CAPÍTULO 1

ANTECEDENTES

Este capítulo tiene como objetivo abordar los conceptos necesarios relacionados con los sistemas de detección de tráfico malicioso, monitoreo de redes y análisis de tráfico de red en general. Éstos son necesarios para contextualizar el desarrollo del proyecto presentado en esta tesis, ya que como se verá más adelante, el esquema de funcionamiento general del motor de detección de tráfico malicioso toma como base a determinadas herramientas y técnicas vistas en este primer apartado.

1.1 MECANISMOS PARA DETECCIÓN DE TRÁFICO DE RED MALICIOSO

La actividad del flujo de datos a través de la red puede ser monitoreada y medida de diversas maneras. Este monitoreo incluye la clasificación del tipo de tráfico en cuanto a su naturaleza y es importante porque muchas de las funciones vitales de la red pueden estar en juego cuando una amenaza potencial se transforma en un incidente.

Se define como "tráfico de red malicioso" a cualquier tipo de tráfico que sea originado por algún agente o entidad maliciosa ya sea lógica (malware) o real (humanware) y que tenga la finalidad de lograr algún objetivo que atente contra alguno de los pilares de la seguridad informática: disponibilidad, integridad o confidencialidad.

A pesar de que el origen del tráfico puede no ser realmente malicioso (por ejemplo usuarios incautos), a final de cuentas se convertirá en tráfico malicioso "accidental" cuando de manera intencional o no, atente contra los pilares de la seguridad.

La importancia de poder identificar y detectar el tráfico malicioso se justifica con el hecho de que este tipo de tráfico es el que puede alterar el funcionamiento de una red o, en el peor de los casos, causar tal impacto que interrumpa por completo la actividad general del entorno. El SANS Institute[1] a través del Internet Storm Center (ISC), cuenta con un cálculo denominado "Survival Time"[2] el cual consiste en medir el tiempo promedio que tarda un equipo de cómputo en ser atacado o alcanzado por algún tipo de malware en propagación, considerando que se expone a una red pública sin

[1] Organización internacional dedicada a la investigación, capacitación y publicación de recursos relacionados con seguridad en cómputo.
[2] http://isc.sans.edu/survivaltime.html

restricciones. Si se diera el caso que el equipo no contara con los parches adecuados, entonces esta medición significaría el tiempo en que el equipo sería infectado o vulnerado. En la figura 1 se puede observar que las últimas mediciones indican que el "survival time" de un equipo Unix es de aproximadamente 3700 minutos, mientras que el de un equipo Windows es casi de 450 minutos. Esto nos da una muestra del verdadero problema con el tráfico de red malicioso.

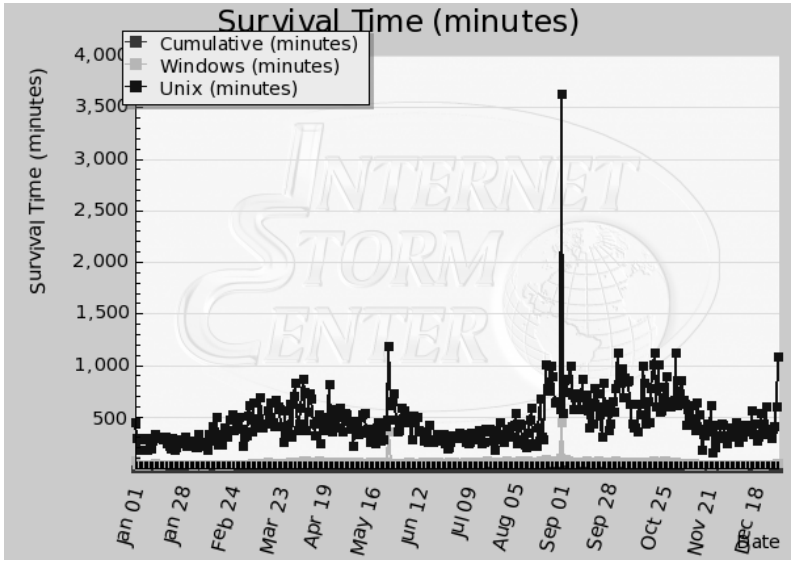

Figura 1. SANS ISC-Survival time (enero 2005 - agosto de 2010)

¿Cómo saber si determinado tráfico es anormal? o ¿cómo definir si el tráfico monitoreado es malicioso? Este tipo de cuestiones pueden ser respondidas abordando distintos puntos de vista. Por un lado se tiene la definición de políticas de una organización en las cuales se puede establecer lo que se considera como anormal y por otro lado se tiene el comportamiento general de determinadas amenazas que definen muy bien a un evento como un incidente de seguridad. Tomando en cuenta lo anterior, un ejemplo es que quizá para una organización el utilizar software P2P no representa mayor amenaza, sin embargo, las consecuencias indirectas de este hecho sí pueden ocasionar cambios en el entorno tales como afectación en la velocidad de la red, la exposición a las redes P2P potencialmente dañinas o fuga de información. Lo anterior demuestra que también es importante identificar la diferencia entre tráfico malicioso y tráfico anómalo.

El tráfico anómalo se define de acuerdo a las políticas o consideraciones de una organización, por lo que su existencia en una red será considerada de manera diferente independientemente de que pueda ser o no malicioso. Entonces, en los términos de la seguridad informática, podemos tratar de una manera más específica al tráfico malicioso relacionándolo con amenazas bien definidas que según su naturaleza puede tratarse de algún agente de malware o de una técnica o mecanismo especializado.

Algunos métodos para detectar "tráfico anormal" se basan en la comparación y análisis del comportamiento "esperado" de cierto tipo de protocolos o aplicaciones. Es decir, si se tiene bien identificada la estructura, forma y comportamiento del tráfico según su naturaleza, entonces cualquier patrón fuera de ella representa un factor para poder identificarlo como anormal. Algunos mecanismos para poder lograr dicha identificación y detección son las técnicas basadas en análisis de patrones, las herramientas especializadas y el análisis de bitácoras (log analysis). Ejemplos de algunos de ellos y que se abordarán a manera de antecedentes necesarios son:

- Firewalls
- Sistemas de detección de intrusos (IDS)
- Sistemas de prevención de intrusos (IPS)
- Analizadores de paquetes
- Analizadores de flujos
- Bitácoras de aplicaciones
- Telescopios de red o de seguridad

1.1.1 FIREWALLS

Básicamente, un firewall es un dispositivo físico o lógico que brinda la capacidad de controlar el flujo del tráfico de red a través de un sistema, es decir, funciona como un mecanismo de control de acceso a diferentes niveles. Estos dispositivos permiten controlar mediante políticas o reglas establecidas todo lo que fluye en el sistema tomando en cuenta factores como dirección, origen, destino, protocolos, aplicaciones,

etc. y brindan la capacidad de mantener aislado, hasta cierto punto, el sistema de potencial tráfico malicioso.

Fundamentalmente, un firewall necesita tener la capacidad de realizar las siguientes tareas:

- Administrar y controlar el tráfico de red
- Control de acceso
- Actuar como intermediario
- Proteger los recursos
- Registrar y reportar los eventos

La figura 2 muestra diversos tipos de firewalls a partir de los cuales se definen varias clasificaciones. Algunas de ellas son:

- Por modo de funcionamiento o tecnología
    - Firewall de filtrado de paquetes
    - Firewall de estado
    - Firewall de aplicación
    - Servidor proxy
- Por esquema de funcionamiento
    - Firewall de host
    - Firewall de red

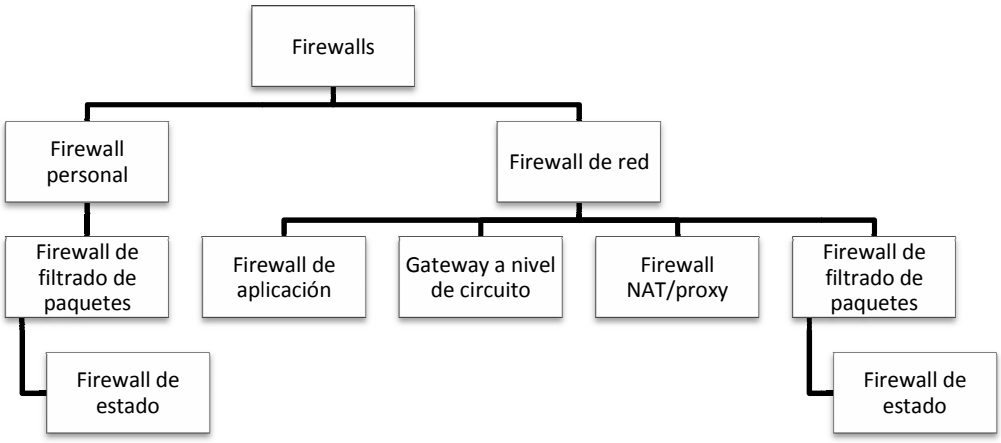

Figura 2. Tipos de firewalls

La principal diferencia entre un firewall de host y uno de red es que este último tiene la capacidad de analizar el tráfico de varios equipos en un entorno, a diferencia del de host que solo abarcará el tráfico que entra y sale de un equipo sencillo.

1.1.2 SISTEMAS DE DETECCIÓN DE INTRUSOS (IDS)

Los sistemas de detección de intrusos son dispositivos físicos o lógicos que permiten analizar el tráfico de red para identificar posibles paquetes maliciosos o anómalos. Son un mecanismo pasivo de detección ya que su tarea fundamental es alertar pero no actuar, sin embargo pueden proveer de información muy específica sobre la actividad detectada debido a que el análisis de un IDS sobre los paquetes de red es más específico que el que realiza un firewall. Esto convierte a los IDS en herramientas muy poderosas para conocer el panorama de la actividad en la red.

El funcionamiento general de los IDS se basa en detectar tráfico malicioso mediante firmas o anomalías.

La detección por firmas consiste en la definición de un patrón con características específicas las cuales comúnmente se basan en patrones de amenazas conocidas. Las firmas contienen características como tipo de tráfico, dirección de flujo, protocolo, direcciones IP, puertos o incluso el contenido de datos en el paquete.

Tabla 1. Ejemplo de firmas del IDS Snort

alert tcp $EXTERNAL_NET any -> $HOME_NET 22 (msg:"SCAN SSH Version map attempt"; flow:to_server,established; content:"Version_Mapper";fast_pattern:only; classtype:network-scan; sid:1638; rev:6;)
alert tcp $EXTERNAL_NET any -> $HOME_NET any (msg:"SCAN SYN FIN"; flow:stateless; flags:SF,12; reference:arachnids,198; classtype:attempted-recon; sid:624; rev:8;)

En la Tabla 1 se pueden observar dos firmas, la primera define la alerta “SCAN SSH Version map attemp” cuando el protocolo es TCP y va hacia cualquier equipo de la red interna hacia el puerto 22, además debe establecerse la conexión y en el contenido del paquete debe contener la cadena “Version Mapper”. Cuando un paquete de red

coincida con este patrón, entonces se levantará la alerta proporcionando la información relacionada como marca de tiempo (timestamp), direcciones IP y puertos origen y destino, etc. Los desarrolladores de IDS comúnmente liberan nuevas firmas para poder detectar nuevas amenazas.

Por otro lado está la detección basada en anomalías. Este esquema funciona definiendo ciertos criterios base o *baselines* que suponen un funcionamiento normal del sistema o de la red. Cuando se detecta cierta actividad que no corresponde con el baseline en un cierto rango, entonces el IDS lo puede interpretar como una anomalía y en su caso identificarlo como tráfico malicioso. Al igual que las firmas, los baselines deben estar actualizados para poder ser más eficaces y detectar nuevas posibles amenazas.

Con los IDS existen dos problemas principales: los falsos positivos y los falsos negativos. Los primeros se refieren a todos los eventos levantados como alertas pero que en realidad no se trataban de tráfico malicioso. Esto puede ser causado por varios motivos pudiendo ser que alguna situación o aplicación llevaron a generar el paquete con las características de la firma y un aspecto muy fuerte a tomar en cuenta es que mientras más general se defina la regla, más falsos positivos se pueden tener.

El segundo problema es la generación de falsos negativos. Estos consisten en todos los eventos que a pesar de presentarse en la red no son alertados por el IDS. En realidad este tipo de omisiones puede implicar un mayor riesgo puesto que desde cierto punto de vista es más conveniente detectar algo que no existe, que no detectar algo que en verdad existe y que es malicioso, aunque debe tenerse presente que también contar con un número significativo de falsos positivos puede llegar a ser contraproducente. La detección y manejo de falsos positivos es un área de estudio específica en el campo de los IDS.

Existen varios tipos de IDS:

- IDS basados en host

Funciona monitoreando la actividad de un sistema local. Por su esquema de funcionamiento (figura 3), analiza el tráfico de red que entra y sale de dicho equipo, así como los cambios en el sistema de archivo y actividad del sistema en general.

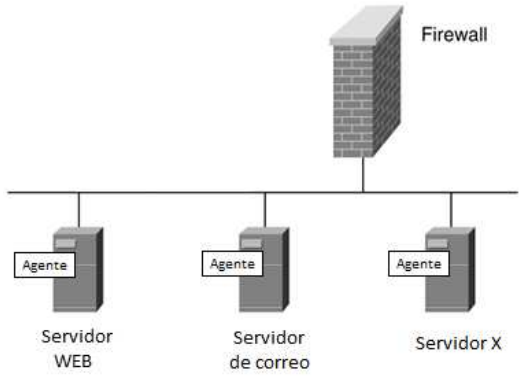

Figura 3. Esquema de un IDS de host

- IDS basados en red

Analiza el tráfico de un equipo o red. Puede instalarse en un equipo analizando sólo el tráfico que fluye a través de él, sin embargo, para que funcione ampliamente debe implementarse un esquema en donde reciba el tráfico de todos los equipos conectados a la red (figura 4), comúnmente llamado "port mirror". Generalmente se instala en el perímetro de la red o subred para poder monitorear el tráfico de entrada y salida de la misma. El éxito en su funcionamiento depende de su correcta ubicación.

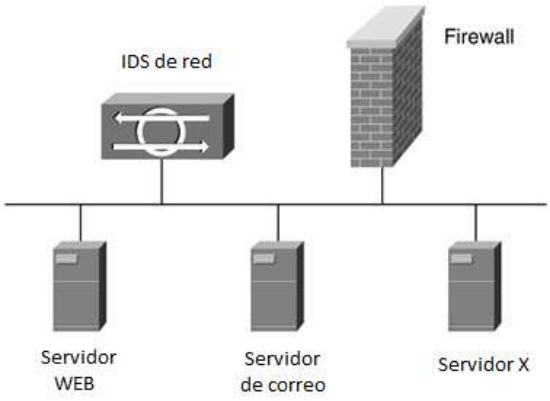

Figura 4. Esquema de un IDS de red

- IDS distribuidos

Es un esquema de varios IDS desplegados a lo largo de una red los cuales centralizan la información (figura 5). Este tipo de esquemas puede ser útil en redes de gran tamaño, sin embargo debido a la gran cantidad de información que implica, necesita un monitoreo y mantenimiento constante.

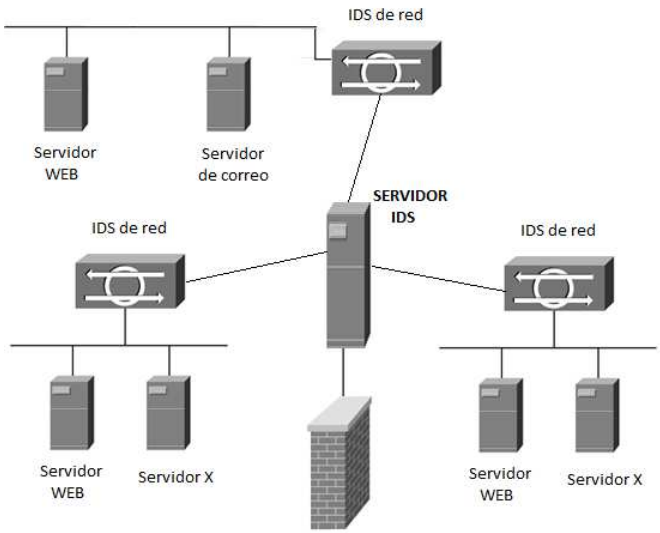

Figura 5. Esquema de un IDS distribuido

Los factores clave para el buen funcionamiento de los IDS son entonces: ubicación, buenas firmas, baselines y mantenimiento.

1.1.3 SISTEMAS DE PREVENCIÓN DE INTRUSOS

Los sistemas de prevención de intrusos o IPS son, al igual que los IDS, mecanismos físicos o lógicos para la detección de tráfico malicioso basándose en firmas o anomalías. La principal diferencia con los IDS es que los sistemas de prevención de intrusos son dispositivos activos que tienen la característica de actuar bajo demanda según las alertas detectadas, a esto se le conoce como "inline". Lo anterior significa que a partir de que un evento es detectado, el sistema puede aplicar automáticamente una medida de mitigación lo cual implica que los IPS tengan capacidades de firewall.

Por estas características, a este tipo de dispositivos también se les conoce como IDP o sistemas de detección y prevención de intrusos.

Por ejemplo, si un IDS alerta sobre un escaneo o ataque de fuerza bruta al servicio de SSH generalmente configurado en el puerto TCP/22, entonces solo alertará. Por su parte, el IPS adicionalmente creará una regla en su sistema de filtrado para bloquear los paquetes provenientes del origen del ataque o alguna otra medida configurada previamente.

Con este esquema podría pensarse que los IPS son mejores que los IDS, o incluso que podrían sustituir a los firewalls. Desde hace algunos años, firmas internacionales de seguridad y otras organizaciones de desarrollo e investigación, han abierto un gran debate sobre si los IDS han quedado obsoletos o representan una tecnología que será sustituida por los IPS. Gartner, organización mundialmente conocida dedicada a la tecnología y negocios, publicó en 2003 una declaración donde afirmaba "Intrusion Detection Systems a Market Failure"[3] mientras que otras opiniones expresaban lo contrario. El argumento era que los IPS no podría aún representar la nueva generación de dispositivos de detección debido a que aún eran inmaduros. Lo que es evidente es que aún existen consideraciones que deben tomarse en cuenta para la implementación de un IPS.

Primeramente está el problema de los falsos positivos. Con los IDS este problema implica solamente realizar un proceso de discriminación de alertas, el cual en el peor de los casos representa un trabajo adicional y excesivo para el administrador de la red. En cambio, con un IPS un falso positivo podría representar una auto-negación de servicio o problema general con la red puesto que de manera automática aplicaría las reglas en su firewall interno para poder mitigar "el ataque". Es por eso que las firmas de detección deben definirse con el criterio más acertado posible para minimizar el número de falsos positivos.

Debido a su naturaleza de mecanismo activo, un IPS debe:

[3] Gartner Information Security Hype Cycle Declares Intrusion Detection Systems a Market Failure. http://www.gartner.com/5_about/press_releases/pr11june2003c.jsp

- Dar respuesta inmediata a un evento
- Mantener el estado de las conexiones
- Tener conocimiento del protocolo o comportamiento de la aplicación

Existen dos tipos de IPS:

- NIPS – Network IPS: Monitorear tráfico de varios equipos en la red.
- HIPS – Host IPS: Monitorean el tráfico y actividad del equipo local.

1.1.4 ANALIZADORES DE PROTOCOLOS

Este tipo de herramientas también se conoce como analizadores de paquetes o "sniffers". Su principal objetivo es proveer la capacidad de capturar los paquetes que circulan por la red para poder analizarlos.

En ocasiones son considerados como "armas de doble filo" debido a que pueden ser utilizados con fines maliciosos, de manera que puedan capturar tráfico y obtener información privada atentando contra la confidencialidad de la información, sin embargo, su función consiste en un análisis a fondo del tráfico de la red.

Para que estas herramientas puedan ser utilizadas, las interfaces o dispositivos de captura son configuradas en modo "promiscuo" el cual consiste en activar la capacidad de poder recibir tráfico de red incluso si no va dirigido a dicha interfaz.

El utilizar este tipo de herramientas también implica la necesidad de mayor experiencia y conocimiento a fondo para un verdadero aprovechamiento, esto se debe a que un analizador de paquetes puede ser explotado solamente si se tienen los conocimientos adecuados de análisis de datagramas, es decir, conocer las técnicas y teoría necesaria para poder interpretar el contenido "crudo" que consiste en datos con formato hexadecimal. Entonces, la tarea fundamental es mapear la información a partir de los modelos de referencia de cada protocolo para saber qué datos del paquete corresponden a cada campo de la cabecera de determinado protocolo.

La relevancia de los analizadores de paquetes está en poder distinguir el tráfico potencialmente malicioso a partir de la definición de expresiones de filtrado o mediante procesamiento cuyo fin es agrupar información estadística e interpretarla. Por ejemplo, al capturarse todo el tráfico de una organización, puede notarse que existe tráfico excesivo hacia un determinado puerto o dirección IP, o que existe tráfico de servicios o aplicaciones no permitidas o utilizadas en la organización. El punto esencial es que el administrador de la red tenga la capacidad de interpretar estos patrones y poder indagar en algo más específico hasta poder detectar una amenaza o anomalía.

En concepto, muchos de los sistemas IDS, IPS o incluso firewalls tienen este tipo de herramientas como motor de captura de tráfico de red a un nivel más bajo, lo cual les permite analizarlo en tiempo real, sin embargo la diferencia radica en que generalmente solo almacenan los paquetes de red por los cuales se detectaron determinadas alertas, es decir, los paquetes que coinciden con las firmas o anomalías definidas por el IDS o IPS.

Varias de las herramientas más conocidas están basadas en la librería libpcap[4] y han sido portadas a varios sistemas operativos. Algunos ejemplos son Tcpdump, Windump, Wireshark, Snort, Tshark, entre otras.

Un punto muy importante a tomar en cuenta es que debido a que estas herramientas pueden capturar todo el tráfico de la red, el equipo en donde se instala debe contar con las características necesarias en cuanto a procesamiento y sobre todo, en capacidad de almacenamiento el cual depende totalmente de la cantidad de tráfico que circula en la red.

[4] Packet Capture Library http://www.tcpdump.org/pcap3_man.html

1.1.5 ANÁLISIS DE FLUJOS

Un flujo está definido como un grupo unidireccional de paquetes que comparten las siguientes características generales:

- Protocolo
- Dirección IP origen
- Puerto origen
- Dirección IP destino
- Puerto destino

El análisis de flujos es una técnica para detección de tráfico malicioso y monitoreo de red muy eficaz que contiene información de la capa de red y de transporte (modelo de referencia OSI). A diferencia del análisis de paquetes o la detección por medio de un IDS/IPS, este tipo de análisis no se basa en la definición de firmas o anomalías, sino que se encarga de estudiar los patrones en los flujos de la red tomando en cuenta métricas como las que se listan a continuación:

- Protocolo
- Direcciones IP: Origen y destino
- Marcas de tiempo (timestamp): inicio y final
- Conteo y estadísticas de paquetes
- Conteo y estadísticas de bytes
- TTL (para paquetes IP)
- Banderas TCP (para tráfico TCP/IP)
- Entre otras

En la tabla 2 se muestra un ejemplo de flujo de red donde se puede apreciar las características mencionadas y cuyos datos son suficientes para generar estadísticas importantes sobre el tráfico de red:

Tabla 2. Ejemplos de flujos de red

TOTAL PAQUETES	TOTAL BYTES	PROTOCOLO	IP ORIGEN	PUERTO ORIGEN	IP DESTINO	PUERTO DESTINO	TIMESTAMP
355	2312	TCP	1.1.1.1	34531	2.2.2.2	22	1282700162
1231	12223	TCP	3.3.3.3	139	4.4.4.4	139	1282700398
313	987	UDP	5.5.5.5	32212	6.6.6.6	21	1282701091

El objetivo de analizar este conjunto de datos busca lograr tareas como:

- Identificar y cuantificar las sesiones de red o conjuntos de sesiones que parezcan sospechosas.
- Obtener un monitoreo general de la actividad de la red mediante estadísticas del tráfico de red incluyendo conteos por protocolo, puertos, direcciones, estados, consumo de recursos, transferencias, etc.
- Facilitar la detección de patrones de tráfico maliciosos como accesos no autorizados, anomalías en servicios, propagación de malware, escaneos, tráfico en segmentos no utilizados, etc.
- Evitar el análisis del contenido completo del tráfico de red por lo que este procedimiento es más fácil de realizar e implica menor cantidad de recursos de almacenamiento.
- En caso de un incidente, mediante una auditoria se puede obtener una fuente importante de evidencia para reconstruir los hechos mediante logs.

Debido a que el análisis de flujos de red es más común en entornos de mediano o gran tamaño, la consideración principal es la cantidad de datos que se recolectan, no obstante, al poder colapsar en sesiones a todos los paquetes de red que compartan la información de protocolo y direccionamiento omitiendo el almacenamiento del contenido de todo el tráfico de red, lo convierte en un método muy eficaz para detección de tráfico malicioso que facilita mantener gran cantidad de información por largos periodos de tiempo haciendo muy conveniente su manejo mediante bases de datos y su explotación a través de técnicas de minería de datos. La figura 6 muestra un

ejemplo de la información de las sesiones de los paquetes de red y sus parámetros relacionados.

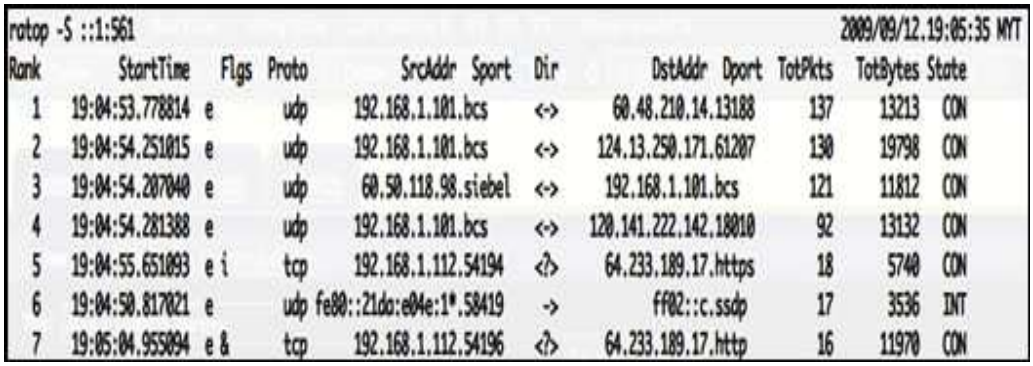

Figura 6. Información proporcionada por el cliente de argus (ratop).

Uno de los pioneros en el desarrollo y utilización de esta tecnología de análisis fue Cisco el cual desde hace algunos años ha agregado esta funcionalidad en sus dispositivos de enrutamiento. Sin embargo, también se han desarrollado tecnologías similares por otras compañías y a su vez, se han creado varias herramientas, formatos y protocolos que implementan el análisis de flujos. La tabla 3 muestra una relación de los diferentes formatos para análisis de flujos desarrollados por algunas compañías:

Tabla 3. Formatos de flujos de tráfico de red

COMPAÑIA	FORMATO FLUJO
Cisco Systems	NetFlow
Extreme Networks	CLEAR-Flow and sFlow
Foundry Networks	sFlow
Huawei Technology	NetStream
Juniper Networks	J-Flow, cFlowd (NetFlow v5)
Bluecoat Systems	Packeteer-2

1.1.6 ANÁLISIS DE BITÁCORAS

El análisis de bitácoras es todo un campo de estudio no solo en la seguridad informática, sino dentro de todas las tecnologías de la información. Es tal su

importancia que muchas de los nuevos mecanismos de monitoreo se han desarrollado basados en técnicas de análisis de bitácoras.

Su mayor importancia está en poder interpretar información almacenada de bitácoras de un sistema para poder identificar algún evento o situación específica. Abarca la utilización de varias herramientas como fuentes de datos y de otras para recopilación, procesamiento y explotación de la información por lo que se entiende como un mecanismo de detección pasivo posterior a la actividad detectada, compartiendo varios aspectos de un análisis forense.

Considerando el campo de la seguridad informática y específicamente la parte de detección de tráfico malicioso o monitoreo de redes, se puede obtener gran cantidad de datos a partir de fuentes como:

- Firewalls
- IDS/IPS
- Analizadores de tráfico y de flujos
- Sistemas de enrutamiento o conectividad (switches, routers, etc)
- Bitácoras del sistema operativo
- Bitácoras de aplicaciones de monitoreo
- Bitácoras de aplicaciones de propósito específico

Esto implica la utilización de algoritmos, técnicas o métodos de ordenamiento, clasificación y almacenamiento de la información para facilitar su procesamiento.

Debido a la gran cantidad de información que puede ser generada según el tamaño de la red o del nivel de detalle de las bitácoras, es conveniente dividir los datos fuente y preprocesarlos para una mayor eficiencia, sin embargo estas tareas pueden resultar demasiado complicadas si no se cuenta con las herramientas necesarias o las habilidades suficientes para poder manejar la información.

Saber si la red sufrió un ataque, conocer las estadísticas sobre el flujo de tráfico de red, detectar si un equipo o usuario específico intenta recurrentemente acceder a un

recurso no autorizado o incluso si se han logrado dichos accesos, entre otras actividades, son eventos que pueden ser detectados mediante el análisis de bitácoras.

Para definir un mecanismo de análisis de bitácoras es necesario tomar en cuenta:

- Cantidad de datos fuente
- Recursos de procesamiento (equipos, aplicaciones, etc.)
- Tiempos de respuesta

Una vez considerado esto, se debe definir de la manera más clara posible:

- Información esperada
- Formatos de salida

Para ello, existen diversas herramientas que permiten procesar datos:

- Herramientas nativas de los sistemas operativos (basados en UNIX/GNU Linux): grep, sort, awk, etc.
- Lenguajes de programación (scripting): Perl, Python, PHP, etc.
- Software especializado

De cualquier manera, una de las principales consideraciones en el análisis de logs será la cantidad de datos, puesto que implica procesamiento, lo cual a su vez se traduce en recursos y tiempo. El Dr. Anton Chuvakin[5] publicó en su blog[6] una medición trivial pero muy clara de la cantidad de recursos de almacenamiento que puede implicar un análisis de bitácoras:

> *100,000 log messages / second x 300 bytes / log message ~ 28.6 MB*
> *x 3600 seconds ~ 100.6 GB / hour*
> *x 24 hours ~ 2.35 TB / day*
> *x 365 days ~ 860.5 TB / year*
> *x 3 years ~ 2.52 PB*
>
> *Oops! Now you know what is a petabyte.*
> *And, BTW, you also now what is a trillion – of log messages.*

[5] Reconocido experto en el campo de la seguridad y análisis de logs, autor de los libros "Security Warrior", "PCI Compliance" y contribuidor de "Know Your Enemy II", "Information Security Management Handbook" y una docena de artículos sobre "log management", SIEM, seguridad y análisis de datos, administración de la seguridad, entre otros.

[6] http://chuvakin.blogspot.com/

Además, dependiendo de las necesidades de tiempo de respuesta ante un evento dado, pueden generarse problemas en la capacidad de procesar los datos cuando se llega a niveles de varios GB al día, no obstante, la correcta aplicación de técnicas y soluciones de análisis de bitácoras permitirá detectar, clasificar, almacenar y obtener interpretaciones sobre la información relacionada con la posible actividad maliciosa en el entorno de la red.

1.2 MECANISMOS ALTERNATIVOS PARA LA DETECCIÓN DE TRÁFICO MALICIOSO

Adicionalmente a los métodos comunes para detección de tráfico malicioso, desde hace varios años han surgido tecnologías con un enfoque de análisis más profundo. Algunas de estas tecnologías se basan en el concepto de emulación de servicios y redes y su objetivo principal es recibir, bajo un ambiente controlado, todo el tráfico malicioso que sea posible para ser almacenado y estudiado. Aunque existen distintas variantes de métodos alternativos de detección de tráfico malicioso, aquí se abordarán los temas relacionados con la propuesta original de la tesis.

1.2.1 TECNOLOGÍAS HONEYPOT

Este tipo de tecnologías surgieron a finales de los 90's bajo el concepto de sistemas honeypot. Su objetivo primordial es recibir el tráfico malicioso de manera intencional y poder interactuar con él. En el contexto más general, un honeypot es un equipo señuelo que se instala en un punto de la red para poder recibir, y por lo tanto detectar, tráfico malicioso o patrones relacionados. Una de las referencias más importantes en este campo es el proyecto "The Honeynet Project"[7], organización internacional fundada por Lance Spitzner en 1999 cuya actividad fundamental es la investigación en el campo de las tecnologías honeypots y las amenazas a la seguridad de la información. Fue Lance Spitzner quién introdujo por primera vez el término honeynet el cual define como una red de honeypots de alta interacción que simulan una red en

[7] Portal web del proyecto internacional http://www.honeynet.unam.mx

producción, configurada de manera que la actividad sea monitoreada, registrada y discretamente regulada.

Conforme han pasado los años, las características de los honeypots han cambiado pero el concepto inicial sigue siendo el mismo. Pueden clasificarse por su entorno y por su funcionamiento.

Por su entorno se dividen en:

- Honeypots de producción: Son equipos situados en el entorno real de la red junto con otros equipos en producción como servidores. Comúnmente son de baja interacción y funcionan como un complemento a la detección de amenazas, mitigando de alguna manera los riesgos en la actividad de la red.
- Honeypots de investigación: Son equipos instalados con el solo propósito de estudiar el comportamiento y las tendencias del tráfico malicioso causado por intrusos o ataques automatizados, situados en entornos exclusivos de prueba y generalmente con fines académicos. Tienen la característica de capturar y analizar la información por lo que su configuración y administración puede ser más compleja.

Por su modo de funcionamiento se dividen en:

- Honeypots de baja interacción: Son equipos cuya característica principal es emular los servicios de un sistema real para poder interactuar lo suficiente con los intrusos o amenazas automatizadas y recopilar datos. Al tratarse de una emulación, la efectividad y cantidad de información recopilada depende de la complejidad con la que el honeypot interactúe. Esto implica un menor riesgo pero los hace detectables más fácilmente. Son muy eficientes para la detección de patrones maliciosos, captura de malware y generación de estadísticas de tráfico en general.
- Honeypots de alta interacción: Son equipos reales, es decir, no implementan la emulación de ningún servicio sino que la interacción se da realmente por el software instalado en el sistema. Esto permite que los intrusos tengan una

interacción y control real del sistema señuelo lo cual implica un mayor riesgo y por lo tanto la necesidad de un sistema de control externo que permita monitorear, almacenar y procesar la información capturada. Este tipo de honeypot puede implementarse junto con mecanismos capaces de obtener información muy detallada de la actividad que el intruso realiza en el sistema (por ejemplo es posible monitorear los comandos que escribe en una terminal), por lo que de manera general el enfoque es más para un ámbito de investigación de las técnicas y tendencias de ataques de seguridad en cómputo.

La única diferencia entre una red de honeypots y una honeynet estrictamente se refiere a que las honeynets están formadas por uno o más honeypots de alta interacción desplegados bajo un esquema de control y análisis de tráfico de red como se observa en la figura 7.

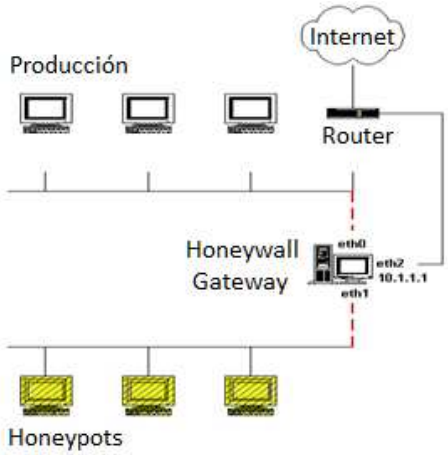

Figura 7. Esquema de conectividad de una honeynet

Las honeynets han evolucionado durante los últimos años implementando nuevas características, principalmente enfocada en el manejo y control de la información. Los tipos de honeynets que existen pueden englobarse en:

- Honeynets Generación I (1999)
    - Arquitectura inicial y básica de control y captura de datos

- Honeynets Generación II (2002)
    - Arquitectura mejorada en el control y captura de los datos: bridge capa 2, filtrado, sistema de alertas, módulo sebek v2.x.
- Honeynets Generación III (2005)
    - Arquitectura con análisis de datos
- Honeynets distribuidas
    - Honeynets desplegadas a lo largo de Internet
- Honeynets virtuales
    - Honeynets desplegadas en un mismo equipo

Cada módulo implementado en una honeynet cumple funciones específicas. La evolución de cada uno de ellos se ha dado según la necesidad de un mejor manejo y control de los datos. Las honeynets actuales permiten obtener información muy detallada de las actividades de los honeypots y de los aspectos del tráfico malicioso en ellos. La tabla 4 muestra las herramientas utilizadas en cada subsistema del Honeywall:

Tabla 4. Subsistemas del honeywall

MÓDULO	CARACTERÍSTICAS
Control de datos	Iptables – Firewall Rate-limitting – Límites de conexiones Snort-inline – Tecnologías IPS
Captura de datos	Logs de iptables – Información de firewall Alertas de IDS – Información de Snort p0f – Identificación pasiva de S.O. Sebek – Captura avanzada de datos Tcpdump – Captura de tráfico
Análisis de datos	MySQL - Almacenamiento de información en BD Argus y Hflow – Análisis de flujos Swatch – Análisis de logs de alertas del IDS y firewall Walleye – Interfaz gráfica Web

Con el avance de las características de emulación de los honeypots de baja interacción y los sistemas de administración como honeywall, la utilización de tecnologías honeypot ha permitido obtener buenos resultados y sobre todo sigue siendo un campo amplio en la investigación, tanto para fines académicos como para la industria. En la UNAM existe desde hace varios años el Proyecto Honeynet UNAM[8], el cual realiza investigaciones sobre este tipo de tecnologías y sobre mecanismos de detección y análisis de tráfico malicioso en general.

1.2.2 DARKNETS

El término de Darknet se refiere a todo el conjunto de direcciones o segmentos IP que no han sido asignados para algún servicio o dispositivo específico en una red, y que son destinadas a ser utilizados por uno o varios equipos con características especiales para identificación de posibles amenazas en la red. Partiendo de este concepto, todo el tráfico que circule por la red y que de algún modo se relacione con alguna de estas direcciones o segmentos IP, adquiere la característica de ser potencialmente anómalo o malicioso debido a que en teoría no debería existir tráfico dirigido hacia ellos.

Este concepto ha sido utilizado para poder identificar tráfico malicioso y fallas en dispositivos de red aprovechando su baja probabilidad de que se trate de un falso positivo. Posee las características adecuadas para combinarla con tecnologías honeypot y poder capturar, procesar y analizar muestras de tráfico malicioso en honeynets y redes en producción.

Debido a que es una parte fundamental en la propuesta de este trabajo de tesis, el concepto de Darknet se abordará de una manera más detallada en el capítulo 2.

[8] Proyecto dependiente de la Subdirección de Seguridad de la Información/UNAM-CERT en la Dirección General de Cómputo y de Tecnologías de la Información y Comunicación (DGTIC) y participante como UNAM-Chapter en The Honeynet Project. http://www.honeynet.unam.mx

1.2.3 TELESCOPIOS DE RED

De manera general, este concepto se refiere a un mecanismo de detección y monitoreo basado en el despliegue de sensores distribuidos a lo largo de un entorno de red o bien mediante el análisis de datos de varios dispositivos de conectividad. De cualquier forma, una de las características principales es que permite conjuntar la información recopilada por diversas fuentes de un espacio grande de monitoreo.

Su objetivo principal es detectar tráfico malicioso y monitorear la actividad general de la red combinando diversas tecnologías como las mencionadas en secciones anteriores (IDS, honeypots, Darknets, etc.), sin embargo, los telescopios de seguridad pueden ir más allá que los sistemas convencionales. Debido a que este tipo de mecanismo está enfocado a entornos de gran escala, adicionalmente de obtener un esquema de la actividad de la red, dan la posibilidad de generar muestras de información de tráfico malicioso de redes externas y dependiendo de su tamaño y complejidad, pueden llegar a ser referencias importantes para conocer las tendencias globales en la actividad del tráfico de Internet.

Algunas características de este tipo de tecnología son:

- Modelo de detección distribuido
- Modelo de administración centralizada o distribuida
- Entornos de gran escala
- Gran cantidad de información recopilada, procesada y almacenada
- Altamente demandante en recursos de hardware
- Generadores de información estadística importante
- Monitores de las tendencias de tráfico en espacios grandes de Internet
- Identificación de anomalías a nivel global

Quizá en un concepto muy general, la diferencia formal entre una Darknet y un telescopio de red es su tamaño y complejidad.

En apartados posteriores se mencionan características específicas del Telescopio de seguridad de la UNAM, cuyo principal motor de detección de tráfico de red malicioso es la Darknet-UNAM, propuesta de esta tesis.