Universidade Estadual da Paraíba
Departamento de Matemática, Estatística e Computação
Disciplina: Redes de Computadores
Professor: Edmar José do Nascimento

2º Exercício Prático: HTTP

Introdução

Vamos agora utilizar o Wireshark para analisar o funcionamento do protocolo HTTP. Serão observados a interação pedido/resposta, os formatos das mensagens, recuperação de grandes arquivos, obtenção de arquivos HTML com objetos embutidos e HTML com autenticação.

A interação HTTP pedido/resposta básica

Iniciemos explorando o acesso a um arquivo HTML que contém somente texto e, portanto é composto de só um objeto. Para isso siga os seguintes passos:

1. Inicie o navegador.
2. Inicie o Wireshark sem iniciar a captura de pacotes. Entre com *http* no filtro de pacotes, isso fará com que somente mensagens *http* sejam mostradas.
3. Espere alguns instantes (um minuto, por exemplo), e então inicie a captura de pacotes.
4. No seu navegador digite o seguinte endereço: http://gaia.cs.umass.edu/ethereal-labs/HTTP-ethereal-file1.html , seu navegador vai mostrar uma página web muito simples.
5. Pare a captura de pacotes.

Nesse momento você deverá observar uma página bem parecida com a que está sendo mostrada na Figura 1.

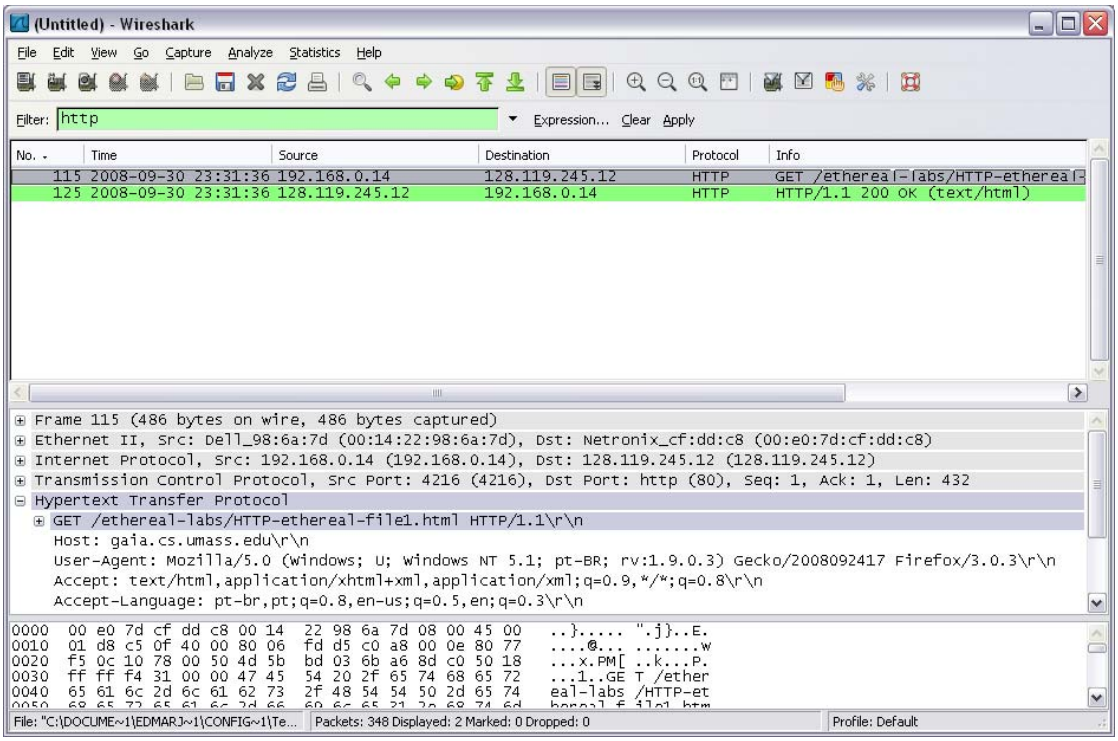

Figura 1 - Tela do Wireshark

O exemplo da Figura 1 mostra na tela de listagem de pacotes que duas mensagens HTTP foram capturadas: uma mensagem GET do navegador que está na máquina *192.168.0.14* para o servidor *gaia.cs.umass.edu* que está na máquina *128.119.245.12* e a reposta do servidor para o navegador.

Indo na janela de conteúdo dos pacotes e expandindo as mensagens *http*, para isso basta selecionar uma mensagem e clicar no símbolo + ao lado de *Hypertext Transfer Protocol*, e obtém-se a tela mostrada na Figura 1. Essa janela mostra os detalhes da mensagem selecionada (no caso uma mensagem *http* GET).

Observando a mensagem http de resposta, responda as seguintes questões:

1. Seu navegador usa a versão 1.0 ou 1.1 do *http*? Qual a versão do *http* que está rodando no servidor?
2. Quais linguagens o navegador indica que aceita?
3. Qual o endereço IP do seu computador? E do servidor?
4. Qual o código de status que o servidor retornou para o seu navegador?
5. Quando foi a última vez que o arquivo *html* que você baixou do navegador foi alterado?
6. Quantos bytes de conteúdo são retornados para o seu navegador?

A interação HTTP Condicional pedido/resposta

Muitos navegadores fazem *cache* de objetos e utilizam GET condicional quando estão solicitando objetos. Antes fazer os passos a seguir verifique que o cache do seu navegador está vazio. (Para fazer isso no Firefox vá a Ferramentas – Limpar dados pessoais – Limpar Cache) . Agora faça o seguinte:

a) Inicie o navegador,
b) Inicie o Wireshark
c) Entre o seguinte endereço no seu navegador http://gaia.cs.umass.edu/ethereal-labs/HTTP-ethereal-file2.html
   Seu navegador irá mostrar uma página web simples.
d) Rapidamente entre com o mesmo endereço (ou simplesmente faça o *refresh* (F5))

Para a captura de pacotes no Wireshark e digite *http* na janela de filtro e peça para aplicar. Responda as seguintes questões:

7. Inspecione o conteúdo do primeiro pedido HTTP GET do seu navegador para o servidor. Você vê uma linha "IF-MODIFIED-SINCE" nesse HTTP GET?
8. Inspecione o conteúdo da resposta do servidor. Ele retornou explicitamente o conteúdo do arquivo?
9. Agora inspecione o conteúdo da segunda mensagem HTTP GET do seu navegador para o servidor. Você vê uma linha "IF-MODIFIED-SINCE:" nesse HTTP GET? Se vê, qual informação segue o cabeçalho "IF-MODIFIED-SINCE:"?
10. Qual o código HTTP de status e a frase retornada pelo servidor em resposta ao segundo HTTP GET? O servidor retornou explicitamente o conteúdo do arquivo?

Obtendo Grandes Documentos

Nos exemplos anteriores trabalhamos com pequenos arquivos html, agora vamos observar a transferência de grandes arquivos *html*. Para isso faça o seguinte:

a) Inicie o navegador, limpe o *cache* como foi feito no exercício anterior.
b) Inicie a captura de pacotes no Wireshark
c) Entre no seu navegador com o seguinte endereço http://gaia.cs.umass.edu/ethereal-labs/HTTP-ethereal-file3.html
   Seu navegador irá mostrar uma página contendo um texto muito longo.
d) Pare a captura de pacotes no Wireshark, e entre com o filtro *http* para que somente as mensagens *http* sejam mostradas.

Na janela de listagem de pacotes você pode ver sua mensagem HTTP GET seguida por uma resposta de múltiplos pacotes. Relembre que a mensagem de resposta *http* consiste de uma linha de status, seguido pelas linhas de cabeçalho, seguido por uma linha em branco, seguido pelo corpo. No caso do nosso HTTP GET, o corpo na resposta é o arquivo HTML solicitado. Esse arquivo é muito grande para caber em um único pacote TCP, logo ele é enviado pelo servidor em vários pedaços, cada um deles segue em um pacote em separado.
Responda as seguintes questões:

11. Quantas mensagens HTTP GET forma enviadas pelo seu navegador?
12. Quantos segmentos TCP contendo dados foram necessários para transportar uma única resposta *http*?
13. Qual o código de status e a frase associados com a reposta ao pedido HTTP GET?

Documentos HTML com Objetos

Agora vamos observar a transferência de arquivos HTML com objetos, i.e., arquivos que contém imagens. Para tanto faça o seguinte:

a) Inicie o navegador e limpe o cache
b) Inicie a captura de pacotes no Wireshark
c) Entre com o seguinte endereço http://gaia.cs.umass.edu/ethereal-labs/HTTP-ethereal-file4.html
   Seu navegador deverá mostrar uma página HTML com algumas imagens. As imagens apresentadas são referenciadas no arquivo HTML, como foi visto seu navegador irá fazer o *download* dessas imagens.
d) Para a captura de pacotes e entre com http no filtro.

Responda as seguintes questões:

14. Quantas mensagens HTTP GET seu navegador enviou? Para quais endereços essas mensagens foram enviadas?
15. Você pode dizer se o seu navegador está fazendo o download das imagens em paralelo ou não? Explique.

Autenticação HTTP

Finalmente vamos visitar um site que é protegido por senha e verificar a seqüência de mensagens *http* trocadas com o site. Para tanto faça o seguinte:

a) Limpe o cache do seu navegador, feche-o e então reinicie.
b) Inicie a captura de pacotes no Wireshark
c) Vá ao site: http://gaia.cs.umass.edu/ethereal-labs/protected_pages/HTTP-ethereal-file5.html
d) Entre com o login **eth-students** e a senha **network**
e) Para a captura de pacotes com o Wireshark

Agora vamos examinar a saída. Inicialmente filtre para que somente as mensagens HTTP sejam mostradas, responda as seguintes questões:

16. Qual a reposta do servidor (código de status e frase) em resposta à mensagem http GET inicial?
17. Quando o seu navegador envia a mensagem http GET pela segunda vez, quais novos campos são incluídos nessa mensagem?

O login (eth-students) e a senha (network) que você forneceu são codificados na seqüência de caracteres (ZXRoLXN0dWRlbnRzOm5ldHdvcmtz) seguindo o cabeçalho "*Authorization: Basic*" na mensagem HTTP GET. Apesar da aparência, esses dados não estão criptografados, eles estão apenas codificados no formato Base64. Para ver isso, vá ao site http://www.securitystats.com/tools/base64.php e digite a cadeia ZXRoLXN0dWRlbnRzOm5ldHdvcmtz e pressione *decode*.