

MassSiteProxy. Простой на функциональный скрипт массового копирования сайтов. Внимание!!! Это не законно, скрипт использовать только в ознакомительных целях. Автор не несет ответственности от любых ущербов и/или абуз владельцев оригинальных сайтов.

Внимание!!! ~10-15% сайтов скрипт не сможет полностью кешировать. Это связано с большим количеством нюансов. Скрипт разрабатывался для больших объемов, некоторыми мелочами пришлось пренебречь.

Внимание!!! Скрипт не копирует сайт, который лежит за Cloudflare. Да и не один не копирует.

Системные требования: (на 90% хостингов будет работать)

- PHP 5.6, для максимальной производительности PHP 7.1(Файлы под нужную версию PHP лежат в корне скрипта, переименуйте в index.php)
- Apache, для NginX или других серверов, может понадобится дополнительная настройка <https://www.slimframework.com/docs/v3/start/web-servers.html>
- CURL с включенным curl_exec
- xml
- md_string
- iconv, не обязательно, можно отключить в config.php
- sqlite3, не обязательно, можно отключить в config.php, сменив тип кеширования на файлов или другой
- **ioncube_loader** - обязательно

Установка:

- Скопировать на сервер в корень домена
- в **config.php** в секцию **userid** прописать свою почту или логин для идентификации лицензии на скрипт,
- **Для версий MassSiteProxy <1.14**, в файл **proxy.txt** добавить список сайтов для клонирования. **Скрипт возьмет случайный**
- Для версий **MassSiteProxy > 1.14** сайты добавляюся **в файл proxy.csv в формате:**
<http://proxy.domain;https://ifirestarter.ru>
или
<http://proxy.domain;https://ifirestarter.ru;ru;en>
Отныне с версии 1.14 скрипт НЕ!!! бурет случайный сайт а только

соответствие акцептор:донор

- Открыть сайт в браузере

Подбор доноров:

Тут все довольно просто: **не берите доров из топ 10 ПС**. Как правило, особенно, в конкурентных нишах, за клон топовых доноров очень быстро последуют санкции как от РКН так и от ПС.

Не берите один донора, берите много и тестируйте их, чтобы проверить, хорошо ли и правильно скрипт копирует их.

Не берите большие сайты, есть вероятность что закончится место на диске, прежде чем вы получите трафик.

Не берите сайты за Cloudflare.com. Как бы вы не хотели, скрипт их не сможет скопировать. Да и не один не копирует.

Настройка:

Все настройки очень детально прокомментированы в файле **config.php**. Изначально ничего менять не нужно)) Перед запуском скрипта внимательно его изучите.

CURL полная эмуляция запроса браузера

Добавлено в MassSiteProxy > 1.15

Для того чтобы использовать, в вашем браузере (Opera, Chrome и т.д):

Перейдите на сайт донора, в консоли вебмастера, на вкладке **network**, в таблице на URL донора правой кнопкой **“Copy->Copy as cURL (cmd)”**

Вставить скопированную информацию между скобками

'curl' => ",

Перевод доноров(изначально выключен):

- в файл **proxy.txt** добавить список сайтов в формате: **http://donor.ru|ru|en**, где ru - язык с которого переводить, en- язык на который переводить.
- в **config.php** в настройки translate **isEnable** установите true
- **yandexApi** - добавьте Api ключ, можно взять тут <https://tech.yandex.ru/translate/doc/dg/concepts/api-keys-docpage/>
- для работы через несколько ключей, укажите массив из них
- **Внимание: для версии 1.14 файл переименован в proxy.csv и формат для переводчика:**
- **http://proxy.domain;https://ifirestarter.ru;ru;en**

Работа через Прокси сервера(изначально выключен):

Настройки работы через прокси, лежат в config.php в секции proxies.

isEnabled может быть false(выключены) или true(включить)

proxyType может быть **CURLPROXY_HTTP** , **CURLPROXY SOCKS4** , **CURLPROXY SOCKS5**, **CURLPROXY SOCKS4A** , **CURLPROXY SOCKS5_HOSTNAME**

proxyPath пример строки socks5://127.0.0.1:9150, для работы через авторизацию - socks5://127.0.0.1:9150;proxyUser:proxyPass

Очистка контента от рекламы или поиск и замена информации на доноре:

- Изначально скрипт сам чистит контент от рекламы, счетчиков и т.д. Но правила можно добавлять или править через файл **replace.php**. Также, в **replace.php**, можно использовать любые PHP функций над текстом страницы.
- Еще, можно чистить рекламу через простые **НАЙТИ|ЗАМЕНИТЬ** правила в файле **replace.txt**(Один файл на всю пачку.)
- в версии **MassSiteProxy > 1.15** можно **менять информацию на доноре с помощью js**. **НАЙТИ|ЗАМЕНИТЬ** правила в файле **replace.js.txt** (Один файл на всю пачку.)

Добавление своего js css и кодов счетчиков:

- чтобы подключить свой js код донору, положите его в файл **head.txt** скрипт автоматом его добавит в раздел head сайта.
- чтобы добавить код счетчиков, например LI, просто положите его код в файл **counters.txt**, скрипт автоматом его добавит в конец body сайта

Работа с формами на сайте:

Для отлова заполненных форм, в **config.php** в параметр **catchForms** установите **isEnabled** и впишите адрес электронной почты, чтобы скрипт передавал формы на нее, так же скрипт дублирует заполненные формы в файл **forms.txt**

Слив трафика через ТДС:

- для слива трафика через встроенную ТДС, установите в config.php в секцию **isRedirect true**, а в **redirectUrl** урл вида <http://redirect.url/redirect.php?query={keyword}> чтобы перенаправить трафик
- для слива трафика через **Ztds**, положите в корень сайта файл **api.php**, который предоставляет ТДСка, и настраивайте его по мануалу к **Ztds**.

Поддержка контента в актуальном состоянии:

в config.php в секцию **lifeTime** установите время в секундах для того чтобы скрипт обновлял контент через этот промежуток. Рекомендуемое значение 0(не обновлять), но если очень нужно, то 24*60*60 что соответствует обновлению раз за день.

Синонимизация контента:

В **config.php** в секцию **synonymizer**, установите **isEnabled = true**. Файл с базой для синонимайзинга лежит **synonym.txt** в корне скрипта.

Перелинковка сетки сайтов:

В **config.php** в секцию **linking(количество ссылок для переинковки)**, установите значение **> 0** а в файл **links.txt** впишите свои ссылки в формате **тут анкор** с новой строки и скрипт будет добавлять на страницы сайта внешние ссылки.

Управление Кешированием сайта:

- **Для полного отключения**, установите настройки **cache type** в config.php параметр **array_cache**
- Для отключения кеша отдельных страниц установите регулярные выражения для идентификации страниц в **noCachePatterns**
- **Для Сжатия места на диске**, установите **compressCache** параметр **true**

Подтверждение сайтов в панелях VM

Все сайты автоматом проходят подтверждения в панелях вебмастера Гугла и Яндекса, выбрав тип подтверждения Файл.

Быстроиндекс для Яндекса:

Добавить в VM -> Отправить на переобход)))

Быстроиндекс для Гугла:

Трастовые дропы и/или Хрумер, через пару недель после создания)))

Если что-то пошло не так(пустая страница):

в **config.php**, в секцию, **debugMode** установите true и анализируйте логи в папке */data/*