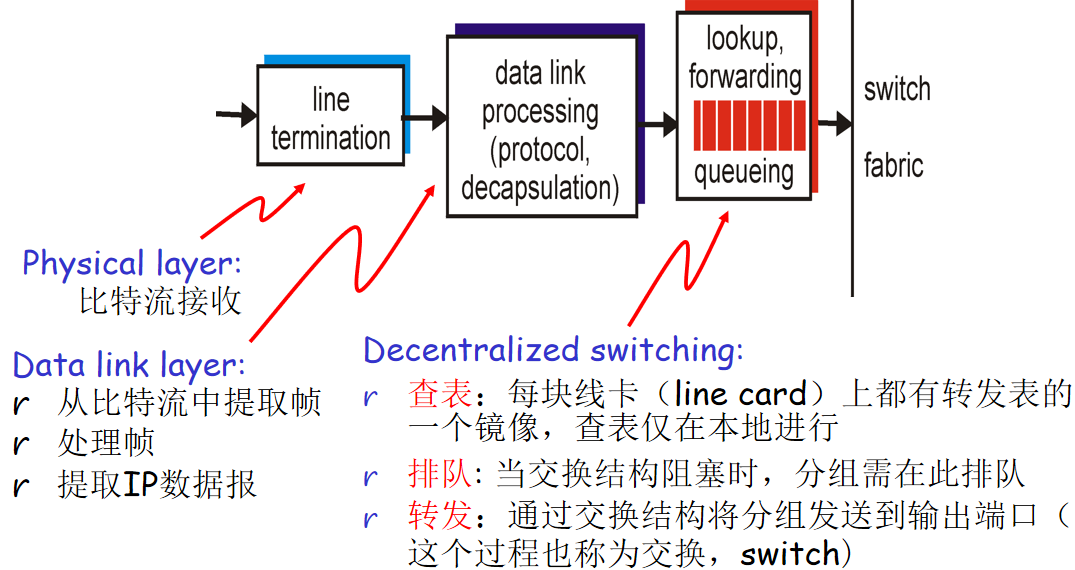
***网络层：数据平面***

**转发和路由选择：**网络层的***转发***功能和***路由选择*功能**区分：**转发**是路由器根据选定的路由，将分组从输入端口转移到输出端口，根据转发表转运分组。涉及分组在*单一*的路由器中从一条入链路到一条出链路的传送。**路由选择**确定去往目的路由器的路由，计算转发表，涉及一个网络的*所有*路由器，它们经路由选择协议共同交互，以决定分组从源到目的地结点所采用的路径。（**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：**执行数据传输的功能，转发、在路由器内部实施分组转运，是路由器本地功能。**控制面：**控制数据传输的功能，选路、确定分组如何去往目的节点，是网络范围的功能。**两种控制面实现方法：传统寻路算法：**在路由器中实现；**软件定义网络SDN：**在服务器中实现。（Per-router control plane & Logically centralized control plane）路由选择设备仅执行转发，远程控制器计算并分发转发表。**路由器和远程控制器**通过交换包含转发表和其他路由选择信息的报文来进行**通信**。**网络服务模型**:定义了分组在发送终端与接收终端之间的传输特性。**可能的网络服务：**（当运输层向网络层*传递一个分组时*）保证交付，具有时延上界的保证交付，有序分组交付，保证最小带宽，安全性。不同架构的网络提供的网络层服务可能不同。同一个网络也可以提供不同的网络层服务。**因特网提供单一的服务，即尽力而为的服务**，但仍尽最大努力解决分组传输时的问题：数据报过大无法通过某个中间网络（措施：对数据报进行分片），数据报可能因寻路错误在网络中循环（措施：设定数据报的最大转发次数），报头出错可能导致误投递（措施：对报头检错）。***网络层的作用***即将报文段从一台发送主机移动到一台接收主机。***分组交换机***是指一台通用分组交换设备，它根据分组首部字段中的值，从输入链路接口到输出链路接口转移分组。分为**(1)*链路层交换机***，基于*链路层帧中字段*值做转发决定；**(2)*路由器***,基于*网络层字段*中的值做转发决定。

***虚电路和数据报网络：***

***网络层***能在*两台主机*之间提供***无连接服务***或***连接服务***。**数据报网络**：提供网络层无连接服务；**虚电路网络**：提供网络层面向连接服务。**与运输层类似：**网络层连接服务以源和目的主机间的握手开始；网络层无连接服务则没有任何握手预备步骤。**差异 网络层：**(1)主机-主机 (2)一个网络不能同时提供无连接服务和面向连接的服务 (3)在网络核心实现；**传输层：**(1)进程-进程 (2)可同时提供无连接服务和面向连接的服务 (3)在网络边缘实现。**虚电路VC：*因特网是一个数据报网络，面向连接服务，类似于电话电路***。网络层连接被称为***虚电路***。虚电路是一条*端到端路径*：(1)传输分组前需建立虚电路，传输结束后应拆除虚电路 (2)所有分组在这条虚电路上传输（分组传输是保序的） (3)如果将路由器资源（带宽、缓存等）预先分配给虚电路(非必须分配带宽)，则虚电路可以提供可预期的网络服务。***建立虚电路的本质***是预先选好源主机到目的主机的路径，此后分组仅沿选好的路径传输（建立虚电路时需要进行路由选择），是否分配资源是可选的。***一条虚电路的组成：***源和目的主机之间的路径（即一系列链路和路由器）；VC号，沿着该路径的每段链路的一个号码（用于区分经过该链路的不同虚电路，仅有*本地意义*）；沿着该路径的每台路由器中的**转发表表项（<进入端口，进入VC号>-<输出端口，输出VC号>**）。**在虚电路上传输分组时，分组只需携带VC号，不需携带目的地址。**属于一条虚电路的分组将在它的***首部携带一个VC号***。因为一条虚电路在*每条链路上可能具有不同的VC号*，每台中间路由器必须*用一个新的VC号替代*每个传输分组的VC号。该*新的VC号从转发表*获得。无论何时跨越一台路由器***创建***一条新的虚电路，***转发表就增加***了一个新表项。无论何时***终止***一条虚电路，沿着该路径每个表中的相应项将被***删除***。***一个分组沿着其路由在每条链路上。不保持相同的VC号的原因：***减少了在分组首部中VC字段的长度；简化了虚电路的建立，否则路由器将不得不交换并处理相当大量的报文以约定一个共同的VC号。***信令报文：***专门用于建立、维护、拆除虚电路的控制报文。***信令协议：***交换信令报文的协议。**数据报网络：**分组携带*目的主机地址*，路由器按目的地址*转发*分组；路由器中的*转发表*记录目的地址->输出链路的映射；转发表被*选路模块*修改，约1～5分钟更新一次。***同一对主机****之间传输的****分组****可能走****不同的路径****，从而可能****无序到达***。虽然在数据报网络中的路由器*不维持连接状态信息*，但它们*在其转发表*中*维持了转发状态信息*。**虚电路和数据报网络的比较：*ATM（虚电路网络）***由电信网发展而来；注重用户体验（用户付费），追求高质量服务；终端无智能或很少智能；复杂工作由网络完成，以保持终端简单。***Internet（数据报网络）***为计算机通信而设计；早期的网络应用均为弹性应用，对网络服务没有严格要求；用户免费使用网络；终端（计算机）具有智能；可将复杂的工作（如差错控制）推到网络边缘，以保持网络核心简单。***数据报网络只提供最小服务的好处：***IP可运行在各种链路之上；增加新服务只涉及终端；任何应用可运行在IP上。**数据报（如RIP）网络**（顺序表，查找复杂度为O(N)）**与虚电路（如MPLS）网络**(索引或Hash表，查找复杂度为O(1))**转发路由表的区别。**

***路由器工作原理：路由器的4个组成部分：***（输入、输出、交换结构都是硬件实现）***(1)输入端口：***性能要求“线速”；***采取其他动作：***检查分组的版本号、检验和以及寿命字段，并且*重写后两个字段*；更新用于网络管理的计数器。

***(2)交换结构：***交换结构将路由器的输入端口与输出端口相连接，这种交换结构完全包含在路由器中， 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；交换速率是输入/输出链路速率的若干倍。**经内存交换：早期：**在CPU的直接控制下完成交换，**数据包拷贝**到系统内存中进行交换，交换速率受限于内存带宽（每个数据包穿过系统总线2次）。**现代：**每个端口使用一个内存接口硬件连接到存储系统，一个控制器硬件在端口之间传输控制消息（不需要CPU参与）。输入端口将一个包放入内存后，接口硬件通过控制器向输出端口发送一个消息，输出端口从内存指定位置读取包，发回响应消息。*转发吞吐量是内存写入或读出速率的一半*并且*不能同时转发两个分组*。**经总线交换：**共享总线；让输入端口为分组预先计划一个交换机内部标签（首部），指示本地输出端口。该分组能由所有输出端口收到，但只有与该标签匹配的端口才能保存该分组。一次只有一个分组能跨越总线。**经互联网交换：**交换结构控制器通过控制交叉点的开、闭，在输入端口与输出端口间建立内部专用电路。能并行转发多个分组。非阻塞（只要没有其他分组当前被转发到该输出端口，转发到输出端口的分组将不会被阻塞）&阻塞。

**（3）*输出端口：***输出端口*存储*从交换结构接收的分组；*组装：*若需要，将交换结构输出的信元组装成分组；*排队：*若输出端口来不及发送，分组在此排队；*调度：*若有多个等待队列，选择一个队头分组发送；并通过执行*必要的链路层*（执行链路层协议，封装）*和物理层功能*（将比特流转换成物理信号）在输入链路上传输这些分组。***路由选择处理器（软件）：***执行*路由选择协议*，*维护路由选择表*以及连接的*链路状态信息*，并为路由器*计算转发表*，执行网络管理功能。一台路由器的输入端口、输出端口和交换结构***共同实现了这种转发功能***，并且***总是用硬件实现。***这些转发功能有时总称为***路由器转发平面。路由器控制平面***通常用***软件实现***并在路由选择处理器（CPU）上执行。**排队：**在输入端口和输出端口处都能够排队。***排队的位置和程度取决于***流量负载、交换结构的相对速率和线路速率。当交换结构不能及时将输入端口的分组转移到输出端口，***输入端口处形成排队。排队带来的问题：***队头阻塞：队头分组阻塞其后分组的转发；***丢包***：当输入队列溢出时，发生丢包；当交换结构速率至少为端口速率的n倍时（n为输入端口数量），可以消除输入端口的排队，但路由器成本提高了。***输入排队 线路前部(HOL)阻塞***，即在一个输入队列中排队的分组必须等待通过交换结构发送（即使输出端口是空闲的），因为它被位于线路前部的另一个分组所阻塞。**输出排队** 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。*当输出队列满时，发生丢包；输出端口排队是不可避免的，***设置多大的输出队列***是一个问题：增大输出队列：可以减少丢包的发生，但会增加内存消耗，并增大分组延迟（延迟太大的分组最终被重传，浪费资源）；输出队列并不是越长越好！***分组丢弃策略：(1)弃尾：***队满时丢弃到达的分组；(2)***主动队列管理：***在队列满之前就开始丢弃分组，如***随机最早检测RED算法：***设计为和TCP拥塞控制机制一起使用；路由器在每个端口上维护输出队列的加权平均长度：***AvgLen = (1- Weight)×AvgLen + Weight×SampleLen；***正常时接纳进入队列；当平均队列长度达到***第一个阈值minth时***，按照丢弃概率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，丢弃间隔越大，p也越大。**N条TCP流流过一条链路时，缓存数量B=RTT•链路容量C/**。**分组调度：(1)先进先出FIFO(FCFS) (2)优先级调度**（设置多个优先级类，每类有自己的优先级和队列，同一类采用FIFO，非抢占式） **(3)轮询调度：**轮流为每个队列队头提供服务 **(4)加权公平排队WFQ：**每个队列分配一个带宽权重（比例），按其选择一定数量（）的分组发送。

***网际协议：因特网中的转发和编址：***因特网编址和转发是***网际协议(IP)的重要组件***。***因特网的网络层有三个主要组件***：**IP协议（数据面）**:编址规则，数据报格式，数据报处理规则；**选路协议（控制面）：**路径选择，RIP/ OSPF/BGP；**ICMP协议（控制面）**：错误报告，路由器信令。**数据报格式：**网络层分组被称为***数据报***。

**IP数据报格式**：***IPv4数据报中的关键字段如下：•版本号***。***•首部长度*（单位：4字节）**IPv4 数据报可包含一些可变数址的选项；大多数IP数据报不包含选项，一般IP数据报有20字节的首部。***•服务类型***。***•数据报长度*（单位：字节）**总长度（*首部加上数据*），该字段长16比特，IP数据报的理论最大长度65535字节。***•标识、标志、片偏移***这三个字段与分片有关。***•寿命(TTL)***用来确保数据报不会永远（如由于长时间的路由选择环路）在网络中循环，每当数据报由一台路由器处理时，该字段的值减1；TTL字段减为0,该数据报丢弃。***•协议***到达其最终目的地才会有用；指示了数据报的数据部分应交给哪个特定的运输层协议；例如，值为6交给TCP,而值为17交给UDP；类似多路分解。***•首部检验和。***是这样计算的：将首部中的每2个字节当作一个数，用反码运算对这些数求和，该和的反码（被称为因特网检验和）存放在检验和字段中；路由器要对每个收到的IP数据报计算其首部检验和，不一致则检测出是个差错，一般会丢弃；注意到在每台路由器上必须重新计算检验和并再次存放到原处，因为TTL字段以及可能的选项字段会改变。***为什么TCP/IP 在运输层与网络层都执行差错检测：***在IP层只对IP首部计算了检验和，TCP/UDP检验和是对整个TCP/UDP报文段进行的；TCP/UDP与IP不一定都必须属于同一个协议栈；原则上TCP能运行在一个不同的协议（如ATM)上，而IP能够携带不一定要传递给TCP/UDP的数据。***•源和目的IP地址•选项•数据（有效载荷）***如果数据报承载一个TCP报文段，则每个（无分片的）数据报共承载了总长40字节的首部(20字节的IP首部加上20字节的TCP首部）以及应用层报文。**链路层帧能承载的最大数据字节数称为MTU**，不同类型的链路可能具有不同的MTU。传输过程中，***较大的IP数据报可以被分片***：将数据报载荷划分为若干较小的数据块，每个数据块封装成一个独立的数据报传输。数据报在传输的过程中可以被多次分片，但仅在目的端系统上而非路由器重组。***分片的报头***取自原始数据报；与分片有关的字段：***标识：***每个分片必须携带与原始数据报相同的标识；***偏移量（单位：8字节）：***指示分片中的数据在原始数据报载荷中的开始位置；***标志位：***MF（more fragments）：最后一个分片的MF=0，其余分片的MF=1；DF（don’t fragment）：DF=1表示不允许对数据报分片。

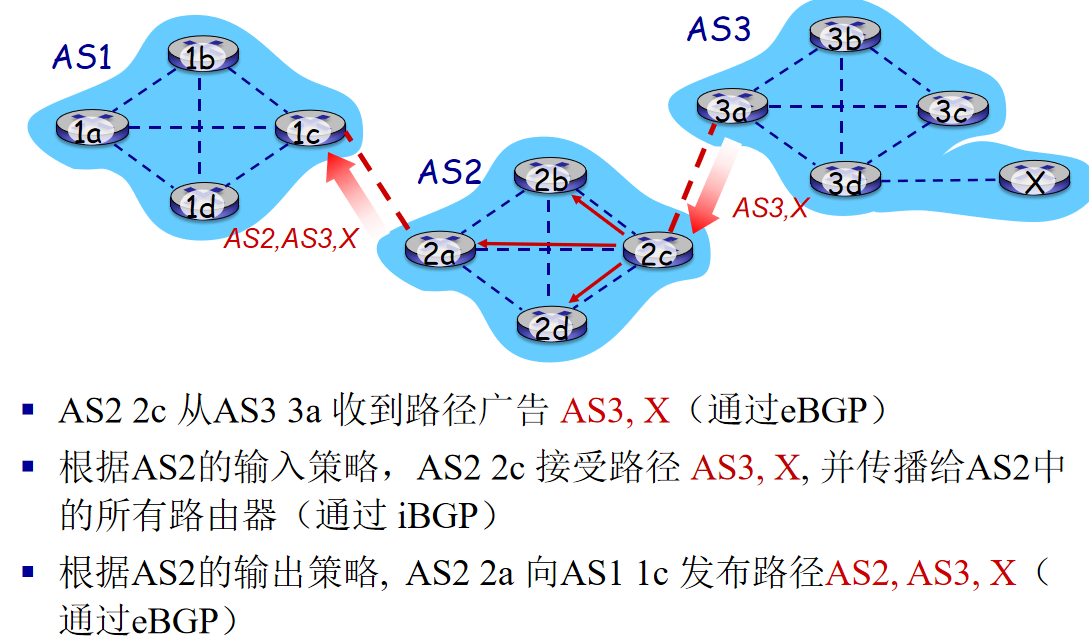
分片报头中的***以下字段需要修改：***总长度，偏移量，MF，TTL，头部检查和。由于偏移量只有13比特，除最后一个分片外，其余分片的数据长度应为8字节的整倍数，假设原始数据报的报头长度为H，则分片的数据长度N应为满足以下条件的最大整数：N ≦ MTU – H，N为8的倍数。***数据报分片的处理过程：***根据报头长度H和输出线路的MTU，确定分片的最大数据长度N；将数据报的载荷划分成长度为N的若干数据块（最后一个数据块可能不足N字节）；将原始报头加到每一个数据块的前面，修改报头中的以下字段：总长度 ＝ H + 数据块长度，最后一个报头的MF位置0，其余报头的MF位置1，**偏移量 ＝ 数据块在原始数据报载荷中的字节开始序号/8**，TTL=TTL-1，计算头部检查和。***重组：***将收到的分片重新组装成原始数据报，在**目的主机**中进行：收集分片：目的主机使用 <源IP地址，标识> 确定属于同一个数据报的分片；利用最后一个分片计算原始数据报长度：原始数据报长度=偏移量×8 + 最后分片总长度；组装：将各分片中的数据块按照其在原始数据报载荷中的偏移量重组。***分片的问题：***分片的开销：降低了路由器的吞吐量；消耗了目的主机的资源，每个重组的数据报需要一个重组缓冲区和一个重组定时器；使得系统更为复杂。针对分片的DoS攻击：攻击者发送一系列奇怪的分片，消耗目的主机的资源。***IPv6取消了路由器分片的功能：***源主机发送探测报文，确定路径上的最小MTU；源主机构造的数据报大小不超过最小MTU；路由器丢弃超大的数据报，并发送错误报告。

**IPv4编址：*接口interface:***主机/路由器与物理链路的边界；路由器有多个接口，主机通常只有一或两个接口（以太网接口、wifi接口）。***IP address:***每个网络接口对应一个IP地址，一个32位的二进制数，通常用点分十进制数（四段）表示。***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，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，从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**：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**：指示本机（仅用作源地址）。**网络号为0、主机号有效的地址**：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：接口地址数：***A类地址：2^24-2 = 16777214；B类地址：2^16-2 = 65534；C类地址：2^8-2 = 254。***地址的个数：***A类地址：2^7-2 = 126；B类地址：2^14-2 = 16382；C类地址：2^21-2 = 2097152。***子网（subnet）：***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。***子网编址：***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；如255.255.255.0；***如何从IP地址中获取子网地址？***将IP地址与子网掩码做“与”运算，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:***子网地址：对应掩码中“1”的部分；主机地址：对应掩码中“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。**路由器的每个端口连接一个子网，不同的端口连接不同的子网，单个路由器一般是一个子网。**路由器是在子网之间转发数据包的设备。子网内部通信不需要通过路由器，子网之间通信必须通过路由器。***网络层转发数据报的两种情形：*直接交付：**同意给子网内，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：**不在同一个子网内，节点将数据包转发给一个路由器去处理。***如何判断使用直接交付还是间接交付？***直接交付：数据包的目的地址与节点的某一端口在同一个子网中；间接交付：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。***间接交付的实现：***节点查找转发表，将数据包发送给下一个路由器。***转发表：***记录目的地址到输出端口的映射；取决于目的地址类型的不同，有***三类转发表项：***(1)目的地址是一个***子网地址***，称地址前缀表项；(2)目的地址是一个***特定的网络接口地址***，称特定主机表项；(3)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。**逐跳转发：**每次只转发到下一跳，下一跳必须是直接可达的（在同一个子网内）。***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）。***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节点）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then把数据包发送到表中指定的下一跳elseif表中包含到N的一个地址前缀表项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，向数据包的源地址发送一条**错误报告消息（ICMP)**。***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；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。***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，表示前21比特是子网地址。***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。***地址聚合：***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：***仍然可以在转发表中给出一条聚合表项；***同时***给出不能被聚合的表项；***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）进行比较；***引入的问题：***地址前缀的长度prefix\_len可以是任意值，Prefix\_len无法从地址本身得到，只能从转发表项中得到，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）。

***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；主机使用***动态主机配置协议DHCP*（应用层）**获取IP地址、子网掩码、缺省路由器、本地DNS服务器等配置信息。***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）。***DHCP目标：***允许主机加入网络时自动获取配置信息。DHCP不只返回分配的IP地址，还包括:第一跳路由器的IP地址（缺省网关、边缘路由器），本地DNS服务器的IP地址，地址掩码。DHCP是一个***客户/服务器模式***的应用协议，***子网中***应有一个DHCP服务器或一个DHCP代理。***DHCP为主机发现新IP地址的四个步骤：*主机广播“DHCP discover”报文**（src:0.0.0.0,68 \dest:255.255.255.255,67 \yiaddr（被分配给该客户的地址）:0.0.0.0（本主机源地址） \transaction ID:654），**寻找子网中的DHCP服务器**；**DHCP服务器用“DHCP offer”报文**（src:223.1.2.5,67 \dest:255.255.255.255,68 \yiaddrr:223.1.2.4 \transaction ID:654 \Lifetime:3600secs）**进行响应，给出推荐的IP地址及租期、其它配置信息**；**主机用“DHCPrequest”报文**（src:0.0.0.0,68 \dest:255.255.255.255,67 \yiaddrr:223.1.2.4 \transaction ID:655 \Lifetime:3600 secs）请求IP地址，主机选择一个DHCP服务器，向其请求IP地址；**DHCP服务器用“DHCP ack”报文**（src:223.1.2.5,67 \dest:255.255.255.255,68 \yiaddrr:223.1.2.4 \transaction ID:655 \Lifetime:3600 secs）**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67，客户使用UDP端口68**。***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之间的TCP连接。

***网络地址转换（NAT）：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）。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***。***NAT实现：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））***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。***16比特端口号: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:***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）。***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。***使用UPnP实现NAT穿越：***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。***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。**NAT主要依据**是数据包转发时主要考虑将其转发到目的网络。**CIDR、DHCP、NAT均可以部分解决IP地址不足的问题，但角度不同。**

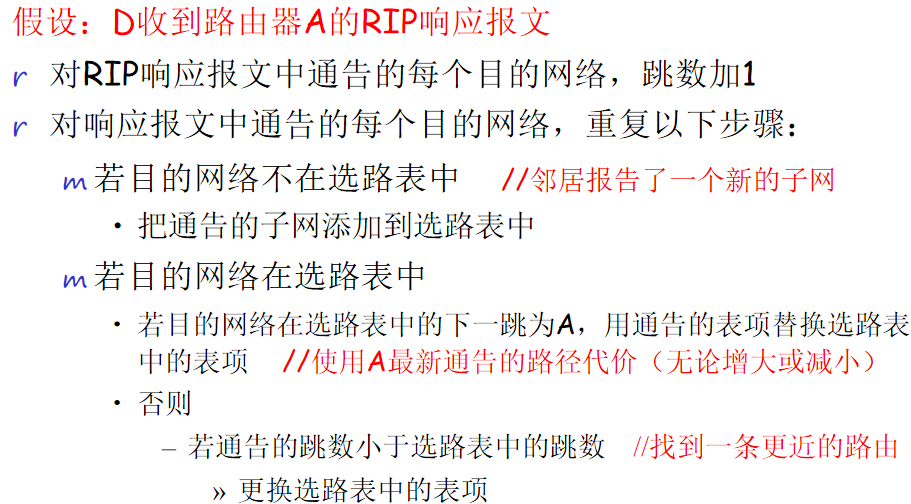
**IPv6：IPv4与IPv6不兼容！*最初的动机：***IPv4地址将很快耗尽；***进一步的动机：***简化头部格式，加快数据报处理和转发；支持服务质量；支持多播；支持移动性；增强安全性。IPv6与IPv4不兼容，但与其它所有因特网协议都兼容。***IPv6地址：***128位，使用冒号十六进制表示，每16位以十六进制的形式写成一组，组之间用冒号分隔，如“8000: 0: 0: 0: 0123: 4567: 89AB: CDEF”；***地址表示的零压缩技术：***可将连续的多组0压缩为一对冒号，如以上地址可表示为：“8000:: 0123: 4567: 89AB: CDEF”。IPv6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（anycast）：一组网络接口中的任意一个（通常是最近的一个）。***IPv6数据报格式：***以一个40字节的基本头开始，后面跟零个或多个扩展头，然后是数据。***PRI（或traffic class）：作用：***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，尚无标准，可以按照用户要求的服务质量等级定义。***流（flow）：***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，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。***IPv6包***与IPv4固定头相比，IPv6的基本头中***去掉了以下一些字段：***选项：IPv6的基本头总是40字节长（但可以通过“下一个首部”）；与分片相关的字段：IPv6路由器不负责分片；头校验：计算校验和太花时间，现在的网络非常可靠，并且链路层和传输层上往往又都有校验和。IPv6基本头中***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。***从IPv4过渡到IPv6：(1)双协议栈方案：***支持IPv6的主机和路由器同时运行IPv4和IPv6；双栈节点同时拥有IPv4和IPv6地址。***(2)建立隧道：***IPv6/IPv4边界路由器将IPv6包(原src/dest)封装到一个IPv4包(src/dest=隧道两端点)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。**本质是封装数据包，IPv6包在IPv4网络中传输时，没有被IPv4路由器检查和处理，报头中的TTL未减1，由隧道端点路由器看来包是直接从另一端路由器直接得到的。**

***网络层：控制平面***

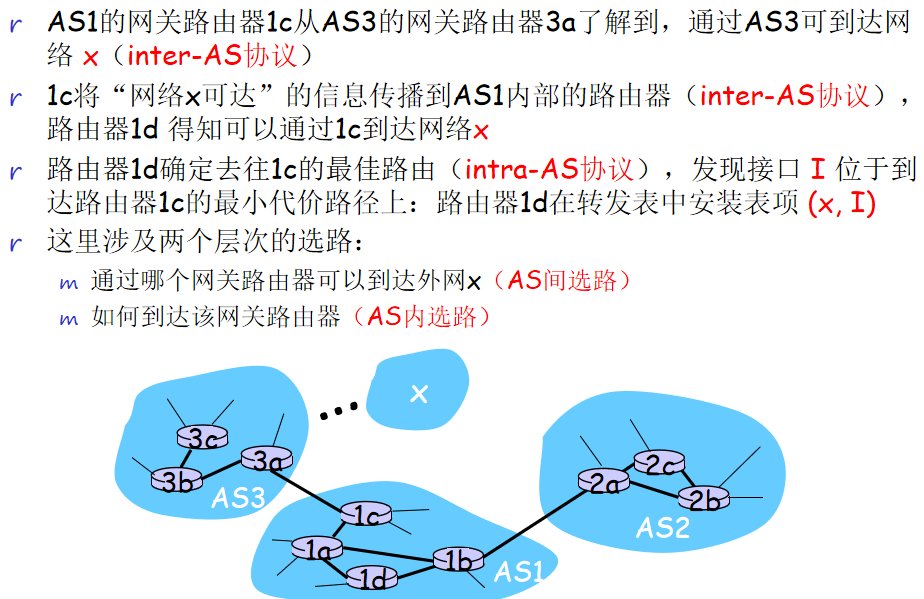
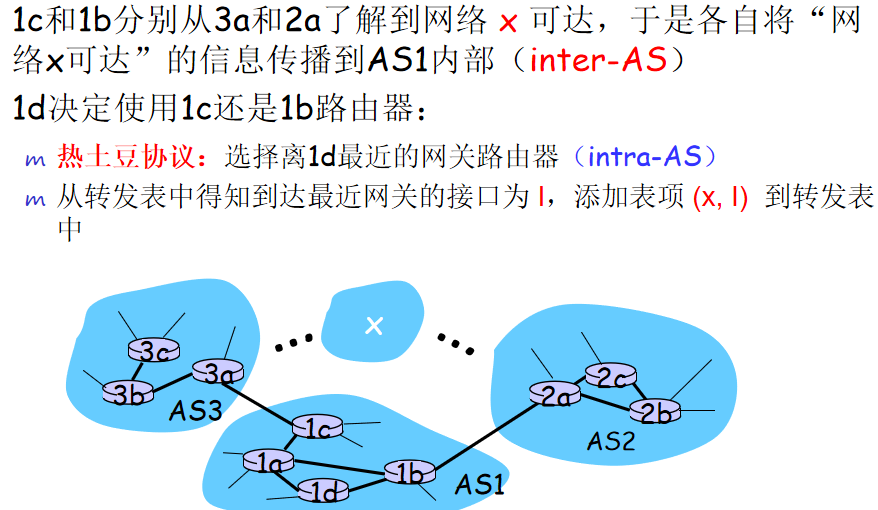
**(1)每路由器控制**(OSPF/BGP) **(2)逻辑集中式控制：**控制器与每台路由器中的一个控制代理CA交互以配置和管理该路由器转发表，CA不能直接相互交互，也不能主动参与计算转发表。***路由选择算法：什么是最佳路径：***应用评价：路径长度、数据速率、分组延迟、通信费用、安全性等；ISP关心：网络吞吐量最大、平均包延迟最小、平均通信费用最低、网络负载均衡、路由稳定、健壮等;路由评价指标通常是矛盾的，需要折衷。***选路算法分类：全局算法 or 分布式算法？全局算法:***所有路由器具有关于拓扑和链路代价的全部信息，集中式计算；***分布式算法:***路由器仅知道邻居节点以及到邻居节点的链路代价，通过与邻居交换信息，进行迭代计算。***静态算法 or 动态算法？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。***负载敏感算法 or 负载迟钝算法？*1.链路状态（LS）选路算法：*全局算法***，其***基本思想***为***：***每个节点利用可靠方法获得全网拓扑信息，抽象成一个s带权拓扑图，计算到各个节点的最短路径。链路状态选路算法包括***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，***构造链路状态（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(***链路状态广播算法***)，***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（Dijkstra）。***可能出现的问题：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(1)***强制费用不依赖拥塞***（2）***确保并非所有路由器都运行LS算法。**2.距离向量（DV）路由选择算法：*距离矢量算法：***利用***Bellman-Ford方程***求解任意两个节点之间的最小代价路径，***分布式（迭代、异步地）***。***算法的思想：***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：Dx(y)←minv{c(x,v)+Dv(y)}。***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)发生了变化或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)有变化时通知其邻居。***链路代价变化：好消息传播快，坏消息传播慢（***路由选择环路和无穷计数问题***）。*毒性逆转：**如果z经过y路由选择到目的地x，z告诉y自己到x的距离是无穷大，只要路径一直没改变，就一直这样告诉y。节点数比较多的环路不能由此检测到。***LS算法和DV算法的比较：链路状态LS：***链路状态信息在全网传播；节点仅传播可靠的信息（LS更健壮）：亲自测量的本地链路代价；节点计算的路由不传播，错误不扩散；***收敛速度：***O(|N||E|)个报文，O(N2)次计算。***距离矢量DV：***距离矢量仅在发生变化时向邻居发送；节点传播的信息可能不正确：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；收敛较慢，还可能出现路由环路、计算至无穷问题。

**层次路由选择：*平面结构的网络不具有扩放性:***路由器数目扩大，路由表规模，信息交换开销；***网络管理员希望对于网络有更多的控制权（管理自治）：***选路算法的选择，隐藏网络内部组织。***自治系统（AS）：***自治系统（AS）是由处于同一个管理域下的网络和路由器组成的集合；每个AS被赋予一个AS编号，由ICANN分配；同一个AS中的路由器运行相同的选路协议（称***Intra-AS*选路协议IGP**（如RIP一般由较低层ISP和企业网使用，OSPF较顶层ISP使用，EGRP））；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；网关路由器之间运行***Inter-AS*选路协议EGP**（如BGP），所有AS之间必须运行相同的Inter-AS选路协议。

***Intra-AS*选路协议IGP：(1)RIP** RIP采用***距离矢量***选路算法。***距离（代价）：***跳数（hop count）。***跳（hop）：***相邻路由器之间的链路为一跳。***路径的跳数：***从源路由器到目的子网（含）经过的子网数量。限定一条路径的最大代价为15跳(不可达设置距离为16)。***RIP通告（RIP响应报文）：***距离向量：路由器到AS内各个子网的最短路径的跳数（估计值）；构造RIP响应报文：**距离向量封装在RIP响应报文中传输**，称为RIP通告，每个报文携带一个目的子网列表（最多包含25个子网），以及到每个目的子网的最短距离；发送RIP响应报文：RIP报文封装在**UDP**报文中发送，使用UDP端口520**（RIP是一个应用层协议！）**，相邻路由器之间大约**每30秒交换一次RIP响应报文**。每台路由器维护一张称为***路由选择表***的RIP表，包括该路由器的距离向量和该路由器的转发表（下一跳路由器）。***RIP链路失效与恢复：***若超过180秒未收到某个邻居的RIP通告，认为该邻居不可达：令通过该邻居的路径失效（距离设为16），发送RIP通告；采用***毒性逆转***解决计数至无穷问题：若选路表中到目的网络 x 的路由是 A 通告的，则向A通告该路由时，到 x 的距离设为16（阻止A使用这条路由）。

**RIP缺点：**更新周期过于频繁；缺少分层，对大规模网络无法适用；收敛性问题（坏消息传得慢）；未实用实际代价表示距离（用跳数表示距离）

**(2)OSPF** OSPF采用***链路状态选路算法***，***链路代价：*由管理员配置**（反映了管理员的选路策略）。***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分组被封装在IP包中传输，协议号为89。路由器周期性地、或在链路状态改变时发送***OSPF链路通告（面对所有其他路由器）。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。***路由器利用链路状态数据库及Dijkstra算法，计算以本路由器为根的最短路径树。***OSPF最重要的优点是支持AS内部的分层选路***；一个OSPF自治系统可以配置成多个***区域（area）：***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。所有区域必须连接到**主干**上；每一个区域都有***区域标识***，主干的区域标识为0；***路由器：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；***内部路由器：***AS内部的非区域边界路由器。***分层的OSPF：两个选路层次：***本地区域，主干；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），通告给其它区域，将收到的其它区域的选路信息（子网及路径代价）通告给本区域的内部路由器；***对于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的其他优点：***安全，可以同时使用多条代价相同的路径（RIP仅一条），对单播和多播路由选择的充分支持。**OSPF封装在IP中发送。OSPF缺点**：采用可靠（或带确认）洪泛（或广播）开销过大；协议复杂（或多种链路状态，多种链路）。

***Inter-AS*选路协议EGP：**

***边界网关协议BGP：***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（可达性；撤销路由或不可达性；路径属性）**。***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***。不同AS的两个边界路由器之间建立的BGP会话，称为***外部BGP（eBGP）会话；***一个AS可能有多个边界路由器，这些边界路由器必须通过半永久TCP连接构成全连通，它们之间的BGP会话称为***内部BGP（iBGP）会话***。**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）。**基于策略的路由：(1)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）**(2)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 **(3)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。 ***Intra-AS和Inter-AS选路协议：***Intra-AS选路协议:用于在AS内部交换选路信息，如OSPF、RIP，使用某个路由测度（代价）选择到目的节点的最优路径；Inter-AS选路协议:用于在不同的AS之间交换选路信息，如BGP，主要依据策略而不是路由测度去寻找可达路径（不追求最佳路径）。

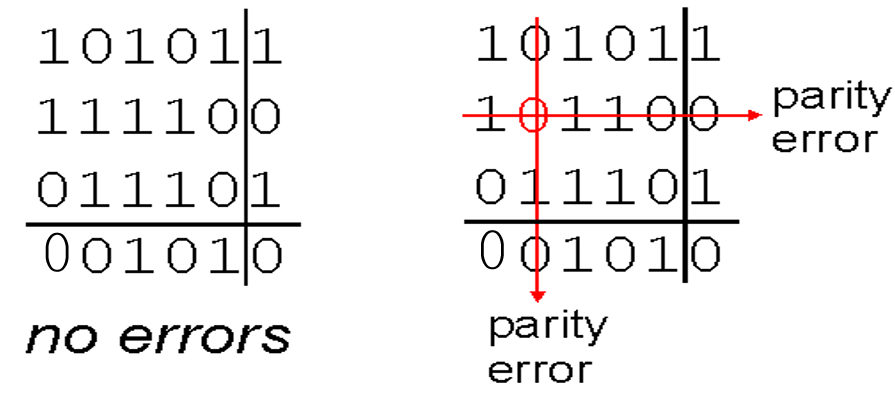
**广播选路：**用***N次单播***实现广播（在源节点上复制分组）：低效：相同的分组在某些链路上可能重复传输，需其它机制支持：源节点需知道所有目的节点的地址。***理想的广播选路：***源节点不需知道其它节点的地址，只需将分组的目的地址设置为广播地址，路由器负责转发到全网（在网络中复制分组）；网络中产生的分组拷贝最少。***在网络中复制分组：洪泛（flooding）：***节点收到广播分组后，向所有邻居节点（分组到来的链路除外）发送该分组的拷贝；缺点：在有环的网络中，广播分组在网络中无休止地循环，浪费资源。***受控洪泛：***目标：每个路由器仅转发它之前未转发过的广播分组；***两种方法：(1)***节点记录之前转发过的分组ID，不重复转发分组（OSPF使用此方法：源地址+分组ID）；***(2)反向路径转发:***利用节点内部的单播转发表，仅转发从本节点->源节点的最短路径的反向路径上到来的广播分组（该方法使用最多）。***反向路径转发RPF基本思想：***当广播分组到达路由器时，路由器检查分组的源地址与输入端口；用分组的源地址查找单播路由表，找到去往该源地址的输出端口；若分组的输入端口与去往该地址的输出端口相同，则扩散该分组，否则丢弃分组；优点：算法合理、易于实现且开销不大。***生成树方法：***使用生成树转发广播分组：路由器知道自己的哪几个端口在生成树上；当从一个端口收到广播分组后，只在属于生成树的其它端口上转发该分组；***基于生成树的广播不会产生冗余的分组拷贝。生成树的构造：基于核心的方法：***选择一个节点作为核心（也称汇聚点）；其它节点向核心发送单播的加入报文：路由器利用单播转发表向核心转发加入报文时，记录报文的输入端口及输出端口，这些端口就是位于生成树上的端口；当加入报文到达生成树上的一个节点时，报文经过的路径被添加到生成树上。

**多播：*如何标识多播通信的接收者：***因特网为这组接收者分配一个标识（多播组标识），使用D类地址作为多播组标识；***如何设置多播分组的接收者：***将分组的目的地址设置为其接收者的多播组地址；发往该地址的数据报要发给组中所有成员。***如何将接收者的IP地址与多播组地址关联起来：***接收者的IP地址与多播组地址无关，接收者可以在任何时候加入或离开一个组，***多播组管理协议（IGMP）***允许主机向本地路由器申请加入或离开一个组；***如何将多播分组交付给每一个接收者：*多播选路协议**协调多播路由器建立到达所有接收者的路径树。**建立多播树的两种方法：基于源的树**（源节点建立一棵到多播组所有成员的最短路径树，源节点S和组G的每一种组合<S,G>构成一棵树）**，组共享树**（每个多播组使用一棵树，树根为该多播组的核心）。**多播分组在单播网络中的传输：隧道技术的应用（把多播分组封装在单播分组中传播）。**

**ICMP两类：询问（回声请求/响应、地址掩码请求/响应、路由器请求、路由器发现）：**用来请求一些信息，通常采用请求-响应模式；**错误报告：（源抑制、超时、目的不可达、重定向**（不丢弃数据包）**、参数错误）**发现错误的节点向源节点报告错误信息，不需响应。**ICMP报文封装在IP包中传输！ICMP报文格式：**type：报文类型，共定义了15种；code：对某类报文作进一步的区分；Checksum：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。**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；**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，IP报头的TTL设为1；第一跳路由器对TTL减1，发现TTL变为0，向源主机发送一个TTL expired报文（源主机从IP报头中发现路由器IP地址）；Traceroute记录第一跳路由器的IP地址，然后向目的主机发送第二个Echo Request报文，IP报头的TTL设为2 ；若收到第二跳路由器的TTL expired报文，记录第二跳路由器的IP地址；接着发送一个TTL为3的Echo Request报文；该过程不断重复，直至收到目的主机的Echo Response报文。

***链路层：链路、接人网和局域网：***

***网络层和链路层的关系：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：将数据报从一个结点传输到相邻的下一个结点，***如：源主机→源路由器，路由器→下一跳路由器，目的路由器→目的主机。***链路层概述：节点：***运行链路层协议的任何设备。***链路：***连接相邻节点的通信信道。***帧：***链路层分组称为***链路层帧***。发送主机与Wifi接入点之间的Wifi链路，接入点和链路层交换机之间的以太网链路，链路层交换机与路由器之间的链路，两台路由器之间的链路，交换机和服务器之间的以太网链路。**链路层提供的服务：*成帧（基本服务）****发送：讲数据报封装到帧中（加上帧头和帧尾）；接收：*从原始比特流中提取出完整的帧；***链路接入****媒体访问控制MAC*在点对点链路上无论何时链路空闲发送方都能发送帧，在广播链路上协调各个节点的发送行为；***差错检测（基本服务）***硬件实现；***差错纠正（有些提供）***:检测并纠正传输错误（不使用重传）；***可靠交付（部分协议提供）*确认、重传**确保接收节点正确收到每一个帧（停-等、GBN、SR），低误码率链路（如光纤、某些双绞线）上很少使用，高误码率链路（如无线链路）应当使用；***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***（提供收/发转换）***&全双工***。**链路层在何处实现：*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（网卡）***中实现。线卡/网络适配器连接物理媒体，还实现物理层的功能。***链路层由硬件和软件实现***：网卡中的控制器芯片：组帧、链路接入、检错、可靠交付、流量控制等；主机上的链路层软件：与网络层接口，激活控制器硬件、响应控制器中断，处理差错条件和将数据报向上传递给网络层。***网络适配器之间的通信：发送侧:***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；***接收侧：***提取帧，检测传输错误；执行可靠传输和流量控制；解封装数据报，交给上层协议。

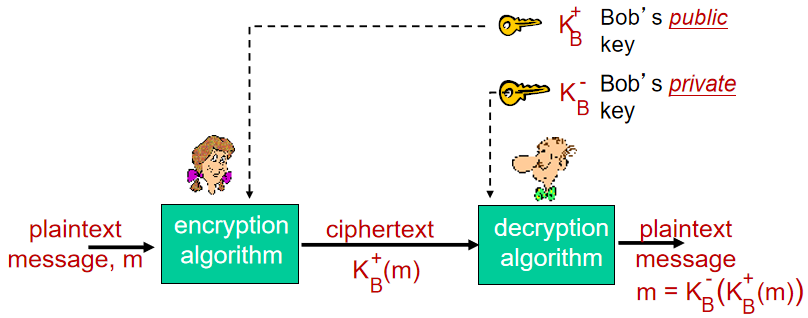
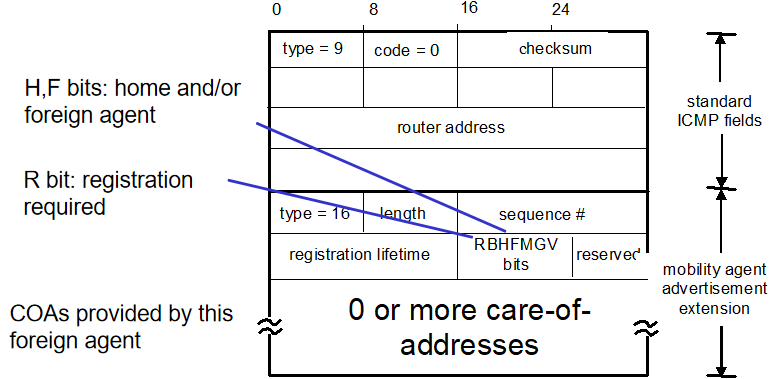
***差错检测和纠正技术：*** 编码集的海明距离：编码集中任意两个有效码字的海明距离的最小值；为检测出所有d比特错误，编码集的海明距离至少应为d+1；为纠正所有d比特错误，编码集的海明距离至少应为2d+1

***EDC（差错检测和纠正比特）*** **~~100%检测出错！~~奇偶校验：*单比特奇偶校验：***在偶(奇)校验中，发送方只需包含一个附加比特，使**d+1**比特中**1**的个数总是偶(奇)数。可*检测奇数个*比特错误，检错率为50%。***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。***前向纠错(FEC)：***是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的次数，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。***因特网检验和(checksum)：***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，接收方将所有接收数据（包括检验和）的和取反码全1；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。***为什么运输层使用检验和而链路层使用CRC：***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；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。**循环冗余校验（CRC）：*CRC***是一种***多项式编码***，将一个位串看成是某个一元多项式的系数。***信息多项式D(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 \* D(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（可检测小于r+1比特的突发错误）。**编码方法：*R(x) = x^r \* D(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，问是否有错？解答：1001001÷1001的余式为001（不为0），有传输错误。***

***多路访问链路和协议：链路的两种类型：(1)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；

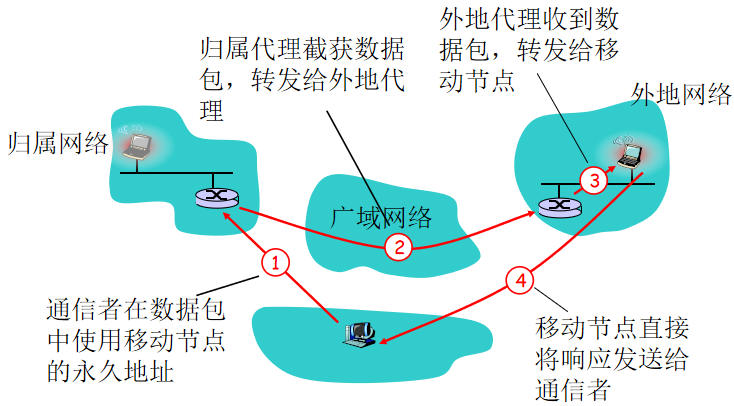
***(2)广播链路：***连接了许多节点的单一共享链路，任何一个节点发送的数据可被链路上的其它节点接收到。***冲突：***在广播链路上，若两个或多个节点同时发送，发送的信号会发生干扰，导致接收失败。***多址接入协议（多路访问协议）：***规定节点共享信道（谁可以发送）的方法，多址接入协议也称***媒体接入控制（MAC）协议***。***理想的多址接入协议：***在速率为R bps的广播信道上，**1.** 当只有一个节点发送时，它应能以速率R发送（信道利用率高）**2.** 当有M个节点发送时，每个节点应能以R/M的平均速率发送（公平性好、信道利用率高）**3.** 协议是完全分布式的: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）。**ALOHA和CSMA 协议具备第一个特性，但不具备第二个特性。*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。**1.信道划分协议：*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（R/N***带宽***），若节点不发送，其子频带空闲。***TDM和FDM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）：***将每个比特时间进一步划分为m个微时隙chip，每个节点被分配一个惟一的m比特码序列chip code；每个节点用它唯一的编码来对它发送的数据编码。**发送方编码：**发送“1”=发送chip code；发送“0”=发送chip code的反码。**信号叠加：**多个节点发送的信号在信道中线性相加。**接收方解码：**用发送方的chip code与信道中收到的混合信号计算内积，恢复出原数据。**前提条件：**任意两个chip code必须是相互正交的；CDMA允许所有节点同时使用整个信道！**2.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；涉及碰撞的节点反复地重发帧，直到该帧无碰撞通过为止。每次碰撞不必立刻重发，而是每个节点独立选择随机时延，时延后重发***。随机接入MAC协议的例子：发送前不监听信道：***ALOHA家族；***发送前监听信道：***CSMA家族。***(1)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：**单个活跃节点可以信道速率连续发送，分布式：节点自行决定什么时候发送，简单。**缺点：**发生冲突的时隙被浪费了，由于概率重传，有些时隙被闲置，需要时钟同步。***时隙多路访问协议的效率：***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***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；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；最大效率: 找到令Np(1-p)N-1最大的概率p\*；令N趋向于无穷，得到：**最大效率 = 1/e = 0.37**。***(2)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，立即以概率P重传，以概率（1-P）等待一个帧时后再决定。（***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。*（3）载波侦听多址接入（CSMA）：CSMA/CD适用于有线网络（802.3），CSMA/CA广泛应用于无线网络（802.11）。*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，它就停止传输，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 ：***若在发送的过程中检测到冲突，怎么办？继续发送余下的部分（浪费带宽）or停止发送余下的部分。CSMA/CD的**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**；立即启动冲突解决的过程。***以太网采用CSMA/CD协议：*1.**NIC从网络层接收数据报，构造以太帧；**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。***希望时间间隔：***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，对于以太网，一个结点等待的实际时间是K\*512比特时间，(512bit是最小以太帧长度)，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够立刻插入一次成功的传输。***CSMA/CD效率：*Tprop =** 以太网中任意两个节点之间传播延迟的最大值；**Ttrans =** 最长帧的传输时间；**efficiency = 1 / (1 + 5\*Tprop/Ttrans)。在以下情况下，以太网的效率趋近于1：**Tprop 趋近于0，或Ttrans趋向于无穷。**结论：**应控制以太网的规模。**3.轮流MAC协议：*中心节点轮询（蓝牙）：***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；**缺点(1)**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；**(2)**是单点失效(主节点)。***令牌传递协议（FDDI、IBM令牌环、令牌总线）：***没有主结点，一个称为**令牌(token)**的小的特殊帧在结点之间以预定次序进行传递；当一个结点收到令牌时，仅当它有帧要发送时，它才持有这个令牌；否则，它立即向下一个结点转发该令牌；一个结点收到令牌时，若它确实有帧要传输，它发送最大数目的帧数，然后把令牌转发给下一个结点。**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来；令牌传递延迟；令牌单点失效。***MAC协议比较：*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：用于电缆因特网接入的链路层协议** 用到三类协议。

***交换局域网：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（**没有MAC地址**）；***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（一百公里以上），规模和容量可任意扩大。**链路层寻址和ARP：**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***等。***MAC地址***长6个字节；**由IEEE负责分配**，每块适配器的地址是全球唯一的。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，即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（无层次结构）**，而且不论适配器到哪里用都不会变化；**IP地址**是**有层次的**，移动时发生改变。***目的MAC地址***有三种类型：**单播地址：**适配器的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（FF-FF-FF-FF-FF-FF）**。网络适配器仅将发送给本节点的帧交给主机：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）。***发送节点如何获得接收节点的物理地址？地址解析问题：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。**ARP（将IP地址解析为MAC地址）与DNS（将主机名解析为IP地址）区别：**DNS为在因特网中任何地方的主机解析主机名，ARP只为在**同一个子网**上的主机和路由器接口解析IP地址。***ARP报文格式：*硬件类型：**硬件接口类型；对于以太网，该值为“1”。**协议类型：**高层协议地址类型。对于IP地址，该值为080016。**操作：**ARP请求为1，ARP响应为2。**在以太网上，ARP报文封装在以太帧中传输**。***地址解析的过程：***A想知道B的MAC地址：**1.** A构造一个ARP请求，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响应）。**4.** B构造一个ARP响应，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。**ARP是即插即用的**，就是一个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，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**，即不完全符合简单的分层协议栈。**改进ARP的措施(1)ARP缓存：ARP表：**每台主机和路由器在其内存中具有一个ARP表，包含IP地址到MAC地址的映射关系和TTL（从表中删除每个映射的时间）。**(2)主动学习：**从ARP请求中获取地址绑定信息；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。***数据报到达子网之外：***数据报从A（子网1）经过R（路由器）到达B（子网2）：A知道下一跳地址为111.111.111.110（R-1），R知道B从其端口R-2直接可达；A创建IP数据报，src IP＝A, dest IP＝B；A利用ARP获得下一跳111.111.111.110对应的MAC地址（R-1）（而非获得B的MAC）；A创建链路层帧，封装IP数据包，src MAC =A, dest 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，src MAC=R-2, dest MAC = B，发送；B的网卡接收帧，取出IP数据报，交给网络层。

****以太网：**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：**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；**集线器（hub，物理层）：**一个物理层中继器，相当于共享电缆，**不分割冲突域**，因为集线器不具有选路功能，本质上是将其受到的信息广播出去。***星型拓扑：交换式以太网：*交换机：**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），端口有物理地址进行选路，每个端口是一个冲突域，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。***以太网帧结构：*前同步码(8字节)：**7个10101010字节，后跟一个10101011字节，用于在发送方和接收方之间建立时钟同步。**目的地址（6字节）+源地址（6字节）**。**Type（2字节）：**指出Data所属的高层协议（如IP、ARP等），每个协议有一个编号，用于多路分解（和网络层数据报中的协议字段、运输层报文段的端口字段类似）。**Data：**46～1500字节，不足46字节填充至46字节；以太网的最大传输单元(MTU)是1500字节，这意味若如果IP数据报超过了1500字节，则主机必须将该数据报分片；如果IP数据报小于46字节，数据报必须被填充到46字节，当采用填充时，传递到网络层的数据包括IP数据报和填充部分，网络层使用IP数据报首部中的长度字段来去除填充部分。**CRC（4字节，循环冗余检测）：**对dest addr.、src addr.、type和data四个字段计算得到的CRC码。**所有的以太网技术都向网络层提供无连接服务**（没有握手）&**不可靠服务**（接收方网卡不发送确认，接收方网卡丢弃CRC错误的帧，依靠上层协议（TCP或应用）进行错误恢复）。***为什么有最小帧长的要求？***为确保节点在发送结束前检测到冲突，帧的发送时间必须足够长：节点检测冲突需要时间，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；**为什么最小帧长为64字节（不包括前导码）：**根据早期以太网的最大直径（2500米）和数据速率（10Mbps）计算得到。***共享式以太网和交换式以太网：*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：**交换机的每个端口为一个冲突域，多对端口可以同时通信，网络的集合带宽=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***是接收入链路层帧并将它们转发到出链路。交换机自身对子网中的主机和路由器是***透明的。即插即用****（不需要配置）****，自主学习***。**双工**（任何交换机接口能够同时发送和接收）**。交换机输出接口设有缓存**。***过滤***是决定一个帧应该转发到某个接口还是应当将其丢弃的***交换机功能***。***转发***是决定一个帧应该被导向哪个接口，并把该帧移动到那些接口的***交换机功能***。交换机的过滤和转发**借助于*交换机表(switch table)*完成：**包含某局域网上某些主机和路由器的但不必是全部的表项，交换机表中的**一个表项包含：**一个MAC地址，通向该MAC地址的交换机接口，表项放置在表中的时间。**帧转发的描述类似于数据转发，重要差异**是交换机转发分组**基于MAC地址**而不是基于IP地址&&交换机表与路由器的转发表的构造方式有很大差别。***交换机收到帧的处理过程：*用帧的目的地址查找转发表（转发决策）：**若目的地址所在端口 = 帧的进入端口，丢弃帧；若目的地址所在端口 ≠ 帧的进入端口，转发帧；若目的地址不在转发表中，扩散帧。**用帧的源地址查找转发表（更新转发表）：**若找到地址，将对应表项的生存期设为最大值；若没有找到该地址，添加源地址和进入端口到转发表，设置表项的生存期为最大值。***交换机和路由器比较：***交换机工作于链路层，根据MAC地址存储转发帧；路由器工作于网络层，根据IP地址存储转发数据报。交换机不能连接异构链路（即MAC协议不同的网络），因为交换机只是按原样转发帧；路由器可以连接异构链路，因为路由器需重新封装链路层帧。交换机不能阻断广播帧的传播，路由器可以。***冲突域：***共享同一条广播链路的主机集合；任何一个主机发送的帧（各种帧），可被冲突域中的其它主机接收到。**广播域：**广播帧能够到达的主机集合。广播风暴：广播帧在网络中大量传播，消耗大量资源。**三层交换机&专业路由器**：路由器可分隔二层网络，但转发速度慢、成本高；**三层交换机**：具有部分路由功能、又有二层转发速度的交换机；专为加快大型局域网内部的数据交换而设计；但在安全、协议支持等方面不如专业路由器。**路由器转发IP包的过程：**用目的IP地址查找转发表，获得下一跳IP地址及端口；利用ARP获得下一跳MAC地址；用下一跳MAC地址构造链路层帧，发送。**三层交换机转发IP包的过程：**将以上第1、第2步的结果缓存到本地三层转发表中；用目的IP地址查找三层转发表：1）若命中，直接用下一跳MAC地址构造链路层帧，发送2）若未命中，执行以上第1、2、3步。三层交换机转发速度快的原因：一次选路，多次转发。三层交换机：通常用在机构网络的核心层，连接不同的子网或虚拟局域网；专业路由器：连接机构网络与外网。***链路虚拟化：网络作为链路层：*在物理网络上增加一个逻辑层次（IP层）**，在逻辑层上统一编址、统一包格式，互联在一起的网络看起来像一个网络；**用网关连接不同的物理网络：**在逻辑层上选路到下一个网关；将IP包封装在本地网络帧中，发送到下一个网关。***Cerf & Kahn’s Internetwork Architecture：*两级编址:** IP网络，物理网络；**IP层提供统一的网络视图：**地址，包格式；**底层可以是任意的物理网络，**物理网络对于IP层是不可见的，对于IP来说物理网络只是一条虚拟链路而已！

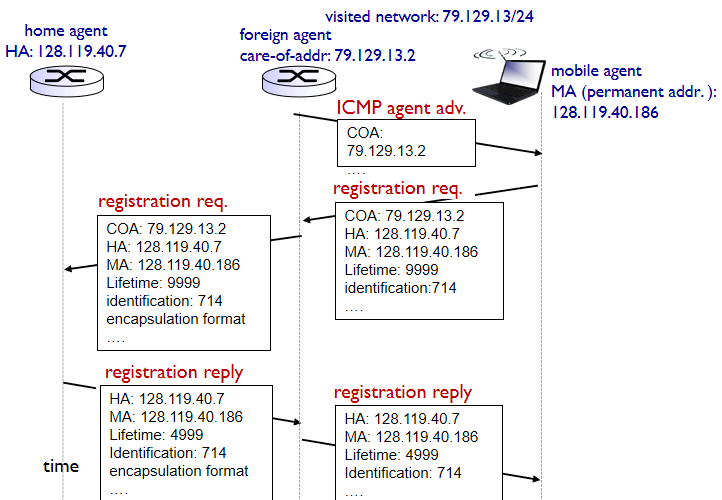
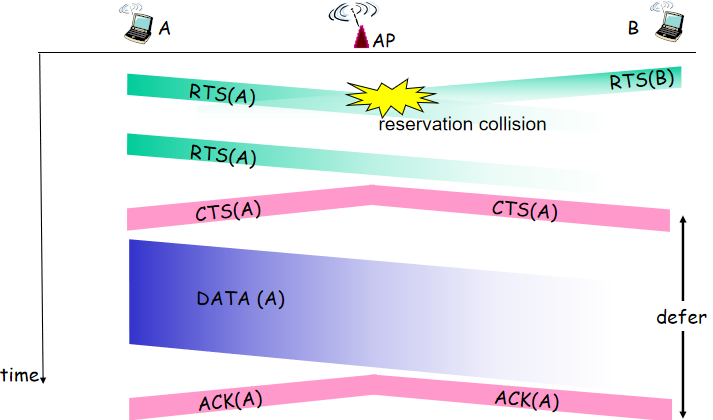
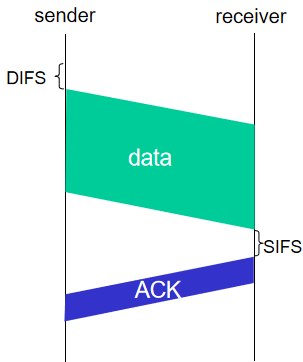
***无线和移动网络：***

****与有线固定网络相比，无线移动上网*增加了以下两方面的问题：*无线：**使用无线链路通信，给物理层和数据链路层带来很多问题；**移动：**终端改变网络接入点，给网络层带来很大问题。***概述：无线网络的组成：(1)*无线终端**，运行网络应用，可能静止或移动。**(2)基站，**通常连接到固定网络，在无线终端和固定网络之间中继数据包；如**802.11AP接入点，蜂窝塔**；通常负责协调与之关联的多个无线主机的传输。**(3)无线链路：**连接无线终端和基站，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速率和传输距离。***无线网络的运行模式：*(1)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），所有传统的网络服务由固定网络提供；***切换：***无线终端接入到不同基站的过程。**(2)自组织模式：**网络中没有基站，节点只能与其通信范围内的节点通信，节点相互帮助转发分组，每个节点既是终端又是路由器。***无线网络的分类：*1. 单跳有基础设施：**主机连接到基站，基站连接到固定网络(如WiFi802.11，蜂窝3G,4GLTE)；**2. 多跳有基础设施：**主机通过多个无线节点的中继才能到达固定网络（如无线网状网络，无线传感网）**3. 单跳无基础设施：**无基站，不连接到固定网络，节点间通信不需要中继（如蓝牙网络）**4. 多跳无基础设施：**无基站，不连接到固定网络，节点间通信需要通过其它节点中继（如移动自组织网，车载网）。

***无线链路和网络特征：无线链路的特性*(1)信号衰减：**信号在传播过程中能量逐渐减少**（路径损耗）(2)干扰：**受到其它信号源的干扰 **(3)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。**以上特性导致无线链路的传输距离受限、误码率很高。信噪比SNR**：所收到的信号和噪声强度的相对测量。**比特差错率BER**：接收方收到的有错传输比特的概率。**(1)对于给定的调制方案，SNER越高，BER越低。**(2)对于给定的SNR，具有较高比特传输率的调制技术将具有较高BER。(3)物理层调制技术的动态选择能用于适配对信道条件的调制技术。 ***无线网络的特性：*(a) 隐藏终端问题：**C正在向B发送；A(在B的传播范围之外)监听到信道空闲，A向B发送；A和C的信号在B冲突；**(b) 暴露节点问题：**B准备向C发送，B监听到信道忙（A在发送），B不发送，其实B可以发送（A和B的信号不会在C冲突）**CSMA/CD不适用！CSMA在多跳无线网络中作用受限：**通过载波侦听，发送节点只能知道其周围是否有节点在发送；但真正影响此次通信的是接收节点周围是否有节点在发送。**隐藏节点：**不在发送节点的通信范围内、但在接收节点通信范围内的活跃节点（发送节点听不到，但影响接收）；**暴露节点：**在发送节点的通信范围内、但不在接收节点通信范围内的活跃节点（发送节点能听到，但不影响接收）。

***WIFI：802.11无线LAN：均使用CSMA/CA*** 作为MAC协议；都***支持基站模式和自组织模式***；物理层不同。**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，一个无线接入点AP（一个中央**基站**）；每个**无线接口**（终端及AP）**均有一个全局唯一的MAC地址**。AP与路由器相连的**有线端口没有MAC地址**（AP对于路由器是透明的）。***802.11信道与关联：***802.11将通信频段划分成若干信道（**信道有部分相互重叠**），每个BSS**分配一个信道**（85MHz频段内定义了11个部分重叠的信道，当且仅当两个信道由四个或更多信道隔开时它们才无重叠，信道1、6、11是唯一的3个非重叠信道的集合）**。**管理员安装AP时，为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）**。**主机必须与一个AP关联：**扫描11个信道，监听各个AP发送的信标帧（包含AP的SSID和MAC地址），选择一个AP进行关联（可能需要身份鉴别），使用DHCP获得AP所在子网中的一个IP地址。***802.11主动/被动扫描：*被动扫描：**主机扫描信道和监听AP发送的信标帧，主机选择一个AP发送关联请求帧，AP向主机发送关联响应帧；**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，AP发送关联响应帧。

**802.11 MAC协议：带碰撞避免的CSMA/CA。*以太网和802.11都使用载波侦听随机接入，但这两种MAC协议有重要的区别：***802.11使用**碰撞避免**而非碰撞检测；由于无线信道相对较高的误比特率，802.11(不同于以太网）使用**链路层确认**（接收方收到帧后需发送确认帧，以便让发送方知道是否发送成功）**／重传(ARQ)方案**。**发送前监听信道，发送过程中不检测冲突的原因？**发送过程中检测冲突很困难（接收信号的强度远小于发送信号的强度）；不能检测出所有的冲突（隐藏节点）。

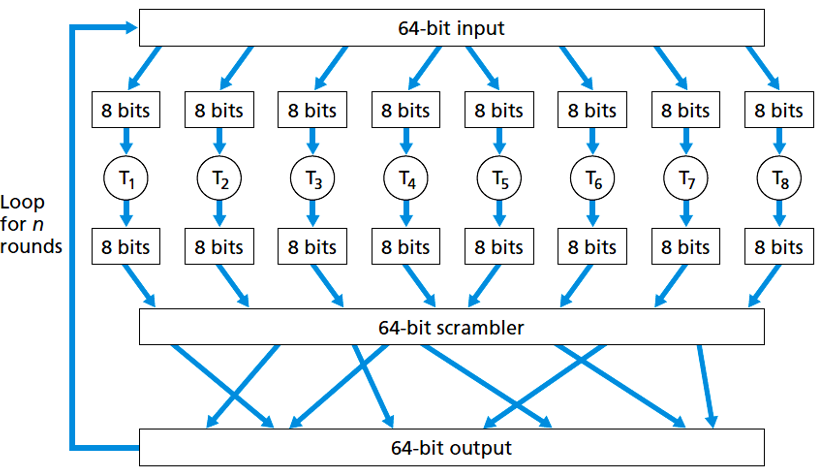
*********(1)使用信道预约机制的CSMA/CA（可选）：***•A向AP发送一个RTS帧：帧中给出随后要发送的数据帧及ACK帧需要的总时间。•AP收到后回复一个CTS帧：帧中给出同样的时间。•A收到CTS帧后：发送数据帧。•AP收到帧后：发送一个ACK帧进行确认。•收到RTS帧（A附近）及CTS帧（AP附近）的节点均沉默指定的时间，让出信道让A和AP完成发送。•若A和B同时发送RTS帧，产生冲突，不成功的发送方随机等待一段时间后重试。**帧间距机制（使得802.11允许DCF和PCF在一个单元内共存）**：**SIFS：**允许正处于会话中的节点优先发送，如收到RTS的节点发送一个CTS，收到数据帧的节点允许发送一个ACK帧；**PIFS：**如果在SIFS后没有节点发送，PCF模式的基站可以发送一个信标帧或一个轮询帧；**DIFS：**如果PIFS后没有基站发送，任何节点可以竞争信道；**EIFS：**如果以上间隔都没有发送，收到坏帧或未知帧的节点可以发送一个错误报告。帧。当要发送的数据较少时，一般不使用信道预约机制。***（2）不使用信道预约机制的CSMA/CA（所有实现必须支持）：***•如果某站点最初监听到信道空闲，它将在分布式帧间间隔DIFS的短时间段后发送该帧•否则该站点选择一个随即回退值并且在侦听信道空闲时递减该值。当侦听信道忙时，计数值保持不变。•当计数值减为0时该站点发送整个数据帧并等待确认•如果收到了确认，发送站点知道它的帧已被目的站正确接收了。如果该站点要发送另一帧，它将从**第二步**开始CSMA/CA协议。如果未收到确认，发送站点将重新进入第二部中的回退阶段，并从一个更大范围选取随机值。如果有k个节点等待发送，它们随机选取的回退值确定了它们的发送顺序。

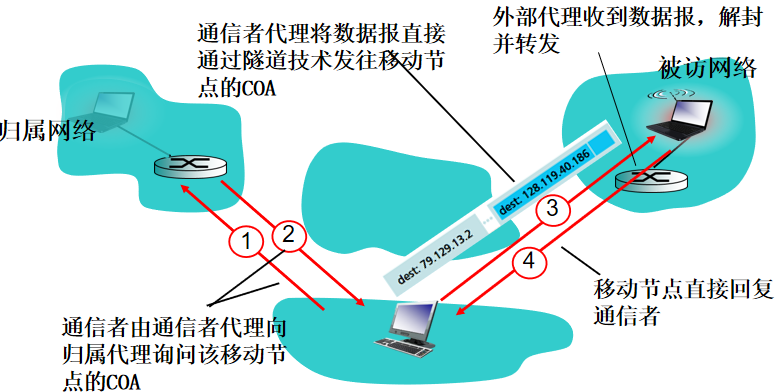
***CSMA/CA与CSMA/CD的不同：*最根本的不同：CSMA/CD在发送过程中检测冲突，不使用确认机制；CSMA/CA在发送过程中不检测冲突，使用确认机制**；由此带来的**协议处理方面的不同：**在CSMA/CD中，节点侦听到信道空闲时立即发送；在CSMA/CA中，节点侦听到信道空闲后要随机回退；**原因：**冲突对无线网络损害很大，要尽可能避免。**处理隐藏终端：请求发送控制帧RTS&允许发送控制帧CTS** 802.11MAC协议包含**预约方案**，帮助在出现隐藏终端的情况下避免碰撞。当AP收到RTS帧后，它广播一个CTS帧作为响应。该CTS帧有两个目的：给发送方明确的发送许可，也指示其他站点在预约期内不要发送。**RTS-CTS能完全避免碰撞，且只在传输长数据帧时使用。**

**802.11帧格式：*（1）四个地址字段：*地址1：**帧的目的无线站点MAC地址；**地址 2：**帧的源无线站点MAC地址；**地址 3：**连接AP的路由器接口的MAC地址；**地址 4：**只在自组织模式中使用。***802.11帧寻址举例：***无线终端H1向路由器R1发送帧，它的AP已知：H1构造一个address 1 = AP MAC，address 2 = H1 MAC，address 3 = R1 MAC，将该帧发给AP；AP将这个802.11帧转换为802.3帧（有线），后者的dest addr = R1 MAC，source addr = H1 MAC。**AP连接路由器的有线端口没有MAC地址！AP仅对无线终端可见，对于固定网络上的设备是不可见的**。***（2）*RTS, CTS, ACK, 数据帧，类型和子类型**（区分关联）**，To和From**（定义不同地址字段含义）**，WEP**（定义是否加密）***。（3）*有效载荷和CRC。**

**802.11终端在子网内移动：*切换：***终端从一个BSS移动到另一个BSS。**(1)发生切换时，终端要关联到新的AP上：**当H1检测到来自AP1的信号逐渐减弱时，开始扫描新的信标帧；当H1收到来自AP2、信号更强的信标帧时，先解除与AP1的关联，然后关联到AP2。**(2)发生切换时，交换机（连接AP1和AP2）中的转发表也需要更新。**交换机***通过自主学习更新转发表：***交换机收到H1发送的帧时，更新H1所在的端口；若转发表未及时更新，可能产生丢包。**主机停留在同一个IP子网中，因而IP地址保持不变，**切换过程中，终端上的应用正常运行：由于IP地址没变，网络层及以上层次感觉不到这个移动，切换过程中产生的延迟及丢包，在上层协议看来是正常的。(1)互联设备：交换机，则可以保持IP地址和TCP连接；(2)互联设备：路由器，新IP且TCP连接中断。**802.11先进功能：*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）；***功率管理：*节点设置功率管理比特，告知AP它将进入休眠状态：**D络WPAN）**，ZigBee**（低功率、低数据率、低工作周期）

**802.11*终端在IP子网间移动：***终端进入到一个新的子网后，必须分配该子网上的一个地址（DHCP），并使用新的地址通信，不能保留其IP地址。但终端改变IP地址后，终端上正在运行的应用将中断：通信的对方不知道终端的新地址，无法与其通信；即使对方获知了终端的新地址，应用必须重新建立连接，因为通信的端点（套接字）变了。**归属网络：**移动节点的永久“居所”。**永久地址：**移动节点在归属网络中的地址，总是可以使用这个地址与移动节点通信，即使移动也保持不变。**归属代理：**当移动节点在外地时，为移动节点执行移动管理功能的实体。**外部网络：**移动节点当前所在的网络。**外部代理：**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（外部地址）：**移动节点在外地网络上的地址。**通信者：**希望与移动节点通信的节点。***移动节点注册：***移动节点进入外地网络后，向外地代理注册；移动节点通过外地代理向归属代理注册；**最终结果：**外地代理知道移动节点在本地网络上；归属代理知道移动节点的转交地址，记录到地址绑定表中。***(1)间接选路到移动节点：间接选路：***

**移动节点到外部代理的协议**：需注册和注销；**外部代理到归属代理的注册协议**：需注册，不需显式注销；**归属代理数据报封装协议**：将通信者的原始数据报封装在一个目的地址为COA的数据报内并转发之；**外部代理拆封协议**：从封装好的数据报中取出通信者的原始数据报并转发之。移动节点**使用两个地址：永久地址**（映射到归属代理的MAC地址）**&转交地址**。节点移动及变换外地网络**对通信者都是透明的：正在进行的通信可以保持！*三角选路问题：*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很低效。

***(2)******直接选路到移动节点：***克服了三角选路的问题；对通信者不透明：通信者需要**移动用户定位协议**知道移动节点的转交地址，通信者（包括固定节点）需要增加对移动通信的支持（首次发现移动节点的外部网络代理标识为锚外部代理，当节点到达新外部网络时，新外部代理向锚外部代理提供移动节点的新COA）。**直接路由vs间接路由：(1)透明性：**间接(是，只需知道永久地址)&直接(不，需要隧道) **(2)效率：**间接(低效，三角路由)&直接(高效) **(3)会话移动性：**间接(简单，只需更新归属代理COA)&直接(复杂)

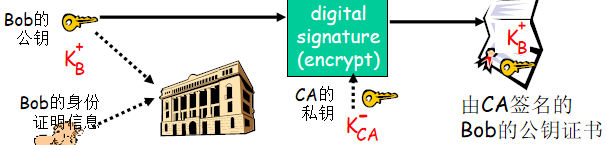
***移动 IP（RFC 3344）：*特性：**归属代理，外地代理，永久地址，转交地址，移动节点注册；**标准化了三个部分：**代理发现，移动节点注册，数据报间接选路。***代理发现：*代理通告：**愿意充当归属代理或外地代理的路由器定期在网络上发送代理通告，宣布自己的存在及IP地址，type=9；**代理请求：**一个想知道代理的移动节点不必等待接收代理通告，就能广播一个代理请求报文，type=10。***移动节点注册：*** 移动节点向外地代理发送一个注册请求，给出自已的永久地址、转交地址、归属代理地址以及认证信息等；外地代理记录相关信息，向归属代理转发注册请求；归属代理处理注册请求，若认证通过，将移动节点的永久地址及转交地址保存在绑定表中，发回一个注册响应；外地代理收到有效的注册响应后，将移动节点记录在自己的转发表中，向移动节点转发注册响应；当移动节点回到归属网络时，要向归属代理注销。***数据报间接选路：***数据包首先被归属代理得到；归属代理查找地址绑定表，获得移动节点当前的转交地址；归属代理将数据包发送到转交地址；外地代理将数据包转发给移动节点。

***无线和移动对上层协议的影响：无线链路带来的问题：***误码率、丢包率、延迟增大。***节点移动带来的问题：***丢包、延迟增大。**逻辑上，没什么影响：**为上层协议提供的仍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：**丢包率高，传输延迟增大；TCP将丢包（长延迟也当作丢包）解释为拥塞，不必要地减小拥塞窗口，导致应用吞吐率很低；延迟实时传输的修复；无线链接的有限带宽。

***计算机网络中的安全：***

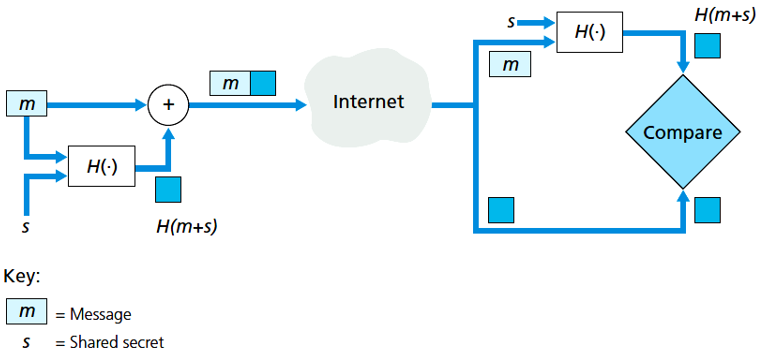
***什么是网络安全：***

***安全通信特性***：**• 机密性：**报文内容的机密性，通信活动的机密性；**• 报文完整性：**报文来自真实的源，且传输过程中未被修改。**• 端点鉴别：**发送者和接收者能够证实对方的身份；**• 运行安全性：**网络不受攻击，网络服务可用。***安全攻击的类型：被动攻击：***试图从系统中获取信息，但不对系统产生影响；**两种类型：*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；**四种类型：*伪装：***一个实体假冒另一个实体；***重放：***从网络中被动地获取一个数据单元，经过一段时间后重新发送到网络中；***报文修改：***改变报文的部分内容、推迟发送报文或改变报文的发送顺序；***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：***通过报文交换确信一个实体的身份，以防假冒；***数据完整性：***用于保护数据单元或数据单元流的完整性，以防报文修改；***数字签名：***附加在一个数据单元后面的数据，用来证明数据单元的起源及完整性，以防伪造及抵赖；***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。

***密码学的原则：明文（plaintext）：***欲加密的原始数据；***密文（ciphertext）：***明文经加密算法作用后的输出；***密钥（key）：***加密和解密时需要使用的参数；***密码分析：***破译密文；***密码学：***设计密码和破译密码。***按照加密密钥与解密密钥是否相同，加密算法分为：*对称加密算法**：加密密钥与解密密钥相同，存在密钥传递问题；**非对称**加密算法：加密密钥与解密密钥不同。发送者使用加密密钥，接收者使用解密密钥，不存在密钥传递问题：加密密钥公开，解密密钥私有。***按照明文被处理的方式，加密算法分为：*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。

**对称密钥密码体制：(密钥管理和分发困难)*替换密码：***用密文字母替换明文字母，字母位置不变；eg：凯撒密码(字母表后第k个字母)，单码代替密码（词频统计攻击，密钥：字母映射表），多码代替密码（密钥=多个凯撒密码密钥+其次序循环模式）；***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（对密码分析者最不利）；***已知明文攻击：***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。***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**加密方法也使用替换和换位两种基本手段，但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，密码的安全性由算法的复杂性和密钥的长度保证。一个加密算法被称为是***计算安全***的，如果由该算法产生的**密文满足以下两个条件之一：**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。**一个安全的加密系统必须能抵御选择明文攻击。全表块密码难以实现。**

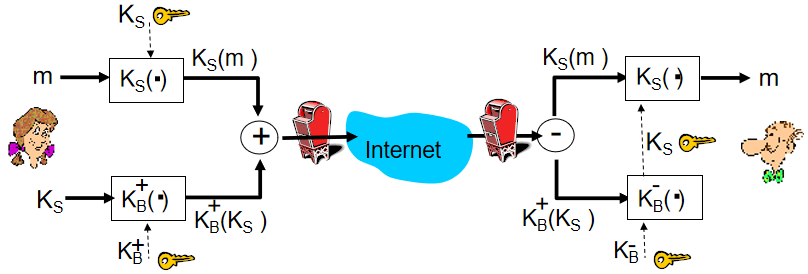
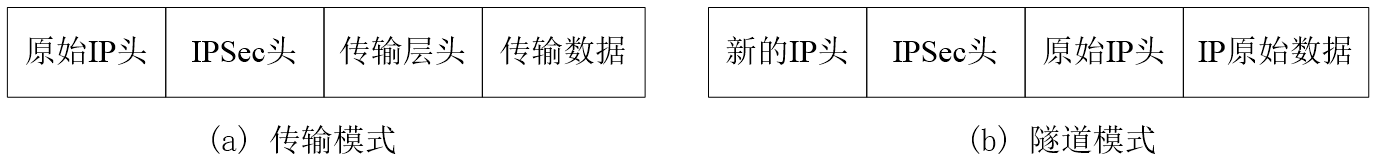
**块密码1.*DES：*** 64比特明文块输入，64比特密文块输出；初始换位；每一轮迭代（共16轮）执行相同的替换和换位操作，但使用不同的密钥；再次换位（与初始换位相反）。主密钥56比特，每一轮迭代使用的子密钥（48比特）由主密钥产生；一种对称加密算法，加密和解密使用相同的函数，两者的不同只是子密钥的次序刚好相反；缺点：密钥长度不够长，迭代次数不够多。***3DES***使用两个密钥进行**三轮DES计算：**第一轮加密模式，使用密钥K1对明文进行变换；第二轮解码模式，使用密钥K2对第一轮的输出进行变换；第三轮加密模式，用密钥K1对第二轮的输出进行变换，输出密文。为什么使用两个密钥而不是三个密钥？112比特的密钥已经足够长。为什么不使用两重DES（EE模式）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），即有C1＝EK2 (EK1 (P1 ) )，令X＝EK1(P1)＝DK2(C1)，攻击者分别计算EK1(P1)和DK2(C1)，并寻找使它们相等的K1和K2，则穷尽整个密钥空间只需2^56的攻击量而不是2^112。（中途攻击）为什么是EDE而不是EEE？为了与单次DES兼容。3DES用户解密单次DES用户加密的数据，只需令K1＝K2就行了。**2.*AES高级加密标准：*** 128比特明文块输入，128比特密文块输出；密钥长度可以是128、192或256比特。**若每个明文块被独立加密，相同的明文块生成相同的密文块，易被重放攻击。*密码块链接CBC：***发送方生成一个随机的初始向量c(0)，用明文发送给接收者；每一个明文块加密前，先与前一个密文块进行异或，然后再加密;c(1)=Ks(m(1)⊕c(0)),c(i)=Ks(m(i)⊕c(i-1));后果：需要在协议中提供机制分发初始向量（IV）C0。**公开密钥加密：*公开密钥算法的使用：***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，解密密钥妥善保管。**公开密钥：**加密密钥，由发送者使用；**私有密钥：**解密密钥，由接收者使用。***要求：***KB-(KB+(m)) = m，给定公钥KB+，不可能计算出私钥KB-。***公开密钥算法应满足的条件：***计算上易生成一对加密密钥和解密密钥；已知加密密钥，易从明文计算出密文；已知解密密钥，易从密文计算出明文，从加密密钥推出解密密钥是不可能的，从加密密钥和密文计算出原始明文是不可能的。***公开密钥算法两个问题：***(1)入侵者知道该公钥和加密算法，可发起选择明文攻击 (2)加密密钥公开，任何人都可能向Bob发送一个已加密的报文，在单一共享密钥情况下，发送方知道共享秘密密钥的事实就已经向接收方隐含地证实了发送方的身份，然而在公钥体制中，任何人都可向Bob发送使用Bob的公开可用密钥加密的报文，这就需用数字签名把发送方和报文绑定。***RSA算法：(1)生成密钥：***选择两个大(越大破解难度越大)素数ｐ和ｑ（典型值为p\*q是1024数量级）；计算n＝p\*q和z＝(p-1)\*(q-1)；选择一个与z互质的数，令其为e<n(用于加密)；找到一个d(用于解密)使满足e\*d＝1 (mod z)；**公开密钥为(e, n)，私有密钥为 (d, n)**。***(2)*加密：**将明文看成是一个比特串，将其划分成若干数据块M（看作整数），且有0≤M＜n；对每个M，计算C＝M^e (mod n)，c即为M的密文；**解密：**对每个密文块C，计算M＝C^d (mod n)，M即要求的明文。***RSA特性：*先用公钥再用私钥和先用私钥再用公钥结果相同。*RSA的特点***：优点：安全性好，建立在难以对大数提取因子的基础上，此目前数学家尚未解决的难题；使用方便：无需传递密钥。缺点：计算开销大，速度慢。***RSA的应用：***加密少量数据，如用于鉴别、数字签名或发送一次性**会话密钥Ks（加密数据本身**，对称密码中的共享密钥，发送方发送c=mod n，接收方得到Ks）等。

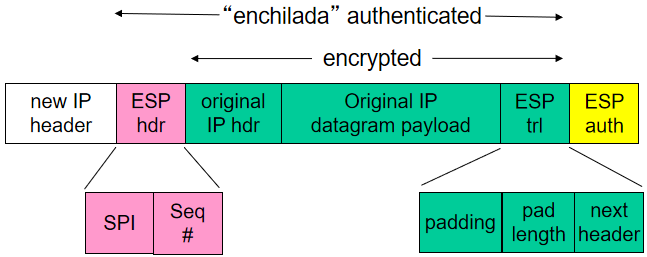
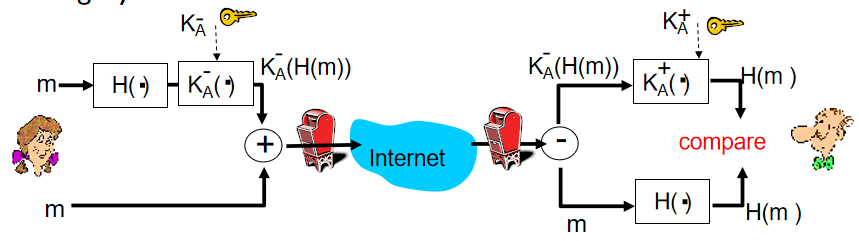
***报文完整性和数字签名：报文完整性（报文鉴别码MAC）：***验证报文是否**可信**。**报文鉴别涉及*两个方面：起源鉴别：***报文是否来自声称的源；***完整性检查：***报文是否被修改过。***报文鉴别方法一：对整个报文加密：***缺点：混淆了机密性和报文鉴别两个概念，有时我们只想知道报文是否可信，而报文本身并不需要保密；加密整个报文会带来不必要的计算开销。***将报文鉴别与数据机密性分开：*设想：**发送者用明文发送报文，并在报文后附上一个允许接收者鉴别报文真伪的标签；**鉴别报文的标签必须满足两个条件：**能够验证报文的完整性（是否被修改），不能被伪造。***数据完整性和报文摘要：报文摘要（数字指纹）：***将一个散列函数作用到一个任意长的报文m上，生成一个固定长度的散列值H(m)即报文摘要。***使用报文摘要验证报文的完整性：***发送者对发送的报文计算一个报文摘要，作为标签和报文一起发给接收者；接收者对收到的报文也计算一个报文摘要，和收到的标签进行比较。***如何保证报文摘要不被修改？报文鉴别方法二：***发送方计算报文摘要，然后用与接收方共享的密钥加密报文摘要，形成报文鉴别标签（***报文鉴别码***）；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：需使用加密算法。**密码散列函数：*为什么要开发一个不需要加密算法的报文鉴别技术？***加密软件运行慢；加密硬件代价是不能忽略的；加密算法可能受专利保护（如RSA），使用代价很高；加密算法可能受到出口控制（如DES），有些组织可能无法得到加密算法。***报文鉴别方法三：使用密码散列函数H生成报文鉴别码MAC：***(1)MAC=H(m+s)，s是**鉴别密钥**(共享比特串) (2)发送方生成扩展报文(m,H(m+s)) (3)接收方收到扩展报文(m,h)，计算出H(m+s)，比较H(m+s)和h是否相等。***散列函数H应满足的特性：***能作用于任意长度的数据块，生成固定长度的输出；对任意给定的数据块x，易计算H(x)；对于任意给定的值h，要找x 满足H(x)=h，在计算上是不可能的（单向性）；对于任意x,y且y≠x 使得H(y)=H(x)，在计算上是不可能的（**checksum不满足**）。***散列函数标准：*** MD5（散列码长度为128bit），SHA-1（散列码长度为160bit），HMAC。

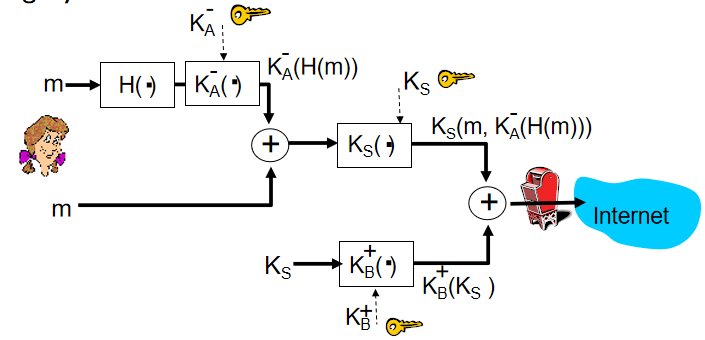
**数字签名：*必须满足三个条件：***接收方通过文档中的数字签名能够鉴别发送方的身份（起源鉴别），发送方过后不能否认发送过签名的文档（防抵赖），接收方不可能伪造被签名文档的内容。MAC无法胜任这项工作，因为有两个人拥有它。***用私钥加密报文摘要：***发送方先计算报文摘要，然后用自己的私钥加密报文摘要形成数字签名，数字签名附加在报文后面一起发送；接收方拷贝一份数字签名，妥善保存，以备将来需要时使用；接收方用发送方的公钥得到原始的报文摘要，对收到的报文计算摘要，如果两者相符，表明报文是真实的。***数字签名与MAC进行比较：***数字签名和MAC都以一个报文（或一个文档）开始，数字签名是一种“技术含量更高的”技术，因为它需要具有认证中心支撑的公钥基础设施(PKI)。***公钥认证****（****数字签名应用）：***为使公钥密码体系有实际应用，每个实体必须能够确认它得到的公钥确实来自声称的实体；解决方案是引入证书机制：使用证书来证明某个主体拥有某个公钥，证书由一个可信任的第三方机构认证权威CA颁布；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：Bob向CA提供身份证明；CA验证了Bob的身份后使用CA的私钥创建证书，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：获取Bob的证书；使用CA的公钥验证Bob的证书，得到Bob的公钥。常用证书标准：***X.509证书***。***X.509定义了三种鉴别程序：***单向鉴别：涉及一个用户到另一个用户的一次报文传输（接收方鉴别发送方）；双向鉴别：通信双方相互鉴别；三向鉴别：通信双方相互鉴别，**并提供报文同步机制**。***谁可以运行CA？世界上有几个CA？***使用一个CA签发全世界所有的证书?流量压力，单点失效。由一个组织运行多个CA?密钥泄露，信任问题。***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拥有自己的私钥，负责为一部分用户签发证书；用户自己决定使用哪一个CA。***信任锚与信任链 信*任锚：**信任的起点，系统中的所有实体都以根CA的公钥作为它们的信任锚，必须通过安全的物理途径获取；**信任链：**从叶结点到根CA的证书序列；**根CA的选择：**有很多根CA，每个根CA都有自己的分级结构，所有根CA间可以进行交叉认证；用户自行决定信任哪个根CA；许多根CA的公钥被预装在浏览器上，这些根CA由浏览器厂商认证并嵌入到软件中，随软件一起发布。***证书的撤销：***每个证书都有有效期，过期后证书自动失效；CA也可以显式地撤销证书，这要求CA定期地发布证书撤销列表CRL给出已经撤销的证书序列号；每个用户在使用一个证书前都要去获取CRL，检查该证书是否在CRL中。***证书存放在哪里？***使用**DNS**作为***证书目录，***标准为DNSSEC；使用专门的目录服务器存放证书，标准为LDAP；证书撤销列表通常与证书存放在一起，由目录服务器负责将CRL中列出的证书清除掉。

***端点鉴别：端点鉴别：***一个实体经过计算机网络向另一个实体证明其身份的过程。鉴别应当在报文和数据交换的基础上，作为某***鉴别协议***的一部分独立完成。**鉴别协议**通常在两个通信实体运行其他协议（例如，可靠数据传输协议、路由选择信息交换协议或电子邮件协议）**之前**运行。**鉴别协议ap1.0：**直接发送一个报文说自己是本人。**鉴别协议ap2.0：**有一个总是用于通信的周知网络地址（IP），利用网络数据报的源地址鉴别。Trudy用Alice的IP地址创建一个数据包（IP地址欺骗）。**鉴别协议ap3.0：**Alice向Bob发送口令证明自己，口令是鉴别者和被鉴别者之间共享的秘密。Trudy监听到Alice发送的明文口令，过后发送给Bob。**鉴别协议ap3.1：**Alice将口令加密，发送给Bob。Trudy截获数据包，直接使用加密口令，不用破译也能用，发送给Bob（重放攻击）**鉴别协议ap4.0：*目标：***避免重放攻击。***失败的情况是因为***Bob不能区分Alice的初始鉴别报文和后来入侵者回放的Alice的初始鉴别报文所致；Bob无法判断Alice是否还活跃（即当前是否还在连接的另一端），或他接收到的报文是否就是前面鉴别Alice时录制的回放。***不重数(nonce)***：在一个协议的生存期中只使用一次的数，也就是说，一旦某协议使用了一个不重数，就永远不会再使用那个数字了。***协议ap4.0使用一个不重数：*1)** Alice向Bob发送报文“我是Alice''；**2)** Bob选择一个不重数R，然后把这个值发送给Alice；**3)** Alice使用她与Bob共享的对称秘密密钥KA-B来加密这个不重数，然后把加密的不重数KA-B(R)发回给Bob。这个不重数用于确定Alice是活跃的；**4)** Bob解密接收到的报文，如果解密得到的不重数等于他发送给Alice的那个不重数，则可鉴别Alice的身份。**缺点：**需要一个共享的对称密钥。**鉴别协议ap5.0：采用公开密钥算法加密不重数：**“I am Alice”→R→KA-(R)→“send me your public key”→KA+。**Bob计算：KA+(KA-(R)) = R**，只有Alice拥有这个私钥，因而一定是Alice！

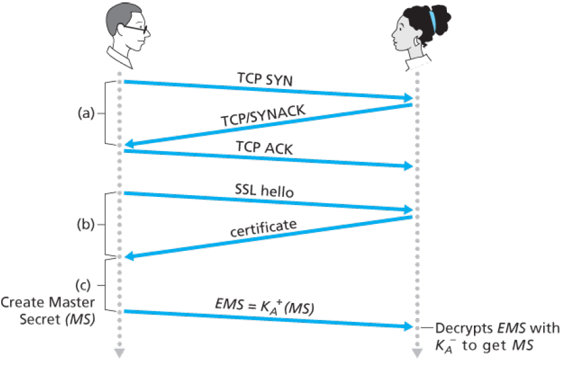
***安全电子邮件：为什么要在因特网的多个层次上提供安全性功能呢？仅在网络层提供安全性功能并加以实施还不足够吗？***(1) 尽管通过加密数据报中所有数据（即运输层报文段），和鉴别所有数据报的源IP地址，在网络层能够提供“地毯式覆盖“安全性，但是却并不能提供用户级的安全性。例如商业站点不能依赖IP层安全性鉴别购买其商品的顾客。(2) 在协议栈的较高层上部署新的因特网服务（包括安全性服务）较为容易，而在网络层上广泛部署安全性较难。

***电子邮件安全***包括**真实性**和**机密性**，**发送方鉴别**，**报文完整性**，**接收方鉴别**。***提供机密性的方式：*对称密钥算法？**仅有Alice和Bob具有该密钥的副本，这使得分发对称密钥非常困难。**公开密钥密码？**效率相对低下，尤其对于长报文更是如此。**为**克服效率问题，利用***会话密钥*：机密性**

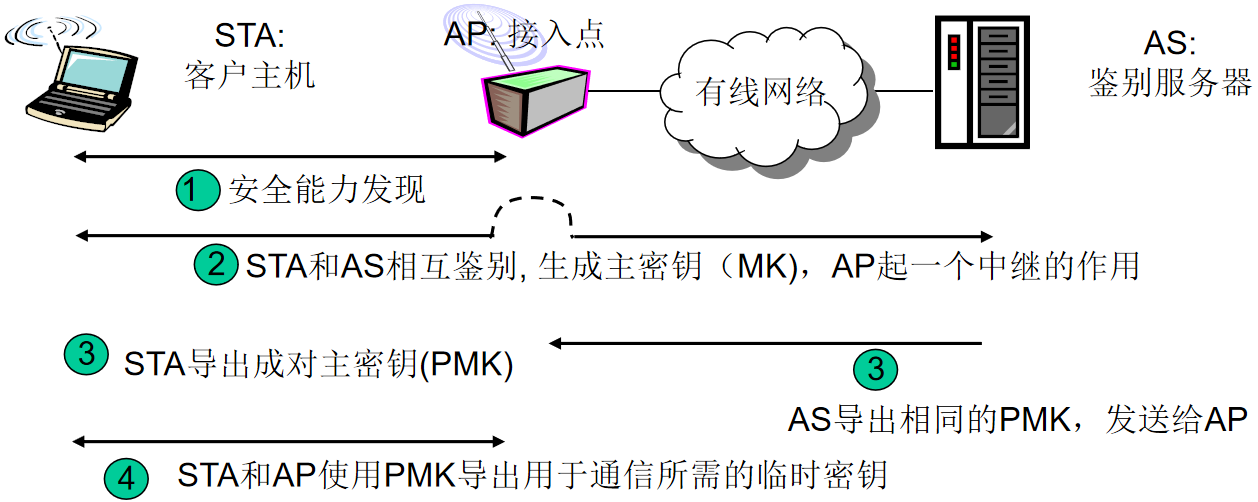
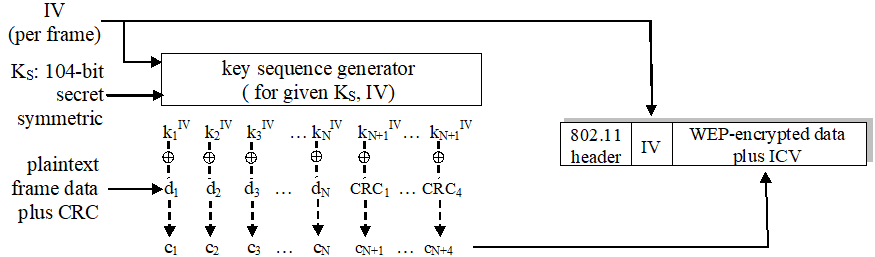
*******只关心发送方鉴别和报文完整性、无机密：***使用数字签名和报文摘要。

***设计一个提供机密性、发送方鉴别和报文完整性的电子邮件系统***Alice两次使用了公开密钥密码：一次用她的私钥，另一次用Bob的公钥。同样，Bob也两次使用了公开密钥密码：一次用了他的私钥，一次用了Alice的公钥。

***目前最流行的两个安全电子邮件协议：*PGP：**一个开放源码的安全电子邮件软件包，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。（因特网安全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，作为工业标准被商业组织或机构使用。***PGP提供五种服务：***鉴别，机密性，压缩(在用Ks加密前进行Zip压缩，既可减少要加密的数据量，又使冗余消息变少，增加密码分析难度)，兼容电子邮件**Encode\_Base64(KB+(Ks) || Ks (Zip (Sgn || Data)))**，分段（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**）。

***使TCP连接安全：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）。***安全服务:*** 服务器鉴别，数据加密，客户鉴别（可选）。**SSL建立在TCP之上，依靠TCP提供可靠的端到端连接。SSL（有连接，有状态）**是***涉及到两个层次（应用和运输）的一组协议：***SSL记录协议：为各种高层协议（如HTTP）提供基本的安全服务；其它三个高层协议：用于SSL交换管理。**类SSL（简化SSL）的三个阶段：**握手，密钥导出，数据传输。**（1）握手：**Bob需要(a)与Alice创建一条TCP连接 (b)验证Alice是真实的Alice (c)发送给Alice一个主密钥，Bob与Alice持用该主密钥生成SSL会话所需的所有对称密钥。**(2)密钥导出：MS生成四个密钥**，两个加密密钥用于加密数据（用于从Bob发送到Alice的数据的会话加密密钥，用于从Alice发送到Bob的数据的会话加密密钥），两个MAC密钥用于验证数据的完整性（用于从Bob发送到Alice的数据的会话MAC密钥，用于从Alice发送到Bob的数据的会话MAC密钥）。 **(3)数据传输:**为了防止等到TCP会话结束时才验证所有数据的完整性，SSL将数据流分割成记录，对每个记录附加一个MAC用于完整性检查，然后加密“记录+MAC”包。MAC=，再。***阻止插入、删除和代替报文段：***使用序号；Bob维护一个序号计数器，计数器开始为0，Bob每发送的一个SSL记录它都增加1；Bob并不实际在记录中包括一个序号，但当他计算MAC时，他把该序号包括在MAC的计算中；所以，该MAC现在是数据加MAC密钥MB加当前序号的散列；Alice跟踪Bob的序号，通过在MAC的计算中包括适当的序号，使她验证一条记录的数据完整性。***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算法及密钥等。**握手协议由客户和服务器之间的一系列报文交换组成：（1）**客户向服务器发送支持的密码算法列表以及客户的不重数Rc **(2)**服务器从列表中选择一种对称算法（AES）、一种公钥算法（例如具有特定密钥长度的RSA）和一种MAC算法，与服务器选择的不重数Rs和证书一起发送给客户 **(3)**客户验证该证书，提取服务器的公钥，生成前主密钥PMS，用服务器的公钥加密该PMS，并将加密的PMS发送给服务器 **(4)**使用相同的密钥导出函数，客户和服务器独立地从PMS和不重数中计算出主密钥MS，然后该MS被切片以生成两个MAC密钥。（若选择的对称密码应用于CBC，则两个初始化向量IV也从该MS获得。）此后客户和服务器之间发送的所有报文均被加密和鉴别（使用MAC） **(5)**客户发送所有握手报文的一个MAC **(6)**服务器发送所有握手报文的一个MAC （步骤5，6使握手免受篡改危害，Trudy能从以明文形式发送的算法列表中删除较强的算法。故步骤5中发送一个级联它已发送和接收的所有握手报文的MAC，服务器能比较这个MAC与它已接收和已发送的握手报文的MAC，如有不一致，服务器能终止连接；类似，服务器也发送MAC，允许客户检查不一致性）**SSL中，不重数用于防御“连接重放”，序号防御在一个进行中的会话中重放个别分组** ***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：分片**(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214字节的数据块（可选）)**->压缩**(对数据块进行压缩；使用密码散列函数对数据块计算报文鉴别码)**->加MAC**（使用对称密钥算法对（压缩的）数据块）**->加密**(报文鉴别码进行加密，**->加头部**(在处理完的数据块前加上SSL头)。**SSL记录字段**：内容类型、SSL版本号、压缩数据块的长度、数据、MAC（数据和MAC用加密）。**连接关闭：**在类型字段指出该记录是否用于终止该SSL会话（类型包含在MAC，保证了不会被篡改，抵御**截断攻击**，防止Trudy介入一个进行中的SSL会话，并用TCP FIN过早结束该会话）。

***网络层安全性：IPsec和虚拟专用网：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。***IP安全协议（IPSec）：***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。**IPsec是面向连接的。IPsec和虚拟专用网（VPN）：*专用网：***通过电信专线将分散在各地的计算机（网络）连接而成的单独的物理网络；安全性好，但代价高。***虚拟专用网VPN：***建立在公用网上的一个覆盖网络，在逻辑上与其它流量隔离；数据在发送到公用网之前进行加密。**VPN的典型结构：**在每个局域网上设置一个安全网关，在每一对安全网关间创建一条**穿过因特网的隧道**，在隧道中使用IPSec。**VPN优点：**可以在一对局域网间提供完整性控制及机密性服务，对流量分析有抵御能力；对因特网中的路由器及用户软件透明，只要系统管理员设置好安全网关就可以了。**传进公用网的IPsec被IPv4封装，拥有IPv4首部**，因此在公共因特网中的路由器处理该数据报时视为普通数据报，但IPsec数据包的载荷包含一个IPsec首部，用于IPsec处理，且载荷是加密的。***IPSec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：IPSec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它部分的实现；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。**IPSec安全协议：**包括**AH**（不提供机密性）**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；**安全关联SA：**在从源实体向目的实体发送IPsec数据报前，源和目的实体创建了一个网络层的逻辑连接。**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。*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，并不对发出的所有数据加密（如访问公共资源）。【一个总部、一个分支机构和n个旅行销售员组成的机构VPN的SA有(2+2n)个】一个IPsec实体在他的安全关联**数据库SAD**中存储其所有**SA状态信息**（SPI，SA初始接口和目的接口，加密类型，加密密钥，完整性检查类型，鉴别密钥），SAD是实体操作系统内核中的一个数据结构。**IPsec数据报：运输模式：**路由器根据原始IP头转发数据包；**隧道模式：**路由器根据外层IP头（源和目的设置为隧道端点网关的接口）的信息转发数据包，隧道的端点（外层IP头中的地址）通常是一个支持IPSec的安全网关。***两种模式的比较：***传输模式比隧道模式占用较少的带宽，隧道模式更安全：隐藏内部网络的细节（原始IP头不可见）；内部网络上的主机可以不运行IPSec，由安全网关来保证安全性；隧道模式可以将一对端点间的通信聚合成一个加密流，从而有效地防止入侵者进行流量分析。**IPsec数据包格式(隧道模式，封装安全载荷ESP)：**新IP头：位于隧道两个端点的源和目的路由器接口（初始源和目的IP加密包装在载荷中），协议号设置为50；ESP头：包含SPI和SeqNum（SPI指示属于哪个SA；序号字段防止重放攻击）；载荷：原始数据包中被加密部分的密文；ESP尾：包括填充（需要的话，成为完整的块方便加密）、填充的长度和下一个头（包含在载荷字段中的数据类型），ESP尾也要被加密；鉴别数据：覆盖enchilada的报文鉴别码。**安全策略库SPD：**IPsec维护的数据结构，指示哪些类型的数据报（作为源IP地址、目的IP地址和协议类型的函数）将被IPsec处理，并对这些将被IPsec处理的数据报应当使用哪个SA。**SPD指示数据报“做什么“，SAD指示”怎样做“。** ***ESP协议提供的安全服务：***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被加密，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；原始IP头也被HMAC覆盖，因而ESP隧道模式可提供数据起源鉴别；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。**因特网密钥交换IKE：**自动生成SA的机制，类似于SSL握手，让两个实体交换证书，协商鉴别和加密算法，并安全地交换用于在IPsecSA中生成会话密钥的密钥材料。但它分成两个阶段执行（与SSL不同）：(1)建立双向IKE SA连接（不同于IPsec SA）(2)ISAKMP用于安全协商SAs对。第一阶段有两种模式：aggressive mode（使用更少信息） & main mode（提供身份保护且更灵活）。

*******使无线LAN安全：802.11 WEP有线等效保密：***最初的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。***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。**有线等效保密：*主机鉴别过程：***无线主机向AP请求鉴别；AP向主机发送一个128比特的不重数；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AP发送给主机的不重数相同，完成主机鉴别。**利用主机与基站共享密钥这个事实鉴别主机。*WEP数据加密：***主机与AP共享一个40比特的对称密钥KS（半永久）；对于每个帧，发送方生成一个24比特的初始向量IV，添加到KS后面，形成一个64比特的密钥 (KS, IV)（每一帧不同，**要求同一个64比特密钥不能使用超过1次**）。***WEP加密的安全漏洞：***每个KS只有224个(KS, IV) 可用：IV会被重复使用；IV用明文传输：攻击者可以观察到IV的重用。**攻击：**Trudy（可能通过欺骗方式）让Alice加密他选择的明文(d1 d2 d3 d4 … )；Trudy能够获得Alice加密的密文: ci = di XOR kiIV；Trudy知道ci和di, 就可以计算出kiIV： di XOR ci = kiIV；Trudy得到了加密所用的密钥流 k1IV k2IV k3IV……；当过后观察到IV被重用时，Trudy就可以破解密文了！***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。***802.11i的操作：*1) 发现。**在发现阶段，AP通告它的存在以及它能够向无线客户结点提供的鉴别和加密形式，客户则请求它希望的特定鉴别和加密形式。**2)相互鉴别和主密钥(MK)生成。**鉴别发生在无线客户和鉴别服务器之间。在这个阶段，接入点基本是起中继的作用，在客户和鉴别服务器之间转发报文。**3)成对主密钥PMK生成。**MK是一个仅为客户和鉴别服务器所知的共享密钥，它们都使用MK来生成一个次密钥，即成对主密钥(PMK)。鉴别服务器则向AP发送该PMK。客户和AP 现在具有一个共享的密钥。**可扩展鉴别协议EAP**：定义了客户和鉴别服务器之间交互使简单的请求/响应模式中使用的端到端报文格式。**EAPoL**封装，通过802.11无线链路发送，在接入点拆封，再使用**RADIUS**协议重新封装，经UDP/IP传输到鉴别服务器。**4)临时密钥TK生成。**使用PMK, 无线客户和AP现在能够生成附加的、将用于通信的密钥。关键是临时密钥TK，TK将被用于执行经无线链路向任意远程主机发送数据的链路级的加密。

***运行安全性：防火墙和入侵检测系统：* 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（硬软结合体）；**目的是**保护内部网络免受来自外部网络的攻击。***防火墙的类型：*(1)*传统分组过滤器：***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包还是丢弃包：源IP地址，目的IP地址；TCP/UDP源端口号、目的端口号； ICMP 报文类型；TCP SYN标志和 ACK标志。***包过滤策略的例子：***不允许访问外部Web网站→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→丢弃进入的TCP SYN包，除非去往130.207.244.203的端口80；防止因特网广播吞噬网络带宽→除DNS包和路由器广播包，丢弃其它进入的UDP包；防止网络拓扑被探测→丢弃所有外出的ICMP TTL expired包；阻止外部客户发起到内部服务器的连接→过滤进入的所有ACK=0的报文段，（去除了所有从外部发起的所有TCP连接，但是允许内部发起TCP连接）。***访问控制列表ACL：***是规则表，包含一系列（动作，匹配条件）；对于每个进出的包，从上到下地匹配规则。***包过滤防火墙孤立地过滤每个包，仍会允许一些异常的包进入：***如允许dest port = 80, ACK=1的包进入，哪怕并没有相应的连接存在。**(2)*状态检测防火墙***可以跟踪TCP连接的状态:跟踪连接的建立（SYN）和关闭（FIN）等状态，判断收到的包**是否有意义**；扩展ACL，指示在允许放行一个包前需检查连接的状态。**(3)*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telnet登录外部主机；所有telnet用户必须连接到应用网关；对于授权用户，鉴定用户之后，应用网关建立与目的主机的telnet会话，并在2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telnet连接。***防火墙的局限性：***无法抵御IP欺骗攻击: 路由器无法知道包是否来自声称源；**应用网关缺陷**：处理开销大，速度慢；每个被代理的应用都需要一个应用网关；应用网关对于用户不透明：客户软件必须设置应用网关的IP地址；对于UDP包，过滤器或者全部允许，或全部禁止；和外界的通信强度与网络安全等级矛盾；许多受到高度保护的站点仍然遭到攻击。**入侵检测系统：**防火墙不检查数据包之间的**关联。入侵检测系统IDS*：***观察到潜在恶意流量时能产生警告的设备；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：端口扫描、DoS攻击。**入侵防止系统IPS：**滤除可疑流量；**网络中可以设置多个IDS:** 在不同位置进行不同类型的检查。