

Руководство по настройке контент-фильтра **DANSGUARDIAN.** **Часть 1.**

Вступление.

С этой части начинается разбор по кирпичикам настройки контент-фильтра Dansguardian.

Предполагается, что читатель имеет базовые навыки работы с консолью Linux и любым консольным текстовым редактором.

Так же потребуется предварительно настроенный прокси-сервер SQUID, достаточно его минимальной рабочей конфигурации.

Здесь не рассматриваются вопросы установки и удаления Dansguardian, так как его дистрибутив есть во многих репозиториях, и устанавливается он обычными для конкретного дистрибутива способами.

Что касается самостоятельной сборки, то здесь должен быть некоторый опыт в сборке программ, иначе корректно собрать не получится.

Я не ставил своей целью сделать перевод всех опций конфигурационных файлов, здесь приведены наиболее важные и критичные опции.

Многие из неописанных мной опций могут остаться со значениями по умолчанию, поскольку их использование требуется очень редко в исключительных случаях.

Куда более важно на мой взгляд, научиться работать с черными и белыми списками, файлами фраз и им подобным, ведь корректная работа фильтра будет зависеть в первую очередь от них.

Тема создания и настройки групп будет затронута в отдельной части.

Не забываем перезапускать Dansguardian после редактирования конфигурационных файлов.

Где находятся конфигурационные файлы Dansguardian.

Путь к конфигурационным файлам задается при сборке Dansguardian, однако может быть использован другой для этого используется запуск Dansguardian с опцией **-c** .

Наиболее вероятные месторасположение конфигурационных файлов ***/etc/dansguardian/*** или ***/usr/local/etc/dansguardian/*** .

Если не нашли их, то запуск dansguardian с опцией **-v** выведет опции сборки , эта строка в выводе `—sysconfdir=/etc` говорит о том, что конфигурационные файлы должны находится в директории `/etc` , поддиректории dansguardian.

В общем, когда мы знаем, где конфигурационные файлы можно приступить к их рассмотрению.

Некоторые особенности формата файлов Dansguardian.

Все файлы конфигурации Dansguardian имеют обычный текстовый формат.

Символ решетки **#** в начале строки означает, что эта строка является комментарием и не используется.

Однако из этого правила есть исключения, то есть символ решетки в начале строки может использоваться для задания специфичных опций. О таких случаях и этих опциях будет рассказано отдельно.

Первый конфигурационный файл *dansguardian.conf*

Настройку мы начнем с этого файла. Открываем его в текстовом редакторе. Да, опций много, но с основными сейчас разберемся. Зеленым отмечены рекомендуемые значения, синими которые возможно надо изменить под ваш сервер.

reportinglevel = 3

Эта опция задает вид веб-страницы запрета (не влияет на запись в лог) Ее возможные значения.

-1 = Только записывать в лог, не блокировать. (Скрытый режим).

0 = Вывести только "Доступ запрещен".

1 = Вывод страницы с запретом без показа найденных запрещенных фраз.

2 = Полный отчет (Рекомендуется на время проверки и настройки фильтра)

3 = Использовать HTML-шаблон страницы запрета (Рекомендуется использовать по умолчанию)

language_dir = '/usr/share/dansguardian/languages'

Расположение директории с HTML-шаблонами для разных языков.

HTML-шаблон используется, когда предыдущая опция имеет параметр 3.

language = 'russian-koi8-r'

HTML-шаблон из директории с шаблонами.

HTML-шаблон написан на PERL, и может быть модифицирован под конкретные задачи.

loglevel = 2

Настройка ведения лога

Возможные значения:

0 = не записывать в лог

1 = записывать только запрещенные

2 = записывать все текстовые запросы (по умолчанию)

3 = записывать всё

logfileformat = 1

Формат лога.

Возможные значения:

1 = оригинальный DansGuardian format (разделитель пробел)

2 = формат файлов CSV с разделителем ;

3 = формат логов прокси-сервера Squid

4 = Tab delimited (разделитель TAB)

Собственно рекомендую использовать оригинальный формат, так как иначе не будут работать специализированные анализаторы логов.

anonymizelogs = off

Не записывать в лог имена пользователей и их IP-адреса.

Возможные значения:

on = включить

off = выключить (записывать в лог имена пользователей и их IP-адреса)

loglocation = '/var/log/dansguardian/access.log'

Расположение файла лога

filterip = 192.168.0.1

IP-адрес интерфейса на котором Dansguardian принимает соединения от клиентов.

Пустое значение прослушивать все сетевые интерфейсы.

Если необходимо задать два сетевых интерфейса просто добавляем еще одну строку, например так

filterip = 192.168.0.1

filterip = 192.168.5.1

filterport = 8080

Порт на котором Dansguardian принимает соединения от клиентов.

proxyip = 127.0.0.1

IP-адрес прокси-сервера, по умолчанию локальный интерфейс, т.е. 127.0.0.1

proxyport = 3128

порт прокси сервера, принимающий соединения, по умолчанию 3128

accessdeniedaddress = 'http://192.168.0.1/cgi-bin/dans.pl'

адрес скрипта, выводящего страницу запрета при опции ***reportinglevel*** равной 1 или 2.

Здесь маленькое отступление.

Скрипт должен располагаться на локальном веб-сервере с поддержкой PERL скриптов. Доступ к этому веб-серверу, желательно должен быть прямой, этого можно достичь, если использовать настройку браузера, **не использовать прокси-сервер для..**, куда ввести адрес вашего веб-сервера.

Есть и другой путь адрес вашего веб-сервера надо добавить в exceptionsitelist.

showweightedfound = on

Запись в лог найденных фраз и вывод в странице запрета при reportinglevel = 2

Возможные значения:

on = включить

off = выключить

weightedphrasemode = 2

Значение этой опции будет рассмотрено далее.

Ниже приведенные опции должны быть такими как приведены, для корректной поддержки кириллицы (комбинация определена опытным путем), хотя все равно придется пойти на хитрость, которая будет описана позже.

phrasefiltermode = 2

preservecase = 2

hexdecodecontent = on

forcequicksearch = on

Вот собственно и все с настройкой этого конфигурационного файла, может некоторые опции остались за кадром, но основное рассмотрено.