IA

L'art de se la jouer en faisant des statistiques

Fabrice Prigent

Université Toulouse 1 Capitole

Capitoul Jeudi 23 juin 2022

Pourquoi ?

Pourquoi ?
Je suis informaticien, j'ai donc 3 qualités:

- la paresse,
- l'orgueil,
- et l'impatience.

Les attaques

A l'UT1 c'est

- entre 1000 et 450 000 tests de port par seconde
- un quadruplement des attaques web en 4 mois.
- un contexte "on veut des solutions, pas des problèmes".

Contexte

- Beaucoup d'étudiants étrangers (donc dégager les russes, les chinois, les irakiens n'est pas possible).
- Des outils de détection approximativement fiables

Alors on fait comment ?

Idée : des maths

Ouais mais les maths ne sont pas vendeurs, donc on va dire que l'on fait de l'IA. C'est pareil, mais c'est plus "startup nation".

- Quelles mathématiques ? Celle du révérend Bayes:
  - On créé des tokens suivant nos besoins
  - On comptabilise quand on bloque (méchant)
  - On comptabilise quand on laisse passer (gentil)
  - On incrémente pour chaque token le nombre de méchants et le nombre de gentils

Les tokens

- basé sur du simple: le log iptables
- on comptabilise en asynchrone
- on évite de faire trop de tokens (pas 1 token par IP par exemple)

Log IPTables

Exemples:

```
Jun 23 08:00:52 fenrir kernel: FIREWALL_DENIED IN=eth0 OUT= MAC=82:45:d3:c8:e9:04:64:c3:d6:55:4f:43:08:00
SRC=222.187.196.46 DST=194.254.255.73 LEN=40 TOS=0x00 PREC=0x00 TTL=240 ID=13627 PROTO=TCP SPT=43788
DPT=2375 WINDOW=1024 RES=0x00 SYN URGP=0

Jun 23 08:00:06 fenrir kernel: FIREWALL_ACCEPT IN=eth0 OUT=eth1 MAC=82:45:d3:c8:e9:04:64:c3:d6:55:4f:43:08
SRC=213.228.60.130 DST=193.49.48.250 LEN=83 TOS=0x00 PREC=0x00 TTL=54 ID=58780 PROTO=UDP SPT=44104
DPT=53 LEN=63 MARK=0x1
```

Ce qui peut être utile

```
Jun 23 08:00:52 SRC=222.187.196.46 DST=194.254.255.73 LEN=40 TOS=0x00 PREC=0x00 TTL=240 ID=13627 PROTO=TCP
SPT=43788 DPT=2375 WINDOW=1024 RES=0x00 SYN URGP=0

Jun 23 08:00:06 SRC=213.228.60.130 DST=193.49.48.250 LEN=83 TOS=0x00 PREC=0x00 TTL=54 ID=58780 PROTO=UDP
SPT=44104 DPT=53 LEN=63
```

UNIVERSITÉ TOULOUSE 1 CAPITOLE

Les tokens: exemples

Quelques types de tokens:

- le CIDR 24 de l'IP d'origine (SRC_)
- la taille de la fenêtre TCP (WIN_)
- le TTL (TTL_)
- le port destination
  - haut ($>$ 49152) $=>$ DPT_HIGH
  - bas ($<$1024) $=>$ DPT_LOW
  - moyen (DPT_MID)
- la longueur du paquet en TCP ou UDP (UDP-LEN_ ou TCP-LEN_)
- ou la combinaison de 2 tokens (SRC_TTL_, WIN_TTL_)

Les tokens: exemples 2

Quelques types de tokens avec leur note de nocivité, leur nombre d'attaques, leur nombre de passages.

Token	**Nocivité**	**Refus**	**Acceptation**
SRC_185.150.190.0-TTL_50	100.00 %	28665088	0
DPT_HIGH	100.00 %	613516725	6787
WIN_1024	99.94 %	3121978586	1977108
TTL_247	99.88 %	27652903	33431
TTL_60	1.76 %	3481873	197878241
WIN_22240	0.81 %	30176	3750560
SRC_172.217.41.0	0.00 %	0	1463525

Note de nocivité ?

Il y a 2 notes de nocivité:

- Celle du token
  - basé sur le théorème de Bayes

Théorème de Bayes

$$proba(Danger \mid Token) = \frac{prop(Token \mid Danger).prop(Danger)}{prop(Token)}$$

- Celle de la communication, basée
  - soit sur la loi du khi 2 (en gros, moindres carrés)
  - soit sur le théorème de Bayes généralisé
  - et appliqué sur l'ensemble des tokens

UNIVERSITÉ TOULOUSE 1 CAPITOLE

Super ! et on en fait quoi ?

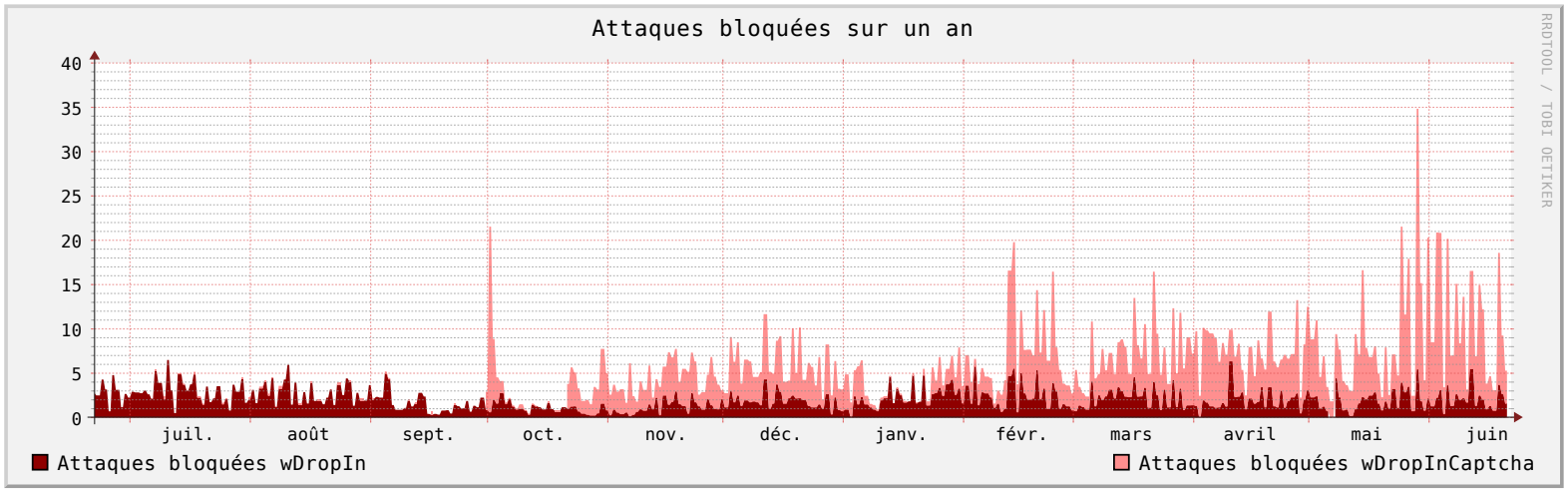
- En théorie: A chaque paquet on calcule dynamiquement la loi du khi2.
- En pratique:
  - On constitue une liste des CIDR trop méchants (IA_HIGH) et une liste des CIDR super lapins trop choupinoux (IA_LOW)
  - On récupère les autres tokens super méchants (TTL, WIN avec plus de 99,9% de nocivité)
  - On teste pour chaque paquet s'il fait l'un ou l'autre.
  - Et on l'envoie sur un reverse-proxy à captcha

Résultats: Nan mais sérieux !?!

Bah oui...

- Ca marche plutôt bien: les attaques web qui auraient du passer, sont à 80% redirigées sur le captcha.
- Les listes (IA_HIGH et IA_LOW) en IP sont distribuées sur d'autres serveurs
  - sur notre serveur de messagerie, pour être plus ou moins restrictif en terme de réception
  - sur nos services non web, pour limiter le nombre de connexion.
- Les "trop choupinoux" ne passent pas dans notre honeypot: S'ils font une action illégale, ils sont "simplement" refusés, on ne va leur mentir.

Vision


Attaques bloquées sur un an
40
35
30
25
20
15
10
5
0
juil.
août
sept.
oct.
nov.
déc.
janv.
févr.
mars
avril
mai
juin
Attaques bloquées wDropIn
Attaques bloquées wDropInCaptcha
RRDTOOL / TOBI OETIKER

Pourquoi tu bloques pas alors ?

- A cause des faux positifs
  - Un serveur OVH a peu de raisons légitimes de venir nous voir...
    - Sauf quand il héberge un serveur "shibollethisé"
    - Sauf quand il vient récupérer nos blacklists
    - Sauf quand il vient moissonner notre OAI
    - Sauf quand...
- Mais on le fera.. plus tard

Les pistes d'amélioration

- Appliquer le calcul complet multi-token.
- Faire un premier filtre, qui renverrait sur un calcul complet de bayésien multi-token (la base est prête et fait 850 Mo).

Schéma IA: constitution de la DB

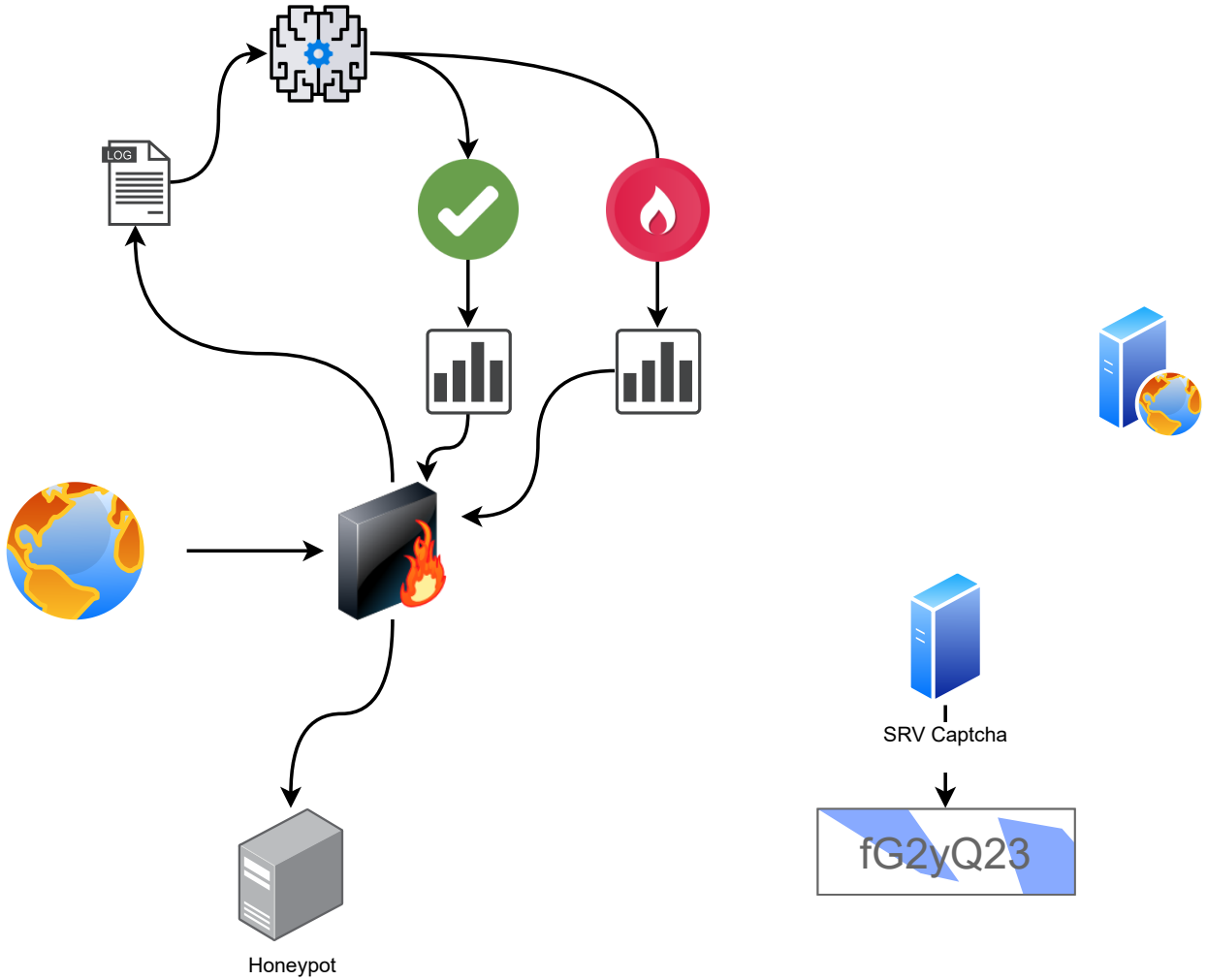


Schéma IA: effets du filtrage

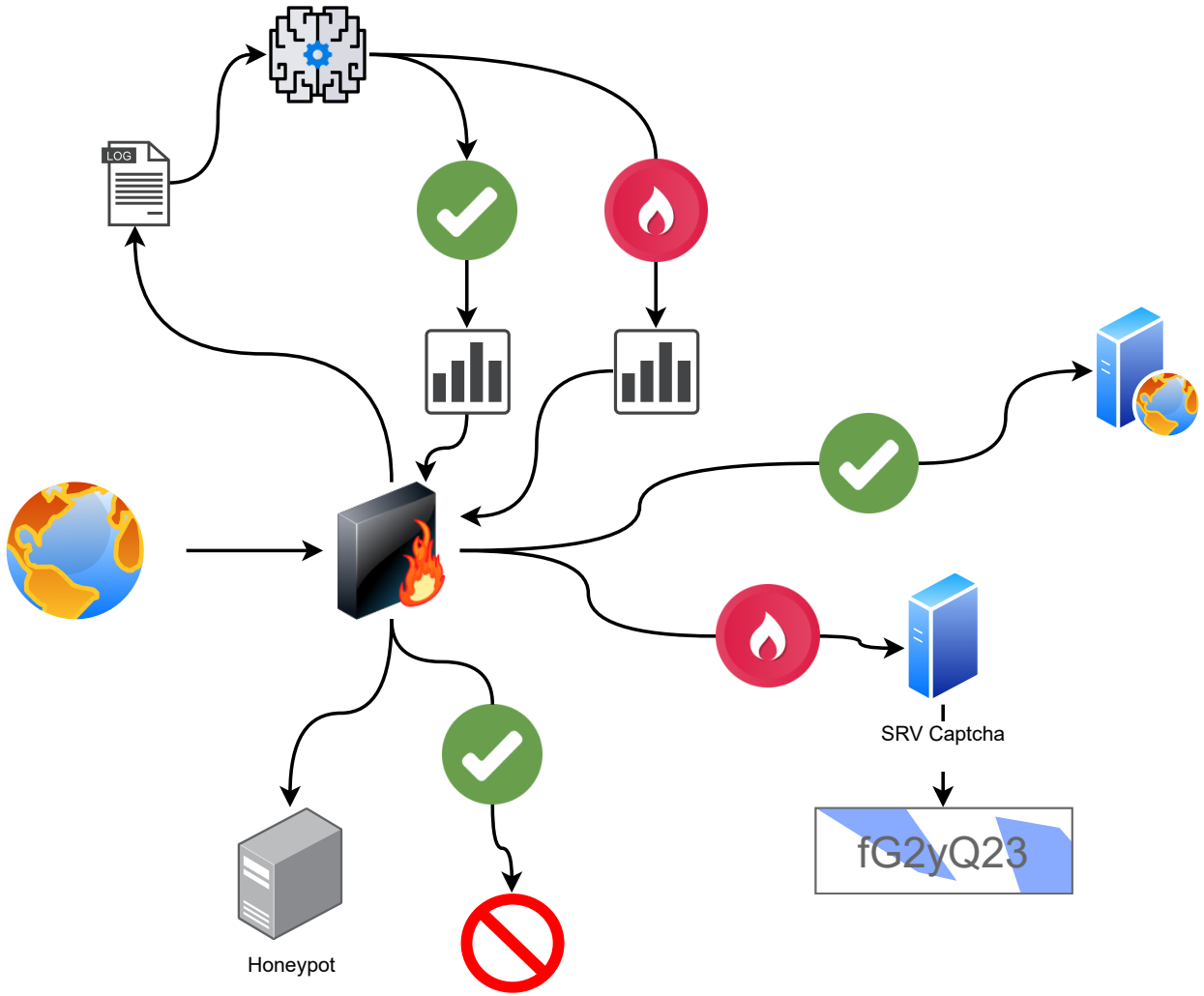


Schéma IA: le Captcha

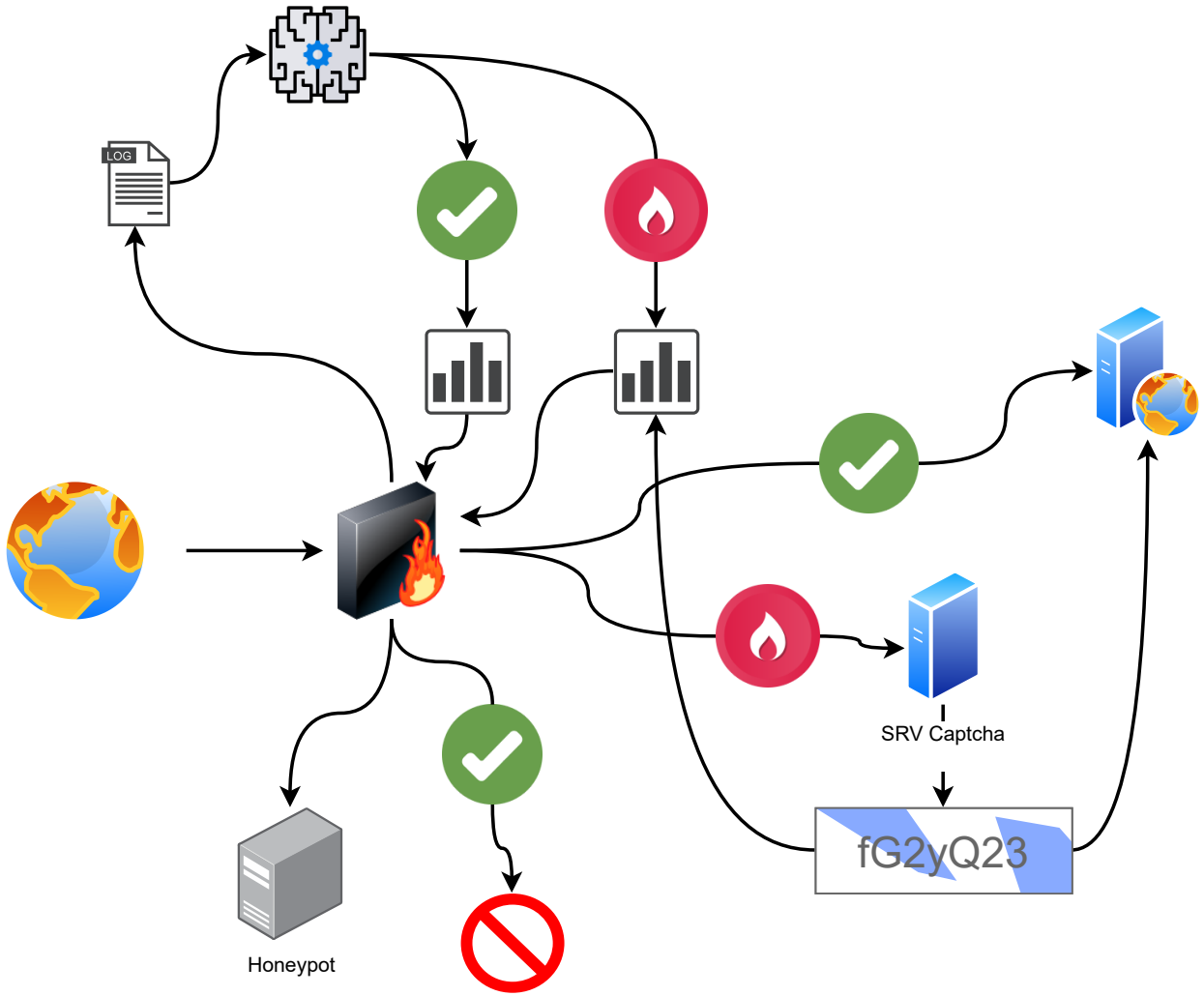
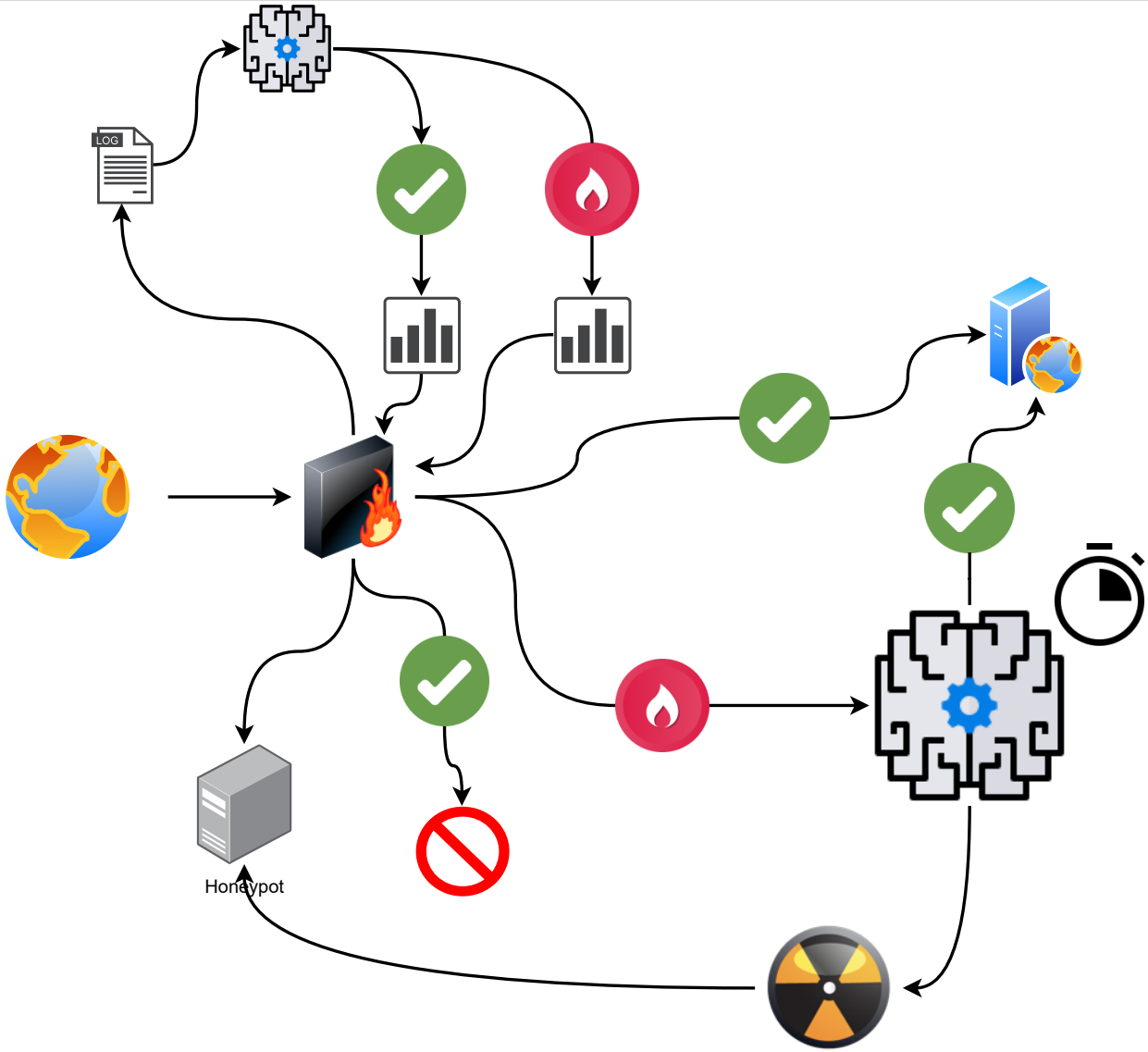


Schéma: quand on sera en full IA



Des questions ?

Des questions ?