Subject: Http file download question Posted by koldo on Wed, 23 May 2012 13:33:55 GMT

View Forum Message <> Reply to Message

Hello all

I am an absolute dummy in Http.

Based in nice GuiWebDownload demo I wanted to download a file. It works for default address "http://downloads.sourceforge.net/project/upp/upp/4179/upp-x1 1-src-4179.tar.gz" but not the intranet address I am trying. I this case it just returns .html text enclosed below.

Is there perhaps a way to do something else to inform the http server that I want to download that file?

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN"
"http://www.w3.org/TR/html4/frameset.dtd">
<!-- turing cluster prod -->
<html>
 <head>
  <title>myintranet.com</title>
  <meta name="keywords" content="myintranet.com">
  <meta name="description" content="myintranet.com">
  <meta name="robots" content="INDEX, FOLLOW">
  <meta name="revisit-after" content="10">
  <meta http-equiv="Content-Type" content="text/html; charset=utf-8">
  <script type="text/javascript">
cookie callback = function(val) {
  var exdate=new Date();
  exdate.setFullYear(exdate.getFullYear() + 1);
  document.cookie = "Spusr=" + escape(val) +
   "; expires=" + exdate.toUTCString();
  document.cookie = "isc=1";
}
</script> <script
src="http://dsparking.com/?epl=K1r1PWFL4whaUopc-k sfvl965SQUDhFchf sgtFFDZmEM9xLcw
w6D2D4Fog9H3fDGpm0AEDjY5GA22FLci0zrMUnBwoBQN6Tt3TMBIFCvuC9zXtFdo4BnVJ0lZSo
kakthURfpJVnIYDJnsBrmd8X_jATD0AqH6q6ZmoaZqmn6pBGgYa0MQ0MuWpBpUAIGDc778AA
OB AQAAQIBbCgAA-X3rIFITJIIBMTZoWkKdAAAA8A">
</script>
 </head>
 <frameset rows="100%,*" frameborder="no" border="0" framespacing="0">
src="http://www.myintranet.com?epl=VFsGxNHfvy2uTVDVdXhhUFxPrIIgoXCK5C6-cs2zArDoQJ
QGnMlOuLYGd9Q I1m4dj9N7K7fm5b0TV4HE0dOXJ8yKBCygFhuCJ27BoeKHlQw2lYtFz PpTW
```

ijmozkHpqHtVo4umpGgAAANAocgAgENynvwAAoH8BAABAgNsHAACuJ1ZBWVMmWUExNmha QnIAAADw" name="myintranet.com">

</frameset>
<noframes>
<body>

Click here to go to myintranet.com .

</body>
</noframes>

</html>

Subject: Re: Http file download question

Posted by mirek on Wed, 23 May 2012 14:58:25 GMT

View Forum Message <> Reply to Message

Well, my first guess is that the download link does not actually lead to file to be downloaded, but is redirected via some javascript voodoo.

General advice to debug such problems: HttpRequest::Trace(), then look into .log... (or post here)

Subject: Re: Http file download question

Posted by koldo on Wed, 23 May 2012 16:51:23 GMT

View Forum Message <> Reply to Message

Thank you Mirek

Log is here:

HTTP START

Starting status 2 'Resolving host name', url: myintranet

HTTP Execute: Resolving host name

HTTP StartConnect HTTP AfterConnect

Starting status 6 'Sending request', url: myintranet

HTTP REQUEST myintranet:0

HTTP request:

GET http://myintranet/data/myfile/myfile.pdf HTTP/1.1

URL: http://myintranet/data/myfile/myfile.pdf

Host: myintranet Connection: close

Accept: */*

Accept-Encoding: gzip

User-Agent: U++ HTTP request

Proxy-Authorization: Basic MjM4NTM6c29ycHJlc2E=

HTTP Execute: Sending request

Starting status 7 'Receiving header', url: myintranet

HTTP Execute: Receiving header

HTTP Header received:

HTTP/1.1 307 Temporarily Moved for Domain Name Expansion

Mime-Version: 1.0

Date: Wed, 23 May 2012 18:45:30 CEST

Connection: close

Proxy-Connection: close

Location: http://www.myintranet.com/data/myfile/myfile.pdf

Content-Length: 1670

HTTP status code: 307

Starting status 8 'Receiving content', url: myintranet

HTTP Execute: Receiving content

HTTP reading data 1670

HTTP Out 1670

--- HTTP redirect http://www.myintranet.com/data/myfile/myfile.pdf

HTTP START

Starting status 2 'Resolving host name', url: www.myintranet.com

HTTP Execute: Resolving host name

HTTP StartConnect HTTP AfterConnect

Starting status 6 'Sending request', url: www.myintranet.com

HTTP REQUEST www.myintranet.com:0

HTTP request:

GET http://www.myintranet.com/data/myfile/myfile.pdf HTTP/1.1

URL: http://www.myintranet.com/data/myfile/myfile.pdf

Host: www.myintranet.com

Connection: close

Accept: */*

Accept-Encoding: gzip

User-Agent: U++ HTTP request

Proxy-Authorization: Basic MjM4NTM6c29ycHJlc2E=

HTTP Execute: Sending request

Starting status 7 'Receiving header', url: www.myintranet.com

HTTP Execute: Receiving header

HTTP Header received: HTTP/1.1 200 OK

Cache-Control: private, no-cache, must-revalidate

Pragma: no-cache

Server: Oversee Turing v1.0.0

Content-Encoding: gzip Content-Type: text/html

Expires: Mon, 26 Jul 1997 05:00:00 GMT

Keep-Alive: timeout=3, max=91

P3P: policyref="http://www.dsparking.com/w3c/p3p.xml", CP="NOI DSP COR ADMa OUR NOR

STA"

Set-Cookie: parkinglot=1; domain=.myintranet.com; path=/; expires=Thu, 24-May-2012 16:45:30

GMT

Content-Length: 947

Date: Wed, 23 May 2012 16:45:30 GMT

Via: 1.1 ironport.myhost.com:8080 (IronPort-WSA/7.1.3-021)

Connection: close

Proxy-Connection: close

HTTP status code: 200

Starting status 8 'Receiving content', url: www.myintranet.com

HTTP Execute: Receiving content

HTTP reading data 947

HTTP Out 1744

Subject: Re: Http file download question

Posted by mirek on Sun, 03 Jun 2012 10:44:09 GMT

View Forum Message <> Reply to Message

Well, I suspect there is some javascript voodoo involved.

I would try putting the url into Google Chrome and use its excellent logging to find out what is the hell going on.

Subject: Re: Http file download question

Posted by koldo on Sun, 03 Jun 2012 13:39:28 GMT

View Forum Message <> Reply to Message

Thank you Mirek

Unfortunately it is an intranet address.

Subject: Re: Http file download guestion

Posted by tojocky on Mon, 04 Jun 2012 00:58:35 GMT

View Forum Message <> Reply to Message

Hi Koldo,

Welcome to web development.

When you reply to server, you have to add in response package header the correct mime type and 'Content-Disposition: attachment;'

for more details you can find on Wikipedia website.

Also very good lessons you can take from Node.JS + Connect and Express.JS packages.

I had a requirement to tell to client that it should be downloaded with a specific file name.

I hope if this a solution what are you looking for!

Update: Would you like to implement this job on client side or on server side?

Subject: Re: Http file download question

Posted by koldo on Mon, 04 Jun 2012 06:59:27 GMT

View Forum Message <> Reply to Message

Thank you lon

This is a kind of black magic for me.

I want to download and upload files from client side. I expected that there would be a direct way to do it (and in addition, a way to list the available files)

I do not know what is the "content type" to get a file. And now HttpRequest does not have a way to set "Content-Disposition".

Subject: Re: Http file download question

Posted by mirek on Mon. 04 Jun 2012 09:45:51 GMT

View Forum Message <> Reply to Message

koldo wrote on Mon, 04 June 2012 02:59

I do not know what is the "content type" to get a file. And now HttpRequest does not have a way to set "Content-Disposition".

.Header("Content-Disposition", "attachement");

However, I doubt it is the problem.

The most direct way is to use chrome for http log (it will show you all http requests made by browser to download the file).

Subject: Re: Http file download question

Posted by koldo on Mon, 04 Jun 2012 10:54:23 GMT

View Forum Message <> Reply to Message

Hello Mirek

I have tried it with Chrome. I have entered in Developer Tools. However I do not know how to log html.

Subject: Re: Http file download question

Posted by tojocky on Mon, 04 Jun 2012 20:08:09 GMT

View Forum Message <> Reply to Message

it exact what I expected.

Try this link. it is dirrect link.

Ok.

- 1. When you open the link a JavaScript event On page load is attached.
- 2. After the time is elapsed, dynamically is created an hidden control (frame with link... or can be something elese. it less important) and it is fired.
- 3. It open the link: http://downloads.sourceforge.net/project/upp/upp/4179/upp-x11-src-4179.tar.gz?r=http%3A%2F%2Fwww.ultimatepp.org%2Fforum%

2Findex.php%3Ft%3Dmsg%26th%3D6738%26start%3D0%26&ts=1338

838532&use_mirror=switch

with header:

Request

URL:http://downloads.sourceforge.net/project/upp/upp/4179/upp-x11-src-4179.tar.gz?r=http%3A %2F%2Fwww.ultimatepp.org%2Fforum%2Findex.php%3Ft%3Dmsg%26th%3D6738%26start%3 D0%26&ts=1338838532&use mirror=switch

Request Method:GET

Status Code:302 Found

Request Headersview source

Accept:text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8

 $Accept-Charset: ISO-8859-1, utf-8; q=0.7, ^*; q=0.3$

Accept-Encoding:gzip,deflate,sdch

Accept-Language:en-US,en;q=0.8,ru;q=0.6,en-GB;q=0.4

Connection:keep-alive

Cookie:__utmx=191645736.00024592574091748957:1:1;

- __utmxx=191645736.00024592574091748957:1338838533:2592000;
- __utma=191645736.866235001.1338670652.1338810779.1338838044.5;
- utmb=191645736.8.9.1338838539021; utmc=191645736;
- __utmz=191645736.1338838044.5.5.utmcsr=ultimatepp.org|utmccn=(referral)|utmcmd=referral|utmcct=/forum/index.php; __utmv=191645736.|5=AB%20Test=4091748957%3A1=1

Host:downloads.sourceforge.net

Referer: http://sourceforge.net/projects/upp/files/upp/4179/upp-x11-src-4179.tar.gz/download?use

mirror=switch

User-Agent: Mozilla/5.0 (X11; Linux i686) AppleWebKit/536.5 (KHTML, like Gecko)

Chrome/19.0.1084.52 Safari/536.5

Query String Parametersview URL encoded

r:http://www.ultimatepp.org/forum/index.php?t=msg&th=6738&start=0&

ts:1338838532

use_mirror:switch

Response Headersview source

Access-Control-Allow-Origin:*

Content-Disposition:attachment; filename="upp-x11-src-4179.tar.gz"

Content-Length:0

Content-type:text/html

Date:Mon, 04 Jun 2012 19:35:39 GMT

Location:http://switch.dl.sourceforge.net/project/upp/upp/4179/upp-x11-src-4179.tar.gz

Server:lighttpd/1.4.26

Set-cookie:sf_mirror_attempt=upp:switch:/upp/upp/4179/upp-x11-src-4179.tar.gz; expires=Mon,

4-Jun-2012 19:37:39 GMT; Path=/

4. browser redirects to the advised link:

http://switch.dl.sourceforge.net/project/upp/upp/4179/upp-x1 1-src-4179.tar.gz with header:

Request URL:http://switch.dl.sourceforge.net/project/upp/upp/4179/upp-x11-src-4179.tar.gz

Request Method:GET Status Code:200 OK

Request Headersview source

Accept:text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8

Accept-Charset:ISO-8859-1,utf-8;q=0.7,*;q=0.3

Accept-Encoding:gzip,deflate,sdch

Accept-Language:en-US,en;q=0.8,ru;q=0.6,en-GB;q=0.4

Connection:keep-alive

Cookie: utmx=191645736.00024592574091748957:1:1;

utmxx=191645736.00024592574091748957:1338838533:2592000:

__utma=191645736.866235001.1338670652.1338810779.1338838044.5;

__utmb=191645736.8.9.1338838539021; __utmc=191645736;

utmz=191645736.1338838044.5.5.utmcsr=ultimatepp.org|utmccn=(referral)|utmcmd=referral|ut

mcct=/forum/index.php; utmv=191645736.|5=AB%20Test=4091748957%3A1=1

Host:switch.dl.sourceforge.net

Referer:http://sourceforge.net/projects/upp/files/upp/4179/upp-x11-src-4179.tar.gz/download?use

_mirror=switch

User-Agent: Mozilla/5.0 (X11; Linux i686) AppleWebKit/536.5 (KHTML, like Gecko)

Chrome/19.0.1084.52 Safari/536.5

Response Headersview source

Accept-Ranges:bytes Connection:Keep-Alive

Content-Length:38392395

Content-Type:application/x-gzip

Date:Mon, 04 Jun 2012 19:35:39 GMT

ETag:"1e99f7-249d24b-4b20c1b1bc8c0" Keep-Alive:timeout=15, max=100 Last-Modified:Sat, 19 Nov 2011 01:09:31 GMT Server:Apache/2.2.16 (Debian)

be careful to the response Content-Disposition:attachment; filename="upp-x11-src-4179.tar.gz"

this tag server tells to browser that this should be downloaded.

Hope if helped you.

Subject: Re: Http file download question Posted by mirek on Tue, 05 Jun 2012 06:21:41 GMT

View Forum Message <> Reply to Message

tojocky: With all respect, I believe your response is offtopic for this case.

koldo: In Chrome, enter "www.google.com", then click on white area and chosse "Inspect element" from local menu. Bottom pane should be displayed. Choose "Network", then enter your intranet url, Chrome then will log all http traffic.

Subject: Re: Http file download question Posted by koldo on Tue. 05 Jun 2012 07:01:30 GMT

View Forum Message <> Reply to Message

Hello

I have tried to follow Mirek instructions:

Request URL:http://myintranet/myfile.pdf

Request Method:GET

Status Code: 304 Not Modified

Request Header

Accept:text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8

Accept-Charset:ISO-8859-1,utf-8;q=0.7,*;q=0.3

Accept-Encoding:gzip,deflate,sdch Accept-Language:es-ES,es;q=0.8

Authorization: Negotiate

TIRMTVNTUAADAAAAGAAYAGoAAAAYABgAggAAAAoACgBIAAAACgAKAFIAAAAOAA4AXAA AAAAAAACaAAAABYKIogUBKAoAAAAPTgBUAEIASQBPADIANAAzADMAMABCAFAAQwA4A

==

Cache-Control:max-age=0
Connection:keep-alive

Cookie:UserType=P; idioma=es%2DES;

ASPSESSIONIDQABDBSDD=DCILLDNCIFEHFLHEEILJJEAD

Host:myintranet

If-Modified-Since:Thu, 10 Nov 2011 14:49:40 GMT

If-None-Match: "3a5dcbf9b79fcc1:0"

Range:bytes=0-1048575

User-Agent: Mozilla/5.0 (Windows NT 5.1) AppleWebKit/536.5 (KHTML, like Gecko)

Chrome/19.0.1084.52 Safari/536.5

Response Headers Accept-Ranges:bytes

Date:Tue, 05 Jun 2012 06:50:15 GMT

ETag: "3a5dcbf9b79fcc1:0"

Last-Modified:Thu, 10 Nov 2011 14:49:40 GMT

MS-Author-Via:MS-FP/4.0

MicrosoftOfficeWebServer:5.0_Pub

Server:Microsoft-IIS/7.0 X-Powered-By:ASP.NET

In addition I have found that when entering the address directly in the browser I get the file downloaded if I use Internet Explorer and Chrome. However if I use Firefox I get the same html page I get when using the U++ program.

Subject: Re: Http file download question

Posted by tojocky on Tue, 05 Jun 2012 21:02:05 GMT

View Forum Message <> Reply to Message

Mirek: Thank you. maybe I'm wrong.

Koldo: if I understood you correctly, you want to download a file from a server. You want to make the sever to communicate with browser for downloading a file. If yes, then I think you should put in response header package tags (you missed):

- 1. Content-Type: more detaild is described here
- 2. Content-Disposition:attachment; filename="<wanted file name for save>"
- 3. Content-Length: <your file length>

Let me know if I'm wrong.

Sorry for confuse you if I was.

Subject: Re: Http file download question

Posted by koldo on Wed, 06 Jun 2012 06:29:00 GMT

View Forum Message <> Reply to Message

Thank you lon

I wanted to do a program in U++ to download a file. Because of it I used GuiWebDownload as a base.

About this:

Quote: 1. Content-Type: more detailed is described here

I have used this: http.Header("Content-Type", "application/octet-stream");

Quote:2. Content-Disposition:attachment; filename="<wanted file name for save>"

I have used this: http.Header("Content-Disposition", "attachment; filename=<fileName>");

Quote:3. Content-Length: <your file length>

I do not know file length before downloading it. Does it serve when uploading files?

Subject: Re: Http file download question

Posted by mirek on Fri, 08 Jun 2012 13:52:43 GMT

View Forum Message <> Reply to Message

koldo wrote on Tue, 05 June 2012 03:01

In addition I have found that when entering the address directly in the browser I get the file downloaded if I use Internet Explorer and Chrome. However if I use Firefox I get the same html page I get when using the U++ program.

Well, obviously your intranet is using some really specific tricks...

I would concentrate now on real meaning of http codes 304 and 307.

Subject: Re: Http file download question

Posted by koldo on Mon, 25 Jun 2012 16:34:39 GMT

View Forum Message <> Reply to Message

Hello Mirek

I have done it successfully using WGET. The output is:

C:\>wget http://myintranet/myfile.pdf --http-user=MYUSERNAME

--http-password=MYPASSWORD -Omyfile.pdf

--2012-06-25 18:17:31-- http://myintranet/myfile.pdf

Resolving myintranet... 122.11.4.32

Connecting to myintranet 122.11.4.32 : 80... connected.

HTTP request sent, awaiting response... 401 Unauthorized

Connecting to myintranet|122.11.4.32|:80... connected.

HTTP request sent, awaiting response... 401 Unauthorized

Reusing existing connection to myintranet:80. HTTP request sent, awaiting response... 200 OK

Length: 2127943 (2.1M) [application/pdf]

Saving to: `myfile.pdf'

100%[========] 2,127,943 --.-K/s in 0.04s

2012-06-25 18:17:31 (73.0 MB/s) - `mydoc.pdf' saved [1707881/1707881]

Subject: Re: Http file download question

Posted by koldo on Tue, 26 Jun 2012 09:49:25 GMT

View Forum Message <> Reply to Message

Hello Mirek

It also works using CURL. The trick is to add --ntlm option (from CURL doc: The NTLM authentication method was designed by Microsoft and is used by IIS web servers. It is a proprietary protocol, reverse-engineered by clever people and implemented in curl based on their efforts):

C:\>curl -u myusername:mypassword --ntlm -O http://myintranet/myfile.pdf -v

- * About to connect() to myintranet port 80 (#0)
- * Trying 152.41.2.31... % Total % Received % Xferd Average Speed Time Time Time Current

Dload Upload Total Spent Left Speed

nnected

- * Connected to myintranet (152.41.2.31) port 80 (#0)
- * Server auth using NTLM with user 'myusername'
- > GET /myfile.pdf HTTP/1.1
- > Authorization: NTLM

- > User-Agent: curl/7.21.4 (i386-pc-win32) libcurl/7.21.4 OpenSSL/0.9.8r zlib/1.2 .5
- > Host: myintranet
- > Accept: */*

>

- < HTTP/1.1 401 Unauthorized
- < Content-Type: text/html; charset=us-ascii
- < Server: Microsoft-HTTPAPI/2.0
- < WWW-Authenticate: NTLM

TIRMTVNTUAACAAAACgAKADgAAAA1gonippnZh0iAu7MAAAAAAAAAAIQ

AhABCAAAABgByFwAAAA9OAFQAQgBJAE8AAgAKAE4AVABCAEkATwABAAoAQgBOAFQAN gAyAAQAGABiAGk

AbwAuAHMAZQBuAGUAcgAuAGUAcwADACQAQgBOAFQANgAyAC4AYgBpAG8ALgBzAGUAbgBiAHIALgBiAHM

ABQAQAHMAZQBuAGUAcgAuAGUAcwAHAAgAFiLFkH1TzQEAAAAA

```
< Content-Length: 341
* Ignoring the response-body
{ [data not shown]
100 341 100 341 0
                        0 21312
                                   0 --:--:- 333k*
Connection #0 to host myintranet left intact
* Issue another request to this URL: 'http://myintranet/myfile.pdf'
* Re-using existing connection! (#0) with host myintranet
* Connected to myintranet (152.41.2.31) port 80 (#0)
* Server auth using NTLM with user 'myusername'
> GET /myfile.pdf HTTP/1.1
> Authorization: NTLM
TIRMTVNTUAADAAAAGAAYAGoAAAAYABgAggAAAAoACgBIAAAACgAKAFIAAA
AOAA4AXAAAABAAEACaAAAANYKI4gUBKAoAAAAPUwBFAE4ARQBSADIAMwA4ADUAMwB
CAFAAQwA2ADUAOA
AxAFaQ7c+fdOfVAAAAAAAAAAAAAAAAAAAAAAAALjckl63UGzaggVOVuI+10LsuaHg+iRwvcfBV7+I
IKyYII
pwvm0d4iU=
> User-Agent: curl/7.21.4 (i386-pc-win32) libcurl/7.21.4 OpenSSL/0.9.8r zlib/1.2
.5
> Host: myintranet
> Accept: */*
< HTTP/1.1 200 OK
< Content-Type: application/pdf
< Last-Modified: Thu, 10 Nov 2011 14:49:40 GMT
< Accept-Ranges: bytes
< ETag: "3a5dcbf9b79fcc1:0"
< Server: Microsoft-IIS/7.0
< X-Powered-Bv: ASP.NET
< MicrosoftOfficeWebServer: 5.0 Pub
< MS-Author-Via: MS-FP/4.0
< Date: Tue, 26 Jun 2012 09:25:02 GMT
< Content-Length: 1707881
<
{ [data not shown]
100 1667k 100 1667k 0
                          0 34.6M
                                     0 --:--:- 34.6M*
Connection #0 to host myintranet left intact
* Closing connection #0
```

Subject: Re: Http file download question

Posted by mirek on Tue, 26 Jun 2012 10:37:12 GMT

View Forum Message <> Reply to Message

< Date: Tue, 26 Jun 2012 09:25:02 GMT

Maybe it just needs password... Have you provided some?

(Using User or Digest)

Subject: Re: Http file download question

Posted by koldo on Tue, 26 Jun 2012 11:03:11 GMT

View Forum Message <> Reply to Message

Hello Mirek

Yes of course. If I do not put any password it does not download anything.

With the password it downloads the html page in the first post.

It is funny that I get the same html page using Firefox but I get the file using Chrome or Explorer.

Subject: Re: Http file download question

Posted by mirek on Tue, 26 Jun 2012 11:15:48 GMT

View Forum Message <> Reply to Message

koldo wrote on Tue, 26 June 2012 07:03Hello Mirek

Yes of course. If I do not put any password it does not download anything.

With the password it downloads the html page in the first post.

It is funny that I get the same html page using Firefox but I get the file using Chrome or Explorer.

Well, at least we are as good as Firefox..

I guess it has something to with that NTLM thing...