

# IP 协议详解

## IP 协议概述

- (1). IP 是 TCP/IP 协议族中最核心的协议，不管是 TCP、UDP、ICMP 数据最终都是以 IP 数据报格式传输。
- (2). IP 提供不可靠、无连接、无状态的数据报传输服务。

## IP 首部格式



### 版本:

占 4 位, 指 IP 协议的版本目前的 IP 协议版本号为 4 (即 IPv4)

### 首部长度:

占 4 位, 可表示的最大数值是 15 个单位 (一个单位为 4 字节) 因此 IP 的首部长度的最大值是 60 字节

### 区分服务:

占 8 位, 用来获得更好的服务, 在旧标准中叫做服务类型, 但实际上一直未被使用过. 1998 年这个字段改名为区分服务. 只有在使用区分服务 (DiffServ) 时, 这个字段才起作用. 一般的情况下都不使用这个字段

### 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 65535 字节. 总长度必须不超过最大传送单元 MTU, 长度超过 MTU 的数据报将会分片传输, 所以实际传输的 IP 数据报长度远远没有达到最大值。以下 3 个字段描述了如何分片;

### 标识:

占 16 位, 它是一个计数器, 用来产生数据报的标识. 唯一标识主机发送的每一个数据报, 其初始值是系统随机生成: 每发送一个数据报, 其值加 1. 该值在分片时被复制到每一个分

片中，因此同一个数据报的所有分片都有相同的标识值。

#### **标志(flag):**

占 3 位, 目前只有两位有意义:

(1) MF, 标志字段的最低位是 MF (More Fragment), MF=1 表示后面“还有分片”。MF=0 表示最后一个分片。

(2) DF, 标志字段中间的一位是 DF (Don't Fragment), 只有当 DF=0 时才允许分片。否则 IP 模板将不对数据报分片, 超过 MTU 的数据将会丢弃并返回一个 ICMP 差错报文。

#### **分片偏移:**

占 13 位, 是分片相对原始 IP 数据报开始处的偏移(仅指数据部分)。实际偏移值是该值左移 3 位得到的。所以除最后一个分片, 其他分片的数据部分长度必须是 8 的整数倍。

#### **生存时间(TTL):**

占 8 位, 设置了数据可以经过的最多的路由器数(一般是 64), 每经过一次路由器, 该值减 1, 如果该值减为 0 依旧没有到达目的主机, 就丢弃改数据报, 发送 ICMP 差错报文(目标不可达)

#### **协议:**

占 8 位, 用于区分上层协议。其中 ICMP 为 1, TCP 为 6, UDP 为 17 等等。

#### **首部校验和:**

占 16 位, 由发送端填充, 接收端对其使用 CRC 算法检验 IP 数据报头部在传输过程中是否损坏(只检查头部, 不管数据部分)

#### **源端 IP 地址和目的 IP 地址:**

各占 32 位, 用来指定发送端和接收端的。

## **IP 分片细节**

### **MTU 介绍:**

在 TCP/IP 分层中, 数据链路层用 MTU (Maximum Transmission Unit, 最大传输单元) 来限制所能传输的数据包大小, MTU 是指一次传送的数据最大长度, 不包括数据链路层数据帧的帧头, 如以太网的 MTU 为 1500 字节, 实际上数据帧的最大长度为 1512 字节, 其中以太网数据帧的帧头为 12 字节。

当发送的 IP 数据报的大小超过了 MTU 时, IP 层就需要对数据进行分片, 否则数据将无法发送成功。

### **IP 分片的实现**

IP 分片发生在 IP 层, 不仅源端主机可能会进行分片, 中间的路由器也有可能分片, 因为不同的网络的 MTU 是不一样的, 如果传输路径上的某个网络的 MTU 比源端网络的 MTU 要小, 路由器就可能对 IP 数据报再次进行分片。而分片数据的重组只会发生在目的端的 IP 层。

在 IP 首部有 4 个字节是用于分片的, 如下图所示。前 16 位是 IP 数据报的标识, 同一个数据报的各个分片的标识是一样的, 目的端会根据这个标识来判断 IP 分片是否属于同一个 IP 数据报。中间 3 位是标志位, 其中有 1 位用来表示是否有更多的分片, 如果是最后一个分片, 该标志位为 0, 否则为 1。后面 13 位表示分片在原始数据的偏移, 这里的原始数据是 IP 层收到的传输的 TCP 或 UDP 数据, 不包含 IP 首部。



需要注意的，在分片的数据中，传输层的首部只会出现在第一个分片中，因为传输层的数据格式对 IP 层是透明的，传输层的首部只有在传输层才会有它的作用，IP 层不知道也不需要保证在每个分片中都有传输层首部。所以，在网络上传输的数据包是有可能没有传输层首部的。