

Relatório 04

Vinícius de Oliveira Peixoto Rodrigues (245294)

Setembro de 2022

Questão 1

Como o site do DCA redireciona para a versão HTTPS independentemente do request, foi usado aqui o site www.neverssl.com.

Item (a)

Request HEAD:

```
HTTP/1.1 200 OK
Date: Wed, 21 Sep 2022 12:50:27 GMT
Server: Apache/2.4.53 ()
Upgrade: h2,h2c
Connection: Upgrade
Last-Modified: Wed, 29 Jun 2022 00:23:33 GMT
ETag: "f79-5e28b29d38e93"
Accept-Ranges: bytes
Content-Length: 3961
Vary: Accept-Encoding
Content-Type: text/html; charset=UTF-8
```

Percebe-se que para o OPTIONS foi recebido o campo **Allow**, indicando os requests HTTP aceitos; foram também omitidos vários campos presentes no HEAD.

Item (b)

```
HTTP/1.1 200 OK
Date: Wed, 21 Sep 2022 12:54:27 GMT
Server: Apache/2.4.53 ()
Transfer-Encoding: chunked
Content-Type: message/http
```

```
2c
TRACE / HTTP/1.1
```

Host: www.neverssl.com

0

Percebe-se que, como esperado, o servidor recebeu o request (200 OK) e enviou de volta a mensagem que recebeu (TRACE / HTTP/1.1...).

Item (c)

```
HTTP/1.1 200 OK
Date: Wed, 21 Sep 2022 12:57:13 GMT
Server: Apache/2.4.53 ()
Upgrade: h2,h2c
Connection: Upgrade
Last-Modified: Wed, 29 Jun 2022 00:23:33 GMT
ETag: "f79-5e28b29d38e93"
Accept-Ranges: bytes
Content-Length: 3961
Vary: Accept-Encoding
Content-Type: text/html; charset=UTF-8
```

```
<html>
  <! CÓDIGO-FONTE DA PÁGINA INICIAL >
</html>
```

Como o request de POST foi enviado vazio, o servidor retornou de volta a página inicial (o recurso pro qual mandamos o POST, localizado em /).

Item (d)

```
HTTP/1.1 405 Method Not Allowed
Date: Wed, 21 Sep 2022 13:03:25 GMT
Server: Apache/2.4.53 ()
Allow: OPTIONS,HEAD,GET,POST,TRACE
Content-Length: 223
Content-Type: text/html; charset=iso-8859-1
```

```
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
<html><head>
<title>405 Method Not Allowed</title>
</head><body>
<h1>Method Not Allowed</h1>
<p>The requested method DELETE is not allowed for this URL.</p>
</body></html>
```

O servidor retornou 405 `Method Not Allowed`, como era de se esperar (visto que o request `DELETE` não estava listado como disponível no `OPTIONS` do item (a)).

Item (e)

```
HTTP/1.1 501 Not Implemented
Date: Wed, 21 Sep 2022 13:05:01 GMT
Server: Apache/2.4.53 ()
Allow: OPTIONS,HEAD,GET,POST,TRACE
Content-Length: 208
Connection: close
Content-Type: text/html; charset=iso-8859-1

<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
<html><head>
<title>501 Not Implemented</title>
</head><body>
<h1>Not Implemented</h1>
<p>AGORA_VAI not supported for current URL.<br />
</p>
</body></html>
```

O servidor retornou um 501 `Not Implemented`, também conforme o esperado, visto que o `AGORA_VAI` não é um request válido do HTTP.

Questão 2

A linha `Host` serve para se referir a hosts *virtuais* (i.e., configurados na máquina host para a qual mandamos o request). Caso o servidor não possua hosts virtuais, a linha não faz efeito:

```
> telnet www.neverssl.com 80
Trying 34.223.124.45...
Connected to www.neverssl.com.
Escape character is '^]'.
GET / HTTP/1.1
Host: www.example.com

HTTP/1.1 200 OK
Date: Wed, 21 Sep 2022 13:14:32 GMT
Server: Apache/2.4.53 ()
Upgrade: h2,h2c
Connection: Upgrade
Last-Modified: Wed, 29 Jun 2022 00:23:33 GMT
ETag: "f79-5e28b29d38e93"
```

```
Accept-Ranges: bytes
Content-Length: 3961
Vary: Accept-Encoding
Content-Type: text/html; charset=UTF-8
```

```
<html>
  <!-- CÓDIGO-FONTE DA PÁGINA INICIAL -->
</html>
```

```
> telnet www.dca.fee.unicamp.br 80
Trying 143.106.148.94...
Connected to www.dca.fee.unicamp.br.
Escape character is '^]'.
GET / HTTP/1.1
Host: www.unicamp.br
```

```
HTTP/1.1 302 Found
Date: Wed, 21 Sep 2022 13:17:49 GMT
Server: Apache
Location: https://www.dca.fee.unicamp.br/
Content-Length: 215
Content-Type: text/html; charset=iso-8859-1
```

```
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
<html><head>
<title>302 Found</title>
</head><body>
<h1>Found</h1>
<p>The document has moved <a href="https://www.dca.fee.unicamp.br/">here</a>.</p>
</body></html>
```

Questão 3

- (a) 200 OK: indica que não houveram erros com o request
- (b) 400 Bad Request: indica que houve algum erro (da parte do cliente) com o request, de modo que o servidor se recusa a responder
- (c) 401 Unauthorized: indica que as credenciais do cliente são inválidas
- (d) 403 Forbidden: indica que o cliente não tem privilégio suficiente para fazer um request
- (e) 404 Not Found: indica que um recurso não foi encontrado no servidor
- (f) 405 Method Not Allowed: indica que o servidor não aceita requests com o método utilizado

- (g) **500 Internal Server Error**: indica que houve um erro interno da parte do servidor que impossibilitou o processamento do request
- (h) **501 Not Implemented**: indica que o servidor não reconheceu um método (seja porque é inválido ou porque o servidor não tem suporte para ele)
- (i) **Accept-Ranges**: especifica se o servidor suporta enviar um range específico de bytes do resource solicitado (útil para fazer download de um arquivo em chunks, por exemplo)
- (j) **Allow**: lista os métodos suportados por um recurso
- (k) **ETag (Entity Tag)**: indica a versão específica de um recurso
- (l) **X-Pad**: um header que faz um padding da resposta do servidor como uma forma de contornar um bug relacionado ao tamanho do response header em alguns browsers antigos
- (m) **Content-Length**: o tamanho da mensagem de resposta, em bytes
- (n) **Content-Type**: indica o tipo do recurso (normalmente usado para mídia)
- (o) **Date**: indica a data em que a mensagem de resposta foi criada
- (p) **Last-Modified**: indica a última data em que um recurso foi modificado (útil para determinar se é necessário fazer refetch de um recurso ou se é possível usar a versão armazenada em cache)
- (q) **Server**: indica o software utilizado pelo servidor que gerou a resposta
- (r) **Transfer-Encoding**: especifica o encoding utilizado no corpo da resposta
- (s) **WWW-Authenticate**: define os métodos de autenticação utilizados para ganhar acesso a um recurso
- (t) **Chunked**: modo de transferência em stream onde um recurso grande é enviado em chunks (blocos pequenos) em vez de por completo

Questão 4

```
> telnet www.example.com 80
DELETE / HTTP/1.1
Host: www.example.com

HTTP/1.1 405 Method Not Allowed
Cache-Control: max-age=604800
Content-Type: text/html; charset=UTF-8
Date: Wed, 21 Sep 2022 13:47:24 GMT
Expires: Wed, 28 Sep 2022 13:47:24 GMT
Server: EOS (vny/0453)
```

Content-Length: 0

GET /aljkdsajsd HTTP/1.1
Host: www.example.com

HTTP/1.1 404 Not Found
Cache-Control: max-age=604800
Content-Type: text/html; charset=UTF-8
Date: Wed, 21 Sep 2022 13:48:33 GMT
Expires: Wed, 28 Sep 2022 13:48:33 GMT
Server: EOS (vny/044E)
Vary: Accept-Encoding
Content-Length: 1256

klasjdasd

HTTP/1.0 501 Not Implemented
Content-Type: text/html
Content-Length: 357
Connection: close
Date: Wed, 21 Sep 2022 13:49:34 GMT
Server: ECSF (mic/9AE2)

GET / HTTP/1.1

HTTP/1.1 400 Bad Request
Content-Type: text/html
Content-Length: 349
Connection: close
Date: Wed, 21 Sep 2022 13:50:12 GMT
Server: ECSF (mic/9B4F)

GET / HTTP/123.2

HTTP/1.0 505 HTTP Version Not Supported
Content-Type: text/html
Content-Length: 379
Connection: close
Date: Wed, 21 Sep 2022 14:01:21 GMT
Server: ECSF (mic/9AF8)

Questão 5

O servidor retorna o seguinte:

GET / HTTP/1.1

Host: www.dca.fee.unicamp.br

HTTP/1.1 302 Found

Date: Wed, 21 Sep 2022 14:11:47 GMT

Server: Apache

Location: https://www.dca.fee.unicamp.br/

Content-Length: 215

Content-Type: text/html; charset=iso-8859-1

```
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
```

```
<html><head>
```

```
<title>302 Found</title>
```

```
</head><body>
```

```
<h1>Found</h1>
```

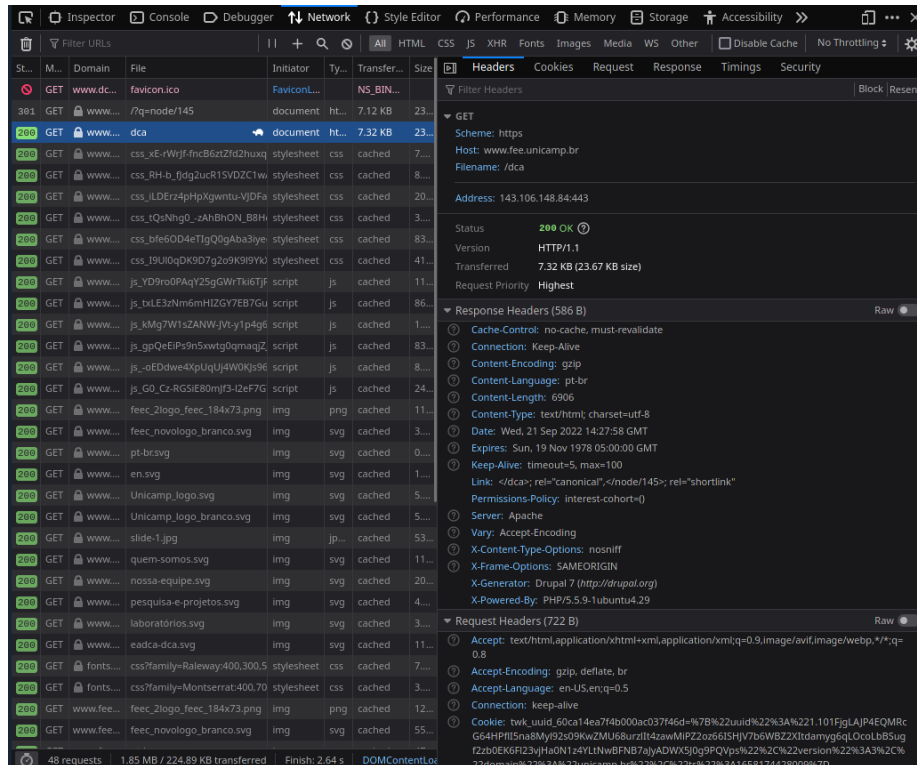
```
<p>The document has moved <a href="https://www.dca.fee.unicamp.br/">here</a>.</p>
```

```
</body></html>
```

De modo a indicar para o navegador que o recurso foi movido para um outro endereço que só aceita o protocolo HTTPS.

Item (b)

Como não é possível se comunicar via telnet com o servidor para fazer os vários GET que montam a página completa, estou deixando documentado aqui os vários requests de recursos diferentes na página que o navegador faz para montar a versão final:



The screenshot shows the Network tab of a web browser's developer tools. The top panel lists 48 requests, including various CSS files, JavaScript files, and images. The bottom panel shows the details of a selected GET request to `www.fee.unicamp.br/dca`. The status is 200 OK, and the response headers include `Cache-Control: no-cache, must-revalidate`, `Connection: Keep-Alive`, `Content-Encoding: gzip`, `Content-Language: pt-br`, `Content-Length: 6906`, `Content-Type: text/html; charset=utf-8`, `Date: Wed, 21 Sep 2022 14:27:58 GMT`, `Expires: Sun, 19 Nov 1978 05:00:00 GMT`, `Keep-Alive: timeout=5, max=100`, `Link: </dca>; rel="canonical", </node145>; rel="shortlink"`, `Permissions-Policy: interest-cohort=()`, `Server: Apache`, `Vary: Accept-Encoding`, `X-Content-Type-Options: nosniff`, `X-Frame-Options: SAMEORIGIN`, `X-Generator: Drupal 7 (http://drupal.org)`, and `X-Powered-By: PHP/5.5.9-1ubuntu4.29`. The request headers show `Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image/webp,*/*;q=0.8`, `Accept-Encoding: gzip, deflate, br`, `Accept-Language: en-US;q=0.5`, `Connection: keep-alive`, and a long cookie string.

St...	M...	Domain	File	Initiator	Ty...	Transfer...	Size
GET	www.d...		favicon.i...		NS_BIN...		
301	GET	www...	/q=node/145	document	ht...	7.12 KB	23...
200	GET	www...	dca	document	ht...	7.32 KB	23...
200	GET	www...	css_xE-rWjJf-fmcB6ztZfd2huxq	stylesheet	css	cached	7...
200	GET	www...	css_RH-b_fdg2ucR15VDZC1w	stylesheet	css	cached	8...
200	GET	www...	css_iLDERz4pHpXgwntu-VJdFa	stylesheet	css	cached	20...
200	GET	www...	css_tQaNHg0_zAhBhON_B8Hi	stylesheet	css	cached	3...
200	GET	www...	css_bfe6QD4eTlQ0gAba3Iye	stylesheet	css	cached	83...
200	GET	www...	css_l9UIdQK9D7g2o9K9I9Yk	stylesheet	css	cached	41...
200	GET	www...	js_YD9ro0PAqY25gGwTki6TjF	script	js	cached	11...
200	GET	www...	js_tLLE3zNm6mHIZGY7EB7G4	script	js	cached	86...
200	GET	www...	js_kMg7W1sZANW3JVe-y1p4g6	script	js	cached	1...
200	GET	www...	js_gpQeEIPs9n5xwtg0qmaqjZ	script	js	cached	83...
200	GET	www...	js_eEdDwe4XpUqUj4W0KjS96	script	js	cached	8...
200	GET	www...	js_G0_Cz-RGSE80mfJ3-I2eF7G	script	js	cached	24...
200	GET	www...	feec_2logo_feec_184x73.png	img	png	cached	11...
200	GET	www...	feec_novologo_branco.svg	img	svg	cached	3...
200	GET	www...	pt-br.svg	img	svg	cached	0...
200	GET	www...	en.svg	img	svg	cached	1...
200	GET	www...	Unicamp_logo.svg	img	svg	cached	5...
200	GET	www...	Unicamp_logo_branco.svg	img	svg	cached	5...
200	GET	www...	slide-1.jpg	img	jp...	cached	53...
200	GET	www...	quem-somos.svg	img	svg	cached	11...
200	GET	www...	nossa-equipe.svg	img	svg	cached	20...
200	GET	www...	pesquisa-e-projetos.svg	img	svg	cached	4...
200	GET	www...	laboratorios.svg	img	svg	cached	3...
200	GET	www...	eadca-dca.svg	img	svg	cached	11...
200	GET	fonts...	css?family=Raleway:400,300,5	stylesheet	css	cached	7...
200	GET	fonts...	css?family=Montserrat:400,70	stylesheet	css	cached	3...
200	GET	www.fee...	feec_2logo_feec_184x73.png	img	png	cached	12...
200	GET	www.fee...	feec_novologo_branco.svg	img	svg	cached	55...

48 requests | 1.85 MB / 224.89 KB transferred | Finish: 2.64 s | DOMContentLoaded

Questão 6

A estrutura do parser do laboratório passado é significativamente diferente dos requisitos desse laboratório, de modo que foi feita uma adaptação um tanto "hackish" para conseguir parsear requests HTTP. As mudanças mais significativas foram:

1. O lexer trata a primeira linha de forma especial, enviando para o parser todos os tokens nela e, ao encontrar um newline, envia um token especial `REQUEST_DESC_END`, de modo a sinalizar que os tokens pertencem à request line. Para todas as linhas subsequentes, é os newlines são reportados normalmente.
2. O parser, por sua vez, guarda o valor do request em si (i.e., `GET`, `POST`, etc) como um comando, e todos os outros tokens em seguida (caminho do recurso, versão do protocolo HTTP, etc) como argumentos
3. Além disso, foi adicionada uma regra no parser para guardar comandos na forma `<token>:<token>` como um comando só

Segue o resultado de dois requests enviados pelo Firefox e pelo Chromium:

```
GET / HTTP/1.1
Host: localhost:5555
User-Agent: Mozilla/5.0 (X11; Linux x86_64; rv:103.0)
          Gecko/20100101 Firefox/103.0
Accept: text/html,application/xhtml+xml,application/
        xml;q=0.9,image/avif,image/webp,*/*;q=0.8
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate, br
Connection: keep-alive
Upgrade-Insecure-Requests: 1
Sec-Fetch-Dest: document
Sec-Fetch-Mode: navigate
Sec-Fetch-Site: none
Sec-Fetch-User: ?1
```

```
GET / HTTP/1.1
Host: localhost:5555
Connection: keep-alive
Cache-Control: max-age=0
sec-ch-ua: " Not A;Brand";v="99", "Chromium";v="104"
sec-ch-ua-mobile: ?0
sec-ch-ua-platform: "Linux"
Upgrade-Insecure-Requests: 1
User-Agent: Mozilla/5.0 (X11; Linux x86_64)
          AppleWebKit/537.36 (KHTML, like Gecko) Chrome
          /104.0.5112.101 Safari/537.36
```

```

Accept: text/html,application/xhtml+xml,application/
      xml;q=0.9,image/avif,image/webp,image/apng,*/*;q
      =0.8,application/signed-exchange;v=b3;q=0.9
Sec-Fetch-Site: none
Sec-Fetch-Mode: navigate
Sec-Fetch-User: ?1
Sec-Fetch-Dest: document
Accept-Encoding: gzip, deflate, br
Accept-Language: en-US,en;q=0.9

```

E em seguida o teste:

```

$ ./run-tests.sh
*** request_chromium.txt ***
command: "GET / HTTP/1.1", args: (none)
command: "Host", args: " localhost:5555"
command: "Connection", args: " keep-alive"
command: "Cache-Control", args: " max-age=0"
command: "Sec-ch-ua", args: " " Not A;Brand";v="99", " "Chromium";v="104""
command: "Sec-ch-ua-mobile", args: " ?0"
command: "Sec-ch-ua-platform", args: " "Linux""
command: "Upgrade-Insecure-Requests", args: " 1"
command: "User-Agent", args: " Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, " like Gecko) Chrome/104.0.5112.101 Safari/5
37.36"
command: "Accept", args: " text/html", "application/xhtml+xml", "application/xml;q=0.9", "image/avif", "image/webp", "image/apng", "*/
;q=0.8", "application/signed-exchange;v=b3;q=0.9"
command: "Sec-Fetch-Site", args: " none"
command: "Sec-Fetch-Mode", args: " navigate"
command: "Sec-Fetch-User", args: " ?1"
command: "Sec-Fetch-Dest", args: " document"
command: "Accept-Encoding", args: " gzip", " deflate", " br"
command: "Accept-Language", args: " en-US", "en;q=0.9"

*** request_firefox.txt ***
command: "GET / HTTP/1.1", args: (none)
command: "Host", args: " localhost:5555"
command: "User-Agent", args: " Mozilla/5.0 (X11; Linux x86_64; rv:103.0) Gecko/20100101 Firefox/103.0"
command: "Accept", args: " text/html", "application/xhtml+xml", "application/xml;q=0.9", "image/avif", "image/webp", "*/;q=0.8"
command: "Accept-Language", args: " en-US", "en;q=0.5"
command: "Accept-Encoding", args: " gzip", " deflate", " br"
command: "Connection", args: " keep-alive"
command: "Upgrade-Insecure-Requests", args: " 1"
command: "Sec-Fetch-Dest", args: " document"
command: "Sec-Fetch-Mode", args: " navigate"
command: "Sec-Fetch-Site", args: " none"
command: "Sec-Fetch-User", args: " ?1"

```