ESCOLA DE APERFEIÇOAMENTO DE OFICIAIS

Cap QCO Infor RICARDO GARCIA GIORDANO

ANÁLISE DA PILHA *ELASTIC* PARA O GERENCIAMENTO DE REGISTRO DE EVENTOS: UM ESTUDO DE CASO NO 6º CTA

Rio de Janeiro
2016

Cap QCO Infor RICARDO GARCIA GIORDANO

ANÁLISE DA PILHA *ELASTIC* PARA O GERENCIAMENTO DE REGISTRO DE EVENTOS: UM ESTUDO DE CASO NO 6º CTA

Trabalho de Conclusão de Curso apresentado à Escola de Formação Complementar do Exército / Escola de Aperfeiçoamento de Oficiais como requisito parcial para a obtenção do Grau Especialização em Ciências Militares.

Aprovado em

COMISSÃO DE AVALIAÇÃO

_____

CARLOS EDUARDO ARRUDA DE SOUZA – Maj QCO – Presidente
Escola de Formação Complementar do Exército

_____

MAXLI BARROSO CAMPOS – Cap QCO – Membro
Escola de Formação Complementar do Exército

R893 Giordano, Ricardo Garcia
Análise da pilha *Elastic* para o gerenciamento de registro de eventos: um estudo de caso no 6º CTA / Ricardo Garcia Giordano. – 2016.
XX f. ; 30 cm

TCC (Especialização) – Escola de Aperfeiçoamento de Oficiais, Rio de Janeiro, 2016.

Bibliografia: f. 78 - 81.

1. gerenciamento de registro de eventos. 2. Syslog. 3. Proteção dos sistemas de logs. 4. A pilha *Elastic*. 5. Estudo de caso. 6. Resultados. I. Título.

CDD 355.5

Dedico este trabalho à minha esposa Danielly e aos meus filhos Gabriel e Nícolas. Todos são razão da minha existência e persistência.

AGRADECIMENTOS

À minha esposa Danielly pelo companheirismo e compreensão durante toda a fase do CAM/2016.

Ao amigo e orientador Maxli Barroso Campos pelo direcionamento.

Ao amigo Hudson da Silva Alves pela colaboração durante o estudo de caso.

Para uma nau sem rumo, não existe vento favorável (Montesquieu).

RESUMO

O gerenciamento dos eventos (*logs*) de dispositivos de tecnologia da informação (TI) críticos necessitam de tratamento adequado. Dessa forma, a adoção de um processo claro e definido para a realização dessa importante atividade de segurança da informação requer uma agregação de dois fatores: Conhecimento técnico e solução adequada. A utilização de ferramentas que somente realizam a coleta de registros, mas não possuem recursos propícios ao tratamento dos eventos armazenados, de modo que seja possível a indexação para a otimização de pesquisas e o correlacionamento para a vinculação dos eventos, não proporcionará qualquer resultado efetivo para a garantia da disponibilidade, integridade e confidencialidade dos dados que são armazenados ou que estejam em trânsito na infraestrutura de TI. Quanto a esse aspecto, a pilha *Elastic*, solução amplamente utilizada para o gerenciamento de *logs*, agrega três poderosas ferramentas (*Elasticsearch*, *Logstash* e *Kibana*) que, juntas, formam uma única solução capaz de atender plenamente aos anseios de um gerenciamento de eventos eficiente e confiável. Capacidade de geração de gráficos de alto nível, mecanismos de tratamento dos dados de alto desempenho, utilização de expressões regulares nas buscas, emprego de criptografia para a transmissão dos registros, armazenamento indexado dos eventos e baixo consumo de recursos de máquina são algumas das características dessa versátil solução que foram identificadas, durante a aplicação de um estudo de caso no 6º CTA, e descritas nesse trabalho. A fim de atestar a sua eficiência no tratamento de *logs* e visando à validação da proposta, a pilha *Elastic* fora implementada em um ambiente de produção crítico (servidores de páginas de Internet), do qual foram coletados registros para a realização de testes que

pudessem corroborar a viabilidade de adoção dessa solução para o gerenciamento dos registros, gerados por todos os dispositivos e serviços que necessitam de monitoramento. Visando subsidiar o trabalho, foi aplicado um questionário amplo em todas as Organizações Militares de TI que compõem o Sistema de Telemática do Exército (SisTEx) com a finalidade de identificar a solução adotada pelas OM para o gerenciamento de eventos. Nesse sentido, o objetivo desse trabalho consiste em identificar, por meio da aplicação de um estudo de caso, a viabilidade de adoção da pilha *Elastic* como solução para o gerenciamento de registro de eventos por todos aqueles Centros de Telemática que não possuem um mecanismo efetivo para o desenvolvimento dessa tarefa. A análise das respostas da pesquisa, que buscou identificar as soluções, utilizadas pelas demais OM de TI do SisTEx, e seus aspectos funcionais, e dos resultados, coletados durante o estudo de caso, permite concluir que se trata de uma ferramenta de alto nível para o gerenciamento de registro de eventos, uma vez que, a partir dos *logs* coletados e tratados, será exequível a obtenção de informações gerenciais e operacionais que indicarão a ocorrência de possíveis problemas de segurança da informação ou nos serviços e recursos computacionais da infraestrutura de TI.

Palavras-chave: Gerenciamento de eventos. Registros. *Logs*. *Elastic*. *Elasticsearch*. *Logstash*. *Kibana*.

ABSTRACT

The event (log) management of critical information technology (IT) devices require appropriate handling. Thus, the adoption of a clear and accurate process to carry out this important information security activity requires an aggregation of two factors: Technical knowledge and appropriate solution. The use of tools that only perform the collection records, but do not have adequate resources to the processing of stored events, so it is possible indexing for search optimization and correlating to the linked events, will not provide any effective result for ensuring the availability, integrity and confidentiality of data that is stored or in transit across the IT infrastructure. In this respect, the Elastic stack adds three powerful tools (ElasticSearch, Logstash and Kibana) which together form a single solution able to fully meet the aspirations of an security information and event management efficient and reliable. High level graphics generation capacity, high performance data handling mechanisms, using regular expressions in searches, encryption adoption for the transmission of records, indexed storage of events and low consumption of machine resources are some of the characteristics of this versatile solution that have been identified during the implementation of a case study on the 6th CTA. In order to attest to its effectiveness in the logs handling and aimed at validation of the proposal, the Elastic stack was implemented in a critical production environment (web pages servers), witch were collected records for testing that could corroborate the feasibility of adoption of this solution for record management generated by all deices and services that require monitoring. In order to support the work, it applied the comprehensive questionnaire in all military IT organization that make up the Sistema de Telemática do Exército (SisTEx) in order to identify the solution adopted by each OM for event management.

In this sense, the objective this work is to identify , through the application of a case study, the adoption of viability of Elastic stack as a solution for event log management for all those Telematic Centers that not have an effective mechanism for the development of this task. The analysis of the survey responses, witch sought to identify the solutions used by other IT OM of the SisTEx, and its functional aspects, and the results collected during the case study shows that it is a high-level tool for the event log management, since, from collected and treated logs, it will be feasible to obtain managerial and operational information that indicate the occurrence of possible information security problems or in services and computing resources of the IT infrastructure.

Keywords: Event management. Records. Elastic. Logs. Elasticsearch. Logstash. Kibana.

LISTA DE GRÁFICOS

LISTA DE FIGURAS

SUMÁRIO

1 **INTRODUÇÃO**

O advento do circuito integrado e a acelerada evolução dos recursos computacionais impuseram à humanidade uma severa dependência de utilização, cada vez mais elevada, de sistemas microprocessados para o planejamento e o desenvolvimento de grande parte das atividades, relacionada ao trabalho e entretenimento da civilização do século XXI.

Tal dependência é tanta que muitas pessoas simplesmente reduzem, ou até mesmo abdicam, o inter-relacionamento pessoal para viverem vidas virtuais, propiciadas por aplicações que integram as chamadas “redes sociais”.

Contudo, mais preponderante ainda é a sujeição das empresas aos recursos tecnológicos; todavia, nesse caso, tal submissão não ocorre por renuncia a este ou aquele estilo de vida, mas sim por necessidade da mantença das informações de negócio disponíveis, seguros e íntegros, de modo que esses possam ser utilizados por todos os colaboradores para o desempenho das suas atividades para o cumprimento da missão da corporação.

Entretanto, com a mesma velocidade que ocorreu o aumento da produção e consumo de informação, houve, também, a amplificação dos incidentes de segurança de computador. Uma situação que, outrora, era inimaginável passou a fazer parte do cotidiano dos usuários dos recursos tecnológicos: O acesso indevido aos dados, armazenados em sistemas computacionais e hospedados nos mais distantes países.

> “Diferentes empresas têm diferentes tipos de dados sensíveis que precisam de uma atenção especial quando for considerada a melhor forma de proteger e garantir os seus sistemas de informação e ativos de informação.”(MILLER *et al.*, 2011, p. 20, tradução nossa).

Para a proteção de toda essa informação, foram desenvolvidos recursos de proteção que visam à redução das possibilidades de uma intrusão, agregando, assim, camadas e camadas de proteção que geram outras informações que, do ponto de vista da administração e da salvaguarda de um dos bens mais valiosos do atual século, são de suma importância para a prevenção da perda, roubo ou exposição indevida das informações pessoais e corporativas.

Todos os equipamentos que armazenam e transferem dados também geram dados, conhecidos como registros ou simplesmente *logs*. Tais registros são

utilizados para a verificação da situação do funcionamento do *hardware*, para a identificação de acessos de usuários, para a análise do comportamento dos sistemas, para a validação do desempenho dos *links* contratados e para obtenção de uma vasta gama de informações, que são importantíssimas para a garantia da disponibilidade, integridade e confidencialidade de toda informação que necessita de proteção.

Chuvakin, Schmidt e Phillips (2013) registram que *logs*, embora muitas vezes relegados, são uma fonte de informação muito útil para o gerenciamento de recursos computacionais, aplicações de usuário e segurança da informação. Nesse contexto, não é difícil de se imaginar que o gerenciamento de *logs* é uma tarefa de suma importância para a garantida da tríade que rege a segurança da informação (disponibilidade, integridade e confidencialidade), uma vez que os fatores de risco a que estão submetidos os dados armazenados e em trânsito transcendem a indisponibilidade e integridade, visto que a omissão no devido trato acarretará consequências imprevisíveis no caso das suas violações.

Gordon (2015) aponta que *logs* devem ser protegidos contra leitura e escrita não autorizadas por que podem conter informações sensíveis de usuários e, também, pela necessidade de utilização desses registros em questões judiciais.

A grande dúvida para essa incumbência é a escolha da ferramenta que propiciará o melhor resultado na coleta, armazenamento, tratamento e exibição de todos os eventos exportados pelos equipamentos monitorados. A utilização de mecanismos eficientes é fator decisivo para a identificação oportuna de problemas que possam comprometer os fundamentos da segurança da informação, bem como para a proteção da integridade de todos os registros coletados.

O emprego de uma aplicação para o gerenciamento de registro de eventos que seja dotada de poderosos filtros que possibilitem a classificação, relação e correlação de todos os *logs* coletados acaba por se tornar o âmago da segurança da informação, uma vez que só é possível considerar que o ambiente esteja livre de ameças se, de fato, houver o monitoramento efetivo de todos os mecanismos, utilizados na disponibilização e acesso à informação.

> “A abordagem fundamental para registros é que eles devem ser um reflexo preciso da atividade do sistema e, como tal, devem ser protegidos e mantidos por um período de tempo adequado, a fim de fornecer um ponto de referência para a atividade de investigação futura.” (GORDON, 2015, 4

ed., p. 192, tradução nossa).

Destarte, será apresentada neste trabalho a análise da pilha *Elastic* para a atividade de gerenciamento de registro de eventos, que é dotada de três poderosas ferramentas (*Elasticsearch*, *Logstash*, *Kibana*) que propiciam um efetivo gerenciamento de eventos, o que permite a rápida identificação de desvios de padrão, que são indicativos de que um incidente de computador possa estar em curso.

Para a obtenção de resultados reais, a pilha fora implementada num ambiente de produção, onde é gerada uma quantidade significativa de *logs* em razão da grande quantidade de ativos de rede e, também, do elevado acesso às informações armazenadas, consistindo-se, assim, no cenário adequado para a prova de conceito dessa renomada ferramenta, baseada em *software* livre.

Antes da apresentação dos resultados, serão abordados os conceitos intrínsecos ao gerenciamento de eventos e informações de segurança (que inclui o gerenciamento de registro de eventos), uma vez que essas definições são de elevada relevância para o bom entendimento do resultado final.

O gerenciamento de ambientes complexos requer a adoção e utilização de tecnologias capazes de fornecer informações exatas que permitam o emprego de medidas efetivas para a proteção da informação, contudo tais tecnologias não são panaceias que substituirão a expertise daqueles que tem a missão de garantir a segurança da informação.

1.1 PROBLEMA

A análise de registros, coletados e tratados por ferramentas que não possuam uma interface gráfica amigável e com recursos avançados de busca torna o trabalho do analista cansativo, pouco produtivo e suscetível a falhas, uma vez que muitas informações importantes deixam de ser verificas por não serem apresentadas de forma objetiva e clara, fazendo com que muitas ações importantes para a prevenção de incidentes ou correção de vulnerabilidades deixem de ser adotadas.

A utilização de ferramentas de coleta e análise de *logs* que tão simplesmente apresentam os registros, enviados pelos equipamentos e serviços monitorados, de forma ordenada acaba por se constituir em mais um serviço a ser administrado,

deixando de desempenhar a importante função de indicar, de forma precisa e oportuna, os eventos críticos que necessitam de investigação imediata.

No ambiente em que será realizado o estudo de caso com a pilha *Elastic* está em produção o serviço *LogAnalyzer*, que não possui uma interface amigável e nem funcionalidades que permitem a identificação eficiente das razões que geraram o evento, constituindo-se num serviço pouco utilizado para o tratamento de incidentes de segurança.

Em virtude dessas deficiências, esse serviço não fornece o panorama, em tempo real, da conjuntura dos estados dos serviços e equipamentos (parado, sobrecarregado, falhando, etc.) o que impossibilita a rápida visualização da sanidade do nível de segurança, uma vez que não fornece índices estatísticos de quaisquer eventos coletados (erros de acesso, bloqueio de acesso, tentativa de acesso a arquivos e diretórios proibidos, tentativa de acesso a endereços não permitidos, etc.), além de não trabalhar com protocolos mais sofisticados que possibilitam a obtenção de elevado grau de inteligência cibernética.

Restringe-se, assim, a um mero coletor e organizador de registros, exportados pelos equipamentos monitorados, sem contribuir significativamente para a implementação efetiva de mecanismos que contribuam para a melhoria da segurança da informação.

1.2 OBJETIVO

O presente trabalho consiste num estudo de caso, aplicado no 6º CTA, que objetivou a avaliação dos recursos da ferramenta, destinada ao gerenciamento e análise de registro de eventos (*event logs*), *Elastic stack* (pilha *Elastic*), de modo que se pudesse identificar os aspectos que a tornam diferenciada para a coleta, armazenamento e processamento de registros (*logs*), gerados por equipamentos que compõem uma infraestrutura crítica de tecnologia da informação.

De modo a viabilizar a consecução do objetivo geral de estudo, foram formulados os objetivos específicos, abaixo relacionados, que permitirão o encadeamento lógico do raciocínio descritivo apresentado nesta pesquisa:

a) identificar a necessidade do gerenciamento e análise de registro de

eventos;

b) identificar as ameaças a que estão sujeitas os registros de eventos;

c) identificar os mecanismos de exportação dos registros;

d) analisar os resultados obtidos; e

e) concluir quanto à importância da implementação e à efetividade da pilha *Elastic* na coleta, armazenamento, tratamento e exibição dos registros de eventos no contexto do 6º CTA.

1.3 QUESTÕES DE ESTUDO

Se os sistemas operacionais, serviços e *firmwares* geram registros de eventos para relatar uma falha ou qualquer outra exceção, então tudo deverá ser coletado e tratado? A resposta a essa pergunta depende estritamente da Política de Segurança da Informação da organização, uma vez que ela definirá quais os serviços e dispositivos que são críticos e que deverão ser monitorados.

Entretanto, independentemente da existência de uma Política de Segurança da Informação, no que tange à resposta a incidentes de segurança da informação, certos registros são de fundamental importância que sejam coletados, uma vez que sem esses será muito difícil o tratamento de qualquer arbitrariedade que envolva a segurança da informação. Assim, as mensagens que possibilitem o direcionamento da investigação sobre as causas que levaram a um incidente devem ser coletadas, devendo a aquisição abarcar os seguintes tópicos:

a) Identificação de usuários;

b) data e horários de conexão e desconexão bem-sucedidas aos sistemas e dispositivos;

c) endereço IP do dispositivo que originou a conexão;

d) tentativas de acesso a serviços e dispositivos negadas;

e) data, hora e usuário que realizou modificações nos arquivos e configurações dos sistemas e dispositivos;

f) data, hora e usuário que fez tentativa frustrada de modificação de arquivos e configurações dos sistemas e dispositivos;

g) escalação bem-sucedida de privilégios; e

h) tentativas frustradas de escalação de privilégios.

No caso de sistemas de controle de acesso a conteúdo disponível na Internet (*proxies* e filtros *web*), devem ser exportados os registros que informem:

a) A identificação do usuário;

b) endereço IP do dispositivo utilizado para o acesso;

c) data e hora do acesso;

c) endereço acessado;

d) tempo da conexão; e

e) métodos de requisição nos casos em que são utilizados *proxies* reversos.

Todas essas questões representam os eventos que poderão ser coletados e tratados pelo estudo de caso abordado nesse trabalho, uma vez que se consistem em informações mínimas que devem ser analisadas para o esclarecimento de incidentes de rede.

Com o intento de delinear a solução para o objetivo proposto, algumas questões de estudo foram elaboradas:

a) Quais os dispositivos de rede e serviços que devem ser monitorados?

b) Quais os registros (*logs*) que deverão ser exportados para o servidor remoto?

c) Por quanto tempo esses registros (*logs*) deverão ser armazenados?

d) Quais os requisitos computacionais que o servidor de *log* deve ter para que forneça os resultados desejados no menor tempo possível?

e) A solução adotada para a coleta, armazenamento e processamento dos registros (*logs*) deve possuir interface gráfica intuitiva que seja acessível por navegador?

f) A ferramenta escolhida para o gerenciamento de registro de eventos deve possuir recursos gráficos, elaborados a partir de sentenças de busca, que permitam a rápida e fácil identificação de possíveis problemas, relacionados a incidentes de segurança ou a indisponibilidade de *hardware*?

g) É desejável que a solução adotada tenha recursos de *geolocation* para identificação das origens dos acessos aos serviços disponibilizados na Internet?

h) A remessa dos *logs* para o servidor remoto deve fazer uso de mecanismos de proteção que impeçam a interceptação dos mesmos?

As respostas a essas indagações nortearão o desenvolvimento deste trabalho, consistindo-se em verdadeiras sinalizadoras para a obtenção do objetivo proposto.

1.4 METODOLOGIA

Esta seção tem por objetivo descrever o caminho a ser percorrido para que seja alcançado o objetivo proposto.

Serão explicitados os procedimentos adotados nos levantamentos documentais e bibliográficos, bem como na seleção das amostras submetidas a instrumento de coleta de dados utilizado, o qual, para este trabalho, fora escolhido o questionário.

Também, serão apresentados os instrumentos e procedimentos utilizados na análise dos dados, visando ao esclarecimento dos resultados obtidos no presente estudo.

Assim, para facilitar a compreensão do assunto, esta seção foi dividida nos seguintes tópicos: Objeto formal de estudo, amostra e delineamento de pesquisa.

1.4.1 **Objeto formal de estudo**

Este trabalho busca apontar as qualidades e benefícios da utilização de um servidor remoto, implementado com a pilha *Elastic*, destinado ao armazenamento e gerenciamento de registros de eventos (*logs*), gerados por serviços e dispositivos de rede críticos para o provimento de serviços de tecnologia da informação pelo 6º CTA, o qual é dotado de recursos visuais que permitem a rápida e fácil identificação de anomalias de tráfego e de *hardware* que possam implicar incidentes de segurança ou indisponibilidade de serviços publicados na rede corporativa do Exército ou na Internet.

De modo a se evitar a produção de pesquisa redundante, fora elaborado e distribuído a todos os Centros de Telemática e Centros de Telemática de Área uma pesquisa que buscou identificar as soluções adotadas por cada uma das Organizações Militares de TIC do EB e se as soluções utilizadas atendiam,

efetivamente, às necessidades técnicas para a atividade de gerenciamento de registros.

Para que houvesse amplitude e consenso das respostas, o questionário fora respondido, em conjunto, por integrantes das áreas de operação e segurança, uma vez que é de interesse desses dois setores a utilização de uma solução para o gerenciamento de *logs* que produza resultados efetivos, de forma eficiente e com a celeridade desejada.

1.4.2 **Amostra**

A amostra selecionada para o preenchimento do questionário é composta de sete Centros de Telemática de Área e cinco Centros de Telemática, os quais compõem, juntamente e sob o comando Centro Integrado de Telemática do Exército (CITEx), o Sistema de Telemática do Exército (SisTEx).

A utilização de um espaço amostral pequeno se justifica em face da importância desses Centros para as atividades de tecnologia de informação, desenvolvidas pelo Exército Brasileiro, uma vez que é por meio do SisTEx que todas as organizações militares da Força Terrestre se conectam e transmitem as suas informações. Assim, em razão da atribuição de provedor de serviços de TI e em função do grande volume de eventos, gerados pelos recursos tecnológicos que compõem toda a infraestrutura de tecnologia da informação que necessita ser gerenciada, as respostas desse pequeno universo permitirão a identificação de possíveis deficiências no gerenciamento de registro de eventos, as quais poderão ser corrigidas com o objetivo deste trabalho.

Outro fator que torna o espaço amostral muito útil é a dispersão geográfica desses Centros, que se encontram distribuídos por todas as regiões do Brasil, o que possibilita a busca por uma solução que possa atender a todas as demandas regionais sem que seja necessária a readequação de quaisquer recursos da pilha *Elastic*.

Dessa forma, o questionário foi distribuído ao:

a) 1º Centro de Telemática de Área, localizado em Porto Alegre, RS;

b) 2º Centro de Telemática de Área, sediado no Rio de Janeiro, RJ;

c) 3º Centro de Telemática de Área, situado em São Paulo, SP;

d) 4º Centro de Telemática de Área, estabelecido em Manaus, AM;

e) 5º Centro de Telemática de Área, fixado em Recife, PE;

f) 6º Centro de Telemática de Área, localizado em Campo Grande, MS;

g) 7º Centro de Telemática de Área, presente em Brasília, DF;

h) 11º Centro de Telemática, sediado em Curitiba, PR;

i) 21º Centro de Telemática, localizado em Belo Horizonte, MG;

j) 41º Centro de Telemática, estabelecido em Belém, PA;

k) 51º Centro de Telemática, situado em Salvador, BA; e

l) 52º Centro de Telemática, presente em Fortaleza, CE.

1.4.3 **Delineamento de pesquisa**

O delineamento de pesquisa abrangerá as seguintes etapas: busca e seleção da bibliografia existente, coleta e análise dos dados, leituras para aprofundamento do tema, adoção de um estudo de caso no qual ocorrerá a implementação de solução *Elastic Stack* para avaliação da sua efetividade e argumentação e discussão dos resultados.

1.4.3.1 Procedimentos para a revisão da literatura

A fim de se definir alguns conceitos, possibilitar a escrituração do trabalho e, ainda, a fundamentação de um texto argumentativo para se alcançar o objetivo proposto, foi realizada a revisão de literatura nos seguintes moldes:

a) **Fontes de busca**

- Artigos científicos das bases de dados do *Scholar* Google, do LILACS, do SCIELO e do ISI;

- livros do acervo pessoal;

- artigos e *papers* emitidos pelos organismos internacionais de referência no estudo e pesquisa da segurança da informação (OWASP, CERT.CC, NIST, MITRE, OSSEC, SANS *Institute*);

- artigos e *papers* emitidos por empresas líderes no desenvolvimento de

soluções na área de segurança da informação (CISCO, DELL, McAfee,TripWire).

b) **Estratégia de busca para as bases de dados eletrônicas**

A fim de realizar a busca a respeito do assunto, será utilizada a localização de dados eletrônicos por meio de motores de busca na internet. Para otimização da pesquisa, serão utilizados os seguintes termos descritores: *"Log management*",”ELK *stack*", “*Elastic Stack*”, “SIEM”, “*Netflow with* ELK *stack*”, “*Syslog*”, “*Syslog*-ng”, “ELK *stack with* SNMP”, “*Elasticsearch*”, “*Logstash*”, “*Kibana*”, “*Filebeat*”, “GeoIP *with Kibana*”, “GeoIP *with Logstash*”.

c) **Critérios de inclusão:**

- Estudos publicados em português;
- estudos publicados em inglês; e
- discussões de fóruns.

d) **Critérios de exclusão:**

Estudos que não estejam relacionados ao gerenciamento registro de eventos e, também, os que não apresentem relação com os objetivos deste trabalho.

1.4.3.2 Procedimentos Metodológicos

Quanto à natureza, o presente estudo caracteriza-se por ser uma pesquisa do tipo aplicada, por ter por objetivo gerar conhecimentos a partir de um estudo de caso prático, no qual fora realizada a implementação da pilha *Elastic*, os quais serão dirigidos à solução de problemas específicos, relacionados à melhoria da análise e gerenciamento de registros de eventos, coletados dos dispositivos e serviços críticos do 6º CTA, valendo-se para tal do método indutivo como forma de viabilizar a tomada de decisões acerca dos procedimentos a serem adotados para a correção de falhas de operação e de vulnerabilidades de serviços, bem como a para mitigação dos riscos da disponibilização de serviços na Internet.

Trata-se de estudo bibliográfico que, para sua consecução, terá por método a leitura exploratória e seletiva do material de pesquisa, bem como sua revisão integrativa, contribuindo para o processo de síntese e análise dos resultados de vários estudos, de forma a consubstanciar um corpo de literatura atualizado e compreensível.

1.4.3.3 Instrumentos

De modo a propiciar maior respaldo ao trabalho e para que se pudesse identificar os pontos fortes da solução analisada em relação às utilizadas pelos demais Centros, fora realizada uma pesquisa, que visou à identificação das ferramentas utilizadas para o gerenciamento de *logs*, com ênfase na efetividade do tratamento dos registros e na satisfação dos profissionais quanto aos resultados obtidos após o processamento das informações coletadas, ou seja, buscou-se verificar quão eficientes são as soluções empregadas na análise e gerenciamento de registro de eventos.

1.4.3.4 Análise dos dados

A análise dos dados, a partir dos questionários recebidos, foi realizada a partir da tabulação dos quesitos para posterior elaboração de gráficos comparativos. Não houve separação em grupos, uma vez que as perguntas se referem a uma atividade desenvolvida por todos os Centros e para a qual é necessária a utilização de ferramenta específica.

Cabe ressaltar que a utilização dos resultados do instrumento de coleta foi muito útil para que fosse possível a identificação da atual situação do SisTEx na atividade de gerenciamento de *logs* e, também, para que, a partir dos resultados obtidos, houvesse a escolha de uma solução que pudesse ser empregada pelos Centros de Telemática que não dispõem de qualquer ferramenta ou que não estejam satisfeitos com solução adotada.

Todos os questionários remetidos foram devolvidos devidamente preenchidos.

1.5 JUSTIFICATIVA

O crescente uso dos recursos de tecnologia da informação (TI) para o desempenho de atividades operacionais e administrativas fez com que o Exército Brasileiro, por meio do Sistema de Telemática do Exército (SisTEx), composto pelo

Centro Integrado de Telemática do Exército (CITEx) e pelos Centros de Telemática de Área (CTA) e Centros de Telemática (CT), buscasse a centralização dos recursos estratégicos de TI, de modo que fosse possível a implementação de uma gestão mais acurada e, também, a adoção de recursos de segurança, voltados ao aumento dos índices da disponibilidade, da integridade e da confidencialidade – tríade da segurança da informação – dos serviços e recursos de TI, utilizados pela Força para as suas atividades administrativas e operacionais.

A concentração nos CTA e CT, além de visar à economicidade dos recursos financeiros, objetiva dar maior proteção aos dados que trafegam pelos equipamentos de conectividade ou que são armazenados nos servidores de aplicação.

De modo a ter maior controle sobre o acesso a conteúdo disponível na Internet, à medida que os contratos locais de *links* de Internet fossem vencendo, as OM passariam a acessar a rede mundial de computadores por meio dos Provedores Regionais de Internet (PRI), estabelecidos nos CTA e CT e que fazem uso de equipamentos para o controle, filtragem e monitoramento dos acessos.

OS PRI, tal qual sugere o nome, são os serviços que disponibilizam acesso ao conteúdo da Internet de forma segura e controlada com a utilização de recursos modernos de segurança que visam à proteção de toda a infraestrutura de TI pela adoção de filtros de conteúdo e controle de acesso a material contraproducente ao trabalho ou que não estejam em adequação com os valores e costumes defendidos pelo Exército Brasileiro.

Além de prover a conectividade com o mundo exterior, o SisTEx também tem a responsabilidade de hospedar todos os serviços *web*, pertencentes ao domínio eb.mil.br, que são disponibilizados na Internet.

Todas essas atividades, indubitavelmente, geram uma grande quantidade de *logs* que informam anomalias nos serviços e dispositivos de rede, empregados no provimento dos serviços.

Para que sejam cumpridas todas as boas práticas de segurança da informação, as legislações da Força, que tratam sobre a utilização e proteção dos recursos tecnológicos, e também as exigências legais, todos os registros gerados

devem ser coletados e mantidos de forma segura, de maneira que seja garantida a disponibilidade e a integridade dessas informações.

Entretanto, a coleta de *logs* não deve servir somente para cumprir imposições de legislações, mas também para a identificação de problemas no correto funcionamento de equipamentos e serviços e de tentativas de se burlar regras de segurança implementadas.

Na árdua e difícil tarefa de se identificar arbitrariedades nos acessos a serviços disponibilizados nos servidores dos CTA e CT e, também, na utilização da rede mundial de computadores, os técnicos das Divisões de Operação e Seções de Segurança necessitam de mecanismos que lhes forneçam informações precisas do momento dos acontecimentos dos erros e das ilegalidades, de modos que medidas preventivas e corretivas possam ser adotadas a fim de se eliminar o problema de operação ou de se mitigar o risco da ocorrência de um incidente de segurança.

Isso posto, a adoção de ferramentas eficientes e eficazes que façam uso de ambientes gráficos ricos em detalhes que possibilitam a fácil e correta identificação de problemas de funcionamento de equipamentos e de ações arbitrárias é de fundamental importância para o provimento de serviços de qualidade pelos CTA e CT, uma vez que, com a utilização de recursos computacionais robustos, voltados à identificação de ameaças, essas OM de TI, além da adequação legal, poderão garantir maiores índices de disponibilidade, integridade e confidencialidade das conexões e dos serviços disponibilizados.

2 **REFERENCIAL TEÓRICO**

Registros de eventos sempre foram gerados pelos sistemas operacionais, *firmwares* e serviços da camada de aplicação do modelo TCP/IP. Contudo, de maneira geral, esses *logs* somente eram visualizados, *in loco*, quando algum problema ocorria em algum dispositivo ou serviço, de modo que se pudesse identificar a causa da falha.

Com o franco crescimento da Rede Mundial de Computadores e o advento dos serviços baseados na *web*, elevou-se, também, o nível dos riscos da exposição de servidores na Internet, nascendo, dessa forma, a necessidade de se exportar e armazenar os registros em outros servidores que não os que os geraram.

Os sistemas operacionais (SO), baseados em *Unix*, possuem mecanismos nativos para geração e armazenamento de registros, que são guardados em arquivos específicos no próprio servidor. Um dos maiores representantes desses mecanismos é o comando *history*, que lista todos os comandos digitados no terminal pelo usuário e que são mantidos no arquivo *.bash_history*. Esses mecanismos, embora muito úteis para a identificação e solução de problemas locais, não oferecem nenhuma proteção contra adulteração ou remoção, visto que podem ser forjados ou totalmente apagados.

As questões legais também devem ser fortemente consideradas no que tange ao armazenamento e proteção dos registros de eventos, uma vez que, no caso de cometimento de crimes com uso de recursos de tecnologia da informação (TI), a justiça requisitará, ao administrador do serviço, os *logs* necessários para a investigação.

> "*Logs* de computador são importantes em muitos aspectos do mundo de TI. Eles geralmente são usados para solucionar um problema ou para tentar compreender os acontecimentos que tiveram lugar em um momento específico no tempo. Quando os *logs* de computadores são utilizados como prova em tribunal, eles devem ser coletados durante as operações normais dos negócios." (HARRIS, 2013, 6 ed., p. 1053, tradução nossa).

Destarte, como garantir a integridade e a disponibilidade desses registros?

2.2 GERENCIAMENTO DE EVENTOS E INFORMAÇÕES DE SEGURANÇA (SIEM – *SECURITY INFORMATION AND EVENT MANAGEMENT*)

O gerenciamento de eventos e informações de segurança é uma atividade desempenhada por uma solução que tenha a capacidade de coletar, armazenar, tratar e correlacionar eventos, ou registros (*logs*), gerados por serviços e dispositivos de rede monitorados, possibilitando uma rápida e fácil identificação de mudanças de padrão que possam indicar possíveis problemas de incidentes de segurança ou de *hardware*, que poderão conduzir a uma indisponibilidade de serviços.

O termo SIEM é uma junção dos nomes SIM (*Security Information Management*), que é a atividade voltada à análise histórica de registros que já foram coletados e correlacionados sem que fossem acompanhados em tempo real, e SEM (*Security Event Management*), que se refere às soluções que possibilitam o monitoramento, em tempo real, dos registros, gerados pelos serviços e dispositivos monitorados. A junção das atividades, desempenhadas pelas soluções integrantes dessas duas categorias, originou as poderosas ferramentas de SIEM, utilizadas para o monitoramento em tempo real de todos os eventos gerados, com a possibilidade de aplicação de sentenças de buscas (*queries*) otimizadas e mais complexas diretamente na base de registros coletados, o que permite a localização e identificação de quaisquer termos que possam estar relacionados a incidentes de segurança.

> “gerenciamento de eventos e informações de segurança (SIEM) é um termo usado para descrever um grupo de tecnologias que reúne informações sobre controles de acesso e atividades selecionadas do sistema para armazenamento para que possam ser analisados e correlacionados.” (GORDON, 2015, 4 ed., p. 1055, tradução nossa).

Eventos de sistemas e dispositivos de rede devem ser coletados por diversas razões, dentre elas:

a) Atendimento às questões legais e corporativas;

b) para fins de auditoria e funções de gerenciamento de riscos;

c) monitoramento de desempenho dos serviços e dispositivos de rede;

d) monitoramento de tendências;

e) correlação de eventos e análise da causa de origem;

f) investigação e resposta a incidentes de segurança.

Em razão da importância, ferramentas de SIEM e análise de *logs* devem ser dotadas, segundo Gordon (2015), de características como:

a) Armazenar informações brutas a partir de vários sistemas de *logs*;

b) reunir as informações em um único repositório;

c) normalizar as informações para fazer comparações mais significativas;

d) possuir ferramentas analíticas que podem processar, mapear e extrair informações de um alvo; e

e) ser provida de ferramentas de alerta e de relatórios.

2.2.1 **Registro de eventos (*logs*)**

Como parte central da temática do trabalho, os registros de eventos (*logs*) precisam ser definidos e classificados, uma vez que, para a correta compreensão das ações adotadas para o gerenciamento dos *logs*, é necessário o entendimento dos conceitos que envolvem esse tipo de informação.

> "Uma mensagem de *log* é o que um sistema de computador, dispositivo, *software*, etc. gera em resposta a algum tipo de estímulo. O que exatamente os estímulos são muito depende da fonte da mensagem de registo. Por exemplo, sistemas Unix terão mensagens de *login* e *logout* dos usuários, *firewalls* terão mensagens relativas à aceitação ou negação das ACL, sistemas de armazenamento em discos gerarão mensagens de *log* quando ocorrem falhas ou, em alguns casos, quando o sistema percebe uma falha iminente." (CHUVAKIN; SCHMIDT; PHILLIPS, 2013, p. 3, tradução nossa).

De forma mais simples e objetiva, *logs* são os registros de eventos ocorridos dentro da rede ou de um sistema de uma organização, os quais eram, inicialmente, utilizados para a solução de problemas, mas que hoje possuem múltiplas funcionalidades, servindo também para a otimização de sistemas, melhoria do desempenho de redes, registro de ações de usuários e para o fornecimento de informações úteis que possibilitam as investigações de atividades maliciosas.

2.2.1.1 Gerenciamento de registro de eventos (*logs*)

"Gerenciamento de *log* é o processo de geração, transmissão, armazenamento, análise e eliminação de dados de *log* de segurança do computador." (GORDON, 2015, 4 ed., p. 976, tradução nossa).

Gerenciamento de *logs* é essencial e de fundamental importância para garantir que os registros de eventos, gerados por sistemas, serviços e dispositivos de rede, sejam armazenados, com detalhes suficientes, de forma segura e por um período de tempo apropriado, de modo que possam ser utilizados para auditorias, que podem demonstrar a ocorrência de incidentes de segurança, violação de políticas, atividades fraudulentas e problemas operacionais.

2.2.1.2 Classificação dos registros de eventos (*logs*)

As mensagens de *logs* podem ser classificados nas seguintes categorias:

a) Informação: Classe de mensagem que tem por finalidade informar a ocorrência de uma anomalia, como, por exemplo, a reinicialização de um equipamento;

b) depuração: Esse tipo de *log* é gerado pelos *softwares* para indicar aos desenvolvedores e administradores problemas de execução;

c) atenção: Mensagens de atenção são emitidas para informar que há ausência de algo importante, mas que essa falta não implicará impacto no funcionamento do sistema ou equipamento;

d) erro: *Logs* de erros são gerados para informar que erros estão ocorrendo num determinado sistema ou dispositivo, mas, geralmente, sem indicar a causa;

e) alerta: Mensagens de alerta são emitidas quando é identificado algum evento que mereça a atenção do administrador. Esse tipo de mensagens, geralmente, é emitido por equipamentos de segurança quando alguma anomalia é detectada.

2.2.1.3 Como os *logs* são coletados?

> A coleta e transmissão de dados de *log* é conceitualmente simples. Um computador ou dispositivo implementa um subsistema de *log* que pode gerar uma mensagem a qualquer momento quando determinar necessário. A forma exata de como a mensagem será gerada depende do dispositivo. Por exemplo, você pode ter a opção de configurar o dispositivo ou mesmo pode ter sido codificado com uma lista pré-definida de mensagens. Por outro lado, você tem que ter um lugar para onde as mensagens de *log* serão enviadas e coletadas. Esse lugar é geralmente referenciado como um servidor de *logs*. Um servidor de *logs* é um sistema de computador,

> geralmente um sistema Unix ou Windows, onde as mensagens de *log* são coletadas de forma centralizada. (CHUVAKIN; SCHMIDT; PHILLIPS, 2013, p. 4, tradução nossa).

De maneira lacônica, os registros são gerados por subsistemas de dispositivos e aplicações que determinam o momento em que deverão informar algum evento anormal. Já os servidores de *log* são computadores remotos que fazem a coleta centralizada dos registros, gerados e enviados pelos dispositivos e serviços configurados para essa finalidade.

2.2.1.4 Como os *logs* são transmitidos?

A forma mais comum de transmissão de *logs* é por meio de procolos projetados para essa finalidade. Nos sistemas baseados em Unix, essa tarefa é desempenhada pelo *Syslog*, que é o protocolo nativo e padrão para a transmissão de mensagens de registro de eventos. Há, também, os protocolos *Rsyslog* e *Syslog*-ng que podem ser instalados a partir de repositório da versão Linux utilizada. Basicamente, esses protocolos, que pertencem à camada de aplicação, são implementados com base no protocolo UDP (*User Datagram Protocol*), o que proporciona, aos serviços que fazem uso deles, menor sobrecarga para a transmissão. Embora exista versão do *Syslog* para as plataformas *Windows*, a *Microsoft* possui seu sistema proprietário para a transmissão de *logs*, o qual é conhecido como *Windows Event Log*. O empecilho da coleta de registros de eventos com uso desse mecanismo proprietário é o armazenamento em formato restrito à aplicação. Contudo, há aplicações comerciais e de código aberto que fazem a conversão desse formato para o padrão *Syslog*, permitindo, assim, que esses *logs* sejam armazenados num servidor central, utilizado para a coleta de outros serviços e sistemas que façam uso do *Syslog* e seus derivados.

Quanto à transmissão de mensagens de *log* de dispositivos de rede, essa tarefa pode ser desempenhada pelo SNMP (*Simple Netwotk Management Protocol*), que é o padrão para essa atividade quando o dispositivo monitorado não se tratar de serviço. A maioria esmagadora dos dispositivos de rede suporta a transmissão de *logs* com uso do SNMP.

Há também a transmissão de fluxo de dados, que são produzidos por

dispositivos intermediários de rede e fornecem informações sobre a transmissão de dados entre dois pontos finais.

> "Por exemplo, um sistema de cliente em sua rede que solicitou uma página *web* de um servidor na Internet pode produzir centenas de mensagens *syslog* em um roteador existente no caminho, o qual lida com cada pacote individual, mas que produz apenas uma única mensagem de fluxo que inclui informações sobre os dois dispositivos (os endereços IP do cliente e do servidor), a quantidade de dados transmitidos, e o serviço que a conexão utiliza (como HTTP pela porta 80). Fluxo de dados é um método muito útil para reunir informações de alto nível do tráfego que transita em sua rede." (MILLER *et al.*, 2011, p. 56, tradução nossa).

Para a transmissão desse fluxo, pode ser utilizado o Cisco *Netflow*, que é um formato amplamente usado para a coleta de informações de trafego de rede, gerados por uma infraestrutura de dispositivos. A grande maioria das aplicações de SIEM reconhecem o formato desse protocolo sem a necessidade de conversão, o que possibilita a sua utilização para obtenção de informações sobre a sanidade dos dispositivos e, também, sobre as conexões que estão sendo estabelecidas a partir de uma rede.

2.2.1.5 Necessidade da análise de *logs*

A análise de *logs* é uma atividade extremamente necessária e útil para o gerenciamento de serviços, uma vez que a sua adequada realização provê informações relevantes e primordiais para:

a) Depuração de problemas de aplicações e serviços;

b) análise de desempenho;

c) análise de segurança; e

d) análise preditiva.

Sem que a análise de todas essas questões seja realizada, é praticamente impossível identificar problemas que possam impactar na qualidade nos níveis de segurança dos serviços prestados.

2.2.2 **Processo de gerenciamento de registros de eventos**

O gerenciamento dos registros de eventos é um processo que, como tal, deve

atender a uma metodologia que defina claramente as etapas a serem desenvolvidas para a sua execução.

Miller *et al.* (2011) enfatizam que “gerenciamento de *log* é a primeira chave para qualquer solução de SIEM” e assevera que:

> “Se você não estiver coletando pelo menos alguns dos eventos que a rede produz, você não será capaz de extrair qualquer informação a partir desses eventos e, portanto, não conseguirá qualquer gerenciamento de sua segurança, a qual será impossível de se alcançar sem a funcionalidade de SIEM.” (MILLER e*t al*., 2011, p. 55, tradução nossa).

Antes mesmo de se iniciar a implementação de qualquer solução para o gerenciamento de eventos, é necessário que se tenha a definição clara dos limites dentro dos quais a ferramenta irá operar. Isso significa delinear, claramente, não só os serviços e dispositivos que terão os seus registros coletados, mas também quais os tipos de *logs* que serão enviados para a solução de SIEM. Assim, fica evidente que a coleta não ocorrerá em todos os serviços e dispositivos de uma infraestrutura e, dos selecionados, não serão todos os tipos de registros que serão encaminhados para o servidor remoto. Assim, as questões básicas que devem ser respondidas para que sejam identificados os limites de emprego da solução são:

a) Por quanto tempo os registros deverão ser mantidos?

Essa pergunta visa determinar por quanto tempo os *logs* serão retidos. Para a resposta a esse quesito, deverão ser consideradas as questões legais e regulatórias, assim como as políticas de segurança da informação da corporação, de modo que o rotacionamento não prejudique o levantamento de informações no caso de possíveis investigações e respostas a incidentes.

Quanto a essa questão, o Brasil já possui regulação específica que regula o período mínimo de retenção.

> De acordo com o artigo 15 da Lei nº 12.965 (BRASIL, 2014), “O provedor de aplicações de internet constituído na forma de pessoa jurídica e que exerça essa atividade de forma organizada, profissionalmente e com fins econômicos deverá manter os respectivos registros de acesso a aplicações de internet, sob sigilo, em ambiente controlado e de segurança, pelo prazo de 6 (seis) meses, nos termos do regulamento.”

Embora o texto preveja a obrigatoriedade do armazenamento dos registros de acesso para pessoas jurídicas que exerçam a atividade de provedor de aplicações com finalidade econômica, os Centros de Telemática acabam por ter que cumprir a

mesma determinação, uma vez que essas OM do Exército desempenham a atividade de Provedor Regional de Internet (PRI) e, também, pelo contido nos parágrafos 1º e 2º do artigo 15 da mesma Lei, que preveem a mesma obrigatoriedade, quando determinado por ordem judicial ou requisitado por autoridade policial ou administrativa ou o Ministério Público, a todos os provedores que não estiverem sujeitos ao disposto do *caput* do artigo.

b) Quanta informação de registros deve ser retida?

O objetivo dessa questão é a limitação do tamanho dos arquivos de log. É necessário que seja limitado o tamanho dos arquivos que armazenarão os registros coletados dos dispositivos e serviços monitorados, visto que uma pequena quantidade de ativos monitorados pode gerar uma grande quantidade de eventos que, rapidamente, podem ocupar uma porção considerável de um disco rígido.

c) Qual o tipo de *log* que será necessário coletar?

Ativos de rede, sistemas operacionais e serviços geram uma grande variedade de registros (*logs*), que possuem causas variadas, que, de forma geral, informam a situação do *hardware*, falhas de aplicação, *status* de conexão de clientes, escalação de privilégio, conexão em conta de administrador, indisponibilidade de serviços dentre outros eventos.

De acordo com o controle A.12.4.1 da Norma ABNT NBR ISO/IEC 27001 (2013), “registros de eventos (*logs*) das atividades do usuário, exceções, falhas e eventos de segurança da informação devem ser produzidos, mantidos e analisados criticamente, a intervalos regulares.”

> “Os *logs* são os registros mais valiosos que um administrador pode ter, porque neles é registrado tudo o que acontece no sistema. Portanto, um administrador sem seus *logs* não pode determinar o que realmente aconteceu.” (TRIGO; MELO, 2004, p. 58).

Definir os tipos de registros que serão coletados é importante para que haja uma análise eficiente, uma vez que a dificuldade de se analisar registros coletados é diretamente proporcional à quantidade de eventos armazenados, ou seja, quanto maior a quantidade e o tipo de registros coletados, maior será o tamanho do arquivo e maior será a dificuldade de obtenção de resultados precisos que possam elucidar uma determinada questão, seja ela qual for, durante um processo de investigação ou tratamento de incidentes.

A atividade de gerenciamento de *logs* pode se tornar tão complexa quanto à

gestão dos recursos de TI à medida que diversas origens de coleta de registros são adicionados à solução SIEM. Assim, estabelecer limites é de fundamental importância para a obtenção de resultados precisos que possibilitem um gerenciamento efetivo dos eventos e da segurança da informação.

2.2.2.1 Coleta de registro de eventos (*logs*)

Delimitados os limites de operação da solução escolhida, é preciso que sejam definidas as questões que intrínsecas à coleta dos *logs*. Para que ocorra uma coleta eficiente que possibilite a produção de resultados efetivos, Miller *et al*. (2011) propõem alguns quesitos básicos que devem ser respondidos:

a) A partir de quais dispositivos os eventos serão coletados?

b) Quais tipos de eventos serão coletados?

c) Por quanto tempo esses eventos serão mantidos?

d) Onde os eventos serão armazenados?

As respostas a essas questões direcionarão a atividade de coleta de eventos, uma vez que exigem a identificação dos dispositivos e aplicações mais importantes de uma infraestrutura de TI que deverão ser monitorados, o que impedirá a coleta indiscriminada de eventos sem qualquer critério.

2.3 SYSLOG

Em função de o *syslog* ser utilizado pelo *kernel* dos sistemas *Unix* e por muitas outras aplicações tê-lo adotado como padrão para geração de mensagens de eventos, esse mecanismo foi escolhido, combinado com os *logs* do *Squid*, para a elaboração deste trabalho. Assim, para melhor entendimento da sua utilização no estudo de caso, é necessário o conhecimento acerca da sua estrutura e possibilidades de emprego.

Syslog é um método padronizado pela indústria para o registro e informações de eventos, gerados por dispositivos e aplicações. Originalmente especificado pela *Request for Comments* (RFC) 3194 e atualizado pela RFC 5424, foi construído com o propósito de coletar informações sobre depuração de aplicações.

Consequentemente, possui limitações que o tornam inadequado para a coleta de *logs*, destinados a análise de segurança. Independentemente das limitações, esse protocolo se tornou o método mais comum de registro de eventos para sistemas baseados em *Unix*.

Como se tornou um padrão da indústria, o *syslog* é utilizado, também, em muitos sistemas operacionais de dispositivos de rede (*switches*, roteadores, *firewalls*, etc.), que são capazes de encaminhar os eventos gerados para servidores remotos.

Syslog consiste em um *deamon* (*syslogd*) que é inicializado durante a carga do sistema operacional e paralisado aquando do seu desligamento. O *syslogd* é utilizado na classificação de informações, avisos, alertas e demais tipos de mensagens que são geradas pelo *kernel* e outras aplicações.

A comunicação com o *syslogd* é realizada através de chamadas às bibliotecas do *syslog*, que recebe registros de aplicações. Os eventos, gerados pelo *kernel*, são encaminhados por um *socket* do domínio *Unix*.

Dispositivos de rede que fazem uso desse protocolo podem ser configurados para gerar mensagens em função do nível estabelecido. Tipicamente, podem ser definidos os níveis alto e baixo para comunicação de eventos, os quais gerarão, respectivamente, grande e pouca quantidade de mensagens.

O *syslogd*, embora utilizado para o gerenciamento de *logs* locais de aplicações *Unix*, também pode ser utilizado para a coleta de eventos, gerados por outros dispositivos, uma vez que esses são encapsulados em datagramas do protocolo UDP e remetidos para a porta 514 do *daemon* remoto.

O comportamento do *syslogd* é controlado pelo arquivo de configuração, usualmente localizado em /etc/*syslog.conf*. O *daemon* interpreta os eventos a partir dos *sockets* Unix, localizados em /dev/log, e os escreve em arquivos de saída ou os encaminha, via UDP, para o servidor remoto.

Cada mensagem do *syslog* deve seguir o formato previsto na RFC. Essa padronização possibilita a construção de filtros, no servidor de *log*, que permitem uma fácil identificação e correlação de eventos.

Miller e*t al*. (2011) registra que muitos desenvolvedores, mantendo o formato original do cabeçalho, agregam ao corpo do *syslog* outros campos que atendem às

necessidades de suas aplicações, criando mensagens que, na prática, transmitem as mesmas informações, porém, de uma forma diferenciada.

2.4 PROTEÇÃO DOS SISTEMAS DE *LOGS*

Aparentemente, um servidor de *log* não necessitaria de qualquer proteção extraordinária, uma vez que não possui informações importantes que poderiam despertar o interesse de um atacante. Da mesma forma, qual seria a razão para proteger a transmissão de registros de uma origem até o servidor remoto se as informações transmitidas podem ser acessadas localmente?

De fato, um servidor de *log* não possui nada mais que uma quantidade imensa de informações que foram coletadas de outros dispositivos e esses registros, na maioria das vezes, podem ser acessados nas suas origens. Contudo, uma das primeiras coisas que um atacante faz, quando realiza um acesso indevido, é a exclusão ou adulteração de todos os registros que podem indicar a sua ação e a sua localização.

> "Um sistema de análise de *log* robusto depende da integridade dos dados de registro que serão analisados. O sistema tem que ser resiliente às tentativas de modificação ou de exclusão de dados. Além de tudo isso, ele também tem que permitir o controle de acesso granular aos dados de *log*. Se os dados de registros serão usados como evidência em um contexto legal, a capacidade de demonstrar a integridade dos *logs* poderá influenciar na aceitação ou não desses registros como evidência." (CHUVAKIN; SCHMIDT; PHILLIPS, 2013, p. 305, tradução nossa).

Toda essa questão, que é corroborada por Trigo e Melo (2004, p. 59) ao afirmarem que "o administrador deve ter 100% de certeza de que seus *logs* são verdadeiros", conduz a duas situações:

a) O atacante poderá buscar informações acerca do gerenciamento de registros e procurar identificar o servidor remoto para atacá-lo e tentar apagar todos os registros coletados a fim de eliminar, definitivamente, os seus rastros; e

b) a interceptação de *logs*, enviados por dispositivos que não estejam na mesma infraestrutura para um servidor remoto.

Essas questões deixam evidente a necessidade da proteção dos serviços de *log*, de modo que a disponibilidade, a integridade e a confidencialidade dessas informações possam ser garantidas.

> “Há uma variedade de razões para que o atacante objetive os *logs*. Primeiramente, o atacante inteligente quer evitar ser pego. Como os dados de *log* fornecerão evidências de sua atividade, ele deseja evitar que a informação seja encontrada. Um atacante ainda mais experiente desejará esconder seus rastros por meio da desorientação, fazendo com que o observador pense que algo esteja acontecendo. Além de esconder seus rastros, os atacantes podem encontrar informações nos *logs* que sejam úteis por si só, como dados transacionais ou informações que possam auxiliar no ataque a outros sistemas, tais como senhas ou números de conta. Assim, há motivação para que não seja registrado, pois destruirá *logs*, modificará registros, bem como se esconderá e vasculhará *logs*, originados de vários sistemas, para obtenção de conhecimento.” (CHUVAKIN; SCHMIDT; PHILLIPS, 2013, p. 306, tradução nossa).

2.4.1 **Alvos do ataque**

Ataques contra uma infraestrutura de gerenciamento de registro de eventos pode ocorrer em qualquer um dos componentes da estrutura. Assim, são alvos potenciais:

a) A origem dos registros: Corresponde aos dispositivos em que os eventos são gerados;

b) o trânsito: Refere-se aos dispositivos de rede que carreiam os eventos da sua origem até os servidores de *log*;

c) o coletor ou agentes de coleta: Designa os mecanismos utilizados para a coleta dos registros em suas origens;

d) o servidor de *log*: Relaciona-se com os dispositivos, utilizados para o armazenamento dos registros coletados de outros dispositivos;

e) a análise: Contempla os sistemas onde as análises são realizadas e processadas; e

f) o analista: Profissional que verifica os registros, identifica os problemas e toma as decisões.

2.4.2 **Classificação dos ataques contra os registros**

Os ataques, direcionados aos *logs*, podem ser classificados quanto às suas características e finalidades. Dessa forma, as investidas podem ser contra:

a) a integridade, que é a capacidade de um atacante alterar ou corromper informações por meio de manipulação e, também, pela capacidade de inserção de

dados previamente construídos pelo arbitrário;

b) a disponibilidade, que ocorre quando o resultado de um ataque é a exclusão de arquivos de *log*, ou a desativação do mecanismo de registro ou de coleta ou, ainda, a indisponibilidade dos mecanismos de análise; e

c) a confidencialidade, que é a capacidade de um atacante ler os arquivos de *log* na origem, em trânsito ou no destino.

2.4.2.1 Ataques contra a integridade dos registros

Ataques contra a integridade dos *logs* visam ao corrompimento dos dados armazenados. A modificação pode ocorrer pela sobrescrita ou por meio da inserção de dados falsificados.

Chuvakin, Schmidt e Phillips (2013, p. 313) afirmam que atacantes mais experientes removem somente as mensagens de *log* que indicam suas atividades no sistema ou simplesmente as modificam, de modo que informem uma situação normal para que não atraiam a atenção dos analistas.

2.4.2.2 Ataques contra disponibilidade dos registros

Ataques contra a disponibilidade dos registros objetivam a negação de acesso de usuários legítimos aos dados ou aplicações. Alguns ataques contra a disponibilidade são muito semelhantes aos ataques contra a integridade, uma vez que produzem os mesmos efeitos e fazem uso dos mesmos vetores. Em muitas circunstâncias, essa categoria de ataque se caracteriza por uma negação de serviço, visto que acaba por causar falhas nos mecanismos de autenticação, controle de acesso ou no próprio sistema acessado.

> “Muitas ferramentas de *hackers* são escritas para ‘sanitizar’ registros de acesso, o que significa remover secretamente o indesejável, incluindo registros. […]. Essas ferramentas operam de duas distintas maneiras: Ou elas zeram/substituem os registros binários de *log* (enchendo assim os arquivos com zeros, o que é suspeito) ou elas os apagam (tornando os arquivos de *log* mais curtos, o que também é suspeito). Ambos os métodos têm deficiências e podem ser detectados.” (CHUVAKIN; SCHMIDT; PHILLIPS, 2013, p. 319, tradução nossa).

2.4.2.3 Ataques contra a confidencialidade dos registros

> “Ataques à confidencialidade em relação aos dados de *log* não estão relacionados a um intruso escondendo seus rastros, mas sobre a obtenção de inteligência. Essa inteligência pode ser informações sobre sistemas e redes para serem usados em outro ataque, ou pode ser a coleta de informações em si, que é o próprio ataque. Mesmo que os dados em si, obviamente, não sejam sensíveis, a Lei ou a política da corporação podem considerar a exposição dos dados como uma violação.” (CHUVAKIN; SCHMIDT; PHILLIPS, 2013, p. 307, tradução nossa).

O ataque a confidencialidade busca localizar:

a) Quais os tipos de aplicação que estão em produção e em quais dispositivos estão sendo utilizadas. Os *logs* frequentemente informam as versões das aplicações, o que pode informar as configurações adotadas. Essas informações podem ser utilizadas para buscar dispositivos vulneráveis na rede;

b) o que está sendo registrado e o que não está sendo registrado, de modo que seja possível determinar como não ser identificado durante uma intrusão;

c) informações sobre credenciais e privilégios de usuários;

d) localização de dados que podem ser roubados ou mesmo dos registros de transações do sistema.

Cabe ressaltar que são suscetíveis a todos esses ataques todos os integrantes da infraestrutura de gerenciamento de registro de eventos, com exceção do analista, uma vez que, em qualquer ponto da cadeia, pode haver comprometimento das informações geradas, coletadas, armazenadas, tratadas e analisadas.

Logo, a adoção de mecanismos mitigatórios em todo o caminho da informação (utilização de recursos criptográficos para a coleta dos eventos, adoção de senhas robustas para acesso aos dispositivos, atualização dos sistemas e aplicações, etc.) é uma necessidade de primeira grandeza, uma vez que a violação de qualquer um dos pontos comprometerá – e consequentemente invalidará – toda a sistemática de gerenciamento de *logs*.

2.5 A PILHA ELASTIC

A pilha *Elastic*, outrora conhecida como pilha ELK, é um instrumento completo

para análise de *logs* e é constituída pela combinação de três ferramentas de código aberto (*software* livre) – *Elasticsearch*, *Logstash* e *Kibana*. Essa solução busca resolver grande parte dos problemas e desafios inerentes ao gerenciamento de registros (formatos de *logs* não consistentes, *logs* descentralizados e exigências de conhecimento especializado) através da integração dessas três ferramentas de renomado conceito e alto desempenho, que se interagem de forma eficiente, para formar uma única aplicação, conforme demonstra a Figura 1, que disponibiliza recursos avançados para o gerenciamento de registro de eventos.

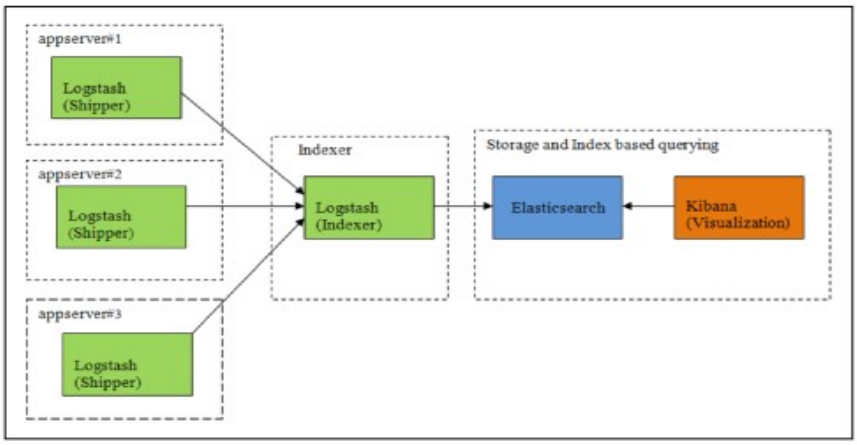


Figura 1 – Esquema de coleta, indexação, armazenamento e busca da pilha *Elastic*
Fonte: Chhajed (2015, p. 8).

2.5.1 ***Elasticsearch***

Elasticsearch é um motor de busca de código aberto, baseado em Apache *Lucene* e liberado sob a licença do Apache 2.0.

Esse sistema fornece escalabilidade horizontal, confiabilidade e alta capacidade para buscas em tempo quase real. Os seus recursos de pesquisa são apoiados no motor sem esquema do Apache *Lucene*. Assim, o *Elasticsearch* é capaz de obter rápidas respostas às pesquisas, uma vez que se utiliza de indexação para a busca e identificação de textos.

O funcionamento dessa ferramenta está baseada nos conceitos *Near Realtime* (NRT), *Cluster*, *Node*, *Index*, *Type*, *Document* e *Shards* & *Replicas*.

2.5.1.1 *Near Realtime*

Elasticsearch é uma plataforma de busca que se aproxima do tempo real. Essa designação é conferida à aplicação em virtude do atraso existente entre o instante em que um texto é indexado e o momento em que o mesmo estará disponível para pesquisa. Segundo o desenvolvedor, a latência existente entre as duas atividades é de um segundo, que não se consiste num tempo que possa comprometer a realização de pesquisas e apresentação de resultados.

2.5.1.2 *Cluster*

Cluster é o agrupamento de um ou mais nós, os quais, no caso do *Elasticsearch*, referem-se aos servidores, que juntos armazenam os dados e fornecem os recursos necessários à realização da indexação e pesquisa unificada em todos os servidores.

No contexto de uso dessa aplicação, cada *cluster* deve ser nomeado de forma única e cada servidor deve pertencer somente a um *cluster* se o mesmo estiver configurado para integrar o *cluster* por meio do seu nome.

É válida a utilização de somente um nó na composição de um *cluster*, assim como a utilização de vários *clusters* independentes com nomes exclusivos.

2.5.1.3 *Node*

Um *node* (nó) é um servidor único que compõe um *cluster*, o qual participa das atividades de armazenamento de dados, indexação e pesquisas, realizadas pelo *cluster*.

Tal qual o *cluster*, um *node* é identificado por um nome exclusivo, que o distinguirá dos demais aquando da composição do *cluster* e, também, para as questões de administração.

2.5.1.4 *Index*

Index (índice) é um conjunto de documentos que possuem características similares, utilizado para a vinculação das informações.

O índice é identificado por um nome, que deve ser escrito com letras minúsculas, sendo o nome atribuído utilizado para se referir ao *index* criado quando for executada a indexação, busca, atualização e exclusão de operações em um documento.

Em um *cluster*, pode haver uma quantidade indiscriminada de índices.

2.5.1.5 *Type*

Type (tipo) é uma categoria ou porção lógica do *index* cujo sentido é definido pelo seu criador. Geralmente, um tipo é definido para um documento que possui uma quantidade de campos em comum. Caso todos esses dados sejam armazenados em um único índice, poderão ser definidos tipo de dados diferenciados para o agrupamento de informações correlatas.

2.5.1.6 *Document*

Document (documento) é a denominação de uma unidade básica de informação que possa ser indexada, cuja representação é feita pelo formato JSON (*JavaScript Object Notation*), que é um formato de dados leve para troca de informações entre aplicações.

Dentro de um índice, é possível o armazenamento de inúmeros documentos, os quais devem ser indexados por meio da utilização de um tipo (*type*) pertencente a um índice (*index*), independentemente de residirem fisicamente no *index*.

2.5.1.7 *Shards* & *Replica*

Um índice pode armazenar uma quantidade muito grande de dados (documentos) que pode comprometer os recursos de *hardware* de um nó. Se um único índice, existente em um nó isolado, que contenha centenas de milhões de documentos exceder a capacidade de um disco rígido de grande densidade, um Tera bytes, por exemplo, poderá ocorrer grande lentidão durante o atendimento de solicitações de pesquisa.

Para que esse problema seja resolvido, o *Elasticsearch* possui a capacidade de subdividir um índice em vários pedaços, chamados *shards* (fragmentos), cuja quantidade a ser utilizada pelo índice é definida durante a sua criação. Cada fragmento é um índice totalmente funcional e independente que pode ser hospedado em qualquer nó de um *cluster*.

Fragmentos são importantes por duas razões:

a) Permite a divisão ou escalação horizontal do conteúdo de um volume;

b) permite a distribuição e paralelização de operações em todos os fragmentos, o que aumenta o desempenho durante o processamento de pesquisas.

Réplicas (*replica*) são cópias de fragmentos (*shards*) de um índice (*index*), que podem ser utilizadas como mecanismo de contorno na ocorrência de alguma falha que ocasione a perda de um fragmento ou que impeça o trabalho de um nó (*node*) de forma *on-line*.

Réplicas são importantes por:

a) Permitirem alta disponibilidade na ocorrência de perda de fragmentos ou falha de nós;

b) possibilitarem o dimensionamento do volume de pesquisa para que sejam executas, de forma paralela, em todas as réplicas existentes.

Por padrão, cada índice no *Elasticsearch* é alocado em cinco fragmentos e uma réplica por nó.

2.5.2 ***Logstash***

Logstash é um motor de coleta de dados, baseado em *software* livre, que possui a capacidade de coletar e analisar, em tempo real, uma grande variedade de dados e eventos, gerados em sistemas estruturados e não estruturados.

Possui a capacidade de unificar, dinamicamente, todos os dados de origens distintas, normalizando-os para um destino especificado, que o faz por meio de *plugins* para conexão a vários tipos de fontes de entrada e plataformas.

É projetado para processar, de forma eficiente, os *logs*, eventos e fontes de dados não estruturados para que sejam distribuídos em uma variedade de saídas por meio da utilização dos seus *plugins* de saída.

Seus recursos-chave são:

a) Processamento centralizado de dados: Auxilia na construção de um canalizador de dados que pode centralizar os seus processamentos. Com a utilização de uma variedade de *plugins* de entrada e saída, o *Logstash* pode converter uma grande quantidade de diferentes fontes de entrada de dados para um único formato comum.

b) Suporte para formatos de dados personalizados: Os diferentes formatos de *logs*, escritos para diferentes aplicações, podem ser analisados e processados em larga escala pela ferramenta.

c) Desenvolvimento de *plugin*: É possível a criação de *plugins* personalizados para a realização de tarefas específicas. Esses *plugins* podem ser publicados, o que garante a existência de uma grande variedade desses mecanismos disponíveis para uso.

2.5.2.1 Configuração genérica de *plugins*

A configuração genérica de um *plugin Logstash* consiste na estruturação de uma série de entradas (*input*), filtros (*filters*) e saídas (*output*), combinada com as suas opções que executam as ações.

Cada *plugin* desenvolve uma função específica na análise, no processamento e na disposição dos dados no formato especificado. De forma simplista, *plugin input* coleta os dados, que serão modificados pelo *plugin filter* e enviados para o seu destino pelo *plugin output* conforme representado na Figura 2.

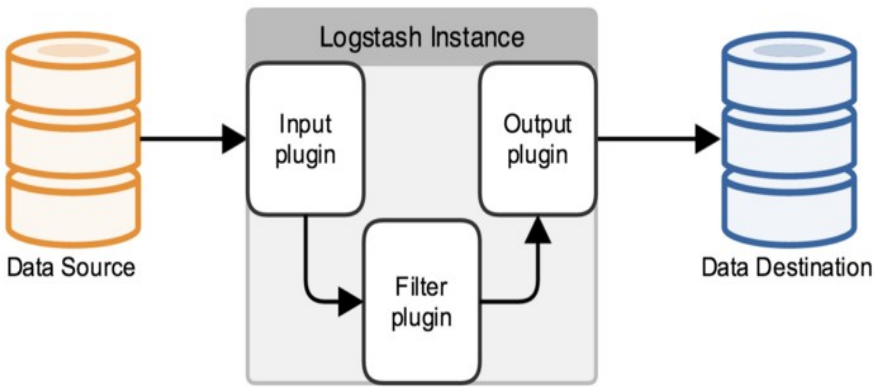


Figura 2 – Esquema geral de funcionamento dos *plugins* do *Logstash*
Fonte: https://www.elastic.co/guide/en/logstash/current/advanced-pipeline.html

Embora os *plugins* estejam relacionados de forma sequencial, não é necessário que esses estejam definidos no mesmo arquivo. É possível a definição da estrutura de cada tipo de *plugin* (Figura 3) em arquivos diferentes. Essa divisão permite a utilização de diversos *plugins* de filtro, de acordo com a necessidade, sem que seja necessária a reestruturação de arquivos que já se encontram em utilização.

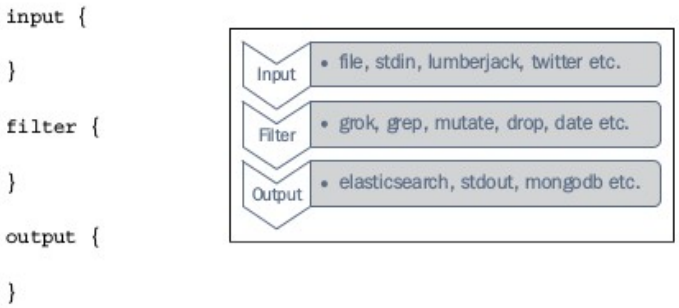


Figura 3 – Estrutura dos *plugins* e suas opções
Fonte: Chhajed (2015, p. 44).

Todos os arquivos que contêm estruturas de *plugin* devem ser armazenados no diretório /etc/logstash/conf.d/ e, com base no conteúdo de cada um, o *Logstash* definirá as suas finalidades. Embora a estrutura do arquivo defina a sua finalidade, o seu nome deve ser exclusivo.

2.5.3 ***Kibana***

Kibana é uma plataforma de código fonte aberto, destinada à visualização e à análise de eventos, projetada para trabalhar em conjunto com *Elasticsearch*.

Essa interface gráfica de alto desempenho possibilita o acesso, interação e busca dos dados, armazenados nos índices *Elasticsearch*, além da elaboração de uma variedade de gráficos, tabelas e mapas com base em pesquisas avançadas que podem ser salvas para utilização futura.

Não bastassem os avançados recursos que propiciam o rápido e fácil acesso aos registros coletados, *Kibana* também oferece a possibilidade de criação de *dashboards* personalizados, nos quais podem ser adicionados todos os recursos gráficos, criados a partir de sentenças de buscas.

A interface gráfica da solução possui três importantes menus (representados

nas figuras 4 e 5), pelos quais se dá o acesso a todos os recursos para manipulação dos eventos armazenados no Elasticsearch: *Discover*, *Visualize* e *Dashboard*.

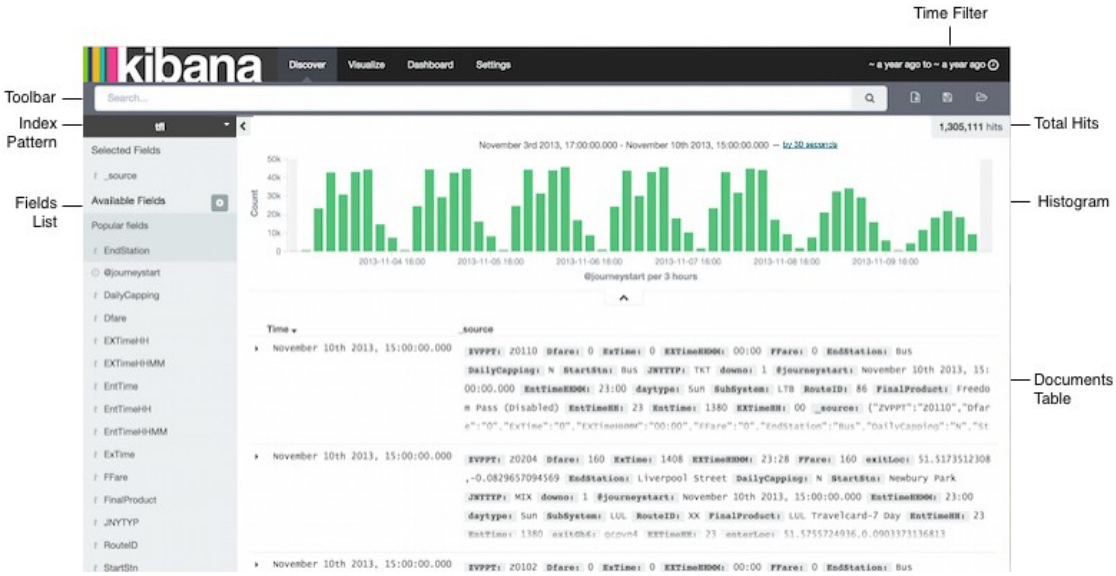


Figura 4 – Representação da tela do menu *Discover*
Fonte: https://www.elastic.co/guide/en/kibana/current/discover.html

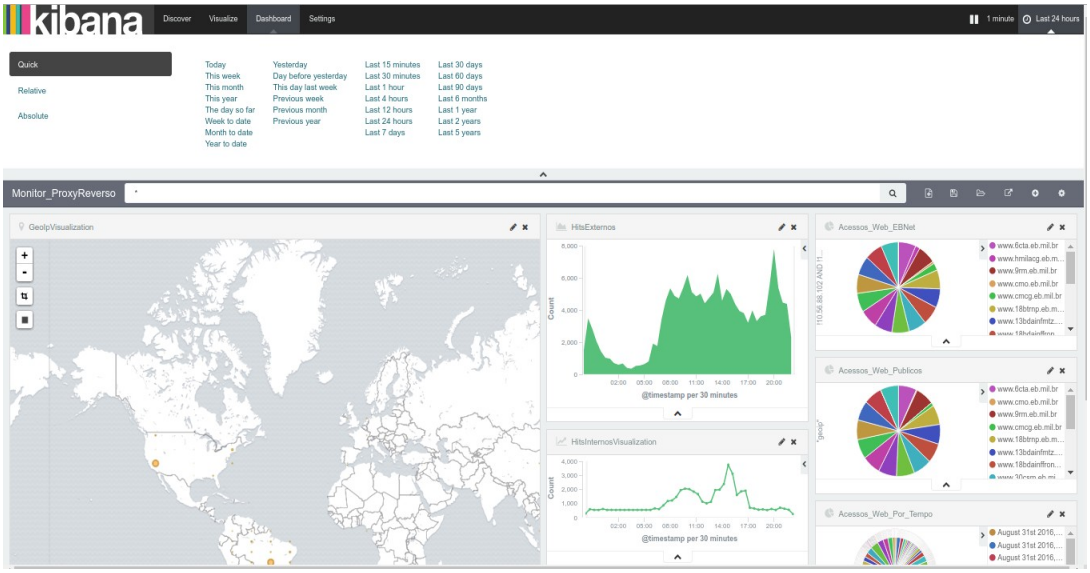


Figura 5 – Representação de um *dashboard* com visualizações gráficas agrupadas
Fonte: O autor.

3 **DESENVOLVIMENTO**

A fim de validar os recursos da pilha *Elastic*, solução escolhida para a implementação de um mecanismo de gerenciamento de registro de eventos, fora elaborado um estudo de caso, no qual fora implementada e colocada em produção a aplicação em questão.

A motivação para a elaboração do estudo de caso advém da necessidade de utilização de uma ferramenta que pudesse substituir a atual solução de gerenciamento de *logs*, adotada pelo 6º CTA, a qual não possui recursos otimizados de busca e nem é dotada de interface gráfica que possibilite a utilização de georreferenciamento (também conhecido como geoIP ou *geolocation*).

Na busca da solução mais adequada para as necessidades do Centro, optou-se pela pilha *Elastic* em virtude da gama de recursos, destinados ao gerenciamento de eventos, provida pelas ferramentas que integram a pilha.

Além do grande poder de processamento e tratamento de *logs* por meio de mecanismos altamente potentes e flexíveis, outro fator que ponderou na escolha da pilha foi a possibilidade de sua utilização sem limitação de quantidade de eventos ou quantidade de tráfego. Muitas outras soluções oferecem recursos semelhantes; contudo, para que seja possível a utilização plena das funcionalidades descritas, é necessária aquisição de licença, uma vez que as versões livres, disponíveis para uso sem custo, possuem restrições que impedem a utilização de todos os seus recursos ou que restringem a quantidade de dados analisados por dia.

Nesta seção, será tratado sobre o cenário, utilizado para o estudo de caso, sobre as questões afetas à implementação da pilha *Elastic* e sobre os resultados obtidos.

3.1 CENÁRIO DO ESTUDO DE CASO

De modo que se pudesse avaliar a pilha *Elastic* quanto às suas capacidades de tratamento de eventos, a sua implementação ocorreu num ambiente de produção que gera centenas de milhares de eventos por dia.

Para um entendimento pleno das circunstâncias nas quais a solução está operando e de modo a se ter uma noção da quantidade de eventos que são coletados, fora elaborado um diagrama, que representa com fidedignidade a topologia de rede, construída para o estudo.

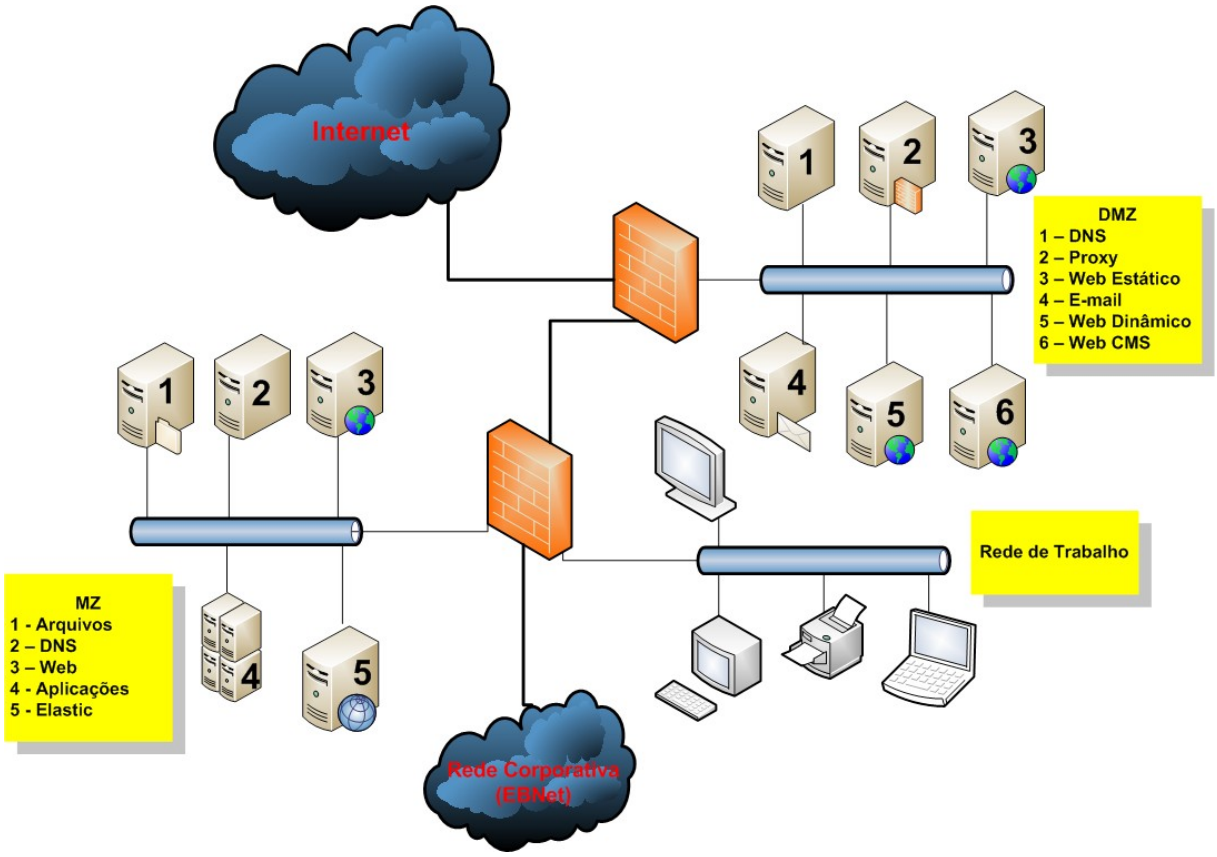


Figura 6 – Cenário de implementação da pilha *Elastic*
Fonte: O autor.

Conforme apresentado no cenário da Figura 6, é possível identificar, claramente, a segregação de redes existentes. Na rede DMZ, encontram-se todos os serviços que estão publicados na Internet e que podem ser acessados por qualquer usuário. Na MZ, estão alocados os servidores que hospedam serviços que são acessados somente a partir da rede corporativa. Na rede de trabalho, estão localizadas as estações, utilizadas pelos militares do 6º CTA. A nuvem da Rede Corporativa (EBNet) representa toda a rede do Exército Brasileiro, pela qual todas as organizações militares estão interconectadas.

Para que haja mitigação dos riscos da disponibilização de serviços na Internet e para a melhoria da segurança da informação, nenhum dos servidores de página (servidores *web* 3, 5 e 6 da DMZ) atendem às requisições dos clientes. Todas as

solicitações de acesso são feitas para o servidor *proxy* (servidor 2 da DMZ), que trabalha de forma reversa, assim, todas as requisições são direcionadas para esse servidor, que as encaminha para os servidores *web*. Uma vez que as conexões dos clientes são estabelecidas com o *proxy* reverso, todas as informações inerentes aos acessos (IP de origem, método HTTP utilizado, código de resposta HTTP, endereço requisitado, arquivos acessados, etc.) estão localizadas nesse servidor, o qual, por sua fez, estabelece uma nova conexão com o respectivo servidor *web* e repassa a solicitação no caso de essa não ter sido bloqueada por alguma das suas regras de segurança.

Em virtude de o servidor de gerenciamento de eventos ser um serviço de acesso restrito, o mesmo fora alocado na MZ (servidor 5 – *Elastic*), que é acessada somente a partir da rede corporativa.

Com a utilização dessa estrutura, fora necessária a adição, nos *firewalls* existentes, de regras que permitissem a passagem do tráfego de encaminhamento dos registros, coletados no servidor *proxy*, uma vez que não é permitido, aos servidores da DMZ, a inicialização de qualquer conexão com outros servidores.

3.2 ASPECTOS DA IMPLEMENTAÇÃO

Durante a fase de implementação da solução, foram constatados alguns aspectos que merecem destaque e registro, uma vez que se constituem em questões determinantes para a adoção da ferramenta por outros interessados na utilização da pilha *Elastic* para o gerenciamento de registro de eventos.

3.2.1 **Óbices**

A instalação de todas as ferramentas que compõem a solução completa é bastante simples e pode ser realizada através de repositórios – mantidos pela *Elastic*, desenvolvedor do *Elasticsearch*, *Logstash* e *Kibana* – bastando para tal somente a adição, no arquivo de informações de repositório da distribuição Linux utilizada, do endereço do servidor e da versão a ser instalada.

Após a instalação do *Logstash*, fora averiguado que a ferramenta

disponibilizava a porta, definida em seu arquivo de configuração, no modo de escuta somente para a versão 6 do protocolo de Internet (IPv6). Tal situação está demostrada na figura 7.

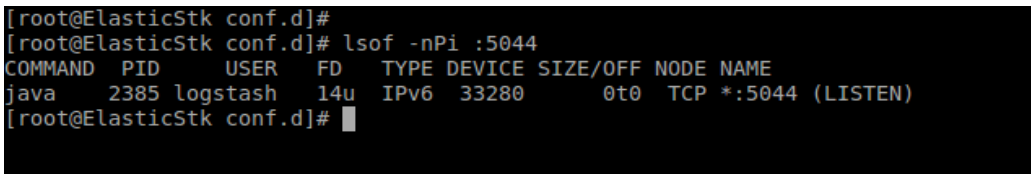


Figura 7 – Porta no estado “escuta” somente para IPv6
Fonte: O autor.

Para solucionar essa questão, fora necessário remover o módulo IPv6 da distribuição Linux utilizada (CentOS), o que possibilitou a disponibilização da porta para o IPv4 (Figura 8), uma vez que é a versão amplamente utilizada para o endereçamento dos dispositivos de rede. Ressalta-se que a remoção do citado módulo não gerou nenhum impacto ao correto funcionamento do sistema operacional.

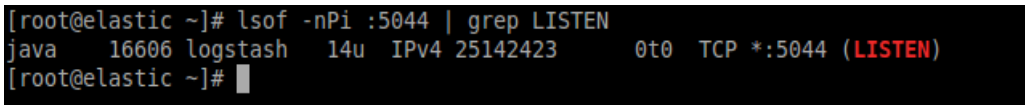


Figura 8 – Porta no estado de “escuta” para o IPv4
Fonte: O autor.

Outro aspecto que requer bastante atenção é a geração dos certificados SSL. Durante essa fase, é possível optar pela geração de certificados com base no nome do serviço ou no seu endereço IP. Evidentemente, por questões de administração de rede, utilizar-se da primeira opção durante a geração do certificado seria a opção mais sensata. Todavia, caso ocorra algum problema com o servidor de tradução de nomes (DNS), o dispositivo monitorado – no qual será utilizado o certificado para a encriptação dos registros a serem enviados ao servidor remoto – não conseguirá enviar os *logs* em razão da impossibilidade da tradução do nome constante no certificado utilizado, visto que não poderá obter o IP do servidor remoto. Assim, a escolha da forma de geração de certificados dependerá muito do índice de disponibilidade e da confiabilidade do serviço DNS.

3.2.2 **Proveitos**

Implementada e colocada em produção, logo no início de sua utilização, foi possível atestar o poder e a efetividade da pilha *Elastic* para o gerenciamento de registro de eventos em razão das funcionalidades disponibilizadas. As vantagens de utilização da solução ficam mais evidentes ainda quando comparadas com a antiga solução utilizada (*LogAnalyzer* – Figura 9).

Indubitavelmente, as maiores vantagens da utilização da pilha *Elastic* é a rapidez das consultas e a forma de exibição dos eventos coletados, proporcionadas pela maneira como os dados são tratados conforme descrito no tópico 2.5.

Fruto do armazenamento dos *logs* no *cluster* de forma indexada, a busca por registros é muito mais fácil e rápida, uma vez que a indexação prévia pela ferramenta *Logstash* agiliza a identificação dos registros que correspondem ao termo, utilizado na pesquisa.

Através da tela inicial do *Elastic* (Figura 10), é possível visualizar, sem a implementação de qualquer regra específica ou ajustes extras, a periodicidade com que os dispositivos monitorados estão enviando os *logs*. Esse comportamento permite a rápida identificação de problemas no correto funcionamento de um serviço, uma vez que, havendo a interrupção da transmissão dos registros de eventos, o histograma inicial deixará de exibir a relação temporal dos *logs* coletados.

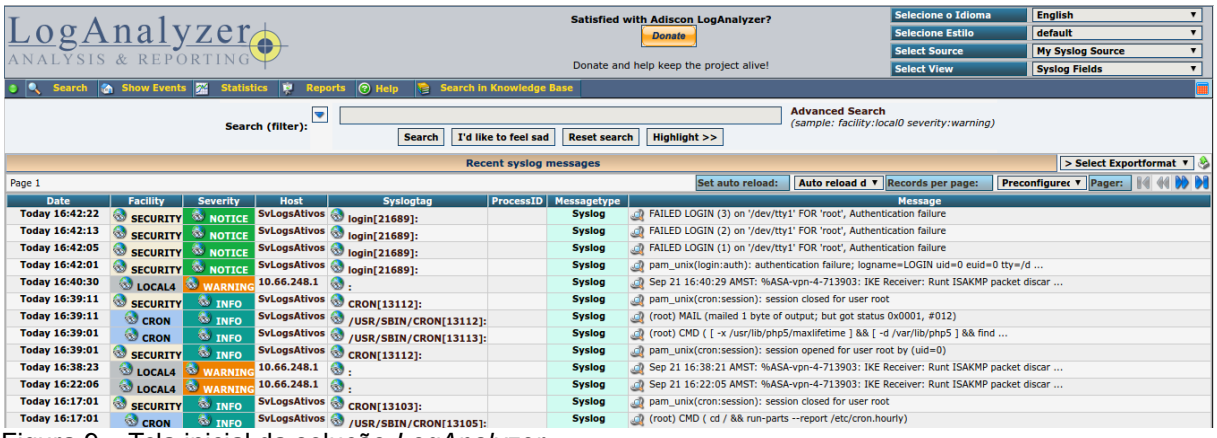


Figura 9 – Tela inicial da solução *LogAnalyzer*
Fonte: O autor.

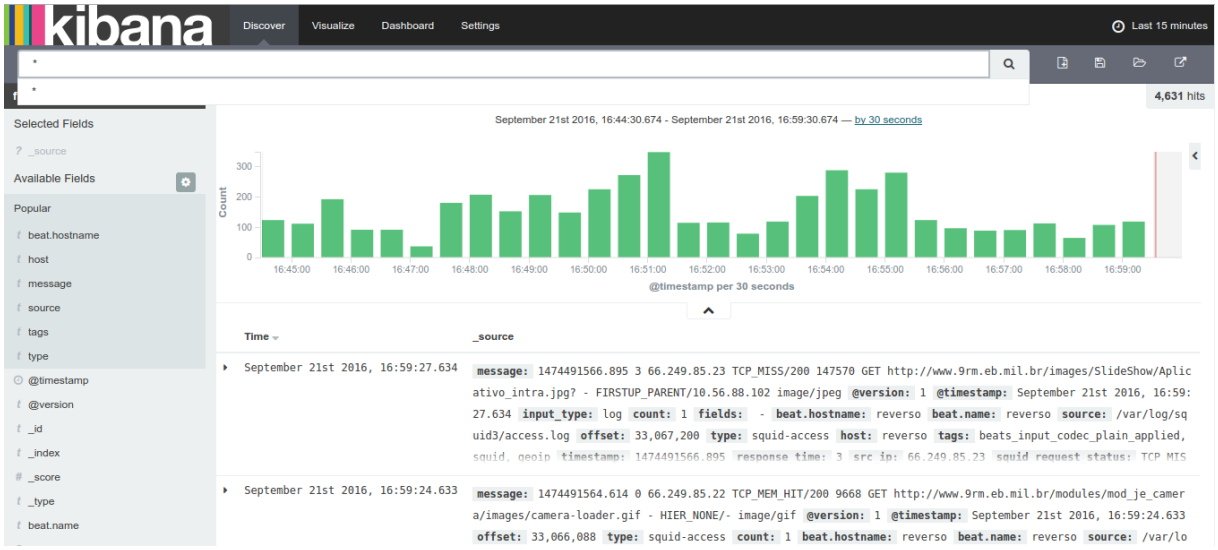


Figura 10 – Tela inicial da solução *Elastic Stack*
Fonte: O autor.

A disposição dos detalhes dos registros de eventos coletados é outro diferencial. A pilha *Elastic* dispõe, de forma clara, todas as informações constantes nos campos do protocolo de coleta utilizado e, com a possibilidade de utilização de georreferenciamento (Figura 12), é possível identificar as origens das conexões. Já no *LogAnalyzer*, as informações dos *logs* não são apresentadas de forma clara e, também, não está disponível a utilização do recurso de GeoIP (Figura 11).

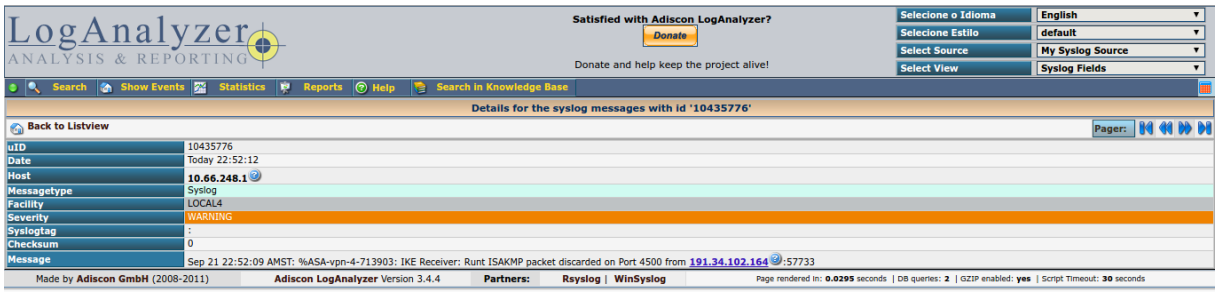
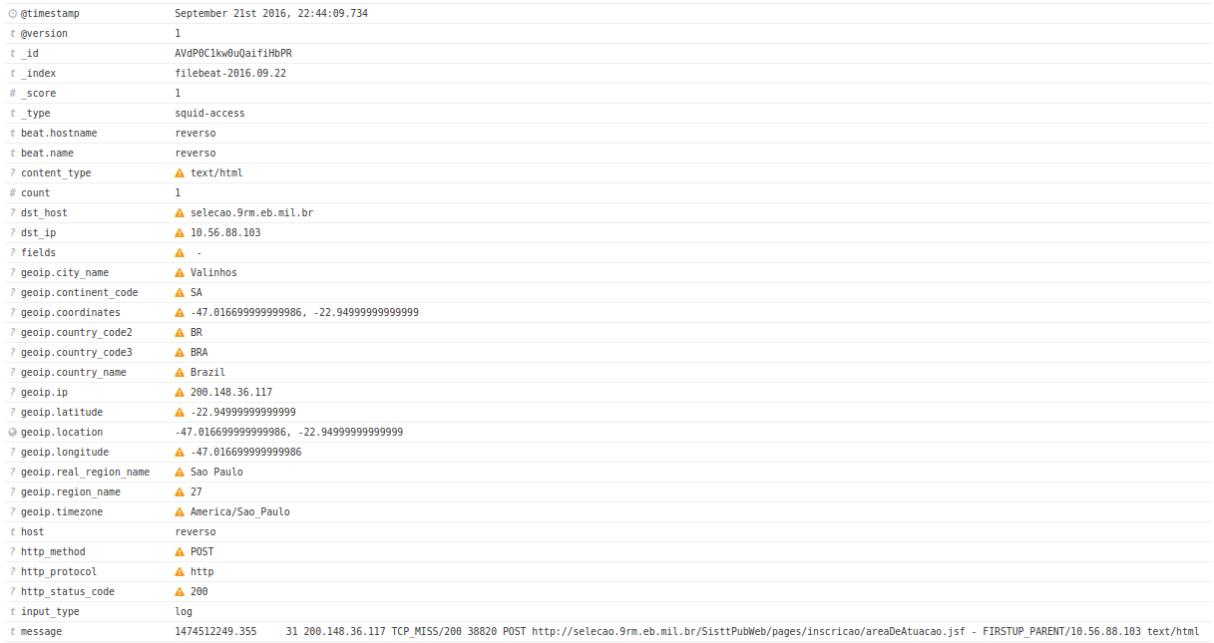


Figura 11 – Tela do *LogAnalyzer* com detalhamento do evento coletado
Fonte: O autor.



@timestamp	September 21st 2016, 22:44:09.734
@version	1
_id	AVdP0C1kw0uQaifiHbPR
_index	filebeat-2016.09.22
_score	1
_type	squid-access
beat.hostname	reverso
beat.name	reverso
content_type	text/html
count	1
dst_host	selecao.9rm.eb.mil.br
dst_ip	10.56.88.103
fields	-
geoip.city_name	Valinhos
geoip.continent_code	SA
geoip.coordinates	-47.016699999999986, -22.94999999999999
geoip.country_code2	BR
geoip.country_code3	BRA
geoip.country_name	Brazil
geoip.ip	200.148.36.117
geoip.latitude	-22.94999999999999
geoip.location	-47.016699999999986, -22.94999999999999
geoip.longitude	-47.016699999999986
geoip.real_region_name	Sao Paulo
geoip.region_name	27
geoip.timezone	America/Sao_Paulo
host	reverso
http_method	POST
http_protocol	http
http_status_code	200
input_type	log
message	1474512249.355 31 200.148.36.117 TCP_MISS/200 38820 POST http://selecao.9rm.eb.mil.br/SisttPubWeb/pages/inscricao/areaDeAtuacao.jsf - FIRSTUP_PARENT/10.56.88.103 text/html

Figura 12 – Tela do *Elastic* com detalhamento do evento coletado
Fonte: O autor.

Outra vantagem muito importante sobre a antiga solução é a utilização de *dashboards*. A possibilidade de criação de múltiplos painéis de instrumentos, que podem agregar diversos tipos de gráficos e mapas, permite que cada profissional faça o monitoramento dos dispositivos sob a sua responsabilidade da forma que melhor lhe convier, uma vez que não há nenhuma obrigatoriedade de que *dashboards* tenham os mesmos conteúdos conforme apresentado na Figura 13.

Quanto aos gráficos e mapas, que podem ser originados a partir de pesquisas específicas que podem ser salvas para uso futuro, cabe ressaltar a interatividade desses mecanismos gráficos de visualização circunstancial, que ocorre através do clique do mouse. Após selecionada uma área de um determinado gráfico, serão exibidas as informações acerca dos eventos que são representados pela visualização gráfica para aquele período de tempo. Esse recurso permite refinar ainda mais uma pesquisa, em função de um lapso temporal, sem a necessidade de sua edição, o que facilita a identificação de distorções e disparidades em relação a uma linha base estabelecida.

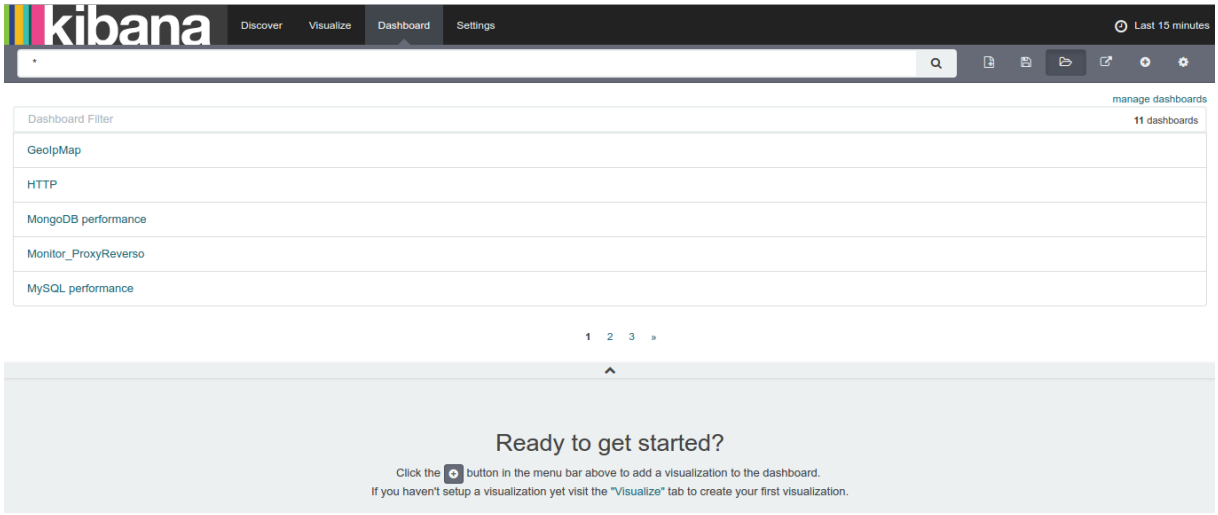


Figura 13 – Tela para escolha dos *dashboards* disponíveis
Fonte: O autor.

Todos esses recursos conduzem ao amplo atendimento das necessidades dos envolvidos na atividade de monitoramento com a utilização de coleta de registros de eventos, visto que, a partir de uma base única de *logs*, é possível a obtenção da consciência situacional do serviço ou dispositivo monitorados de forma personalizada e tempestiva.

3.2.3 **Benefícios**

Os benefícios advindos com a introdução da pilha *Elastic* para o gerenciamento de registro de eventos foram imensamente positivos, uma vez as funcionalidades oferecidas, aliadas à sua robustez e estabilidade, possibilitaram a sua utilização para o monitoramento dos dispositivos e serviços de maior criticidade, existentes na infraestrutura de TI do 6º CTA.

Todos os profissionais envolvidos com a operação passaram a fazer uso do *Elastic*, com utilização de *dashboards* individuais, para o acompanhamento dos serviços e dispositivos de suas responsabilidades.

O pessoal da área de segurança passou a contar com uma solução que permite a rápida identificação de desvios de padrões de conexão nos serviços disponibilizados para acesso público.

Outro grande benefício foi a maior integração de todos os profissionais na discussão de problemas, identificados a partir da análise dos registros coletados.

Uma vez que todos os implicados possuem o rápido e fácil acesso ao *log* que indica alguma anomalia, o debate em torno das possíveis causas e das soluções exequíveis se torna mais célere e efetiva, promovendo uma maior sinergia entre as equipes de trabalho.

O aumento da capacidade de gerenciamento e análise dos registros de eventos coletados foi exponencial, visto que, com a utilização da pilha *Elastic*, a atividade se tornou mais precisa e eficiente, uma vez que a solução dispõe de todos os recursos necessários para o adequado desempenho dessas atividades.

4 **DISCUSSÕES E RESULTADOS**

Nesta seção, serão discutidos os resultados obtidos com a implementação da pilha *Elastic* no estudo de caso.

As considerações abordarão as questões afetas aos recursos descritos nos tópicos relacionados à solução, assim, serão apresentados os resultados quanto à facilidade de implementação e utilização, ao mecanismo de busca de registros armazenados, à forma de apresentação dos eventos coletados, à segurança na coleta e no tratamento dos *logs* e ao desempenho e recursos da interface gráfica.

Todavia, para que seja possível a mensuração da importância da atividade de gerenciamento de registro de eventos, será apresentado um levantamento, obtido a partir da aplicação de um questionário nos demais CTA e CT, das soluções adotadas para o gerenciamento de registro de eventos e algumas considerações pertinentes aos diferentes aspectos de implementação, o que possibilitará a identificação dos fatores que tornam a pilha *Elastic* viável de implementação nas diferentes OM de TI do EB como solução padronizada para a atividade de gerenciamento de eventos.

4.1 DISCUSSÕES ACERCA DA UTILIZAÇÃO DE OUTRAS SOLUÇÕES NOS DEMAIS CTA E CT PARA O GERENCIAMENTO DE REGISTRO DE EVENTOS

Neste tópico, serão apresentados os gráficos, originados a partir das respostas do questionário constante no Apêndice A, e discutidas as circunstâncias a que estão relacionados.

O primeiro quesito trata da existência de solução para o gerenciamento de eventos. De acordo com o Gráfico 1, 75% dos Centros fazem uso de uma solução e 25% não possuem nenhum sistema de coleta de *logs*. Embora a quantidade dos que não possuem solução para o gerenciamento de eventos seja pequena (3), essa situação é particularmente preocupante, uma vez que todos os CTA e CT possuem infraestruturas de TI, voltadas à disponibilização de serviços que interagem com a Internet, situação essa que demanda ainda maior necessidade do gerenciamento dos eventos para a identificação de problemas, afetos à segurança da informação.

Gráfico 1 – Utilização de solução para o gerenciamento de registro de eventos
Fonte: O autor.

Ainda no quesito 1, fora requisitada a informação da solução adotada pelo respectivo Centro. As respostas revelaram, conforme demostradas no Gráfico 2, que somente duas soluções estão em uso: *LogAnalyzer* e OSSIM (*Open Source Security Information Management*).

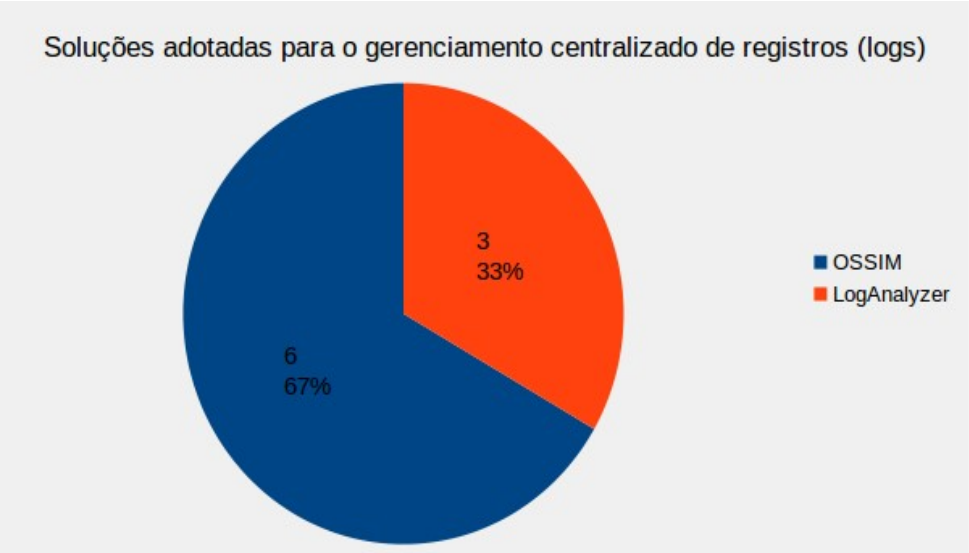


Gráfico 2 – Soluções adotadas para o gerenciamento de registro de eventos
Fonte: O autor.

O elevado número de Centros que fazem uso do OSSIM é fruto das experiências adquiridas pelo longo tempo de utilização da ferramenta, assim, a sua ampla utilização decorre da possibilidade de troca de informações, entre os seus usuários, das melhores práticas para a sua implementação, configuração e aprendizado de funcionamento. Se por um lado a utilização em massa de uma

determinada solução para o gerenciamento de *logs* possibilita a criação de uma rede de suporte para a elucidação de dúvidas, por outro, novas tecnologias deixam de ser testadas, impedindo, assim, a adoção de soluções mais eficientes, o que impacta diretamente na melhoria contínua da atividade de gerenciamento de registro de eventos.

A segunda pergunta aborda a existência de uma interface gráfica intuitiva que permita o rápido e fácil acesso aos recursos, disponibilizados pela solução, para o gerenciamento dos registros. Quanto a esse aspecto, conforme demonstrado no Gráfico 3, 56% dos usuários consideram que a solução utilizada seja dotada de tal recurso e 44% acreditam que a interface gráfica das suas ferramentas não atendem a esse quesito. No tocante à disponibilidade de ambientes gráficos que facilitam o acesso aos recursos da solução utilizada, cabe ressaltar que a existência dessa facilidade é um grande fator de motivação para a exploração dos potenciais da aplicação, o que permite a identificação das melhores formas de análise dos registros coletados.

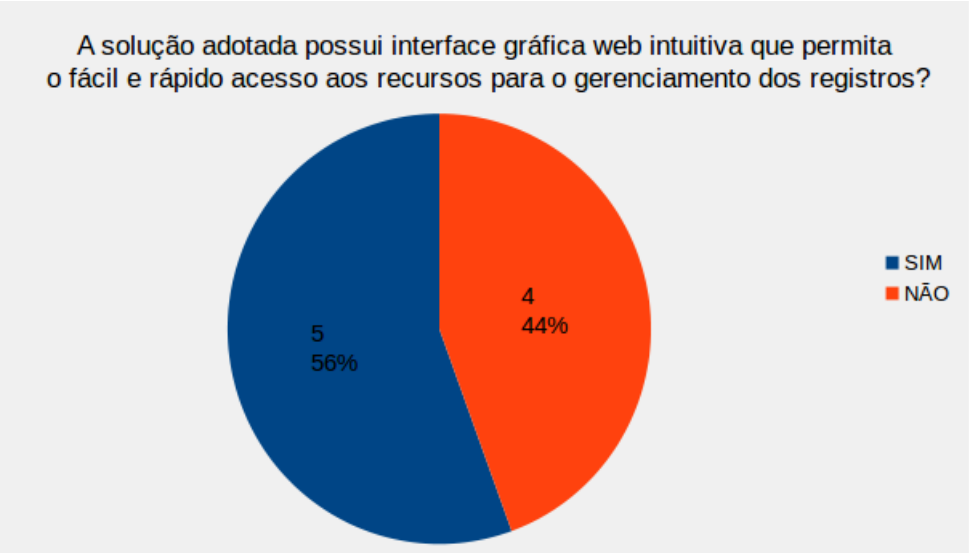


Gráfico 3 – Utilização, pela solução adotada, de interface gráfica intuitiva
Fonte: O autor.

No que concerne a terceira questão (Gráfico 4), apenas 3 (33%) Centros afirmaram que as suas soluções para o gerenciamento de *logs* fazem uso de certificados SSL para a proteção dos registros durante a transmissão para o servidor remoto. Em relação às ferramentas que formam esse percentual, duas são OSSIM e uma é o *LogAnalyzer*. A disparidade do percentual das respostas negativas quanto à

disponibilidade desse recurso na ferramenta mais utilizada chama atenção, uma vez que essa situação evidencia a baixa exploração e emprego desse importante recurso de segurança.

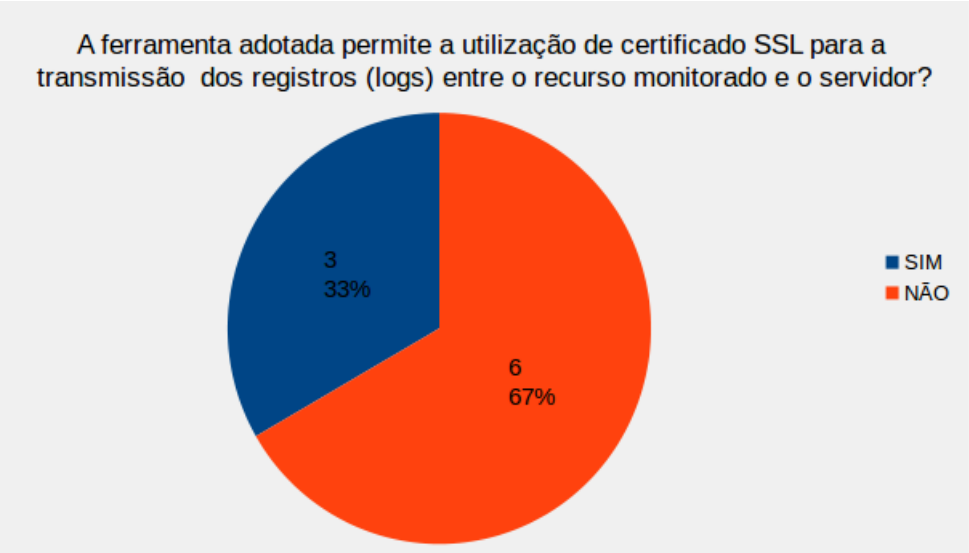


Gráfico 4 – Utilização de recursos criptográficos para proteção dos *logs*
Fonte: O autor.

Quanto maior as possibilidades existentes para a identificação de um determinado evento, maiores serão as chances de obtenção de resultados mais precisos acerca das causas que o originaram. Destarte, se a solução para gerenciamento de registro de eventos dispõe de várias formas de pesquisa, múltiplas também serão as possibilidades de se lograr resultados mais confiáveis.

Nesse contexto, consoante o gráfico 5, 67% das soluções adotadas para o gerenciamento de *logs* não permitem a utilização de outros mecanismos de busca senão os disponibilizados pela própria ferramenta. Essa situação limita a ação de análise dos eventos às circunstâncias estabelecidas pela solução, impedindo, assim, que ocorra a exploração dos registros de acordo com a expertise do profissional que a utiliza para a atividade de gerenciamento e análise de registro de eventos.

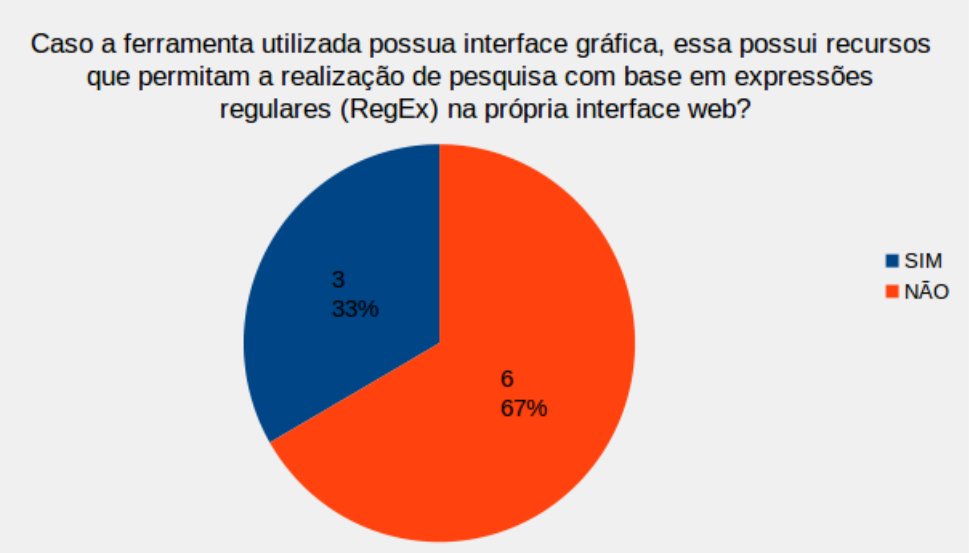


Gráfico 5 – Utilização de RegEx na interface *web* da solução adotada
Fonte: O autor.

A pergunta 5 se refere à possibilidade de persistência de sentenças de busca (*queries*), utilizadas para a localização de eventos específicos, para uso futuro. A ausência desse artifício impõe o trabalho redobrado, já que, na necessidade de realização de um mesmo tipo de pesquisa mais de uma vez, toda a expressão terá que ser reconstruída, o que pode conduzir a resultados diferentes num eventual erro de sintaxe ou de passagem de parâmetros.

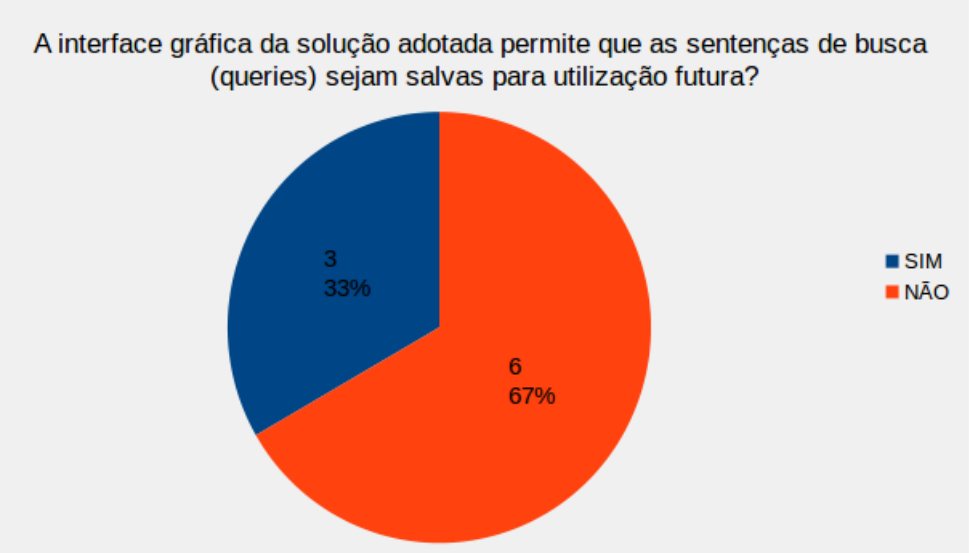


Gráfico 6 – Possibilidade de disponibilização das *queries* para uso futuro
Fonte: O autor.

Em relação a essa questão, segundo é explicitado pelo Gráfico 6, 67% das ferramentas não são dotadas desse recurso. O elevado percentual é indicador de

que outros mecanismos (editores de texto por exemplo) estão sendo utilizados para registro das sentenças construídas para uma circunstância.

O quesito seis busca identificar a possibilidade de utilização de georreferenciamento para identificação das origens das conexões públicas. De acordo com o Gráfico 7, somente 33% das ferramentas não possuem esse recurso. Esse percentual se refere à solução *LogAnalyzer*.

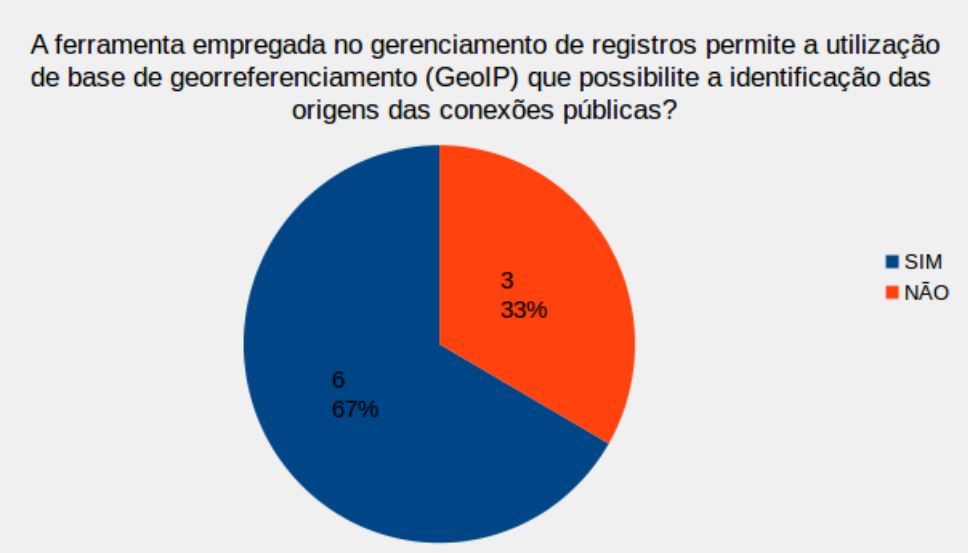


Gráfico 7 – Utilização do recurso de GeoIP pela solução adotada
Fonte: O autor.

O reconhecimento de diversos formatos de registro de eventos é fator decisivo para a escolha de uma solução, visto que, na grande maioria das infraestruturas de TI, há uma miscelânea de equipamentos que geram *logs* nos mais diferentes formatos. A pergunta 8 visa à identificação da capacidade das soluções adotas de reconhecerem os mais variados formatos de registros. Embora o Gráfico 8 indique que 11% das ferramentas não possuem a capacidade de reconhecer os mais variados formatos em razão do número absoluto que representa esse percentual, é possível inferir que a ferramenta que representa o valor do gráfico não tem a capacidade de reconhecer algum formato muito específico, uma vez que a resposta à questão é de um Centro que indicou utilizar o *LogAnalyzer* para o gerenciamento de registro de eventos. Essa possibilidade aumenta quando se analisa o percentual das ferramentas que reconhecem os principais formatos (89%), no qual está incluso também a supramencionada ferramenta.

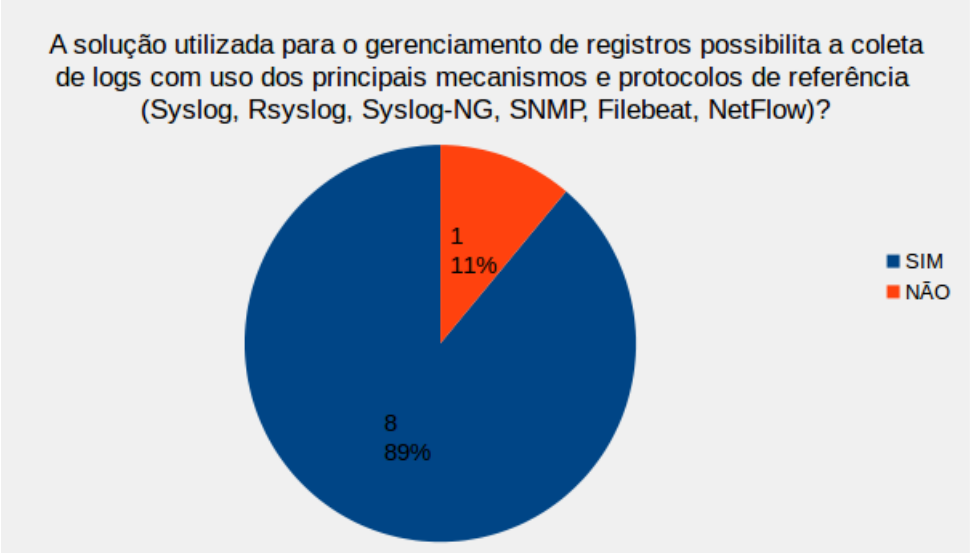


Gráfico 8 – Capacidade de reconhecimento de formatos variados de *logs*
Fonte: O autor.

Gráficos são recursos essenciais para a identificação rápida de desvios de padrão. A elaboração de representações gráficas a partir dos registros coletados permite o acompanhamento em tempo real da situação de um recurso monitorado. A pergunta oito busca identificar se as ferramentas adotas permitem a elaboração desses artifícios a partir de sentenças de busca que foram construídas pelos profissionais, o que permite a criação de outros gráficos além dos que já são oferecidos pela ferramenta. Do total das respostas, de acordo com o Gráfico 9, 56% declararam que as soluções adotadas possuem a capacidade em questão. Todavia quatro Centros (46%) declararam que as suas ferramentas não possuem tal recurso, entretanto, constatou-se que, na composição do percentual, estão as ferramentas *LogAnalyzer* e OSSIM, que são as duas soluções adotadas por todos os CTA e CT. É sabido, através da experiência de uso, que a primeira não permite a elaboração de gráficos de qualquer natureza, logo, o Centro que afirmou que a sua solução não possui tal recurso tem, no seu quadro, profissionais que não possuem pleno entendimento das possibilidades de emprego da ferramenta adotada, uma vez que os demais usuários do OSSIM afirmaram ser possível a elaboração de gráficos a partir de pesquisas personalizadas.

Gráfico 9 – Possibilidade de elaboração de gráficos a partir de *queries*
Fonte: O autor.

A pergunta nove objetiva a identificação da possibilidade de as ferramentas fazerem uso de *dashboards* independentes. Esse recurso é importante por se tratar de um meio que permite que cada profissional personalize o seu painel de instrumentos da forma que melhor lhe convier e com base nos *logs* dos recursos que monitore. O Gráfico 10 indica que 56% das soluções oferecem essa possibilidade. Dentre os 44% (quatro soluções) que não permitem a utilização de *dashboards* personalizados consta novamente o OSSIM, o que remete a mesma situação da possibilidade de criação de gráficos personalizados.

Gráfico 10 – Possibilidade de utilização de *dashboards* personalizados e independentes
Fonte: O autor.

No que se refere à utilização de soluções livres (Gráfico 11), o resultado da tabulação da questão dez indicou que todos os Centros fazem uso de ferramentas de código aberto, conhecidas como *software* livre.

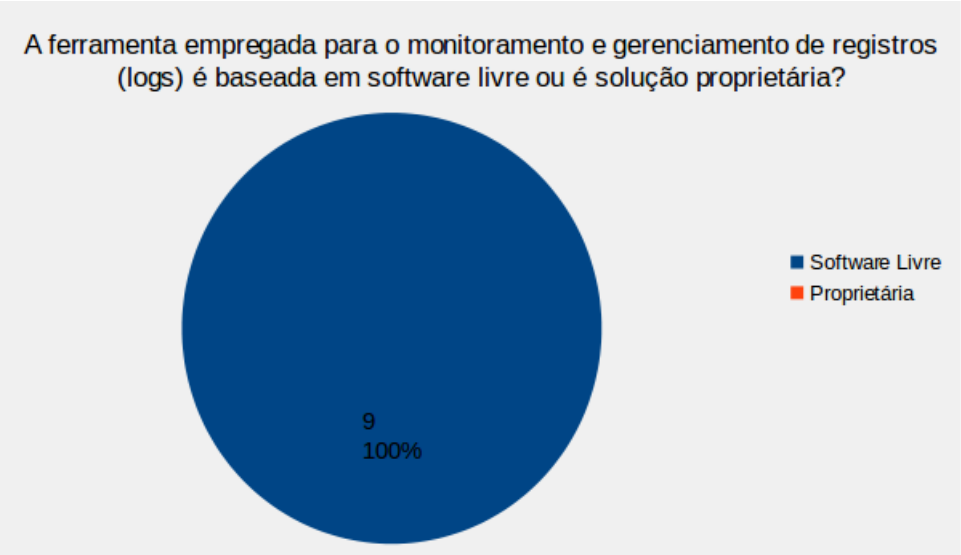


Gráfico 11 – Utilização de solução livre *versus* solução proprietária
Fonte: O autor.

Quanto ao pleno atendimento das necessidades de gerenciamento de registro de eventos e à satisfação dos profissionais pelos resultados produzidos pelas soluções adotadas, aspecto abarcado pela pergunta 11, consoante o Gráfico 12, 56% das ferramentas atendem e satisfazem as necessidades dos seus usuários. Os 46% (quatro soluções) restantes representam os três Centros que fazem uso do *LogAnalyzer* e um que se utiliza do OSSIM para o gerenciamento e análise de *logs*.

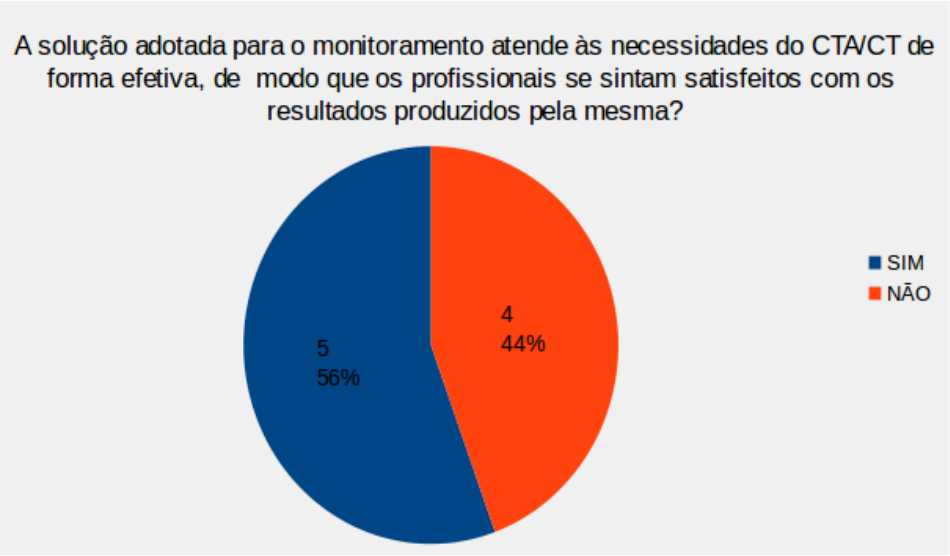


Gráfico 12 – Atendimento de necessidades e satisfação com os resultados produzidos
Fonte: O autor.

4.2 RESULTADO DA ANÁLISE DA PILHA *ELASTIC*

Neste tópico, será disposta a análise da solução pilha *Elastic* com ênfase na facilidade de implementação e utilização, no mecanismo de busca de registros armazenados, na forma de apresentação dos eventos coletados, na segurança na coleta e no tratamento dos *logs* e no desempenho e recursos da interface gráfica.

4.2.1 **Facilidade de implementação e utilização**

A instalação da solução não se constitui num problema, uma vez que todos os pacotes para instalação das ferramentas integrantes da pilha estão disponíveis no repositório. O único trabalho, nessa etapa, é a criação de um arquivo, relacionado a cada uma das ferramentas, que contenha as informações do endereço do repositório e da versão do pacote a ser utilizado.

No que tange à documentação, o desenvolvedor demonstra muita preocupação com a escrituração do material de informação. Da instalação à configuração, tudo está documentado. O acesso a todo o material, necessário à implementação da ferramenta, é muito fácil, não necessitando nem mesmo de cadastro prévio. Todos manuais de utilização estão disponíveis na página do desenvolvedor e estão agrupados em seções que estão vinculadas à ferramenta, ou seja, tudo o que é necessário para o seu aprendizado e instalação está num mesmo local.

Embora a fase de instalação seja muito simples, a curva de aprendizado da ferramenta é bastante ascendente, uma vez que requer um bom entendimento da sua forma de funcionamento. Todavia, isso não se constitui num problema em virtude da grande quantidade de material de consulta disponível, o que é de grande importância para o sucesso dessa atividade.

Cabe ressaltar que o maior entrave na etapa de instalação e configuração ocorreu por conta da utilização, pelo sistema operacional, do módulo IPv6 em detrimento do IPv4, o que fez com que a porta definida no arquivo do *plugin input* do *Logstash* somente ficasse em modo de escuta para o IPv6. A desativação do módulo IPv6 resolveu o problema.

De modo a contribuir para a difusão de conhecimento e melhoria do Sistema de Telemática do Exército e com base no abordado nas seções 1.4.2 e 1.4.3.3, a utilização da pilha *Elastic* poderá ser adotada pelos Centros de Telemática que não possuem nenhum mecanismo de gerenciamento de registros, cobrindo, dessa forma os óbices identificados pela primeira questão da pesquisa, constante no Anexo A.

Quanto à facilidade de implementação e utilização, os resultados foram muito positivos e satisfatórios.

4.2.2 **Mecanismo de busca de registros armazenados**

Quanto a esse aspecto, a solução é dotada de recursos muito avançados para a busca dos registros coletados e tratados.

Em razão da utilização de indexadores, a realização de busca e correlação dos eventos com base num determinado parâmetro é muito fácil e rápido.

Sem que haja qualquer dificuldade, no menu *Discover*, após a inserção de qualquer sentença, a busca ocorrerá e, havendo resultados, os mesmos serão apresentados na área definida como tabela de documentos e, também, no formato gráfico por meio dos histogramas. Caso seja necessário fazer algum refinamento em função de um lapso temporal, basta definir o período de pesquisa no filtro de tempo. A restrição de um período de pesquisa também pode ocorrer por meio da seleção de um histograma. Após a seleção de uma barra do gráfico, serão exibidas todas as ocorrências para aquele momento selecionado

Na realização de pesquisas, é possível a utilização de expressões regulares, as quais são muito importantes para a elaboração de buscas que necessitam de emprego de operadores lógicos. Buscas dessa natureza também são realizadas na mesma caixa de pesquisa.

A Figura 14 representa parte dos recursos anteriormente descritos. Nela é possível notar que fora realizada uma pesquisa com a utilização dos argumentos “GET && geoip” e com determinação do período de tempo (de 4 de setembro, das 8:30:00.000 às 8:35:00.000). Não só todos os resultados que atenderam aos parâmetros utilizados, como a quantidade de *hits* (conexões) foram apresentados nas áreas da tabela de documentos, do histograma e do total de *hits*.

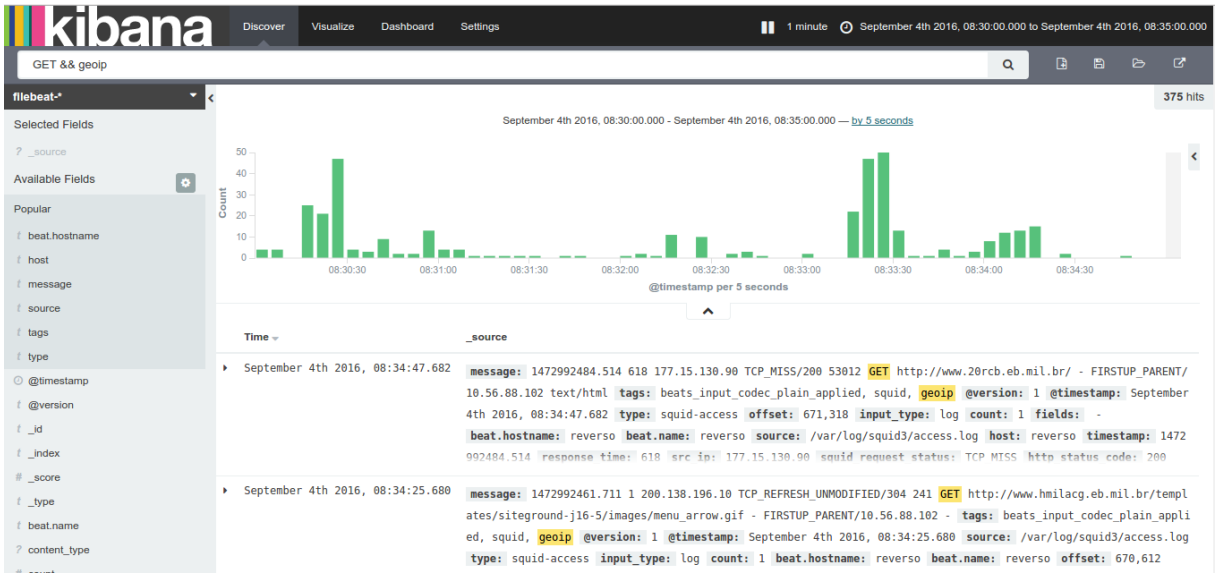


Figura 14 – Representação de uma busca com uso de operador lógico
Fonte: O autor.

Outro recurso muito importante é a possibilidade de armazenamento persistente de pesquisas realizadas, o que pode ser feito através da opção *Save Search*, disponível no próprio ambiente de busca.

Quanto a esse aspecto e conforme demostrado nos Gráficos 3, 4 e 5, grande parte das soluções adotadas pelos CTA e CT possui alguma limitação de recurso que restringe o efetivo gerenciamento dos eventos coletados e tratados. Assim, as limitações identificadas pelas questões 2, 3 e 4 poderão ser corrigidas com a adoção da pilha *Elastic*.

4.2.3 **Apresentação dos eventos coletados**

A mera coleta e armazenamento de registros, por vezes, não é suficiente para o atendimento de necessidades inerentes ao gerenciamento de registro de eventos.

A utilização de recursos gráficos é de fundamental importância na detecção de anomalias e na prevenção de incidentes de segurança da informação ou de problemas que levem à indisponibilidade de serviços.

Para atender a essa necessidade, a pilha *Elastic* possui recursos muito poderosos que permitem a elaboração de diversos tipos de gráficos com base em sentenças de busca.

A formulação de elementos gráficos a partir de *queries* é realizada por meio

do menu *Visualize*, no qual são oferecidas diversas opções de modelo de gráficos, que podem ser formados a partir de novas pesquisas ou de buscas previamente salvas. Os recursos gráficos, gerados nesse ambiente, podem ser utilizados somente para uma dada circunstância ou gravadas para que possam ser empregadas na composição de *dashboards*.

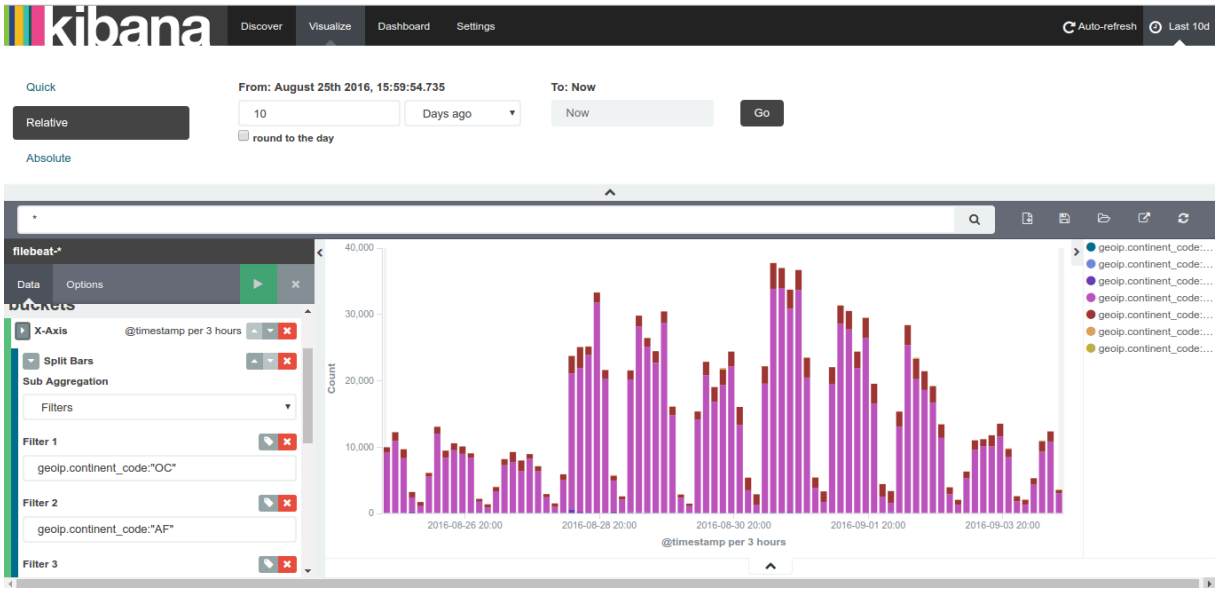


Figura 15 – Representação da criação de gráfico pelo *Visualize*
Fonte: O autor.

A Figura 15 representa um gráfico de histogramas com divisão da barra, criado pelo *Visualize*, elaborado a partir de uma pesquisa com base no campo *geoip.continent_code* dos registros coletados e dentro de um período de 10 dias.

A utilização dessa visualização em outra ocasião é possível: Basta que a mesma seja salva.

Esse tópico aborda as limitações identificadas pelas questões cinco e seis da pesquisa.

4.2.4 **Segurança na coleta e no tratamento dos *logs***

Garantir que os registros, gerados pelos dispositivos monitorados, sejam enviados para o servidor remoto sem que tenham sido alterados é uma preocupação que deve ser constante, uma vez que, nem sempre, os equipamentos gerenciados estarão numa infraestrutura próxima ou que seja dotada de todos os controles de acesso físicos. A adoção de recursos de segurança que garantam a integridade e a

confidencialidade dos dados em trânsito, bem como a autenticidade dos dispositivos envolvidos na transação, requer a utilização de soluções que sejam capazes de implementar procedimentos de proteção de forma bidirecional, ou seja, o cliente possui mecanismos para atestar a legitimidade do servidor para o qual enviará seus registros e o servidor tem a garantia de que os registros recebidos não foram adulterados em algum momento da transmissão.

Consoante o abordado na seção 2.4.2, há muitas formas de ataque à integridade dos registros que podem ser originadas por diversas motivações. Para proteger os dados dos registros em todo o seu caminho, a pilha *Elastic* faz uso de certificado SSL, gerado no servidor e exportado para os clientes, para a criptografia dos *logs* transmitidos e para a autenticação dos equipamentos envolvidos.

Esse tópico compreende o impedimento constatado pela pergunta sete do questionário da pesquisa.

4.2.5 **Desempenho e recursos da interface gráfica**

As questões que mais chamam atenção na utilização da pilha *Elastic* é o desempenho do seu mecanismo de busca, o seu baixo consumo de recursos do servidor e o alto nível da sua interface gráfica. Pesquisas complexas são realizadas rapidamente, independentemente da quantidade de eventos que estejam armazenadas nos arquivos indexados do *Elasticsearch* e da utilização de ambiente gráfico. Essa característica é reflexo do tratamento dos registros pelo *Logstash* e do armazenamento de forma indexada pelo *Elasticsearch*, o que evita elevado consumo de recursos computacionais no gerenciamento de eventos.

A interface gráfica do *Kibana* é dotada de recursos que permitem, de forma rápida e fácil, o total acesso aos registros armazenados sem qualquer perda de desempenho. O menu *Dashboards* possibilita a agregação de gráficos, gerados pelo *Visualize*, numa única tela, de modo que seja possível a visualização de diversos indicadores através de elementos gráficos, dispostos de acordo com a vontade do usuário.

O recurso de filtragem de tempo pode ser ajustado para qualquer período, fazendo com que os gráficos do *dashboard* exibam os resultados das buscas,

utilizadas nas suas elaborações, de acordo com o período definido, independentemente da construção da sentença de busca com uso de parâmetros de tempo. Além do ajuste do tempo de exibição, é possível, também, definir o período de atualização da tela (*refreshing*). Esse recurso permite o monitoramento em tempo quase real dos registros coletados.

Um recurso muito útil é a possibilidade de criação de múltiplos *dashboards*. Essa funcionalidade permite que cada profissional, envolvido na atividade de gerenciamento e análise de registro de eventos, crie seu próprio painel de instrumentos com os gráficos que melhor lhe convier para a realização do seu trabalho sem que haja impacto no desempenho do servidor ou interferência em outros *dashboards*. Um gráfico pode estar presente em vários painéis e um único painel de instrumentos pode ser acessado por todos os usuários da solução.

A Figura 16 representa o *dashboard* produzido no estudo de caso. Para a sua construção, foi utilizado um mapa (*tile map*) para a indicação das origens das conexões externas, um gráfico de área (*area chart*) para a representação dos *hits* externos, um gráfico de linha (*line chart*) para mostrar os *hits* internos (originados na rede corporativa) e quatro gráficos no formato de pizza (*pie chart*), utilizados para apontar as ocorrências de *status* do protocolo TCP, os métodos HTTP das conexões, a quantidade de requisições por endereço de página a partir da EBNet e a partir da Internet e os acessos às páginas por fatia de tempo.

Todos esses gráficos são igualmente importantes por possibilitarem a rápida e fácil identificação de qualquer anomalia que possa acarretar incidente de segurança da informação ou indisponibilidade de serviço, uma vez que os desvios de padrão serão rapidamente identificados. Esse tópico abrange as limitações identificadas pelas questões oito, nove, dez e onze da pesquisa constante no Anexo A.

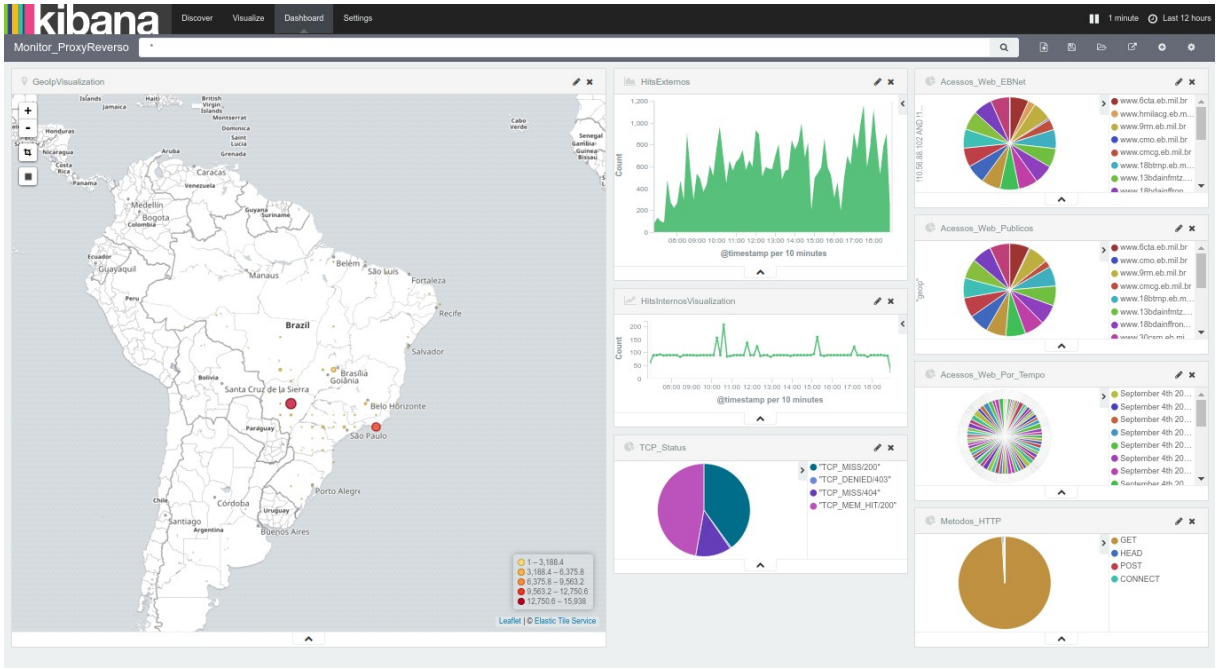


Figura 16 – Representação de *dashboard* com mapa e gráficos
Fonte: O autor.

Com base nos experimentos realizados no estudo de caso, é possível afirmar que a solução avaliada pode ser adotada e implementada facilmente naqueles Centros de Telemática que não dispõem de qualquer mecanismo para o gerenciamento de registro de eventos, uma vez que a pilha *Elastic* é dotada de grandes recursos e funcionalidades que permitem o gerenciamento completo e efetivo dos eventos de todos os recursos de uma infraestrutura de TI crítica, uma vez que os seus recursos permitem a identificação de grande parte dos fatores que possam impactar na segurança da informação ou indisponibilidade dos serviços publicados.

Em razão da grande quantidade de recursos para a coleta e tratamento de eventos, dos mecanismos de criptografia, utilizados para a garantia da integridade, e da grande flexibilidade para a elaboração de buscas e gráficos, a pilha *Elastic* se constitui, também, numa vantajosa opção para a homogeneização, em todo o SisTEx, da solução para o gerenciamento de *logs*, uma vez que tal situação faria com que fosse formada uma rede de suporte interno para essa nova tecnologia de monitoramento, o que tornaria menos íngreme o aprendizado da solução.

Conforme citado anteriormente, embora a padronização possa se constituir num fator limitante para a pesquisa de novas soluções, a uniformização em torno da

pilha *Elastic* poderá eliminar as deficiências de recursos de segurança e monitoramento, apresentadas pelas outras duas ferramentas em uso pelos demais CTA e CT e identificadas pelo questionário aplicado, sem grandes óbices em razão da farta disponibilidade de fontes de consulta e ampla utilização mundial da solução avaliada.

O resultado positivo, decorrido do estudo da pilha *Elastic* como ferramenta para gerenciamento de registro de eventos (ou simplesmente gerenciamento de *logs*), possibilita asseverar que a mesma possui todas as características, recursos e funcionalidades necessários ao efetivo monitoramento dos eventos, gerados por dispositivos de tecnologia da informação, sem qualquer necessidade de adoção de recursos complementares.

5 **CONCLUSÃO**

O gerenciamento de registro de eventos é uma atividade que, aparentemente, não requer um grande aparato tecnológico ou elevado conhecimento para a identificação de problemas, uma vez que as ferramentas mais atuais permitem a coleta, o tratamento, o armazenamento, a correlação e a exibição de todos os registros, gerados pelos dispositivos monitorados.

Tal situação pode ser evidenciada durante a elaboração deste trabalho, uma vez que pode ser constatado que o grande diferencial entre as soluções, adotadas para o gerenciamento de eventos, está na flexibilidade e nas funcionalidades que são oferecidas pelas mesmas, possibilitando a obtenção da efetividade no gerenciamento de registros.

Durante a etapa de avaliação da pilha *Elastic*, ficou latente que a ferramenta possui funcionalidades e características singulares que se adequaram perfeitamente às necessidades da área de operação do 6º CTA, visto que, com a grande gama de opções de pesquisa e recursos gráficos, viabilizou o controle mais concreto e confiável dos registros, coletados dos dispositivos críticos que integram a infraestrutura de TI do Centro.

Embora a instalação seja fácil e rápida, a curva de aprendizado da solução é acentuada, posto que é necessário o pleno entendimento da forma de funcionamento dos *plugins* do *Logstash*, que é o cerne do tratamento dos *logs* coletados, uma vez que, caso esses não estejam devida e corretamente configurados para o tipo de registro que é coletado, o resultado final, tal qual esperado, não ocorrerá, podendo conduzir à percepção de que a ferramenta não atende ao propósito de gerenciamento de eventos, enquanto, na verdade, toda a falha está nas instruções passadas a mesma.

O pleno entendimento da correta forma de construção dos *plugins* permitiu a construção de filtros eficientes e eficazes que viabilizaram a correta indexação dos arquivos pelo *Elasticsearch* sem que o desempenho do servidor fosse comprometido. A relação intrínseca existente entre os *plugins* do *Logstash* e a forma de armazenamento do *Elasticsearch* é um dos pontos fortes da pilha, uma vez que, após todo esse tratamento, a busca por eventos com a utilização de expressões

regulares, operadores lógicos ou mesmo por um campo específico se torna fácil e rápida, sendo os resultados apresentados de forma clara numa interface gráfica de alto nível.

O emprego da solução trouxe reflexos imediatos para a área da operação, uma vez que o monitoramento dos serviços passou a ser mais confiável e tempestivo em razão do acompanhamento, em tempo quase real, de todos os registros exportados pelos dispositivos monitorados. A utilização de georreferenciamento permitiu a identificação de todas as origens das conexões e o correlacionamento dos *logs* possibilitou a identificação de comportamentos arbitrários que pudessem ocasionar a indisponibilidade dos serviços.

A análise dos resultados da pesquisa, que visou à identificação das soluções, adotadas pelos demais CTA e CT para a gerência de registros, possibilizou a obtenção de referenciais comparativos que permitiram aferir quão poderosa é a pilha *Elastic*, visto que muitos dos seus recursos que asseguram a proteção dos registros durante todo o processo de coleta e análise, bem como a flexibilidade para a construção de expressões de busca dinâmicas e de múltiplos gráficos a partir dessas, não estão presentes nas ferramentas identificadas, embora uma delas seja muito conhecida e utilizada pelos profissionais de tecnologia da informação.

Outras questões que tornam a solução analisada neste trabalho diferenciada em relação as outras duas, utilizadas pelos demais Centros, é a possibilidade de criação de inúmeros *dashboards*, que podem ser utilizados de forma independente e conter elementos gráficos específicos, elaborados de acordo com as necessidades de cada profissional que desempenhe atividades de monitoramento de serviços e dispositivos de TI.

Em razão dos recursos e qualidades descritos nesse trabalho, a pilha *Elastic* é uma solução factível de adoção por todos os CTA e CT, independentemente da existência de qualquer mecanismo para o gerenciamento de eventos. Outro motivo que torna exequível o seu uso é a disponibilização, pelo desenvolvedor, de uma base de consulta muito bem detalhada, que é acessível por meio da rede mundial de computadores sem a necessidade de qualquer cadastro. Somada a esses fatores, está a experiência dos profissionais do 6º CTA na utilização da solução, o que é indispensável na elucidação de dúvidas durante a fase de implementação e

aprendizado da pilha *Elastic*.

A melhora substancial no gerenciamento de *logs* no 6º CTA, advinda da adoção da solução em questão, permite asseverar que os objetivos propostos nesse trabalho foram alcançados, servindo o seu embasamento teórico e análise da solução como um arcabouço de consulta para a implementação de um mecanismo efetivo para o gerenciamento de registros de eventos.

REFERÊNCIAS

Associação Brasileira de Normas Técnicas. ABNT NBR ISO/IEC 27001: Tecnologia da informação – Técnicas de segurança – Sistemas de gestão da segurança da informação – Requisitos. Rio de Janeiro, 2013.

BRASIL. Lei 12965, de 23 de abril de 2014. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm>. Acesso em: 17 jul. 2016.

CHHAJED, Saurabh. **Learning ELK Stack**. Birmingham: Packt Publishing, 2015.

CHUKAVIN, Anton A.; SCHMIDT, Kevin J.; PHILLIPS, Christopher. **Logging and Log Management**. The Authoritative Guide to Understanding the Concepts Sorrouding Logging and Log Management. Waltham: Elsevier, 2013.

ELASTIC. Elastic stack and product documentation. Disponível em: <https://www.elastic.co/guide/index.html>. Acesso em: 26 ago. 2016.

FERREIRA, Rubem E. **LINUX: guia do administrador do sistema**. São Paulo: Novatec Editora, 2008.

GORDON, Adam. **Official (ISC)^2 Guide to the CISSP CBK**. Boca Raton: Taylor & Francis Group, LLC, 2015.

HARRIS, Shon. **All in One CISSP Exam Guide**. New York: MacGraw-Hill, 2013.

MELO, Sandro. **Exploração de Vulnerabilidades em Redes TCP/IP**. 2 ed. Rio de Janeiro: Alta Books, 2006.

MELO, Sandro; DOMINGOS, Cesar; CORREIA, Lucas; MARUYAMA, Tiago. **BS7799: Da tática à Prática em Servidores Linux**. Rio de Janeiro: Alta Books, 2006.

MILLER, David R.; HARRIS, Shon; HARPER, Alen A.; VANDYKE, Stephen; BLASK, Chris. **Security Information and Event Management (SIEM) Implementation**. New

York: McGraw-Hill, 2011.

TRIGO, Clodonil Honório; MELO, Sandro Pereira. **Projeto de Segurança em Software Livre**. Rio de Janeiro: Alta Books, 2004.

APÊNDICE A – QUESTIONÁRIO DE PESQUISA

QUESTIONÁRIO DE LEVANTAMENTO DE SOLUÇÃO ADOTADA PARA O GERENCIAMENTO DE REGISTROS

1 – O CTA/CT possui alguma ferramenta para o gerenciamento centralizado de registros (*logs*)?
( ) Sim.
( ) Não.
Considerações acerca da questão:
_____
_____

2 – A solução adotada possui interface gráfica *web* intuitiva que permita o fácil e rápido acesso aos recursos para o gerenciamento dos registros?
( ) Sim.
( ) Não.
( ) Não há ferramenta para essa finalidade.
Considerações acerca da questão:
_____
_____

3 – A ferramenta adotada permite a utilização de certificado X.509 para a transmissão dos registros (*logs*) entre o recurso monitorado e o servidor?
( ) Sim.
( ) Não.
( ) Não há ferramenta para essa finalidade.
Considerações acerca da questão:
_____
_____

4 – Caso a ferramenta utilizada possua interface gráfica, essa possui recursos que permitam a realização de pesquisa com base em expressões regulares (RegEx) na própria interface *web*?
( ) Sim.
( ) Não.
( ) Não há ferramenta para essa finalidade.
Considerações acerca da questão:
_____
_____

5 – A interface gráfica da solução adotada permite que as sentenças de busca (*queries*) sejam salvas para utilização futura?
( ) Sim.
( ) Não.
( ) Não há ferramenta para essa finalidade.
Considerações acerca da questão:

_____
_____

6 – A ferramenta empregada no gerenciamento de registros permite a utilização de base de georreferenciamento (GeoIP) que possibilite a identificação das origens das conexões públicas?
( ) Sim.
( ) Não.
( ) Não há ferramenta para essa finalidade.
Considerações acerca da questão:

_____
_____

7 – A solução utilizada para o gerenciamento de registros possibilita a coleta de *logs* com uso dos principais mecanismos e protocolos de referência (*Syslog*, *Rsyslog*, *Syslog*-NG, SNMP, *Filebeat*, *NetFlow*)?
( ) Sim.
( ) Não.
( ) Não há ferramenta para essa finalidade.
Considerações acerca da questão:

_____
_____

8 – A ferramenta utilizada para o gerenciamento de registros permite a criação e utilização de gráficos, elaborados a partir de sentenças de buscas (*queries*), aplicadas nos *logs* coletados?
( ) Sim.
( ) Não.
( ) Não há ferramenta para essa finalidade.
Considerações acerca da questão:

_____
_____

9 – A ferramenta empregada no gerenciamento de registros permite a criação de *Dashboards* personalizados que atendam às necessidades individuais de cada profissional para a administração dos serviços, recursos computacionais e de rede ou tratamento de incidentes e que possam ser utilizados de forma independente?
( ) Sim.
( ) Não.
( ) Não há ferramenta para essa finalidade.
Considerações acerca da questão:

_____
_____

10 – A ferramenta empregada para o monitoramento e gerenciamento de registros (*logs*) é baseada em *software* livre ou é solução proprietária?
( ) Baseada em *software* livre.
( ) Solução proprietária.

( ) Não há ferramenta para essa finalidade.
Considerações acerca da questão:

_____
_____

11 – A solução adotada para o monitoramento atende às necessidades do X CTA de forma efetiva, de modo que os profissionais se sintam satisfeitos com os resultados produzidos pela mesma?
( ) Sim.
( ) Não.
( ) Não há ferramenta para essa finalidade.
Considerações acerca da questão:

_____
_____