Faut-il brûler vos certificats ?

Serge Aumont
Comité Réseau des Universités
Campus Beaulieu 35042 Rennes Cedex
serge.aumont AT cru.fr

Octobre 2003

Résumé

Lorsque l'on découvre les applications des certificats X509 on se prend à rêver d'une solution universelle aux problèmes de confidentialité, d'authentification et de signature posés par les systèmes d'informations de nos établissements.
Cet article tente d'expliquer l'infléchissement de ces derniers mois du discours ambiant sur les IGC ainsi que le contraste entre les slogans sur "l'IGC solution universelle" et la modestie des déploiements opérationnels (en particulier dans la communauté Rénater) de cette technologie pourtant disponible depuis quelques années. Il montre les difficultés de mise en œuvre dues à de multiples facteurs : piètre qualité des implémentations disponibles, complexité du modèle, contraintes pesant sur les IGC, contexte juridique incertain, nécessité d'une implication plus grande de nos tutelles, ...

Mots clefs

IGC, PKI, Certificat X509, HTTPS, S/Mime, signature électronique, dématérialisation

1 Introduction

Les technologies basées sur des certificats X509 ne manquent pas d'atouts. Au premier chef, elles sont normalisées et très universelles. Les certificats permettent de faire bien plus que SSL (HTTPS, LDAPS, IMAPS, SMTP/TLS, …). La même technologie permet des choses aussi variées que sécuriser IP (IPSec), signer des documents ou des logiciels, contrôler des licences ou même sécuriser un système d'exploitation (projet Palladium). De plus, elles sont déjà très anciennes, donc probablement sont-elles mûres. Ainsi, la recommandation CCITT X509 (« The Directory - Authentication Framework ») date de 1988, la spécification par Netscape de SSL v2 de 1994, le RFC2246[1] sur TLS date seulement de 99, mais il reprend mot pour mot la spécification de SSLV3 de 96. Quant à S/MIME, le premier RFC1847[2] a été publié en 95. Les produits supportant ces standards sont eux aussi très anciens, par exemple, le projet Openssl a démarré au printemps 98. Outlook et Netscape supportent SSL et S/MIME depuis un grand nombre de versions.

Pourtant, la généralisation de ces technologies, annoncée avec tambours et trompettes, se fait toujours attendre.
Ainsi, même si cette édition de JRES est la troisième réservant une part significative de son programme à ces questions[1][4], et même si la plupart des techniciens de nos établissements a effectivement eu l'occasion d'installer quelques certificats et d'établir des sessions SSL, on ne peut certainement pas parler de déploiement massif d'applications utilisant les services de signature, d'intégrité et de non répudiation qui semblent pourtant à notre portée. Les travaux du CRU sur ce sujet ne sont que des gouttes d'eau comparés à la variété et aux volumes des autres usages de nos réseaux.
Il en va de même au-delà de notre communauté. Certes, les serveurs HTTPS sont aujourd'hui légion, mais voici plusieurs années que l'on brandit toujours les mêmes exemples de mise en œuvre des technologies de signature : télé-déclaration de la TVA, télé-déclaration de l'impôt sur le revenu, carte santé, etc. Il est tout à fait remarquable que ces grandes applications soient issues de la sphère des administrations et non de celle de l'entreprise comme si ces technologies ne s'imposaient pas d'elles même du fait de leur pertinence technique, mais bien au contraire, à la condition d'être soutenues par une volonté politique active.
Depuis un an la presse économique distille un message contrasté sur le sujet. On y lit toujours que ces technologies d'avenir représentent de juteux marchés et que les leaders de ce secteur ont un fort potentiel de développement. On y découvre aussi que beaucoup d'entre elles ne font pas de bénéfices et que le marché, très en dessous des prévisions, n'en finit pas d'être sur le point de décoller. Quelques acteurs du secteur seraient en difficulté. Certains analystes se risquent même à avancer l'idée que comme pour toute infrastructure, les profits ne seront jamais au rendez-vous de la prestation d'IGC du fait du montant des investissements et des coups de maintenance. Bien que cette presse économique nous ait habitués à raconter tout et son contraire, on peut sans conteste retenir que l'unanimité autour de cette question s'effrite.

Plusieurs articles [10][11] ont décrit il y a déjà trois ans les insuffisances des IGC et certains auteurs ont déjà enterré les IGC[12][13][14]. Cet article tentera plus modestement de montrer l'écart considérable entre nos pratiques actuelles et le modèle théorique très sophistiqué sur lequel nos usages devraient s'appuyer. Après une énumération de défauts constatés

de nos clients et de nos serveurs, l'article aborde aussi les questions plus théoriques sur la révocation de certificats et l'exploitation d'une IGC. Puis il aborde le champ de la signature électronique et son contexte juridique.

2 Une technologie toujours immature

La longue litanie des insuffisances des technologies basées sur les certificats de cette section vise à démontrer l'écart considérable entre les pratiques actuelles et les modèles théoriques très sophistiqués sur lesquels elles sont construites.

La première hypothèse qui pourrait expliquer ce retard de déploiement concerne les réserves sur la maturité de ces technologies. Comme cela a été évoqué dans l'introduction de cet article, les grands produits du marché implémentent depuis déjà longtemps des applications à base de certificats. Mais il est assez étonnant de constater que malgré cette ancienneté, la qualité des implémentations est souvent peut satisfaisante. Pire, elle n'évolue pas toujours favorablement.

2.1 Le contrôle des autorités de confiance dans les clients

Rappelons que la validation d'un certificat répond à plusieurs critères qui dépendent de l'application concernée. Schématiquement, le certificat doit avoir été émis par une autorité de certification aussi appelée autorité de confiance (AC) pour l'usage prévu, il ne doit pas être périmé ou avoir été révoqué. En ajoutant une AC dans la liste des autorités de confiance d'un navigateur, on accorde une confiance dans tous les certificats émis par cette autorité. Connaître et contrôler la liste des AC d'un navigateur est donc un enjeu fondamental pour la sécurité des opérations qui seront faites avec ce navigateur.

Netscape, Mozilla, Internet Explorer représentent l'énorme majorité des clients utilisés dans notre communauté. Tous permettent d'installer des certificats personnels ou d'autorité de certification. A l'installation, ils sont préconfigurés avec une liste d'autorités de certification ; celles qui ont un poids commercial ou des relations privilégiées avec les éditeurs de ces navigateurs. Elles sont au nombre de 80 dans Internet Explorer et 83 dans Mozilla.

Toutes les AC de confiance sont égales vis-à-vis des navigateurs dans lesquels elles sont installées : on distingue seulement l'usage de S/MIME, SSL et des applets signées ce qui est singulièrement insuffisant… On peut par exemple légitimement se poser la question de l'intérêt des deux AC « Thawte premium » et « Thawte freemail » puisque toutes deux ont dans les mêmes paramètres de vérification des signatures S/MIME dans les navigateurs. (le seul élément de la politique de certification de l'AC gratuite « Thawte freemail » est une vérification de l'adresse email par retour de courrier).
Ce point est particulièrement préoccupant car tout le bel édifice de sécurité (les politiques de certification, les déclarations des pratiques de certification, les accords de confiances entre AC, etc) n'est que du décorum si l'on ne peut ni choisir les autorités de référence ni pondérer leur validité en fonction des usages.

Mozilla permet en apparence de supprimer des AC par défaut, mais celles-ci réapparaissent subrepticement lors du premier redémarrage de l'application ! Seule solution éditer une par une ces autorités et décocher les paramètres de confiance « `this certificate can identify web site` », « `this certificate can identify mail user` », « `this certificate can identify software maker` »… Quant à Internet Explorer, sous Windows XP, l'opération « Windows update » réinstalle à votre insu deux certificats Microsoft. Internet Explorer permettrait-il d'ajouter d'autres AC de confiance dans un système Windows XP à l'insu de son utilisateur ?

Ce problème du contrôle des AC pré-installées est d'autant plus préoccupant qu'il n'est pas le fruit d'une bogue amenée à disparaître mais au contraire, il résulte d'un choix déterminé des éditeurs de ces navigateurs.

Le modèle de confiance d'une AC nous apprend que la confiance est relative à un usage or les navigateurs ne nous permettent pas pour chaque AC cette association entre la confiance dans AC et certaines applications. Par exemple, il est impossible de définir que les serveurs de l'intranet de l'établissement doivent présenter un certificat émis par l'AC du CRU, ceux liés aux mises à niveaux du système Windows des certificats Microsoft, ceux du monde commercial par d'autres AC. En gros, accepter la confiance d'une AC se fait pour toutes les applications ou aucune. Ce problème ne concerne pas que les sessions SSL ; dans le cas de Mozilla, il concerne aussi la vérification de signature S/MIME et dans le cas des systèmes Windows, toutes les applications Microsoft partagent les ACs stockées dans les « magasins de certificats » du système.

S'il est difficile de supprimer des autorités de certification, il est possible d'en installer de nouvelles. Malheureusement, l'ergonomie et les bogues de ce processus rendent cette opération difficilement gérable par une population de non spécialistes. Cette opération se fait en général directement dans une session HTTP ; le serveur offrant l'interface utilisateur

de l'IGC envoie une entête appropriée qui déclenche sur le client le lancement d'assistant d'installation du certificat d'autorité. Malheureusement, certaines combinaisons de version d'Internet Explorer/Windows ne reconnaissent pas le « `Content-type` » utilisé. Dans certains cas liés à l'acceptation d'un certificat serveur, Mozilla installe le certificat d'autorité correspondant mais n'autorise aucune opération basée sur ce certificat. L'utilisateur doit alors parcourir une série de menus très mal documentés pour éditer les usages autorisés du certificat, aucune opération du serveur ne peut venir au secours de l'utilisateur confronté à des diagnostics très souvent incompréhensibles ou donnant des informations erronées. Ces difficultés ne sont certes pas des obstacles théoriques au bon fonctionnement mais elles peuvent s'avérer extrêmement consommatrices en assistance.

La seule issue crédible serait-elle l'emploi exclusif de certificats commerciaux pré installés dans les navigateurs du marché ?

2.2 Le contrôle des certificats corrompus par les clients

Pour une opération sensible, la validation d'un certificat impose la vérification du statut du certificat vis-à-vis de la liste de référence des certificats corrompus (l'équivalent des listes de cartes bancaires volées). Cette question de la révocation fera l'objet d'un paragraphe spécifique. Il n'est question ici que de la qualité de traitement des listes de révocations par les clients IE, Netscape, Mozilla. Netscape 4.X et 6.X nous obligent à recharger manuellement la liste de révocation car aucun dispositif de mise à jour automatique n'y a été prévu. Si l'utilisateur oublie cette opération, toute vérification de certificat devient impossible lors de l'expiration de la validité de la liste de révocation alors qu'au contraire, elle reste possible si l'utilisateur néglige les listes de révocation et choisit de ne pas les installer…
Même si l'AC a pris soin d'indiquer dans chaque certificat que la vérification des listes de révocation en cours de validité est obligatoire, cet attribut X509 est superbement ignoré par les navigateurs Netscape, Mozilla et Internet Explorer.

2.3 Certificats de signature versus certificats de chiffrement

Plusieurs notions de base dans le domaine des applications à base de certification ne sont pas implémentées ou le sont très mal. Ainsi, on sait parfaitement qu'on ne peut pas assurer avec l'usage classique d'un bi-clé unique la non répudiation d'une signature et le recouvrement d'un document chiffré. Le recouvrement d'un document suppose un service de recouvrement de la clé ce qui est incompatible avec le « contrôle exclusif de la clé privée » propriété indispensable pour assurer la non répudiation. Depuis peu Mozilla, Netscape 7 et Outlook implémentent la possibilité d'utiliser deux certificats distincts pour le chiffrement et la signature mais aucune de ces applications n'empêche une personne de chiffrer des messages en utilisant un certificat de signature.

Les trois clients SSL référencés comportent leur lot de bogues comme tout autre produit informatique, mais dans ce cas, la qualité des implémentations n'évolue pas toujours favorablement. Ainsi, Netscape 6 ne comportait plus de client S/MIME alors que Netscape 4.7X disposait de ce service depuis longtemps. Mozilla 1.4 et Netscape 7 ne peuvent traiter correctement les certificats dans lesquels le DN contient des lettres accentuées, etc.

2.4 Conclusion sur la qualité des clients

Devant la difficulté de contrôler la mise en œuvre effective des éléments techniques d'une politique de confiance dans chaque navigateur, il convient de s'interroger. Est-il possible de mettre en pratique les éléments d'une politique de confiance concernant les clients dans l'ensemble des navigateurs d'un établissement ? La partie semble moins délicate si l'on décide de diffuser un navigateur préalablement configuré avec les bonnes AC sur tous les postes de l'établissement. On en profitera pour y adjoindre bon nombre de réglages initiaux pour donner au navigateur une couleur locale, mais si l'effort pour y arriver est assez important, on devine que celui-ci se heurtera à l'impossibilité de s'immiscer dans la configuration d'un grand nombre de postes de travail, en particulier les PC portables et assistants personnels, sans oublier les modifications que les utilisateurs ne manqueront pas de faire dans la configuration initiale.

Ce n'est donc pas un hasard si devant cet ensemble de difficultés, des travaux de standardisation (RFC3029[5]) et certains éditeurs ont mis au point des serveurs de validation qui permettent de s'affranchir des médiocres fonctionnalités des navigateurs dans ce domaine. Cette approche permet de s'assurer du respect de la politique de validation d'une signature grâce à une administration centralisée.

2.5 De la qualité des serveurs HTTPS

Il est bien plus facile de choisir et de faire évoluer quelques serveurs que de déployer un parc de poste de travail. Cependant, nos serveurs HTTPS, eux non plus, ne sont pas au dessus de tout soupçon.

Sauf à y adjoindre un dispositif matériel spécifique, de par son architecture, un serveur HTTPS (IIS ou Apache) doit avoir accès à la clé privée associée à son certificat. Dès lors, l'administrateur doit choisir entre protéger cette clé par une « passphrase » ou installer celle-ci en clair sur le disque du système. Dans le premier cas, tout redémarrage du serveur ou du démon http requiert l'intervention d'un opérateur. C'est bien entendu, une contrainte d'exploitation intenable et l'énorme majorité des serveurs HTTPS fonctionne avec leur clé privée stockée sur le disque, non chiffrée, accessible en lecture par le démon httpd. Cette faiblesse devrait au minimum nous imposer de n'utiliser que des machines dédiées. Elle peut en outre être potentiellement exploitée par diverses méthodes en particulier en forçant l'exécution d'un code hostile interprété (perl, php). Ce code est en général soumis au serveur via les données d'un formulaire CGI mal écrit dans le but de récupérer cette clé privée (nous aborderons aussi les interrogations sur la protection des clefs privées personnelles).

Le couple Apache/ModSSL représente un très fort pourcentage des serveurs HTTPS. Il ne fournit quasiment aucun outil pour la gestion des listes de révocations (CRL) ou des serveurs « On line Certificate Status Protocol » (OCSP) (. Il est incapable de rapatrier les listes de révocations dont l'URL figure pourtant dans les certificats X509 des AC de confiance et ignore complètement l'attribut qui indique que l'usage d'une CRL valide est requis : à chacun de bricoler à coup de « crontab » une gestion des CRL ! Sur ce point le serveur Microsoft IIS est nettement supérieur puisque l'exploitation des CRL est intégrée et ne réclame aucune configuration.

2.6 HTTPS illusion ou gain de sécurité ?

Toutes les objections de cette première partie sur les qualités et les défauts des serveurs et surtout des clients HTTPS sont-elles si graves ? Pour répondre à cette question, il faut préalablement savoir quels objectifs l'administrateur poursuit en installant une solution SSL. Trois réponses sont affichées :

- identifier les utilisateurs dotés de certificats
- contrer la menace de la mise en place d'un faux serveur. Cette attaque est peut probable car elle demande préalablement d'installer de fausses informations dans le DNS puis l'installation d'un faux serveur à l'apparence crédible mais qui serait facile à détecter
- contrer la menace du « reniflage » des mots de passe

Le premier objectif est très difficile à atteindre sauf dans des communautés réduites d'utilisateurs car le déploiement de certificats personnels est une véritable gageure. Notre expérience au sein de la communauté des RSSI ou des correspondants logiciels d'établissement, bien ciblée sur une population d'informaticiens et d'experts sécurité montre de très grosses difficultés et un besoin d'assistance important. La menace du faux serveur est surtout théorique, de plus un certificat ne saurait nous protéger contre le détournement d'un nom de domaine dans un autre « top level domain » . Exemple le serveur de JRES est-il www.jres.fr, www.jres.org, jres.edu ou www.jres.info ?

C'est donc bien le chiffrement de la session HTTP qui est la raison du succès de HTTPS. A cet égard, un certificat serveur auto-signé est tout à fait suffisant.

Dans la sphère des serveurs commerciaux, nombre d'administrateurs affichent fièrement l'emploi de session SSL comme un véritable label de sécurité du serveur. En effet, les utilisateurs ont vite fait l'amalgame entre un « serveur sécurisé » et une « session sécurisée » vers un serveur qui peut-être contient lui-même de nombreuses failles et accueille quelques pirates. A défaut d'une politique de sécurité, le chiffrement des sessions HTTPS peut n'être qu'un gadget.

2.7 La protection des clés privées

Nous avons abordé le problème des clés privées de serveurs. Dans le cas de la diffusion de certificats de personne, la protection des clefs privées est un élément très important de la politique de certification de l'AC. En effet, nous savons tous que les magasins de certificats de Windows, de même que la base de certificats de Mozilla, sont exposés parce que stockés

sur le disque dur du poste de travail (bien entendu, les clés privées des certificats sont chiffrées avec un mot de passe choisi par l'utilisateur). Dans ce cas, il est relativement facile de s'emparer des clés sous leur forme chiffrée, c'est en particulier le cas lors du vol d'un PC portable, ou de l'intrusion d'un pirate sur le poste. L'attaquant peut alors appliquer toutes les méthodes d'attaque de mot de passe usuelles en particulier l'attaque par dictionnaire. Les chances de succès de telles attaques sont réelles car :

- les tentatives sont faites « offline » et ne laissent donc aucune trace dans les logs d'aucun serveur sans limitation de temps.
- les mots de passe ont de grandes chances d'être fragiles car il n'est pas possible d'appliquer une politique centralisée de gestion des mots de passe (longueur et complexité minimale, test d'existence dans un dictionnaire, changement périodique obligatoire, …).

(Notre expérience montre un autre effet désastreux de cette absence de gestion centralisée des mots de passe : un nombre significatif d'utilisateurs perdent ce mot de passe. Il faut alors révoquer le certificat et en émettre un autre.)

Il semble donc particulièrement important de s'appuyer sur une interface « PKCS#11 » aussi appelé *Cryptoki* [7] qui permet la génération du bi-clé initial, le stockage et la rétention de la clé privée sur un support actif (token USB, carte à puce, …). Les services de chiffrement et de signature sont alors accessibles via une API dont le code s'exécute sur le dispositif de protection de clefs. Ainsi organisé il est possible de rendre totalement inaccessible la clé privée en dehors de son support sécurisé. Même si certains spécialistes décrivent des attaques de ces supports de clés, le niveau de protection atteint est très élevé au regard des pratiques actuelles par mot de passe, mais à quel prix ?

Les tokens USB sont à la mode car contrairement aux cartes à puce ils ne requièrent pas l'installation de lecteur spécifique. Les promoteurs de cette solution font facilement l'impasse sur les inconvénients de ce support :

- l'installation d'un driver spécifique
- la connectique USB inadaptée à des usages instantanés : on ouvre facilement une porte avec une carte à puce, cela parait impossible avec un token USB !
- le caractère totalement impersonnel du support ce qui est paradoxal pour supporter une donnée éminemment personnelle (une carte à puce sera le plus souvent gravée avec le nom de son titulaire, voir avec sa photo). Une conséquence observée de ce manque de personnalisation est la tendance des utilisateurs à prêter leur token ! Adieu l'imputabilité des transactions !
- nos références culturelles associent la carte à puce et la carte bancaire, objet qu'il convient de protéger alors que les cartes mémoires USB qui se généralisent ressemblent comme deux gouttes d'eaux aux tokens USB cryptoki créant la confusion.

Que le choix se porte sur une carte à puce ou sur un token USB, le déploiement impose de :

- initialiser électroniquement le support (et dans le cas d'une carte à puce le personnaliser physiquement) ;
- le faire parvenir à son titulaire ;
- s'assurer que le titulaire fait bien sa demande de certificat sur le support et non via son navigateur. Cela ne semble possible que par un contrôle visuel d'un opérateur !
- déployer les pilotes et dans le cas de cartes à puce les lecteurs ;
- gérer les pertes de mot de passe (reformater le support, révoquer le certificat) ;
- gérer les pertes de support (révoquer le certificat, détecter les utilisateurs récidivistes pour gérer le coût induit par les pertes) ;
- récupérer les supports lors du départ de son titulaire (mutation, etc).

Les supports sécurisés de clés privées apportent non seulement un gain sécurité, mais aussi un lot de nouveaux problèmes de gestion et de sécurité principalement dus au fait qu'après avoir été distribués, il n'est plus possible de faire des opérations de gestion centralisée de ces supports PKCS11. Ces difficultés sont suffisamment significatives pour que une société (*Cryptolog*) ait choisi de promouvoir un serveur centralisé de « cartes à puce virtuelles ».

2.8 Mobilité et certificat

Les techniques basées sur SSL sont un apport remarquable pour ouvrir l'accès à un certain nombre de services à des personnes à l'extérieur du réseau interne de l'établissement. On pense bien entendu à la mise en œuvre de VPN permettant de prolonger le réseau local jusque sur le portable de l'utilisateur nomade. Cette approche n'est pas sans risque car si le

VPN protège le lien entre le poste nomade et le réseau local, il raccorde au cœur sensible du réseau local des machines vulnérables comme les PC familiaux dont les règles d'usage ne sont pas celles de l'établissement. Aussi, les usages les plus répandus sont moins ambitieux ; ainsi par exemple, les protocoles IMAPS et POPS permettent d'ouvrir raisonnablement l'accès au service de messagerie en installant simplement un certificat sur le serveur de courrier. Cependant, ces personnes de l'établissement connectées hors du réseau local sont bloquées quand elles souhaitent poster des courriers par les règles d'anti-relais. L'utilisation de SMTP sur TLS constitue une solution élégante en permettant de lever cette restriction mais dans ce cas un certificat serveur n'est pas suffisant, il faut aussi distribuer des certificats aux personnes car le but n'est plus de protéger un mot passe mais d'identifier le poste client pour autoriser le relais.
Il convient donc d'examiner comment un utilisateur doté d'un certificat peut utiliser celui-ci dans des cas aussi variés que l'usage d'un PC portable ou la connexion depuis un cyber café. Dans le premier cas, le PC portable peut être le support du certificat de l'usager (et de la clé privée associée), dans le deuxième cas, l'utilisateur n'a aucun contrôle sur le poste client ; le « token usb » avec son API cryptoki (PKCS#11) n'est d'aucun secours car il ne pourra probablement pas installer le pilote pour ce matériel. Il peut à la rigueur installer son certificat personnel à partir d'un fichier au format PKCS#12 qu'il transporte avec lui sur un support quelconque (cette opération est possible si la clé privée est exportable, attention, dans ce cas, un pirate peut obtenir après une intrusion une copie en clair de cette clé). Cette opération est lourde, elle est aussi particulièrement risquée car l'utilisateur ne dispose pas de garanties suffisantes quant à l'effacement réel de sa clé privée lorsqu'il quitte son poste de travail occasionnel.

Assurément, SSH et le webmail ont encore de beaux jours devant eux.

3 La révocation

La révocation est souvent présentée comme le talon d'Achille de l'IGC. Elle pose en effet de nombreuses difficultés théoriques et pratiques qu'on ne saurait évacuer simplement en arguant que la révocation est un évènement exceptionnel. En effet, même si la corruption d'un certificat par vol de la clef privée est un évènement rarissime, on ne peut accepter ce risque sans évaluer les procédures à mettre en œuvre a posteriori.

La première question, probablement la plus préoccupante est celle de la détection qu'une clé à été compromise ou risque d'avoir été compromise. Bien entendu, il y a des cas où la réponse est évidente. En cas de vol d'un PC portable, d'un support USB ou d'une carte à puce, il ne fait pas de doute qu'il faille immédiatement demander la révocation des certificats correspondants. En dehors du cas évident d'un vol du support physique du certificat, comment savoir si une personne hostile s'est emparée d'une copie des clés privées de mes certificats si je réalise que je me suis absenté de mon bureau en oubliant de verrouiller l'accès à mon navigateur ? Quelles sont les conséquences d'un virus ou d'une faille d'Internet Explorer sur la protection d'une clé ?

Bien entendu, ce n'est pas la seule difficulté de la révocation, on trouve de nombreux argumentaires théoriques qui discutent de ce point. La technique de base est de mettre à disposition un fichier signé contenant les numéros de série des certificats révoqués (appelé dans le jargon des IGC CRL). On est alors confronté au problème de diffusion de cette CRL. La fenêtre de vulnérabilité est la période pendant laquelle un usurpateur peut utiliser un certificat corrompu. Ceux qui gèrent des antivirus connaissent bien cette notion. Cette fenêtre comprend le délai entre le vol et la diffusion sur les applications de la CRL. Elle dépend de :

1. la capacité à détecter la corruption d'une clé ;
2. la réactivité de l'IGC responsable de la révocation ;
3. la diffusion de la CRL jusqu'aux applications.

Pour diminuer la fenêtre de vulnérabilité, il faut impérativement maintenir un haut niveau de disponibilité du service de révocation. C'est une contrainte très forte pour l'exploitation d'une IGC et son impact sur le coût d'usage de chaque certificat est un facteur important qui pousse à mutualiser fortement l'IGC, donc à en élargir la couverture. Le deuxième facteur qui permet de limiter la fenêtre de vulnérabilité est d'augmenter la fréquence de diffusion de la CRL. Dans le cadre de la technique classique des CRLs, au lieu de maîtriser directement cette fréquence on doit définir la durée de validité de la CRL. Les utilisateurs doivent rafraîchir la CRL avant son expiration. Le choix de la durée de validité de la CRL est un compromis statique entre la charge induite sur le serveur et la réactivité du service. La durée choisie ne peut être révisée après émission de la CRL même si une avalanche exceptionnelle de révocations intervient. Imaginez une faille dans le navigateur le plus utilisé tel qu'on puisse s'emparer à distance des clés privées stockées sur ce navigateur. Dans une telle hypothèse, non seulement il faudrait un grand nombre de certificats, mais en outre, faute d'un dispositif de « push », la propagation de l'information de révocation prendrait plusieurs jours.

Une autre limitation de la technique de CRL est liée à la taille des CRL. Le nombre de certificats révoqués, donc la taille d'une CRL sont strictement croissants puisqu'une révocation est définitive. Ce nombre est proportionnel au nombre de certificats distribués (plus il y a d'utilisateurs, plus il y a de risques de compromission). Le temps de transfert et d'analyse d'une CRL lors de la vérification d'un certificat peut devenir significatif quand la CRL est importante. Par ailleurs, le nombre d'accès au serveur délivrant la CRL est proportionnel au nombre d'utilisateurs. Les durées de validité des CRL rencontrées actuellement vont de une heure à un mois. Cet écart illustre la difficulté de choisir un compromis satisfaisant.

Le volume des échanges serait donc proportionnel au carré du nombre d'utilisateurs. C'est donc un facteur de blocage théorique si les usages venaient à se développer comme on nous le promet. Pour répondre à ce problème deux solutions ont été imaginées : les listes de révocations incrémentales et les serveurs de validation de certificat utilisant le protocole OCSP (Online Certificat Status Protocol). La première solution est partielle puisqu'elle vise simplement à diminuer la taille des fichiers transférés, en outre elle est très peu implémentée. L'autre solution permet de tester la validité d'un certificat au moment où l'on y fait référence et donc de supprimer un des trois facteurs qui constituent la fenêtre de validité : le délai de diffusion de l'information de révocation. Les promoteurs d'OCSP arguent en outre du gain de volume lié à l'absence de transfert de la liste des certificats révoqués. Ce gain est discutable car il est compensé par le nombre d'accès au service qui n'est plus lié au nombre d'utilisateurs mais au nombre d'opérations de validation de certificats. Enfin, une attaque par déni de service sur un serveur OCSP pourrait momentanément paralyser les usages d'une IGC. Le RFC 2560 qui défini OCSP date de 1999 mais les implémentations ne sont pas encore déployées et les usages sont encore très largement centrés sur la technique des CRL. Mozilla et Netscape 6 disposent d'un client OCSP, pas Internet Explorer.

4 Exploiter une IGC ou comment vivre dangereusement et sous la contrainte

Il est d'usage de répéter que le travail autour d'une IGC « c'est 80% d'organisationnel et 20% de technique ». Ce slogan est très inquiétant pour qui s'est frotté aux contraintes techniques d'exploitation d'une IGC dont nous ne donnerons qu'un exemple. Ces contraintes sont souvent très fortes et contradictoires. Ainsi on doit d'une part assurer la disponibilité des clés privées des ACs et d'autre part assurer une protection très forte de ces clés. Le premier objectif ne saurait être négligé car perdre la clé de l'AC racine serait presque aussi grave qu'une compromission de celle-ci (rappelons qu'une AC est tenue de maintenir la CRL qu'elle signe jusqu'à l'expiration du dernier certificat émis par cette AC). Il convient donc de dupliquer cette clé et d'en stoker les copies dans des endroits indépendants (un sinistre ne doit pas pouvoir détruire toutes les copies de cette clé). De même, on prendra soin de donner les procédures d'accès de ces clés à plusieurs personnes. Tout ce que nous pouvons faire pour nous protéger contre la perte des clés suite à un sinistre ou au départ d'un opérateur de l'AC rend plus difficile l'objectif de protection des clés. Ce deuxième objectif nous conduira par exemple à ne pas laisser à une personne seule les moyens d'utiliser cette clé. Pour satisfaire ces deux contraintes, on a recours à des algorithmes de type « *n parmi m* » qui impose l'accord de *n* opérateurs parmi une équipe de m personnes habilitées pour tout usage de la clé.

Devinette : considérant les règles de la RTT et le délai de réactivité de 24 heures spécifié dans la politique de certification de l'IGC, comment fixer l'effectif minimum de l'équipe pour ne jamais laisser les clés de l'IGC à un seul agent ?

Un des enjeux primordiaux de l'exploitation d'une AC est bien entendu de maintenir le niveau de confidentialité de la clé privée de l'AC pendant toute la durée de validité du certificat associé. Cette durée de vie est fixée lors de la cérémonie de génération des clés, elle peut être prolongée par renouvellement du certificat mais ne peut être diminuée. Elle est en général assez longue car il est extrêmement difficile de changer le bi-clé d'une AC. Une durée de validité typique pour un certificat d'AC se compte en dizaines d'années, par exemple le certificat de l'AC racine de la direction générale des impôts est valide jusqu'en 2013, celui de l'AC « Certiposte » jusqu'en 2018, celui de « VeriSign Trust Network » jusqu'en 2028 etc.
Il ne faut pas confondre le niveau de protection mis en œuvre pour la protection de la clé privée et le niveau de confiance dans la confidentialité de celle-ci même si la première notion conditionne la seconde. En effet, le niveau de confiance dans la confidentialité d'une clé privée est fonction de la protection la **moins** élevée jamais appliquée à la gestion de cette clé durant son cycle de vie. On se doit de renforcer les mesures de protection qui entourent une clé privée, ne serait-ce que parce que les menaces qui pèsent sur le système risquent de s'intensifier. Pour autant, de nouvelles mesures de protection ne sauraient effacer un doute, aussi faible soit-il, qui serait né d'un incident ou d'un manquement aux règles de gestion édictées dans la politique de certification. Autrement dit, la confiance dans une clé est une fonction décroissante du temps, toute erreur constitue une tache indélébile.

Un certificat doit contenir une référence à la politique de certification en vigueur lors de l'émission de ce certificat. Valider un certificat pour un usage consiste à accorder ou non une certaine confiance dans celui-ci. Cette opération se fait donc normalement en fonction de la politique de certification en cours lors de l'émission de ce certificat en testant l'extension

X509 V3 « certificate policies » contenant un identificateur unique d'une politique de certification. En théorie, cela permet de faire évoluer la politique de certification dans le courrant de la vie d'une IGC. Nous avons montré qu'en pratique les clients et les serveurs qui acceptent ou refusent un certificat sont très loin de ce niveau de finesse dans le processus de validation. Aussi faut-il peser chaque engagement figurant dans une politique de certification au regard de la durée de vie du certificat correspondant. Ainsi par exemple, s'engager à prendre en compte une demande de révocation dans un délai de 24 heures pendant une durée de 20 ans ne saurait être fait à la légère.

L'émission d'un certificat pour une durée de validité de plusieurs années est donc un véritable pari sur la capacité de la structure à respecter la politique de certification avec toutes les contraintes qu'elle suppose. A cela s'ajoute un régime législatif qui rend l'autorité administrative de l'IGC financièrement responsable des conséquences d'une anomalie dans la gestion de l'IGC.

C'est la difficulté de gagner un tel pari qui explique les précautions extrêmes qui entourent habituellement le cœur d'une autorité de certification. L'exemple authentique d'une cérémonie de clé mise en œuvre dans une cage de Faraday et à l'issue de laquelle on broie le disque dur ayant été utilisé illustre la paranoïa qui est de mise ! Le niveau de confiance atteint est-il à la hauteur de ces contraintes ?

5 Un attirail juridique inabouti

La transcription dans le droit et la réglementation française des directives européennes sur la signature électronique est maintenant aboutie, du moins peut-on l'espérer. Le droit de la preuve reconnaît à l'écrit électronique la même valeur probante qu'à l'écrit papier. Les non juristes pourront se reporter à une synthèse des textes sur la signature électronique[6] sur le site du CRU. L'arsenal législatif distingue trois formes de signature :

- la signature électronique ;
- la signature électronique sécurisée ;
- la signature électronique sécurisée présumée fiable.

Les trois signatures sont recevables. Pour les deux premiers types de signature, en cas de contestation d'une signature par une partie, il appartient à l'autre partie d'apporter la preuve de la fiabilité de cette signature. Pour une signature présumée fiable la contestation est possible en apportant la preuve d'une faille du dispositif de signature ou des dispositifs de protection des clés (notion d'inversion de la charge de la preuve).

On devine que le législateur a voulu un cadre qui rassure les acteurs quant au risque de contestation systématique : il serait si facile de semer le doute sur l'infaillibilité d'une signature devant un tribunal. Pourtant le terme de «signature sécurisée présumée fiable » est particulièrement malheureux parce qu'il sous entend à tort qu'en dehors de ce cadre une signature ne saurait être fiable ou sécurisée.

Une signature est présumée fiable si :

1. elle est basée sur un dispositif de signature et de vérification de signature *certifié* ;
2. elle utilise des certificats *qualifiés.*

La certification de chaque application et la qualification d'un prestataire de certification sont deux processus d'une extrême complexité technique, organisationnelle et administrative, la figure 1 en donne un aperçu. La colonne de droite de cette figure concerne la qualification des certificats, celle de gauche la certification des dispositifs de signature. Trois années pleines après la loi reconnaissant l'écrit électronique, il n'existe toujours pas d'organisme accrédité pouvant évaluer un prestataire de certification pour le qualifier. Bien entendu, à ce jour, aucun service de certification ne peut délivrer de certificat qualifié.

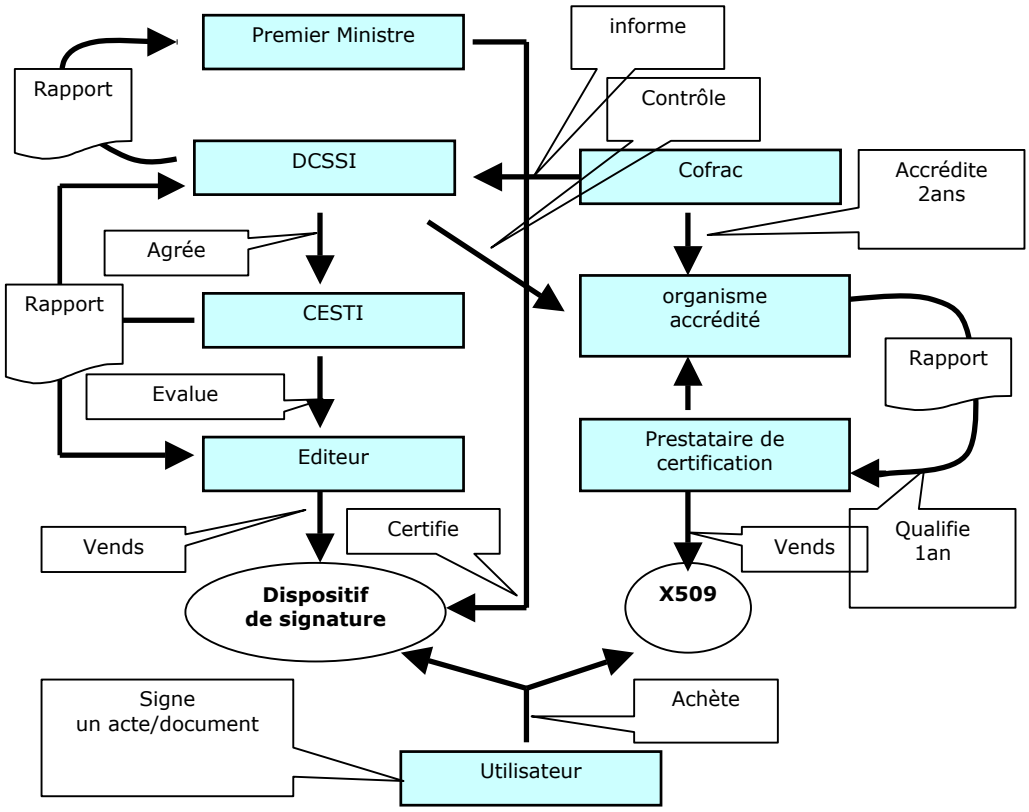

Figure 1 *Processus de certification d'une application de signature et de qualification d'une IGC*

Malgré tout cela, on trouve facilement des experts et des juristes pour affirmer que le dernier mot reviendra de toute façon à la jurisprudence. Qui est volontaire pour contribuer ainsi à l'écriture de celle-ci ?
L'économie de la signature électronique est entravée par cette complexité et devant ce constat, la commission européenne envisage d'assouplir les dispositions actuelles. Aussi n'est-il pas impossible qu'il faille ouvrir nouveau chantier juridique.

6 La signature : un concept extrêmement complexe

Pour aborder la question de la signature électronique, il faut considérer entre autre les conséquences d'une faille qui permettrait de fabriquer de fausses signatures électroniques. La version numérique des publicités non sollicitées qui débordent des boites aux lettres, c'est le spam. On n'ose pas imaginer un système de signature électronique qui aboutisse au même parallèle entre les fausses signatures manuscrites et les fausses signatures électroniques. Aussi, signer ou vérifier la signature d'un document dans des conditions qui satisfassent les normes et les exigences légales est une opération très sophistiquée. Nous abordons ici deux points parmi d'autres pour montrer cette complexité.

6.1 What you see is what you sign ?

Lorsque l'on signe un document papier on appréhende facilement la notion de ce qui a été signé. Dans le cas d'un document électronique, il faut s'interroger sur toutes les évidences. Quelles sont les garanties que l'application de signature affiche le même contenu que l'application de vérification de signature ? On montre qu'on peut fabriquer un document tel que « Word » et « Word pad » l'interprètent de façon différente. Imaginez que vous signez un contrat financier mais que la somme concernée varie avec l'application de visualisation ! De même, comment s'assurer que le document n'utilise pas, par le biais d'une macro, des informations contenue dans un autre fichier ? Dans cette hypothèse, lors de l'affichage ou de l'impression du document le contenu de celui-ci peut varier sans que l'intégrité du fichier signé ne soit affectée. Dès lors, est-il vraiment possible de signer des documents codés dans des formats permettant l'emploi de « macros » ou « d'includes » [8] ?

La certification d'un dispositif de signature doit offrir des garanties sur ce concept de « what you see is what you sign ». Cette certification englobe donc une large partie de l'application et de l'interface homme machine. La cible d'évaluation (Target Of Evaluation) est donc forcément très large.

6.2 Qu'elle est l'utilité de l'horodatage dans le processus de signature ?

Quand on signe un document, par exemple un contrat, on peut avoir besoin d'en vérifier la signature longtemps après. La révocation ou la fin de validité d'un certificat ne doit pas remettre en cause la signature des documents antérieurs. Cela signifie que l'on teste la validité d'un certificat non pas à l'instant de la vérification de signature mais à la date auquel celui-ci a été utilisé pour signer le document. Il en résulte qu'au delà de la période de validité du certificat, il est possible pour le titulaire d'un certificat d'anti-dater un document et de le signer avec le certificat expiré. De même, si la clé privée de ce certificat a été révélée, même si le certificat est révoqué, il est possible de faire une signature valide en anti-datant le document à une date antérieure à la date de révocation.
Il est donc fondamental de dater par une méthode sécurisée l'acte de signature. Ce service est confié à l'Autorité d'Horodatage qui signe une empreinte du document en y ajoutant le tampon d'horodatage.

Le droit français ne fait aucune référence aux autorités d'horodatage.

7 La dématérialisation, dans quel but ?

La dématérialisation de procédures administratives comme par exemple la déclaration d'impôt ou de TVA, est la principale justification de la notion de signature. Le droit national et européen ne se contente pas de reconnaître la signature électronique, il impose la dématérialisation d'un certain nombre d'actes administratifs. La dématérialisation des procédures est systématiquement avancée comme un atout pour rationaliser l'organisation interne d'une administration et augmenter sa productivité. Cette idée semble fondée et se vérifie pour un certain nombre d'applications existantes. Pour autant, peut-on la poser comme un axiome ? Au contraire, existe-t-il des cas ou la dématérialisation des procédures entraîne surtout une plus grande rigidité et des coûts conséquents ?

En effet, pour transformer une procédure administrative en un « workflow » intégrant la signature, il convient de formaliser complètement cette procédure. A chaque étape, on analysera les prérequis pour passer à l'étape suivante du processus. Ce qui est gagné en rigueur risque fort d'être perdu en souplesse en supprimant toute possibilité d'initiative des acteurs de la procédure.

Dans ses travaux préparatoires, le Plan Stratégique pour l'Administration Electronique (PSAE) pour la période 2003/2007 fixe des objectifs en matière de dématérialisation des processus administratifs. Les documents préparatoires du PSAE mettent l'accent sur les deux écueils les plus dangereux :

1. les risques pour les libertés individuelles ;
2. l'accroissement des inégalités des citoyens devant l'accès aux services publics.

Il est difficile de mettre en place des garde-fous contre ces dérives. Le PSAE prévoit à cet effet de maintenir un accès non dématérialisé aux services mais il fixe aussi des objectifs d'économie et des objectifs quantitatifs très ambitieux sur le taux d'opérations dématérialisées. On peut légitimement s'interroger sur le maintien de la qualité d'accès aux services par la voie non dématérialisée. L'équipement Internet des foyers et plus encore, la culture informatique, pourraient devenir alors une nouvelle frontière sociale.

8 Finalement, faut-il brûler votre certificat ?

Le discours ambiant sur les IGC a été pendant plusieurs années très simplifié. Il a porté l'image d'une solution universelle et apportant une sécurité absolue ou presque. Cette image s'est répandue parce que les porteurs de cette technologie voulaient emporter l'adhésion des utilisateurs potentiels et surtout des décideurs en délivrant un message simplifié pour présenter une problématique très complexe. Il est pourtant possible de porter un regard critique sur l'IGC en examinant les apports des technologies à base de certificats, usage par usage, et en les comparant à leur alternative si elle existe.

Le chiffrement et l'authentification entre machines sont faciles à mettre en œuvre dès lors qu'on dispose d'une source de certificats (contrôle de la configuration coté client et serveur, pas d'assistance car pas d'utilisateurs directs). Le gain de sécurité est significatif. C'est le cas des VPN, des serveurs d'authentification LDAP ou des différents composants de système de SSO et des architectures multi tiers.

L'utilisation de HTTPS (sans certificat client) est difficile du fait de la procédure d'enrôlement des AC de confiance. Cette difficulté peut être contournée en utilisant un certificat serveur commercial. Faut-il pour autant accepter ce dictat des éditeurs de navigateurs ? La réponse à cette question dépend de la communauté des utilisateurs du serveur ; plus celle-ci est étendue et éloignée de la structure qui gère le serveur, plus le recours à un certificat commercial s'impose. Même si le gain de sécurité est relatif, le coût de ce genre d'installation reste mesuré.

Dans le cas de serveurs HTTPS avec authentification par contrôle du certificat client, la distribution de certificats à une population d'utilisateurs et l'assistance qui accompagne cette opération est de toute façon très coûteuse, quelque soit l'AC qui distribue les certificats aux utilisateurs. Ne pas négliger des systèmes de Single Sign On [9] du type CAS qui sont des alternatives offrant un niveau de sécurité très intéressant.

La signature de messages est souvent confondue avec la signature des documents que celui-ci contient. Concernant S/MIME, il serait préférable de parler d'authentification (et de chiffrement). Même si S/MIME ne permet pas de signer les entêtes d'un message, ce standard peut nous protéger contre des cas d'usurpation d'adresse de messagerie. Pourquoi ne pas envisager l'usage de messages signés si l'on estime qu'il existe une véritable menace de malversation par la diffusion de fausses informations. On peut ainsi sécuriser avec un investissement mesuré des publications internes et des notes de service dont les auteurs possibles sont en nombre limité. Encore faut-il le faire systématiquement.

Quant à la signature électronique, si l'on peut en espérer un large champ d'application, la mise en œuvre d'une IGC et la distribution de certificats ne sont certainement pas des conditions suffisantes pour entamer ce processus. La signature requiert :

- un cadre juridique plus abouti ;
- des applications clientes et des « workflow » repensés pour intégrer la notion de signature ;
- des services d'horodatage ;
- des services d'archivage.

Cependant, il n'existe pas d'alternative aux techniques à base de certificats pour les besoins de signature.

Les normes et les concepts de chiffrement asymétrique, de certificats X509, d'autorités de certification et d'IGC sont incontournables. Malgré de graves difficultés qu'il convient de ne pas négliger, et à condition d'examiner systématiquement les alternatives possibles, les apports pour la sécurisation des infrastructures réseaux et pour l'authentification justifient nos efforts de déploiement de certificats. Le doute est de rigueur si l'on introduit la notion de signature qui change radicalement la problématique.

Si à la lecture de cet article vous cédez à la tentation de détruire vos certificats, interrogez vous sur une dernière évidence qui peut être trompeuse : comment s'assurer qu'un certificat est effectivement éradiqué ?

Références

[1] RFC 2246 : The TLS Protocol Version 1
[2] RFC 1847 : Security Multiparts for MIME: Multipart/Signed and Multipart/Encrypted
[3] Serge Aumont, Roland Direlewanger et Olivier Porte Accès sécurisé aux données. *Tutoriel de JRES99*
[4] Serge Aumont, Claude Gross et Philippe Leca Certificats X509 et infrastructure de gestion de clés. *Tutoriel de JRES2001*
[5] RFC 3029 : Internet X.509 Public Key Infrastructure Data Validation and Certification Server Protocols.
[6] Florent Guilleux, Textes de loi relatifs à la signature électronique en France http://www.cru.fr/igc/signature_electronique.pdf Mars 2003
[7] RSA, PKCS #11 - Cryptographic Token Interface Standard http://www.rsasecurity.com/rsalabs/pkcs/pkcs-11
[8] CERTA, Visualisation incorrecte de document word, http://www.certa.ssi.gouv.fr/site/CERTA-2001-REC-001/index.html.2.html
[9] Olivier Salaün, Introduction aux architectures de Single Sign On web, http://www.cru.fr/sso/introduction.pdf
[10] G Ellisson, B Schneier Ten Risks of PKI: What You're Not Being Told About Public Key Infrastructure 2000
[11] Roger Clarke The fundamental Inadequacies of Conventional PKI http://www.anu.edu.au/people/Roger.Clarke/II/ECIS2001.html
[12] Scott Berinato Only Mostly Dead : *RIP PKI. Why a security platform never took off*
[13] Gene Schultz 2002 will be the year that public Key Infrastructure dies http://www.arcsight.com/graphics/solutions/Arcticle02_march_2002.pdf
[14] Mickael Meehan To late for digital certificates 2001 (computerworld)