

MODULO 05 • S2.1 • AGOSTO 2026

Network e Cloud Forensics

Traffico, log e infrastrutture distribuite:
quando la prova non risiede in un solo sistema

pcap • NetFlow • Zeek • proxy • IDS • audit log

Corso di Laurea Magistrale in Sicurezza Informatica

Al termine del modulo lo studente sarà in grado di

01 • RICOSTRUIRE

Sessioni e flussi da catture PCAP: protocolli, conversazioni, trasferimenti di file e canali di esfiltrazione

02 • CORRELARE

Log eterogenei (proxy, firewall, IDS, DNS) in un quadro probatorio unitario, gestendone sincronizzazione e lacune

03 • ACQUISIRE

Evidenze da ambienti cloud e virtualizzati, nel rispetto dei vincoli di giurisdizione e cooperazione internazionale

Limite strutturale. Con la generalizzazione del TLS, il contenuto del traffico è di norma inaccessibile: l'analisi si sposta sui **metadati** (chi parla con chi, quanto, quando, con quale regolarità), che spesso bastano a rispondere al quesito.

I quattro livelli dell'analisi di rete



Le quattro aree si susseguono nelle slide del modulo nello stesso ordine dello schema: ciascuna presuppone la precedente.

Tre livelli di osservazione e i loro strumenti

Livello	Strumenti	Che cosa restituisce
Pacchetto (full capture)	Wireshark, tcpdump	Il contenuto integrale: ricostruzione di sessioni, file e credenziali ove in chiaro
Sessione e metadati	Zeek, NetworkMiner, Arkime	Chi, con chi, quando, quanto: log strutturati per DNS, HTTP, TLS (JA3), SMB
Flusso aggregato	NetFlow / IPFIX	Volumi e direzioni su lunghi periodi: la vista strategica per individuare anomalie

Pattern di esfiltrazione

Upload anomali verso domini ignoti, DNS tunneling, beaconing verso C2: emergono dalle ricorrenze temporali.

Catena di custodia del PCAP

Punto di cattura, eventuali filtri applicati, perdita di pacchetti e hash del file: senza questi dati la cattura è contestabile.

TLS: che cosa resta leggibile quando il contenuto non lo è

Il TLS cifra il contenuto della comunicazione; dall'handshake e dal flusso restano osservabili:

- **Indirizzi e porte:** chi parla con chi, quando, per quanto
- **SNI:** il nome del server richiesto (salvo Encrypted Client Hello)
- **Certificato** presentato dal server e sua catena
- **Fingerprint del client** (es. JA3/JA4): l'impronta della configurazione TLS, che distingue un browser da un malware
- **Volumi e cadenze:** dimensioni e tempi dei flussi

ESEMPIO • JA3 AL LAVORO

Su una workstation, tutto il traffico TLS mostra il fingerprint tipico del browser aziendale, tranne un flusso ricorrente con un JA3 diverso, associato in letteratura a una famiglia di malware.

Il contenuto resta illeggibile; l'impronta del client, da sola, indirizza l'indagine verso il processo giusto, da confermare poi in memoria (modulo 4).

Cautela peritale: i fingerprint sono indicatori probabilistici (librerie comuni producono impronte comuni) e il quadro evolve con ECH e TLS 1.3. In relazione valgono come indizi orientativi.

Il registro DNS come fonte di prova

Prima di quasi ogni comunicazione c'è una risoluzione DNS: il registro delle query è una cronologia delle **intenzioni di connessione** dell'intera rete, che occupa poco spazio e si presta a interrogazioni molto efficaci.

- **Domini mai visti** o registrati da pochi giorni: tipici dell'infrastruttura d'attacco
- **DGA** (domain generation algorithm): raffiche di nomi pseudo-casuali in gran parte inesistenti (risposte NXDOMAIN)
- **DNS tunneling**: dati esfiltrati dentro le query stesse, con sottodomini lunghi e ad alta entropia, record TXT anomali e volumi di query spropositati verso un solo dominio

ESEMPIO • TUNNELING

```
aGVsbG8taG9zdDEt...  
.x9k2.esempio-c2.net  
(migliaia di query/ora)
```

Sottodomini da decine di caratteri apparentemente casuali, cadenza regolare, sempre lo stesso dominio di coda: il firewall vede «solo DNS» e lascia passare, mentre il log DNS documenta un'esfiltrazione a pacchetti di pochi byte.

Nota evolutiva: DoH/DoT cifrano le query verso resolver esterni, sottraendole all'osservazione di rete. Resta un ambito in più da coprire con log sull'endpoint o resolver aziendali obbligati.

Il ritmo come firma: riconoscere il beaconing

Un impianto malevolo deve «chiamare casa» (il server di **command & control**) per ricevere ordini. Lo fa a intervalli, spesso con un **jitter** (variazione casuale) per non essere troppo regolare. Il contenuto è cifrato, perciò la firma va cercata nel **pattern temporale**:

- Connessioni brevi e piccole, ripetute verso la stessa destinazione
- Intervalli quasi costanti (es. $60\text{ s} \pm 10\%$): la distribuzione dei tempi tra connessioni è stretta attorno a un valore
- Attività anche quando nessun utente è presente (notti, festivi)
- Risposte di dimensione quasi identica: il «nulla da fare» del C2

ESEMPIO • DAI NUMERI

Da conn.log: 1.389 connessioni in 24 ore verso lo stesso host, mediana dell'intervallo 62 s, deviazione minima; payload in uscita sempre 312–340 byte.

Un utente umano non produce questa regolarità; un aggiornamento legittimo può produrla, ma verso domini noti del fornitore. L'inventario del software permette di distinguere i due casi.

Ponte al modulo 8: questo è il pattern che l'anomaly detection trova su scala, con gli stessi falsi positivi (NTP, telemetrie, monitoraggi) che qui si imparano a escludere a mano.

NetFlow: osservare la rete dai metadati

Quando la cattura completa (PCAP) è impraticabile per volumi e durata, i flussi **NetFlow/IPFIX** registrano solo i metadati di ogni conversazione: IP e porte, protocollo, byte, pacchetti, orari. Poiché escludono il contenuto, permettono di conservare a costi ridotti una **memoria storica di mesi** dell'intera rete.

Per l'analista forense i flussi rispondono alle domande strategiche: **quando è iniziato il contatto** con l'infrastruttura ostile, quali host interni hanno parlato con essa, quanti dati sono usciti e verso dove. Il PCAP, quando esiste, si usa poi in modo chirurgico sugli intervalli che i flussi hanno indicato.

Sul piano giuridico, i metadati di flusso pongono meno problemi di proporzionalità e privacy della cattura integrale del contenuto (modulo 11).

ESEMPIO • QUANTIFICARE L'ESFILTRAZIONE

Dopo un'intrusione, l'azienda deve notificare la violazione al Garante descrivendone natura e portata. Il PCAP copre solo le ultime 48 ore, mentre i flussi coprono sei mesi.

Dai flussi: 12 sessioni verso l'IP ostile in tre settimane, 4,3 GB in uscita dal file server, sempre fra le 2 e le 4 di notte. La quantificazione finisce nella notifica ex art. 33 GDPR e nella relazione tecnica.

Un evento, molte registrazioni indipendenti

Un accesso illecito lascia tracce su firewall, proxy, controller di dominio, applicazione e DNS. La forza probatoria sta nella **convergenza di fonti indipendenti**: il singolo log, isolato, è falsificabile.

Questioni metodologiche

- Sincronizzazione (NTP) e normalizzazione dei fusi orari
- Retention: i log scadono: acquisire subito, congelare formalmente
- Integrità: log firmati o esportati con calcolo dell'hash
- Lacune: l'assenza di log è essa stessa un dato da spiegare

Piattaforme di correlazione

Splunk, ELK/OpenSearch: interrogazione su grandi volumi, dashboard, ricerche ripetibili ed esportabili a verbale.

Timesketch: la super-timeline del modulo 3 accoglie anche i log di rete, ordinati sullo stesso asse temporale.

Le query utilizzate vanno riportate in relazione, perché fanno parte del metodo e quindi della prova.

Un accesso illecito ricostruito su quattro fonti

Ora (UTC)	Fonte	Evento registrato
02:14:07	VPN gateway	Autenticazione riuscita dell'account «m.rossi» da IP estero mai visto prima
02:14:31	Controller di dominio	Logon dell'account su un server non assegnato a quell'utenza
02:16–02:41	File server	Lettura massiva di 1.204 documenti dalla cartella commerciale
02:43:12	Proxy in uscita	Upload di 850 MB verso un servizio di storage estraneo all'azienda

Perché la correlazione regge

I quattro sistemi sono indipendenti e amministrati separatamente, e la sequenza è coerente sul piano temporale e logico: falsificarla richiederebbe il controllo di tutte le fonti.

Che cosa NON prova

Che al terminale ci fosse Rossi: la sequenza dimostra soltanto l'uso delle *sue credenziali*. L'attribuzione personale richiede altro (accessi fisici, dispositivo, orari compatibili) o la pista alternativa del furto di credenziali.

Header e autenticazione: SPF, DKIM, DMARC

Gli header documentano il percorso

Ogni server attraversato aggiunge in testa un header **Received:** con IP e timestamp; letti dal basso verso l'alto, questi header ricostruiscono il tragitto. Il campo **From:** è invece una dichiarazione del mittente, liberamente modificabile.

Primi indicatori di anomalia: Message-ID, X-Mailer, fusi orari.

I tre controlli di autenticazione

SPF: autorizzazione dell'IP mittente da parte del dominio.

DKIM: firma del dominio e integrità del contenuto firmato.

DMARC: allineamento con il From visibile e politica in caso di esito negativo.

Esiti: **Authentication-Results** del messaggio originale.

ESEMPIO • Frode «CEO fraud»: il CFO riceve un ordine di bonifico da un dominio quasi identico a quello aziendale (rn al posto di m). SPF e DKIM risultano validi perché autenticano il dominio contraffatto. L'analisi forense documenta la somiglianza tipografica, la registrazione del dominio 48 ore prima e il percorso dei server.

La navigazione nei log di proxy e web server

Il log del **proxy aziendale** è la controparte server-side della cronologia del browser: utente autenticato, URL, metodo, byte, user-agent, esito. Sopravvive alla pulizia del client e copre tutti i dispositivi della rete. I log del **web server** (access log, error log) documentano invece chi ha visitato un sito.

La lettura richiede cautele: il timestamp segue il fuso del server, il **NAT** può nascondere molti utenti dietro un solo IP e l'user-agent è facilmente falsificabile. I pattern di attacco (scansioni, path traversal, SQL injection nei parametri) emergono con filtraggio e aggregazione, oggi assistiti da strumenti di anomaly detection (moduli 7 e 8).

I log di terzi (hosting, CDN) si richiedono con i canali giuridici del modulo 11; la conservazione breve impone richieste tempestive di congelamento.

ESEMPIO • L'INSIDER E IL PROXY

Un dipendente nega di aver caricato documenti su un servizio di file sharing personale. Il suo portatile risulta «ripulito».

Il log del proxy mostra, con la sua utenza autenticata, POST multipli verso il servizio per 2,1 GB complessivi, in orari compatibili con i badge di ingresso. Tre fonti indipendenti (proxy, autenticazione, badge) convergono: la pulizia del client non ha toccato i log server-side.

La prova presso terzi: tecnica e giurisdizione

SaaS

Il provider detiene dati e log: acquisizione via **API ufficiali** o export certificati (audit log Microsoft 365, Google Workspace), con documentazione di ambito, momento e account utilizzati.

IaaS e container

Snapshot di VM e volumi come equivalente funzionale della copia forense; per i container, effimeri per natura, contano immagini, orchestratore e log centralizzati.

Giurisdizione

Dati fisicamente all'estero: ordine europeo di produzione (regolamento e-Evidence), MLAT, cooperazione diretta col provider. Le procedure giuridiche condizionano i tempi tecnici.

Nel cloud la nozione di «supporto originale» perde senso: l'integrità si fonda su export riproducibili, hash, marche temporali e sulla documentazione del procedimento di estrazione, che sostituisce il riferimento al reperto fisico.

Ricostruire un'esfiltrazione dal traffico di rete

Laboratorio

- Analisi di un PCAP di esfiltrazione con **Zeek** e **Wireshark**
- Ricostruzione con **NetworkMiner**
- Correlazione con log proxy e DNS

DELIVERABLE: ricostruzione dell'incidente

Lecture obbligatorie

- Davidoff & Ham, *Network Forensics*
- Documentazione dei log Zeek
- Regolamento e-Evidence: nota di sintesi

Verifica

La ricostruzione deve rispondere a cinque domande: chi, cosa, quando, come, verso dove, ciascuna con l'evidenza che la sostiene.

Prossimo modulo: S2.2 Cloud, identità ed e-evidence.

In rete la prova è distribuita: nessuna fonte è autosufficiente e le conclusioni richiedono correlazioni verificabili.



Grazie per l'attenzione

Vincenzo Calabrò

Domande, materiali del corso e contatti
sul sito del docente



vincenzocalabro.it