BENEFÍCIOS DA UTILIZAÇÃO DE UMA APLICAÇÃO PARA CONCENTRAÇÃO E ANÁLISE EM TEMPO REAL DE LOGS DOS SERVIDORES WEB DE UMA INSTITUIÇÃO DE ENSINO SUPERIOR[1]

Fernando da Rocha[2]
Horácio Dutra Mello[3]

Resumo: Os arquivos de logs refletem a saúde das aplicações e fornecem insumos para os tomadores de decisão a partir da transformação dos dados em informações. Por causa desta importância, a concentração de logs é uma prática que ganha cada vez mais espaço dentro da infraestrutura de TI das instituições, com o intuito de preservar a robustez dos negócios e tornar os processos de TI mais ágeis, contribuindo assim para uma equipe de TI ser mais proativa. Como consequência da concentração e análise de dados em série temporal, as informações geradas e descritas no presente trabalho permitem que a equipe de TI consiga prever um comportamento anormal das aplicações e, por conseguinte, agir imediatamente e diretamente no agente deste comportamento. Portanto, nesta pesquisa, a concentração e análise destes arquivos buscam contribuir com a segurança da informação na Instituição de Ensino Superior do mesmo modo que seu objetivo é apresentar os benefícios da utilização desta aplicação.

Palavras-chave: Segurança da informação. Arquivos de logs. Concentração de arquivos.

1 INTRODUÇÃO

Enorme quantidade de dados são gerados e gravados em logs de diversos servidores web, a partir dos acessos aos recursos tecnológicos e sistemas disponíveis em uma Instituição de Ensino Superior. Os logs são fontes de informações preciosas e refletem a saúde e segurança das aplicações. Outrossim, possuem papel importante no monitoramento de ameaças e auditorias de segurança da informação. Entretanto, consultar esses dados em busca de informações sobre o comportamento dos serviços disponibilizados, a identificação e resolução de problemas, o monitoramento de anomalias e possíveis tentativas de invasão é tarefa difícil e de sobremaneira demorada, devido à complexidade destes arquivos e sua distribuição entre os diversos recursos Institucionais.

[1] Artigo apresentado como Trabalho de Conclusão do Curso de Especialização em Gestão de Segurança da Informação, da Universidade do Sul de Santa Catarina, como requisito parcial para a obtenção do título de Especialista em Gestão da Segurança da Informação.

[2] Graduado em Gestão de Tecnologia da Informação, fernandoroch4@gmail.com
[3] Orientador, mestre em educação, horaciomello@gmail.com

Neste sentido, para Sêmola (2003, p. 108), “realizar uma análise de segurança já é prioridade para a grande maioria das empresas, o que vem demonstrar a percepção da necessidade de diagnosticar os riscos”.

A centralização de logs faz com que a busca seja mais rápida, podendo ajudar na resolução de problemas de forma mais ágil, pois não é preciso decifrar qual o servidor teve ou está com problemas, devido ao fato de todos os logs estarem em único lugar. Além disso, é possível utilizar poderosas ferramentas para analisá-los, incluindo soluções de gerenciamento de logs que podem transformá-los de texto simples em campos que podem ser facilmente pesquisados e examinados. Desta maneira, a concentração e o monitoramento em tempo real dos logs de servidores web fornecerá, além de agilidade aos processos de busca e identificação, uma vantagem competitiva para organização, pois a segurança da informação será elevada com o emprego dos mecanismos de busca centralizada de dados e alertas de anomalias (Loggly, 2015).

Segundo Sêmola (2003, p. 43), podemos definir segurança da informação “como uma área do conhecimento dedicada à proteção de ativos da informação contra acessos não autorizados, alterações indevidas ou sua indisponibilidade”. Nesta perspectiva, um dos principais objetivos da segurança da informação é reduzir os impactos adversos na organização para um nível aceitável de risco, assim como proteger os ativos de informações contra o risco de perda, descontinuidade operacional, uso indevido, divulgação não autorizada, inacessibilidade e danos (ITGI, 2006, p. 15).

A realização desta pesquisa partiu do desejo de contribuir com a segurança da informação na Instituição de Ensino Superior e vai ao encontro do planejamento da equipe de TI em aprimorá-la, por isso a pesquisa e construção deste artigo teve como objetivo analisar e apresentar para diretoria e equipe de TI os benefícios da utilização de uma aplicação para concentração e análise em tempo real de logs dos servidores web. Logo, este trabalho apresentou o estudo de caso como método utilizado para retratar a realidade antes e após a implementação da aplicação para concentração e análise em tempo real de logs e os instrumentos utilizados foram a observação direta e a execução de testes em laboratório.

No item dois deste trabalho, o autor retrata a fundamentação teórica, preparando o leitor para acompanhar a análise dos resultados descritos no item três, assim como seus

respectivos subitens, escolha da aplicação, centralização dos logs e análise em tempo real de logs coletados. No item quatro, o autor expõe suas considerações acerca do trabalho, descrevendo em subitens a conclusão, as dificuldades encontradas e os trabalhos futuros. Por fim, é possível constatar as referências utilizadas neste trabalho.

Para resguardar a Instituição abordada neste trabalho de pesquisa optou-se por não identificá-la. Sendo assim, o artigo apenas tratará como Instituição de Ensino Superior.

2 FUNDAMENTAÇÃO TEÓRICA

A ausência de um grupo responsável pela segurança da informação, a complexidade e a ampla quantidade de recursos tecnológicos direcionam os esforços da diretoria e equipe de TI para a atuação reativa em questões relacionadas à segurança da informação na Instituição.

Desta maneira, Sêmola (2003) destaca que, elaborar um plano de ações orientados à reatividade, definir investimentos subestimados e limitados à abrangência da diretoria de TI, não perceber a interferência direta da segurança com o negócio, tratar a segurança da informação como um projeto e não como um processo, adotar ferramentas pontuais como medida paliativa e tratar as atividades como despesa e não como investimento podem refletir negativamente no negócio.

Como descrito por Ferreira (2003, p. 1) “a segurança da informação protege a informação de diversos tipos de ameaças garantindo a continuidade dos negócios, minimizando os danos e maximizando o retorno dos investimentos e das oportunidades”.

Neste entendimento, uma equipe de segurança da informação é essencial para preservar a saúde do negócio, disseminar boas práticas e recomendações e garantir que as propriedades básicas da segurança da informação - confidencialidade - disponibilidade - integridade - sejam preservadas.

Conforme Starti (2015, p. 1):

> Na era em que a virtualização cresce a cada dia, é fundamental investir em uma equipe que entenda tudo sobre ameaças virtuais. Eles conseguem diagnosticar quais os melhores softwares e hardwares para a empresa e elaboram estratégias para protegê-la, deixando-a mais segura e com melhor custo-benefício.

Neste contexto, os servidores web são equipamentos que requerem atenção especial desta equipe por disponibilizarem aplicações indispensáveis para o negócio da

companhia. Essas máquinas são obrigatórias para o funcionamento de serviços online e sua manutenção e segurança devem ser garantidas pela diretoria de TI, do mesmo modo que os cuidados dedicados aos serviços precisam ser replicados aos servidores com a finalidade de identificar possíveis problemas, realizar ações preventivas e corretivas quando necessário. Ademais, são instrumentos que registram em arquivos de logs todas requisições de acessos aos recursos tecnológicos institucionais. Desta maneira, esses documentos são fonte útil para acompanhar se as propriedades básicas da segurança da informação estão em ameaça de violação.

Segundo Lockhart (2006), os logs são importantes para manutenção da segurança da informação, porque podem ajudar desde a identificação de um ataque iminente até problemas de configuração e performance dos servidores. Entretanto, cada servidor web é responsável por armazenar seus próprios arquivos de logs e registrar todas requisições de acesso ao sistema como, por exemplo, a origem do acesso e o que se está tentando acessar.

Ressalta-se, em concordância com a ABNT NBR ISO/IEC 27002:2013 (2013, p. 65), que:

> Registros (logs) de sistema frequentemente contêm um grande volume de informações, muitas das quais não dizem respeito ao monitoramento de segurança da informação. Para ajudar a identificar eventos significativos para o propósito de monitoramento de segurança da informação, convém considerar que seja feita a cópia automática dos tipos apropriados de mensagens para um segundo registro, ou que sejam utilizados utilitários de sistemas adequados ou ferramentas de auditoria para realizar a racionalização e investigação do arquivo.

Sendo assim, a busca e análise de informações sobre a saúde e segurança das aplicações é prejudicada pela quantidade de registros e pela descentralização desses arquivos. Essa complicação reflete na demora e baixa assertividade na pesquisa por ocorrências. Uma forma de agilizar e aumentar a efetividade desse processo é utilizar ferramentas para concentração de logs de servidores web em um único lugar e a análise em tempo real desses logs durante o arquivamento.

Para Anussoya, Rajan e StyaMurty (2015, p. 2224, tradução nossa), “um concentrador de logs opera para centralizar e armazenar os registros em tempo real de

vários servidores web e esses registros são referenciados para solução de problemas, monitoramento de recursos e análise de segurança".

Assim, o monitoramento é um processo ativo que pode operar em tempo real e é encarregado de identificar e relatar eventos de segurança da informação que podem ser prejudiciais para os aplicativos e componentes da rede (Cole, 2008). Para Grubor e Barać (2014, p. 696-698, tradução nossa) "um software de análise de log está se tornando uma das ferramentas mais importantes para segurança dos sistemas, reduzindo o tempo de inatividade de uma aplicação e ajudando administradores de redes".

Neste panorama, o gerenciamento de log pode beneficiar uma organização/instituição de várias maneiras. Isso ajuda a garantir que os registros de segurança sejam armazenados em detalhes suficientes por um período de tempo apropriado. Além disso, as análises são benéficas para identificar incidentes de segurança, violações de políticas, atividades fraudulentas e problemas operacionais pouco depois de terem ocorrido e fornecer informações úteis para resolver esses problemas. Ainda, os logs também podem ser úteis para realizar auditoria e análise forense, apoiar as investigações internas da organização, estabelecer linhas de base e identificar tendências operacionais e problemas a longo prazo (NIST, 2006).

Para Fontes (2006, p. 54):

> Precisamos entender que ações preventivas, detectáveis e corretivas existem porque a organização deve tratar problemas de uma maneira profissional. Mesmo que pareça para você que essas ações atrapalham um pouco e tomam um tempo não previsto para as suas atividades, entenda que elas protegem o negócio da sua organização.

Diante disso, para agregar valor ao negócio, se faz necessária a utilização de uma aplicação para concentração e análise em tempo real dos logs de servidores web. Portanto, a busca por informações nesses arquivos, a partir da implementação, tendem à assertividade e rapidez, agilizando o processo de identificação e reação da equipe de TI. Essa medida é importante para aprimorar e reproduzir a segurança da informação nos processos institucionais, do mesmo modo que uma reação proativa pode beneficiar a manutenção dos sistemas e diminuir o risco de falhas.

3 ANÁLISE DOS RESULTADOS

3.1 ESCOLHA DA APLICAÇÃO

Inúmeras soluções foram desenvolvidas para ajudar as companhias a suportarem um amplo volume de informação criada diariamente, otimizando seus processos tecnológicos para analisar, armazenar e recuperar estes dados. Por outro lado, essa diversidade de softwares pode gerar um atraso na adoção de um deles, devido a sua complexidade de implementação e pelo fato de poucas soluções disporem de um conjunto completo e integrado de ferramentas.

Para escolha da aplicação a ser explorada nesta pesquisa, a diretoria e a equipe de TI elencaram dois requisitos importantes a serem considerados, para garantir o alinhamento da solução com o seu planejamento. Os requisitos descritos foram:

a) O conjunto de ferramentas deve respeitar as necessidades básicas da implementação, como a concentração, o armazenamento e a análise de logs; e
b) As ferramentas selecionadas devem fornecer suporte nativo para integração com as demais.

Esses requisitos são importantes para garantir a integração entre os mecanismos tecnológicos, além de facilitar a execução do projeto, a procura por documentação e a disponibilidade de suporte profissional que atenda a todo o conjunto de ferramentas. Após a manifestação dos requisitos, foram elencados sete critérios de avaliação para validar se a aplicação escolhida cumpriria com as expectativas da Instituição. Deste modo, criou-se uma tabela contendo os itens para serem avaliados, o peso de cada item e a nota atribuída durante a pesquisa.

As soluções Elastic Stack, TICK Stack, Apache Kafka, Splunk, Loggly, Scalyr, Scribe, Apache Flume, Heka, Apache Chukwa, Fluentd, Graylog, Logsign, Logzilla, Loganlyze, entre outras são exemplos da disponibilidade de diversas alternativas para estudo e resolução de problemas. Entretanto, para este trabalho, a aplicação Elastic Stack [4] foi escolhida por oferecer um conjunto de ferramentas que atende aos requisitos necessários para o sucesso deste projeto. Ademais, o pouco tempo para execução da

[4] https://www.elastic.co/

pesquisa e a familiaridade com as tecnologias envolvidas nesta solução foi determinante para essa seleção.

No Quadro 1 o leitor pode observar o parecer deste processo e o resultado obtido a partir do cálculo da média ponderada simples.

Quadro 1 - avaliação de cada critério

Critérios	**Peso**	**Nota**
Alerta de notificação na interface	2,5	5
Comunidade ativa	1	9
Documentação extensa	2,5	9
Envio de notificação por e-mail	1	7
Interface amigável	1	8
Software livre	1	10
Suporte profissional	1	10
$F = (n1 * p1 + n2 * p2 + n3 * p3 + n4 * p4 + n5 * p5 + n6 * p6 + n7 * p7) / (p1 + p2 + p3 + p4 + p5 + p6 + p7)$		7.9

Fonte: Elaborado pelo autor, 2018.

A Elastic Stack foi desenvolvida para solucionar os problemas de análise, de armazenamento e de coleta de dados, ajudando as companhias a transformarem em tempo real seus dados em informações relevantes para o negócio (Elastic, 2018). Desta maneira, o autor apresenta no Quadro 2 o kit de instrumentos que foram utilizados neste trabalho para a conclusão dos objetivos propostos.

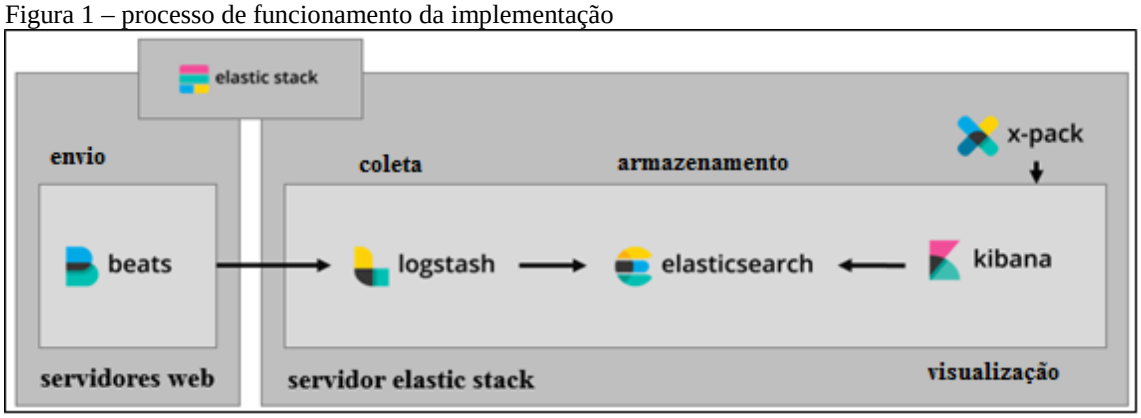
Quadro 2 – instrumentos e sua classificação

Instrumento	**Classificação**
Kibana	Visualização dos dados armazenados.
Elasticsearch	Análise, armazenamento e procura de dados.
Logstash	Coleta dados de diferentes fontes e tipos.
Beats	Envia dados de diferentes fontes e tipos.
X-pack	Estende as funcionalidades dos demais instrumentos.

Fonte: Adaptado de Elastic (2018).

O processo de funcionamento da implementação proposta neste trabalho é observado na Figura 1, no qual é possível entender as ações do projeto como envio, coleta, armazenamento e visualização dos dados utilizando-se as ferramentas descritas acima.

Figura 1 – processo de funcionamento da implementação



Fonte: Adaptado de Elastic (2018).

3.2 CENTRALIZAÇÃO DOS LOGS

A coleta de dados foi realizada em dois ambientes distintos, mas com configurações semelhantes. O atual cenário da Instituição foi denominado de modelo convencional enquanto o cenário de testes foi chamado de modelo centralizado. Outro detalhe importante é que o tempo considerado para comparação foi iniciado no modelo convencional, a partir da abertura no terminal de conexão com o servidor web, e finalizado após a exibição da primeira ocorrência pesquisada. Já no modelo centralizado, se iniciou a partir da abertura do navegador de internet e foi finalizado após a exibição da primeira ocorrência pesquisada. Deste modo, no Quadro 3 o autor apresenta os testes executados e os resultados obtidos.

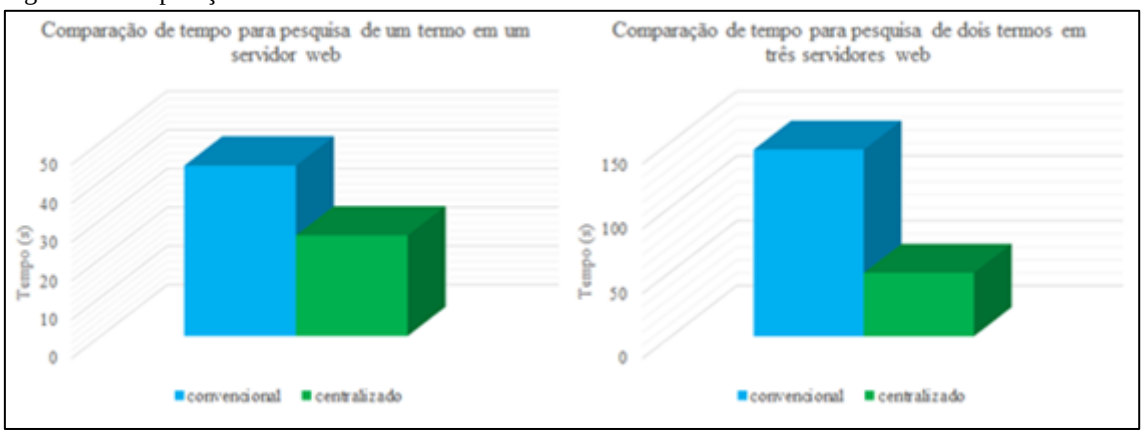
Quadro 3 – comparação entre modelo convencional de centralizado

Modelo	**Nº de servidores**	**Nº de termos**	**Termo pesquisado**	**Tempo**
Convencional	1	1	IP: 192.168.215.101	44 s
Centralizado	1	1	IP: 192.168.215.101	26 s
Convencional	3	2	IP: 192.168.215.101 e Método: GET	2 m 24 s
Centralizado	3	2	IP: 192.168.215.101 e Método: GET	49 s

Fonte: Elaborado pelo autor, 2018.

Fundamentando-se na análise dos dados coletados, é possível afirmar que encontrar uma ocorrência no modelo centralizado é aproximadamente 50 % mais rápido quando comparado ao modelo convencional. Outrossim, a eficácia do modelo centralizado tende a aumentar conforme a quantidade de servidores e termos pesquisados, assim como a complexidade da pesquisa, como constatado na Figura 2.

Figura 2 - comparação entre o modelo convencional e centralizado



Fonte: Elaborado pelo autor, 2018.

Sendo assim, é correto afirmar que o tempo para encontrar uma ocorrência nos registros de logs é menor a partir da utilização do modelo centralizado. Sem dúvida, os dados apresentam um benefício valioso para a diretoria e equipe de TI, promovendo maior agilidade nos processos, uma vez que as informações sobre os servidores web estão centralizadas em único local. Desta forma, não há necessidade de acessar os arquivos de logs em cada servidor web para encontrar uma ocorrência.

A aplicação Elastic Stack reúne diversas fontes de dados em único lugar para facilitar a busca por ocorrências, do mesmo modo que seu indexador agiliza o processo de busca (Lei, Wang e He, 2015).

As funcionalidades disponibilizadas pelo ambiente gráfico são outro ponto positivo, pois permitem a realização de buscas a partir da utilização de consultas no formato *JSON*[5]; de operadores lógico; de expressões regulares; e da seleção dos campos disponíveis para filtro; entre outras. Na Figura 3, imagem a, é possível observar a tela

[5] https://www.json.org/json-pt.html

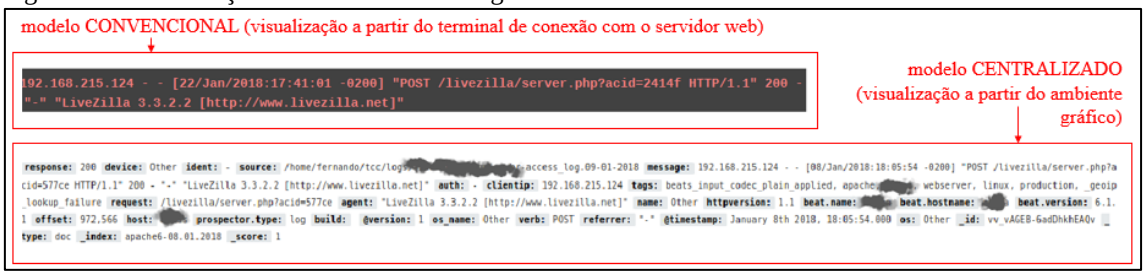
inicial do ambiente gráfico, assim como a visualização do campo de busca, os campos disponíveis para filtro e o detalhamento de uma ocorrência.

Figura 3 – console web de administração e visualização dos dados

Fonte: Elaborado pelo autor, 2018.

Com a utilização do console web é possível visualizar que os dados são distribuídos e separados em diferentes campos, como constatado na imagem b da Figura 3. Desta maneira, a equipe de TI encontra e identifica uma ocorrência de maneira mais eficiente quando comparado com o modelo convencional. Essa justificativa pode ser observada na Figura 4, a qual exibe a demonstração de uma entrada de log no modelo convencional e centralizado.

Figura 4 - demonstração de uma entrada de log no modelo convencional e centralizado



Fonte: Elaborado pelo Autor, 2018.

Conforme apurado, a aplicação de uma solução para centralização dos logs de servidores web é fundamental para reter e pesquisar informações essenciais da saúde

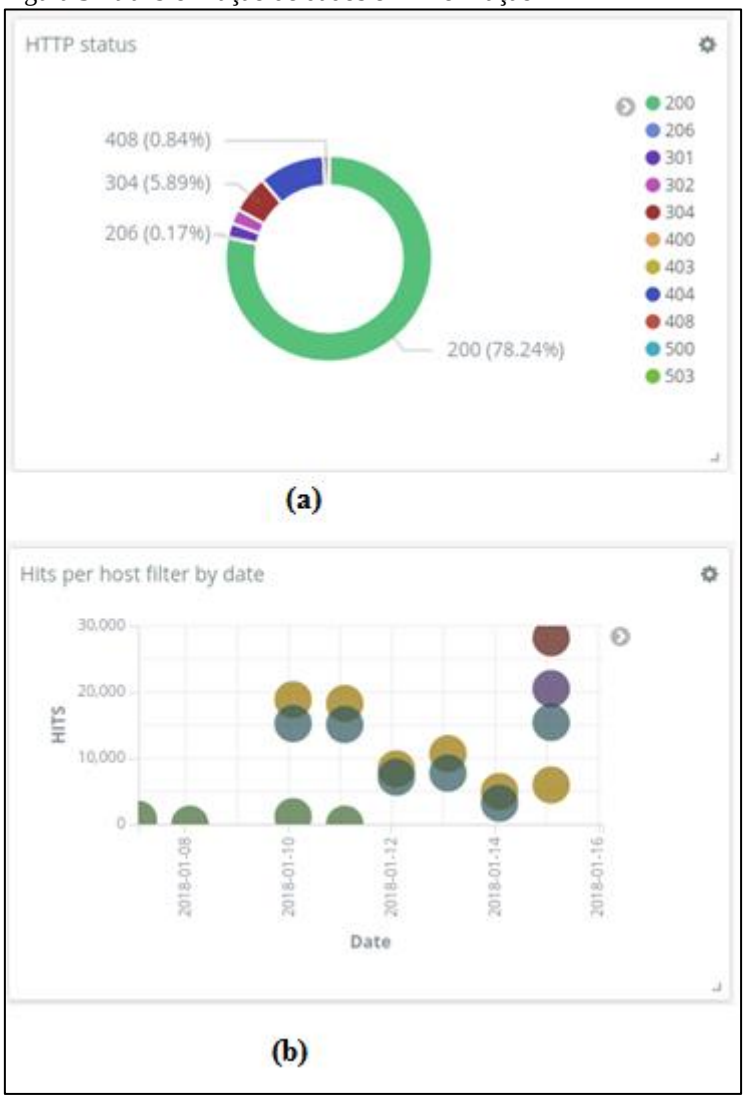
desses equipamentos. Os dados gerados podem, de maneira rápida, serem transformados em informações para toda equipe de TI a partir das necessidades institucionais,

Como descrito por Söderströma e Moradiana (2013, p. 1249, tradução nossa):

> Gerenciamento e análise de log é uma parte vital da administração de rede e dos sistemas da organização. Arquivos de logs indicam o estado atual dos sistemas e incluem informações que se referem a diferentes eventos de segurança da informação. Os logs são usados para diferentes fins, como a gravação de atividades do usuário, acompanhamento de tentativas de autenticação e outros eventos de segurança.

Na Figura 5 há um exemplo de transformação de dados em informações valiosas para a equipe de TI.

Figura 5 - transformação de dados em informação



Fonte: Elaborado pelo autor, 2018.

A partir da análise da Figura 5 é viável conhecer o percentual de códigos de retorno das requisições de acesso aos servidores monitorados, do mesmo modo que identificar qual deles foi o mais acessado. Ao analisar o gráfico a, a equipe de TI pode observar se há um número anormal de códigos de erro sendo retornados dos servidores web, se esse comportamento incomum pode ser originado do funcionamento inadequado de algum servidor web, sistema e/ou até mesmo de uma tentativa de acesso não autorizado. No gráfico b, a equipe de TI consegue perceber qual dos servidores web teve maior quantidade de solicitações de acesso e verificar caso o número seja demasiadamente fora dos padrões Institucionais.

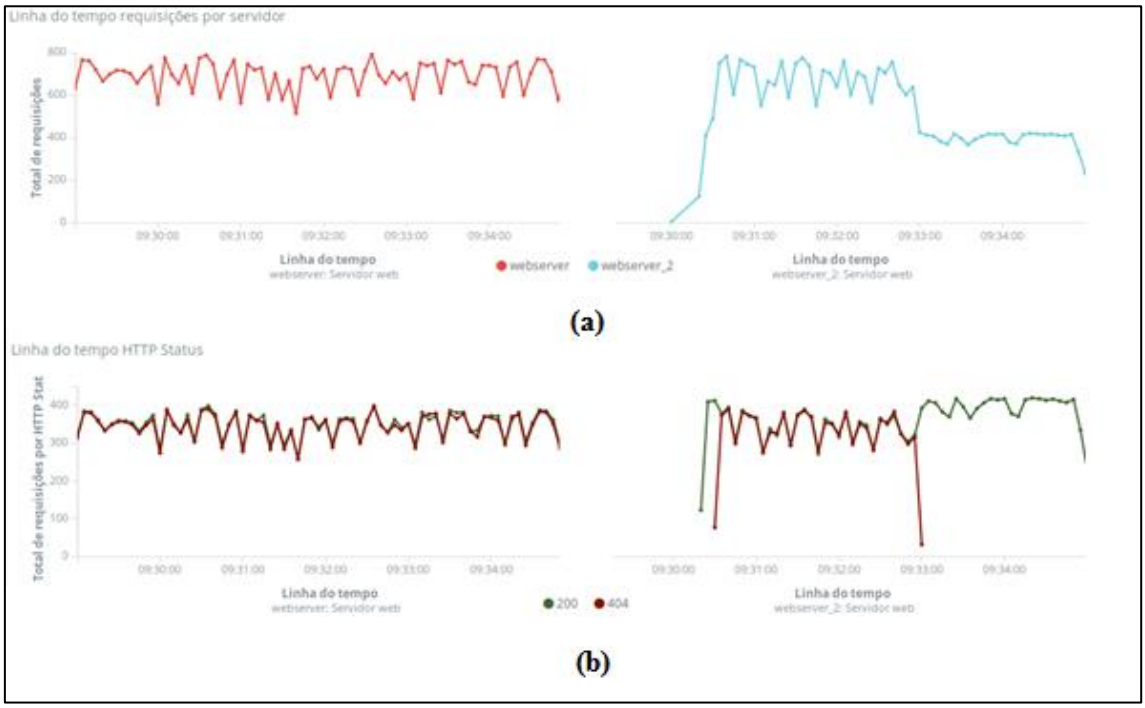
3.3 ANÁLISE EM TEMPO REAL DOS LOGS COLETADOS

Umas das funcionalidades da aplicação Elastic Stack é a possiblidade de visualização dos dados em série temporal. Deste modo, os dados coletados são transformados, em tempo real, em informações relevantes para a equipe de TI que apoiada nestas referências é capaz de prever um comportamento anormal dos servidores web.

Nesta linha, de acordo com Kleindienst (2016, p. 7, tradução nossa) “somos capazes de identificar rapidamente o motivo de um comportamento anormal e tomar medidas antes que o sistema fique indisponível”.

Na sequência, é notado na Figura 6 dois conjuntos de gráficos sendo processados pelo sistema.

Figura 6 – conjunto de gráficos sendo processados pelo sistema



Fonte: Elaborado pelo Autor, 2018.

Ao analisar os gráficos a da Figura 6 é possível identificar que nos últimos cinco (05) minutos de operação dos servidores web, o *webserver* recebeu um número maior de acessos do que o *webserver_2*. Entretanto, ao examinar os gráficos b, é plausível afirmar que há um comportamento anormal por parte de ambos servidores web no período entre 9 h 30 min e 9 h 33 min. Nesta série temporal eles responderam as tentativas de acesso em média 50 % de maneira correta e 50 % de maneira incorreta. Diante destas informações, a equipe de TI pode realizar uma análise rápida e detalhada diretamente nos ativos afetados, agindo de maneira proativa para evitar que as propriedades básicas da segurança da informação sejam prejudicadas.

Desta maneira e, segundo a RNP (1999), "os logs, com certeza, desempenham um papel imprescindível no processo de detecção de intrusão. A auditoria e avaliação destes devem se tornar rotineiras e uma constante na vida dos administradores a fim de evitar surpresas desagradáveis".

Para o CERT.br:

> Logs são essenciais para notificação de incidentes, pois permitem que diversas informações importantes sejam detectadas, como por exemplo: a data e o horário em que uma determinada atividade ocorreu, o fuso horário do log, o endereço IP de origem da atividade, as portas envolvidas e o protocolo utilizado no ataque

> (TCP, UDP, ICMP, etc.), os dados completos que foram enviados para o computador ou rede e o resultado da atividade (se ela ocorreu com sucesso ou não).

Ademais, conforme a necessidade, é possível elaborar em tempo real diferentes conversões de dados para informações úteis para a diretoria e equipe de TI. Essa transformação é um benefício que agrega valor ao negócio da Instituição, uma vez que permite o processamento de dados de diferentes origens, fornecendo um panorama amplo de informações sobre a infraestrutura de TI.

4 CONSIDERAÇÕES

4.1 CONCLUSÃO

Este trabalho foi proposto para exibir os benefícios da utilização de uma aplicação para concentração e análise em tempo real dos logs de servidores web. Neste caso, buscou-se validar a utilização da aplicação Elastic Stack como solução que atendesse aos requisitos elencados pela diretoria e equipe de TI.

A escolha pela Elastic Stack demonstrou que a solução atende aos critérios definidos e cumpre de maneira satisfatória o objetivo do trabalho. Assim sendo, é possível afirmar, com base nas coletas e análise dos resultados, que a solução proposta agrega valor ao negócio da companhia por ampliar a segurança da informação.

A busca por informações/ocorrências foi enriquecida pela concentração de dados que tornou o processo de pesquisa mais ágil e confiável, devido ao fato de disponibilizar um ambiente gráfico intuitivo que fornece funcionalidades avançadas como os filtros e a interpretação de operadores lógicos e de expressões regulares no campo de busca. Além disto, todos os dados provenientes dos arquivos de logs são concentrados em uma única aplicação, fato que amplia a segurança da infraestrutura, uma vez que minimiza possíveis perdas de dados e torna o processo de investigação mais assertivo.

O agrupamento dos dados em série temporal é um benefício que proporciona um monitoramento ativo dos recursos tecnológicos a partir da conversão em tempo real de dados em informações sobre a saúde das aplicações, permitindo que a equipe de TI acompanhe isto de maneira online, por meio da visualização de gráficos, a disposição das informações.

Resumindo, os testes efetuados neste trabalho, a partir da observação direta, demostraram que há uma série de benefícios com a utilização da aplicação proposta e, apesar das dificuldades encontradas durante o processo de implementação, a solução é altamente recomendada e assegura oportunidades para as Instituições, desde que haja planejamento e, por conseguinte, investimentos adequados.

Portanto, conclui-se que a concentração e análise de logs de servidores web é algo latente e deve ser incluído no planejamento das Instituições, a fim de tornar os processos de TI mais ágeis e seguros, assim como fornecer informações relevantes sobre os recursos tecnológicos em operação, ajudando no planejamento, na correção de falhas, na tomada de decisão e, principalmente, na manutenção do negócio.

4.2 DIFICULDADES ENCONTRADAS

Os testes executados em laboratório revelaram algumas dificuldades. A primeira foi observada durante o armazenamento dos logs por meio do software Elasticsearch. Naquele momento, identificou-se que os arquivos de logs que ocupavam 144 MB de espaço em disco aumentaram 463 % a utilização deste recurso após o processamento dos dados, totalizando 667 MB utilizados. Desta maneira, é importante que o planejamento projete um crescimento na utilização de espaço em disco de aproximadamente 500 %. Outra alternativa, é utilizar-se de filtros para permitir que somente dados importantes para o projeto sejam armazenados.

Neste sentido, segundo Virtanen (2017, p. 26, tradução nossa), "o desafio é incluir o registro importante e excluir o que não é importante para o projeto a partir de milhões de linhas de registro de log".

Em concordância com a RNP (1999):

> A maior dificuldade em um processo de detecção de intrusão decorre da imensa quantidade de informação gerada pelos eventos diários de um sistema. Separar os dados relevantes dos comuns é algo que requer tempo, atenção e, até mesmo, experiência.

Muthurajkumara, Ganapathyb, Vijayalakshmia e Kannana (2015, p.595, tradução nossa) afirmam que "nos tempos recentes, o tamanho dos arquivos de logs cresce

rapidamente. A concentração e o gerenciamento local de uma grande quantidade de logs envolvem enorme investimento e significantes recursos de *hardware*".

Outrossim, a alta quantidade de recursos de *hardware* necessários para o funcionamento da Elastic Stack pode comprometer a qualidade dos testes, uma vez que as requisições demoram mais que o previsto para serem processadas pela aplicação. Do mesmo modo que é necessário planejar questões físicas, é indispensável planejar o que se pretende alcançar com a implantação desta solução, em virtude de sua complexidade no momento da transformação de dados em informação.

4.3 TRABALHOS FUTUROS

A aplicação testada fornece inúmeras possibilidades de envios, de concentração, de agregação, de filtros e de análise de dados, portanto, como sugestão, os trabalhos futuros podem explorar mais o universo destas funcionalidades, uma vez que há consideráveis benefícios com a utilização deste conjunto de ferramentas.

Igualmente, um ambiente de teste com maior capacidade de processamento é indicado para processar uma amostra maior de logs, bem como enviá-los de diferentes servidores e/ou data centers é essencial para analisar se a latência de rede influenciará na tomada de decisão, a partir da visualização das informações geradas pela aplicação.

REFERÊNCIAS

ANUSSOYA, R; RAJAN, J; STYAMURTY, S.A.V. Importance of Centralized Log Server and Log Analyzer Software for an Organization. **International Research Journal of Engineering and Technology (IRJET).** v. 2, p. 2244-2249, jun. 2015. Disponível em: <https://www.irjet.net/archives/V2/i3/Irjet-v2i3365.pdf>. Acesso em: 03 nov. 2017.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO/IEC 27002:2013**: tecnologia da informação - técnicas de segurança: código de Prática para controles de segurança da informação. Rio de Janeiro, 2013.

CERT.br. Cartilha de Segurança para Internet: mecanismos de segurança. Disponível em: <https://cartilha.cert.br/mecanismos/>. Acesso em: 14 fev. 2018.

COLE, Eric et al. **Network security fundamentals**. Hoboken: Wiley, 2008.

ELASTIC. **The Open Source Elastic Stack**. Disponível em: https://www.elastic.co. Acesso em: 19 jan. 2018.

FERREIRA, Fernando Nicolau Freitas. **Segurança da informação**. Rio de Janeiro: Editora Ciência Moderna Ltda, 2003.

FONTES, Eduardo. **Segurança da informação:** o usuário faz a diferença. São Paulo: Saraiva, 2006.

GRUBOR, Gojko e BARAĆ, Ivan et al. Integrated Proactive Forensics Model in Network Information Security. In: International Scientific Conference, p. 696-698, 2014, Serbia. **Impact of the Internet on Business Activities in Serbia and Worldwide**. Disponível em: <http://portal.sinteza.singidunum.ac.rs/Media/files/2014/693-699.pdf>. Acesso em: 14 fev. 2018.

ITGI. **Information Security Governance**: Guidance for Boards of Directors and Executive Management, 2nd Edition. Disponível em: <http://www.isaca.org/knowledge-center/research/documents/information-security-govenance-for-board-of-directors-and-executive-management_res_eng_0510.pdf>. Acesso em: 20 out. 2017.

KLEINDIENST, Patrick. **Building a real-world logging infrastructure with Logstash, Elasticsearch and Kibana**. Disponível em: <https://hdms.bsz-bw.de/files/5021/elk_paper_patrick_kleindienst.pdf>. Acesso em: 15 jan. 2018.

LEI, Xiafei; WANG, Zhe; HE, Yuzhen et al. Log Real-time Management Scheme Based on LEK. P. 210-2013, 2015, Jinan. **International Workshop on Materials Engineering and Computer Sciences**. Disponível em: <https://www.atlantis-press.com/php/download_paper.php?id=25840584>. Acesso em: 20 fev. 2018.

LOCKHART, Andrew. **Network security hacks**. 2nd ed. Beijing: O'Reilly, 2006

LOGGLY. **Benefits of Centralizing Logs**. Disponível em: <https://www.loggly.com/ultimate-guide/managing-linux-logs/>. Acesso em: 26 out. 2017.

MUTHURAJKUMARA, S; GANAPATHYB, S; VIJAYALAKSHMIA, M e KANNANA, A et al. Secured Temporal Log Management Techniques for Cloud, p. 589-595, 2014, London. **International Conference on Information and Communication Technologies**. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1877050915001623

>. Acesso em 14 fev. 2018.

NIST. **National Institute of Standards and Technology**: Guide to Computer Security Log Management. Disponível em: <https://www.nist.gov/publications/guide-computer-security-log-management>. Acesso em: 8 nov. 2017.

RNP. **Rede Nacional de Ensino e Pesquisa**: Os Logs como Ferramenta de Detecção de Intrusão. Disponível em: <http://memoria.rnp.br/newsgen/9905/logs.html>. Acesso em: 10 fev. 2018.

SÊMOLA, Marcos. **Gestão da Segurança da Informação**: uma visão executiva. Rio de Janeiro: Elsevier, 2003.

SÖDERSTRÖM, Olof e MORADIAN, Esmiralda et al. Secure Audit Log Management. p. 1249-1258, 2013, Kitakyushu. **International Conference in Knowledge Based and Intelligent Information and Engineering Systems**. Disponível em: <

https://www.sciencedirect.com/science/article/pii/S1877050913010053>. Acesso em: 2 mar. 2018.

STARTI. **O que é segurança da informação e por que você deve saber?.** Disponível em: <http://www.starti.com.br/blog/o-que-e-seguranca-da-informacao-e-porque-voce-deve-saber/>. Acesso em: 08 nov. 2017.

VIRTANEN, Henri. **Implementing Automated Log Based Alerts in a Patient Information System**. Disponível em: < https://www.theseus.fi/bitstream/handle/10024/128480/Virtanen_Henri.pdf?sequence=1 >. Acesso em: 10 mar. 2018.