

RSS TAG 邮件列表

关于我们/About 广告合作/AD



帐号  ☐ 自动登录 找回密码

密码   骑士注册

技术 ◆ 学习 新闻 ◆ 快讯 观点 ◆ 热议 软件 ◆ 分享 论坛 投稿

Locez 新手指南： [下载 Linux »](#) [安装 Linux »](#) [安装软件 »](#) [基础命令 »](#)

请注册后再搜索 搜索

**七牛云，实践40万用户，针对七大行业推出一站式数据服务**

技术 ◆ 学习 查看内容

## 自己动手开发一个 Web 服务器（一）

2016-1-1 10:00 评论: 9 收藏: 4

来源：编程派 参考原文：<http://ruslanspivak.com/lsbaws-part1/>  
编译文章：<http://codingpy.com/article/build-a-simple-web-server-part-one/>  
文章地址：<https://linux.cn/article-6815-1.html>

作者：Ruslan  
译者：Flask

有一天，一位女士散步时经过一个工地，看见有三个工人在干活。她问第一个人，“你在做什么？”第一个人有点不高兴，吼道“难道你看不出来我在砌砖吗？”女士对这个答案并不满意，接着问第二个人他在做什么。第二个人回答道，“我正在建造一堵砖墙。”然后，他转向第一个人，说道：“嘿，你砌的砖已经超过墙高了。你得把最后一块砖拿下来。”女士对这个答案还是不满意，她接着问第三个人他在做什么。第三个人抬头看着天空，对她说：“我在建造这个世界上有史以来最大的教堂”。就在他望着天空出神的时候，另外两个人已经开始争吵多出的那块砖。他慢慢转向前两个人，说道：“兄弟们，别管那块砖了。这是一堵内墙，之后还会被刷上石灰的，没人会注意到这块砖。接着砌下层吧。”

这个故事的寓意在于，当你掌握了整个系统的设计，明白不同的组件是以何种方式组合在一起的（砖块，墙，教堂）时候，你就能够更快地发现并解决问题（多出的砖块）。



但是，这个故事与从头开发一个 Web 服务器有什么关系呢？

在我看来，要成为一名更优秀的程序员，你**必须**更好地理解自己日常使用的软件系统，而这就包括了编程语言、编译器、解释器、数据库与操作系统、Web 服务器和网络开发框架。而要想更好、更深刻地理解这些系统，你**必须**从头重新开发这些系统，一步一个脚印地重来一遍。

孔子曰：不闻不若闻之，闻之不若见之，见之不若知之，知之不若行之。

不闻不若闻之



### 热点评论

来自北京的 Chrome 47.0|Windows 8.1 用户 2016-1-2 21:43

单epoll的socket通讯就要写200行,加上http协议解析要1000左右,不是我说,这20行构建的web服务器太low了 [赞 2](#)

来自浙江的 Chrome 47.0|Windows 8.1 用户 2016-1-1 11:43

20行的代码说这么复杂,得有多新手才愿意看.....建议从c语言开始实现,才能体会到网络通信、HTTP报文解析、业务层实现、异步同步操作等细节 [赞 2](#)

来自上海的 Sina Weibo 5.7|Android 4.4 用户 2016-1-1 16:02

写的真好 [赞 2](#)

来自山东青岛的 Chrome Mobile iOS 47.0|iOS 8.1 用户 2016-1-2 02:06

你以为c语言实现会比这20行多多少么 [赞 1](#)

来自湖北鄂州的 Mobile Safari 7.0|iOS 7.1 用户 2016-1-2 08:24

嘿，把图纸藏起来 [赞 1](#)

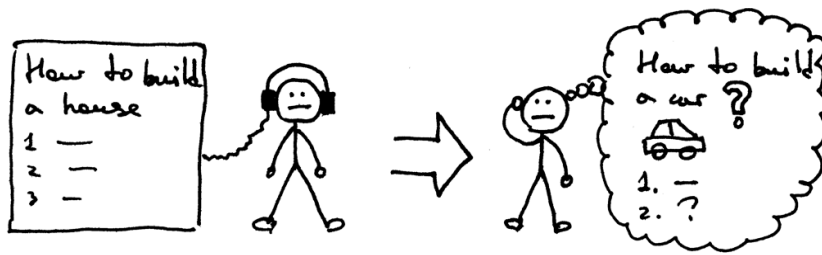
告别盲目投简历 让海量好机会主动来找你 [了解更多 >](#)

### 相关阅读

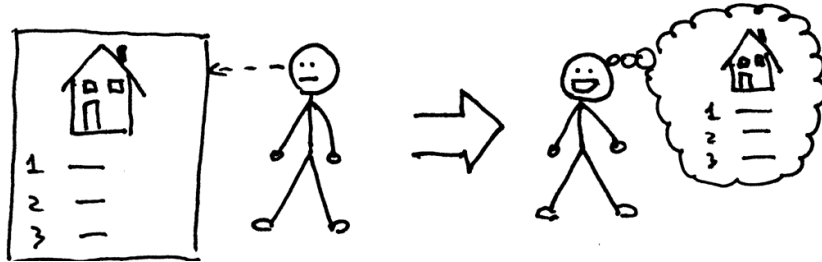
服务器	web
淘宝Web服务器, Tengine-1.2.3 正式发布	2012-3-1
4月全球Web服务器份额: Apache居首 Nginx	2012-4-16
5月全球Web服务器市场份额: Nginx升至	2012-5-18
淘宝web服务器Tengine-1.3.0 版本发布	2012-5-29

6月全球Web服务器市场份额：Apache升至 2012-6-14

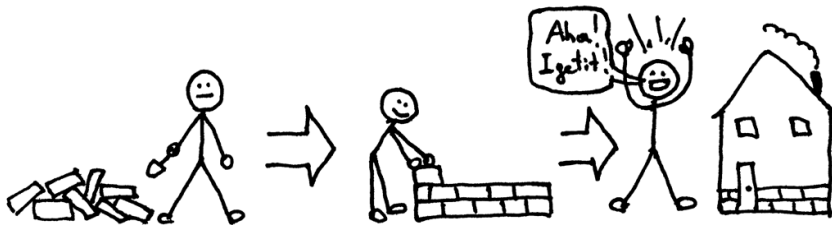
IETF正式开始开发HTTP 2.0标准 2012-10-12



闻之不若见之



见之不若知之，知之不若行之。

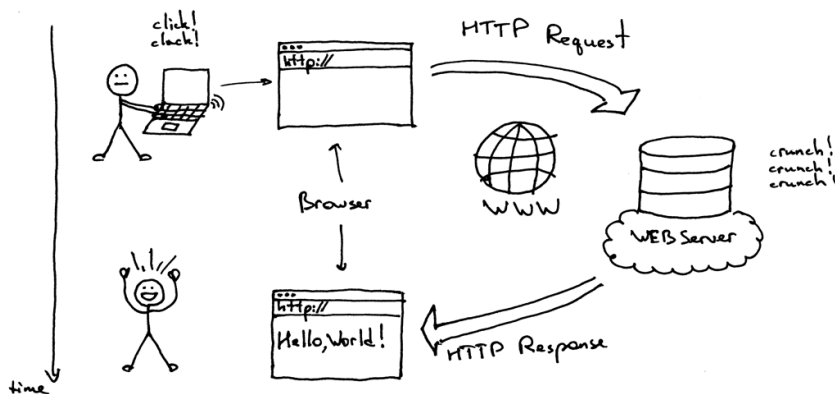


译者注：上面原作者所引用的那段话在国外的翻译是：I hear and I forget, I see and I remember, I do and I understand。外国人普遍认为出自孔子，但在查找这句英文的出处时，查到有篇博文 [http://blog.sina.com.cn/s/blog\\_60ebcd1d0100f4tv.html](http://blog.sina.com.cn/s/blog_60ebcd1d0100f4tv.html) 称这句话的中文实际出自荀子的《儒效篇》，经查确实如此。

我希望你读到这里的时候，已经认可了通过重新开发不同软件系统来学习其原理这种方式。

《自己动手开发 Web 服务器》会分为三个部分，将介绍如何从头开发一个简易 Web 服务器。我们这就开始吧。

首先，到底什么是 Web 服务器？



简而言之，它是在物理服务器上搭建的一个网络连接服务器（networking server），永久地等待客户端发送请求。当服务器收到请求之后，它会生成响应并将其返回至客户端。客户端与服务器之间的通信，是以HTTP协议进行的。客户端可以是

浏览器，也可以是任何支持HTTP协议的软件。

那么，Web 服务器的简单实现形式会是怎样的呢？下面是我对此的理解。示例代码使用Python语言实现，不过即使你不懂Python语言，你应该也可以从代码和下面的解释中理解相关的概念：

```
import socket

HOST, PORT = '', 8888

listen_socket = socket.socket(socket.AF_INET, socket.SOCK_STREAM)
listen_socket.setsockopt(socket.SOL_SOCKET, socket.SO_REUSEADDR, 1)
listen_socket.bind((HOST, PORT))
listen_socket.listen(1)
print 'Serving HTTP on port %s ...' % PORT
while True:
    client_connection, client_address = listen_socket.accept()
    request = client_connection.recv(1024)
    print request

    http_response = """\
    HTTP/1.1 200 OK

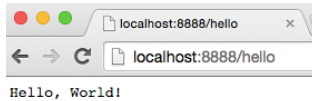
    Hello, World!
    """

    client_connection.sendall(http_response)
    client_connection.close()
```

将上面的代码保存为 `webserver1.py`，或者直接从我的 [Github仓库](https://github.com/rspivak/lbaweb/blob/master/part1/webserver1.py) <<https://github.com/rspivak/lbaweb/blob/master/part1/webserver1.py>> 下载，然后通过命令行运行该文件：

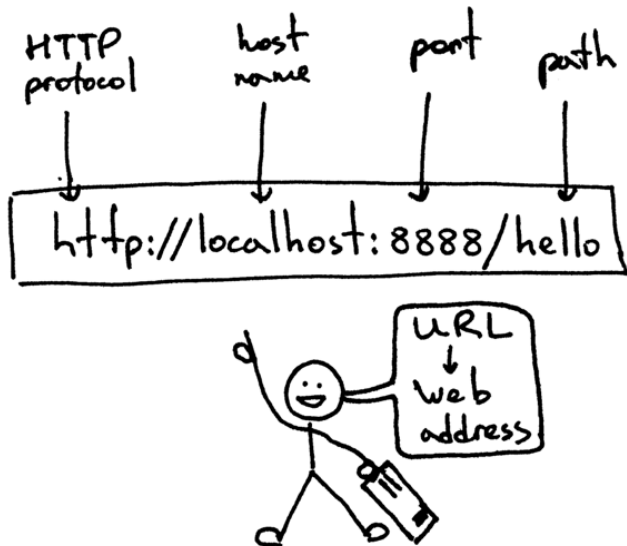
```
$ python webserver1.py
Serving HTTP on port 8888 ...
```

接下来，在浏览器的地址栏输入这个链接：<http://localhost:8888/hello>，然后按下回车键，你就会看见神奇的一幕。在浏览器中，应该会出现“Hello, World!”这句话：



是不是很神奇？接下来，我们来分析背后的实现原理。

首先，我们来看你所输入的网络地址。它的名字叫URL（<sup>Uniform Resource Locator</sup>统一资源定位符），其基本结构如下：



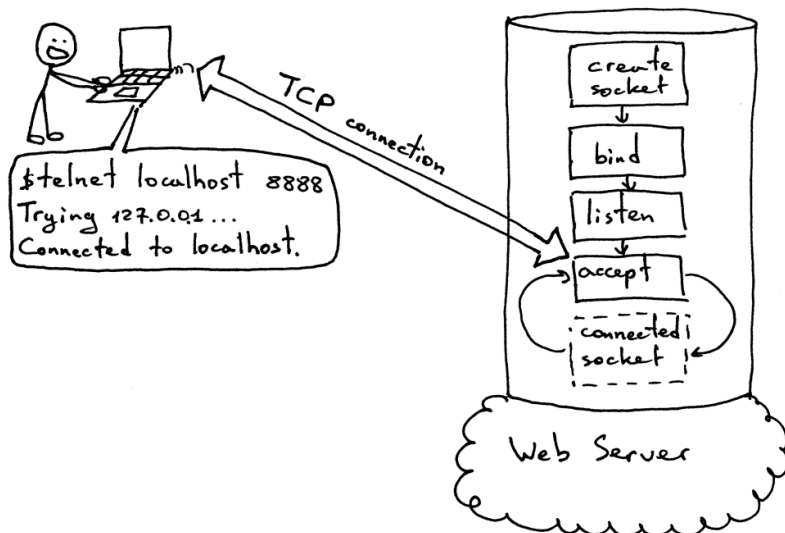
通过URL, 你告诉了浏览器它所需要发现并连接的 Web 服务器地址, 以及获取服务器上的页面路径。不过在浏览器发送 HTTP 请求之前, 它首先要与目标 Web 服务器建立 TCP 连接。然后, 浏览器再通过 TCP 连接发送 HTTP 请求至服务器, 并等待服务器返回 HTTP 响应。当浏览器收到响应的时候, 就会在页面上显示响应的内容, 而在上面的例子中, 浏览器显示的就是 “Hello, World!” 这句话。

那么, 在客户端发送请求、服务器返回响应之前, 二者究竟是如何建立起 TCP 连接的呢? 要建立起 TCP 连接, 服务器和客户端都使用了所谓的套接字 (socket)。接下来, 我们不直接使用浏览器, 而是在命令行使用 `telnet` 手动模拟浏览器。

在运行 Web 服务器的同一台电脑商, 通过命令行开启一次 `telnet` 会话, 将需要连接的主机设置为 `localhost`, 主机的连接端口设置为 `8888`, 然后按回车键:

```
$ telnet localhost 8888
Trying 127.0.0.1 ...
Connected to localhost.
```

完成这些操作之后, 你其实已经与本地运行的 Web 服务器建立了 TCP 连接, 随时可以发送和接收 HTTP 信息。在下面这张图片里, 展示的是服务器接受新 TCP 连接所需要完成的标准流程。

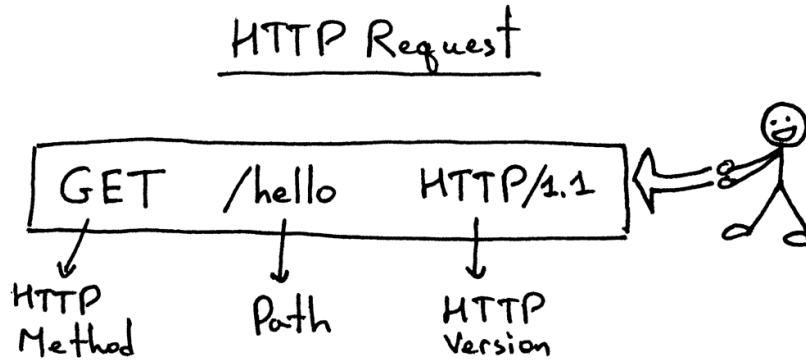


在上面那个 `telnet` 会话中, 我们输入 `GET /hello HTTP/1.1`, 然后按下回车:

```
$ telnet localhost 8888
Trying 127.0.0.1 ...
```

```
Connected to localhost.  
GET /hello HTTP/1.1  
  
HTTP/1.1 200 OK  
Hello, World!
```

你成功地手动模拟了浏览器！你手动发送了一条HTTP请求，然后收到了HTTP响应。下面这幅图展示的是HTTP请求的基本结构：

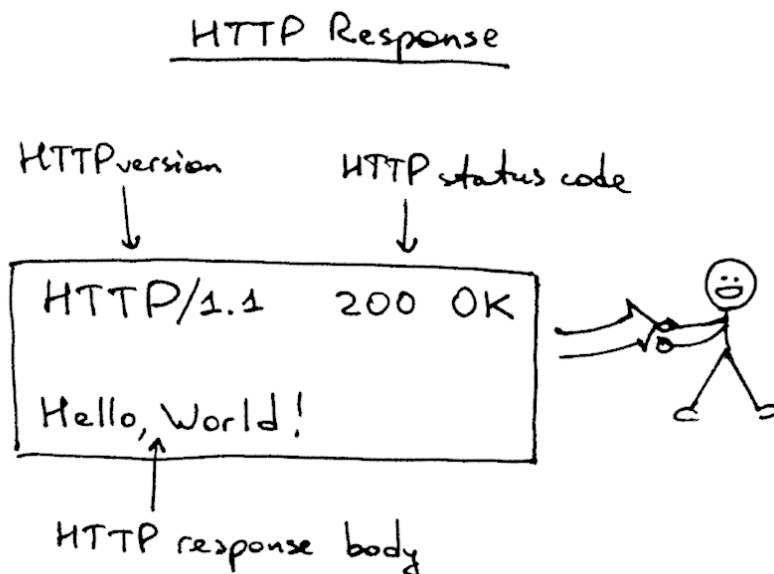


HTTP请求行包括了HTTP方法（这里使用的是 `GET` 方法，因为我们希望从服务器获取内容），服务器页面路径（`/hello`）以及HTTP协议的版本。

为了尽量简化，我们目前实现的 Web 服务器并不会解析上面的请求，你完全可以输入一些没有任何意义的代码，也一样可以收到“Hello, World!”响应。

在你输入请求代码并按下回车键之后，客户端就将该请求发送至服务器了，服务器则会解析你发送的请求，并返回相应的HTTP响应。

下面这张图显示的是服务器返回至客户端的HTTP响应详情：



我们来分析一下。响应中包含了状态行 `HTTP/1.1 200 OK`，之后是必须的空行，然后是HTTP响应的正文。

响应的状态行 `HTTP/1.1 200 OK` 中，包含了HTTP版本、HTTP状态码以及与状态码相对应的原因短语（Reason Phrase）。浏览器收到响应之后，会显示响应的正文，这就是为什么你会在浏览器中看到“Hello, World!”这句话。

这就是 Web 服务器基本的工作原理了。简单回顾一下：Web 服务器首先创建一个侦听套接字（listening socket），并开启一个永续循环接收新连接；客户端启动一个与服务器的TCP连接，成功建立连接之后，向服务器发送HTTP请求，之后服务器返回HTTP响应。要建立TCP连接，客户端和服务端都使用了套接字。

现在，你已经拥有了一个基本可用的简易 Web 服务器，你可以使用浏览器或其他HTTP客户端进行测试。正如上文所展示

的，通过 `telnet` 命令并手动输入HTTP请求，你自己也可以成为一个HTTP客户端。

下面给大家布置一道思考题：如何在不对服务器代码作任何修改的情况下，通过该服务器运行Djando应用、Flask应用和Pyramid应用，同时满足这些不同网络框架的要求？

答案将在《自己动手开发 Web 服务器》系列文章的第二部分揭晓。

来源：编程派 参考原文：<http://ruslanspivak.com/lsbaws-part1/>  
编译文章：<http://codingpy.com/article/build-a-simple-web-server-part-one/>

作者：Ruslan  
译者：Flask

本文为转载，如需再次转载，请查看源站“编程派”的要求。如果我们的工作有侵犯到您的权益，请及时联系我们。  
文章仅代表作者的知识和看法，如有不同观点，请楼下排队吐槽 :D

上一篇：[黑客利用 Wi-Fi 攻击你的七种方法](#)

发表评论

评论

最新评论

我也要发表评论

来自四川绵阳的 Chrome 47.0|Mac 10.9 用户 2016-1-3 12:58

其实我认为解析HTTP数据包就已经够繁琐了吧

赞 回复

来自北京的 Chrome 47.0|Windows 8.1 用户 2016-1-2 21:43

来自浙江的 Chrome 47.0|Windows 8.1 用户 发表于 2016-1-1 11:43 的评论：  
20行的代码说这么复杂，得有多新手才愿意看.....  
建议从c语言开始实现，才能体会到网络通信、HTTP报文解析、业务层实现、异步同步操作等细节

赞 2 回复

来自山东青岛的 Chrome Mobile iOS 47.0|iOS 8.1 用户 发表于 2016-1-2 02:06 的评论：

你以为c语言实现会比这20行多多么

2

	赞 1 回复
单epoll的socket通讯就要写200行,加上http协议解析要1000左右，不是我说，这20行构建的web服务器太low了	赞 2 回复
来自湖北鄂州的 Mobile Safari 7.0 iOS 7.1 用户 2016-1-2 08:24	
嘿，把图纸藏起来	赞 1 回复
来自山东青岛的 Chrome Mobile iOS 47.0 iOS 8.1 用户 2016-1-2 02:06	
来自浙江的 Chrome 47.0 Windows 8.1 用户 发表于 2016-1-1 11:43 的评论： 20行的代码说这么复杂，得有多新手才愿意看..... 建议从c语言开始实现，才能体会到网络通信、HTTP报文解析、业务层实现、异步同步操作等细节	1
	赞 2 回复
你以为c语言实现会比这20行多多少么	赞 1 回复
linux [Chrome 47.0 Mac 10.11] 2016-1-1 23:27	
darKnight [Firefox 43.0 Ubuntu] 发表于 2016-1-1 19:42 的评论： 能看到我的ip么？	1
	赞 回复
不能	赞 回复
darKnight [Firefox 43.0 Ubuntu] 2016-1-1 19:42	
能看到我的ip么？	赞 回复
来自上海的 Sina Weibo 5.7 Android 4.4 用户 2016-1-1 16:02	
写的真好	赞 2 回复
来自河南郑州的 Iceweasel 38.2 GNU/Linux 用户 2016-1-1 12:19	
看看	赞 1 回复
来自浙江的 Chrome 47.0 Windows 8.1 用户 2016-1-1 11:43	
20行的代码说这么复杂，得有多新手才愿意看..... 建议从c语言开始实现，才能体会到网络通信、HTTP报文解析、业务层实现、异步同步操作等细节	赞 2 回复