

# Relatório do Trabalho Final de Teleinformática e Redes 2

*Jônatas Rocha de Paiva – 140177043*

*Matheus Oliveira Braga – 140155571*

## I. Introdução

- Contextualização

O relatório contém uma breve explicação do que foi realizado no trabalho prático da disciplina de Teleinformática e Redes 2, incluindo, no capítulo 1, o problema abordado e o objetivo visado, em seguida, no capítulo 2, é explicitado como o software foi feito, sua arquitetura e utilização. No capítulo 3, valida-se os resultados por meio da experimentação, e por fim, no capítulo 4, tira-se as conclusões.

- Questão

O problema proposto neste trabalho é a construção de um inspetor de tráfego utilizando um servidor proxy, similar ao conceito de ataque man-in-the-middle, onde as requisições e respostas HTTP serão repassadas para um intermediário. E também, a implementação de um web crawler, que faz a indexação dos sites, comumente utilizado para manter a base de dados dos motores de busca atualizados ou para mineração de dados.

- Objetivo

Sendo um trabalho final de disciplina, seu objetivo é a consolidação dos conceitos estudados em sala, por meio da construção de um software que aborde o problema proposto, colocando a teoria em prática.

## II. Metodologia

- Procedimentos metodológicos.

Como a matéria é focada apenas na parte teórica de redes, foi feito primeiramente, um estudo da linguagem, quais os recursos, bibliotecas, arquiteturas e soluções genéricas poderiam ser usados na implementação.

- Arquitetura

Utilizando os recursos de modularização da linguagem C++, os módulos foram separados da seguinte forma:

- Strings: Para facilitar a manipulação de strings.
- HTML\_Parser: Para transformar um HTML em componentes utilizáveis.
- Request: Gerar uma requisição padrão HTTP 1.1.
- Response: Gerar uma resposta padrão HTTP.
- Spider: Realizar o crawler, dump e gerar o output.
- Proxy\_Server: Criação e gerência dos sockets e seu tráfego.

- Utilização

Considerando que o código foi feito baseado no Linux, deve-se rodar em um sistema operacional compatível. No terminal, entre na pasta do projeto e execute o comando

“make” para gerar os objetos com suas dependências, execute arquivo aracne com “./aracne” caso a porta de interesse seja a porta 8228, ou “./aracne -p porta” para selecionar uma porta diferente. Selecione 1 para inspecionar o tráfego, 2 para gerar a árvore hipertextual, 3 para gerar o dump, ou 4 para sair do programa.

### III. Resultados

- Exposição dos Resultados

Resultados da inspeção de tráfego:

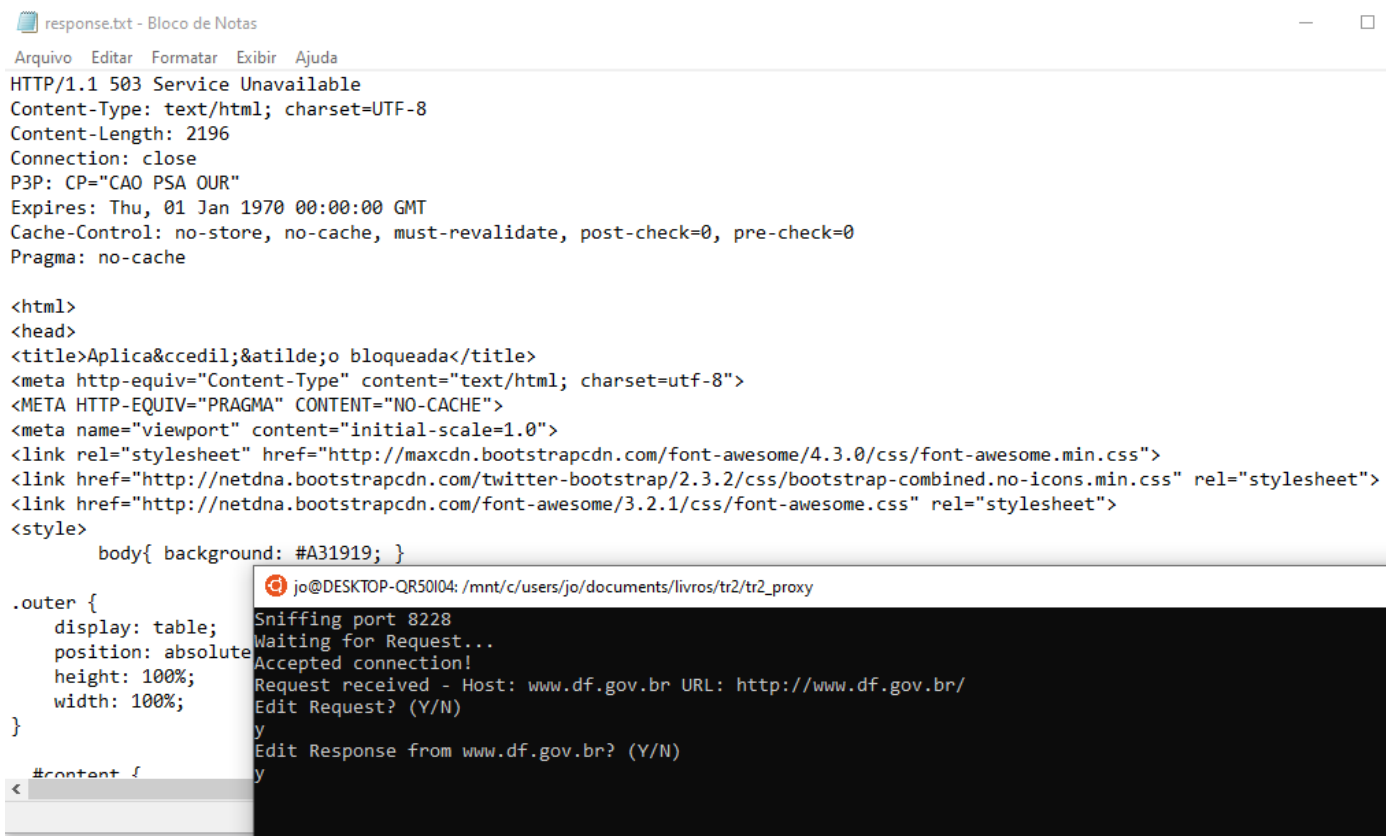


The image shows a Notepad window titled "request.txt - Bloco de Notas" with a menu bar (Arquivo, Editar, Formatar, Exibir, Ajuda). The text contains an HTTP GET request to "http://www.df.gov.br/" with various headers like "Accept-Encoding: identity", "Accept-Language: pt-BR,pt;q=0.8,en-US;q=0.5,en;q=0.3", "Accept: text/html,application/xhtml+xml,application/xml;q=0.9,\*/\*;q=0.8", "Connection: close", "Host: www.df.gov.br", "Upgrade-Insecure-Requests: 1", and "User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:67.0) Gecko/20100101 Firefox/67.0". Below the Notepad window is a terminal window with a black background and white text. It shows the command "jo@DESKTOP-QR50I04: /mnt/c/users/jo/documents/livros/tr2/tr2\_proxy", followed by "Sniffing port 8228", "Waiting for Request...", "Accepted connection!", "Request received - Host: www.df.gov.br URL: http://www.df.gov.br/", "Edit Request? (Y/N)", and the user input "y".

```
request.txt - Bloco de Notas
Arquivo  Editar  Formatar  Exibir  Ajuda
GET http://www.df.gov.br/ HTTP/1.1
Accept-Encoding: identity
Accept-Language: pt-BR,pt;q=0.8,en-US;q=0.5,en;q=0.3
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Connection: close
Host: www.df.gov.br
Upgrade-Insecure-Requests: 1
User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:67.0) Gecko/20100101 Firefox/67.0

jo@DESKTOP-QR50I04: /mnt/c/users/jo/documents/livros/tr2/tr2_proxy
Sniffing port 8228
Waiting for Request...
Accepted connection!
Request received - Host: www.df.gov.br URL: http://www.df.gov.br/
Edit Request? (Y/N)
y
```

Img1. Requisição



The image shows a Notepad window titled "response.txt - Bloco de Notas" with a menu bar (Arquivo, Editar, Formatar, Exibir, Ajuda). The text contains an HTTP 503 Service Unavailable response with headers like "Content-Type: text/html; charset=UTF-8", "Content-Length: 2196", "Connection: close", "P3P: CP="CAO PSA OUR"", "Expires: Thu, 01 Jan 1970 00:00:00 GMT", "Cache-Control: no-store, no-cache, must-revalidate, post-check=0, pre-check=0", and "Pragma: no-cache". Below the Notepad window is a terminal window with a black background and white text. It shows the command "jo@DESKTOP-QR50I04: /mnt/c/users/jo/documents/livros/tr2/tr2\_proxy", followed by "Sniffing port 8228", "Waiting for Request...", "Accepted connection!", "Request received - Host: www.df.gov.br URL: http://www.df.gov.br/", "Edit Request? (Y/N)", the user input "y", "Edit Response from www.df.gov.br? (Y/N)", and the user input "y".

```
response.txt - Bloco de Notas
Arquivo  Editar  Formatar  Exibir  Ajuda
HTTP/1.1 503 Service Unavailable
Content-Type: text/html; charset=UTF-8
Content-Length: 2196
Connection: close
P3P: CP="CAO PSA OUR"
Expires: Thu, 01 Jan 1970 00:00:00 GMT
Cache-Control: no-store, no-cache, must-revalidate, post-check=0, pre-check=0
Pragma: no-cache

<html>
<head>
<title>Aplica&ccedil;&atilde;o bloqueada</title>
<meta http-equiv="Content-Type" content="text/html; charset=utf-8">
<META HTTP-EQUIV="PRAGMA" CONTENT="NO-CACHE">
<meta name="viewport" content="initial-scale=1.0">
<link rel="stylesheet" href="http://maxcdn.bootstrapcdn.com/font-awesome/4.3.0/css/font-awesome.min.css">
<link href="http://netdna.bootstrapcdn.com/twitter-bootstrap/2.3.2/css/bootstrap-combined.no-icons.min.css" rel="stylesheet">
<link href="http://netdna.bootstrapcdn.com/font-awesome/3.2.1/css/font-awesome.css" rel="stylesheet">
<style>
  body{ background: #A31919; }

.outer {
  display: table;
  position: absolute;
  height: 100%;
  width: 100%;
}

#content {

jo@DESKTOP-QR50I04: /mnt/c/users/jo/documents/livros/tr2/tr2_proxy
Sniffing port 8228
Waiting for Request...
Accepted connection!
Request received - Host: www.df.gov.br URL: http://www.df.gov.br/
Edit Request? (Y/N)
y
Edit Response from www.df.gov.br? (Y/N)
y
```

Img2. Resposta

Resultados da árvore hipertextual:

```
URL crawled:
/www.ba.gov.br/ -->
-/cidadao
-/
-/cidadao
-/index.php/node/5639
-/investidor
-/node/1050
-/node/525
-/noticias/novo-sac-conquista-ii-inclui-servicos-do-procon-e-ciretran-e-amplia-capacidade-de
-/noticias/procon-movel-realiza-atendimentos-na-arena-fonte-nova-ate-sexta-feira
-/noticias/rede-sac-leva-atendimento-catorze-municipios-baianos-em-setembro
-/noticias/rede-sac-nao-funcionara-neste-sabado
-/paginas_especificas/152
-/servidor
-/sitemap
-/themes/secom/favicon.ico
-/turismo
-/galeria-multimedia
-/
-/cidadao
-/galeria-multimedia/51
-/galeria-multimedia/52
-/galeria-multimedia/53
-/galeria-multimedia/biblioteca-anisio-teixeira-promove-curso-de-libras-para-ouvintes
-/galeria-multimedia/biblioteca-anisio-teixeira-promove-curso-de-libras-para-ouvintes-0
-/galeria-multimedia/desfile-alusivo-independencia-do-brasil-leva-mensagem-de-prevencao-ao-suicidio
-/galeria-multimedia/desfile-civico-marca-comemoracoes-ao-7-de-setembro
-/galeria-multimedia/governador-rui-costa-participa-de-desfile-oficial-do-7-de-setembro
-/galeria-multimedia/governo-entrega-plano-plurianual-participativo-2020-2023-para-analise-da-alba
-/galeria-multimedia/governo-entrega-plano-plurianual-participativo-2020-2023-para-analise-da-alba-0
-/galeria-multimedia/governo-investe-r-67-milhoes-em-aco-es-contra-incendios-florestais-na-bahia
-/galeria-multimedia/lancamento-do-transformae-reforca-protagonismo-estudantil
-/galeria-multimedia/protagonismo-estudantil-e-reforcado-com-o-lancamento-do-transformae
-/galeria-multimedia/rui-defende-justica-na-distribuicao-de-receitas-no-brasil
-/galeria-multimedia/setembro-verde-incentiva-doacao-de-orgaos-no-estado
-/investidor
-/node/1050
-/node/525
-/servidor
-/sitemap
-/themes/secom/favicon.ico
-/turismo
```

Ativar o V  
Navegar com

Img3. Crawler

- Validação

Os resultados obtidos não foram os ideais por não se apoiarem em uma interface gráfica porém, as funcionalidades básicas foram implementadas conforme as especificações.