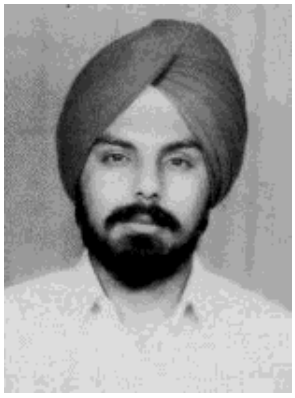
LinuxFocus article number 235
http://linuxfocus.org

Een Squid-Proxy Server opzetten



door D.S. Oberoi
<ds_oberoi/at/yahoo.com>

Over de auteur:
D.S. Oberoi woont in Jammu, India en heeft op dit moment wat problemen met zijn internet verbinding door de politieke inrust in het gebied.

Vertaald naar het Nederlands door:
Hendrik-Jan Heins
<hjh/at/passys.nl>

Kort:

Linux is synoniem voor Netwerken. Het wordt zowel in kantoren al thuis gebruikt als bestands-, print- e-mail- en applicatie server daarnaast wordt het ook steeds vaker gebruikt als proxy server.

Een proxy-server biedt internet toegang aan meerdere gebruikers tegelijk door het delen van een enkele internet verbinding. Een goede proxy server biedt ook de mogelijkheid om aanvragen te cachen, hierdoor wordt de verbinding minder belast doordat bij een tweede aanvraag eerst de lokale data wordt doorgegeven. Squid is een software pakket dat zowel proxy, HTTP caching, ftp, gopher enz. ondersteunt. Bovendien ondersteunt het SSL, toegangs controles, cachen van DNS en onderhoudt het een log bestand waarin alle aanvragen worden bewaard. Squid is ook beschikbaar voor Windows-NT, op Logi Sense.

Dit artikel is gericht op de basisprincipes van het opzetten van een proxy server en de mogelijkheid om gebruikers een gecontroleerde webtoegang te bieden.

Is Squid geïnstalleerd ?

Squid's rpm pakket wordt meegeleverd bij RedHat 7.1 en wordt automatisch geïnstalleerd als u kiest voor de "Network OS Installation" optie. U kunt met het volgende commando controleren of het wel geïnstalleerd is:

```
rpm -q squid
```

De nieuwste versie van Squid kn altijd worden gevonden op de Squid Homepage en andere mirror sites. Squid kan met het volgende rpm commando worden geïnstalleerd op een systeem:

```
rpm -ivh squid-2.3.STABLE4-10.i386.rpm
```

Squid Configureren

De werking en het gedrag van Squid worden aangegeven door de instellingen in het configuratie bestand, hier is dat "squid.conf". Dit bestand kan meestal worden gevonden in de directory /etc/squid. Het configuratie bestand van Squid is een nogal lang verhaal, het is vele pagina's lang, maar het positieve hieraan is wel dat alle opties duidelijk worden aangegeven en verklaard.

Het eerste dat moet worden ingesteld is de http_port, deze geeft aan op welk socket adres Squid aanvragen van client systemen moet aannemen; standaard staat dit op 3128, maar het mag door de gebruiker in een andere waarde worden veranderd. Naast een poort waarde kan ook het IP adres van de machine waar Squid op draait worden aangegeven:

http_port 192.168.0.1:8080

Met de bovenstaande invulling, zit Squid op IP adres 192.168.0.1 en op poort adres 8080. Hier kan ieder ander poort adres worden opgegeven, maar zorg er wel voor dat er geen andere applicatie op dat poortadres zit. Met ongeveer glijkwaardige configuratie regels kunnen ook de poorten voor andere services worden opgegeven.

Toegangscontrole

Door middel van toegangscontrole mechanismen, kan de toegang tot internet worden gecontroleerd, vooral met betrekking tot tijdsduur, tijdstip, cachen, toegang tot populaire websites of website groepen, enz. De Squid toegangscontrole bestaat uit twee onderdelen, ACL elementen en een toegangslijst. Een toegangslijst maakt toegang tot de service mogelijk of juist onmogelijk.

Enkele belangrijke soorten ACL elementen worden hieronder genoemd

- src : Source (bron), ofwel het IP adres van clients
- dst : Destination(doel), ofwel de IP adressen van de betreffende server(s)
- srcdomain : Source, ofwel de domeinnaam van de client
- dstdomain : Destination, ofwel de domeinnaam van de server
- time : Tijd, dag, week, datum...
- url_regex : URL standaard uitdrukkingspatroon koppeling
- urlpath_regex: URL-path standaard uitdrukkingspatroon koppeling, zonder protocol en naam van de gastheer
- proxy_auth : Gebruikers authorisatie met behulp van externe processen
- maxconn : Maximum aantal verbindingen vanaf een enkel client IP adres

Om deze controles te activeren, moet eerst een ACL worden gedefinieerd, en daarna kan deze worden toegepast. Het format voor een ACL statement is:

acl acl_element_name type_of_acl_element values_to_acl

Opmerking:

1. acl_element_name kan een door de gebruiker gedefinieerde naam zijn die gegeven is aan een ACL element.
2. Geen twee ACL elementen mogen dezelfde naam hebben.
3. Iedere ACL bestaat uit een lijst met waarden. Wanneer je op zoek bent naar een overeenkomst, wordt voor meerdere waarden gebruik gemaakt van de "OR"-logica. Met andere woorden: een ACL element wordt gezien als overeenkomstig zodra een van de opgegeven waarden overeenkomt.
4. Niet alle ACL elementen kunnen worden gebruikt in combinatie met alle soorten toegangslijsten.
5. Verschillende ACL elementen worden aangegeven op verschillende regels en Squid combineerd deze en maakt er een lijst van.

Er zijn een aantal verschillende toegangslijsten beschikbaar. Degene die we hier gaan gebruiken, zijn hieronder te vinden:

- **http_access:** Laat HTTP clients een verbinding maken via de HTTP poort. Dit is de primaire toegangscontrole lijst.
- **no_cache:** Definieert het cachen van antwoorden op aanvragen.

Een toegangslijst regel bestaat uit "keywords" zoals "allow" of "deny" die toegang tot een bepaald ACL element voor de service of een aantal services regelt.
Opmerking:

1. De regels worden gecontroleerd in de volgorde waarop ze zijn opgegeven en de controle stopt zodra er een overeenkomst is gevonden.
2. Een toegangslijst kan bestaan uit meerdere regels.
3. Zodra geen van de regels overeenkomt, is de standaard actie het omgekeerde van de laatste regel in de lijst; het is dus verstandig om duidelijk te zijn over de standaard actie.
4. Alle elementen van een toegangsregel zijn samengevoegd met "AND" en worden op de volgende manier uitgevoerd:
   http_access Action statement1 AND statement2 AND statement OR.
   http_access Action statement3
   Multiple http_access statements worden gekoppeld via "OR" terwijl elementen van een toegangsregel met "AND" worden gekoppeld
5. Denk eraan dat regels altijd van boven naar beneden worden gelezen.

Terug naar de Configuratie

Standaard zal Squid geen toegang geven aan clients, om dit mogelijk te maken moeten de toegangscontrole lijsten worden aangepast. Je moet je eigen regels opstellen voor toegang. Bekijk het bestand squid.conf en voeg de volgende regels toe boven "http_access deny all"

acl mynetwork 192.168.0.1/255.255.255.0
http_access allow mynetwork

mynetwork is de naam van de ACL en de volgende regel is de regel die toe moet worden gepast op een

bepaalde ACL, geheten "mynetwork". 192.168.0.1 is een referentie naar het netwerkadres met het netmask 255.255.255.0. mynetwork is in feite niets meer dan een naam voor een groep machines in het netwerk en de volgende regel geeft de clients toegang. De bovenstaande veranderingen samen met "http_port" is al voldoende om Squid aan het werk te zetten. Na de veranderingen kan Squid met het volgende commando worden gestart:

service squid start

Opmerking:
Squid kan ook automatisch worden gestart bij het opstarten door het aan te zetten in ntsysv of setup (System Service Menu). Na iedere verandering in het configuratie bestand, moet het draaiende Squid proces worden gestopt om de nieuwe instellingen te activeren bij een herstart. Deze twee stappen kunnen met de volgende commando's worden uitgevoerd:

1. service squid restart of
2. /etc/rc.d/init.d/squid restart

Client Machine Configuratie

Aangezien de client aanvraag vanaf een bepaalde poort van de proxy server moet komen, moet de client zo worden ingesteld dat hij een aanvraag doet op de juiste poort. Er wordt nu vanuit gegaan dat de machines al aangesloten zijn op het lokale netwerk (LAN) (en al geldige IP adressen hebben) en de Linux server kunnen pingen.
Voor Internet Explorer

1. Ga naar "Tools" -> "Internet Options"
2. Kies het "Connection" Tab-blad en klik op "LAN Setting"
3. Vink de "Proxy Server" box aan en geef het IP adres van de proxy server en het poortadres op (http_port address).

Voor Netscape Navigator

1. Ga naar "Edit" -> "Preference" -> "Advanced" -> "Proxies".
2. Kies de "Manual Proxy Configuration" knop.
3. Klik op de "View" knop &
4. Geef het IP adres en de poort van de proxy server op (http_port address).

Toegangs Controle Gebruiken

Meervoudige toegangscontrole en regels bieden een zeer goede en flexibele manier om de client toegang tot internet te beheren. Voorbeelden van de meest voorkomende controles worden hieronder gegeven; dit betekent echter niet dat dit de enige mogelijke toegangs controle is.

1. Laat alleen bepaalde machines het internet op

   acl allowed_clients src 192.168.0.10 192.168.0.20 192.168.0.30
   http_access allow allowed_clients
   http_access deny !allowed_clients

   Dit laat alleen machines met de volgende IP adressen op internet: 192.168.0.10, 192.168.0.20 en 192.168.0.30 De andere IP adressen (niet aangegeven) kunnen internet niet op.
2. Beperk de toegang tot bepaalde uren en voor bepaalde tijd

   acl allowed_clients src 192.168.0.1/255.255.255.0
   acl regular_days time MTWHF 10:00-16:00
   http_access allow allowed_clients regular_days
   http_access deny allowed_clients

   Hiermee mogen de clients in netwerk 192.168.0.1 van maandag tot vrijdag van 10:00 uur tot 16:00 uur het internet op.
3. Meervoudige toegang voor verschillende clients

   acl hosts1 src192.168.0.10
   acl hosts2 src 192.168.0.20
   acl hosts3 src 192.168.0.30
   acl morning time 10:00-13:00
   acl lunch time 13:30-14:30
   acl evening time 15:00-18:00
   http_access allow host1 morning
   http_access allow host1 evening
   http_access allow host2 lunch
   http_access allow host3 evening
   http_access deny all

   De bovenstaande regel laat host1 zowel gedurende de ochtend als de avonduren toe; terwijl host2 en host3 alleen tijdens de lunch en ’s avonds toegang hebben.

   Opmerking:
   Alle elementen van een toegangsregel worden met "AND" gekoppeld en op de volgende manier uitgevoerd:

   http_access Action statement1 AND staement2 AND statement OR.

   Meervoudige http_access statements worden met "OR" gekoppeld, terwijl elementen van een toegangs regel met "AND" worden gekoppeld; hierdoor kan

   http_access allow host1 morning evening

   nooit werken als ochtend en avond statement (morning AND evening), aangezien dat nooit waar zou zijn en er daardoor geen actie zou worden ondernomen.
4. Sites Blokkeren

Squid kan de toegang tot een bepaalde site of meerdere sites die een bepaald woord bevatten, verbieden. Dit kan op de volgende manier worden gedaan:

```
acl allowed_clients src 192.168.0.1/255.255.255.0
acl banned_sites url_regex abc.com *()(*.com
http_access deny banned_sites
http_access allow allowed_clients
```

Ditzelfde mechanisme kan ook worden gebruikt om toegang tot sites die een bepaald woord bevatten, bijvoorbeeld "dummy" , "fake"...

```
acl allowed_clients src 192.168.0.1/255.255.255.0
acl banned_sites url_regex dummy fake
http_access deny banned_sites
http_access allow allowed_machines
```

Het is niet handig om alle woorden en site namen op te geven in het configuratie bestand. Dit kan ook in een ander bestand worden ondergebracht (bijvoorbeeld banned.list in de /etc directory) en ACL kan de gegevens uit dit bestand halen en hiermee de toegang tot sites ontzeggen.

```
acl allowed_clients src 192.168.0.1/255.255.255.0
acl banned_sites url_regex "/etc/banned.list"
http_access deny banned_sites
http_access allow allowed_clients
```

5. Om het gebruik te optimaliseren
   kan Squid het aantal verbindingen dat client machines mogen maken, beperken met behulp van het "maxconn" element. Om gebruik te kunnen maken van deze optie, moet "client_db feature" eerst worden ingeschakeld.

   ```
   acl mynetwork 192.168.0.1/255.255.255.0
   acl numconn maxconn 5
   http_access deny mynetwork numconn
   ```

   Opmerking:
   maxconn ACL maakt gebruik van een "minder dan"-vergelijking. Deze ACL wordt bereikt zodra het aantal verbindingen groter is dan de aangegeven waarde. Dit is de belangrijkste reden waarom deze ACL niet wordt gebruikt in samenwerking met de "http_access allow" regel.
6. De gegevens Cachen
   Antwoord op de aanvragen wordt meteen gecached, dit is handig voor de statistische gegevens. Het is niet nodig om cgi-bin of Servlet te cachen en dit kan worden voorkomen door gebruik te maken van het "no_cache" ACL element.

   ```
   acl cache_prevent1 url_regex cgi-bin /?
   acl cache_prevent2 url_regex Servlet
   no_cache deny cache_prevent1
   no_cache deny cache_prevent2
   ```

7. Maak je eigen Foutmeldingen
   Het is mogelijk om je eigen foutmeldingen te maken met behulp van een "deny"-regel, het gaat hier om de "deny_info" optie. Standaard staan alle foutmeldingen van Squid in de directory /etc/squid/errors. De foutmeldings directory kan worden ingesteld met behulp van de "error_directory" optie. Je kunt zelfs de bestaande foutmeldingen aanpassen.

   acl allowed_clients src 192.168.0.1/255.255.255.0
   acl banned_sites url_regex abc.com *()(*.com
   http_access deny banned_sites
   deny_info ERR_BANNED_SITE banned_sites
   http_access allow allowed_clients

   In het bovenstaande voorbeeld wordt er een speciaal bericht weergegeven wanneer een gebruiker probeert een website te bereiken die niet getoond mag worden. De bestandsnaam achter de optie "ERR_BANNED_SITE" moet wel bestaan in de bovenstaande foutmeldingsdirectory. Het foutmeldings bestand moet opgemaakt zijn in HTML format. De hierboven genoemde voorbeelden zijn slechts enkele van de mogelijkheden, en capaciteiten van ACL. Je kunt meer hierover vinden in de FAQ sectie op de Squid Home Pagina. Er is hier ook meer uitleg over breder gebruik te vinden en een uitleg van andere ACL elementen en toegangselementen.

Log Bestanden

Alle log bestanden van Squid staan in de directory /var/log/squid; hierin staan de cache log, toegangs (access) logs en het store.log. Het bestand access.log bevat informatie over de aanvragen van clients, activiteit en het beheert de opgave voor iedere HTTP & ICP query die ontvangen wordt door de proxy server, het IP adres van de clients, de aanvraag methode, het aangevraagde URL, enz... De gegevens in dit bestand kunnen worden gebruikt om de toegangs aanvragen te analyseren. Er zijn veel programma’s als sarg, calamaris en Squid-Log-Analyzer beschikbaar die de gegevens kunnen analyseren en rapporten kunnen genereren (in HTML format). Er kunnen rapporten worden gegenereerd over gebruikers, IP adressen, bezochte sites enz...

De locatie van deze bestanden kan ook worden veranderd met behulp van de volgende opties:

```
cache_access_log      For access.log
cache_log             For cache.log
cache_store_log       For store.log (Store manager)
pid_filename          Squid process ID file name
```

Autorisatie Methoden

Squid kan met de standaard instelling gebruikers toelaten zonder authorisatie. Om gebruikers te kunnen authoriseren (en wel of niet toe te laten), om gebruik te maken van internet, biedt Squid de mogelijkheid van een authorisatie, maar dan wel via een extern programma, waar een gebruikersnaam en wachtwoord

voor vereist is. Hiervoor moet er gebruik worden gemaakt van de "proxy_auth" ACL en van "authenticate_program"; dit zorgt ervoor dat een gebruiker zich moet authoriseren met gebruikersnaam en wachtwoord, voordat hij gebruik kan maken van de verbinding. Er zijn verscheidene authorisatie programma's beschikbaar die werken met Squid, dit zijn:

1. LDAP : Gebruikt het Linux Lightweight Directory Access Protocol
2. NCSA : Gebruikt een NCSA-achtig gebruikersnaam en wachtwoordbestand
3. SMB : Gebruikt een SMB server zoals SAMBA of Windows NT
4. MSNT : Gebruikt het Windows NT authentication domain
5. PAM : Gebruikt Linux Pluggable Authentication Modules
6. getpwam : Gebruikt het Linux wachtwoord bestand.

Je moet aangeven welk authorisatie programma je wilt gebruiken en dat kan via de optie "authenticate_program". Zorg ervoor dat het het authorisatie programma dat je hier wilt gebruiken al is geïnstalleerd en werkt.

De veranderingen in het squid.conf bestand moeten ook terug te vinden zijn in "authenticate_program", /usr/local/bin/pam_auth.

```
acl pass proxy_auth REQUIRED
acl mynetwork src 192.168.0.1/255.255.255.0
http_access deny !mynetwork
http_access allow pass
http_access deny all
```

Hier wordt gebruik gemaakt van de PAM authorisatie module en alle gebruikers moeten zich identificeren voordat ze toegang krijgen tot internet.

Opties als "authenticate_ttl" en "authenticate_ip_ttl" kunnen ook worden gebruikt om het authorisatie gedrag aan te passen, en wel door na bepaalde tijd opnieuw om een naam en wachtwoord te vragen.

Bronnen

Dit artikel is slechts het topje van de Squid-ijsberg; kijk voor meer informatie op de volgende websites:

- Squid Home pagina, www.squid-cache.org
- Squid Documentation Project, squid-docs.sourceforge.net
- visolve.com
- Voor Proxy Authorisatie, home.iae.nl/users/devet/squid/proxy_auth

Site onderhouden door het LinuxFocus editors team © D.S. Oberoi "some rights reserved" see linuxfocus.org/license/ http://www.LinuxFocus.org	Vertaling info: en --> -- : D.S. Oberoi <ds_oberoi/at/yahoo.com> en --> nl: Hendrik-Jan Heins <hjh/at/passys.nl>

2005-01-14, generated by lfparser_pdf version 2.51