

Icarus(BATIZE SUA SOLUCAO): Uma solução para coleta de evidências forenses de ataques de injeção na nuvem

Hamilton Fonte II

Escola Politécnica – Universidade de São Paulo (USP) São Paulo, SP, Brasil

Email: hamiltonii@gmail.com

Marcos A. Simplicio Jr.

Escola Politécnica – Universidade de São Paulo (USP) São Paulo, SP, Brasil

Email: mjunior@larc.usp.br

Abstract—A adoção de arquiteturas em nuvem aumenta a cada dia, e com ela também o número de casos em que esse tipo de tecnologia é usada para fins ilícitos. Infelizmente, devido à natureza volátil da nuvem, a tarefa de coletar evidências para análise forense nesse ambiente tem esbarrado em desafios práticos e legais. Este trabalho analisa propostas na literatura voltadas a resolver os principais desafios existentes na coleta de evidências na nuvem, discutindo suas limitações, e então propõe uma solução que cobre coleta, transporte e armazenamento da evidência visando suplantá-las. A solução aqui proposta tem como focos (1) a reprodutibilidade do processo de coleta e (2) a garantia de custódia da evidência; em suma, a abordagem proposta provê uma forma de correlacionar evidências e sua origem virtual, permitindo transportar e armazenar tais dados sem afetar sua credibilidade.

I. INTRODUÇÃO

Técnicas de virtualização, replicação de serviços e compartilhamento de recursos entre múltiplos usuários (multi-inquilinato) proveem a nuvens computacionais uma elevada escalabilidade [26]. Ao mesmo tempo, tais mecanismos também criam uma elevada volatilidade das máquinas virtuais que executam aplicações em nuvem. Afinal, uma aplicação hospedada na nuvem, quando submetida a um pico de uso, pode criar clones das máquinas virtuais que a hospedam e balancear a carga entre elas, de modo a atender à demanda sem prejuízos na qualidade do serviço oferecido. Após esse pico, e com o objetivo de não incorrer em custos desnecessários, as máquinas que foram clonadas são normalmente desativadas, seus recursos liberados e o sistema retorna à capacidade anterior.

Embora interessante do ponto de vista de eficiência e custos, do ponto de vista forense a volatilidade da nuvem traz problemas em caso de ataques. Por exemplo, caso uma das instâncias de máquina virtual criadas temporariamente seja alvo de ameaças que atuam diretamente na sua memória, sem deixar rastros em discos ou em arquivos de log, as evidências desse evento podem ser completamente perdidas. Essa dificuldade é ainda agravada por aspectos como multi-inquilinato e multi-jurisdição típicas de soluções em nuvem [2]. Especificamente, o aspecto multi-inquilino dificulta a obtenção do hardware

que executa as aplicações de interesse, pois, como ele é compartilhado por vários usuários, removê-los para análise poderia levar a uma violação de privacidade dos usuários não relacionados à investigação. Já a característica distribuída da nuvem pode levar à alocação de informações relevante à investigação em vários países, dificultando a obtenção das mesmas em especial quando não existem acordos de cooperação entre as instituições e/ou países envolvidos [3]. Combinadas, tais características da nuvem dificultam a coleta de evidências com a credibilidade necessária para que elas possam ser usadas em processos legais, o que costuma exigir o respeito à privacidade, à jurisdição e à cadeia de custódia, bem como a reprodutibilidade do processo de coleta [27].

Embora existam soluções na literatura que abordam a coleta de informações de nuvem com o propósito de análise forense, a grande maioria aborda a coleta, o transporte e o armazenamento o faz de forma isolada. Por exemplo, trabalhos como [6] e [7] tratam de fatores como multi-inquilinato e multi-jurisdição, discutindo formas de coleta e preservação da evidência fora da nuvem. Já estudos como [4] focam na forense ao vivo para a coleta de evidência das máquinas virtuais, enquanto trabalhos como [8] aborda a questão de processos de garantia de cadeia de custódia em ambientes de nuvem para transporte da evidência. Por outro lado, não foram identificados na literatura propostas de solução que (1) descrevam como o dado é coletado e armazenado observando a cadeia de custódia, e (2) permitam garantir que, mesmo que uma máquina virtual seja desalocada, haja condições de se reproduzir o processo de coleta de evidências.

O presente trabalho visa suplantiar tais limitações, por meio de uma proposta que tem como focos (1) a reprodutibilidade do processo de coleta e (2) a garantia de custódia da evidência. Em suma, a abordagem aqui descrita provê uma forma de correlacionar evidências e sua origem virtual, permitindo transportar e armazenar tais dados de modo a preservar sua credibilidade. Esta proposta supõe que o sistema sendo monitorado é executado dentro de um container em nuvem. O foco da solução em containers se justifica pelo crescimento

da adoção de containerização nos últimos anos e a previsão de que este será o mais usado modelo de implementação [28]. A solução tem como algo específico ataques de injeção de código [17], pois estes, quando usados contra uma arquitetura em nuvem, não deixam rastros quando máquinas virtuais são desativadas e seus recursos liberados [29], [17]. Em particular, têm especial interesse quatro tipos específicos dessa família de ameaças [17]: **MARCOS: TODO: talvez seja prematuro listar aqui os tipos de ataques (pode valer a pena mover para seção posterior). Por enquanto, vamos deixar aqui mesmo.-** **HAMILTON: OK**

- **Injeção remota de bibliotecas:** Um processo malicioso força o processo alvo a carregar uma biblioteca em seu espaço de memória. Como resultado, o código da biblioteca carregada executa com os mesmos privilégios do executável em que ela foi injetada. Esta estratégia, comumente usada para instalar malwares, pode fazer com que uma biblioteca maliciosa armazenada no sistema seja distribuída por vários processos de uma mesma máquina, dificultando sua remoção [24].
- **Inline Hooking:** Um processo malicioso escreve código como uma sequência de bytes diretamente no espaço de memória de um processo alvo e força este último a executá-lo. O código pode ser, por exemplo, um script de shell.
- **Injeção reflexiva de biblioteca:** Um processo malicioso acessa diretamente a memória de um processo alvo, inserindo nela o código de uma biblioteca na forma de uma sequência de bytes, e então força o processo a executar essa biblioteca. Nesta forma de ataque, a biblioteca maliciosa não existe fisicamente; isso torna esta estratégia de injeção de código potencialmente mais atrativa, pois o carregamento da biblioteca não é registrado no sistema operacional e, portanto, o ataque torna-se mais difícil de ser detectado [25].
- **Injeção de processo vazio:** Um processo malicioso dispara uma instância de um processo legítimo no estado “suspensão”; a área do executável é então liberada e realocada com código malicioso.

O restante deste documento está organizado da seguinte forma. A Seção II discute brevemente soluções em nuvem e suas características. A Seção III apresenta a evolução da computação forense e os desafios que soluções em nuvem trouxeram para a área. A Seção IV analisa os trabalhos relacionados na área de forense de memória. A Seção V detalha a solução proposta e avalia como ela trata os desafios descritos na Seção III. Finalmente, a Seção VI apresenta algumas considerações finais e discute ideias para trabalhos futuros.

II. ADOÇÃO DE ARQUITETURAS EM NUVEM

MARCOS: Discutir a importância da nuvem faz mais sentido na Introdução, no “Contexto Geral”. Se o leitor chegou até aqui, ou ele está convencido da importância da nuvem (=seu contexto geral) e não é necessário lembrá-lo, ou ele parou de ler lá atrás, por não ter gostado do seu contexto

geral... Informações sobre modelos de nuvem (deixando claro seu foco em IaaS) e containers são interessantes de incluir no documento, mas não acho que condizem com o título da seção... SUGESTÃO: Renomeie a seção para “Fundamentação Teórica” e explique nela, com sub-seções sucintas, os principais conceitos que interessam para o seu trabalho. NOTA: Isso normalmente é a última seção que se escreve, pois só no final você vai saber todos os conceitos necessários para o seu documento

Nuvem computacional é um modelo de infra estrutura onde recursos compartilhados configuráveis acessíveis via rede são provisionados e descartados com esforço mínimo de gerenciamento ou recurso de um provedor de serviço. [30] Há três modelos principais de comercialização de uso da nuvem [30]: plataforma como serviço (PaaS) onde se provê infra estrutura para que o cliente instale seu software, software como serviço(SaaS) onde se provê o software que será usado pelo cliente e, o tipo mais pertinente para este trabalho, Infraestrutura como serviço (IaaS) onde se provê recursos computacionais fundamentais.

Containers Linux (LXC), uma tecnologia para auxiliar no gerenciamento de containers para isolamento de recursos **MARCOS: Aliás, acho que o termo correto em português é “contêiner”, não? - HAMILTON: o dicionário tem as 2 grafias** introduzida em 2008, proveram uma série de ferramentas para tirar vantagens das funcionalidades de cgroups e namespaces do kernel do Linux. Este conceito tem três implementações, LMCTFY **MARCOS: REFERÊNCIA**, Rocket **MARCOS: REFERÊNCIA** e Docker **MARCOS: REFERÊNCIA**. A adoção de containers para a implantação de software tem crescido muito para nas aplicações baseadas em nuvem. Segundo o “Container Market Adoption Survey 2016”, realizado pelas empresas DevOps.com (<https://devops.com/>) e ClusterHQ (<https://clusterhq.com>) com 235 empresas que tem desenvolvimento de software suportando ou como sua atividade fim, 76% dos respondentes utilizam containers para melhorar a eficiência do processo de desenvolvimento e em suas arquiteturas de micro serviços em nuvem.

III. FORENSE DE MEMÓRIA EM NUVEM

MARCOS: Hum... o que você de fato uso desses conceitos no seu trabalho? Pergunto porque está muito com cara de “aula de história”, e você pode imaginar o quanto um revisor da área de exatas gostava de aula de história no colégio... Na boa, vai direto aos pontos que interessam, deixando esse blá-blá-blá para outro lugar (pode incluir na dissertação, que não tem limite de páginas, mas não em um artigo) A evolução da forense digital pode ser descrita em 3 fases [19]. A primeira ad-hoc se caracteriza pela falta de estrutura, processos, ferramentas e objetivos. Nesta fase, evidências apresentadas em processos legais eram descartadas com base em erros procedurais e falta de garantias de acurácia ou cadeia de custódia. Foi nesta fase que acusados de crimes usando um computador eram soltos sob justificativa de terem tido sua privacidade violada pelos investigadores. Na segunda fase começa a estruturação da forense, surgem as primeiras políticas e processos de coleta,

armazenamento, transporte e análise da evidência. O primeiro processo nesta fase foi proposto por Palmer G. em 2001 na primeira Digital Forensics Research Conference e ficou conhecido como DIP **MARCOS: É uma sigla? Se for, dê o nome completo.** Em 2002 Reith M. propôs o Abstract Digital Forensics Model que adicionava ao DIP estágios que faltava a este. Em 2003 Carrier e Spafford propuseram o Integrated Digital Investigation Process baseado nas técnicas e teorias da forense física. Finalmente em 2004 Baryamureeba e Tushabe propuseram o Enhanced Digital Investigation Process Model, uma evolução de Carrier e Spafford. Importante notar que esses modelos foram todos propostos antes que a forma atual de computação em nuvem estivesse disponível. [20] Nesta fase surgiram as ferramentas que tinham por objetivo principal coletar evidências de forma que fossem legalmente aceitáveis, para isso precisavam atender requisitos de reprodutibilidade, confiabilidade e preservação da evidência. A terceira fase se caracteriza pela migração do ferramental de soluções pontuais para soluções empresariais. Conceitos como coleta em tempo real e forense como serviço emergem nesta fase.

MARCOS: Mas de novo? Esse texto traz alguma coisa de novidade com relação à introdução? A utilização crescente de virtualização, ferramentas online e hospedagem em nuvem [1], está criando dificuldades para a coleta de informações, análise e utilização em processos legais [21]. A funcionalidade de elasticidade de carga oferecida pelos provedores de nuvem por meio da qual infraestrutura pode ser alocada e desalocada dinamicamente, trouxe o problema da volatilidade dos dados nas máquinas virtuais. Com algumas ameaças que não deixam evidências em disco [22], a memória de uma máquina virtual despejada de um pool e seus recursos liberados serão para sempre perdidos e com ela evidências importantes. Neste cenário, o simples armazenamento do conteúdo da memória não satisfaz o requisito de reprodutibilidade do processo pois a máquina virtual não existe mais. Tentou-se a abordagem de armazenar constantemente todas as alterações da memória para não perder informações importantes mas tal abordagem agravou o problema do crescente backlog de dados que os investigadores tem para analisar [23] pois salva todo o histórico de alterações da memória da máquina virtual.

O ferramental forense disponível hoje está pouco adaptado a desafios trazidos pela nuvem [3] pois focam em completude e poucos geram evidências aceitáveis em um processo jurídico [7] já que não satisfazem os três requisitos citados anteriormente. A cadeia de custódia é um processo de coleta e armazenamento de evidências que visa garantir que a evidência não foi alterada, destruída ou manipulada por pessoas não autorizadas, é pouco abordada nas soluções existentes hoje.

IV. TRABALHOS RELACIONADOS

Analisamos a literatura voltada a análise forense na nuvem foi analisada a luz dos seguintes conceitos pertinentes a este trabalho.

A. Acessar e coletar as informações de memória das máquinas virtuais em nuvem

Referente a coleta de informações, os autores [7], [5], [6], [4] e [8] focam em coleta "após o fato" pois ela acontece apenas após a intrusão ser detectada. Os processos de coleta descritos nos trabalhos são iniciados de forma manual ou automática via integração com um mecanismo de detecção de intrusão. No caso específico de memória volátil, tal forma de coleta não consegue descrever como era a memória antes da intrusão pois o processo só é acionado depois. A capacidade de saber como era a memória antes do fato é descrita por [17] como necessária para viabilizar a abordagem de coletar o suficiente para realizar a investigação pois permite comparar dois instantâneos de memória e minimizar o volume coletado antes do fato. A única proposta encontrada que leva tal necessidade em consideração é [9] mas propõe que o dado seja armazenado no próprio dispositivo porém essa abordagem não é aplicável ao cenário em nuvem pois leva a perda de informações importantes caso a máquina virtual seja despejada e seus recursos liberados.

Ainda na coleta de informações, os autores [7] e [4] sugerem a abordagem de forense ao vivo onde os dados são constantemente coletados sem distinção do antes ou depois do fato. Os autores [5], [6] e [8] adotam a estratégia de isolar e parar a máquina virtual para em seguida realizar o processo de coleta. Nas duas estratégias citadas anteriormente, o problema do grande volume de informações coletadas não é abordado pelo autores nem o cenário onde é necessário coletar evidências de uma máquina virtual que já foi despejada do pool e os recursos liberados. Atender este último cenário é importante pois com as soluções em nuvem que escalam automaticamente, as evidências de uma máquina vítima de um ataque que foi despejada de um pool com a diminuição da demanda serão para sempre perdidas. Analisando a proposta de [5], parece ser possível cobrir o cenário mencionado mas ele não dá detalhes da implementação suficiente para termos certeza.

B. Capacidade de reproduzir o processo e obter os mesmos resultados

A reprodutibilidade do processo de coleta é uma dos requisitos para garantir a cadeia de custódia da evidência e sua aceitação em um processo legal. Cadeia de custódia esta relacionado a credibilidade e ter dois analistas reproduzindo o processo de coleta de memória chegando ao mesmo conjunto de evidências tem um peso muito forte em termos de credibilidade. Neste tópico, nenhuma das propostas encontradas até o momento consegue reproduzir os mesmos resultados ao repetir o processo no cenário em que uma máquina virtual é despejada da nuvem e seus recursos liberados pois todas elas dependem da existência da máquina virtual para a repetição da coleta. Analisando a proposta de [4] parece que é possível mas o autor não dá detalhes de implementação suficientes para termos certeza.

C. Não violar privacidade ou jurisdição das partes não envolvidas na investigação

No caso das soluções em nuvem, não é possível remover o hardware para análise pois ele contém informações de vários usuários, alguns dos quais não estão envolvidos na investigação em curso, fazê-lo levaria a violações de privacidade, o que diminui a credibilidade da evidência. A maioria dos autores resolve este problema adequadamente e podemos listar duas estratégias usadas. Os autores [7], [4], [5] e [6] usam estratégias de coletar dados pertinentes a investigação e armazená-los fora da nuvem enquanto que [8] e um caso específico de [4] dependem da cooperação do provedor de serviços de nuvem para conseguir as informações necessárias à investigação. Dependendo do provedor de serviços de nuvem é considerada uma estratégia fraca pela comunidade forense pois o foco do provedor de nuvem é garantir a continuidade do serviço e não a coleta de evidências.

D. Garantir a cadeia de custódia da evidência

Na garantia da cadeia de custódia apenas [8] aborda a questão, mas toma cuidados somente para garantir que a evidência não foi destruída ou alterada através do cálculo de hashing da mesma mas não explica como impede o acesso não autorizado. As propostas dos outros autores estão focadas apenas no aspecto técnico da coleta, nenhum deles menciona garantia de custódia, apenas que as evidências são coletadas de forma "forensicamente aceitável".

A Tabela 1 mostra um comparativo das soluções estudadas.

TABLE I
COMPARATIVO DE SOLUÇÕES MARCOS: DE QUÊ? LEGENDAS
DEVEM SER DESCRITIVAS O SUFICIENTE PARA QUE SEJAM
ENTENDIDAS SEM QUE SEQUER EU TENHA LIDO O SEU
TEXTO... MARCOS: INCLUA A SOLUÇÃO AQUI PROPOSTA
COMO A ÚLTIMA OU PRIMEIRA LINHA DESTA TABELA

	Coleta é contínua?	Reproduz o processo sem a VM?	Garante cadeia de custódia?	Preserva jurisdição e privacidade?
[4]	✗	✗	✗	✓
[5]	✗	✗	✗	✓
[6]	✗	✗	✗	✓
[14]	✗	✗	✗	✓
[7]	✗	✗	✓	✓
[8]	✓	✓	✗	✓
[10]	✗	✗	✗	✓
[13]	✗	✗	✗	✓
[9]	✓	✗	✗	✓
[11]	✓	✗	✓	✓

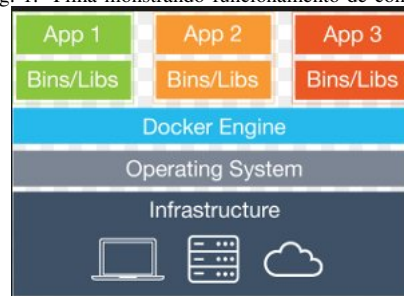
V. SOLUÇÃO PROPOSTA: COLETOR DE MEMÓRIA DE CONTAINER

O presente proposta tem como objetivo principal coletar memória de uma máquina virtual de modo a conseguir: (1) identificar os quatro tipos de ataque listados anteriormente; (2) identificar sua fonte, mesmo se a máquina virtual não existir mais; (3) descrever o sistema antes e depois do incidente. Além disso, deve-se armazenar a memória coletada de modo a garantir sua integridade e confidencialidade, sem violar a jurisdição cabível ou a privacidade de outros usuários no mesmo servidor físico.

A. Descrição

Nas soluções com infra-estrutura física a máquina é persistente, associar uma cópia da memória, a imagem de um disco ou pacotes trafegando na rede a sua origem é uma tarefa simples. Com as soluções de infra virtual, em especial as auto-escaláveis, a máquina deixou de ser persistente e tornou-se volátil. Para resolver o problema da identificação da fonte precisamos encontrar outra forma, persistente, para identificar a fonte da evidência coletada, para isto usaremos containers. Embora o container seja uma peça de software e por consequência também é volátil, a imagem compilada e sua execução na forma de container estão atrelados a um hash que os identificam. A pilha de um container pode ser visto na Figura 1.

Fig. 1. Pilha mostrando funcionamento de container



A solução proposta por este trabalho, para resolver o problema de associação da evidência a sua origem de modo que o processo seja reproduzível, pausa a execução do container e coleta um instantâneo da memória dos processos sob sua execução. Este processo é executado em intervalos de tempo conhecidos de modo a se ter uma evolução da história da memória dos processos. Em um sistema derivado do linux (Ubuntu 14.04) isso foi atingido via cópia do diretório "proc" relacionado aos processos sob o "cgroup" associado ao container e salvo em disco. Para relacionar o instantâneo a sua origem, assinamos o arquivo em que salvamos o instantâneo da memória com o hash da imagem como mostrado na Figura 2.

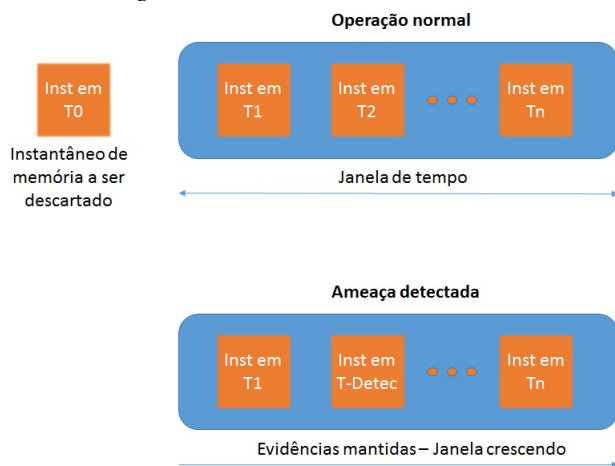
As técnicas forenses praticadas hoje estão voltadas para a obtenção da informação em sua totalidade, seja via cópia bit a

Fig. 2. Evidência salva - hash do container e imagem

```
rw-r--r-- 1 root root 13376 Jul 18 22:40 4b0953884935d80421133400130290429778acc85dfed7366e23a9d19425d1d-8fa80cdda11002f45c835254343bce274fa
27e1136708a8e4c13ecf57d6053-3522-18-07-2016-10-40.mem
rw-r--r-- 1 root root 12782 Jul 18 22:40 4b0953884935d80421133400130290429778acc85dfed7366e23a9d19425d1d-8fa80cdda11002f45c835254343bce274fa
27e1136708a8e4c13ecf57d6053-3522-18-07-2016-10-40.mem
rw-r--r-- 1 root root 12782 Jul 18 22:40 4b0953884935d80421133400130290429778acc85dfed7366e23a9d19425d1d-8fa80cdda11002f45c835254343bce274fa
27e1136708a8e4c13ecf57d6053-3522-18-07-2016-10-40.mem
rw-r--r-- 1 root root 12782 Jul 18 22:40 4b0953884935d80421133400130290429778acc85dfed7366e23a9d19425d1d-8fa80cdda11002f45c835254343bce274fa
27e1136708a8e4c13ecf57d6053-3522-18-07-2016-10-40.mem
rw-r--r-- 1 root root 12782 Jul 18 22:40 4b0953884935d80421133400130290429778acc85dfed7366e23a9d19425d1d-8fa80cdda11002f45c835254343bce274fa
27e1136708a8e4c13ecf57d6053-3522-18-07-2016-10-40.mem
rw-r--r-- 1 root root 13354 Jul 18 22:40 87f609cc438812334017a9211230b36c3b71b5e8d0808040631ee1c8625e142d-8fa80cdda11002f45c835254343bce274fa
27e1136708a8e4c13ecf57d6053-3522-18-07-2016-10-40.mem
```

bit, seja por remoção do hardware [15] [16]. Tais práticas tem levado ao crescente volume de dados que os investigadores tem que analisar. Há uma vertente na comunidade chamada “sniper forensics” onde se coleta e armazena o suficiente para a investigação. A solução proposta por este trabalho acompanha esta tendência. A questão foi definir a quantidade de dados “suficiente” para uma investigação. De acordo com [17], detectar intrusões na memória de processos depende de termos uma descrição da memória antes e depois da intrusão. Com base nisso decidimos que “suficiente” seria a quantidade necessária para descrever o sistema antes e depois do ataque. A idéia é implementar um log rotativo de instantâneos de memória cobrindo uma quantidade de tempo configurável, integrar a solução com algum sistema de detecção de ameaça de modo que, ao detectar um ataque, o log passa de rotativo a completo assim permitindo que se conheça o sistema antes e depois do ataque como mostrado na Figura 3.

Fig. 3. Janela deslizante de coleta de evidência



De modo a não violar a jurisdição de outros países ou a privacidade de outros usuários por causa do caráter multi-inquilino e multi-jurisdição das arquiteturas em nuvem pública, a solução proposta por este trabalho foi o de armazenar a evidência em um local físico fora da nuvem, utilizando como transporte conexão segura. Outro ponto importante é garantir a cadeia de custódia da evidência ou seja, garantir que a evidência não foi destruída, alterada ou acessada por qualquer pessoa. Assim a solução proposta por este trabalho usará de armazenamento físico fora da nuvem, o transporte será feito por TLS, no momento do armazenamento calcularemos o hash da evidência e o acesso ao mesmo será

controlado.

Tendo a implementação sido bem sucedida conseguiremos analisar e identificar as formas de ataque enumeradas nos objetivos.

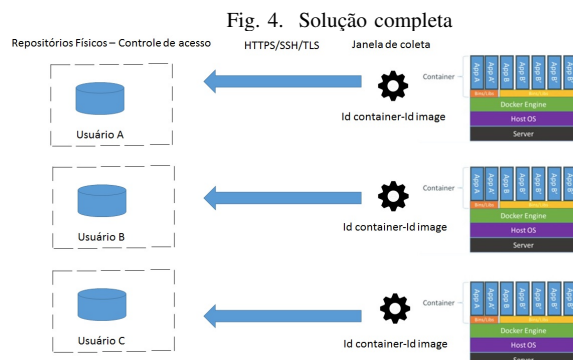
B. Implementação

A implementação da solução foi realizada em um notebook intel I5 de 2.30Mhz e 4Gb de RAM com sistema operacional de 64 bits. Nele, usando Oracle Virtual Box 5.0 criamos uma máquina virtual com 2 Gb de memória RAM emulando apenas 1 processador. Na máquina virtual instalamos a versão 1.10 do Docker engine e 1.21 da API, criamos 3 containers, cada um rodando um nginx 1.0 em diferentes portas. Foi escrita uma aplicação em JAVA que descobre qual o PID associado a cada container e salva o `/proc/pid/numa_maps` em um arquivo. A cópia e gravação do arquivo acontece da seguinte forma: a cada minuto a aplicação pausa o container em questão, tira uma cópia do `numa_maps`, salva em um arquivo .mem e concatena com o hash de identificação da imagem. Em seguida verifica qual o arquivo .mem mais antigo em disco, se for mais velho que o tempo 't', o arquivo é descartado.

C. Limitações

A solução esta focada em coletar informações de memória do user space, ela não enxerga o kernel space. Técnicas de investigação de malware que se baseam em informações do kernel space como por exemplo, a comparação de informações do Process Environment Block (PEB), que ficam no user space, com informações do Virtual Address Descriptor (VAD), que fica no Kernel space não são possíveis. Outro exemplo é a análise de ameaças que realizam manipulação direta dos objetos do kernel (*D.K.O.M. - Direct Kernel Object Manipulation*) também não se beneficiam de associação com o container.

A solução completa com todos os elementos descritos anteriormente pode ser visto na figura 4



VI. CONCLUSOES FINAIS

Até o momento a presente proposta teve sucesso em relacionar o instantâneo de memória a sua origem através do hash de identificação da imagem, usou-se a versão 1.10 do Docker para tal fim. A versão é importante pois até a 1.9.x, o

identificador da imagem era apenas um randômico. Na versão 1.10 ele passou a ser um hash calculado a partir da imagem.

Apesar do sucesso em salvar a memória relacionadas ao container e sua associação com a origem, não foi possível até o momento relizar a análise. As ferramentas de leitura de memória disponíveis no mercado dependem que todo o conteúdo da memória da máquina esteja disponível para realização da análise. Como coletamos apenas a memória relacionada aos processos, o ferramental não funciona. É necessário o desenvolvimento de uma ferramenta que trabalhe sem a memória completa da máquina, ou no caso da ferramenta *Volatility* é necessário a criação de um profile para a memória do processo.

REFERENCES

- [1] AMAZON. *Amazon Media Room Press Release*. [S.l.], 2016. 2 p.
- [2] KEYUN, R. et al. *Advances in Digital Forensics IV*. 7. ed. Orlando: [s.n.], 2011. 35–46 p. ISSN 1098-6596. ISBN 9788578110796.
- [3] DYKSTRA, J.; SHERMAN, A. T. Acquiring forensic evidence from infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques. *Digital Investigation*, Elsevier Ltd, v. 9, n. SUPPL., p. S90–S98, 2012. ISSN 17422876.
- [4] GEORGE, S.; VENTER, H.; THOMAS, F. Digital Forensic Framework for a Cloud Environment. In: CUNNINGHAM, P.; CUNNINGHAM, M. (Ed.). *IST Africa 2012*. Tanzania: International Information Management Corporation, 2012. p. 1–8. ISBN 9781905824342.
- [5] POISEL, R.; MALZER, E.; TJOA, S. Evidence and cloud computing: The virtual machine introspection approach. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, v. 4, n. 1, p. 135–152, 2013. ISSN 20935374 (ISSN).
- [6] DYKSTRA, J.; SHERMAN, A. T. Design and implementation of FROST: Digital forensic tools for the OpenStack cloud computing platform. *Digital Investigation*, Elsevier Ltd, v. 10, n. SUPPL., p. S87–S95, 2013. ISSN 17422876.
- [7] REICHERT, Z.; RICHARDS, K.; YOSHIGOE, K. Automated forensic data acquisition in the cloud. *Proceedings - 11th IEEE International Conference on Mobile Ad Hoc and Sensor Systems, MASS 2014*, p. 725–730, 2015.
- [8] SANG, T. A log-based approach to make digital forensics easier on cloud computing. *Proceedings of the 2013 3rd International Conference on Intelligent System Design and Engineering Applications, ISDEA 2013*, p. 91–94, 2013.
- [9] DEZFOULI, F. N. et al. Volatile memory acquisition using backup for forensic investigation. *Proceedings 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic, CyberSec 2012*, p. 186–189, 2012.
- [10] DOLAN-GAVITT, B. et al. Virtuoso: Narrowing the semantic gap in virtual machine introspection. *Proceedings - IEEE Symposium on Security and Privacy*, p. 297–312, 2011. ISSN 10816011.
- [11] BAAR, R. B. van; BEEK, H. M. A. van; EIJK, E. J. van. Digital Forensics as a Service: A game changer. *Digital Investigation*, Elsevier Ltd, v. 11, p. S54–S62, 2014. ISSN 17422876.
- [12] ZHANG, L.; ZHANG, D.; WANG, L. Live Digital Forensics in a Virtual Machine. In: *2010 International Conference on Computer Application and System Modelling (ICCASM 2010)*. [S.l.: s.n.], 2010. v. 6, p. 328–332.
- [13] ALJAEDI, A. et al. Comparative Analysis of Volatile Memory Forensics. *IEEE International Conference on Privacy, Security, Risk and Trust (PASSAT) and IEEE International Conference on Social Computing (SocialCom)*, p. 1253–1258, 2011.
- [14] BARBARA, D. *Desafios da perícia forense em um ambiente de computação nas nuvens*. [S.l.], 2014.
- [15] SIMOU, S. et al. Cloud forensics: Identifying the major issues and challenges. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 8484 LNCS, p. 271–284, 2014. ISSN 16113349.
- [16] BEM, D. et al. Computer Forensics - Past, Present and Future. *Journal of Information Science and Technology*, v. 5, n. 3, p. 43–59, 2008.
- [17] CASE, A. et al. *The Art of Memory Forensics: Detecting malware and threats in Windows, Linux and Mac memory*. Kindle edi. [S.l.]: Wiley, 2014.
- [18] SOUSA, F. R. C.; MOREIRA, L. O.; MACHADO, J. C. Computação em Nuvem: Conceitos, Tecnologias, Aplicações e Desafios. *II Escola Regional de Computação, Ceara, Maranhão, Piauí (ERCEMAPI)*, v. 1, n. EDUFPI, p. 150–175, 2009.
- [19] CHARTERS, I.; SMITH, M.; MCKEE, G. The Evolution of Digital Forensics. In: *Techno Forensics 2008 Conference*. [S.l.: s.n.], 2008. p. 1–39.
- [20] GRISPOS, G.; STORER, T.; GLISSON, W. Calm before the storm: the challenges of cloud computing in digital forensics. *International Journal of Digital Crime and Forensics*, v. 4, n. 2, p. 28–48, 2012. ISSN 1466640073.
- [21] SHARMA, H.; SABHARWAL, N. Investigating the Implications of Virtual Forensics. *Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on*, p. 617–620, 2012.
- [22] RAFIQUE, M.; KHAN, M. N. A. Exploring Static and Live Digital Forensics: Methods, Practices and Tools. *International Journal of Scientific & Engineering Research*, v. 4, n. 10, p. 1048–1056, 2013.
- [23] QUICK, D.; CHOO, K. K. R. Impacts of increasing volume of digital forensic data: A survey and future research challenges. *Digital Investigation*, Elsevier Ltd, v. 11, n. 4, p. 273–294, 2014. ISSN 17422876.
- [24] Miller, M. and Turkulainen, J. Remote Library Injection. www.nologin.org, 2004.
- [25] Fewer, Stephen Reflective DLL Injection. harmonysecurity.com, 2008.
- [26] MORSY, A. M.; GRUNDY, J.; MULLER, I. An Analysis of the Cloud Computing Security Problem. *APSEC Cloud Workshop. Sydney, Australia: [s.n.], 2010*.
- [27] RAHMAN, S.; KHAN, M. N. A. Review of live forensic analysis techniques. *International Journal of Hybrid Information Technology*, v. 8, n. 2, p. 379–388, 2015.
- [28] PIRAGHAJ, S. F. et al. A Framework and Algorithm for Energy Efficient Container Consolidation in Cloud Data Centers. *Proceedings - 2015 IEEE International Conference on Data Science and Data Intensive Systems; 8th IEEE International Conference Cyber, Physical and Social Computing; 11th IEEE International Conference on Green Computing and Communications and 8th IEEE International Conference on Internet of Things, DSDIS/CPSCoM/GreenCom/iThings 2015*, p. 368–375, 2016.
- [29] VÖMEL, S.; STÜTTGEN, J. An evaluation platform for forensic memory acquisition software. In: *The Digital Forensic Research Conference*. Monterey, Ca: [s.n.], 2013. v. 10, n. SUPPL., p. S30–S40. ISBN 1742-2876. ISSN 17422876.
- [30] MELL, P.; GRANCE, T. The NIST definition of cloud computing. *NIST Special Publication*, v. 145, p. 7, 2011. ISSN 00845612.