Sécurité système

Rémi Dubois • STI 2011

v.20221019

$ whoami

Rémi Dubois

- Ingénieur diplômé STI – ENSIB 2011
- Titre ESSI – ANSSI 2019
- Ingénieur SSI – cybersécurité & Responsable de SOC

Plan des cours

1. Principes généraux
2. Installer un système sécurisé
3. Cloisonnement et isolation des processus
4. Journalisation et centralisation
5. Travaux dirigés

Sécurité système

Principes généraux

Plan

1. Moindre privilèges
2. Défense en profondeur

Propriétés de sécurité recherchées

Il s'agit des 4 propriétés classiques :

- Confidentialité
- Intégrité
- Disponibilité
- Traçabilité / auditabilité

Moindre privilèges

Moindre privilèges

Définition

Ne disposer que des droits strictement nécessaires à l'exécution, et rien de plus.

Corollaires :

- Minimalisation (KISS) : réduction de la surface d'attaque
  - N'installer que ce qui est nécessaire
  - N'activer et n'exposer que ce qui est nécessaire
- Restriction au besoin d'en connaître
  Un des deux principes du classifié de défense

Moindre privilèges

Application

- Ne pas utiliser un compte d'administration au quotidien
  `Sudo`er ou `Administrateur`, ou pire, `root` ou `SYSTEM`
- Utiliser un poste dédié à l'administration
- Configurer précisément et durcir le système
  Montage des partitions, AppArmor, SELinux...
- Diviser, cloisonner et isoler les services

Défense en profondeur

Défense en profondeur

Définition

> *défense globale et dynamique, coordonnant plusieurs lignes de défense couvrant toute la profondeur du système.*
>
> Mémento de la DCSSI, 2004

- Principe de construction des forteresses
- Chaque ligne de défense ralentit l'adversaire
- Classique en matière industrielle (multiples « barrières »)

Défense en profondeur

Application

- Appréhender l'ensemble du système
- Son environnement (physique et réseau)
- Son matériel
- Son noyau et couches basses
- Ses logiciels et couches hautes
- Sa maintenance

Défense en profondeur

Application

- Filtrer les flux réseau globalement **et** localement
- Diversifier : utiliser plusieurs types de pare-feu, moteurs antiviraux, logiciels socles (DNS, mail…)
- Journaliser : la confiance n'exclue pas le contrôle
- Être à jour : installer systématiquement les correctifs de sécurité
- Un peu de paranoïa : **tout** ce qui n'est pas normal **est** anormal
- Ne jamais faire confiance à l'utilisateur

Défense en profondeur

L'attaquant peut-il ?

- Accéder sans limite au matériel ?
  Matériel sans surveillance et sans verrouillage :
  - douanes dans certains pays
  - casier à l'entrée d'une salle de réunion
- Analyser *a posteriori* le contenu du support de stockage ?
  Disque dur accessible sans chiffrement
- Maîtriser la chaîne de démarrage ?
  SecureBoot, options du noyau modifiables
- Injecter du code dans le noyau ?
  Drivers

Questions ?

Sécurité système

Installer un système sécurisé

Plan

1. Orientations techniques
2. Partitionnement
3. Configuration noyau (`sysctl`)

Défense en profondeur

Choisir le système en fonction :

- des services qu'il va héberger
  Linux = polyvalent mais gros, *BSD plus minimaliste ?
  Disponibilité des fonctionnalités souhaitées ?
- du besoin d'avoir un support commercial (RHEL, SLES...)
  parfois nécessaire pour le support des logiciels (Oracle...)
- de la durée d'utilisation et la capacité de le maintenir à jour
  RIP CentOS et bof pour *BTW I use Arch*
- de la capacité à l'administrer !

Orientations techniques

Orientations techniques

Configuration matérielle

- Protéger le BIOS / UEFI
  Mot de passe administrateur ?
- Privilégier UEFI : SecureBoot
- Mode 64 bits
- Bit VT-x / AMD V uniquement si virtualisation nécessaire

Orientations techniques

Configuration système

- Partitionner finement
- N'installer que le **strict minimum** nécessaire
  Pas d'interface graphique sur un serveur !
- Durcir les options noyau (`sysctl`)
- Durcir le système (règles `sudoers` granulaires…)
- Durcir les services (empêcher les connexions `root` sur SSH…)
- Filtrer les accès (Netfilter `iptables` / `nftables`)
- Mot de passe `root` complexe et spécifique à ce système
  Au coffre-fort ou dans un gestionnaire de mot de passe

Orientations techniques

Chiffrement de disque

- Utiliser un secret distinct (*passphrase* ou *token*)
  LUKS / LUKS2 (Linux) ou BitLocker (Windows **Pro**)
- Impératif pour une **station nomade** !
- Envisageable pour une station fixe
  Permet d'éviter de disséminer des documents sensibles lors de la mise au rebus
- Plus discutable pour un serveur
  Comment saisir la *passphrase* ?
- Prévoir un clé de séquestre (accident, réquisition judiciaire...)
  Au coffre-fort ou dans un gestionnaire de mot de passe

Partitionnement

Partitionnement I

Point de montage	Options	Description
`/`	<sans option>	Racine, contient le reste
`/boot`	`nosuid,nodev,noexec` (`noauto` optionnel)	Contient le noyau et le bootloader, pas d'accès après démarrage sauf mise à jour
`/tmp`	`nosuid,nodev,noexec`	Fichiers temporaires, ne doit pas contenir d'exécutables. Nettoyé après redémarrage, préférer du `tmpfs`
`/proc`	`hidepid=2`	Informations sur les processus et le système

Partitionnement II

Point de montage	Options	Description
`/home`	`nosuid,nodev,noexec` (`ro` optionnel)	Répertoires utilisateurs. `ro` si non utilisé. `noexec` à discuter sur station de travail
`/usr`	`nodev`	Majorité des utilitaires (exécutables et bibliothèques)
`/var`	`nosuid,nodev,noexec`	Fichiers variables pendant la vie du système (mails, BDD...)
`/var/log`	`nosuid,nodev,noexec`	Journaux du système
`/var/tmp`	`nosuid,nodev,noexec`	Fichiers temporaires conservés après extinction

Partitionnement III

Point de montage	Options	Description
`/opt`	`nosuid,nodev` (`ro` optionnel)	Packages additionnels. `ro` si non utilisé
`/srv`	`nosuid,nodev` (`noexec,ro` optionnels)	Fichiers servis par un service

D'après les « Recommandations de sécurité relatives à un système GNU/Linux », ANSSI, www.ssi.gouv.fr/reco-securite-systeme-linux/

Nota : utiliser correctement le système de fichier et se référer au *Filesystem Hierarchy Standard*, `hier(7)`

Partitionnement

- `/var/log` peut vous exploser à la figurer et bloquer le démarrage d'applications (comme X.org)
- `nodev` empêche de jouer avec les **fichiers spéciaux** pour accéder sans droit ni titre à une partition autrement protégée (avec `mknod` utilisé par exemple sur une clé USB) :

```
brw-rw---- 1 root disk     8,   0 18 mars  16:29 sda
brw-rw---- 1 root disk     8,   1 18 mars  16:29 sda1
brw-rw---- 1 root disk     8,   2 18 mars  16:29 sda2
brw-rw---- 1 root disk     8,   3 18 mars  16:29 sda3
```

Configuration noyau (`sysctl`)

Configuration noyau (sysctl)

Se référer aux recommandations GNU/Linux de l'ANSSI, www.ssi.gouv.fr/reco-securite-systeme-linux/ :

- pour la pile réseau, durcir les configuration IPv4 et IPv6 face à des paquets peu commun ou non standards
  ANSSI : serveur sans routage et avec adresse IPv6 statique
  Pas d'IPv6 : `net.ipv6.conf.all.disable_ipv6 = 1`
- pour le reste du système : paramétrage de la mémoire, des processus, du système de fichiers...

Configuration noyau (sysctl) - Options réseau I

```
# Pas de routage entre les interfaces
net.ipv4.ip_forward = 0
# Filtrage par chemin inverse
net.ipv4.conf.all.rp_filter = 1
net.ipv4.conf.default.rp_filter = 1
# Ne pas envoyer de redirections ICMP
net.ipv4.conf.all.send_redirects = 0
net.ipv4.conf.default.send_redirects = 0
# Refuser les paquets de source routing
net.ipv4.conf.all.accept_source_route = 0
net.ipv4.conf.default.accept_source_route = 0
# Ne pas accepter les ICMP de type redirect
net.ipv4.conf.all.accept_redirects = 0
net.ipv4.conf.all.secure_redirects = 0
net.ipv4.conf.default.accept_redirects = 0
net.ipv4.conf.default.secure_redirects = 0
# Loguer les paquets ayant des IPs anormales
net.ipv4.conf.all.log_martians = 1
```

Configuration noyau (sysctl) - Options réseau II

```
# RFC 1337
net.ipv4.tcp_rfc1337 = 1
# Ignorer les réponses non conformes à la RFC 1122
net.ipv4. icmp_ignore_bogus_error_responses = 1
# Augmenter la plage pour les ports éphémères
net.ipv4.ip_local_port_range = 32768 65535
# Utiliser les SYN cookies
net.ipv4.tcp_syncookies = 1
# Désactiver le support des "router solicitations"
net.ipv6.conf.all.router_solicitations = 0
net.ipv6.conf.default.router_solicitations = 0
# Ne pas accepter les "router preferences" par "router
   advertisements"
net.ipv6.conf.all.accept_ra_rtr_pref = 0
net.ipv6.conf.default.accept_ra_rtr_pref = 0
# Pas de configuration auto des prefix par "router advertisements"
net.ipv6.conf.all.accept_ra_pinfo = 0
net.ipv6.conf.default.accept_ra_pinfo = 0
```

Configuration noyau (sysctl) - Options réseau III

```
# Pas d'apprentissage du routeur par défaut par "router
   advertisements"
net.ipv6.conf.all.accept_ra_defrtr = 0
net.ipv6.conf.default.accept_ra_defrtr = 0
# Pas de configuration auto des adresses à partir des "router
   advertisements"
net.ipv6.conf.all.autoconf = 0
net.ipv6.conf.default.autoconf = 0
# Ne pas accepter les ICMP de type redirect
net.ipv6.conf.all.accept_redirects = 0
net.ipv6.conf.default.accept_redirects = 0
# Refuser les packets de source routing
net.ipv6.conf.all.accept_source_route = 0
net.ipv6.conf.default.accept_source_route = 0
# Nombre maximal d'adresses autoconfigurées par interface
net.ipv6.conf.all.max_addresses = 1
net.ipv6.conf.default.max_addresses = 1
```

Configuration noyau (sysctl) - Options système I

```
# Désactivation des SysReq
kernel.sysrq = 0
# Pas de core dump des exécutables setuid
fs.suid_dumpable = 0
# Interdiction de déréférencer des liens vers des fichiers dont
# l'utilisateur courant n'est pas le propriétaire
# Peut empêcher certains programmes de fonctionner correctement
fs.protected_symlinks = 1
fs.protected_hardlinks = 1
# Activation de l'ASLR
kernel.randomize_va_space = 2
# Interdiction de mapper de la mémoire dans les adresses basses
   (0)
vm.mmap_min_addr = 65536
```

Configuration noyau (sysctl) - Options système II

```
# Espace de choix plus grand pour les valeurs de PID
kernel.pid_max = 65536
# Obfuscation des adresses mémoire kernel
kernel.kptr_restrict = 1
# Restriction d'accès au buffer dmesg
kernel.dmesg_restrict = 1
# Restreint l'utilisation du sous système perf
kernel.perf_event_paranoid = 2
kernel. perf_event_max_sample_rate = 1
kernel.perf_cpu_time_max_percent = 1
```

Questions ?

Sécurité système

Cloisonnement et isolation des processus

Plan

1. Contrôle d'accès UNIX (ACL)
2. Mécanismes anciens
3. Cgroups et *Namespaces* Linux
4. Conteneurs et virtualisation

Pourquoi cloisonner ?

- Séparer les fonctionnalités d'une application
- Isoler des actions potentielles dangereuses
- Limiter les privilèges requis (ports privilégiés...)
- Garder les secrets... secrets

Exemple : service de cachet serveur

Contrôle d’accès UNIX (ACL)

Contrôle d'accès UNIX (ACL)

Rappels

Système de contrôle d'accès discrétionnaire (DAC) où tout est fichier :

- Utilisateur propriétaire (`u`)
- Groupe propriétaire (`g`)
- Autres utilisateurs (`o`)
- Lecture (`r`)
- Écriture (`w`)
- Exécution / traversée (`x`)

Ne pas oublier sur Linux :

- ACL étendues (`getfacl` / `setfacl`)
- On peut ainsi (un peu) s'orienter vers du contrôle d'accès par rôles (RBAC)

Contrôle d'accès UNIX (ACL)

Concept d'utilisation

- Utilisateur spécifique par processus/fonctionnalité
- Groupe commun
- Contrôle plus fin avec des ACL étendues

Exemple : Serveur web (`httpd:http`) et PHP FPM (`php:http`)

```
srw-rw---- 1 httpd http    0  5 mars  16:21 /run/php-fpm.sock
-r-------- 1 php   http 3272 13 févr.  2021 /etc/myapp/privkey.pem
```

Contrôle d'accès UNIX (ACL)

Exemple : Répertoire de numérisation pour photocopieur

```
ls -lR /srv/partage
drwxrwx---+ 7 root secretariat 4096 15 janv.  2019 Secrétariat

/srv/partage/Secrétariat
total 38
drwxrwx-w-  2 root secretariat 4096 15 janv.  2019 Numérisation
```

```
getfacl /srv/partage/Secrétariat
# file: srv/partage/Secrétariat
# owner: root
# group: secretariat
user::rwx
group::rwx
other::---
user:photocopieur:--x
```

SELinux

DAC souffre de nombreuses limitations :

- les utilisateurs ont un contrôle complet de leurs ressources (CRUD + nouveaux droits)
- les programmes ont les mêmes droits que leur utilisateur

→ Si un programme en `UID 0` est compromis, le système entier est compromis

Security-Enhanced Linux apporte un système de contrôle d'accès mandataire (MAC) où le noyau applique une politique de contrôle complémentaire.

SELinux

SELinux utilise des contextes de sécurité composés :

- d'une identité (`_u`, propriétaire du contexte)
- d'un rôle (`_r`)
- d'un domaine pour les sujet (utilisateurs ou processus) ou d'un type pour les objets (ressources passives) (`_t`).

Le contexte d'un objet peut être vu avec `ls -Z` :

```
drwxr-xr-x root root system_u:object_r:etc_t  /etc
```

Écrire une politique de sécurité est **difficile**.
Certaines distributions appliquent par défaut une politique SELinux, comme RHEL qui se base sur le FHS.

Mécanismes anciens

Mécanismes anciens

`chroot` / `jail`

`chroot` (appel système et commande) simule une autre racine à l'intérieur du système de fichiers

- seuls les fichiers sous la nouvelle racine seront visibles
- nécessite de recréer une arborescence FHS (**`/`**, **`/dev`**, **`/proc`**, **`/sys`**…), et d'exposer les bibliothèques nécessaires
- nécessite des privilèges élevés pour être utilisé (ou **`CAP_SYS_CHROOT`**), mais utiliser **`root`** ne respecte pas le principe de moindre privilèges (→ **`setuid`**)
- *chroot is not and never has been a security tool* (Alan Cox)
- *This call […] is not intended to be used for any kind of security purpose* (**`chroot(2)`**)

Mécanismes anciens

`chroot` / `jail`

`jail` est une alternative plus robuste sous FreeBSD avec :

- un appel système
- un jeu de commandes (**`jail`**, **`jls`** et **`jexec`**)

Un **`jail`** a aussi son propre environnement réseau (nom d'hôte et adresse IP.)

Faute de mieux, on a le programme **`firejail`** sous Linux, utilisant les *namespaces* et *seccomp-bpf*.

Mécanismes anciens

`setuid` et `setgid`

Ces deux appels systèmes changent l'utilisateur et le groupe principal du processus :

- fait par la plupart des services *sérieux* lancés en **`root`**
- sinon, écrire un service systemd adapté (**`systemd.exec(5)`**)
- y dédier un compte de service spécifique
- y dédier un groupe spécifique à un ensemble de services (cf. exemple ACL)

Pour limiter les privilèges requis sur certaines actions sensibles, Linux propose les **`capabilities(7)`**

Mécanismes anciens

`seccomp(-bpf)`

À l'origine, **`seccomp(2)`** est un appel système pour jeter ses droits et se restreindre à la manipulation de fichiers déjà ouverts (`SECCOMP_SET_MODE_STRICT`).

Effectuer un autre appel système que **`exit`**, **`sigreturn`**, **`read`** ou **`write`** conduit à se faire tuer par le noyau.

Le mode filtrer *Berkley Packet Filter* (`SECCOMP_SET_MODE_FILTER`) permet de préciser la liste des appels systèmes autorisés.

`seccomp-bpf` a ainsi été particulièrement utilisé par Chromium OS dont le code source peut servir de référence d'implémentation.

Cgroups et *Namespaces* Linux

Cgroups et *Namespaces* Linux

La fonctionnalité Linux des *Control groups* (**`cgroups(7)`**) permet de :

- organiser les processus en groupes hiérarchiques
- limiter leurs ressources (**`systemd-cgtop`**)
- les superviser

Ils sont présentés à l'utilisateur comme un simple système de fichiers monté (**`/sys/fs/cgroup`**) : *tout est fichier*.

systemd crée par défaut un cgroup sous **`system.slice`** pour chaque service (→ **`systemd-cgls`**).

Cgroups et *Namespaces* Linux

Les `(*_)namespaces(7)` permettent de cloisonner les applications sur plusieurs types de ressources :

Type	Fonctionnalité
Mount	Isolation des points de montage `chroot` en *beaucoup* mieux
UTS	Isolation du nom d'hôte et de domaine
IPC	Isolation des communications interprocessus (IPC System V, files de messages POSIX)
User	Isolation des identifiants utilisateurs et des groupes Permet d'avoir un `UID 0` sans être vraiment privilégié
PID	Isolation des identifiants des processus
Network	Isolation des ressources réseau (interfaces, ports...)
Cgroup	Isolation des Cgroups
Time	Isolation des horloges système

Conteneurs et virtualisation

Conteneurs

Avec les fonctionnalités précédentes, il est possible sur un même système hôte :

- de mettre en place des sous-systèmes Linux (conteneurs)
- d'isoler les processus et les ressources associées

Ces conteneurs peuvent être gérés :

- bas niveau avec les commandes LXC (`lxc-*`)
- avec `systemd-nspawn` / `systemd-machined` (`machinectl`)
- plus haut niveau avec Docker ou containerd pour faciliter la diffusion et le déploiement

Conteneurs ±

Avantages

- isolation des composants applicatifs au sein des conteneurs
- réduction des privilèges (`user_namespaces`...)
- scalabilité horizontale facilitée par la répétabilité du déploiement

Inconvénients

- complexité accrue de l'architecture applicative
  (quel conteneur héberge quoi ?)
- complexité accrue de l'administration (socle + gestionnaire)
- jonction (plus) forte entre administration système et développement (*devops*)
- est-ce adapté pour assurer une sécurité élevée ?
  (confiance dans l'isolation offerte par le noyau)

Virtualisation

Lorsqu'une segmentation plus importante ou des fonctionnalités spécifiques du noyau sont requises (comme le filtrage réseau), l'isolation par conteneurs ne suffit plus.

On peut alors employer la virtualisation où un **hyperviseur** (comme KVM, VirtualBox, Xen ou VMware) émule ou **partage les ressources physiques** (CPU, mémoire, stockage, I/O) de l'hôte.

Le matériel doit être configuré pour supporter la virtualisation (bits Intel VT-x ou AMD-V...)

Virtualisation ±

Avantages

- accès aux fonctionnalités spécifiques du noyau
- isolation plus marquée, reposant sur des fonctionnalités du CPU
- performances accrues grâce aux ressources matérielles pouvant être dédiées

Inconvénients

- matériel dédié pour un hyperviseur type 1 (*bare metal*)
- possible complexité de l'architecture de virtualisation (cluster d'hyperviseurs, stockage SAN...)
- coût des potentielles licences (VMware...)
- est-ce adapté pour assurer une sécurité très élevée ? (confiance dans l'isolation offerte par l'hyperviseur)

Questions ?

Sécurité système

Journalisation et centralisation

Plan

1. Journaliser ?
2. Collecter et centraliser ?

Pourquoi journaliser ?

- Aider à comprendre ce qui se passe pour débugger
- Assurer une traçabilité légale ou réglementaire
- Contrôler : la confiance n'exclue pas le contrôle
- Détecter et analyser les incidents de sécurité

La suite est librement inspirée des « Recommandations de sécurité pour l'architecture d'un système de journalisation », ANSSI, www.ssi.gouv.fr/journalisation/

Journaliser ?

Que journaliser ?

- Authentifications et élévations de privilèges
- Gestion des comptes, rôles (groupes) et privilèges
- Modification des stratégies de sécurité (dont journalisation)
- Accès aux ressources sensibles (RWX + suppresion)
- Activité des processus ⚠ volumétrie
- Activité des systèmes (démarrage, modules noyau, matériel...)

- Événements de sécurité pare-feu

Comment journaliser ?

- Prévoir la fonctionnalité dans les applications
- Activité système, processus... : `auditd(8)`
- Fonctionnalité native ou `syslog`/`journald` + `logrotate`
- Identifier la source des journaux (application & système)
- Horodater ! (avec une même source de temps : NTP)
- Données structurées

Collecter et centraliser ?

Pourquoi collecter et centraliser ?

- Assurer le stockage dans la durée (plusieurs mois)
  Espace disque local probablement faible
- Assurer l'intégrité des journaux face à un attaquant
  - Contrôle de l'équipement $\neq$ contrôle des journaux exportés
  - Privilégier l'export en temps-réel !
- Faciliter la recherche avec un outil adapté (`grep` = 💀)
- Corréler les événements de plusieurs sources (SIEM)

Comment centraliser ?

- Utiliser des protocoles sécurisés (confidentialité, intégrité, authentification) natifs ou avec TLS, à défaut SSH ou IPSec
- Durcir et cloisonner la plateforme de centralisation (+ MCS)
- Prévoir une partition dédiée (cf. principes généraux)
- Superviser la volumétrie peut vite déborder
- Préférer une base de données structurée et indexée à des fichiers à plat
  - comme Elasticsearch + Logstash
  - à défaut, `syslog-ng` / `rsyslog`

Comment l'exploiter ?

- Recherche des anomalies « évidentes »
  (utilisateurs ou horaires atypiques, erreurs)
- Superviser les scénarios de risques ou de menaces identifiés
  (analyse de risques)
- Leitmotiv : **tout** ce qui n'est pas normal **est** anormal !

Questions ?

Sécurité système

Travaux dirigés

Installer un système sécurisé sur base Debian

- Chiffrement des disques
- Partitionnement et montage adaptés
- Durcissement noyau (`sysctl`)
- Durcissement OpenSSH (`PermitRootLogin no` au minimum)
- Filtrage Netfilter (⬇ SSH – ⬆ DHCP, DNS, HTTP[S])

→ Recommandations de l'ANSSI

- GNU/Linux www.ssi.gouv.fr/reco-securite-systeme-linux/ (~~`nosuid,noexec`~~ sur `/var` pour Debian)
- (Open)SSH www.ssi.gouv.fr/nt-ssh/

Cloisonner et isoler

Intégrer une application avec des composants isolés

- Nginx, uWSGI, Python et systemd
- Service HTTP (Flask) (`api.py`)
- Service cryptographique (`crypto.py`)
- Isolation des deux services
- Communication par socket UNIX

→ Recommandations de l'ANSSI

- Cloisonnement www.ssi.gouv.fr/guide-cloisonnement-systeme/
- Virtualisation www.ssi.gouv.fr/virtualisation/

Journaliser et centraliser

Ajouter des fonctionnalités de journalisation centralisée

- Journaliser les requêtes et les opérations cryptographiques
- Conserver localement sur 7 jours (`systemd` + `logrotate`)
- Centraliser sur un serveur voisin (`rsyslog`) et conserver 1 mois

Pour aller plus loin :

- Identifier les ressources sensibles et les surveiller avec `auditd`
- Mettre en place une pile Elasticsearch

→ Recommandations de l'ANSSI

- Journalisation www.ssi.gouv.fr/journalisation/
- Journalisation Windows www.ssi.gouv.fr/journalisation-windows/

En cas de besoin :

remi@rdubois.fr