Telecooperation: Internet Lab.
3rd Assignment – Caching Web Proxy
Fernando Ruiz Vera - 1177767 - fruizvera@yahoo.com
You Yue - 1282746 - you@stud.tu-darmstadt.de

1 GENERAL DESCRIPTION OF THE SYSTEM.

The class *VBProxyServer* corresponds to the Graphic User Interface (GUI). This class is responsible for starting the Proxy Server on a port number specified by the user. Addtionally, the graphic user interface allows specifying the cache directory where the files will be located. As another additional feature, the request and response headers are shown in separated textareas. To store the path where headers and their corresponding body contents are located, two HashTables with identical key are provided.

In a separated Thread runs the class *ProxyServer* whose responsibility is listen to incoming request and attend each request into a separated Thread. The HttpRequest class is in charge of this task. A different *HttpCommand* is created according to the request method (GET or POST), either *GetCommand* or *PostCommand* respectively.

Both command inherit from the *HttpCommand* interface the *Execute()* method. The *GetCommand* verify if the requested URL is already stored at the cache. In positive case, the response message is re-constructed here and sends to the Browser. In case that the requested URL has not been already cached, the request is sent to the corresponding web server and an instance of *HttpResponse* is created that will wait, process and storage the response message. For the case of *PostCommand*, the instance of *HttpResponse* waits for the response and simply sends to the client browser.

The *HttpResponse* class provides three public methods. *SetCacheable()* to decide when a response is cacheable, in the case of this implementation any response to a GET request. A second method *setRequestPath()* that work as interface to establish the URL to be cached. At last but not least, *processResponse()* is responsible to wait for the answer, sent to the client browser and according to the class attribute *bCacheable* store or not this response.

The private method *storeURL()* at first calculate the hashcode of the requested URL, this one will be the key for two Hash Tables. One hash table is used to stored address the headers files and the other one for the content body files. The key coincide in both tables, whereas the value is calculated as follows:

```
Random rndIndex = new Random();
String sURL =  serverSocket.getInetAddress().getHostName() + httpPath;
hash = sURL.hashCode();

UID = System.currentTimeMillis();
index = rndIndex.nextInt(128);

fBody = new File(strRoot + "\\B" + UID + index);
fHeader = new File(strRoot + "\\H" + UID + index);
```

With the use of *Random.nextInt(128)* is it tried to eliminate the possible coincidence in two subsequent requests.

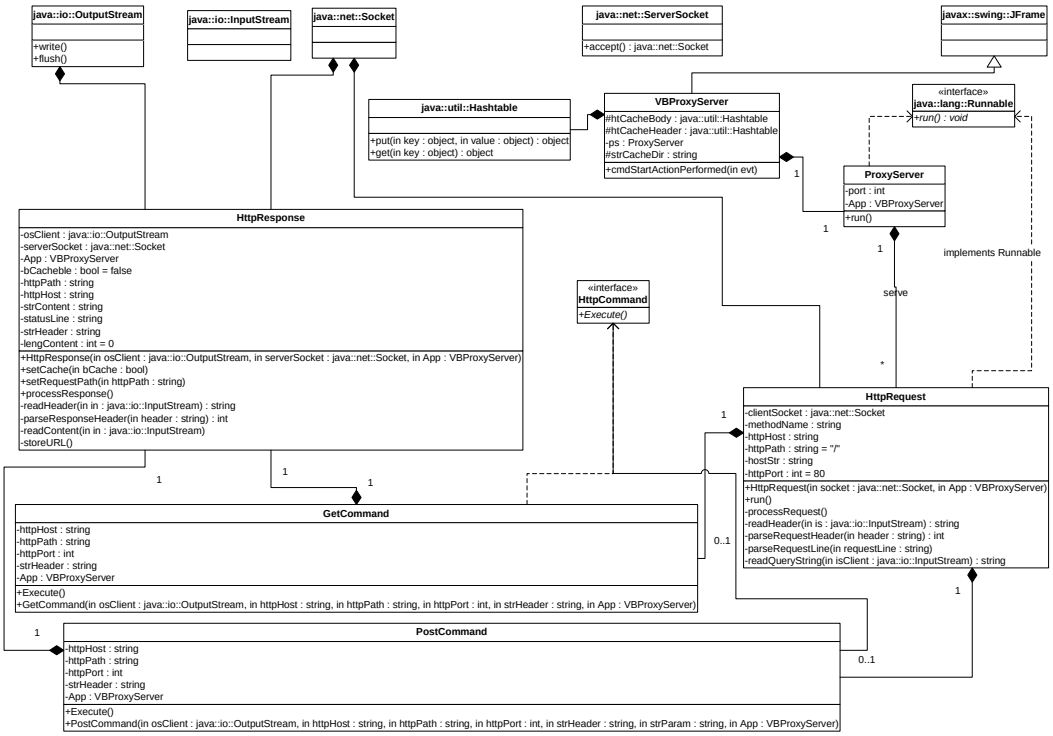


Figura 1. Static Structure - Class diagram

Figure 2 shows a GET request when there is no cache hit and the request must be sent to the Web Server, received by the Proxy Server, send it back to the browser and stored it at the Cache. Figure 3 shows the case when a Cache hit happens for a GET request and there is no need to send any request to the Web Server. A third case is a POST request where always the request is pass to the Web Server and there is no need to ask if the response is located at the cache, see figure 4.

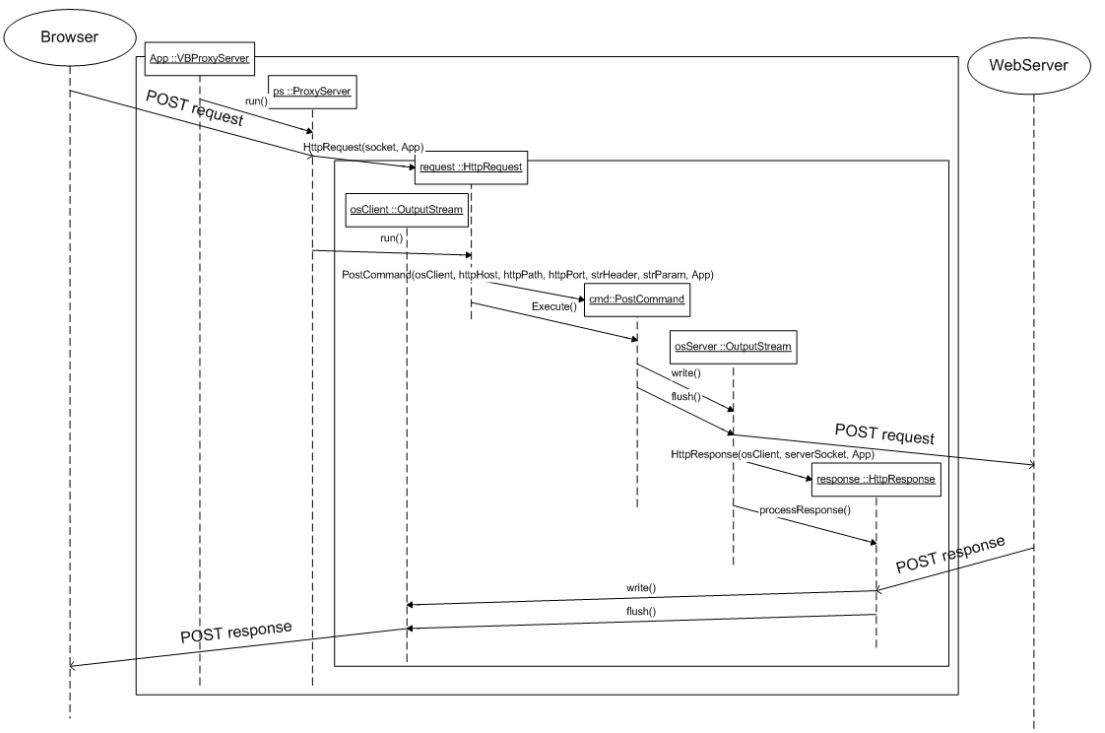


Figura 2. GET request - Cache missing object case

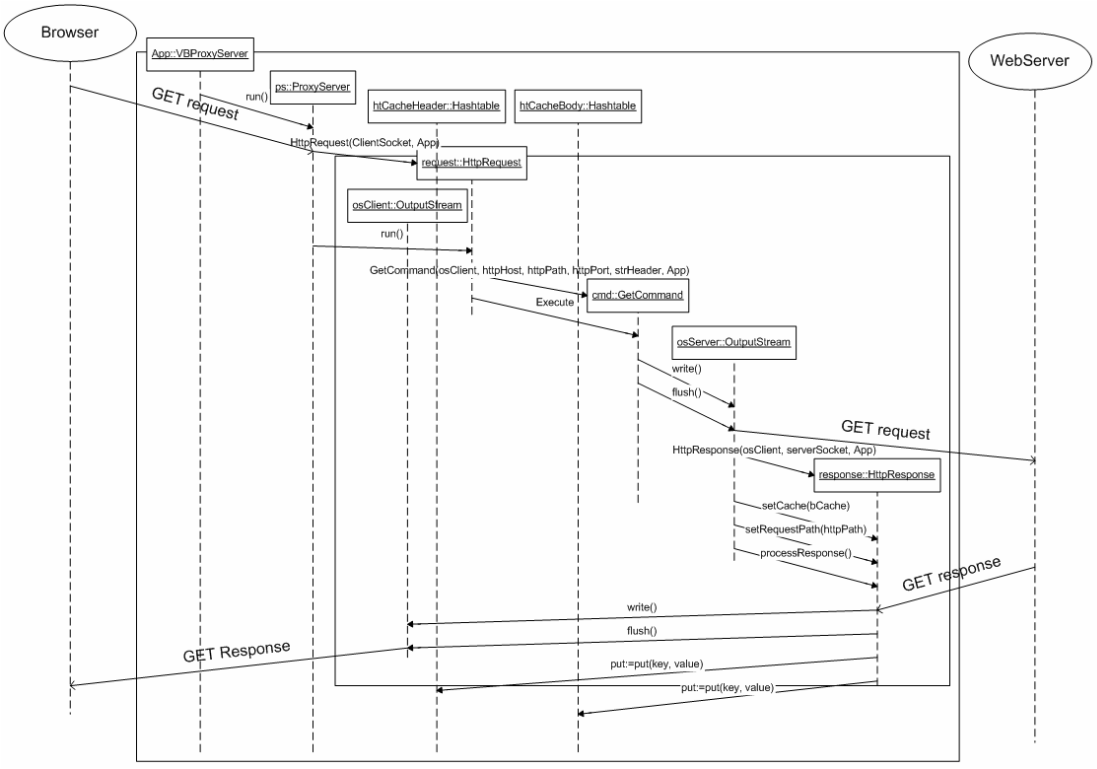


Figura 3. GET request - Cache hit object case

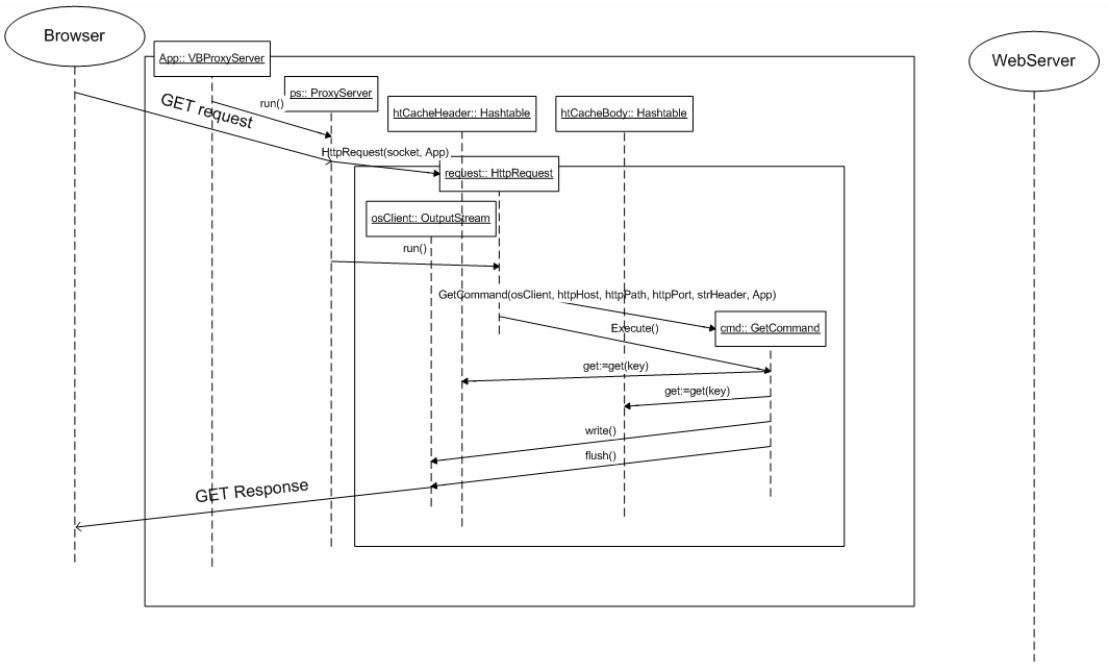


Figura 4. non cache - POST request case

2 WORKING WITH THE GRAPHIC USER INTERFACE.

The package com.vbcolombia.proxy contains the files required by the application.

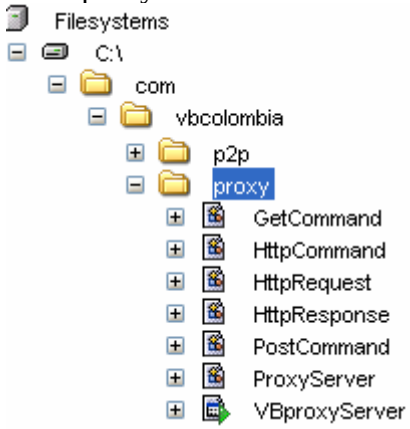


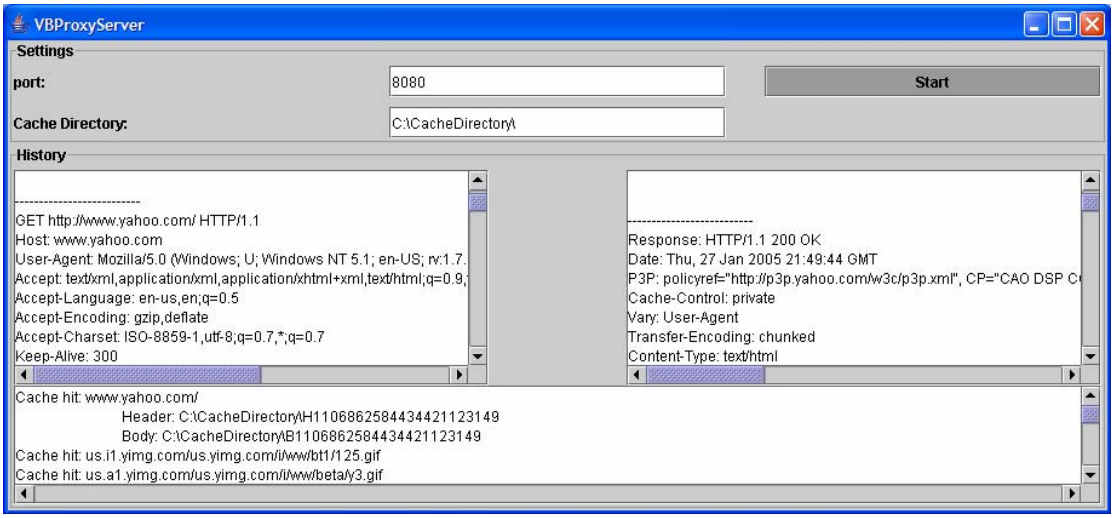
Figura 5. Package structure

From the command line it can be executed as:

> java -classpath "C:\" com.vbcolombia.proxy.VBProxyServer

Before starting the Proxy Server, the listening port number and the directory where the cache files will be located need to be specified.

Every request message received by the ProxyServer is showed at the left textarea. At the right side the response message from the Web Server are showed and in the button textarea a historical log of cache hits is displayed.


VBProxyServer
Settings
port:
8080
Start
Cache Directory:
C:\CacheDirectory\
History

GET http://www.yahoo.com/ HTTP/1.1
Host: www.yahoo.com
User-Agent: Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.7.
Accept: text/xml,application/xml,application/xhtml+xml,text/html;q=0.9,
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300

Response: HTTP/1.1 200 OK
Date: Thu, 27 Jan 2005 21:49:44 GMT
P3P: policyref="http://p3p.yahoo.com/w3c/p3p.xml", CP="CAO DSP C
Cache-Control: private
Vary: User-Agent
Transfer-Encoding: chunked
Content-Type: text/html
Cache hit: www.yahoo.com/
Header: C:\CacheDirectory\H110686258443442112314 9
Body: C:\CacheDirectory\B110686258443442112314 9
Cache hit: us.i1.yimg.com/us.yimg.com/i/ww/bt1/125.gif
Cache hit: us.a1.yimg.com/us.yimg.com/i/ww/beta/y3.gif