前一篇文章学习了IP报文结构，我们有三个字段并未详细说明，即标识、标志和片偏移，本篇文章来学习一下。

一、数据分片原理说明

标识、标志、片偏移是关于IP数据报分片的。那为什么会出现分片呢?

> 什么叫分片?
>
> 答：将报文分割成多个片段的过程叫做分片。

为什么要分片?

我们知道，数据链路层中最大的帧长为1518，IP报文1500byte +帧头18byte = 1518byte

那么也就是说，一帧的数据部分最大是1500字节，这个其实就是以太网协议中规定的最大传输单元MTU。

以太网规定其最大传送单元MTU的值是1500字节，如果从网络层传输下来的数据报长度超过MTU值，就必须把过长的数据报进行分片处理。

很容易可以想到，分完片后接收方肯定要重组，那么我们的分片就得有依据，否则报文重组就头大了。IP报文中用标识、标志、片偏移来用于数据报文分片。

（1）标识：占16位，所有分片的数据报的标识必须要和原数据报的标识相同。假如一个数据报的标识是12345，这个数据报过大，分片后将它分为3个小的数据报，这3个较小的数据报的标识也必须是12345，**可以理解这3个数据报是一个家族的。相同的标识字段的值可以使分片后的各个数据报最后能正确的重装成原来的数据报。**

（2）标志：占3位，目前只有两位有意义。

> 1. 最低位即第3位记为 `MF` （ `More Fragment` ），意思是是否还有更多分片。当值为1时，表示该分片不是最后一片，后面还有分片，当值为0时，表示这是原数据报分片后的最后一片数据报，后面已经没有更多的分片了。
> 2. 中间位即第2位记为 `DF` （ `Don't Fragment` ），意思是原数据报能否分片。当值为1时，表示该数据报不允许分片，当值为0时，表示该数据报允许分片。

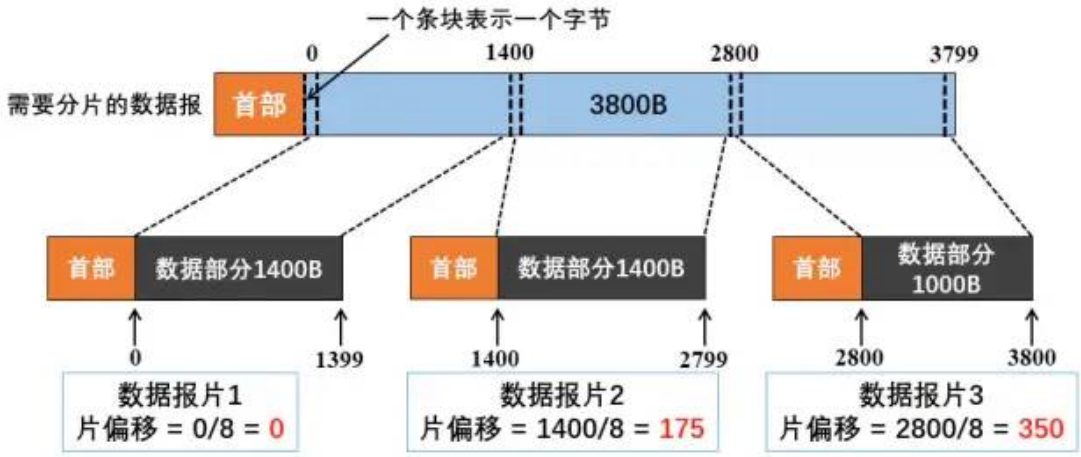
（3）片偏移：占13位，以8B为单位。其表示较长分组分片后，某一片在原分组中的相对位置，也就是说相对于用户数据字段的起点，该片从何处开始。这也就是说，除了最后一个分片，每个分片的长度一定是8B的整数倍。

下面跟着笔者结合具体实例来看一看这三个字段到底如何填充的。

二、分片例子理解

假设一个数据报的总长度是3820个字节，其数据部分为3800字节长（首部仅仅使用固定部分），需要分片为长度不超过1420字节的数据报片。

因固定首部长度为20字节，因此每个数据报片的长度不超过1400字节。于是分为3个报片，其数据部分的长度分别为1400、1400、1000字节。原始数据报首部被复制为各个数据报的首部，但是必须修改有关字段。

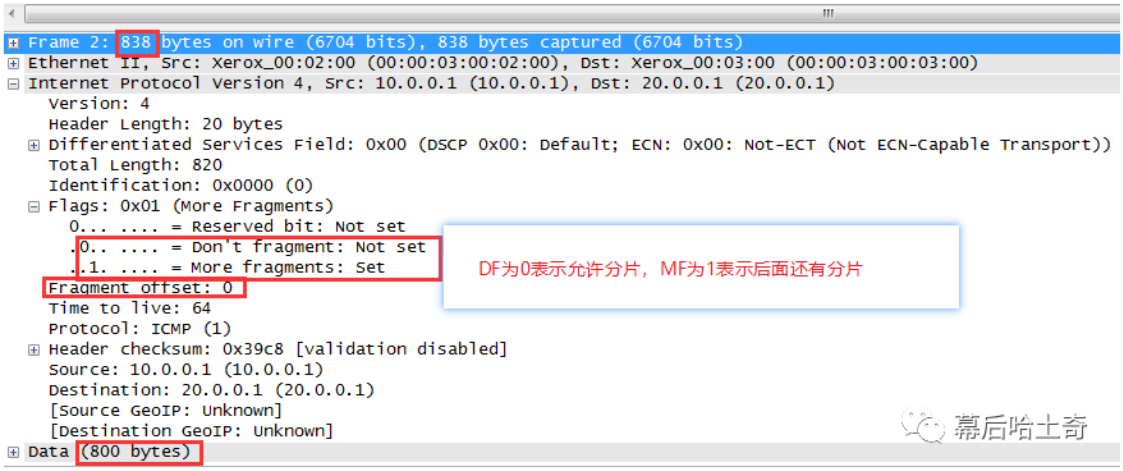


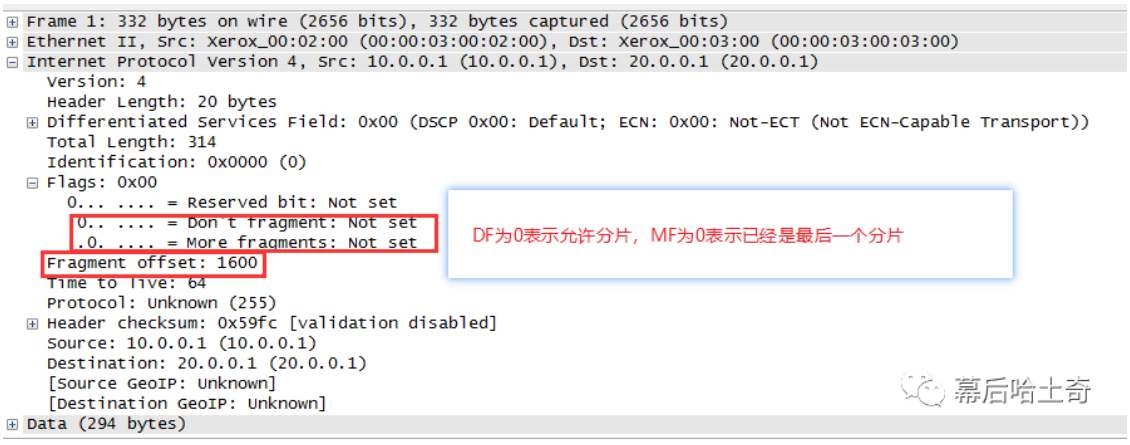
对于原始数据报、数据报片1、2、3的首部部分信息如下图，（原始数据报的标识取12345）

	总长度	标识	MF	DF	片偏移
原始数据报	**3820**	**12345**	**0**	**0**	**0**
数据报片1	1420	12345	1	0	0
数据报片2	1420	12345	1	0	175
数据报片3	1020	12345	0	0	350

三、分片实际抓包举例

以下三张图，就是一个大报文被分片成了三份传输的抓包报文：





我们来简单分析，我们可以看到在数据链路层传递的帧的大小分别为：838字节、838字节、332字节。

首先我们可以看到，标识字段即Identification都是一样的，表示这3个数据报是一个家子的，到时候接收端接收的时候就还把这一家子的人整合起来即可。

并且标志中的DF和MF字段我在图中做了解释，比较简单。

最重要的是偏移量，我们看下第一个报文，总长度为838字节，去除以太网的帧头18个字节，实际上以太网上传输的数据长度为820字节，这个与抓包里面的Total Length长度是吻合的。

那么由于IP报文的首部固定部分有20字节长度，那么无可变部分，那么实际数据长度就是800字节。结合片偏移的单位是8B，那么显然第二个分片的片偏移应该是800/8=100。（只是wireshark抓包后又乘上了8来表示实际的偏移字节），所以第二个显示800，第三个显示1600。可见wireshark的贴心。

你也许会问，为啥不采用MTU为1400字节来分片呢?

实际上，当两台远程PC互联的时候，它们的数据需要穿过很多的路由器和各种各样的网络媒介才能到达对端，网络中不同媒介的MTU各不相同，就好比一长段的水管，由不同粗细的水管组成（MTU不同）通过这段水管最大水量就是由中间最细的水管决定。

对于网络层的上层协议而言，它们对水管粗细不在意，它们认为这是网络层的事情。网络层IP协议会检查每个从上层协议下来的数据包大小，并根据本机MTU的大小决定是否作“分片”处理。**分片最大的坏处就是降低了传输性能，本来可以一次搞定的事情，分多次搞定**，所以在网络层更高一层（即传输层）的实现中往往会对此加以注意。有些高层因为某些原因就会要求不能分层，因此会在IP数据包包头加上一个DF标签。这样当这个数据包在一大段网络里传输的时候，如果遇到MTU小于IP数据包的情况，转发设备就会根据要求丢弃这个数据包。然后返回一个错误信息给发送者。

后续学习TCP的时候我们还将接触到MSS的概念，到时候我们回过头来结合MTU进行对比。

分析了这种分片的报文，那么不分片的报文就更简单了，不再赘述了。下面进入本篇文章的总结阶段。

四、小结

关于IP协议就学习这么多，我们对这两篇文章的内容做一次重点知识的提取。

首先注意区分几个字段的单位：

1. 首部长度：单位是4B，表示数据报的首部的长度。
2. 总长度：单位是B，标识整个数据报的长度。
3. 片偏移量：单位是8B，表示某一分片相对于用户数据字段的起点。

重点学习了IP的报文首部各个字段的位置和含义，关键知识点总结如下：

1. IP数据包由首部+数据组成，其中首部分为固定的20字节+可选部分（长度可变且非必须，最小为0字节，最大为40字节）
2. 首部包含很多字段，比以太网层的帧结构复杂多了

3. a. 版本：IP协议版本，IPV4/IPV6
   b. 首部长度：数据报首部长度，单位是4B
   c. 区分服务：一般情况下不用
   d. 总长度：数据报总长度，即首部+数据部分
   e. 标识：分片的标识要和原数据的标识一致
   f. 标志：MF（是否还有分片，1表示还有分片，0表示已经是最后一个分片）、DF（能否分片，0表示允许分片，1表示不允许分片）
   g. 片偏移：分片相对于用户数据字段的起点，单位是8字节
   h. 生存时间TTL：数据报寿命，每经过一个路由器生存时间减一
   i. 协议：数据部分使用的协议类型，6标识TCP，17标识UDP
   j. 首部校验和：只校验数据报首部
   k. 源地址和目的地址
4. wireshark很人性化，很有必要学习下哦

好了，至此，，我们现在知道了 IP 数据报的头部各个元素的含义及其位置。我们之前提到过，一个网络到另一个网络，中间要穿过很多不同的网络，那么就可能存在多条路径，这就是路由，并且涉及OSI第三层最重要的硬件：路由器，这就是我们下面要学习的重点，本文完。