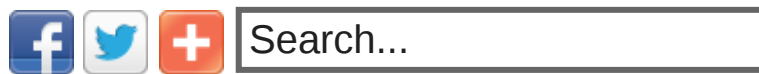


TECHSLIDES

Data, Maps, Usability, and Performance



Curl with NodeJS

Last updated on September 3, 2012 in [Development](#)



If you are familiar with [Curl](#), the command line tool for transferring data with URL syntax, you might wonder

how can that be done with [NodeJS](#). Well, looking at the [nodeJS API](#), it's clear that [HTTP.request](#) gives you something that looks like a Curl equivalent. But, I want to write the response to a file and do some looping so that I can use an array of URLs instead of just making a request to one URL.

Looking around, I found a great nodeJS module called [request](#) which is a small abstraction on top of `http.request` and makes writing my script much easier. First, I added the `request` and `filesystem` module to my script plus the array of URLs I want to scrape. Then I did a for-loop to iterate through the array and inside the loop, I put the `request` function, passing the file and url parameters, and some `console.log()` calls to print what's happening:

```
1  var fs = require('fs');
2  var request = require('request');
3
4  var urls = new Array("http://www.yahoo.com", "http://www.bing.com");
5
6  for (var i = 0; i < urls.length; i++) {
7
8      var file = 'log'+[i]+' .txt';
9      var url = urls[i];
10
11     console.log(url);
12     console.log(file);
13
14     request(url, function (error, response, body) {
15         if (!error && response.statusCode == 200) {
16             console.log('request url: '+url);
17             console.log('request file: '+file);
18             fs.writeFile(file, body);
19         }
20     })
21 }
```

```
20         });
21
22     }
23
24     /*
25     Output:
26
27     http://www.yahoo.com
28     log0.txt
29     http://www.bing.com
30     log1.txt
31     request url: http://www.bing.com
32     request file: log1.txt
33     request url: http://www.bing.com
34     request file: log1.txt
35
36     */
```

nodejs-curl-bad.js hosted with ♥ by GitHub

[view raw](#)

As you can see from the output, only the second url and file was scraped. This is because, in NodeJS, everything is asynchronous, so we need to pass the url and file from the loop to a separate function:

```
1  var fs = require('fs');
2  var request = require('request');
3
4  var urls = new Array("http://www.yahoo.com","http://www.bing.com");
5
6  function scrape(url,file){
7
8      request(url, function (error, response, body) {
9          if (!error && response.statusCode == 200) {
10             console.log('request url: '+url);
11             console.log('request file: '+file);
12             fs.writeFile(file, body);
13         }
14     });
15
16 }
17 for (var i = 0; i < urls.length; i++) {
18
19     var file = 'log'+[i]+' .txt';
```

```
20     var url = urls[i];
21
22     console.log(url);
23     console.log(file);
24
25     scrape(url, file);
26
27 }
28
29
30 /*
31 http://www.yahoo.com
32 log0.txt
33 http://www.bing.com
34 log1.txt
35 request url: http://www.bing.com
36 request file: log1.txt
37 request url: http://www.yahoo.com
38 request file: log0.txt
39 */
```

curl-nodejs.js hosted with ♥ by GitHub

[view raw](#)

Another good nodeJS module for doing curl like web scraping is [curlrequest](#). Of course, for some serious web scraping, I recommend using PhantomJS.

Tags: [JavaScript](#), [nodejs](#), [scraping](#)



0 Comments TechSlides

 Login ▾

 Recommend  Share

Sort by Best ▾






Start the discussion...

LOG IN WITH

OR SIGN UP WITH DISQUS 

Name

Be the first to comment.

 Subscribe  Add Disqus to your siteAdd DisqusAdd  Privacy

Archives

- August 2016
- July 2016
- June 2016
- July 2015
- June 2015
- May 2015
- April 2015
- March 2015
- February 2015
- January 2015
- December 2014
- November 2014
- October 2014
- September 2014
- August 2014

Categories

- Data Visualization
- Design
- Development
- Domains
- SEO
- Social Media
- Uncategorized

April 2014
March 2014
February 2014
January 2014
December 2013
November 2013
October 2013
September 2013
August 2013
July 2013
June 2013
May 2013
April 2013
March 2013
February 2013
January 2013
December 2012
November 2012
October 2012
September 2012
August 2012
July 2012
June 2012
May 2012

Copyright ©2016 TechSlides, All Rights Reserved