Sentinel by CryptoITData

Votre serveur est attaqué en ce moment. Voyez par qui.

Un agent de sécurité qui tourne 24/7 sur un seul serveur Linux : il collecte depuis six sources, ouvre des incidents, met le bouton de blocage sur votre téléphone, scanne les vulnérabilités et génère des procédures de correctif avec un chemin de retour vérifié.

Les chiffres ci-dessous sont réels, issus d'une installation en production — AlmaLinux 9, 14 applications surveillées, les dernières 24 heures.

Détecte

Six sources de collecte, des règles déterministes toutes les 10 secondes, des incidents dédupliqués par empreinte.

Alerte

L'incident sur Telegram en quelques secondes, avec le bouton de blocage dessous et un verdict IA à côté du verdict déterministe.

Répare

Des plans de correctif priorisés par KEV et exposition, avec sauvegarde vérifiée et retour arrière automatique.

1 055	2 765	6	0	14
attaquants uniques / 24h	événements hostiles	sources de collecte	vulnérabilités ouvertes	services actifs

LE TABLEAU DE BORD DE SÉCURITÉ

La première ligne répond à une seule question : est-ce que ça va ?

L'ordre de lecture est le design. Le verdict, puis les chiffres qui le soutiennent, puis l'interprétation, puis les tableaux d'où il sort. Celui qui ne lit que la première ligne a déjà la bonne réponse.

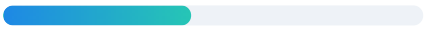
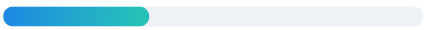
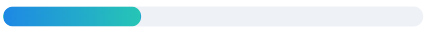
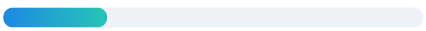
sentinel.example.com — dernières 24 h

Sous attaque constante, la défense tient

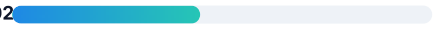
1 055 adresses hostiles distinctes ont généré 2 765 événements sur les dernières 24 h. 5 avertissements à vérifier.

Attaquants uniques	Événements hostiles	Incidents ouverts	Vulnérabilités	IP bloquées
1055	2765	554	0	5
↑ 41 %	↑ 62 %	13 graves	aucune exploitée activement	en registre
dernières 24 h	sur 7 031 au total			

ORIGINE DES ATTAQUES · 7 JOURS

DE		1207 ev · 52 IP
US		970 ev · 402 IP
SG		545 ev · 15 IP
HK		428 ev · 26 IP
CN		394 ev · 16 IP
AD		299 ev · 8 IP

OPÉRATEURS RÉSEAU

AS401116		666 ev · 3 IP
AS13220		306 ev · 1 IP
AS45102		305 ev · 25 IP

SOURCES DE DONNÉES

nginx	3242	ACTIVE
suricata	1465	ACTIVE
auditd	1364	ACTIVE
sshd	907	ACTIVE
sudo	52	ACTIVE

QUI VOUS ATTAQUE · 48 H

ADRESSE	PAYS	SOURCES	VUE PAR	ÉV.	ÉTAT
203.0.113.41	HK	3	auditd+sshd+suricata	281	BLOQUÉE
203.0.113.42	DE	3	auditd+sshd+suricata	113	MULTI-VECTOR
203.0.113.43	DE	3	auditd+sshd+suricata	68	MULTI-VECTOR
203.0.113.44	RO	3	auditd+sshd+suricata	33	MULTI-VECTOR

Le vrai tableau de bord, avec les données de production du jour. « Multi-vecteur » signifie que l'adresse a été vue par trois collecteurs indépendants — un scanner de masse touche une seule surface ; qui en essaie trois vous a choisi.

INTERPRÉTATION

Un nombre n’est pas une observation.

« 2 765 événements » est un chiffre. « root a été visé 1 086 fois depuis 93 adresses — désactivez l’authentification par mot de passe pour root » est une observation. Le moteur applique des règles déterministes aux données et ne sort que ce sur quoi vous pouvez agir.

5 adresses vous attaquent sur plusieurs fronts à la fois

Elles apparaissent dans 3 sources indépendantes sur les dernières 48 h. Un scanner de masse touche une seule surface ; qui en essaie trois vous a choisi.

```
À FAIRE /block 203.0.113.41 24h
```

root est la cible n° 1 — 1 086 tentatives depuis 93 adresses

Tant que l’authentification SSH par mot de passe pour root est possible, chacune de ces tentatives a une chance.

```
À FAIRE grep PermitRootLogin /etc/ssh/sshd_config
```

Infrastructure d’attaque concentrée : Netiface America, Inc.

666 événements hostiles depuis seulement 3 adresses du même réseau. C’est de l’infrastructure d’attaque louée, pas des machines compromises éparpillées — et la prochaine adresse viendra du même endroit.

5 blocages sont au registre mais plus dans le pare-feu

Les blocages ne survivent délibérément pas à un redémarrage — c’est le filet anti-verrouillage. Le tableau de bord vous le dit, au lieu de vous laisser croire que vous êtes protégé.

Campagne active contre GLPI : 560 sondages

Persistante et ciblée — mais vous n’utilisez pas GLPI, donc elle ne peut pas vous atteindre. C’est du bruit de fond, utile seulement pour savoir dans quelle vague vous êtes.

Zéro vulnérabilité de sécurité ouverte

4 716 ont été corrigées et confirmées par le scan suivant.

Les 11 observations ci-dessus sont générées à partir des données réelles du serveur, sans appel à un modèle — la page s’affiche instantanément et dit la même chose deux fois de suite.

LE CANAL D'URGENCE

L'alerte arrive sur votre téléphone avec le bouton d'arrêt dessous.

Long-polling : aucun port ouvert en plus et aucun endpoint public à falsifier. Ça marche même si nginx ou le certificat sont cassés — exactement quand vous en avez besoin.

Le bouton n'est pas l'autorité

L'appui passe par le même contrôle de rôle qu'une commande tapée, et l'exécuteur refuse de toute façon de bloquer l'adresse d'administration — même si le bouton est pressé par erreur.

Il survit à la panne

Le bot tire les messages au lieu de les recevoir : aucun port ouvert ne pointe vers lui et il n'y a pas d'endpoint public à falsifier. Si nginx, TLS ou le tableau de bord tombent, le canal tient.

Lire est vert, modifier est rouge

Les commandes qui changent quelque chose demandent une confirmation séparée, et les destructrices une seconde qui redit la cible.

Un vrai transcript : les messages, les libellés des boutons et les noms de commandes sont ceux qu'envoie le bot, pas une maquette.

DU PAQUET À LA DÉCISION

Ce qui se passe entre l'attaque et le bouton sur votre téléphone.

Cinq étapes, chacune avec un travail qu'elle ne délègue pas à la suivante. La partie déterministe décide ; le modèle explique ensuite.

01 · COLLECTE

Six sources, filtrées au collecteur

sshd, sudo/su, nginx, auditd, Suricata. Le bruit est rejeté avant la base — 96 000 diagnostics moteur par jour n'y arrivent jamais.

02 · DÉTECTION

Des règles déterministes, toutes les 10 secondes

Des seuils honnêtes sur du SQL indexé : brute-force, énumération web, alertes IDS, anomalie de volume. Coût nul par événement.

03 · INCIDENT

Dédupliqué par empreinte

400 tentatives depuis une adresse font un incident, pas 400. L'acteur reçoit un pays et un opérateur depuis des bases géo locales.

04 · TRIAGE IA

Du jugement, pas des scores

Le modèle recalibre la sévérité et rédige le résumé. Le verdict déterministe reste à côté — le désaccord est visible.

05 · DÉCISION

La vôtre, d'une pression

Le blocage automatique existe et est testé, mais il est livré **désactivé**. Les 72 premières heures vous montrent ce qu'il *aurait* bloqué.

× Le modèle ne décide pas des blocages

La détection, le score et la décision sont déterministes et tournent à coût nul. Une API indisponible ou un budget épuisé ne change rien à ce qui est bloqué.

× Le modèle n'exécute rien

Le plan qu'il écrit passe par un validateur déterministe qui le refuse s'il n'est pas sûr. Le modèle rédige, le validateur décide.

× Le modèle n'est pas sur le chemin critique

API en panne ou budget épuisé : la détection, le blocage et l'alerte continuent, marqués « analyse IA indisponible ».

× Le modèle ne dépasse pas le budget

Le plafond est vérifié avant chaque appel, et le coût par incident est visible dans le tableau de bord.

La détection tourne à coût nul par événement ; le modèle n'est appelé que pour les incidents qui dépassent le seuil de sévérité — quelques dizaines par jour en général.

VULNÉRABILITÉS ET CORRECTIFS

Scanner est la partie facile. Ce que vous corrigez en premier est la difficile.

Priorisation par **CVSS × EPSS × KEV × exposition × criticité**. Un CVSS 6.5 sur la liste CISA des failles activement exploitées passe devant un 9.8 que personne n'exploite.

sentinel.example.com/patches/3 — **Plan de correctif #3** — VALIDÉ

Niveau de risque	Impact	Indisponibilité estimée	Réversible	Redémarrage requis
élevé	un seul service	0s	oui	non

CIBLE · EMPREINTE

CIBLE `postgresql · hash 909a1d35a818de31…`

PREFLIGHT — VÉRIFICATIONS AVANT

check_disk	no_incident	curl_installed
disk_free	no_open_incident	pkg_version

APPLICATION · RETOUR ARRIÈRE

APPLIQUER `dnf -y update curl`

RETOUR `dnf -y downgrade curl-7.76.1-29.el9`

SAUVEGARDE `rpm_state · curl`

× Exploité l'emporte sur théorique

Une vulnérabilité sur la liste CISA des failles activement exploitées passe devant un score plus élevé que personne n'utilise réellement.

× L'exposition l'emporte sur l'inventaire

Ce qui est joignable depuis internet est corrigé avant ce qui est derrière le pare-feu avec la même CVE.

× Les actifs critiques ne sont jamais automatiques

Les bases de données et tout ce que vous avez marqué critique ne reçoivent jamais de correctif généré automatiquement — elles reçoivent un plan que vous approuvez.

Un vrai plan, généré automatiquement et validé du premier coup. Le retour arrière fige correctement l'ancienne version — la solution ennuyeuse qui marche.

COMMENT SE CONSTRUIT UN CORRECTIF

Six étapes, et le modèle n’est présent qu’à la première.

01 **Génération**

À partir des faits déterministes du scan. Aucun shell n’est donné au modèle.

02 **Validation**

Chaque commande est une liste, le premier élément vient d’une liste autorisée, `rm` interdit. Invalide → un nouvel essai → rejeté.

03 **Approbation**

Deux confirmations sur Telegram, un jeton à usage unique lié à l’empreinte.

04 **Sauvegarde**

Un point vérifié en relisant les archives. Sans lui, le correctif ne démarre pas.

05 **Application**

Étape par étape, écrite en base *avant* chaque commande.

06 **Retour arrière**

Automatique à la moindre panne. Un retour arrière raté est un état distinct, plus grave.

> Le modèle est présent à l’étape 01 et nulle part ailleurs. Tout ce qui suit est déterministe : un validateur qui refuse, un jeton à usage unique, une sauvegarde relue avant d’être comptée, et un retour arrière qui s’exécute à la moindre panne.

× **Aucun shell pour le modèle**

Il écrit un plan à partir des faits déterministes du scan. Il n’exécute jamais de commande, sur aucun hôte.

× **Un nouvel essai, puis un refus**

Un plan invalide est régénéré une seule fois. Le second échec est un refus, pas un contournement.

× **Écrit avant de s’exécuter**

Chaque étape arrive en base avant l’exécution de la commande, pour qu’une exécution interrompue reste lisible ensuite.

Les correctifs sont générés automatiquement ; ils ne sont jamais appliqués sans deux confirmations explicites.

APPROBATION DES CORRECTIFS

La première pression n’applique rien.

Un bouton vit pour toujours dans l’historique d’une conversation. Donc le bouton *n’est pas* l’autorité — il porte une référence vers elle : un jeton à usage unique, lié à la conversation et à l’empreinte du plan, avec expiration.

Le second écran apporte de l’information

Sinon « êtes-vous sûr ? » n’est qu’une boîte de dialogue qu’on ferme par réflexe. Il redit la cible, l’indisponibilité et ce qui est créé avant que quoi que ce soit ne soit touché.

Un plan régénéré tue les anciens boutons

Si le plan est régénéré entre-temps, l’empreinte change et tous les boutons précédents meurent.

Le jeton est à usage unique

Lié à la conversation et à l’empreinte du plan, avec expiration. Transférer le message ne transfère rien d’utilisable.

La génération du plan est automatique. L’application demande deux confirmations explicites, depuis le téléphone, par quelqu’un ayant le bon rôle.

ARCHITECTURE DE SÛRETÉ

Un agent capable de bloquer internet se juge à ce qu'il refuse.

Tout ce qui est privilégié passe par un seul processus root de quelques centaines de lignes, sans dépendances et sans import du reste de l'application. Le reste tourne sans privilèges.

× Il ne peut pas vous verrouiller dehors

Deux couches indépendantes : la table nftables a `policy accept` avec une liste blanche avant les blocages, et l'exécuteur refuse directement l'adresse d'administration. Vérifié en vrai — le refus fonctionne.

× Le tableau de bord ne peut bloquer personne

L'interface web peut seulement *annuler* un blocage, jamais en créer un. Un tableau de bord compromis reste une fuite de données, pas une prise de contrôle.

× `rm` n'est pas sur la liste

Absent des binaires autorisés pour les correctifs. La suppression passe par une seule opération, strictement limitée au répertoire de sauvegarde, le chemin étant revérifié après résolution.

× Le script de restauration s'écrit lui-même

Généré *à l'intérieur* de l'exécuteur, à partir de ce qu'il voit sur le disque. Du texte venu du côté non privilégié ne devient jamais une ligne exécutable lancée en root.

× Une sauvegarde vide n'est pas une sauvegarde

Les archives sont relues et re-hachées au scellement. Un artefact de taille nulle est rejeté — sinon il ressemble à un chemin de retour jusqu'au moment précis où vous en avez besoin.

× On ne peut pas le tromper par les journaux

Les lignes de log sont encapsulées comme données non fiables, et le modèle ne peut renvoyer qu'un verdict structuré. Une injection réussie change une étiquette, jamais le système.

× La rétention ne supprime pas le dernier chemin de retour

Le dernier point de restauration d'un actif n'est jamais supprimé, aussi ancien soit-il. Une politique de rétention capable de le faire est une perte de données à retardement.

× Vidage d'urgence

`touch /etc/sentinel/PANIC` vide la liste de blocage toutes les 10 secondes tant que le fichier existe. Les blocages ne survivent pas à un redémarrage — redémarrer est toujours une sortie.

INSTALLATION

Une commande, et rien ne change tant que vous n’avez pas dit oui.

L’assistant devine tout seul où il tourne : sur votre portable il demande une cible et installe via SSH ; sur le serveur il saute la moitié SSH.

```
$ git clone https://github.com/dragosnimu/sentinel.git
$ cd sentinel && ./scripts/wizard.sh --dry-run
✓ AlmaLinux 9.8 (famille rhel)
✓ python3.12 — 5208 Mo de mémoire — 94 Go libres
✓ nginx détient 80/443 → le mode « shared » est disponible
! le port 8443 est pris par autre chose
--dry-run : rien n’a été modifié.
```

01 · DEMANDE

Rien n’est deviné en silence

La cible, le domaine, le mode nginx, votre adresse d’administration, Telegram, la clé IA. Là où une valeur par défaut raisonnable existe, vous la voyez avant de l’accepter.

02 · VÉRIFIE

En lecture seule

Distribution, mémoire, disque, qui détient 80/443, si le port demandé est libre, si le domaine résout ici. Rien n’est touché.

03 · RÉSUME

La dernière sortie

Tout ce qui va changer, à l’écran, avant la confirmation. Un installateur qui a déjà édité nginx au moment où il vous demande vous ment.

04 · INSTALLE

Numéroté et reprenable

Chaque étape a un marqueur sur le disque. Si un de vos services s’est arrêté, l’installation revient en arrière toute seule.

× Les secrets n’atteignent jamais la table des processus

Le jeton du bot et la clé d’API sont lus sans écho et voyagent sur stdin. En arguments de ligne de commande, `ps` les montrerait à tous les comptes de l’hôte.

× Deux familles de distributions, un fichier

RHEL (AlmaLinux, Rocky, CentOS Stream, Fedora) et Debian (Debian, Ubuntu). Tout ce qui diffère se trouve dans `deploy/lib/distro.sh`.

× Ce qui n’est pas supporté est refusé nommément

`unsupported distribution` — un hôte où l’installation a échoué visiblement est récupérable. Un hôte qui semble protégé sans l’être ne l’est pas.

Les réponses peuvent être enregistrées et rejouées sans surveillance — un second serveur, ou une reconstruction, ne pose aucune question.

ÉTAT

Ce qui est livré et tourne en production.

514	31	7	13	0
tests automatisés	invariants de sûreté	services systemd	migrations de schéma	dépendances CDN

Collecte depuis six sources, détection déterministe, incidents avec triage IA, blocage par bouton sur Telegram, scan de vulnérabilités avec priorisation KEV, génération et application de correctifs avec sauvegarde vérifiée et retour arrière automatique. Une interface web sans JavaScript tiers, avec une politique de sécurité stricte et zéro ressource externe.

Le blocage automatique est livré **désactivé** — la mécanique est complète et testée, mais les 72 premières heures servent à observer, pour que vous trouviez les faux positifs avant de couper quelqu'un à 3 h du matin. Les correctifs sont générés automatiquement ; ils ne sont jamais appliqués sans deux confirmations explicites.

× Ce qui n'est pas encore couvert

Les journaux des conteneurs ne sont pas collectés, il n'y a pas de moniteur d'intégrité de fichiers dédié au-delà d'auditd, et l'installation Debian est couverte par des tests mais pas encore vérifiée sur un hôte Debian réel.

× Comment l'obtenir

Installation, Infogéré ou Flotte — la licence est par serveur, et la configuration initiale plus le calibrage de 72 heures sont inclus dans chaque niveau.

PROCHAINE ÉTAPE

Trente minutes, gratuites, sans engagement.

Vous nous dites quel serveur vous avez et ce qui tourne dessus. Nous vous disons ce qui est exposé, ce qui vous attaque en ce moment et si Sentinel a du sens pour vous — y compris quand la réponse est « pas encore ».

cryptoitdata.eu/contact
réserver un échange

cryptoitdata.eu/sentinel/fr
la page produit