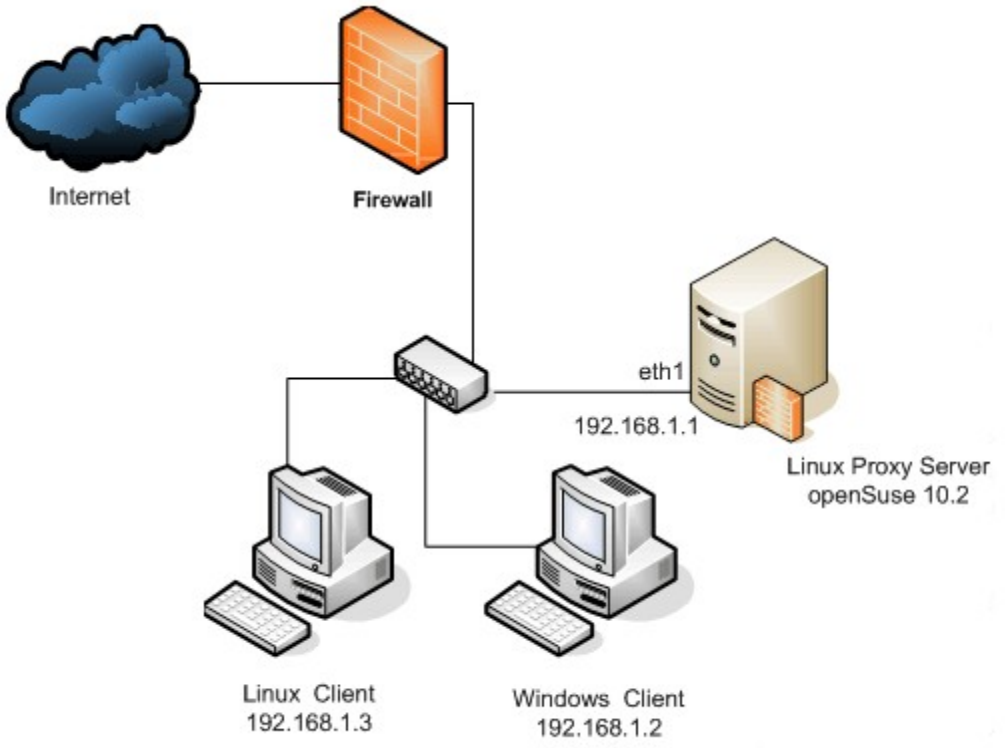
TEMA 9
Proxy

1. Proxy

El término en ingles «**Proxy**» tiene un significado muy general y al mismo tiempo ambiguo, aunque invariablemente se considera un sinónimo del concepto de «**Intermediario**». Se suele traducir, en el sentido estricto, como **delegado** o **apoderado** (el que tiene el que poder sobre otro).

Un **Servidor Intermediario (Proxy)** se define como un servicio de red que permite a los clientes realizar conexiones de red indirectas hacia otros servicios de red. Su finalidad por tanto, es interceptar las conexiones de red que un cliente hace a un servidor de destino y puede ser realizada por un programa o dispositivo.



Los motivos por los que se introduce un proxy en la red son seguridad, rendimiento, anonimato, etc.

Si tenemos un proxy en la red el proceso que ocurre es el siguiente:

- Cliente se conecta hacia el Servidor Intermediario (Proxy).
- Cliente solicita al proxy una conexión, fichero u otro recurso disponible en un servidor distinto.
- El servidor Intermediario (Proxy) proporciona el recurso ya sea conectándose al servidor especificado o sirviendo éste desde su caché.
- En algunos casos el Servidor Intermediario (Proxy) puede alterar la solicitud del cliente o bien la respuesta del servidor para diversos propósitos.

Los Servidores Intermediarios (Proxies) se distinguen de los **cortafuegos** en que los proxys trabajan a nivel de aplicación, mientras que los cortafuegos operan en el Nivel de Red. Los firewalls deniegan el acceso a ciertas IP y puertos TCP o UDP, y los proxys deniegan el acceso a ciertas páginas web o usuarios de la red.

Una aplicación común de los Servidores Intermediarios (Proxies) es funcionar como **caché** de contenido de Red (principalmente HTTP),

proporcionando en la proximidad de los clientes un caché de páginas y ficheros disponibles a través de la Red en servidores HTTP remotos, permitiendo a los clientes de la red local acceder hacia éstos de forma más rápida y confiable.

Cuando se recibe una petición para un recurso de Red especificado en un **URL** (Uniform Resource Locator) el Servidor Intermediario busca el resultado del URL dentro del caché. Si éste es encontrado, el Servidor Intermediario responde al cliente proporcionado inmediatamente el contenido solicitado. Si el contenido solicitado no estuviera disponible en el caché, el Servidor Intermediario lo traerá desde servidor remoto, entregándolo al cliente que lo solicitó y guardando una copia en el caché.

El contenido en la caché es sustituido de acuerdo con el algoritmo de expiración elegido. Este podrá basarse en la antigüedad de la información almacenada, su tamaño o el historial de solicitudes recibidas.

El algoritmo de expiración se puede configurar en el proxy, siendo los más habituales:

LRU (Least Recently Used o Menos Recientemente Utilizado). En este algoritmo los objetos que no han sido accedidos en mucho tiempo son eliminados primero, manteniendo siempre en el caché a los objetos más recientemente solicitados. Ésta política es la utilizada por Squid (el proxy de Linux) de modo predefinido.

LFUDA (Least Frequently Used with Dynamic Aging o Menos Frecuentemente Utilizado con Envejecimiento Dinámico). En este algoritmo los objetos más veces solicitados permanecen en la caché, eliminándose los que menos veces se han solicitado.

GDSF (GreedyDual Size Frequency o Frecuencia de tamaño GreedyDual). Este algoritmo, además de tener en cuenta la frecuencia con la que se reciben solicitudes, también tendrá en cuenta el tamaño de los objetos solicitados. De este modo, se da prioridad a los objetos pequeños (que ocupan menos memoria) con muchas solicitudes.

2. Servidor proxy en Linux: Squid

Squid es un Servidor Intermediario (Proxy) de alto desempeño que se ha venido desarrollando desde hace varios años y es hoy en día un muy popular y ampliamente utilizado entre los sistemas operativos como GNU/Linux y derivados de Unix®. Es muy confiable, robusto y versátil y se distribuye bajo los términos de la Licencia Pública General GNU (GNU/GPL). Siendo sustento lógico libre, está disponible el código fuente para quien así lo requiera.

Entre otras cosas, Squid puede funcionar como Servidor Intermediario (Proxy) y caché de contenido de Red para los protocolos HTTP, FTP, GOPHER y WAIS, Proxy de SSL, caché transparente, WWCP, aceleración HTTP, caché de consultas DNS y otras muchas más como filtración de contenido y control de acceso por IP y por usuario.

2.1. Instalación y arranque de Squid

Para instalar squid, una vez comprobada la configuración de red tan sólo tendremos que abrir el terminal y escribir:

```
sudo apt-get install squid
```

2.2. Inicio y parada del servicio

Para arrancar squid escribimos:

```
service squid start
```

Para pararlo:

```
service squid stop
```

Y para reiniciarlo:

```
service squid restart
```

También podemos ver si el servicio está activo o no con el siguiente comando:

```
service squid status
```

2.3. Configuración básica de Squid

El archivo de configuración se encuentra en ***/etc/squid/*** y se llama ***squid.conf***. Antes de modificarlo, conviene realizar una copia de seguridad.

Los espacios en blanco al principio de una línea puede dar problemas. Así que, los evitaremos siempre.

Una configuración básica debe incluir, al menos, los parámetros que se indican a continuación:

- **http-port**: Establece el puerto de escucha para squid (por defecto puerto 3128).

```
http_port 3128
```

Si nuestro proxy va a trabajar de modo transparente, en esta línea se debe añadir el parámetro "transparent":

```
http_port 3128 transparent
```

- **visible_hostname**: nombre del equipo que queremos que aparezca en los mensajes de error.
- **acl**: a cada ACL o lista de control de acceso se le hace corresponder una regla de control de acceso (http_access) que es la que permite o deniega las conexiones definidas en cada acl (más adelante veremos que son las acl).

Otros parámetros importantes son:

- **cache_dir**. Establece la localización y el tamaño de la caché en el disco duro. Ejemplo: **cache_dir ufs /var/spool/squid 100 16 256**. ufs es el sistema de almacenamiento que utiliza squid. 100 el tamaño en megas de la caché, 16 el nivel de subdirectorios de primer nivel y 256 el segundo nivel de subdirectorios por cada directorio de primer nivel.
- **cache_replacement_policy**: se indica el algoritmo de sustitución que va a utilizar la memoria cache cuando se llena. Squide permite los 3 algoritmos siguientes: LRU, LFUDA eta GDSF. Por defecto utilizará LRU.

2.3.1. Las listas de control de acceso o ACL

Mediante las listas de control de acceso se asignan nombres a grupos de maquinas y a otros conceptos para luego asignar o denegar permiso de acceso mediante las reglas de control de acceso.

- **acl nombre src [IP] [IPRED/MASCARA] [ARCHIVO]**

```
acl equipo1 src  192.168.1.23
acl redlocal src 192.168.1.0/255.255.255.0
acl varios src “/etc/squid/varios”
```

Contenido del archivo /etc/squid3/varios

```
192.168.2.45
192.168.2.78
192.168.2.46
```

- **acl nombre dstdomain [nombre de dominio] [archivo]**

```
acl prohibidas dstdomain www.google.es www.hotmail.com
acl prohibidas dstdomain "/etc/squid/prohibidas"
```

Contenido del archivo /etc/squid/prohibidas

```
www.google.es
www.hotmail.es
```

- **acl nombre time [días][rango horario] (S-domingo, M-lunes, T-martes, W-miercoles, H-jueves, F-viernes, A-sabado)**

```
#sabados y domingos de 8h a 24h
acl horario1 time AS 08:00-24:00
```

- **acl nombre url_regex [caracteres] [archivo]**

Permite especificar expresiones regulares como la extensión de los sitios Web:

```
acl extensiones url_regex .com .es
```

- **acl nombre urlpath_regex [caracteres] [archivo]**

Permite especificar la extensión de ficheros a descargar:

```
acl descargas urlpath_regex \.pdf$ \.doc$
```

- **ACLs predefinidas**

En el archivo de configuración ya existen unas acl creadas con nombre **localhost**, **all** (cualquier equipo), **localnet** (red a la que pertenece el servidor).

Una vez hemos creado los elementos que intervienen en la red, vamos a definir las reglas que van a determinar el funcionamiento del proxy. En concreto las reglas que van a controlar las solicitudes HTTP a servidores Web. El proxy analizará estas reglas y decidirá si responde a la solicitud del cliente o no.

2.3.2. Las reglas de acceso http_access

```
http_access allow/deny !nombreAcl
```

Allow: permite el acceso a las solicitudes que cumplan las condiciones de los acl (si hay varios comprueba que se cumpla las condiciones de todos)

Deny :deniega el acceso a las solicitudes que cumplan las condiciones de los acl.

Se puede utilizar el operador ¡ delante del nombreAcl con lo que negamos las condiciones

```
#permiso de acceso a todos los equipos de la redlocal
http_access allow redlocal
#se deniega el permiso al equipo1
http_access deny equipo1
#se deniega el permiso a las prohibidas desde cualquier equipo
http_access deny prohibidas
#se deniega el permiso en ese horario1 desde cualquier equipo
http_access deny horario1
#se deniega el acceso a las paginas incluidas en prohibidos al equipo1
http_access deny equipo1 prohibidos
```

Hay que tener muy en cuenta el orden de las reglas de acceso. Se deben colocar de más concretas a mas generales. Cuando al proxy le llega una petición, se comprueba en orden cada regla hasta que alguna se aplicable, a partir de ahí no se comprueban mas reglas. Por defecto se permite todo lo que no esté denegado explícitamente.

```
http_access allow equipo1
http_access deny prohibidos redlocal
http_access allow horario redlocal
http_access deny all ¡localhost
```

El equipo1 no tiene ninguna restricción, los equipos de la red local no pueden acceder a las paginas prohibidas y solo pueden acceder en el horario indicado.

Con la última restricción se deniega el acceso en cualquier otro caso (excepto a localhost)

2.4. Mensajes de error y registro

Por lo general Squid viene preconfigurado con mensajes en inglés, podemos modificarlo para que estos mensajes de error salgan en español o poner los nuestros propios. Si queremos que aparezcan en español en el archivo de configuración de squid pondremos:

```
error_directory /usr/share/squid/errors/Spanish
```

Si queremos modificar por ejemplo el mensaje de error que aparece al impedir el acceso a una determinada página editaremos el archivo */usr/share/squid/errors/Spanish/ERR_ACCESS_DENIED*.

Los ficheros donde se guarda el registro de incidencias son:

/var/log/messages o /var/log/syslog), en el almacena mensajes de error

/***var/log/squid/Access.log***, en el almacena información sobre los accesos

/var/log/squid/cache.log, información sobre el cache de objetos.

2.5. Varios servidores proxy en la misma red

El parámetro cache_peer se utiliza para especificar otros Servidores Intermediarios (Proxies) con caché en una jerarquía como padres o como hermanos. Es decir, definir si hay un Servidor Intermediario (Proxy) adelante o en paralelo. La sintaxis básica es la siguiente:

```
cache_peer servidor tipo http_port icp_port opciones
```

Ejemplo: Si su caché va a estar trabajando detrás de otro servidor cache, es decir un caché padre, y considerando que el caché padre tiene una IP 192.168.1.1, escuchando peticiones HTTP en el puerto 8080 y peticiones ICP en puerto 3130 (puerto utilizado de modo predefinido por Squid) ,especificando que no se almacenen en caché los objetos que ya están presentes en el caché del Servidor Intermediario (Proxy) padre, utilice la siguiente línea:

```
cache_peer 192.168.1.1 parent 8080 3130 proxy-only
```

Cuando se trabaja en redes muy grandes donde existen varios Servidores Intermediarios (Proxy) haciendo caché de contenido de Internet, es una buena idea hacer trabajar todos los caché entre si. Configurar caches vecinos como sibling (hermanos) tiene como beneficio el que se consultarán estos caches localizados en la red local antes de acceder hacia Internet y consumir ancho de banda para acceder hacia un objeto que ya podría estar presente en otro caché vecino.

Ejemplo: Si su caché va a estar trabajando en paralelo junto con otros caches, es decir caches hermanos, y considerando los caches tienen IP 10.1.0.1, 10.2.0.1 y 10.3.0.1, todos escuchando peticiones HTTP en el puerto 8080 y peticiones ICP en puerto 3130, especificando que no se almacenen en caché los objetos que ya están presentes en los caches hermanos, utilice las siguientes líneas:

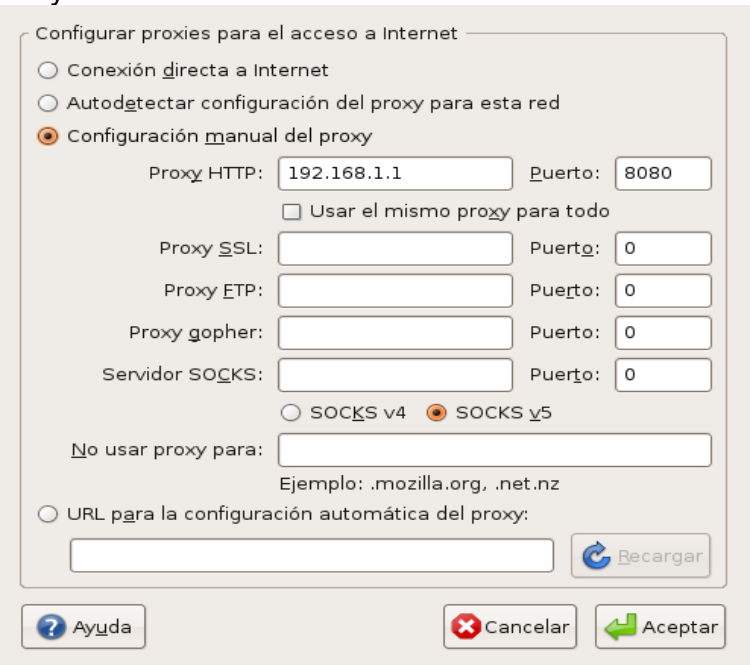
```
cache_peer 10.1.0.1 sibling 8080 3130 proxy-only
cache_peer 10.2.0.1 sibling 8080 3130 proxy-only
cache_peer 10.3.0.1 sibling 8080 3130 proxy-only
```

Pueden hacerse combinaciones que de manera tal que se podrían tener caches padres y hermanos trabajando en conjunto en una red local. Ejemplo:

```
cache_peer 10.0.0.1 parent 8080 3130 proxy-only
cache_peer 10.1.0.1 sibling 8080 3130 proxy-only
cache_peer 10.2.0.1 sibling 8080 3130 proxy-only
cache_peer 10.3.0.1 sibling 8080 3130 proxy-only
```

3. Configuración de los clientes

Para que cada uno de los clientes de la red pueda comunicarse con Squid, debemos configurar el navegador en cada uno de ellos para que salgan a Internet a través del proxy. En Mozilla Firefox, por ejemplo, debemos ir a Editar/Preferencias/Avanzado/Red/Configuración. Seleccionamos a continuación "Configuración manual del proxy" e introducimos la IP y puerto de escucha del proxy.



4. Ejercicios

1. Crea un archivo de configuración para denegar el acceso a todos los equipos a la dirección www.google.es
2. Crea un archivo de configuración que deniegue el acceso a las direcciones www.google.es y www.hotmail.com
3. Crea en tu carpeta personal un archivo llamado no_permitidos que contenga las direcciones de los tres siguientes dominios:
   - www.google.es
   - http://es.yahoo.com/
   - http://es.msn.com/

   A continuación crea un archivo de configuración squid.conf que deniegue las conexiones a las direcciones que se encuentran en el archivo no_permitidos.
4. Se dispone de una red local con dirección 192.168.1.0 y máscara 255.255.255.0. Crear un archivo de configuración squid.conf que permita el acceso a Squid a todos los ordenadores de la red y no lo permita a los restantes.

5. Se dispone de una red de área local con dirección 192.168.1.0 y máscara 255.255.255.0. Se desea permitir el acceso a Squid a los ordenadores con las IP que están comprendidas en el rango 192.168.1.1 y 192.168.1.10 (ambas incluidas). Crea en tu directorio personal un fichero llamado ip_permitidas que tenga estas direcciones (cada dirección en una línea diferente). A continuación indica que fichero de configuración para Squid crearías para permitir el acceso a Squid a todas estas direcciones y denegar el acceso a las restantes.
6. Deniega las conexiones a todos los equipos en horario de 18:00 a 21:00 horas.
7. Deniega las conexiones a todos los equipos en horario de 18:00 a 21:00 horas, pero sólo los lunes, martes y miércoles.
8. Deniega el acceso a Squid al equipo con IP 192.168.1.5. Permite el resto de accesos al servidor proxy.
9. Deniega el acceso a Squid al equipo con IP 192.168.1.5 en horario de 18:00 a 21:00 horas. Permite el resto de accesos a Squid.
10. Deniega el acceso a Squid al equipo con IP 192.168.1.5 en horario de 18:00 a 21:00 horas. Para el resto de equipos permitir el acceso sólo en horario de 10:00 a 14:00 horas. Se supone que los equipos pertenecen a la red 192.168.1.0 con máscara 255.255.255.0.
11. En el archivo /etc/squid/permitidos se tiene una lista de todas las direcciones IP de la red local. El equipo10 tiene la dirección IP 192.168.1.10. Se permite el acceso a Internet al equipo10 de lunes a miércoles de 9:00 a 14:00 horas. También se permite el acceso a los equipos de la red local de lunes a miércoles. Se prohíbe el acceso en el resto de casos.
12. Restringe el acceso a todo el contenido con extensión .mp3 a los ordenadores del aula.
13. Restringe el acceso a todo el contenido con extensión .mp3 a los ordenadores del aula en horario de 9:00 a 14:00 horas.
14. A los equipos comprendidos en el rango de direcciones IP 192.168.1.1-192.168.1.5 se les quiere denegar el acceso a las páginas web que están en el archivo /etc/squid/sitios_permitidos los fines de semana. Se supone que las direcciones de ese rango están en el archivo /etc/squid/permitidos. El resto de equipos de la red tienen acceso a todas las páginas pero no los fines de semana, en el que no tienen acceso a ninguna página.
15. Define un fichero de configuración para que los equipos de la red (192.168.35.0/24) solo puedan acceder a internet de las 14h a las 17h de lunes a viernes. Que no puedan acceder a los dominios www.yonkys.com y www.marca.es y que no se puedan descargar archivos mp3 ni avi. El director (IP 192.168.35.2) no estará bajo ninguna restricción.
16. Explica que utilidad tiene y como funciona un servidor proxy-cache
17. ¿Como podemos conseguir que los mensajes del servidor al cliente sean en castellano?
18. Instala dos servidores proxy y haz que uno de ellos sea el padre
19. Instala dos servidores proxy y configuralos como hermanos
20. En una red local tenemos 5 servidores proxy:
    - PR_N: servidor principal (de primer nivel), Tiene la IP 10.0.0.1
    - PR_1: servidor de segundo nivel, Tiene la IP 10.0.0.2
    - PR_2: servidor de segundo nivel, Tiene la IP 10.0.0.3

- PR_1_1: servidor de tercer nivel, hijo de PR_1. Tiene la IP 10.0.0.4
- PR_1_2: servidor de segundo nivel, hijo de PR_1. Tiene la IP 10.0.0.5

Todos los proxys utilizan el puerto 3128 como puerto HTTP y el 3130 como puerto ICP. Teniendo en cuenta todas las relaciones existentes entre ellos, dibuja el esquema jerárquico, y define qué líneas habría que escribir en el archivo de configuración de cada proxy para que el esquema funcione.

5. Prácticas en Ubuntu

1. Deniega el acceso a Internet al equipo Windows 7
2. Deniega el acceso a Internet a todo los equipos de la red local y permite sólo al equipo Windows 7
3. Deniega a todos los equipos de la red el acceso a la página del marca
4. Deniega el acceso del equipo Windows 7 a las siguientes páginas que estarán indicadas en el fichero /etc/squid/paginas_web
5. Permite el acceso a Internet a todo los equipos de la red local de lunes a viernes pero deniega al equipo Windows 7 de 10 a 12. Además, deniega la descarga de ficheros pdf entre semana de 9 a 13 a toda la red local.
6. Deniega el acceso a Internet a todo los equipos de la red local de lunes a viernes pero permite al equipo Windows 7 de 9 a 11 visitar la página de moodle. Además, permite a toda la red local que hoy puedan visitar las página .org.
7. Leen la información que nos dan los apuntes sobre los mensajes del proxy y configúralos en castellano.
8. Modifica el mensaje de error para que de el nombre y el teléfono del administrador