Gerenciamento e Correlação de Eventos de Segurança: Concepção de Uma Alternativa para Sistemas SIEM

Ricardo Almeida[1], Victor Junes[1], Thiago Cardoso[1], Diórgenes da Rosa[1], Rafael Padilha[1], Thomas Oliveira[1], Roger Machado[1]

[1]Universidade Federal de Pelotas (UFPel), Pelotas – RS – Brasil

{ricardo.almeida,vrcjunes,thiago.cardoso,diorgenes.yuri,

rafael.padilha,thomas.aguiar,rmachado.ifm}@ufpel.edu.br

Resumo. *Os eventos de segurança, geralmente registrados em logs, podem fornecer informações valiosas sobre o ambiente monitorado, sendo úteis para avaliação do estado da segurança da infraestrutura, e também servindo de evidências para auditoria, análise forense e entre outras atividades. Sendo assim, o objetivo central deste trabalho é de apresentar uma proposta de solução de gerenciamento e correlação de eventos de segurança sem custo com licenças de software. Para avaliação, a proposta foi colocada em produção gerando como resultado inicial, a possibilidade de identificação e reação à atividades suspeitas em tempo de execução, bem como o fornecimento de uma visão holística sobre o ambiente distribuído analisado.*

1. Introdução

Nos atuais ambientes computacionais, onde novas ameaças e vulnerabilidades surgem em um ritmo acelerado, a velocidade para detectar e responder à incidentes de segurança é fundamental. Com isso, as tecnologias de segurança devem ingerir e processar eventos de segurança o mais próximo possível do tempo real, procurando evitar a sobrecarrega das equipes de resposta com um número elevado de alertas de segurança [Shackleford 2017].

As plataformas de Gerenciamento de Eventos e Informações de Segurança (SIEM) destinam-se a consumir eventos produzidos na infraestrutura computacional provenientes de diferentes aplicações, incluindo logs de softwares e dispositivos de segurança. Além de fornecerem um painel para os analistas avaliarem os eventos e usarem as informações disponíveis para responder às atividades suspeitas identificadas, elas também podem correlacionar os eventos em tempo de execução na busca por padrões suspeitos, e quando oportuno, realizarem as devidas ações automaticamente. Não obstante, essas plataformas também são úteis para retenção dos eventos, permitindo a conformidade com legislações.

Tendo isto em vista, o objetivo central deste trabalho consiste de apresentar um relato de experiência e a concepção de uma proposta com funcionalidades de sistemas SIEM, a qual visa fornecer correlação de eventos em tempo de execução e uma visão holística sobre a infraestrutura distribuída monitorada.

2. Métodos

Após a revisão conceitual em torno de sistemas SIEM, foram considerados os desafios impostos pelo ambiente computacional na Universidade Federal de Pelotas (UFPel), onde

esta proposta foi desenvolvida. Sendo assim, primeiramente será descrito brevemente o ambiente computacional utilizado.

No que diz respeito aos requisitos impostos pelo ambiente da UFPel, é importante salientar que a UFPel conta com mais de 400 prédios distribuídos, tendo dois principais *data centers* - um no Campus Anglo (CA) e outro no Campus Capão do Leão (CCL). Destaca-se que o campis são geograficamente distribuídos, o que implica na dependência da Internet para intercomunicação. Desta forma, os seguintes requisitos foram definidos e considerados durante a definição da proposta:

- sem custo com licenças (obrigatório) - considerando os altos custos das soluções comerciais dos sistemas SIEM disponíveis no mercado e limitações orçamentárias, a solução deve considerar a inexistência de investimento monetário em licenças. Por outro lado, além da existência de hardware disponível, para este projeto existe a possibilidade de alocação de tempo e recurso humano para pesquisa e trabalho exclusivo com a estratégia adotada, o que também seria necessário com a aquisição de uma SIEM.
- resiliência (obrigatório) - devem haver estratégias para eventuais quedas nos links de comunicação entre os prédios. Adicionalmente, espera-se que a solução seja capaz de lidar com formato de eventos inesperados e com paradas anormais dos nodos e dos serviços que constituem a solução.
- flexível (obrigatório) - este conceito deve estar empregado de forma ampla na solução, buscando atender as diferentes necessidades e restrições do ambiente, incluindo a coleta com e sem agente, a possibilidade personalização das regras de correlação, a criação de gráficos e painéis para auxílio na detecção de desvios em métricas, a ativação e desativação de módulos, etc.
- escalável (obrigatório) - a solução deve permitir a escalabilidade vertical (aumentando a capacidade de um único nodo) e horizontal (adicionando novos nodos e dividindo o processamento).
- suporte a heterogeneidade (obrigatório) - coleta e tratamento de eventos em diferentes formatos, gerados por aplicações, sistemas e dispositivos, como *switches* e *Access Points* (AP's), incluindo eventos armazenados em bancos de dados.
- correlação de eventos (obrigatório) - correlação de eventos baseada em regras, preferencialmente com sintaxe similar à SQL, a qual ainda que de acordo com a situação desejada possa alcançar um nível considerável de complexidade, é considerada pelos autores uma sintaxe mais simples que o tradicional uso de expressões regulares (como adotado pelo HIDS OSSEC[1] para correlação, por exemplo).
- instalação facilitada (opcional) - a solução adotada deve considerar características que auxiliem na instalação e configuração de agentes para coleta dos eventos.
- código aberto (opcional) - a proposta deve evitar "caixas pretas", ou seja, espera-se a maior transparência possível. O benefício disto é o conhecimento dos algoritmos e da forma de implementação e comunicação entre os componentes, além da maior flexibilidade e customização por parte dos administradores para melhor atender o ambiente monitorado.

Após a experiência frustada com a solução *Free and Open Source Software* (OSSIM) por questões de estabilidade e escalabilidade, confirmando as análises disponíveis

[1] `https://ossec.github.io`

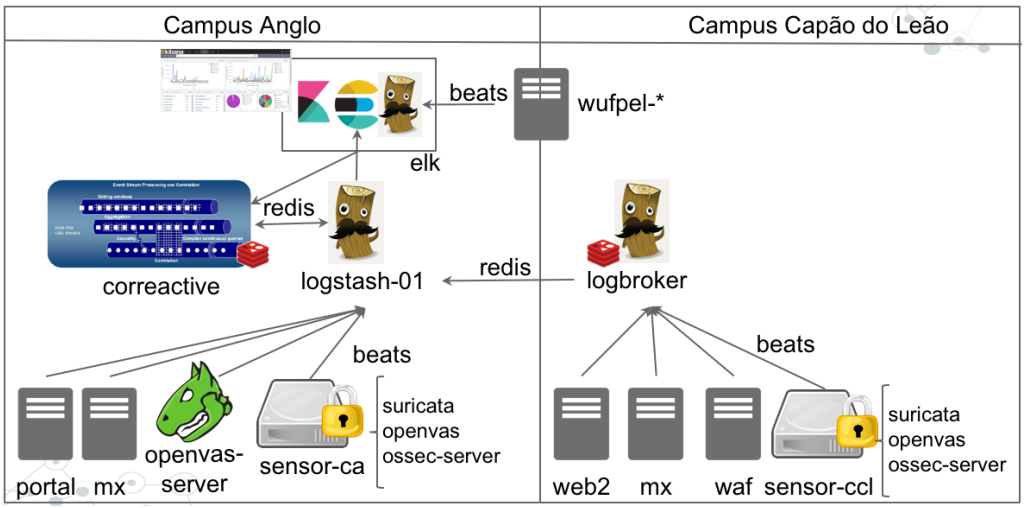

Figura 1. Cenário de implantação da proposta concebida na UFPel

em [Rochford and Kavanagh 2015] e [Shankar 2014], e do estudo de outras possíveis soluções, como por exemplo a SIEMonster[2], a qual não oferece correlação em tempo de execução, foi decidido realizar a concepção de uma nova solução.

A solução adotada teve como base o Elastic Stack[3] (Elasticsearch, Logstash, Kibana e a plataforma Beats). Além de atender boa parte dos requisitos especificados, alguns dos fatores que motivaram a sua adoção são: ser empregada em diferentes soluções de segurança[2]; o crescimento de sua utilização, inclusive por empresas reconhecidas da área de tecnologia como Microsoft, Facebook e Mozilla[4].

Para suprir o requisito de correlação de eventos com sintaxe similar à SQL, foi empregada a solução CorReactive[5], a qual é baseada na biblioteca de Processamento de Eventos Complexos (CEP) denominada Esper (empregada pela solução RSA Analytics), sendo suas únicas desvantagens o fato de não ser de código aberto e de possuir limitações de escalabilidade. Já para contemplar a instalação facilitada, em especial do Filebeat e do agente do HIDS OSSEC, o software de automação de provisionamento Ansible[6] foi empregado. Esta escolha justifica-se pois o Ansible permite o gerenciamento centralizado de dispositivos através de Secure Shell (SSH) e autenticação por chave compartilhada, se adequando bem a infraestrutura já existente na UFPel.

A figura 1 apresenta uma visão geral do ambiente, onde é exibida a disposição dos componentes. Nela, é possível observar a existência de dois servidores denominados de sensores (sensor-ca e sensor-ccl) alocados em cada um dos principais campis. Os mesmos consistem de instalações do suricata, do OSSEC server e instâncias do openvas scanner, sendo responsáveis por coordenar os servidores em seu campus, bem como eventuais servidores em alguns dos demais prédios distribuídos da UFPel.

O logbroker é um servidor que possui como principal função, receber os even-

[2]http://siemonster.com
[3]https://www.elastic.co/products
[4]https://www.elastic.co/use-cases/industry/high-tech
[5]https://sourceforge.net/p/correactive/wiki/FAQ/
[6]https://www.ansible.com/

tos dos dispositivos alocados no CCL e armazená-los em uma fila de eventos suportada pelo Redis. Esta abordagem visa evitar possíveis quedas no link de comunicação entre os prédios. O logstash-01 recebe os eventos dos servidores no CA, busca os eventos do logbroker via Redis e encaminha estes tanto para o Redis que opera no servidor correactive como para o Elasticsearch alocado no servidor elk. Desta forma, o CorReactive é responsável por buscar os eventos do Redis e executar a correlação baseada em regras.

Por meio do CorReactive, as situações podem ser contextualizadas usando a anotação "@Enrichment". Além disso, ele executa ações em função das situações detectadas pelas regras definidas. No cenário da UFPel, alguns scripts foram desenvolvidos para executar modificações temporárias em regras de firewall, sendo executadas por meio de um SSH para o sensor que coordena o servidor que necessita das alterações. Com isso, a atuação é realizadas por meio do componente agent_control do OSSEC server, desonerando o administrador de controlar possíveis "*timeouts*" ou estratégias para recuperação do estado anterior à atuação.

A organização do Kibana e do Elasticsearch no CA foi realizada visando permitir que a equipe de segurança interaja com o componente independentemente da comunicação pela Internet. Adicionalmente, observa-se que no servidor elk foi implantado um logstash para atender os eventos de diferentes servidores que constituem a solução institucional de rede sem fio WUFPel. Isto ocorreu pois nesta instância houve a necessidade de serializar o processamento destes eventos para utilização do plugin aggregate, o qual agrega os dados (como CPF e nome) dos eventos de autenticação do portal de captura junto aos eventos registrados pelo proxy squid.

3. Resultados

Atualmente, a proposta concebida está em produção por aproximadamente 2 anos, onde a última atualização contemplou a integração com o CorReactive, sendo capaz de indexar em torno de 10Gb à 20Gb, o que representa uma média de 30 milhões de eventos por dia. Estes eventos são produzidos por aproximadamente 120 servidores, 45 AP's e 91 *switches*. Os eventos produzidos nos *switches* são coletados via protocolo syslog, enquanto que para coleta das estatísticas dos AP's foi desenvolvido um script em Python que executa comandos em shell em cada dispositivo.

Para coleta das vulnerabilidades identificadas também foi desenvolvido um código que interage com o *OpenVAS Management Protocol* (OMP). Dados armazenados em banco de dados como MySQL (usado no OSTicket) e DB2 (aplicado em um sistema legado) também foram coletados para fins de métricas. O ambiente gera em torno de 25 tipos diferentes de eventos possuindo regras específicas para normalização e contextualização. Alguns arquivos de configuração, incluindo exemplos de regras de correlação do CorReactive podem ser observadas em `https://github.com/exehda-issa/technologies`.

Como resultados práticos, além da conformidade com a legislação vigente (norma complementar 21) e de facilidades para auditoria e análise forense, destaca-se principalmente a detecção e resposta às atividades suspeitas de forma antecipada. Além disso, foi possível a elaboração e visualização tanto de métricas da qualidade dos serviços de rede, como a medição do desempenho/delay do postfix, do antispam e entre outros. Em especial, foi possível gerar métricas de segurança, como por exemplo, os falsos positivos

e negativos gerados pelo antispam através da coleta dos eventos do Webmail utilizado.

4. Conclusão

Com o desenvolvimento desta proposta, conclui-se que ela é capaz de operar considerando os desafios atuais dos ambientes distribuídos, esforçando-se para contemplar os requisitos de ser uma solução sem custo de licenças, e oferecer resiliência, flexibilidade, suporte à heterogeneidade, correlação de eventos com sintaxe similar à SQL e, instalação e configuração facilitada. Apesar do CorReactive não atender diretamente a escalabilidade, foi possível observar que o seu desempenho no ambiente de produção atendeu as demandas da UFPel, e ainda assim, por meio do logstash é possível distribuir os eventos entre diferentes instâncias do CorReactive. Ainda que esta tecnologia não atenda o requisito de ser de código aberto, seu modo de operação e algoritmos envolvidos são conhecidos (basicamente, Esper e Redis).

A transparência resultante da exploração das tecnologias destacadas, ao contrário dos sistemas SIEM disponíveis no mercado, oferece suporte a auditoria, flexibilidade e maior controle sobre o funcionamento da solução, um ponto essencial em se tratando de uma aplicação de segurança. Adicionalmente, por evitar os custos com licenças de software, a solução proposta permite que investimentos no desenvolvimento dos recursos humanos, o que se reflete no amadurecimento da organização de forma geral.

Como pontos interessantes da experiência adquirida com esta proposta, ressalta-se a importância no desenvolvimento adequado das expressões para normalização, uma vez que estas podem prejudicar consideravelmente o desempenho do logstash. Observa-se também que a flexibilidade oferecida pode implicar em problemas caso a mesma não seja empregada de maneira eficaz. Considerando isso, o ambiente computacional e as tecnologias utilizadas devem ser estudadas com cuidado. Para complementar isso, de acordo com a literatura, o monitoramento de segurança é uma tarefa intensiva, e nenhum fornecedor no mercado alega reduzir a demanda por analistas humanos qualificados.

No atual momento, está sendo realizada a migração do OSSEC para o Wazuh, o qual inclui facilidades de gerenciamento por meio de sua integração com o Elastic Stack. Como trabalhos futuros, inicialmente pretende-se integrar a proposta com uma solução que ofereça resposta à incidentes automatizada por meio de workflows/playbooks. Na sequência, será desenvolvida a avaliação de risco combinando a prioridade dos eventos, com a confiabilidade das regras, a criticidade do ativo e informações de *Cyber Threat Intelligence*. Além da geração de relatórios, também espera-se desenvolver suporte a escalabilidade para correlação de eventos (integrando o Esper com Apache Storm ou Spark), e avaliar a integração com algoritmos de inteligência artificial.

Referências

Rochford, O. and Kavanagh, K. M. (2015). Magic quadrant for security information and event management. Technical report, Gartner Group.

Shackleford, D. (2017). Speed and scalability matter: Review of logrhythm 7 siem and analytics platform. *SANS Institute - A SANS Product Review*.

Shankar, V. (2014). Acesso em: Abril de 2018. Clash of the titans - Arcsight vs QRadar. Disponível em: `http://infosecnirvana.com/clash-titans-arcsight-vs-qradar/`.