

wget命令8种实用用法



良许

2021-10-22 14:38

已关注



文章配图

大家好，我是良许。

wget 是一个可以从网络上下载文件的免费实用程序，它的工作原理是从 Internet 上获取数据，并将其保存到本地文件中或显示在你的终端上。

这实际上也是大家所使用的浏览器所做的，例如 Firefox 或 Chrome，其实在内部也是调用了 wget 程序进行数据下载。

本文介绍 8 个 wget 命令常见使用方式，希望对小伙伴们有所帮助。

1. 使用 wget 命令下载文件

你可以使用 `wget` 命令来下载指定链接的文件。默认情况下，下载的文件将保存到当前工作目录中的同名文件中。

```
$ wget http://www.linux.net--2021-09-20 17:23:47-- http://www.linux.n
```

如果你不想将下载的文件保存在本地，而只是想将其显示在标准输出（`stdout`）里，可以使用 `--output-document` 选项，同时后面再跟上一个 `-` 号。

```
$ wget http://www.linux.net --output-document - | head -n4<!doctype html
```

如果你想给下载的文件重新命名，那么可以使用 `--output-document` 选项（或者更简便，直接使用 `-O` 即可）：

```
$ wget http://www.linux.net --output-document newfile.html
```

2. 断点续传

如果你要下载的文件非常大，因为网络的原因有可能出现一次性无法下载完全的情况。如果每次都要重新下载，那么都不知道要等到猴年马月。

这种情况下，就可以使用 `--continue` 选项（或者 `-c`）实现断点续传。也就是说，如果因为各种原因导致下载中断，使用了这个选项，就可以继续上次的下载，而不需要重新进行下载。

```
$ wget --continue https://www.linux.net/linux-distro.iso
```

3. 下载一系列文件

如果你下载的不是一个大文件，而是需要很多个小文件，那么 `wget` 命令也可以帮你轻松实现。

但是，这里还需要使用一些 `bash` 语法来实现目的。一般来讲，这些文件的名称都有一定的规律，比如：`file_1.txt`，`file_2.txt`，`file_3.txt`，等等，那么你就需要使用这

样的命令：

```
$ wget http://www.linux.net/file_{1..4}.txt
```

4. 镜像整个站点

如果你想对某个网站进行整站下载，包括其目录结构，那么你就需要使用 `--mirror` 选项。

这个选项等同于 `--recursive --level inf --timestamping --no-remove-listing`，这意味着它是无限递归的，因此你可以下载到指定域上的所有内容。

如果你使用 `wget` 存档站点，那么这些选项 `--no-cookies --page-requisites --convert-links` 也可用于确保每个页面都是最新的、最完整的。

5. 修改 HTML 请求标头

学过网络通讯的小伙伴都知道，HTTP 数据包里包含了非常多的元素。其中，HTTP 标头是数据包初始的组成部分。

当你使用浏览器浏览网页时，你的浏览器会向服务器发送 HTTP 请求标头。具体发了些什么东西呢？可以使用 `--debug` 选项查看 `wget` 每次请求发送的标头信息：

```
$ wget --debug www.linux.net---request begin---GET / HTTP/1.1User-Agent
```



你可以使用 `--header` 选项修改请求标头。为什么要这么做呢？其实有很多使用场景的。例如，有时候为了测试，需要模拟特定的浏览器发出的请求。

比如，你想模拟 Edge 浏览器发出来的请求，可以这样操作：

```
$ wget --debug --header="User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64;
```



除此之外，你还可以伪装成特定的移动设备（比如 iPhone）：

```
$ wget --debug \--header = "User-Agent: Mozilla/5.0 (iPhone; CPU iPhone 0
```

6. 查看响应标头

与浏览器请求发送标头信息的方式相同，标头信息也包含在响应中。同样地，你可以使用 `--debug` 选项查看响应标头：

```
$ wget --debug www.linux.net[...]---response begin---HTTP/1.1 200 OKAcc
```

7. 响应 301 响应

熟悉网络协议的都知道，200 响应码意味着一切都按预期进行。而 301 响应则意味着 URL 已经指向不同的网站。

这种情况下，如果你需要下载文件，那么就需要使用 `wget` 的重定向功能。所以，如果你遇到 301 响应时，就需要使用 `--max-redirect` 选项。

如果你不想要重定向，那么可以将 `--max-redirect` 设置为 0。

```
$ wget --max-redirect 0 http://www.linux.net--2021-09-21 11:01:35-- htt
```

或者，你还可以将其设置为其他数字以控制 `wget` 跟随的重定向数量。

8. 展开短链接

有时候，我们需要将一个长链接转为短链接，例如在文本框里填写信息时，有时候文本框对字符长度有限制，这时短链就可以大大减少字符数。

除了使用第三方平台，其实我们可以直接使用 `wget` 命令来将短链接还原为长链接。这里依然使用 `--max-redirect` 选项：

```
$ wget --max-redirect 0 "https://bit.ly/2yDyS4T"--2021-09-21 11:32:04-- h
```

输出的倒数第二行，在 Location 部分，后面你就看到了短链展开之后的真面目。

[举报/反馈](#)