

11. 下列描述中, () 不是软件定义网络 (SDN) 的特点。
- A. 控制与转发功能分离
 - B. 控制平面集中化
 - C. 接口开放可编程
 - D. Openflow 取代了路由协议
12. 【2022 统考真题】在 SDN 网络体系结构中, SDN 控制器向数据平面的 SDN 交换机下发流表时所使用的接口是 ()。
- A. 东向接口
 - B. 南向接口
 - C. 西向接口
 - D. 北向接口

4.1.6 答案与解析

单项选择题

01. C

选项 A、B 不是网络层的目的, IP 提供的是不可靠的服务, 因此选项 D 错误。

02. C

网络的异构性是指传输介质、数据编码方式、链路控制协议及不同的数据单元格式和转发机制, 这些特点分别在物理层和数据链路层协议中定义。

03. D

拥塞现象是指到达通信子网中某一部分的分组数量过多, 使得该部分网络来不及处理, 以致引起这部分乃至整个网络性能下降的现象, 严重时甚至会导致网络通信业务陷入停顿, 即出现死锁现象。选项 A 的网络性能显然是提高的, 选项 B、C 中网络结点接收和发出的分组多少与网络的吞吐量并不呈正比关系, 不能确定网络是否拥塞。

04. C

路由器是第三层设备, 向传输层及以上层次隐藏下层的具体实现, 所以物理层、数据链路层、网络层协议可以不同。而网络层之上的协议数据是路由器所不能处理的, 因此网络层以上的高层协议必须相同。本题容易误选 B, 主要原因是在目前的互联网中广泛使用的是 TCP/IP 协议族, 在网络层用的多是 IPv4, 所以误认为网络层协议必须相同。而实际上, 使用特定的路由器连接 IPv4 与 IPv6 网络, 就是典型的网络层协议不同而实现互连的例子。

05. C

路由器工作在网络层, 不转发广播包 (目的地址为 255.255.255.255 的 IP 包), 因此能够分隔广播域, 抑制网络风暴。交换机工作在数据链路层, 能够分隔冲突域, 但不能分隔广播域。集线器和中继器是物理层设备, 既不能分隔广播域又不能分隔冲突域。

06. C

路由器是网络层设备, 其任务是转发分组。每个路由器都维护一个路由表以决定分组的转发。为了提高路由器的查询效率并减少路由表维护的内容, 路由表只保留到达目的地址的下一个路由器的地址, 而不保留整个传输路径的信息。另外, 采用目的网络可使每个路由表项包含很多目的主机 IP 地址, 这样可减少路由表中的项目。因此, 路由表通常包含目的网络和到达该目的网络路径上的下一个路由器的 IP 地址。

07. C

路由器是网络层设备, 网络层通过 IP 地址标识主机, 所以路由器根据 IP 地址转发分组。

08. A

路由器转发一个分组的过程如下: 先接收整个分组, 然后对分组进行错误检查, 如果出错, 那么丢弃错误的分组; 否则存储该正确的分组。最后根据路由选择协议, 将正确的分组转发到合适的端口, 这种机制称为存储转发机制。

09. B

每个路由器都根据路由表选择 IP 分组的下一跳地址，只有到了下一跳路由器，才能知道再下一跳应当怎样走。主机仅知道到达本地网络的路径，到达其他网络的 IP 分组均转发到路由器。而源主机也只把 IP 分组发给网关，所以路由器和源主机都不知道 IP 分组要经过的完整路径。

10. D

TCP 属于传输层协议，FTP 属于应用层协议，只有 IP 和 ICMP 属于网络层协议。

11. D

选项 A、B 和 C 都是 SDN 的特点。Openflow 协议是控制平面和数据平面之间的接口。在 SDN 中，路由器之间不再相互交换路由信息，由远程控制器计算出最佳路由。

12. B

SDN 对上层开发者提供的编程接口称为北向接口，而南向接口则负责控制平面和数据平面间的通信，所以 SDN 控制器向数据平面的 SDN 交换机下发流表时使用南向接口。

4.2 路由算法

4.2.1 静态路由与动态路由

路由器转发分组是通过路由表转发的，而路由表是通过各种算法得到的。从能否随网络的通信量或拓扑自适应地进行调整变化来划分，路由算法可以分为如下两大类。

静态路由算法（又称非自适应路由算法）。指由网络管理员手工配置的路由信息。当网络的拓扑结构或链路的状态发生变化时，网络管理员需要手工去修改路由表中相关的静态路由信息。它不能及时适应网络状态的变化，对于简单的小型网络，可以采用静态路由。

动态路由算法（又称自适应路由算法）。指路由器上的路由表项是通过相互连接的路由器之间彼此交换信息，然后按照一定的算法优化出来的，而这些路由信息会在一定时间间隙里不断更新，以适应不断变化的网络，随时获得最优的寻路效果。

静态路由算法的特点是简便和开销较小，在拓扑变化不大的小网络中运行效果很好。动态路由算法能改善网络的性能并有助于流量控制；但算法复杂，会增加网络的负担，有时因对动态变化的反应太快而引起振荡，或反应太慢而影响网络路由的一致性，因此要仔细设计动态路由算法，以发挥其优势。常用的动态路由算法可分为两类：距离-向量路由算法和链路状态路由算法。

4.2.2 距离-向量路由算法

在距离-向量路由算法中，所有结点都定期地将它们的整个路由选择表传送给所有与之直接相邻的结点。这种路由选择表包含：

- 每条路径的目的地（另一结点）。
- 路径的代价（也称距离）。

注意：这里的距离是一个抽象的概念，如 RIP 就将距离定义为“跳数”。跳数指从源端口到达目的端口所经过的路由器个数，每经过一个路由器，跳数加 1。

在这种算法中，所有结点都必须参与距离向量交换，以保证路由的有效性和一致性，也就是说，所有的结点都监听从其他结点传来的路由选择更新信息，并在下列情况下更新它们的路由选择表：

关注公众号【乘龙考研】
一手更新 稳定有保障

- 1) 被通告一条新的路由, 该路由在本结点的路由表中不存在, 此时本地系统加入这条新的路由。
- 2) 发来的路由信息中有一条到达某个目的地的路由, 该路由与当前使用的路由相比, 有较短的距离(较小的代价)。此种情况下, 就用经过发送路由信息的结点的新路由替换路由表中到达那个目的地的现有路由。

距离-向量路由算法的实质是, 迭代计算一条路由中的站段数或延迟时间, 从而得到到达一个目标的最短(最小代价)通路。它要求每个结点在每次更新时都将它的全部路由表发送给所有相邻的结点。显然, 更新报文的大小与通信子网的结点个数成正比, 大的通信子网将导致很大的更新报文。由于更新报文发给直接邻接的结点, 所以所有结点都将参加路由选择信息交换。基于这些原因, 在通信子网上传递的路由选择信息的数量很容易变得非常大。

最常见的距离-向量路由算法是 RIP 算法, 它采用“跳数”作为距离的度量。

4.2.3 链路状态路由算法

链路状态路由算法要求每个参与该算法的结点都具有完全的网络拓扑信息, 它们执行下述两项任务。第一, 主动测试所有邻接结点的状态。两个共享一条链接的结点是相邻结点, 它们连接到同一条链路, 或者连接到同一广播型物理网络。第二, 定期地将链路状态传播给所有其他结点(或称路由结点)。典型的链路状态算法是 OSPF 算法。

在一个链路状态路由选择中, 一个结点检查所有直接链路的状态, 并将所得的状态信息发送给网上的所有其他结点, 而不是仅送给那些直接相连的结点。每个结点都用这种方式从网上所有其他的结点接收包含直接链路状态的路由选择信息。

每当链路状态报文到达时, 路由结点便使用这些状态信息去更新自己的网络拓扑和状态“视野图”, 一旦链路状态发生变化, 结点就对更新的网络图利用 Dijkstra 最短路径算法重新计算路由, 从单一的源出发计算到达所有目的结点的最短路径。

链路状态路由算法主要有三个特征:

- 1) 向本自治系统中所有路由器发送信息, 这里使用的方法是洪泛法, 即路由器通过所有端口向所有相邻的路由器发送信息。而每个相邻路由器又将此信息发往其所有相邻路由器(但不再发送给刚刚发来信息的那个路由器)。
- 2) 发送的信息是与路由器相邻的所有路由器的链路状态, 但这只是路由器所知道的部分信息。所谓“链路状态”, 是指说明本路由器与哪些路由器相邻及该链路的“度量”。对于 OSPF 算法, 链路状态的“度量”主要用来表示费用、距离、时延、带宽等。
- 3) 只有当链路状态发生变化时, 路由器才向所有路由器发送此信息。

由于一个路由器的链路状态只涉及相邻路由器的连通状态, 而与整个互联网的规模并无直接关系, 因此链路状态路由算法可以用于大型的或路由信息变化聚敛的互联网环境。

链路状态路由算法的主要优点是, 每个路由结点都使用同样的原始状态数据独立地计算路径, 而不依赖中间结点的计算; 链路状态报文不加改变地传播, 因此采用该算法易于查找故障。当一个结点从所有其他结点接收到报文时, 它可以在本地立即计算正确的通路, 保证一步汇聚。最后, 由于链路状态报文仅运载来自单个结点关于直接链路的信息, 其大小与网络中的路由结点数目无关, 因此链路状态算法比距离-向量算法有更好的规模可伸展性。

距离-向量路由算法与链路状态路由算法的比较: 在距离-向量路由算法中, 每个结点仅与它的直接邻居交谈, 它为它的邻居提供从自己到网络中所有其他结点的最低费用估计。在链路状态路由算法中, 每个结点通过广播的方式与所有其他结点交谈, 但它仅告诉它们与它直接相连的链

路的费用。相较之下，距离-向量路由算法有可能遇到路由环路等问题。

4.2.4 层次路由

当网络规模扩大时，路由器的路由表成比例地增大。这不仅会消耗越来越多的路由器缓冲区空间，而且需要用更多CPU时间来扫描路由表，用更多的带宽来交换路由状态信息。因此路由选择必须按照层次的方式进行。

因特网将整个互联网划分为许多较小的自治系统（注意一个自治系统中包含很多局域网），每个自治系统有权自主地决定本系统内应采用何种路由选择协议。如果两个自治系统需要通信，那么就需要一种在两个自治系统之间的协议来屏蔽这些差异。据此，因特网把路由选择协议划分为两大类：

- 1) 一个自治系统内部所使用的路由选择协议称为内部网关协议（IGP），也称域内路由选择，具体的协议有RIP和OSPF等。
- 2) 自治系统之间所使用的路由选择协议称为外部网关协议（EGP），也称域间路由选择，在不同自治系统的路由器之间交换路由信息，并负责为分组在不同自治系统之间选择最优的路径。具体的协议有BGP。

使用层次路由时，OSPF将一个自治系统再划分为若干区域（Area），每个路由器都知道在本区域内如何把分组路由到目的地的细节，但不用知道其他区域的内部结构。

采用分层次划分区域的方法虽然会使交换信息的种类增多，也会使OSPF协议更加复杂。但这样做却能使每个区域内部交换路由信息的通信量大大减小，因而使OSPF协议能够用于规模很大的自治系统中。

4.2.5 本节习题精选

单项选择题

01. 动态路由选择和静态路由选择的主要区别是()。
- A. 动态路由选择需要维护整个网络的拓扑结构信息，而静态路由选择只需要维护部分拓扑结构信息
 - B. 动态路由选择可随网络的通信量或拓扑变化而自适应地调整，而静态路由选择则需要手工去调整相关的路由信息
 - C. 动态路由选择简单且开销小，静态路由选择复杂且开销大
 - D. 动态路由选择使用路由表，静态路由选择不使用路由表
02. 下列关于路由算法的描述中，()是错误的。
- A. 静态路由有时也被称为非自适应的算法
 - B. 静态路由所使用的路由选择一旦启动就不能修改
 - C. 动态路由也称自适应算法，会根据网络的拓扑变化和流量变化改变路由决策
 - D. 动态路由算法需要实时获得网络的状态
03. 关于链路状态协议的描述，()是错误的。
- A. 仅相邻路由器需要交换各自的路由表
 - B. 全网路由器的拓扑数据库是一致的
 - C. 采用洪泛技术更新链路变化信息
 - D. 具有快速收敛的优点
04. 在链路状态路由算法中，每个路由器都得到网络的完整拓扑结构后，使用()算法来找出它到其他路由器的路径长度。
- A. Prim最小生成树算法
 - B. Dijkstra最短路径算法

关注公众号【乘龙考研】
一手更新 稳定有保障

- C. Kruskal 最小生成树算法 D. 拓扑排序
05. 在距离-向量路由协议中, () 最可能导致路由回路的问题。
- A. 由于网络带宽的限制, 某些路由更新数据报被丢弃
 - B. 由于路由器不知道整个网络的拓扑结构信息, 当收到一个路由更新信息时, 又将该更新信息发回自己发送该路由信息的路由器
 - C. 当一个路由器发现自己的一条直接相邻链路断开时, 未能将这个变化报告给其他路由器
 - D. 慢收敛导致路由器接收了无效的路由信息
06. 下列关于分层路由的描述中, () 是错误的。
- A. 采用分层路由后, 路由器被划分成区域
 - B. 每个路由器不仅知道如何将分组路由到自己区域的目标地址, 而且知道如何路由到其他区域
 - C. 采用分层路由后, 可以将不同的网络连接起来
 - D. 对于大型网络, 可能需要多级的分层路由来管理

4.2.6 答案与解析

单项选择题

01. B

静态路由选择使用手动配置的路由信息, 实现简单且开销小, 需要维护整个网络的拓扑结构信息, 但不能及时适应网络状态的变化。动态路由选择通过路由选择协议, 自动发现并维护路由信息, 能及时适应网络状态的变化, 实现复杂且开销大。动态路由选择和静态路由选择都使用路由表。

02. B

静态路由又称非自适应算法, 它不会估计流量和结构来调整其路由决策。但这并不说明路由选择是不能改变的, 事实上用户可以随时配置路由表。而动态路由也称自适应算法, 需要实时获取网络的状态, 并根据网络的状态适时地改变路由决策。

03. A

在链路状态路由算法中, 每个路由器在自己的链路状态变化时, 将链路状态信息用洪泛法传送给网络中的其他路由器。发送的链路状态信息包括该路由器的相邻路由器及所有相邻链路的状态, 选项 A 错误。链路状态协议具有快速收敛的优点, 它能够在网络拓扑发生变化时, 立即进行路由的重新计算, 并及时向其他路由器发送最新的链路状态信息, 使得各路由器的链路状态表能够尽量保持一致, 选项 B、C、D 正确。

04. B

在链路状态路由算法中, 路由器通过交换每个结点到邻居结点的延迟或开销来构建一个完整的网络拓扑结构。得到完整的拓扑结构后, 路由器就使用 Dijkstra 最短路径算法来计算到所有结点的最短路径。

05. D

在距离-向量路由协议中, “好消息传得快, 而坏消息传得慢”, 这就导致了当路由信息发生变化时, 该变化未能及时地被所有路由器知道, 而仍然可能在路由器之间进行传递, 这就是“慢收敛”现象。慢收敛是导致发生路由回路的根本原因。

06. B

采用分层路由后, 路由器被划分为区域, 每个路由器知道如何将分组路由到自己所在区域内的目标地址, 但对于其他区域内的结构毫不知情。当不同的网络相互连接时, 可将每个网络当作一个

独立的区域，这样做的好处是一个网络中的路由器不必知道其他网络的拓扑结构。

4.3 IPv4

关注公众号【乘龙考研】
一手更新稳定有保障

4.3.1 IPv4 分组

IPv4 即现在普遍使用的 IP 协议（版本 4）。IP 协议定义数据传送的基本单元——IP 分组及其确切的数据格式。IP 协议也包括一套规则，指明分组如何处理、错误怎样控制。特别是 IP 协议还包含非可靠投递的思想，以及与此关联的分组路由选择的思想。

1. IPv4 分组的格式

一个 IP 分组由首部和数据部分组成。首部前一部分的长度固定，共 20B，是所有 IP 分组必须具有的。在首部固定部分的后面是一些可选字段，其长度可变，用来提供错误检测及安全等机制。IP 数据报的格式如图 4.4 所示。

IP 首部的部分重要字段含义如下：

- 1) 版本。指 IP 协议的版本，目前广泛使用的版本号为 4。
- 2) 首部长度。占 4 位，可以表示的最大十进制数是 15。以 32 位为单位，最大值为 60B (15×4B)。最常用的首部长度是 20B，此时不使用任何选项（即可选字段）。



图 4.4 IP 数据报的格式

- 3) 总长度。占 16 位。指首部和数据之和的长度，单位为字节，因此数据报的最大长度为 $2^{16} - 1 = 65535$ B。以太网帧的最大传送单元 (MTU) 为 1500B，因此当一个 IP 数据报封装成帧时，数据报的总长度（首部加数据）一定不能超过下面的数据链路层的 MTU 值。
- 4) 标识。占 16 位。它是一个计数器，每产生一个数据报就加 1，并赋值给标识字段。但它并不是“序号”（因为 IP 是无连接服务）。当一个数据报的长度超过网络的 MTU 时，必须分片，此时每个数据报片都复制一次标识号，以便能正确重装成原来的数据报。
- 5) 标志。占 3 位。标志字段的最低位为 MF，MF=1 表示后面还有分片，MF=0 表示最后一个分片。标志字段中间的一位是 DF，只有当 DF=0 时才允许分片。
- 6) 片偏移。占 13 位。它指出较长的分组在分片后，某片在原分组中的相对位置。片偏移以 8 个字节为偏移单位。除最后一个分片外，每个分片的长度一定是 8B 的整数倍。

- 7) 生存时间 (TTL)。占 8 位。数据报在网络中可通过的路由器数的最大值，标识分组在网络中的寿命，以确保分组不会永远在网络中循环。路由器在转发分组前，先把 TTL 减 1。若 TTL 被减为 0，则该分组必须丢弃。
- 8) 协议。占 8 位。指出此分组携带的数据使用何种协议，即分组的数据部分应上交给哪个协议进行处理，如 TCP、UDP 等。其中值为 6 表示 TCP，值为 17 表示 UDP。
- 9) 首部校验和。占 16 位。首部校验和只校验分组的首部，而不校验数据部分。
- 10) 源地址字段。占 4B，标识发送方的 IP 地址。
- 11) 目的地址字段。占 4B，标识接收方的 IP 地址。

注意：在 IP 数据报首部中有三个关于长度的标记，首部长度、总长度、片偏移，基本单位分别为 4B、1B、8B（需要记住）。题目中经常会出现这几个长度之间的加减运算。另外，读者要熟悉 IP 数据报首部的各个字段的意义和功能，但不需要记忆 IP 数据报的首部，正常情况下如果需要参考首部，题目都会直接给出。第 5 章学到的 TCP、UDP 的首部也是一样的。

2. IP 数据报分片

一个链路层数据报能承载的最大数据量称为最大传送单元 (MTU)。因为 IP 数据报被封装在链路层数据报中，因此链路层的 MTU 严格地限制着 IP 数据报的长度，而且在 IP 数据报的源与目的地路径上的各段链路可能使用不同的链路层协议，有不同的 MTU。例如，以太网的 MTU 为 1500B，而许多广域网的 MTU 不超过 576B。当 IP 数据报的总长度大于链路 MTU 时，就需要将 IP 数据报中的数据分装在多个较小的 IP 数据报中，这些较小的数据报称为片。

片在目的地的网络层被重新组装。目的主机使用 IP 首部中的标识、标志和片偏移字段来完成对片的重组。创建一个 IP 数据报时，源主机为该数据报加上一个标识号。当一个路由器需要将一个数据报分片时，形成的每个数据报（即片）都具有原始数据报的标识号。当目的主机收到来自同一发送主机的一批数据报时，它可以通过检查数据报的标识号来确定哪些数据报属于同一个原始数据报的片。IP 首部中的标志位占 3 位，但只有后 2 位有意义，分别是 MF 位 (More Fragment) 和 DF 位 (Don't Fragment)。只有当 DF=0 时，该 IP 数据报才可以被分片。MF 则用来告知目的主机该 IP 数据报是否为原始数据报的最后一个片。当 MF=1 时，表示相应的原始数据报还有后续的片；当 MF=0 时，表示该数据报是相应原始数据报的最后一个片。目的主机在对片进行重组时，使用片偏移字段来确定片应放在原始 IP 数据报的哪个位置。

IP 分片涉及一定的计算。例如，一个长 4000B 的 IP 数据报（首部 20B，数据部分 3980B）到达一个路由器，需要转发到一条 MTU 为 1500B 的链路上。这意味着原始数据报中的 3980B 数据必须被分配到 3 个独立的片中（每片也是一个 IP 数据报）。假定原始数据报的标识号为 777，那么分成的 3 片如图 4.5 所示。可以看出，由于偏移值的单位是 8B，所以除最后一个片外，其他所有片中的有效数据载荷都是 8 的倍数。

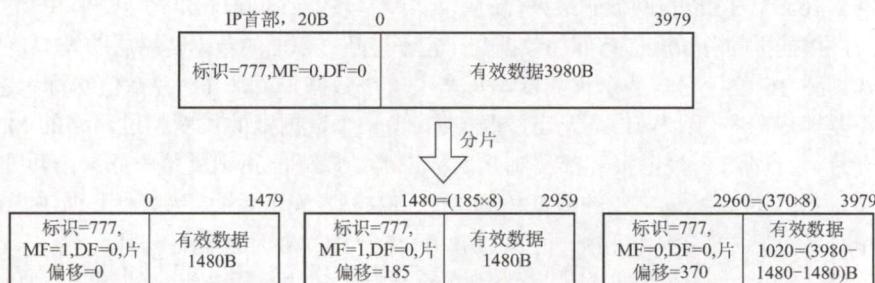


图 4.5 IP 分片的例子

4.3.2 IPv4 地址与 NAT

1. IPv4 地址

连接到因特网上的每台主机（或路由器）都分配一个 32 比特的全球唯一标识符，即 IP 地址。IP 地址由互联网名字和数字地址分配机构 ICANN 进行分配。

互联网早期采用的是分类的 IP 地址，如图 4.6 所示。

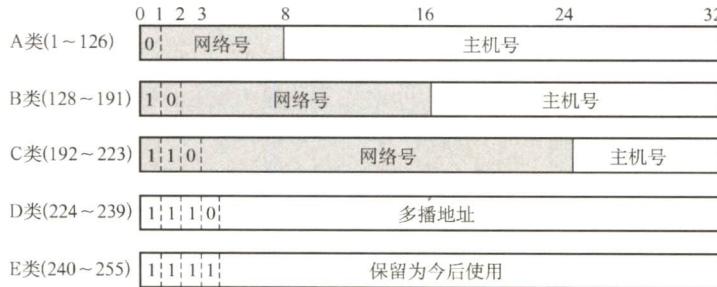


图 4.6 分类的 IP 地址

无论哪类 IP 地址，都由网络号和主机号两部分组成。即 IP 地址 $::=\{\langle\text{网络号}\rangle,\langle\text{主机号}\rangle\}$ 。其中网络号标志主机（或路由器）所连接到的网络。一个网络号在整个因特网范围内必须是唯一的。主机号标志该主机（或路由器）。一台主机号在它前面的网络号所指明的网络范围内必须是唯一的。由此可见，一个 IP 地址在整个因特网范围内是唯一的。

在各类 IP 地址中，有些 IP 地址具有特殊用途，不用做主机的 IP 地址：

- 主机号全为 0 表示本网络本身，如 202.98.174.0。
- 主机号全为 1 表示本网络的广播地址，又称直接广播地址，如 202.98.174.255。
- 127.x.x.x 保留为环回自检（Loopback Test）地址，此地址表示任意主机本身，目的地址为环回地址的 IP 数据报永远不会出现在任何网络上。
- 32 位全为 0，即 0.0.0.0 表示本网络上的本主机。
- 32 位全为 1，即 255.255.255.255 表示整个 TCP/IP 网络的广播地址，又称受限广播地址。实际使用时，由于路由器对广播域的隔离，255.255.255.255 等效为本网络的广播地址。

常用的三种类别 IP 地址的使用范围见表 4.1。

表 4.1 常用的三种类别 IP 地址的使用范围^①

网络类别	最大可用网络数	第一个可用的网络号	最后一个可用的网络号	每个网络中的最大主机数
A	$2^7 - 2$	1	126	$2^{24} - 2$
B	2^{14}	128.0	191.255	$2^{16} - 2$
C	2^{21}	192.0.0	223.255.255	$2^8 - 2$

在表 4.1 中，A 类地址可用的网络数为 $2^7 - 2$ ，减 2 的原因是：第一，网络号字段全为 0 的 IP 地址是保留地址，意思是“本网络”；第二，网络号为 127 的 IP 地址是环回自检地址。

IP 地址有以下重要特点：

- 1) 每个 IP 地址都由网络号和主机号两部分组成，因此 IP 地址是一种分等级的地址结构。分等级的好处是：①IP 地址管理机构在分配 IP 地址时只分配网络号，而主机号则由得到该

^① B 类网络地址 128.0.0.0 和 C 类网络地址 192.0.0.0 早期标准规定不能指派，但现在都能指派[RFC6890]。

网络的单位自行分配，方便了 IP 地址的管理；②路由器仅根据目的主机所连接的网络号来转发分组（而不考虑目标主机号），从而减小了路由表所占的存储空间。

- 2) IP 地址是标志一台主机（或路由器）和一条链路的接口。当一台主机同时连接到两个网络时，该主机就必须同时具有两个相应的 IP 地址，每个 IP 地址的网络号必须与所在网络的网络号相同，且这两个 IP 地址的主机号是不同的。因此 IP 网络上的一个路由器必然至少应具有两个 IP 地址（路由器每个端口必须至少分配一个 IP 地址）。
 - 3) 用转发器或桥接器（网桥等）连接的若干 LAN 仍然是同一个网络（同一个广播域），因此该 LAN 中所有主机的 IP 地址的网络号必须相同，但主机号必须不同。
 - 4) 在 IP 地址中，所有分配到网络号的网络（无论是 LAN 还是 WAN）都是平等的。
 - 5) 在同一个局域网上的主机或路由器的 IP 地址中的网络号必须是一样的。路由器总是具有两个或两个以上的 IP 地址，路由器的每个端口都有一个不同网络号的 IP 地址。
- 近年来，由于广泛使用无分类 IP 地址进行路由选择，这种传统分类的 IP 地址已成为历史。

2. 网络地址转换 (NAT)

网络地址转换 (NAT) 是指通过将专用网络地址（如 Intranet）转换为公用地址（如 Internet），从而对外隐藏内部管理的 IP 地址。它使得整个专用网只需要一个全球 IP 地址就可以与因特网连通，由于专用网本地 IP 地址是可重用的，所以 NAT 大大节省了 IP 地址的消耗。同时，它隐藏了内部网络结构，从而降低了内部网络受到攻击的风险。

此外，为了网络安全，划出了部分 IP 地址为私有 IP 地址。私有 IP 地址只用于 LAN，不用于 WAN 连接（因此私有 IP 地址不能直接用于 Internet，必须通过网关利用 NAT 把私有 IP 地址转换为 Internet 中合法的全球 IP 地址后才能用于 Internet），并且允许私有 IP 地址被 LAN 重复使用。这有效地解决了 IP 地址不足的问题。私有 IP 地址网段如下：

A 类：1 个 A 类网段，即 **10.0.0.0~10.255.255.255**。

B 类：16 个 B 类网段，即 **172.16.0.0~172.31.255.255**。

C 类：256 个 C 类网段，即 **192.168.0.0~192.168.255.255**。

在因特网中的所有路由器，对目的地址是私有地址的数据报一律不进行转发。这种采用私有 IP 地址的互联网络称为专用互联网或本地互联网。私有 IP 地址也称可重用地址。

使用 NAT 时需要在专用网连接到因特网的路由器上安装 NAT 软件，NAT 路由器至少有一个有效的外部全球 IP 地址。使用本地地址的主机和外界通信时，NAT 路由器使用 NAT 转换表进行本地 IP 地址和全球 IP 地址的转换。NAT 转换表中存放着{本地 IP 地址：端口}到{全球 IP 地址：端口}的映射。通过这种映射方式，可让多个私有 IP 地址映射到一个全球 IP 地址。

表 4.2 一个典型的 NAT 转换表

NAT 转换表	
WAN 端	LAN 端
138.76.29.7, 5001	192.168.0.2, 2233
138.76.29.7, 5060	192.168.0.3, 1234
...	...

以宿舍共享上网为例，假设某个宿舍办理了 2Mb/s 的电信宽带，那么这个宿舍就获得了一个全球 IP 地址（如 138.76.29.7），而宿舍内 4 台主机使用私有地址（如 192.168.0.0 网段）。宿舍的网关路由器应开启 NAT 功能，并且某时刻路由器上的 NAT 转换表见表 4.2。那么，当路由器从 LAN 端口收到源 IP 及源端口号为 192.168.0.2, 2233 的数据报时，就将其映射成 138.76.29.7, 5001，然后从 WAN 端口发送到因特网上。当路由器从 WAN 端口收到目的 IP 及目的端口号为 138.76.29.7, 5060 的数据报时，就将其映射成 192.168.0.3, 1234，然后从 LAN 端口发送给相应的本地主机。这样，只需要一个全球地址，就可以让多台主机同时访问因特网。

下面以图 4.7 为例来说明 NAT 路由器的工作原理：①假设用户主机 10.0.0.1（随机端口 3345）向 Web 服务器 128.119.40.186（端口 80）发送请求。②NAT 路由器收到 IP 分组后，为该 IP 分组生成一个新端口号 5001，将 IP 分组的源地址更改为 138.76.29.7（即 NAT 路由器的全球 IP 地址），将源端口号更改为 5001。NAT 路由器在 NAT 转换表中增加一个表项。③Web 服务器并不知道刚抵达的 IP 分组已被 NAT 路由器进行了改装，更不知道用户的专用地址，它响应的 IP 分组的目的地址是 NAT 路由器的全球 IP 地址，目的端口号是 5001。④响应分组到达 NAT 路由器后，通过 NAT 转换表将 IP 分组的目的 IP 地址更改为 10.0.0.1，将目的端口号更改为 3345。

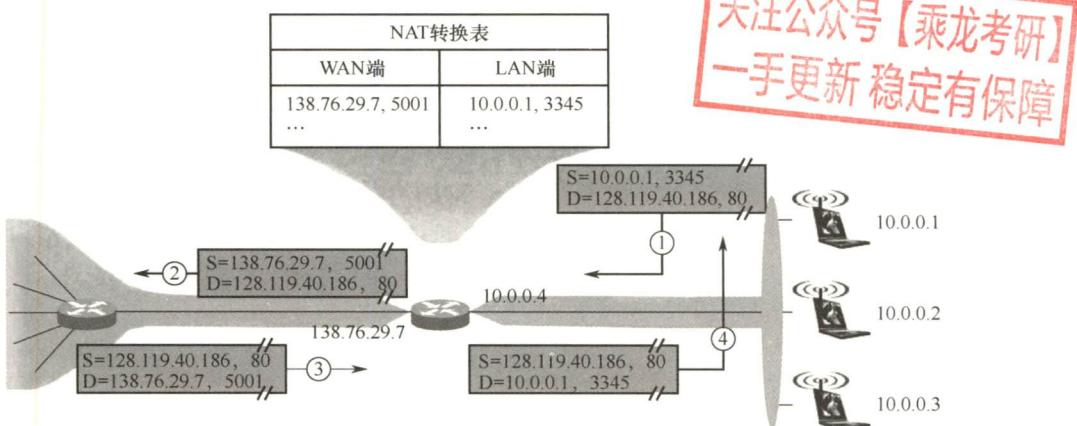


图 4.7 NAT 路由器的工作原理

注意：普通路由器在转发 IP 数据报时，不改变其源 IP 地址和目的 IP 地址。而 NAT 路由器在转发 IP 数据报时，一定要更换其 IP 地址（转换源 IP 地址或目的 IP 地址）。普通路由器仅工作在网络层，而 NAT 路由器转发数据报时需要查看和转换传输层的端口号。

4.3.3 子网划分与子网掩码、CIDR

1. 子网划分

两级 IP 地址的缺点：IP 地址空间的利用率有时很低；给每个物理网络分配一个网络号会使路由表变得太大而使网络性能变坏；两级的 IP 地址不够灵活。

从 1985 年起，在 IP 地址中又增加了一个“子网号字段”，使两级 IP 地址变成了三级 IP 地址。这种做法称为子网划分。子网划分已成为因特网的正式标准协议。

子网划分的基本思路如下：

- 子网划分纯属一个单位内部的事情。单位对外仍然表现为没有划分子网的网络。
- 从主机号借用若干比特作为子网号，当然主机号也就相应减少了相同的比特。三级 IP 地址的结构如下：IP 地址 = {<网络号>, <子网号>, <主机号>}。
- 凡是从其他网络发送给本单位某台主机的 IP 数据报，仍然是根据 IP 数据报的目的网络号，先找到连接到本单位网络上的路由器。然后该路由器在收到 IP 数据报后，按目的网络号和子网号找到目的子网。最后把 IP 数据报直接交付给目的主机。

注意：

- 1) 划分子网只是把 IP 地址的主机号这部分进行再划分，而不改变 IP 地址原来的网络号。因此，从一个 IP 地址本身或 IP 数据报的首部，无法判断源主机或目的主机所连接的网络是否进行了子网划分。

2) RFC 950 规定, 对分类的 IPv4 地址进行子网划分时, 子网号不能为全 1 或全 0。但随着 CIDR 的广泛使用, 现在全 1 和全 0 的子网号也可使用, 但一定要谨慎使用, 要弄清你的路由器所用的路由选择软件是否支持全 0 或全 1 的子网号。

3) 不论是分类的 IPv4 地址还是 CIDR, 其子网中的主机号为全 0 或全 1 的地址都不能被指派。子网中主机号全 0 的地址为子网的网络号, 主机号全 1 的地址为子网的广播地址。

2. 子网掩码

为了告诉主机或路由器对一个 A 类、B 类、C 类网络进行了子网划分, 使用子网掩码来表达对原网络中主机号的借位。

子网掩码是一个与 IP 地址相对应的、长 32bit 的二进制串, 它由一串 1 和跟随的一串 0 组成。其中, 1 对应于 IP 地址中的网络号及子网号, 而 0 对应于主机号。计算机只需将 IP 地址和其对应的子网掩码逐位“与”(逻辑 AND 运算), 就可得出相应子网的网络地址。

现在的因特网标准规定: 所有的网络都必须使用子网掩码。如果一个网络未划分子网, 那么就采用默认子网掩码。A、B、C 类地址的默认子网掩码分别为 255.0.0.0、255.255.0.0、255.255.255.0。例如, 某主机的 IP 地址 192.168.5.56, 子网掩码为 255.255.255.0, 进行逐位“与”运算后, 得出该主机所在子网的网络号为 192.168.5.0。

由于子网掩码是一个网络或一个子网的重要属性, 所以路由器在相互之间交换路由信息时, 必须把自己所在网络(或子网)的子网掩码告诉对方。路由表中的每个条目, 除要给出目的网络地址和下一跳地址外, 还要同时给出该目的网络的子网掩码。

在使用子网掩码的情况下:

- 1) 一台主机在设置 IP 地址信息的同时, 必须设置子网掩码。
- 2) 同属于一个子网的所有主机及路由器的相应端口, 必须设置相同的子网掩码。
- 3) 路由器的路由表中, 所包含信息的主要内容有目的网络地址、子网掩码、下一跳地址。

3. 无分类编址 CIDR

无分类域间路由选择 CIDR 是在变长子网掩码的基础上提出的一种消除传统 A、B、C 类网络划分, 并且可以在软件的支持下实现超网构造的一种 IP 地址的划分方法。

例如, 如果一个单位需要 2000 个地址, 那么就给它分配一个 2048 地址的块(8 个连续的 C 类网络), 而不是一个完全的 B 类地址。这样可以大幅度提高 IP 地址空间的利用率, 减小路由器的路由表大小, 提高路由转发能力。

CIDR 消除了传统 A、B、C 类地址及划分子网的概念, 因而可以更有效地分配 IPv4 的地址空间。CIDR 使用“网络前缀”的概念代替子网络的概念, 与传统分类 IP 地址最大的区别就是, 网络前缀的位数不是固定的, 可以任意选取。CIDR 的记法是:

$$\text{IP} ::= \{\langle \text{网络前缀} \rangle, \langle \text{主机号} \rangle\}.$$

CIDR 还使用“斜线记法”(或称 CIDR 记法), 即 IP 地址/网络前缀所占比特数。其中, 网络前缀所占比特数对应于网络号的部分, 等效于子网掩码中连续 1 的部分。例如, 对于 128.14.32.5/20 这个地址, 它的掩码是 20 个连续的 1 和后续 12 个连续的 0, 通过逐位相“与”的方法可以得到该地址的网络前缀(或直接截取前 20 位):

$$\begin{aligned} \text{逐位与} & \left\{ \begin{array}{l} \text{IP} = 10000000.00001110.00100000.00000101 \\ \text{掩码} = 11111111.11111111.11110000.00000000 \end{array} \right. \\ \text{网络前缀} & = 10000000.00001110.00100000.00000000 \quad (128.14.32.0) \end{aligned}$$

CIDR 虽然不使用子网, 但仍然使用“掩码”一词。“CIDR 不使用子网”是指 CIDR 并没有

在 32 位地址中指明若干位作为子网字段。但分配到一个 CIDR 地址块的组织，仍可以在本组织内根据需要划分出一些子网。例如，某组织分配到地址块/20，就可以再继续划分为 8 个子网（从主机号中借用 3 位来划分子网），这时每个子网的网络前缀就变成了 23 位。全 0 和全 1 的主机号地址一般不使用。

将网络前缀都相同的连续 IP 地址组成“CIDR 地址块”。一个 CIDR 地址块可以表示很多地址，这种地址的聚合称为路由聚合，或称构成超网。路由聚合使得路由表中的一个项目可以表示多个原来传统分类地址的路由，有利于减少路由器之间的信息的交换，从而提高网络性能。

例如，在如图 4.8 所示的网络中，如果不使用路由聚合，那么 R1 的路由表中需要分别有到网络 1 和网络 2 的路由表项。不难发现，网络 1 和网络 2 的网络前缀在二进制表示的情况下，前 16 位都是相同的，第 17 位分别是 0 和 1，并且从 R1 到网络 1 和网络 2 的路由的下一跳皆为 R2。若使用路由聚合，在 R1 看来，网络 1 和网络 2 可以构成一个更大的地址块 206.1.0.0/16，到网络 1 和网络 2 的两条路由就可以聚合成一条到 206.1.0.0/16 的路由。

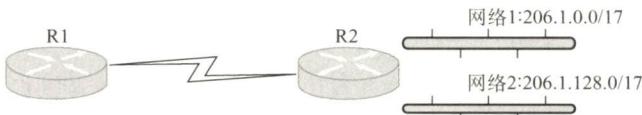


图 4.8 路由聚合的例子

CIDR 地址块中的地址数一定是 2 的整数次幂，实际可指派的地址数通常为 $2^N - 2$ ， N 表示主机号的位数，主机号全 0 代表网络号，主机号全 1 为广播地址。网络前缀越短，其地址块所包含的地址数就越多。而在三级结构的 IP 地址中，划分子网使网络前缀变长。

CIDR 的优点在于网络前缀长度的灵活性。由于上层网络的前缀长度较短，因此相应的路由表的项目较少。而内部又可采用延长网络前缀的方法来灵活地划分子网。

最长前缀匹配（最佳匹配）：使用 CIDR 时，路由表中的每个项目由“网络前缀”和“下一跳地址”组成。在查找路由表时可能会得到不止一个匹配结果。此时，应当从匹配结果中选择具有最长网络前缀的路由，因为网络前缀越长，其地址块就越小，因而路由就越具体。

CIDR 查找路由表的方法：为了更加有效地查找最长前缀匹配，通常将无分类编址的路由表存放在一种层次式数据结构中，然后自上而下地按层次进行查找。这里最常用的数据结构就是二叉线索。

4. 网络层转发分组的过程

分组转发都是基于目的主机所在网络的，这是因为互联网上的网络数远小于主机数，可以极大地压缩转发表的大小。当分组到达路由器后，路由器根据目的 IP 地址的网络前缀来查找转发表，确定下一跳应当到哪个路由器。因此，在转发表中，每条路由必须有下面两条信息：

(目的网络, 下一跳地址)

这样，IP 数据报最终一定可以找到目的主机所在目的网络上的路由器（可能要通过多次间接交付），当到达最后一个路由器时，才试图向目的主机进行直接交付。

采用 CIDR 编址时，如果一个分组在转发表中可以找到多个匹配的前缀，那么应当选择前缀最长的一个作为匹配的前缀，称为最长前缀匹配。网络前缀越长，其地址块就越小，因而路由就越精准。为了更快地查找转发表，可以按照前缀的长短，将前缀最长的排在第 1 行，按前缀长度的降序排列。这样，从第 1 行最长的开始查找，只要检索到匹配的，就不必再继续查找。

此外，转发表中还可以增加两种特殊的路由：

- 1) 主机路由：对特定目的主机的 IP 地址专门指明一个路由，以方便网络管理员控制和测试网络。若特定主机的 IP 地址是 a.b.c.d，则转发表中对应项的目的网络是 a.b.c.d/32。/32 表示的子网掩码没有意义，但这个特殊的前缀可以用在转发表中。
- 2) 默认路由：用特殊前缀 0.0.0.0/0 表示默认路由，全 0 掩码和任何目的地址进行按位与运算，结果必然为全 0，即必然和转发表中的 0.0.0.0/0 相匹配。只要目的网络是其他网络（不在转发表中），就一律选择默认路由。

综上所述，归纳出路由器执行的分组转发算法如下：

- 1) 从收到的 IP 分组的首部提取目的主机的 IP 地址 D（即目的地址）。
- 2) 若查找到特定主机路由（目的地址为 D），就按照这条路由的下一跳转发分组；否则从转发表中的下一条（即按前缀长度的顺序）开始检查，执行步骤 3)。
- 3) 将这一行的子网掩码与目的地址 D 进行按位与运算。若运算结果与本行的前缀匹配，则查找结束，按照“下一跳”指出的进行处理（或者直接交付本网络上的目的主机，或通过指定接口发送到下一跳路由器）。否则，若转发表还有下一行，则对下一行进行检查，重新执行步骤 3)。否则，执行步骤 4)。
- 4) 若转发表中有一个默认路由，则把分组传送给默认路由；否则，报告转发分组出错。

值得注意的是，转发表（或路由表）并未给分组指明到某个网络的完整路径（即先经过哪个路由器，然再经过哪个路由器等）。转发表指出，到某个网络应当先到某个路由器（即下一跳路由器），在到达下一跳路由器后，再继续查找其转发表，知道下一步应当到哪个路由器。这样一步一步地查找下去，直到最后到达目的网络。

注意：得到下一跳路由器的 IP 地址后，并不是直接将该地址填入待发送的数据报，而是将该 IP 地址转换成 MAC 地址（通过 ARP），将此 MAC 地址放到 MAC 帧首部中，然后根据这个 MAC 地址找到下一跳路由器。在不同网络中传送时，MAC 帧中的源地址和目的地址要发生变化，但是网桥在转发帧时，不改变帧的源地址，请注意区分。

4.3.4 ARP、DHCP 与 ICMP

1. IP 地址与硬件地址

IP 地址是网络层使用的地址，它是分层次等级的。硬件地址是数据链路层使用的地址（MAC 地址），它是平面式的。在网络层及网络层之上使用 IP 地址，IP 地址放在 IP 数据报的首部，而 MAC 地址放在 MAC 帧的首部。通过数据封装，把 IP 数据报分组封装为 MAC 帧后，数据链路层看不见数据报分组中的 IP 地址。

由于路由器的隔离，IP 网络中无法通过广播 MAC 地址来完成跨网络的寻址，因此在网络层只使用 IP 地址来完成寻址。寻址时，每个路由器依据其路由表（依靠路由协议生成）选择到目标网络（即主机号全为 0 的网络地址）需要转发到的下一跳（路由器的物理端口号或下一网络地址），而 IP 分组通过多次路由转发到达目标网络后，改为在目标 LAN 中通过数据链路层的 MAC 地址以广播方式寻址。这样可以提高路由选择的效率。

- 1) 在 IP 层抽象的互联网上只能看到 IP 数据报。
- 2) 虽然在 IP 数据报首部中有源 IP 地址，但路由器只根据目的 IP 地址进行转发。
- 3) 在局域网的链路层，只能看见 MAC 帧。IP 数据报被封装在 MAC 帧中，通过路由器转发 IP 分组时，IP 分组在每个网络中都被路由器解封装和重新封装，其 MAC 帧首部中的源地址和目的地址会不断改变。这也决定了无法使用 MAC 地址跨网络通信。

- 4) 尽管互连在一起的网络的硬件地址体系各不相同，但 IP 层抽象的互联网却屏蔽了下层这些复杂的细节。只要我们在网络层上讨论问题，就能够使用统一的、抽象的 IP 地址研究主机与主机或路由器之间的通信。

注意：路由器由于互连多个网络，因此它不仅有多个 IP 地址，也有多个硬件地址。

2. 地址解析协议（ARP）

无论网络层使用什么协议，在实际网络的链路上传送数据帧时，最终必须使用硬件地址。所以需要一种方法来完成 IP 地址到 MAC 地址的映射，这就是地址解析协议（Address Resolution Protocol, ARP）。每台主机都设有一个 ARP 高速缓存，用来存放本局域网上各主机和路由器的 IP 地址到 MAC 地址的映射表，称 ARP 表。使用 ARP 来动态维护此 ARP 表。

ARP 工作在网络层，其工作原理如下：主机 A 欲向本局域网上的某台主机 B 发送 IP 数据报时，先在其 ARP 高速缓存中查看有无主机 B 的 IP 地址。如果有，就可查出其对应的硬件地址，再将此硬件地址写入 MAC 帧，然后通过局域网将该 MAC 帧发往此硬件地址。如果没有，那么就通过使用目的 MAC 地址为 FFFF-FF-FF-FF-FF 的帧来封装并广播 ARP 请求分组（广播发送），使同一个局域网里的所有主机都收到此 ARP 请求。主机 B 收到该 ARP 请求后，向主机 A 发出 ARP 响应分组（单播发送），分组中包含主机 B 的 IP 与 MAC 地址的映射关系，主机 A 收到 ARP 响应分组后就将此映射写入 ARP 缓存，然后按查询到的硬件地址发送 MAC 帧。ARP 由于“看到了”IP 地址，所以它工作在网络层，而 NAT 路由器由于“看到了”端口，所以它工作在传输层。对于某个协议工作在哪个层次，读者应该能通过协议的工作原理进行猜测。

注意：ARP 用于解决同一个局域网上的主机或路由器的 IP 地址和硬件地址的映射问题。如果所要找的主机和源主机不在同一个局域网上，那么就要通过 ARP 找到一个位于本局域网上的某个路由器的硬件地址，然后把分组发送给这个路由器，让这个路由器把分组转发给下一个网络。剩下的工作就由下一个网络来做，尽管 ARP 请求分组是广播发送的，但 ARP 响应分组是普通的单播，即从一个源地址发送到一个目的地址。

使用 ARP 的 4 种典型情况总结如下（见图 4.9）：

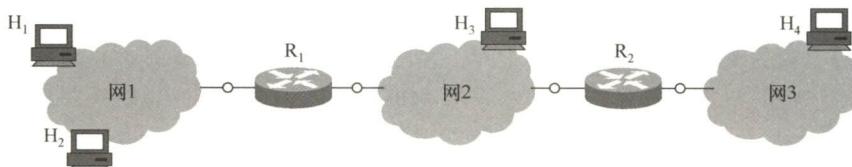


图 4.9 使用 ARP 的 4 种典型情况

- 1) 发送方是主机（如 H₁），要把 IP 数据报发送到本网络上的另一台主机（如 H₂）。这时 H₁ 在网 1 用 ARP 找到目的主机 H₂ 的硬件地址。
- 2) 发送方是主机（如 H₁），要把 IP 数据报发送到另一个网络上的一台主机（如 H₃）。这时 H₁ 用 ARP 找到与网 1 连接的路由器 R₁ 的硬件地址，剩下的工作由 R₁ 来完成。
- 3) 发送方是路由器（如 R₁），要把 IP 数据报转发到与 R₁ 连接的网络（网 2）上的一台主机（如 H₃）。这时 R₁ 在网 2 用 ARP 找到目的主机 H₃ 的硬件地址。
- 4) 发送方是路由器（如 R₁），要把 IP 数据报转发到网 3 上的一台主机（如 H₄）。这时 R₁ 在网 2 用 ARP 找到与网 2 连接的路由器 R₂ 的硬件地址，剩下的工作由 R₂ 来完成。

从 IP 地址到硬件地址的解析是自动进行的，主机的用户并不知道这种地址解析过程。只要主机或路由器和本网络上的另一个已知 IP 地址的主机或路由器进行通信，ARP 就会自动地将这个 IP 地址解析为数据链路层所需要的硬件地址。

3. 动态主机配置协议 (DHCP)

动态主机配置协议 (Dynamic Host Configuration Protocol, DHCP) 常用于给主机动态地分配 IP 地址，它提供了即插即用的联网机制，这种机制允许一台计算机加入新的网络和获取 IP 地址而不用手工参与。DHCP 是应用层协议，它是基于 UDP 的。

DHCP 的工作原理如下：使用客户/服务器模式。需要 IP 地址的主机在启动时就向 DHCP 服务器广播发送发现报文，这时该主机就成为 DHCP 客户。本地网络上所有主机都能收到此广播报文，但只有 DHCP 服务器才回答此广播报文。DHCP 服务器先在其数据库中查找该计算机的配置信息。若找到，则返回找到的信息。若找不到，则从服务器的 IP 地址池中取一个地址分配给该计算机。DHCP 服务器的回答报文称为提供报文。

DHCP 服务器和 DHCP 客户端的交换过程如下：

- 1) DHCP 客户机广播“DHCP 发现”消息，试图找到网络中的 DHCP 服务器，以便从 DHCP 服务器获得一个 IP 地址。源地址为 0.0.0.0，目的地址为 255.255.255.255。
- 2) DHCP 服务器收到“DHCP 发现”消息后，广播“DHCP 提供”消息，其中包括提供给 DHCP 客户机的 IP 地址。源地址为 DHCP 服务器地址，目的地址为 255.255.255.255。
- 3) DHCP 客户机收到“DHCP 提供”消息，如果接受该 IP 地址，那么就广播“DHCP 请求”消息向 DHCP 服务器请求提供 IP 地址。源地址为 0.0.0.0，目的地址为 255.255.255.255。
- 4) DHCP 服务器广播“DHCP 确认”消息，将 IP 地址分配给 DHCP 客户机。源地址为 DHCP 服务器地址，目的地址为 255.255.255.255。

DHCP 允许网络上配置多台 DHCP 服务器，当 DHCP 客户机发出“DHCP 发现”消息时，有可能收到多个应答消息。这时，DHCP 客户机只会挑选其中的一个，通常挑选最先到达的。

DHCP 服务器分配给 DHCP 客户的 IP 地址是临时的，因此 DHCP 客户只能在一段有限的时间内使用这个分配到的 IP 地址。DHCP 称这段时间为租用期。租用期的数值应由 DHCP 服务器自己决定，DHCP 客户也可在自己发送的报文中提出对租用期的要求。

DHCP 的客户端和服务端需要通过广播方式来进行交互，原因是在 DHCP 执行初期，客户端不知道服务器端的 IP 地址，而在执行中间，客户端并未被分配 IP 地址，从而导致两者之间的通信必须采用广播的方式。采用 UDP 而不采用 TCP 的原因也很明显：TCP 需要建立连接，如果连对方的 IP 地址都不知道，那么更不可能通过双方的套接字建立连接。

DHCP 是应用层协议，因为它是通过客户/服务器模式工作的，DHCP 客户端向 DHCP 服务器请求服务，而其他层次的协议是没有这两种工作方式的。

4. 网际控制报文协议 (ICMP)

为了提高 IP 数据报交付成功的机会，在网络层使用了网际控制报文协议 (Internet Control Message Protocol, ICMP) 来让主机或路由器报告差错和异常情况。ICMP 报文作为 IP 层数据报的数据，加上数据报的首部，组成 IP 数据报发送出去。ICMP 是网络层协议。

ICMP 报文的种类有两种，即 ICMP 差错报告报文和 ICMP 询问报文。

ICMP 差错报告报文用于目标主机或到目标主机路径上的路由器向源主机报告差错和异常情况。共有以下 5 种常用的类型：

- 1) 终点不可达。当路由器或主机不能交付数据报时，就向源点发送终点不可达报文。

- 2) 源点抑制^②。当路由器或主机由于拥塞而丢弃数据报时，就向源点发送源点抑制报文，使源点知道应当把数据报的发送速率放慢。
- 3) 时间超过。当路由器收到生存时间（TTL）为零的数据报时，除丢弃该数据报外，还要向源点发送时间超过报文。当终点在预先规定的时间内不能收到一个数据报的全部数据报片时，就把已收到的数据报片都丢弃，并向源点发送时间超过报文。
- 4) 参数问题。当路由器或目的主机收到的数据报的首部中有的字段的值不正确时，就丢弃该数据报，并向源点发送参数问题报文。
- 5) 改变路由（重定向）。路由器把改变路由报文发送给主机，让主机知道下次应将数据报发给另外的路由器（可通过更好的路由）。

不应发送 ICMP 差错报告报文的几种情况如下：

- 1) 对 ICMP 差错报告报文不再发送 ICMP 差错报告报文。
- 2) 对第一个分片的数据报片的所有后续数据报片都不发送 ICMP 差错报告报文。
- 3) 对具有组播地址的数据报都不发送 ICMP 差错报告报文。
- 4) 对具有特殊地址（如 127.0.0.0 或 0.0.0.0）的数据报不发送 ICMP 差错报告报文。

ICMP 询问报文有 4 种类型：回送请求和回答报文、时间戳请求和回答报文、地址掩码请求和回答报文、路由器询问和通告报文，最常用的是前两类。

ICMP 的两个常见应用是分组网间探测 PING（用来测试两台主机之间的连通性）和 Traceroute（UNIX 中的名字，在 Windows 中是 Tracert，可以用来跟踪分组经过的路由）。其中 PING 使用了 ICMP 回送请求和回答报文，Traceroute（Tracert）使用了 ICMP 时间超过报文。

注意：PING 工作在应用层，它直接使用网络层的 ICMP，而未使用传输层的 TCP 或 UDP。
Traceroute/Tracert 工作在网络层。

关注公众号【乘龙考研】
一手更新 稳定有保障

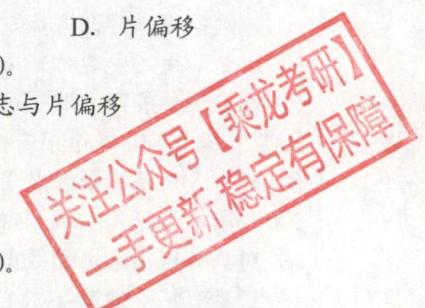
4.3.5 本节习题精选

一、单项选择题

01. Internet 的网络层含有 4 个重要的协议，分别为（ ）。
 - A. IP, ICMP, ARP, UDP
 - B. TCP, ICMP, UDP, ARP
 - C. IP, ICMP, ARP, RARP
 - D. UDP, IP, ICMP, RARP
02. 以下关于 IP 分组结构的描述中，错误的是（ ）。
 - A. IPv4 分组头的长度是可变的
 - B. 协议字段表示 IP 的版本，值为 4 表示 IPv4
 - C. 分组头长度字段以 4B 为单位，总长度字段以字节为单位
 - D. 生存时间字段值表示一个分组可以经过的最多的跳数
03. IPv4 分组首部中有两个有关长度的字段：首部长度和总长度，其中（ ）。
 - A. 首部长度字段和总长度字段都以 8bit 为计数单位
 - B. 首部长度字段以 8bit 为计数单位，总长度字段以 32bit 为计数单位
 - C. 首部长度字段以 32bit 为计数单位，总长度字段以 8bit 为计数单位
 - D. 首部长度字段和总长度字段都以 32bit 为计数单位
04. IP 分组中的检验字段检查范围是（ ）。
 - A. 整个 IP 分组
 - B. 仅检查分组首部

② 最新的 ICMP 标准[RFC6633]已不再使用“源点抑制报文”。

- C. 仅检查数据部分 D. 以上皆检查
05. 当数据报到达目的网络后，要传送到目的主机，需要知道 IP 地址对应的（ ）。
 A. 逻辑地址 B. 动态地址 C. 域名 D. 物理地址
06. 如果 IPv4 的分组太大，会在传输中被分片，那么在（ ）将对分片后的数据报重组。
 A. 中间路由器 B. 下一跳路由器 C. 核心路由器 D. 目的主机
07. 在 IP 首部的字段中，与分片和重组无关的字段是（ ）。
 A. 总长度 B. 标识 C. 标志 D. 片偏移
08. 以下关于 IP 分组的分片与组装的描述中，错误的是（ ）。
 A. IP 分组头中与分片和组装相关的字段是：标识、标志与片偏移
 B. IP 分组规定的最大长度为 65535B
 C. 以太网的 MTU 为 1500B
 D. 片偏移的单位是 4B
09. 以下关于 IP 分组分片基本方法的描述中，错误的是（ ）。
 A. IP 分组长度大于 MTU 时，就必须对其进行分片
 B. DF=1，分组的长度又超过 MTU 时，则丢弃该分组，不需要向源主机报告
 C. 分片的 MF 值为 1 表示接收到的分片不是最后一个分片
 D. 属于同一原始 IP 分组的分片具有相同的标识
10. 路由器 R0 的路由表见右表。若进入路由器 R0 的分组的目的地址为 132.19.237.5，该分组应该被转发到（ ）
 下一跳路由器。
 A. R1 B. R2 C. R3 D. R4
- | 目的网络 | 下一跳 |
|-----------------|-----|
| 132.0.0.0/8 | R1 |
| 132.0.0.0/11 | R2 |
| 132.19.232.0/22 | R3 |
| 0.0.0.0/0 | R4 |
11. IP 规定每个 C 类网络最多可以有（ ）台主机或路由器。
 A. 254 B. 256 C. 32 D. 1024
12. 下列地址中，属于子网 86.32.0.0/12 的地址是（ ）。
 A. 86.33.224.123 B. 86.79.65.126 C. 86.79.65.216 D. 86.68.206.154
13. 下列地址中，属于单播地址的是（ ）。
 A. 172.31.128.255/18 B. 10.255.255.255
 C. 192.168.24.59/30 D. 224.105.5.211
14. 下列地址中，属于本地回路地址的是（ ）。
 A. 10.10.10.1 B. 255.255.255.0 C. 192.0.0.1 D. 127.0.0.1
15. 访问因特网的每台主机都需要分配 IP 地址（假定采用默认子网掩码），下列可以分配给主机的 IP 地址是（ ）。
 A. 192.46.10.0 B. 110.47.10.0 C. 127.10.10.17 D. 211.60.256.21
16. 为了提供更多的子网，为一个 B 类地址指定了子网掩码 255.255.240.0，则每个子网最多可以有的主机数是（ ）。
 A. 16 B. 256 C. 4094 D. 4096
17. 不考虑 NAT，在 Internet 中，IP 数据报从源结点到目的结点可能需要经过多个网络和路由器。在整个传输过程中，IP 数据报头部中的（ ）。
 A. 源地址和目的地址都不会发生变化
 B. 源地址有可能发生变化而目的地址不会发生变化





- C. 源地址不会发生变化而目的地址有可能发生变化
D. 源地址和目的地址都有可能发生变化
18. 把 IP 网络划分成子网，这样做的好处是（ ）。
 A. 增加冲突域的大小 B. 增加主机的数量
 C. 减少广播域的大小 D. 增加网络的数量
19. 一个网段的网络号为 198.90.10.0/27，最多可以分成（ ）个子网，每个子网最多具有（ ）个有效的 IP 地址。
 A. 8, 30 B. 4, 62 C. 16, 14 D. 32, 6
20. 一台主机有两个 IP 地址，一个地址是 192.168.11.25，另一个地址可能是（ ）。
 A. 192.168.11.0 B. 192.168.11.26 C. 192.168.13.25 D. 192.168.11.24
21. CIDR 技术的作用是（ ）。
 A. 把小的网络汇聚成大的超网 B. 把大的网络划分成小的子网
 C. 解决地址资源不足的问题 D. 由多台主机共享同一个网络地址
22. CIDR 地址块 192.168.10.0/20 所包含的 IP 地址范围是（①）。与地址 192.16.0.19/28 同属于一个子网的主机地址是（②）。
 ① A. 192.168.0.0 ~ 192.168.12.255 B. 192.168.10.0 ~ 192.168.13.255
 C. 192.168.10.0 ~ 192.168.14.255 D. 192.168.0.0 ~ 192.168.15.255
 ② A. 192.16.0.17 B. 192.16.0.31 C. 192.16.0.15 D. 192.16.0.14
23. 路由表错误和软件故障都可能使得网络中的数据形成传输环路而无限转发环路的分组，IPv4 协议解决该问题的方法是（ ）。
 A. 报文分片 B. 设定生命期
 C. 增加校验和 D. 增加选项字段
24. 为了解决 IP 地址耗尽的问题，可以采用以下一些措施，其中治本的是（ ）。
 A. 划分子网 B. 采用无类比编址 CIDR
 C. 采用网络地址转换 NAT D. 采用 IPv6
25. 下列对 IP 分组的分片和重组的描述中，正确的是（ ）。
 A. IP 分组可以被源主机分片，并在中间路由器进行重组
 B. IP 分组可以被路径中的路由器分片，并在目的主机进行重组
 C. IP 分组可以被路径中的路由器分片，并在中间路由器上进行重组
 D. IP 分组可以被路径中的路由器分片，并在最后一跳的路由器上进行重组
26. 一个网络中有几个子网，其中一个已分配了子网号 74.178.247.96/29，则下列网络前缀中不能再分配给其他子网的是（ ）。
 A. 74.178.247.120/29 B. 74.178.247.64/29
 C. 74.178.247.96/28 D. 74.178.247.104/29
27. 主机 A 和主机 B 的 IP 地址分别为 216.12.31.20 和 216.13.32.21，要想让 A 和 B 工作在同一个 IP 子网内，应该给它们分配的子网掩码是（ ）。
 A. 255.255.255.0 B. 255.255.0.0
 C. 255.255.255.255 D. 255.0.0.0
28. 某单位分配了 1 个 B 类地址，计划将内部网络划分成 35 个子网，将来可能增加 16 个子网，每个子网的主机数目接近 800 台，则可行的掩码方案是（ ）。
 A. 255.255.248.0 B. 255.255.252.0

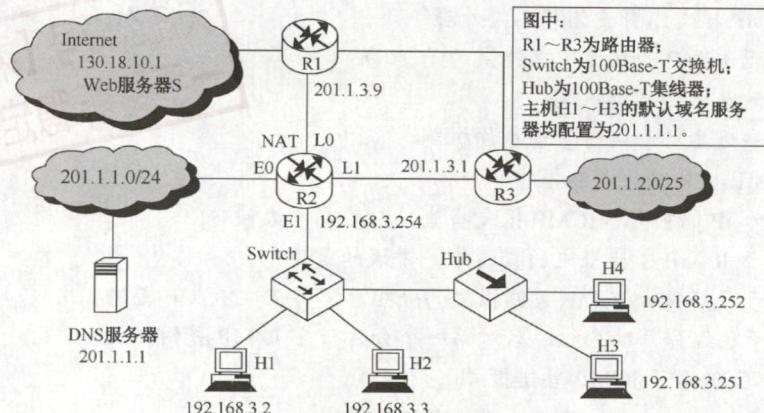
- B. ICMP 报文用于报告 IP 数据报发送错误
C. ICMP 报文封装在 IP 数据报中发送
D. ICMP 报文本身出错将不再处理
39. 以下关于 ICMP 的描述中，错误的是（ ）。
A. IP 缺乏差错控制机制
B. IP 缺乏主机和网络管理查询机制
C. ICMP 报文分为差错报告和查询两类
D. 作为 IP 的补充，ICMP 报文将直接封装在以太帧中
40. 以下关于 ICMP 差错报文的描述中，错误的是（ ）。
A. 对于已经携带 ICMP 差错报文的分组，不再产生 ICMP 差错报文
B. 对于已经分片的分组，只对第一个分片产生 ICMP 差错报文
C. PING 使用了 ICMP 差错报文
D. 对于组播的分组，不产生 ICMP 差错报文
41. 【2010 统考真题】某网络的 IP 地址空间为 192.168.5.0/24，采用定长子网划分，子网掩码为 255.255.255.248，则该网络中的最大子网个数、每个子网内的最大可分配地址个数分别是（ ）。
A. 32, 8 B. 32, 6 C. 8, 32 D. 8, 30
42. 【2010 统考真题】若路由器 R 因为拥塞丢弃 IP 分组，则此时 R 可向发出该 IP 分组的源主机发送的 ICMP 报文类型是（ ）。
A. 路由重定向 B. 目的不可达 C. 源点抑制 D. 超时
43. 【2011 统考真题】在子网 192.168.4.0/30 中，能接收目的地址为 192.168.4.3 的 IP 分组的最大主机数是（ ）。
A. 0 B. 1 C. 2 D. 4
44. 【2012 统考真题】某主机的 IP 地址为 180.80.77.55，子网掩码为 255.255.252.0。若该主机向其所在子网发送广播分组，则目的地址可以是（ ）。
A. 180.80.76.0 B. 180.80.76.255 C. 180.80.77.255 D. 180.80.79.255
45. 【2012 统考真题】ARP 的功能是（ ）。
A. 根据 IP 地址查询 MAC 地址 B. 根据 MAC 地址查询 IP 地址
C. 根据域名查询 IP 地址 D. 根据 IP 地址查询域名
46. 【2012 统考真题】在 TCP/IP 体系结构中，直接为 ICMP 提供服务的协议是（ ）。
A. PPP B. IP C. UDP D. TCP
47. 【2015 统考真题】某路由器的路由表如下所示：

目的网络	下一跳	接口
169.96.40.0/23	176.1.1.1	S1
169.96.40.0/25	176.2.2.2	S2
169.96.40.0/27	176.3.3.3	S3
0.0.0.0/0	176.4.4.4	S4

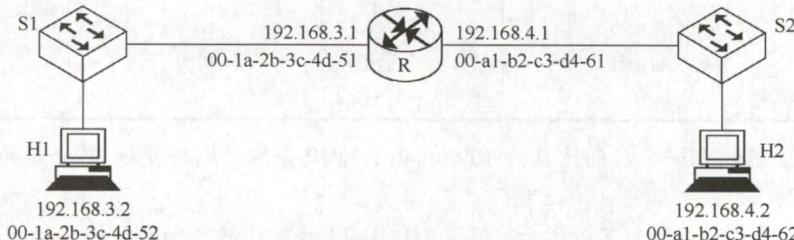
- 若路由器收到一个目的地址为 169.96.40.5 的 IP 分组，则转发该 IP 分组的接口是（ ）。
A. S1 B. S2 C. S3 D. S4
48. 【2016 统考真题】如下图所示，假设 H1 与 H2 的默认网关和子网掩码均分别配置为 192.

关注公众号【乘龙考研】
一手更新 稳定有保障

168.3.1 和 255.255.255.128, H3 和 H4 的默认网关和子网掩码均分别配置为 192.168.3.254 和 255.255.255.128, 则下列现象中可能发生的是 ()。



- A. H1 不能与 H2 进行正常 IP 通信 B. H2 与 H4 均不能访问 Internet
 C. H1 不能与 H3 进行正常 IP 通信 D. H3 不能与 H4 进行正常 IP 通信
49. 【2016 统考真题】在题 48 图中，假设连接 R1、R2 和 R3 之间的点对点链路使用地址 201.1.3.x/30, 当 H3 访问 Web 服务器 S 时, R2 转发出去的封装 HTTP 请求报文的 IP 分组是源 IP 地址和目的 IP 地址, 它们分别是 ()。
 A. 192.168.3.251, 130.18.10.1 B. 192.168.3.251, 201.1.3.9
 C. 201.1.3.8, 130.18.10.1 D. 201.1.3.10, 130.18.10.1
50. 【2017 统考真题】若将网络 21.3.0.0/16 划分为 128 个规模相同的子网, 则每个子网可分配的最大 IP 地址个数是 ()。
 A. 254 B. 256 C. 510 D. 512
51. 【2017 统考真题】下列 IP 地址中, 只能作为 IP 分组的源 IP 地址但不能作为目的 IP 地址的是 ()。
 A. 0.0.0.0 B. 127.0.0.1 C. 200.10.10.3 D. 255.255.255.255
52. 【2018 统考真题】某路由表中有转发接口相同的 4 条路由表项, 其目的网络地址分别为 35.230.32.0/21、35.230.40.0/21、35.230.48.0/21 和 35.230.56.0/21, 将该 4 条路由聚合后的目的网络地址为 ()。
 A. 35.230.0.0/19 B. 35.230.0.0/20 C. 35.230.32.0/19 D. 35.230.32.0/20
53. 【2018 统考真题】路由器 R 通过以太网交换机 S1 和 S2 连接两个网络, R 的接口、主机 H1 和 H2 的 IP 地址与 MAC 地址如下图所示。若 H1 向 H2 发送一个 IP 分组 P, 则 H1 发出的封装 P 的以太网帧的目的 MAC 地址、H2 收到的封装 P 的以太网帧的源 MAC 地址分别是 ()。



- A. 00-a1-b2-c3-d4-62, 00-1a-2b-3c-4d-52 B. 00-a1-b2-c3-d4-62, 00-a1-b2-c3-d4-61
 C. 00-1a-2b-3c-4d-51, 00-1a-2b-3c-4d-52 D. 00-1a-2b-3c-4d-51, 00-a1-b2-c3-d4-61
54. 【2019 统考真题】若将 101.200.16.0/20 划分为 5 个子网，则可能的最小子网的可分配 IP 地址数是（ ）。
 A. 126 B. 254 C. 510 D. 1022
55. 【2021 统考真题】现将一个 IP 网络划分为 3 个子网，若其中一个子网是 192.168.9.128/26，则下列网络中，不可能是另外两个子网之一的是（ ）。
 A. 192.168.9.0/25 B. 192.168.9.0/26
 C. 192.168.9.192/26 D. 192.168.9.192/27
56. 【2021 统考真题】若路由器向 MTU=800B 的链路转发一个总长度为 1580B 的 IP 数据报（首部长度为 20B）时，进行了分片，且每个分片尽可能大，则第 2 个分片的总长度字段和 MF 标志位的值分别是（ ）。
 A. 796, 0 B. 796, 1 C. 800, 0 D. 800, 1
57. 【2022 统考真题】若某主机的 IP 地址是 183.80.72.48，子网掩码是 255.255.192.0，则该主机所在网络的网络地址是（ ）。
 A. 183.80.0.0 B. 183.80.64.0 C. 183.80.72.0 D. 183.80.192.0
58. 【2022 统考真题】下图所示网络中的主机 H 的子网掩码与默认网关分别是（ ）。



关注公众号【乘龙考研】
一手更新 稳定有保障

- A. 255.255.192.192, 192.168.1.1 B. 255.255.192.192, 192.168.1.62
 C. 255.255.224, 192.168.1.1 D. 255.255.224, 192.168.1.62

二、综合应用题

01. 一个 IP 分组报头中的首部长度字段值为 101（二进制），而总长度字段值为 10100（二进制）。请问该分组携带了多少字节的数据？
02. 一个数据报长度为 4000B（固定头部长度）。现在经过一个网络传送，但此网络能够传送的最大数据长度为 1500B。试问应当划分为几个短一些的数据报片？各数据片段的数据字段长度、片段偏移字段和 MF 标志应为何值？
03. 某网络的一台主机产生了一个 IP 数据报，头部长度为 20B，数据部分长度为 2000B。该数据报需要经过两个网络到达目的主机，这两个网络所允许的最大传输单位（MTU）分别为 1500B 和 576B。问原 IP 数据报到达目的主机时分成了几个 IP 小报文？每个报文的数据部分长度分别是多少？
04. 如果到达的分组的片偏移值为 100，分组首部中的首部长度字段值为 5，总长度字段值为 100，那么数据部分第一个字节的编号是多少？能够确定数据部分最后一个字节的编号吗？
05. 设目的地址为 201.230.34.56，子网掩码为 255.255.240.0，试求子网地址。
06. 在 4 个“/24”地址块中进行最大可能的聚合：212.56.132.0/24、212.56.133.0/24、212.56.134.0/24、

212.56.135.0/24。

07. 现有一公司需要创建内部网络，该公司包括工程技术部、市场部、财务部和办公室 4 个部门，每个部门有 20~30 台计算机。试问：
- 1) 若要将几个部门从网络上分开，如果分配给该公司使用的地址为一个 C 类地址，网络地址为 192.168.161.0，那么如何划分网络？可以将几个部门分开？
 - 2) 确定各部门的网络地址和子网掩码，并写出分配给每个部门网络中的主机 IP 地址范围。
08. 某路由器具有右表所示的路由表项。
- 1) 假设路由器收到两个分组：分组 A 的目的地址为 131.128.55.33，分组 B 的目的地址为 131.128.55.38。确定路由器为这两个分组选择的下一跳，并加以说明。
 - 2) 在路由表中增加一个路由表项，它使以 131.128.55.33 为目的地址的 IP 分组选择“A”作为下一跳，而不影响其他目的地址的 IP 分组的转发。
 - 3) 在路由表中增加一个路由表项，使所有目的地址与该路由表中任何路由表项都不匹配的 IP 分组被转发到下一跳“E”。
 - 4) 将 131.128.56.0/24 划分为 4 个规模尽可能大的等长子网，给出子网掩码及每个子网的可分配地址范围。

网络前缀	下一跳
131.128.56.0/24	A
131.128.55.32/28	B
131.128.55.32/30	C
131.128.0.0/16	D

09. 下表是使用无类别域间路由选择（CIDR）的路由选择表，地址字段是用十六进制表示的，试指出具有下列目标地址的 IP 分组将被投递到哪个下一站？

网络/掩码长度	下一站地	C4.68.0.0/14	D
C4.50.0.0/12	A	80.0.0.0/1	E
C4.5E.10.0/20	B	40.0.0.0/2	F
C4.60.0.0/12	C	00.0.0.0/2	G

1) C4.5E.13.87 2) C4.5E.22.09 3) C3.41.80.00 4) 5E.43.91.12

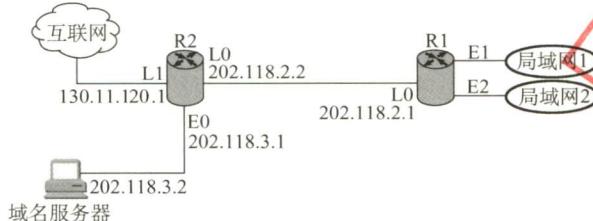
10. 一个自治系统有 5 个局域网，如下图所示，LAN2 至 LAN5 上的主机数分别为 91、150、3 和 15，该自治系统分配到的 IP 地址块为 30.138.118.0/23，试给出每个局域网的地址块（包括前缀）。



11. 某个网络地址块 192.168.75.0 中有 5 台主机 A、B、C、D 和 E，主机 A 的 IP 地址为 192.168.75.18，主机 B 的 IP 地址为 192.168.75.146，主机 C 的 IP 地址为 192.168.75.158，主机 D 的 IP 地址为 192.168.75.161，主机 E 的 IP 地址为 192.168.75.173，共同的子网掩码是 255.255.255.240。请回答：

- 1) 5 台主机 A、B、C、D、E 分属几个网段？哪些主机位于同一网段？主机 D 的网络地址为多少？
- 2) 若要加入第 6 台主机 F，使它能与主机 A 属于同一网段，其 IP 地址范围是多少？
- 3) 若在网络中另加入一台主机，其 IP 地址为 192.168.75.164，它的广播地址是多少？哪些主机能够收到？

12. 【2009 统考真题】某网络拓扑图如下图所示，路由器 R1 通过接口 E1、E2 分别连接局域网 1、局域网 2，通过接口 L0 连接路由器 R2，并通过路由器 R2 连接域名服务器与互联网。R1 的 L0 接口的 IP 地址是 202.118.2.1；R2 的 L0 接口的 IP 地址是 202.118.2.2，L1 接口的 IP 地址是 130.11.120.1，E0 接口的 IP 地址是 202.118.3.1；域名服务器的 IP 地址是 202.118.3.2。



关注公众号【乘龙考研】
一手更新稳定有保障

R1 和 R2 的路由表结构如下：

目的网络 IP 地址	子网掩码	下一跳 IP 地址	接口
------------	------	-----------	----

- 将 IP 地址空间 202.118.1.0/24 划分为两个子网，分别分配给局域网 1 和局域网 2，每个局域网需分配的 IP 地址数不少于 120 个。请给出子网划分结果，说明理由或给出必要的计算过程。
 - 请给出 R1 的路由表，使其明确包括到局域网 1 的路由、局域网 2 的路由、域名服务器的主机路由和互联网的路由。
 - 请采用路由聚合技术，给出 R2 到局域网 1 和局域网 2 的路由。
13. 一个 IPv4 分组到达一个结点时，其首部信息（以十六进制表示）为：0x45 00 00 54 00 03 58 50 20 06 FF F0 7C 4E 03 02 B4 0E 0F 02。请回答：
- 分组的源 IP 地址和目的 IP 地址各是什么（点分十进制表示法）？
 - 该分组数据部分的长度是多少？
 - 该分组是否已经分片？如果有分片，那么偏移量是多少？
14. 主机 A 的 IP 地址为 218.207.61.211，MAC 地址为 00:1d:72:98:1d:fc。A 收到一个帧，该帧的前 64 个字节的十六进制形式和 ASCII 形式如下图所示。

0000	00	1d	72	98	1d	fc	00	00	5e	00	01	01	88	64	11	00	..r.....	^....d..
0010	75	89	01	92	00	21	45	00	01	90	f9	bf	40	00	33	06	u....!E.	...@.3.
0020	f3	15	da	c7	66	28	da	cf	3d	d3	00	50	c4	8f	dc	a6f(..	=..P....
0030	a2	96	23	4c	44	69	50	18	00	0f	76	3d	00	00	90	b5	#LDiP.	..v=....

IP 分组首部如右图所示。

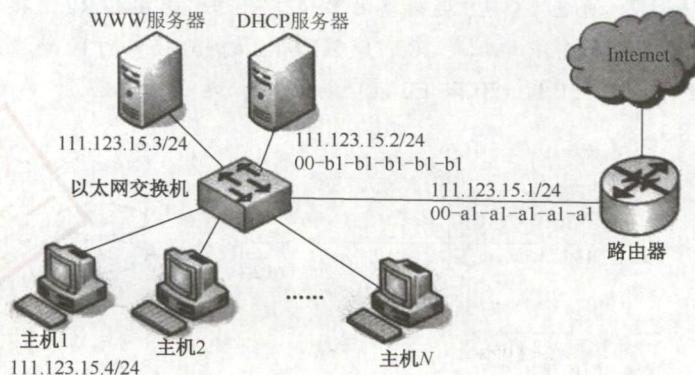
以太网帧的结构参见第 3 章。问：

- 主机 A 所在网络的网关路由器的相应端口的 MAC 地址是多少？
- 该 IP 分组所携带的数据量为多少字节？
- 若该分组需要被路由器转发到一条 MTU 为 380B 的链路上，则路由器将做何种操作？

版本	首部 长度	服务类型	总长度
标识		标志	片偏移
生存时间(TTL)	协议	首部校验和	
		源IP地址	
		目的IP地址	

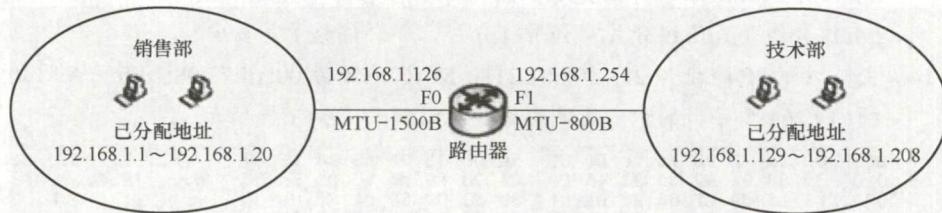
15. 【2015 统考真题】某网络拓扑如下图所示，其中路由器内网接口、DHCP 服务器、WWW 服务器与主机 1 均采用静态 IP 地址配置，相关地址信息见图中标注；主机 2~主机 N 通

过 DHCP 服务器动态获取 IP 地址等配置信息。



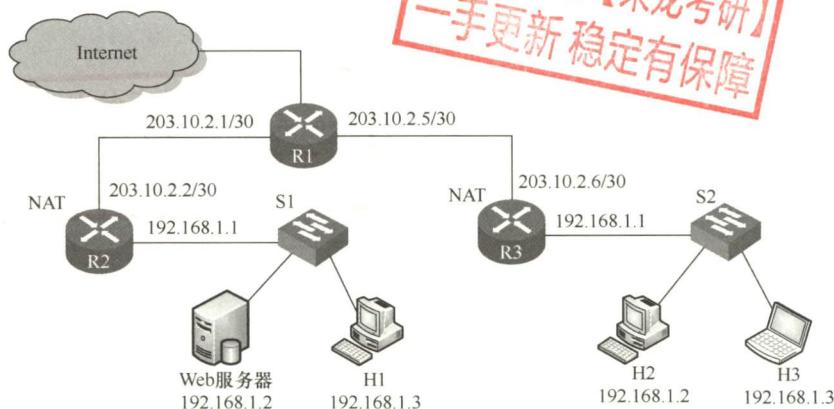
回答下列问题：

- 1) DHCP 服务器可为主机 2 ~ N 动态分配 IP 地址的最大范围是什么？主机 2 使用 DHCP 获取 IP 地址的过程中，发送的封装 DHCP Discover 报文的 IP 分组的源 IP 地址和目的 IP 地址分别是多少？
 - 2) 若主机 2 的 ARP 表为空，则该主机访问 Internet 时，发出的第一个以太网帧的目的 MAC 地址是什么？封装主机 2 发往 Internet 的 IP 分组的以太网帧的目的 MAC 地址是什么？
 - 3) 若主机 1 的子网掩码和默认网关分别配置为 255.255.255.0 和 111.123.15.2，则该主机是否能访问 WWW 服务器？是否能访问 Internet？请说明理由。
16. 【2018 统考真题】某公司的网络如下图所示。IP 地址空间 192.168.1.0/24 均分给销售部和技术部两个子网，并已分别为部分主机和路由器接口分配了 IP 地址，销售部子网的 MTU=1500B，技术部子网的 MTU=800B。



回答下列问题：

- 1) 销售部子网的广播地址是什么？技术部子网的子网地址是什么？若每台主机仅分配一个 IP 地址，则技术部子网还可以连接多少台主机？
 - 2) 假设主机 192.168.1.1 向主机 192.168.1.208 发送一个总长度为 1500B 的 IP 分组，IP 分组的头部长度为 20B，路由器在通过接口 F1 转发该 IP 分组时进行了分片。若分片时尽可能分为最大片，则一个最大 IP 分片封装数据的字节数是多少？至少需要分为几个分片？每个分片的片偏移量是多少？
17. 【2020 统考真题】某校园网有两个局域网，通过路由器 R1、R2 和 R3 互连后接入 Internet，S1 和 S2 为以太网交换机。局域网采用静态 IP 地址配置，路由器部分接口以及各主机的 IP 地址如下图所示。



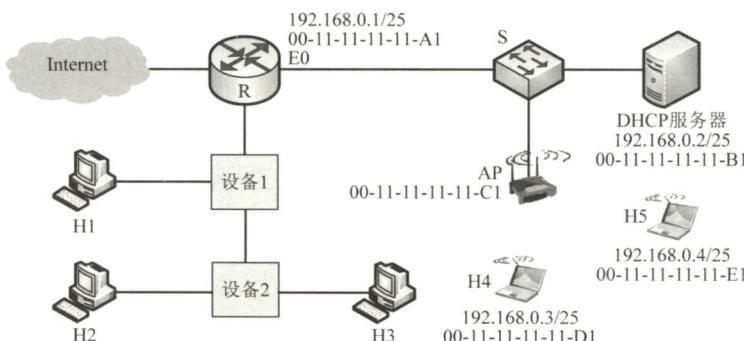
假设 NAT 转换表结构为

外网		内网	
IP 地址	端口号	IP 地址	端口号

请回答下列问题：

- 1) 为使 H2 和 H3 能够访问 Web 服务器（使用默认端口号），需要进行什么配置？
- 2) 若 H2 主动访问 Web 服务器时，将 HTTP 请求报文封装到 IP 数据报 P 中发送，则 H2 发送的 P 的源 IP 地址和目的 IP 地址分别是什么？经过 R3 转发后，P 的源 IP 地址和目的 IP 地址分别是什么？经过 R2 转发后，P 的源 IP 地址和目的 IP 地址分别是什么？

18. 【2022 统考真题】某网络拓扑如下图所示，R 为路由器，S 为以太网交换机，AP 是 802.11 接入点，路由器的 E0 接口和 DHCP 服务器的 IP 地址配置如图中所示；H1 与 H2 属于同一个广播域，但不属于同一个冲突域；H2 和 H3 属于同一个冲突域；H4 和 H5 已经接入网络，并通过 DHCP 动态获取了 IP 地址。现有路由器、100BaseT 以太网交换机和 100BaseT 集线器（Hub）三类设备各若干。



请回答下列问题。

- 1) 设备 1 和设备 2 应该分别选择哪类设备？
- 2) 若信号传播速度为 2×10^8 m/s，以太网最小帧长为 64B，信号通过设备 2 时会产生额外的 $1.51\mu s$ 的时间延迟，则 H2 与 H3 之间可以相距的最远距离是多少？
- 3) 在 H4 通过 DHCP 动态获取 IP 地址过程中，H4 首先发送了 DHCP 报文 M，M 是哪种 DHCP 报文？路由器 E0 接口能否收到封装 M 的以太网帧？S 向 DHCP 服务器转发的封装 M 的以太网帧的目的 MAC 地址是什么？

- 4) 若 H4 向 H5 发送一个 IP 分组 P, 则 H5 收到的封装 P 的 802.11 帧的地址 1、地址 2 和地址 3 分别是什么?

4.3.6 答案与解析

一、单项选择题

关注公众号【乘龙考研】
一手更新 稳定有保障

01. C

TCP 和 UDP 是传输层协议, IP、ICMP、ARP、RARP(逆地址解析协议)是网络层协议。

02. B

协议字段表示使用 IP 的上层协议, 如值为 6 表示 TCP, 值为 17 表示 UDP。版本字段表示 IP 的版本, 值为 4 表示 IPv4, 值为 6 表示 IPv6。

03. C

在首部中有三个关于长度的标记: 首部长度、总长度和片偏移, 基本单位分别为 4B、1B 和 8B。IP 分组的首部长度必须是 4B 的整数倍, 取值范围是 5~15(默认值是 5)。由于 IP 分组的首部长度是可变的, 因此首部长度字段必不可少。总长度字段给出 IP 分组的总长度, 单位是字节, 包括分组首部和数据部分的长度。数据部分的长度可以从总长度减去分组首部长度计算。

04. B

IP 分组的校验字段仅检查分组的首部信息, 不包括数据部分。

05. D

在数据链路层, MAC 地址用来标识主机或路由器, 数据报到达具体的目的网络后, 需要知道目的主机的 MAC 地址才能成功送达, 因此需要将 IP 地址转换成对应的 MAC 地址, 即物理地址。

06. D

数据报被分片后, 每个分片都将独立地传输到目的地, 其间有可能会经过不同的路径, 而最后在目的端主机分组才能被重组。

07. A

在 IP 首部中, 标识字段的用途是让目标机器确认一个新到达的分片是否属于同一个数据报, 用于重组分片后的 IP 数据报。标志字段中的 DF 表示是否允许分片, MF 表示后面是否还有分片。片偏移则指出分组在分片后某片在原分组中的相对位置。

08. D

片偏移标识该分片所携带数据在原始分组所携带数据中的相对位置, 以 8B 为单位。

09. B

如果分组长度超过 MTU, 那么当 DF=1 时, 丢弃该分组, 并且要用 ICMP 差错报文向源主机报告; 当 DF=0 时, 进行分片, MF=1 表示后面还有分片。

10. B

对于 A 选项, 132.19.237.5 的前 8 位与 132.0.0.0/8 匹配。而 B 选项中, 132.19.237.5 的前 11 位与 132.0.0.0/11 匹配。C 选项中, 132.19.237.5 的前 22 位与 132.19.232.0/22 不匹配。根据“最长前缀匹配原则”, 该分组应该被转发到 R2。D 选项为默认路由, 只有当前面的所有目的网络都不能和分组的目的 IP 地址匹配时才使用。

11. A

在分类的 IP 网络中, C 类地址的前 24 位为网络位, 后 8 位为主机位, 主机位全“0”表示网络号, 主机位全“1”表示广播地址, 因此最多可以有 $2^8 - 2 = 254$ 台主机或路由器。

12. A

CIDR 地址块 86.32.0.0/12 的网络前缀为 12 位，说明第 2 个字节的前 4 位在前缀中，第 2 个字节 32 的二进制形式为 00100000。给出的 4 个地址的前 8 位均相同，而第 2 个字节的前 4 位分别是 0010, 0100, 0100, 0100，所以本题答案为 A。

13. A

10.255.255.255 为 A 类地址，主机号全 1，代表网络广播，为广播地址。192.168.24.59/30 为 CIDR 地址，只有后面 2 位为主机号，而 59 用二进制表示为 00111011，可知主机号全 1，代表网络广播，为广播地址。224.105.5.211 为 D 类组播地址。

14. D

所有形如 127.xx.yz.zz 的 IP 地址，都作为保留地址，用于回路测试。

15. B

A 是 C 类地址，掩码为 255.255.255.0，由此得知 A 地址的主机号为全 0（未使用 CIDR），因此不能作为主机地址。C 是为回环测试保留的地址。D 是语法错误的地址，不允许有 256。B 为 A 类地址，其网络号是 110，主机号是 47.10.0。

16. C

由于 $240_{10} = 11110000_2$ ，所以共有 12 比特位用于主机地址，且主机位全 0 和全 1 不能使用，所以最多可以有的主机数为 $2^{12} - 2 = 4094$ 。

17. A

在 Internet 中，IP 数据报从源结点到目的结点可能需要经过多个网络和路由器。当一个路由器接收到一个 IP 数据报时，路由器根据 IP 数据报首部的目的 IP 地址进行路由选择，并不改变源 IP 地址的取值。即使 IP 数据报被分片时，原 IP 数据报的源 IP 地址和目的 IP 地址也将复制到每个分片的首部，因此在整个传输过程中，IP 数据报首部的源 IP 地址和目的 IP 地址都不发生变化。

18. C

划分子网可以增加子网的数量，子网之间的数据传输需要通过路由器进行，因此自然就减少了广播域的大小。另外，划分子网，由于子网号占据了主机号位，所以会减少主机的数量；划分子网仅提高 IP 地址的利用率，并不增加网络的数量。

19. A

由题可知，主机号有 5 位，若主机号只占 1 位，则没有有效的 IP 地址可供分配（除去 0 和 1），因此最少 2 位表示主机号，还剩 3 位表示子网号，所以最多可以分成 8 个子网。而当 5 位都表示主机数，即只有 1 个子网时，每个子网最多具有 30 个有效的 IP 地址（除去全 0 和全 1）。

20. C

如果一台主机有两个或两个以上的 IP 地址，那么说明这台主机属于两个或两个以上的逻辑网络。值得注意的是，在同一时刻一个合法的 IP 地址只能分配给一台主机，否则就会引起 IP 地址冲突。IP 地址 192.168.11.25 属于 C 类 IP 地址，所以 A、B、D 同属于一个逻辑网络，只有 C 的网络号不同，表明它在不同的逻辑网络。

21. A

CIDR 是一种归并网络的技术，CIDR 技术的作用是把小的网络汇聚成大的超网。

22. D、A

CIDR 地址由网络前缀和主机号两部分组成，CIDR 将网络前缀都相同的连续 IP 地址组成“CIDR 地址块”。网络前缀的长度为 20 位，主机号为 12 位，因此 192.168.0.0/20 地址块中的地址数为 2^{12} 个。其中，当主机号为全 0 时，取最小地址 192.168.0.0。当主机号全为 1 时，取最大地址 192.168.15.255。注意，这里并不是指可分配的主机地址。

关注公众号【乘龙考研】
一手更新 稳定有保障

对于 192.16.0.19/28，表示子网掩码为 255.255.255.240。IP 地址 192.16.0.19 与 IP 地址 192.16.0.17 所对应的前 28 位数相同，都是 11000000 00010000 00000000 0001，所以 IP 地址 192.16.0.17 是子网 192.16.0.19/28 的一台主机地址。注意，主机号全 0 和全 1 的地址不使用。

23. B

为每个 IP 分组设定生存时间 (TTL)，每经过一个路由器，TTL 减 1，TTL 为 0 时，路由器就不再转发该分组。因此可以避免分组在网络中无限循环下去。

24. D

最初设计的分类 IP 地址，由于每类地址所能连接的主机数大大超过一般单位的需求量，从而造成了 IP 地址的浪费。划分子网通过从网络的主机号借用若干比特作为子网号，从而使原来较大规模的网络细分为几个规模较小的网络，提高了 IP 地址的利用率。

CIDR 是比划分子网更为灵活的一种手段，它消除了 A、B、C 类地址及划分子网的概念。使用各种长度的网络前缀来代替分类地址中的网络号和子网号，将网络前缀都相同的 IP 地址组成“CIDR 地址块”。网络前缀越短，地址块越大。因特网服务提供者再根据客户的具体情况，分配合适大小的 CIDR 地址块，从而更加有效地利用 IPv4 的地址空间。

采用网络地址转换 (NAT)，可以使一些使用本地地址的专用网连接到因特网上，进而使得一些机构的内部主机可以使用专用地址，只需给此机构分配一个 IP 地址即可，并且这些专用地址是可重用的——其他机构也可使用，所以大大节省了 IP 地址的消耗。

尽管以上三种方法可以在一定阶段内有效缓解 IP 地址耗尽的危机，但无论是从计算机本身发展来看还是从因特网的规模和传输速率来看，现在的 IPv4 地址已很不适用，所以治本的方法还是使用 128bit 编址的 IPv6 地址。

25. B

当路由器准备将 IP 分组发送到网络上，而该网络又无法将整个分组一次发送时，路由器必须将该分组分片，使其长度能满足这一网络对分组长度的限制。IP 分片可以独立地通过各个路径发送，而且在传输过程中仍然存在分片的可能（不同网络的 MTU 可能不同），因此不能由中间路由器进行重组。分片后的 IP 分组直至到达目的主机后才能汇集在一起，并且甚至不一定以原先的次序到达。这样，进行接收的主机都要求支持重组能力。

26. C

“/29”表明前 29 位是网络号，4 个选项的前 3 个字节均相同。A 中第 4 个字节 120 为 01111000，前 5 位为 01111；B 中第 4 个字节 64 为 01000000，前 5 位为 01000；C 中第 4 个字节 96 为 01100000，前 4 位为 0110；D 中第 4 个字节 104 为 01101000，前 5 位为 01101。由于已经分配的子网 74.178.247.96/29 的第 4 字节的前 5 位为 01100，这与 C 中第 4 字节的前 4 位重叠。因此 C 中的网络前缀不能再分配给其他子网。

27. D

本题实际上就是要求找一个子网掩码，使得 A 和 B 的 IP 地址与该子网掩码逐位相“与”之后得到相同的结果。D 选项与 A、B 相“与”的结果均为 216.0.0.0。

28. B

未进行子网划分时，B 类地址有 16 位作为主机位。由于共需要划分 51 个子网， $2^5 < 51 < 2^6$ ，那么需要从主机位划出 6 位作为子网号，剩下的 10 位主机位可容纳的主机数为 1022（即 $2^{10} - 2$ ）台主机，满足题目要求。因此子网掩码为 255.255.252.0。

29. A

4 条路由的前 24 位（3 个字节）为网络前缀，前 2 个字节都一样，因此只需要比较第 3 个字节。

节即可， $129 = 10000001$, $130 = 10000010$, $132 = 10000100$, $133 = 10000101$ 。前 5 位是完全相同的，因此聚合后的网络的掩码中，1 的数量应该是 $8 + 8 + 5 = 21$ ，聚合后的网络的第 3 个字节应该是 $10000000 = 128$ ，因此答案为 $172.18.128.0/21$ 。

30. B

本题中，某主机不能正常通信意味着它与其他三台主机不在同一个子网，只需判断哪个选项和其他选项不在同一个子网即可。子网掩码为 $255.255.255.224$ 表示前 27 位是网络号，可以看出选项 B 属于子网 $202.3.1.64/27$ ，其他三项属于子网 $202.3.1.32/27$ 。或者，后 5 位是主机号，前 3 个子网的地址范围为 $202.3.1.1 \sim 30, 33 \sim 62, 65 \sim 94$ （排除全 0 或全 1），据此也能选出答案。

31. A

IP 数据报的首部既有源 IP 地址，又有目的 IP 地址，但在通信中路由器只会根据目的 IP 地址进行路由选择。IP 数据报在通信过程中，首部的源 IP 地址和目的 IP 地址在经过路由器时不会发生改变。ARP 广播只在子网中传播，由于相互通信的主机不在同一个子网内，因此不可以直接通过 ARP 广播得到目的站的硬件地址。硬件地址只具有本地意义，因此每当路由器将 IP 数据报转发到一个具体的网络中时，都需要重新封装源硬件地址和目的硬件地址。

注意：路由器在接收到分组后，剥离该分组的数据链路层协议头，然后在分组被转发之前，给分组加上一个新的链路层协议头。

32. C

NAT 协议保留了 3 段 IP 地址供内部使用，这 3 段地址如下：

A 类：1 个 A 类网段，即 $10.0.0.0 \sim 10.255.255.255$ ，主机数 16777216。

B 类：16 个 B 类网段，即 $172.16.0.0 \sim 172.31.255.255$ ，主机数 1048576。

C 类：256 个 C 类网段，即 $192.168.0.0 \sim 192.168.255.255$ ，主机数 65536。

所以只有 C 选项是内部地址，不允许出现在因特网上。

33. C

NAT 的表项需要管理员添加，这样才能控制一个内网到外网的网络连接。题目中主机发送的分组在 NAT 表项中找不到（端口 80 从源端口而非转换端口找），所以服务器不转发该分组。

34. C

当源主机要向本地局域网上的某主机发送 IP 数据报时，先在其 ARP 高速缓存中查看有无目的 IP 地址与 MAC 地址的映射。若有，就把这个硬件地址写入 MAC 帧，然后通过局域网把该 MAC 帧发往此硬件地址；若没有，则先通过广播 ARP 请求分组，在获得目的主机的 ARP 响应分组后，将目的主机的 IP 地址与硬件地址的映射写入 ARP 高速缓存。如果目的主机不在本局域网上，那么将 IP 分组发送给本局域网上的路由器，当然要先通过同样的方法获得路由器的 IP 地址和硬件地址的映射关系。

35. C、A

由于不知道目标设备在哪里，所以 ARP 请求必须使用广播方式。但是 ARP 请求包中包含有发送方的 MAC 地址，因此应答时应该使用单播方式。

36. B

主机先使用 ARP 来查询本网络路由器的地址，然后每个路由器使用 ARP 来寻找下一跳路由的地址，总共使用了 4 次 ARP 从主机 A 网络的路由器到达主机 B 网络的路由器。然后，主机 B 网络的路由器使用 ARP 找到主机 B，所以总共使用了 $1 + 4 + 1 = 6$ 次 ARP。

37. C

DHCP 提供了一种机制，使得使用 DHCP 可自动获得 IP 的配置信息而无须手工干预。

38. A

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

39. D

ICMP 是一个网络层协议，但是其文仍然要封装在 IP 分组中发送。

40. C

PING 使用了 ICMP 的询问报文中的回送请求和回答报文。

关注公众号【乘龙考研】
一手更新 稳定有保障

41. B

由于该网络的 IP 地址为 192.168.5.0/24，网络号为前 24 位，后 8 位为子网号 + 主机号。子网掩码为 255.255.255.248，第 4 个字节 248 转换成二进制为 11111000，因此后 8 位中，前 5 位用于子网号，在 CIDR 中可以表示 $2^5=32$ 个子网；后 3 位用于主机号，除去全 0 和全 1 的情况，可以表示 $2^3-2=6$ 台主机地址。

42. C

ICMP 差错报告报文有 5 种：终点不可达、源点抑制、时间超过、参数问题、改变路由（重定向），其中源点抑制是指在路由器或主机由于拥塞而丢弃数据报时，向源点发送源点抑制报文，使源点知道应当把数据报的发送速率放慢。

43. C

首先分析 192.168.4.0/30 这个网络，主机号只占 2 位，地址范围为 192.168.4.0~192.168.4.3，主机号全 1 时，即 192.168.4.3 是广播地址，因此可容纳 $4-2=2$ 台主机。

44. D

子网掩码的第 3 个字节为 11111100，可知前 22 位为子网号、后 10 位为主机号。IP 地址的第 3 个字节为 01001101（下画线为子网号的一部分），将主机号（即后 10 位）全置为 1，可以得到广播地址为 180.80.79.255。

45. A

在实际网络的数据链路层上传送数据时，最终必须使用硬件地址，ARP 将网络层的 IP 地址解析为数据链路层的 MAC 地址。

46. B

ICMP 报文作为数据字段封装在 IP 分组中，因此 IP 直接为 ICMP 提供服务。UDP 和 TCP 都是传输层协议，为应用层提供服务。PPP 是数据链路层协议，为网络层提供服务。

47. C

根据“最长前缀匹配原则”，169.96.40.5 与 169.96.40.0 的前 27 位匹配最长，因此选 C。选项 D 为默认路由，只有当前面的所有目的网络都不能和分组的目的 IP 地址匹配时才使用。

48. C

从子网掩码可知 H1 和 H2 处于同一网段，H3 和 H4 处于同一网段，分别可以进行正常的 IP 通信，A 和 D 错误。因为 R2 的 E1 接口的 IP 地址为 192.168.3.254，而 H2 的默认网关为 192.168.3.1，所以 H2 不能访问 Internet，而 H4 的默认网关为 192.168.3.254，所以 H4 可以正常访问 Internet，B 错误。由 H1、H2、H3 和 H4 的子网掩码可知 H1、H2 和 H3、H4 处于不同的网段，需通过路由器才能进行正常的 IP 通信，而这时 H1 和 H2 的默认网关为 192.168.3.1，但 R2 的 E1 接口的 IP 地址为 192.168.3.254，无法进行通信，从而 H1 不能与 H3 进行正常的 IP 通信。C 正确。

49. D

由题意知连接 R1、R2 和 R3 之间的点对点链路使用 201.1.3.x/30 地址，其子网掩码为 255.255.255.252，R1 的一个接口的 IP 地址为 201.1.3.9，转换为对应的二进制的后 8 位为 0000 1001（由 201.1.3.x/30 知，IP 地址对应的二进制的后两位为主机号，而主机号全为 0 表示本网络本身，主机号全为 1 表示本网络的广播地址，不用于源 IP 地址或目的 IP 地址），那么除 201.1.3.9 外，只有 IP 地址为 201.1.3.10 可以作为源 IP 地址使用（本题为 201.1.3.10）。Web 服务器的 IP 地址为 130.18.10.1，作为 IP 分组的目的 IP 地址。综上可知，D 正确。

50. C

这个网络有 16 位的主机号，平均分成 128 个规模相同的子网，每个子网有 7 位的子网号，9 位的主机号。除去一个网络地址和广播地址，可分配的最大 IP 地址个数是 $2^9 - 2 = 512 - 2 = 510$ 。

51. A

0.0.0.0/32 可以作为本主机在本网络上的源地址。127.0.0.1 是回送地址，以它为目的 IP 地址的数据将被立即返回本机。200.10.10.3 是 C 类 IP 地址。255.255.255.255 是广播地址。

52. C

对于此类题目，先分析需要聚合的 IP 地址。观察发现，题中的四个路由地址，前 16 位完全相同，不同之处在于第 3 段的 8 位中，将这 8 位展开写成二进制，分别如下：

	7	6	5	4	3	2	1	0
32	0	0	1	0	0	0	0	0
40	0	0	1	0	1	0	0	0
48	0	0	1	1	0	0	0	0
56	0	0	1	1	1	0	0	0

观察发现，四个地址的第 3 段中，从前向后最多有 3 位相同，因此这 3 位是能聚合的最大位数。将这些相同的位都保留，将第 3 段第 3 位之后的所有位都置 0，就得到了聚合后的 IP 地址：35.230.32.0，其网络前缀为 $16 + 3$ ，也即前 19 位，因此聚合后的网络地址为 35.230.32.0/19。

53. D

在网络的信息传递中，会经常用到两个地址：MAC 地址和 IP 地址。其中，MAC 地址会随着信息被发往不同的网络而改变，但 IP 地址当且仅当信息在私人网络中传递时才会改变。分组 P 在如题图所示的网络中传递时，首先由主机 H1 将分组发往路由器 R，此时源 MAC 地址为 H1 主机本身的 MAC 地址，即 00-1a-2b-3c-4d-52，目的 MAC 地址为路由器 R 的 MAC 地址，即 00-1a-2b-3c-4d-51。路由器 R 收到分组 P 后，根据分组 P 的目的 IP 地址，得知应将分组从另一个端口转发出去，于是会给分组 P 更换新的 MAC 地址，此时由于从另外的端口转发出去，因此 P 的新源 MAC 地址变为负责转发的端口 MAC 地址，即 00-a1-b2-c3-d4-61，目的 MAC 地址应为主机 H2 的 MAC 地址，即 00-a1-b2-c3-d4-62。根据分析过程，题目所问的 MAC 地址应为路由器 R 两个端口的 MAC 地址，因此选 D。

54. B

网络前缀为 20 位，将 101.200.16.0/20 划分为 5 个子网，为了保证有子网的可分配 IP 地址数尽可能小，即要让其他子网的可分配 IP 地址数尽可能大，不能采用平均划分的方法，而要采用变长的子网划分方法，也就是最大子网用 1 位子网号，第二大子网用 2 位子网号，以此类推。

子网 1：101.200.00010000.00000001~101.200.00010111.11111110；地址范围为 101.200.16.1/21~101.200.23.254/21；可分配的 IP 地址数为 2046 个。

子网 2：101.200.00011000.00000001~101.200.00011011.11111110；地址范围为 101.200.24.1/22~

101.200.27.254/22；可分配的 IP 地址数为 1022 个。

子网 3：101.200.00011100.00000001~101.200.00011101.11111110；地址范围为 101.200.28.1/23~101.200.29.254/23；可分配的 IP 地址数为 510 个。

子网 4：101.200.00011110.00000001~101.200.00011110.11111110；地址范围为 101.200.30.1/24~101.200.30.254/24；可分配的 IP 地址数为 254 个。

子网 5：101.200.00011111.00000001~101.200.00011111.11111110；地址范围为 101.200.31.1/24~101.200.31.254/24；可分配的 IP 地址数为 254 个。

综上所述，可能的最小子网的可分配 IP 地址数是 254 个。

55. B

根据题意，将 IP 网络划分为 3 个子网。其中一个是 192.168.9.128/26。可以简写成 x.x.x.10/26（其中 10 是 128 的二进制 1000 0000 的前两位，因为 $26 - 24 = 2$ ）。

A 选项可以简写成 x.x.x.0/25；

B 选项可以简写成 x.x.x.00/26；

C 选项可以简写成 x.x.x.11/26；

D 选项可以简写成 x.x.x.110/27。

关注公众号【乘龙考研】
一手更新 稳定有保障

对于 A 和 C，可以组成 x.x.x.0/25、x.x.x.10/26、x.x.x.11/26 这样 3 个互不重叠的子网。

对于 D，可以组成 x.x.x.10/26、x.x.x.110/27、x.x.x.111/27 这样 3 个互不重叠的子网。

但对于 B，要想将一个 IP 网络划分为几个互不重叠的子网，3 个是不够的，至少需要划分为 4 个子网：x.x.x.00/26、x.x.x.01/26、x.x.x.10/26、x.x.x.11/26。

56. B

链路层 MTU=800B。IP 分组首部长 20B。片偏移以 8 个字节为偏移单位，因此除了最后一个分片，其他每个分片的数据部分长度都是 8B 的整数倍。所以，最大 IP 分片的数据部分长度为 776B。在总长度为 1580B 的 IP 数据报中，数据部分占 1560B， $1560B/776B=2.01\dots$ ，需要分成 3 片。故第 2 个分片的总长度字段为 796，MF 为 1（表示还有后续的分片）。

57. B

主机所在网络的网络地址可以通过主机的 IP 地址和子网掩码逐位相与得到。子网掩码 255.255.192.0 的二进制前 18 位为 1、后 14 位为 0，把主机 IP 地址的后 14 位变为 0，得到的结果为 183.80.64.0，即为主机所在网络的网络地址。

58. D

默认网关可以理解为离当前主机最近的路由器的端口地址，所以是 192.168.1.62，而该主机的子网掩码和网关的子网掩码也相同，/27 即为 255.255.255.224。

二、综合应用题

01. 【解答】

要求出分组所携带数据的长度，就需要分别知道首部的长度和分组的总长度。解题的关键在于弄清首部长度的字段和总长度字段的单位。由于首部长度字段的单位是 4B，101 的十进制为 5，所以首部长度= $5\times4=20$ B。而总长度字段的单位是字节，101000 的十进制为 40，所以总长度为 40B，因此分组携带的数据长度为 $40-20=20$ B。

02. 【解答】

数据报长度为 4000B，有效载荷为 $4000-20=3980$ B。网络能传送的最大有效载荷为 $1500-20=1480$ B，因此应分为 3 个短些的片，各片的数据字段长度分别为 1480、1480 和 1020B。片

段偏移字段的单位为 8B, $1480/8=185$, $(1480 \times 2)/8=370$, 因此片段偏移字段的值分别为 0、185、370。MF=1 时, 代表后面还有分片; MF=0 时, 代表后面没有分片, 因此 MF 字段的值分别为 1、1 和 0 (注意, MF=0 不能确定是独立的数据报, 还是分片得来的, 只有当 MF=0 且片段偏移字段 >0 时, 才能确定是分片的最后一个分片)。

03. 【解答】

在 IP 层下面的每种数据链路层都有自己的帧格式, 其中包括帧格式中的数据字段的最大长度, 这称为最大传输单位 (MTU)。 $1500 - 20 = 1480$, $2000 - 1480 = 520$, 所以原 IP 数据报经过第一个网络后分成了两个 IP 小报文, 第一个报文的数据部分长度是 1480B, 第二个报文的数据部分长度是 520B。

(除最后一个报片外的)所有报片的有效载荷都是 8B 的倍数。 $576 - 20 = 556$, 但 556 不能被 8 整除, 所以分片时的数据部分最大只能取 552。第一个报文经过第 2 个网络后, $1480 - 552 \times 2 = 376 < 576$, 变成数据长度分别为 552B、552B、376B 的 3 个 IP 小报文; 第 2 个报文 $520 < 552$, 因此不用分片。因此到达目的主机时, 原 2000B 的数据被分成数据长度分别为 552B、552B、376B、520B 的 4 个小报文。

04. 【解答】

分片的片偏移值表示其数据部分首字节在原始分组的数据部分中的相对位置, 单位为 8B。首部长度字段以 4B 为单位, 总长度字段以字节为单位。题目中, 分组的片偏移值为 100, 因此其数据部分第一个字节的编号是 800。因为分组的总长度为 100B, 首部长度为 $4 \times 5 = 20$ B, 所以数据部分长度为 80B。因此该分组的数据部分的最后一个字节的编号是 879。

05. 【解答】

通过将目的地址和子网掩码换算成二进制, 并进行逐位“与”就可得到子网地址。但是通常在目的地址中, 子网掩码为 255 所对应的部分在子网地址中不变, 子网掩码为 0 所对应的部分在子网地址中为 0, 其他部分按二进制逐位“与”求得(也可直接截取)。本题中, 子网掩码的前两部分为 255.255, 因此子网地址的前两部分为 201.230; 子网掩码最后一部分为 0, 因此子网地址的最后一部分为 0; 子网地址的第三部分为 240, 进行换算有 $240 = (11110000)_2$, $34 = (00100010)_2$, 逐位相“与”得 $(00100000)_2 = 32$ 。因此子网地址为 201.230.32.0。

06. 【解答】

由于一个 CIDR 地址块中可以包含很多地址, 所以路由表中就利用 CIDR 地址块来查找目的网络, 这种地址的聚合常称为路由聚合。

本题已知有 212.56.132.0/24、212.56.133.0/24、212.56.134.0/24、212.56.135.0/24 地址块, 可知第 3 字节前 6 位相同, 因此共同前缀为 $8 + 8 + 6 = 22$ 位, 由于这 4 个地址块的第 1、2 个字节相同, 考虑它们的第 3 个字节: $132 = (10000100)_2$, $133 = (10000101)_2$, $134 = (10000110)_2$, $135 = (10000111)_2$, 所以共同的前缀有 22 位, 即 1101010000111000100001 , 聚合的 CIDR 地址块是 212.56.132.0/22。

07. 【解答】

- 1) 可以采用划分子网的方法对题公司的网络进行划分。由于该公司包括 4 个部门, 共需要划分为 4 个子网。
- 2) 已知网络地址 192.168.161.0 是一个 C 类地址, 所需的子网数为 4, 每个子网的主机数为 20~30。子网号的比特数为 3, 即最多有 $2^3 = 8$ 个可分配的子网, 主机号的比特数为 5, 由于主机号不允许为全 0 或全 1, 因此每个子网最多有 $2^5 - 2 = 30$ 个可分配的 IP 地址。
- 3) 4 个部门子网的子网掩码均为 255.255.255.224, 各部门的网络地址与部门主机的 IP 地址范围可分配如下:

部 门	部门网络地址	主机 IP 地址范围
工程技术部	192.168.161.32	192.168.161.33~192.168.161.62
市场部	192.168.161.64	192.168.161.65~192.168.161.94
财务部	192.168.161.96	192.168.161.97~192.168.161.126
办公室	192.168.161.128	192.168.161.129~192.168.161.158

08. 【解答】

1) 使用 CIDR 时，可能会导致有多个匹配结果，应当从当前匹配结果中选择具有最长网络前缀的路由。下面来一一分析分组 A 与表中这四项的匹配性：

- ① 131.128.56.0/24 与 31.128.55.33 不匹配，因为前 24 位不同。
- ② 131.128.55.32/28 与 131.128.55.33 的前 24 位匹配，只需看后面 4 位是否匹配，32 转换为二进制为 **0010 0000**，33 转换为二进制为 **0010 0001**，匹配，且匹配了 28 位。
- ③ 131.128.55.32/30 与 131.128.55.33 的前 24 位匹配，只需要看后面 6 位是否匹配，32 转换为二进制为 **0010 0000**，33 转换为二进制为 **0010 0001**，匹配，且匹配了 30 位。
- ④ 131.128.0.0/16 与 131.128.55.33 匹配，且匹配了 16 位。

综上，对于分组 A，第 2、3、4 项都能与之匹配，但根据最长网络前缀匹配原则，应该选择网络前缀为 131.128.55.32/30 的表项进行转发，下一跳路由器为 C。

同理，对于分组 B，路由表中第 2 和 4 项都能与之匹配，但是根据最长网络前缀匹配原则，应该选择第 2 个路由表项转发，下一跳路由器为 B。

- 2) 要想该路由表项使得以 131.128.55.33 为目的地址的 IP 分组选择“A”作为下一跳，而不影响其他目的地址的 IP 分组转发，只需构造 1 条网络前缀和该地址匹配 32 位的项即可。增加的表项为：网络前缀 131.128.55.33/32；下一跳 A。
- 3) 增加 1 条默认路由：网络前缀 0.0.0.0/0；下一跳 E。
- 4) 要划分成 4 个规模尽可能大的子网，需要从主机位中划出 2 位作为子网位 ($2^2=4$)，CIDR 广泛使用之后允许子网位可以全 0 和全 1。子网掩码均为 11111111 11111111 11111111 11000000，即 255.255.255.192。而地址范围内不能包含主机位全 0 或全 1 的地址。

子 网	子 网 掩 码	地 址 范 围
131.128.56.0/26	255.255.255.192	131.128.56.1~131.128.56.62
131.128.56.64/26	255.255.255.192	131.128.56.65~131.128.56.126
131.128.56.128/26	255.255.255.192	131.128.56.129~131.128.56.190
131.128.56.192/26	255.255.255.192	131.128.56.193~131.128.56.254

09. 【解答】

- 1) 网络号 C4.5E.10.0/20（下一站地是 B）的第 3 字节可以用二进制表示成 0001 0000。目标地址 C4.5E.13.87 的第 3 字节可以用二进制表示成 0001 0011，显然取 20 位掩码与网络号 C4.5E.10.0/20 相匹配，所以具有该目标地址的 IP 分组将被投递到下一站地 B。
- 2) 网络号 C4.50.0.0/12（下一站地是 A）的第 2 字节可以用二进制表示成 0101 0000。目标地址 C4.5E.22.09 的第 2 字节可以用二进制表示成 0101 1110，显然取 12 位掩码与网络号 C4.50.0.0/12 相匹配，所以具有该目的地址的 IP 分组将被投递到下一站地 A。
- 3) 网络号 80.0.0.0/1（下一站地是 E）的第 1 字节可以用二进制表示成 1000 0000。目标地址 C3.41.80.02 的第 1 字节可以用二进制表示成 1100 0011，显然取 1 位掩码与网络号 80.0.0.0/1 相匹配，所以具有该目标地址的 IP 分组将被投递到下一站地 E。

- 4) 网络号 40.0.0.0/2 (下一站地是 F) 的第 1 字节可以用二进制表示成 0100 0000。目标地址 5E.43.91.12 的第 1 字节可以用二进制表示成 0101 1110，显然取 2 位掩码与网络号 40.0.0.0/2 相匹配，所以具有该目标地址的 IP 分组将被投递到下一站地 F。

10. 【解答】

分配网络前缀应先分配地址数较多的前缀。已知该自治系统分配到的 IP 地址块为 30.138.118/23 (注意：①一个路由器端口也需要占用一个 IP 地址；②子网划分的答案不唯一)。

LAN3：主机数 150，由于 $(2^7 - 2) < 150 + 1 < (2^8 - 2)$ ，所以主机号为 8bit，网络前缀为 24。取第 24 位为 0，分配地址块 30.138.118.0/24。

LAN2：主机数 91，由于 $(2^6 - 2) < 91 + 1 < (2^7 - 2)$ ，所以主机号为 7bit，网络前缀为 25。取第 24、25 位为 10，分配地址块 30.138.119.0/25。

LAN5：主机数为 15，由于 $(2^4 - 2) < 15 + 1 < (2^5 - 2)$ ，所以主机号为 5bit，网络前缀 27。取第 24、25、26、27 位为 1110，分配地址块 30.138.119.192/27。

LAN1：共有 3 个路由器，再加上一个网关地址，至少需要 4 个 IP 地址。由于 $(2^2 - 2) < 3 + 1 < (2^3 - 2)$ ，所以主机号为 3bit，网络前缀为 29。取第 24、25、26、27、28、29 位为 111101，分配地址块 30.138.119.232/29。

LAN4：主机数为 3，由于 $(2^2 - 2) < 3 + 1 < (2^3 - 2)$ ，所以主机号为 3bit，网络前缀 29。取第 24、25、26、27、28、29 位为 111110，分配地址块 30.138.119.240/29。

11. 【解答】

- 1) 共同的子网掩码为 255.255.255.240，表示前 28 位为网络号，同一网段内的 IP 地址具有相同的网络号。主机 A 的网络号为 192.168.75.16；主机 B 的网络号为 192.168.75.144；主机 C 的网络号为 192.168.75.144；主机 D 的网络号为 192.168.75.160；主机 E 的网络号为 192.168.75.160。因此 5 台主机 A、B、C、D、E 分属 3 个网段，主机 B 和 C 在一个网段，主机 D 和 E 在一个网段，A 主机在一个网段。主机 D 的网络号为 192.168.75.160。
- 2) 主机 F 与主机 A 同在一个网段，所以主机 F 所在的网段为 192.168.75.16，第 4 个字节 16 的二进制表示为 0001 0000，最后边的 4 位为主机位，去掉全 0 和全 1。则其 IP 地址范围为 192.168.75.17~192.168.75.30，并且不能为 192.168.75.18。
- 3) 由于 164 的二进制为 1010 0100，将最右边的 4 位全置为 1，即 1010 1111，则广播地址为 192.168.75.175。主机 D 和主机 E 可以收到。

12. 【解答】

- 1) CIDR 中的子网号可以全 0 或全 1，但主机号不能全 0 或全 1。

因此若将 IP 地址空间 202.118.1.0/24 划分为 2 个子网，且每个局域网需分配的 IP 地址个数不少于 120 个，则子网号至少要占用一位。

由 $2^6 - 2 < 120 < 2^7 - 2$ 可知，主机号至少要占用 7 位。

由于源 IP 地址空间的网络前缀为 24 位，因此主机号位数 + 子网号位数 = 8。

综上可得主机号位数为 7，子网号位数为 1。

因此子网的划分结果为子网 1：202.118.1.0/25，子网 2：202.118.1.128/25。

地址分配方案：子网 1 分配给局域网 1，子网 2 分配给局域网 2；或子网 1 分配给局域网 2，子网 2 分配给局域网 1。

- 2) 由于局域网 1 和局域网 2 分别与路由器 R1 的 E1、E2 接口直接相连，因此在 R1 的路由表中，目的网络为局域网 1 的转发路径是直接通过接口 E1 转发的，目的网络为局域网 2 的转发路径是直接通过接口 E2 转发的。由于局域网 1、2 的网络前缀均为 25 位，因此它

们的子网掩码均为 255.255.255.128。

R1 专门为域名服务器设定了一个特定的路由表项，因此该路由表项中的子网掩码应为 255.255.255.255（只有和全 1 的子网掩码相与时，才能完全保证和目的 IP 地址一样，从而选择该特定路由）。对应的下一跳转发地址是 202.118.2.2，转发接口是 L0。

R1 到互联网的路由实质上相当于一个默认路由（即当某一目的网络 IP 地址与路由表中其他任何一项都不匹配时，匹配该默认路表项），默认路由一般写为 0/0，即目的地址为 0.0.0.0，子网掩码为 0.0.0.0。对应的下一跳转发地址是 202.118.2.2，转发接口是 L0。

综上可得到路由器 R1 的路由表如下：

(若子网 1 分配给局域网 1，子网 2 分配给局域网 2)

目的网络 IP 地址	子网掩码	下一跳 IP 地址	接口
202.118.1.0	255.255.255.128	—	E1
202.118.1.128	255.255.255.128	—	E2
202.118.3.2	255.255.255.255	202.118.2.2	L0
0.0.0.0	0.0.0.0	202.118.2.2	L0

(若子网 1 分配给局域网 2，子网 2 分配给局域网 1)

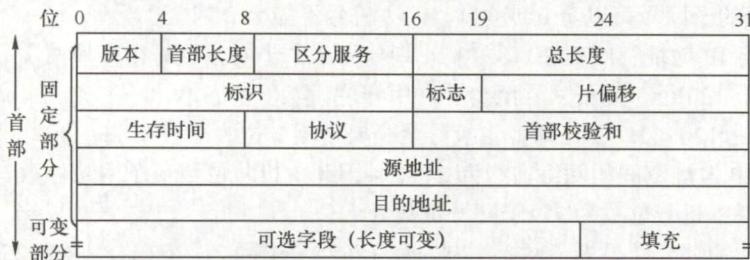
目的网络 IP 地址	子网掩码	下一跳 IP 地址	接口
202.118.1.128	255.255.255.128	—	E1
202.118.1.0	255.255.255.128	—	E2
202.118.3.2	255.255.255.255	202.118.2.2	L0
0.0.0.0	0.0.0.0	202.118.2.2	L0

- 3) 局域网 1 和局域网 2 的地址可以聚合为 202.118.1.0/24，而对于路由器 R2 来说，通往局域网 1 和局域网 2 的转发路径都是从 L0 接口转发的，因此采用路由聚合技术后，路由器 R2 到局域网 1 和局域网 2 的路由如下：

目的网络 IP 地址	子网掩码	下一跳 IP 地址	接口
202.118.1.0	255.255.255.0	202.118.2.1	L0

13. 【解答】

IPv4 的首部格式如下，然后根据首部格式来解析首部各个字段的含义。



- 由上图可知，源 IP 地址为 IP 首部的第 13、14、15、16 字节，即 7C 4E 03 02，转换为点分十进制表示可得源 IP 地址为 124.78.3.2。目的 IP 地址为 IP 首部的第 17、18、19、20 字节，即 B4 0E 0F 02，转换为点分十进制表示可得目的 IP 地址为 180.14.15.2。
- 分组总长度是 IP 首部的第 3、4 字节，即 00 54，转换为十进制得该分组总长度为 84，单位为字节。而首部长度是 IP 首部的第 5~8 位，值为 5，单位为 4B，因此首部长度为 $4B \times 5 = 20B$ 。

数据部分长度 = 总长度 - 首部长度 = $84 - 20 = 64B$ 。

- 3) 该分组首部的片偏移字段为第 7、8 字节（除去第 7 字节的前 3 位），不等于 0，而是二进制值 1 1000 0101 0000，即十进制数 6224，单位是 8B。

另外，分组的标志字段为第 7 字节的前 3 位，即 010，中间位 DF=1 表示不可分片，最后位 MF=0 表示后面没有分片。IP 规范规定，所有主机和网关至少能支持 576B 的分组长度。在 576B 的数据报中，512B 用于存放数据，64B 用作分组头。由于本报片的数据部分的长度只有 64B，所以不会再次被分割。

14. 【解答】

- 1) MAC 地址只具有本地意义（ARP 也只能工作在同一局域网中）。该帧为 A 收到的帧，因此目的 MAC 地址为 A 的 MAC 地址，源 MAC 地址为网关路由器端口的 MAC 地址（若为 A 发出的帧，则目的 MAC 地址为默认网关的 MAC 地址）。首先找到目的 MAC 地址 00:1d:72:98:1d:fc 的位置（在下图中的位置 1 标出），根据以太网帧的结构，目的 MAC 地址后面紧邻的是源 MAC 地址，因此源 MAC 地址为 00:00:5e:00:01:01。

0000	1	00	1d	72	98	1d	fc	00	00	5e	00	01	01	88	64	11	00	^ . . . d . .
0010	75	89	01	92	00	21	45	00	01	90	f9	bf	40	00	33	06	u ! E @ . 3 .		
0020	f3	15	da	c7	66	28	da	cf	3d	d3	00	50	c4	8f	dc	a6	. . . f (. . . = . . P . . .		
0030	a2	96	23	4c	44	69	25	18	00	0f	76	3d	00	00	90	b5	. . # L D i P . . v = . . .		

- 2) 要求得 IP 分组所携带的数据量，需要知道首部长度和总长度。218.207.61.211 表示成十六进制是 da.cf.3d.d3，并且作为分组中的目的 IP 地址。在图中确定目的 IP 地址的位置（位置 2），再根据 IP 首部的结构，分别从目的 IP 的位置向前数 14 和 16 个字节，即可找到总长度和首部长度字段的位置。但首部长度字段所在的字节值为 0x45，首部长度字段只有 4 位，前 4 位是版本号。因此首部字段的值为 5，单位为 4B，所以首部长度为 20B。总长度字段值为 0x0190，十进制为 400B。因此分组携带的数据长度为 380B。
- 3) 由于整个 IP 分组的长度是 400B，大于输出链路 MTU（380B）。这时需要考虑分片，但是否能够分片还得看 IP 首部中的标志位。IP 首部中的标志字段占 3 位，从前到后依次为保留位、DF 位、MF 位。根据 IP 首部结构找到标志字段所在的字节，其值为 0x40，二进制表示为 01000000，那么 DF=1，不能对该 IP 分组进行分片。此时，路由器应进行的操作是丢弃该分组，并用 ICMP 差错报文向源主机报告。

15. 【解答】

- 1) DHCP 服务器可为主机 2~N 动态分配 IP 地址的最大范围是 111.123.15.5~111.123.15.254；主机 2 发送的封装 DHCP Discover 报文的 IP 分组的源 IP 地址和目的 IP 地址分别是 0.0.0.0 和 255.255.255.255。
- 2) 主机 2 发出的第一个以太网帧的目的 MAC 地址是 ff-ff-ff-ff-ff-ff；封装主机 2 发往 Internet 的 IP 分组的以太网帧的目的 MAC 地址是 00-a1-a1-a1-a1-a1。
- 3) 主机 1 能访问 WWW 服务器，但不能访问 Internet。由于主机 1 的子网掩码配置正确而默认网关 IP 地址被错误地配置为 111.123.15.2（正确 IP 地址是 111.123.15.1），所以主机 1 可以访问在同一个子网内的 WWW 服务器，但当主机 1 访问 Internet 时，主机 1 发出的 IP 分组会被路由到错误的默认网关（111.123.15.2），从而无法到达目的主机。

16. 【解答】

- 1) 广播地址是网络地址中主机号全 1 的地址（主机号全 0 的地址代表网络本身）。销售部和技术部均分配了 192.168.1.0/24 的 IP 地址空间，IP 地址的前 24 位为子网的网络号。于是在后 8 位中划分部门的子网，选择前 1 位作为部门子网的网络号。令销售部子网的网络

号为 0，技术部子网的网络号为 1，则技术部子网的完整地址为 192.168.1.128；令销售部子网的主机号全 1，可以得到该部门的广播地址为 192.168.1.127。

每台主机仅分配一个 IP 地址，计算目前还可以分配的主机数，用技术部可以分配的主机数减去已分配的主机数，技术部总共可以分配的计算机主机数为 $2^7 - 2 = 126$ （减去全 0 和全 1 的主机号）。已经分配了 $208 - 129 + 1 = 80$ 台，此外还有 1 个 IP 地址（192.168.1.254）分配给了路由器的端口，因此还可以分配 $126 - 80 - 1 = 45$ 台。

- 2) 判断分片的大小，需要考虑各个网段的 MTU，而且注意分片的数据长度必须是 8B 的整数倍。由题可知，在技术部子网内， $MTU = 800B$ ，IP 分组头部长 20B，最大 IP 分片封装数据的字节数为 $\lfloor (800 - 20)/8 \rfloor \times 8 = 776$ 。至少需要的分片数为 $\lceil (1500 - 20)/776 \rceil = 2$ 。第 1 个分片的偏移量为 0；第 2 个分片的偏移量为 $776/8 = 97$ 。

17. 【解答】

- 1) 两个子网使用了相同的网段，且路由器开启了 NAT 功能，加上题干给出了 NAT 表结构，因此需要配置 NAT 表。路由器 R2 开启 NAT 服务，当路由器 R2 从 WAN 口收到来自 H2 或 H3 发来的数据时，根据 NAT 表发送给 Web 服务器的对应端口。外网 IP 地址应该为路由器的外端 IP 地址，内网 IP 地址应该为 Web 服务器的地址，Web 服务器默认端口为 80，因此内网端口号固定为 80，当其他网络的主机访问 Web 服务器时，默认访问的端口应该也是 80，但是访问的目的 IP 是路由器的 IP 地址，因此 NAT 表中的外部端口最好也统一为 80。题中并未要求对 H1 进行访问，因此 H1 的 NAT 表项可以不写。R2 的 NAT 表配置如下：

外网		内网	
IP 地址	端口号	IP 地址	端口号
203.10.2.2	80	192.168.1.2	80

- 2) 由于启用了 NAT 服务，H2 发送的 P 的源 IP 地址应该是 H2 的内网地址，目的地址应该是 R2 的外网 IP 地址，源 IP 地址是 192.168.1.2，目的 IP 地址是 203.10.2.2。R3 转发后，将 P 的源 IP 地址改为 R3 的外网 IP 地址，目的 IP 地址仍然不变，源 IP 地址是 203.10.2.6，目的 IP 地址是 203.10.2.2。R2 转发后，将 P 的目的 IP 地址改为 Web 服务器的内网地址，源地址仍然不变，源 IP 地址是 203.10.2.6，目的 IP 地址是 192.168.1.2。

18. 【解答】

- 1) 设备 1 选择 100BaseT 以太网交换机，设备 2 选择 100BaseT 集线器。因为物理层设备既不能隔离冲突域，又不能隔离广播域，链路层设备可以隔离冲突域但不能隔离广播域。
- 2) 假设 H2 与 H3 之间的最远距离是 D，根据 CSMA/CD 协议的工作原理有

$$\text{最短帧长} = \text{总线传播时延} \times \text{数据传输速率} \times 2$$

由于使用 100BaseT 局域网标准，数据传输率为 100Mb/s，总线传播时延由两部分组成，一部分是信号传播时延，另一部分是信号通过设备 2 时产生的额外 $1.51\mu s$ 延迟。代入公式为 $64B = (1.51\mu s + D/(2 \times 10^8 m/s)) \times 100Mb/s \times 2$ ，注意单位换算，最终解得 $D = 210m$ 。

- 3) M 是 DHCP 发现报文（DISCOVER 报文）。路由器 E0 接口能收到封装 M 的以太网帧，由于 H4 发送的 DHCP 发现报文是广播的形式，所以同一个广播域内的所有设备和接口都可以收到该以太网帧。由于是广播帧，所以目的 MAC 地址是全 1，S 向 DHCP 服务器转发的封装 M 的以太网帧的目的 MAC 地址是 FF-FF-FF-FF-FF-FF。
- 4) 在 H5 收到的帧中，地址 1、地址 2 和地址 3 分别是 00-11-11-11-11-E1、00-11-11-11-11-C1

和 00-11-11-11-11-D1。该帧来自 AP，地址 1 代表接收端的地址，地址 2 代表 AP 的地址，地址 3 是发送端的地址。

4.4 IPv6

4.4.1 IPv6 的主要特点

解决“IP 地址耗尽”问题的措施有以下三种：①采用无类别编址 CIDR，使 IP 地址的分配更加合理；②采用网络地址转换（NAT）方法以节省全球 IP 地址；③采用具有更大地址空间的新版本的 IPv6。其中前两种方法只是延长了 IPv4 地址分配完毕的时间，只有第三种方法从根本上解决了 IP 地址的耗尽问题。

IPv6 的主要特点如下：

- 1) 更大的地址空间。IPv6 将地址从 IPv4 的 32 位增大到了 128 位。IPv6 的字节数（16B）是 IPv4 字节数（4B）的平方。
- 2) 扩展的地址层次结构。
- 3) 灵活的首部格式。
- 4) 改进的选项。
- 5) 允许协议继续扩充。
- 6) 支持即插即用（即自动配置）。
- 7) 支持资源的预分配。
- 8) IPv6 只有在包的源结点才能分片，是端到端的，传输路径中的路由器不能分片，所以从一般意义上说，IPv6 不允许分片（不允许类似 IPv4 的路由分片）。
- 9) IPv6 首部长度必须是 8B 的整数倍，而 IPv4 首部是 4B 的整数倍。
- 10) 增大了安全性。身份验证和保密功能是 IPv6 的关键特征。



虽然 IPv6 与 IPv4 不兼容，但总体而言它与所有其他的因特网协议兼容，包括 TCP、UDP、ICMP、IGMP、OSPF、BGP 和 DNS，只是在少数地方做了必要的修改（大部分是为了处理长的地址）。IPv6 相当好地满足了预定的目标，主要体现在：

- 1) 首先也是最重要的，IPv6 有比 IPv4 长得多的地址。IPv6 的地址用 16 个字节表示，地址空间是 IPv4 的 $2^{128-32} = 2^{96}$ 倍，从长远来看，这些地址是绝对够用的。
- 2) 简化了 IP 分组头，它包含 8 个域（IPv4 是 12 个域）。这一改变使得路由器能够更快地处理分组，从而可以改善吞吐率。
- 3) 更好地支持选项。这一改变对新的分组首部很重要，因为一些从前必要的段现在变成了可选段。此外，表示选项的方式的改变还能加快分组的处理速度。

4.4.2 IPv6 地址

IPv6 数据报的目的地址可以是以下三种基本类型地址之一：

- 1) 单播。单播就是传统的点对点通信。
- 2) 多播。多播是一点对多点的通信，分组被交付到一组计算机的每台计算机。
- 3) 任播。这是 IPv6 增加的一种类型。任播的目的站是一组计算机，但数据报在交付时只交付其中的一台计算机，通常是距离最近的一台计算机。

IPv4 地址通常使用点分十进制表示法。如果 IPv6 也使用这种表示法，那么地址书写起来将会相当长。在 IPv6 标准中指定了一种比较紧凑的表示法，即把地址中的每 4 位用一个十六进制数表示，并用冒号分隔每 16 位，如 4BF5:AA12:0216:FEBC:BA5F:039A:BE9A:2170。

通常可以把 IPv6 地址缩写成更紧凑的形式。当 16 位域的开头有一些 0 时，可以采用一种缩写表示法，但在域中必须至少有一个数字。例如，可以把地址 4BF5:0000:0000:0000:BA 5F:039A:000A:2176 缩写为 4BF5:0:0:0:BA5F:39A:A:2176。

当有相继的 0 值域时，还可以进一步缩写。这些域可以用双冒号缩写 (::)。当然，双冒号表示法在一个地址中仅能出现一次，因为 0 值域的个数没有编码，需要从指定的总的域的个数来推算。这样一来，前述地址可被更紧凑地书写成 4BF5::BA5F:39A:A:2176。

IPv6 扩展了 IPv4 地址的分级概念，它使用以下 3 个等级：第一级（顶级）指明全球都知道的公共拓扑；第二级（场点级）指明单个地点；第三级指明单个网络接口。IPv6 地址采用多级体系主要是为了使路由器能够更快地查找路由。

从 IPv4 向 IPv6 过渡只能采用逐步演进的办法，同时还必须使新安装的 IPv6 系统能够向后兼容。IPv6 系统必须能够接收和转发 IPv4 分组，并且能够为 IPv4 分组选择路由。

从 IPv4 向 IPv6 过渡可以采用双协议栈和隧道技术两种策略：双协议栈是指在一台设备上同时装有 IPv4 和 IPv6 协议栈，那么这台设备既能和 IPv4 网络通信，又能和 IPv6 网络通信。如果这台设备是一个路由器，那么在路由器的不同接口上分别配置了 IPv4 地址和 IPv6 地址，并很可能分别连接了 IPv4 网络和 IPv6 网络；如果这台设备是一台计算机，那么它将同时拥有 IPv4 地址和 IPv6 地址，并具备同时处理这两个协议地址的功能。隧道技术的要点是在 IPv6 数据报要进入 IPv4 网络时，把整个 IPv6 数据报封装到 IPv4 数据报的数据部分，使得 IPv6 数据报就好像在 IPv4 网络的隧道中传输。

4.4.3 本节习题精选

单项选择题

01. 下一代因特网核心协议 IPv6 的地址长度是 ()。
 - A. 32bit
 - B. 48bit
 - C. 64bit
 - D. 128bit
02. 与 IPv4 相比，IPv6 ()。
 - A. 采用 32 位 IP 地址
 - B. 增加了头部字段数目
 - C. 不提供 QoS 保障
 - D. 没有提供校验和字段
03. 以下关于 IPv6 地址 1A22:120D:0000:0000:72A2:0000:00C0 的表示中，错误的是()。
 - A. 1A22:120D::72A2:0000:00C0
 - B. 1A22:120D::72A2:0:0:C0
 - C. 1A22::120D::72A2::00C0
 - D. 1A22:120D:0:0:72A2::C0
04. 下列关于 IPv6 的描述中，错误的是 ()。
 - A. IPv6 的首部长度是不可变的
 - B. IPv6 不允许分片
 - C. IPv6 采用了 16B 的地址，在可预见的将来不会用完
 - D. IPv6 使用了首部校验和来保证传输的正确性
05. 如果一个路由器收到的 IPv6 数据报因太大而不能转发到链路上，那么路由器将把该数据报 ()。
 - A. 丢弃
 - B. 暂存
 - C. 分片
 - D. 转发至能支持该数据报的链路上

4.4.4 答案与解析

单项选择题

01. D

IPv6 的地址用 16B (即 128bit) 表示, 比 IPv4 长得多, 地址空间是 IPv4 的 2^{96} 倍。

02. D

IPv6 采用 128 位地址, 所以选项 A 错。IPv6 减少了头部字段数目, 仅包含 8 个字段, 选项 B 错。IPv6 支持 QoS, 以满足实时、多媒体通信的需要, 选项 C 错。由于目前网络传输介质的可靠性较高, 出现比特错误的可能性很低, 且数据链路层和传输层有自己的校验, 为了效率, IPv6 没有校验和字段。

03. C

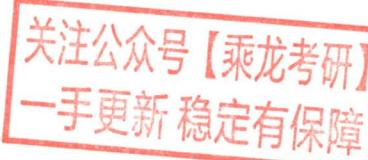
使用零压缩法时, 双冒号 “::” 在一个地址中只能出现一次。也就是说, 当有多处不相邻的 0 时, 只能用 “::” 代表其中的一处。

04. D

IPv6 的首部长度是固定的, 因此不需要首部长度字段。IPv6 取消了校验和字段, 这样就加快了路由器处理数据报的速度。我们知道, 数据链路层会丢弃检测出差错的帧, 运输层也有相应的差错处理机制, 因此网络层的差错检测可以精简掉。

05. A

IPv6 中不允许分片。因此, 如果路由器发现到来的数据报太大而不能转发到链路上, 那么丢弃该数据报, 并向发送方发送一个指示分组太大的 ICMP 报文。



4.5 路由协议

4.5.1 自治系统

自治系统 (Autonomous System, AS): 单一技术管理下的一组路由器, 这些路由器使用一种 AS 内部的路由选择协议和共同的度量来确定分组在该 AS 内的路由, 同时还使用一种 AS 之间的路由选择协议来确定分组在 AS 之间的路由。

一个自治系统内的所有网络都由一个行政单位 (如一家公司、一所大学、一个政府部门等) 管辖, 一个自治系统的所有路由器在本自治系统内都必须是连通的。

4.5.2 域内路由与域间路由

自治系统内部的路由选择称为域内路由选择, 自治系统之间的路由选择称为域间路由选择。因特网有两大类路由选择协议。

1. 内部网关协议 (Interior Gateway Protocol, IGP)

内部网关协议即在一个自治系统内部使用的路由选择协议, 它与互联网中其他自治系统选用什么路由选择协议无关。目前这类路由选择协议使用得最多, 如 RIP 和 OSPF。

2. 外部网关协议 (External Gateway Protocol, EGP)

若源站和目的站处在不同的自治系统中, 当数据报传到一个自治系统的边界时 (两个自治系统可能使用不同的 IGP), 就需要使用一种协议将路由选择信息传递到另一个自治系统中。这样的

协议就是外部网关协议（EGP）。目前使用最多的外部网关协议是 BGP-4。

图 4.10 是两个自治系统互连的示意图。每个自治系统自己决定在本自治系统内部运行哪个内部路由选择协议（例如，可以是 RIP，也可以是 OSPF），但每个自治系统都有一个或多个路由器（图中的路由器 R1 和 R2）。除运行本系统的内部路由选择协议外，还要运行自治系统间的路由选择协议（如 BGP-4）。



图 4.10 自治系统和内部网关协议、外部网关协议

4.5.3 路由信息协议（RIP）

路由信息协议（Routing Information Protocol, RIP）是内部网关协议（IGP）中最先得到广泛应用的协议。RIP 是一种分布式的基于距离向量的路由选择协议，其最大优点就是简单。

1. RIP 规定

- 1) 网络中的每个路由器都要维护从它自身到其他每个目的网络的距离记录（因此这是一组距离，称为距离向量）。
- 2) 距离也称跳数（Hop Count），规定从一个路由器到直接连接网络的距离（跳数）为 1。而每经过一个路由器，距离（跳数）加 1。
- 3) RIP 认为好的路由就是它通过的路由器的数目少，即优先选择跳数少的路径。
- 4) RIP 允许一条路径最多只能包含 15 个路由器（即最多允许 15 跳）。因此距离等于 16 时，它表示网络不可达。可见 RIP 只适用于小型互联网。距离向量路由可能会出现环路的情况，规定路径上的最高跳数的目的是为了防止数据报不断循环在环路上，减少网络拥塞的可能性。
- 5) RIP 默认在任意两个使用 RIP 的路由器之间每 30 秒广播一次 RIP 路由更新信息，以便自动建立并维护路由表（动态维护）。
- 6) 在 RIP 中不支持子网掩码的 RIP 广播，所以 RIP 中每个网络的子网掩码必须相同。但在新的 RIP2 中，支持变长子网掩码和 CIDR。

2. RIP 的特点（注意与 OSPF 的特点比较）

- 1) 仅和相邻路由器交换信息。
- 2) 路由器交换的信息是当前路由器所知道的全部信息，即自己的路由表。
- 3) 按固定的时间间隔交换路由信息，如每隔 30 秒。

RIP 通过距离向量算法来完成路由表的更新。最初，每个路由器只知道与自己直接相连的网络。通过每 30 秒的 RIP 广播，相邻两个路由器相互将自己的路由表发给对方。于是经过第一次 RIP 广播，每个路由器就知道了与自己相邻的路由器的路由表（即知道了距离自己跳数为 1 的网络的路由）。同理，经过第二次 RIP 广播，每个路由器就知道了距离自己跳数为 2 的网络的路由……因此，经过若干 RIP 广播后，所有路由器都最终知道了整个 IP 网络的路由表，称为 RIP 最终是收敛的。通过 RIP 收敛后，每个路由器到每个目标网络的路由都是距离最短的（即跳数最少，最短路由），哪怕还存在另一条高速（低时延）但路由器较多的路由。

3. 距离向量算法

每个路由表项目都有三个关键数据： \langle 目的网络 N , 距离 d , 下一跳路由器地址 X \rangle 。对于每个相邻路由器发送过来的 RIP 报文，执行如下步骤：

- 1) 对地址为 X 的相邻路由器发来的 RIP 报文，先修改此报文中的所有项目：把“下一跳”字段中的地址都改为 X ，并把所有“距离”字段的值加 1。
- 2) 对修改后的 RIP 报文中的每个项目，执行如下步骤：
 - ① 当原来的路由表中没有目的网络 N 时，把该项目添加到路由表中。
 - ② 当原来的路由表中有目的网络 N ，且下一跳路由器的地址是 X 时，用收到的项目替换原路由表中的项目。
 - ③ 当原来的路由表中有目的网络 N ，且下一跳路由器的地址不是 X 时，如果收到的项目中的距离 d 小于路由表中的距离，那么就用收到的项目替换原路由表中的项目；否则什么也不做。
- 3) 如果 180 秒 (RIP 默认超时时间为 180 秒) 还没有收到相邻路由器的更新路由表，那么把此相邻路由器记为不可达路由器，即把距离设置为 16 (距离为 16 表示不可达)。
- 4) 返回。

RIP 最大的优点是实现简单、开销小、收敛过程较快。RIP 的缺点如下：

- 1) RIP 限制了网络的规模，它能使用的最大距离为 15 (16 表示不可达)。
- 2) 路由器之间交换的是路由器中的完整路由表，因此网络规模越大，开销也越大。
- 3) 网络出现故障时，会出现慢收敛现象（即需要较长时间才能将此信息传送到所有路由器），俗称“坏消息传得慢”，使更新过程的收敛时间长。

RIP 是应用层协议，它使用 UDP 传送数据（端口 520）。RIP 选择的路径不一定是时间最短的，但一定是具有最少路由器的路径。因为它是根据最少跳数进行路径选择的。

4.5.4 开放最短路径优先 (OSPF) 协议

1. OSPF 协议的基本特点

开放最短路径优先 (OSPF) 协议是使用分布式链路状态路由算法的典型代表，也是内部网关协议 (IGP) 的一种。OSPF 与 RIP 相比有以下 4 点主要区别：

- 1) OSPF 向本自治系统中的所有路由器发送信息，这里使用的方法是洪泛法。而 RIP 仅向自己相邻的几个路由器发送信息。
- 2) 发送的信息是与本路由器相邻的所有路由器的链路状态，但这只是路由器所知道的部分信息。“链路状态”说明本路由器和哪些路由器相邻及该链路的“度量”（或代价）。而在 RIP 中，发送的信息是本路由器所知道的全部信息，即整个路由表。
- 3) 只有当链路状态发生变化时，路由器才用洪泛法向所有路由器发送此信息，并且更新过程收敛得快，不会出现 RIP “坏消息传得慢”的问题。而在 RIP 中，不管网络拓扑是否发生变化，路由器之间都会定期交换路由表的信息。
- 4) OSPF 是网络层协议，它不使用 UDP 或 TCP，而直接用 IP 数据报传送（其 IP 数据报首部的协议字段为 89）。而 RIP 是应用层协议，它在传输层使用 UDP。

除以上区别外，OSPF 还有以下特点：

- 1) OSPF 对不同的链路可根据 IP 分组的不同服务类型 (TOS) 而设置成不同的代价。因此，OSPF 对不同类型的业务可计算出不同的路由，十分灵活。
- 2) 如果到同一个目的网络有多条相同代价的路径，那么可以将通信量分配给这几条路径。这称为多路径间的负载平衡。

关注公众号【乘龙考研】
一手更新 稳定有保障

- 3) 所有在 OSPF 路由器之间交换的分组都具有鉴别功能，因而保证了仅在可信赖的路由器之间交换链路状态信息。
- 4) 支持可变长度的子网划分和无分类编址 CIDR。
- 5) 每个链路状态都带上一个 32 位的序号，序号越大，状态就越新。

2. OSPF 的基本工作原理

由于各路由器之间频繁地交换链路状态信息，因此所有路由器最终都能建立一个链路状态数据库。这个数据库实际上就是全网的拓扑结构图，它在全网范围内是一致的（称为链路状态数据库的同步）。然后，每个路由器根据这个全网拓扑结构图，使用 Dijkstra 最短路径算法计算从自己到各目的网络的最优路径，以此构造自己的路由表。此后，当链路状态发生变化时，每个路由器重新计算到各目的网络的最优路径，构造新的路由表。

注意：虽然使用 Dijkstra 算法能计算出完整的最优路径，但路由表中不会存储完整路径，而只存储“下一跳”（只有到了下一跳路由器，才能知道再下一跳应当怎样走）。

为使 OSPF 能够用于规模很大的网络，OSPF 将一个自治系统再划分为若干更小的范围，称为区域。划分区域的好处是，将利用洪泛法交换链路状态信息的范围局限于每个区域而非整个自治系统，减少了整个网络上的通信量。在一个区域内部的路由器只知道本区域的完整网络拓扑，而不知道其他区域的网络拓扑情况。这些区域也有层次之分。处在上层的域称为主干区域，负责连通其他下层的区域，并且还连接其他自治域。

3. OSPF 的五种分组类型

OSPF 共有以下五种分组类型：

- 1) 问候分组，用来发现和维持邻站的可达性。
- 2) 数据库描述分组，向邻站给出自己的链路状态数据库中的所有链路状态项目的摘要信息。
- 3) 链路状态请求分组，向对方请求发送某些链路状态项目的详细信息。
- 4) 链路状态更新分组，用洪泛法对全网更新链路状态。
- 5) 链路状态确认分组，对链路更新分组的确认。

通常每隔 10 秒，每两个相邻路由器要交换一次问候分组，以便知道哪些站可达。在路由器刚开始工作时，OSPF 让每个路由器使用数据库描述分组和相邻路由器交换本数据库中已有的链路状态摘要信息。然后，路由器使用链路状态请求分组，向对方请求发送自己所缺少的某些链路状态项目的详细信息。经过一系列的这种分组交换，就建立了全网同步的链路数据库。图 4.11 给出了 OSPF 的基本操作，说明了两个路由器需要交换的各种类型的分组。

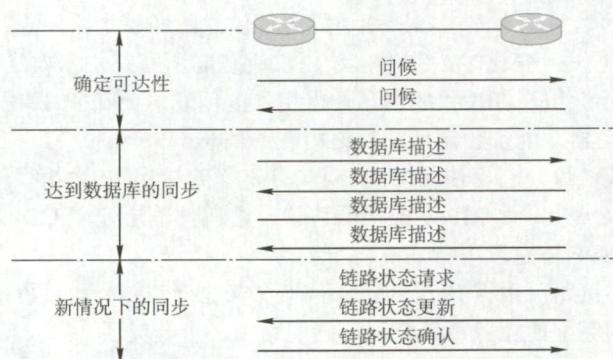


图 4.11 OSPF 的基本操作

在网络运行的过程中，只要一个路由器的链路状态发生变化，该路由器就要使用链路状态更新分组，用洪泛法向全网更新链路状态。其他路由器在更新后，发送链路状态确认分组对更新分组进行确认。

为了确保链路状态数据库与全网的状态保持一致，OSPF 还规定每隔一段时间（如 30 分钟）就刷新一次数据库中的链路状态。由于一个路由器的链路状态只涉及与相邻路由器的连通状态，因而与整个互联网的规模并无直接关系。因此，当互联网规模很大时，OSPF 要比 RIP 好得多，而且 OSPF 协议没有“坏消息传播得慢”的问题。

注意：教材上说 OSPF 协议不使用 UDP 数据报传送，而是直接使用 IP 数据报传送，在此解释一下什么称为用 UDP 传送，什么称为用 IP 数据报传送。用 UDP 传送是指将该信息作为 UDP 报文的数据部分，而直接使用 IP 数据报传送是指将该信息直接作为 IP 数据报的数据部分。RIP 报文是作为 UDP 数据报的数据部分。

4.5.5 边界网关协议（BGP）

边界网关协议（Border Gateway Protocol, BGP）是不同自治系统的路由器之间交换路由信息的协议，是一种外部网关协议。边界网关协议常用于互联网的网关之间。

内部网关协议主要设法使数据报在一个 AS 中尽可能有效地从源站传送到目的站。在一个 AS 内部不需要考虑其他方面的策略。然而 BGP 使用的环境却不同，主要原因如下：

- 1) 因特网的规模太大，使得自治系统之间路由选择非常困难。
- 2) 对于自治系统之间的路由选择，要寻找最佳路由是很不现实的。
- 3) 自治系统之间的路由选择必须考虑有关策略。

边界网关协议（BGP）只能力求寻找一条能够到达目的网络且比较好的路由（不能兜圈子），而非寻找一条最佳路由。BGP 采用的是路径向量路由选择协议，它与距离向量协议和链路状态协议有很大的区别。BGP 是应用层协议，它是基于 TCP 的。

BGP 的工作原理如下：每个自治系统的管理员要选择至少一个路由器（可以有多个）作为该自治系统的“BGP 发言人”。一个 BGP 发言人与其他自治系统中的 BGP 发言人要交换路由信息，就要先建立 TCP 连接（可见 BGP 报文是通过 TCP 传送的，也就是说 BGP 报文是 TCP 报文的数据部分），然后在此连接上交换 BGP 报文以建立 BGP 会话，再利用 BGP 会话交换路由信息。当所有 BGP 发言人都相互交换网络可达性的信息后，各 BGP 发言人就可找出到达各个自治系统的较好路由。

每个 BGP 发言人除必须运行 BGP 外，还必须运行该 AS 所用的内部网关协议，如 OSPF 或 RIP。BGP 所交换的网络可达性信息就是要到达某个网络（用网络前缀表示）所要经过的一系列 AS。图 4.12 给出了一个 BGP 发言人交换路径向量的例子。



图 4.12 主干网与自治系统间路径向量的交换

BGP 的特点如下：

- 1) BGP 交换路由信息的结点数量级是自治系统的数量级，比这些自治系统中的网络数少很多。
- 2) 每个自治系统中 BGP 发言人（或边界路由器）的数目是很少的。这样就使得自治系统之间的路由选择不致过分复杂。
- 3) BGP 支持 CIDR，因此 BGP 的路由表也就应当包括目的网络前缀、下一跳路由器，以及到达该目的网络所要经过的各个自治系统序列。
- 4) 在 BGP 刚运行时，BGP 的邻站交换整个 BGP 路由表，但以后只需在发生变化时更新有变化的部分。这样做对节省网络带宽和减少路由器的处理开销都有好处。

BGP-4 共使用 4 种报文：

- 1) 打开（Open）报文。用来与相邻的另一个 BGP 发言人建立关系。
- 2) 更新（Update）报文。用来发送某一路由的信息，以及列出要撤销的多条路由。
- 3) 保活（Keepalive）报文。用来确认打开报文并周期性地证实邻站关系。
- 4) 通知（Notification）报文。用来发送检测到的差错。

RIP、OSPF 与 BGP 的比较如表 4.3 所示。

表 4.3 三种路由协议的比较

协议	RIP	OSPF	BGP
类型	内部	内部	外部
路由算法	距离-向量	链路状态	路径-向量
传递协议	UDP	IP	TCP
路径选择	跳数最少	代价最低	较好，非最佳
交换结点	和本结点相邻的路由器	网络中的所有路由器	和本结点相邻的路由器
交换内容	当前本路由器知道的全部信息，即自己的路由表	与本路由器相邻的所有路由器的链路状态	首次 整个路由表 非首次 有变化的部分

4.5.6 本节习题精选

一、单项选择题

01. 以下关于自治系统的描述中，不正确的是（ ）。
 - A. 自治系统划分区域的好处是，将利用洪泛法交换链路状态信息的范围局限在每个区域内，而不是整个自治系统
 - B. 采用分层划分区域的方法使交换信息的种类增多，同时也使 OSPF 协议更加简单
 - C. OSPF 协议将一个自治系统再划分为若干更小的范围，称为区域
 - D. 在一个区域内部的路由器只知道本区域的网络拓扑，而不知道其他区域的网络拓扑的情况
02. 在计算机网络中，路由选择协议的功能不包括（ ）。

A. 交换网络状态或通路信息	B. 选择到达目的地的最佳路径
C. 更新路由表	D. 发现下一跳的物理地址
03. 用于域间路由的协议是（ ）。

A. RIP	B. BGP	C. OSPF	D. ARP
--------	--------	---------	--------
04. 在 RIP 中，到某个网络的距离值为 16，其意义是（ ）。

A. 该网络不可达	B. 存在循环路由
C. 该网络为直接连接网络	D. 到达该网络要经过 15 次转发
05. 在 RIP 中，假设路由器 X 和路由器 K 是两个相邻的路由器，X 向 K 说：“我到目的网络

Y 的距离为 N”，则收到此信息的 K 就知道：“若将到网络 Y 的下一个路由器选为 X，则我到网络 Y 的距离为（ ）。”（假设 N 小于 15）

- A. N B. N-1 C. 1 D. N+1

06. 以下关于 RIP 的描述中，错误的是（ ）。

- A. RIP 是基于距离-向量路由选择算法的
B. RIP 要求内部路由器将它关于整个 AS 的路由信息发布出去
C. RIP 要求内部路由器向整个 AS 的路由器发布路由信息
D. RIP 要求内部路由器按照一定的时间间隔发布路由信息

07. 对路由选择协议的一个要求是必须能够快速收敛，所谓“路由收敛”是指（ ）。

- A. 路由器能把分组发送到预定的目标
B. 路由器处理分组的速度足够快
C. 网络设备的路由表与网络拓扑结构保持一致
D. 能把多个子网聚合成一个超网

08. 下列关于 RIP 和 OSPF 协议的叙述中，错误的是（ ）。

- A. RIP 和 OSPF 协议都是网络层协议
B. 在进行路由信息交换时，RIP 中的路由器仅向自己相邻的路由器发送信息，OSPF 协议中的路由器向本自治系统中的所有路由器发送信息
C. 在进行路由信息交换时，RIP 中的路由器发送的信息是整个路由表，OSPF 协议中的路由器发送的信息只是路由表的一部分
D. RIP 的路由器不知道全网的拓扑结构，OSPF 协议的任何一个路由器都知道自己所在区域的拓扑结构

09. OSPF 协议使用（ ）分组来保持与其邻居的连接。

- A. Hello B. Keepalive
C. SPF (最短路径优先) D. LSU (链路状态更新)

10. 以下关于 OSPF 协议的描述中，最准确的是（ ）。

- A. OSPF 协议根据链路状态法计算最佳路由
B. OSPF 协议是用于自治系统之间的外部网关协议
C. OSPF 协议不能根据网络通信情况动态地改变路由
D. OSPF 协议只适用于小型网络

11. 以下关于 OSPF 协议特征的描述中，错误的是（ ）。

- A. OSPF 协议将一个自治域划分成若干域，有一种特殊的域称为主干区域
B. 域之间通过区域边界路由器互连
C. 在自治系统中有 4 类路由器：区域内部路由器、主干路由器、区域边界路由器和自治域边界路由器
D. 主干路由器不能兼作区域边界路由器

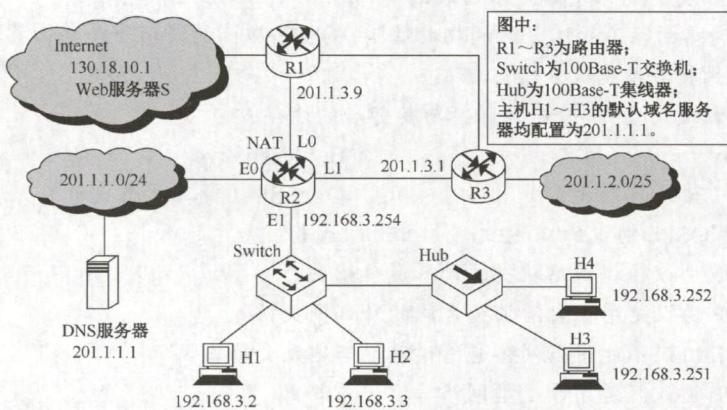
12. BGP 交换的网络可达性信息是（ ）。

- A. 到达某个网络所经过的路径 B. 到达某个网络的下一跳路由器
C. 到达某个网络的链路状态摘要信息 D. 到达某个网络的最短距离及下一跳路由器

13. RIP、OSPF 协议、BGP 的路由选择过程分别使用（ ）。

- A. 路径向量协议、链路状态协议、距离向量协议
B. 距离向量协议、路径向量协议、链路状态协议





目的网络	A 的距离向量	B 的距离向量	C 的距离向量	D 的距离向量
Net1	1	23	20	22
Net2	12	35	30	28
Net3	24	18	16	36
Net4	36	30	8	24

- A. 9, 10, 12, 6 B. 9, 10, 28, 20 C. 9, 20, 12, 20 D. 9, 20, 28, 20

二、综合应用题

01. RIP 使用 UDP, OSPF 使用 IP, 而 BGP 使用 TCP。这样做有何优点? 为什么 RIP 周期性地和邻站交换路由信息而 BGP 却不这样做?
02. 在某个使用 RIP 的网络中, B 和 C 互为相邻路由器, 其中表 1 为 B 的原路由表, 表 2 为 C 广播的距离向量报文<目的网络, 距离>。

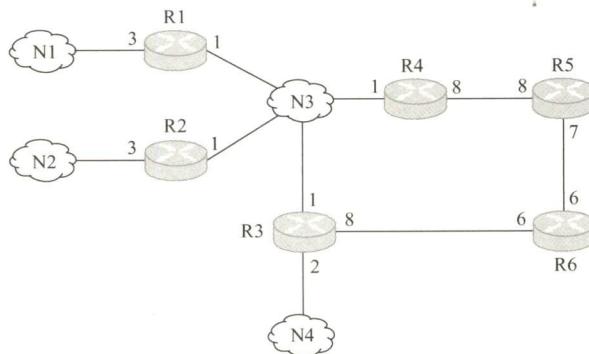
表 1

目的网络	距离	下一跳
N1	7	A
N2	2	C
N6	8	F
N8	4	E
N9	4	D

表 2

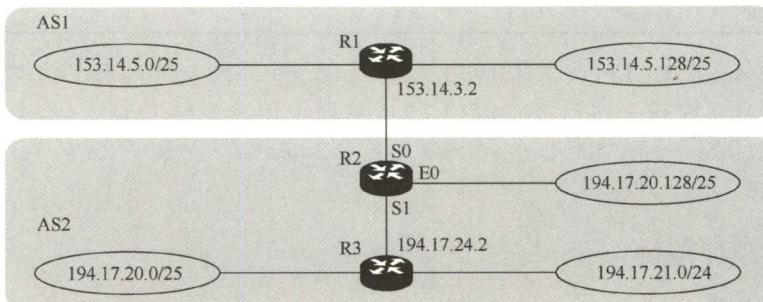
目的网络	距离
N2	15
N3	2
N4	8
N8	2
N7	4

- 1) 试求路由器 B 更新后的路由表并说明主要步骤。
 2) 当路由器 B 收到发往网络 N2 的 IP 分组时, 应该做何处理?
 03. 因特网中的一个自治系统的内部结构如下图所示。路由选择协议采用 OSPF 协议时, 计算 R6 的关于网络 N1、N2、N3、N4 的路由表。



注: 端口处的数字指该路由器向该链路转发分组的代价。

04. 【2013 统考真题】假设 Internet 的两个自治系统构成的网络如下图所示, 自治系统 AS1 由路由器 R1 连接两个子网构成; 自治系统 AS2 由路由器 R2、R3 互连并连接 3 个子网构成。各子网地址、R2 的接口名、R1 与 R3 的部分接口 IP 地址如下图所示。



请回答下列问题:

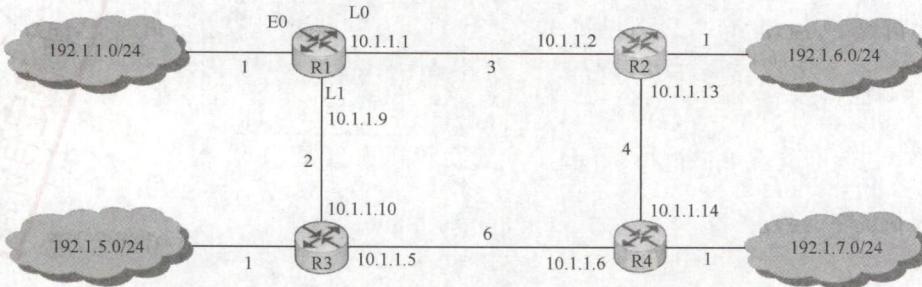
- 1) 假设路由表结构如下表所示。利用路由聚合技术，给出 R2 的路由表，要求包括到达图中所有子网的路由，且路由表中的路由项尽可能少。

目的网络	下一跳	接口
------	-----	----

- 2) 若 R2 收到一个目的 IP 地址为 194.17.20.200 的 IP 分组，R2 会通过哪个接口转发该 IP 分组？
- 3) R1 与 R2 之间利用哪个路由协议交换路由信息？该路由协议的报文被封装到哪个协议的分组中进行传输？

05. 【2014 统考真题】某网络中的路由器运行 OSPF 路由协议，下表是路由器 R1 维护的主要链路状态信息 (LSI)，下图是根据该表及 R1 的接口名构造的网络拓扑。

		R1 的 LSI	R2 的 LSI	R3 的 LSI	R4 的 LSI	备注
Router ID		10.1.1.1	10.1.1.2	10.1.1.5	10.1.1.6	标识路由器的 IP 地址
Link1	ID	10.1.1.2	10.1.1.1	10.1.1.6	10.1.1.5	所连路由器的 Router ID
	IP	10.1.1.1	10.1.1.2	10.1.1.5	10.1.1.6	Link1 的本地 IP 地址
	Metric	3	3	6	6	Link1 的费用
Link2	ID	10.1.1.5	10.1.1.6	10.1.1.1	10.1.1.2	所连路由器的 Router ID
	IP	10.1.1.9	10.1.1.13	10.1.1.10	10.1.1.14	Link2 的本地 IP 地址
	Metric	2	4	2	4	Link2 的费用
Net1	Prefix	192.1.1.0/24	192.1.6.0/24	192.1.5.0/24	192.1.7.0/24	直连网络 Net1 的网络前缀
	Metric	1	1	1	1	到达直连网络 Net1 的费用



请回答下列问题：

- 1) 假设路由表结构如下表所示，给出图中 R1 的路由表，要求包括到达图中子网 192.1.x.x 的路由，且路由表中的路由项尽可能少。

目的网络	下一跳	接口
------	-----	----

- 2) 当主机 192.1.1.130 向主机 192.1.7.211 发送一个 TTL=64 的 IP 分组时，R1 通过哪个接口转发该 IP 分组？主机 192.1.7.211 收到的 IP 分组的 TTL 是多少？
- 3) 若 R1 增加一条 Metric 为 10 的链路连接 Internet，则表中 R1 的 LSI 需要增加哪些信息？

4.5.7 答案与解析

一、单项选择题

01. B

划分区域的好处是，将利用洪泛法交换链路状态信息的范围局限在每个区域内，而不是整个自治系统。因此，在一个区域内部的路由器只知道本区域的网络拓扑，而不知道其他区域的网络拓扑情况。采用分层次划分区域的方法虽然使交换信息的种类增多了，同时也使 OSPF 协议更加复杂了，但这样做却能使每个区域内部交换路由信息的通信量大大减少，进而使 OSPF 协议能够用于规模很大的自治系统中。

02. D

路由选择协议的功能通常包括：获取网络拓扑信息、构建路由表、在网络中更新路由信息、选择到达每个目的网络的最优路径、识别一个网络的无环通路等。发现下一跳的物理地址一般是通过其他方式（如 ARP）来实现的，不属于路由选择协议的功能。

03. B

BGP（边界网关协议）是域间路由协议。RIP 和 OSPF 是域内路由协议，ARP 不是路由协议。

04. A

RIP 规定的最大跳数为 15，16 表示网络不可达。

05. D

RIP 规定，每经过一个路由器，距离（跳数）加 1。

06. C

RIP 规定一个路由器只向相邻路由器发布路由信息，而不像 OSPF 那样向整个域洪泛。

07. C

所谓收敛，是指当路由环境发生变化后，各路由器调整自己的路由表以适应网络拓扑结构的变化，最终达到稳定状态（路由表与网络拓扑状态保持一致）。收敛越快，路由器就能越快适应网络拓扑结构的变化。

08. A

RIP 是应用层协议，它使用 UDP 传送数据，OSPF 才是网络层协议。A 错误。

09. A

此题属于记忆性题目，OSPF 协议使用 Hello 分组来保持与其邻居的连接。

10. A

OSPF 协议是一种用于自治系统内的路由协议，选项 B 错误。它是一种基于链路状态路由选择算法的协议，能适用大型全局 IP 网络的扩展，支持可变长子网掩码，所以 OSPF 协议可用于管理一个受限地址域的中大型网络，选项 D 错误。OSPF 协议维护一张它所连接的所有链路状态信息的邻居表和拓扑数据库，使用组播链路状态更新（Link State Update, LSU）报文实现路由更新，并且只有当网络已经发生变化时才传送 LSU 报文，选项 C 错误。OSPF 协议不传送整个路由表，而传送受影响的路由更新报文。

11. D

主干区域中，用于连接主干区域和其他下层区域的路由器称为区域边界路由器。只要是在主干区域中的路由器，就都称为主干路由器，因此主干路由器可以兼作区域边界路由器。

12. A

由于 BGP 仅力求寻找一条能够到达目的网络且较好的路由（不能兜圈子），而并非寻找一条最佳路由，因此 D 选项错误。BGP 交换的路由信息是到达某个目的网络所要经过的各个自治系统序列而不仅仅是下一跳，因此选项 A 正确。

13. D

RIP 是一种分布式的基于距离向量的路由选择协议，它使用跳数来度量距离。RIP 选择的路径不一定是时间最短的，但一定是具有最小距离（最少跳数）的路径。

关注公众号【乘龙考研】
一手更新 稳定有保障

OSPF 协议使用分布式的链路状态协议，通过与相邻路由器频繁交流链路状态信息，来建立全网的拓扑结构图，然后使用 Dijkstra 算法计算从自己到各目的网络的最优路径。

由于 BGP 仅力求寻找一条能够到达目的网络且较好的路由（不能兜圈子），而并非寻找一条最佳路由，因此它采用的是路径向量路由选择协议。在 BGP 中，每个自治系统选出一个 BGP 发言人，这些发言人通过相互交换自己的路径向量（即网络可达性信息）后，就可找出到达各自治系统的较好路由。

14. B

距离-向量路由算法要求每个路由器维护一张路由表，该表给出了到达每个目的地址的已知最佳距离（最小代价）和下一步的转发地址。算法要求每个路由器定期与所有相邻路由器交换整个路由表，并更新自己的路由表项。注意从邻接结点接收到路由表不能直接进行比较，而要加上相邻结点传输消耗后再进行计算。C 到 B 的距离是 6，那么从 C 开始通过 B 到达各结点的最短距离向量是(11, 6, 14, 18, 12, 8)。同理，通过 D 和 E 的最短距离向量分别是(19, 15, 9, 3, 12, 13)和(12, 11, 8, 14, 5, 9)。那么 C 到所有结点的最短路径应该是(11, 6, 0, 3, 5, 8)。

15. D

R1 在收到信息并更新路由表后，若需要经过 R2 到达 net1，则其跳数为 17，由于距离为 16 表示不可达，因此 R1 不能经过 R2 到达 net1，R2 也不可能到达 net1。选项 B、C 错误，选项 D 正确。而题目中并未给出 R1 向 R2 发送的信息，因此选项 A 也不正确。

16. B

因为 R3 检测到网络 201.1.2.0/25 不可达，因此将到该网络的距离设置为 16（距离为 16 表示不可达）。当 R2 从 R3 收到路由信息时，因为 R3 到该网络的距离为 16，则 R2 到该网络也不可达，但此时记录 R1 可达（由于 RIP 的特点是“坏消息传得慢”，R1 并未收到 R3 发来的路由信息），R1 到该网络的距离为 2，再加上从 R2 到 R1 距离的 1，得 R2 到该网络的距离为 3。

17. D

RIP 是一种分布式的基于距离向量的路由选择协议，它通过广播 UDP 报文来交换路由信息。OSPF 是一个内部网关协议，要交换的信息量较大，应使报文的长度尽量短，所以不使用传输层协议（如 UDP 或 TCP），而直接采用 IP。BGP 是一个外部网关协议，在不同的自治系统之间交换路由信息，由于网络环境复杂，需要保证可靠传输，所以采用 TCP。因此，答案为选项 D。

18. D

根据距离向量路由算法，E 收到相邻路由器的距离向量后，更新它的路由表：

- ① 当原路由表中没有目的网络时，把该项目添加到路由表中。
- ② 发来的路由信息中有一条到达某个目的网络的路由，该路由与当前使用的路由相比，有较短的距离，就用经过发送路由信息的结点的新路由替换。

分析题意可知，E 与邻居路由器 A、B、C 和 D 之间的直接链路距离分别是 8, 10, 12 和 6。到达 Net1~Net4 没有直接链路，需要通过邻居路由器。从上述算法可知，E 到达目的网络一定是经过 A、B、C 和 D 中距离最小的。根据题中所给的距离信息，计算 E 经邻居路由器到达目的网络 Net1~Net4 的距离，如下表所示，选择到达每个目的网络距离的最短值。

目的网络	经过 A 需要的距离	经过 B 需要的距离	经过 C 需要的距离	经过 D 需要的距离
Net1	<u>9</u>	33	32	28
Net2	<u>20</u>	45	42	34
Net3	32	<u>28</u>	<u>28</u>	42
Net4	44	40	<u>20</u>	30

所以距离分别是 9, 20, 28, 20。

二、综合应用题

01. 【解答】

RIP 处于 UDP 的上层, RIP 所接收的路由信息都封装在 UDP 的数据报中; OSPF 的位置位于网络层, 由于要交换的信息量较大, 因此应使报文的长度尽量短, 因此采用 IP; BGP 要在不同的自治系统之间交换路由信息, 由于网络环境复杂, 需要保证可靠的传输, 所以选择 TCP。

内部网关协议主要设法使数据报在一个自治系统中尽可能有效地从源站传送到目的站, 在一个自治系统内部并不需要考虑其他方面的策略, 然而 BGP 使用的环境却不同。主要有以下三个原因: 第一, 因特网规模太大, 使得自治系统之间的路由选择非常困难; 第二, 对于自治系统之间的路由选择, 要寻找最佳路由是不现实的; 第三, 自治系统之间的路由选择必须考虑有关策略。由于上述情况, BGP 只能力求寻找一条能够到达目的网络且较好的路由, 而并非寻找一条最佳路由, 所以 BGP 不需要像 RIP 那样周期性地和邻站交换路由信息。

02. 【解答】

1) 根据 RIP 算法, 首先将从 C 收到的路由信息的下一跳改为 C, 并且将每个距离都加 1, 得右表。

将题中表 2 与原路由表项进行比较, 根据更新路由表项的 规则: ①如果目的网络相同, 且下一跳路由器相同, 直接更新; ②如果是新的目的网络地址, 那么增加表项; ③若目的网络相同, 且下一跳路由器不同, 而距离更短, 则更新; ④否则, 无操作。更新后的路由表见下表。

目的网络	距离	下一跳
N2	16	C
N3	3	C
N4	9	C
N8	3	C
N7	5	C

目的网络	距离	下一跳路由器	目的网络	距离	下一跳路由器
N1	7	A	N6	8	F
N2	16	C	N7	5	C
N3	3	C	N8	3	C
N4	9	C	N9	4	D

2) 在更新后的路由表中, 路由器 B 到 N2 的距离为 16 (网络拓扑结构变化导致), 这意味着 N2 网络不可达, 这时路由器 B 应该丢弃该 IP 分组并向源主机报告目的不可达。

03. 【解答】

根据 Dijkstra 的最短路径算法, 加入结点的次序之一为 (R6, R5, R3, N3, R4, R1, R2, N4, N1, N2), 可以得到 R6 的路由表如下表所示。

目的网络	距离	下一跳路由器	目的网络	距离	下一跳路由器
N1	10	R3	N3	7	R3
N2	10	R3	N4	8	R3

04. 【解答】

1) 要求 R2 的路由表能到达图中的所有子网, 且路由项尽可能少, 则应对每个路由接口的子网进行聚合。在 AS1 中, 子网 153.14.5.0/25 和子网 153.14.5.128/25 可聚合为子网 153.14.5.0/24; 在 AS2 中, 子网 194.17.20.0/25 和子网 194.17.21.0/24 可聚合为子网 194.17.20.0/23; 子网 194.17.20.128/25 单独连接到 R2 的接口 E0。

于是可以得到 R2 的路由表如下：

目的 网 络	下 一 跳	接 口
153.14.5.0/24	153.14.3.2	S0
194.17.20.0/23	194.17.24.2	S1
194.17.20.128/25	—	E0

- 2) 该 IP 分组的目的 IP 地址 194.17.20.200 与路由表中 194.17.20.0/23 和 194.17.20.128/25 两个路由表项均匹配，根据最长匹配原则，R2 将通过 E0 接口转发该 IP 分组。
- 3) R1 和 R2 属于不同的自治系统，因此应使用边界网关协议（BGP 或 BGP4）交换路由信息；BGP 是应用层协议，它的报文被封装到 TCP 段中进行传输。

05. 【解答】

- 1) 因为题目要求路由表中的路由项尽可能少，所以这里可以把子网 192.1.6.0/24 和 192.1.7.0/24 聚合为子网 192.1.6.0/23，其他网络照常，可得到路由表如下：

目的 网 络	下 一 跳	接 口
192.1.1.0/24	—	E0
192.1.6.0/23	10.1.1.2	L0
192.1.5.0/24	10.1.1.10	L1

- 2) 通过查路由表可知：R1 通过 L0 接口转发该 IP 分组。因为该分组要经过 3 个路由器（R1、R2、R4），所以主机 192.1.7.211 收到的 IP 分组的 TTL 是 $64 - 3 = 61$ 。
- 3) R1 的 LSI 需要增加一条特殊的直连网络，网络前缀 Prefix 为 “0.0.0.0/0”，Metric 为 10。

4.6 IP 组播

4.6.1 组播的概念

为了能够支持像视频点播和视频会议这样的多媒体应用，网络必须实施某种有效的组播机制。使用多个单播传送来仿真组播总是可能的，但这会引起主机上大量的处理开销和网络上太多的交通量。人们所需要的组播机制是让源计算机一次发送的单个分组可以抵达用一个组地址标识的若干目标主机，并被它们正确接收。

组播一定仅应用于 UDP，它对将报文同时送往多个接收者的应用来说非常重要。而 TCP 是一个面向连接的协议，它意味着分别运行于两台主机（由 IP 地址来确定）内的两个进程（由端口号来确定）之间存在一条连接，因此会一对一直地发送。

使用组播的缘由是，有的应用程序要把一个分组发送给多个目的地主机。不是让源主机给每个目的地主机都发送一个单独的分组，而是让源主机把单个分组发送给一个组播地址，该组播地址标识一组地址。网络（如因特网）把这个分组的副本投递给该组中的每台主机。主机可以选择加入或离开一个组，因此一台主机可以同时属于多个组。

因特网中的 IP 组播也使用组播组的概念，每个组都有一个特别分配的地址，要给该组发送的计算机将使用这个地址作为分组的目标地址。在 IPv4 中，这些地址在 D 类地址空间中分配，而 IPv6 也有一部分地址空间保留给组播组。

主机使用一个称为 IGMP（因特网组管理协议）的协议加入组播组。它们使用该协议通知本地网络上的路由器关于要接收发送给某个组播组的分组的愿望。通过扩展路由器的路由选择和转发功能，可以在许多路由器互连的支持硬件组播的网络上面实现因特网组播。

需要注意的是，主机组播时仅发送一份数据，只有数据在传送路径出现分岔时才将分组复制后继续转发。因此，对发送者而言，数据只需发送一次就可发送到所有接收者，大大减轻了网络的负载和发送者的负担。组播需要路由器的支持才能实现，能够运行组播协议的路由器称为组播路由器。单播与组播的比较如图 4.13 所示。

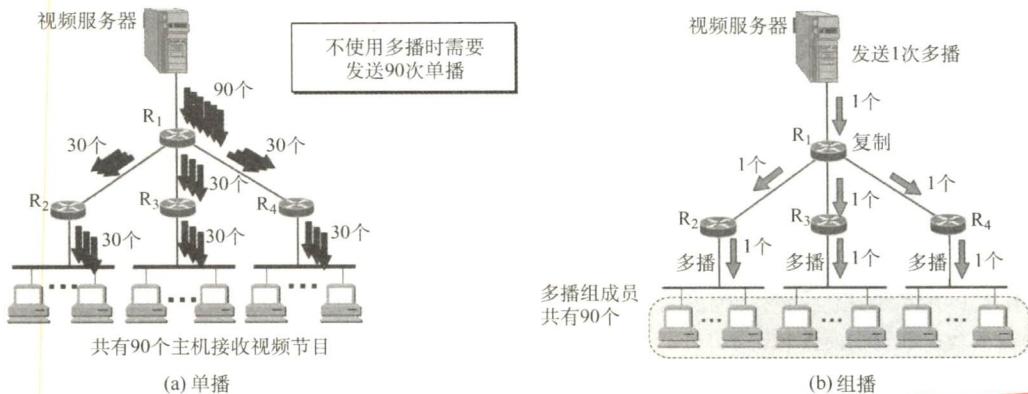


图 4.13 单播与组播的比较

4.6.2 IP 组播地址

IP 组播使用 D 类地址格式。D 类地址的前四位是 1110，因此 D 类地址范围是 224.0.0.0~239.255.255.255。每个 D 类 IP 地址标志一个组播组。

组播数据报和一般的 IP 数据报的区别是，前者使用 D 类 IP 地址作为目的地址，并且首部中的协议字段值是 2，表明使用 IGMP。需要注意的是：

- 1) 组播数据报也是“尽最大努力交付”，不提供可靠交付。
- 2) 组播地址只能用于目的地址，而不能用于源地址。
- 3) 对组播数据报不产生 ICMP 差错报文。因此，若在 PING 命令后面键入组播地址，将永远不会收到响应。
- 4) 并非所有的 D 类地址都可作为组播地址。

IP 组播可以分为两种：一种只在本局域网上进行硬件组播；另一种则在因特网的范围内进行组播。在因特网上进行组播的最后阶段，还是要把组播数据报在局域网上用硬件组播交付给组播组的所有成员〔见图 4.13(b)〕。下面讨论这种硬件组播。

IANA 拥有的以太网组播地址的范围是从 01-00-5E-00-00-00 到 01-00-5E-FF-FF-FF。不难看出，在每个地址中，只有 23 位可用作组播。这只能和 D 类 IP 地址中的 23 位有一一对应关系。D 类 IP 地址可供分配的有 28 位，可见在这 28 位中，前 5 位不能用来构成以太网的硬件地址，如图 4.14 所示。

例如，IP 组播地址 224.128.64.32（即 E0-80-40-20）和另一个 IP 组播地址 224.0.64.32（即 E0-00-40-20）转换成以太网的硬件组播地址都是 01-00-5E-00-40-20。由于组播 IP 地址与以太网硬件地址的映射关系不是唯一的，因此收到组播数据报的主机，还要在 IP 层利用软件进行过滤，把不是本主机要接收的数据报丢弃。

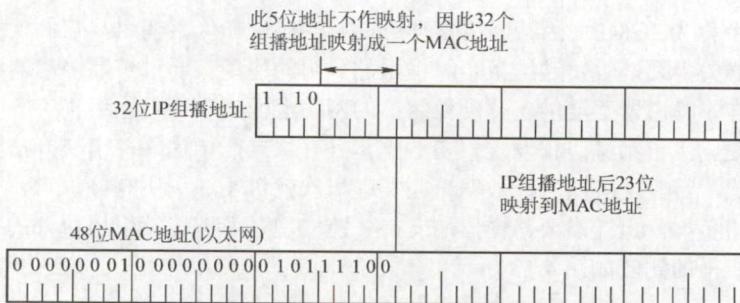


图 4.14 D类 IP 地址与以太网组播地址的映射关系

4.6.3 IGMP 与组播路由算法

要使路由器知道组播组成员的信息，需要利用因特网组管理协议（Internet Group Management Protocol, IGMP）。连接到局域网上的组播路由器还必须和因特网上的其他组播路由器协同工作，以便把组播数据报用最小代价传送给所有组成员，这就需要使用组播路由选择协议。

IGMP 并不是在因特网范围内对所有组播组成员进行管理的协议。IGMP 不知道 IP 组播组包含的成员数，也不知道这些成员分布在哪些网络上。IGMP 让连接到本地局域网上的组播路由器知道本局域网上是否有主机参加或退出了某个组播组。

IGMP 应视为网际协议 IP 的一个组成部分，其工作可分为两个阶段。

第一阶段：当某台主机加入新的组播组时，该主机应向组播组的组播地址发送一个 IGMP 报文，声明自己要成为该组的成员。本地的组播路由器收到 IGMP 报文后，将组成员关系转发给因特网上的其他组播路由器。

第二阶段：因为组成员关系是动态的，本地组播路由器要周期性地探询本地局域网上的主机，以便知道这些主机是否仍继续是组的成员。只要对某个组有一台主机响应，那么组播路由器就认为这个组是活跃的。但一个组在经过几次的探询后仍然没有一台主机响应时，则不再将该组的成员关系转发给其他的组播路由器。

组播路由选择实际上就是要找出以源主机为根结点的组播转发树，其中每个分组在每条链路上只传送一次（即在组播转发树上的路由器不会收到重复的组播数据报）。不同的多播组对应于不同的多播转发树；同一个组播组，对不同的源点也会有不同的多播转发树。

在许多由路由器互连的支持硬件多点传送的网络上实现因特网组播时，主要有三种路由算法：第一种是基于链路状态的路由选择；第二种是基于距离-向量的路由选择；第三种可以建立在任何路由器协议之上，因此称为协议无关的组播（PIM）。

4.6.4 本节习题精选

一、单项选择题

01. 以下关于组播概念的描述中，错误的是（ ）。
 - A. 在单播路由选择中，路由器只能从它的一个接口转发收到的分组
 - B. 在组播路由选择中，路由器可以从它的多个接口转发收到的分组
 - C. 用多个单播仿真一个组播时需要更多的带宽
 - D. 用多个单播仿真一个组播时延基本上是相同的
02. 在设计组播路由时，为了避免路由环路，（ ）。
 - A. 采用了水平分割技术
 - B. 构造组播转发树

- C. 采用了 IGMP D. 通过生存时间 (TTL) 字段
- 03.** 以太网组播 IP 地址 224.215.145.230 应该映射到的组播 MAC 地址是 ()。
- A. 01-00-5E-57-91-E6 B. 01-00-5E-D7-91-E6
 C. 01-00-5E-5B-91-E6 D. 01-00-5E-55-91-E6
- 04.** 下列地址中, () 是组播地址。
- A. 10.255.255.255 B. 228.47.32.45
 C. 192.32.44.59 D. 172.16.255.255

二、综合应用题

01. 因特网的组播是怎样实现的? 为什么因特网上的组播比以太网上的组播复杂得多?

4.6.5 答案与解析

一、单项选择题

01. D

多个单播可以仿真组播,但是一个组播所需要的带宽要小于多个单播带宽之和;用多个单播仿真一个组播时,路由器的时延将很大,而处理一个组播分组的时延是很小的。

02. B

由于树具有不存在环路的特性,因此构造一个组播转发树,通过该转发树既能将组播分组传送到组内的每台主机,又能避免环路[见图 4.13(b)]。水平分割用于避免距离-向量路由算法中的无穷计数问题。TTL 字段用于防止 IP 分组由于环路而在网络中无限循环。

03. A

以太网组播地址块的范围是 01-00-5E-00-00-00~01-00-5E-7F-FF-FF,而且在每个地址中,只有后 23 位可用组播。这样,只能和 D 类 IP 地址中的后 23 位有一一对应关系。D 类 IP 地址可供分配的有 28 位,可见这 28 位中的前 5 位不能用来构成以太网硬件地址。215 的二进制为 11010111,其中,在映射过程中最高位为 0,因此 215.145.230 映射的二进制为 01010111.10010001.11100110,对应的十六进制数是 57-91-E6。

04. B

组播地址使用点分十进制表示的范围是 224.0.0.0~239.255.255.255,这 4 个选项中,只有选项 B 在这个区间内。

二、综合应用题

01. 【解答】

因特网的组播是靠路由器来实现的,这些路由器必须增加一些能够识别组播的软件。能够运行组播协议的路由器可以是一个单独的路由器,也可以是运行组播软件的普通路由器。因特网上的组播比以太网上的组播复杂得多,因为以太网本身支持广播和组播,而因特网上当前的路由器和许多物理网络都不支持广播和组播。

关注公众号【乘龙考研】
一手更新稳定有保障

4.7 移动 IP

4.7.1 移动 IP 的概念

移动 IP 技术是指移动站以固定的网络 IP 地址实现跨越不同网段的漫游功能,并保证基于网络 IP 的网络权限在漫游过程中不发生任何改变。移动 IP 的目标是把分组自动地投递给移动站。

一个移动站是把其连接点从一个网络或子网改变到另一个网络或子网的主机。

移动 IP 定义了三种功能实体：移动节点、本地代理（也称归属代理）和外地代理。

1) 移动节点。具有永久 IP 地址的移动站。

2) 本地代理。通常就是连接在归属网络（原始连接到的网络）上的路由器。

3) 外地代理。通常就是连接在被访网络（移动到另一地点所接入的网络）上的路由器。

值得注意的是，某用户将笔记本关机后从家里带到办公室重新上网，在办公室能很方便地通过 DHCP 自动获取新的 IP 地址。虽然笔记本移动了，更换了地点及所接入的网络，但这并不是移动 IP。但如果我们在移动中进行 TCP 传输，在移动站漫游时，应一直保持这个 TCP 连接，否则移动站的 TCP 连接就会断断续续的。可见，若要使移动站在移动中的 TCP 连接不中断，就必须使笔记本的 IP 地址在移动中保持不变。这就是移动 IP 要研究的问题。

4.7.2 移动 IP 通信过程

用一个通俗的例子来描述移动 IP 的通信原理。例如，在以前科技不那么发达的年代，本科毕业时都将走向各自的工作岗位。由于事先并不知道自己未来的准确通讯地址，那么怎样继续和同学们保持联系呢？实际上也很简单。彼此留下各自的家庭地址（即永久地址）。毕业后若要和某同学联系，只要写信寄到该同学的永久地址，再请其家长把信件转交即可。

在移动 IP 中，每个移动站都有一个原始地址，即永久地址（或归属地址），移动站原始连接的网络称为归属网络。永久地址和归属网络的关联是不变的。归属代理通常是连接到归属网络上的路由器，然而它实现的代理功能是在应用层完成的。当移动站移动到另一地点，所接入的外地网络也称被访网络。被访网络中使用的代理称为外地代理，它通常是连接在被访网络上的路由器。外地代理有两个重要功能：①要为移动站创建一个临时地址，称为转交地址。转交地址的网络号显然和被访网络一致。②及时把移动站的转交地址告诉其归属代理。

移动 IP 技术的基本通信流程如下：

1) 移动站在归属网络时，按传统的 TCP/IP 方式进行通信。

2) 移动站漫游到外地网络时，向外地代理进行登记，以获得一个临时的转交地址。外地代理要向移动站的归属代理登记移动站的转交地址。

3) 归属代理知道移动站的转交地址后，会构建一条通向转交地址的隧道，将截获的发送给移动站的 IP 分组进行再封装，并通过隧道发送给被访网络的外地代理。

4) 外地代理把收到的封装的数据报进行拆封，恢复成原始的 IP 分组，然后发送给移动站，这样移动站在被访网络就能收到这些发送给它的 IP 分组。

5) 移动站在被访网络对外发送数据报时，仍然使用自己的永久地址作为数据报的源地址，此时显然无须通过 A 的归属代理来转发，而是直接通过被访网络的外部代理。

6) 移动站移动到另一外地网络时，在新外地代理登记后，然后新外地代理将移动站的新转交地址告诉其归属代理。无论如何移动，移动站收到的数据报都是由归属代理转发的。

7) 移动站回到归属网络时，移动站向归属代理注销转交地址。

请注意两点：转交地址是供移动站、归属代理及外地代理使用的，各种应用程序都不会使用。外地代理要向连接在被访网络上的移动站发送数据报时，直接使用移动站的 MAC 地址。

4.7.3 本节习题精选

单项选择题

01. 以下关于移动 IP 基本工作原理的描述中，错误的是（ ）。

- A. 移动 IP 的基本工作过程可以分为代理发现、注册、分组路由与注销 4 个阶段

- B. 结点在使用移动 IP 进行通信时，归属代理和外部代理之间需要建立一条隧道
 C. 移动结点到达新的网络后，通过注册过程把自己新的可达信息通知外部代理
 D. 移动 IP 的分组路由可以分为单播、广播与组播
02. 一台主机移动到了另一个 LAN 中，如果一个分组到达了它原来所在的 LAN 中，那么分组会被转发给（ ）。
 A. 移动 IP 的本地代理 B. 移动 IP 的外部代理
 C. 主机 D. 丢弃
03. 移动 IP 为移动主机设置了两个 IP 地址：主地址和辅地址，（ ）。
 A. 这两个地址都是固定的 B. 这两个地址随主机的移动而动态改变
 C. 主地址固定，辅地址动态改变 D. 主地址动态改变，辅地址固定
04. 如果一台主机的 IP 地址为 160.80.40.20/16，那么当它移动到了另一个不属于 160.80/16 子网的网络中时，它将（ ）。
 A. 可以直接接收和直接发送分组，没有任何影响
 B. 既不可以直接接收分组，也不可以直接发送分组
 C. 不可以直接发送分组，但可以直接接收分组
 D. 可以直接发送分组，但不可以直接接收分组

4.7.4 答案与解析

关注公众号【乘龙考研】
一手更新 稳定有保障

单项选择题

01. C

选项 C 把移动结点新的可达信息（转交地址）通知归属代理。这样，归属代理就可将发往移动结点的分组通过隧道转到转交地址（外部代理），再由外部代理交付给移动结点。

02. A

当一个分组到达用户的本地 LAN 时，它被转发给某一台与本地 LAN 相连的路由器。该路由器寻找目的主机，这时本地代理响应该请求，将这些分组封装到一些新 IP 分组的载荷，并将新分组发送给外部代理，外部代理将原分组解出来后，移交给移动后的主机。

03. C

移动主机在原始本地网时，获得的是主地址，当它移动到一个外地网络中时，需获得一个新的临时辅地址，主地址保持不变；当它移动到另一个外地网络或返回本地网络时，辅地址改变或撤销，而主地址仍然保持不变。选项 C 正确。

04. B

因为所有路由器都是按照子网来安排路由的，因此所有发往主机 160.80.40.20/16 的分组都会被发送到 160.80/16 子网中，当主机离开了这个子网时，自然就不能直接接收和直接发送分组，但可以通过转交地址来间接接收和发送分组。

4.8 网络层设备

4.8.1 冲突域和广播域

这里的“域”表示冲突或广播在其中发生并传播的区域。

1. 冲突域

冲突域是指连接到同一物理介质上的所有结点的集合，这些结点之间存在介质争用的现象。

在 OSI 参考模型中，冲突域被视为第 1 层概念，像集线器、中继器等简单无脑复制转发信号的第 1 层设备所连接的结点都属于同一个冲突域，也就是说它们不能划分冲突域。而第 2 层（网桥、交换机）、第 3 层（路由器）设备都可以划分冲突域。

2. 广播域

广播域是指接收同样广播消息的结点集合。也就是说，在该集合中的任何一个结点发送一个广播帧，其他能收到这个帧的结点都被认为是该广播域的一部分。在 OSI 参考模型中，广播域被视为第 2 层概念，像第 1 层（集线器等）、第 2 层（交换机等）设备所连接的结点都属于同一个广播域。而路由器，作为第 3 层设备，则可以划分广播域，即可以连接不同的广播域。

通常所说的局域网（LAN）特指使用路由器分割的网络，也就是广播域。

4.8.2 路由器的组成和功能

路由器是一种具有多个输入/输出端口的专用计算机，其任务是连接不同的网络（连接异构网络）并完成路由转发。在多个逻辑网络（即多个广播域）互连时必须使用路由器。

当源主机要向目标主机发送数据报时，路由器先检查源主机与目标主机是否连接在同一个网络上。如果源主机和目标主机在同一个网络上，那么直接交付而无须通过路由器。如果源主机和目标主机不在同一个网络上，那么路由器按照转发表（路由表）指出的路由将数据报转发给下一个路由器，这称为间接交付。可见，在同一个网络中传递数据无须路由器的参与，而跨网络通信必须通过路由器进行转发。例如，路由器可以连接不同的 LAN，连接不同的 VLAN，连接不同的 WAN，或者把 LAN 和 WAN 互连起来。路由器隔离了广播域。

从结构上看，路由器由路由选择和分组转发两部分构成，如图 4.15 所示。而从模型的角度看，路由器是网络层设备，它实现了网络模型的下三层，即物理层、数据链路层和网络层。

注意：如果一个存储转发设备实现了某个层次的功能，那么它就可以互连两个在该层次上使用不同协议的网段（网络）。如网桥实现了物理层和数据链路层，那么网桥可以互连两个物理层和数据链路层不同的网段；但中继器实现了物理层后，却不能互连两个物理层不同的网段，这是因为中继器不是存储转发设备，它属于直通式设备。

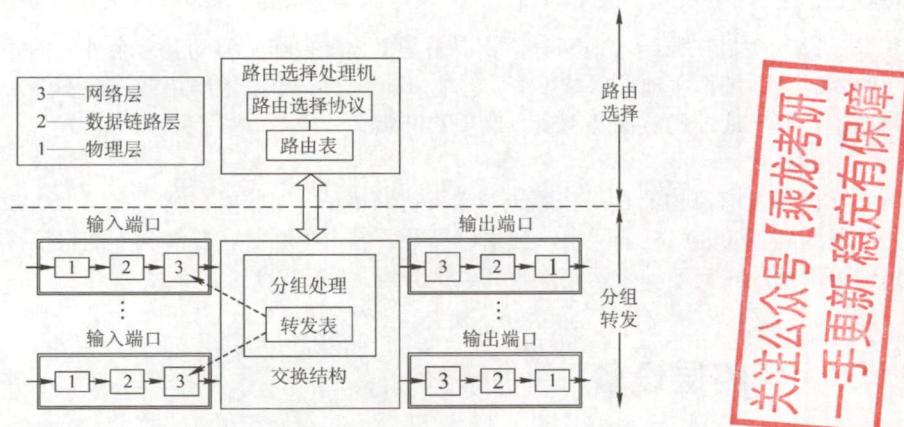


图 4.15 路由器体系结构

路由选择部分也称控制部分，其核心构件是路由选择处理机。路由选择处理机的任务是根据所选定的路由选择协议构造出路由表，同时经常或定期地和相邻路由器交换路由信息而不断更新和维护路由表。

分组转发部分由三部分组成：交换结构、一组输入端口和一组输出端口。输入端口在从物理

层接收到的比特流中提取出数据链路层帧，进而从帧中提取出网络层数据报，输出端口则执行恰好相反的操作。交换结构是路由器的关键部件，它根据转发表对分组进行处理，将某个输入端口进入的分组从一个合适的输出端口转发出去。有三种常用的交换方法：通过存储器进行交换、通过总线进行交换和通过互联网络进行交换。交换结构本身就是一个网络。

路由器主要完成两个功能：一是分组转发，二是路由计算。前者处理通过路由器的数据流，关键操作是转发表查询、转发及相关的队列管理和任务调度等；后者通过和其他路由器进行基于路由协议的交互，完成路由表的计算。

路由器和网桥的重要区别是：网桥与高层协议无关，而路由器是面向协议的，它依据网络地址进行操作，并进行路径选择、分段、帧格式转换、对数据报的生存时间和流量进行控制等。现今的路由器一般都提供多种协议的支持，包括 OSI、TCP/IP、IPX 等。

4.8.3 路由表与路由转发

路由表是根据路由选择算法得出的，主要用途是路由选择。从历年统考真题可以看出，标准的路由表有 4 个项目：目的网络 IP 地址、子网掩码、下一跳 IP 地址、接口。在如图 4.16 所示的网络拓扑中，R1 的路由表见表 4.4，该路由表包含到互联网的默认路由。

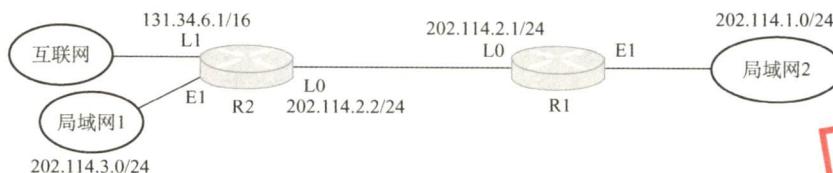


图 4.16 一个简单的网络拓扑

表 4.4 R1 的路由表

目的网络 IP 地址	子网掩码	下一跳 IP 地址	接口
202.114.1.0	255.255.255.0	Direct	E1
202.114.2.0	255.255.255.0	Direct	L0
202.114.3.0	255.255.255.0	202.114.2.2	L0
0.0.0.0	0.0.0.0	202.114.2.2	L0

转发表是从路由表得出的，其表项和路由表项有直接的对应关系。但转发表的格式和路由表的格式不同，其结构应使查找过程最优化（而路由表则需对网络拓扑变化的计算最优化）。转发表中含有一个分组将要发往的目的地址，以及分组的下一跳（即下一步接收者的目的地址，实际为 MAC 地址）。为了减少转发表的重复项目，可以使用一个默认路由代替所有具有相同“下一跳”的项目，并将默认路由设置得比其他项目的优先级低，如图 4.17 所示。路由表总是用软件来实现的；转发表可以用软件来实现，甚至也可以用特殊的硬件来实现。

目的站	下一跳
1	直接
2	3
3	2
4	3

(a) 未使用默认路由

目的站	下一跳
1	直接
3	2
默认	3

(b) 使用了默认路由

图 4.17 未使用默认路由的转发表和使用了默认路由的转发表的对比

注意转发和路由选择的区别：“转发”是路由器根据转发表把收到的 IP 数据报从合适的端口转发出去，它仅涉及一个路由器。而“路由选择”则涉及很多路由器，路由表是许多路由器协同工作的结果。这些路由器按照复杂的路由算法，根据从各相邻路由器得到的关于网络拓扑的变化情况，动态地改变所选择的路由，并由此构造出整个路由表。

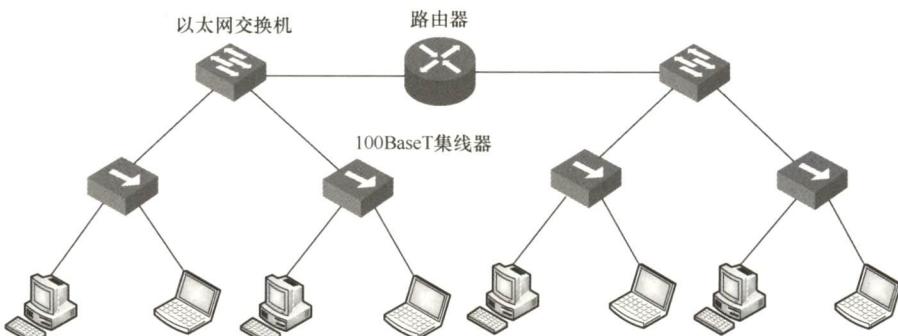
注意，在讨论路由选择的原理时，往往不去区分转发表和路由表的区别，但要注意路由表不等于转发表。分组的实际转发是靠直接查找转发表，而不是直接查找路由表。

4.8.4 本节习题精选

一、单项选择题

01. 要控制网络上的广播风暴，可以采用的方法是（ ）。
 - A. 用网桥将网络分段
 - B. 用路由器将网络分段
 - C. 将网络转换成 10Base-T
 - D. 用网络分析仪跟踪正在发送广播信息的计算机
02. 一个局域网与在远处的另一个局域网互连，则需要用到（ ）。
 - A. 物理通信介质和集线器
 - B. 网间连接器和集线器
 - C. 路由器和广域网技术
 - D. 广域网技术
03. 路由器主要实现（ ）的功能。
 - A. 数据链路层、网络层与应用层
 - B. 网络层与传输层
 - C. 物理层、数据链路层与网络层
 - D. 物理层与网络层
04. 关于路由器的下列说法中，正确的是（ ）。
 - A. 路由器处理的信息量比交换机少，因而转发速度比交换机快
 - B. 对于同一目标，路由器只提供延迟最小的最佳路由
 - C. 通常的路由器可以支持多种网络层协议，并提供不同协议之间的分组转发
 - D. 路由器不但能够根据 IP 地址进行转发，而且可以根据物理地址进行转发
05. 下列关于路由器交付的说法中，错误的是（ ）。
 - I. 路由选择分直接交付和间接交付
 - II. 直接交付时，两台机器可以不在同一物理网段内
 - III. 间接交付时，不涉及直接交付
 - IV. 直接交付时，不涉及路由器
 - A. I 和 II
 - B. II 和 III
 - C. III 和 IV
 - D. I 和 IV
06. (未使用 CIDR) 当一个 IP 分组进行直接交付时，要求发送方和目的站具有相同的（ ）。
 - A. IP 地址
 - B. 主机号
 - C. 端口号
 - D. 子网地址
07. 一个路由器的路由表通常包含（ ）。
 - A. 需要包含到达所有主机的完整路径信息
 - B. 需要包含所有到达目的网络的完整路径信息
 - C. 需要包含到达目的网络的下一跳路径信息
 - D. 需要包含到达所有主机的下一跳路径信息
08. 决定路由器转发表中的值的算法是（ ）。
 - A. 指数回退算法
 - B. 分组调度算法
 - C. 路由算法
 - D. 拥塞控制算法
09. 路由器中计算路由信息的是（ ）。
 - A. 输入队列
 - B. 输出队列
 - C. 交换结构
 - D. 路由选择处理机
10. 路由表的分组转发部分由（ ）组成。

- A. 交换结构 B. 输入端口 C. 输出端口 D. 以上都是
11. 路由器的路由选择部分包括()。
 A. 路由选择处理机 B. 路由选择协议
 C. 路由表 D. 以上都是
12. 在下列网络设备中, 传输延迟时间最大的是()。
 A. 局域网交换机 B. 网桥 C. 路由器 D. 集线器
13. 在路由表中设置一条默认路由, 则其目的地址和子网掩码应分别置为()。
 A. 192.168.1.1、255.255.255.0 B. 127.0.0.0、255.0.0.0
 C. 0.0.0.0、0.0.0.0 D. 0.0.0.0、255.255.255.255
14. 【2010统考真题】下列网络设备中, 能够抑制广播风暴的是()。
 I 中继器 II 集线器 III 网桥 IV 路由器
 A. 仅I和II B. 仅III C. 仅III和IV D. 仅IV
15. 【2011统考真题】某网络拓扑如下图所示, 路由器R1只有到达子网192.168.1.0/24的路由。为使R1可以将IP分组正确地路由到图中的所有子网, 则在R1中需要增加的一条路由(目的网络, 子网掩码, 下一跳)是()。
-
- A. 192.168.2.0 255.255.255.128 192.168.1.1
 B. 192.168.2.0 255.255.255.0 192.168.1.1
 C. 192.168.2.0 255.255.255.128 192.168.1.2
 D. 192.168.2.0 255.255.255.0 192.168.1.2
16. 【2012统考真题】下列关于IP路由器功能的描述中, 正确的是()。
 I. 运行路由协议, 设备路由表
 II. 监测到拥塞时, 合理丢弃IP分组
 III. 对收到的IP分组头进行差错校验, 确保传输的IP分组不丢失
 IV. 根据收到的IP分组的目的IP地址, 将其转发到合适的输出线路上
 A. 仅III、IV B. 仅I、II、III C. 仅I、II、IV D. I、II、III、IV
17. 【2020统考真题】下图所示的网络中, 冲突域和广播域的个数分别是()。

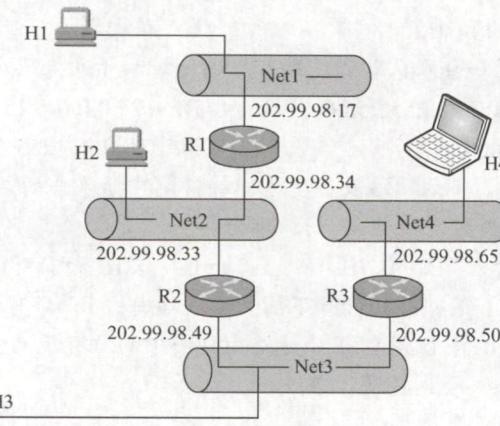


关注公众号【乘龙考研】
一手更新 稳定有保障

- A. 2, 2 B. 2, 4 C. 4, 2 D. 4, 4

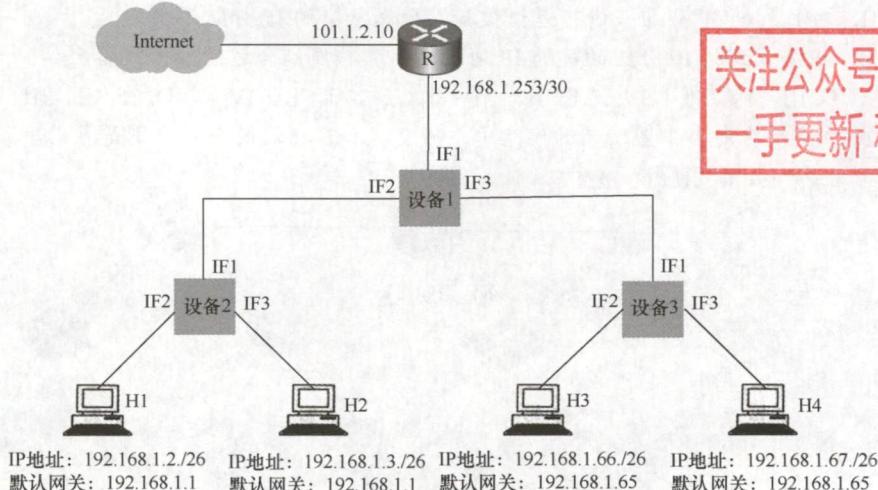
二、综合应用题

01. 某个单位的网点由 4 个子网组成，结构如下图所示，其中主机 H1、H2、H3 和 H4 的 IP 地址和子网掩码见下表。



主 机	IP 地址	子 网 掩 码
H1	202.99.98.18	255.255.255.240
H2	202.99.98.35	255.255.255.240
H3	202.99.98.51	255.255.255.240
H4	202.99.98.66	255.255.255.240

- 1) 请写出路由器 R1 到 4 个子网的路由表。
 2) 试描述主机 H1 发送一个 IP 数据报到主机 H2 的过程（包括物理地址解析过程）。
 02. 试简述路由器的路由功能和转发功能。
 03. 【2019 统考真题】某网络拓扑如下图所示，其中 R 为路由器，主机 H1 ~ H4 的 IP 地址配置以及 R 的各接口 IP 地址配置如图中所示。现有若干以太网交换机（无 VLAN 功能）和路由器两类网络互连设备可供选择。



关注公众号【乘龙考研】
一手更新 稳定有保障

请回答下列问题：

- 1) 设备 1、设备 2 和设备 3 分别应选择什么类型的网络设备？
- 2) 设备 1、设备 2 和设备 3 中，哪几个设备的接口需要配置 IP 地址？为对应的接口配置正确的 IP 地址。
- 3) 为确保主机 H1 ~ H4 能够访问 Internet，R 需要提供什么服务？
- 4) 若主机 H3 发送一个目的地址为 192.168.1.127 的 IP 数据报，网络中哪几个主机会接收该数据报？

4.8.5 答案与解析

一、单项选择题



01. B

网桥和交换机是第二层设备，能够分割冲突域，但不能分割广播域。路由器是第三层设备，不转发全网广播（目的地 255.255.255.255），因此可以分割广播域。

02. C

局域网的互连需要路由器作为连接设备，同时是远程的局域网，因此要用到广域网技术。

03. C

路由器是网络层设备，所以它也必须要处理网络层以下的功能，即物理层和数据链路层。而传输层和应用层是网络层之上的，它们使用网络层的接口，路由器不实现它们的功能。

04. C

路由器是第三层设备，要处理的内容比第二层设备交换机更多，因而转发速度比交换机慢，选项 A 错误。虽然一些路由协议也将延迟等作为参数进行路由选择，但路由协议使用得最多的参数是传输距离，此外还有一些其他参数，选项 B 错误。路由器只能根据 IP 地址进行转发，选项 D 错误。

05. B

路由选择分为直接交付和间接交付，当发送站与目的站在同一网段内时，就使用直接交付，反之使用间接交付，因此 I 正确、II 错误。间接交付的最后一个路由器肯定直接交付，III 错误。直接交付在同一网段内，因此不涉及路由器，IV 正确。

06. D

判断一个 IP 分组的交付方式是直接交付还是间接交付，路由器需要根据分组的目的 IP 地址和该路由器接收端口的 IP 地址是否属于同一个子网来进行判断。具体来说，将该分组的源 IP 地址和目的 IP 地址分别与子网掩码进行“与”操作，如果得到的子网地址相同，那么该分组就采用直接交付方式，否则采用间接交付方式。

07. C

路由表中包含到目的网络的下一跳路径信息。由路由表表项的组成也不难得出正确答案为选项 C。路由表也不可能包含到达所有主机的下一跳信息，否则路由转发将是不可想象的。

08. C

由于转发表是根据路由表生成的，而路由表又是由路由算法得到的，因此路由算法决定了转发表中的值。

09. D

路由选择处理机的任务是根据所选定的路由选择协议构造路由表，同时经常或定期地与相邻路由器交换路由信息而不断地更新和维护路由表。

10. D

分组转发部分包括 3 部分：①交换结构，根据转发表对分组进行处理，将某个输入端口进入

的分组从一个合适的输出端口转发出去。②输入端口，包括物理层、数据链路层和网络层的处理模块。③输出端口，负责从交换结构接收分组，再将其发送到路由器外面的线路上。

11. D

路由器的路由选择部分包括 3 部分：①路由选择处理机，它根据所选定的路由选择协议构造路由表，同时和相邻路由器交换路由信息。②路由选择协议，用来更新路由表的算法。③路由表，它是根据路由算法得出的，一般包括从目的网络到下一跳的映射。

12. C

由于路由器是网络层设备，在路由器上实现了物理层、数据链路层和网络层的功能，因此路由器的传输延迟时间最长。

13. C

路由表中默认路由的目的地址和子网掩码都是 0.0.0.0。

14. D

中继器和集线器工作在物理层，既不隔离冲突域也不隔离广播域。为了解决冲突域的问题，人们利用网桥和交换机来分隔互联网的各个网段中的通信量，建立多个分离的冲突域，但当网桥和交换机接收到一个未知转发信息的数据帧时，为了保证该帧能被目的结点正确接收，将该帧从所有的端口广播出去，可以看出网桥和交换机的冲突域等于端口个数，广播域为 1。路由器可以隔离广播域和冲突域，要屏蔽数据链路层的广播帧，当然应该是网络层设备路由器。在此题的选项中，路由器是其中最高层的网络设备，其他设备能隔离的，路由器一定能隔离。

15. D

要使 R1 能够正确地将分组路由到所有子网，R1 中需要有到 192.168.2.0/25 和 192.168.2.128/25 的路由，分别转换成二进制如下：

192.168.2.0: 11000000 10101000 00000010 00000000

192.168.2.128: 11000000 10101000 00000010 10000000

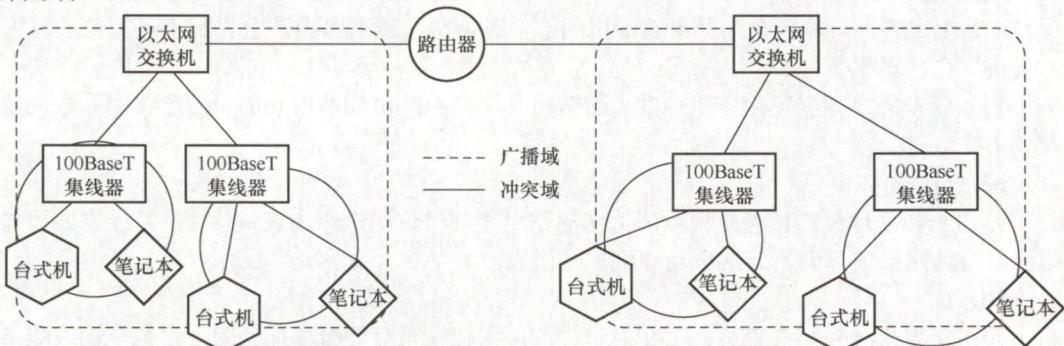
前 24 位都是相同的，于是可以聚合成超网 192.168.2.0/24，子网掩码为前 24 位，即 255.255.255.0。下一跳是与 R1 直接相连的 R2 的地址，因此是 192.168.1.2。

16. C

I 和 IV 显然是 IP 路由器的功能。对于 II，当路由器监测到拥塞时，可合理丢弃 IP 分组，并向发出该 IP 分组的源主机发送一个源点抑制的 ICMP 报文。对于 III，路由器对收到的 IP 分组首部进行差错检验，丢弃有差错首部的报文，但不保证 IP 分组不丢失。

17. C

网络层设备路由器可以隔离广播域和冲突域；数据链路层设备普通交换机只能隔离冲突域；物理层设备集线器、中继器既不能隔离冲突域，又不能隔离广播域。因此，题中共有 2 个广播域，4 个冲突域。



关注公众号【乘龙考研】
一手更新 稳定有保障

二、综合应用题

01. 【解答】

- 1) 将 H1、H2、H3、H4 的 IP 地址分别与它们的子网掩码进行“与”操作，可得到 4 个子网的网络地址，分别为 202.99.98.16、202.99.98.32、202.99.98.48、202.99.98.64，因此路由器 R1 到 4 个子网的路由表见下表。

目的 网络	子 网 掩 码	下 一 跳
202.99.98.16	255.255.255.240	直接
202.99.98.32	255.255.255.240	直接
202.99.98.48	255.255.255.240	202.99.98.33
202.99.98.64	255.255.255.240	202.99.98.33

- 2) 主机 H1 向主机 H2 发送一个 IP 数据报的过程如下：

- ① 主机 H1 首先构造一个源 IP 地址为 202.99.98.18、目的 IP 地址为 202.99.98.35 的 IP 数据报，主机 H1 先把本子网的子网掩码与 H2 的 IP 地址逐位相与，所得结果不等于 H1 的网络地址，因此 H1 与 H2 不在同一子网，无法直接交付，然后将该数据报传送给数据链路层。
- ② 主机 H1 通过 ARP 获得路由器 R1（202.99.98.17）对应的 MAC 地址，并将其作为目的 MAC 地址，将 H1 的 MAC 地址作为源 MAC 地址填入封装有 IP 数据报的帧，然后将该帧发送出去。
- ③ 路由器 R1 收到该帧后，去除帧头与帧尾，得到 IP 数据报，然后根据 IP 数据报中的目的 IP 地址（202.99.98.35）去查找路由表，得到下一跳地址为直接相连。
- ④ 路由器 R1 通过 ARP 得到主机 H2 的 MAC 地址，并将其作为目的 MAC 地址，将 R1 的 MAC 地址作为源 MAC 地址填入封装有 IP 数据报的帧，然后将该帧发送到子网 Net2 上。
- ⑤ 主机 H2 将收到的帧，去除帧头与帧尾，并最终得到从主机 H1 发来的 IP 数据报。

注意：在②中（发出的帧），帧的目的地 MAC 地址为默认网关的 MAC 地址；在④中（接收的帧），帧的源 MAC 地址为默认网关的 MAC 地址。

02. 【解答】

转发即当一个分组到达时所采取的动作。在路由器中，每个分组到达时对它进行处理，它在路由表中查找分组所对应的输出线路。通过查得的结果，将分组发送到正确的线路上。

路由算法是网络层软件的一部分，它负责确定一个进来的分组应该被传送到哪条输出线路上。路由算法负责填充和更新路由表，转发功能则根据路由表的内容来确定当每个分组到来时应该采取什么动作（如从哪个端口转发出去）。

03. 【解答】

- 1) 以太网交换机（无 VLAN 功能）连接的若干 LAN 仍然是一个网络（同一个广播域），路由器可以连接不同的 LAN、不同的 WAN 或把 WAN 和 LAN 互连起来，隔离了广播域。IP 地址 192.168.1.2/26 与 192.168.1.3/26 的网络前缀均为 192.168.1.0，视为 LAN1。IP 地址 192.168.1.66/26 与 192.168.1.67/26 的网络前缀均为 192.168.1.64，视为 LAN2。所以设备 1 为路由器，设备 2、3 为以太网交换机。
- 2) 设备 1 为路由器，其接口应配置 IP 地址。IF1 接口与路由器 R 相连，其相连接口的 IP 地

址为 192.168.1.253/30, 253 的二进制表示形式为 11111101, 因此 IF1 接口的网络前缀也应为 192.168.1.111111, 已分配 192.168.1.253, 去除全 0 全 1, IF1 接口的 IP 地址应为 192.168.1.254。LAN1 的默认网关为 192.168.1.1, LAN2 的默认网关为 192.168.1.65, 网关的 IP 地址是具有路由功能的设备的 IP 地址, 通常默认网关地址就是路由器中的 LAN 端口地址, 设备 1 的 IF2、IF3 接口的 IP 地址分别设置为 192.168.1.1 和 192.168.1.65。

- 3) 私有地址段: C 类 192.168.0.0~192.168.255.255, 即 H1~H4 均为私有 IP 地址, 若要能够访问 Internet, R 需要提供 NAT 服务, 即网络地址转换服务。
- 4) 主机 H3 发送一个目的地址为 192.168.1.127 的 IP 数据报, 主机号全为 1, 为本网络的广播地址, 由于路由器可以隔离广播域, 只有主机 H4 会接收到数据报。

4.9 本章小结及疑难点

1. “尽最大努力交付”有哪些含义?

- 1) 不保证源主机发送的 IP 数据报一定无差错地交付到目的主机。
- 2) 不保证源主机发送的 IP 数据报都在某一规定的时间内交付到目的主机。
- 3) 不保证源主机发送的 IP 数据报一定按发送时的顺序交付到目的主机。
- 4) 不保证源主机发送的 IP 数据报不会重复交付给目的主机。
- 5) 不故意丢弃 IP 数据报。丢弃 IP 数据报的情况是: 路由器检测出首部校验和有错误; 或由于网络中通信量过大, 路由器或目的主机中的缓存已无空闲空间。

但要注意, IP 数据报的首部中有一个“首部校验和”。当它检验出 IP 数据报的首部出现了差错时, 就丢弃该数据报。因此, 凡交付给目的主机的 IP 数据报都是 IP 首部没有差错的或没有检测出差错的。也就是说, 在传输过程中, 出现差错的 IP 数据报都被丢弃了。

现在因特网上绝大多数的通信量都属于“尽最大努力交付”。如果数据必须可靠地交付给目的地, 那么使用 IP 的高层软件必须负责解决这一问题。

2. “IP 网关”和“IP 路由器”是否为同义语? “互连网”和“互联网”有没有区别?

当初发明 TCP/IP 的研究人员使用 IP Gateway 作为网际互连的设备, 可以认为“IP 网关”和“IP 路由器”是同义词。“互连网”和“互联网”都是推荐名, 都可以使用, 不过建议优先使用“互联网”。

3. 在一个互联网中, 能否用一个很大的交换机(switch)来代替互联网中很多的路由器?

不行。交换机和路由器的功能是不相同的。

交换机可在单个网络中与若干计算机相连, 并且可以将一台计算机发送过来的帧转发给另一台计算机。从这一点上看, 交换机具有集线器的转发帧的功能, 但交换机比集线器的功能强很多。在同一时间, 集线器只允许一台计算机发送数据。

路由器连接两个或多个同构的或异构的网络, 在网络之间转发分组(即 IP 数据报)。因此, 如果许多相同类型的网络互连时, 那么用一个很大的交换机(如果能够找其他计算机进行通信, 交换机允许找得到)代替原来的一些路由器是可行的。但若这些互连的网络是异构的网络, 那么就必须使用路由器来进行互连。

4. 网络前缀是指网络号字段(net-id)中前面的几个类别位还是指整个的网络号字段?

是指整个的网络号字段, 包括最前面的几个类别位在内。网络前缀常常简称为前缀。例如一

个B类地址10100000 00000000 00000000 00010000，其类别位就是最前面的两位：10，而网络前缀就是前16位：10100000 00000000。

5. IP有分片的功能，但广域网中的分组则不必分片，这是为什么？

IP数据报可能要经过许多个网络，而源结点事先并不知道数据报后面要经过的这些网络所能通过的分组的最大长度是多少。等到IP数据报转发到某个网络时，中间结点可能才发现数据报太长了，因此在这时就必须进行分片。但广域网能够通过的分组的最大长度是该广域网中所有结点都事先知道的，源结点不可能发送网络不支持的过长分组。因此广域网没有必要将已经发送出的分组再进行分片。

6. 数据链路层广播和IP广播有何区别？

数据链路层广播是用数据链路层协议（第二层）在一个以太网上实现的对该局域网上的所有主机进行广播MAC帧，而IP广播则是用IP通过因特网实现的对一个网络（即目的网络）上的所有主机进行广播IP数据报。

7. 主机在接收一个广播帧或组播帧时，其CPU所要做的事情有何区别？

在接收广播帧时，主机通过其适配器〔即网络接口卡（NIC）〕接收每个广播帧，然后将其传递给操作系统。CPU执行协议软件，并界定是否接收和处理该帧。在接收组播帧时，CPU要对适配器进行配置，而适配器根据特定的组播地址表来接收帧。凡与此组播地址表不匹配的帧都将被NIC丢弃。因此在组播的情况下，是适配器NIC而不是CPU决定是否接收一个帧。

8. 假定在一个局域网中计算机A发送ARP请求分组，希望找出计算机B的硬件地址。这时局域网上的所有计算机都能收到这个广播发送的ARP请求分组。试问这时由哪个计算机使用ARP响应分组将计算机B的硬件地址告诉计算机A？

这要区分两种情况。第一，如果计算机B和计算机A都连接在同一个局域网上，那么就是计算机B发送ARP响应分组。第二，如果计算机B和计算机A不连接在同一个局域网上，那么就必须由一个连接计算机A所在局域网的路由器来转发ARP请求分组。这时，该路由器向计算机A发送ARP回答分组，给出自己的硬件地址。

9. 路由器实现了物理层、数据链路层、网络层，这句话的含义是什么？

第1章中提到了网络中的两个通信结点利用协议栈进行通信的过程。发送方一层一层地把数据“包装”，接收方一层一层地把“包装”拆开，最后上交给用户。路由器实现了物理层，数据链路层和网络层的含义是指路由器有能力对这三层协议的控制信息进行识别、分析以及转换，直观的理解是路由器有能力对数据“包装”这三层协议或者“拆开”这三层协议。自然，路由器就有能力互连这三层协议不同的两个网络。



第 5 章 传输层

关注公众号【乘龙考研】
一手更新 稳定有保障

【考纲内容】

(一) 传输层提供的服务

传输层的功能；传输层寻址与端口；无连接服务和面向连接服务

(二) UDP

UDP 数据报；UDP 校验

(三) TCP

TCP 段；TCP 连接管理；TCP 可靠传输；TCP 流量控制与拥塞控制

扫一扫



视频讲解

【复习提示】

传输层是整个网络体系结构中的关键层次。要求掌握传输层在计算机网络中的地位、功能、工作方式及原理等，掌握 UDP 及 TCP（如首部格式、可靠传输、流量控制、拥塞控制、连接管理等）。其中，TCP 报文分析、流量控制与拥塞控制机制，出选择题、综合题的概率均较大，因此要将其工作原理透彻掌握，以便能在具体的题目中灵活运用。

5.1 传输层提供的服务

5.1.1 传输层的功能

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

传输层位于网络层之上，它为运行在不同主机上的进程之间提供了逻辑通信，而网络层提供主机之间的逻辑通信。显然，即使网络层协议不可靠（网络层协议使分组丢失、混乱或重复），传输层同样能为应用程序提供可靠的服务。

从图 5.1 可以看出，网络的边缘部分的两台主机使用网络核心部分的功能进行端到端的通信时，只有主机的协议栈才有传输层和应用层，而路由器在转发分组时都只用到下三层的功能（即在通信子网中没有传输层，传输层只存在于通信子网以外的主机中）。

传输层的功能如下：

- 1) 传输层提供应用进程之间的逻辑通信（即端到端的通信）。与网络层的区别是，网络层提供的是主机之间的逻辑通信。

从网络层来说，通信的双方是两台主机，IP 数据报的首部给出了这两台主机的 IP 地址。但“两台主机之间的通信”实际上是两台主机中的应用进程之间的通信，应用进程之间的通信又称端到端的逻辑通信。这里“逻辑通信”的意思是：传输层之间的通信好像是沿水平方向传送数据，但事实上这两个传输层之间并没有一条水平方向的物理连接。

- 2) 复用和分用。复用是指发送方不同的应用进程都可使用同一个传输层协议传送数据；分用是指接收方的传输层在剥去报文的头部后能够把这些数据正确交付到目的应用进程。

注意：网络层也有复用分用的功能，但网络层的复用是指发送方不同协议的数据都可以封装成IP数据报发送出去，分用是指接收方的网络层在剥去首部后把数据交付给相应的协议。

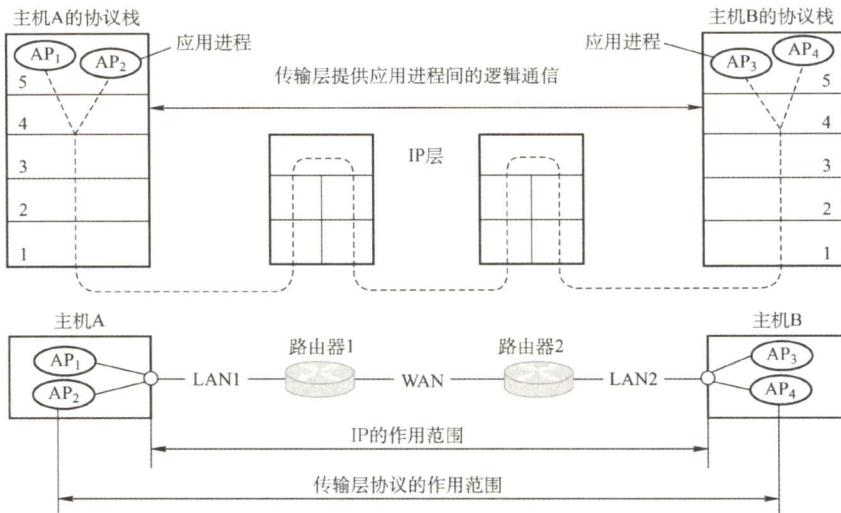


图 5.1 传输层为相互通信的进程提供逻辑通信

- 3) 传输层还要对收到的报文进行差错检测（首部和数据部分）。而网络层只检查IP数据报的首部，不检验数据部分是否出错。
- 4) 提供两种不同的传输协议，即面向连接的TCP和无连接的UDP。而网络层无法同时实现两种协议（即在网络层要么只提供面向连接的服务，如虚电路；要么只提供无连接服务，如数据报，而不可能在网络层同时存在这两种方式）。

传输层向高层用户屏蔽了低层网络核心的细节（如网络拓扑、路由协议等），它使应用进程看见的是在两个传输层实体之间好像有一条端到端的逻辑通信信道，这条逻辑通信信道对上层的表现却因传输层协议不同而有很大的差别。当传输层采用面向连接的TCP时，尽管下面的网络是不可靠的（只提供尽最大努力的服务），但这种逻辑通信信道就相当于一条全双工的可靠信道。但当传输层采用无连接的UDP时，这种逻辑通信信道仍然是一条不可靠信道。

5.1.2 传输层的寻址与端口

1. 端口的作用

端口能够让应用层的各种应用进程将其数据通过端口向下交付给传输层，以及让传输层知道应当将其报文段中的数据向上通过端口交付给应用层相应的进程。端口是传输层服务访问点（TSAP），它在传输层的作用类似于IP地址在网络层的作用或MAC地址在数据链路层的作用，只不过IP地址和MAC地址标识的是主机，而端口标识的是主机中的应用进程。

数据链路层的SAP是MAC地址，网络层的SAP是IP地址，传输层的SAP是端口。

在协议栈层间的抽象的协议端口是软件端口，它与路由器或交换机上的硬件端口是完全不同的概念。硬件端口是不同硬件设备进行交互的接口，而软件端口是应用层的各种协议进程与传输实体进行层间交互的一种地址。传输层使用的是软件端口。

2. 端口号

应用进程通过端口号进行标识，端口号长度为 16bit，能够表示 65536 (2^{16}) 个不同的端口号。端口号只具有本地意义，即端口号只标识本计算机应用层中的各进程，在因特网中不同计算机的相同端口号是没有联系的。根据端口号范围可将端口分为两类：

- 1) 服务器端使用的端口号。它又分为两类，最重要的一类是熟知端口号，数值为 0~1023，IANA（互联网地址指派机构）把这些端口号指派给了 TCP/IP 最重要的一些应用程序，让所有的用户都知道。另一类称为登记端口号，数值为 1024~49151。它是供没有熟知端口号的应用程序使用的，使用这类端口号必须在 IANA 登记，以防止重复。
一些常用的熟知端口号如下：

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

- 2) 客户端使用的端口号，数值为 49152~65535。由于这类端口号仅在客户进程运行时才动态地选择，因此又称短暂端口号（也称临时端口）。通信结束后，刚用过的客户端口号就不再存在，从而这个端口号就可供其他客户进程以后使用。

3. 套接字

在网络中通过 IP 地址来标识和区别不同的主机，通过端口号来标识和区分一台主机中的不同应用进程，端口号拼接到 IP 地址即构成套接字 Socket。在网络中采用发送方和接收方的套接字来识别端点。套接字，实际上是一个通信端点，即

$$\text{套接字 } \text{Socket} = (\text{IP 地址: 端口号})$$

它唯一地标识网络中的一台主机和其上的一个应用（进程）。

在网络通信中，主机 A 发给主机 B 的报文段包含目的端口号和源端口号，源端口号是“返回地址”的一部分，即当 B 需要发回一个报文段给 A 时，B 到 A 的报文段中的目的端口号便是 A 到 B 的报文段中的源端口号（完全的返回地址是 A 的 IP 地址和源端口号）。

5.1.3 无连接服务与面向连接服务

面向连接服务就是在通信双方进行通信之前，必须先建立连接，在通信过程中，整个连接的情况一直被实时地监控和管理。通信结束后，应该释放这个连接。

无连接服务是指两个实体之间的通信不需要先建立好连接，需要通信时，直接将信息发送到“网络”中，让该信息的传递在网上尽力而为地往目的地传送。

TCP/IP 协议族在 IP 层之上使用了两个传输协议：一个是面向连接的传输控制协议（TCP），采用 TCP 时，传输层向上提供的是一条全双工的可靠逻辑信道；另一个是无连接的用户数据报协议（UDP），采用 UDP 时，传输层向上提供的是一条不可靠的逻辑信道。

TCP 提供面向连接的服务，在传送数据之前必须先建立连接，数据传送结束后要释放连接。TCP 不提供广播或组播服务。由于 TCP 提供面向连接的可靠传输服务，因此不可避免地增加了许多开销，如确认、流量控制、计时器及连接管理等。这不仅使协议数据单元的头部增大很多，还要占用许多的处理机资源。因此 TCP 主要适用于可靠性更重要的场合，如文件传输协议（FTP）、超文本传输协议（HTTP）、远程登录（TELNET）等。

UDP 是一个无连接的非可靠传输层协议。它在 IP 之上仅提供两个附加服务：多路复用和对数据的错误检查。IP 知道怎样把分组投递给一台主机，但不知道怎样把它们投递给主机上的具体应用。UDP 在传送数据之前不需要先建立连接，远程主机的传输层收到 UDP 报文后，不需要给

出任何确认。由于 UDP 比较简单，因此执行速度比较快、实时性好。使用 UDP 的应用主要包括小文件传送协议（TFTP）、DNS、SNMP 和实时传输协议（RTP）。

注意：

- 1) IP 数据报和 UDP 数据报的区别：IP 数据报在网络层要经过路由的存储转发；而 UDP 数据报在传输层的端到端的逻辑信道中传输，封装成 IP 数据报在网络层传输时，UDP 数据报的信息对路由是不可见的。
- 2) TCP 和网络层虚电路的区别：TCP 报文段在传输层抽象的逻辑信道中传输，对路由器不可见；虚电路所经过的交换结点都必须保存虚电路状态信息。在网络层若采用虚电路方式，则无法提供无连接服务；而传输层采用 TCP 不影响网络层提供无连接服务。

5.1.4 本节习题精选

单项选择题

01. 下列不属于通信子网的是（ ）。
 - A. 物理层
 - B. 数据链路层
 - C. 网络层
 - D. 传输层
02. OSI 参考模型中，提供端到端的透明数据传输服务、差错控制和流量控制的层是（ ）。
 - A. 物理层
 - B. 网络层
 - C. 传输层
 - D. 会话层
03. 传输层为（ ）之间提供逻辑通信。
 - A. 主机
 - B. 进程
 - C. 路由器
 - D. 操作系统
04. 关于传输层的面向连接服务的特性是（ ）。
 - A. 不保证可靠和顺序交付
 - B. 不保证可靠但保证顺序交付
 - C. 保证可靠但不保证顺序交付
 - D. 保证可靠和顺序交付
05. 在 TCP/IP 参考模型中，传输层的主要作用是在互联网的源主机和目的主机对等实体之间建立用于会话的（ ）。
 - A. 操作连接
 - B. 点到点连接
 - C. 控制连接
 - D. 端到端连接
06. 可靠传输协议中的“可靠”指的是（ ）。
 - A. 使用面向连接的会话
 - B. 使用尽力而为的传输
 - C. 使用滑动窗口来维持可靠性
 - D. 使用确认机制来确保传输的数据不丢失
07. 以下（ ）能够唯一确定一个在互联网上通信的进程。
 - A. 主机名
 - B. IP 地址及 MAC 地址
 - C. MAC 地址及端口号
 - D. IP 地址及端口号
08. 在（ ）范围内的端口号被称为“熟知端口号”并限制使用。这就意味着这些端口号是为常用的应用层协议如 FTP、HTTP 等保留的。
 - A. 0~127
 - B. 0~255
 - C. 0~511
 - D. 0~1023
09. 以下哪个 TCP 熟知端口号是错误的？（ ）
 - A. TELNET:23
 - B. SMTP:25
 - C. HTTP:80
 - D. FTP:24
10. 关于 TCP 和 UDP 端口的下列说法中，正确的是（ ）。
 - A. TCP 和 UDP 分别拥有自己的端口号，它们互不干扰，可以共存于同一台主机
 - B. TCP 和 UDP 分别拥有自己的端口号，但它们不能共存于同一台主机
 - C. TCP 和 UDP 的端口没有本质区别，但它们不能共存于同一台主机
 - D. 当一个 TCP 连接建立时，它们互不干扰，不能共存于同一台主机
11. 以下说法错误的是（ ）。

关注公众号【乘龙考研】
一手更新 稳定有保障

- A. 传输层是 OSI 参考模型的第四层
 - B. 传输层提供的是主机间的点到点数据传输
 - C. TCP 是面向连接的，UDP 是无连接的
 - D. TCP 进行流量控制和拥塞控制，而 UDP 既不进行流量控制，又不进行拥塞控制
12. 假设某应用程序每秒产生一个 60B 的数据块，每个数据块被封装在一个 TCP 报文中，然后再封装在一个 IP 数据报中。那么最后每个数据报所包含的应用数据所占的百分比是（ ）。(注意：TCP 报文和 IP 数据报文的首部没有附加字段。)
- A. 20%
 - B. 40%
 - C. 60%
 - D. 80%
13. 若用户程序使用 UDP 进行数据传输，则（ ）层协议必须承担可靠性方面的全部工作。
- A. 数据链路层
 - B. 网际层
 - C. 传输层
 - D. 应用层

5.1.5 答案与解析

单项选择题

01. D

通信子网包括物理层、数据链路层和网络层，主要负责数据通信。资源子网是 OSI 参考模型的上三层，传输层的主要任务是向高层用户屏蔽下面通信子网的细节（如网络拓扑、路由协议等）。

02. C

端到端即是进程到进程，物理层只提供在两个结点之间透明地传输比特流，网络层提供主机到主机的通信服务，主要功能是路由选择。此题的条件若换成“TCP/IP 参考模型”，答案依然是 C。

03. B

传输层提供是端到端服务，为进程之间提供逻辑通信。

04. D

面向连接服务是指通信双方在进行通信之前，要先建立一个完整的连接，在通信过程中，整个连接一直可以被实时地监控和管理。通信完毕后释放连接。面向连接的服务可以保证数据的可靠和顺序交付。

05. D

TCP/IP 模型中，网络层及其以下各层所构成的通信子网负责主机到主机或点到点的通信，而传输层的主要作用是在源主机进程和目的主机进程之间提供端到端的数据传输。一般来说，端到端通信是由一段段的点到点信道构成的，端到端协议建立在点到点协议之上（正如 TCP 建立在 IP 之上），提供应用进程之间的通信手段。所以答案为选项 D。

06. D

如果一个协议使用确认机制对传输的数据进行确认，那么可以认为它是一个可靠的协议；如果一个协议采用“尽力而为”的传输方式，那是不可靠的。例如，TCP 对传输的报文段提供确认，因此是可靠的传输协议；而 UDP 不提供确认，因此是不可靠的传输协议。

07. D

要在互联网上唯一地确定一个进程，就要使用 IP 地址和端口号的组合，通常称为套接字（Socket），IP 地址确定某主机，端口号确定该主机上的某进程。

08. D

熟知端口号的数值为 0~1023，登记端口号的数值是 1024~49151，客户端使用的端口号的数值是 49152~65535。

09. D

FTP 控制连接的端口是 21，数据连接的端口是 20。

10. A

端口号只具有本地意义，即端口号只标识本计算机应用层中的各个进程，且同一台计算机中 TCP 和 UDP 分别拥有自己的端口号，它们互不干扰。

11. B

传输层是 OSI 参考模型中的第 4 层，TCP 是面向连接的，它提供流量控制和拥塞控制，保证服务可靠；UDP 是无连接的，不提供流量控制和拥塞控制，只能做出尽最大努力的交付。传输层提供的是进程到进程间的传输服务，也称端到端服务。

12. C

此题中，一个 TCP 报文的首部长度是 20B，一个 IP 数据报的首部长度也是 20B，再加上 60B 的数据，一个 IP 数据报的总长度为 100B，可知数据占 60%。

13. D

传输层协议需要具有的主要功能包括：创建进程到进程的通信；提供流量控制机制。UDP 在一个低的水平上完成以上功能，使用端口号完成进程到进程的通信，但在传送数据时没有流量控制机制，也没有确认，而且只提供有限的差错控制。因此 UDP 是一个无连接、不可靠的传输层协议。如果用户应用程序使用 UDP 进行数据传输，那么必须在传输层的上层即应用层提供可靠性方面的全部工作。

5.2 UDP 协议

5.2.1 UDP 数据报



1. UDP 概述

UDP 仅在 IP 的数据报服务之上增加了两个最基本的服务：复用和分用以及差错检测。如果应用开发者选择 UDP 而非 TCP，那么应用程序几乎直接与 IP 打交道。为什么应用开发者宁愿在 UDP 之上构建应用，也不选择 TCP？既然 TCP 提供可靠的服务，而 UDP 不提供，那么 TCP 总是首选吗？答案是否定的，因为有很多应用更适合用 UDP，主要因为 UDP 具有如下优点：

- 1) UDP 无须建立连接。因此 UDP 不会引入建立连接的时延。试想如果 DNS 运行在 TCP 而非 UDP 上，那么 DNS 的速度会慢很多。HTTP 使用 TCP 而非 UDP，是因为对于基于文本数据的 Web 网页来说，可靠性是至关重要的。
- 2) 无连接状态。TCP 需要在端系统中维护连接状态。此连接状态包括接收和发送缓存、拥塞控制参数和序号与确认号的参数。而 UDP 不维护连接状态，也不跟踪这些参数。因此，某些专用应用服务器使用 UDP 时，一般都能支持更多的活动客户机。
- 3) 分组首部开销小。TCP 有 20B 的首部开销，而 UDP 仅有 8B 的开销。
- 4) 应用层能更好地控制要发送的数据和发送时间。UDP 没有拥塞控制，因此网络中的拥塞不会影响主机的发送效率。某些实时应用要求以稳定的速度发送，能容忍一些数据的丢失，但不允许有较大的时延，而 UDP 正好满足这些应用的需求。
- 5) UDP 支持一对一、一对多、多对一和多对多的交互通信。

UDP 常用于一次性传输较少数据的网络应用，如 DNS、SNMP 等，因为对于这些应用，若

采用 TCP，则将为连接创建、维护和拆除带来不小的开销。UDP 也常用于多媒体应用（如 IP 电话、实时视频会议、流媒体等），显然，可靠数据传输对这些应用来说并不是最重要的，但 TCP 的拥塞控制会导致数据出现较大的延迟，这是它们不可容忍的。

UDP 不保证可靠交付，但这并不意味着应用对数据的要求是不可靠的，所有维护可靠性的工
作可由用户在应用层来完成。应用开发者可根据应用的需求来灵活设计自己的可靠性机制。

UDP 是面向报文的。发送方 UDP 对应用层交下来的报文，在添加首部后就向下交付给 IP 层，一次发送一个报文，既不合并，也不拆分，而是保留这些报文的边界；接收方 UDP 对 IP 层交上来 UDP 数据报，在去除首部后就原封不动地交付给上层应用进程，一次交付一个完整的报文。因此报文不可分割，是 UDP 数据报处理的最小单位。因此，应用程序必须选择合适大小的报文，若报文太长，UDP 把它交给 IP 层后，可能会导致分片；若报文太短，UDP 把它交给 IP 层后，会使 IP 数据报的首部的相对长度太大，两者都会降低 IP 层的效率。

2. UDP 的头部格式

UDP 数据报包含两部分：UDP 首部和用户数据。UDP 首部有 8B，由 4 个字段组成，每个字
段的长度都是 2B，如图 5.2 所示。各字段意义如下：

- 1) 源端口号。源端口号。在需要对方回信时选用，不需要时可用全 0。
- 2) 目的端口号。目的端口号。这在终点交付报文时必须使用到。
- 3) 长度。UDP 数据报的长度（包括首部和数据），其最小值是 8（仅有首部）。
- 4) 校验和。检测 UDP 数据报在传输中是否有错。有错就丢弃。该字段是可选的，当源主机
不想计算校验和时，则直接令该字段为全 0。

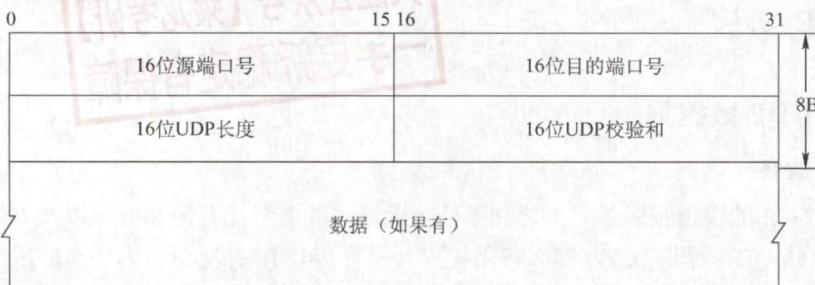


图 5.2 UDP 数据报格式

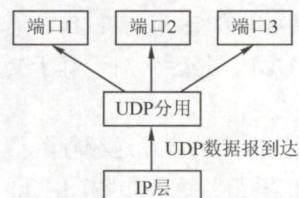


图 5.3 UDP 基于端口的分用

当传输层从 IP 层收到 UDP 数据报时，就根据首部中的目的端口，把 UDP 数据报通过相应的端口上交给应用进程，如图 5.3 所示。

如果接收方 UDP 发现收到的报文中的目的端口号不正确（即不存在对应于端口号的应用进程），那么就丢弃该报文，并由 ICMP 发送“端口不可达”差错报文给发送方。

5.2.2 UDP 校验

在计算校验和时，要在 UDP 数据报之前增加 12B 的伪首部，伪首部并不是 UDP 的真正首部。只是在计算校验和时，临时添加在 UDP 数据报的前面，得到一个临时的 UDP 数据报。校验和就是按照这个临时的 UDP 数据报来计算的。伪首部既不向下传送又不向上递交，而只是为了计算校验和。图 5.4 给出了 UDP 数据报的伪首部各字段的内容。

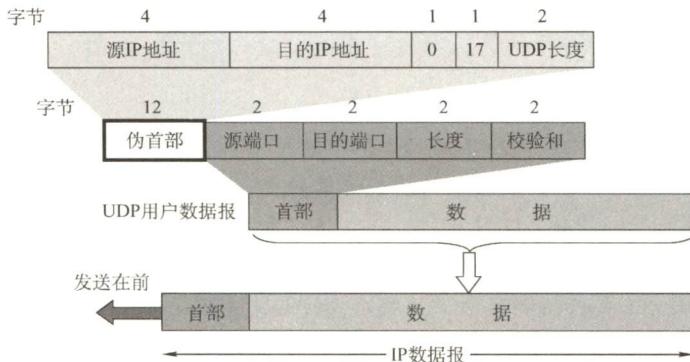


图 5.4 UDP 数据报的首部和伪首部

UDP 校验和的计算方法和 IP 数据报首部校验和的计算方法相似。但不同的是，IP 数据报的校验和只检验 IP 数据报的首部，但 UDP 的校验和则检查首部和数据部分。

发送方首先把全零放入校验和字段并添加伪首部，然后把 UDP 数据报视为许多 16 位的字串接起来。若 UDP 数据报的数据部分不是偶数个字节，则要在数据部分末尾填入一个全零字节（但此字节不发送）。然后按二进制反码计算出这些 16 位字的和，将此和的二进制反码写入校验和字段，并发送。接收方把收到的 UDP 数据报加上伪首部（如果不为偶数个字节，那么还需要补上全零字节）后，按二进制反码求这些 16 位字的和。当无差错时其结果应为全 1，否则就表明有差错出现，接收方就应该丢弃这个 UDP 数据报。

图 5.5 给出了一个计算 UDP 校验和的例子。本例中，UDP 数据报的长度是 15B（不含伪首部），因此需要添加一个全 0 字节。



图 5.5 计算 UDP 校验和的例子

注意：

- 1) 校验时，若 UDP 数据报部分的长度不是偶数个字节，则需填入一个全 0 字节，如图 5.5 所示。但是此字节和伪首部一样，是不发送的。
- 2) 如果 UDP 校验和校验出 UDP 数据报是错误的，那么可以丢弃，也可以交付给上层，但是需要附上错误报告，即告诉上层这是错误的数据报。
- 3) 通过伪首部，不仅可以检查源端口号、目的端口号和 UDP 用户数据报的数据部分，还可以检查 IP 数据报的源 IP 地址和目的地址。

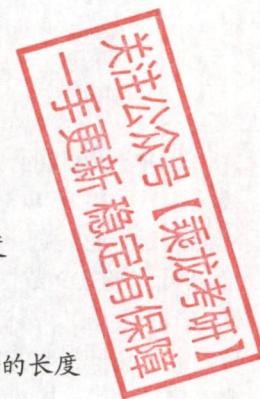


这种简单的差错校验方法的校错能力并不强，但它的好处是简单、处理速度快。

5.2.3 本节习题精选

一、单项选择题

01. 使用 UDP 的网络应用，其数据传输的可靠性由（ ）负责。
 - A. 传输层
 - B. 应用层
 - C. 数据链路层
 - D. 网络层
02. 以下关于 UDP 协议的主要特点的描述中，错误的是（ ）。
 - A. UDP 报头主要包括端口号、长度、校验和等字段
 - B. UDP 长度字段是 UDP 数据报的长度，包括伪首部的长度
 - C. UDP 校验和对伪首部、UDP 报文头及应用层数据进行校验
 - D. 伪首部包括 IP 分组报头的一部分
03. UDP 数据报首部不包含（ ）。
 - A. UDP 源端口号
 - B. UDP 校验和
 - C. UDP 目的端口号
 - D. UDP 数据报首部长度
04. UDP 数据报中的长度字段（ ）。
 - A. 不记录数据的长度
 - B. 只记录首部的长度
 - C. 只记录数据部分的长度
 - D. 包括首部和数据部分的长度
05. UDP 数据报比 IP 数据报多提供了（ ）服务。
 - A. 流量控制
 - B. 拥塞控制
 - C. 端口功能
 - D. 路由转发
06. 下列关于 UDP 的描述，正确的是（ ）。
 - A. 给出数据的按序投递
 - B. 不允许多路复用
 - C. 拥有流量控制机制
 - D. 是无连接的
07. 接收端收到有差错的 UDP 用户数据时的处理方式是（ ）。
 - A. 丢弃
 - B. 请求重传
 - C. 差错校正
 - D. 忽略差错
08. 以下关于 UDP 校验和的说法中，错误的是（ ）。
 - A. UDP 的校验和功能不是必需的，可以不使用
 - B. 如果 UDP 校验和计算结果为 0，那么在校验和字段填充 0
 - C. UDP 校验和字段的计算包括一个伪首部、UDP 首部和携带的用户数据
 - D. UDP 校验和的计算方法是二进制反码运算求和再取反
09. 下列关于 UDP 校验的描述中，（ ）是错误的。
 - A. UDP 校验和段的使用是可选的，若源主机不想计算校验和，则该校验和段应为全 0
 - B. 在计算校验和的过程中，需要生成一个伪首部，源主机需要把该伪首部发送给目的主机
 - C. 如果数据报在传输过程中被破坏，那么就把它丢弃
 - D. UDP 数据报的伪首部包含了 IP 地址信息
10. 下列网络应用中，（ ）不适合使用 UDP 协议。
 - A. 客户机/服务器领域
 - B. 远程调用
 - C. 实时多媒体应用
 - D. 远程登录
11. 【2014 统考真题】下列关于 UDP 协议的叙述中，正确的是（ ）。
 - I. 提供无连接服务
 - II. 提供复用/分用服务
 - III. 通过差错校验，保障可靠数据传输



- A. 仅 I B. 仅 I、II C. 仅 II、III D. I、II、III
12. 【2018 统考真题】 UDP 协议实现分用时所依据的头部字段是()。
- A. 源端口号 B. 目的端口号 C. 长度 D. 校验和

二、综合应用题

01. 为什么要使用 UDP? 让用户进程直接发送原始的 IP 分组不就足够了吗?
02. 使用 TCP 对实时语音数据的传输是否有问题? 使用 UDP 传送数据文件时有什么问题?
03. 一个应用程序用 UDP, 到了 IP 层将数据报再划分为 4 个数据报片发送出去。结果前两个数据报片丢失, 后两个到达目的站。过了一段时间应用程序重传 UDP, 而 IP 层仍然划分为 4 个数据报片来传送。结果这次前两个到达目的站而后两个丢失。试问: 在目的站能否将这两次传输的 4 个数据报片组装成为完整的数据报? 假定目的站第一次收到的后两个数据片仍然保存在目的站的缓存中。
04. 一个 UDP 首部的信息(十六进制表示)为 0xF7 21 00 45 00 2C E8 27。UDP 数据报的格式如下图所示。试问:



- 1) 源端口、目的端口、数据报总长度、数据部分长度分别是什么?
- 2) 该 UDP 数据报是从客户发送给服务器还是从服务器发送给客户? 使用该 UDP 服务的程序使用的是哪个应用层协议?
05. 一个 UDP 用户数据报的数据字段为 8192B, 要使用以太网来传送。假定 IP 数据报无选项。试问应当划分为几个 IP 数据报片? 说明每个 IP 数据报片的数据字段长度和片段偏移字段的值。

5.2.4 答案与解析

一、单项选择题

01. B

UDP 本身是无法保证传输的可靠性的, 并且 UDP 是基于网络层的 IP 的, IP 的特点是尽最大努力交付, 因此无法在网络层及数据链路层提供可靠传输。因此, 只能通过应用层协议来实现可靠传输。

02. B

伪首部只是在计算校验和时临时添加的, 不计入 UDP 的长度。对于选项 D, 伪首部包括源 IP 和目的 IP, 这是 IP 分组报头的一部分。

03. D

UDP 数据报的格式包括 UDP 源端口号、UDP 目的端口号、UDP 报文长度和校验和, 但不包括 UDP 数据报首部长度。因为 UDP 数据报首部长度是固定的 8B, 所以没有必要再设置首部长度字段。

04. D

长度字段记录 UDP 数据报的长度(包括首部和数据部分), 以字节为单位。

关注公众号【乘龙考研】
一手更新 稳定有保障

05. C

虽然 UDP 协议和 IP 协议都是数据报协议，但是它们之间还是存在差别。其中，最大的差别是 IP 数据报只能找到目的主机而无法找到目的进程，UDP 提供端口功能及复用和分用功能，可以将数据报投递给对应的进程。

06. D

UDP 是不可靠的，所以没有数据的按序投递，排除选项 A；UDP 只在 IP 的数据报服务上增加了一点功能，即复用和分用功能及差错检测功能，排除选项 B；显然 UDP 没有流量控制，排除选项 C；UDP 是传输层的无连接协议，答案为 D。

07. A

接收端通过校验发现数据有差错，就直接丢弃该数据报，仅此而已。

08. B

UDP 的校验和不是必需的，如果不使用校验和，那么将校验和字段设置为 0，而如果校验和的计算结果恰好为 0，那么将校验和字段置为全 1。

09. B

UDP 数据报的伪首部包含了 IP 地址信息，目的是通过数据校验保证 UDP 数据报正确地到达目的主机。该伪首部由源和目的主机仅在校验和计算期间建立，并不发送。

10. D

UDP 的特点是开销小，时间性能好且易于实现。在客户/服务器模式中，它们之间的请求都很短，使用 UDP 不仅编码简单，而且只需要很少的消息；远程调用使用 UDP 的理由和客户/服务器模式一样；对于实时多媒体应用，需要保证数据及时传送，而比例不大的错误是可以容忍的，所以使用 UDP 也是合适的，而且使用 UDP 协议可以实现多播，给多个客户端服务；而远程登录需要依靠一个客户端到服务器的可靠连接，使用 UDP 是不合适的。

11. B

UDP 提供的是无连接服务，I 正确；同时 UDP 也提供复用/分用服务，II 正确；UDP 虽然有差错校验机制，但 UDP 的差错校验只是检查数据在传输的过程中有没有出错，出错的数据直接丢弃，并没有重传等机制，不能保证可靠传输。使用 UDP 协议时，可靠传输必须由应用层实现，III 错误。答案选 B。

12. B

传输层分用的定义是，接收方的传输层剥去报文首部后，能把这些数据正确交付到目的进程。选项 C 和 D 显然不符。端口号是传输层服务访问点（TSAP），用来标识主机中的应用进程。对于选项 A 和 B，源端口号在需要对方回信时选用，不需要时可用全 0。目的端口号在终点交付报文时使用，符合题意，因此答案为选项 B。

二、综合应用题

01. 【解答】

仅仅使用 IP 分组还不够。IP 分组包含 IP 地址，该地址指定一个目的机器。一旦这样的分组到达目的机器，网络控制程序如何知道把它交给哪个进程呢？UDP 分组包含一个目的端口，这一信息是必需的，因为有了它，分组才能被投递给正确的进程。此外，UDP 可以对数据报做包括数据段在内的差错检测，而 IP 只对其首部做差错检测。

02. 【解答】

如果语音数据不实时播放，那么可以使用 TCP，因为 TCP 有重传机制，传输可靠。接收端

用 TCP 将语音数据接收完毕后，可以在以后的任何时间进行播放。若假定是实时传输，不宜重传，则必须使用 UDP。UDP 不保证可靠递交，没有重传机制，因此在传输数据时可能会丢失数据，但 UDP 比 TCP 的开销要小很多，实时性好。

03. 【解答】

不行。重传时，IP 数据报的标识字段会有另一个标识符。仅当标识符相同的 IP 数据报片才能组装成一个 IP 数据报。前两个 IP 数据报片的标识符与后两个 IP 数据报片的标识符不同，因此不能组装成一个 IP 数据报。

04. 【解答】

- 1) 第 1、2 个字节为源端口，即 F7 21，转换成十进制数为 63265。第 3、4 个字节为目的端口，即 00 45，转换成十进制数为 69。第 5、6 个字节为 UDP 长度（包含首部和数据部分），即 00 2C，转换成十进制数为 44，数据报总长度为 44B，数据部分长度为 $44 - 8 = 36$ B。
- 2) 由 1) 可知，该 UDP 数据报的源端口号为 63265，目的端口号为 69，前一个为客户端使用的端口号，后一个为熟知的 TFTP 协议的端口，可知该数据报是客户发给服务器的。

05. 【解答】

以太网帧的数据段的最大长度是 1500B，UDP 用户数据报的首部是 8B。假定 IP 数据报无选项，首部长度都是 20B。IP 数据报的片段偏移指出一个片段在原 IP 分组中的相对位置，偏移的单位是 8B。UDP 用户数据报的数据字段为 8192B，加上 8B 的首部，总长度是 8200B。应当划分为 6 个 IP 报片。各 IP 报片总长度、数据长度和片偏移如下表所示。

	1	2	3	4	5	6
IP 报片总长度	1500B	1500B	1500B	1500B	1500B	820B
数据长度	1480B	1480B	1480B	1480B	1480B	800B
片偏移	0	185	370	555	740	925

5.3 TCP 协议

5.3.1 TCP 协议的特点

关注公众号【乘龙考研】

一手更新 稳定有保障

TCP 是在不可靠的 IP 层之上实现的可靠的数据传输协议，它主要解决传输的可靠、有序、无丢失和不重复问题。TCP 是 TCP/IP 体系中非常复杂的一个协议，主要特点如下：

- 1) TCP 是面向连接的传输层协议，TCP 连接是一条逻辑连接。
 - 2) 每一条 TCP 连接只能有两个端点，每一条 TCP 连接只能是点到点的（一对一）。
 - 3) TCP 提供可靠交付的服务，保证传送的数据无差错、不丢失、不重复且有序。
 - 4) TCP 提供全双工通信，允许通信双方的应用进程在任何时候都能发送数据，为此 TCP 连接的两端都设有发送缓存和接收缓存，用来临时存放双向通信的数据。
发送缓存用来暂时存放以下数据：①发送应用程序传送给发送方 TCP 准备发送的数据；②TCP 已发送但尚未收到确认的数据。接收缓存用来暂时存放以下数据：①按序到达但尚未被接收应用程序读取的数据；②不按序到达的数据。
 - 5) TCP 是面向字节流的，虽然应用程序和 TCP 的交互是一次一个数据块（大小不等），但 TCP 把应用程序交下来的数据仅视为一连串的无结构的字节流。
- TCP 和 UDP 在发送报文时所采用的方式完全不同。UDP 报文的长度由发送应用进程决定，

而 TCP 报文的长度则根据接收方给出的窗口值和当前网络拥塞程度来决定。如果应用进程传送到 TCP 缓存的数据块太长，TCP 就把它划分得短一些再传送；如果太短，TCP 也可以等到积累足够多的字节后再构成报文段发送出去。关于 TCP 报文的长度问题，后面会详细讨论。

5.3.2 TCP 报文段

TCP 传送的数据单元称为报文段。TCP 报文段既可以用来运载数据，又可以用来建立连接、释放连接和应答。一个 TCP 报文段分为首部和数据两部分，整个 TCP 报文段作为 IP 数据报的数据部分封装在 IP 数据报中，如图 5.6 所示。其首部的前 20B 是固定的。TCP 首部最短为 20B，后面有 $4N$ 字节是根据需要而增加的选项，长度为 4B 的整数倍。

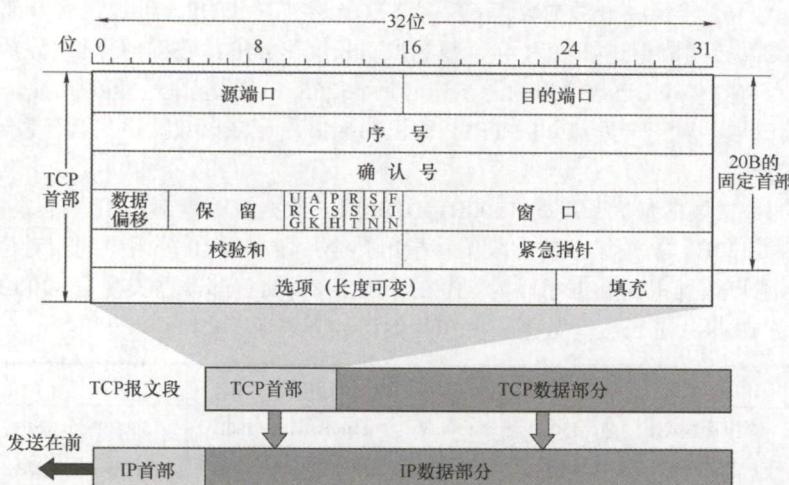


图 5.6 TCP 报文段

TCP 的全部功能体现在其首部的各个字段中，各字段意义如下：

- 1) 源端口和目的端口。各占 2B。端口是传输层与应用层的服务接口，传输层的复用和分用功能都要通过端口实现。
- 2) 序号。占 4B，范围为 $0 \sim 2^{32} - 1$ ，共 2^{32} 个序号。TCP 是面向字节流的（即 TCP 传送时是逐个字节传送的），所以 TCP 连接传送的字节流中的每个字节都按顺序编号。序号字段的值指的是本报文段所发送的数据的第一个字节的序号。
例如，一报文段的序号字段值是 301，而携带的数据共有 100B，表明本报文段的数据的最后一个字节的序号是 400，因此下一个报文段的数据序号应从 401 开始。
- 3) 确认号。占 4B，是期望收到对方下一个报文段的第一个数据字节的序号。若确认号为 N ，则表明到序号 $N-1$ 为止的所有数据都已正确收到。
例如，B 正确收到了 A 发送过来的一个报文段，其序号字段是 501，而数据长度是 200B（序号 501~700），这表明 B 正确收到了 A 发送的到序号 700 为止的数据。因此 B 期望收到 A 的下一个数据序号是 701，于是 B 在发送给 A 的确认报文段中把确认号置为 701。
- 4) 数据偏移（即首部长度）。占 4 位，这里不是 IP 数据报分片的那个数据偏移，而是表示首部长度（首部中还有长度不确定的选项字段），它指出 TCP 报文段的数据起始处距离 TCP 报文段的起始处有多远。“数据偏移”的单位是 32 位（以 4B 为计算单位）。由于 4 位二进制数能表示的最大值为 15，因此 TCP 首部的最大长度为 60B。

- 5) 保留。占 6 位, 保留为今后使用, 但目前应置为 0。
- 6) 紧急位 URG。当 $URG=1$ 时, 表明紧急指针字段有效。它告诉系统此报文段中有紧急数据, 应尽快传送(相当于高优先级的数据)。但 URG 需要和首部中紧急指针字段配合使用, 即数据从第一个字节到紧急指针所指字节就是紧急数据。
- 7) 确认位 ACK。仅当 $ACK=1$ 时确认号字段才有效。当 $ACK=0$ 时, 确认号无效。
TCP 规定, 在连接建立后所有传送的报文段都必须把 ACK 置 1。
- 8) 推送位 PSH (Push)。接收方 TCP 收到 $PSH=1$ 的报文段, 就尽快地交付给接收应用进程, 而不再等到整个缓存都填满了后再向上交付。
- 9) 复位位 RST (Reset)。当 $RST=1$ 时, 表明 TCP 连接中出现严重差错(如主机崩溃或其他原因), 必须释放连接, 然后再重新建立运输连接。
- 10) 同步位 SYN。当 $SYN=1$ 时表示这是一个连接请求或连接接受报文。
当 $SYN=1$, $ACK=0$ 时, 表明这是一个连接请求报文, 对方若同意建立连接, 则应在响应报文中使用 $SYN=1$, $ACK=1$ 。
- 11) 终止位 FIN (Finish)。用来释放一个连接。当 $FIN=1$ 时, 表明此报文段的发送方的数据已发送完毕, 并要求释放运输连接。
- 12) 窗口。占 2B, 范围为 $0 \sim 2^{16}-1$ 。它指出现在允许对方发送的数据量, 接收方的数据缓存空间是有限的, 因此用窗口值作为接收方让发送方设置其发送窗口的依据。
例如, 设确认号是 701, 窗口字段是 1000。这表明, 从 701 号算起, 发送此报文段的一方还有接收 1000 字节数据(字节序号为 701~1700)的接收缓存空间。
- 13) 校验和。占 2B。校验和字段检验的范围包括首部和数据两部分。在计算校验和时, 和 UDP 一样, 要在 TCP 报文段的前面加上 12B 的伪首部(只需将 UDP 伪首部的协议字段的 17 改成 6, UDP 长度字段改成 TCP 长度, 其他的和 UDP 一样)。
- 14) 紧急指针。占 2B。紧急指针仅在 $URG=1$ 时才有意义, 它指出在本报文段中紧急数据共有多少字节(紧急数据在报文段数据的最前面)。
- 15) 选项。长度可变。TCP 最初只规定了一种选项, 即最大报文段长度 (Maximum Segment Size, MSS)。MSS 是 TCP 报文段中的数据字段的最大长度(注意仅仅是数据字段)。
- 16) 填充。这是为了使整个首部长度是 4B 的整数倍。

5.3.3 TCP 连接管理

TCP 是面向连接的协议, 因此每个 TCP 连接都有三个阶段: **连接建立、数据传递和连接释放**。TCP 连接的管理就是使运输连接的建立和释放都能正常进行。

在 TCP 连接建立的过程中, 要解决以下三个问题:

- 1) 要使每一方能够确知对方的存在。
- 2) 要允许双方协商一些参数(如最大窗口值、是否使用窗口扩大选项、时间戳选项及服务质量等)。
- 3) 能够对运输实体资源(如缓存大小、连接表中的项目等)进行分配。

TCP 把连接作为最基本的抽象, 每条 TCP 连接有两个端点, TCP 连接的端点不是主机, 不是主机的 IP 地址, 不是应用进程, 也不是传输层的协议端口。TCP 连接的端口即为套接字(Socket)或插口, 每条 TCP 连接唯一地被通信的两个端点(即两个套接字)确定。

TCP 连接的建立采用客户/服务器模式。主动发起连接建立的应用进程称为客户 (Client), 而被动等待连接建立的应用进程称为服务器 (Server)。

关注公众号【乘龙考研】

一手更新 稳定有保障

1. TCP 连接的建立

连接的建立经历以下 3 个步骤，通常称为三次握手，如图 5.7 所示。

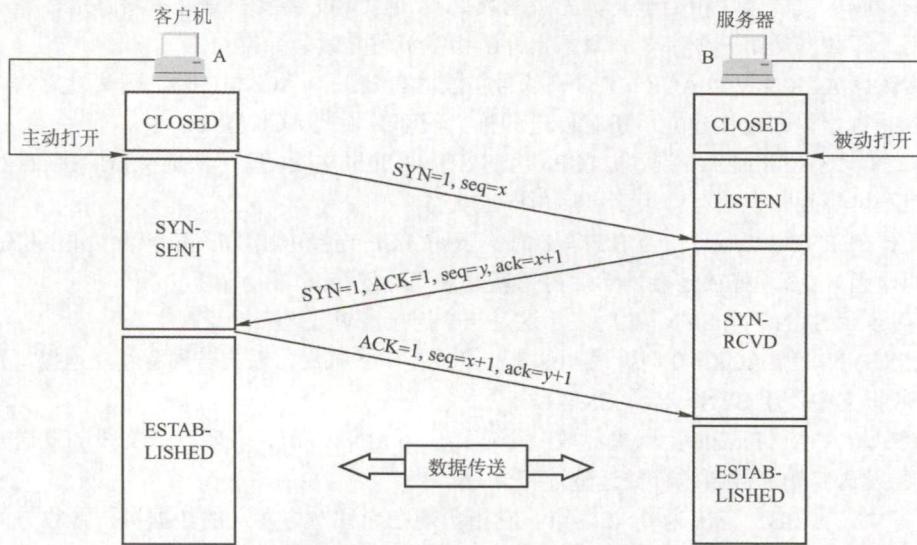


图 5.7 用“三次握手”建立 TCP 连接

连接建立前，服务器进程处于 LISTEN（收听）状态，等待客户的连接请求。

第一步：客户机的 TCP 首先向服务器的 TCP 发送连接请求报文段。这个特殊报文段的首部中的同步位 SYN 置 1，同时选择一个初始序号 $seq=x$ 。TCP 规定，SYN 报文段不能携带数据，但要消耗掉一个序号。这时，TCP 客户进程进入 SYN-SENT（同步已发送）状态。

第二步：服务器的 TCP 收到连接请求报文段后，如同意建立连接，则向客户机发回确认，并为该 TCP 连接分配缓存和变量。在确认报文段中，把 SYN 位和 ACK 位都置 1，确认号是 $ack=x+1$ ，同时也为自己选择一个初始序号 $seq=y$ 。注意，确认报文段不能携带数据，但也要消耗掉一个序号。这时，TCP 服务器进程进入 SYN-RCVD（同步收到）状态。

第三步：当客户机收到确认报文段后，还要向服务器给出确认，并为该 TCP 连接分配缓存和变量。确认报文段的 ACK 位置 1，确认号 $ack=y+1$ ，序号 $seq=x+1$ 。该报文段可以携带数据，若不携带数据则不消耗序号。这时，TCP 客户进程进入 ESTABLISHED（已建立连接）状态。

成功进行以上三步后，就建立了 TCP 连接，接下来就可以传送应用层数据。TCP 提供的是全双工通信，因此通信双方的应用进程在任何时候都能发送数据。

另外，值得注意的是，服务器端的资源是在完成第二次握手时分配的，而客户端的资源是在完成第三次握手时分配的，这就使得服务器易于受到 SYN 洪泛攻击。

2. TCP 连接的释放

天下没有不散的筵席，TCP 同样如此。参与 TCP 连接的两个进程中的任何一个都能终止该连接。TCP 连接释放的过程通常称为四次握手，如图 5.8 所示。

第一步：客户机打算关闭连接时，向其 TCP 发送连接释放报文段，并停止发送数据，主动关闭 TCP 连接，该报文段的终止位 FIN 置 1，序号 $seq=u$ ，它等于前面已传送过的数据的最后一个字节的序号加 1，FIN 报文段即使不携带数据，也消耗掉一个序号。这时，TCP 客户进程进入 FIN-WAIT-1（终止等待 1）状态。TCP 是全双工的，即可以想象为一条 TCP 连接上有两条数据通路，发送 FIN 的一端不能再发送数据，即关闭了其中一条数据通路，但对方还可以发送数据。

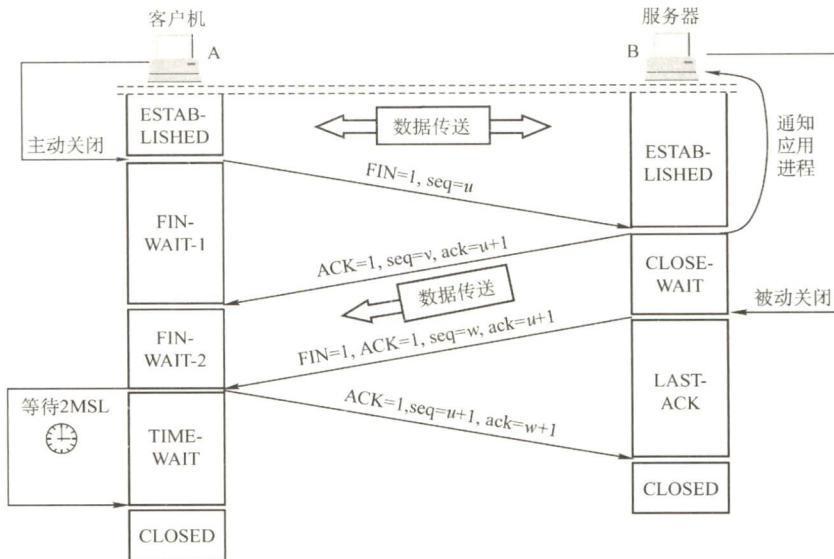


图 5.8 用“四次握手”释放 TCP 连接

第二步：服务器收到连接释放报文段后即发出确认，确认号 $ack = u + 1$ ，序号 $seq = v$ ，等于它前面已传送过的数据的最后一个字节的序号加 1。然后服务器进入 CLOSE-WAIT（关闭等待）状态。此时，从客户机到服务器这个方向的连接就释放了，TCP 连接处于半关闭状态。但服务器若发送数据，客户机仍要接收，即从服务器到客户机这个方向的连接并未关闭。

第三步：若服务器已经没有要向客户机发送的数据，就通知 TCP 释放连接，此时，其发出 $FIN = 1$ 的连接释放报文段。设该报文段的序号为 w （在半关闭状态服务器可能又发送了一些数据），还须重复上次已发送的确认号 $ack = u + 1$ 。这时服务器进入 LAST-ACK（最后确认）状态。

第四步：客户机收到连接释放报文段后，必须发出确认。把确认报文段中的确认位 ACK 置 1，确认号 $ack = w + 1$ ，序号 $seq = u + 1$ 。此时 TCP 连接还未释放，必须经过时间等待计时器设置的时间 2MSL（最长报文段寿命）后，客户机才进入 CLOSED（连接关闭）状态。

对上述 TCP 连接建立和释放的总结如下：

1) 连接建立。分为 3 步：

- ① $SYN = 1, seq = x$ 。
- ② $SYN = 1, ACK = 1, seq = y, ack = x + 1$ 。
- ③ $ACK = 1, seq = x + 1, ack = y + 1$ 。

2) 释放连接。分为 4 步：

- ① $FIN = 1, seq = u$ 。
- ② $ACK = 1, seq = v, ack = u + 1$ 。
- ③ $FIN = 1, ACK = 1, seq = w, ack = u + 1$ 。
- ④ $ACK = 1, seq = u + 1, ack = w + 1$ 。

选择题喜欢考查（关于连接和释放的题目，ACK、SYN、FIN 一定等于 1），请牢记。

关注公众号【乘龙考研】
一手更新 稳定有保障

5.3.4 TCP 可靠传输

TCP 的任务是在 IP 层不可靠的、尽力而为服务的基础上建立一种可靠数据传输服务。TCP 提供的可靠数据传输服务保证接收方进程从缓存区读出的字节流与发送方发出的字节流完全一

样。TCP 使用了校验、序号、确认和重传等机制来达到这一目的。其中，TCP 的校验机制与 UDP 校验一样，这里不再赘述。

1. 序号

TCP 首部的序号字段用来保证数据能有序提交给应用层，TCP 把数据视为一个无结构但有序的字节流，序号建立在传送的字节流之上，而不建立在报文段之上。

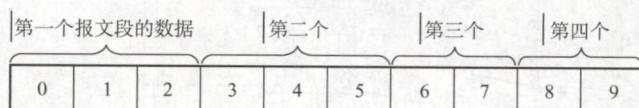


图 5.9 A 的发送缓存区中的数据划分成 TCP 段

缓存区中共有 10B，序号从 0 开始标号，第一个报文包含第 0~2 个字节，第二个报文段的序号是 1，第三个报文段的序号是 3。

TCP 连接传送的数据流中的每个字节都编上一个序号。序号字段的值是指本报文段所发送的数据的第一个字节的序号。如图 5.9 所示，假设 A 和 B 之间建立了一条 TCP 连接，A 的发送

2. 确认

TCP 首部的确认号是期望收到对方的下一个报文段的数据的第一个字节的序号。在图 5.9 中，如果接收方 B 已收到第一个报文段，此时 B 希望收到的下一个报文段的数据是从第 3 个字节开始的，那么 B 发送给 A 的报文中的确认号字段应为 3。发送方缓存区会继续存储那些已发送但未收到确认的报文段，以便在需要时重传。

TCP 默认使用累积确认，即 TCP 只确认数据流中至第一个丢失字节为止的字节。例如，在图 5.9 中，接收方 B 收到了 A 发送的包含字节 0~2 及字节 6~7 的报文段。由于某种原因，B 还未收到字节 3~5 的报文段，此时 B 仍在等待字节 3（及其后面的字节），因此 B 到 A 的下一个报文段将确认号字段置为 3。

3. 重传

有两种事件会导致 TCP 对报文段进行重传：超时和冗余 ACK。

(1) 超时

TCP 每发送一个报文段，就对这个报文段设置一次计时器。计时器设置的重传时间到期但还未收到确认时，就要重传这一报文段。

由于 TCP 的下层是一个互联网环境，IP 数据报所选择的路由变化很大，因而传输层的往返时延的方差也很大。为了计算超时计时器的重传时间，TCP 采用一种自适应算法，它记录一个报文段发出的时间，以及收到相应确认的时间，这两个时间之差称为报文段的往返时间 (Round-Trip Time, RTT)。TCP 保留了 RTT 的一个加权平均往返时间 RTT_s，它会随新测量 RTT 样本值的变化而变化。显然，超时计时器设置的超时重传时间 (Retransmission Time-Out, RTO) 应略大于 RTT_s，但也不能大太多，否则当报文段丢失时，TCP 不能很快重传，导致数据传输时延大。

(2) 冗余 ACK (冗余确认)

超时触发重传存在的一个问题就是超时周期往往太长。所幸的是，发送方通常可在超时事件发生之前通过注意所谓的冗余 ACK 来较好地检测丢包情况。冗余 ACK 就是再次确认某个报文段的 ACK，而发送方先前已经收到过该报文段的确认。例如，发送方 A 发送了序号为 1、2、3、4、5 的 TCP 报文段，其中 2 号报文段在链路中丢失，它无法到达接收方 B。因此 3、4、5 号报文段对于 B 来说就成了失序报文段。TCP 规定每当比期望序号大的失序报文段到达时，就发送一个冗余 ACK，指明下一个期待字节的序号。在本例中，3、4、5 号报文到达 B，但它们不是 B 所期望收到的下一

关注公众号【乘龙考研】
一手更新 稳定有保障

个报文，于是 B 就发送 3 个对 1 号报文段的冗余 ACK，表示自己期望接收 2 号报文段。TCP 规定当发送方收到对同一个报文段的 3 个冗余 ACK 时，就可以认为跟在这个被确认报文段之后的报文段已经丢失。就前面的例子而言，当 A 收到对于 1 号报文段的 3 个冗余 ACK 时，它可以认为 2 号报文段已经丢失，这时发送方 A 可以立即对 2 号报文执行重传，这种技术通常称为快速重传。当然，冗余 ACK 还被用在拥塞控制中，这将在后面的内容中讨论。

5.3.5 TCP 流量控制

TCP 提供流量控制服务来消除发送方（发送速率太快）使接收方缓存区溢出的可能性，因此可以说流量控制是一个速度匹配服务（匹配发送方的发送速率与接收方的读取速率）。

TCP 提供一种基于滑动窗口协议的流量控制机制，滑动窗口的基本原理已在第 3 章的数据链路层介绍过，这里要介绍的是 TCP 如何使用窗口机制来实现流量控制。

在通信过程中，接收方根据自己接收缓存的大小，动态地调整发送方的发送窗口大小，这称为接收窗口 rwnd，即调整 TCP 报文段首部中的“窗口”字段值，来限制发送方向网络注入报文的速率。同时，发送方根据其对当前网络拥塞程度的估计而确定的窗口值，这称为拥塞窗口 cwnd（后面会讲到），其大小与网络的带宽和时延密切相关。

例如，在通信中，有效数据只从 A 发往 B，而 B 仅向 A 发送确认报文，这时 B 可以通过设置确认报文段首部的窗口字段来将 rwnd 通知给 A。rwnd 即接收方允许连续接收的最大能力，单位是字节。发送方 A 总是根据最新收到的 rwnd 值来限制自己发送窗口的大小，从而将未确认的数据量控制在 rwnd 大小之内，保证 A 不会使 B 的接收缓存溢出。当然，A 的发送窗口的实际大小取 rwnd 和 cwnd 中的最小值。图 5.10 中的例子说明了如何利用滑动窗口机制进行流量控制。设 A 向 B 发送数据，在连接建立时，B 告诉 A：“我的接收窗口 $rwnd = 400$ ”。接收方主机 B 进行了三次流量控制，这三个报文段都设置了 $ACK = 1$ ，只有在 $ACK = 1$ 时确认号字段才有意义。第一次把窗口减小到 $rwnd = 300$ ，第二次又减到 $rwnd = 100$ ，最后减到 $rwnd = 0$ ，即不允许发送方再发送数据。这使得发送方暂停发送的状态将持续到 B 重新发出一个新的窗口值为止。

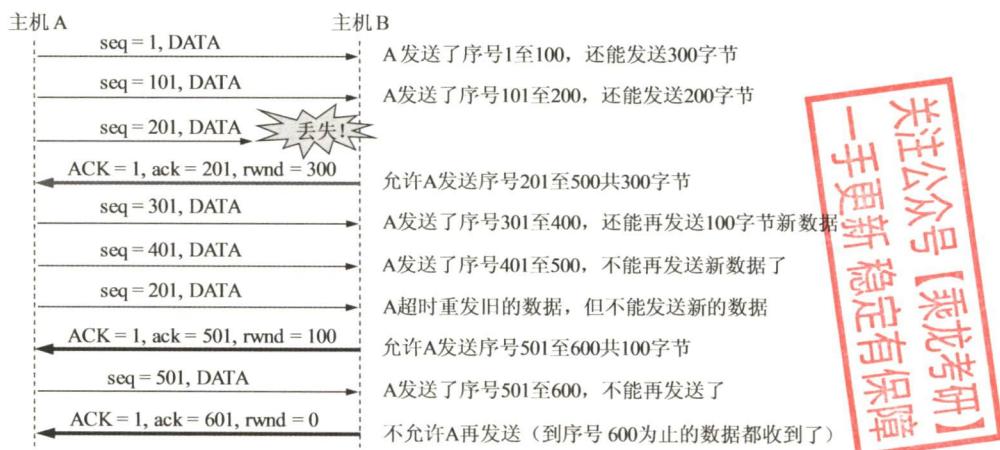


图 5.10 利用可变窗口进行流量控制举例

传输层和数据链路层的流量控制的区别是：传输层定义端到端用户之间的流量控制，数据链路层定义两个中间的相邻结点的流量控制。另外，数据链路层的滑动窗口协议的窗口大小不能动态变化，传输层的则可以动态变化。

5.3.6 TCP 拥塞控制

拥塞控制是指防止过多的数据注入网络，保证网络中的路由器或链路不致过载。出现拥塞时，端点并不了解拥塞发生的细节，对通信连接的端点来说，拥塞往往表现为通信时延的增加。

拥塞控制与流量控制的区别：拥塞控制是让网络能够承受现有的网络负荷，是一个全局性的过程，涉及所有的主机、所有的路由器，以及与降低网络传输性能有关的所有因素。相反，流量控制往往是指点对点的通信量的控制，是个端到端的问题（接收端控制发送端），它所要做的是抑制发送端发送数据的速率，以便使接收端来得及接收。当然，拥塞控制和流量控制也有相似的地方，即它们都通过控制发送方发送数据的速率来达到控制效果。

例如，某个链路的传输速率为 10Gb/s，某大型机向一台 PC 以 1Gb/s 的速率传送文件，显然网络的带宽是足够大的，因而不存在拥塞问题，但如此高的发送速率将导致 PC 可能来不及接收，因此必须进行流量控制。但若有 100 万台 PC 在此链路上以 1Mb/s 的速率传送文件，则现在的问题就变为网络的负载是否超过了现有网络所能承受的范围。

因特网建议标准定义了进行拥塞控制的 4 种算法：慢开始、拥塞避免、快重传和快恢复。

发送方在确定发送报文段的速率时，既要根据接收方的接收能力，又要从全局考虑不要使网络发生拥塞。因此，TCP 协议要求发送方维护以下两个窗口：

- 1) 接收窗口 $rwnd$ ，接收方根据目前接收缓存大小所许诺的最新窗口值，反映接收方的容量。
由接收方根据其放在 TCP 报文的首部的窗口字段通知发送方。
- 2) 拥塞窗口 $cwnd$ ，发送方根据自己估算的网络拥塞程度而设置的窗口值，反映网络的当前容量。只要网络未出现拥塞，拥塞窗口就再增大一些，以便把更多的分组发送出去。但只要网络出现拥塞，拥塞窗口就减小一些，以减少注入网络的分组数。

发送窗口的上限值应取接收窗口 $rwnd$ 和拥塞窗口 $cwnd$ 中较小的一个，即

$$\text{发送窗口的上限值} = \min[rwnd, cwnd]$$

接收窗口的大小可根据 TCP 报文首部的窗口字段通知发送方，而发送方如何维护拥塞窗口呢？这就是下面讲解的慢开始和拥塞避免算法。

注意：这里假设接收方总是有足够的缓存空间，因而发送窗口大小由网络的拥塞程度决定，也就是说，可以将发送窗口等同为拥塞窗口。

1. 慢开始和拥塞避免

(1) 慢开始算法

在 TCP 刚刚连接好并开始发送 TCP 报文段时，先令拥塞窗口 $cwnd=1$ ，即一个最大报文段长度 MSS。每收到一个对新报文段的确认后，将 $cwnd$ 加 1，即增大一个 MSS。用这样的方法逐步增大发送方的 $cwnd$ ，可使分组注入网络的速率更加合理。

例如，A 向 B 发送数据，发送方先置拥塞窗口 $cwnd=1$ ，A 发送第一个报文段，A 收到 B 对第一个报文段的确认后，把 $cwnd$ 从 1 增大到 2；于是 A 接着发送两个报文段，A 收到 B 对这两个报文段的确认后，把 $cwnd$ 从 2 增大到 4，下次就可一次发送 4 个报文段。

慢开始的“慢”并不是指拥塞窗口 $cwnd$ 的增长速率慢，而是指在 TCP 开始发送报文段时先设置 $cwnd=1$ ，使得发送方在开始时只发送一个报文段（目的是试探一下网络的拥塞情况），然后再逐渐增大 $cwnd$ ，这对防止网络出现拥塞是一个非常有力的措施。使用慢开始算法后，每经过一个传输轮次（即往返时延 RTT）， $cwnd$ 就会加倍，即 $cwnd$ 的值随传输轮次指数规律增长。这样，慢开始一直把 $cwnd$ 增大到一个规定的慢开始门限 $ssthresh$ （阈值），然后改用拥塞避免算法。

(2) 拥塞避免算法

拥塞避免算法的思路是让拥塞窗口 $cwnd$ 缓慢增大，具体做法是：每经过一个往返时延 RTT 就把发送方的拥塞窗口 $cwnd$ 加 1，而不是加倍，使拥塞窗口 $cwnd$ 按线性规律缓慢增长(即加法增大)，这比慢开始算法的拥塞窗口增长速率要缓慢得多。

根据 $cwnd$ 的大小执行不同的算法，可归纳如下：

- 当 $cwnd < ssthresh$ 时，使用慢开始算法。
- 当 $cwnd > ssthresh$ 时，停止使用慢开始算法而改用拥塞避免算法。
- 当 $cwnd = ssthresh$ 时，既可使用慢开始算法，又可使用拥塞避免算法（通常做法）。

(3) 网络拥塞的处理

无论在慢开始阶段还是在拥塞避免阶段，只要发送方判断网络出现拥塞（未按时收到确认），就要把慢开始门限 $ssthresh$ 设置为出现拥塞时的发送方的 $cwnd$ 值的一半（但不能小于 2）。然后把拥塞窗口 $cwnd$ 重新设置为 1，执行慢开始算法。这样做的目的是迅速减少主机发送到网络中的分组数，使得发生拥塞的路由器有足够时间把队列中积压的分组处理完。

慢开始和拥塞避免算法的实现过程如图 5.11 所示。

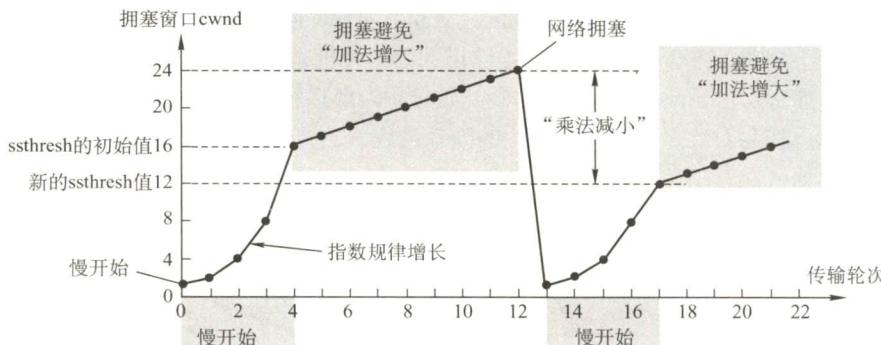


图 5.11 慢开始和拥塞避免算法的实现过程

- 初始时，拥塞窗口置为 1，即 $cwnd=1$ ，慢开始门限置为 16，即 $ssthresh=16$ 。
- 慢开始阶段， $cwnd$ 的初值为 1，以后发送方每收到一个确认 ACK， $cwnd$ 值加 1，也即经过每个传输轮次 (RTT)， $cwnd$ 呈指数规律增长。当拥塞窗口 $cwnd$ 增长到慢开始门限 $ssthresh$ 时（即当 $cwnd=16$ 时），就改用拥塞避免算法， $cwnd$ 按线性规律增长。
- 假定 $cwnd=24$ 时网络出现超时，更新 $ssthresh$ 值为 12（即变为超时时 $cwnd$ 值的一半）， $cwnd$ 重置为 1，并执行慢开始算法，当 $cwnd=12$ 时，改为执行拥塞避免算法。

注意：在慢开始(指数级增长)阶段，若 $2cwnd > ssthresh$ ，则下一个 RTT 后的 $cwnd$ 等于 $ssthresh$ ，而不等于 $2cwnd$ ，即 $cwnd$ 不能跃过 $ssthresh$ 值。如图 5.11 所示，在第 16 个轮次时 $cwnd=8$ 、 $ssthresh=12$ ，则在第 17 个轮次时 $cwnd=12$ ，而不等于 16。

在慢开始和拥塞避免算法中使用了“乘法减小”和“加法增大”方法。“乘法减小”是指不论是在慢开始阶段还是在拥塞避免阶段，只要出现超时（即很可能出现了网络拥塞），就把慢开始门限值 $ssthresh$ 设置为当前拥塞窗口的一半（并执行慢开始算法）。当网络频繁出现拥塞时， $ssthresh$ 值就下降得很快，以大大减少注入网络的分组数。而“加法增大”是指执行拥塞避免算法后，在收到对所有报文段的确认后（即经过一个 RTT），就把拥塞窗口 $cwnd$ 增加一个 MSS 大小，使拥塞窗口缓慢增大，以防止网络过早出现拥塞。

拥塞避免并不能完全避免拥塞。利用以上措施要完全避免网络拥塞是不可能的。拥塞避免是指在拥塞避免阶段把拥塞窗口控制为按线性规律增长，使网络比较不容易出现拥塞。

2. 快重传和快恢复

快重传和快恢复算法是对慢开始和拥塞避免算法的改进。

(1) 快重传

在上一节介绍的 TCP 可靠传输机制中，快重传技术使用了冗余 ACK 来检测丢包的发生。同样，冗余 ACK 也用于网络拥塞的检测（丢了包当然意味着网络可能出现了拥塞）。快重传并非取消重传计时器，而是在某些情况下可更早地重传丢失的报文段。

当发送方连续收到三个重复的 ACK 报文时，直接重传对方尚未收到的报文段，而不必等待那个报文段设置的重传计时器超时。

(2) 快恢复

快恢复算法的原理如下：当发送方连续收到三个冗余 ACK（即重复确认）时，执行“乘法减小”算法，把慢开始门限 ssthresh 设置为此时发送方 cwnd 的一半。这是为了预防网络发生拥塞。但发送方现在认为网络很可能没有发生（严重）拥塞，否则就不会有几个报文段连续到达接收方，也不会连续收到重复确认。因此与慢开始不同之处是它把 cwnd 值设置为慢开始门限 ssthresh 改变后的数值，然后开始执行拥塞避免算法（“加法增大”），使拥塞窗口缓慢地线性增大。

由于跳过了拥塞窗口 cwnd 从 1 起始的慢开始过程，所以被称为快恢复。快恢复算法的实现过程如图 5.12 所示，作为对比，虚线为慢开始的处理过程。

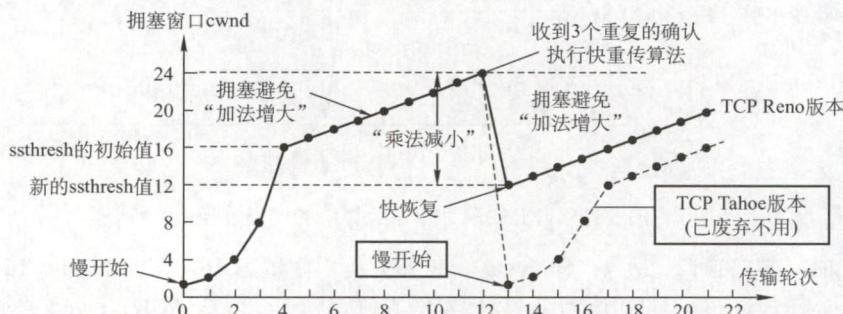


图 5.12 快恢复算法的实现过程

在流量控制中，发送方发送数据的量由接收方决定，而在拥塞控制中，则由发送方自己通过检测网络状况来决定。实际上，慢开始、拥塞避免、快重传和快恢复几种算法是同时应用在拥塞控制机制中。四种算法使用的总结：在 TCP 连接建立和网络出现超时时，采用慢开始和拥塞避免算法；当发送方接收到冗余 ACK 时，采用快重传和快恢复算法。

在本节的最后，再次提醒读者：接收方的缓存空间总是有限的。因此，发送方发送窗口的实际大小由流量控制和拥塞控制共同决定。当题目中同时出现接收窗口 (rwnd) 和拥塞窗口 (cwnd) 时，发送方实际的发送窗口大小是由 rwnd 和 cwnd 中较小的那个确定的。

5.3.7 本节习题精选

一、单项选择题

01. 下列关于传输层协议中面向连接的描述，() 是错误的。

- A. 面向连接的服务需要经历 3 个阶段：连接建立、数据传输及连接释放

- B. 当链路不发生错误时，面向连接的服务可以保证数据到达的顺序是正确的
C. 面向连接的服务有很高的效率和时间性能
D. 面向连接的服务提供了一个可靠的数据流
02. TCP 协议规定 HTTP () 进程的端口号为 80。
A. 客户机 B. 解析 C. 服务器 D. 主机
03. 下列 () 不是 TCP 服务的特点。
A. 字节流 B. 全双工 C. 可靠 D. 支持广播
04. () 字段包含在 TCP 首部中，而不包含在 UDP 首部中。
A. 目的端口号 B. 序列号 C. 校验和 D. 目的 IP 地址
05. 以下关于 TCP 报头格式的描述中，错误的是 ()。
A. 报头长度为 20~60B，其中固定部分为 20B
B. 端口号字段依次表示源端口号与目的端口号
C. 报头长度总是 4 的倍数个字节
D. TCP 校验和伪首部中 IP 分组头的协议字段为 17
06. 在采用 TCP 连接的数据传输阶段，如果发送端的发送窗口值由 1000 变为 2000，那么发送端在收到一个确认之前可以发送 ()。
A. 2000 个 TCP 报文段 B. 2000B
C. 1000B D. 1000 个 TCP 报文段
07. A 和 B 建立了 TCP 连接，当 A 收到确认号为 100 的确认报文段时，表示 ()。
A. 报文段 99 已收到
B. 报文段 100 已收到
C. 末字节序号为 99 的报文段已收到
D. 末字节序号为 100 的报文段已收到
08. 为保证数据传输的可靠性，TCP 采用了对 () 确认的机制。
A. 报文段 B. 分组 C. 字节 D. 比特
09. 在 TCP 协议中，发送方的窗口大小取决于 ()。
A. 仅接收方允许的窗口
B. 接收方允许的窗口和发送方允许的窗口
C. 接收方允许的窗口和拥塞窗口
D. 发送方允许的窗口和拥塞窗口
10. 滑动窗口的作用是 ()。
A. 流量控制 B. 拥塞控制 C. 路由控制 D. 差错控制
11. 以下关于 TCP 工作原理与过程的描述中，错误的是 ()。
A. TCP 连接建立过程需要经过“三次握手”的过程
B. TCP 传输连接建立后，客户端与服务器端的应用进程进行全双工的字节流传输
C. TCP 传输连接的释放过程很复杂，只有客户端可以主动提出释放连接的请求
D. TCP 连接的释放需要经过“四次挥手”的过程
12. TCP 的滑动窗口协议中，规定重传分组的数量最多可以 ()。
A. 是任意的
B. 1 个
C. 大于滑动窗口的大小
D. 等于滑动窗口的大小
13. TCP 中滑动窗口的值设置太大，对主机的影响是 ()。

关注公众号【乘龙考研】
一手更新 稳定有保障

- A. 由于传送的数据过多而使路由器变得拥挤，主机可能丢失分组
 B. 产生过多的 ACK
 C. 由于接收的数据多，而使主机的工作速度加快
 D. 由于接收的数据多，而使主机的工作速度变慢
14. 以下关于 TCP 窗口与拥塞控制概念的描述中，错误的是（ ）。
 A. 接收端窗口 (rwnd) 通过 TCP 首部中的窗口字段通知数据的发送方
 B. 发送窗口确定的依据是：发送窗口 = min[接收端窗口, 拥塞窗口]
 C. 拥塞窗口是接收端根据网络拥塞情况确定的窗口值
 D. 拥塞窗口大小在开始时可以按指数规律增长
15. TCP 使用三次握手协议来建立连接，设 A、B 双方发送报文的初始序列号分别为 X 和 Y ，
 A 发送 (①) 的报文给 B，B 接收到报文后发送 (②) 的报文给 A，然后 A 发送一个确认报文给 B 便建立了连接 (注意，ACK 的下标为捎带的序号)。
 ① A. SYN=1, 序号= X B. SYN=1, 序号= $X+1$, ACK _{X} =1
 C. SYN=1, 序号= Y D. SYN=1, 序号= Y , ACK _{$Y+1$} =1
 ② A. SYN=1, 序号= $X+1$ B. SYN=1, 序号= $X+1$, ACK _{X} =1
 C. SYN=1, 序号= Y , ACK _{$X+1$} =1 D. SYN=1, 序号= Y , ACK _{$Y+1$} =1
16. TCP “三次握手”过程中，第二次“握手”时，发送的报文段中（ ）标志位被置为 1。
 A. SYN B. ACK
 C. ACK 和 RST D. SYN 和 ACK
17. A 和 B 之间建立了 TCP 连接，A 向 B 发送了一个报文段，其中序号字段 seq=200，确认号字段 ack=201，数据部分有 2 个字节，那么在 B 对该报文的确认报文段中（ ）。
 A. seq=202, ack=200 B. seq=201, ack=201
 C. seq=201, ack=202 D. seq=202, ack=201
18. TCP 的通信双方，有一方发送了带有 FIN 标志的数据段后，表示（ ）。
 A. 将断开通信双方的 TCP 连接
 B. 单方面释放连接，表示本方已经无数据发送，但可以接收对方的数据
 C. 中止数据发送，双方都不能发送数据
 D. 连接被重新建立
19. 一个 TCP 连接的数据传输阶段，如果发送端的发送窗口值由 2000 变为 3000，那么意味着发送端可以（ ）。
 A. 在收到一个确认之前可以发送 3000 个 TCP 报文段
 B. 在收到一个确认之前可以发送 1000B
 C. 在收到一个确认之前可以发送 3000B
 D. 在收到一个确认之前可以发送 2000 个 TCP 报文段
20. 在一个 TCP 连接中，MSS 为 1KB，当拥塞窗口为 34KB 时发生了超时事件。如果在接下来的 4 个 RTT 内报文段传输都是成功的，那么当这些报文段均得到确认后，拥塞窗口的大小是（ ）。
 A. 8KB B. 9KB C. 16KB D. 17KB
21. 设 TCP 的拥塞窗口的慢开始门限值初始为 8 (单位为报文段)，当拥塞窗口上升到 12 时发生超时，TCP 开始慢启动和拥塞避免，那么第 13 次传输时拥塞窗口的大小为（ ）。
 A. 4 B. 6 C. 7 D. 8

22. 在一个 TCP 连接中, MSS 为 1KB, 当拥塞窗口为 34KB 时收到了 3 个冗余 ACK 报文。如果在接下来的 4 个 RTT 内报文段传输都是成功的, 那么当这些报文段均得到确认后, 拥塞窗口的大小是 ()。
- A. 8KB B. 16KB C. 20KB D. 21KB
23. A 和 B 建立 TCP 连接, MSS 为 1KB。某时, 慢开始门限值为 2KB, A 的拥塞窗口为 4KB, 在接下来的一个 RTT 内, A 向 B 发送了 4KB 的数据 (TCP 的数据部分), 并且得到了 B 的确认, 确认报文中的窗口字段的值为 2KB。在下一个 RTT 中, A 最多能向 B 发送() 数据。
- A. 2KB B. 8KB C. 5KB D. 4KB
24. 假设在没有发生拥塞的情况下, 在一条往返时延 RTT 为 10ms 的线路上采用慢开始控制策略。如果接收窗口的大小为 24KB, 最大报文段 MSS 为 2KB。那么发送方能发送出第一个完全窗口 (也就是发送窗口达到 24KB) 需要的时间是 ()。
- A. 30ms B. 40ms C. 50ms D. 60ms
25. 如果主机 1 的进程以端口 x 和主机 2 的端口 y 建立了一条 TCP 连接, 这时如果希望再在这两个端口间建立一个 TCP 连接, 那么会 ()。
- A. 建立失败, 不影响先建立连接的传输
B. 建立成功, 且两个连接都可以正常传输
C. 建立成功, 先建立的连接被断开
D. 建立失败, 两个连接都被断开
26. 【2009 统考真题】主机甲与主机乙之间已建立一个 TCP 连接, 主机甲向主机乙发送了两个连续的 TCP 段, 分别包含 300B 和 500B 的有效载荷, 第一个段的序列号为 200, 主机乙正确接收到这两个数据段后, 发送给主机甲的确认序列号是 ()。
- A. 500 B. 700 C. 800 D. 1000
27. 【2009 统考真题】一个 TCP 连接总以 1KB 的最大段长发送 TCP 段, 发送方有足够的数据要发送, 当拥塞窗口为 16KB 时发生了超时, 如果接下来的 4 个 RTT 时间内的 TCP 段的传输都是成功的, 那么当第 4 个 RTT 时间内发送的所有 TCP 段都得到肯定应答时, 拥塞窗口大小是 ()。
- A. 7KB B. 8KB C. 9KB D. 16KB
28. 【2010 统考真题】主机甲和主机乙之间已建立一个 TCP 连接, TCP 最大段长为 1000B。若主机甲的当前拥塞窗口为 4000B, 在主机甲向主机乙连续发送两个最大段后, 成功收到主机乙发送的第一个段的确认段, 确认段中通告的接收窗口大小为 2000B, 则此时主机甲还可以向主机乙发送的最大字节数是 ()。
- A. 1000 B. 2000 C. 3000 D. 4000
29. 【2011 统考真题】主机甲向主机乙发送一个 ($SYN=1$, $seq=11220$) 的 TCP 段, 期望与主机乙建立 TCP 连接, 若主机乙接受该连接请求, 则主机乙向主机甲发送的正确的 TCP 段可能是 ()。
- A. ($SYN=0$, $ACK=0$, $seq=11221$, $ack=11221$)
B. ($SYN=1$, $ACK=1$, $seq=11220$, $ack=11220$)
C. ($SYN=1$, $ACK=1$, $seq=11221$, $ack=11221$)
D. ($SYN=0$, $ACK=0$, $seq=11220$, $ack=11220$)
30. 【2011 统考真题】主机甲与主机乙之间已建立一个 TCP 连接, 主机甲向主机乙发送

关注公众号【乘龙考研】
一手更新 稳定有保障

了3个连续的TCP段，分别包含300B、400B和500B的有效载荷，第3个段的序号为900。若主机乙仅正确接收到第1个段和第3个段，则主机乙发送给主机甲的确认序号是()。

- A. 300 B. 500 C. 1200 D. 1400

31. 【2013统考真题】主机甲与主机乙之间已建立一个TCP连接，双方持续有数据传输，且数据无差错与丢失。若甲收到一个来自乙的TCP段，该段的序号为1913、确认序号为2046、有效载荷为100B，则甲立即发送给乙的TCP段的序号和确认序号分别是()。

- A. 2046、2012 B. 2046、2013 C. 2047、2012 D. 2047、2013

32. 【2014统考真题】主机甲和乙建立了TCP连接，甲始终以MSS=1KB大小的段发送数据，并一直有数据发送；乙每收到一个数据段都会发出一个接收窗口为10KB的确认段。若甲在t时刻发生超时的时候拥塞窗口为8KB，则从t时刻起，不再发生超时的情况下，经过10个RTT后，甲的发送窗口是()。

- A. 10KB B. 12KB C. 14KB D. 15KB

33. 【2015统考真题】主机甲和主机乙新建一个TCP连接，甲的拥塞控制初始阈值为32KB，甲向乙始终以MSS=1KB大小的段发送数据，并一直有数据发送；乙为该连接分配16KB接收缓存，并对每个数据段进行确认，忽略段传输延迟。若乙收到的数据全部存入缓存，不被取走，则甲从连接建立成功时刻起，未出现发送超时的情况下，经过4个RTT后，甲的发送窗口是()。

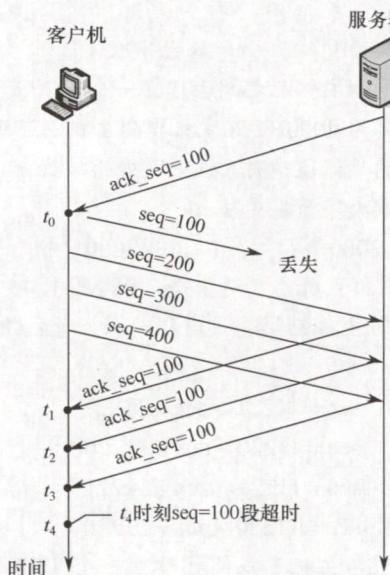
- A. 1KB B. 8KB C. 16KB D. 32KB

34. 【2017统考真题】若甲向乙发起一个TCP连接，最大段长MSS=1KB，RTT=5ms，乙开辟的接收缓存为64KB，则甲从连接建立成功至发送窗口达到32KB，需经过的时间至少是()。

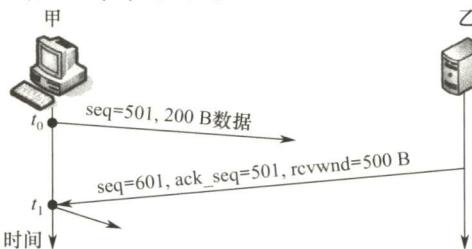
- A. 25ms B. 30ms C. 160ms D. 165ms

35. 【2019统考真题】某客户通过一个TCP连接向服务器发送数据的部分过程如下图所示。客户在 t_0 时刻第一次收到确认序列号ack_seq=100的段，并发送序列号seq=100的段，但发生丢失。若TCP支持快速重传，则客户重新发送seq=100段的时刻是()。

- A. t_1 B. t_2 C. t_3 D. t_4



36. 【2019 统考真题】若主机甲主动发起一个与主机乙的 TCP 连接，甲、乙选择的初始序列号分别为 2018 和 2046，则第三次握手 TCP 段的确认序列号是（ ）。
- A. 2018 B. 2019 C. 2046 D. 2047
37. 【2020 统考真题】若主机甲与主机乙已建立一条 TCP 连接，最大段长 (MSS) 为 1KB，往返时间 (RTT) 为 2ms，则在不出现拥塞的前提下，拥塞窗口从 8KB 增长到 32KB 所需的最长时间是（ ）。
- A. 4ms B. 8ms C. 24ms D. 48ms
38. 【2020 统考真题】若主机甲与主机乙建立 TCP 连接时，发送的 SYN 段中的序号为 1000，在断开连接时，甲发送给乙的 FIN 段中的序号为 5001，则在无任何重传的情况下，甲向乙已经发送的应用层数据的字节数为（ ）。
- A. 4002 B. 4001 C. 4000 D. 3999
39. 【2021 统考真题】若客户首先向服务器发送 FIN 段请求断开 TCP 连接，则当客户收到服务器发送的 FIN 段并向服务器发送 ACK 段后，客户的 TCP 状态转换为（ ）。
- A. CLOSE_WAIT B. TIME_WAIT C. FIN_WAIT_1 D. FIN_WAIT_2
40. 【2021 统考真题】若大小为 12B 的应用层数据分别通过 1 个 UDP 数据报和 1 个 TCP 段传输，则该 UDP 数据报和 TCP 段实现的有效载荷（应用层数据）最大传输效率分别是（ ）。
- A. 37.5%, 16.7% B. 37.5%, 37.5% C. 60.0%, 16.7% D. 60.0%, 37.5%
41. 【2021 统考真题】设主机甲通过 TCP 向主机乙发送数据，部分过程如下图所示。甲在 t_0 时刻发送一个序号 seq=501、封装 200B 数据的段，在 t_1 时刻收到乙发送的序号 seq=601、确认序号 ack_seq=501、接收窗口 rwnd=500B 的段，则甲在未收到新的确认段之前，可以继续向乙发送的数据序号范围是（ ）。



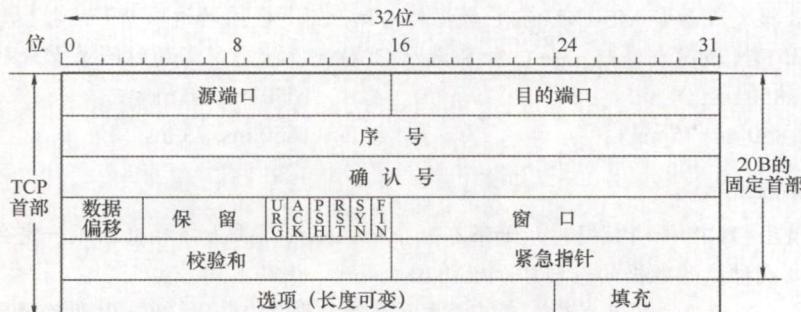
- A. 501 ~ 1000 B. 601 ~ 1100 C. 701 ~ 1000 D. 801 ~ 1100
42. 【2022 统考真题】假设主机甲和主机乙已建立一个 TCP 连接，最大段长 MSS=1KB，甲一直向乙发送数据，当甲的拥塞窗口为 16KB 时，计时器发生了超时，则甲的拥塞窗口再次增长到 16KB 所需要的时间至少是（ ）。
- A. 4 RTT B. 5 RTT C. 11 RTT D. 16 RTT
43. 【2022 统考真题】假设客户 C 和服务器 S 已建立一个 TCP 连接，通信往返时间 RTT=50ms，最长报文寿命 MSL=80ms，数据传输结束后，C 主动请求断开连接。若从 C 主动向 S 发出 FIN 段时刻算起，则 C 和 S 进入 CLOSED 状态所需的时间至少分别是（ ）。
- A. 850 ms, 50 ms B. 1650 ms, 50 ms
C. 850 ms, 75 ms D. 1650 ms, 75 ms

二、综合应用题

01. 在使用 TCP 传输数据时，如果有一个确认报文段丢失，那么也不一定会引起与该确认报文段对应的数据的重传。试说明理由。
02. 如果收到的报文段无差错，只是报文段失序，那么 TCP 对此未做明确规定，而是让 TCP

的实现者自行确定。试讨论两种可能的方法的优劣：

- 1) 将失序报文段丢弃。
- 2) 先将失序报文段暂存于接收缓存内，待所缺序号的报文段收齐后再一起上交应用层。
03. 一个 TCP 连接要发送 3200B 的数据。第一个字节的编号为 10010。如果前两个报文各携带 1000B 的数据，最后一个携带剩下的数据，请写出每个报文段的序号。
04. 设 TCP 使用的最大窗口尺寸为 64KB，TCP 报文在网络上的平均往返时间为 20ms，问 TCP 协议所能得到的最大吞吐量是多少？（假设传输信道的带宽是不受限的。）
05. 网络允许的最大报文段的长度为 128B，序号用 8 位表示，报文段在网络中的寿命为 30s。求每条 TCP 连接所能达到的最高数据率。
06. 在一个 TCP 连接中，信道带宽为 1Gb/s，发送窗口固定为 65535B，端到端时延为 20ms。可以取得的最大吞吐率是多少？线路效率是多少？（发送时延忽略不计，TCP 及其下层协议首部长度忽略不计，最大吞吐率 = 一个 RTT 传输的有效数据/一个 RTT 的时间。）
07. 主机 A 基于 TCP 向主机 B 连续发送 3 个 TCP 报文段。第 1 个报文段的序号为 90，第 2 个报文段的序号为 120，第 3 个报文段的序号为 150。
 - 1) 第 1、2 个报文段中有多少数据？
 - 2) 假设第 2 个报文段丢失而其他两个报文段到达主机 B，在主机 B 发往主机 A 的确认报文中，确认号应是多少？
08. 考虑在一条具有 10ms 来回路程时间的线路上采用慢启动拥塞控制而不发生网络拥塞情况下的效应，接收窗口为 24KB，且最大段长为 2KB。那么需要多长时间才能发送第一个完全窗口？
09. 设 TCP 的拥塞窗口的慢开始门限值初始为 12（单位为报文段），当拥塞窗口达到 16 时出现超时，再次进入慢启动过程。从此时起若恢复到超时时刻的拥塞窗口大小，需要的往返次数是多少？
10. 假定 TCP 报文段载荷是 1500B，最大分组存活时间是 120s，那么要使得 TCP 报文段的序列号不会循环回来而重叠，线路允许的最快速度是多大？（不考虑帧长限制。）
11. 一个 TCP 连接使用 256kb/s 的链路，其端到端时延为 128ms。经测试发现吞吐率只有 128kb/s。问窗口是多少？忽略 PDU 封装的协议开销及接收方应答分组的发送时间（假定应答分组长度很小）。
12. 假定 TCP 最大报文段的长度是 1KB，拥塞窗口被置为 18KB，并且发生了超时事件。如果接着的 4 次逆发量传输都是成功的，那么该窗口将是多大？
13. 一个 TCP 首部的数据信息（十六进制表示）为 0x0D 28 00 15 50 5F A9 06 00 00 00 00 70 02 40 00 C0 29 00 00。TCP 首部的格式如下图所示。请回答：



- 1) 源端口号和目的端口号各是多少?
- 2) 发送的序列号是多少? 确认号是多少?
- 3) TCP 首部的长度是多少?
- 4) 这是一个使用什么协议的 TCP 连接? 该 TCP 连接的状态是什么?
14. 【2012 统考真题】主机 H 通过快速以太网连接 Internet, IP 地址为 192.168.0.8, 服务器 S 的 IP 地址为 211.68.71.80。H 与 S 使用 TCP 通信时, 在 H 上捕获的其中 5 个 IP 分组如表 1 所示。

表 1

编 号	IP 分组的前 40B 内容 (十六进制)				
1	45 00 00 30	01 9b 40 00	80 06 1d e8	c0 a8 00 08	d3 44 47 50
	0b d9 13 88	84 6b 41 c5	00 00 00 00	70 02 43 80	5d b0 00 00
2	45 00 00 30	00 00 40 00	31 06 6e 83	d3 44 47 50	c0 a8 00 08
	13 88 0b d9	e0 59 9f ef	84 6b 41 c6	70 12 16 d0	37 e1 00 00
3	45 00 00 28	01 9c 40 00	80 06 1d ef	c0 a8 00 08	d3 44 47 50
	0b d9 13 88	84 6b 41 c6	e0 59 9f f0	50 f0 43 80	2b 32 00 00
4	45 00 00 38	01 9d 40 00	80 06 1d de	c0 a8 00 08	d3 44 47 50
	0b d9 13 88	84 6b 41 c6	e0 59 9f f0	50 18 43 80	e6 55 00 00
5	45 00 00 28	68 11 40 00	31 06 06 7a	d3 44 47 50	c0 a8 00 08
	13 88 0b d9	e0 59 9f f0	84 6b 41 d6	50 10 16 d0	57 d2 00 00

回答下列问题:

- 表 1 中的 IP 分组中, 哪几个是由 H 发送的? 哪几个完成了 TCP 连接建立过程? 哪几个在通过快速以太网传输时进行了填充?
- 根据表 1 中的 IP 分组, 分析 S 已经收到的应用层数据字节数是多少。
- 若表 1 中的某个 IP 分组在 S 发出时的前 40B 如表 2 所示, 则该 IP 分组到达 H 时经过了多少个路由器?

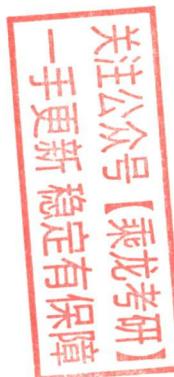
表 2

来自 S 的分组	45 00 00 28	68 11 40 00	40 06 ec ad	d3 44 47 50	ca 76 01 06
	13 88 a1 08	e0 59 9f f0	84 6b 41 d6	50 10 16 d0	b7 d6 00 00

IP 分组头和 TCP 段头结构分别如图 1 和图 2 所示。



图 1 IP 分组头结构



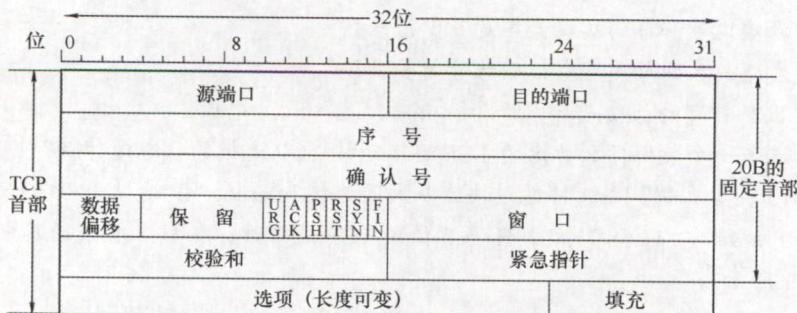
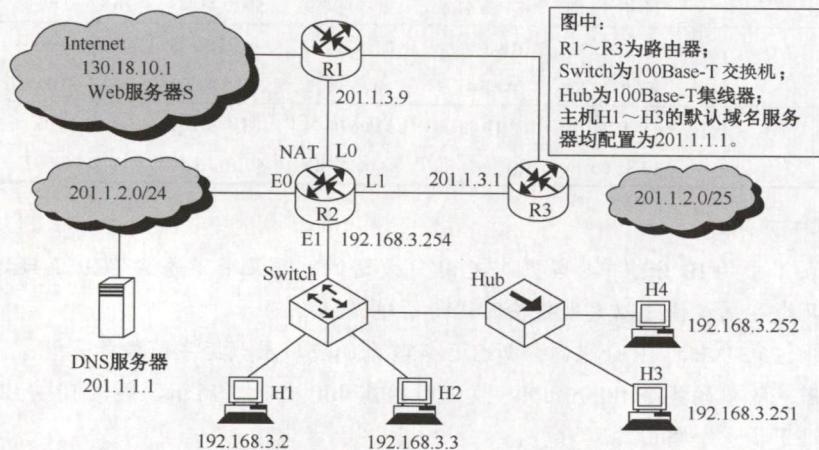


图 2 TCP 段头结构

15. 【2016 统考真题】假设下图中的 H3 访问 Web 服务器 S 时，S 为新建的 TCP 连接分配了 20KB (K=1024) 的接收缓存，最大段长 MSS=1KB，平均往返时间 RTT=200ms。H3 建立连接时的初始序号为 100，且持续以 MSS 大小的段向 S 发送数据，拥塞窗口初始阈值为 32KB；S 对收到的每个段进行确认，并通告新的接收窗口。假定 TCP 连接建立完成后，S 端的 TCP 接收缓存仅有数据存入而无数据取出。请回答下列问题：



- 1) 在 TCP 连接建立过程中，H3 收到的 S 发送过来的第二次握手 TCP 段的 SYN 和 ACK 标志位的值分别是多少？确认序号是多少？
- 2) H3 收到的第 8 个确认段所通告的接收窗口是多少？此时 H3 的拥塞窗口变为多少？H3 的发送窗口变为多少？
- 3) H3 的发送窗口等于 0 时，下一个待发送的数据段序号是多少？H3 从发送第 1 个数据段到发送窗口等于 0 时刻为止，平均数据传输速率是多少？(忽略段的传输延时。)
- 4) 若 H3 与 S 之间通信已经结束，在 t 时刻 H3 请求断开该连接，则从 t 时刻起，S 释放该连接的最短时间是多少？

5.3.8 答案与解析

一、单项选择题

01. C

由于面向连接的服务需要建立连接，且需要保证数据的有序性和正确性，因此它比无连接的服务开销大，而速度和效率方面也要比无连接的服务差一些。

02. C

TCP 中端口号 80 标识 Web 服务器端的 HTTP 进程，客户端访问 Web 服务器的 HTTP 进程的端口号由客户端的操作系统动态分配。因此答案为选项 C。

03. D

TCP 提供的是一对一全双工可靠的字节流服务，所以 TCP 并不支持广播。

04. B

TCP 报文段和 UDP 数据报都包含源端口、目的端口、校验号。由于 UDP 提供不可靠的传输服务，不需要对报文编号，因此不会有序列号字段，而 TCP 提供可靠的传输服务，因此需要设置序列号字段。目的 IP 地址属于 IP 数据报中的内容。

05. D

TCP 伪首部与 UDP 伪首部一样，包括 IP 分组首部的一部分。IP 首部中有一个协议字段，用于指明上层协议是 TCP 还是 UDP。17 代表 UDP，6 代表 TCP，所以 D 错误。对于 A 选项，由于数据偏移字段的单位是 4B，也就是说当偏移取最大时 TCP 首部长度为 $15 \times 4 = 60$ B。由于使用填充，所以长度总是 4B 的倍数，选项 C 正确。

06. B

TCP 使用滑动窗口机制来进行流量控制。在 ACK 应答信息中，TCP 在接收端用 ACK 加上接收方允许接收数据范围的最大值回送给发送方，发送方把这个最大值当作发送窗口值，表明发送端在未收到确认之前可以发送的最大字节数，即 2000B。

07. C

TCP 的确认号是指明接收方下一次希望收到的报文段的数据部分第一个字节的编号，可以看出，前一个已收到的报文段的最后一个字节的编号为 99，所以选项 C 正确。报文段的序号是其数据部分第一个字节的编号。选项 A、B 不正确，因为有可能已收到的这个报文的数据部分不止一个字节，那么报文段的编号就不为 99，但可以说编号为 99 的字节已收到。

08. A

TCP 是面向字节的。对每个字节进行编号，但并不是接收到每个字节都要发回确认，而是在发送一个报文段的字节后才发回一个确认，所以 TCP 采用的是对报文段的确认机制。

09. C

TCP 让每个发送方仅发送正确数量的数据，保持网络资源被利用但又不会过载。为了避免网络拥塞和接收方缓冲区溢出，TCP 发送方在任意时刻可以发送的最大数据流是接收方允许的窗口和拥塞窗口中的最小值。

10. A

TCP 采用大小可变的滑动窗口进行流量控制。

11. C

参与 TCP 连接的两个进程中的任何一个都能提出释放连接的请求。

12. D

TCP 滑动窗口协议中发送方滑动窗口的大小规定了发送方最多能够传送的分组数目，只有窗口滑动了，才能往后继续发送。分组重传的最大值也是发送方能发送数据的最大值，因而重传分组的数量最多也不能超过滑动窗口的大小。

13. A

TCP 使用滑动窗口机制来进行流量控制，其窗口尺寸的设置很重要，如果滑动窗口值设置得太小，那么会产生过多的 ACK（因为窗口大可以累积确认，因此会有更少的 ACK）；如果设置得



太大，那么又会由于传送的数据过多而使路由器变得拥挤，导致主机可能丢失分组。

14. C

拥塞窗口是发送端根据网络拥塞情况确定的窗口值。

15. A、C

TCP 使用三次握手来建立连接，第一次握手 A 发给 B 的 TCP 报文中应置其首部 SYN 位为 1，并选择序号 $seq = X$ ，表明传送数据时的第一个数据字节的序号是 X ；在第二次握手中，即 B 接收到报文后，发给 A 的确认报文段中应使 $SYN = 1$ ，使 $ACK = 1$ ，且确认号 $ACK = X + 1$ ，即 $ACK_{X+1} = 1$ （ACK 的下标为捎带的序号），同时告诉自己选择的序号 $seq = Y$ 。

16. D

在 TCP 的“三次握手”中，第二次握手时，SYN 和 ACK 均被置为 1。

17. C

在 A 发向 B 的报文中， seq 表示发送的报文段中数据部分的第一个字节在 A 的发送缓存区中的编号， ack 表示 A 期望收到的下一个报文段的数据部分的第一个字节在 B 的发送缓存区中的编号。因此，同一个 TCP 报文中的 seq 和 ack 的值是没有联系的。在 B 发给 A 的报文（捎带确认）中， seq 值应和 A 发向 B 的报文中的 ack 值相同，即 201； ack 值表示 B 期望下次收到 A 发出的报文段的第一个字节的编号，应是 $200 + 2 = 202$ 。

18. B

FIN 位用来释放一个连接，它表示本方已没有数据要传输。然而，在关闭一个连接后，对方还可以继续发送数据，所以还有可能接收到数据。

19. C

TCP 提供的是可靠的字节流传输服务，使用窗口机制进行流量控制与拥塞控制。TCP 的滑动窗口机制是面向字节的，因此窗口大小的单位为字节。假设发送窗口的大小为 N ，这意味着发送端可以在没有收到确认的情况下连续发送 N 个字节。

20. C

在拥塞窗口为 34KB 时发生了超时，那么慢开始门限值（ $ssthresh$ ）就被设定为 17KB，并且在第一个 RTT 中拥塞窗口（ $cwnd$ ）置为 1KB。按照慢开始算法，第二个 RTT 中 $cwnd = 2KB$ ，第三个 RTT 中 $cwnd = 4KB$ ，第四个 RTT 中 $cwnd = 8KB$ 。当第四个 RTT 中发出去的 8 个报文段的确认报文收到后， $cwnd = 16KB$ （此时还未超过慢开始门限值）。所以选 C。本题中“这些报文段均得到确认后”这句话很重要。

21. C

在慢开始和拥塞避免算法中，拥塞窗口初始为 1，窗口大小开始按指数增长。当拥塞窗口大于慢开始门限后停止使用慢开始算法，改用拥塞避免算法。此处慢开始的门限值初始为 8，当拥塞窗口增大到 8 时改用拥塞避免算法，窗口大小按线性增长，每次增加 1 个报文段，当增加到 12 时，出现超时，重新设门限值为 6（12 的一半），拥塞窗口再重新设为 1，执行慢开始算法，到门限值 6 时执行拥塞避免算法。

这样，拥塞窗口的变化就为 1, 2, 4, 8, 9, 10, 11, 12, 1, 2, 4, 6, 7, 8, 9, …，其中第 13 次传输时拥塞窗口的大小为 7。

22. D

条件“收到了 3 个冗余 ACK 报文”说明此时应执行快恢复算法，因此慢开始门限值设为 17KB，并在接下来的第一个 RTT 中 $cwnd$ 也被设为 17KB，第二个 RTT 中 $cwnd = 18$ ，第三个 RTT 中 $cwnd = 19KB$ ，第四个 RTT 中 $cwnd = 20KB$ ，第四个 RTT 中发出的报文全部得到确认后， $cwnd$ 再

增加1KB，变为21KB。注意cwnd的增加都发生在收到确认报文后。

23. A

本题中出现了拥塞窗口和接收端窗口，为了保证B的接收缓存不发生溢出，发送窗口应该取两者的最小值。先看拥塞窗口，由于慢开始门限值为2KB，第一个RTT中A拥塞窗口为4KB，按照拥塞避免算法，收到B的确认报文后，拥塞窗口增长为5KB。再看接收端窗口，B通过确认报文中窗口字段向A通知接收端窗口，那么接收端窗口为2KB。因此在下一次发送数据时，A的发送窗口应该为2KB，即一个RTT内最多发送2KB。所以选项A正确。

24. B

按照慢开始算法，发送窗口的初始值为拥塞窗口的初始值，即MSS的大小2KB，然后依次增大为4KB、8KB、16KB，然后是接收窗口的大小24KB，即达到第一个完全窗口。因此达到第一个完全窗口所需要的时间为 $4\text{RTT} = 40\text{ms}$ 。

25. A

一条连接使用它们的套接字来表示，因此 $(1, x) - (2, y)$ 是在两个端口之间唯一可能的连接。而后建立的连接会被阻止。

26. D

返回的确认序列号是接收方期待收到对方下一个报文段数据部分的第一个字节的序号，因此乙在正确接收到两个段后，返回给甲的确认序列号是 $200 + 300 + 500 = 1000$ 。

27. C

发生超时后，慢开始门限 $ssthresh$ 变为 $16\text{KB}/2 = 8\text{KB}$ ，拥塞窗口变为1KB。在接下来的3个RTT内，执行慢开始算法，拥塞窗口大小依次为2KB、4KB、8KB，由于慢开始门限 $ssthresh$ 为8KB，因此之后转而执行拥塞避免算法，即拥塞窗口开始“加法增大”。因此第4个RTT结束后，拥塞窗口的大小为9KB。

28. A

发送方的发送窗口的上限值取接收方窗口和拥塞窗口这两个值中的较小一个，于是此时发送方的发送窗口为 $\min\{4000, 2000\} = 2000\text{B}$ ，由于发送方还未收到第二个最大段的确认，所以此时主机甲还可以向主机乙发送的最大字节数为 $2000 - 1000 = 1000\text{B}$ 。

29. C

在确认报文段中，同步位SYN和确认位ACK必须都是1；返回的确认号ack是甲发送的初始序号 $seq = 11220$ 加1，即 $ack = 11221$ ；同时乙也要选择并消耗一个初始序号seq，seq值由乙的TCP进程任意给出，它与确认号、请求报文段的序号没有任何关系。

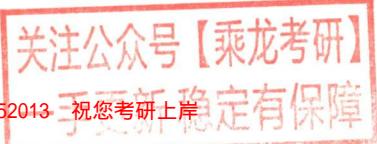
30. B

TCP首部的序号字段是指本报文段数据部分的第一个字节的序号，而确认号是期待收到对方下一个报文段的第一个字节的序号。第三个段的序号为900，则第二个段的序号为 $900 - 400 = 500$ ，现在主机乙期待收到第二个段，因此发给甲的确认号是500。

31. B

确认序号ack是期望收到对方下一个报文段的数据的第一个字节的序号，序号seq是指本报文段所发送的数据的第一个字节的序号。甲收到一个来自乙的TCP段，该段的序号 $seq = 1913$ 、确认序号 $ack = 2046$ 、有效载荷为100B，表明到序号 $1913 + 100 - 1 = 2012$ 为止的所有数据甲均已收到，而乙期望收到下一个报文段的序号从2046开始。因此甲发给乙的TCP段的序号 $seq_1 = ack = 2046$ 和确认序号 $ack_1 = seq + 100 = 2013$ 。

32. A



当 t 时刻发生超时时，把 ssthresh 设为 8 的一半，即 4，把拥塞窗口设为 1KB。然后经历 10 个 RTT 后，拥塞窗口的大小依次为 2, 4, 5, 6, 7, 8, 9, 10, 11, 12，而发送窗口取当时的拥塞窗口和接收窗口的最小值，接收窗口始终为 10KB，所以此时的发送窗口为 10KB，答案为选项 A。

实际上该题接收窗口一直为 10KB，可知不管何时，发送窗口一定小于或等于 10KB，选项中只有选项 A 满足条件，可直接得出答案为选项 A。

33. A

发送窗口的上限值 = $\min\{\text{接收窗口}, \text{拥塞窗口}\}$ 。4 个 RTT 后，乙收到的数据全部存入缓存，不被取走，接收窗口只剩下 1KB ($16 - 1 - 2 - 4 - 8 = 1$) 缓存，使得甲的发送窗口为 1KB。

34. A

按照慢开始算法，发送窗口 = $\min\{\text{拥塞窗口}, \text{接收窗口}\}$ ，初始的拥塞窗口为最大报文段长度 1KB。每经过一个 RTT，拥塞窗口翻倍，因此需至少经过 5 个 RTT，发送窗口才能达到 32KB，所以选 A。这里假定乙能及时处理接收到的数据，空闲的接收缓存 $\geq 32\text{KB}$ 。

35. C

TCP 规定当发送方收到对同一个报文段的 3 个重复确认时，就可以认为跟在这个被确认报文段之后的报文已丢失，立即执行快速重传算法。 t_3 时刻连续收到来自服务器的三个确认序列号 $\text{ack_seq} = 100$ 的段，发送方认为 $\text{seq} = 100$ 的段已经丢失，执行快速重传算法，重新发送 $\text{seq} = 100$ 的段。

36. D

根据 TCP 连接建立的“三次握手”原理，第三次握手时甲发出的确认序列号应为第二次握手时乙发出的序列号 +1，即 2047。

37. D

由于慢开始门限 ssthresh 可以根据需求设置，为了求拥塞窗口从 8KB 增长到 32KB 所需的最长时间，可以假定慢开始门限小于或等于 8KB，只要不出现拥塞，拥塞窗口就都是加法增大，每经历一个传输轮次 (RTT)，拥塞窗口逐次加 1，因此所需的最长时间为 $(32 - 8) \times 2\text{ms} = 48\text{ms}$ 。

38. C

甲与乙建立 TCP 连接时发送的 SYN 段中的序号为 1000，则在数据传输阶段所用起始序号为 1001，在断开连接时，甲发送给乙的 FIN 段中的序号为 5001，在无任何重传的情况下，甲向乙已经发送的应用层数据的字节数为 $5001 - 1001 = 4000$ 。

39. B

TCP 连接释放的过程在 5.3.3 节中介绍。当客户机收到服务器发送的 FIN 段并向服务器发送 ACK 段时，客户机的 TCP 状态变为 TIME_WAIT，此时 TCP 连接还未释放，必须经过时间等待计时器设置的时间 2MSL（最长报文段寿命）后，客户机才进入 CLOSED（连接关闭）状态。

40. D

当应用层数据交给传输层时，放在报文段的数据部分。UDP 首部有 8B，TCP 首部最短有 20B。为了达到最大传输效率，通过 UDP 传输时，总长度为 20B，最大传输效率是 $12\text{B}/20\text{B} = 60\%$ 。通过 TCP 传输时，总长度为 32B，最大传输效率是 $12\text{B}/32\text{B} = 37.5\%$ 。

41. C

依题意，甲发送 200B 报文后，继续发送的报文段中序号字段 $\text{seq} = 701$ 。由于乙被告知接收窗口为 500，且甲未收到乙对 $\text{seq} = 501$ 报文段的确认，甲还能发送的报文段字节数为 $500 - 200 = 300\text{B}$ ，因此甲在未收到新的确认段之前，还能发送的数据序号范围是 701~1000。

42. C

时刻 0 发生了超时，门限值 ssthresh 变为拥塞窗口 cwnd 的一半即 8，同时 cwnd 置为 1，执行慢开始算法，cwnd 指数增长，经过 3 个 RTT，增长到 ssthresh 值；之后执行拥塞避免算法，cwnd 线性增长，再经过 8 个 RTT，增长到 16，共花费 11 个 RTT，如下表所示。

时刻	0	1	2	3	4	5	6	7	8	9	10	11
拥塞窗口	1	2	4	8	9	10	11	12	13	14	15	16

43. D

TCP 连接的释放过程如图 5.8 所示。题目问的是最少时间，所以当服务器 S 收到客户 C 发送的 FIN 请求后不再发送数据，而立马发送 FIN 请求（即第②步和第③步同时发生，忽略 FIN-WAIT-2 和 CLOSE-WAIT 状态）。C 收到 S 发来的 FIN 报文段后，进入 CLOSED 状态还需等到 TIME-WAIT 结束，总用时至少为 $1RTT + 2MSL = 50 + 800 \times 2 = 1650ms$ 。S 进入 CLOSED 状态需要经过 3 次报文段的传输时间，即 $1.5RTT = 75ms$ 。

二、综合应用题

01. 【解答】

这是因为发送方可能还未重传时，就收到了对更高序号的确认。例如主机 A 连续发送两个报文段（SEQ=92，DATA 共 8B）和（SEQ=100，DATA 共 20B），均正确到达主机 B。B 连续发送两个确认（ACK=100 和 ACK=120），但前一个确认帧在传送时丢失。例如 A 在第一个报文段（SEQ=92，DATA 共 8B）超时之前收到了对第二个报文段的确认（ACK=120），此时 A 知道，119 号和在 119 号之前的所有字节（包括第一个报文段中的所有字节）均已被 B 正确接收，因此 A 不会再重传第一个报文段。

02. 【解答】

第一种方法将失序报文段丢弃，会引起被丢弃报文段的重复传送，增加对网络带宽的消耗，但由于用不着将该报文段暂存，可避免对接收方缓冲区的占用。

第二种方法先将失序报文段暂存于接收缓存内，待所缺序号的报文段收齐后再一起上交应用层；这样有可能避免发送方对已被接收方收到的失序报文段的重传，减少对网络带宽的消耗，但增加了接收方缓冲区的开销。

03. 【解答】

TCP 为传送的数据流中的每个字节都编上一个序号。报文段的序号指的是本报文段所发送的数据的第一个字节的序号。因此第一个报文段的序号为 10010，第二个报文段的序号为 $10010 + 1000 = 11010$ ，第三个报文段的序号为 $11010 + 1000 = 12010$ 。

04. 【解答】

最大吞吐量表明在一个 RTT 内将窗口中的字节全部发送完毕。在平均往返时间 20ms 内，发送的最大数据量为最大窗口值，即 $64 \times 1024B$ ，

$$64 \times 1024 \times 8 / (20 \times 10^{-3}) \approx 26.2 \text{ Mb/s}$$

因此，所能得到的最大吞吐量是 26.2Mb/s。

05. 【解答】

具有相同编号的报文段不应同时在网络中传输，必须保证当序列号循环回来重复使用时，具有相同序列号的报文段已从网络中消失，类似于 GBN 原理 ($2^n - 1$)。现在序号用 8 位表示，报文段的寿命为 30s，那么在 30s 的时间内发送方发送的报文段的数目不能多于 255 个，

$$255 \times 128 \times 8 / 30 = 8704 \text{ b/s}$$

所以，每条 TCP 连接所能达到的最高数据率为 8704b/s。

关注公众号【乘龙考研】

一手更新 稳定有保障

06. 【解答】

由于收到接收方的确认至少需要一个 RTT，因此在一个 RTT 内，发送的数据量不能超过发送窗口大小，所以吞吐率 = 发送窗口大小/RTT。题目中告诉的是端到端时延， $RTT = 2 \times$ 端到端时延，因此 $RTT = 2 \times 20 = 40\text{ms}$ ，所以吞吐率 = $65535 \times (8/0.04) = 13.107\text{Mb/s}$ 。

线路效率 = 吞吐率/信道带宽。本题中，线路效率 $(13.107\text{Mb/s})/(1000\text{Mb/s}) = 1.31\%$ 。本题在计算时要特别注意单位（是 b 还是 B），要区分 Gb/s 和 GB/s。

07. 【解答】

- 1) 注意，TCP 传送的数据流中的每个字节都有一个编号，而 TCP 报文段的序号为其数据部分第一个字节的编号。因此第 1 个报文中的数据有 $120 - 90 = 30\text{B}$ ，第 2 个报文中的数据有 $150 - 120 = 30\text{B}$ 。
- 2) 由于 TCP 使用累积确认策略，因此当第 2 个报文段丢失后，第 3 个报文段就成了失序报文，B 期望收到的下一个报文段是序号为 120 的报文段，所以确认号为 120。

08. 【解答】

慢启动拥塞控制考虑了两个潜在的问题，即网络容量和接收方容量，并且分别处理每个问题。为此，每个发送方都维持两个窗口，即接收方准许的窗口和拥塞窗口。发送方可以发送的字节数是这两个窗口中的最小值。

建立一条连接时，发送方把拥塞窗口初始化为在该连接上使用的 1 个最大报文段尺寸。然后它发送一个最大报文段。如果这个报文段在超时之前得到确认，那么发送方就把拥塞窗口增加到 2 个最大报文段长，并发送两个报文段。发出去的每个报文段被确认后，拥塞窗口都要增加 1 个最大报文段。因此，当拥塞窗口是 n 个报文段时，如果所有 n 个报文段都及时得到确认，那么拥塞窗口将增加 n 个最大报文段，变成 $2n$ 个最大报文段。事实上，每一次突发性连续报文段都会使拥塞窗口加倍。

拥塞窗口继续按指数型增长，直到超时发生，或者到了接收方窗口的边界。其思想是如果突发量 1024B、2048B 和 4096B 工作得很好，但 8192B 的突发量引起超时，那么拥塞窗口应该设置成 4096B 以避免拥塞。只要拥塞窗口保持在 4096B，不管接收方准许什么样的窗口空间，都不会发送大于 4096B 的突发量。这种算法称为慢启动。现在所有的 TCP 实现都需要支持这个算法。

现在，最大的段长是 2KB，开始的突发量分别是 2KB, 4KB, 8KB 和 16KB，下面是 24KB，即第一个完全窗口。 $10\text{ms} \times 4 = 40\text{ms}$ ，因此需要 40ms 才能发送第一个完全窗口。

09. 【解答】

在慢启动和拥塞避免算法中，拥塞窗口初始为 1，窗口大小开始按指数增长。当拥塞窗口大于慢开始门限后停止使用慢启动算法，改用拥塞避免算法。此处慢开始的门限值初始为 12，当拥塞窗口增大到 12 时改用拥塞避免算法，窗口大小按线性增长，每次增加 1 个报文段，当增加到 16 时，出现超时，重新设门限值为 8（16 的一半），拥塞窗口再重新设为 1，执行慢启动算法，到门限值 8 时执行拥塞避免算法。

这样，拥塞窗口的变化就为 1, 2, 4, 8, 12, 13, 14, 15, 16, 1, 2, 4, 8, 9, 10, 11, 12, 13, 14, 15, 16, …。可见从出现超时时拥塞窗口为 16 到恢复拥塞窗口大小为 16，需要的往返时间次数是 11。注意，发现超时时，拥塞窗口从 16 变为 1 是立即进行的，不会间隔一个 RTT。

10. 【解答】

目标在 120s 内最多发送 2^{32}B （序列号为 32 位），即 35791394B/s 的载荷。TCP 报文段载荷是 1500B，因此可以发送 23861 个报文段。TCP 开销是 20B，IP 开销是 20B，以太网开销是 26B（18B 的首部和尾部，7B 的前同步码，1B 的帧开始定界符）。这就意味着对于 1500B 的载荷，必

须发送 1566B。 $1566 \times 8 \times 23861 = 299\text{Mb/s}$ ，因此允许的最快线路速率是 299Mb/s。比这一速度更快时，就会冒在同一时间内不同的 TCP 报文段具有相同序号的风险。

11. 【解答】

来回路程的时延 $128\text{ms} \times 2 = 256\text{ms}$ 。设窗口值为 X （注意：单位为字节）。

假定一次最大发送量等于窗口值，且发送时间等于 256ms，那么每发送一次都得停下来期待再次得到下一个窗口的确认，以得到新的发送许可。这样，发送时间等于停止等待应答的时间，结果测到的平均吞吐率就等于发送速率的一半，即 128kb/s，

$$8X/(128 \times 2 \times 1000) = 256 \times 0.001 \Rightarrow X = 256 \times 1000 \times 256 \times 0.001 / 8 = 256 \times 32 = 8192$$

所以，窗口值为 8192。

12. 【解答】

在 TCP 的拥塞控制算法中，除使用慢启动的接收窗口和拥塞窗口外，还使用第 3 个参数，即门槛值。发生超时的时候，该门槛值被设置成当前拥塞窗口值的一半即 9KB，而拥塞窗口则重置成一个最大报文段长。然后再使用慢启动的算法决定网络可以接受的进发量，一直增长到门槛值为止。从这一点开始，成功的传输线性地增加拥塞窗口，即每次进发传输后只增加一个最大报文段，而不是每个报文段传输后都增加一个最大报文段的窗口值。现在由于发生了超时，下一次传输将是 1 个最大报文段，然后是 2 个、4 个和 8 个最大报文段，第四次发送成功，且门限为 9KB，所以在 4 次进发量传输后，拥塞窗口将增加为 9KB。

13. 【解答】

- 1) 源端口号为第 1、2 个字节，即 0D 28，转换为十进制数为 3368。目的端口号为第 3、4 个字节，即 00 15，转换为十进制数为 21。
- 2) 第 5~8 个字节为序列号，即 50 5F A9 06。第 9~12 个字节为确认号，即 00 00 00 00，也即十进制数 0。
- 3) 第 13 个字节的前 4 位为 TCP 首部的长度，这里的值是 7（以 4B 为单位），因此乘以 4 后得到 TCP 首部的长度为 28B，说明该 TCP 首部还有 8B 的选项数据。
- 4) 根据目的端口是 21 可知这是一条 FTP 连接，而 TCP 的状态则需要分析第 14 个字节。第 14 个字节的值为 02，即 SYN 置为 1，而且 ACK=0 表示该数据段没有捎带的确认，这说明是第一次握手时发出的 TCP 连接。

14. 【解答】

- 1) 由图 1 看出，源 IP 地址为 IP 分组头的第 13~16 个字节。在表 1 中，1、3、4 号分组的源 IP 地址均为 192.168.0.8 (c0a80008H)，所以 1、3、4 号分组是由 H 发送的。
在表 1 中，1 号分组封装的 TCP 段的 SYN=1，ACK=0，seq=846b 41c5H；2 号分组封装的 TCP 段的 SYN=1，ACK=1，seq=e059 9fefH，ack=846b 41c6H；3 号分组封装的 TCP 段的 ACK=1，seq=846b 41c6H，ack=e059 9ff0H，所以 1、2、3 号分组完成了 TCP 连接的建立过程。
由于快速以太网数据帧有效载荷的最小长度为 46B，表 1 中 3、5 号分组的总长度为 40 (28H) 字节，小于 46B，其余分组总长度均大于 46B。所以 3、5 号分组通过快速以太网传输时需要填充。
- 2) 由 3 号分组封装的 TCP 段可知，发送应用层数据初始序号为 seq=846b 41c6H，由 5 号分组封装的 TCP 段可知，ack 为 seq=846b 41d6H，所以 S 已经收到的应用层数据的字节数为 846b 41d6H - 846b 41c6H = 10H = 16B。
- 3) 由于 S 发出的 IP 分组的标识=6811H，所以该分组所对应的是表 1 中的 5 号分组。S 发出

的 IP 分组的 TTL=40H=64，5 号分组的 TTL=31H=49， $64 - 49 = 15$ ，所以可以推断该 IP 分组到达 H 时经过了 15 个路由器。

15. 【解答】

- 1) 第二次握手 TCP 段的 SYN=1, ACK=1; 确认序号是 101。
- 2) H3 收到的第 8 个确认段所通告的接收窗口是 12KB; 此时 H3 的拥塞窗口变为 9KB; H3 的发送窗口变为 9KB。
- 3) H3 的发送窗口等于 0 时, 下一个待发送段的序号是 $20K + 101 = 20 \times 1024 + 101 = 20581$; H3 从发送第 1 个段到发送窗口等于 0 时刻为止, 平均数据传输速率是 $20KB/(5 \times 200ms) = 20KB/s = 20.48k \times 8b/s = 163.84kb/s$ 。

注意: K 表示文件大小或描述存储空间时等于 1024, 这里通常用大写的 K; k 表示传输速率或描述网络通信时等于 1000, 这里通常用小写的 k。注意区分和转换。

- 4) 从 t 时刻起, S 释放该连接的最短时间是 $1.5 \times 200ms = 300ms$ 。

5.4 本章小结及疑难点



1. MSS 设置得太大或太小会有什么影响?

规定最大报文段 MSS 的大小并不是考虑到接收方的缓存可能放不下 TCP 报文段。实际上, MSS 与接收窗口没有关系。TCP 的报文段的数据部分, 至少要加上 40B 的首部 (TCP 首部至少 20B 和 IP 首部至少 20B), 才能组装成一个 IP 数据报。若选择较小的 MSS 值, 网络的利用率就很低。设想在极端情况下, 当 TCP 报文段中只含有 1B 的数据时, 在 IP 层传输的数据报的开销至少有 40B。这样, 网络的利用率就不会超过 $1/41$ 。到了数据链路层还要加上一些开销, 网络的利用率进一步降低。但反过来, 若 TCP 报文段很长, 那么在 IP 层传输时有可能要分解成多个短数据报片, 在终端还要把收到的各数据报片装配成原来的 TCP 报文段。传输有差错时, 还要进行重传。这些都会使开销增大。

因此, MSS 应尽量大一些, 只要在 IP 层传输时不要再分片就行。由于 IP 数据报所经历的路径是动态变化的, 在一条路径上确定的不需要分片的 MSS, 如果改走另一条路径, 就可能需要进行分片。因此, 最佳的 MSS 是很难确定的。MSS 的默认值为 536B, 因此在因特网上的所有主机都能接收的报文段长度是 $536 + 20 (\times \text{TCP 固定首部长度}) = 556B$ 。

2. 为何不采用“三次握手”释放连接, 且发送最后一次握手报文后要等待 2MSL 的时间呢? 原因有两个:

- 1) 保证 A 发送的最后一个确认报文段能够到达 B。如果 A 不等待 2MSL, 若 A 返回的最后确认报文段丢失, 则 B 不能进入正常关闭状态, 而 A 此时已经关闭, 也不可能再重传。
- 2) 防止出现“已失效的连接请求报文段”。A 在发送最后一个确认报文段后, 再经过 2MSL 可保证本连接持续的时间内所产生的所有报文段从网络中消失。造成错误的情形与下文 (疑难点 6) 不采用“两次握手”建立连接所述的情形相同。

注意: 服务器结束 TCP 连接的时间要比客户机早一些, 因为客户机最后要等待 2MSL 后才可进入 CLOSED 状态。

3. 如何判定此确认报文段是对原来的报文段的确认, 还是对重传的报文段的确认?

由于对于一个重传报文的确认来说, 很难分辨它是原报文的确认还是重传报文的确认, 使用

修正的 Karn 算法作为规则：在计算平均往返时间 RTT 时，只要报文段重传了，就不采用其往返时间样本，且报文段每重传一次，就把 RTO 增大一些。

4. TCP 使用的是 GBN 还是选择重传？

这是一个有必要弄清的问题。前面讲过，TCP 使用累积确认，这看起来像是 GBN 的风格。但是，正确收到但失序的报文并不会丢弃，而是缓存起来，并且发送冗余 ACK 指明期望收到的下一个报文段，这是 TCP 方式和 GBN 的显著区别。例如，A 发送了 N 个报文段，其中第 k ($k < N$) 个报文段丢失，其余 $N-1$ 个报文段正确地按序到达接收方 B。使用 GBN 时，A 需要重传分组 k ，及所有后继分组 $k+1, k+2, \dots, N$ 。相反，TCP 却至多重传一个报文段，即报文段 k 。另外，TCP 中提供一个 SACK（Selective ACK）选项，即选择确认选项。使用选择确认选项时，TCP 看起来就和 SR 非常相似。因此，TCP 的差错恢复机制可视为 GBN 和 SR 协议的混合体。

5. 为什么超时事件发生时 cwnd 被置为 1，而收到 3 个冗余 ACK 时 cwnd 减半？

大家可以从如下角度考虑。超时事件发生和收到 3 个冗余 ACK，哪个意味着网络拥塞程度更严重？通过分析不难发现，在收到 3 个冗余 ACK 的情况下，网络虽然拥塞，但至少还有 ACK 报文段能被正确交付。而当超时发生时，说明网络可能已经拥塞得连 ACK 报文段都传输不了，发送方只能等待超时后重传数据。因此，超时事件发生时，网络拥塞更严重，那么发送方就应该最大限度地抑制数据发送量，所以 cwnd 置为 1；收到 3 个冗余 ACK 时，网络拥塞不是很严重，发送方稍微抑制一下发送的数据量即可，所以 cwnd 减半。

6. 为什么不采用“两次握手”建立连接呢？

这主要是为了防止两次握手情况下已失效的连接请求报文段突然又传送到服务器而产生错误。考虑下面这种情况。客户 A 向服务器 B 发出 TCP 连接请求，第一个连接请求报文在网络的某个结点长时间滞留，A 超时后认为报文丢失，于是再重传一次连接请求，B 收到后建立连接。数据传输完毕后双方断开连接。而此时，前一个滞留在网络中的连接请求到达服务器 B，而 B 认为 A 又发来连接请求，此时若使用“三次握手”，则 B 向 A 返回确认报文段，由于是一个失效的请求，因此 A 不予理睬，建立连接失败。若采用的是“两次握手”，则这种情况下 B 认为传输连接已经建立，并一直等待 A 传输数据，而 A 此时并无连接请求，因此不予理睬，这样就造成了 B 的资源白白浪费。

7. 是否 TCP 和 UDP 都需要计算往返时间 RTT？

往返时间 RTT 仅对传输层 TCP 协议才很重要，因为 TCP 要根据 RTT 的值来设置超时计时器的超时时间。UDP 没有确认和重传机制，因此 RTT 对 UDP 没有什么意义。

因此，不能笼统地说“往返时间 RTT 对传输层来说很重要”，因为只有 TCP 才需要计算 RTT，而 UDP 不需要计算 RTT。

8. 为什么 TCP 在建立连接时不能每次都选择相同的、固定的初始序号？

- 1) 假定主机 A 和 B 频繁地建立连接，传送一些 TCP 报文段后，再释放连接，然后又不断地建立新的连接、传送报文段和释放连接。
- 2) 假定每次建立连接时，主机 A 都选择相同的、固定的初始序号，如选择 1。
- 3) 假定主机 A 发出的某些 TCP 报文段在网络中会滞留较长时间，以致主机 A 超时重传这些 TCP 报文段。
- 4) 假定有一些在网络中滞留时间较长的 TCP 报文段最后终于到达主机 B，但这时传送该报文段的那个连接早已释放，而在到达主机 B 时的 TCP 连接是一条新的 TCP 连接。

这样，工作在新的 TCP 连接的主机 B 就有可能会接收在旧的连接传送的、已无意义的、过时的 TCP 报文段(因为这个 TCP 报文段的序号有可能正好处在当前新连接所用的序号范围之中)，结果产生错误。因此，必须使得迟到的 TCP 报文段的序号不处在新连接所用的序号范围之中。

这样，TCP 在建立新的连接时所选择的初始序号一定要和前面的一些连接所用过的序号不同。因此，不同的 TCP 连接不能使用相同的初始序号。

9. 假定在一个互联网中，所有链路的传输都不出现差错，所有结点也都不会发生故障。试问在这种情况下，TCP 的“可靠交付”的功能是否就是多余的？

不是多余的。TCP 的“可靠交付”功能在互联网中起着至关重要的作用。至少在以下的情况下，TCP 的“可靠交付”功能是必不可少的。

- 1) 每个 IP 数据报独立地选择路由，因此在到达目的主机时有可能出现失序。
- 2) 由于路由选择的计算出现错误，导致 IP 数据报在互联网中转圈。最后数据报首部中的生存时间（TTL）的数值下降到零。这个数据报在中途就被丢失。
- 3) 某个路由器突然出现很大的通信量，以致路由器来不及处理到达的数据报。因此有的数据报被丢弃。

以上列举的问题表明：必须依靠 TCP 的“可靠交付”功能才能保证在目的主机的目的进程中接收到正确的报文。

第 6 章 应用层

关注公众号【乘龙考研】
一手更新 稳定有保障

【考纲内容】

(一) 网络应用模型

客户/服务器模型；P2P 模型

(二) 域名系统 (DNS)

层次域名空间；域名服务器；域名解析过程

(三) 文件传输协议 (FTP)

FTP 的工作原理；控制连接与数据连接

(四) 电子邮件 (E-mail)

电子邮件系统的组成结构；电子邮件格式与 MIME；SMTP 与 POP3

(五) 万维网 (WWW)

WWW 的概念与组成结构；HTTP

【复习提示】

本章内容既可以以选择题的形式考查，也可以结合其他章节的内容出综合题。所以牢固掌握本章的几个典型应用层协议是关键。我们生活中的很多网络应用都是建立在这些协议的基础上的，因此在学习时要注意联系实际，提高学习的兴趣，才会获得更好的效果。

6.1 网络应用模型

6.1.1 客户/服务器模型

在客户/服务器 (Client/Server, C/S) 模型中，有一个总是打开的主机称为服务器，它服务于许多来自其他称为客户机的主机请求。其工作流程如下：

- 1) 服务器处于接收请求的状态。
- 2) 客户机发出服务请求，并等待接收结果。
- 3) 服务器收到请求后，分析请求，进行必要的处理，得到结果并发送给客户机。

客户程序必须知道服务器程序的地址，客户机上一般不需要特殊的硬件和复杂的操作系统。而服务器上运行的软件则是专门用来提供某种服务的程序，可同时处理多个远程或本地客户的请求。系统启动后即自动调用并一直不断地运行着，被动地等待并接收来自各地客户的请求。因此，服务器程序不需要知道客户程序的地址。

客户/服务器模型最主要的特征是：客户是服务请求方，服务器是服务提供方。如 Web 应用程序，其中总是打开的 Web 服务器服务于运行在客户机上的浏览器的请求。当 Web 服务器接收到来自客户机对某对象的请求时，它向该客户机发送所请求的对象以做出响应。常见的使用客户/服务

扫一扫



视频讲解

器模型的应用包括 Web、文件传输协议（FTP）、远程登录和电子邮件等。

客户/服务器模型的主要特点还有：

- 1) 网络中各计算机的地位不平等，服务器可以通过对用户权限的限制来达到管理客户机的目的，使它们不能随意存储/删除数据，或进行其他受限的网络活动。整个网络的管理工作由少数服务器担当，因此网络的管理非常集中和方便。
- 2) 客户机相互之间不直接通信。例如，在 Web 应用中两个浏览器并不直接通信。
- 3) 可扩展性不佳。受服务器硬件和网络带宽的限制，服务器支持的客户机数有限。

6.1.2 P2P 模型

不难看出，在 C/S 模型中（见图 6.1），服务器性能的好坏决定了整个系统的性能，当大量用户请求服务时，服务器就必然成为系统的瓶颈。P2P 模型（见图 6.2）的思想是整个网络中的传输内容不再被保存在中心服务器上，每个结点都同时具有下载、上传的功能，其权利和义务都是大体对等的。

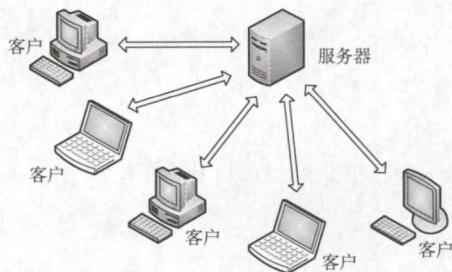


图 6.1 C/S 模型

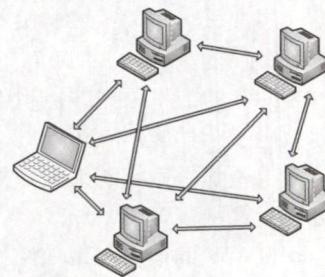


图 6.2 P2P 模型

在 P2P 模型中，各计算机没有固定的客户和服务器划分。相反，任意一对计算机——称为对等方（Peer），直接相互通信。实际上，P2P 模型从本质上来看仍然使用客户/服务器模式，每个结点既作为客户访问其他结点的资源，也作为服务器提供资源给其他结点访问。当前比较流行的 P2P 应用有 PPStream、BitTorrent 和电驴等。

与 C/S 模型相比，P2P 模型的优点主要体现如下：

- 1) 减轻了服务器的计算压力，消除了对某个服务器的完全依赖，可以将任务分配到各个结点上，因此大大提高了系统效率和资源利用率（例如，播放流媒体时对服务器的压力过大，而通过 P2P 模型，可以利用大量的客户机来提供服务）。
- 2) 多个客户机之间可以直接共享文档。
- 3) 可扩展性好，传统服务器有响应和带宽的限制，因此只能接受一定数量的请求。
- 4) 网络健壮性强，单个结点的失效不会影响其他部分的结点。

P2P 模型也有缺点。在获取服务的同时，还要给其他结点提供服务，因此会占用较多的内存，影响整机速度。例如，经常进行 P2P 下载还会对硬盘造成较大的损伤。据某互联网调研机构统计，当前 P2P 程序已占互联网 50%~90% 的流量，使网络变得非常拥塞，因此各大 ISP（互联网服务提供商，如电信、网通等）通常都对 P2P 应用持反对态度。

6.1.3 本节习题精选

单项选择题

01. 服务程序在 Windows 环境下工作，并且运行该服务器程序的计算机也作为客户访问其他计算机上提供的服务。那么，这种网络应用模型属于（ ）。

关注公众号【乘龙考研】

一手更新 稳定有保障

- A. 主从式 B. 对等式
 C. 客户/服务器模型 D. 集中式
02. 在客户/服务器模型中，客户指的是（ ）。
 A. 请求方 B. 响应方 C. 硬件 D. 软件
03. 用户提出服务请求，网络将用户请求传送到服务器；服务器执行用户请求，完成所要求的操作并将结果送回用户，这种工作模式称为（ ）。
 A. C/S 模型 B. P2P 模型
 C. CSMA/CD 模式 D. 令牌环模式
04. 下面关于客户/服务器模型的描述，（ ）存在错误。
 I. 客户端必须提前知道服务器的地址，而服务器则不需要提前知道客户端的地址
 II. 客户端主要实现如何显示信息与收集用户的输入，而服务器主要实现数据的处理
 III. 浏览器显示的内容来自服务器
 IV. 客户端是请求方，即使连接建立后，服务器也不能主动发送数据
 A. I、IV B. III、IV C. 只有 IV D. 只有 III
05. 下列关于客户/服务器模型的说法中，不正确的是（ ）。
 A. 服务器专用于完成某些服务，而客户机则作为这些服务的使用者
 B. 客户机通常位于前端，服务器通常位于后端
 C. 客户机和服务器通过网络实现协同计算任务
 D. 客户机是面向任务的，服务器是面向用户的
06. 以下关于 P2P 概念的描述中，错误的是（ ）。
 A. P2P 是网络结点之间采取对等方式直接交换信息的工作模式
 B. P2P 通信模式是指 P2P 网络中对等结点之间的直接通信能力
 C. P2P 网络是指与互联网并行建设的、由对等结点组成的物理网络
 D. P2P 实现技术是指为实现对等结点之间直接通信的功能所需要设计的协议、软件等
07. 【2019 统考真题】下列关于网络应用模型的叙述中，错误的是（ ）。
 A. 在 P2P 模型中，结点之间具有对等关系
 B. 在客户/服务器（C/S）模型中，客户与客户之间可以直接通信
 C. 在 C/S 模型中，主动发起通信的是客户，被动通信的是服务器
 D. 在向多用户分发一个文件时，P2P 模型通常比 C/S 模型所需的时间短

6.1.4 答案与解析

单项选择题

01. B

在 P2P 模型中，各用户计算机共享资源，从而提供比单个用户所能提供的多得多的资源。这里，各个计算机没有固定的客户和服务器划分，任意一对计算机称为对等方。

02. A

客户机既不是硬件又不是软件，只是服务的请求方，服务器才是响应方。

03. A

用户提出服务请求，网络将用户请求传送到服务器；服务器执行用户请求，完成所要求的操作并将结果送回用户，这种工作模型称为客户/服务器模型。

04. C



在连接未建立前，服务器在某一个端口上监听。客户端是连接的请求方，客户端必须事先知道服务器的地址才能发出连接请求，而服务器则从客户端发来的数据包中获取客户端的地址。一旦连接建立，服务器就能响应客户端请求的内容，服务器也能主动发送数据给客户端，用于一些消息的通知，如一些错误的通知。所以只有 IV 错误。

05. D

客户机的作用是根据用户需求向服务器发出服务请求，并将服务器返回的结果呈现给用户，因此客户机是面向用户的，服务器是面向任务的。

06. C

选项 C 中“P2P 网络是一种物理网络”的描述是错误的。P2P 网络是指在互联网中由对等结点组成的一种覆盖网络（Overlay Network），是一种动态的逻辑网络。另外，对等结点之间具有直接通信的能力是 P2P 的显著特点。

07. B

在 P2P 模型中，每个结点的权利和义务是对等的。在 C/S 模型中，客户是服务发起方，服务器被动接受各地客户的请求，但客户之间不能直接通信，例如 Web 应用中两个浏览器之间并不直接通信。P2P 模型减轻了对某个服务器的计算压力，可以将任务分配到各个结点上，极大提高了系统效率和资源利用率。

6.2 域名系统（DNS）

域名系统（Domain Name System, DNS）是因特网使用的命名系统，用来把便于人们记忆的具有特定含义的主机名（如 www.cskaoyan.com）转换为便于机器处理的 IP 地址。相对于 IP 地址，人们更喜欢使用具有特定含义的字符串来标识因特网上的计算机。值得注意的是，DNS 系统采用客户/服务器模型，其协议运行在 UDP 之上，使用 53 号端口。

从概念上可将 DNS 分为 3 部分：层次域名空间、域名服务器和解析器。

6.2.1 层次域名空间

因特网采用层次树状结构的命名方法。采用这种命名方法，任何一个连接到因特网的主机或路由器，都有一个唯一的层次结构名称，即域名（Domain Name）。域（Domain）是名字空间中一个可被管理的划分。域还可以划分为子域，而子域还可以继续划分为子域的子域，这样就形成了

了顶级域、二级域、三级域等。每个域名都由标号序列组成，而各标号之间用点（“.”）隔开。一个典型的例子如图 6.3 所示，它是王道论坛用于提供 WWW 服务的计算机（Web 服务器）的域名，它由三个标号组成，其中标号 com 是顶级域名，标号 cskaoyan 是二级域名，标号 www 是三级域名。



图 6.3 一个域名的例子

关于域名中的标号有以下几点需要注意：

- 1) 标号中的英文不区分大小写。
- 2) 标号中除连字符（-）外不能使用其他的标点符号。
- 3) 每个标号不超过 63 个字符，多标号组成的完整域名最长不超过 255 个字符。
- 4) 级别最低的域名写在最左边，级别最高的顶级域名写在最右边。

顶级域名（Top Level Domain, TLD）分为如下三大类：

- 1) 国家(地区)顶级域名(nTLD)。国家和某些地区的域名,如“.cn”表示中国,“.us”表示美国,“.uk”表示英国。
- 2) 通用顶级域名(gTLD)。常见的有“.com”(公司)、“.net”(网络服务机构)、“.org”(非营利性组织)和“.gov”(国家或政府部门)等。
- 3) 基础结构域名。这种顶级域名只有一个,即arpa,用于反向域名解析,因此又称反向域名。

国家(地区)顶级域名下注册的二级域名均由该国家(地区)自行确定。图6.4展示了域名空间的树状结构。

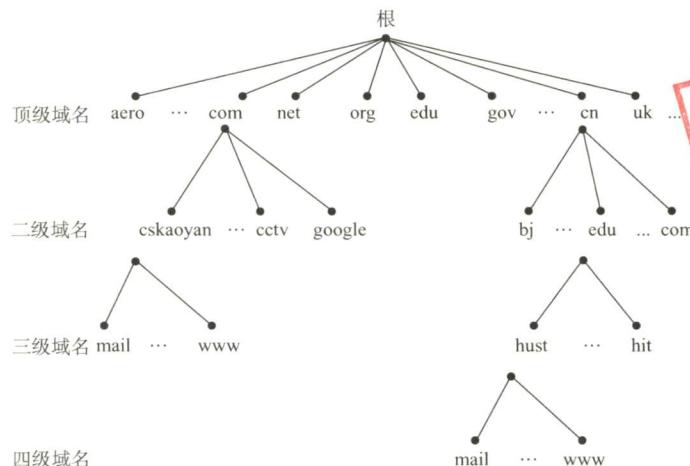


图6.4 域名空间的树状结构

在域名系统中,每个域分别由不同的组织进行管理。每个组织都可以将它的域再分成一定数目的子域,并将这些子域委托给其他组织去管理。例如,管理cn域的中国将edu.cn子域授权给中国教育和科研计算机网(CERNET)来管理。

6.2.2 域名服务器

因特网的域名系统被设计成一个联机分布式的数据库系统,并采用客户/服务器模型。域名到IP地址的解析是由运行在域名服务器上的程序完成的,一个服务器所负责管辖的(或有权限的)范围称为区(不以“域”为单位),各单位根据具体情况来划分自己管辖范围的区,但在一个区中的所有结点必须是能够连通的,每个区设置相应的权限域名服务器,用来保存该区中的所有主机的域名到IP地址的映射。每个域名服务器不但能够进行一些域名到IP地址的解析,而且还必须具有连向其他域名服务器的信息。当自己不能进行域名到IP地址的转换时,能够知道到什么地方去找其他域名服务器。

DNS使用了大量的域名服务器,它们以层次方式组织。没有一台域名服务器具有因特网上所有主机的映射,相反,该映射分布在所有的DNS上。采用分布式设计的DNS,是一个在因特网上实现分布式数据库的精彩范例。主要有4种类型的域名服务器。

1. 根域名服务器

根域名服务器是最高层次的域名服务器,所有的根域名服务器都知道所有的顶级域名服务器的IP地址。根域名服务器也是最重要的域名服务器,不管是哪个本地域名服务器,若要对因特网

上任何一个域名进行解析，只要自己无法解析，就首先要求助于根域名服务器。因特网上有 13 个根域名服务器，尽管我们将这 13 个根域名服务器中的每个都视为单个服务器，但每个“服务器”实际上是冗余服务器的集群，以提供安全性和可靠性。需要注意的是，根域名服务器用来管辖顶级域（如.com），通常它并不直接把待查询的域名直接转换成 IP 地址，而是告诉本地域名服务器下一步应当找哪个顶级域名服务器进行查询。

2. 顶级域名服务器

这些域名服务器负责管理在该顶级域名服务器注册的所有二级域名。收到 DNS 查询请求时，就给出相应的回答（可能是最后的结果，也可能是下一步应当查找的域名服务器的 IP 地址）。

3. 授权域名服务器（权限域名服务器）

每台主机都必须在授权域名服务器处登记。为了更加可靠地工作，一台主机最好至少有两个授权域名服务器。实际上，许多域名服务器都同时充当本地域名服务器和授权域名服务器。授权域名服务器总能将其管辖的主机名转换为该主机的 IP 地址。

4. 本地域名服务器

本地域名服务器对域名系统非常重要。每个因特网服务提供者（ISP），或一所大学，甚至一所大学中的各个系，都可以拥有一个本地域名服务器。当一台主机发出 DNS 查询请求时，这个查询请求报文就发送给该主机的本地域名服务器。事实上，我们在 Windows 系统中配置“本地连接”时，就需要填写 DNS 地址，这个地址就是本地 DNS（域名服务器）的地址。

DNS 的层次结构如图 6.5 所示。

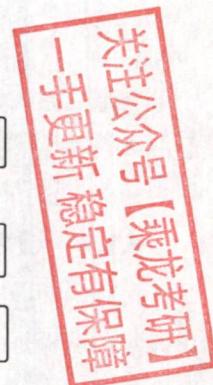
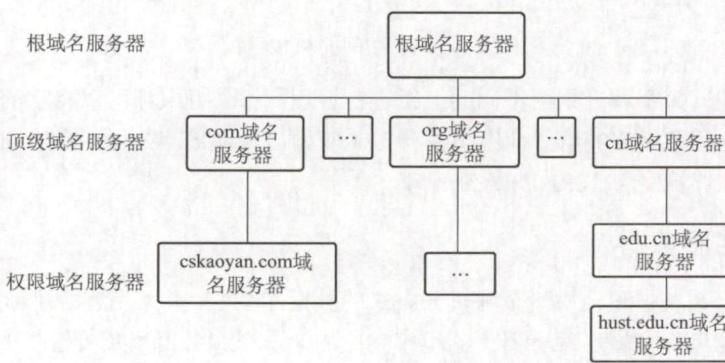


图 6.5 DNS 的层次结构

6.2.3 域名解析过程

域名解析是指把域名映射成为 IP 地址或把 IP 地址映射成域名的过程。前者称为正向解析，后者称为反向解析。当客户端需要域名解析时，通过本机的 DNS 客户端构造一个 DNS 请求报文，以 UDP 数据报方式发往本地域名服务器。

域名解析有两种方式：递归查询和递归与迭代相结合的查询。

递归查询的过程如图 6.6(a)所示，本地域名服务器只需向根域名服务器查询一次，后面的几次查询都是递归地在其他几个域名服务器之间进行的〔步骤③～⑥〕。在步骤⑦中，本地域名服务器从根域名服务器得到了所需的 IP 地址，最后在步骤⑧中，本地域名服务器把查询结果告诉发起查询的主机。由于该方法给根域名服务造成的负载过大，所以在实际中几乎不使用。

常用递归与迭代相结合的查询方式如图 6.6(b)所示，该方式分为两个部分。



图 6.6 两种域名解析方式工作原理

(1) 主机向本地域名服务器的查询采用的是递归查询

也就是说，如果本地主机所询问的本地域名服务器不知道被查询域名的 IP 地址，那么本地域名服务器就以 DNS 客户的身份，向根域名服务器继续发出查询请求报文（即替该主机继续查询），而不是让该主机自己进行下一步的查询。两种查询方式的这一步是相同的。

(2) 本地域名服务器向根域名服务器的查询采用迭代查询

当根域名服务器收到本地域名服务器发出的迭代查询请求报文时，要么给出所要查询的 IP 地址，要么告诉本地域名服务器：“你下一步应当向哪个顶级域名服务器进行查询”。然后让本地域名服务器向这个顶级域名服务器进行后续的查询，如图 6.6(b)所示。同样，顶级域名服务器收到查询报文后，要么给出所要查询的 IP 地址，要么告诉本地域名服务器下一步应向哪个权限域名服务器查询。最后，知道所要解析的域名的 IP 地址后，把这个结果返回给发起查询的主机。

下面举例说明域名解析的过程。假定某客户机想获知域名为 y.abc.com 主机的 IP 地址，域名解析的过程（共使用了 8 个 UDP 报文）如下：

- ① 客户机向其本地域名服务器发出 DNS 请求报文（递归查询）。
- ② 本地域名服务器收到请求后，查询本地缓存，若没有该记录，则以 DNS 客户的身份向根域名服务器发出解析请求报文（迭代查询）。
- ③ 根域名服务器收到请求后，判断该域名属于.com 域，将对应的顶级域名服务器 dns.com 的 IP 地址返回给本地域名服务器。
- ④ 本地域名服务器向顶级域名服务器 dns.com 发出解析请求报文（迭代查询）。
- ⑤ 顶级域名服务器 dns.com 收到请求后，判断该域名属于 abc.com 域，因此将对应的授权域名服务器 dns.abc.com 的 IP 地址返回给本地域名服务器。
- ⑥ 本地域名服务器向授权域名服务器 dns.abc.com 发起解析请求报文（迭代查询）。
- ⑦ 授权域名服务器 dns.abc.com 收到请求后，将查询结果返回给本地域名服务器。
- ⑧ 本地域名服务器将查询结果保存到本地缓存，同时返回给客户机。

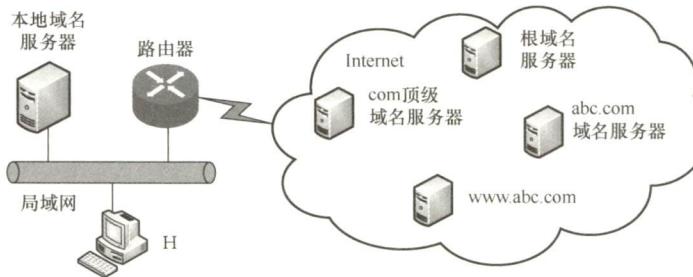
为了提高 DNS 的查询效率，并减少因特网上的 DNS 查询报文数量，在域名服务器中广泛地使用了高速缓存。当一个 DNS 服务器接收到 DNS 查询结果时，它能将该 DNS 信息缓存在高速缓存中。这样，当另一个相同的域名查询到达该 DNS 服务器时，该服务器就能够直接提供所要求的 IP 地址，而不需要再去向其他 DNS 服务器询问。因为主机名和 IP 地址之间的映射不是永久的，所以 DNS 服务器将在一段时间后丢弃高速缓存中的信息。

6.2.4 本节习题精选

一、单项选择题

01. 域名与（ ）具有一一对应的关系。
A. IP 地址 B. MAC 地址 C. 主机 D. 以上都不是
02. 下列说法错误的是（ ）。
A. Internet 上提供客户访问的主机一定要有域名
B. 同一域名在不同时间可能解析出不同的 IP 地址
C. 多个域名可以指向同一台主机 IP 地址
D. IP 子网中的主机可以由不同的域名服务器来维护其映射
03. DNS 是基于（ ）模型的分布式系统。
A. C/S B. B/S C. P2P D. 以上均不正确
04. 域名系统（DNS）的组成不包括（ ）。
A. 域名空间 B. 分布式数据库
C. 域名服务器 D. 从内部 IP 地址到外部 IP 地址的翻译程序
05. 互联网中域名解析依赖于由域名服务器组成的逻辑树。在域名解析过程中，主机上请求域名解析的软件不需要知道（ ）信息。
I. 本地域名服务器的 IP
II. 本地域名服务器父结点的 IP
III. 域名服务器树根结点的 IP
A. I 和 II B. I 和 III C. II 和 III D. I、II 和 III
06. 在 DNS 的递归查询中，由（ ）给客户端返回地址。
A. 最开始连接的服务器 B. 最后连接的服务器
C. 目的地所在服务器 D. 不确定
07. 一台主机要解析 www.cskaoyan.com 的 IP 地址，如果这台主机配置的域名服务器为 202.120.66.68，因特网顶级域名服务器为 11.2.8.6，而存储 www.cskaoyan.com 的 IP 地址对应关系的域名服务器为 202.113.16.10，那么这台主机解析该域名通常首先查询（ ）。
A. 202.120.66.68 域名服务器
B. 11.2.8.6 域名服务器
C. 202.113.16.10 域名服务器
D. 可以从这 3 个域名服务器中任选一个
08. （ ）可以将其管辖的主机名转换为主机的 IP 地址。
A. 本地域名服务器 B. 根域名服务器
C. 授权域名服务器 D. 代理域名服务器
09. 【2010 统考真题】若本地域名服务器无缓存，则在采用递归方法解析另一网络某主机域名时，用户主机和本地域名服务器发送的域名请求条数分别为（ ）。
A. 1 条，1 条 B. 1 条，多条 C. 多条，1 条 D. 多条，多条
10. 【2016 统考真题】假设所有域名服务器均采用迭代查询方式进行域名解析。当主机访问规范域名为 www.abc.xyz.com 的网站时，域名服务器在完成该域名解析的过程中，可能发出 DNS 查询的最少和最多次数分别是（ ）。
A. 0, 3 B. 1, 3 C. 0, 4 D. 1, 4

11. 【2018 统考真题】下列 TCP/IP 应用层协议中，可以使用传输层无连接服务的是（ ）。
- A. FTP B. DNS C. SMTP D. HTTP
12. 【2020 统考真题】假设下图所示网络中的本地域名服务器只提供递归查询服务，其他域名服务器均只提供迭代查询服务；局域网内主机访问 Internet 上各服务器的往返时间 (RTT) 均为 10ms，忽略其他各种时延。若主机 H 通过超链接 <http://www.abc.com/index.html> 请求浏览纯文本 Web 页 index.html，则从单击超链接开始到浏览器接收到 index.html 页面为止，所需的最短时间与最长时间分别是（ ）。



- A. 10ms, 40ms B. 10ms, 50ms C. 20ms, 40ms D. 20ms, 50ms

二、综合应用题

01. 一台具有单个 DNS 名称的机器可以有多个 IP 地址吗？为什么？
02. 一台计算机可以有两个属于不同顶级域的 DNS 名字吗？如果可以，试举例说明。
03. DNS 使用 UDP 而非 TCP，如果一个 DNS 分组丢失，没有自动恢复，那么这会引起问题吗？如果会，应该如何解决？
04. 为什么要引入域名的概念？举例说明域名转换过程。域名服务器中的高速缓存有何作用？

6.2.5 答案与解析

一、单项选择题

01. D

如果一台主机通过两块网卡连接到两个网络（如服务器双线接入），那么就具有两个 IP 地址，每个网卡对应一个 MAC 地址，显然这两个 IP 地址可以映射到同一个域名上。此外，多台主机也可以映射到同一个域名上（如负载均衡），一台主机也可以映射到多个域名上（如虚拟主机）。因此，选项 A、B、C 和域名均不具有一一对应的关系。

02. A

Internet 上提供访问的主机一定要有 IP 地址，而不一定要有域名，选项 A 错。域名在不同的时间可以解析出不同的 IP 地址，因此可以用多台服务器来分担负载，选项 B 对。也可以把多个域名指向同一台主机 IP 地址，选项 C 对。IP 子网中主机也可以由不同的域名服务器来维护其映射，选项 D 对。

03. A

域名系统（DNS）是一个基于客户/服务器模型的分布式数据库系统，主要作用是进行域名和 IP 地址之间的相互映射。

04. D

DNS 提供从域名到 IP 地址或从 IP 地址到域名的映射服务。它被设计成为一个联机分布式数据库系统，并采用客户/服务器方式。域名的解析是由若干域名服务器程序完成的。从内部 IP 地

关注公众号【乘龙考研】
一手更新 稳定有保障

址到外部 IP 地址的映射是由 NAT 实现的，用于缓解 IPv4 地址紧缺的问题，与域名系统无关。

05. C

正常情况下，客户机只需把域名解析请求发往本地域名服务器，其他事情都由本地域名服务器完成，并把最后结果返回给客户机。所以主机只需要知道本地域名服务器的 IP。

06. A

在递归查询中，每台不包含被请求信息的服务器都转到其他地方去查找，然后它再往回发送结果，所以客户端最开始连接的服务器最终将返回正确的信息。

07. A

当这台主机发出对 www.cskaoyan.com 的 DNS 查询报文时，这个查询报文首先被送往该主机的本地域名服务器 202.120.66.68。本地域名服务器不能立即回答该查询时，就以 DNS 客户的身份向某一根域名服务器查询。但不管采用何种查询方式，首先都要查询本地域名服务器。

08. C

每台主机都必须在授权域名服务器处注册登记，授权域名服务器一定能够将其管辖的主机名转换为该主机的 IP 地址。

09. A

采用递归查询时，如果主机所询问的本地域名服务器不知道被查询域名的 IP 地址，那么本地域名服务器就以 DNS 客户的身份，向根域名服务器继续发出查询请求报文，而不是让该主机自己进行下一步的查询。因此，采用这种方法时，用户主机和本地域名服务器发送的域名请求条数均为 1。因此答案为选项 A。

10. C

最少情况：当本地域名服务器中有该域名的 DNS 信息时，不需要查询任何其他域名服务器，最少发出 0 次 DNS 查询。最多情况：因为均采用迭代查询方式，在最坏情况下，本地域名服务器需要依次迭代地向根域名服务器、顶级域名服务器 (.com)、权限域名服务器 (xyz.com)、权限域名服务器 (abc.xyz.com) 发出 DNS 查询请求，因此最多发出 4 次 DNS 查询。

11. B

FTP 用来传输文件，SMTP 用来发送电子邮件，HTTP 用来传输网页文件，都对可靠性的要求较高，因此都用传输层有连接的 TCP 服务。无连接的 UDP 服务效率更高、开销小，DNS 在传输层采用无连接的 UDP 服务。

12. D

题中 RTT 均为局域网内主机（主机 H、本地域名服务器）访问 Internet 上各服务器的往返时间，且忽略其他时延，因此主机 H 向本地域名服务器的查询时延忽略不计。最短时间：本地主机中有该域名到 IP 地址对应的记录，因此不需要 DNS 查询时延，直接和 www.abc.com 服务器建立 TCP 连接再进行资源访问，TCP 连接建立需要 1 个 RTT，接着发送访问请求并收到服务器资源响应需要 1 个 RTT，共计 2 个 RTT，即 20ms；最长时间：本地主机递归查询本地域名服务器（延时忽略），本地服务器依次迭代查询根域名服务器、.com 顶级域名服务器、abc.com 域名服务器，共 3 个 RTT，查询到 IP 地址后，将该映射返回给主机 H，主机 H 和 www.abc.com 服务器建立 TCP 连接再进行资源访问，共 2 个 RTT，因此最长时间需要 $3 + 2 = 5$ 个 RTT，即 50ms。

二、综合应用题

01. 【解答】

可以，IP 地址由网络号和主机号两部分构成。如果一台机器有两个以太网卡，它就可以同时连到两个不同的网络上（网络号不能相同，否则发生冲突）；如果是这样，它就需要两个 IP 地址。

02. 【解答】

可以，例如 www.cskaoyan.com 和 www.cskaoyan.cn 属于不同的顶级域 (.com 和.cn)，但它们可以有同样的 IP 地址。用户输入这两个不同的 DNS 名字，访问的都是同一台服务器。

03. 【解答】

DNS 使用传输层的 UDP 而非 TCP，因为它不需要使用 TCP 在发生传输错误时执行的自动重传功能。实际上，对于 DNS 服务器的访问，多次 DNS 请求都返回相同的结果，即做多次和做一次的效果一样。因此 DNS 操作可以重复执行。当一个进程做一次 DNS 请求时，它启动一个定时器。如果定时器计满而未收到回复，那么它就再请求一次，这样做不会有害处。

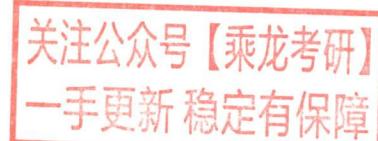
04. 【解答】

IP 地址很难记忆，引入域名是为了便于人们记忆和识别。

域名解析可以把域名转换成 IP 地址。域名转换过程是向本地域名服务器申请解析，如果本地域名服务器查不到，那么向根域名服务器进行查询。如果根域名服务器中也查不到，那么向根域名服务器中保存的顶级域名服务器和相应授权域名服务器进行查询，一定可以查找到。

域名服务器中高速缓存的作用：将近期访问过的域名信息保存在高速缓存，再次访问时会从缓存中读取，不需要重新解析，这样就可以加快域名解析的响应速度。

6.3 文件传输协议（FTP）



6.3.1 FTP 的工作原理

文件传输协议（File Transfer Protocol, FTP）是因特网上使用得最广泛的文件传输协议。FTP 提供交互式的访问，允许客户指明文件的类型与格式，并允许文件具有存取权限。它屏蔽了各计算机系统的细节，因而适合于在异构网络中的任意计算机之间传送文件。

FTP 提供以下功能：

- ① 提供不同种类主机系统（硬、软件体系等都可以不同）之间的文件传输能力。
- ② 以用户权限管理的方式提供用户对远程 FTP 服务器上的文件管理能力。
- ③ 以匿名 FTP 的方式提供公用文件共享的能力。

FTP 采用客户/服务器的工作方式，它使用 TCP 可靠的传输服务。一个 FTP 服务器进程可同时为多个客户进程提供服务。FTP 的服务器进程由两大部分组成：一个主进程，负责接收新的请求；另外有若干从属进程，负责处理单个请求。其工作步骤如下：

- ① 打开熟知端口 21（控制端口），使客户进程能够连接上。
- ② 等待客户进程发连接请求。
- ③ 启动从属进程来处理客户进程发来的请求。主进程与从属进程并发执行，从属进程对客户进程的请求处理完毕后即终止。
- ④ 回到等待状态，继续接收其他客户进程的请求。

FTP 服务器必须在整个会话期间保留用户的状态信息。特别是服务器必须把指定的用户账户与控制连接联系起来，服务器必须追踪用户在远程目录树上的当前位置。

6.3.2 控制连接与数据连接

FTP 在工作时使用两个并行的 TCP 连接（见图 6.7）：一个是控制连接（服务器端口号 21），一个是数据连接（服务器端口号 20）。使用两个不同的端口号可以使协议更容易实现。

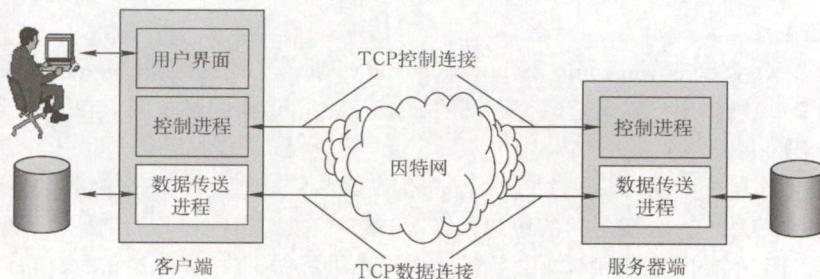


图 6.7 控制连接和数据连接

1. 控制连接

服务器监听 21 号端口，等待客户连接，建立在这个端口上的连接称为控制连接，控制连接用来传输控制信息（如连接请求、传送请求等），并且控制信息都以 7 位 ASCII 格式传送。FTP 客户发出的传送请求，通过控制连接发送给服务器端的控制进程，但控制连接并不用来传送文件。在传输文件时还可以使用控制连接（如客户在传输中途发一个中止传输的命令），因此控制连接在整个会话期间一直保持打开状态。

2. 数据连接

服务器端的控制进程在接收到 FTP 客户发来的文件传输请求后，就创建“数据传送进程”和“数据连接”。数据连接用来连接客户端和服务器端的数据传送进程，数据传送进程实际完成文件的传送，在传送完毕后关闭“数据传送连接”并结束运行。

数据连接有两种传输模式：主动模式 PORT 和被动模式 PASV。PORT 模式的工作原理：客户端连接到服务器的 21 端口，登录成功后要读取数据时，客户端随机开放一个端口，并发送命令告知服务器，服务器收到 PORT 命令和端口号后，通过 20 端口和客户端开放的端口连接，发送数据。PASV 模式不同点是，客户端要读取数据时，发送 PASV 命令到服务器，服务器在本地随机开放一个端口，并告知客户端，客户端再连接到服务器开放的端口进行数据传输。可见，是用 PORT 模式还是 PASV 模式，选择权在客户端。简单概括为，主动模式传送数据是“服务器”连接到“客户端”的端口；被动模式传送数据是“客户端”连接到“服务器”的端口。

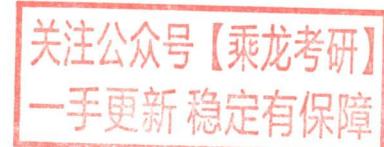
因为 FTP 使用了一个分离的控制连接，所以也称 FTP 的控制信息是带外（Out-of-band）传送的。使用 FTP 时，若要修改服务器上的文件，则需要先将此文件传送到本地主机，然后再将修改后的文件副本传送到原服务器，来回传送耗费很多时间。网络文件系统（NFS）采用另一种思路，它允许进程打开一个远程文件，并能在该文件的某个特定位置开始读写数据。这样，NFS 可使用户复制一个大文件中的一个很小的片段，而不需要复制整个大文件。

6.3.3 本节习题精选

一、单项选择题

01. 文件传输协议（FTP）的一个主要特征是（ ）。
 - A. 允许客户指明文件的类型但不允许指明文件的格式
 - B. 不允许客户指明文件的类型但允许指明文件的格式
 - C. 允许客户指明文件的类型与格式
 - D. 不允许客户指明文件的类型与格式
02. 以下关于 FTP 工作模型的描述中，错误的是（ ）。
 - A. FTP 使用控制连接、数据连接来完成文件的传输

- B. 用于控制连接的 TCP 连接在服务器端使用的熟知端口号为 21
 C. 用与控制连接的 TCP 连接在客户端使用的端口号为 20
 D. 服务器端由控制进程、数据进程两部分组成
03. 控制信息是带外传送的协议是()。
 A. HTTP B. SMTP C. FTP D. POP
04. 下列关于 FTP 连接的叙述中，正确的是()。
 A. 控制连接先于数据连接被建立，并先于数据连接被释放
 B. 数据连接先于控制连接被建立，并先于控制连接被释放
 C. 控制连接先于数据连接被建立，并晚于数据连接被释放
 D. 数据连接先于控制连接被建立，并晚于控制连接被释放
05. FTP 客户发起对 FTP 服务器连接的第一阶段是建立()。
 A. 传输连接 B. 数据连接 C. 会话连接 D. 控制连接
06. 一个 FTP 用户发送了一个 LIST 命令来获取服务器的文件列表，这时服务器应通过()端口来传输该列表。
 A. 21 B. 20 C. 22 D. 19
07. 下列关于 FTP 的叙述中，错误的是()。
 A. FTP 可以在不同类型的的操作系统之间传送文件
 B. FTP 并不适合用在两个计算机之间共享读写文件
 C. 控制连接在整个 FTP 会话期间一直保持
 D. 客户端默认使用端口 20 与服务器建立数据传输连接
08. 当一台计算机从 FTP 服务器下载文件时，在该 FTP 服务器上对数据进行封装的 5 个转换步骤是()。
 A. 比特，数据帧，数据报，数据段，数据
 B. 数据，数据段，数据报，数据帧，比特
 C. 数据报，数据段，数据，比特，数据帧
 D. 数据段，数据报，数据帧，比特，数据
09. 匿名 FTP 访问通常使用()作为用户名。
 A. guest B. E-mail 地址 C. anonymous D. 主机 id
10. 【2009 统考真题】FTP 客户和服务器间传递 FTP 命令时，使用的连接是()。
 A. 建立在 TCP 之上的控制连接 B. 建立在 TCP 之上的数据连接
 C. 建立在 UDP 之上的控制连接 D. 建立在 UDP 之上的数据连接
11. 【2017 统考真题】下列关于 FTP 的叙述中，错误的是()。
 A. 数据连接在每次数据传输完毕后就关闭
 B. 控制连接在整个会话期间保持打开状态
 C. 服务器与客户端的 TCP 20 端口建立数据连接
 D. 客户端与服务器的 TCP 21 端口建立控制连接



二、综合应用题

01. 文件传输协议的主要工作过程是怎样的？主进程和从属进程各起什么作用？
02. 为什么 FTP 要使用两个独立的连接，即控制连接和数据连接？
03. 主机 A 想下载文件 `ftp://ftp.abc.edu.cn/file`，大致描述下载过程中主机和服务器的交互过程。

6.3.4 答案与解析

一、单项选择题

01. C

FTP 提供交互式访问，允许客户指明文件的类型与格式，并允许文件具有存取权限。所以选项 C 为正确答案。

02. C

在服务器端，控制连接使用 TCP 的 21 号端口，数据连接使用 TCP 的 20 号端口；而在客户端，控制连接和数据连接的 TCP 端口号都是由客户端系统自动分配的。需要注意的是，当我们说 FTP 使用 20、21 号端口，HTTP 使用 80 号端口，SMTP 使用 25 号端口时，都是指相应协议的服务器端所使用的端口号，而客户端使用系统自动分配的端口号向这些服务的熟知端口发起连接。

03. C

FTP 传输控制信息使用的是数据连接外的控制连接，因此 FTP 的控制信息是带外传送的。

04. C

FTP 客户首先连接服务器的 21 号端口，建立控制连接（控制连接在整个会话期间一直保持打开），然后建立数据连接，在数据传送完毕后，数据连接最先释放，控制连接最后释放。

05. D

FTP 工作时使用两个连接：控制连接和数据连接。FTP 客户对 FTP 服务器发起连接时，首先建立控制连接，即向服务器的 21 号 TCP 端口发起连接；然后再建立数据连接（20 号 TCP 端口）。FTP 并没有传输连接和会话连接的说法。

06. B

FTP 中数据连接的端口是 20，而文件的列表是通过数据连接来传输的。

07. D

控制连接建立后，服务器进程用自己传送数据的熟知端口 20 与客户进程所提供的端口号建立数据传输连接（默认为 PORT 模式），即客户进程的端口号是客户进程自己提供的。

08. B

FTP 服务器的数据要经过应用层、传输层、网络层、数据链路层及物理层。因此，对应的封装是数据、数据段、数据报、数据帧，最后是比特。

09. C

针对文件传输 FTP，系统管理员建立了一个特殊的用户 ID，名为 anonymous，即匿名用户。Internet 上的任何人在任何地方都可以使用该用户 ID，只是在要求提供用户 ID 时必须输入 anonymous，该用户 ID 的密码可以是任何字符串。

10. A

对于 FTP 文件传输，为了保证可靠性，选择 TCP，排除 C、D。FTP 的控制信息是带外传送的，即 FTP 使用了一个分离的控制连接来传送命令，因此答案为选项 A。

11. C

FTP 使用控制连接和数据连接，控制连接存在于整个 FTP 会话过程中，数据连接在每次文件传输时才建立，传输结束就关闭，选项 A 和 B 正确。默认情况（PORT 模式）下 FTP 服务器使用 TCP 20 端口进行数据连接，使用 TCP 21 端口进行控制连接，这里的端口号是指 FTP 服务器的端口号，选项 C 错误，选项 D 正确。此外还需要注意的是，FTP 服务器是否使用 TCP 20 端口建立数据连接与传输模式有关，PORT 模式使用 TCP 20 端口，PASV 模式由服务器和客户端协商决定。

二、综合应用题

01. 【解答】

FTP的主要工作过程如下：在进行文件传输时，FTP客户所发出的传送请求通过控制连接发送给服务器端的控制进程，并在整个会话期间一直保持打开，但控制连接不用来传送文件。服务器端的控制进程在接收到FTP客户发送来的文件传输请求后，就创建数据传送进程和数据连接，数据连接用来连接客户端和服务器端的数据传输进程，数据传送进程实际完成对文件的传送，在传送完毕后关闭“数据传送连接”，并结束运行。

FTP的服务器进程由两大部分组成：一个主进程，负责接收新的请求；若干从属进程，负责处理单个请求。

02. 【解答】

在FTP的实现中，客户与服务器之间采用了两条传输连接，其中控制连接用于传输各种FTP命令，而数据连接用于文件的传送。之所以这样设计，是因为使用两条独立的连接可使FTP变得更加简单、更容易实现、更有效率。同时在文件传输过程中，还可以利用控制连接控制传输过程，如客户可以请求终止、暂停传输等。

03. 【解答】

大致过程如下：

- ① 建立一个TCP连接到服务器ftp.abc.edu.cn的21号端口，然后发送登录账号和密码。
- ② 服务器返回登录成功信息后，主机A打开一个随机端口，并将该端口号发送给服务器。
- ③ 主机A发送读取文件命令，内容为get file，服务器使用20号端口建立一个TCP连接到主机A的随机打开的端口。
- ④ 服务器把文件内容通过第二个连接发送给主机A，传输完毕后连接关闭。

关注公众号【乘龙考研】
一手更新 稳定有保障

6.4 电子邮件

6.4.1 电子邮件系统的组成结构

自从有了因特网，电子邮件就在因特网上流行起来。电子邮件是一种异步通信方式，通信时不需要双方同时在场。电子邮件把邮件发送到收件人使用的邮件服务器，并放在其中的收件人邮箱中，收件人可以随时上网到自己使用的邮件服务器进行读取。

一个电子邮件系统应具有图6.8所示的三个最主要的构件，即用户代理（User Agent）、邮件服务器和电子邮件使用的协议，如SMTP、POP3（或IMAP）等。



图6.8 电子邮件系统最主要的组成构件

用户代理（UA）：用户与电子邮件系统的接口。用户代理向用户提供一个很友好的接口来发送和接收邮件，用户代理至少应当具有撰写、显示和邮件处理的功能。通常情况下，用户代理就是一

一个运行在 PC 上的程序（电子邮件客户端软件），常见的有 Outlook 和 Foxmail 等。

邮件服务器：它的功能是发送和接收邮件，同时还要向发信人报告邮件传送的情况（已交付、被拒绝、丢失等）。邮件服务器采用客户/服务器方式工作，但它必须能够同时充当客户和服务器。例如，当邮件服务器 A 向邮件服务器 B 发送邮件时，A 就作为 SMTP 客户，而 B 是 SMTP 服务器；反之，当 B 向 A 发送邮件时，B 就是 SMTP 客户，而 A 就是 SMTP 服务器。

邮件发送协议和读取协议：邮件发送协议用于用户代理向邮件服务器发送邮件或在邮件服务器之间发送邮件，如 SMTP；邮件读取协议用于用户代理从邮件服务器读取邮件，如 POP3。注意，SMTP 用的是“推”（Push）的通信方式，即用户代理向邮件服务器发送邮件及在邮件服务器之间发送邮件时，SMTP 客户将邮件“推”送到 SMTP 服务器。而 POP3 用的是“拉”（Pull）的通信方式，即用户读取邮件时，用户代理向邮件服务器发出请求，“拉”取用户邮箱中的邮件。

电子邮件的发送、接收过程可简化为如图 6.9 所示。

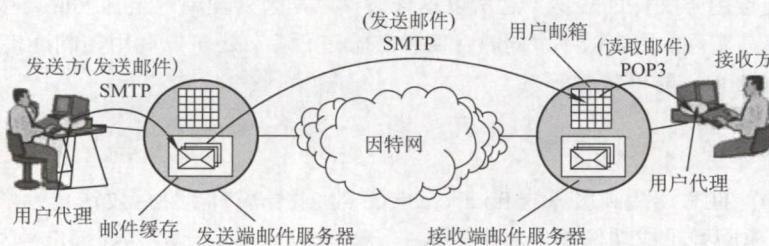


图 6.9 电子邮件的发送、接收过程

下面简单介绍电子邮件的收发过程。

- ① 发信人调用用户代理来撰写和编辑要发送的邮件。用户代理用 SMTP 把邮件传送给发送端邮件服务器。
- ② 发送端邮件服务器将邮件放入邮件缓存队列中，等待发送。
- ③ 运行在发送端邮件服务器的 SMTP 客户进程，发现邮件缓存中有待发送的邮件，就向运行在接收端邮件服务器的 SMTP 服务器进程发起建立 TCP 连接。
- ④ TCP 连接建立后，SMTP 客户进程开始向远程 SMTP 服务器进程发送邮件。当所有待发送邮件发完后，SMTP 就关闭所建立的 TCP 连接。
- ⑤ 运行在接收端邮件服务器中的 SMTP 服务器进程收到邮件后，将邮件放入收信人的用户邮箱，等待收信人在方便时进行读取。
- ⑥ 收信人打算收信时，调用用户代理，使用 POP3（或 IMAP）协议将自己的邮件从接收端邮件服务器的用户邮箱中取回（如果邮箱中有来信的话）。

6.4.2 电子邮件格式与 MIME

1. 电子邮件格式

一个电子邮件分为信封和内容两大部分，邮件内容又分为首部和主体两部分。RFC 822 规定了邮件的首部格式，而邮件的主体部分则让用户自由撰写。用户写好首部后，邮件系统自动地将信封所需的信息提取出来并写在信封上，用户不需要亲自填写信封上的信息。

邮件内容的首部包含一些首部行，每个首部行由一个关键字后跟冒号再后跟值组成。有些关键字是必需的，有些则是可选的。最重要的关键字是 To 和 Subject。

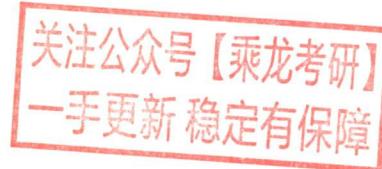
To 是必需的关键字，后面填入一个或多个收件人的电子邮件地址。电子邮件地址的规定格式

为：收件人邮箱名@邮箱所在主机的域名，如 abc@cskaoyan.com，其中收信人邮箱名即用户名，abc 在 cskaoyan.com 这个邮件服务器上必须是唯一的。这也就保证了 abc@cskaoyan.com 这个邮件地址在整个因特网上是唯一的。

Subject 是可选关键字，是邮件的主题，反映了邮件的主要内容。

当然，还有一个必填的关键字是 **From**，但它通常由邮件系统自动填入。首部与主体之间用一个空行进行分割。典型的邮件内容如下：

```
From:hoopdog@hust.edu.cn
To:abc@cskaoyan.com
Subject:Say hello to Internet
}
blahblah...
...
```



2. 多用途国际邮件扩充 (MIME)

由于 SMTP 只能传送一定长度的 ASCII 码邮件，许多其他非英语国家的文字（如中文、俄文，甚至带重音符号的法文或德文）就无法传送，且无法传送可执行文件及其他二进制对象，因此提出了多用途网络邮件扩充（Multipurpose Internet Mail Extensions, MIME）。

MIME 并未改动 SMTP 或取代它。MIME 的意图是继续使用目前的格式，但增加了邮件主体的结构，并定义了传送非 ASCII 码的编码规则。也就是说，MIME 邮件可在现有的电子邮件程序和协议下传送。MIME 与 SMTP 的关系如图 6.10 所示。

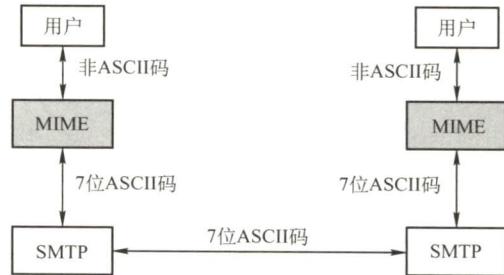


图 6.10 SMTP 与 MIME 的关系

MIME 主要包括以下三部分内容：

- ① 5 个新的邮件首部字段，包括 MIME 版本、内容描述、内容标识、传送编码和内容类型。
- ② 定义了许多邮件内容的格式，对多媒体电子邮件的表示方法进行了标准化。
- ③ 定义了传送编码，可对任何内容格式进行转换，而不会被邮件系统改变。

6.4.3 SMTP 和 POP3

1. SMTP

简单邮件传输协议（Simple Mail Transfer Protocol, SMTP）是一种提供可靠且有效的电子邮件传输的协议，它控制两个相互通信的 SMTP 进程交换信息。由于 SMTP 使用客户/服务器方式，因此负责发送邮件的 SMTP 进程就是 SMTP 客户，而负责接收邮件的 SMTP 进程就是 SMTP 服务器。SMTP 用的是 TCP 连接，端口号为 25。SMTP 通信有以下三个阶段。

(1) 连接建立

发件人的邮件发送到发送方邮件服务器的邮件缓存中后，SMTP 客户就每隔一定时间对邮件缓存扫描一次。如发现有邮件，就使用 SMTP 的熟知端口号（25）与接收方邮件服务器的 SMTP 服务器建立 TCP 连接。连接建立后，接收方 SMTP 服务器发出 220 Service ready（服务就绪）。然后 SMTP 客户向 SMTP 服务器发送 HELO 命令，附上发送方的主机名。

SMTP 不使用中间的邮件服务器。TCP 连接总是在发送方和接收方这两个邮件服务器之间直接建立，而不管它们相隔多远，不管在传送过程中要经过多少个路由器。当接收方邮件服务器因

故障暂时不能建立连接时，发送方的邮件服务器只能等待一段时间后再次尝试连接。

(2) 邮件传送

连接建立后，就可开始传送邮件。邮件的传送从 MAIL 命令开始，MAIL 命令后面有发件人的地址。如 MAIL FROM: <hoopdog@hust.edu.cn>。若 SMTP 服务器已准备好接收邮件，则回答 250 OK。接着 SMTP 客户端发送一个或多个 RCPT (收件人 recipient 的缩写) 命令，格式为 RCPT TO: <收件人地址>。每发送一个 RCPT 命令，都应有相应的信息从 SMTP 服务器返回，如 250 OK 或 550 No such user here (无此用户)。

RCPT 命令的作用是，先弄清接收方系统是否已做好接收邮件的准备，然后才发送邮件，以便不至于发送了很长的邮件后才知道地址错误，进而避免浪费通信资源。

获得 OK 的回答后，客户端就使用 DATA 命令，表示要开始传输邮件的内容。正常情况下，SMTP 服务器回复的信息是 354 Start mail input; end with <CRLF>.<CRLF>。<CRLF> 表示回车换行。此时 SMTP 客户端就可开始传送邮件内容，并用<CRLF>.<CRLF>表示邮件内容的结束。

(3) 连接释放

邮件发送完毕后，SMTP 客户应发送 QUIT 命令。SMTP 服务器返回的信息是 221 (服务关闭)，表示 SMTP 同意释放 TCP 连接。邮件传送的全部过程就此结束。

2. POP3 和 IMAP

邮局协议 (Post Office Protocol, POP) 是一个非常简单但功能有限的邮件读取协议，现在使用的是它的第 3 个版本 POP3。POP3 采用的是“拉”(Pull) 的通信方式，当用户读取邮件时，用户代理向邮件服务器发出请求，“拉”取用户邮箱中的邮件。

POP 也使用客户/服务器的工作方式，在传输层使用 TCP，端口号为 110。接收方的用户代理上必须运行 POP 客户程序，而接收方的邮件服务器上则运行 POP 服务器程序。POP 有两种工作方式：“下载并保留”和“下载并删除”。在“下载并保留”方式下，用户从邮件服务器上读取邮件后，邮件依然会保存在邮件服务器上，用户可再次从服务器上读取该邮件；而使用“下载并删除”方式时，邮件一旦被读取，就被从邮件服务器上删除，用户不能再次从服务器上读取。

另一个邮件读取协议是因特网报文存取协议 (IMAP)，它比 POP 复杂得多，IMAP 为用户提供了创建文件夹、在不同文件夹之间移动邮件及在远程文件夹中查询邮件等联机命令，为此 IMAP 服务器维护了会话用户的状态信息。IMAP 的另一特性是允许用户代理只获取报文的某些部分，例如可以只读取一个报文的首部，或多部分 MIME 报文的一部分。这非常适用于低带宽的情况，用户可能并不想取回邮箱中的所有邮件，尤其是包含很多音频或视频的大邮件。

此外，随着万维网的流行，目前出现了很多基于万维网的电子邮件，如 Hotmail、Gmail 等。这种电子邮件的特点是，用户浏览器与 Hotmail 或 Gmail 的邮件服务器之间的邮件发送或接收使用的是 HTTP，而仅在不同邮件服务器之间传送邮件时才使用 SMTP。

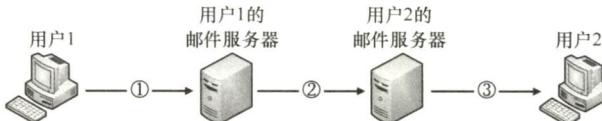
6.4.4 本节习题精选

一、单项选择题

01. 因特网用户的电子邮件地址格式必须是 ()。
 - A. 用户名@单位网络名
 - B. 单位网络名@用户名
 - C. 邮箱所在主机的域名@用户名
 - D. 用户名@邮箱所在主机的域名

02. SMTP 基于传输层的 () 协议，POP3 基于传输层的 () 协议。
 - A. TCP, TCP
 - B. TCP, UDP
 - C. UDP, UDP
 - D. UDP, UDP

03. 用 Firefox 在 Gmail 中向邮件服务器发送邮件时，使用的是（ ）协议。
 A. HTTP B. POP3 C. P2P D. SMTP
04. 用户代理只能发送而不能接收电子邮件时，可能是（ ）地址错误。
 A. POP3 B. SMTP C. HTTP D. Mail
05. 不能用于用户从邮件服务器接收电子邮件的协议是（ ）。
 A. HTTP B. POP3 C. SMTP D. IMAP
06. 下列关于电子邮件格式的说法中，错误的是（ ）。
 A. 电子邮件内容包括邮件头与邮件体两部分
 B. 邮件头中发信人地址（From:）、发送时间、收信人地址（To:）及邮件主题（Subject:）是由系统自动生成的
 C. 邮件体是实际要传送的信函内容
 D. MIME 允许电子邮件系统传输文字、图像、语音与视频等多种信息
07. 下列关于 POP3 协议的说法，（ ）是错误的。
 A. 由客户端而非服务器选择接收后是否将邮件保存在服务器上
 B. 登录到服务器后，发送的密码是加密的
 C. 协议是基于 ASCII 码的，不能发送二进制数据
 D. 一个账号在服务器上只能有一个邮件接收目录
08. 【2012 统考真题】若用户 1 与用户 2 之间发送和接收电子邮件的过程如下图所示，则图中①、②、③阶段分别使用的应用层协议可以是（ ）。



- A. SMTP、SMTP、SMTP B. POP3、SMTP、POP3
 C. POP3、SMTP、SMTP D. SMTP、SMTP、POP3
09. 【2013 统考真题】下列关于 SMTP 的叙述中，正确的是（ ）。
- I. 只支持传输 7 比特 ASCII 码内容
 II. 支持在邮件服务器之间发送邮件
 III. 支持从用户代理向邮件服务器发送邮件
 IV. 支持从邮件服务器向用户代理发送邮件
 A. 仅 I、II 和 III B. 仅 I、II 和 IV C. 仅 I、III 和 IV D. 仅 II、III 和 IV
10. 【2015 统考真题】通过 POP3 协议接收邮件时，使用的传输层服务类型是（ ）。
 A. 无连接不可靠的数据传输服务
 B. 无连接可靠的数据传输服务
 C. 有连接不可靠的数据传输服务
 D. 有连接可靠的数据传输服务
11. 【2018 统考真题】无须转换即可由 SMTP 直接传输的内容是（ ）。
 A. JPEG 图像 B. MPEG 视频 C. EXE 文件 D. ASCII 文本

关注公众号【乘龙考研】
一手更新 稳定有保障

二、综合应用题

01. 电子邮件系统使用 TCP 传送邮件，为什么有时会遇到邮件发送失败的情况？为什么有时对方会收不到发送的邮件？

02. MIME 与 SMTP 的关系是怎样的？

03. 下面列出的是使用 TCP/IP 通信的两台主机 A 和 B 传送邮件的对话过程，请根据这个对话回答问题。

A: 220 beta.gov simple mail transfer service ready

B: HELO alpha.edu

A: 250 beta.gov

B: MAIL FROM:<smith@alpha.edu>

A: 250 mail accepted

B: RCPT TO:<jones@beta.gov>

A: 250 recipient accepted

B: RCPT TO:<green@beta.gov>

A: 550 no such user here

B: RCPT TO:brown@beta.gov

A: 250 recipient accepted

B: DATA

A: 354 start mail input; end with <CR><LF>.<CR><LF>

B: Date:Fri 27 May 2011 14:16:21 BJ

B: From:smith@alpha.edu

B: ...

B: ...

B: .

A: 250 OK

B: QUIT

A: 221 beta.gov service closing transmission channel.

问题：

1) 邮件接收方和发送方机器的全名是什么？发邮件的用户名是什么？

2) 发送方想把邮件发给几个用户？它们的名字各是什么？

3) 哪些用户能收到该邮件？

4) 传送邮件所使用的传输层协议的名称是什么？

5) 为了接收邮件，接收方机器上等待连接的端口号是多少？

关注公众号【乘龙考研】
一手更新 稳定有保障

6.4.5 答案与解析

一、单项选择题

01. D

电子邮件是因特网最基本、最常用的服务功能。要使用电子邮件服务，首先要拥有自己的电子邮件地址，其格式为：用户名@邮箱所在主机的域名。

02. A

SMTP 和 POP3 都是基于 TCP 的协议，提供可靠的邮件通信。

03. A

在基于万维网的电子邮件中，用户浏览器与 Hotmail 或 Gmail 的邮件服务器之间的邮件发送或接收使用的是 HTTP，而仅在不同邮件服务器之间传送邮件时才使用 SMTP。

04. A

用户代理使用 POP3 协议接收邮件。通常用户在配置电子邮件用户代理时需要设置邮件服务器的 POP3 地址（如 pop3.gmail.com），若这个地址设置错误，则会导致用户无法接收邮件。用户代理中的 SMTP 地址错误时会导致无法发送邮件。收件人 E-mail 地址错误时，可能会发错人，也可能会导致投递失败（不存在的地址）。

05. C

SMTP 是一种“推”协议，用于发送方用户代理与发送方服务器之间及发送方服务器与接收方服务器之间，不能用于接收方用户从服务器上读取邮件。常用的邮件读取协议有 POP3、HTTP 和 IMAP。大家平时通过浏览器登录 163 邮箱、Gmail 邮箱时，使用的邮件读取协议就是 HTTP。IMAP 是另一个专用于读取邮件的协议，它要比 POP3 复杂得多，功能也更为强大。

06. B

邮件头是由多项内容构成的，其中一部分是由系统自动生成的，如发信人地址（From:）、发送时间；另一部分是由发件人输入的，如收信人地址（To:）、邮件主题（Subject:）等。

07. B

POP3 协议在传输层是使用明文来传输密码的，并不对密码进行加密。所以 B 选项错误。POP3 协议基于 ASCII 码，如果要传输非 ASCII 码的数据，那么要使用 MIME 将数据转换成 ASCII 码形式。

08. D

SMTP 采用“推”的通信方式，即用户代理向邮件服务器及邮件服务器之间发送邮件时，SMTP 客户主动将邮件“推”送到 SMTP 服务器。而 POP3 采用“拉”的通信方式，即用户读取邮件时，用户代理向邮件服务器发出请求，“拉”取用户邮箱中的邮件。

09. A

根据 6.4.1 节可知，SMTP 用于用户代理向邮件服务器发送邮件，或在邮件服务器之间发送邮件。SMTP 只支持传输 7 比特的 ASCII 码内容。

10. D

POP3 建立在 TCP 连接上，使用的是有连接可靠的数据传输服务。

11. D

电子邮件出现得较早，当时的数据传输能力较弱，使用者往往也不需要传输较大的图片、视频等，因此 SMTP 具有一些目前来看较为老旧的性质，如限制所有邮件报文的体部分只能采用 7 位 ASCII 码来表示。在如今的传输过程中，如果传输了非文本文件，那么往往需要将这些多媒体文件重新编码为 ASCII 码再传输。因此无须转换即可传输的是 ASCII 文本，答案为选项 D。

二、综合应用题

01. 【解答】

有时对方的邮件服务器不工作，邮件就发送不出去。对方的邮件服务器出故障也会使邮件丢失。有时网络非常拥塞，路由器丢弃大量的 IP 数据报，导致通信中断。

02. 【解答】

由于 SMTP 存在着一些缺点和不足，通过 MIME 并非改变或取代 SMTP。MIME 继续使用 RFC 822 格式，但增加了邮件主体的结构，并定义了传送非 ASCII 码的编码规则。也就是说，MIME 邮件可在已有的电子邮件和协议下传送。

03. 【解答】

- 1) 邮件接收方机器的全名是 beta.gov，邮件发送方机器的全名是 alpha.edu，发邮件的用户名是 smith。

- 2) 发送方想把该邮件发给三个用户，它们的名字分别是 jones、green 和 brown。
- 3) 用户 jones 和 brown 能收到邮件，beta.gov 上不存在用户 green。
- 4) 传送邮件所用的传输层协议称为 TCP（传输控制协议）。
- 5) 为了接收邮件，接收方服务器上等待连接的端口号是 25。

6.5 万维网（WWW）

6.5.1 WWW 的概念与组成结构

万维网（World Wide Web, WWW）是一个分布式、联机式的信息存储空间，在这个空间中：一样有用的事物称为一样“资源”，并由一个全域“统一资源定位符”（URL）标识。这些资源通过超文本传输协议（HTTP）传送给使用者，而后者通过单击链接来获取资源。

万维网使用链接的方法能非常方便地从因特网上的一个站点访问另一个站点（即“链接到另一个站点”），从而主动地按需获取丰富的信息。超文本标记语言（HyperText Markup Language, HTML）使得万维网页面的设计者可以很方便地用一个超链接从本页面的某处链接到因特网上的任何一个万维网页面，并能够在自己的计算机屏幕上显示这些页面。

万维网的内核部分是由三个标准构成的：

- 1) 统一资源定位符（URL）。负责标识万维网上的各种文档，并使每个文档在整个万维网的范围内具有唯一的标识符 URL。
- 2) 超文本传输协议（HTTP）。一个应用层协议，它使用 TCP 连接进行可靠的传输，HTTP 是万维网客户程序和服务器程序之间交互所必须严格遵守的协议。
- 3) 超文本标记语言（HTML）。一种文档结构的标记语言，它使用一些约定的标记对页面上的各种信息（包括文字、声音、图像、视频等）、格式进行描述。

URL 是对可以从因特网上得到的资源的位置和访问方法的一种简洁表示。URL 相当于一个文件名在网络范围的扩展。URL 的一般形式是：

<协议>://<主机>:<端口>/<路径>。

<协议>指用什么协议来获取万维网文档，常见的协议有 http、ftp 等；<主机>是存放资源的主机在因特网中的域名或 IP 地址；<端口>和<路径>有时可省略。在 URL 中不区分大小写。

万维网以客户/服务器方式工作。浏览器是在用户主机上的万维网客户程序，而万维网文档所驻留的主机则运行服务器程序，这台主机称为万维网服务器。客户程序向服务器程序发出请求，服务器程序向客户程序送回客户所要的万维网文档。工作流程如下：

- 1) Web 用户使用浏览器（指定 URL）与 Web 服务器建立连接，并发送浏览请求。
- 2) Web 服务器把 URL 转换为文件路径，并返回信息给 Web 浏览器。
- 3) 通信完成，关闭连接。

万维网是无数个网络站点和网页的集合，它们在一起构成了因特网最主要的部分（因特网也包括电子邮件、Usenet 和新闻组）。

6.5.2 超文本传输协议（HTTP）

HTTP 定义了浏览器（万维网客户进程）怎样向万维网服务器请求万维网文档，以及服务器怎样把文档传送给浏览器。从层次的角度看，HTTP 是面向事务的（Transaction-oriented）应用层

协议，它规定了在浏览器和服务器之间的请求和响应的格式与规则，是万维网上能够可靠地交换文件（包括文本、声音、图像等各种多媒体文件）的重要基础。

1. HTTP 的操作过程

从协议执行过程来说，浏览器要访问 WWW 服务器时，首先要完成对 WWW 服务器的域名解析。一旦获得了服务器的 IP 地址，浏览器就通过 TCP 向服务器发送连接建立请求。

万维网的大致工作过程如图 6.11 所示。每个万维网站点都有一个服务器进程，它不断地监听 TCP 的端口 80（默认），当监听到连接请求后便与浏览器建立 TCP 连接。然后，浏览器就向服务器发送请求获取某个 Web 页面的 HTTP 请求。服务器收到请求后，将构建所请求 Web 页的必需信息，并通过 HTTP 响应返回给浏览器。浏览器再将信息进行解释，然后将 Web 页显示给用户。最后，TCP 连接释放。

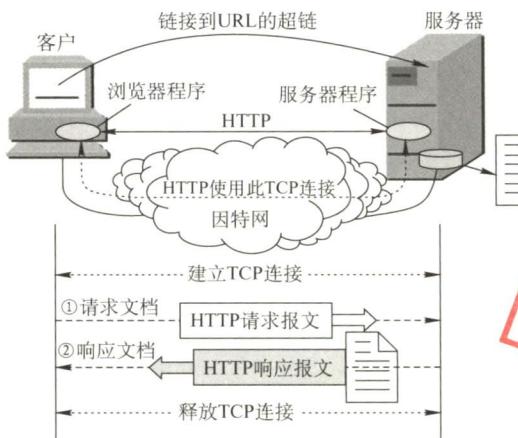


图 6.11 万维网的工作过程

在浏览器和服务器之间的请求与响应的交互，必须遵循规定的格式和规则，这些格式和规则就是 HTTP。因此 HTTP 有两类报文：请求报文（从 Web 客户端向 Web 服务器发送服务请求）和响应报文（从 Web 服务器对 Web 客户端请求的回答）。

用户单击鼠标后所发生的事件按顺序如下（以访问清华大学的网站为例）：

- 1) 浏览器分析链接指向页面的 URL (<http://www.tsinghua.edu.cn/chn/index.htm>)。
- 2) 浏览器向 DNS 请求解析 www.tsinghua.edu.cn 的 IP 地址。
- 3) 域名系统 DNS 解析出清华大学服务器的 IP 地址。
- 4) 浏览器与该服务器建立 TCP 连接（默认端口号为 80）。
- 5) 浏览器发出 HTTP 请求：GET /chn/index.htm。
- 6) 服务器通过 HTTP 响应把文件 index.htm 发送给浏览器。
- 7) 释放 TCP 连接。
- 8) 浏览器解释文件 index.htm，并将 Web 页显示给用户。

2. HTTP 的特点

HTTP 使用 TCP 作为传输层协议，保证了数据的可靠传输。HTTP 不必考虑数据在传输过程中被丢弃后又怎样被重传。但是，HTTP 本身是无连接的（务必注意）。也就是说，虽然 HTTP 使用了 TCP 连接，但通信的双方在交换 HTTP 报文之前不需要先建立 HTTP 连接。

HTTP 是无状态的。也就是说，同一个客户第二次访问同一个服务器上的页面时，服务器的

响应与第一次被访问时的相同。因为服务器并不记得曾经访问过的这个客户，也不记得为该客户曾经服务过多少次。

HTTP 的无状态特性简化了服务器的设计，使服务器更容易支持大量并发的 HTTP 请求。在实际应用中，通常使用 Cookie 加数据库的方式来跟踪用户的活动（如记录用户最近浏览的商品等）。Cookie 的工作原理：当用户浏览某个使用 Cookie 的网站时，该网站服务器就为用户产生一个唯一的识别码，如“123456”，接着在给用户的响应报文中添加一个 Set-cookie 的首部行“Set cookie: 123456”。用户收到响应后，就在它管理的特定 Cookie 文件中添加这个服务器的主机名和 Cookie 识别码，当用户继续浏览这个网站时，会取出这个网站的识别码，并放入请求报文的 Cookie 首部行“Cookie: 123456”。服务器根据请求报文中的 Cookie 识别码就能从数据库中查询到该用户的活动记录，进而执行一些个性化的工作，如根据用户的历史浏览记录向其推荐新产品等。

HTTP 既可以使用非持久连接，也可以使用持久连接（HTTP/1.1 支持）。

对于非持久连接，每个网页元素对象（如 JPEG 图形、Flash 等）的传输都需要单独建立一个 TCP 连接，如图 6.12 所示（第三次握手的报文段中捎带了客户对万维网文档的请求）。请求一个万维网文档所需的时间是该文档的传输时间（与文档大小成正比）加上两倍往返时间 RTT（一个 RTT 用于 TCP 连接，另一个 RTT 用于请求和接收文档）。每个对象引用都导致 $2 \times \text{RTT}$ 的开销，此外每次建立新的 TCP 连接都要分配缓存和变量，使万维网服务器的负担很重。

所谓持久连接，是指万维网服务器在发送响应后仍然保持这条连接，使同一个客户（浏览器）和该服务器可以继续在这条连接上传送后续的 HTTP 请求和响应报文，如图 6.13 所示。

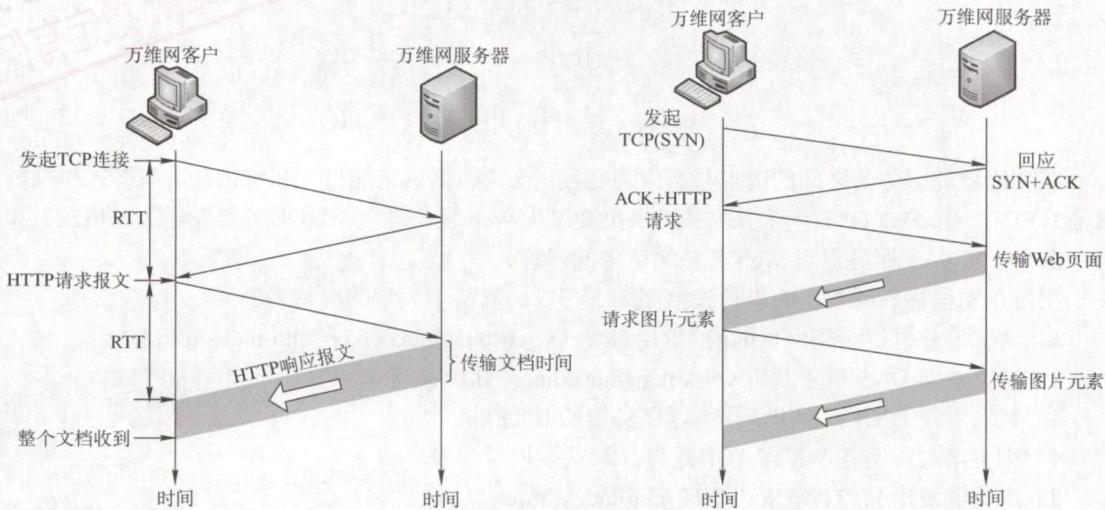


图 6.12 请求一个万维网文档所需的时间

图 6.13 使用持久连接（非流水线）

持久连接又分为非流水线和流水线两种方式。对于非流水线方式，客户在收到前一个响应后才能发出下一个请求，服务器发送完一个对象后，其 TCP 连接就处于空闲状态，浪费了服务器资源。HTTP/1.1 的默认方式是使用流水线的持久连接，这种情况下，客户每遇到一个对象引用就立即发出一个请求，因而客户可以逐个地连续发出对各个引用对象的请求。如果所有的请求和响应都是连续发送的，那么所有引用的对象共计经历 1 个 RTT 延迟，而不是像非流水线方式那样，每个引用都必须有 1 个 RTT 延迟。这种方式减少了 TCP 连接中的空闲时间，提高了效率。

3. HTTP 的报文结构

HTTP 是面向文本的 (Text-Oriented)，因此报文中的每个字段都是一些 ASCII 码串，并且每个字段的长度都是不确定的。有两类 HTTP 报文：

- **请求报文：**从客户向服务器发送的请求报文，如图 6.14(a)所示。
- **响应报文：**从服务器到客户的回答，如图 6.14(b)所示。

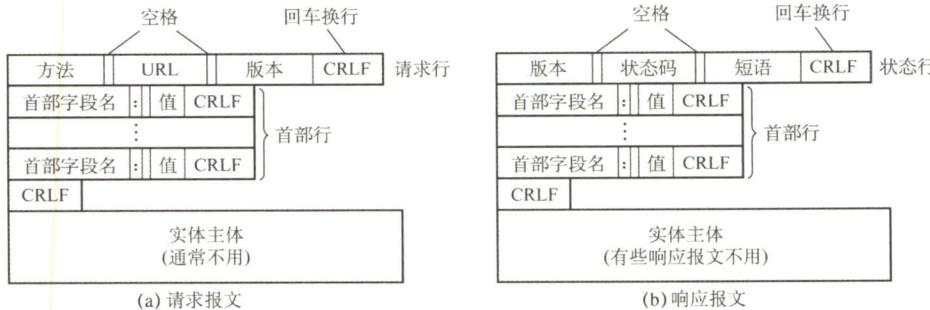


图 6.14 HTTP 的报文结构

HTTP 请求报文和响应报文都由三个部分组成。从图 6.14 可以看出，这两种报文格式的区别就是开始行不同。

开始行：用于区分是请求报文还是响应报文。在请求报文中的开始行称为请求行，而在响应报文中的开始行称为状态行。开始行的三个字段之间都以空格分隔，最后的“CR”和“LF”分别代表“回车”和“换行”。请求报文的“请求行”有三个内容：方法、请求资源的 URL 及 HTTP 的版本。其中，“方法”是对所请求对象进行的操作，这些方法实际上也就是一些命令。表 6.1 给出了 HTTP 请求报文中常用的几个方法。

首部行：用来说明浏览器、服务器或报文主体的一些信息。首部可以有几行，但也可以不使用。在每个首部行中都有首部字段名和它的值，每一行在结束的地方都要有“回车”和“换行”。整个首部行结束时，还有一空行将首部行和后面的实体主体分开。

实体主体：在请求报文中一般不用这个字段，而在响应报文中也可能没有这个字段。

图 6.15 所示为使用 Wireshark 捕获的 HTTP 请求报文的示例，下面结合前几章的内容对请求报文（图中下部分）进行分析。

根据帧的结构定义，在图 6.15 所示的以太网数据帧中，第 1~6 个字节为目的 MAC 地址（默认网关地址），即 00-0f-e2-3f-27-3f；第 7~12 个字节为本机 MAC 地址，即 00-27-13-67-73-8d；第 13~14 个字节 08~00 为类型字段，表示上层使用的是 IP 数据报协议。第 15~34 个字节（共 20B）为 IP 数据报的首部，其中第 27~30 个字节为源 IP 地址，即 db-df-d2-70，转换成十进制为 219.223.210.112；第 31~34 个字节为目的 IP 地址，即 71-69-4e-0a，转换成十进制为 113.105.78.10。第 35~54 个字节（共 20B）为 TCP 报文段的首部。

从第 55 个字节开始才是 TCP 数据部分（阴影部分），即从应用层传递下来的数据（本例中即请求报文），GET 对应请求行的方法，/face/20.gif 对应请求行的 URL，HTTP/1.1 对应请求行的版本，左边数字是对应字符的 ASCII 码，如'G'=0x47、'E'=0x45、'T'=0x54 等。图 6.15 的请求报文中首部行字段内容的含义，建议读者自行了解，也可以自己动手抓包分析。

表 6.1 HTTP 请求报文中常用的几个方法

方法（操作）	意 义
GET	请求读取由 URL 标识的信息
HEAD	请求读取由 URL 标识的信息的首部
POST	给服务器添加信息（如注释）
CONNECT	用于代理服务器

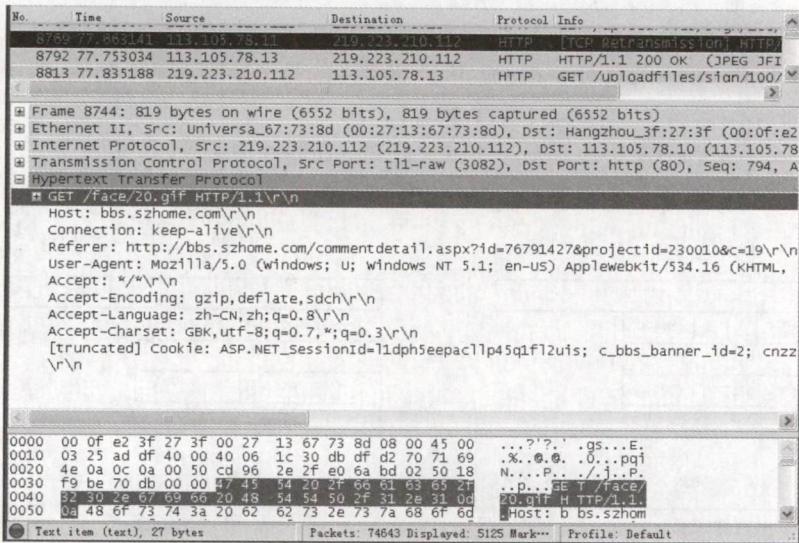


图 6.15 使用 Wireshark 捕获的 HTTP 请求报文的示例

右下角开始的“…??.gs…E..@.0…pgi”等是上面介绍过的第 1~54 个字节中对应的 ASCII 码字符，而这些字符在这里不代表任何意义。

常见应用层协议小结如表 6.2 所示。

表 6.2 常见应用层协议小结

应用程序	FTP 数据连接	FTP 控制连接	TELNET	SMTP	DNS	TFTP	HTTP	POP3	SNMP
使用协议	TCP	TCP	TCP	TCP	UDP	UDP	TCP	TCP	UDP
熟知端口号	20	21	23	25	53	69	80	110	161

6.5.3 本节习题精选

一、单项选择题

01. 下面的（ ）协议中，客户机与服务器之间采用面向无连接的协议进行通信。
 - A. FTP
 - B. SMTP
 - C. DNS
 - D. HTTP
02. 从协议分析的角度，WWW 服务的第一步操作是浏览器对服务器的（ ）。
 - A. 请求地址解析
 - B. 传输连接建立
 - C. 请求域名解析
 - D. 会话连接建立
03. TCP 和 UDP 的一些端口保留给一些特定的应用使用。为 HTTP 保留的端口号为（ ）。
 - A. TCP 的 80 端口
 - B. UDP 的 80 端口
 - C. TCP 的 25 端口
 - D. UDP 的 25 端口
04. 从某个已知的 URL 获得一个万维网文档时，若该万维网服务器的 IP 地址开始时并不知道，则需要用到的应用层协议有（ ）。
 - A. FTP 和 HTTP
 - B. DNS 和 FTP
 - C. DNS 和 HTTP
 - D. TELNET 和 HTTP
05. 万维网上的每个页面都有一个唯一的地址，这些地址统称为（ ）。
 - A. IP 地址
 - B. 域名地址
 - C. 统一资源定位符
 - D. WWW 地址

06. 使用鼠标单击一个万维网文档时，若该文档除有文本外，还有三幅 gif 图像，则在 HTTP/1.0 中需要建立（ ）次 UDP 连接和（ ）次 TCP 连接。
 A. 0, 4 B. 1, 3 C. 0, 2 D. 1, 2
07. 仅需 Web 服务器对 HTTP 报文进行响应，但不需要返回请求对象时，HTTP 请求报文应该使用的方法是（ ）。
 A. GET B. PUT C. POST D. HEAD
08. HTTP 是一个无状态协议，然而 Web 站点经常希望能够识别用户，这时需要用到（ ）。
 A. Web 缓存 B. Cookie C. 条件 GET D. 持久连接
09. 下列关于 Cookie 的说法中，错误的是（ ）。
 A. Cookie 存储在服务器端 B. Cookie 是服务器产生的
 C. Cookie 会威胁客户的隐私 D. Cookie 的作用是跟踪用户的访问和状态
10. 以下关于非持续连接 HTTP 特点的描述中，错误的是（ ）。
 A. HTTP 支持非持续连接与持续连接
 B. HTTP/1.0 使用非持续连接，而 HTTP/1.1 的默认方式为持续连接
 C. 非持续连接中对每次请求/响应都要建立一次 TCP 连接
 D. 非持续连接中读取一个包含 100 个图片对象的 Web 页面，需要打开和关闭 100 次 TCP 连接
11. 【2014 统考真题】使用浏览器访问某大学的 Web 网站主页时，不可能使用到的协议是（ ）。
 A. PPP B. ARP C. UDP D. SMTP
12. 【2015 统考真题】某浏览器发出的 HTTP 请求报文如下：

```
GET /index.html HTTP/1.1
Host: www.test.edu.cn
Connection: Close
Cookie: 123456
```



- 下列叙述中，错误的是（ ）。
- A. 该浏览器请求浏览 index.html
 B. index.html 存放在 www.test.edu.cn 上
 C. 该浏览器请求使用持续连接
 D. 该浏览器曾经浏览过 www.test.edu.cn
13. 【2022 统考真题】假设主机 H 通过 HTTP/1.1 请求浏览某 Web 服务器 S 上的 Web 页 news408.html, news408.html 引用了同目录下的 1 幅图像，news408.html 文件大小为 1MSS (最大段长)，图像文件大小为 3MSS，H 访问 S 的往返时间 RTT=10 ms，忽略 HTTP 响应报文的首部开销和 TCP 段传输时延。若 H 已完成域名解析，则从 H 请求与 S 建立 TCP 连接时刻起，到接收到全部内容止，所需的时间至少是（ ）。
- A. 30ms B. 40ms C. 50ms D. 60ms

二、综合应用题

01. 在浏览器中输入 http://cskaoyan.com 并按回车，直到王道论坛的首页显示在其浏览器中，请问在此过程中，按照 TCP/IP 参考模型，从应用层到网络层都用到了哪些协议？
02. 在如下条件下，计算使用非持续方式和持续方式请求一个 Web 页面所需的时间：
 1) 测试的 RTT 的平均值为 150ms，一个 gif 对象的平均发送时延为 35ms。

2) 一个 Web 页面中有 10 幅 gif 图片, Web 页面的基本 HTML 文件、HTTP 请求报文、TCP 握手报文大小忽略不计。

3) TCP 三次握手的第三步中捎带一个 HTTP 请求。

4) 使用非流水线方式。

03. 用户主机上的电子邮件用户代理与邮件服务器建立了连接, 现截获一个 TCP 报文段, 如下图所示。图中显示了该报文段的前 126 个字节的十六进制及 ASCII 码内容。TCP 首部长度为 20B。请回答:

0020	c0 e6 00 19 b0 ca	d5 6f eb c9 10 e9 50 18O....P.
0030	f9 98 51 bd 00 00 4d 65	73 73 61 67 65 2d 49 44	.Q...Me ssage-ID
0040	3a 20 3c 34 44 43 45 39	32 42 41 2e 32 30 31 30	: <4DCE9 2BA.2010
0050	39 30 32 40 31 36 33 2e	63 6f 6d 3e 0d 0a 44 61	902@163. com>..Da
0060	74 65 3a 20 53 61 74 2c	20 31 34 20 4d 61 79 20	te: Sat, 14 May
0070	32 30 31 31 20 32 32 3a	33 33 3a 33 30 20 2b 30	2011 22: 33:30 +0
0080	38 30 30 0d 0a 46 72 6f	6d 3a 20 63 73 6b 61 6f	800..Fro m: cskao
0090	79 61 6e 32 30 31 32 40	31 36 33 2e 63 6f 6d 0d	yan2012@ 163.com.

1) 用户代理和服务器之间使用的应用层协议是什么?

2) 用户代理使用的端口号是多少?

3) 该邮件的发件人邮箱是什么?

04. 【2011 统考真题】某主机的 MAC 地址为 00-15-C5-C1-5E-28, IP 地址为 10.2.128.100 (私有地址)。图 1 是网络拓扑, 图 2 是该主机进行 Web 请求的一个以太网数据帧前 80B 的十六进制及 ASCII 码内容。

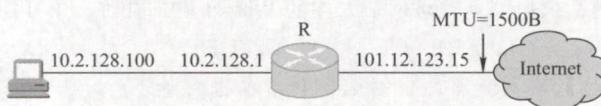


图 1 网络拓扑

0000 00 21 27 21 51 ee 00 15	c5 c1 5e 28 08 00 45 00	.! !Q... ..^(..E.
0010 01 ef 11 3b 40 00 80 06	ba 9d 0a 02 80 64 40 aa	...:@...d@..
0020 62 20 04 ff 00 50 e0 e2	00 fa 7b f9 f8 05 50 18	b ...P... ...{..P.
0030 fa f0 1a c4 00 00 47 45	54 20 2f 72 66 63 2e 68GE T /rfc.h
0040 74 6d 6c 20 48 54 54 50	2f 31 2e 31 0d 0a 41 63	tml HTTP /1.1..Ac

图 2 以太网数据帧 (前 80B)

请参考图中的数据回答以下问题。

- 1) Web 服务器的 IP 地址是什么? 该主机的默认网关的 MAC 地址是什么?
- 2) 该主机在构造图 2 的数据帧时, 使用什么协议确定目的 MAC 地址? 封装该协议请求报文的以太网帧的目的 MAC 地址是什么?
- 3) 假设 HTTP/1.1 协议以持续的非流水线方式工作, 一次请求-响应时间为 RTT, rfc.html 页面引用了 5 幅 JPEG 小图像。问从发出图 2 中的 Web 请求开始到浏览器收到全部内容为止, 需要多少个 RTT?
- 4) 该帧封装的 IP 分组经过路由器 R 转发时, 需修改 IP 分组头中的哪些字段?

注: 以太网数据帧结构和 IP 分组头结构分别如图 3 和图 4 所示。

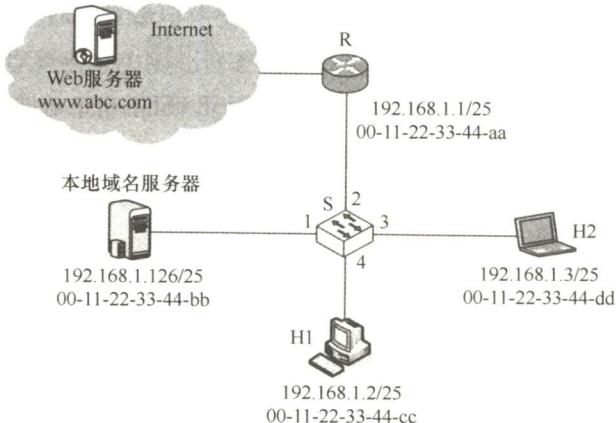
6B	6B	2B	46~1500B	4B
目的 MAC 地址	源 MAC 地址	类型	数据	CRC

图 3 以太网帧结构



图 4 IP 分组头结构

05. 【2021 统考真题】某网络拓扑如下图所示，以太网交换机 S 通过路由器 R 与 Internet 互连。路由器部分接口、本地域名服务器、H1、H2 的 IP 地址和 MAC 地址如图中所示。在 t_0 时刻 H1 的 ARP 表和 S 的交换表均为空，H1 在此刻利用浏览器通过域名 www.abc.com 请求访问 Web 服务器，在 t_1 时刻 ($t_1 > t_0$) S 第一次收到了封装 HTTP 请求报文的以太网帧，假设从 t_0 到 t_1 期间网络未发生任何与此次 Web 访问无关的网络通信。



请回答下列问题。

- 1) 从 t_0 到 t_1 期间，H1 除了 HTTP，还运行了哪个应用层协议？从应用层到数据链路层，该应用层协议报文是通过哪些协议进行逐层封装的？
- 2) 若 S 的交换表结构为<MAC 地址，端口>，则 t_1 时刻 S 交换表的内容是什么？
- 3) 从 t_0 到 t_1 期间，H2 至少接收到几个与此次 Web 访问相关的帧？接收的是什么帧？帧的目的 MAC 地址是什么？

6.5.4 答案与解析

一、单项选择题

01. C

DNS 采用 UDP 来传送数据，UDP 是一种面向无连接的协议。

02. C

建立浏览器与服务器之间的连接需要知道服务器的 IP 地址和端口号（80 端口是熟知端口），而访问站点时浏览器从用户那里得到的是 WWW 站点的域名，所以浏览器必须首先向 DNS 请求域名解析，获得服务器的 IP 地址后，才能请求建立 TCP 连接。

03. A

关注公众号【乘龙考研】
一手更新 稳定有保障

HTTP 在传输层使用 TCP，端口号为 80。TCP 的 25 号端口是为 SMTP 保留的。

04. C

由于不知道服务器的 IP 地址，因此先要用 DNS 进行域名解析，然后使用 HTTP 进行用户和服务器之间的交互。

05. C

统一资源定位符负责标识万维网上的各种文档，并使每个文档在整个万维网的范围内具有唯一的标识符 URL。

06. A

HTTP 在传输层用的是 TCP，所以无须建立 UDP 连接；HTTP 1.0 只支持非持久连接，所以每请求一个对象需要建立一次 TCP 连接，在本题的情景中，共需要传输 1 个基本 HTML 对象和 3 个 gif 对象，所以共需建立 4 次 TCP 连接。

07. D

使用 HEAD 方法时服务器可对 HTTP 报文进行响应，但不会返回请求对象，其作用主要是调试。另外三个选项中的方法的作用请查看本章中的表 6.1。

08. B

可以在 HTTP 中使用 Cookie 保存 HTTP 服务器和客户之间传递的状态信息。

09. A

Cookie 是一个存储在用户主机中的文本文件。它由服务器产生，作为识别用户的手段。由于服务器的后端数据库记录了用户在 Web 站点上的活动，这些信息（如用户的个人信息及购物的偏好等）有可能被出卖给第三方，从而威胁到了用户的隐私。

10. D

非持续连接对每次请求/响应都建立一次 TCP 连接。在浏览器请求一个包含 100 个图片对象的 Web 页面时，服务器需要传输 1 个基本 HTML 文件和 100 个图片对象，因此一共是 101 个对象，需要打开和关闭 TCP 连接 101 次。

11. D

接入网络时可能会用到 PPP，A 可能用到；计算机不知道某主机的 MAC 地址时，用 IP 地址查询相应的 MAC 地址会用到 ARP，B 可能用到；访问 Web 网站时，若 DNS 缓冲没有存储相应域名的 IP 地址，用域名查询相应的 IP 地址时要使用 DNS，而 DNS 是基于 UDP 的，所以 C 可能用到；SMTP 只有使用邮件客户端发送邮件，或邮件服务器向其他邮件服务器发送邮件时才会用到，单纯地访问 Web 网页不可能用到，选 D。

12. C

Connection:连接方式，Close 表明为非持续连接方式，keep-alive 表示持续连接方式。Cookie 值由服务器产生，HTTP 请求报文中有 Cookie 报头表示曾经访问过 www.test.edu.cn 服务器。

13. B

HTTP/1.1 默认使用流水线的持久连接，所有请求都是连续发送的。要求最少时间，最理想的情况是 TCP 在第三次握手的报文段中捎带 HTTP 请求，以及 TCP 连接后慢开始阶段不考虑拥塞。假设接收方有足够的缓存空间，即发送窗口等同于拥塞窗口，共需要经过：第 1 个 RTT，进行 TCP 连接，此时服务器 S 的发送窗口 = 1MSS，并在第三次握手时捎带 HTTP 请求；第 2 个 RTT，服务器 S 发送大小为 1MSS 的 html 文件，主机 C 确认后服务器 S 的发送窗口变为 2MSS；第 3 个 RTT，服务器 S 发送大小为 2MSS 的图像文件，主机 C 确认后服务器 S 的发送窗口变为 4MSS；第 4 个 RTT，服务器 S 发送剩下的 1MSS 图像文件，完成传输，总共需要 4 个 RTT，即 40ms。

二、综合应用题

01. 【解答】

- 1) 应用层。HTTP: WWW 访问协议; DNS: 域名解析服务。
- 2) 传输层。TCP: HTTP 提供可靠的数据传输; UDP: DNS 使用 UDP 传输。
- 3) 网络层。IP: IP 包传输和路由选择; ICMP: 提供网络传输中的差错检测; ARP: 将本机的默认网关 IP 地址映射成物理 MAC 地址。

02. 【解答】

每次进行 TCP 三次握手时, 前两次握手消耗一个 $RTT = 150ms$, 第 3 次握手的报文段捎带客户对 HTML 文件的请求, 因此请求和接收基本 HTML 文件耗时一个 $RTT = 150ms$ (其大小忽略不计时, 发送时延为 0ms)。

在非持久连接方式下:

第一次建立 TCP 连接并传送 html 文件所需的时间为 $t_{html} = (150 + 150)ms = 300ms$;

每次建立 TCP 连接并传送一个 gif 文件所需的时间为 $t_{gif} = (150 + 150 + 35)ms = 335ms$;

所以总时间 $t_{\text{总}} = t_{html} + t_{gif} \times 10 = (300 + 335 \times 10)ms = 3650ms$ 。

在持久连接方式下:

只需要建立一次 TCP 连接, 然后传送 html 文件和 10 个 gif 文件。

总时间 $t_{\text{总}} = t_{\text{建立 TCP}} + t_{html} + t_{gif} \times 10 = 150 + 150 + (150 + 35) \times 10 = 2150ms$ 。

关注公众号【乘龙考研】
一手更新 稳定有保障

03. 【解答】

- 1) 本题中并未明确告诉这个报文段是从用户代理发往服务器还是从服务器发往用户代理。分析 TCP 首部格式可知, 源端口为 49382 (0xc0e6), 目的端口为 25 (0x0019), 因此该应用层协议为 SMTP。
- 2) 由于使用的是 SMTP, 且服务器端口 25 作为目的端口, 因此源端口 49382 为用户代理所使用的端口。
- 3) 由于 SMTP 的协议字段都是用 ASCII 码表示的, 发件人的关键字是 FROM, 从截图右侧的 ASCII 形式中直接找到答案 FROM: cskaoyan2012@163.com。

04. 【解答】

- 1) 以太网帧的数据部分是 IP 数据报, 只要数出相应字段所在的字节即可。由图 3 可知以太网帧头部有 $6 + 6 + 2 = 14B$, 由图 4 可知 IP 数据报首部的目的 IP 地址字段前有 $4 \times 4 = 16B$, 从图 2 的帧第 1 字节开始数 $14 + 16 = 30B$, 得到目的 IP 地址为 40.aa.62.20 (十六进制), 转换成十进制为 64.170.98.32。由图 2 可知以太网帧的前 6 字节 00-21-27-21-51-ee 是目的 MAC 地址, 即为主机的默认网关 10.2.128.1 端口的 MAC 地址。
- 2) ARP 用于解决 IP 地址到 MAC 地址的映射问题。主机的 ARP 进程在本以太网以广播形式发送 ARP 请求分组, 在以太网上广播时, 以太网帧的目的地址为全 1, 即 FF-FF-FF-FF-FF-FF。
- 3) HTTP/1.1 协议以持续的非流水线方式工作时, 服务器发送响应后仍在一段时间内保持这段连接, 客户机在收到前一个请求的响应后才能发出下一个请求。第一个 RTT 用于请求 Web 页面, 客户机收到第一个请求的响应后 (还有五个请求未发送), 每访问一次对象就用去一个 RTT。因此共需 $1 + 5 = 6$ 个 RTT 后浏览器收到全部内容。
- 4) 私有地址和 Internet 上的主机通信时, 须由 NAT 路由器进行网络地址转换, 把 IP 数据报的源 IP 地址 (本题为私有地址 10.2.128.100) 转换为 NAT 路由器的一个全球 IP 地址 (本

题为 101.12.123.15)。因此，源 IP 地址字段 0a 02 80 64 变为 65 0c 7b 0f。IP 数据报每经过一个路由器，TTL 值就减 1，并重新计算首部校验和。若 IP 分组的长度超过输出链路的 MTU，则总长度字段、标志字段、片偏移字段也会发生变化。

05. 【解答】

1) 从 t_0 到 t_1 期间，除了 HTTP，H1 还运行了 DNS 应用层协议，以将域名转换为 IP 地址。DNS 运行在 UDP 之上，UDP 将应用层交付的 DNS 报文添加首部后，向下交付给 IP 层，IP 层使用 IP 数据报进行封装，封装好后，向下交付给数据链路层，数据链路层使用 CSMA/CD 帧进行封装。因此，逐层封装关系如下：DNS 报文 → UDP 数据报 → IP 数据报 → CSMA/CD 帧。

2) t_0 时刻，H1 的 ARP 表和 S 的交换表为空。H1 利用浏览器通过域名请求访问 Web 服务器。由于要先解析域名，会发送 DNS 报文到本地域名服务器，查询该域名对应的 IP 地址，所以要先向本地域名服务器发送请求。ARP 表为空，所以需要先发送 ARP 请求分组，查询本地域名服务器对应的 MAC 地址。这些帧的目的 MAC 地址均是 FF-FF-FF-FF-FF-FF。S 接收到这个帧，在交换表中记录 MAC 地址为 00-11-22-33-44-cc，位于端口 4，然后广播该帧。当本地域名服务器接收到 ARP 请求后，向 H1 发送响应 ARP 分组。S 接收到这个帧，在交换表中记录 MAC 地址为 00-11-22-33-44-bb，位于端口 1，然后将该帧从端口 4 发送出去。

得到了域名对应的 IP 地址，发现不在本局域网中，需要通过路由表转发。

H1 的 ARP 表中并没有路由器对应的 MAC 地址，因此需要先发送 ARP 请求分组，查询路由器对应的 MAC 地址。这些帧的目的 MAC 地址均是 FF-FF-FF-FF-FF-FF。S 接收到这个帧，广播该帧。当路由器收到 ARP 请求后，向 H1 发送响应 ARP 分组。S 接收到这个帧，在交换表中记录 MAC 地址为 00-11-22-33-44-aa，位于端口 2，然后将该帧从端口 4 发送出去。现在，H1 就能将数据发送给路由器了。在整个过程中，并没有涉及 H2，H2 没有主动发送数据。所以 S 不会记录 H2 的 MAC 地址和端口，所以 S 在 t_1 时刻的交换表如下表所示。

MAC 地址	端口
00-11-22-33-44-cc	4
00-11-22-33-44-bb	1
00-11-22-33-44-aa	2

3) 由步骤 2) 的分析可知，H2 至少会接收到 2 个和此次 Web 访问相关的帧。接收到的均是封装 ARP 查询报文的以太网帧；这些帧的目的 MAC 地址均是 FF-FF-FF-FF-FF-FF。

6.6 本章小结及疑难点

关注公众号【乘龙考研】
一手更新 稳定有保障

1. 如何理解客户进程端口号与服务器进程端口号？

通常我们所说的熟知端口号是指应用层协议在服务器端的默认端口号，而客户端进程的端口号是由客户端进程任意指定的（临时的）。

当客户进程向服务器进程发出建立连接请求时，要寻找连接服务器进程的熟知端口号，同时还要告诉服务器进程自己的临时端口号。接着，服务器进程就用自己的熟知端口号与客户进程所提供的端口号建立连接。

2. 互联网、因特网和万维网的区别是什么？

互联网（internet）泛指由多个计算机网络按照一定的通信协议相互连接而成的一个大型计算机网络。

因特网（Internet）是指在 ARPA 网基础上发展而来的世界上最大的全球性互连网络。因特网和其他类似的由计算机相互连接而成的大型网络系统，都可算是“互联网”，因特网只是互联网中最大的一个。

万维网是无数个网络站点和网页的集合，它们一起构成了因特网最主要的部分（因特网也包括电子邮件、Usenet 和新闻组）。

3. 域名的高速缓存是什么？

每个域名服务器都维护一个高速缓存，存放最近用过的名字以及从何处获得名字映射信息的记录，可大大减轻根域名服务器的负荷，使因特网上的 DNS 查询请求和回答报文的数量大为减少。为保持高速缓存中的内容正确，域名服务器应为每项内容设置计时器，并处理超过合理时间的项（如每个项目只存放两天）。当权限域名服务器回答一个查询请求时，在响应中都指明绑定有效存在的时间值。增加此时间值可减少网络开销，减少此时间值可以提高域名转换的准确性。



参 考 文 献

- [1] 谢希仁. 计算机网络(第8版) [M]. 北京: 电子工业出版社, 2021.
- [2] James F. Kurose, Keith W. Ross. 计算机网络: 自顶向下方法[M]. 北京: 机械工业出版社, 2019.
- [3] 本书编写组. 计算机专业基础综合考试大纲解析[M]. 北京: 高等教育出版社, 2009.
- [4] 黄传河. 计算机网络考研指导[M]. 北京: 机械工业出版社, 2009.
- [5] 鲁士文. 计算机网络习题与解析[M]. 北京: 清华大学出版社, 2005.
- [6] 张沪寅, 黄传河等. 计算机网络考研指导[M]. 北京: 清华大学出版社, 2010.
- [7] 翔高教育. 计算机学科专业基础综合复习指南[M]. 上海: 复旦大学出版社, 2009.
- [8] 崔魏等. 计算机学科专业基础综合辅导讲义[M]. 北京: 原子能出版社, 2011.

王道论坛&网易慕课考研

经久才是王道

十五年考研磨练

2024年王道计算机考研课程

- ◆ 阶段一：计算机考研零基础入门
- ◆ 阶段二：大纲考点、经典习题精讲
- ◆ 阶段三：暑期强化提升训练
- ◆ 阶段四：真题、模拟题、考前冲刺
- ◆ 阶段五：复试机试课程



加客服领券购课

抄底价/299元起

- 2024年数据结构考研复习指导
- 2024年计算机组成原理考研复习指导
- 2024年操作系统考研复习指导
- 2024年计算机网络考研复习指导
- 2024年计算机专业基础综合考试历年真题解析
- 2024年计算机专业基础综合考试冲刺模拟题
- 计算机考研——机试指南（第2版）



责任编辑：谭海平
封面设计：张昱

ISBN 978-7-121-44473-9



9 787121 444739 >

定价：69.00 元