

# 面试必问之HTTP协议

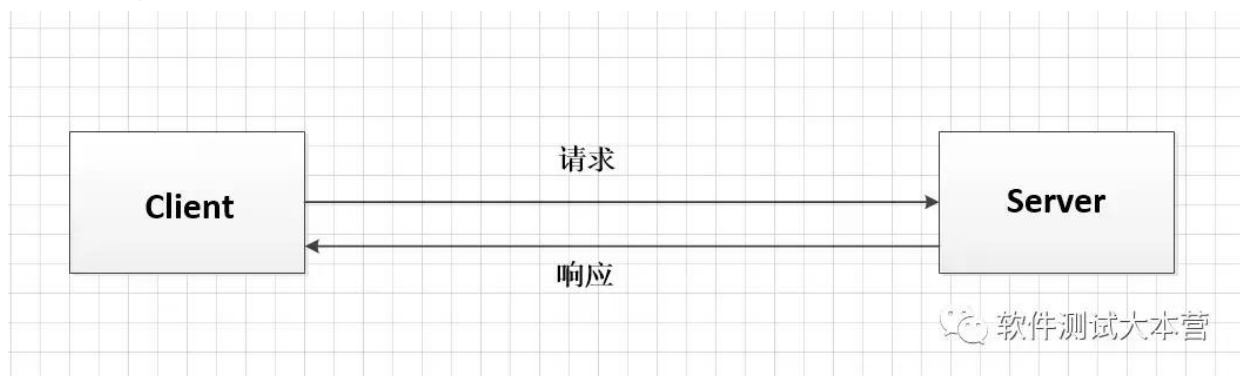
## HTTP协议是什么？

HTTP协议是超文本传输协议的缩写，英文是Hyper Text Transfer Protocol。是从万维网服务器传输超文本到本地浏览器的传送协议。

## 原理

HTTP是一个基于TCP/IP通信协议来传递数据（HTML 文件，图片文件，查询结果等）的请求应答协议。

HTTP协议工作于客户端-服务端架构为上。浏览器作为HTTP客户端通过URL向HTTP服务端即WEB服务器发送所有请求。Web服务器根据接收到的请求后，向客户端发送响应信息。



## HTTP特点：

- 1、简单快速：客户向服务器请求服务时，只需传送请求方法和路径。请求方法常用的有GET、HEAD、POST。
- 2、灵活：HTTP允许传输任意类型的数据对象。传输的类型由Content-Type加以标记。
3. 无连接：限制每次连接只处理一个请求。服务器处理完请求，并收到客户的应答后，即断开连接。采用这种方式可以节省传输时间。
4. 无状态：无状态是指协议对于事务处理没有记忆，后续处理需要前面的信息，则必须重传。

## http与https的区别？

HTTPS是HTTP协议的安全版本，HTTP协议的数据传输是明文的，是不安全的，HTTPS使用了SSL/TLS协议进行了加密处理。

常用的请求方法：

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。

HEAD:类似于get请求，只不过返回的响应中没有具体的内容，用于获取报头

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。

请求报文构成

http请求由请求行、消息报头、请求正文构成。

请求行以一个方法符号开头，以空格分开，后面跟着请求的URL和协议的版本。

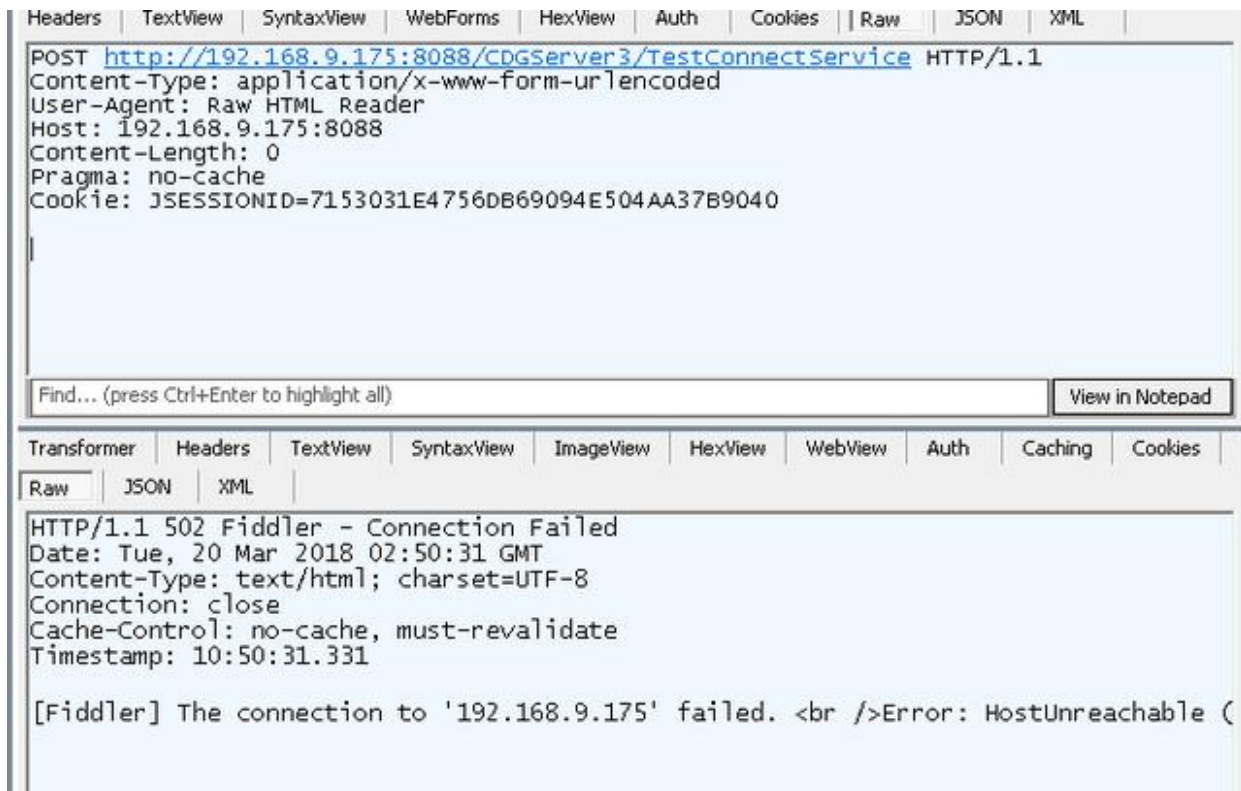
响应报文构成

HTTP响应也由状态行、消息报头、空行和响应正文构成。

**get请求**

```
GET https://www.baidu.com/s?ie=utf-8&csq=1&pstq=22&mod=2&isbd=1&cqid=920fb16000045f9e&istc=1030&ver
is_referer: https://www.baidu.com/
Accept: */*
is_xhr: 1
X-Requested-With: XMLHttpRequest
Referer: https://www.baidu.com/s?ie=utf-8&f=8&rsv_bp=0&rsv_idx=1&tn=baidu&wd=http%E5%8D%8F%E8%AE%AE
Accept-Language: zh-CN
Accept-Encoding: gzip, deflate
User-Agent: Mozilla/5.0 (Windows NT 6.1; WOW64; Trident/7.0; rv:11.0) like Gecko
Host: www.baidu.com
Connection: Keep-Alive
Cookie: BAIDUID=53B3E948288DEA6921575516A8E3C7BD; FG=1; BIDUPSID=53B3E948288DEA6921575516A8E3C7BD; P
```

**post请求及响应**



**post和get的区别：**都包含请求头请求行，**post**多了请求body。**get**多用来查询，请求参数放在url中，不会对服务器上的内容产生作用。**post**用来提交，如把账号密码放入body中。**GET**是直接添加到URL后面的，直接就可以在URL中看到内容，而**POST**是放在报文内部的，用户无法直接看到。**GET**提交的数据长度是有限制的，因为URL长度有限制，具体的长度限制视浏览器而定。而**POST**没有。

## 响应状态码

访问一个网页时，浏览器会向web服务器发出请求。此网页所在的服务器会返回一个包含HTTP状态码的信息头用以响应浏览器的请求。

## 状态码分类

- 1XX- 信息型，服务器收到请求，需要请求者继续操作。
- 2XX- 成功型，请求成功收到，理解并处理。
- 3XX - 重定向，需要进一步的操作以完成请求。
- 4XX - 客户端错误，请求包含语法错误或无法完成请求。
- 5XX - 服务器错误，服务器在处理请求的过程中发生了错误。

## 常见状态码

200 OK - 客户端请求成功  
301 - 资源（网页等）被永久转移到其它URL  
302 - 临时跳转  
400 Bad Request - 客户端请求有语法错误，不能被服务器所理解  
401 Unauthorized - 请求未经授权，这个状态代码必须和WWW-Authenticate报头域一起使用  
403 禁止访问  
404 - 请求资源不存在，可能是输入了错误的URL  
500 - 服务器内部发生了不可预期的错误  
503 Server Unavailable - 服务器当前不能处理客户端的请求，一段时间后可能恢复正常。

## http报头

请求报头由浏览器间接地设定，并跟在初始的GET和POST请求行之后发送，比如：

```
GET /servlet/Search?keywords=servlet+jsp HTTP/1.1
Accept: image/gif, image/jpg
Connection: Keep-Alive
Cookie: userID=id456578
Host: www.somebookstore.com
Referer: http://www.somebookstore.com/findbooks.html
User-Agent: Mozilla/4.0
```

对请求报头的访问，使得servlet能够执行许多优化，并提供大量特性

Accept: 指定浏览器或其他客户程序能够处理的MIME类型: `request.getHeader("Accept");`  
Accept-Charset: 使用的字符集，如ISO-8859-1  
Accept-Encoding: 客户端能够处理的编码类型，如gzip或compress  
Accept-Language: 客户端的首选语言  
Authorization: 客户用这个报头来标识自己的身份  
Connection: 标明客户是否能够处理持续性HTTP连接。持续性连接允许客户或者浏览器在单个socket中读取多个文件，从而节省协商几个独立连接所需的开销  
Content-Length: 只适用于POST请求，用来给定POST数据的大小，以字节为单位:  
`request.getContentLength`

Cookie: 向服务器返回cookie, 这些cookie是之前由服务器发送给浏览器的:

request.getCookies

Host: 标明原始URL中给出的主机名和端口号

If-Modified-Since: 仅当页面在指定日期之后发生改变的情况下, 客户程序才希望获取该页面。如果没有更新的结果, 则服务器发送304报头。这个选项十分有用, 因为使用它, 浏览器可以缓存文档, 只在它们发生改变时才通过网络重新载入它们

Referer: 标明引用Web页面的URL

User-Agent: 请求的浏览器或者其他客户程序

## 关于响应报头:

Web服务器对请求的响应, 一般由一个状态行, 一些响应报头, 一个空行和响应的文档组成

HTTP/1.1 200 OK

Content-Type: text/html

Header 2: ...

Header N: ...

(Blank Line)

<!DOCTYPE>

<HTML>

<HEAD>...</HEAD>

<BODY>...</BODY>

</HTML>

首先看HTTP状态代码:

200: OK

302: sendRedirect

404: sendError

## 关于响应报头的其他部分:

Allow: 指定服务器支持的请求方法 (GET, POST等)

Cache-Control: 告诉浏览器或者其他客户, 什么环境可以安全地缓存文档

Connection: close值, 指定浏览器不用使用持续性的HTTP连接

Content-Disposition: 要求浏览器询问客户, 将响应存储在磁盘上给定名称的文件中

Content-Encoding: 标明页面在传输过程中所使用的编码方式

Content-Language: 文档使用的语言

Content-Length: 响应中的字节数

Content-Type: MIME

Expires: 规定内容的过期时间, 从而不再需要继续缓存:

```
response.setHeader("Expires", Time)
```

Last-Modified: 标明文件最后的修改时间

Location: 300-399之间的所有响应都应该包括这个报头, 它通知浏览器文档的地址

Refresh: 标明浏览器应该多长时间之后请求最新的页面:

```
response.setIntHeader("Refresh", 30)
```

Set-Cookie: 指定一个同页面相关的cookie

有的细节还不是非常的清楚, 以后慢慢再了解

还有一点想说一下, 也是Core Servlets and JavaServer Pages中的内容, 对于执行耗时较长的计算任务的处理方法:

只需启动一个Thread, 将该线程的优先级设为一个较低的值, 这样才不致于影响整个服务器的运行, 当然需要不断使用

```
If(!isLastResult) {  
    response.setIntHeader("Refresh", 5)  
}
```

对于具体的运算类, 可以这样定义:

```
Thread t =new Thread(this);  
t.setPriority(Thread.MIN_PRIORITY);  
t.start();
```

这样做的好处是: 不让用户等待过长的时间

## MIME类型是什么? 包含哪些类型?

摘自: <http://www.tuidc.com/idczixun/newsx/newsidc/3479.html>

### **问: MIME类型是什么?**

答: MIME(Multipurpose Internet Mail Extensions)多用途互联网邮件扩展类型。是设定某种扩展名的文件用一种应用程序来打开的方式类型, 当该扩展名文件被访问的时候, 浏览器会自动使用指定应用程序来打开。

### **问: MIME类型有哪些?**

答: 常用Mime类型:

文件后缀	Mime类型	说明

.flv	flv/flv-flash	在线播放
.html或.htm	text/html	超文本标记语言文本
.rtf	application/rtf	RTF文本
.gif	image/gif	GIF图形
.jpeg或.jpg	image/jpeg	JPEG图形
.au	audio/basic	au声音文件
.mid或.midi	audio/midi或audio/x-midi	MIDI音乐文件
.ra或.ram或.rm	audio/x-pn-realaudio	RealAudio音乐文件
.mpg或.mpeg或.mp3	video/mpeg	MPEG文件
.avi	video/x-msvideo	AVI文件
.gz	application/x-gzip	GZIP文件
.tar	application/x-tar	TAR文件
.exe	application/octet-stream	下载文件类型
.rmvb	video/vnd.rn-realvideo	在线播放
.txt	text/plain	普通文本
.mrp	application/octet-stream	MRP文件（国内普遍的手机）
.ipa	application/iphone-package-archive	IPA文件(IPHONE)
.deb	application/x-debian-package-archive	DED文件(IPHONE)
.apk	application/vnd.android.package-archive	APK文件(安卓系统)
.cab	application/vnd.cab-com-archive	CAB文件(Windows Mobile)
.xap	application/x-silverlight-app	XAP文件(Windows Phone 7)
.sis	application/vnd.symbian.install-archive	SIS文件(symbian平台)
.jar	application/java-archive	JAR文件(JAVA平台手机通用格)
.jad	text/vnd.sun.j2me.app-descriptor	JAD文件(JAVA平台手机通用格)
.sisx	application/vnd.symbian.epoc/x-sisx-app	SISX文件(symbian平台)

