
This is the **published version** of the bachelor thesis:

Rivera Vila, Roc. *Disseny i implementació d'un sistema d'automatització per la gestió d'alertes en un SOC de Nivell 1*. Treball de Final de Grau (Universitat Autònoma de Barcelona), 2026 (Enginyeria Informàtica)

This version is available at <https://ddd.uab.cat/record/334430>

under the terms of the  license.

Disseny i implementació d'un sistema d'automatització per a la gestió d'alertes en un SOC de Nivell 1

Roc Rivera Vila

29 de juny de 2026

Resum – Els analistes de Nivell 1 d'un Centre d'Operacions de Seguretat (SOC) dediquen bona part de la jornada a gestionar alertes repetitives, fet que augmenta el temps de resposta (MTTR) i provoca fatiga. Aquest treball dissenya i implementa un sistema que automatitza de punta a punta la resposta a les alertes més habituals de Nivell 1, sobre Microsoft Sentinel com a plataforma integrada de SIEM i SOAR. La solució s'estructura en una arquitectura modular d'agents especialitzats —detecció, orquestració, enriquiment, identitat i notificació— que col·laboren mitjançant APIs seguint els principis de mínim privilegi i traçabilitat. S'ha muntat un laboratori al núvol amb ingesta d'atacs simulats, regles de detecció i playbooks de resposta per a tres escenaris: phishing, atacs de força bruta i malware, incloent-hi accions de contenció reals sobre un directori d'identitat dedicat. Els resultats confirmen que el cicle complet de detecció, contenció, notificació i documentació s'executa de forma autònoma i reproduïble, sense intervenció humana.

Paraules clau – SOC de Nivell 1, automatització de seguretat, SIEM, SOAR, Microsoft Sentinel, resposta a incidents, playbooks, orquestració de seguretat

Abstract – Tier 1 analysts in a Security Operations Center (SOC) spend much of their time handling repetitive alerts, which increases response time (MTTR) and causes fatigue. This work designs and implements a system that fully automates the response to the most common Tier 1 alerts, built on Microsoft Sentinel as an integrated SIEM and SOAR platform. The solution follows a modular architecture of specialized agents —detection, orchestration, enrichment, identity and notification— that collaborate through APIs under the principles of least privilege and traceability. A cloud laboratory was deployed with simulated attack ingestion, detection rules and response playbooks for three scenarios: phishing, brute-force attacks and malware, including real containment actions on a dedicated identity directory. Results confirm that the full detect–contain–notify–document cycle runs autonomously and reproducibly, without human intervention.

Keywords – Tier 1 SOC, security automation, SIEM, SOAR, Microsoft Sentinel, incident response, playbooks, security orchestration



1 INTRODUCCIÓ

EN el panorama actual de la ciberseguretat, els Centres d'Operacions de Seguretat (SOC) representen la primera línia de defensa de les organitzacions davant de ciberatacs. L'estructura operativa d'un SOC es

divideix en tres nivells d'especialització per optimitzar la resposta als incidents:

Nivell 1 (Triatge): És el punt d'entrada de totes les alertes. Els analistes monitoritzen les consoles, filtren falsos positius i resolen incidents coneguts o de baixa complexitat seguint procediments establerts.

Nivell 2 (Resposta): S'encarrega de la investigació profunda i resolució d'incidents escalats pel Nivell 1 que requereixen un anàlisi més tècnic.

Nivell 3 (Cerca d'amenaces): Especialistes dedicats a la recerca proactiva d'amenaces no detectades (Threat Hunting).

Els analistes de Nivell 1 d'aquests centres inverteixen una

- E-mail de contacte: 1532654@uab.cat
- Menció: Tecnologies de la Informació
- Treball tutoritzat per: Carles Garrigues Olivella (DEIC)
- Curs 2025/26

gran part de la seva jornada laboral en la gestió i resolució d'alertes repetitives i de poc valor afegit. Aquestes tasques manuals generen un increment significatiu en el temps de resposta (MTTR - Mean Time To Respond) i produeixen una sobrecàrrega operativa i fatiga en el personal tècnic.

Per fer front a aquesta problemàtica, aquest projecte proposa el disseny i la implementació d'un sistema d'automatització avançat orientat a la gestió d'alertes d'un SOC de Nivell 1, utilitzant Microsoft Sentinel com a plataforma integral de Gestió d'Esdeveniments i Informació de Seguretat (SIEM) i d'Orquestració, Automatització i Resposta (SOAR).

Aquest projecte es desenvolupa en col·laboració amb l'empresa Minsait, una consultora que disposa d'un equip especialitzat en ciberseguretat, el qual ha identificat aquesta problemàtica com una necessitat recurrent en múltiples organitzacions i clients.

Per tal de demostrar l'eficàcia d'aquesta solució, el projecte se centrarà en l'automatització de la resposta davant de tres de les amenaces més recurrents i crítiques que col·lapsen el Nivell 1: els atacs de phishing, les anomalies d'autenticació (com ara els inicis de sessió fallits o login failures) i la detecció de malware.

El projecte preveu la configuració de regles analítiques i el desenvolupament de fluxos de treball automatitzats dins de l'entorn de Sentinel. Aquestes automatitzacions interactuaran mitjançant APIs amb l'ecosistema de seguretat per executar de forma autònoma accions de contenció, tals com l'aïllament de màquines compromeses per codi maliciós, el bloqueig de credencials davant d'accessos anòmals i la interacció directa amb els usuaris afectats per campanyes de phishing.

El propòsit final d'aquest treball és demostrar que la transformació d'aquests procediments manuals en processos completament autònoms permet absorbir la pràctica totalitat d'aquestes alertes sense intervenció humana. Això facilitarà l'evolució del model operatiu del SOC, permetent que els analistes es dediquin a la investigació d'incidents complexos i a la millora de la postura de seguretat global de l'organització.

2 OBJECTIUS

2.1 Objectiu Principal

Dissenyar, desenvolupar i implementar un sistema d'automatització i orquestració de seguretat basat en Microsoft Sentinel, capaç d'analitzar i respondre de forma autònoma a les alertes de Nivell 1 eliminant la necessitat d'intervenció manual en les tasques de triatge i contenció bàsica.

2.2 Objectius Específics

Per assolir l'objectiu principal, es defineixen els següents objectius operatius en ordre cronològic:

1. **Definir el model de resposta i matriu de responsabilitats:** Identificar les accions tècniques necessàries per a cada tipus d'atac i definir els agents encarregats d'executar-les.
2. **Desplegar la infraestructura de Microsoft Sentinel:** Configurar l'entorn de treball i connectar les fonts de

dades a Sentinel, i establir els requeriments d'arquitectura per connectar el SOAR amb els sistemes de contenció definits.

3. **Desenvolupar les regles de detecció i correlació:** Implementar regles de detecció personalitzades a Microsoft Sentinel per garantir una identificació precisa de cadascun dels tres escenaris (phishing, anomalies d'autenticació i malware), minimitzant els falsos positius per assegurar un llançament fiable de les automatitzacions.
4. **Implementar els fluxos d'orquestració:** Crear els fluxos lògics automatitzats perquè cada agent executi les seves accions definides per cada escenari sense necessitat de validació humana per a les alertes de Nivell 1.
5. **Validar l'eficàcia mitjançant simulació d'atacs:** Realitzar proves controlades de phishing, intents d'autenticació mitjançant força bruta i execució de fitxers maliciosos per verificar que el sistema respon en temps real i que les accions de contenció s'executen correctament segons la matriu definida (Fig. 7 de l'apèndix).
6. **Analitzar la millora operativa:** Avaluar la reducció de la càrrega de treball dels analistes i el temps de resposta (MTTR).

3 METODOLOGIA

Aquest projecte es desenvolupa seguint una metodologia en cascada (Waterfall). Aquesta metodologia es basa en una estratègia lineal on les diferents etapes del projecte s'organitzen de manera seqüencial, de forma que cada fase s'inicia un cop finalitzada l'anterior. Aquest enfocament s'ajusta als requeriments d'implementació del projecte, ja que permet definir prèviament l'arquitectura, els fluxos de treball i les responsabilitats abans de procedir amb el desenvolupament.

En aquest projecte, un dels principals riscos és fer una implementació prematura sense tenir el flux i les responsabilitats ben definits. La metodologia en cascada permet estructurar el desenvolupament de manera ordenada i progressiva, assegurant que cada etapa queda correctament validada abans de continuar amb la següent.

Les fases del projecte són les següents:

1. **Anàlisi i Disseny:** En aquesta fase inicial es defineixen els requeriments del sistema i es determinen els casos d'ús que seran automatitzats. A continuació es dissenyen els fluxos lògics (playbooks) a seguir per tots els agents implicats. Un playbook és una seqüència d'accions automàtiques predefinida que el sistema executa pas a pas davant d'un determinat tipus d'incident. Aquesta etapa és crítica per establir les responsabilitats de cada component.
2. **Desplegament de la infraestructura:** Un cop definit el disseny, es procedeix al muntatge de l'entorn tècnic. Això inclou la preparació de la infraestructura virtual, la instal·lació i configuració de la plataforma SOAR i la configuració dels connectors i APIs necessaris per a la comunicació entre les diferents eines de seguretat implicades en el projecte.

3. **Desenvolupament:** Durant aquesta fase s'implementen les funcionalitats definides en la fase de disseny. També es duen a terme les tasques d'integració entre les diferents tecnologies.
4. **Proves:** L'última etapa consisteix en la validació del sistema. Es realitzen simulacions d'incidents per verificar que els playbooks i scripts executen les accions establertes i no s'activen davant de falços positius.

4 PLANIFICACIÓ

Per poder complir amb les fites, els objectius i la metodologia definida, es defineix la següent organització, amb tasques dins de cada fase del projecte.

Fase	Tasca	Període
Anàlisi i disseny	Definició dels casos d'ús (phishing, login i malware)	Setmana 3 - 4 febrer
Anàlisi i disseny	Definició dels playbooks	Setmana 2 - 3 de març
Anàlisi i disseny	Disseny de la infraestructura	Setmana 2 - 3 de març
Desplegament de la infraestructura	Desplegament del laboratori	Setmana 2 - 4 de març
Desplegament de la infraestructura	Configuració d'APIs	Setmana 4 de març
Desenvolupament	Script phishing	Setmana 1 - 2 d'abril
Desenvolupament	Script login	Setmana 2 - 3 d'abril
Desenvolupament	Script malware	Setmana 3 - 4 d'abril
Desenvolupament	Desenvolupament i integració dels components	Setmana 1 - 3 de maig
Desenvolupament	Redacció progressiva de la memòria	Abril - juny
Proves	Execució de proves	Setmana 4 de maig

Fig. 1: Planificació inicial de les fases del projecte.

Arrel de dificultats i inconvenients que han sorgit en la fase de Desplegament de la infraestructura, principalment una alta complexitat a l'hora d'otorgar permisos i retards en la sincronització i propagació de canvis en l'entorn, hem hagut d'actualitzar la planificació del projecte.

A la taula següent es veuen els canvis respecte a la planificació inicial. S'han actualitzat els períodes de la fase de Desplegament de la infraestructura i de Desenvolupament.

Fase	Tasca	Període
Anàlisi i disseny	Definició dels casos d'ús (phishing, login i malware)	Setmana 3 - 4 febrer
Anàlisi i disseny	Definició dels playbooks	Setmana 2 - 3 de març
Anàlisi i disseny	Disseny de la infraestructura	Setmana 2 - 3 de març
Desplegament de la infraestructura	Desplegament del laboratori	Setmana 2 - 4 de març
Desplegament de la infraestructura	Configuració d'APIs	Setmana 4 de març - 1 d'abril
Desenvolupament	Script phishing	Setmana 2 - 3 d'abril
Desenvolupament	Script login	Setmana 3 - 4 d'abril
Desenvolupament	Script malware	Setmana 4 d'abril - 1 maig
Desenvolupament	Desenvolupament i integració dels components	Setmana 2 - 4 de maig
Desenvolupament	Redacció progressiva de la memòria	Abril - juny
Proves	Execució de proves	Setmana 4 de maig

Fig. 2: Replanificació de les fases del projecte.

5 DISSENY

Aquesta secció descriu el treball central del projecte: com s'ha concebut el sistema d'automatització. En l'apartat *Disseny de la solució* s'explica l'arquitectura plantejada, quins components la formen i, sobretot, com es relacionen entre ells per resoldre una alerta sense intervenció humana. Tot seguit, a la secció *Desplegament de la infraestructura*, es detalla com aquest disseny s'ha materialitzat en un entorn real i funcional sobre el núvol de Microsoft.

5.1 Disseny de la solució

5.1.1 Plantejament general i necessitat a cobrir

Un Centre d'Operacions de Seguretat (SOC) rep cada dia un gran volum d'alertes. Les de Nivell 1 són, en la seva majoria, repetitives i segueixen sempre el mateix guió: quan apareix una alerta coneguda, l'analista fa unes quantes comprovacions, decideix si és una amenaça real i, si ho és, executa unes accions de contenció predefinides (avisar l'usuari, bloquejar un compte o aïllar un equip). Aquesta feina, tot i ser crítica, és mecànica i consumeix una part molt important del temps de l'equip, cosa que endarrereix la resposta i provoca fatiga en els analistes.

La solució dissenyada parteix d'una idea senzilla: si aquestes tasques segueixen sempre el mateix procediment, es poden encarregar a un sistema automàtic. L'objectiu del projecte és, precisament, construir aquesta "central intel·ligent" per a les alertes de Nivell 1.

Tècnicament, el sistema es recolza en dos grans blocs que treballen junts:

- Un **SIEM** (*Security Information and Event Management*), que és la peça encarregada de recollir tota l'activitat de l'entorn, analitzar-la i detectar les situacions sospitoses.
- Un **SOAR** (*Security Orchestration, Automation and Response*), que és la peça encarregada de reaccionar quan el SIEM detecta una amenaça, el SOAR executa de manera automàtica el conjunt de passos de resposta.

Tots dos blocs s'han implementat sobre **Microsoft Sentinel**, una plataforma de núvol que integra les funcions de SIEM i de SOAR en un mateix entorn.

5.1.2 Arquitectura general

L'arquitectura s'ha concebut de manera modular, de forma que cada funció del sistema recau en un component especialitzat i independent. Aquesta separació no és casual: permet que cada peça tingui assignats únicament els permisos imprescindibles per a la seva tasca (principi de mínim privilegi) i facilita que el sistema es pugui ampliar o modificar sense haver de refer-lo sencer.

A alt nivell, la informació circula sempre en la mateixa direcció: l'entorn que es vigila genera activitat → aquesta activitat s'envia al SIEM → el SIEM la analitza i, si detecta una amenaça, genera un *incident* → l'incident desencadena el SOAR → el SOAR executa la resposta i deixa constància de tot el que ha fet (Fig. 3). La comunicació entre les diferents peces es fa íntegrament mitjançant APIs, és a dir, interfícies estandarditzades que permeten que els serveis s'enviïn ordres i informació entre ells de manera segura.

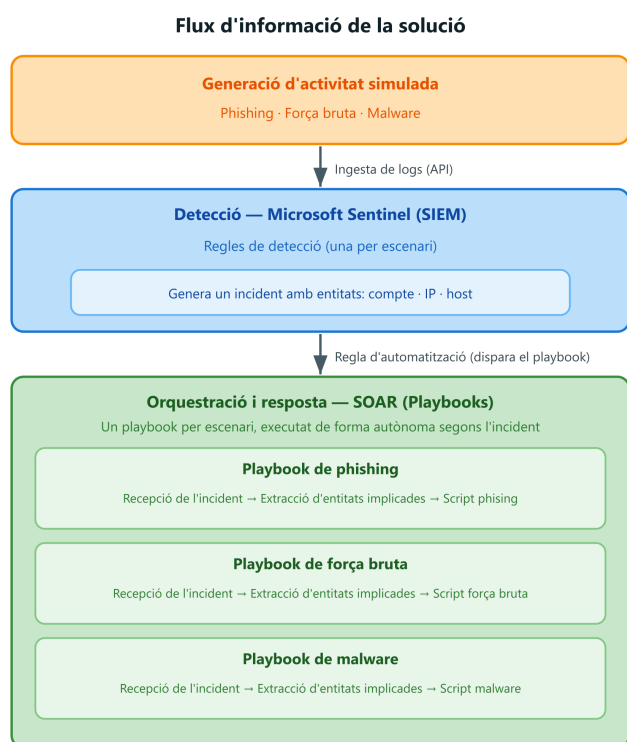


Fig. 3: Flux d'informació de la solució: la informació circula des de la generació d'activitat fins a la resposta, passant pel SIEM (detecció) i el SOAR (orquestració amb un playbook per escenari).

A l'apèndix (Fig. 8) es pot consultar l'arquitectura completa amb més detall, incloent-hi els agents especialitzats que intervenen en cada fase.

5.1.3 Els agents del sistema i les seves responsabilitats

Anomenem **agents** els components que tenen una responsabilitat concreta dins del procés de resposta. Cadascun fa una única funció i col·labora amb la resta a través de les APIs. El sistema s'ha estructurat al voltant dels agents següents:

- **Agent de generació d'activitat simulada.** Reprodueix de manera controlada l'activitat que es vol vigilar (correus de phishing, intents d'autenticació i deteccions de malware) i la injecta a l'entorn. Permet disposar d'atacs reproduïbles amb què validar el sistema tantes vegades com calgui.
- **Agent de detecció (SIEM – Microsoft Sentinel).** És el “cervell vigilant”. Recull de manera contínua els registres d'activitat de l'entorn, els guarda de forma centralitzada i hi aplica un conjunt de *regles de detecció*. Quan una regla es compleix, genera una alerta i, a partir d'ella, un incident que agrupa tota la informació rellevant (qui, què, on i quan), incloses les entitats implicades, com ara el compte d'usuari, l'adreça IP o l'equip afectat.
- **Agent d'orquestració (SOAR – playbooks).** És el “coordinador” de la resposta. Quan es crea un incident, aquest agent s'activa automàticament, interpreta la informació de l'incident i decideix i executa la seqüència d'accions corresponent. S'ha implementat

mitjançant *playbooks*, que són fluxos de treball automatitzats; cada *playbook* representa el procediment de resposta d'un tipus d'amenaça concret.

- **Agent d'identitat (proveïdor d'identitat – Microsoft Entra ID).** És el responsable de les accions sobre els comptes d'usuari. Quan l'orquestrador ho ordena, deshabilita el compte compromès i en revoca les sessions actives, de manera que un atacant que n'hagi obtingut les credencials en perdi l'accés immediatament.
- **Agent de notificació.** S'encarrega d'informar les persones implicades. Avisa l'usuari afectat i/o l'equip de seguretat del que ha passat i de les accions que s'han pres, a través de correu electrònic.
- **Agent d'enriquiment.** Aporta context addicional sobre els indicadors d'una amenaça (per exemple, una adreça IP, un domini o un fitxer sospitos). Consulta serveis externs de reputació i intel·ligència d'amenaçes per determinar si aquests elements ja són coneguts com a maliciosos, cosa que ajuda a confirmar la gravetat de l'incident.

5.1.4 Flux d'interacció entre agents

El valor del sistema rau en la manera en què aquests agents col·laboren de forma encadenada i sense intervenció humana. El recorregut complet d'una alerta, des que es produeix fins que es resol, és el següent (vegeu la Fig. 8 de l'apèndix):

1. **Generació d'activitat.** L'entorn vigilat genera logs de la seva activitat (per exemple, intents d'inici de sessió o correus rebuts).
2. **Detecció.** L'agent de detecció (Sentinel) analitza aquests registres amb les seves regles. Si una regla identifica un patró d'amenaça, crea un incident amb les entitats implicades i la informació necessària per actuar posteriorment.
3. **Activació de la resposta.** La creació de l'incident desencadena automàticament l'agent d'orquestració, mitjançant una regla d'automatització que connecta la detecció amb la resposta.
4. **Anàlisi i enriquiment.** L'orquestrador llegeix les dades de l'incident i, si cal, consulta l'agent d'enriquiment per obtenir més context sobre els indicadors implicats.
5. **Contenció.** Segons el tipus d'amenaça, l'orquestrador ordena a l'agent corresponent que executi l'acció de contenció.
6. **Documentació i notificació.** Finalment, l'orquestrador deixa constància de totes les accions dins del mateix incident i n'informa, a través de l'agent de notificació, els usuaris i analistes pertinents.

D'aquesta manera, una alerta que tradicionalment requeriria diverses comprovacions manuals d'un analista queda resolta de forma autònoma en qüestió de segons.

5.1.5 Escenaris automatitzats

Per demostrar l'eficàcia de la solució, el disseny s'ha centrat en els tres tipus d'amenaça que més col·lapsen el Nivell 1, cadascun amb el seu propi flux de resposta:

- **Phishing.** Davant d'un correu maliciós detectat, el sistema enriqueix els indicadors (remitent, domini, enllaços), notifica l'usuari afectat i documenta l'incident.
- **Anomalies d'autenticació.** Davant d'un patró d'inicis de sessió fallits (indici de credencials compromeses), el sistema deshabilita el compte i en revoca les sessions per tallar l'accés de l'atacant.
- **Malware.** Davant de la detecció de codi maliciós en un equip que no ha estat neutralitzat, el sistema inicia l'aïllament de l'equip per evitar que l'amenaça es propagui.

5.1.6 Principis de disseny

El disseny s'ha guiat per quatre principis fonamentals:

- **Autonomia:** les alertes de Nivell 1 es resolen sense validació humana.
- **Mínim privilegi:** cada agent disposa només dels permisos estrictament necessaris.
- **Modularitat:** cada funció és independent i substitüïble.
- **Traçabilitat:** cada acció automàtica queda registrada a l'incident, de manera que tota la cadena de decisions és auditable a posteriori.

6 DESPLEGAMENT DE LA INFRAESTRUCTURA

Un cop definit el disseny, s'ha procedit a construir l'entorn real on s'executa el sistema. Tota la infraestructura s'ha desplegat sobre el núvol de Microsoft, fet que ha permès muntar un laboratori complet i funcional sense necessitat de maquinari propi.

6.1 Entorn base i SIEM

El nucli del sistema és un espai de treball de registres (*Log Analytics workspace*) sobre el qual s'ha activat **Microsoft Sentinel**. Aquest espai actua com a magatzem centralitzat de tota l'activitat i com a motor d'anàlisi. Sobre ell s'executen les regles de detecció i es generen els incidents que alimenten la resposta automàtica.

6.2 Fonts de dades i simulació d'atacs

En tractar-se d'un laboratori controlat, l'activitat que vigila el sistema s'ha generat de manera simulada però realista. S'ha desenvolupat un mecanisme d'ingesta que injecta a l'espai de treball registres dels tres escenaris (phishing, força bruta i malware), reproduint el comportament que tindrien aquestes dades en un entorn real. L'avantatge clau és que els atacs són reproduïbles i permeten validar el sistema de manera fiable.

6.3 Capa de detecció

Sobre les dades ingerides s'han desplegat les regles de detecció, una per escenari (Fig. 4). Cada regla cerca periòdicament el patró de la seva amenaça i, en trobar-lo, genera un incident. En cada regla s'ha definit el mapatge d'entitats (compte, IP i host), imprescindible perquè la resposta automàtica pugui actuar sobre els elements concrets.

Severity	Name	Rule type	Status
☐ High	Malware	Scheduled	Enabled
☐ High	Brute force	Scheduled	Enabled
☐ Medium	Phishing	Scheduled	Enabled

Fig. 4: Regles de detecció desplegades a Microsoft Sentinel, una per a cada escenari (phishing, força bruta i malware).

Quan l'activitat de l'entorn vigilat arriba al SIEM, no s'emmagatzema com un llistat únic i indiferenciat, sinó que es classifica per tipus en estructures anomenades taules. Cada taula es pot imaginar com un full de càlcul: cada fila és un esdeveniment concret que ha ocorregut (un correu rebut, un intent d'inici de sessió, una alerta de l'antivirus) i cada columna n'és una propietat (qui, què, quan, des d'on, amb quin resultat, etc.).

En aquest projecte conviuen tres taules, una per a cada escenari; **PhishingEmails**, que recull els correus rebuts amb el resultat del filtratge de phishing, **SigninEvents**, que recull els intents d'inici de sessió contra el sistema i **MalwareEvents**, que recull les deteccions de codi maliciós als equips. Cada regla de detecció opera exclusivament sobre la taula del seu àmbit i hi cerca el patró que la caracteritza.

A continuació es descriu cadascuna de les tres regles desplegades, indicant tant el patró que cerca dins de les dades com l'incident que genera quan dispara:

- **Regla de phishing:** La primera regla vigila la taula d'esdeveniments de correu (PhishingEmails) i identifica els missatges que, havent estat classificats com a phishing per les eines de filtratge, han arribat efectivament a la safata d'entrada de l'usuari. La regla selecciona, per tant, els correus que compleixen alhora dues condicions: que tinguin el tipus d'amenaça Phish i que l'acció d'entrega resultant sigui Delivered. Es tracta dels casos crítics, en què el correu maliciós no ha estat bloquejat i, per tant, requereix una resposta del SOC. Quan dispara, la regla genera un incident de severitat mitjana i hi associa tres entitats que descriuen el cas: el compte del destinatari afectat, la IP del remitent del correu i la URL o contingut maliciós del missatge.

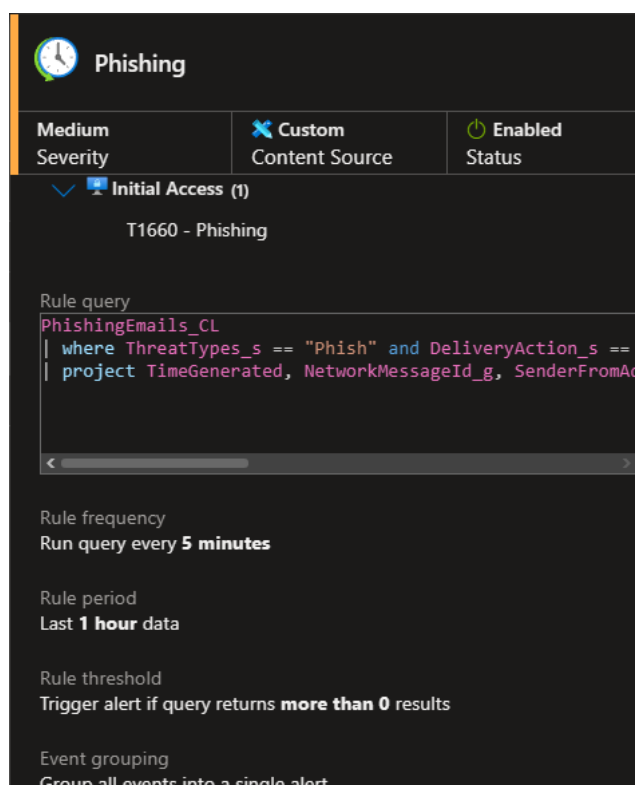


Fig. 5: Configuració de la regla de detecció de phishing.

- **Regla de força bruta:** La segona regla és la més elaborada de les tres, ja que no busca un esdeveniment aïllat sinó un patró de comportament sobre la taula d'inicis de sessió (SignInEvents). Per a cada combinació d'usuari i adreça IP d'origen, la regla agrega els intents d'inici de sessió i compta per separat quants han fallat per credencials incorrectes i quants han tingut èxit. La regla considera que s'ha produït un possible compromís de credencials quan, des d'una mateixa IP i contra un mateix usuari, es donen com a mínim deu intents fallits seguits. L'incident generat té severitat alta. Les dues entitats que s'hi associen són el compte d'usuari presumptament compromès i la IP d'origen des de la qual s'ha dut a terme l'atac. Aquesta informació és la que permetrà al playbook corresponent deshabilitar el compte i revocar-ne les sessions actives de manera precisa. La configuració d'aquesta regla es pot consultar a la Fig. 9 de l'apèndix.
- **Regla de malware:** La tercera regla vigila la taula d'esdeveniments de codi maliciós (MalwareEvents) i identifica les deteccions que no han estat neutralitzades per l'antivirus del propi equip. Concretament, selecciona aquelles entrades en què l'acció de remediació no figura com a cap dels estats segurs (Quarantined, Remediated o Blocked): són els casos en què la detecció s'ha produït però el fitxer maliciós continua actiu al dispositiu, de manera que la màquina segueix compromesa. L'incident corresponent té severitat alta. Incorpora tres entitats: el host on s'ha detectat el codi maliciós, el compte d'usuari associat a la sessió i el hash SHA-256 del fitxer detectat. Aquestes entitats permeten al playbook iniciar l'aïllament de l'equip a la xarxa i compartir el hash amb la resta del sistema com a indicador conegut. La configuració d'aquesta regla

es pot consultar a la Fig. 10 de l'apèndix.

Les tres regles s'executen totes amb la mateixa cadència (una passada cada cinc minuts sobre les dades de l'última hora) i s'han configurat perquè agrupin tots els resultats coincidents en un únic incident per disparada, per evitar generar soroll innecessari. En conjunt, aquestes tres regles constitueixen el pont entre les dades en brut que arriben al SIEM i les accions de resposta que executa el SOAR: cada incident que generen porta incorporada tota la informació estructurada necessària perquè la resposta automàtica pugui actuar de manera precisa i sense intervenció humana.

6.4 Capa d'orquestració i resposta (SOAR)

La resposta s'ha implementat amb **playbooks**, desplegats com a fluxos de treball dins de la mateixa plataforma. Cada playbook recull l'incident, n'extreu les entitats i executa la seqüència d'accions de resposta del seu escenari. La connexió entre la detecció i la resposta es fa mitjançant **regles d'automatització**. Una regla d'automatització és, essencialment, un distribuïdor (dispatcher) que vigila contínuament la cua d'incidents que va generant el SIEM: cada vegada que n'apareix un de nou, en comprova un conjunt de condicions i, si aquestes condicions coincideixen amb les que té configurades, dispara automàticament l'acció corresponent (en el nostre cas, executar un playbook concret). Així, mentre el playbook és el responsable de què fer davant d'una amenaça, la regla d'automatització és la responsable de decidir a quin playbook li toca actuar.

Aquesta separació entre la detecció (quan saltar) i la resposta (què fer), és la que dona flexibilitat al sistema: afegir un nou escenari en el futur es redueix a crear una nova regla de detecció, un nou playbook i una nova regla d'automatització que els lligui, sense haver de modificar res del que ja existeix.

Cadascun dels tres playbooks aplica aquesta mateixa estructura amb les particularitats pròpies del seu escenari: el de força bruta executa la contenció real sobre el directori d'identitats (deshabilita el compte compromès i en revoca les sessions actives), el de phishing es focalitza en l'enriquiment dels indicadors del correu i en la notificació immediata a l'usuari afectat, i el de malware sol·licita l'aïllament de l'equip a la xarxa per evitar la propagació del codi maliciós. El comportament concret d'aquests tres playbooks, les seves execucions sobre incidents reals i les sortides que se'n deriven es presenten i s'analitzen en detall a la secció de Desenvolupament. L'arquitectura d'aquesta capa de resposta es pot veure a la Fig. 6 de la pàgina següent.

6.5 Entorn d'identitat per a la contenció

Per poder executar accions reals sobre comptes d'usuaris, s'ha utilitzat un **directori d'identitats dedicat** amb usuaris de prova. Aquest directori representa la base d'usuaris de l'organització i és l'entorn sobre el qual el sistema executa les accions de contenció d'identitat (deshabilitar comptes i revocar sessions). Treballar sobre un directori propi i aïllat permet validar aquestes accions de forma segura, sense cap risc per a usuaris reals.

Arquitectura de la capa de detecció i resposta SOAR

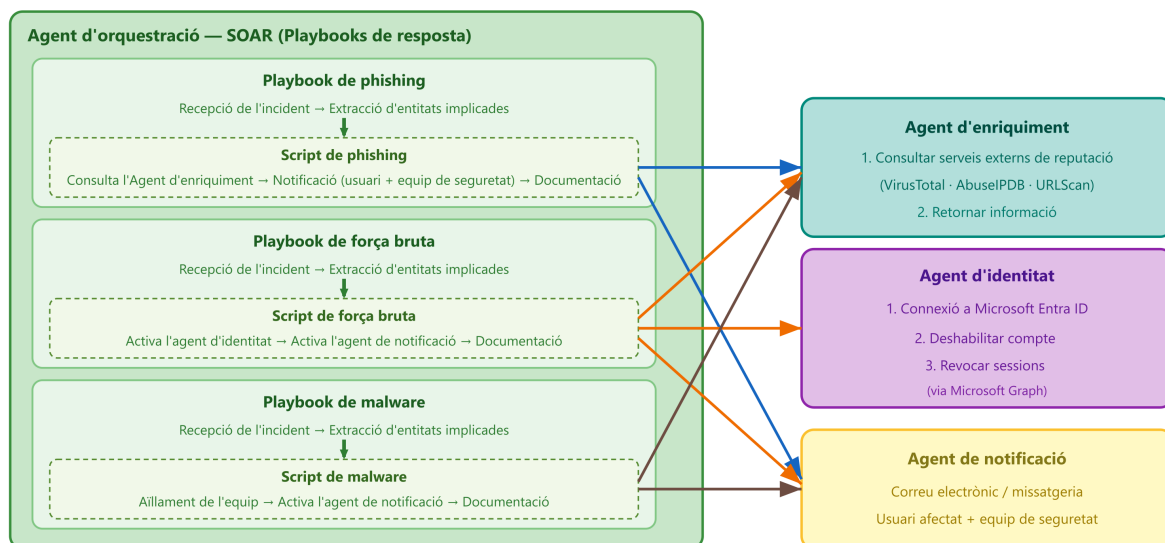


Fig. 6: Arquitectura de la capa d'orquestració i resposta (SOAR): cada playbook rep l'incident, n'extreu les entitats i executa el seu script de resposta, que activa els agents especialitzats corresponents (enriquiment, identitat o notificació).

6.6 Integració, autenticació i permisos

L'aspecte més delicat del desplegament ha estat la integració segura entre tots els components. Les APIs de Microsoft són estrictes pel que fa a l'autenticació i als permisos, de manera que cada comunicació entre agents ha requerit establir una identitat i unes credencials pròpies, autoritzades específicament per a l'acció concreta que ha de dur a terme. Aquest treball s'ha aprofitat per aplicar de manera rigorosa el principi de mínim privilegi. Aquesta configuració acurada de permisos i connectors és la que garanteix que la infraestructura resultant sigui robusta i segura. Com a resultat, el sistema queda desplegat de punta a punta i és capaç d'executar el cicle complet de detecció i resposta de forma autònoma i traçable.

7 DESENVOLUPAMENT

El principal resultat assolit és un sistema operatiu de punta a punta en l'entorn de laboratori. En llançar cadascun dels atacs simulats, s'ha verificat que es desencadena de manera autònoma la cadena esperada: l'activitat genera registres (Fig. 11 de l'apèndix), l'agent de detecció crea l'incident (Fig. 12 de l'apèndix) amb el mapeig d'entitats corresponent (Fig. 4) i l'agent d'orquestració executa el playbook adequat fins a la contenció, la notificació i la documentació.

Concretament, s'han desenvolupat tres playbooks de resposta, un per a cada escenari. A continuació es detalla el comportament de cadascun.

Playbook de phishing. Activat per l'incident Phishing, recupera de l'incident les entitats associades (remitent, IP del remitent i URL maliciosa), crida al **Script de phishing**. Aquest, consulta l'Agent d'enriquiment, que consulta la reputació contra serveis externs de threat intelligence: VirusTotal, AbuseIPDB i URLScan. El veredictat recopilat (si la URL figura com a maliciosa coneguda, si la IP apareix en llistes de reputació) s'incorpora com a comentari a l'inci-

dent, dotant l'analista d'un context immediat. A continuació, notifica per correu electrònic l'usuari afectat advertint-lo del missatge sospitos i, en paral·lel, n'informa l'equip de seguretat. Finalment elimina el correu de la bústia, marca l'incident amb el resultat obtingut i el tanca.

Playbook de força bruta. Activat per l'incident Força bruta, n'extreu el compte compromès i la IP atacant, i activa el **Script de força bruta**, que a través de l'Agent d'identitat executa la contenció contra el directori d'identitat dedicat: mitjançant Microsoft Graph, deshabilita el compte (PATCH /users/id amb accountEnabled=false) i tot seguit en revoca totes les sessions actives (POST /users/id/revokeSignInSessions), invalidant qualsevol token vàlid que l'atacant pogués haver obtingut. Posteriorment enriqueix l'incident amb la reputació de la IP atacant consultada a AbuseIPDB a través de l'Agent d'enriquiment, afegeix un comentari amb el detall de les accions executades i n'informa per correu l'equip de seguretat.

Playbook de malware. Activat per l'incident Malware, llegeix de l'incident l'equip afectat, l'usuari propietari de la sessió i el hash SHA-256 del fitxer detectat. Tot seguit, activa el **Script de malware** que consulta el hash a VirusTotal utilitzant l'agent d'enriquiment per validar i contextualitzar l'amenaça i, a continuació, procedeix amb l'aïllament de l'equip a la xarxa per evitar la propagació del codi maliciós. En el present laboratori, aquesta darrera acció s'ha modelat de manera simulada perquè la seva execució real requeriria una eina EDR (Endpoint Detection and Response), és a dir, una plataforma de detecció i resposta davant d'incidents als equips finals, amb endpoints connectats. Tot i això, el flux està dissenyat de manera que en un entorn real n'hi hauria prou amb substituir aquest pas pel connector real sense modificar la resta de la lògica. Finalment, el playbook documenta totes les accions a l'incident i notifica l'equip de seguretat.

Aquests resultats són reproduïbles: en repetir els atacs si-

mulats, el sistema hi respon de manera consistent. Resten pendents, per a la fase de validació, les mètriques quantitatives —principalment la reducció del temps de resposta (MTTR) i el percentatge d'alertes resoltes sense intervenció humana.

Aquests resultats s'alineen amb la direcció actual del sector, que avança cap a operacions de seguretat orquestrades i, cada cop més, assistides per agents autònoms [2][4], i amb les bones pràctiques d'enginyeria de deteccions documentades a la literatura [3]. De fet, les plataformes comercials de SOAR ja integren capacitats equivalents a les implementades aquí [1]; la diferència és que en aquest treball s'han construït i comprès de zero en un entorn controlat, cosa que aporta una comprensió més profunda del seu funcionament intern.

8 PROVES

L'objectiu d'aquesta fase és validar que el sistema respon de manera correcta, autònoma i reproduïble davant de cadascun dels escenaris automatitzats, i comprovar que les accions executades es corresponen amb les definides a la matriu de responsabilitats. Les proves s'han dut a terme sobre el mateix laboratori controlat descrit a la secció de desplegament, aprofitant que els atacs simulats són reproduïbles i permeten repetir cada cas tantes vegades com calgui.

8.1 Proves d'activació

S'ha seguit una estratègia de proves funcionals de punta a punta: per a cada escenari s'injecta l'activitat maliciosa simulada i s'observa tota la cadena de resposta, des de la generació del registre fins al tancament de l'incident. Es considera que una prova és satisfactòria quan es compleixen tots els criteris següents:

- La regla corresponent dispara i genera l'incident amb la severitat prevista.
- L'incident incorpora correctament les entitats necessàries per actuar (compte, IP, host o hash, segons l'escenari).
- El playbook executa l'acció de resposta definida per a aquell escenari.
- L'incident queda documentat amb el detall de les accions i s'emet la notificació pertinent.

S'ha verificat que per cada cas s'executa la cadena esperada de forma autònoma i completa.

8.2 Proves de no-activació (falsos positius)

Per comprovar que el sistema no actua quan no ha de fer-ho, s'han executat casos de control negatius: un correu de phishing *bloquejat* abans d'arribar a l'usuari, una seqüència d'inicis de sessió fallits per sota del llindar de deu intents, i una detecció de malware ja *posada en quarantena*. En cap d'aquests casos s'ha generat incident ni s'ha disparat cap playbook, fet que valida la precisió de les regles de detecció.

Cada cas s'ha repetit diverses vegades i el sistema hi ha respost de manera consistent, cosa que confirma la reproduïbilitat de la solució.

9 CONCLUSIONS

L'aportació principal d'aquest treball ha estat demostrar, sobre un entorn real, que és viable automatitzar de punta a punta la resposta a les alertes més habituals d'un SOC de Nivell 1. Concretament, s'ha aportat: (i) un disseny d'arquitectura modular basat en agents especialitzats amb una matriu de responsabilitats clara, que serveix com a model reutilitzable per a futurs escenaris; (ii) un laboratori funcional sobre Microsoft Sentinel que integra SIEM i SOAR i que permet generar atacs reproduïbles; i (iii) la implementació de fluxos de resposta autònoms per als tres escenaris prioritzats —phishing, força bruta i malware—, incloent-hi accions de contenció reals sobre un directori d'identitat dedicat. Tot plegat confirma que la transformació de procediments manuals i repetitius en processos autònoms és tècnicament assolible amb les eines actuals.

Pel que fa als objectius plantejats, s'han assolit els relatius al disseny del model, el desplegament de la infraestructura, el desenvolupament de les regles de detecció i la implementació dels fluxos d'orquestració. En canvi, l'objectiu d'analitzar la millora operativa encara no s'ha completat: la fase de validació mitjançant simulació controlada i la recollida de mètriques quedaven, segons la replanificació, per a l'etapa final del projecte. Aquest endarreriment és conseqüència directa de la complexitat de la integració segura entre components —especialment la gestió de permisos i la propagació de canvis a l'entorn—, que ha consumit més temps del previst.

A partir d'aquí, s'obren diverses línies de continuació. La més immediata és tancar la validació quantitativa, mesurant la reducció del temps de resposta (MTTR) i de la càrrega manual dels analistes, per demostrar numèricament l'eficàcia de la solució. Una segona línia, orientada a acostar el sistema a un entorn de producció, és refinar les regles de detecció perquè incorporin més context i no es limitin a patrons simples. Per exemple, un inici de sessió fallit procedent d'una IP de confiança (inclosa a la llista blanca de l'organització) no s'hauria de comptabilitzar com a intent maliciós; de la mateixa manera, caldria tenir en compte la ubicació geogràfica, els horaris habituals de l'usuari o la correlació entre diversos senyals abans de generar un incident. Aquestes millores reduirien els falsos positius i augmentarien la fiabilitat de les automatitzacions. Finalment, el caràcter modular de l'arquitectura facilita ampliar el sistema amb nous escenaris d'amenaça i nous agents de contenció sense haver de redissenyar-lo.

ÚS DE LA IA GENERATIVA

S'ha utilitzat IA generativa com a suport per la redacció i el format del document, la resolució d'alguns problemes de configuració de l'entorn de desenvolupament relacionats amb tecnologies de Microsoft i, de manera molt puntual, en la proposta o revisió de petits fragments de codi. El disseny de l'arquitectura del sistema, el desenvolupament de la solució, la implementació de la lògica de funcionament i les decisions tècniques adoptades, són responsabilitat exclusiva de l'estudiant.

AGRAÏMENTS

Vull expressar el meu agraïment a totes les persones que han fet que estigui on sóc a dia d'avui i han fet possibles la realització d'aquest Treball de Final de Grau.

Al meu tutor de TFG, Carles Garrigues Olivella, per acceptar el repte proposat i per tot el seguiment i orientació rebuts durant el desenvolupament d'aquest treball.

Finalment, a la meva família, per tot el suport incondicional durant la carrera i la realització d'aquest projecte.

REFERÈNCIES

- [1] BlinkOps. (s.f.). Partners & Integrations. (2026). Última visita el 23/02/2026
- [2] MJ. (2024, 23 d'octubre). Building an Agentic SOC on Elastic. Medium. Última visita el 04/03/2026
- [3] Bailey, K. (2023). Automating Security Detection Engineering. O'Reilly Media. [Llibre]
- [4] Google Cloud. (2024). The future of security operations is powered by AI agents [Infografia].
- [5] Microsoft, "What is Microsoft Sentinel?," Microsoft Learn, 2024. [En línia]. Disponible a: <https://learn.microsoft.com/azure/sentinel/>. Última visita: 06/2026.
- [6] Gartner, "Market Guide for Security Orchestration, Automation and Response Solutions," Gartner Research, 2022.
- [7] P. Cichonski, T. Millar, T. Grance i K. Scarfone, "Computer Security Incident Handling Guide," NIST Special Publication 800-61 Rev. 2, 2012.
- [8] The MITRE Corporation, "ATT&CK Matrix for Enterprise," 2024. [En línia]. Disponible a: <https://attack.mitre.org/>. Última visita: 06/2026.
- [9] Verizon Business, "2024 Data Breach Investigations Report," 2024. [Informe tècnic].
- [10] Microsoft, "user: revokeSignInSessions — Microsoft Graph REST API," Microsoft Learn, 2024. [En línia]. Última visita: 06/2026.
- [11] Ponemon Institute, "The Economics of Security Operations Centers: What Is the True Cost for Effective Results?," 2023. [Informe].
- [12] Microsoft, "Kusto Query Language (KQL) overview," Microsoft Learn, 2024. [En línia]. Última visita: 06/2026.
- [13] VirusTotal (Google), AbuseIPDB i urlscan.io, "Serveis públics d'anàlisi de reputació i intel·ligència d'amenaçes." [En línia]. Última visita: 06/2026.

APÈNDIX

A.1 Matriu de responsabilitats

Matriu de responsabilitats: acció que executa cada agent en cada escenari			
Agent	Phishing	Força bruta	Malware
Detecció (SIEM)	Detecta el correu de phishing i crea l'incident (severitat mitjana) amb compte, IP i URL	Detecta ≥10 intents fallits des d'una mateixa IP contra un mateix usuari i crea l'incident (severitat alta) amb compte i IP	Detecta una remediació no segura i crea l'incident (severitat alta) amb host, compte i hash
Orquestració (SOAR)	Rep l'incident, n'extreu les entitats, executa el playbook corresponent i documenta totes les accions a l'incident		
Enriquiment	Reputació de remitent, IP i URL (VirusTotal, AbuseIPDB, URLScan)	Reputació de la IP atacant (AbuseIPDB)	Reputació del hash SHA-256 (VirusTotal)
Identitat (Entra ID)	—	Deshabilita el compte i revoca les sessions actives (Microsoft Graph)	—
Contenció (aïllament d'equip)	—	—	Aïlla l'equip de la xarxa (simulat)
Notificació	Avisa l'usuari afectat i l'equip de seguretat	Notifica l'equip de seguretat	Notifica l'equip de seguretat

Fig. 7: Matriu de responsabilitats: acció que executa cada agent en cada escenari (phishing, força bruta i malware).

A.2 Arquitectura del sistema d'automatització de respostes (SOC Nivell 1)

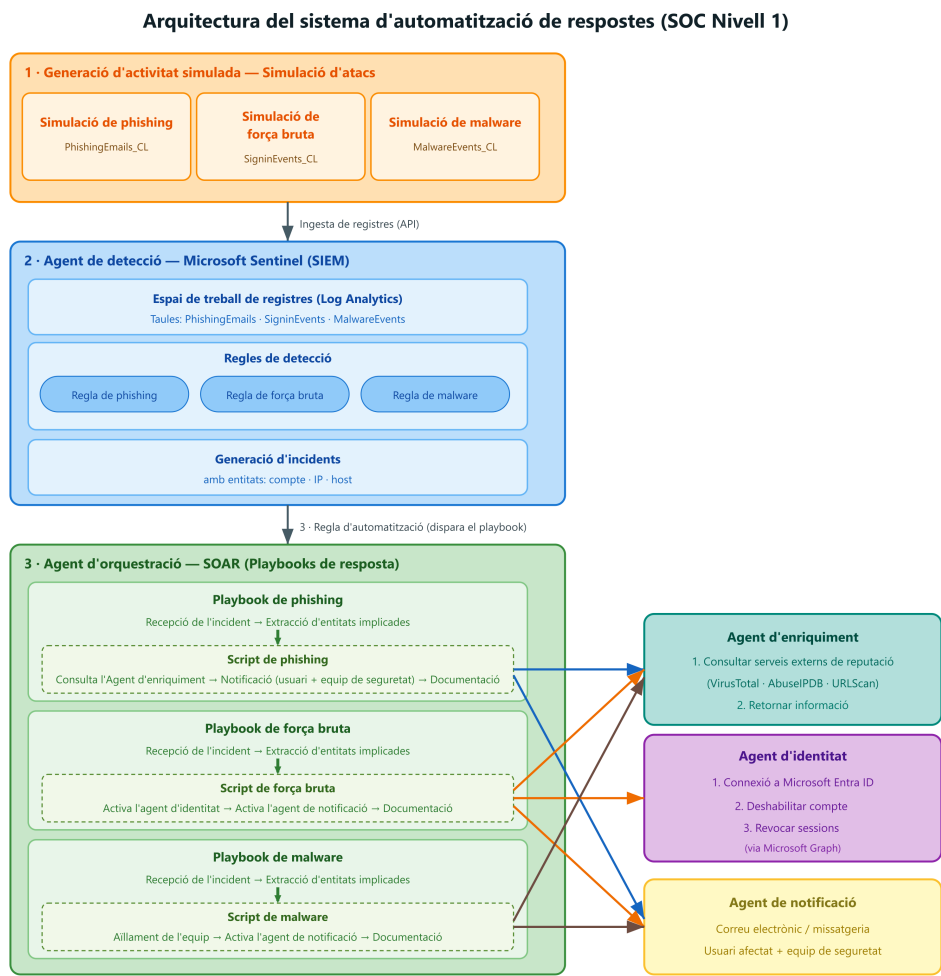
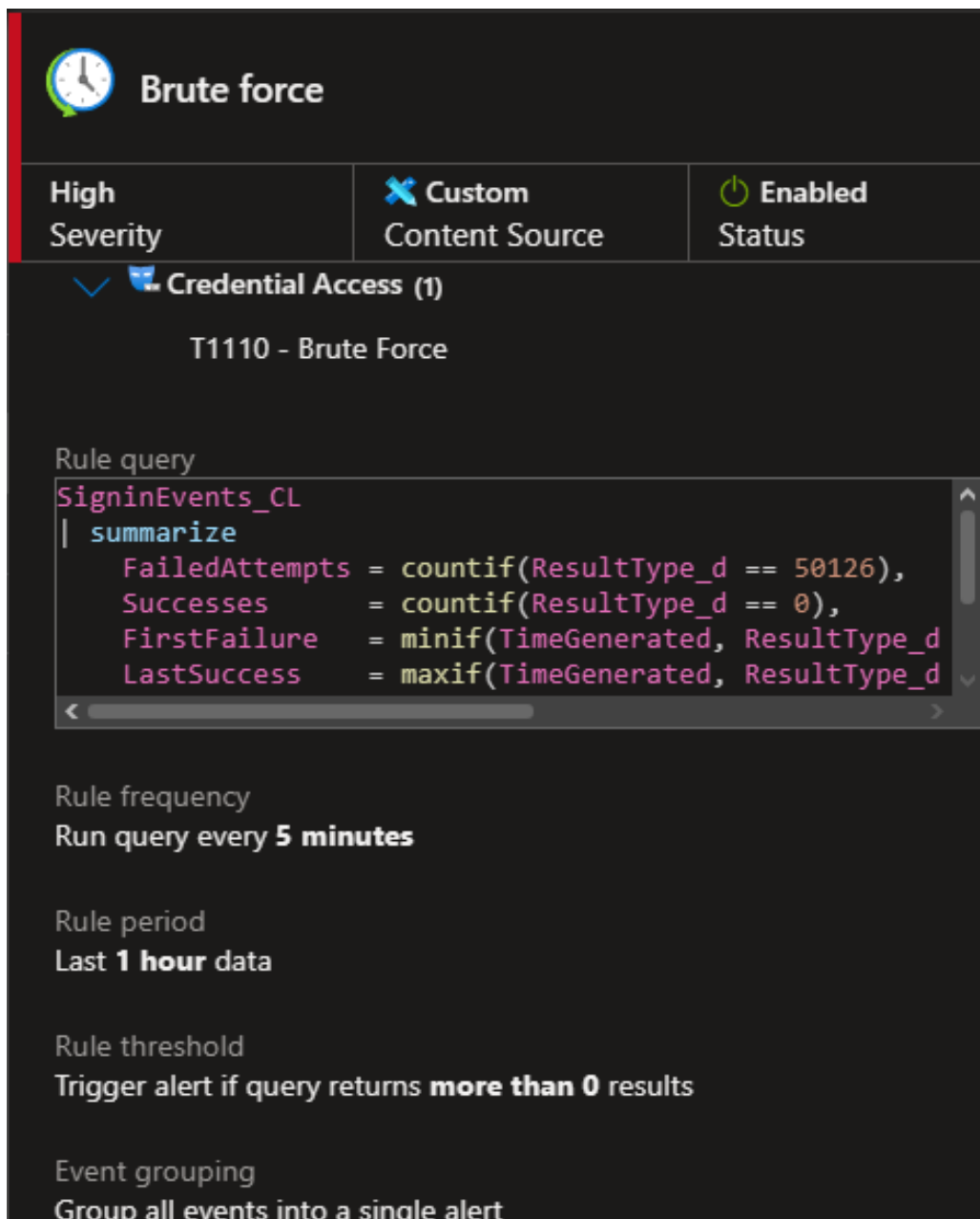


Fig. 8: Arquitectura general del sistema d'automatització de respostes per a un SOC de Nivell 1, amb els agents especialitzats i el flux d'informació entre la generació d'activitat, la detecció (SIEM) i l'orquestració de la resposta (SOAR).

A.3 Configuració de la regla de detecció de força bruta



The screenshot displays the configuration interface for a "Brute force" rule in Microsoft Sentinel. The rule is categorized under "Credential Access (1)" and is identified as "T1110 - Brute Force". The configuration includes a "Rule query" section with a KQL query, a "Rule frequency" of "Run query every 5 minutes", a "Rule period" of "Last 1 hour data", a "Rule threshold" of "Trigger alert if query returns more than 0 results", and an "Event grouping" setting of "Group all events into a single alert".

Brute force

High Severity	Custom Content Source	Enabled Status
----------------------	------------------------------	-----------------------

Credential Access (1)

T1110 - Brute Force

Rule query

```
SigninEvents_CL
| summarize
    FailedAttempts = countif(ResultType_d == 50126),
    Successes      = countif(ResultType_d == 0),
    FirstFailure   = minif(TimeGenerated, ResultType_d),
    LastSuccess    = maxif(TimeGenerated, ResultType_d)
```

Rule frequency
Run query every **5 minutes**

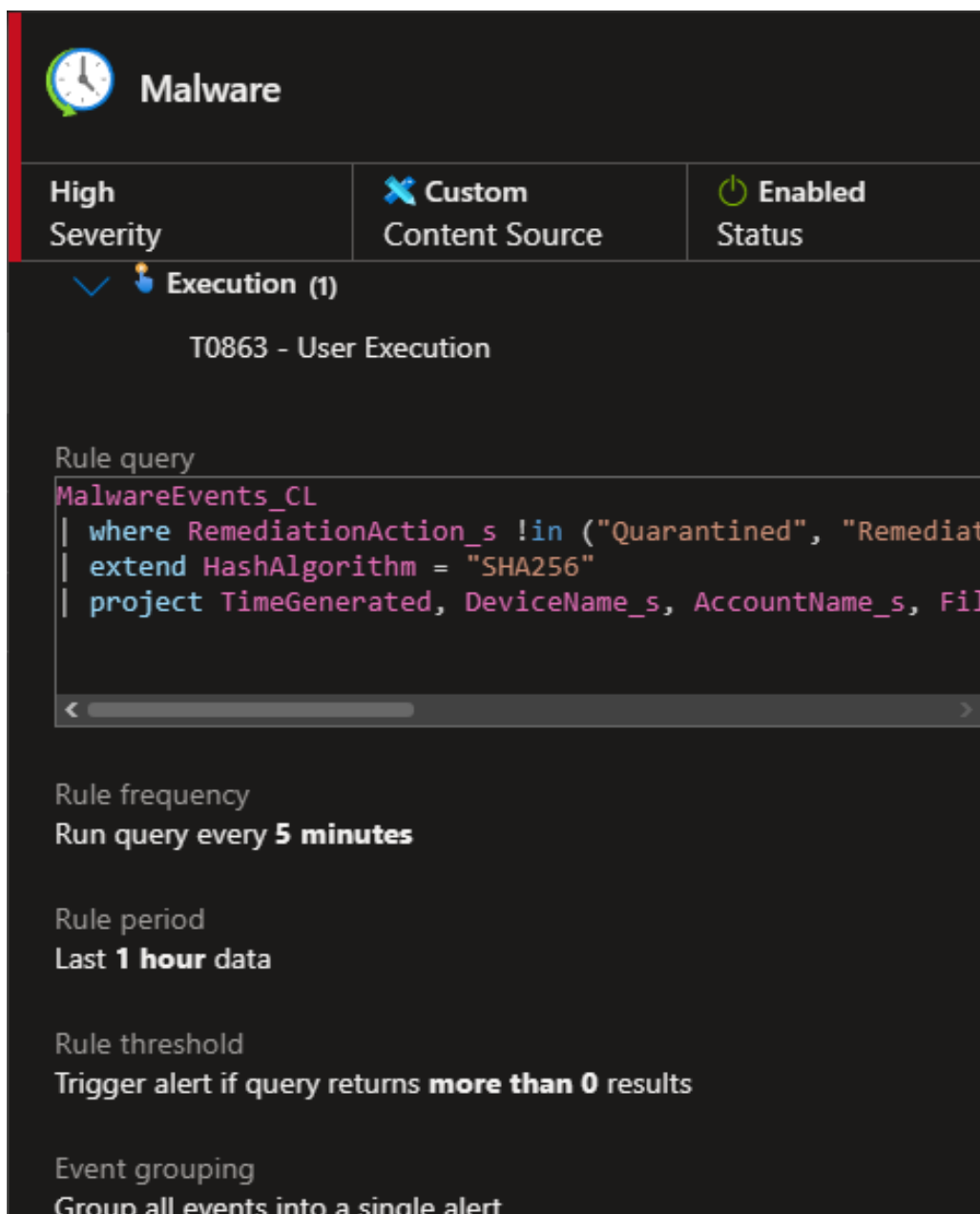
Rule period
Last **1 hour** data

Rule threshold
Trigger alert if query returns **more than 0** results

Event grouping
Group all events into a single alert

Fig. 9: Configuració de la regla de detecció d'atacs de força bruta a Microsoft Sentinel.

A.4 Configuració de la regla de detecció de malware



The screenshot displays the configuration page for a 'Malware' rule in Microsoft Sentinel. The interface is dark-themed. At the top, there's a header with a clock icon and the word 'Malware'. Below this is a table with three columns: 'Severity' (set to 'High'), 'Content Source' (set to 'Custom'), and 'Status' (set to 'Enabled'). Under the 'Execution (1)' section, the rule is identified as 'T0863 - User Execution'. The 'Rule query' section contains a Kusto query:

```
MalwareEvents_CL  
| where RemediationAction_s !in ("Quarantined", "Remediat  
| extend HashAlgorithm = "SHA256"  
| project TimeGenerated, DeviceName_s, AccountName_s, Fil
```

. Below the query, the 'Rule frequency' is set to 'Run query every 5 minutes'. The 'Rule period' is set to 'Last 1 hour data'. The 'Rule threshold' is set to 'Trigger alert if query returns more than 0 results'. Finally, the 'Event grouping' is set to 'Group all events into a single alert'.

Malware

High Severity	 Custom Content Source	 Enabled Status
-------------------------	--	---

Execution (1)

T0863 - User Execution

Rule query

```
MalwareEvents_CL  
| where RemediationAction_s !in ("Quarantined", "Remediat  
| extend HashAlgorithm = "SHA256"  
| project TimeGenerated, DeviceName_s, AccountName_s, Fil
```

Rule frequency
Run query every **5 minutes**

Rule period
Last **1 hour** data

Rule threshold
Trigger alert if query returns **more than 0** results

Event grouping
Group all events into a single alert

Fig. 10: Configuració de la regla de detecció de malware a Microsoft Sentinel.

A.5 Registres d'activitat generats

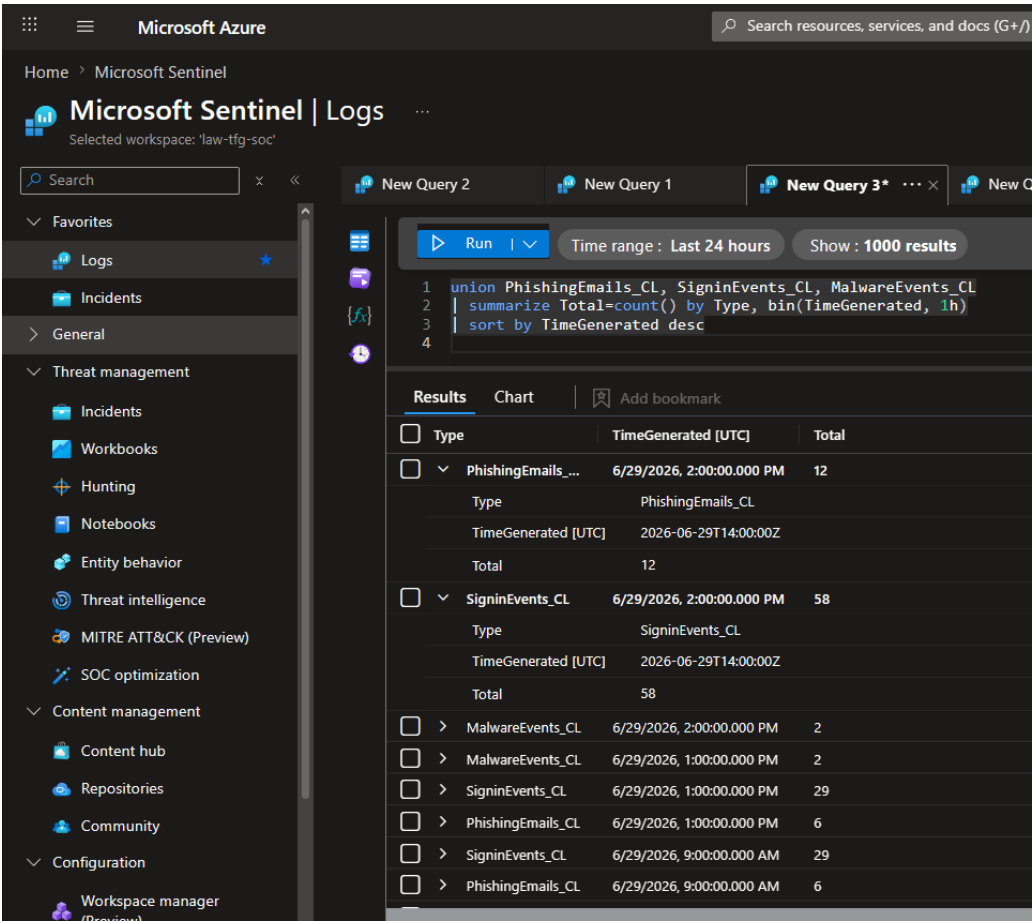


Fig. 11: Registres d'activitat generats pels atacs simulats i ingerits a l'espai de treball de Microsoft Sentinel.

A.6 Incidents generats

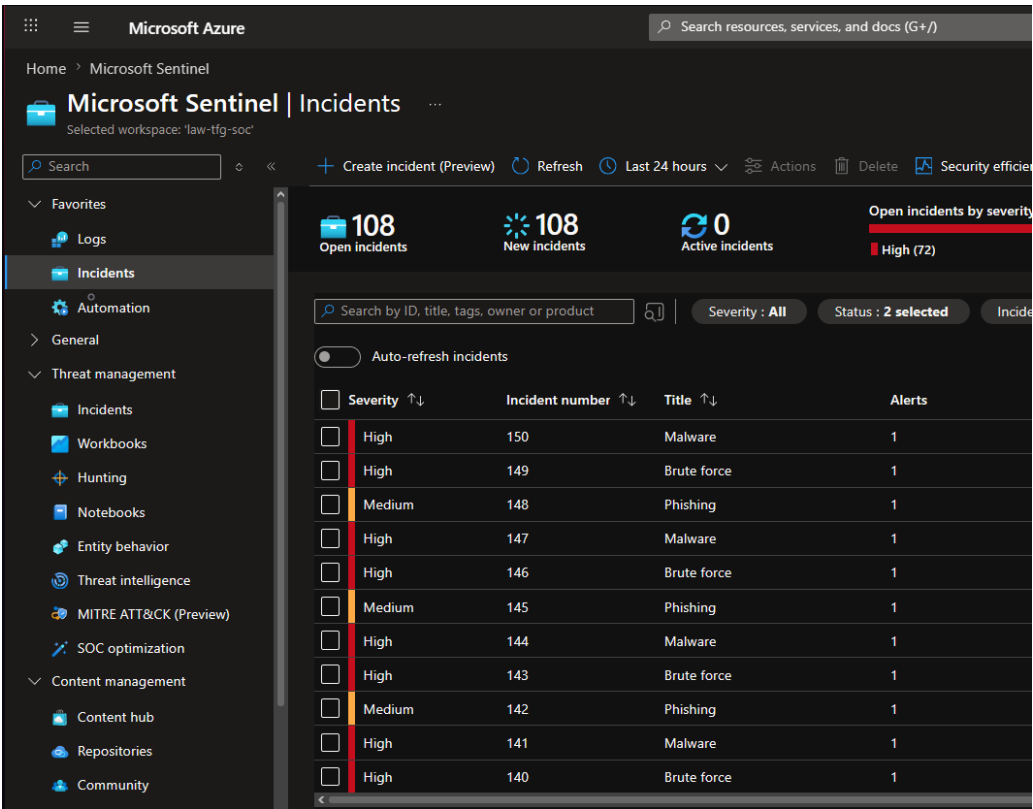


Fig. 12: Incidents creats per l’agent de detecció a Microsoft Sentinel a partir dels registres d’activitat.