逻辑互连网络**
  - 它的意思就是互连起来的各种物理网络的异构性本来是客观存在的，但是我们利用 IP 协议就可以使这些性能各异的网络从用户看起来好像是一个统一的网络。
  - 使用 IP 协议的虚拟互连网络可简称为 **IP 网**。
  - 使用虚拟互连网络的好处是：
  - 当互联网上的主机进行通信时，就好像在一个网络上通信一样，而看不见互连的各具体的网络异构细节。
- 从网络层看 IP 数据报的传送
- 如果我们只从网络层考虑问题，那么 IP 数据报就可以想象是在**网络层**中传送。

4.2.7 分类的 IP 地址

- 我们把整个因特网看成为一个单一的、抽象的网络。
  - IP 地址就是给每个连接在因特网上的主机（或路由器）分配一个在全世界范围是唯一的 32 位的标识符。

♦ IP 地址现在由因特网名字与号码指派公司 **ICANN** (Internet Corporation for Assigned Names and Numbers)进行分配

▷ 分类的 IP 地址。

♦ 这是最基本的编址方法

♦ 在 1981 年就通过了相应的标准协议。

▷ 子网的划分。

♦ 这是对最基本的编址方法的改进

♦ 其标准[RFC 950]在 1985 年通过。

▷ 构成超网。

♦ 这是比较新的无分类编址方法。

♦ 1993 年提出后很快就得到推广应用。

▷ 每一类地址都由两个固定长度的字段组成:

♦ 其中一个字段是**网络号 net-id**：它标志主机（或路由器）所连接到的网络，

♦ 而另一个字段则是**主机号 host-id**：它标志该主机（或路由器）。

♦ 两级的 IP 地址可以记为：

♦ **IP 地址 ::= { <网络号>, <主机号>}** (4-1)

▷ IP 地址的一些重要特点

♦ IP 地址是一种分等级的地址结构。

◆ 分两个等级的好处是：

◆ 第一，IP 地址管理机构在分配 IP 地址时只分配网络号，而剩下的主机号则由得到该网络号的单位自行分配。这样就方便了 IP 地址的管理。

◆ 第二，路由器仅根据目的主机所连接的网络号来转发分组（而不考虑目的主机号），这样就可以使路由表中的项目数大幅度减少，从而减小了路由表所占的存储空间。

♦ 实际上 IP 地址是标志一个主机（或路由器）和一条链路的接口。

◆ 当一个主机同时连接到两个网络上时，该主机就必须同时具有两个相应的 IP 地址，其网络号 net-id 必须是不同的。这种主机称为**多归属主机**(multihomed host)。

◆ 由于一个路由器至少应当连接到两个网络（这样它才能将 IP 数据报从一个网络转发到另一个网络），因此一个路由器至少应当有两个不同的 IP 地址。

♦ 用转发器或网桥连接起来的若干个局域网仍为一个网络，因此这些局域网都具有同样的网络号 net-id。

♦ 所有分配到网络号 net-id 的网络，范围很小的局域网，还是可能覆盖很大地理范围的广域网，都是平等的。

4.2.8 地址解析协议 ARP

▷ **管网络层使用的是什么协议，在实际网络的链路上传送数据帧时，最终还是必须使用硬件地址。**

♦ 每一个主机都设有一个 ARP 高速缓存(ARP cache)，里面有<u>所在的局域网上</u>的各主机和路由器的 IP 地址到硬件地址的映射表。

♦ 当主机 A 欲向本局域网上的某个主机 B 发送 IP 数据报时:

♦ 就先在其 ARP 高速缓存中查看有无主机 B 的 IP 地址。

- 如有，就可查出其对应的硬件地址，
- 再将此硬件地址写入 MAC 帧，然后通过局域网将该 MAC 帧发往此硬件地址。

▷ ARP 高速缓存的作用

- 为了减少网络上的通信量，
  - 主机 A 在发送其 ARP 请求分组时，就将自己的 IP 地址到硬件地址的映射写入 ARP 请求分组。
- 当主机 B 收到 A 的 ARP 请求分组时，
  - 就将主机 A 的这一地址映射写入主机 B 自己的 ARP 高速缓存中。这对主机 B 以后向 A 发送数据报时就更方便了。

▷ ARP 是解决**同一个局域网上**的主机或路由器的 IP 地址和硬件地址的映射问题。

- 如果所要找的主机和源主机不在同一个局域网上，
- 那么就要通过 ARP 找到一个位于本局域网上的某个路由器的硬件地址，
- 然后把分组发送给这个路由器，让这个路由器把分组转发给下一个网络。
- 剩下的工作就由下一个网络来做。
- 从 IP 地址到硬件地址的解析是自动进行的，主机的用户对这种地址解析过程是不知道的。
- 只要主机或路由器要和本网络上的另一个已知 IP 地址的主机或路由器进行通信，ARP 协议就会自动地将该 IP 地址解析为链路层所需要的硬件地址。

▷ 使用 ARP 的四种典型情况

- 发送方是主机，要把 IP 数据报发送到本网络上另一个主机
  - 这时用 ARP 找到目的主机的硬件地址。
- 发送方是主机，要把 IP 数据报发送到另一个网络上的一个主机。
  - 这时用 ARP 找到本网络上的一个路由器的硬件地址。
  - 剩下的工作由这个路由器来完成。
- 发送方是路由器，要把 IP 数据报转发到本网络上的一个主机。
  - 这时用 ARP 找到目的主机的硬件地址。
- 发送方是路由器，要把 IP 数据报转发到另一个网络上的一个主机。
  - 这时用 ARP 找到本网络上另一个路由器的硬件地址。
  - 剩下的工作由这个路由器来完成。

4.2.9 IP 数据报的格式

▷ 一个 IP 数据报由首部和数据两部分组成。

- 首部的前一部分是固定长度，共 20 字节，是所有 IP 数据报必须具有的。
- 在首部的固定部分的后面是一些可选字段，其长度是可变的。

▷ IP 层转发分组的流程

- 有四个 A 类网络通过三个路由器连接在一起。每一个网络上都可能有成千上万个主机。

--可以想像，若按目的主机号来制作路由表，则所得出的路由表就会过于庞大。

--但若按主机所在的网络地址来制作路由表，那么每一个路由器中的路由表就只包含 4 个项目。

--这样就可使路由表大大简化。

▷ 查找路由表

♦ 根据目的网络地址就能确定下一跳路由器，这样做的结果是：

--IP 数据报最终一定可以找到目的主机所在目的网络上的路由器（可能要通过多次的间接交付）。

--只有到达最后一个路由器时，才试图向目的主机进行直接交付。

♦

▷ 特定主机路由

♦ 这种路由是为特定的目的主机指明一个路由。

♦ 采用特定主机路由可使网络管理人员能更方便地控制网络和测试网络，同时也可在需要考虑某种安全问题时采用这种特定主机路由。

▷ 默认路由(default route)

♦ 路由器还可采用默认路由以减少路由表所占用的空间和搜索路由表所用的时间。

--这种转发方式在一个网络只有很少的对外连接时是很有用的。

--默认路由在主机发送 IP 数据报时往往更能显示出它的好处。

--如果一个主机连接在一个小网络上，而这个网络只用一个路由器和因特网连接，那么在这种情况下使用默认路由是非常合适的。

▷ 分组转发算法

♦ (1) 从数据报的首部提取目的主机的 IP 地址 *D*, 得出目的网络地址为 *N*。

♦ (2) 若网络 *N* 与此路由器直接相连，则把数据报直接交付目的主机 *D*；否则是间接交付，执行(3)。

♦ (3) 若路由表中有目的地址为 *D* 的特定主机路由，则把数据报传送给路由表中所指明的下一跳路由器；否则，执行(4)。

♦ (4) 若路由表中有到达网络 *N* 的路由，则把数据报传送给路由表指明的下一跳路由器；否则，执行(5)。

♦ (5) 若路由表中有一个默认路由，则把数据报传送给路由表中所指明的默认路由器；否则，执行(6)。

♦ (6) 报告转发分组出错。

4.3 划分子网 和 构造超网

4.3.1 划分子网

▷ 1. 从两级 IP 地址到三级 IP 地址

♦ 在 ARPANET 的早期，IP 地址的设计确实不够合理。

♦ IP 地址空间的利用率有时很低。

♦ 给每一个物理网络分配一个网络号会使路由表变得太大因而使网络性能变坏。

♦ 两级的 IP 地址不够灵活。

♦ 从 1985 年起在 IP 地址中又增加了一个“子网号字段”，使两级的 IP 地址变成为三级的 IP 地址。

♦ 这种做法叫作划分子网(subnetting) 。划分子网已成为因特网的正式

标准协议。

▷ 划分子网的基本思路

♦ 划分子网纯属一个单位内部的事情。单位对外仍然表现为没有划分子网的网络。

♦ 从主机号借用若干个位作为**子网号** subnet-id，而主机号 host-id 也就相应减少了若干个位。

♦ IP 地址 ::= {<网络号>, <子网号>, <主机号>}

♦ 凡是从其他网络发送给本单位某个主机的 IP 数据报，仍然是根据 IP 数据报的目的网络号 net-id，先找到连接在本单位网络上的路由器。

♦ 然后此路由器在收到 IP 数据报后，再按目的网络号 net-id 和子网号 subnet-id 找到目的子网。

♦ 最后就将 IP 数据报直接交付目的主机。

♦ 划分子网后变成了三级结构

--当没有划分子网时，IP 地址是两级结构。

--划分子网后 IP 地址就变成了三级结构。

--划分子网只是把 IP 地址的主机号 host-id 这部分进行再划分，而不改变 IP 地址原来的网络号 net-id。

--

▷ 子网掩码

♦ 从一个 IP 数据报的首部并无法判断源主机或目的主机所连接的网络是否进行了子网划分。

♦ 使用子网掩码(subnet mask)可以找出 IP 地址中的子网部分。

♦ 子网掩码是一个网络或一个子网的重要属性。

--路由器在和相邻路由器交换路由信息时，必须把自己所在网络（或子网）的子网掩码告诉相邻路由器。

--路由器的路由表中的每一个项目，除了要给出目的网络地址外，还必须同时给出该网络的子网掩码。

--若一个路由器连接在两个子网上就拥有两个网络地址和两个子网掩码。

♦ 使用子网掩码的分组转发过程

--在不划分子网的两级 IP 地址下，从 IP 地址得出网络地址是个很简单的事。

--但在划分子网的情况下，从 IP 地址却不能唯一地得出网络地址来，这是因为网络地址取决于那个网络所采用的子网掩码，但数据报的首部并没有提供子网掩码的信息。

--因此分组转发的算法也必须做相应的改动

♦ 在划分子网的情况下路由器转发分组的算法

--(1) 从收到的分组的首部提取目的 IP 地址 *D*。

--(2) **先用各网络的子网掩码和 *D* 逐位相“与”，**看是否和

-- 相应的网络地址匹配。若匹配，则将分组直接交付。

-- 否则就是间接交付，执行(3)。

--(3) 若路由表中有目的地址为 *D* 的特定主机路由，则将

-- 分组传送给指明的下一跳路由器；否则，执行(4)。

--(4) 对路由表中的每一行的子网掩码和 *D* 逐位相“与”，

-- 若其结果与该行的目的网络地址匹配，则将分组传送
-- 给该行指明的下一跳路由器；否则，执行(5)。
--(5) 若路由表中有一个默认路由，则将分组传送给路由表
-- 中所指明的默认路由器；否则，执行(6)。
--(6) 报告转发分组出错。

4.3.2 无分类编址 CIDR

▷ 网络前缀

♦ 划分子网在一定程度上缓解了因特网在发展中遇到的困难。然而在 1992 年因特网仍然面临三个必须尽早解决的问题，这就是：

--B 类地址在 1992 年已分配了近一半，眼看就要在 1994 年 3 月全部分配完毕！

--因特网主干网上的路由表中的项目数急剧增长（从几千个增长到几万个）。

--整个 IPv4 的地址空间最终将全部耗尽。

♦ 1987 年，RFC 1009 就指明了在一个划分子网的网络中可同时使用几个不同的子网掩码。

♦ 使用变长子网掩码 VLSM (Variable Length Subnet Mask)可进一步提高 IP 地址资源的利用率。

♦ 在 VLSM 的基础上又进一步研究出无分类编址方法，它的正式名字是无分类域间路由选择 **CIDR** (Classless Inter-Domain Routing)。

▷ CIDR 最主要的特点

♦ CIDR 消除了传统的 A 类、B 类和 C 类地址以及划分子网的概念，因而可以更加有效地分配 IPv4 的地址空间。

♦ **CIDR 使用各种长度的“网络前缀”(network-prefix)来代替分类地址中的网络号和子网号**。

♦ IP 地址从三级编址（使用子网掩码）又回到了两级编址。

♦ 分类的两级编址的记法是：

--IP 地址 ::= {<网络前缀>, <主机号>} (4-3)

♦

♦ CIDR 还使用“斜线记法”(slash notation)，它又称为 CIDR 记法，即在 IP 地址面加上一个斜线“/”，然后写上网络前缀所占的位数（这个数值对应于三级编址中子网掩码中 1 的个数）。

♦ CIDR 把网络前缀都相同的连续的 IP 地址组成“CIDR 地址块”。

♦ 128.14.32.0/20 表示的地址块共有 212 个地址（因为斜线后面的 20 是网络前缀的位数，所以这个地址的主机号是 12 位）。

♦ 这个地址块的起始地址是 128.14.32.0。

♦ 在不需要指出地址块的起始地址时，也可将这样的地址块简称为“/20 地址块”。

♦ 128.14.32.0/20 地址块的最小地址：128.14.32.0

♦ 128.14.32.0/20 地址块的最大地址：128.14.47.255

♦ **全 0 和全 1 的主机号地址一般不使用。**

▷ 路由聚合(route aggregation)

♦ 一个 CIDR 地址块可以表示很多地址，这种地址的聚合常称为**路由聚合**

--它使得路由表中的一个项目可以表示很多个（例如上千个）原来传统分类地址的路由。

--路由聚合也称为构成超网(supernetting)。

--CIDR 虽然不使用子网了，但仍然使用“掩码”这一名词（但不叫子网掩码）。

--对于 /20 地址块，它的掩码是 20 个连续的 1。 斜线记法中的数字就是掩码中 1 的个数。

♦ 10.0.0.0/10 可简写为 10/10，也就是把点分十进制中低位连续的 0 省略。

♦ 10.0.0.0/10 隐含地指出 IP 地址 10.0.0.0 的掩码是 255.192.0.0。此掩码可表示为

-- 11111111 11000000 00000000 00000000

♦ 10.0.0.0/10 可简写为 10/10，也就是将点分十进制中低位连续的 0 省略。

♦ 10.0.0.0/10 相当于指出 IP 地址 10.0.0.0 的掩码是 255.192.0.0，即

♦ 11111111 11000000 00000000 00000000

♦ 网络前缀的后面加一个星号 * 的表示方法

--如 00001010 00*，

--在星号 * 之前是网络前缀，

--而星号 * 表示 IP 地址中的主机号，可以是任意值。

▷ 构成超网

♦ 前缀长度不超过 23 位的 CIDR 地址块都包含了多个 C 类地址。

--这些 C 类地址合起来就构成了超网。

--CIDR 地址块中的地址数一定是 2 的整数次幂。

♦ **网络前缀越短，其地址块所包含的地址数就越多**。

--而在三级结构的 IP 地址中，划分子网是使网络前缀变长。

▷ 最长前缀匹配

♦ 使用 CIDR 时，路由表中的每个项目由“网络前缀”和“下一跳地址”组成。

--在查找路由表时可能会得到不止一个匹配结果。

♦ 应当从匹配结果中选择具有最长网络前缀的路由：**最长前缀匹配** (longest-prefix matching)。

--网络前缀越长，其地址块就越小，因而路由就越具体(more specific) 。

♦ 最长前缀匹配又称为最长匹配或最佳匹配。

4.4 网际控制报文协议 ICMP

4.4.1 为了提高 IP 数据报交付成功的机会，在网际层使用了网际控制报文协议 **ICMP** (*Internet Control Message Protocol*)。

▷ ICMP 允许主机或路由器报告差错情况和提供有关异常情况的报告。

▷ ICMP 不是高层协议，而是 IP 层的协议。

▷ ICMP 报文作为 IP 层数据报的数据，加上数据报的首部，组成 IP 数据报发送出去。

4.4.2 ICMP 报文的种类

- **ICMP 差错报告报文 和**
- **ICMP 询问报文。**

4.4.3 ICMP 差错报告报文共有 5 种

- **终点不可达**
- **源点抑制(Source quench)**
- **时间超过**
- **参数问题**
- **改变路由（重定向）(Redirect)**

4.4.4 ICMP 询问报文有两种

- 回送请求和回答报文
- 时间戳请求和回答报文

4.4.5

4.5 因特网的路由选择协议

4.5.1 有关路由选择协议的几个基本概念

- 理想的路由算法
  - 算法必须是正确的和完整的。
  - 算法在计算上应简单。
  - 算法应能适应通信量和网络拓扑的变化(自适应性) 。
  - 算法应具有稳定性。
  - 算法应是公平的。
  - 算法应是最佳的。
- 不存在一种绝对的最佳路由算法。
  - 所谓“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。
  - 实际的路由选择算法，应尽可能接近于理想的算法。
- 路由选择是个非常复杂的问题
  - 它是网络中的所有结点共同协调工作的结果。
  - 路由选择的环境往往是不断变化的，而这种变化有时无法事先知道。
- 从路由算法的自适应性考虑
  - **静态路由选择策略——**
    --即非自适应路由选择，
    --其特点是简单和开销较小，但不能及时适应网络状态的变化。
  - **动态路由选择策略——**
    --即自适应路由选择，
    --其特点是能较好地适应网络状态的变化，但实现起来较为复杂，开销也比较大。
- 分层次的路由选择协议
  - 因特网采用分层次的路由选择协议。
    --因特网的规模非常大。
    - 如果让所有的路由器知道所有的网络应怎样到达，则这种路由表将非常大，处理起来也太花时间。
    - 而所有这些路由器之间交换路由信息所需的带宽就会使因特网的通信链路饱和。

--许多单位不愿意外界了解自己单位网络的布局细节和本部门所采用的路由选择协议

- 这属于本部门内部的事情
- 但同时还希望连接到因特网上

▷ 自治系统 AS

(Autonomous System)

- ♦ 自治系统 AS 的定义：

--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 之间的路由选择协议用以确定分组在 AS 之间的路由。

- ♦ 现在对自治系统 AS 的定义是强调下面的事实：

--尽管一个 AS 使用了多种内部路由选择协议和度量，但重要的是一个 AS 对其他 AS 表现出的是一个单一的和一致的路由选择策略。

▷ 因特网有两大类路由选择协议

- ♦ 内部网关协议 **IGP** (*Interior Gateway Protocol*)

--即在一个自治系统内部使用的路由选择协议。

--目前这类路由选择协议使用得最多

- 如 RIP 和 OSPF 协议。

- ♦ 外部网关协议 **EGP** (*External Gateway Protocol*)

--若源站和目的站处在不同的自治系统中，当数据报传到一个自治系统的边界时，就需要使用一种协议将路由选择信息传递到另一个自治系统中。

--这样的协议就是外部网关协议 EGP。

- 在外部网关协议中目前使用最多的是 BGP-4。

4.5.2 内部网关协议 RIP

(Routing Information Protocol)

▷ 工作原理

- ♦ 路由信息协议 RIP 是内部网关协议 IGP 中最先得到广泛使用的协议。
- ♦ RIP 是一种分布式的基于**距离向量**的路由选择协议。
- ♦ RIP 协议要求网络中的每一个路由器都要维护从它自己到其他每一个目的网络的距离记录。

▷ 从一路由器到直接连接的网络的距离定义为 1。

- ♦ 从一个路由器到非直接连接的网络的距离定义为所经过的路由器数加 1。
- ♦ RIP 协议中的“距离”也称为“跳数”(hop count)，因为每经过一个路由器，跳数就加 1。
- ♦ 这里的“距离”实际上指的是“最短距离”，

▷ RIP 认为一个好的路由就是它通过的路由器的数目少，即“距离短”。

- ♦ RIP 允许一条路径最多只能包含 **15** 个路由器。
- ♦ “距离”的最大值为 **16** 时即相当于不可达。可见 RIP 只适用于小型互联网。

- ♦ RIP **不能在两个网络之间同时使用多条路由**。
- ♦ RIP 选择一个具有最少路由器的路由（即最短路由），哪怕还存在另一条高速(低时延)但路由器较多的路由。

▷ RIP 协议的三个要点

- ♦ 仅和**相邻路由器**交换信息。
- ♦ 交换的信息是当前本路由器所知道的**全部信息**，即自己的路由表。
- ♦ 按固定的时间间隔**交换路由信息**，例如，每隔 30 秒。

▷ **路由表的建立**

- ♦ 路由器在刚刚开始工作时，只知道到直接连接的网络的距离（此距离定义为 1）。
- ♦ 以后，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，所有的路由器最终都会知道到达本自治系统中任何一个网络的最短距离和下一跳路由器的地址。
- ♦ RIP 协议的**收敛**(convergence)过程较快，即在自治系统中所有的结点都得到正确的路由选择信息的过程。

▷ **距离向量算法**

- ♦

收到相邻路由器（其地址为 X）的一个 RIP 报文：
(1) 先修改此 RIP 报文中的所有项目：
把“下一跳”字段中的地址都改为 X，并把所有的“距离”字段的值加 1。
(2) 对修改后的 RIP 报文中的每一个项目，重复以下步骤：
若 项目中的目的网络不在路由表中，则把该项目加到路由表中
否则
若 下一跳字段给出的路由器地址是同样的，则把收到的项目替换原路由表中的项目。
否则
若 收到项目中的距离小于路由表中的距离，则进行更新，
否则，什么也不做。
(3) 若 3 分钟还没有收到相邻路由器的更新路由表，则把此相邻路由器记为不可达路由器，即将距离置为16（距离为16表示不可达）。
(4) 返回。

▷ 路由器之间交换信息

- ♦ RIP 协议让互联网中的所有路由器都和自己的相邻路由器不断交换路由信息，并不断更新其路由表，使得从每一个路由器到每一个目的网络的路由都是最短的（即跳数最少）。
- ♦ 虽然所有的路由器最终都拥有了整个自治系统的全局路由信息，但由于每一个路由器的位置不同，它们的路由表当然也应当是不同的。

4.5.3 内部网关协议 OSPF

(Open Shortest Path First)

▷ OSPF 协议的基本特点

- ♦ “开放”表明 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 只是一个协议的名字，它并不表示其他的路由选择协议不是

“最短路径优先”。

- ♦ 是分布式的**链路状态协议**。

▷ 三个要点

- ♦ 向本自治系统中所有路由器发送信息，这里使用的方法是洪泛法。
- ♦ 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，但这只是路由器所知道的部分信息。
  - --“链路状态”就是说明本路由器都和哪些路由器相邻，以及该链路的“度量”(metric)。
  - --只有当链路状态发生变化时，路由器才用洪泛法向所有路由器发送此信息。

▷ 链路状态数据库

- ♦ 由于各路由器之间频繁地交换链路状态信息，因此所有的路由器最终都能建立一个链路状态数据库。
- ♦ 这个数据库实际上就是**全网**的拓扑结构图，它在全网范围内是一致的（这称为链路状态数据库的同步）。
- ♦ OSPF 的链路状态数据库能较快地进行更新，使各个路由器能及时更新其路由表。OSPF 的更新过程收敛得快是其重要优点

▷ OSPF 的区域(area)

- ♦ 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，叫作区域。
  - --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。
- ♦ 划分区域的好处:
  - --将利用洪泛法交换链路状态信息的范围局限于每一个区域而不是整个的自治系统，这就减少了整个网络上的通信量。
- ♦ 在一个区域内部的路由器只知道本区域的完整网络拓扑，而不知道其他区域的网络拓扑的情况。
- ♦ OSPF 使用层次结构的区域划分。
  - --在上层的区域叫作**主干区域**(***backbone area***)。
  - --主干区域的标识符规定为 **0.0.0.0**。
  - --主干区域的作用是用来连通其他在下层的区域。

▷ OSPF 直接用 IP 数据报传送

- ♦ OSPF 不用 UDP 而是直接用 IP 数据报传送。
- ♦ OSPF 构成的数据报很短。
  - --这样做可减少路由信息的通信量。
  - --数据报很短的另一好处是可以不必将长的数据报分片传送。
  - --分片传送的数据报只要丢失一个，就无法组装成原来的数据报，而整个数据报就必须重传。
- ♦ 所有在 OSPF 路由器之间交换的分组都具有鉴别的功能。
- ♦ 支持可变长度的子网划分和无分类编址 CIDR。
- ♦ 每一个链路状态都带上一个 32 位的序号，序号越大状态就越新。

▷ OSPF 的五种分组类型

- ♦ 类型 1——问候(Hello)分组。
- ♦ 类型 2 ——数据库描述(Database Description)分组。

- ♦ 类型 3 ——链路状态请求(Link State Request)分组。
- ♦ 类型 4 ——链路状态更新(Link State Update)分组,
  --用洪泛法对全网更新链路状态。
- ♦ 类型 5 ——链路状态确认(Link State Acknowledgment)分组

▷ OSPF 还规定每隔一段时间，如 30 分钟，要刷新一次数据库中的链路状态。

♦ 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。

♦ OSPF 没有“坏消息传播得慢”的问题，据统计，其响应网络变化的时间小于 100 ms。

♦

▷ 指定的路由器
(designated router)

♦ 多点接入的局域网采用了指定的路由器的方法，使广播的信息量大大减少。

♦ 指定的路由器代表该局域网上所有的链路向连接到该网络上的各路由器发送状态信息。

4.5.4 外部网关协议 BGP

▷ BGP 是不同自治系统的路由器之间交换路由信息的协议。

♦ BGP 较新版本是 2006 年 1 月发表的 BGP-4（BGP 第 4 个版本），即 RFC 4271～4278。

♦ 可以将 BGP-4 简写为 BGP。

▷ 因特网的规模太大，使得自治系统之间路由选择非常困难。对于自治系统之间的路由选择，要寻找最佳路由是很不现实的。

♦ 当一条路径通过几个不同 AS 时，要想对这样的路径计算出有意义的代价是不太可能的。

♦ 比较合理的做法是在 AS 之间交换“可达性”信息。

▷ 自治系统之间的路由选择必须考虑有关策略。

▷ 因此，边界网关协议 BGP 只能是力求寻找一条能够到达目的网络且**比较好的**路由（不能兜圈子），而并非要寻找一条最佳路由。

▷ BGP 发言人

♦ 每一个自治系统的管理员要选择至少一个路由器作为该自治系统的“**BGP 发言人**”。

♦ 一般说来，两个 BGP 发言人都是通过一个共享网络连接在一起的

♦ 而 BGP 发言人往往就是 BGP 边界路由器，但也可以不是 BGP 边界路由器。

♦ 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，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。

▷ BGP 交换路由信息

♦ 使用 TCP 连接能提供可靠的服务，也简化了路由选择协议。

♦ 使用 TCP 连接交换路由信息的两个 BGP 发言人，彼此成为对方的邻站或对等站。

♦ BGP 所交换的网络可达性的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的信息后，各 BGP 发言人就根据所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。

▷ 特点

♦ BGP 协议交换路由信息的结点数量级是自治系统数的量级

♦ 这要比这些自治系统中的网络数少很多。

♦ 每一个自治系统中 BGP 发言人（或边界路由器）的数目是很少的。

♦ 这样就使得自治系统之间的路由选择不致过分复杂。

▷ 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方面都有好处。

4.5.5 路由器在网际互连中的作用

▷ 路由器的结构

♦ 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。

--也就是说，将路由器某个输入端口收到的分组，按照分组要去的目的地（即目的网络），把该分组从路由器的某个合适的输出端口转发给下一跳路由器。

--下一跳路由器也按照这种方法处理分组，直到该分组到达终点为止。

▷ “转发”和“路由选择”的区别

♦ “转发”(forwarding)

--就是路由器根据转发表将用户的 IP 数据报从合适的端口转发出去。

♦ “路由选择”(routing)

--则是按照分布式算法，根据从各相邻路由器得到的关于网络拓扑的变化情况，动态地改变所选择的路由。

♦ 路由表是根据路由选择算法得出的。而转发表是从路由表得出的。

--在讨论路由选择的原理时，往往不去区分转发表和路由表的区别

4.6 IP 多播

4.6.1 IP 多播的基本概念

▷ (1) 多播使用组地址—— IP 使用 **D 类**地址支持多播。多播地址只能用于目的地址，而不能用于源地址。

▷ (2) 永久组地址——由因特网号码指派管理局 IANA 负责指派。

▷ (3) 动态的组成员

▷ (4) 使用硬件进行多播

4.6.2 在局域网上进行硬件多播

▷ 因特网号码指派管理局 IANA 拥有的以太网地址块的高 24 位为 00-00-5E。

▷ 因此 TCP/IP 协议使用的**以太网多播地址块的范围**是：从 **00-00-5E**-00-00-00

▷ 到 **00-00-5E**-FF-FF-FF

▷ D 类 IP 地址可供分配的有 28 位，在这 28 位中的前 5 位不能用来构成以太网硬件地址。

4.6.3 网际组管理协议 IGMP 和多播路由选择协议

▷ IP 多播需要两种协议

♦ 为了使路由器知道多播组成员的信息，需要利用网际组管理协议 **IGMP** (Internet Group Management Protocol)。

♦ 连接在局域网上的多播路由器还必须和因特网上的其他多播路由器协同工作，以便把多播数据报用最小代价传送给所有的组成员。这就需要使用多播路由选择协议。

▷ IGMP 的**本地**使用范围

♦ IGMP 并非在因特网范围内对所有多播组成员进行管理的协议。

♦ IGMP 不知道 IP 多播组包含的成员数，也不知道这些成员都分布在哪些网络上。

♦ IGMP 协议是让连接在本地局域网上的多播路由器知道本局域网上是否有主机（严格讲，是主机上的某个进程）参加或退出了某个多播组。

▷ 多播路由选择协议比单播路由选择协议复杂得多

▷ IGMP 可分为两个阶段

♦ 第一阶段：

--当某个主机加入新的多播组时，该主机应向多播组的多播地址发送 IGMP 报文，声明自己要成为该组的成员。

--本地的多播路由器收到 IGMP 报文后，将组成员关系转发给因特网上的其他多播路由器。

♦ 第二阶段：

--因为组成员关系是动态的，因此本地多播路由器要周期性地探询本地局域网上的主机，以便知道这些主机是否还继续是组的成员。

--只要对某个组有一个主机响应，那么多播路由器就认为这个组是活跃的。

--但一个组在经过几次的探询后仍然没有一个主机响应，则不再将该组的成员关系转发给其他的多播路由器。

▷ **在主机和多播路由器之间的所有通信都是使用 IP 多播**。

♦ 多播路由器在探询组成员关系时，只需要对所有的组发送一个请求信息的询问报文，而不需要对每一个组发送一个询问报文。

--默认的询问速率是每 125 秒发送一次。

♦ 当同一个网络上连接有几个多播路由器时，它们能够迅速和有效地选择其中的一个来探询主机的成员关系。

♦ 在 IGMP 的询问报文中有一个数值 *N*，它指明一个最长响应时间（默认值为 10 秒）。当收到询问时，主机在 0 到 *N* 之间随机选择发送响应所需经过的时延。

--对应于最小时延的响应最先发送。

♦ 同一个组内的每一个主机都要监听响应，只要有本组的其他主机先发送了响应，自己就可以不再发送响应了。

- 多播路由选择
  - 多播路由选择协议尚未标准化。
    --一个多播组中的成员是动态变化的，随时会有主机加入或离开这个多播组。
    --多播路由选择实际上就是要找出以源主机为根结点的多播转发树。
    --在多播转发树上的路由器不会收到重复的多播数据报。
    --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。

4.7 课后练习

- **完成课后练习题，并在课堂进行答案校对、剖析**

第五章 运输层

❒ 课时安排：

8 学时

❒ 教学目的：

1. 认识网络边缘的第四个层——传输层、运输层
2. 传输层的功能
3. 传输层两个出名的协议
4. UDP 协议
5. TCP 协议

❒ 教学重、难点：

- 传输层的两个协议特点
- UDP 协议与 TCP 协议各自适用的范围
- UDP 协议的动作方式和优缺点
- TCP 协议工作的过程，与 UDP 的工作不同点
- 传输层的可靠或不可靠协议的交付与 IP 层的可靠和不可靠传送的区别
- 传输层里数据格式是——数据报文

❒ 教学内容：

5.1 运输层协议概述

5.1.1 进程之间的通信

▷ 从通信和信息处理的角度看，运输层向它上面的应用层提供通信服务，它属于面向通信部分的最高层，同时也是用户功能中的最低层。

▷ 当网络的边缘部分中的两个主机使用网络的核心部分的功能进行端到端的通信时:

- ♦ 只有位于网络边缘部分的主机的协议栈才有运输层
- ♦ 而网络核心部分中的路由器在转发分组时都只用到下三层的功能。

▷ 应用进程之间的通信

- ♦ 两个主机进行通信实际上就是两个主机中的应用进程互相通信。
- ♦ 应用进程之间的通信又称为**端到端**的通信。
- ♦ 运输层的一个很重要的功能就是**复用**和**分用**。

  --应用层不同进程的报文通过不同的端口向下交到运输层，再往下就共用网络层提供的服务。
- ♦ “运输层提供应用进程间的逻辑通信”。

  --“逻辑通信”的意思是：
  - 运输层之间的通信好像是沿水平方向传送数据。
  - 但事实上这两个运输层之间并没有一条水平方向的物理连

接。

▷ 运输层的主要功能

♦ 运输层为应用进程之间提供端到端的逻辑通信（但网络层是为主机之间提供逻辑通信）。

♦ 运输层还要对收到的报文进行差错检测。

♦ 运输层需要有两种不同的运输协议，

--面向连接的 TCP

--无连接的 UDP

▷ 两种不同的运输协议

♦ 运输层向高层用户屏蔽了下面网络核心的细节（如网络拓扑、所采用的路由选择协议等），它使应用进程看见的就是好像在两个运输层实体之间有一条端到端的逻辑通信信道。

--当运输层采用面向连接的 TCP 协议时

● 尽管下面的网络是不可靠的（只提供尽最大努力服务），但这种逻辑通信信道就相当于一条全双工的可靠信道。

--当运输层采用无连接的 UDP 协议时

● 这种逻辑通信信道是一条不可靠信道。

5.1.2 运输层的两个主要协议

▷ TCP 与 UDP

♦ 两个对等运输实体在通信时传送的数据单位叫作运输协议数据单元 TPDU (Transport Protocol Data Unit)。

--TCP 传送的数据单位协议是 TCP 报文段(segment)

--UDP 传送的数据单位协议是 UDP 报文或用户数据报。

♦ UDP 在传送数据之前不需要先建立连接。

--对方的运输层在收到 UDP 报文后，不需要给出任何确认。

--虽然 UDP 不提供可靠交付，但在某些情况下 UDP 是一种最有效的工作方式。

♦ TCP 则提供面向连接的服务。

--TCP 不提供广播或多播服务。

--由于 TCP 要提供可靠的、面向连接的运输服务，因此不可避免地增加了许多的开销。这不仅使协议数据单元的首部增大很多，还要占用许多的处理机资源。

♦ 运输层的 UDP 用户数据报与网际层的 IP 数据报有很大区别。

--IP 数据报要经过互连网中许多路由器的存储转发

--UDP 用户数据报是在运输层的端到端抽象的逻辑信道中传送的。

♦ TCP 报文段是在运输层抽象的端到端逻辑信道中传送，这种信道是可靠的全双工信道。

--但这样的信道却不知道究竟经过了哪些路由器

--这些路由器也根本不知道上面的运输层是否建立了 TCP 连接。

5.1.3 运输层的端口

▷ 运行在计算机中的进程是用**进程标识符**来标志的。

▷

▷ 运行在应用层的各种应用进程却不应当让计算机操作系统指派它的进程标识符。

♦ 这是因为在因特网上使用的计算机的操作系统种类很多，而不同的操作系统又使用不同格式的进程标识符。

♦ 为了使运行不同操作系统的计算机的应用进程能够互相通信，就必须用统一的方法对 TCP/IP 体系的应用进程进行标志。

♦ 问题：

--由于进程的创建和撤销都是动态的，发送方几乎无法识别其他机器上的进程。

--有时我们会改换接收报文的进程，但并不需要通知所有发送方。

--我们往往需要利用目的主机提供的功能来识别终点，而不需要知道实现这个功能的进程。

▷ 解决这个问题的方法就是在运输层使用协议端口号(protocol port number)，或通常简称为**端口**(port)。

♦ 虽然通信的终点是应用进程，但我们可以把端口想象是通信的终点

♦ 因为我们只要把要传送的报文交到目的主机的某一个合适的目的端口，剩下的工作（即最后交付目的进程）就由 TCP 来完成。

▷ 在协议栈层间的抽象的协议端口是软件端口。

▷ 路由器或交换机上的端口是硬件端口。

▷ **区别：**

♦ 硬件端口是不同硬件设备进行交互的接口

♦ 而软件端口是应用层的各种协议进程与运输实体进行层间交互的一种地址。

▷ TCP 的端口

♦ 端口用一个 16 位端口号进行标志。

♦ 端口号只具有**本地**意义

--即端口号只是为了标志**本**计算机应用层中的各进程。

--在因特网中不同计算机的相同端口号是没有联系的。

▷ 三类端口

♦ **熟知端口**

--数值一般为 0~1023。

♦ **登记端口号**

--数值为 1024~49151，为没有熟知端口号的应用程序使用的。使用这个范围的端口号必须在 IANA 登记，以防止重复。

♦ **客户端口号或短暂端口号**

--数值为 49152~65535，留给客户进程选择暂时使用。

--当服务器进程收到客户进程的报文时，就知道了客户进程所使用的动态端口号。

--通信结束后，这个端口号可供其他客户进程以后使用。

5.2 用户数据报协议 UDP

5.2.1 UDP 概述

▷ UDP 只在 IP 的数据报服务之上增加了很少一点的功能——

♦ 即端口的功能和差错检测的功能。

▷ 虽然 UDP 用户数据报只能提供不可靠的交付，但 UDP 在某些方面有其特殊的优点。

▷ 主要特点

- ♦ UDP 是无连接的，即发送数据之前不需要建立连接。
- ♦ UDP 使用尽最大努力交付，即不保证可靠交付，同时也不使用拥塞控制。
- ♦ UDP 是面向报文的。UDP 没有拥塞控制，很适合多媒体通信的要求。
- ♦ UDP 支持一对一、一对多、多对一和多对多的交互通信。
- ♦ UDP 的首部开销小，只有 8 个字节。

▷ 面向报文的 UDP

- ♦ 发送方 UDP 对应用程序交下来的报文，在添加首部后就向下交付 IP 层。

--UDP 对应用层交下来的报文，既不合并，也不拆分，而是保留这些报文的边界。

--应用层交给 UDP 多长的报文，UDP 就照样发送，即一次发送一个报文。

- ♦ 接收方 UDP 对 IP 层交上来的 UDP 用户数据报，在去除首部后就原封不动地交付上层的应用进程，一次交付一个完整的报文。
- ♦ **应用程序必须选择合适大小的报文**。

5.3 传输控制协议 TCP 概述

5.3.1 TCP 最主要的特点

- ♦ TCP 是面向连接的运输层协议。
- ♦ 每一条 TCP 连接只能有两个端点(endpoint)，每一条 TCP 连接只能是点对点的（一对一）。
- ♦ TCP 提供可靠交付的服务。
- ♦ TCP 提供全双工通信。
- ♦ 面向字节流。

▷ TCP 连接是一条虚连接而不是一条真正的物理连接。

▷ TCP 对应用进程一次把多长的报文发送到 TCP 的缓存中是不关心的。

▷ TCP 根据对方给出的窗口值和当前网络拥塞的程度来决定一个报文段应包含多少个字节（UDP 发送的报文长度是应用进程给出的）。

▷ TCP 可把太长的数据块划分短一些再传送。TCP 也可等待积累有足够多的字节后再构成报文段发送出去。

5.3.2 TCP 的连接

▷ TCP 把连接作为最基本的抽象。

▷ 每一条 TCP 连接有两个端点。

- ♦ TCP 连接的端点不是主机，不是主机的 IP 地址，不是应用进程，也不是运输层的协议端口。TCP 连接的端点叫做套接字(socket)或插口。
- ♦ 端口号拼接到(contatenated with) IP 地址即构成了套接字。

▷ 套接字 **socket = (IP 地址: 端口号)** (5-1)

- ♦ 每一条 **TCP** 连接唯一地被通信两端的两个端点（即两个套接字）所确定。即：
- ♦ TCP 连接 ::= {socket1, socket2}
- ♦ = {(IP1: port1), (IP2: port2)} (5-2)

5.4 可靠传输的工作原理

5.4.1 可靠传输

▷ 使用确认和重传机制，我们就可以在不可靠的传输网络上实现可靠的通

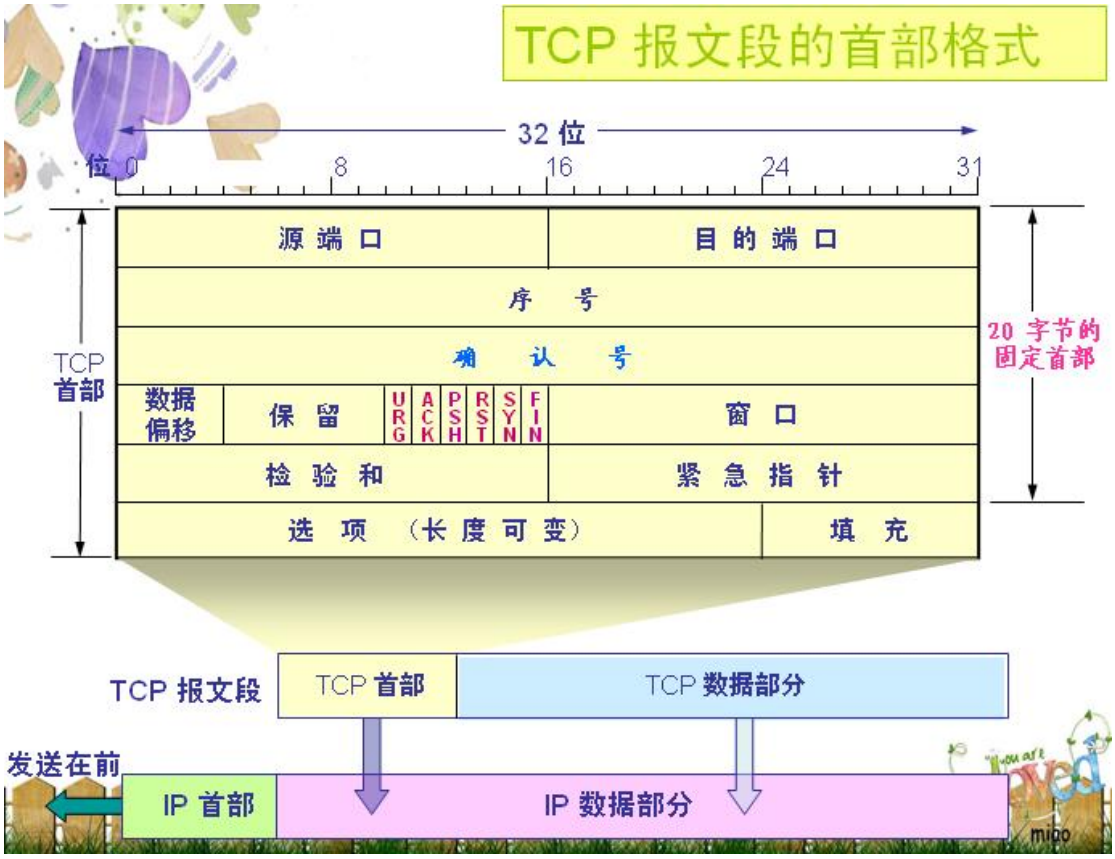
信。

- ▷ 这种可靠传输协议常称为自动重传请求 **ARQ** (Automatic Repeat reQuest)。
- ▷ ARQ 表明重传的请求是自动进行的。接收方不需要请求发送方重传某个出错的分组 。

5.4.2 连续 ARQ 协议

- ▷ 接收方一般采用累积确认的方式。
  - ♦ 即不必对收到的分组逐个发送确认，而是对按序到达的最后一个分组发送确认
  - ♦ 这样就表示：到这个分组为止的所有分组都已正确收到了。
- ▷ 累积确认优点是：
  - ♦ 容易实现，即使确认丢失也不必重传。
- ▷ 缺点是：
  - ♦ 不能向发送方反映出接收方已经正确收到的所有分组的信息。
- ▷ Go-back-N（
  - ♦ 如果发送方发送了前 5 个分组，而中间的第 3 个分组丢失了。
    --这时接收方只能对前两个分组发出确认。
    --发送方无法知道后面三个分组的下落，
    --只好把后面的三个分组都再重传一次。
  - ♦ 这就叫做 Go-back-N（回退 N），表示需要再退回来重传已发送过的 *N* 个分组。
  - ♦ 可见当通信线路质量不好时，连续 ARQ 协议会带来负面的影响。
- ▷ TCP 连接的每一端都必须设有两个窗口——一个发送窗口和一个接收窗口。
  - ♦ TCP 的可靠传输机制用字节的序号进行控制。TCP 所有的确认都是基于序号而不是基于报文段。
  - ♦ TCP 两端的四个窗口经常处于动态变化之中。
  - ♦ TCP 连接的往返时间 RTT 也不是固定不变的。需要使用特定的算法估算较为合理的重传时间。

5.5 TCP 报文段的首部格式



5.5.1

5.5.2 发送缓存与接收缓存的作用

- ▷ 发送缓存用来暂时存放：
  - ♦ 发送应用程序传送给发送方 TCP 准备发送的数据；
  - ♦ TCP 已发送出但尚未收到确认的数据。
- ▷ 接收缓存用来暂时存放：
  - ♦ 按序到达的、但尚未被接收应用程序读取的数据；
  - ♦ 不按序到达的数据。

5.5.3 超时重传时间的选择

- ▷ 重传机制是 TCP 中最重要和最复杂的问题之一。
  - ♦ TCP 每发送一个报文段，就对这个报文段设置一次计时器。
  - ♦ 只要计时器设置的重传时间到但还没有收到确认，就要重传这一报文段。

5.6 选择确认 SACK

- ▷ 接收方收到了和前面的字节流不连续的两个字节块。
- ▷ 如果这些字节的序号都在接收窗口之内，那么接收方就先收下这些数据，但要把这些信息准确地告诉发送方，使发送方不要再重复发送这些已收到的数据。

5.7 TCP 的流量控制

5.7.1 利用滑动窗口实现流量控制

- ▷ 一般说来，我们总是希望数据传输得更快一些。但如果发送方把数据发送得过快，接收方就可能来不及接收，这就会造成数据的丢失。
- ▷ **流量控制**(flow control)就是让发送方的发送速率不要太快，既要让接收方来得及接收，也不要使网络发生拥塞。

- ♦ 利用滑动窗口机制可以很方便地在 TCP 连接上实现流量控制。

5.7.2 持续计时器

- ▷ 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（仅携带 1 字节的数据），而对方就在确认这个探测报文段时给出了现在的窗口值。
- ▷ 若窗口仍然是零，则收到这个报文段的一方就重新设置持续计时器。
- ▷ 若窗口不是零，则死锁的僵局就可以打破了。

5.7.3 必须考虑传输效率

- ▷ 可以用不同的机制来控制 TCP 报文段的发送时机:
- ▷ 第一种机制是 TCP 维持一个变量，它等于最大报文段长度 MSS。只要缓存中存放的数据达到 MSS 字节时，就组装成一个 TCP 报文段发送出去。
- ▷ 第二种机制是由发送方的应用进程指明要求发送报文段，即 TCP 支持的推送(push)操作。
- ▷ 第三种机制是发送方的一个计时器期限到了，这时就把当前已有的缓存数据装入报文段（但长度不能超过 MSS）发送出去。

5.8 TCP 的拥塞控制

5.8.1 拥塞控制的一般原理

- ▷ 在某段时间，若对网络中某资源的需求超过了该资源所能提供的可用部分，网络的性能就要变坏——产生**拥塞**(congestion)。
- ▷ 出现资源拥塞的条件：
- ▷ **对资源需求的总和 > 可用资源** (5-7)
- ▷ 若网络中有许多资源同时产生拥塞，网络的性能就要明显变坏，整个网络的吞吐量将随输入负荷的增大而下降。
- ▷ **拥塞控制**
  - ♦ 都有一个前提
  - ♦ 就是网络能够承受现有的网络负荷。
  - ♦ 拥塞控制是一个全局性的过程
  - ♦ 涉及到所有的主机、所有的路由器，以及与降低网络传输性能有关的所有因素。
- ▷ **流量控制**
  - ♦ 往往指在给定的发送端和接收端之间的点对点通信量的控制。
  - ♦ 流量控制所要做的就是抑制发送端发送数据的速率，以便使接收端来得及接收。
- ▷ 拥塞控制是很难设计的:
  - --因为它是一个动态的（而不是静态的）问题。
  - --当前网络正朝着高速化的方向发展，这很容易出现缓存不够大而造成分组的丢失。
  - --但分组的丢失是网络发生拥塞的征兆而不是原因。
  - ♦ 在许多情况下，甚至正是拥塞控制本身成为引起网络性能恶化甚至发生死锁的原因。
  - --这点应特别引起重视。
- ▷ 开环控制方法

♦ 就是在设计网络时事先将有关发生拥塞的因素考虑周到，力求网络在工作时不产生拥塞。

▷ 闭环控制

♦ 是基于反馈环路的概念。

♦ 属于闭环控制的有以下几种措施：

♦ 监测网络系统以便检测到拥塞在何时、何处发生。

♦ 将拥塞发生的信息传送到可采取行动的地方。

♦ 调整网络系统的运行以解决出现的问题。

▷

5.8.2 几种拥塞控制方法

▷ **慢开始和拥塞避免**

♦ 发送方维持一个叫做拥塞窗口 cwnd (congestion window)的状态变量。

♦ 拥塞窗口的大小取决于网络的拥塞程度，并且动态地在变化。

♦ 发送方让自己的发送窗口等于拥塞窗口。

♦ 如再考虑到接收方的接收能力，则发送窗口还可能小于拥塞窗口。

♦ 发送方控制拥塞窗口的原则是：

♦ 只要网络没有出现拥塞，拥塞窗口就再增大一些，以便把更多的分组发送出去。

♦ 但只要网络出现拥塞，拥塞窗口就减小一些，以减少注入到网络中的分组数。

♦ 在主机刚刚开始发送报文段时可先设置拥塞窗口 cwnd = 1，即设置为一个最大报文段 MSS 的数值。

♦ 在每收到一个对新的报文段的确认后，将拥塞窗口加 1，即增加一个 MSS 的数值。

♦ 用这样的方法逐步增大发送端的拥塞窗口 cwnd，可以使分组注入到网络的速率更加合理。

♦ **传输轮次(transmission round)**

--使用慢开始算法后，每经过一个传输轮次，拥塞窗口 cwnd 就加倍。

--一个传输轮次所经历的时间其实就是往返时间 RTT。

--“传输轮次”更加强调：

--把拥塞窗口 cwnd 所允许发送的报文段都连续发送出去，并收到了对已发送的最后一个字节的确认。

--例如:

--拥塞窗口 cwnd = 4，这时的往返时间 RTT 就是发送方连续发送 4 个报文段，并收到这 4 个报文段的确认，总共经历的时间。

♦ **设置慢开始门限状态变量 ssthresh**

--慢开始门限 ssthresh 的用法如下：

--当 cwnd < ssthresh 时，使用慢开始算法。

--当 cwnd > ssthresh 时，停止使用慢开始算法改用拥塞避免算法。

--当 cwnd = ssthresh 时，既可使用慢开始算法，也可使用拥塞避免算法。

--拥塞避免算法的思路是

--让拥塞窗口 cwnd 缓慢地增大，即每经过一个往返时间 RTT 就把

发送方的拥塞窗口 cwnd 加 1，而不是加倍，使拥塞窗口 cwnd 按线性规律缓慢增长。

- **当网络出现拥塞时**

--无论在慢开始阶段还是在拥塞避免阶段，只要发送方判断网络出现拥塞（其根据就是没有按时收到确认），就要把慢开始门限 ssthresh 设置为出现拥塞时的发送方窗口值的一半（但不能小于 2）。

--然后把拥塞窗口 cwnd 重新设置为 1，执行慢开始算法。

--这样做的目的就是要迅速减少主机发送到网络中的分组数，使得发生拥塞的路由器有足够时间把队列中积压的分组处理完毕。

▷ **快重传和快恢复**

♦ 快重传算法首先要求接收方每收到一个失序的报文段后就立即发出重复确认。这样做可以让发送方及早知道有报文段没有到达接收方。

♦ 发送方只要一连收到三个重复确认就应当立即重传对方尚未收到的报文段。

♦ 不难看出，快重传并非取消重传计时器，而是在某些情况下可更早地重传丢失的报文段。

♦ (1) 当发送端收到连续三个重复的确认时，就执行“乘法减小”算法，把慢开始门限 ssthresh 减半。

♦ 但接下去不执行慢开始算法。

♦ (2) 由于发送方现在认为网络很可能没有发生拥塞，因此现在不执行慢开始算法，即拥塞窗口 cwnd 现在不设置为 1，而是设置为慢开始门限 ssthresh 减半后的数值，然后开始执行拥塞避免算法（“加法增大”），使拥塞窗口缓慢地线性增大。

- **发送窗口的上限值**

--发送方的发送窗口的上限值应当取为接收方窗口 rwnd 和拥塞窗口 cwnd 这两个变量中较小的一个，即应按以下公式确定：

-- 发送窗口的上限值 = Min [rwnd, cwnd] (5-8)

--当 rwnd < cwnd 时，是接收方的接收能力限制发送窗口的最大值。

--当 cwnd < rwnd 时，则是网络的拥塞限制发送窗口的最大值。

5.9 TCP 的运输连接管理

5.9.1 传输阶段

▷ 运输连接就有三个阶段，

♦ 即：

♦ 连接建立、

♦ 数据传送

♦ 和连接释放。

▷ 运输连接的管理就是使运输连接的建立和释放都能正常地进行。

▷ 连接建立过程中要解决以下三个问题：

♦ 要使每一方能够确知对方的存在。

♦ 要允许双方协商一些参数（如最大报文段长度，最大窗口大小，服务质量等）。

♦ 能够对运输实体资源（如缓存大小，连接表中的项目等）进行分配。

▷ TCP 连接的建立都是采用**客户服务器**方式。

- 主动发起连接建立的应用进程叫做客户(client)。
- 被动等待连接建立的应用进程叫做服务器(server)。

第六章 应用层

❒ 课时安排：

4 学时

❒ 教学目的：

1.
2.

❒ 教学重、难点：

1.
2.

❒ 教学内容：

第七章 网络安全

❒ 课时安排：

4 学时

❒ 教学目的：

1.
2.

❒ 教学重、难点：

1.
2.

❒ 教学内容：