## 4. 1 概述 网络层的主要功能：选路(ms)、转发(ns)

选路:确定往目标路由器的路由(计算转发表)；转发:路由器根据转发表将分组从输入端口转移到输出端口；转发表:记录分组头中某个字段与路由器输出端口之间的映射关系

**[数据面与控制面](部分模型中划分管理面)**

• 数据面负责数据传输(转发)，在路由器内部进行分组转运(路由器本地功能)<唯一功能>

• 控制面负责控制数据传输(选路)，确定分组如何前往目的节点(网络范围的功能)

**[网络服务模型：定义了分组在发送终端与接受终端之间的端到端传输特性]**

网络层能够提供的服务: (具有延迟上界的)确定交付，对于分组流，网络层提供有序分组交付、保证最小带宽(低于带宽发送则保证交付)、保证最大延迟抖动、安全性服务等。不同架构网络提供的网络层服务可能不同，同一网络也可提供不同网络层服务。

一些例子：恒定比特率 ATM 网络服务 CBR、可用比特率 ATM ABR

**第四章 网络层：数据面**

## 4. 2 路由器工作原理(匹配-动作抽象)

**[路由器的主要功能]**选路功能：运行选路协议，计算转发表。控制面(软件)，毫秒级；

转发功能：依据转发表转发数据报。数据面(硬件)，纳秒级

**[路由器的组成]**

**输入端口**:物理层负责将物理信号转化为比特流，链路层负责从比特流中提取帧并解封 IP 数据包，网络层负责后续处理:【查表:每块线卡上都有转发表的一个镜像，以及相应的处理器缓存。路由器仅在本地查表】【排队:当交换结构阻塞时，输入分组在这里排队】【转发:通过交换结构将分组发送到各个输出端口(也就是交换)】

**交换结构**:路由器中的互联网络，用于在输入端口、输出端口和选路处理器之间转运分组。速率通常是输入输出链路的若干倍。【经内存交换:早期路由器在 CPU 控制下，将数据包拷贝到内存中进行交换。**交换速率受限于内存带宽:每个数据包都需要传输两次**。现代路由器每个端口使用一个内存接口硬件连接到存储系统，采用控制器硬件在端口之间传输控制消息(而不是 CPU)。输入端口将数据包放入内存后，由接口硬件告知输出端口从指定位置获取数据包并返回响应消息。】【经总线交换:数据包经由共享总线，从输入端口缓存转移到输出端口缓存。每个端口由接口硬件连接到总线上，并分配一个内部标签。**交换速率受限于总线带宽**，常用于接入路由器、总线路由器。引入的问题:总线竞争，输入端口在总线上轮流广播分组、多路复用】【经互联网络交换:交换结构控制器通过控制交叉点的开闭，在输入输出端口之间建立专属电路，多对端口之间可以并行传输。分为阻塞与非阻塞类型。先进设计:将分组划分为若干定长信元，同时送入若干并行交换结构，离开后再组装成分组】

**输出端口**:网络层负责预先处理:【组装:根据需要将交换结构输出的信元组专程分组】【排队:若输出端口来不及传输，则输出分组在此排队】【调度:输出端口每次选择一个分组发送】。链路层封装并执行链路层协议，物理层将比特流转化为物理信号。

**输入输出端口在物理上是合二为一的:全双工设计**

**[排队与丢包]**

**输入端口**:当交换结构不能及时将输入端口分组转移到输出端口分组时便会产生排队。带来的问题有:队头阻塞(队头分组阻塞后续分组原本的正常转发)、丢包(输入队列满后发生丢包)。当交换结构速率为端口速率的 n 倍(n 为输入端口数目)可以消除输入端口的排队，但提高了路由器成本

**输出端口**:当多个输入端口同时向一个输出端口发送时，在输出端口形成排队。**这是不可避免的**:若增大输出队列，可以减少丢包的发生，但会增加内存消耗，增大分组延迟(排队时间过长的分组最终会被重传)。因此输出队列不是越长越好

**[输出端口的排队缓解策略]** 非抢占式排队:分组一旦开始传输，就不可被中断

• 分组丢弃策略(当输出队列满时丢弃哪个分组):弃尾(丢弃到来的分组)、优先级丢弃(丢弃低优先级分组)、随机(随机丢弃分组)

• 主动队列管理 AQM: 当网络出现某些拥塞征兆时就主动丢弃到达的分组。随机早检测算法 RED:路由器在每个端口上维护输出队列的平均长度: AvgLen = (1-Weight)×AvgLen +Weight×SampleLen。当平均队列长 度达到第一个阈值 minth 时，按照概率 p 丢弃到来的分组；当平均队列长度达到第二个阈值 maxth 时，丢弃每一个到达的分组。概率 p 是平均队列长度和上一次丢弃距当前时间的函数，分组队列长度越大，p 也越大。实际 RED 可与 TCP 拥塞控制一起使用：达到低阈值时，丢弃一个包而触发快速重传，发送速率降至一半；达到高阈值时，丢弃每个包触发超时，发送速率降至最低。

**[调度策略：输出端口如何从多个队列中选择下一个发送的分组]**先来先服务 FIFO、加权公平排队 WFQ(按照每个队列的带宽权重选择一定数量的分组发送)、优先级调度(根据分组头中的某些域 IP、port 设置优先级，同一优先级 FIFO)、轮询调度(若干队列之间轮流服务，每次选择一个队头分组发送)

**[动态主机配置协议 DHCP：应用层协议]**

• 主机/路由器如何获得 IP 地址？路由器：管理员手工配置路由器各个接口的 IP 地址。主机：管理员手工配置主机 IP 地址，服务器通常采用这种方法(问题：更换网络时需要重新配置)、使用动态主机配置协议 DHCP 获取 IP 地址、子网掩码、第一跳路由器(默认网关)、本地 DNS 服务器等配置信息，每次主机被分配的地址可以相同或不同(个人终端通常采用)。使用 DHCP 的好处：免去手工配置的麻烦(即插即用)可用少量的 IP 地址服务较多的客户(地址重用)，适用于主机频繁加入和离开的情况。

**一个子网中可能有多个 DHCP 服务器，也可能没有 DHCP 服务器(管理员手动配置)**

**• DHCP** 目标：允许主机加入网络时自动获取配置信息。DHCP 是一个 CS 架构的应用协议，子网中应有一个 DHCP 服务器或一个 DHCP 中继代理(通常为路由器)，该代理知道用于该网络的 DHCP 服务器的地址。DHCP 协议的四个步骤：**1.DHCP 服务器发现**：主机广播“DHCP discover”报文，寻找子网中的 DHCP 服务器，链路层将该帧广播到所有与该子网连接的子网；**2.DHCP 服务器提供**：DHCP 服务器用“DHCP offer”报文进行响应；**3.DHCP 请求**：主机用“DHCPrequest”报文选择一个 DHCP 服务器请求 IP 地址，回显配置参数 **4.DHCP ACK：**DHCP 服务器用“DHCP ack”报文发送 IP 地址，服务器响应客户的请求，确认所要求的参数。DHCP服务器使用UDP端口68，客户使用UDP端口67，初始地址 0.0.0.0:xx，广播地址 255.255.255.255:xx。DHCP同时返回第一跳路由器地址(缺省网关、边缘路由器)、本地 DNS 服务器地址、地址掩码。DHCP允许客户更新它对一个IP地址的租用。DHCP被封装到UDP/IP/以太帧中。DHCP不足之处：从移动性角度看，结点移动时不能维持与远程应用间的 TCP 连接

## 4.3 IPv4 : 网络层数据面协议 (网络层协议：选路协议、IP 数据报协议、ICMP 控制协议)

IP 负责将数据报交付到目的地址和目的协议。其本身并**不提供任何服务承诺**，但仍然努力解决分组在网络中传输可能遇到的一些问题: 数据报过大-对数据报分片、数据报在网络在循环-设定最大转发次数、报头出错导致误投递-对报头检错。

**[IPv4 分片与重组] 【MTU:** 链路层帧能承载的最大数据字节数。不同类型链路 MTU 可能不同，如以太网 MTU=1500 B,有些广域网 MTU=576 B。】【**分片:** 传输过程中较大的 IP 数据报可被分片，将**载荷**划分为若干较小数据块，每个数据块封装成一个独立的数据报传输。数据在传输的过程中可被**多次分片**，但**仅在目的主机上重组**。分片报头: 大部分内容取自原始数据报。**以下字段**:**标识**(每个分片携带与原始数据报相同的标识)、**偏移量**(分片中数据在原始数据报载荷中的位置)、**标志位**(MF 更多分片，最后一个分片为 0，DF=1 表示不许分片，对大小超过 MTU 且 DF=1 的分组直接丢弃)、**TTL**(原数据报 TTL-1)、**头部检查和**(重新计算)。**分片的数据长度**:假设原始数据报报头长度为 H，分片的长度为 N，则 H+N≤MTU。由于偏移量域只有 13 b，除最后一个分片外，其余分片的数据长度应为 8B 的整数倍(报文长度域 16 b，因此对应 2^3=8)。N 应当为满足上述条件的最大整数。】【**重组:** 目的主机将收到的分片重新组装成原始数据报。可根据[源 IP,标识]确定属于同一数据报的分片，利用最后一个分片计算原始数据报总长度(L = Offset\_last \* 8 + L\_last，L 为分片长度)。为了防止分片迟迟未到，每个重组的数据报都需要重组缓冲区与重组定时器。】分片与重组引入的问题: 增大了开销(降低路由器吞吐量、消耗目的主机的资源)、针对分片的 DoS 攻击(攻击者发送一系列分片消耗目的主机资源)。为此 IPv6 取消了路由器分片，确保源主机发送数据报不超过 MTU。

**[IPv4 编址]** 网络接口：是主机/路由器与物理链路的边界。路由器有多个接口，主机一般有一至两个：以太网接口通过以太网交换机连接，无线 WIFI 接口通过基站连接。IP 地址：每个网络接口关联一个 IP 地址(32 位二进制数，通常用点分十进制数表示)。

**<基于类的编址(早期方式)>** A 类 8b(0-)子网地址+24b 网内地址，B 类 16b(10-)+16b，C 类 24b(110-)+8b，D 类(1110-)+广播地址，E 类(1111-)+保留后续使用。

• **单播地址结构**：单播地址除类别标识外，其余比特被划分成**网络号**和**主机号**两部分。**网络号：**标识一个物理网络；**主机号：**标识该物理网络上的一个网络接口。同一个物理网络上的网络接口，它们的 IP 地址具有相同的网络号。

**• 地址分配：**因特网中的每个接口必须具有唯一的 IP 地址(在 NAT 后面的接口除外)，为在因特网范围内保证 IP 地址的全局唯一性：网络号由 ICANN 统一分配，主机号由网络管理员统一分配。建立私有网络的组织可以自己选择网络号，但同样必须保证**每个网络号在私有网络内的唯一性**。

• **特殊的地址：**全 0 或全 1 的网络号及主机号是特殊地址，不分配给特定的网络接口；

网络号有效、主机号全为 0 的地址保留给网络本身；**网络号有效、主机号全为 1 的地址保留作为定向广播，即在网络号指定的网络中广播(仅用作目的地址)**；32 位全 1 的地址代表本地广播地址，表示仅在发送节点所在的网络中广播(仅用作目的地址)；32 位全 0 的地址用于指示本机(仅用作源地址)；网络号为 0、主机号有效的地址用于指代本网中的主机；形如 127.xx.yy.xx 的地址保留作为回路测试，发送到这个地址的分组不输出到线路上，而是送回内部的接收端。 • **网络数量与地址数量**：A 类地址个数 2^7-2=126，提供接口数 2^24-2=16777214，B 类地址个数 2^14-2=16382，提供接口数 2^16-2=65534，C 类地址个数 2^21-2=2097152，提供接口数 2^8-2=254

**[子网] 具有相同子网地址、不需要通过路由器就可以相互到达的网络接口构成一个子网。子网内部没有路由器连接的接口，而是以太网交换机或无线接入点。**

**• 子网编址：**从概念上说，引入子网仅略微改变了 IP 地址的解释。主机号被进一步划分成子网号和主机号两部分；子网号标识网络内的一个子网，主机号标识子网中的一个网络接口；管理员可以根据需要，如子网的数量和规模，选择合适的子网号长度。

**• 子网掩码：**用于指示 IP 地址中子网号与主机号的边界。子网掩码是一个 32 比特的数，其中对应主机号的比特为 0，其余比特为 1；子网掩码也采用点分十进制表示，如 255.255.252.0。注意：子网地址≠子网号，子网地址包括主机号之前的所有比特。如何从 IP 地址中获取子网地址(网络地址)：将 IP 地址与子网掩码做“与”运算，例如 128.10.1.1 AND 255.255.255.0 = 128.10.1.0。子网掩码将 IP 地址划分为两部分：子网地址(对应子网掩码中“1”的部分)、主机地址(对应子网掩码中“0”的部分)。

**• 如何确定子网？**将网络接口与主机/路由器分开，形成一些分离的网络岛，每个网络岛就是一个子网。每个子网内部具有相同的子网地址，子网内部不包括路由器，之间被路由器隔开。路由器的每个端口连接一个子网，不同的端口连接不同子网。路由器是在子网间转发数据包的设备。子网内通信不需要通过路由器，子网间通信必须通过路由器。

**[IP 数据报转发]**

**• 网络层转发数据报的两种情形：**直接交付：节点(源主机或目的路由器)将数据包直接发送给目的主机(不需要其它路由器转发)、间接交付：节点查找转发表，将数据包转发给下一个路由器去处理。数据包的目的地址与节点的某一端口在同一个子网中-直接交付；数据包的目的地址不与节点的任何一个端口在同一个子网中-间接交付

**• 转发表：**转发表记录目的地址到输出端口的映射关系。每个转发表项包括：目的地址/掩码、下一跳地址、输出端口等。下一跳地址必须与输出端口在同一个子网中(不需要通过其它路由器就可以直接到达)。根据目的地址类型不同，分为三种转发表项：地址前缀表项(目的地址是一个子网地址)、特定主机表项(目的地址是一个特定的网络接口地址)、缺省项(不匹配所有其它表项的地址，这些地址被映射到一个默认的路由器端口)。IP 采用逐跳选路，每个转发表项只记录去往目的地址的下一跳信息(下一个要到达的路由器端口)，而不是一条完整的端到端路由。

**• IP 数据报的转发过程(主机/路由器)：**从数据报中提取目的 IP 地址 D，利用子网掩码计算地址前缀 N。 if D 与自己的任何一个 IP 地址匹配(本节点是数据报的目的点)then 将数据包交给 protocol 域指定的协议实体处理； elseif N 与自己的任何一个直连网络的地址匹配(直接交付)then 通过该直连网络把数据包直接交付到目的节点 D； elseif 表中包含到 D 的特定主机表项或包含到 N 的一个地址前缀表项(间接交付) then 把数据包发送到表中指定的下一跳； elseif 表中包含一个缺省项(间接交付) then 把数据包发送到表中指定的默认路由器端口； else 宣告选路出错，向数据包源地址发送一条错误报告消息(ICMP)

**[CIDR：无类域间路由]** 分类编址的缺点:只能按照三种固定的大小分配地址空间，地址浪费严重(尤其是 A、B 类地址)、转发表必须记录每个已分配的网络，转发表规模爆炸式增长(C 类地址网络非常多)。**CIDR 按照实际需要的地址数量分配地址空间，提高地址使用效率；允许将若干条转发表项进行聚合，减小转发表规模。**

**• 按照实际需要分配地址：**若一个网络需要 2000 个地址，可为其分配一个具有 2048 个连续地址的地址块。这些地址的前 21 位必须相同，从而可将其看成是一个具有 21 位子网地址的网络。CIDR 地址分配的原则：地址块的长度 L 必须是 2 的幂次；所有地址的前(32-log\_2\_L)位必须相同。网络地址的表示方法：用掩码指示网络地址的长度，如 194.24.0.0，255.255.248.0；用“/长度”指示子网地址的长度，如 194.24.0.0/21。组织网络外部的路由器仅考虑前面的前缀比特 x，减少了转发表的长度

*IP 广播地址：255.255.255.255，可能会被转发到临近的子网。*

机构如何获得网络地址？机构通常从 ISP 的地址空间中分配地址。原则：优先确定前缀部分的内容，保证同一划分前缀部分相同。 *ICANN 组织：分配 IP 地址；管理 DNS 根服务器；分配域名与解决域名纷争*

**• 地址聚合**：使用单个网络前缀通告多个网络的能力。转发表中符合以 下条件的若干个表项可以合并成一个表项：这些表项的目的地址可以聚合成一个前缀更短的地址；这些表项使用相同的下一跳。地址聚合的过程可以递归进行。若个别表项不满足路由聚合的条件，可以单独给出不能被聚合的表项，并聚合其他能聚合的表项。

**最长前缀匹配：在所有匹配的路由表项中，选择前缀最长的表项。**查找转发表：使用分类地址：转发表分为 A、B、C 三张表，分别记录 A、B、C 三类地址的转发表项，用哈希表组织。路由器收到数据报后：根据目的地址的类型确定要查找的转发表，根据目的地址的类型提取网络地址，用网络地址在相应的转发表中进行哈希查找(精确匹配)。

• 采用 CIDR 后出现的问题：为与某个转发表项 Dest\_addr/prefix\_len 进行匹配运算，路由器需要先从表项中读出地址掩码(或 prefix-len 值)，计算包的目的地址前缀(用地址掩码和包的目的地址相与)，与 Dest\_addr 的地址前缀(Dest\_addr 与地址掩码相与)比较

引入的问题：地址前缀的长度 prefix\_len 可以是任意值，其法从地址本身得到，只能从转发表项中得到，且必须从所有匹配的表项中选择前缀最长的表项。在大规模转发表中进行快速查找是一个难题(已经解决)

## 4.4 IPv6：改进 引入动机：IPv4 地址将很快耗尽；同时借机简化头部格式，加快数据报处理和转发；支持服务质量；支持多播；支持移动性；增强安全性。IPv6 与 IPv4 不兼容，但与其它所有因特网协议都兼容。IPv6 地址：128 位，使用冒号十六进制表示，每 16 位以十六进制的形式写成一组，组之间用冒号分隔。可将连续的多组 0 压缩为一对冒号，但一个地址仅压缩一次。IPv6 定义了三种地址类型：单播地址：一个特定的网络接口；多播地址：一组网络接口；任播地址：一组网络接口中的任意一个(通常是最近的一个)

**[从 IPv4 过渡到 IPv6] 双协议栈方案：**支持 IPv6 的主机和路由器同时运行 IPv4 和 IPv6；运行双栈的源节点先对目的节点查询 DNS,若 DNS 返回 IPv4 地址发送 IPv4 分组；若返回 IPv6 地址发送 IPv6 分组。双栈节点同时拥有 IPv4 和 IPv6 地址。

IPv6 数据包如何穿越 IPv4 网络？【报头转换：双栈节点在将数据报传递给 IPv4 路由器之前将 IPv6 报头转换成 IPv4 报头；缺点：报头转换不完全，有信息丢失(数据字段可以直接复制，并作适当的地址映射)】【建立隧道：IPv6/IPv4 边界路由器将 IPv6 包**封装**到一个 IPv4 包中送入 IPv4 网络，目的边界路由器取出 IPv6 包继续传输；优点：保留原始数据报的全部信息。IPv6 包在 IPv4 网络中传输时，没有被 IPv4 路由器检查和处理，在隧道出口路由器看来，IPv6 包是从隧道入口路由器直接到来的，报头中的 Hop Limit 不会减小】

## 4.5 其他网络层协议

**[网络地址转换(NAT)] 引入动机：**使用一个公用 IP 地址支持许多用户同时上网，仅为公共可访问的节点分配公用 IP 地址(减少需要的公用 IP 地址数)，网络内部节点对外是不可见的(安全考虑)。例如，地址空间 10.0.0.0/8 用于家庭网络等专用网络或具有专用地址的地域(其地址仅对该网络中的设备有意义的网络)。

• **NAT 实现：**NAT 使能路由器两侧接口对应广域和局域 IP。外出的数据报：将数据报中的(源 IP 地址，源端口号)替换为(NAT IP 地址(不变)，NAT 端口号(不断分配))，同时增加表项。NAT 转换表：记录每个(源 IP 地址，源端口号)与(NAT IP 地址，NAT 端口号)的转换关系。进入的数据报：取出数据报中的(目的 IP 地址，目的端口号)查找 NAT 转换表，然后用转换表中对应的 (IP 地址，端口号)进行替换。16b 端口号：允许一个 NAT IP 地址同时支持65535个对外连接。

• **NAT 的使用有争议：**端口号是用于进程编址而非主机编址；路由器应当只处理三层以下的包头(端口号在传输层)；违反端到端原则(节点介入修改 IP 地址和端口号)；应使用 IPv6 解决 IP 地址短缺问题。**NAT 妨碍 P2P 应用：**NAT 只允许内部主动发起的通信，位于 NAT 后面的主机对外是不可见的；但 P2P 应用要求任何对等方可以向任何其它(参与的)对等方发起通信。**问题的本质：**如果 B 在 NAT 后面，他不能充当服务器并接收 TCP 连接：如果 A 不在 NAT 后，则可先与不位于 NAT 后且与 B 已经创建了一条进行中的 TCP 连接 C 联系，再与 B 建立 TCP，即连接反转，被许多 P2P 程序用于 NAT 穿越。

• **使用 UPnP 实现 NAT 穿越：**UPnP 是一种允许主机发现并配置临近 NAT 的协议，要求主机和临近 NAT 都是 UPnP 兼容的。假设主机在端口 3345 上运行一个 BT 程序：BT 程序请求 NAT 产生一个“洞”，将<10.0.0.1,3345>映射到<138.76.29.7,5001>上；BT 程序向追踪器通告它在<138.76.29.7,5001>上可用；其它主机通过追踪器可以看到该主机，并能向<138.76.29.7, 5001>发起 TCP 连接；NAT 将<138.76.29.7,5001>上收到的 SYN 包转发给主机。总之，UPnP 允许外部主机使用 TCP 或 UDP 向 NAT 化的主机发起通信会话。

• **使用中继服务器实现 NAT 穿越：**在 Skype 中使用：NAT 后面的服务器与中继器建立连接，外部客户与中继器建立连接，中继器在两个连接之间转发分组。

**[因特网报文控制协议 ICMP] 见第五章，[安全网络层协议 IPsec] 见第八章**

**第五章 网络层：控制面**

## 5.1 路由选择算法**[选路问题]** 给定一组路由器与连接路由器的链路，寻找一条由源路由器到目的路由器的最佳路径。路径是分组从源节点到目的节点所经过的节点序列。路径的评价指标包括：路径长度、数据速率、分组延迟、通信费用、安全性等；ISP 关心一些全局性指标：网络吞吐量最大、平均包延迟最小、平均通信费用最低、网络负载均衡、路由稳定、健壮等。路由评价指标通常是矛盾的，需要折中。选路算法——寻找从源节点到目的节点代价最小的路径。分类包括【全局算法：所有路由器具有关于拓扑和链路代价的全部信息，采用集中式计算(也可在多点重复进行)；分布式算法：路由器仅知道邻居节点以及到邻居节点的链路代价，通过与邻居交换信息进行迭代计算】【静态算法：路由随时间不变或缓慢变化(手工配置)；动态算法：路由器根据拓扑及链路代价的变化而自动更新路由】负载敏感or迟钝算法：链路费用是否反应底层链路的拥塞水平(动态)

## 5.2 LS(链路状态)与 DV(距离矢量)算法

**[LS 算法：全局算法]** 基本思想：每个节点利用可靠方法(每个结点向网络中所有其他结点广播链路状态分组，包括所连接的链路特征和代价(一般与流量相关))获得**全网拓扑信息**，抽象成一个带权拓扑图，计算到各个节点的最短路径。**五个步骤：**发现邻居(有链路直接相连的节点为邻居)、探测链路代价(探测到每个邻居的代价)、构造 LS 分组(利用邻居及链路代价信息)、扩散 LS 分组(向网络中所有节点发送 LS 分组)、计算路由(利用收到的 LS 分组构造网络拓扑，计算从本节点到其它各个节点的最短路径)示例：Dijkstra 算法。**可能出现的问题：**拥塞敏感的选路震荡(可能出现在任何使用拥塞或基于时延的链路测度算法中)，解决方案一个是强制规定开销不依赖于所承载的流量，一个是确保并非所有路由器都同时运行 LS 算法(避免路由器间的自同步：让每台路由器发送链路通告的时间随机化)

**[DV 算法：异步、迭代、分布式、自我终止]** Bellman-Ford 方程：分布式求解任意两个节点之间的最小代价路径 【d\_x(y)=min\_p[c(x,p)}d\_p(y)], p\in Neighbor(x)】基本思想：节点 x 测量其到各个邻居 v 的链路代价 c(x,v)、节点 x 估计其到达各个节点 y 的最小代价 Dx(y)，这些代价构成了自己的距离矢量 Dx=[Dx(y):y∈N]、每个节点周期性地将它的距离矢量 Dx 发送给邻居、节点 x 拥有每个邻居 v 的距离矢量 Dv=[Dv(y):y∈N]、当节点 x 从各个邻居收到它们的距离矢量后，利用 B-F 方程更新自己的距离矢量(最小的 v 被更新到 y 的转发表项)。**节点的本地计算由以下两种事件引起：本地某条链路的代价 c(x,p)发生了变化、收到了某个邻居节点的距离矢量 dp(y)。**节点仅在发现距离矢量 dx(y)有变化时通知其每个邻居。**链路代价变化：**好消息(费用降低)传播快，坏消息传播慢(路由选择环路和无穷计数问题。解决：毒性逆转：如果 z 通过 y 路由到 x，则 z 通告 y：Dz(x)=无穷；涉及3个及以上结点的环路无法使用该方法)。

**[LS 算法和 DV 算法的比较]** 通信复杂度方面：LS 链路状态信息在全网传播；DV 距离矢量仅发送给邻居。收敛速度方面：LS 发送 O(|N||E|)个报文，O(N^2)次计算；DV 的速度差异大，且可能出现环路、无穷计数的问题。健壮性方面：LS 节点仅传播可靠的信息(亲自测量的本地链路代价；一个节点也可损坏或丢弃收到的任何 LS 广播分组，每个节点仅计算自己的转发表节点计算的路由不传播，错误不扩散)；DV 邻居的距离矢量是“道听途说”的；节点计算的路由要传播，会造成错误扩散。

其他路由选择算法：将分组流量看作网络中源和目的之间的流；电路交换路由选择算法

## 5.3 内部网关协议 Intra-AS：使用路由测度(代价)选择最佳路径

**[因特网中的路由选择]** 先前的假设：所有路由器运行相同的选路算法、网络是平面结构的(选路信息在全网扩散)。然而，实际的网络环境中有数十亿的目的节点，无法放入同一个选路表中，也无法在全网中交换所有的路由信息(平面结构不具有扩放性)。同时，网络管理员希望对于网络有更多的控制权：选路算法的选择，隐藏网络内部组织(管理的自治性)

**[自治系统 AS]** AS 是由处于同一个管理域下的网络和路由器组成的集合；ICANN 为每个 AS 分配一个 AS 编号。同一个 AS 中的路由器运行相同的选路协议(Intra-AS 选路协议)；不同 AS 中的路由器可以运行不同的 Intra-AS 选路协议。**网关路由器：在一个 AS 内直接连接到其它 AS的路由器**，负责向本AS之外的目的地转发分组，可能有多个；网关路由器之间运行 Inter-AS 选路协议，所有AS必须运行相同的 Inter-AS 选路协议(因特网中为 BGP4)。

**[选路算法和选路协议]**选路算法是选路协议的一部分；选路协议还包括路径代价的定义、报文格式、报文传输、报文处理、异常事件处理等具体实施方面的问题。**因特网中的选路协议：**Intra-AS 选路协议，也称内部网关协议 IGP。常见的有：路由选择信息协议 RIP(较低层 ISP 和企业网使用，采用DV)、开放最短路优先 OSPF (较顶层 ISP 使用，采用LS)、IGRP、EIGRP(DV)。Inter-AS 选路协议，也称外部网关协议 EGP。目前只有：BGP。

**[因特网中自治系统内部的路由选择：OSPF]** OSPF 采用 LS 算法，没有规定代价怎么取，而是由管理员配置。为使路由相对稳定，通常选取不随负载而改变的链路特性作为代价。链路代价通常反映了管理员的选路策略，例如对于偏好的链路可将代价设置较小，实际应用中通常还会选若干条路径，每条路径设置一定的使用概率。**OSPF 分组：**OSPF 协议定义了 5 种分组类型，分别用于探测邻居、通告链路状态等。OSPF 报文被封装在 IP 包中传输，协议号为 89。路由器周期性地(一般至少 30min 一次，这样增加了算法健壮性)、或在链路状态改变时发送 OSPF 链路通告。OSPF 协议负责链路通告分组在网络中的广播及可靠传输。路由器根据收到的链路通告分组构造链路状态数据库，利用链路状态数据库及 Dijkstra 算法，计算以本路由器为根所有子网的的最短路径树。

**<AS 内部的分层选路：OSPF 最大优点>** 一个 OSPF 自治系统可以配置成多个区域：每个区域运行自己的 OSPF 协议，区域内部的链路状态仅在本区域内广播，区域边界路由器负责区域间的选路。一个特殊区域称为主干，所有区域必须连接到主干上。主干为 AS 内其他区域之间的流量提供路由选择，包含AS内的所有区域边界路由器，可能还包含一些非边界路由器。每一个区域都有区域标识，主干为 0。**路由器包括：**区域边界路由器(连接本地区域和主干的路由器)、主干路由器(主干上的路由器，可以同时是区域边界路由器，可以有多个)、内部路由器(AS 内部的非区域边界路由器)。**分层的 OSPF 实现：**每个区域(包括主干)运行自己的 OSPF 协议；每个区域边界路由器将本区域的选路信息汇总(子网及路径代价)并通告给其它区域，将收到的其它区域选路信息通告给本区域的内部路由器；去往其它区域的分组首先转发到本地区域边界路由器，在主干上转发到目的区域边界路由器，然后再转发到目的子网。**OSPF 的优点：**AS 内部分层选路、安全(能鉴别 OSPF 路由器间的交换，简单(相同口令)和MD5(共享密钥)，防止恶意入侵者将不正确的信息注入路由器表内)、可以同时使用多条费用相同的路径(RIP只允许一条，可以按一定比例将流量分配到多条路由上)、对单播和多播路由选择的综合支持(MOSPF 和 OSPF 使用相同的拓扑数据库)。

## 5.4 外部网关协议 Inter-AS：依据策略寻找可达路径

**[AS 与分层路由]** 当 AS 内部路由器想要发送数据报到 AS 外，要求 AS 的网关路由器必须了解**哪些目的网络由自己可达**，并**将可达信息传播到 AS 内部所有路由器。**如果存在多条外部 AS 可达路径，则优先选择经过 AS 数较少的、距离源路由器最近的网关路由器所对应道路的路径。

**[边界网关协议 BGP：应用层协议、分布式、异步] AS 间选路的困难与目标：**因特网规模极其庞大且结构非常复杂；每个 AS可运行自己的内部路由协议，使用自己的路由测度确定到目的网络的最佳路由，不同网络判断最佳路由的标准不同；一个 AS可能不信任来自某个 AS的选路信息；一个AS可能不愿意为其它 AS转发数据包；AS间选路试图找到能够到达目的网络的路由，**但不试图(也不可能)找到最佳路由**

**BGP 为 AS 提供：**每个 AS 可以从邻居 AS 那儿获得到其它子网的可达性信息、可以将其它网络的可达性信息传播到 AS 所有的内部路由器、可以基于可达性信息和**策略**确定到其它网络的“好”路由。*重要：使得每个子网向因特网的其余部分通告它的存在。* **BGP 对等方：**当一对 AS 同意交换选路信息时，每个 AS 指定一个接近 AS 边缘的路由器(或主机)，使用 BGP 协议交换选路信息。运行 BGP 协议的边界路由器(或主机)称为 BGP speaker。一对 BGP speaker 通过一条半永久的 TCP 连接(端口179)建立 BGP 会话，交换 BGP 报文(BGP 是应用层协议)，BGP 会话的两个端点互为 BGP 对等方

**[边界网关协议 BGP](续) BGP 会话：**沿着 BGP 连接发送所有 BGP 报文的 TCP 连接称为 BGP 会话；跨越 2 个 AS 建立的 BGP 会话称为外部 BGP(eBGP)会话；相同 AS 内的两台路由器之间的 BGP 会话称为内部 BGP(iBGP)会话。一个 AS 可能有多个网关路由器，这些网关路由器必须通过半永久 TCP 连接构成全连通。BGP 会话交换的最重要内容是可达性信息：路由器收到相邻 AS 的路由通告，在向下一个 AS 发送该路由之前修改报文，将自己的标识及 AS 号加入到完全路径中。边界路由器的选路表中，每个表项包含目的前缀(允许聚合)、下一跳路由器以及到达目的前缀的 AS 序列。以路径广告通报完全路径。**路径广告：**以 AS 枚举形式通告的、到达目的网络前缀的完全路径(为便于检测出路径环路)。例如 AS2,AS3,X(2 为 1 转发)。BGP 会话线并不总是与物理链路对应。BGP 中目的地为子网或子网集合，BGP 会聚合前缀。在任何 AS 中的网关路由器接收到 eBGP 学习到的前缀后该网关路由器使用其 iBGP 向该AS中的其他路由器发布这些前缀；当一个路由器得到一个新前缀时便创建表项。**基于策略的路由：**入境过滤(根据输入策略，对每条新的路由进行入境过滤。结果可能是丢弃，按原样接受，接受但修改某些属性(如偏好度))、路径选择(对于每一个目的前缀，从所有可达的路径中按照 BGP 指定的决策顺序确定一条最佳路由[前缀及其属性])、出境过滤(根据输出策略，决定是否向其邻居AS发布某条路径)。缺点：贵、无法感知实时丢包与延迟

## 5.5 因特网报文控制协议 ICMP

**[ICMP 概述]**主机或路由器使用 ICMP 协议传递网络层上的一些信息。ICMP报文有询问和差错报告两类：**询问**(用来请求信息，通常采用请求-响应模式交互。例如：回声请求/响应用于发现对方是否在线，地址掩码请求/响应用于获得本网的地址掩码，路由器请求用于请求本网路由器的信息，路由器发现用于定期发送消息通告自己的信息)、**错误报告**(发现错误的节点向源节点报告错误信息，不需响应。例如：源抑制4-路由器缓存满，无法容纳新到来的数据报、超时11-数据报的TTL减为0，或主机的重组定时器超时、目的不可达3-路由器判断一个数据报不可能到达它的最终目的地、重定向-路由器发现主机使用的路由有错或不是最佳路由，报文不丢弃)。**与IP的关系：**ICMP报文是封装在IP分组中传输的，这是因为ICMP报文可能需要经过几个网络才能到达源节点，而跨网络传输必须借助IP。ICMP通常被认为是IP协议的一部分，因为IP协议使用ICMP向源节点发送错误报告，但体系结构上它位于IP之上

**[ICMP 应用] Ping 与 ICMP：**Ping 利用 ICMP 报文测试目的主机是否活跃，以及去往目的主机的路径是否正常：源主机发送 Type=8，Code=0 的 Echo Request 报文；若目的主机收到，发送 Type=0，Code=0 的 Echo Response 报文；源主机计算 RTT，并报告；若源主机连续几次超时(收不到 Echo Response)，向调用者报告目的不可达。**源抑制报文：**type=4，code=0 很少使用，最初目的是执行拥塞控制，使得拥塞的路由器向一台主机发送一个 ICMP 源抑制报文，强制该主机减小发送速率。TCP 需要网络层反馈信息。**Traceroute 与 ICMP：**Traceroute 测试到达目的主机的路由(经过的路由器)：源主机的 Traceroute 程序向目的主机发送一个 Echo Request 报文(书上：普通的 IP 数据报，每个包含了一个具有不可达端口号的 UDP 报文段)，IP 报头的 TTL 设为 1。源主机也为每个数据报启动定时器。第一跳路由器对 TTL 减 1，发现 TTL 变为 0(路由器丢弃该数据报)，向源主机发送一个 TTL expired(ICMP 告警)报文(type=11，code=0，IP 报头中有路由器的 IP 地址，还包含名字)。Traceroute 从定时器获得往返时延，从 ICMP 报文中得到第一跳路由器的名字和 IP 地址，接着发送一个 TTL 为 2 的 Echo Request 报文…不断重复，直至收到目的主机的 Echo Response 报文(该目的主机向源发送一个端口不可达(type=3,code=3)的 ICMP 报文)。实际情况中相同的 TTL 会发送 3 个一组的分组。**ICMPv6：**合并了 IPV4 中的 ARP 与多播组管理，并取消了RARP(该协议的功能已被其它协议取代)。ICMPv6 仍然使用差错报告和查询两类报文。差错报告：去掉了源抑制报文，优先级和流标签允许路由器控制拥塞，丢弃不太重要的数据包；增加了 Packet Too Big 报文：路由器丢弃长度超过MTU的包，并报告错误。信息查询：去掉了一些不必要的查询报文；增加了一些查询报文，用于实现 ARP(地址解析)和IGMP(多播组管理)的功能。

**第六章 链路层**

## 6.1 概述

**[网络层、链路层与物理层的功能]**网络层：选路(路由器确定去往目的节点的下一跳)、转发(在路由器内部将数据报从输入端口转移到输出端口)；链路层：将数据报从一个节点传输到相邻的下一个节点；物理层：多种类型的传输媒体传输原始比特流(无结构)

• 节点：执行链路层协议的任何设备(主机、路由器、交换机、WiFi 接入点)，在链路层上设备之间已经没有区别。• 链路：连接相邻节点的通信信道。有线链路、无线链路、局域网

**[链路层服务]【**组帧(基本服务)：发送(将数据报加上帧头帧尾封装到帧中)、接收(从原始比特流中提取出完整的帧，要求链路层正确识别帧的边界)】【链路接入(广播链路需要)：媒体访问控制协议 MAC 规定了帧在链路上传输规则，在点对点链路上 MAC 简单或不存在；在多个结点共享单个广播链路时(多路访问)MAC 协调多个结点的帧传输】【差错检测(基本服务)：检测传输错误，通过让发送结点在帧中包括差错检测比特，让接收结点进行差错检查，比运输层网络层的检测更复杂，采用硬件实现】【差错纠正(有些提供)：检测(差错出现的位置)并纠正传输错误(不使用重传)】【可靠交付(部分协议提供)：通过确认、重传等机制确保接收节点正确收到每个帧(停-等、GBN、SR)，在低误码率链路(如光纤、同轴电缆、许多双绞铜线链路)上很少使用，高误码率链路(如无线链路)应当使用，目的是在差错发生的链路上纠正一个差错，而不是在更高层进行重传】【流量控制：调节发送速度，避免接收节点缓存溢出(提供可靠交付的链路层协议不需要专门的流量控制，不提供可靠交付的链路层协议需要流量控制机制)】【半双工和全双工：半双工通信时，提供收/发转换】

**[链路层的实现]** 路由器中链路层在线卡中实现，主机链路层主体部分在网络适配器(网络接口卡)中实现，与PCI相连，其核心是链路层控制器(实现了许多链路层服务的专用芯片)。线卡/网络适配器连接物理媒体，并实现物理层的功能。链路层由硬件和软件实现(协议栈中软硬件交接的地方)：网卡中控制器芯片：组帧、链路接入、检错、可靠交付、流量控制等；主机上链路层软件(高层链路层功能)：与网络层接口，组装链路层寻址信息和激活控制器硬件，在接收端响应控制器中断，处理差错条件和将数据报向上传给网络层。**节点间的通信：**发侧-将数据报封装到帧中，生成校验比特，执行可靠传输和流量控制；接侧-提取帧，检测传输错误；(可选)执行可靠传输和流量控制；解封装数据报，交给上层协议。

**[随机接入的 MAC 协议](续)**

• 效率计算：当网络中存在大量活跃节点总有帧发送时时，长期运行过程中成功时隙(刚好有一个结点传输的时隙)所占的比例。**时隙 Aloha 的效率：**假设: 有 N 个活跃节点，每个节点在每个时隙开始时以概率 p 发送(重传和新发送)；给定节点在一个时隙中发送成功的概率 = p(1-p)^(N-1)；给定时隙中有节点发送成功的概率 = Np(1-p)^(N-1)(也即效率)；最大效率: 找到令 Np(1-p)^(N-1)最大的概率 p\*；代入 Np\*(1-p\*)^(N-1)，并令 N 趋向于无穷，得到：最大效率 = 1/e = 0.37(另外 0.37 空闲，0.26 碰撞)。

• **纯 ALOHA：**基本思想：取消同步时钟，任何节点有数据发送就可以立即发送，节点通过监听信道判断本次传输是否成功，若不成功(过程中发生碰撞)，以概率 P 立即重传，以概率(1-P)等待一个帧时后再决定，等待后重复此过程。(帧时：发送一帧的时间，假设帧长度相同)。发生冲突的情形：在时刻 t0 发送的帧与在 [t0 - 1,t0 + 1] 时段内发送的其它帧冲突。**纯 Aloha 的效率：**P(给定节点发送成功)=P(节点发送)\*P(无其它节点在[t0-1,t0]内发送)\*P(无其它节点在[t0, t0+1]内发送)=p\*(1-p)^(N-1)\*(1-p)^(N-1)=p(1-p)^2(N-1)；求出令节点发送成功概率 Np(1-p)^2(N-1)最大的 p\*，并令 N→infty：最大效率 = 1/(2e) = 0.18。效率更低的原因：冲突的区间更大了(两倍帧时大小的冲突区间)

ALOHA 家族效率较低的原因：没有提前监听信道是否被占用。

• **载波侦听多址接入(CSMA)**：两个重要的规则：发送前监听信道(载波侦听)，信道空闲发送整个帧，信道忙则等待至少一段时间没有传输再发送(推迟发送)；碰撞检测，即当一个传输结点在传输时一直在侦听此信道，如果它检测到另一个结点正在传输干扰帧，就立即停止传输，在重复＂侦听－当空闲时传输＂循环之前等待一段随机时间。这两个规则包含在载波侦听多路访问(CSMA)和具有碰撞检测的 CSMA(CSMA/CD)协议族中。冲突仍可能发生：**由于存在端到端信道传播时延，节点可能没有监听到其它节点正在发送**；即使忽略传输延迟，当两个(或多个)节点同时发现信道由忙变为空闲、并都决定立即发送时，仍会发生冲突。信道传播时延在决定其性能方面起关键作用。

• **CSMA/CD (Collision Detection)：**基本思想：在发送的过程中检测冲突(发生冲突时信号较强)；检测到冲突后，立即停止发送剩余的部分；并立即启动冲突解决的过程。

以太网采用 CSMA/CD 协议：1.NIC(网卡)从网络层接收数据报，构造以太帧(链路层帧)，并将其放入帧适配器缓存中；2.若 NIC 监听到信道空闲，立即发送帧；若信道忙，坚持监听直至发现信道空闲，然后发送帧；3.若 NIC 发送完整个帧而没有检测到冲突，认为发送成功；4.若 NIC 在传输过程中检测到冲突，立即停止发送帧，并发送一个阻塞信号(加强冲突)；5.中止传输后等待一个随机时间量(指数回退)然后返回步骤 2。我们希望时间间隔应该这样：当碰撞结点数最较少时，时间间隔较短；当碰撞结点数量较大时，时间间隔较长。二进制指数后退算法：当传输一个给定帧时，在该帧经历了一连串的 n 次碰撞后，结点随机地从{ 0, 1, 2, 3,…, 2^n-1 }中选择一个 K 值，(一个帧经历的碰撞越多，K 选择的间隔越大)对于以太网，一个结点等待的实际时间是 K\*512 比特时间(512 比特为以太帧最小长度)，n 能够取的最大值在 10 以内。指数回退的目的是根据网络负载调整重传时间：负载越重(冲突次数越多)，重传时间的选择范围越大，再次发生冲突的可能性越小。每次适配器准备传输一个新的帧时，它要运行 CSMA/CD 算法，不考虑近期过去的时间内可能已经发生的任何碰撞，因此，当几个其他适配器处于指数后退状态时，有可能一个具有新帧的结点能立刻插入一次成功的传输。

• CSMA/CD 效率：t\_prop = 以太网中任意两个节点间传播延迟的最大值；t\_trans = 最长帧的传输时间；efficiency = 1 / (1 + 5\*t\_prop/t\_trans)。以下情况以太网的效率趋近于 1：t\_prop 趋近于 0，或 t\_trans 趋向于无穷。结论：应控制以太网的规模。

**[轮流 MAC 协议]** 多路访问协议的两个理想特性是：1. 当只有一个结点活跃时，该活跃结点具有 R bps 的吞吐量；2. 当有 M 个结点活跃时，每个活跃结点的吞吐撞接近 R/M bps。ALOHA 和 CSMA 协议具备第一个特性，但不具备第二个特性。

• **轮询协议：**结点之一被指定为主结点，主结点以循环的方式轮询每个结点，主结点首先向结点 1 发送一个报文，告诉它能够传输的帧的最多数量，在结点 1 传输了某些帧后，主结点告诉结点 2 它能够传输的帧的最多数量(主结点能够通过观察在信道上是否缺乏信号，来决定一个结点何时完成了帧的发送)。优劣：消除了碰撞和空时隙，使得轮询取得高得多的效率；第一个缺点是引入了轮询时延，即通知一个结点“它可以传输”所需的时间，例如，如果只有一个结点是活跃的，那么这个结点将以小于 R bps 的速率传输，因为每次活跃结点发送了它最多数址的帧时，主结点必须依次轮询每一个非活跃的结点；第二个缺点是单点失效，主结点有故障，整个信道都变得不可操作。

**• 令牌传递协议：**没有主结点，一个称为令牌(token)的小的特殊帧在结点之间以某种固定的次序进行交换；当一个结点收到令牌时，仅当有一些帧要发送时，它才持有这个令牌；否则，它立即向下个结点转发令牌；一个结点收到令牌时，若确实有帧要传输，则它发送最大数目的帧数，然后把令牌转发给下个结点。优劣：令牌传递是分散的，并有很高的效率；一个结点的故障可能会使整个信道崩溃；如果一个结点偶然忘记了释放令牌，则必须调用某些恢复步骤使令牌返回到循环中；令牌传递延迟；令牌单点失效。

**[MAC 协议的比较] 信道划分 MAC 协议：**重负载下高效(没有冲突，节点公平使用信道)；轻负载下低效(即使只有一个活跃节点也只能使用 1/N 的带宽)。**随机接入 MAC 协议：**轻负载时高效(单个活跃节点可以使用整个信道)；重负载时低效(频繁发生冲突，信道使用效率低)。**轮流协议(试图权衡以上两者)：**按需使用信道(避免轻负载下固定分配信道的低效)；消除竞争(避免重负载下的发送冲突)。

• DOCSIS(综合例子)：使用 FDM 将下行(CMTS 到调至解调器)和上行网络段划分为多个频率信道；上下行均为广播信道；下行方向不存在多路访问问题；上行信道被划分为时间间隔(TDM)，通过下行信道上的 MAP 控制报文使得无碰撞；电缆调制解调器向 CMTS 发送请求帧，采用随机接入，根据 MAP 响应检测碰撞，如果有碰撞，采用二进制指数回退；当上行信道流量少时可能微时隙请求帧可能被用来实际传输数据

**6.3 多址接入(多路访问)协议/媒体接入控制协议 Medium Access Control，MAC**

**[概述]** 链路的两种类型：点到点链路(仅连接了一个发送方和一个接收方的链路，一条全双工链路可以看成是由两条单工链路组成)、广播链路(连接了许多节点的单一共享链路，任何一个节点发送的数据可被链路上的其它节点接收到)。

• 冲突(collision)：在**广播链路**上，若两个或多个节点同时发送，发送的信号会发生干扰，导致接收失败。多址接入协议(多路访问协议)：规定节点共享信道(谁可以发送)的方法，多址接入协议也称媒体接入控制(MAC)协议。理想的多址接入协议：在速率为 R bps 的广播信道上，1. 当只有一个节点发送时，它应能以速率 R 发送(信道利用率高)2. 当有 M 个节点发送时，每个节点应能以 R/M 的平均速率发送(公平性好、信道利用率高)3. 协议是完全分布式的:不需要一个特殊的节点来协调发送(健壮性好)，不需要时钟或时隙同步(不需要额外的机制)4. 简单(实现和运行开销小)。

**[MAC 协议的分类]**

**信道划分**：将信道划分为若干子信道，每个节点固定分配一个子信道，不会发生冲突，关注公平性，轻负载时信道利用率不高；**随机接入(竞争)：**不划分信道，每个节点自行决定何时发送，出现冲突后设法解决，轻负载时信道利用率高，重负载时冲突严重；**轮流使用信道：**不划分信道，有数据的节点轮流发送，不会出现冲突，信道利用率是以上两种方法的折衷，引入额外机制。

**[信道划分的 MAC 协议]**

• TDMA(时分多址)：将信道的使用时间划分成帧，每个节点在帧中被分配一个固定长度的时间片(一个时间帧 N 个时隙)，每个时间片可以发送一个分组，节点只能在分配给自己的时间片内发送，若节点不发送，其时间片轮空。

• FDMA(频分多址)：将信道频谱划分为若干子频带，每个节点被分配一个固定的子频带(具有 R/N 带宽)，若节点不发送，其子频带空闲。TDM 和 FDM 都有的优劣：消除碰撞而且非常公平；结点被限制于 R/N bps 的平均速率，必须总是等待它在传输序列中的轮次。

• CDMA(码分多址)：将一个 bit 时间划分为 m 个较短的间隔，每个节点被分配一个不同的 m-bit 编码，编码之间彼此正交；每个结点用其唯一的编码来对它发送的数据进行编码。当原始数据为 0 时发送编码，为 1 时发送编码的反码。CDMA 允许所有节点同时使用整个信道，它们各自相应的接收方仍能正确接收发送方编码的数据比特，不在乎其他结点的干扰传输。接收方解码：用各发送方的编码与信道中收到的混合信号计算内积，根据结果恢复出原数据。

**[随机接入的 MAC 协议]**

随机接入的基本思想：当节点有数据要发送时，以信道速率 R 发送，发送前不需要协调；当有碰撞时，涉及碰撞的每个结点反复地重发它的分组，到该帧无碰撞地通过为止；经历一次碰撞时，在重发之前等待一个随机时延，涉及碰撞的每个结点独立地选择随机时延。随机接入 MAC 协议规定如何检测冲突，以及如何从冲突中恢复。

发送前不监听信道：ALOHA 家族；发送前监听信道：CSMA 家族

• **时隙ALOHA：**假设：所有帧长度相同、时间被划分为等长的时隙、每个时隙传一帧、节点只能在时隙开始时发送、节点是时钟同步的(知道时隙何时开始)、所有节点可在时隙结束前检测到是否有冲突发生。操作：节点从上层收到数据后，在下一个时隙发送；若时隙结束前未检测到冲突，节点可在下一个时隙发送新的帧；若检测到冲突，节点在随后的每一个时隙中以概率 P 重传，直至发送成功。**优点：**单个活跃节点可以完整信道速率连续发送；分布式：每个结点检测碰撞并独立地决定什么时候重传；简单。**缺点：**发生冲突的时隙被浪费；由于概率重传，有些时隙被闲置；需要时钟同步(时隙 ALOHA 要求在结点中对时隙进行同步，而 ALOHA 和 CSMA 不需要)。

## 6.2 差错检测与纠正技术

**[检错与纠错]** 传输出错的类型：单个错(由随机的信道热噪声引起，一次只影响 1 位)、突发错(由瞬间的脉冲噪声引起，一次影响许多位，使用突发长度表示突发错影响的最大数据位数)。差错控制编码的类型：检错码(只能检测出传输错误的编码，不能确定出错位置，通常与反馈重传机制结合进行差错恢复)、纠错码(能够确定错误位置并自行纠正的编码)

**<编码纠错理论>** 码字：由 m 比特的数据加上 r 比特的冗余位(校验位)构成；有效编码集：由 2^m 个符合编码规则的码字组成；检错：若收到的码字为无效码字，判定出现传输错误；海明距离：两个码字的对应位取值不同的位数；纠错：将收到的无效码字纠正到距其最近的有效码字。检错码与纠错码的能力都是有限的；编码集的海明距离：编码集中任意两个有效码字的海明距离的最小值；检错能力：为检测出所有d比特错误，编码集的海明距离至少应为d+1；纠错能力：为纠正所有d比特错误，编码集的海明距离至少应为 2d+1；差错检测的实施：发送端对要保护的数据 D(包括帧头字段)生成校验位EDC(差错检测和纠正比特)添加在帧头中；接收端对收到的数据 D\_计算校验位 EDC\_，根据 EDC\_判定是否有错。

**[奇偶校验]** 单比特奇偶校验：(1 的总数是奇数还是偶数，校验比特本身也包含在内) 可检测单比特错误，检错率为 50%(在突发差错的情况下)，编码集海明距离为 2。(在偶校验方案中发现奇数个 1，则出现了奇数个比特差错)；二维奇偶校验：可检测 2 比特错和纠正单比特错，编码集海明距离为 3，有利于检测突发错误；前向纠错(FEC)：接收方检测和纠正差错的能力；优点：减少所需的发送方重发次数，允许在接收方立即纠正差错，避免了不得不等待的往返时延，而这些时延是发送方收到 NAK 分组并向接收方重传分组所需要的。

**[CRC 与检验和]**

CRC 是一种多项式编码，它将一个位串看成是某个一元多项式的系数，如 1011 看成是一元多项式 X^3 + X + 1 的系数。信息多项式 M(x)：由 m 个信息比特为系数构成的多项式。冗余多项式 R(x)：由 r 个冗余比特为系数构成的多项式。码多项式 T(x)：在 m 个信息比特后加上 r 个冗余比特构成的码字所对应的多项式，表达式为 T(x) = x^r \* M(x) + R(x)。生成多项式 G(x)：双方确定用来计算 R(x)的一个 r+1 比特多项式，最高位为 1。编码方法：R(x) = x^r \* M(x) ÷ G(x) 的余式(减法运算定义为异或操作)。检验方法：若 T(x)÷G(x)的余式为 0，判定传输正确。CRC 码检错能力极强，可用硬件实现，是应用最广泛的检错码。**CRC 举例：**取 G(X) = X^3 + 1，对信息比特 101110 计算 CRC 码。101110000 ÷ 1001 的余式为 R=011 (CRC code)，码字：101110011；取 G(X) = X^3 + 1，接收端收到比特串1001001，是否有错？解答：1001001÷1001 的余式为 001(不为 0)，有传输错误。CRC 能检测小于 r+1 比特的突发错误，能检测任何奇数个比特差错。**因特网检验和：**数据的字节作为 16 比特的整数对待并求和(溢出回卷)，这个和的反码形成了携带在报文段首部的因特网检验和；TCP 和 UDP 对所有字段(包括首部和数据字段)计算因特网检验和。优劣：检验和方法需要相对小的分组开销，与 CRC 相比提供相对弱的差错保护。

**为什么传输层使用检验和而链路层使用 CRC：**传输层通常是在主机中作为用户操作系统的一部分用软件实现的。因为运输层差错检测用软件实现，采用简单而快速如检验和这样的差错检测方案是重要的。另一方面，链路层的差错检测在适配器中用专用的硬件实现，它能够快速执行更复杂的 CRC 操作。

## 6.4 交换局域网 LANs

**[局域网、城域网和广域网]** 局域网 LAN：将小范围内的计算机及外设连接起来的网络，范围在几公里以内，通常为个人或机构所有；城域网 MAN：通常覆盖一个城市的范围(几十公里)，要能支持数据、音频和视频在内的综合业务，服务质量好，支持用户数量多；广域网 WAN：通常覆盖一个国家或一个洲(百公里以上)，规模和容量可任意扩大

**[链路层编址]** 早期的局域网多采用广播信道，节点如何判断收到的帧是给自己的？为每一块网络适配器(网卡)固定分配一个地址，称为物理地址(硬件地址、链路层地址、MAC 地址、LAN 地址)。MAC 地址长 6 个字节(1A-2B-3C-CD-06-9B)；由 IEEE 负责分配，每块适配器的地址是全球唯一的：网卡生产商向 IEEE 购买一块 MAC 地址空间(前 3 字节)，生产商确保生产的每一块网卡有不同的 MAC 地址；MAC 地址固化在网卡的 ROM 中；现在用软件改变网卡的 MAC 地址也是可能的。**主机或路由器的适配器(即网络接口)具有链路层地址**，因此，具有多个网络接口的主机或路由器将具有与之相关联的多个链路层地址；链路层交换机并不具有与它们的接口(这些接口是与主机和路由器相连的)相关联的链路层地址，这是因为链路层交换机的任务是在主机与路由器之间承载数据报；**交换机透明地执行该项任务**，这就是说，主机或路由器不必明确地将帧寻址到其间的交换机。适配器的 MAC 地址具有扁平结构(这与层次结构相反)，而且不论适配器到哪里用都不会变化。**MAC 地址有三种类型：**单播地址(配器的 MAC 地址，地址最高比特为 0)；多播地址(标识一个多播组的逻辑地址，地址最高比特为 1)；广播地址(全 1)。网络适配器仅将发送给本节点的帧交给主机：(一台交换机偶尔将一个入帧广播到它的所有接口，因此一块适配器可以接受到一个并非向它寻址的帧)目的地址为适配器 MAC 地址的单播帧、所有广播帧、指定接收的多播帧(若将适配器设置成混收模式，适配器将收到的所有帧交给主机)。**世界上先有MAC地址，后有IP地址。**在TCP/IP(互联网)出现之前，只使用 MAC 地址在单个的物理网络中寻址。为什么有了MAC地址，还要IP地址？因为MAC地址是扁平结构的，无法在因特网范围内快速确定接口的位置；IP地址是有结构的，可以在因特网范围内快速确定网络接口的位置。此外，局域网是为任意网络层协议而设计的，而不只是用于IP和因特网；如果适配器使用网络层地址而不是 MAC 地址的话，网络层地址必须存储在适配器的 RAM 中，并且在每次适配器移动(或加电)时要重新配置。另一种选择是在适配器中不使用任何地址，让每个适配器将它收到的每帧数据(通常是 IP 数据报)沿协议栈向上传递，然后网络层则能够核对网络地址层是否匹配，这种选择带来的一个问题是，主机将被局域网上发送的每个帧中断，包括被目的地是在相同广播局域网上的其他结点的帧中断。为了使网络体系结构中各层次成为独立的构建模块，不同的层次需要有它们自己的寻址方案。**IP 地址与 MAC 地址没有固定的关联关系**：MAC 地址与网卡绑定，与节点在哪个子网无关；IP 地址与所在子网有关，与网卡没有关系。

**[地址解析] MAC 地址：Media Access Control Address**

**如何将数据报发送到下一跳？**当发送节点 A、接收节点 B 位于同一个物理网络上时，数据报可从 A 直接交付给 B：A 的网络层将数据报、以及 B 的物理地址交给链路层；链路层将数据报封装在一个链路层帧中，帧的目的地址=B 的物理地址，B 的适配器收到帧，根据目的地址判断是发给本机的，取出数据报交给网络层。**已知 IP 地址，如何得到对应的 MAC 地址？**静态映射 IP 地址-MAC 地址的缺点：主机每次使用的 IP 地址可能不同(DHCP)、主机可能更换网卡。地址解析协议(ARP)用于动态获得 IP 地址-MAC 地址映射：若节点 A 希望获得节点 B 的 MAC 地址，节点 A 广播 B 的 IP 地址(地址解析请求)，节点 B 用自己的 MAC 地址进行响应。在以太网上，ARP 报文被封装在以太帧中传输。每台主机或路由器在其内存中具有一个 ARP 表：IP 到 MAC 的映射以及 TTL(从表中删除的时间)，(表中不包含所有的主机和路由器，它们可能还没进入或已过期，过期时间为 20min)。**地址解析过程**：A 想知道 B 的 MAC 地址(且 A 的 APR 表中无对应表项)：1.A 构造一个 ARP 查询分组，在发送方字段填入自己的 MAC 地址和 IP 地址，在目标字段填入 B 的 IP 地址；2.A 将 ARP 请求封装在广播帧中发送；3.每个收到 ARP 请求的节点用目标 IP 地址与自己的 IP 地址比较，地址相符的节点进行响应(B)。4.B 构造一个 ARP 响应分组(格式和查询分组相同)，交换发送方与目标字段内容，在发送方硬件地址字段填入自己的 MAC 地址，修改操作字段为 2；5.B 将 ARP 响应封装在单播帧(目的地址为 A 的 MAC 地址)中发送 6.发送方更新 ARP 表并发送其 IP 数据报。

**改进 ARP 的措施：**ARP 缓存：每个节点在内存中维护一个地址映射(绑定)表，称 ARP 缓存；每次发送数据报前先查询 ARP 缓存，若找不到则发送 ARP 请求，并在收到 ARP 响应后将地址映射缓存起来；ARP 缓存中的信息，在超时(20 分钟)后删除。主动学习：从 ARP 请求中获取地址绑定信息：每个节点可以收到全部的ARP请求报文，可将发送节点的地址映射缓存到自己的ARP表中；节点在启动时自动广播自己的地址映射：节点A在启动时主动广播一个ARP请求，在目标字段内填入自己的IP地址，收到ARP请求的节点将 A的地址映射缓存起来。若 A 收到ARP响应，报告IP地址重复错误(有其他节点与 A IP 相同)。

**ARP 是即插即用的，即一个 ARP 表是自动建立的。**一个 ARP 分组封装在链路层帧中，因而在体系结构上位于链路层之上。然而，一个 ARP 分组具有包含链路层地址的字段，因而可认为是链路层协议，但它也包含网络层地址，因而也可认为是为网络层协议。所以，可能最好把 ARP 看成是跨越链路层和网络层边界两边的协议。

• 数据报到达子网之外：路由器的每个接口都有一个适配器和 ARP 模块。数据报从 A(子网 1)经过 R(路由器)到达 B(子网 2)，A 知道下一跳 IP 地址为 R，R 知道 B 从其端口 R-2 直接可达。首先，A 创建 IP 数据报，src IP＝A, dest IP＝B；A 利用 ARP 获得下一跳 R 对应的 MAC 地址(R-1)(而非获得 B 的 MAC，否则 IP 数据报无法到达网络层)；A 创建链路层帧，封装 IP 数据包，s MAC =A, d MAC = R-1。发送；R 接收帧，取出 IP 数据报，发现目的地址为 B；R 利用 ARP 获得 B 的 MAC 地址；R 创建链路层帧，封装 IP 数据报，s MAC=R-2, d MAC = B，发送；B 的网卡接收帧，取出 IP 数据报，交给网络层

• ARP 与 DNS：相同：都是将一给定的输入实体转换为另一实体，DNS 将域名转换为IP，而 ARP 则将 IP 转换为物理地址；不同：两者的输入对象不一致，且两者在网络层次结构所在的位置不同，DNS 工作在应用层，ARP 工作在网络层。此外，DNS 为因特网中任何地方的主机解析主机名，而 ARP 只为同一个子网上的主机和路由器接口解析 IP 地址。

**[以太网]** 以太网：第一个广泛应用的高速局域网技术，也是目前占主导地位的有线局域网技术；与其它的局域网技术相比，技术简单、成本低；为提高速率，以太网技术不断演化和发展。**总线拓扑-共享式以太网**：总线(1970s 中期)：以同轴电缆作为共享传输媒体(总线)，所有节点通过特殊接口连接到这条总线上，是一种广播局域网；集线器(hub，物理层，作用于比特而非帧)(1990s 后期)：一个物理层中继器，从一个端口进入的物理信号(光，电)放大后立即从其它所有端口输出；集线器相当于共享电缆。主机和路由器直接用双绞对铜线与一台集线器相连；是一种广播局域网；如果某集线器同时从两个不同的接口收到帧，将出现一次冲突，需重传。**星型拓扑**-**交换式以太网**：交换机(21 世纪早期)：主机通过双绞线或光纤连接到交换机，交换机在端口之间存储转发帧(链路层设备)，主机与交换机之间为全双工链路，**交换式以太网不会产生冲突，不需使用 CSMA/CD 协议**。星型拓扑：各节点仅与中心节点直接通信，各节点之间不直接通信；不同于基于集线器的星型连接。

**以太网的特征：无连接**(所有的以太网技术都向网络层提供无连接服务，双方网卡之间没有握手)、 **不可靠**(接收方网卡无ACK NAK、接收方网卡丢弃 CRC 错误的帧、依靠上层协议(TCP or App)进行纠错))。缺乏可靠的传输有助于使得以太网简单和便宜。但是它也意味着传递到网络层的数据报流存在间隙。若应用使用 UDP，则会看到数据中的间隙；若使用 TCP 则不会确认包含在丢弃帧中的数据，从而引起重传，消除间隙。

## 8.8 链路层：安全无线局域网

IEEE 802.11 Security 包括 802.11 WEP(最初 802.11 规范使用的安全协议，在主机和基站之间提供较弱的加密及鉴别服务。没有密钥分发机制)和 802.11i(具有更强安全机制的 802.11 版本，提供较强的加密机制及鉴别机制，提供密钥分发机制)

**[802.11 WEP] 主机鉴别：**利用主机与 AP 共享密钥这个事实鉴别主机。无线主机向 AP 请求鉴别，AP 向无线主机发送一个 128b 的不重数，主机使用与 AP 共享的对称密钥加密不重数发送给 AP，AP 解密不重数，若与自己发送的相同则完成主机鉴别。**数据加密：**主机与 AP 共享一个 40b 都对称密钥 KS(半永久)。对于每个帧，发送方生成 24b 的初始向量 IV 添加到 KS 后面，形成一个密钥。该密钥用于生成一个密钥流 ki，用来加密帧中的第 i 个字节 di：ci=di XOR ki，将 IV 与加密后的字节 ci 放在帧中传输。接收方收到后使用相同密钥流进行解密：di=ci XOR ki。**安全漏洞：**IV 可能被重复利用或被重放攻击。例如：Trudy 通过欺骗让 Alice 加密他选择的明文 d，得到密文 c。此时可计算出 ki=ci XOR di，便得到了密钥流 k。当观察到 IV 被重用时便可破译。

**[802.11i]** 增强的安全性，可以使用各种较强的加密算法，提供了密钥分发机制，使用专门的鉴别服务器(而不是AP )来提供鉴别服务。

## 8.3 报文完整性和数字签(续) 报文完整性(又称报文鉴别)是用于验证报文是否可信的技术，如果一个报文来自声称的源并且没有被修改，则称该报文是可信的。报文鉴别涉及两个方面：来源鉴别(报文是否来自身声称的源)、内容完整性检查(报文是否被修改过)

**[朴素的设想]** 如果发送方和接收方有一个共享的密钥，可以通过加密整个报文来提供报文鉴别：发送方用共享密钥加密整个报文发给接收方，如果接收方能够正确解密收到的报文，则报文必是可信的。**问题：**混淆了机密性和报文鉴别两个概念。有时我们只想知道报文是否可信，而报文本身并不需要保密，加密整个报文会带来不必要的计算开销。将报文鉴别与数据机密性分开：发送者用明文发送报文，并在报文后附上一个标签，允许接收者利用这个标签来鉴别报文的真伪。用于鉴别报报文的标签必须满足两个条件：能够验证报文内容的完整性(是否被修改)、不能被伪造。

**[报文摘要(数字指纹)与报文鉴别码 MAC]** 将一个散列函数 H 作用到一个任意长的报文 m 上，生成固定长度的散列值 H(m)，这个散列值被称为该报文的报文摘要，也称数字指纹。使用报文摘要验证报文内容的完整性：发送方对发送的报文计算一个报文摘要，作为标签和报文一起发送给接收者，接收者对收到的报文也计算一个报文摘要，和收到的标签(发送方计算的报文摘要)进行比较。如何保证报文摘要不被修改？**基于加密算法的报文鉴别：**发送方计算报文摘要，用与接收方共享的密钥加密报文摘要，形成**报文鉴别标签(码)**。接收方用共享的密钥解密报文鉴别码，得到发送方计算的报文摘要，与自己计算的报文摘要进行比较。缺点：需要使用加密算法(加密软件通常运行的很慢，即使只加密少量的数据、加密硬件的代价是不能忽略的、加密算法可能受专利保护，如 RSA，因而使用代价很高、加密算法可能受到出口控制，如 DES，因此有些组织可能无法得到加密算法)。**基于哈希运算的报文鉴别：**使用散列函数计算报文摘要时需要包含一个密钥，但它并不用来做加密运算。发送方用双方共享的一个秘密密钥 KS 添加到报文 m 之前，然后计算报文摘要 H(KS || m)形成**报文鉴别码**，接收方收到后用报文 m 和共享密钥 KS 计算报文摘要 H(KS || m)，并与发送方发送的报文摘要比较。

**[密码散列函数]** 不同于普通的散列函数，信安领域使用的散列函数 H(称密码散列函数)必须具有以下特性：对任意给定的数据块 x，H(x) 易于计算；对任意给定的 h，要找到 x 满足 H(x)=h 在计算上是不可能的(单向性，该特性对基于哈希运算的报文鉴别很重要，如果根据 H(KS||m)=h 找到一个 x 使得 H(x)=h，则根据 x 和 m 可推出 KS)；对于任意给定的数据块 x，找到一个 y≠x 且 H(y)=H(x) 在计算上是不可能的(该特性对于使用加密算法的报文鉴别很重要，如果能够找到一个不同于 x 的数据块 y，使得 H(y)=H(x)，就可以用 y 替换 x 而不被接收方察觉；要找到一对(x,y)满足H(y)=H(x)在计算上是不可能的(抵抗生日攻击)。称满足前三个特性的散列函数为弱散列函数，满足所有特性的散列函数为强散列函数。散列函数的标准：目前使用最多的两种密码散列函数是 MD5(128b)和SHA-1(160b)，目前获得最多支持的报文鉴别方案为HMAC，可以与MD5和SHA-1一起使用。

**[数字签名]** 一个可以代替手写签名的数字签名必须满足以下三个条件：接收方通过文档中的数字签名能够鉴别发送方的身份(起源鉴别)、发送方不能否认发送过有自己签名的文档(防抵赖)、接收方不可能伪造被签名文档的内容。**通用的方法：用私钥加密报文摘要。**发送方先计算报文摘要，然后用自己的私钥加密形成数字签名，附在报文后面一起发送。接收方拷贝一份数字签名并妥善保存，以备将来需要时使用。接收方用发送方的公钥解密得到原始的报文摘要，再对收到的报文计算摘要，如果二者相符表明报文是真实的。**问题：如何可靠地获取公钥？**Bob 将公钥 EB 发布在自己的主页上，Alice获取Bob主页的请求被 Trudy截获，Trudy将假冒的Bob主页发送给Alice，主页中的公钥是 Trudy 的公钥 ET。Alice 使用Trudy的公钥加密会话密钥发送给 Bob，Trudy截获会话密钥，用Bob的公钥加密后再发送给Bob。则Alice和Bob之间所有的通信报文都被Trudy破译。根本问题：当Alice从公开的途径获取Bob公钥时，如何确认自己得到的是Bob的公钥而不是其他人的？需要公钥证书。

**[公钥证书]** *为使公钥密码体系有实际应用，每个实体必须能确认它得到的公钥确实来自声称的实体*

使用证书来证明某个主体拥有某个公钥。证书由一个可信第三方机构颁发，该机构被称为认证权威 CA。证书包含主体的公钥和 CA 的签名，任何人无法伪造或篡改证书的内容。当一个主体获得其公钥证书后，可将证书放在任何一个可公开访问的地方。证书的获取：Bob 向 CA 注册其公钥，先向 CA 提供身份证明，CA 验证了 Bob 的身份后创建证书，绑定 Bob 及其公钥，证书包含 Bob 的公钥并有 CA 的签名。证书的验证：当 Alice 需要 Bob 的公钥时，首先获取 Bob 的证书，并使用 CA 的公钥验证 Bob 的证书，得到 Bob 的公钥。证书标准：X.509，目前最常用的证书标准。X.509 建立在公钥算法和数字签名的基础上：CA 对证书内容先进行 SHA-1 散列，然后用 CA 的私钥对报文摘要进行加密，形成数字签名。为验证公钥证书的真实性，验证方用 CA 的公钥解开证书的签名，得到报文摘要，再与由证书内容计算的报文摘要比较。CA 的组织架构：不能仅使用一个 CA(流量压力，单点失效)，不能仅有一个组织运行 CA(密钥泄露，信任问题)。采用分布式公钥基础设施 PKI：提供公钥加密和数字签名服务的系统或平台，包含不同组织运行的 CA，每个 CA 有着自己的私钥，负责为一部分用户签发证书。用户自己决定使用哪一个 CA，在获取 CA 的公钥时，也需要查询该公钥的证书。一种树状结构：[Root-RA1, RA2-CA1, CA2, ...]，RA 为证书注册审批机构，CA 为证书签发机构。系统中所有实体都有根 CA 的公钥(安全的物理途径获取)，所有的根 CA 之间可以交叉验证。许多根 CA 的公钥被预装在浏览器上，由浏览器厂商认证并嵌入到软件中用于验证。证书撤销：每个证书都有有效期，过期后证书自动失效。CA 也可以显式撤销证书，这要求 CA 定期发布证书撤销列表 CRL，给出已经撤销的证书序列号。每个用户在使用证书前都需要查询 CRL。证书目录：使用 DNS 作为证书目录(标准 DNSSEC)，使用专门的目录服务器存放证书(标准 LDAP)，。证书撤销列表通常与证书存放在一起，CA 定期将 CRL 推入目录服务器，由目录服务器将CRL中列出的证书清除掉。

## 7. 4 移动用户的地址与选路准则 (目标：在移动到过程中维持正在进行的连接)

**[终端在 IP 子网之间移动]**终端进入到一个新的子网后，必须为其分配子网上的一个地址(DHCP)，并使用新的地址通信，而不保留原 IP 地址。然而，当 IP 地址改变后，终端上正在运行的应用被打断，通信的对方不知道新的 IP 地址无法通信，且由于套接字改变，在知道新地址后应用必须重新建立连接。

**[归属网络与永久地址]**归属网络：移动节点的永久居所；归属代理：移动节点在外地时为移动节点执行移动管理功能的实体；永久地址：移动节点在归属网络中的地址，外界总可以用这个地址与移动节点通信；外地网络：移动节点当前所在的网络；外地代理：外地网络上为移动节点执行移动管理功能的实体；转交地址(COA)：节点在外地网络上的地址。

**移动节点注册：**移动终端需要向归属代理通报其在外地网络的新地址。外地网络方面，移动节点在进入外地网络后，向外地代理注册，并通过外地代理向归属代理注册；归属网络方面，归属代理记录移动节点的转交地址。注册完成后，外地代理知道移动节点在自己的网络上，归属代理知道移动节点的转交地址，并记录到地址绑定表中。

**移动节点离开：**移动终端离开外地网络时，会向外地代理取消注册。外地代理并不需要显示向归属代理注销 COA，因为会有新的外地代理注册新的 COA。代理转发数据报时采用隧道技术进行封装，在移动节点看来数据报是由通信者直接发送过来的

**[间接选路]** 流程：通信者在数据报中使用移动节点的永久地址→归属代理截获数据报，转发给外地代理→外地代理收到数据报，转发给移动节点→移动节点**直接**将响应发送给通信者。**为支持移动性网络层所需要的协议：移动节点向外地代理注册协议、外地代理向归属代理注册协议、数据报封装解封装协议。三角选路问题：**移动节点会使用两个地址，永久地址(通信者用来向移动节点发送数据报，移动节点的位置对于通信者是未知的)、转交地址(归属代理用来向移动节点转发数据报)。这样就构成了三角选路：通信者-归属网络-移动节点。当通信者和移动节点在同一个网络中时就会很低效。**优势：**终端在外地网络间移动时，永久地址保持不变，因此可以维持连接，节点移动以及变换外地网络在通信者看来都是未知的，因此正在进行的通信可以保持。

**[直接选路]** 流程：通信者向归属代理请求，获知移动代理的转交地址→通信者将包发送给外地代理→外地代理将包转发给移动节点→移动节点直接向通信者发送。特征：克服了三角选路问题，但节点移动与网络变换对通信者不透明(通信者需要知道移动节点的转交地址，所有通信者(包括固定节点)都需要增加对移动通信的支持。这里的支持包括：通信者需要新的协议查询移动节点的 COA，需要新的协议告知通信者移动节点 COA 的改变)

• **最小帧长的要求(为了避免共享式以太网中广播链路的冲突)**：CSMA/CD协议规定，发送方仅在发送的过程中检测冲突；为保证在发送结束前检测到冲突，帧的发送时间必须足够长。假设信号在相距最远的两个适配器之间的往返延迟为2τ，则帧的发送时间不应小于2τ，即帧的最小长度≧链路速率×2τ(极限情况下帧到达接收方时恰好产生冲突，则再经过τ发送方才能检测到冲突)。根据早期以太网的结点最远距离(2500m)和速率(10Mbps)计算得到最小帧长为64B。802.3以太网标准(链路层相同，物理层不同)：10Mbps以太网(早期以太网)、100Mbps以太网(快速以太网)、千兆万兆以太网。

• **共享式以太网与交换式以太网** 共享式以太网：集线器的所有端口位于同一个冲突域，任一时刻(与有线区别)最多只允许一个主机发送，网络规模(节点数量)与网络性能的矛盾无法解决；交换式以太网：使用存储转发分组交换，交换机的每个端口为一个冲突域；交换机协调输出，不会向相同的接口转发超过一帧；且为全双工；多对端口可以同时通信，网络的聚合带宽=各个端口的带宽之和，**从根本上解决了网络规模与网络性能的矛盾。**交换式以太网的最小帧长及规模：交换式以太网不再使用 CSMA/CD 协议，理论上不再需要限制帧的最小长度；但为了向后兼容，帧格式及最小帧长度的限制仍然保持不变；由于交换式以太网不再使用 CSMA/CD 协议，网络直径不再受到信号最大往返时间的限制。

**[链路层交换机]** 交换机的任务是接收链路层帧，并将它们转发到链路上。**交换机自身对子网中的主机和路由器是透明的。交换机转发：用帧的目的地址查找转发表。**每个交换机内部有一张端口转发表，每个表项记录(MAC 地址、到达该 MAC 地址的端口)。当帧到来时，检查输入帧的 MAC 地址，根据表项将帧转发到一个或多个输出链路(目的 MAC 地址位于转发表中，且不需要丢弃，则转发帧到表项记录的端口；目的 MAC 地址不在转发表中，则向输入端口以外的所有端口转发该帧(扩散)) **交换机更新：用帧的源地址查找转发表。**对于到来帧，检索源地址，若没有找到则添加(源 MAC 地址、输入端口)到转发表，并设置该表项的生命周期为最大值；若找到则更新相应的表项。**交换机过滤：**若目的 MAC 地址位于转发表中，但输出端口 = 到来端口，则丢弃该帧。**【交换机的特征：**透明(**交换机没有 MAC 地址**，主机和路由器并不需要了解交换机的任何信息，也感受不到交换机的存在)、自主学习(初始交换机表为空。当一个帧到达时，交换机从源 MAC 地址了解发送节点，从到来端口了解发送节点从该端口可达，并进行记录。一段时间后，若交换机没有收到来自某地址的帧，则从表项中删除该地址)、即插即用(不需要配置)、双工、消除碰撞(可以缓存帧并保证不会同时传输多于一个帧)、异构链路支持(链路彼此隔离，从而允许不同链路以不同速率在不同媒体上运行)、可管理(提供强化安全性，可手机带宽使用的统计数据、冲突率与流量类型)】【**交换机(链路层分组交换机)与路由器(网络层分组交换机) 工作位置：**路由器工作于网络层，根据 IP 地址存储转发数据报；交换机工作于链路层，根据 MAC 地址存储转发帧。路由器工作于更高的层次，因此对每个分组的处理时间往往比交换机更长。**异构链路：**交换机无法连接异构链路(MAC 协议不同的网络)，因为交换机只是**原样转发**帧；路由器可以连接异构链路，因为路由器需要重新封装链路层帧。**广播帧：**交换机无法组织广播帧的传播，因为交换机只能学习单播 MAC 地址，而所有广播帧都会扩散发送，且通过交换机连接的所有主机位于同一个广播域中；路由器可以阻断广播帧的传播，因为路由器根据 IP 地址转发，每一个路由器端口都是一个独立的广播域。广播域：广播帧能够到达的主机集合。广播风暴：若某主机故障且无线发送以太网广播帧，则会使得整个以太网崩溃。交换机无法对其处理，而路由器则可以建立更为丰富的拓扑结构，从而为广播风暴提供防火墙保护。冲突域：共享同一条广播链路的主机集合。任何主机发送的各种帧，都可被冲突域中的其他主机接收到】【**有环网络与生成树算法：**实际的网络中不采用树状结构，而是存在冗余链路(有环)。此时交换机扩散帧会造成冗余传输，且无法停止。一种解决方案是在有环物理网络上建立无环的网络拓扑(生成树)，正常情况下根据无环拓扑转发帧。IEEE 802.1D 标准化了构造生成树的分布式算法：首先选举具有最小序列号的交换机作为生成树的根，并按照根到各个交换机的最短路径构造生成树。只有位于生成树中的交换机可以在属于生成树的边上发送给帧。当有节点或链路发生故障时，启用冗余链路，重新计算无环拓扑】【**级联交换机：**多个交换机可以级联在一起，形成更大范围的局域网。级联交换机采用与单交换机相同的自主学习策略与转发表记录操作】【**减小广播域：三层交换机与路由器(二层：链路层，三层：网络层)。**路由器可以分隔二层网络，但转发速度慢、成本高。为此引入三层交换机：具有部分路由功能，同时具有二层转发速度、专为加快大型局域网内部的数据交换而设计、在安全、协议支持等方面不如专业路由器的交换机。转发过程：路由器(获得从链路层解包的 IP 报文、用目的 IP 地址查找转发表，获得下一跳 IP 地址与端口、利用 ARP 获得下一跳 MAC 地址、用下一跳 MAC 地址构造链路层帧并发送)；三层交换机(与路由器类似，但缓存<目的 IP - 下一跳 MAC> 的映射关系。得到目的 IP 时，先查找缓存，若没有命中则执行与路由器相同的步骤)一次选路，多次转发。在机构网络中，三层交换机往往用于机构网络的核心层。连接不同的子网或虚拟局域网。而专业路由器往往用于连接机构网络与外网】

**[链路虚拟化：网络作为链路层]** 问题：早期网络间彼此不互联，且在编制方式、报文格式、选路策略、差错检测等方面不同。为此在物理网络上增加逻辑 IP 层，并在该逻辑层上统一编址，统一包格式。这样 IP 之上的层看来互联在一起的网络就是一个整体。采用网关连接不同的物理网络(逻辑层上选路到下一个网关，并将 IP 包封装在本地网络帧中发送到下一个网关)。物理网络对于 IP 层是不可见的，仅是一条虚拟链路而已

## 8.9 运营安全：防火墙与应用网关

**[防火墙]**在可信的内部网络与不可信的外部网络之间执行访问控制策略的硬件或软件系统，目的是保护内部网络免受来自外部网络的攻击。包过滤防火墙(内部网络通过有包过滤功能的路由器连接到因特网上，路由器对数据包进行逐包过滤，基于以下字段决定转发还是丢弃数据包：源 IP、目的 IP、TCP/UDP 源端口号、目的端口号、ICMP 报文类型、TCP SYN 和 ACK 标识。例如：不允许访问外部 web 网站，则丢弃所有外出的、目的端口为 80 的包；不允许外部发起的 TCP 连接，除非访问的是内网的公共 web 服务器，则丢弃进入的 TCP SYN 包，除非去往服务器的端口 80；防止因特网广播吞噬网络带宽，则除 DNS 包和路由器广播包外丢弃其他进入的 UDP 包；防止网络拓扑被探测，则丢弃所有外出的 ICMP TTL expired 包)；状态检测防火墙(包过滤防火墙孤立地过滤每一个包，但仍会允许一些异常的包进入，例如 允许 dest port=80，ack=1 的包进入，哪怕并没有相应的连接存在。状态检测防火墙可跟踪 TCP 连接的建立 syn 和关闭 fin 等状态，判断收到的包是否有意义)；应用网关(应用网关除了检查网络层及传输层协议头，还检查应用层数据。例如：允许特定的内部用户使用 telent 登录外部主机。则所有的 telent 用户必须连接到应用网关；对于授权用户，应用网关建立与目的主机的 telent 对话，并在两个连接之间中继数据；包过滤防火墙阻塞所有不源自应用网关的 telent 连接。局限性：应用网关处理开销大，速度慢，且每个被代理的应用都需要一个应用网关)

**[IDS：Intrusion detection system]** 深度数据包检查：查看包内容，如检查包中是否包含已知的病毒特征，攻击特征等；检查多个包之间的关联性，防止端口扫描或 DoS 攻击。网络中可设置多个 IDS，在不同位置进行不同类型的检查。

## 7.5 移动 IP 协议 移动 IP 是指支持移动性的因特网体系结构与协议，具有归属代理、外地代理、永久地址、转交地址、移动节点注册等特性。标准化了三个部分：代理发现、移动节点注册、数据报间接选路。

**<代理发现>** 愿意充当外地代理的路由器定期在网络上发送代理通告(ICMP)，宣布自己的存在与 IP 地址。愿意充当外地代理的路由器在代理通告中会提供一个或多个转交地址(通常使用自己的 IP 地址作为转交地址)。移动节点通过接收和分析代理通告，判断自己是否处于外地网络以及是否切换了网络。若发现自己在外地网络上，则从外地代理提供的转交地址中选择一个作为自己的转交地址。移动节点也可以主动广播请求报文(ICMP)，收到请求的代理向移动节点单播代理通告。**<移动主机注册>** 移动节点向外地代理发送一个注册请求，给出自己的永久地址、转交地址、归属代理地址以及认证信息等(UDP, 434)。外地代理记录相关信息，并向归属代理转发注册请求。归属代理收到后处理注册请求，若认证通过，则将移动节点的永久地址及转交地址保存在绑定表中，并发回一个注册响应。外地代理收到有效的注册响应后(正确性)，将移动节点记录在自己的转发表中，并向移动节点转发注册响应。当移动节点回到归属网络时，需要向归属代理隐式注销。**<数据报间接选路>** 数据报首先被归属代理得到。随后，归属代理查找地址绑定表，获得移动节点当前的转交地址。归属代理将数据报发送到转交地址，再由外地代理将数据报转发给移动节点。

**[间接选路下数据报的传输]** 归属代理如何得到数据报？若通信者不在归属网络上，则数据报首先到达移动节点归属网络上的路由器，归属网络路由器查表得知可以直接交付给归属代理，于是查找 ARP 缓存或者发送 ARP 请求,得到与移动节点永久地址对应的 MAC 地址，并将数据报封装到链路层帧中发送给归属代理。若通信者在归属网络上，则通信者查表可知移动节点的归属代理直接可达，于是查找 ARP 缓存或者发送 ARP 请求，得到与移动节点永久地址对应的 MAC 地址，并与封装数据报发送。ARP 代理：归属代理作为移动主机的 ARP 代理，代替位于外地网络的移动主机，用自己的 MAC 地址进行 ARP 响应。也就是说，**将移动主机的永久地址映射到归属代理的 MAC 地址**。**当收到移动主机的注册请求后，归属代理主动发送 ARP 请求，更新其他节点的 ARP 缓存**，防止因未更新而导致的丢包。**数据报如何到达转交地址？**归属代理收到的数据报，目的地址为移动节点的永久地址，而移动节点的转交地址位于外地网络。不能直接修改目的地址=转交地址，因为数据报的最终目的地是移动节点而不是外地代理。应当采用隧道技术：归属代理将数据包封装在更大的数据报中，发送给外地代理(Src IP = 归属代理地址，Dst IP = 转交地址，封装的数据包的 Dst IP = 永久地址)；外地代理解封装收到的数据包得到原始数据报。在移动节点注册阶段，外地代理获知了移动节点的永久地址和 MAC 地址，并将其记录在转发表中。这样，外地代理利用原始数据报中的目的 IP 地址(也就是永久地址)查找转发表，得到移动节点的 MAC 地址，并据此将数据报封装到链路层中发送给移动节点。移动节点如何发送数据报？将数据报发送给外地代理(构建 Src IP = 永久地址， Dst IP = 通信者地址，Src MAC = 移动节点 MAC 地址， Dst MAC = 外地代理 MAC 地址的帧)。外地代理在收到链路层帧后按照正常方式转发数据报。移动节点可以从代理通告报文中得到外地代理 MAC 地址。

**报文格式**

**[UDP 报文格式]**

• 伪头（12 B）：32b 源 IP 地址 + 32b 目的 IP + 8b 全 0 + 8b 协议编码 + 16b UDP

报文段长度 • 报头（8 B）：16b 源端口号 + 16b 目的端口号 + 16b UDP 报文段长度 + 16b 检查和 • 有效载荷：来自应用层的数据

**[TCP 报文格式] （1 行 4 Bytes）**

• 源/目的端口号（2B）：多路复用/分解 • 序号：首字节在字节流中的序号，非报文段序号 • 确认号：期望收到的下一字节序号，采用累计确认 • 首部长度(4b)：32bits 为单位的首部长度 通常是 20B • 标志位(6b)：URG：紧急数据指针有效、ACK 确认收到、PSH：立即交给上层、RST：不接受连接、SYN：建立连接、FIN：结束连接 • 接受窗口字段（2B）：用于流量控制 • (全部)检验和、紧急数据指针（指向最后一个字节) • 选项字段：最大段长度(MSS)：不包括TCP首部，基于路径MTU设置、窗口比例因子：实际窗口为 win-size \* 2^win-scale、选择确认(允许接收端指出缺失的数据字节) • 数据段

**[IPv4 报文格式] 报头 20B**

• IP 版本号(4b)、包含选项的报头总长度(4b，单位 4B)、服务类型(1B)(时延性要求、非实时要求等)、数据报总长度(2B，单位 1B) • 分组标识(2B)、标志位(3b，单位 8B)、偏移量(13b)，三者用于分片 • TTL(1B)、载荷的上层协议(1B)、头部检查和(2B) • 源 IP 地址(4B) • 目的 IP 地址(4B) • 选项字段，如时间戳、采用的路由、待访问路由器列表等 • 数据载荷，通常是 TCP 或 UDP 数据报

**[IPv6 报文格式] 基本头 40B + 零个或多个扩展头**

• 版本号 4b；(流量类型)PRI 8b：发送方在该域定义数据报的优先级，路由器发现网络拥塞时，按优先级从低到高的顺序丢弃包。IPv6 将网络流量划分为两大类：受拥塞控制的流(非实时流属于这一类，优先级 0～7，按照重要性及用户体验设定)；不受拥塞控制的流(实时多媒体流属于这一类，优先级 8～15。可以按照用户要求的服务质量等级定义)；流标签 20b：用于标识一条数据报的流(给属于特殊流的分组加标签)一般而言，流是具有相同传输特性(源/目的、优先级、选项等)、并要求相同处理(使用相同的路径和资源、具有相同的服务质量和安全要求等)的一系列数据包；流由源地址和流标签(flow label)唯一标识；流标签由发送方分配，由[源地址，流标签]唯一标识。不支持流的节点忽略该域；支持流的路由器维护一张流表(flow table)，记录每一个流需要的处理；收到数据包后，根据源地址和流标签查找流表，进行相应的处理。流的引入使得 IPv6 具备了对数据包进行区分处理的能力，极大降低数据包分类的复杂度 • 有效载荷长度 20b：无符号整数，给出 40B 首部后的字节数；扩展头(下一个首部) 8b；跳数限制 8b • 源地址与目标地址各128b；数据。

**与 IPv4 固定头相比，IPv6 的基本头中去掉了以下一些字段：**IHL(IPv6 的基本头总是 40 字节长、与分片相关的字段(IPv6 路由器不负责分片与重组)、头校验(计算校验和太花时间，TTL 导致每次都要重算)。增加了：流标签(支持对数据包区分处理)。改变了以下字段的作用：Type of Service：代之以 Traffic Class；总长度：代之以载荷长度；Protocol：代之以 Next header，允许任意扩展选项。

**[ICMP 报文格式]**

• Type 8b，共定义了 15 种；•Code 8b，用于对某类报文作进一步的区分；•Checksum 16b，ICMP 报文的检查和；•内容：与报文类型有关，报告错误的ICMP报文包含触发该错误的数据报的头部和前 8个数据字节。

**[ARP 报文格式]：跨越网络层与链路层的协议**

• 硬件类型 2B：硬件接口类型。对于以太网，该值为“1”；• 协议类型 2B：高层协议地址类型。对于IP地址，该值为 0800；• 硬件地址长度与软件地址长度各 1B；• 操作 2B：ARP 请求为 1，ARP 响应为 2；

• 发送方硬件地址 6B、发送方 IP 地址 4B、目标硬件地址 6B、目标 IP 地址 4B

**[以太帧结构]**

• 前导码 8B：7个10101010，后跟10101011，用于在双方之间建立时钟同步，一般不计入以太帧长度。

• 目的 MAC 地址、源 MAC 地址 各 6B • 类型 2B：用于指出数据所对应的高层协议(ARP、IP)的编号

• 数据 46~1500B，不足 46B 填充至 46B，超出需要分片

• CRC 校验码 4B，对从双地址到数据之间部分的内容计算得到

**[802.11 帧结构]**

• 帧控制 2B，细分为类型 2b(RTS CTS ACK data)、子类型 4b、协议信息等

• 持续期 2B，用于 RTS/CTS 的传输预留时间(传输时间 + 收到确认时间)，RTS、CTS、data 均有

• 地址字段 1 6B，为接收该帧的无线站点的 MAC 地址(目的地址)

• 地址字段 2 6B，为传输该帧的无线站点的 MAC 地址(源地址)

• 地址字段 3 6B，为连接 AP 的路由器接口的 MAC 地址，在有线局域网与 BSS 互联中起关键作用(AP 仅对无线终端可见，对于固定网络上的设备时不可见的。AP 连接路由器的有线端口没有 MAC 地址，所以 AP 需要将路由器与自己接口的 MAC 地址作为目的地址，无线终端的 MAC 地址作为源地址重新构造链路层帧。期间涉及到 802.11 无线帧与 802.3 以太网帧的转换)

• 帧序号 2B，用于可靠传输 • 地址字段 4 6B，用于自组织模式中的互相转发 • 有效载荷0-2312B、CRC 4B

**[ESP 数据包格式]**

• ESP 头：包含 SPI(32b 数，和目的 IP、安全协议结合起来唯一标识数据报的 SA)、SeqNum(32b，对 SA 上发送的数据包进行编号，共接收端检测重放攻击。为防止重复，在流量到达 2^32 之前需要重新协商 SA 和密钥) • 载荷：原始数据包中被加密部分的密文 • ESP 尾，包括填充(若有)、填充长度、下一个头。ESP 尾也需要加密 • Authentication Data：包含报文鉴别码的可变长度区域

## 补充问题

**互联网中的可靠传输：**网络层是不可靠的，只负责把一个数据包从发送主机运动到接收主机，数据包丢失是不管的。可靠传输是由传输层实现的，在发送完一个分组后，发送方须暂时保留已发送的分组的副本以备重传。部分链路层协议实现了可靠传输，但不是强制实现。链路层可靠传输需要增加很多机制来保证数据正确性，这是由于早期链路质量很差。现在的高可靠性高带宽的链路因为错误很少，所以可以把纠错工作放到高层处理。因此，在数据链路层只是检查错误并不纠错，而在传输层纠错，所以数据链路层也是不可靠的。

## **常用协议与端口号：**HTTP 80、FTP 20(数据)/21(控制) 、SMTP 25、POP3 110、DNS 53、DHCP 67(客户)/68(服务器)、BGP 179(半永久TCP)

**拥有 n 台主机的子网**需要 n+3 个 IP 地址(主机号为0、为1以及网关路由器)

**基于下一跳路由器转发的优缺点：**优点-每个路由表项只需保留下一跳的地址，无需给出完整的路由(路径)；缺点-要求下一跳路由器知道剩余的路径信息或网络中的所有路由器信息保持一致

## **解决 IPv4 地址不足的方法：**采用无类别编址 CIDR 使 IP 地址的分配更加合理、采用网络[地址转换](https://so.csdn.net/so/search?q=%E5%9C%B0%E5%9D%80%E8%BD%AC%E6%8D%A2&spm=1001.2101.3001.7020" \t "https://blog.csdn.net/xxxli_/article/details/_blank) NAT 以节省全球 IP 地址，实现复用、采用动态地址分配 DHCP 以及时腾出不被使用的地址、采用具有更大[地址空间](https://so.csdn.net/so/search?q=%E5%9C%B0%E5%9D%80%E7%A9%BA%E9%97%B4&spm=1001.2101.3001.7020" \t "https://blog.csdn.net/xxxli_/article/details/_blank)的新版本的 IPv6 [治本方案]

**ISP 与 AS：**因特网服务提供商(ISP)是多个分组交换机与多段通信链路组成的网络，自治系统(AS)是一组通常处在相同管理控制下的路由器组成的网络。一般 ISP 内部可以被拆分为多个互联的 AS，也可以使用一个庞大的 AS。

**RIP、OSPF 协议的缺点：**RIP[更新周期(30s)过短、未进行区域划分]；OSPF 缺点[用可靠广播方式在整个区域广播所有节点的链路状态，开销过大]

**实际的路由选择算法：**路由的属性包括<NEXT HOP,AS PATH>，记录了一条经过一系列外部 AS 的路径以及路径起始路由器端口的 IP 地址。BGP按照如下顺序依次进行选择：对于经过AS的本地偏好、距离当前AS最短的AS PATH、距离发送方最近的NEXT HOP、BGP标识符。

**链路层的信道：**广播信道(用于连接有线局域网、卫星网和混合光纤同轴电缆接入网中的多台主机，需要媒体访问协议协调帧传输)、点对点通信链路(用于长距离链路连接的终端间)

**移动设备与服务器之间通信的链路：**移动主机与基站之间的 WIFI 链路、基站与交换机间的以太网链路、交换机与路由器、路由器之间的链路、路由器与服务器之间的以太网链路

**交换机毒化：**攻击者发送大量具有伪造 MAC 地址的帧，填满了交换机的表项。此时大部分合法来源的帧无法在表项中存储，都需要广播发送，从而被攻击者捕获。

**集线器、路由器与交换机：**• 路由器(Router)是连接因特网中各局域网、广域网的设备，工作在网络层。在路由器中记录着路由表，它会根据信道的情况自动选择和设定路由，以最佳路径、按前后顺序发送数据包。路由器没有生成树限制，可以以丰富的拓扑结构构建因特网。此外，路由器可以避免广播风暴的产生。但是路由器的使用需要配置转发表，且对于分组的处理时间更长。• 交换机(Switch)是一种用于电(光)信号转发的网络设备，工作在链路层。它可以为接入交换机的任意两个网络节点提供独享的电信号通路，把传输的信息送到符合要求的相应路由上。交换机是即插即用、无需配置、自主学习的，且具有相对较高的转发速率。然而，为了限制广播帧的循环，交换机会将路径限制在一棵生成树上。此外，对于大型交换网络需要保存庞大的 ARP 表，这将带来十分可观的ARP广播帧流量。而交换机对广播帧没有任何保护措施。• 集线器(Hub)是指将多条以太网双绞线或光纤集合连接在同一段物理介质下的设备，工作在物理层，即插即用。集线器传输数据的时候，如果局域网中的一台主机要发送消息，则局域网内的所有主机都可以接收到这个消息，安全性较差，而且每一次只能有一台主机发送，否则会产生冲突。

**移动网络的寻址问题：**当移动结点位于一个外部网络时，所有指向此结点永久地址的流量需要导向外部网络。解决方法一：外部网络通过向所有其他网络发通告，告诉它们该移动结点正在它的网络中(不适用于大规模网络，否则百万级表项使得网络核心路由器的负载过大)；方法二：将移动性功能从网络核心搬到网络边缘，由该移动结点的归属网络来实现。

**转交地址(COA)：**注册时外地代理提供的一系列 IP 地址。移动主机选择其中一个作为自己的转交地址，外地代理将接收发送给该 COA 的数据报，然后转发给移动节点。

**网络安全的四大属性：**机密性(为了使信息不泄露给非授权用户、非授权实体或非授权过程，或供其利用，防止用户非法获取关键的敏感 信息或机密信息。通常采用加密来保证数据的机密性)、完整性(是为了使数据未经授权不能被修改，即信息在存储或传输过程中保持不被修改、不被破坏和不被丢失。主要包括软件的完整性和数据的完整性两个方面的内容。软件完整性是为了防止对程序的修改，如病毒。数据完整性是为了保证存储在计算机系统中或在网络上传输的数据不受非法删改或意外事件的破坏，保持数据整体的完整)、可用性(为了被授权实体访问并按需求使用，即当用户需要时能够在提供服务的服务器上进行所需信息的存取。例如：网络环境下拒绝服务、破坏网络和破坏有关系统的正常运行等，都属于对可用性的攻击)、可控性(是为了对信息的传播及内容具有控制能力。任何信息都要在一定传输范围内可控，如密码的托管政策等)。

**数字签名是一种可提供发送方身份鉴别、报文完整性和防发送方抵赖的安全机制。请给出数字签名最常见的构造方法：**当实体 A 需要为报文 M 生成数字签名时，A 首先用一个

散列函数计算 M 的报文摘要，然后用 A 的私钥加密该报文摘要，生成数字签名。**说明数字签名为何可以提供以上安全服务：**A 的私钥是只有 A 知道的秘密，任何其它实体无法得到，因而一个有效的数字签名可提供发送方身份鉴别。报文摘要可用于检测报文的完整性，对报文内容的任何修改将产生不同的报文摘要。用 A 的私钥加密后的报文摘要是不可伪造的，从而数字签名就将 A 与报文 M 紧密关联在一起，既能提供报文完整性服务，也能防止发送方抵赖。

## 8.7 网络层：IP 安全协议(IPSec) Ipv4 在设计时没有考虑安全性，缺少对通信双方身份的鉴别，容易遭受地址欺骗攻击；缺少对网络中数据的完整性和机密性的保护，数据很容易被窃听、修改甚至劫持。IPSec 是 IETF 以 RFC 形式公布的一组安全协议集，目标是把安全特征集成到 IP 层，以便对因特网中的安全业务提供低层的支持

**[专用网与虚拟专用网]** 专用网是通过电信专线将分散在各地的计算机(网络)连接而成的网络，安全性好，但代价高。虚拟专用网(VPN)是建立在公用网上的一个覆盖网络，在逻辑上与其他流量隔离，数据在发送到公用网之前会进行加密。VPN 的典型结构：在每一个区域网上设置一个安全网关，在每一对安全网关之间创建一条穿过因特网的隧道，在隧道中使用 IPSec。VPN 的优点：可以在一对局域网间提供完整性控制及机密信服务，甚至对流量分析也有相当的抵御能力。对因特网中的路由器及用户软件是透明的，只要系统管理员设置好安全网关就可以了

**[IPSec]安全体系框架**：IPSec 在提供了一个用于集成多种安全服务、加密算法及安全控制粒度的安全体系框架。提供的安全服务包括：访问控制、无连接完整性、数据起源认证、抗重放攻击、机密性等。IPSec 的安全机制独立于算法，因此在选择和改变算法时不会影响其他部分的实现。IPSec 提供多种安全控制粒度，包括一条 TCP 连接上的通信，一对主机之间的通信，一对安全网关之间的所有通信。用户可以为数据通信选择合适的安全服务、算法、协议和控制粒度。**组成：**从技术上说，IPSec 主要包括两个部分：IPSec 安全协议，包括 AH 和 ESP 两个安全协议，定义了用于安全通信的 IP 扩展头和字段，以提供机密性、完整性和源鉴别服务；密钥管理协议，定义了通信实体之间进行身份鉴别、协商加密算法及生成共享会话密钥的方法。将以上两部分绑定在一起的是称为安全关联(SA)的抽象。 SA 是通信对等实体之间对某些要素的协定，如使用的安全协议、协议的操作模式、使用的密码算法、密钥及密钥的生存期等。SA 是两个通信端点间的一个单工连接，由一个安全参数索引(SPI)唯一标识，如果在两个方向上都需要安全通信则需要建立两个 SA。SPI 携带在数据包中，由数据包的处理进程用来查找密钥及相关信息。SA 可以建立在一对主机之间，一台主机与一个安全网关之间或一对安全网关之间。**使用模式：**传输模式(IPSec 头被插入到原始 IP 头和传输层头之间，路由器根据原始 IP 头转发数据包)、隧道模式(原始数据包被封装在一个新的 IP 包中，IPSec 被放在新的 IP 头和原始 IP 头之间，路由器根据外层 IP 头的信息转发数据包。隧道的端点(外层 IP 头的地址)通常是一个支持 IPSec 的安全网关)。传输模式比隧道模式占用更少的带宽，隧道模式则更加安全(内部细节，如原始 IP 头被隐藏；可将一对端点之间的通信聚合成一个加密流，从而有效防止入侵者进行流量分析) **封装安全载荷 ESP：**ESP 协议提供数据机密性、无连接完整性、抗重放攻击、数据起源鉴别和有限的数据流机密性服务。原始数据包的载荷部分为加密，因而可提供数据机密性；HMAC 覆盖数据包载荷部分，可提供无连接完整性服务；ESP 头中有序号，且被 HMAC 覆盖，可抵抗重放攻击；ESP 隧道模式中原始 IP 头也未 HMAC 覆盖，因而可提供数据起源鉴别；隧道模式中原始 IP 头也被加密，路由器只能看到外层 IP 头，因而 ESP 隧道模式可提供数据流机密性服务。

## 8.6 传输层：安全套接字层 SSL SSL 建立在 TCP 之上，依靠 TCP 提供可靠的端到端连接。SSL 向基于 TCP 的网络应用提供安全的传输层服务，如支持 web 浏览器和服务器之间的安全通信 https，是涉及到两个层次的一组协议。SSL 位于应用层与传输层之间，包含三个高层协议(用于 SSL 交换管理)和一个记录协议(为各种上层协议如 HTTP 提供基本的安全服务)。安全服务：机密性、数据完整性、服务器鉴别、客户鉴别(可选)**SSL 握手协议：**允许服务器和客户之间相互鉴别并协商加密算法、MAC 算法及密钥的握手协议，由客户和服务器之间的一系列报文交换组成。包括：**能力协商**(浏览器向服务器发送建立 SSL 会话的请求报文，说明可支持的 SSL 协议最高版本、支持的加密算法(按优先级从高到低排列)和压缩方法等，以及浏览器选择的一个随机数 Rc。服务器从浏览器给出的选择中确定合适的 SSL 版本号、加密算法和压缩方法，与服务器选择的一个随机数 Rs 一起发送给浏览器)、**服务器鉴别**(服务器向浏览器发送它的公钥证书(和必要的证书链)以及其他信息，浏览器检查签发证书的 CA 是否在浏览器的可信 CA 列表中，若不再则向用户警告该问题，若在则使用该 CA 的公钥验证证书，得到服务器的公钥。如果客户也需要被鉴别(收到服务器的证书请求)，则浏览器向服务器发送他的公钥证书)、**生成密钥**(浏览器生成一个 48B 的随机数，称预密钥，用服务器的公钥加密后发给服务器。客户和服务器各自从预密钥、Rc 和 Rs 中计算加密数据所需要的会话密钥，以及计算 MAC 所需要的密钥)。双方发送给所有握手报文(级联)的 MAC。**SSL 记录协议：**SSL记录协议为SSL连接提供两种服务：机密性，通过加密 SSL 载荷实现；完整性，通过报文鉴别码保护。SSL记录协议的操作过程：从上层接收一个要传输的应用报文，将报文划分成长度不超过 2^14B的数据块；(可选)对数据块进行压缩；对数据块生成基于哈希运算的报文鉴别码；使用对称密钥算法对(压缩的)数据块及报文鉴别码进行加密，加密算法可以是DES、3DES、IDEA、RC等；在处理完的数据块前加上SSL头，包括内容类型，SSL版本号，压缩数据块的长度等

## 8.5 应用层：电子邮件安全 电子邮件安全包括真实性和机密性两个方面。目前最流行的两个电子邮件安全协议是 PGP(一个开放源码的电子安全邮件软件包，提供对邮件的保密、鉴别、数字签名和压缩服务。PGP 较多地用于个人电子邮件安全，是因特网电子邮件的事实标准)、S/MIME(基于公钥加密技术对 MIME 所做的安全扩展。S/MIME 可能作为一种工业标准被商业组织或一些机构使用)。**PGP 提供五种服务：鉴别**(PGP 使用基于公开密钥算法的数字签名提供鉴别服务。生成可供鉴别的电子邮件：发送方创建电子邮件，用 SHA-1 计算邮件的报文摘要，然后用发送者的私钥加密报文摘要，形成数字签名，将数字签名附在报文的前面与报文一起发送 Sig||Data)、**机密性**(PGP 使用对称密钥算法保护邮件的机密性，发送方将选择的一次性会话密钥用接收方的公钥加密后，与报文一起发送给接收方。仅使用机密性服务的过程：发送方 A 生成一个报文和随机的 128b 数(一次性会话密钥)，先用会话密钥加密报文，在用接收方 B 的公钥加密会话密钥，将加密后的会话密钥放在报文前面，与报文一起发送。 KB+(KA-B)||KA-B(Data) )、**压缩**(PGP 在完成签名后，在加密报文之前会对报文进行压缩，压缩算法采用 ZIP。完整形式为 KB+(KA-B)||KA-B(ZIP(Sig||Data))。在加密报文前进行压缩，一方面可以减少要加密的数据量，另一方面压缩后消息冗余很少，增加密码分析的困难)、**兼容电子邮件**(PGP 使用 base64 编码，将二进制数据转化为简单 ASCII 文本以解决邮件的传输问题，PGP 可被配置为仅对报文的某些部分，如签名部分进行 base64 编码转换)同时使用以上种服务的过程：发送方先对明文报文计算签名，将签名放在报文前面，签名与明文一起被压缩，用会话密钥对压缩后的数据块进行加密，再用接收方的公钥加密会话密钥放在报文的前面，将整个数据块转化为 base64 编码格式。Base64[ KB+(KA-B)||KA-B(ZIP(Sig||Data))]。**分段**(许多电子邮件系统能够接收的最大报文程度不超过 5wB，PGP 在完成对报文的全部处理后，自动将超过长度的报文分成小块传输，会话密钥和签名只在第一个片段中出现。接收端去掉每个片段的头部，然后将所有的片段重新组装成一个数据块)

## 8.4 端点鉴别 **[鉴别协议 ap]** 目标：Bob 希望 Alice 能够确认自己的身份

ap1.0 Alice 声称自己是 Alice。问题：其他人 Trudy 也可以声称自己是 Alice；ap2.0 Alice 用自己的 IP 地址与声明封装成数据包发送。问题：Trudy 可以用 Alice 的 IP 地址创建一个数据包(IP 地址哄骗)； ap3.0 Alice 向 Bob 发送明文口令证明自己(口令只有 Alice 知道)。问题：Trudy 监听到 Alice 发送的明文口令，再伪造声明发送给 Bob； ap3.1 Alice 将口令加密后再发给 Bob(对称密钥)。问题：Trudy 截获数据包中口令，待下次通信时发送给 Bob(重放攻击)； ap4.0 Bob 向 Alice 发送不重数 R，Alice 用共享密钥加密 R 回送给 Bob。缺点：需要共享对称密钥； ap5.0 采用公开密钥算法加密不重数

**[X.509 的单向鉴别服务]** 单向鉴别：涉及发送方到接收方的一次报文传输。**内容：**t\_A||r\_A||ID\_B||Data||Kb+(Ka-b)||Signature\_A。**其中：**t\_A 为时间戳，由报文的产生时间和到期时间组成；r\_A 为 A 随机选择的一个不重数，供接收者检测重放攻击；ID\_B 为 B 的标识，指示报文的接收者；Data 为数据信息；Ka-b 为当数据需要保密时用来加密 Data 的对称密钥；Kb+ 为 B 的公开密钥，用于加密对称密钥 Ka-b；Signature\_A 为 A 的数字签名，对 t\_A、r\_A、ID\_B、Data 四个部分的明文生成，并用自己的密钥 Ka- 加密。**接收方鉴别：**B 用自己的私钥 Kb- 解出 Ka-b，再用 Ka-b 解出 Data，并计算报文摘要；再用 Ka+ 从 Signature\_A 中得到原始报文摘要进行比较。

## 8.2 基础密码学原则 基本术语：明文(欲加密的原始数据)、密文(明文经加密算法作用后的输出)、密钥(加密和解密时使用的参数)、密码分析(破译密文)、密码学(设计密码和破译密码的技术统称为密码学)

**[加密算法] 分类：**按照加密密钥与解密密钥是否相同来区分：对称加密算法(加密密钥与解密密钥相同)、非对称加密算法(加密密钥与解密密钥不同，也称公开密钥算法)；按照明文被处理的方式来分：块密码(分组密码)(每次处理一个明文块，生成一个密文块)、流密码(处理连续输入的明文流，并生成连续输出的密文流)。**传统加密算法：**替换(用密文字母替换明文字母，但字母位置不变)，例如凯撒密码、单字母表替换、多字母表替换，密钥为字母映射表；换位(保留明文字母不变，但改变字母的位置)，例如列替换密码。

**针对加密系统的密码分析攻击：**唯密文攻击(密码分析者仅能根据截获的密文进行分析，以得到明文或密钥，是对密码分析者最不利的情况)、已知明文分析(密码分析者除了有截获的密文外，还有一些已知的明文-密文对来帮助破译密码，以得出密钥)、选择明文攻击(密码分析者可以任意选择一定数量的明文，用被攻击的加密算法加密得到相应的密文，以利于将来更有效的破解由同样加密算法以及相关密钥加密的信息)。一个安全的加密系统必须能够抵御选择明文攻击。

**[密码的安全性]** 传统加密方法的安全性建立在算法保密的基础上。现代密码学的基本原则是，加密与解密的算法是公开的，只有密钥是需要隐藏的，因为日常通信的群体数量太大，很难做到算法保密，且计算机需要标准化、自动化算法。如果一个加密算法的密钥满足以下两个条件之一，则该算法被称为是**计算安全的：**破译密文的代价，超过信息本身的价值；破译密文所需的时间超过信息的有效生命期。现代密码学中密码的安全性是通过算法的复杂性和密钥的长度来保证的。

**[现代密码学：对称加密算法]** <DES：Data Encryption Standard> 1997 年成为美国的数据加密标准。DES 是块加密算法，每次输入 64b 的明文块，输出 64b 的密文块。DES 是基于迭代的算法,每一轮迭代执行相同的替换和换位操作，但使用不同的密钥(使用一个 56b 的主密钥，每一轮迭代的 48b 子密钥由主密钥产生)。DES 是对称加密算法，加密和解密使用相同的函数，两者的不同只是子密钥的次序刚好相反。缺点：密钥长度不够长，迭代次数不够多。<Triple DES (3DES)> 使用两个密钥进行三轮 DES 计算：第一轮密钥 K1 加密(E)，第二轮密钥 K2 解密(D)，第三轮密钥 K1 加密(E)，最终密文为 K1+(K2-(K1+(m)))。解码过程为 DED。一些额外问题：为什么使用两个密钥而不是三个？因为 112b 的密钥已经足够长；为什么不用两重 DES 模式(EE)而是三重 DES？ 考虑采用 EE 模式的两重 DES，且攻击者已经拥有了一个匹配的明文-密文对 P1,C1，即 C1 = EK2(EK1(P1))。令 X = EK1(P1)=DK2(C1)，则只需要计算使得 EK1(P1)=DK2(C1) 的 K1 和 K2，密钥空间为 2^56 而非 2^112；为什么是 EDE 不是 EEE？为了与单次 DES 兼容，只需令 K1=K2 即可。<AES> 2001 年代替 DES 称为新的对称加密标准，输入输出变为 128b，密钥长度为 128b 或 192b 或 256b。<密码块链接 CBC> 若每个明文块被独立加密，则相同的明文块生成相同的密文块，容易被重放攻击利用。CBC 操作让发送方生成一个初始点随机向量 c(0)，以明文发送给接受者。第一个明文块加密前与 C(0) 异或，随后每个明文块加密前与前一个密文块先异或。相同明文块几乎不可能生成相同密文块。

**[现代密码学：非对称加密算法(密钥公开算法)]** 对称加密算法的问题：要求发送方和接收方使用相同密钥，则在发送方选择密钥后应如何安全传递给接收方？采用非对称加密算法，双方不共享密钥，则不存在密钥传递问题：加密密钥公开，解密密钥私有。密钥应当满足两项基本要求：K-(K+(m))=m(交换成立)、给定公钥 K+，不可能计算出 K-。

**公开密钥算法：**每个用户生成一对加密密钥和解密密钥，加密密钥放在一个公开的文件中，解密密钥自己妥善保管。当 Alice 希望向 Bob 发送加密信息时，Alice 从公开的文件中查询到 Bob 的加密密钥，并用其加密信息发送给 Bob，Bob 用自己的解密密钥解密信息。Alice 也可以用自己的私钥加密信息，Bob 收到后用 Alice 的公钥解密信息，从而确保该信息来源于 Alice。**公开密钥算法应满足的条件：**已知加密密钥从明文计算出密文是容易的，已知解密密钥从密文计算出明文是容易的；从计算上来说，根据加密密钥推出解密密钥是不可能的，根据加密密钥和密文计算出原始明文是不可能的。**公开密钥算法的问题：**首先，入侵者知道公钥与加密算法，可以据此发起选择明文攻击，故密钥选择和加密算法必须保证任何一个入侵者都不能确定 Bob 的私钥，或以某种方式解密或猜出对应的明文；其次，加密密钥是公开的，任何人都可能向 Bob 发送一个用 Bob 公钥加密的报文，而在单一共享密钥情况下，发送方知道共享密钥的事实，就已经向接收方隐含地证实了自己的身份。此时就需要数字签名把发送方和报文绑定起来。**RSA 算法：**生成密钥：选择两个大素数 p 和 q (典型值为 > 10^100)，计算n=p×q 和 z=(p-1)×(q-1)，选择一个与 z 互质的数，令其为 d，找到一个 e 使其满足e×d≡1(mod z)，则公开密钥为 (e,n)，私有密钥为(d,n)。加密：将明文看成一个比特串，划分为一个个数据块 M，满足 0<M<n，对每个数据块 M 计算C=M^e(mod n)，则 C 即为 M 的密文。解密：对每个密文块C，计算 M=C^d(mod n)，则 M 即为要求的明文。特性：公钥和私钥的先后顺序对结果无关，即K-(K+(m))=m=K+(K-(m))；优点：安全性好(RSA的安全性建立在难以对大数提取因子的基础上，这是数学界尚未解决的难题)、使用方便(免除传递密钥的麻烦)；缺点：计算开销大，速度慢。应用：RSA一般用来加密少量数据，如鉴别、数字签名或发送一次性会话密钥等

## 8.1 网络安全概述 网络安全指网络系统的硬件、软件及其系统中的数据受到保护，不受偶然的或者恶意的原因而遭到破坏，更改、泄露，系统连续可靠运行，网络服务不中断。

**[安全通信的特征] 机密性：**报文内容机密性(仅发送方和希望的接收方能够理解报文的内容)、通信活动的机密性(通信活动或其特征不被外界察觉)；**端点鉴别：**发送方和接收方都能够证实通信过程中涉及的另一方；**报文完整性：**报文来自真实的来源，且传输过程中未被修改；**运行安全性：**网络系统正常运行，网络服务可用。

**[网络安全攻击] 被动攻击:**试图从系统中获取信息但不对系统产生影响。例如偷听(监听并记录网络中传输的内容)、流量分析(从通信频度、报文长度等流量模式推断通信的性质) **主动攻击:**试图改变系统资源或影响系统的操作。例如伪装(一个实体假冒另一个实体)、重放(从网络中被动地捕获一个数据单元经过一段时间后重新发送到网络中)、报文修改(修改 插入 删除报文或报文部分内容)、拒绝服务(阻止通信设施的正常使用或管理)

**[常见的安全机制] 加密：**使用数学算法对数据进行变换，使其不易理解；**鉴别：**通过报文交换证实一个实体的身份，以防假冒；**报文完整性：**验证一个报文是否可信，包括来源和内容；**数字签名：**附加在一个数据单元后的数据，用来证明数据单元的来源及完整性，以防伪造以及抵赖；**流量填充：**在数据流间隙中插入比特，以挫败流量分析的企图；**访问控制**：通过授权机制限制用户对资源的访问，防止越权行为。

**第八章 网络安全**

## 7.6 无线和移动对上层协议的影响

## **[引入的问题] 无线方面：**误码率高(传输距离、环境干扰、多径传输造成信噪比较低)，丢包率高(由于误码、发送冲突造成传输失败)，延迟增大(由于冲突重传造成延迟增大)；

**移动方面：**丢包增加(切换，包括 AP 切换、交换机更新转发表、移动 IP 注册，造成丢包)，延迟增大(切换及间接选路带来延迟增大)。

**[带来的影响](应用层表现为带宽较低，以下为传输层层面)** 逻辑上没什么影响：为上层协议提供的依然是尽力而为的服务，因此 TCP 和 UDP 也可以运行在无线网络上；性能上有很大影响：丢包率升高，传输延迟增大。TCP 将丢包(高延迟也视作丢包)解释为拥塞，因而会不必要地减小拥塞窗口，导致应用吞吐率很低。无线链路、有线/无线混合链路上的 TCP 拥塞控制是一个研究问题。**TCP 拥塞问题可能的解决方法：**本地恢复(在比特差错出现的当时当地恢复，例如 ARQ 协议)、 TCP 发送方知晓无线电路，并进行区分处理(如只对优先网络中的拥塞性丢包采用拥塞控制)、分离连接方法(移动用户和其他端点之间的端到端连接被切分为两个传输层连接，一个从移动主机到无线接入点，一个从无线接入点到其他通信端点。移动用户-无线接入点之间为新的选择性重传协议，而无线接入点和其他通信端点之间为标准的 TCP 或 UDP 协议)。

## 7.3 IEEE 802.11 wireless LANs (Wi-Fi) 标准

**[802.11 组成]** 802.11b(2.4~5 GHz,up to 11 Mbps)、802.11a(5~6 GHz,up to 54 Mbps)、802.11g(2.4~5 GHz,up to 54 Mbps)、802.11n(多天线，.4~5 GHz,up to 200 Mbps，多输入输出)。均使用 CSMA/CA 作为 MAC 协议，链路层帧格式相同，都具有降低传输速率以到达更远距离的能力，都支持基站模式和自组织模式，在物理层上不同

**[802.11 无线局域网架构]** 802.11 无线 LAN 的基本组成单元是基本服务集 BSS。一个 BSS 包括若干无线终端和一个无线接入点 AP(AP 连接到路由器或交换机从而接入因特网)。每个无线接口均有全局唯一的 MAC 地址，AP 与路由器相连的端口没有 MAC 地址(AP 对于路由器是透明的)。**信道与关联：**802.11 将通信频段划分成若干信道，每个 BSS 分配一个信道。管理员安装 AP 时，为每个 AP 分配一个服务集标识符 SSID，并选择 AP 使用的信道。相邻 AP 使用的信道可能相互干扰。主机需要与一个 AP 关联(主机扫描信道，监听各个 AP 发送到信标帧(包含 AP 的 SSID 和 MAC 地址)，并选择一个 AP 进行关联(可能需要身份鉴别)，最后使用DHCP获得 AP 所在子网中的一个 IP 地址)。**主动扫描**(主机广播探测请求帧，AP 发送探测响应帧，主机从收到的探测响应中选择一个 AP 发送关联请求，该 AP 发送关联响应帧。是与DHCP类似的二次握手)；**被动扫描**(主机监听 AP 发送的信标帧，并选择一个AP发送关联请求，AP发送关联响应帧。缺点：速度较慢)

**[802.11 MAC 协议]** 多个无线站点或 AP 希望同时经相同信道传输数据帧。**目标：避免冲突**。采用 CSMA(发送前监听信道，不与当前正在进行的发送冲突)。发送过程中不检测冲突，一方面是因为接收信号的强度远小于发送信号，因此很难检测，另一方面存在无法检测的隐藏结点。使用链路层确认(接收方成功收到帧后需要发送确认帧，以便让发送方知道是否发送成功)。**操作模式：**PCF (使用较少)，该模式只能用于有基础设施的无线网络，由基站控制单元内的所有通信活动。采用轮询模式：基站依次询问单元中的结点，被询问到的结点可以发送它们的帧，不会有冲突发生。DCF 可用于无基础设施的无线网络，所有结点(AP 和无线终端)使用 CSMA/CA 协议竞争信道。DCF 支持信道预约机制(可选)和无信道预约的机制(必须支持)。

**不使用信道预约机制的 CSMA/CA：**当结点有帧要发送时，首先侦听信道。若一开始就侦听到信道空闲，则等待 DIFS 时间后发送帧；否则，选择一个随机回退值，在侦听到信道空闲时递减，在信道忙时冻结。当计数值减为 0 时，发送整个帧并等待确认。若收到确认帧，则表明发送成功，此时重新帧听。若未收到确认帧，则在更大的范围内随机选择回退值并等待。不同结点随机选择的回退值决定了其发送顺序。此时可能存在碰撞：两个站点对 AP 不隐藏但彼此相互隐藏、两个节点选择了非常接近的随机回退值等。

**使用信道预约机制的 CSMA/CA：(缓解隐藏节点问题)**采用短请求发送控制帧 RTS 和短允许发送控制帧 CTS 处理隐藏终端。假设 A 将向 AP 发送一个数据帧，则 A 首先发送 RTS 帧，帧中给出随后要发送的数据帧及收到 ACK 帧的总时间。AP 收到后回复一个 CTS 帧，帧中给出同样的帧。A 收到 CTS 帧后发送数据帧，AP 收到数据帧后发送 ACK 帧。收到 RTS(A 附近)和 CTS(AP 附近)的结点均沉默指定的时间，让出信道。若收到 ACK 帧，则可以提前结束沉默。若 A 和 B 同时发送 RTS 帧，则产生冲突，不成功的发送方随机等待一段时间后重试。**当要发送的数据较少时，一般不使用信道预约机制。RTS 和 CTS 很短，因此冲突的损耗小，但引入了时延消耗信道资源。*帧间距机制(允许 DCF 和 OCF 共存)：***在一次成功的发送后，时间段划分为 0-SIFS (谁都不能发送)、SIFS-PIFS (允许正处于会话中的结点优先发送。例如收到 RTS 的结点发送一个 CTS，收到数据帧的结点发送一个 ACK)、PIFS-DIFS(SIFS 后没有结点发送，则 PCF 模式的基站可以发送一个信标帧或轮询帧)、DIFS-EIFS(PIFS 后没有基站发送，则任何结点均可竞争信道)、EIFS-(DIFS 后没有发送，则收到坏帧或未知帧的结点可以发送一个错误报告帧)***链路层重传：***一旦结点开始发送一个帧，它就完全地发送该帧(尽管可能发生冲突)。目标节点收到通过CRC检测的帧后，等待SIFS时间发送一个确认帧。若发送结点在给定时间内没有收到确认帧则重传，若多次重传依然没有收到则放弃发送。**CSMA/CA与CSMA/CD的不同：最主要**(CSMA/CD发送时检测冲突，而CSMA/CA发送时不检测冲突)。因此在CSMA/CD 中，节点帧听到信道空闲时立即发送(不怕冲突，冲突后立即停发，损失不大)，而在 CSMA/CA中，节点帧听到信道空闲后随机回退(冲突对无线网络损害很大，应尽量避免)。

**[802.11 终端在子网中移动]** *切换：终端从一个 BSS 移动到另一个 BSS。*发生切换时，终端要关联到新的 AP 上(当终端检测到来自 AP1 的信号逐渐减弱时，便开始扫描新的信标帧。当收到来自 AP2 的信号更强的信标帧时，就先解除 AP1 的关联，然后关联到AP2)。交换机中的转发表也要更新(交换机通过自主学习更新转发表，当收到移动终端(或者AP2代替)发送的帧时，更新其所在的端口信息。若不及时更新，则可能导致丢包)。切换过程中，终端上的应用正常运行(若主机停留在一个IP子网中，则IP地址不会改变。因此网络层及以上感觉不到这个移动，切换过程中产生的延迟和丢包在上层协议看来是正常的)

**[802.11 高级特色] 速率适应：**当主机移动或信噪比变化时，基站和主机动态改变传输速率(物理层调制技术)。例如如果一个节点连续发送两个帧而没有得到确认，则降低发送速率；如果连续 10 个帧均被确认或用于记录上次降速以来时间的定时器超时，则提高发送速率。**功率管理：**节点设置功率管理比特，告知 AP 它将要进入休眠状态。此时节点进入休眠，并在下一个信标帧到达之前醒来(100ms)。在节点休眠期间，AP缓存发往该节点的帧。AP在发送的信标帧中包含一个移动节点列表，列表中的节点表明有帧缓存在 AP中。收到列表且在其中的节点向 AP 请求帧，其他节点重新进入休眠。

## 7. 2 无线网络和网络特征

**[无线链路的特性]** 信号衰减：信号在传播过程中能量逐渐减少(路径损耗，包括穿透和扩散)；干扰：在同一个频段发送信号的电波源将相互干扰，环境中的电磁噪声也能形成干扰；多径传播：由于地面或物体的反射作用，信号沿多条不同长度的路径到达接收端，使得接收方收到的信号变得模糊。由于以上特性，无线链路的传输距离受限，且误码率很高。**因此无线链路协议采用有效的 CRC 错误检测码和链路层 ARQ 协议重传受损的帧**。

**[一些无线网络的问题](多径访问的复杂性)**隐藏结点问题：A-B-C，C 正向 B 发送，A 监听到信道空闲，便向 B 发送。A、C 信号在 B 发生冲突；暴露结点问题：B 准备向 C 发送，B 监听到信道忙(A 在发送)，所以 B 不发送，但实际上可以发送。CSMA 在多跳网络中作用受限：通过载波侦听，发送结点只知道其周围是否有结点在发送，但真正影响此次通信的是接收结点周围是否有结点在发送。隐藏结点：不在发送结点的通信范围内、但在接收结点通信范围内的活跃节点(发送结点听不到但影响接收)暴露结点：在发送结点的通信范围内、但不在接收结点通信范围内的活跃结点(发送结点能听到但不影响接收)

## 7. 1 概述 与有线网络相比，无线网络引入：无线(使用无线链路通信，为物理层和链路层带来问题)、移动(终端改变 IP，为网络层带来问题)

**[无线网络组成]** 无线终端：笔记本智能手机等，可运行网络应用，可静止或移动(无线≠移动)；基站：通常连接到固定网络，在无线终端与固定网络间中继数据包(如802.11AP 、蜂窝塔)基站是无线网络基础设施的关键部分，负责协调与之相关联的多个无线主机的传输；无线链路：连接无线终端和基站。通常需要 MAC 协议协调无线链路的使用。不同的无线链路具有不同的数据传输速率和传输距离。无线链路将位于网络边缘的主机连接到庞大的网络基础设施中，也可在一个网络中连接路由器、交换机和其他网络设备。

**[无线网络的运行模式]** 基础设施模式：无线终端通过基站连接到固定网络(网络基础设施)，所有传统网络服务由固定网络提供。切换：无线终端连接到不同基站的过程。自组织模式：网络中没有基站，主机本身必须提供如路由选择，地址分配，类似于 DNS 的名字转换等服务。节点只能与其通信范围内的节点通信。节点之间相互帮助转发分组，每个节点既是终端又是路由器。

**[无线网络的分类]** 基于基础设施的单跳：主机连接到基站，基站连接到固定网络。如 WIFI(802.11)、 3G 蜂窝网络；无基础设施的单跳：无基站，不连接到固定网络，节点通信不需要中继。如蓝牙；基于基础设施的多跳：主机通过多个节点的转发连接到固定网络。如无线网状网络；无基础设施的多跳：无基站，不连接到固定网络，节点通信需要其他节点转发。如移动自组织网络 MANET、车载自组织网络 VANET

**第七章 无线移动网络**