

# 시스템프로그래밍 과제 보고서

## Proxy 2-3

수업 명: 시스템프로그래밍 월5수6

담당 교수: 김태석 교수님

학과: 컴퓨터정보공학부

학번: 2023202070

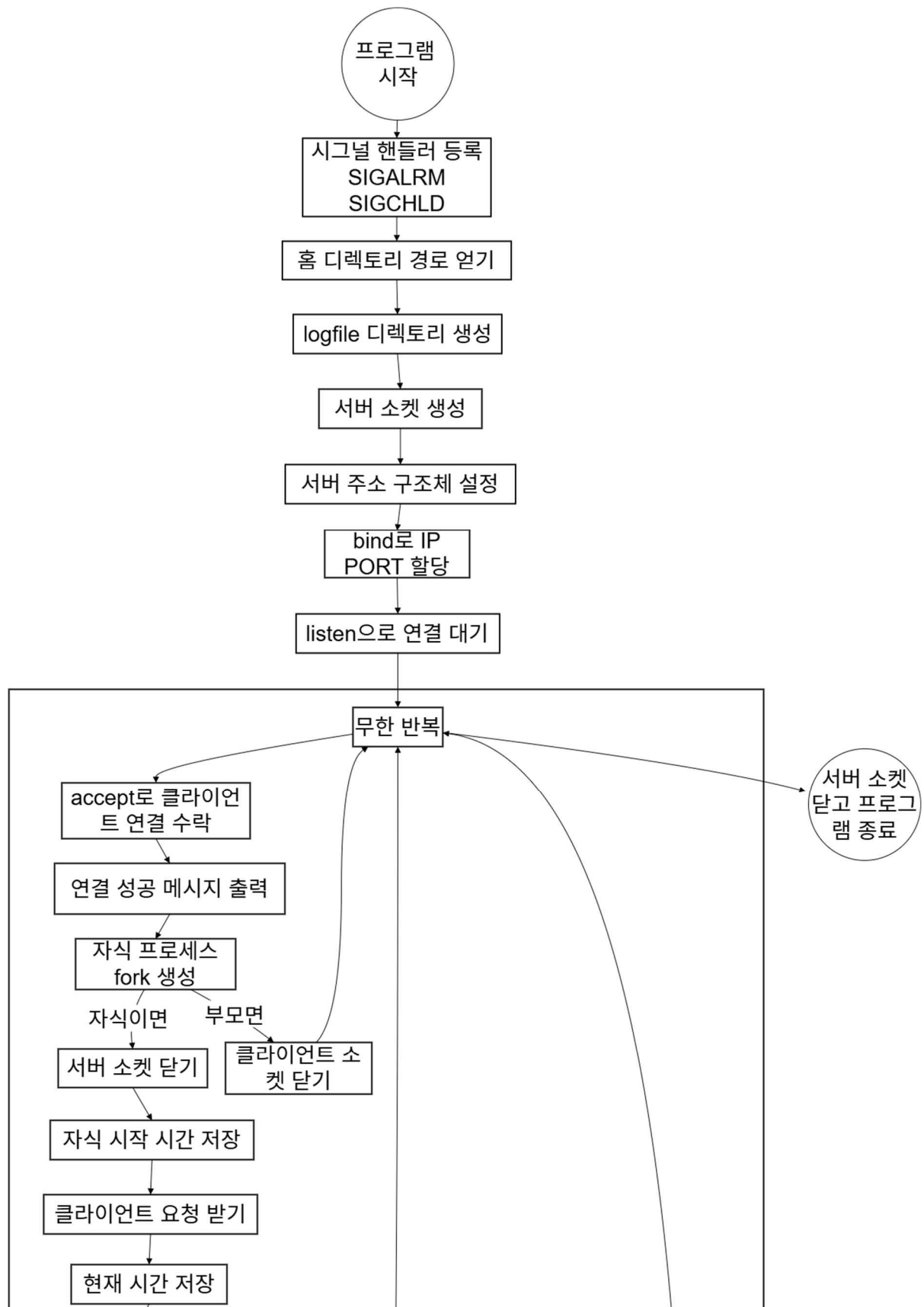
이름: 최현진

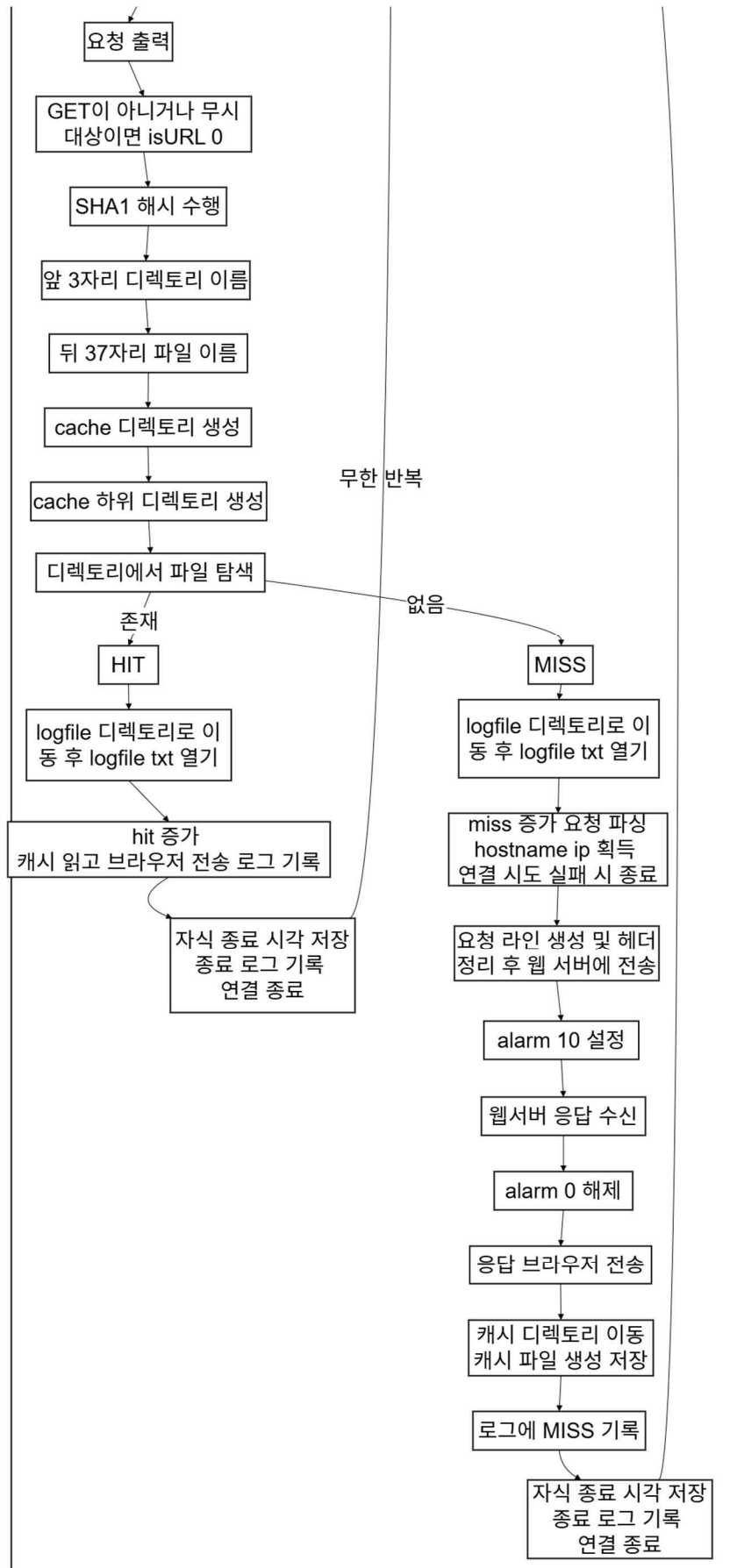
제출일: 2025.05.15

## Introduction

이번 Proxy 2-3 과제는 이전 Proxy 2-2 과제에서 구현한 기능을 기반으로, 오리진 서버(웹 서버)와의 통신 기능을 포함한 프록시 서버를 구현하는 실습 과제이다. 프록시 서버는 웹 브라우저로부터 HTTP 요청을 수신하고, 해당 요청을 기반으로 HIT 또는 MISS를 판별한다. MISS인 경우, 오리진 서버에 직접 HTTP 요청을 전송하고, 응답을 수신하여 브라우저에 전달한 뒤 캐시에 저장한다. 응답이 10초 이내에 수신되지 않을 경우, SIGALRM 시그널을 통해 응답 없음 메시지를 출력하고 자식 프로세스를 종료한다. 자식 프로세스의 종료 시에는 SIGCHLD 시그널을 통해 좀비 프로세스를 정리한다. 모든 요청 결과(HIT/MISS)와 자식 프로세스의 종료 시점 정보를 로그 파일에 기록한다. 본 과제를 통해 시스템 프로그래밍의 주요 주제인 멀티프로세스, 네트워크 통신, 시그널 제어, 파일 입출력에 대해 통합적으로 학습할 수 있다.

## Flow Chart





## Pseudo code

main:

프로그램 시작

    시그널 핸들러 등록(SIGALRM, SIGCHLD)

    홈 디렉토리 경로 얻기

    ~/logfile 디렉토리 생성

    서버 소켓 생성

    서버 주소 구조체 설정

    bind()로 IP/PORT 할당

    listen()으로 연결 대기

무한 반복:

    accept()로 클라이언트 연결 수락

    연결 성공 메시지 출력

    fork()로 자식 프로세스 생성

자식이면:

    서버 소켓 닫기

    자식 시작 시간 저장

    클라이언트 요청 받기

    현재 시간 저장

    요청 출력

    method가 GET이 아니거나 URL이 없거나 ico/css/txt 파일이면 isURL = 0

    SHA1 해시 수행

        앞 3자리 → 캐시 디렉토리 이름

        뒤 37자리 → 캐시 파일 이름

    ~/cache 디렉토리 생성

    ~/cache/xxx 디렉토리 생성

isURL일 경우:

디렉토리에서 파일 탐색

존재 → HIT

없으면 → MISS

~/logfile 디렉토리로 이동

logfile.txt 열기

HIT이면:

hit++

캐시 디렉토리로 이동 후 캐시 파일 열기

파일 읽어 브라우저로 전송

로그에 HIT 기록

MISS이면:

miss++

요청에서 hostname, port, path 추출

hostname으로 오리진 서버 IP 얻기

소켓 생성 및 오리진 서버에 연결 시도

실패 시 종료

HTTP/1.0 요청 라인 생성

브라우저 요청 헤더들 정제하여 재조합

오리진 서버에 요청 전송

alarm(10) 설정

오리진 서버로부터 응답 수신

alarm(0)으로 해제

응답을 브라우저로 전송

~/cache/xxx 디렉토리로 이동 후 캐시 파일 생성

응답 내용을 캐시 파일에 저장

로그에 MISS 기록

자식 종료 시간 저장

종료 로그 기록

연결 종료 메시지 출력

클라이언트 소켓 닫기

자식 종료

부모면:

클라이언트 소켓 닫기

서버 소켓 닫고 프로그램 종료

## 결과 화면

### 1. proxy 2-2

#### 1. make

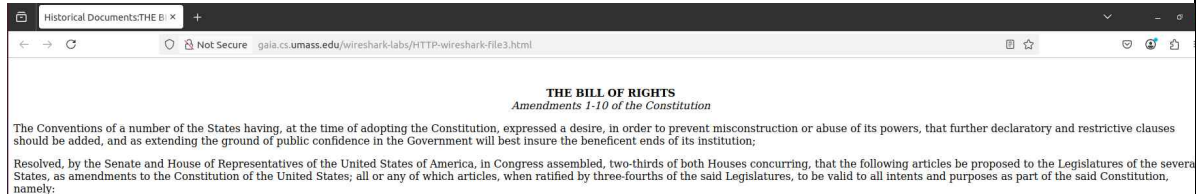
```
kw2023202070@ubuntu:~$ ls
Desktop  Documents  Downloads  Makefile  Music  Pictures  proxy_cache.c  Public  Templates  Videos
kw2023202070@ubuntu:~$ make
gcc proxy_cache.c -o proxy_cache -lcrypto
kw2023202070@ubuntu:~$ ls
Desktop  Documents  Downloads  Makefile  Music  Pictures  proxy_cache  proxy_cache.c  Public  Templates  Videos
kw2023202070@ubuntu:~$
```

\$ ls: Makefile과 소스 코드를 작성하였다.

\$ make: make를 실행하여 컴파일을 수행하였고, gcc 컴파일 명령어가 정상적으로 실행되었다.

\$ ls: 컴파일 결과, 실행 파일 proxy\_cache가 생성된 것을 확인했다.

#### 2. no response 동작 확인



Historical Documents: THE B. x

← → ↻ Not Secure gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html

THE BILL OF RIGHTS  
Amendments 1-10 of the Constitution

The Conventions of a number of the States having, at the time of adopting the Constitution, expressed a desire, in order to prevent misconstruction or abuse of its powers, that further declaratory and restrictive clauses should be added, and as extending the ground of public confidence in the Government will best insure the beneficent ends of its institution;

Resolved, by the Senate and House of Representatives of the United States of America, in Congress assembled, two-thirds of both Houses concurring, that the following articles be proposed to the Legislatures of the several States, as amendments to the Constitution of the United States; all or any of which articles, when ratified by three-fourths of the said Legislatures, to be valid to all intents and purposes as part of the said Constitution, namely:

<http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html> 입력후 바로 가상머신의 입력을 끊었다.

```
[192.168.245.128 : 55358] client was connected
=====
Request from [192.168.245.128 : 55358]
GET http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html HTTP/1.1
Host: gaia.cs.umass.edu
User-Agent: Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:136.0) Gecko/20100101 Firefox/136.0
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate
Connection: keep-alive
Upgrade-Insecure-Requests: 1
Priority: u=0, i

=====
===== MISS =====
===== [192.168.245.128 : 55358] connect with origin server =====
===== no response =====
█
```

터미널 출력 확인 결과 child process는 browser request의 url을 확인하고 MISS임을 판단했다. 이후 web server로부터 response를 받기 위해 origin server와 연결했다. 연결 후 response를 read()를 통해 받아오는 동안, 인터넷 연결이 끊겼기에 no response가 출력되고 child process가 종료됐다.

```
kw2023202070@ubuntu:~$ cat ~/logfile/logfile.txt
kw2023202070@ubuntu:~$ █
```



MISS이지만 response를 10초 동안 받지 못하여 child process가 종료됐고 생성할 캐시 파일 data가 없으므로, log는 기록되지 않았다.

### 3. MISS, HIT 동작 확인

<http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html> 을 브라우저 상에 입력했다.

```
[192.168.245.128 : 58580] client was connected
=====
Request from [192.168.245.128 : 58580]
GET http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html HTTP/1.1
Host: gaia.cs.umass.edu
User-Agent: Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:136.0) Gecko/20100101 Firefox/136.0
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate
Connection: keep-alive
Upgrade-Insecure-Requests: 1
Priority: u=0, i

=====
===== MISS =====
===== [192.168.245.128 : 58580] connect with origin server =====
===== response buffer =====
HTTP/1.1 200 OK
Date: Wed, 14 May 2025 14:47:28 GMT
Server: Apache/2.4.6 (CentOS) OpenSSL/1.0.2k-fips PHP/7.4.33 mod_perl/2.0.11 Perl/v5.16.3
Last-Modified: Wed, 14 May 2025 05:59:01 GMT
ETag: "1194-635123eee37eb"
Accept-Ranges: bytes
Content-Length: 4500
Keep-Alive: timeout=5, max=100
Connection: Keep-Alive
Content-Type: text/html; charset=UTF-8

<html><head>
<title>Historical Documents:THE BILL OF RIGHTS</title></head>

<body bgcolor="#ffffff" link="#330000" vlink="#666633">
<p><br>
</p>
<p></p><center><b>THE BILL OF RIGHTS</b><br>
  <em>Amendments 1-10 of the Constitution</em>
</center>

<p>The Conventions of a number of the States having, at the time of adopting
</p><p><a name="9"><strong><h3>Amendment IX</h3></strong></a>

<p></p><p>The enumeration in the Constitution, of certain rights, shall
not be construed to deny or disparage others retained by the people.

</p><p><a name="10"><strong><h3>Amendment X</h3></strong></a>

<p></p>
<p>The powers not delegated to the United States by the Constitution, nor prohibited
  by it to the states, are reserved to the states respectively, or to the people.</p>
</body></html>
[192.168.245.128 : 58580] client was disconnected
```

58580 client가 요청을 보내오고 child process는 MISS임을 판단했다. 이후 58580 client가 origin web server로 연결하고 read()를 통해 받은 응답 버퍼를 cmd에 출력해봤다. (너무 길어서 중간에 잘랐다.) 해당 응답은 캐시 파일에 저장된다. 이후 web server와의 연결이 끊기고 child

process가 종료됐다.

```
[192.168.245.128 : 55072] client was connected
=====
Request from [192.168.245.128 : 55072]
GET http://gaia.cs.umass.edu/favicon.ico HTTP/1.1
Host: gaia.cs.umass.edu
User-Agent: Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:136.0) Gecko/20100101 Firefox/136.0
Accept: image/avif,image/webp,image/png,image/svg+xml,image/*;q=0.8,*/*;q=0.5
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate
Connection: keep-alive
Referer: http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html
Priority: u=6

=====
[192.168.245.128 : 55072] client was disconnected
```

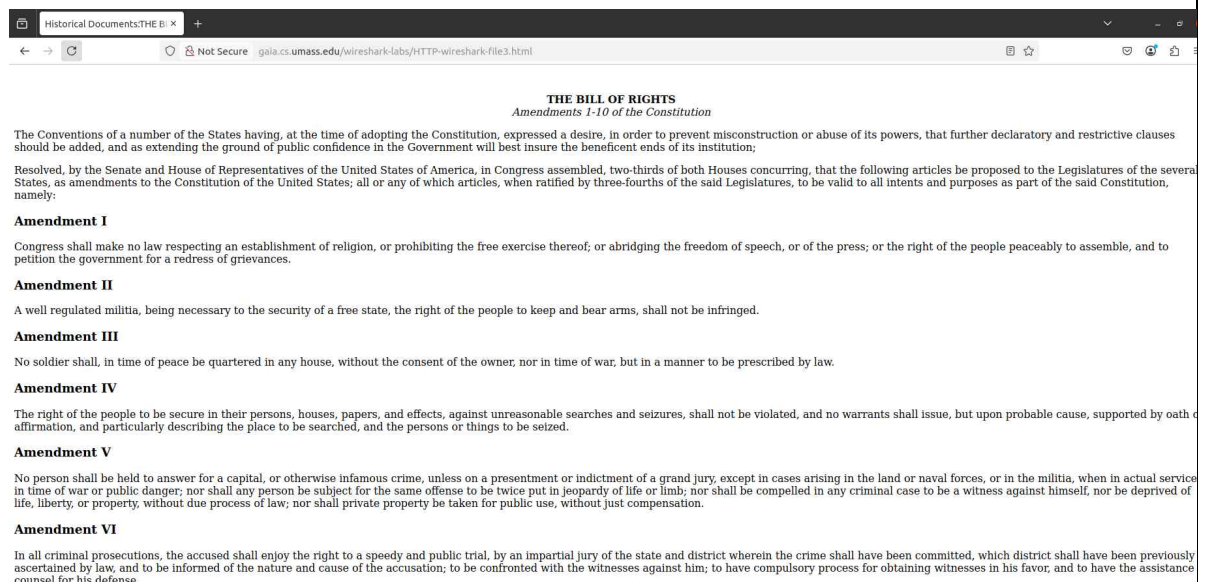
중간에 위와 같은 favicon.ico 요청이 많이 들어오지만, HIT/MISS 판단에서 제외된다. 따라서 log나 캐시되지 않고, client 연결과 child process가 종료된다.

```
[192.168.245.128 : 55158] client was connected
=====
Request from [192.168.245.128 : 55158]
GET http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html HTTP/1.1
Host: gaia.cs.umass.edu
User-Agent: Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:136.0) Gecko/20100101 Firefox/136.0
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate
Connection: keep-alive
Upgrade-Insecure-Requests: 1
If-Modified-Since: Wed, 14 May 2025 05:59:01 GMT
If-None-Match: "1194-635123eee37eb"
Priority: u=0, i

=====
===== HIT =====
===== send cache to browser =====
[192.168.245.128 : 55158] client was disconnected
```

<http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html> 을 브라우저 상에 재입력했다.

55158 client가 요청을 보내오고 child process는 HIT임을 판단했다. 이후 child process는 캐시 파일에 저장되어 있는 response를 client(browser)로 보낸다. client 연결과 child process가 종료된다.

A screenshot of a web browser window. The address bar shows the URL "gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html". The page content is titled "THE BILL OF RIGHTS" and "Amendments 1-10 of the Constitution". It includes the text of the first three amendments: Amendment I (Congress shall make no law respecting an establishment of religion...), Amendment II (A well regulated militia, being necessary to the security of a free state...), and Amendment III (No soldier shall, in time of peace be quartered in any house...). The page is formatted with a serif font and a light background.

#### Amendment VII

In suits at common law, where the value in controversy shall exceed twenty dollars, the right of trial by jury shall be preserved, and no fact tried by a jury, shall be otherwise reexamined in any court of the United States, than according to the rules of the common law.

#### Amendment VIII

Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishments inflicted.

#### Amendment IX

The enumeration in the Constitution, of certain rights, shall not be construed to deny or disparage others retained by the people.

#### Amendment X

The powers not delegated to the United States by the Constitution, nor prohibited by it to the states, are reserved to the states respectively, or to the people.

MISS일 때 browser에 web server가 보내온 응답과 같은 화면이 HIT일 때도 잘 출력된다.

```
kw2023202070@ubuntu:~$ cat ~/logfile/logfile.txt
[MISS] ServerPID : 5389 | http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html - [2025/05/14, 07:47:27]
[Terminated] ServerPID : 5389 | run time : 6 sec. #request hit : 0, miss : 1
[HIT] ServerPID : 5405 | 17f/dfdf9b8d3e1680f5ce00777638c24b8dd4b4a - [2025/05/14, 07:47:38]
[HIT] http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html
[Terminated] ServerPID : 5405 | run time : 0 sec. #request hit : 1, miss : 0
```

<http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html> url에 대한 한 child process에서 MISS 1번, 또다른 child process에서 HIT 1번일 때의 로그 파일 출력 결과를 확인했다.

#### 4. cache directory

```
kw2023202070@ubuntu:~$ ls -R ~/cache
/home/kw2023202070/cache:
17f
dfdf9b8d3e1680f5ce00777638c24b8dd4b4a
```

\$ ls -R ~/cache: ls 명령어를 -R 옵션을 통해 재귀적으로 실행하여 ~/cache 디렉토리 및 그 하위 디렉토리까지 출력하여 캐시 디렉토리 및 파일이 생성됨을 확인했다.

```
kw2023202070@ubuntu:~$ tree ~/cache/
/home/kw2023202070/cache/
├── 17f
│   └── dfdf9b8d3e1680f5ce00777638c24b8dd4b4a
1 directory, 1 file
kw2023202070@ubuntu:~$
```

\$ tree ~/cache/: ~/cache 구조 확인 결과, SHA1 해시된 url의 앞 3글자를 이름으로 하여 디렉토리가 생성되었다. 그 디렉토리의 하위에는 나머지 37글자 이름으로 파일이 생성되었다.

cmd에 출력된 바와 같이 다양한 브라우저 요청이 들어오지만, 예외 처리를 통해 실제로 입력된 <http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html> url에 대해서만 해싱하고 캐시한 결과를 나타낸다.



Open

dfdf9b8d3e1680f5ce00777638c24b8dd4b4a  
~/cache/111

```
1 HTTP/1.1 200 OK
2 Date: Wed, 14 May 2025 14:47:28 GMT
3 Server: Apache/2.4.6 (CentOS) OpenSSL/1.0.2k-fips PHP/7.4.33 mod_perl/2.0.11 Perl/v5.16.3
4 Last-Modified: Wed, 14 May 2025 05:59:01 GMT
5 ETag: "1194-635123eee37eb"
6 Accept-Ranges: bytes
7 Content-Length: 4500
8 Keep-Alive: timeout=5, max=100
9 Connection: Keep-Alive
10 Content-Type: text/html; charset=UTF-8
11
12 <html><head>
13 <title>Historical Documents:THE BILL OF RIGHTS</title></head>
14
15
16 <body bgcolor="#ffffff" link="#330000" vlink="#666633">
17 <p><br>
18 </p>
19 <p><center><b>THE BILL OF RIGHTS</b><br>
20 <em>Amendments 1-10 of the Constitution</em>
21 </center>
22
23 <p>The Conventions of a number of the States having, at the time of adopting
24 the Constitution, expressed a desire, in order to prevent misconstruction
25 or abuse of its powers, that further declaratory and restrictive clauses
26 should be added, and as extending the ground of public confidence in the
27 Government will best insure the beneficent ends of its institution; </p><p> Resolved, by the Senate and House of Representatives of the United
28 States of America, in Congress assembled, two-thirds of both Houses concurring,
29 that the following articles be proposed to the Legislatures of the several
30 States, as amendments to the Constitution of the United States; all or any
31 of which articles, when ratified by three-fourths of the said Legislatures,
32 to be valid to all intents and purposes as part of the said Constitution,
33 namely: </p><p><a name="1"><strong><h3>Amendment I</h3></strong></a>
34
35 <p></p><p>Congress shall make no law respecting an establishment of
36 religion, or prohibiting the free exercise thereof; or
37 abridging the freedom of speech, or of the press; or the
38 right of the people peaceably to assemble, and to petition
39 the government for a redress of grievances.
40
41 </p><p><a name="2"><strong><h3>Amendment II</h3></strong></a>
42
43 <p></p><p>A well regulated militia, being necessary to the security
44 of a free state, the right of the people to keep and bear
45 arms, shall not be infringed.
46
47 </p><p><a name="3"><strong><h3>Amendment III</h3></strong></a>
48
49 <p></p><p>No soldier shall, in time of peace be quartered in any house,
50 without the consent of the owner, nor in time of war, but
51 in a manner to be prescribed by law.
52
53 </p><p><a name="4"><strong><h3>Amendment IV</h3></strong></a>
54
55 <p></p><p>The right of the people to be secure in their persons, houses,
56 papers, and effects, against unreasonable searches and seizures,
57 shall not be violated, and no warrants shall issue, but upon
58 probable cause, supported by oath or affirmation, and
59 particularly describing the place to be searched, and the
60 persons or things to be seized.
61
62 </p><p><a name="5"><strong><h3>Amendment V</h3></strong></a>
63
64 <p></p><p>No person shall be held to answer for a capital, or otherwise
65 infamous crime, unless on a presentment or indictment of a grand
66 jury, except in cases arising in the land or naval forces,
67 or in the militia, when in actual service in time of war
68 or public danger; nor shall any person be subject for the
69 same offense to be twice put in jeopardy of life or limb;
70 nor shall be compelled in any criminal case to be a witness
71 against himself, nor be deprived of life, liberty, or property,
72 without due process of law; nor shall private property be
73 taken for public use, without just compensation.
74
75 </p><p><a name="6"><strong><h3>Amendment VI</h3></strong></a>
76
77 <p></p><p>In all criminal prosecutions, the accused shall enjoy the right
78 to a speedy and public trial, by an impartial jury of the state
79 and district wherein the crime shall have been committed, which
80 district shall have been previously ascertained by law, and
81 to be informed of the nature and cause of the accusation;
82 to be confronted with the witnesses against him; to have
83 compulsory process for obtaining witnesses in his favor,
84 and to have the assistance of counsel for his defense.
85
86 </p><p><a name="7"><strong><h3>Amendment VII</h3></strong></a>
87
88 <p></p><p>In suits at common law, where the value in controversy shall
89 exceed twenty dollars, the right of trial by jury shall be
90 preserved, and no fact tried by a jury, shall be otherwise
91 reexamined in any court of the United States, than according
92 to the rules of the common law.
93
94 </p><p><a name="8"><strong><h3>Amendment VIII</h3></strong></a>
95
96 <p></p><p>Excessive bail shall not be required, nor excessive fines
97 imposed, nor cruel and unusual punishments inflicted.
98
99 </p><p><a name="9"><strong><h3>Amendment IX</h3></strong></a>
100
101 <p></p><p>The enumeration in the Constitution, of certain rights, shall
102 not be construed to deny or disparage others retained by the people.
103
104 </p><p><a name="10"><strong><h3>Amendment X</h3></strong></a>
105
106 <p></p>
```

캐시 파일을 열어 MISS일 때 web server로부터의 response가 저장된 것을 확인했다.

## 고찰

이번 Proxy 2-3 과제를 수행하면서, 시그널 처리와 HTTP 요청 구조에 대한 실질적인 이해를 얻을 수 있었다. 기존 Proxy 2-2 과제까지는 프록시서버와 클라이언트와의 통신에 집중했다면, 이번 구현에서는 프록시 서버가 오리진 웹 서버와도 직접 통신해야 하므로 클라이언트와 서버의 역할을 모두 수행해야 했다.

우선, 이번 과제에서 새롭게 적용한 기능 중 하나는 SIGALRM과 SIGCHLD 시그널 핸들러였다. 이론 시간에 배운 내용을 실제로 코드에 적용해보며, 비동기 이벤트 기반의 흐름 제어에 대해 더 구체적으로 체감할 수 있었다. 특히 오리진 서버로 요청을 보낸 뒤 일정 시간 안에 응답이 도착하지 않으면 자식 프로세스를 종료시키는 구조를 alarm()과 SIGALRM 시그널로 구현하면서, 단순한 프로세스 포크를 넘어서 시그널 기반의 비동기 제어 방식을 익힐 수 있었다.

또한, 응답 처리 과정에서 read()와 write()에 사용하는 버퍼의 크기가 충분하지 않으면 브라우저 화면에서 응답이 깨지거나, 콘솔 출력에 쓰레기 값이 포함되는 문제가 발생했다. 처음에는 응답이 잘못 왔다고 생각했지만, 응답 자체는 정상이었고 내가 설정한 버퍼가 Content-Length보다 작았기 때문에 발생한 문제였다. 이를 해결하기 위해 기존에 사용하던 BUF\_SIZE \* 4에서 BUF\_SIZE \* 10 수준으로 버퍼 크기를 확장했고, 이로 인해 응답 데이터가 잘리는 문제를 해결할 수 있었다. 실제 응답 길이를 정확히 계산하고 그에 맞는 크기를 확보하는 것으로 발전하고 싶다.

가장 어려웠던 부분은 오리진 서버에 HTTP 요청을 보낼 때의 요청 구성 방식이었다. 처음에는 브라우저로부터 받은 요청을 그대로 오리진 서버에 전송했지만, 응답이 오지 않거나 408 Timeout이라는 현상이 반복되었다. 구글링을 통해 다양한 자료를 찾아보면서 HTTP/1.0 형식으로 요청 라인을 다시 구성해야 한다는 것을 알게 되었고, 그에 따라 요청 라인을 GET /path HTTP/1.0 형식으로 수정하고 필수적인 헤더를 포함시키는 방식으로 코드를 변경했다. 단순한 요청 전달이 아니라, 서버가 인식할 수 있는 형식으로 정확히 재조립하는 작업이 필요하다는 점에서, HTTP 프로토콜의 다시 한 번 학습할 수 있었다.

## Reference

<https://kimjingyu.tistory.com/entry/proxy-%EC%84%9C%EB%B2%84-%EB%A7%8C%EB%93%A4%EA%B8%B0> – proxy server 만들기

시스템프로그래밍 이론 및 실습 자료 참고하였습니다.