AYI NEDJIMI Consultants

Threat Hunting : Méthodologie, Outils et Pratique pour

Catégorie : SOC et Detection Lecture : 8 min Publié le : 08/03/2026 Auteur : Ayi NEDJIMI

Guide complet threat hunting : modèle PEAK, hypothèses de chasse, techniques d'analyse (stack counting, long tail, frequency analysis), outils.

2.1 Le spectre de la détection

Pour comprendre la place du threat hunting, il faut le situer sur le **spectre de la détection** qui va du purement réactif au pleinement proactif :

Approche	Déclencheur	Automatisation	Exemple
Détection par signature	Pattern connu	100% automatique	Règle Snort, antivirus
Détection par règles SIEM	Corrélation d'événements	95% automatique	5 échecs auth en 2 min
Détection par anomalie (ML)	Déviation du baseline	90% automatique	UEBA, traffic anomaly
Threat hunting guidé	IOC / renseignement	50% manuel	Recherche d'un hash IOC
Threat hunting proactif	Hypothèse analyste	80% manuel	"Un adversaire utilise le DNS tunneling pour exfiltrer"

Le threat hunting **guidé par les IOC** (aussi appelé "IOC sweeping") est le niveau d'entrée : on recherche des indicateurs spécifiques (hashes, IPs, domaines) issus du renseignement sur les menaces. C'est utile mais limité, car les IOC ont une durée de vie courte -- un attaquant change d'infrastructure en quelques heures.

Le threat hunting **proactif basé sur les hypothèses** est la forme la plus mature. L'analyste formule une hypothèse sur le comportement d'un adversaire (basée sur les TTPs MITRE ATT&CK, le renseignement sectoriel ou l'intuition opérationnelle), puis conçoit une requête ou une analyse pour tester cette hypothèse contre les données télémétriques de l'organisation. Cette approche cible les **comportements** plutôt que les indicateurs, ce qui la rend résistante au changement d'infrastructure de l'adversaire.

2.2 Prérequis pour le threat hunting

Le threat hunting ne peut fonctionner sans une base solide. Les prérequis sont :

- **Visibilité (télémétrie)** : vous ne pouvez chasser que ce que vous pouvez voir. Les données essentielles incluent : logs d'authentification, événements de processus (Sysmon), logs DNS, flux réseau (NetFlow/Zeek), logs proxy/firewall, événements FIM et logs cloud. Un déploiement Wazuh avec agents sur les endpoints fournit une base solide.
- **Rétention** : le dwell time moyen étant de 10+ jours, il faut au minimum 30 à 90 jours de rétention sur les données détaillées. Les hunts sur les APT peuvent nécessiter des recherches sur 6 à 12 mois.
- **Capacité d'interrogation** : un SIEM ou un data lake (Elasticsearch, Splunk, Azure Data Explorer) avec des performances de requête suffisantes pour des recherches ad-hoc sur de gros volumes.
- **Connaissance des menaces** : compréhension des TTPs adverses, du framework MITRE ATT&CK, des rapports de threat intelligence sectoriels et des techniques offensives documentées.
- **Compétences analytiques** : capacité à formuler des hypothèses, à construire des requêtes complexes, à interpréter des données statistiques et à identifier les anomalies significatives dans le bruit.

Notre avis d'expert
La fatigue d'alerte est l'ennemi silencieux des SOC modernes. Quand les analystes traitent des centaines de faux positifs par jour, les vraies menaces passent inaperçues. La priorisation intelligente et l'automatisation des tâches de triage sont essentielles.

Exemples d'hypothèses bien formulées :

- *"Un adversaire utilise le protocole DNS pour exfiltrer des données en encodant le payload dans les sous-domaines de requêtes vers un domaine contrôlé."* -- Technique : DNS Tunneling (T1071.004).
- *"Un implant C2 communique avec son infrastructure via des requêtes HTTPS périodiques avec un intervalle régulier (beaconing), détectable par l'analyse de fréquence des connexions sortantes."* -- Technique : Application Layer Protocol (T1071.001).
- *"Un attaquant ayant compromis un poste utilisateur utilise des outils natifs Windows (PsExec, WMI, WinRM) pour se déplacer latéralement vers d'autres systèmes du réseau."* -- Technique : Lateral Movement via Living off the Land (T1021).
- *"Des credentials volés par un infostealer sont utilisés pour accéder à nos applications cloud depuis des géolocalisations inhabituelles."* -- Technique : Valid Accounts (T1078), lien avec les infostealers.

Pour chaque hypothèse, documentez : les **sources de données nécessaires** (logs DNS, NetFlow, événements Sysmon, etc.), les **requêtes ou analyses prévues**, le **scope** (quels segments du réseau, quelle période temporelle), et les **critères de succès/échec** (quand considère-t-on que l'hypothèse est validée ou invalidée).

3.3 Phase Execute : mener la chasse

La phase Execute est le coeur opérationnel du hunt. L'analyste exécute les requêtes planifiées, examine les résultats, affine les recherches et investigue les anomalies. Cette phase est itérative : chaque résultat peut mener à une nouvelle requête plus ciblée, dans un processus de "pivoting" analytique similaire au pivoting technique en pentest.

La durée d'un hunt varie de quelques heures (IOC sweep simple) à plusieurs jours (investigation comportementale complexe). Il est recommandé de timeboxer les hunts : si aucun résultat significatif n'émerge après le temps alloué, documentez les résultats négatifs (qui ont aussi de la valeur) et passez à l'hypothèse suivante.

3.4 Phase Act : transformer les découvertes en action

Lorsqu'un hunt identifie une menace réelle ou potentielle, la phase Act déclenche les actions nécessaires :

- **Réponse immédiate** : si une compromission active est détectée, déclencher le processus d'incident response (containment, eradication, recovery). La chaîne de preuve numérique doit être préservée.
- **Nouvelles règles de détection** : transformer le pattern identifié en règle SIEM/XDR automatisée (Sigma, Wazuh, Splunk SPL, KQL). Le hunt ponctuel devient une détection permanente.
- **Amélioration de la visibilité** : si le hunt a révélé des lacunes de télémétrie (logs manquants, sources non collectées), initier les actions pour combler ces lacunes.
- **Mise à jour du modèle de menace** : intégrer les découvertes dans l'évaluation des risques de l'organisation.

3.5 Phase Knowledge : capitaliser et partager

La phase Knowledge est souvent négligée mais elle est fondamentale pour la maturité du programme. Chaque hunt, qu'il ait trouvé quelque chose ou non, produit de la connaissance qui doit être documentée et partagée :

- **Hunt report** : hypothèse, méthodologie, requêtes utilisées, résultats, conclusions et recommandations. Format standardisé pour faciliter la réutilisation.
- **Hunt playbook** : procédure reproductible qui peut être exécutée par un autre analyste. Inclut les requêtes, les critères d'analyse et les seuils de décision.
- **Analytics** : requêtes ou détections réutilisables produites par le hunt (au format Sigma pour la portabilité).
- **Formation** : partager les enseignements avec l'équipe SOC pour élever le niveau collectif.

Cas concret
L'attaque SolarWinds a démontré comment un adversaire sophistiqué peut contourner la détection SIEM en mimant les communications réseau légitimes. Le backdoor SUNBURST communiquait avec son C2 via des requêtes DNS apparemment normales, soulignant la nécessité d'une détection basée sur l'analyse comportementale.

Pratiquez-vous le threat hunting proactif ou attendez-vous que les alertes se déclenchent ?

L'outil **RITA** (Real Intelligence Threat Analytics) automatise cette analyse sur les logs Zeek. Il calcule automatiquement les scores de beaconing pour chaque paire source-destination et fournit un dashboard des résultats. Nous y reviendrons dans la section outils.

4.4 TTP-Based Hunting

Le hunting basé sur les TTP (Tactics, Techniques and Procedures) utilise le framework MITRE ATT&CK comme guide structuré. Pour chaque technique ciblée, le hunter identifie les **data sources** pertinentes, les **indicateurs comportementaux** attendus et construit les requêtes correspondantes. Cette approche est systématique et reproductible.

Exemple de hunt TTP-based pour la technique **T1053.005 - Scheduled Task** (création de tâches planifiées pour la persistance) :

```
# Hunt: Scheduled Tasks suspectes (T1053.005)
# Data sources: Sysmon Event ID 1 (Process Create), Windows Event ID 4698 (Task Created)

# Étape 1: Identifier toutes les tâches créées récemment
GET wazuh-alerts-*/_search
{
  "query": {
    "bool": {
      "must": [
        { "match": { "data.win.system.eventID": "4698" }},
        { "range": { "timestamp": { "gte": "now-30d" }}}
      ]
    }
  },
  "aggs": {
    "task_creators": {
      "terms": { "field": "data.win.eventdata.subjectUserName.keyword", "size": 100 },
      "aggs": {
        "tasks": {
          "terms": { "field": "data.win.eventdata.taskName.keyword", "size": 50 }
        }
      }
    }
  }
}

# Étape 2: Filtrer les créateurs de tâches inhabituels
# - Comptes de service créant des tâches → normal
# - Comptes utilisateur standard créant des tâches → suspect
# - Tâches avec des actions pointant vers %TEMP%, %APPDATA%, C:\Users\Public → suspect
# - Tâches avec des déclencheurs "AtLogon" ou "AtStartup" → persistance probable

# Étape 3: Pour chaque tâche suspecte, pivoter vers le processus créateur
# Correler avec Sysmon EventID 1 pour identifier la chaîne de processus complète
```

Astuce : utilisez la matrice ATT&CK comme backlog de hunts
Parcourez systématiquement les techniques ATT&CK pertinentes pour votre environnement. Pour chaque technique, évaluez : (1) avez-vous les données nécessaires pour la détecter ? (2) avez-vous des règles de détection automatiques ? (3) quand avez-vous mené un hunt ciblé pour la dernière fois ? Cette approche garantit une couverture progressive et identifie les lacunes de visibilité. Consultez notre article sur les top techniques MITRE ATT&CK pour prioriser.

```
# Installation et utilisation de RITA
# 1. Installer RITA (nécessite MongoDB)
wget https://github.com/activecm/rita/releases/latest/download/install.sh
chmod +x install.sh && sudo ./install.sh

# 2. Importer des logs Zeek
rita import /opt/zeek/logs/2026-03-01/ dataset_20260301

# 3. Afficher les résultats de beaconing
rita show-beacons dataset_20260301

# Score | Source          | Destination     | Connections | Avg Bytes | Ts Score | Ds
Score
# 0.987 | 10.10.5.42     | 185.141.25.168  | 8432        | 128       | 0.99     | 0.95
# 0.912 | 10.10.12.88    | cdn-static.xyz  | 2156        | 256       | 0.94     | 0.88
# 0.034 | 10.10.1.15     | update.msft.com | 48          | 4096      | 0.12     | 0.02

# 4. Afficher les connexions longues (potential tunnels)
rita show-long-connections dataset_20260301

# 5. Analyser les requêtes DNS (domains suspects, exfiltration)
rita show-exploded-dns dataset_20260301
```

Un score de beaconing supérieur à **0.80** est généralement suspect et justifie une investigation approfondie. RITA est particulièrement efficace pour détecter les C2 qui utilisent des intervalles réguliers (Cobalt Strike avec son jitter par défaut, par exemple). Pour les C2 avec un jitter élevé, l'analyse manuelle de fréquence décrite dans la section 4.3 reste nécessaire.

5.4 Jupyter Notebooks : l'environnement d'analyse avancée

Les **Jupyter Notebooks** sont devenus l'environnement de prédilection des threat hunters avancés. Ils permettent de combiner code Python, visualisations, requêtes SIEM et documentation dans un document unique et reproductible. Le projet **MSTIC Jupyter Notebooks** de Microsoft et la bibliothèque **MSTICPy** fournissent des composants prêts à l'emploi pour le hunting :

- **Connecteurs data** : interrogation directe d'Elasticsearch, Splunk, Sentinel, VirusTotal, Shodan depuis le notebook
- **Enrichissement** : résolution IP/domaine, géolocalisation, réputation, whois automatisés
- **Visualisations** : timelines d'événements, graphes de processus, heatmaps de connexions
- **Analyse statistique** : bibliothèques pandas, scipy, scikit-learn pour le clustering, la détection d'anomalies et l'analyse de séries temporelles

Hypothèse : *"Un attaquant ayant compromis un compte utilisateur utilise PsExec, WMI ou WinRM pour se déplacer latéralement vers des serveurs du réseau interne, en dehors des patterns d'administration normaux."*

Execute

```
# Étape 1: Baseline - Qui administre normalement quoi ?
# Identifier les paires (compte, machine cible) normales via les 30 derniers jours

# WinRM (Event ID 4624, LogonType 3, AuthPackage Kerberos/NTLM via port 5985/5986)
GET wazuh-alerts-*/_search
{
  "size": 0,
  "query": {
    "bool": {
      "must": [
        { "match": { "data.win.system.eventID": "4624" }},
        { "match": { "data.win.eventdata.logonType": "3" }},
        { "range": { "timestamp": { "gte": "now-30d", "lt": "now-7d" }}}
      ]
    }
  },
  "aggs": {
    "admin_pairs": {
      "composite": {
        "size": 10000,
        "sources": [
          { "user": { "terms": { "field": "data.win.eventdata.targetUserName.keyword" }}},
          { "target": { "terms": { "field": "agent.name.keyword" }}}
        ]
      }
    }
  }
}

# Étape 2: Rechercher les nouvelles paires dans les 7 derniers jours
# Toute paire (compte, machine) apparaissant dans les 7 derniers jours
# mais ABSENTE du baseline des 30 jours précédents = anomalie

# Étape 3: Filtrer les résultats
# - Exclure les comptes de service connus (svc_*, SYSTEM)
# - Exclure les accès depuis les jump servers / PAW officiels
# - Exclure les logons liés au patching (SCCM, WSUS)
# - Se concentrer sur: comptes utilisateur standard → serveurs
#   et comptes admin → machines hors de leur périmètre
```

La clé de ce hunt est la **comparaison baseline vs actuel**. Un mouvement latéral crée par définition de nouvelles paires d'accès qui n'existaient pas auparavant. L'attaquant peut utiliser des outils légitimes, mais il ne peut pas empêcher la création de nouveaux patterns d'accès.

6.3 Hunt #3 : Data staging avant exfiltration

Avant d'exfiltrer des données, un adversaire doit les **collecter et préparer** (data staging). Cette phase génère des indicateurs détectables : compression de gros volumes, copie de fichiers sensibles vers un répertoire de staging, chiffrement de données.

Pour la plupart des organisations, le **modèle hybride** est le plus réaliste. Un hunt lead expérimenté définit les hypothèses, structure le programme et mentore les analystes SOC qui participent aux hunts en rotation. Cette approche permet de faire monter en compétence l'ensemble de l'équipe tout en maintenant la continuité du programme.

8.3 Montée en maturité progressive

La mise en place d'un programme de hunting suit généralement un parcours en 4 niveaux de maturité :

- **Niveau 1 - IOC Sweeping (0-3 mois)** : recherche réactive d'IOC issus du threat intelligence dans les logs SIEM. Pas d'hypothèse originale, mais développe la compétence de requêtage et la familiarité avec les données.
- **Niveau 2 - Hunts structurés (3-6 mois)** : formulation d'hypothèses basées sur ATT&CK, exécution de hunts avec le modèle PEAK, documentation systématique. Focus sur les techniques les plus courantes (T1059, T1053, T1021).
- **Niveau 3 - Hunts avancés (6-12 mois)** : utilisation de Jupyter notebooks, analyses statistiques avancées (clustering, anomaly detection), corrélation multi-sources, hunts proactifs basés sur le renseignement sectoriel.
- **Niveau 4 - Programme mature (12+ mois)** : couverture ATT&CK systématique, métriques de programme, création automatisée de détections, partage inter-organisations, contribution aux communautés (règles Sigma, threat intelligence).

Pour approfondir ce sujet, consultez notre outil open-source threat-hunting-queries qui facilite le threat hunting proactif.

Questions frequentes

Comment mettre en place Threat Hunting dans un environnement de production ?

La mise en place de Threat Hunting en production necessite une planification rigoureuse, incluant l'evaluation des prerequis techniques, la definition d'une architecture cible, des tests de validation approfondis et un plan de deploiement progressif avec des points de controle a chaque etape.

Pourquoi Threat Hunting est-il essentiel pour la securite des systemes d'information ?

Threat Hunting constitue un element fondamental de la securite des systemes d'information car il permet de reduire significativement la surface d'attaque, d'ameliorer la detection des menaces et de renforcer la posture globale de securite de l'organisation face aux cybermenaces actuelles.

Combien de règles de détection faut-il pour démarrer avec Threat Hunting : Méthodologie, Outils et Pratique ?

Commencez par 20 à 30 règles alignées sur les techniques MITRE ATT&CK les plus courantes. Mieux vaut peu de règles bien calibrées que des centaines qui génèrent du bruit.

Sources et références : MITRE ATT&CK · MITRE CAR

> Points clés à retenir
>
> • Questions frequentes
>
> • 9. Conclusion

9. Conclusion

Le threat hunting représente l'évolution nécessaire du SOC face à des adversaires qui opèrent sous le radar des détections automatisées. En passant d'une posture purement réactive à une capacité de chasse proactive, les organisations réduisent drastiquement le dwell time, découvrent des compromissions invisibles aux outils et construisent progressivement une défense plus robuste et plus adaptée à leur environnement spécifique.

Les clés du succès sont claires : une **méthodologie structurée** (le modèle PEAK), des **techniques d'analyse maîtrisées** (stack counting, long tail, beaconing detection), des **outils adaptés** (SIEM + Velociraptor + RITA + Jupyter) et surtout des **analystes compétents et motivés** par la chasse. Le threat hunting n'est pas un projet avec un début et une fin -- c'est un programme continu qui s'améliore à chaque itération.

Commencez petit : un analyste, un hunt par semaine, basé sur les techniques ATT&CK les plus courantes dans votre secteur. Documentez tout, capitalisez les enseignements, transformez chaque découverte en détection automatisée. En 12 mois, vous aurez construit une capacité qui transforme fondamentalement la posture de sécurité de votre organisation.

En résumé : Le threat hunting est l'art de formuler la bonne question, de la poser aux bonnes données et d'agir sur la réponse. C'est la discipline qui transforme un SOC réactif en un SOC proactif -- et un défenseur passif en un chasseur qui prend l'initiative.

Ayi NEDJIMI Consultants — Expert cybersécurité offensive & intelligence artificielle

ayinedjimi-consultants.fr · ayi@ayinedjimi-consultants.fr