链滴

www 开头的域名和 Cookie-free

作者：88250
原文链接：https://ld246.com/article/1445234634721
来源网站：链滴
许可协议：署名-相同方式共享 4.0 国际 (CC BY-SA 4.0)

本文我们会探索两个问题：

1. 网站是应该选择 www 开头的域名作为用户入口还是应该使用不带 www 的域名（裸域）？
2. Cookie-free 是什么？为什么很重要？

www 开头的域名

国内外大多数知名站点都是使用带 www 的二级域名作为用户入口，其中有通过 301 将裸域重定向到 www 二级域名的，也有通过 302 的（较少）。

为什么这些大站（流量大/子域名众多）都要使用带 www 的域名呢，主要原因应该有以下几点：

- 对用户来说识别度高，带了 www 后（即使不加 http://）用户就知道这是一个可以用浏览器打开 URL
- 避免将 Cookie 设置在裸域上（后面我们会讨论这样做的坏处）
- 在其他地方引用带 www 的 URL 更容易被识别（比如会自动加上链接）

以上主要是对大站来说，我们经常使用的小站（流量小/子域名较少）入口很多是不带 www 的（两都允许以及类似 Twitter 这样的个例除外），这样做原因应该主要只有一点：

- 域名更短，突出简约的个性，特别是一些个性后缀的域名不带 www 时更有利于用户记忆

301 or 302

无论我们选择带 www 的域名作为用户入口还是选择裸域，都需要做跳转。这个时候是选择使用 301 久跳转还是使用 302 临时跳转呢？

从大多数网站来看，**选择 301 永久跳转的居多**，主要原因是 SEO：

302重定向很容易被搜索引擎误认为是利用多个域名指向同一网站，那么你的网站就会被封掉，罪名“利用重复的内容来干扰Google搜索结果的网站排名”。因为302重定向经常被用于做url劫持，黑帽eo技术中，而且百度在处理302重定向技术还不成熟，经常将它纳入到黑帽seo的范畴中，而google这方面识别处理就完善了许多。所以302重定向在现阶段的搜索引擎技术中，还是容易导致网站降权，尽量不用。但从seo、网站优化方面来说是弊大于利。302 重定向 - 百度百科

Cookie 规则

我们先简单明确一下 Cookie 在浏览器上的规则，以域名 hacpai.com 为例：

- domain=.hacpai.com 时，Cookie 对裸域 hacpai.com 以及所有子域名 **.hacpai.com 是可用的
- domain=hacpai.com 时效果同上
- domain=hacpai.com 时对于非 hacpai.com 的其他域名无效
- **.hacpai.com 可以将 Cookie 设置在裸域上
- **.hacpai.com 不可以在其他子域上设置 Cookie

具体细节请参考RFC 6265。

遵守了以上规则后，我们来看看“从输入 URL 到页面加载完成的过程中都发生了什么事情？”

- 以计算机组成原理和网络通信的视角
- 以应用层协议和 Web 开发者的视角

我们这里的重点是返回 HTML 到浏览器后开始加载静态资源（js/css/images），此时浏览器会根据则带上 Cookie 发送这些请求。很显然，这样会**占用额外的网络带宽**，因为每个静态源的请求都会带上 Cookie 数据，特别是当 Cookie 较大的时候。

Cookie-free 域名

Cookie-free 域名指的是请求时不发送 Cookie 的域名，要解决前面我们提到的静态资源请求时带 Cookie 的方法就是使用 Cookie-free 域名来 serve 静态资源文件。

这个方法是最目前最普遍和有效的做法，大多数站点都是使用了和用户入口所在主域名完全独立的另个域名来提供静态资源：

- 无论用户入口是 hacpai.com 还是 www.hacpai.com，静态资源如果放在另一个域名上（比如 symphony-static.b3log.org），那么浏览器在请求这些静态资源时都不会带上 Cookie
- symphony-static.b3log.org 可以非常方便的做 CDN 加速，因为都是静态资源，不需要配置复杂匹配策略或重写规则

另外，Cookie 最好只设置在当前域名下，不要设置在裸域上，避免不必要的“污染”。当然，有的景下，“污染”是为了以最简单的方式实现某些特性，比如多个子站通过裸域会话追踪的 Cookie 实会话共享。

结论

- 尽量以 www 作为站的用户入口，避免将 Cookie 设置在裸域上
- 尽量考虑 Cookie-free 问题，使用其他域名进行静态资源 Serve 是最好选择

如果以上两点在站点初期都没有考虑好，那也不用担心。因为我们自己的站点一般达不到“大站”的级，但此时我们依然需要：

- 尽量安全地使用 Cookie，避免混乱和漏洞（比如子站 A 漏洞会影响子站 B）
- 做好 Cache 控制（200 cache），减少不必要请求的发送

参考

- 为什么越来越多的网站域名不加 ‘www’ 前缀?
- Best Practices for Speeding Up Your Web Site
- How do browser cookie domains work?
- 浅谈跨域WEB攻击