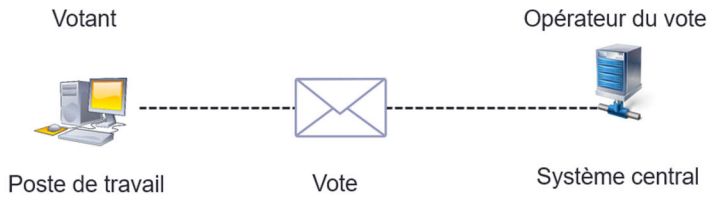
Une architecture robuste pour un système de vote électronique (et pour tout autre service du genre)

septembre 2019

Les systèmes de vote électronique ont été initiés dans la mouvance internet du début des années 2000. Voilà donc 20 ans que ces projets ont été élaborés, autant dire l'âge de la pierre à l'échelle de l'informatique.

Durant ces 20 ans les concepteurs de ces systèmes ont mobilisé beaucoup d'énergie pour sécuriser leur solution et prouver qu'elle était absolument fiable, avec des résultats pas toujours probants.
En matière de vote, outil de la démocratie, la confiance dans le système est déterminante, à l'inverse, le doute est un poison qui fait énormément de dégâts.

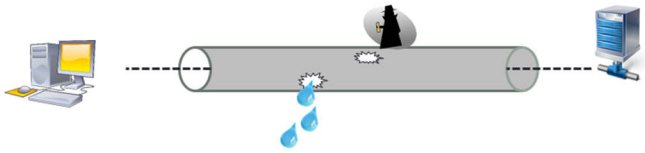


Un système non fiable par nature

Les concepteurs de ces systèmes ont défini une architecture technique sur les principes en usage à l'époque, soit un poste de travail qui communique avec un ordinateur central au travers du réseau.
L'hypothèse de travail est que le réseau internet fonctionne sur le même principe que le réseau privé.

La solution part du principe qu'un tunnel sécurisé est établi entre les deux parties. Cependant, il faut noter les éléments suivants :

1) Le réseau transite dans le domaine public. Il est exposé, donc attaquable.
2) Le protocole de communication IP établi un lien qui est structurellement discontinu, il est réputé non fiable.



Par conséquent, une solution qui fait l'hypothèse que le canal de communication est sécurisé est appelé à quelques déconvenues.

Un changement de paradigme

Nous devons imaginer une solution qui utilise un lien potentiellement corrompu. La problématique consiste à prouver aux deux parties que le paquet transmis est identique des deux côtés du lien.

Le paquet de données (le vote) est l'unité à préserver. Il s'agit de prouver aux deux parties le vote stocké par l'opérateur n'a pas été modifié ou altéré tout du long de l'opération.

L'architecture de la solution proposée est baptisée "*vérifiabilité par symétrie des empreintes*".

La solution repose sur deux outils cryptographiques : le *chiffrement* et la *fonction de hachage*.

Définitions : le chiffrement[1]

Chiffrement RSA
Procédé de cryptographie qui rend la compréhension d'un document impossible à toute personne qui n'a pas la clé de déchiffrement

Algorithme de cryptographie asymétrique
Utilise une paire de clés composée :
- d'une clé publique pour chiffrer
- d'une clé privée pour déchiffrer des données confidentielles

Définition : fonction de hachage[2]

SHA (Secure Hash Algorithm)
Fonction cryptographique qui produit une empreinte numérique.
- Pour une valeur de hachage donnée (l'empreinte), il est impossible de construire un bloc de données ayant cette valeur de hachage.
- Impossible de modifier un bloc de données sans changer sa valeur de hachage.
- Impossible de trouver deux messages différents ayant la même valeur de hachage.

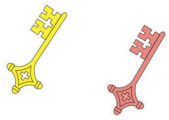
SHA-3 256
Fonction de décrite en août 2015 par la publication FIPS-202
Plus robuste que SHA-1 et SHA-2

Description d'une séquence de vote.

a) Préparation de l'opération de vote

#1 L'initialisation de l'urne.

Création des clés de chiffrages (asymétriques).
Les clés permettront de chiffrer le paquet de données (le vote, représenté par la balle jaune dans les schémas suivants) et non pas la base de données qui stocke l'ensemble des votes.



#2 Création d'un registre des identités (ID) sous forme d'empreinte

L'opérateur du vote (généralement un état) définit un ensemble de données à renseigner dont il considère qu'elles sont suffisantes pour authentifier une personne. Par exemple à Genève, le numéro de carte de vote (reçue par courrier), la date de naissance et la commune d'origine sont réputés suffisant pour authentifier un votant.

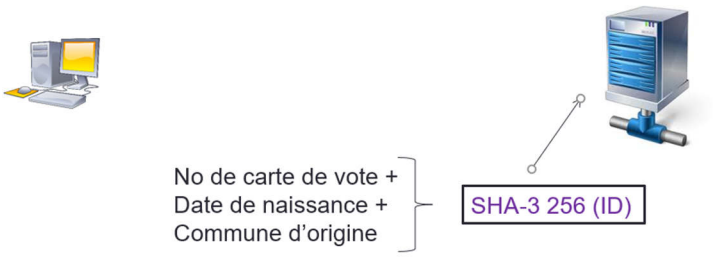
L'opérateur du vote établi deux listes :

1) Registre des noms et numéros de carte de vote
2) Numéro de carte de vote + date de naissance + commune d'origine

La première liste est destinée à l'impression des cartes de vote et l'envoi au votant.

La seconde liste est destinée à générer l'empreinte ID (hash) de chaque votant. L'ensemble des empreintes sont ensuite stockées dans un registre des identités (ID).
La seconde liste ayant servi à générer les empreintes est détruite. Ainsi seul le registre des empreintes ID subsiste. Rappelons qu'il est impossible de recréer l'identité du votant sur la base de son empreinte.

[1] https://fr.wikipedia.org/wiki/Chiffrement_RSA
[2] https://fr.wikipedia.org/wiki/Fonction_de_hachage_cryptographique



Ainsi le système a créé l'identification d'une personne sans faire de lien avec son identité. Il n'est donc pas possible de reconstituer l'identité d'une personne sur la base de son empreinte ID.

b) Opération de vote

#3 Téléchargement du logiciel et de la description

Le votant télécharge sur le site de l'opérateur de vote un paquet d'outils qui va lui permettre de voter.

Le paquet d'outils contient :

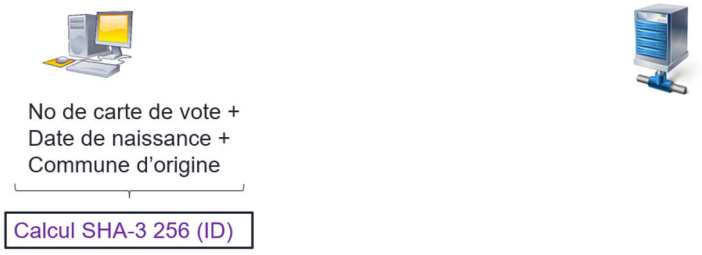
Les programmes : la couche de présentation, les outils de hash et de chiffrage.
La description : l'objet du vote et clé de chiffrage.

La couche de présentation forme l'interface et le système de vote du coté de l'utilisateur. L'objet du vote décrit la liste d'objets ou de personnes qui constituent le scrutin.

Afin de garantir que le paquet d'outils reçu est le bon, l'empreinte du paquet d'outils téléchargé est calculée sur le poste de travail. L'empreinte résultante est ensuite comparée avec celle contenue dans le serveur de l'opérateur. Elle doit être identique.

#4 Vote : calcul de l'identifiant

Le votant renseigne les données nécessaires à son authentification. Ces données sont hashées avec les logiciels qui lui sont mis à disposition à l'étape précédente. Ainsi il génère sa propre empreinte SHA-3 (ID) d'identité.

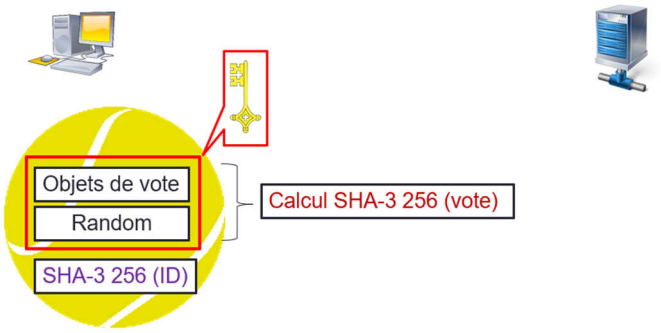


#5 Vote : prise de position

L'interface de prise de position est générée sur la base de la description des objets de vote reçus dans le paquet décrit dans la séquence #3.
La prise position est faite par le votant.

Le système du votant prépare un paquet de vote.

- Les données de vote + une chaine aléatoire (padding) sont chiffrés avec la clé publique de l'opération. Ensuite une empreinte SHA-3 (vote) est calculée pour cet élément
- Le paquet de vote comprend : les données de vote chiffées + l'empreinte de l'ID du votant.



Le votant conserve son empreinte de vote et son empreinte d'identité (SHA-3 ID + SHA-3 vote).

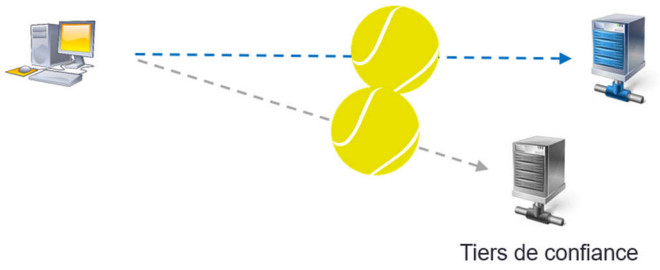
La chaine aléatoire (padding) insérée dans la prise de vote permet à deux même prises de position de ne pas avoir la même empreinte. L'empreinte est unique pour chaque vote, ce qui va être utile pour la preuve (vérifiabilité).

Ainsi le paquet de vote comprend la prise de position chiffrée ainsi que l'empreinte de l'ID du votant.
La prise de position ne peut être déchiffrée qu'à l'aide de la clé privée de l'opérateur de vote. L'empreinte de l'ID du votant est transmise sans qu'on ne puisse reconstituer son identité.

#6 Vote : envoi

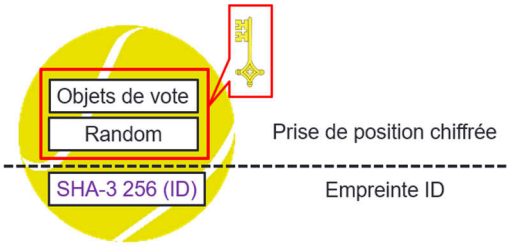
Le paquet de vote est envoyé au serveur de l'opérateur.

Il est possible de mettre en place un serveur tiers de confiance.
Dans ce cas, le paquet sera alors systématiquement envoyé aux deux instances.
Ainsi, une autorité tierce qui dispose de la totalité des prises de position pourra dérouler l'opération pour son propre compte et comparer le résultat.



#7 Traitement du vote, ID

A la réception du paquet de vote, le serveur sépare l'empreinte ID de la prise de position chiffrée.



Le serveur recherche la correspondance entre l'empreinte ID reçue avec l'empreinte identique contenue dans le registre des empreintes ID.

- Si l'ID est retrouvé, il brûle l'ID du registre.
- Si l'ID est invalide, le vote est rejeté et envoi d'une notification à l'utilisateur pour un ID invalide.

#8 Traitement du vote, prise de position

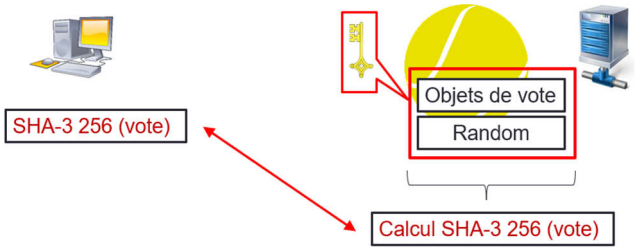
La prise de position reçue est chiffrée, il n'est pas possible d'en lire le contenu.

- Le système calcul l'empreinte SHA-3 (vote) de la prise de position.
- La prise de position chiffrée et son empreinte (vote) sont stockées sur le serveur.
- L'empreinte calculée par le serveur est renvoyée au votant.

L'empreinte calculée sur le poste du votant est comparée avec l'empreinte calculée par le serveur.

- Si les empreintes sont identiques, le vote est accepté. L'opération est terminée.
- Si les empreintes sont différentes, une alarme est inscrite dans le serveur et le vote est refusé.

Le traitement du vote est une opération atomique.

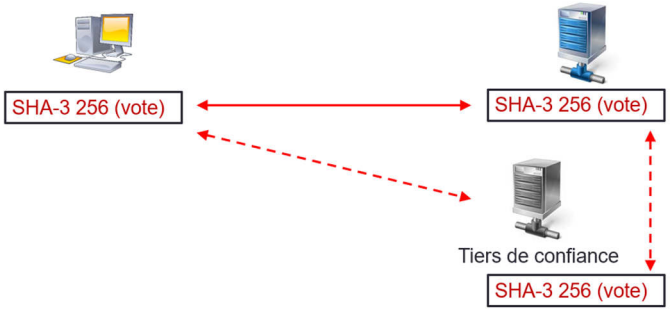


c) Vérifiabilité

Le votant peut à n'importe quel moment recontrôler que son vote est bien existant sur le serveur.
Il lui suffit d'envoyer l'empreinte (vote) locale au serveur. Le serveur contrôle l'existence de cette empreinte (vote) dans le stockage des votes et confirme être dépositaire de cette même empreinte.
Par ailleurs, cette opération de contrôle est systématiquement exécutée lors de l'envoi du bulletin.

Si les empreintes, qui sont uniques, sont strictement identiques chez le votant et sur le serveur, le votant à la confirmation que son vote est bien pris en compte, ce qui constitue la vérifiabilité du vote.
C'est le principe de la "*vérifiabilité par symétrie des empreintes*".

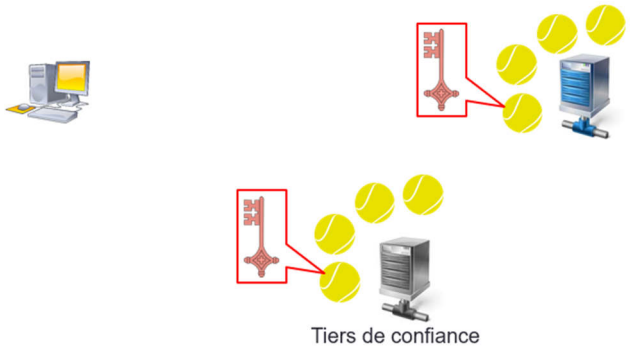
De plus, si un tiers de confiance est mis en place, la vérifiabilité peut être effectuée auprès du tiers indépendant.



d) Ouverture de l'urne

Lors de l'ouverture de l'urne, chaque prise de position individuelle est déchiffrée avec la clé privée de l'opération de vote.

Les prises de position sont maintenant disponibles en clair et le serveur peut effectuer l'opération de comptage des résultats.



Si un tiers de confiance a été mis en place, la clé de déchiffrage lui est communiquée à l'ouverture de l'urne. Le tiers dispose ainsi des votes et de la clé de déchiffrage. Il peut produire son propre décompte des bulletins. Ce décompte indépendant renforce la confiance sur le traitement du comptage des bulletins.

Avantage du système reposant sur la *vérifiabilité par symétrie des empreintes*

Le système de "*vérifiabilité par symétrie des empreintes*" a quelques avantages parmi lesquels :

- Ergonomie simplifiée pour l'utilisateur, pas de tableau de vérification et de codes multiples à introduire.
- Compréhension et clarté des principes.
- Fonctionnement technique autonome.
- Transaction atomique.
- Ne nécessite pas le chiffrement de la ligne. Toutefois un lien HTTPS standard est conservé pour les quittances.
- Architecture robuste.
- Résistant aux attaques «man in the middle».
- Vérifiabilité implicite.
- Permet d'intégrer un tiers de confiance.

Conclusions

La proposition comporte de nombreux avantages, elle mérite certainement d'être débattue afin d'affiner la démarche, challenger la sécurité et valider le concept.

La seule certitude que je puisse avoir est que les solutions disponibles sur le marché sont techniquement obsolètes et lacunaires du point de vue de la sécurité. Elles devront être repensées dans leur intégralité avec des solutions techniques en phase avec les connaissances techniques actuelles.

Pascal Rulfi, ingénieur consultant à Genève.