

湖南大学

HUNAN UNIVERSITY

云计算技术 性能报告

题 目:	Your Own HTTP Server
	Lab 2
组长姓名:	龙思宇-201708010823
学生姓名:	刘梦迪-201708010826
学生姓名:	高文静-201708010827
学生姓名:	吕 喆-201708010830
完成日期:	2020/4/27



云计算技术性能报告

目 录

一、 实验概要.....	2
二、 性能测试.....	3



一、实验概要

超文本传输协议（HTTP）是当今 Internet 上最常用的应用程序协议。像许多网络协议一样，HTTP 使用客户端-服务器模型。HTTP 客户端打开与 HTTP 服务器的网络连接，并发送 HTTP 请求消息。HTTP 消息是简单的格式化数据块。所有 HTTP 消息分为两种：请求消息和响应消息。请求消息请求来自 Web 服务器的操作。响应消息将请求的结果返回给客户端。请求和响应消息都具有相同的基本消息结构。

库网址：<https://github.com/longsiyu-1999/CloudComputingLabs/tree/master/Lab2>

最终实验结果路径：CloudComputingLabs/Lab2/

1.1 HTTP 服务器功能说明

每个 TCP 连接只能同时发送一个请求。客户端等待响应，当客户端获得响应时，也许将 TCP 连接重新用于新请求（或使用新的 TCP 连接）。这也是普通 HTTP 服务器支持的内容。

- a) 处理 HTTP GET 请求
- b) 处理 HTTP POST 请求
- c) 其他请求不处理

1.2 编译背景

Socket C++

1.3 使多线程来提高并发性

多线程设计：

每连接一个客户端，就创建一个线程解决所有该客户端发出的请求。在创建线程前上锁，进入线程回调函数后解锁。

1.4 指定参数

程序应启用长选项以接受参数并在启动期间指定这些参数。它们是--ip，--port。

- 1. --ip -指定服务器 IP 地址。
- 2. --port -选择 HTTP 服务器侦听传入连接的端口。

如果未指定端口号，则默认为端口 9999。（实验指导书的默认端口 80 无法使用）

1.5 运行 HTTP 服务器

创建套接字 sockfd，并绑定 bind 套接字，监听 listen 套接字。循环接受不同的客户端，每连接一个客户端则创建一个线程，客户端关闭则线程关闭。最后关闭监听的套接字。

运行方法：./httpserver（默认 ip 为 127.0.0.1，port 为 9999）

若要更改 ip 和 port，则可指定参数更改。

1.6 请求处理大致思路

1. 处理 GET 请求

判断是报文是 GET 请求：请求路径与 html 页面文件相对应→回复 200 OK 和文件的完整内容；否则回复 404 Not Found。

2. 处理 POST 请求



云计算技术性能报告

判断报文是 POST 请求：判断 URL 为 /Post_show，并且键为 Name 和 ID→则回复 200 OK 和 Name-ID 对；否则回复 404 Not Found。

3. 其他方法

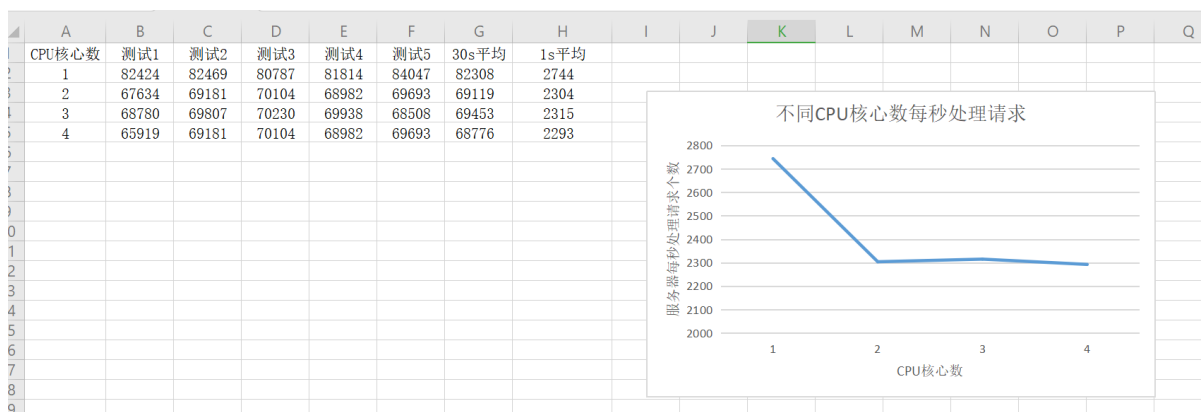
回复 501 Not Implemented。

1.7 实验环境

Linu 内核版本为 3.13.0-32-generic; 2GB 内存; CPU 型号为 Intel(R) Core(TM) i5-7200U CPU @ 2.5GHz 2.7GHz，共有 1 个核心 CPU（CPU 核心数在性能测试 1 有改动）；不使用超线程技术。

二、性能测试

1. 更改服务器 CPU 内核数，测试服务器每秒可以处理多少个 HTTP 请求。



测试两个客户端同时发送请求到不同核心数的服务器的情况。

手动设置虚拟机 CPU 核心数，对每一个环境测试 30s 内服务器可以处理的请求数，共测试五组，取平均值，再计算 1s 内服务器可以处理的请求数。

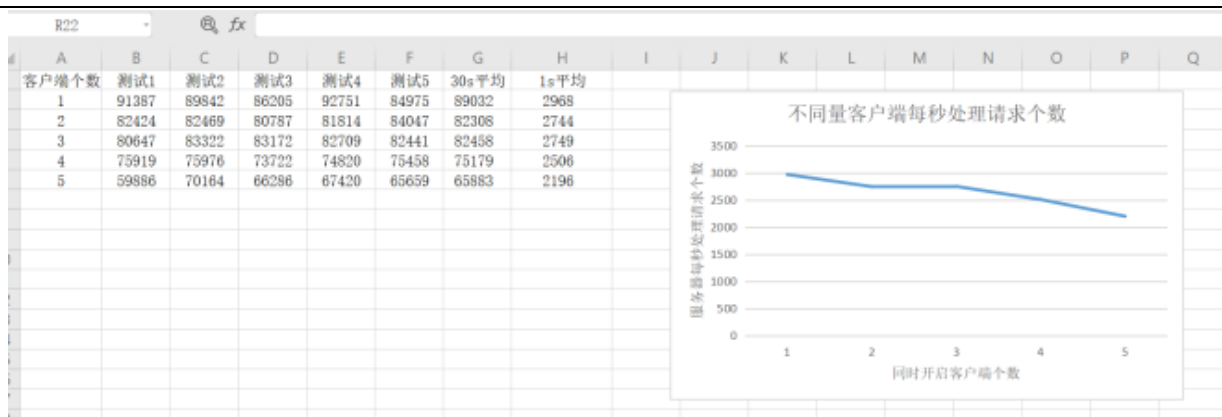
如上图，当 CPU 核心数为 1 的时候，每秒处理的请求数最多，核心数增加后所处理请求数明显减少，测试主机最多只支持 4 核 CPU 虚拟机。

但是理论上来说，应该是 CPU 内核数越多，处理请求的速度会越快。但是实际的测试结果却是 CPU 核心数为 1 的时候，处理请求速度越快，测试了多次都是一样的结果。考虑有可能是线程数量开的比较少，使多开的 CPU 内核可能利用率还没有单线程的轮转调度高，导致速度反而下降了。可能有我们多线程设计的问题，我们在这次测试只开启了两个客户端，所以只有两个线程。

2. 通过更改同时发送请求到服务器的并发客户端数，测试服务器每秒可以处理多少个 HTTP 请求。



云计算技术性能报告



手动编写客户端，在虚拟机 CPU 核心数为 1 的情况下测试服务器处理不同客户端发送请求的情况。

同样是测试 30s 服务器处理请求个数，同时开启 1 到 5 个客户端，保证客户端开启时间相差不超过 1s，测试 5 组取平均值，最终计算 1s 内服务器可以处理的请求个数。

如上图，当只有一个客户端发送请求时服务器处理请求数最多，两到三个客户端发送请求时处理稍慢，四到五个客户端时请求数量急剧减少。

这个测试结果就比较贴近理论了，对于单核 CPU 来说，客户端越多即线程越多，增加的上下文切换越多，使最终处理请求个数的速度越慢。