Capítulo 4

Monitorización estadística de tráfico

Índice del capítulo

4.1. Introducción a *NetFlow*

NetFlow es un protocolo de comunicaciones ideado por *Cisco Systems*. En primer lugar fue concebido como un protocolo para que distintos encaminadores y conmutadores de la red pudiesen intercambiar rápidamente información de su estado, notificando a los demás conmutadores si estaban en estado de congestión y, teniendo la información del estado de los demás conmutadores, poder decidir realizar el encaminamiento del tráfico por otro segmento de la red.

NetFlow evoluciona y comienza a ser utilizado no sólo como una herramienta para la comunicación entre los encaminadores sino que también se comienza a plantear su uso como un sistema de recogida de estadísticas sobre tráfico [2]. Con *NetFlow*, es posible la recolección de información del estado de los encaminadores y máquinas de la red, sin necesidad de pasar luego a ejercer una acción de control sobre los dispositivos.

En este capítulo del documento nos vamos a centrar sobre el uso de *NetFlow* para la recogida de la información del estado de la red en sistemas remotos y

su envío al sistema donde se desea procesar la información. En primer lugar haremos una breve visión sobre algunas sondas disponibles y sus principios de funcionamiento. Tras ello hablaremos sobre la *sonda* que hemos utilizado, que será la utilidad que nos genera la información. Estudiaremos las posibilidades de generación de información que ofrece la sonda, su configuración y haremos un comentario referente a sus posibilidades de seguridad. En tercer lugar trataremos sobre el *recolector* que hemos elegido, que es la utilidad que recoge o recopila la información de las sondas. El cuarto punto versará sobre algunas cuestiones a tener en cuenta acerca de la seguridad de la información de *NetFlow*. Por último, en el apéndice A, "*Monitorización estadística de tráfico*", veremos, entre otras, cuestiones referentes a instalaciones y configuraciones de las distintas utilidades.

4.2. Sonda de *NetFlow*: *fprobe-ng*

4.2.1. Sondas de *NetFlow*

En el punto 4.1 hemos mencionado ya el concepto de *sonda*. Lo que aquí denominamos *sonda* es un programa que se ejecuta en un sistema remoto (sistema que tiene acceso físico al tráfico que deseamos monitorizar), elabora la información estadística y la envía a un segundo sistema que denominaremos *recolector* (sistema que será nuestro sistema de monitorización). Para identificar con claridad la ubicación que tendrían las distintas sondas y el recolector en la ilustración de red que hemos estado siguiendo mostramos la Figura 4.1.

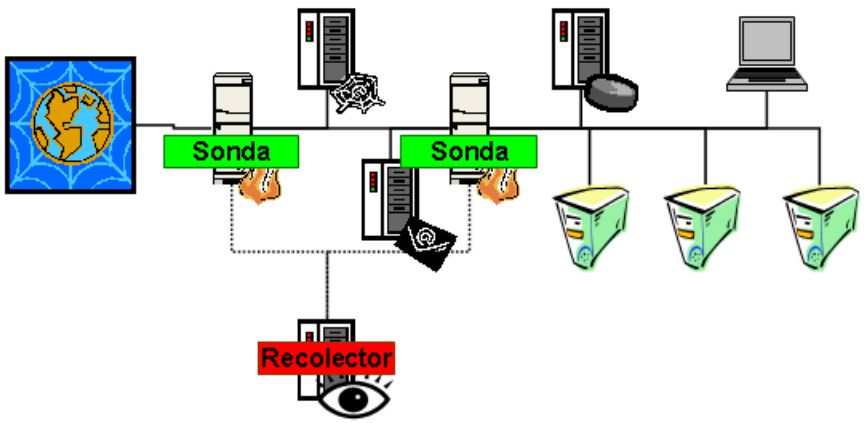

Figura 4.1: Ubicación de las sondas y el recolector

Para los sistemas GNU/Linux existen distintas sondas que permiten la recogida de la información en los sistemas remotos. Algunas de ellas son:

- *softflowd*: una sonda que escuchará en modo promiscuo en un interfaz del sistema en el que se encuentre instalada. Soporta IPv6 y es capaz de leer archivos *pcap* grabados por *tcpdump* y sistemas compatibles. No se

proporciona como *paquete Debian*. Para más información acerca de esta sonda véase [6].

- *pcNetFlow*: sonda que actualmente se encuentra sin desarrollo. Utiliza el principio de funcionamiento de la anterior: capturar los datos al funcionar en modo promiscuo en un interfaz, elaborando la información y transmitiéndola vía *NetFlow* hacia un recolector de destino. Para más información acerca de esta sonda véase [7].

- *fprobe*: una sonda de funcionamiento similar a la anterior. Exportará la información capturada en formato *NetFlow* v5. Actualmente no se encuentra muy actualizada a pesar de que se proporciona como *paquete Debian*. Para más información acerca de esta sonda véase [8].

- *fprobe-ng*: se presenta como una sonda alternativa para reemplazar a la anterior. Basada en *libpcap* también actúa como un *sniffer* en un interfaz para recopilar la información del tráfico que lo atraviesa. Soporta *NetFlow* versiones 1, 5 y 7, tiene más opciones de configuración que su predecesora. Se encuentra disponible como *paquete Debian*. Para más información acerca de esta sonda véase [9].

- *fprobe-ulog*: un proyecto paralelo al anterior que utiliza un método distinto para la captura de la información. En vez de basarse en *libpcap* utilizará el núcleo de *netfilter*, el cortafuegos del sistema, para identificar con cadenas de *iptables* el tráfico que queremos analizar por lo que requiere de una configuración adicional y compatibilidad del sistema con el módulo *ULOG* del *kernel*, resultando en una utilización bastante más compleja. También se ofrece como *paquete Debian*. Para más información acerca de esta sonda véanse [9, 11, 12].

De la lista anterior de sondas nosotros hemos optado por *fprobe-ng* dado la facilidad de su uso al basarse en *libpcap*, sus posibilidades de configuración y al ofrecerse como *paquete Debian*.

4.2.2. Introducción al paquete *fprobe-ng*

Nosotros hemos trabajado sobre una distribución **Debian Sarge** que actualmente se encuentra en estado estable, y nos hemos centrado la sonda ***fprobe-ng***, que está disponible en el momento de desarrollo de este Proyecto como *paquete Debian* en su versión 1.1-2.

fprobe-ng funcionará en un interfaz capturando en modo promiscuo los datos que atraviesen dicho interfaz y formando con los datos capturados la información que luego será enviada mediante el protocolo *NetFlow* hacia el recolector. Mostramos la ficha sobre el paquete ***fprobe-ng*** en el Cuadro 4.1.

En el apéndice A.2, "*Instalación y configuración de la sonda*", se verán en detalle los aspectos de instalación y configuración de esta aplicación.

Paquete	***fprobe-ng***
Autor	Slava Astashonok < sla@0n.ru >
Versión	1.1
Mantenedor	Radu Spineanu < radu@timisoara.roedu.net >
Versión del *paquete Debian*	1.1-2
Lenguaje	C
Web	http://fprobe.sourceforge.net

Cuadro 4.1: Ficha de *fprobe-ng*

4.2.3. La sonda y el protocolo *NetFlow*

Es importante hacer una reflexión sobre el comportamiento de la sonda con respecto al protocolo *NetFlow*.

La sonda *fprobe-ng* es, funcionalmente, un *sniffer* que captura en modo promiscuo la información que atraviese la interfaz donde la propia sonda esté escuchando. La sonda, en vez de representar la información capturada de una manera gráfica o almacenarla en un archivo, mandará la información a un recolector a través del protocolo *NetFlow* usando las capacidades que este ofrece.

NetFlow, al ser ideado en principio como un protocolo de intercambio de información de encaminadores que posteriormente pasó a ser utilizado para la recogida de información estadística, está limitado en la cantidad de información que porta en su interior. *NetFlow* está orientado a conocer los volúmenes de tráfico y las cargas de los interfaces o segmentos de red, y esa es la información que *fprobe-ng* mandará hacia el recolector al usarlo. Podremos ver el formato de los mensajes de *NetFlow* en el apéndice A.1, "*NetFlow versión 7*", de este documento.

Por tanto, dado lo limitado de la información estadística que se transporta en el protocolo no es posible que en el recolector se haga un análisis a nivel de aplicación de los datos capturados por la sonda en el sistema remoto, y esta es una acción que tampoco realizará la sonda. Como se verá en el apéndice A.1 del presente documento, la información que recoge la sonda y que puede enviar sobre *NetFlow* es información propia de los niveles de Red y Transporte, no del nivel de Aplicación, por eso hablamos de una *monitorización estadística* del tráfico.

Por último, el protocolo *NetFlow* en su definición esta pensado para transportar informaciones de encaminamiento que la sonda no utilizará en ningún caso. Como hemos dicho la sonda sólo es un *sniffer* y por tanto no atenderá los campos de *NetFlow* relativos a `nexthop`, `src_as`, etc, como se explicará en el apéndice A.1 de este documento.

4.3. Recolector de *NetFlow*: *flow-tools*

4.3.1. Recolectores de *NetFlow*

Como hemos adelantado en la sección 4.2.1, "*Sondas de NetFlow*", el *recolector* de *NetFlow* es el programa que recibirá la información generada en un sistema remoto por la *sonda* y procederá a su tratamiento.

Al igual que existen diversas sondas (véase el punto 4.2.1 de este documento) existen también distintos recolectores de *NetFlow*. Algunos de ellos son:

- *cflowd*: un recolector desarrollado para recoger y analizar información de *NetFlow*. Permite al usuario almacenar la información capturada y ofrece distintas formas de visualizarla. Actualmente no se prosigue su desarrollo y sus creadores recomiendan utilizar *flow-tools* como sustituto. Para más detalle acerca de este recolector véase [13].
- *flow-tools*: es un conjunto de herramientas para el tratamiento de la información de *NetFlow*. Consta de un recolector que almacena la información comprimida en el disco duro del sistema, algunas herramientas para el tratamiento de la información, retransmisores de la información, exportadores a bases de datos y alguna otra utilidad más. Es un conjunto de utilidades configurable y con bastantes opciones, y se encuentra disponible como *paquete Debian*. Para más detalle acerca de este recolector véase [14].
- *pmacct*: un paquete de utilidades reciente y actualmente en versión no definitiva. Dispone de unas capacidades similares a las de la herramienta anterior, además de herramientas para facilitar la exportación a bases de datos *RRD*. Para más detalle acerca de este recolector véase [23].

De la lista de herramientas anteriores nosotros hemos elegido *flow-tools*, por su combinación de utilidades, sencillez, estado estable y amplia utilización, capacidades de configuración y encontrarse disponible como *paquete Debian*.

4.3.2. Introducción al paquete *flow-tools*

Nosotros, al trabajar sobre **Debian Sarge** nos hemos centrado en las utilidades ofrecidas dentro del paquete ***flow-tools*** en su versión 0.67-8, entre las que se encuentran una serie de herramientas para el estudio de la información estadística, su almacenamiento, y por supuesto su captura con la ayuda del recolector ***flow-capture***.

Mostramos la ficha del paquete ***flow-tools*** en el Cuadro 4.2.

Dentro de las muchas utilidades que contiene el paquete ***flow-tools*** trataremos ahora sólo la utilidad *flow-capture*. *flow-capture* atenderá la llegada de paquetes del protocolo *NetFlow* en nuestro sistema de monitorización y procederá a su almacenaje de manera estática y en un formato propio comprimido en el disco duro del sistema de monitorización.

Paquete	***flow-tools***
Autor	Mark Fullmer < maf@splintered.net >
Versión	0.67
Mantenedor	Anibal Monsalve Salazar < anibal@debian.org >
Versión del *paquete Debian*	0.67-8
Lenguaje	C
Web	http://www.splintered.net/sw/flow-tools

Cuadro 4.2: Ficha de *flow-tools*

En el apéndice A.3, "*Instalación y configuración del recolector*", se verán en detalle los aspectos de instalación y configuración de esta utilidad.

4.3.3. El recolector y el protocolo *NetFlow*

Al contrario que la sonda comentada en la sección 4.2.3, "*La sonda y el protocolo NetFlow*", el recolector de *NetFlow* aquí usado, *flow-capture*, sí es capaz de hacer un uso pleno de la información transportada por el protocolo.

Con esto queremos decir que aunque la sonda no introduzca determinados tipos de información en el protocolo, el recolector sí va a leer dicha información del protocolo, por lo que tendremos que filtrar con posterioridad a la captura la información que queremos que luego se procese.

El recolector recogerá la información de las sondas de la red y las almacenará de forma estática en el disco duro del sistema donde se ejecute, y podrá crear estructura de directorios que posteriormente permitirán el acceso a la información de forma cómoda y ordenada. Además, el almacenamiento de la información de *NetFlow* de forma comprimida permitirá a este recolector almacenar mayor cantidad de información que los otros recolectores vistos en el punto 4.3.1.

4.4. Consideraciones de seguridad para *NetFlow*

Hemos tratado de encontrar durante el estudio de las utilidades antes mecionadas una solución en lo que a la seguridad del transporte de la información de *NetFlow* se refiere.

Es importante garantizar la seguridad de la información de monitorización en su viaje por la red desde el sistema remoto en el que la sonda captura la información hasta el sistema de monitorización en el que el recolector recibe dicha información. *NetFlow* por sí mismo no da una seguridad a los datos, no están cifrados, y estos son capturables e interpretables por cualquier *sniffer* que esté capturando información en la red como por ejemplo *ethereal*.

Por ello hemos tratado de encontrar una solución a este transporte inseguro de la información de monitorización a través de la red, donde debemos hacer

las siguientes consideraciones:

1. *NetFlow* se transporta sobre UDP, por lo que las soluciones basadas en túneles SSL como la aplicación *stunnel* no son utilizables.
2. No existen *paquetes Debian* de utilidades para el transporte seguro de UDP, por lo que su instalación en un elevado número de sistemas puede resultar una complicación además de un gran gasto de tiempo.
3. Las herramientas existentes para el transporte seguro de UDP sobre una conexión SSL no se encuentran todavía en un estado de desarrollo finalizado. Una herramienta de estas características es *Zebedee* [24].

Por tanto, se recomienda que el transporte de la información de monitorización se produzca en una red distinta, independiente a la red de datos e inaccesible, como se muestra en la Figura 4.2. De esta manera, las máquinas en las que escuchen las sondas *fprobe-ng* deberán estar conectadas con los recolectores de información de la red mediante interfaces de red separados y sobre una estructura de red independiente a la cual no se deberá poder acceder desde la red de datos.

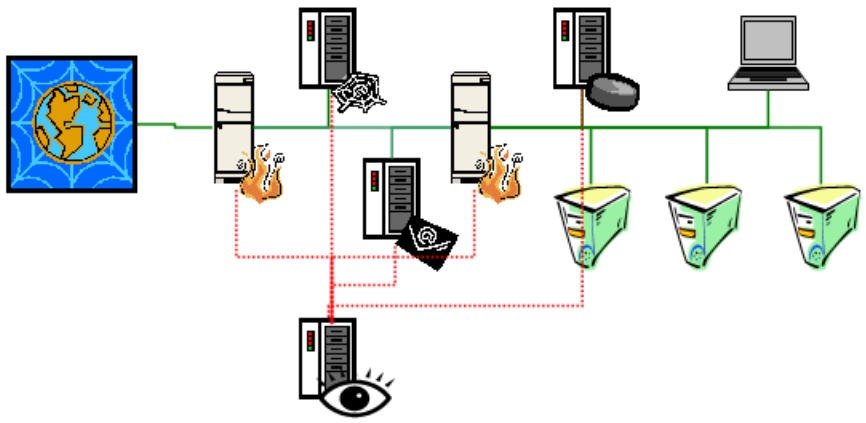

Figura 4.2: Separación de la red de datos de la red de datos de gestión