Implementación de Security Data Lake con Splunk

Creación de reglas de correlación y modelos para la detección avanzada de amenazas

Tomás García Hidalgo

Grado de ingeniería informática

Seguridad informática

Consultor: Jorge Miguel Moneo

Profesora: Helena Rifà Pous

Universitat Oberta de Catalunya

uoc.edu

Introducción

Objetivos Académicos

- Exponer problemáticas y limitaciones actuales de los sistemas de recolección y análisis de eventos
- Revisión del estado de arte en los métodos de detección de amenazas y las plataformas utilizadas en la actualidad
- Estudio de los diferentes productos del mercado

Objetivos de implementación

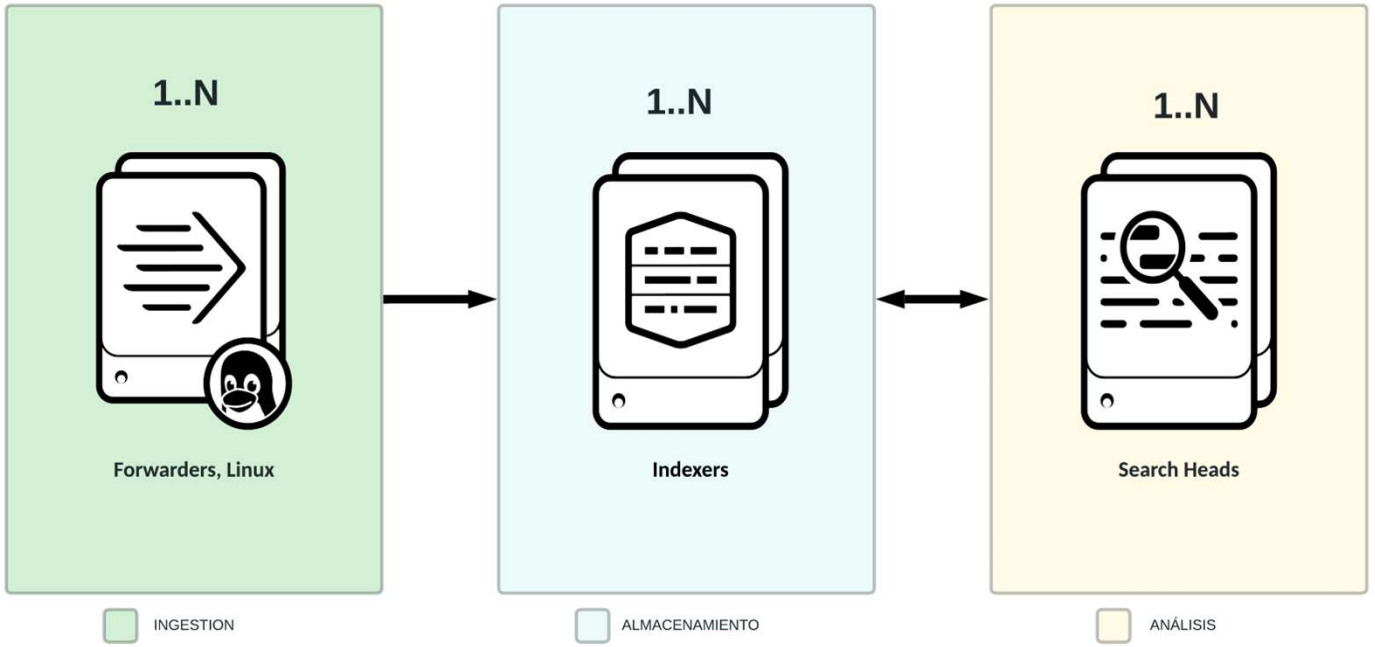
- Implementación de una plataforma *security data lake* distribuida
- Configuración de un entorno de pruebas
- Desarrollo de las diferentes reglas de correlación y modelos para la detección de anomalías

Arquitectura del sistema propuesto

Modelo distribuido de tres capas

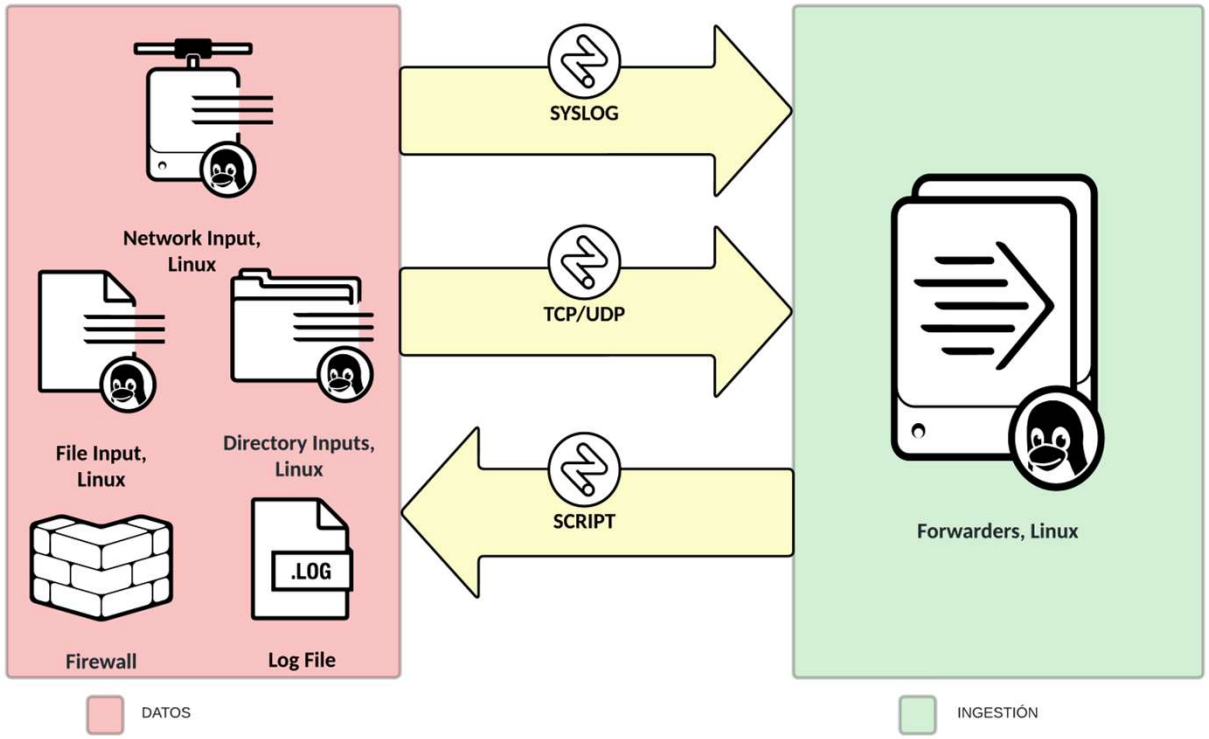
- Capa de Ingestión – Recolección de todos los eventos a través de los diferentes protocolos
- Capa de Almacenamiento – Almacenamiento de todos los eventos indexados
- Capa de Análisis – Procesamiento de las diferentes consultas y transformaciones sobre los datos

Arquitectura del sistema propuesto



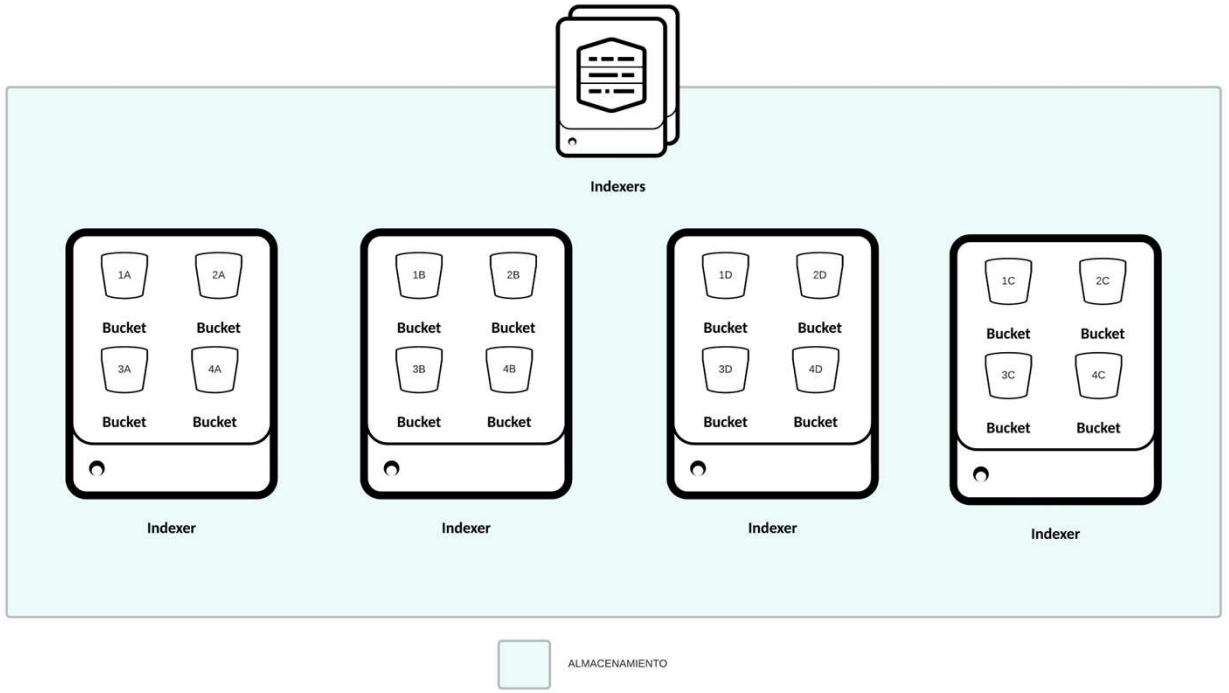
Arquitectura del sistema propuesto

Capa de Ingestión



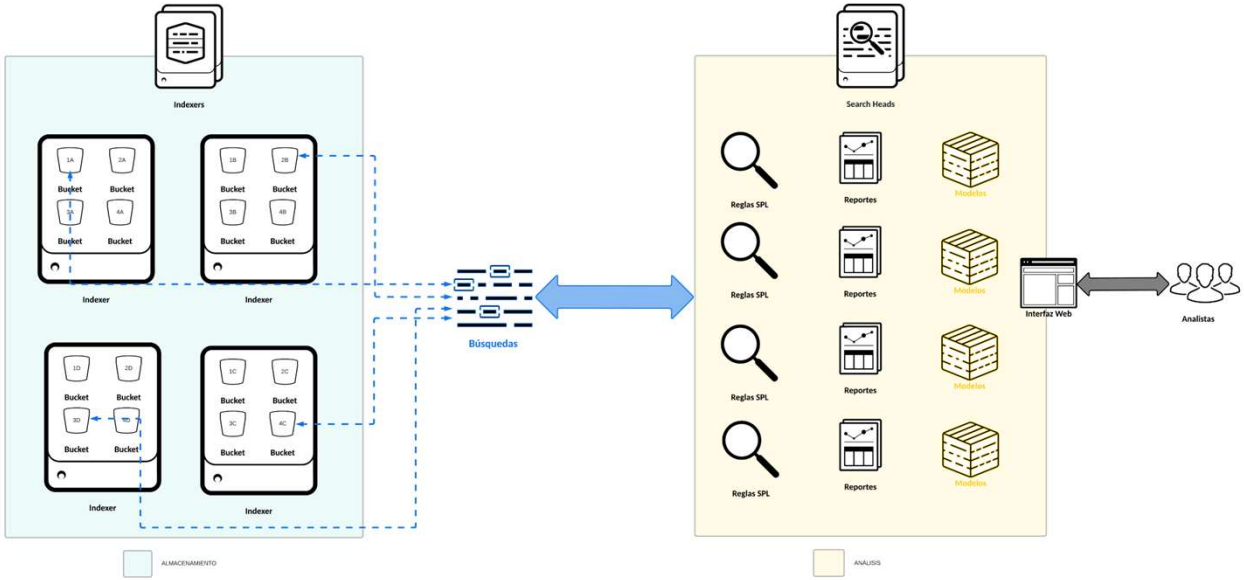
Arquitectura del sistema propuesto

Capa de Almacenamiento



Arquitectura del sistema propuesto

Capa de Análisis



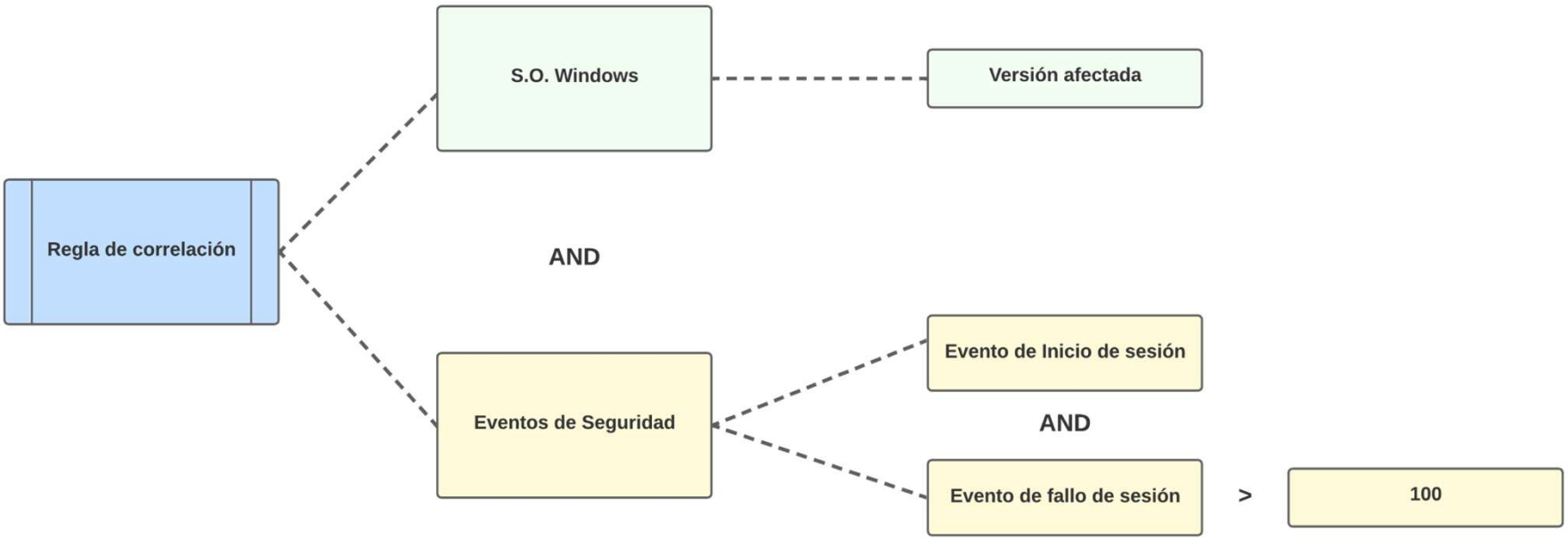
Reglas de Correlación y Modelos de Detección

Existen dos técnicas principales para la detección de patrones de actividad sospechosa

- Reglas de correlación – Son un conjunto de reglas lógicas que producen un subconjunto de eventos
- Modelos de detección – Algoritmos utilizados para detectar patrones o comportamientos anómalos

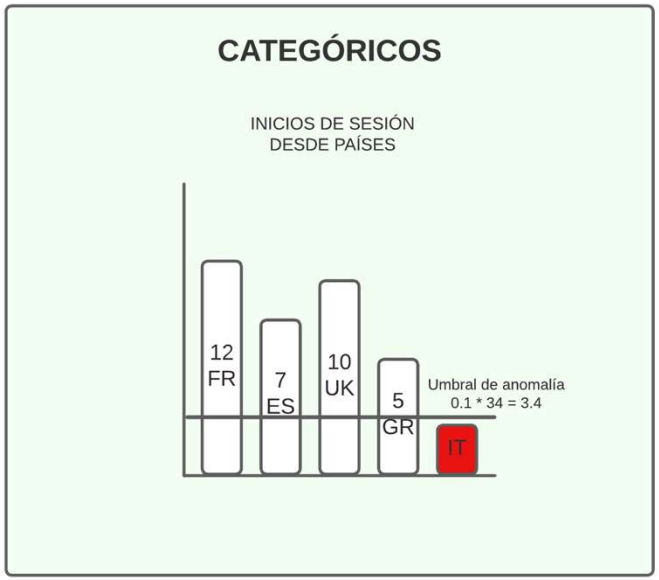
Reglas de Correlación

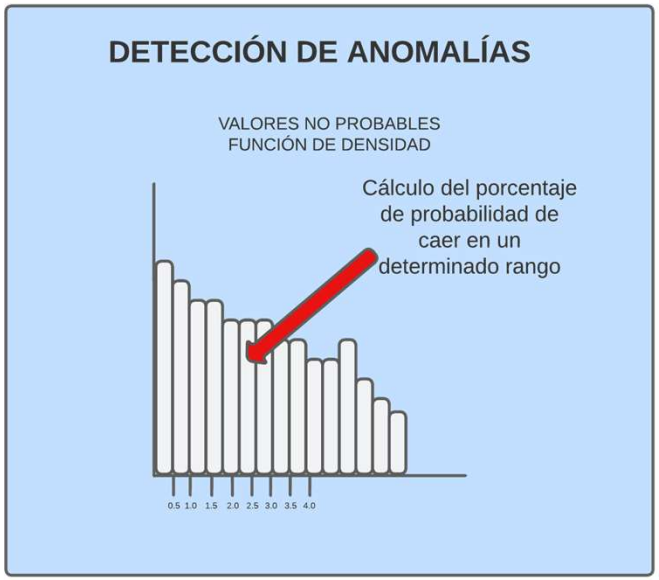
Ejemplo



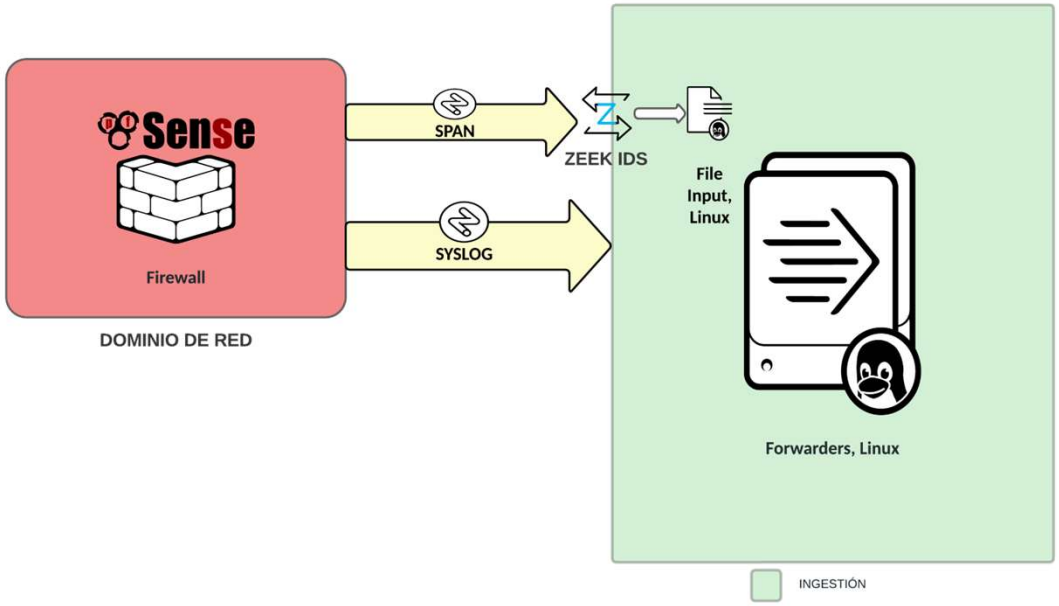
Modelos de Detección

Ejemplos

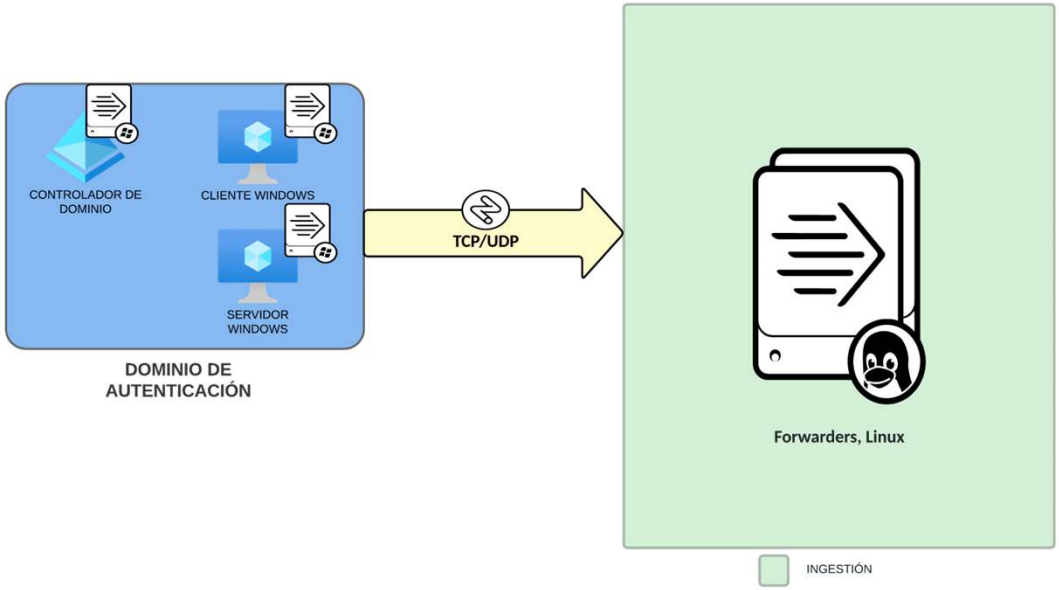




Arquitectura del dominio de red



Arquitectura del dominio de autenticación



Conclusiones

- Se ha presentado una propuesta viable de implementación de un entorno *security data lake* distribuido que resuelve las problemáticas de los SIEM actuales
- Reglas de correlación – Son más efectivas para la identificación de patrones de amenazas conocidas pero poco flexibles
- Modelos de detección – Algoritmos con beneficios para la detección de patrones anómalos en nuestros datos pero muchas veces difíciles de implementar