Universidad de Alcalá

Escuela Politécnica Superior

Universidad de Alcalá

PECL2

Arquitectura y Diseño de Sistemas Web y C/S

Protocolo Http

Laboratorio Jueves 10:00 – 12:00

Grado en Ingeniería Informática & Ingeniería en Sistemas de Información – Curso 2019/2020

Marcos Barranquero Fernández – 51129104N

Daniel Manzano Estébanez – 03220212M

CONTENIDO

INTRODUCCIÓN

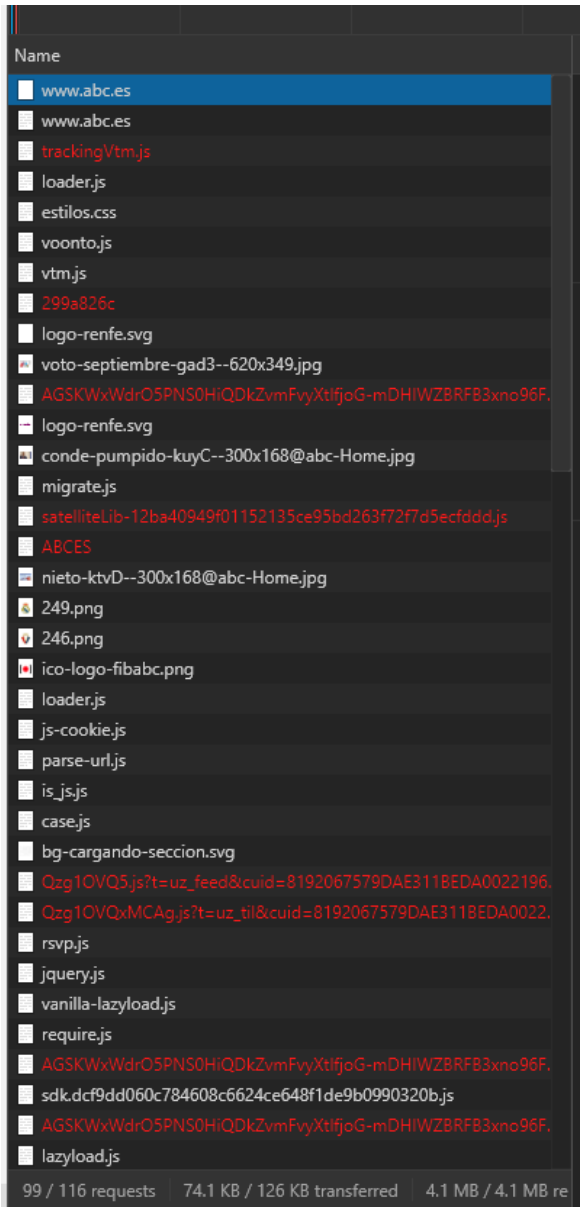
En esta práctica se procede al estudio de las cabeceras del protocolo HTTP, interactuando con las distintas webs propuestas.

Para el desarrollo de la práctica se ha elegido utilizar Google Chrome y su herramienta para desarrollador que permite **loguear** los paquetes e interacciones con servidores por el protocolo **http**.

SOLICITUDES – GETS

CONEXIÓN A ABC

Nada más realizar la solicitud, observamos que recibimos una gran cantidad de datos. En total, vemos que se han recibido **4.1 Mb**. Además, se han realizado 116 solicitudes de las cuales se han completado satisfactoriamente 99.

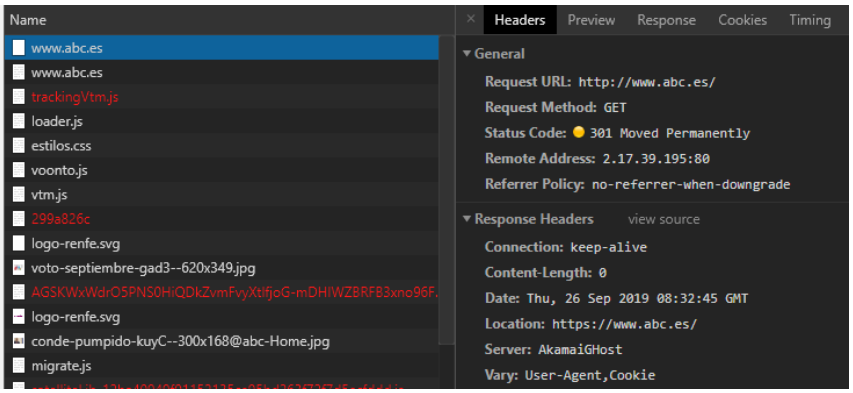


La primera solicitud es la de la propia página web. Es interesante observar que hemos sido redirigidos de la versión **http** a la versión **https**, debido a la instalación de la aplicación httpsEverywhere en el navegador.

Observamos además que se solicitan gran cantidad de archivos:

- Los archivos HTML con el contenido de la web.
- Los archivos .css correspondientes con el estilo de la página.
- Archivos png, jpg, y svg correspondientes a las imágenes mostradas en la web.
- Archivos javascript (.js) que contienen los diferentes scripts que se ejecutan sobre la web.
- Ciertos archivos correspondientes a la **SSID** y **Key de la sesión**, posiblemente para uso en monitoreo y análisis de datos por parte de la web.

Es interesante observar la cabecera de la primera petición:

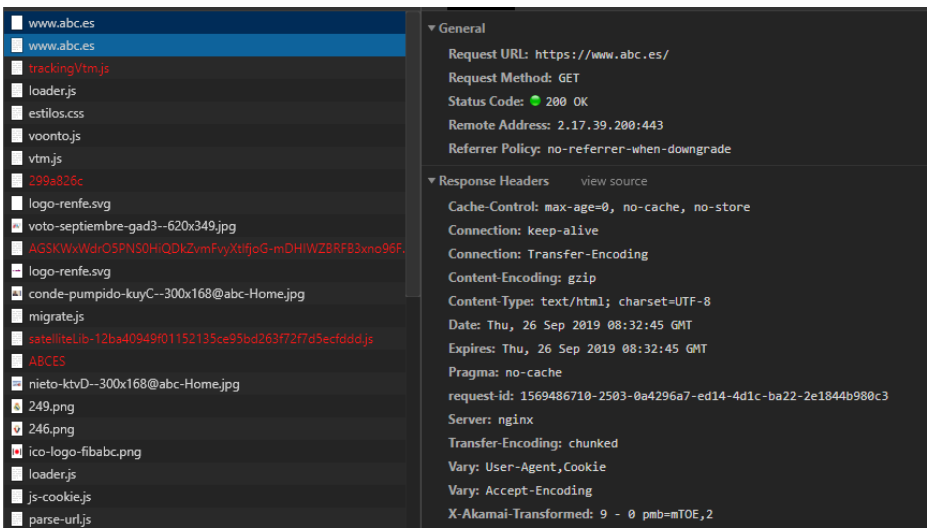


En ella, vemos que se ha realizado una solicitud por parte del cliente del tipo Get al dominio de abc.

Vemos también la cabecera de solicitud del cliente. En ella se especifican datos como el idioma, que tipo de archivos acepta, el host solicitado, y cookies identificativas de la sesión y del navegador.

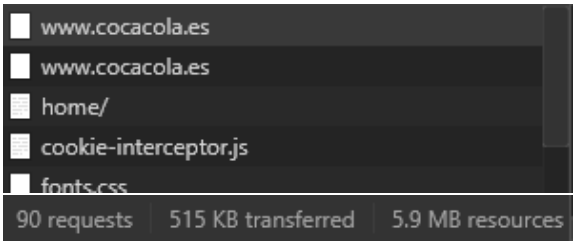
La respuesta del servidor es una redirección de ip y puerto:

1. Inicialmente se solicita la web (el archivo html) sobre el puerto 80, correspondiente al protocolo http. La web contesta que el navegador debe redirigir a su versión https, sobre otra IP y puerto.
2. El navegador solicita la web sobre la IP devuelta y el puerto 443. El servidor responde con status ok, y devolviendo la versión sobre el protocolo https.

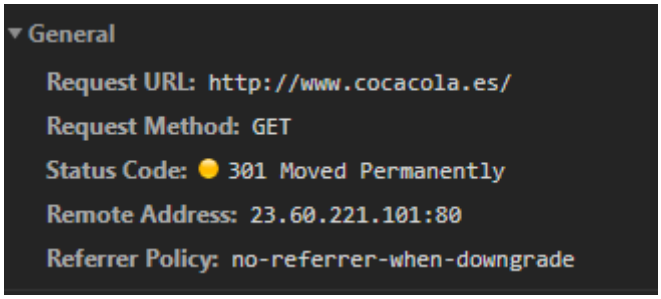


CONEXIÓN A WWW.COCACOLA.ES

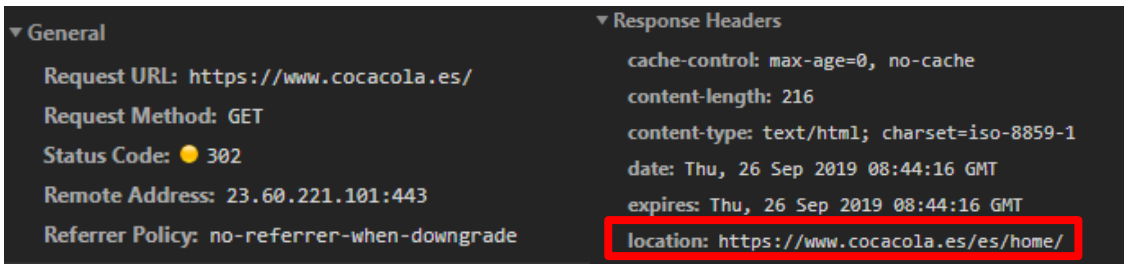
Del mismo modo que en la página anterior, al realizar la solicitud recibimos muchos datos, 9,5MB, que contiene el código de la página y diversos archivos de estilo. Si miramos a la primera petición (www.cocacola.es), podemos observar su cabecera.



El navegador completa la dirección con los elementos que faltan, dejándola como http://www.cocacola.es/, y adjuntando la orden GET. La respuesta que obtiene es 301 (redireccionamiento), y se informa de que ese dominio está movido permanentemente. Esto se debe a que el dominio no es "http", sino "https". Entonces el navegador inicia una nueva petición a la página a la que ha sido redireccionado.

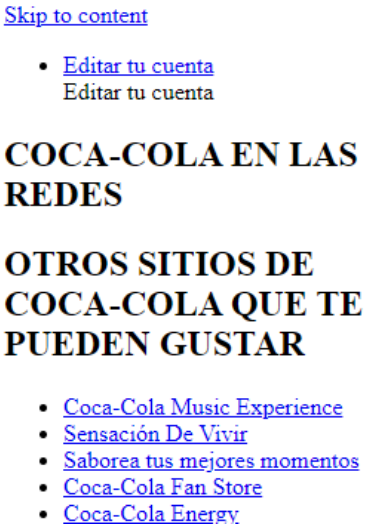


Se realiza de nuevo la petición, pero al https, con un GET. Sin embargo, también se produce un redireccionamiento, pues el servidor de Coca Cola quiere que la página de inicio sea "https://www.cocacola.es/es/home/".



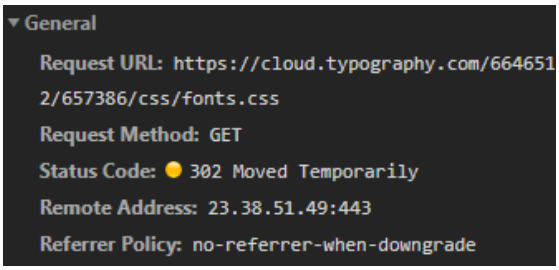
El navegador repite la petición para obtener esa página, obteniendo en este caso un código 200 (petición correcta).

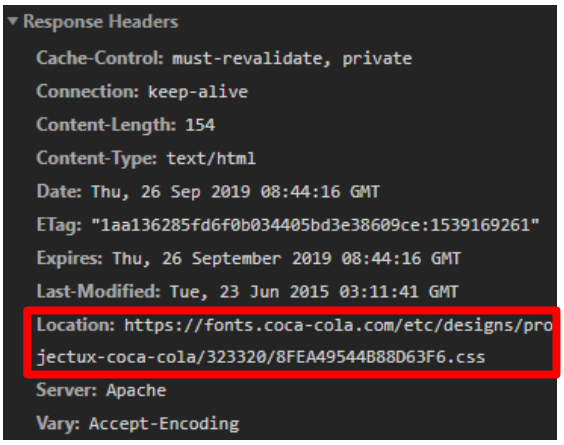
Con la respuesta a esta petición se incluye el código fuente de la página web y, tras recibirlo completamente, se desencadenan una serie de GETs para obtener todos los archivos de estilo y demás componentes de la web.



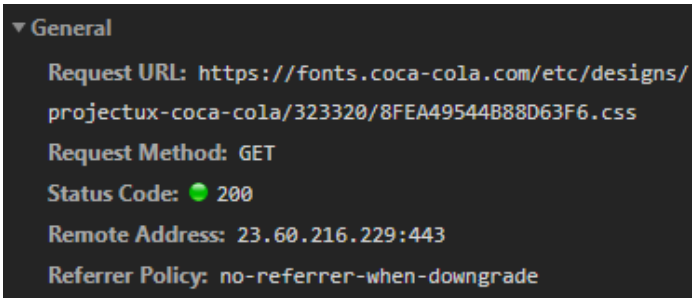
La gran mayoría de peticiones obtienen un código de respuesta 200, a excepción de alguna petición que es redireccionada (300). En esos casos, se puede encontrar la nueva petición realizada de nuevo a la nueva dirección más adelante.

Por ejemplo, al solicitar "fonts.css" recibe esta respuesta.





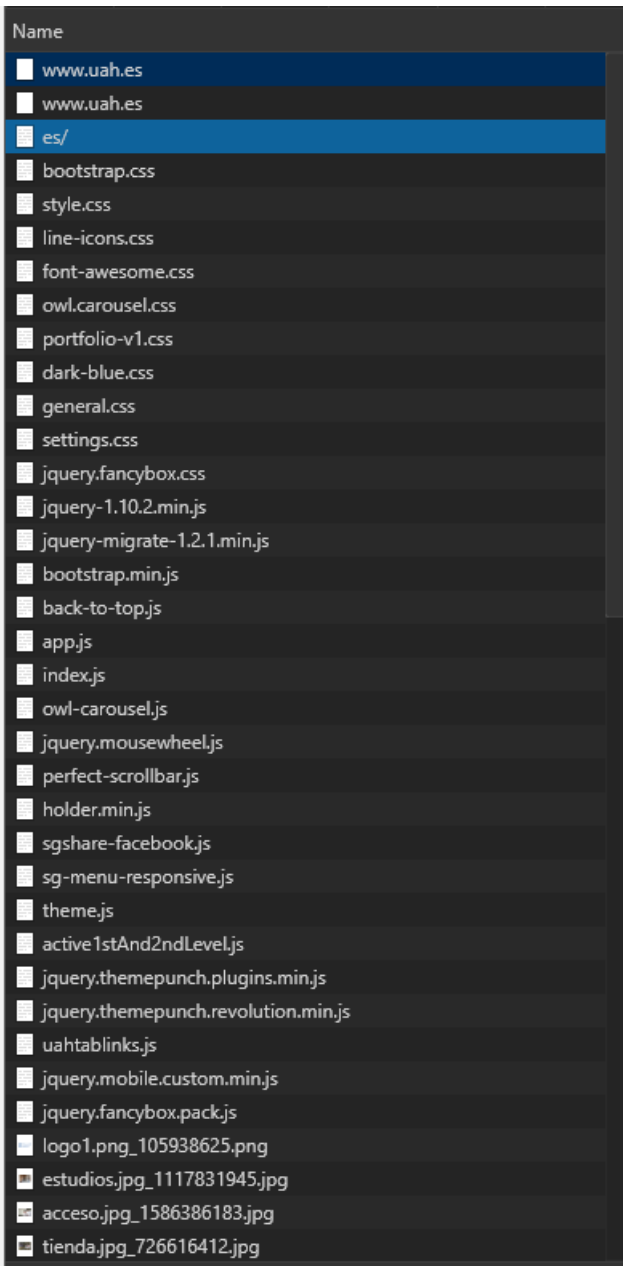
Posteriormente, realiza la solicitud para obtener el archivo al que se ha redireccionado, la cual se completa correctamente:



En este caso, ninguna petición realizada ha obtenido error.

CONEXIÓN A UAH

La conexión a la web de la Universidad de Alcalá genera el siguiente log:



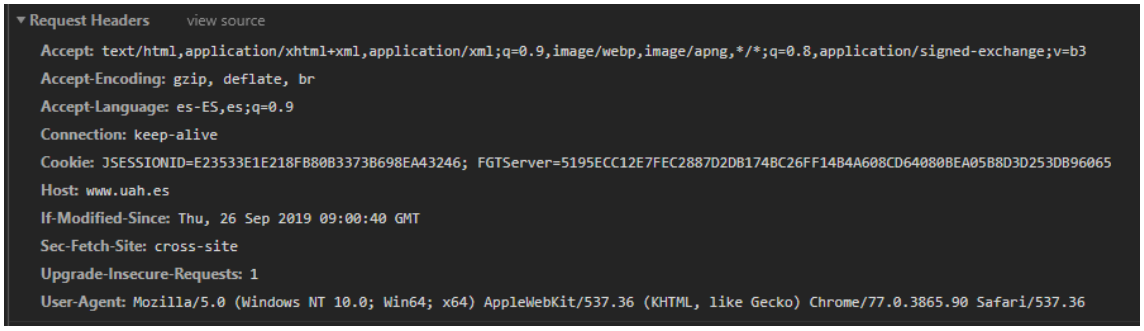
Observamos que es bastante similar a las solicitudes anteriores. Distinguimos archivos de HTML, archivos de estilos .css, varios scripts en JavaScript, y archivos de imágenes png.

Respecto al inicio de la conexión, vemos que nos ha redirigido a la versión https española.

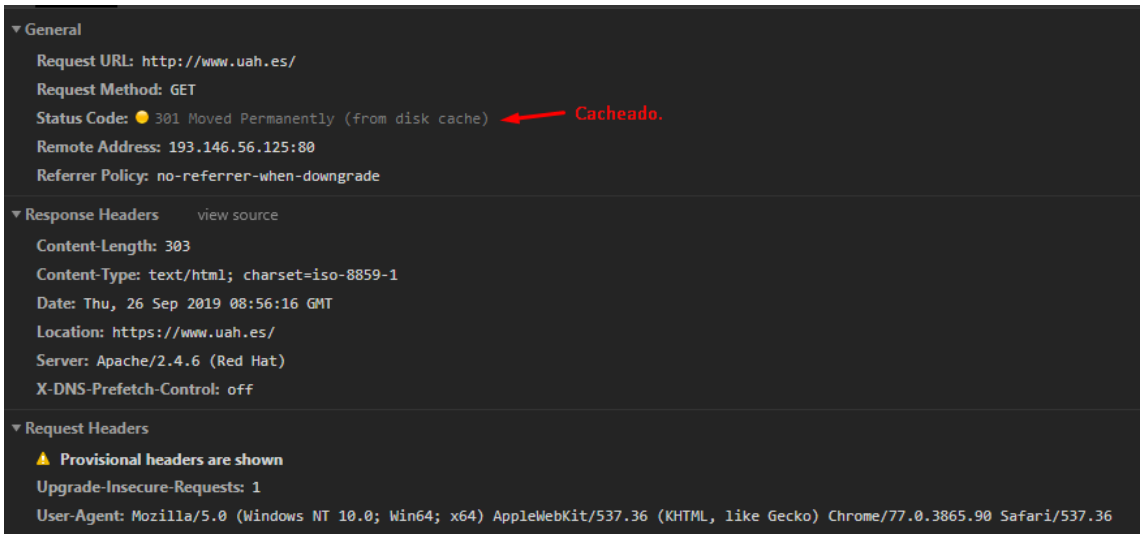
El proceso seguido es el siguiente:

1. El cliente ha enviado un request a la URL de la Uah por el protocolo http.
2. El servidor ha respondido con código 301, que significa que el recurso ha sido desplazado. En la cabecera vemos que la nueva localización es la misma url pero por conexión de https.
3. El cliente solicita el recurso web por protocolo https y puerto 80.
4. El servidor responde con código 301, explicando que el recurso se ha desplazado a la misma url solicitando el recurso "/es/", refiriéndose a la versión en español de la web.
5. El cliente solicita la URL con la extensión "/es/" y puerto 443 de protocolo https.
6. El servidor responde con status 200, indicando que se ha completado la request satisfactoriamente y devolviendo el archvio html y todos los recursos asociados.

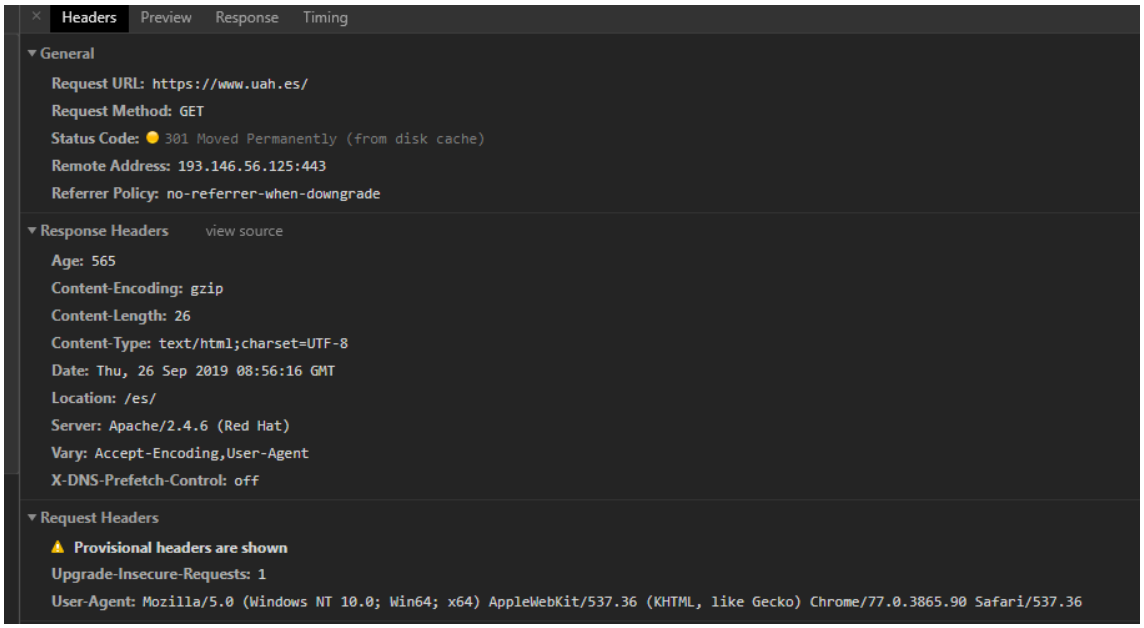
Pero ¿cómo sabe el servidor que debe redirigir a la versión española de la web? En el primer paquete, en la cabecera del cliente se especifica la localización, que es Spain:



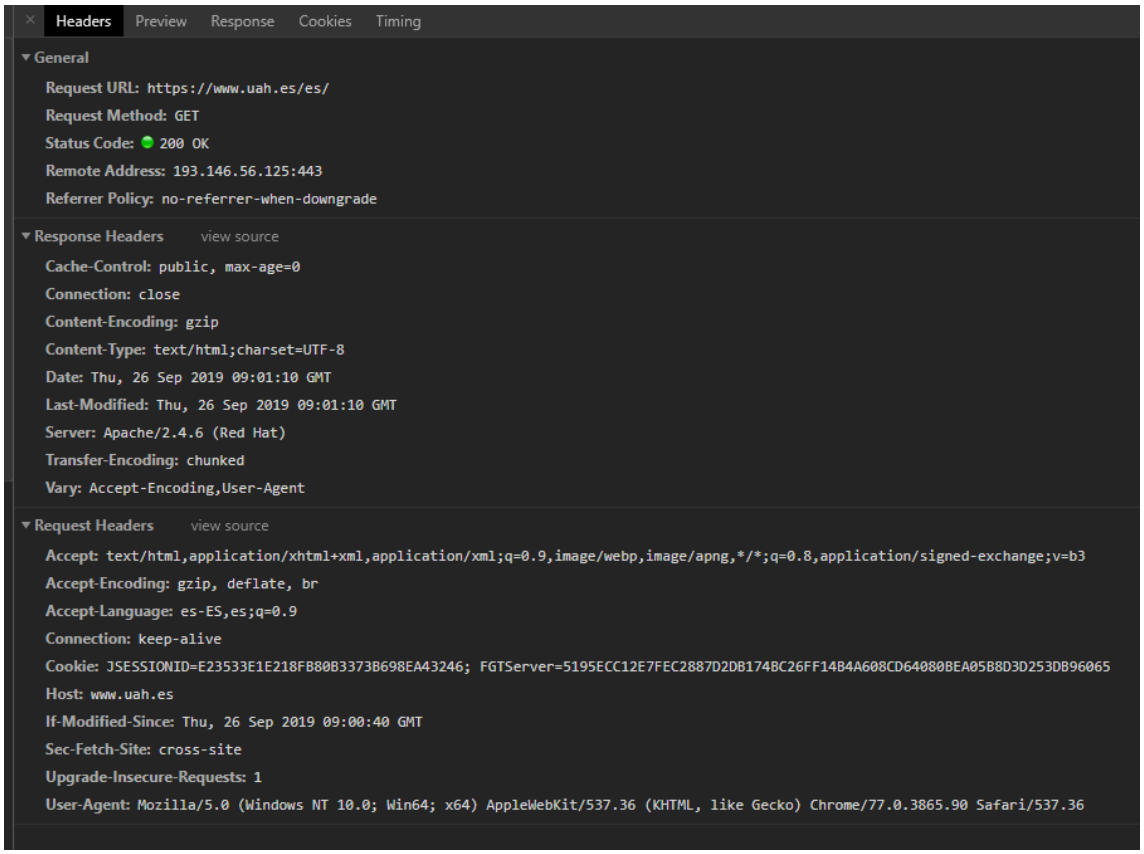
En la captura de la solicitud inicial no se muestra dicho header debido a que la solicitud se encuentra en caché, debido a accesos previos a la web de la universidad:



Segundo paquete y redirección a la web en su versión española:



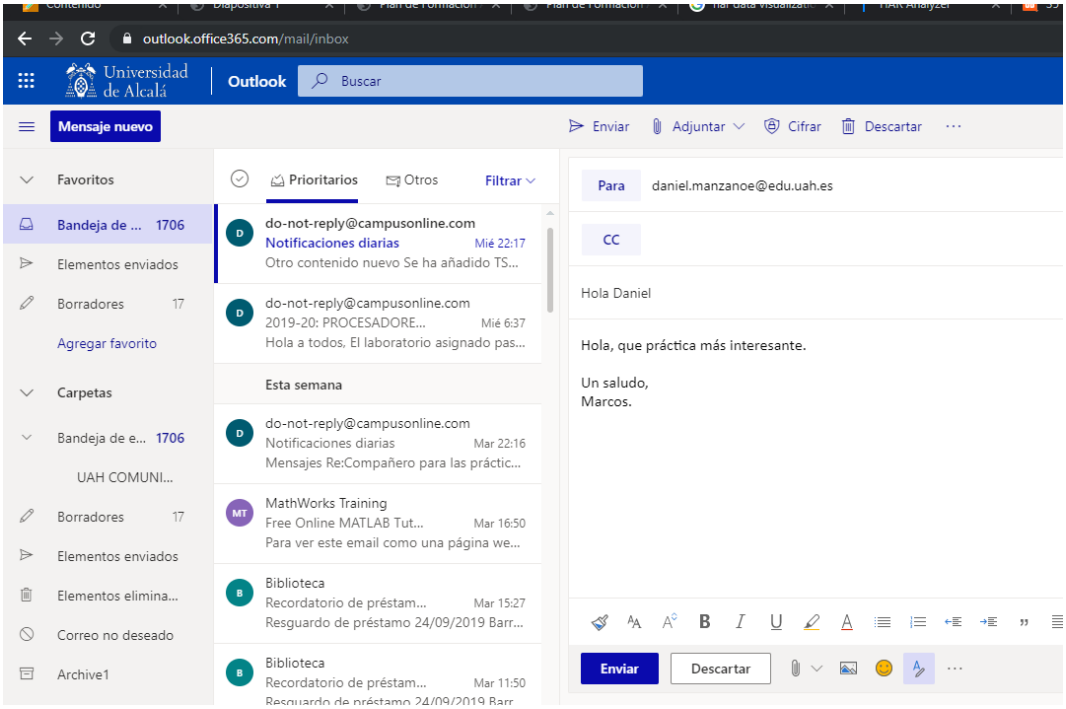
Finalmente, se devuelve la web (el html y los archivos consiguientes) en española.



ENVÍOS – POSTS

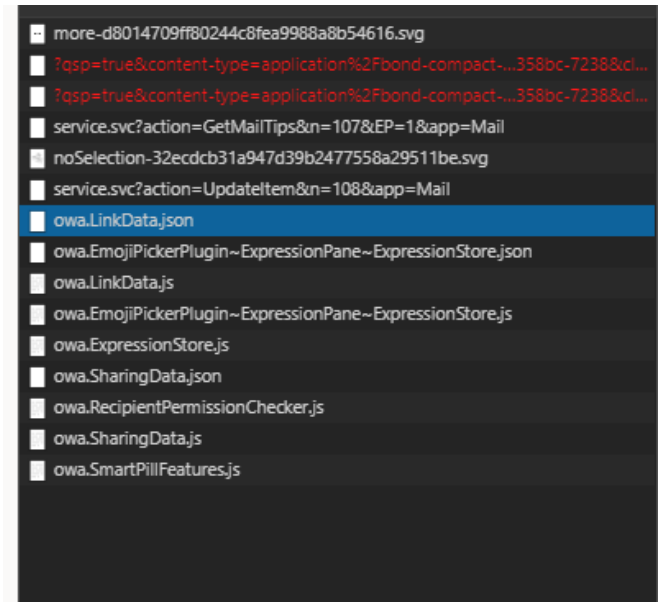
CORREO OUTLOOK ASOCIADO A LA UNIVERSIDAD

Se ha redactado y enviado el siguiente correo a mi compañero:



Una vez abierta la extensión de log. de la red, se pulsa el botón enviar y se obtienen los siguientes paquetes:

Vemos varios archivos .json, js, .svg de imágenes, etc.



También hay varios archivos .html con comandos de acciones insertados. Observamos el GetMailTips que corresponde al envío de la información del correo al servidor de Outlook, y UpdateItems, que recoge los nuevos mails tras volver a la sección de mails recibidos tras enviar el mail.

Además, vemos dos solicitudes denegadas, que han sido bloqueadas por una extensión del navegador que bloquea pop-ups y anuncios.

En ningún momento vemos que se utilice protocolos de correo, esa parte corresponde a los servidores de Outlook debido a la arquitectura cliente – servidor. Por nuestra parte solo se envía la información mediante el protocolo https.

Como ya nos hemos logueado, el proceso de envio del mail es el siguiente:

1. El cliente envía mediante un POST el mail encriptado en la siguiente cabecera:

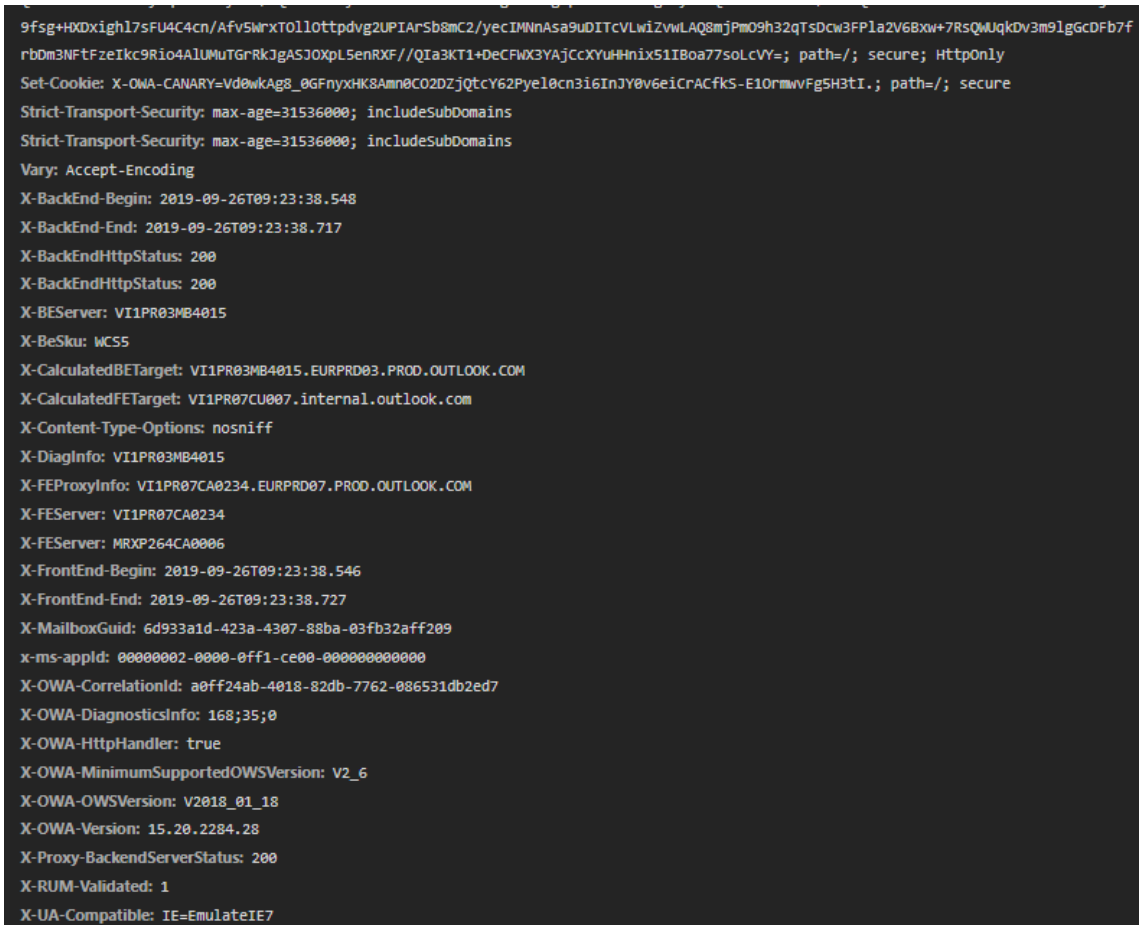
```
▼ General
    Request URL: https://outlook.office365.com/owa/service.svc?action=GetMailTips&n=107&EP=1&app=Mail
    Request Method: POST
    Status Code: 200 OK
    Remote Address: 40.101.92.194:443
    Referrer Policy: no-referrer

▼ Request Headers   view source
    Accept: */*
    Accept-Encoding: gzip, deflate, br
    Accept-Language: es-ES,es;q=0.9
    action: GetMailTips
    Connection: keep-alive
    Content-Length: 0
    content-type: application/json; charset=utf-8
    Cookie: ClientId=9FD5B6AE2C4443C4917AB8D775227730; OIDC=1; as=Y2RgYGAEYmYgZjE0M7cEAA; DefaultAnchorMailbox=marcos.barranquero@edu.uah.es; SuiteServi
    ceProxyKey=VwsOxiDWcpb5/ewTH38eEVnGwK0xdL+adQ5HDgczjZ4=&kSwTeMG0ZtkE1S8TKkVHeQ==; UC=d2fdd6dc4a7f4931b999d1325e8ca1a7; RoutingKeyCookie=v2:u4%2bu%2
    bH4ayWRKlzL5QYkhvR9poyus0IDSw0WEDKznoMU%3d:6d933a1d-423a-4307-88ba-03fb32aff209@edu.uah.es; domainName=uah.es; HostSwitch=1; HostSwitchPrg=1; OptIn
    H=QpEL8hrZ9elvS5wPCKrV99J_8aN-Tt7nl3NjU1pfBTA; X-OWA-RedirectHistory=AtKNjSIB3HvB32JC1wg|AvAZGHMBq7R-32JC1wg|ArLym14BfgwC2mJC1wg; X-OWA-JS-PSD=2; A
    ppcacheVer=16.3278.2.2710622:es-esbase; X-OWA-CANARY=ijrIKfifHkeoxsQ6Pg6hhICvpAZjQtcY_UCowOEEss00dq-IovaD5HonEPdZfKT3drqd8v2O3v0.; OpenIdConnect.to
    ken.v1=AQAAAPADAADyOFNfuPUzeT3s8l56i/f9a4+HqOS8UBbwYfPM7jVdUTOq2d/Av1af+TwxuizpjKq2yCZDBCMokoO7wMIbdjDyLORR++nHgnGyyMiSV8z78uJtinvJ3cFTQWJnCeXeAhg9
    PFS4e6JddWPBIfw+nUBzmEWZU3y8C5aMnWDdG1k2RTkNei4AbRtGFCMyD0UtjBk10MnRxH6oUth/I7DCsuPBwLyO3dy0iKGQ/9jrMeXjlcKFWjppt/Z4CEenpqv23W32bU5BD322LSkYdWVN40g
    TH6thbxRmXrluDUQ8Y42rP+K1SEfC9tz6UKvOzPlK+tWr96hyAeGIZ3RNUIjpxMUkhlDRpMAsOKuCttQhEB3tPSWx11xawKcnYCv86QYbhFjX8fD7wwfsn5S1CHQO5myvtLoBDOJZx3awuqYhwQ
    HrfyYrqDbBSC6eAxl5G/3emzZI3ftr0M65lUGhXh0EXBXnAPdViSCD401y5+sysZXB8Y3OmXQYeFa1MOc31PFfM/alkZrjDEvqOaMG3XTeNXTSjG8QHuZKde0+Zdx8Co4oTCBxGA0SA/CfwgXZZ
    aQV6XnNH0kGc/3w2/aducfbpLUNF/CGNtZWIdcnPSqqypJm5ZFI03PbRbT0ewsIRsABuh7DBet7F676KOUOfDm9NAHmf0ukXrqbUFw8DgDR2oDPOiQ5iyMKdBqST4IzfLxz2ppj1XDUD9yVJqmO
    CWKgb2+7tBy8qPlFQLdAZLTvkIFlsNN0MSoT2LD43kF7vKikJyUAhPAjdhO+xiciNM91D55Ig+npKLrtUgKd+c5aBwlhsfPLvGNYu18Ee5VVPX+DnxmD3NhCDbm6JyCVC5Y1OLNCJ33uzE8i5nu
    64QwdMJzxqOe1S1XSLBl7jyF4JdUHIwOe6gsTsVu2xapHqpuTOxMiuF/jNhe+wc/hL4endHLIFJ3TLQQU2LifzdHgUr1Y7+42NLvj2uWUo/43d0bDWNMhs7eGsgfK3yVaIcgqt/MhT3vTsaown5
    dsQ+bO0u/M1nNLfhzR/Fa92Zj1pFBdJjNnSiU8Zp7Mq1dr5n+Y3QO57s5Q5lbeCU0oAMuI61Z/eJnuEEQw27u0kjRcBK0iuIPyTZSDTQa3dNPUjR+bdqigjtgNkWq0+xfkJda0EO1R0l3bFkH/Y
    3xdZ9upyKIrQMQmkvvw+SYHIn6c8OuB5hUmMzwT2qRrm+c9xPZ6A9cXX2jMyH1DSE+lWeuol3UATJNMwC9+2QCuCA/X78W7IyLgfKsxqfG297PRHISXaaNJMe3gENJ4/eqlxf8f5HY6yAXl7+zD
    JRth86R+xuw4+L8WnZpbOc0jWGxPGP8Xltk2zNgAAQAABJMIrWadHFiaA4Em6wgc0ArWaHSGJm6PVKFOSKnZlFopFoVv3g325P9ggaxp9ASGb4xk1L5PHIYKl6f7ZalQWzpEhQv9ltwTQ2MGLg/
    xnAZ5A3i7cPRfRfFq0sTiMmG3bIe3RmKB+Ir9OHGzLj86Q/FOZi1jWo3mjQzOQcnmTRk1MU6yyIYpJMee58eDg6zUNNIz1DGiGdOm+QU20aPb5KAyy1W8BmtlmmN9xmOQgk8KZGVherYXVKVTUf
    DM8XQ3XrZwJGhhc5qUOoPoa9kJc31RxwoF0mmMW0TJgNQdEYTvwuf/PuhjmAjS08Lm2vrV7Qgx1fmL4LWTlS8qF7rqxwABAABTZ9IjMENgQh2WPooSbsoOsXV7Aec1FqX5r5uYBLUoGfq/tBEJR
    XD2xbQ2DxEffLMhy4Xl92BZ8pIcPD12bzF2tWRw43OrvbQyLpX+2/wfZrt269y0OEyL6gIc5tWGabvKVvMtAjuBQl2iRtJ6aHubEgRKb20IZ/az/zdxQZjV/V/LNM9EZnZwTPz78ukexe9V7cK4
    NKUf/eA+74/+q2SJRkctvhK/j20J4qxbwRG71QsY+zm4TQ3FqUEsZeTzZNPqUzOsCTL6jSGs9c0kk+6FFYWk37+8nzTuIxoQW47COBoL/Nilc5XY8p4aXdRBEnsuzDlO2JCBz1PA6gLPRuaWAAE
    AAIF9MhXnLpcwEEuvymnyccSy5qb1BRq+z18wuPuHA1WmdGbrTbuVe/AySE+P+kF9Svdr1qMNzPYYjvo53KuMWsqGfaOudFTap3GstS9My3T8I33rEXhTV0hC/7Wbtbxh5GnvqEfHjB3Y7MaSUA
    Ah/CN0O+MokiPCRQnhvqCIvUXaD/vefpx1VqVfN4VB6LVAABycsyAi5en5WN/Gr/2o2QjmcWXZUm3bejPx6tr5jLJDAfaKQEyJRVFJ/SF3v0l6oI26NqR3cJYDzJQZ7gRNXXzwPZ97lKspNdBRm
    BlhWkAQ/ayLRN2nbqcDObyVrvWnXgG0+IlSy55taQA5t90ixNM=
    Host: outlook.office365.com
    ms-cv: BePnGpbCVG/AqClP8drAYT.107
    Origin: https://outlook.office365.com
    Sec-Fetch-Mode: cors
    Sec-Fetch-Site: same-origin
    User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/77.0.3865.90 Safari/537.36
    x-owa-canary: ijrIKfifHkeoxsQ6Pg6hhICvpAZjQtcY_UCowOEEss00dq-IovaD5HonEPdZfKT3drqd8v2O3v0.
    x-owa-correlationid: a0ff24ab-4018-82db-7762-086531db2ed7
    x-owa-urlpostdata: %7B%22__type%22%3A%22GetMailTipsJsonRequest%3A%23Exchange%22%2C%22Header%22%3A%7B%22__type%22%3A%22JsonRequestHeaders%3A%23Exchang
    e%22%2C%22RequestServerVersion%22%3A%22V2018_01_08%22%2C%22TimeZoneContext%22%3A%7B%22__type%22%3A%22TimeZoneContext%3A%23Exchange%22%2C%22TimeZone
    Definition%22%3A%7B%22__type%22%3A%22TimeZoneDefinitionType%3A%23Exchange%22%2C%22Id%22%3A%22Romance%20Standard%20Time%22%7D%7D%2C%22Body%22%3A%
    7B%22__type%22%3A%22GetMailTipsRequest%3A%23Exchange%22%2C%22MailTipsRequested%22%3A1%2C%22SendingAs%22%3A%7B%22EmailAddress%22%3A%22marcos.barranq
    uero%40edu.uah.es%22%7D%2C%22Recipients%22%3A%5B%7B%22EmailAddress%22%3A%22daniel.manzanoe%40edu.uah.es%22%7D%5D%7D%7D
    x-req-source: Mail
```

Vemos información de telemetría (x-owa), cookies, la acción realizada, codificación del texto, etc.

2. El servidor responde diciendo que ha aceptado la solicitud, además de proporcionar información sobre la infraestructura.

```
▾ Response Headers    view source
  Cache-Control: no-cache, no-store
  Content-Encoding: gzip
  Content-Length: 656
  Content-Type: application/json; charset=utf-8
  Date: Thu, 26 Sep 2019 09:23:38 GMT
  Expires: -1
  P3P: CP="ALL IND DSP COR ADM CONo CUR CUSo IVAo IVDo PSA PSD TAI TELo OUR SAMo CNT COM INT NAV ONL PHY PRE PUR UNI"
  Pragma: no-cache
  request-id: 24933d8f-adfa-483d-bd60-47181637dbcb
  Server: Microsoft-IIS/10.0
  Set-Cookie: OpenIdConnect.token.v1=AQAAAPADAADaNdmn0OK0rNpmgG0YUtyTqRxEAiTbRA+uSAKb6YBYarikJfW84gWK3phOylwJ1ooAjgqSUpC9ZucFJhlPd31pM
  TncATbqkPK8Wusz3FmRy8w4VdnBuJW5Dc46927sGbMZDG1pMmtVpAwCRNg+UAITlNCM4uZAjTr4+Me2iqm5szuTvgOyD//yJ2BVNtMi3xZUtSYvYjfq9hJF0wEqbob8Cfm
  QFco46iWYzNOVHEasB/ygloWs7alEHSQSGi0K/8V03tI/kNgWAnylf0Lqf3GFKSkG/0R6+Xe9+fqzkJ9cgJZD5ydfUDQSvrbI49OL6/DiCt4pV/Y2ZFifPKCa9CFp4S/5g
  tvOqToKlJXyRqZ8cUZGogmb3I+Bl/bEFO8o6dGn9otmZCwMTL59jPAPeqMFwfmOW/ZaJzh5KMVIejd5HnxgHkCfTqeRXqiSIOX3Mp17fGF/3aJZgYvQ1HlzJfilt/5WTVY
```

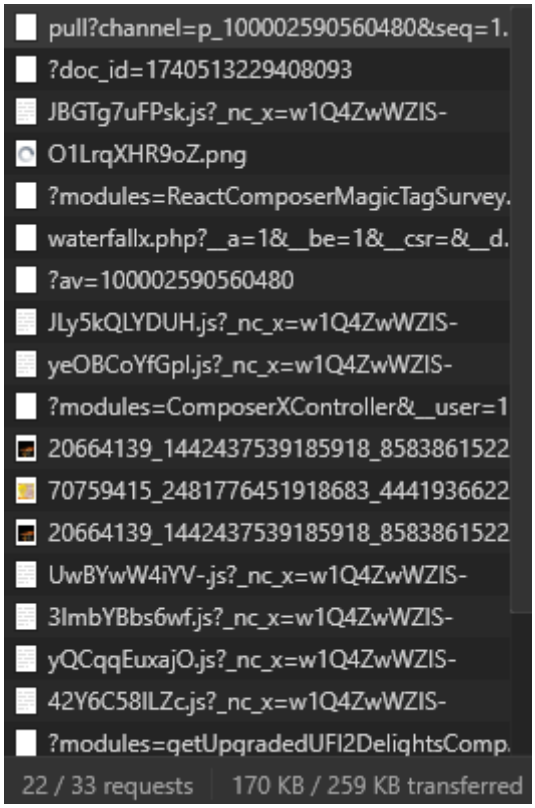


```
9fsg+HXDxighl7sFU4C4cn/Afv5WrxTOllOttpdvg2UPIArSb8mC2/yecIMNnAsa9uDITcVLWiZvwLAQ8mjPmO9h32qTsDcw3FPla2V6Bxw+7RsQWUqkDv3m9lgGcDFb7f
rbDm3NFtFzeIkc9Rio4AlUMuTGrRkJgASJOXpL5enRXF//QIa3KT1+DeCFWX3YAjCcXYuHHnix51IBoa77soLcVY=; path=/; secure; HttpOnly
Set-Cookie: X-OWA-CANARY=Vd0wkAg8_0GFnyxHK8Amn0CO2DZjQtcY62Pyel0cn3i6InJY0v6eiCrACfkS-E1OrmwvFg5H3tI.; path=/; secure
Strict-Transport-Security: max-age=31536000; includeSubDomains
Strict-Transport-Security: max-age=31536000; includeSubDomains
Vary: Accept-Encoding
X-BackEnd-Begin: 2019-09-26T09:23:38.548
X-BackEnd-End: 2019-09-26T09:23:38.717
X-BackEndHttpStatus: 200
X-BackEndHttpStatus: 200
X-BEServer: VI1PR03MB4015
X-BeSku: WCS5
X-CalculatedBETarget: VI1PR03MB4015.EURPRD03.PROD.OUTLOOK.COM
X-CalculatedFETarget: VI1PR07CU007.internal.outlook.com
X-Content-Type-Options: nosniff
X-DiagInfo: VI1PR03MB4015
X-FEProxyInfo: VI1PR07CA0234.EURPRD07.PROD.OUTLOOK.COM
X-FEServer: VI1PR07CA0234
X-FEServer: MRXP264CA0006
X-FrontEnd-Begin: 2019-09-26T09:23:38.546
X-FrontEnd-End: 2019-09-26T09:23:38.727
X-MailboxGuid: 6d933a1d-423a-4307-88ba-03fb32aff209
x-ms-appId: 00000002-0000-0ff1-ce00-000000000000
X-OWA-CorrelationId: a0ff24ab-4018-82db-7762-086531db2ed7
X-OWA-DiagnosticsInfo: 168;35;0
X-OWA-HttpHandler: true
X-OWA-MinimumSupportedOWSVersion: V2_6
X-OWA-OWSVersion: V2018_01_18
X-OWA-Version: 15.20.2284.28
X-Proxy-BackendServerStatus: 200
X-RUM-Validated: 1
X-UA-Compatible: IE=EmulateIE7
```

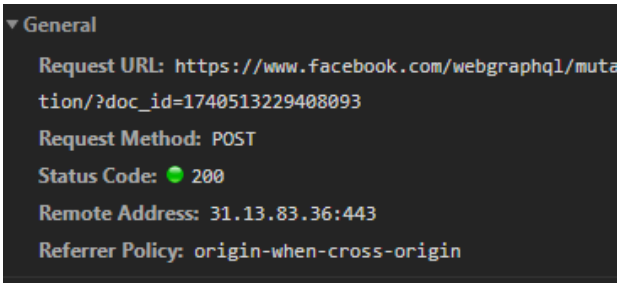
El otro "action" está asociado a la actualización de mails en la bandeja de entrada, y no forma parte del proceso de envío del email.

POST EN FACEBOOK

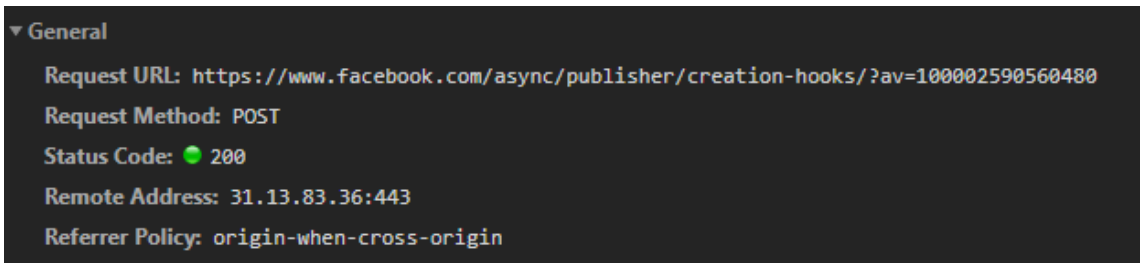
Ahora probaremos con un post en Facebook. Se crea una entrada y se publica, mientras observamos el registro del navegador.

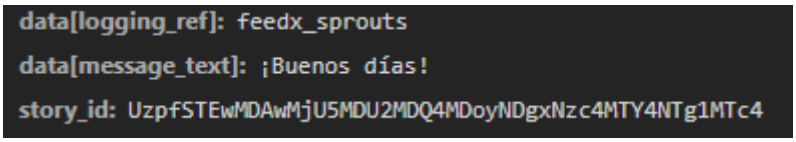


En este caso, se registran muchas solicitudes con diversos nombres, así que se busca en ellas las que sean del tipo POST.



Encontramos una primera solicitud, pero observando los datos no se puede encontrar ningún dato relevante relacionado con el contenido de la publicación hecha. Sin embargo, más adelante, aparece otra solicitud POST, también correcta, donde se puede encontrar el texto que se ha publicado dentro de "Form data".





El texto de la publicación aparece con la etiqueta "data[message_text]", y se muestra directamente. También aparecen otras etiquetas que Facebook utiliza para gestionar las publicaciones, como el ID del post.

CONCLUSIONES

Aunque el protocolo **HTTP** ha recibido varias revisiones a lo largo de los años, con esta práctica hemos observado que las bases se mantienen. Es interesante ver que los métodos de petición o GETs y los códigos de respuesta se han mantenido idénticos y funcionales a lo largo de los años. También cabe resaltar la incorporación del protocolo **HTTPs**, que incorpora filtros y medidas de seguridad entre la conexión cliente/servidor, y que en más de una web a la que hemos accedido por http nos ha redirigido a su versión https.