TECNICHE DI CLOUD FORENSICS

La quantità di servizi basati su cloud per il consumatore ha aumentato notevolmente la capacità di elaborazione e di archiviazione. Molte applicazioni basate sul modello client-server sono in realtà servizi PaaS su cloud, utilizzati in modo trasparente al consumatore. Oggi la Cloud Forensics assume pertanto un ruolo importante nelle indagini elettroniche, ma può potenzialmente portare ad elaborare un'enorme quantità di dati e quindi ad un rallentamento nella conduzione delle indagini stesse. Una soluzione proposta è l'uso della riduzione dell'evidenza per formare un processo parallelo attraverso la rimozione di dati noti irrilevanti.

Fabio MASSA è un graduato dell'Arma dei Carabinieri e Digital Forensic Examiner. È Vice Presidente Nazionale dell'Associazione Nazionale Giuristi e Informatici Forensi ANGIF. Ha maturato la sua esperienza decennale nell'ambito delle investigazioni tecnico-scientifiche presso la Sezione Investigazioni Scientifiche del capoluogo Ligure. È esperto in indagini digitali, attualmente svolge il ruolo di consulente tecnico in materia di informatica forense presso il Tribunale di Genova, Chiavari e per le Forze dell'Ordine. Ha conseguito le certificazioni statunitensi presso The American Institute of Forensic Science – Youngsville, NC, in qualità di Evidence Collector Specialist and Crime Scene Tecnology, Drug Identification Analisys, Ufed Physical pro certified Examiner, Accessdata Certified Examiner and Certified Trainer ACE, BloodStain Pattern Documentation Expert, Forensic Examiner AISF, Bloodstain Pattern Analysis advanced Expert, Blasting Insider IRE.

1. Introduzione

La definizione più affidabile di *cloud computing* è stata fornita dal *National Institute of Standards and Technology* (NIST), che afferma che il cloud computing è un modello per consentire l'accesso *on-demand* a un pool condiviso di risorse informatiche configurabili, ad esempio reti, server, storage, applicazioni e servizi, di cui è possibile eseguire rapidamente il *provisioning* e il rilascio con il minimo sforzo di gestione o interazione del fornitore di servizi. Questa definizione descrive il modo in cui opera un sistema di cloud computing piuttosto che definire una tecnologia esatta, architettura o insieme specifico di servizi forniti dal fornitore. Questa generalizzazione è indicativa delle sfide che l'investigatore forense deve affrontare in qualsiasi indagine che coinvolga un sistema con *provisioning cloud* o che utilizzi un servizio basato su *provisioning cloud*. Il NIST divide il cloud computing in tre modelli di servizio complementari, delineato nelle seguenti articolazioni:

1) **Infrastructure as a Service (IaaS)**: IaaS, fa riferimento alla fornitura di hardware virtuale, come server, storage e networking, che consente ai clienti di costruire un'infrastruttura virtuale che imita l'hardware fisico del computer tradizionale. Il provider IaaS più popolare è Amazon Web Services (AWS) che include l'offerta di istanze di *cloud computing on demand Elastic Compute Cloud* (EC2) e di *Storage as a Service* (StaaS), Simple Storage Service (S3).

2) **Piattaforma come servizio (PaaS)**: PaaS, opera a un livello sopra l'hardware di elaborazione grezzo visualizzato. La PaaS fornisce metodi per lo sviluppo di strumenti che interagiscono facilmente con servizi come database, server Web e archiviazione di file. Questi servizi sono sottratti ai vincoli dello spazio di archiviazione fisico sottostante, alla pianificazione della replica e della ridondanza e al bilanciamento del carico. Alcuni esempi di servizi PaaS includono Google App Engine e Force.com.

3) **Software come servizio (SaaS)**: Considerato come il più probabile *entry level del cloud computing* per la maggior parte degli utenti e delle aziende, SaaS è definito dal NIST come "*la capacità fornita al consumatore di utilizzare le applicazioni del provider in esecuzione su un'infrastruttura cloud*". Forse il provider SaaS più noto è Salesforce.com, specializzato in suite di vendita e marketing ospitate nei loro *data center* ma a cui si accede, tramite connessioni con licenza, dai sistemi del cliente. I clienti sfruttano le capacità dei servizi back-end Oracle e SAP senza dover pagare per l'installazione, la licenza o la manutenzione di queste soluzioni software e hardware. Più di recente, Salesforce.com ha aggiunto Data Analytics *as a service* alla propria linea di prodotti. La maggior parte dei modelli SaaS sono pubblicizzati con un "*pay for what you use*". ubblicizzati con un "*pay for what you use*".

Un altro tipo di SaaS è presente anche sotto forma di *Application Service Provider* (ASP) che sub-licenzia software ai clienti che possono accedervi sull'hardware dell'ASP. Un esempio di ciò potrebbe configurarsi nei servizi forniti da alcuni provider di *web hosting*, dove i server di posta elettronica che eseguono Microsoft Exchange possono essere noleggiati con prezzi stabiliti dal numero di licenze richieste. L'ASP fornisce il portale di accesso al server e allo spazio di archiviazione, ma lascia l'amministrazione dell'applicazione al client. Il client, tuttavia, ha un controllo limitato o nullo sull'infrastruttura o sui componenti del server sottostanti e potrebbe richiedere il supporto dell'ASP per apportare modifiche significative al di là della gestione dei contenuti.

Fig. 1: IaaS vs PaaS vs SaaS: livello di controllo.

Il *cloud computing mobile*, molto comune nelle attuali indagini forensi, invece, offre agli utenti mobili la potenza di elaborazione, archiviazione e funzionalità aggiuntive normalmente non disponibili localmente sul dispositivo stesso. Con l'avvento delle connessioni dati sempre attive nel mondo del mobile computing, la maggior parte delle indagini forensi eseguite sui dispositivi mobili, possono essere aiutate con la raccolta di evidenze digitali nel cloud.

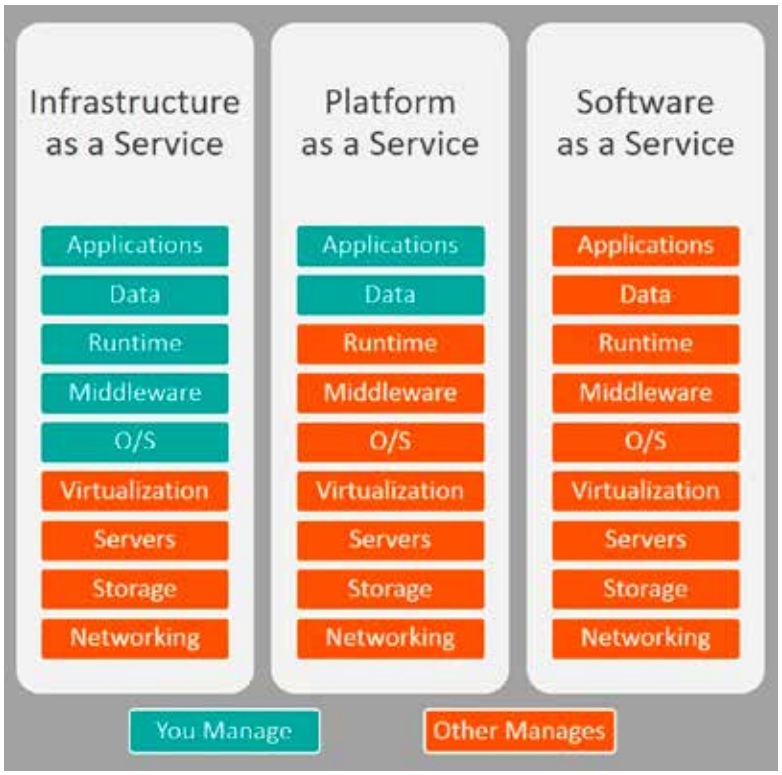

Fig. 2: IaaS vs PaaS vs SaaS: controllo e management.

Le prove digitali recuperate dal cloud dei backup mobili possono contenere messaggi di testo, foto, video, dati delle applicazioni, e il recupero di questi dati di backup mobili potrebbe essere l'unico metodo per accedere a dati da un dispositivo mobile compromesso a causa della crittografia o del degrado del dispositivo. Le tecniche di *cloud computing forensic* richiedono una combinazione di molte diverse competenze forensi digitali a seconda del tipo di cloud oggetto di indagine.
Per fornire flessibilità e scalabilità, i sistemi cloud sono generalmente costruiti su un ambiente virtualizzato con risorse allocate dinamicamente. Per utilizzare in modo efficiente l'hardware sottostante, gli ambienti virtuali, anche quelli non utilizzati per i sistemi cloud, vengono generalmente sovrascritti.
Quando un'indagine coinvolge un IaaS, StaaS, PaaS o SaaS, la metodologia per esaminare una rete, una condivisione di file, una workstation o un'applicazione diventa più difficile da applicare poiché l'ambito dell'indagine raggiunge una serie di potenziali massimi anziché un valore effettivo. In un sistema cloud la quantità di spazio su disco rigido allocata in un volume si espanderà e si ridurrà in modo dinamico, la quantità di RAM disponibile nel pool di risorse varierà e non sempre corrisponderà alla RAM totale disponibile e così via. Questa incredibile complessità, alla base del meccanismo che fornisce un'allocazione e una gestione dinamica delle risorse, è impercettibile all'utente finale.

Lo stesso, tuttavia, non si può dire per il *digital forensic examiner*, poiché, nella maggior parte dei casi, solo nel corso di un'indagine egli scopre che lo spazio di archiviazione, come ad esempio quello di uno smartphone sospetto, è in realtà un *front-end* per un contenitore di numerosi gigabyte o terabyte fornito su un servizio cloud. Lo smartphone potrebbe essere semplicemente utilizzato come punto di ingresso per un'intera rete di server e client ospitati in più data center distribuiti in tutto il mondo. Un decreto per una semplice perquisizione e sequestro di un dispositivo potrebbe trasformarsi in un incubo giurisdizionale internazionale.
Dal punto di vista relativo alla Digital Forensics, il cloud computing può essere considerato un'arma a doppio taglio. Da un lato, la raccolta di prove digitali da fonti cloud può comportare complicate sfide tecniche e legali trasversali, dall'altro, l'impiego di funzionalità di archiviazione ed elaborazione su cloud possono accelerare il processo di indagine forense utile a concentrare la stessa esclusivamente sui dati pertinenti.
In un ambiente di *cloud computing*, i processi forensi digitali, tradizionali metodi per indagare sui cyber crimini perpetrati in un ambiente di cloud computing, sono limitati a causa della complessità dell'ambiente cloud stesso. I sistemi di cloud computing che utilizzano file system distribuiti dispongono di aree di

archiviazione di grandi dimensioni distribuite fisicamente in molte posizioni geografiche. L'applicazione degli attuali metodi forensi non può essere utilizzata perché è impossibile creare immagini e ricostruire repliche separate di ciascun nodo del disco. I tradizionali processi di acquisizione digitale includono il mantenimento del controllo della catena di custodia dei dati delle prove digitali, controllo che avviene nella fase di raccolta delle prove attraverso l'imaging di un sistema. La tecnologia applicata nel *cloud computing,* nella conduzione di una indagine forense digitale, interrompe questo passaggio iniziale e presenta un problema per gli investigatori poiché non è possibile creare un'immagine forense di un ambiente così ampio.
È necessario sviluppare un'architettura forense per ambienti di cloud computing, poiché molti degli aspetti chiave della corretta acquisizione, gestione e analisi delle prove, come il controllo delle prove, le capacità di acquisizione e gli strumenti forensi, necessitano di ulteriore sviluppo per soddisfare i requisiti per acquisire correttamente le prove digitali negli ambienti di cloud computing. Gli attuali strumenti, processi e competenze maturi per le indagini digitali si concentrano su piccoli ambienti individuali. Gli strumenti forensi diventano instabili quando i file del caso diventano troppo grandi e settimane o mesi di lavoro vengono annullati se il file del caso creato non risponde costantemente perché la capacità di dati è troppo grande per essere gestita dallo strumento.

2. Cloud Forensics

La Cloud Forensics è definita come l'applicazione della Digital Forensics negli ambienti di cloud computing. Tecnicamente, consiste in un approccio forense ibrido alla generazione di prove digitali. Dal punto di vista organizzativo, coinvolge le interazioni tra gli attori del cloud allo scopo di facilitare le indagini sia interne che esterne. Legalmente implica spesso situazioni multigiurisdizionali e *multi-tenant*. Attualmente la raccolta di evidenze digitali cloud implica prima l'identificazione della posizione in cui sono archiviati i dati e il sequestro dell'hardware di archiviazione pertinente per clonare la macchina virtuale e, successivamente, l'esecuzione di analisi "tradizionali" sui dati acquisiti. Tuttavia, i sistemi cloud possono ospitare più reti di *tenant* IaaS/PaaS o archiviare archivi di dati di diverse organizzazioni. Ciò può potenzialmente portare a un'enorme quantità di dati da elaborare che a sua volta può portare a massicci rallentamenti della produttività nelle indagini.
Una soluzione proposta è l'uso della riduzione dell'evidenza per formare un processo parallelo attraverso la rimozione di dati noti irrilevanti e la deduplicazione. Questo parallelo tratterebbe un volume di dati inferiore e tenterebbe di rispondere a domande meno ambigue e quindi fungerebbe da forma di *triage* per il processo investigativo nel suo insieme. Solo se il *triage* risultasse un'attività potenzialmente sospetta, l'intero set di dati sarebbe stato esaminato. Questo processo di riduzione presenta una nuova serie di sfide come il metodo di raccolta, il processo di riduzione stesso per garantire la minor perdita di dettagli possibile e quale tecnica di *data mining* può essere utilizzata che richieda una minima inferenza possibile. La raccolta di dati da un'istanza cloud è stata definita come la richiesta della collaborazione di terze parti e possibilmente di software specializzato fornito dal CSP tramite un'API.
La fiducia di questa API è una preoccupazione giustificabile in quanto è l'unico modo per riunire tutte le cache di dati disparate che formano un'istanza guest VM sul cloud, sia che si tratti di un intero sistema operativo, di un dispositivo di archiviazione o di un'istanza di un software applicazione. In generale, una tecnica accettata per l'acquisizione forense di un sistema basato su cloud consiste nell'eseguire uno snapshot del sistema sospetto e utilizzare tale snapshot come fonte di prova, mitigando potenzialmente la necessità di eventuali tempi di inattività del sistema durante il processo investigativo.

A - Analisi forense sul cloud remoto

L'analisi forense remota è il recupero di dati ospitati da una terza parte su un sistema che può essere o meno controllato direttamente dal proprietario dei dati. Le analisi forensi eseguite su sistemi remoti vengono generalmente eseguite solo quando l'investigatore non può ottenere l'accesso fisico al sistema che ospita i dati. Ciò potrebbe essere dovuto al fatto di non sapere dove si trova il sistema (come può essere il caso in indagini che coinvolgono reti anonime) o per motivi logistici come la pratica di distribuzione dei dati da parte del CSP. La raccolta remota di prove digitali può essere utile per estendere la finestra di acquisizione delle prove digitali, in cui i dati locali non sono più recuperabili a causa di corruzione, sovrascrittura o crittografia.

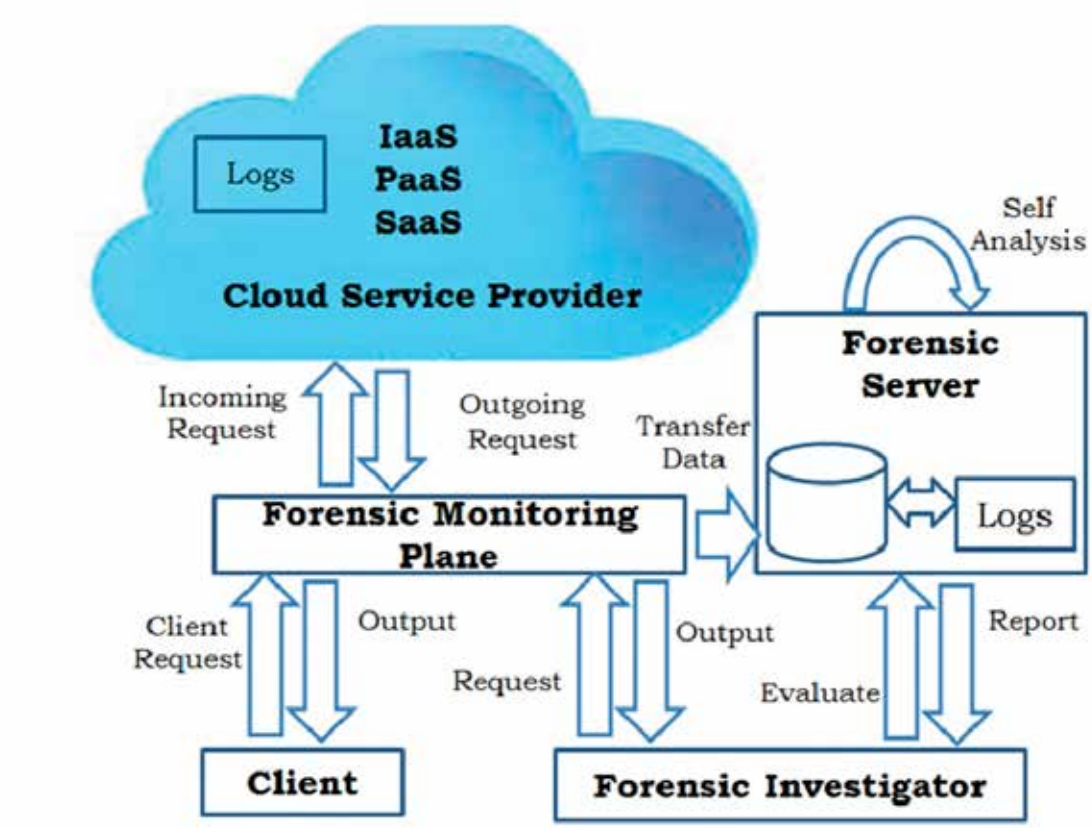

Fig. 3: Architettura di Cloud Forensics.

Tale acquisizione remota richiede la stessa attenzione ai dettagli delle normali indagini forensi sul cloud, poiché è necessario prestare attenzione per garantire che solo i dati sospetti vengano estratti dalla posizione remota. Qualsiasi altra cosa può essere considerata una ricerca ingiustificata o addirittura una violazione della privacy di un individuo. L'alternativa più sicura è richiedere che il CSP fornisca i dati dai propri sistemi in un formato che possa essere utilizzato dall'esaminatore. Questo è attualmente il modo in cui vengono condotte le indagini forensi che coinvolgono gli ISP. Una richiesta di informazioni viene effettuata nell'ambito di un quadro concordato e i dati richiesti vengono restituiti, se detenuti, dall'ISP.

B. Analisi forense delle macchine virtuali

Un Cloud è descritto come un insieme di servizi allocati dinamicamente ad un cliente abbonato, e presentato come un prodotto o servizio finale senza l'esposizione dei meccanismi sottostanti. Attualmente, il modo più efficiente per gestire questa allocazione è attraverso l'uso della virtualizzazione. La virtualizzazione implica l'inserimento di uno strato di astrazione tra il "*bare metal*" e il sistema operativo che il client vede (es. una raccolta di sistemi che l'abbonato può organizzare e collegare (IaaS)). L'esecuzione di una indagine forense in un tale contesto può essere un compito impegnativo, con una serie indefinita di restrizioni e limitazioni, non ultima la questione delle risorse condivise.

Un altro problema si verifica quando l'investigatore non ha accesso al sistema operativo guest, oppure quando il sistema sospettato di attività criminali è stato crittografato in qualche modo. Se il sistema è stato fornito dal CSP ma il sistema operativo dell'utente finale è stato configurato dal Cliente, allora non c'è modo di leggere i dati contenuti all'interno dell'istanza virtuale. Tutto ciò che esiste indipendentemente dal sistema operativo sarebbe accessibile, come la RAM e la connettività di rete (ma non il contenuto del traffico a meno che non venga inviato in un formato non crittografato). In tal caso, l'investigatore potrebbe essere in grado di utilizzare un metodo noto come Introspezione della macchina virtuale in cui il VMM viene utilizzato per fornire una modalità ristretta di input e output da e verso il sistema operativo guest. Un uso di questo percorso di I/O basato su VMM è eseguire l'iniezione del kernel che può indurre il sistema operativo guest a eseguire un comando. In un'indagine

forense quel comando potrebbe essere quello di aprire l'accesso a una porta, avviare un servizio o persino fornire informazioni sul suo stato attuale e sui suoi contenuti.

C. Analisi forense sul cloud privato
Un'indagine forense digitale che coinvolge un cloud privato non è limitata nella sua metodologia come quelle che coinvolgono implementazioni di cloud pubblico, ibrido o di comunità. Il proprietario del cloud è anche l'abbonato e quindi la cooperazione dell'uno significa la cooperazione dell'altro che può essere volontaria o per ordine d tribunale. Inoltre, a seconda delle dimensioni dell'organizzazione proprietaria del cloud, è probabile che la distribuzione tra i *datastore* sia molto meno dispersiva e potrebbe anche essere limitata a due o tre siti (un sito principale, un sito secondario per disponibilità elevata e ripristino di emergenza/continuità operativa e un sito finale ridondante per scopi di backup e di emergenza) e questi siti possono o meno essere ospitati nelle immediate vicinanze.

Una metodologia forense per indagare, ad esempio, su ownCloud, un sistema cloud open source destinato a essere distribuito come StaaS privato, il client utilizzato per accedere al sistema ownCloud, viene sfruttato per enumerare i contenuti della condivisione StaaS fornita e l'accesso al servizio di cloud storage gestito attraverso un utilizzo controllato del client in indagine.

L'enumerazione viene eseguita utilizzando la sincronizzazione lato client e i metadati di gestione dei file ed è possibile creare una sequenza temporale dalla creazione/modifica e dai *timestamp* di sincronizzazione registrati. Inoltre, è possibile raccogliere evidenze da file memorizzati nella cache nonché eventuali log o artefatti di autenticazione con il servizio di archiviazione stesso (ricerche DNS, cronologia URL, cookie, certificati ecc.). Per il server stesso, nonostante il moniker cloud, dovrebbe essere affrontato come verrebbe gestito qualsiasi hypervisor standard. Se sufficientemente piccolo, sarebbe necessario eseguire una copia forense dell'intero server, altrimenti i file di registro devono essere esportati dal VMM e dall'hypervisor e i sistemi sospetti devono essere clonati su un'unità esterna prima di essere sospesi, oppure potrebbe essere necessario acquisire un'istantanea e quindi esportarli come immagine per l'analisi *post mortem*.

Nel caso di ownCloud, i metadati da estrarre dall'hypervisor e dal VMM vengono archiviati come database SQLite o MySQL e devono essere raccolti utilizzando le tecniche forensi dedicate.

D. Archiviazione come servizio (StaaS Forensics)
Sebbene sia la soluzione privata più importante, ownCloud non è l'unico sistema StaaS disponibile e l'accesso mobile a dati coerenti o l'accesso collaborativo alle risorse ha stimolato la disponibilità di soluzioni StaaS basate su cloud come Dropbox, in cui i file possono essere archiviati e condivisi con altri utenti tramite un'applicazione o una console basata sul Web. Molto interessante, è la caratteristica che identifica lo "spazio slack online" simile allo *slack space* di una unità fisica, in cui i dati risiedono nella parte inutilizzata di un singolo blocco su un disco rigido ove, nel caso specifico, i dati vengono archiviati in blocchi su un sistema Dropbox senza attribuzione a un sistema di proprietà. Proprio come il cloud è in continua evoluzione, così anche il software. Gli sviluppatori cambiano e alterano le strutture per fornire funzionalità aggiuntive, chiudere le falle di sicurezza identificate o semplicemente adattarsi ai cambiamenti tecnologici del backend e del client.

3. Conclusioni

Il cloud computing offre sfide tecniche e legali significative agli investigatori digitali e le soluzioni a molti di questi problemi sono ancora agli albori. Molte delle soluzioni esistenti proposte a questi problemi si basano sulla cooperazione dei Cloud Service Provider. I CSP sono in grado di fornire strumenti e sistemi per l'applicazione della legge attraverso la cooperazione di tipo forense, ovvero metodi di acquisizione dei dati, gestione completa dei registri, provenienza dei dati sicura e affidabile, ecc.

La fornitura di questi sistemi potrebbe non apportare vantaggi ai profitti per il CSP, ma dovrebbe rimanere una priorità da fornire sotto un punto di vista etico e legale. Attualmente esistono numerosi strumenti a disposizione delle FF.OO. per eseguire, in particolar modo sull'ambiente cloud mobile, l'acquisizione dei contenuti di cloud backup di dispositivi mobili sottoposti ad indagine, come ad esempio Ufed Cellebrite Cloud Analyzer, Oxygen Forensics Cloud Extractor, Magnet Forensics Cloud. ©