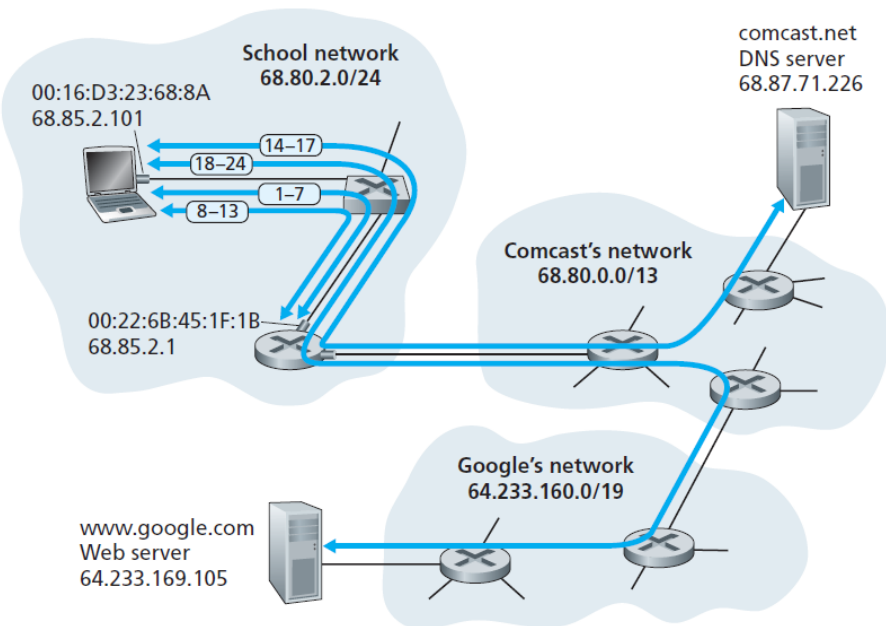


## 浏览器输入一个URL后发生了什么

原创 ACMFans Club Coder Club 2021-06-23

浏览器输入一个URL之后发生了什么？这个问题一直是一个非常经典的问题，透过这个问题，我们可以从数据包的角度，自顶向下（也可以自底向上）地理解计算机网络各层发生的事情，同时也是一个面试的热点问题。但很多的回答都是从应用层和传输层的角度做一个浅显的回答，没有继续对下面的层次进行挖掘。但这篇文章将不仅局限于最上面的两层，而是会涉及到每一层的传输细节。相信只要把这个问题理解思考透彻，对计算机网络的各层协议将会有有一个更加深层次的理解。

假设现在在局域网下，通过以太网线访问google服务器，有如下的网络拓扑关系



为了方便查看，下面将不时出现上面这张图

下面我们沿着协议栈自顶向下，以一个数据包的角度，来观察这个过程发生了什么事情。每个发生事情的背后都有面试或者研究生笔试对应的考点

### 一、DHCP UDP IP 以太网

每个能够上网的主机都需要有一个ip地址，用来唯一标志网络中的主机。【考点：IPv4地址的划分<附录1>、IPv6协议<2>、CIDR、IP数据报格式、子网掩码】

#### 1. 生成DHCP请求报文<3>

本地主机需要请求DHCP服务器（这里假定了DHCP服务器运行在路由器中）分配得到一个ip地址，因此，本地主机发出DHCP请求报文，这个报文被封装在UDP报文中。不难想到，本地主机一开始不知道路由器的ip地址，所以它也不知道应该将这个DHCP请求包往哪发，所以只能将这个UDP报文封装在一个目的ip地址为255.255.255.255（广播）和源ip地址为0.0.0.0（本地）的IP数据报中。【考点：单播、多播（组播）、广播】这个数据帧接着翻译成比特流放置到物理层进行传输。【考点：物理层编解码】

#### 2. 继续向下，封装以太帧传输至DHCP服务器

以太网帧的目的MAC地址为FF:FF:FF:FF:FF:FF，源MAC地址为00:16:D3:23:68:6A，这里就还需要思考MAC帧的格式及传输特点<4> DHCP请求包到达交换机后【考点：交换机】，在所有的出口广播入帧，这样也必定能有一帧能够发到DHCP服务器。

#### 3. DHCP服务器响应报文传输至本地主机

DHCP服务器通过报文的请求IP地址为255.255.255.255判定这是一个DHCP请求报文，因此交给上层协议处理。DHCP服务器确定要将地址68.85.2.101分配给本地主机，因此生成包含68.85.2.101和DNS服务器的IP地址、默认网关的IP地址、网络掩码的DHCP ACK报文，又经过与上述过程的层层封装发送给本地主机（这个时候不需广播，因为DHCP服务器已经能从DHCP请求报文提取出源主机的MAC地址）

#### 4. 本地主机对DHCP ACK报文的处理

本地主机从DHCP ACK报文中提取出上面的种种信息进行初始化配置。

## 二、DNS ARP

输入google服务器的URL，本地浏览器得先对URL进行解析【考点：URL知识点】<5> 这里最重要的是找到主机host name，之后又开启了一段新的旅程。

#### 1. 生成DNS查询报文

主机将要访问的host name放到DNS查询报文中<6>，使用UDP封装，下层目的IP地址填入68.87.71.226，源IP地址填入68.85.2.101。

2. ARP请求得到默认网关的MAC地址

生成一个目的IP地址为68.85.2.1的ARP查询报文，放置在一个广播帧中发送给交换机，交换机洪泛发送给所有设备，默认网关必然也会收到一个ARP请求包。

3. 默认网关ARP响应

默认网关从ARP请求报文中抽取源MAC地址作为目的MAC地址，并将自己的MAC地址发还给本地主机。

4. 将DNS查询报文发往默认网关

本地主机处理完ARP响应报文后 (主要是提取默认网关的MAC地址)，将DNS报文中的目的MAC字段填充上默认网关的MAC地址。现在这个DNS数据帧，目的IP地址是DNS服务器，目的MAC地址的默认网关的MAC地址。本地主机只需将这个DNS数据包发送出去，这个数据包便会经过交换机的转发来到默认网关。

三、IGP EGP 域内域间路由选择

1. 默认网关转发DNS查询报文

默认网关根据目的地址68.87.71.226根据转发表决定发给Comcast网络中最左边的边界路由器【考点：BGP协议】

2. 域内路由选择

Comcast网络边界路由器接收到数据帧之后，抽取IP数据包，提取出目的IP地址为68.87.71.226，根据域内路由选择协议发送给下一跳路由器，由下一跳路由器存储转发【考点：RIP协议、OSPF协议】

3. DNS查询过程

数据包到达DNS服务器之后，开始做分析和查询。这个时候得看DNS服务器采用的是迭代查询还是递归查询【考点：DNS查询】，总之无论经过哪种方式，DNS最终一定可以获得google服

务器的IP地址，封装在DNS响应报文中，发还回本地主机，本地主机就可以使用这个ip地址访问google服务器了。此外，本地主机会将这个<域名，ip地址>缓存在本地中，下次直接用ip发起访问就好。

四、TCP HTTP

下面相信对各位经历过面试的读者，已经非常熟悉了，从简。

1. 建立TCP连接

使用TCP三次握手建立一次持久连接。【考点：TCP三次握手机制】

2. 本地主机发送HTTP GET报文

主机浏览器生成包含要获取的 URL 的HTTP GET报文，写入套接字，成为TCP报文段的一个有效载荷。【考点：HTTP协议(重要!!!)】

3. google服务器发送HTTP RESPONSE报文

google服务器接收到HTTP GET报文，处理并生成一个HTTP响应报文，发还给本地主机。

4. 数据传输

WEB数据传输【考点：HTTP文档传输次数计算】

5. 断开TCP连接

通过TCP四次挥手结束TCP持久连接【考点：TCP四次挥手机制】

学习附录

- 1. IPv4地址分类及特征
- 2. IPv6原理、应用与实践
- 3. 网络协议篇之DHCP协议（一）—— DHCP协议基础
- 4. MAC帧格式，IP数据报格式和ARP数据报格式
- 5. URL格式

6. 超全总结！关于 DNS 看这一篇就够了

7. UDP协议格式及特点

## 参考文献

- 计算机网络 自顶向下方法
- 计算机网络 考研复习指导
- 计算机网络 第7版 (谢希仁)
- 计算机网络与下一代互联网

## 写在最后

这篇文章简要自顶向下漫游了计算机网络的协议栈，虽然涉及面非常广，但还是有挺多的细节是值得我们关注的。比如，数据包丢失时，还需要网际控制报文协议ICMP协助处理；网络拥塞时，还需要有传输层的拥塞控制机制；发送和接收速度不对等时，还需要有传输层的流量控制机制；数据链路层发送数据帧，有停等协议、ARQ协议、滑动窗口协议；保证数据的可靠性，有序号、确认、重传、校验，校验又有奇偶校验、CRC循环冗余校验、海明码；以及通信界最伟大的定理——香农定理等..... 面试中，我们一般常说诸如“服务器将数据发送给本地主机”这样的话，这里的“发送”包含了太多太多，我们一直站在传输层以上的角度学习计算机网络，其实这样是学得不深刻的，这也是我们今天总结回顾传输层底下的目的所在。作为计算机科班学生，学习计算机网络，除了理论课本之外，用WireShark抓取分析各种包、使用Chrome开发者工具分析HTTP请求和响应报文、多用一些网络命令行指令跟踪网络、玩一玩云服务器、捣弄捣弄路由器和交换机，这些方式是实践并深化计算机网络理论的最佳实践。关于网络抓包分析，将在下期推出。