Arquitectura

Moisés Gautier Gómez

2 de junio de 2015

Vamos a analizar la posible estructura funcional del sistema para desgranar cada uno de los componentes el mismo en profundidad.

1. Monitor

Habrá una capa de monitorización por cada asset que reciba del sistema.

- Este monitor podría ser un demonio que se ejecutara en background y que pudiera levantarse cuando se quisiera hacer uso real del mismo.
- También podría ser un sistema sincrono restfull/nodejs o servicio web que solicitase información al dispositivo de red.
- Sistema de multiprocesamiento en el que cada hilo del proceso principal fuera para un asset de información distinta que manipula y obtiene los resultados deseados para que el padre sea capaz de compactarlos y enviarlos a su capa superior de monitorización.
- ¿Recibirá la información comprimida por tiempo, momento del día o constantemente? La idea más simple podría ser un socket de comunicación con el monitor en tiempo real que vaya haciendo minería sobre los datos extraídos.
- Data Warehouse, ¿se podría aplicar esto cómo metaconcepto? (Extraído de la wikipedia)
    - Bill Inmon1 fue uno de los primeros autores en escribir sobre el tema de los almacenes de datos, define un data warehouse (almacén de datos) en términos de las características del repositorio de datos:

- Orientado a temas.- Los datos en la base de datos están organizados de manera que todos los elementos de datos relativos al mismo evento u objeto del mundo real queden unidos entre sí.
- Variante en el tiempo.- Los cambios producidos en los datos a lo largo del tiempo quedan registrados para que los informes que se puedan generar reflejen esas variaciones.
- No volátil.- La información no se modifica ni se elimina, una vez almacenado un dato, éste se convierte en información de sólo lectura, y se mantiene para futuras consultas.
- Integrado.- La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes.

Inmon defiende una metodología descendente (top-down) a la hora de diseñar un almacén de datos, ya que de esta forma se considerarán mejor todos los datos corporativos. En esta metodología los Data marts se crearán después de haber terminado el data warehouse completo de la organización.

- Simplemente puede ser algo que analice algo y nada más. Extraer información de unos ficheros con unos determinados formatos y extraer rasgos característicos.

2. Assets

Las distintas formas que tienen los assets de generar logs o información.

2.1. Firewall logs

- Iptables: Fuente Coge la información de sus logs mediante el registo del núcleo syslog que puede ser leía mediante dmesg. (Mirar también syslog.conf)
- Ipcop: Fuente Ofrece monitorización de elementos de la red así cómo del pc en el que se encuentra instalado Fuente. La forma en cómo obtiene registros del sistema viene descrita aquí: Fuente y es usando una vez más el demonio syslogd.
- Más firewalls: Fuente

2.2. IDS logs

Fuente

- Suricata: Fuente Guía de usuario: Fuente Tipo de logs que nos ofrece suricata:
  - syslog alerting:Fuente
  - json.log output: Fuente
  - Suricata with OOSIM: Fuente
- Snort: Fuente - RSyslog rate limiting configuration: Fuente
- Para hosts: OSSEC - Fuente
- Otros:
  - Snorby: Fuente
  - Sguil: Fuente

2.3. CPU, processes, memory, IO

- Hardinfo: Fuente Generar informes en html del sistema.

2.4. Netflow

Descripción general: NetFlow es un protocolo de red desarrollado por Cisco Systems para recolectar información sobre tráfico IP. Se ha definido el flujo de network de numerosas maneras. La definición tradicional de Cisco implica una clave séptuple en que el flujo se define como una secuencia unidireccional de paquetes que comparten los siguientes 7 valores:

- Dirección IP de origen.
- Dirección IP de destino.
- Puerto UDP o TCP de origen.
- Puerto UDP o TCP de destino.
- Protocolo IP.
- Interfaz (SNMP ifIndex)

- Tipo de servicio IP

[Extraído de la wikipedia]

- Netflow Analyzer: Fuente
- NFDump: Fuente
- Más info: SWITCH - NetFlow

2.5. Antivirus logs

- Symantec logs: Fuente
- Kaspersky (como obtener logs): Fuente

2.6. SNMP info

Descripción aquí.

2.7. Nessus

Configuraciones de Nessus (para logs): Fuente
Comentan que los logs de nessus son muy pesados de procesar y manipular.

3. SIS format

Se intentará buscar un formato de tablas o base de datos que pueda ir desde la granularidad más fina a la más alta posible dentro del sistema de gestión de incidencias.

3.1. BBDD

Primer tabla: Monitores

ID_monitor	ID_monitor_process	IP_host	IP_host	MAC_origen	Tipo_log	ID_incidencia
1.1	1.2	1.3	1.4	1.5	1.6	1.7

1. Identificador del monitor para dicha tabla

2. Identificador del monitor cómo proceso dentro del sistema desde dónde lo invoca

3. Dirección ip del host desde dónde opera el monitor

4. Dirección mac del dispositivo de origen

5. Tipo de log: Intrusión, Tráfico Red, Benchmak, Puertos, etc (Podría ser otra tabla en dónde se podría incluir dato anómalos)

6. Identificador de la incidencia en el caso que fuera (primary_key de el id de la tabla que haga referencia)

¿Incluir el nombre del archivo log en un registro o campo de bd? ¿Se va almacenar los archivos dentro de la bd o cómo índices que podremos consultar a la bd para buscarlos dentro del sistema de archivos?

Segunda tabla: Incidencia

ID	Descripción	Tipo_Incidencia	Hora y Fecha	IP_host	IP_origen	MAC_origen	Traceo_IP
2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8

1. Identificador de la incidencia para dicha tabla

2. Descripción de la incidencia (ver el tamaño máximo que se podría poner como string)

3. Tipo de Incidencia: Intrusión, Tráfico Red, Virus, etc

4. Hora y fecha de la incidencia en el sistema origen

5. Dirección ip del host dónde se ha encontrado la incidencia

6. Dirección ip del nodo o dispositivo dónde está la incidencia

7. Dirección mac del dispositivo origen

8. Traceo de la ip con algún software del estilo (border-check)

4. Local Visualization

4.1. Web UI

4.2. Local visualization algorithms

5. Manager

5.1. SIS Protocol

5.1.1. PCA algorithm

5.1.2. BBDD management

5.1.3. Watchdog of monitors

5.1.4. Communication with other managers