Privoxy tutorial

by Chen Sanon 2014-09-24

theme: #Privoxy

contents

The Privoxy version involved in this article is 3.0.26, which was released on 2016-08-27.

Privoxy is a web proxy software:

> Privoxy is a non-caching web proxy with advanced filtering capabilities for enhancing privacy, modifying web page data and HTTP headers, controlling access, and removing ads and other obnoxious Internet junk.

Simply put, it's the gatekeeper of traffic in and out of your computer. With Privoxy, we can control the outgoing request and rewrite the returned response. Unnecessary requests - such as address address, image ads, video ads, we can directly `block` out; unnecessary response content - such as text ads on the page, we can borrow from the `filter` off filter, let them be displayed on the browser page .

Install Privoxy

Privoxy supports many platforms :

> At present, Privoxy is known to run on Windows 95 and later versions (98, ME, 2000, XP, Vista, Windows 7 etc.), GNU/Linux (RedHat, SuSE, Debian, Fedora, Gentoo, Slackware and others), Mac OS X (10.4 and upwards on PPC and Intel processors), OS/2, Haiku, DragonFly, ElectroBSD, FreeBSD, NetBSD, OpenBSD, Solaris, and various other flavors of Unix.

For each platform, Privoxy provides installation instructions .

A few platforms are briefly mentioned below.

1. Windows platform

   The installation under the Windows platform is usually visualized. Download the installation package, double-click, and then follow the prompts to install by clicking all the way.

2. Linux platform

   It can be installed through the warehouse most of the time.

   For example, Ubuntu:

```
sudo apt-get install privoxy
```

```
sudo apt-get install privoxy
```

Another example is openSUSE:

```
sudo zypper install privoxy
```

3. mac platform

   If you have a mac install homebrew, you can perform `brew install privoxy` to install Privoxy.

If you really like toss, just download the source code and compile and install it yourself.

Start Privoxy

After installing Privoxy, you need to start it, because the conditions of each system under each platform are different, so I won't introduce them one by one here. Please see the instructions for starting Privoxy .

How do I know whether Privoxy has started successfully? Please see the next step.

Configure browser

Privoxy after startup, it runs by default `localhost:8118` on. So we only need to point the proxy of the browser to it.

After configuring the browser, open the http://pp URL in the browser to see if the following content is displayed:

> This is Privoxy 3.0.26 on localhost (127.0.0.1), port 8118, enabled

If so, it means that Privoxy has started successfully and the browser is configured correctly-at this time, all traffic from the browser will pass through the Privoxy proxy.

Set up Privoxy

After starting Privoxy and configuring the browser, you can start customizing our Privoxy.

Profiles from the core `config` to start.

`config` The file location and name may be different under various systems. For example, under Windows system, it is actually called `config.txt` . Under openSUSE system, the directory where it is located is `/etc/privoxy` , this directory is a soft link and points to it `/var/lib/privoxy/etc` .

However, under normal circumstances, we do not need to modify `config` the file.

Usually, we modify the other two types of files:

1. action file
    1. match-all.action
    2. default.action
    3. user.action
2. filter file
    1. default.filter

1. default.filter
2. user.filter

`match-all.action`, `default.action`, `default.filter` These documents, it is recommended not to modify, because Privoxy upgrades will be overwritten. We need self-configuration definition content written `user.action` or `user.filter`.

action file

As the name suggests, action file defines Privoxy action, for example, we `user.action` add a file `{+block}`:

```
{+block{禁止访问陈三的博客}}
.zfanw.com
```

Analyze:

1. `{+block}` Is an action, `block` after the `{}` written comments, write from time to write.
2. `.zfanw.com` One effect is the object of the operation, is divided into two portions, a host, a path, host portion support portion wildcard, for example `*`, , `?`, `[0-9]`; `[a-z]` first path portion means a `/` portion of the URL, support POSIX 1003.2 regular expressions, than The host part is flexible. See the actions document for details .

In this way, Privoxy put my website stopped by: Any `zfanw.com` request will return 403, the content probably as follows:

> The proxy server is refusing connections

So, having said so much, where can the `user.action` files be modified? Do I need to restart Privoxy after modification?

After configuring Privoxy's browser access `http://config.privoxy.org/show-status` , you can see the path of all configuration files. By default, we can only view the configuration, but in fact, the editing function can be enabled . Usually I use Vim to edit directly. After editing and saving, you can see the effect without restarting Privoxy.

filter file

The filter file defines rules for filtering **responses** , such as:

```
FILTER: replaceText 替换文本
s|陈三|陈四|g
```

The first row, uppercase `FILTER` represents a filter rule is defined, `replaceText` represents the name of the rule, and then followed by the comment.

The second line is to modify the returned page. For example you used, etc. Vi / Vim or sed tool, should be `s` very familiar with this substitution command. Simply put, the above statement is to replace the code in the page-the principle of Privoxy to remove ads is here.

However, `user.filter` the only defined filtering rules, the application of the rules, or to the action file, so the `user.action` file, add the following configuration:

```
{+filter{replaceText}}
.zfanw.com
```

I guess you already know the meaning of the above configuration: `filter` a directive requiring execution `user.filter` defined `replaceText`, `.zfanw.com` the definition of `filter` the role of the object - that URL.

But you may find that during the test, the above `FILTER` rule does not take effect.

No, no, this is certainly not the author's spoof. This is because my website has https already enabled, and Privoxy can do very little on https-encrypted pages -the only exception is the previous one `{+block}`. That's because the domain name part of the https request is still exposed to Privoxy, so there is a way to intercept it.

This means that under the trend that https becomes more and more popular in the future, the demand for Privoxy will become smaller and smaller. Take Baidu as an example. Now it has enabled https across the board. Privoxy, which was able to block ads on Baidu pages, is now powerless.

Proxy forwarding

The current mainstream browsers only support http proxy, yes, they don't even support https proxy, let alone any socks proxy.

Now suppose you have started a socks proxy locally, running it `127.0.0.1:9999`, and want the browser to use it. How to do?

One way is to use Privoxy forwarding .

Privoxy open the `config` file, add the following:

```
forward-socks5 / 127.0.0.1:9999 .
```

In this way, all traffic passing through the Privoxy proxy will be forwarded to the local socks5 proxy.

Of course, you can also split traffic according to the domain name. For example, Google's website needs to be forwarded to the sock5 proxy. Domestic websites don't need it. Just forward Google's domain name:

```
forward-socks5 .google.com 127.0.0.1:9999 .
forward-socks5 google.com 127.0.0.1:9999 .
```

Edit this page

see more

- Privoxy proxy sharing
- Privoxy cannot start
- Privoxy blocks ads

New article
Srcset and sizes

Old one
Github best practices