Análisis de flujos de red en entornos virtualizados

Said Carlos Gerardo; Pompei Sabrina
Escuela de Tecnología, Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA), Buenos Aires, Argentina

carlossaid@gmail.com, sabrinapompei@unnoba.edu.ar

Resumen

Las redes actuales son segmentadas y descentralizadas lo cual plantea desafíos a la forma tradicional de supervisión. ¿Cómo dar cobertura de seguridad a los segmentos internos y externos? ¿Las tecnologías SIEM son suficientes? Frecuentemente la respuesta parece estar en recurrir a los flujos de datos en la red, estos nos proporcionan respuestas para el ¿Quién?, ¿Qué?, ¿Cuándo?, ¿Dónde? y ¿Cómo? de las potenciales amenazas informáticas.

Por otra parte la virtualización está transformando el centro de datos de manera sustancial. Se ha demostrado que esta tecnología genera ahorro en los costos de operación pero si no es cuidadosamente planificada, diseñada e implementada puede ser un riesgo informático asociado a la pérdida de la visibilidad. ¿Qué hacer con el análisis de flujos de paquetes de red, en estos entornos? ¿Cómo se obtiene el mismo nivel de visibilidad en un entorno virtual y en uno físico?

Nuestro trabajo pretende identificar el valor agregado de los flujos de datos en la red y el análisis de las ventajas y desventajas de las distintas tecnologías de análisis de flujos de red, bajo la perspectiva de incrementar la visibilidad de eventos vinculados a la seguridad informática y el análisis de la llamada 'capa 7'.

Palabras clave: seguridad informática, flujo de red, visibilidad

Contexto

El presente artículo trata sobre las líneas a desarrollar en el marco de un proyecto de investigación de seguridad informática desarrollado por alumnos y docentes, del área informática, de la Escuela de Tecnología de la UNNOBA.

Introducción

A medida que las amenazas de seguridad informática que enfrentan las organizaciones crecen de manera exponencial, la necesidad de una mayor visibilidad de la actividad de la red se convierte en un factor esencial. Los ataques y las amenazas se han vuelto más sofisticados. Las llamadas 'amenazas avanzadas persistentes' (APTs) son una evidencia de ello.

Los responsables de seguridad de la información se enfrentan a una variedad de ataques informáticos, agravados por el cambio frecuente de las mismas. Estos tienen la tarea de proteger el activo más crítico de la organización, la información.

Las redes actuales son segmentadas y descentralizadas, lo cual plantea desafíos a la forma tradicional de supervisión (basada en análisis de registros de eventos, SNMP/RMON, firewalls, y pruebas de detección de intrusos).

La confusión sobre las diferentes categorías de tecnologías que intentan resolver o mitigar estas amenazas ha hecho que sea difícil definir los requisitos necesarios.

Un elemento importante de la confusión está en las definiciones. Los acrónimos SEM, SIM y SIEM se han utilizado indistintamente, sin embargo, existen diferencias en el significado y las capacidades.

Un aspecto de la gestión de la seguridad que se ocupa de la supervisión en tiempo real, correlación de eventos, notificaciones, generación de informes y visualización de eventos se conoce comúnmente como gestión de eventos de seguridad (SEM – Security Event Management).

Otro aspecto, relacionado con el almacenamiento de mayor plazo, el análisis y la comunicación de los datos existentes en los registros, se conoce como gestión de la información de seguridad (SIM – Security Information Management). SIM requiere un repositorio para los datos registrados, técnicas eficientes de indexación y herramientas de consulta y gestión de informes.

Cuando SEM y SIM se combinan, se convierten en Seguridad de la Información y Gestión de Eventos (SIEM – Security Information and Event Management).

Los eventos de seguridad recolectados a partir de múltiples fuentes u orígenes de datos tienen que ser filtrados para reducir el esfuerzo necesario para administrarlos y poder dar prioridad a las actividades de respuesta correspondientes a estos eventos. El análisis y las consultas deben identificar aquellas anomalías que deberían ser efectivamente analizadas.

Las prestaciones de los productos SIEM incluyen la recolección, análisis y presentación de información de la red y los dispositivos de seguridad, las aplicaciones de gestión de identidad y acceso, herramientas de gestión de vulnerabilidades y soluciones vinculadas al cumplimiento de políticas/estándares normativos; registros del sistema operativo, eventos en bases de datos y aplicaciones, y datos provenientes de amenazas externas.

La correlación de eventos aparece como una característica esencial de la tecnología SIEM. La correlación establece relaciones entre las entradas o eventos que son generados por los dispositivos, sistemas o aplicaciones, basándose en características como el origen, destino, protocolo o tipo de evento.

Las tecnologías de correlación de eventos se basan en reglas o correlación estadística. La correlación basada en reglas permite establecer el patrón de eventos.

Respecto al flujo de red y detección de anomalías:

¿Cómo dar cobertura de seguridad a los segmentos internos y externos? ¿Las tecnologías SIEM son suficientes? Con mayor frecuencia la respuesta parece estar en recurrir a los flujos de datos en la red. El flujo de datos en la red proporciona respuestas para el ¿Quién?, ¿Qué?, ¿Cuándo?, ¿Dónde? y ¿Cómo? de las potenciales amenazas informáticas".

Los sistemas tales como sistema de prevención de intrusiones (IPS – Intrusion Prevention Systems), clasifican la colección de datos que recogen de la red y hacen un trabajo de detección de amenazas conocidas. Pero, por su naturaleza, proporcionan datos de microanálisis y no pueden proporcionar el contexto adecuado, o visión más amplia del problema.

Nuestro trabajo pretende identificar el valor agregado de los flujos de datos en la red y el valor que ofrecen en ayudar a

proteger la misma. El flujo de datos en la red y su recolección para el análisis de seguridad, es la 'puerta de ingreso' a un campo denominado detección de anomalías o análisis de comportamiento.

¿Pero qué es un flujo de datos en la red? Un flujo de datos en la red es un grupo unidireccional de paquetes que comparten características comunes.

El análisis de estos flujos puede brindar visibilidad en la capa 7 (capa de aplicaciones) para ayudar a comprender acabadamente las potenciales amenazas y responder en función de la actividad que tiene lugar dentro de la red. Esto facilita detectar las amenazas que otras tecnologías no logran detectar, asegurar el cumplimiento de las políticas y marcos normativos, reducir los riesgos para los servicios de misión crítica, datos y activos digitales. Los agentes recolectores de flujo de red proporcionan esta visibilidad de la capa de aplicación (Capa 7) , así como la clasificación de las aplicaciones y protocolos tales como: voz sobre IP (VoIP), multimedia, planificación de recursos empresariales (ERP – Enterprise Resource Planning), base de datos, y otros protocolos y aplicaciones. Los datos de flujo de las aplicaciones se obtiene a partir de un examen de cada paquete, lo que también permite la detección de amenazas informáticas a través del análisis del contenido de 'la carga útil' (payload) de los paquetes.

Hablar de datos de flujo de red cubre el conjunto de intercambios de paquetes o "conversaciones" entre los dispositivos de una red. Un registro de flujo de red proporciona información acerca de una conversación específica entre dos dispositivos que utilizan un protocolo específico, incluye campos/datos que describen la interacción. Estos datos incluyen direcciones de origen y destino IP, el protocolo de transporte (por ejemplo User Datagram Protocol (UDP)), el Protocolo de Control de la Transmisión (TCP – Transmission Control Protocol), los puertos de origen y destino, información de la aplicación, estadísticas de tráfico, Tipo de Servicio (ToS), calidad del servicio y en algunos casos, la 'carga útil del paquete' o 'payload'.

Por lo general las tecnologías se detienen en la Capa 4 y proporcionan información de direcciones IP a nivel de red y puertos TCP / UDP. Este nivel de prestación es útil para obtener una idea general de las conversaciones que ocurren en protocolos bien definidos, sin embargo, los datos resumidos y estáticos procedentes de diversas fuentes no proporcionan visibilidad acabada de la actividad de la red y las aplicaciones.

La correlación de esta información con el flujo de la red, los eventos de seguridad, las vulnerabilidades, la información de identidad y la inteligencia aplicada a las amenazas es una manera más certera de obtener una visión completa y precisa del estado de seguridad de una organización.

Debido a que el tráfico de los servidores virtualizados no puede ser recuperada utilizando tecnologías tradicionales de monitoreo, debemos buscar soluciones que den visibilidad a los mismo para monitoreo de los entornos virtuales. Es decir proporcionar visibilidad de la de capa de aplicación a todo el tráfico de red virtual.

¿Qué sucede en entornos con Virtualización creciente?

Uno de los desarrollos más interesantes en TI es la virtualización. Aunque existente en mainframes desde hace años, la virtualización es una de las corrientes principales de TI y está transformando el centro de datos de una manera sustancial. Los estudios muestran que la tecnología de virtualización genera

ahorro sustancial en los costos de operación (OPEX – Operational Expenses). Pero si no es cuidadosamente planificada, diseñada e implementada, la virtualización puede ser un riesgo informático asociado a la pérdida de la visibilidad y la consiguiente complejidad en la solución de problemas donde las herramientas y metodologías tradicionales no funcionan como lo hacían antes.

¿Qué hacer con el análisis de flujos de paquetes de red, en estos entornos? ¿Cómo se obtiene el mismo nivel de visibilidad en un entorno virtual y en uno físico? Muchos enfoques tradicionales de la seguridad de la información, no se transfieren de manera directa al mundo virtual.

Una red física de cualquier tamaño por lo general tiene puntos adecuados para la inspección de tráfico. Esto hace que sea relativamente simple obtener muestras.

Esto no es igual en redes virtuales, las cuales normalmente tiene varias tarjetas de interfaz de red física (NIC), switches virtuales y tarjetas de red virtuales.

Una opción es monitorear el tráfico que egresa de los NIC del servidor físico en la red física.

Esto podría ser aceptable en función de las expectativas y necesidades, pero es posible que sólo se vea una pequeña parte del tráfico real que los servidores virtuales crean: solo los que salen del servidor físico. ¿Qué pasa con el tráfico que es interno a la caja física? ¿Cómo monitorear el tráfico VM-a-VM que es interna al servidor físico y no llega a la red física? Es probable que esta sea la mayor parte del tráfico, especialmente si las aplicaciones de 'n-capas' están virtualizadas en la misma máquina física.

Exportar los flujos de red en la capa de conmutación virtual (o virtual switch) es una solución a este problema. Proporciona visibilidad sobre el tráfico que atraviesa el conmutador virtual, pero se queda dentro del mismo servidor físico, así como el tráfico que fluye de la NIC física a la red física.

Nuestra propuesta es el análisis de las ventajas y contras de las distintas tecnologías de análisis de flujos de red, bajo la perspectiva de incrementar la visibilidad de eventos vinculados a la seguridad informática, en las capas 3,4 y 7 del modelo OSI: Netflow, sFlow, Qflow, Vflow, IPFIX, jFlow, entre otras.

Líneas de Investigación, Desarrollo e Innovación

- La relación entre SIEM, la trazabilidad de eventos y el análisis forense digital.
- Análisis de flujos de paquetes de red en entornos virtuales. Visibilidad en entorno virtual y físico.
- Visibilidad de eventos en las capas 3,4 y 7 del modelo OSI: Netflow, sFlow, Qflow, Vflow, IPFIX, jFlow, entre otras.

Resultados y Objetivos

El objetivo general del presente proyecto consiste en comparar tecnologías de análisis de flujos de red, a la luz de los cambios tecnológicos y sociales existentes (redes sociales, obsolescencia de tecnologías SIEM, entornos virtualizados, dispositivos móviles, grandes volúmenes de datos (en Inglés: Big Data), tecnologías BYOD (bring your own device)).

Formación de Recursos Humanos

Integran el equipo de trabajo docentes investigadores, docentes investigadores

en formación y alumnos de las carreras informáticas de la Escuela de Tecnología de la UNNOBA.

Se trabaja en la presentación de un trabajo de tesis en el marco del presente proyecto.

Referencias

- Estan Cristian (2002). New directions in traffic measurement and accounting. Recuperado el 3 de Marzo de 2014, de http://pages.cs.wisc.edu/~estan/publications/newdirstechrep.html.
- http://www.securitywizardry.com/radar.htm
- http://www.ntop.org/