Websites para bajar diferentes componentes de software:

- **Squid : http://www.squid-cache.org/**
- **IRCache: http://www.ircache.net/**
- **Guía de Squid: http://www.buanzo.com.ar/ver_articulo.html?n=9**
- **Shorewall : http://www.shorewall.net/**

Proxy Transparente con Squid y Netfilter

Lo crean o no esta relativamente simple tarea puede provocar más de un dolor de cabeza,

ya que consta de varios componentes que deben entenderse al 100%.

En principio, hay ciertos conceptos que vamos a tener que aclarar, para poder entender correctamente qué significa "Proxy Transparente". Estos términos, y su significado, son los siguientes:

Netfilter

Funcionalidad y esquema interno del Núcleo Linux en las versiones 2.4 y 2.6 que proveen Firewall con conocimiento de Estado (Stateful Firewall). Un firewall habilita la capacidad de aplicar políticas sobre los paquetes, como por ejemplo "permitir acceso desde cualquier IP al puerto 80 de la interfaz eth0", o tal vez "denegar el acceso al puerto 22, excepto a las IP 1, 2 y 3". La cuestión de conocimiento de estado está vinculada conque Netfilter mantiene una tabla de las conecciones entrantes y salientes, y de esta forma nos permite armar reglas en base a parámetros como conección establecida, relacionada o nueva. Por ejemplo, supongan que de 8 a 19hs se permiten nuevas conecciones salientes hacia Internet. Pasado este horario podríamos armar una regla que especifique que solamente las conecciones **relacionadas o establecidas,** pero no **nuevas** puedan seguir saliendo a Internet. De esta forma, la descarga de un archivo continuará hasta su fin, incluso pasado este horario, pero no se podrá ingresar a nuevos sitios. ¿Qué tiene esto que ver con un Proxy Transparente? Simple: Netfilter también permite aplicar ciertas reglas de redireccionamiento, no solo las clásicas de "ACEPTAR" y "RECHAZAR". En resumen, vamos a tener que aplicar un par de reglas de Netfilter (quizá a la mayoría les suene mas por su conjunto de utilidades, Iptables) para lograr nuestro acometido. Les recomiendo la lectura de los artículos sobre Iptables de la primer época de Linux USERS, divididos en 3 partes.

Proxy

Mucha gente confunde el término Proxy con el de Gateway (o "puerta de enlace predeterminada", según la traducción de cierta empresa de software privativo). En toda red TCP/IP, por ejemplo en una red privada 192.168.0.0 Clase C (Máscara 255.255.255.0 o '/24'), se necesita un gateway si es que deseamos llegar a otras redes, como Internet. Dicho Gateway poseerá la cantidad de interfaces necesarias y rutas establecidas y políticas de acceso que permitirán o no el acceso a ciertos destinos desde esta red interna. Por supuesto, en este caso hablamos de acceso "transparente" (por así decirlo) a la red destino en cuestión. Esto significa que el Gateway no tiene en cuenta el protocolo de aplicación (HTTP, FTP, etc) o mejor dicho que "no los entiende ni tiene en cuenta excepto por puerto

de origen o destino". Por ejemplo, se puede asumir que en el puerto 80 de cierta IP de destino habrá un servicio que entienda HTTP... pero el gateway no puede asegurarlo.

Un proxy actúa como gateway pero a un nivel más alto, en la llamada "capa de aplicación". Significa que entiende HTTP, FTP, u algún otro protocolo de alto nivel, y que acepta por parte de un cliente (de la red interna, por ejemplo) solicitudes vinculadas a dicho protocolo. El proxy realizará, a su vez, la solicitud al servidor de destino, tomará el resultado y lo devolverá. Al tener conocimiento del protocolo se pueden aplicar reglas mucho más interesantes, como restricciones basadas en contenido, partes del nombre de un sitio, usuario, grupo al que un usuario pertenece, IP de origen, etc. **Squid** es un proxy de HTTP y FTP, y a su vez provee la funcionalidad de Cache: guarda copias de las páginas y archivos visitados. De esta forma, cada vez que un usuario vuelve a acceder a cierto sitio, sólo el contenido que haya cambiado será transferido, logrando una reducción de la utilización del ancho de banda disponible.

NAT

Corresponde a Network Address Translation o Traducción de Dirección de Red. Las direcciones IP de una red privada no son direccionables en Internet, por lo que el Gateway suele aplicar lo que se llama comunmente "enmascaramiento" (Masquerading) de la IP de origen, reemplazando la interna por la correspondiente a la interfaz de red pública del Gateway. Por supuesto, se mantiene una tabla con los datos necesarios para poder relacionar las respuestas que provengan de internet con su destino "real" en la red privada.

Transparente

Bien, ya sabemos que es un Gateway, Netfilter y un Proxy. El hecho de que sea transparente permite al administrador lograr que toda solicitud HTTP (puerto de destino 80/tcp) realizada por un cliente de la red interna sea automáticamente redirigida al Proxy, evitando la salida directa. Los motivos para realizar ésto pueden depender del administrador, pero seguramente tengan que ver con políticas de administración de recursos, seguridad, performance, etc. Esto se realiza, como ya dijimos, mediante reglas de redireccionamiento de Netfilter (aplicadas con la utilidad Iptables).

Es importante aclarar que cierta funcionalidad del protocolo HTTP se pierde al utilizar un proxy transparente en vez de uno debidamente configurado en los clientes. Va más atllá del propósito de éste artículo el explicar dichos problemas. A modo general podemos responder "probablemente no tengas problemas notables". La práctica será lo mejor.

Ahora, el Squid

En principio instalen el paquete Squid desde el sistema de administración de paquetes de su distribución favorita (emerge, yast2, apt-get) o bajen la última versión (STABLE, o sino DEVEL si quieren probar las versiones en aún en proceso de desarrollo) del sitio de Squid (ver recuadro). El proceso de configuración no lo vamos a detallar, ya que ya lo hemos hecho en ediciones anteriores. De todas formas, el archivo /etc/squid/squid.conf (o, si instalan de fuentes sin cambiar el parametro **sysconfdir** del script **configure)**, / usr/local/squid/etc/squid.conf está impecablemente comentado, aunque en idioma Inglés. En el recuadro pueden encontrar la URL a una guía de configuración.

Una vez configurado el Squid, debemos probarlo en formato "no-transparente", configurando un navegador para que lo utilice. Recuerden que Squid utiliza el puerto 3128/tcp para recibir las solicitudes. Si funciona, podemos pasar a la siguiente etapa: la transparentización de Squid.

Squid en el Medio

En principio, son sólo 4 los parámetros los que debemos configurar. Uno de estos parámetros no lo encontrarán comentado y con un valor por defecto, sino que deberán tipearlo completo. Los parámetros y sus correspondientes valores, son los siguientes:

```
httpd_accel_host virtual
httpd_accel_port 80
httpd_accel_with_proxy on
httpd_accel_uses_host_header on
```

¿Qué significa cada uno de ellos?

httpd_accel_host : Squid puede configurarse como cache, como acelerador de navegación o como ambos. Este parámetro indica el nombre de host o IP de un Squid configurado como acelerador. En nuestro caso usaremos el valor "**virtual**", ya que vamos a usar el modo transparente.

httpd_accel_port : Este es el puerto donde generalmente los servidores web esperan solicitudes. En el caso de que usemos Squid como frente de un grupo de servidores web locales, habria que hacer que Squid escuche en el 80 y no en el 3128, y que los servidores web escuchen en otro puerto, y NO en el 80.

httpd_accel_with_proxy : Al habilitar httpd_accel_host la función de cache deja de funcionar. Para que vuelva a funcionar, debemos forzarlo con el valor "**on**".

httpd_accel_users_host_header : En modo transparente, las solicitudes salientes al puerto 80 de algun servidor externo son redirigidas al puerto 3128 del Gateway, en el que Squid espera solicitudes. Si no se habilitara esta opción los sitios que manejen sitios web virtuales, o sea, mas d eun sitio por dirección IP, no serían correctamente cacheados. La cabecera Host definida en HTTP 1.1 siempre tiene como valor el nombre de Host de la URL a donde estemos intentando navegar. Squid, si se habilita este parámetro, la utiliza y nos entrega el contenido requerido.

Reglas de Netfilter – Ultimo paso

Una vez configurado Squid con las opciones de proxy transparente requeridas, tan solo nos queda armar la regla de Iptables necesaria. Si utilizan Shorewall como solución de Firewall (muy buena, por cierto), en su FAQ y Documentación encontrarán como crear esta regla desde su sistema de configuración. (ver recuadro). La regla en cuestión es la siguiente:

```
iptables -t nat -A PREROUTING -i eth1 -p tcp --dport 80 -j REDIRECT --to-port 3128
```

Recuerden que el archivo access.log (definido en squid.conf) nos irá mostrando las solicitudes que se realizan, y así podremos verificar el correcto funcionamiento de la regla de direccionamiento.

Optimización Básica

Como una reglá mínima de performance podemos setear:

cache_store_log al valor **none**
half_closed_clientes al valor **off**
cache_swap_high al valor **100%**
cache_swap_low al valor **80%**

De la misma forma se puede establecer correctamente el parámetro cache_dir con la siguiente fórmula:

Siendo:
x = Tamaño_de_cache_en_KB
y = Tamaño_de_objeto_promedio_en_KB (aprox. 18Kb)

El valor MD será (((x / y) / 256) / 256) * 2

La línea cache_dir, entonces, será, por ejemplo, para 6Gb de cache:

```
cache_dir ufs /var/cache/squid 6144 14 256
```

Saludos, y hasta la próxima entrega! Envíen sus solicitudes a linux@tectimes.com

©Arturo A. Busleiman 2004
e-mail: buanzo@buanzo.com.ar