
Content Filtering with DansGuardian

Created May 14, 2005
by Bruce A. Westbrook

Last Revision: June 12, 2005 - BAW

Introduction

This document describes the step-by-step process of installing and configuring DansGuardian content filtering on a SUSE Professional v9.2 box or Red Hat Fedora Core 3, using Squid as a proxy server on the same box.

This will work by building a Linux box that will run the three necessary services: Squid Proxy, DansGuardian and Apache. Apache is needed simply to serve up a couple of graphics

Installing / Configuring Linux

√	Description
Install SUSE Linux	<p>You can use whatever version of Linux you want, but this document will assume either SUSE Professional v9.2 or Red Hat Fedora Core 3.</p> <ol style="list-style-type: none">1. Boot with a SUSE Pro v9.2 boot CD<ol style="list-style-type: none">a. You can create the CD from an iso image located at: ftp://ftp.suse.com/pub/suse/i386/current/iso/SUSE-Linux-9.2-mini-installation.iso2. At the boot menu screen use the arrow keys to select Installation and then type: <code>install=ftp://140.221.37.133/pub/suse/i386/9.2</code><ol style="list-style-type: none">a. Other ftp sites can be found at: http://www.novell.com/products/linuxprofessional/downloads/ftp/int_mirrors.html (be sure to use the IP address though, not the name)3. Press ENTER to begin initializing the installation4. After a few minutes (it takes a minute to load the CD image into RAM disk – you can press ESC and watch it load) you will be at the “Welcome to YaST2” screen.<ol style="list-style-type: none">a. <u>Interesting installation information</u> – hit [CTRL]+[ALT]+[F5] to open a root prompt. You can then use the [ALT]+[F]unction keys to look at things like:<ol style="list-style-type: none">i. [ALT] + [F1] = Boot screenii. [ALT] + [F2] = root promptiii. [ALT] + [F3] = hardware device infoiv. [ALT] + [F4] = kernel messagesv. [ALT] + [F5] = root promptvi. [ALT] + [F6] = root promptvii. [ALT] + [F7] = switches you back to GUIviii. [ALT] + [F8] = more hardware infoix. [ALT] + [F9] = root promptx. [ALT] + [F10] = IP information5. Select your language (English-US) and click Accept6. On the Installation Settings screen, verify that the Software section does NOT have an error. If it states that

		<p>“Cannot read package data from Installation media...” good luck resolving that ☺</p> <ol style="list-style-type: none">7. Now click the Partitioning section8. SUSE will have automatically determined the partitioning. You can accept it the way it is, which is probably just two partitions – a swap that equals your RAM and a root with the rest of the drive space. A smarter way is to create a separate var partition. This will allow you to still boot the box if the log files ever fill up the partition. Do this as follows:<ol style="list-style-type: none">a. Select Create custom partition setup and click Nextb. Select Custom partitioning - for experts and click Nextc. Now delete all the partitions on the drive (if this is server hardware there is probably a utility partition. You don't want to delete the utility partition)d. Create partitions as follows:<ol style="list-style-type: none">i. Click Create.ii. Select Primary partition and click OKiii. Under Format, File System, choose Swap.iv. Calculate the amount of RAM in the system times 1.5 (if the machine has 1GB of RAM, you'll want a swap partition that is approx. 1.5GB)v. Under Size, End, enter your calculation in the following syntax: +#GB, where # is 1.5 times your RAM (for instance, for a swap partition of 1.5GB, enter +1.5GB)vi. Click OKvii. Now Create another Primary partitionviii. Under Format, File System, leave the default (Reiser)ix. Under Size, End, enter +4GB for the partition size.x. Under Mount Point select /var and click OKxi. Now Create another Primary partitionxii. Under Format, File System, leave the default (Reiser)xiii. Under Size, End, you can leave the defaults which will use the rest of the drive for this partition.xiv. Under Mount Point select / and click OKxv. Click Finish to return to the Installation Settings screen9. Now let's configure our software:<ol style="list-style-type: none">a. Click the Software sectionb. Click the Detailed Selection buttonc. Under Simple Webserver with Apache2, select the following:<ol style="list-style-type: none">i. apache2ii. apache2-dociii. apache2-mod_php4iv. apache2-preforkv. gd
--	--	--

		<ul style="list-style-type: none">d. Under Network/Server, select the following:<ul style="list-style-type: none">i. etherealii. perl-libwww-perliii. squide. Under Experienced User, select the following:<ul style="list-style-type: none">i. findutils-locateii. MozillaFirefoxiii. nmapiv. xpdff. Click Acceptg. You will be prompted with an Automatic Changes screen that details packages changes that were made to resolve conflicts. Click Continue <p>10. Back at the Installation Settings screen let's configure the time</p> <ul style="list-style-type: none">a. Click the Time Zone sectionb. Configure the time zone appropriatelyc. Change the date and/or time if necessaryd. Under the Hardware clock set to pulldown, select local timee. Click Accept <p>11. At the Installation Settings screen we need to configure SUSE to startup in non-GUI mode</p> <ul style="list-style-type: none">a. Click the Default Runlevel sectionb. At the pulldown select 3: Full multiuser with network and click OK <p>12. At the Installation Settings screen go ahead and click Accept and let's get this install moving! Click through the license agreement and then confirm the installation. The installer will begin formatting, downloading files from the FTP server and installing. We're looking at 30 minutes plus depending on the machine speed and Internet connectivity.</p>
--	--	--

	<p>Post-SuSE Linux Install Configuration</p>	<p>Once the installation is complete the system will reboot automatically. You can remove the CD at this point, but if you miss the opportunity don't worry – the system will boot from your newly installed SUSE linux installation after the CD boot doesn't receive any input.</p> <ol style="list-style-type: none"> 1. After the reboot, proceed through any popups regarding hardware detection 2. Enter the appropriate Root Password – click Next 3. Make the following changes on the Network Configuration screen: <ol style="list-style-type: none"> a. Click Firewall <ol style="list-style-type: none"> i. Select Enable Firewall – click Next ii. The external interface should be your one and only configured interface and internal interface should be blank – click Next iii. For Services, select: <ol style="list-style-type: none"> 1. HTTP 2. HTTPS 3. SSH iv. Click Expert v. Under TCP, type in 8080 – click OK vi. Click Next vii. For Features, only Allow Traceroute should be selected – click Next viii. For Logging, only the Standard options should be selected – click Finish b. Click Network Interfaces <ol style="list-style-type: none"> i. For your Already Configured Devices, click Change ii. Select your network card and click Edit iii. Set the static IP address and mask iv. Click the Host name and name server button– set the host name, domain and name servers (DNS) – click OK v. Click the Routing button – set the default gateway – click OK vi. Click Next, then Finish c. Click Next 4. On the Test Internet connection screen, select Yes and click Next <ol style="list-style-type: none"> a. After the result comes back as Success, click Next b. At the Online Updates Available popup, select Yes and click OK c. Accept the defaults on the YaST Online Update (YOU) screen and click Next d. After a moment you will be presented with the update screen – note the color scheme of patches: <ol style="list-style-type: none"> i. Red = Security ii. Blue = Recommended iii. Black = Optional e. Accept the defaults and click Accept f. If presented with a Kernel patch popup, click Skip Patch – we'll patch it after the installation is completed g. Sit back and wait for updates to be downloaded and installed
--	---	---

		<ul style="list-style-type: none">h. Once completed, you can click the Remove Sources Packages after Update to save disk space – Click Finishi. The process will then write system configuration data <ol style="list-style-type: none">5. At the User Authentication Method screen, set to Local – click Next6. At the Add a New Local User screen, add the Console account with the appropriate password and deselect Auto Login – click Next7. At the Release Notes screen read the entire screen and take notes as there will be a test at the end – click Next8. At the Hardware Configuration screen, make the following changes:<ol style="list-style-type: none">a. ?????? – need to complete this section, my test box “blackedscreened” and I had to reboot, which brought it straight into the GUI login.
--	--	---

	Install Red Hat Fedora Core 3	<p>Ok, so you want to use Red Hat Fedora instead (good idea!). Here we go then...</p> <ol style="list-style-type: none"> 1. Boot with CD 1 of Fedora <ol style="list-style-type: none"> a. You can download the Fedora ISO images from: http://fedora.redhat.com/download/ 2. You can skip the CD-ROM test 3. Welcome to Fedora – click Next 4. Select your language – click Next 5. Select your keyboard – click Next 6. For Installation type select Custom – click Next 7. For Disk Partitioning select Manually Partition with Disk Druid – click Next 8. Setup partitions as follows: <ol style="list-style-type: none"> a. Select your hard-drive (typically <code>/dev/hda</code>) and click Delete – this will delete all partitions on the drive. If this is server hardware, you'll want to delete any partitions individually and leave the utility partition. b. Now click the New button c. Calculate the amount of RAM in the system times 1.5 (if the machine has 1024MB of RAM, you'll want a swap partition that is approx. 1536MB) d. For File System Type select Swap e. For Size (MB) enter the size of the swap partition (RAM times 1.5) in megabytes f. Click OK g. Click New h. For Mount Point, pulldown and select <code>/var</code> i. For File System Type leave as ext3 j. For Size (MB) enter 4096 k. Click OK l. Click New m. For Mount Point, pulldown and select <code>/</code> n. For File System Type leave as ext3 o. For Size (MB) enter 4096 p. Click the checkbox for Force to be a primary partition q. Click OK r. Click Next 9. For Boot Loader Configuration click Next 10. For the Network Devices screen, set your static IP, your FQDD hostname, gateway and DNS servers. 11. Click Next 12. For Firewall Configuration, select to Enable the firewall and then allow the following services: <ol style="list-style-type: none"> a. Remote Login (SSH) b. Web Server (HTTP, HTTPS) 13. Leave the Enable SELinux as Active 14. Click Next 15. For Additional Languages – click Next 16. For Time Zone, set your time zone. Do <u>not</u> enable the system clock to use UTC – click Next 17. Set your root password and click Next 18. Now make the following software changes at the Package Group Selection screen:
--	--------------------------------------	--

		<ul style="list-style-type: none">a. Under Applications select the following:<ul style="list-style-type: none">i. Editors (defaults)ii. Office / Productivity (defaults)<ul style="list-style-type: none">1. xpdfb. Under Servers select the following:<ul style="list-style-type: none">i. Server Configuration Tools (defaults)ii. Web Server (defaults – includes squid)c. Under Development select the following:<ul style="list-style-type: none">i. Development Tools (defaults)d. Under System select the following:<ul style="list-style-type: none">i. Administration Tools (defaults)ii. System Tools (defaults)<ul style="list-style-type: none">1. ethereal-gnome2. nmap-frontendiii. Deselect Printing Support <p>19. Click Next, Next and Continue to begin loading the system</p> <p>20. When the installation is complete, you will be prompted as such. Click the Reboot button</p>
--	--	---

<p>Post-Red Hat Linux Install Configuration</p>	<p>After the installation and the initial reboot you will be walked through a post-installation wizard.</p> <ol style="list-style-type: none"> 1. At the welcome screen – Next 2. Accept the license agreement – Next 3. Set the date/time – Next 4. Set your display as appropriate – Next 5. Create a user account – Next 6. Test your audio device – Next 7. For Additional CDs, just click Next 8. Finish – Next <p>Now let's login as root and update the system using yum.</p> <ol style="list-style-type: none"> 1. First, we need to replace the configuration file with a current one for the list of mirror servers: <pre>wget http://www.fedorafaq.org/samples/yum.conf</pre> 2. Copy this file to <code>/etc/yum.conf</code> (overwrite your existing file) 3. Install GPG keys for package checking (taken from http://www.fedorafaq.org/#gpgsig) <p>Red Hat: <pre>rpm --import /usr/share/doc/fedora-release-3/RPM-GPG-KEY*</pre> </p> <p>Extras: <pre>rpm --import http://download.fedora.redhat.com/pub/fedora/linux/extras/RPM-GPG-KEY-Fedora-Extras</pre> </p> <p>rpm.livonia.org: <pre>rpm --import http://rpm.livna.org/RPM-LIVNA-GPG-KEY</pre> </p> <p>FreshRPMS: <pre>rpm --import http://freshrpms.net/packages/RPM-GPG-KEY.txt</pre> </p> <p>DAG: <pre>rpm --import http://dag.wieers.com/packages/RPM-GPG-KEY.dag.txt</pre> </p> <p>ATrpms: <pre>rpm --import http://atrpms.net/RPM-GPG-KEY.atrpms</pre> </p> <p>NewRPMS: <pre>rpm --import http://newrpms.sunsite.dk/gpg-pubkey-newrpms.txt</pre> </p> <p>Dries: <pre>rpm --import http://apt.sw.be/dries/RPM-GPG-KEY.dries.txt</pre> </p> <p>JPackage: <pre>rpm --import http://www.jpackage.org/jpackage.asc</pre> </p>
--	--

kde-redhat:

```
rpm --import http://kde-  
redhat.sourceforge.net/gpg-pubkey-ff6382fa-  
3e1ab2ca
```

4. Now update your packages by running:
`yum -y update`
5. You can also run yum in various ways:
 - a. To see a list of what's available:
`yum list available`
 - b. To install a software package:
`yum install packagename`
 - c. To update a software package:
`yum update packagename`
 - d. To see what updates are available:
`yum check-update`
 - e. To search for a package:
`yum search`
6. When yum is completed it will tell you.

Once you've updated the system, let's configure Red Hat to boot up into text mode, not GUI. No reason to boot into the GUI by default on this server. To do this, simply `vi /etc/inittab` and change the line
`id:5:initdefault`
to
`id:3:initdefault`

Finally, let's open port 8080 in the firewall so the proxy will work.

1. **Application > System Settings > Security Level**
2. In the **Other Ports** dialog box, type `8080:tcp`
3. Click **OK**

Now reboot the system which will probably boot into a new kernel that was installed with yum and you'll end up at a text prompt. Yeah!

Content Filtering Installation & Configuration

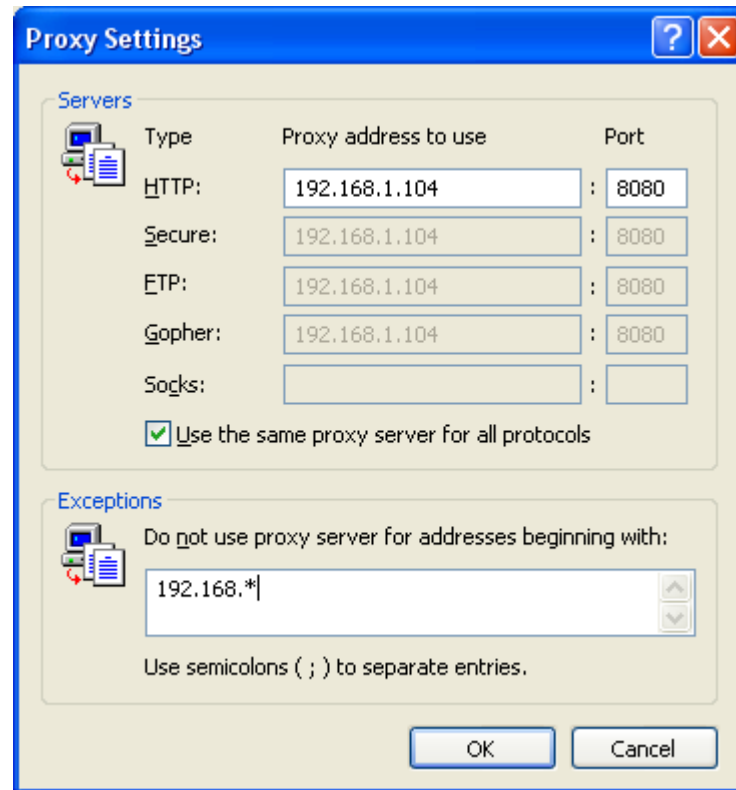
√	Description
Configure Squid	<p>To configure squid to listen on the network we need to add three simple lines to the configuration.</p> <ol style="list-style-type: none">1. <code>vi /etc/squid/squid.conf</code>2. Locate the line <code>http_access allow our_networks</code>3. Add the following two lines immediately after the above line, substituting your own network schema as appropriate: <code>acl localnet src 192.168.1.0/255.255.255.0</code> <code>http_access allow localnet</code>4. Save and exit the file5. Set squid to start automagically at boot<ol style="list-style-type: none">a. <code>chkconfig squid on</code>b. <code>chkconfig --list squid</code>6. Start Squid<ol style="list-style-type: none">a. <code>service squid start</code>
DansGuardian	<p>Download the DansGuardian .rpm: <code>wget http://usmirror.dansguardian.org/downloads/2/Stable/FedoraCore3/dansguardian-2.8.0.4-1.fc3.i386.rpm</code></p> <p>Other download mirrors can be found at: http://dansguardian.org/?page=download</p> <ol style="list-style-type: none">1. Install DansGuardian <code>rpm dansguardian-2.8.0.4-1.fc3.i386.rpm -Uvh</code>2. Configure DansGuardian as follows:<ol style="list-style-type: none">a. <code>vi /etc/dansguardian/dansguardian.conf</code>b. Set <code>reportinglevel = 3</code> (use HTML template file)c. Set <code>loglevel = 2</code> (all requests)d. Set <code>logfileformat = 3</code> (Squid Log File format)e. Save and exit the file3. Blacklists – If you want to use blacklists in addition to content filtering, perform the following. Note that this initial download of the blacklist is free. After your first download you must pay to continue to update it. But even if you don't update the list it's a great list to add in.<ol style="list-style-type: none">a. <code>wget "http://urlblacklist.com/cgi-bin/commercialdownload.pl?type=download&file=bigblacklist" -O blacklist.tar.gz</code> – download a DansGuardian compatible blacklists (I used the one from http://urlblacklist.com/ but it is a paid service to keep updated)b. <code>tar -zxvf blacklist.tar.gz -C /etc/dansguardian</code> – unzip the file to <code>/etc/dansguardian</code> so it will create the blacklist directory and populate it.c. Now you need to turn on the blacklist functionality by configuring the various banned lists as you desire. To do this you edit the

		<p><code>/etc/dansguardian/bannedurllist</code> and the <code>/etc/dansguardian/bannedsitelist</code>. Uncomment the files you want to use. Note that if you configure to use a file that doesn't exist DansGuardian will fail to start.</p> <p>4. Customize the denied page – the default page that users see when they are denied access is located in <code>/etc/dansguardian/language/ukenglish/template.html</code>. You can leave this as is or customize it to your desire. If you want to use my customized version you can download the files from: <code>wget http://www.thewestbrooks.com/downloads/dansguardian/template.tar.gz</code>. Then replace the original <code>template.html</code> with the customized version and place the <code>stop.gif</code> file in <code>/var/www/html/</code> to be served up. Finally you will also need to edit the <code>template.html</code> file to set your web server's IP address for the location of the <code>stop.gif</code> file, your company/organization name and your IT contact. Look for these three keywords to replace:</p> <ol style="list-style-type: none"> <code>SERVER_IP</code> <code>COMPANY_NAME_HERE</code> <code>IT_CONTACT_HERE</code> <p>5. Unfilter the webserver – let's add the server IP address to an exception list that will not filter the server when it comes to sarg. <code>vi /etc/dansguardian/exceptioniplist</code> and then add the IP address of the server to the list.</p> <p>6. Now set DansGuardian and Apache to startup automatically at boot:</p> <ol style="list-style-type: none"> <code>chkconfig dansguardian on</code> <code>chkconfig --list dansguardian</code> <code>chkconfig httpd on</code> <code>chkconfig --list httpd</code> <p>7. Start the services</p> <ol style="list-style-type: none"> <code>service httpd start</code> <code>service dansguardian start</code>
	<p>Block ALL Sites Except...</p>	<p>If the client wants to be able to block ALL Internet sites with the exception of specific sites, this can be done by enabling what's called Blanket Block. To do this edit the <code>/etc/dansguardian/bannedsitelist</code> and uncomment both <code>**</code> and <code>*ip</code>.</p> <p>Then edit the <code>/etc/dansguardian/exceptionsitelist</code> or <code>/etc/dansguardian/greysitelist</code> (depends on whether you want content still filtered or not) and type in the sites the client wants access to.</p>

Clients

Now to force the clients to go through DansGuardian for content filtering you simply set all of the client's browser's to proxy through the IP address of DansGuardian using port 8080 (by default, DansGuardian uses port 8080 while Squid uses port 3128).

NOTE: If you turn on the **Blanket Block** or **Blanket IP Block** in the bannedsitelist file, be sure to add an **Exception** for the local network, or at least the content filtering server that's running Apache so that the stop.gif will show up in the blocked page template.



To mitigate someone circumventing the proxy, setup the Internet border firewall or NAT device to ONLY allow traffic outbound from the Squid proxy IP address. To further define this you can configure Squid to only allow access from the DansGuardian IP address. This way outbound traffic must go first to DansGuardian, be passed to Squid and then go through the firewall.

If you don't take these measures a user can simply turn off proxying on the browser. Or, if you only setup the firewall to allow outbound traffic from Squid, a smart user might try to change their proxy port to 3128 which would allow them to bypass DansGuardian and only go through the proxy which would let them out to the Internet.

	<p>Enforce Access through DansGuardian</p>	<p>What good is a content filter if users can simply bypass it? Not much! So be sure to setup your firewall with egress rules that allow web browsing ONLY from the proxy box – and maybe your administrator box. ☺</p> <p>For a pix this means adding an access-level that has port 80 allowed from the IP address of the proxy, NOT from “any” – for instance:</p> <pre>access-level outbound permit tcp host 192.168.0.1 255.255.255.255 any eq 80</pre>
	<p>Log Analyzer</p>	<p>If you want to see where your users are going and what sites are being denied, the SARG log analyzer works great with DansGuardian when its set to log as a Squid logfile.</p> <ol style="list-style-type: none"> 1. Download SARG RPM: <pre>wget http://mack.ro/linux/sarg/sarg-2.0.7-1.fc3.mack.i386.rpm</pre> <p>(the main SARG page is http://sarg.sourceforge.net/sarg.php)</p> 2. Install the rpm – <pre>rpm sarg-2.0.7-1.fc3.mack.i386.rpm -Uvh</pre> 3. Edit the sarg configuration file – <pre>vi /etc/sarg/sarg.conf</pre> – locate these sections and adjust as follows: <ol style="list-style-type: none"> a. Set <code>access_log</code> <pre>/var/log/dansguardian/access.log</pre> (DansGuardian's log file) b. Set <code>title</code> “DansGuardian Web Filtering Reports” (or whatever you prefer the title to be) c. Set <code>denied_report_limit</code> 0 (to report on ALL denied sites) 4. Edit the apache sarg configuration file – <pre>vi /etc/httpd/conf.d/sarg.conf</pre> – to allow any remote workstations that you want to be able to access the sarg logs. <p><u>Examples:</u></p> <ol style="list-style-type: none"> a. <code>Allow from all</code> – will allow access from all IP addresses b. <code>Allow from xxx.xxx.xxx.xxx</code> – will allow only the specific IP address 5. Restart the apache service – <pre>service httpd restart</pre> 6. Run sarg for the first time – <pre>sarg</pre> 7. Now bring up the logs in a browser – http://SERVER/sarg 8. Tada!

	Secure Log Analyzer Website	<p>You don't want everyone being able to see the log results? Ok, let's lock it down with a username/password.</p> <ol style="list-style-type: none">1. Create a .htaccess file to control access <code>vi /var/www/sarg/.htaccess</code>2. Input the following information into the file <code>AuthType Basic</code> <code>AuthName "Sarg"</code> <code>AuthUserFile /var/www/.htpasswd</code> <code>require valid-user</code>3. Create the user/password for the site by issuing the following command. If you want more then one user, do not use the "-c" (which creates the file) after the first user. <code>/usr/bin/htpasswd -c /var/www/.htpasswd sarg</code> > New password: <i>(enter a password to use)</i>4. Now we need to configure Apache to allow use of the .htaccess file: <code>vi /etc/httpd/conf/httpd.conf</code>5. Find the line <Directory />. Right after this is the line <code>AllowOverride None</code>. Change the None to All, as follows: <code>AllowOverride All</code>6. Save and exit the file7. Restart Apache <code>service httpd restart</code>8. Test the site with your new password9. Tada!
--	------------------------------------	---

	<p>Log Rotation</p>	<p>If you think you want to keep all your logfiles, then the first choice is the best. It is a script that will zip and date-stamp each day's logfile for you.</p> <p>If you only want to keep a certain number of logfiles, say the most recent 31 days worth, then use the second choice. It is a script that ships with DansGuardian and will number each logfile for X number of days, based on your settings.</p> <p>First Choice – Keep 'em All</p> <ol style="list-style-type: none"> 1. <code>vi /etc/dansguardian/logrotate.sh</code> – then input the following script: <pre>#!/bin/sh # DansGuardian logrotation script for version 2.8.x # Concept by Don Vosburg # Modified by NexGen Systems, Inc. to use # more specific date and time stamp ## Stop the DG service /etc/rc.d/init.d/dansguardian stop ## If the access.log file exists... if [-f /var/log/dansguardian/access.log]; then ## ...zip it... gzip /var/log/dansguardian/access.log ## ...check if the archive directory exists and if ## note, create it... [-d /var/log/dansguardian/archive <code><line wrap></code> mkdir /var/log/dansguardian/archive ## ...then move the zipped up access.log file to the ## archive directory and rename it with yesterday's ## date (the day of the file contents...) mv /var/log/dansguardian/access.log.gz <code><line wrap></code> /var/log/dansguardian/arvhive/\$(date <code><line wrap></code> --date "1 day ago" +%Y-%m-%d).gz fi ## Pause for a few seconds before restarting DG sleep 10 ## Start the DG service again /etc/rc.d/init.d/dansguardian start</pre> 2. <code>crontab -e</code> – set a cron job to run at 5am in order to zip the data AFTER Sarg runs it's daily cronjob @ 4:02am every day (default time in RedHat): <pre>00 05 * * * /etc/dansguardian/logrotate.sh</pre> 3. OR if you prefer you can edit the time that ALL daily cron jobs run and then run the logrotate file after your new time – I leave the specifics up to you, but to change the default times you need to edit the <code>/etc/crontab</code> file. <p>Second Choice – Just Keep Some</p> <ol style="list-style-type: none"> 1. <code>vi /etc/dansguardian/logrotation</code> – Verify the <code>LOG DIR</code> is correct, then set the <code>NUM LOGS</code> to however many
--	----------------------------	---

		<p>you'd like to keep</p> <p>2. <code>crontab -e</code> – set a cron job to run at 5am in order to zip the data AFTER Sarg runs it's daily cronjob @ 4:02am every day (default time in RedHat):</p> <pre>00 05 * * * /etc/dansguardian/logrotation</pre>
--	--	--

	Filter Tweaking	<p>There are several files affiliated with how the content filtering works. The next section shows all the files and many of the configuration options.</p> <p>The biggest issue will probably be unblocking sites or parts of sites, or blocking other sites that are making it past the filtering. The files that are concerned with these issues are all in <code>/etc/dansguardian</code>.</p> <p>There are two types of lists that will bypass filtering; exception lists and grey lists. Both will bypass the URL filtering. The difference is the exception lists will bypass all filtering while the grey lists will still perform content filtering.</p> <p>Within these two types there are two sub-types; site lists and url lists. Site lists refer to an entire site (e.g. <code>playboy.com</code>) while url lists refer to allowing specific portions of a site while still filtering other parts (e.g. <code>playboy.com/quiz/health</code>).</p> <ul style="list-style-type: none"> • <code>exceptionsitelist</code> - Unblock ALL of a site • <code>exceptionurllist</code> - Unblock PART of a site • <code>greysitelist</code> - Unblock ALL of a site but still filter on content • <code>greyurllist</code> - Unblock PART of a site but still filter on content <p>Here's some typical examples of how to use these files:</p> <p>Q1. My user is blocked from a site that should be allowed. A1. Edit the <code>/etc/dansguardian/exceptionsitelist</code> and add the domain</p> <p>Q2. My user is blocked from a page that they need to see but I still want the rest of the domain blocked A2. Edit the <code>/etc/dansguardian/exceptionurllist</code> and add the URL</p> <p>Q3. There is a site that I want unblocked as a banned URL but still filtered for content. A3. Edit the <code>/etc/dansguardian/greysitelist</code> and add the doman</p> <p>Q4. There is a page(s) I want my users to see but I don't want them to access any other parts of that site. A4. Edit the <code>/etc/dansguardian/greyurllist</code> and add the URL</p>
--	------------------------	---

DansGuardian Configuration Information:

The following files make up the overall configuration of DansGuardian:

<i>exceptioniplist</i>	This file contains a list of client IP addresses that you wish to allow unrestricted access (no filtering).
<i>exceptionphraselist</i>	This file contains a list of phrases that, if they appear in a web page, will bypass filtering. You may want to use the <i>weightedphraselist</i> instead, as this can result in a lot of pages not being blocked.
<i>exceptionsitelist</i>	This file contains a list of domain endings that if found in the requested URL, will not be filtered.
<i>exceptionurllist</i>	This file contains a list of URL parts for sites where filtering should be turned off.
<i>exceptionuserlist</i>	This file contains a list of usernames that will not be filtered (you must use basic authentication or ident must be enabled for this to work).
<i>bannedextensionlist</i>	This file contains a list of file extensions that will be banned. This can be used to restrict users from downloading screen savers, executable files, viruses, and so forth.
<i>bannediplist</i>	This file contains a list of client IP addresses that will not get web access at all.
<i>bannedmimetyplist</i>	This file contains a list of MIME-types that will be banned. If a URL request returns a MIME-type in this list, DansGuardian will block it. This can be used to block movies, but shouldn't be used to graphic image files or text/html, etc.
<i>bannedphraselist</i>	This file contains a list of phrases that will result in banning a page. Each phrase must be enclosed between < and > characters, and they may contain spaces. You can also use a combination of phrases that, if all are found in a page, will result in it being blocked.
<i>bannedregexpurllist</i>	This file contains a list of regular expression URLs that will be banned. that will be banned. This can be used to restrict users from downloading screen savers, executable files, viruses, and so forth.
<i>bannedsitelist</i>	This file contains a list of sites that are to be banned. You can use IP addresses here as well as domain names, and can even include stock SquidGuard blacklists as well.
<i>bannedurllist</i>	This file contains a list of URL parts to block, which allows you to block parts of a site rather than the entire site. You can also use SquidGuard lists here as well.
<i>banneduserlist</i>	This file contains a list of usernames to whom, if basic proxy authentication is enabled, access will be denied automatically.
<i>weightedphraselist</i>	This file contains a list of phrases with a corresponding positive or negative value. As phrases are encountered in a page, the total "value" of the page will be calculated based on these values; good phrases will have negative values and bad phrases will have positive values. One the Naughtyness Limit has been reached (defined in <i>dansguardian.conf</i> , the page will be blocked.
<i>pics</i>	This file contains a number of PICS sections that allow you to fine-tune your PICS filtering. The defaults for DansGuardian are for young children (mild profanity, artistic nudity, etc.).

Each of these configuration files are very straightforward and are basically one item per line (ie. a URL or IP address, etc.).

The [dansguardian.conf](#) file is the primary configuration file for DansGuardian. It is here that you will configure things like logging, where to redirect users when attempting to access a denied page, and so forth. The file is heavily commented and fairly straightforward.

This is a fairly standard configuration one might have; you can even use it verbatim provided you change IP addresses and port settings to match your own system.

The [reportinglevel](#) setting tells DansGuardian to fully report why access was denied (ie. give the denied phrase). You may choose to use a level of [1](#) instead, or [3](#) to use the HTML template file. If you use the HTML template file, then the [htmltemplate](#) file needs to be set to the full path and filename of the template file you wish to use. If you use a setting of [0](#) through [2](#) you will need to set the [accessdeniedaddress](#) keyword. In this case, it's pointing to the internal IP address of our firewall (10.0.5.1), and the port it listens to (in this case port 8444). It also contains the full path to the [dansguardian.pl](#) CGI script.

The `loglevel` keyword is to determine what gets logged to the `/var/log/dansguardian/access.log` logfile.

The `filterip` determines what IP address that DansGuardian will listen on. If left blank, all IPs will be listened on. The `filterport` keyword is the port that DansGuardian will bind to. The `proxyip` is the IP address of the proxy; usually the localhost. The `proxyport` is the port to use to connect to the proxy (in this case, 3128, which is the port that Squid is listening on).

The keywords following are all related to the various configuration files discussed earlier, and simply include them to be a part of the configuration.

The `weightedphrasemode` determines how weighted phrases are used. A setting of `1` is for normal operation. The `naughtynesslimit` keyword sets the limit over which a page will be blocked. This is based on the values of the `weightedphraselist` file and each "hit" on a page will modify the naughtyness of the page. The higher the rating, the "naughtier" the page. As a general rule of thumb, with the default settings, a limit of 50 is suitable for young children, 100 for older children, and 160 for young adults.

The `showweightedfound` keyword determines whether the phrases found that made up the total that exceeds the naughtyness limit will be logged, and if `reportinglevel` is set to `2`, reported.

The `reverseaddresslookups` keyword determines whether or not DansGuardian will look up the forward DNS for an IP URL address and search for both the banned site and URL lists. This is useful for preventing a user from simply entering the IP address for a banned site. It can also have an impact on the searching speed, however.

The `createlistcachefiles` keyword determines whether or not the `bannedsitelist` and `bannedurl` files will be cached. Fast computers do not need this, but on slower computers this could result in a significant process start speed increase.

The `maxuploadsize` keyword is used for POST protection on web upload forms. A setting of `-1` disables, a setting of `0` blocks completely, and any other value sets the file upload size in kilobytes (after MIME encoding and headers).

The `forward_for` keyword, if enabled, will add an X-Forwarded-For to the HTTP request header. This may be required for some sites that need to know the source IP.

The `maxchildren` keyword sets the maximum number of processes to spawn to handle incoming connections. This can be used to prevent DoS attacks from killing the server by maxing out spawned processes.

The `log_connection_handling_errors` keyword is used to determine if DansGuardian will log debug info to syslog.

Extracted from: <http://linsec.ca/bin/view/Main/DansGuardian>