Capítulo 7

Tratamiento de la información

Índice del capítulo

7.1. Tratamiento de la información de *NetFlow*

La información que *NetFlow* ofrece no es demasiado alta, tan sólo es información de los niveles de Red y Transporte, pero lo que sí ocurre es que podemos disponer de un gran volumen de esa información si disponemos de un elevado número de sondas y recolectores en nuestra red, o bien si se trata una red con una carga alta, con una alto volumen de tráfico. En cualquiera de estos casos las sondas de *NetFlow* que estén capturando información de manera continuada estarán mandando continuos flujos de información hacia nuestros recolectores, y se hace necesario realizar una presentación de la información ya capturada de una forma eficiente y cómoda.

Dentro de esta sección nos acercaremos a tres posibles maneras de visualizar la información recogida por nuestros recolectores. El primer punto versa sobre una visualización simple en una consola, con todas las limitaciones que ello conlleva. En el segundo punto de la sección veremos cómo es posible reenviar la información ya capturada a otros recolectores con mejores capacidades de representación gráfica de la información, como por ejemplo *ntop*. Como tercera posibilidad veremos las capacidades de exportación de la información recogida a bases de datos *SQL* y *RRD*. Finalmente, en el apéndice D.1, "*Tratamiento de la información de NetFlow*", veremos los detalles más técnicos de estos aspectos.

7.1.1. Presentación mediante consola

El primero de los métodos de presentación que detalladamente se verán en el apéndice D.1, "*Tratamiento de la información de NetFlow*", será la presentación en modo consola o terminal, en función de donde apliquemos el método.

La presentación en modo consola está limitada en las posibilidades que ofrece, y citamos las siguientes ventajas e inconvenientes:

- La primera limitación es la caducidad de la información, puesto que en los sistemas de consola más comunes la información más reciente desplaza a la anterior, quedando esta imposible de acceder en períodos cortos de tiempo. Si esto se añade a la posibilidad de usar el sistema sobre una red cargada de tráfico, las sondas nos inundarán con tanta información que poco podremos ver sobre la consola que estemos usando.
- La segunda limitación, que más bien es una molestia, es que en las máquinas donde estemos almacenando la información podrá no haber modo gráfico alguno en buena parte de las ocasiones y tendremos como única opción de consola a los terminales *tty* tradicionales, muy limitados en resolución y en número de filas y de columnas. En estos terminales las líneas de información que querremos comprobar, si son extensas, estarán truncadas y resultarán en un galimatías de direcciones IP y números difíciles de entender.
- Por otro lado, la presentación en modo consola tiene la ventaja de tener una latencia nula entre la recogida de la información y su presentación si se configura adecuadamente su funcionamiento, haciendo que cada paquete de *NetFlow* que se reciba en el recolector sea enviado al terminal.

Por tanto, la presentación mediante consola es poco recomendable y está dirigida más a los administradores que necesiten hacer alguna comprobación en el sistema cuando los demás sistemas fallan o verificar que las sondas funcionan durante un proceso de instalación.

En el apéndice D.1.1, "*Presentación mediante consola*", veremos con detalle como utilizar las utilidades incluidas dentro del paquete ***flow-tools*** con estos propósitos [59].

7.1.2. Exportando a *ntop*

ntop es una conocida utilidad de monitorización de redes. Se trata de un *sniffer* en modo promiscuo que captura la información que pasa por la interfaz donde se encuentra instalado. Presenta la información recogida mediante formato *Web*, en una serie de páginas donde podemos consultar los datos en función de los *hosts*, de los protocolos, etc.

ntop dispone de una serie de *plug-ins* que le permiten funcionalidades adicionales más especializadas. Entre ellas se encuentra una *plug-in* llamada *NetFlow* que activa un recolector de *NetFlow* dentro del propio *ntop* [60], como se muestra en la Figura 7.1.

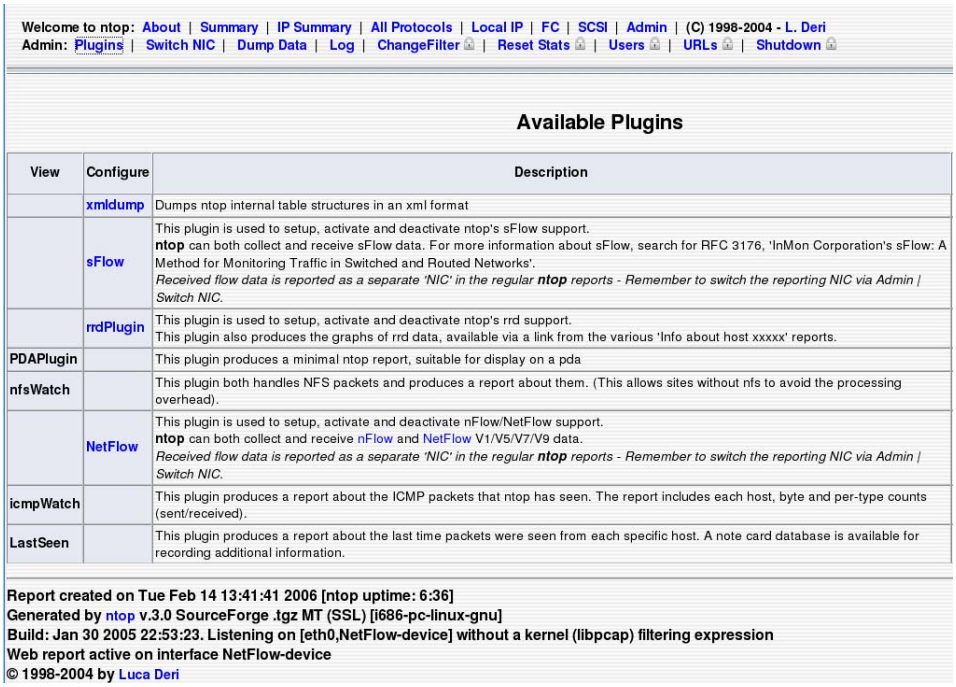

Figura 7.1: Plugin de *ntop* para activar el soporte de *NetFlow*

De esta forma, con la configuración adecuada, *ntop* puede estar funcionando en su manera habitual (como *sniffer*) y a la vez puede recoger información de una serie de sondas remotas instaladas por nuestra red. *ntop* permite diferenciar qué información proviene de la captura local y qué otra información proviene de las capturas de las sondas, por lo que se evitan los problemas de interpretaciones incorrectas.

Veremos con detalle la configuración que debemos dar al sistema para producir este efecto en el apéndice D.1.2, "*Exportando a ntop*".

7.1.3. Exportando a una base de datos

Como hemos dicho, la información de *NetFlow* puede ser no muy variada pero sí tendremos gran cantidad de registros sobre la misma. Por ello, es lógico pensar en la posibilidad o necesidad de exportar la información capturada a una base de datos.

Una posibilidad es la exportación a bases de datos *MySQL* que están ampliamente extendidas en la actualidad. Podremos elegir qué información de *NetFlow* queremos almacenar en la base de datos, y posteriormente podremos manejar la información ya almacenada en la base de datos de las múltiples maneras que se pueden hacer en la actualidad: con consultas a través de una interfaz *Web*, a través de otros programas, etc. Estudiaremos este aspecto en más detalle en el apéndice D.1.3, "*Exportando a una base de datos MySQL*", a través de la utilidad ***flow-export*** [22] incluída en el paquete de herramientas ***flow-tools***.

El problema sin embargo radica en la cantidad de información que se almacena en la base de datos, que en algunos momentos puede ser desbordante y dificultará su manejo por parte de cualquier servidor de bases de datos. Debemos pensar entonces cómo utilizar adecuadamente la base de datos. Un uso lógico es utilizarla para almacenamiento temporal de la información de gestión durante un período de seguridad, para su estudio en caso de necesidad. Esto nos permitirá acceder a la información estadística de tráfico ya almacenada si la situación actual nos impone la necesidad de estudiar el estado de la red durante un período pasado de tiempo. Un período habitual de salvaguarda de los datos es de tres días, lo que es un punto de compromiso entre el tamaño de la información almacenada y el período de tiempo capturado.

Una solución a este problema del almacenamiento y que hemos tratado en este documento es la exportación a bases de datos *Round Robin* o *RRD*, bases de datos circulares y de tamaño fijo que permiten abarcar cómodamente períodos largos de tiempo a costa de bajar la resolución de la información paulatinamente al incremento del período temporal abarcado, pero ofrecen buenas capacidades para el almacenamiento de extensos períodos de tiempo. Veremos estos aspectos con más detenimiento en el apéndice D.1.4, "*Exportando a bases de datos Round Robin*". Concretamente, veremos las capacidades de la utilidad ***FlowScan*** [58] y de algunas extensiones de la misma, ***CUFlow***, ***CUGrapher*** [66] y ***FlowMonitor*** [67] para el almacenamiento de la información de *NetFlow* y su manejo en dichas bases de datos *RRD*. No todas estas herramientas se presentan en *paquetes Debian*, y las que lo hacen ni siquiera están disponibles en el momento de realización de este Proyecto en la versión *stable*. Mostramos las fichas de los paquetes disponibles en versión *texting* en los Cuadros 7.1 y 7.2, y la información equivalente de *FlowMonitor* en el Cuadro 7.3.

Paquete	***flowscan***
Autor	Dave Plonka < plonka@doit.wisc.edu >
Versión	1.006
Mantenedor	Anibal Monsalve Salazar < anibal@debian.org >
Versión del *paquete Debian*	1.006-8
Lenguaje	Perl
Web	http://net.doit.wisc.edu/ plonka/FlowScan/

Cuadro 7.1: Ficha de *flowscan*

Paquete	***flowscan-cuflow*** y ***flowscan-cugrapher***
Autor	Johan Andersen < johan@columbia.edu > y Matt Selsky < selsky@columbia.edu >
Versión	1.7
Mantenedor	Russell Stuart < russell-debian@stuart.id.au >
Versión del *paquete Debian*	1.7-1
Lenguaje	Perl
Web	http://www.columbia.edu/acis/networks/advanced/CUFlow/

Cuadro 7.2: Ficha de *flowscan-cuflow* y *flowscan-cugrapher*

Utilidad	***flowscan-flowmonitor***
Autor	Johan Andersen < johan@columbia.edu >
Versión	1.2
Lenguaje	Perl
Web	http://www.columbia.edu/acis/networks/advanced/FlowMonitor/FlowMonitor.html

Cuadro 7.3: Ficha de *flowscan-flowmonitor*

7.2. Tratamiento de la información de los registros

7.2.1. Visualización en consola

Similar a lo que ocurre con la presentación de la información de *NetFlow* mediante consola, como se indicó en el punto 7.1.1, la presentación en modo consola es un sistema algo tosco y primitivo. Presentará unas ventajas e inconvenientes similares a las vistas con anterioridad:

- Con respecto a la limitación de la caducidad de la información, la diferencia con la presentación de la información de *NetFlow* radica en la velocidad con la que esta se genera, que será muy inferior a la de *NetFlow*. Esto hará que la visualización en modo consola pueda tener mejores resultados que la visualización que se obtuvo con *NetFlow*.

- En cuanto a la segunda limitación mencionada, acerca de la limitación en resolución y en número de filas y columnas de los terminales sin modo gráfico, el problema también se ve menos acusado dada la naturaleza del contenido de la información. Las líneas más legibles y cortas de los mensajes del sistema tendrán una mejor cabida en los terminales que la información de *NetFlow*.

- Igual que en el caso anterior, la presentación en modo consola tiene la ventaja de tener una latencia nula entre la recogida de los mensajes y su

presentación si se configura adecuadamente su funcionamiento, obteniendo en el terminal y «en tiempo real» los mensajes del sistema.

La representación de los registros en modo consola está ideada para que quede de manera permanente en un terminal, reflejando los informes de los registros para que un administrador los pueda ver al momento de su generación. Esto permite la más rápida actuación por parte de los administradores para solventar los problemas que se generen especialmente en las fases de instalación y configuración de las aplicaciones *syslog-ng* o cualquier otra utilidad que vaya a afectar a los registros del sistema.

7.2.2. Visualización mediante página *Web*

La visualización de los registros del sistema mediante página *Web* es muy similar a la presentación en modo consola, si bien ofrece una serie de ventajas y beneficios que la hacen más interesante:

- En primer lugar, si la información se ofrece mediante una página *Web* esta puede ser accedida desde múltiples lugares, desde todos aquellos desde los que se pueda acceder al servidor que la aloje.
- En segundo lugar, soluciona los problemas de visualización caduca y limitada en resolución de la consola, al ser dependiente ahora de la configuración del cliente.

Por otra parte plantea también unos nuevos problemas:

- Como contrapartida a la primera ventaja, será necesario establecer una política de seguridad en nuestro servidor que impida a usuarios no adecuados acceder a la información, dado que ahora será accesible desde todos los puntos desde los cuales se pueda acceder al servidor.
- Referente a la segunda, implica la necesidad de máquinas con entorno gráfico, lo cual hace pensar en la necesidad de más máquinas para proceder a la monitorización si no queremos mezclar máquinas con entorno gráfico con máquinas sin el mismo.

Para resolver este punto y el anterior (7.2.1) hemos utilizado la utilidad ***ccze*** [68]. *ccze* es un *colorizador* de registros como algunas opciones adicionales como son su generación de código *HTML*, por lo que además de permitir una visualización cómoda en consola es capaz de hacer una exportación idéntica a una página *Web*. ***ccze*** se distribuye también como *paquete Debian*, y mostramos su ficha en el Cuadro 7.4.

Paquete	***ccze***
Autor	Gergely Nagy < algernon@bonehunter.rulez.org >
Versión	0.2.1
Mantenedor	Gergely Nagy < algernon@bonehunter.rulez.org >
Versión del *paquete Debian*	0.2.1-1
Lenguaje	C
Web	http://bonehunter.rulez.org/CCZE.html

Cuadro 7.4: Ficha de *ccze*