No log, no crime.

0- Intro

Falar em logging é sempre assunto divertido e ao mesmo tempo caótico. Embora seja recurso que nos acompanha desde os primórdios da telemática, ainda é difícil colocar um grupo de pessoas para conversar onde exista coesão sobre o assunto. Os motivos, os mais variados possíveis. Se você clicou no link a procura da confirmação do conteúdo do título, recomendamos que pule para a **seção 3** e depois leia a **seção 5**. Nos da **OmniSec** esperamos, de verdade, que tenhamos errado no prognóstico, mas normalmente ele se confirma em alguma plataforma crítica. Se esse for seu caso e esteja precisando de ajuda, conte conosco.

Explicaremos de uma maneira ordenada alguns pontos fundamentais que versam sobre esse importante componente da disciplina de Segurança, entre elas:

(i) apresentar um mecanismo mínimo e fundamental de rastreabilidade, atributos, carimbo de tempo e trilhas de auditoria visando suprir questões de monitoramento, auditoria, proteção e a classificação das informações;

(ii) descrever boas-práticas no que concerne à confidencialidade, integridade, disponibilidade, autenticidade, posse e utilidade (Parkerian Hexad sempre);

(iii) estabelecer diretrizes gerais para desenvolvedores, usuários, plataformas e recursos tecnológicos de forma a garantir que questões legais, de auditoria, controle, monitoração, retenção e regulatórias cobertas tenham correlação e sejam referenciadas, a partir dos seguintes macro-assuntos:

1. Atributos técnicos fundamentais
2. Considerações importantes e não-técnicas
3. Considerações extras
4. Registros de logs importantes para Segurança
5. Exceções: O que não logar
6. O que mais considerar para o futuro?

Esqueleto estabelecido, vamos ao conteúdo e discussão que interessa.

1- Atributos técnicos fundamentais

Independentemente do tipo de seu negócio, todo ativo, aplicação, sistema ou componente gerador de eventos precisa de um livro de registro de acontecimentos técnicos numa base temporal comum que possa ajudar o desenvolvedor, o time de Segurança, de Redes, de Auditoria ou outros times na difícil missão de identificar anormalidades, até porque, poder voltar no tempo e checar se uma situação em específico ocorreu pode ser um recurso muito útil em situações mais delicadas - principalmente as contraditórias ou que envolvam questões legais.

Abaixo, apresentaremos algumas perguntas e sugestões comuns de controles que podem ser utilizadas nesse momento inicial de padronização e considere que, padronização é a chave para o sucesso futuro da sua empreitada – independente do seu estágio ou desorganização inicial.

	Controle	**Descrição**
Quando?	Data e hora	É esperado que seja definida uma entidade oficial para o fornecimento do carimbo atômico de tempo (data e hora), em formato do país interessado, considerando GMT, *timezone* e horário de verão automático - se aplicável. Avançar com o tema sem essa premissa inicial é começar bem errado e produzindo conteúdo/ruído ruim.
Quem?	Identificar/Ator	Agente, ambiente, módulo, ferramenta, URL, interface, MAC, navegador, objeto ou elemento gerador do “evento”, preferencialmente com a sua

		versão. Um *host* ou elemento nominativo, deve ter registrado seu nome/identificador comum e preferencialmente único.
Onde?	Endereço Origem	Endereço do recurso tecnológico (IPv4 ou IPv6) que está realizando a requisição/pedido, sua identificação original (algum ID que retorne esse valor), seja este endereço físico ou lógico.
	Endereço Destino	Endereço do recurso tecnológico (IPv4 ou IPv6) que foi acionado/servido (o destinatário). A identificação destino (algum ID que retorne esse valor), seja este endereço físico ou lógico. Fontes estáticas e com menor facilidade de adulteração são bem-vindas.
Como?	Porta/Protocolo	O número da porta, do serviço ou protocolo utilizado na ação. Considere avaliar a existência de outros protocolos que não necessariamente forneçam essa e as demais informações, mas sejam importantes (RAW, *BUS, o quase extinto Appletalk, legado de IPX/SPX e quaisquer outras variações não-TCP/IP).
	Status da Operação	Armazenar a falha e o sucesso das operações relevantes, principalmente falhas que sejam anomalias/anormalidades importantes para o negócio de acordo com cada contexto de uso.
Por quê?	Tipo de Acesso ou Atividade executada	Uma descrição do evento, da situação que resultou o registro (acesso a um objeto, a um recurso, escrita, criação, modificação e/ou exclusão).
	Criticidade	De acordo com a plataforma e modelos utilizados, mencionar a criticidade, o nível da ocorrência (*emergential*, *alert*, *critical*, *error*, *warning*, *notice information* e *debug*). Considere usar os padrões oferecidos de "facilities e level" se estiver usando o syslog ou padrão similar (acredite em padrões, eles irão salvar o mundo, economizar milhares de horas e retrabalho).
Quanto?	Em volume e valor	Controlar um ecossistema de logging ao longo do tempo exige espaço e, naturalmente, mesmo usando uma plataforma open, haverá custo financeiro associado. Considere desenhar sua estrutura olhando não somente para as questões de momento, mas principalmente de longo prazo. Faça cálculos que possam lhe ajudar no provisionamento dos recursos e que lhe forneça uma referência volumétrica ao longo do tempo. Atenção especial aqui para provedores de serviços que ofertam soluções de correlação sem custo e você só paga pelo armazenamento.
O quê?	Sobre atributos, padronização e normalização	Tão importante quanto considerar cada uma dessas caixinhas anteriores é ter conhecimento para definir, desde a largada do seu projeto, como iremos iniciar as atividades de padronização (mínima) entre os diversos tipos de ativos que participarão do arranjo. É comum, por exemplo, ativos de redes e servidores de aplicações registrarem de forma diferente (em qualidade e quantidade) seus logs e ações. Lembre-se que o syslog (e outros) permitem configurações distintas por contexto, podem mandar dados iguais para servidores diferentes etc. Padronizar e normalizar seus dados (pelo menos o mínimo) nessa fase inicial vai evitar ruído, volume desnecessário e diversos problemas ao longo das novas integrações (detalhes na sequência). Essa atividade junto com o NTP sãos as mais importantes antes de avançar com as demais personalizações.

2- Considerações importantes e não técnicas

Normalizar as entradas, possuir uma fonte de tempo fidedigna como base, se preocupar com o armazenamento seguro (e redundante em locais diferentes), além de outros quesitos é de suma importância para manter seu ambiente a prova de falhas e questionamentos. Considere o mantra 'no log no crime' e invista tempo desenhando uma infraestrutura robusta a prova de problemas antes da materialização de qualquer incidente. Use e abuse de "Chaos Engineering" neste momento e garanta toda a segurança e redundância necessária que seu bolso possa patrocinar, afinal, na hora de um incidente ou que necessite dar satisfações a algum regulador ou entidade legal, a única coisa que você não vai querer dizer é que as "provas foram comprometidas".

Controle	**Descrição**
Padrão de Logging	Deve ser estabelecido um sistema de padronização comum para armazenamento dos *logs*, tanto para o conteúdo (*syslog*, *eventviewer etc*), quanto para a comunicação, protocolos utilizados e nomenclatura dos arquivos.
Timestamp oficial e comum	É esperado que seja definido uma entidade oficial para o fornecimento do carimbo atômico de tempo (data e hora), em formato do país interessado ou padrão global, considerando GMT, *timezone* e horário de verão automático (quando aplicável).
Armazenamento	Localização final onde os logs e informações estão efetivamente armazenadas (local ou remoto), sendo padrão desejado remoto com backup e redundância (pelo menos 2 servidores/nós). Deve considerar rotação, prazos (para cada contexto regulatório, sugestões mais a frente neste artigo) e tamanhos, além do uso em partições separadas do sistema operacional para evitar problemas com falta de espaço em partições compartilhadas. Dependendo do volume transacionado, pode ser interessante pensar em padrões de armazenamento (em arquivo, disco ou banco) que performem melhor com alta volumetria.
Validações Periódicas	Tempo entre verificações temporais nas condições do ambiente. Pode ser via demanda interna, requisição pontual ou pelo menos a cada 90 dias. Ideal que seja realizada por área diferente da proprietária dos loggings ou pela área de Segurança (quem gera o loggging não pode ser responsável por sua validação).
Privacidade e Segurança	Sendo conteúdo com possíveis informações sensíveis, classificável ou de uso restrito, é fortemente recomendado que a proteção, backup, criptografia, rotulagem e segurança sejam considerados nestes repositórios. Não esqueça que o próprio repositório deve ter condições de segurança tão adequadas quanto o solicitado pelo rótulo de maior peso/proteção. No caso de rótulos importantes, pode ser interessante codificar ou definir mecanismos que facilitem a identificação do contexto no payload ou na trilha a ser armazenada em sua ferramenta centralizadora de logs (SIEM ou similar).
Dados críticos	Dados que contenham Informações de identificação pessoal (*Personally Identifiable Information* – PII, em inglês, que impactam diretamente na LGPD/GDPR), o Padrão de Segurança de Dados da Indústria de Cartões de Pagamento (*Payment Card Industry Data Security Standard* – PCI-DSS), dados sensíveis, tokens, informações de sigilo bancário não devem ser expostos nos registros (*logs*). Em caso de processamento, o Valor Verificador do Cartão (*Card Verification Value* – CVV, em inglês) não deve ser armazenado em hipótese alguma por questões de aderência ao PCI com pena de invalidar sua certificação.

3- Considerações extras

Tão importante quanto o armazenamento correto e seguro de seus logs está o período de retenção por contexto. Além de ajudar no atendimento de pontos de auditoria, fraudes, checagens de segurança, suporte a situação de não-repúdio, este elemento pode ser essencial em ações legais, álibis ou mesmo evidências criminais, caso alguma anormalidade ou incidente tenha reflexos jurídicos entre os envolvidos.

Um ponto de atenção que, por diversas vezes acaba sendo esquecido e não considerado é que, uma única aplicação pode ter necessidade de tempo de retenção diferente para módulos ou componentes de diferentes contextos. Apenas como exemplo, considere uma solução de ERP corporativo ou uma plataforma de BaaS (Banking as a Service) onde há módulos de pessoas (Dados Pessoais) que coexistem com módulos, por exemplo, de contas (Sigilo Bancário) e pagamentos (Cartão de Crédito). Em diversos casos, a regra de retenção está unificada e atende unicamente um contexto (e rezemos para que seja o mais restritivo).

Neste caso, o correto armazenamento pode ajudar sua empresa a dimensionar custos (para mais ou para menos), atender questões regulatórias, se livrar de problemas futuros e/ou consumir estas fontes como ótimos conectores (datasources) para, por exemplo, suportar os times de resposta a incidente, fraude corporativa ou de inteligência.

Contexto		**Retenção**	
Regulador	**Setor**	**Online**	**Offline**
ANBIMA	Financeiro/Investimentos	N/A	5 anos
B3	Ordem, custódia, trilhas etc.	N/A	5 anos*
BaCen 4.474	Transações Financeiras	N/A	5 anos
California CCPA	Privacidade	N/A	2 anos
CVM	OMS - Order Mgmt System	N/A	5 anos**
EU Data Retention	Telecomunicações & Internet	N/A	6 meses - 2 anos
HIPAA	Médica/Saúde	N/A	6 anos
LGPD/GDPR	Proteção a Dados Pessoais	N/A	7 anos – 10 anos
Marco Civil Internet	Internet/Provedores	N/A	6 meses - 1 ano***
Master SDP	Meio de Pagamentos	N/A	< a definir >
OpenBanking	Transações Financeiras	N/A	5 anos
PCI-DSS	Cartões de crédito	3 meses	1 ano
Ministério Trabalho	RH (Contratações)	N/A	5 anos****
RSFN/SFN/BC	Transações Financeiras	N/A	5 anos -10 anos*****
Sarbanes Oxley	Financeiro	N/A	7 anos
The Basell II	Financeiro	N/A	7 anos
UK - DRIP	Telecomunicações & Internet	N/A	Até 2 anos******

*Com replicação diária fora do ambiente de produção.
**Ou por prazo superior por determinação expressa do regulador.
***Todos os logs 6 meses, exceto entidades administradoras de ASs (Autonomous System).
****Após finalização das relações trabalhistas, demais finalidades de acordo com contexto regulador.
*****Contexto de Know Your Customer, Lavagem de Dinheiro e Terrorismo explicitamente é 10 anos.
******Proporcional ao volume de dados gerados pelo provedor (maior volume, menor retenção).

4- Registros de logs importantes para Segurança

A listagem abaixo não estressa todas as possibilidades e pode variar bastante de acordo com o ambiente, tipo de negócio, obrigações da sua companhia e tecnologia utilizada pelo interessado. Utilize-a como uma referência-base para montar sua própria lista e não esqueça de envolver setores diferentes (além de Tecnologia e Segurança) durante processo de construção. Algumas fontes e provedores de registros importantes podem ser:

- Quaisquer entidades autoritativas presentes em seu ambiente (OpenLDAP, Active Directory, SAP, Samba, NIS, YellowPages, RADIUS, TACACS+, NDS, DNS...)
- Alertas de antivírus/anti-spywares indicando anormalidades;
- Alertas críticos de IDS/IPS e outros endpoints;
- Alertas críticos de firewall, proxy e dispositivos de borda;
- Alterações nos logs de segurança, sistemas e/ou eventos;
- Alertas de alterações em integridade de arquivos (FIM);
- Alterações, adições a grupos de usuários administrativos;
- Serviços essenciais (stringent assets), como AD, VPN, DHCP, DNS e JumpBoxes;
- Eventos de sistema ou segurança reconhecidamente como não comuns;
- Alarmes e traps gerados por infraestrutura em nuvem e outros CSPs.

Em posse da sua lista e uma vez que os ativos estejam conectados e gerando informações, defina os próximos passos (limpeza, correlação, agrupamentos de eventos repetíveis, recorrência das análises e gatilhos).

5- Exceções: O que não logar

Sim, exceções existem e mantê-las fora de controle e acesso restrito pode lhe gerar um grande risco ou não-conformidade. Abaixo algumas sugestões do que você não deveria estar registrando em seus sistemas de monitoramento e quaisquer operações envolvendo AAA (Authentication, Authorization & Accounting).

Controle	**Descrição**
Código ou *Paylod*	A não ser em situações de *debbuging*, nunca armazene o código-fonte processado ou o conteúdo do *payload* transportado na área de dados (*payload* de protocolos de comunicação e sinalização por vezes podem ser necessários).
Tokens, chaves, credenciais ou *tickets* de sessão	Considere substituir por um valor de *hash forte*, se necessário, para rastrear eventos específicos da sessão. Credenciais, cookies e senhas nunca devem ser armazenadas nem expostas e transferidas (quando necessário) em texto puro. Chaves criptográficas e outros segredos devem ter o mesmo tratamento.
Dados protegidos por rótulos/*tags* ou legislação aplicável	Dados pessoais (PII, documentos, endereços, telefones...), dados sensíveis (orientação sexual, religiosa, política, etnia, condições de saúde, identificadores de governo, pessoas vulneráveis) e dados classificados como sigilo bancário (dados do cliente, da conta bancária, transações financeiras que possam expor/inferir seu proprietário, sua situação ou condição), protegidos por política de classificação específica da cia (uso restrito, interno, sistemas *outside-border*s ou confidencial), conteúdo com classificação de Segurança mais

	alto que o suportado pelo sistema de rastreabilidade/*logging* ou informações protegidas de serem capturadas por serem consideradas ilegais em seu país de origem. Num primeiro momento, não é tão trivial e imediato a identificação deste contexto, mas recomendamos fortemente que esse escrutínio seja realizado com calma.

6- O que mais considerar para o futuro?

Caso esteja trabalhando com o tema pela primeira vez, separe um tempo previamente a sua decisão e faça uma leitura sobre pontos importantes que lhe ajudarão em muito acertar a sua jornada de curto, médio e longo prazos (do ponto de vista de organização, crescimento vegetativo, custos, hierarquia, relacionamento e modelos funcionais). Quais são esses pontos?

- Avalie construir um modelo aproveitando o data warehouse da sua companhia ou criar seu próprio (e específico) datamart;
- Leia sobre BigData, SmallData e SmartData, principalmente sobre o último;
- Invista um tempo avaliando as proposições e diferenças de arquitetura sugeridas pelos mestres Inmon e Kimball;
- Considere utilizar e priorizar produtos e serviços que ofereçam compatibilidade com ETL ou modelo similar. Tornar-se dependente de console proprietária de fornecedor é de longe uma péssima ideia;
- Considere fortemente criar “stage-areas” para suas extrações, servindo como base das transformações (evite acesso diretamente ao dado em real time a não ser que seja algo realmente necessário);
- Consuma somente o que for necessário para seu “contexto”. Se alguém abriu uma discussão dizendo que vai lhe “enviar tudo”, recomece a conversa;
- Nas primeiras extrações, o lixo/resíduo/sujeira pode ser seu melhor amigo (e do pessoal de fraudes/pesquisa). Gaste um tempo com transações nulas, campos em bancos, erros não definidos, eventos cifrados, pacotes fragmentados (esse último, por vezes dão mais medo que qualquer operação volumosa de DDoS), com erros de CRC, offset etc.
- Normalize, padronize, corrija sempre na fonte (MDM – Master Data Mgmt), teste e melhore. Considere fazer a menor quantidade de transformação e conserto de informação possível ao longo do caminho. Se está errado, corrija na origem/fonte. Isso evita processamento desnecessário e que outras pessoas, consumindo da mesma origem, tenham interpretação/dados diferentes para o mesmo conteúdo;
- Considere sempre que possível a premissa do “baixo acoplamento com alta relevância” deve vencer.

A **OmniSec Intelligence & Security** deseja que sua leitura tenha sido prazerosa e enriquecedora. A colaboração é a essência do crescimento, por isso, encorajamos a compartilhar suas ideias, dúvidas, sugestões ou mesmo este conteúdo com pessoas que possam ser beneficiadas com o material. Seja parte do nosso universo de conhecimento, conecte-se e contribua para a expansão coletiva, aberta e contínua. Estamos à disposição para receber suas contribuições, críticas, recomendações e fortalecer todas as nossas produções.

Partial © 2023, OmniSec Intelligence & Security.

Creative Commons. O conteúdo desta página pode ser reproduzido por quaisquer meios. Esta licença permite que interessados remixem, adaptem e criem a partir do trabalho original, mesmo para fins comerciais, desde que citado este link, sua origem e que os interessados licenciem novas criações e alterações sob termos idênticos.

Esta licença costuma ser comparada com as licenças de software livre e de código aberto “copyleft”. Novos trabalhos terão a mesma licença e/ou menos restritiva, portanto, quaisquer trabalhos derivados também permitirão o uso comercial. Esta é a licença usada pela Wikipédia e é recomendada para materiais que seriam beneficiados com a incorporação de conteúdos da Wikipédia e de outros projetos com licenciamento semelhante.

Esta licença permite que interessados aproveitem, adaptem e criem a partir deste trabalho, mesmo para fins comerciais, desde que atribuído o devido crédito e que licenciem as novas criações sob termos idênticos ou menos restritivos. Todos os trabalhos baseados neste conteúdo deverão ter a mesma licença, portanto, quaisquer trabalhos derivados herdarão automaticamente a permissão de uso e compartilhamento.

https://www.omniseccorp.com/

info@omniseccorp.com

+55(19)2042.3240