

AAD

Sessió de problemes: HTTP i WEBDAV

(Laboratori. 1 hora)

Autors: Joan Manuel Marquès i Leandro Navarro.

Introducció

Aquesta sessió de problemes ens ajudarà a entendre el funcionament d'HTTP i WEBDAV.

Objectius

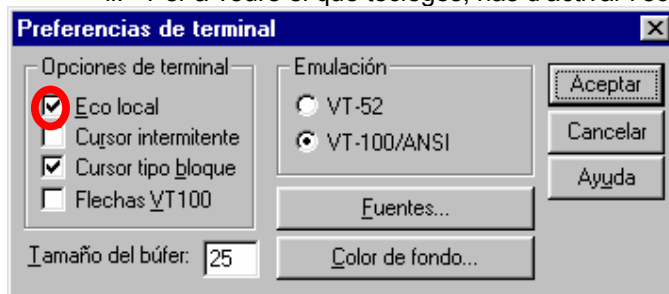
- Veure i entendre el funcionament dels missatges de petició i de resposta d'HTTP
- Veure el funcionament de WEBDAV
- Conèixer el funcionament bàsic d'un analitzador de xarxes per a examinar els missatges HTTP i WEBDAV

Tasques

HTTP

Per a veure el funcionament de l'HTTP ens connectarem via telnet al port 80 de servidors web. Un cop connectats podrem interactuar amb el servidor com si fossim un client web (navegador)

1. Arranca un client de telnet i connectat al port 80 de `www.upc.es`
 - a. Pots utilitzar qualsevol programa que faci telnet. Si utilitzes MS-Windows pots utilitzar el telnet que porta per defecte:
 - i. P.ex. obre una finestra ms-dos i executa: `telnet www.upc.es 80`
 - ii. Per a veure el que tecleges, has d'activar l'echo local:



1.- Missatge HTTP de petició.

```
% telnet www.upc.es 80
```

```
GET / HTTP/1.1 ← petició de l'objecte "/" (vol dir l'objecte per defecte que hi ha a l'arrel d'aquest servidor). Usem la versió d'HTTP 1.1
Host: www.upc.es
User-agent: telnet ← User-agent és l'agent que estem fent servir per a connectar-nos al servidor Web
<CR+LF>
```

Format missatge de petició:

línia de petició	→	GET / HTTP/1.0	MÈTODE espai URL espai VERSIÓ CR LF
capçaleres	→	Host: www.upc.es	CAPÇALERA: espai VALOR CR LF
		...	
		User-agent: telnet	CAPÇALERA: espai VALOR CR LF
línia en blanc	→	Cr lf	CR+LF
entity body	→		En aquest cas està buit. S'utilitza amb el mètode POST (mètode usat sovint per a omplir formularis)

GET: mètode per a demanar un objecte al servidor

2.- Analitzem la resposta...:

Versió de protocol de la resposta

```

HTTP/1.1 200 OK ← 200 indica que la petició ha tingut èxit. La informació sol·licitada es torna com a resposta
Date: Tue, 12 Nov 2002 20:10:57 GMT
Server: Apache/1.3.24 (Unix) ApacheJServ/1.1.2 PHP/4.2.0rc1
Connection: close
Content-Type: text/html ← indica que l'objecte que rebem com a resposta és tipus html

<HTML>
  <HEAD>
    <META HTTP-EQUIV="Content-Type" CONTENT="text/html; charset=iso-8859-1">
    <META NAME="GENERATOR" CONTENT="Mozilla/4.03 [es] (Win95; I) [Netscape]">
    <META NAME="Author" CONTENT="Oficina de Premsa">
  <TITLE>Universitat Politecnica de Catalunya</TITLE>
  ...
  
```

Format missatge de resposta:

línia d'estat	→	HTTP/1.1 200 OK	VERSIÓ espai CODI_ESTAT espai FRASE cr lf
capçaleres	→	Date:	CAPÇALERA: espai VALOR cr lf
		...	
		Content-Type:	CAPÇALERA: espai VALOR cr lf
línia en blanc	→	cr lf	CR+LF
entity body	→		Conté l'objecte sol·licitat

3.- HTTP/1.1 implementa per defecte connexions persistents.

% telnet www.upc.es 80

```

GET / HTTP/1.1
Host: www.upc.es
<CR+LF>
  
```

Resposta

Un cop rebuda la resposta, la connexió amb el servidor continua oberta. A continuació podem fer una altra petició i aprofitem la mateixa connexió (no cal tornar a establir la connexió TCP)

```
GET /catala/webs/webs.htm HTTP/1.1
Host: www.upc.es
<salt de línia>
Resposta .....
```

En la segona petició, estem aprofitant la mateixa connexió TCP.

4.- Si volem una connexió no persistent quan utilitzem HTTP/1.1, cal afegir la capçalera: Connection:close

```
% telnet www.upc.es 80
```

```
GET / HTTP/1.1
Host: www.upc.es
Connection: close ← quan el servidor acabi d'enviar-nos la resposta tancarà la connexió TCP
<CR+LF>
Resposta .....
```

5.- Altres capçaleres: Accept-Language:.

```
% telnet www.google.com 80
```

[Aquest és un exemple del camp per a seleccionar l'idioma de preferència que es pot incloure en les capçaleres HTTP. Si ho proveu vosaltres us donarà un error perquè darrerament google redirecciona automàticament www.google.com a www.google.es.]

```
GET / HTTP/1.1
Host: www.google.com
User-Agent: telnet
Accept-language: ca ←
HTTP/1.1 200 OK
Content-Length: 2935
Server: GWS/2.0
Date: Tue, 12 Nov 2002 21:43:44 GMT
Content-Type: text/html
Cache-control: private
Set-Cookie: PREF=ID=56ed23a1346aff10:TM=1037137424:LM=1037137424:S=71z7-hB2OmYp1VVT; expires=Sun, 17-Jan-2038 19:14:07 GMT; path=/; domain=.google.com

<html><head><meta http-equiv="content-type" content="text/html; charset=ISO-8859-1"><title>Google</title><style><!--
      body,td,a,p,.h{font-family:arial,sans-serif;} .h{font-size:
20px;} .h{color:} .q{text-decoration:none; color:#0000cc;}
//--></style>
<script>
<!--
function sf(){document.f.q.focus();}
// -->
</script>
</head><body bgcolor=#ffffff
text=#000000 link=#0000cc vlink=#551a8b alink=#ff0000 onLoad=sf()><center><table
cellpadding=0 cellspacing=0 border=0><tr><td align=right valign=bottom><img
src=images/hp0.gif width=158 height=78 alt="Google"></td><td valign=bottom><img
src=images/hp1.gif width=50 height=78 alt=""></td><td valign=bottom><img src=images/hp2.gif
width=68 height=78 alt=""></td></tr><tr><td class=h align=right valign=top><b></b></td><td
valign=top><img src=images/hp3.gif width=50 height=32 alt=""></td><td valign=top
```

Indiquem al servidor la preferència en l'idioma de la resposta (català). Si no té una versió en aquest idioma ens contesta amb la versió per defecte


```
Server: Apache
Last-Modified: Tue, 12 Nov 2002 02:02:54 GMT
ETag: "15758e-1ad3-3dd0614e"
Accept-Ranges: bytes ← ens indica que accepta peticions per parts del fitxer. En el proper
apartat veurem un exemple
Content-Length: 6867
Connection: close
Content-Type: text/html; charset=iso-8859-1

<!DOCTYPE HTML PUBLIC "-//W3C//DTD html 4.0 transitional//EN">
<html>

<head>

<meta http-equiv="Content-Type" content="text/html; charset='iso-8859-
1'">

<meta name="GENERATOR" content="HMS 2.01">
.....
.....
```

Quan accedim a una pàgina, el navegador en desa una còpia a la memòria cau (caché). Si al cap d'un temps tornem a demanar el mateix objecte, només hem de comprovar si la versió que tenim ha canviat:

- En cas que no hagi canviat, el servidor ens notifica que no ha canviat.
- Si hagués canviat, rebríem la versió nova de l'objecte.

```
GET /presentacio/ HTTP/1.0
Host: www.ac.upc.es
If-modified-since: Tue, 12 Nov 2002 02:02:54 GMT

HTTP/1.1 304 Not Modified
Date: Tue, 12 Nov 2002 19:44:24 GMT
Server: Apache
Connection: close
ETag: "15758e-1ad3-3dd0614e"
```

6.- Obtenir trossos d'una pàgina:

```
% telnet www.ac.upc.es 80
```

```
GET / HTTP/1.1
Host: www.ac.upc.es
User-Agent: telnet
Accept: */*
Range: bytes=10-20,30-35

HTTP/1.1 206 Partial Content
Date: Tue, 12 Nov 2002 20:15:25 GMT
Server: Apache
Last-Modified: Tue, 12 Nov 2002 02:02:28 GMT
ETag: "37e4c-13d3-3dd06134"
Accept-Ranges: bytes
Content-Length: 233
```

```
Content-Type: multipart/byteranges; boundary=3dd1615d69b5
```

```
--3dd1615d69b5
```

```
Content-type: text/html; charset=iso-8859-1  
Content-range: bytes 10-20/5075
```

```
HTML_PUBLI
```

```
--3dd1615d69b5
```

```
Content-type: text/html; charset=iso-8859-1  
Content-range: bytes 30-35/5075
```

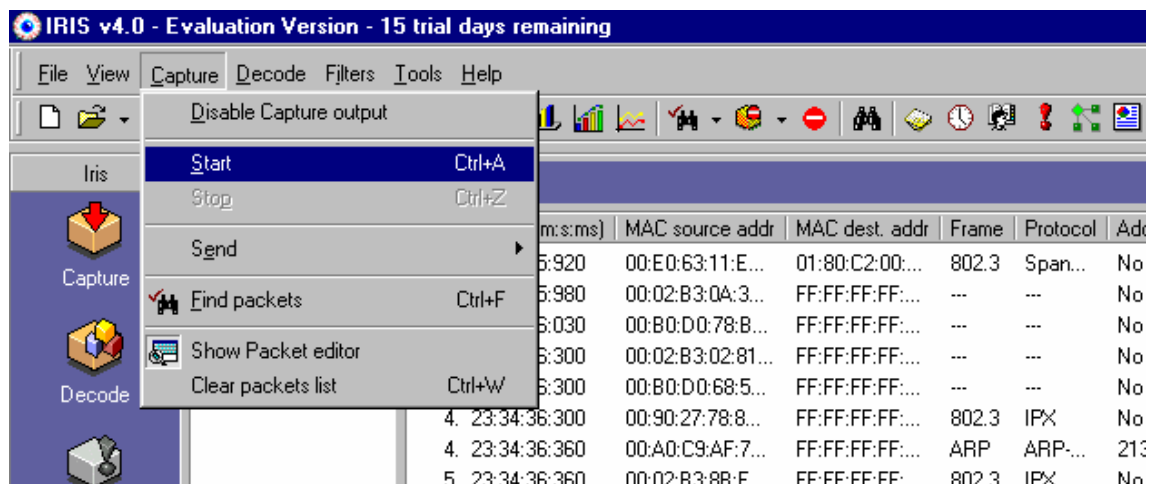
```
//DTD
```

```
--3dd1615d69b5--
```

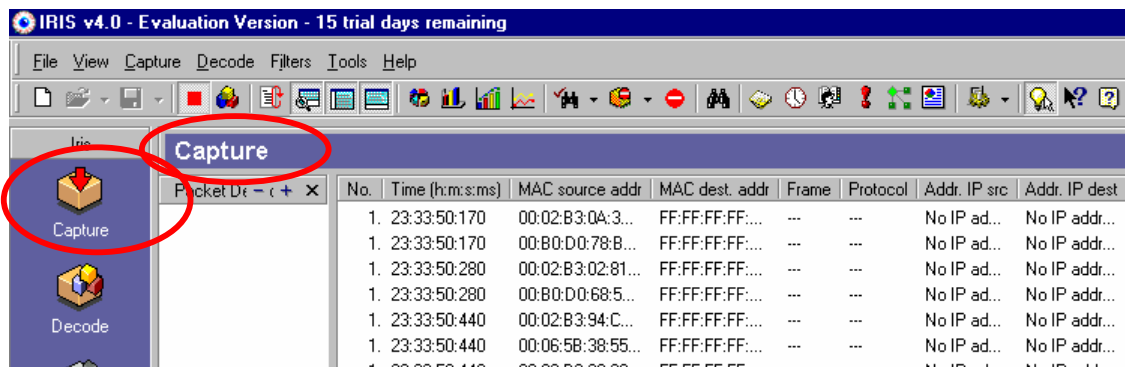
ANALITZADOR

Com analitzador de xarxa utilitzarem Iris (<http://www.eeye.com/html/Products/Iris/>). Un cop engegat,

- a) **Capture – start** per a que l'analitzador comenci a capturar trames. En el mateix menú podreu aturar la captura amb l'opció **Stop**.



L'analitzador comença a capturar les trames que passen per la xarxa a la que esteu connectats. Si voleu veure les trames que captura heu d'estar en la vista **Capture**



- b) El que ens interessa és veure els paquets HTTP que s'han intercanviat.
 - a. Engegueu la captura
 - b. Aneu amb el navegador a algun pàgina web. P.ex. a la pàgina web de la UPC (<http://www.upc.es>) o de la FIB (<http://www.fib.upc.es>)
 - c. Atureu el captador.

Per a veure els missatges HTTP:

- a. Heu d'anar a la vista **Decode**
- b. Expandir la informació del vostre ordinador (l'identificareu per l'adreça IP)
- c. Entre la informació que hi trobareu hi ha informació sobre HTTP
- d. Selecciónt-la, us apareixerà les connexions HTTP que heu realitzat
- e. Per cada entrada a l'analitzador hi ha missatge petició i resposta.

The screenshot shows the IRIS v4.0 interface. The 'Decode' button in the left sidebar is circled in red. The main window displays a list of network packets. One packet is selected and highlighted with a red box, showing its details: 'TCP-> HTTP (80)'. Below this, the decoded content of the packet is shown, including the request and response headers. A red box with the text 'Missatge petició i missatge resposta de l'entrada seleccionada' points to this decoded content. The status bar at the bottom shows system information like CPU usage and IP address.

- c) Mireu les diferents entrades i vegeu quina informació s'han intercanviat.
- d) Buida la MEMÒRIA CAU (*caché* en anglès). En el Netscape ho pots fer anant a **Edita – Preferències – Avançat – Memòria Cau**.
- e) Ara captura la connexió a <http://www.w3.org>. Mira quins han estat els missatges de petició i de resposta.
 - a. Examinant la capçalera hauríeu d'entendre el format de la petició i la resposta
 - b. Adonar-vos que en una mateixa connexió s'ha fet més d'una petició i una resposta (persistència de la connexió)
- f) Repeteix la captura a <http://www.w3.org>. Veuràs que:

- a. tant el volum de bytes sortints com entrants és diferent de la petició anterior
- b. hi ha pàgines que no s'ha descarregat perquè ja en tenia una còpia a la memòria cau

Qüestió 1:

- a) En la connexió www.fib.upc.es: Quantes connexions s'han establert? Quantes peticions i respostes per sessió?
- b) I en la connexió www.w3.org?
- c) I en el cas de la segona petició a www.w3.org en que ja teníem coses a la memòria cau del navegador?

GALETES (cookies en anglès)

Les galetes són un mecanisme que alguns llocs web utilitzen per a seguir els usuaris que s'hi connecten. No tots els portals utilitzen galetes, però els principals portals n'utilitzen (p.ex. Yahoo, Netscape), també els portals de comerç electrònic (Amazon); fins i tot n'hi ha en portals com softcatlà (<http://www.softcatla.org>) que és un portal sense ànim de lucre que recull el programari en català.

La tecnologia que hi ha darrera les galetes es basa en quatre components:

- a) una capçalera de galeta en un missatge de resposta HTTP
- b) una capçalera de galeta en un missatge HTTP de petició
- c) un fitxer de les galetes desat al sistema de l'usuari i gestionat per navegador de l'usuari
- d) una base de dades al lloc web

A través d'un exemple veurem com es fan servir les galetes. Ho farem en diferents etapes:

- a) Mireu com teniu configurat el vostre navegador respecte a les galetes.
 - Heu d'anar a les preferències del navegador que esteu utilitzant (Netscape: **Edita – preferències**) i busqueu on parla de les galetes (Netscape: **galetes**(o *cookies* en anglès).
 - Activeu l'opció d'**Avisa'm abans d'emmagatzemar una galeta**.
 - Mireu quines galetes teniu en aquest moment i esborreu les que hi hagi.
- b) Accediu a una pàgina que té galetes i veureu com el navegador les detecta
 - Connecteu-vos a <http://www.softcata.org>. Veureu que el navegador us notifica que els servidor vol emmagatzemar una galeta (o modificar-ne alguna si ja teniu aquesta galeta del servidor) i us demana autorització per a fer-ho.
- c) Elimineu les galetes del navegador
 - Elimineu les galetes del navegador.
- d) Activeu el mode captura de l'analitzador de xarxa
- e) Torneu a accedir a la pàgina que tenia les galetes
- f) Accepteu les galetes
- g) Navegueu una mica per la pàgina on us heu connectat
- h) Atureu l'analitzador de xarxa
- i) Examineu les interaccions per veure les galetes intercanviades. Entre altres coses veureu que:
 - a) el primer missatge de petició no porta cap galeta
 - b) la resposta del servidor conté una galeta

- c) si el navegador accepta la galeta, afegeix una entrada al fitxer del navegador que gestiona les galetes
- d) en les successives consultes a pàgines d'aquest lloc web, el navegador afegeix als missatges de petició una capçalera que inclou les galetes d'aquest lloc web.
- e) El navegador inclourà les galetes cada cop que es torni a connectar a aquest lloc web, encara que hagi passat un cert temps. D'aquesta manera un lloc de comerç electrònic pot recomanar a un usuari uns productes enlloc d'altres pel coneixement de les seves compres en el passat
- f) Si, a més, l'usuari es registra al lloc web (dóna informació del seu nom, adreça, tarja de crèdit, ...), aquest pot associar aquesta informació amb l'identificador de la galeta. D'aquesta manera és com els llocs web proporcionen "compra amb un clic" (tant per aquella sessió com per futures sessions).
- g) Les galetes també s'usen per autenticar o per a crear sessions sobre un nivell sense estat com és l'HTTP. P.ex. quan ens connectem a un correu web, el navegador envia informació de les galetes al servidor, permetent al servidor identificar l'usuari durant tota la sessió de l'usuari amb l'aplicació.

De tot això ja es veu que, tot i que les galetes simplifiquen certes interaccions amb llocs web (p.ex. en els llocs web comercials), generen controvèrsia pel fet que es poden veure com una violació de la privacitat. Com ja hem vist, usant una combinació de les galetes i d'informació aportada per l'usuari, un lloc web pot aprendre molt d'un usuari i vendre aquesta informació a un tercer. Les galetes també es poden usar per a recollir informació sobre el comportament d'un usuari concret a través de molts llocs web. Les pàgines web que mostren anuncis (banners ads) usen missatges de petició HTTP per a obtenir l'anunci (que són GIF o JPEG) del servidor HTTP de l'empresa anunciadora. Cadascuna de les peticions al servidor HTTP de l'empresa que s'anuncia pot portar una galeta que gestiona l'empresa anunciadora. Donat que les principals empreses anunciadores a Internet proporciona anuncis a molts llocs web, l'empresa pot construir un perfil de comportaments de navegació d'un individu a través de varis llocs web.

Si t'interessa analitzar més interaccions amb galetes pots connectar-te a <http://www.amazon.com>, <http://home.netscape.com> o <http://www.uoc.edu> (utilitza una galeta per a saber quin és l'idioma en que l'usuari vol veure una pàgina. Les opcions són: a (català), b (castellà) o c (anglès)).

WEBDAV

a) Per a connectar-vos a un servidor WEBDAV utilitzant el protocol WEBDAV cal que creeu una '**carpeta web**' des del MS-Windows que tingueu instal·lat. Aquesta carpeta web s'ha de relacionar amb el servidor que us indicarem a classe. Aquest servidor WEBDAV només funcionarà durant la sessió de problemes (si algú ho vol provar posteriorment haurà de tenir accés a algun servidor WEBDAV o instal·lar-se'n un (l'apache v2.0 porta un mòdul de DAV)).

b) Poseu en marxa l'analitzador

c) Accediu al servidor WEBDAV per posar o obtenir documents. També accediu-hi per a veure les propietats d'un objecte...

d) Atureu l'analitzador

f) Mireu quines missatges s'han intercanviat. Els missatges amb el servidor han de contenir peticions i respostes de l'estil de les que us posem a continuació:

8.1.1 Example - Retrieving Named Properties ← extret del text de l'RFC 2518 (WEBDAV)

>>Request

PROPFIND /file HTTP/1.1

Host: www.foo.bar

Content-type: text/xml; charset="utf-8"

Content-Length: xxxx

```
<?xml version="1.0" encoding="utf-8" ?>
<D:propfind xmlns:D="DAV:">
  <D:prop xmlns:R="http://www.foo.bar/boxschema/">
    <R:bigbox/>
    <R:author/>
    <R:DingALing/>
    <R:Random/>
  </D:prop>
</D:propfind>
```

>>Response

HTTP/1.1 207 Multi-Status

Content-Type: text/xml; charset="utf-8"

Content-Length: xxxx

```
<?xml version="1.0" encoding="utf-8" ?>
<D:multistatus xmlns:D="DAV:">
  <D:response>
    <D:href>http://www.foo.bar/file</D:href>
    <D:propstat>
      <D:prop xmlns:R="http://www.foo.bar/boxschema/">
        <R:bigbox>
          <R:BoxType>Box type A</R:BoxType>
        </R:bigbox>
        <R:author>
          <R:Name>J.J. Johnson</R:Name>
        </R:author>
      </D:prop>
      <D:status>HTTP/1.1 200 OK</D:status>
    </D:propstat>
    <D:propstat>
      <D:prop><R:DingALing/><R:Random/></D:prop>
      <D:status>HTTP/1.1 403 Forbidden</D:status>
      <D:responsedescription> The user does not have access to
the DingALing property.
      </D:responsedescription>
    </D:propstat>
  </D:response>
  <D:responsedescription> There has been an access violation error.
</D:responsedescription>
</D:multistatus>
```

In this example, PROPFIND is executed on a non-collection resource <http://www.foo.bar/file>. The propfind XML element specifies the name of four properties whose values are being requested. In this case only two properties were returned, since the principal issuing the request did not have sufficient access rights to see the third and fourth properties.

Apart d'usar els mètodes GET i PUT d'HTTP també veureu que els missatges de petició porten altres mètodes (si voleu més informació, consulteu el RFC 2518):

- PROPFIND method: retrieves properties defined on the resource identified by the Request-URI
- PROPPATCH method: processes instructions specified in the request body to set and/or remove properties defined on the resource identified by the Request-URI.
- ...

Enviament solució

Envieu la resposta a les qüestions plantejades a: aad@ac.upc.es (en un missatge de text per grup indicant a més els noms dels membres del grup). CAL QUE EL **TEMA** DEL MISSATGE SIGUI: **AAD-problemes sessio HTTP i WEBDAV**.

Bibliografia

- HTTP: RFC 2616 (HTTP/1.1) i RFC 1945 (HTTP/1.0): <http://www.ietf.org/rfc.html>
- Galetes: RFC 2109: <http://www.ietf.org/rfc.html>
- WEBDAV: RFC 2518: <http://www.ietf.org/rfc.html>
- Analitzador de xarxa: <http://www.eeye.com/html/Products/Iris/>