Лабораториска Вежба 1

HTTP Протокол

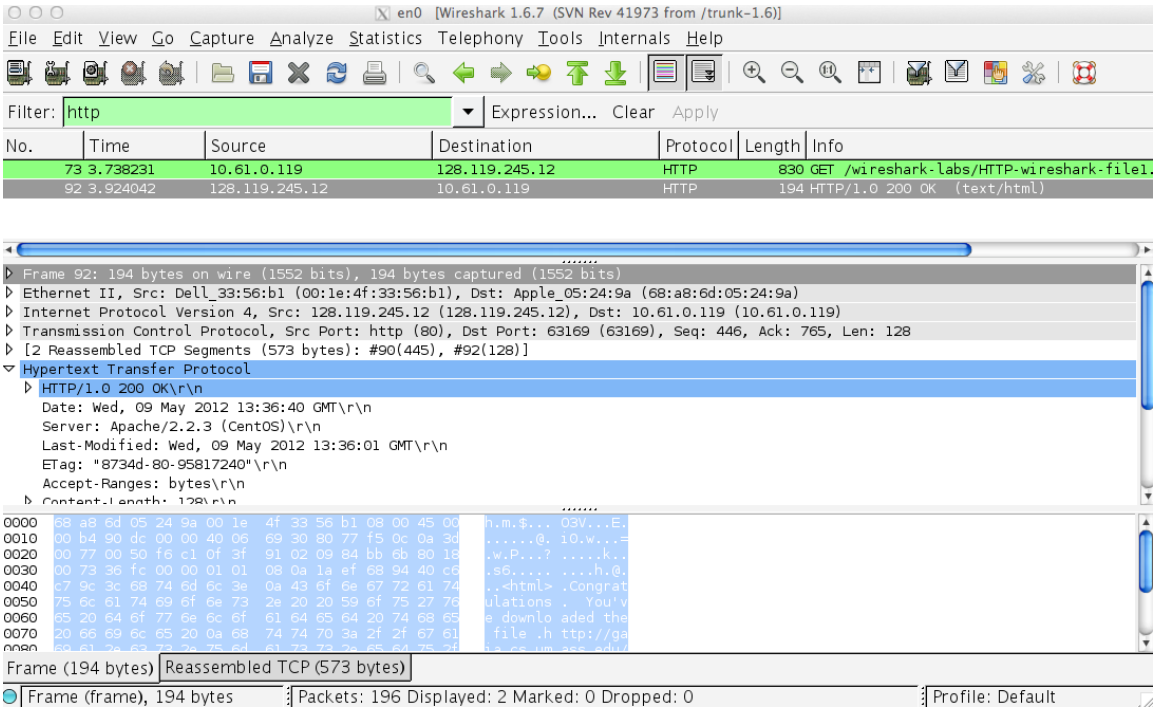
Во оваа лабораториска вежба, ќе разгледаме неколку аспекти на HTTP протоколот: основна интеракција со GET/одговор, формати на HTTP пораки, преземање големи HTML датотеки, преземање на HTML-датотеки со вградени објекти и автентикација и безбедност на HTTP. Пред да започнете со овие лабораториски вежби, можеби ќе сакате да го разгледате Дел 2.2 од книгата.

1. Основната HTTP GET/response интеракција

Да ги започнеме нашите истражувања на HTTP со преземање на многу едноставна HTML-датотека - таква која е многу кратка и не содржи вградени предмети. Направете го следново:

- Стартувајте го вашиот интернет пребарувач
- Стартувајте го Wireshark без да почнете со снимање на пакети. Внесете "http" во прозорецот за филтрирање на пакети, така што ќе се листаат само HTTP пакетите.
- Почекајте малку повеќе од една минута (ќе видиме зошто наскоро), а потоа започнете со фаќање на пакетите Wireshark.
- Внесете го следново во прелистувачот
  http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file1.html
  Вашиот прелистувач треба да ја прикаже многу едноставната HTML-датотека.
- Престанете да фаќате пакети во Wireshark.

Вашиот Wireshark прозорец треба да изгледа слично на прозорецот прикажан на Слика 1.



Слика 1: Wireshark прозорецот откако http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file1.html е преземен од вашиот прелистувач

Примерот на Слика 1 покажува во прозорецот за список со пакети дека се фатени две HTTP пораки: GET пораката (од вашиот прелистувач до веб-серверот gaia.cs.umass.edu) и пораката за одговор од серверот до вашиот прелистувач. Прозорецот со содржина на пакетите ги покажува деталите за избраната порака (во овој случај HTTP-пораката OK, која е истакната во прозорецот за список со пакети). Потсетете се дека бидејќи HTTP-пораката се носи во TCP-сегмент, кој се носи во IP-датаграм, кој е спроведен во рамка на Ethernet, Wireshark ги прикажува и Frame, Ethernet, IP и TCP деловите. Ние сакаме да ја минимизираме количината на не-HTTP прикажани податоци (ние сме заинтересирани за HTTP овде, и ќе ги испитуваме овие други протоколи во други лабораториски вежби), затоа проверете дали линиите од левата страна на Рамката, Етернет, IP и Информациите за TCP имаат знак плус или триаголник со десно-покажувач (што значи дека има скриени, необјаснети информации), а линијата HTTP има минус знак или триаголник надолу (што значи дека се прикажани сите информации за HTTP-пораката) .

(Забелешка: Треба да ги игнорирате сите HTTP GET и одговор за favicon.ico. Ако видите повикување на оваа датотека, вашиот прелистувач автоматски го прашува серверот дали тој (серверот) има мала икона што треба да се прикаже веднаш до прикажаната URL-адреса во прелистувачот. Ние ќе ги игнорираме упатствата за оваа досадна датотека во оваа лабораториска вежба.)

Гледајќи ги информациите во HTTP GET и пораките за одговор, одговорете на следниве прашања. Кога одговарате на следниве прашања, треба да ги отпечатите GET и response пораките (видете во упатството за Wireshark - "Вовед во Wireshark" за објаснување како да го направите ова) и да наведете каде во пораката сте пронашле информација што одговара на следниве прашања.

1. Дали вашиот прелистувач работи со HTTP верзија 1.0 или 1.1? Која верзија на HTTP работи на серверот?
2. На кои јазици (ако има) вашиот прелистувач означува дека може да прифати од серверот?
3. Која е IP адресата на вашиот компјутер? На серверот gaia.cs.umass.edu?
4. Кој е кодот на статусот вратен од серверот до вашиот прелистувач?
5. Кога последно е изменета на серверот HTML-датотеката што беше добиена?
6. Колку бајти со содржина се враќаат во прелистувачот?
7. Преку преглед на сурови податоци во прозорецот за содржина на пакети, дали гледате заглавија во податоците кои не се прикажани во прозорецот за огласување пакети? Ако е така, наведете еден.

Во вашиот одговор на прашањето 5 погоре, можеби сте изненадени кога откривте дека документот што штотуку го симнавте последен пат е изменет во рок од една минута пред да го преземете документот. Тоа е затоа што (за оваа конкретна датотека), серверот gaia.cs.umass.edu го поставува последното модифицирано време на датотеката како тековното време и тоа го прави еднаш во минута. Така, ако почекате една минута помеѓу пристапи, датотеката се чини дека е неодамна изменета, и оттука вашиот прелистувач ќе преземе „нова" копија на документот.

2. HTTP CONDITIONAL GET/одговор интеракција

Потсетете се од Дел 2.2.5 во книгата, дека повеќето веб прелистувачи вршат меморирање на објекти и на тој начин вршат условен GET при пребарување на HTTP објект. Пред да ги извршите чекорите подолу, проверете дали кешот на прелистувачот е празен. (За да го направите ова под Firefox, изберете Алатки-> Избриши ја неодамнешната историја и проверете го полето Кеш, или за Internet Explorer, изберете Алатки-> Интернет-опции-> Избриши датотека; овие активности ќе ги отстранат зачуваните датотеки од кешот на прелистувачот.) Сега направете го следното:

- Стартувајте го вашиот веб-прелистувач и проверете дали е исчистена кешот на прелистувачот, како што беше дискутирано погоре.
- Стартувајте го снимањето на пакети во Wireshark
- Внесете ја следнава URL-адреса во прелистувачот

http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file2.html
Вашиот прелистувач треба да прикаже многу едноставна датотека со пет линии HTML.

- Брзо внесете ја истата URL повторно во вашиот прелистувач (или едноставно изберете го копчето за освежување на прелистувачот)
- Запрете го фаќањето на пакетите Wireshark и внесете „http“ во прозорецот за спецификација на филтерот за приказ, така што само снимените HTTP-пораки ќе бидат прикажани во прозорецот за список со пакети.

Одговори ги следниве прашања:

8. Проверете ја содржината на првото барање HTTP GET од вашиот прелистувач до серверот. Дали гледате линија „IF-MODIFIED-SINCE“ во HTTP GET?
9. Проверете ја содржината на одговорот на серверот. Дали серверот експлицитно ја вратил содржината на датотеката? Како можеш да препознаеш?
10. Сега проверете ја содржината на второто барање за HTTP GET од вашиот прелистувач до серверот. Дали гледате линија „IF-MODIFIED-SINCE:“ во HTTP GET? Ако е така, кои информации следат по заглавието „IF-MODIFIED-SINCE:“?
11. Кој е кодот и фразата за статус на HTTP вратени од серверот како одговор на овој втор HTTP GET? Дали серверот експлицитно ја вратил содржината на датотеката? Објаснете.

3. Преземање на големи документи

Во досегашните примери, преземените документи беа едноставни и кратки HTML-датотеки. Ајде да видиме што се случува кога ќе преземеме долга HTML-датотека. Направете го следново:

- Стартувајте го вашиот веб-прелистувач и проверете дали е исчистен кешот на прелистувачот, како што беше дискутирано погоре.
- Стартувајте го снимањето на пакети во Wireshark
- Внесете ја следнава URL-адреса во прелистувачот
http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html
Вашиот прелистувач треба да го прикаже прилично долгиот Закон за права на САД.
- Запрете го снимањето на пакетот Wireshark и внесете „http“ во прозорецот за спецификација на филтерот за приказ, така што ќе се прикажат само снимените HTTP-пораки.

Во прозорецот за список со пакети, треба да ја видите вашата HTTP GET порака, проследена со одговор во повеќе TCP сегменти на вашето HTTP GET барање. Овој одговор од повеќе сегменти заслужува малку објаснување. Потсетете се од Дел 2.2 (видете слика 2.9 во текстот) дека пораката за HTTP одговор се состои од статусна линија, проследена со линии

од заглавја, проследено со празна линија, проследено со телото на ентитетот. Во случајот со нашиот HTTP GET, телото на субјектот како одговор е целата побарана HTML-датотека. Во нашиот случај овде, HTML датотеката е 4500 бајти, преголема за да се вклопи во еден TCP сегмент. Во тој случај, HTTP response пораката се разделува на неколку делови, при што секој дел се става во посебен TCP сегмент (види слика 1.24 во текстот). Во последните верзии на Wireshark, Wireshark го означува секој TCP сегмент како посебен пакет, а фактот дека единечниот HTTP одговор е фрагментиран низ повеќе TCP пакети е означен со "TCP segment of a reassembled PDU" (TCP-сегментот од склопени PDU). Претходните верзии на Wireshark ја користеа фразата „Продолжување" за да посочат дека целата содржина на HTTP-пораката е расцепкана во повеќе TCP сегменти. Овде потенцираме дека нема порака за „Продолжување" во HTTP!

Одговори ги следниве прашања:

12. Колку пораки за барање HTTP GET испрати вашиот прелистувач? Кој број на пакет во трагата ја содржи GET пораката за "the Bill of Rights"?
13. Кој број на пакет во трагата содржи код за статус и фраза поврзана со одговорот на барањето на HTTP GET?
14. Кој е кодот за статус и фразата во одговорот?
15. Колку TCP-сегменти што содржат податоци беа потребни за да се донесе единствениот одговор на HTTP и текстот за "the Bill of Rights"?

4. HTML Документи со Вграден Објекти

Сега кога видовме како Wireshark го прикажува снимениот пакет сообраќај за големи HTML-датотеки, можеме да погледнеме што се случува кога прелистувачот презема датотека со вградени објекти, т.е. датотека што вклучува други предмети (на пример подолу, датотеки со слики) што се чуваат на други сервери.

Направете го следново:

- Стартувајте го вашиот веб-прелистувач и проверете дали е исчистен кешот на прелистувачот, како што беше дискутирано погоре.
- Стартувајте го снимањето на пакети во Wireshark
- Внесете ја следнава URL-адреса во прелистувачот
  http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file4.html
  Вашиот прелистувач треба да прикаже кратка HTML-датотека со две слики. Овие две слики се повикуваат во основната HTML-датотека. Тоа значи, самите слики не се содржани во HTML; наместо тоа, URL-то за слики се содржани во преземената HTML-датотека. Како што беше дискутирано во учебникот, вашиот прелистувач ќе мора да ги преземе овие логоа од наведените веб-страници. Логото на нашиот издавач е преземено од веб-страницата gaia.cs.umass.edu. Сликата на насловната

страница за нашето 5-то издание е зачувано на серверот caite.cs.umass.edu. (Овие се два различни веб-сервери во cs.umass.edu).

- Запрете го снимањето на пакетот Wireshark и внесете „http“ во прозорецот за спецификација на филтерот за приказ, така што ќе се прикажат само снимените HTTP-пораки.

Одговори ги следниве прашања:

16. Колку пораки за барање HTTP GET испрати вашиот прелистувач? На кои Интернет адреси беа испратени овие GET барања?
17. Можете ли да кажете дали прелистувачот ги преземал двете слики сериски, или дали паралелно се преземени од двете веб-страници? Објаснете

5. HTTP Автентикација

Конечно, да се обидеме да посетиме веб-страница заштитена со лозинка и да ја испитаме низата на HTTP пораки, разменета за таква страница.

URL-то http://gaia.cs.umass.edu/wireshark-labs/protected_pages/HTTP-wireshark-file5.html е заштитено со лозинка. Корисничкото име е „wireshark-students“ (без наводници), а лозинката е „network“ (повторно, без наводници). Значи, да дозволиме пристап до оваа „безбедна“ страница заштитена со лозинка. Направете го следново:

- Осигурете се дека кешот на прелистувачот е исчистен, како што беше дискутирано погоре, и затворете го прелистувачот. Потоа, стартувајте го вашиот прелистувач
- Стартувајте трагач на пакетите Wireshark
- Внесете ја следнава URL-адреса во прелистувачот http://gaia.cs.umass.edu/wireshark-labs/protected_pages/HTTP-wireshark-file5.html Внесете ги бараните корисничко име и лозинка во полето што се појавува.
- Запрете го фаќањето на пакетите со Wireshark и внесете „http“ во прозорецот за спецификација на филтерот за приказ, така што само снимените HTTP-пораки ќе бидат прикажани подоцна во прозорецот за список со пакети.
- (Забелешка: Ако не можете да ја стартувате Wireshark на живо мрежна врска, можете да користите трага за пакетите http-ethereal-trace-5 за да одговорите на прашањата подолу; видете ја фуснотата 2. Оваа датотека во трага беше собрана при извршувањето на горенаведените чекори на еден од авторите.)

Сега да го разгледаме излезот од Wireshark. (Можеби прво сакате да прочитате за HTTP автентикација на http://frontier.userland.com/stories/storyReader$2159)

Одговори ги следниве прашања:

18. Кој е одговорот на серверот (код за статус и фраза) како одговор на првичната порака HTTP GET од вашиот прелистувач?
19. Кога вашиот прелистувач ќе ја испрати пораката HTTP GET по втор пат, кое ново поле е вклучено во пораката HTTP GET?

Корисничкото име (wireshark-students) и лозинката (network) што сте ги внеле се кодирани во низата карактери (d2lyZXNoYXJrLXN0dWRlbnRzOm5ldHdvcms=) по насловот „Authorization: Basic“ во HTTP GET пораката на клиентот. Иако се чини дека вашето корисничко име и лозинка се криптирани, тие едноставно се кодираат во формат познат како Base64. Корисничкото име и лозинката не се шифрирани! За да го видите ова, одете на http://www.motobit.com/util/base64-decoder-encoder.asp и внесете ја основната енкодирана низа d2lyZXNoYXJrLXN0dWRlbnRz и декодирајте ја. Вие преведовте од Base64 кодирање во ASCII кодирање, и со тоа треба да го видите вашето корисничко име! За да ја видите лозинката, внесете го остатокот од низата Om5ldHdvcms= и притиснете декодирање. Бидејќи секој може да преземе алатка како Wireshark и да снима пакети (не само негови) кои минуваат покрај мрежниот адаптер, и секој може да преведе од Base64 во ASCII (вие едноставно го сторивте тоа!), треба да ви биде јасно дека едноставните лозинки на WWW страниците не се безбедни, освен ако не се преземат дополнителни мерки.

Како што ќе видиме во Поглавје 8, постојат начини да се направи посигурен пристап до WWW. Сепак, јасно ќе ни треба нешто што ја надминува основната рамка за автентикација на HTTP!