

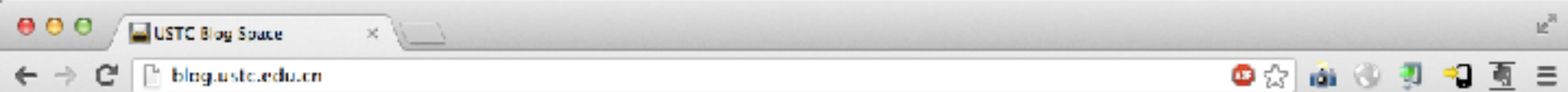


# HTTP 协议简介 - 上

VinsOn 2016.10.28

\*修改自USTC LUG





[Introduction](#) [Easy](#) [Reliable](#) [Help](#)

[Login](#)

# This is USTC.

USTC Cloud, solid as stone.

Your person blog data will be stored in USTC Cloud in secure.

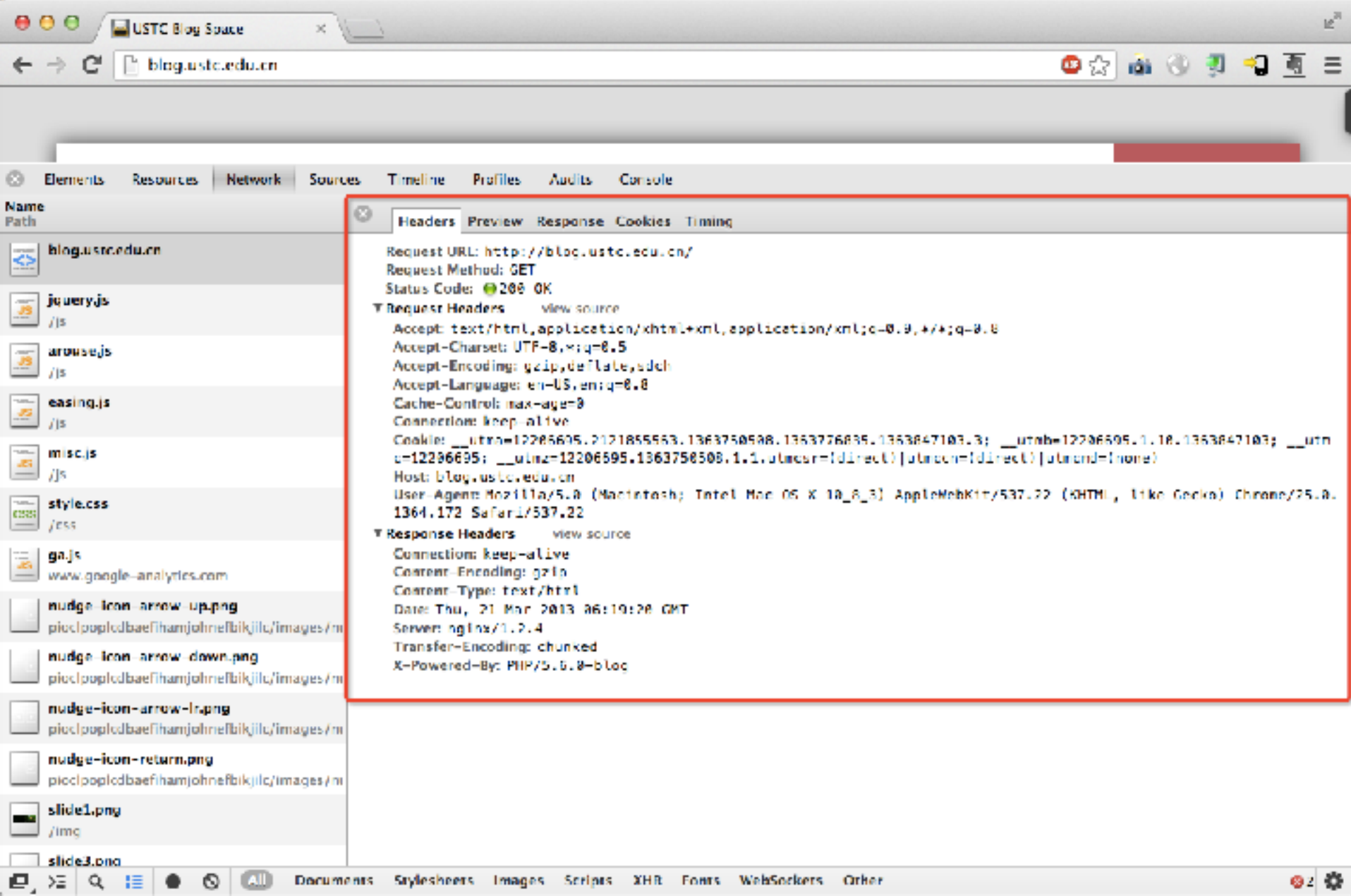
We promise your 99% blog uptime, and you will never worry about data loss because your person data will be backed up daily.

[Take the tour](#)

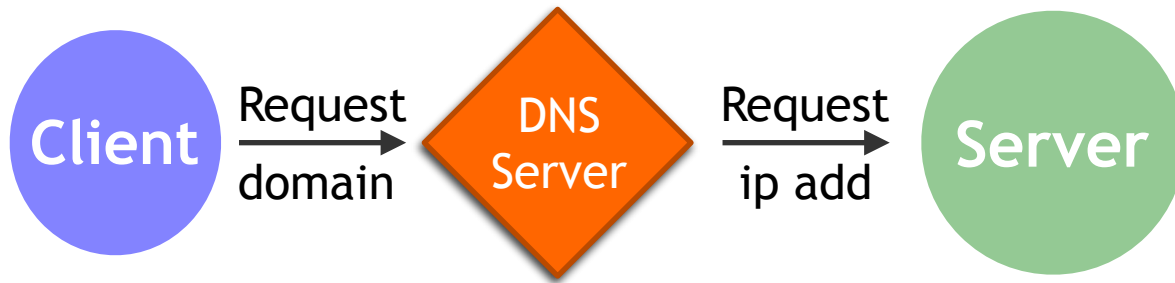


There are **416** users, contributing **lots of** blogs. Want to join us?

[Register now!](#)



# 一个最简单的HTTP请求过程



www.baidu.com->DNS->111.13.100.92

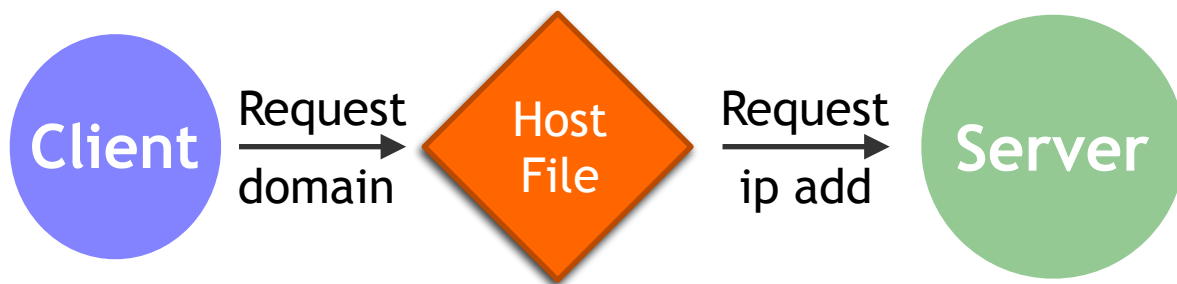
## DNS挟持与DNS污染

DNS挟持：运营商挟持、路由器挟持

DNS污染：通过关键词检测而返回虚假结果（GFW）

The feature of /etc/hosts ??

# 通过host文件自定义Domain to IP



www.baidu.com->Host->0.0.0.0

## DNS挟持与DNS污染

DNS挟持：运营商挟持、路由器挟持


DNS污染：通过关键词检测而返回虚假结果（GFW）

# 一个最简单的HTTP请求

GET / HTTP/1.1


Host: mirrors.ustc.edu.cn

␣



等等...

这个请求发给谁？



```
$ http 202.38.95.110
```

```
GET / HTTP/1.1
```

```
Host: mirrors.ustc.edu.cn
```

```
↵
```

```
HTTP/1.1 200 OK
```

```
Server: nginx/0.7.67
```

```
Date: Thu, 21 Mar 2013 06:54:58 GMT
```

```
Content-Type: text/html
```

```
Content-Length: 14457
```

```
Last-Modified: Thu, 21 Mar 2013 06:50:03  
GMT
```

```
Connection: keep-alive
```

```
Accept-Ranges: bytes
```

```
↵
```

```
{{body}}
```





```
(! 9370)-> nc 202.38.95.110 80
```

```
GET /status/ HTTP/1.1
```

```
Host: mirrors.ustc.edu.cn
```

Request

```
HTTP/1.1 200 OK
```

```
Server: nginx/0.7.67
```

```
Date: Thu, 21 Mar 2013 07:14:20 GMT
```

```
Content-Type: text/html
```

```
Content-Length: 5098
```

```
Last-Modified: Tue, 05 Mar 2013 09:51:02 GMT
```

```
Connection: keep-alive
```

```
Accept-Ranges: bytes
```

Response Header

```
<html>
```

```
  <head>
```

```
    <title>MIRRORS.USTC.EDU.CN Sync Status</title>
```

```
    <script type="text/javascript" src="jquery-1.6.2.min.js"></script>
```

```
    <script type="text/javascript" src="jquery.tablesorter.min.js"></script>
```

```
    <meta content="text/html; charset=utf-8" http-equiv="Content-Type" />
```

```
...
```


```
... This content has been removed to save space
```

```
...
```

```
  </body>
```

```
</html>
```

Response Body



(! 9370)-> nc 202.38.95.110 80

GET /status/ HTTP/1.1

Host: mirrors.ustc.edu.cn

HTTP/1.1 200 OK

Server: nginx/0.7.67

Date: Thu, 21 Mar 2013 07:14:20 GMT

Content-Type: text/html

Content-Length: 5098

Last-Modified: Tue, 05 Mar 2013 09:51:02 GMT

Connection: keep-alive

Accept-Ranges: bytes



```
<method> <url> <proto-ver>  
<header1>: <header1-value>  
<header2>: <header2-value>  
...  
⌋
```

```
<proto-ver> <resp-code> <resp-message>  
<header1>: <header1-value>  
<header2>: <header2-value>  
...  
⌋
```

```
[ body ]
```



# HTTP Method

HTTP1.0定义了三种请求方法： GET, POST 和 HEAD方法。

HTTP1.1新增了五种请求方法： OPTIONS, PUT, DELETE, TRACE 和 CONNECT 方法。

**GET**      请求指定的页面信息，并返回实体主体。

**POST**      向指定资源提交数据进行处理请求（例如提交表单或者上传文件）。数据被包含在请求体中。POST请求可能会导致新的资源的建立和/或已有资源的修改。

**PUT**      从客户端向服务器传送的数据取代指定的文档的内容。

**DELETE**   请求服务器删除指定的页面。

# 常见的HTTP Request Headers

- **Host:** 这是必选的，服务器需要靠这个头来区分Virtual Host，可以认为它的作用跟网络通信中的端口类似。
- **User-Agent:** 耍流氓用的，可“伪装成其他设备”。
- **Referer:** 表示当前请求的URL由哪个URL的内容发起。一般服务器端靠这个来做防盗链。**站点统计必用**
- **Range:** 断点续传、多线程下载全靠它了。
- **Accept-Encoding:** 网页接收的编码方式
- **Accept-Language:** 网页接收的语言
- **Cookie:** 小甜饼很废带宽。。。 (众所周知)
- **If-Modified-Since:** 这个东西节省了大量的带宽，大大提高了网页打开的速度
- **Connection-keep-alive:** 当出现对服务器的后继请求时，Keep-Alive功能避免了建立或者重新建立连接。



# 常见的HTTP Response Headers

- **Server:** 客户端有User-Agent, 服务器端就要有Server嘛
- **Date:** 服务器响应该请求的时间
- **Content-Type:** 返回内容的类型, 浏览器依此判断如何处理
- **Content-Length:**
- **Transfer-Encoding:chunked** 分块传输
- **Content-Encoding:(gzip, deflate)** 压缩传输
- **Content-Language:**
- **Cache-Control:** 与cache有关的头有很多

# User-Agent

- 1.判断Device，自适应网站UI
- 2.UA统计某网站访客用何种Device登录，作出适应调整
- 3.限制设备访问网站，例如微信内置浏览器



请在微信客户端打开链接

# 常见的HTTP Response Code含义

200

500

萌喵图

206

401

301

405

302

451-中国内地限定

304

2015年12月18日，此代码由国际网络工程研究团队通过  
HTTP 451 Unavailable For Legal Reasons

403

404

502

503





# HTTP调试神器之 curl

- H 附带request header
- I http response
- V http 通讯过程包括request

上面三个参数几乎足够了



感谢，期待下集！

<https://github.com/qducc>