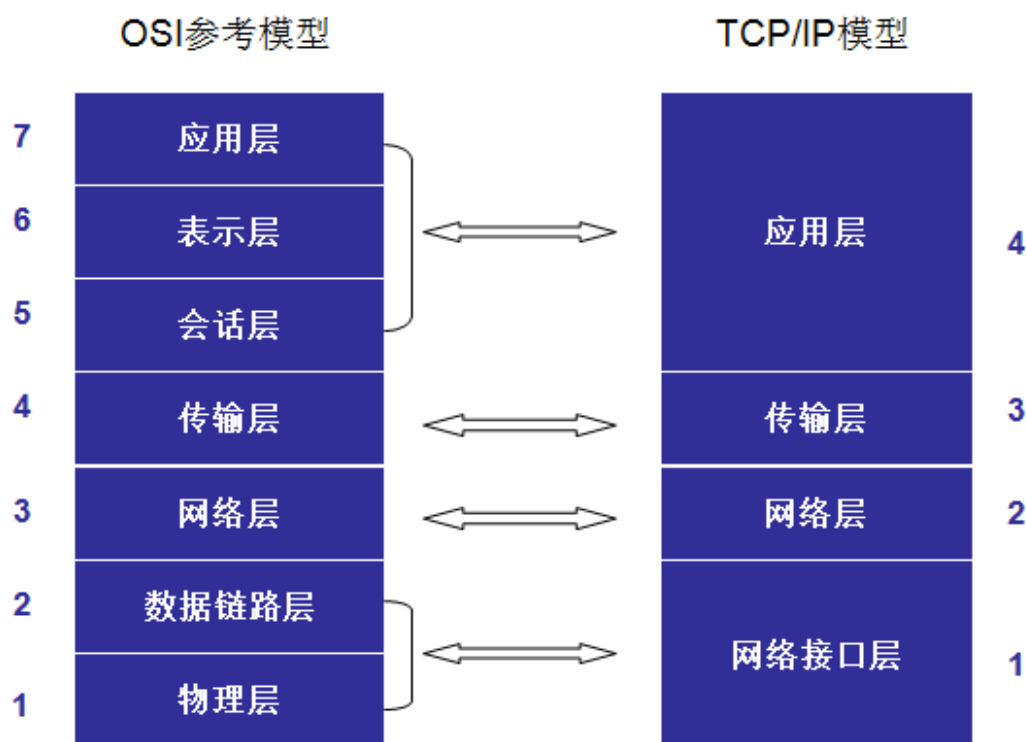


## [面试·网络] TCP/IP（一）：数据链路层

这一系列的文章主要是为一般的、非专业开发岗位(如移动端)的工程师准备，一方面可以对网络的基本知识有基本的了解，另一方面不至于面试中被问到相关问题时束手无策。知识以 TCP/IP 协议簇为主，也会有应用层和数据链路层的简单介绍。文章内容不会很难，也不会过多讨论各种算法，目标是以最快的速度达到最深的理解。内容肯定比直接百度搜索“TCP/IP协议”，然后随便看一篇文章要丰富得多，但也不足以让读者凭此就能胜任网络开发的工作。

诚然，面试以 TCP/UDP/HTTP 等协议为主，IP 协议都涉及甚少，更遑论数据链路层等。但我希望可以从原理上理解那些问题，而不是临时抱佛脚，背了一些答案然后在面试后忘干净。不要为面试而准备面试，为了完善自己的知识体系而准备。如果你觉得这正是你需要的，Let's Begin!

## OSI七层模型和协议



在这一节中，我们不谈这些层和协议的具体作用，目前只要知道 OSI 模型中，网络被分为七层，由底层向高层依次是：物理层，数据链路层，网络层，传输层，会话层，表示层和应用层。

协议是一个 **Big** 很高，出现很频繁的词。其实它很好理解，它实际上是一种通信双方共同遵守的规范。比如我需要把性别和年龄传递给另外一台主机，那么我可以定义一个"A 协议"，协议规定数据的前 4 个字节表示性别，后四个字节表示年龄。这样对方主机接收时就知道前 4 个字节是性别，而不会错把它当成年龄来处理。

整个互联网世界能够运行，完全得益于各个软件、硬件厂商严格遵守现有的协议。以 IP 协议为例，你可以随便修改它，然后自己弄出一个 IP2 协议，只不过没有人认可、遵守这个协议，所以它毫无用武之地。

## 物理层

物理层处于 OSI 七层模型的最底端，它的主要任务是将比特流与电子信号进行转换。

在计算机的世界中，一切都由 0 和 1 组成。你看到的这篇文章，在通过网络传输到你电脑的过程中，自然也是以 0 和 1 的形式存在。但是网络传输的介质(比如光纤，双绞线，电波等)中显然是不存在 0 和 1 的。比如在光线中，数据通过光的形式传递。0 和 1 以光的亮灭表示，其中的转换由物理层完成。

如果没有物理层，由 0 和 1 构成的比特流就无法在物理介质中传播。

## 数据链路层

数据链路层处于 OSI 七层模型的第二层，它定义了通过通信介质相互连接的设备之间，数据传输的规范。

在数据链路层中，数据不再以 0、1 序列的形式存在，它们被分割为一个一个的“帧”，然后再进行传输。

数据链路层中有两个重要的概念：**MAC** 地址和分组交换。

## MAC地址

MAC 地址是被烧录到网卡 ROM 中的一串数字，长度为 48 比特，它在世界范围内唯一(不考虑虚拟机自定义 MAC 地址)。由于 MAC 地址的唯一性，它可以被用来区分不同的节点，一旦指定了 MAC 地址，就不可能出现不知道往哪个设备传输数据的情况。

## 分组交换

分组交换是指将较大的数据分割为若干个较小的数据，然后依次发送。使用分组交换的原因是不同的数据链路有各自的最大传输单元(MTU: Maximum Transmission Unit)。不同的数据链路就好比不同的运输渠道，一辆卡车(对应通信介质)的载重量为 5 吨。那么通过卡车运送 20 吨的货物就需要把这些货物分成四部分，每份重 5 吨。如果运输机的载重量是 30 吨，那么这些货物不需要分割，直接一架运输机就可以拉走。

以以太网(一种数据链路)为例，它的MTU是 1500 字节，也就是通过以太网传输的数据，必须分割为若干帧，每个帧的数据长度不超过 1500 字节。如果上层传来的数据超过这个长度，数据链路层需要分割后再发送。

## 以太网帧

我们用以太网举例，介绍一下以太网帧的格式。

以太网帧的开头是“前导码(Preamble)”，长度为 8 字节，这一段没什么用，重点在于以太网帧的本体。

本体由首部，数据和 FCS 三部分组成：

以太网帧体格式

目标MAC地址 (6字节)	源MAC地址 (6字节)	类型 (2字节)	数据 (46~1500字节)	FCS (4字节)
------------------	-----------------	-------------	-------------------	--------------

自学过程

类型部分存储了上层协议的编号，比如上层是 IP 协议，则编号为 0800。

FCS 表示帧校验序列(Frame Check Sequence)，用于判断帧是否在传输过程中有损坏

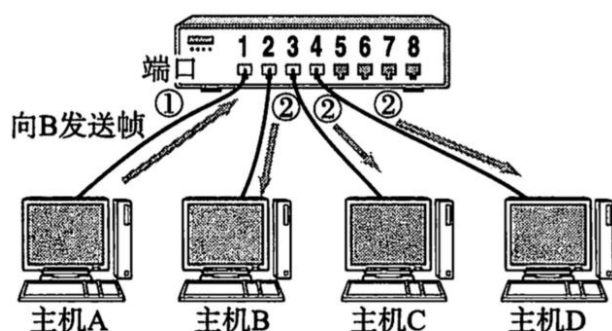
(比如电子噪声干扰)。FCS 保存着发送帧除以某个多项式的余数，接收到的帧也做相同计算，如果得到的值与 FCS 相同则表示没有出错。

## 交换机

交换机是一种在数据链路层工作的网络设备，它有多个端口，可以连接不同的设备。交换机根据每个帧中的目标 MAC 地址决定向哪个端口发送数据，此时它需要参考“转发表”

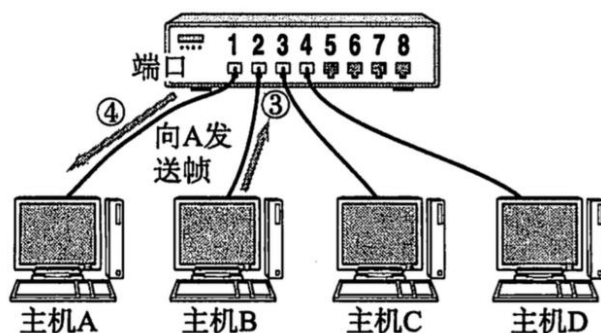
转发表并非手动设置，而是交换机自动学习得到的。当某个设备向交换机发送帧时，交换机将帧的源 MAC 地址和接口对应起来，作为一条记录添加到转发表中。

下图描述了交换机自学过程的原理



① 从源MAC地址可以获知主机A与端口1相连接。

② 拷贝那些以“未知”MAC地址为目标的帧给所有的端口。



③ 从源MAC地址可以获知主机B与端口2相连接。

④ 由于已经知道主机A与端口1相连接，那么发给主机A的帧只拷贝给端口1。

以后，主机A与主机B的通信就只在它们各自所连接的端口之间进行。

自学过程

关于数据链路层，最重要的一点还是它的定义：“通过通信介质相互连接的设备之间，数据传输的规范”。这说明数据链路层的协议适用于**处于同一种通信介质两端的节点**。如果不能理解这一点，就无法理解网络层和 IP 协议。

数据链路层的意义在于，如果没有数据链路层，数据只能以流的形式存在与通信介质

中，不知道该发送往哪里，过长的数据流可能无法在通信介质中传输。