

Forum:[Hack Forums](#)

Thread: [Downloading a Webpage's Source \(C++\)](#)

dotcppfile - May 5, 2014, 11:31 am

Hello everyone,

Well the title says it all.

I am actually working on a simple injector to be released sooner or later for personal and educational purposes.

I chose to download the source of a page to check any kind of change, now some people may find us kind of odd but I do really think that getting about 2 kb of text and do whatever I want to do with them is actually a pretty good thing and a fast one as well.

Anyways, I managed to work with some libraries in c++ such as urlmon.lib and I wasn't really sure I was doing it the right way because with urlmon you have to download the source to a text file(I am pretty sure that's the only way, not sure though) but now I kept on moving with it and I made it work amazingly but I am looking for something much faster; probably something like downloading the source and putting it in a string instantly with no further do which will increase the speed by some milliseconds which does actually matter in such a project.

I tried working with sockets but it's much slower and a waste of computer power etc...

I am not looking for any source or anything similar but I am asking for ideas.

What do you think?

What should I use instead? (By the way, I am happy working with the webpage's source itself so if you're willing to help me out in this then please stick to the same method)

If you have any other ideas, libs or anything similar that may be useful then it would be nice to share it with me, again, I am looking for the fastest way to do achieve my goal.

Thanks in advance,
dotcppfile.

matty9 - May 5, 2014, 2:24 pm

Check out [Curl](#), ive used it and works very well

Cryptodemon - May 5, 2014, 2:27 pm

ws2_32.dll

send/recv

[http://msdn.microsoft.com/en-us/library/windows/desktop/ms740149\(v=vs.85\).aspx](http://msdn.microsoft.com/en-us/library/windows/desktop/ms740149(v=vs.85).aspx)

Obergruppenführer - May 5, 2014, 4:11 pm

(May 5, 2014 02:27 PM)Cryptodemon Wrote: [\[CITE\]](#) ws2_32.dll

send/recv

[http://msdn.microsoft.com/en-us/library/windows/desktop/ms740149\(v=vs.85\).aspx](http://msdn.microsoft.com/en-us/library/windows/desktop/ms740149(v=vs.85).aspx)

He stated that he doesn't want to use sockets.

Look up Wininet.

danyllo - May 5, 2014, 4:33 pm

I tried working with sockets but it's much slower and a waste of computer power etc...

Just quit.

I tried working with sockets but it's much slower and a waste of computer power etc...

Just quit.

0xDEAD10CC - May 5, 2014, 6:40 pm

You don't need curl, you don't need winsock, if you're using Windows, WinInet should work.

InternetOpen
InternetConnect
HttpOpenRequest
HttpSendRequest
InternetReadFile
InternetCloseHandle

There are others (querying request status, etc), but these will get you as far getting the source of a webpage from a GET request (HttpOpenRequest), unless you need to POST data, which you can do too with the same function.

It's actually pretty simple. Even if you wanted to use winsock, it's not that difficult once you understand what each of the core/'main' functions do.

dotcppfile - May 5, 2014, 8:49 pm

(May 5, 2014 06:40 PM)0xDEAD10CC Wrote: [\[CITE\]](#) You don't need curl, you don't need winsock, if you're using Windows, WinInet should work.

InternetOpen
InternetConnect
HttpOpenRequest
HttpSendRequest
InternetReadFile
InternetCloseHandle

There are others (querying request status, etc), but these will get you as far getting the source of a webpage from a GET request (HttpOpenRequest), unless you need to POST data, which you can do too with the same function.

It's actually pretty simple. Even if you wanted to use winsock, it's not that difficult once you understand what each of the core/'main' functions do.

I did work with winsock as well but I realised it's kind of slow when it comes to urlmon and what I am using it for. Anyways, thanks for sharing your idea, I will give it a shot today.

(May 5, 2014 02:24 PM)matty9 Wrote: [\[CITE\]](#) Check out [Curl](#), ive used it and works very well I heard about Curl but I never used it, will give it a shot today once I get back to my desk, thanks for sharing.

(May 5, 2014 02:27 PM)Cryptodemon Wrote: [\[CITE\]](#) ws2_32.dll

send/recv

[http://msdn.microsoft.com/en-us/library/windows/desktop/ms740149\(v=vs.85\).aspx](http://msdn.microsoft.com/en-us/library/windows/desktop/ms740149(v=vs.85).aspx)

True, it does the trick but I used it and it was kind of slow when it comes to other libs.

_SLQ - May 6, 2014, 5:16 am

Winsock is slow? Sure, if you fuck up in how to program correctly with it.

You can achieve performance beyond what any library provides you, if you know what you are doing.

For your information, WinInet, libCurl and whatever external source that exists on Windows - it's just a wrapper around Winsock.

One exception is CLR .NET. I'm not exactly sure if managed environments either internally make native calls to already existing API's, or they're re-written completely.

OP: Just connect() to a DNS/IP address, send() a GET request to the server, recv() the HTTP header & parse for the Content-Length and recv() again that much. Your buffer has your page. Don't forget to closesocket().

Quote:HTTP/1.1 200 OK

Date: Mon, 23 May 2005 22:38:34 GMT

Server: Apache/1.3.3.7 (Unix) (Red-Hat/Linux)

Last-Modified: Wed, 08 Jan 2003 23:11:55 GMT

ETag: "3f80f-1b6-3e1cb03b"

Content-Type: text/html; charset=UTF-8

Content-Length: 131

Accept-Ranges: bytes

Connection: close

protocolback - May 6, 2014, 5:58 am

I say, go with socket. cURL sucks, believe me

I wrote this for a project and works pretty well

Spoiler ([Click to View](#))Code #ifndef WEBCLIENT_HPP

#define WEBCLIENT_HPP

```
#include <sys/socket.h>
```

```
#include <arpa/inet.h>
```

```
#include <netinet/in.h>
```

```
#include <sys/types.h>
```

```
int main_socket = 0;
```

```
void use(int sock)
```

```
{
```

```
    if(main_socket==0)
```

```
        main_socket = sock;
```

```
}
```

```
bool socket_connect(const char* ip, int port)
```

```
{
```

```
    if(main_socket != 0)
```

```
    {
```

```
        struct sockaddr_in sin;
```

```
        sin.sin_family=AF_INET;
```

```

    sin.sin_port=htons(80);
    sin.sin_addr.s_addr=inet_addr(ip);

    return (connect(main_socket, (struct sockaddr*)&sin, sizeof(sin)) < 0) ? false : true;
}
else
    return false;
}

std::string getString(const char* ip, int port=80)
{
    if(socket_connect(ip, port))
    {
        std::string data = "GET / HTTP/1.1\r\nHost: webclient\r\nConnection: keep-alive\r\n\r\n";
        if(send(main_socket, data.c_str(), data.length(), 0) > 0)
        {
            char buffer[1024*1024];
            if(recv(main_socket, buffer, 4096, 0) > 0)
            {
                return buffer;
            }
        }
    }
}

#endif

```

dotcppfile - May 6, 2014, 6:46 am

(May 6, 2014 05:58 AM)protocolback Wrote: [\[CITE\]](#) I say, go with socket. cURL sucks, believe me I wrote this for a project and works pretty well

Spoiler ([Click to View](#))Code #ifndef WEBCLIENT_HPP

#define WEBCLIENT_HPP

```

#include <sys/socket.h>
#include <arpa/inet.h>
#include <netinet/in.h>
#include <sys/types.h>

```

```

int main_socket = 0;

```

```

void use(int sock)
{
    if(main_socket==0)
        main_socket = sock;
}

```

```

bool socket_connect(const char* ip, int port)
{
    if(main_socket != 0)
    {
        struct sockaddr_in sin;
        sin.sin_family=AF_INET;
        sin.sin_port=htons(80);
    }
}

```

```

sin.sin_addr.s_addr=inet_addr(ip);

return (connect(main_socket, (struct sockaddr*)&sin, sizeof(sin)) < 0) ? false : true;
}
else
return false;
}

std::string getString(const char* ip, int port=80)
{
if(socket_connect(ip, port))
{
std::string data = "GET / HTTP/1.1\r\nHost: webclient\r\nConnection: keep-alive\r\n\r\n";
if(send(main_socket, data.c_str(), data.length(), 0) > 0)
{
char buffer[1024*1024];
if(recv(main_socket, buffer, 4096, 0) > 0)
{
return buffer;
}
}
}
}
}

```

#endif

True, but in my case this have to be called over and over and the same connection, sockey etc will be created and openned multiply times since the Link changes over and over. I am not sure it will be fast enough but I will give it a try, thanks for sharing.

(May 6, 2014 05:16 AM)_SLQ Wrote: [CITE](#) Winsock is slow? Sure, if you fuck up in how to program correctly with it.

You can achieve performance beyond what any library provides you, if you know what you are doing.

For your information, WinInet, libCurl and whatever external source that exists on Windows - it's just a wrapper around Winsock.

One exception is CLR .NET. I'm not exactly sure if managed environments either internally make native calls to already existing API's, or they're re-written completely.

OP: Just connect() to a DNS/IP address, send() a GET request to the server, recv() the HTTP header & parse for the Content-Length and recv() again that much. Your buffer has your page. Don't forget to closesocket().

Yh but same reply for protocal when it comes to it.

Okay, I guess I will just go for wininet for now. I will get back to curl or anything similar that would make this project work over any platform later.

dotcppfile - May 7, 2014, 10:49 am

Released [\[Release\] sqlinjector v0.1 beta \[C++ | Free\]+ Video](#) and Special thanks to all of you.