# Seguridad y Alta Disponibilidad

## Instalación y Configuración de servidores proxy

Javier Ayllón Pérez
Escuela Superior de Informática
Universidad de Castilla-La Mancha

## Contenidos

- Introducción teórica
  - Conceptos, relaciones y protocolos
- Demostración práctica
  - Clientes, Servidores.
  - Instalación y configuración



## PRIMERA PARTE

INTRODUCCIÓN TEÓRICA



## Motivación

- El problema es la velocidad.
  - Líneas de comunicaciones
  - Servidores saturados
- Contenidos estáticos vs dinámicos
- Técnicas
  - Navegadores
  - ISP's: Caches y réplicas



## Definiciones

- RFC 3040 de la IETF:
  - Cliente: Programa que establece comunicación con el propósito de realizar peticiones.
  - Servidor: Programa que acepta conexiones con el propósito de servir peticiones mediante el envió de respuestas.
  - Proxy: "Programa intermediario que actúa como cliente y servidor con el propósito de realizar peticiones en nombre de otros clientes. Las peticiones serán servidas internamente o pasadas a terceros, quizás con modificación de la petición"

## Definiciones

- Proxy transparente: No modifica la petición o respuesta
- Proxy no-transparente: Modifica la petición o respuesta con la finalidad de proveer servicios añadidos.
- Cache: Sistema de almacenamiento local para respuestas. También es parte de la cache el subsistema que lo organiza, salva y borra todos estos mensajes. Un cache salva mensajes “almacenables” para reducir los tiempos de respuesta.
- Almacenable: Mensaje susceptible de ser salvado por un sistema cache. Si es o no posible esto depende de muchos factores.



## Cacheabilidad

- Susceptibilidad de un contenido de ser salvado en un sistema cache
- Cuestiones técnicas:
  - Http 1.1 Contiene mecanismos para controlar el contenido de las caches
  - “Frescura” de contenidos
  - Una sesión encriptada es dificilmente cachable.
  - Servicios de streaming en tiempo real
- Cuestiones legales:
  - Contenidos protegidos por copyright.
  - Accesos a sistemas donde los logs son requeridos



## Servidor proxy

- Concepto mal usado para referirse a servidor proxy de cacheo “caching proxy”, también es correcto llamarlo “proxy cache”.
- Actúa como servidor para los clientes y como cliente para los servidores.
- “Proxy inverso”: Servidor proxy que responde a peticiones como si fuera el destinatario final de la petición y que colabora estrechamente con éste.



## Agrupaciones de proxy’s

- Matriz de proxys, también conocido como “cluster de proxy’s”. Es un conjunto de servidores que actúan como uno solo particionando el espacio de nombres.
- Malla de proxys. Conjunto de servidores que actúan independientemente pero que comparten contenidos de cache mediante protocolos de comunicación entre caches.
- Estos protocolos de comunicación permiten que las peticiones sean enrutadas a los servidores proxy adecuados según el propietario de respuestas.



## ICP

- Internet Cache Protocol. Es un protocolo de comunicación entre servidores proxy que permite que se pregunten unos a otros sobre los contenidos de sus caches.
- RFC 2186 Describe una agrupación jerárquica de servidores. Cuando una petición no puede ser atendida por un servidor, éste pasa a los hijos la petición, cuando esta es encontrada es pasada directamente al cliente.



## Replicación

- Replicación de contenidos en internet. Mirror.
- Un servidor copia el contenido completo de otros sitio.
- No existe protocolo transparente al cliente
- El cliente debe saber explícitamente donde se encuentra ej. http://ftp.rediris.es/mirror/
- Contenidos volcados a intervalos fijos de tiempo.
  - Ej Diariamente.
- Protocolos de copia
  - RDIST.



## Configuración clientes

- Agente de usuario: navegador web, aplicación cliente o robot
- Configuración para uso de proxy.
  - Manual: El usuario configura cual será el proxy a usar
  - Automático: El agente descubre el proxy sin intervención del usuario
  - Interceptación de tráfico. Elementos de red activos se encargan de la inspección de las peticiones.
  - Redirección de tráfico. Las peticiones inspeccionadas y determinadas ser susceptibles de cacheo son reencaminadas a los servidores proxy.
  - Proxy de interceptación. Proxy que recibe peticiones de un cliente en el que no se ha realizado configuración de proxy.



## Configuración manual

- El administrador de la red provee los datos del servidor proxy
- Estos datos son distribuidos a todos los usuarios en forma escrita.
- Cada usuario configura, de acuerdo a la documentación de su “user-agent” los datos recibidos.
- El “user-agent” funciona redireccionando todas las peticiones de usuario al proxy.
- Ej Internet Explorer:
  - Herramientas, Opciones de Internet, Configuración, Servidor Proxy
  - Configuramos dirección y puerto.

# Configuración automática

- Auto-configuración de Proxy.
- Manualmente se configura una URL en la que se encuentra la configuración.
- Esta URL es descargada cada vez que se arranca el navegador.
- Formato PAC. Es un archivo JavaScript.
- Facilita que un administrador pueda cambiar las políticas de forma centralizada.
- Permite el uso de Cache Array Routing Protocol
- Ej Internet Explorer:
  - Herramientas, Opciones de Internet, Configuración, Servidor Proxy
  - Usar Scripts de configuración automática



# CARP

- Protocolo para el rutado de peticiones de cache a clústers.
- Poner inteligencia en la petición de contenidos a arrays de cache.
- Usar códigos hash para direcciones URL
- Aún en “draft” de la IETF pero ya implementado por Microsoft y Squid



# Configuración transparente

- También llamado auto-configuración.
- Este mecanismo trata de localizar automáticamente el servidor de configuración PAC, no el servidor proxy.
- Usa protocolo WPAD Web Proxy Auto Discovery Protocol
- Problema: Diferentes implementaciones
  - DHCP Dynamic Host Configuration Protocol
  - SLP Service Location Protocol
  - DNS. Domain Name Services. Mediante un Registro bien conocido, puntero SRV o Registro TXT
- Internet Explorer:
  - Herramientas, Opciones de Internet, Configuración, Servidor Proxy
  - Detectar la configuración automáticamente.



# WPAD

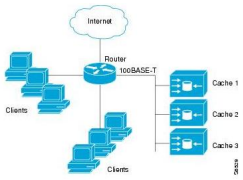
- Internet Explorer
  - Usa la opción DHCP 252
  - En ésta se indica la URL donde localizar el archivo wpad.dat
  - Si la respuesta DHCP no es correcta usa DNS
- Firefox
  - Solo usa DNS
  - Se hace la petición http://wpad.<nombre_dominio>/wpad.dat
- Ejemplo archivo wpad.dat

```
function FindProxyForURL(url, host)
  { return "PROXY proxy.example.com:8080; DIRECT"; }
```



# Interceptación de tráfico

- Un elemento de red inspecciona el tráfico y lo reencamina
- Según la configuración de proxy, la petición será encaminada al servidor origen (Internet) o proxys.
- Protocolo WCCP Web Cache Communication Protocol






# WCCP

- En estado “borrador” de la IETF.
- Protocolo incluido en los enrutadores Cisco
- Los servidores proxy son clientes y se registran
- El router es servidor y acepta registros de proxy’s
- Las peticiones son reescritas, cambiando la MAC de destino. Alternativamente se usa GRE.
- Los servidores proxy devuelven el contenido a los clientes como si fueran el propio servidor origen.
- Soportado por fabricantes de software cliente.



# Referencias


- IETF RFC 3040 Taxonomía de replicación y cacheo web
- IETF RFC 2616 HTTP /1.1
- IETF RFC 2186 Internet Cache Protocolo
- Cisco WCCP http://www.cisco.com/en/US/docs/ios/12_0t/12_0t3/feature/guide/wccp.html




# SEGUNDA PARTE

DEMOSTRACIÓN PRÁCTICA

## Clientes

- Internet Explorer
- Firefox
- Safari
- Opera



## Servidores

- Microsoft Forefront TMG Treat Managment Gateway
- SQUID



## SQUID - Intro

- Servidor proxy “Open Source y Free”. Licencia GNU
- Fundación nacional para la Ciencia EEUU. Principalmente universidad de California en San Diego
- Hoy mantenido por voluntarios, pero con donaciones importantes de empresas como Kaperski o SGI.
- Posibilidad de contratos de mantenimiento con empresas
- “Fork” comercial es netcache hoy propiedad de netapp.
  - Este se vende en formato appliance



## SQUID - Requerimientos

- Requiere un PC con linux (compatible con otros SSOO)
- Disco duro rápido y grande. Cache
- Gran ancho de banda. Proxy
  - Normalmente dos interfaces de red
- Potencia de CPU, procesamiento de peticiones.



## SQUID - Instalacion

- Configuración de sistema operativo dedicado
  - Ejemplo para familia Red Hat (Fedora)
- Actualización de parches
- Configuración de Red, IP’s fijas.
- Muy conveniente registro en DNS
  - Opcionalmente con registro para WPAD
- Instalación mediante repositorio yum
  - yum install squid
- Configuración de firewall.



## SQUID – Arranque

- Comprobamos que el servicio está configurado para comoienzo en arranque.
  - chkconfig –list squid
  - chkconfig squid on
- Arrancamos
  - Service squid start
- Comprobamos archivos de log
  - /var/log/squid
    - access.log
    - cache.log



## SQUID – Comprobación de arranque

- Vemos los procesos levantados
  - ps -ef | grep squid
- Vemos los puertos abiertos
  - netstat -anp | grep squid



## SQUID – Prueba

- Configuramos clientes de forma manual
  - Internet Explorer
  - Firefox
- En un terminal examinamos el log de accesos
  - tail –f /var/log/squid/access.log
- Realizamos conexión, ej www.google.es
- ¿Funciona?
- Analizamos log

# SQUID – Configuración

- Archivo de configuración de text
  - /etc/squid/squid.conf
- Dirección y puerto de escucha. http_port.
- Por defecto completamence cerrado
  - Excepto localnet, localhost
- Añadimos nuestro host
  - acl cliente  src 161.67.X.X/32
  - http_access allow cliente
- Reinicialmos el servidor
  - service squid reload



# SQUID – Configuración II

- Configuración de la cache.
  - Cache_dir <tipo> <max_tam_ob> <dir> <tam> <L1> <L2>..
    - <tipo> Es el tipo de sistema de archivos, se usa normalmente ufs
    - <max_tam_ob> Tamaño máximo para un objeto en esta cache
    - <dir> Directorio del servidor donde se sitúan los objetos de cache
    - <tam> El espacio máximo en disco que se usará en total
    - <L1><L2>… Son el número máximo de directorios creados en cada nivel



# SQUID – Configuración III

- Squid es muy flexible en cuanto a configuración
  - Seguridad de autenticación
  - Seguridad en restricción de acceso a clientes
  - Seguridad en restricción de contenidos accesibles
  - Soporte SSL como proxy-inverso.
    - Solo actúa como punto final de conexión https, después crea un túnel hacia el servidor del que saca los contenidos.
    - Se puede usar con aceleradores SSL



# SQUID – Configuración III

- Restricción de acceso según aplicación cliente
- Restricción de acceso a sitios web



# SQUID – Depuración

- A veces es complicado encontrar problemas de configuración
- Puede haber matching de reglas inesperadas o lo contrario.
- Primitiva de configuración
  - debug_options ALL,1
  - Vuelca depuración en cache.log