INSA – 5° R.T.

CORRIGE

Examen du 19 janvier 2005

Sécurité des systèmes d'information

2ème partie

Exercice 1 *(3 points)*

Rédigez des exemples fictifs mais réalistes de règles de sécurité pouvant figurer dans chacun des documents suivants (6 règles en tout) :

A) 2 règles de la « Politique de Sécurité du Système d'Information » (PSSI)
B) 2 règles appartenant aux « Spécifications de sécurité réseau » (B.1) ou « Spécifications de sécurité des systèmes » (B.2)
C) 2 règles pour les « Cahier de recette de configuration d'Apache v.1.3 » (C.1) ou « Cahier de recette de configuration de Windows 2000 » (C.2)

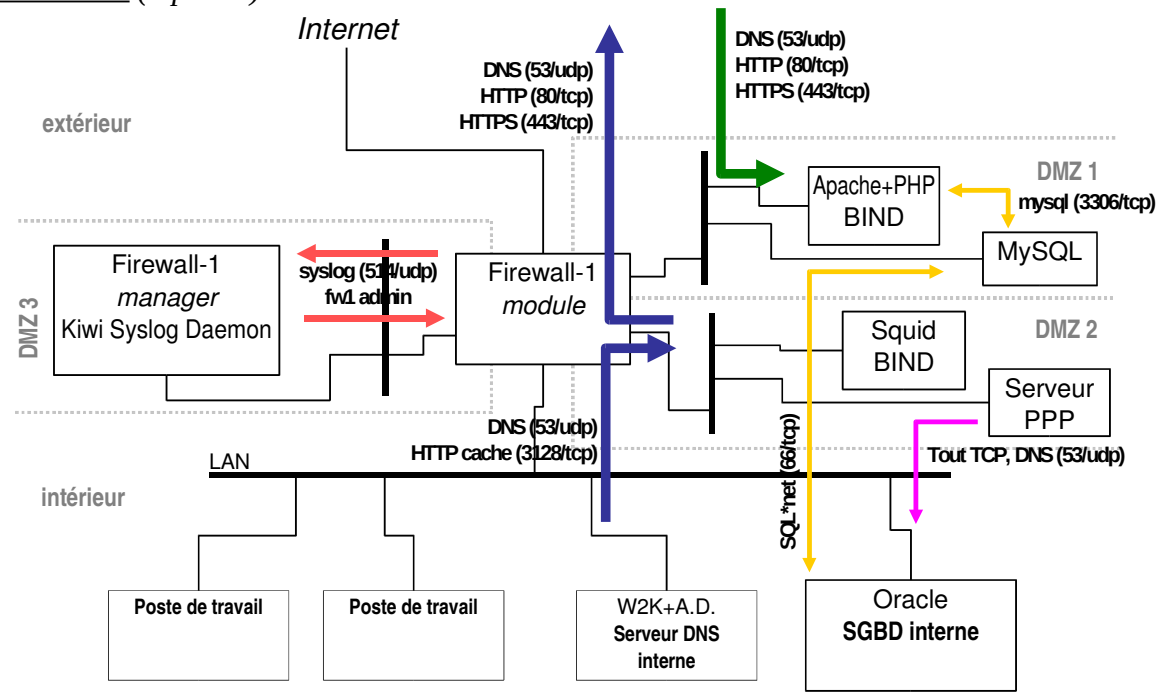
Identifiez bien dans lequel de ces 5 documents figure selon vous chacune de vos 6 propositions.

Exemples :

- *« Politique de Sécurité du Système d'Information » (PSSI)*
  1. *Tous les utilisateurs permanents du système d'information doivent être authentifiés par un mécanisme d'authentification basé sur un algorithme de cryptographie asymétrique (type RSA) et un dispositif portable et sécurisé de stockage de la clef publique (type carte à puce) avec des longueurs de clefs suffisantes.*
  2. *Tous les projets de développement d'applications à vocation généraliste doivent inclure un cahier des charges spécifique pour la définition des besoins de sécurité de ces applications dès la phase de conception.*
  3. *Il est interdit de désactiver les systèmes de sécurité du système d'information.*
  4. *Toute évolution de la PSSI doit être approuvée par le comité de direction.*
  5. *Seul le service HAL/2005 est autorisé à réaliser ou faire réaliser des recherches de vulnérabilités ou des tests d'intrusion sur le système d'information.*
  6. *Tout composant du système d'information sortant de la zone LOIN/51 doit occuper un volume de moins de 1 cm^3.*
- *« Spécifications de sécurité réseau » (B.1)*
  7. *Un sous-réseau sécurisé ne doit comporter que des hôtes dont le système est sécurisé (au sens de B.2) et n'être relié à d'autres sous-réseaux que par des équipements réseau sécurisés (au sens de B.1).*
  8. *L'utilisation des protocoles WiFi IEEE 802.11a et 802.11b est prohibée.*
  9. *L'utilisation d'un protocole de communication normalisé sur un numéro de port de communication non-standard est interdite sans autorisation.*
  10. *Il est interdit de connecter un système informatique sur le réseau opérationnel d'un pas de tir entre H-10 et H+1.*
  11. *Le réseau commercial d'un avion ne doit pas être connecté au(x) réseau(x) des systèmes de pilotage.*

- *« Spécifications de sécurité des systèmes » (B.2)*
  12. *L'utilisation des systèmes d'exploitation Windows 95, Windows 98 et DOS3.3 est interdite.*
  13. *L'utilisation de RSH est interdite.*
  14. *L'utilisation de Telnet et FTP est déconseillée. Elle est interdite sur les systèmes sécurisés.*
  15. *L'utilisation de SSH est recommandée. Elle est obligatoire sur les systèmes sécurisés.*
- *« Cahier de recette de configuration d'Apache v.1.3 » (C.1)*
  16. *Vérifier que la configuration d'Apache réside dans le répertoire /etc/apache/*
  17. *Le fichier /etc/apache/httpd.conf doit avoir les droits d'accès suivants : `rw- r-- ---`*
  18. *La directive BindAddress ne doit pas être utilisée.*
      (grep BindAddress /etc/apache/httpd.conf ne doit rien afficher.)
- *« Cahier de recette de configuration de Windows 2000 » (C.2)*
  19. *L'audit des modifications de stratégie de sécurité locale doit être activé pour les échecs et les réussites.*
  20. *L'audit des créations de comptes doit être activé pour les échecs et les réussites.*
  21. *La taille maximale du fichier de conservation de l'historique des évènements sécurité doit faire au moins 5 Mo.*
  22. *HKLM>Microsoft>Windows>CurrentVersion>AMoinsQue>Cenesoit>Ailleurs>PourInterdireLe>Routage doit valoir False.*
  23. *Le service WindowsUpdate doit être activé et utiliser le serveur XYZ.*
  24. *Vérifier que le poste de travail appartient à l'UO "Postes de travail" du domaine Active Directory.*

Exercice 2 *(4 points)*



Question 1 *(2 points)* : Compte tenu du mode de fonctionnement suggéré par le schéma, indiquez l'usage de chacune des DMZ et la (les) fonction(s) qu'elle apporte à l'organisation utilisant cette architecture.

Question 2 *(2 points)* : Critiquez (avantages/inconvénients) l'architecture utilisée (notamment en terme de sécurité). Proposez des évolutions de l'architecture permettant de pallier certains inconvénients et éventuellement précisez les nouveaux inconvénients associés à vos préconisations.

Description des DMZ (question 1) :

- *DMZ 1 : zone de mise à disposition de serveurs publics accessibles depuis Internet. On y voit apparaître un serveur Web (HTTP et HTTPS) intégrant un langage de programmation de pages Web dynamiques ou d'applications Web (PHP). Ces applications interagissent avec une base de données MySQL située sur une machine séparée. Ce serveur de données accède également à un autre SGBD interne basé sur Oracle situé sur le réseau local en utilisant le protocole SQL*net.*
  Par exemple, le serveur MySQL peut contenir le « panier » des clients d'un site Web marchand en train d'effectuer leurs achats, le catalogue des produits disponibles via Internet, les relevés de comptes à J-1, etc. Les commandes effectuées peuvent être consolidées une fois par jour sur le système de gestion interne appuyé sur Oracle.
  On identifie également la présence d'un serveur DNS, probablement utilisé seulement pour publier le nom du serveur Web et des machines accessibles depuis Internet (réponse seulement aux réquêtes itératives venant d'autres serveurs DNS sur Internet).
- *DMZ 2 : C'est d'abord une zone d'installation des relais utilisés pour les flux sortants du réseau local vers Internet. On y voit tout d'abord un relais HTTP (squid). Les machines du réseau local utilisent le protocole de cache HTTP pour accéder à ce relais et ainsi aux pages d'Internet. Ce mode de fonctionnement est une base permettant par exemple de réaliser le contrôle des pages accédées (filtrage d'URL).*
  Un deuxième composant, un serveur DNS (BIND) est également présent sur la même machine. Il est très probablement utilisé en mode relais seulement par le serveur DNS interne afin de statisfaire à la résolution générale des noms sur Internet (prise en charge des requêtes récursives issues du LAN via le serveur DNS interne).
  Cette DMZ est aussi la zone d'arrivée de connexions externes par modem permettant d'accéder au réseau local (nomades, itinérants, etc.).
- *DMZ 3 : DMZ d'administration du firewall. Celui-ci est administré depuis cette zone via les protocoles propriétaires du constructeur (CheckPoint dans ce cas). (Il est très probable qu'on a interdit de l'administrer depuis une autre de ses interfaces réseau.)*
  La machine hébergeant le manager du firewall peut également stocker des traces émises par d'autres machines présentes dans l'architecture via le protocole syslog.

*La DMZ 1 permet donc à l'entreprise de fournir un **service sur Internet** via un serveur Web.*
*La DMZ 2 offre à la fois un service **d'accès HTTP sortants** pour les postes de travail du réseau local et un service **d'accès** au réseau local de l'entreprise via le RTC (réseau téléphonique commuté).*
*Enfin, la DMZ 3 offre une zone **d'administration protégée** de l'architecture de communication.*

<u>Critique des DMZ et évolutions</u> (question 2) :

- *DMZ 1 :*
  - *Avantages :*
    - *La proximité du serveur Apache+PHP et du serveur MySQL facilite le développement des applications.*
  - *Inconvénients :*
    - *L'ensemble du service offert via Internet peut être compromis en cas de vulnérabilité du serveur Web ou de l'application qu'il exécute. Notamment, ceci pourrait permettre alors de rebondir depuis le serveur Apache vers le serveur MySQL ou tout simplement d'usurper son adresse IP ; celui-ci disposant d'un accès à un composant peut-être essentiel du système informatique interne (le SGBD), les conséquences peuvent être importantes.*
- *DMZ 2 :*
  - *Avantages :*
    - *Possibilité de faire une authentification des utilisateurs au niveau du relais Squid.*
    - *Possibilité de faire du filtrage d'URL au niveau de ce relais.*
    - *Les flux d'accès téléphoniques transitent par le firewall : on peut faire un contrôle des protocoles, ou une écoute du réseau.*
  - *Inconvénients :*
    - *L'infrastructure d'accès à Internet de l'architecture (le serveur Squid+BIND) n'est pas protégée par rapport aux machines utilisant l'accès téléphonique, alors qu'elle est protégée du réseau local.*
    - *Il n'y a pas de contrôle des protocoles autorisés à entrer sur le réseau local via le réseau téléphonique (tous les protocoles TCP sont autorisés). Il est souhaitable de limiter les protocoles TCP utilisables par les ordinateurs accédant via la serveur PPP.*
- *DMZ 3 :*
  - *Avantages :*
    - *Le manager peut facilement être utilisé pour gérer des modules Firewall-1 additionnels (par exemple dans des sites distants : agence, filiale, etc.). Ce manager pourra donc ultérieurement prendre en charge une architecture plus étendue avec une administration centralisée.*
    - *Le point de gestion du ou des firewall est clairement identifié, ainsi que le point de centralisation des traces.*
  - *Inconvénients :*
    - *On doit créer et protéger physiquement un réseau Ethernet isolé (problème d'organisation physique, de cablage, d'accessibilité, etc.)*
    - *L'architecture n'identifie pas précisément la "Console" d'administration du firewall (l'IHM utilisée par les administrateurs eux-mêmes). Il faut espérer que des restrictions d'accès par rapport au LAN ont été mises en place ; ou que le mot de passe d'accès au manager de Firewall-1 est extrèmement fort et très secret.*

La première évolution envisageable consiste à dissocier le serveur PPP du relais Squid dans des DMZ séparées, pour séparer les deux types de fonction.
Une autre évolution serait d'isoler le serveur MySQL dans une DMZ spécifique isolée d'Internet et relativement protégée d'intrusions réussies sur le serveur public.
L'architecture finale compterait alors 2 DMZ supplémentaires, soit un total de 7 interfaces sur le firewall. Les principaux inconvénients de cette évolution sont la complexité accrue de l'administration (surtout en présence de translation d'adresse) et éventuellement le coût.

Exercice 3 *(3 points)*

Voici 3 exemples de méthode utilisables par un utilisateur pour choisir un mot de passe :

1. Utilisation d'un mot de passe de 7 symboles exactement choisis au hasard parmi les **36** symboles alphanumériques (ce qui suppose que le système d'exploitation ne fait *pas* de distinction entre majuscules et minuscules).
   Exemple : `8PER2ZZ`
2. Utilisation de 2 mots du dictionnaire français courant accolés ou éventuellement séparés ou suivis par des caractères spéciaux choisis parmi : `,;:!?-`
   Exemples : `EMPORTE,VENT` ou `VIVE-MOI!`
3. Utilisation de la première lettre d'une phrase comptant au moins douze mots. On fera aussi l'hypothèse que toutes les lettres sont mises en minuscule pour constituer le mot de passe et qu'on n'inclut pas les signes de ponctuation.
   Exemple : `udlpldupcamdm` (1ère phrase ci-dessus)

Question 1 *(2 points)* : Comparez ces méthodes en évaluant *quantitativement* la taille de l'espace de recherche associé à chaque méthode, en vous servant si besoin des informations indiquées ci-dessous. Il n'est pas indispensable de calculer de manière exacte la taille de l'espace considéré, mais essayez d'obtenir un *ordre de grandeur* utile et justifié du nombre de mots de passe possibles (ou d'un minorant).
Question 2 *(1 point)*: Ensuite, en vous appuyant sur les informations expérimentales fournies ci-dessous, calculez (en jours) la durée espérée de résistance d'un mot de passe de type 1, 2 ou 3 face à un outil d'attaque par dictionnaire, suite à un vol de la forme chiffrée du mot de passe[1]. Considérez : un système Unix utilisant un chiffrement DES classique, un système Windows utilisant NT LM (DES) et un système OpenBSD utilisant Blowfish et un *salt* de grande taille.

Informations utiles :

- On pourra considérer qu'un lexique de français courant compte environ 4000 mots[2] (méthode 2 ou 3).
- Pour plus de réalisme, on pourra considérer (méthode 3) qu'une phrase est composée à 50% de « mots outils » qui ne représentent que 0,5% du lexique total mais surtout qui ne représentent que 25% des lettres de l'alphabet (pour leur première lettre), soit seulement 6 lettres.

[1] Contenue dans /etc/passwd ou /etc/shadow sur un système Unix, ou dans la base SAM d'une machine Windows, ou transitant sur le réseau pour certaines applications (VNC, etc.).
[2] Le lexique Dubois-Buyse des mots français courants (enfants entre 0 et 16 ans) compte 3726 mots.

- Le plus sûr est souvent d'évaluer expérimentalement le nombre de combinaisons par seconde (c/s) qu'une machine peut essayer. Voici, sur une machine moderne (P4 2,66 GHz) les performances obtenues par un outil librement disponible :

```
$ john -test
Benchmarking: Standard DES [48/64 4K]... DONE
Many salts:     192409 c/s real, 193182 c/s virtual

Benchmarking: OpenBSD Blowfish (x32) [32/32]... DONE
Raw:    380 c/s real, 381 c/s virtual

Benchmarking: NT LM DES [48/64 4K]... DONE
Raw:    1808998 c/s real, 1816263 c/s virtual
```

Pour les 3 méthodes, le nombre de mots de passe possibles est le suivant :

1. $36^7 = 78\,364\,164\,096 \simeq 7{,}8 \cdot 10^{10}$ *possibilités.*
2. *On* 4000^2 *possibilités pour le choix des 2 mots, et ensuite* $(6+1)^2$ *possibilités pour choisir, soit un des 6 symboles, soit rien du tout, entre ou après les 2 mots ; ce qui donne au total :* $4000^2 \times (6+1)^2 = 784\,000\,000 \simeq 7{,}8.10^8$ *possibilités.*
3. *On peut considérer qu'un caractère sur deux du mot de passe est choisi parmi les 26 lettres de l'alphabet des minuscules, et un caractère sur deux parmi un alphabet de 6 lettres (associé aux premières lettres des mots outils). Soit au total :* $6^6 \times 26^6 = 14\,412\,774\,445\,056 \simeq 1{,}4 \cdot 10^{13}$ *possibilités*[3].

En divisant le nombre total de mots de passe possibles par la vitesse d'essai de l'outil, on obtient alors les résultats suivants (valeurs approchées minorées).

	Cas 1	*Cas 2*	*Cas 3*
DES	*4,7 j*	*1,1 h*	*867 j*
NT LM	*12 h*	*7,2 min*	*92 j*
OpenBSD	*6,5 an*	*23 j*	*1200 an*

En conclusion, comme on peut le voir, les méthodes 1 et surtout 2 restent largement vulnérables à ce type d'attaque, surtout dans les cas où l'algorithme de stockage du mot de passe sur le système d'exploitation utilise des techniques « anciennes ». La méthode 3 semble fournir un meilleur niveau de sécurité, essentiellement dû à l'augmentation de la longueur du mot de passe (et au fait qu'on n'utilise pas directement les mots d'un dictionnaire).
On voit également l'avantage que peut présenter un système d'exploitation où la sécurité est prise en compte dès la conception, même pour des utilisateurs utilisant la méthode la plus faible.[4] *L'avantage d'OpenBSD tient en effet essentiellement à l'utilisation d'une fonction de chiffrement lente en même temps qu'un salt de très grande taille (128 bits).*
Dans tous les cas, compte tenu de l'efficacité de cette attaque dans la pratique, il faut souligner l'importance de protéger les serveurs contenant la forme chiffrée des mots de passe

[3] *En fait, cette valeur est trop optimiste. La fréquence d'apparition des différentes lettres de l'alphabet en première position sur les mots usuels est loin d'être égale. Il faudrait en tenir compte pour mieux évaluer la vitesse de parcours de l'espace de recherche par un outil essayant en premier les combinaisons les plus probables. Pour cela, il faut donc tout d'abord établir une statistique à partir d'un lexique du français. On comprend donc pourquoi certaines techniques de cryptanalyse conduisent des gens très sérieux à compter les lettres dans le dictionnaire...*

[4] *Et l'inconvénient dual...*

(utilisée en entrée pour l'outil d'attaque) : par exemple les contrôleurs de domaine, ou les gestionnaires NIS ou LDAP.