|  |
| --- |
| **IP概述**  **一些概念：**   1. [IP](https://tools.ietf.org/html/rfc791)是TCP/IP协议族中最为核心的协议。所有的TCP，UDP，ICMP，IGMP数据都是以IP数据报格式传输。IP协议就是定义数据如何从源地址传送到目的地址，同时也定义在传输过程中的数据分片和重新组装 2. IP协议是被它的上层协议（TCP，UDP）调用，同时它也调用它的下层协议（以太网协议，IEEE802），把IP数据报传到下一个网关（gateway）或目的地址 3. IP协议主要实现两个基本功能：寻址和数据分片   **IP运作模式（model of operation）**  http://images.cnitblog.com/blog/322405/201301/30215049-d38a67b062f44ba587b7b9960ed035a0.jpg   1. 网络1中的应用程序把要数据交给IP模块处理，加上目标，源地址和相关参数等信息放到IP头中 2. IP模块根据特定的链路层网络协议，进行相应的协议转换。例如如果现在是以太网局域网，则根据目的IP地址得到相应的MAC地址，并封装IP数据包为特定的链路层数据帧格式进行传输到局域肉的网关 3. 网关接收到数据后，去掉相关的链路层头信息，交由IP模块处理 4. IP模块去掉链路层相关数据头信息后，根据IP地址和特定链路层协议，把数据报传到网络2中的网关。 5. 网络2中的网关根据目的IP地址得到对应硬件地址MAC地址，并将数据传输到指定目的主机上 6. 目的主机IP模块将得到的数据去掉底层协议头信息后将给相应的应用程序进程以供调用处理   **IP头部格式定义（IP Header Format）**  http://images.cnitblog.com/blog/322405/201301/30232516-5ef120a7fcc14f0fabb87c96b966dbce.jpg  我们从上到下，从左到右介绍下各字段含义：   * version(版本)： 指IP协议版本号，目前是4，也称IPv4,将来（现在已经在用了）会是IPv6 * IHL(internet header length)：指头部长度数值，以4字节（32比特）为单位。因为IHL占四比特位置，所以最大值是：2^4 -1 = 15，则头部可能最大为：15 ＊ 4字节 ＝ 60字节，一般没有选项(options)数据时值为5（即5 \* 4字节＝20字节） * Type of service(服务类型)： 前三个比特指优先权字段（已被忽略）。中间四个比特表示：最小时延，最大吞吐量，最高可靠性，最小费用。0表时一般，1表示最优。四个中只能一个置1。最后一比特预留，总置为0 * Total length(IP数据总长度) ： IP数据报内容的长度，包括IP头部部分，以字节为单位。因为占16比特，所以最长为65535（2^16-1）。尽管IP数据报最长可以达65535字节，但是我们知道有些链路层有最大传输单元的限制（MTU），如以太网就是1500，这时候IP数据报就要分片（Fragmentation）传输。 * Identification(唯一标识)： 每一个IP数据报都要有唯一的标识字段，这样在分片和重组时不至于混淆两个完全不同的分片数据。这在后面分组，重组中会讲到 * Flages(标志)： 这个字段也是为IP数据报的分片和重组作用的。用于标识收到的分片后续是否还有分片待接收。 * Fragment offset(分片偏移)：指当前分片数据在整体IP数据报中所处的偏移量，以8字节（64比特）为单位。由此可见,「identification」，「Flages」,「Fragment offset」这三个字段是为IP数据报分片和重组功能作用的字段 * Time to live(生存时间) ： 设置IP数据报可以经过的最多路由器数目。每次经过一次处理它的路由器，它的值就减1，值到为0时，该数据报丢弃，并发送ICMP报文到源主机。 * Protocol(协议) ： 指定它的上层协议，如TCP。 * Header checksum(头检验和) ： 为了验证IP头数据的完整性，进行一系列计算求出的检验和值，当接收方收到后也要进行同样的计算并把求出的检验和与收到的相比较，以判断IP数据在传输的过程中是无误传输的。 关于计算检验和的算法，在[RFC1071](https://tools.ietf.org/html/rfc1071)里定义。我没有详细看这份rfc，但是这篇[文章](http://www.thegeekstuff.com/2012/05/ip-header-checksum/)介绍得挺详细的。 * source address 和 destination address 分别指源IP地址和目的IP地址。 * options(头部选项)：一些特定应用时定义的选项值。只需要注意几点：「选项」是可选的（optional），且如果有选项值的话，长度值必需以32比特为上界，不足的话补0。这是因为IP头部始终以32比特为单元。   **IP路由选择**  我们自己的电脑和一台主机通信时，如访问某一个网站的web服务，映射到底层后其实就是IP数据报(ip datagrams)的传输。那么一个IP数据报（假如一个http地址请求的数据）是怎么的方式从我的电脑达到web站点的呢？这里我们设置一个具体网络环境，然后基于这个环境分析下。    http://images.cnitblog.com/blog/322405/201302/02161513-61f5c02217f545939072fd2b339526ae.png    我的电脑主机名是「host」，IP是「192.168.1.121」，假如我现在要访问[www.baidu.com的页面，ping](http://www.baidu.com的页面，ping)得baidu.com的服务器IP是220.181.111.53。则IP数据报的传输过程如下描述：   1. 由于目的地址IP是「220.181.111.53」，与要机IP地址「192.168.1.121」不在一个网络号，因此数据发送到本地局域网的网关上，即本地路由器「192.168.1.1」。 2. 本地路由器接收到数据后，要进行转发，由于本地路由器是通过电信ADSL拨号上网，分配的IP是「113.89.201.22」，与目的IP，即baidu的服务器IP 「220.181.111.53」也不是在一个网络上，因此把数据发到默认网关上转发，即发到电信网关「113.89.201.1」 3. 电信网关「113.89.201.1」通过查找路由表记录，找到一条路由记录项记录的下一站路由IP「220.181.111.12」，它和我们的目的IP地址「220.181.111.53」处在同一个网络号（network address）上 4. 当数据传输到baidu网关「220.181.111.12」时，发现目的地址「220.181.111.53」处在同一个网络中，则通过ARP得到目的主机的硬件地址（MAC），封装成以太网数据帧格式后传输到目的主机「220.181.111.53」上。 5. 至此，整个数据报从源主机「192.168.1.121」到「220.181.111.53」传输完毕   通过上面的一个实例分析，总结下IP路由的一些规范：   * 搜索路由表，寻找与目的IP地址完全匹配的表目：网络号和主机号都要匹配。如果存在，则把报文直接发送到目的IP * 搜索路由表，寻找与目的IP地址相匹配的网络号表目，如果存在，则把报文发送到与之匹配网络号的路由器上 * 如果上述两步中都没有查找成功，则把报文发送到路由器设置的默认下一站路网器上。 * 如果到这里还没有成功，则当前数据报就不能传送了（因为现在都不知道下一站去哪里），则表明发送失败，那么需要向源发送端发送一个“错误”讯号。   **子网**  我们知道，IP地址是由网络号和主机号组成的，如下图所示一个B类地址：  http://images.cnitblog.com/blog/322405/201302/03005526-3c9b674e58444aa3ab02dd270b5608b2.jpg  网络号一般是由一个国际组织机构（InterNIC）颁发的，像上面的140.152；但是后面的子网号（5）和主机号（121）则是由拥有该网络号的管理员自己分配。  假如我们看到一个IP地址，可以马上知道它的网络号（因为可以根据IP地址的类型：A，B，C就可以得知网络号），但是还要知道后面的比特值有多少是为子网号占有，有多少是主机号占有。像上图所示的我们把8比特分给了子网号，另外8比特分给了主机号，但是这只是我们可选的一种分配方式，并没有规定必需这样，也可以按3：13这个比例来分配。所以这个时间，我们就需要设置规定这样的一个分配比较规则，即就是「子网掩码」这个概念。子网掩码可以怎么定义：把32比特的地址值中网络号和子网号全部置为1，主机号置为0，像上面的IP：140.152.5.121的子网掩码格式如下：  http://images.cnitblog.com/blog/322405/201302/03011910-a787de6b2d98421097cd882af722dd7e.jpg  所以给定一个IP地址，加上一个子网掩码，我们就可以等到如下信息：   1. IP地址类型：根据最高位可知，如A，B，C类IP 2. 网络号      ：根据IP类型后，就知道网络号（每个IP类型明确知道网络号占的比特位数） 3. 子网号      ：根据子网掩码的定义，对应的网络号和子网号的位置都是1，所以我们知道了子网号所占的比特值 4. 主机号      ：知道了子网号，同时也知道了主机号   -----------------------------------EOF-----------------------------------------------------------------------------------------------          http://www.cnblogs.com/jcli/aggbug/2880671.html?type=1来自：http://www.cnblogs.com/jcli/archive/2013/02/03/2880671.html |