Módulo 1. Loggin 101

CÓRDOBA CREA FUTURO

1. Formatos y fuentes

En los entornos tecnológicos actuales, cada sistema, aplicación o dispositivo conectado genera registros que describen lo que ocurre durante su funcionamiento. Estos registros —conocidos como logs— documentan eventos como accesos de usuarios, errores de aplicaciones, cambios de configuración, comunicaciones entre sistemas o ejecuciones de procesos. En contextos de monitoreo y seguridad informática, estos datos permiten reconstruir actividades, analizar comportamientos del sistema y detectar patrones que podrían indicar fallas operativas o incidentes de seguridad.

La gestión estructurada de logs constituye una práctica ampliamente utilizada en infraestructuras digitales modernas. Según el National Institute of Standards and Technology, los registros de eventos proporcionan evidencia fundamental para el análisis forense, la detección de intrusiones y la verificación del cumplimiento de políticas organizacionales (Kent & Souppaya, 2006). En consecuencia, la capacidad de recolectar, almacenar y analizar registros provenientes de múltiples sistemas se convierte en una condición necesaria para comprender lo que sucede dentro de una infraestructura tecnológica compleja.

En ese contexto, el trabajo con logs plantea una serie de cuestiones operativas que orientan la práctica cotidiana del monitoreo: ¿de qué sistemas provienen los registros que se analizan? ¿cómo se estructuran los eventos generados por distintos sistemas operativos? ¿qué formatos permiten representar esos datos de manera comprensible y procesable por herramientas de análisis? Asimismo, cuando múltiples sistemas interactúan entre sí, surge la necesidad de interpretar correctamente el momento exacto en que ocurrió cada evento. Esta cuestión introduce la importancia de los timestamps estandarizados y de la gestión de zonas horarias, elementos que permiten ordenar cronológicamente los registros producidos en entornos distribuidos.

A su vez, la acumulación constante de eventos plantea decisiones relacionadas con la retención de logs. Determinar cuánto tiempo deben conservarse estos registros implica considerar factores técnicos, operativos y regulatorios. Una política de retención adecuada permite mantener disponible la información necesaria para auditorías, investigaciones de incidentes y análisis históricos del comportamiento de los sistemas.

A partir de este marco, en esta unidad se abordarán los principales formatos y fuentes de registros utilizados en entornos de monitoreo y gestión de logs. Se analizarán los sistemas de registro basados en

Syslog y los eventos generados por Windows, el uso del formato JSON para estructurar eventos, la estandarización temporal mediante timestamps, y los criterios que orientan las políticas de retención de logs. Estos elementos constituyen la base técnica sobre la cual se organizan posteriormente los procesos de análisis, correlación y monitoreo en plataformas de gestión de eventos y sistemas SIEM.

Syslog / Windows

Los sistemas informáticos registran continuamente eventos que describen su funcionamiento. Estos registros permiten documentar actividades del sistema operativo, acciones de usuarios, errores de aplicaciones o interacciones entre distintos componentes de una infraestructura tecnológica. En el ámbito del monitoreo y la seguridad informática, estos registros constituyen una fuente de información que permite analizar comportamientos del sistema, detectar anomalías y reconstruir secuencias de eventos relevantes (Kent & Souppaya, 2006).

Entre las fuentes más utilizadas para generar registros se encuentran los sistemas basados en Syslog y los registros de eventos de Windows. Ambos mecanismos permiten registrar eventos del sistema y ponerlos a disposición de herramientas de monitoreo o plataformas de análisis, aunque se originan en entornos tecnológicos diferentes y utilizan estructuras de registro distintas.

El protocolo Syslog se utiliza ampliamente en sistemas Unix, dispositivos de red y diversos servicios de infraestructura. Este protocolo define un método estandarizado para generar y transmitir mensajes de registro entre sistemas. El funcionamiento del protocolo se describe en el documento RFC 5424, que establece la estructura de los mensajes, los campos que los componen y las reglas para su intercambio entre dispositivos (Gerhards, 2009).

Los mensajes Syslog contienen información que permite identificar el contexto del evento registrado. Entre los elementos incluidos en el mensaje se encuentran el momento en que ocurrió el evento, el sistema que generó el registro y la aplicación responsable de producirlo. Esta estructura facilita que los sistemas receptores puedan interpretar los eventos y clasificarlos según su origen o su nivel de severidad.

Tabla 1.Componentes de un mensaje Syslog y su función dentro del registro

Campo del mensaje	Qué representa	Función en el análisis del log
PRI	Prioridad del evento	Permite identificar nivel de severidad

VERSION	Versión del protocolo	Define formato del mensaje
TIMESTAMP	Momento del evento	Permite ordenar cronológicamente
HOSTNAME	Sistema que envía el log	Identifica el origen del evento
APP-NAME	Aplicación que genera el evento	Permite identificar el servicio involucrado
PROCID	Identificador del proceso	Permite rastrear procesos
MSGID	Tipo de evento	Facilita clasificación
MSG	Descripción del evento	Contiene el detalle del registro

Fuente: elaboración propia.

La estructura estandarizada del protocolo Syslog permite que distintos dispositivos puedan enviar sus registros hacia un mismo servidor recolector. Routers, firewalls, servidores o aplicaciones pueden transmitir eventos hacia un sistema centralizado de logs, donde posteriormente pueden ser analizados por herramientas de monitoreo o plataformas SIEM.

En los entornos basados en Microsoft Windows, los eventos del sistema se registran mediante el mecanismo conocido como Windows Event Log. Este sistema documenta diferentes tipos de actividades relacionadas con el funcionamiento del sistema operativo y de las aplicaciones instaladas. Los eventos registrados pueden incluir acciones de autenticación de usuarios, cambios en la configuración del sistema o errores producidos por aplicaciones.

Según la documentación técnica de Microsoft, los eventos generados por los equipos Windows pueden ser recopilados mediante mecanismos como Windows Event Forwarding, que permite enviar los registros desde múltiples equipos hacia un servidor centralizado para su posterior análisis (Microsoft, 2023). Este mecanismo facilita la supervisión de eventos de seguridad y la recopilación de registros provenientes de distintos sistemas dentro de una infraestructura.

Los registros de eventos de Windows se organizan en categorías que agrupan diferentes tipos de actividad del sistema.

A continuación se presenta un cuadro comparativo que resume los principales tipos de registros del sistema Windows Event Log y el tipo de eventos que cada uno almacena.

Tabla 2. Registros principales de Windows Event Log

Tipo de registro	Tipo de eventos registrados	Ejemplo de uso
Security	Autenticaciones y eventos de seguridad	Intentos de login
System	Eventos del sistema operativo	Inicio de servicios
Application	Eventos generados por aplicaciones	Error de software
Setup	Instalación y configuración	Instalación de componentes

Fuente: elaboración propia.

La centralización de registros provenientes de múltiples fuentes constituye una práctica habitual en la gestión de logs. Al reunir eventos generados por servidores, estaciones de trabajo y dispositivos de red dentro de un repositorio común, se facilita el análisis conjunto de la actividad de la infraestructura tecnológica. Esta integración permite identificar relaciones entre eventos que, analizados de forma aislada, resultarían difíciles de interpretar.

Comprender cómo funcionan estas fuentes de registro permite reconocer de qué manera se generan los eventos que posteriormente serán analizados en sistemas de monitoreo o plataformas SIEM. A partir de estas fuentes, los registros pueden representarse mediante formatos estructurados que facilitan su procesamiento por herramientas de análisis. En el siguiente apartado se abordará uno de los formatos más utilizados para este propósito: JSON.

JSON

A medida que las infraestructuras tecnológicas comenzaron a integrar múltiples sistemas, aplicaciones y dispositivos, surgió la necesidad de representar los registros de eventos mediante formatos que permitieran organizar la información de forma clara y estructurada. En este contexto, el formato JSON (JavaScript Object Notation) se consolidó como un mecanismo ampliamente utilizado para el intercambio y representación de datos en distintos entornos tecnológicos.

JSON se define como un formato ligero de intercambio de datos diseñado para representar información mediante estructuras simples y fácilmente interpretables por aplicaciones informáticas (JSON.org, s. f.). Su estructura se basa en la organización de datos a partir de pares clave–valor, lo que permite describir atributos específicos dentro de un mismo registro de manera explícita y comprensible.

En los sistemas de registro de eventos, esta característica resulta especialmente útil. Cada evento generado por un sistema puede representarse mediante un conjunto de campos que describen distintas dimensiones del evento, como el sistema que lo generó, el momento en que ocurrió, el tipo de actividad registrada o los actores involucrados. Esta estructura facilita que las herramientas de análisis puedan acceder directamente a cada atributo del evento sin necesidad de interpretar cadenas de texto complejas.

Según la especificación del formato JSON, los datos se organizan principalmente a partir de dos estructuras: objetos y arreglos. Los objetos agrupan pares clave–valor dentro de una estructura delimitada por llaves, mientras que los arreglos permiten almacenar múltiples valores dentro de una misma categoría (JSON.org, s. f.). Esta combinación de estructuras permite representar información jerárquica y compleja manteniendo una organización clara de los datos.

En el contexto de la gestión de logs, un evento puede representarse como un objeto JSON que contiene distintos atributos relacionados con la actividad registrada. Cada uno de estos atributos describe un aspecto específico del evento, lo que facilita su interpretación por sistemas de monitoreo o plataformas de análisis.

Tabla 3. Ejemplo simplificado de evento representado en JSON

Campo	Ejemplo de valor
timestamp	2024-05-10T14:22:31Z
host	servidor01
event_type	login_attempt
user	admin
source_ip	192.168.10.25
status	failed

Fuente: elaboración propia.

En este ejemplo, cada campo representa un atributo del evento registrado. El campo timestamp indica el momento en que ocurrió la actividad, mientras que host identifica el sistema que generó el registro. Otros campos describen el tipo de evento, el usuario involucrado o la dirección desde la cual se originó la acción.

El uso de estructuras JSON facilita el procesamiento automatizado de registros de eventos. Las herramientas de análisis pueden interpretar cada atributo del evento como un campo independiente, lo que permite realizar consultas específicas sobre los datos almacenados. Por ejemplo, un sistema de monitoreo podría buscar todos los eventos asociados a una dirección IP específica o identificar intentos de autenticación fallidos dentro de un intervalo temporal determinado.

Otra característica relevante de JSON es su flexibilidad para representar distintos tipos de datos. Los sistemas que generan logs pueden incluir diferentes campos según la naturaleza del evento registrado. De este modo, un evento relacionado con la autenticación de usuarios puede contener atributos distintos a los de un evento generado por una aplicación o por un dispositivo de red. Esta flexibilidad permite adaptar la estructura de los registros a distintos contextos operativos sin modificar el formato general del registro.

La adopción de formatos estructurados como JSON también facilita la integración de datos provenientes de múltiples fuentes dentro de un mismo entorno de análisis. Cuando los registros se organizan mediante estructuras basadas en clave–valor, resulta más sencillo combinar eventos generados por distintos sistemas y analizarlos de forma conjunta. Esta característica resulta especialmente útil en plataformas de monitoreo que procesan grandes volúmenes de eventos provenientes de diferentes dispositivos y servicios.

En los entornos de monitoreo de seguridad, la representación estructurada de los registros también contribuye a mejorar los procesos de correlación de eventos. Al disponer de campos claramente identificados —como direcciones IP, identificadores de usuario o nombres de procesos—, las herramientas de análisis pueden establecer relaciones entre eventos generados por diferentes sistemas dentro de una misma infraestructura.

En este sentido, el uso de JSON forma parte de una tendencia más amplia hacia la estandarización y estructuración de los registros de eventos. A medida que las organizaciones integran múltiples sistemas dentro de sus infraestructuras tecnológicas, la representación clara de los datos se convierte en un requisito para facilitar su análisis y su utilización en procesos de monitoreo.

Comprender cómo se organizan los registros mediante formatos como JSON permite interpretar de manera más precisa la información generada por los sistemas. Sobre esta base, resulta posible analizar otro elemento fundamental para el estudio de los eventos registrados: la representación temporal de los registros, que será abordada en el siguiente apartado a partir del uso de timestamps y la estandarización temporal mediante el formato ISO 8601.

Timestamps y zona horaria

En los sistemas de registro de eventos, cada actividad documentada se asocia con un momento específico en el tiempo. Esta referencia temporal se expresa mediante lo que se denomina timestamp, es decir, una marca temporal que indica cuándo ocurrió un evento dentro de un sistema informático. La presencia de timestamps permite ordenar cronológicamente los registros y reconstruir secuencias de actividad dentro de una infraestructura tecnológica.

En los entornos de monitoreo y análisis de logs, la precisión temporal resulta especialmente relevante. Cuando múltiples sistemas generan eventos simultáneamente —como servidores, aplicaciones, dispositivos de red o estaciones de trabajo— resulta necesario contar con una referencia temporal coherente que permita comparar registros provenientes de diferentes fuentes. La correcta interpretación del tiempo en que ocurrió cada evento constituye un elemento central para comprender el comportamiento de los sistemas y analizar incidentes de seguridad.

Para facilitar esta interpretación, muchos sistemas utilizan estándares de representación temporal que permiten expresar fechas y horas de forma consistente. Uno de los estándares más utilizados para este propósito

es ISO 8601, que define una forma estructurada de representar información temporal en entornos informáticos (Ionos, s. f.).

El estándar ISO 8601 establece un formato que organiza la información temporal de manera jerárquica, comenzando por el año, seguido del mes, el día y, posteriormente, la hora, los minutos y los segundos. Esta estructura facilita la ordenación automática de registros y evita ambigüedades que pueden surgir cuando diferentes sistemas utilizan formatos de fecha distintos.

Tabla 4. Ejemplo de estructura temporal según ISO 8601

Elemento	Ejemplo
Fecha	2024-05-10
Hora	14:22:31
Zona horaria	Z
Representación completa	2024-05-10T14:22:31Z

Fuente: elaboración propia.

En esta representación, la letra T separa la fecha de la hora, mientras que el indicador Z señala que el tiempo se expresa en UTC (Coordinated Universal Time). Esta convención permite interpretar los registros generados por sistemas ubicados en distintas regiones geográficas sin generar inconsistencias en la lectura temporal de los eventos.

El uso de un estándar temporal común resulta especialmente importante en infraestructuras distribuidas. En una red corporativa o en entornos de computación en la nube, los sistemas pueden operar en diferentes zonas horarias o ubicarse en centros de datos situados en distintos países. Si cada sistema registrara los eventos utilizando su hora local sin una referencia común, la reconstrucción cronológica de las actividades podría resultar ambigua o imprecisa.

El estándar ISO 8601 permite resolver esta dificultad mediante la incorporación explícita de la zona horaria dentro de la representación temporal. Esto permite identificar la relación entre la hora registrada y el tiempo universal coordinado, lo que facilita la comparación entre registros generados en distintos sistemas (ISO, 2004).

En los sistemas de gestión de logs, la utilización de timestamps estandarizados permite ordenar los eventos de manera precisa y reconstruir secuencias de actividad dentro de la infraestructura tecnológica. Esta capacidad resulta particularmente relevante cuando se analizan incidentes de seguridad, ya que permite identificar la

relación temporal entre distintos eventos registrados en múltiples sistemas.

Asimismo, la estandarización temporal facilita el trabajo de las herramientas de análisis automatizado. Los sistemas de monitoreo pueden ordenar los eventos según su marca temporal, identificar patrones de actividad dentro de intervalos específicos y correlacionar eventos que ocurren en distintos sistemas dentro de una misma ventana temporal.

La correcta gestión de los timestamps también requiere considerar la sincronización de los relojes de los sistemas que generan registros. Cuando los sistemas presentan diferencias significativas en sus relojes internos, los eventos pueden registrarse con marcas temporales inconsistentes, lo que dificulta la reconstrucción cronológica de las actividades. Por esta razón, muchas infraestructuras utilizan mecanismos de sincronización temporal para mantener alineados los relojes de los distintos dispositivos.

En el análisis de logs, la combinación de timestamps precisos, representaciones temporales estandarizadas y sincronización entre sistemas permite construir una línea temporal confiable de los eventos registrados. Esta línea temporal constituye una herramienta fundamental para comprender el comportamiento de la infraestructura tecnológica y analizar incidentes dentro de entornos complejos.

A partir de esta base, el siguiente apartado abordará otro aspecto central de la gestión de logs: la retención de registros, es decir, las decisiones relacionadas con cuánto tiempo deben conservarse los logs y bajo qué criterios se almacenan. En ese contexto se presentará una checklist de verificación para políticas de retención de logs, que permitirá sistematizar los criterios técnicos utilizados en la gestión de registros.

Retención

En los sistemas de gestión de logs, la generación de registros constituye solo una parte del proceso de administración de eventos. Una vez que los registros son producidos por los sistemas, las organizaciones deben establecer criterios que determinen cómo se almacenan, durante cuánto tiempo se conservan y en qué condiciones pueden ser consultados. Estas decisiones forman parte de lo que se denomina política de retención de logs.

La retención de registros permite mantener disponible información histórica sobre el funcionamiento de los sistemas informáticos. Esta información resulta útil para diferentes actividades organizacionales, como auditorías de seguridad, análisis forense de incidentes o revisiones operativas de los sistemas. Según el National Institute of Standards and Technology, la gestión adecuada de los logs implica

definir políticas que establezcan cómo se recolectan, almacenan y preservan los registros dentro de una infraestructura tecnológica (Kent & Souppaya, 2006).

Uno de los desafíos asociados a la retención de logs se relaciona con el volumen de datos que los sistemas generan diariamente. Los servidores, aplicaciones, dispositivos de red y sistemas de seguridad producen eventos de forma continua, lo que implica que los repositorios de logs pueden crecer rápidamente si no se establecen mecanismos de gestión adecuados. En este contexto, las políticas de retención permiten definir criterios que equilibran la disponibilidad de información histórica con las capacidades de almacenamiento disponibles.

La guía NIST SP 800-92 señala que las organizaciones deben considerar diversos factores al establecer sus políticas de retención. Entre estos factores se encuentran las necesidades operativas de monitoreo, los requisitos de auditoría, las normativas regulatorias aplicables y las capacidades técnicas de almacenamiento de la infraestructura (Kent & Souppaya, 2006). Estos elementos permiten determinar cuánto tiempo deben conservarse los registros antes de ser archivados o eliminados.

En muchos entornos de seguridad informática, la retención de logs también cumple una función relevante en los procesos de investigación de incidentes. Cuando ocurre una actividad sospechosa dentro de una infraestructura tecnológica, los registros históricos permiten reconstruir la secuencia de eventos que condujo al incidente. La disponibilidad de estos registros facilita identificar qué sistemas estuvieron involucrados, qué acciones se realizaron y en qué momento ocurrieron.

La gestión de la retención también implica considerar la forma en que los registros se almacenan y protegen. Los logs contienen información sensible sobre el funcionamiento de los sistemas, por lo que su almacenamiento debe garantizar la integridad de los datos y evitar modificaciones no autorizadas. Las prácticas recomendadas incluyen el uso de sistemas de almacenamiento centralizado, controles de acceso y mecanismos de verificación de integridad.

En este contexto, las organizaciones suelen establecer procedimientos que regulan el ciclo de vida de los registros. Este ciclo puede incluir diferentes etapas, como la recolección inicial de los eventos, su almacenamiento en repositorios activos para análisis operativo y su posterior traslado a sistemas de archivo para almacenamiento a largo plazo.

Para facilitar la revisión de estos aspectos, es posible utilizar herramientas de verificación que permitan evaluar si los distintos criterios de retención han sido contemplados dentro de la organización. A continuación, se presenta una checklist de verificación que resume algunos de los elementos considerados en la gestión de retención de logs.

Checklist para evaluar una política de retención de logs

- [ ] Se identificaron todas las fuentes de logs críticas (servidores, dispositivos de red, aplicaciones y sistemas de seguridad).
- [ ] Se definió un período de retención para cada tipo de registro según necesidades operativas o regulatorias.
- [ ] Se dispone de infraestructura de almacenamiento suficiente para conservar los logs durante el período definido.
- [ ] Los logs se almacenan en sistemas centralizados que permiten su consulta y análisis posterior.
- [ ] Se aplican mecanismos que preservan la integridad de los registros y evitan modificaciones no autorizadas.
- [ ] Se establecieron controles de acceso para limitar quién puede consultar o administrar los logs.
- [ ] Existen procedimientos de archivado para trasladar registros antiguos a almacenamiento histórico.
- [ ] La política de retención contempla requisitos de auditoría, cumplimiento normativo o investigaciones de incidentes.

Esta checklist permite revisar de forma sistemática los principales elementos que intervienen en una política de retención de logs. Al utilizar este tipo de herramientas de verificación, las organizaciones pueden evaluar si sus prácticas de almacenamiento de registros contemplan los requisitos operativos, técnicos y de seguridad asociados a la gestión de eventos.

La definición de políticas de retención constituye, por lo tanto, una práctica que complementa los procesos de recolección y análisis de logs. Mantener registros disponibles durante períodos adecuados permite sostener la capacidad de análisis histórico de los sistemas y contribuye a mejorar la respuesta ante incidentes de seguridad.

Al concluir esta unidad, se han abordado distintos aspectos relacionados con las fuentes y formatos de registros utilizados en entornos de monitoreo, incluyendo los sistemas Syslog, los registros de eventos de Windows, el uso de JSON para estructurar eventos, la estandarización temporal mediante timestamps y las políticas de retención de logs. Estos elementos constituyen la base técnica sobre la cual se construyen posteriormente los procesos de normalización, clasificación y correlación de eventos, que serán analizados en la siguiente unidad dedicada a taxonomías y normalización de logs.

CONTINUAR

2. Taxonomías y normalización

En los entornos de monitoreo de seguridad y gestión de logs, la recolección de registros provenientes de múltiples sistemas constituye solo una parte del proceso de análisis de eventos. Los servidores, aplicaciones, dispositivos de red y sistemas operativos generan continuamente registros que describen actividades del sistema, acciones de usuarios y eventos de seguridad. Sin embargo, estos registros suelen producirse utilizando estructuras, formatos y denominaciones diferentes, lo que introduce dificultades cuando se intenta analizarlos de forma conjunta dentro de una infraestructura tecnológica compleja.

A medida que las organizaciones incorporan nuevos sistemas y servicios digitales, el volumen y la diversidad de los logs aumentan considerablemente. En estos escenarios, los equipos de monitoreo deben interpretar registros provenientes de múltiples fuentes que pueden utilizar terminologías distintas para describir eventos similares. Esta heterogeneidad dificulta la lectura de los datos y puede limitar la capacidad de los sistemas de monitoreo para identificar patrones de actividad relevantes.

Frente a este escenario, los procesos de taxonomía y normalización de logs se convierten en prácticas utilizadas para organizar la información

registrada por los sistemas. Las taxonomías permiten clasificar los eventos dentro de categorías que facilitan su interpretación, mientras que la normalización busca transformar los registros provenientes de distintas fuentes para que compartan una estructura coherente. Estas prácticas permiten que los datos puedan analizarse dentro de un mismo marco conceptual, independientemente del sistema que los haya generado.

En este contexto aparecen preguntas relevantes para el trabajo cotidiano en monitoreo: ¿qué información dentro de un registro resulta verdaderamente significativa para interpretar un evento? ¿cómo se identifican los campos que describen la actividad registrada? ¿de qué manera pueden clasificarse los eventos para facilitar su análisis dentro de herramientas de monitoreo o plataformas SIEM? Asimismo, cuando múltiples sistemas generan registros simultáneamente, resulta necesario establecer mecanismos que permitan relacionar eventos entre sí y mantener consistencia en los datos.

En esta unidad se abordarán los conceptos de campos clave, categorización de eventos, correlaciones simples y consistencia en la estructura de los logs. Estos elementos permiten organizar los registros generados por distintos sistemas y facilitan su análisis dentro de herramientas de monitoreo y plataformas de gestión de eventos de seguridad.

Campos clave

En los sistemas de monitoreo y análisis de eventos, los registros generados por los distintos componentes de una infraestructura tecnológica contienen múltiples atributos que describen lo ocurrido. Sin embargo, no todos los elementos presentes en un log poseen el mismo nivel de relevancia para el análisis. En este contexto, resulta necesario identificar los campos clave, es decir, aquellos atributos del registro que permiten describir de forma clara el evento registrado y facilitan su interpretación dentro de herramientas de monitoreo o plataformas SIEM.

Los campos clave representan información esencial sobre el evento que se registra. Entre estos datos se incluyen, por ejemplo, el momento en que ocurrió la actividad, el sistema que generó el registro, el tipo de acción que se realizó o el usuario asociado al evento. Cuando estos atributos se identifican y se organizan de manera consistente, los registros pueden compararse entre sí y analizarse dentro de un mismo entorno de monitoreo, incluso cuando provienen de sistemas distintos.

En los entornos modernos de análisis de logs, la identificación y organización de estos campos suele apoyarse en esquemas de normalización que definen cómo deben representarse los atributos de los eventos. Uno de los modelos más utilizados para este propósito es el Elastic Common Schema (ECS), un esquema que propone una estructura común para describir eventos provenientes de distintas

fuentes de datos (Elastic, s. f.). El objetivo de este enfoque consiste en facilitar la integración y el análisis de registros dentro de plataformas de observabilidad y sistemas de monitoreo.

El modelo ECS organiza los eventos a partir de campos estandarizados que describen distintos aspectos del evento registrado. Estos campos permiten representar información sobre el sistema que genera el evento, la actividad que se registra y el contexto en el que ocurre. Al utilizar una estructura común para describir los eventos, se facilita la correlación y el análisis conjunto de registros provenientes de múltiples fuentes.

Tabla 5. Ejemplos de campos clave en registros de eventos

Campo	Descripción
@timestamp	Momento en que ocurrió el evento
host.name	Nombre del sistema que generó el registro
event.action	Acción específica registrada
event.category	Categoría general del evento

user.name	Usuario asociado al evento
source.ip	Dirección IP de origen
destination.ip	Dirección IP de destino

Fuente: elaboración propia.

La identificación de estos campos permite que los sistemas de monitoreo interpreten los eventos de manera estructurada. Cuando los registros incluyen atributos claramente definidos, las herramientas de análisis pueden realizar búsquedas más precisas, filtrar eventos según determinados criterios y establecer relaciones entre registros generados por distintos sistemas.

Otra ventaja de la utilización de campos clave consiste en que facilita la comparación entre eventos generados por fuentes heterogéneas. Por ejemplo, un intento de autenticación fallido puede registrarse de manera distinta en un sistema operativo, en una aplicación o en un dispositivo de red. Sin embargo, si todos los registros incluyen campos equivalentes —como el usuario involucrado, el momento del evento y el sistema de origen— resulta posible analizar estos eventos dentro de una misma estructura.

El uso de esquemas de campos normalizados también contribuye a mejorar la capacidad de las plataformas de monitoreo para procesar grandes volúmenes de datos. Al contar con una estructura de campos consistente, los motores de indexación pueden organizar la información de forma más eficiente, lo que permite realizar consultas sobre grandes cantidades de registros sin perder claridad en la interpretación de los eventos.

En el contexto de los sistemas SIEM, los campos clave constituyen uno de los elementos fundamentales para la normalización de logs. Estos campos permiten transformar registros provenientes de distintos sistemas para que compartan una estructura común, lo que facilita su análisis dentro de plataformas de seguridad. Cuando los eventos utilizan una estructura consistente, las herramientas de monitoreo pueden identificar patrones de actividad y generar alertas basadas en condiciones definidas sobre los distintos atributos del evento.

La identificación y utilización de campos clave constituye, por lo tanto, una práctica que contribuye a organizar la información registrada por los sistemas y a facilitar su análisis posterior. A partir de esta base, resulta posible avanzar hacia un segundo nivel de organización de los eventos: la categorización, que permite clasificar los registros según el tipo de actividad que representan. Este aspecto será desarrollado en el siguiente apartado dedicado a las categorías de eventos.

Categorías

En los entornos de monitoreo y análisis de logs, los registros generados por distintos sistemas describen una amplia variedad de actividades. Estos eventos pueden representar acciones de usuarios, operaciones del sistema operativo, comunicaciones de red o actividades de seguridad. Para facilitar la interpretación de esta información, resulta útil agrupar los eventos según el tipo de actividad que representan. En este contexto surge el concepto de categorías de eventos, que permite clasificar los registros dentro de grupos que comparten características similares.

La categorización de eventos constituye una práctica que facilita la organización de los registros dentro de sistemas de monitoreo y plataformas SIEM. Cuando los eventos se clasifican dentro de categorías comunes, resulta más sencillo analizar patrones de actividad, identificar comportamientos anómalos y construir consultas que permitan explorar los registros de forma estructurada. Esta organización también permite reducir la complejidad asociada al análisis de grandes volúmenes de datos provenientes de múltiples fuentes.

Dentro de los modelos de normalización de logs, las categorías permiten describir el tipo general de actividad que representa un evento. En el marco del Elastic Common Schema (ECS), las categorías constituyen uno de los mecanismos utilizados para estructurar la información de los registros. Este esquema propone un conjunto de categorías que permiten agrupar eventos según la naturaleza de la actividad registrada (Elastic, s. f.).

Las categorías de eventos permiten representar acciones relacionadas con diferentes aspectos del funcionamiento de los sistemas. Algunos eventos se relacionan con procesos de autenticación, otros con comunicaciones de red, actividades de archivos o interacciones con aplicaciones. Al clasificar los eventos dentro de estas categorías, se facilita el análisis conjunto de registros que describen actividades similares.

La utilización de categorías facilita la organización de los registros dentro de los sistemas de análisis de eventos. Cuando los logs incluyen una categoría claramente definida, las herramientas de monitoreo pueden agrupar eventos similares y analizar tendencias dentro de cada tipo de actividad. Por ejemplo, un sistema de seguridad podría examinar todos los eventos relacionados con autenticación para identificar intentos de acceso fallidos o patrones de uso inusuales.

Además de facilitar el análisis de los registros, la categorización también contribuye a mejorar los procesos de correlación de eventos. Cuando los registros se encuentran clasificados según su tipo de actividad, las plataformas de monitoreo pueden establecer relaciones entre eventos pertenecientes a categorías específicas. Esto permite construir reglas de análisis que consideren la secuencia de eventos dentro de un mismo contexto operativo.

Otra ventaja de utilizar categorías consiste en que permite mantener consistencia en la forma en que los eventos se describen dentro de los sistemas de monitoreo. Cuando distintas fuentes de logs utilizan un mismo sistema de categorización, los registros pueden interpretarse dentro de un marco común, lo que facilita su análisis y reduce las ambigüedades en la interpretación de los eventos.

En los entornos de seguridad informática, la categorización de eventos constituye un paso importante dentro de los procesos de normalización de logs. Al agrupar los registros según el tipo de actividad que representan, se crea una base estructurada que facilita la construcción de consultas, la generación de alertas y el análisis de comportamientos dentro de la infraestructura tecnológica.

La identificación de categorías de eventos complementa el uso de campos clave analizado en el apartado anterior. Mientras los campos clave permiten describir atributos específicos del evento, las categorías permiten comprender qué tipo de actividad representa el registro dentro del sistema. Esta combinación de atributos estructurados y clasificación por categorías permite organizar los logs de forma más coherente dentro de las plataformas de monitoreo.

A partir de esta base, es posible avanzar hacia un nivel adicional de análisis de los registros: la correlación de eventos. Este proceso permite establecer relaciones entre distintos registros para identificar secuencias de actividad dentro de la infraestructura tecnológica. En el siguiente apartado se abordará el concepto de correlaciones simples, que constituye uno de los primeros pasos para relacionar eventos dentro de los sistemas de monitoreo.

Correlaciones simples

En los sistemas de monitoreo de seguridad, los registros de eventos describen actividades individuales que ocurren dentro de una infraestructura tecnológica. Cada log representa una acción específica

del sistema, como un intento de autenticación, la ejecución de un proceso o una comunicación de red. Sin embargo, muchos comportamientos relevantes dentro de un sistema solo pueden interpretarse adecuadamente cuando se analizan múltiples eventos en conjunto. En este contexto surge el concepto de correlación de eventos, que consiste en establecer relaciones entre distintos registros para interpretar una secuencia de actividades dentro de la infraestructura.

Las correlaciones permiten identificar patrones de comportamiento que no resultan evidentes cuando los eventos se observan de forma aislada. Un único registro puede describir una actividad legítima dentro del sistema, pero una serie de eventos relacionados puede indicar una situación que requiere análisis adicional. Por ejemplo, múltiples intentos fallidos de autenticación seguidos por un acceso exitoso desde la misma dirección IP pueden representar un patrón que merece ser examinado con mayor atención dentro de un entorno de monitoreo.

En las plataformas de análisis de logs, la correlación de eventos se basa en la comparación de atributos presentes en los registros. Los sistemas de monitoreo pueden relacionar eventos que comparten determinados campos, como el usuario asociado a la actividad, la dirección IP de origen o el sistema que generó el registro. Cuando estos atributos se encuentran organizados de forma consistente, resulta posible

identificar relaciones entre eventos generados por diferentes sistemas.

La documentación de Wazuh describe que los registros recolectados por los sistemas de monitoreo son procesados mediante mecanismos que permiten interpretar la estructura del log y extraer información relevante del evento. Una vez interpretados los registros, el motor de análisis puede aplicar reglas que permiten identificar patrones dentro de los datos y relacionar eventos que comparten características comunes (Wazuh, s. f.). Este proceso permite transformar registros individuales en información contextual sobre la actividad que ocurre dentro de la infraestructura tecnológica.

Las correlaciones simples representan una forma inicial de establecer relaciones entre eventos. Este tipo de correlación se basa en condiciones directas que permiten relacionar registros a partir de atributos comunes o secuencias temporales. Por ejemplo, un sistema de monitoreo puede correlacionar eventos que comparten la misma dirección IP de origen, el mismo identificador de usuario o el mismo sistema generador del registro.

En muchos casos, las correlaciones simples también consideran el orden temporal de los eventos. Cuando distintos registros ocurren dentro de un mismo intervalo de tiempo y comparten ciertos atributos, los sistemas de monitoreo pueden interpretarlos como parte de una misma secuencia de actividad. Este tipo de análisis permite comprender mejor el comportamiento de los sistemas y detectar situaciones que podrían pasar desapercibidas si los eventos se analizaran de forma individual.

Las correlaciones simples constituyen uno de los primeros niveles de análisis dentro de los sistemas de monitoreo de seguridad. A medida que las plataformas de análisis integran más fuentes de datos y reglas de detección más complejas, las correlaciones pueden incorporar múltiples condiciones, secuencias de eventos y relaciones entre distintos sistemas. Sin embargo, incluso en estos escenarios más avanzados, las correlaciones iniciales suelen basarse en la comparación de atributos básicos presentes en los registros.

Para que estos procesos de correlación funcionen de manera confiable, los registros deben mantener una estructura coherente en la forma en que representan la información. Cuando los logs utilizan campos inconsistentes o estructuras diferentes, los sistemas de análisis pueden encontrar dificultades para relacionar eventos provenientes de distintas fuentes. Por esta razón, la consistencia en los registros de eventos constituye un aspecto importante dentro de los procesos de normalización de logs.

Este aspecto será abordado en el siguiente apartado, donde se analizará cómo la consistencia en los datos de los registros permite mejorar la interpretación de los eventos y facilita su análisis dentro de sistemas de monitoreo y plataformas SIEM.

Consistencia

En los sistemas de monitoreo de eventos, la capacidad de analizar registros provenientes de múltiples fuentes depende en gran medida de que los datos mantengan una estructura consistente. La consistencia en los logs se refiere a la forma en que los eventos representan la información utilizando campos, formatos y convenciones similares, lo que permite que los sistemas de análisis interpreten los registros de manera uniforme.

Cuando los logs se generan utilizando estructuras diferentes o campos inconsistentes, los sistemas de monitoreo pueden encontrar dificultades para interpretar correctamente los eventos. En estos casos, la información registrada puede resultar ambigua o incompleta, lo que reduce la capacidad de las herramientas de análisis para establecer relaciones entre eventos o detectar patrones de actividad dentro de la infraestructura tecnológica.

Las plataformas de monitoreo de seguridad utilizan procesos de procesamiento de logs que dependen de la consistencia en la forma en que se representan los datos. Según la documentación de Wazuh, los registros recolectados por los agentes o integraciones del sistema son procesados mediante componentes que interpretan la estructura del log y extraen información relevante del evento (Wazuh, s. f.). Este proceso requiere que los registros mantengan formatos reconocibles que permitan identificar correctamente los distintos atributos del evento.

Dentro de este flujo de procesamiento, los sistemas de análisis utilizan mecanismos conocidos como decoders o analizadores de logs, que permiten identificar la estructura de los registros y extraer campos específicos como direcciones IP, identificadores de usuario o tipos de actividad. Cuando los logs presentan una estructura consistente, estos mecanismos pueden interpretar los eventos con mayor precisión y facilitar su posterior análisis.

La consistencia también resulta importante para los procesos de correlación de eventos. Como se analizó en el apartado anterior, las plataformas de monitoreo establecen relaciones entre registros que comparten atributos comunes. Si los eventos utilizan estructuras diferentes para representar la misma información, los sistemas de análisis pueden encontrar dificultades para identificar estas relaciones.

En este contexto, la consistencia en los registros constituye un elemento central dentro de los procesos de normalización de logs. Al asegurar que los eventos utilicen campos y formatos coherentes, se facilita la interpretación de los registros y se mejora la capacidad de las herramientas de monitoreo para analizar grandes volúmenes de datos provenientes de múltiples sistemas.

La gestión de la consistencia en los logs también implica revisar cómo se recolectan los registros dentro de la infraestructura tecnológica. En muchos sistemas de monitoreo, los logs son generados por diferentes

componentes —como servidores, aplicaciones o dispositivos de red— y posteriormente enviados hacia una plataforma central de análisis. En este proceso, mantener una estructura homogénea en los registros permite simplificar el procesamiento de los datos y mejorar la eficiencia de los mecanismos de análisis.

Para facilitar la revisión de estos aspectos, es posible utilizar herramientas de verificación que permitan evaluar si los registros cumplen con criterios básicos de consistencia antes de ser utilizados en procesos de análisis o correlación de eventos. A continuación, se presenta una checklist que resume algunos de los elementos que pueden considerarse al revisar la consistencia de los logs dentro de un sistema de monitoreo.

Checklist para verificar la consistencia de logs antes de su análisis

- [ ] Los registros utilizan un formato temporal consistente (por ejemplo, ISO 8601).
- [ ] Los eventos incluyen campos básicos que permiten identificar el origen del registro.
- [ ] Los logs utilizan nombres de campos consistentes para representar información similar.

- [ ] Los registros contienen información suficiente para interpretar el evento registrado.
- [ ] Los eventos mantienen una estructura reconocible para los mecanismos de procesamiento de logs.
- [ ] Los logs pueden ser interpretados correctamente por los decoders o analizadores del sistema de monitoreo.
- [ ] Los registros provenientes de distintas fuentes pueden integrarse dentro de un mismo esquema de análisis.

La aplicación de criterios de consistencia en los registros permite mejorar la calidad de los datos utilizados por las plataformas de monitoreo. Cuando los logs mantienen estructuras coherentes, los sistemas de análisis pueden interpretar los eventos con mayor precisión y facilitar la detección de patrones de actividad dentro de la infraestructura tecnológica.

A lo largo de esta unidad se han analizado distintos aspectos relacionados con la organización y normalización de los registros de eventos, incluyendo la identificación de campos clave, la clasificación de eventos mediante categorías, el uso de correlaciones simples para relacionar registros y la importancia de mantener consistencia en los datos de los logs. Estos elementos permiten estructurar la información generada por los sistemas y facilitan su análisis dentro de herramientas de monitoreo y plataformas SIEM.

CONTINUAR

Referencias

Elastic. (s. f.). *Elastic Common Schema (ECS) reference*. https://www.elastic.co/guide/en/ecs/current/ecs-reference.html

Gerhards, R. (2009). *The Syslog protocol* (RFC 5424). Internet Engineering Task Force. https://datatracker.ietf.org/doc/html/rfc5424

Ionos. (s. f.). *ISO 8601 date and time format*. https://www.ionos.es/digitalguide/paginas-web/desarrollo-web/iso-8601/

JSON.org. (s. f.). *Introducing JSON*. https://www.json.org/json-en.html

Kent, K., & Souppaya, M. (2006). *Guide to computer security log management* (NIST Special Publication 800-92). National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-92.pdf

Microsoft. (2023). *Use Windows Event Forwarding to assist in intrusion detection.* Microsoft Learn. https://learn.microsoft.com/en-us/windows/security/operating-system-security/device-management/use-windows-event-forwarding-to-assist-in-intrusion-detection

Wazuh. (s. f.). *Log data collection: How it works.* https://documentation.wazuh.com/current/user-manual/capabilities/log-data-collection/how-it-works.html

CONTINUAR