Simple HTTP Proxy

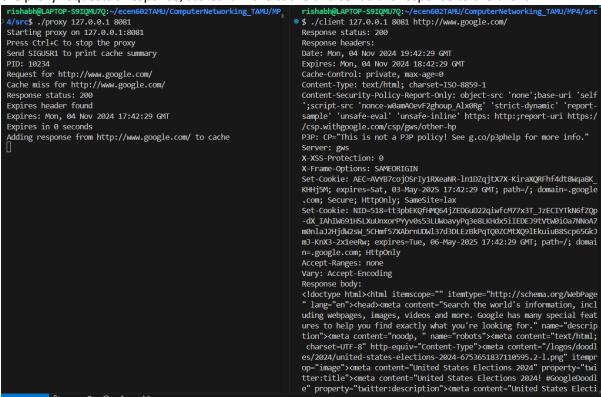
ECEN 602 Network Programming Assignment 4

Team ID: 9

Below are the tests performed on the Simple HTTP Proxy:

1. A request that is not in the cache is proxied, saved in the cache, and returned to the requester –

We see that there is a cache miss when we request for the URL http://www.google.com/ and the proxy requests a response, stores it in the cache and sends the response to the client.



Below is the response from the above URL viewed in a web browser after saving as HTML file (All the header data + HTML content):

← → C O File wsl.localhost/Ubuntu-20.04/home/rishabit/ecen602TAMU/ComputerNetworking_TAMU/MP4/src/httpф_www.google.com_html ☆ む		
Search Images Maps Play YouTube News Gmail Drive More -		Sign in 🌣
hp P3P: CP=This is not a P3P policy! See g_corp3phetp for more into.* Server: give X-XSS-Protection: 0 X-Frame-Options: SAMEORIGIN Set-Cookie: AEC=AVYB7cojOSrty1RXeaNR-In1DZqjtX7X-KiraXQRF-Infed8Wqa8K_KHHj5M; exprise=Sat, 03-May-2025 17-42:29 GMT; path+; domain=.google.com; Secure; HttpOnly; SameSite=lax Set-Cookie: NID=518-tt3pbECMpHX04EyEDGdbZe2pidwFM7X3T_JzECTWN86ZQp- dX_IAniW691HSLXuUnxxPYyv0s53LUWoavyPq3e6LKHdx5ilEDEJ9fVW0iOa7NNoA7m0nlaJ2HjdWzsW_5CHmf57XAbmUDWi37d3DLezBkPqTQ0ZCMIXQ9fEkuiuB8Scp65GkJmJ-KnX3-2x1eeRw; expires=Tue, 06-May-2025 17'42:29 GMT; path=/; domain=.google.com; HttpOnly Accept-Ranges: none Vary: Accept-Encoding		
	United States Elections 2024	Advanced search
	Google Search I'm Feeling Lucky	
Election Day is tomorrow. Find where to vote		
	Advertising Business Solutions About Google	
	© 2024 - Privacy - Terms	

2. A cache hit returns the saved data to the requester -

We see that the cache misses the first time when we request http://www.gmail.com/ and the response is stored by the HTTP Proxy. When we request for the same URL a second time, there is a cache hit and the proxy re-sends the response stored in the cache.

```
02TAMU/ComputerNetworking_TAMU/MP
                                                                           <HTML><HEAD><meta http-equiv="content-type" content="text/html;char</pre>
4/src$ ./proxy 127.0.0.1 8081
Starting proxy on 127.0.0.1:8081
Press Ctrl+C to stop the proxy
Send SIGUSR1 to print cache summary
                                                                           set=utf-8">
                                                                           <TITLE>301 Moved</TITLE></HEAD><BODY>
                                                                           <H1>301 Moved</H1>
Request for http://www.gmail.com/
                                                                           <A HREF="https://mail.google.com/mail/u/0/">here</A>.
Cache miss for http://www.gmail.com/
                                                                           </BODY></HTML>
Response status: 301
Expires header found
                                                                          rishabh@LAPTOP-S9IQMU7Q:~/ecen602TAMU/ComputerNetworking TAMU/MP4/s
Expires: Mon, 04 Nov 2024 18:56:33 GMT
                                                                         • rc$ ./client 127.0.0.1 8081 http://www.gmail.com/
Expires in 5300 seconds
                                                                          Response status: 301
Adding response from http://www.gmail.com/ to cache
                                                                          Response headers:
                                                                          Location: https://mail.google.com/mail/u/0/
Request for http://www.gmail.com/
Cache hit for http://www.gmail.com/
                                                                          Cross-Origin-Resource-Policy: cross-origin
                                                                          X-Content-Type-Options: nosniff
Re-send the response
Expires in 5296 seconds
                                                                           Content-Length: 230
                                                                           X-XSS-Protection: 0
                                                                          Date: Mon, 04 Nov 2024 19:26:33 GMT
                                                                          Expires: Mon, 04 Nov 2024 19:56:33 GMT Cache-Control: public, max-age=1800
                                                                           Content-Type: text/html; charset=UTF-8
                                                                           Age: 100
                                                                           Response body:
                                                                           <HTML><HEAD><meta http-equiv="content-type" content="text/html;char</pre>
                                                                           set=utf-8">
                                                                           <TITLE>301 Moved</TITLE></HEAD><BODY>
                                                                           <H1>301 Moved</H1>
                                                                           The document has moved
                                                                           <A HREF="https://mail.google.com/mail/u/0/">here</A>.
                                                                           </BODY></HTML>
                                                                         rishabh@LAPTOP-S9IQMU7Q:~/ecen602TAMU/ComputerNetworking_TAMU/MP4/s
orcs
```

We also perform the test using an HTTP server running on the localhost on port 8080. The HTTP server sends a response which expires in 60 seconds.

```
rishabh@APTOP-S9IQWI7Q:-/ecen602TAWI/ComputerNetworking_TAWI/MP4/s

0 4/srcs /proxy 127.0.0.1 8881

Starting proxy on 127.0.0.1:8081

Press Ctrlc to stop the proxy

Send SIGUSRI to print cache summary

PID: 97/3

Request for http://www.jgmail.com/

Response status: 200

Response for inttp://lowa.lgmil.com/

Cortent-Tuppel.Tuppel.Tuppel.Tuppel.Tuppel.
```

3. A cache miss with 10 items already in the cache is proxied, saved in the LRU location in cache, and the data is returned to the requester –

We see that the HTTP Proxy declares "Cache is full" and then goes on to replace the oldest

cache entry (LRU) with the response from the latest URL requested.

```
Request for http://ucsd.edu/
                                                                                    h@LAPTOP-S9IQMU7Q:~/ecen602TAMU/ComputerNetworking_TAMU/MP4/src
Cache hit for http://ucsd.edu/
                                                                            • $ ./client 127.0.0.1 8081 http://www.mit.edu/
Re-send the response
                                                                             Response status: 301
Expires in 3593 seconds
                                                                             Response headers:
Request for http://www.utdallas.edu/
                                                                             Server: AkamaiGHost
Cache miss for http://www.utdallas.edu/
                                                                             Content-Length: 0
                                                                             Location: https://www.mit.edu/
Response status: 301
Date header found
Date: Mon, 04 Nov 2024 18:54:19 GMT
                                                                             Date: Mon, 04 Nov 2024 19:54:38 GMT
                                                                             Connection: close
Expires in 3600 seconds
                                                                             Content-Security-Policy: frame-ancestors 'self' https://web.mit.edu h
Cache is full, removing oldest entry
Adding response from http://www.utdallas.edu/ to cache
                                                                             ttps://www.mit.edu http://web.mit.edu http://www.mit.edu
                                                                             X-Frame-Options: SAMEORIGIN rishabh@LAPTOP-S9IQMU7Q:~/ecen602TAMU/ComputerNetworking_TAMU/MP4/src
Request for http://www.nyu.edu/
Cache miss for http://www.nyu.edu/
                                                                            • $ ./client 127.0.0.1 8081 http://www.tamu.edu
Response status: 301
                                                                             Response status: 301
Date header found
Date: Mon, 04 Nov 2024 18:54:29 GMT
                                                                             Response headers:
                                                                             Server: CloudFront
Expires in 3600 seconds
                                                                             Date: Mon, 04 Nov 2024 19:54:48 GMT
                                                                             Content-Type: text/html
Cache is full, removing oldest entry
Adding response from http://www.nyu.edu/ to cache
Request for http://www.mit.edu/
Cache miss for http://www.mit.edu/
                                                                             Content-Length: 167
                                                                             Connection: close
                                                                             Location: https://www.tamu.edu/
Response status: 301
                                                                             X-Cache: Redirect from cloudfront
Date header found
                                                                             Via: 1.1 6cb2fa1d5087faf89d959c36e8f2fbfa.cloudfront.net (CloudFront)
                                                                             X-Amz-Cf-Pop: DFW59-P2
Expires in 3600 seconds
                                                                             X-Amz-Cf-Id: LHt9Ll_32bZ2jeHscab1kainQHgYEHWeEJsdZVNlhPUwBS72W9t2-A==
Cache is full, removing oldest entry
                                                                             Response body:
Adding response from http://www.mit.edu/ to cache
                                                                             <html>
Request for http://www.tamu.edu
Cache miss for http://www.tamu.edu
                                                                             <head><title>301 Moved Permanently</title></head>
                                                                             <body>
                                                                             <center><h1>301 Moved Permanently</h1></center>
Response status: 301
Date header found
                                                                             <hr><center>CloudFront</center>
Date: Mon, 04 Nov 2024 18:54:48 GMT
                                                                             </body>
Expires in 3600 seconds
                                                                             </html>
Cache is full, removing oldest entry
Adding response from http://www.tamu.edu to cache
                                                                             rishabh@LAPTOP-S9IQMU7Q:~/ecen602TAMU/ComputerNetworking_TAMU/MP4/src
```

4. (Bonus Implemented) A stale 'Expires' header in the cache is accessed, the cache entry is replaced with a fresh copy, and the fresh data is delivered to the requester – We see that the HTTP Proxy declares that there is a stale cache entry for the requested URL. Then it sends 'If-Modified-Since' header and replaces the cache entry with a fresh copy of data.

```
\label{twed} $$ tved="+f+"\&jsname="+(a||""))} else $F=a,E=[c]\} window.document.addEvent Listener("DOMContentLoaded",function(){document.body.addEventListener} $$ for the property of the pr
                  @LAPTOP-S9IQMU7Q:~/ecen602TAMU/ComputerNetworking_TAMU/MP
4/src$ ./proxy 127.0.0.1 8081
Starting proxy on 127.0.0.1:8081
Press Ctrl+C to stop the proxy
                                                                                                                                                               ("click",G)});}).call(this);</script></body></html>
rishabh@LAPTOP-S9IQMU7Q:~/ecen602TAMU/ComputerNetworking_TAMU/MP4/src
                                                                                                                                                            • $ ./client 127.0.0.1 8081 http://www.google.com/
Send SIGUSR1 to print cache summary
PID: 11806
                                                                                                                                                               Response status: 200
Request for http://www.tamu.edu/
                                                                                                                                                               Response headers:
Cache miss for http://www.tamu.edu/
                                                                                                                                                               Date: Mon, 04 Nov 2024 20:08:48 GMT
                                                                                                                                                               Expires: Mon, 04 Nov 2024 19:08:49 GMT
Cache-Control: private, max-age=0
Content-Type: text/html; charset=ISO-8859-1
Response status: 301
Date header found
Date: Mon, 04 Nov 2024 19:08:34 GMT
                                                                                                                                                               Content-Security-Policy-Report-Only: object-src 'none';base-uri 'self 
';script-src 'nonce-lzQPU9AVuo2LuTKAnGlWfA' 'strict-dynamic' 'report-
sample' 'unsafe-eval' 'unsafe-inline' https: http:;report-uri https:/
Expires in 3600 seconds
Adding response from http://www.tamu.edu/ to cache
Request for http://www.google.com/
Cache miss for http://www.google.com/
                                                                                                                                                                /csp.withgoogle.com/csp/gws/other-hp
Response status: 200
                                                                                                                                                                P3P: CP="This is not a P3P policy! See g.co/p3phelp for more info."
Expires header found
                                                                                                                                                               Server: gws
Expires: Mon, 04 Nov 2024 18:08:46 GMT
                                                                                                                                                               X-XSS-Protection: 0
                                                                                                                                                               X-Frame-Options: SAMEORIGIN
Expires in 0 seconds
Adding response from http://www.google.com/ to cache
                                                                                                                                                               Set-Cookie: AEC=AVYB7crWZCXWsEBXmiRUK8irSxTR69B8kCBr2JNCOcVkjOtxhhtWH
Request for http://www.google.com/
                                                                                                                                                               eVEgQ; expires=Sat, 03-May-2025 18:08:48 GMT; path=/; domain=.google.
Stale cache entry for http://www.google.com/
Sending if modified since header
                                                                                                                                                               com; Secure; HttpOnly; SameSite=lax
Set-Cookie: NID=518=KjomQbdsE0Ze_LPBOIBLDHH-Pz4_8tLw8j_nKkLbTGFqJWMfy
                                                                                                                                                                ah9PTPoBoTWaMOpLGUBFt0ex2GBfkWn2V2cRNolhcFI3F0402Wup4HvEinIfz1gh9VUge
Response status: 200
Expires header found
                                                                                                                                                               yLfivmbCWizxLZgvdrRnaCevrmQxE3evSUZT-nzoce5L8ybk_IDtP5r_asZV6Fz_u0y2_
                                                                                                                                                                 5V8XkDDph; expires=Tue, 06-May-2025 18:08:48 GMT; path=/; domain=.goo
Expires: Mon, 04 Nov 2024 18:08:49 GMT
Expires in 1 seconds
                                                                                                                                                               gle.com; HttpOnly
Adding response from http://www.google.com/ to cache
                                                                                                                                                               Accept-Ranges: none
                                                                                                                                                               Vary: Accept-Encoding
                                                                                                                                                               Response body:
                                                                                                                                                               </doctype html><html itemscope="" itemtype="http://schema.org/WebPage
" lang="en"><head><meta content="Search the world's information, incl</pre>
                                                                                                                                                              uding webpages, images, videos and more. Google has many special feat ures to help you find exactly what you're looking for." name="descrip tion"><meta content="noodp, " name="robots"><meta content="text/html; charset=UFF-8" http-equiv="Content-Type"><meta content="/logos/doodl
                                                                                                                                                               es/2024/united-states-elections-2024-6753651837110595.2-l.png" itempr
```

HTTP server sends a response which expires immediately. When we request the same URL the second time, we see that the HTTP Proxy declares that the cache entry is stale and sends the 'If-Modified-Since' header, replaces the cache entry with a fresh one and sends it to the client.

```
rishabh@LAPTOP-S9IQMJQ:-/ecen602TAMU/ComputerNe
tworking_TAMU/MPA/srcf. /proxy 177.0.0.1 8081
Starting proxy on 127.0.0.1:8081
Press CtrluC to stop the proxy
Send SIGUSRI to print cache summary
PID: 14456
Request for http://localhost:8080/test4
Response status: 200
Response status: 200
Expires header found
Expires: Mon, 04 Nov 2024 20:08:05 GMT
Expires in -10 seconds
Adding response from http://localhost:8080/test4
Sending if modified since header
Response status: 200
Response round http://localhost:8080/test4
Sending if modified since header
Response status: 200
Response round http://localhost:8080/test4
Sending if modified since header
Response from http://localhost:8080/test4
Sending response from http://localhost:8080/test4 to cache
Response status: 200
Response round response from http://localhost:8080/test4
Sending if modified since header
Response status: 200
Response round response from http://localhost:8080/test4
Sending if modified since header
Response status: 200
Response round response from http://localhost:8080/test4
Sending if modified since header
Response round response from http://localhost:8080/test4
Sending if modified since header
Response round response from http://localhost:8080/test4
Sending if modified since header
Response round response from http://localhost:8080/test4
Sending if modified since header
Res
```

5. **(Bonus implemented)** A stale entry in the cache without an 'Expires' header is determined by sending an 'If-Modified-Since' header to use Conditional GET, the Web server responds with 'NOT_MODIFIED' (304) response, and the old cache entry is used to deliver data to the client –

We see that HTTP proxy detects that the expires header was missing, sends an 'If-Modified-Since' header and as the response is NOT_MODIFIED (304), it serves the data from the cache entry.

```
rishabh@LAPTOP-S9IQWI7Q:-/ecen602TAMU/ComputerNetworking_TAMU/MP4/src

4/srcs /proxy 127.0.0.1 8081
Press Ctrl+C to stop the proxy
Send SIGUSRI to print cache summary
PID: 15870
Request for http://localhost:8080/test5
Cache miss for http://localhost:8080/test5
Response status: 200
Last-Modified: Mon, 04 Nov 2024 21:16:22 GMT
Expires in -10 seconds
Adding response from http://localhost:8080/test5 to cache
Request for http://localhost:8080/test5 to cache
Reponse status: 304
Serving from cache

Prishabh@LAPTOP-S9IQWI7Q:-/ecen602TAMU/ComputerNetworking_TAMU/MP4/src

* /client 127.0.0.1 8081 http://localhost:8080/test5

Response status: 200
Response headers:
Content-Type: text/plain
Date: Mon, 04 Nov 2024 21:16:32 GMT
Last-Modified: Mon, 04 Nov 2024 21:16:22 GMT
Content-Length: 9
Response body:
Test 5: 0
```

6. (Bonus Implemented) A cache entry without an 'Expires' header that has been modified since the last request is determined by sending 'If-Modified-Since' header, the cache entry is replaced with a fresh copy, and the fresh data is delivered to the requester – We see that when we make the request for the first time, we receive response without an 'Expires' header but 'Last-Modified' header is found, the HTTP Proxy stores the data in the cache. When we request the same URL a second time, the HTTP Proxy sends the 'If-Modified-Since' header, receives a fresh response from the HTTP Server, the cache entry is replaced with a fresh copy, and the fresh data is delivered to the requester.

```
rishabh@LAPTOP-S9IQWI7Q:-/ecen602TAWU/ComputerNetworking_TAWU/MP4/src
2 Afarcs ./proxy 127.0.0.1 8081
Starting proxy on 127.0.0.1:80801
Press Ctrl.ct to stop the proxy
Send SIGUSRI to print cache summary
PID: 16324
Request for http://localhost:8080/test6
Response status: 200
Response status: 200
Response status: 200
Response status: 200
Last-Modified: sat, 65 Oct 2024 21:33:48 GMT
Content-Length: 9
Response body:
Text 6: 0
Prishabh@LAPTOP-S9IQWI7Q:-/ecen602TAWU/ComputerNetworking_TAWU/MP4/src
Pib: 16324
Request for http://localhost:8080/test6
Response status: 200
Response status: 200
Response status: 200
Response status: 200
Response body:
Text 6: 0
Prishabh@LAPTOP-S9IQWI7Q:-/ecen602TAWU/ComputerNetworking_TAWU/MP4/src
Pib: 10246
Response status: 200
Resp
```

7. Three clients can simultaneously access the proxy server and get the correct data — We see that 3 clients are able to access the proxy server and get the requested data from the respective URLs.

