

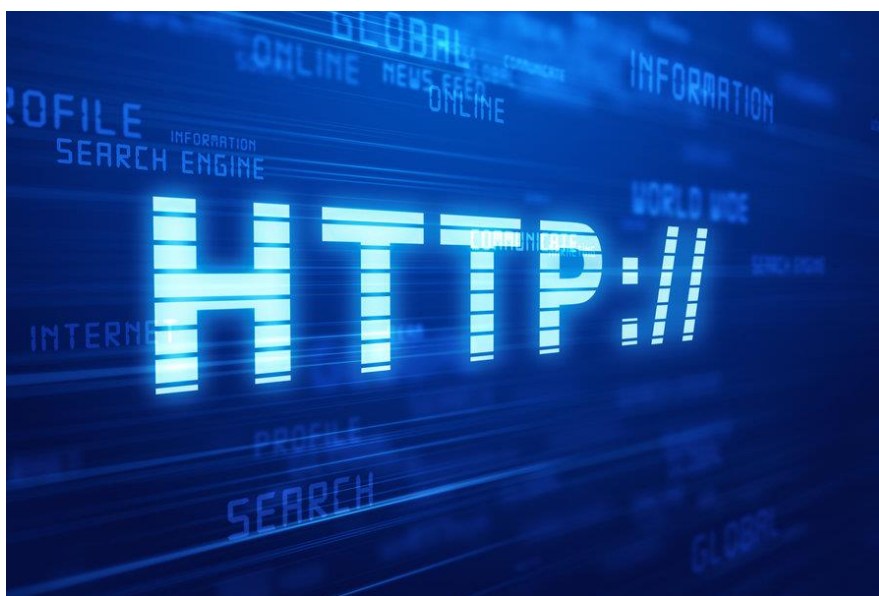


Universidade do Minho
Escola de Engenharia

Comunicações e Redes

TP5

Análise do protocolo HTTP



João Miguel da Silva Alves (83624)

Paulo Jorge Alves (84480)

MESTRADO INTEGRADO EM ENGENHARIA BIOMÉDICA
INFORMÁTICA MÉDICA 2020/2021

HTTP (HyperText Transfer Protocol) é um protocolo da camada de aplicação do TCP/IP cuja função é de proporcionar a transferência de hipertexto. É implementado em dois programas: Cliente e Servidor. É o HTTP quem define a estrutura da mensagem que o cliente vai trocar com o servidor e utiliza TCP como seu protocolo de transporte. Além disso, o HTTP não mantém memória sobre as suas ações, ou seja, se um cliente fizer uma requisição idêntica a uma anterior a qualquer momento, o HTTP não sabe informar sobre esse histórico.^[2]

Os passos para uma comunicação HTTP são:^[4]

- cliente inicia conexão TCP (cria socket) com servidor, porta 80.
- servidor aceita conexão TCP do cliente
- mensagens HTTP (do protocolo da camada de aplicação) trocadas entre navegador (cliente HTTP) e servidor Web (servidor HTTP)
- conexão TCP fechada

No início, o HTTP usava um único modelo para lidar com estas conexões. Um modelo de curta duração em que, no máximo, é enviado um objeto por cada conexão TCP, ou seja, é criada uma cada vez que uma mensagem é enviada e esta é fechada mal a mensagem seja recebida. A este modelo dá-se o nome de HTTP não persistente (figura 1). Mais tarde, visto como uma melhoria ao método de conexão anterior, surgiu o HTTP persistente (figura 2), onde múltiplas mensagens podem ser enviadas numa única conexão TCP entre cliente e servidor. Ao contrário do HTTP não persistente, este modelo reduz o número de RTT's (tempo para um pacote ir do cliente ao servidor e retornar), uma vez que mantém as conexões abertas entre mensagens sucessivas.^[1]

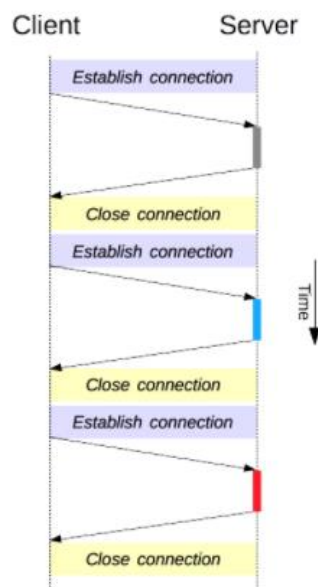


Figura 2 - HTTP não persistente

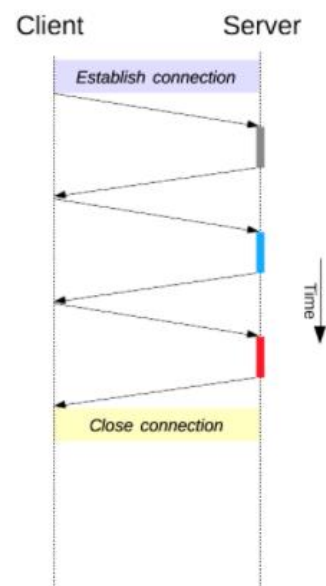


Figura 1 - HTTP persistente

O modelo original de HTTP é o padrão HTTP/1.0 (conexões de curta duração). Cada solicitação HTTP é concluída na sua própria conexão. Este modelo tem um grande problema: o tempo necessário para estabelecer uma nova conexão é significativo. Para aliviar este problema, surgiu o HTTP/1.1 (conexão *keep-alive*). É uma conexão persistente que permanece aberta durante um período de tempo (especificado pelo *keep_alive*) que pode ser usada para vários pedidos. ^[1]

Embora, atualmente, o HTTP/2.0 adicione mais complexidade, embutindo mensagens HTTP em quadros para melhorar a performance, a estrutura básica das mensagens continua a mesma desde o HTTP/1.0. ^[3]

1. Aceda ao servidor gcom.di.uminho.pt, na porta 80, utilizando telnet <destination-server> 80, usando HTTP / 1.0:

- a) Obtenha o arquivo `index.html` e a partir da resposta retornada identifique a versão http do servidor, a identificação do daemon de software e data do serviço.

```
paulojorge@PAULOJ-PC:/mnt/c/Users/paulo$ telnet gcom.di.uminho.pt 80
Trying 193.136.9.240...
Connected to gcom.di.uminho.pt.
Escape character is '^]'.
GET /index.html HTTP/1.0

HTTP/1.1 200 OK
Date: Sat, 21 Nov 2020 22:41:20 GMT
Server: Apache/2.2.15 (CentOS)
Last-Modified: Tue, 03 Jul 2012 20:54:01 GMT
ETag: "60932-5cb-4c3f31d300c40"
Accept-Ranges: bytes
Content-Length: 1483
Connection: close
Content-Type: text/html
```

VERSÃO DO HTTP: HTTP 1.0 (apesar de o software ter capacidade na versão 1.1)

DAEMON DE SOFTWARE: Apache/2.2.15 (CentOS)

DATA DO SERVIÇO: 3 de julho de 2012 22:54:01 GMT

- b) Tente obter o arquivo `maybe-index2.html` e interpretar a mensagem de resposta.

```
paulojorge@PAULOJ-PC:/mnt/c/Users/paulo$ telnet gcom.di.uminho.pt 80
Trying 193.136.9.240...
Connected to gcom.di.uminho.pt.
Escape character is '^]'.
GET /maybe-index2.html HTTP/1.0

HTTP/1.1 404 Not Found
Date: Sat, 21 Nov 2020 23:12:53 GMT
Server: Apache/2.2.15 (CentOS)
Content-Length: 296
Connection: close
Content-Type: text/html; charset=iso-8859-1
```

A resposta obtida indica que o ficheiro não foi encontrado na localização dada, ou seja, o ficheiro não existe nesse caminho. Contudo, o ficheiro pode existir noutra localização distinta (servidor diferente).

- c) Obtenha o arquivo / figura UMEnglog.jpg (conforme referenciado em index.html);

```

```

No *index.html* conseguimos encontrar o caminho para a imagem, “UMEnglog.jpg”. Com este caminho, obtém-se:

[illegible]

Este é o aspeto de parte da imagem vista no terminal. Não a conseguimos visualizar porque para isso teríamos de recorrer a um interpretador *jpeq*.

2. Aceda ao servidor `gcom.di.uminho.pt` na porta 80 e force a utilização de HTTP / 1.1 para obter o arquivo `index.html` e obter a figura `UMEnglogo.jpg`.

```
paulojorge@PAULO3-PC:/mnt/c/Users/paulo$ telnet gcom.di.uminho.pt 80
Trying 193.136.9.240...
Connected to gcom.di.uminho.pt.
Escape character is '^]'.
GET /index.html HTTP/1.1
Host: gcom.di.uminho.pt

HTTP/1.1 200 OK
Date: Sat, 21 Nov 2020 23:51:32 GMT
Server: Apache/2.2.15 (CentOS)
Last-Modified: Tue, 03 Jul 2012 20:54:01 GMT
ETag: "60932-5cb-4c3f31d300c40"
Accept-Ranges: bytes
Content-Length: 1483
Content-Type: text/html
```

[illegible]

As diferenças encontradas entre o protocolo HTTP/1.0 e HTTP/1.1 são o facto de o segundo ser persistente (existe uma espera para a realização de novos pedidos, tal como foi dito anteriormente), e a existência de um Host. Além disto, a possibilidade da utilização do parâmetro *KEEPALIVE*: “x” faz com que a ligação aguarde por mais pedidos “x” segundos, antes de fechar a ligação.

3. Agora acesse o servidor `marco.uminho.pt`. Obtenha o ficheiro `/index.html` em `marco.uminho.pt`, usando HTTP / 1.1 e Host: `marco.uminho.pt`

```
paulojorge@PAULOJ-PC:/mnt/c/Users/paulo$ telnet gcom.di.uminho.pt 80
Trying 193.136.9.240...
Connected to gcom.di.uminho.pt.
Escape character is '^]'.
GET /index.html HTTP/1.1
Host: marco.uminho.pt

HTTP/1.1 200 OK
Date: Sun, 22 Nov 2020 00:04:52 GMT
Server: Apache/2.2.15 (CentOS)
Last-Modified: Tue, 03 Jul 2012 20:54:01 GMT
ETag: "60932-5cb-4c3f31d300c40"
Accept-Ranges: bytes
Content-Length: 1483
Content-Type: text/html
```

Foi procurado o ficheiro `index.html` no servidor `marco.uminho.pt`. O ficheiro foi encontrado, e a data da última modificação diz respeito ao dia 3 de julho de 2012.

É o mesmo ficheiro encontrado no ponto 2, visto que o IP dos dois servidores é o mesmo.

```
paulojorge@PAULOJ-PC:/mnt/c/Users/paulo$ traceroute marco.uminho.pt
traceroute to marco.uminho.pt (193.136.9.240), 30 hops max, 60 byte packets
 1 PAULOJ-PC.mshome.net (172.26.80.1)  0.978 ms  0.892 ms  0.922 ms
 2 172.26.254.254 (172.26.254.254)  13.998 ms  13.985 ms  13.966 ms
 3 172.16.2.1 (172.16.2.1)  13.905 ms  12.920 ms  13.916 ms
 4 * 172.16.115.252 (172.16.115.252)  13.863 ms  13.869 ms
 5 marco.uminho.pt (193.136.9.240)  13.814 ms !X  13.842 ms !X  13.823 ms !X
paulojorge@PAULOJ-PC:/mnt/c/Users/paulo$ traceroute gcom.di.uminho.pt
traceroute to gcom.di.uminho.pt (193.136.9.240), 30 hops max, 60 byte packets
 1 PAULOJ-PC.mshome.net (172.26.80.1)  0.955 ms  0.902 ms  0.904 ms
 2 172.26.254.254 (172.26.254.254)  17.326 ms  17.367 ms  17.358 ms
 3 172.16.2.1 (172.16.2.1)  17.272 ms  17.261 ms  17.201 ms
 4 172.16.115.252 (172.16.115.252)  17.342 ms  17.315 ms  17.314 ms
 5 marco.uminho.pt (193.136.9.240)  17.178 ms !X  17.167 ms !X  13.576 ms !X
```

4. Entre no servidor HTTP `gcom.di.uminho.pt` (use `telnet <destination-server> 80`).

Discuta os resultados obtidos.

- a) Obtenha o arquivo `index.html` usando:

- `GET /index.html HTTP/1.1`
- `Host: gcom.di.uminho.pt`
- `If-Modified-Since: Tue, 03 Jul 2012 19:50:00 GMT`

```
paulojorge@PAULOJ-PC:/mnt/c/Users/paulo$ telnet gcom.di.uminho.pt 80
Trying 193.136.9.240...
Connected to gcom.di.uminho.pt.
Escape character is '^]'.
GET /index.html HTTP/1.1
Host: gcom.di.uminho.pt
If-Modified-Since: Tue, 03 Jul 2012 19:50:00 GMT

HTTP/1.1 200 OK
Date: Sun, 22 Nov 2020 00:25:09 GMT
Server: Apache/2.2.15 (CentOS)
Last-Modified: Tue, 03 Jul 2012 20:54:01 GMT
ETag: "60932-5cb-4c3f31d300c40"
Accept-Ranges: bytes
Content-Length: 1483
Content-Type: text/html
```

Ao aplicar a condição “If-Modified-Since: Tue, 03 Jul 2012 19:50:00 GMT” verifica-se que a resposta é idêntica às respostas obtidas nas perguntas anteriores. O “If” faz com que apenas se transfira o ficheiro se este sofreu alguma alteração depois da data pretendida.

Uma vez que, a última modificação ocorreu a 3 de julho de 2012 pelas 20:54:01 GMT, depois da hora usada no “If”, então ocorreu a transferência do ficheiro.

b) Repetir a obtenção do arquivo index.html modificando a data.

- GET /index.html HTTP / 1.1
- Host: gcom.di.uminho.pt
- If-Modified-Since: Tue, 04 Jul 2012 20:20:20 GMT

```
paulojorge@PAULOJ-PC:/mnt/c/Users/paulo$ telnet gcom.di.uminho.pt 80
Trying 193.136.9.240...
Connected to gcom.di.uminho.pt.
Escape character is '^]'.
GET /index.html HTTP/1.1
Host: www.di.uminho.pt
If-Modified-Since: Tue, 04 Jul 2012 20:20:20 GMT

HTTP/1.1 304 Not Modified
Date: Sun, 22 Nov 2020 00:32:30 GMT
Server: Apache/2.2.15 (CentOS)
ETag: "60932-5cb-4c3f31d300c40"
```

Neste segundo caso, aplicando a condição “If-Modified-Since: Tue, 04 Jul 2012 20:20:20 GMT”, verifica-se que o ficheiro não é transferido, isto porque, ao contrário do caso anterior, o ficheiro foi não foi modificado desde dia 4 de julho de 2012.

Bibliografia:

- [1] https://developer.mozilla.org/pt-BR/docs/Web/HTTP/Gerenciamento_de_Conex%C3%A3o_em_HTTP_1.x
- [2] https://pt.wikiversity.org/wiki/Introdu%C3%A7%C3%A3o_%C3%A0s_Redes_de_Computadores/WWW_e_HTTP
- [3] <https://developer.mozilla.org/pt-BR/docs/Web/HTTP/Overview>
- [4] https://gaia.cs.umass.edu/kurose_ross/ppt.htm