综合课程设计报告

姓名：万笛文 学号：2013060102008 班级：计科二班

# 项目名称

Web服务器的设计与实现。

# 项目主要任务

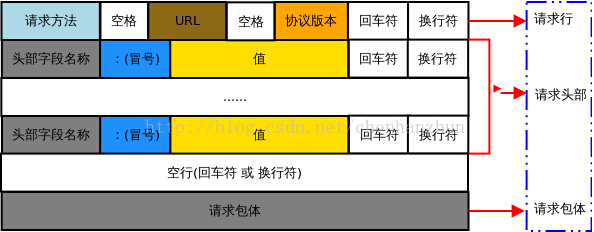
实现一简单的Web服务器，并实现相关的请求处理机制

# 项目详细功能描述

1. 学习理解HTTP 协议原理及协议格式
2. 设计Web服务器
3. 开发多线程服务器通信程序，可接收远程数据，每个线程对应一个TCP链接；
4. 定义HTTP请求/响应数据传输格式；
5. 开发数据解析模块；
6. 开发与HTTP的GET和POST等方法对应的服务器处理程序；
7. 根据请求报文处理完成后，发送响应报文到客户端，若服务器出错也发送相应错误响应。
8. 利用Web浏览器测试服务器，并可在线演示。

# HTTP协议原理及协议格式

1. 目前最广泛应用的HTTP协议版本是HTTP/1.1[RFC 2616]，因此基于文档RFC 2616来学习HTTP协议的原理及协议格式。
2. 总体操作：HTTP协议是一种请求/响应型的协议。客户端向服务器发出HTTP请求报文，服务器接收到请求报文并发出相应的响应报文进行响应。
3. HTTP请求报文格式



1. 协议版本HTTP-Version

协议版本的格式是HTTP/<major>.<minor>, 如：HTTP/1.1, 指定协议的版本号。在本次设计的Web服务器，实现的是HTTP/1.1。

1. 请求方法Method

请求方法指明了在被Request-URL指定的资源上执行的方法。常见的请求方法有：（大小写敏感的）

* 1. GET: 获取被请求URL指定的信息（以实体的格式），如果包含了If-Modified-Since,If-Unmodified-Since或If-Range等请求头就成了条件GET
  2. HEAD：同GET方法一致，除了服务器不能再响应里返回消息主体。
  3. POST：被用于请求源服务器接受请求中的实体作为请求资源的一个新的从属物。POST 方法的实际功能是由服务器决定的，并且经常依赖于请求URL。
  4. PUT：PUT 方法请求服务器去把请求里的实体存储在请求 URI标识下。
  5. DELETE：DELETE 方法请求源服务器删除请求 URI 指定的资源。
  6. 其他方法如：TRACT，CONNECT， OPTIONS（选项）等。

1. 其他请求头域
   1. Accept：被用于指定哪些媒体类型的响应对请求端是可接受的。ccept 头域被用于指明请求只对某些期望的媒体类型有效。
   2. Accept-Charset: 用来指名哪些字符集的响应对请求端是可接受的。
   3. Accept-Encoding: Accept-Encoding 是限定服务器返回给客户端可以接受的内容编码
   4. Accept-Language: 返回给客户端喜爱的自然语言。
2. 请求URL及HOST请求头

请求URL一般有两种格式：

一种格式为绝对URL，即"http:" "//" host [ ":" port ] [ abs\_path [ "?" query ]]的格式，例：

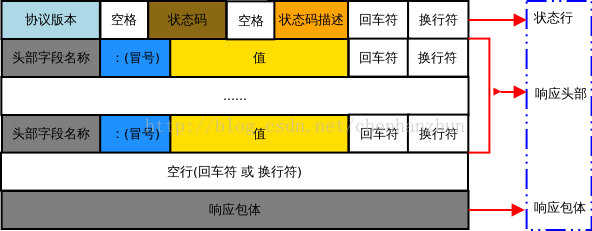
GET http://www.w3.org/pub/www/TheProject.html HTTP/1.1

另一种格式：URL为绝对路径abs\_path,以”/”（服务器的根）开头。在HOST请求头中指出URL的网络地址，例：

GET /pub/www/TheProject.html HTTP/1.1

HOST: www.w3.org

1. HTTP响应报文格式



1. 状态码与原因短语

状态码（Status-Code）是三位数字整数码，原因短语（Reason-Phrase）是关于状态码的简单文本描述。

常用状态码如下

* 200 OK ：成功 客户端的请求已经成功
* 204 No Content ：无内容 服务器已经满足了请求但并没有返回一个实体而是更新的元信息
* 205 Reset Content ： 重置内容 服务器告诉用户代理应该重置引起请求被发送的文档视图。此响应主要的目的是清空文档视图表单里的输入框以便用户能输入其它信息。此响应不能包含一个实体。
* 301 Moved Permanently ：永久移动 请求资源被赋予了一个新的永久的URL。
* 304 Not Modified :没有改变 条件GET时，所请求的资源没有改变。需要包含Date等头域。
* 400 Bad Request : 坏请求 请求不能被服务器理解。
* 403 Forbidden : 禁用 服务器理解此请求，但拒绝满足此请求。
* 404 Not Found : 没有找到 服务器没有找到所请求的资源
* 500 Internal Server Error : 服务器内部错误
* 503 Service Unavailable ： 服务不能获得 由于服务器暂时地过载或维护，服务器不能处理请求。

1. 响应头域
   1. Date: 表明产生消息的日期和时间，格式：wkday "," SP DD SP month SP YYYY SP HH:MM:SS SP "GMT" , wkday是星期的前三位，month是月份的前三位

例：Date:Tue,15 Nov 1994 08:12:31GMT

1. 实体头域
   1. Content-Encoding:是对媒体类型的修饰。当此头域出现时，其值表明对实体主体采用了何种内容编码，从而可以知道采用何种解码机制以获取 Content-Type 头域中指出的媒体类型。 Content-Encoding 头域主要目的是可以在不丢失下层媒体类型的标识下对文档进行压缩。
   2. Content-Language: 描述了实体面向用户的自然语言
   3. Content-Length: 用于指明发送给接收者实体主体（entity-bidy）的大小（以十进制的字节数表示），或是在使用 HEAD 方法时，指明实体主体本应在 GET 方法时发送实体主体的大小。
   4. Content-Type: 指明发给接收者的实体主体的媒体类型，或在 HEAD 方法中指明若请求为 GET 时将发送的媒体类型。

例：Content-Type: text/html; charset=ISO-8859-4

# 程序框图

开启服务器

开始监听

TCP连接

打开一个线程

接收请求报文

解析请求报文

发送响应报文