

SimplePie 1.x - Feature Request # 10: Re-write our HTTP support from scratch

Status:	New	Priority:	High
Author:	Ryan Parman	Category:	HTTP
Created:	2008-02-03	Assigned to:	Geoffrey Sneddon
Updated:	2008-12-20	Due date:	
Affected Version:	1.1.3		
PHP Version:	5.2.4		
mbstring enabled:	No		
iconv enabled:	No		
cURL enabled:	No		
zlib enabled:	No		
Subject:	Re-write our HTTP support from scratch		
Description:	Re-write our HTTP support from scratch, which will allow for things like SSL support, HTTP Proxies, HTTP 401: Digest Authentication, obey any cache expiry set in the HTTP headers, HTTP status codes, ability to see if a remote file exists without actually downloading it, etc.		

History

2008-02-04 08:58 - Geoffrey Sneddon

- Assigned to set to Geoffrey Sneddon
- Target version set to 1.3

Assigning this to myself, and setting to version 1.3.

2008-02-04 13:38 - Geoffrey Sneddon

- Subject changed from *Re-write SimplePie_File() class from scratch* to *Re-write our HTTP support from scratch*
- PHP Version set to 5.2.4

2008-03-26 01:52 - Rasmus Schultz

Are you writing your own HTTP client from the ground up?

You might want to take a look at this class:

[[<http://code.google.com/p/php-httpclient/>]]

It's a self-contained HttpClient with a clean PHP5-interface. I don't know if it does everything you need, but even if it doesn't, it might make a good starting point and save you a lot of time.

As per issue #8, it is a problem for me that the `subscribe_url()` method does not return the accurate, final URL of the feed. For example, setting the URL to "`http://www.devmaster.net/rss`" should have this function return "`http://www.devmaster.net/rss/`" (with the trailing slash), but doesn't.

So in your application, when two users request this feed with and without the trailing slash, you get two distinct feeds with individual URLs, which are actually one and the same.

My workaround thus far is to use another HTTP Client first, which follows redirects, to determine the final URL of a feed, before using SimplePie to fetch and parse it. It's twice the effort, as the feed has to be downloaded twice, first by the HTTP client, then by SimplePie. Can't think of any other solution at this point though...?

2008-03-26 09:16 - Geoffrey Sneddon

- Affected Version set to 1.1.1
- PHP Version set to 5.2.4
- mbstring enabled set to No
- iconv enabled set to No
- cURL enabled set to No

- *zlib enabled set to No*

@Rasmus: That doesn't seem that much better than what we currently have, really. Also, it has the issue for 1.3 of not supporting PHP 4 (and this, and what this blocks, is one of the main reasons to have a 1.3 at all).

2008-03-26 12:08 - Rasmus Schultz

@Geoff: well, it follows redirects correctly, which was the problem with the current implementation?

Using the `getRequestURL()` method after a request will give you the correct URL, e.g. `"http://devmaster.net/rss"`, `"http://devmaster.net/rss/"`, `"http://www.devmaster.net/rss"` and `"http://www.devmaster.net/rss/"` will all give you `"http://www.devmaster.net/rss/"`, I tested it.

This class is actually a PHP5 upgrade of the older PHP4-compatible version - it wouldn't be difficult to diff from the previous version and retrograde it for PHP4 compatibility.

Just trying to help :-)

2008-03-26 14:17 - Geoffrey Sneddon

@Rasmus: Just blindly returning the last URL you requested isn't right either: some responses are cacheable, others are not. Notably, you have permanent redirects and temporary ones, only the former of which can be cached. If we get into a complex edge case, with a permanent redirect leading to a temporary redirect leading to a permanent redirect, only the first is cacheable (as the second permanent redirect is reliant on a non-cacheable temporary redirect). Also, it makes the same mistake we currently do with Content-Coding: it actually trusts the header. This cannot be done in the real world. When it claims to be GZIP or DEFLATE, it can be any of GZIP, ZLIB, or DEFLATE.

2008-11-16 09:27 - Geoffrey Sneddon

- *Affected Version changed from 1.1.1 to 1.1.2*

2008-12-20 09:02 - Geoffrey Sneddon

- *Affected Version changed from 1.1.2 to 1.1.3*

2008-12-20 09:03 - Geoffrey Sneddon

- *Tracker changed from Defect to Feature Request*