Ruby master - Feature #6546
Net::HTTP to check for HTTP_PROXY environment setting.
06/05/2012 08:06 AM - dekz (Jacob Evans)

Status: Closed
Priority: Normal
Assignee: naruse (Yui NARUSE)
Target version: 2.0.0

Description

nahi (Hirosi Nakamura) suggested I request the OS environment HTTP_PROXY be honoured.

Open-uri checks for this environment setting and acts according whereas Net::HTTP does not. Not having Net::HTTP check for this condition sets precedence and stops the implementation of other environment settings (such as JRuby and JVM -Dhttp.proxyHost).

Having this functionality will greatly enable testing with proxies rather then monkey patching or rewriting to always use HTTP.Proxy.

Related issues:

- Related to Ruby master - Bug #4388: open-uriで環境変数http_proxyを使うときに認証付きのProxyが使...
  Rejected
- Related to Ruby master - Feature #8771: Start does not use proxy configuratio...
  Closed
- Has duplicate Ruby master - Feature #12921: Retrieve user and password for pr...
  Closed

Associated revisions

Revision 9416becd - 07/20/2012 01:56 AM - drbrain (Eric Hodel)

- lib/net/http.rb: Net::HTTP now automatically detects and uses proxies from the environment. A proxy may also be specified as before.

Net::HTTP::Proxy still creates anonymous classes, but these classes are only used to store configuration information. When an HTTP instance is created the configuration is now copied.

Additionally, Net::HTTP::ProxyDelta is no longer used by Net::HTTP

[Feature #6546]

- lib/open-uri.rb: Moved URI::Generic#find_proxy to uri/generic.
- lib/uri/generic.rb: Imported find_proxy from open-uri.
- test/open-uri/test_open-uri.rb: Moved proxy-discovery tests to URI.
- test/net/http/test_http.rb: Added tests for proxy behavior.

git-svn-id: svn+ssh://ci.ruby-lang.org/ruby/trunk@36476 b2dd03c8-39d4-4d8f-98ff-823fe69b080e

Revision 36476 - 07/20/2012 01:56 AM - drbrain (Eric Hodel)

- lib/net/http.rb: Net::HTTP now automatically detects and uses proxies from the environment. A proxy may also be specified as before.

Net::HTTP::Proxy still creates anonymous classes, but these classes are only used to store configuration information. When an HTTP instance is created the configuration is now copied.

Additionally, Net::HTTP::ProxyDelta is no longer used by Net::HTTP

[Feature #6546]

- lib/open-uri.rb: Moved URI::Generic#find_proxy to uri/generic.
- lib/uri/generic.rb: Imported find_proxy from open-uri.
- test/open-uri/test_open-uri.rb: Moved proxy-discovery tests to URI.
- test/net/http/test_http.rb: Added tests for proxy behavior.

Revision 36476 - 07/20/2012 01:56 AM - drbrain (Eric Hodel)

- lib/net/http.rb: Net::HTTP now automatically detects and uses proxies from the environment. A proxy may also be specified as before.
Net::HTTP::Proxy still creates anonymous classes, but these classes are only used to store configuration information. When an HTTP instance is created the configuration is now copied.

Additionally, Net::HTTP::ProxyDelta is no longer used by Net::HTTP

[Feature #6546]

- lib/open-uri.rb: Moved URI::Generic#find_proxy to uri/generic.
- lib/uri/generic.rb: Imported find_proxy from open-uri.
- test/open-uri/test_open-uri.rb: Moved proxy-discovery tests to URI.
- test/net/http/test_http.rb: Added tests for proxy behavior.

Revision 36476 - 07/20/2012 01:56 AM - drbrain (Eric Hodel)

- lib/net/http.rb: Net::HTTP now automatically detects and uses proxies from the environment. A proxy may also be specified as before.

Net::HTTP::Proxy still creates anonymous classes, but these classes are only used to store configuration information. When an HTTP instance is created the configuration is now copied.

Additionally, Net::HTTP::ProxyDelta is no longer used by Net::HTTP

[Feature #6546]

- lib/open-uri.rb: Moved URI::Generic#find_proxy to uri/generic.
- lib/uri/generic.rb: Imported find_proxy from open-uri.
- test/open-uri/test_open-uri.rb: Moved proxy-discovery tests to URI.
- test/net/http/test_http.rb: Added tests for proxy behavior.

Revision 36476 - 07/20/2012 01:56 AM - drbrain (Eric Hodel)

- lib/net/http.rb: Net::HTTP now automatically detects and uses proxies from the environment. A proxy may also be specified as before.

Net::HTTP::Proxy still creates anonymous classes, but these classes are only used to store configuration information. When an HTTP instance is created the configuration is now copied.

Additionally, Net::HTTP::ProxyDelta is no longer used by Net::HTTP

[Feature #6546]

- lib/open-uri.rb: Moved URI::Generic#find_proxy to uri/generic.
- lib/uri/generic.rb: Imported find_proxy from open-uri.
- test/open-uri/test_open-uri.rb: Moved proxy-discovery tests to URI.
- test/net/http/test_http.rb: Added tests for proxy behavior.

Revision 36476 - 07/20/2012 01:56 AM - drbrain (Eric Hodel)

- lib/net/http.rb: Net::HTTP now automatically detects and uses proxies from the environment. A proxy may also be specified as before.

Net::HTTP::Proxy still creates anonymous classes, but these classes are only used to store configuration information. When an HTTP instance is created the configuration is now copied.

Additionally, Net::HTTP::ProxyDelta is no longer used by Net::HTTP

[Feature #6546]

- lib/open-uri.rb: Moved URI::Generic#find_proxy to uri/generic.
- lib/uri/generic.rb: Imported find_proxy from open-uri.
- test/open-uri/test_open-uri.rb: Moved proxy-discovery tests to URI.
- test/net/http/test_http.rb: Added tests for proxy behavior.

Revision 36476 - 07/20/2012 01:56 AM - drbrain (Eric Hodel)

- lib/net/http.rb: Net::HTTP now automatically detects and uses proxies from the environment. A proxy may also be specified as before.
Net::HTTP::Proxy still creates anonymous classes, but these classes are only used to store configuration information. When an HTTP instance is created the configuration is now copied.

Additionally, Net::HTTP::ProxyDelta is no longer used by Net::HTTP

[Feature #6546]
- lib/open-uri.rb: Moved URI::Generic#find_proxy to uri/generic.
- lib/uri/generic.rb: Imported find_proxy from open-uri.
- test/open-uri/test_open-uri.rb: Moved proxy-discovery tests to URI.
- test/net/http/test_http.rb: Added tests for proxy behavior.

Revision 36477 - 07/20/2012 02:05 AM - drbrain (Eric Hodel)
- NEWS: Updated net/http for automatic proxy detection (#6546) and automatic gzip and deflate compression (#6492, #6494).

Revision 36477 - 07/20/2012 02:05 AM - drbrain (Eric Hodel)
- NEWS: Updated net/http for automatic proxy detection (#6546) and automatic gzip and deflate compression (#6492, #6494).

Revision 36477 - 07/20/2012 02:05 AM - drbrain (Eric Hodel)
- NEWS: Updated net/http for automatic proxy detection (#6546) and automatic gzip and deflate compression (#6492, #6494).

Revision 36477 - 07/20/2012 02:05 AM - drbrain (Eric Hodel)
- NEWS: Updated net/http for automatic proxy detection (#6546) and automatic gzip and deflate compression (#6492, #6494).

Revision 36477 - 07/20/2012 02:05 AM - drbrain (Eric Hodel)
- NEWS: Updated net/http for automatic proxy detection (#6546) and automatic gzip and deflate compression (#6492, #6494).

Revision 36477 - 07/20/2012 02:05 AM - drbrain (Eric Hodel)
- NEWS: Updated net/http for automatic proxy detection (#6546) and automatic gzip and deflate compression (#6492, #6494).

History
#1 - 06/07/2012 03:51 AM - naruse (Yui NARUSE)
- Status changed from Open to Feedback

I agree the concept, a patch is welcome.

#2 - 06/09/2012 10:00 AM - drbrain (Eric Hodel)
- File net.http.proxy_from_env.patch added
- Category set to lib
- Assignee set to naruse (Yui NARUSE)
- Target version set to 2.0.0

Here is a patch, it is based on code from net-http-persistent.
It tries both http_proxy and HTTP_PROXY to support windows users (who sometimes use the lowercase version).

#3 - 06/09/2012 10:02 AM - drbrain (Eric Hodel)
I did not add it to the patch, but if this patch is accepted should Net::HTTP.new be changed so that p_addr = :ENV by default?
This would allow users to use their proxy by default.

#4 - 06/12/2012 02:34 AM - naruse (Yui NARUSE)
drbrain (Eric Hodel) wrote:
I did not add it to the patch, but if this patch is accepted should Net::HTTP.new be changed so that p_addr = :ENV by default?

This would allow users to use their proxy by default.

The patch still needs a fix to use HTTP_PROXY, so it should be. But setting password through environment variable is not acceptable, because they are visible from other users on some OSs as #4398 says (in Japanese).

#5 - 06/12/2012 09:17 AM - drbrain (Eric Hodel)
- File net.http.proxy_from_env.2.patch added

OK.

This updated patch uses :ENV by default in Net::HTTP.new and Net::HTTP::Proxy.

This updated patch does not use HTTP_PROXY_USER, etc.

#6 - 06/12/2012 10:42 AM - naruse (Yui NARUSE)
- Status changed from Feedback to Assigned
- Assignee changed from naruse (Yui NARUSE) to drbrain (Eric Hodel)

Drbrain (Eric Hodel) wrote:

OK.

This updated patch uses :ENV by default in Net::HTTP.new and Net::HTTP::Proxy.

This updated patch does not use HTTP_PROXY_USER, etc.

OK, please commit it.

#7 - 06/12/2012 10:52 AM - akr (Akira Tanaka)

Do you consider that HTTP_PROXY environment variable can be set from outside in CGI by Proxy: header?

In that case, HTTP_PROXY should be ignored.

Also, platforms which environment variables are case insensitive (Windows), http_proxy is also dangerous.

#8 - 06/13/2012 06:36 AM - drbrain (Eric Hodel)
- File net.http.proxy_from_env.2.no_env_by_default.patch added

Here is a patch which does not use HTTP_PROXY by default (most of the changes are to documentation).

Please tell me which patch I should apply.

#9 - 06/13/2012 08:44 AM - akr (Akira Tanaka)

net.http.proxy_from_env.2.no_env_by_default.patch seems have following line.

env_proxy = ENV['http_proxy'] || ENV['HTTP_PROXY']

It may refer the environment variable set by Proxy: header sent from client in CGI.

I think it is vulnerable, as I said in [ruby-core:45579].

As far as I know, this issue is found for libwww-perl in 2001.

http://cpansearch.perl.org/src/GAAS/libwww-perl-6.04/Changes:
| 2001-03-14  Gisle Aas  gisle@ActiveState.com  |
|    Release 5.51 |
| SECURITY FIX: If LWP::UserAgent::env_proxy is called in a CGI |
| environment, the case-insensitivity when looking for "http_proxy" |
| permits "HTTP_PROXY" to be found, but this can be trivially set by the |
| web client using the "Proxy:" header. The fix applied is that |
$ENV{HTTP_PROXY} is not longer honored for CGI scripts.
The CGI_HTTP_PROXY environment variable can be used instead.
Problem reported by Randal L. Schwartz.

#10 - 06/13/2012 08:48 AM - nahi (Hiroshi Nakamura)
drbrain (Eric Hodel) lib/open-uri.rb has battle-tested proxy configuration logic in it. If we really want to add proxy configuration to net/http, we'd better start from the logic.

#11 - 06/28/2012 09:42 AM - drbrain (Eric Hodel)
- File net.http.proxy_from_env.3.patch added

=begin
Here is an updated patch that uses the battle-tested code from open-uri. This patch reads the proxy from (((ENV))) by default like open-uri. It is safe in a CGI environment just like open-uri.

Details:
URI::Generic#find_proxy was moved from lib/open-uri.rb to lib/uri/generic.rb.
Proxy configuration is now stored in the Net::HTTP instance, not the class. This allows (((no_proxy))) to work at connect time.
For compatibility, Net::HTTP.Proxy still returns an anonymous subclass of Net::HTTP, but this subclass only stores proxy configuration information. When new is called the proxy configuration is copied from the class to the instance.
Removed use of Net::HTTP::ProxyDelta. The module still exists for use with lib/net/http/backward.rb.
Net::HTTP.new no longer calls Net::HTTP::Proxy.
Net::HTTP.newobj has been removed.
Net::HTTP#connect no longer calls #conn_address or #conn_port. Instead it calls the new method #proxy?. If the proxy information is being determined from (((ENV))) and the current connection is listed in (((no_proxy))) a direct connection is made, if it is not listed a proxy connection is made. If a manual proxy has been configured a proxy connection is made. If no proxy is configured a direct connection is made.
#proxy_address and #proxy_port may return nil if called when #proxy? returns false. Since these methods were previously called only via ProxyDelta and #conn_address and #conn_port I don't think this should be a problem.
#conn_address and #conn_port always return #address and #port now. Since these methods are private I don't think this should be a problem.
#edit_path has been updated to behave properly based on #proxy? and #use_ssl?
I have not altered open-uri to remove it's built-in proxy handling.
I ran the tests with (((http_proxy))) pointing to a valid and invalid proxy with no problem.
=end

#12 - 06/29/2012 04:56 AM - drbrain (Eric Hodel)
- Assignee changed from drbrain (Eric Hodel) to naruse (Yui NARUSE)

#13 - 07/19/2012 03:16 PM - naruse (Yui NARUSE)
OK, commit it.

#14 - 07/20/2012 10:56 AM - drbrain (Eric Hodel)
- Status changed from Assigned to Closed
- % Done changed from 0 to 100

This issue was solved with changeset r36476.
Jacob, thank you for reporting this issue.
Your contribution to Ruby is greatly appreciated.
May Ruby be with you.

- lib/net/http.rb: Net::HTTP now automatically detects and uses proxies from the environment. A proxy may also be specified as before.
Net::HTTP::Proxy still creates anonymous classes, but these classes are only used to store configuration information. When an HTTP instance is created the configuration is now copied.
Additionally, Net::HTTP::ProxyDelta is no longer used by Net::HTTP
- Moved URI::Generic#find_proxy to uri/generic.
- Imported find_proxy from open-uri.
- Moved proxy-discovery tests to URI.
- Imported proxy-discovery tests from open-uri.
- Added tests for proxy behavior.

**Files**

<table>
<thead>
<tr>
<th>File Name</th>
<th>Size</th>
<th>Date</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>net.http.proxy_from_env.patch</td>
<td>5.95 KB</td>
<td>06/09/2012</td>
<td>drbrain (Eric Hodel)</td>
</tr>
<tr>
<td>net.http.proxy_from_env.2.patch</td>
<td>8.01 KB</td>
<td>06/12/2012</td>
<td>drbrain (Eric Hodel)</td>
</tr>
<tr>
<td>net.http.proxy_from_env.2.no_env_default.patch</td>
<td>7.42 KB</td>
<td>06/13/2012</td>
<td>drbrain (Eric Hodel)</td>
</tr>
<tr>
<td>net.http.proxy_from_env.3.patch</td>
<td>23.6 KB</td>
<td>06/28/2012</td>
<td>drbrain (Eric Hodel)</td>
</tr>
</tbody>
</table>