Introduzione

I dati non sono più dove ci si aspettava che fossero

> ... Quando qualche anno fa parlai per la prima volta di Digital Forensics a una conferenza italiana, notai delle espressioni di compatimento, tipo quelle che di solito si riservano a un parente noto per non essere troppo sano di mente e che, nel mezzo di un matrimonio, si lancia in una delle sue filippiche. Un magistrato mi disse che se mi interessava questo campo avrei dovuto "andarmene in America". Inutile dire che quella conferenza non va annoverata tra i miei successi...

Quando l'editore mi ha chiesto di pensare alla terza edizione del libro, per prima cosa ho riletto l'edizione precedente da zero. Pur contenendo principi e metodologie che rimangono tutt'ora validi (per fortuna!), il testo sembra scritto in un'altra era.
Solo dopo questa analisi mi sono reso conto di quanto sia cambiato il mondo dell'informatica in così pochi anni.
L'uomo della strada ha sentito più volte il termine *cloud*, e probabilmente non ha il concetto così chiaro nella propria mente. Il tecnico magari ha pensato all'inizio che "cloud" fosse solo una ripacchettizzazione di molte tecnologie sotto un nuovo nome più attrattivo a livello di marketing.
Invece è stata davvero una rivoluzione culturale. Ed è stata così silenziosa e pervasiva che in poco tempo tutti noi ci siamo trovati a usare paradigmi nuovi senza nemmeno accorgercene.
E così il mondo è cambiato e noi ci siamo sorpresi nel duplice ruolo di spettatori e consumatori a vedere tramontare un'era, quella dell'informatica personale.
Questo libro è stato scritto per intero tramite un PC in cloud (Iaas), una suite OpenOffice installata sotto EyeOS (Iaas), ascoltando musica in streaming tramite Google Music (Saas) e sfruttando come device di data entry una decina di sistemi diversi tra cui notebook, sistemi desktop, workstation Unix, tablet e cellulari.
Apple con iPhone e iPad ha lanciato su scala mondiale una rivoluzione. Chi ha potuto ha inseguito e cercato di giocarsela ad armi pari (Google con Android), altri sono arrivati tardivamente (Microsoft con Windows RT e Windows Phone) pur con prodotti apprezzabili; altri sono collassati pur essendo stati i leader precedenti (RIM, Palm e Nokia).
La profezia di Gates, che sostiene che nessun leader di un'era informatica ha mai dominato anche la successiva, si è dimostrata ancora una volta drammaticamente vera.

Se questo incipit può sembrare fuori tema rispetto a una pubblicazione che parla di Digital Forensics, si rifletta sulle implicazioni di questa rivoluzione.
I sistemi personali e aziendali stanno subendo una notevole trasformazione, trasferendosi su fronti totalmente opposti.
Vi è chi ha deciso di esternalizzare tutto, spostando quanto possibile in cloud decentrati, e chi invece ha seguito l'idea di accorpare tutto il possibile in sistemi cloud privati dove client e server sono virtualizzati e dove i dati risiedono tutti su SAN o NAS che servono centinaia di sistemi.
Qualunque sia la strada intrapresa il risultato è sempre lo stesso: i dati non sono più dove ci si aspettava che fossero.
Senza entrare in un ambito enterprise facciamo un esempio banale. Tutta la musica che avevo in casa è passata dal supporto fisico (CD) al file (MP3), per poi essere stata trasferita nella sua interezza su Google Music (circa 30 GB). Tutta la mail è passata da un account tradizionale gestito tramite POP3 su un MacBook, a un servizio cloud (Gmail). Le fotografie sono state caricate su un server privato virtuale con ownCloud.
I documenti sono gestiti tramite Google Drive e Google Docs, o al limite tramite OpenOffice su EyeOS su un server privato.
Certo, ho ancora a casa un piccolo datacenter (è indispensabile per fare analisi forense) ma la maggior parte dei miei dati non è più né su un server né su un NAS all'interno della mia rete.
L'accesso a tutte queste informazioni avviene tramite dispositivi leggeri (tablet e smartphone), apposite app o attraverso servizi web.
Anche la parte di analisi forense è passata da una serie di file/application server fisici a un cluster di due workstation HP e di uno storage Synology che mi esporta i volumi condivisi via iSCSI. Programmi, tool, macchine specializzate, sono tutti installati su una serie di Virtual Machine KVM gestite da due nodi Proxmox VE.
A dispetto di quanto qualcuno possa pensare io, pur essendo un appassionato di tecnologia, non sono certo un *early adopter*, anzi tendo a essere piuttosto conservativo, almeno con le mie cose.
Per quanto riguarda le organizzazioni che seguo nel mio lavoro di tutti i giorni, posso affermare che ormai nessuno mi chiede più un deploy su server fisico. I vantaggi di una struttura virtuale sono tali e tanti che è del tutto normale virtualizzare tutto il back-end, e molti sono i progetti in cantiere per fornire servizi terminal e VDI al fine di eliminare i PC dalle scrivanie.
Tutto questo ha un fortissimo impatto nel mondo della Digital Forensics. Ho la netta sensazione di essere tornato ai primi anni dopo il 2000, quando il campo era pionieristico e non vi erano tool adatti ad affrontare le situazioni che si ponevano di fronte all'investigatore che si accingeva ad affrontare la materia.
Certo, ora ci sono molto ottimi pacchetti commerciali, come X-Ways Forensics, FTK, Encase (continuo a mal sopportarlo ma è uno standard *de facto*), ma al momento sono tutti fermi al concetto di fonte di prova che risiede, perlomeno, su supporto fisico.
Molti di questi pacchetti ora supportano l'analisi di macchine virtuali gestite dagli hypervisor più comuni ma difettano totalmente nella gestione di un'indagine che includa elementi forniti tramite un servizio di tipo cloud.
Anche il fronte mobile presenta seri problemi. In questi anni ho avuto modo di apprezzare strumenti quali UFED di Cellbrite o Oxygen Forensics Suite. Ma per quanto dotati di un ottimo supporto e di funzioni molto potenti questi strumenti esaminano i terminali

mobili a livello di supporto. Permettono di estrarre immagini, SMS, MMS, e-mail, contatti, artifact nella cache del brower, dati e file cancellati, ma peccano miseramente sia nell'analisi dei dati delle varie app presenti sul dispositivo (e non può essere altrimenti, visto che esistono centinaia di migliaia di app per ognuno dei top player del mercato) sia nell'analisi dell'iterazione delle app con i server di back-end presenti in cloud.
Per rendere il quadro più veritiero non si può non parlare di quanto sta accadendo lato back-end o a livello di datacenter. Il deployment massivo di macchine virtuali (che si tratti di macchine private o di servizi di public cloud) ha di fatto ridisegnato tutta la logica dei servizi server. Non si trovano più decine e decine di server specializzati. Troviamo piuttosto macchine con due o quattro processori multicore, con capacità di memoria RAM nell'ordine (per gli ultimi modelli presenti sul mercato all'atto della scrittura di questo libro) di 768 GB o di 1,5 TB.
Un sistema blade può contenere in 10 U fino a 18 server, con 288 core totali e 13,5 TB di RAM. Presupponendo di infilare due sistemi blade e uno storage in un singolo Rack 42U, si può ben immaginare come si possa disporre della potenza di calcolo necessaria per gestire un paio di migliaia di server virtuali.
Infrastrutture di questo genere non si possono affrontare con i comuni strumenti di analisi forense. Le normali metodologie di acquisizione e analisi non si possono applicare in ambienti di questo genere. Si può tranquillamente dire addio alle tradizionali acquisizioni effettuate tramite il boot di una Live distro come DEFT o CAINE. Vi sono problemi legali (il datacenter si può trovare all'estero, l'infrastruttura può essere contenuta all'interno di un tenant in un cloud pubblico, i dati possono risiedere all'interno di un servizio PaaS o SaaS condiviso con migliaia di altri utenti non oggetto di indagine, il contenuto può essere a sua volta fornito da una CDN che risponde in maniera diversa a seconda del paese di origine del fruitore del servizio), procedurali (affrontare uno storage da 2 PB richiede metodi mai esaminati finora) e metodologici (estrarre i dati da un database condiviso tra migliaia di utenze, recuperare i dati da uno snapshot e sfruttare un datacenter in disaster recovery geografico richiedono nel contempo fantasia, adattabilità ma anche un rigore ferreo per rispettare tutte le normative).
Normalmente quando inizio a trattare queste problematiche nei corsi tenuti da IISFA, vedo facce sempre più sgomente. Purtroppo non si parla di fantascienza, quanto piuttosto di una realtà presente e pervasiva.
Nel prossimo futuro aleggia già lo "spettro del Big Data". Ho avuto modo di ascoltare il dottor Joseph Reger, CTO di Fujitsu, al Fujitsu Forum 2012 a Monaco. È il secondo anno che seguo con molta attenzione il suo intervento. Non sarà tra i più noti guru dell'informatica ma ho già avuto modo di apprezzare il suo acume e le sue doti di preveggenza in relazione ai trend dell'informatica attuale e futura. Parlando della sicurezza dei dati nel prossimo Big Cloud ha espresso questo concetto:

> Qualcuno potrebbe pensare che depositare petabyte e petabyte di dati in un unico punto renda più facile gestire la loro sicurezza. Obiettivamente solo il tempo di trasferimento in un'altra locazione da parte di un attacker rende un furto tradizionale praticamente impossibile. Ma è bene considerare un fattore. Possedere miliardi di informazioni provenienti da fonti diverse ma ricercabili tramite un sistema strutturato quale potrebbe essere un database a oggetti, hash table, object file system o altro, può permettere una cosa del tutto nuova, ovvero il furto di informazioni che nemmeno si sa di possedere. Perché ciò che importa in questi ambiti non è tanto il dato grezzo ma la capacità di correlare tra loro le diverse informazioni.

(A questo punto deve essermi sfuggita un'imprecazione ad alta voce dato che le persone a me vicine si sono girate a guardarmi con aria torva.)
Alla luce di questa rivelazione non ho potuto non pensare al fatto che in Italia la legge 48/2008 sia arrivata con sette anni di ritardo rispetto alla conferenza di Budapest del 2001 che ne aveva fissato i principi, che ancora i giudici fatichino a capire il concetto di furto di un file se non legato a un supporto fisico, e al fatto che la nostra professione come specialisti di quella che è una disciplina (per quanto bistrattata) della polizia scientifica non sia ancora riconosciuta a livello di codice penale (e che quindi ci costringe a essere pagati meno di una donna delle pulizie e per di più con tempi biblici). In questa situazione l'idea di entrare in un'aula di tribunale e dover spiegare che l'attacker ha sottratto della conoscenza che il possessore dei dati neanche sapeva di avere è cosa degna di un romanzo di fantascienza.
Ci avviamo verso un mondo dove le informazioni sono distribuite, universalmente accessibili e salvate nel back-end in grandi datacenter dove, nel contempo, sono riversate informazioni provenienti da migliaia di utenti e fonti diverse. La fine dell'informatica personale e l'inizio di quella pervasiva e accentrata.

Andrea Ghirardini
CTO BE.iT

Struttura del libro

Oltre al solito compendio su come muoversi nel panorama della Digital Forensics tradizionale, troverete molte riflessioni, dubbi e stimoli per ricercare nuove soluzioni ai problemi che via via ci verranno posti nel corso del prossimo futuro.
Per molte cose ammetto che non ho una risposta definitiva. Siamo, di nuovo, in un periodo in cui la Digital Forensics torna a essere una scienza di frontiera. Un periodo dove molto è lasciato alla conoscenza approfondita e alla capacità di adattamento del singolo. Sicuramente ci si porranno delle sfide stimolanti dove sarà necessario ripensare a nuovi tool e metodi per poter gestire situazioni completamente nuove e, per certi versi, contrarie a quelle trovate finora sul campo.
Il volume si compone di 18 capitoli. La lettura sequenziale è consigliata ma non obbligatoria: essendo il volume pensato non tanto come la "Bibbia della Digital Forensics" quanto come uno strumento per valorizzare le proprie competenze informatiche nella prassi investigativa, ogni lettore è libero di focalizzarsi sulle parti più attinenti alle proprie necessità che, come ogni Digital Forensics expert sa bene, variano da caso a caso.
Dopo il Capitolo 1, scritto per fare luce su alcuni concetti base, con il Capitolo 2 si presenta una panoramica sulla situazione giuridica italiana inerente la materia.
I Capitoli 3 e 4 sono quindi dedicati all'importante fase dell'acquisizione del dato, mentre il Capitolo 5 apre le porte del laboratorio del Digital Forensics expert.
Nei Capitoli 6 e 7 si apre invece un'ampia parentesi sull'analisi dei file system, su cui si concentra buona parte della prassi investigativa.
Con i Capitoli 8, 9 e 10 si ritorna invece all'illustrazione di importanti strumenti di lavoro.
Giunti a questo punto, il Capitolo 11 propone una metodologia di analisi generale.
I Capitoli 12, 13 e 14 sono rispettivamente dedicati all'analisi dei sistemi Windows, OS X e Linux.
Il Capitolo 15 tratta la gestione e l'analisi dei file di log.
In conclusione, i Capitoli 16, 17 e 18 affrontano le parti più avanzate, ivi comprese la virtualizzazione e il cloud, l'hardware di tipo enterprise e il mondo mobile.

Requisiti per la lettura

Questo volume è stato pensato per chi si avvicina alla pratica investigativa sui sistemi informatici e informativi. Nozioni di programmazione e robuste esperienze di naviga-

zione in Rete sono pertanto assolutamente necessarie, così come una buona confidenza con i principali protocolli di comunicazione e una solida conoscenza delle architetture dei più diffusi sistemi operativi (specialmente Unix). A monte di tutto questo vi è, imprescindibile, la curiosità e un reale interesse verso la scienza nota come *Digital Forensics*.

Convenzioni utilizzate nel testo

Leggendo questo volume vi capiterà di incontrare due particolari box di approfondimento.

Diario di un Digital Forensics expert

Qui Andrea Ghirardini apre le porte dei suoi archivi, integrando quanto in discussione con esempi tratti dai casi di cui si compone la sua esperienza investigativa.

Cosa dice la legge

Qui Gabriele Faggioli fissa l'attenzione su quanto previsto dalla normativa in merito al tema trattato, perché la giurisprudenza gioca sempre un ruolo importante.

Contatti

È possibile contattare gli autori presso i seguenti indirizzi di posta elettronica:
`ghirardini@beitsa.ch` (Andrea Ghirardini)
`gf@gabrielefaggioli.it` (Gabriele Faggioli)

Ringraziamenti

Come sempre nella stesura di un volume, vi è una numerosa serie di persone cui sono dovuti i più sentiti ringraziamenti.
Il gruppo BIG e in particolare Massimiliano Marchese. Potrei dilungarmi a dire tantissime cose su di lui, tutte positive. Alla fine mi limito a un ringraziamento per tutto quello che ha fatto per mettere in piedi BE.iT. Un grazie anche Claudio Cesare Secchi. Mi auguro di poter fare molta strada assieme.
Alessandro Caviglione. Di buono posso dire che se qualcosa ha a che fare con il mondo virtuale, lui lo sa. Punto e senza esclusione alcuna. Di meno buono posso dire che mi ha fatto rivalutare i prodotti Microsoft. Non gliela perdonerò mai.
Fabio Brivio. Doveva essere un aggiornamento. Non lo è stato. Doveva essere un lavoro di pochi mesi. Non lo è stato. Doveva chiudersi eoni fa. Non è successo. Nonostante tutto non ha perso pazienza e speranza. Ultimamente però le nostre conversazioni sono prevedibili. Mi dice sempre: "Andrea, chiudi!". Grazie per tutto.
Corrado Giustozzi. No, dico, si è offerto lui di scrivere la Prefazione! Quando avevo pochi anni leggevo *MCmicrocomputer* e lo idolatravo. Poi mi sono connesso a McLink e nella settimana successiva al mio primo collegamento mi ha fatto una lavata di capo sulla netiquette per un mio cross post un po' invasivo. Ma aveva ragione. Poi mi ha chiesto di scrivere per Byte Italia (mai troppo compianta). E me ne sono vantato per anni. Ora il fatto che il sommo NightGaunt si sia offerto di scrivere la Prefazione mi farà diventare ossessivo con tutti i miei amici. Anzi, lo racconterò ai miei nipoti.
Luigi Ranzato. Secondo i giapponesi un allievo supera il maestro dopo quarant'anni di apprendistato. Non è il suo caso, ci ha messo molto ma molto meno. Difatti ho dovuto ricorrere ai suoi consigli e al suo sapere per lo studio delle shell bag. Grazie.
Stefano Fratepietro e Nanni Bassetti. Per aver reinventato il "made in Italy". Nella pratica della Digital Forensics semplicemente hanno dimostrato che "Italians do it better".
Davide Gabrini. Un hacker. Ma così hacker che anche nei peggiori raduni underground lo accettano di buon grado anche se è un poliziotto. Beh, gioca molto a suo favore anche la simpatia e la propensione per tutto ciò che è buono e/o alcolico. Mi ha fatto scoprire Prezi e tante altre cose. Lo scambio di battute "Rebus tu quante slide hai fatto", "Una, però grande" rimarrà nella storia.
Matteo LK Flora. La dimostrazione più palese che genio e follia possono andare assolutamente d'accordo in una stessa persona. Non solo è bravo nel suo lavoro, non solo sa

fare bene le fotografie, ma nei raduni hacker europei lui non è una primadonna, è una vera diva. Ah, giusto per precauzione, la prossima volta guido io.
Litiano Piccin. *Chapeau*. Le sue conoscenze sul mobile sono semplicemente troppo avanti. Il resto è superfluo.
"Il Lo". L'entusiasmo fatta persona. L'unica persona che pur essendo un nerd allucinante ha il coraggio di dirmi che io lo sono più di lui. Ogni tanto ci sentiamo e mi tira fuori qualche argomento di cui non so nulla, ma con tale entusiasmo che mi costringe a studiarlo. A cosa mi servirà non lo saprò mai, ma intanto è fatta. Ah, ti serve una sedia?
Federica Bertoni. Discutiamo animatamente su tutto. Però ha il suo perché. O no? Discutiamone, però ho ragione io. Grazie per avermi fatto usare (e abusare) del tuo laboratorio e il contributo con le bozze e la parte sul mobile.

Infine, un "grazie" particolare e personale.

A Elga. Dopo quasi 20 anni di matrimonio continua a sopportarmi. E lo fa con amore, suscitando reazioni perplesse in tutti i miei amici nerd. La via per la santità passa anche per questo. E io intanto sono una persona migliore grazie a lei.
L'incontro più fortunato della mia vita e la mia luce del mattino.
A Sirio per allietarmi la vita con le sue molte chitarre e semplicemente per la sua presenza, il che non è poca cosa. Perché ogni volta che fai sentire il tuo humor sagace tua madre si gira e mi dice "Me lo stai rovinando!"?
A Fiamma per essere una ragazzina dolce ma determinata ma soprattutto unica. All'inizio saliva su e mi chiedeva "Papà tutto bene con il libro?", poi "Papà continui a scrivere, posso fare qualcosa?", successivamente "Papà ancora a scrivere?" e ora "Papà ma ancora con 'sto libro???". Si l'ho tirata molto per le lunghe. Scusa se ti ho trascurata.
Ad Alessia. Sei veramente una bella persona (ti convincerai mai di questo?). La prova vivente che l'amicizia esiste, trascende anche la distanza, e che quando c'è è destinata a durare per anni. Mi auguro ancora moltissimi.
A Lina e Antonio. Grazie di tutto. Lina, posso chiamarti "mamma", vero?

Di nuovo a Fabio Brivio per non aver ingaggiato un killer per uccidermi.