第一章 概述

一、先行课程

通信原理、操作系统

二、解决的主要问题 本章主要涉及计算机网络的概念、计算机网络的应用、计算机网络的组成和分类方法；

讨论计算机网络层次化的设计方法及计算机网络的体系结构；研究 OSI 参考模型和 TCP/IP 模型。

三、本章知识点

1、基本概念

计算机网络的概念、计算机网络与分布式系统的区别 计算机网络的应用：客户端/服务器（CS）模式、浏览器 / 服务器（B/S)模式、对等网络、Web 服务 、网格计算、三网融合等。

计算机网络：由自主计算机互联起来的集合体，与主从计算的的区别，每台计算机都有独立的操作系统，不从属某台主机

分布式系统：对用户看起来是一个单独系统的计算机集合。存在着一个能为用户自动管理资源的网络操作系统。计算机网络是分布式系统的技术基础，而分布式系统是计算机网发展得高级阶段。www 是原型在 Internet 的分布式系统。

2、计算机网络的组成与分类

分类：传输技术和网络尺度（拓扑结构，交换功能）

计算机网络的传输技术：广播式的传输技术（Broadcast）、点到点的传输技术（Point-to-Point）；

计算机网络的组成和分类：个人网络（PAN）、局域网（LAN）、城域网（MAN）、广域网（WAN）、因特网/互联网（Internet）等。

局域网：广播式，总线型或者环形，覆盖有限的地理范围，提供高数据率，低误码率

城域网：结余广域网和个域网之间的一种高速网络（局域网的扩张，广域网的渗透）

广域网：主机+通信子网（主要使用分组交换），覆盖范围从几十公里到几千公里

局域网和广域网具有不同的侧重点：局域网侧重减少冲突，减少错误，关于王侧重路由算法，找出一条最佳路由

Internet：网络的网络，全球独一无二的，internet：互联的网络的集合体，十分通用

3、计算机网络的体系结构 协议分层技术：层、协议、接口、面向连接服务与无连接服务、服务与服务原语、服务与协议关系等。

层：使用了信息隐蔽、抽象数据类型以及面向对象的设计方法； 目的是向上层提供服

务，上层可以使用其提供的服务，但对于其内部的状态和算法不可见。

服务：下层（n-1 层）向上层（n 层）提供的功能，方向是垂直的。

接口：存在于每一对相邻层之间的临界处，下层通过接口向上层提供服务

协议：对等层关于如何进行通信的一种规则约定，是对该层功能如何实现的一种定义

实体(entity) 表示任何可发送或接收信息的硬件或软件进程。

协议是控制两个对等实体进行通信的规则的集合。

在协议的控制下，两个对等实体间的通信使得本层能够向上一层提供服务。

要实现本层协议，还需要使用下层所提供的服务。

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。

协议是“水平的”，即协议是控制对等实体之间通信的规则。

服务是“垂直的”，即服务是由下层向上层通过层间接口提供的。

同一系统相邻两层的实体进行交互的地方，称为服务访问点 SAP (Service Access Point)。

面向连接放服务：面向连接服务具有连接建立、数据传输和连接释放这三个阶段

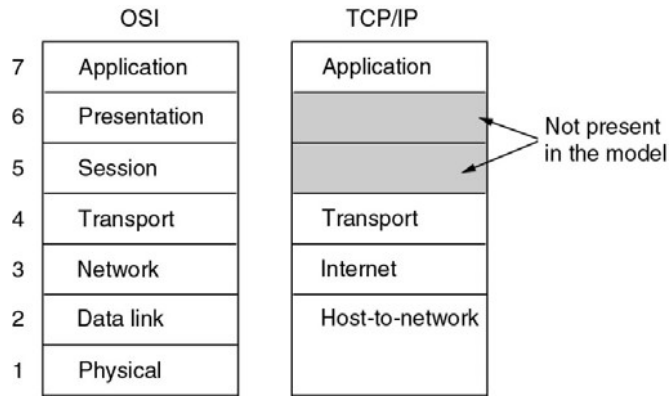
无连接服务：两个实体之间的通信不需要先建立好连接。

- 每个报文都携带了完整的目标地址，因此可以被系统独立地路由

服务原语：一个服务通常是由一组原语操作描述，用户进程通过这些操作访问该服务

4、参考模型

两个著名的参考模型 OSI 参考模型和 TCP/IP 模型如下图所示。



OSI 模型：
应用层：网络应用进程，为应该进程提供了网络服务，使得应用程序间同步
表示层：数据表示，处理两个通信系统交换信息的表示方式

会话层：主机间通信，位数两个节点间的传输爱你节，确保传输不中断
传输层：端对端的连接，处理数据包的错误，向高层屏蔽了下层
网络层：选择最优路由，是想拥塞控制，网络互联等功能
数据链路层：接入介质，在同性链路实体间建立数据链路连接：传输以帧为单位，采用差错控制和流量控制，使有差错的物理链路变成无差错的数据链路
物理层：二进制传输，实现比特流的透明传输，为数据链路层提供数据传输服务，建立，管理和释放物理连接

TCP/IP
网络接口层：发送，接收 IP 数据报
网际层:　无连接网络服务，处理互连的路由选择、流控与拥塞问题，尽力而为
传输层：TCP UDP
5、网络举例和网络标准化

Internet 与下一代 Internet、帧中继、ATM、Ethernet、无线 LAN；
国际化标准化组织。

四、重点和难点 1、计算机网络的概念与分布式系统的区别

2、按照距离尺度和传输技术进行计算机网络分类的方法

3、局域网、城域网和广域网的工作原理及其区别

4、协议分层技术

5、OSI 参考模型和 TCP/IP 参考模型

第二章 物理层

一、先行课程

概率论与随机过程、计算机组成原理、通信原理

二、解决的主要问题 本章主要涉及物理层的基本概念和功能、信道极限容量的概念以及传输速率的计算方法；

研究数据通信的基础理论，包括传输介质的选择、调制解调的方法、传输技术和物理层接口 协议等。

三、本章知识点

1、数据通信的理论基础

傅立叶分析、有限带宽信号、信道的最大数据传输速率分析，包括奈奎斯特准则、香农 定理。

2、传输介质

双绞线（RJ-45、现场测试参数）、同轴电缆、光纤、无线传输。

3、无线传输

电磁波谱、微波、卫星。

4、电话系统

本地回路（含调制解调器及其压缩功能、FTTH、FTTC、FHC 和 ADSL）、多路复用（FDM、TDM、WDM 与 DWDM）、交换。

PCM：脉冲编码调制：以一组代码序列表示模拟信号抽样值（模拟信号脉冲振幅调制，抽样，编码 8 位）

T1 系统;8000HZ,8 位二进制码元（7 语音编码+1 同步码），一个话路是 64kb/s，中的数据率是 1.544Mb/s 8000*193

E1 2.408Mb/s(30*8+2*8/125us)

5、调制解调技术 xDSL、ADSL

调制解调技术：数字数据在模拟信道中传输的方法（调制数字信号模拟化，解调模拟信号数据化）

Modem 包括调制器（数字信号转化为 300-3400hz 的模拟信号在电话用户线上传送），解调器（恢复为数字信号）

AM FM PM 载波的幅度，频率，相位随着基带数字信号而变化

基本 sonet 帧 810 字节/125us 传输速率为 51.84Mbps

调制

四、相关协议和设备 调制解调器

五、重点和难点 1、信道极限速率公式：奈奎斯特公式和香农定理

2、多路复用技术

3、模拟传输和数字传输物理层标准

第三章 数据链路层

一、先行课程 计算机组成原理、操作系统、数据结构

二、解决的主要问题 本章主要解决相邻两站点间可靠数据通信问题。包括： 1、成帧
2、差错控制
3、流量控制

三、本章知识点
1、 概念
（1）帧：是数据链路层传输数据的单位，分为数据帧和控制帧。
（2）汉明距离：码表中两个编码不同比特的数量称为这两个编码的距离。码表中最小的距离值为该码表的汉明距离。
（3）捎带应答：接收端收到数据帧时，不适用 ACK 进行确认，而是通过后续的数据帧将应答信息回送给另一端。
（4）发送窗口：滑动窗口协议中发送方保持的一组序号，对应可发送的数据帧。
（5）接收窗口：滑动窗口协议中接收方保持的一组序号，对应可接收的数据帧。
2、 提供的服务类型 数据链路层提供的服务有三种：无连接且无确认的服务（以太网实时通信）、有确认+面向连接（无线通信）、有确认+无连接（通信要求较高，长距离，不可靠，卫星通信）
连接的可靠的服务。
3、 成帧的方法
数据链路层的成帧方法有 4 种：字符计数法、字符填充法、比特填充法和物理层编码违例法。
4、 差错控制技术 差错控制技术用于检测物理层的传输误码并进行纠正，包括两种：纠错码和检错码。 纠错码是指在发送端在发送的信息中采用相关算法加上足够多的冗余信息，使接收端不
仅可以判断接收的信息有误码，还可纠正出现的错误。纠错码的实例：纠正单比特差错的汉明码。
检错码是指在发送端在发送的信息中采用相关算法加上有限的冗余信息，使接收端仅可判断接收的信息有误码。检错码的实例包括：奇偶校验、校验和、循环冗余校验码（CRC）等。
CRC 放在帧尾可以边发送边计算使得效率更高
纠错和检错能力与汉明距离有关。若码表的汉明距离为 d+1，则可检查出 d 比特的错误，汉明距离为 2d+1,则可纠正 d 比特的错误。

5、 ARQ 技术
ARQ (Automatic Repeat reQuest)即自动重传请求。若数据链路层向网络层提供可靠的数据传输服务，当物理链路有传输误码时，出现的传输错误需数据链路层进行恢复。目前大部分数据链路层协议采用的差错控制技术为检错码，对于差错的恢复必须由发送端对出错的数据帧进行重传，且重传是由发送端自主完成的，因而这种技术称作 ARQ。

6、 滑动窗口协议 流量控制问题是数据链路层需处理的另一重要问题，解决的是接收端处理能力不能及时
处理到达的帧而导致帧丢失的问题。简单的方法是采用停等的方式，发送端每次只发送一个数据帧，收到确认后再发送下一个数据帧。这种方法的缺点是，当收发两端距离较远（传播时间较大）时，系统的性能较差。

为改善停等协议的性能问题，可采用的策略是每次发送的帧数不是一个而是若干个，这种方案即为滑动窗口协议。滑动窗口协议又分为两种：接收窗口为 1 时，为 Go Back N 协议， 接收窗口大于 1 时为选择重传协议（Selective Repeat）。

发送窗口和接收窗口的大小与帧序号的比特数有关，通常情况下，Go Back N 协议的发送窗口最大值为 2^n-1，SR 协议的发送窗口最大值为 2^{n-1}，

7、 协议性能

滑动窗口协议的效率与发送窗口的大小有关，若帧发送时间为 Tf，两站间的传播时间为 Tp，设发送窗口的大小为 W，a=Tp/Tf，则信道利用率为： U=100%，W≥1+2a

U=W/(1+2a)，W＜1+2a

采用捎带应答时，信道利用率为： U=100%，W≥2+2a U=W/(2+2a)，W＜2+2a

四、相关协议和设备

1、协议

PPP、HDLC、PPPOE、SLIP

2、设备

路由器

五、重点和难点

1、成帧方法、各种成帧方法的特点和应用环境

2、汉明吗、CRC 循环冗余校验码

3、滑动窗口协议的工作原理及性能评价

六、实验内容

在数据链路层的实验中，设计并实现了涵盖物理层和数据链路层功能，并存在信道错误和延时的仿真平台环境，通过安排数据链路层的协议设计实验，让学生使用仿真通信平台，实现一个数据链路层的滑动窗口协议，通过该实验学生可掌握计算机网络协议的一般实现和调试方法并对网络中讲授的流量控制算法进行验证。该仿真平台提供基本的网络层和物理层功能，并可对通信环境进行仿真，包括数据流模型、传输延迟和误码率等。

第四章 介质访问控制子层

一、先行课程

概率论与随机过程、计算机组成原理、通信原理

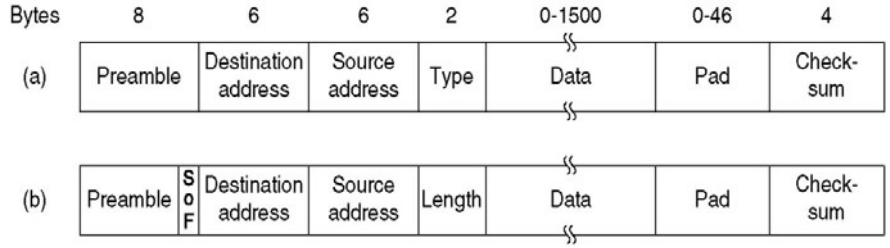
二、解决的主要问题 本章主要解决共享通信环境下介质访问控制问题，共享通信环境设计有线环境和无线环境。

三、本章知识点

1、概念

（1） 以太网帧

以太网帧结构如下图所示，图(a)为 DIX Ethernet 帧结构，图(b)为 IEEE 802.3 帧结构。



（2） MAC 地址

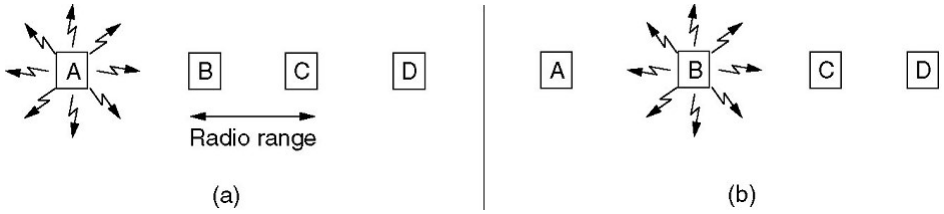
MAC 也称物理地址或以太网地址，是局域网设备的唯一标识，由 6 个 16 进制数表示，例如 00-41-43-00-80-0c。特殊的 MAC 地址 FF-FF-FF-FF-FF-FF 不能分配给设备使用，该地址是广播地址，向网内所有站点发送数据时使用该地址作为目的地址。

（3） 共享信道的性能

共享信道的性能主要包括两个指标：轻负载情况下的响应时间和重负载下的吞吐量。

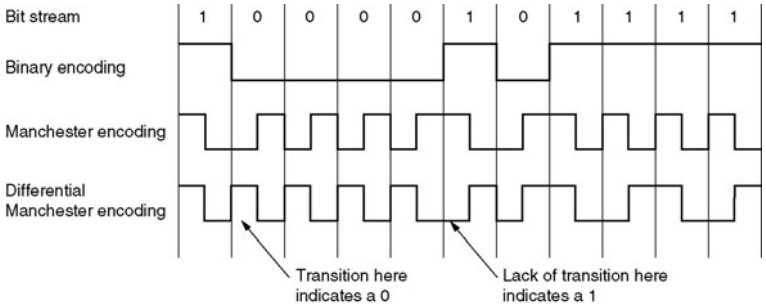
（4） 站隐藏和站暴漏

无线通信环境下的两个特殊问题。站隐藏问题是与正在通信的工作站距离较远时无法判断其正在通信，此时发数据会产生干扰。站暴露问题是与正在通信的工作站距离较近但与接收站点距离较远，此时发数据不会产生干扰



（5） 曼彻斯特编码

一种编码方式，传统以太网使用该编码。



（6） 冲突域

两个站点不能同时发送数据，则这两个站点属于一个冲突域。连接在一个集线器的所有站点属于一个冲突域。

（7） 广播域

站点发送广播帧时，收到该帧的站点与其处在同一个广播域。连接在一个交换机或集线器的所有站点同处一个广播域。

（8） 广播风暴

局域网中产生大量广播帧时，会对接收站点的性能产生影响，这种情况称为广播风暴。

2、 信道共享算法

（1） 静态信道分配算法

将信道静态划分为多个子信道使用，例如 PCM。

（2） 动态信道分配算法-受控多路访问控制 轮询、令牌

（3） 动态信道分配算法-随机多路访问控制

ALOHA：想发就发，两个帧时内无冲突 eˆ-2G，吞吐量：S=G*eˆ-2G

G=0.5 时，最好的信道利用率为 18.4%（G 每帧时的平均帧数）

S-ALOHA ：等到时间槽的开始才能发，无冲突的概率为 eˆ-G，G=1 时效率最高位 36.8%

CSMA ：载波侦听多路访问协议，发送前先侦听。

1 坚持：忙时一直监听，空闲就发，冲突使用二进制回避算法，减少空闲时间，增加冲突概率

非坚持，P 坚持

CSMA/CD ：（半双工）以太网使用，使用 CSMA 协议，发送期间检查到冲突就停止，随机等待一段时间

CSMA/CA：无线网使用

（4） 无冲突协议

bit-map：效率，轻负载为 d/(N+d),重负载下为 d/(d+1)

binary countdown

（5） 有限竞争协议

重负载时采用无冲突方法，而轻负载时采用竞争方式。

减少竞争站的数目，获取信道的概率

思路：将站分组，组内竞争

3、 局域网协议

（1）802.3 10M 以太网 传输介质、拓扑结构、最大帧长、二进制指数退避算法

传统以太网： 总线型

使用了曼彻斯特码

传输介质：10base5（粗同轴电缆，基带传输，500 米） 10base2 （细同轴电缆，185 米）10base-T（双绞线，100 米） 10base-F（2000）

数据大小：46-1500B 头部：18B 总长度：64（2t 限制）-1518B（6+6+2+4+数据）

CSMA/CD 流程：先听后发 边听边发 冲突停止 延迟重发

二进制指数退避算法（冲突后的等待事件，随机的选择）

小于 10 次 0-2^i-1，大于 10 次固定为 0-1023 16 次冲突后 放弃努力

（2）百兆以太网

传输介质、半双工和全双工工作方式

传输介质：双绞线或光纤（100Base-T4 100 米，100Base-Tx，100 米全双工，1000base-Fx，2000 米）

改进：帧格式，接口，规程不变 比特时间从 100ns 缩短到 10ns，缩短电缆长度

100Base-T4 8B6T 一对是 33.3Mbps 3 对则 100M

（3）G 比特（千兆）以太网 传输介质、扩大网络规模的方法、流控方法

允许全双工和半双工，使用 802.3 协议规定的帧格式，半双工需要使用 CSMA/CD

半双工模式：载波扩展：添加填充位 帧中继：多个待发帧级联（512 字节）

全双工模式不需要

传输介质：1000BASE-SX SX 表示短波长（550） 1000BASE-LX LX 表示长波长（5000）1000BASE-CX CX 表示铜线（屏蔽双绞线 25 米） 1000BASE-T（100 米，5 类 utp）

（4）10G 比特（万兆）以太网 传输介质、应用环境

（5）无线局域网 802.11 4、 局域网互联

重点是：CSMA/CA（物理侦听+NAV（网络分配向量）） + 确认机制

使用二进制指数退避算法

4、局域网互联

（1）网桥及其工作原理

逆向学习+扩散算法

见书 P259

（2）透明网桥

（3）交换机

5、 VLAN

VLAN 原理及协议

四、相关协议和设备

1、协议

802.1q、802.3、802.11

2、设备

集线器、交换机、网桥

五、重点和难点

1、CSMA 系列协议的工作原理及其性能

2、10Mbps、100Mbps、1000Mbps 以太网协议和无线局域网协议

3、网桥工作原理及协议

4、VLAN 技术

第五章 网络层

一、先行课程

概率论与随机过程、算法与数据结构

二、解决的主要问题

本章主要解决异构的物理网络如何通过软件协议（如 IP 协议）进行互联和控制的问题，即通过知晓通信子网的拓扑结构，通过子网选择适当的路径；同时，在选择路径时避免一部分通信链路和路由器超载，而另一部分链路和节点空闲；以及当源端和目的端处于不同网络时，处理它们之间的差异，并解决由此带来的问题。上述问题的解决包括了路由选择协议、网络控制协议的设计，以及广域网的规划和设置方法。

三、本章知识点

1、概念

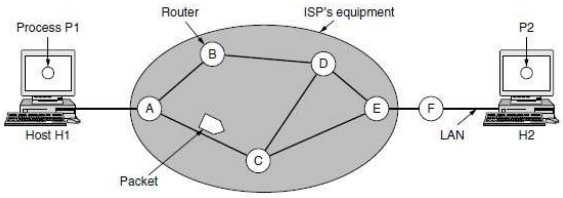
（1） 网络层的功能：**网络层的任务是将源计算机发出的数据分组（数据报）经过适当的路径送到目的地**

为什么要创建网络层？1. 为了实现端到端通信，最底层(链路层只是从介质的一端到另外一端，是点到点通信)

2. 解决路由选择 拥塞控制 网络互联

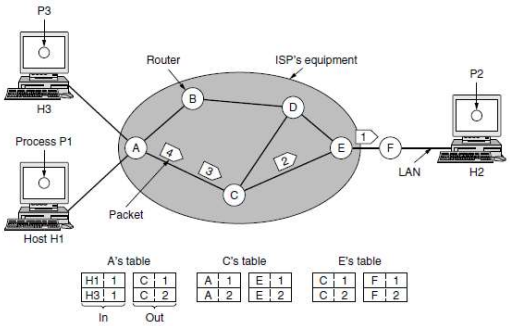
计算机，从源端到目的端可能要经过若干中间节点。这一功能与数据链路层有很大的区别，数据链路层仅把数据帧从线缆或信道的一端传到另一端。因此，网络层是处理计算机网络中端到端数据传输的最低层。

（2） 存储转发分组交换机制（Store-and-Forward Packet Switching） 如下图所示，一台主机要发送一个分组（数据报），那么它将分组传送给最近的路由器，该路由器或者在它自己的 LAN 上，或者在一条通向承运商的点到点链路上。该分组被存储在路由器上，一直到它完全到达路由器为止，所以路由器可以验证它的校验和。然后它被沿路转发到下一台路由器，直到到达目标主机为止，最后在目标主机上它被递交给相应的进程。

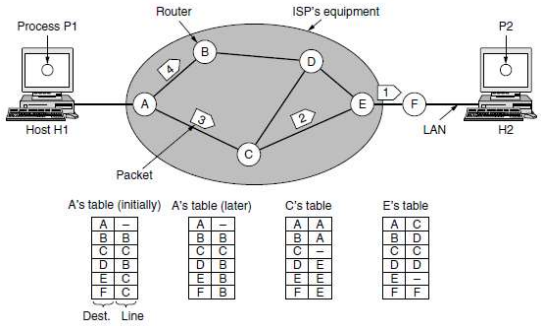


（3） 向传输层提供的服务：虚电路和数据报

- 虚电路：网络层向传输层提供了面向连接的服务，在发送分组（数据报）之前，必须首先建立起一条从源路由器到目标路由器之间的路径。该连接成为一个 VC（virtual circuit，虚电路）。



- 分组（数据报）：网络层向传输层提供了无连接的服务，所有的分组（数据报）都被独立地传送到子网中，并且独立于路由，不需要提前建立任何辅助设施。



虚电路和数据报的比较：电路建立，寻址，（连接）状态信息，路由方式，服务质量，拥塞控制，路由器失效影响

2、 路由选择算法（产生路由表的算法，网络层软件的一部分）

（1） 最优化原则

如果路由器 J 是在从路由器 I 到路由器 K 的最优路径上，那么，从 J 到 K 的最优路径也必定沿着同样的路由路径。

（2） 最短路径选择算法（Dijkstra 最短路由搜索算法，静态路由算法，非自适应算法）

首先为通信子网建立一个子网图，图中的每个节点代表一个网络节点（路由器），每条弧代表一条通信新路（链路），弧上的标注代表两个相邻节点之间的权值。然后把每个节点用从源节点沿已知最佳路径到本节点的或距离来标注。

- 算法思想：设 G=(V,E)是一个带权有向图，把图中顶点集合 V 分成两组，第一组为已求出最短路径的顶点集合（用 S 表示，初始时 S 中只有一个 源点，以后每求得一条最短路径 ，就将加入到集合 S 中，直到全部顶点都加 入到 S 中，算法就结束了），第二组为其余未确定最短路径的顶点集合（用 U 表示），按最短路径长度的递增次序依次把第二组的顶点加入 S 中。在加入的 过程中，总保持从源点 v 到 S 中各顶点的最短路径长度不大于从源点 v 到 U 中任何顶点的最短路径长度。此外，每个顶点对应一个距离，S 中的顶点的 距离就是从 v 到此顶点的最短路径长度，U 中的顶点的距离，是从 v 到此顶 点只包括 S 中的顶点为中间顶点的当前最短路径长度。
- 算法步骤：

A. 初始时，S 只包含源点，即 S＝{v}，v 的距离为 0。U 包含除 v 外的其他顶点，即:U={其余顶点}，若 v 与 U 中顶点 u 有边，则<u,v>正常有权值，若 u 不是 v 的出边邻接点，则<u,v>权值为∞。

B. 从 U 中选取一个距离 v 最小的顶点 k，把 k，加入 S 中（该选定的距离就是 v 到 k 的最短路径长度）。

C. 以 k 为新考虑的中间点，修改 U 中各顶点的距离；若从源点 v 到顶点 u 的距离（经过顶点 k）比原来距离（不经过顶点 k）短，则修改顶点 u 的距离值， 修改后的距离值的顶点 k 的距离加上边上的权。

D. 重复步骤 B 和 C 直到所有顶点都包含在 S 中。

（3） 扩散算法（洪泛、Flooding） 是一种静态路由算法。在扩散法中，每一个进来的分组将被发送到除了它进来的 那条路线之外的每一条输出线路上。 由于扩散算法

会产生大量的重复分组，需要改进扩散算法，避免重复的分组。

优点：不用事先知道拓扑信息，所有的路径都被尝试，第一个可能最优

解决方法：1. 头部加一个计数器，没经过一个计数器减一，为 0 时丢弃

2. 路由器记录已经扩散过的分组，源端放置一个序号，首端路由器记录{源地址，序号}，大于分组的记录，小于分组的丢弃

（4） **距离矢量路由选择**（Distance Vector routing） 是一种动态路由算法。在一个距离矢量路由选择算法中，所有的节点都**定期**地将它们的**距离表**（没更新前的表）传送给所有与之直接邻接的节点，包括：

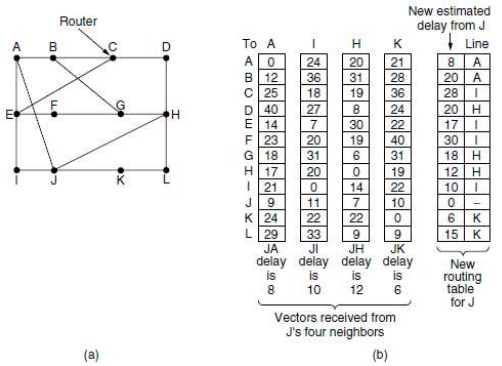
- 每条路径的目的地（另一节点）
- 路径的代价（距离）

如下图所示，所有的节点都监听从其他节点传送来的路由选择更新信息，并在下列情况下更新它们的路由选择表：

A. 被通告一条新的路径，该路由在本节点的路由表中部存在，此时本地系统加入这条新的路由；

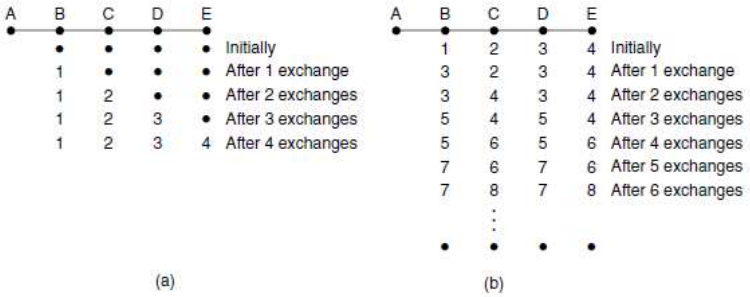
B. 通过发送来路由信息的节点有一条到达某个目的地的路由，该路由比当前使用的路由有较短的距离（较小的代价）。在这种情况下，就用经过发送路由信息的节点的新路由替换路由表中到达那个目的地的现有路由。

C. 在本节点的现有路由表中为了到达某一目的地首先应前往的下一节点如果通告了一个较高的代价，就要使用这一新的代价更新从本节点前往同一目的地的代价。



具体过程：周期性的测到邻居的距离，同时向相邻节点发送他到每个目的节点的距离（老表去掉下一跳（只包含目的地，花费），同时接受邻居发来的表）

距离矢量路由算法的主要缺点是网络规模的伸展性差。它对链路状态变化的响应慢，需要大尺寸的路由信息报文交换，并且报文的长度与通信子网内的个数成正比。由于距离矢量协议需要每个存储转发的节点都参与路由信息的交换，因而交换信息的交通量也可能巨大，并产生无穷计算（Count-to-infinity）问题。如下图所示。

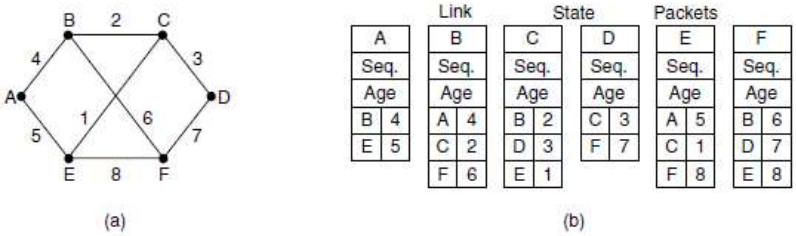


（5） **链路状态路由选择（OSPF 和 IS-IS 使用）**

是一种动态路由算法，通常包括以下 5 个步骤： A. 发现邻居节点 发现它的邻居节点，并知道其网络地址（发送一个 hello 包）；

B. 测量线路开销 测量到各邻居节点的延迟或者开销（发送一个 echo 包要求对方返回，往返时间/2）；

C. 创建链路状态分组构造一个分组，分组中包含所有它刚刚知道的信息；这个分组的内容包含了**发送方的标识**，以及是一个**序列号**（Seq）和**年龄**（Age），以及一个邻居列表。对于每个邻居，同时也要给出到这个邻居的延迟。两种构建时间：周期性，有重要事情发生。如下图所示。

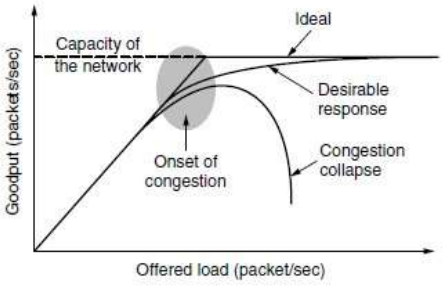


D. 发布链路状态分组使用**改进的扩散算法**，将这个分组发送给所有其他的路由器； E. 计算新的路由路径 计算出到每一个其他路由器的最短路径。

注释：序号是为了解决前面发生的浪费太多资源的问题，每发一包，序号加 1 年龄是为了解决路由器崩溃，序号发生错误等情况，每秒年龄减 1，减到 0 后丢弃

3、 **拥塞控制算法（对需求的总和>可用资源）**

（1）拥塞（congestion） 当一个子网或子网的一部分中出现太多分组导致数据报延迟和丢失，网络的性能开始下降。这种情况称之为拥塞，如下图所示，网络资源上有太多的分组时，将会导致网络性能下降，即对资源需求的总和大于可用资源。



产生拥塞的原因：

- 低带宽线路
- 多个输入对应一个输出
- 节点缓冲容量太小
- 结点处理机速度不高

拥塞控制和流量控制的区别：拥塞控制是一个全局过程，设计主机，路由器很多因素，流量控制，主要是解决发送方和接收方的问题，是局部过程，基于反馈进行控制。

（2）拥塞控制的通用原则

开环：预防性放大，闭环 闭环：网络供给，流量感知路由，准入控制，流量限流，负载丢弃。

包括开环的（open loop）和闭环的（close loop）方法。 完成开环的控制的手段：确定何

时接受新的流量、确定何时丢弃分组及丢弃那些分组， 以及在网络的不同点上执行闭环方案包括以下三个部分：

- 监视系统，检测到何时何地发生了拥塞
- 将该信息传递到能够采取行动的地方
- 调整系统的运行，以改正问题。

（3）虚电路中的拥塞控制

- 准入控制（admission control）：降低负载，在虚电路中控制新的连接建立
- 资源预留（traffic-aware routing）

（4）数据报子网中的拥塞控制

- 警告位：dnew=a*dold + (1-a)s s 是瞬时队列长度，d 大于某个阈值就需要注意拥塞（a=0.875）

  ECN（显示拥塞通知，警告位），输出达到告警状态，头部设置告警位，在目的地，告警位被复制到 ACK 分组，发给源端。 源端收到后降低传输速率，带有警告位的确认分组减少到规定值是，源端增加它的传输速率（沿途所以路由器排除问题后，传输速率才能增加上去）
- 抑制分组（choke packet）：路由器给远端返回一个抑制分组，并指出元分组的目标地址，原来的分组被打上一个标记，防止沿途其他路由器又重复产生抑制分组；元足迹收到抑制分组后，减少流量，一段时间后监听有无抑制分组，有则，进一步讲得，无则增加流量
- 逐跳（hop-by-hop）抑制分组：抑制包对经过的每个路由器都起作用，能缓冲拥塞除的拥塞，上游路由器要求有更多的缓冲区。

（5）负载丢弃 当路由器来不及处理分组而被淹没时，只要将这些分组丢弃即可。 负载丢弃采用葡萄酒（wine，旧的比新的好，比如文件传输）或牛奶（milk，新的比旧的好，比如视频传输）策略。

随机的早期检测（Random Early Detection，RED）：在实际耗尽所有的缓冲区空间之间就开始随机丢弃分组。RED 用在主机不能接收显示信号的环境里，如果 ECN 可用，有限选择它，提供了一个显示的拥塞信号，而不用丢包来识别。但 ECN 能提升性能。

（6）抖动控制

- 抖动（jitter）：分组到达的时间的变化量。
- 抖动控制：路由器通过计算出沿途每一跳的期望传输时间，就可以对抖动加以控制：当一个分组到达一台路由器的时候，该路由器对它进行检查，看这个分组比预定的时间来早了，还是来晚了。这些信息被保存在分组之中，每一跳都相应地更新。如果一个分组提早到达，则它尽可能多停留一会，以便回到预定的时间点上；如果一个分组比预期的时间到达得晚的话，则路由器尽可能快地将它转发出去。

4、 服务质量

（1）服务质量的评价标准

- 流（flow）：从一个源到一个目标的分组流（steam）称之为流。
- 评价标准：可靠性、延迟、抖动和带宽

（2）流量整形

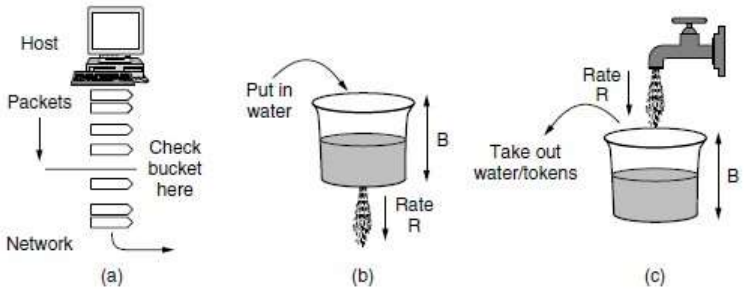
- 流量整形（traffic shaping）：在服务器端对流量进行平滑处理，即调节数据传输的平均速率（以及突发性）。
- 漏桶和令牌桶算法

漏桶（leaky bucket）算法：每个主机连接到网络的接口中都包含一个漏桶，即含有一个有限长度的内部队列。如果当该队列满的时候，又有一个分组到来，那么该分组将被丢弃，亦即是，如果在一台主机上，队列中的分组数目已经达到了最大值，这是又有

一个或者多个进程要发送分组，那么新发送的分组将被丢弃。

令牌桶（token bucket）算法：桶中保存的是令牌，每个令牌包含一定数量的字节的分组，令牌桶允许主机积累发送全，直至到达桶的最大尺寸，因而相对于漏桶，允许有一定的突发流量。

有一个计算令牌桶突发速度传送时间：C/(v1-v2)



（3）分组调度（预约 3 种资源：带宽 缓冲区 CPU 资源）

在同一个流的数据包之间以及在竞争流之间分配路由器资源

FIFO（FIFS，尾丢弃）

- 公平排队（fair queueing）：路由器为每一条输出线路使用一组单独的队列，每个流一个队列。当一条线路空闲的时候，路由器轮流扫描这些队列，从下一个队列中取出第一个分组。使得所有的流量以相同的速率发送数据包（缺陷：他给使用大数据包的主机比使用小数据包的主机提供了更多的带宽）。
- 加权的公平排队（weighted fair queueing）：通过不同的主机赋予不同的优先级进行加权调节，从而提高算法的效率。

Fi=max(Ai，Fi-1)+Li/W

5、 网络互联

（1）互联网络的分组转发

级联虚电路：优点：

路由器预留缓冲区等资源，保证服务质量

分组按序号传输；

分组头短

缺点：

路由器需要大量内存，存储虚电路信息；

一旦发生拥塞，没有其它路由；

健壮性差；

如果网络中有一个不可靠的数据报子网，级连虚电路很难实现。

无连接网络互联：优点：

能够容忍拥塞，并能适应拥塞；

健壮性好；

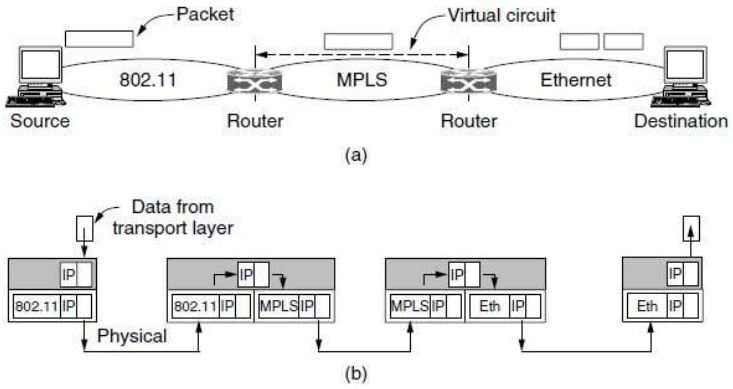
可用于多种网络互连。

缺点：

长包头；

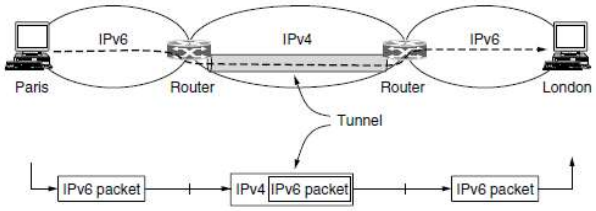
包不能保证按序号到达；

不能保证服务质量。



（2）隧道技术

隧道（tunneling）：当源和目标主机位于相同类型的网络中，中间的网络属于不同类型时使用隧道方案。



（3）互联网路由

- 互联网定义了两级路由算法：在每个网络内部使用的内部网关协议（interior gateway protocol）和在网络之间使用的外部网关协议（exterior gateway protocol）。
- 自治系统（AS，Autonomous System） 自治系统：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。

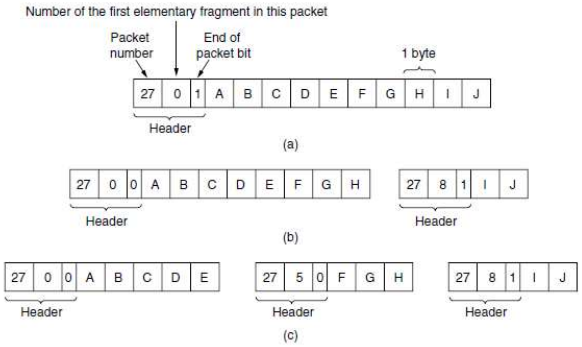
- 因特网有两大类路由选择协议

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）分段和重组 分段：互联网协议定义一个足够小的基本分段长度值，以便于基本的分段能够通过 每一个网络。当一个分组被分段的时候，除了最后一个分段以外其他所有的分段都 等于基本分段长度，而最后一个分段只会更短。 重组：在互联网分组头部有两个序列号域：原始的分组号和分段号（利用分组好区分是不是同一个原始报文）。并且还有一位

来指明：该互联网分组中包含的最后一个基本分段是否为原始分组的最后一个分段。 重组的过程如下图所示。

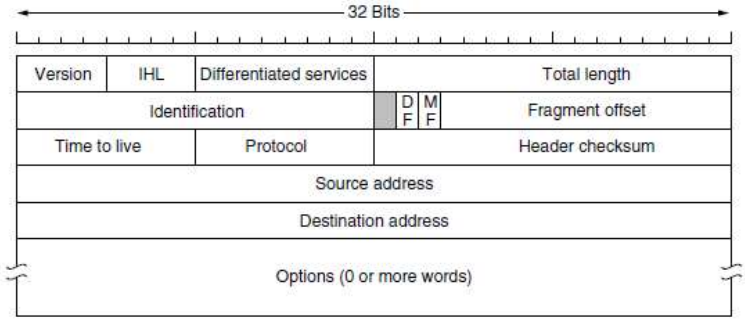


6、 Internet 上的网络层

（1）IP 协议

- IP 分组（包）

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 分组（数据报）结构如下图所示。



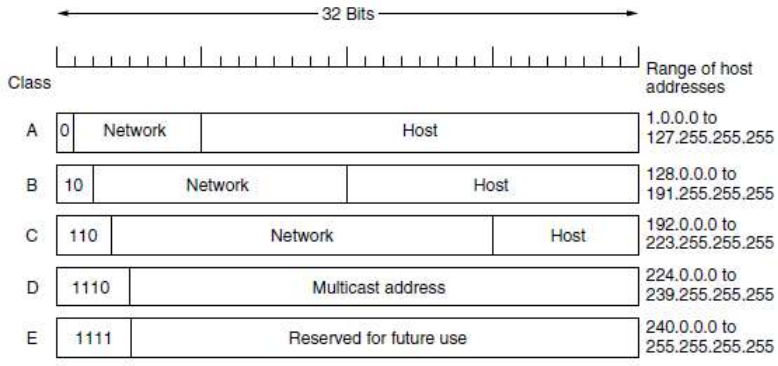
（2）IP 地址

- 分类 IP 地址

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）。

IP 地址 ::= { <网络号>, <主机号>}

如下图，分类 IP 地址的结构



（3）地址解析协议 ARP

不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)，里面有所在的局域网上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。

（4）子网

●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。

●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

A. 从收到的分组的首部提取目的 IP 地址 D。

B. 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)。

C. 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)。

D. 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)。

E. 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)。

F. 报告转发分组出错。

（5）无分类编址 CIDR

- CIDR 最主要的特点

A. 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。

B. 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。

IP 地址 ::= {<网络前缀>, <主机号>}

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。

- 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。

四、相关协议和设备

1、协议

IP 协议、IP 地址、ICMP 协议、OSPF、BGP、IPv6、IGMP、ARP 协议、DHCP 协议

IP: version;4bit

IHL：首部长度，单位 4 字节

Total length：小于 65535，包括首部+数据

Identification：16bit，数据包的标识（数据报分片的时候需要考虑）

基本是 20 字节：

标志：3bit 中间 1 位是 DF（为 0 允许分片） 最后一位是 MF（为 1 代表还有分片，为 0 代表最后一个分片）

Fragment offset：13bit，最多支持 8192 段，单位为 8 字节，不包括头部！！！

TTl：生存时间每经过一个路由器减 1，（导致每次校验和和生存时间都会变）

Header Checksum：只校验首部（所以 TCP，UDP 提供了校验数据的功能）

Source IP address：32bit，8 个 16 进制位，最后一道题的倒数第二个分组，路由器只看源地址，与掩码做与运算

Destination IP address：

NAT（网络地址转换）：解决 IP 地址不够的一种方法，一个全球地址，一个本地地址
违反了 IP 结构，打破端到端模型，必须维护一个映射关系，违

背分层协议 10.0.0.0-10.255.255.255/8 172.16.0-172.31. 192.168.0-

ICMP：因特网控制报文协议，是网络层的协议，作为 IP 层数据包的数据，加上数据报的首部，组成 IP 数据报发送出去。分为 ICMP 差错报告报文和 ICMP 询问报文
5 中类型：重点不可达，源站抑制，时间超过，参数问题，改变路由
发过一次就不再发了，第一个分片以后的都不再发了，多播不发，特殊地址不发
Ping 用来测试连通性，使用了 ICMP 回送请求和回送回答报文。是应用层直接使用网络层 ICMP 的例子，没有通过 TCP 或 UDP

OSPF：开放最短路径有限，分布式链路状态协议
1. 使用洪泛 2 发送与本路由器相邻的所有路由器的链路状态 3 只有链路状态发生变化才发送次信息

BGP：发言人要交换路由需要先建立 TCP 连接

解决 IP 地址耗尽：使用 CIDR，使用网络地址转换 NAT 方法节省全球 IP 地址，使用 IPv6

2、设备 路由器

五、重点和难点
1、路由算法和路由协议
2、IP 协议和 IP 地址
3、拥塞控制
4、服务质量
5、网络互联

第六章 传输层

一、 先行课程

计算机组成原理、数据结构、操作系统

二、解决的主要问题 本章描述传输层的基本功能和服务、协议实现的要素、以及因特网的传输层协议——TCP 和 UDP 的主要原理。

三、本章知识点

1、主要概念

引入网络层的原因：消除网络层的不可靠性；提供源主机到端主机的可靠地，如实际使用网络无关的信息传输；提高服务质量，进行差错检测。

与网络层的区别：传输层提供应用进程间的逻辑通信，而网络层只是实现主机到主机之间通信

（1）端到端通信：传输层通过使用端口号做标识，向应用层提供了进程间的逻辑通信的功能。

（2）端口号：16bit,因特网的传输层地址，区分应用层的不同进程，端口号能复用，服务器端采用固定的常用端口，客户端采用由操作系统动态分配的短暂端口。常见的端口（熟知端口 0-1023（不能被使用）HTTP：80 FTP：21 DNS：53 SMTP：25 TELNET：23） 联系：MAC 48 位标识主机 IPV4 32 位，IPV6 128 标识网络

套接字，IP 地址+端口号，越界的使用了其它层

寻址：定义传输服务访问点，将应用进程与这些 TSAP 相连（1.广为人知的 TSAP 2.使用初始连接 ICP（进程服务器作为不常使用的服务进程代理） 3.端口映射器（目录服务器））

（3）三次握手：传输层的连接建立和连接释放方式。

三次握手方案解决了由于网络层会丢失，储存和重复包带来的问题。

建立链接：前两次 SYN 为 1，后两次 ACK 为 1

释放连接：本来是 4 次握手，由于第一个 ACK 和第二个 FIN 可以合并，可以缩减到 3 个（考试用 3 个）+定时器

连接释放：三次握手+定时器

（4）拥塞控制：当网络中的负载超过网络资源（路由器的处理能力和缓存）时，将会出现传输时延过长甚至丢包的情况，即网络拥塞。拥塞控制是指通过采用某种策略，避免拥塞，或者在拥塞出现时缓解拥塞。

（5）最大最小公平性：在端到端的共享链路上，最大限度保证每个数据流的最小带宽需求。

公平性定义很多，

尽可能给某个用户分配速率，但是不能从速率比这个用户低的其他用户身上剥夺速率来分配给这个用户

因此，为了增加某个用户的速率，只能降低另外一个速率比它高

的用户的速率

（6）伪报头（Pseudo-header）：TCP 和 UDP 在计算校验和时，需增加 12 个字节的伪报头（包含源 IP 地址和目的 IP 地址等），伪报头不会传输到网络中。

端口和 IP 地址多路复用

（7）（向上）多路复用：多个端口使用一个 IP 地址传输 TPDU

向下（逆向）多路复用：一个端口使用多个 IP 地址进行传输 TPDU（流量均衡，不常用，流控制传输协议 SCTP）

（8）崩溃恢复：N 层崩溃，N+1 层恢复

（9）传输策略：
基于确认和可变窗口大小
窗口为 0，紧急数据可以发送，，为防止死锁可以发送一字节的 TCP 段
提高 TCP 传输效率的方法：1. 发送发缓存应用程序的数据，等到形成一个比较大的段再发出 2.接受法延迟发送确认段 3. 发送发使用 Nagle 算法（发送方每次发的数据比较小，避免发送太小的数据段）
（1）当应用程序每次向传输实体发出一个字节时，传输实体发出第一个字节并缓存所有其后的字节直至收到来自于接收方的对第一个字节的确认；
（2）然后将已缓存的所有字节组段发出并对再收到的字节缓存，直至收到下一个确认；如果缓存的数据填满一半的窗口或是最大数据段时也可发送；
愚笨(低能)窗口综合症：症状，应用程序一个一个的读出，传输实体产生一个一字节的窗口更新段，使得发送发只能发一个字节，解决使用 Clark 算法：限制受访只有具备一般的空缓存或者最大段长的空缓存时，才产生一个窗口更新段。（）
避免亲戚也太小的数据段
2、算法
（1）TCP 的拥塞控制算法

A. 发现拥塞

- 隐式拥塞通知：源 TCP 通过丢包或者收到三个重复的 ACK 来判断出现拥塞
- 显式拥塞通知：TCP with ECN，路由器设置 ECN 通知目的 TCP，目的 TCP 设置 ECE 通知源 TCP

B. 拥塞控制

- 慢启动：启动速率很低，拥塞窗口初始值为 1 个最大报文段长度（MSS）。在达到阈值或者发现丢包之前，拥塞窗口按指数级增长
- AI：当拥塞窗口达到阈值时，按线性增长
- MD：出现重发定时器超时，拥塞窗口降为最小值（1 个 MSS），阈值改为当前窗口的一半，重新开始新的慢启动
- 快速恢复：收到三个重复的 ACK 时，拥塞窗口降为当前窗口的一半，开始新的 AI。

（2）Nagle 算法
TCP 的发送端使用此算法来提高传输效率，预防大量短报文段引起拥塞，一次尽可能发送较大的数据量（1 个 MSS）。
（3）Clark 算法
TCP 的接收端使用此算法来提高传输效率，预防大量短报文段引起拥塞，只在有较大缓存（至少为 1/2 空闲缓存或者一个 MSS）时，才发送窗口更新通知。
（4）Jacobson 算法 用于估算端到端环回时延 RTT（往返时延）。

RTT 估值 $= \alpha \times$ RTT 估值的历史值 $+ (1-\alpha) \times$ RTT 的测量值（典型情况 a=7/8）

RTO=β*RTT（RTO 超时重传时间，β 推荐取 2）

Karn 算法：及孙平均返回时延 RTT 时，只要报文段重传了，就不采用其样本，连续重传的超时间隔加倍。

持续计时器（接收端发送一个窗口大小为 0 的确认），保活计时器（半链接，当连接空闲了一段较长的时间）TIMED WAIT state(连接释放，两倍最大数据包生存周期)

3、传输层协议

（1）TCP 提供面向连接的、可靠的字节流服务，不保证应用层的消息边界；全双工点对点通信，无丢失无重复，无乱序；采用三次握手建立连接、 三步释放连接；默认采用 Go-back-N

ARQ 协议进行差错控制，通过选项 SACK 实现选择重传 ARQ；采用动态的滑动窗口进行流量控制；提供拥塞控制功能，发送窗口取决于接收窗口和 拥塞窗口的较小值。

TCP 的头部（重点的重点），基本头部是 20 字节：

source/destination port 2B（端口，传输层与应用层的 TSAP（传输层的服务访问点））

Sequence number：4 字节（保证不会出现序号回转），是第一个字节的编号（每个字节都编号，等于收到的上一个 ACK 序号，双方商定初始序号）

Acknowledgement number：4 字节期望收的下一个报文段的第一字节（等于收到的最后一个+1），采用捎带确认和累计确认

TCP header length：占 4 位，它指出数据起始距离离 TCP 报文段的起始处有多远（头部长度），4 字节为单位（和 IP 一样）

保留：6 位！！！注意（看 ACK 的时候注意清除）

URG：URG=1 表示紧急指针字段有效，此报文有紧急数据，应该尽快传送。

ACK：为 1，代表前面的确认号有效。

Psh：为 1，尽快交给应用程序

RST=1，表明 TCP 连接中出现严重差错，需要释放连接，然后重新建立运输连接。

SYN：为 1 代表这是连接请求或者连接接收报文（三次握手中前两次为 1）

FIN：为 1，代表发送端的数据已经完毕，并要求释放运输连接。

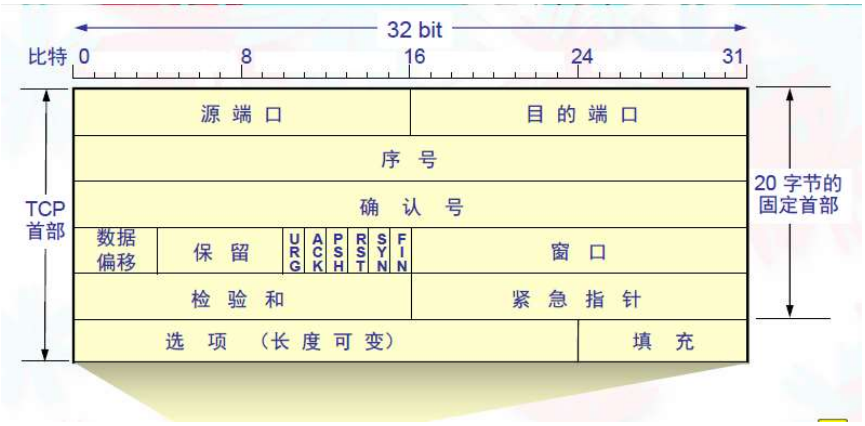
Window size：2B，窗口字段用来控制对方发送的数据量，单位为字节。TCP 连接的一端根据设置的缓存空间大小确定自己的接收窗口大小，然后通知对方以确定对方的发送窗口的上限

正常情况是 2^16-1=65535 字节 利用扩展项：65535*2^14=1073725440 字节。

校验和：2B，校验范围包括 首部+数据+伪首部 补码相加 和求补 必须做

Urgent pointer：与 URG 配合使用，指出紧急数据的最后一个字节的序号。

Options：TCP 中即为最大报文段长度 MSS！！告诉对方 TCP，我能接收的最大报文段的数据字段的最大长度是 MSS 字节。数据字段加上 TCP 首部才等于整个的 TCP 报文段。所有主机都要求能接受 536+20=556 字节的 TCP 段



（2）UDP 提供无连接、尽力而为的传输服务，采用校验和方式进行差错检测，不提供差错恢复功能。只提供端口的功能和差错检测的功能。基本头部为 8 字节

基本格式：

源端口：2B

目的端口：2B

长度：2B

检验和：2B（头+数据+伪首部（源 IP，目的 IP，0,协议号，UDP 长度（包括头的字节计数））

UDP 的典型应用：DNS，DHCH，RIP，RPC（远程过程调用），RTP/RTCP（IP/多媒体传输）

NFS（远程文件服务器），p2p

四、相关协议和设备

1、协议 TCP、UDP

2、设备 无

五、重点和难点

1. TCP 的连接建立和连接释放

2. TCP 的拥塞控制原理

3. 最大最小公平性