C&ESAR 2020

BEEZH: une plateforme de détonation réaliste pour l'analyse des modes opératoires d'attaquants

Frédéric Guihery
Alban Siffer
Joseph Paillard

www.amossys.fr

Contexte de la détonation

Objectifs : mieux comprendre la menace à laquelle une organisation fait face

- Permettre l'identification des **TTP** (Techniques, Tactiques et Procédures) et **IoC** (Indicateurs de Compromission) de groupes d'attaquants ciblant l'organisation
- Production de **renseignements** exploitables et partageables auprès d'équipes SOC/CERT

Contexte de la détonation

Approches actuelles

- Détonation (exécution) d'un artefact dans un environnement de type sandbox, la plupart du temps limité à un OS
- Pas ou peu d'activité utilisateur sur cet OS
- Pas ou peu d'adaptation de l'OS à l'environnement opérationnel
  - Les quelques fois où ce travail de personnalisation est réalisé, celui-ci est fait manuellement

Contexte de la détonation

Approches actuelles

- Détonation (exécution) d'un artefact dans un environnement de type sandbox, la plupart du temps limité à un OS
- Pas ou peu d'activité utilisateur sur cet OS
- Pas ou peu d'adaptation de l'OS à l'environnement opérationnel
  - Les quelques fois où ce travail de personnalisation est réalisé, celui-ci est fait manuellement

=> Ces approches ne permettent pas d'observer tout le spectre des TTP : limitation à quelques étapes de la *kill chain*

=> Leurrage limité à cause du faible réalisme de l'environnement

Notre approche de la détonation

Le cahier des charges

- Plateforme de type honeynet pour simuler un SI complet, laissant l'opportunité à l'attaquant d'exposer l'ensemble de ses TTP
- Automatisation de la construction du SI et de personnalisation à l'organisme ciblé
- Simulation de comportements utilisateurs réalistes
- Assister l'analyste sur l'extraction des traces d'intérêts pour la production de renseignements

www.amossys.fr

Notre approche de la détonation

Le cahier des charges

- Plateforme de type honeynet pour simuler un SI complet, laissant l'opportunité à l'attaquant d'exposer l'ensemble de ses TTP
- Automatisation de la construction du SI et de personnalisation à l'organisme ciblé
- Simulation de comportements utilisateurs réalistes
- Assister l'analyste sur l'extraction des traces d'intérêts pour la production de renseignements

Contributions présentées à C&ESAR et s'appuyant sur BEEZH

www.amossys.fr

BEEZH PLATFORM

AMOSSYS
EXPERT IN CYBERSECURITY

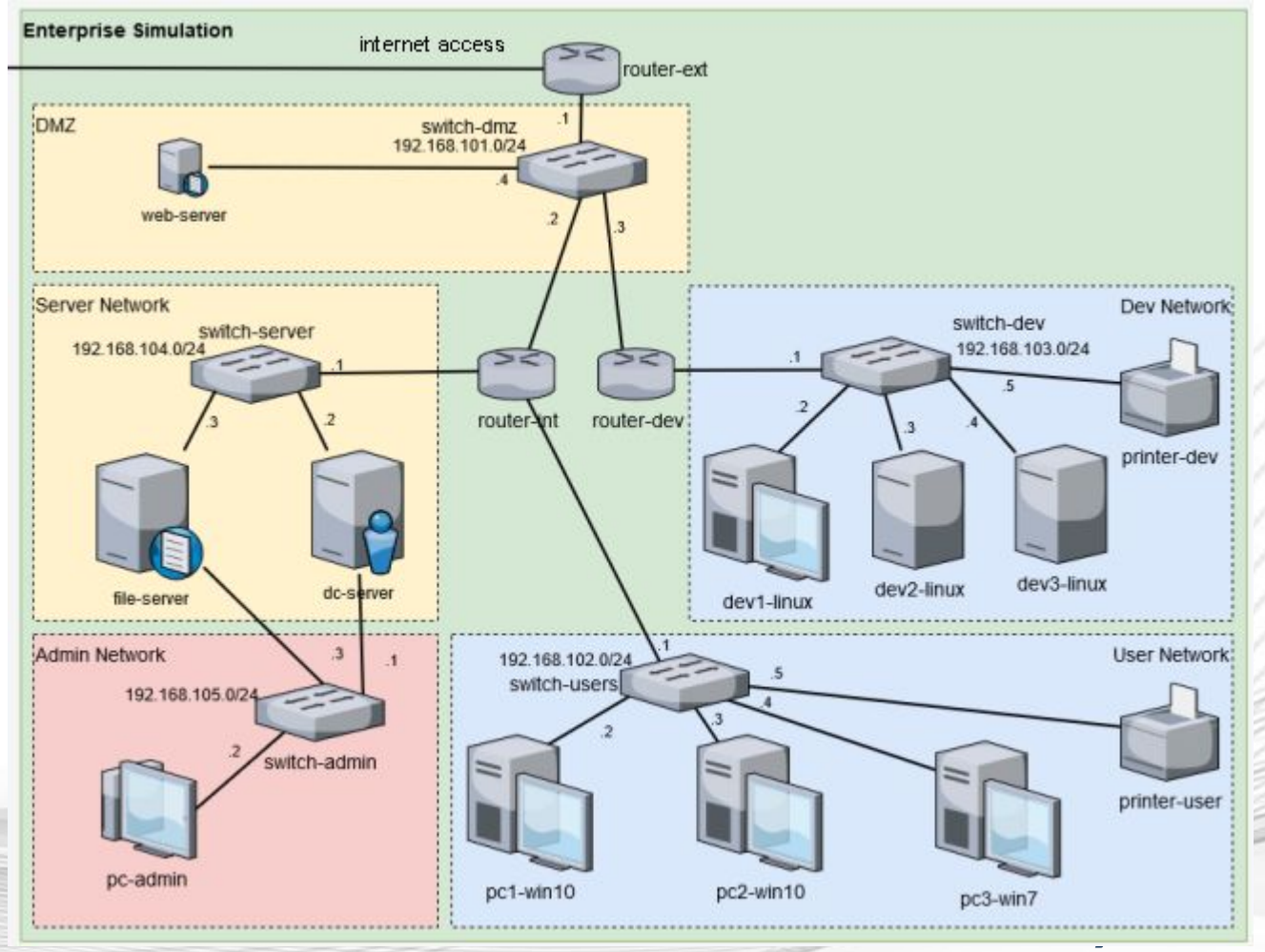
La plateforme BEEZH

www.amossys.fr

La plateforme BEEZH

Innovations

- ❏ Simulation de SI et d'arborescence Active Directory
- ❏ Personnalisation des ressources déployées dans l'environnement et capacité de vieillissement
- ❏ Simulation de comportements utilisateurs et administrateurs sans agent



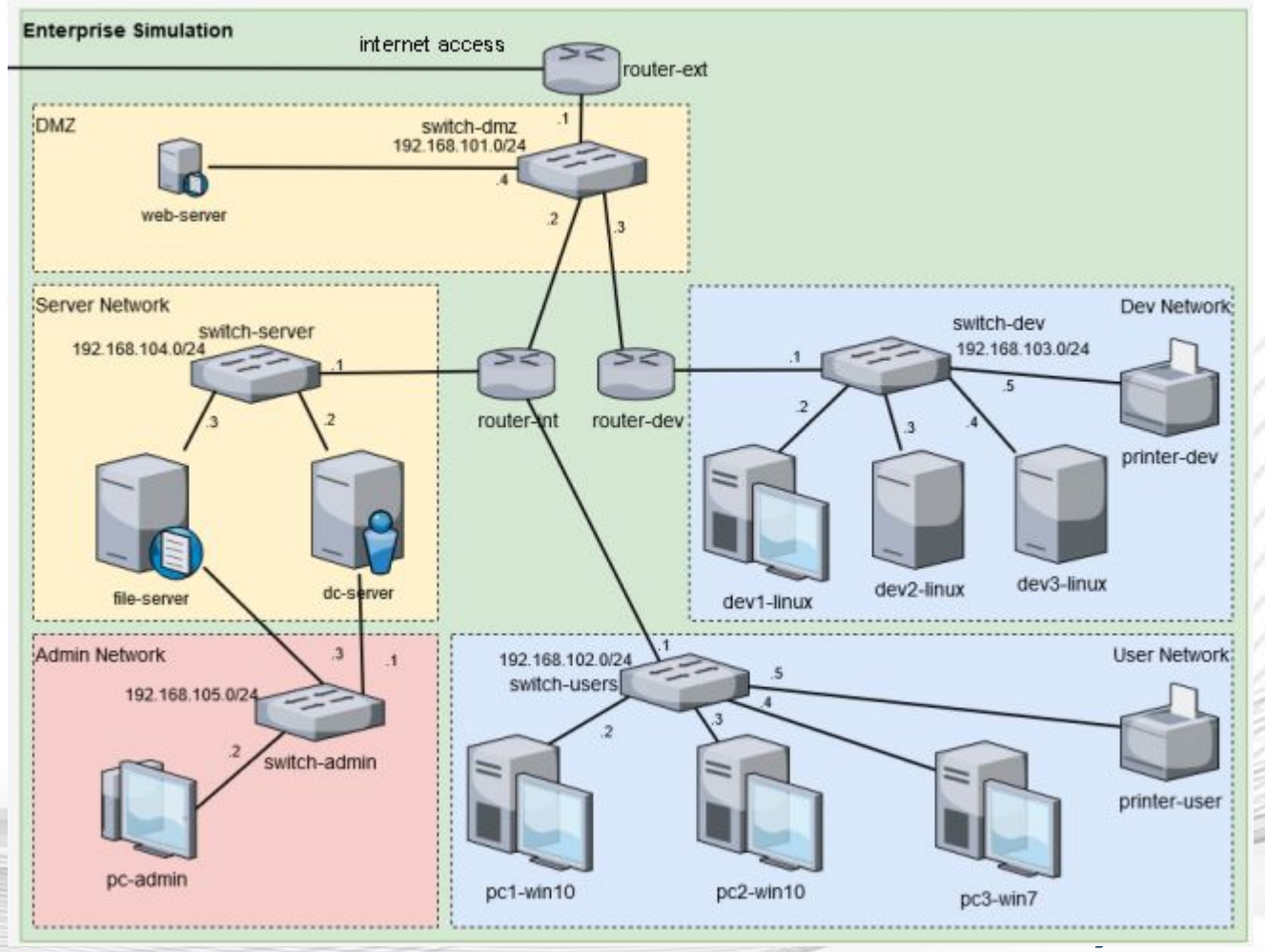
La plateforme BEEZH

Innovations

- ❏ Simulation de SI et d'arborescence Active Directory
- ❏ Personnalisation des ressources déployées dans l'environnement et capacité de vieillissement
- ❏ Simulation de comportements utilisateurs et administrateurs sans agent

A reçu en janvier 2020 le premier prix du défi ***Deceptive Security*** *organisé par*

- *la DGA*
- *le Commandement de la Cyberdéfense*
- *l'Innovation Défense Lab*



La plateforme BEEZH

Automatisation de la génération du SI

- Modélisation dans un format pivot de l'environnement matériel, système et logiciel
- Modélisation dans un format pivot de la topologie réseau (switchs, routeurs, VLAN, adressage statique et dynamique, QOS pour forcer des bandes passantes, latences ou pertes de paquets, ...)
- Modélisation dans un format pivot des directives de paramétrage et de vieillissement du SI

www.amossys.fr

AMOSSYS

EXPERT IN CYBERSECURITY

Simulation de comportements utilisateur réalistes

Le besoin de génération de vie réaliste

- Existence de codes malveillants disposant de routines de détection d'environnement
  - Zebrocy s'exécute uniquement à la fermeture d'un document
  - BaneChan exploite les interactions de l'utilisateur pour échapper à la détection : il attend un nombre donné de clics souris avant de commencer à s'exécuter
- RATs disposant de capacité de capture d'écran et d'analyse de l'activité souris/clavier
- Connexions malveillantes sur déport d'affichage (RDP, VNC, ...) et visualisation des actions légitimes

Le besoin de génération de vie réaliste

- Existence de codes malveillants disposant de routines de détection d'environnement
  - Zebrocy s'exécute uniquement à la fermeture d'un document
  - BaneChan exploite les interactions de l'utilisateur pour échapper à la détection : il attend un nombre donné de clics souris avant de commencer à s'exécuter
- RATs disposant de capacité de capture d'écran et d'analyse de l'activité souris/clavier
- Connexions malveillantes sur déport d'affichage (RDP, VNC, ...) et visualisation des actions légitimes

=> La génération d'actions utilisateur a pour objectif de leurrer ces phases de reconnaissance

Approches pour la génération d'activité utilisateur

- Etat de l'art : orienté réseau majoritairement
  - Génération pure (outils: Iperf, Ostinato)
  - Rejeu de captures (Tcpreply, TCPivo)
  - Génération de données à partir de captures (Harpoon, Swing)

- Quelques travaux utilisent des frameworks de tests graphiques
  - Génération d'activités hôte+réseau cohérentes
  - Utilisation d'un agent sur la machine cible => génère du bruit

Notre approche pour la génération d'activité utilisateur

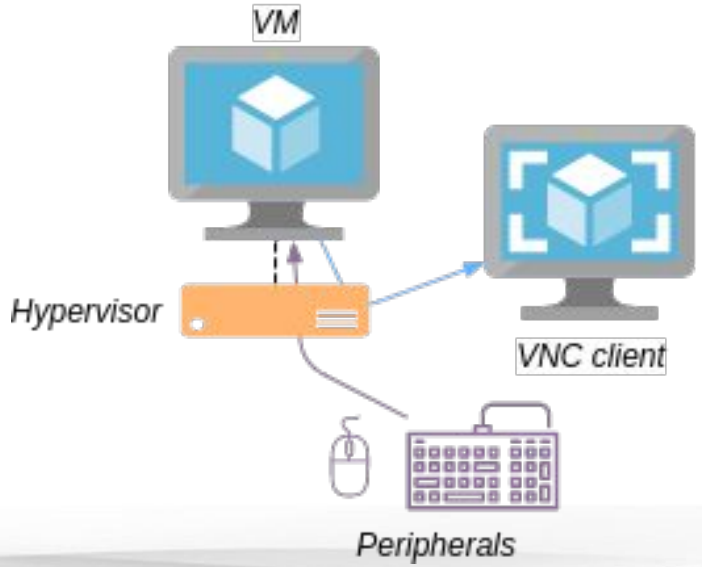
- **Reproduction d'un comportement humain : tâches multiples et difficiles**
  - Simuler la vision: détection des zones d'intérêt
  - Simuler les entrées utilisateurs
    - souris : déplacements, clics/double-clics réalistes
    - clavier : frappes réalistes (vitesse, erreurs), génération de contenu cohérent
  - Simuler l'activité : création d'un scénario cohérent et réaliste mêlant les fonctions précédentes

www.amossys.fr

Notre approche pour la génération d'activité utilisateur

En pratique

- ❏ Pas d'agent dans l'OS : interactions via déport d'affichage exposé par l'hyperviseur:
- ❏ Interactions avec la bibliothèque Sikulix



www.amossys.fr

Notre approche pour la génération d'activité utilisateur

Combinaison

- ❑ Approche déterministe : framework de création de scénarios utilisateurs sur la base d'u un catalogue d'actions
- ❑ Approche opportuniste : moteur d'apprentissage des actions par réseau de neurones

Composants	Framework de création de scénarios	Moteur d'apprentissage
Détail	• actions unitaires relatives à des OS • actions unitaires relatives à des applications	• reconnaissance d'images • détection/génération de texte • détection de zones d'intérêt
Atouts	Actions maîtrisées	Capacité de généralisation (environnements inconnus)
Faiblesses	Faible capacité de généralisation	Prédictions non maîtrisées

www.amossys.fr

Création d'une bibliothèque dédiée à l'apprentissage

Bibliothèque python : `desker` (***desk**top **e**lements **r**ecognition*)

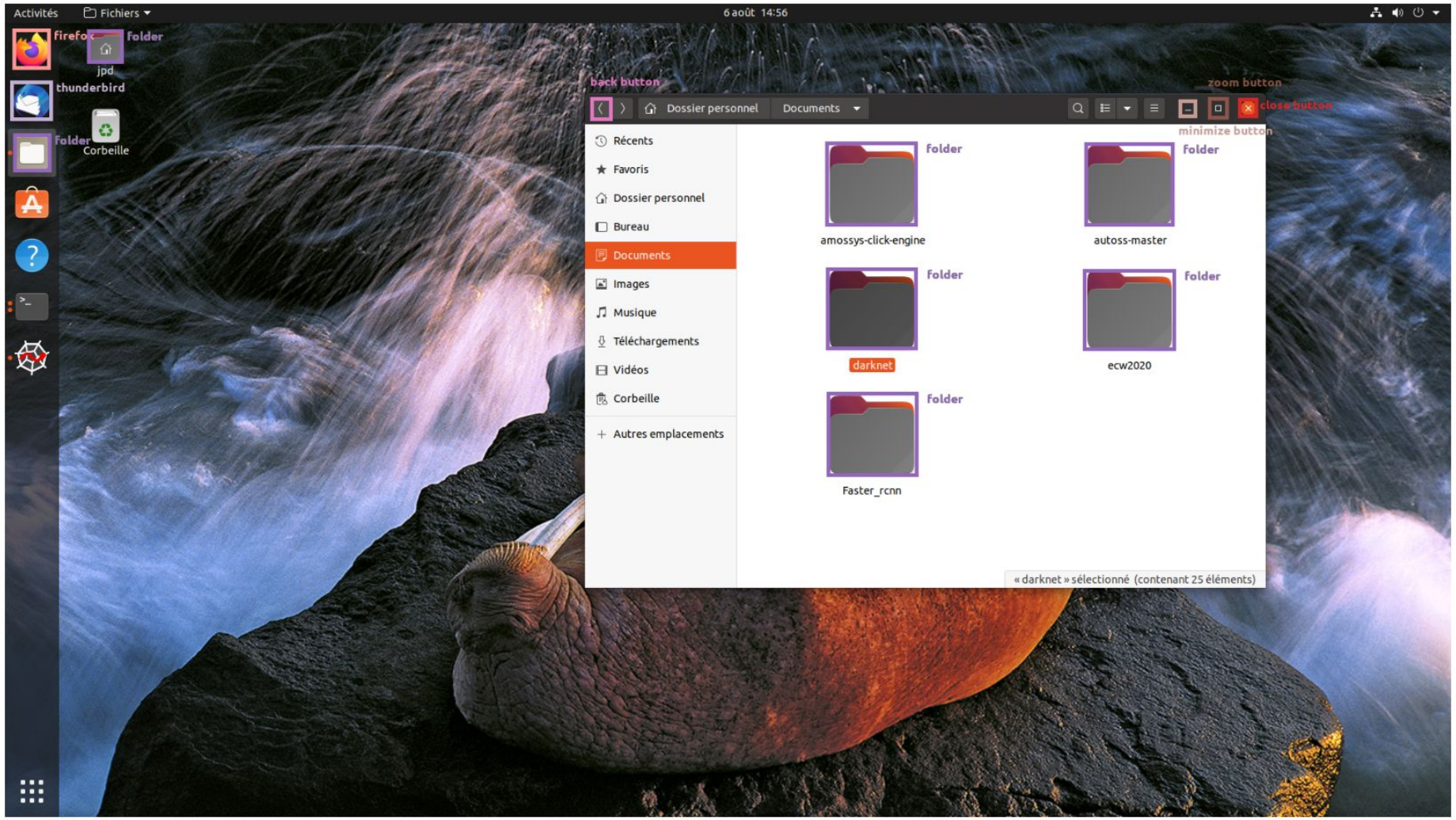
- **Fonctionnalités**
  - détection d'objets (déjà vu)
  - segmentation d'images (détection de fenêtres, paragraphes…)
  - extraction de texte
  - génération de texte
  - détection de liens
  - …
- **Backend**
  - Torch : pour le calcul scientifique
  - Tesseract : pour l'OCR (*optical character recognition*)
  - markovify : pour la génération de texte NLG (*neural langage generation*)

www.amossys.fr

Détection d'objets avec Desker

- **Algorithme: Faster R-CNN**
  - Autres possibilités: YOLO (moins précis) et SSD (plus gourmand)
  - Backbone pré-entraîné: resnet50
  - Fine-tuning des dernières couches pour l'adaptation à notre problème
- **Données**
  - ~100 captures d'écran annotées (images HD majoritairement)
  - ~10 classes (firefox, thunderbird, folder, close button…)
- **Quelques chiffres (avec GPU GTX 1660)**
  - labellisation d'images: travail actuellement manuel, donc assez long
  - entraînement: 15 min (batch réduit, ~15-25 epochs, 4.5Go de mémoire)
  - prédiction: 5 images/s (<250 Mo de mémoire)

Desker en action



Objets annotés

ww.amossys.fr

Desker en action

Identification de liens web

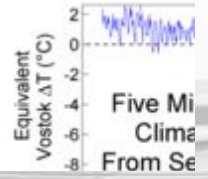
La sédimentologie contribue à l'étude des climats : les échantillons de sédiments marins, précisément géoréférencés, sont l'une des sources majeures d'information sur les climats passés

Méthodes

L'étude des flores et des faunes fossiles en tant que paléoformes a est à l'origine de la paléoclimatologie, et en reste la base principale. La géochimie et les analyses isotopiques y jouent aussi un rôle croissant, de même que la modélisation informatique.

Différents paramètres, d'origine externe au système climatique, sont à l'origine des variations climatiques (notion de forçage (forçage radiatif) ajoutant ses effets à ceux du forçage volcanique et à ceux ayant pour origine l'expansion et l'évolution de la vie (production d'oxygène, albédo modifiée par la couverture végétale, etc.).

Les variations d'insolation liées aux paramètres de l'orbite de la Terre (théorie astronomique des paléoclimats) sont l'un des forçages que les modèles doivent prendre en compte (pouvant être facilement reliées à des observations géologiques).



www.amossys.fr

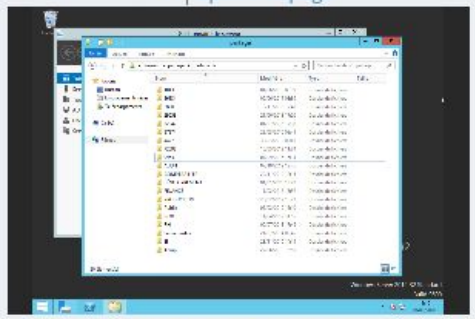
Le réalisme des comportements

- Réalisation de comportements habituels
  - Identification de liens dans les pages web, et clic sur ces liens
  - Notion de taux d'erreurs de frappes clavier (revenir en arrière, mettre le mot de passe à la place du login, . . .)
  - Mouvements de souris intégrant des tics d'usage habituels (réalisation machinale de courbes, surlignage de morceaux de texte sans réfléchir, . . .)
  - Zoom en arrivant sur une page (action classique réalisé par un humain dès lors que le site web est présenté dans une taille de texte trop petite)
  - Scroll haut/bas pendant lecture d'un document
  - Copier-coller de texte dans la barre de recherche
  - Sélection d'un mot dans un texte

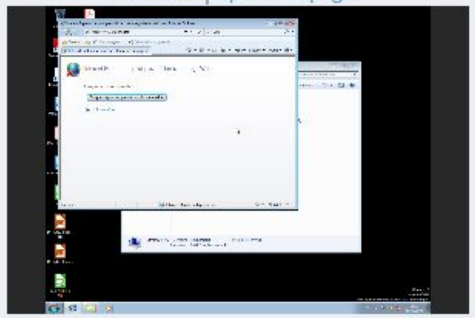
Simulation de vie sur l'ensemble du parc

AMOSSYS

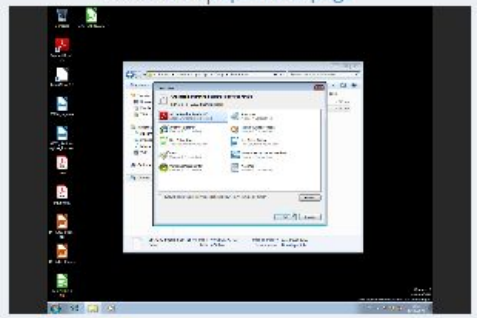
- AD1 | Open VNC page



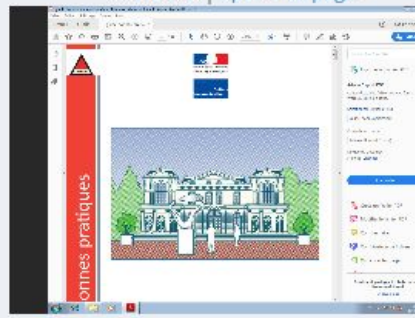
- CLIENT1 | Open VNC page



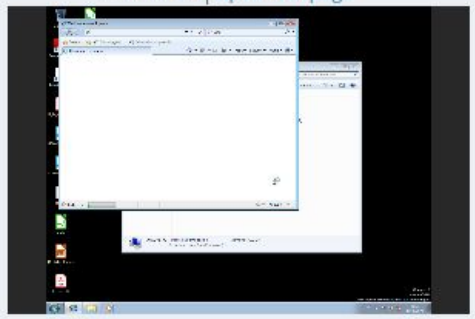
- CLIENT2 | Open VNC page



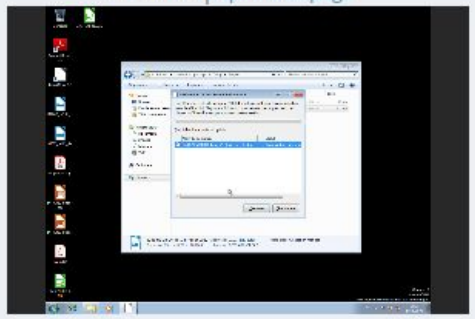
- CLIENT3 | Open VNC page



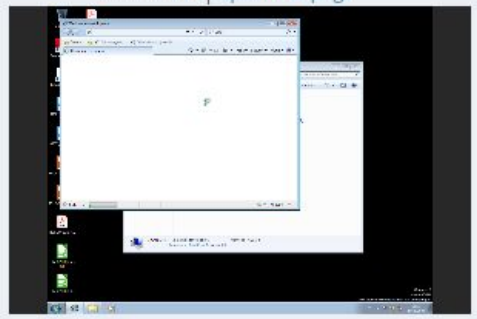
- CLIENT4 | Open VNC page



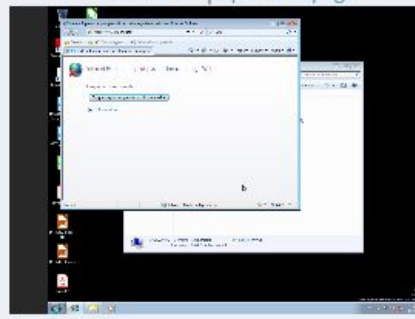
- CLIENT5 | Open VNC page



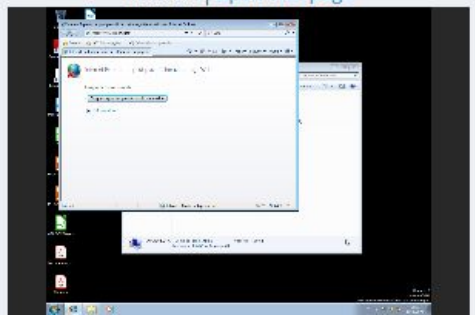
- CLIENT6 | Open VNC page



- CLIENT7 | Open VNC page

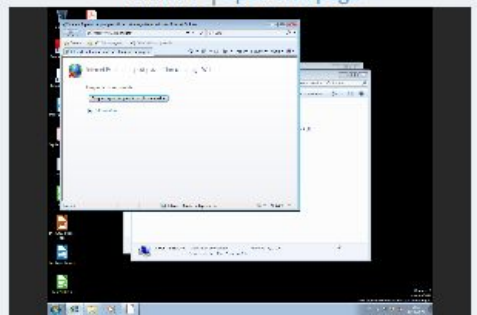


- CLIENT8 | Open VNC page

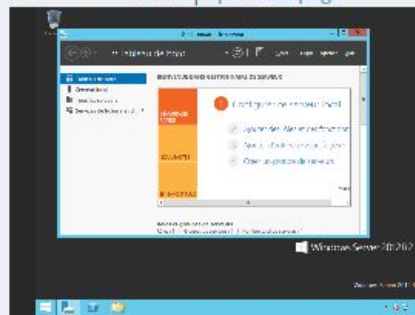


- CLIENT9 | Open VNC page

- CLIENT10 | Open VNC page



- FILE1 | Open VNC page



Simulation de vie réaliste

[DEMO]

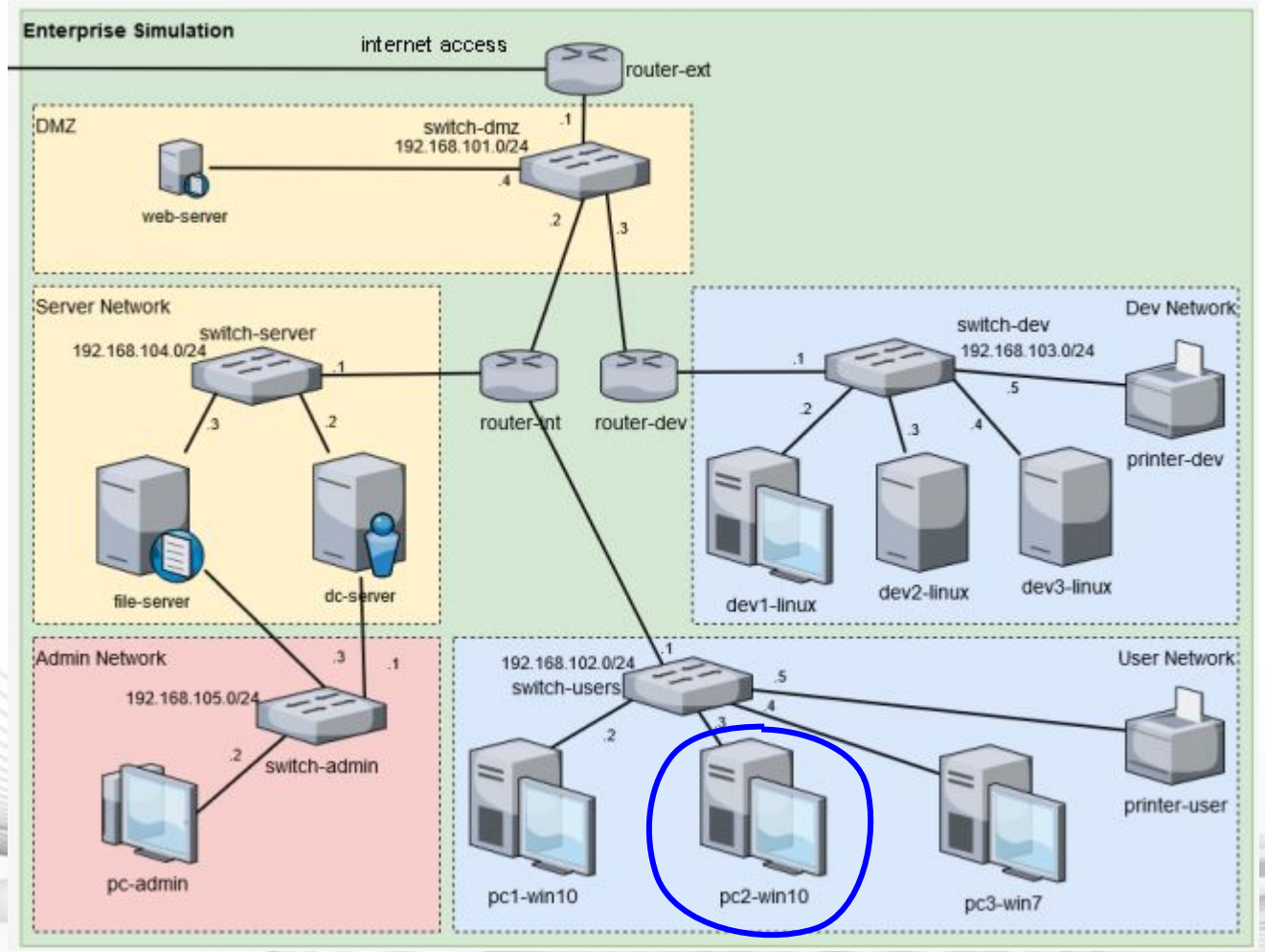
www.amossys.fr

AMOSSYS
EXPERT IN CYBERSECURITY

Extraction des traces d'intérêts pour la production de renseignements

www.amossys.fr

Scénario d'une détonation



1. détonation

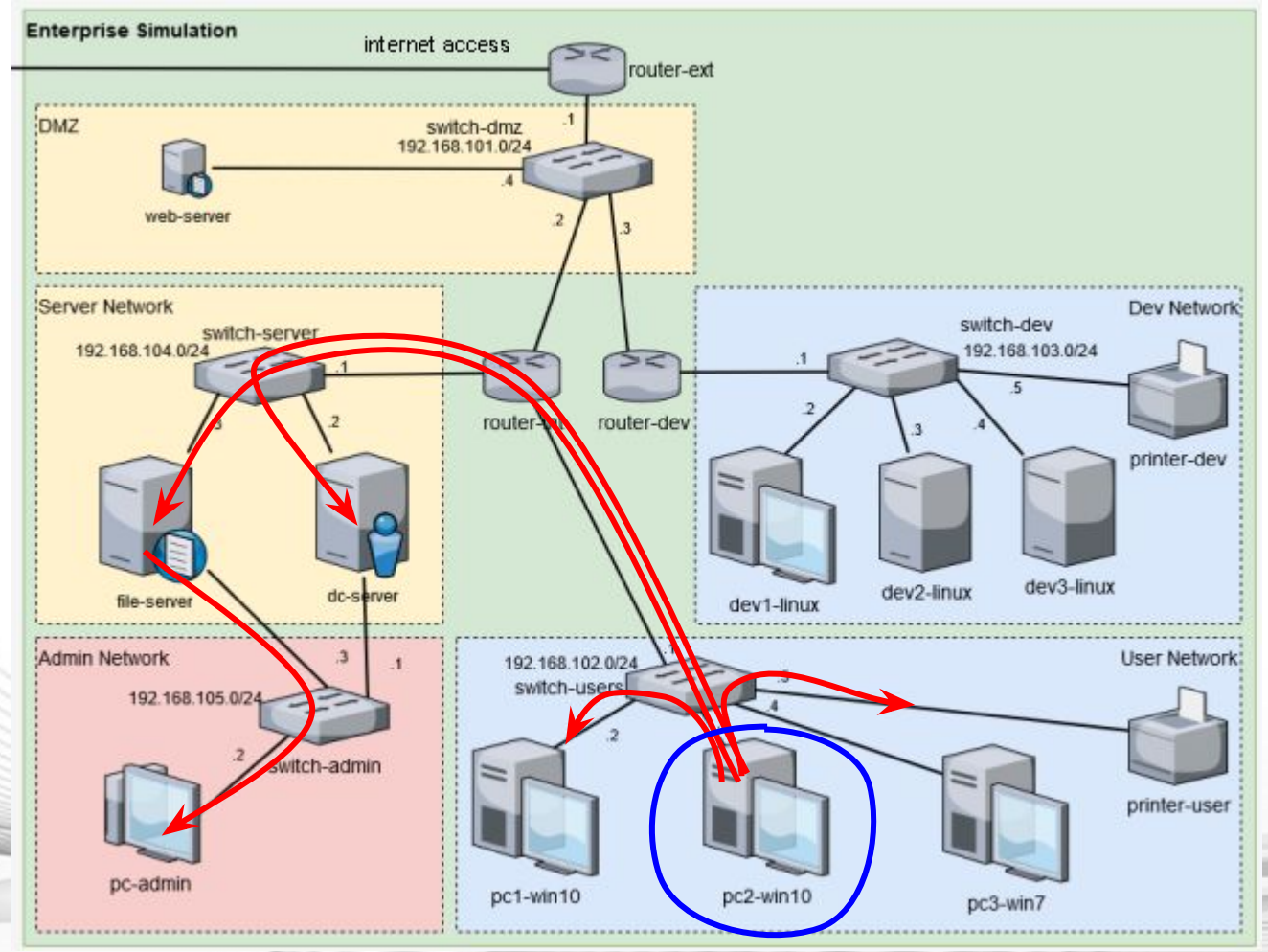
- PJ mail
- lien mail
- exécutable
- archive
- document
- ...

www.amossys.fr

Scénario d'une détonation

2. l'attaquant déroule sa kill chain sur le SI

- exécution
- reconnaissance
- privesc
- persistence
- rebond latéral
- compromission assets critiques
- impact
- ...



1. détonation

- PJ mail
- lien mail
- exécutable
- archive
- document
- ...

www.amossys.fr

Extraction des traces d'intérêts pour la production de renseignements

Processus construit autour de trois étapes

1. Production des métadonnées des actions de vie
2. Production du dataset des traces système et réseau
3. Extraction des traces d'attaque

www.amossys.fr

1. Production des métadonnées des actions de vie

- Le moteur de génération d'actions de vie est en mesure de produire une scénario séquentiel maîtrisé
- Scénario produit sur chaque poste client (Windows)
- Connaissance des timestamps de début et de fin de chaque action
- Connaissance des données de paramétrage de chaque action (login, applications lancées, fichiers écrits, …)


```
{
  "name": "open_session",
  "parameters": {
    "login": "rguillois"
  },
  "start_time": "2020-12-05 16:35:18.199",
  "end_time": "2020-12-05 16:35:31.614"
},
{
  "name": "open_app",
  "parameters": {
    "name": "firefox"
  },
  "start_time": "2020-12-05 16:35:51.163",
  "end_time": "2020-12-05 16:36:09.679"
},
{
  "name": "close_app",
  "parameters": {
    "name": "firefox"
  },
  "start_time": "2020-12-05 16:37:09.820",
  "end_time": "2020-12-05 16:37:28.469"
},
{
  "name": "open_app",
  "parameters": {
    "name": "evolution"
  },
  "start_time": "2020-12-05 16:37:28.474",
  "end_time": "2020-12-05 16:37:53.976"
},
{
  "name": "open_last_message",
  "start_time": "2020-12-05 16:37:53.979",
  "end_time": "2020-12-05 16:37:55.341"
},
{
  "name": "reply_message",
  "start_time": "2020-12-05 16:37:55.379",
  "end_time": "2020-12-05 16:38:27.241"
},
{
  "name": "send_message",
  "start_time": "2020-12-05 16:38:27.244",
  "end_time": "2020-12-05 16:38:32.322"
}
```


Métadonnées des actions de vie en JSON

www.amossys.fr

1. Production des métadonnées des actions de vie

- Le moteur de génération d'actions de vie est en mesure de produire une scénario séquentiel maîtrisé
- Scénario produit sur chaque poste client (Windows)
- Connaissance des timestamps de début et de fin de chaque action
- Connaissance des données de paramétrage de chaque action (login, applications lancées, fichiers écrits, …)
- **Actions de vie spécifiques pour la détonation (ouverture de PJ, clic sur URL, exécution d'un binaire, ...)**


```
{
  "name": "open_session",
  "parameters": {
    "login": "rguillois"
  },
  "start_time": "2020-12-05 16:35:18.199",
  "end_time": "2020-12-05 16:35:31.614"
},
{
  "name": "open_app",
  "parameters": {
    "name": "firefox"
  },
  "start_time": "2020-12-05 16:35:51.163",
  "end_time": "2020-12-05 16:36:09.679"
},
{
  "name": "close_app",
  "parameters": {
    "name": "firefox"
  },
  "start_time": "2020-12-05 16:37:09.820",
  "end_time": "2020-12-05 16:37:28.469"
},
{
  "name": "open_app",
  "parameters": {
    "name": "evolution"
  },
  "start_time": "2020-12-05 16:37:28.474",
  "end_time": "2020-12-05 16:37:53.976"
},
{
  "name": "open_last_message",
  "start_time": "2020-12-05 16:37:53.979",
  "end_time": "2020-12-05 16:37:55.341"
},
{
  "name": "reply_message",
  "start_time": "2020-12-05 16:37:55.379",
  "end_time": "2020-12-05 16:38:27.241"
},
{
  "name": "send_message",
  "start_time": "2020-12-05 16:38:27.244",
  "end_time": "2020-12-05 16:38:32.322"
}
```


Métadonnées des actions de vie en JSON

www.amossys.fr

2. Production du dataset des traces système et réseau

- ❑ Sysmon déployé sur chaque poste
- ❑ Collecte des journaux Events Logs Windows suivants
  - ❑ Application/*
  - ❑ System/*
  - ❑ Security/*
  - ❑ Microsoft-Windows-Sysmon/Operational/*
  - ❑ Microsoft-Windows-Windows Defender/Operational/*

Types d'événements pris en compte

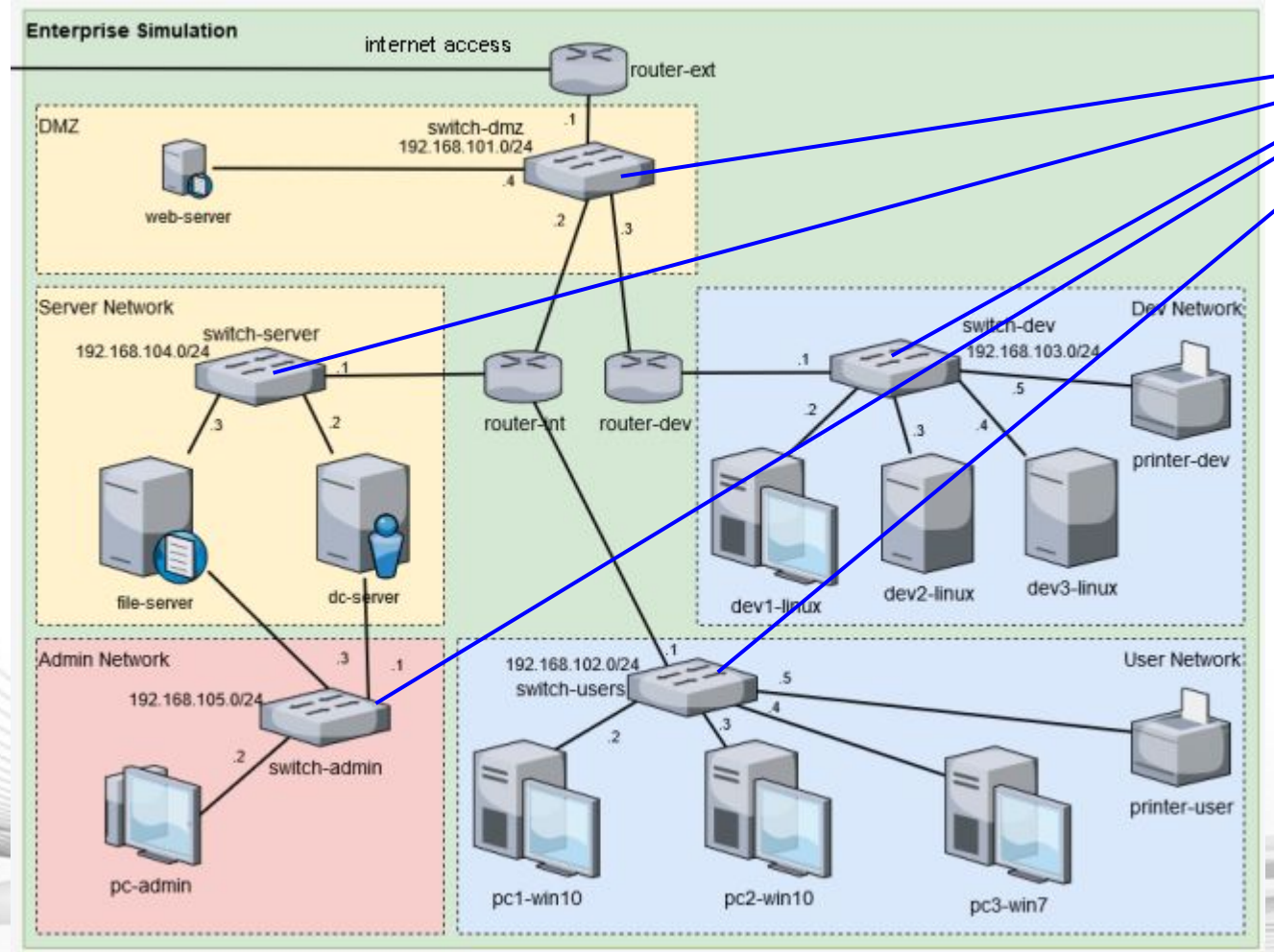
- ouverture/fermeture de session
- accès distants
- création de processus
- chargement de DLL
- création de thread distant
- accès au système de fichiers
- résolutions DNS
- écriture dans la base de registre
- alertes de sécurité
- ...

www.amossys.fr

2. Production du dataset des traces système et réseau

2. l'attaquant déroule sa kill chain sur le SI

- exécution
- reconnaissance
- privesc
- persistence
- rebond latéral
- compromission assets critiques
- impact
- ...



PCAPs

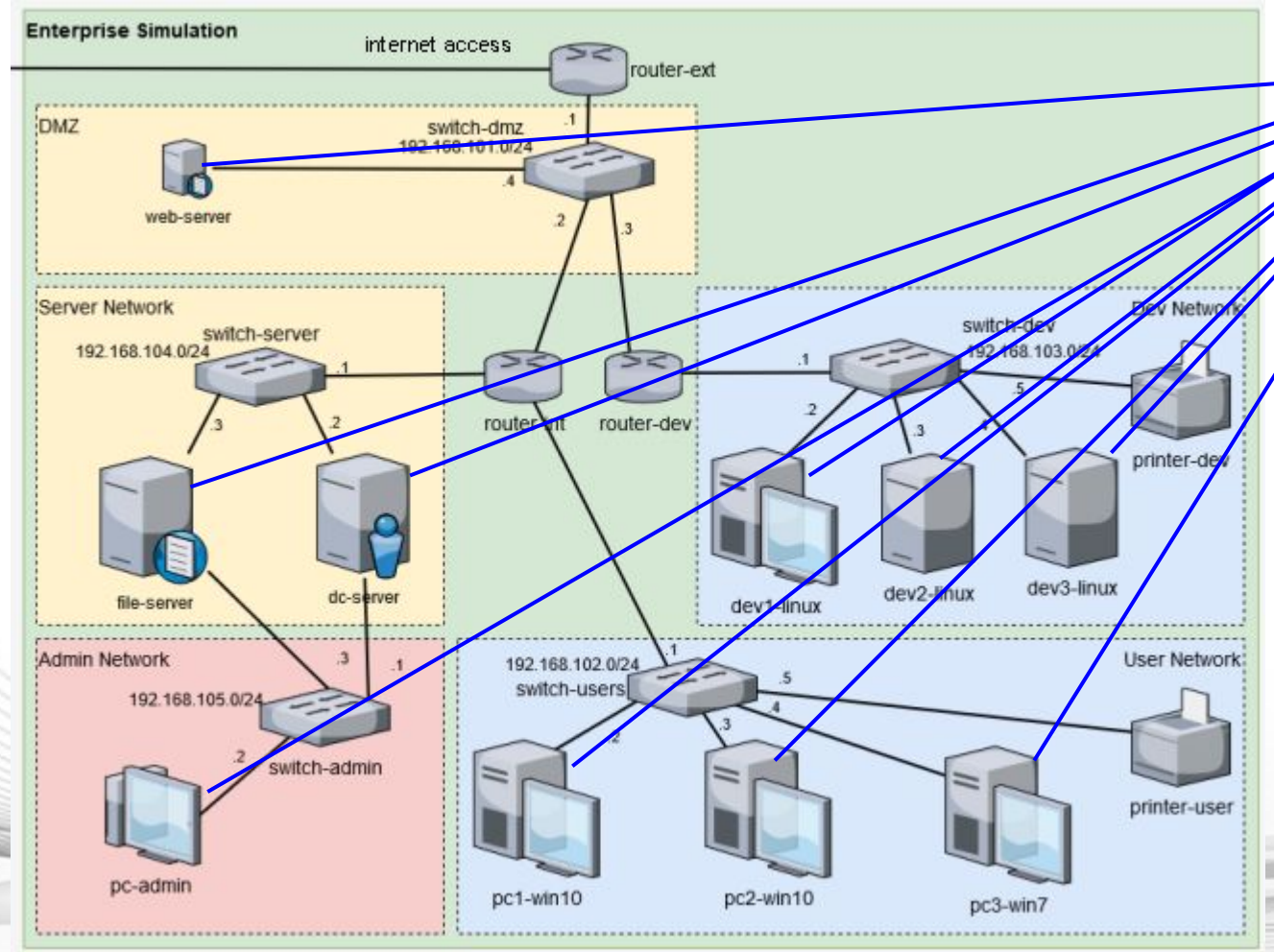
3. capture des traces réseau

www.amossys.fr

2. Production du dataset des traces système et réseau

2. l'attaquant déroule sa kill chain sur le SI

- exécution
- reconnaissance
- privesc
- persistence
- rebond latéral
- compromission assets critiques
- impact
- ...



Puit de logs

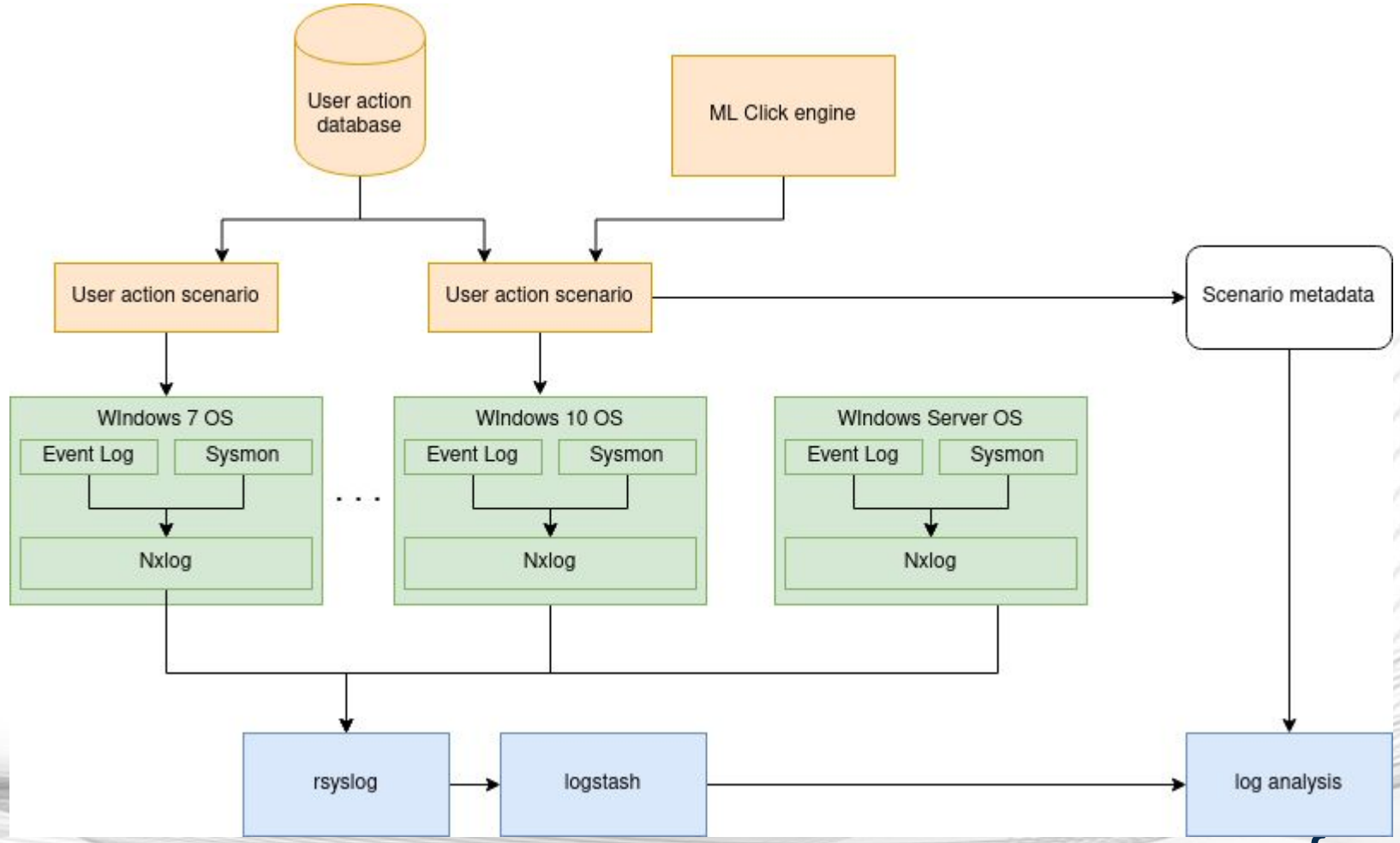
3. capture des traces réseau

4. capture des traces système

www.amossys.fr

2. Production du dataset des traces système et réseau

- Journaux collectés sur un puit de logs (rsyslog)
- Evènements normalisés en JSON avec logstash, avant analyse



www.amossys.fr

3. Extraction des traces d'attaque

Objectif : séparer les logs liés à l'activité de vie des logs liés à l'attaque

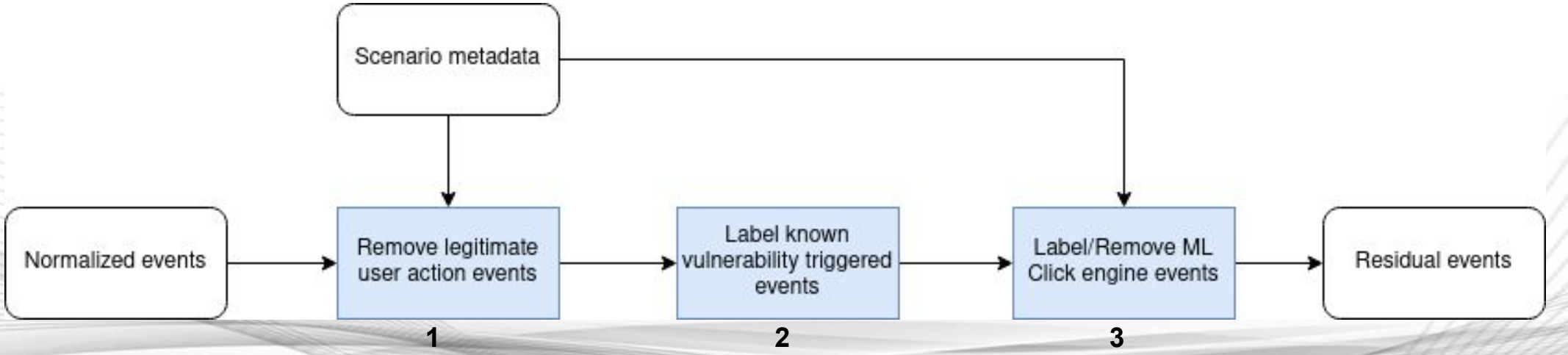
Hypothèse : cette séparation est possible car les actions de vie sont maîtrisées

3. Extraction des traces d'attaque

Objectif : séparer les logs liés à l'activité de vie des logs liés à l'attaque

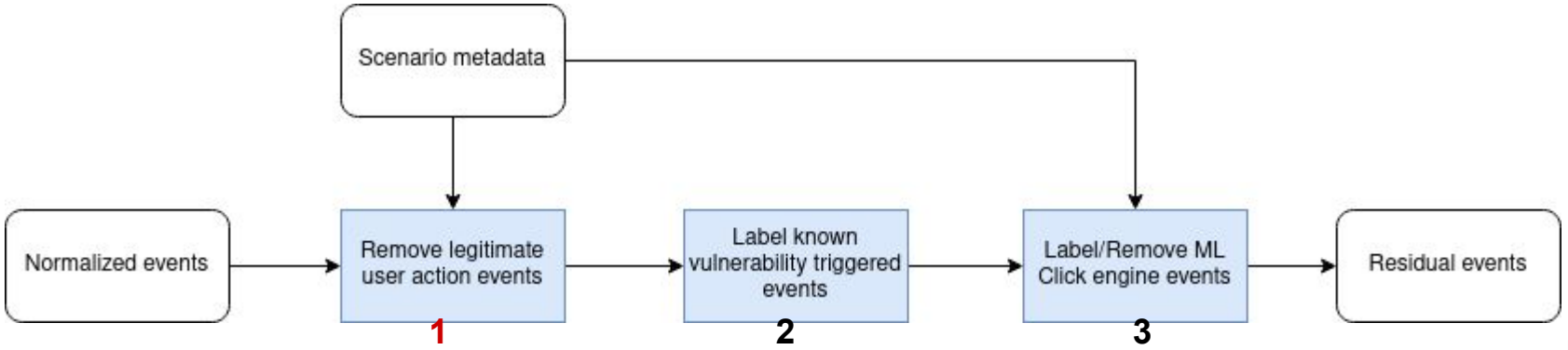
Hypothèse : cette séparation est possible car les actions de vie sont maîtrisées

Phase d'analyse en trois étapes



www.amossys.fr

3. Extraction des traces d'attaque

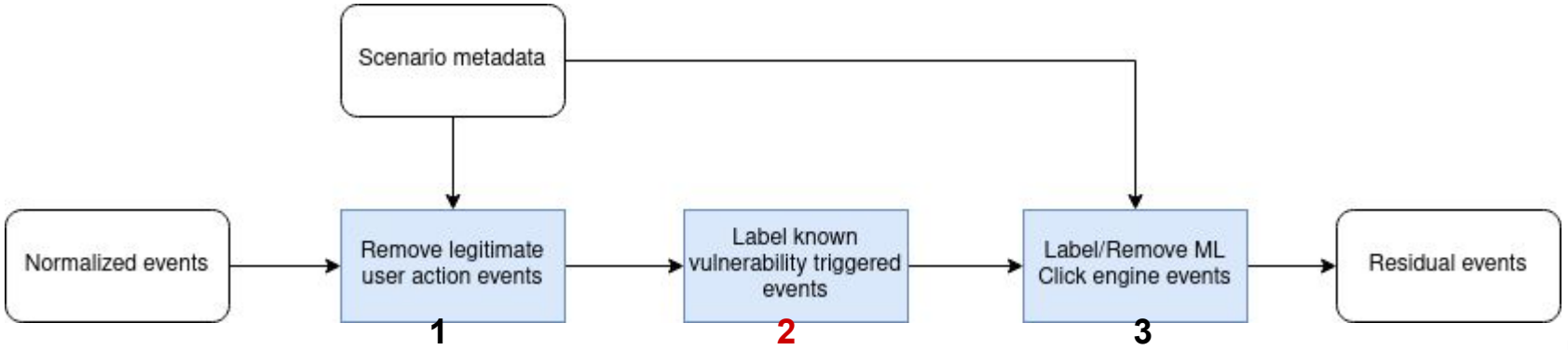


Etape 1 : suppression des logs légitimes liés aux scénarios de vie joués

- ❑ Labellisation des logs liés aux actions de vie du catalogue, puis suppression de ces logs
- ❑ Usage des timestamp de début et de fin et du paramétrage de chaque action
- ❑ Usage de la connaissance à priori des logs produits par chaque action, lors d'une phase d'apprentissage en amont
  - ❑ Lors de cette phase d'apprentissage, les évènements des services systèmes et de l'OS sont également traités
- ❑ Suppression des logs labellisés

www.amossys.fr

3. Extraction des traces d'attaque

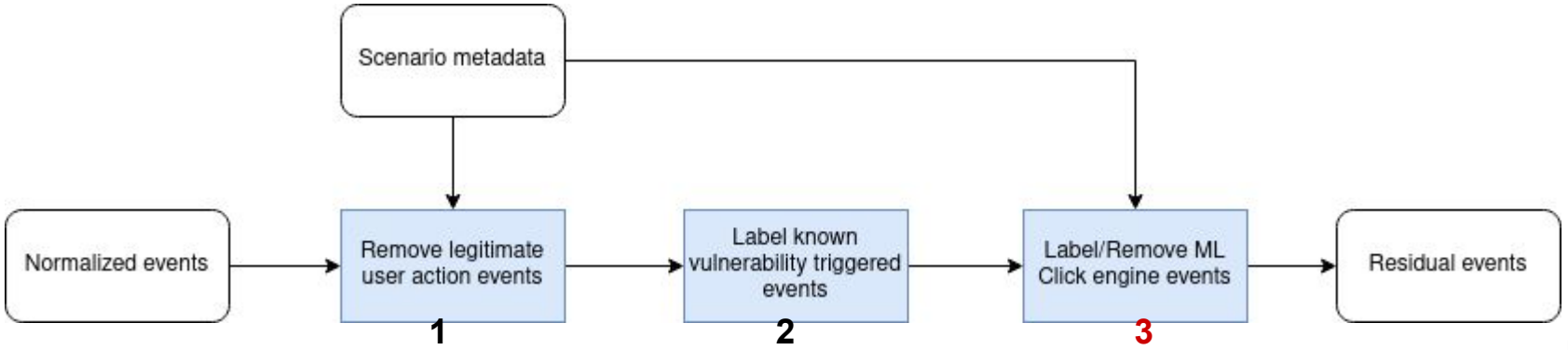


Etape 2 : identification de l'exploitation des vulnérabilités et faiblesses connues

- Labellisation des logs liés à l'exploitation des vulnérabilités et faiblesses volontairement déployées sur le SI
- Exemples :
  - objets Active Directory accédés
  - fichiers leurres accédés
  - alertes AV, NIDS connues
  - ...

www.amossys.fr

3. Extraction des traces d'attaque

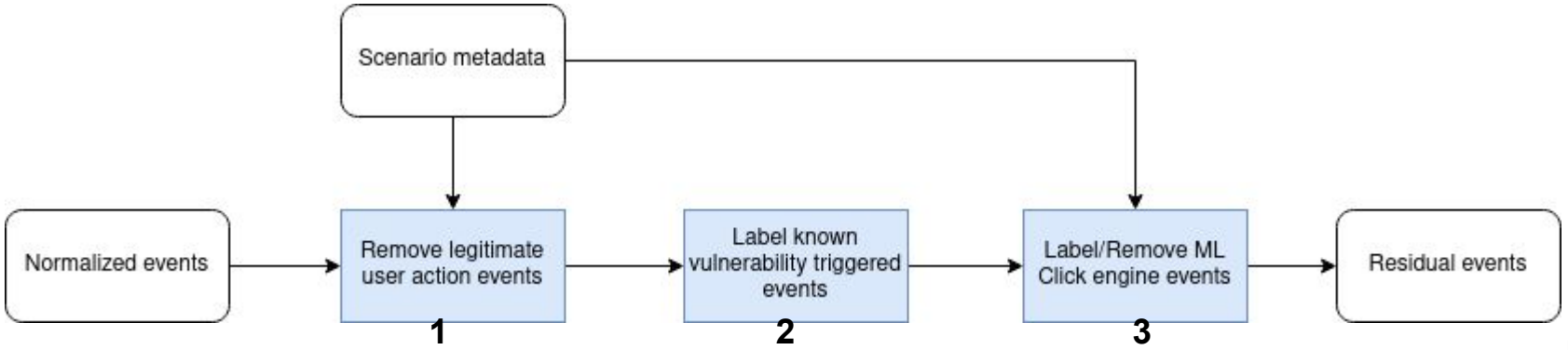


Etape 3 : suppression des logs légitimes liés au moteur d'apprentissage

- ❏ Labellisation des logs liés aux actions du moteur d'apprentissage
- ❏ Suppression des logs labellisés, suivant un niveau de confiance
  - ❏ Seuil reposant sur le délai avec le timestamp de l'action de référence (mouvement de souris, frappe clavier)

www.amossys.fr

3. Extraction des traces d'attaque



En sortie

- Ensemble de traces systèmes et réseau résiduelles correspondant à une sur-approximation des actions d'attaque
- L'analyste peut alors prendre la main est affiner le résultat
  - Identification des TTP sur la base des évènements labellisés avec les vulnérabilités/faiblesses connues
  - Investigation pour comprendre le mode opératoire complet

www.amossys.fr

Résultats

```
2020-09-22 18:23:17: gmichel-pc - C:\Windows\System32\smss.exe -> C:\Windows\System32\winlogon.exe
2020-09-22 18:23:17: gmichel-pc - C:\Windows\System32\winlogon.exe -> C:\Windows\System32\LogonUI.exe
2020-09-22 18:24:05: gmichel-pc - C:\Windows\System32\services.exe -> C:\Windows\System32\taskhost.exe
2020-09-22 18:24:05: gmichel-pc - C:\Windows\System32\winlogon.exe -> C:\Windows\System32\slui.exe
2020-09-22 18:24:06: gmichel-pc - C:\Windows\System32\winlogon.exe -> C:\Windows\System32\userinit.exe
2020-09-22 18:24:06: gmichel-pc - C:\Windows\System32\userinit.exe -> C:\Windows\explorer.exe
2020-09-22 18:27:35: gmichel-pc - C:\Windows\explorer.exe -> C:\Windows\explorer.exe
2020-09-22 18:27:49: gmichel-pc - C:\Windows\explorer.exe -> C:\Program Files\LibreOffice\program\scalc.exe
2020-09-22 18:27:49: gmichel-pc - C:\Program Files\LibreOffice\program\scalc.exe -> C:\Program Files\LibreOffice\program\soffice.exe
2020-09-22 18:27:49: gmichel-pc - C:\Program Files\LibreOffice\program\soffice.exe -> C:\Program Files\LibreOffice\program\soffice.bin
2020-09-22 18:30:44: gmichel-pc - C:\Windows\explorer.exe -> C:\Program Files (x86)\Internet Explorer\iexplore.exe
2020-09-22 18:31:15: gmichel-pc - C:\Program Files (x86)\Internet Explorer\iexplore.exe -> C:\Windows\System32\cmd.exe
2020-09-22 18:31:17: gmichel-pc - C:\Windows\System32\cmd.exe -> C:\Users\gmichel\AppData\Local\Temp\pesec.exe
```

Résultats

Logs système sur ouverture de session

Logs de vie

Logs résiduels

```
2020-09-22 18:23:17: gmichel-pc - C:\Windows\System32\smss.exe -> C:\Windows\System32\winlogon.exe
2020-09-22 18:23:17: gmichel-pc - C:\Windows\System32\winlogon.exe -> C:\Windows\System32\LogonUI.exe
2020-09-22 18:24:05: gmichel-pc - C:\Windows\System32\services.exe -> C:\Windows\System32\taskhost.exe
2020-09-22 18:24:05: gmichel-pc - C:\Windows\System32\winlogon.exe -> C:\Windows\System32\slui.exe
2020-09-22 18:24:06: gmichel-pc - C:\Windows\System32\winlogon.exe -> C:\Windows\System32\userinit.exe
2020-09-22 18:24:06: gmichel-pc - C:\Windows\System32\userinit.exe -> C:\Windows\explorer.exe
2020-09-22 18:27:35: gmichel-pc - C:\Windows\explorer.exe -> C:\Windows\explorer.exe
2020-09-22 18:27:49: gmichel-pc - C:\Windows\explorer.exe -> C:\Program Files\LibreOffice\program\scalc.exe
2020-09-22 18:27:49: gmichel-pc - C:\Program Files\LibreOffice\program\scalc.exe -> C:\Program Files\LibreOffice\program\soffice.exe
2020-09-22 18:27:49: gmichel-pc - C:\Program Files\LibreOffice\program\soffice.exe -> C:\Program Files\LibreOffice\program\soffice.bin
2020-09-22 18:30:44: gmichel-pc - C:\Windows\explorer.exe -> C:\Program Files (x86)\Internet Explorer\iexplore.exe
2020-09-22 18:31:15: gmichel-pc - C:\Program Files (x86)\Internet Explorer\iexplore.exe -> C:\Windows\System32\cmd.exe
2020-09-22 18:31:17: gmichel-pc - C:\Windows\System32\cmd.exe -> C:\Users\gmichel\AppData\Local\Temp\psexec.exe
```

AMOSSYS
EXPERT IN CYBERSECURITY

Conclusion et perspectives

Ce qu'il faut retenir

- Importance du caractère réaliste d'un environnement honeynet, notamment dans un cas d'usage détonation
  - Ce réalisme sert à exposer les techniques et modes opératoires (TTP) des attaquants
- Avec la plateforme BEEZH, il est possible de construire de manière automatisée un environnement SI crédible
- Sur cette base, nous avons proposé deux contributions :
  - Simulation de comportements utilisateur réalistes
  - Procédé permettant d'extraire une approximation des traces d'attaque
    - Objectif de faciliter le travail des analystes pour la compréhension du mode opératoire

www.amossys.fr

Perspectives

- Améliorations techniques de la plateforme, dont notamment
  - la furtivité de la virtualisation
  - le passage à l'échelle
- Meilleure prise en compte des événements réseau dans le processus d'extraction des traces d'intérêts
- Ouverture vers une plateforme Deceptive
  - Rapprochement envisagé entre la capacité de honeynet de la plateforme BEEZH et son utilisation comme outil de détection sur un SI opérationnel
  - Levée d'alerte auprès d'un SIEM lorsqu'un attaquant pénètre sur le périmètre du honeynet
- Sur ce sujet, plusieurs problématiques se posent, dont notamment la difficulté d'intégration du SI simulé au sein d'un parc de grande dimension

www.amossys.fr

Perspectives

Au delà du cas d'usage de la détonation, l'intérêt à produire de l'activité utilisateur réaliste est multiple

- Test d'efficacité de produits de sécurité
  - Création de datasets labellisés avec actions de vie et d'attaque
  - Cibles : sondes hôte et réseau, EDR, SIEM, ...
- Entraînement et formation
  - Environnement immersif avec comportements de vie et d'attaque
  - Formations Blue Team, Red Team, …
- ...

Merci...

Et si vous avez des questions => **frederic.guihery@amossys.fr**

www.amossys.fr